Dipl.-Math. Johannes Cuno

# Combinatorial and Probabilistic Aspects of Discrete Groups

DISSERTATION

zur Erlangung des akademischen Grades

Doktor der technischen Wissenschaften

eingereicht an der

**Technischen Universität Graz**

Betreuer:

Univ.-Prof. Dipl.-Ing. Dr. rer. nat. Wolfgang Woess

Institut für Mathematische Strukturtheorie

Graz, August 2015

**EIDESSTATTLICHE ERKLÄRUNG**

***AFFIDAVIT***

Ich erkläre an Eides statt, dass ich die vorliegende Arbeit selbstständig verfasst, andere als die angegebenen Quellen/Hilfsmittel nicht benutzt, und die den benutzten Quellen wörtlich und inhaltlich entnommenen Stellen als solche kenntlich gemacht habe. Das in TUGRAZonline hochgeladene Textdokument ist mit der vorliegenden Dissertation identisch.

*I declare that I have authored this thesis independently, that I have not used other than the declared sources/resources, and that I have explicitly indicated all material which has been quoted either literally or by content from the sources used. The text document uploaded to TUGRAZonline is identical to the present doctoral dissertation.*

Datum/Date

Unterschrift/Signature

# Contents

# A gentle introduction written in layman's terms

A couple of days after Emmy Noether, one of the leading mathematicians of her time, had passed away, Albert Einstein wrote a letter to the editor of the New York Times. It was published as an obituary on Sunday, the 5th of May 1935, and contained the words

*"Pure mathematics is, in its way, the poetry of logical ideas."*

This is a good description of what is driving people like me. In pure mathematics, one often starts with a set of axioms and wonders about the objects that satisfy these axioms. More precisely, one seeks to understand the objects and their properties. The method to achieve this goal is logical conclusion, which makes research engaging and, occasionally, extremely frustrating. But, every single thought experiment expels the fog a little bit and improves the view on hidden structures of intrinsic perfection and beauty. Unfortunately, this view is hard to communicate without using a particular language. So, in case you want to have a closer look into my thesis, which you are cordially invited to, please keep in mind that the content is not necessarily difficult but formulated in the language of mathematics which takes a while to get used to. Nevertheless, let me try to give you an impression of the projects that constitute my thesis. We shall do this in reverse order and start with the third one.

### Project C: Distinguishing graphs with infinite motion and non-linear growth

The third project of my thesis is about a conjecture in graph theory. Let me emphasise that the graphs under investigation are by no means graphs of functions as they are studied at school. For us, a graph is nothing but a vertex set and a collection of two-element subsets, the edges. An example is the graph with vertex set $\{a,b,c,d\}$ and edges $\{a,b\}$, $\{b,c\}$, $\{c,d\}$, $\{d,a\}$. If you want to, you may think of this graph as illustrated in the first part of Figure 1. But, keep in mind that this is only a geometric model of the abstract data.

Let us make use of this geometric model and think of the vertices as balls. We may thus take the vertices in hands and later return them to possibly different positions. This is a permutation of the vertices, see the second part of Figure 1. If you have followed me until now, you have started playing with an abstract object, which is the key to pure mathematics. Let us go on. Among the permutations there are good and bad ones. Consider for example the permutation that swaps the vertices $a$ and $b$ and leaves the vertices $c$ and $d$ at their original position. After this permutation, go to the vertex $a$ and have a look into the neighbourhood. The vertex $d$ is not a neighbour any more and the vertex $c$ has become a new neighbour. We consider this as a bad permutation. Good permutations are the ones that preserve neighbourhoods in the sense that every vertex has precisely the same neighbours as before. These permutations are called automorphisms. An example is the identity, which leaves every vertex

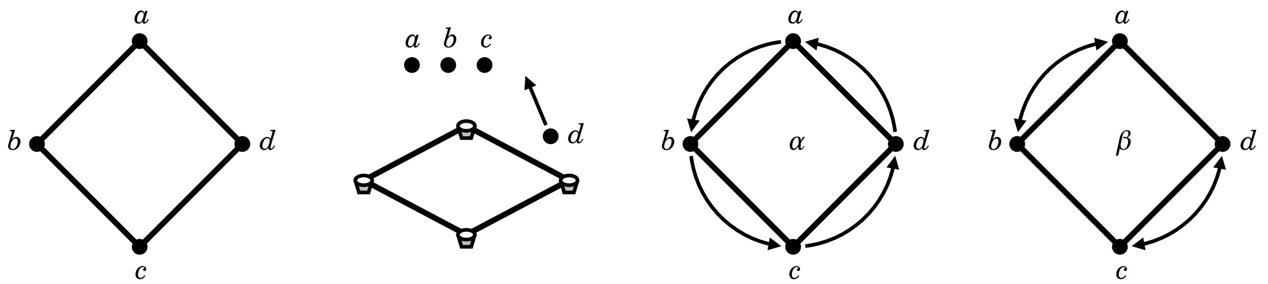


Figure 1: The graph with vertex set $\{a,b,c,d\}$ and edges $\{a,b\}$, $\{b,c\}$, $\{c,d\}$, $\{d,a\}$.

at its original position. Two slightly more instructive examples are illustrated in the third and fourth part of Figure 1. We will come back to them later. For now, let me tell you that our graph has eight automorphisms. Can you find them all?

We seek to develop an understanding of the interplay between automorphisms and colourings. Imagine we assigned a colour to every vertex. Then, we say that an automorphism or, more generally, a permutation preserves this colouring if it can be obtained by permuting the vertices of each colour among themselves. I recommend to digest this definition for a second and to observe that, no matter which colouring we have chosen, the identity will certainly preserve it. Here is a challenge: *"Assume we are given a graph and a number of colours, can we find a colouring such that the identity is the only automorphism that preserves it?"*

For the graph with vertex set $\{a,b,c,d\}$ and edges $\{a,b\}$, $\{b,c\}$, $\{c,d\}$, $\{d,a\}$ we will certainly master the challenge with four colours, simply by giving each vertex an individual colour. Notice that we will even master the challenge with three colours. In order to do so, we give the vertices $a$ and $b$ the first colour, the vertex $c$ the second colour, and the vertex $d$ the third colour. Then, the only two permutations that preserve this colouring are the identity and the one that swaps the vertices $a$ and $b$ and leaves the vertices $c$ and $d$ at their original position. But, as discussed above, the latter is not an automorphism. So, we are done. However, we will not be able to master the challenge with two or fewer colours. Can you give an argument?

The minimal number of colours needed to master the challenge is called the distinguishing number of the graph. With your help, we have just proved that the graph with vertex set $\{a,b,c,d\}$ and edges $\{a,b\}$, $\{b,c\}$, $\{c,d\}$, $\{d,a\}$ has distinguishing number three. The third project of my thesis focuses on graphs with infinitely many vertices ("infinite") and two more properties. First, every vertex has only finitely many neighbours ("locally finite") and, second, one can connect every pair of vertices by a sequence of edges ("connected"). An example of such an infinite, locally finite, connected graph is illustrated in Figure 2. In light of this figure, it should not surprise you that this graph in an example of what is called a tree. In the year 2007, Watkins and Zhou showed that many trees, including our one, are two-distinguishable, i. e. have distinguishing number one or two, see [WZ07, Theorem 3.1].

Another property of our tree is that, as soon as an automorphism is not the identity, it has to move infinitely many vertices.[1] This property is called infinite motion and, roughly speaking, it seems to facilitate two-distinguishability. Tucker even conjectured that every infinite, locally finite, connected graph with infinite motion is two-distinguishable. His conjecture is still open but, in the third project

[1] I will not give a rigorous proof of this statement. But, we should at least have a look at Figure 2 and observe that if an automorphism swaps the vertices $a$ and $b$, then it has to leave the vertex $c$, being the only common neighbour of $a$ and $b$, at its original position. Therefore, it has to move all the infinitely many vertices from the box $A$ to the box $B$, and vice versa.

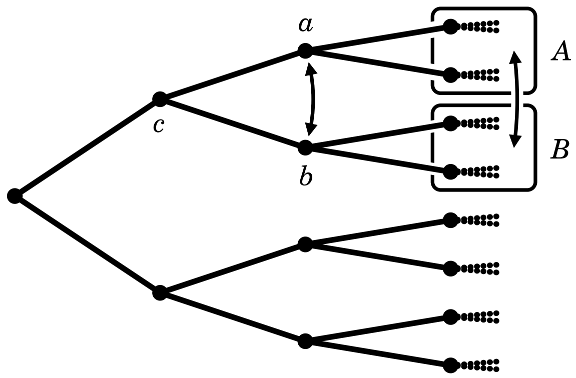


Figure 2: This tree has infinite motion.

of my thesis, my coauthors and I prove its validity under the assumption that the graph does not grow too fast. If you want to know what growing means, then I succeeded in making you curious and refer to page 80 where Sections 2.3 and 2.4 provide you with the necessary definitions.

## Project B: The Tits alternative for non-spherical triangles of groups

Recall the graph with vertex set $\{a,b,c,d\}$ and edges $\{a,b\}$, $\{b,c\}$, $\{c,d\}$, $\{d,a\}$. As I told you and you might have checked, it has eight automorphisms. A very nice property of automorphisms is that we can multiply them. More precisely, we can take any two automorphisms and compose them to get another one. Take, for example, the automorphisms $\alpha$ and $\beta$ illustrated in the third and fourth part of Figure 1, and apply first $\alpha$ and then $\beta$. The resulting automorphism, which we denote by $\alpha * \beta$, swaps the vertices $b$ and $d$ and leaves the vertices $a$ and $c$ at their original position. Another way to get the same automorphism is $\beta * \alpha * \alpha * \alpha$.

Let me give you three properties of the automorphisms and their composition. First, the composition is associative, i. e. rearranging parentheses as in $(\beta * \alpha) * (\alpha * \alpha)$ and $\beta * (\alpha * (\alpha * \alpha))$ does not change the resulting automorphism. Second, the identity is a neutral element, i. e. for every automorphism $\varphi$ the equations $\mathrm{id} * \varphi = \varphi$ and $\varphi * \mathrm{id} = \varphi$ hold. Third, every automorphism has an inverse, i. e. for every automorphism $\varphi$ there is an automorphism $\widetilde{\varphi}$ such that the equations $\widetilde{\varphi} * \varphi = \mathrm{id}$ and $\varphi * \widetilde{\varphi} = \mathrm{id}$ hold.

These properties are very important, and one takes them as axioms: *"A group is a set equipped with a multiplication that is associative, has a neutral element, and has the property that every element has an inverse."* There are a lot of examples of groups, not only the automorphisms and their composition but also the integers and their addition and many more. In fact, group theory is an active field of research and the second project of my thesis belongs into this field. More precisely, it is about a construction principle for groups. The fundamental ingredient is a triangle of groups. For now, you may think of a triangle of groups as a collection of three groups $A$, $B$, $C$ that intersect[2] as illustrated in Figure 3. Based on such a triangle of groups, we may construct its colimit group $\mathfrak{G}$. This is another group into which the elements of the triangle of groups are mapped.[3]

[2]This deserves an explanation. Recall that each group comes with its own multiplication. So, if we take two elements that are both in the group $A$ and both in the group $B$, then there are a priori two ways to multiply them. We shall only consider situations where the two products are the same. In other words, we assume that the multiplications of the three groups agree on the intersections. Moreover, we assume that each intersection, on which we now have a unique way to multiply two elements, forms again a group.

[3]This mapping is even compatible with the multiplications. More precisely, if we take two elements from the same group,

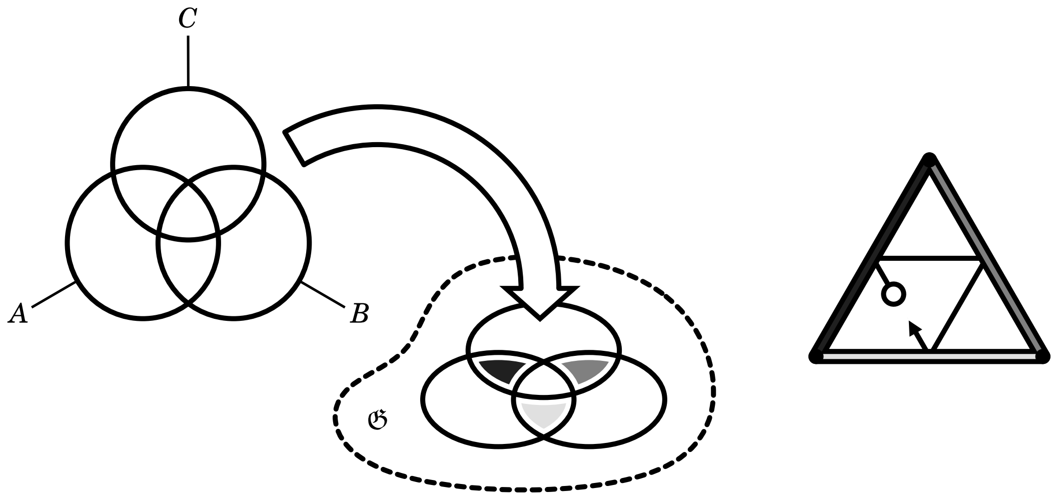


Figure 3: A triangle of groups with its colimit and a billiard shot.

One may ask different questions about the colimit group $\mathfrak{G}$. Here is one of them: *"If we take two distinct elements from the same group, for example from the group A, are they always mapped to distinct elements of the colimit group $\mathfrak{G}$?"*

There are examples for which the answer is negative. But, in the year 1991, Gersten and Stallings found a condition under which the answer is positive, see [Sta91]. It is called non-sphericity. In the second project of my thesis, my coauthor and I prove that, for non-spherical triangles of groups, two distinct elements, be they from the same group or not, are always mapped to distinct elements of the colimit group $\mathfrak{G}$.[4] It is therefore legitimate to say that the colimit group $\mathfrak{G}$ contains a copy of the triangle of groups.

We were originally interested in a different question. Let me explain. The automorphisms $\alpha$ and $\beta$ illustrated in the third and fourth part of Figure 1 satisfy the equation $\alpha * \beta * \alpha * \beta = \mathrm{id}$. So, products sometimes turn out to be the neutral element, and it is desirable to have criteria that help to decide in advance whether a product can be the neutral element or not.[5] We have found such a criterion for the colimit group $\mathfrak{G}$. Given a non-spherical triangle of groups, we construct a triangular billiard table. Each of the three cushions corresponds to one of the three sets that are indicated in different shades of grey in Figure 3. The billiard table helps us to analyse products whose factors are taken from these sets. More precisely, we show that if there is a billiard shot that hits at least one cushion and has the property that the sequence of cushions agrees with the sequence of sets where the factors are taken from, then the product cannot be the neutral element.

Assume, for example, we are given a non-spherical triangle of groups with the property that the billiard table is an equilateral triangle. Now, consider the product $\varphi_1 * \varphi_2 * \varphi_3$ with factors $\varphi_1$ taken from the darkest set, $\varphi_2$ taken from the intermediate set, $\varphi_3$ taken from the brightest set, then the billiard shot on the right-hand side of Figure 3 shows that $\varphi_1 * \varphi_2 * \varphi_3$ cannot be the neutral element.

for example from the group $A$, then it does not matter whether we first multiply them using the multiplication of $A$ and then map the product to the colimit group $\mathfrak{G}$ or first map them to the colimit group $\mathfrak{G}$ and then multiply the images using the multiplication of $\mathfrak{G}$.

[4]As illustrated in Section 3.1 on page 51, this is not a consequence of the result obtained by Gersten and Stallings.

[5]A wonderful example of such a criterion is Britton's Lemma, see also the second paragraph in Section 1.1.1 on page 15.

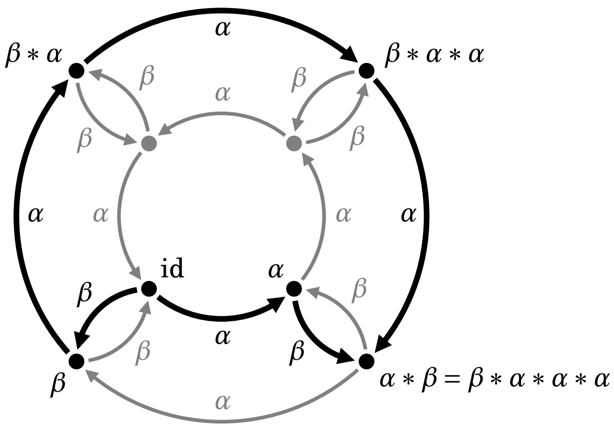


Figure 4: The Cayley graph with respect to the generators $\alpha$ and $\beta$.

Moreover, by hitting the ball a bit stronger, it shows that the same holds for $\varphi_1 * \varphi_2 * \varphi_3 * \ldots * \varphi_1 * \varphi_2 * \varphi_3$.

So, this criterion translates between a geometric and an algebraic property. And, in some way, geometry is the way we think. However, we succeed in applying this criterion to prove a theorem about the structure of the colimit group $\mathfrak{G}$, essentially by playing billiards, see e. g. Figure 16 on page 74.

## Project A: Random walks on Baumslag-Solitar groups

After these abstract considerations let us, once again, return to our initial example of the graph with vertex set $\{a,b,c,d\}$ and edges $\{a,b\}$, $\{b,c\}$, $\{c,d\}$, $\{d,a\}$. We have introduced its automorphism group and already mentioned that the equations $\alpha * \beta = \beta * \alpha * \alpha * \alpha$ and $\alpha * \beta * \alpha * \beta = \mathrm{id}$ hold. In order to understand all possible ways to compose $\alpha$ and $\beta$ and the automorphisms they produce, we may draw the diagram given in Figure 4. The eight vertices represent the eight automorphisms. Each automorphism $\varphi$ has two outgoing arrows; one is labelled by $\alpha$ and goes to $\varphi * \alpha$ and the other one is labelled by $\beta$ and goes to $\varphi * \beta$. Once we have drawn this diagram, we can actually read off the equation $\alpha * \beta = \beta * \alpha * \alpha * \alpha$ because, starting at the identity, the sequences $\xrightarrow{\alpha}\xrightarrow{\beta}$ and $\xrightarrow{\beta}\xrightarrow{\alpha}\xrightarrow{\alpha}\xrightarrow{\alpha}$ both end up at the same vertex. Similarly, we can read off the equation $\alpha * \beta * \alpha * \beta = \mathrm{id}$ because, starting at the identity, the sequence $\xrightarrow{\alpha}\xrightarrow{\beta}\xrightarrow{\alpha}\xrightarrow{\beta}$ returns to its origin. The diagram is called the Cayley graph with respect to the generators $\alpha$ and $\beta$, even though it is rather a quiver than a graph. Such a Cayley graph can be constructed for any group with respect to any set of generators.

The first project of my thesis considers random walks on Baumslag-Solitar groups. Instead of giving an abstract definition of these groups, I recommend to have a look on page 17 where Figure 1 illustrates a part of the Cayley graph of the Baumslag-Solitar group BS(1,2). Before running a random walk on this group, let me give you an introductory example.

Consider the graph with vertex set $\{1,2,3,4,5,6\}$ and edges $\{1,2\}$, $\{2,3\}$, $\{3,4\}$, $\{4,5\}$, $\{5,6\}$. Let us now define a random walk on this graph. Imagine we were starting at vertex 5 and proceeding according to the following rule. From time to time, we select one of the two neighbours and move there, say with probability $2/3$ to the smaller one and with probability $1/3$ to the larger one. This procedure goes on until we reach vertex 1 or vertex 6.

A natural question is: *"How likely is it to end up at vertex* 1*?"* In fact, the answer to this question is a classic. Before reading it, have a guess. Is it smaller or larger than $1/2$? To find the answer, we make a thought experiment.

Let $p(n)$ be the probability to end up at vertex 1 when starting at vertex $n$. In the two cases $n=1$ and $n=6$, we would stop immediately. Therefore, $p(1)=1$ and $p(6)=0$. But, if $n$ is between 2 and 5, we can make a first step decomposition. More precisely, starting at $n$, the first step goes either to $n-1$ or to $n+1$. After this first step, the probability to end up at vertex 1 has become $p(n-1)$ or $p(n+1)$, respectively. So, we obtain the formula $p(n) = 2/3 \cdot p(n-1) + 1/3 \cdot p(n+1)$. Hence, we know that ① $p(1)=1$, ② $p(2) = 2/3 \cdot p(1) + 1/3 \cdot p(3)$, ③ $p(3) = 2/3 \cdot p(2) + 1/3 \cdot p(4)$, ④ $p(4) = 2/3 \cdot p(3) + 1/3 \cdot p(5)$, ⑤ $p(5) = 2/3 \cdot p(4) + 1/2 \cdot p(6)$, ⑥ $p(6)=0$. This system of equations can be solved using the techniques we have learnt at school, and it turns out that $p(5) = 16/31 \approx 52\%$. Record from the above that, if $n$ is between 2 and 5, in which case $n$ is called an interior vertex, we obtained the formula $p(n) = 2/3 \cdot p(n-1) + 1/3 \cdot p(n+1)$. Functions with this property are called harmonic on the interior vertices. As soon as it is clear which vertices are the interior ones, we just call them harmonic. The probability $p(n)$ is not the only harmonic function. Another one is the probability $q(n)$ to end up at vertex 6 when starting at vertex $n$. The probabilities $p(n)$ and $q(n)$ allow us to master the following challenge: *"Assume we are given two arbitrary real numbers a and b, can we find a harmonic function $h : \{1,\dots,6\} \to \mathbb{R}$ with $h(1)=a$ and $h(6)=b$?"* The answer is positive, and $h(n) := a \cdot p(n) + b \cdot q(n)$ is the unique solution.

As already mentioned, we are interested in random walks on Baumslag-Solitar groups. So, recall the Cayley graph of the Baumslag-Solitar group BS(1,2) and imagine that we start a random walk at the neutral element. Again, from time to time, we select for example one of the four neighbours and move there, say to each neighbour with probability $1/4$. In this model, we do not stop when reaching some vertex such as 1 or 6 in the introductory example. So, there is no obvious analogue to the probabilities $p(n)$ and $q(n)$. Nevertheless, we may still wonder about harmonic functions, i. e. functions with the property that the value at each vertex is the average of the values at the four neighbours.

It turns out that one can describe all bounded harmonic functions by constructing a new space called the Poisson-Fürstenberg boundary. Its elements are, roughly speaking, all possible kinds of long-time behaviour of the random walk. And, while we previously ended up in a random element of the set $\{1,6\}$, we now end up in a random element of the Poisson-Fürstenberg boundary.

The latter has the nice property that we can describe all bounded harmonic functions on the group in terms of the probabilities to hit certain parts of the Poisson-Fürstenberg boundary. The problem is to understand this boundary and to identify it geometrically. For the Baumslag-Solitar group BS(1,2), this has been done by Kaĭmanovich in [Kaĭ91, Theorem 5.1]. We extend his result and investigate all non-amenable Baumslag-Solitar groups such as BS(2,3). Even though these groups are substantially different, my coauthor and I obtain similar results.

## Acknowledgements

First of all, I would like to thank my parents and my sister. They provided me with two of the most valuable things a mathematician can have, namely time and moral support. By accepting my individual speed of learning and consequently supporting my decisions, they contributed more than they may guess to the existence of this document.

Mathematically, my advisor Wolfgang Woess played an important role. He carefully introduced me to random walks on groups and, whenever my problems seemed to drive me into despair, he was available as a source of knowledge and creativity. He encouraged me to work on other projects, too, which brings me to my coauthors. Each of them taught me a lot and thus contributed substantially to my mathematical education. My colleagues at TU Graz and the DK Discrete Mathematics provided an inspiring environment. This includes the guests at the department, in particular Vadim Kaĭmanovich, with whom I had a couple of valuable discussions. Last but not least, I would like to mention the students that I had the pleasure to teach over the years. Their enthusiasm reminded me every week how awesome even the basics of mathematics are. I am deeply grateful to all of these people.

# Summary

This thesis consists of three projects, each of which focuses on a different aspect of infinite graphs and groups. The first project is joint with Ecaterina Sava-Huss. We consider random walks on non-amenable Baumslag-Solitar groups and give an extrinsic description of the Poisson-Fürstenberg boundary. First, we recall Baumslag-Solitar groups and their geometry. Then, we change our focus to random walks on groups. The Poisson-Fürstenberg boundary is, roughly speaking, a probability space that serves as a model to describe the long-time behaviour of the random walk. For random walks on non-amenable Baumslag-Solitar groups we describe it in terms of the boundary of the hyperbolic plane, denoted by $\partial\mathbb{H}$, and the space of ends of the associated Bass-Serre tree, denoted by $\partial\mathbb{T}$. Our main result is:

**Theorem 2.12 (see p. 39)** *Let $Z = (Z_0, Z_1, \dots)$ be a random walk on a non-amenable Baumslag-Solitar group $G = \mathrm{BS}(p,q)$ with $1 < p < q$. Assume there is an $\varepsilon > 0$ such that the increment $X_1$ has finite $(2+\varepsilon)$-th moment. If the vertical drift is non-negative, i. e. $\delta \geq 0$, then the Poisson-Fürstenberg boundary is isomorphic to $(\partial\mathbb{T}, \mathscr{B}_{\partial\mathbb{T}}, \nu_{\partial\mathbb{T}})$. On the other hand, if the vertical drift is negative, i. e. $\delta < 0$, then the Poisson-Fürstenberg boundary is isomorphic to $(\partial\mathbb{H} \times \partial\mathbb{T}, \mathscr{B}_{\partial\mathbb{H}\times\partial\mathbb{T}}, \nu_{\partial\mathbb{H}\times\partial\mathbb{T}})$.*

The second project is joint with Jörg Lehnert. A triangle of groups is a commutative diagram of groups and injective homomorphisms of the following form. For every subset $J \subseteq \{1,2,3\}$ with $|J| \leq 2$ there is a group $G_J$ and for every two subsets $J_1 \subset J_2 \subseteq \{1,2,3\}$ with $|J_2| \leq 2$ there is a homomorphism $\varphi_{J_1 J_2} : G_{J_1} \to G_{J_2}$. We shall always assume that in each $G_{\{i,j\}}$ the images of $G_{\{i\}}$ and $G_{\{j\}}$ intersect precisely along the image of $G_\varnothing$. In analogy to the pushout, which yields the amalgamated free product of two groups, we consider the generalised pushout of the diagram, also known as the colimit.

The fundamental question is: *"What can be said about the structure of the colimit group?"* Gersten and Stallings introduced the notion of curvature showed that, for non-spherical triangles of groups, the homomorphisms from the groups $G_J$ with $J \subseteq \{1,2,3\}$ and $|J| \leq 2$ to the colimit group are all injective. We first give an example showing that, despite of the fact that these homomorphisms are injective, the intersection of the images of $G_{\{1,2\}}$ and $G_{\{1,3\}}$ may be strictly larger than the image of $G_{\{1\}}$. However, our example happens to be a spherical triangle of groups and we can prove an intersection theorem saying that for non-spherical triangles of groups and, more generally, for non-spherical Corson diagrams this surprising behaviour is not possible.

Bridson's theory of two-dimensional metric simplicial complexes allows us to construct non-abelian free subgroups in the colimit groups of many non-spherical triangles of groups. The language we use is the one of billiards. In the end, we give two natural conditions, each of which ensures that in those cases where our construction does not produce a non-abelian free subgroup, the colimit group is already virtually solvable. More precisely:

**Theorem 4.24 (see p. 75)** *The Tits alternative holds for the class of colimit groups of non-degenerate non-spherical triangles of groups with the property that the group $G_\varnothing$ either has a non-abelian free subgroup or is virtually solvable.*

**Theorem 4.25 (see p. 75)** *The Tits alternative holds for the class of colimit groups of non-spherical triangles of groups with the property that every group $G_J$ with $J \subseteq \{1,2,3\}$ and $|J| \leq 2$ either has a non-abelian free subgroup or is virtually solvable.*

The third project is joint with Wilfried Imrich and Florian Lehner. It asks for the minimal number of colours needed to colour the vertices of a given graph in such a way that the trivial automorphism is the only one that preserves the colouring. The infinite motion conjecture concerns infinite, locally finite, connected graphs and states that if such a graph has the property that every non-trivial automorphism moves infinitely many vertices, then the minimal number of colours needed is either 1 or 2. It holds for graphs with linear growth. In our project, we obtained the first result going beyond the realm of linear growth:

**Corollary 3.5 (see p. 85)** *Let $G$ be an infinite, locally finite, connected graph with infinite motion and $v_0 \in V(G)$. Moreover, let $\varepsilon > 0$. If there are infinitely many $n \in \mathbb{N}$ such that*

$$\left|B_{v_0}(n)\right| \leq \frac{n^2}{(2+\varepsilon)\log_2(n)},$$

*then the distinguishing number $D(G)$ is either 1 or 2. In particular, Tom Tucker's infinite motion conjecture holds for all graphs of growth $o(n^2/\log_2(n))$.*

The project has stimulated further research. Inspired by our joint project, Florian Lehner extended the result to graphs with a higher growth rate, including all graphs of polynomial growth. Moreover, interesting new questions have come up. For example, the one asking for a two-colouring that makes mainly use of one colour and rarely use of the other one, and still breaks all non-trivial automorphisms. Also the question of endomorphism distinguishability seems to be a promising playground.

# Project A

# Random walks on Baumslag-Solitar groups

**( with Ecaterina Sava-Huss )**

## 1 Introduction and preliminaries

### 1.1 Baumslag-Solitar groups

#### 1.1.1 Introduction

For any two non-zero integers $p$ and $q$, the Baumslag-Solitar group $\mathrm{BS}(p,q)$ is given by the presentation $\mathrm{BS}(p,q) = \langle a,b \mid ab^p a^{-1} = b^q \rangle$. These groups were introduced by Baumslag and Solitar in [BS62], who identified $\mathrm{BS}(2,3)$ as the first example of a two-generator one-relator non-Hopfian group and thus answered a question by B. H. Neumann, see [Neu54]. Later on, was shown that $\mathrm{BS}(p,q)$ is Hopfian if and only if $|p| = 1$ or $|q| = 1$ or $\mathscr{P}(p) = \mathscr{P}(q)$, where $\mathscr{P}(x)$ denotes the set of prime divisors of $x$, see [BS62] and [Mes72].

The structure of Baumslag-Solitar groups can be studied by means of HNN extensions. Indeed, $\mathrm{BS}(p,q)$ is precisely the HNN extension $\mathbb{Z}*_\varphi$ with isomorphism $\varphi : p\mathbb{Z} \to q\mathbb{Z}$ given by $\varphi(p) := q$. This fact 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 $1 \in \mathrm{BS}(p,q)$ if it contains $ab^r a^{-1}$ with $p \mid r$ or $a^{-1}b^r a$ with $q \mid r$ as a subword.

Now, one can easily conclude that, if neither $|p| = 1$ nor $|q| = 1$, the elements $x := a$ and $y := bab^{-1}$ generate a non-abelian free subgroup. So, $\mathrm{BS}(p,q)$ is non-amenable. On the other hand, if $|p| = 1$ or $|q| = 1$, a simple calculation shows that the normal subgroup $\langle\langle b \rangle\rangle \trianglelefteq \mathrm{BS}(p,q)$ is abelian with quotient isomorphic to $\mathbb{Z}$. So, in this case, $\mathrm{BS}(p,q)$ is solvable and therefore amenable. As we will discuss, the distinction between these two cases is of importance when working with random walks.

This paper is organised as follows. For the remainder of Section 1 and for Section 2, we will assume that the two non-zero integers $p$ and $q$ satisfy $1 \le p < q$. In Lemma 2.7 and in the subsequent results, we will restrict ourselves to the non-amenable subcase $1 < p < q$. Later, in the Appendix, we will investigate the remaining non-amenable cases.

### 1.1.2 Projection to the Bass-Serre tree

So, let us first assume that $1 \le p < q$. The Cayley graph $\Gamma$ of the group $G := \mathrm{BS}(p,q)$ with respect to the standard generators $a$ and $b$ is the quiver (= directed multigraph) with vertex set $G$, edge set $G \times \{a,b\}$, source function $s : G \times \{a,b\} \to G$ given by $s(g,x) := g$, and target function $t : G \times \{a,b\} \to G$ given by $t(g,x) := gx$. As usual, we label the edges by $a$ and $b$, respectively.

**Remark 1.1 ("quiver" vs. "graph")** *In contrast to a quiver, a graph is just a pair consisting of a vertex set and an edge set with the property that every edge is a two-element subset of the vertex set. Every quiver can be converted into a 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 graph, and we shall tacitly do so.*

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 $B := \langle b \rangle \le G$ and let $\mathbb{T}$ be the graph with vertex set $G/B = \{gB \mid g \in G\}$ and edge set $\{\{gB, gaB\} \mid g \in G\}$. This graph is actually a tree; it is obviously connected and, by Britton's Lemma, it does not contain a cycle. Notice that the canonical projection $\pi_{\mathbb{T}} : G \to G/B$ given by $\pi_{\mathbb{T}}(g) := gB$ is a weak graph homomorphism from $\Gamma$ to $\mathbb{T}$, i. e. whenever the vertices $g$ and $h$ are adjacent in $\Gamma$, their images $gB$ and $hB$ either agree or they are adjacent in $\mathbb{T}$.

**Remark 1.2 ("levels")** *Consider the infinite cyclic group $\mathbb{Z}$ and the map $\lambda : \{a,b\} \to \mathbb{Z}$ given by $\lambda(a) := 1$ and $\lambda(b) := 0$. The latter can be uniquely extended to a group homomorphism $\lambda : G \to \mathbb{Z}$. Indeed, the equation $\lambda(a) + p \cdot \lambda(b) - \lambda(a) = q \cdot \lambda(b)$ holds in $\mathbb{Z}$ so that we can apply von Dyck's Theorem to extend $\lambda$, see e. g. [Rot95, p. 346, fn. 2]. Since $\lambda(b) = 0$, the group homomorphism $\lambda : G \to \mathbb{Z}$ is constant on the cosets from $G/B$ and therefore induces a well-defined map $\widetilde{\lambda} : G/B \to \mathbb{Z}$ given by $\widetilde{\lambda}(gB) := \lambda(g)$. We shall think of $\lambda$ and $\widetilde{\lambda}$ as level functions, they assign a level to every vertex of $\Gamma$ and $\mathbb{T}$, respectively.*

**Lemma 1.3** *Every vertex $gB$ of $\mathbb{T}$ has exactly $p+q$ neighbours, $p$ of them are located one level below the vertex $gB$ and $q$ of them are located one level above it.*

*Proof.* By construction, the levels of two adjacent vertices always differ exactly by 1. The defining relation $ab^pa^{-1} = b^q$ ("$\Leftarrow$") and Britton's Lemma ("$\Rightarrow$") imply that $gaB = gb^raB$ if and only if $q \mid r$, whence the vertex $gB$ has exactly $q$ neighbours above. Similarly, it has exactly $p$ neighbours below because $ga^{-1}B = gb^ra^{-1}B$ if and only if $p \mid r$. $\square$

### 1.1.3 Projection to the hyperbolic plane

The second projection captures the information that is obtained by looking at $\Gamma$ from the front. In order to construct it, we introduce another group. Let $\mathrm{Aff}^+(\mathbb{R})$ be the set of all affine transformations of the real line that preserve the orientation, i. e. all maps $\varphi : \mathbb{R} \to \mathbb{R}$ of the form $\varphi(x) = \alpha x + \beta$ with $\alpha, \beta \in \mathbb{R}$ and $\alpha > 0$. This set, endowed with the composition $(\varphi_2 \circ \varphi_1)(x) := \varphi_2(\varphi_1(x)) = \alpha_2(\alpha_1 x + \beta_1) + \beta_2 = (\alpha_2\alpha_1)x + (\alpha_2\beta_1 + \beta_2)$, forms a group. Similarly to the construction of the level function $\lambda : G \to \mathbb{Z}$ in Remark 1.2, consider the map $\pi_{\mathrm{Aff}^+(\mathbb{R})} : \{a,b\} \to \mathrm{Aff}^+(\mathbb{R})$ given by $\pi_{\mathrm{Aff}^+(\mathbb{R})}(a) := (x \mapsto {}^q\!/_p \cdot x)$ and $\pi_{\mathrm{Aff}^+(\mathbb{R})}(b) := (x \mapsto x+1)$. Again, due to von Dyck's Theorem, it can be uniquely extended to a group homomorphism $\pi_{\mathrm{Aff}^+(\mathbb{R})} : G \to \mathrm{Aff}^+(\mathbb{R})$. The group $\mathrm{Aff}^+(\mathbb{R})$ has a geometric interpretation. In order to describe it, let $\mathbb{H}$ be the hyperbolic plane as per the Poincaré half-plane model, i. e. $\mathbb{H} = \{z \in \mathbb{C} \mid \Im(z) > 0\}$, endowed with the standard metric

$$d_{\mathbb{H}}(z_1,z_2) := \ln\left(\frac{|z_1 - \overline{z_2}| + |z_1 - z_2|}{|z_1 - \overline{z_2}| - |z_1 - z_2|}\right) = \operatorname{arcosh}\left(1 + \frac{|z_1 - z_2|^2}{2\Im(z_1)\Im(z_2)}\right),$$

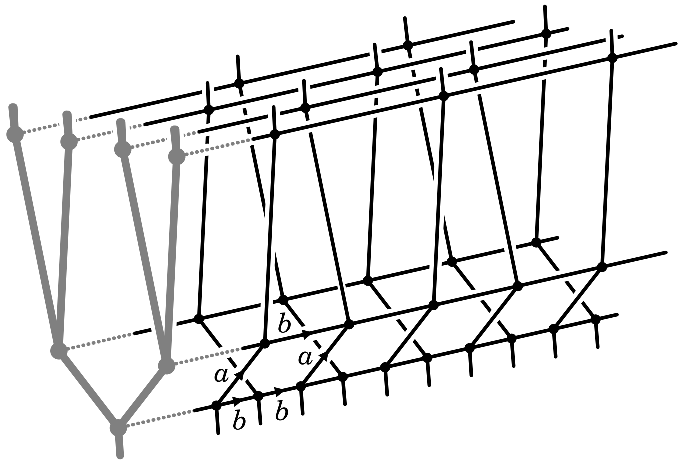


Figure 1: The Cayley graph $\Gamma$ of BS(1,2) with respect to the standard generators $a$ and $b$.

and notice that every $\varphi \in \mathrm{Aff}^+(\mathbb{R})$ can be thought of as an isometry of $\mathbb{H}$. Indeed, the isometries of $\mathbb{H}$ are precisely the maps $\varphi : \mathbb{H} \to \mathbb{H}$ of the form

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

see e. g. [Bea83, Theorem 7.4.1]. Now, we are prepared to construct the second projection $\pi_{\mathbb{H}} : G \to \mathbb{H}$. Pick an element $g \in G$, map it via $\pi_{\mathrm{Aff}^+(\mathbb{R})}$ to $\mathrm{Aff}^+(\mathbb{R})$, think of the latter as an isometry of $\mathbb{H}$, and evaluate it at $i \in \mathbb{H}$. Here are two lemmas to illustrate this construction and to record some of its properties.

**Lemma 1.4** *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 is $\ell_a := \ln q/p$. Similarly, for every $g \in G$ the point $\pi_{\mathbb{H}}(gb) \in \mathbb{H}$ is right from the point $\pi_{\mathbb{H}}(g) \in \mathbb{H}$; the two points have the same imaginary part and their distance is $\ell_b := \ln \frac{3+\sqrt{5}}{2}$. So, in some way, we are actually looking at $\Gamma$ from the front.*

*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 applying $\pi_{\mathrm{Aff}^+(\mathbb{R})}(g)$ to the points $\pi_{\mathbb{H}}(a) \in \mathbb{H}$ and $\pi_{\mathbb{H}}(1) \in \mathbb{H}$.[1] But, since $\pi_{\mathrm{Aff}^+(\mathbb{R})}(g)$ is the composition of a dilation $z \mapsto \alpha z$ and a translation $z \mapsto z + \beta$, the relative position of the two points is preserved. The same argument works for the second assertion, which completes the proof. $\square$

**Lemma 1.5** *The projection $\pi_{\mathbb{H}} : G \to \mathbb{H}$ is injective if and only if $p = 1$.*

*Proof.* If $p = 1$, it follows from the defining relation $aba^{-1} = b^q$ that every $g \in G$ can be expressed as $a^{-k} b^l a^m$ with exponents $k, m \in \mathbb{N}_0$ $(= \mathbb{N} \cup \{0\})$ and $l \in \mathbb{Z}$ such that $q \mid l \Rightarrow k = 0 \vee m = 0$. Now, it is not hard to see that the knowledge of the point $\pi_{\mathbb{H}}(g) = \pi_{\mathbb{H}}(a^{-k} b^l a^m) = q^m/q^k \cdot i + l/q^k \in \mathbb{H}$ allows us to reconstruct the exponents $k$, $l$, and $m$. Hence, we can also reconstruct $g \in G$, and the projection $\pi_{\mathbb{H}} : G \to \mathbb{H}$ is injective. On the other hand, if $p \geq 2$, we can make use of the fact that the Bass-Serre tree $\mathbb{T}$ branches both in upward and downward direction. More precisely, let $g := a^{-1} b^{-1} a b^{-1} a^{-1} b a b \in G$. Then, both $\pi_{\mathbb{H}}(g) = i$ and $\pi_{\mathbb{H}}(1) = i$. But, by Britton's Lemma, $g \neq 1$ in $G$. So, the projection $\pi_{\mathbb{H}} : G \to \mathbb{H}$ is not injective. $\square$

[1] Notice that the equation $(\varphi_2 \circ \varphi_1)(x) = \varphi_2(\varphi_1(x))$ remains true when replacing $x \in \mathbb{R}$ by $z \in \mathbb{H}$. Therefore, we may actually conclude that $\pi_{\mathbb{H}}(ga) = \pi_{\mathrm{Aff}^+(\mathbb{R})}(ga)(i) = (\pi_{\mathrm{Aff}^+(\mathbb{R})}(g) \circ \pi_{\mathrm{Aff}^+(\mathbb{R})}(a))(i) = \pi_{\mathrm{Aff}^+(\mathbb{R})}(g)(\pi_{\mathrm{Aff}^+(\mathbb{R})}(a)(i)) = \pi_{\mathrm{Aff}^+(\mathbb{R})}(g)(\pi_{\mathbb{H}}(a))$.

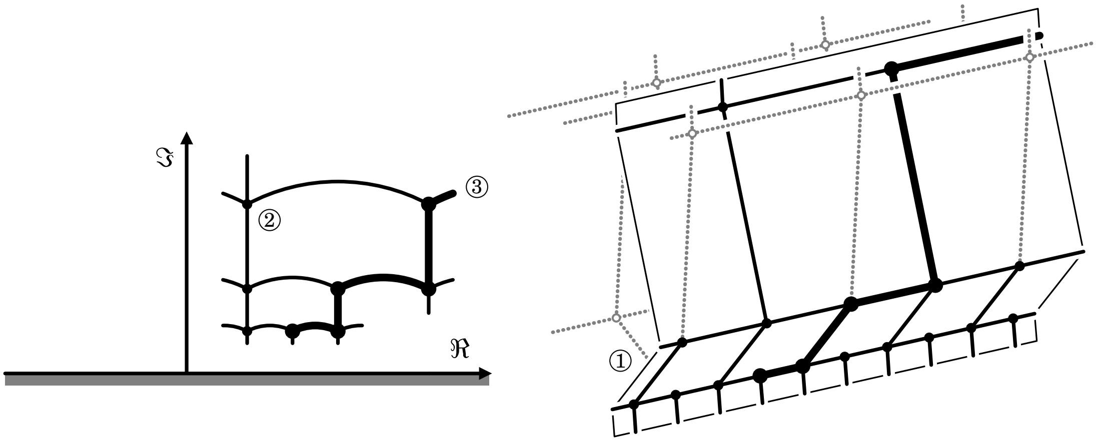


Figure 2: A part of a discrete hyperbolic plane $\Gamma_v$ and its projection to $\mathbb{H}$.

### 1.1.4 Discrete hyperbolic plane

**Definition 1.6 ("path", "reduced path")** *Given a graph with vertex set $V$, we consider finite paths $v : \{0,1,\dots,n\} \to V$, infinite paths $v : \mathbb{N}_0 \to V$, and doubly infinite paths $v : \mathbb{Z} \to V$. In any case, being a path means that for every possible $k$ the vertices $v(k)$ and $v(k+1)$ are adjacent in the graph. Moreover, we say that a path is reduced if for every possible $k$ the vertices $v(k)$ and $v(k+2)$ are distinct.*

Fix an ascending doubly infinite path $v : \mathbb{Z} \to G/B$ in the tree $\mathbb{T}$. Here, the word ascending refers to the level function defined in Remark 1.2, and it means that for every $k \in \mathbb{Z}$ the vertex $v(k)$ is located above the preceding vertex $v(k-1)$. Now, let $G_v$ be the full $\pi_{\mathbb{T}}$-preimage of $v$, i. e. the set consisting of all $g \in G$ such that the image $\pi_{\mathbb{T}}(g)$ is traversed by $v$. The subgraph $\Gamma_v \le \Gamma$ spanned by $G_v$, see ① in Figure 2, is obviously connected so that the graph distance $d_{\Gamma_v}$ becomes a metric. This subgraph is sometimes referred to as discrete hyperbolic plane or plane of bricks, which makes particular sense in light of Lemma 1.7.

**Lemma 1.7** *The restriction $\pi_{\mathbb{H}}|_{G_v} : G_v \to \mathbb{H}$ is a quasi-isometry between the graph $\Gamma_v$, endowed with the graph distance $d_{\Gamma_v}$, and the hyperbolic plane $\mathbb{H}$, endowed with the standard metric $d_{\mathbb{H}}$.*

*Proof.* We realise the edges of the graph $\Gamma_v$ geometrically. Whenever two vertices $g, h \in G_v$ are adjacent, we connect their images $\pi_{\mathbb{H}}(g) \in \mathbb{H}$ and $\pi_{\mathbb{H}}(h) \in \mathbb{H}$ by a geodesic in $\mathbb{H}$. In order to avoid confusion, we refer to these images as $\mathbb{H}$-vertices and to the geodesics between them as $\mathbb{H}$-edges. By the proof of Lemma 1.3 and by Lemma 1.4, the $\mathbb{H}$-vertices and $\mathbb{H}$-edges define a tessellation of the hyperbolic plane with isometric bricks of the following shape. The $\mathbb{H}$-vertices of each brick are located on two distinct horizontal lines; on the upper one there are $p+1$ of them and on the lower one there are $q+1$ of them. In either case, the $\mathbb{H}$-vertices are connected by $\mathbb{H}$-edges of length $\ell_b$ to form a chain (= piecewise geodesic curve). Due to the curvature, both the two leftmost and the two rightmost $\mathbb{H}$-vertices are located exactly above each other and connected by vertical $\mathbb{H}$-edges of length $\ell_a$, see ② in Figure 2 and Figure 3. Since the bricks are uniformly bounded and cover the hyperbolic plane $\mathbb{H}$, the restriction $\pi_{\mathbb{H}}|_{G_v} : G_v \to \mathbb{H}$ is certainly quasi-surjective.

