On the coarse geometry of solvable Baumslag–Solitar groups and relatively hyperbolic groups

Patrick S. Nairne
St Edmund Hall
University of Oxford

A thesis submitted for the degree of
Doctor of Philosophy
Trinity 2024

Dedicated to my flatmates through my first two years at Oxford and through the COVID months when we first arrived: Faiz, Nick, Sam, Toby.

Abstract

Firstly, we prove that the solvable Baumslag–Solitar groups are rigid under quasiisometric embeddings, strengthening a classic result of Farb and Mosher [FM98]. These same ideas allow us to give a quasiisometric classification of treebolic spaces, confirming a conjecture of Woess [Woe13]. Both of these results are proved by relating the boundedness of a novel integer sequence $\mathcal{X}(p, q, p', q')$ to a pair of treebolic spaces.

In the second chapter, which is joint with Sam Hughes and Davide Spriano, we prove that the language of $(\lambda', 0)$-quasigeodesics in a non-hyperbolic group is not regular for $\lambda' > 54$. This is a strong converse to a result of Holt and Rees [HR03] in which they prove that in a hyperbolic group the (λ, ϵ)-quasigeodesics are regular whenever λ is rational. So we have provided a new characterisation of hyperbolic groups in terms of whether their quasigeodesics form regular languages.

In the next chapter, inspired by Buyalo, Dranishnikov and Schroeder's *Alice's Diary* [BDS07], we develop a general theory of *diaries* and *linear statistics.* These notions provide a powerful framework by which one can take a quasiisometric embedding of a metric space into a product of infinite-valence trees and upgrade it to a quasiisometric embedding into a product of *binary* trees.

Consequently, in the final chapter, we use diaries and linear statistics to prove that if a group G is relatively hyperbolic with respect to virtually abelian peripheral subgroups then G quasiisometrically embeds into a product of binary trees. This extends the result of Buyalo, Dranishnikov and Schroeder in which they prove that a hyperbolic group quasiisometrically embeds into a product of binary trees. To prove this result, we rely on the machinery of projection complexes and quasi-trees of metric spaces developed by Bestvina, Bromberg, Fujiwara and Sisto [Sis13] [BBF15] [BBFS19] [BBF21]. We build on this theory by proving that one can remove certain edges from the quasi-tree of metric spaces, and be left with a *tree* of metric spaces which is quasiisometric to the quasi-tree of metric spaces. In particular, this reproves a result of Hume [Hum17].

Introduction

This thesis consists of four pieces of research, which are covered in Chapter 1, Chapter 2, Chapter 3 and Chapter 4. In the first chapter we prove that the solvable Baumslag–Solitar groups are rigid under quasiisometric embeddings, in the second chapter we provide a characterisation of hyperbolic groups in terms of regular languages of quasigeodesics, the third chapter outlines a theory of *diaries* and *statistics* which is used in the fourth chapter to construct quasiisometric embeddings from certain relatively hyperbolic groups into products of binary trees. Since the research contained in Chapter 1 is published [Nai23], I have chosen not to present the detailed arguments but rather to refer the reader to the proofs in the published paper. This allows us to focus primarily on the ideas involved instead. At the time of submission, the research in Chapter 3 and Chapter 4 is being put together into a preprint that will be uploaded to arXiv.

The research on solvable Baumslag–Solitar groups and regular languages of quasigeodesics was completed in 2021. The research on diaries and relatively hyperbolic groups was completed in 2022, 2023 and 2024.

There is also Chapter 0 which contains certain basic definitions so that the later chapters of the thesis can focus more on proofs and ideas. Finally, there are Appendices A, B and C, in which a few results, definitions and examples have been placed if they are not entirely necessary for the comprehension of the previous chapters.

The topics of this thesis are quite varied; the research on solvable Baumslag–Solitar groups sits most naturally within research on quasiisometric rigidity (for example, Eskin, Fisher and Whyte's quasiisometric classification of the Sol groups [EFW12]), the research on regular languages of quasigeodesics sits within the intersection of computer science and group theory (as outlined in [ECH$^+$92]) and the research on relatively hyperbolic groups is at home among the mathematics of metric embeddings and their obstructions (for example [BS00] or [HMT22]). However, maybe all mathematicians are drawn to certain ideas, perhaps to the ones with which they have the natural ability to manipulate and experiment, and with no conscious intention I seem to have been drawn to trees, words and quasiisometric embeddings. So the reader should hopefully appreciate thematic links between the different chapters.

A defence of coarse geometry

On the topic of Cayley graphs, Gromov [Gro93] famously wrote the following paragraph.

> This space may appear boring and uneventful to the geometer's eye since it is discrete and the traditional local (e.g. topological and infinitesimal) machinery does not run in Γ. To regain the geometric perspective one has to change his/her position and move the observation point far away from Γ. Then the metric in Γ seen from the distance d becomes the original distance divided by d and for $d \to \infty$ the points in Γ coalesce into a connected continuous solid unity which occupies the visual horizon without any gaps or holes and fills our geometer's heart with joy.
>
> *Mikhael Gromov, Asymptotic Invariants of Infinite Groups*

Gromov is implicitly referring to the *asymptotic cone* in these sentences, a formalisation of the idea of looking at a metric space from infinitely far away. The asymptotic cone is, without a doubt, the mathematical idea discovered during my DPhil that I think most special. I think it perfectly captures the playful spirit that coarse geometry can possess. We begin with a fanciful idea—that if we look at a jaggedy shape filled with holes from really far away it should begin to appear smooth and the holes will disappear—and then we find the correct mathematical formalisation of this idea via ultralimits. But the formalisation is secondary; once the geometric idea has been sparked, everything else follows.

I do not know if the same can be said of many other mathematical disciplines. The very fact that our metric spaces are *coarse*, meaning they are to some extent fuzzy and only meaningful up to quasiisometry or some other coarse equivalence, allows us the flexibility to apply to them any geometric construction that comes to mind. The disadvantage of this fuzziness is that many constructions or ideas are immediately dismissed as utterly meaningless. Any finite alteration is equivalent to simply doing nothing, and there is often no distinction between constructing a metric space one way as opposed to another. The silver lining to this slightly dismissive attitude of coarse geometry is that the concepts which remain, and which are meaningful in this context, seem to capture the most fundamental aspects of a given infinite shape. A coarse invariant is a property so fundamental that vast amounts of distortion and destruction will fail to diminish it. A coarse geometric property reflects the large-scale *essence* of an object and so, to me, never fails to have a sort of grandeur.

Hence the interest in quasiisometries and quasiisometric embeddings. If one wishes to obstruct the existence of a quasiisometric embedding $X \to Y$ between metric spaces, then one needs to prove that X possesses a property so utterly incompatible with the geometry of Y that it cannot be placed inside Y even with vast amounts of distortion. If one wishes to construct a quasiisometric embedding $X \to Y$, then the coarse nature of the question gives us the flexibility to be experimental and bold in our choice of function $X \to Y$.

A note on the writing style

It is typical in mathematical writing to avoid using the words *I* and *me* in single author papers and instead to use the royal *we*. This is now so commonplace within mathematics that one forgets how strange it seems to outsiders. From my experience of reading a few academic books in other fields, I don't believe this is a common writing style outside of the sciences. I appreciate *we* when it is used to lead the reader through a proof; it makes the reading experience feel like a journey through which you are being led by the author. However I am sceptical when *we* is used in the introduction or conclusion of a paper to refer to results or ideas of the single author. I think this would often be justified on the grounds of formality. But sometimes I think it is instead used to evoke authority.

I think the apt use of *I*, instead of being in some way egoistic, announces to the reader that the author is a human (who you can therefore email freely!), who is necessarily fallible and who produced this research as a result of hours of hard work in libraries and offices, as opposed to some sort of equation-fabricating machine. I think it is good to reflect, even slightly, the human and community effort that the mathematics required. If Gauss believed (referring to mathematical proof) that "no self-respecting architect leaves the scaffolding in place after completing the building" then I am suggesting that on the contrary one ought to leave some of this authorial and historical scaffolding in place. So I use the first-person periodically throughout this thesis.

Acknowledgements

Thank you to my family for finding ways to remain excited and curious about what I was doing for almost four years. Thank you to Sam Fisher, Monika Kudlinska and Filippos Sytilidis for being a lovely group of people to start my DPhil with. Thank you to Davide Spriano for unreservedly and charismatically explaining ideas from geometric group theory to me when I knew nothing. Thank you to Harry Petyt for discussing the possibility of quasiisometrically embedding mapping class groups into products of binary trees with me. And thank you to my supervisor Cornelia Druțu for her sage advice and wisdom throughout my DPhil.

Contents

Chapter 0

Definitions and conventions

0.1 Numbers

Notation 0.1. We use $\mathbb{N}_0$ to denote the non-negative integers and $\mathbb{N}$ to denote the positive integers. Notation such as $\mathbb{R}_{>0}$ means the set $\{x \in \mathbb{R} : x > 0\}$.

0.2 Coarse geometry

Throughout this section, let X and Y be a pair of metric spaces and let $\lambda \geq 1$, $\mu \geq 0$ be a pair of constants.

Definition 0.2. We say that $f : X \to Y$ is (λ, μ)-*coarsely Lipschitz* if

$$d(f(x), f(x')) \leq \lambda d(x, x') + \mu$$

A $(\lambda, 0)$-coarsely Lipschitz function $f : X \to Y$ is λ-*Lipschitz*.

Definition 0.3. A (λ, μ)-*quasiisometric embedding* $f : X \to Y$ is a function which satisfies

$$\frac{1}{\lambda} d(x, x') - \mu \leq d(f(x), f(x')) \leq \lambda d(x, x') + \mu$$

for all $x, x' \in X$. It is a (λ, μ)-*quasiisometry* if in addition, for all $y \in Y$ there exists $x \in X$ such that $d(y, f(x)) \leq \mu$. A $(\lambda, 0)$-quasiisometric embedding $f : X \to Y$ is a λ-*bilipschitz embedding*. A $(\lambda, 0)$-quasiisometry $f : X \to Y$ is a λ-*bilipschitz homeomorphism*. A $(1, \mu)$-quasiisometric embedding $f : X \to Y$ is a μ-*rough isometric embedding*. A $(1, \mu)$-quasiisometry $f : X \to Y$ is a μ-*rough isometry*.

Definition 0.4. A metric space X has *bounded growth at some scale* if there are constants $0 < r < R$ and $N \in \mathbb{N}$ such that all open balls of radius R can be covered by N open balls of radius r.

The following proposition will be used, briefly, in Chapter 4.

Proposition 0.5. *If a metric space X admits a quasiisometric embedding into a uniformly bounded degree connected graph Γ then X has bounded growth on some scale.*

Proof. Suppose we have some (λ, μ)-quasiisometric embedding $\phi : X \to \Gamma$. Suppose that D is a uniform bound on the degree of vertices of Γ. We may assume that ϕ takes elements of X to vertices of Γ. Consider an open ball B of radius R in X. Then $\phi(B)$ is contained with an open ball B' of radius $\lambda R + \mu$ in Γ. B' contains at most $D^{\lambda R+\mu}$ vertices. For each vertex $v \in B'$ with $\phi^{-1}(v) \cap B$ non-empty, choose an element $x_v \in B$ with $\phi(x_v) = v$. If we consider balls of radius μ around all of the x_v, we can see that we have covered B by $D^{\lambda R+\mu}$ balls of radius μ. □

Convention 0.6. *Let X be a metric space. Sometimes, but not always, we will denote the metric on X by d_X, with a subscript to indicate the metric space. Other times, when the metric space is clear, we will just write d.*

0.3 Groups

Let G be a group and let S be a finite generating set (meaning that S is finite and every element of g can be written as a finite product of elements of S and S^{-1}). We use the notation $\Gamma(G, S)$ to indicate the *Cayley graph* of G with respect to S, which is the graph with vertex set G and an edge $\{g, gs\}$ for all $g \in G$ and $s \in S$.

Notation 0.7. Let $\Gamma(G, S)$ be a Cayley graph and suppose a subgroup $H \leq G$ is generated by some finite set $S_H \subset S$. We may think of the Cayley graph $\Gamma(H, S_H)$ as being a connected subgraph of $\Gamma(G, S)$. Given $g \in G$, we use the notation

$$g\Gamma(H, S_H)$$

to refer to the image of the subgraph $\Gamma(H, S_H)$ under the isometric action of g. So $g\Gamma(H, S_H)$ has vertex set gH and the edges correspond to the generators S_H.

0.4 Trees

Definition 0.8. Let T be a simplicial rooted tree with root vertex v_0.

- Say that two vertices of T are *adjacent* if they are distinct and connected by an edge.

- The *children* of a vertex $v \in T$ are precisely the vertices $v' \in T$ which are adjacent to v and for which $d_T(v_0, v') > d_T(v_0, v)$.
- We say that a vertex $v' \in T$ *descends* from a vertex $v \in T$ if the unique geodesic from v_0 to v' passes through v. If v' descends from v then v is an *ancestor* of v'.
- Given two vertices $u, v \in T$, the *lowest common ancestor* of u and v, denoted $\mathrm{LCA}(u, v)$, is the common ancestor of u and v with the greatest distance from the basepoint v_0.
- We say that T has *uniformly bounded valence* or *uniformly bounded degree* or that it is *uniformly locally finite* if there is a constant $M \in \mathbb{N}$ such that every vertex of T has degree at most M.

Now suppose T and T' are simplicial rooted trees with root vertices v_0 and v_0' respectively.

- We say that a map $f : T \to T'$ is *height-preserving* if for all vertices $v \in T$ we have $d(v_0, v) = d(v_0', f(v))$.
- We say that $f : T \to T'$ is *order-preserving* if the following implication holds: if v' descends from v in T then $f(v')$ descends from $f(v)$ in T'.

0.5 Words and sentences

Words

Definition 0.9. An alphabet is just a set of *letters*. We will often denote alphabets by A, B or C and we will often denote letters by a and b. We say that the alphabet is finite if there are finitely many letters.

Definition 0.10. A *word* is an empty, finite, or infinite string of letters from an alphabet A. The *length* of a word u is the number of letters in it. We denote this by $\mathrm{length}(u)$.

For example, if $A = \{a, b\}$ then $u = abab$ is a word of length 4 on the alphabet A. We will often denote by W the set of all finite words on an alphabet A. We will often denote words themselves by the letters u, v or w. To indicate the concatenation of a pair of words into a longer word, we simply write the two words next to each other. For example, if $u = abab$ and $v = aaa$ then

$$uv = ababaaa$$

Notation 0.11. We write $\overleftarrow{w}$ to denote the word w written backwards. So $\overleftarrow{abcd} = dcba$ and so on.

Further, we sometimes need to refer to the position of a letter a within a word u. To do this, we might refer to the *distance* of a from the start or end of u. If we say a has distance d from the start of u that means a is the d'th letter in the word u. If we say that a has distance d from the end of u that means a is the $(\text{length}(u) - d + 1)$'th letter of u, or, equivalently, the d'th letter of $\overleftarrow{u}$. For example, in the word $u = abaca$, the letter c has distance 2 from the end of u and distance 4 from the start of u.

Sentences

Definition 0.12. A *sentence* is a string of words. We will often denote sentences by the Greek letters α and β. In order to describe the words which make up a sentence, we will use *overlines* to indicate where one word ends and another begins. For example, if $A = \{a, b\}$ then

$$\alpha = \overline{abab}\,\overline{aaa}\,\overline{ba}$$

is a sentence that consists of the three words $abab$, aaa and ba. If $u = abab$, $v = aaa$ and $w = ba$ we can also write

$$\alpha = \overline{u}\ \overline{v}\ \overline{w}$$

The purpose of the overlines is so that we can distinguish between the single word uvw and the three word sentence $\overline{u}\ \overline{v}\ \overline{w}$. The *length* of a sentence is the number of words in it. We denote this by $\text{length}(\alpha)$.

Sentences are, of course, just words on an alphabet of words. To indicate the concatenation of a pair of sentences into a longer sentence, we simply write the two sentences next to each other. For example, if $\alpha = \overline{abab}\,\overline{aaa}\,\overline{ba}$ and $\beta = \overline{ab}\,\overline{a}$ then

$$\alpha\beta = \overline{abab}\ \overline{aaa}\ \overline{ba}\ \overline{ab}\ \overline{a}$$

Notation 0.13. Given a sentence α and $m \in \mathbb{N}$, we can write α^m to indicate the sentence α written m times. For example, $(\overline{a}\overline{b})^3 = \overline{a}\overline{b}\overline{a}\overline{b}\overline{a}\overline{b}$.

The following notation will be convenient in Chapter 4.

Notation 0.14. Suppose we have a group G generated by a finite set S and suppose $A = S \cup S^{-1}$ is our alphabet. If $q \in \mathbb{Z}$ and $q < 0$, and if $s \in S$, then $(\overline{s})^q$ indicates the sentence $(\overline{s^{-1}})^{-q}$. For example, $(\overline{s})^{-3} = \overline{s^{-1}}\ \overline{s^{-1}}\ \overline{s^{-1}}$.

Word-trees and sentence-trees

Word-trees and *sentence-trees* (both of these are my own terminology) are used frequently in Chapter 3 and Chapter 4.

Definition 0.15. Given an alphabet A, we will use the notation T_A to indicate *the word-tree* on A. The tree T_A is defined precisely so that its vertices are in bijection with finite words on the alphabet A. Every edge of T_A will be labelled by an element of A. The word-tree T_A is defined as follows. As a set, T_A is precisely the rooted tree where every vertex has $|A|$ children. Further, given a vertex $v \in T_A$, the $|A|$ edges which go between v and its children are labelled by the elements of A, with every such edge having a label in A, and no two such edges having the same label. We will now describe the natural bijection between the vertices of T_A and finite words on the alphabet A. First, we associate the empty word $\emptyset$ to the root vertex of T_A. Now let v be a vertex in T_A. The geodesic from $\emptyset$ to v in T_A travels along a series of labelled edges. These labels, read in order, make up the word that we associate to v. Thus every vertex in T_A has a corresponding word on A. *Indeed, we will often identify them.* The vertices at distance n from the root of T_A correspond to words of length n on the alphabet A.

The following notation for the rooted binary tree will be used frequently in Chapter 3 and Chapter 4.

Example 0.16. If $A = \{0, 1\}$ then $T_A = T_{\{0,1\}}$ is a rooted binary tree whose vertices correspond to finite strings of zeros and ones.

Definition 0.17. Let A be a finite alphabet and let W be the set of non-empty finite words on A. Analogous to the word-tree on A, we have a *sentence-tree* on A which is exactly the tree T_W, i.e. the word-tree on W. So the vertices of T_W are in bijection with sentences on A, and the edges of T_W are labelled by finite words on A. The vertices at distance n from the root vertex of T_W correspond to sentences with n words.

Note that if A is a finite alphabet then T_A is a uniformly bounded valence tree, whereas T_W is a tree with countable valence at every vertex.

Notation for pairs of sentences

Let T_W be a sentence-tree and suppose we have a pair of sentences α, $\beta \in T_W$. In Chapter 3 and Chapter 4, this situation will occur so frequently that it is useful to have some standard notation for the words that make up these sentences.

Notation 0.18. Given sentences $\alpha, \beta \in T_W$, we can always write

$$\alpha = \overline{u_1}\ \overline{u_2} \dots \overline{u_p}\ \overline{u_{p+1}} \dots \overline{u_{p+m}} \tag{1}$$

and

$$\beta = \overline{u_1}\ \overline{u_2} \dots \overline{u_p}\ \overline{u'_{p+1}} \dots \overline{u'_{p+n}} \tag{2}$$

where $u_{p+1} \neq u'_{p+1}$ and $m, n \geq 0$. In (1), u_1, u_2, $\ldots$, u_p, u_{p+1}, $\ldots$, u_{p+m-1}, u_{p+m} are the words that make up the sentence α. In (2), u_1, u_2, $\ldots$, u_p, u'_{p+1}, $\ldots$, u'_{p+n-1}, u'_{p+n} are the words that make up the sentence β. Therefore, the sentence $\overline{u_1}\,\overline{u_2}\ldots\overline{u_p}$ is the lowest common ancestor of the sentences α and β in the tree T_W and $d_{T_W}(\alpha, \beta) = m + n$.

Chapter 1

Quasiisometric embeddings of solvable Baumslag–Solitar groups

To what extent can some algebraic property of finitely generated groups be characterised by a geometric property? Suppose we have a group G satisfying some algebraic property P and we have another group H that is quasiisometric to G. Does H also satisfy P? If this holds for all such groups G, H then we say that the class of groups satisfying P is *quasiisometrically rigid*. It suggests that the algebraic property P in fact corresponds to some geometric feature of the groups. For example, Gromov [Gro81] proved that virtual nilpotency is invariant under quasiisometries and corresponds to the geometric property of polynomial growth. In contrast, Erschler [Dyu00] proved that virtual solvability is not geometric; there exist groups G and H, where G is solvable and H is not virtually solvable, such that G and H have a common Cayley graph. All is not lost - we can limit our focus to subclasses of solvable groups. Farb and Mosher [FM98] proved that the solvable Baumslag–Solitar groups $BS(1, m)$ exhibit a fascinating rigidity; $BS(1, m)$ is quasiisometric to $BS(1, n)$ if and only if m, n are powers of a common integer (which holds if and only if they are commensurable). This chapter concerns a stronger notion of rigidity. We will characterise when you can quasiisometrically embed $BS(1, m)$ into $BS(1, n)$.

1.1 Baumslag–Solitar groups

Let m and n be non-zero integers with $|m| \leq |n|$. The *Baumslag–Solitar group* $BS(m, n)$ is defined by the group presentation

$$BS(m, n) = \langle a, t \mid ta^m t^{-1} = a^n \rangle$$

It is known that $BS(m,n)$ is solvable if and only if $|m| = 1$; if $|m| = 1$ then $BS(m,n) \cong \mathbb{Z}[\frac{1}{n}] \rtimes \mathbb{Z}$ and if $|m| > 1$ then $BS(m,n)$ has a rank 2 free abelian subgroup. If n is a non-zero integer, we thus refer to the groups

$$BS(1,n) = \langle a, t \mid tat^{-1} = a^n \rangle$$

as the *solvable Baumslag–Solitar groups*. The group $BS(1,n)$ is quasiisometric to the group $BS(1,-n)$ and $BS(1,1) \cong \mathbb{Z}^2$ so we focus on the case when $n \geq 2$. Farb and Mosher [FM98] proved the following rigidity result for these groups.

Theorem 1.1 (Farb–Mosher)**.** *If m,n are integers and $m,n \geq 2$ then $BS(1,m)$ is quasiisometric to $BS(1,n)$ if and only if they are commensurable which occurs if and only m,n are powers of a common integer.*

In other words, if $BS(1,m)$ and $BS(1,n)$ are geometrically alike—i.e. they are quasiisometric—then they are algebraically similar by being commensurable. The full quasiisometric classification of (not necessarily solvable) Baumslag–Solitar groups $BS(m,n)$ was provided afterwards by Whyte [Why01].

I was was able to strengthen Farb and Mosher's rigidity theorem even further [Nai23].

Theorem 1.2. *If m,n are integers and $m,n \geq 2$ then $BS(1,m)$ admits a quasiisometric embedding into $BS(1,n)$ if and only if they are commensurable which occurs if and only m,n are powers of a common integer.*

In this chapter we will discuss the ideas behind the proof of this result.

1.2 Treebolic spaces

Instead of considering the geometry of a Cayley graph of $BS(1,m)$ directly, it is convenient to define a model space that is quasiisometric to $BS(1,m)$ and to work with that instead. This model space will be the *treebolic space* $HT(m,m)$; one can show [FM98, Section 3] that $BS(1,m)$ acts isometrically and cocompactly on $HT(m,m)$. However, the model spaces $HT(m,m)$ are only very specific examples of treebolic spaces: indeed, one can define a treebolic space $HT(p,q)$ for every integer $p \geq 2$ and every real number $q > 1$. These spaces were introduced by Bendikov, Saloff–Coste, Salvatori and Woess [BSCSW11] and they are natural geometric generalisations of the solvable Baumslag–Solitar groups.

Definition 1.3. Let $T(p,q)$ be the infinite metric tree with all edges of length $\log(q)$ and with valency $p+1$ at every vertex. Let $\gamma : [0,\infty) \to T(p,q)$ be some geodesic ray based

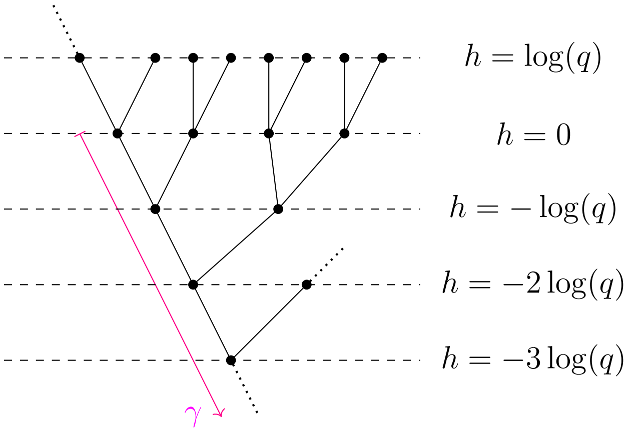

Figure 1.1: The height function induced by the geodesic ray γ. Drawn for $p = 2$. Made using TikZiT.

at a vertex b of $T(p, q)$. The ray γ induces a *height function* $h : T(p, q) \to \mathbb{R}$ on $T(p, q)$ via

$$h(x) = d(b, x) - 2 \cdot \text{length}(\gamma_x \cap \gamma)$$

where γ_x is the geodesic segment from b to x. See Figure 1.1. Now consider the upper half plane model of the hyperbolic plane: $\mathbb{H}^2 = \{(x, y) \in \mathbb{R} \times \mathbb{R}_{>0}\}$ with the metric $ds^2 = \frac{dx^2+dy^2}{y^2}$. We can put a height function on this model of $\mathbb{H}^2$ via

$$h(x, y) = \log(y)$$

Both of these height functions coincide with the classical notion of a Busemann function (see [BH99, Definition II.8.17]). By a *horostrip* in $\mathbb{H}^2$ we mean a subset of the form $\{(x, y) \in \mathbb{H}^2 : h(x, y) \in [a, b]\}$ for some real interval $[a, b]$. The *treebolic space* $HT(p, q)$ is formed by gluing horostrips onto every edge of $T(p, q)$ in a height-preserving manner. More precisely, let $HT(p, q) = T(p, q) \times \mathbb{R}$ as a set. Let $e = [v, w]$ be some edge of $T(p, q)$ such that $h(v) < h(w)$. We can put a metric on $e \times \mathbb{R}$ by identifying it with the horostrip

$$\{(x, y) \in \mathbb{H}^2 : h(x, y) \in [h(v), h(w)]\}$$

If we do this for every edge of $T(p, q)$ then we will have produced a metric on the whole of $HT(p, q)$ by taking the shortest path metric.

Informally, $HT(p, q)$ is the metric space formed by branching the hyperbolic plane $\mathbb{H}^2$ along a horocycle in p directions every time you walk upwards by a distance of $\log(q)$.

The key to understanding the geometry of $HT(p, q)$ is the following fact. Let $\pi : HT(p, q) \to T(p, q)$ be the projection. Let $L : (-\infty, \infty) \to T(p, q)$ be a height-increasing bi-infinite geodesic in $T(p,q)$. Then $\pi^{-1}(L) \subset HT(p,q)$ is an isometrically embedded copy of $\mathbb{H}^2$.

For more details on treebolic spaces see [Woe13, Section 2B].

Wolfgang Woess conjectured [Woe13, Question 2.15] the following quasiisometric classification of treebolic spaces. Since $HT(m,m)$ is quasiisometric to $BS(1,m)$, Woess' conjecture generalises the quasiisometric classification of solvable Baumslag–Solitar groups provided by Farb and Mosher.

Conjecture 1.4 (Woess). *Suppose $p, p' \geq 2$ are integers and $q, q' > 1$ are real numbers. Then $HT(p,q)$ is quasiisometric with $HT(p',q')$ if and only if p, p' are powers of a common integer and $\log(p)/\log(q) = \log(p')/\log(q')$.*

The treebolic space $HT(p,q)$ is an example of a *horocyclic product* (for a definition see [Woe13, Section 2]). Indeed, $HT(p,q)$ is the horocyclic product of the hyperbolic plane $\mathbb{H}^2$ and the regular tree $T(p,q)$. Other examples include the Diestel–Lieder graphs $\mathrm{DL}(p,q)$ (these are horocyclic products of pairs of regular trees and were introduced in [DL01]) and the Lie groups $\mathrm{Sol}(p,q)$ (which are horocyclic products of pairs of hyperbolic spaces). The quasiisometric classifications of the Diestel–Lieder graphs and the Sol groups were provided by Eskin, Fisher and Whyte [EFW12] [EFW13].

Theorem 1.26 below confirms Conjecture 1.4 to be correct. We would like to kill two birds with one stone by proving Woess' conjecture and Theorem 1.2 at the same time. So from this point onwards we will discuss quasiisometries *and* quasiisometric embeddings between treebolic spaces.

1.3 Trees and Cantor sets

We will understand quasiisometric embeddings and quasiisometries between treebolic spaces by relating them to morphisms in distinct, but related categories.

Definition 1.5. Let $p \geq 2$ be an integer and let $q > 1$ be a real number.

- The *regular tree* $T(p,q)$ is the metric tree with all edges of length $\log(q)$ and with valency $p+1$ at every vertex.
- The *regular rooted tree* $R(p,q)$ is the rooted metric tree with all edges of length $\log(q)$ and with valency $p+1$ at every vertex apart from the basepoint which has valency p.

- The *symbolic Cantor set* $\mathbb{Z}(p,q) = \{0,1,\dots,p-1\}^{\mathbb{N}}$ is the space of infinite sequences on p letters. We can put a metric ρ on $\mathbb{Z}(p,q)$ as follows. Given sequences $(a_n),(b_n) \in \mathbb{Z}(p,q)$, set $\rho((a_n),(b_n)) = q^{-N}$ where $a_n = b_n$ for $n \leq N$ and $a_{N+1} \neq b_{N+1}$.

I was able to prove the following two propositions.

Proposition 1.6. *Suppose $p, p' \geq 2$ are integers and suppose $q, q' > 1$ are real numbers. The following are equivalent.*

1. *There exists a quasiisometric embedding $HT(p,q) \to HT(p',q')$;*
2. *There exists a rough isometric embedding $T(p,q) \to T(p',q')$;*
3. *There exists a rough isometric embedding $R(p,q) \to R(p',q')$;*
4. *There exists a bilipschitz embedding $\mathbb{Z}(p,q) \to \mathbb{Z}(p',q')$.*

Proof. See [Nai23, Theorem 1 and Propositions 18,19,20]. ☐

Proposition 1.7. *Suppose $p, p' \geq 2$ are integers and suppose $q, q' > 1$ are real numbers. The following are equivalent.*

1. *$HT(p,q)$ is quasiisometric to $HT(p',q')$;*
2. *$T(p,q)$ is rough isometric to $T(p',q')$;*
3. *$R(p,q)$ is rough isometric to $R(p',q')$;*
4. *$\mathbb{Z}(p,q)$ is bilipschitz homeomorphic to $\mathbb{Z}(p',q')$.*

Proof. See [Nai23, Theorem 2 and Propositions 18,19,20]. ☐

Rough isometries of trees notably appeared in a recent work of Kerr [Ker23] in which it is proved that any geodesic metric space which is *quasiisometric* to a simplicial tree is in fact *rough isometric* to a simplicial tree.

A remark on the proofs of the two propositions above. *Most of the implications required to prove the two propositions above are relatively simple. For Proposition 1.6 and Proposition 1.7, the equivalence of (2) and (3) follows, as you might imagine, from elementary arguments involving trees. For both Proposition 1.6 and Proposition 1.7, the equivalence of (3) and (4) follows from the work of Bonk and Schramm [BS00] and the fact that $\mathbb{Z}(p,q)$ is essentially the boundary of $R(p,q)$. Proving that (2) implies (1) for Proposition 1.6 and Proposition 1.7 is more interesting: given a map $f : T(p,q) \to T(p',q')$, you can define an associated map $\hat{f} : HT(p,q) \to HT(p',q')$ called the horocyclic extension which will be a quasiisometric embedding (resp. quasiisometry) if f was a* *coarsely height-preserving rough isometric embedding (resp. coarsely height-preserving rough isometry).*

But of all the implications required to prove the two propositions above, the most significant is proving that (1) of Proposition 1.6 implies (4) of Proposition 1.6. In [FM98], Farb and Mosher prove that a quasiisometry $HT(m,m) \to HT(n,n)$ induces a bilipschitz embedding $\mathbb{Z}(m,m) \to \mathbb{Z}(n,n)$ onto a clopen. Then, in the appendix, they enlist Daryl Cooper to prove that a bilipschitz embedding $\mathbb{Z}(m,m) \to \mathbb{Z}(n,n)$ onto a clopen implies that the number theoretic condition of Theorem 1.1 holds. I was able to prove that a quasiisometric embedding $HT(p,q) \to HT(p',q')$ induces a bilipschitz embedding $\mathbb{Z}(p,q) \to \mathbb{Z}(p',q')$ which may or may not be onto a clopen. The key step was to show that [FM98, Lemma 5.1] does not actually require the map to be a quasiisometry, and that a quasiisometric embedding suffices. This improvement appears as [Nai23, Lemma 12]. Once this key step is proven, the remainder of Farb and Mosher's argument goes through verbatim.

However, since the map $\mathbb{Z}(p,q) \to \mathbb{Z}(p',q')$ is no longer onto a clopen, Cooper's argument no longer works. For this reason, it is necessary to search for an alternate argument.

Finally, the reason why (1) of Proposition 1.7 implies (4) of Proposition 1.7 is rather unsatisfactory; indeed, I don't know how to prove it directly, and it instead follows from the implication (1) $\implies$ (4) of Proposition 1.6 and the fact that if you have bilipschitz embeddings $\mathbb{Z}(p,q) \to \mathbb{Z}(p',q')$ and $\mathbb{Z}(p',q') \to \mathbb{Z}(p,q)$ then $\mathbb{Z}(p,q)$ and $\mathbb{Z}(p',q')$ are bilipschitz homeomorphic. See the proof of Theorem 1.26.

It follows from Proposition 1.6 and Proposition 1.7 that if we are interested in the existence or non-existence of quasiisometries and quasiisometric embeddings between treebolic spaces, it suffices to consider rough isometries and rough isometric embeddings between regular rooted trees $R(p,q)$. This will now be our focus. The big idea is to associate an integer sequence $\mathcal{X}(p,q,p',q')$ to every (ordered) pair of regular rooted trees $R(p,q)$ and $R(p',q')$ so that the boundedness of $\mathcal{X}$ captures the existence of rough isometric embeddings between the rooted trees.

1.4 Rough isometric embeddings between regular rooted trees

Definitions

Definition 1.8. Given a regular rooted tree $R(p,q)$ with basepoint b, we define a height function $h : R(p,q) \to \mathbb{R}_{\geq 0}$ on $R(p,q)$ via $h(x) = d(b,x)$.

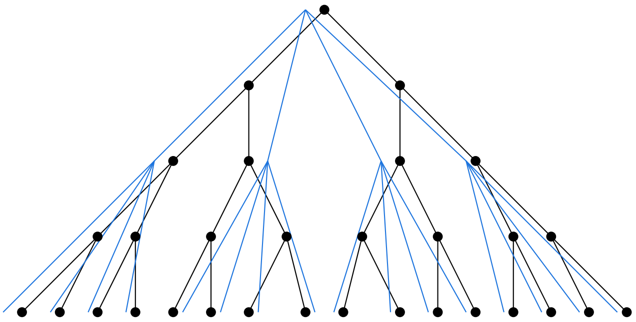

Figure 1.2: A rough isometry $R(4,4) \to R(2,2)$. Made using TikZiT.

So a map $f : R(p,q) \to R(p',q')$ is height-preserving if $h(f(x)) = h(x)$ for all $x \in R(p,q)$.

Definition 1.9. Given a regular rooted tree $R(p,q)$ and some height $h \geq 0$ we use the notation $R(p,q)\big|_h$ to denote the set of all points in $R(p,q)$ of height h; this is always a finite set.

Definition 1.10. Given $x \in R(p,q)$, let $\tau(x) \subset R(p,q)$ denote the *subtree based at* x, in other words, the set of all points $x' \in R(p,q)$ which descend from x.

Definition 1.11. Given a vertex v of $R(p,q)$, we denote by $\mathcal{C}(v)$ the set of *children* of v. That is, $\mathcal{C}(v) = \{w_1, \ldots, w_p\}$ where $w_1, \ldots, w_p$ are the vertices adjacent to v such that $h(w_i) > h(v)$.

Let $R(p,q)$ and $R(p',q')$ be regular rooted trees with basepoints b and b' respectively. We are interested in the following question: when does there exist a rough isometric embedding $R(p,q) \to R(p',q')$?

Existence

Consider the following two number theoretic conditions on p, p', q, q'.

(C1) $\log(p)/\log(q) < \log(p')/\log(q')$;

(C2) $\log(p)/\log(q) = \log(p')/\log(q')$ and p, p' are powers of a common integer.

Proposition 1.12. *If (C2) holds then $R(p,q)$ is rough isometric to $R(p',q')$.*

Proof. See [Nai23, Proposition 10]. As an illustrative example of why this holds consider the rough isometry $R(4,4) \to R(2,2)$ depicted in Figure 1.2. □

Proposition 1.13. *If (C1) or (C2) holds then there exists a rough isometric embedding* $R(p, q) \to R(p', q')$*.*

Proof. See [Nai23, Propositions 10 and 11]. Of course, if (C2) holds then there exists a rough isometric embedding due to Proposition 1.12. If (C1) holds, then one might justifiably say that $R(p, q)$ is smaller than $R(p', q')$—indeed one can show that the Hausdorff dimension of $\mathbb{Z}(p, q)$ is exactly $\log(p)/\log(q)$. $\square$

Obstructions

We are interested in whether the rough isometric embeddings and rough isometries that arise from Proposition 1.12 and Proposition 1.13 are the only possibilities. Throughout this subsection, I recommend that you consider in particular whether there exists a rough isometric embedding $R(2, 2) \to R(3, 3)$.

I believe there are two natural candidates of map $R(p, q) \to R(p', q')$ which might coarsely preserve the metric.

The first, which is perhaps the most obvious, is an order-preserving map $R(p, q) \to R(p', q')$ which sends basepoints to basepoints, vertices to vertices, edges to edges, stretching or squishing the the edges of $R(p, q)$ slightly (which have length $\log(q)$) so that they now have length $\log(q')$ (the length of edges in $R(p', q')$). If $p \le p'$, this map may be chosen to be injective. However, of course, this sort of function is never a rough isometric embedding when $q \neq q'$ because it stretches edges by a factor of $\log(q')/\log(q)$ and so the quasiisometric inequality would require a multiplicative constant.

The second candidate map seeks to address the issues with the first candidate by ensuring that there is no stretching of edges at all. We call these functions *waterfall maps.*

Definition 1.14. A *waterfall map* is a height-preserving and continuous map $f : R(p, q) \to R(p', q')$.

The following proposition is more of an observation.

Proposition 1.15. *Let* $f : R(p, q) \to R(p', q')$ *be a function. The following are equivalent.*

1. f *is a waterfall map;*
2. f *is height-preserving and order-preserving;*
3. $f(b) = b'$ *and* $f\big|_{\gamma}$ *is an isometry for all geodesic rays* γ *based at* b*.*

So a waterfall map $R(p,q) \to R(p',q')$ is one that places $R(p,q)$ inside $R(p',q')$ without stretching anything. They are called waterfall maps because it's as if we are letting $R(p,q)$ fall through $R(p',q')$ like water.

There is no chance of a waterfall map being injective unless $p \leq p'$ and $\log(q)$ is a multiple of $\log(q')$. Indeed, a waterfall map $R(p,q) \to R(p',q')$ might be metrically terrible. For example, it could hypothetically map every geodesic ray based at b onto a single geodesic ray based at b'. But this is clearly a bad choice of waterfall map if our aim is to preserve the metric on $R(p,q)$ as much as possible. There is a natural subclass of waterfall maps called *distributive* waterfall maps which attempt to preserve the metric on $R(p,q)$ as much as possible. A distributive waterfall map attempts (but generally fails) to be injective by dividing itself evenly at the vertices of $R(p',q')$. We shall see that the failure of a distributive waterfall map $R(p,q) \to R(p',q')$ to be a rough isometric embedding will actually imply that there are no rough isometric embeddings $R(p,q) \to R(p',q')$ *at all*. This is because, in a sense, a distributive waterfall map is the best map $R(p,q) \to R(p',q')$ possible.

Definition 1.16. Suppose $f : R(p,q) \to R(p',q')$ is a waterfall map and suppose w is a vertex of $R(p',q')$ such that $f^{-1}(w)$ is non-empty. Write $f^{-1}(w) = \{x_1, \ldots, x_m\}$. Consider the subsets

$$\begin{aligned} B_\epsilon(w) &= \tau(w) \cap \{x \in R(p',q') : d(w,x) \leq \epsilon\} \\ B_\epsilon(x_i) &= \tau(x_i) \cap \{x \in R(p,q) : d(x_i,x) \leq \epsilon\} \qquad (1 \leq i \leq m) \end{aligned}$$

and choose ϵ small enough that $B_\epsilon(w)$ is isometric to a star with p' arms and $B_\epsilon(x_i)$ is isometric to a star with p arms (if the x_i are vertices) or a star with one arm (if the x_i are not vertices). If the x_i are all vertices then set $P = mp$, if the x_i are not vertices then set $P = m$. Since f is a waterfall map, an arm of $B_\epsilon(x_i)$ emanating from some x_i is mapped isometrically onto one of the p' arms of $B_\epsilon(w)$ emanating from w. Thus, we get a map from a set of cardinality P (the set of arms emanating from *all* the x_i) to a set of cardinality p' (the set of arms emanating from w). We say that f is *distributive* if at most $\lceil \frac{P}{p'} \rceil$ arms emanating from the x_i can be mapped to the same arm emanating from w. See Figure 1.3.

The point is this: a waterfall map $f : R(p,q) \to R(p',q')$ is completely determined by how it chooses to divide its image at the vertices of $R(p',q')$. At a vertex $w \in R(p',q')$, it might choose to send the image down a single arm emanating from w, which would be a bad choice metrically. But if f is distributive, then it chooses to divide the image as evenly as possible among the arms emanating from w.

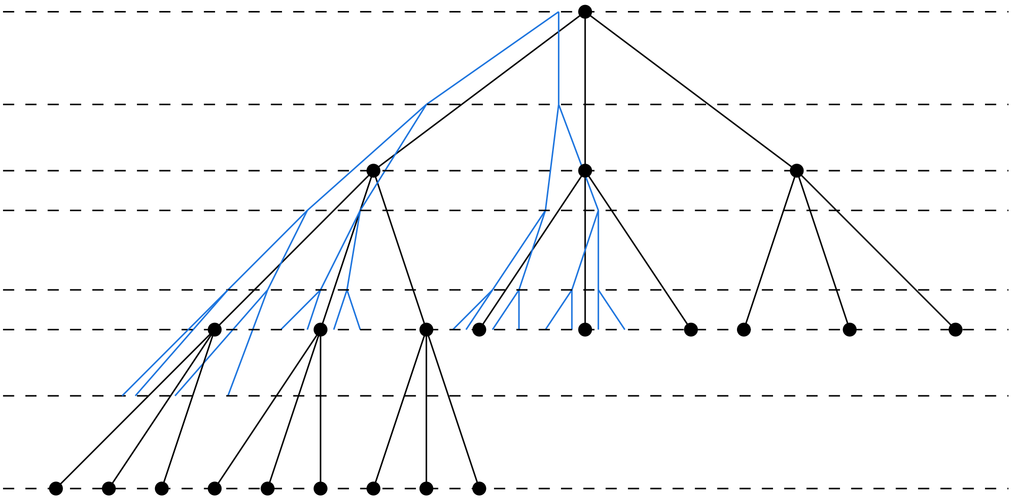

Figure 1.3: A distributive waterfall map $R(2,2) \to R(3,3)$. The dashed lines indicate multiples of $\log(2)$ and $\log(3)$ where there is branching. If you look at the vertex on the far left at height $2\log(3)$, you can see that the preimage has cardinality 4 and so the blue lines are divided as $2, 1, 1$ between the edges beneath the vertex. Made using TikZiT.

Example 1.17. Let P and p' be as in Definition 1.16 and suppose $P = 8$ and $p' = 3$. In this case, the only ways to divide the 8 arms of the $B_\epsilon(x_i)$ evenly between the 3 arms of $B_\epsilon(w)$ is as $3, 3, 2$ or $3, 2, 3$ or as $2, 3, 3$.

If a waterfall map $f : R(p, q) \to R(p', q')$ is map which lets $R(p, q)$ fall within $R(p', q')$ like water, then a distributive waterfall map is one which lets $R(p, q)$ expand within $R(p', q')$ *like a gas.*

Now, let $f : R(p, q) \to R(p', q')$ be a distributive waterfall map. Let us consider whether f is a rough isometric embedding. Consider the following function $\mathcal{X} : \mathbb{R}_{\geq 0} \to \mathbb{N}$

$$h \mapsto \max\{|f^{-1}(x)| : x \in R(p', q')\Big|_h\}$$

where $|f^{-1}(x)|$ indicates the cardinality of $f^{-1}(x) \subset R(p, q)\big|_h$. So $\mathcal{X}$ captures the maximal cardinality of point preimages at a given height. Here are five key observations we can make about the function $\mathcal{X}$.

- (O1) $\mathcal{X}$ does not depend on the choice of distributive waterfall map. In other words, the function $\mathcal{X}$ only depends on the pair of trees $R(p, q)$ and $R(p', q')$;
- (O2) $\mathcal{X}$ is a piecewise constant map which changes value exclusively at multiples of $\log(q)$ and $\log(q')$;
- (O3) when h is a multiple of $\log(q)$, the function $\mathcal{X}$ multiplies by p';

(O4) when h is a multiple of $\log(q')$, the function $\mathcal{X}$ divides itself by p' and then applies the ceiling function;

(O5) when h is a multiple of $\log(q)$ *and* $\log(q')$, $\mathcal{X}$ multiplies itself by p, then divides itself by p' and applies the ceiling function.