Pick any two vertices $g, h \in G_v$. We aim to estimate the distances $d_{\Gamma_v}(g,h)$ and $d_{\mathbb{H}}(\pi_{\mathbb{H}}(g), \pi_{\mathbb{H}}(h))$ by multiples of each other. So, choose a path of minimal length from $g$ to $h$ in $\Gamma_v$. It corresponds to a chain of $\mathbb{H}$-edges from $\pi_{\mathbb{H}}(g)$ to $\pi_{\mathbb{H}}(h)$, see ③ in Figure 2. This chain consists of $d_{\Gamma_v}(g,h)$ many $\mathbb{H}$-edges, each of

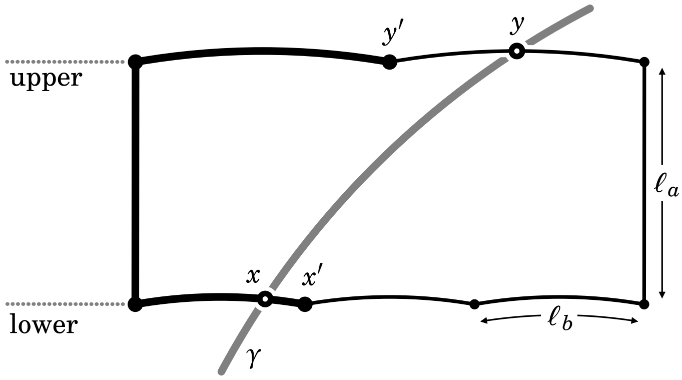


Figure 3: Approximation of the geodesic $\gamma$ in the case $p = 2$ and $q = 3$.

which has length at most $\max\{\ell_a, \ell_b\}$. Hence,

$$d_{\mathbb{H}}(\pi_{\mathbb{H}}(g), \pi_{\mathbb{H}}(h)) \leq d_{\Gamma_v}(g,h) \cdot \max\{\ell_a, \ell_b\}.$$

On the other hand, let us make the following auxiliary definition. Every point $x \in \mathbb{H}$ that is not in the interior of a brick is either an $\mathbb{H}$-vertex, in which case we define $x'$ to be $x$, or it is in the interior of an $\mathbb{H}$-edge, in which case we define $x'$ to be one of the endpoints of the $\mathbb{H}$-edge, whichever is closer. In the case that $x$ is exactly in the middle of the $\mathbb{H}$-edge, we choose the left endpoint rather than the right one and the lower endpoint rather than the upper one. With this notion in mind, consider the geodesic $\gamma$ from $\pi_{\mathbb{H}}(g)$ to $\pi_{\mathbb{H}}(h)$, see Figure 3. Whenever $\gamma$ traverses the interior of a brick $B$, it enters the interior at some point $x \in \partial B$ and leaves it at some other point $y \in \partial B$. In this situation, approximate the part of $\gamma$ from $x$ to $y$ by a chain of $\mathbb{H}$-edges from $x'$ to $y'$. We may choose this chain such that, whenever $x' = y'$, the chain has no $\mathbb{H}$-edge at all and, otherwise, the number of $\mathbb{H}$-edges in the chain is at most $c := \lfloor 1/2 \cdot (p+q+2) \rfloor$. But, by a compactness argument, there is an $\varepsilon > 0$ such that if the part of $\gamma$ has length smaller than $\varepsilon$, then $x' = y'$ and the chain has no $\mathbb{H}$-edge at all. Therefore, we may conclude that

$$\text{number of } \mathbb{H}\text{-edges in the chain} \leq \frac{c}{\varepsilon} \cdot \text{length of the part of } \gamma.$$

It is not hard to see that if we do this for every brick $B$ whose interior is traversed by $\gamma$, we finally obtain a chain of $\mathbb{H}$-edges from $\pi_{\mathbb{H}}(g)$ to $\pi_{\mathbb{H}}(h)$. Depending on whether a part of $\gamma$ originally traversed the interior of a brick or ran along an $\mathbb{H}$-edge, we may estimate the number of $\mathbb{H}$-edges approximating it by $c/\varepsilon$, by $1/\ell_a$, or by $1/\ell_b$ times its length. Hence,

$$d_{\Gamma_v}(g,h) \leq d_{\mathbb{H}}(\pi_{\mathbb{H}}(g), \pi_{\mathbb{H}}(h)) \cdot \max\left\{\frac{c}{\varepsilon}, \frac{1}{\ell_a}, \frac{1}{\ell_b}\right\}.$$

□

**Remark 1.8** *Notice that the horizontal lines mentioned in the proof of Lemma 1.7 are horospheres, and by no means geodesics. For example, one may pick such a horizontal line and observe that the part of the line contained in the ball $B(i,n) \subseteq \mathbb{H}$ with centre $i \in \mathbb{H}$ and radius $n \in \mathbb{N}$ has a length growing exponentially in $n$, see also Figure 12 on page 42.*

**Remark 1.9** *The projections $\pi_{\mathbb{T}} : G \to G/B$ and $\pi_{\mathbb{H}} : G \to \mathbb{H}$ are certainly not independent, e.g. the level of a vertex $g \in G$ can be reconstructed both from $\pi_{\mathbb{T}}(g) \in G/B$ and from $\pi_{\mathbb{H}}(g) \in \mathbb{H}$. In fact, the image of $\pi_{\mathbb{T}} \times \pi_{\mathbb{H}} : G \to G/B \times \mathbb{H}$ is contained in the horocyclic product of the tree $\mathbb{T}$ and the hyperbolic plane $\mathbb{H}$, which is sometimes referred to as treebolic space, see [BSCSW] for details.*

### 1.1.5 Compactifications

Both the tree $\mathbb{T}$ and the hyperbolic plane $\mathbb{H}$ have a natural compactification. In case of $\mathbb{T}$ it is the end compactification, which can be constructed as follows. Fix a base point, say $B \in G/B$, and consider the set $\widehat{\mathbb{T}}$ 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 in $\widehat{\mathbb{T}}$ and the vertices $G/B$. We may therefore think of $G/B$ as a subset of $\widehat{\mathbb{T}}$. The set $\widehat{\mathbb{T}}$ can be endowed with the metric

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

Here, the symbol $|x \wedge y|$ denotes the number of edges the two paths run together until they separate, i. e. $|x \wedge y| = \max\{n \in \mathbb{N}_0 \mid x(n) \text{ and } y(n) \text{ are both defined and } x(n) = y(n)\}$, see ① in Figure 4. Hence, the later the paths separate the closer they are. The set $\widehat{\mathbb{T}}$, endowed with the metric $d_{\widehat{\mathbb{T}}}$, 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 in $\widehat{\mathbb{T}}$, it is usually denoted by $\partial\mathbb{T}$ and called the space of ends.

In case of $\mathbb{H}$, we could simply construct the closure of $\mathbb{H}$ after identifying it as a subspace of the Riemann sphere $\mathbb{C} \cup \{\infty\}$, which is the one-point compactification of the complex plane $\mathbb{C}$. But, in order to visualise the compactification of $\mathbb{H}$ a little better, we temporarily switch to the Poincaré disc model. More precisely, instead of working in the half-plane $\mathbb{H} = \{z \in \mathbb{C} \mid \Im(z) > 0\}$, we consider the open unit disc $\mathbb{D} := \{z \in \mathbb{C} \mid |z| < 1\}$. The Cayley transform $W : \mathbb{H} \rightarrowtail\!\!\!\!\twoheadrightarrow \mathbb{D}$ given by $W(z) := \frac{z-i}{z+i}$ is one possibility to convert between the two models. Since we are currently interested in the topological structure, let us underline that the hyperbolic topology on $\mathbb{D}$ is the one induced by the Cayley transform, i. e. the one that turns the Cayley transform into a homeomorphism. It agrees with the standard topology on $\mathbb{D}$. So, 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 $\widehat{\mathbb{D}} := \{z \in \mathbb{C} \mid |z| \leq 1\}$. 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 $W : \mathbb{H} \cup \mathbb{R} \cup \{\infty\} \rightarrowtail\!\!\!\!\twoheadrightarrow \widehat{\mathbb{D}}$, and then apply its inverse. The resulting 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 usually 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.

**Lemma 1.10** *A sequence $(x_0, x_1, \ldots)$ in $\mathbb{H}$ converges to $\infty \in \partial\mathbb{H}$ if and only if the absolute values $|x_n|$ tend to infinity. Moreover, it converges to a point $\xi \in \partial\mathbb{H} \smallsetminus \{\infty\}$ if and only if it does with respect to the standard topology on the complex plane $\mathbb{C}$.*

*Proof.* By definition, a sequence $(x_0, x_1, \ldots)$ in $\mathbb{H}$ converges to a point $\xi \in \partial\mathbb{H}$ if and only if its pointwise image $(W(x_0), W(x_1), \ldots)$ in $\mathbb{D}$ converges to $W(\xi) \in \partial\mathbb{D}$. But, as mentioned above, the latter can be verified using the standard topology on the complex plane $\mathbb{C}$. In particular, the sequence $(x_0, x_1, \ldots)$ converges to $\infty \in \partial\mathbb{H}$ if and only if

$$|W(x_n) - W(\infty)| = \left| \frac{x_n - i}{x_n + i} - \frac{\infty - i}{\infty + i} \right| = \frac{2}{|x_n + i|} \xrightarrow{n \to \infty} 0 ,$$

which holds if and only if the absolute values $|x_n|$ tend to infinity. In order to prove the second assertion, let us fix, once and for all, the standard topology on the complex plane $\mathbb{C}$ and observe that we may again adjust both the domain and the codomain of the Cayley transform so that we obtain a bijection $W : \mathbb{C} \smallsetminus \{-i\} \rightarrowtail\!\!\!\!\twoheadrightarrow \mathbb{C} \smallsetminus \{1\}$. This is a homeomorphism between subspaces of the complex plane $\mathbb{C}$, whence $W(x_n) \to W(\xi)$ if and only if $x_n \to \xi$. $\square$

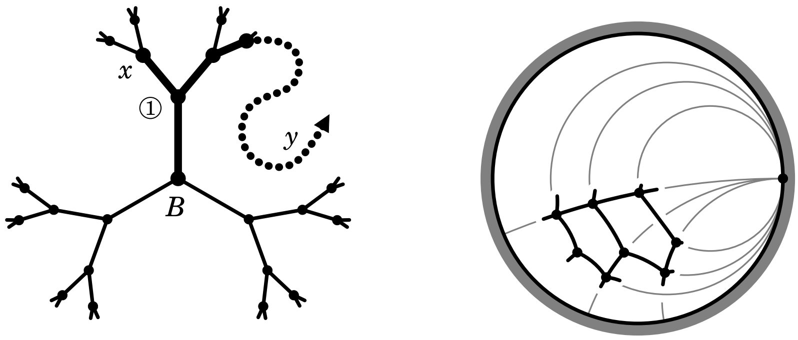


Figure 4: The space of ends and the hyperbolic boundary in the Poincaré disc model.

## 1.2 Random walks on groups

### 1.2.1 Discrete Markov chains

In this paper, we aim to study random walks on Baumslag-Solitar groups. Let us therefore recall some basic notions from the theory of random walks. Given a countable state space $X$, an initial probability measure[2] $\vartheta : X \to [0,1]$, and transition probabilities $p : X \times X \to [0,1]$, we are interested in the Markov chain $Z = (Z_0, Z_1, \dots)$ that starts according to $\vartheta$ and proceeds according to $p$.

In order to construct $Z$ rigorously, take the space of trajectories $\Omega := \{(x_0, x_1, \dots) \mid \forall n \in \mathbb{N}_0 : x_n \in X\}$. For every $n \in \mathbb{N}_0$ there is a projection $Z_n : \Omega \to X$ given by $Z_n(x_0, x_1, \dots) := x_n$. Now, endow $\Omega$ with the smallest $\sigma$-algebra $\mathscr{A}$ having the property that all projections $Z_n$ are measurable. It is called the product $\sigma$-algebra. An equivalent way to define $\mathscr{A}$ is via cylinder sets. First, fix some time $t \in \mathbb{N}_0$. Given a tuple $a = (a_0, a_1, \dots, a_t) \in X^{t+1}$, the associated cylinder set $\mathrm{C}(a)$ consists of all trajectories that start according to $a$ and continue arbitrarily, i. e. $\mathrm{C}(a) := \{(a_0, a_1, \dots, a_t, x_{t+1}, x_{t+2}, \dots) \mid \forall n \geq t+1 : x_n \in X\}$. Let $\mathscr{A}_t$ be the $\sigma$-algebra generated by all cylinder sets that are determined until time $t$, i. e. $\mathscr{A}_t := \sigma(\{\mathrm{C}(a) \mid a \in X^{t+1}\})$. Observe that $\mathscr{A}_0 \subseteq \mathscr{A}_1 \subseteq \dots$ is an increasing sequence of $\sigma$-algebras. Their union $\mathscr{F} := \bigcup_{t \in \mathbb{N}_0} \mathscr{A}_t$ is an algebra; it generates the $\sigma$-algebra $\mathscr{A}$, i. e. $\mathscr{A} := \sigma(\mathscr{F})$. The latter definition is certainly more involved but it facilitates the construction of a probability measure. For each time $t \in \mathbb{N}_0$ there is a probability measure $\nu_t$ on the $\sigma$-algebra $\mathscr{A}_t$ given by

$$\nu_t(A) := \sum_{\substack{a=(a_0,a_1,\dots,a_t) \\ \text{with } \mathrm{C}(a) \in A}} \vartheta(a_0) \cdot p(a_0, a_1) \cdot p(a_1, a_2) \cdot \ldots \cdot p(a_{t-1}, a_t).$$

These probability measures $\nu_t$ are compatible in the sense that they assemble to a content $\nu$ with total mass 1 on the algebra $\mathscr{F}$. One can even show $\sigma$-additivity, whence $\nu$ is a premeasure and, by Carathéodory's extension theorem, uniquely extends to a probability measure $\mathbb{P}$ on $\mathscr{A}$.

We may therefore consider the probability space $(\Omega, \mathscr{A}, \mathbb{P})$. Doing so, the projections $Z_n : \Omega \to X$ become random variables that constitute the Markov chain. For details on the terminology used above, see e. g. [Kle14, §1], and for a gentle introduction to discrete Markov chains, see e. g. [Woe09, §1]. From now on, we will always use the term random walk instead of Markov chain.

[2] Here, and throughout the paper, we use the term probability measure to denote probability measures and probability mass functions on discrete spaces.

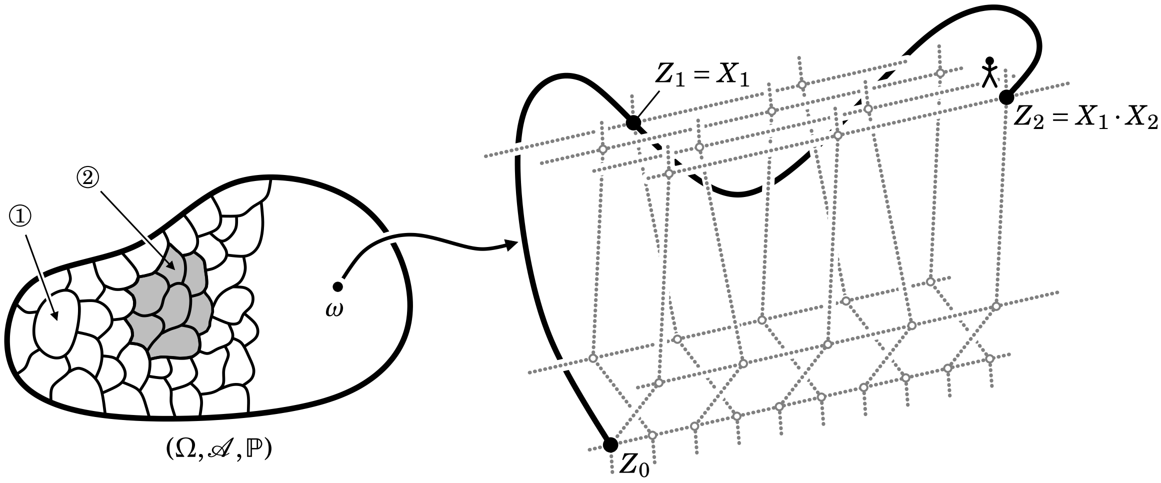


Figure 5: The first steps of a random walk on BS(1,2).

### 1.2.2 From countable sets to countable groups

In Section 1.2.1, we considered a countable state space $X$ without any structure. Now, we assume that $X$ is a countable group $G$, in which case we may study random walks whose transition probabilities are adapted to the group structure. In order to do so, we first pick a probability measure $\mu : G \to [0,1]$ whose support $\mathrm{supp}(\mu) = \{g \in G \mid \mu(g) > 0\}$ generates $G$ as a semigroup. Then, we consider the random walk given by the following data. The initial probability measure $\vartheta : G \to [0,1]$ puts all mass on the identity element $1 \in G$ and the transition probabilities $p : G \times G \to [0,1]$ 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 a. s. (= almost surely) at the identity element and has independent $\mu$-distributed increments each of which is multiplied from the right to the current state. Therefore, $Z_0 = 1$ a. s. and for every $n \in \mathbb{N}$ we may decompose $Z_n = X_1 \cdot \ldots \cdot X_n$, where $X_1, X_2, \ldots$ is a sequence of independent $\mu$-distributed random variables, see the right-hand side of Figure 5.

**Remark 1.11** *Since we assume that* $\mathrm{supp}(\mu)$ *generates G as a semigroup, the random walk is irreducible, i. e. any two states can be reached from each other with positive probability. In particular, the following dichotomy holds. Either every state is recurrent, i. e. the return probability is equal to 1, or every state is transient, i. e. the return probability is smaller than 1. In the latter case, the probability that every finite set of states will eventually be left and the random walker escapes to infinity is equal to 1. Roughly speaking, the Poisson-Fürstenberg boundary, which we are about to identify in this paper, serves as a probabilistic model to describe the long-time behaviour of the random walk in more detail.*

### 1.2.3 Moment conditions

As soon as $\mathrm{supp}(\mu)$ is infinite, it is harder to understand the behaviour of the random walk. In this situation, we often need to assume that the probability of huge jumps is sufficiently low. The notion of moments helps us to make this assumption rigorous. Recall that, given a probability space, e. g. $(\Omega, \mathscr{A}, \mathbb{P})$ constructed in Section 1.2.1, and a random variable $X : \Omega \to \mathbb{R}$, the latter has finite first moment if $\int |X| \,\mathrm{d}\mathbb{P} < \infty$. In this case both $\int X^+ \,\mathrm{d}\mathbb{P} < \infty$ and $\int X^- \,\mathrm{d}\mathbb{P} < \infty$, and we are able to define the expectation $\mathbb{E}(X) := \int X^+ \,\mathrm{d}\mathbb{P} - \int X^- \,\mathrm{d}\mathbb{P}$. Of course, the difference would still make sense if only one of the two integrals was finite. But this is not of relevance for us. So, when writing "$\mathbb{E}(X)$" we implicitly mean that $-\infty < \mathbb{E}(X) < \infty$. More generally, given any non-negative $k \in \mathbb{R}$, the random variable $X : \Omega \to \mathbb{R}$ has finite $k$-th moment if $\int |X|^k \,\mathrm{d}\mathbb{P} < \infty$. In our setting, the increments $X_1, X_2, \ldots$ take values in $G$, whence we need to define the hugeness of a jump before talking about their moments.

**Definition and Lemma 1.12 ("word metric", "finite k-th moment")** *If $G$ is a finitely generated group and $S \subseteq G$ is a finite generating set, then the word metric $d_{\mathrm{S}}$ on $G$ is given by*

$$d_{\mathrm{S}}(g,h) := \min\{n \in \mathbb{N}_0 \mid \exists s_1,\ldots,s_n \in S : \exists \varepsilon_1,\ldots,\varepsilon_n \in \{1,-1\} : g^{-1}h = {s_1}^{\varepsilon_1} \cdot \ldots \cdot {s_n}^{\varepsilon_n}\}.$$

*Notice that the word metric coincides with the distance in the respective Cayley graph. Using the word metric $d_{\mathrm{S}}$ on $G$, a random variable $X : \Omega \to G$ has finite $k$-th moment if the image $d_{\mathrm{S}}(1,X) : \Omega \to \mathbb{R}$ has finite $k$-th moment in the classical sense, i.e. if $\int |d_{\mathrm{S}}(1,X)|^k \,\mathrm{d}\mathbb{P} = \int d_{\mathrm{S}}(1,X)^k \,\mathrm{d}\mathbb{P} < \infty$. This property does not depend on the choice of $S \subseteq G$.*

*Proof.* Let $S, T \subseteq G$ be finite generating sets and assume that $X$ has finite $k$-th moment with respect to $S$. It is a basic fact from geometric group theory, see e.g. [Mei08, Lemma 11.37], that the identity map between the metric spaces $(G, d_{\mathrm{S}})$ and $(G, d_{\mathrm{T}})$ is a quasi-isometry. In particular, there are $a, b \in \mathbb{R}$ such that for all $g, h \in G$ the inequality $d_{\mathrm{T}}(g,h) \le a\, d_{\mathrm{S}}(g,h) + b$ holds. Now, given $a$ and $b$, we can also find $A, B \in \mathbb{R}$ such that for all $g, h \in G$ the inequalities $d_{\mathrm{T}}(g,h)^k \le (a\, d_{\mathrm{S}}(g,h) + b)^k \le A\, d_{\mathrm{S}}(g,h)^k + B$ hold. Hence,

$$\int d_{\mathrm{T}}(1,X)^k \,\mathrm{d}\mathbb{P} \le \int A\, d_{\mathrm{S}}(1,X)^k + B \,\mathrm{d}\mathbb{P} \le A \cdot \underbrace{\int d_{\mathrm{S}}(1,X)^k \,\mathrm{d}\mathbb{P}}_{<\infty} + B \cdot \underbrace{\int \mathrm{d}\mathbb{P}}_{=1} < \infty.$$

$\square$

### 1.2.4 Further notions concerning random walks on Baumslag-Solitar groups

Let us now return to the situation we are interested in, namely that $G = \mathrm{BS}(p,q)$ with $1 \le p < q$. When working with the projection $\pi_{\mathbb{H}} : G \to \mathbb{H}$, we often consider the imaginary parts $\Im(\pi_{\mathbb{H}}(g))$ and the real parts $\Re(\pi_{\mathbb{H}}(g))$ separately, and it is convenient to abbreviate the former by $A_g$ and the latter by $B_g$. Occasionally, we do not need to assume that $X_1$ has some finite moment but impose this assumption only on the images $\ln(A_{X_1})$ and $\ln(1 + |B_{X_1}|)$. The following lemma relates these two situations.

**Lemma 1.13** *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 1.14** *Before we prove Lemma 1.13, notice that for every $g \in G$ the imaginary part $A_g$ can be expressed in terms of the level $\lambda(g)$, namely by the formula $A_g = (q/p)^{\lambda(g)}$. Taking the logarithm on both sides yields $\ln(A_g) = \ln(q/p) \cdot \lambda(g)$. So, instead of thinking of $\ln(A_g)$ we may think of a multiple of $\lambda(g)$.*

*Proof of Lemma 1.13.* Let $S := \{a, b\} \subseteq G$ be the standard generating set. We may now estimate

$$\int |\ln(A_{X_1})|^k \,\mathrm{d}\mathbb{P} = \ln(q/p)^k \cdot \int |\lambda(X_1)|^k \,\mathrm{d}\mathbb{P} \le \ln(q/p)^k \cdot \underbrace{\int d_{\mathrm{S}}(1,X_1)^k \,\mathrm{d}\mathbb{P}}_{<\infty} < \infty.$$

Concerning the second assertion, observe that $d_{\mathbb{H}}(\pi_{\mathbb{H}}(1), \pi_{\mathbb{H}}(g)) \le \max\{\ell_a, \ell_b\} \cdot d_{\mathrm{S}}(1,g)$, which can be shown by the same argument as in the proof of Lemma 1.7. This observation allows us to estimate $\ln(1 + |B_g|)$ by a multiple of $d_{\mathrm{S}}(1,g)$. Indeed,

$$\ln(1 + |B_g|) \le \ln\left(1 + 1/2 \cdot |B_g|^2 + \sqrt{(1 + 1/2 \cdot |B_g|^2)^2 - 1}\right) = \operatorname{arcosh}\left(1 + 1/2 \cdot |B_g|^2\right) = d_{\mathbb{H}}(i, i + B_g)$$

$$\le d_{\mathbb{H}}(i, A_g \cdot i + B_g) + d_{\mathbb{H}}(A_g \cdot i + B_g, i + B_g) = d_{\mathbb{H}}(\pi_{\mathbb{H}}(1), \pi_{\mathbb{H}}(g)) + |\ln(A_g)|$$

$$\le \max\{\ell_a, \ell_b\} \cdot d_S(1,g) + \ln(q/p) \cdot |\lambda(g)| \le \max\{\ell_a, \ell_b\} \cdot d_S(1,g) + \ln(q/p) \cdot d_S(1,g).$$

Therefore,

$$\int \ln(1 + |B_{X_1}|)^k \, \mathrm{d}\mathbb{P} \le (\max\{\ell_a, \ell_b\} + \ln(q/p))^k \cdot \underbrace{\int d_S(1, X_1)^k \, \mathrm{d}\mathbb{P}}_{<\infty} < \infty.$$

□

In addition to the moments of $\ln(A_{X_1})$ and $\ln(1 + |B_{X_1}|)$, we will use the notion of vertical drift. Consider a random walk $Z = (Z_0, Z_1, \dots)$ on $G$ and its pointwise projection $\lambda(Z) = (\lambda(Z_0), \lambda(Z_1), \dots)$ to the levels. Since $\lambda(Z_n) = \lambda(X_1 \cdot \ldots \cdot X_n) = \lambda(X_1) + \ldots + \lambda(X_n)$, these projections constitute a random walk on the set of integers with i. i. d. (= independent and identically distributed) increments.

**Definition 1.15 (“vertical drift”)** *If $\ln(A_{X_1})$ has finite first moment, then $\lambda(X_1)$ has finite first moment and we are able to define the expectation $\mathbb{E}(\lambda(X_1))$. The latter is called the vertical drift and denoted by $\delta$. We will distinguish between positive vertical drift, i. e. $\delta > 0$, negative vertical drift, i. e. $\delta < 0$, and no vertical drift, i. e. $\delta = 0$, which is the most subtle of the three cases.*

## 1.3 Poisson-Fürstenberg boundary

### 1.3.1 Definition and basic examples of Lebesgue-Rohlin spaces

As mentioned in Remark 1.11, the Poisson-Fürstenberg boundary serves as a probabilistic model to describe the long-time behaviour of a random walk. In order to define it, we need to ensure that we are working with Lebesgue-Rohlin spaces. In this section, we recall their definition and basic examples. Further details can be found e. g. in [Roh52], [Hae73], [Rud90]. Moreover, let us mention the collection of facts in [KKR04, Appendix] and the more informal introduction in [CK12].

**Definition 1.16 (“separable probability space”)** *A probability space $(\Omega, \mathscr{A}, \mathbb{P})$ is called separable if it is ① complete, ② countably generated, i. e. there is a subset $A = \{A_1, A_2, \dots\} \subseteq \mathscr{A}$ such that the $\sigma$-algebra generated by $A$ and completed with respect to $\mathbb{P}$ is the $\sigma$-algebra $\mathscr{A}$, and ③ separable in the narrow sense, i. e. for any two $x, y \in \Omega$ there is an index $k \in \mathbb{N}$ such that either $x \in A_k$ and $y \notin A_k$ or $x \notin A_k$ and $y \in A_k$. In this case, the subset $A \subseteq \mathscr{A}$ is called a base of $(\Omega, \mathscr{A}, \mathbb{P})$.*

Let us fix a separable probability space $(\Omega, \mathscr{A}, \mathbb{P})$ with base $A = \{A_1, A_2, \dots\} \subseteq \mathscr{A}$ and consider the map $i_A : \Omega \to \{0,1\}^{\mathbb{N}}$ given by $i_A(x) := (\mathbb{1}_{A_1}(x), \mathbb{1}_{A_2}(x), \dots)$. Here, $\mathbb{1}_{A_k}$ denotes the characteristic function of the set $A_k$. Let $\mathscr{B}$ be the product $\sigma$-algebra on $\{0,1\}^{\mathbb{N}}$. The map $i_A$ is $\sigma(A)$-$\mathscr{B}$-measurable, i. e. $i_A$ is a measurable map from $(\Omega, \sigma(A))$ to $(\{0,1\}^{\mathbb{N}}, \mathscr{B})$, see [Hae73, Proposition 1]. We may therefore consider the pushforward probability measure $Q := i_A(\mathbb{P})$, and construct the completion $\mathscr{B}_Q$ of the $\sigma$-algebra $\mathscr{B}$ with respect to $Q$. It is not hard to see that the map $i_A$ is also $\mathscr{A}$-$\mathscr{B}_Q$-measurable, see [Hae73, Proposition 2]. Moreover, by separability in the narrow sense, it is certainly injective. If it is also surjective, then the separable probability space $(\Omega, \mathscr{A}, \mathbb{P})$ with base $A \subseteq \mathscr{A}$ is called “complete”. A weaker notion is the one of “almost completeness”.

**Definition 1.17 (“almost complete”)** *A separable probability space $(\Omega, \mathscr{A}, \mathbb{P})$ with base $A \subseteq \mathscr{A}$ is called almost complete if the image $i_A(\Omega) \subseteq \{0,1\}^{\mathbb{N}}$ is $\mathscr{B}_Q$-measurable, i. e. if $i_A(\Omega) \in \mathscr{B}_Q$.*

It turns out that almost completeness is an intrinsic property of the separable probability space, and does not depend on the choice of the base, see [Hae73, Proposition 4].

**Definition 1.18 ("Lebesgue-Rohlin space")** *A separable probability space $(\Omega, \mathscr{A}, \mathbb{P})$ which is almost complete is called a Lebesgue-Rohlin space.*

The most prominent examples of Lebesgue-Rohlin spaces are discrete probability spaces and the unit interval $[0,1]$ equipped with the Lebesgue $\sigma$-algebra $\mathscr{L}$ and the Lebesgue measure $\lambda$. In fact, every Lebesgue-Rohlin space is isomorphic[3] either to one of these examples or to the disjoint union of an interval $[0,\alpha]$ with $0 < \alpha \leq 1$ and countably many atoms with total mass $1-\alpha$, see [Roh52, §2.4].

**Remark 1.19 ("Polish spaces")** *A Polish space is a topological space that is separable, i. e. contains a countable and dense subset, and completely metrisable, i. e. there is a metric that induces the topology and turns the space into a complete metric space. All Polish spaces equipped with the Borel $\sigma$-algebra $\mathscr{B}$ and a Borel measure $\mu$ become after completion examples of Lebesgue-Rohlin spaces, see [Roh52, §2.7] and [Hae73, p. 248, Example 1].*

**Example 1.20 ("Variation of the perforated interval")** *A separable probability space that is not a Lebesgue-Rohlin space can be obtained as follows. Consider the half-closed interval $[0,1)$ and fix an irrational number $\alpha \in \mathbb{R} \setminus \mathbb{Q}$. The transformation $\tau : [0,1) \to [0,1)$ given by $\tau(x) := x + \alpha \mod 1$ has the property that for every point $x \in [0,1)$ the bidirectional orbit $\{\tau^n(x) \mid n \in \mathbb{Z}\}$ is countably infinite. Consider all these orbits and use the axiom of choice to pick one representative from each of them. Let A be the union of these representatives. We claim that A is non-measurable with respect to the Lebesgue measure. Indeed, if A was measurable, then the translation invariance would imply that for every $n \in \mathbb{Z}$ the values $\lambda(\tau^n(A))$ and $\lambda(A)$ agree. But, by construction, the interval $[0,1)$ is the disjoint union of all $\tau^n(A)$, whence $1 = \lambda([0,1)) = \sum_{n\in\mathbb{Z}} \lambda(\tau^n(A)) = \sum_{n\in\mathbb{Z}} \lambda(A)$, and each of the assumptions $\lambda(A) = 0$ and $\lambda(A) > 0$ yields an immediate contradiction. The same argument shows that A has inner Lebesgue measure $\lambda_*(A) = 0$. Hence, we know that the complement $B := [0,1) \setminus A$ has outer Lebesgue measure $\lambda^*(B) = 1$. Now, equip the set B with the induced $\sigma$-algebra $\mathscr{L}_B$ and the measure $\beta$ that assigns to every $S' \in \mathscr{L}_B$ the Lebesgue measure of an arbitrary $S \in \mathscr{L}$ with $S \cap B = S'$. Since $\lambda^*(B) = 1$, the outcome does not depend on the choice of $S \in \mathscr{L}$. The probability space $(B, \mathscr{L}_B, \beta)$ constructed this way can be equipped with the induced base and therefore inherits separability from $([0,1), \mathscr{L}, \lambda)$. But, since B is non-measurable in $[0,1)$, its image cannot be measurable in $\{0,1\}^{\mathbb{N}}$.*

**Remark 1.21** *As illustrated on the left-hand side of Figure 6, we may use $x \mapsto \exp(2\pi i x)$ to identify the half-closed interval $[0,1)$ with the sphere $\mathbb{S}^1$, in which case the transformation $\tau$ becomes a rotation by an irrational multiple of $2\pi$.*

### 1.3.2 Definition of the Poisson-Fürstenberg boundary

In light of Remark 1.19, we may observe that the space of trajectories $\Omega$ introduced in Section 1.2.1 is the product $X^{\mathbb{N}_0}$ and can therefore be equipped with the product topology. One can show that the latter is actually a Polish space, see e. g. [Wil70, Theorem 24.11]. Since its Borel $\sigma$-algebra agrees with the $\sigma$-algebra $\mathscr{A}$, the completion of $(\Omega, \mathscr{A}, \mathbb{P})$ is a Lebesgue-Rohlin space. From now on, let us assume

[3] We shall consider probability spaces up to subsets of measure 0. So, we actually mean isomorphic mod 0. Recall that two probability spaces $(\Omega_1, \mathscr{A}_1, \mathbb{P}_1)$ and $(\Omega_2, \mathscr{A}_2, \mathbb{P}_2)$ are isomorphic mod 0 if there are null sets $N_k \subseteq \Omega_k$ with $k \in \{1,2\}$ and a bijection $\varphi : \Omega_1 \setminus N_1 \to \Omega_2 \setminus N_2$ which is measurable and measure preserving in both directions.

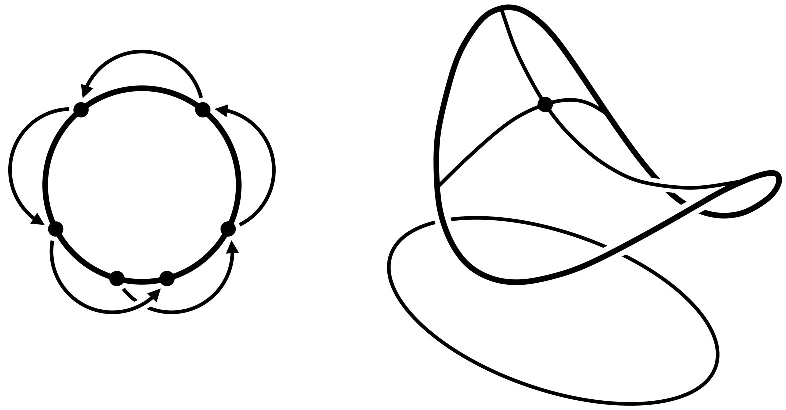

Figure 6: A rotation (left) and a harmonic extension of a continuous function (right).

that, as soon as a measurable space is equipped with a probability measure, we are working with the completion. So, abusing notation, we say that $(\Omega,\mathscr{A},\mathbb{P})$ is a Lebesgue-Rohlin space.

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

$$x\sim y \;:\Longleftrightarrow\; \exists\, t_1,t_2\in\mathbb{N}_0 : \forall\, n\in\mathbb{N}_0 : x_{t_1+n}=y_{t_2+n}\,.$$

Notice that we allow the times $t_1$ and $t_2$ to be different. If we did not, we would end up with the tail boundary instead of the Poisson-Fürstenberg boundary. Consider the partition $\zeta$ of $\Omega$ into equivalence classes mod $\sim$, see ① in Figure 5. This partition induces a sub-$\sigma$-algebra $\mathscr{A}_\zeta$ of $\mathscr{A}$, consisting of all $A\in\mathscr{A}$ which are compatible with the partition $\zeta$, i. e. which are unions of equivalence classes mod $\sim$, see ② in Figure 5. The resulting probability space $(\zeta,\mathscr{A}_\zeta,\mathbb{P}|_{\mathscr{A}_\zeta})$ is not necessarily a Lebesgue-Rohlin space. We solve this issue by taking the measurable hull $\zeta_1$ of $\zeta$, see [Roh52, §3.3] and [CK12, §1.4]. This is a coarsening of $\zeta$, characterised mod 0, with the property that the induced sub-$\sigma$-algebras $\mathscr{A}_{\zeta_1}$ and $\mathscr{A}_\zeta$ agree, again mod 0, and the probability space $(\zeta_1,\mathscr{A}_{\zeta_1},\mathbb{P}|_{\mathscr{A}_{\zeta_1}})$ is a Lebesgue-Rohlin space. We denote the latter by $(B,\mathscr{B},\nu)$ and call it the Poisson-Fürstenberg boundary. The map from the trajectory space $\Omega$ to the Poisson-Fürstenberg boundary $B$ that assigns to every trajectory $x\in\Omega$ the respective element of the partition $\zeta_1$ is called the boundary map $\mathrm{bnd}:\Omega\to B$. Notice that is not the only possible definition of the Poisson-Fürstenberg boundary, further equivalent ones are given in [KV83].

### 1.3.3 Representation of bounded harmonic functions

**Motivation 1.22 ("Classical Poisson integral representation formula")** *The Dirichlet problem considers a region $A\subseteq\mathbb{R}^n$ and a continuous function $f:\partial A\to\mathbb{R}$. It asks whether $f$ can be extended to a function $\varphi: A\cup\partial A\to\mathbb{R}$ which is continuous in the closure $A\cup\partial A$ and harmonic in the interior $A$, i. e. twice continuously differentiable and satisfying the Laplace equation $\Delta\varphi=0$. In the special case that $A$ is the open unit disc in $\mathbb{R}^2$, the Dirichlet problem can be solved explicitly by the Poisson formula. An example is illustrated on the right-hand side of Figure 6. In order to write the unique solution $\varphi$ down,* *let us identify $\mathbb{R}^2$ with the complex plane $\mathbb{C}$. Then, $\varphi$ assigns to every $a \in A$ the value*

$$\varphi(a) = \int_0^1 f(\exp(2\pi x i)) \underbrace{\frac{1-|a|^2}{|\exp(2\pi x i) - a|^2}\,\mathrm{d}x}_{=:\,\mathrm{d}\nu_a(x)} .$$

One important feature of the Poisson-Fürstenberg boundary is that it can be used to describe all bounded harmonic functions on the state space $X$. Let us therefore translate the notion of harmonic functions to our situation, where no derivatives are available. So, assume we are given a countable state space $X$ and transition probabilities $p : X \times X \to [0,1]$ as introduced in Section 1.2.1.

**Definition 1.23 ("harmonic", "superharmonic")** *A function $\varphi : X \to \mathbb{R}$ is called harmonic if for every element $x \in X$ the equation $\varphi(x) = \sum_{y \in X} p(x,y)\varphi(y)$ holds. In other words, being at $x \in X$, the value of $\varphi$ today is precisely as large as the expected value of $\varphi$ tomorrow. A function called is superharmonic if the same holds with "$\geq$" instead of "$=$".*

The initial probability measure of a random walk is denoted by $\vartheta : X \to [0,1]$. First, we pick some reference measure $\vartheta$ with $\mathrm{supp}(\vartheta) = X$. Then, we consider the random walk $Z = (Z_0, Z_1, \dots)$ that starts according to $\vartheta$ and has probability measure $\mathbb{P}_\vartheta$ and Poisson-Fürstenberg boundary $(B, \mathscr{B}, \nu_\vartheta)$.

All other initial probability measures, in particular the Dirac measures $\delta_x$ at points $x \in X$, are absolutely continuous with respect to $\vartheta$. Therefore, the measures $\mathbb{P}_x := \mathbb{P}_{\delta_x}$ are absolutely continuous with respect to $\mathbb{P}_\vartheta$, which implies that we may equip $(B, \mathscr{B})$ with measures $\nu_x := \nu_{\delta_x}$ in order to obtain the respective Poisson-Fürstenberg boundaries. From this point of view, it would have made sense to define the Poisson-Fürstenberg boundary as a measurable space $(B, \mathscr{B})$ equipped with a family of measures.

Now, a first step decomposition shows that the equation $\nu_x = \sum_{y \in X} p(x,y) \cdot \nu_y$ holds for every two points $x, y \in X$. Hence, given an essentially bounded function $f$ mapping from the Poisson-Fürstenberg boundary $(B, \mathscr{B}, \nu_\vartheta)$ to the real numbers $\mathbb{R}$, we can construct a bounded harmonic function $\varphi : X \to \mathbb{R}$ given by the Poisson integral representation formula $\varphi(x) := \int f \,\mathrm{d}\nu_x$.

There is also a way back from $\varphi$ to $f$ using martingale convergence so that, in the end, one obtains a one-to-one correspondence, even an isometry of Banach spaces, between the space $L^\infty(B, \mathscr{B}, \nu_\vartheta)$ of equivalence classes of essentially bounded functions and the space $H^\infty(X, \mu)$ of bounded harmonic functions, see e. g. [KV83]. Let us illustrate this by an example.

**Example 1.24 ("Non-triviality of the boundary implies transience")** *It is known that the random walk $Z = (Z_0, Z_1, \dots)$ is transient if and only if there is a non-constant non-negative superharmonic function, see e.g. [Woe00, Theorem 1.16]. Now, assume that the Poisson-Fürstenberg boundary $B$ is non-trivial with respect to some $\nu_x$. Then, it is non-trivial with respect to $\nu_\vartheta$. So, there are non-constant functions $f \in L^\infty(B, \mathscr{B}, \nu_\vartheta)$ and $\varphi \in H^\infty(X, \mu)$. But, then, every $\varphi + c$ with $c := -\inf\{\varphi(x) \mid x \in X\}$ is a non-constant non-negative superharmonic function, and the random walk $Z$ is transient.*

### 1.3.4 The triviality and identification problem

Given a random walk, be it on a set or on a group, a challenging problem is to decide whether the Poisson-Fürstenberg boundary is trivial or not. In the latter case, one may go on and wonder how to identify it geometrically. We shall only outline a few results about the Poisson-Fürstenberg boundary of random walks on countable groups. A recent survey has been given by Erschler in [Ers10].

As before, let $Z = (Z_0, Z_1, \dots)$ be a random walk on a countable group $G$ driven by the probability measure $\mu$. We assume that the support $\mathrm{supp}(\mu)$ generates $G$ as a semigroup, see Section 1.2.2.

If $G$ is abelian, then the Poisson-Fürstenberg boundary is always trivial, see [Bla55] and [CD60]. The same holds true for all groups of polynomial growth, and for groups of subexponential growth with a probability measure $\mu$ of finite first moment. It has been shown in [Ers04], that the assumption of finite first moment cannot be dropped.

If $G$ is amenable, then one can show that there is at least one symmetric probability measure $\mu$ such that the Poisson-Fürstenberg boundary is trivial, see the conjecture in [Fü73, §9]. The latter has been proved in [Ros81] and [KV83]. For example, in case of the Baumslag-Solitar group $G = \mathrm{BS}(1,2)$, the Poisson-Fürstenberg boundary may or may not be trivial depending on the vertical drift $\delta$. More precisely, for random walks on $G = \mathrm{BS}(1,2)$ with finite first moment the Poisson-Fürstenberg boundary is isomorphic to $\mathbb{R}$ for $\delta < 0$ and trivial for $\delta = 0$ and isomorphic to $\mathbb{Q}_2$ for $\delta > 0$, see [Kaĭ91, Theorem 5.1]. We may think of $\mathbb{Q}_2$ as the space of upper ends of the corresponding Bass-Serre tree $\mathbb{T}$. Further results about random walks on rational affinities are given in [Bro06]. Finally, if $G$ is non-amenable, then the Poisson-Fürstenberg boundary is always non-trivial[4], see [Fü73, §9]. This holds in particular for random walks on non-amenable Baumslag-Solitar groups $G = \mathrm{BS}(p,q)$, also for $\delta = 0$.

### 1.3.5 Statement of the strip criterion

Kaĭmanovich's strip criterion is a tool to identify the Poisson-Fürstenberg boundary geometrically. The strategy is to guess a suitable candidate. In our situation, this candidate will be given in terms of the boundaries $\partial\mathbb{H}$ and $\partial\mathbb{T}$. The strip criterion enables us to prove that this candidate is actually the Poisson-Fürstenberg boundary. We will first state it and then, in Sections 1.3.6 and 1.3.7, discuss the notions that have not been mentioned so far. A proof can be found in [Kaĭ00, §6.4].

**Theorem 1.25 ("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 $(B_-, \mathscr{B}_-, \nu_-)$ and $(B_+, \mathscr{B}_+, \nu_+)$ be $\check{\mu}$- and $\mu$-boundaries, respectively. If there exist a gauge $\mathscr{G} = (\mathscr{G}_1, \mathscr{G}_2, \dots)$ on $G$ with associated gauge function $|\cdot| = |\cdot|_{\mathscr{G}}$ and a measurable $G$-equivariant map $S$ assigning to pairs of points $(b_-, b_+) \in B_- \times B_+$ non-empty strips $S(b_-, b_+) \subseteq G$ such that for every $g \in G$ and $\nu_- \otimes \nu_+$-almost every $(b_-, b_+) \in B_- \times B_+$*

$$1/n \cdot \ln\left(\mathrm{card}\left(S(b_-, b_+)g \cap \mathscr{G}_{|Z_n|}\right)\right) \xrightarrow{n\to\infty} 0 \;\; \textit{in probability},$$

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

**Remark 1.26** *The proof shows that, under certain conditions, it is not even necessary to verify the convergence for every $g \in G$ and it suffices to consider the special case $g = 1$. The only thing we have to ensure is that a random strip contains the identity element $1 \in G$ with positive probability, i. e. that*

$$\nu_- \otimes \nu_+ \{(b_-, b_+) \in B_- \times B_+ \mid 1 \in S(b_-, b_+)\} > 0\,.$$

### 1.3.6 Remark about Entropy

Roughly speaking, the entropy of the probability measure $\mu$ is the expected amount of information contained in the outcome of a random variable that is distributed according to $\mu$. More precisely, it is the real number given by $H(\mu) := \sum_{g\in G} -\log_2(\mu(g)) \cdot \mu(g)$. Here, as usual, one defines $-\log_2(0)\cdot 0 := 0$. For us, the assumption of finite entropy will be no issue because Baumslag-Solitar groups are finitely generated and the increments under investigation have finite first moment. This implies that their probability measures $\mu$ have finite entropy, as shown by the following well-known lemma.

[4] By Example 1.24, this tells us in particular that all random walks on non-amenable groups are transient.