To see the truth of the first observation, recall that a waterfall map f is determined by how it divides its image at vertices $w \in R(p', q')$. But if f is distributive, then the way in which it divides its image between the edges emanating downwards from w is essentially predetermined. For example, in the case of Example 1.17, from the perspective of the function $\mathcal{X} : \mathbb{R}_{\geq 0} \to \mathbb{N}$ each of the three possibilities $3, 3, 2$ or $3, 2, 3$ or $2, 3, 3$ are equivalent. Indeed, given two distributive waterfall maps $f, g : R(p, q) \to R(p', q')$, there exists a graph automorphism $\psi \in \mathrm{Aut}(R(p', q'))$ such that $f(x) = \psi(g(x))$. Since the function $\mathcal{X}$ is piecewise constant and only changes value at multiples of $\log(q)$ and $\log(q')$ we may associate to it an integer sequence $\mathcal{X} : \mathbb{N}_{\geq 0} \to \mathbb{N}$ which simply captures all the different integer values that the function $\mathcal{X} : \mathbb{R}_{\geq 0} \to \mathbb{N}$ takes in order. It follows from observations (O3), (O4) and (O5) that the sequence $\mathcal{X} : \mathbb{R}_{\geq 0} \to \mathbb{N}$ is *exactly* the following inductively defined sequence.

Definition 1.18. We have the sets of vertex heights

$$\mathcal{H}(R(p, q)) = \{a \log(q) : a \in \mathbb{N}_0\} \quad \text{and} \quad \mathcal{H}(R(p', q')) = \{b \log(q') : b \in \mathbb{N}_0\}$$

Now let $\mathcal{H} = \mathcal{H}(R(p, q)) \cup \mathcal{H}(R(p', q'))$ and write $\mathcal{H} = \{h_0, h_1, h_2, \dots\}$ such that $h_n < h_{n+1}$ for $n \in \mathbb{N}_0$. We can now define the infinite integral sequence $\mathcal{X} = \mathcal{X}(p, q, p', q') : \mathbb{N}_{\geq 0} \to \mathbb{N}$. Set $\mathcal{X}_0 = 1$. If $\mathcal{X}_n$ has already been defined then we set

$$\mathcal{X}_{n+1} = \begin{cases} p\mathcal{X}_n, & \text{if } h_n \in \mathcal{H}(R(p, q)) \setminus \mathcal{H}(R(p', q')) \\ \lceil \frac{\mathcal{X}_n}{p'} \rceil, & \text{if } h_n \in \mathcal{H}(R(p', q')) \setminus \mathcal{H}(R(p, q)) \\ \lceil \frac{p\mathcal{X}_n}{p'} \rceil, & \text{if } h_n \in \mathcal{H}(R(p, q)) \cap \mathcal{H}(R(p', q')) \end{cases}$$

Remark 1.19. *Another definition of the sequence $\mathcal{X}_n$ can be found by observing that the progression of the sequence $\mathcal{X}_n$ depends only on the order of the multiples of $\log(q)$ and $\log(q')$ within the sequence (h_n) rather than on the values themselves. Hence, we could replace $\mathcal{H}$ with the set*

$$\mathcal{S} = \mathcal{S}(R(p, q)) \cup \mathcal{S}(R(p', q')) = \{q^a : a \in \mathbb{N}_0\} \cup \{(q')^b : b \in \mathbb{N}_0\} = \{s_0, s_1, s_2, \dots\}$$

where $h_n = \log(s_n)$ and $\mathcal{X}_n$ would not change.

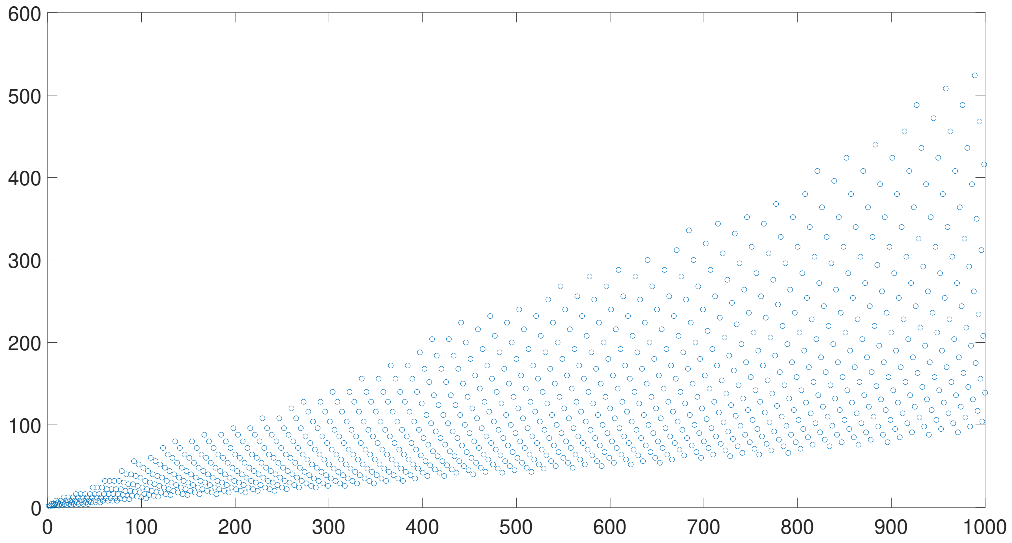

Figure 1.4: The sequence $\mathcal{X}(2,2,3,3)$. Plotted up to $n = 1000$. Made with MATLAB.

Example 1.20. Consider the following description of the sequence $\mathcal{X}(2,2,3,3)$. We imagine that we are about to go for a walk along $\mathbb{R}_{\geq 1}$ (which we think of as being a long pebbly beach) beginning at 1. We suppose that we begin with a single pebble in our possession. We then imagine that we walk forwards along $\mathbb{R}_{\geq 1}$ obeying the following rule as we go: each time we pass a power of 2, we multiply the amount of pebbles in our possession by 2; each time we pass a power of 3, we divide our pebbles evenly into 3 groups and keep only one of the larger groups (e.g. we divide 8 pebbles into 3,3,2 and then keep only 3 pebbles). The changing quantity of pebbles in our possession as we walk along the number line forms the integral sequence $\mathcal{X}(2,2,3,3)$. So the sequence $\mathcal{X}(2,2,3,3)$ begins

$$1, 1, 2, 1, 2, 4, 2, 4, 2, 4, 8, 3, 6, 2, 4, 8, 3, 6, 12, 4, \ldots$$

See Figure 1.4.

Observation 1.21. *If the sequence* $\mathcal{X}(p,q,p',q')$ *is unbounded then a distributive waterfall map* $f : R(p,q) \to R(p',q')$ *is not a rough isometric embedding.*

The observation follows from the fact that if $\mathcal{X}(p,q,p',q')$ is unbounded then we have preimages whose cardinalities are arbitrarily large. Since distributive waterfall maps are height-preserving, this means points (with the same height) in $R(p,q)$ which are arbitrarily far apart are being mapped to the same point in $R(p',q')$.

Now for the punchline.

Theorem 1.22. *If $\mathcal{X}(p,q,p',q')$ is unbounded then there does not exists a rough isometric embedding $f : R(p,q) \to R(p',q')$.*

Proof. See [Nai23, Theorem 5]. But the idea is this: we suppose there exists a rough isometric embedding $f : R(p,q) \to R(p',q')$. One can then budge f by a bounded distance to create a height-preserving rough isometric embedding $g : R(p,q) \to R(p',q')$. Further, g will be *coarsely* order-preserving. In other words, g is almost a waterfall map. If we cleverly analyse the cardinalities of preimages (we will actually have to consider the preimages of certain uniformly bounded subsets of $R(p',q')$ as opposed to single points) of g, we will find the sequence $\mathcal{X}(p,q,p',q')$ appear as a lower bound on these cardinalities. It follows that $\mathcal{X}(p,q,p',q')$ is bounded. □

We need the following lemma in order to analyse when $\mathcal{X}$ is unbounded.

Lemma 1.23. *Suppose that $\log(q')/\log(q)$ is irrational. Let $a\log(q)$ be a multiple of $\log(q)$ and let $b\log(q')$ be the least multiple of $\log(q')$ that is greater than $a\log(q)$. Then there exists $a' > a$ such that, if $b'\log(q')$ is the least multiple of $\log(q')$ greater than $a'\log(q)$, we have*

$$b'\log(q') - a'\log(q) < b\log(q') - a\log(q)$$

Proof. See [Nai23, Lemma 8]. The lemma follows from the fact that the rotation of the circle $S^1 = \mathbb{R}/\mathbb{Z}$ by the irrational angle of $\log(q')/\log(q)$ has dense orbits. □

Theorem 1.24. *If $\mathcal{X}(p,q,p',q')$ is bounded then (C1) or (C2) holds.*

Proof. See [Nai23, Theorem 9].

The proof has two steps. The first step is to show that if $\log(p)/\log(q) > \log(p')/\log(q')$ then $\mathcal{X}$ is unbounded. The second step is to show that if $\log(p)/\log(q) = \log(p')/\log(q')$ and p, p' are not powers of a common integer then $\mathcal{X}$ is unbounded.

I will provide a heuristic argument for the first step. Consider the function $\mathcal{X} : \mathbb{R}_{\geq 0} \to \mathbb{N}$. At height $h \in \mathbb{R}_{\geq 0}$, we might expect the sequence to have multiplied by p on $h/\log(q)$ occasions. Similarly, we might expect the sequence to have divided by p' (and taken the ceiling) $h/\log(q')$ times. Thus we might estimate

$$\mathcal{X}(h) \approx p^{h/\log(q)} \left(\frac{1}{p'}\right)^{h/\log(q')} = e^{h(\log(p)/\log(q) - \log(p')/\log(q'))}$$

The second step has an elegant argument, that I will provide in full. So suppose that $\log(p)/\log(q) = \log(p')/\log(q')$ and p, p' are not powers of a common integer. This implies that $\log(q')/\log(q)$ is irrational. Let $h_l = a\log(q)$ be some arbitrary element of $\mathcal{H}(R(p,q))$.

Let $b\log(q')$ be the least multiple of $\log(q')$ greater than $a\log(q)$. By Lemma 1.23 we can find some $a' > a$ such that if $b'\log(q')$ is the least multiple of $\log(q')$ greater $a'\log(q)$ then

$$b'\log(q') - a'\log(q) < b\log(q') - a\log(q)$$

which rearranges to

$$b' - b < (a' - a)\frac{\log(q)}{\log(q')}$$

If $h_m = a'\log(q) \in \mathcal{H}(R(p,q))$ then

$$\mathcal{X}_m \geq p^{a'-a}\left(\frac{1}{p'}\right)^{b'-b}\mathcal{X}_l > p^{a'-a}\left(\frac{1}{p'}\right)^{(a'-a)\frac{\log(q)}{\log(q')}}\mathcal{X}_l = \mathcal{X}_l$$

So $\mathcal{X}_m > \mathcal{X}_l$. But both $\mathcal{X}_m$ and $\mathcal{X}_l$ are integral and so in fact $\mathcal{X}_m \geq \mathcal{X}_l + 1$. We have shown that given $l \in \mathbb{N}_0$ such that $h_l \in \mathcal{H}(R(p,q))$ we can find $m > l$ such that $h_m \in \mathcal{H}(R(p,q))$ and $\mathcal{X}_m \geq \mathcal{X}_l + 1$. Repeating this process gives arbitrarily large values of $\mathcal{X}_n$ and so the sequence is unbounded. □

1.5 Conclusions

Combining our work in previous sections gives us the following theorems.

Theorem 1.25. *Suppose $p, p' \geq 2$ are integers and suppose $q, q' > 1$ are real numbers. The following are equivalent.*

1. *There exists a quasiisometric embedding $HT(p,q) \to HT(p',q')$;*
2. *There exists a rough isometric embedding $T(p,q) \to T(p',q')$;*
3. *There exists a rough isometric embedding $R(p,q) \to R(p',q')$;*
4. *There exists a bilipschitz embedding $\mathbb{Z}(p,q) \to \mathbb{Z}(p',q')$;*
5. *$\mathcal{X}(p,q,p',q')$ is bounded;*
6. *Either (C1) or (C2) hold.*

Proof. The first four equivalences are given by Proposition 1.6. Theorem 1.22 gives that (3) implies (5). Theorem 1.24 gives that (5) implies (6). Proposition 1.13 gives that (6) implies (3). □

Theorem 1.26. *Suppose $p, p' \geq 2$ are integers and suppose $q, q' > 1$ are real numbers. The following are equivalent.*

1. *$HT(p,q)$ is quasiisometric to $HT(p',q')$;*
2. *$T(p,q)$ is rough isometric to $T(p',q')$;*
3. *$R(p,q)$ is rough isometric to $R(p',q')$;*
4. *$\mathbb{Z}(p,q)$ is bilipschitz homeomorphic to $\mathbb{Z}(p',q')$;*
5. *(C2) holds.*

Proof. The first four equivalences are given by Proposition 1.7. If $R(p,q)$ is rough isometric to $R(p',q')$ then there exist rough isometric embeddings $R(p,q) \to R(p',q')$ and $R(p',q') \to R(p,q)$. It then follows from Theorem 1.25 that (C2) holds. Then Proposition 1.12 gives that (C2) implies (3). □

Theorem 1.25 implies that there exists a quasiisometric embedding $HT(m,m) \to HT(n,n)$ if and only if (C2) holds. Since $HT(m,m)$ is quasiisometric to $BS(1,m)$, we have proved Theorem 1.2.

Further, the equivalence of (1) and (5) in Theorem 1.26 provides the quasiisometric classification of treebolic spaces conjectured by Woess [Woe13, Question 2.15].

1.6 Discussion

The integer sequence $\mathcal{X}(p,q,p',q')$

Interestingly enough, despite the relative simplicity of its definition, the integer sequence $\mathcal{X}(p,q,p',q')$ does not seem to have been defined previously. In particular, the sequences $\mathcal{X}(2,2,3,3)$ and $\mathcal{X}(3,3,2,2)$ do not appear in the On-Line Encyclopedia of Integer Sequences (`https://oeis.org`).

Theorem 1.25 characterises when the sequence $\mathcal{X}_n$ is bounded. In the case when the sequence is unbounded, if we plot $\mathcal{X}_n(p,q,p',q')$ up to large values of $n \in \mathbb{N}$, it looks as if it is bounded above and below by linear functions (see Figure 1.4).

Question 1.27. *Is the limit inferior of the sequence $\mathcal{X}_n$ finite or infinite? Are the limit inferior and limit superior of the sequence $\mathcal{X}_n/n$ finite? If so, what are they?*

Generalisation

Theorem 1.22 can be generalised to a broader class of trees than just the regular rooted trees.

Definition 1.28. Suppose we have an integer sequence $(p_n)_{n\in\mathbb{N}}$ where $p_n \geq 2$ and a real sequence $(q_n)_{n\in\mathbb{N}}$ where $q_n \in \mathbb{R}_{>1}$. Suppose also that p_n and q_n are bounded sequences and that $\inf_n q_n > 1$. We can construct a rooted metric tree $R((p_n),(q_n))$ as follows. We start with a basepoint b. The basepoint b has p_1 edges emanating from it of length $\log(q_1)$. The terminal vertices of these edges have p_2 edges emanating from them of length $\log(q_2)$. Then the terminal vertices of those edges have p_3 edges emanating from them of length $\log(q_3)$. Continuing in this way we will have constructed an infinite rooted tree $R((p_n),(q_n))$. We call a rooted tree constructed in this manner *spherically homogeneous.* The regular rooted tree $R(p,q)$ is precisely the spherically homogeneous tree associated to the constant sequences $p_n = p$, $q_n = q$.

Given two spherically homogeneous trees $R = R((p_n),(q_n))$ and $R' = R((p'_n),(q'_n))$, one can define a distributive waterfall map $f : R \to R'$ analogously to Definition 1.16. Hence, we can also define an integer sequence $\mathcal{X}((p_n),(q_n),(p'_n),(q'_n))$ analogously to Definition 1.18. I was able [Nai23, Theorem 5] to prove the following generalisation of Theorem 1.22.

Theorem 1.29. *Suppose we have a pair of spherically homogeneous trees* $R((p_n),(q_n))$, $R((p'_n),(q'_n))$ *such that* (p'_n) *is a constant sequence. If* $\mathcal{X}((p_n),(q_n),(p'_n),(q'_n))$ *is unbounded then there does not exist a rough isometric embedding* $R((p_n),(q_n)) \to R((p'_n),(q'_n))$.

The requirement that (p'_n) is constant is frustrating, and one wonders whether…

Question 1.30. *Can the condition that* (p'_n) *is constant be removed from Theorem 1.29? This would follow from a proof that a rough isometric embedding* $R((p_n),(q_n)) \to R((p'_n),(q'_n))$ *is at bounded distance from a waterfall map.*

Remark 1.31. *Given sequences* $(p_n)_{n\in\mathbb{N}}$ *and* $(q_n)_{n\in\mathbb{N}}$ *as above, one can define a symbolic Cantor set* $\mathbb{Z}((p_n),(q_n))$ *in a natural way. It is highly likely that a bilipschitz embedding* $\mathbb{Z}((p_n),(q_n)) \to \mathbb{Z}((p'_n),(q'_n))$ *is equivalent to a rough isometric embedding* $R((p_n),(q_n)) \to R((p'_n),(q'_n))$. *By attaching horocycles to the trees* $R((p_n),(q_n))$, *it is possible one could also deduce statements about generalisations of treebolic spaces.*

Relationship with the work of other mathematicians

The category of metric trees and rough isometric embeddings is evidently very similar to the category of Cantor sets and bilipschitz embeddings. For this reason, the results stated above no doubt intersect with the wealth of research on bilipschitz embeddings and bilipschitz homeomorphism between Cantor sets. As a geometric group theorist, I cannot fully grasp how my work on trees sits within the existing research on Cantor sets, however

I am aware of a few results which intersect with it. I discovered these results after having already completed most of the work in this section.

Firstly, Falconer and Marsh [FM92] describe some necessary conditions for two self-similar Cantor sets to be bilipschitz homeomorphic. Indeed, it follows from [FM92, Theorem 3.3] that (4) implies (5) of Theorem 1.26. This is because the Falconer–Marsh result implies that we have the following equality of fields:

$$\mathbb{Q}(\left(\frac{1}{q}\right)^{\frac{\log(p)}{\log(q)}}) = \mathbb{Q}(\left(\frac{1}{q'}\right)^{\frac{\log(p')}{\log(q')}})$$

Secondly, Deng, Wen, Xiong and Xi [DWXX11] consider bilipschitz embeddings of self-similar sets satisfying a "strong separation condition"; the symbolic Cantor sets $\mathbb{Z}(p,q)$ discussed in this thesis fall within this category. [DWXX11, Theorem 1] immediately gives us that (C1) implies (4) in Theorem 1.25. Further, [DWXX11, Theorem 2] tells us that there exists a bilipschitz embedding $\mathbb{Z}(p,q) \to \mathbb{Z}(p',q')$ if and only if $\mathbb{Z}(p,q)$ and $\mathbb{Z}(p',q')$ are bilipschitz homeomorphic. Hence Theorem 1.2 follows from a combination of [FM92, Theorem 3.3], [DWXX11, Theorem 2] and the implication (1) $\implies$ (4) of Proposition 1.6.

Both the Falconer–Marsh and Deng–Wen–Xiong–Xi papers concern *self-similar* sets. However, Theorem 1.29 concerns spherically homogeneous trees which generally have no self-similarity whatsoever. By passing to the boundary Cantor set $\mathbb{Z}((p_n),(q_n))$ (see Remark 1.31), we will find a necessary condition for there to exist a bilipschitz embedding between the *non-self-similar* Cantor sets $\mathbb{Z}((p_n),(q_n))$ and $\mathbb{Z}((p'_n),(q'_n))$. Thus the sequences $\mathcal{X}((p_n),(q_n),(p'_n),(q'_n)$ ought to nonetheless provide novel obstructions to various metric embeddings.

Chapter 2

Regular languages of quasigeodesics in groups

joint with Sam Hughes and Davide Spriano

> If I make no mistake, I have heard some people say this: there is no definition of the primary elements—so to speak—out of which we and everything else are composed; for everything that exists in its own right can only be named, no other determination is possible, neither that it is nor that it is not . . . But what exists in its own right has to be . . . named without any other determination. In consequence it is impossible to give an account of any primary element; for it, nothing is possible but the bare name; its name is all it has. But just as what consists of these primary elements is itself complex, so the names of the elements become descriptive language by being compounded together. For the essence of speech is the composition of names.
> *Socrates, from Plato's Theaetetus*

And so from the complexity of the world, built out of primary elements that cannot be described with anything but a name, arises the complexity of language built from those names.

In this section, we characterise hyperbolic groups in terms of whether the languages formed out of their quasigeodesics are regular or not. One direction of this characterisation is due to Holt and Rees [HR03]; myself, Sam Hughes and Davide Spriano prove a strong converse to their result.

2.1 Geodesic and quasigeodesic words in a Cayley graph

Let G be a group with finite generating set S. Let $\Gamma(G, S)$ denote the Cayley graph of G with respect to S. Let d denote the path metric on the Cayley graph. Let W be the set of finite or infinite words on the alphabet $S \cup S^{-1}$. We may identify elements of W with paths in $\Gamma(G, S)$ based at the identity $e \in G$. The *length* of a word $w \in W$ is the number of letters in w. If $w \in W$ has finite length then we say it is a *finite* word. Finite words in W also correspond to group elements, the group element being the terminal vertex of the path based at e corresponding to w. We denote by $w_G \in G$ the group element corresponding to a finite word $w \in W$.

Given two words $u, v \in W$, we denote their concatenation, which is also an element of W, by uv. We say that $y \in W$ is a *subword* of a word $w \in W$ if w can be written as $w = xyz$ for words $x, z \in W$.

Definition 2.1. Let $\lambda \geq 1$ and $\epsilon \geq 0$. We say that a word $w \in W$ is a (λ, ϵ)*-quasigeodesic* if

$$\text{length}(u) \leq \lambda d(e, u_G) + \epsilon$$

for all subwords u of w. A $(1, 0)$-quasigeodesic word is known as a *geodesic* word.

Remark 2.2. *We might also define a* (λ, ϵ)*-quasigeodesic word* $w \in W$ *as follows. Let* $w : [0, \mathit{length}(w)] \to \Gamma(G, S)$ *be the natural arclength parametrisation of the path* w *in the Cayley graph. We say that* w *is a* (λ, ϵ)*-quasigeodesic if*

$$|t - s| \leq \lambda d(w(s), w(t)) + \epsilon$$

for all $s, t \in [0, \mathit{length}(w)]$*. This definition is equivalent to Definition 2.1.*

Definition 2.3. A finite word $w \in W$ is a *loop* if $w =_G e$. A loop corresponds to a graph homomorphism $\phi : C_n \to \Gamma(G, S)$ of the circular graph C_n into the Cayley graph $\Gamma(G, S)$.

Definition 2.4. Suppose we have a loop $w \in W$ of length n and suppose $\phi : C_n \to X$ is the corresponding graph homomorphism. Let L be a natural number with $L \leq n$, and let $\lambda \geq 1$ and $\epsilon \geq 0$ be real. We say that w is a L*-locally* (λ, ϵ)*-quasigeodesic* if the restriction of ϕ to any subpath of C_n of length L corresponds to a (λ, ϵ)-quasigeodesic in $\Gamma(G, S)$.

2.2 Regular languages of geodesics

Regular languages of geodesics are usually defined via *finite state automaton*. Personally, I find this definition unwieldy, and believe it is preferable to give a definition in terms of cone types in the Cayley graph. The equivalence of these two perspectives is perhaps more of an observation: a cone type is nothing more than a state of the automata.

Let G be a group with finite generating set S. Let $\Gamma(G, S)$ denote the Cayley graph of G with respect to S. Let W be the set of finite or infinite words on the alphabet $S \cup S^{-1}$.

Definition 2.5. Let w be a finite geodesic word on the alphabet $S \cup S^{-1}$. We define the *geodesic cone* of w in $\Gamma(G, S)$ to be

$$\mathrm{Cone}(w) = \{w' \in W : ww' \text{ is a geodesic word in } \Gamma(G, S)\}$$

Example 2.6. If $G = \mathbb{Z}^2 = \langle a, b \mid [a, b]\rangle$ and $S = \{a, b\}$ then

$$\mathrm{Cone}(ab) = \mathrm{Cone}(a^3b^4) = \{\text{finite or infinite words on the letters } a, b\}$$

and

$$\mathrm{Cone}(a^{-3}) = \mathrm{Cone}(a^{-3}b^2) \cup \mathrm{Cone}(a^{-3}b^{-1})$$

Definition 2.7. We define

$$\mathrm{Cones}(G, S) = \{\mathrm{Cone}(w) : w \text{ is a finite geodesic word in } W\}$$

to be the set of *geodesic cone types* in $\Gamma(G, S)$. If $\mathrm{Cones}(G, S)$ is a finite set, then we say that $\Gamma(G, S)$ has a *regular language of geodesics.*

Example 2.8. If $G = \mathbb{Z}^2 = \langle a, b \mid [a, b]\rangle$ and $S = \{a, b\}$ then $\Gamma(G, S)$ has a regular language of geodesics. Indeed, it has exactly 9 geodesic cone types. These are

$$\begin{aligned}
&\mathrm{Cone}(ab) = \{\text{finite or infinite words on the letters } a, b\}\\
&\mathrm{Cone}(ab^{-1}) = \{\text{finite or infinite words on the letters } a, b^{-1}\}\\
&\mathrm{Cone}(a^{-1}b) = \{\text{finite or infinite words on the letters } a^{-1}, b\}\\
&\mathrm{Cone}(a^{-1}b^{-1}) = \{\text{finite or infinite words on the letters } a^{-1}, b^{-1}\}\\
&\mathrm{Cone}(a) = \mathrm{Cone}(ab) \cup \mathrm{Cone}(ab^{-1})\\
&\mathrm{Cone}(a^{-1}) = \mathrm{Cone}(a^{-1}b) \cup \mathrm{Cone}(a^{-1}b^{-1})\\
&\mathrm{Cone}(b) = \mathrm{Cone}(ab) \cup \mathrm{Cone}(a^{-1}b)\\
&\mathrm{Cone}(b^{-1}) = \mathrm{Cone}(ab^{-1}) \cup \mathrm{Cone}(a^{-1}b^{-1})\\
&\mathrm{Cone}(\emptyset) = \mathrm{Cone}(ab) \cup \mathrm{Cone}(ab^{-1}) \cup \mathrm{Cone}(a^{-1}b) \cup \mathrm{Cone}(a^{-1}b^{-1})
\end{aligned}$$

In 1984, Cannon [Can84] proved the following theorem.

Theorem 2.9 (Cannon)**.** *If G is a hyperbolic group and S is any finite generating set then $\Gamma(G, S)$ has a regular language of geodesics.*

Similarly, Neumann and Shapiro [NS95] were able to prove that it does not matter which finite generating set we choose for the free abelian groups $\mathbb{Z}^n$.

Theorem 2.10 (Neumann–Shapiro)**.** *If S is any finite generating set for $\mathbb{Z}^n$ then $\Gamma(\mathbb{Z}^n, S)$ has a regular language of geodesics.*

Having a regular language of geodesics is a rather flimsy concept, as the following example shows [NS95, Remark after Theorem 4.3].

Theorem 2.11 (Cannon)**.** *There exists a finite generating set S for a semidirect product $\mathbb{Z}^2 \rtimes \mathbb{Z}_2$ such that $\Gamma(G, S)$ does not have a regular language of geodesics.*

It is currently unknown whether hyperbolic groups and the free abelian groups $\mathbb{Z}^n$ are the only examples of groups which have regular languages of geodesics for *all* their finite generating sets. However there are several examples of groups for which *there exists* a finite generating set S which gives a regular language of geodesics: virtually abelian groups [NS95], Garside groups [CM04], groups hyperbolic relative to virtually abelian peripheral subgroups [AC16].

2.3 Regular languages of quasigeodesics

Definition 2.12. Let $\lambda \geq 1$ and $\epsilon \geq 0$. Let w be a finite (λ, ϵ)-quasigeodesic word on the alphabet $S \cup S^{-1}$. We define the (λ, ϵ)-*quasigeodesic cone* of w in $\Gamma(G, S)$ to be

$$\mathrm{Cone}_{\lambda,\epsilon}(w) = \{w' \in W : ww' \text{ is a } (\lambda, \epsilon)\text{-quasigeodesic word in } \Gamma(G, S)\}$$

Definition 2.13. We define

$$\mathrm{Cones}_{\lambda,\epsilon}(G, S) = \{\mathrm{Cone}_{\lambda,\epsilon}(w) : w \text{ is a finite } (\lambda, \epsilon)\text{-quasigeodesic word in } W\}$$

to be the set of (λ, ϵ)-*quasigeodesic cone types* in $\Gamma(G, S)$. If $\mathrm{Cones}_{\lambda,\epsilon}(G, S)$ is a finite set, then we say that $\Gamma(G, S)$ has a *regular language of* (λ, ϵ)-*quasigeodesics.*

Definition 2.14. We say that a finitely generated group G is $\mathbb{Q}$**REG** if it has a regular language of (λ, ϵ)-quasigeodesics for all finite generating sets S, for all rational $\lambda \geq 1$ and for all real $\epsilon \geq 0$. In other words, $\mathrm{Cones}_{\lambda,\epsilon}(G, S)$ is finite whenever λ is rational.

Holt and Rees [HR03] were able to prove the following remarkable theorem.

Theorem 2.15 (Holt–Rees)**.** *If G is a hyperbolic group then G is* $\mathbb{Q}$**REG**.

In the same paper, Holt and Rees also prove that if G is hyperbolic, then for all finite generating sets S, all *irrational* $\lambda \geq 1$ and all real $\epsilon \geq 2\lambda$, the set of (λ, ϵ)-quasigeodesics do not form a regular language.

Now, although Theorem 2.9 tells us that hyperbolic groups have regular languages of geodesics, it follows from Theorem 2.10 that this is not a sufficient condition. It is reasonable to ask whether the stronger condition $\mathbb{Q}$**REG** suffices to guarantee hyperbolicity. Indeed, it does.

Theorem 2.16 (Hughes–Nairne–Spriano)**.** *A finitely generated group G is hyperbolic if and only if it is* $\mathbb{Q}$**REG**.

The forward implication of Theorem 2.16 is exactly Theorem 2.15 of Holt and Rees. Sam Hughes, Davide Spriano and I proved that all non-hyperbolic groups fail to be $\mathbb{Q}$**REG**. In other words, given a non-hyperbolic group G, there exists a choice of generating set S, a choice of rational number $\lambda \geq 1$, and a choice of real $\epsilon \geq 0$, such that the (λ, ϵ)-quasigeodesics are not regular in $\Gamma(G, S)$.

Characterisations of groups in terms of their language theoretic properties are limited. A classical example is the result of Muller and Schupp [MS83] in which they prove that virtually free groups are exactly those for which the language of words representing the identity is context-free. Theorem 2.16 provides another such characterisation.

2.4 Why $\mathbb{Z}^2$ is not $\mathbb{Q}$REG

The inspiration for the general case comes from the following sketch proof of the fact that $\mathbb{Z}^2$ is not $\mathbb{Q}$**REG**.

Example 2.17. Let $G = \mathbb{Z}^2 = \langle a, b \mid [a, b] \rangle$ and $S = \{a, b\}$. We claim that $\Gamma(G, S)$ does not have a regular language of $(\lambda', 0)$-quasigeodesics for any $\lambda' \geq 3$.

Consider for each $n \in \mathbb{N}$, the increasing sequence of loops based at e in the Cayley graph $\gamma_n = a^n b^n a^{-n} b^{-n}$. We may parametrise these loops as $\gamma_n : [0, 4n] \to \Gamma(G, S)$.

- Let t_n be the maximal natural number such that $\gamma_n\big|_{[0,t_n]}$ is a geodesic (so $t_n = 2n$ and $\gamma_n\big|_{[0,t_n]} = a^n b^n$);
- let T_n be the minimal natural number such that $\gamma_n\big|_{[0,T_n]}$ is *not* a $(\lambda', 0)$–quasigeodesic.

Suppose we have two natural numbers $m, n \in \mathbb{N}$ where $n \geq 1000m$. Now, by definition,

$$\gamma_n \Big|_{[t_n,T_n]} \notin \mathrm{Cone}_{\lambda',0}(a^n b^n)$$

By contrast, we claim that

$$\gamma_n \Big|_{[t_n,T_n]} \in \mathrm{Cone}_{\lambda',0}(a^m b^m)$$

We may write $\gamma_n \big|_{[t_n,T_n]} = a^{-n}b^{-k}$ for some $1 \leq k \leq n$. We claim that the path $\eta = a^m b^m a^{-n} b^k$ is a $(\lambda', 0)$-quasigeodesic. It holds for this reason: we have chosen n so much larger than m that the end of η, where plausibly the quasiisometric inequality could break, is forced to be far away from its start. Assuming this to be true, it follows that $\mathrm{Cone}_{\lambda',0}(a^m b^m) \neq \mathrm{Cone}_{\lambda',0}(a^n b^n)$ whenever $n \geq 1000m$. So we have proved that all the words in the set $\{a^{1000^k} b^{1000^k} : k \in \mathbb{N}\}$ have distinct $(\lambda', 0)$-cone types.

See Example 2.27 for a formal proof that the $(\lambda', 0)$-quasigeodesics in $\mathbb{Z}^2$ with the standard generating set are not regular for all $\lambda' > 1$.

2.5 Why non-hyperbolic groups are not $\mathbb{Q}$REG

If we would like to mimic the argument in Example 2.17 for an arbitrary non-hyperbolic group G, then as a first step we would like to find a sequence of locally quasigeodesic loops γ_n in the Cayley graph. Consider the following proposition of Hume and Mackay [HM20, Proposition 5.1].

Proposition 2.18 (Hume–Mackay)**.** *Let X be a connected graph. X is hyperbolic if and only if there is some N such that every 18-bilipschitz embedded cyclic subgraph in X has length at most N.*

By an "18-bilipschitz embedded cyclic subgraph in X", Hume and Mackay mean an injective graph homomorphism $\phi : C_n \to X$ of the circular graph C_n into the graph X such that $d_{C_n}(x, y) \leq 18 d_X(\phi(x), \phi(y))$ for all vertices $x, y \in C_n$. The construction of simple locally $(100, 1)$-quasigeodesics loops in non-hyperbolic groups had appeared previously in the work of Papasoglu [Pap05, Lemma 3.8].

Corollary 2.19. *Let $K > 2$ be some constant. If G is a non-hyperbolic group, then for any finite generating set S, there exists an increasing sequence of natural numbers $l_n \to \infty$ and a sequence of L_n-locally $(18, 0)$-quasigeodesic loops γ_n (in the Cayley graph $\Gamma(G, S)$) of length l_n where $l_n \leq K L_n$.*

Proof. Proposition 2.18 tells us that there exists an increasing sequence of natural numbers $l_n \to \infty$ and a sequence of 18-bilipschitz embedded cyclic subgraphs in $\Gamma(G, S)$ of length l_n. We may write these cyclic subgraphs as injective graph homomorphisms $\phi_n : C_{l_n} \to \Gamma(G, S)$. If $L_n = \lfloor l_n/2 \rfloor$ then any subpath in C_{l_n} of length L_n is a geodesic. So C_{l_n} is L_n-locally $(1, 0)$-quasigeodesic, from which it follows that its image in $\Gamma(G, S)$ is L_n-locally $(18, 0)$-quasigeodesic. For l_n sufficiently large we have $l_n \leq KL_n$ and the result follows. □

Definition 2.20. Let $\Gamma(G, S)$ be a Cayley graph and consider the following condition:

> There exists an increasing sequence of positive numbers $l_n \to \infty$ and a pair of constants $K, \lambda \geq 1$ such that for every n there exists an L_n-locally $(\lambda, 0)$–quasigeodesic loop γ_n of length l_n in $\Gamma(G, S)$ with $l_n \leq KL_n$. $(\star)$

At times, it is convenient to specify the values of the constants K, λ. In that case we say that $\Gamma(G, S)$ satisfies $(\star)$ with constants (K, λ).

Remark 2.21. *Without the insistence that the local constant L_n is linearly comparable to l_n (i.e. that $l_n \leq KL_n$ for all n), the condition would be almost vacuous since every simple loop is 1-locally geodesic.*

The statement of Corollary 2.19 is exactly that for any non-hyperbolic group G, and for any finite generating set S, the Cayley graph $\Gamma(G, S)$ satisfies $(\star)$ with constants $(K, 18)$ where K is any constant greater than 2.

Remark 2.22. *In our preprint [HNS22], we prove a version of Proposition 2.18 for non-hyperbolic geodesic metric spaces. We prove that a geodesic metric space is non-hyperbolic if and only if there exists a constant $K \geq 1$ and a sequence of L_n-locally $(3, 0)$-quasigeodesic loops γ_n of length l_n where $l_n \to \infty$ and $l_n \leq KL_n$. We proved this prior to discovering Proposition 2.18. Our theorem applies to all geodesic metric spaces as opposed to the Hume–Mackay result which applies only to graphs. However, in our result, the constant K is not explicitly defined and, a priori, it depends on the choice of non-hyperbolic metric space.*

In order to prove that a non-hyperbolic geodesic metric space X admits an increasing sequence of locally quasigeodesic loops γ_n as above, we used the fact that X has an asymptotic cone which is not a real tree, and this asymptotic cone therefore contains a simple triangle. If we have a simple triangle "at infinity", one might expect that X itself admits an expanding sequence of simple polygons whose limit is the triangle. This is the content of [DS05, Proposition 3.29(c)]. By "smoothing out" the corners of these polygons we will get the sequence of locally quasigeodesic loops γ_n that we desire.

Theorem 2.23. *Suppose* $\Gamma(G,S)$ *is a Cayley graph that satisfies* ($\star$) *with constants* (K,λ)*. Then for all* $\lambda' > (2K-1)\lambda$ *the set of* $(\lambda',0)$*-quasigeodesics do not form a regular language.*

Proof. Suppose we have a Cayley graph $\Gamma(G,S)$ satisfying ($\star$) with constants K,λ. Let $\lambda' > (2K-1)\lambda$. We need to choose the parametrisations of our loops γ_n thoughtfully.

Claim. There exists an arclength parametrisation $\gamma_n : [0,l_n] \to \Gamma(G,S)$ of the loop γ_n such that if $T_n \in \mathbb{N}$ is the minimal natural number such that $\gamma_n\big|_{[0,T_n]}$ is *not* a $(\lambda',0)$-quasigeodesic, then $\gamma_n\big|_{[1,T_n]}$ *is* a $(\lambda',0)$-quasigeodesic.

Proof of claim. To begin with, suppose $\gamma_n' : [0,l_n] \to \Gamma(G,S)$ is some arbitrary parametrisation by arclength of the loop γ_n. Let $T_n' \in \mathbb{N}$ be the minimal natural number such that $\gamma_n'\big|_{[0,T_n']}$ is *not* a $(\lambda',0)$-quasigeodesic. It follows that there exists a non-empty collection of non-negative integers

$$\mathcal{T}_n' = \{T \in \mathbb{N}_0 : T \leq T_n' \text{ and } \lambda'|\gamma_n'(T)^{-1}\gamma_n'(T_n')| < T_n' - T\}$$

Let $S_n' = \max \mathcal{T}_n'$. Consider now the alternate parametrisation $\gamma_n : [0,l_n] \to \Gamma(G,S)$ of our loop defined by $\gamma_n(t) = \gamma_n'(t+S_n')$. It follows that $\gamma_n(0) = \gamma_n'(S_n')$ and $\gamma_n(T_n'-S_n') = \gamma_n'(T_n')$. If we define $T_n = T_n' - S_n'$ then

- $\gamma_n\big|_{[0,T_n-1]}$ is a $(\lambda',0)$-quasigeodesic by the minimality of T_n';
- $\gamma_n\big|_{[0,T_n]}$ is not a $(\lambda',0)$-quasigeodesic since $S_n' \in \mathcal{T}_n'$;
- $\gamma_n\big|_{[1,T_n]}$ is a $(\lambda',0)$-quasigeodesic by the maximality of S_n';

and the claim follows. □

You might think of the claim above in terms of *rotation* of our loops. We may also assume that $\gamma_n(0) = e$ for all n by considering *translation*. For each n, we fix the following notation:

- Let t_n be the maximal natural number such that $\gamma_n\big|_{[0,t_n]}$ is a $(\lambda,0)$–quasigeodesic;
- let T_n be the minimal natural number such that $\gamma_n\big|_{[0,T_n]}$ is *not* a $(\lambda',0)$–quasigeodesic;
- let $g_n \coloneqq \gamma_n(t_n)$;
- let $h_n \coloneqq \gamma_n(t_n)^{-1}\gamma_n(T_n)$. So $\gamma_n(T_n) = g_n h_n$.

Now, let $m \in \mathbb{N}$ be arbitrary. Since $l_n \to \infty$ and $L_n \geq l_n/K$, we may choose n such that

$$L_n > \frac{2KL_m}{\kappa}. \tag{2.1}$$

where κ is the positive constant

$$\kappa := \frac{1}{\lambda} - \frac{2K-1}{\lambda'}. \tag{2.2}$$

We will show that the $(\lambda', 0)$-cone types of the $(\lambda', 0)$–quasigeodesics $\gamma_m\big|_{[0,t_m]}$ and $\gamma_n\big|_{[0,t_n]}$ are distinct.

Let $\eta : [0, t_m + T_n - t_n]$ denote the concatenation of $\gamma_m\big|_{[0,t_m]}$ with the path $\gamma_n\big|_{[t_n,T_n]}$.

Suppose first that $\eta\big|_{[0,t_m+T_n-t_n-1]}$ is *not* a $(\lambda', 0)$-quasigeodesic. Then we are done since we know that $\gamma_n\big|_{[0,t_n]}$ concatenated with $\gamma_n\big|_{[t_n,T_n-1]}$ *is* a $(\lambda', 0)$-quasigeodesic (by the minimality of T_n) whereas $\gamma_m\big|_{[0,t_m]}$ concatenated with the same path is *not* a $(\lambda', 0)$-quasigeodesic. So we may assume that $\eta\big|_{[0,t_m+T_n-t_n-1]}$ is a $(\lambda', 0)$-quasigeodesic.

Suppose we have proven that $\eta\big|_{[0,t_m+T_n-t_n]}$ *is* a $(\lambda', 0)$-quasigeodesic. Then $\gamma_n\big|_{[0,t_n]}$ concatenated with the path $\gamma_n\big|_{[t_n,T_n]}$ *is not* a $(\lambda', 0)$–quasigeodesic (by the definition of T_n), but $\gamma_m\big|_{[0,t_m]}$ concatenated with the same path *is* a $(\lambda', 0)$–quasigeodesic. It would follow that the $(\lambda', 0)$–quasigeodesics $\gamma_m\big|_{[0,t_m]}$ and $\gamma_n\big|_{[0,t_n]}$ have distinct $(\lambda', 0)$-cone types. So we would like to prove that $\eta\big|_{[0,t_m+T_n-t_n]}$ is a $(\lambda', 0)$-quasigeodesic. Looking for a contradiction, suppose this is false. Then there exists some integer t with $0 \leq t \leq t_m + T_n - t_n$ such that

$$\lambda'|\eta(t)^{-1}\eta(t_m + T_n - t_n)| < t_m + T_n - t_n - t$$

If $t \geq t_m$, then $\gamma_n\big|_{[t,T_n]}$ is not a $(\lambda', 0)$-quasigeodesic. However, by our assumption on the parametrisations of the loops, we know this is not the case. So we may assume that $t \leq t_m$. Hence,

$$\lambda'|\gamma_m(t)^{-1}g_m h_n| < t_m + T_n - t_n - t \tag{2.3}$$

The following six inequalities are easily verified

$$L_m \leq t_m \leq KL_m; \tag{2.4}$$

$$L_n \leq t_n; \tag{2.5}$$

$$T_n \leq KL_n \leq Kt_n; \tag{2.6}$$

$$|\gamma_m(t)^{-1}g_m| \leq t_m - t; \tag{2.7}$$

$$\frac{t_n}{\lambda} \leq |g_n|; \tag{2.8}$$

$$|\gamma_n(T_n)| < \frac{T_n}{\lambda'}. \tag{2.9}$$

where the final inequality follows from the fact that $\gamma_n\big|_{[0,T_n-1]}$ and $\gamma_n\big|_{[1,T_n]}$ are both $(\lambda', 0)$-quasigeodesics yet $\gamma_n\big|_{[0,T_n]}$ is not a $(\lambda', 0)$-quasigeodesic.

We have $|h_n| \geq |g_n| - |\gamma_n(T_n)|$, so by (2.8) and (2.9) we see that $|h_n| \geq \frac{t_n}{\lambda} - \frac{T_n}{\lambda'}$. It then follows from (2.6) that

$$|h_n| \geq \left(\frac{1}{\lambda} - \frac{K}{\lambda'}\right) t_n. \tag{2.10}$$

Now, $|\gamma_m(t)^{-1}g_m h_n| \geq |h_n| - |\gamma_m(t)^{-1}g_m|$, so by (2.10) and (2.7) we obtain

$$|\gamma_m(t)^{-1}g_m h_n| \geq \left(\frac{1}{\lambda} - \frac{K}{\lambda'}\right) t_n - (t_m - t). \tag{2.11}$$

Combining our assumption (2.3) with (2.6) we obtain

$$\lambda'|\gamma_m(t)^{-1}g_m h_n| \leq t_m - t + (K-1)t_n. \tag{2.12}$$

Next, combining (2.11) and (2.12) yields

$$\frac{t_m - t + (K-1)t_n}{\lambda'} \geq \left(\frac{1}{\lambda} - \frac{K}{\lambda'}\right) t_n - (t_m - t).$$

This rearranges to

$$\begin{aligned} 0 &\geq \left(\frac{1}{\lambda} - \frac{(2K-1)}{\lambda'}\right) t_n - \left(1 + \frac{1}{\lambda'}\right)(t_m - t); \\ &\geq \kappa t_n - 2(t_m - t), \end{aligned}$$

where κ is defined in (2.2). Now,

$$t_n \leq \frac{2(t_m - t)}{\kappa} \leq \frac{2t_m}{\kappa},$$

and so by (2.4) and (2.5) we have

$$L_n \leq \frac{2KL_m}{\kappa}$$

which contradicts (2.1). So $\eta\big|_{[0,t_m+T_n-t_n]}$ is a $(\lambda',0)$-quasigeodesic and so the $(\lambda',0)$–quasigeodesics $\gamma_m\big|_{[0,t_m]}$ and $\gamma_n\big|_{[0,t_n]}$ have distinct $(\lambda',0)$-cone types.

Let $\xi : \mathbb{N} \to \mathbb{N}$ be the function

$$\xi(m) = \min\left\{n \in \mathbb{N} : L_n > \frac{2KL_m}{\kappa}\right\}.$$

Let $(n_i)_{i\in\mathbb{N}}$ be the integer sequence defined inductively by $n_1 = 1$, $n_{i+1} = \xi(n_i)$. By the above, we know that $\gamma_{n_i}\big|_{[0,t_{n_i}]}$ and $\gamma_{n_j}\big|_{[0,t_{n_j}]}$ have distinct $(\lambda',0)$-cone types whenever $i \neq j$. It follows that there are infinitely many different $(\lambda',0)$-cone types. Hence, the $(\lambda',0)$–quasigeodesics in $\Gamma(G,S)$ cannot form a regular language. ☐

2.6 Conclusion and discussion

Corollary 2.24. *Let G be a non-hyperbolic group and let S be any finite generating set. Then for all $\lambda' > 54$, the set of $(\lambda',0)$-quasigeodesics in $\Gamma(G,S)$ do not form a regular language. In particular, G is not* $\mathbb{Q}$**REG**.

Proof. This is Corollary 2.19 combined with Theorem 2.23. ☐

Theorem 2.16 follows from Theorem 2.15 and Corollary 2.24. As you can see, there is quite dramatic behaviour among groups. Indeed, for any finitely generated group G, one of two things can occur:

1. for all finite generating sets S and for all rational $\lambda' \geq 1$, the $(\lambda',0)$-quasigeodesics in $\Gamma(G,S)$ form a regular language;

2. for all finite generating sets S and for all $\lambda' > 54$, the $(\lambda',0)$-quasigeodesics in $\Gamma(G,S)$ do not form a regular language.

Corollary 2.24 suggests that we ask the following question.

Question 2.25. *Given a non-hyperbolic group G, what is the infimal value of $\lambda \geq 1$ such that for all $\lambda' > \lambda$ we have that the $(\lambda',0)$-quasigeodesics are not regular?*

I would conjecture the following.

Conjecture 2.26. *If G is a non-hyperbolic finitely generated group, then for all generating sets S, and for all $\lambda' > 1$, the $(\lambda', 0)$-quasigeodesics in $\Gamma(G, S)$ do not form a regular language.*

Indeed, for $\mathbb{Z}^2$ with the standard generating set, one can show that the $(\lambda', 0)$-quasigeodesics are not regular for $\lambda' > 1$. This is clearly an improvement over what is proved in Example 2.17, however it is not obvious how the argument would generalise to an arbitrary non-hyperbolic group.

Example 2.27. Let $G = \mathbb{Z}^2 = \langle a, b \mid [a, b] \rangle$ and $S = \{a, b\}$. Let $\lambda' > 1$. We will prove that $\Gamma(G, S)$ does not have a regular language of $(\lambda', 0)$-quasigeodesics.

Suppose, for a contradiction, that the $(\lambda', 0)$-quasigeodesics do form a regular language. Then there are $M < \infty$ possible $(\lambda', 0)$-cone types.

Let k, m, t be natural numbers and suppose $\gamma_{k,m,t} : [0, k + m + t] \to \Gamma(G, S)$ is the path in the Cayley graph given by $a^k b^m a^{-t}$.

Claim. *If $\gamma_{k,m,t}$ is a $(\lambda', 0)$-quasigeodesic and if $k \geq M$ then there exists some $k' > k$ such that $\gamma_{k',m,t}$ is a $(\lambda', 0)$-quasigeodesic.*

Proof of claim. Since $k \geq M$, we know that there exist $0 \leq s_1 < s_2 \leq k$ such that a^{s_1} and a^{s_2} have the same $(\lambda', 0)$-cone type. Since $\gamma_{k,m,t}$ is a $(\lambda', 0)$-quasigeodesic, we know that $a^{k-s_1} b^m a^{-t} \in \mathrm{Cone}_{\lambda',0}(a^{s_1})$. Hence $a^{k-s_1} b^m a^{-t} \in \mathrm{Cone}_{\lambda',0}(a^{s_2})$ and so $a^{s_2} a^{k-s_1} b^m a^{-t} = a^{k+s_2-s_1} b^m a^{-t} = \gamma_{k+s_2-s_1,m,t}$ is a $(\lambda', 0)$-quasigeodesic. By setting $k' = k + s_2 - s_1$, we are done. ☐

A similar argument proves the following claim.

Claim. *If $\gamma_{k,m,t}$ is a $(\lambda', 0)$-quasigeodesic and if $t \geq M$ then there exists some $t' > t$ such that $\gamma_{k,m,t'}$ is a $(\lambda', 0)$-quasigeodesic.*

Now, let $m \in \mathbb{N}$ be large enough that the path $\gamma_{M,m,M}$ is a $(\lambda', 0)$-quasigeodesic (choosing $m \geq 2M/(\lambda' - 1)$ will suffice). By repeatedly applying the first claim above, we see that there exist arbitrarily large values of k such that $\gamma_{k,m,M}$ is a $(\lambda', 0)$-quasigeodesic. In particular, there exists $k \geq \lambda' m$ such that $\gamma_{k,m,M}$ is a $(\lambda', 0)$-quasigeodesic. Then, by repeatedly applying the second claim, we can find $t \geq k$ such that $\gamma_{k,m,t}$ is a $(\lambda', 0)$-quasigeodesic. Consider the subpath $a^k b^m a^{-k}$ of $\gamma_{k,m,t}$. Since $2k + m > \lambda' m = \lambda' d(e, a^k b^m a^{-k})$ we see that this subpath is not a $(\lambda', 0)$-quasigeodesic. So $\gamma_{k,m,t}$ is not a $(\lambda', 0)$-quasigeodesic and we have a contradiction.

Chapter 3

Diaries

In *100 Years of Solitude* a plague of insomnia afflicts the town of Macondo. The most troubling effect of the insomnia plague is that it ultimately leads to memory loss.

> One day he was looking for the small anvil that he used for laminating metals and he could not remember its name. His father told him: "Stake." Aureliano wrote the name on a piece of paper that he pasted to the base of the small anvil: *stake.* In that way he was sure of not forgetting it in the future. It did not occur to him that this was the first manifestation of a loss of memory, because the object had a difficult name to remember. But a few days later he discovered that he had trouble remembering almost every object in the laboratory. Then he marked them with their respective names so that all he had to do was read the inscription in order to identify them. When his father told him about his alarm at having forgotten even the most impressive happenings of his childhood, Aureliano explained his method to him, and José Arcadio Buendía put it into practice all through the house and later on imposed it on the whole village. With an inked brush he marked everything with its name: *table, chair, clock, door, wall, bed, pan.* He went to the corral and marked the animals and plants: *cow, goat, pig, hen, cassava, caladium, banana.* Little by little, studying the infinite possibilities of a loss of memory, he realized that the day might come when things would be recognized by their inscriptions but that no one would remember their use. Then he was more explicit. The sign that he hung on the neck of the cow was an exemplary proof of the way in which the inhabitants of Macondo were prepared to fight against loss of memory: *This is the cow. She must be milked every morning so that she will produce milk, and the milk must be boiled in order to be mixed with coffee to make coffee and milk.* Thus they went on living in a reality that was slipping away, momentarily captured by words, but which would escape

> irremediably when they forgot the values of the written letters.
>
> *100 Years of Solitude, Gabriel García Márquez*

The human effort to record daily events in journals, databases and diaries is an attempt to record as effectively as possible the infinite detail of day-to-day life in finite packets of information. To record everything is a lost cause: you can only hope to record the key features of an event. For example, it is a waste of time to write in your diary *The sun rose today.* because this always happens. Far better to take note of the events that distinguish the day from every other.

3.1 Diaries

A *diary*, as defined in this chapter, and perhaps in its normal sense as well, is some system of converting infinite quantities of data that occur on days $1, 2, 3, \dots, i$ into finite diary entries at the end of each day.

Definition 3.1. A *diary* is a height-preserving and order-preserving map (with respect to some chosen basepoints) from an infinite-valence tree to a uniformly locally finite tree.

Remark 3.2. *Since a diary is a graph homomorphism, it is always* 1*-Lipschitz.*

Throughout this chapter, we will assume that the infinite-valence trees have some extra structure; we will assume that they are *sentence-trees* as defined in Definition 0.17. So, for the remainder of this chapter, we suppose that A is a finite alphabet of letters, we suppose that W is the set of finite but non-empty words on A and we suppose that T_W is the associated sentence-tree. We will also assume that the codomain of a diary is a word-tree T_Ω for some finite set Ω.

Example 3.3. Let $\mathrm{Last} : W \to A$ be the function which outputs the last letter of a word $w \in W$. Then there exists a diary $\mathrm{LD} : T_W \to T_A$ which just outputs the last letter of every word in $\alpha \in T_W$. More precisely,

$$\mathrm{LD}(\overline{w_1} \dots \overline{w_i}) = \mathrm{Last}(w_1) \dots \mathrm{Last}(w_i)$$

For example, $\mathrm{LD}(\overline{abc}\ \overline{bc}\ \overline{aa}) = cca$.

The term *diary* comes from the following metaphor. Every day, our protagonist Alice receives some sort of input data (it could be an image, or a voice note, or a text message) which may be very long, in the form of a finite word in W. Every day, she has only a finite amount of time to record the data in her diary, and there are only finite many letters in the English alphabet, and she can only write at a certain pace, and therefore the daily entry in her diary may be summed up as an element of some finite set Ω. Suppose that on days $1, 2, \dots, i$ Alice receives the input data $w_1, w_2, \dots, w_i \in W$ respectively and in response to this data Alice records $v_1, v_2, \dots, v_i \in \Omega$ in her diary. We have thus defined a map $D : T_W \to T_\Omega$ by $D(\overline{w_1} \dots \overline{w_i}) = \overline{v_1} \dots \overline{v_i}$ and it is a diary: it is height-preserving because Alice writes exactly one diary entry per day and it is order-preserving because Alice does not "re-write" diary entries.

We will be interested in understanding whether a given diary $D : T_W \to T_\Omega$ can successfully distinguish between two sentences $\alpha, \beta \in T_W$ (meaning that $D(\alpha) \neq D(\beta)$). As the following example demonstrates, this theory can be applied to understanding when we can upgrade quasiisometric embeddings into products of infinite-valence trees to quasiisometric embeddings into products of binary trees. The following example explains how the theorems in this chapter are used to prove Theorem 4.1 in Chapter 4 and Theorem C.4 in Appendix C.

Example 3.4. Suppose X is a metric space and suppose we have a (λ, μ)-quasiisometric embedding $F : X \to \prod_{q=1}^{Q} T_W$. Let P be a property of *pairs* of sentences in T_W (or, what is the same thing, P is a subset of the product $T_W \times T_W$) and suppose we have a diary $D : T_W \to T_\Omega$ for which there exists a constant $M \geq 1$ such that

$$d_{T_\Omega}(D(\alpha), D(\beta)) \geq \frac{1}{M} d_{T_W}(\alpha, \beta) \tag{3.1}$$

for all pairs α and β satisfying P. We are interested in whether the composition

$$X \xrightarrow{F} \prod_{q=1}^{Q} T_W \xrightarrow{D \times D \times \cdots \times D} \prod_{q=1}^{Q} T_\Omega \tag{3.2}$$

is a quasiisometric embedding. Suppose products are given the L^1 metric.

Let $x, z \in X$ and write $F(x) = (\alpha_1, \alpha_2, \dots, \alpha_Q)$ and $F(z) = (\beta_1, \beta_2, \dots, \beta_Q)$. Then $d(F(x), F(z)) = \sum_{q=1}^{Q} d_{T_W}(\alpha_q, \beta_q)$. Suppose $\mathfrak{q} \in \{1, 2, \dots, Q\}$ is such that $d_{T_W}(\alpha_{\mathfrak{q}}, \beta_{\mathfrak{q}}) = \max_q d_{T_W}(\alpha_q, \beta_q)$. Then $d_{T_W}(\alpha_{\mathfrak{q}}, \beta_{\mathfrak{q}}) \geq \frac{1}{Q} d(F(x), F(z))$.

Suppose we manage to prove that the pair $\alpha_{\mathfrak{q}}, \beta_{\mathfrak{q}} \in T_W$ arising in this manner satisfy the property P. Then I claim that (3.2) would be a quasiisometric embedding. The diary D is 1-Lipschitz and the map F is a quasiisometric embedding so the composition (3.2) is coarsely Lipschitz. For the lower bound of the quasiisometric inequality we have

$$d_{T_\Omega}(D(\alpha_{\mathfrak{q}}), D(\beta_{\mathfrak{q}})) \geq \frac{1}{M} d_{T_W}(\alpha_{\mathfrak{q}}, \beta_{\mathfrak{q}}) \geq \frac{1}{QM} d(F(x), F(z)) \geq \frac{1}{\lambda Q M} d_X(x, z) - \mu$$

where the first lower bound follows from the fact that $\alpha_\mathfrak{q}$ and $\beta_\mathfrak{q}$ satisfy P. As Ω is finite, T_Ω is quasiisometric to a binary tree (assuming $|\Omega| > 1$).

In the remainder of this chapter, we will provide four criteria (♈), (♌), (♍) and (♉) on pairs of sentences, analogous to the property P in Example 3.4, which imply that a certain diary D coarsely preserves the metric on those sentences as in (3.1). The ideas and definitions in this chapter were created in order to quasiisometrically embed relatively hyperbolic groups into products of binary trees as described in Theorem 4.1. However, since all the definitions in this chapter are independent of the work in Chapter 4 and since they have a different mathematical flavour, I thought it would help the clarity of this thesis by placing them in a separate chapter. Possibly, it also makes this work more accessible to anyone who might want to use it in the future.

Remark 3.5. *I do not know how to approach the situation described in Example 3.4 without assuming that the infinite-valence tree is a sentence-tree. It is for this reason that it is necessary in the proof of Theorem 4.1 to find a quasiisometric embedding of G into a product of sentence-trees, as opposed to some abstract infinite-valence trees. This is why the quasiisometric embedding in Theorem 4.70 needs to be chosen very carefully.*

The ideas behind these criteria stem from the work of Buyalo, Dranishnikov and Schroeder [BDS07] on *Alice's Diary.* Alice's Diary allows one to quasiisometrically embed subsets of products of infinite-valence trees into products of uniformly bounded valence trees in certain circumstances. However, the results that Buyalo, Dranishikov and Schroeder prove on Alice's Diary (in particular [BDS07, Proposition 5.5]) are phrased in a rather *ad hoc* manner. It was my aim to unravel the ideas and theorems which are latent in their work, and to provide easy to use and flexible criteria so that one may find quasiisometric embeddings into products of binary trees in many contexts. The definitions that are created to achieve this aim, i.e. diaries, finite statistics and linear statistics, are quite abstract. Indeed, a reader solely interested in the proof of the main result of Chapter 4 might reasonably ask whether so many abstract definitions were needed for the proof of this one theorem. However I can guarantee a concerned reader that although an *ad hoc* proof, forged by bending Alice's Diary to the specific case of relatively hyperbolic groups, might have made for a marginally shorter proof, it would have been less coherent and the ideas involved would have been buried. Earlier drafts of this work can attest to this. An *ad hoc* proof could only be understood on a line-by-line basis, as if we were staring at a painting with the tip of our nose a centimetre from the canvas. A more theoretical approach allows for the broader strokes of the proof to be seen.

I would also like to express my appreciation of Buyalo, Dranishnikov and Schroeder's paper. They chose to use metaphor to help the reader understand the Alice's Diary function, and it was this metaphor and the questions it raised that sparked my curiosity on the topic.

3.2 Finite statistics

Definition 3.6. A *finite statistic* is a function $\mathrm{stat} : T_W \to \Omega$ where Ω is some finite set. If we wish to specify the domain of stat we say that it is a finite statistic *on* T_W.

You might think of $\mathrm{stat}(\alpha)$ as being a fact about α that contains a uniformly bounded amount of information. It is a *statistic* in the sense that it is a value which can be calculated from some input data α. We can use finite statistics to create diaries as follows.

Definition 3.7. Suppose we have a finite statistic $\mathrm{stat} : T_W \to \Omega$ and suppose T_Ω is the word-tree on Ω. Then the diary $D : T_W \to T_\Omega$ defined by

$$D(\overline{w_1}\ \overline{w_2}\dots\overline{w_i}) = \mathrm{stat}(\overline{w_1})\mathrm{stat}(\overline{w_1}\ \overline{w_2})\dots\mathrm{stat}(\overline{w_1}\ \overline{w_2}\dots\overline{w_i})$$

is the *diary associated to stat* : $T_W \to \Omega$.

Example 3.8. Let $\kappa \in \mathbb{N}$ be a constant and let Ω_κ denote the set of all words in W of length at most κ. Consider the map $\mathrm{trunc}_\kappa : T_W \to \Omega_\kappa$ which is such that $\mathrm{trunc}_\kappa(\overline{w_1}\ \overline{w_2}\dots\overline{w_i})$ is equal to the final κ letters of the word w_i. Then trunc_κ is a finite statistic on T_W. The associated diary simply returns the final κ letters of each word in the sentence. If the letters in A are thought of as *events*, and the words in W as *sequences of events that occur on a given day*, then this diary is the one in which Alice records at the end of the day the κ most recent events.

Hypothetically, Alice could be more imaginative in her choice of diary.

Example 3.9. Suppose $A = \{0, 1\}$ and let Ω be the set $\{n/10^5 : 0 \leq n \leq 10^5\}$. Let $\mathrm{av}_5 : T_W \to \Omega$ be defined so that $\mathrm{av}_5(\overline{w_1}\dots\overline{w_i})$ is the element of Ω closest to the the average value of the string $w_1w_2\dots w_i$, rounding up if necessary. Let $D : T_W \to T_\Omega$ be the associated diary. In other words, on day i Alice records the average of $w_1w_2\dots w_i$ up to 5 decimal places.

We may now provide our first criterion for when there exists a diary that coarsely preserves a pair of sentences.

Let $\mathcal{S}_{\mathrm{finite}}$ be a finite collection of finite statistics on the same sentence-tree T_W (but the finite statistics are allowed to have different codomains). Let $0 \leq \delta < 1$, $J \in \mathbb{N}$ be constants. Consider the following property, pronounced *Aries*, that a pair of sentences α and β in the forms described by Notation 0.18 might possess.

Definition 3.10. We say that α, β satisfy $\Upsilon(\mathcal{S}_{\text{finite}}, \delta, J)$ if there exists some $1 \leq j \leq \delta \min(m, n) + J$ and $\text{stat} \in \mathcal{S}_{\text{finite}}$ such that

$$\text{stat}(\overline{u_1} \dots \overline{u_p}\ \overline{u_{p+1}} \dots \overline{u_{p+j}}) \neq \text{stat}(\overline{u_1} \dots \overline{u_p}\ \overline{u'_{p+1}} \dots \overline{u'_{p+j}})$$

Property (Υ) is saying that there exists a *fact* about α and β which distinguishes them and which only requires a finite amount of information to state. The purpose of the constants δ and J is to ensure that this distinguishing fact separates α and β sufficiently close to their lowest common ancestor $\overline{u_1} \dots \overline{u_p}$.

Theorem 3.11. *There exists a finite set $\Omega = \Omega(\mathcal{S}_{\mathit{finite}})$, a diary $D = D(\mathcal{S}_{\mathit{finite}})$ which is a function $D : T_W \to T_\Omega$, and a constant $M = M(\delta, J) \geq 1$, such that*

$$\frac{1}{M} d_{T_W}(\alpha, \beta) \leq d_{T_\Omega}(D\alpha, D\beta) \leq d_{T_W}(\alpha, \beta)$$

for all α, β satisfying $\Upsilon(\mathcal{S}_{\mathit{finite}}, \delta, J)$.

Proof. Write $\mathcal{S}_{\text{finite}} = \{\text{stat}^k : T_W \to \Omega^k : 1 \leq k \leq K\}$. We begin by combining all the finite statistics in $\mathcal{S}_{\text{finite}}$ into a single finite statistic $\text{STAT} : T_W \to \Omega$. Set $\Omega = \prod_{k=1}^{K} \Omega^k$. We define $\text{STAT} : T_W \to \Omega$ by

$$\text{STAT}(\overline{w_1} \dots \overline{w_i}) = \begin{pmatrix} \text{stat}^1(\overline{w_1} \dots \overline{w_i}) \\ \text{stat}^2(\overline{w_1} \dots \overline{w_i}) \\ \vdots \\ \text{stat}^K(\overline{w_1} \dots \overline{w_i}) \end{pmatrix}$$

Let $D : T_W \to T_\Omega$ be the diary associated to STAT. More explicitly

$$D(\overline{w_1} \dots \overline{w_i}) = \text{STAT}(\overline{w_1})\text{STAT}(\overline{w_1}\ \overline{w_2}) \dots \text{STAT}(\overline{w_1}\ \overline{w_2} \dots \overline{w_i})$$

Suppose α, β satisfy $\Upsilon(\mathcal{S}_{\text{finite}}, \delta, J)$. Then there exists $1 \leq j \leq \delta \min(m, n) + J$ and $1 \leq k \leq K$ such that $\text{stat}^k(\overline{u_1} \dots \overline{u_p} \dots \overline{u_{p+j}}) \neq \text{stat}^k(\overline{u_1} \dots \overline{u_p} \dots \overline{u'_{p+j}})$. It follows that $\text{STAT}(\overline{u_1} \dots \overline{u_p} \dots \overline{u_{p+j}}) \neq \text{STAT}(\overline{u_1} \dots \overline{u_p} \dots \overline{u'_{p+j}})$ and so

$$d_{T_\Omega}(D\alpha, D\beta) \geq (p + m) - (p + j - 1) + (p + n) - (p + j - 1) = m + n - 2j + 2$$

We have two cases: either $m + n < 4J/(1 - \delta)$ or $m + n \geq 4J/(1 - \delta)$. In the first case, we have

$$d_{T_\Omega}(D\alpha, D\beta) \geq m + n - 2j + 2 \geq 2 = \frac{1 - \delta}{2J} \frac{4J}{1 - \delta} \geq \frac{1 - \delta}{2J}(m + n)$$

In the second case, we have

$$d_{T_\Omega}(D\alpha, D\beta) \geq m+n-2j+2 \geq m+n-\delta(m+n)-2J \geq \frac{1-\delta}{2}(m+n)$$

So if we set $M = 2J/(1-\delta)$ then we are done. □

So we may now address the situation described in Example 3.4: if the sentences $\alpha_{\mathsf{q}}, \beta_{\mathsf{q}} \in T_W$ satisfy $\Upsilon(\mathcal{S}_{\text{finite}}, \delta, J)$ for some uniform choices of $\mathcal{S}_{\text{finite}}$, δ and J, then the diary $D = D(\mathcal{S}_{\text{finite}})$ given by Theorem 3.11 will be such that (3.2) is a quasiisometric embedding.

Definition 3.12. We say that $\alpha, \beta \in T_W$ satisfy $♌(\mathcal{S}_{\text{finite}}, J)$ if they satisfy $\Upsilon(\mathcal{S}_{\text{finite}}, 0, J)$. This property is pronounced *Leo*.

In the next section, we are going to introduce Buyalo, Dranishnikov and Schroeder's *Alice's Diary*. Afterwards, in Section 3.4, Alice's Diary will be used to create diaries that are more powerful than the ones which are made via finite statistics.

3.3 Alice's Diary

Let us add some more texture to our diary metaphor. We refer to a letter chosen from the finite alphabet A as an *event*. On each day, several events occur in some order (thereby forming a word in W), and at the end of each day Alice dedicates $\kappa \in \mathbb{N}$ minutes to diary writing. Alice can write down one letter per minute and so records at most κ letters in her diary each day. On each page she writes down only one letter, and so on each day she fills in at most κ pages of her diary. The pages of her diary corresponding to a single day form a *chapter*. Alice wants to record the events on days $1, 2, 3, 4, 5, 6, 7, \ldots$ to the best of her ability. She has perfect memory of past events but seems to always be paranoid that one day she might forget everything and only her diary will be left behind to remind her of what occurred.

If Ω_κ is the set of all words of length at most κ on the alphabet A, then Alice's diary writing process defines a diary $D : T_W \to T_{\Omega_\kappa}$ (assuming she responds deterministically to possible sequences of events). A vertex in T_W corresponds to some possible sequence of events over the course of several days, and a vertex in T_{Ω_κ} corresponds to a possible sequence of chapters in the diary.

The diary described in Example 3.8, in which Alice simply records the final κ events of the day provides an example of this sort of diary. However, this truncation diary might be very inefficient.

Example 3.13. Let $\text{trunc}_\kappa : T_W \to \Omega_\kappa$ be the truncation statistic from Example 3.8 and suppose $\kappa = 4$. Let $D : T_W \to T_{\Omega_\kappa}$ denote the associated truncation diary. Suppose we have some sentence $\alpha = \overline{w_1}\ \overline{w_2} \dots \overline{w_i} \in T_W$ such that half of the words in α have length 2, and the other half of the words in α have length 8. Then the diary $D : T_W \to T_{\Omega_4}$ described above is very wasteful: on the days when only two events occur Alice records only two letters in her diary as opposed to her maximum of 4 - she could have used that spare time to write about previous days on which eight events occurred! And indeed, the total amount of wastefulness (or viewed another way, the potential for gains in efficiency) is in the order of i where i is the length of the word α.

The inefficiency of the truncation diary motivates the definition of *Alice's Diary.* Alice's Diary was invented by Buyalo, Dranishnikov and Schroeder [BDS07].

Definition 3.14. Alice's Diary $\text{AD}_\kappa : T_W \to T_{\Omega_\kappa}$ is characterised by the following rule: *Alice always records the most recent unrecorded event first.*

It is easiest to understand Alice's Diary through an example. I have also provided a formal and non-metaphorical definition in the appendix, see Definition A.1.

Example 3.15. Let $A = \{a, b, c\}$ and $\kappa = 3$. Suppose that on day one, at first event a occurs, then event b, then event a again and then event c. So the associated word is $abac$. Since $\kappa = 3$, Alice only has time to write down 3 letters at the end of the day, and so she records cab. That's because event c occured most recently, then event a, then event b. Suppose that on day two, the associated word is cb. Then Alice writes down in her diary bca. The final a she writes down was the first event of day one. Now suppose that on days three, four and five, the associated events are $accc$, $bcbc$ and a respectively. Then the chapters that Alice records in her diary at the end of days three, four and five are ccc, cbc, aba. In the chapter corresponding to day five, the first page records an event from day five, the second page records an event from day four, and the third page records an event from day three. So we have shown that $\text{AD}_3 : T_W \to T_\Omega$ satisfies $\text{AD}_3(\overline{abac}\ \overline{cb}\ \overline{accc}\ \overline{bcbc}\ \overline{a}) = \overline{cab}\ \overline{bca}\ \overline{ccc}\ \overline{cbc}\ \overline{aba}$.

So Alice's Diary is a greedy algorithm for creating diaries.

We would now like to understand how much information AD_κ captures about the original sentence. What does this mean? One interpretation is that we would like to understand when $\alpha \neq \beta \implies \text{AD}_\kappa(\alpha) \neq \text{AD}_\kappa(\beta)$. The statement that $\alpha \neq \beta$ roughly amounts to saying that the two sentences have letters which are in the same position yet correspond to a different element of A. You might call these "clashing" letters. So we would like to understand when the existence of clashing letters implies that $\text{AD}_\kappa(\alpha) \neq \text{AD}_\kappa(\beta)$. If we take the contrapositive of this statement, we find ourselves studying whether $\mathrm{AD}_\kappa(\alpha) = \mathrm{AD}_\kappa(\beta)$ implies that two letters in the same position in α, β must be the same element of A. Hopefully this justifies why the conclusion of Theorem 3.20 below is significant.

But first, we want to assume that our sentences α have some further structure.

Definition 3.16. Suppose that our alphabet A contains a special character $\star$. We say that a word $u \in W$ is *starred* if $u = \star w$ for some word $w \in W$ that does not contain the letter $\star$. We say that a sentence $\alpha \in T_W$ is *starred* if all the constituent words of α are starred. Maybe $\star$ corresponds to the event of sunrise.

The primary reason to work with starred words is so that you can say the following: if u and u' are distinct starred words then there exist distinct letters $a \in u$ and $a' \in u'$ with the same distance to the end of their words.

In Lemma 3.17, Lemma 3.18, Lemma 3.19 and Theorem 3.20 that follow, we will always assume that we have two starred sentences $\alpha = \overline{\star u_1}\,\overline{\star u_2} \dots \overline{\star u_m}$ and $\beta = \overline{\star u'_1}\,\overline{\star u'_2} \dots \overline{\star u'_n}$ in T_W and that $\mathrm{AD}_\kappa(\alpha) = \overline{v_1}\,\overline{v_2} \dots \overline{v_m}$ and $\mathrm{AD}_\kappa(\beta) = \overline{v'_1}\,\overline{v'_2} \dots \overline{v'_n}$.

Lemma 3.17. *Suppose the diary chapter v_i has the form $v_i = u \star v$ where u does not contain the letter $\star$. Then $u_i = \overleftarrow{u}$.*

Proof. At the end of day i, Alice sets out writing her diary. She starts by recording the events of day i. If she does not fully record the events of day i, then there can be no $\star$ in that day's chapter. Since $v_i = u \star v$, it follows that Alice records day i in its entirety. Since Alice always records the most recent events first, her recording of day i must come at the start of the word v_i and the $\star$ in $\star u_i$ must be the first $\star$ of chapter i. It follows that $u_i = \overleftarrow{u}$. □

Lemma 3.18. *Suppose that $AD_\kappa(\alpha) = AD_\kappa(\beta)$. If a word $\star u_i$ has length at most κ then $u'_i = u_i$. Similarly, if a word $\star u'_j$ has length at most κ then $u_j = u'_j$.*

Proof. Suppose $\star u_i$ has length at most κ. It follows that v_i is of the form $\overleftarrow{u_i} \star v$ for some word v. Since $v'_i = v_i$, by Lemma 3.17 we deduce that $u'_i = u_i$. The case when $\star u'_j$ has length at most κ follows similarly. □

Lemma 3.19. *Let $i \leq \min(m, n)$ and suppose $a \in \star u_i$ and $a' \in \star u'_i$ are letters such that*

1. *a, a' have the same distance from the end of their words;*

2. *there exists $i \leq j \leq \min(m, n)$ and $1 \leq k \leq \kappa$ such that event a is recorded on page k of diary chapter v_j;*

3. *u_l has the same length as u'_l for $i + 1 \leq l \leq j$.*

Then a' is recorded on page k of diary chapter v'_j

Proof. Write $\star u_i = vaw$ and $\star u'_i = v'a'w'$. The content of the first condition is that w has the same length as w'. The position of a within the diary words $\overline{v_i}\ \overline{v_{i+1}} \dots \overline{v_j}$ depends only on the lengths of the words $w, u_{i+1}, \dots, u_j$. Similarly, the position of a' within the diary words $\overline{v'_i}\ \overline{v'_{i+1}} \dots \overline{v'_j}$ depends only on the lengths of the words $w', u'_{i+1}, \dots, u'_j$. It follows that a and a' appear on the same page of v_j and v'_j respectively. $\square$

Theorem 3.20. *Suppose that $AD_\kappa(\alpha) = AD_\kappa(\beta)$. Let a be a letter in the word $\star u_i$ and let a' be a letter in the word $\star u'_i$ and suppose*

1. *a, a' have the same distance from the end of their words;*
2. *The events a and a' are recorded in the diaries of α and β.*

Then there exists $i \leq j \leq m$ and $1 \leq k \leq \kappa$ such that a, a' appear on page k of chapters v_j and v'_j respectively. Consequently, a, a' correspond to the same element of A.

Proof. Write $\star u_i = vaw$ and $\star u'_i = v'a'w'$. The content of the first condition on a, a' is that w has the same length as w'. If a is recorded within v_i then it follows that aw and $a'w'$ both have length at most κ and the conclusion clearly holds. So assume that a is recorded within v_j for some $i+1 \leq j \leq m$. Then clearly $\star u_j$ has length at most κ and so Lemma 3.18 implies that $u_j = u'_j$. We will prove by induction that $u_l = u'_l$ for $i+1 \leq l \leq j$. So fix $i+1 \leq l \leq j$ and suppose $u_{l+1} = u'_{l+1}, u_{l+2} = u'_{l+2}, \dots, u_j = u'_j$. Looking for a contradiction, suppose that $u_l \neq u'_l$. Then we can find distinct letters $b \in \star u_l$ and $b' \in \star u'_l$ with the same distance to the end of their words. We also know that b must appear in one of the diary words $v_l, v_{l+1}, \dots, v_j$ because for Alice to record event a in chapter v_j, she must already have recorded event b on some earlier page of the diary. It then follows from Lemma 3.19, and the fact that the diaries are equal, that b and b' correspond to the same element of A, which is a contradiction. Thus $u_l = u'_l$.

So we have shown that $u_l = u'_l$ for all $i+1 \leq l \leq j$. We can now apply Lemma 3.19 again to conclude that a and a' appear on the same page of $v_j = v'_j$ and thus are equal. $\square$

In my view, the above theorem is the fundamental observation on Alice's Diary. To paraphrase it: if you know two diaries are equal, and if you have two events that are

- recorded in the diary;
- occur on the same day;
- are in the same position (with respect to the ends of their words);

then those two events correspond to the same element of A.

It leads us naturally to trying to understand when we can guarantee that an event is recorded in the diary. We need some further definitions.

Definition 3.21. Suppose we have a sentence $\alpha = \overline{u_1}\,\overline{u_2}\dots\overline{u_m} \in T_W$ and suppose a is some letter in u_i. We can write $u_i = vaw$ where v and w are some words on the alphabet A. Then the sentence

$$\overline{aw}\,\overline{u_{i+1}}\dots\overline{u_m}$$

is the *tail-sentence* of a. Similarly, the sentence

$$\overline{u_1}\dots\overline{u_{i-1}}\,\overline{va}$$

is the *head-sentence* of a.

In terms of our analogy, the tail-sentence of a consists of a and every event that happens after a. The head-sentence of a consists of a and every event that happens before a.

Definition 3.22. The *average word length (AWL)* of a sentence $\alpha = \overline{u_1}\,\overline{u_2}\dots\overline{u_m} \in T_W$ is the amount of letters in $u_1u_2\dots u_m$ divided by m. We will often write $\mathrm{AWL}(\alpha)$ to denote the average word length of α.

Proposition 3.23. *Suppose $\alpha = \overline{u_1}\,\overline{u_2}\dots\overline{u_m} \in T_W$ and let a be some letter in u_i. Suppose that the tail-sentence of a has average word length $N \geq 1$. If $\kappa \geq N$ then a is recorded in $AD_\kappa(\alpha)$.*

Proof. We will prove the contrapositive. Suppose that a is not recorded in $\mathrm{AD}_\kappa(\alpha)$. Write $\mathrm{AD}_\kappa(\alpha) = \overline{v_1}\,\overline{v_2}\dots\overline{v_m}$. A chapter only has fewer than κ pages when Alice has successfully recorded all past events already. It follows that v_i, v_{i+1}, …, v_m all have κ pages. Since Alice would only record an event that happened prior to a if a had already been recorded, we know that every page of the chapters v_i, v_{i+1}, …, v_m corresponds to an event that happened after a. It follows that the tail-sentence of a has at least $\kappa(m-i+1)+1$ letters. It consists of $m-i+1$ words and hence the average word length of the tail-sentence is greater than κ. □

Corollary 3.24. *Suppose α, β are starred sentences in T_W. Write*

$$\alpha = \overline{u_1}\,\overline{u_2}\dots\overline{u_p}\,\overline{u_{p+1}}\dots\overline{u_{p+m}}$$

and

$$\beta = \overline{u_1}\,\overline{u_2}\dots\overline{u_p}\,\overline{u'_{p+1}}\dots\overline{u'_{p+n}}$$

where $u_{p+1} \neq u'_{p+1}$. *Write* $AD_\kappa(\alpha) = \overline{v_1}\ \overline{v_2}\ \dots\ \overline{v_{p+m}}$ *and* $AD_\kappa(\beta) = \overline{v'_1}\ \overline{v'_2}\ \dots\ \overline{v'_{p+n}}$. *Suppose we have letters* $a \in u_{p+j}$ *and* $a' \in u'_{p+j}$ *which satisfy the following*

- a *and* a' *correspond to different elements of the alphabet* A*;*
- a *and* a' *are at equal distance from the ends of their words;*
- *the AWL of the tail-sentence of* a *is* N *and the AWL of the tail-sentence of* a' *is* N'*;*

Let $i \in \mathbb{N}_0$ *be such that* $p + j + i \leq p + \min(m, n)$. *If* $\kappa \geq N\frac{m-j+1}{i+1}$ *and* $\kappa \geq N'\frac{n-j+1}{i+1}$ *then*

$$d(AD_\kappa(\alpha), AD_\kappa(\beta)) \geq d(\alpha, \beta) - 2j - 2i$$

Proof. Suppose the amount of letters in the tail-sentence of a is l and the amount of letters in the tail-sentence of a' is l'. Then $N = \frac{l}{m-j+1}$ and $N' = \frac{l'}{n-j+1}$. Consider the truncated sentences $\alpha_{p+j+i} = \overline{u_1} \dots \overline{u_{p+j+i}}$ and $\beta_{p+j+i} = \overline{u_1} \dots \overline{u'_{p+j+i}}$. The AWL of the tail-sentence of $a \in \alpha_{p+j+i}$ is at most $\frac{l}{i+1} = N\frac{m-j+1}{i+1}$. So, by Proposition 3.23, if $\kappa \geq N\frac{m-j+1}{i+1}$ then a is recorded in $\mathrm{AD}_\kappa(\alpha_{p+j+i})$. Similarly, if $\kappa \geq N'\frac{n-j+1}{i+1}$ then a' is recorded in $\mathrm{AD}_\kappa(\beta_{p+j+i})$. It then follows from Theorem 3.20 that $\mathrm{AD}_\kappa(\alpha_{p+j+i}) \neq \mathrm{AD}_\kappa(\beta_{p+j+i})$ and the corollary follows. ☐

3.4 Linear statistics

Theorem 3.11 provides a way of roughly preserving the metric on pairs of sentences $\alpha, \beta \in T_W$ (of the forms described in Notation 0.18) when there is a uniform bound L on the length of the words $u_{p+1}, u_{p+2}, \dots, u_{p+m}$ and $u'_{p+1}, u'_{p+2}, \dots, u'_{p+n}$. When this is the case, we may choose the finite statistic $\mathrm{trunc}_L : T_W \to \Omega_L$ and the constants $\delta = 0$, $J = 1$. With these choices, the sentences α, β will satisfy $\Upsilon(\mathrm{trunc}_L, 0, 1)$. However, in the "real-life" situations that arise from attempting to quasiisometrically embed groups into products of binary trees, we will need to deal with situations in which there is no uniform bound on the words $u_{p+1}, u_{p+2}, \dots, u_{p+m}, u'_{p+1}, u'_{p+2}, \dots, u'_{p+n}$. However, we will often find that these words are short *on average*. It is for these situations that we will need linear statistics and their associated diaries.

Intuitively, where a finite statistic is some uniformly bounded quantity that can be calculated from a sentence $\overline{w_1}\ \overline{w_2} \dots \overline{w_i}$, a linear statistic is a quantity that can be calculated from $\overline{w_1}\ \overline{w_2} \dots \overline{w_i}$ that contains a linear amount of information with respect to some variable $c \in \mathbb{N}_0$.

Definition 3.25. Suppose, as ever, we have a finite alphabet A and an associated sentence-tree T_W. Suppose we have another, potentially distinct, finite alphabet B and let T_B be the word-tree on B. Given a number $c \in [0, \infty]$ let Ω_c denote all words on the alphabet B of length at most c. So we can identify the sets Ω_c with subsets of the vertices of T_B. A *linear statistic* on T_W consists of a constant $\tau \geq 1$ and a collection of finite statistics $\{\text{stat}_c : T_W \to \Omega_{\tau c} : c \in \mathbb{N}_0\}$ such that if $c \leq c'$ then $\text{stat}_{c'}(\overline{w_1} \dots \overline{w_i})$ descends from $\text{stat}_c(\overline{w_1} \dots \overline{w_i})$ in T_B. We will typically denote a linear statistic just by $\text{stat}_c : T_W \to \Omega_{\tau c}$ as opposed to $\{\text{stat}_c : T_W \to \Omega_{\tau c} : c \in \mathbb{N}_0\}$. Associated to every linear statistic $\text{stat}_c : T_W \to \Omega_{\tau c}$ is an *infinite statistic* $\text{stat}_\infty : T_W \to \Omega_\infty$ defined by $\text{stat}_\infty(\overline{w_1} \dots \overline{w_i}) = \lim_{c \to \infty} \text{stat}_c(\overline{w_1} \dots \overline{w_i})$.

So $\text{stat}_c(\overline{w_1} \dots \overline{w_i})$ is a collection of facts about the sentence $\overline{w_1} \dots \overline{w_i}$ that contain steadily more information as c increases. Since the image of stat_c is in $\Omega_{\tau c}$, the amount of information contained in the word $\text{stat}_c(\overline{w_1} \dots \overline{w_i})$ is comparable in size to c. If $\text{stat}_{c'}(\overline{w_1} \dots \overline{w_i})$ descends from $\text{stat}_c(\overline{w_1} \dots \overline{w_i})$ that means that the former contains at least as much information as the latter. This happens when when $c \leq c'$.

Let's now consider some examples of linear statistics.

Example 3.26. Let $B = A$. If we define $\text{trunc}_c(\overline{w_1} \dots \overline{w_i})$ to be the final τc letters of the word u_i written backwards (i.e. starting from the final letter) then $\text{trunc}_c : T_W \to \Omega_{\tau c}$ is a linear statistic.

Example 3.27. Let $B = A$. Let $l = \text{length}(w_1 w_2 \dots w_i)$. Let $a_1, a_2, \dots, a_l$ denote all the letters in the word $w_1 w_2 \dots w_i$. Let $\sigma \in S_l$ be a permutation which we call the *order of priority*. We can define $\text{oop}(\sigma)_c : T_W \to \Omega_{\tau c}$ by letting $\text{oop}(\sigma)_c(\overline{w_1} \dots \overline{w_i})$ be the first τc letters of the word $a_{\sigma(1)} a_{\sigma(2)} \dots a_{\sigma(l)}$. Then $\text{oop}(\sigma)$ is a linear statistic.

Example 3.28. Let $B = \{0\}$. We can define $\text{howmany}_c : T_W \to \Omega_c$ by setting

$$\text{howmany}_c(\overline{w_1} \dots \overline{w_i}) = \begin{cases} \emptyset & \text{if length}(w_i) > c \\ 0 & \text{if length}(w_i) \leq c \end{cases}$$

Then howmany_c is a linear statistic. You might even just identify howmany_c with the *question* "does w_i have at most c letters?".

Non-example 3.29. Let $B = A$. Let $\overline{w_1} \dots \overline{w_i} \in T_W$ and let $a_1, a_2, \dots, a_l$ denote all the letters in the word $w_1 w_2 \dots w_i$. Suppose we define $\text{periodic}_c : T_W \to \Omega_c$ by setting $\text{periodic}_c(\overline{w_1} \dots \overline{w_i}) = a_1 a_{\lfloor l/c \rfloor} a_{2\lfloor l/c \rfloor} a_{3\lfloor l/c \rfloor} \dots a_{c\lfloor l/c \rfloor}$. So $\text{periodic}_c(\overline{w_1} \dots \overline{w_i})$ outputs the word formed by taking every $\lfloor l/c \rfloor$'th letter of the word $w_1 w_2 \dots w_i$. One might reasonably think that this is a linear statistic on T_W since $\text{periodic}_c(\overline{w_1} \dots \overline{w_i})$ is a word that contains c letters. But it fails to satisfy the defining condition of a linear statistic: there is no reason to think that $\text{periodic}_{c'}(\overline{w_1} \dots \overline{w_i})$ will descend from $\text{periodic}_c(\overline{w_1} \dots \overline{w_i})$ when $c' > c$.

We now come to our second criterion. Let $\mathcal{S}_{\text{linear}}$ be a finite collection of linear statistics on the same tree T_W (but the linear statistics are allowed to have different codomains). Let $0 \leq \delta < 1$, $J \in \mathbb{N}$, $N > 0$, $\epsilon > 0$ be constants. Consider the property, pronounced *Virgo*, that a pair of vertices $\alpha, \beta \in T_W$ of the form described in Notation 0.18 might possess. It is somewhat analogous to (♈).