**Lemma 1.27** *Let $G$ be a finitely generated group*[5] *and let $\mu : G \to [0,1]$ be a probability measure. If a random variable $X : \Omega \to G$ distributed according to $\mu$ has finite first moment, then $\mu$ has finite entropy.*

*Proof.* Let $S \subseteq G$ be a non-empty finite generating set. Moreover, let $b := 2 \cdot |S| + 1$, whence $b \geq 3$. We have to show that the entropy $H(X) = \sum_{g \in G} -\log_2(\mu(g)) \cdot \mu(g)$ is finite. First, we change the base of the logarithm

$$H(X) = \sum_{g \in G} -\log_2(\mu(g)) \cdot \mu(g) = \log_2(b) \cdot \sum_{g \in G} -\log_b(\mu(g)) \cdot \mu(g),$$

and split the summands appropriately

$$\ldots = \log_2(b) \cdot \left( -\log_b(\mu(1)) \cdot \mu(1) + \sum_{\substack{g \in G \setminus \{1\} \text{ with} \\ \mu(g) \leq b^{-d_S(1,g)}}} -\log_b(\mu(g)) \cdot \mu(g) + \sum_{\substack{g \in G \setminus \{1\} \text{ with} \\ \mu(g) > b^{-d_S(1,g)}}} -\log_b(\mu(g)) \cdot \mu(g) \right).$$

Then, we recall that the function $x \mapsto -\log_b(x) \cdot x$ is increasing on the interval $[0, 1/e]$, and conclude that

$$\begin{aligned}
\sum_{\substack{g \in G \setminus \{1\} \text{ with} \\ \mu(g) \leq b^{-d_S(1,g)}}} -\log_b(\mu(g)) \cdot \mu(g) &\leq \sum_{\substack{g \in G \setminus \{1\} \text{ with} \\ \mu(g) \leq b^{-d_S(1,g)}}} -\log_b\left(b^{-d_S(1,g)}\right) \cdot b^{-d_S(1,g)} \\
&= \sum_{\substack{g \in G \setminus \{1\} \text{ with} \\ \mu(g) \leq b^{-d_S(1,g)}}} d_S(1,g) \cdot b^{-d_S(1,g)} \\
&\leq \sum_{g \in G} d_S(1,g) \cdot b^{-d_S(1,g)} \\
&\leq \sum_{n=0}^{\infty} (2 \cdot |S|)^n \cdot n \cdot b^{-n} < \infty
\end{aligned}$$

On the other hand, since $X$ has finite first moment,

$$\sum_{\substack{g \in G \setminus \{1\} \text{ with} \\ \mu(g) > b^{-d_S(1,g)}}} -\log_b(\mu(g)) \cdot \mu(g) \leq \sum_{\substack{g \in G \setminus \{1\} \text{ with} \\ \mu(g) > b^{-d_S(1,g)}}} d_S(1,g) \cdot \mu(g) \leq \sum_{g \in G} d_S(1,g) \cdot \mu(g) < \infty.$$

So, both sums are finite, whence $H(X)$ must be finite, too. □

### 1.3.7 Further remarks about $\mu$-boundaries, gauges, ...

- The notion of a $\mu$-boundary goes back to Fürstenberg. Two equivalent definitions can be found in [Kaĭ00, §1.5]. For us, it suffices to record that every Lebesgue-Rohlin space $(B_+, \mathscr{B}_+, \nu_+)$ equipped with a left $G$-action and a boundary map $\mathrm{bnd}_+ : \Omega \to B_+$ that is ① measurable, ② $\sim$-invariant, and ③ $G$-equivariant is a $\mu$-boundary.

  Here, ① measurable means being a homomorphism (= measurable and measure preserving map) between Lebesgue-Rohlin spaces. It turns out that there is a natural correspondence between these homomorphisms and the measurable partitions of their domain, see [Roh52, §3.2] and [Hae73, p. 255, Remark] for details.

  The other two properties are standard. ② $\sim$-invariant means constant on the equivalence classes mod $\sim$ and ③ $G$-equivariant means that for every $g \in G$ and every trajectory $\omega = (x_0, x_1, \ldots) \in \Omega$ the equation $\mathrm{bnd}'(g.\omega) := \mathrm{bnd}'(g.(x_0, x_1, \ldots)) := \mathrm{bnd}'(gx_0, gx_1, \ldots) = g.\,\mathrm{bnd}'(\omega)$ holds.

[5] This assumption is necessary. For example, imagine the group $(\mathbb{Q}, +)$ and a probability measure $\mu$ with infinite entropy supported on the generating set $S := \{1, -1, 1/2, -1/2, 1/3, -1/3, \ldots\}$. Then, $X$ has finite first moment with respect to the word metric $d_S : \mathbb{Q} \times \mathbb{Q} \to \mathbb{R}$, but it has infinite entropy.

- A $\check{\mu}$-boundary is the space that arises when the probability measure $\mu$ is replaced by the reflected probability measure $\check{\mu}$ given by $\check{\mu}(g) := \mu(g^{-1})$.

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

- The power set $\{0,1\}^G$ is naturally equipped with the product $\sigma$-algebra, which enables us to talk about measurability of the map $S : B_- \times B_+ \to \{0,1\}^G$.

- The conclusion of the strip criterion, namely that the $\mu$-boundary $(B_+, \mathscr{B}_+, \nu_+)$ is maximal, is equivalent to saying that it is isomorphic to the Poisson-Fürstenberg boundary.

# 2 Identification of the Poisson-Fürstenberg boundary

In this section, we still assume $1 \le p < q$ and consider a random walk $Z = (Z_0, Z_1, \dots)$ on $G = \mathrm{BS}(p,q)$. Also, recall the abbreviations $A_{Z_n} := \Im(\pi_{\mathbb{H}}(Z_n))$ and $B_{Z_n} := \Re(\pi_{\mathbb{H}}(Z_n))$ introduced in Section 1.2.4.

## 2.1 Convergence to the boundary of the hyperbolic plane

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

*Proof.* We can use the strong law of large numbers, see e. g. [Kle14, Theorem 5.17], to obtain

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

Hence, the projections $\lambda(Z_n)$ tend a. s. to infinity and, by Remark 1.14, the imaginary parts $A_{Z_n}$ do. This, of course, implies that the absolute values $|\pi_{\mathbb{H}}(Z_n)|$ tend a. s. to infinity, and Lemma 1.10 completes the proof. ☐

**Lemma 2.2** *Assume that $\ln(A_{X_1})$ and $\ln(1+|B_{X_1}|)$ have finite first moment. If the vertical drift is negative, i. e. $\delta < 0$, then the projections $\pi_{\mathbb{H}}(Z_n)$ converge a. s. to a random element $\xi \in \partial\mathbb{H} \setminus \{\infty\}$.*

*Proof.* Notice that the argument given in the proof of Lemma 2.1 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 $B_{Z_n}$. The equation $\pi_{\mathbb{H}}(Z_n) = A_{Z_n} \cdot i + B_{Z_n}$ yields $\pi_{\mathrm{Aff}^+(\mathbb{R})}(Z_n)(z) = A_{Z_n} \cdot z + B_{Z_n}$, and in light of the multiplication in $\mathrm{Aff}^+(\mathbb{R})$ we obtain

$$\begin{aligned}
\pi_{\mathbb{H}}(Z_n) &= \pi_{\mathrm{Aff}^+(\mathbb{R})}(Z_n)(i) = \pi_{\mathrm{Aff}^+(\mathbb{R})}(X_1 \cdot \ldots \cdot X_n)(i) \\
&= \big(\pi_{\mathrm{Aff}^+(\mathbb{R})}(X_1) \circ \ldots \circ \pi_{\mathrm{Aff}^+(\mathbb{R})}(X_n)\big)(i) \\
&= A_{X_1} \cdot \ldots \cdot A_{X_n} \cdot i + \sum_{k=1}^{n} A_{X_1} \cdot \ldots \cdot A_{X_{k-1}} \cdot B_{X_k}\,.
\end{aligned}$$

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

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

For the first factor we can use the strong law of large numbers, for the second one the Borel-Cantelli Lemma. Indeed, let us write $Q_k$ for the quotient $1/k \cdot \ln(1 + |B_{X_k}|)$. In order to show that $Q_k \to 0$ a. s., recall that $\ln(1 + |B_{X_1}|)$ has finite first moment. For every $\varepsilon > 0$ we may thus estimate

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

Therefore, the Borel-Cantelli Lemma yields $\mathbb{P}(\{\omega \in \Omega \mid \exists \text{ infinitely many } k \in \mathbb{N} \text{ such that } Q_k(\omega) > \varepsilon\}) = 0$. Replacing $\varepsilon$ by $1, 1/2, 1/3, \ldots$, we obtain a countable family of null sets whose union is, of course, again a null set that consists of all $\omega \in \Omega$ with $Q_k(\omega) \not\to 0$. Hence, $Q_k \to 0$ a. s., see also [Kle14, Exercise 5.1.3]. So, we have finally convinced ourselves that $\limsup_{k \to \infty} |C_k|^{1/k} < 1$ a. s., whence $\sum_{k=1}^{\infty} C_k$ converges a. s. to a random element $\xi \in \mathbb{R}$. □

The natural question that remains is the one asking for the driftless case. An answer has been given by Brofferio in [Bro03, Theorem 1]. It says that under the same mild assumptions, namely that $\ln(A_{X_1})$ and $\ln(1 + |B_{X_1}|)$ have finite first moment, the projections $\pi_{\mathbb{H}}(Z_n)$ converge a. s. to $\infty \in \partial\mathbb{H}$. But, for us, a result of slightly different flavour will be of relevance.

**Lemma 2.3** *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 there is no vertical drift, i. e. $\delta = 0$, then the projections $\pi_{\mathbb{H}}(Z_n)$ have sublinear speed, i. e.*

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

The proof is based on ideas that go back to Élie in [Éli82, Lemme 5.49] and have also been used by Cartwright, Kaĭmanovich, and Woess in [CKW94, Proposition 4b]. We first adapt these ideas to our situation by stating and proving Lemma 2.4, and then deduce Lemma 2.3.

By assumption, there is no vertical drift so that $\lambda(Z) = (\lambda(Z_0), \lambda(Z_1), \ldots)$ is recurrent. Indeed, see Pólya's Theorem in [Pól21] for recurrence of the simple random walk and the Chung-Fuchs Theorem in [CF51] for the general case. In particular, we know that there is a. s. a strictly increasing sequence $\tau(0), \tau(1), \ldots$ given by $\tau(0) := 0$ and $\tau(n) := \inf\{k \in \mathbb{N} \mid \tau(n-1) < k \text{ and } \lambda(Z_{\tau(n-1)}) < \lambda(Z_k)\}$. We call $\tau(n)$ the $n$-th ladder time, see Figure 7 for an illustration of the first ladder times $\tau(0)$ and $\tau(1)$. The following lemma concerns the random variable $\ln(1 + \sum_{k=1}^{\tau} |B_{X_k}|)$ with $\tau := \tau(1)$.

**Lemma 2.4** *Under the same assumptions as in Lemma 2.3, namely 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 and, of course, that there is no vertical drift, the random variable $\ln(1 + \sum_{k=1}^{\tau} |B_{X_k}|)$ has finite first moment.*

*Proof.* Adapting the proof of [Éli82, Lemme 5.49], we begin with some preliminaries. Pick an $\varepsilon > 0$ that satisfies the requirements of Lemma 2.4 and let $\beta := \frac{1}{2+\varepsilon}$. Since $\ln(A_{X_1})$ has finite second moment, we

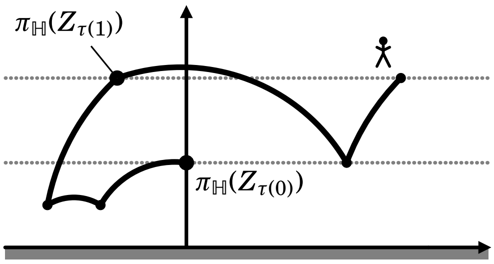


Figure 7: The first ladder times $\tau(0)$ and $\tau(1)$.

know that also $\lambda(X_1)$ has finite second moment and $\mathbb{P}(\{\omega\in\Omega \mid \tau(\omega)>k\}) \sim \text{const}\cdot k^{-1/2}$, see [Éli82, §5.44] referring to [Fel71, p. 415]. Using this asymptotics, we obtain that

$$\int \tau^\beta \,\mathrm{d}\mathbb{P} \le \int \left\lceil \tau^\beta \right\rceil \mathrm{d}\mathbb{P} = \sum_{k=1}^{\infty} \mathbb{P}\left(\left\{\omega\in\Omega \,\middle|\, \left\lceil \tau(\omega)^\beta \right\rceil \ge k\right\}\right) = \sum_{k=0}^{\infty} \underbrace{\mathbb{P}\left(\left\{\omega\in\Omega \,\middle|\, \tau(\omega) > k^{1/\beta}\right\}\right)}_{\sim\, \text{const}\cdot k^{-(1+\varepsilon/2)}}.$$

In particular, there is a $k_0\in\mathbb{N}$ such that for all $k\ge k_0$ the summands $\mathbb{P}(\{\omega\in\Omega\mid\tau(\omega)>k^{1/\beta}\})$ are strictly smaller than $k^{-(1+\varepsilon/4)}$. And, since $\sum_{k=k_0}^{\infty} k^{-(1+\varepsilon/4)}<\infty$, we know that $\int\tau^\beta\,\mathrm{d}\mathbb{P}<\infty$. Moreover, notice that, by construction of the ladder times $\tau(0),\tau(1),\dots$, the differences $\tau(1)-\tau(0),\tau(2)-\tau(1),\dots$ are i. i. d., whence the fact that $0<\beta<1$, $\Rightarrow (x+y)^\beta\le x^\beta+y^\beta$, and the strong law of large numbers yield

$$\frac{\tau(n)^\beta}{n} \le \frac{(\tau(1)-\tau(0))^\beta+\ldots+(\tau(n)-\tau(n-1))^\beta}{n} \xrightarrow[\text{a. s.}]{n\to\infty} \mathbb{E}\left(\tau^\beta\right), \implies \limsup_{n\to\infty}\frac{\tau(n)^\beta}{n}<\infty \text{ a. s.} \qquad (*)$$

Now, we are prepared for the main argument. Recall that we aim to show that $\ln(1+\sum_{k=1}^{\tau}|B_{X_k}|)$ has finite first moment. The sums $\sum_{k=\tau(0)+1}^{\tau(1)}|B_{X_k}|$, $\sum_{k=\tau(1)+1}^{\tau(2)}|B_{X_k}|$, ... are i. i. d. and non-negative with the additional property that they are not a. s. equal to 0. Hence, by [Éli82, Lemme 5.23],

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

In order to verify the right-hand side, we would like to estimate

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

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

$$L=\limsup_{n\to\infty}\frac{\ln\left(1+\sum_{k=1}^{\tau(n)}|B_{X_k}|\right)}{\tau(n)^\beta}\le\limsup_{n\to\infty}\frac{\ln\left(1+\tau(n)\cdot\max_{1\le k\le\tau(n)}\{|B_{X_k}|\}\right)}{\tau(n)^\beta}$$

$$\le\underbrace{\limsup_{n\to\infty}\frac{\ln(\tau(n))}{\tau(n)^\beta}}_{=0}+\limsup_{n\to\infty}\frac{\ln\left(1+\max_{1\le k\le\tau(n)}\{|B_{X_k}|\}\right)}{\tau(n)^\beta}$$

$$= \limsup_{n\to\infty} \left( \frac{\max_{1\le k\le \tau(n)} \left\{ \ln\left(1+|B_{X_k}|\right)^{1/\beta} \right\}}{\tau(n)} \right)^{\beta}$$

$$\le \limsup_{n\to\infty} \left( \underbrace{\frac{\sum_{k=1}^{\tau(n)} \ln\left(1+|B_{X_k}|\right)^{1/\beta}}{\tau(n)}}_{=:M_n} \right)^{\beta}.$$

Now, recall that $1/\beta = 2+\varepsilon$. So, by the strong law of large numbers, the sequence $M_n$ converges a. s. to the expectation $\mathbb{E}(\ln(1+|B_{X_1}|)^{1/\beta})$, which implies that $L \le \limsup_{n\to\infty} {M_n}^{\beta} < \infty$ a. s., and completes the proof. $\square$

**Remark 2.5** *Élie makes a conclusion concerning the horizontal displacement at the first ladder time, see [Éli82, Lemme 5.49]. The corresponding result in our situation would be that* $\ln(1+|B_{Z_\tau}|)$ *has finite first moment. But, whereas Élie assumes that the horizontal increments are i. i. d., this is not the case in our situation. Just recall from the proof of Lemma 2.2 that our horizontal increments amount to* $C_k = A_{X_1}\cdot\ldots\cdot A_{X_{k-1}}\cdot B_{X_k}$. *One brutal way to solve this issue is to ignore the factors* $A_{X_1}\cdot\ldots\cdot A_{X_{k-1}}$. *And, indeed, this turns out to work. We actually obtain that* $\ln(1+|B_{Z_\tau}|)$ *has finite first moment. The only thing we have to do is to observe that*

$$|B_{Z_\tau}| = \left|\sum_{k=1}^{\tau} C_k\right| = \left|\sum_{k=1}^{\tau} A_{X_1}\cdot\ldots\cdot A_{X_{k-1}}\cdot B_{X_k}\right| \le \sum_{k=1}^{\tau} \underbrace{A_{X_1}\cdot\ldots\cdot A_{X_{k-1}}}_{\le 1\ a.\,s.}\cdot|B_{X_k}| \le \sum_{k=1}^{\tau}|B_{X_k}|\ a.\,s. \qquad (**)$$

*Proof of Lemma 2.3.* Recall from [Éli82, §5.44] and [Fel71, p. 415] that $\mathbb{P}(\{\omega\in\Omega \mid \tau(\omega)>k\}) \sim \mathrm{const}\cdot k^{-1/2}$ with a strictly positive constant. In particular, there is a $k_0\in\mathbb{N}$ such that for all $k\ge k_0$ the summands $\mathbb{P}(\{\omega\in\Omega\mid\tau(\omega)>k\})$ are larger than $k^{-1}$ and we obtain that

$$\int \tau\,\mathrm{d}\mathbb{P} = \sum_{k=1}^{\infty}\mathbb{P}(\{\omega\in\Omega\mid\tau(\omega)\ge k\}) = \sum_{k=0}^{\infty}\mathbb{P}(\{\omega\in\Omega\mid\tau(\omega)>k\}) \ge \sum_{k=k_0}^{\infty} k^{-1} = \infty.$$

As we have noticed in the proof of Lemma 2.4, the differences $\tau(1)-\tau(0), \tau(2)-\tau(1), \ldots$ are i. i. d., and they are non-negative. So, we may deduce from the strong law of large numbers that

$$\frac{\tau(n)}{n} = \frac{(\tau(1)-\tau(0))+(\tau(2)-\tau(1))+\ldots+(\tau(n)-\tau(n-1))}{n} \xrightarrow[\text{a. s.}]{n\to\infty} \infty \quad\text{and}\quad \frac{n}{\tau(n)} \xrightarrow[\text{a. s.}]{n\to\infty} 0.$$

This can be used to estimate the distance between $\pi_{\mathbb{H}}(Z_0)$ and $\pi_{\mathbb{H}}(Z_n)$ from above. First, for every $n\in\mathbb{N}_0$ let $m = m(n)\in\mathbb{N}_0$ be the unique element with $\tau(m)\le n<\tau(m+1)$. This element exists a. s. because the ladder times $\tau(0), \tau(1), \ldots$ do. Now, we may estimate

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

The numbers refer to Figure 8. We will consider the three summands separately and show that each of them converges a. s. to 0. For ① and ③ this is straightforward. Indeed,

$$① = \frac{|\ln(A_{Z_{\tau(m)}})|}{n} \le \frac{|\ln(A_{Z_{\tau(m)}})|}{\tau(m)} = \ln(q/p)\cdot\left|\frac{\lambda(X_1)+\ldots+\lambda(X_{\tau(m)})}{\tau(m)}\right| \xrightarrow[\text{a. s.}]{n\to\infty} \ln(q/p)\cdot|\mathbb{E}(\lambda(X_1))| = \ln(q/p)\cdot|\delta| = 0$$

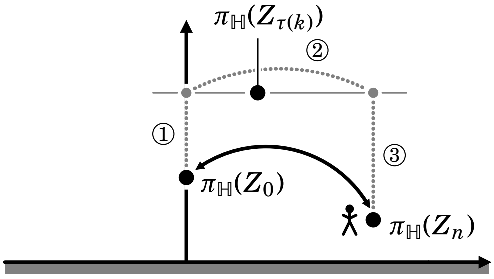


Figure 8: Estimate of the distance between $\pi_{\mathbb{H}}(Z_0)$ and $\pi_{\mathbb{H}}(Z_n)$.

and similarly

$$③ \le \frac{d_{\mathbb{H}}(A_{Z_{\tau(m)}}\cdot i, i)}{n} + \frac{d_{\mathbb{H}}(i, A_{Z_n}\cdot i)}{n} = ① + \frac{|\ln(A_{Z_n})|}{n} = ① + \ln(q/p)\cdot\left|\frac{\lambda(X_1)+\ldots+\lambda(X_n)}{n}\right| \xrightarrow[\text{a. s.}]{n\to\infty} 0.$$

For ② recall again from the proof of Lemma 2.2 that $B_{Z_n} = \sum_{k=1}^n A_{X_1}\cdot\ldots\cdot A_{X_{k-1}}\cdot B_{X_k}$ and observe that not just for the case $m=0$ and $\ell=\tau$, which has been studied in $(**)$ in Remark 2.5, but for every $m,\ell\in\mathbb{N}_0$ with $\tau(m)\le\ell\le\tau(m+1)$ the following estimate holds

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

Hence, using that $A_{Z_{\tau(0)}} < A_{Z_{\tau(1)}} < \ldots < A_{Z_{\tau(m)}}$ a. s. and that $n<\tau(m+1)$ a. s., we obtain

$$\begin{aligned}② &= 1/n\cdot\operatorname{arcosh}\left(1+1/2\cdot\left(\frac{|B_{Z_n}|}{A_{Z_{\tau(m)}}}\right)^2\right) = 1/n\cdot\ln\left(1+1/2\cdot\left(\frac{|B_{Z_n}|}{A_{Z_{\tau(m)}}}\right)^2+\sqrt{\left(1+1/2\cdot\left(\frac{|B_{Z_n}|}{A_{Z_{\tau(m)}}}\right)^2\right)^2-1}\right)\\
&\le 1/n\cdot\left(\ln(2)+\ln\left(1+\left(\frac{|B_{Z_n}|}{A_{Z_{\tau(m)}}}\right)^2\right)\right) \le 1/n\cdot\left(\ln(2)+2\cdot\ln\left(1+\frac{|B_{Z_n}|}{A_{Z_{\tau(m)}}}\right)\right)\\
&\le 1/n\cdot\left(\ln(2)+2\cdot\ln\left(1+\frac{|B_{Z_{\tau(1)}}-B_{Z_{\tau(0)}}|}{A_{Z_{\tau(0)}}}+\frac{|B_{Z_{\tau(2)}}-B_{Z_{\tau(1)}}|}{A_{Z_{\tau(1)}}}+\ldots+\frac{|B_{Z_{\tau(m)}}-B_{Z_{\tau(m-1)}}|}{A_{Z_{\tau(m-1)}}}+\frac{|B_{Z_n}-B_{Z_{\tau(m)}}|}{A_{Z_{\tau(m)}}}\right)\right)\\
&\le 1/n\cdot\left(\ln(2)+2\cdot\ln\left(1+\sum_{k=1}^{n}|B_{X_k}|\right)\right) \le 1/n\cdot\left(\ln(2)+2\cdot\ln\left(1+\sum_{k=1}^{\tau(m+1)}|B_{X_k}|\right)\right)\\
&\le \underbrace{\frac{\ln(2)}{n}}_{\to 0}+2\cdot\underbrace{\frac{\ln\left(1+\sum_{k=\tau(0)+1}^{\tau(1)}|B_{X_k}|\right)+\ldots+\ln\left(1+\sum_{k=\tau(m)+1}^{\tau(m+1)}|B_{X_k}|\right)}{m+1}}_{\to\mathbb{E}(\ln(1+\sum_{k=1}^{\tau}|B_{X_k}|))\text{ a. s. by Lemma 2.4}}\cdot\underbrace{\frac{m+1}{\tau(m)}}_{\substack{\to 0\\ \text{a. s.}}}\cdot\underbrace{\frac{\tau(m)}{n}}_{\substack{\le 1\\ \text{a. s.}}} \xrightarrow[\text{a. s.}]{n\to\infty} 0.\end{aligned}$$

□

## 2.2 Convergence to the space of ends of the Bass-Serre tree

Unlike the ones considered in Section 2.1, the projections $\pi_{\mathbb{T}}(Z_n)$ do not need to satisfy the Markov property. Consider, for example, the random walk $Z = (Z_0, Z_1, \ldots)$ driven by the uniform measure on the standard generators and their formal inverses. Then, given $\pi_{\mathbb{T}}(Z_{k-2}) = B$ and $\pi_{\mathbb{T}}(Z_{k-1}) = a^{-1}B$, the projection $\pi_{\mathbb{T}}(Z_k)$ comes back to $B$ with probability $1/4$. On the other hand, coming back to $B$ would not be possible if the history was $\pi_{\mathbb{T}}(Z_{k-3}) = B$ and $\pi_{\mathbb{T}}(Z_{k-2}) = \pi_{\mathbb{T}}(Z_{k-1}) = a^{-1}B$. Despite of this subtlety, the following lemmas yield almost sure convergence of the projections $\pi_{\mathbb{T}}(Z_n)$ to a random end.

**Lemma 2.6** *Assume that $X_1$ has finite first moment. If the vertical drift is different from 0, i. e. $\delta \neq 0$, then the projections $\pi_{\mathbb{T}}(Z_n)$ converge a. s. to a random end $\xi \in \partial\mathbb{T}$.*

We give an argument using the notion of regular sequences, see [CKW94, §2.C]. One difference is that we do not fix any particular end $\omega \in \partial\mathbb{T}$. Therefore, we replace the Busemann function $h$, which depends on the choice of $\omega \in \partial\mathbb{T}$, by the graph distance to the basepoint $B$. The other difference is that we work with the limit inferior instead of the limit in order to be prepared to deal with the driftless case, too.

*Proof of Lemma 2.6.* Let $d_{\mathbb{T}}$ be the graph distance in the tree $\mathbb{T}$. Accordingly, the symbol $|x|$ denotes the graph distance $d_{\mathbb{T}}(B, x)$ from the basepoint $B$ to the vertex $x$. We call a sequence $(x_0, x_1, \ldots)$ of vertices regular if

$$\text{①} \quad \liminf_{n\to\infty} \frac{|x_n|}{n} > 0 \quad \text{and} \quad \text{②} \quad \frac{d_{\mathbb{T}}(x_n, x_{n+1})}{n} \xrightarrow{n\to\infty} 0 .$$

In order to prove Lemma 2.6, we pursue a two-step strategy. First, we show that every regular sequence converges to an end and, second, that the projections $\pi_{\mathbb{T}}(Z_n)$ constitute a. s. a regular sequence.

Concerning the first part, we pick an arbitrary regular sequence $(x_0, x_1, \ldots)$ and claim that there is an end $\xi \in \partial\mathbb{T}$ such that for every $\varepsilon > 0$ the open ball $B_\varepsilon(\xi) := \{x \in \widehat{\mathbb{T}} \mid d_{\widehat{\mathbb{T}}}(\xi, x) < \varepsilon\}$ contains infinitely many $x_k$. Assume there was no such end. Then, we know that for every $\xi \in \partial\mathbb{T}$ there is an $\varepsilon_1 = \varepsilon_1(\xi) > 0$ such that $B_{\varepsilon_1}(\xi)$ contains only finitely many $x_k$, whence there is also an $\varepsilon_2 = \varepsilon_2(\xi) > 0$ such that $B_{\varepsilon_2}(\xi)$ does not contain any $x_k$ at all. The open balls $B_{\varepsilon_2}(\xi)$ with $\xi \in \partial\mathbb{T}$ and the singletons $\{x\}$ with $x \in \mathbb{T}$ form an open covering of $\widehat{\mathbb{T}}$. By compactness, it contains a finite subcovering. But, since the constants $\varepsilon_2 = \varepsilon_2(\xi) > 0$ have been chosen in such a way that the sequence $(x_0, x_1, \ldots)$ does not enter any of the open balls $B_{\varepsilon_2}(\xi)$, it must remain in a finite subset of the tree, which contradicts ①.

Next, we pick such an end $\xi \in \partial\mathbb{T}$ and claim that the sequence $(x_0, x_1, \ldots)$ converges to $\xi$. Let $\varepsilon > 0$. We define $\alpha := 1/3 \cdot \liminf_{n\to\infty} |x_n|/n$. In this case, all but finitely many $|x_n|$ are strictly greater than $2\alpha n$. The elements in $B_\varepsilon(\xi)$ are characterised by the property that the paths starting in $B$ and representing them must have a certain finite initial piece. Let $m$ be the length of this piece, i. e. $m := \max\{0, \lfloor 1 - \log_2(\varepsilon) \rfloor\}$. By ① and ②, there is an $n_0 \in \mathbb{N}$ such that for all $n \geq n_0$ the inequalities $|x_n| > \alpha n + m$ and $d_{\mathbb{T}}(x_n, x_{n+1}) < \alpha n$ hold. Since $B_\varepsilon(\xi)$ contains infinitely many $x_k$, we can even find an $n_1 \geq n_0$ such that $x_{n_1} \in B_\varepsilon(\xi)$. It turns out that not just for $n_1$ but for all $n \geq n_1$ we have $x_n \in B_\varepsilon(\xi)$. Indeed, if there was an $n \geq n_1$ such that $x_n \in B_\varepsilon(\xi)$ and $x_{n+1} \notin B_\varepsilon(\xi)$, we know that

$$d_{\mathbb{T}}(x_n, x_{n+1}) \geq d_{\mathbb{T}}(x_n, x_n \wedge x_{n+1}) = |x_n| - |x_n \wedge x_{n+1}| \geq |x_n| - m > \alpha n .$$

The latter, of course, contradicts $d_{\mathbb{T}}(x_n, x_{n+1}) < \alpha n$, see Figure 9. So, the two claims show that every regular sequence converges to an end. Concerning the second part, we aim to prove that

$$\text{①} \quad \liminf_{n\to\infty} \frac{|\pi_{\mathbb{T}}(Z_n)|}{n} > 0 \ \text{ a. s.} \quad \text{and} \quad \text{②} \quad \frac{d_{\mathbb{T}}(\pi_{\mathbb{T}}(Z_n), \pi_{\mathbb{T}}(Z_{n+1}))}{n} \xrightarrow[\text{a. s.}]{n\to\infty} 0 .$$

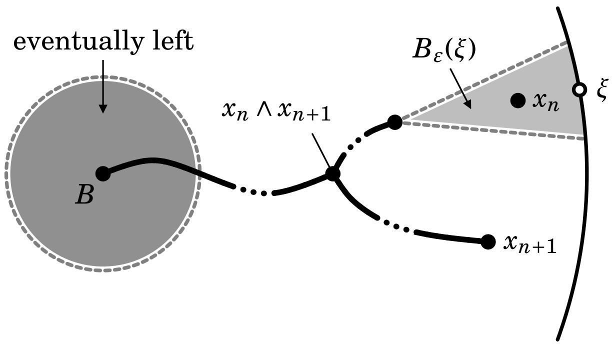


Figure 9: Jumping away from $B_\varepsilon(\xi)$.

Recall from Lemma 1.13 and Remark 1.14 that not just $X_1$ but also $\lambda(X_1)$ has finite first moment. So, the strong law of large numbers yields

$$\frac{|\pi_{\mathbb{T}}(Z_n)|}{n} \geq \frac{|\lambda(Z_n)|}{n} = \frac{|\lambda(X_1)+\ldots+\lambda(X_n)|}{n} \xrightarrow[\text{a.s.}]{n\to\infty} |\mathbb{E}(\lambda(X_1))| = |\delta| > 0\,, \implies \liminf_{n\to\infty} \frac{|\pi_{\mathbb{T}}(Z_n)|}{n} > 0 \text{ a.s.}$$

Next, let $S := \{a,b\} \subseteq G$ be the standard generating set. The numerators $d(\pi_{\mathbb{T}}(Z_n), \pi_{\mathbb{T}}(Z_{n+1}))$ of the fraction considered in ② are i. i. d., and the first one satisfies

$$\int d_{\mathbb{T}}(\pi_{\mathbb{T}}(Z_0), \pi_{\mathbb{T}}(Z_1))\,\mathrm{d}\mathbb{P} \leq \int d_{\mathrm{S}}(Z_0, Z_1)\,\mathrm{d}\mathbb{P} = \int d_{\mathrm{S}}(1, X_1)\,\mathrm{d}\mathbb{P} < \infty\,.$$

So, again, by the strong law of large numbers

$$\frac{d_{\mathbb{T}}(\pi_{\mathbb{T}}(Z_0), \pi_{\mathbb{T}}(Z_1)) + \ldots + d_{\mathbb{T}}(\pi_{\mathbb{T}}(Z_n), \pi_{\mathbb{T}}(Z_{n+1}))}{n+1} \xrightarrow[\text{a.s.}]{n\to\infty} \mathbb{E}(d_{\mathbb{T}}(\pi_{\mathbb{T}}(Z_0), \pi_{\mathbb{T}}(Z_1)))\,.$$

Now, a simple calculation yields

$$\frac{d_{\mathbb{T}}(\pi_{\mathbb{T}}(Z_n), \pi_{\mathbb{T}}(Z_{n+1}))}{n} \xrightarrow[\text{a.s.}]{n\to\infty} 0\,.$$

□

For the driftless case, the situation is not as easy and in order to show almost sure convergence of the projections $\pi_{\mathbb{T}}(Z_n)$ to a random end, we restrict ourselves to the non-amenable subcase $1 < p < q$.

**Lemma 2.7** *Let $1 < p < q$. Assume that $X_1$ has finite first moment and there is an $\varepsilon > 0$ such that $\ln(1+|B_{X_1}|)$ has finite $(2+\varepsilon)$-th moment. If there is no vertical drift, i. e. $\delta = 0$, then the projections $\pi_{\mathbb{T}}(Z_n)$ converge a. s. to a random end $\xi \in \partial\mathbb{T}$.*

*Proof.* Again, we claim that the projections $\pi_{\mathbb{T}}(Z_n)$ constitute a. s. a regular sequence. But, since $\delta = 0$, we need to modify the argument from the proof of Lemma 2.6 that showed ①. By assumption, $G$ is non-amenable and, in particular, the spectral radius $\varrho(\mu)$ of the random walk $Z = (Z_0, Z_1, \ldots)$ is strictly smaller than 1, see e. g. [Woe00, Corollary 12.5]. This, together with the fact that the random walk is uniformly irreducible, yields that

$$\liminf_{n\to\infty} \frac{d_{\mathrm{S}}(Z_0, Z_n)}{n} > 0\,.$$

For a proof, see e. g. [Woe00, Proposition 8.2]. In order to estimate the numerators $d_{\mathrm{S}}(Z_0, Z_n)$ from above, we apply an auxiliary result: *There are $\alpha, \beta > 0$ such that for every element $g \in G$ the inequality*

$d_{\mathrm{S}}(1,g) \le \alpha \cdot |\pi_{\mathbb{T}}(g)| + \beta \cdot d_{\mathbb{H}}(\pi_{\mathbb{H}}(1),\pi_{\mathbb{H}}(g))$ *holds.* Let us postpone the proof and record that, using this auxiliary result and Lemma 2.3, we obtain

$$\liminf_{n\to\infty} \frac{|\pi_{\mathbb{T}}(Z_n)|}{n} \ge \frac{1}{\alpha} \cdot \liminf_{n\to\infty} \left( \frac{d_{\mathrm{S}}(Z_0,Z_n)}{n} - \beta \cdot \underbrace{\frac{d_{\mathbb{H}}(\pi_{\mathbb{H}}(Z_0),\pi_{\mathbb{H}}(Z_n))}{n}}_{\to 0 \text{ a. s. by Lemma 2.3}} \right) > 0 \text{ a. s.}$$

This is ①. Concerning ②, notice that the respective argument from the proof of Lemma 2.6 did not use the assumption that $\delta \neq 0$, and therefore works also for $\delta = 0$. So, we know that the projections $\pi_{\mathbb{T}}(Z_n)$ constitute a. s. a regular sequence, which converges to a random end by the proof of Lemma 2.6.

It remains to show the auxiliary result. In order to do so, we construct a path from 1 to $g$ in the Cayley graph $\Gamma$ with at most $\alpha \cdot |\pi_{\mathbb{T}}(g)| + \beta \cdot d_{\mathbb{H}}(\pi_{\mathbb{H}}(1),\pi_{\mathbb{H}}(g))$ many edges, where the values of $\alpha$ and $\beta$ are to be determined uniformly, i. e. not depending on $g$. First, we aim to adjust the tree component. Either combinatorially using the defining relation $ab^pa^{-1} = b^q$ or geometrically using the properties of the Cayley graph $\Gamma$, we can find a path from 1 to the coset $gB$ with at most $(\lfloor q/2 \rfloor + 1) \cdot |\pi_{\mathbb{T}}(g)|$ many edges. Let $h \in gB$ be the endpoint of this path. Next, recall the notion of a discrete hyperbolic plane from Section 1.1.4. We pick an arbitrary ascending doubly infinite path $v : \mathbb{Z} \to G/B$ in the tree $\mathbb{T}$ that traverses the vertex $gB$ and follow a shortest path from $h$ to $g$ in the discrete hyperbolic plane $\Gamma_v$. By the proof of Lemma 1.7, its length $d_{\Gamma_v}(h,g)$ can be estimated from above by $\kappa \cdot d_{\mathbb{H}}(\pi_{\mathbb{H}}(h),\pi_{\mathbb{H}}(g))$ with $\kappa := \max\{c/\varepsilon, 1/\ell_a, 1/\ell_b\} > 0$. We may continue this estimate and finally obtain

$$\begin{aligned} d_{\Gamma_v}(h,g) &\le \kappa \cdot d_{\mathbb{H}}(\pi_{\mathbb{H}}(h),\pi_{\mathbb{H}}(g)) \\ &\le \kappa \cdot d_{\mathbb{H}}(\pi_{\mathbb{H}}(h),\pi_{\mathbb{H}}(1)) + \kappa \cdot d_{\mathbb{H}}(\pi_{\mathbb{H}}(1),\pi_{\mathbb{H}}(g)) \\ &\le \kappa \cdot \max\{\ell_a,\ell_b\} \cdot (\lfloor q/2 \rfloor + 1) \cdot |\pi_{\mathbb{T}}(g)| + \kappa \cdot d_{\mathbb{H}}(\pi_{\mathbb{H}}(1),\pi_{\mathbb{H}}(g)). \end{aligned}$$

So, the concatenation of the two paths considered above has at most $\alpha \cdot |\pi_{\mathbb{T}}(g)| + \beta \cdot d_{\mathbb{H}}(\pi_{\mathbb{H}}(1),\pi_{\mathbb{H}}(g))$ many edges with $\alpha := (1 + \kappa \cdot \max\{\ell_a,\ell_b\}) \cdot (\lfloor q/2 \rfloor + 1) > 0$ and $\beta := \kappa > 0$. □

## 2.3 Construction of the Poisson-Fürstenberg boundary

Resuming Sections 2.1 and 2.2, we may formulate the following theorem.

**Theorem 2.8 ("Convergence theorem")** *Let $Z = (Z_0, Z_1, \dots)$ be a random walk on a non-amenable Baumslag-Solitar group $G = \mathrm{BS}(p,q)$ with $1 < p < q$. If there is an $\varepsilon > 0$ such that the increment $X_1$ has finite $(2+\varepsilon)$-th moment, then the projections $\pi_{\mathbb{H}}(Z_n)$ converge a. s. to a random element in $\partial\mathbb{H}$ and the projections $\pi_{\mathbb{T}}(Z_n)$ converge a. s. to a random element in $\partial\mathbb{T}$. Moreover, if there is no vertical drift, i. e. $\delta = 0$, then the projections $\pi_{\mathbb{H}}(Z_n)$ have sublinear speed.*

*Proof.* Recall from Lemma 1.13 that not only the increment $X_1$ but also $\ln(A_{X_1})$ and $\ln(1 + |B_{X_1}|)$ have finite $(2+\varepsilon)$-th moment, which implies that for every $k \le 2 + \varepsilon$ they also have finite $k$-th moment. Now, apply the lemmas from Sections 2.1 and 2.2, and Brofferio's result mentioned on page 31. □

The boundaries $\partial\mathbb{H}$ and $\partial\mathbb{T}$ are equipped with their Borel $\sigma$-algebras $\mathscr{B}_{\partial\mathbb{H}}$ and $\mathscr{B}_{\partial\mathbb{T}}$. Under the assumptions of the convergence theorem, there are boundary maps $\mathrm{bnd}_{\partial\mathbb{H}} : \Omega \to \partial\mathbb{H}$ and $\mathrm{bnd}_{\partial\mathbb{T}} : \Omega \to \partial\mathbb{T}$, defined almost everywhere, assigning to a trajectory $\omega = (x_0, x_1, \dots) \in \Omega$ the limits

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

Even though they are only defined almost everywhere, the boundary maps are measurable in the sense that the preimages of measurable sets are measurable. Their product $\mathrm{bnd}_{\partial\mathbb{H}\times\partial\mathbb{T}} : \Omega \to \partial\mathbb{H}\times\partial\mathbb{T}$ is measurable with respect to the product $\sigma$-algebra $\mathscr{B}_{\partial\mathbb{H}} \otimes \mathscr{B}_{\partial\mathbb{T}}$. Since both $\partial\mathbb{H}$ and $\partial\mathbb{T}$ are metrisable and separable topological spaces, it is not hard to see that the product $\sigma$-algebra $\mathscr{B}_{\partial\mathbb{H}} \otimes \mathscr{B}_{\partial\mathbb{T}}$ agrees with the Borel $\sigma$-algebra $\mathscr{B}_{\partial\mathbb{H}\times\partial\mathbb{T}}$, see e. g. [Bil99, Appendix M.10].

**Definition 2.9 ("hitting measures")** *The three pushforward probability measures $\nu_{\partial\mathbb{H}} := \mathrm{bnd}_{\partial\mathbb{H}}(\mathbb{P})$, $\nu_{\partial\mathbb{T}} := \mathrm{bnd}_{\partial\mathbb{T}}(\mathbb{P})$, $\nu_{\partial\mathbb{H}\times\partial\mathbb{T}} := \mathrm{bnd}_{\partial\mathbb{H}\times\partial\mathbb{T}}(\mathbb{P})$ on the measurable spaces $(\partial\mathbb{H},\mathscr{B}_{\partial\mathbb{H}})$, $(\partial\mathbb{T},\mathscr{B}_{\partial\mathbb{T}})$, $(\partial\mathbb{H}\times\partial\mathbb{T},\mathscr{B}_{\partial\mathbb{H}\times\partial\mathbb{T}})$ are called the hitting measures. Notice that we may again, tacitly, complete the probability spaces with respect to $\nu_{\partial\mathbb{H}}$, $\nu_{\partial\mathbb{T}}$, $\nu_{\partial\mathbb{H}\times\partial\mathbb{T}}$.*

Each of the boundaries $\partial\mathbb{H}$ and $\partial\mathbb{T}$ is equipped with a left $G$-action. The one on $\partial\mathbb{H}$ is induced by the action $g.z := \pi_{\mathrm{Aff}^+(\mathbb{R})}(g)(z)$ on $\mathbb{H}$ and the one on $\partial\mathbb{T}$ is induced by the action $g.hB := ghB$ on $\mathbb{T}$. Let us describe them in more detail. The former is an action by isometries, and in light of their classification mentioned in Section 1.1.3, we can also evaluate them at $\partial\mathbb{H}$. For the latter, recall that ends are infinite reduced paths that start in $B$. The coordinatewise action on the ends maps every such path $\xi \in \partial\mathbb{T}$ to some other path that need not start in $B$ any more. The end $g.\xi \in \partial\mathbb{T}$ is obtained by connecting $B$ with the initial vertex of this path and reduce the concatenation. This way, it is not hard to see that we can map every $\xi \in \partial\mathbb{T}$ to an end with an arbitrarily chosen finite initial piece[6]. In particular, every orbit $\{g.\xi \mid g \in G\}$ is infinite and dense in $\partial\mathbb{T}$.

By changing the initial probability measure of $Z = (Z_0, Z_1, \dots)$ as in Section 1.3.3, we may obtain stationarity of the measure $\nu_{\partial\mathbb{T}}$. More precisely, for every measurable set $A \subseteq \partial\mathbb{T}$

$$\nu_{\partial\mathbb{T}}(A) = \nu_{\partial\mathbb{T},1}(A) = \sum_{g\in G} \mu(g)\cdot \nu_{\partial\mathbb{T},g}(A) = \sum_{g\in G} \mu(g)\cdot \nu_{\partial\mathbb{T},1}(g^{-1}.A) = \sum_{g\in G} \mu(g)\cdot \nu_{\partial\mathbb{T}}(g^{-1}.A). \qquad (*)$$

The same result holds true for $\partial\mathbb{H}$ and the product $\partial\mathbb{H}\times\partial\mathbb{T}$, which is equipped with the componentwise left $G$-action. These observations will be helpful in a moment, when we show that the hitting measures are either Dirac measures or non-atomic. Our proof is based on [Woe89, Lemma 3.4], which is much more general. The original idea for our special case might be older.

**Lemma 2.10** *The hitting measure $\nu_{\partial\mathbb{T}}$ is non-atomic. Moreover, if $\delta \geq 0$, then the hitting measure $\nu_{\partial\mathbb{H}}$ is the Dirac measure at $\infty \in \partial\mathbb{H}$ and, if $\delta < 0$, then it is non-atomic as well.*

*Proof.* Let us first consider the hitting measure $\nu_{\partial\mathbb{T}}$. Suppose, it contained elements of positive measure. Then, we may choose such an element $\xi \in \partial\mathbb{T}$ with maximal measure $a$. In particular, for every element $\eta \in \{g.\xi \mid g \in G\}$ we know that $\nu_{\partial\mathbb{T}}(\eta) \leq a$. We claim that $\nu_{\partial\mathbb{T}}(\eta) = a$. Indeed, let us first suppose that there was an element $h \in \mathrm{supp}(\mu) \subseteq G$ with $\nu_{\partial\mathbb{T}}(h^{-1}.\xi) < a$. Then,

$$a = \nu_{\partial\mathbb{T}}(\xi) \overset{(*)}{=} \sum_{g\in G} \mu(g)\cdot \nu_{\partial\mathbb{T}}(g^{-1}.\xi) = \underbrace{\mu(h)\cdot \nu_{\partial\mathbb{T}}(h^{-1}.\xi)}_{<\,\mu(h)\cdot a} + \underbrace{\sum_{g\in G\smallsetminus\{h\}} \mu(g)\cdot \nu_{\partial\mathbb{T}}(g^{-1}.\xi)}_{\leq\,(1-\mu(h))\cdot a}.$$

This is a contradiction. Due to the irreducibility of the random walk, $\nu_{\partial\mathbb{T}}(h^{-1}.\xi) = a$ does not only hold for all $h \in \mathrm{supp}(\mu) \subseteq G$ but inductively for all $h \in G$, which proves our claim. But, the orbit $\{g.\xi \mid g \in G\}$ is infinite, so $1 = \nu_{\partial\mathbb{T}}(\partial\mathbb{T}) \geq |\{g.\xi \mid g \in G\}| \cdot a = \infty$. And this is, again, a contradiction.