Definition 3.30. We say that α, β satisfy ♍$(\mathcal{S}_{\text{linear}}, \delta, J, N, \epsilon)$ if there exists some $1 \leq j \leq \delta \min(m, n) + J$ and $\text{stat}_c \in \mathcal{S}_{\text{linear}}$ such that

(♍1) $\text{AWL}(\overline{u_{p+j+1}} \dots \overline{u_{p+m}})$ and $\text{AWL}(\overline{u'_{p+j+1}} \dots \overline{u'_{p+n}})$ are both at most N;

(♍2) $\text{length}(u_{p+j}) \geq \epsilon(m+n)$ or $\text{length}(u_{p+j}) \geq \epsilon(m+n)$ or $u_{p+j} \neq u'_{p+j}$;

(♍3) $\text{stat}_{m+n}(\overline{u_1} \dots \overline{u_p}\ \overline{u_{p+1}} \dots \overline{u_{p+j}}) \neq \text{stat}_{m+n}(\overline{u_1} \dots \overline{u_p}\ \overline{u'_{p+1}} \dots \overline{u'_{p+j}})$.

(♍) is saying that there exists some fact about α and β which distinguishes them. However, unlike in the case of (♈) and (♌), this fact can now contain an unbounded amount of information, as long as the amount of information is at most linear with respect to $m + n = d_{T_W}(\alpha, \beta)$. The cost of this upgrade is that you very much need the AWL of the tail-sentences to be uniformly bounded.

Theorem 3.31. *There exists a finite set* $\Omega = \Omega(\mathcal{S}_{\textit{linear}}, \delta, J, N, \epsilon)$, *a diary* $D = D(\mathcal{S}_{\textit{linear}}, \delta, J, N, \epsilon)$ *which is a function* $D : T_W \to T_\Omega$, *and a constant* $M = M(\delta, J) \geq 1$, *such that*

$$\frac{1}{M} d_{T_W}(\alpha, \beta) \leq d_{T_\Omega}(D\alpha, D\beta) \leq d_{T_W}(\alpha, \beta)$$

for all α, β *satisfying* ♍$(\mathcal{S}_{\textit{linear}}, \delta, J, N, \epsilon)$.

The theorem is powerful for this reason: when $m + n = d_{T_W}(\alpha, \beta)$ is very large, $\text{stat}_{m+n}(\overline{u_1} \dots \overline{u_p}\ \overline{u_{p+1}} \dots \overline{u_{p+j}})$ and $\text{stat}_{m+n}(\overline{u_1} \dots \overline{u_p}\ \overline{u'_{p+1}} \dots \overline{u'_{p+j}})$ contain vast amounts of information about our sentences and so (♍3) is more likely to hold.

Proof of Theorem 3.31. Now,

- write $\mathcal{S}_{\text{linear}} = \{\text{stat}^k_c : 1 \leq k \leq K\}$;
- let $\tau^k \geq 1$ be the constant associated to the linear statistic stat^k_c;
- let B^k denote the finite alphabet associated to the linear statistic stat^k_c;
- let stat^k_∞ denote the infinite statistic associated to stat^k_c;
- let $\tau = \max_k \tau^k$;
- let ω be the smallest natural number which is at least τ/ϵ;

- let $\star$ and $\bigstar$ be two special letters.

Let $u = b_1 b_2 b_3 \dots$ be a word on some alphabet, and let $l = \text{length}(u)$. For $r \in \mathbb{N}$, we define

$$\text{norm}_r(u) = \begin{cases} b_1 b_2 \dots b_r & r \leq l \\ b_1 b_2 \dots b_l (\star)^{r-l} & r > l \end{cases}$$

So $\text{norm}_r(u)$ is the same as u except it has been either truncated or otherwise extended by $\star$'s so that it has exactly r letters (the word has been "normalised"). Recall that $\overleftarrow{u}$ is the word u written backwards.

If $\overline{w_1} \dots \overline{w_i}$ is an element of T_W then we define a function $\widehat{\text{stat}^k_\infty}$ with domain T_W by

$$\widehat{\text{stat}^k_\infty}(\overline{w_1} \dots \overline{w_i}) = \overleftarrow{\text{norm}_{\omega\text{length}(w_i)}(\text{stat}^k_\infty(\overline{w_1} \dots \overline{w_i}))}$$

So $\widehat{\text{stat}^k_\infty}(\overline{w_1} \dots \overline{w_i})$ is a word on the alphabet $B^k \cup \star$ of length exactly $\omega\text{length}(w_i)$. Let W' denote the set of finite words on the finite alphabet

$$A' = \bigstar \cup \left(A \times \prod_{k=1}^{K} (B^k \cup \star) \right)$$

Then we have a map $I : T_W \to T_{W'}$ given by

$$I(\overline{w_1} \dots \overline{w_i}) = \bigstar \overline{\begin{pmatrix} (w_1)^\omega \\ \widehat{\text{stat}^1_\infty}(\overline{w_1}) \\ \widehat{\text{stat}^2_\infty}(\overline{w_1}) \\ \vdots \\ \widehat{\text{stat}^K_\infty}(\overline{w_1}) \end{pmatrix}} \bigstar \overline{\begin{pmatrix} (w_2)^\omega \\ \widehat{\text{stat}^1_\infty}(\overline{w_1}\ \overline{w_2}) \\ \widehat{\text{stat}^2_\infty}(\overline{w_1}\ \overline{w_2}) \\ \vdots \\ \widehat{\text{stat}^K_\infty}(\overline{w_1}\ \overline{w_2}) \end{pmatrix}} \dots \bigstar \overline{\begin{pmatrix} (w_i)^\omega \\ \widehat{\text{stat}^1_\infty}(\overline{w_1}\ \overline{w_2} \dots \overline{w_i}) \\ \widehat{\text{stat}^2_\infty}(\overline{w_1}\ \overline{w_2} \dots \overline{w_i}) \\ \vdots \\ \widehat{\text{stat}^K_\infty}(\overline{w_1}\ \overline{w_2} \dots \overline{w_i}) \end{pmatrix}}$$

It can be checked that I is an isometric embedding: evidently it is height-preserving and order-preserving, and it is injective since if $I(\overline{w_1} \dots \overline{w_i}) = I(\overline{w'_1} \dots \overline{w'_i})$ then $(w_l)^\omega = (w'_l)^\omega$ for all $1 \leq l \leq i$ and so $w_l = w'_l$ for all $1 \leq l \leq i$.

Let

$$U = \frac{12\tau J}{1-\delta} + \omega N + 1 \quad \text{and} \quad V = \frac{12(\tau+\epsilon)J}{1-\delta} + \omega N + 1$$

and let κ be the smallest natural number that is larger than $\frac{16U}{1-\delta}$ and $\frac{64J\tau}{1-\delta}$ and $\frac{16V}{1-\delta}$ and $\frac{64J(\tau+\epsilon)}{1-\delta}$. Let $\text{AD}_\kappa : T_{W'} \to T_{\Omega_\kappa}$ be Alice's Diary for the value of κ given above. Let $D = \text{AD}_\kappa \circ I$. The map $D : T_W \to T_{\Omega_\kappa}$ is a diary since it is height-preserving and order-preserving. Suppose $\alpha, \beta \in T_W$ satisfy $\virgo(\mathcal{S}_{\text{linear}}, \delta, J, N, \epsilon)$, and suppose, as ever, that they have the form given by Notation 0.18. We can assume that

$$m \leq 2\min(m,n) \quad \text{and} \quad n \leq 2\min(m,n) \tag{3.3}$$

otherwise, we would have $|m-n| \geq (m+n)/3$ and so $d(D\alpha, D\beta) \geq (m+n)/3$ and we would be done by choosing $M(\delta, J) \geq 3$. Let $1 \leq j \leq \delta\min(m,n) + J$ and $\mathrm{stat}^k_c \in \mathcal{S}_{\text{linear}}$ be those given by condition $\text{♍}(\mathcal{S}_{\text{linear}}, \delta, J, N, \epsilon)$. We have three cases: $\mathrm{length}(u_{p+j}) \geq \epsilon(m+n)$ or $\mathrm{length}(u'_{p+j}) \geq \epsilon(m+n)$ or $u_{p+j} \neq u'_{p+j}$.

Case 1: We suppose first that $\mathrm{length}(u_{p+j}) \geq \epsilon(m+n)$.

Claim. *Either (1) or (2) below occurs.*

1. *There exist letters*
$$b \in \widehat{\mathrm{stat}^k_\infty}(\overline{u_1}\dots\overline{u_p}\dots\overline{u_{p+j}})$$
and
$$b' \in \widehat{\mathrm{stat}^k_\infty}(\overline{u_1}\dots\overline{u_p}\dots\overline{u'_{p+j}})$$
which are distinct and at the same distance $d \leq \tau(m+n)$ from the end of their words;

2. *length*$(\widehat{\mathrm{stat}^k_\infty}(\overline{u_1}\dots\overline{u_p}\dots\overline{u_{p+j}})) \geq \tau(m+n)$ *but length*$(\widehat{\mathrm{stat}^k_\infty}(\overline{u_1}\dots\overline{u_p}\dots\overline{u'_{p+j}})) < \tau(m+n)$.

Proof of claim. First, note that $\mathrm{length}(\widehat{\mathrm{stat}^k_\infty}(\overline{u_1}\dots\overline{u_p}\dots\overline{u_{p+j}})) \geq \tau(m+n)$ since $\mathrm{length}((u_{p+j})^\omega) = \omega\mathrm{length}(u_{p+j}) \geq \omega\epsilon(m+n) \geq \tau(m+n)$.

Now suppose that (2) doesen't occur, i.e. $\mathrm{length}(\widehat{\mathrm{stat}^k_\infty}(\overline{u_1}\dots\overline{u_p}\dots\overline{u'_{p+j}})) \geq \tau(m+n)$. So $\tau(m+n) \leq \omega\mathrm{length}(u'_{p+j})$. For a contradiction, suppose that (1) also doesen't occur. Then the final $\tau(m+n)$ letters of $\widehat{\mathrm{stat}^k_\infty}(\overline{u_1}\dots\overline{u_p}\dots\overline{u_{p+j}})$ and $\widehat{\mathrm{stat}^k_\infty}(\overline{u_1}\dots\overline{u_p}\dots\overline{u'_{p+j}})$ must be the same. It follows that the first $\tau(m+n)$ letters of $\mathrm{norm}_{\omega\mathrm{length}(u_{p+j})}(\mathrm{stat}^k_\infty(\overline{u_1}\dots\overline{u_{p+j}}))$ and $\mathrm{norm}_{\omega\mathrm{length}(u'_{p+j})}(\mathrm{stat}^k_\infty(\overline{u_1}\dots\overline{u'_{p+j}}))$ must be the same. But $\mathrm{norm}_{\omega\mathrm{length}(u_{p+j})}(\mathrm{stat}^k_\infty(\overline{u_1}\dots\overline{u_{p+j}}))$ contains $\mathrm{stat}^k_{m+n}(\overline{u_1}\dots\overline{u_{p+j}})$ as an initial subword since $\tau(m+n) \leq \omega\mathrm{length}(u_{p+j})$. Similarly, $\mathrm{norm}_{\omega\mathrm{length}(u'_{p+j})}(\mathrm{stat}^k_\infty(\overline{u_1}\dots\overline{u'_{p+j}}))$ contains $\mathrm{stat}^k_{m+n}(\overline{u_1}\dots\overline{u'_{p+j}})$ as an initial subword since $\tau(m+n) \leq \omega\mathrm{length}(u'_{p+j})$. We have a contradiction since $\mathrm{stat}^k_{m+n}(\overline{u_1}\dots\overline{u_{p+j}}) \neq \mathrm{stat}^k_{m+n}(\overline{u_1}\dots\overline{u'_{p+j}})$. □

It follows from the claim that there exist letters

$$b \in \bigstar \begin{pmatrix} (u_{p+j})^{\omega} \\ \widehat{\mathrm{stat}}^{1}_{\infty}(\overline{u_1}\dots\overline{u_p}\dots\overline{u_{p+j}}) \\ \widehat{\mathrm{stat}}^{2}_{\infty}(\overline{u_1}\dots\overline{u_p}\dots\overline{u_{p+j}}) \\ \vdots \\ \widehat{\mathrm{stat}}^{K}_{\infty}(\overline{u_1}\dots\overline{u_p}\dots\overline{u_{p+j}}) \end{pmatrix} \quad \text{and} \quad b' \in \bigstar \begin{pmatrix} (u'_{p+j})^{\omega} \\ \widehat{\mathrm{stat}}^{1}_{\infty}(\overline{u_1}\dots\overline{u_p}\dots\overline{u'_{p+j}}) \\ \widehat{\mathrm{stat}}^{2}_{\infty}(\overline{u_1}\dots\overline{u_p}\dots\overline{u'_{p+j}}) \\ \vdots \\ \widehat{\mathrm{stat}}^{K}_{\infty}(\overline{u_1}\dots\overline{u_p}\dots\overline{u'_{p+j}}) \end{pmatrix}$$

which are distinct and both at distance $d \leq \tau(m+n)$ from the end of their words (in case (2) you choose b' to be the $\bigstar$ at the beginning of the word). Now, suppose that the sentence $\overline{u_{p+j+1}}\dots\overline{u_{p+m}}$ has exactly l letters. This implies that the tail-sentence of

$$b \in I(\alpha)$$

has at most $\tau(m+n)+\omega l+(m-j)$ letters. So it has average word length at most

$$\frac{\tau(m+n)+\omega l+(m-j)}{m-j+1} \leq \tau\frac{m+n}{m-j}+\omega N+1 \leq \tau\frac{3m}{m-j}+\omega N+1$$

where we have used that $\mathrm{AWL}(\overline{u_{p+j+1}}\dots\overline{u_{p+m}}) = l/(m-j) \leq N$ together with (3.3). We now need to divide into two subcases: either $J \leq \frac{1-\delta}{16}\min(m,n)$ or $J > \frac{1-\delta}{16}\min(m,n)$.

Case 1a: We suppose $J \leq \frac{1-\delta}{16}\min(m,n)$. In particular, we have $J \leq \frac{1-\delta}{2}\min(m,n)$ and $2 \leq \frac{1-\delta}{8}(m+n)$. Then the average word length of $b \in I(\alpha)$ is at most

$$\tau\frac{3m}{m-j}+\omega N+1 \leq \tau\frac{3m}{m-\delta m-J}+\omega N+1 \leq \tau\frac{6}{1-\delta}+\omega N+1 \leq U$$

By similar arguments, the average word length of the tail-sentence of $b' \in I(\beta)$ is also at most U. In order to apply Corollary 3.24 to the starred sentences $I(\alpha)$ and $I(\beta)$, we choose

$$i = \left\lceil \frac{\min(m,n)-j}{2} \right\rceil$$

from which it follows that

$$U\frac{m-j+1}{i+1} \leq 4U\frac{m-j}{\min(m,n)-j} \leq 4U\frac{2\min(m,n)}{\min(m,n)-\delta\min(m,n)-J} \leq \frac{16U}{1-\delta} \leq \kappa$$

and similarly $U\frac{n-j+1}{i+1} \leq \kappa$. Hence, Corollary 3.24 tells us that

$$\begin{aligned}
d(\mathrm{AD}_\kappa \circ I(\alpha), \mathrm{AD}_\kappa \circ I(\alpha)) &\geq d_{T_{W'}}(I(\alpha), I(\beta)) - 2j - 2i \\
&= (m - j - i) + (n - j - i) \\
&\geq ((m - j)/2 - 1) + ((n - j)/2 - 1) \\
&\geq ((m - \delta m - J)/2) + ((n - \delta n - J)/2) - 2 \\
&\geq \frac{(1-\delta)m}{4} + \frac{(1-\delta)n}{4} - 2 \\
&\geq \frac{(1-\delta)(m+n)}{4} - \frac{(1-\delta)(m+n)}{8} \\
&= \frac{1-\delta}{8} d_{T_W}(\alpha, \beta)
\end{aligned}$$

If we choose $M(\delta, J) \geq 8/(1-\delta)$ then we are done in this case.

Case 1b: We suppose $J > \frac{1-\delta}{16} \min(m, n)$. Then, using (3.3), we have $m, n \leq \frac{32J}{1-\delta}$ and so $m + n \leq \frac{64J}{1-\delta}$. Recall that $b \in I(\alpha)$ and $b' \in I(\beta)$ are both at distance $d \leq \tau(m+n)$ from the ends of their words. So, since $\kappa \geq \frac{64J\tau}{1-\delta} \geq \tau(m+n) \geq d$, they are distance at most κ from the ends of their words. It follows that Alice records the events b and b' on page d of chapter $p + j$ of the diary. It follows that chapter $p + j$ of $\mathrm{AD}_\kappa \circ I(\alpha)$ is distinct from chapter $p + j$ of $\mathrm{AD}_\kappa \circ I(\beta)$. Hence,

$$\begin{aligned}
d(\mathrm{AD}_\kappa \circ I(\alpha), \mathrm{AD}_\kappa \circ I(\alpha)) &\geq (p + m) - (p + j - 1) + (p + n) - (p + j - 1) \\
&= m + n - 2j + 2 \\
&\geq 2 \\
&= \frac{1-\delta}{32J} \frac{64J}{1-\delta} \\
&\geq \frac{1-\delta}{32J}(m+n)
\end{aligned}$$

So if we choose $M(\delta, J) \geq \frac{32J}{1-\delta}$ then we are done in this case too.

Case 2: We are done for symmetrical reasons in the case when $\mathrm{length}(u'_{p+j}) \geq \epsilon(m+n)$.

Case 3: Finally, we suppose that $\mathrm{length}(u_{p+j}) \leq \epsilon(m+n)$, $\mathrm{length}(u'_{p+j}) \leq \epsilon(m+n)$ and $u_{p+j} \neq u'_{p+j}$. Two possibilities can occur: (1) there exist distinct letters $a \in (u_{p+j})^\omega$ and $a' \in (u'_{p+j})^\omega$ which are at the same distance from the end of their words and this distance is at most $\epsilon\omega(m+n) \leq (\tau + \epsilon)(m+n)$ or (2) $\mathrm{length}((u_{p+j})^\omega) \neq \mathrm{length}((u'_{p+j})^\omega)$. In either case, it follows that there exist distinct letters $b \in I(\alpha)$ and $b' \in I(\beta)$ which occur on the $(p+j)$'th day and which have equal distance $d \leq (\tau + \epsilon)(m+n)$ from the ends of their words. The proof now proceeds along the same lines as Case 1 and Case 2; a similar argument, splitting into cases depending on the magnitude of J, will give us that $d(\mathrm{AD}_\kappa \circ I(\alpha), \mathrm{AD}_\kappa \circ I(\alpha)) \geq \frac{1-\delta}{32J} d_{T_W}(\alpha, \beta)$. $\square$

We will now define another useful criterion, pronounced *Taurus*, which makes use of the fact that we always know that $u_{p+1} \neq u'_{p+1}$ when α and β have the form given by Notation 0.18. It is somewhat analogous to condition $\Omega(\mathcal{S}_{\text{finite}}, J)$. As before, suppose that $\mathcal{S}_{\text{linear}}$ is some finite collection of linear statistics on T_W (with possibly different codomains) and that $J \in \mathbb{N}$, $N > 0$ and $\epsilon > 0$ are constants.

Definition 3.32. We say that α, β satisfy $\taurus(\mathcal{S}_{\text{linear}}, J, N, \epsilon)$ if there exists some $1 \leq j \leq J$ such that

($\taurus$1) $\mathrm{AWL}(\overline{u_{p+j+1}} \dots \overline{u_{p+m}})$ and $\mathrm{AWL}(\overline{u'_{p+j+1}} \dots \overline{u'_{p+n}})$ are both at most N;

($\taurus$2) if $1 \leq j' \leq j$ is such that we have $\mathrm{length}(u_{p+j''}) \leq \epsilon(m+n)$ and $\mathrm{length}(u'_{p+j''}) \leq \epsilon(m+n)$ for all $j' < j'' \leq j$ then there exists $\mathrm{stat}_c \in \mathcal{S}_{\text{linear}}$ such that $\mathrm{stat}_{m+n}(\overline{u_1} \dots \overline{u_p}\ \overline{u_{p+1}} \dots \overline{u_{p+j'}}) \neq \mathrm{stat}_{m+n}(\overline{u_1} \dots \overline{u_p}\ \overline{u'_{p+1}} \dots \overline{u'_{p+j'}})$.

There is no denying that ($\taurus$2) is ugly. But you can observe that if we replaced ($\taurus$2) with the following more friendly condition

($\taurus$2′) for all $1 \leq j' \leq j$ there exists $\mathrm{stat}_c \in \mathcal{S}_{\text{linear}}$ such that $\mathrm{stat}_{m+n}(\overline{u_1} \dots \overline{u_p}\ \overline{u_{p+1}} \dots \overline{u_{p+j'}}) \neq \mathrm{stat}_{m+n}(\overline{u_1} \dots \overline{u_p}\ \overline{u'_{p+1}} \dots \overline{u'_{p+j'}})$.

then we would have created a new property, called ($\taurus'$), which implies ($\taurus$). So ($\taurus$) is a weaker (and thus easier to prove) condition. And indeed the weaker form will be needed in the proof of Theorem 4.1. It is also worth noting that ($\taurus$2) implies that there exists a linear statistic $\mathrm{stat}_c \in \mathcal{S}_{\text{linear}}$ such that $\mathrm{stat}_{m+n}(\overline{u_1} \dots \overline{u_p}\ \overline{u_{p+1}} \dots \overline{u_{p+j}}) \neq \mathrm{stat}_{m+n}(\overline{u_1} \dots \overline{u_p}\ \overline{u'_{p+1}} \dots \overline{u'_{p+j}})$.

Theorem 3.33. *There exists a finite set* $\Omega = \Omega(\mathcal{S}_{linear}, J, N, \epsilon)$, *a diary* $D = D(\mathcal{S}_{linear}, J, N, \epsilon)$ *which is a function* $D : T_W \to T_\Omega$, *and a constant* $M = M(J) \geq 1$, *such that*

$$\frac{1}{M} d_{T_W}(\alpha, \beta) \leq d_{T_\Omega}(D\alpha, D\beta) \leq d_{T_W}(\alpha, \beta)$$

for all α, β *satisfying* $\taurus(\mathcal{S}_{linear}, J, N, \epsilon)$.

Proof. Suppose that α, β satisfy $\taurus(\mathcal{S}_{\text{linear}}, J, N, \epsilon)$. We let the map $D : T_W \to T_\Omega$ be the diary $D(\mathcal{S}_{\text{linear}}, 0, J, N + 6J^2\epsilon, \epsilon)$ given by Theorem 3.31.

Note that if $m < (m+n)/3$ then $n > 2(m+n)/3$ and so $|m-n| > (m+n)/3$, which implies $d_{T_\Omega}(D\alpha, D\beta) \geq (m+n)/3$. So we are done when $m < (m+n)/3$ by choosing $M = M(J) \geq 3$. Similarly we are done if $n < (m+n)/3$. So we can assume that $m, n \geq (m+n)/3$.

Claim. *α and β satisfy $\virgo(\mathcal{S}_{linear}, 0, J, N + 6J^2\epsilon, \epsilon)$*

Proof. Since α, β satisfy $\taurus(\mathcal{S}_{\text{linear}}, J, N, \epsilon)$, we know that there exists $1 \leq j \leq J$ such that ($\taurus$1) and ($\taurus$2) hold. In particular, $\mathrm{AWL}(\overline{u_{p+j+1}} \ldots \overline{u_{p+m}}) \leq N \leq N + 6J^2\epsilon$, $\mathrm{AWL}(\overline{u'_{p+j+1}} \ldots \overline{u'_{p+n}}) \leq N \leq N + 6J^2\epsilon$ and there exists $\mathrm{stat}_c \in \mathcal{S}$ such that $\mathrm{stat}_{m+n}(\overline{u_1} \ldots \overline{u_p} \ldots \overline{u_{p+j}}) \neq \mathrm{stat}_{m+n}(\overline{u_1} \ldots \overline{u_p} \ldots \overline{u'_{p+j}})$. Hence, if $\mathrm{length}(u_{p+j}) \geq \epsilon(m + n)$ or $\mathrm{length}(u_{p+j}) \geq \epsilon(m + n)$ then α, β satisfy $\virgo(\mathcal{S}_{\text{linear}}, 0, J, N + 6J^2\epsilon, \epsilon)$ and the claim is proved. So we can assume both lengths are at most $\epsilon(m + n)$. But this implies that

$$\begin{aligned}
\mathrm{AWL}(\overline{u_{p+j}} \ldots \overline{u_{p+m}}) &= \frac{\mathrm{length}(u_{p+j+1} \ldots u_{p+m}) + \mathrm{length}(u_{p+j})}{m - j + 1} \\
&\leq \mathrm{AWL}(\overline{u_{p+j+1}} \ldots \overline{u_{p+m}}) + \frac{\epsilon(m+n)}{m - j + 1} \\
&\leq N + \frac{3\epsilon m}{m - J + 1} \\
&\leq \begin{cases} N + 6\epsilon & \text{when } m \geq 2J \\ N + 6J\epsilon & \text{when } m \leq 2J \end{cases} \\
&\leq N + 6J^2\epsilon
\end{aligned}$$

and similarly $\mathrm{AWL}(\overline{u'_{p+j}} \ldots \overline{u'_{p+n}}) \leq N + 6J^2$. It also implies, since α and β satisfy $\taurus(\mathcal{S}_{\text{linear}}, J, N, \epsilon)$, that there exists $\mathrm{stat}_c \in \mathcal{S}_{\text{linear}}$ such that $\mathrm{stat}_{m+n}(\overline{u_1} \ldots \overline{u_p} \ldots \overline{u_{p+j-1}}) \neq \mathrm{stat}_{m+n}(\overline{u_1} \ldots \overline{u_p} \ldots \overline{u'_{p+j-1}})$. So if $\mathrm{length}(u_{p+j-1}) \geq \epsilon(m+n)$ or $\mathrm{length}(u'_{p+j-1}) \geq \epsilon(m+n)$ then the claim is proved. So we can assume that they are both at most $\epsilon(m + n)$. Continuing in this way, we can assume that $\mathrm{length}(u_{p+2})$, $\mathrm{length}(u_{p+3})$, $\ldots$, $\mathrm{length}(u_{p+j})$ and $\mathrm{length}(u'_{p+2})$, $\mathrm{length}(u'_{p+3})$, $\ldots$, $\mathrm{length}(u'_{p+j})$ are all at most $\epsilon(m + n)$. It follows that $\mathrm{AWL}(\overline{u_{p+2}} \ldots \overline{u_{p+m}})$ and $\mathrm{AWL}(\overline{u'_{p+2}} \ldots \overline{u'_{p+n}})$ are both at most $N + 6J^2\epsilon$. Since $u_{p+1} \neq u'_{p+1}$, α and β satisfy $\virgo(\mathcal{S}_{\text{linear}}, 0, J, N + 6J^2\epsilon, \epsilon)$ and the claim is proved. ☐

It follows that if $M = M(0, J)$ is the constant given by Theorem 3.31 then we have

$$\frac{1}{M} d_{T_W}(\alpha, \beta) \leq d_{T_\Omega}(D\alpha, D\beta) \leq d_{T_W}(\alpha, \beta)$$

and we are done. ☐

Remark 3.34. *A reader might have noticed that in the proof of Theorem 3.20 we actually only needed one of the events a and a′ to be recorded in the diary since that is all that is asked for in Lemma 3.19. I believe this means that Corollary 3.24 only requires κ to be larger than one of $N^{\frac{m-j+1}{i+1}}$ and $N'^{\frac{n-j+1}{i+1}}$, and, carrying this logic forward, I believe this means there are versions of ($\virgo$) and ($\taurus$) which only require one, and not both, of the average word lengths to be bounded by N. That would imply we would be able to deduce the* *conclusions of Theorem 3.31 and Theorem 3.33 with weaker assumptions. However I have not pursued this line of reasoning as it does not greatly affect the proof of Theorem 4.1 and because it would marginally complicate the proofs of Theorem 3.31 and Theorem 3.33.*

Recall the definition of $\leo(\mathcal{S}_{\text{finite}}, J)$ from Definition 3.12. The following corollary combines conditions $\leo(\mathcal{S}_{\text{finite}}, J)$ and conditions $\taurus(\mathcal{S}_{\text{linear}}, J, N, \epsilon)$ so that we can use them simultaneously.

Corollary 3.35. *Let $\mathcal{S}_{finite}$ be a finite collection of finite statistics, let $\mathcal{S}_{linear}$ be a finite collection of linear statistics, let $J_{finite} \in \mathbb{N}$, let $J_{linear} \in \mathbb{N}$, let $N \geq 1$, and finally let $\epsilon > 0$.*

Then there exists a finite set $\Omega = \Omega(\mathcal{S}_{finite}, \mathcal{S}_{linear}, J_{linear}, N, \epsilon)$, a diary $D = D(\mathcal{S}_{finite}, \mathcal{S}_{linear}, J_{linear}, N, \epsilon)$ which is a function $D : T_W \to T_\Omega$, and a constant $M = M(J_{finite}, J_{linear})$, such that

$$\frac{1}{M} d_{T_W}(\alpha, \beta) \leq d_{T_\Omega}(D\alpha, D\beta) \leq d_{T_W}(\alpha, \beta)$$

for all α, β satisfying $\leo(\mathcal{S}_{finite}, J_{finite})$ or $\taurus(\mathcal{S}_{linear}, J_{linear}, N, \epsilon)$.

Proof. Let $D_{\text{finite}} : T_W \to T_{\Omega_{\text{finite}}}$ be the diary $D(\mathcal{S}_{\text{finite}})$ given by Theorem 3.11, and let $M_{\text{finite}} = M_{\text{finite}}(0, J_{\text{finite}})$ be the associated constant. Let $D_{\text{linear}} : T_W \to T_{\Omega_{\text{linear}}}$ be the diary $D(\mathcal{S}_{\text{linear}}, J_{\text{linear}}, N, \epsilon)$ given by Theorem 3.33, and let $M_{\text{linear}} = M_{\text{linear}}(J_{\text{linear}})$ be the associated constant. Let $\Omega = \Omega_{\text{finite}} \times \Omega_{\text{linear}}$; we think of Ω as being a finite alphabet. Define $M = \max(M_{\text{finite}}, M_{\text{linear}})$. We define the diary $D : T_W \to T_\Omega$ by

$$D(\overline{w_1} \dots \overline{w_i}) = \begin{pmatrix} D_{\text{finite}}(\overline{w_1}) \\ D_{\text{linear}}(\overline{w_1}) \end{pmatrix} \begin{pmatrix} D_{\text{finite}}(\overline{w_1}\ \overline{w_2}) \\ D_{\text{linear}}(\overline{w_1}\ \overline{w_2}) \end{pmatrix} \cdots \begin{pmatrix} D_{\text{finite}}(\overline{w_1}\ \overline{w_2} \dots \overline{w_i}) \\ D_{\text{linear}}(\overline{w_1}\ \overline{w_2} \dots \overline{w_i}) \end{pmatrix}$$

Suppose that $\alpha, \beta \in T_W$ satisfy $\leo(\mathcal{S}_{\text{finite}}, J_{\text{finite}})$ or $\taurus(\mathcal{S}_{\text{linear}}, J_{\text{linear}}, N, \epsilon)$. Since D is a diary, we have $d_{T_\Omega}(D\alpha, D\beta) \leq d_{T_W}(\alpha, \beta)$. If α, β satisfy $\leo(\mathcal{S}_{\text{finite}}, J_{\text{finite}})$ then

$$d_{T_\Omega}(D\alpha, D\beta) \geq d_{T_{\Omega_{\text{finite}}}}(D_{\text{finite}}\alpha, D_{\text{finite}}\beta) \geq \frac{1}{M_{\text{finite}}} d_{T_W}(\alpha, \beta)$$

If α, β satisfy $\taurus(\mathcal{S}_{\text{linear}}, J_{\text{linear}}, N, \epsilon)$ then

$$d_{T_\Omega}(D\alpha, D\beta) \geq d_{T_{\Omega_{\text{linear}}}}(D_{\text{linear}}\alpha, D_{\text{linear}}\beta) \geq \frac{1}{M_{\text{linear}}} d_{T_W}(\alpha, \beta)$$

In either case,

$$d_{T_\Omega}(D\alpha, D\beta) \geq \frac{1}{M} d_{T_W}(\alpha, \beta)$$

and so we are done. □

Remark 3.36. *It should hopefully be clear that the above argument can be repeated in order to combine any of the criteria (♈), (♌), (♍), (♉).*

Corollary 3.35 is used repeatedly in the proof of Theorem 4.1. Example 3.4 gives an idea as to how Corollary 3.35 is used in the proof of Theorem 4.1. First, we need to thoughtfully choose the data $\mathcal{S}_{\text{finite}}, \mathcal{S}_{\text{linear}}, J_{\text{linear}}, N, \epsilon$ at the start of the proof in order to define the map $D(\mathcal{S}_{\text{finite}}, \mathcal{S}_{\text{linear}}, J_{\text{linear}}, N, \epsilon)$ given by Corollary 3.35 which is a function $D : T_W \to T_\Omega$. We then analyse all the possible pairs α_{q}, β_{q} to ensure that they satisfy ♌$(\mathcal{S}_{\text{finite}}, J_{\text{finite}})$ or ♉$(\mathcal{S}_{\text{linear}}, J_{\text{linear}}, N, \epsilon)$. You might think of this as *finding a finite or linear (with respect to $m + n = d_{T_W}(\alpha_{\mathsf{q}}, \beta_{\mathsf{q}})$) fact which distinguishes α_{q} from β_{q}*. Once this is done, we have proved that there exists a quasiisometric embedding $X \to \prod_{q=1}^{Q} T_\Omega$.

If a reader would like to the see the criteria (♌) and (♍) in action in a relatively simple setting then they should read Appendix C in which these ideas are applied in order to quasiisometrically embed the hyperbolic plane into a product of binary trees. The theorems in this current chapter and also Section 4.2 (which is short) in the next chapter suffice for the understanding of this example.

3.5 Discussion

The theory in this chapter, and the proof structure outlined in Example 3.4, can only be applied if we have a quasiisometric embedding of a metric space into a product of sentence-trees. This feels very strange to me—if we are doing coarse geometry, why is it necessary to insist that our quasiisometric embeddings are into products of sentence-trees as opposed to products of abstract infinite-valence trees (which are quasiisometric to the sentence-trees)? Further, the application of the criteria (♈), (♌), (♍) and (♉) requires that we have a very precise understanding of the sentences α and β—it generally does not suffice to imagine that α or β are just vertices *near* a certain a vertex or roughly in some area, as is the case in most coarse geometry proofs. The result of this is that Theorem 4.1 requires a great deal of preciseness in the definition of the sentences α and β which again clashes with the typical flavour of coarse geometry. For example, if you consider the proof given in Appendix C that the hyperbolic plane $\mathbb{H}^2$ quasiisometrically embeds into a product of two binary trees, it is necessary to first pass to a specific Cayley graph that is quasiisometric to $\mathbb{H}^2$. Indeed, it is the exact opposite of the philosophy of geometric group theory—we are passing from the geometry of $\mathbb{H}^2$ to the discrete combinatorics of a Cayley graph.

Question 3.37. *Can the theory given above be given a more geometric basis? Is that even desirable?*

Chapter 4

Quasiisometric embeddings of (relatively) hyperbolic groups into products of binary trees

4.1 Introduction

Figure 4.1: Margate where some of this chapter was written.

Results

In this section, we prove the following result.

Theorem 4.1 (Main theorem)**.** *Let G be a finitely generated group that is relatively hyperbolic with respect to virtually abelian peripheral subgroups $H_1, H_2, \ldots, H_T$. Let R_t denote* *the rank of a finite index abelian subgroup of H_t, and define $R = \max(R_1, R_2, \ldots, R_T)$. Then G quasiisometrically embeds into a product of $R + \max(asdim(G), R+1) + 1$ binary trees. Since $asdim(G)$ is finite [Osi05], this is a finite product of binary trees.*

In particular, the fundamental group of a finite volume real hyperbolic manifold quasiisometrically embeds into a finite product of binary trees.

Theorem 4.1 should be compared with the result of Mackay and Sisto below [MS13].

Theorem 4.2 (Mackay–Sisto)**.** *Let G be a finitely generated group. Suppose G is relatively hyperbolic with respect to peripheral subgroups $H_1, H_2, \ldots, H_T$. Suppose also that each H_t quasiisometrically embeds into a product of m_t trees, and define $m = \max(m_1, m_2, \ldots, m_T)$. Then G quasiisometrically embeds into a product of $m + \max(asdim(G), m+1) + 1$ trees. Since $asdim(G)$ is finite [Osi05], this is a finite product of trees.*

If a metric space X quasiisometrically embeds into a finite product of trees then X has finite asymptotic dimension. Hence the Mackay–Sisto result is itself a refinement of a result of Osin [Osi05]: if a finitely generated group G is relatively hyperbolic with respect to peripheral subgroups with finite asymptotic dimension, then G has finite asymptotic dimension.

It is unclear to me whether the virtually abelian condition on the peripheral subgroups in Theorem 4.1 can be weakened. The obvious conjecture is that G quasiisometrically embeds into a product of binary trees if and only if the peripheral subgroups do as well. In this paper, the virtual nilpotence of the peripheral subgroups is used in order to quasiisometrically embed the Bowditch space $X(G)$ into a product of binary trees (see Corollary 4.44).

Question 4.3. *Can the virtually abelian condition be removed? If the peripherals quasiisometrically embed into products of binary trees, does G do the same?*

It follows from [Pau01] that a virtually nilpotent but not virtually abelian finitely generated group cannot quasiisometrically embed into a product of trees. Hence, since peripheral subgroups are undistorted in a relatively hyperbolic group, if a single peripheral subgroup H of a relatively hyperbolic group is virtually nilpotent but not virtually abelian, then G cannot quasiisometrically embed into a product of trees. So, thanks to [Pau01] and Theorem 4.1, the picture is fully understood when the peripheral subgroups are virtually nilpotent.

Corollary 4.4. *Let G be finitely generated and relatively hyperbolic with respect to virtually nilpotent peripheral subgroups. Then G quasiisometrically embeds into a product of binary trees if and only if all the peripheral subgroups are virtually abelian.*

Theorem 4.1 and Theorem 4.2 should themselves be compared with the two results below [BL07] [BDS07].

Theorem 4.5 (Buyalo–Lebedeva)**.** *A hyperbolic group G admits a quasiisometric embedding into a product of $n+1$ trees where n is the topological dimension of the boundary.*

Theorem 4.6 (Buyalo–Dranishnikov–Schroeder)**.** *A hyperbolic group admits a quasiisometric embedding into a product of $n+1$ binary trees where n is the topological dimension of the boundary.*

Buyalo and Lebedeva [BL07] also prove that for a hyperbolic group G we have $\operatorname{asdim}(G) = n+1$ where n is the topological dimension of the boundary.

Theorem 4.1 sits within the circle of ideas in coarse geometry and geometric group theory that involve constructing and obstructing metric embeddings between groups and spaces. Sometimes this is justified via Guoliang Yu's proof [Yu98] that finitely generated groups with finite asymptotic dimension satisfy Novikov's conjecture. Indeed, a group that admits a coarse embedding into a finite product of trees has finite asymptotic dimension. Yu's result proves that, in principle, one can draw algebraic conclusions from metric embedding results for groups. However I am currently unaware of any algebraic deductions that stem specifically from a quasiisometric embedding into a product of binary trees. As an aside, it would be interesting to know whether there existed a quasiisometry invariant, analogous to the asymptotic dimension or linearly controlled asymptotic dimension (which is also known as the asymptotic Assouad–Nagata dimension), which could distinguish between a product of infinite-valence trees and a product of binary trees.

Perhaps the best justification for Theorem 4.1 is that it sits at the top of a collection of results which attempt to bound the size and dimension of hyperbolic and relatively hyperbolic groups. For Theorem 4.1 implies that hyperbolic groups quasiisometrically embed into products of binary trees [BDS07] which implies that hyperbolic groups quasiisometrically embed into products of infinite-valence trees [BL07] which implies that hyperbolic groups have finite asymptotic dimension [Gro93]. With that said, it would not be accurate to call Theorem 4.1 a generalisation of these results since it relies on [BDS07, Theorem 1.2] in an essential way. As far as I'm aware, Theorem 4.1, the results of [BDS07] and the thesis of Alina Rull [Rul08], in which Rull proves that an n-coloured right-angled Artin group or n-coloured right-angled Coxeter group admits a quasiisometric embedding into a product of n binary trees, are the only results on the topic of quasiisometric embeddings into products of binary trees.

Ideas

The initial stages of the proof of Theorem 4.1 mimic the proof of Mackay and Sisto's Theorem 4.2. Using the projection complex machinery of [BBF15] [BBFS19] [BBF21], and the distance formula for relatively hyperbolic groups of [Sis13], one can prove that there exists a quasiisometric embedding

$$G \to \mathcal{C}_K(\mathbb{G}) \times X(G)$$

where $\mathcal{C}_K(\mathbb{G})$ is a quasi-tree of peripheral subgroups and $X(G)$ is a hyperbolic metric space called the Bowditch space. Using Theorem 4.6 and a theorem of Dahmani and Yaman [DY05], one can show that $X(G)$ quasiisometrically embeds into a product of binary trees when the peripherals are virtually abelian (see Corollary 4.44). Hence, it only remains to consider the structure of the quasi-tree of peripheral subgroups $\mathcal{C}_K(\mathbb{G})$. There are two significant novel ideas that I develop in order to complete the proof of Theorem 4.1.

The first new idea is to prove that by simply removing some edges from the quasi-tree of peripheral subgroups $\mathcal{C}_K(\mathbb{G})$ we can form a *tree* of peripheral subgroups $\mathcal{C}_K^{\mathcal{T}}(\mathbb{G})$ which is quasiisometric to $\mathcal{C}_K(\mathbb{G})$. This upgrade simplifes the metric structure of $\mathcal{C}_K(\mathbb{G})$ greatly. This will be covered in Section 4.4. Indeed, this reproves a result of Hume (see [Hum17, Proposition 2.8 and Theorem 1]) who proved that the quasi-tree of metric spaces $\mathcal{C}_K(\mathbb{Y})$, as defined by Bestvina, Bromberg and Fujiwara, is quasiisometric to a tree-graded space. However, in Hume's work this quasiisometry is defined abstractly, following from the fact that $\mathcal{C}_K(\mathbb{Y})$ satisfies a *relative bottleneck property*. In contrast, the quasiisometry given in Section 4.4 is a very explicit inclusion map.

The second new idea is to use the theory of diaries and statistics described in Chapter 3. We will use in particular the criteria (♌) and (♉).

Outline of this chapter

- In Section 4.2 and Section 4.3, we provide some preliminary results on the Morse–Thue sequence and regular maps into products of binary trees. The reader might want to skip these at first and then come back to them when the ideas become relevant in the proof of Theorem 4.1.
- In Section 4.4, the background on projection complexes and quasi-trees of metric spaces is provided. We also prove in particular that a quasi-tree of metric spaces $\mathcal{C}_K(\mathbb{Y})$ is quasiisometric to a tree of metric spaces $\mathcal{C}_K^{\mathcal{T}}(\mathbb{Y})$.
- In Section 4.5, we cover the relevant background on relatively hyperbolic groups.

- In Section 4.6, we define *standard paths* in a group that is relatively hyperbolic with respect to virtually abelian peripheral subgroups. We prove that these standard paths are quasigeodesics.
- In Section 4.7, we give the first stage of the proof of Theorem 4.1 by finding a quasiisometric embedding of G into a product of infinite-valence sentence-trees.
- In Section 4.8, using the theory developed in Chapter 3, we upgrade this quasiisometric embedding so that it is into a product of uniformly bounded valence trees.

Outline of the proof

Let G be relatively hyperbolic with respect to virtually abelian peripheral subgroups $H_1, \ldots H_T$. Let R be the maximal rank of finite index abelian subgroups of the peripheral subgroups. Let S be a finite generating set of G. Let A be the finite alphabet $S \cup S^{-1}$, let W be the set of finite words on A and let T_W be the sentence-tree. Using work of Dahmani and Yaman [DY05] and work of Buyalo, Dranishnikov and Schroeder [BDS07], we can prove there exists a regular map $\phi : G \to \prod_{q=1}^{Q} T_{\{0,1\}}$ where $Q \leq \max(\text{asdim}(G), R+1)+1$ (see Definition 4.12 and Corollary 4.46). Recall from Example 0.16 that $T_{\{0,1\}}$ denotes the rooted binary tree.

(Step 1) For each $1 \leq r \leq R$ and $x, z \in G$, we define a canonical standard path (which is a quasigeodesic) from x to z called the *r'th ordered standard path*. This is Definition 4.64.

(Step 2) By partitioning the r'th ordered standard path from $e \in G$ to $x \in G$ (which corresponds to an element of W) into several words, we may define a map $F_r : G \to T_W$.

(Step 3) We prove that $F_1 \times \cdots \times F_R \times \phi : G \to \prod_{r=1}^{R} T_W \times \prod_{q=1}^{Q} T_{\{0,1\}}$ is a quasiisometric embedding. The proof of this step uses the fact that the quasi-tree of metric spaces is quasiisometric to a tree of metric spaces (which is proved in Section 4.4).

(Step 4) We choose several finite statistics and linear statistics, and constants J_{finite}, J_{linear}, N and ϵ so that Corollary 3.35 induces a diary $D : T_W \to T_\Omega$.

(Step 5) We prove that

$$G \xrightarrow{F_1 \times \cdots \times F_R \times \phi} \prod_{r=1}^{R} T_W \times \prod_{q=1}^{Q} T_{\{0,1\}} \xrightarrow{D \times \ldots D \times \text{id}} \prod_{r=1}^{R} T_\Omega \times \prod_{q=1}^{Q} T_{\{0,1\}}$$

is a quasiisometric embedding. We follow the proof structure described in Example 3.4; in other words, we reduce the problem to proving that a pair of sentences $\alpha, \beta \in T_W$ satisfy either $\leo(\mathcal{S}_{\text{finite}}, J_{\text{finite}})$ or $\taurus(\mathcal{S}_{\text{linear}}, J_{\text{linear}}, N, \epsilon)$.

Since the uniformly bounded valence tree T_Ω is quasiisometric to $T_{\{0,1\}}$, we are done.

4.2 The Morse–Thue sequence

A remarkable feature of the Buyalo, Dranishnikov and Schroeder paper is their use of the *Morse–Thue sequence* in the proof of their result. We will also use the Morse–Thue sequence when proving Theorem 4.1.

Definition 4.7 (Quoted from [BDS07])**.** Consider the substitution rule $0 \mapsto 01$ and $1 \mapsto 10$. Then start from 0 and perform the substitutions

$$0 \mapsto 01 \mapsto 0110 \mapsto 01101001 \mapsto \dots$$

to obtain a nested family of sequences t_k of length 2^k in the alphabet $\{0, 1\}$. The resulting limit sequence is called the *Morse–Thue sequence.*

Notation 4.8. We will use the notation $\mathrm{MT}(n)$ to refer to the nth term of the Morse–Thue sequence. We will use the notation $\mathrm{MT}[a, b]$ to refer to the substring of the Morse–Thue sequence between the ath term and the bth term (inclusive). Given a finite word $u = a_1a_2 \dots a_l$ on some alphabet A, we define $\mathrm{MT}(u) := \mathrm{MT}(1)\mathrm{MT}(2)\dots\mathrm{MT}(l) = \mathrm{MT}[1, l]$.

An amazing property of the sequence is the following.

Theorem 4.9 (See, for example, [Hed67])**.** *The Morse–Thue sequence is cube-free: it contains no string of the form www where w is any word on the alphabet $\{0, 1\}$.*

Theorem 4.9 has the following corollary.

Corollary 4.10 (Rigidity of the Morse–Thue sequence)**.** *Suppose we have two identical substrings of the Morse–Thue sequence of length $3k$. Denote these substrings by $MT[a - 3k + 1, a]$ and $MT[b - 3k + 1, b]$. Then*

$$|a - b| \leq k \implies a = b$$

The above corollary says that if you take a substring w of the Morse–Thue sequence, then you cannot shift w a small amount and land on an identical substring.

Proof of Corollary 4.10. Suppose $|a-b| \leq k$ and $a \neq b$. Suppose, without loss of generality, that $a > b$. The key observation is that $\mathrm{MT}(n) = \mathrm{MT}(n+a-b)$ for $b-3k+1 \leq n \leq b$. Hence,

$$w = \mathrm{MT}[b-3k+1, b-3k+1+(a-b-1)]$$

is the same string as

$$w' = \mathrm{MT}[b-3k+1+(a-b), b-3k+1+(a-b-1)+(a-b)]$$

which in turn is the same string as

$$w'' = \mathrm{MT}[b-3k+1+2(a-b), b-3k+1+(a-b-1)+2(a-b)]$$

The substrings w, w', w'' are adjacent and equal in the Morse–Thue sequence and so we have a contradiction. $\square$

The following corollary describes how we will apply the Morse–Thue sequence.

Corollary 4.11. *Let A be a finite alphabet and let T_A be the word-tree on A (whose vertices are in bijection with the free monoid A^*). Suppose $w, w' \in A^*$ are distinct words on A that satisfy $d_{T_A}(w, w') \leq k$. Then either the final $3k$ letters of w and w' are distinct or the final $3k$ letters of $MT(w)$ and $MT(w')$ are distinct.*

Proof. If $\mathrm{length}(w) < 3k$ or $\mathrm{length}(w') < 3k$ then obviously the final $3k$ letters of w and the final $3k$ letters of w' are distinct. So we may assume that $\mathrm{length}(w) \geq 3k$ and $\mathrm{length}(w') \geq 3k$. Suppose that the final $3k$ letters of $\mathrm{MT}(w)$ are the same as the final $3k$ letters of $\mathrm{MT}(w')$. Since $|\mathrm{length}(w) - \mathrm{length}(w')| \leq k$, it follows from Corollary 4.10 that $\mathrm{length}(w) = \mathrm{length}(w')$. Since $d_{T_A}(w, w') \leq k$, it follows that there exists letters $a \in w$ and $a' \in w'$ which are distinct and at the same distance $d \leq k/2$ from the end of their words. It follows that the final $3k$ letters of w and w' are distinct. $\square$

4.3 Regular maps and binary trees

The notion of a *regular map* notably appears in the work of Benjamini, Schramm and Timár [BST12] on separation profiles since separation profiles are monotone under regular maps.

Definition 4.12. Let $\phi : V(\Gamma) \to V(\Gamma')$ be a map between the vertex sets of connected bounded degree graphs Γ, Γ'. We say that ϕ is *regular* if ϕ is Lipschitz and if there is a uniform bound on the cardinality of $\phi^{-1}(v)$.

Two important examples of regular maps are coarse embeddings and subgroup inclusion. It is possible to define a regular map between arbitrary metric spaces, but we will not do so here. In the lemma below, the maps ϕ and Φ should be interpreted as having vertex sets as domain and codomain, although this is dropped from the notation.

Lemma 4.13. *Suppose $\phi : \Gamma \to \prod_{q=1}^{Q} T_{\{0,1\}}$ is a regular map of a bounded degree graph (with base vertex $v_0 \in \Gamma$) into the product of Q binary trees. Then there exists a finite alphabet C, with associated word-tree T_C, and a Lipschitz embedding $\Phi : \Gamma \to \prod_{q=1}^{Q} T_C$ such that*

- *$\Phi(v)$ is a vertex for all vertices $v \in \Gamma$;*
- *Φ is injective on the vertices of Γ;*
- *the basepoint $v_0 \in \Gamma$ is mapped by Φ to the product of the roots $\emptyset = (\emptyset, \emptyset, \dots, \emptyset)$.*

Proof. Let $m \in \mathbb{N}$ be a uniform upper bound on the amount of vertices in $\phi^{-1}(v)$. Let $n \in \mathbb{N}$ and let T denote the n-ary rooted tree. That is, T is the rooted tree such that every vertex has n children. I think it is clear that if $n \in \mathbb{N}$ is much larger than m then by adjusting ϕ by a uniformly bounded amount, we can define an *injective* and coarsely Lipschitz map $\Phi : \Gamma \to \prod_{q=1}^{Q} T$ which takes $v_0 \in \Gamma$ to the basepoint of $\prod_{q=1}^{Q} T$. Further, a coarsely Lipschitz map whose domain is the vertex set of a graph is in fact Lipschitz. So Φ is Lipschitz on the vertices of Γ.

Finally, by setting $C = \{1, 2, 3, \dots, n\}$, we can identify T with the word-tree T_C. This completes the proof of the lemma. □

4.4 Projection complexes

Background

Various metric spaces can be understood by considering projections of subspaces onto other subspaces. As we shall see in Section 4.5, due to work of Sisto, the metric of a relatively hyperbolic group can be understood by considering projections of peripheral cosets onto other peripheral cosets [Sis13]. Similarly, the metric on the mapping class group $MCG(\Sigma)$ can be understood by considering the projection of curves in Σ onto curve complexes of subsurfaces of Σ [MM99] [MM00]. Bestvina, Bromberg and Fujiwara [BBF15] innovated two structures that you can associate to such metric spaces: the *projection complex* $\mathcal{P}_K(\mathbb{Y})$ and the *quasi-tree of metric spaces* $\mathcal{C}_K(\mathbb{Y})$. Bestvina, Bromberg and Fujiwara also found a general set of conditions that a set of projections must satisfy such that $\mathcal{P}_K(\mathbb{Y})$ and $\mathcal{C}_K(\mathbb{Y})$ exist.

Definition 4.14. Let $\mathbb{Y}$ be an indexing set such that for each $Y \in \mathbb{Y}$ we have an associated geodesic metric space $\mathcal{C}(Y)$. Suppose also that for each $Y \in \mathbb{Y}$ we have a function

$$\pi_Y : \mathbb{Y} \setminus \{Y\} \to \text{non-empty subsets of } \mathcal{C}(Y)$$

We call these functions *projections.* We then define $d_Y^\pi : \mathbb{Y} \setminus \{Y\} \times \mathbb{Y} \setminus \{Y\} \to [0, \infty]$ by

$$d_Y^\pi(X, Z) = \operatorname{diam}(\pi_Y(X) \cup \pi_Y(Z))$$

We call these functions the *projection distances.* The collection of maps $\{\pi_Y\}$ satisfy the *projection axioms* if there exists a constant $\xi < \infty$, known as the *projection constant*, such that

(P3) $\operatorname{diam}(\pi_Y(X)) \leq \xi$ for all distinct $X, Y \in \mathbb{Y}$;

(P4) if $X, Y, Z \in \mathbb{Y}$ are distinct and such that $d_Y^\pi(X, Z) > \xi$ then $d_X^\pi(Y, Z) \leq \xi$;

(P5) $\{Y : d_Y^\pi(X, Z) > \xi\}$ is finite for each pair $X, Z \in \mathbb{Y} \setminus \{Y\}$.

Example 4.15. Due to [Sis13], if $\mathbb{Y}$ is the set of cosets of peripheral subgroups in a relatively hyperbolic group, if $\mathcal{C}(Y)$ is the coset $Y = gH$ with the word metric arising from some finite set of generators of H, and if $\pi_Y(X)$ is the nearest point projection of the coset X onto the coset Y in the Cayley graph, then the collection of maps $\{\pi_Y\}$ satisfy axioms (P3) - (P5) for some choice of ξ.

The axioms (P3) - (P5) are those that appear in "real-life", however we need them to satisfy an extra axiom in order to conduct the theory of projection complexes. The collection of maps $\{\pi_Y\}$ satisfy the *strong projection axioms* if, in addition to (P3) - (P5), the following condition also holds

(P4') if $X, Y, Z \in \mathbb{Y}$ are distinct and such that $d_Y^\pi(X, Z) > \xi$ then $\pi_X(Y) = \pi_X(Z)$.

Example 4.16. Let $G = F(a, b)$ be the free group on two generators, and consider the standard Cayley graph $\Gamma = \Gamma(G, \{a, b\})$. We let $\mathbb{Y}$ be the set of cosets of the subgroup $\langle a \rangle$. Given a coset $Y \in \mathbb{Y}$, let $\mathcal{C}(Y)$ be the associated "horizontal" line (which is isometric to $\mathbb{R}$) in Γ. Let $\pi_Y(X) \subset \mathcal{C}(Y)$ be the nearest point projection in the Cayley graph. Then one can check that the collection $\{\pi_Y\}$ satisfy the strong projection axioms (P3) - (P5) and (P4') with respect to the projection constant $\xi = 0$.

Bestvina, Bromberg and Fujiwara prove that, given projections $\{\pi_Y\}$ satisfying the projection axioms, we can adjust the projections $\pi_Y(X)$ by a uniformly bounded amount so that the projections $\{\pi_Y\}$ now satisfy the strong projection axioms with respect to some possibly larger projection constant θ. This is the content of the following theorem (see [BBFS19, Theorem 4.1]).

Theorem 4.17 (Bestvina–Bromberg–Fujiwara)**.** *Suppose we have a collection of maps* $\{\pi_Y\}$ *satisfying the projection axioms with projection constant* ξ*. Then there exists a new collection of projections*

$$\{\pi'_Y : \mathbb{Y} \setminus \{Y\} \to \textit{non-empty subsets of } \mathcal{C}(Y)\}$$

such that if $d_Y(X,Z) := \mathit{diam}(\pi'_Y(X) \cup \pi'_Y(Z))$ *then*

- $\pi'_Y(X) \subseteq N_\xi(\pi_Y(X))$*;*
- $\{\pi'_Y\}$ *now satisfy the strong projection axioms (P3) - (P5) and (P4') with respect to the projection constant* $\theta = 11\xi$*;*
- $d^\pi_Y - 2\xi \leq d_Y \leq d^\pi_Y + 2\xi$.

Further, if all the metric spaces $\mathcal{C}(Y)$ *are graphs, and all the projections* $\pi_Y(X) \subset \mathcal{C}(Y)$ *are collections of vertices, then* $\pi'_Y(X) \subset \mathcal{C}(Y)$ *is also a collection of vertices.*

The *Further* part of the above theorem may be deduced from the definition of $\pi'_Y(X)$ in [BBFS19, Definition 4.11]. Indeed, one may write $\pi'_Y(X)$ as a union $\bigcup_W \pi_Y(W)$ where the union is over a non-empty set.

Given a collection of projections $\{\pi_Y : Y \in \mathbb{Y}\}$ with projection distances $d_Y : \mathbb{Y}\setminus\{Y\} \times \mathbb{Y} \setminus \{Y\} \to [0,\infty)$ and projection constant θ satisfying the *strong* projection axioms, we can now define standard paths, the projection complex and the quasi-tree of metric spaces.

Definition 4.18. For a constant $K \geq 0$, define

$$\mathbb{Y}_K(X,Z) = \{Y \in \mathbb{Y} : d_Y(X,Z) > K\}$$

and

$$\mathbb{Y}_K[X,Z] = \mathbb{Y}_K(X,Z) \cup \{X,Z\}$$

In what follows, we will also use the notation

$$L(X,Z) = |\mathbb{Y}_K(X,Z)|$$

and

$$L[X,Z] = |\mathbb{Y}_K[X,Z]|$$

One should interpret the statement $Y \in \mathbb{Y}_K(X,Z)$ as saying that Y is *between* X and Z.

Definition 4.19. The *projection complex* is the graph $\mathcal{P}_K(\mathbb{Y})$ with vertex set $\mathbb{Y}$ and an edge between two vertices if $\mathbb{Y}_K(X,Z) = \emptyset$.

We now state three results which describe the structure of the sets $\mathbb{Y}_K[X,Z]$. These are Lemma 2.3, Proposition 2.4 and Corollary 2.6 of [BBFS19] respectively.

Lemma 4.20 (Bestvina–Bromberg–Fujiwara)**.** *For $K \geq 2\theta$, and $Y, Y' \in \mathbb{Y}_K(X,Z)$, the following conditions are equivalent*

1. $d_Y(X, Y') > \theta$;
2. $d_{Y'}(Y, Z) > \theta$;
3. $d_Y(Y', Z) \leq \theta$;
4. $d_{Y'}(X, Y) \leq \theta$;
5. $d_Y(Y', W) = d_Y(Z, W)$ *for all* $W \neq Y$;
6. $d_{Y'}(Y, W) = d_{Y'}(X, W)$ *for all* $W \neq Y'$.

Given $K \geq 2\theta$ and two distinct elements $Y, Y' \in \mathbb{Y}_K(X,Z)$, we write $Y < Y'$ if any of the above equivalent conditions hold. Write $X < Y$ and $Y < Z$ for all $Y \in \mathbb{Y}_K(X,Z)$ and write $X < Z$.

Proposition 4.21 (Bestvina–Bromberg–Fujiwara)**.** *Suppose $K \geq 2\theta$. The relation $<$ on $\mathbb{Y}_K[X,Z]$ defines a total order with least element X and greatest element Z. Further, if $Y < Y' < Y''$ in $\mathbb{Y}_K[X,Z]$ then $d_{Y'}(Y, Y'') = d_{Y'}(X,Z)$.*

Corollary 4.22 (Bestvina–Bromberg–Fujiwara)**.** *Let $K \geq 2\theta$. Suppose $Y, Y'' \in \mathbb{Y}_K[X,Z]$ and for some $Y' \in \mathbb{Y}$ we have $d_{Y'}(Y, Y'') > K$. Then $Y' \in \mathbb{Y}_K(X,Z)$.*

We now state some important facts about the graph $\mathcal{P}_K(\mathbb{Y})$. These are Lemma 3.1, Theorem 3.5, Lemma 3.6 and Corollary 3.7 from [BBFS19] respectively.

Lemma 4.23 (Bestvina–Bromberg–Fujiwara)**.** *Let $K \geq 2\theta$. Suppose $\mathbb{Y}_K[X,Z] = \{X < X_1 < \cdots < X_k < Z\}$. Then*

$$X \to X_1 \to \cdots \to X_k \to Z$$

is a path in $\mathcal{P}_K(\mathbb{Y})$.

We refer to the sets $\mathbb{Y}_K[X,Z]$ as *standard paths* in $\mathcal{P}_K(\mathbb{Y})$. Note that for any $X, Z \in \mathbb{Y}$ we have $\mathbb{Y}_K[X,Z] = \mathbb{Y}_K[Z,X]$ as a set, and $\mathbb{Y}_K[Z,X]$ is the path $\mathbb{Y}_K[X,Z]$ walked in reverse.

Theorem 4.24 (Bestvina–Bromberg–Fujiwara)**.** *For $K \geq 3\theta$, $\mathcal{P}_K(\mathbb{Y})$ is a quasi-tree.*

Lemma 4.25 (Bestvina–Bromberg–Fujiwara)**.** *Suppose $K \geq 2\theta$. Let X, Y, Z be distinct elements of $\mathbb{Y}$. Write $\mathbb{Y}_K[X,Y] = \{X = Y_0 < Y_1 < \cdots < Y_k < Y_{k+1} < Y\}$ and $\mathbb{Y}_K[Y,Z] = \{Y = Z_0 < Z_1 < \cdots < Z_l < Z_{l+1} < Z\}$. Then the standard path $\mathbb{Y}_K[X,Z]$ is the concatenation of three disjoint segments:*

- *it begins with an initial segment $X \to Y_1 \to \cdots \to Y_i$ in $\mathbb{Y}_K[X,Y]$ for some $0 \leq i \leq k$;*
- *it then has a (possibly empty) middle segment with at most 2 vertices which are in neither $\mathbb{Y}_K[X,Y]$ or $\mathbb{Y}_K[Y,Z]$;*
- *finally, it ends with a terminal segment $Z_j \to Z_{j+1} \to \cdots \to Z_l \to Z$ for some $1 \leq j \leq l+1$.*

The above lemma is the key to understanding the geometry of the projection complex. It says that the standard paths $\mathbb{Y}_K[X,Z]$, $\mathbb{Y}_K[X,Y]$, $\mathbb{Y}_K[Y,Z]$ *almost* form a tripod. Instead, the three standard paths have a cycle of length at most 9 in their centre. See Figure 4.2.

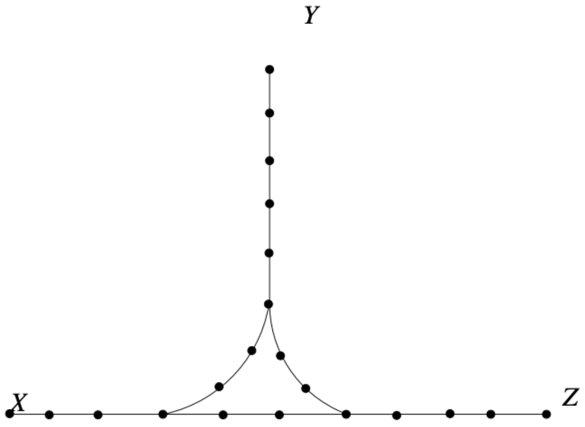

Figure 4.2: A triangle of standard paths. Taken from [BBFS19].

Corollary 4.26 (Bestvina–Bromberg–Fujiwara)**.** *For $K \geq 3\theta$, standard paths are quasi-geodesics; for all $X, Z \in \mathbb{Y}$ we have*

$$(L[X,Z] - 1)/2 \leq d_{\mathcal{P}_K(\mathbb{Y})}(X,Z) \leq L[X,Z] - 1$$

Quasi-trees of metric spaces

Suppose again that $\{\pi_Y : Y \in \mathbb{Y}\}$ satisfies (P3) - (P5) and (P4') with respect to the projection constant θ. Suppose also that $K \geq 3\theta$.

Definition 4.27. The *quasi-tree of metric spaces* $\mathcal{C}_K(\mathbb{Y})$ is built by taking the disjoint union of the metric spaces $\mathcal{C}(Y)$ and then adding an edge of length K between every element of $\pi_X(Y)$ and every element of $\pi_Y(X)$ whenever there is an edge between X and Y in $\mathcal{P}_K(\mathbb{Y})$. We will often refer to one of the edges between $\pi_X(Y)$ and $\pi_Y(X)$ as a *transverse edge.*

The name of $\mathcal{C}_K(\mathbb{Y})$ comes from the fact that crushing all the disjoint metric spaces $\mathcal{C}(Y)$ that make up $\mathcal{C}_K(\mathbb{Y})$ to points leaves us with the quasi-tree $\mathcal{P}_K(\mathbb{Y})$.

It is useful to extend the definitions of π_Y, d_Y and $\mathbb{Y}_K(\cdot,\cdot)$ over to elements $x \in \mathcal{C}(X) \subseteq \mathcal{C}_K(\mathbb{Y})$.

Definition 4.28. For $x \in \mathcal{C}(X) \subseteq \mathcal{C}_K(\mathbb{Y})$ and $z \in \mathcal{C}(Z) \subseteq \mathcal{C}_K(\mathbb{Y})$ and $Y \in \mathbb{Y}$

- we define $\pi_Y(x) = \begin{cases} \pi_Y(X) & \text{if } Y \neq X \\ \{x\} & \text{if } Y = X \end{cases}$;
- we define $d_Y(x,z) = \operatorname{diam}(\pi_Y(x) \cup \pi_Y(z))$ where the diameter is taken with respect to the metric on $\mathcal{C}(Y)$;
- we define $\mathbb{Y}_K(x,z) = \{Y \in \mathbb{Y} : d_Y(x,z) > K\}$.

We then have the following distance formula for $\mathcal{C}_K(\mathbb{Y})$ [BBFS19, Theorem 6.4]. It says that we can approximate distances in $\mathcal{C}_K(\mathbb{Y})$ by just looking at the projections.

Theorem 4.29 (Bestvina–Bromberg–Fujiwara)**.** *If $K \geq 4\theta$, then for all $x \in \mathcal{C}(X) \subseteq \mathcal{C}_K(\mathbb{Y})$ and $z \in \mathcal{C}(Z) \subseteq \mathcal{C}_K(\mathbb{Y})$ we have*

$$\frac{1}{4} \sum_{Y \in \mathbb{Y}_K(x,z)} d_Y(x,z) \leq d_{\mathcal{C}_K(\mathbb{Y})}(x,z) \leq 2 \sum_{Y \in \mathbb{Y}_K(x,z)} d_Y(x,z) + 3K$$

Constructing a tree of metric spaces

Suppose we have a collection of projections $\{\pi_Y : Y \in \mathbb{Y}\}$ with projection distances $d_Y : \mathbb{Y} \setminus \{Y\} \times \mathbb{Y} \setminus \{Y\} \to [0,\infty)$ and projection constant θ that satisfy axioms (P3) - (P5) and (P4'). Let $K \geq 4\theta$.

Definition 4.30. Let $B \in \mathbb{Y}$ be a fixed basepoint. This corresponds to a vertex of $\mathcal{P}_K(\mathbb{Y})$. We can define the following subgraph of $\mathcal{P}_K(\mathbb{Y})$.

$$\mathcal{T}_K(\mathbb{Y}, B) := \bigcup_{X \in \mathbb{Y}} \mathbb{Y}_K[B, X]$$

We will generally write $\mathcal{T} = \mathcal{T}_K(\mathbb{Y}, B)$ for simplicity. In the above, $\mathbb{Y}_K[B, X]$ should be interpreted as a path that includes the relevant edges as described in Lemma 4.23.

By definition $\mathcal{T}$ is a connected subgraph of $\mathcal{P}_K(\mathbb{Y})$ that must contain all the vertices of $\mathcal{P}_K(\mathbb{Y})$.

Theorem 4.31. *The subgraph $\mathcal{T} \subseteq \mathcal{P}_K(\mathbb{Y})$ is a tree.*

First we need a preliminary lemma.

Lemma 4.32. *Let $\mathbb{Y}_K[X, Z] = \{X = X_0 < X_1 < \cdots < X_k < X_{k+1} = Z\}$ be the standard path from X to Z. Let $0 \leq i < j \leq k+1$. Then $\{X_i < X_{i+1} < \cdots < X_j\}$ is the standard path from X_i to X_j.*

Proof. Let P denote the path $\{X_i < X_{i+1} < \cdots < X_j\}$. Consider l such that $i < l < j$. Then, since $X_i < X_l < X_j$ in $\mathbb{Y}_K[X, Z]$, Proposition 4.21 implies that $d_{X_l}(X_i, X_j) = d_{X_l}(X, Z) > K$. Hence, $X_l \in \mathbb{Y}_K[X_i, X_j]$. So $P \subseteq \mathbb{Y}_K[X_i, X_j]$.

Now suppose that $Y \in \mathbb{Y}_K(X_i, X_j)$. Then $d_Y(X_i, X_j) > K$ which, by Corollary 4.22, implies that $Y \in \mathbb{Y}_K(X, Z)$. It also implies that $X_i < Y < X_j$ in $\mathbb{Y}_K[X, Z]$ and hence $Y \in P$. Hence $\mathbb{Y}_K[X_i, X_j] \subseteq P$. So P and $\mathbb{Y}_K[X_i, X_j]$ share the same vertex set.

We therefore just need to check that the order in which the vertices are traversed in $\mathbb{Y}_K[X_i, X_j]$ is consistent with the order $X_i < X_{i+1} < \cdots < X_j$ in P. Suppose that $i < l < m < j$. Then $d_{X_l}(X_i, X_m) > K$ since $X_i < X_l < X_m$ in $\mathbb{Y}_K[X, Z]$. But this implies that $X_l < X_m$ in $\mathbb{Y}_K[X_i, X_j]$ as well. ☐

Proof of Theorem 4.31. We begin by proving the following claim.

Claim. *Suppose we have a standard path $\mathbb{Y}_K[B, X] = \{B = X_0 < X_1 < \cdots < X_k < X_{k+1} = X\}$ and suppose there exists another vertex Z connected by an edge (in $\mathcal{T}$) to X such that $Z \neq X_k$. Then $\{B < X_1 < \cdots < X_k < X < Z\}$ is the standard path from B to Z.*

Proof of claim. Since the edge from X to Z is in $\mathcal{T}$, we know that there exists some vertex $U \in \mathbb{Y}$ such that $\mathbb{Y}_K[B, U] = \{B = X'_0 < X'_1 < \cdots < X'_l < X'_{l+1} = U\}$ and $\{X, Z\} = \{X'_j, X'_{j+1}\}$. Suppose for a contradiction that $X = X'_{j+1}$ and $Z = X'_j$. Then, by Lemma 4.32, $B < X'_1 < X'_2 < \cdots < X'_{j-1} < Z < X$ is the standard path from B to X. It follows that $Z = X_k$ which contradicts our assumption on Z. So $X = X'_j$ and $Z = X'_{j+1}$. By Lemma 4.32, we know that $B < X'_1 < \cdots < X'_j$ is the standard path from B to X. Lemma 4.32 also implies that $B < X'_1 < \cdots < X'_{j+1}$ is the standard path from B to Z. Hence $B < X_1 < \cdots < X_k < X < Z$ is the standard path from B to Z. $\square$

Suppose in $\mathcal{T}$ we have a cycle $C = \{Y_0, Y_1, \ldots, Y_{l-1}\}$ where there is an edge between Y_i and Y_{i+1} (taken modulo l). Let $Y \in C$ and suppose that the path $\mathbb{Y}_K[B, Y]$ first intersects C at the vertex $Y' \in C$. It follows that the standard path $\mathbb{Y}_K[B, Y'] = \{B < X_1 < \cdots < X_k < Y'\}$ only intersects C at Y'. Suppose, for simplicity, that $Y' = Y_0$. By the claim, we know that $B < X_1 < \cdots < X_k < Y_0 < Y_1$ is the standard path from B to Y_1. Similarly, we know that $B < X_1 < \cdots < X_k < Y_0 < Y_1 < Y_2$ is the standard path from B to Y_2. Continuing in this way, we find that $B < X_1 < \cdots < X_k < Y_0 < Y_1 < Y_2 < \cdots < Y_l < Y_0$ is the standard path from B to Y_0 which is of course nonsense. $\square$

Theorem 4.33. *The inclusion map $\mathcal{T} \hookrightarrow \mathcal{P}_K(\mathbb{Y})$ is a quasiisometry.*

Proof. Evidently the inclusion map is Lipschitz; it remains to prove the lower bound of the quasiisometric inequality.

Consider the triple B, X, Z. By Lemma 4.25, the path $\mathbb{Y}_K[X, Z]$ can be described as follows: an initial segment that is also initial within $\mathbb{Y}_K[X, B]$, a middle segment of length at most 2, and a terminal segment that is also terminal in $\mathbb{Y}_K[B, Z]$. Let Y' be the vertex at which $\mathbb{Y}_K[X, Z]$ diverges from $\mathbb{Y}_K[X, B]$ and let Y'' be the vertex at which $\mathbb{Y}_K[X, Z]$ joins $\mathbb{Y}_K[B, Z]$. Let Y be the vertex at which $\mathbb{Y}_K[B, X]$ and $\mathbb{Y}_K[B, Z]$ diverge. By Lemma 4.25 we have

$$L[X, Y] \leq L[X, Y'] + 3 \tag{4.1}$$

and

$$L[Y, Z] \leq L[Y'', Z] + 3 \tag{4.2}$$

Now, by the definition of $\mathcal{T}$, we have that

$$d_{\mathcal{T}}(X, Z) = d_{\mathcal{T}}(X, Y) + d_{\mathcal{T}}(Y, Z) = L[X, Y] - 1 + L[Y, Z] - 1 \tag{4.3}$$

We also know that

$$L[X, Z] \geq L[X, Y'] + L[Y'', Z] - 1 \tag{4.4}$$

Hence

$$\begin{aligned} d_{\mathcal{P}_K(\mathbb{Y})}(X,Z) &\geq \frac{1}{2}(L[X,Z]-1) && \text{by Corollary 4.26} \\ &\geq \frac{1}{2}(L[X,Y']+L[Y'',Z]-1) && \text{by (4.4)} \\ &\geq \frac{1}{2}(L[X,Y]-3+L[Y,Z]-3-1) && \text{by (4.1) and (4.2)} \\ &= \frac{1}{2}(d_{\mathcal{T}}(X,Z)-5) && \text{by (4.3)} \end{aligned}$$

and so we are done. □

A key fact about the projection complex $\mathcal{P}_K(\mathbb{Y})$ is that it is a quasi-tree (see Theorem 4.24). Theorem 4.33 above provides a very explicit quasiisometry from $\mathcal{P}_K(\mathbb{Y})$ to a tree. We can use $\mathcal{T}=\mathcal{T}(\mathbb{Y},B)$ to define a *tree* of metric spaces $\mathcal{C}_K^{\mathcal{T}}(\mathbb{Y})$ as follows.

Definition 4.34. For each pair $X,Y\in\mathbb{Y}$ which are connected by an edge in $\mathcal{P}_K(\mathbb{Y})$, fix an arbitrary element $p(X,Y)\in\pi_X(Y)\subset\mathcal{C}(X)$ and an arbitrary element $p(Y,X)\in\pi_Y(X)\subset\mathcal{C}(Y)$. In order to define $\mathcal{C}_K^{\mathcal{T}}(\mathbb{Y})$, we take the disjoint union of the metric spaces $\mathcal{C}(X)$ and draw an edge of length K between $p(X,Y)$ and $p(Y,X)$ whenever there is an edge between X and Y in $\mathcal{T}$.

Remark 4.35. *By all rights, we should be using the notation $p_X(Y)$ instead of $p(X,Y)$ in order to properly mimic the existing notation $\pi_X(Y)$. However, in the proof of Theorem 4.1 we will need to consider elements of $\mathbb{Y}$ with subscripts in their notation—having the additional subscript in $p_X(Y)$ would be painful to read.*

Evidently, there exists a natural inclusion map $\mathcal{C}_K^{\mathcal{T}}(\mathbb{Y})\hookrightarrow\mathcal{C}_K(\mathbb{Y})$.

Theorem 4.36. *The inclusion $\mathcal{C}_K^{\mathcal{T}}(\mathbb{Y})\hookrightarrow\mathcal{C}_K(\mathbb{Y})$ is a quasiisometry.*

We first need a lemma about the projection complex.

Lemma 4.37. *Let X,Y,Z be distinct vertices of $\mathcal{P}_K(\mathbb{Y})$.*

1. *If W is an element of $\mathbb{Y}_K[X,Y]$ and $\mathbb{Y}_K[X,Z]$ but not $\mathbb{Y}_K[Y,Z]$ then $d_W(Y,Z)\leq K$.*
2. *If W is an element of $\mathbb{Y}_K[X,Y]$ but not $\mathbb{Y}_K[X,Z]$ and not $\mathbb{Y}_K[Y,Z]$ then $d_W(X,Y)\leq 2K$.*

Proof. The first statement is immediate. For the second, note that $d_W(X,Y)\leq d_W(X,Z)+d_W(Z,Y)\leq 2K$. □

Proof of Theorem 4.36. Coarse surjectivity is almost immediate: every element $x \in \mathcal{C}_K(\mathbb{Y})$ is at distance at most $K/2$ from a metric space $\mathcal{C}(Y) \subset \mathcal{C}_K^{\mathcal{T}}(\mathbb{Y})$. Further, it is clear that the inclusion map is Lipschitz. It remains to prove the lower bound of the quasiisometric inequality.

Let $x \in \mathcal{C}(X)$ and $z \in \mathcal{C}(Z)$. To produce a lower bound on $d_{\mathcal{C}_K(\mathbb{Y})}(x,z)$, it is sufficient to find a nice path from x to z in $\mathcal{C}_K^{\mathcal{T}}(\mathbb{Y})$ whose length is comparable to the distance from x to z in $\mathcal{C}_K(\mathbb{Y})$.

We begin by finding a quasigeodesic from x to z in $\mathcal{C}_K(\mathbb{Y})$ (we will then find a path from x to z in $\mathcal{C}_K^{\mathcal{T}}(\mathbb{Y})$ of a comparable length). Suppose that $\mathbb{Y}_K[X,Z] = \{X = X_0 < X_1 < \cdots < X_{n-1} < X_n = Z\}$. We have the following path γ from x to z in $\mathcal{C}_K(\mathbb{Y})$

$$\begin{tikzcd}
x \arrow[r] & p(X_0,X_1) \arrow[d] & & & & & \\
& p(X_1,X_0) \arrow[r] & p(X_1,X_2) \arrow[d] & & & & \\
& & p(X_2,X_1) \arrow[r] & \dots \arrow[r] & p(X_{n-2},X_{n-1}) \arrow[d] & & \\
& & & & p(X_{n-1},X_{n-2}) \arrow[r] & p(X_{n-1},X_n) \arrow[d] & \\
& & & & & p(X_n,X_{n-1}) \arrow[r] & z
\end{tikzcd}$$

where each arrow either corresponds to a geodesic in $\mathcal{C}(X_i)$ or to a transverse edge between pairs of disjoint metric spaces. A segment $p(X_i,X_{i+1}) \to p(X_{i+1},X_i)$ has length K. Each segment $p(X_i,X_{i-1}) \to p(X_i,X_{i+1})$ has length at most $d_{X_i}(x,z)$. Similarly, $x \to p(X_0,X_1)$ has length at most $d_{X_0}(x,z)$ and $p(X_n,X_{n-1}) \to z$ has length at most $d_{X_n}(x,z)$. Thus the length of γ is at most

$$\sum_i d_{X_i}(x,z) + nK$$

It is shown in the proof of [BBFS19, Theorem 6.3] that

$$\sum_i d_{X_i}(x,z) + nK \leq 2 \sum_{Y \in \mathbb{Y}_K(x,z)} d_Y(x,z) + 3K$$

Applying Theorem 4.29 we see that

$$2 \sum_{Y \in \mathbb{Y}_K(x,z)} d_Y(x,z) + 3K \leq 8 d_{\mathcal{C}_K(\mathbb{Y})}(x,z) + 3K$$

and so γ has length at most $8d_{\mathcal{C}_K(\mathbb{Y})}(x,z) + 3K$. So γ is an $(8,3K)$-quasigeodesic in $\mathcal{C}_K(\mathbb{Y})$.

We will now find a path in $\mathcal{C}_K^{\mathcal{T}}(\mathbb{Y})$ that is very close to γ. Let X_i be the vertex of $\mathcal{P}_K(\mathbb{Y})$ at which $\mathbb{Y}_K[X, Z]$ and $\mathbb{Y}_K[X, B]$ diverge (if they never diverge, then we are done, as γ will lie inside $\mathcal{C}_K^{\mathcal{T}}(\mathbb{Y})$). Let X_j be the vertex of $\mathcal{P}_K(\mathbb{Y})$ at which $\mathbb{Y}_K[Z, X]$ and $\mathbb{Y}_K[Z, B]$ diverge (if they never diverge, then we are done, as γ will lie inside $\mathcal{C}_K^{\mathcal{T}}(\mathbb{Y})$). Finally, let Y be the vertex of $\mathcal{P}_K(\mathbb{Y})$ at which $\mathbb{Y}_K[B, X]$ and $\mathbb{Y}_K[B, Z]$ diverge (if they never diverge, then we are done, as γ will lie inside $\mathcal{C}_K^{\mathcal{T}}(\mathbb{Y})$). We can also assume that X_i, X_j and Y are distinct: if any of them are equal then γ lies inside $\mathcal{C}_K^{\mathcal{T}}(\mathbb{Y})$. By Lemma 4.25, we have a geodesic of length at most 6 from X_i to X_j in $\mathcal{T}$. This geodesic in $\mathcal{T}$ has the form $X_i \to \cdots \to Y \to \cdots \to X_j$. Now let η be a geodesic in $\mathcal{C}_K^{\mathcal{T}}(\mathbb{Y})$ from $p(X_i, X_{i+1})$ to $p(X_j, X_{j-1})$. The path η is a "diversion" from γ between $p(X_i, X_{i+1})$ and $p(X_j, X_{j-1})$. If we can show that η has uniformly bounded length then we will be done, as the path which travels along γ but takes the diversion η will be a path in $\mathcal{C}_K^{\mathcal{T}}(\mathbb{Y})$ of comparable length to γ.

The only metric spaces $\mathcal{C}(W)$ that η travels through are those corresponding to the geodesic $X_i \to \cdots \to Y \to \cdots \to X_j$ in $\mathcal{T}$ described previously. The intersection of η with $\mathcal{C}(X_i)$ corresponds to case (1) of Lemma 4.37 and hence has length at most K. Similarly, the intersection of η with $\mathcal{C}(Y)$ or $\mathcal{C}(X_j)$ has length K. The intersection of η with a metric space other than $\mathcal{C}(X_i), \mathcal{C}(Y), \mathcal{C}(X_j)$ corresponds to case (2) of Lemma 4.37 and hence has length at most $2K$. We conclude, by also taking into account the length of the transverse edges that η travels along, that η has length at most $17K$. $\square$

4.5 Relatively hyperbolic groups

Background

Let G be a group with finite generating set S and let $\Gamma(G, S)$ denote the Cayley graph of G with respect to S. When we speak of the metric on G, we are actually referring to the metric induced by $\Gamma(G, S)$, but we will generally denote it by d_G nonetheless. Suppose that $\mathbb{H}$ is some collection of subgroups of G. Elements of $\mathbb{H}$ are called *peripheral subgroups*. Let $\mathbb{G}$ be the collection of all cosets of peripheral subgroups. In other words, $\mathbb{G} = \{gH : H \in \mathbb{H}\}$. Elements of $\mathbb{G}$ are called *peripheral cosets*. Let us suppose as well that for each $H \in \mathbb{H}$, $S \cap H$ generates the subgroup H.

Definition 4.38. The *coned-off graph* of G with respect to $\mathbb{H}$ is the graph formed by taking the Cayley graph $\Gamma(G, S)$ and then adding an edge between g and g' (if it's not already there) whenever g and g' are contained in the same peripheral coset. Although the coned-off graph depends on $\mathbb{H}$, this is omitted from the notation and the coned-off graph is simply denoted by $\hat{G}$.

The coned-off graph is (coarsely) what you get if you crush every peripheral coset to a point. We will now define the *Bowditch space* $X(G)$, a metric space closely related to $\hat{G}$.

Definition 4.39. Suppose that Γ is a connected graph with vertex set V_Γ and edge set E_Γ. Suppose also that every edge has length 1. The *horoball* $\mathcal{H}(\Gamma)$ is the graph with vertex set

$$V_{\mathcal{H}(\Gamma)} = V_\Gamma \times \mathbb{N}_0$$

and edge set

$$E_{\mathcal{H}(\Gamma)} = \{\{(v,n),(v',n)\} : \{v,v'\} \in E_\Gamma, n \in \mathbb{N}_0\} \sqcup \{\{(v,n),(v,n+1)\} : v \in V_\Gamma, n \in \mathbb{N}_0\}$$

We insist that edges of the form $\{(v,n),(v',n)\}$ have length e^{-n} and edges of the form $\{(v,n),(v,n+1)\}$ have length 1.

So the horoball $\mathcal{H}(\Gamma)$ consists of countably many copies of Γ stacked on top of each other, getting smaller and smaller as one travels higher.

Definition 4.40. The *Bowditch space* of G with respect to $\mathbb{H}$ is formed by taking $\Gamma(G,S)$ and then, for every $gH \in \mathbb{G}$, attaching a copy of $\mathcal{H}(\Gamma(H, S \cap H))$ to the subgraph $g\Gamma(H, S \cap H)$. The Bowditch space is denoted by $X(G)$.

The Bowditch space, like the coned-off graph, also involves contracting the peripheral cosets. However it does so less brutally; where the coned-off graph created a path of length 1 between any two vertices in the same peripheral coset, the Bowditch space creates a path between the vertices whose length is the log of their distance as elements of $g\Gamma(H, S \cap H)$.

There are now several definitions of a relatively hyperbolic group:

- the Farb definition [Far98] where G is relatively hyperbolic if it is weakly relatively hyperbolic and satisfies the bounded coset penetration property;
- the Bowditch definition [Bow12] where G is relatively hyperbolic if it is weakly relatively hyperbolic and if the coned-off graph is *fine*;
- the Osin definition [Osi06b] where G is relatively hyperbolic if it has a finite *relative* presentation and if the relative Dehn function is linear.

The definition we use below is equivalent to all the other definitions of relative hyperbolicity. It is due to Groves and Manning [GM08].

Definition 4.41. We say that G is *relatively hyperbolic* with respect to $\mathbb{H}$ if $X(G)$ is hyperbolic.

Remark 4.42. *Due to [Osi06b, Theorem 1.1], if G is relatively hyperbolic we may always assume that $\mathbb{H}$ is finite and that every $H \in \mathbb{H}$ is finitely generated.*

Strangely, I cannot find it recorded in the literature that [BDS07, Theorem 1.2] is really an *if and only if*. I'll record it here, since we have it almost for free.

Corollary 4.43 (Buyalo–Dranishnikov–Schroeder, Bonk–Schramm)**.** *If X is a visual hyperbolic metric space then X admits a quasiisometric embedding into a finite product of binary trees if and only if the boundary of X is a doubling metric space. Further, if the boundary of X is a doubling metric space, then X admits a quasiisometric embedding into a finite product of $n+1$ binary trees, where n is the linearly controlled metric dimension of the boundary.*

Proof. The backwards implication of the if and only if, and the *Further* part of the corollary, are exactly [BDS07, Theorem 1.2].

So suppose X is a visual hyperbolic metric space that admits a quasiisometric embedding into a finite product of binary trees. Then obviously X admits a quasiisometric embedding into a uniformly bounded degree graph. So Proposition 0.5 tells us that X has bounded growth on some scale. Hence, [BS00, Theorem 9.2] tells us that the boundary of X is doubling. ☐

Corollary 4.44 (Buyalo–Dranishnikov–Schroeder, Dahmani–Yaman, Bonk–Schramm)**.** *Suppose H is relatively hyperbolic with respect to $\mathbb{H}$. Then $X(G)$ quasiisometrically embeds into a finite product of binary trees if and only if all the subgroups in $\mathbb{H}$ are virtually nilpotent. Further, if all the subgroups in $\mathbb{H}$ are virtually nilpotent then $X(G)$ quasiisometrically embeds into a product of $n+1$ binary trees where n is the linearly controlled metric dimension of the boundary of $X(G)$.*

Proof. Mackay and Sisto [MS20, Proposition 4.5] use the main result of Dahmani and Yaman's paper [DY05] to prove that all the peripheral subgroups are virtually nilpotent if and only if the boundary of $X(G)$ is doubling. It then follows from Corollary 4.43 that all the peripheral subgroups are virtually nilpotent if and only if $X(G)$ quasiisometrically embeds into a finite product of binary trees.

The *Further* part of the corollary follows from Corollary 4.43. ☐

Recall from Example 0.16 that $T_{\{0,1\}}$ denotes the rooted binary tree.

Corollary 4.45. *Suppose $\mathbb{H} = \{H_1, H_2, \ldots, H_T\}$, suppose each peripheral subgroup H_t is virtually abelian, and define $R = \max_t R_t$ where R_t is the rank of a finite index abelian subgroup of H_t. Then there exists a quasiisometric embedding $\phi' : X(G) \to \prod_{q=1}^{Q} T_{\{0,1\}}$ where Q is at most $\max(asdim(G), R+1)+1$. We may assume that the image of ϕ' is contained in the vertex set of $\prod_{q=1}^{Q} T_{\{0,1\}}$.*

Proof. Corollary 4.44 implies that $X(G)$ quasiisometrically embeds into a product of $n+1$ binary trees where n is the linearly controlled metric dimension of the boundary of $X(G)$. Proposition 3.6 and Proposition 3.4 of [MS13] combine to prove that $n \leq \max(\operatorname{asdim}(G), R+1)$ as desired. $\square$

Corollary 4.46. *Let G be as in Corollary 4.45. Then there exists a regular map $\phi : G \to \prod_{q=1}^{Q} T_{\{0,1\}}$ where $Q \leq \max(\mathit{asdim}(G), R+1) + 1$.*

Proof. Let ϕ be the composition $G \hookrightarrow X(G) \to \prod_{q=1}^{Q} T_{\{0,1\}}$ where the second map is ϕ' from Corollary 4.45. By the assumption on ϕ', the image of ϕ is in the vertex set of $\prod_{q=1}^{Q} T_{\{0,1\}}$.

Since $G \to X(G)$ is Lipschitz and $\phi' : X(G) \to \prod_{q=1}^{Q} T_{\{0,1\}}$ is coarsely Lipschitz we conclude that ϕ is Lipschitz.