[6] Consider such a finite initial piece, i. e. a finite reduced path from $B$ to some vertex $gB$. Given the end $\xi \in \partial\mathbb{T}$, we construct its image $g.\xi \in \partial\mathbb{T}$. This image will already have the correct finite initial piece unless cancellation takes place. It this case, consider the image $gb.\xi \in \partial\mathbb{T}$ instead. Since $|p| \neq 1$ and $|q| \neq 1$, cancellation will take place in at most one of the two cases.

If $\delta \geq 0$, the fact that the hitting measure $\nu_{\partial\mathbb{H}}$ is the Dirac measure at $\infty \in \partial\mathbb{H}$ is an immediate consequence of Lemma 2.1 and Brofferio's result mentioned on page 31. On the other hand, if $\delta < 0$, then $\nu_{\partial\mathbb{H}}(\infty) = 0$ by Lemma 2.2. Now, we can repeat the above argument. Suppose, there was an element of positive measure. Then, we may choose such an element $\xi \in \partial\mathbb{H} \setminus \{\infty\}$ of maximal measure. Again, all elements in its orbit $\{g.\xi \mid g \in G\}$ must have the same measure and, since the orbit is infinite, this yields a contradiction. □

**Lemma 2.11** *The hitting measure $\nu_{\partial\mathbb{T}}$ has full support, i.e. every non-empty open subset $A \subseteq \partial\mathbb{T}$ has positive measure. Moreover, if $\delta < 0$, the hitting measure $\nu_{\partial\mathbb{H}\times\partial\mathbb{T}}$ on the Cartesian product $\partial\mathbb{H} \times \partial\mathbb{T}$ has full support.*

*Proof.* We classify the ends of the tree $\mathbb{T}$ according to which neighbour of the vertex $B$ they first traverse. This yields a partition of the the space of ends into $p+q$ open subsets, i.e. the open balls of radius 1. At least one of them must have positive measure, call it $P \subseteq \partial\mathbb{T}$. Now, let $A \subseteq \partial\mathbb{T}$ be a non-empty open subset. In particular, there is a vertex $gB$, such that all ends traversing $gB$ belong to $A$. Similarly to the argument given in Footnote 6, either $g.P$ or $gb.P$ is contained in $A$, see Figure 10. We may assume w.l.o.g. that $g.P \subseteq A$. Moreover, due to the irreducibility of the random walk, there is an $n \in \mathbb{N}$ such that $\mu^{(n)}(g) > 0$, i.e. the probability to reach $g$ in precisely $n$ steps is positive. Hence,

$$\nu_{\partial\mathbb{T}}(A) \geq \mu^{(n)}(g) \cdot \nu_{\partial\mathbb{T},g}(A) \geq \mu^{(n)}(g) \cdot \nu_{\partial\mathbb{T},g}(g.P) = \mu^{(n)}(g) \cdot \nu_{\partial\mathbb{T}}(P) > 0\,.$$

The proof of the second assertion is similar. Since we have to keep track of two components, it is slightly more technical so that we give only a proof sketch. Recall from above the set $P \subseteq \partial\mathbb{T}$. Given the random variable $\mathrm{bnd}_{\partial\mathbb{T}}$ takes a value in $P$, at least one open interval $(k, k+1) \subseteq \partial\mathbb{H}$ with $k \in \mathbb{Z}$ will be hit by the random variable $\mathrm{bnd}_{\partial\mathbb{H}}$ with positive probability, call it $Q \subseteq \partial\mathbb{H}$. So, we know that $\nu_{\partial\mathbb{H}\times\partial\mathbb{T}}(Q \times P) > 0$. Now, let $A \subseteq \partial\mathbb{H} \times \partial\mathbb{T}$ be a non-empty open subset. By definition of the product topology, $A$ contains a rectangle of open sets $A_{\partial\mathbb{H}} \subseteq \partial\mathbb{H}$ and $A_{\partial\mathbb{T}} \subseteq \partial\mathbb{T}$. Now, we seek to construct an element $h \in G$ such that $h.Q \subseteq A_{\partial\mathbb{H}}$ and $h.P \subseteq A_{\partial\mathbb{T}}$, from where we may finally conclude as above that $\nu_{\partial\mathbb{H}\times\partial\mathbb{T}}(A) > 0$.

Again, there is a vertex $gB$ of the tree, such that all ends traversing $gB$ belong to $A_{\partial\mathbb{T}}$. Moreover, there is a number $r \in \mathbb{R}$ and an $\varepsilon > 0$ such that the open interval $(r, r+\varepsilon)$ is contained in $A_{\partial\mathbb{H}}$. Based on this data, we construct an element $h \in G$ of the form $h = gb^{k_1}a^{-k_2}b^{k_3}$ with the desired properties.

Let us first look at the tree component. The exponent $k_1$ is either 0 or 1, whichever ensures that the reduced path from $B$ to $gb^{k_1}a^{-1}B$ traverses $gB$. Now, let us turn to the hyperbolic component. The image $gb^{k_1}.Q \subseteq \partial\mathbb{H}$ is a bounded open interval. The exponent $k_2 \in \mathbb{N}$ is chosen in such a way that the length of the image $gb^{k_1}a^{-k_2}.Q \subseteq \partial\mathbb{H}$ is at most $\varepsilon/3$. Finally, there is an integer $k \in \mathbb{Z}$ such that both images $gb^{k_1}a^{-k_2}b^{k}.Q \subseteq \partial\mathbb{H}$ and $gb^{k_1}a^{-k_2}b^{k+1}.Q \subseteq \partial\mathbb{H}$ are contained in the interval $(r, r+\varepsilon)$ and therefore both belong to $A_{\partial\mathbb{H}}$. The exponent $k_3$ will be either $k$ or $k+1$. Let us return to the tree and choose it in such a way that all ends in the image $gb^{k_1}a^{-k_2}b^{k_3}.P$ traverse $gb^{k_1}a^{-k_2}B$. Then, by construction, they also traverse the vertex $gB$ and belong to $A_{\partial\mathbb{T}}$. □

We claim that these spaces are the Poisson-Fürstenberg boundaries of the random walk.

**Theorem 2.12 ("Identification theorem")** *Let $Z = (Z_0, Z_1, \dots)$ be a random walk on a non-amenable Baumslag-Solitar group $G = \mathrm{BS}(p,q)$ with $1 < p < q$. Assume there is an $\varepsilon > 0$ such that the increment $X_1$ has finite $(2+\varepsilon)$-th moment. If the vertical drift is non-negative, i.e. $\delta \geq 0$, then the Poisson-Fürstenberg boundary is isomorphic to $(\partial\mathbb{T}, \mathscr{B}_{\partial\mathbb{T}}, \nu_{\partial\mathbb{T}})$. On the other hand, if the vertical drift is negative, i.e. $\delta < 0$, then the Poisson-Fürstenberg boundary is isomorphic to $(\partial\mathbb{H} \times \partial\mathbb{T}, \mathscr{B}_{\partial\mathbb{H}\times\partial\mathbb{T}}, \nu_{\partial\mathbb{H}\times\partial\mathbb{T}})$.*

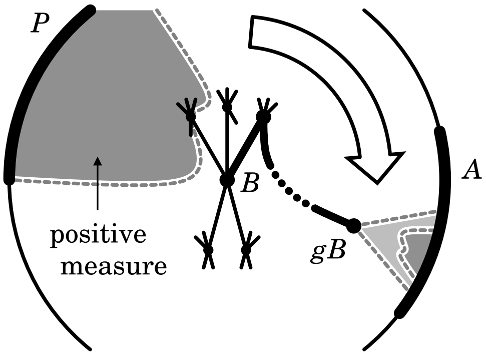


Figure 10: The hitting measure $\nu_{\partial\mathbb{T}}$ has full support.

*Proof.* As already mentioned, we seek to apply the strip criterion, see Theorem 1.25. By Lemma 1.27, the probability measure $\mu$ driving the random walk has finite entropy. Moreover, it is not hard to see that $\partial\mathbb{H}$ and $\partial\mathbb{T}$ are Polish spaces, and so is their product $\partial\mathbb{H}\times\partial\mathbb{T}$. Therefore, by Remark 1.19, the probability spaces $(\partial\mathbb{T},\mathscr{B}_{\partial\mathbb{T}},\nu_{\partial\mathbb{T}})$ and $(\partial\mathbb{H}\times\partial\mathbb{T},\mathscr{B}_{\partial\mathbb{H}\times\partial\mathbb{T}},\nu_{\partial\mathbb{H}\times\partial\mathbb{T}})$ are Lebesgue-Rohlin spaces. They are equipped with a left $G$-action and boundary maps $\mathrm{bnd}_{\partial\mathbb{T}}:\Omega\to\partial\mathbb{T}$ and $\mathrm{bnd}_{\partial\mathbb{H}\times\partial\mathbb{T}}:\Omega\to\partial\mathbb{H}\times\partial\mathbb{T}$, defined almost everywhere[7]. In order to show that they are $\mu$-boundaries, we have to ensure that the boundary maps are ① measurable, ② $\sim$-invariant, and ③ $G$-equivariant. But, all three properties are immediate by construction, compare also [Kaĭ00, End of §1.5].

If the vertical drift is negative, i. e. $\delta<0$, take the $\mu$-boundary $(\partial\mathbb{H}\times\partial\mathbb{T},\mathscr{B}_{\partial\mathbb{H}\times\partial\mathbb{T}},\nu_{\partial\mathbb{H}\times\partial\mathbb{T}})$ and the $\check{\mu}$-boundary $(\partial\mathbb{T},\mathscr{B}_{\partial\mathbb{T}},\check{\nu}_{\partial\mathbb{T}})$. Here, $\check{\nu}_{\partial\mathbb{T}}$ denotes the hitting measure of the random walk $\check{Z}=(\check{Z}_0,\check{Z}_1,\dots)$ driven by the probability measure $\check{\mu}$. Since $\delta<0$, the random walk $\check{Z}$ has positive vertical drift.

Next, we need to define gauges and strips. Let $S:=\{a,b\}\subseteq G$ be the standard generating set and define gauges $\mathscr{G}_k:=\{g\in G\mid d_S(1,g)\le k\}$. In other words, the gauges exhaust the group $G$ with balls centred at the identity $1\in G$, and the gauge function $|\cdot|=|\cdot|_{\mathscr{G}}$ is nothing but the distance to 1 with respect to the word metric $d_S$.

By Lemma 2.10, we know that $\check{\nu}_{\partial\mathbb{T}}\otimes\nu_{\partial\mathbb{H}\times\partial\mathbb{T}}$-almost every pair of points $(\xi_-,(r_+,\xi_+))\in\partial\mathbb{T}\times(\partial\mathbb{H}\times\partial\mathbb{T})$ has distinct ends $\xi_-,\xi_+\in\partial\mathbb{T}$ and a boundary value $r_+\in\mathbb{R}$. In this situation, we may connect $\xi_-$ and $\xi_+$ by a unique doubly infinite reduced path $v:\mathbb{Z}\to\mathbb{T}$ and define the strip $S(\xi_-,(r_+,\xi_+))$ as follows. It consists of all group elements $g\in G$ that are contained in the $\pi_{\mathbb{T}}$-preimage of $v$, i. e. their image $\pi_{\mathbb{T}}(g)$ is traversed by $v$, and have the property that the real part $\Re(\pi_{\mathbb{H}}(g))$ has minimal distance to $r_+\in\mathbb{R}$ among all real parts $\Re(\pi_{\mathbb{H}}(h))$ with $h\in gB$, see the left-hand side of Figure 11. To all remaining pairs we assign the whole of $G$ as a strip. This way, the map $S$ becomes measurable and $G$-equivariant. By Lemma 2.11, a random strip contains the identity element $1\in G$ with positive probability, i. e. the map $S$ satisfies the inequality of Remark 1.26. So, it suffices to verify the following convergence for an arbitrary pair $(\xi_-,(r_+,\xi_+))\in\partial\mathbb{T}\times(\partial\mathbb{H}\times\partial\mathbb{T})$ of the first kind,

$$
{}^1\!/\!_n\cdot\ln\big(\mathrm{card}\big(S(\xi_-,(r_+,\xi_+))\cap\mathscr{G}_{|Z_n|}\big)\big)\xrightarrow[\text{a. s.}]{n\to\infty}0\,.
$$

But, the strip $S(\xi_-,(r_+,\xi_+))$ intersects the gauge $\mathscr{G}_{|Z_n|}$ in at most $2\cdot|Z_n|+1$ many cosets of the form $G/B$,

[7] To be precise, we should declare what to do with the remaining trajectories. Since $(\Omega,\mathscr{A},\mathbb{P})$ is complete, we may extend the definition arbitrarily. However, for the following, it will be more convenient to add another point † equipped with the trivial left $G$-action to each codomain onto which all remaining trajectories are mapped. It has hitting measure 0 and, in the end, can be removed without changing the isomorphism type of the Poisson-Fürstenberg boundary.

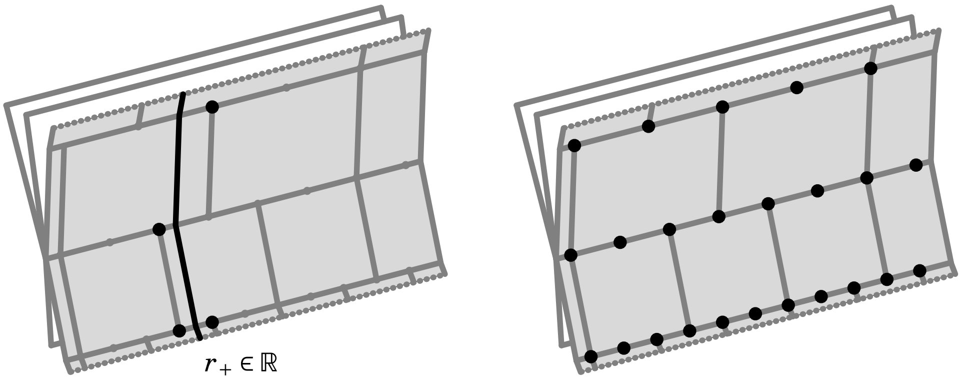


Figure 11: Strips for the cases $\delta \neq 0$ (left) and $\delta = 0$ (right).

and each of them contains at most two elements of the strip. Therefore,

$$^1\!/_n \cdot \ln\left(\operatorname{card}\left(S(\check{\xi},(r,\xi)) \cap \mathscr{G}_{|Z_n|}\right)\right) \leq \frac{\ln((2\cdot|Z_n|+1)\cdot 2)}{n} = \frac{\ln((2\cdot d_S(1,Z_n)+1)\cdot 2)}{n} \xrightarrow[\text{a.s.}]{n\to\infty} 0\,.$$

In the final step of the above calculation, we use again that the increments $X_1$ have finite first moment. Indeed, $^1\!/_n \cdot d_S(1,Z_n) = {}^1\!/_n \cdot d_S(1, X_1 \cdot \ldots \cdot X_n) \leq {}^1\!/_n \cdot \sum_{k=1}^n d_S(1,X_k) \to \mathbb{E}(d_S(1,X_1))$ a. s., from where we may conclude that the sequence $^1\!/_n \cdot d_S(1,Z_n)$ is a. s. bounded and that the above calculation is correct.

So, we can finally apply the strip criterion and obtain that $(\partial\mathbb{H}\times\partial\mathbb{T}, \mathscr{B}_{\partial\mathbb{H}\times\partial\mathbb{T}}, \nu_{\partial\mathbb{H}\times\partial\mathbb{T}})$ is isomorphic to the Poisson-Fürstenberg boundary. Vice versa, if the vertical drift is positive, i. e. $\delta > 0$, the same argument yields that $(\partial\mathbb{T}, \mathscr{B}_{\partial\mathbb{T}}, \nu_{\partial\mathbb{T}})$ is isomorphic to the Poisson-Fürstenberg 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 allows us to solve this issue. More precisely, take the $\mu$-boundary $(\partial\mathbb{T}, \mathscr{B}_{\partial\mathbb{T}}, \nu_{\partial\mathbb{T}})$ and the $\check{\mu}$-boundary $(\partial\mathbb{T}, \mathscr{B}_{\partial\mathbb{T}}, \check{\nu}_{\partial\mathbb{T}})$. Now, define gauges

$$\mathscr{G}_k := \left\{ g \in G \,\middle|\, d_{\mathbb{T}}(\pi_{\mathbb{T}}(1), \pi_{\mathbb{T}}(g)) \leq k^2 \text{ and } d_{\mathbb{H}}(\pi_{\mathbb{H}}(1), \pi_{\mathbb{H}}(g)) \leq k \right\}.$$

Again, we know that $\check{\nu}_{\partial\mathbb{T}} \otimes \nu_{\partial\mathbb{T}}$-almost every pair of points $(\xi_-, \xi_+) \in \partial\mathbb{T}\times\partial\mathbb{T}$ has distinct ends $\xi_-, \xi_+ \in \partial\mathbb{T}$, which we may connect by a unique doubly infinite reduced path $v : \mathbb{Z} \to \mathbb{T}$. Let $S(\xi_-,\xi_+)$ be the full $\pi_{\mathbb{T}}$-preimage of $v$, i. e. the set of all group elements $g \in G$ such that the image $\pi_{\mathbb{T}}(g)$ is traversed by $v$, see the right-hand side of Figure 11. Again, to all remaining pairs we assign the whole of $G$ as a strip. This way, the map $S$ becomes measurable, $G$-equivariant, and satisfies the inequality of Remark 1.26. Now, pick an arbitrary pair $(\xi_-,\xi_+) \in \partial\mathbb{T}\times\partial\mathbb{T}$ of the first kind. We claim that

$$^1\!/_n \cdot \ln\left(\operatorname{card}\left(S(\xi_-,\xi_+) \cap \mathscr{G}_{|Z_n|}\right)\right) \leq \frac{\ln\left((2\cdot|Z_n|^2+1)\cdot\exp(|Z_n|+2)\right)}{n} = \underbrace{\frac{\ln\left(2\cdot|Z_n|^2+1\right)}{n}}_{①} + \underbrace{\frac{|Z_n|+2}{n}}_{②}.$$

Indeed, the inequality holds for a similar reason as above; the strip $S(\xi_-,\xi_+)$ intersects the gauge $\mathscr{G}_{|Z_n|}$ in at most $2\cdot|Z_n|^2+1$ many cosets of the form $G/B$. 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 the horizontal line $L \subseteq \mathbb{H}$ with imaginary part $y := \Im(\pi_{\mathbb{H}}(g))$. One necessary condition for such an element $h \in gB$ to be contained in the gauge $\mathscr{G}_{|Z_n|}$ is that the projection $\pi_{\mathbb{H}}(h)$ is contained in the ball $B := \{z \in \mathbb{H} \mid d_{\mathbb{H}}(i,z) \leq |Z_n|\} \subseteq \mathbb{H}$. If $L \cap B$ is empty, then the coset $gB$ does

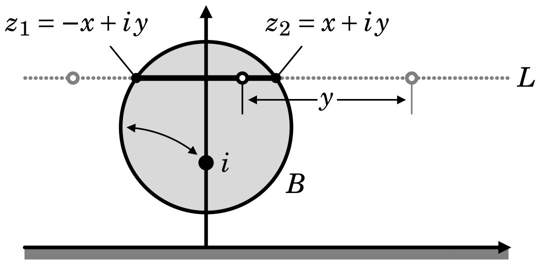


Figure 12: The horizontal line $L$, the ball $B$, and their intersection $L \cap B$.

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 $L \cap B$ is the horizontal line between $z_1 := -x + iy$ to $z_2 := x + iy$, see Figure 12. The projections $\pi_{\mathbb{H}}(h)$ with $h \in gB$ have the property that the real parts $\Re(\pi_{\mathbb{H}}(h))$ and $\Re(\pi_{\mathbb{H}}(hb))$ differ precisely by $y$. So, the horizontal line $L \cap B$ contains at most $1 + 2x/y$ many of them. Let us now estimate $1 + 2x/y$ in terms of $|Z_n|$. Since $z_1$ and $z_2$ are both contained in $B$, their distance is at most $2 \cdot |Z_n|$. Therefore,

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

And, in particular,

$$\exp(2 \cdot |Z_n|) \geq 1 + \frac{2x^2}{y^2}, \implies \exp(2 \cdot |Z_n|) \geq \frac{2x^2}{y^2}, \implies \exp(2 \cdot |Z_n| + \ln(2)) \geq \frac{4x^2}{y^2},$$
$$\implies \exp(|Z_n| + 1/2 \cdot \ln(2)) \geq \frac{2x}{y}, \implies \exp(|Z_n| + 2) \geq 1 + \frac{2x}{y}.$$

So, the coset $gB$ contains at most $\exp(|Z_n| + 2)$ many elements of the gauge. We will now show that both summands ① and ② converge a. s. to 0, which completes the proof. Concerning ①, it suffices to observe that $|Z_n| \leq \max\{\ell_a, \ell_b, 1\} \cdot d_{\mathrm{S}}(1, Z_n) + 1$ and therefore, as above,

$$① = \frac{\ln\left(2 \cdot |Z_n|^2 + 1\right)}{n} \leq \frac{\ln\left(2 \cdot (\max\{\ell_a, \ell_b, 1\} \cdot d_{\mathrm{S}}(1, Z_n) + 1)^2 + 1\right)}{n} \xrightarrow[\text{a. s.}]{n \to \infty} 0.$$

On the other hand, concerning ②, we observe that

$$|Z_n| - 1 \leq \max\left\{d_{\mathbb{H}}(\pi_{\mathbb{H}}(1), \pi_{\mathbb{H}}(Z_n)), \sqrt{d_{\mathbb{T}}(\pi_{\mathbb{T}}(1), \pi_{\mathbb{T}}(Z_n))}\right\} \leq \max\left\{d_{\mathbb{H}}(\pi_{\mathbb{H}}(1), \pi_{\mathbb{H}}(Z_n)), \sqrt{d_{\mathrm{S}}(1, Z_n)}\right\}.$$

Hence, by Lemma 2.3,

$$② = \frac{|Z_n| + 2}{n} \leq \frac{\max\left\{d_{\mathbb{H}}(\pi_{\mathbb{H}}(1), \pi_{\mathbb{H}}(Z_n)), \sqrt{d_{\mathrm{S}}(1, Z_n)}\right\} + 3}{n} \xrightarrow[\text{a. s.}]{n \to \infty} 0.$$

□

# A Appendix: The remaining non-amenable cases

Recall from Section 1.1.1 that a Baumslag-Solitar group $\mathrm{BS}(p, q)$ is non-amenable if and only if neither $|p| = 1$ nor $|q| = 1$. Until now, we have only considered non-amenable Baumslag-Solitar groups $\mathrm{BS}(p, q)$ with $1 < p < q$. Replacing one of the generators by its inverse, it is easy to see that

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

So, in order to investigate the remaining non-amenable Baumslag-Solitar groups, the only cases that we have to consider are $1 < p < -q$ and $1 < p = |q|$. In this appendix, we shall review our methods from Section 2 and explain how to adjust the arguments in order to obtain similar results for these cases.

## A.1 Action by suitable isometries on the hyperbolic plane

Let us first assume that $G = \mathrm{BS}(p,q)$ with $1 < p < -q$. In order to define the projection $\pi_{\mathbb{H}} : G \to \mathbb{H}$ back in Section 1.1.3, we considered the map $\pi_{\mathrm{Aff}^+(\mathbb{R})} : \{a,b\} \to \mathrm{Aff}^+(\mathbb{R})$ given by $\pi_{\mathrm{Aff}^+(\mathbb{R})}(a) := (x \mapsto {}^{q}/_{p} \cdot x)$ and $\pi_{\mathrm{Aff}^+(\mathbb{R})}(b) := (x \mapsto x+1)$, and extended it to a homomorphism.

Now, we are assuming that $1 < p < -q$, in which case the transformation $x \mapsto {}^{q}/_{p} \cdot x$ is not orientation preserving any more. If we replaced $q$ by $|q|$ in the definition, then $\pi_{\mathrm{Aff}^+(\mathbb{R})}(a) := (x \mapsto {}^{|q|}/_{p} \cdot x)$ would be orientation preserving but $\pi_{\mathrm{Aff}^+(\mathbb{R})}(a) \circ \pi_{\mathrm{Aff}^+(\mathbb{R})}(a)^p \circ \pi_{\mathrm{Aff}^+(\mathbb{R})}(a)^{-1} \neq \pi_{\mathrm{Aff}^+(\mathbb{R})}(b)^q$, whence we could not apply von Dyck's theorem any more. So, we have to change the approach.

Let $A$ be the set of all maps $\varphi : \mathbb{C} \to \mathbb{C}$ either of the form $\varphi(z) = \alpha z + \beta$ or of the form $\varphi(z) = \alpha \cdot (-\overline{z}) + \beta$ with $\alpha, \beta \in \mathbb{R}$ and $\alpha > 0$. This set, endowed with the composition, forms again a group. Consider the map $\pi_A : \{a,b\} \to A$ given by $\pi_A(a) := (z \mapsto {}^{|q|}/_{p} \cdot (-\overline{z}))$ and $\pi_A(b) := (z \mapsto z+1)$. With this map, it is possible to apply von Dyck's theorem and to extend it uniquely to a group homomorphism $\pi_A : G \to A$. Finally, as in the case of $\mathrm{Aff}^+(\mathbb{R})$, every $\varphi \in A$ can be thought of as an isometry of $\mathbb{H}$. So, we may consider the projection $\pi_{\mathbb{H}} : G \to \mathbb{H}$ given by $\pi_{\mathbb{H}}(g) := \pi_A(g)(i)$. The following lemma illustrates this definition.

**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 is $\ell_a := \ln {}^{|q|}/_{p}$. But, for every $g \in G$ the point $\pi_{\mathbb{H}}(gb) \in \mathbb{H}$ is either right or left from 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 is $\ell_b := \ln \frac{3+\sqrt{5}}{2}$.*

*Proof sketch.* The proof is similar to the one of Lemma 1.4. So, we only discuss the differences. Let us consider the two points $\pi_{\mathbb{H}}(1) \in \mathbb{H}$ and $\pi_{\mathbb{H}}(b) \in \mathbb{H}$. If $\pi_A(g)$ is of the form $z \mapsto \alpha z + \beta$, then it is again the composition of a dilation $z \mapsto \alpha z$ and a translation $z \mapsto z + \beta$, whence the relative position of the two points is preserved. On the other hand, if $\pi_A(g)$ is of the form $z \mapsto \alpha \cdot (-\overline{z}) + \beta$, then it is the composition of a reflection at the imaginary axis $z \mapsto -\overline{z}$, a dilation $z \mapsto \alpha z$, and a translation $z \mapsto z + \beta$, in which case the relative position of the two points is preserved with the exception that right and left are switched.

In order to decide whether $\pi_A(g)$ is of the first or the second form, we can write the element $g \in G$ as a product over $a^{\pm 1}$ and $b^{\pm 1}$. Since $\pi_A$ is a homomorphism, the image $\pi_A(g)$ can be written as the respective product over $\pi_A(a^{\pm 1})$ and $\pi_A(b^{\pm 1})$. But, each occurrence of $\pi_A(a^{\pm 1})$ yields one reflection. So, the image $\pi_A(g)$ is of the first form if and only if $\lambda(g)$ is even. □

Using this projection $\pi_{\mathbb{H}} : G \to \mathbb{H}$ and, of course, replacing $q$ by $|q|$ wherever it is necessary, we can repeat the arguments from Section 2. The definition of the tree $\mathbb{T}$ and the level functions $\lambda$ and $\widetilde{\lambda}$, including Lemma 1.3, as well as the definition of the discrete hyperbolic plane $\Gamma_v$, including Lemma 1.7, can be adapted. Recall that, in Section 1.2.4, we considered the imaginary and real parts of $\pi_{\mathbb{H}}(g) \in \mathbb{H}$ separately, and introduced the shorthand notation $A_g := \Im(\pi_{\mathbb{H}}(g))$ and $B_g := \Re(\pi_{\mathbb{H}}(g))$. Let us highlight that we now have $\ln(A_g) = \ln({}^{|q|}/_{p}) \cdot \lambda(g)$, which allows us to adapt the proof of Lemma 1.13.

In order to identify the Poisson-Fürstenberg boundary geometrically, we have to ensure convergence to the boundaries $\partial\mathbb{H}$ and $\partial\mathbb{T}$. Let us first consider the boundary $\partial\mathbb{H}$. The proof of Lemma 2.1 for $\delta > 0$ can be adapted. The proof of Lemma 2.2 for $\delta < 0$, in turn, deserves a bit of work. We have to show that the real parts $B_{Z_n}$ converge a. s. to a random element $\xi \in \partial\mathbb{H} \setminus \{\infty\}$. In the original proof, we observed that $A_{Z_n} = A_{X_1} \cdot \ldots \cdot A_{X_n}$ and $B_{Z_n} = \sum_{k=1}^{n} C_k$ with $C_k := A_{X_1} \cdot \ldots \cdot A_{X_{k-1}} \cdot B_{X_k}$. While the first formula remains true, the second one does not. 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. However, instead of the above, we obtain that $C_k := \varepsilon_{X_1} \cdot A_{X_1} \cdot \ldots \cdot \varepsilon_{X_{k-1}} \cdot A_{X_{k-1}} \cdot B_{X_k}$ with $\varepsilon_g := 1$ if $\lambda(g)$ is even and $\varepsilon_g := -1$ if $\lambda(g)$ is odd. This observation allows us to apply Cauchy's root test precisely as in the proof of Lemma 2.2. For the same reason, namely because all the estimates are not in terms of the the actual horizontal increments but of their absolute values, the proofs of Lemmas 2.3 and 2.4 for $\delta = 0$ can be adapted. The same holds, concerning the boundary $\partial\mathbb{T}$, for the proofs of Lemma 2.6 for $\delta \neq 0$ and Lemma 2.7 for $\delta = 0$.

**Remark A.2** *Notice that we do not claim that Brofferio's result still holds. It said that, if there is no vertical drift, i. e. $\delta = 0$, then the projections $\pi_{\mathbb{H}}(Z_n)$ converge a. s. to $\infty \in \partial\mathbb{H}$. However, this point is not of relevance for us. If there is no vertical drift, we do not need to take the hyperbolic component into account because our candidate for the Poisson-Fürstenberg boundary is $(\partial\mathbb{T}, \mathscr{B}_{\partial\mathbb{T}}, \nu_{\partial\mathbb{T}})$.*

From these observations, we may deduce the following results.

**Theorem A.3 (Version of the "convergence theorem")** *Let $Z = (Z_0, Z_1, \ldots)$ be a random walk on a non-amenable Baumslag-Solitar group $G = \mathrm{BS}(p,q)$ with $1 < p < -q$. Assume there is an $\varepsilon > 0$ such that the increment $X_1$ has finite $(2+\varepsilon)$-th moment. If the vertical drift is different from 0, i. e. $\delta \neq 0$, then the projections $\pi_{\mathbb{H}}(Z_n)$ converge a. s. to a random element in $\partial\mathbb{H}$ and the projections $\pi_{\mathbb{T}}(Z_n)$ converge a. s. to a random element in $\partial\mathbb{T}$. If there is no vertical drift, i. e. $\delta = 0$, then the projections $\pi_{\mathbb{H}}(Z_n)$ have sublinear speed and the projections $\pi_{\mathbb{T}}(Z_n)$ converge a. s. to a random element in $\partial\mathbb{T}$.*

**Theorem A.4 (Version of the "identification theorem")** *Let $Z = (Z_0, Z_1, \ldots)$ be a random walk on a non-amenable Baumslag-Solitar group $G = \mathrm{BS}(p,q)$ with $1 < p < -q$. Assume there is an $\varepsilon > 0$ such that the increment $X_1$ has finite $(2+\varepsilon)$-th moment. If the vertical drift is non-negative, i. e. $\delta \geq 0$, then the Poisson-Fürstenberg boundary is isomorphic to $(\partial\mathbb{T}, \mathscr{B}_{\partial\mathbb{T}}, \nu_{\partial\mathbb{T}})$. On the other hand, if the vertical drift is negative, i. e. $\delta < 0$, then the Poisson-Fürstenberg boundary is isomorphic to $(\partial\mathbb{H} \times \partial\mathbb{T}, \mathscr{B}_{\partial\mathbb{H}\times\partial\mathbb{T}}, \nu_{\partial\mathbb{H}\times\partial\mathbb{T}})$.*

## A.2 Action by isometries on the Euclidean plane

Let us now assume that $G = \mathrm{BS}(p,q)$ with $1 < p = |q|$. This situation differs fundamentally from the ones discussed so far.[8] We use the Euclidean plane $\mathbb{R}^2$ instead on the hyperbolic plane $\mathbb{H}$. In order to construct a projection $\pi_{\mathbb{R}^2} : G \to \mathbb{R}^2$, let $A := \mathrm{Isom}(\mathbb{R}^2)$ and consider the map $\pi_A : \{a, b\} \to A$ given by

$$\pi_A(b) := \left( \begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} x+1 \\ y \end{pmatrix} \right) \quad \text{and} \quad \pi_A(a) := \begin{cases} \left( \begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} x \\ y+1 \end{pmatrix} \right) & \text{if } q > 0 \\ \left( \begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} -x \\ y+1 \end{pmatrix} \right) & \text{if } q < 0. \end{cases}$$

In both cases, it is possible to apply von Dyck's theorem and to extend the map uniquely to a group homomorphism $\pi_{\mathbb{R}^2} : G \to A$. So, we may consider the projection $\pi_{\mathbb{R}^2} : G \to \mathbb{R}^2$ given by $\pi_{\mathbb{R}^2}(g) := \pi_A(g)(0)$.

The definition of the tree $\mathbb{T}$ and the level functions $\lambda$ and $\widetilde{\lambda}$, including Lemma 1.3, remain the same. But, instead of the discrete hyperbolic plane, we now obtain a discrete Euclidean plane $\Gamma_v$. The proof of Lemma 1.7 can be adapted to the new situation and shows that the graph $\Gamma_v$, endowed with the graph distance $d_{\Gamma_v}$, is quasi-isometric to the Euclidean plane $\mathbb{R}^2$, endowed with the standard metric $d_{\mathbb{R}^2}$.

[8] Let us motivate the new construction by a heuristic argument. Consider the brick illustrated in Figure 3. If $1 < p = |q|$, then the brick would have to have the same number of $\mathbb{H}$-edges on the upper and the lower horizontal line. This would be possible if the horizontal lines agreed or we worked in a space without curvature. And the latter is what we are going to do.

We aim to show that, as soon as the projections converge to a random element in $\partial\mathbb{T}$, independently of the vertical drift, the Poisson-Fürstenberg boundary is $(\partial\mathbb{T}, \mathscr{B}_{\partial\mathbb{T}}, \nu_{\partial\mathbb{T}})$. So, we don't need to worry about a boundary for the projections $\pi_{\mathbb{R}^2}(Z_n)$. Concerning the projections $\pi_{\mathbb{T}}(Z_n)$, we have to distinguish between two cases. If the vertical drift is different from 0, i. e. $\delta \neq 0$, then the proof of Lemma 2.6 can be adapted and we obtain that the projections $\pi_{\mathbb{T}}(Z_n)$ still converge a. s. to a random end in $\partial\mathbb{T}$. But, if there is no vertical drift, then the proof of Lemma 2.7 cannot be adapted because it was based on the fact that the projections $\pi_{\mathbb{H}}(Z_n)$ had sublinear speed. Now, the projections $\pi_{\mathbb{R}^2}(Z_n)$ do not need to have this property any more. However, we conjecture that if the probability measure $\mu$ driving the random walk $Z = (Z_0, Z_1, \dots)$ has finite support, the projections $\pi_{\mathbb{T}}(Z_n)$ eventually leave every finite ball and hence converge a. s. to a random end in $\partial\mathbb{T}$.

As soon as they do, we can show as in Lemmas 2.10 and 2.11 that the hitting measure $\nu_{\mathbb{T}}$ is again non-atomic and has full support. Moreover, we obtain the following version of the identification theorem.

**Theorem A.5 (Version of the "identification theorem")** *Let $Z = (Z_0, Z_1, \dots)$ be a random walk on a non-amenable Baumslag-Solitar group $G = \mathrm{BS}(p,q)$ with $1 < p = |q|$. Assume that the increment $X_1$ has finite first moment. If the projections $\pi_{\mathbb{T}}(Z_n)$ converge a. s. to a random end in $\partial\mathbb{T}$, e. g. if the vertical drift is different from 0, i. e. $\delta \neq 0$, then the Poisson-Fürstenberg boundary is isomorphic to $(\partial\mathbb{T}, \mathscr{B}_{\partial\mathbb{T}}, \nu_{\partial\mathbb{T}})$.*

*Proof sketch.* Again, we apply the strip criterion. As in the original proof of the identification theorem, take the $\mu$-boundary $(\partial\mathbb{T}, \mathscr{B}_{\partial\mathbb{T}}, \nu_{\partial\mathbb{T}})$ and the $\check{\mu}$-boundary $(\partial\mathbb{T}, \mathscr{B}_{\partial\mathbb{T}}, \check{\nu}_{\partial\mathbb{T}})$. Next, we define gauges

$$\mathscr{G}_k := \left\{ g \in G \,\middle|\, d_{\mathbb{T}}(\pi_{\mathbb{T}}(1), \pi_{\mathbb{T}}(g)) \leq k \text{ and } d_{\mathbb{R}^2}(\pi_{\mathbb{R}^2}(1), \pi_{\mathbb{R}^2}(g)) \leq k \right\}.$$

We know that $\check{\nu}_{\partial\mathbb{T}} \otimes \nu_{\partial\mathbb{T}}$-almost every pair of points $(\xi_-, \xi_+) \in \partial\mathbb{T} \times \partial\mathbb{T}$ has distinct ends $\xi_-, \xi_+ \in \partial\mathbb{T}$, which we may connect by a unique doubly infinite reduced path $v : \mathbb{Z} \to \mathbb{T}$.

Let $S(\xi_-, \xi_+)$ be the full $\pi_{\mathbb{T}}$-preimage of $v$. To all remaining pairs we assign the whole of $G$ as a strip. This way, the map $S$ becomes measurable, $G$-equivariant, and satisfies the inequality of Remark 1.26. Now, pick an arbitrary pair $(\xi_-, \xi_+) \in \partial\mathbb{T} \times \partial\mathbb{T}$ of the first kind. We claim that

$$1/n \cdot \ln\left(\mathrm{card}\left(S(\xi_-, \xi_+) \cap \mathscr{G}_{|Z_n|}\right)\right) \leq \frac{\ln\left((2 \cdot |Z_n| + 1) \cdot (2 \cdot |Z_n| + 1)\right)}{n} = \dots$$

Indeed, the inequality is easy to see. The strip $S(\xi_-, (r_+, \xi_+))$ intersects the gauge $\mathscr{G}_{|Z_n|}$ in at most $2 \cdot |Z_n| + 1$ many cosets of the form $G/B$, 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 $|Z_n| \leq d_{\mathrm{S}}(1, Z_n) + 1$. Then, as above, using the fact that $1/n \cdot d_{\mathrm{S}}(1, Z_n)$ is a. s. bounded, we may conclude that

$$\dots = \frac{\ln\left((2 \cdot |Z_n| + 1)^2\right)}{n} \leq \frac{\ln\left((2 \cdot d_{\mathrm{S}}(1, Z_n) + 3)^2\right)}{n} \xrightarrow[\text{a. s.}]{n \to \infty} 0,$$

which allows us to apply the strip criterion. $\square$

# Project B

# The Tits alternative for non-spherical triangles of groups

**( with Jörg Lehnert )**

*The following text is essentially the corresponding publication: Transactions of the London Mathematical Society, to appear.*


### Abstract

Triangles of groups were introduced by Gersten and Stallings in 1991. They are, roughly speaking, a generalisation of the amalgamated free product of two groups and occur in the framework of Corson diagrams. First, we prove an intersection theorem for Corson diagrams. Then, we focus on triangles of groups. It has been shown by Howie and Kopteva that the colimit of a hyperbolic triangle of groups contains a non-abelian free subgroup. We give two natural conditions, each of which ensures that the colimit of a non-spherical triangle of groups either contains a non-abelian free subgroup or is virtually solvable.



**Keywords:** Tits alternative, triangles of groups, disc pictures, metric simplicial complexes, non-positive curvature.

**MSC classes:** 20E06 (primary), 20F05, 20F65, 20E05, 20E07 (secondary).


## 1 Introduction

Given a commutative diagram of groups and injective homomorphisms, we may construct its colimit (in the category of groups). The colimit, or, more precisely, the colimit group, can be obtained by taking the free product of the groups and identifying the factors according to the homomorphisms. A good example is the amalgamated free product $X *_A Y$, which is the colimit group of the diagram $X \leftarrow A \rightarrow Y$.

We are interested in Corson diagrams. A Corson diagram is based on a set $I$. For every subset $J \subseteq I$ with $|J| \leq 2$ there is a group $G_J$ and for every two subsets $J_1 \subset J_2 \subseteq I$ with $|J_2| \leq 2$ there is a homomorphism $\varphi_{J_1 J_2} : G_{J_1} \rightarrow G_{J_2}$, see Figure 1. Notice that both Artin groups and Coxeter groups have a natural interpretation as colimit groups of Corson diagrams. A triangle of groups is nothing but a Corson diagram based on a set $I$ with $|I| = 3$.

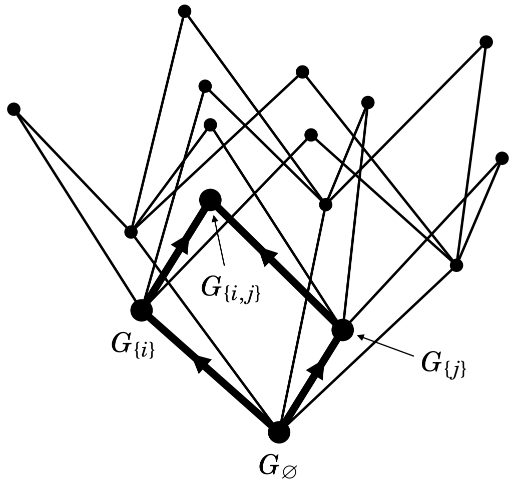


Figure 1: A Corson diagram based on a set $I$ with $|I| = 5$. For simplicity, the homomorphisms $\varphi_{\varnothing J} : G_{\varnothing} \to G_J$ with $J \subseteq I$ and $|J| = 2$ have been omitted in this figure. Notice that, if $|I| = n$, then the graphical representation of the Corson diagram has exactly $1/2 \cdot (n^2 + n + 2)$ vertices.

Gersten and Stallings introduced the notion of curvature and proved that for non-spherical triangles of groups the natural homomorphisms $\nu_J$ from the groups $G_J$ to the colimit group $\mathfrak{G}$ are injective, see [Sta91]. A similar result holds for non-spherical Corson diagrams, see [Cor96]. While these two results can be proved by nice arguments based on Euler's formula for planar graphs, spherical Corson diagrams are much harder to investigate, see e. g. [Che95] and [All12].

In Section 2, we introduce the basic notions for this paper. Then, in Section 3, we give an example of a spherical triangle of groups showing that, even though the natural homomorphisms $\nu_J : G_J \to \mathfrak{G}$ are injective, the intersections of their images may be larger than the amalgamated subgroups. But this can only happen in the spherical realm. For non-spherical triangles of groups and, more generally, non-spherical Corson diagrams, there are no large intersections, see Theorem 3.8.

It seems worth mentioning that the absence of large intersections shall not be confused with the developability of complexes of groups, which is implied by the injectivity of the natural homomorphisms $\nu_J : G_J \to \mathfrak{G}$, see [BH99, Corollary III.C.2.15].

Howie and Kopteva showed that, under mild assumptions, the colimit group of a hyperbolic triangle of groups has a non-abelian free subgroup, see [HK06]. In Section 4, we focus on the Euclidean case and discuss the following version of the Tits alternative: *A class $\mathscr{C}$ of groups satisfies the Tits alternative if each $G \in \mathscr{C}$ either has a non-abelian free subgroup or is virtually solvable.* The Tits alternative is named after Jacques Tits, who proved in 1972 that the class of finitely generated linear groups has this property, see [Tit72, Corollary 1]. Since then, the Tits alternative has been proved for many other classes of groups. For a list of results and open problems we refer to [SW05].

As indicated above, we are interested in Euclidean triangles of groups. In the case that none of the Gersten-Stallings angles is 0, we may follow Bridson's construction of a simplicial complex $\mathscr{X}$, see [Bri91], and use billiards on a suitable triangle in the Euclidean plane to obtain geodesics in the geometric realisation $|\mathscr{X}|$. These geodesics allow us to prove that, as soon as the simplicial complex $\mathscr{X}$ branches, the colimit group has a non-abelian free subgroup, see Theorems 4.18 and 4.19.

The remaining cases can be analysed with quotients and amalgamated free products. In the end, we generalise the result by Howie and Kopteva mentioned above and prove that the Tits alternative holds for the class of colimit groups of non-spherical triangles of groups with the property that none of the Gersten-Stallings angles is 0 and the group $G_\varnothing$ either has a non-abelian free subgroup or is virtually solvable, see Theorem 4.24, or with the property that every group $G_J$ with $J \subseteq I$ and $|J| \leq 2$ either has a non-abelian free subgroup or is virtually solvable, see Theorem 4.25.

### Acknowledgements

This paper originates in a Diplomarbeit, the equivalent of a master's thesis, under supervision of Robert Bieri at Goethe-Universität Frankfurt am Main. Over the last couple of years, various improvements have been made. But, in light of this beginning, we would like to thank Robert Bieri for his enthusiasm, advice, and patience. Moreover, we would like to thank the referee for making many comments that helped us to brush up this paper.

## 2 Preliminaries

### 2.1 Corson diagrams and their colimits

Let $I$ be an arbitrary set. Assume we are given for every subset $J \subseteq I$ with $|J| \leq 2$ a group $G_J$ and for every two subsets $J_1 \subset J_2 \subseteq I$ with $|J_2| \leq 2$ an injective homomorphism $\varphi_{J_1 J_2} : G_{J_1} \to G_{J_2}$. Moreover, assume the resulting diagram to be *commutative*, i. e. for every sequence of subsets $\varnothing = J_1 \subset J_2 \subset J_3 \subseteq I$ with $|J_3| = 2$ the equation $\varphi_{J_1 J_3} = \varphi_{J_2 J_3} \circ \varphi_{J_1 J_2}$ holds.

Since these diagrams were introduced by Corson in [Cor96], we refer to them as *Corson diagrams*. In the case $|I| = 3$, Corson diagrams are known as *triangles of groups*. Whenever we consider a triangle of groups, we may assume w. l. o. g. that $I = \{1,2,3\}$.