Let v be a vertex of $\prod_{q=1}^{Q} T_{\{0,1\}}$. Since $\phi' : X(G) \to \prod_{q=1}^{Q} T_{\{0,1\}}$ is coarsely Lipschitz, the preimage of v in $X(G)$ is contained in a ball of uniformly bounded radius in $X(G)$. But a ball of uniformly bounded radius in $X(G)$ can contain only a uniformly bounded amount of elements of $G \subset X(G)$. So $\phi^{-1}(v) \subset G$ has a uniformly bounded cardinality. $\square$

We now state some results on the metric structure of relatively hyperbolic groups. These are [DS05, Lemma 4.15] and [Sis13, Lemma 1.13, Proposition 1.14 and Lemma 1.15].

Proposition 4.47 (Druțu–Sapir)**.** *Peripheral cosets are undistorted in G. That is, the natural map $gH \subseteq G \to g\Gamma(H, S \cap H)$ is a quasiisometry.*

Given $gH \in \mathbb{G}$ and $x \in G$, we define $\pi_{gH}(x) \subseteq gH$ to be the nearest point projection of x onto $gH \subseteq \Gamma(G, S)$, i.e.

$$\pi_{gH}(x) = \{y \in gH : d_G(x, y) = \min_{z \in gH} d_G(x, z)\}$$

Lemma 4.48 (Sisto)**.** *There exists M such that if $\hat{\gamma}$ is a geodesic in the coned-off graph $\hat{G}$ connecting $x \in G$ to $y \in gH$ then the first point in $\hat{\gamma} \cap gH$ is at distance (taken with respect to the metric on $\Gamma(G, S)$) at most M from $\pi_{gH}(x)$.*

The *lift* of a geodesic $\hat{\gamma}$ in $\hat{G}$ is a path γ in G obtained by substituting edges of $\hat{\gamma}$ that go between $g, g' \in kH$ with a geodesic in the subgraph $k\Gamma(H, S \cap H)$ from g to g'.

Proposition 4.49 (Sisto)**.** *There exist $\lambda \geq 1$, $\mu \geq 0$ so that if $\hat{\gamma}$ is a geodesic in $\hat{G}$ then its lifts are (λ, μ)-quasigeodesics.*

Let $\bar{B}_r(x)$ denote the closed r-ball around some point $x \in G$.

Lemma 4.50 (Sisto)**.** *There exists L so that if $d_G(\pi_{gH}(x), \pi_{gH}(z)) \geq L$ for some $gH \in \mathbb{G}$ and $x, z \in G$ then*

1. *all (λ, μ)-quasigeodesics in $\Gamma(G, S)$ connecting x to z intersect $\bar{B}_R(\pi_{gH}(x))$ and $\bar{B}_R(\pi_{gH}(z))$ where $R = R(\lambda, \mu)$;*
2. *all geodesics in $\hat{G}$ connecting x to z contain an edge between two elements of gH.*

Relatively hyperbolic groups as projection complexes

As before, let G be finitely generated with respect to a finite generating set S and relatively hyperbolic with respect to a set of peripheral subgroups $\mathbb{H}$ (where $S \cap H$ generates each $H \in \mathbb{H}$) and let $\mathbb{G}$ be the set of peripheral cosets.

The set of peripheral cosets $\mathbb{G}$ will be our indexing set, taking the place of $\mathbb{Y}$ in Section 4.4. I have chosen to use the notation $\mathbb{G}$ in this context so that there is a clear demarcation between the general theory of projection complexes given in Section 4.4 and the specific case of relatively hyperbolic groups given here. However, for ease of notation, we will frequently refer to cosets by the notation X, Y, Z, W as in Section 4.4.

We associate to each $gH \in \mathbb{G}$ the metric space $\mathcal{C}(gH) = g\Gamma(H, S \cap H)$. Given $Y = gH \in \mathbb{G}$ and $x \in \mathcal{C}(X)$, we define, as in the previous subsection, $\pi_Y(x) \subseteq \mathcal{C}(Y)$ to be the nearest point projection of x onto $gH \subseteq \Gamma(G, S)$. Then we have the projection data given by

$$\pi_Y(X) = \{\pi_Y(x) : x \in \mathcal{C}(X)\}$$

There is a slight subtlety in the above definitions; the nearest-point projections π_Y are determined by the metric in $\Gamma(G, S)$, but we think of them as being subsets of $\mathcal{C}(Y) = g\Gamma(H, S \cap H)$.

Mackay and Sisto [MS13] notice that, by applying the work of Sisto [Sis13], the following result holds.

Theorem 4.51 (Mackay–Sisto)**.** *The metric spaces $\mathcal{C}(Y) = g\Gamma(H, S \cap H)$ together with the projection data*

$$\pi_Y : \mathbb{G} \setminus Y \to \textit{non-empty subsets of } \mathcal{C}(Y)$$

satisfy axioms (P3) - (P5) for some projection constant ξ.

It follows that we can apply all the theory developed in Section 4.4 on the projection complex $\mathcal{P}_K(\mathbb{Y})$, the quasi-tree of metric spaces $\mathcal{C}_K(\mathbb{Y})$ and the tree of metric spaces $\mathcal{C}_K^{\mathcal{T}}(\mathbb{Y})$ to relatively hyperbolic groups. So let us use Theorem 4.17 to upgrade our projections $\{\pi_Y\}$ so that they now satisfy the *strong* projection axioms with respect to a constant θ, and let us define $\mathcal{P}_K(\mathbb{G})$, $\mathcal{C}_K(\mathbb{G})$ and $\mathcal{C}_K^{\mathcal{T}}(\mathbb{G})$ for some $K \geq 4\theta$ accordingly.

Suppose $\mathbb{H} = \{H_1, H_2, \dots, H_T\}$. A group element $g \in G$ has T possible images in the quasi-tree of metric spaces $\mathcal{C}_K(\mathbb{G})$. Therefore, to properly specify a vertex of $\mathcal{C}_K(\mathbb{G})$ we need specify the group element *and* the coset. We write $g \in gH_t$ to mean the copy of g that sits within the coset $gH_t \subset \mathcal{C}_K(\mathbb{G})$. Let's define $\iota_t : G \to \mathcal{C}_K(\mathbb{G})$ by $\iota_t(g) = g \in gH_t$. It is not difficult to prove that the embeddings $\iota_t : G \to \mathcal{C}_K(\mathbb{G})$ are the same up to bounded distance; for lack of a better choice, we will always consider the embedding $\iota_1 : G \to \mathcal{C}_K(\mathbb{G})$.

Notation 4.52. A group element $g \in G$ has a natural image in $\hat{G}$ and $X(G)$. We will denote both of these images in $\hat{G}$ and $X(G)$ by g.

Sisto [Sis13] proves that the distance in a relatively hyperbolic group can be understood in terms of projections onto peripheral cosets and the distance in the coned-off graph. In order to state Sisto's theorem, we need some notation.

Notation 4.53. Let $A, B, L \in \mathbb{R}$, let $\lambda \geq 1$ and let $\mu \geq 0$.

- We write $A \approx_{\lambda,\mu} B$ if $A/\lambda - \mu \leq B \leq \lambda A + \mu$. If A and B are quantities which depend on a pair of group elements g and h (for example if $A = d_G(g, h)$ and $B = d_X(f(x), g(x))$ for some $f : G \to X$ and metric space X) then we might write $A \approx B$ if $A \approx_{\lambda,\mu} B$ for some $\lambda \geq 1$ and $\mu \geq 0$ which don't depend on g and h (in other words, λ, μ are *uniform* constants).
- We define $\{\{A\}\}_L$ to be 0 if $A \leq L$ and to be A if $A > L$.

Let $Y = gH$ and let $\rho_Y : G \to Y$ be *some* choice of nearest-point projection onto Y.

Theorem 4.54 (Sisto)**.** *There exists $L_0 \in \mathbb{R}$ such that for each $L \geq L_0$ there exist λ, μ such that the following holds. If $g, h \in G$ then*

$$d_G(g, h) \approx_{\lambda,\mu} \sum_{Y \in \mathbb{G}} \{\{d_G(\rho_Y(g), \rho_Y(h))\}\}_L + d_{\hat{G}}(g, h)$$

It follows from Theorem 4.54 and Theorem 4.29 that you may approximate distances in G via the images of group elements in $\mathcal{C}_K(\mathbb{G})$, $\hat{G}$ and $X(G)$. I have relegated the proofs of the following two corollaries to the appendix because, though intuitive, they are slightly technical. Further, they appear as [MS13, Theorem 4.1]. Mackay and Sisto [MS13] prove this result by simply comparing the distance formula for relatively hyperbolic groups (Theorem 4.54) directly with the distance formula for $\mathcal{C}_K(\mathbb{Y})$ (Theorem 4.29). However, I believe that if one is to be completely rigorous then some further checks have to be made since Theorem 4.54 and Theorem 4.29 are not automatically directly comparable (for example, Theorem 4.54 concerns the nearest point projection, whereas Theorem 4.29

concerns the *modified* projections given by Theorem 4.17). See Appendix B for the proofs of these corollaries.

Corollary 4.55. *For sufficiently large K, the map $G \to \mathcal{C}_K(\mathbb{G}) \times \hat{G}$ given by*

$$g \mapsto (\iota_1(g), g)$$

is a quasiisometric embedding.

Corollary 4.56. *For sufficiently large K, the map $G \to \mathcal{C}_K(\mathbb{G}) \times X(G)$ given by*

$$g \mapsto (\iota_1(g), g)$$

is a quasiisometric embedding.

4.6 Standard paths in groups that are relatively hyperbolic with respect to virtually abelian peripheral subgroups

In this section, we will use the theory developed in Section 4.5 to define a collection of canonical quasigeodesic paths between two group elements $x, z \in G$ when G is relatively hyperbolic with respect to virtually abelian peripheral subgroups.

So suppose G is relatively hyperbolic with respect to virtually abelian peripheral subgroups $H_1, H_2, \ldots, H_T$. Let R_t denote the rank of a finite index abelian subgroup of H_t, and define $R = \max(R_1, R_2, \ldots, R_T)$. Let $A_t \cong \mathbb{Z}^{R_t}$ be a finite index subgroup of H_t. A virtually abelian group always contains a normal, finite index abelian subgroup. Therefore without loss of generality we may assume that A_t is normal in H_t. Suppose G is generated by a finite set S. We may assume that $S_{H_t} := S \cap H_t$ generates H_t and that $S_{A_t} := S_{H_t} \cap A_t$ generates A_t. We write $S_{H_t} = S_{A_t} \cup \{b_{t,1}, b_{t,2}, \ldots, b_{t,R'_t}\}$ where $b_{t,r} \notin A_t$ for $1 \leq r \leq R'_t$. Further, we may assume the following: that $S_{A_t} = \{a_{t,1}, a_{t,2}, \ldots, a_{t,R_t}\}$ and that identifying

$$a_{t,1}^{q_1} a_{t,2}^{q_2} \ldots a_{t,R_t}^{q_{R_t}} \in \Gamma(A_t, S_{A_t})$$

with

$$(q_1, q_2, \ldots, q_{R_t}) \in \Gamma(\mathbb{Z}^{R_t}, \{(1, 0, \ldots, 0), (0, 1, \ldots, 0), \ldots, (0, 0, \ldots, 1)\})$$

gives a graph isomorphism. In other words, we assume that the subgraph $g\Gamma(A_t, S_{A_t})$ is a *grid* for all $g \in G$. Let I_t denote the cardinality of the quotient group H_t/A_t and write $I = \max_t I_t$.

As in Section 4.5, let $\mathbb{G} = \{gH_t : g \in G, 1 \leq t \leq T\}$ and define the projections $\{\pi_Y\}$ using the nearest point projections. We then modify $\{\pi_Y\}$ by a uniformly bounded amount using Theorem 4.17 so that they now satisfy axioms (P3)-(P5) and (P4') with respect to the projection constant θ. Since the original projections are collections of vertices, the modified projections will be too, as described in Theorem 4.17. The new projection distances will be denoted by d_Y. We choose K to be at least $4\theta + 2I$, and large enough that Corollary 4.56 holds. In particular, since $K \geq 4\theta$, all the results in Section 4.4 hold and we can define $\mathcal{P}_K(\mathbb{G})$ and $\mathcal{C}_K(\mathbb{G})$. By setting $B = H_1$, we can also define $\mathcal{T} = \mathcal{T}_K(\mathbb{G}, B)$ as in Definition 4.30. In order to define $\mathcal{C}_K^{\mathcal{T}}(\mathbb{G})$ as in Definition 4.34, we arbitrarily choose vertices $p(X,Y) \in \pi_X(Y)$ and $p(Y,X) \in \pi_Y(X)$ for every distinct pair $X, Y \in \mathbb{G}$ that are connected by an edge in $\mathcal{P}_K(\mathbb{G})$.

Standard non-abelian paths between cosets of A_t in H_t

Write $H_t/A_t = \{h_{t,1}A_t, h_{t,2}A_t, \ldots, h_{t,I_t}A_t\}$.

Definition 4.57. For $1 \leq i \leq I_t$, let $\psi_{t,i}$ be some fixed path in $\Gamma(G,S)$ of length at most I_t, which is a word on the generators $\{b_{t,1}, b_{t,2}, \ldots, b_{t,R'_t}\} \subset H_t$ and their inverses, and that travels from e to the coset $h_{t,i}A_t$. This is possible since H_t/A_t is a group of cardinality I_t that is generated by the set $\{b_{t,r}A_t : 1 \leq r \leq R'_t\}$.

Given two group elements k, k' in the same coset of H_t, say $k, k' \in gH_t$, there exists a unique $1 \leq i \leq I_t$ such that for all group elements $l \in kA_t$ we have that $l\psi_{t,i} \in k'A_t$. We refer to $\psi_{t,i}$ as *the standard non-abelian path* from the coset kA_t to the coset $k'A_t$.

Standard paths in $g\Gamma(H_t, S_{H_t})$

Definition 4.58. Given a coset $Y \in \mathbb{G}$, write

$$\mathbb{G}_K[H_1, Y] = \{H_1 = Y_1 < Y_2 < \cdots < Y_{L-1} < Y_L = Y\}$$

The *basepoint* of Y, which we denote by $p_0(Y)$, is defined to be $p(Y, Y_{L-1})$ if $Y \neq H_1$ and defined to be the identity e if $Y = H_1$.

Definition 4.59. Suppose k and k' are vertices of $\mathcal{C}(Y) = g\Gamma(H_t, S_{H_t})$. Let $\psi_{t,i}$ be the standard non-abelian path from the coset kA_t to the coset $p_0(Y)A_t$. Let $\psi_{t,j}$ be the standard non-abelian path from the coset $p_0(Y)A_t$ to the coset $k'A_t$. A *standard path* from k to k' in $\mathcal{C}(Y)$ is any path which is formed by concatenating the following three paths

- The standard non-abelian path $\psi_{t,i}$;

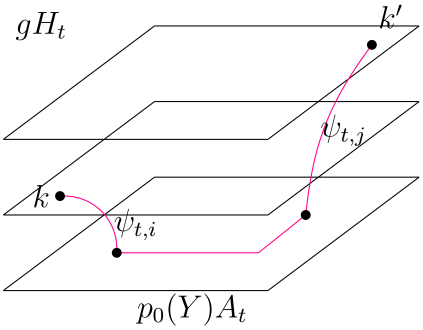

Figure 4.3: An ordered standard path from k to k' in the coset gH_t. Made using TikZiT.

- A geodesic in $p_0(Y)\Gamma(A_t, S_{A_t})$ formed of the abelian generators $\{a_{t,1}, \dots, a_{t,R_t}\}$ that takes us from $k\psi_{t,i}$ to $k'\psi_{t,j}^{-1}$;
- The standard non-abelian path $\psi_{t,j}$.

So the standard path from k to k' in $\mathcal{C}(Y)$ has the label $\psi_{t,i}\nu\psi_{t,j}$ where ν is a geodesic in the generators $\{a_{t,1}, \dots, a_{t,R_t}\}$ from $k\psi_{t,i}$ to $k'\psi_{t,j}^{-1}$. So we have forced the standard path from k to k' to go via the coset of A_t that contains $p_0(Y)$.

Sometimes, we would like to make choices of standard path which reflect the ordering of the abelian generators $\{a_{t,r} : 1 \le r \le R_t\}$. We will refer to these special choices of standard path as *ordered* standard paths. See Figure 4.3.

Definition 4.60. Recall that $R = \max_t R_t$ and suppose $1 \le r \le R$. As before, suppose k, k' are vertices of $\mathcal{C}(Y) = g\Gamma(H_t, S_{H_t})$. The *$r$'th ordered standard path* from k to k' in $\mathcal{C}(Y)$ is the unique standard path from k to k' which travels the abelian generators of A_t in the cyclic order

$$\begin{aligned} &a_{t,r} < a_{t,r+1} < \dots < a_{t,R_t} < a_{t,1} < a_{t,2} < \dots < a_{t,r-1} && \text{if } r \le R_t \\ &a_{t,1} < a_{t,2} < \dots < a_{t,R_t} && \text{if } r > R_t \end{aligned}$$

So the r'th ordered standard path from k to k' has a label of the form

$$\begin{aligned} &\psi_{t,i} a_{t,r}^{q_r} a_{t,r+1}^{q_{r+1}} \dots a_{t,R_t}^{q_{R_t}} a_{t,1}^{q_1} a_{t,2}^{q_2} \dots a_{t,r-1}^{q_{r-1}} \psi_{t,j} && \text{if } r \le R_t \\ &\psi_{t,i} a_{t,1}^{q_1} a_{t,2}^{q_2} \dots a_{t,R_t}^{q_{R_t}} \psi_{t,j} && \text{if } r > R_t \end{aligned}$$

Proposition 4.61. *There exist $\lambda_1 \ge 1$ and $\mu_1 \ge 0$ such that for all $1 \le t \le T$ and $g \in G$, every standard path in $g\Gamma(H_t, S_{H_t})$ is a (λ_1, μ_1)-quasigeodesic with respect to the metric on $g\Gamma(H_t, S_{H_t})$.*

Proof. This follows from the fact that $A_t \hookrightarrow H_t$ is a quasiisometry (since A_t has finite index in H_t) and the fact that $\mathrm{length}(\psi_{t,i}) \leq I$ for all t and i. $\square$

Lemma 4.62. *Suppose Y, Y', Y'' are distinct cosets in $\mathbb{G}$ and that $\{Y, Y'\}$ and $\{Y', Y''\}$ are edges in $\mathcal{P}_K(\mathbb{G})$. Suppose also that $d_{Y'}(Y, Y'') > K$. Then a standard path from $p(Y', Y)$ to $p(Y', Y'')$ contains at least one abelian generator $a_{t,r}^{\pm 1}$.*

Proof. Let $\zeta = \psi_{t,i}\nu\psi_{t,j}$ denote a standard path from $p(Y', Y)$ to $p(Y', Y'')$. We have that

$$\begin{aligned}
\mathrm{length}(\zeta) &\geq d(p(Y', Y), p(Y', Y'')) \\
&\geq \mathrm{diam}(\pi_{Y'}(Y) \cup \pi_{Y'}(Y'')) - 2\theta \\
&= d_{Y'}(Y, Y'') - 2\theta \\
&> K - 2\theta \\
&> 2I
\end{aligned}$$

where we have used $K \geq 4\theta + 2I$ in the final inequality. Since $\mathrm{length}(\psi_{t,i}) \leq I$ and $\mathrm{length}(\psi_{t,j}) \leq I$ we deduce that $\mathrm{length}(\nu) \geq 1$. $\square$

Quasigeodesics between cosets of H_t

Let $\lambda_2 \geq 1$ and $\mu_2 \geq 0$ be constants of our choosing.

Definition 4.63. Given $X, Y \in \mathbb{G}$ that are connected by an edge in $\mathcal{P}_K(\mathbb{G})$, let $\gamma(X, Y)$ be a fixed choice of (λ_2, μ_2)-quasigeodesic in $\Gamma(G, S)$ from $p(X, Y) \in G$ to $p(Y, X) \in G$. We sometimes refer to $\gamma(X, Y)$ as a *transverse quasigeodesic*.

It does not exactly matter what λ_2 and μ_2 are as long as they are uniform over all the choices of quasigeodesic. For example, you could always choose the transverse quasigeodesics to simply be geodesics from $p(X, Y) \in G$ to $p(Y, X) \in G$.

Standard paths in $\Gamma(G, S)$

Let $x, z \in G$ and write $X = xH_1$ and $Z = zH_1$. Let $L = L[X, Z]$ and suppose

$$\mathbb{G}_K[X, Z] = \{X = Y_1 < Y_2 < \cdots < Y_{L-1} < Y_L = Z\}$$

Definition 4.64. A *standard path* from x to z is formed by concatenating standard paths within cosets Y_l (as defined in Definition 4.59) together with the transverse quasigeodesics

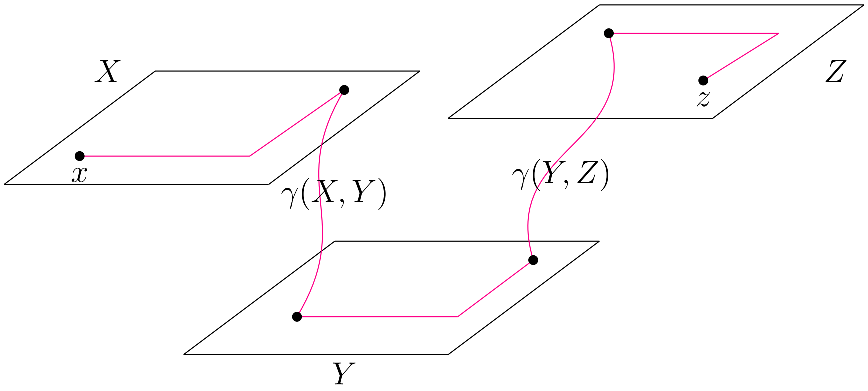

Figure 4.4: An ordered standard path from x to z in $\Gamma(G, S)$. It is drawn as if $\mathbb{G}_K[X, Z] = \{X < Y < Z\}$ and as if $A_t = H_t = \mathbb{Z}^2$ for all $1 \leq t \leq T$. Made using TikZiT.

$\gamma(Y_l, Y_{l+1})$ between these cosets (as defined in Definition 4.63). More formally, a standard path ζ from x to z is defined to be any concatenation

$$\zeta = \zeta_1 \ \gamma(Y_1, Y_2) \ \zeta_2 \ \gamma(Y_2, Y_3) \ \zeta_3 \ \dots \ \zeta_{L-1} \ \gamma(Y_{L-1}, Y_L) \ \zeta_L$$

where

- ζ_1 is a standard path from $x \in Y_1$ to $p(Y_1, Y_2) \in Y_1$;
- for $2 \leq l \leq L - 1$, ζ_l is a standard path from $p(Y_l, Y_{l-1})$ to $p(Y_l, Y_{l+1})$;
- ζ_L is a standard path from $p(Y_L, Y_{L-1})$ to z.

Given $1 \leq r \leq R$, the *r'th ordered standard path from x to z* is the unique standard path such that for $1 \leq l \leq L$, we have that ζ_l is the r'th ordered standard path between its initial and terminal vertices. See Figure 4.4.

Proposition 4.65. *For sufficiently large K, there exist constants $\lambda_3 \geq 1$ and $\mu_3 \geq 0$, neither of which depend on x or z, such that a standard path from x to z is a (λ_3, μ_3)-quasigeodesic in G.*

Proof. We use the notation for a standard path from x to z that is given in Definition 4.64.

Let $\hat{\gamma}$ denote a geodesic from x to z in the coned-off graph $\hat{G}$. By Lemma 4.50 (2), the geodesic $\hat{\gamma}$ must pass through the cosets $Y_1, Y_2, \dots, Y_L$ and the intersection of $\hat{\gamma}$ with one of these cosets is either an edge or possibly a single vertex (in the case of Y_1 and Y_L). For $1 \leq l \leq L$, let p_l be the vertex where $\hat{\gamma}$ enters Y_l (or let $p_l = x$ if $l = 1$). For $1 \leq l \leq L$, let q_l be the vertex where $\hat{\gamma}$ leaves Y_l (or let $q_l = z$ if $l = L$). Now the vertices $V := \{p_l : 1 \leq l \leq L\} \cup \{q_l : 1 \leq l \leq L\}$ admit a natural ordering of the following form:

$$p_1 \leq q_1 \leq p_2 \leq q_2 \leq p_3 \leq \dots \leq p_L \leq z$$

and given $v, v' \in V$ we say that $v < v'$ if $v \leq v'$ and $v \neq v'$. Another possible ordering on V, which we denote by $<<$, is to say that $v << v'$ if $\hat{\gamma}$ hits v before v' on its journey from x to z. We write $v \leq\leq v'$ if $v << v'$ or $v = v'$.

Claim. *The total orders $<$ and $<<$ are the same.*

Proof. First, note that by definition $p_l \leq\leq q_l$. Second, recall that the intersection of $\hat{\gamma}$ with a coset Y_l is either an edge or a vertex. It follows that there cannot exist a $v \in V$ satisfying $p_l << v << q_l$. Third, note that we clearly have $p_1 << p_l$ for all $2 \leq l \leq L$ and $q_l << q_L$ for all $1 \leq l \leq L-1$. Fourth, suppose for a contradiction that $2 \leq l < l' \leq L-1$ and yet $p_{l'} << p_l$. This implies that $\hat{\gamma}$ enters the coset $Y_{l'}$ before it enters the coset Y_l. Lemma 4.48 implies that $\hat{\gamma}$ first enters Y_l close to $\pi_{Y_l}(X)$. But it also implies that $\hat{\gamma}$ enters Y_l close to $\pi_{Y_l}(Y_{l'})$, since $\hat{\gamma}$ passes through $Y_{l'}$ beforehand. This implies that $\pi_{Y_l}(X)$ and $\pi_{Y_l}(Y_{l'})$ are close. But $d_{Y_l}(X, Y_{l'}) > K$ since $X < Y_l < Y_{l'}$ in $\mathbb{G}_K[X, Z]$, and so by choosing K sufficiently large (i.e. much larger than the constant M given by Lemma 4.48) we have a contradiction. So if $l < l'$ then $p_l \leq\leq p_{l'}$.

The claim follows from the four statements above. □

So $\hat{\gamma}$ moves through the cosets $Y_1, Y_2, \ldots, Y_L$ in the natural order. Proposition 4.49 tells us that a lift of $\hat{\gamma}$ in G is a (λ, μ)-quasigeodesic for some constants $\lambda \geq 1$ and $\mu \geq 0$. We denote some chosen lift by γ. The claim implies that

$$\operatorname{length}(\gamma) \geq \sum_{l=1}^{L} d_G(p_l, q_l) + \sum_{l=1}^{L-1} d_G(q_l, p_{l+1})$$

Further, Lemma 4.48 implies that there exists some uniform constant M such that for $1 \leq l \leq L$ we have $d_G(p_l, p(Y_l, Y_{l-1})) \leq M$ and $d_G(q_l, p(Y_l, Y_{l+1})) \leq M$.

Since peripheral cosets are undistorted in G (see Proposition 4.47), and since standard paths *within* cosets are quasigeodesics (see Proposition 4.61), there exist constants $\lambda' \geq 1$ and $\mu' \geq 0$ such that for $1 \leq l \leq L$ we have

$$\operatorname{length}(\zeta_l) \leq \lambda' d_G(p(Y_l, Y_{l-1}), p(Y_l, Y_{l+1})) + \mu'$$

We may assume for simplicity that $\lambda' \geq \lambda_2$ and $\mu' \geq \mu_2$. It follows that

$$\begin{aligned}
\text{length}(\zeta) &= \sum_{l=1}^{L} \text{length}(\zeta_l) + \sum_{l=1}^{L-1} \text{length}(\gamma(Y_l, Y_{l+1})) \\
&\leq \lambda' \sum_{l=1}^{L} d_G(p(Y_l, Y_{l-1}), p(Y_l, Y_{l+1})) + \lambda_2 \sum_{l=1}^{L-1} d_G(p(Y_l, Y_{l+1}), p(Y_{l+1}, Y_l)) + L\mu' + (L-1)\mu_2 \\
&\leq \lambda' \sum_{l=1}^{L} (d_G(p_l, q_l) + 2M) + \lambda_2 \sum_{l=1}^{L-1} (d_G(q_l, p_{l+1}) + 2M) + L\mu' + (L-1)\mu_2 \\
&\leq \lambda' \text{length}(\gamma) + 2M\lambda' L + 2M\lambda'(L-1) + L\mu' + (L-1)\mu' \\
&= \lambda' \text{length}(\gamma) + (2L-1)(2M\lambda' + \mu')
\end{aligned}$$

Since $\hat{\gamma}$ contains an edge in Y_l for each $2 \leq l \leq L-1$, we have $\text{length}(\hat{\gamma}) \geq L-2$. So certainly $\text{length}(\gamma) \geq L-2$. Hence,

$$\text{length}(\zeta) \leq \lambda' \text{length}(\gamma) + (2\text{length}(\gamma) + 3)(2M\lambda' + \mu')$$

and so we are done since $\text{length}(\gamma) \leq \lambda d_G(x, z) + \mu$. □

4.7 Proof of the main theorem part one

Let G be relatively hyperbolic with respect to virtually abelian peripheral subgroups as in Theorem 4.1. Let us retain all the notation and assumptions described at the beginning of Section 4.6.

However, in contrast to Section 4.6 where the transverse quasigeodesics $\gamma(X, Y)$ were chosen arbitrarily, for this proof, we want to choose them so that the following two conditions hold. Let $X, Y \in \mathbb{G}$ and suppose $\{X, Y\}$ is an edge in $\mathcal{P}_K(\mathbb{G})$.

($\natural$1) We choose $\gamma(X, Y)$ from $p(X, Y) \in G$ to $p(Y, X) \in G$ to be such that $\gamma(X, Y)$ is non-empty and not entirely contained in any coset subgraph $\mathcal{C}(W) \subset \Gamma(G, S)$.

($\natural$2) We choose $\gamma(X, Y)$ in such a way that we can deduce from the final χ letters of $\gamma(X, Y)$ which peripheral subgroup $H_1, H_2, \ldots, H_T$ is such that $Y = gH_t$, where χ is a constant that only depends on T.

It is always possible to choose $\gamma(X, Y)$ such that ($\natural$1) holds (for uniform choices of $\lambda_2 \geq 1$, $\mu_2 \geq 0$) since we can force $\gamma(X, Y)$ to take a small diversion out of any coset subgraph $\mathcal{C}(W)$. It is possible to choose $\gamma(X, Y)$ such that ($\natural$2) holds since there are only T peripheral subgroups. One solution would be this: if Y is a coset of H_t, then you could insist that $\gamma(X,Y)$ consists of a geodesic from $p(X,Y)$ to $p(Y,X)$ followed by the word $(a_{t,1}a_{t,1}^{-1})^t$ - in this case you could choose $\chi = 2T$.

Remark 4.66. *The necessity of ($\natural$1) and ($\natural$2) is an artifact of the fact that distinct peripheral cosets X, Y can have non-empty intersection. If we have a single peripheral subgroup H this cannot happen and we may choose the transverse quasigeodesics $\gamma(X,Y)$ to simply be geodesics.*

The purpose of ($\natural$1) is so that we can recognise, just by looking at the letters of $\gamma(X,Y)$ whether it is a transverse quasigeodesic or contained in a peripheral coset.

The purpose of ($\natural$2) is so that we can recognise, just by looking at the final letters of $\gamma(X,Y)$, which peripheral subgroup Y corresponds to.

Other solutions could have been chosen: I could have added a special character to the alphabet to indicate when you are travelling along a transverse geodesic and T more special characters to indicate which peripheral subgroups $\{H_1, H_2, \dots, H_T\}$ are relevant to the transverse quasigeodesic.

The map $F : G \to T_W \times T_W \times \cdots \times T_W$

Recall that S is the finite generating set of G. Let $A = S \cup S^{-1}$ be our finite alphabet, let W be the set of all finite (but non-empty) words on A and let T_W be the associated sentence-tree. Given $x \in G$, we will now define a sentence $F_r(x) \in T_W$ by partitioning the standard path from e to x into multiple words.

Definition 4.67. We aim to define a function $F_r : G \to T_W$. Let $x \in G$. Write $X = xH_1$. Let $L = L[H_1, X]$ and suppose

$$\mathbb{G}_K[H_1, X] = \{H_1 = X_1 < X_2 < \cdots < X_{L-1} < X_L = X\}$$

Let $1 \leq r \leq R$. Let ζ denote the r'th ordered standard path from e to x in $\Gamma(G,S)$. Recall from Definition 4.64 that we can write ζ as a concatenation of paths in the follow manner

$$\zeta = \zeta_1\ \gamma(X_1, X_2)\ \zeta_2\ \gamma(X_2, X_3)\ \zeta_3\ \ldots\ \zeta_{L-1}\ \gamma(X_{L-1}, X_L)\ \zeta_L$$

Let $1 \leq l \leq L$ and suppose that $X_l = gH_t$. We would like to define a sentence $\alpha_l \in T_W$. It has a different form depending on whether $r \leq R_t$ or $r > R_t$. Note that ζ_l has the form

$$\zeta_l = \begin{cases} a_{t,r}^{q_r} a_{t,r+1}^{q_{r+1}} \ldots a_{t,R_t}^{q_{R_t}} a_{t,1}^{q_1} a_{t,2}^{q_2} \ldots a_{t,r-1}^{q_{r-1}} \psi_{t,j} & \text{if } r \leq R_t \\ a_{t,1}^{q_1} a_{t,2}^{q_2} \ldots a_{t,R_t}^{q_{R_t}} \psi_{t,j} & \text{if } r > R_t \end{cases}$$

Then we define

$$\alpha_l = \begin{cases} (\overline{a_{t,r}})^{q_r}\ \overline{a_{t,r+1}^{q_{r+1}}} \dots \overline{a_{t,R_t}^{q_{R_t}}}\ \overline{a_{t,1}^{q_1}}\ \overline{a_{t,2}^{q_2}} \dots \overline{a_{t,r-1}^{q_{r-1}}}\ \overline{\psi_{t,j}} & \text{if } r \leq R_t \\ \overline{a_{t,1}^{q_1}}\ \overline{a_{t,2}^{q_2}} \dots \overline{a_{t,R_t}^{q_{R_t}}}\ \overline{\psi_{t,j}} & \text{if } r > R_t \end{cases}$$

where the meaning of $(\overline{a_{t,r}})^{q_r}$ when $q_r < 0$ is described in Notation 0.14. Further, for $1 \leq l \leq L-1$, we define a sentence $\alpha_{l,l+1}$ by

$$\alpha_{l,l+1} = \overline{\gamma(X_l, X_{l+1})}$$

Finally, we define

$$F_r(x) = \alpha_1\alpha_{1,2}\alpha_2\alpha_{2,3}\dots\alpha_{L-1,L}\alpha_L$$

Intuitively, $F_r(x)$ "crushes" everything in the r'th ordered standard path from e to x other than the directions corresponding to generators of the form $a_{t,r}$.

Remark 4.68. *We don't allow empty words to appear in sentences. For example, if $\psi_{t,j}$ is empty, then it simply doesen't appear in the sentence α_l and we don't write something like $\overline{\emptyset}$. So the vertices of T_W are in bijection with the free monoid over non-empty words on $S \cup S^{-1}$.*

Recall that $R = \max_t R_t$.

Definition 4.69. We define $F : G \to T_W \times T_W \times \dots \times T_W$ by $F(x) = (F_1(x), F_2(x), \dots, F_R(x))$.

A quasiisometric embedding into a product of trees

By Corollary 4.46, there exists a regular map $\phi : G \to \prod_{q=1}^{Q} T_{\{0,1\}}$ where $Q \leq \max(\text{asdim}(G), R+1) + 1$. Recall as well that ϕ is the composition $G \hookrightarrow X(G) \to \prod_{q=1}^{Q} T_{\{0,1\}}$ where the second map comes from Corollary 4.45.

Theorem 4.70. *There exist $\lambda_4 \geq 1$ and $\mu_4 \geq 0$ such that*

$$F \times \phi : G \to T_W \times T_W \times \dots \times T_W \times \prod_{q=1}^{Q} T_{\{0,1\}}$$

is a (λ_4, μ_4)-quasiisometric embedding.

Remark 4.71. *If all we needed was a quasiisometric embedding of G into a product of trees then we would have an easier time. It follows from Corollary 4.56, Corollary 4.45, Theorem 4.36 that we have a quasiisometric embedding*

$$G \to \mathcal{C}_K(\mathbb{G}) \times X(G) \to \mathcal{C}_K^{\mathcal{T}}(\mathbb{G}) \times \prod_{q=1}^{Q} T_{\{0,1\}}$$

The tree of cosets $\mathcal{C}_K^{\mathcal{T}}(\mathbb{G})$ is then simply a tree of virtually abelian cosets. So $\mathcal{C}_K^{\mathcal{T}}(\mathbb{G})$ is quasiisometric to a tree of flats. It is not difficult to see that there is a quasiisometric embedding of a tree of flats into a product of trees.

However, the problem with the quasiisometric embedding above is twofold. Firstly, the quasiisometric embedding of the tree of flats is into a product of arbitrary infinite-valence trees as opposed to a product of sentence-trees. We need our image to be inside sentence-trees so that we can apply the theory described in Chapter 3. Secondly, and more importantly, the quasiisometry between the tree of virtually abelian cosets and the tree of flats loses lots of structure. The special aspect of the functions $F_r : G \to T_W$ is that we can recover from $F_r(x)$ exactly what $x \in G$ is, and $F_r(x)$ has a very specific structure coming from the structure of ordered standard paths. This would all be lost if we used a messy quasiisometry between the tree of virtually abelian cosets and the tree of flats.

Let me describe why there is a quasiisometric embedding of a tree of two-dimensional flats into a product of two trees. Let us say that each flat has two axes: horizontal and vertical. You may indeed choose these axes arbitarily for each flat. There is a map from the tree of flats onto a tree given by crushing each flat onto the horizontal axis. Similarly, there is a map from the tree of flats onto a tree given by crushing each flat onto the vertical axis. It is not difficult to show that the product of these two maps provides a quasiisometric embedding into the product of two trees.

The map $F_r : G \to T_W$ is meant to imitate one of these horizontal/vertical crushing maps. By fixing a basepoint $p_0(gH_t) \in gH_t$, we are essentially providing a set of R_t canonical axes within gH_t (the first axis being $p_0(gH_t)\langle a_{t,1}\rangle$, the second axis being $p_0(gH_t)\langle a_{t,2}\rangle$ and so on.) The map F_r then crushes every axis within gH_t other than the r'th axis.

So the idea is simple, but the particular nature of the map $F_r : G \to T_W$ makes the proof lengthy.

Proof of Theorem 4.70. Let $x, z \in G$. By Corollary 4.56, Theorem 4.36 and Corollary 4.44, we have

$$d_G(x,z) \approx d_{\mathcal{C}_K^{\mathcal{T}}(\mathbb{G})}(\iota_1(x), \iota_1(z)) + d_{\prod_{q=1}^{Q} T_{\{0,1\}}}(\phi(x), \phi(z))$$

It follows that, in order to prove Theorem 4.70, we simply need to show that

$$d_{\mathcal{C}_K^{\mathcal{T}}(\mathbb{G})}(\iota_1(x), \iota_1(z)) \approx d(F(x), F(z))$$

First, we develop some notation. Write $X := xH_1$, $Z := zH_1$, $L_X := L[H_1, X]$, $L_Z := L[H_1, Z]$, $\mathbb{G}_K[H_1, X] = \{H_1 = X_1 < X_2 < \cdots < X_{L_X} = X\}$, $\mathbb{G}_K[H_1, Z] = \{H_1 = Z_1 < Z_2 < \cdots < Z_{L_Z} = Z\}$ and $l_\dagger = \max\{l : X_l = Z_l\}$. So $l_\dagger$ is the index at which the

$$\begin{tikzcd}
X = X_{L_X} = Y_1 & & Z = Z_{L_Z} = Y_L \\
X_{L_X-1} = Y_2 \arrow[u] & & Z_{L_Z-1} = Y_{L-1} \arrow[u] \\
\vdots \arrow[u] & & \vdots \arrow[u] \\
X_{l_\dagger+1} = Y_{l_*-1} \arrow[u] & & Z_{l_\dagger+1} = Y_{l_*+1} \arrow[u] \\
& X_{l_\dagger} = Y_{l_*} = Z_{l_\dagger} \arrow[ul] \arrow[ur] & \\
& X_{l_\dagger-1} = Z_{l_\dagger-1} \arrow[u] & \\
& \vdots \arrow[u] & \\
& H_1 = X_1 = Z_1 \arrow[u] &
\end{tikzcd}$$

Figure 4.5: The paths $\mathbb{G}_K[H_1, X]$ and $\mathbb{G}_K[H_1, Z]$. Made using `https://q.uiver.app`.

paths $\mathbb{G}_K[H_1, X]$ and $\mathbb{G}_K[H_1, Z]$ diverge in $\mathcal{P}_K(\mathbb{G})$. Further, let

$$X = Y_1 \to Y_2 \to \cdots \to Y_{L-1} \to Y_L = Z$$

denote the unique geodesic from X to Z in $\mathcal{T} = \mathcal{T}_K(\mathbb{G}, H_1)$. Suppose $1 \leq l_* \leq L$ is the index such that $Y_{l_*} = X_{l_\dagger} = Z_{l_\dagger}$. See Figure 4.5.

Let ζ be the composition of the following paths in $\mathcal{C}_K^{\mathcal{T}}(\mathbb{G})$:

a standard path from $x \in \mathcal{C}(Y_1)$ to $p(Y_1, Y_2) \in \mathcal{C}(Y_1)$ (in the sense of Definition 4.59)

the transverse edge between $p(Y_1, Y_2)$ and $p(Y_2, Y_1)$

a standard path from $p(Y_2, Y_1) \in \mathcal{C}(Y_2)$ to $p(Y_2, Y_3) \in \mathcal{C}(Y_2)$ (in the sense of Definition 4.59)

the transverse edge between $p(Y_2, Y_3)$ and $p(Y_3, Y_2)$

a standard path from $p(Y_3, Y_2) \in \mathcal{C}(Y_3)$ to $p(Y_3, Y_4) \in \mathcal{C}(Y_3)$ (in the sense of Definition 4.59)

the transverse edge between $p(Y_3, Y_4)$ and $p(Y_4, Y_3)$

$\vdots$

the transverse edge between $p(Y_{L-1}, Y_L)$ and $p(Y_L, Y_{L-1})$

a standard path from $p(Y_L, Y_{L-1}) \in \mathcal{C}(Y_L)$ to $z \in \mathcal{C}(Y_L)$ (in the sense of Definition 4.59)

Let ζ_l indicate the intersection of ζ with the coset $\mathcal{C}(Y_l)$, in other words, the standard path from $p(Y_l, Y_{l-1}) \in \mathcal{C}(Y_l)$ to $p(Y_l, Y_{l+1}) \in \mathcal{C}(Y_l)$.

Claim 1. *There exists $\lambda \geq 1$ and $\mu \geq 0$, not depending on x or z, such that ζ is a (λ, μ)-quasigeodesic in $\mathcal{C}_K^{\mathcal{T}}(\mathbb{G})$.*

Proof of claim 1. Let γ denote a geodesic from x to z in $\mathcal{C}_K^{\mathcal{T}}(\mathbb{G})$. For $1 \leq l \leq L$, let γ_l indicate the intersection of γ with the coset Y_l. We have that

$$d_{\mathcal{C}_K^{\mathcal{T}}(\mathbb{G})}(x, z) = \text{length}(\gamma) = \sum_{l=1}^{L} \text{length}(\gamma_l) + K(L-1)$$

Now, by Proposition 4.61, since γ_l and ζ_l have the same initial and terminal vertices, we have that $\text{length}(\zeta_l) \leq \lambda_1 \text{length}(\gamma_l) + \mu_1$. Hence,

$$\begin{aligned}\text{length}(\zeta) &= \sum_{l=1}^{L} \text{length}(\zeta_l) + K(L-1) \\ &\leq \lambda_1 \sum_{l=1}^{L} \text{length}(\gamma_l) + L\mu_1 + K(L-1) \\ &\leq \lambda_1 d_{\mathcal{C}_K^{\mathcal{T}}(\mathbb{G})}(x, z) + L\mu_1\end{aligned}$$

If we note that $L \leq d_{\mathcal{C}_K^{\mathcal{T}}(\mathbb{G})}(x, z) + 1$ then we are done. ☐

Given a coset Y_l, where $1 \leq l \leq L$, let $s_r(Y_l)$ denote the amount of times that the path ζ crosses a generator of the form $a_{t,r}^{\pm 1}$ in the coset subgraph $\mathcal{C}(Y_l) \subset \mathcal{C}_K^{\mathcal{T}}(\mathbb{G})$. Further, let us define $s_r = \sum_{l=1}^{L} s_r(Y_l)$ and $s = \sum_{r=1}^{R} s_r$.

Claim 2. *There exist $\lambda' \geq 1$ and $\mu' \geq 0$, not depending on x or z, such that*

$$s \leq \textit{length}(\zeta) \leq \lambda' s + \mu'$$

Proof of claim 2. The lower bound of the inequality follows from the definition of s.

Write $\zeta_l = \psi_l \nu_l \psi_l'$ where ν_l indicates the segment of ζ_l that consists entirely of the abelian generators $a_{t,r}$ and ψ_l, ψ_l' indicate the standard non-abelian paths between cosets.