Given a Corson diagram, we will mostly be interested in its *colimit group*. The colimit group can be obtained by taking the free product of the groups $G_J$ and identifying the factors according to the homomorphisms. Let us make this construction a little more precise. Think of each $G_J$ as a set, and let $R_J$ be the set of all words over the group elements and their formal inverses that represent the identity. Then, the colimit group $\mathfrak{G}$ is given by the following presentation:

$$\mathfrak{G} = \left\langle \bigsqcup_{\substack{J \subseteq I \\ |J| \leq 2}} G_J : \bigsqcup_{\substack{J \subseteq I \\ |J| \leq 2}} R_J, \bigsqcup_{\substack{J_1 \subset J_2 \subseteq I \\ |J_2| \leq 2}} \{g = \varphi_{J_1 J_2}(g) : g \in G_{J_1}\} \right\rangle \tag{$*$}$$

This presentation, though not very economic, turns out to be suitable for our purposes. For every subset $J \subseteq I$ with $|J| \leq 2$ we may consider the natural homomorphism $\nu_J : G_J \to \mathfrak{G}$ given by $g \mapsto g$. The colimit group, equipped with these homomorphisms, is called the *colimit*. For further reading about it we refer to [Tit86, 1.1] and [AHS90, Chapter 3].

### 2.2 Curvature of Corson diagrams

The homomorphisms $\nu_J : G_J \to \mathfrak{G}$ do not need to be injective. An example of a triangle of groups in which they are not has been given by Gersten and Stallings in [Sta91, 1.4]. On the other hand, it turns out that for non-spherical triangles of groups and, more generally, for non-spherical Corson diagrams they are. Let us therefore introduce the notion of curvature. For every two distinct $i, j \in I$ the homomorphisms $\varphi_{\{i\}\{i,j\}}$ and $\varphi_{\{j\}\{i,j\}}$ uniquely determine a homomorphism $\alpha : G_{\{i\}} *_{G_\varnothing} G_{\{j\}} \to G_{\{i,j\}}$. If $\alpha$ is not injective, let $\hat{m}$ denote the minimal length of a non-trivial element in its kernel (in the usual length function on the amalgamated free product). Recall that the homomorphisms $\varphi_{\{i\}\{i,j\}}$ and $\varphi_{\{j\}\{i,j\}}$ are injective, whence the minimal length $\hat{m}$ must be even. The *Gersten-Stallings angle* $\sphericalangle_{\{i,j\}}$ is now defined by:

$$\sphericalangle_{\{i,j\}} = \begin{cases} 2\pi/\hat{m} & \text{if } \alpha \text{ is not injective} \\ 0 & \text{if } \alpha \text{ is injective} \end{cases}$$

Three pairwise distinct elements $i, j, k \in I$ are called a *spherical triple* if $\sphericalangle_{\{i,j\}} + \sphericalangle_{\{i,k\}} + \sphericalangle_{\{j,k\}}$ is strictly larger than $\pi$. The Corson diagram is called *spherical* if it has a spherical triple, and *non-spherical* otherwise.

Consider a non-spherical triangle of groups. Since $|I| = 3$, there is only one set of three pairwise distinct elements. Depending on whether $\sphericalangle_{\{1,2\}} + \sphericalangle_{\{1,3\}} + \sphericalangle_{\{2,3\}}$ is strictly smaller than $\pi$ or equal to $\pi$, the triangle of groups is called *hyperbolic* or *Euclidean*. This distinction is of relevance in Section 4.

### 2.3 Embedding theorems

We are now able to state the theorem about non-spherical triangles of groups that has been mentioned above.

**Theorem 2.1 (Gersten-Stallings)** *For every non-spherical triangle of groups and every subset $J \subseteq I$ with $|J| \leq 2$ the natural homomorphism $\nu_J : G_J \to \mathfrak{G}$ is injective.*

This theorem has been proved in [Sta91]. Later, it has been generalised to non-spherical Corson diagrams in [Cor96]. Even more has been shown in [Cor96]. Not only the groups $G_J$ but also the colimit groups of subdiagrams naturally embed into $\mathfrak{G}$. Let us clarify. Given a Corson diagram and a subset $K \subseteq I$ we may restrict our focus to the subdiagram spanned by the groups $G_J$ with $J \subseteq K$ and $|J| \leq 2$, see the bold vertices and arrows in Figure 1 for an example. The colimit group of such a subdiagram can be obtained by modifying ($*$) as follows:

$$\mathfrak{G}_K = \left\langle \bigsqcup_{\substack{J \subseteq K \\ |J| \leq 2}} G_J : \bigsqcup_{\substack{J \subseteq K \\ |J| \leq 2}} R_J, \bigsqcup_{\substack{J_1 \subset J_2 \subseteq K \\ |J_2| \leq 2}} \{g = \varphi_{J_1 J_2}(g) : g \in G_{J_1}\} \right\rangle$$

Analogously to $\nu_J : G_J \to \mathfrak{G}$ introduced in Section 2.1, we may now consider the natural homomorphisms $\widetilde{\nu}_K : \mathfrak{G}_K \to \mathfrak{G}$ given by $g \mapsto g$.

**Theorem 2.2 (Corson)** *For every non-spherical Corson diagram and every subset $K \subseteq I$ the natural homomorphism $\widetilde{\nu}_K : \mathfrak{G}_K \to \mathfrak{G}$ is injective.*

**Remark 2.3** *For every subset $K \subseteq I$ with $|K| \leq 2$ there is an isomorphism $\mu_K : G_K \to \mathfrak{G}_K$ given by $g \mapsto g$. Hence, the injectivity of $\widetilde{\nu}_K : \mathfrak{G}_K \to \mathfrak{G}$ implies the injectivity of $\nu_K = \widetilde{\nu}_K \circ \mu_K : G_K \to \mathfrak{G}$ and, in particular, Theorem 2.2 implies Theorem 2.1.*

**Remark 2.4** *In order to keep the notation simple, we make the following convention: Whenever we know that the homomorphisms $\widetilde{\nu}_K : \mathfrak{G}_K \to \mathfrak{G}$ are injective, e. g. in case of a non-spherical Corson diagram, we do not need to mention them any more and may tacitly interpret $\mathfrak{G}_K$ as a subgroup of $\mathfrak{G}$. In this case, the symbol $\mathfrak{G}_K$ refers to the subgroup of $\mathfrak{G}$ that is generated by the elements of the groups $G_J$ with $J \subseteq K$ and $|J| \leq 2$. Now, we can easily observe that $K_1 \subseteq K_2$ implies $\mathfrak{G}_{K_1} \subseteq \mathfrak{G}_{K_2}$.*

### 2.4 Standing assumption on the Gersten-Stallings angles

We will have to make one more assumption, which has already been indicated by Gersten and Stallings in [Sta91, p. 493, ll. 4–6] and Corson in [Cor96, p. 567, l. 15], even though Theorems 2.1 and 2.2 hold without it, also compare with [Pri92, p. 210, ll. 18–21] and [KW08, p. 58, ll. 12–16].

**Standing Assumption** *In this paper, we shall always assume, without stating it explicitly, that none of the Gersten-Stallings angles is equal to $\pi$, or, equivalently, that for every two distinct $i,j \in I$ the equation $\varphi_{\{i\}\{i,j\}}(G_{\{i\}}) \cap \varphi_{\{j\}\{i,j\}}(G_{\{j\}}) = \varphi_{\varnothing\{i,j\}}(G_{\varnothing})$ holds.*

## 3 Intersection theorem

Assume we are given a Corson diagram with the property that the homomorphisms $\widetilde{\nu}_K : \mathfrak{G}_K \to \mathfrak{G}$ are injective. One question we are interested in is whether two subgroups $\mathfrak{G}_{K_1}$ and $\mathfrak{G}_{K_2}$ intersect only along the obvious subgroup $\mathfrak{G}_{K_1 \cap K_2}$ or along some larger subgroup of $\mathfrak{G}$. In Section 3.1, we give an example of a spherical Corson diagram in which the homomorphisms $\widetilde{\nu}_K : \mathfrak{G}_K \to \mathfrak{G}$ are injective but there are $K_1, K_2 \subseteq I$ such that $\mathfrak{G}_{K_1} \cap \mathfrak{G}_{K_2} \neq \mathfrak{G}_{K_1 \cap K_2}$. Then, we recall the notion of *disc pictures* and use it to prove an *intersection theorem* showing that this can only happen in the spherical realm.

### 3.1 Example

Let us consider the following Corson diagram: $I = \{1,2,3\}$, $G_{\varnothing} = \{1\}$, $G_{\{1\}} = \langle a : - \rangle$, $G_{\{2\}} = \langle b : - \rangle$, $G_{\{3\}} = \langle c : - \rangle$, $G_{\{1,2\}} = \langle a, b : b^{-1}ab = a^2 \rangle$, $G_{\{1,3\}} = \langle a, c : c^{-1}ac = a^2 \rangle$, and $G_{\{2,3\}} = \langle b, c : bc = cb \rangle$. Here, the homomorphisms $\varphi_{J_1 J_2} : G_{J_1} \to G_{J_2}$ are implicitly given by $a \mapsto a$, $b \mapsto b$, and $c \mapsto c$. Since $G_{\varnothing}$ is trivial, the resulting diagram is commutative. Britton's Lemma [Bri63, Principal Lemma] shows that the homomorphisms $\varphi_{J_1 J_2} : G_{J_1} \to G_{J_2}$ are injective and the Gersten-Stallings angles amount to $\pi/2$ each. So, it is a spherical Corson diagram in the sense of Section 2.2.

**Proposition 3.1** *The natural homomorphisms $\widetilde{\nu}_K : \mathfrak{G}_K \to \mathfrak{G}$ are injective.*

*Proof.* Since the homomorphism $\widetilde{\nu}_{\{1,2,3\}} : \mathfrak{G}_{\{1,2,3\}} \to \mathfrak{G}$ is obviously injective, it suffices to verify the injectivity of the homomorphisms $\widetilde{\nu}_K : \mathfrak{G}_K \to \mathfrak{G}$ with $K \subseteq \{1,2,3\}$ and $|K| \leq 2$. But then, we already know from Remark 2.3 that there are isomorphisms $\mu_K : G_K \to \mathfrak{G}_K$ with $\nu_K = \widetilde{\nu}_K \circ \mu_K$. So, it suffices to verify the injectivity of the homomorphisms $\nu_K : G_K \to \mathfrak{G}$.

Recall the presentation $(*)$ of the colimit group $\mathfrak{G}$ and notice that, in our situation, it can be simplified by deleting superficial generators and relators so that we finally obtain the presentation $\mathfrak{G} = \langle a, b, c : b^{-1}ab = a^2, c^{-1}ac = a^2, bc = cb \rangle$. If $K = \{1,2\}$ or $K = \{1,3\}$, let $\mathfrak{N} := \langle\!\langle bc^{-1} \rangle\!\rangle \trianglelefteq \mathfrak{G}$ and let $\pi : \mathfrak{G} \to \mathfrak{G}/\mathfrak{N}$ be the canonical projection. It is easy to see that both $\pi \circ \nu_{\{1,2\}}$ and $\pi \circ \nu_{\{1,3\}}$ are isomorphisms and, hence, both $\nu_{\{1,2\}}$ and $\nu_{\{1,3\}}$ are injective. If $K = \{2,3\}$, let $\mathfrak{N} := \langle\!\langle a \rangle\!\rangle \trianglelefteq \mathfrak{G}$ instead and proceed analogously. Finally, if $|K| \leq 1$, then $K$ is contained in some $\widetilde{K} \subseteq \{1,2,3\}$ with $|\widetilde{K}| = 2$. By construction of the colimit, we have $\nu_K = \nu_{\widetilde{K}} \circ \varphi_{K\widetilde{K}}$. Since both $\nu_{\widetilde{K}}$ and $\varphi_{K\widetilde{K}}$ are injective, their composition is injective, too. □

**Proposition 3.2** *The equation $\mathfrak{G}_{\{1,2\}} \cap \mathfrak{G}_{\{1,3\}} = \mathfrak{G}_{\{1\}}$ does not hold.*

*Proof.* Use the isomorphism $\mu_{\{1,2\}} : G_{\{1,2\}} \to \mathfrak{G}_{\{1,2\}}$ and Britton's Lemma to show that the word $bab^{-1}$ represents an element in $\mathfrak{G}_{\{1,2\}}$ that is not in $\mathfrak{G}_{\{1\}}$. On the other hand, in the colimit group $\mathfrak{G}$, the equations $bab^{-1} = bca^2c^{-1}b^{-1} = cba^2b^{-1}c^{-1} = cac^{-1}$ hold. So, the words $bab^{-1}$ and $cac^{-1}$ represent

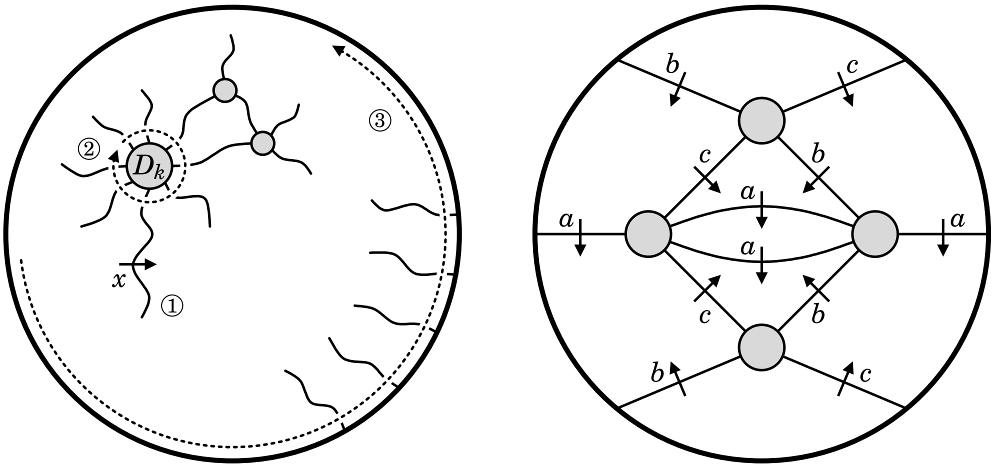


Figure 2: Some elements of a disc picture (left) and an example showing that $bab^{-1} = cac^{-1}$ holds in $G = \langle a, b, c : b^{-1}ab = a^2, c^{-1}ac = a^2, bc = cb \rangle$ (right).

the same element of the colimit group $\mathfrak{G}$, which is in $\mathfrak{G}_{\{1,2\}} \cap \mathfrak{G}_{\{1,3\}}$ but not in $\mathfrak{G}_{\{1\}}$. This calculation is also illustrated in Figure 2, in terms of disc pictures. □

### 3.2 Preliminaries about disc pictures

The proof of the intersection theorem involves *disc pictures*. Let us therefore follow Corson's preliminary section, see [Cor96], and recall some basic notions.

Consider a group $G$ and a presentation $G = \langle X : R \rangle$. A disc picture $\mathscr{P}$ over this presentation consists of the disjoint union of closed discs $D_1, D_2, \ldots, D_n$ in the interior of a closed disc $D$ and a compact 1-manifold $M$ properly embedded into $D \smallsetminus \mathrm{int}(D_1 \cup D_2 \cup \cdots \cup D_n)$. The closed discs $D_1, D_2, \ldots, D_n$ are called *vertices*, the components of $M$ are called *arcs*. Moreover, the components of $\mathrm{int}(D) \smallsetminus (D_1 \cup D_2 \cup \cdots \cup D_n \cup M)$ are called *regions*. Every arc has a transversal orientation and is labelled by a generator, see ① in Figure 2. Every vertex $D_k$ has the property that one can read off a relator *along* its boundary $\partial D_k$, i. e. by starting at some point on $\partial D_k \smallsetminus M$ and going once around $\partial D_k$ in some orientation, see ② in Figure 2. Every word that can be read off along the outer boundary $\partial D$ is called a *boundary word* of the disc picture, see ③ in Figure 2. It is well known, and easy to verify, that a word over the generators and their formal inverses represents the identity of the group $G$ if and only if it is a boundary word of some disc picture over the presentation $G = \langle X : R \rangle$. Disc pictures are, roughly speaking, duals of van Kampen diagrams. For further reading about them we refer to [BP93]. In addition to the above, the following notions will be of relevance for us.

**Definition 3.3 ("subpicture")** *Consider a closed disc $D_{\mathscr{Q}}$ in $D$. If the parts of the disc picture $\mathscr{P}$ that are contained in $D_{\mathscr{Q}}$ assemble to a disc picture $\mathscr{Q}$, we call $\mathscr{Q}$ a subpicture of $\mathscr{P}$. Notice that every boundary word of a subpicture does necessarily represent the identity of the group G. A simple kind of subpicture is a spider. It consists of exactly one vertex $D_k$ and some arcs, each of which connects $D_k$ to the outer boundary $\partial D_{\mathscr{Q}}$ of the subpicture $\mathscr{Q}$.*

Since we are interested in Corson diagrams and their colimit groups, we will focus on disc pictures over (∗). Here, it makes sense to distinguish between local and joining vertices.

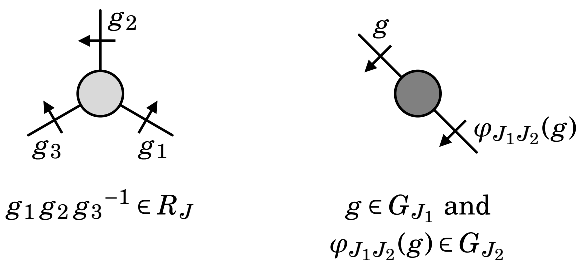


Figure 3: A local vertex (bright) and a joining vertex (dark).

**Definition 3.4 ("local and joining vertices")** *A vertex $D_k$ is called local if one can read off a relator of the form ${g_1}^{\varepsilon_1}{g_2}^{\varepsilon_2}\cdots{g_m}^{\varepsilon_m} \in R_J$ along its boundary. Otherwise, it is called joining, in which case one can read off a relator of the form $g = \varphi_{J_1J_2}(g)$ with $g \in G_{J_1}$ and $\varphi_{J_1J_2}(g) \in G_{J_2}$, see Figure 3.*

**Definition 3.5 ("bridge")** *Let $\mathscr{B}$ be the union of the compact 1-manifold $M$ and the joining vertices. The components of $\mathscr{B}$ are called bridges. Every simply connected bridge has two distinct endpoints, each of which lies either on the boundary of some local vertex or on the outer boundary. Two local vertices, say $D_k$ and $D_l$, are called neighbours if there is a bridge that connects $D_k$ and $D_l$, i. e. a bridge with one endpoint on the boundary $\partial D_k$ and the other endpoint on the boundary $\partial D_l$.*

**Definition 3.6 ("inner and outer")** *A bridge is called inner if it connects two local vertices. Similarly, a region is called* inner *if its closure does not meet the outer boundary $\partial D$. A bridge or a region that is not inner is called outer.*

Let us consider a non-spherical Corson diagram. As stated in Remark 2.4, we may interpret $\mathfrak{G}_K$ as a subgroup of $\mathfrak{G}$. The following lemma uses this interpretation to describe the labels of the arcs of a bridge.

**Lemma 3.7** *Consider a bridge with $m$ arcs labelled by generators $b_1 \in G_{J_1}$, $b_2 \in G_{J_2}, \dots, b_m \in G_{J_m}$. Then, all these generators represent the same element of the colimit group $\mathfrak{G}$. This element, say $b \in \mathfrak{G}$, is called the value of the bridge. It is contained in the subgroup $\mathfrak{G}_{J_1\cap J_2\cap\cdots\cap J_m}$.*

*Proof.* The first assertion is immediate. So, we only need to verify that the value of the bridge is actually contained in $\mathfrak{G}_{J_1\cap J_2\cap\cdots\cap J_m}$. Let us make two observations. First, if one of the sets $J_1, J_2, \dots, J_m$ is empty, say $J_k = \varnothing$, then the value of the bridge can be represented by $b_k \in G_\varnothing$. So, $b \in \mathfrak{G}_\varnothing$. This, of course, can be written as $b \in \mathfrak{G}_{J_1\cap J_2\cap\cdots\cap J_m}$, whence we are done. Therefore, we may assume w. l. o. g. that none of the sets $J_1, J_2, \dots, J_m$ is empty, in which case they must alternately have cardinality 1 and 2. Second, if $m = 1$, there is nothing to show. So, we may assume w. l. o. g. that $m \geq 2$. But then, there must be at least one set of cardinality 1 among $J_1, J_2, \dots, J_m$.

1. If all the sets of cardinality 1 are equal, say equal to $\{i\}$, then $b \in \mathfrak{G}_{\{i\}}$. But, in this case, all sets of cardinality 2 must contain $i$, which implies that $J_1 \cap J_2 \cap \cdots \cap J_m = \{i\}$. Therefore, $b \in \mathfrak{G}_{\{i\}}$ can be written as $b \in \mathfrak{G}_{J_1\cap J_2\cap\cdots\cap J_m}$.

2. If there are two distinct sets of cardinality 1 among $J_1, J_2, \dots, J_m$, say $\{i\}$ and $\{j\}$, then $b \in \mathfrak{G}_{\{i\}} \cap \mathfrak{G}_{\{j\}}$. We claim that $\mathfrak{G}_{\{i\}} \cap \mathfrak{G}_{\{j\}} = \mathfrak{G}_\varnothing$. Once this has been shown, we know that $b \in \mathfrak{G}_\varnothing$. And, again,

since $J_1 \cap J_2 \cap \cdots \cap J_m = \varnothing$, this can be written as $b \in \mathfrak{G}_{J_1 \cap J_2 \cap \cdots \cap J_m}$. So, it remains to show that $\mathfrak{G}_{\{i\}} \cap \mathfrak{G}_{\{j\}} = \mathfrak{G}_\varnothing$. By our standing assumption, the equation $\varphi_{\{i\}\{i,j\}}(G_{\{i\}}) \cap \varphi_{\{j\}\{i,j\}}(G_{\{j\}}) = \varphi_{\varnothing\{i,j\}}(G_\varnothing)$ holds in $G_{\{i,j\}}$. In order to transport this equation to the colimit group $\mathfrak{G}$, we apply the injective homomorphism $\nu_{\{i,j\}} : G_{\{i,j\}} \to \mathfrak{G}$. This yields:

$$\nu_{\{i,j\}}\left(\varphi_{\{i\}\{i,j\}}\left(G_{\{i\}}\right)\right) \cap \nu_{\{i,j\}}\left(\varphi_{\{j\}\{i,j\}}\left(G_{\{j\}}\right)\right) = \nu_{\{i,j\}}\left(\varphi_{\varnothing\{i,j\}}\left(G_\varnothing\right)\right)$$

Using the equations $\nu_{\{i,j\}}(\varphi_{K\{i,j\}}(G_K)) = \nu_K(G_K) = \widetilde{\nu}_K(\mu_K(G_K)) = \widetilde{\nu}_K(\mathfrak{G}_K)$ for the subsets $K \subset \{i,j\}$, we finally obtain $\widetilde{\nu}_{\{i\}}(\mathfrak{G}_{\{i\}}) \cap \widetilde{\nu}_{\{j\}}(\mathfrak{G}_{\{j\}}) = \widetilde{\nu}_\varnothing(\mathfrak{G}_\varnothing)$, which reads as $\mathfrak{G}_{\{i\}} \cap \mathfrak{G}_{\{j\}} = \mathfrak{G}_\varnothing$ in the shorthand notation of Remark 2.4.

□

### 3.3 Statement and proof of the intersection theorem

We are now ready to discuss the intersection theorem. The proof is based on ideas and techniques that go back to Gersten and Stallings in [Sta91] and Corson in [Cor96], but we have to be more careful when counting weights.

**Theorem 3.8** *For every non-spherical Corson diagram and every two subsets $K_1, K_2 \subseteq I$ the equation $\mathfrak{G}_{K_1} \cap \mathfrak{G}_{K_2} = \mathfrak{G}_{K_1 \cap K_2}$ holds.*

*Proof.* The inclusion “$\supseteq$” is a consequence of Remark 2.4. So, we only need to verify the inclusion “$\subseteq$”. Suppose there were a non-spherical Corson diagram and subsets $K_1, K_2 \subseteq I$ with $\mathfrak{G}_{K_1} \cap \mathfrak{G}_{K_2} \nsubseteq \mathfrak{G}_{K_1 \cap K_2}$. Then, we can find an element $g \in \mathfrak{G}$ with $g \in \mathfrak{G}_{K_1} \cap \mathfrak{G}_{K_2}$ but $g \notin \mathfrak{G}_{K_1 \cap K_2}$. Being contained in $\mathfrak{G}_{K_1}$, it can be represented by a word $w_1$ over the generators from the groups $G_J$ with $J \subseteq K_1$ and $|J| \leq 2$, and their formal inverses. On the other hand, being contained in $\mathfrak{G}_{K_2}$, it can also be represented by a word $w_2$ over the generators from the groups $G_J$ with $J \subseteq K_2$ and $|J| \leq 2$, and their formal inverses.

Since $w_1$ and $w_2$ represent the same element of the colimit group $\mathfrak{G}$, there is a disc picture $\mathscr{P}$ over ($*$) with boundary word $w_1 {w_2}^{-1}$. By construction, $g \notin \mathfrak{G}_{K_1 \cap K_2}$. So, it cannot be the identity of the colimit group $\mathfrak{G}$. Therefore, the words $w_1$ and $w_2$ cannot be empty and there are at least two arcs, or one arc twice, incident with the outer boundary $\partial D$.

We may assume w. l. o. g. that the element $g$, the words $w_1$ and $w_2$, and the disc picture $\mathscr{P}$ are chosen in such a way that the *complexity* of the disc picture is minimal, i. e. the number of local vertices is minimal and, among all disc pictures with this minimal number of local vertices, the number of bridges is minimal. This assumption has many consequences on the structure of the disc picture.

(1) *The disc picture $\mathscr{P}$ is connected. In particular, since there are arcs incident with the outer boundary $\partial D$, every local vertex is incident with at least one arc. Moreover, all bridges and regions are simply connected.* We claim that if $\mathscr{P}$ was not connected, we could remove at least one component and, hence, obtain a disc picture with strictly fewer local vertices or with the same number of local vertices but strictly fewer bridges. In other words, we could obtain a disc picture of lower complexity.

First, notice that there are two distinct points $x, y \in \partial D \setminus M$ such that one can read off the words $w_1$ and $w_2$ when going from $x$ to $y$ along the respective side of $\partial D$, see ① in Figure 4. If there is a component of $\mathscr{P}$ that is incident with at most one side of $\partial D$, we can remove it. In this case, the boundary words of the disc picture may change. But the new disc picture gives rise to new words $\widetilde{w}_1$ and $\widetilde{w}_2$. Since the removed component has been incident with at most one side of $\partial D$, at least

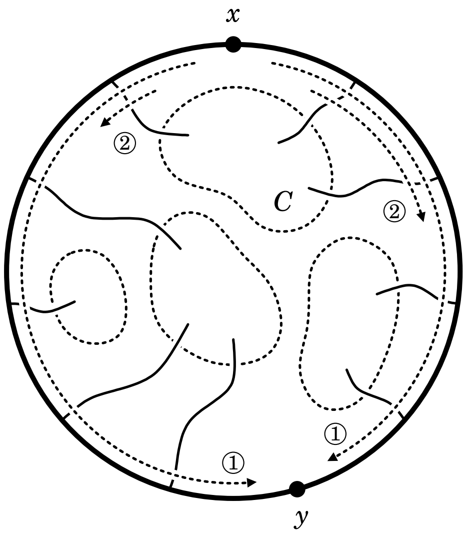


Figure 4: The disc picture $\mathscr{P}$ is connected.

one of the words $\widetilde{w}_i$ remains equal to $w_i$. So, both $\widetilde{w}_1$ and $\widetilde{w}_2$, which represent the same element of the colimit group $\mathfrak{G}$, still represent $g$. *In the following steps, as here, we may obtain new words $\widetilde{w}_1$ and $\widetilde{w}_2$, and sometimes even a new element $\widetilde{g} \in \mathfrak{G}$. But, in each step, it is easy to see that this data could have been chosen right at the beginning.*

By the above, we may assume w. l. o. g. that every component of $\mathscr{P}$ is incident with both sides of $\partial D$. Suppose there is more than one such component and let $C$ be the first one traversed when going from $x$ to $y$ along $\partial D$. For a moment, let us focus on $C$ and ignore all the other components. Now, one can read off new words $\widetilde{w}_1$ and $\widetilde{w}_2$ along the respective sides of $\partial D$ that represent a new element $\widetilde{g} \in \mathfrak{G}$, see ② in Figure 4. By construction, $\widetilde{g} \in \mathfrak{G}_{K_1} \cap \mathfrak{G}_{K_2}$. If $\widetilde{g} \notin \mathfrak{G}_{K_1 \cap K_2}$, the component $C$ is already a suitable disc picture and we can remove the other components completely. On the other hand, if $\widetilde{g} \in \mathfrak{G}_{K_1 \cap K_2}$, we can keep the other components and remove $C$. Then, the words that one can read off along the respective sides of $\partial D$ represent the element $\widetilde{g}^{-1} g \in \mathfrak{G}$, which is the product of an element in $\mathfrak{G}_{K_1 \cap K_2}$ and an element not in $\mathfrak{G}_{K_1 \cap K_2}$. Therefore, $\widetilde{g}^{-1} g \notin \mathfrak{G}_{K_1 \cap K_2}$ and, again, we end up with a suitable disc picture.

**Definition 3.9 ("type of a local vertex")** *Since every local vertex is incident with at least one arc, we can associate a type to every local vertex. More precisely, for every local vertex $D_k$ there is a unique*[1] *subset $J \subseteq I$ with $|J| \leq 2$ such that all arcs incident with $D_k$ are labelled by generators from $G_J$. In this case, we say that $D_k$ is of type $J$.*

(2) *The disc picture $\mathscr{P}$ has at least one local vertex.* Since $w_1$ and $w_2$ cannot be empty, there is at least one arc incident with each side of $\partial D$. Therefore, if $\mathscr{P}$ had no local vertex at all, it would have to be a single bridge $B$ connecting the two sides of $\partial D$. Depending on the transversal orientation of its arcs, the value of $B$ is either $g$ or $g^{-1}$. The extremal arcs of $B$ are labelled by generators, say $b_1 \in G_{J_1}$ and $b_m \in G_{J_m}$ with $J_1 \subseteq K_1$ and $J_m \subseteq K_2$. Using Lemma 3.7 and Remark 2.4, we can now observe that $g \in \mathfrak{G}_{J_1 \cap J_2 \cap \cdots \cap J_m} \subseteq \mathfrak{G}_{J_1 \cap J_m} \subseteq \mathfrak{G}_{K_1 \cap K_2}$, in contradiction to $g \notin \mathfrak{G}_{K_1 \cap K_2}$.

[1] Recall that the set of generators is the disjoint union of all $G_J$ with $J \subseteq I$ and $|J| \leq 2$.

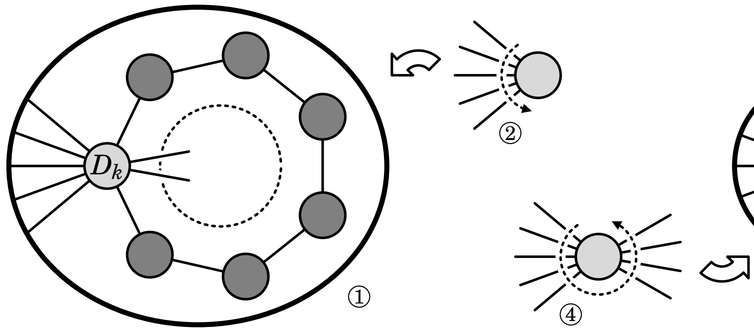


Figure 5: Replace the subpictures $\mathscr{Q}$ by spiders.

(3) *A local vertex cannot be a neighbour of itself.* If there was such a local vertex $D_k$, say of type $J$, we could consider the subpicture $\mathscr{Q}$ consisting of the local vertex $D_k$, one of the bridges that connect $D_k$ with itself, everything that is enclosed by this bridge, and the extremal parts of the remaining arcs incident with $D_k$, see ① in Figure 5. Every boundary word $w$ of the subpicture $\mathscr{Q}$ is a word over generators from $G_J$ and their formal inverses that represents the identity of the colimit group $\mathfrak{G}$. Since the natural homomorphism $\nu_J : G_J \to \mathfrak{G}$ is injective, the word $w$ does not only represent the identity of the colimit group $\mathfrak{G}$ but also the identity of the group $G_J$. Therefore, we can remove the subpicture $\mathscr{Q}$ and replace it by a single spider with boundary word $w$, see ② in Figure 5. After this modification, we obtain a disc picture with at most as many local vertices and stricly fewer bridges, and, hence, of lower complexity.

(4) *Two local vertices of the same type cannot be neighbours.* If there were two such local vertices $D_k$ and $D_l$, w. l. o. g. $D_k \neq D_l$, we could consider the subpicture $\mathscr{Q}$ consisting of the local vertices $D_k$ and $D_l$, one of the bridges that connect $D_k$ and $D_l$, and the extremal parts of the remaining arcs incident with $D_k$ and $D_l$, see ③ in Figure 5. By the same arguments as in (3), we can remove the subpicture $\mathscr{Q}$ and replace it by a single spider with the same boundary word, see ④ in Figure 5. Again, we obtain a disc picture of lower complexity.

(5) *Every bridge has at least two arcs.* If there was a bridge $B$ with only one arc, we could distinguish between three cases. First, if $B$ is connecting two local vertices, say of types $J_1$ and $J_2$, then $J_1 = J_2$, in contradiction to (4). Second, if $B$ is connecting the outer boundary $\partial D$ with itself, then, by (1), $B$ is already the whole disc picture, in contradiction to (2). So, we may assume w. l. o. g. that $B$ is connecting a local vertex $D_k$, say of type $J$, and the outer boundary $\partial D$, say at the side of $\partial D$ along which one can read off the word $w_1$.

By the former, $B$ is labelled by some generator from $G_J$ and, by the latter, $J \subseteq K_1$. Now, consider the subpicture $\mathscr{Q}$ consisting of the local vertex $D_k$, the bridge $B$, and the extremal parts of the remaining arcs incident with $D_k$, see ① in Figure 6. Replace it by a subpicture in which the arcs traversing $\partial D_{\mathscr{Q}}$, which are all labelled by generators from $G_J$, are extended to the outer boundary $\partial D$, see ② in Figure 6. This gives a disc picture with one fewer local vertex and, hence, of lower complexity.

(6) *The two regions on either side of a bridge cannot be the same.* Suppose there was a bridge $B$ having the same region $R$ on either side. Then, we can find a subpicture $\mathscr{Q}$ whose boundary $\partial D_{\mathscr{Q}}$

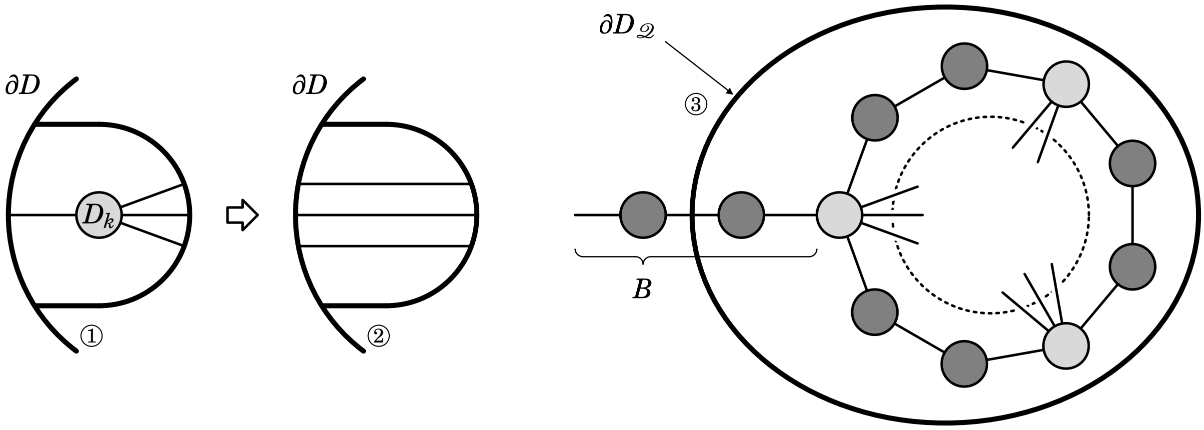


Figure 6: Replace the subpicture $\mathscr{Q}$ by some arcs (left) and remove the bridge $B$ (right).

is contained in $R$, except for one point where it crosses $B$, see ③ in Figure 6. Therefore, the value of $B$ is the identity of the colimit group $\mathfrak{G}$. Since the natural homomorphisms $\nu_J : G_J \to \mathfrak{G}$ are injective, the labels of the arcs of $B$ must also be the identities of the respective groups $G_J$. So, we can remove the bridge $B$ and obtain a disc picture of lower complexity.

(7) *The value of a bridge cannot be an element of $\mathfrak{G}_\varnothing$.* Suppose there was a bridge $B$ with value $b \in \mathfrak{G}_\varnothing$. Since $b \in \mathfrak{G}_\varnothing$, there is a generator $b_\varnothing \in G_\varnothing$ that represents it. In fact, for every $J \subseteq I$ with $1 \leq |J| \leq 2$ there is a generator $b_J \in G_J$ that represents it, namely $b_J := \varphi_{\varnothing J}(b_\varnothing) \in G_J$.

By (6), the two regions on either side of $B$ cannot be the same. Among these two distinct regions choose the region $R$ with the property that the arcs of $B$ are heading away from $R$. Now, remove $B$ and relabel all the remaining arcs in the boundary $\partial R$ as follows: If an arc is labelled by a generator $a \in G_J$ and is heading towards $R$, relabel it by $a \cdot b_J \in G_J$. If it is heading away from $R$, relabel it by ${b_J}^{-1} \cdot a \in G_J$. This guarantees that one can still read off relators along the boundaries of the remaining vertices. Here, we leave the details to the reader, see [Cor96, Appendix] for another description and ①–③ in Figure 7 for some examples. Notice that, in ②, the generator $\varphi_{J_1 J_2}(a) \cdot b_{J_2} \in G_{J_2}$ satisfies the following equation:

$$\begin{aligned}
\varphi_{J_1 J_2}(a) \cdot b_{J_2} &= \varphi_{J_1 J_2}(a) \cdot \varphi_{\varnothing J_2}(b_\varnothing) \\
&= \begin{cases} \varphi_{J_1 J_2}(a) \cdot \varphi_{J_1 J_2}(b_{J_1}) & \text{if } J_1 = \varnothing \\ \varphi_{J_1 J_2}(a) \cdot \varphi_{J_1 J_2}(\varphi_{\varnothing J_1}(b_\varnothing)) & \text{otherwise} \end{cases} \\
&= \varphi_{J_1 J_2}(a) \cdot \varphi_{J_1 J_2}(b_{J_1}) \\
&= \varphi_{J_1 J_2}(a \cdot b_{J_1})
\end{aligned}$$

Therefore, one can actually read off the relator $a \cdot b_{J_1} = \varphi_{J_1 J_2}(a \cdot b_{J_1})$ along the boundary of the respective joining vertex. So, we obtain a disc picture with the same number of local vertices and strictly fewer bridges and, hence, of lower complexity.

(8) *For every bridge there is a unique element $i \in I$ such that the value of the bridge is in $\mathfrak{G}_{\{i\}} \setminus \mathfrak{G}_\varnothing$.* Let $B$ be a bridge with $m$ arcs that are labelled by generators $b_1 \in G_{J_1}, b_2 \in G_{J_2}, \ldots, b_m \in G_{J_m}$. By (7), none of the sets $J_1, J_2, \ldots, J_m$ is empty, which implies that they must alternately have cardinality

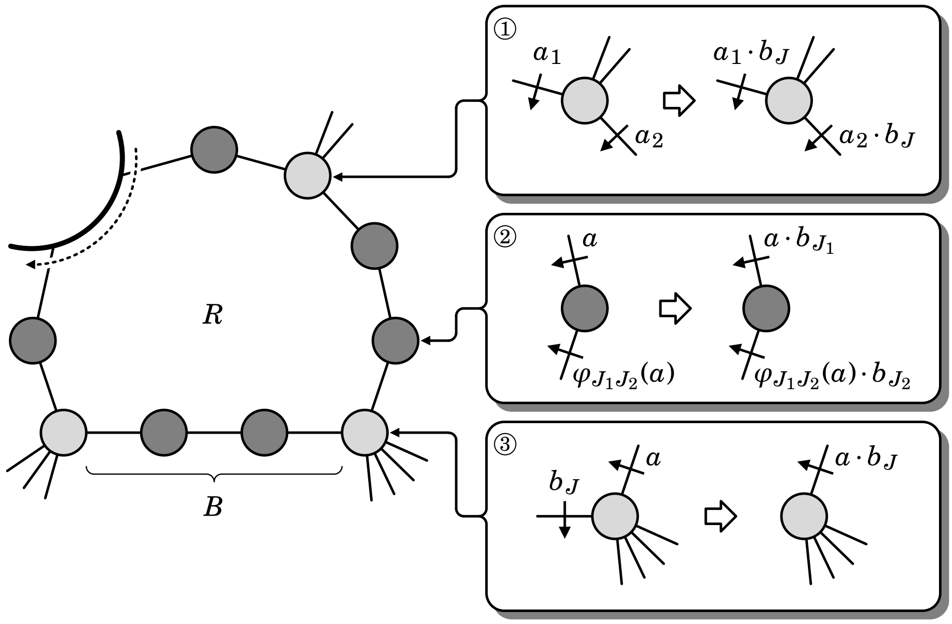


Figure 7: Relabel the remaining arcs in the boundary $\partial R$.

1 and 2. By (5), every bridge has at least two arcs, i. e. $m \geq 2$. Therefore, there must be at least one set $J_k$ of cardinality 1, say $J_k = \{i\}$. So, the value of $B$ is an element of $\mathfrak{G}_{\{i\}}$ and, again by (7), cannot be an element of $\mathfrak{G}_\varnothing$. The element $i \in I$ is unique; in the proof of Lemma 3.7, we have seen that for any two distinct $i, j \in I$ the equation $\mathfrak{G}_{\{i\}} \cap \mathfrak{G}_{\{j\}} = \mathfrak{G}_\varnothing$ holds, whence $B$ cannot have a value that is simultaneously in $\mathfrak{G}_{\{i\}} \setminus \mathfrak{G}_\varnothing$ and $\mathfrak{G}_{\{j\}} \setminus \mathfrak{G}_\varnothing$.

**Definition 3.10 ("type of a bridge")** *In this case, we say that the bridge $B$ is of type $i$.*

(9) *If a bridge of type $i$ is incident with a local vertex of type $J$, then $i \in J$. Similarly, if it is incident with one side of the outer boundary $\partial D$, then $i \in K_1$ or $i \in K_2$ respectively.* We give a proof of the first assertion, the proof of the second one is essentially the same. Let $B$ be a bridge of type $i$ that is incident with a local vertex $D_k$ of type $J$. As we have seen in (8), one of the arcs of $B$ is labelled by a generator from $G_{\{i\}}$. Moreover, the extremal arc of $B$ that is incident with $D_k$ is labelled by a generator from $G_J$. By Lemma 3.7 and Remark 2.4, we can conclude that $B$ has a value in $\mathfrak{G}_{\{i\} \cap J}$. By (7), this value is not in $\mathfrak{G}_\varnothing$. Therefore, $\{i\} \cap J \neq \varnothing$, whence $i \in J$.

(10) *There are no local vertices of type $\varnothing$.* By (1), every local vertex is incident with at least one arc. So, if there was a local vertex of type $\varnothing$, it would have to be incident with an arc that is labelled by some generator $a \in G_\varnothing$. But this arc is part of a bridge with value in $\mathfrak{G}_\varnothing$, in contradiction to (7).

(11) *There are no local vertices of type $\{i\}$ with $i \in I$.* First, observe that if there was such a local vertex $D_k$, it would have to be a neighbour of some other local vertex $D_l$. Suppose it wasn't. Then, all bridges that are incident with $D_k$ must either connect it to itself, which is not possible by (3), or to the outer boundary $\partial D$. But, by (1), the disc picture $\mathscr{P}$ is connected. So, this is already the whole disc picture. In particular, all bridges are incident with $D_k$, which is a local vertex of type $\{i\}$. Therefore, by (9), all bridges are of type $i$. But this means that each letter of $w_1$ and $w_2$

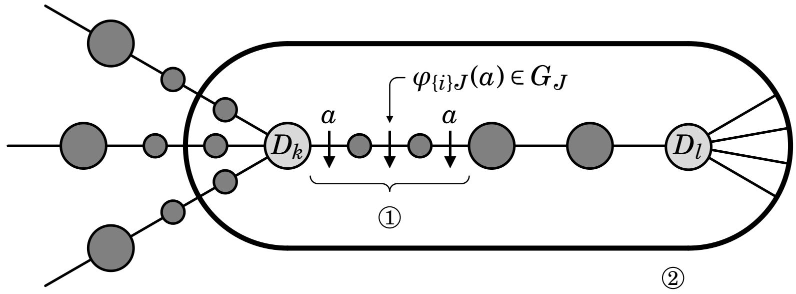


Figure 8: Replace each arc by a sequence of three arcs.

represents an element in $\mathfrak{G}_{\{i\}}$, whence $g \in \mathfrak{G}_{\{i\}}$. On the other hand, since both $w_1$ and $w_2$ are not empty, there is at least one bridge connecting $D_k$ to either side of $\partial D$. Again, by (9), this implies both $i \in K_1$ and $i \in K_2$. But since $\{i\} \subseteq K_1 \cap K_2$, we can use Remark 2.4 to conclude that $g \in \mathfrak{G}_{K_1 \cap K_2}$, in contradiction to $g \notin \mathfrak{G}_{K_1 \cap K_2}$.

So, we may assume w. l. o. g. that $D_k$ is a neighbour of some other local vertex $D_l$. Consider a bridge that connects $D_k$ and $D_l$. Again, by (9), this bridge is of type $i$ and the local vertex $D_l$ is of some type $J$ with $i \in J$, i. e. $\{i\} \subseteq J$. By (4), two local vertices of the same type cannot be neighbours. So, we actually obtain that $\{i\} \subset J$. Now, replace every arc that is incident with $D_k$, say labelled by some generator $a \in G_{\{i\}}$, by a sequence of three arcs with the same transversal orientation. The first and the third are labelled by $a \in G_{\{i\}}$, the second by $\varphi_{\{i\}J}(a) \in G_J$, see ① in Figure 8. Then, consider the subpicture $\mathscr{Q}$ indicated in ② in Figure 8. By the same arguments as in (3) and (4), we can remove the subpicture $\mathscr{Q}$ and replace it by a single spider with the same boundary word. Again, we obtain a disc picture of lower complexity.

**Definition 3.11 ("angle and swap")** *By (10) and (11), we know that every local vertex $D_k$ is of some type $\{i,j\}$ with distinct $i,j \in I$. By (9), such a local vertex is incident with bridges each of which is either of type $i$ or of type $j$. Consider the connected components of $\partial D_k \smallsetminus M$. They are called angles. An angle is called a swap if one of the two bridges enclosing it is of type $i$ and the other one is of type $j$.*

(12) *Every local vertex has at least one swap in its boundary.* Suppose there was a local vertex $D_k$, say as above of type $\{i,j\}$, without any swap in its boundary. Then, all bridges that are incident with $D_k$ are of the same type, say of type $i$.

Since $D_k$ is a local vertex of type $\{i,j\}$, the arcs that are incident with $D_k$ are labelled by generators $a_1, a_2, \ldots, a_m \in G_{\{i,j\}}$. The respective bridges are all of type $i$. So, each of the generators represents an element in $\mathfrak{G}_{\{i\}} \smallsetminus \mathfrak{G}_{\varnothing}$, whence we can even find generators $\widetilde{a}_1, \widetilde{a}_2, \ldots, \widetilde{a}_m \in G_{\{i\}}$ representing the same elements, i. e. satisfying the equations $\nu_{\{i,j\}}(a_s) = \nu_{\{i\}}(\widetilde{a}_s) = \nu_{\{i,j\}}(\varphi_{\{i\}\{i,j\}}(\widetilde{a}_s))$. Now, we can use the injectivity of the homomorphism $\nu_{\{i,j\}} : G_{\{i,j\}} \to \mathfrak{G}$ to conclude that $a_s = \varphi_{\{i\}\{i,j\}}(\widetilde{a}_s)$. Similarly to the modification described above in (11), we replace every arc that is incident with $D_k$ by a sequence of two arcs with the same transversal orientation. If the arc has been labelled by $a_s \in G_{\{i,j\}}$, the new arc that is incident with $D_k$ is labelled by $\widetilde{a}_s \in G_{\{i\}}$ whereas the other one is labelled by $a_s \in G_{\{i,j\}}$, see ① in Figure 9. Now, let $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_m \in \{-1,1\}$ such that the word ${a_1}^{\varepsilon_1}{a_2}^{\varepsilon_2} \cdots {a_m}^{\varepsilon_m}$ has originally been a boundary word of $D_k$. After this modification, one can read

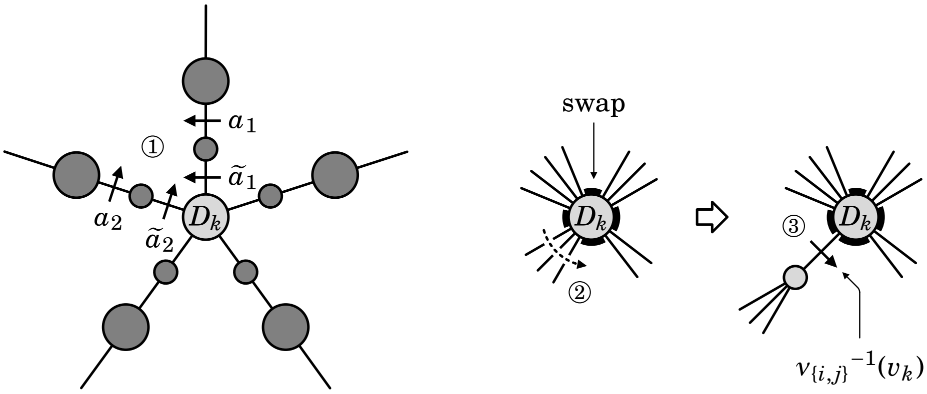


Figure 9: Replace each arc by a sequence of two arcs (left) and create a new local vertex (right).

off the word $\tilde{a}_1{}^{\varepsilon_1}\tilde{a}_2{}^{\varepsilon_2}\cdots\tilde{a}_m{}^{\varepsilon_m}$ along $\partial D_k$. But:

$$\varphi_{\{i\}\{i,j\}}\left(\tilde{a}_1{}^{\varepsilon_1}\tilde{a}_2{}^{\varepsilon_2}\cdots\tilde{a}_m{}^{\varepsilon_m}\right) = a_1{}^{\varepsilon_1}a_2{}^{\varepsilon_2}\cdots a_m{}^{\varepsilon_m} = 1 \text{ in } G_{\{i,j\}}$$

The injectivity of $\varphi_{\{i\}\{i,j\}} : G_{\{i\}} \to G_{\{i,j\}}$ implies that $D_k$ has become a local vertex of type $\{i\}$ and can be removed as in (11). We obtain a disc picture of lower complexity.

**Definition 3.12 ("sufficiently many swaps")** *Assume we are given a local vertex $D_k$ of some type $\{i,j\}$ with distinct $i,j \in I$. Then, the local vertex $D_k$ has sufficiently many swaps in its boundary if the Gersten-Stallings angle $\sphericalangle_{\{i,j\}} \neq 0$ and the number of swaps $m \geq 2\pi/\sphericalangle_{\{i,j\}}$.*

(13) *Every local vertex has sufficiently many swaps in its boundary.* Suppose there was a local vertex $D_k$ without sufficiently many swaps in its boundary. By (12), there is at least one swap. If we start at some swap and go from swap to swap once around $\partial D_k$, then we can read off words $v_1, v_2, \ldots, v_m$, see ② in Figure 9. Each of these words represents an element in $\mathfrak{G}$, so it makes sense to write $v_1, v_2, \ldots, v_m \in \mathfrak{G}$. Their product $v_1 \cdot v_2 \cdot \ldots \cdot v_m$ is the identity element. We may assume w. l. o. g. that $v_1, v_3, \ldots, v_{m-1} \in \mathfrak{G}_{\{i\}}$ and $v_2, v_4, \ldots, v_m \in \mathfrak{G}_{\{j\}}$. In order to show that at least one of these elements is contained in $\mathfrak{G}_\varnothing$, we construct their preimages under the injective homomorphisms $\nu_K : G_K \to \mathfrak{G}$:

$$\begin{aligned}
&\nu_{\{i\}}{}^{-1}(v_1), \nu_{\{i\}}{}^{-1}(v_3), \ldots, \nu_{\{i\}}{}^{-1}(v_{m-1}) \in G_{\{i\}} \\
&\nu_{\{j\}}{}^{-1}(v_2), \nu_{\{j\}}{}^{-1}(v_4), \ldots, \nu_{\{j\}}{}^{-1}(v_m) \in G_{\{j\}}
\end{aligned}$$

The preimages assemble to an element $x := \nu_{\{i\}}{}^{-1}(v_1) \cdot \nu_{\{j\}}{}^{-1}(v_2) \cdot \ldots \cdot \nu_{\{j\}}{}^{-1}(v_m)$ of the amalgamated free product $G_{\{i\}} *_{G_\varnothing} G_{\{j\}}$. Now, recall the definition of the Gersten-Stallings angle from Section 2.2. The homomorphism $\alpha : G_{\{i\}} *_{G_\varnothing} G_{\{j\}} \to G_{\{i,j\}}$ introduced there satisfies:

$$\begin{aligned}
\alpha(x) &= \varphi_{\{i\}\{i,j\}}\left(\nu_{\{i\}}{}^{-1}(v_1)\right) \cdot \varphi_{\{j\}\{i,j\}}\left(\nu_{\{j\}}{}^{-1}(v_2)\right) \cdot \ldots \cdot \varphi_{\{j\}\{i,j\}}\left(\nu_{\{j\}}{}^{-1}(v_m)\right) \\
&= \nu_{\{i,j\}}{}^{-1}(v_1) \cdot \nu_{\{i,j\}}{}^{-1}(v_2) \cdot \ldots \cdot \nu_{\{i,j\}}{}^{-1}(v_m) \\
&= \nu_{\{i,j\}}{}^{-1}(v_1 \cdot v_2 \cdot \ldots \cdot v_m) \\
&= \nu_{\{i,j\}}{}^{-1}(1) \\
&= 1
\end{aligned}$$

So, $x \in \ker(\alpha)$. Since $D_k$ does not have sufficiently many swaps in its boundary, we know that either the Gersten-Stallings angle $\sphericalangle_{\{i,j\}} = 0$ or the length of $x$, which is at most $m$, is strictly smaller than $2\pi/\sphericalangle_{\{i,j\}}$, which is nothing but the minimal length of a non-trivial element in $\ker(\alpha)$. In either case, $x$ must be trivial in $G_{\{i\}} *_{G_\varnothing} G_{\{j\}}$.

It is a consequence of the normal form theorem, see [Mil68, Lemma 1], that there is an index $k \in \{1,2,\dots,m\}$ such that $\nu_{\{i\}}{}^{-1}(v_k) \in \varphi_{\varnothing\{i\}}(G_\varnothing)$ or $\nu_{\{j\}}{}^{-1}(v_k) \in \varphi_{\varnothing\{j\}}(G_\varnothing)$, depending on the parity of $k$. But then:

$$v_k \in \left\{ \begin{array}{ll} \nu_{\{i\}}\left(\varphi_{\varnothing\{i\}}(G_\varnothing)\right) & \text{if } k \text{ is odd} \\ \nu_{\{j\}}\left(\varphi_{\varnothing\{j\}}(G_\varnothing)\right) & \text{if } k \text{ is even} \end{array} \right\} = \nu_\varnothing(G_\varnothing) = \mathfrak{G}_\varnothing$$

In either case, $v_k \in \mathfrak{G}_\varnothing$. Now, we add a new local vertex and a new bridge to the disc picture as illustrated in ③ in Figure 9. The arcs that had been traversed when reading off the word $v_k$ end up at the new local vertex, which is connected to $D_k$ by a single arc labelled by $\nu_{\{i,j\}}{}^{-1}(v_k)$. This increases both the number of local vertices and the number of bridges by 1. But still, some of the properties we have discussed so far, in particular (1), hold true and the bridge connecting the new local vertex and $D_k$, which has a value in $\mathfrak{G}_\varnothing$, can be removed as in (6) or (7). Next, we want to get rid of the new local vertex. If the removal of the bridge has made the disc picture $\mathscr{P}$ disconnected, we can remove one of the two components as in (1). Otherwise, there is still a path from the new local vertex to $D_k$, which implies that the new local vertex is a neighbour of some other local vertex. Once this is clear, we can remove it as in (12) and in the final step of (11). In either case, in particular in the latter, we do not only remove the new local vertex but also at least one more bridge. So, in either case, we obtain a disc picture of lower complexity.

After all these observations, we can give an easy combinatorial argument that yields a contradiction. The principal idea is to distribute weights over certain parts of the disc picture. Every local vertex gets the weight $2\pi$, every inner bridge gets the weight $-2\pi$, and every inner region gets the weight $2\pi$. For the notion of *inner* see Definition 3.6. The weighted parts of the disc picture correspond to vertices, edges, and bounded regions of a planar graph, which is non-empty, finite, and connected. So, we may use Euler's formula for planar graphs to calculate the total weight:

$$2\pi \cdot \#\text{local vertices} - 2\pi \cdot \#\text{inner bridges} + 2\pi \cdot \#\text{inner regions} = 2\pi$$

Let us count again. But, this time, we reallocate the weights to the regions. Every inner bridge distributes its weight $-2\pi$ equally to the two regions on either side and every local vertex distributes its weight $2\pi$ equally to the swaps in its boundary, each of which lets it traverse to the adjacent region. The new total weight of a region $R$ is denoted by $\mathrm{wt}(R)$ and can be estimated from above using (14), (15), and (16):

(14) *The weight traversing each swap is bounded by the Gersten-Stallings angle associated to the type of the local vertex. In particular, it is at most* $\pi/2$. Let $D_k$ be a local vertex of type $J = \{i,j\}$. By (13), $D_k$ has sufficiently many swaps in its boundary. So, the Gersten-Stallings angle $\sphericalangle_{\{i,j\}} \neq 0$ and the number of swaps is at least $2\pi/\sphericalangle_{\{i,j\}}$. Since $D_k$ distributes its weight $2\pi$ equally to the swaps, the weight traversing each swap is at most $\sphericalangle_{\{i,j\}}$.

(15) *There are no inner regions of positive weight.* Let $R$ be an inner region. By (1), $R$ is an open disc. By (6), the boundary $\partial R$ contains some number of inner bridges, say $m$, and the same number of angles, some of which may be swaps. By (3), $m \geq 2$. By (6), each of the $m$ inner bridges contributes

$-\pi$ to $\mathrm{wt}(R)$ and, by (14), each of the at most $m$ swaps contributes at most $\pi/2$. Therefore, we can estimate $\mathrm{wt}(R)$ as follows:

If $m \geq 4$, then $\mathrm{wt}(R) \leq 1 \cdot 2\pi - m \cdot \pi + m \cdot \pi/2 \leq 0$. If $m = 3$ and there are at most two swaps, then $\mathrm{wt}(R) \leq 1 \cdot 2\pi - 3 \cdot \pi + 2 \cdot \pi/2 = 0$. If $m = 3$ and there are exactly three swaps, then there are three pairwise distinct $i, j, k \in I$ such that the local vertices in the boundary $\partial R$ are of types $\{i, j\}$, $\{i, k\}$, $\{j, k\}$. Since we consider a non-spherical Corson diagram, there are no spherical triples. In particular, $\sphericalangle_{\{i,j\}} + \sphericalangle_{\{i,k\}} + \sphericalangle_{\{j,k\}}$ is at most $\pi$, whence $\mathrm{wt}(R) \leq 1 \cdot 2\pi - 3 \cdot \pi + \sphericalangle_{\{i,j\}} + \sphericalangle_{\{i,k\}} + \sphericalangle_{\{j,k\}} \leq 0$. What remains is the case that $m = 2$. But if there were two distinct $i, j \in I$ such that one of the inner bridges is of type $i$ and the other is of type $j$, then both local vertices must be of type $\{i, j\}$, in contradiction to (4). So, in this case, there cannot be any swap, whence $\mathrm{wt}(R) = 1 \cdot 2\pi - 2 \cdot \pi + 0 = 0$.

(16) *There are at most two outer regions of positive weight, and each of them has at most weight $\pi/2$.* Let $R$ be an outer region. By (1), $R$ is an open disc. The boundary $\partial R$ contains some number of bridges, say $m$, and some number of angles. By (1), (2), and (6), exactly two of the bridges are outer, which implies that $m \geq 2$. Moreover, it contains exactly $m-1$ angles, some of which may be swaps.

Again, if $m \geq 3$, then $\mathrm{wt}(R) \leq 0 \cdot 2\pi - (m-2) \cdot \pi + (m-1) \cdot \pi/2 = (-m+3) \cdot \pi/2 \leq 0$. If $m = 2$ and there is no swap, then $\mathrm{wt}(R) = 0 \cdot 2\pi - 0 \cdot \pi + 0 = 0$. What remains is the case that $m = 2$ and there is a swap. Then, $\mathrm{wt}(R)$ might well be positive, but $\mathrm{wt}(R) \leq 0 \cdot 2\pi - 0 \cdot \pi + 1 \cdot \pi/2 = \pi/2$.

Next, we show that this can happen at most twice. Since there is a swap, we know that there are two distinct $i, j \in I$ such that one of the outer bridges is of type $i$ and the other one is of type $j$. If both bridges end up at the same side of $\partial D$, say at the side of $\partial D$ along which one can read off the word $w_1$, then, by (9), both $i \in K_1$ and $j \in K_1$. This allows us to remove the respective local vertex $D_k$ of type $\{i, j\} \subseteq K_1$ as in the final step of (5) and to obtain a disc picture of lower complexity. So, we may assume w. l. o. g. that the two bridges end up at different sides of $\partial D$. But, by (1), we know that the disc picture $\mathscr{P}$ is connected. Hence, this can happen at most twice, namely when the boundary $\partial R$ contains one of the two points $x$ and $y$ that have been chosen in (1).

By (15) and (16), the total weight given to the disc picture is at most $\pi$. This is a contradiction to the above observation that the total weight amounts to $2\pi$, which completes the proof. □

### 3.4 Interpretation

In case of a non-spherical triangle of groups, the intersection theorem says that the groups $\mathfrak{G}_K$ with $K \subseteq \{1, 2, 3\}$ and $|K| \leq 2$ intersect exactly as sketched in Figure 10. One particularly nice way of reading the intersection theorem is to start with such a setting. Let $M$ be the union of three groups with the property that each two of them intersect along a common subgroup, by which we implicitly mean that the two multiplications agree on the subgroup. $M$ is a set equipped with a partial multiplication and the question arises whether $M$ can be homomorphically embedded into a group, i. e. whether there exists an injective map into a group such that the restriction to each of the three groups is a homomorphism. In order to give a partial answer to this question, we may interpret our three groups and their intersections, equipped with the inclusion maps, as a triangle of groups. By construction of the colimit, the natural homomorphisms $\nu_J : G_J \to \mathfrak{G}$ agree on the intersections and, hence, yield a map $\nu : M \to \mathfrak{G}$. This map is injective if and only if the natural homomorphisms $\nu_J : G_J \to \mathfrak{G}$ are injective and the equations $\mathfrak{G}_{K_1} \cap \mathfrak{G}_{K_2} = \mathfrak{G}_{K_1 \cap K_2}$ hold.

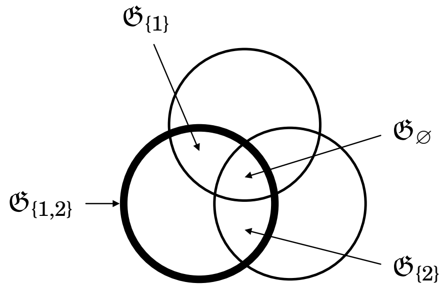


Figure 10: $\mathfrak{G}_{\{1,2\}}$ intersecting $\mathfrak{G}_{\{1,3\}}$ and $\mathfrak{G}_{\{2,3\}}$.

So, if the triangle of groups is non-spherical, the answer is affirmative. On the other hand, it is a consequence of the universal property, see [AHS90, Chapter 3], that if the map $\nu : M \to \mathfrak{G}$ is not injective, then $M$ cannot be homomorphically embedded into any group and the answer is negative.

## 4 Billiards theorem for triangles of groups

In the previous section, we used a combinatorial argument based on Euler's formula for planar graphs to prove the intersection theorem. This kind of argument, be it in the language of homotopies, see e. g. [Sta91], in the language of van Kampen diagrams, see e. g. [ERST00] and [HK06], or in the language of disc pictures, see e. g. [Cor96], turned out to be very powerful in our context, and we highlight the following two results from [ERST00] and [HK06]. In each of them, one considers a non-spherical triangle of groups and assumes that for every $\alpha \in \{1,2,3\}$ there is an element $g_\alpha \in \mathfrak{G}_{\{\alpha\}} \setminus \mathfrak{G}_{\varnothing}$.

**Theorem 4.1 (Edjvet et al.)** *The product $g_1 g_2 g_3 \in \mathfrak{G}$ has infinite order.*

**Theorem 4.2 (Howie-Kopteva)** *If the triangle of groups is hyperbolic, then there is an $n \in \mathbb{N}$ such that the elements $(g_1 g_2 g_3)^n \in \mathfrak{G}$ and $(g_1 g_3 g_2)^n \in \mathfrak{G}$ generate a non-abelian free subgroup.*

**Remark 4.3** *Both, in Theorems 4.1 and 4.2, the authors had the case $G_{\varnothing} = \{1\}$ in mind and, therefore, used $\alpha : G_{\{i\}} * G_{\{j\}} \to G_{\{i,j\}}$ instead of $\alpha : G_{\{i\}} *_{G_{\varnothing}} G_{\{j\}} \to G_{\{i,j\}}$ to define the angle $\sphericalangle_{\{i,j\}}$. Notice that their notion of a group-theoretic angle coincides with Pride's property-$W_k$, which has been introduced in [Pri87] and generalised in [Pri92]. In fact, following their notion, the angle $\sphericalangle_{\{i,j\}}$ is strictly smaller than $\pi/k$ if and only if the group $G_{\{i,j\}}$ has Pride's property-$W_k$. Our notion is reminiscent of Pride's property-$W_k$, but using the amalgamated free product prevents us from measuring the intersection along the group $G_{\varnothing}$. And one can check that Theorems 4.1 and 4.2 also hold in our setting with arbitrary groups $G_{\varnothing}$ and Gersten-Stallings angles as defined in Section 2.2; see [Cun11] for details.*

In the light of Theorem 4.2, one may wonder about the Euclidean case. Let us therefore assume that the triangle of groups is Euclidean and ask under which conditions the colimit group $\mathfrak{G}$ has a non-abelian free subgroup. A first class of examples to look at are *Euclidean triangle groups*:

$$\Delta(k,l,m) = \langle a,b,c : a^2, b^2, c^2, (ab)^k, (ac)^l, (bc)^m \rangle$$
$$\text{with } (k,l,m) \in \{(3,3,3),(2,4,4),(2,3,6)\}$$

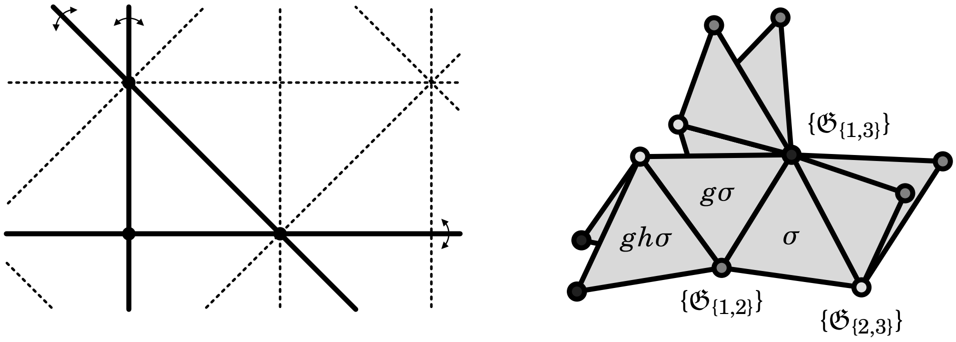


Figure 11: $\Delta(2,4,4)$ realised as $S \leq \operatorname{Isom}(\mathbb{E}^2)$ (left) and a part of the simplicial complex $\mathscr{X}$ including translates of $\sigma$ by group elements $g$ and $gh$ with $g \in \mathfrak{G}_{\{1\}}$ and $h \in \mathfrak{G}_{\{2\}}$ (right).

Each of these groups happens to be the colimit group of the Euclidean triangle of groups based on the following data:

$$G_{\varnothing} = \{1\},\, G_{\{1\}} = \langle a : a^2 = 1 \rangle,\, G_{\{2\}} = \langle b : b^2 = 1 \rangle,\, G_{\{3\}} = \langle c : c^2 = 1 \rangle,$$
$$G_{\{1,2\}} = \langle a, b : a^2 = b^2 = (ab)^k = 1 \rangle,\, G_{\{1,3\}} = \langle a, c : a^2 = c^2 = (ac)^l = 1 \rangle,$$
$$G_{\{2,3\}} = \langle b, c : b^2 = c^2 = (bc)^m = 1 \rangle$$

Here, as in Section 3.1, the homomorphisms $\varphi_{J_1 J_2} : G_{J_1} \to G_{J_2}$ are implicitly given by $a \mapsto a$, $b \mapsto b$, and $c \mapsto c$. It turns out that the Gersten-Stallings angles amount to $\sphericalangle_{\{1,2\}} = \pi/k$, $\sphericalangle_{\{1,3\}} = \pi/l$, and $\sphericalangle_{\{2,3\}} = \pi/m$, whence the triangle of groups is actually Euclidean.

The algebraic structure of the colimit group $\mathfrak{G}$ can be revealed by geometry. Consider three lines in the Euclidean plane $\mathbb{E}^2$ enclosing a triangle with angles $\pi/k$, $\pi/l$, and $\pi/m$, see Figure 11. The reflections along these lines generate a subgroup $S \leq \operatorname{Isom}(\mathbb{E}^2)$ that is isomorphic to $\Delta(k,l,m)$. Since $\operatorname{Isom}(\mathbb{E}^2)$ is solvable, $S \leq \operatorname{Isom}(\mathbb{E}^2)$ must be solvable, too. So, $\Delta(k,l,m)$ is solvable and cannot have a non-abelian free subgroup, see e. g. [Joh80] and [Cun11] for more detailed descriptions of the geometry.

In this section, we generalise the geometric approach and use a construction introduced by Bridson in [Bri91] to study all *non-degenerate* Euclidean triangles of groups, i. e. all Euclidean triangles of groups with the property that each Gersten-Stallings angle is strictly between 0 and $\pi$. Since the strict upper bound $\pi$ is our standing assumption, see Section 2.4, being non-degenerate actually means that none of the Gersten-Stallings angles is 0. But even if there was no such standing assumption, it is clear that a Euclidean triangle of groups is non-degenerate if and only if none of the Gersten-Stallings angles is 0. Given such a triangle of groups, we will construct a simplicial complex $\mathscr{X}$. The action of $\mathfrak{G}$ on $\mathscr{X}$ will give us new insight into the structure of $\mathfrak{G}$. Actually, it turns out that if $\mathscr{X}$ *branches*, i. e. if $|\mathscr{X}|$ is not a topological manifold any more, then the colimit group $\mathfrak{G}$ has a non-abelian free subgroup, see Theorem 4.19. This allows us to give an answer to a problem mentioned by Kopteva and Williams in [KW08, p. 58, l. 24], who wondered if the class of colimit groups of non-spherical triangles of groups satisfies the Tits alternative.

As already mentioned, the construction and the basic properties of $\mathscr{X}$ were introduced by Bridson in [Bri91]. In Section 4.1, we summarise what is relevant for our work, and apply it from Section 4.2 onwards. Our proofs are based on ideas and techniques that go back to two Diplomarbeiten under supervision of Bieri, namely by Lorenz in [Lor95] and Brendel in [Bre04]. Lorenz and Brendel use altitudes in triangles to detect non-abelian free subgroups, but under additional assumptions on the Gersten-Stallings angles. We use the language of billiards instead, which gives us the flexibility we need.

## 4.1 Bridson's simplicial complex

Given a non-degenerate Euclidean triangle of groups, we define an abstract simplicial complex $\mathscr{X}$ as follows (notice that the same definition can be given for any triangle of groups, but we need it only in the non-degenerate Euclidean case):

$$\begin{aligned}
\text{0-simplices} := &\left\{\left\{g\mathfrak{G}_{\{1,2\}}\right\} : g \in \mathfrak{G}\right\} \\
&\sqcup \left\{\left\{g\mathfrak{G}_{\{1,3\}}\right\} : g \in \mathfrak{G}\right\} \\
&\sqcup \left\{\left\{g\mathfrak{G}_{\{2,3\}}\right\} : g \in \mathfrak{G}\right\} \\
\text{1-simplices} := &\left\{\left\{g\mathfrak{G}_{\{1,2\}}, g\mathfrak{G}_{\{1,3\}}\right\} : g \in \mathfrak{G}\right\} \\
&\sqcup \left\{\left\{g\mathfrak{G}_{\{1,2\}}, g\mathfrak{G}_{\{2,3\}}\right\} : g \in \mathfrak{G}\right\} \\
&\sqcup \left\{\left\{g\mathfrak{G}_{\{1,3\}}, g\mathfrak{G}_{\{2,3\}}\right\} : g \in \mathfrak{G}\right\} \\
\text{2-simplices} := &\left\{\left\{g\mathfrak{G}_{\{1,2\}}, g\mathfrak{G}_{\{1,3\}}, g\mathfrak{G}_{\{2,3\}}\right\} : g \in \mathfrak{G}\right\}
\end{aligned}$$

### 4.1.1 Group action and stabilisers

We will use the letter $\sigma$ to denote the 2-simplex that is represented by the identity of the colimit group $\mathfrak{G}$, i. e. $\sigma := \{\mathfrak{G}_{\{1,2\}}, \mathfrak{G}_{\{1,3\}}, \mathfrak{G}_{\{2,3\}}\}$. There is a natural action of $\mathfrak{G}$ on $\mathscr{X}$ given by left-multiplication of each coset. A fundamental domain for this action consists of the 2-simplex $\sigma$ and its faces, see Figure 11.

Bridson mentioned in [Bri91, p. 431, ll. 8–9] that "the pattern of stabilisers in this fundamental domain is precisely the original triangle of groups." Since we distinguish between the diagram and its image in the colimit group, we would replace "original triangle of groups" by "pattern of subgroups $\mathfrak{G}_K$ with $K \subseteq \{1,2,3\}$ and $|K| \leq 2$." Anyway, notice that there is an easy way to prove this observation using the intersection theorem because we can observe that for all pairwise distinct elements $a, b, c \in \{1,2,3\}$:

1. $\operatorname{stab}_{\mathfrak{G}}\left(\left\{\mathfrak{G}_{\{a,b\}}\right\}\right) = \left\{g \in \mathfrak{G} : g\mathfrak{G}_{\{a,b\}} = \mathfrak{G}_{\{a,b\}}\right\} = \mathfrak{G}_{\{a,b\}}$

2. $\operatorname{stab}_{\mathfrak{G}}\left(\left\{\mathfrak{G}_{\{a,b\}}, \mathfrak{G}_{\{a,c\}}\right\}\right)$
   $= \left\{g \in \mathfrak{G} : g\mathfrak{G}_{\{a,b\}} = \mathfrak{G}_{\{a,b\}}, g\mathfrak{G}_{\{a,c\}} = \mathfrak{G}_{\{a,c\}}\right\}$
   $= \mathfrak{G}_{\{a,b\}} \cap \mathfrak{G}_{\{a,c\}} = \mathfrak{G}_{\{a\}}$

3. $\operatorname{stab}_{\mathfrak{G}}\left(\left\{\mathfrak{G}_{\{a,b\}}, \mathfrak{G}_{\{a,c\}}, \mathfrak{G}_{\{b,c\}}\right\}\right)$
   $= \left\{g \in \mathfrak{G} : g\mathfrak{G}_{\{a,b\}} = \mathfrak{G}_{\{a,b\}}, g\mathfrak{G}_{\{a,c\}} = \mathfrak{G}_{\{a,c\}}, g\mathfrak{G}_{\{b,c\}} = \mathfrak{G}_{\{b,c\}}\right\}$
   $= \mathfrak{G}_{\{a,b\}} \cap \mathfrak{G}_{\{a,c\}} \cap \mathfrak{G}_{\{b,c\}} = \mathfrak{G}_{\varnothing}$

Hence, by (1) to (3), the stabilisers of the 0-simplices are the groups $\mathfrak{G}_K$ with $K \subseteq \{1,2,3\}$ and $|K| = 2$. The stabilisers of the 1-simplices and the 2-simplex are their pairwise and triple intersections, which are precisely the groups $\mathfrak{G}_K$ with $K \subseteq \{1,2,3\}$ and $|K| = 1$ and $|K| = 0$.

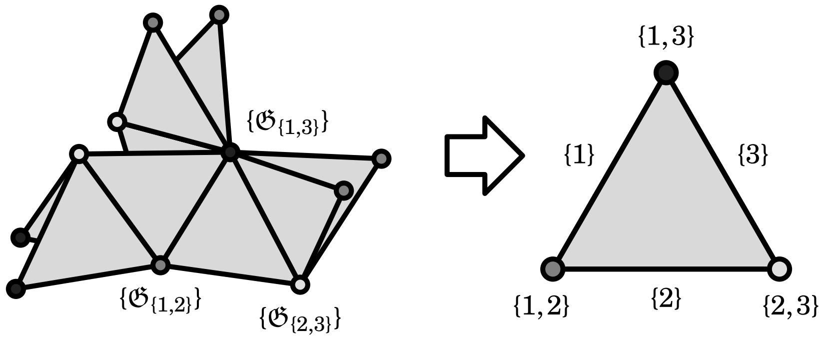


Figure 12: Labelling the triangle $\Delta$ and constructing $h : |\mathscr{X}| \to \Delta$.

#### 4.1.2 Simple connectedness

Let us now consider the geometric realisation $|\mathscr{X}|$. As usual, it is equipped with the *weak topology*. For details about abstract simplicial complexes and their geometric realisations we refer to [Mun84, Sections 1.1–1.3].

**Lemma 4.4 (Behr)** *The geometric realisation $|\mathscr{X}|$ is simply connected.*

Behr has proved a slightly more general version of this lemma in 1975, see [Beh75, Satz 1.2]. Roughly speaking, he translates edge paths in $|\mathscr{X}|$ to products $h_1h_2\cdots h_n$ of elements $h_i \in \mathfrak{G}_{K_i}$ with $K_i \subseteq \{1,2,3\}$ and $|K_i| = 2$, and vice versa.

**Remark 4.5** *In order to keep the notation simple, we use the same symbols to refer to simplices in $\mathscr{X}$ and to their geometric realisations in $|\mathscr{X}|$. Moreover, whenever we talk about a simplex in $|\mathscr{X}|$ without further specification, we mean the closed simplex.*

#### 4.1.3 Metric structure

The geometric realisation $|\mathscr{X}|$ can be equipped with a piecewise Euclidean metric structure. The triangle of groups is non-degenerate and Euclidean. We may therefore pick a closed triangle $\Delta$ in the Euclidean plane $\mathbb{E}^2$ with the property that its angles agree with the Gersten-Stallings angles. For every subset $K \subseteq \{1,2,3\}$ with $|K| = 2$ we label the corresponding vertex of $\Delta$ with angle $\sphericalangle_K$ by $K$. For later purposes, let us label the edges of $\Delta$, too. The edge between the two vertices that are labelled by $K_1$ and $K_2$ is labelled by their intersection $K_1 \cap K_2$.

Next, we construct a continuous map $h : |\mathscr{X}| \to \Delta$. Every 0-simplex is of the form $\{g\mathfrak{G}_K\}$ for some $K \subseteq \{1,2,3\}$ with $|K| = 2$. Map it to the vertex that is labelled by $K$. In order to continue this map to the higher dimensional simplices, map the 1-simplices homeomorphically to the *corresponding* edges, i. e. if a 1-simplex is of the form $\{g\mathfrak{G}_{\{a,b\}}, g\mathfrak{G}_{\{a,c\}}\}$, map it to the edge that is labelled by $\{a,b\} \cap \{a,c\} = \{a\}$. For every 2-simplex $\tau$, use Schoenflies' Theorem to continue the homeomorphism $h|_{\partial\tau} : \partial\tau \to \partial\Delta$ to $h|_\tau : \tau \to \Delta$. The latter ones assemble to the desired continuous map $h : |\mathscr{X}| \to \Delta$. Given two points $x, y \in \tau$ we can now measure their *local distance*:

$$\mathrm{d}_\tau(x,y) := \| h|_\tau(x) - h|_\tau(y) \|$$

The local distance $\mathrm{d}_\tau : \tau \times \tau \to \mathbb{R}$ is a metric on the single 2-simplex $\tau$. The geometric realisation $|\mathscr{X}|$ equipped with the local distances is called an $\mathbb{E}$-*simplicial complex*. For a formal definition, see [Bri91, Section 1.1]. In order to extend the local distances to a metric on $|\mathscr{X}|$, we follow Bridson's work.

**Definition 4.6 ("$m$-chain")** *Let $x, y \in |\mathscr{X}|$. An $m$-chain from $x$ to $y$ is a finite sequence $\mathscr{C} = (x_0, x_1, \ldots, x_m)$ of points in $|\mathscr{X}|$ with the property that $x_0 = x$ and $x_m = y$, and that for every $1 \le i \le m$ both $x_{i-1}$ and $x_i$ are contained in some common 2-simplex $\tau_i$.*

Let $\mathscr{C}$ be an $m$-chain as above. Then, the *length* of $\mathscr{C}$ is defined by:

$$\operatorname{length}(\mathscr{C}) := \sum_{i=1}^{m} \mathrm{d}_{\tau_i}(x_{i-1}, x_i)$$

For every $1 \le i \le m$ there is a unique geodesic from $x_{i-1}$ to $x_i$ in $\tau_i$. We call it the *segment* from $x_{i-1}$ to $x_i$. The concatenation of all segments is called the *path* induced by $\mathscr{C}$. It is denoted by $[\![\mathscr{C}]\!]$. Notice that neither $\operatorname{length}(\lambda)$ nor $[\![\mathscr{C}]\!]$ depends on the choice of $\tau_i$, i. e. if there are two 2-simplices $\tau_i$ and $\widetilde{\tau}_i$ such that $x_{i-1}, x_i \in \tau_i \cap \widetilde{\tau}_i$, then the local distances $\mathrm{d}_{\tau_i}(x_{i-1}, x_i)$ and $\mathrm{d}_{\widetilde{\tau}_i}(x_{i-1}, x_i)$ agree and the segment $[\![x_{i-1}, x_i]\!]$ is well defined.

Let $x, y \in |\mathscr{X}|$. Since $|\mathscr{X}|$ is path connected, there is a path from $x$ to $y$. By the construction given in Behr's proof or by a direct argument, we can even find a path from $x$ to $y$ that is induced by some $m$-chain $\mathscr{C}$. Hence, there is a function $\mathrm{d} : |\mathscr{X}| \times |\mathscr{X}| \to \mathbb{R}$ given by:

$$\mathrm{d}(x, y) := \inf\{\operatorname{length}(\mathscr{C}) : \exists m \text{ such that } \mathscr{C} \text{ is an } m\text{-chain from } x \text{ to } y\}$$

It is easy to see that $\mathrm{d} : |\mathscr{X}| \times |\mathscr{X}| \to \mathbb{R}$ is a pseudometric. On the other hand, *distinguishability*, i. e. $\mathrm{d}(x, y) = 0$ implies $x = y$, is an issue. Here, a lemma from [Bri91, Section 1.2] comes into play. It uses that $|\mathscr{X}|$ is an $\mathbb{E}$-simplicial complex with a finite set of shapes, i. e. with a finite set of isometry classes of simplices.

**Lemma 4.7 (Bridson)** *For every $x \in |\mathscr{X}|$ there is an $\varepsilon(x) > 0$ with the following property: For every $y \in |\mathscr{X}|$ with $\mathrm{d}(x, y) < \varepsilon(x)$ there is a common 2-simplex $\tau$ containing both $x$ and $y$ such that the distances $\mathrm{d}_\tau(x, y)$ and $\mathrm{d}(x, y)$ agree.*

This lemma implies distinguishability, whence the pseudometric $\mathrm{d} : |\mathscr{X}| \times |\mathscr{X}| \to \mathbb{R}$ is actually a metric. Even more, it makes $|\mathscr{X}|$ a complete geodesic metric space, see [Bri91, Theorem 1.1]. As mentioned in [Bri91, p. 381, ll. 25–31], the topology induced by the metric $\mathrm{d} : |\mathscr{X}| \times |\mathscr{X}| \to \mathbb{R}$ is coarser, and may even be strictly coarser, than the weak topology. However, $|\mathscr{X}|$ remains simply connected as a metric space.

**Remark 4.8** *Another important application of Lemma 4.7 concerns the arc length of paths. Given an $m$-chain $\mathscr{C} = (x_0, x_1, \ldots, x_m)$, the metric $\mathrm{d} : |\mathscr{X}| \times |\mathscr{X}| \to \mathbb{R}$ allows us to determine the arc length of the path $[\![\mathscr{C}]\!]$, see e. g. [BH99, Definition I.1.18]. In the light of Lemma 4.7, it is easy to verify that the arc length of the path $[\![\mathscr{C}]\!]$ agrees with* $\operatorname{length}(\mathscr{C})$.

### 4.1.4 CAT(0) property

From now on, we will consider $|\mathscr{X}|$ as a metric space. A crucial observation is that $|\mathscr{X}|$ has the CAT(0) property. In order to prove this, we will verify the *link condition*.

**Definition 4.9 ("geometric link")** *Let $x \in |\mathscr{X}|$. The (geometric) closed star* $\mathrm{St}(x)$ *is the union of the geometric realisations of all simplices that contain $x$. If $y \in \mathrm{St}(x) \setminus \{x\}$, then there is at least one 2-simplex that contains both $x$ and $y$. So, we may consider the segment $[\![x, y]\!]$. Two such segments $[\![x, y]\!]$ and $[\![x, y']\!]$ are called equivalent if one of them is contained in the other. We call the set of equivalence classes the* geometric link *of $x$. It is denoted by* $\mathrm{Lk}(x, |\mathscr{X}|)$.

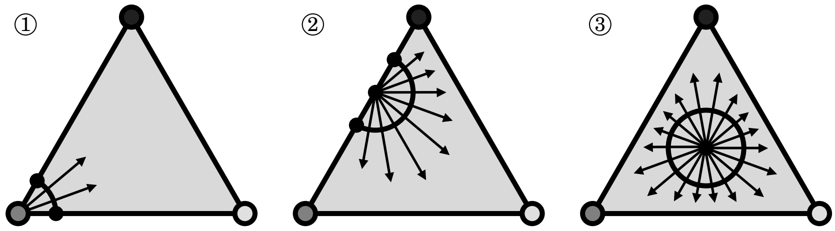


Figure 13: The sets $\mathrm{Lk}(x,\tau)$ for different points $x \in \tau$.

The geometric link $\mathrm{Lk}(x,|\mathscr{X}|)$ can be equipped with a metric structure. First, we consider certain subsets of $\mathrm{Lk}(x,|\mathscr{X}|)$. For every 2-simplex $\tau$ in $\mathrm{St}(x)$ let $\mathrm{Lk}(x,\tau)$ be the subset of all elements of $\mathrm{Lk}(x,|\mathscr{X}|)$ that are represented by segments in $\tau$. Notice that, as soon as one representative has this property, all representatives do. The subset $\mathrm{Lk}(x,\tau) \subseteq \mathrm{Lk}(x,|\mathscr{X}|)$ has a natural metric $\mathrm{d}_{\mathrm{Lk}(x,\tau)} : \mathrm{Lk}(x,\tau) \times \mathrm{Lk}(x,\tau) \to \mathbb{R}$ given by the Euclidean angle:

$$\mathrm{d}_{\mathrm{Lk}(x,\tau)}\left([\![x,y]\!]_\sim, [\![x,y']\!]_\sim\right) := \angle_{h|_\tau(x)}\left(h|_\tau(y), h|_\tau(y')\right) \in [0,\pi]$$

If $x$ is in the 1-skeleton of $|\mathscr{X}|$, then every $(\mathrm{Lk}(x,\tau), \mathrm{d}_{\mathrm{Lk}(x,\tau)})$ is isometrically isomorphic to a closed interval of length $\sphericalangle_{\{1,2\}}$, $\sphericalangle_{\{1,3\}}$, $\sphericalangle_{\{2,3\}}$, or $\pi$, see ① and ② in Figure 13. In particular, we may interpret the subsets $\mathrm{Lk}(x,\tau) \subseteq \mathrm{Lk}(x,|\mathscr{X}|)$ as 1-simplices and, at least after a barycentric subdivision of each simplex, the whole geometric link $\mathrm{Lk}(x,|\mathscr{X}|)$ as a simplicial complex. Even more, it is an $\mathbb{E}$-simplicial complex with a finite set of shapes. As in Section 4.1.3, the connected components of $\mathrm{Lk}(x,|\mathscr{X}|)$ can be equipped with a pseudometric. This pseudometric turns out to be a metric which makes every connected component a complete geodesic metric space. If we set the distance of elements from distinct connected components to $\infty$, we obtain an extended metric $\mathrm{d}_{\mathrm{Lk}(x,|\mathscr{X}|)} : \mathrm{Lk}(x,|\mathscr{X}|) \times \mathrm{Lk}(x,|\mathscr{X}|) \to \mathbb{R} \cup \{\infty\}$.

On the other hand, if $x$ is not in the 1-skeleton of $|\mathscr{X}|$, then it must be in the interior of some 2-simplex $\tau$. In this case, $(\mathrm{Lk}(x,\tau), \mathrm{d}_{\mathrm{Lk}(x,\tau)})$ is isometrically isomorphic to the standard 1-sphere $\mathbb{S}^1$, see ③ in Figure 13. Since $\mathrm{Lk}(x,\tau) = \mathrm{Lk}(x,|\mathscr{X}|)$, the metric $\mathrm{d}_{\mathrm{Lk}(x,\tau)}, \tau) \times \mathrm{Lk}(x,\tau) \to \mathbb{R}$ is already a metric $\mathrm{d}_{\mathrm{Lk}(x,|\mathscr{X}|)} : \mathrm{Lk}(x,|\mathscr{X}|) \times \mathrm{Lk}(x,|\mathscr{X}|) \to \mathbb{R}$ on the geometric link. For more details, in particular for a remark about equivalent definitions, we refer to [BH99, Sections 1.7.14–1.7.15]. Notice that Bridson and Haefliger consider the open star instead of the closed one. But, in the end, this does not make a difference.

**Definition 4.10 ("link condition")** *The geometric realisation $|\mathscr{X}|$ satisfies the link condition if for every $x \in |\mathscr{X}|$ and every pair of points $a, b \in \mathrm{Lk}(x,|\mathscr{X}|)$ with $\mathrm{d}_{\mathrm{Lk}(x,|\mathscr{X}|)}(a,b) < \pi$ there is a unique geodesic from a to b, or, equivalently, if every injective loop $\lambda : \mathbb{S}^1 \to \mathrm{Lk}(x,|\mathscr{X}|)$ has arc length at least $2\pi$.*

The equivalence relies on the fact that $|\mathscr{X}|$ is a 2-dimensional $\mathbb{E}$-simplicial complex with a finite set of shapes. One can either prove it directly or apply [BH99, Theorem I.7.55, (3) ⇔ (1)], [BH99, Theorem II.5.5, (3) ⇔ (2)], and [BH99, Lemma II.5.6].

**Lemma 4.11 (Bridson, Gersten-Stallings)** *The geometric realisation $|\mathscr{X}|$ satisfies the link condition.*

A proof of this lemma has been given by Bridson in [Bri91, p. 431, ll. 19–25] and by Gersten and Stallings in [Sta91, p. 499, ll. 19–28]. The argument is both simple and important, so let us outline the main ideas. The most interesting case occurs when $x$ is a 0-simplex, w. l. o. g. $x \in \{\{g\mathfrak{G}_{\{1,2\}}\} : g \in \mathfrak{G}\}$.

Recall the definition of $\sigma$ from Section 4.1.1, it allows us to describe the set of 2-simplices of $|\mathscr{X}|$ as the orbit $\{g\sigma : g \in \mathfrak{G}\}$. Every injective loop $\lambda : \mathbb{S}^1 \to \mathrm{Lk}(x,|\mathscr{X}|)$ traverses some finite number of 1-simplices, say $m$. If the first one is $\mathrm{Lk}(x, g\sigma)$, the next ones are $\mathrm{Lk}(x, gh_1h_2\cdots h_i\sigma)$ with elements $h_i$ alternately in $\mathfrak{G}_{\{1\}} \setminus \mathfrak{G}_\varnothing$ and $\mathfrak{G}_{\{2\}} \setminus \mathfrak{G}_\varnothing$. At the end, $\lambda$ traverses $\mathrm{Lk}(x, gh_1h_2\cdots h_{m-1}\sigma)$. Since $\lambda$ is a loop, we know that there is an $h_m$ such that $\mathrm{Lk}(x, gh_1h_2\cdots h_m\sigma) = \mathrm{Lk}(x, g\sigma)$, which is equivalent to $h_1h_2\cdots h_m \in \mathfrak{G}_\varnothing$. Let us write $h := h_1h_2\cdots h_m$. We may assume w. l. o. g. that $h_1, h_3, \ldots, h_{m-1} \in \mathfrak{G}_{\{1\}} \setminus \mathfrak{G}_\varnothing$ and $h_2, h_4, \ldots, h_m \in \mathfrak{G}_{\{2\}} \setminus \mathfrak{G}_\varnothing$. Since $h \in \mathfrak{G}_\varnothing$, also $h_mh^{-1} \in \mathfrak{G}_{\{2\}} \setminus \mathfrak{G}_\varnothing$. Now, we proceed similarly to (13) in the proof of the intersection theorem. We construct the preimages under the injective homomorphisms $\nu_K : G_K \to \mathfrak{G}$:

$$\nu_{\{1\}}{}^{-1}(h_1), \nu_{\{1\}}{}^{-1}(h_3), \ldots, \nu_{\{1\}}{}^{-1}(h_{m-1}) \in G_{\{1\}} \setminus \varphi_{\varnothing\{1\}}(G_\varnothing)$$
$$\nu_{\{2\}}{}^{-1}(h_2), \nu_{\{2\}}{}^{-1}(h_4), \ldots, \nu_{\{2\}}{}^{-1}(h_mh^{-1}) \in G_{\{2\}} \setminus \varphi_{\varnothing\{2\}}(G_\varnothing)$$

Again, the preimages assemble to an element $\nu_{\{1\}}{}^{-1}(h_1)\cdot\nu_{\{2\}}{}^{-1}(h_2)\cdot\ldots\cdot\nu_{\{2\}}{}^{-1}(h_mh^{-1})$ of the amalgamated free product $G_{\{1\}} *_{G_\varnothing} G_{\{2\}}$ that is contained in the kernel of the homomorphism $\alpha : G_{\{1\}} *_{G_\varnothing} G_{\{2\}} \to G_{\{1,2\}}$ introduced in Section 2.2. But, by the normal form theorem, see [Mil68, Lemma 1], this element is non-trivial. Therefore, $m \geq 2\pi/\sphericalangle_{\{1,2\}}$. Now, Remark 4.8 also holds for $m$-chains in $\mathrm{Lk}(x,|\mathscr{X}|)$. It implies that the arc length of $\lambda$ is equal to $m \cdot \sphericalangle_{\{1,2\}}$, which can be estimated from below by $2\pi/\sphericalangle_{\{1,2\}} \cdot \sphericalangle_{\{1,2\}} = 2\pi$, whence we are done. The link condition for the other cases, i. e. if $x$ is in the interior of a 1-simplex or in the interior of a 2-simplex, is almost immediate.