For the upper bound of the inequality, note that we have

$$\begin{aligned}\text{length}(\zeta) &= \sum_{l=1}^{L} \text{length}(\zeta_l) + K(L-1) \\ &= s + \sum_{l=1}^{L} (\text{length}(\psi_l) + \text{length}(\psi_l')) + K(L-1) \\ &\leq s + 2LI + K(L-1)\end{aligned}$$

If $1 < l < l_*$, it follows from Lemma 4.62 that the intersection of ζ with the coset Y_l contains at least one abelian generator $a_{t,r}^{\pm 1}$ since $d_{Y_l}(Y_{l-1}, Y_{l+1}) > K$. For the same reason, if $l_* < l < L$ then the intersection of ζ with the coset Y_l also contains at least one abelian generator $a_{t,r}^{\pm 1}$. It follows that $s \geq L - 3$. Hence,

$$\text{length}(\zeta) \leq s + 2(s+3)I + K(s+2) = (1 + 2I + K)s + 6I + 2K$$

and we are done. □

As described in Definition 4.67, we can write $F_r(x)$ and $F_r(z)$ as concatenations of sentences. Let us write

$$F_r(x) = \alpha_1 \alpha_{1,2} \alpha_2 \alpha_{2,3} \dots \alpha_{L_X}$$

and

$$F_r(z) = \beta_1 \beta_{1,2} \beta_2 \beta_{2,3} \dots \beta_{L_Z}$$

Claim 3. *We have*

$$\begin{aligned} d_{T_W}(F_r(x), F_r(z)) &= d_{T_W}(\alpha_1 \alpha_{1,2} \dots \alpha_{l_\dagger}, \beta_1 \beta_{1,2} \dots \beta_{l_\dagger}) \\ &\quad + \sum_{l=l_\dagger+1}^{L_X} \mathit{length}(\alpha_l) + \sum_{l=l_\dagger+1}^{L_Z} \mathit{length}(\beta_l) + L - 1 \end{aligned}$$

Proof of claim 3. Write

$$\begin{aligned} u_1 &= \alpha_1 \alpha_{1,2} \dots \alpha_{l_\dagger} \\ u_2 &= \alpha_1 \alpha_{1,2} \dots \alpha_{l_\dagger} \alpha_{l_\dagger, l_\dagger+1} \\ u_3 &= \alpha_1 \alpha_{1,2} \dots \alpha_{l_\dagger} \alpha_{l_\dagger, l_\dagger+1} \alpha_{l_{\dagger+1}} \\ &\vdots \\ u_m &= \alpha_1 \alpha_{1,2} \dots \alpha_{l_\dagger} \alpha_{l_\dagger, l_\dagger+1} \alpha_{l_{\dagger+1}} \dots \alpha_{L_X} = F_r(x) \end{aligned}$$

Similarly, we choose

$$\begin{aligned}
v_1 &= \beta_1\beta_{1,2}\dots\beta_{l_\dagger}\\
v_2 &= \beta_1\beta_{1,2}\dots\beta_{l_\dagger}\beta_{l_\dagger,l_\dagger+1}\\
v_3 &= \beta_1\beta_{1,2}\dots\beta_{l_\dagger}\beta_{l_\dagger,l_\dagger+1}\beta_{l_{\dagger+1}}\\
&\vdots\\
v_n &= \beta_1\beta_{1,2}\dots\beta_{l_\dagger}\beta_{l_\dagger,l_\dagger+1}\beta_{l_{\dagger+1}}\dots\beta_{L_Z} = F_r(z)
\end{aligned}$$

Suppose, for a contradiction, that v_1 descends from u_2. If v_1 were to descend from u_2, then $\alpha_1\alpha_{1,2}\dots\alpha_{l_\dagger}\alpha_{l_\dagger,l_\dagger+1}$ would be a subsentence at the start of $\beta_1\beta_{1,2}\dots\beta_{l_\dagger}$. Noting that $\alpha_1\alpha_{1,2}\dots\alpha_{l_\dagger-1}\alpha_{l_\dagger-1,l_\dagger} = \beta_1\beta_{1,2}\dots\beta_{l_\dagger-1}\beta_{l_\dagger-1,l_\dagger}$, it follows that $\alpha_{l_\dagger}\alpha_{l_\dagger,l_\dagger+1}$ is a subsentence of $\beta_{l_\dagger}$. However, due to our choice of transverse quasigeodesics described in (♮1), the path corresponding to $\alpha_{l_\dagger}\alpha_{l_\dagger,l_\dagger+1}$ leaves the coset $\mathcal{C}(X_{l_\dagger}) = \mathcal{C}(Z_{l_\dagger})$. In contrast, the path corresponding to $\beta_{l_\dagger}$ remains entirely within $\mathcal{C}(X_{l_\dagger}) = \mathcal{C}(Z_{l_\dagger})$. So we have a contradiction.

For similar reasons, u_1 cannot descend from v_2.

We claim that the path

$$u_m \to u_{m-1} \to \dots \to u_2 \to u_1 \to v_1 \to v_2 \to \dots \to v_{n-1} \to v_n$$

is a geodesic in T_W. If v_1 does not descend from u_1 and u_1 does not descend from v_1 then it is clearly a geodesic. If v_1 descends from u_1 then, since $d_{T_W}(u_1, u_2) = 1$ and since v_1 does not descend from u_2, we can also deduce that it is a geodesic. Similarly, if u_1 descends from v_1 it is a geodesic. Hence,

$$\begin{aligned}
&d_{T_W}(F_r(x), F_r(z)) = d_{T_W}(\alpha_1\alpha_{1,2}\dots\alpha_{L_X}, \beta_1\beta_{1,2}\dots\beta_{L_Z})\\
&= d_{T_W}(\alpha_1\alpha_{1,2}\dots\alpha_{l_\dagger}, \beta_1\beta_{1,2}\dots\beta_{l_\dagger})\\
&+ \sum_{l=l_\dagger+1}^{L_X} d_{T_W}(\alpha_1\alpha_{1,2}\dots\alpha_{l-1}, \alpha_1\alpha_{1,2}\dots\alpha_l\alpha_{l-1,l}) + d_{T_W}(\alpha_1\alpha_{1,2}\dots\alpha_{l-1,l}, \alpha_1\alpha_{1,2}\dots\alpha_{l-1,l}\alpha_l)\\
&+ \sum_{l=l_\dagger+1}^{L_Z} d_{T_W}(\beta_1\beta_{1,2}\dots\beta_{l-1}, \beta_1\beta_{1,2}\dots\beta_{l-1}\beta_{l-1,l}) + d_{T_W}(\beta_1\beta_{1,2}\dots\beta_{l-1,l}, \beta_1\beta_{1,2}\dots\beta_{l-1,l}\beta_l)
\end{aligned}$$

By noting that

$$d_{T_W}(\alpha_1\alpha_{1,2}\dots\alpha_{l-1}, \alpha_1\alpha_{1,2}\dots\alpha_{l-1}\alpha_{l-1,l}) = d_{T_W}(\beta_1\beta_{1,2}\dots\beta_{l-1}, \beta_1\beta_{1,2}\dots\beta_{l-1}\beta_{l-1,l}) = 1$$

and

$$d_{T_W}(\alpha_1\alpha_{1,2}\dots\alpha_{l-1,l}, \alpha_1\alpha_{1,2}\dots\alpha_{l-1,l}\alpha_l) = \text{length}(\alpha_l)$$

and

$$d_{T_W}(\beta_1\beta_{1,2}\dots\beta_{l-1,l}, \beta_1\beta_{1,2}\dots\beta_{l-1,l}\beta_l) = \text{length}(\beta_l)$$

the claim follows. $\square$

Claim 4. *If $l_\dagger < l \le L_X$ then $s_r(X_l) \le$ length$(\alpha_l) \le s_r(X_l) + R$ and if $l_\dagger < l \le L_Z$ then $s_r(Z_l) \le$ length$(\beta_l) \le s_r(Z_l) + R$.*

Proof of claim 4. Let $l_\dagger < l \le L_X$ and suppose X_l is a coset of H_t. We have two cases: either $r \le R_t$ or $r > R_t$. If $r \le R_t$, then α_l has the form

$$(\overline{a_{t,r}})^{q_r}\overline{a_{t,r+1}^{q_{r+1}}}\dots\overline{a_{t,r-1}^{q_{r-1}}}\,\overline{\psi_{t,j}}$$

where

$$(a_{t,r})^{q_r}a_{t,r+1}^{q_{r+1}}\dots a_{t,r-1}^{q_{r-1}}\psi_{t,j}$$

is the r'th ordered standard path from $p(X_l, X_{l-1})$ to $p(X_l, X_{l+1})$. Hence,

$$s_r(X_l) \le \text{length}(\alpha_l) \le s_r(X_l) + R_t \le s_r(X_l) + R$$

where we have used the fact that $s_r(X_l) = |q_r|$. If, $r > R_t$, then $s_r(X_l) = 0$ and length$(\alpha_l) \le R_t + 1 \le R$. So, in either case, we have $s_r(X_l) \le$ length$(\alpha_l) \le s_r(X_l) + R$.

For similar reasons, for $l_\dagger < l \le L_Z$, we have $s_r(Z_l) \le$ length$(\beta_l) \le s_r(Z_l) + R$ and we are done. $\square$

Claim 5. *We have $s_r(Y_{l_*}) \le d_{T_W}(\alpha_1\alpha_{1,2}\dots\alpha_{l_\dagger}, \beta_1\beta_{1,2}\dots\beta_{l_\dagger}) \le s_r(Y_{l_*}) + 2(R+1)$.*

Proof of claim 5. Recall that $Y_{l_*} = X_{l_\dagger} = Z_{l_\dagger}$. Suppose that Y_{l_*} is a coset of H_t. We have two cases: $r \le R_t$ and $r > R_t$.

Suppose first that $r > R_t$. Then $s_r(Y_{l_*}) = 0$. Further, length$(\alpha_{l_\dagger}) \le R_t + 1 \le R + 1$ and length$(\beta_{l_\dagger}) \le R_t + 1 \le R + 1$. Since $\alpha_1\alpha_{1,2}\dots\alpha_{l_\dagger-1,l_\dagger} = \beta_1\beta_{1,2}\dots\beta_{l_\dagger-1,l_\dagger}$, it follows that

$$d_{T_W}(\alpha_1\alpha_{1,2}\dots\alpha_{l_\dagger}, \beta_1\beta_{1,2}\dots\beta_{l_\dagger}) = d_{T_W}(\alpha_{l_\dagger}, \beta_{l_\dagger}) \le 2(R+1)$$

So $s_r(Y_{l_*}) \le d_{T_W}(\alpha_1\alpha_{1,2}\dots\alpha_{l_\dagger}, \beta_1\beta_{1,2}\dots\beta_{l_\dagger}) \le s_r(Y_{l_*}) + 2(R+1)$.

Now suppose that $r \le R_t$. Then $\alpha_{l_\dagger}$ has the form

$$\alpha_{l_\dagger} = (\overline{a_{t,r}})^{q_r}\ \overline{a_{t,r+1}^{q_{r+1}}}\dots\overline{a_{t,R_t}^{q_{R_t}}}\ \overline{a_{t,1}^{q_1}}\ \overline{a_{t,2}^{q_2}}\dots\overline{a_{t,r-1}^{q_{r-1}}}\ \overline{\psi_{t,j}}$$

and $\beta_{l_\dagger}$ has the form

$$\beta_{l_\dagger} = (\overline{a_{t,r}})^{q'_r}\ \overline{a_{t,r+1}^{q'_{r+1}}}\dots\overline{a_{t,R_t}^{q'_{R_t}}}\ \overline{a_{t,1}^{q'_1}}\ \overline{a_{t,2}^{q'_2}}\dots\overline{a_{t,r-1}^{q'_{r-1}}}\ \overline{\psi'_{t,j}}$$

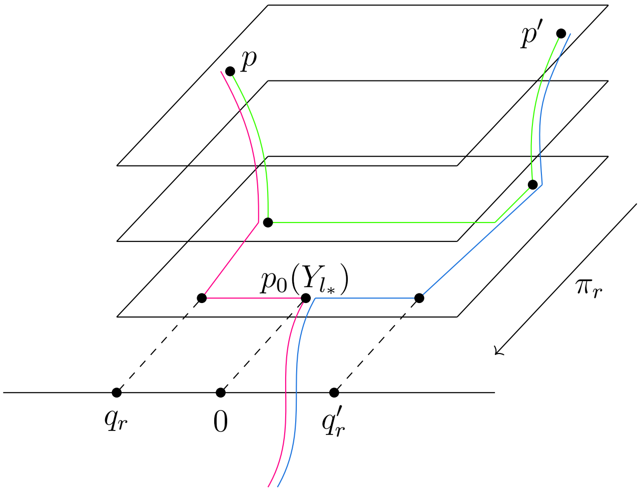

Figure 4.6: The green line is the path ζ_{l_*}. The red line is a section of the r'th ordered standard path from e to x. Similarly, the blue line is a section of the r'th ordered standard path from e to z. Made using TikZiT.

It follows that

$$|q_r - q'_r| \leq d_{T_W}(\alpha_{l_\dagger}, \beta_{l_\dagger}) \leq |q_r - q'_r| + 2R_t \tag{4.5}$$

Now consider the coset $p_0(Y_{l_*})A_t \subset Y_{l_*}$. We identify $p_0(Y_{l_*}) \in p_0(Y_{l_*})A_t$ with $(0, 0, \dots, 0) \in \mathbb{Z}^{R_t}$. More precisely, we identify $p_0(Y_{l_*})a_{t,1}^{q_1}a_{t,2}^{q_2}\dots a_{t,R_t}^{q_{R_t}} \in p_0(Y_{l_*})A_t$ with $(q_1, q_2, \dots, q_{R_t}) \in \mathbb{Z}^{R_t}$. Under this identification, let $\pi_r : \mathbb{Z}^{R_t} \to \mathbb{Z}$ be the projection onto the r'th factor. Let us also write $p := p(Y_{l_*}, Y_{l_*-1})$ and $p' := p(Y_{l_*}, Y_{l_*+1})$. The intersection of ζ with Y_{l_*}, which we have denoted by ζ_{l_*}, is precisely a standard path from p to p' in the coset Y_{l_*}. Now ζ_{l_*} has the form $\psi_{t,j}^{-1}\nu\psi'_{t,j}$ where ν corresponds to a geodesic in $\mathbb{Z}^{R_t}$ from $p\psi_{t,j}^{-1}$ to $p'(\psi'_{t,j})^{-1}$ and $\psi_{t,j}$ and $\psi_{t,j'}$ are the standard non-abelian paths arising from our description of $\alpha_{l_\dagger}$ and $\beta_{l_\dagger}$. It should be clear that

$$s_r(Y_{l_*}) = d_{\mathbb{Z}}(\pi_r(p\psi_{t,j}^{-1}), \pi_r(p'(\psi'_{t,j})^{-1})) = \text{length}(\pi_r(\nu)) \tag{4.6}$$

Finally, observe that $\pi_r(p\psi_{t,j}^{-1}) = q_r$ and $\pi_r(p'(\psi'_{t,j})^{-1}) = q'_r$. See Figure 4.6.

So we conclude from (4.5) and (4.6) that

$$s_r(Y_{l_*}) \leq d_{T_W}(\alpha_{l_\dagger}, \beta_{l_\dagger}) \leq s_r(Y_{l_*}) + 2R_t \leq s_r(Y_{l_*}) + 2R$$

As before, since $\alpha_1\alpha_{1,2}\dots\alpha_{l_\dagger-1,l_\dagger} = \beta_1\beta_{1,2}\dots\beta_{l_\dagger-1,l_\dagger}$, we know that

$$d_{T_W}(\alpha_1\alpha_{1,2}\dots\alpha_{l_\dagger}, \beta_1\beta_{1,2}\dots\beta_{l_\dagger}) = d_{T_W}(\alpha_{l_\dagger}, \beta_{l_\dagger})$$

and so we are done. □

Claim 3, Claim 4 and Claim 5 combine to give us the following claim.

Claim 6. *We have*

$$s_r + (L-1) \leq d_{T_W}(F_r(x), F_r(z)) \leq s_r + (R+1)(L+1)$$

Proof of claim 6. We have the following chain of inequalities.

$$\begin{aligned}
d_{T_W}(F_r(x), F_r(z)) &= d_{T_W}(\alpha_1\alpha_{1,2}\dots\alpha_{l_\dagger}, \beta_1\beta_{1,2}\dots\beta_{l_\dagger}) \\
&\quad + \sum_{l=l_\dagger+1}^{L_X} \text{length}(\alpha_l) + \sum_{l=l_\dagger+1}^{L_Z} \text{length}(\beta_l) + L - 1 \\
&\leq s_r(Y_{l_*}) + 2(R+1) + \sum_{l=l_\dagger+1}^{L_X} (s_r(X_l) + R) + \sum_{l=l_\dagger+1}^{L_Z} (s_r(Z_l) + R) + L - 1 \\
&= s_r + 2(R+1) + (L-1)R + L - 1 \\
&= s_r + (R+1)(L+1)
\end{aligned}$$

Similarly,

$$\begin{aligned}
d_{T_W}(F_r(x), F_r(z)) &= d_{T_W}(\alpha_1\alpha_{1,2}\dots\alpha_{l_\dagger}, \beta_1\beta_{1,2}\dots\beta_{l_\dagger}) \\
&\quad + \sum_{l=l_\dagger+1}^{L_X} \text{length}(\alpha_l) + \sum_{l=l_\dagger+1}^{L_Z} \text{length}(\beta_l) + L - 1 \\
&\geq s_r(Y_{l_*}) + \sum_{l=l_\dagger+1}^{L_X} s_r(X_l) + \sum_{l=l_\dagger+1}^{L_Z} s_r(Z_l) + L - 1 \\
&= s_r + L - 1
\end{aligned}$$

and so we have proved the claim. □

We will now combine Claim 1, Claim 2 and Claim 6 above in order to prove that

$$d_{\mathcal{C}^{\mathcal{T}}_K(\mathbb{G})}(\iota_1(x), \iota_1(z)) \approx d(F(x), F(z))$$

Recalling that $L-3 \leq s$ (by the argument given in the proof of Claim 2), we have

$$\begin{aligned}
d(F(x),F(z)) &= \sum_{r=1}^{R} d_{T_W}(F_r(x),F_r(z)) \\
&\leq \sum_{r=1}^{R}(s_r+(R+1)(L+1)) && \text{(by Claim 6)} \\
&= s+R(R+1)(L+1) \\
&\leq s+R(R+1)(s+4) \\
&= (R^2+R+1)s+4R(R+1) \\
&\leq (R^2+R+1)\mathrm{length}(\zeta)+4R(R+1) && \text{(by Claim 2)} \\
&\leq (R^2+R+1)(\lambda d_{\mathcal{C}_K^{\mathcal{T}}(\mathbb{G})}(x,z)+\mu)+4R(R+1) && \text{(by Claim 1)}
\end{aligned}$$

Similarly,

$$\begin{aligned}
d_{\mathcal{C}_K^{\mathcal{T}}(\mathbb{G})}(x,z) &\leq \mathrm{length}(\zeta) && \text{(by Claim 1)} \\
&\leq \lambda' s+\mu' && \text{(by Claim 2)} \\
&= \lambda' \sum_{r=1}^{R} s_r+\mu' \\
&\leq \lambda' \sum_{r=1}^{R} d_{T_W}(F_r(x),F_r(z))+\mu' && \text{(by Claim 6)} \\
&= \lambda' d(F(x),F(z))+\mu'
\end{aligned}$$

and so we have proved the theorem. □

4.8 Proof of the main theorem part two

A reader may find it helpful at this point to read the proof in Appendix C that the hyperbolic plane quasiisometrically embeds into a product of two binary trees. Several of the ideas in the proof of Theorem 4.1 also appear in this simpler case, however it does nonetheless have its own quirks.

We continue with our proof of Theorem 4.1 and we use all the notation developed in Section 4.6 and Section 4.7.

Let $\phi : G \to \prod_{q=1}^{Q} T_{\{0,1\}}$ be the regular map given by Corollary 4.46. Let $\Phi : G \to \prod_{q=1}^{Q} T_C$ be the injective and Lipschitz map given by Lemma 4.13. Let us say that Φ is λ_5-Lipschitz.

Notation 4.72. A vertex of T_C is naturally identified with a word on the alphabet C. Hence, every vertex of $\prod_{q=1}^{Q} T_C$ naturally corresponds to a vector $(v_1, \ldots, v_Q)$ where v_q is a word on the alphabet C. Let $g \in G$. Then $\Phi(g) \in \prod_{q=1}^{Q} T_C$ corresponds to a vector

$$\begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_Q \end{pmatrix}$$

For $1 \leq q \leq Q$, we let $\Phi_q(g)$ denote the word $v_q \in T_C$.

Choosing our statistics

We would like to apply Corollary 3.35 to the sentence-tree T_W where W is the set of finite words on $S \cup S^{-1}$. In order to do so, we need to make choices for $\mathcal{S}_{\text{finite}}$, $\mathcal{S}_{\text{linear}}$, J_{finite}, J_{linear}, N and ϵ. We make the following choices.

Let $J_{\text{linear}} = 4\lambda_1 K + 2\mu_1 + 3$, $J_{\text{finite}} = \max(J_{\text{linear}}, R)$, $N = 12(R+1)\lambda_4(\lambda_3^2 + \lambda_3\mu_3 + \lambda_3 K + \mu_3)$ and let $\epsilon = 1$. Here, λ_1 and μ_1 are constants arising from Proposition 4.61, λ_3 and μ_3 are the constants arising from Proposition 4.65, and λ_4 is the constant arising from Theorem 4.70.

Let $\mathcal{S}_{\text{finite}}$ be the following collection of finite statistics on T_W (defined with respect to an arbitrary sentence $\overline{w_1}\,\overline{w_2}\ldots\overline{w_i} \in T_W$).

- is the path w_i entirely contained in one of the coset subgraphs $\mathcal{C}(H_1), \mathcal{C}(H_2), \ldots, \mathcal{C}(H_T)$?
- the final $I + 3K$ letters of w_i;
- length(w_i) modulo $(\lambda_1 K + \mu_1 + 1)$;
- if w_j is the last word of $\overline{w_1}\,\overline{w_2}\ldots\overline{w_i}$ that is not entirely contained in one of the subgraphs $\mathcal{C}(H_1), \ldots, \mathcal{C}(H_T)$, then what are the final χ letters of w_i?

So, in total, we have chosen four finite statistics. Note that a question like "is the path w_i entirely contained in one of the coset subgraphs $\mathcal{C}(H_1), \mathcal{C}(H_2), \ldots, \mathcal{C}(H_T)$?" is indeed a finite statistic as it has a yes or no answer. Recall that χ is the constant defined in (♮2).

Let ω be the smallest natural number that is greater than both

$$\lambda_5(J_{\text{linear}}(\lambda_4 + \mu_4) + K + \lambda_3 + \mu_3 + J_{\text{linear}}(\lambda_4 + \mu_4) + K)$$

and

$$\lambda_5(J_{\text{linear}}(\lambda_4 + \mu_4) + K + \lambda_3 + \mu_3 + \lambda_3(\lambda_3 + \mu_3 + K) + \mu_3 + K)$$

Let $\mathcal{S}_{\text{linear}}$ be the following collection of linear statistics on T_W, where we again let $\overline{w_1}\ \overline{w_2} \dots \overline{w_i}$ denote an arbitrary sentence in T_W. In the linear statistics below, note that c is the *variable* required in the definition of a linear statistic (see Definition 3.25) and not a constant like R, λ_4 and ω.

- the final $6(R+1)\lambda_4\omega c$ letters of w_{i-r+1} (for $1 \leq r \leq R$);
- the final $6(R+1)\lambda_4\omega c$ letters of $\mathrm{MT}(w_{i-r+1})$ (for $1 \leq r \leq R$);
- the final $6(R+1)\lambda_4\omega c$ letters of $\Phi_q(w_{i-j+1} \dots w_i)$ (for $1 \leq j \leq J_{\text{linear}}$ and $1 \leq q \leq Q$);
- the final $6(R+1)\lambda_4\omega c$ letters of $\mathrm{MT}(\Phi_q(w_{i-j+1} \dots w_i))$ (for $1 \leq j \leq J_{\text{linear}}$ and $1 \leq q \leq Q$).

So, in total, we have chosen $2R + 2J_{\text{linear}}Q$ linear statistics. To be clear: when we refer to $\Phi_q(w_{i-j+1} \dots w_i)$, we mean the image of $w_{i-j+1}w_{i-j+2} \dots w_i \in G$ under $\Phi_q : G \to T_C$.

The associated diary

Let $D = D(\mathcal{S}_{\text{finite}}, \mathcal{S}_{\text{linear}}, J_{\text{linear}}, N, \epsilon)$ be the diary given by Corollary 3.35. We claim that the composition

$$G \xrightarrow{F \times \phi} T_W \times T_W \times \cdots \times T_W \times \prod_{q=1}^{Q} T_{\{0,1\}} \xrightarrow{D \times D \times \cdots \times D \times \mathrm{id}} T_\Omega \times T_\Omega \times \cdots \times T_\Omega \times \prod_{q=1}^{Q} T_{\{0,1\}}$$

is a quasiisometric embedding. Once this is shown, the proof of the main theorem (Theorem 4.1) is complete since T_Ω is quasiisometric to the rooted binary tree $T_{\{0,1\}}$.

Notation 4.73. Let us refer to the composition $(D \times D \times \cdots \times D \times \mathrm{id}) \circ (F \times \phi)$ as $\mathcal{F} : G \to T_\Omega \times T_\Omega \times \cdots \times T_\Omega \times \prod_{q=1}^{Q} T_{\{0,1\}}$.

Note that $\mathcal{F}$ is coarsely Lipschitz since $F \times \phi$ is a quasiisometric embedding and $D \times D \times \cdots \times D \times \mathrm{id}$ is 1-Lipschitz. It follows that we only need to prove the lower bound of the quasiisometric inequality for $\mathcal{F}$.

Reductions

Let $x, z \in G$. We may assume that

$$d_G(x, z) \geq 2\lambda_4\mu_4 + 12R(R+1)\lambda_4 J_{\text{linear}} \tag{4.7}$$

since we are only interested in the coarse geometry of G. Now,

- define $d := d_G(x, z)$;
- define $d' := d((F \times \phi)(x), (F \times \phi)(z))$;
- define $d'' := d(\mathcal{F}(x), \mathcal{F}(z))$;
- let d'_r be the distance between $(F \times \phi)(x)$ and $(F \times \phi)(z)$ in the r'th factor for $r = 1, 2, \ldots, R+1$;
- let d''_r be the distance between $\mathcal{F}(x)$ and $\mathcal{F}(z)$ in the r'th factor for $r = 1, 2, \ldots, R+1$.

In the above, by the $(R+1)$'st factor, we mean the entirety of $\prod_{q=1}^{Q} T_{\{0,1\}}$. So $d' = d'_1 + d'_2 + \cdots + d'_R + d'_{R+1}$ and $d'' = d''_1 + d''_2 + \cdots + d''_R + d''_{R+1}$. Further, note that $d'_{R+1} = d''_{R+1}$. Since $d \geq 2\lambda_4\mu_4$ by (4.7), and since $F \times \phi$ is a (λ_4, μ_4)-quasiisometric embedding, we have $d' \geq d/2\lambda_4$.

Claim. *We are done if $d'_{R+1} \geq d'/(R+1)$.*

Proof of claim. If $d'_{R+1} \geq d'/(R+1)$ then

$$d'' \geq d''_{R+1} = d'_{R+1} \geq d'/(R+1) \geq \frac{d}{2(R+1)\lambda_4}$$

In other words,

$$d(\mathcal{F}(x), \mathcal{F}(z)) \geq \frac{d}{2(R+1)\lambda_4}$$

Since, as mentioned above, $\mathcal{F}$ is coarsely Lipschitz, we are done. □

So we can assume that $d'_{R+1} \leq d'/(R+1)$. This implies that $d'_1 + d'_2 + \cdots + d'_R \geq Rd'/(R+1)$. Suppose $1 \leq \mathfrak{r} \leq R$ is such that $d'_{\mathfrak{r}} = \max(d'_1, d'_2, \ldots, d'_R)$. It follows that

$$d'_{\mathfrak{r}} \geq \frac{d'_1 + d'_2 + \cdots + d'_R}{R} \geq d'/(R+1) \geq \frac{d}{2(R+1)\lambda_4} \tag{4.8}$$

Let $\alpha \in T_W$ be the $\mathfrak{r}$'th factor of $(F \times \phi)(x)$ and let $\beta \in T_W$ be the $\mathfrak{r}$'th factor of $(F \times \phi)(z)$. We can write

$$\begin{aligned}
&\alpha = \overline{u_1} \ldots \overline{u_p}\ \overline{u_{p+1}} \ldots \overline{u_{p+m}} \\
&\beta = \overline{u_1} \ldots \overline{u_p}\ \overline{u'_{p+1}} \ldots \overline{u'_{p+n}} \\
&u_{p+1} \neq u'_{p+1}
\end{aligned}$$

So $d_{T_W}(\alpha, \beta) = m + n = d'_{\mathfrak{r}}$.

Claim. *We are done if $m \leq (m+n)/3$ or $n \leq (m+n)/3$.*

Proof of claim. If either of these hold, we have $|m-n| \geq d'_{\mathfrak{r}}/3$ and hence

$$d'' \geq d''_{\mathfrak{r}} \geq |m-n| \geq d'_{\mathfrak{r}}/3 \geq \frac{d}{6(R+1)\lambda_4}$$

where the second inequality follows from the fact that D is height-preserving. ☐

So we can assume that

$$m, n \geq (m+n)/3 = d'_{\mathfrak{r}}/3 \geq \frac{d}{6(R+1)\lambda_4} \tag{4.9}$$

In particular, it follows that

$$m \geq \max(R, J_{\text{linear}}) \quad \text{and} \quad n \geq \max(R, J_{\text{linear}}) \tag{4.10}$$

since $d \geq 6R(R+1)\lambda_4$ and $d \geq 6(R+1)\lambda_4 J_{\text{linear}}$ by (4.7).

Claim. *We are done if we manage to prove that $\alpha, \beta \in T_W$ satisfy either $\leo(\mathcal{S}_{finite}, J_{finite})$ or $\taurus(\mathcal{S}_{linear}, J_{linear}, N, \epsilon)$.*

Proof. If $\alpha, \beta \in T_W$ satisfy either $\leo(\mathcal{S}_{\text{finite}}, J_{\text{finite}})$ or $\taurus(\mathcal{S}_{\text{linear}}, J_{\text{linear}}, N, \epsilon)$ then Corollary 3.35 would imply that

$$d''_{\mathfrak{r}} = d_{T_\Omega}(D\alpha, D\beta) \geq d_{T_W}(\alpha, \beta)/M = d'_{\mathfrak{r}}/M$$

where $M = M(J_{\text{finite}}, J_{\text{linear}})$ is the constant given by Corollary 3.35. But then, using (4.8), we would have

$$d'' \geq \frac{d}{2(R+1)M\lambda_4}$$

and the proof of the theorem would be complete. ☐

So we have reduced the problem to proving that

$$\alpha = F_{\mathfrak{r}}(x) = \overline{u_1} \ldots \overline{u_p}\ \overline{u_{p+1}} \ldots \overline{u_{p+m}} \in T_W$$
$$\beta = F_{\mathfrak{r}}(z) = \overline{u_1} \ldots \overline{u_p}\ \overline{u'_{p+1}} \ldots \overline{u'_{p+n}} \in T_W$$

satisfy either $\leo(\mathcal{S}_{\text{finite}}, J_{\text{finite}})$ or $\taurus(\mathcal{S}_{\text{linear}}, J_{\text{linear}}, N, \epsilon)$. This will now be our only goal.

Recall the definition of $F_r : G \to T_W$ given in Definition 4.67. We have several cases depending on the forms of the words u_{p+1} and u'_{p+1}: they can be a single letter of the form $a_{t,\mathfrak{r}}^{\pm 1}$, they can be a word of the form $a_{t,r}^q$ for some $r \neq \mathfrak{r}$ and $q \neq 0$, they can have the form $\psi_{t,i}$ or they can correspond to a transverse quasigeodesic $\gamma(X, Y)$.

Simple cases

If length$(u_{p+1}) \leq I + 3K$ or length$(u'_{p+1}) \leq I + 3K$ then since $u_{p+1} \neq u'_{p+1}$ we know that the finite statistic "the final $I + 3K$ letters of w_i" in $\mathcal{S}_{\text{finite}}$ distinguishes u_{p+1} from u'_{p+1}. So α and β would satisfy $\leo(\mathcal{S}_{\text{finite}}, J_{\text{finite}})$ (with a choice of $j = 1$). Thus u_{p+1} and u'_{p+1} can't have the forms $a_{t,\mathfrak{r}}$ or $a_{t,\mathfrak{r}}^{-1}$ or $\psi_{t,i}$.

Recall ($\natural$1) which implies that the transverse quasigeodesics $\gamma(X, Y)$ cannot be entirely contained in a single coset subgraph $\mathcal{C}(X)$. Thus if u_{p+1} has the form $a_{t,r}^{q}$ (for $r \neq \mathfrak{r}$) and u'_{p+1} has the form $\gamma(X, Y)$ then the the finite statistic "is the path w_i entirely contained in one of the coset subgraphs $\mathcal{C}(H_1), \mathcal{C}(H_2), \ldots, \mathcal{C}(H_T)$?" distinguishes u_{p+1} from u'_{p+1}. It follows that α, β satisfy $\leo(\mathcal{S}_{\text{finite}}, J_{\text{finite}})$. Similarly, if u'_{p+1} has the form $a_{t,r}^{q}$ ($r \neq \mathfrak{r}$) and u_{p+1} has the form $\gamma(X, Y)$ then we are done.

Two cases remain:

(C1) $u_{p+1} = a_{t,r}^{q}$ and $u'_{p+1} = a_{t,r}^{q'}$ where $r \neq \mathfrak{r}$;

(C2) $u_{p+1} = \gamma(X, Y)$ and $u'_{p+1} = \gamma(X', Y')$ for some transverse quasigeodesics $\gamma(X, Y)$ and $\gamma(X', Y')$.

Some more notation

We need some more notation. Write $X = xH_1 \in \mathbb{G}$ and $Z = zH_1 \in \mathbb{G}$. Set $L_X = L[H_1, X]$ and $L_Z = L[H_1, Z]$ and $L = L[X, Z]$. Write

$$\mathbb{G}_K[H_1, X] = \{H_1 = X_1 < X_2 < \cdots < X_{L_X} = X\}$$

and

$$\mathbb{G}_K[H_1, Z] = \{H_1 = Z_1 < Z_2 < \cdots < Z_{L_Z} = Z\}$$

and

$$\mathbb{G}_K[X, Z] = \{X = Y_1 < Y_2 < \cdots < Y_L = Z\}$$

Let ζ denote the $\mathfrak{r}$'th ordered standard path from e to x and write ζ as the concatenation of paths

$$\zeta = \zeta_1 \; \gamma(X_1, X_2) \; \zeta_2 \; \gamma(X_2, X_3) \; \ldots \; \gamma(X_{L_X - 1}, X_{L_X}) \; \zeta_{L_X}$$

as described in Definition 4.64. Similarly, Let ζ' denote the $\mathfrak{r}$'th ordered standard path from e to z and write ζ' as the concatenation of paths

$$\zeta' = \zeta'_1 \; \gamma(Z_1, Z_2) \; \zeta'_2 \; \gamma(Z_2, Z_3) \; \ldots \; \gamma(Z_{L_Z - 1}, Z_{L_Z}) \; \zeta'_{L_Z}$$

Degenerate cases

Claim 1. *If $\mathbb{G}_K[H_1, X]$ and $\mathbb{G}_K[H_1, Z]$ never diverge, equivalently, if one is contained in the other, then we are done.*

Proof of Claim 1. If, say, $\mathbb{G}_K[H_1, X]$ is contained in $\mathbb{G}_K[H_1, Z]$, then that would imply that the ζ and ζ' diverge in the coset subgraph $\mathcal{C}(X)$. This implies, since $x \in X$, that u_{p+1} has the form $a_{t,\mathfrak{r}}^{\pm 1}$ or $\psi_{t,i}$ or $a_{t,r}^{q}$ (for some $r \neq \mathfrak{r}$). As described above, we are done unless u_{p+1} has the form $a_{t,r}^{q}$ ($r \neq \mathfrak{r}$). If that occurs, by considering the structure of $F_{\mathfrak{r}}(x)$, there must exist $2 \leq j \leq R$ such that u_{p+j} has the form $\gamma(X, Y)$ for some $Y \in \mathbb{G}$ (note that $m \geq R$ by (4.10)). But this can't happen because $x \in X$.

We are done for similar reasons if $\mathbb{G}_K[H, Z]$ is contained in $\mathbb{G}_K[H, X]$. $\square$

So we can assume there exists some $l_\dagger < \min(L_X, L_Z)$ for which $X_l = Z_l$ when $l \leq l_\dagger$ and $X_{l_\dagger+1} \neq Z_{l_\dagger+1}$. In other words $X_{l_\dagger} = Z_{l_\dagger}$ is the divergence point of the two paths $\mathbb{G}_K[H_1, X]$ and $\mathbb{G}_K[H_1, Z]$ in $\mathcal{P}_K(\mathbb{G})$.

Claim 2. *We may assume that one of the two following situations occur:*

1. $\zeta_{l_\dagger} \neq \zeta'_{l_\dagger}$;
2. $\zeta_{l_\dagger} = \zeta'_{l_\dagger}$ *and* $\gamma(X_{l_\dagger}, X_{l_\dagger+1}) \neq \gamma(Z_{l_\dagger}, Z_{l_\dagger+1})$.

Proof of Claim 2. If $\zeta_{l_\dagger} = \zeta'_{l_\dagger}$ and $\gamma(X_{l_\dagger}, X_{l_\dagger+1}) = \gamma(Z_{l_\dagger}, Z_{l_\dagger+1})$ then

$$\zeta_1\, \gamma(X_1, X_2)\, \zeta_2\, \gamma(X_2, X_3)\ \ldots \zeta_{l_\dagger}\gamma(X_{l_\dagger}, X_{l_\dagger+1}) = \zeta'_1\, \gamma(Z_1, Z_2)\, \zeta'_2\, \gamma(Z_2, Z_3)\ \ldots \zeta'_{l_\dagger}\gamma(Z_{l_\dagger}, Z_{l_\dagger+1})$$

and so $p(X_{l_\dagger+1}, X_{l_\dagger}) =_G p(Z_{l_\dagger+1}, Z_{l_\dagger})$. It follows that $X_{l_\dagger+1} \cap Z_{l_\dagger+1} \neq \emptyset$ and so they are cosets of different peripheral subgroups $H_1, H_2, \ldots H_T$. But then, due to ($\natural$2), it is possible to distinguish $\gamma(X_{l_\dagger}, X_{l_\dagger+1})$ from $\gamma(Z_{l_\dagger}, Z_{l_\dagger+1})$ just by looking at the final χ letters. In particular, $\gamma(X_{l_\dagger}, X_{l_\dagger+1}) \neq \gamma(Z_{l_\dagger}, Z_{l_\dagger+1})$ which is a contradiction. $\square$

So, since either (C1) or (C2) occurs, and since one of the two cases from Claim 2 occurs, we have reduced the problem to the following two possibilities:

(D1) $\zeta_{l_\dagger} \neq \zeta'_{l_\dagger}$ and $u_{p+1} = a_{t,r}^{q}$ and $u'_{p+1} = a_{t,r}^{q'}$ where $r \neq \mathfrak{r}$;

(D2) $\zeta_{l_\dagger} = \zeta'_{l_\dagger}$ and $u_{p+1} = \gamma(X_{l_\dagger}, X_{l_\dagger+1})$ and $u'_{p+1} = \gamma(Z_{l_\dagger}, Z_{l_\dagger+1})$ and $p(X_{l_\dagger}, X_{l_\dagger+1}) = p(Z_{l_\dagger}, Z_{l_\dagger+1})$.

Claim 3. *When (D1) holds, we have*

$$|q - q'| \leq \mathit{length}(\eta) \leq \lambda_1 d_{\mathcal{C}(X_{l_\dagger})}(p(X_{l_\dagger}, X_{l_\dagger+1}), p(Z_{l_\dagger}, Z_{l_\dagger+1})) + \mu_1$$

where η is a standard path from $p(X_{l_\dagger}, X_{l_\dagger+1})$ to $p(Z_{l_\dagger}, Z_{l_\dagger+1})$ in the coset $\mathcal{C}(X_{l_\dagger}) = \mathcal{C}(Z_{l_\dagger})$ in the sense of Definition 4.59.

Proof of Claim 3. If (D1) holds then $u_{p+1} = a^q_{t,r}$ and $u'_{p+1} = a^{q'}_{t,r}$ correspond to subpaths of $\zeta_{l_\dagger}$ and $\zeta'_{l_\dagger}$ respectively.

We know that $\zeta_{l_\dagger}$ travels from $p_0(X_{l_\dagger})$ to $p(X_{l_\dagger}, X_{l_\dagger+1})$ and has the form $\nu\psi$ where ν is a geodesic in $p_0(X_{l_\dagger})A_t \cong \mathbb{Z}^{R_t}$ from $p_0(X_{l_\dagger})$ to $p(X_{l_\dagger}, X_{l_\dagger+1})\psi^{-1}$ and ψ is one of the non-abelian paths $\psi_{t,i}$. Analogously, we can write $\zeta'_{l_\dagger} = \nu'\psi'$. Now, identifying the abelian coset $p_0(X_{l_\dagger})A_t = p_0(Z_{l_\dagger})A_t$ with $\mathbb{Z}^{R_t}$ by identifying $p_0(X_{l_\dagger}) = p_0(Z_{l_\dagger})$ with $(0, 0, \ldots, 0)$, we know that the r'th coordinate of $p(X_{l_\dagger}, X_{l_\dagger+1})\psi^{-1}$ is q and the r'th coordinate of $p(Z_{l_\dagger}, Z_{l_\dagger+1})(\psi')^{-1}$ is q'. Hence,

$$|q - q'| \leq d_{\mathbb{Z}^{R_t}}(p(X_{l_\dagger}, X_{l_\dagger+1})\psi^{-1}, p(Z_{l_\dagger}, Z_{l_\dagger+1})(\psi')^{-1})$$

But we also know that a standard path η from $p(X_{l_\dagger}, X_{l_\dagger+1})$ to $p(Z_{l_\dagger}, Z_{l_\dagger+1})$ in the coset $X_{l_\dagger} = Z_{l_\dagger}$ contains a geodesic in $p_0(X_{l_\dagger})A_t \cong \mathbb{Z}^{R_t}$ from $p(X_{l_\dagger}, X_{l_\dagger+1})\psi^{-1}$ to $p(Z_{l_\dagger}, Z_{l_\dagger+1})(\psi')^{-1}$. It follows that

$$|q - q'| \leq \text{length}(\eta) \leq \lambda_1 d_{\mathcal{C}(X_{l_\dagger})}(p(X_{l_\dagger}, X_{l_\dagger+1}), p(Z_{l_\dagger}, Z_{l_\dagger+1})) + \mu_1$$

where the second inequality follows from Proposition 4.61. ☐

Claim 4. *We are done if $X_{l_\dagger} = Z_{l_\dagger}$ is an element of $\mathbb{G}_K[X, Z]$.*

Proof of Claim 4. Suppose this is true. It follows from Lemma 4.32 that

$$\mathbb{G}_K[Z, X] = \{Z_{L_Z} < \cdots < Z_{l_\dagger+1} < Z_{l_\dagger} = X_{l_\dagger} < X_{l_\dagger+1} < \cdots < X_{L_X}\} \tag{4.11}$$

Now, if (D1) occurs, then, by considering the structure of $F_\mathfrak{r}(x)$ and $F_\mathfrak{r}(z)$, there exists $2 \leq j \leq R$ such that $u_{p+j} = \gamma(X_{l_\dagger}, X_{l_\dagger+1})$ and there exists $2 \leq j' \leq R$ be such that $u'_{p+j'} = \gamma(Z_{l_\dagger}, Z_{l_{\dagger+1}})$. So if either (D1) or (D2) occurs, there exists $1 \leq j \leq R$ such that $u_{p+j} = \gamma(X_{l_\dagger}, X_{l_\dagger+1})$ and there exists $1 \leq j' \leq R$ be such that $u'_{p+j'} = \gamma(Z_{l_\dagger}, Z_{l_{\dagger+1}})$.

Suppose that $j < j'$. Then $u_{p+j} = \gamma(X_{l_\dagger}, X_{l_\dagger+1})$ yet u'_{p+j} has the form $a^q_{t,r}$ or $\psi_{t,i}$. Recalling that $J_{\text{finite}} \geq R$, it follows that α and β satisfy $\Omega(\mathcal{S}_{\text{finite}}, J_{\text{finite}})$. Similarly, we are done if $j' < j$. So we may assume that $j = j'$.

Now, we can see from (4.11) that $X_{l_\dagger} < X_{l_\dagger+1} < \cdots < X$ is a subpath of $\mathbb{G}_K[Z, X]$ and hence $u_{p+j+1} \ldots u_{p+m}$ is a subpath of the $\mathfrak{r}$'th ordered standard path from z to x. It follows that length$(u_{p+j+1} \ldots u_{p+m}) \leq \lambda_3 d + \mu_3$ by Proposition 4.65. Hence $\overline{u_{p+j+1}} \ldots \overline{u_{p+m}}$ has at most $(\lambda_3 + \mu_3)d$ letters. Recall from (4.9) that $m, n \geq \frac{d}{6(R+1)\lambda_4}$ and so $\overline{u_{p+j+1}} \ldots \overline{u_{p+m}}$

has at least $\frac{d}{6(R+1)\lambda_4} - R$ words. Now, thanks to (4.7), we have $d \geq 12R(R+1)\lambda_4$ and so $R \leq \frac{d}{12(R+1)\lambda_4}$. Therefore $\frac{d}{6(R+1)\lambda_4} - R \geq \frac{d}{12(R+1)\lambda_4}$. It follows that

$$\mathrm{AWL}(\overline{u_{p+j+1}}\dots\overline{u_{p+m}}) \leq \frac{(\lambda_3+\mu_3)d}{(d/12(R+1)\lambda_4)} \leq 12(R+1)\lambda_4(\lambda_3+\mu_3) \leq N$$

For similar reasons,

$$\mathrm{AWL}(\overline{u'_{p+j+1}}\dots\overline{u'_{p+n}}) \leq N$$

In other words, we have proved that $(\mho 1)$ holds for α and β. We now turn to proving that $(\mho 2)$ holds for α and β (with respect to the choices for $\mathcal{S}_{\mathrm{linear}}$, J_{linear}, N, ϵ given above).