Once we have convinced ourselves that the geometric realisation $|\mathscr{X}|$ satisfies the link condition, we may apply Bridson's main theorem [Bri91, Section 2, Main Theorem, (11) ⇒ (2)].

**Theorem 4.12 (Bridson)** *Since $|\mathscr{X}|$ is a simply connected $\mathbb{E}$-simplicial complex with a finite set of shapes that satisfies the link condition, it has the* CAT(0) *property.*

### 4.1.5 Geodesics

Let $[\![\mathscr{C}]\!]$ be the path induced by an $m$-chain $\mathscr{C} = (x_0, x_2, \ldots, x_m)$. There is a necessary and, as we will see in Lemma 4.15, sufficient condition for $[\![\mathscr{C}]\!]$ to be a geodesic, namely that $\mathscr{C}$ is *straight*, i. e. that there are no obvious shortcuts at the points $x_1, x_3, \ldots, x_{m-1}$. Let us make the notion of straightness a little more precise.

**Definition 4.13 ("straight $m$-chain")** *An $m$-chain $\mathscr{C} = (x_0, x_1, \ldots, x_m)$ is called straight if for every $1 \leq i \leq m-1$ the distance between $[\![x_i, x_{i-1}]\!]_\sim$ and $[\![x_i, x_{i+1}]\!]_\sim$ in $\mathrm{Lk}(x_i,|\mathscr{X}|)$ is at least $\pi$.*

**Remark 4.14** *In general, it is not easy to determine the distance between $[\![x_i, x_{i-1}]\!]_\sim$ and $[\![x_i, x_{i+1}]\!]_\sim$ in $\mathrm{Lk}(x_i,|\mathscr{X}|)$. But the link condition ensures that, once we are able to connect them by an injective path of arc length $\pi$, the distance between them is actually equal to $\pi$.*

Now, the CAT(0) property comes into play. It allows us to conclude from the local property "straight $m$-chain" to the global property "geodesic".

**Lemma 4.15 (Bridson)** *If $\mathscr{C}$ is a straight $m$-chain, then $[\![\mathscr{C}]\!]$ is a geodesic.*

This lemma can be proved by showing that every straight $m$-chain induces a local geodesic, which is an easy consequence of [Bri91, Section 2, Main Theorem, (2) ⇒ (5)]. And once we know this, [BH99, Proposition II.1.4 (2)] tells us that every local geodesic is a geodesic.

**Remark 4.16** *By Lemma 4.15, we are now able to construct geodesics easily; and these geodesics are unique, see [Bri91, Section 2, Main Theorem, (2) ⇒ (1)], which will be of relevance in the proof of the billiards theorem.*

## 4.2 Billiards theorem

In this section, we still assume given a non-degenerate Euclidean triangle of groups and consider *billiard shots* and *billiard sequences* on the triangle $\Delta$.

### 4.2.1 Billiard shots and billiard sequences

We choose some point $y_0$ in the interior of $\Delta$ and some direction. Then, we consider the path that starts at $y_0$ and goes in a straight line into the chosen direction. Eventually, this path leaves $\Delta$. Let $y_1 \in \partial\Delta$ be its last point in $\Delta$. If this point is a vertex ("the ball is in the pocket"), we withdraw the path. Otherwise, it is in the interior of an edge ("the ball hits the cushion"), which allows us to reflect the path according to the rule that the *angle of incidence* is equal to the *angle of reflection*. Now, we can go on. Whenever the path leaves $\Delta$ at some point in the interior of an edge, we reflect it again. After some finite number of reflections, say at the points $y_1, y_2, \ldots, y_{m-1} \in \partial\Delta$, we stop at some point $y_m$ in the interior of $\Delta$. The sequence $\mathscr{B} = (y_0, y_1, \ldots, y_m)$ is called a *billiard sequence*, the induced path $[\![\mathscr{B}]\!] = [\![y_0, y_1, \ldots, y_m]\!]$ is called a *billiard shot*.

### 4.2.2 Statement and proof of the billiards theorem

The notion of billiard shots and billiard sequences allows us to prove that certain elements of the colimit group $\mathfrak{G}$ are non-trivial.

**Definition 4.17 ("adapted")** *Given a billiard sequence $\mathscr{B} = (y_0, y_1, \ldots, y_m)$, we call an element $g \in \mathfrak{G}$ adapted to $\mathscr{B}$, if it is a product $g_1 g_2 \cdots g_{m-1}$ such that each $g_i \in \mathfrak{G}_{\{a_i\}} \setminus \mathfrak{G}_\varnothing$, where $\{a_i\}$ is the label of the edge whose interior contains $y_i$.*

**Theorem 4.18** *Assume we are given a non-degenerate Euclidean triangle of groups and a closed triangle $\Delta$ in the Euclidean plane $\mathbb{E}^2$ as constructed in Section 4.1.3. If an element $g \in \mathfrak{G}$ is adapted to a billiard sequence $\mathscr{B} = (y_0, y_1, \ldots, y_m)$ on $\Delta$ with at least one reflection, i. e. with $m \geq 2$, then $g$ is non-trivial.*

*Proof.* The idea is to lift the billiard shot $[\![\mathscr{B}]\!]$ to the geometric realisation $|\mathscr{X}|$. For every $1 \leq i \leq m$ we use $h|_{g_1 g_2 \cdots g_{i-1}\sigma} : g_1 g_2 \cdots g_{i-1}\sigma \to \Delta$ to lift the segment $[\![y_{i-1}, y_i]\!]$. Let us make some observations.

(1) *These lifts assemble to a path in $|\mathscr{X}|$.* For every $1 \leq i \leq m-1$ the first segment $[\![y_{i-1}, y_i]\!]$ is lifted by $h|_{g_1 g_2 \cdots g_{i-1}\sigma}$, the second segment $[\![y_i, y_{i+1}]\!]$ is lifted by $h|_{g_1 g_2 \cdots g_i \sigma}$. To show that these two lifts actually fit together, we convince ourselves that in either case $y_i$ is lifted to the same point. If the edge whose interior contains $y_i$ is labelled by $\{a\}$, then the two adjacent vertices are labelled by $\{a, b\}$ and $\{a, c\}$, where $b$ and $c$ are the remaining two elements of $\{1,2,3\}$. We can therefore observe:
$$\left(h|_{g_1 g_2 \cdots g_{i-1}\sigma}\right)^{-1}(y_i) \in \left\{g_1 g_2 \cdots g_{i-1} \mathfrak{G}_{\{a,b\}}, g_1 g_2 \cdots g_{i-1} \mathfrak{G}_{\{a,c\}}\right\}$$
$$\left(h|_{g_1 g_2 \cdots g_i \sigma}\right)^{-1}(y_i) \in \left\{g_1 g_2 \cdots g_i \mathfrak{G}_{\{a,b\}}, g_1 g_2 \cdots g_i \mathfrak{G}_{\{a,c\}}\right\}$$
Since $g_i \in \mathfrak{G}_{\{a\}}$, the two 1-simplices agree. Call them $\tau$ and observe:
$$\left(h|_{g_1 g_2 \cdots g_{i-1}\sigma}\right)^{-1}(y_i) = \left(h|_\tau\right)^{-1}(y_i) = \left(h|_{g_1 g_2 \cdots g_i \sigma}\right)^{-1}(y_i)$$
So, the lift of $y_i$ is well defined. Let us denote it by $x_i$. For the lifts of the extremal points $y_0$ and $y_m$ we define analogously $x_0 := (h|_\sigma)^{-1}(y_0)$ and $x_m := (h|_{g\sigma})^{-1}(y_m)$, whence the lifts of each two segments $[\![y_{i-1}, y_i]\!]$ and $[\![y_i, y_{i+1}]\!]$ assemble to the path $[\![x_{i-1}, x_i, x_{i+1}]\!]$ and, more general, the lifts of all segments $[\![y_0, y_1]\!], [\![y_1, y_2]\!], \ldots, [\![y_{m-1}, y_m]\!]$ assemble to the path $[\![\mathscr{C}]\!]$ induced by the $m$-chain $\mathscr{C} := (x_0, x_1, \ldots, x_m)$.

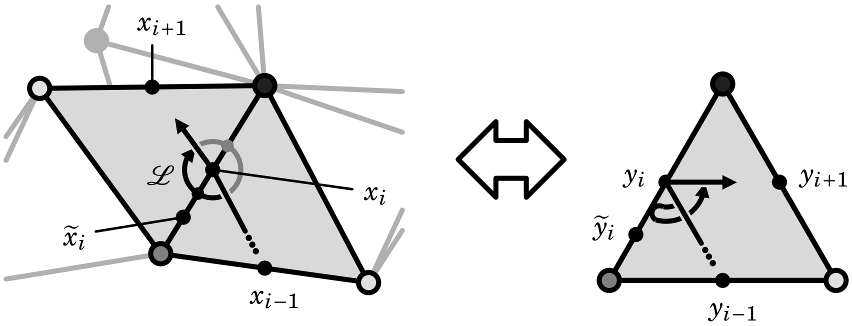


Figure 14: Lift the billiard shot to $g_1g_2\cdots g_{i-1}\sigma$ and $g_1g_2\cdots g_i\sigma$.

(2) *The m-chain $\mathscr{C}$ is straight.* Let $1 \le i \le m-1$. We construct a 2-chain $\mathscr{L}$ from $[\![x_i, x_{i-1}]\!]_\sim$ to $[\![x_i, x_{i+1}]\!]_\sim$ in $\mathrm{Lk}(x_i, |\mathscr{X}|)$ such that $\mathrm{length}(\mathscr{L}) = \pi$ and the path $[\![\mathscr{L}]\!]$ is injective. By Remark 4.8, the path $[\![\mathscr{L}]\!]$ is of arc length $\pi$ and, by Remark 4.14, the distance between $[\![x_i, x_{i-1}]\!]_\sim$ and $[\![x_i, x_{i+1}]\!]_\sim$ in $\mathrm{Lk}(x_i, |\mathscr{X}|)$ is equal to $\pi$. First, consider the edge whose interior contains $y_i$ and choose another point $\widetilde{y}_i$ in the interior of this edge. Let $\widetilde{x}_i := (h|_{g_1g_2\cdots g_{i-1}\sigma})^{-1}(\widetilde{y}_i) = (h|_{g_1g_2\cdots g_i\sigma})^{-1}(\widetilde{y}_i)$ be its lift. Then, move to the geometric link $\mathrm{Lk}(x_i, |\mathscr{X}|)$ and construct the 2-chain illustrated in Figure 14, namely $\mathscr{L} := ([\![x_i, x_{i-1}]\!]_\sim, [\![x_i, \widetilde{x}_i]\!]_\sim, [\![x_i, x_{i+1}]\!]_\sim)$.

Observe that the path $[\![\mathscr{L}]\!]$ traverses the interval $\mathrm{Lk}(x_i, g_1g_2\cdots g_{i-1}\sigma)$ until it reaches its endpoint $[\![x_i, \widetilde{x}_i]\!]_\sim$. Then, it traverses the interval $\mathrm{Lk}(x_i, g_1g_2\cdots g_i\sigma)$. Therefore:

$$\begin{aligned}\mathrm{length}(\mathscr{L}) &= \mathrm{d}_{\mathrm{Lk}(x_i, g_1g_2\cdots g_{i-1}\sigma)}([\![x_i, x_{i-1}]\!]_\sim, [\![x_i, \widetilde{x}_i]\!]_\sim)\\ &\quad + \mathrm{d}_{\mathrm{Lk}(x_i, g_1g_2\cdots g_i\sigma)}([\![x_i, \widetilde{x}_i]\!]_\sim, [\![x_i, x_{i+1}]\!]_\sim)\\ &= \angle_{y_i}(y_{i-1}, \widetilde{y}_i) + \angle_{y_i}(\widetilde{y}_i, y_{i+1}) = \pi\end{aligned}$$

Since $g_i \notin \mathfrak{G}_\varnothing$, the two intervals traversed by $[\![\mathscr{L}]\!]$ are actually not the same. So, $[\![\mathscr{L}]\!]$ must be injective.

With these two observations in mind the final conclusion that $g$ is non-trivial in $\mathfrak{G}$ is almost immediate. By Lemma 4.15, $[\![\mathscr{C}]\!]$ is a geodesic from $x_0 \in \sigma$ to $x_m \in g\sigma$. Before going on, observe that any two points in $\sigma$ can be connected by a 1-chain, which is, of course, straight and therefore induces a geodesic. But, as mentioned in Remark 4.16, geodesics are unique. Hence, the unique geodesic between any two points in $\sigma$ is completely contained in $\sigma$. Let us now go back to our situation. We assume that $m \ge 2$. So, the geodesic $[\![\mathscr{C}]\!]$ leaves $\sigma$ eventually and, therefore, does not end in $\sigma$, i. e. $x_m \notin \sigma$, which implies that $g\sigma \ne \sigma$ and, finally, $g \ne 1$. □

### 4.2.3 A first example

We conclude Section 4.2 with an example that illustrates the application of the billiards theorem. Assume we are given a triangle of groups with Gersten-Stallings angles $\sphericalangle_{\{1,2\}} = \sphericalangle_{\{1,3\}} = \sphericalangle_{\{2,3\}} = \pi/3$. Since none of them is equal to 0, it is easy to see[2] that for every $a \in \{1,2,3\}$ there is an element $g_a \in \mathfrak{G}_{\{a\}} \setminus \mathfrak{G}_\varnothing$.

[2] If there was no element $g_a \in \mathfrak{G}_{\{a\}} \setminus \mathfrak{G}_\varnothing$, then $\mathfrak{G}_{\{a\}} = \mathfrak{G}_\varnothing$, which means $\widetilde{\nu}_{\{a\}}(\mathfrak{G}_{\{a\}}) = \widetilde{\nu}_\varnothing(\mathfrak{G}_\varnothing)$. Now, observe that $\nu_{\{a\}}(G_{\{a\}}) = \widetilde{\nu}_{\{a\}}(\mu_{\{a\}}(G_{\{a\}})) = \widetilde{\nu}_{\{a\}}(\mathfrak{G}_{\{a\}}) = \widetilde{\nu}_\varnothing(\mathfrak{G}_\varnothing) = \widetilde{\nu}_\varnothing(\mu_\varnothing(G_\varnothing)) = \nu_\varnothing(G_\varnothing) = \nu_{\{a\}}(\varphi_{\varnothing\{a\}}(G_\varnothing))$. Since the homomorphism $\nu_{\{a\}} : G_{\{a\}} \to \mathfrak{G}$ is injective, we obtain $G_{\{a\}} = \varphi_{\varnothing\{a\}}(G_\varnothing)$, whence $G_{\{a\}} *_{G_\varnothing} G_{\{b\}}$ is generated by $G_{\{b\}}$. So, the homomorphism $\alpha : G_{\{a\}} *_{G_\varnothing} G_{\{b\}} \to G_{\{a,b\}}$ is injective and the Gersten-Stallings angle $\sphericalangle_{\{a,b\}} = 0$.

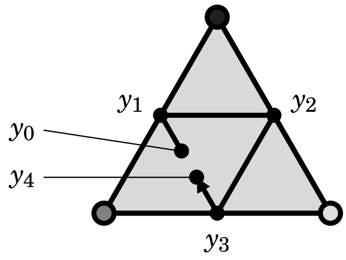


Figure 15: A first billiard shot.

Their product $g := g_1g_2g_3 \in \mathfrak{G}$, which has been considered in Theorem 4.1, is adapted to the billiard sequence $\mathscr{B} = (y_0, y_1, y_2, y_3, y_4)$ drawn in Figure 15 and is therefore non-trivial. We may also continue the billiard shot. This yields billiard sequences of the form $(y_0, y_1, y_2, y_3, y_1, y_2, y_3, \dots, y_1, y_2, y_3, y_4)$. Every *power* of $g$, i. e. every element $g^n \in \mathfrak{G}$ with $n \in \mathbb{N}$, is adapted to such a billiard sequence and is therefore non-trivial. Hence, $g$ has infinite order.

### 4.3 Constructing non-abelian free subgroups

Again, assume we are given a non-degenerate Euclidean triangle of groups. So, for every $a \in \{1,2,3\}$ there is an element $g_a \in \mathfrak{G}_{\{a\}} \setminus \mathfrak{G}_\varnothing$. Let us recall the notion of branching. We say that the simplicial complex $\mathscr{X}$ *branches* if the geometric realisation $|\mathscr{X}|$ is not a topological manifold any more. It is easy to see that $\mathscr{X}$ branches if and only if there is an $a \in \{1,2,3\}$ such that the index of $\mathfrak{G}_\varnothing$ in $\mathfrak{G}_{\{a\}}$ is at least 3 or there are two distinct $a, b \in \{1,2,3\}$ such that $\mathfrak{G}_{\{a,b\}}$ is not generated by $\mathfrak{G}_{\{a\}}$ and $\mathfrak{G}_{\{b\}}$. The following theorem says that branching already implies the existence of non-abelian free subgroups in $\mathfrak{G}$.

**Theorem 4.19** *Assume we are given a non-degenerate Euclidean triangle of groups. If the simplicial complex $\mathscr{X}$ branches, the colimit group $\mathfrak{G}$ has a non-abelian free subgroup.*

**Remark 4.20** *Notice that Theorem 4.19 is the 2-dimensional analogue of a well known fact. Consider an amalgamated free product $X *_A Y$ with the property that the image of $A$ in $X$ and the image of $A$ in $Y$ have index at least 2. The associated Bass-Serre tree $\mathscr{T}$ branches if and only if one of the indices is at least 3. In this case, the amalgamated free product $X *_A Y$ has a non-abelian free subgroup.*

*Proof of Theorem 4.19.* We may assume w. l. o. g. that $\sphericalangle_{\{1,2\}} \geq \sphericalangle_{\{1,3\}} \geq \sphericalangle_{\{2,3\}}$. So, there are exactly three possibilities for the Gersten-Stallings angles, each of which is considered in a separate column in Figure 16.

First, if there is an $a \in \{1,2,3\}$ such that the index of $\mathfrak{G}_\varnothing$ in $\mathfrak{G}_{\{a\}}$ is at least 3, consider the element $h \in \mathfrak{G}$ that is given in the respective entry in Figure 16. It is constructed in such a way that both $h$ and $h^{-1}$ are adapted to a billiard sequence $\mathscr{B}_1$ with the following property: The billiard shot $[\![\mathscr{B}_1]\!]$ starts at some point in the interior of $\Delta$ and goes orthogonally away from the edge labelled by $\{a\}$. After a couple of reflections, it comes back to the starting point, but in the opposite direction, see ① in Figure 16.

Notice that, given an element $g \in \mathfrak{G}$ with a decomposition into factors that are alternately from $\{h, h^{-1}\}$ and $\mathfrak{G}_{\{a\}} \setminus \mathfrak{G}_\varnothing$, we may concatenate the billiard shot $[\![\mathscr{B}_1]\!]$ and the orthogonal reflection at the edge labelled by $\{a\}$, see ② in Figure 16, accordingly. This yields a billiard sequence $\mathscr{B}_2$, which $g$ is adapted to. Hence, we know: *If there is at least one factor in the decomposition of $g$, then there is at least one reflection in the billiard sequence $\mathscr{B}_2$ and, by the billiards theorem, $g$ is non-trivial.*

Since the index of $\mathfrak{G}_\varnothing$ in $\mathfrak{G}_{\{a\}}$ is at least 3, we can find an element $\tilde{g}_a \in \mathfrak{G}_{\{a\}}$ that is neither in $\mathfrak{G}_\varnothing$ nor in $g_a\mathfrak{G}_\varnothing$. In particular, neither $\tilde{g}_a{}^{-1}g_a$ nor $g_a{}^{-1}\tilde{g}_a$ is in $\mathfrak{G}_\varnothing$. Define $x := g_a h \tilde{g}_a{}^{-1} \in \mathfrak{G}$ and $y := hg_a h \tilde{g}_a{}^{-1} h^{-1} \in \mathfrak{G}$. We claim that $x$ and $y$ generate a non-abelian free subgroup of $\mathfrak{G}$. Consider a non-empty freely reduced word over the letters $x$ and $y$ and their formal inverses. The element $g \in \mathfrak{G}$ that is represented by this word has a natural decomposition into factors from $\{h^{\pm 1}, g_a{}^{\pm 1}, \tilde{g}_a{}^{\pm 1}\}$. Cancel each $h^{-1}h$ and subsume each $\tilde{g}_a{}^{-1}g_a$ and each $g_a{}^{-1}\tilde{g}_a$ to a single element in $\mathfrak{G}_{\{a\}} \smallsetminus \mathfrak{G}_\varnothing$. This yields a new decomposition of $g$ into factors that are alternately from $\{h, h^{-1}\}$ and $\mathfrak{G}_{\{a\}} \smallsetminus \mathfrak{G}_\varnothing$. It is easy to see that, despite of the cancellation of each $h^{-1}h$, there is at least one factor left in the new decomposition of $g$. So, by our preliminary discussion, $g$ is non-trivial, which completes the proof that $x$ and $y$ generate a non-abelian free subgroup of $\mathfrak{G}$.

Second, consider the case that there are two distinct $a, b \in \{1,2,3\}$ such that $\mathfrak{G}_{\{a,b\}}$ is not generated by $\mathfrak{G}_{\{a\}}$ and $\mathfrak{G}_{\{b\}}$. In this case, let $X := \mathfrak{G}_{\{a,b\}}$, $A := \langle \mathfrak{G}_{\{a\}}, \mathfrak{G}_{\{b\}} \rangle \leq \mathfrak{G}$, and $Y := \langle \mathfrak{G}_{\{a,c\}}, \mathfrak{G}_{\{b,c\}} \rangle \leq \mathfrak{G}$, where $c$ is the remaining element of $\{1,2,3\}$.

Using the presentation ($*$) one can show that $\mathfrak{G} \cong X *_A Y$. Here, the homomorphisms are the ones induced by the inclusions. By assumption, $|X : A| \geq 2$. On the other hand, we know that there is an element $g_c \in \mathfrak{G}_{\{c\}} \smallsetminus \mathfrak{G}_\varnothing$. By the intersection theorem, $\mathfrak{G}_{\{c\}} \cap \mathfrak{G}_{\{a,b\}} = \mathfrak{G}_\varnothing$. So, $g_c \in \mathfrak{G}_{\{c\}} \smallsetminus \mathfrak{G}_{\{a,b\}} \subseteq Y \smallsetminus A$ and $|Y : A| \geq 2$. If $|Y : A| = 2$, then $A$ is a normal subgroup of $Y$. Again, by the intersection theorem:

$$g_c{}^{-1} g_a g_c \in A \cap \mathfrak{G}_{\{a,c\}} \subseteq \mathfrak{G}_{\{a,b\}} \cap \mathfrak{G}_{\{a,c\}} = \mathfrak{G}_{\{a\}}$$
$$g_c{}^{-1} g_b g_c \in A \cap \mathfrak{G}_{\{b,c\}} \subseteq \mathfrak{G}_{\{a,b\}} \cap \mathfrak{G}_{\{a,b\}} = \mathfrak{G}_{\{b\}}$$

This implies that $\sphericalangle_{\{a,c\}} = \sphericalangle_{\{b,c\}} = \pi/2$. Hence $\sphericalangle_{\{a,b\}} = 0$, which is not possible since we assume the triangle of groups to be non-degenerate. So, $|Y : A| \geq 3$ and, by Remark 4.20, $\mathfrak{G} \cong X *_A Y$ has a non-abelian free subgroup. □

**Remark 4.21** *The idea of Theorems 4.18 and 4.19 is certainly ping-pong-ish. In the above proof, we construct products $g = g_1 g_2 \cdots g_m$ whose factors $g_i$ are alternately from $\{h, h^{-1}\}$ and $\mathfrak{G}_{\{a\}} \smallsetminus \mathfrak{G}_\varnothing$. Therefore, the sequence $\sigma \mapsto g_m\sigma \mapsto g_{m-1}g_m\sigma \mapsto \ldots \mapsto g_1 \cdots g_{m-1} g_m \sigma$ moves the fundamental domain $\sigma$ back and forth through $|\mathscr{X}|$. But, instead of defining ping-pong sets, we construct geodesics to ensure that the final position of $\sigma$ is actually different from the initial one. The language of billiards helps us to see these geodesics without getting unnecessarily confused by the surrounding complex.*

## 4.4 Tits alternative

In this section, we ask about the cases that are not covered by Theorems 4.2 and 4.19, and discuss the following version of the Tits alternative.

**Definition 4.22 ("Tits alternative")** *A class $\mathscr{C}$ of groups satisfies the Tits alternative if each $G \in \mathscr{C}$ either has a non-abelian free subgroup or is virtually solvable.*

**Remark 4.23** *There are groups that neither have a non-abelian free subgroup nor are virtually solvable. For example, take Thompson's group $F$. It has been shown by Brin and Squier in [BS85] that $F \leq \mathrm{PLF}(\mathbb{R})$ doesn't have a non-abelian free subgroup. And if $F$ was virtually solvable, then $[F,F]$ would have to be virtually solvable, too. But $[F,F]$ is infinite and simple, see [CFP96, Section 4], which implies that $[F,F]$ cannot be virtually solvable.*

We may use Thompson's group $F$ to prove that the Tits alternative doesn't hold for the class of colimit groups of non-spherical triangles of groups. For example, let $\Gamma_1$ be the triangle of groups with

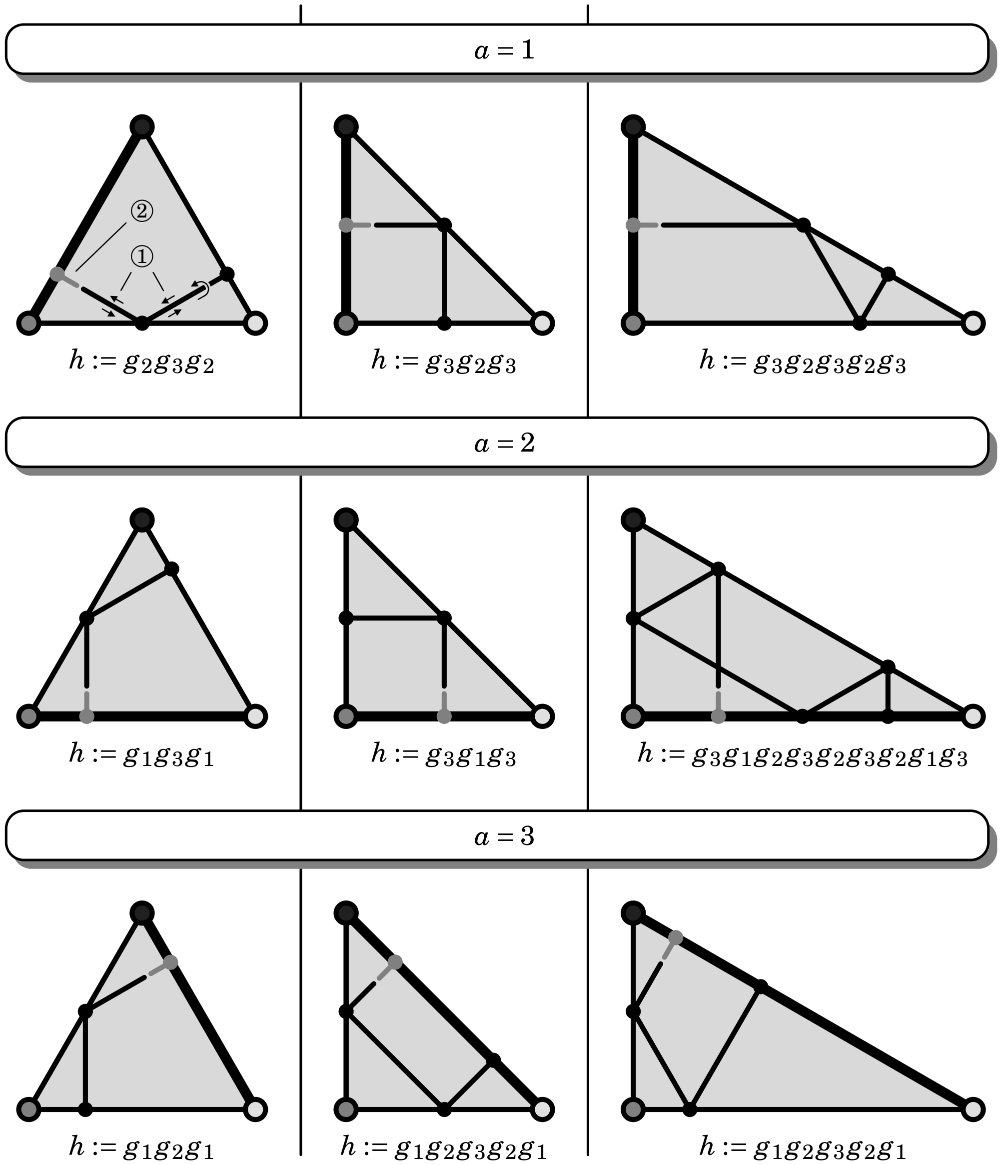


Figure 16: The first column refers to Gersten-Stallings angles $\sphericalangle_{\{1,2\}} = \sphericalangle_{\{1,3\}} = \sphericalangle_{\{2,3\}} = \pi/3$, the second to $\sphericalangle_{\{1,2\}} = \pi/2$, $\sphericalangle_{\{1,3\}} = \sphericalangle_{\{2,3\}} = \pi/4$, the third to $\sphericalangle_{\{1,2\}} = \pi/2$, $\sphericalangle_{\{1,3\}} = \pi/3$, $\sphericalangle_{\{2,3\}} = \pi/6$.

the property that the groups $G_J$ are all equal to $F$ and the injective homomorphisms $\varphi_{J_1J_2}: G_{J_1} \to G_{J_2}$ are all identities. The Gersten-Stallings angles amount to $\sphericalangle_{\{1,2\}} = \sphericalangle_{\{1,3\}} = \sphericalangle_{\{2,3\}} = 0$, whence $\Gamma_1$ is a degenerate hyperbolic triangle of groups. But the colimit group is isomorphic to $F$ and, therefore, neither has a non-abelian free subgroup nor is virtually solvable.

Notice that there are non-degenerate examples, too. Pick one of the three triangles of groups given in the introduction to Section 4 and replace every group $G_J$ by $F \times G_J$ and every injective homomorphism $\varphi_{J_1J_2}: G_{J_1} \to G_{J_2}$ by $\mathrm{id}_F \times \varphi_{J_1J_2}: F \times G_{J_1} \to F \times G_{J_2}$. The Gersten-Stallings angles remain the same, whence the new triangle of groups, call it $\Gamma_2$, is non-degenerate and Euclidean. But the colimit group is isomorphic to $F \times \Delta(k,l,m)$ and, therefore, neither has a non-abelian free subgroup nor is virtually solvable.[3]

Let us now assume that $G_\varnothing$ either has a non-abelian free subgroup or is virtually solvable, which is certainly true if $G_\varnothing = \{1\}$ as in Remark 4.3. In the non-degenerate case, this assumption already implies the Tits alternative.

**Theorem 4.24** *The Tits alternative holds for the class of colimit groups of non-degenerate non-spherical triangles of groups with the property that the group $G_\varnothing$ either has a non-abelian free subgroup or is virtually solvable.*

Interestingly, in the degenerate case, it doesn't. Just consider the triangle of groups $\Gamma_3$ given by the following data:

$$G_\varnothing = \{1\},\ G_{\{1\}} = F,\ G_{\{2\}} = \langle a : a^2 = 1\rangle,\ G_{\{3\}} = \langle b : b^2 = 1\rangle,$$
$$G_{\{1,2\}} = F \times \langle a : a^2 = 1\rangle,\ G_{\{1,3\}} = F \times \langle b : b^2 = 1\rangle,\ G_{\{2,3\}} = \langle a, b : a^2 = b^2 = 1\rangle$$

Here, the homomorphisms $\varphi_{J_1J_2}: G_{J_1} \to G_{J_2}$ are given by $\forall f \in F : f \mapsto (f,1)$, by $a \mapsto (1,a)$ and $a \mapsto a$, and by $b \mapsto (1,b)$ and $b \mapsto b$. The Gersten-Stallings angles amount to $\sphericalangle_{\{1,2\}} = \pi/2$, $\sphericalangle_{\{1,3\}} = \pi/2$, and $\sphericalangle_{\{2,3\}} = 0$, whence $\Gamma_3$ is a degenerate Euclidean triangle of groups. Its colimit group is isomorphic to $F \times (\mathbb{Z}_2 * \mathbb{Z}_2)$ and, therefore, neither has a non-abelian free subgroup nor is virtually solvable. The following theorem is an analogue of Theorem 4.24. It includes the degenerate case.

**Theorem 4.25** *The Tits alternative holds for the class of colimit groups of non-spherical triangles of groups with the property that every group $G_J$ with $J \subseteq \{1,2,3\}$ and $|J| \leq 2$ either has a non-abelian free subgroup or is virtually solvable.*

Given Theorem 4.19, the proofs of Theorems 4.24 and 4.25 are elementary. But we need an auxiliary result. It had been asked by Button and was answered independently by several authors thereafter, see [But10, Problem 3] for details.

**Lemma 4.26 (Linnell, Minasyan, Klyachko, ...)** *Let $G$ be a group and let $N \trianglelefteq G$ be a normal subgroup. If both $N$ and $G/N$ are virtually solvable, then $G$ is virtually solvable, too.*

*Proof of Theorem 4.24.* Consider a non-degenerate non-spherical triangle of groups with the property that the group $G_\varnothing$ either has a non-abelian free subgroup or is virtually solvable. If $G_\varnothing$, and hence $\mathfrak{G}_\varnothing$, has a non-abelian free subgroup, then the colimit group $\mathfrak{G}$ has so, too. So, we may assume w. l. o. g. that $G_\varnothing$, and hence $\mathfrak{G}_\varnothing$, are virtually solvable. The triangle of groups is non-degenerate. So,

[3] Recall the following two results: Let $G$ be a group and let $N \trianglelefteq G$ be a normal subgroup. $G$ has a non abelian free subgroup if and only if $N$ or $G/N$ does. So, $F \times \Delta(k,l,m)$ does not have a non-abelian free subgroup. On the other hand, if $G$ is virtually solvable, then every subgroup of $G$ is virtually solvable, too. So, $F \times \Delta(k,l,m)$ is not virtually solvable.

for every $a \in \{1,2,3\}$ there is an element $g_a \in \mathfrak{G}_{\{a\}} \setminus \mathfrak{G}_\varnothing$. If the triangle of groups is hyperbolic, then, by Theorem 4.2, the colimit group $\mathfrak{G}$ has a non-abelian free subgroup. So, we may assume w. l. o. g. that the triangle of groups is Euclidean. Now, let $a \in \{1,2,3\}$. Since $g_a \in \mathfrak{G}_{\{a\}} \setminus \mathfrak{G}_\varnothing$, the index of $\mathfrak{G}_\varnothing$ in $\mathfrak{G}_{\{a\}}$ is at least 2. If it is strictly larger than 2, then the simplicial complex $\mathscr{X}$ branches and, by Theorem 4.19, $\mathfrak{G}$ has a non-abelian free subgroup. So, we may assume w. l. o. g. that it is equal to 2. In particular, $\mathfrak{G}_\varnothing$ is normal in $\mathfrak{G}_{\{a\}}$. By the same argument, we may assume w. l. o. g. that for every two distinct $a, b \in \{1,2,3\}$ the group $\mathfrak{G}_{\{a,b\}}$ is generated by $\mathfrak{G}_{\{a\}}$ and $\mathfrak{G}_{\{b\}}$. Therefore, $\mathfrak{G}_\varnothing$ is normal in $\mathfrak{G}_{\{1\}}$, $\mathfrak{G}_{\{2\}}$, $\mathfrak{G}_{\{3\}}$, $\mathfrak{G}_{\{1,2\}}$, $\mathfrak{G}_{\{1,3\}}$, $\mathfrak{G}_{\{2,3\}}$, and, finally, in $\mathfrak{G}$.

Notice that this property also holds in the triangle of groups itself, i. e. for every $J \subseteq \{1,2,3\}$ with $1 \le |J| \le 2$ the image $\varphi_{\varnothing J}(G_\varnothing)$ is normal in $G_J$. For a formal proof, apply the natural homomorphism $\nu_J : G_J \to \mathfrak{G}$, which is injective, and observe:

$$\begin{aligned} \nu_J \circ \varphi_{\varnothing J}(G_\varnothing) &= \nu_\varnothing(G_\varnothing) = \widetilde{\nu}_\varnothing \circ \mu_\varnothing(G_\varnothing) = \widetilde{\nu}_\varnothing(\mathfrak{G}_\varnothing) \\ &\trianglelefteq \widetilde{\nu}_J(\mathfrak{G}_J) = \widetilde{\nu}_J \circ \mu_J(G_J) = \nu_J(G_J) \end{aligned}$$

We may therefore construct the *quotient triangle of groups*, which is obtained by replacing the group $G_\varnothing$ by $G_\varnothing/G_\varnothing \cong \{1\}$ and for every $J \subseteq \{1,2,3\}$ with $1 \le |J| \le 2$ the group $G_J$ by $G_J/\varphi_{\varnothing J}(G_\varnothing)$. Here, one needs to verify that every injective homomorphism $\varphi_{J_1 J_2} : G_{J_1} \to G_{J_2}$ induces an injective homomorphism between the quotients and that the Gersten-Stallings angles remain the same. We leave this work to the reader. However, the resulting diagram is a non-degenerate Euclidean triangle of groups. Moreover, using the presentation $(*)$ one can show that its colimit group is isomorphic to $\mathfrak{G}/\mathfrak{G}_\varnothing$. We will now study the quotient triangle of groups in some more detail. For every $a \in \{1,2,3\}$ the index of $\varphi_{\varnothing\{a\}}(G_\varnothing)$ in $G_{\{a\}}$ is equal to 2, i. e. the quotient $G_{\{a\}}/\varphi_{\varnothing\{a\}}(G_\varnothing)$ has two elements. Therefore, the quotient triangle of groups must be isomorphic to one of the three triangles of groups given in the introduction to Section 4 and, in particular, its colimit group $\mathfrak{G}/\mathfrak{G}_\varnothing$ is (virtually) solvable. Moreover, by assumption, $G_\varnothing$, and hence $\mathfrak{G}_\varnothing$, is virtually solvable, too. So, by Lemma 4.26, we may conclude that $\mathfrak{G}$ is virtually solvable. □

*Proof of Theorem 4.25.* Consider a non-spherical triangle of groups with the property that every group $G_J$ with $J \subseteq \{1,2,3\}$ and $|J| \le 2$ either has a non-abelian free subgroup or is virtually solvable. Again, we may assume w. l. o. g. that the groups $G_J$, and hence their images $\mathfrak{G}_J$, are virtually solvable. Moreover, if the triangle of groups is non-degenerate, then we know by Theorem 4.24 that $\mathfrak{G}$ either has a non-abelian free subgroup or is virtually solvable. So, we may assume w. l. o. g. that the Gersten-Stallings angle $\sphericalangle_{\{2,3\}} = 0$, which means that the homomorphism $\alpha : G_{\{2\}} *_{G_\varnothing} G_{\{3\}} \to G_{\{2,3\}}$ induced by $\varphi_{\{2\}\{2,3\}}$ and $\varphi_{\{3\}\{2,3\}}$ is injective. As mentioned in the proof of Theorem 4.19, one can always show that $\mathfrak{G} \cong X *_A Y$ with $X := \mathfrak{G}_{\{2,3\}}$, $A := \langle \mathfrak{G}_{\{2\}}, \mathfrak{G}_{\{3\}} \rangle \le \mathfrak{G}$, $Y := \langle \mathfrak{G}_{\{1,2\}}, \mathfrak{G}_{\{1,3\}} \rangle \le \mathfrak{G}$. Depending on $|X : A|$ and $|Y : A|$, we distinguish between four cases:

(1) If $|X : A| = 1$, then $\mathfrak{G}_{\{2,3\}}$ is generated by $\mathfrak{G}_{\{2\}}$ and $\mathfrak{G}_{\{3\}}$ or, equivalently, $G_{\{2,3\}}$ is generated by $\varphi_{\{2\}\{2,3\}}(G_{\{2\}})$ and $\varphi_{\{3\}\{2,3\}}(G_{\{3\}})$. So, $\alpha$ is bijective, whence $G_{\{2\}} *_{G_\varnothing} G_{\{3\}} \cong G_{\{2,3\}}$. This allows us to simplify the original presentation $(*)$ of the colimit group $\mathfrak{G}$ by deleting superficial generators and relators:

$$\begin{aligned} \mathfrak{G} = \big\langle G_{\{1\}}, G_{\{1,2\}}, G_{\{1,3\}} : R_{\{1\}}, R_{\{1,2\}}, R_{\{1,3\}}, \\ \big\{ g = \varphi_{\{1\}\{1,2\}}(g) : g \in G_{\{1\}} \big\}, \\ \big\{ g = \varphi_{\{1\}\{1,3\}}(g) : g \in G_{\{1\}} \big\} \big\rangle \end{aligned}$$

So, $\mathfrak{G} \cong G_{\{1,2\}} *_{G_{\{1\}}} G_{\{1,3\}}$ or, equivalently, $\mathfrak{G} \cong \mathfrak{G}_{\{1,2\}} *_{\mathfrak{G}_{\{1\}}} \mathfrak{G}_{\{1,3\}}$. Now, we may, again, distinguish between four cases:

(a) If $|\mathfrak{G}_{\{1,2\}} : \mathfrak{G}_{\{1\}}| = 1$, then $\mathfrak{G} \cong \mathfrak{G}_{\{1,3\}}$, which is virtually solvable.

(b) If $|\mathfrak{G}_{\{1,3\}} : \mathfrak{G}_{\{1\}}| = 1$, then $\mathfrak{G} \cong \mathfrak{G}_{\{1,2\}}$, which is virtually solvable.

(c) If $|\mathfrak{G}_{\{1,2\}} : \mathfrak{G}_{\{1\}}| = |\mathfrak{G}_{\{1,3\}} : \mathfrak{G}_{\{1\}}| = 2$, then $\mathfrak{G}_{\{1\}}$ is normal in $\mathfrak{G}_{\{1,2\}}$, $\mathfrak{G}_{\{1,3\}}$, and $\mathfrak{G}$. The quotient $\mathfrak{G}/\mathfrak{G}_{\{1\}} \cong \mathbb{Z}_2 * \mathbb{Z}_2$, which is (virtually) solvable. On the other hand, $\mathfrak{G}_{\{1\}}$ itself is virtually solvable. So, by Lemma 4.26, the colimit group $\mathfrak{G}$ is virtually solvable.

(d) Otherwise, $|\mathfrak{G}_{\{1,2\}} : \mathfrak{G}_{\{1\}}| \geq 2$ and $|\mathfrak{G}_{\{1,3\}} : \mathfrak{G}_{\{1\}}| \geq 2$ and not both equal to 2. But then, by Remark 4.20, $\mathfrak{G} \cong \mathfrak{G}_{\{1,2\}} *_{\mathfrak{G}_{\{1\}}} \mathfrak{G}_{\{1,3\}}$ has a non-abelian free subgroup.

(2) If $|Y : A| = 1$, then $\mathfrak{G} \cong X = \mathfrak{G}_{\{2,3\}}$, which is virtually solvable.

(3) If $|X : A| = |Y : A| = 2$, then $A$ is normal in $X$, $Y$, and $\mathfrak{G}$. The quotient $\mathfrak{G}/A \cong \mathbb{Z}_2 * \mathbb{Z}_2$, which is (virtually) solvable. Let us now study the normal subgroup $A$ in some more detail. Since the homomorphism $\alpha : G_{\{2\}} *_{G_\varnothing} G_{\{3\}} \to G_{\{2,3\}}$ is injective, its image $\langle \varphi_{\{2\}\{2,3\}}(G_2), \varphi_{\{3\}\{2,3\}}(G_3)\rangle \leq G_{\{2,3\}}$, and hence $A = \langle \mathfrak{G}_{\{2\}}, \mathfrak{G}_{\{3\}}\rangle \leq \mathfrak{G}_{\{2,3\}} \leq \mathfrak{G}$, is isomorphic to $G_{\{2\}} *_{G_\varnothing} G_{\{3\}}$. By a case analysis analogue to the one in (1), we can see that $A$ is either virtually solvable, namely in cases (a)–(c), or has a non-abelian free subgroup, namely in case (d). In the former cases, by Lemma 4.26, the colimit group $\mathfrak{G}$ is virtually solvable. In the latter case, the colimit group $\mathfrak{G}$, which contains $A$ as a subgroup, has a non-abelian free subgroup.

(4) Otherwise, $|X : A| \geq 2$ and $|Y : A| \geq 2$ and not both equal to 2. But then, by Remark 4.20, $\mathfrak{G} \cong X *_A Y$ has a non-abelian free subgroup.

□

**Remark 4.27** *In [KW08, Theorem 1], Kopteva and Williams have proved that the Tits alternative holds for the class of non-spherical Pride groups that are based on graphs with at least four vertices. One way to read Theorem 4.24 is the following: The Tits alternative does not hold for the class of non-spherical Pride groups that are based on graphs with three vertices. But once we assume that each edge is* genuine*, i. e. that none of the Gersten-Stallings angles is equal to 0, it does.*

# A Appendix: Further applications

In Sections 4.1 and 4.2, we worked with non-degenerate Euclidean triangles of groups. The construction can be extended to all non-degenerate non-spherical triangles of groups. In the hyperbolic case, one can either pick a triangle $\Delta$ in the hyperbolic plane $\mathbb{H}^2$, as suggested by Bridson in [Bri91, p. 431, ll. 13–16], or we can pick a triangle $\Delta$ in the Euclidean plane $\mathbb{E}^2$ "whose angles are perhaps a little bit larger than the group-theoretic angles," as suggested by Gersten and Stallings in [Sta91, p. 499, ll. 7–9]. Let us sketch an application for each of the two alternatives.

## A.1 Normal forms

In the billiards theorem, we assume given an element $g \in \mathfrak{G}$ that is adapted to a billiard sequence, i. e. that is equipped with a suitable decomposition into factors from $\mathfrak{G}_{\{1\}} \setminus \mathfrak{G}_\varnothing$, $\mathfrak{G}_{\{2\}} \setminus \mathfrak{G}_\varnothing$, $\mathfrak{G}_{\{3\}} \setminus \mathfrak{G}_\varnothing$. But we could also go the other way and use the simplicial complex $\mathscr{X}$ to construct decompositions. More precisely, given a non-degenerate non-spherical triangle of groups, pick a triangle $\Delta$ either in the Euclidean plane $\mathbb{E}^2$ or in the hyperbolic plane $\mathbb{H}^2$ whose angles agree with the Gersten-Stallings angles and construct the simplicial complex $\mathscr{X}$. Notice that all the results from Sections 4.1 and 4.2 still hold true. Given an arbitrary element $g \in \mathfrak{G}$, consider the unique geodesic in $|\mathscr{X}|$, see Remark 4.16, from the barycentre of $\sigma$ to the barycentre of $g\sigma$. As soon as $g \notin \mathfrak{G}_\varnothing$, the geodesic traverses several 2-simplices. First, it traverses $\sigma$. Then, depending on whether it leaves $\sigma$ crossing a 0-simplex or the interior of a 1-simplex, there is an element $g_1 \in \mathfrak{G}_K$ with $K \subseteq \{1,2,3\}$ and $|K| = 2$ or $|K| = 1$ such that the next 2-simplex it traverses is $g_1\sigma$. This procedure goes on. At the end, it traverses $g_1 g_2 \cdots g_m \sigma = g\sigma$, which yields a decomposition of $g$ into factors from the groups $\mathfrak{G}_K$ with $K \subseteq \{1,2,3\}$ and $1 \leq |K| \leq 2$ and one final factor from $\mathfrak{G}_\varnothing$.

Notice that this decomposition is not well defined, even in the case $\mathfrak{G}_\varnothing = \{1\}$. Just imagine the geodesic running along some 1-simplex. Then, there are many possibilities to choose the respective 2-simplex $g_1 g_2 \cdots g_i \sigma$. On the other hand, if we fix a set of coset representatives for each pair of subgroups $\mathfrak{G}_{K_1} \leq \mathfrak{G}_{K_2}$ with $K_1 \subset K_2 \subseteq \{1,2,3\}$ and $|K_2| \leq 2$, there is a well defined decomposition in terms of these coset representatives and one final factor from $\mathfrak{G}_\varnothing$. Even though it seems to be inconvenient to work with, we may call it a normal form.

## A.2 Euclidean domination

The second alternative has the advantage that there are only three different triangles $\Delta$. More precisely, given a non-degenerate non-spherical triangle of groups, the Gersten-Stallings angles are always of the form $2\pi/\hat{m}$, where $\hat{m}$ is even. Let us think of them as $\pi/k$, $\pi/l$, $\pi/m$ with $k,l,m \in \mathbb{N}$ and $k \leq l \leq m$. It is easy to see that either $(\pi/3,\pi/3,\pi/3)$ or $(\pi/2,\pi/4,\pi/4)$ or $(\pi/2,\pi/3,\pi/6)$ *dominates* $(\pi/k,\pi/l,\pi/m)$, i. e. is coordinatewise at least as large as $(\pi/k,\pi/l,\pi/m)$. If we take the dominating triple instead of the original Gersten-Stallings angles, then, again, all the results from Sections 4.1 and 4.2, in particular the link condition and the billiards theorem, hold true. Therefore, the proof of Theorem 4.19 extends to all non-degenerate non-spherical triangles of groups.

**Remark A.1** *Our methods almost yield an alternative proof of Theorem 4.2; we cannot say anything about the generators $(g_1g_2g_3)^n$ and $(g_1g_3g_2)^n$ but we can prove the existence of another non-abelian free subgroup. If the hyperbolic triangle of groups is non-degenerate and the simplicial complex $\mathscr{X}$ branches, then Theorem 4.19 does the job. The remaining cases are elementary. If it is non-degenerate and the simplicial complex $\mathscr{X}$ does not branch, then the quotient $\mathfrak{G}/\mathfrak{G}_\varnothing$ is a hyperbolic triangle group, has a non-abelian free subgroup, and so has $\mathfrak{G}$. Finally, if it is degenerate, say with $\sphericalangle_{\{2,3\}} = 0$, then $\mathfrak{G}$ contains $\mathfrak{G}_{\{1,2\}} *_{\mathfrak{G}_{\{1\}}} \mathfrak{G}_{\{1,3\}}$ as a subgroup. Since there are elements $g_2 \in \mathfrak{G}_{\{2\}} \setminus \mathfrak{G}_\varnothing$ and $g_3 \in \mathfrak{G}_{\{3\}} \setminus \mathfrak{G}_\varnothing$, the indices $|\mathfrak{G}_{\{1,2\}} : \mathfrak{G}_{\{1\}}|$ and $|\mathfrak{G}_{\{1,3\}} : \mathfrak{G}_{\{1\}}|$ are both at least 2. If one of these indices is equal to 2, the respective Gersten-Stallings angle must be equal to $\pi/2$. But the triangle of groups is hyperbolic, so either $\sphericalangle_{\{1,2\}}$ or $\sphericalangle_{\{1,3\}}$ is smaller than $\pi/2$, which implies that either $|\mathfrak{G}_{\{1,2\}} : \mathfrak{G}_{\{1\}}|$ or $|\mathfrak{G}_{\{1,3\}} : \mathfrak{G}_{\{1\}}|$ is larger than 2. So, by Remark 4.20, $\mathfrak{G}$ has a non-abelian free subgroup.*

# Project C

# Distinguishing graphs with infinite motion and non-linear growth

**( with Wilfried Imrich and Florian Lehner )**

*The following text is a selection from the corresponding publication:*
*Ars Mathematica Contemporanea 7 (2014), no. 1, 201 – 213.*

**Abstract**

The distinguishing number $D(G)$ of a graph $G$ is the least cardinal $d$ such that $G$ has a colouring with $d$ colours which is only preserved by the trivial automorphism. We show that the distinguishing number of an infinite, locally finite, connected graph $G$ with infinite motion and growth $o(n^2/\log_2(n))$ is either 1 or 2, which proves Tom Tucker's infinite motion conjecture for this type of graphs.

**Keywords:** Distinguishing number, automorphisms, infinite graphs.

**MSC classes:** 05C25, 05C63, 05C15.

## 1 Introduction

Albertson and Collins, see [AC96], introduced the distinguishing number $D(G)$ of a graph $G$ as the least cardinal $d$ such that $G$ has a colouring with $d$ colours which is only preserved by the trivial automorphism. This seminal concept spawned many papers on finite and infinite graphs.

In this paper, we are mainly interested in infinite, locally finite, connected graphs of polynomial growth, see also [IJK08], [Tuc11], [STW12]. In particular, there is one conjecture on which we focus our attention, namely Tom Tucker's infinite motion conjecture. Before stating it, we introduce the notation $m(\varphi)$ for the number of elements moved by an automorphism $\varphi$. In other words, $m(\varphi)$ is the size of the support $\operatorname{supp}(\varphi)$. We call $m(\varphi)$ the *motion* of $\varphi$.

**Conjecture 1.1 ("Tom Tucker's infinite motion conjecture")** *Let $G$ be an infinite, locally finite, connected graph. If every non-trivial automorphism of $G$ has infinite motion, then the distinguishing number $D(G)$ is either 1 or 2.*

For the origin of the conjecture and partial results, we refer to [STW12]. Here, we only mention that the conjecture is true if the automorphism group of the graph is countable. In this case, the validity of the conjecture follows from either one of two different results in [ISTW15]. Hence, we may concentrate on graphs with uncountable automorphism group.

The first of the above results is [ISTW15, Corollary 3.8]. It replaces the requirement of infinite motion by a lower and upper bound on the size of the automorphism group. More precisely, it asserts that every infinite, locally finite, connected graph $G$ whose automorphism group is infinite, but strictly smaller than $2^{\aleph_0}$, must have countably infinite automorphism group, infinite motion, and distinguishing number 2. The proof is not easy and follows from results of either Halin [Hal73], Trofimov [Tro85], or Evans [Eva87]. On the other hand, the second result is [ISTW15, Lemma 3.3]. It relaxes the condition of local finiteness and requires only that the automorphism group is countably infinite. The proof is short and elementary.

### Acknowledgements

We thank the two referees for their comments and remarks, as they contributed considerably to the readability of the paper. Furthermore, we are grateful to Norbert Sauer and Claude Laflamme for their suggestions.

## 2 Preliminaries

Throughout this paper the symbol $\mathbb{N}$ denotes the set $\{1,2,3,...\}$ of positive integers, whereas $\mathbb{N}_0$ denotes the set $\{0,1,2,3,...\}$ of non-negative integers.

### 2.1 Distinguishing number

Let $G$ be a graph with vertex set $V(G)$, and let $C$ be a set of colours. A *$C$-colouring* $\chi$ of $G$ is a function $\chi : V(G) \to C$. For us, the set of colours $C$ will mostly be the set $\{\text{black}, \text{white}\}$, in which case we speak of a *two-colouring* of $G$.

Assume given a $C$-colouring $\chi$ of $G$. Let $\varphi \in \mathrm{Aut}(G)$. If, for every $v \in V(G)$, the equation $\chi(\varphi(v)) = \chi(v)$ holds, we say that $\chi$ *is preserved by* $\varphi$. Otherwise, we say that $\chi$ *breaks* $\varphi$. A $C$-colouring $\chi$ of $G$ is called *distinguishing* if it is only preserved by the trivial automorphism. The *distinguishing number* $D(G)$ is the least cardinal $d$ such that there is a distinguishing $C$-colouring of $G$ with $d$ colours, i. e. $|C| = d$.

### 2.2 Group action on a graph

Given a group $A$ equipped with a homomorphism $\varphi : A \to \mathrm{Aut}(G)$, we say that $A$ *acts on* $G$. Moreover, we say that $A$ *acts non-trivially* on $G$ if there is an $a \in A$ such that $\varphi(a)$ moves at least one vertex of $G$. By abuse of language, we write $a(v)$ instead of $\varphi(a)(v)$ and say that a $C$-colouring $\chi$ of $G$ *is preserved by* $a \in A$ if it is preserved by $\varphi(a) \in \mathrm{Aut}(G)$.

### 2.3 Balls and spheres

The *ball with centre* $v_0 \in V(G)$ *and radius* $r$ is the set of all vertices $v \in V(G)$ with $d_G(v_0, v) \le r$, it is denoted by $B^G_{v_0}(r)$. On the other hand, $S^G_{v_0}(r)$ stands for the set of all vertices $v \in V(G)$ with $d_G(v_0, v) = r$. It is called the *sphere with centre* $v_0 \in V(G)$ *and radius* $r$. If $G$ is clear from the context, we just write $B_{v_0}(r)$ and $S_{v_0}(r)$ respectively. For terms not defined in Sections 2.1 – 2.3, we refer to [HIK11].

### 2.4 Growth rate of a graph

Although our graphs are infinite, as long as they are locally finite, all balls and spheres of finite radius are finite. The number of vertices in $B^G_{v_0}(r)$ is a monotonically increasing function of $r$. Indeed, notice that

$$\left|B^G_{v_0}(r)\right| = \sum_{i=0}^{r} \left|S^G_{v_0}(i)\right| \quad \text{and} \quad \left|S^G_{v_0}(i)\right| \geq 1.$$

Nonetheless, the precise growth function depends very much on the choice of $v_0 \in V(G)$, and it is helpful to define the growth rate of $G$. We say that an infinite, locally finite, connected graph $G$ has *polynomial growth* if there is a vertex $v_0 \in V(G)$ and a polynomial $p$ such that

$$\forall\, r \in \mathbb{N}_0 : \left|B^G_{v_0}(r)\right| \leq p(r).$$

It is easy to see that this implies that all growth functions $r \mapsto \left|B^G_v(r)\right|$ are bounded by polynomials of the same degree as $p$, independent of the choice of $v \in V(G)$. In this context, it should be clear what we mean by *linear* and *quadratic growth*. For example, observe that the two-sided infinite path has linear growth, and that the growth of the grid of integers in the plane is quadratic. We say that $G$ has *exponential growth* if there is a constant $c > 1$ such that

$$\forall\, r \in \mathbb{N}_0 : \left|B^G_{v_0}(r)\right| \geq c^r.$$

Homogeneous trees of degree $d > 2$, i. e. infinite trees in which every vertex has the same degree $d$, have exponential growth. For the distinguishability of such trees and tree-like graphs, see [WZ07] and [IKT07].

### 2.5 Toolbox

In this paper, we are mainly interested is the distinguishability of infinite, locally finite, connected graphs of polynomial growth. Here, we provide a toolbox that helps us to break automorphisms.

**Lemma 2.1** *Let $A$ be a finite group acting on a graph $G$. If a colouring $\chi : V(G) \to C$ breaks some element of $A$, then it breaks at least half of the elements of $A$.*

*Proof.* The elements of $A$ that preserve a colouring $\chi$ form a subgroup. If some element of $A$ is broken by a colouring $\chi$, then this subgroup is proper. Thus, by Lagrange's theorem, it cannot contain more than half of the elements of $A$. □

**Corollary 2.2** *Let $A$ be a finite group acting non-trivially on a graph $G$. Then, there is a two-colouring of $G$ that breaks at least half of the elements of $A$.*

The proof of Lemma 2.1 is based on the fact that $A$ is a group. But a very similar result holds for any finite family of non-trivial automorphisms, as the following lemma shows.

**Lemma 2.3** *Let $G$ be a finite graph. If $A$ is a finite set equipped with a function $\varphi : A \to \mathrm{Aut}(G) \smallsetminus \{\mathrm{id}\}$, then there is a two-colouring of $G$ that breaks $\varphi(a)$ for at least half of the elements of $A$.*

*Proof.* Let $V(G) = \{v_1, v_2, \dots, v_n\}$. For every $k \in \{1, 2, \dots, n\}$, let $A_k$ be the set of all elements $a \in A$ with $\mathrm{supp}(\varphi(a)) \subseteq \{v_1, v_2, \dots, v_k\}$. We show by induction that the assertion holds for all $A_k$ and, in particular, for $A$. Because $A_1$ is the empty set, the assertion is true for $A_1$. Suppose it is true for $A_{k-1}$. Then, we can choose a two-colouring of $G$ that breaks $\varphi(a)$ for at least half of the elements of $A_{k-1}$. This remains true, even when we change the colour of $v_k$. Notice that, for every $a \in A_k \setminus A_{k-1}$, $\varphi(a)$ either maps $v_k$ into a white vertex in $\{v_1, v_2, \dots, v_{k-1}\}$ or into a black vertex in $\{v_1, v_2, \dots, v_{k-1}\}$. Depending on which of the two alternatives occurs more often, we colour $v_k$ black or white such that this two-colouring also breaks $\varphi(a)$ for at least half of the elements of $A_k \setminus A_{k-1}$ and, hence, for at least half of the elements of $A_k$. □

### 2.6 Graphs with infinite motion

If every non-trivial automorphism of a graph $G$ has infinite motion, we say that $G$ has *infinite motion*. For graphs with infinite motion, the following lemma is easy to see. Nevertheless, a rigorous proof can be found in [Leh14, Corollary 4.8].

**Lemma 2.4** *Let $G$ be an infinite, locally finite, connected graph with infinite motion. If an automorphism $\varphi \in \mathrm{Aut}(G)$ fixes a vertex $v_0 \in V(G)$ and moves at least one vertex in $S_{v_0}(k)$, then, for every $i \geq k$, it moves at least one vertex in $S_{v_0}(i)$.*

## 3 Main result about graphs with non-linear growth

Let us mention that infinite, locally finite, connected graphs with infinite motion and linear growth have countable automorphism group, see e. g. [Leh14, Proof of Theorem 4.5], and therefore distinguishing number either 1 or 2. But, if the growth rate of such graphs becomes non-linear, then the automorphism group can become uncountable. This holds, even if the growth rate becomes only slightly non-linear.

**Example 3.1 ("Stretched tree")** *Let $\varepsilon > 0$. Then, there is an infinite, locally finite, connected graph $G$ with uncountable automorphism group, infinite motion, and non-linear growth function $g : \mathbb{N}_0 \to \mathbb{N}_0$ such that, for sufficiently large $n \in \mathbb{N}_0$, $g(n)$ is bounded from above by $n^{1+\varepsilon}$.*

*Proof.* We construct $G$ from $T_3$, i. e. the tree in which every vertex has degree 3. First, choose an arbitrary vertex $v_0 \in V(T_3)$. Our strategy is to replace the edges of $T_3$ by paths such that, for sufficiently large $n \in \mathbb{N}_0$, $g(n) = |B^G_{v_0}(n)| \leq n^{1+\varepsilon}$. For every $i \in \mathbb{N}_0$, there are $3 \cdot 2^i$ edges from $S^{T_3}_{v_0}(i)$ to $S^{T_3}_{v_0}(i+1)$. If we replace them by paths of the same length, then the cardinality of the balls $B^G_{v_0}(n)$ grows linearly with slope $3 \cdot 2^i$ from $S^{T_3}_{v_0}(i)$ to $S^{T_3}_{v_0}(i+1)$.

Observe that, given any affine linear function $h : \mathbb{N}_0 \to \mathbb{N}_0$, there is a number $n_h \in \mathbb{N}$ such that, for all $n \geq n_h$, $h(n) \leq n^{1+\varepsilon}$. In particular, we may consider the functions $h_i : \mathbb{N}_0 \to \mathbb{N}_0$ defined by $h_i(x) = 3 \cdot 2^i \cdot x + 1$, and choose numbers $n_i \in \mathbb{N}$ such that, for every $n \geq n_i$, $h_i(n) \leq n^{1+\varepsilon}$.

As illustrated in Figure 1, for every $i \in \mathbb{N}_0$, we replace the edges from $S^{T_3}_{v_0}(i)$ to $S^{T_3}_{v_0}(i+1)$ by paths of length $n_{i+1}$. For every $i \in \mathbb{N}$ and every vertex $v \in V(G)$ on such a path from $S^{T_3}_{v_0}(i)$ to $S^{T_3}_{v_0}(i+1)$, we have $d_G(v, v_0) \geq n_i$ and, hence, $g(d_G(v, v_0)) \leq 3 \cdot 2^i \cdot d_G(v, v_0) + 1 = h_i(d_G(v, v_0)) \leq d_G(v, v_0)^{1+\varepsilon}$. So, for every $n \geq n_1$, $g(n)$ is bounded from above by $n^{1+\varepsilon}$.

Every automorphism of $T_3$ that fixes $v_0$ induces an automorphism of $G$. It is easy to see that this correspondence is bijective. Thus, $\mathrm{Aut}(G)$ is uncountable. Furthermore, $G$ inherits infinite motion from $T_3$. Since $\mathrm{Aut}(G)$ is uncountable, the result mentioned at the beginning of Section 3 implies that $G$ cannot have linear growth. □

Though we cannot assume that the automorphism groups of our graphs are countable, we prove that infinite, locally finite, connected graphs with infinite motion and non-linear, but moderate, growth are still *two-distinguishable*, i. e. they have distinguishing number either 1 or 2.

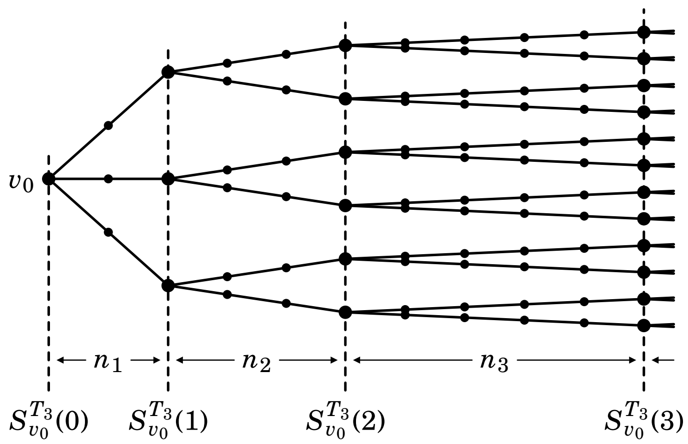


Figure 1: Replacing the edges of $T_3$ by paths.

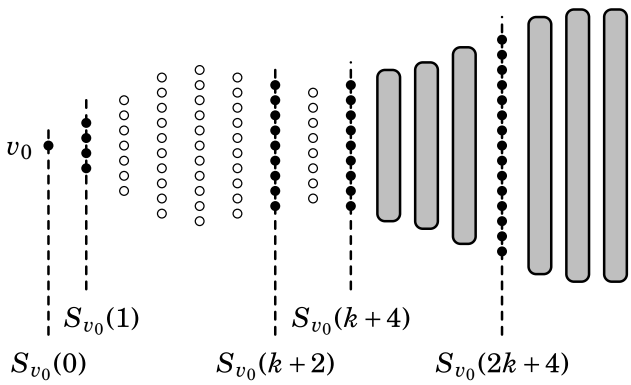


Figure 2: Breaking all automorphisms that move $v_0$.

Our construction of a suitable colouring consists of several steps. In Lemma 3.2, we colour a part of the vertices in order to break all automorphisms that move a distinguished vertex $v_0$. In Lemma 3.3, we show how to colour some of the remaining vertices in order to break more automorphisms. Iteration of this procedure yields a distinguishing colouring, as shown in Theorem 3.4.

**Lemma 3.2** *Let $G$ be an infinite, locally finite, connected graph with infinite motion and $v_0 \in V(G)$. Then, for every $k \in \mathbb{N}$, one can two-colour all vertices in $B_{v_0}(k+3)$ and $S_{v_0}(\lambda k+4)$, $\lambda \in \mathbb{N}$, such that, no matter how one colours the remaining vertices, all automorphisms that move $v_0$ are broken.*

*Proof.* If $k=1$, then we colour $v_0$ black and all $v \in V(G) \setminus \{v_0\}$ white, whence all automorphisms that move $v_0$ are broken. So, let $k \geq 2$. First, we colour all vertices in $S_{v_0}(0)$, $S_{v_0}(1)$, and $S_{v_0}(k+2)$ black and the remaining vertices in $B_{v_0}(k+3)$ white. Moreover, we colour all vertices in $S_{v_0}(\lambda k+4)$, $\lambda \in \mathbb{N}$, black and claim that, no matter how we colour the remaining vertices, $v_0$ is the only black vertex that has only black neighbours and only white vertices at distance $r \in \{2,3,\dots,k+1\}$, see Figure 2.

It clearly follows from this claim that this colouring breaks every automorphism that moves $v_0$. So, it remains to verify the claim. Consider a vertex $v \in V(G) \setminus \{v_0\}$. If $v$ is not in $S_{v_0}(1)$, then it is easy to see that $v$ cannot have the aforementioned properties. So, let $v$ be in $S_{v_0}(1)$ and assume it has only black neighbours and only white vertices at distance 2. Then, it cannot be neighbour to any vertex in

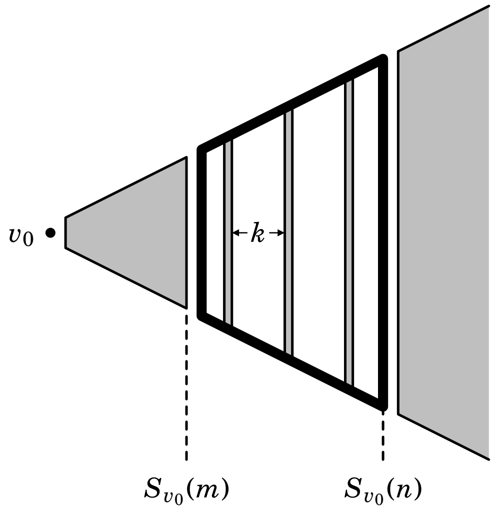


Figure 3: Breaking all automorphisms that fix $v_0$ and act non-trivially on $B_{v_0}(m)$.

$S_{v_0}(2)$, but must be neighbour to all vertices in $B_{v_0}(1)$ except itself. Therefore, the transposition of the vertices $v$ and $v_0$ is a non-trivial automorphism of $G$ with finite support. Since $G$ has infinite motion, this is not possible. $\square$

**Lemma 3.3** *Let $G$ be an infinite, locally finite, connected graph with infinite motion and $v_0 \in V(G)$. Moreover, let $\varepsilon > 0$. Then, there is a $k \in \mathbb{N}$ such that, for every $m \in \mathbb{N}$ and every $n \in \mathbb{N}$ which is sufficiently large and fulfils*

$$\left|S_{v_0}(n)\right| \leq \frac{n}{(1+\varepsilon)\log_2(n)}, \qquad (*)$$

*one can two-colour all vertices in $S_{v_0}(m+1), S_{v_0}(m+2), \ldots, S_{v_0}(n)$, but not those in $S_{v_0}(\lambda k+4)$, $\lambda \in \mathbb{N}$, such that all automorphisms that fix $v_0$ and act non-trivially on $B_{v_0}(m)$ are broken.*

*Proof.* Notice that the meaning of the variables $m$, $n$, and $k$ is illustrated in Figure 3. First, choose a $k \in \mathbb{N}$ that is larger than $1 + \frac{1}{\varepsilon}$. Then

$$\frac{k-1}{k} > \frac{1}{1+\varepsilon}. \qquad (1)$$

Next, let $m \in \mathbb{N}$. By (1), there is an $n_0 \in \mathbb{N}$ such that

$$\forall\, n \geq n_0 : (n-m) \cdot \frac{k-1}{k} \geq n \cdot \frac{1}{1+\varepsilon} + 1. \qquad (2)$$

Now, let $n \in \mathbb{N}$ be sufficiently large, i. e. $n \geq n_0$, and assume it fulfils ($*$). Then, by (2), the number of spheres $S_{v_0}(m+1), S_{v_0}(m+2), \ldots, S_{v_0}(n)$ that are not of type $S_{v_0}(\lambda k+4)$, $\lambda \in \mathbb{N}$, is at least

$$\left\lfloor (n-m) \cdot \frac{k-1}{k} \right\rfloor \geq \left\lfloor n \cdot \frac{1}{1+\varepsilon} + 1 \right\rfloor > \frac{n}{1+\varepsilon}. \qquad (3)$$

Our goal is to two-colour the vertices in these spheres in order to break all automorphisms that fix $v_0$ and act non-trivially on $B_{v_0}(m)$. Let $\mathrm{Aut}(G, v_0)$ be the group of all automorphisms that fix $v_0$. Every $\varphi \in \mathrm{Aut}(G, v_0)$ induces a permutation $\varphi | B_{v_0}(n)$ of the vertices in $B_{v_0}(n)$. These permutations form a group $A$. If $\sigma$ and $\tau$ are different elements of $A$, then $\sigma\tau^{-1} \in A$ acts non-trivially on $B_{v_0}(n)$. By

Lemma 2.4, it also does so on $S_{v_0}(n)$, which means that $\sigma$ and $\tau$ do not agree on $S_{v_0}(n)$. Therefore, the cardinality of $A$ is at most $|S_{v_0}(n)|!$, for which the following rough estimate suffices for our purposes:

$$\begin{aligned} |S_{v_0}(n)|! \leq |S_{v_0}(n)|^{|S_{v_0}(n)|-1} &\leq \left(\frac{n}{(1+\varepsilon)\log_2(n)}\right)^{\frac{n}{(1+\varepsilon)\log_2(n)}-1} \\ &\leq n^{\frac{n}{(1+\varepsilon)\log_2(n)}-1} = 2^{\left(\frac{n}{(1+\varepsilon)\log_2(n)}-1\right)\log_2(n)} \leq 2^{\frac{n}{1+\varepsilon}-1}. \end{aligned} \tag{4}$$

It is clear that, if an element $\sigma \in A$ that acts non-trivially on $B_{v_0}(m)$ is broken by a suitable two-colouring of some spheres in $B_{v_0}(n)$, then all $\varphi \in \operatorname{Aut}(G, v_0)$ with $\varphi|B_{v_0}(n) = \sigma$ are broken at once. So, it suffices to break all $\sigma \in A$ which act non-trivially on $B_{v_0}(m)$ by a suitable two-colouring of some spheres in $B_{v_0}(n)$ in order to ensure that all $\varphi \in \operatorname{Aut}(G, v_0)$ which act non-trivially on $B_{v_0}(m)$ are broken.

Before doing this, let us remark that any element $\sigma \in A$ which acts non-trivially on the ball $B_{v_0}(m)$, also acts non-trivially on every sphere $S_{v_0}(m+1), \ldots, S_{v_0}(n)$. This is a consequence of Lemma 2.4, and implies that we can break $\sigma$ by breaking the action of $\sigma$ on any one of the spheres $S_{v_0}(m+1), \ldots, S_{v_0}(n)$.

Now, consider the subset $S \subseteq A$ of all elements that act non-trivially on $B_{v_0}(m)$. As already remarked, every $\sigma \in S$ acts non-trivially on every sphere $S_{v_0}(m+1), \ldots, S_{v_0}(n)$. Hence, we can apply Lemma 2.3 to break at least half of the elements of $S$ by a suitable colouring of $S_{v_0}(m+1)$. What remains unbroken is a subset $S' \subseteq S$ of cardinality at most $|S|/2$. Now, we proceed to the next sphere. We can break at least half of the elements of $S'$ by a suitable colouring of $S_{v_0}(m+2)$. What still remains unbroken, is a subset $S'' \subseteq S$ of cardinality at most $|S|/4$.

Iterating the procedure, but avoiding spheres of type $S_{v_0}(\lambda k + 4)$, $\lambda \in \mathbb{N}$, we end up with the empty subset $\varnothing \subseteq S$ after at most $\log_2(|S|) + 1 \leq \log_2(|A|) + 1 \leq \frac{n}{1+\varepsilon}$ steps, see (4). This is less than the number of spheres not of type $S_{v_0}(\lambda k + 4)$, $\lambda \in \mathbb{N}$, between $S_{v_0}(m+1)$ and $S_{v_0}(n)$, see (3). Thus, we remain within the ball $B_{v_0}(n)$. Hence, all $s \in S$ and, therefore, all $\varphi \in \operatorname{Aut}(G, v_0)$ that act non-trivially on $B_{v_0}(m)$ are broken, and we are done. $\square$

**Theorem 3.4** *Let $G$ be an infinite, locally finite, connected graph with infinite motion and $v_0 \in V(G)$. Moreover, let $\varepsilon > 0$. If there are infinitely many $n \in \mathbb{N}$ such that*

$$|S_{v_0}(n)| \leq \frac{n}{(1+\varepsilon)\log_2(n)}, \tag{$**$}$$

*then the distinguishing number $D(G)$ is either 1 or 2.*

*Proof.* Consider the number $k \in \mathbb{N}$ provided by Lemma 3.3. First, we use Lemma 3.2 to two-colour all vertices in $B_{v_0}(k+3)$ and in $S_{v_0}(\lambda k+4)$, $\lambda \in \mathbb{N}$, such that, no matter how we colour the remaining vertices, all automorphisms that move $v_0$ are broken.

Let $m_1 = k+3$. Among all $n \in \mathbb{N}$ that satisfy ($**$) we choose a number $n_1 \in \mathbb{N}$ that is larger than $m_1$ and sufficiently large to apply Lemma 3.3. Hence, we can two-colour all vertices in $S_{v_0}(m_1+1)$, $S_{v_0}(m_1+2), \ldots, S_{v_0}(n_1)$, except those in $S_{v_0}(\lambda k+4)$, $\lambda \in \mathbb{N}$, such that all automorphisms that fix $v_0$ and act non-trivially on $B_{v_0}(m_1)$ are broken. Next, let $m_2 = n_1$ and choose an $n_2 \in \mathbb{N}$ to apply Lemma 3.3 again. Iteration of this procedure yields a two-colouring of $G$.

If an automorphism $\varphi \in \operatorname{Aut}(G) \setminus \{\mathrm{id}\}$ moves $v_0$, then it is broken by our colouring. If it fixes $v_0$, consider a vertex $v$ with $\varphi(v) \neq v$. Since $G$ is connected and $m_1 < m_2 < m_3 < \ldots$, there is an $i \in \mathbb{N}$ such that $v$ is contained in $B_{v_0}(m_i)$. Hence, $\varphi$ acts non-trivially on $B_{v_0}(m_i)$ and is again broken by our colouring. $\square$

**Corollary 3.5** *Let $G$ be an infinite, locally finite, connected graph with infinite motion and $v_0 \in V(G)$. Moreover, let $\varepsilon > 0$. If there are infinitely many $n \in \mathbb{N}$ such that*

$$\left|B_{v_0}(n)\right| \leq \frac{n^2}{(2+\varepsilon)\log_2(n)}, \tag{$***$}$$

*then the distinguishing number $D(G)$ is either 1 or 2. In particular, Tom Tucker's infinite motion conjecture holds for all graphs of growth $o(n^2/\log_2(n))$.*

*Proof.* Let $n_1 < n_2 < n_3 < \ldots$ be an infinite sequence of numbers that fulfil ($***$). Notice that, for every $k \in \mathbb{N}$,

$$\sum_{i=1}^{n_k} \frac{i}{(1+\frac{\varepsilon}{2})\log_2(i)} > \frac{{n_k}^2}{(2+\varepsilon)\log_2(n_k)} \geq \left|B_{v_0}(n_k)\right| > \sum_{i=1}^{n_k} \left|S_{v_0}(i)\right|.$$

Since

$$\lim_{k\to\infty} \left( \left( \sum_{i=1}^{n_k} \frac{i}{(1+\frac{\varepsilon}{2})\log_2(i)} \right) - \frac{{n_k}^2}{(2+\varepsilon)\log_2(n_k)} \right) = \infty,$$

we infer that

$$\lim_{k\to\infty} \sum_{i=1}^{n_k} \left( \frac{i}{(1+\frac{\varepsilon}{2})\log_2(i)} - \left|S_{v_0}(i)\right| \right) = \infty,$$

and that, for infinitely many $i \in \mathbb{N}$,

$$\left|S_{v_0}(i)\right| < \frac{i}{(1+\frac{\varepsilon}{2})\log_2(i)}.$$

Hence, we can apply Theorem 3.4 to show that the distinguishing number $D(G)$ is either 1 or 2. $\square$

# Bibliography


[AC96] M. O. Albertson and K. L. Collins, *Symmetry breaking in graphs*, Electron. J. Combin. **3** (1996), no. 1, Research Paper 18, 17 pp. (electronic).

[AHS90] J. Adámek, H. Herrlich, and G. E. Strecker, *Abstract and concrete categories*, Pure and Applied Mathematics (New York), John Wiley & Sons, Inc., New York, 1990, The joy of cats, A Wiley-Interscience Publication.

[All12] D. Allcock, *Triangles of Baumslag-Solitar groups*, Canad. J. Math. **64** (2012), no. 2, 241–253.

[Bea83] A. F. Beardon, *The geometry of discrete groups*, Graduate Texts in Mathematics, vol. 91, Springer-Verlag, New York, 1983.

[Beh75] H. Behr, *Explizite Präsentation von Chevalleygruppen über $\mathbb{Z}$ [Explicit presentations for Chevalley groups over $\mathbb{Z}$]*, Math. Z. **141** (1975), 235–241.

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

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

[Bla55] D. Blackwell, *On transient Markov processes with a countable number of states and stationary transition probabilities*, Ann. Math. Statist. **26** (1955), 654–658.

[BP93] W. A. Bogley and S. J. Pride, *Calculating generators of* $\Pi_2$, Two-dimensional homotopy and combinatorial group theory, London Math. Soc. Lecture Note Ser., vol. 197, Cambridge Univ. Press, Cambridge, 1993, pp. 157–188.

[Bre04] D. Brendel, *Nichtsphärische und sphärische Dreiecke nilpotenter Gruppen [non-spherical and spherical triangles of nilpotent groups]*, Diplomarbeit, Goethe-Universität Frankfurt am Main, 2004.

[Bri63] J. L. Britton, *The word problem*, Ann. of Math. (2) **77** (1963), 16–32.

[Bri91] M. R. Bridson, *Geodesics and curvature in metric simplicial complexes*, Group theory from a geometrical viewpoint (Trieste, 1990), World Sci. Publ., River Edge, NJ, 1991, pp. 373–463.

[Bro03] S. Brofferio, *How a centred random walk on the affine group goes to infinity*, Ann. Inst. H. Poincaré Probab. Statist. **39** (2003), no. 3, 371–384.

[Bro06] ———, *The Poisson boundary of random rational affinities*, Ann. Inst. Fourier (Grenoble) **56** (2006), no. 2, 499–515.

[BS62] G. Baumslag and D. Solitar, *Some two-generator one-relator non-Hopfian groups*, Bull. Amer. Math. Soc. **68** (1962), 199–201.

[BS85] M. G. Brin and C. C. Squier, *Groups of piecewise linear homeomorphisms of the real line*, Invent. Math. **79** (1985), no. 3, 485–498.

[BSCSW] A. Bendikov, L. Saloff-Coste, M. Salvatori, and W. Woess, *Brownian motion on treebolic space: escape to infinity*, Rev. Mat. Iberoam., to appear.

[But10] J. O. Button, *Largeness of LERF and 1-relator groups*, Groups Geom. Dyn. **4** (2010), no. 4, 709–738.

[CD60] G. Choquet and J. Deny, *Sur l'équation de convolution $\mu = \mu * \sigma$*, C. R. Acad. Sci. Paris **250** (1960), 799–801.

[CF51] K. L. Chung and W. Fuchs, *On the distribution of values of sums of random variables*, Mem. Amer. Math. Soc. **1951** (1951), no. 6, 12.

[CFP96] J. W. Cannon, W. J. Floyd, and W. R. Parry, *Introductory notes on Richard Thompson's groups*, Enseign. Math. (2) **42** (1996), no. 3-4, 215–256.

[Che95] A. Chermak, *Triangles of groups*, Trans. Amer. Math. Soc. **347** (1995), no. 11, 4533–4558.

[CK12] V. Climenhaga and A. Katok, *Measure theory through dynamical eyes*, 2012, arXiv:1208.4550.

[CKW94] D. I. Cartwright, V. A. Kaĭmanovich, and W. Woess, *Random walks on the affine group of local fields and of homogeneous trees*, Ann. Inst. Fourier (Grenoble) **44** (1994), no. 4, 1243–1288.

[Cor96] J. M. Corson, *Amalgamated sums of groups*, Proc. Edinburgh Math. Soc. (2) **39** (1996), no. 3, 561–570.

[Cun11] J. Cuno, *Nichtsphärische Dreiecke von Gruppen [Non-spherical triangles of groups]*, Diplomarbeit, Goethe-Universität Frankfurt am Main, 2011.

[Éli82] L. Élie, *Comportement asymptotique du noyau potentiel sur les groupes de Lie*, Ann. Sci. École Norm. Sup. (4) **15** (1982), no. 2, 257–364.

[Ers04] A. Erschler, *Boundary behavior for groups of subexponential growth*, Ann. of Math. (2) **160** (2004), no. 3, 1183–1210.

[Ers10] ———, *Poisson-Furstenberg boundaries, large-scale geometry and growth of groups*, Proceedings of the International Congress of Mathematicians. Volume II, Hindustan Book Agency, New Delhi, 2010, pp. 681–704.

[ERST00] M. Edjvet, G. Rosenberger, M. Stille, and R. M. Thomas, *On certain finite generalized tetrahedron groups*, Computational and geometric aspects of modern algebra (Edinburgh, 1998), London Math. Soc. Lecture Note Ser., vol. 275, Cambridge Univ. Press, Cambridge, 2000, pp. 54–65.

[Eva87] D. M. Evans, *A note on automorphism groups of countably infinite structures*, Arch. Math. (Basel) **49** (1987), no. 6, 479–483.

[Fel71] W. Feller, *An introduction to probability theory and its applications. Vol. II.*, Second edition, John Wiley & Sons, Inc., New York-London-Sydney, 1971.

[Fü73] H. Fürstenberg, *Boundary theory and stochastic processes on homogeneous spaces*, Harmonic analysis on homogeneous spaces (Proc. Sympos. Pure Math., Vol. XXVI, Williams Coll., Williamstown, Mass., 1972), Amer. Math. Soc., Providence, R.I., 1973, pp. 193–229.

[Hae73] J. Haezendonck, *Abstract Lebesgue-Rohlin spaces*, Bull. Soc. Math. Belg. **25** (1973), 243–258.

[Hal73] R. Halin, *Automorphisms and endomorphisms of infinite locally finite graphs*, Abh. Math. Sem. Univ. Hamburg **39** (1973), 251–283.

[HIK11] R. Hammack, W. Imrich, and S. Klavžar, *Handbook of product graphs*, second ed., Discrete Mathematics and its Applications (Boca Raton), CRC Press, Boca Raton, FL, 2011, With a foreword by Peter Winkler.

[HK06] J. Howie and N. Kopteva, *The Tits alternative for generalized tetrahedron groups*, J. Group Theory **9** (2006), no. 2, 173–189.

[IJK08] W. Imrich, J. Jerebic, and S. Klavžar, *The distinguishing number of Cartesian products of complete graphs*, European J. Combin. **29** (2008), no. 4, 922–929.

[IKT07] W. Imrich, S. Klavžar, and V. Trofimov, *Distinguishing infinite graphs*, Electron. J. Combin. **14** (2007), no. 1, Research Paper 36, 12 pp. (electronic).

[ISTW15] W. Imrich, S. M. Smith, T. W. Tucker, and M. E. Watkins, *Infinite motion and 2-distinguishability of graphs and groups*, J. Algebraic Combin. **41** (2015), no. 1, 109–122.

[Joh80] D. L. Johnson, *Topics in the theory of group presentations*, London Mathematical Society Lecture Note Series, vol. 42, Cambridge University Press, Cambridge-New York, 1980.

[Kaĭ91] V. A. Kaĭmanovich, *Poisson boundaries of random walks on discrete solvable groups*, Probability measures on groups, X (Oberwolfach, 1990), Plenum, New York, 1991, pp. 205–238.

[Kaĭ00] ———, *The Poisson formula for groups with hyperbolic properties*, Ann. of Math. (2) **152** (2000), no. 3, 659–692.

[KKR04] V. A. Kaĭmanovich, Y. Kifer, and B.-Z. Rubshtein, *Boundaries and harmonic functions for random walks with random transition probabilities*, J. Theoret. Probab. **17** (2004), no. 3, 605–646.

[Kle14] A. Klenke, *Probability theory*, second ed., Universitext, Springer, London, 2014, A comprehensive course.

[KV83] V. A. Kaĭmanovich and A. M. Vershik, *Random walks on discrete groups: boundary and entropy*, Ann. Probab. **11** (1983), no. 3, 457–490.

[KW08] N. Kopteva and G. Williams, *The Tits alternative for non-spherical Pride groups*, Bull. Lond. Math. Soc. **40** (2008), no. 1, 57–64.

[Leh14] F. Lehner, *Symmetry breaking in graphs and groups*, Dissertation, Technische Universität Graz, 2014.

[Lor95] S. Lorenz, *Freie Untergruppen in Dreiecken von Gruppen [Free subgroups in triangles of groups]*, Diplomarbeit, Goethe-Universität Frankfurt am Main, 1995.

[Mei08] J. Meier, *Groups, graphs and trees*, London Mathematical Society Student Texts, vol. 73, Cambridge University Press, Cambridge, 2008, An introduction to the geometry of infinite groups.

[Mes72] S. Meskin, *Nonresidually finite one-relator groups*, Trans. Amer. Math. Soc. **164** (1972), 105–114.

[Mil68] C. F. Miller, III, *On Britton's theorem* A, Proc. Amer. Math. Soc. **19** (1968), 1151–1154.

[Mun84] J. R. Munkres, *Elements of algebraic topology*, Addison-Wesley Publishing Company, Menlo Park, CA, 1984.

[Neu54] B. H. Neumann, *An essay on free products of groups with amalgamations*, Philos. Trans. Roy. Soc. London. Ser. A. **246** (1954), 503–554.

[Pól21] G. Pólya, *Über eine Aufgabe der Wahrscheinlichkeitsrechnung betreffend die Irrfahrt im Straßennetz*, Math. Ann. **84** (1921), no. 1-2, 149–160.

[Pri87] S. J. Pride, *Groups with presentations in which each defining relator involves exactly two generators*, J. London Math. Soc. (2) **36** (1987), no. 2, 245–256.

[Pri92] ———, *The (co)homology of groups given by presentations in which each defining relator involves at most two types of generators*, J. Austral. Math. Soc. Ser. A **52** (1992), no. 2, 205–218.

[Roh52] V. A. Rohlin, *On the fundamental ideas of measure theory*, Amer. Math. Soc. Translation **1952** (1952), no. 71, 55.

[Ros81] J. Rosenblatt, *Ergodic and mixing random walks on locally compact groups*, Math. Ann. **257** (1981), no. 1, 31–42.

[Rot95] J. J. Rotman, *An introduction to the theory of groups*, fourth ed., Graduate Texts in Mathematics, vol. 148, Springer-Verlag, New York, 1995.

[Rud90] D. J. Rudolph, *Fundamentals of measurable dynamics*, Oxford Science Publications, The Clarendon Press, Oxford University Press, New York, 1990, Ergodic theory on Lebesgue spaces.

[Sta91] J. R. Stallings, *Non-positively curved triangles of groups*, Group theory from a geometrical viewpoint (Trieste, 1990), World Sci. Publ., River Edge, NJ, 1991, pp. 491–503.

[STW12] S. M. Smith, T. W. Tucker, and M. E. Watkins, *Distinguishability of infinite groups and graphs*, Electron. J. Combin. **19** (2012), no. 2, Paper 27, 10 pp. (electronic).

[SW05] M. Sageev and D. T. Wise, *The Tits alternative for* CAT(0) *cubical complexes*, Bull. London Math. Soc. **37** (2005), no. 5, 706–710.

[Tit72] J. Tits, *Free subgroups in linear groups*, J. Algebra **20** (1972), 250–270.

[Tit86] ———, *Buildings and group amalgamations*, Proceedings of groups—St. Andrews 1985, London Math. Soc. Lecture Note Ser., vol. 121, Cambridge Univ. Press, Cambridge, 1986, pp. 110–127.

[Tro85] V. I. Trofimov, *Groups of automorphisms of graphs as topological groups*, Mat. Zametki **38** (1985), no. 3, 378–385, 476.

[Tuc11] T. W. Tucker, *Distinguishing maps*, Electron. J. Combin. **18** (2011), no. 1, Paper 50, 21 pp. (electronic).

[Wil70] S. Willard, *General topology*, Addison-Wesley Publishing Co., Reading, Mass.-London-Don Mills, Ont., 1970.

[Woe89] W. Woess, *Boundaries of random walks on graphs and groups with infinitely many ends*, Israel J. Math. **68** (1989), no. 3, 271–301.

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

[Woe09] ———, *Denumerable Markov chains*, EMS Textbooks in Mathematics, European Mathematical Society (EMS), Zürich, 2009.

[WZ07] M. E. Watkins and X. Zhou, *Distinguishability of locally finite trees*, Electron. J. Combin. **14** (2007), no. 1, Research Paper 29, 10 pp. (electronic).