Suppose first that (D1) holds, i.e. u_{p+1} has the form $a^{q}_{t,r}$ and u'_{p+1} has the form $a^{q'}_{t,r}$ where $r \neq \mathfrak{r}$. We want to bound $|q-q'|$ in terms of $d = d_G(x,z)$. Let ζ denote a standard path from x to z in G. Then the intersection of ζ with $\mathcal{C}(X_{l_\dagger}) = \mathcal{C}(Z_{l_\dagger})$, which we denote by η, is a standard path from $p(X_{l_\dagger}, X_{l_\dagger+1})$ to $p(Z_{l_\dagger}, Z_{l_\dagger+1})$ in the coset $X_{l_\dagger} = Z_{l_\dagger}$. Hence, using Claim 3 and Proposition 4.65, we have

$$|q-q'| \leq \mathrm{length}(\eta) \leq \mathrm{length}(\zeta) \leq \lambda_3 d + \mu_3 \leq \omega d \tag{4.12}$$

So we can deduce, from Corollary 4.11, that either the final $3\omega d$ letters of $a^{q}_{t,r}$ and $a^{q'}_{t,r}$ are distinct or the final $3\omega d$ letters of $\mathrm{MT}(a^{q}_{t,r})$ and $\mathrm{MT}(a^{q'}_{t,r})$ are distinct.

Now let $1 \leq j' \leq j$. Consider the linear statistic "the final $6(R+1)\lambda_4\omega c$ letters of $w_{i-j'+1}$". When applied to the sentences $\overline{u_1}\dots\overline{u_{p+j'}}$ and $\overline{u_1}\dots\overline{u'_{p+j'}}$ with $c = m+n$ this becomes "the final $6(R+1)\lambda_4\omega(m+n)$ letters of $a^{q}_{t,r}$" and "the final $6(R+1)\lambda_4\omega(m+n)$ letters of $a^{q'}_{t,r}$" respectively. Note that $6(R+1)\lambda_4\omega(m+n) \geq 3\omega d$ by (4.8). Hence α and β satisfy $\mho(\mathcal{S}_{\mathrm{linear}}, J_{\mathrm{linear}}, N, \epsilon)$ if the final $3\omega d$ letters of $a^{q}_{t,r}$ and $a^{q'}_{t,r}$ are distinct. For similar reasons, by considering the linear statistic "the final $6(R+1)\lambda_4\omega c$ letters of $\mathrm{MT}(w_{i-j'+1})$", we are done when the final $3\omega d$ letters of $\mathrm{MT}(a^{q}_{t,r})$ and $\mathrm{MT}(a^{q'}_{t,r})$ are distinct. So we are done when (D1) holds.

Now suppose that (D2) holds. In particular this implies that $p(X_{l_\dagger}, X_{l_\dagger+1}) = p(Z_{l_\dagger}, Z_{l_\dagger+1})$. By considering (4.11), we see that u_{p+1} is a subpath of a standard path from z to x and so u_{p+1} has length at most $\lambda_3 d + \mu_3 \leq \omega d$ by Proposition 4.65. Similarly $\mathrm{length}(u'_{p+1}) \leq \omega d$. Since $u_{p+1} \neq u'_{p+1}$, it follows that they are distinguished by the linear statistic "the final $6(R+1)\lambda_4\omega c$ letters of w_i" in $\mathcal{S}_{\mathrm{linear}}$ and so α and β satisfy $\mho(\mathcal{S}_{\mathrm{linear}}, J_{\mathrm{linear}}, N, \epsilon)$. $\square$

So we can assume that $X_{l_\dagger} = Z_{l_\dagger} \notin \mathbb{G}_K[X,Z]$.

Claim 5. *We are done when (D1) holds.*

Proof of Claim 5. Suppose (D1) holds. If $q > 0$ and $q' < 0$ then of course α, β satisfy $\Omega(\mathcal{S}_{\text{finite}}, J_{\text{finite}})$ as $\overline{u_1} \dots \overline{u_p}\ \overline{u_{p+1}}$ and $\overline{u_1} \dots \overline{u_p}\ \overline{u'_{p+1}}$ are distinguished by the finite statistic "the final $I + 3K$ letters of w_i". Similarly we are done if $q < 0$ and $q' > 0$. So $q, q' > 0$ or $q, q' < 0$ and so $\text{length}(u_{p+1}) \neq \text{length}(u'_{p+1})$.

Since $X_{l_\dagger} = Z_{l_\dagger} \notin \mathbb{G}_K[X, Z]$, we know that $d_{X_{l_\dagger}}(X, Z) \leq K$. Recall that $d_{X_{l_\dagger}}(X, Z)$ is by definition the diameter of $\pi_{X_{l_\dagger}}(X) \cup \pi_{X_{l_\dagger}}(Z)$ in $\mathcal{C}(X_{l_\dagger})$. Hence, using Claim 3 we have

$$|q - q'| \leq \lambda_1 d_{\mathcal{C}(X_{l_\dagger})}(p(X_{l_\dagger}, X_{l_\dagger+1}), p(Z_{l_\dagger}, Z_{l_\dagger+1})) + \mu_1 \leq \lambda_1 d_{X_{l_\dagger}}(X, Z) + \mu_1 \leq \lambda_1 K + \mu_1$$

So the sentences $\overline{u_1} \dots \overline{u_p}\ \overline{u_{p+1}}$ and $\overline{u_1} \dots \overline{u_p}\ \overline{u'_{p+1}}$ are distinguished by the finite statistic "$\text{length}(w_i)$ modulo $(\lambda_1 K + \mu_1 + 1)$". □

By Claim 5 we can assume that (D2) holds.

Claim 6. *We may assume that for all $1 \leq k \leq J_{\mathit{finite}}$ the group elements $u_1 \dots u_p u_{p+1} \dots u_{p+k}$ and $u_1 \dots u_p u'_{p+1} \dots u'_{p+k}$ are distinct (as elements of G).*

Proof of Claim 6. Suppose $1 \leq k \leq J_{\text{finite}}$ is such that $u_1 \dots u_p u_{p+1} \dots u_{p+k} =_G u_1 \dots u_p u'_{p+1} \dots u'_{p+k}$.

Firstly, note that $u_1 \dots u_p u_{p+1} \dots u_{p+k}$ is an element of the coset X_l for some $l \geq l_\dagger + 1$ and similarly $u_1 \dots u_p u'_{p+1} \dots u'_{p+k}$ is an element of the coset $Z_{l'}$ for some $l' \geq l_\dagger + 1$. We know that $X_l \neq Z_{l'}$ since $X_l \in \mathbb{G}_K[X_{l_\dagger+1}, X]$ and $Z_l \in \mathbb{G}_K[Z_{l_\dagger+1}, Z]$.

Due to ($\natural$1), we know that the last word of $\overline{u_1} \dots \overline{u_p}\ \overline{u_{p+1}} \dots \overline{u_{p+k}}$ not contained in one of the cosets $\mathcal{C}(H_t)$ has the form $\gamma(X_{l-1}, X_l)$. Similarly, the last word of $\overline{u_1} \dots \overline{u_p}\ \overline{u'_{p+1}} \dots \overline{u'_{p+k}}$ not contained in one of the cosets $\mathcal{C}(H_t)$ has the form $\gamma(Z_{l'-1}, Z'_l)$.

But since $u_1 \dots u_p u_{p+1} \dots u_{p+k} =_G u_1 \dots u_p u'_{p+1} \dots u'_{p+k}$ we know that $X_l \cap Z_{l'} \neq \emptyset$ and so, since $X_l \neq Z_{l'}$, we deduce that X_l and $Z_{l'}$ are cosets of different peripheral subgroups $H_1, \dots, H_T$. It follows from ($\natural$2) that the final χ letters of $\gamma(X_{l-1}, X_l)$ are distinct from the final χ letters of $\gamma(Z_{l'-1}, Z'_l)$. By considering the fourth finite statistic in $\mathcal{S}_{\text{finite}}$, we deduce that α and β satisfy $\Omega(\mathcal{S}_{\text{finite}}, J_{\text{finite}})$. □

The non-degenerate case

We have now reached the heart of the proof.

Consider the path $\mathbb{G}_K[X, Z]$. By Lemma 4.25, we know that $\mathbb{G}_K[X, Z]$ diverges from $\mathbb{G}_K[X, H_1]$ at some vertex $X_{l_\dagger + l_x}$ where $1 \leq l_x \leq 3$ and that it joins $\mathbb{G}_K[H_1, Z]$ at a vertex $Z_{l_\dagger + l_z}$ for some $1 \leq l_z \leq 3$. See Figure 4.7 for a visual depiction of the situation at which we have arrived.

Suppose that $j_x \in \mathbb{N}$ is such that $u_{p+j_x} = \gamma(X_{l_\dagger + l_x - 1}, X_{l_\dagger + l_x})$. Suppose that $j_z \in \mathbb{N}$ is such that $u'_{p+j_z} = \gamma(Z_{l_\dagger + l_z - 1}, Z_{l_\dagger + l_z})$.

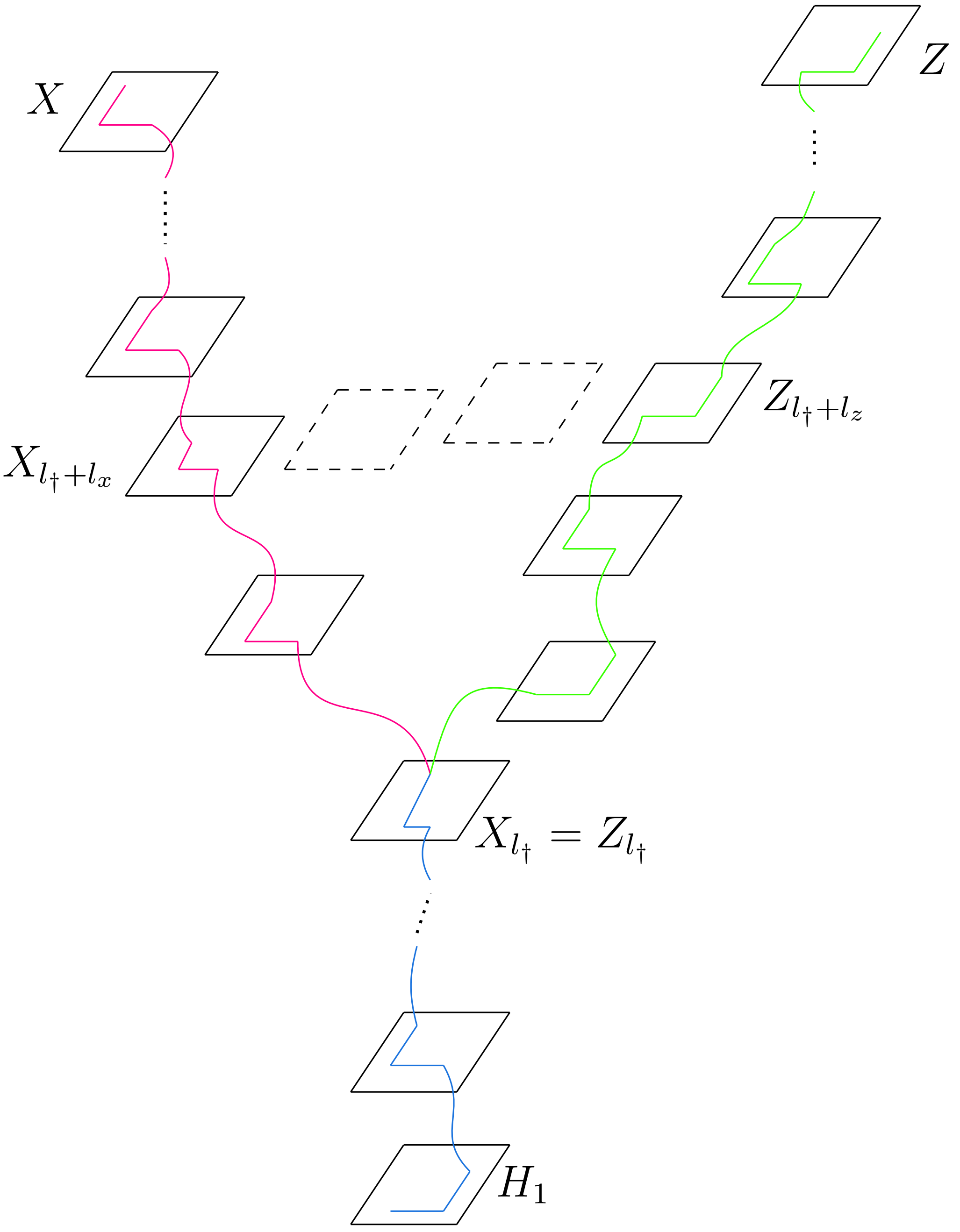

Figure 4.7: This diagram is drawn as if $H_t = A_t = \mathbb{Z}^2$ for all $1 \leq t \leq T$. It is also drawn for $l_x = 2$ and $l_z = 3$. The blue line indicates where the $\mathfrak{r}$'th ordered standard path from e to x (i.e. ζ) and the $\mathfrak{r}$'th ordered standard path from e to z (i.e. ζ') agree. The pink line is the rest of ζ from $p(X_{l_\dagger}, X_{l_\dagger+1}) = p(Z_{l_\dagger}, Z_{l_\dagger+1})$ to x. The green line is the rest of ζ' from $p(X_{l_\dagger}, X_{l_\dagger+1}) = p(Z_{l_\dagger}, Z_{l_\dagger+1})$ to x. The dashed squares are those which are in $\mathbb{G}_K[X, Z]$ but not in $\mathbb{G}_K[H_1, X]$ or $\mathbb{G}_H[H_1, Z]$. Made using TikZiT.

Claim 7. *We have $j_x \leq J_{linear}$ and $j_z \leq J_{linear}$.*

Proof of Claim 7. We begin by proving that $j_x \leq J_{\text{linear}}$. Now, $\overline{u_{p+1}}\,\overline{u_{p+2}}\ldots\overline{u_{p+j_x}}$ has the form

$$\overline{\gamma(X_{l_\dagger}, X_{l_\dagger+1})}$$

or the form

$$\overline{\gamma(X_{l_\dagger}, X_{l_\dagger+1})}\alpha_{l_\dagger+1}\overline{\gamma(X_{l_\dagger+1}, X_{l_\dagger+2})}$$

or the form

$$\overline{\gamma(X_{l_\dagger}, X_{l_\dagger+1})}\alpha_{l_\dagger+1}\overline{\gamma(X_{l_\dagger+1}, X_{l_\dagger+2})}\alpha_{l_\dagger+2}\overline{\gamma(X_{l_\dagger+2}, X_{l_\dagger+3})}$$

depending on whether $l_x = 1$, 2 or 3. In the above, we are using the notation from Definition 4.67.

We will prove that (in the second and third cases above) we have $\text{length}(\alpha_{l_\dagger+1}) \leq 2\lambda_1 K + \mu_1$. First, note that $\text{length}(\alpha_{l_\dagger+1}) \leq \text{length}(\zeta_{l_\dagger+1})$. But we have

$$\text{length}(\zeta_{l_\dagger+1}) \leq \lambda_1 d_{\mathcal{C}(X_{l_\dagger+1})}(p(X_{l_\dagger+1}, X_{l_\dagger}), p(X_{l_\dagger+1}, X_{l_\dagger+2})) + \mu_1 \leq \lambda_1 d_{X_{l_\dagger+1}}(H_1, X) + \mu_1$$

where the first inequality follows from Proposition 4.61. However, and this is really the crucial observation, since $X_{l_\dagger+1} \notin \mathbb{G}_K[H_1, Z]$ and $X_{l_\dagger+1} \notin \mathbb{G}_K[X, Z]$ we have

$$d_{X_{l_\dagger+1}}(H_1, X) \leq d_{X_{l_\dagger+1}}(H_1, Z) + d_{X_{l_\dagger+1}}(Z, X) \leq 2K$$

and so

$$\text{length}(\alpha_{l_\dagger+1}) \leq \text{length}(\zeta_{l_\dagger+1}) \leq 2\lambda_1 K + \mu_1$$

For identical reasons, in the third case above, we have $\text{length}(\alpha_{l_\dagger+2}) \leq 2\lambda_1 K + \mu_1$. It follows that, in all three cases, we have $j_x \leq 2(2\lambda_1 K + \mu_1) + 3 = J_{\text{linear}}$.

We have $j_z \leq J_{\text{linear}}$ for analogous reasons. □

We define $j = \max(j_x, j_z)$. Without loss of generality we can assume that in fact $j = j_x$. Let p_x denote the group element $u_1 \ldots u_p u_{p+1} \ldots u_{p+j_x} = p_0(X_{l_\dagger+l_x}) = p(X_{l_\dagger+l_x}, X_{l_\dagger+l_x-1})$ and let p_z denote the group element $u_1 \ldots u_p u'_{p+1} \ldots u'_{p+j_z} = p_0(Z_{l_\dagger+l_z}) = p(Z_{l_\dagger+l_z}, Z_{l_\dagger+l_z-1})$. Then $u_{p+j_x+1}u_{p+j_x+2}\ldots u_{p+m}$ is a path in $\Gamma(G, S)$ from p_x to x and $u'_{p+j_z+1}u'_{p+j_z+2}\ldots u'_{p+n}$ is a path in $\Gamma(G, S)$ from p_z to z.

Recall that $\mathbb{G}_K[X, Z] = \{X = Y_1 < Y_2 < \cdots < Y_L = Z\}$. Let $1 \leq l \leq L$ be such that $Y_l = X_{l_\dagger+l_x}$. Let $p = p(Y_l, Y_{l+1}) \in X_{l_\dagger+l_x}$. So $p \in \pi_{X_{l_\dagger+l_x}}(Z)$. Now, since $X_{l_\dagger+l_x} \notin \mathbb{G}_K[H_1, Z]$, we have $d_{X_{l_\dagger+l_x}}(H_1, Z) \leq K$ and so

$$d_G(p_x, p) \leq K \tag{4.13}$$

Similarly, let $1 \leq l' \leq L$ be such that $Y_{l'} = Z_{l_\dagger + l_z}$ and let p' be the vertex $p' = p(Y_{l'}, Y_{l'-1})$. So $p' \in \pi_{Z_{l_\dagger + l_z}}(X)$. Since $Z_{l_\dagger + l_z} \notin \mathbb{G}_K[H_1, X]$, we have $d_{Z_{l_\dagger + l_z}}(H_1, X) \leq K$ and so

$$d_G(p_z, p') \leq K \tag{4.14}$$

Claim 8. *$AWL(\overline{u_{p+j+1}}\ \overline{u_{p+j+2}} \dots \overline{u_{p+m}})$ and $AWL(\overline{u'_{p+j+1}}\ \overline{u'_{p+j+2}} \dots \overline{u'_{p+n}})$ are both at most N.*

Proof of Claim 8. We begin by counting the number of letters in $u_{p+j_x+1} u_{p+j_x+2} \dots u_{p+m}$. Now, $u_{p+j_x+1} u_{p+j_x+2} \dots u_{p+m}$ is a subpath of the $\mathfrak{r}$'th ordered standard path from e to x. By Proposition 4.65, this is (λ_3, μ_3)-quasigeodesic and hence

$$\text{length}(u_{p+j_x+1} u_{p+j_x+2} \dots u_{p+m}) \leq \lambda_3 d_G(p_x, x) + \mu_3 \leq \lambda_3 (d_G(p, x) + K) + \mu_3$$

where we have used (4.13). Now, consider a standard path in G from z to x. By definition, this standard path must go through the coset $X_{l_\dagger + l_x}$ since $X_{l_\dagger + l_x} \in \mathbb{G}_K[Z, X]$ and also through the point $p \in X_{l_\dagger + l_x}$. By Proposition 4.65, we know that this standard path has length at most $\lambda_3 d + \mu_3$. Hence,

$$d_G(p, x) \leq \lambda_3 d + \mu_3$$

So we have

$$\text{length}(u_{p+j_x+1} u_{p+j_x+2} \dots u_{p+m}) \leq \lambda_3^2 d + \lambda_3 \mu_3 + \lambda_3 K + \mu_3$$

For similar reasons, we have $\text{length}(u'_{p+j_z+1} u'_{p+j_z+2} \dots u'_{p+n}) \leq \lambda_3^2 d + \lambda_3 \mu_3 + \lambda_3 K + \mu_3$.

Since $j = j_x$, we therefore have

$$\text{length}(u_{p+j+1} u_{p+j+2} \dots u_{p+m}) \leq \lambda_3^2 d + \lambda_3 \mu_3 + \lambda_3 K + \mu_3$$

Further, we know that $u'_{p+j+1} u'_{p+j+2} \dots u'_{p+n}$ is a subword of $u'_{p+j_z+1} u'_{p+j_z+2} \dots u'_{p+n}$ and so

$$\text{length}(u'_{p+j+1} u'_{p+j+2} \dots u'_{p+n}) \leq \lambda_3^2 d + \lambda_3 \mu_3 + \lambda_3 K + \mu_3$$

By (4.7) we have $d \geq 12(R+1)\lambda_4 J_{\text{linear}}$ and hence $J_{\text{linear}} \leq \frac{d}{12(R+1)\lambda_4}$. Therefore, using Claim 7 and (4.9), we have

$$m - j \geq m - J_{\text{linear}} \geq \frac{d}{6(R+1)\lambda_4} - J_{\text{linear}} \geq \frac{d}{12(R+1)\lambda_4}$$

Therefore

$$\begin{aligned}\mathrm{AWL}(\overline{u_{p+j+1}}\;\overline{u_{p+j+2}}\dots\overline{u_{p+m}})) &\leq \frac{\lambda_3^2 d + \lambda_3\mu_3 + \lambda_3 K + \mu_3}{d/12(R+1)\lambda_4} \\ &\leq 12(R+1)\lambda_4(\lambda_3^2 + \lambda_3\mu_3 + \lambda_3 K + \mu_3) \\ &= N\end{aligned}$$

For similar reasons, we also have

$$\mathrm{AWL}(\overline{u'_{p+j+1}}\;\overline{u'_{p+j+2}}\dots\overline{u'_{p+n}})) \leq N$$

and we have proved the claim. □

So (♉1) of ♉$(\mathcal{S}_{\mathrm{linear}}, J, N, \epsilon)$ holds. We now turn to proving (♉2). With this in mind, and recalling that $\epsilon = 1$, let us suppose that $1 \leq j' \leq j$ is such that $\mathrm{length}(u_{p+j''}) \leq m+n$ and $\mathrm{length}(u'_{p+j''}) \leq m+n$ for all $j' < j'' \leq j$.

Claim 9. *The distance between* $\Phi(u_{p+1}\dots u_{p+j'})$ *and* $\Phi(u'_{p+1}\dots u'_{p+j'})$ *in* $\prod_{q=1}^{Q} T_C$ *is at most* ωd.

Proof of Claim 9. Since $\mathrm{length}(u_{p+j''}) \leq m+n$ for all $j' < j'' \leq j$, we have

$$\begin{aligned}d_G(u_1\dots u_p u_{p+1}\dots u_{p+j'}, p_x) &= d_G(u_1\dots u_p u_{p+1}\dots u_{p+j'}, u_1\dots u_p u_{p+1}\dots u_{p+j}) \\ &\leq J_{\mathrm{linear}}(m+n)\end{aligned}$$

So, using (4.13), we have

$$d_G(u_1\dots u_p u_{p+1}\dots u_{p+j'}, p) \leq J_{\mathrm{linear}}(m+n) + K$$

Recall that we have $m+n = d'_{\mathfrak{r}} \leq d' \leq \lambda_4 d + \mu_4$. Hence,

$$d_G(u_1\dots u_p u_{p+1}\dots u_{p+j'}, p) \leq J_{\mathrm{linear}}(\lambda_4 d + \mu_4) + K \tag{4.15}$$

Consider a standard path from x to z to $\Gamma(G,S)$. By Proposition 4.65, the length of this path is at most $\lambda_3 d + \mu_3$. Further, note that this standard path goes through the vertices p and p'. Therefore,

$$d_G(p,p') \leq \lambda_3 d + \mu_3 \tag{4.16}$$

We must now divide into two cases: we can have $j' \leq j_z$ or $j' > j_z$.

In the first case, since $\mathrm{length}(u'_{p+j''}) \leq m+n$ for all $j' < j'' \leq j$, and since $j_z \leq j$, we have

$$\begin{aligned}
d_G(u_1 \dots u_p u'_{p+1} \dots u'_{p+j'}, p_z) &= d_G(u_1 \dots u_p u'_{p+1} \dots u'_{p+j'}, u_1 \dots u_p u'_{p+1} \dots u'_{p+j_z}) \\
&\leq \sum_{j' < j'' \leq j_z} \mathrm{length}(u'_{p+j''}) \\
&\leq J_{\mathrm{linear}}(m+n)
\end{aligned}$$

and so, using (4.14), we have

$$d_G(u_1 \dots u_p u'_{p+1} \dots u'_{p+j'}, p') \leq J_{\mathrm{linear}}(m+n) + K \leq J_{\mathrm{linear}}(\lambda_4 d + \mu_4) + K \tag{4.17}$$

Therefore, combining (4.15), (4.16) and (4.17), and recalling that $\Phi : G \to \prod_{q=1}^{Q} T_C$ is λ_5-Lipschitz, we have

$$\begin{aligned}
&d(\Phi(u_{p+1} \dots u_{p+j'}), \Phi(u'_{p+1} \dots u'_{p+j'})) \leq \lambda_5 d_G(u_{p+1} \dots u_{p+j'}, u'_{p+1} \dots u'_{p+j'}) \\
&= \lambda_5 d_G(u_1 \dots u_p u_{p+1} \dots u_{p+j'}, u_1 \dots u_p u'_{p+1} \dots u'_{p+j'})) \\
&\leq \lambda_5 (J_{\mathrm{linear}}(\lambda_4 d + \mu_4) + K + \lambda_3 d + \mu_3 + J_{\mathrm{linear}}(\lambda_4 d + \mu_4) + K) \\
&\leq \omega d
\end{aligned}$$

as desired.

In the second case, when $j' > j_z$, we know that $u'_{p+j_z+1} u'_{p+j_z+2} \dots u'_{p+n}$ is a subpath of the $\mathfrak{r}$'th ordered standard path from e to z. By Proposition 4.65, this is a (λ_3, μ_3)-quasigeodesic, hence

$$\begin{aligned}
\mathrm{length}(u'_{p+j_z+1} u'_{p+j_z+2} \dots u'_{p+j'}) &\leq \mathrm{length}(u'_{p+j_z+1} u'_{p+j_z+2} \dots u'_{p+n}) \\
&\leq \lambda_3 d_G(p_z, z) + \mu_3 \\
&\leq \lambda_3 (d_G(p', z) + K) + \mu_3 \\
&\leq \lambda_3 (\lambda_3 d + \mu_3 + K) + \mu_3
\end{aligned}$$

where we have used (4.14) and the fact that a standard path in $\Gamma(G, S)$ from x to z must pass through p'. Therefore

$$\begin{aligned}
d_G(u_1 \dots u_p u'_{p+1} \dots u'_{p+j'}, p') &\leq d_G(u_1 \dots u_p u'_{p+1} \dots u'_{p+j'}, p_z) + K \\
&\leq \lambda_3 (\lambda_3 d + \mu_3 + K) + \mu_3 + K \qquad (4.18)
\end{aligned}$$

where we have again used (4.14). Hence, combining (4.15), (4.16) and (4.18), and using the fact that Φ is λ_5-Lipschitz, we have

$$\begin{aligned}
&d(\Phi(u_{p+1}\dots u_{p+j'}), \Phi(u'_{p+1}\dots u'_{p+j'})) \leq \lambda_5 d_G(u_{p+1}\dots u_{p+j'}, u'_{p+1}\dots u'_{p+j'}) \\
&= \lambda_5 d_G(u_1\dots u_p u_{p+1}\dots u_{p+j'}, u_1\dots u_p u'_{p+1}\dots u'_{p+j'})) \\
&\leq \lambda_5(J_{\text{linear}}(\lambda_4+\mu_4) + K + \lambda_3 d + \mu_3 + \lambda_3(\lambda_3 d + \mu_3 + K) + \mu_3 + K) \\
&\leq \omega d
\end{aligned}$$

and we are done. □

Since $j' \leq J_{\text{finite}}$, it follows from Claim 6 that $u_{p+1}\dots u_{p+j'} \neq u'_{p+1}\dots u'_{p+j'}$. Then, since Φ is injective by Lemma 4.13, we have $\Phi(u_{p+1}\dots u_{p+j'}) \neq \Phi(u'_{p+1}\dots u'_{p+j'})$, and so there exists some $1 \leq q \leq Q$ such that $\Phi_q(u_{p+1}\dots u_{p+j'}) \neq \Phi_q(u'_{p+1}\dots u'_{p+j'})$. For simplicity of notation, let us write $v_q = \Phi_q(u_{p+1}\dots u_{p+j'})$ and $v'_q = \Phi_q(u'_{p+1}\dots u'_{p+j'})$. We have $d_{T_C}(v_q, v'_q) \leq \omega d$ by Claim 9. So we deduce from Corollary 4.11 that either the final $3\omega d$ letters of v_q and v'_q are distinct or the final $3\omega d$ letters of $\mathrm{MT}(v_q)$ and $\mathrm{MT}(v'_q)$ are distinct.

Consider the linear statistic "the final $6(R+1)\lambda_4\omega c$ letters of $\Phi_q(w_{i-j'+1}\dots w_i)$". When applied to the sentences $\overline{u_1}\dots\overline{u_{p+j'}}$ and $\overline{u_1}\dots\overline{u'_{p+j'}}$ with $c = m+n$ this becomes "the final $6(R+1)\lambda_4\omega(m+n)$ letters of v_q" and "the final $6(R+1)\lambda_4\omega(m+n)$ letters of v'_q" respectively. Note that $6(R+1)\lambda_4\omega(m+n) \geq 3\omega d$ by (4.8). Hence α and β satisfy $\mho(\mathcal{S}_{\text{linear}}, J_{\text{linear}}, N, \epsilon)$ if the final $3\omega d$ letters of v_q and v'_q are distinct. For similar reasons, by considering the linear statistic "the final $6(R+1)\lambda_4\omega c$ letters of $\mathrm{MT}(\Phi_q(w_{i-j'+1}\dots w_i))$", we are done when the final $3\omega d$ letters of $\mathrm{MT}(v_q)$ and $\mathrm{MT}(v'_q)$ are distinct. So α and β satisfy $\mho(\mathcal{S}_{\text{linear}}, J_{\text{linear}}, N, \epsilon)$. □

4.9 Discussion and questions

It would be interesting to know whether Theorem 4.1 is truly distinct from Theorem 4.2, i.e. whether there exists a relatively hyperbolic group which quasiisometrically embeds into a finite product of trees but does not quasiisometrically embed into a finite product of binary trees.

Question 4.74. *Does there exist a finitely generated relatively hyperbolic group which quasiisometrically embeds into a finite product of trees but which does not quasiisometrically embed into a finite product of binary trees?*

Question 4.74 is highly related to the following question.

Question 4.75. *Does there exist a finitely generated group which quasiisometrically embeds into a finite product of trees but which does not quasiisometrically embed into a finite product of binary trees?*

Indeed, they are equivalent.

Proposition 4.76. *Question 4.74 and Question 4.75 are equivalent.*

Proof. Obviously an affirmative answer to Question 4.74 would give an affirmative answer to Question 4.75. For the reverse implication, suppose we have a finitely generated group H which answers Question 4.75 in the affirmative. Then $G = H * \mathbb{Z}$ is relatively hyperbolic with respect to H [Osi06a, Corollary 1.5]. G quasiisometrically embeds into a finite product of trees by Theorem 4.2, but G does not admit a quasiisometric embedding into a finite product of binary trees since H is undistorted in G by Proposition 4.47. □

If a group G satisfying Question 4.75 was found, then Theorem 4.1 would obstruct G from quasiisometrically embedding into a toral relatively hyperbolic group (whereas Theorem 4.2 alone would fail to provide such an obstruction).

Le Coz [Coz21] proves, using the coarse invariant of *separation profiles*, (introduced by Benjamini, Schramm and Timár [BST12]) that there exist finitely generated groups which have asymptotic dimension one and yet which do not coarsely embed into any finite product of binary trees. It follows that these groups do not quasiisometrically embed into finite products of binary trees (since any quasiisometric embedding is in particular also a coarse embedding) but they do coarsely embed into a finite product of trees (since this condition is equivalent to having finite asymptotic dimension, see [Kas22] and [Dra03]). So one naturally asks whether one can prove that the groups provided by Le Coz also quasiisometrically embed into finite products of trees.

Question 4.77. *Do the groups described by Le Coz admit a quasiisometric embedding into a finite product of trees?*

I spent a fair amount of time attempting to apply the methods of Chapter 3 and the proof structure of Theorem 4.1 to the case of mapping class groups. The hope would be to prove that a mapping class group quasiisometrically embeds into a product of *binary* trees, as opposed to the infinite-valence trees described in [Hum17] and [BBF21]. One might even hope to use this to imitate [Pet22] and prove that the mapping class group is quasiisometric to a uniformly locally finite CAT(0) cube complex. The attempted proof structure was as follows. By applying the results of [BBF21], one can find a quasiisometric embedding of the mapping class group into a product of quasi-trees of pseudo-Anosov axes. Thanks to Theorem 4.36, this can be upgraded to a product of trees of pseudo-Anosov axes. A tree of axes is itself a tree and so we would have a quasiisometric embedding into a product of trees. If the directed edges of these trees are labelled thoughtfully, then, by associating a vertex of each tree with the label of the unique geodesic from a basepoint to that vertex, we have defined an isometric embedding of the tree into a sentence-tree. So we have a quasiisometric embedding of the mapping class group into a product of sentence-trees. Surely we can now apply the theory of Chapter 3 to prove that this can be upgraded to a product of binary trees? However, and this is the major problem, we have no equivalent of Corollary 4.46 for the mapping class group. This regular map of the relatively hyperbolic group G into a product of binary trees was used in an essential way as a linear statistic in the proof of Theorem 4.1. So I ask the following question.

Question 4.78. *Does there exist a regular map of the mapping class group (of, say, some compact hyperbolic surface) into a finite product of binary trees?*

Mackay and Sisto apply Theorem 4.2 in order to fully characterise when a 3-manifold group quasiisometrically embeds into a product of trees (see [MS13, Theorem 5.1]). Since Theorem 4.1 covers all compact hyperbolic manifolds, it takes us closer to providing a similar characterisation for quasiisometric embeddings of 3-manifold groups into product of binary trees.

Question 4.79. *Let $G = \pi_1(M)$, where M is a compact orientable 3-manifold whose (possibly empty) boundary is a union of tori. Is it true that G quasiisometrically embeds into a finite product of binary trees if and only if no manifold in the prime decomposition of M has Nil geometry?*

To prove Question 4.79 in the affirmative, by following the proof structure of [MS13, Theorem 5.1] I believe two steps remain: firstly, to prove that a graph manifold quasiisometrically embeds into a finite product of binary trees (see [HS13] for the proof that a graph manifold quasiisometrically embeds into a finite product of infinite-valence trees) and secondly, to prove that a group which is relatively hyperbolic with respect to graph manifold groups and virtually abelian groups admits a quasiisometric embedding into a finite product of binary trees. So, as a first step, I ask the following.

Question 4.80. *Can you quasiisometrically embed graph manifold groups into finite products of binary trees?*

Appendix A

A formal definition of Alice's Diary

Definition A.1. Let $\kappa \in \mathbb{N}$ and let Ω_κ be the set of all words on A of length at most κ. We will now define the *Alice's Diary* map $\mathrm{AD}_\kappa : T_W \to T_{\Omega_\kappa}$. Let

$$\alpha = \overline{w_1}\ \overline{w_2} \dots \overline{w_i}$$

be a vertex of T_W. The image of α under AD_κ will be a sentence

$$\mathrm{AD}_\kappa(\alpha) = \overline{v_1}\ \overline{v_2} \dots \overline{v_i}$$

We will define the v_j for $1 \leq j \leq i$ inductively. Let l_0 be the empty word. Now, let $r_1 = l_0 w_1$, let v_1 be the first κ letters of $\overleftarrow{r_1}$ and let l_1 be the (possibly empty) string of letters at the start of r_1 that is leftover. Now let $r_2 = l_1 w_2$, let v_2 be the first κ letters of $\overleftarrow{r_2}$ and let l_2 be the (possibly empty) string of letters at the start of r_2 that is leftover. Now let $r_3 = l_2 w_3$, let v_3 be the first κ letters of $\overleftarrow{r_3}$ and let l_3 be the (possibly empty) string of letters at the start of r_3 that is leftover. Continuing in this way, we will have defined $v_1, \dots, v_i$ and hence AD_κ.

Appendix B

Two quasiisometric embeddings of relatively hyperbolic groups

Corollary B.1. *For sufficiently large K, the map $G \to \mathcal{C}_K(\mathbb{G}) \times \hat{G}$ given by*

$$g \mapsto (\iota_1(g), g)$$

is a quasiisometric embedding.

Before we can prove the corollary, we need the following technical lemma.

Lemma B.2. *Suppose that $A_{\mathbb{Y}}$ and $B_{\mathbb{Y}}$ are collections of numbers indexed by a set $\mathbb{Y}$. Suppose also that $A_Y \approx_{\lambda,\mu} B_Y$ for all $Y \in \mathbb{Y}$. Let $L \geq 2\lambda\mu$. Then*

$$\frac{1}{(\lambda+1)} \sum_{Y \in \mathbb{Y}} \{\{B_Y\}\}_{\lambda(L+\mu)} \leq \sum_{Y \in \mathbb{Y}} \{\{A_Y\}\}_L \leq (\lambda+1) \sum_{Y \in \mathbb{Y}} \{\{B_Y\}\}_{L/\lambda-\mu}$$

Proof. We prove the upper bound. We have that

$$\begin{aligned}\sum_{Y \in \mathbb{Y}} \{\{A_Y\}\}_L &= \sum_{A_Y > L} A_Y \leq \sum_{A_Y > L} (\lambda B_Y + \mu) \leq \sum_{B_Y > L/\lambda - \mu} (\lambda B_Y + \mu) \leq \sum_{B_Y > L/\lambda - \mu} (\lambda+1) B_Y \\ &= (\lambda+1) \sum_{Y \in \mathbb{Y}} \{\{B_Y\}\}_{L/\lambda-\mu}\end{aligned}$$

By symmetry, and the fact that $\lambda(L+\mu) \geq 2\lambda\mu$, we have that

$$\sum_{Y \in \mathbb{Y}} \{\{B_Y\}\}_{\lambda(L+\mu)} \leq (\lambda+1) \sum_{Y \in \mathbb{Y}} \{\{A_Y\}\}_L$$

and we are done. □

Proof of Corollary 4.55. Let $g, h \in G$. The distance formula for relatively hyperbolic groups tells us that for sufficiently large L we have

$$d_G(g,h) \approx \sum_{Y \in \mathbb{G}} \{\{d_G(\rho_Y(g), \rho_Y(h))\}\}_L + d_{\hat{G}}(g,h)$$

The distance formula for $\mathcal{C}_K(\mathbb{G})$ tells us that for sufficiently large K we have

$$d_{\mathcal{C}_K(\mathbb{G})}(\iota_1(g), \iota_1(h)) \approx \sum_{Y \in \mathbb{G}_K(\iota_1(g), \iota_1(h))} d_Y(\iota_1(g), \iota_1(h))$$

Suppose $Y \in \mathbb{G}$ and suppose first that $gH_1 \neq Y$ and $hH_1 \neq Y$. Write $Y = gH$ and let $F_Y : gH \to \mathcal{C}(Y) = g\Gamma(H, S \cap H)$ be the obvious map. Then

$$\begin{aligned}
d_G(\rho_Y(g), \rho_Y(h)) &\approx \operatorname{diam}_G(\rho_Y(gH_1) \cup \rho_Y(hH_1)) && \text{(by Theorem 4.51 and (P3))}\\
&\approx \operatorname{diam}_{\mathcal{C}(Y)}(F_Y \circ \rho_Y(gH_1) \cup F_Y \circ \rho_Y(hH_1)) && \text{(by Proposition 4.47)}\\
&\approx \operatorname{diam}_{\mathcal{C}(Y)}(\pi_Y(gH_1) \cup \pi_Y(hH_1)) && \text{(by Theorem 4.17)}\\
&= d_Y(\iota_1(g), \iota_1(h))
\end{aligned}$$

If instead $gH_1 = Y$ but $hH_1 \neq Y$, then

$$\begin{aligned}
d_G(\rho_Y(g), \rho_Y(h)) &= d_G(g, \rho_Y(h))\\
&\approx \operatorname{diam}_G(g, \rho_Y(hH_1))\\
&\approx \operatorname{diam}_{\mathcal{C}(Y)}(F_Y(g) \cup F_Y \circ \rho_Y(hH_1))\\
&\approx \operatorname{diam}_{\mathcal{C}(Y)}(F_Y(g) \cup \pi_Y(hH_1))\\
&= d_Y(\iota_1(g), \pi_Y(hH_1))
\end{aligned}$$

we are done for similar reasons when $gH_1 \neq Y$ and $hH_1 = Y$, and when $gH_1 = Y$ and $hH_1 = Y$.

Let A_Y denote the quantity $d_G(\rho_Y(g), \rho_Y(h))$ and let B_Y denote the quantity $d_Y(\iota_1(g), \iota_1(h))$. We have shown that there exist λ, μ such that $A_Y \approx_{\lambda,\mu} B_Y$ for all $Y \in \mathbb{G}$. Lemma B.2 then implies that (for sufficiently large L and $K = \lambda(L + \mu)$)

$$\sum_{Y \in \mathbb{G}} \{\{A_Y\}\}_L \geq \frac{1}{(\lambda+1)} \sum_{Y \in \mathbb{Y}} \{\{B_Y\}\}_K = \frac{1}{(\lambda+1)} \sum_{Y \in \mathbb{G}_K(\iota_1(g), \iota_1(h))} B_Y \approx d_{\mathcal{C}_K(\mathbb{G})}(\iota_1(g), \iota_1(h))$$

It also implies that for $L' = \lambda(K + \mu)$

$$\sum_{Y \in \mathbb{G}} \{\{A_Y\}\}_{L'} \leq (\lambda+1) \sum_{Y \in \mathbb{Y}} \{\{B_Y\}\}_K = (\lambda+1) \sum_{Y \in \mathbb{G}_K(\iota_1(g), \iota_1(h))} B_Y \approx d_{\mathcal{C}_K(\mathbb{G})}(\iota_1(g), \iota_1(h))$$

But Sisto's formula (Theorem 4.54) tells us that

$$\sum_{Y\in\mathbb{G}}\{\{A_Y\}\}_L \approx d_G(g,h) - d_{\hat{G}}(g,h) \approx \sum_{Y\in\mathbb{G}}\{\{A_Y\}\}_{L'}$$

and so $\sum_{Y\in\mathbb{G}}\{\{A_Y\}\}_L \approx d_{\mathcal{C}_K(\mathbb{G})}(\iota_1(g),\iota_1(h))$. Combining this with Theorem 4.54 gives us our desired approximation. □

Corollary B.3. *For sufficiently large K, the map $G \to \mathcal{C}_K(\mathbb{G}) \times X(G)$ given by*

$$g \mapsto (\iota_1(g), g)$$

is a quasiisometric embedding.

Proof. The inclusion map $G \to \hat{G}$ factors as the pair of 1-Lipschitz maps $G \to X(G) \to \hat{G}$ and so

$$d_{\hat{G}}(g,h) \leq d_{X(G)}(g,h) \leq d_G(g,h)$$

Hence,

$$\begin{aligned} d_{\hat{G}}(g,h) + d_{\mathcal{C}_K(\mathbb{G})}(\iota_1(g),\iota_1(h)) &\leq d_{X(G)}(g,h) + d_{\mathcal{C}_K(\mathbb{G})}(\iota_1(g),\iota_1(h)) \\ &\leq d_G(g,h) + d_{\mathcal{C}_K(\mathbb{G})}(\iota_1(g),\iota_1(h)) \\ &\leq d_G(g,h) + d_{\hat{G}}(g,h) + d_{\mathcal{C}_K(\mathbb{G})}(\iota_1(g),\iota_1(h)) \end{aligned}$$

Corollary 4.55 tells us that $d_{\hat{G}}(g,h) + d_{\mathcal{C}_K(\mathbb{G})}(\iota_1(g),\iota_1(h)) \approx d_G(g,h)$ and so $d_{X(G)}(g,h) + d_{\mathcal{C}_K(\mathbb{G})}(\iota_1(g),\iota_1(h)) \approx d_G(g,h)$. □

Appendix C

A quasiisometric embedding of $\mathbb{H}^2$ into a product of two binary trees

In this section we will prove that there is a quasiisometric embedding of the hyperbolic plane into a product of two binary trees. The proof roughly follows the outline of the proof of the same result given in [BDS07, Section 6], however I use the terminology of diaries and statistics in order to prove it. I hope this section helps clarify the usage of diaries and statistics in a simple case.

One can prove that the hyperbolic plane $\mathbb{H}^2$ is quasiisometric to the hexagonal hyperbolic Coxeter group

$$G = \langle a_1, a_2, a_3, b_1, b_2, b_3 \mid a_i^2 = b_i^2 = e \text{ for } i \in \{1, 2, 3\} \text{ and } [a_k, b_l] = e \text{ for } k \neq l \rangle$$

To see this, one considers the cocompact action of G on $\mathbb{H}^2$, where the generators of G correspond to reflections in the sides of a regular hexagon in $\mathbb{H}^2$. Write $S = \{a_1, a_2, a_3, b_1, b_2, b_3\}$. We consider the word metric on G arising from this generating set.

Let $A = S \cup S^{-1} = S$ and let W be the set of finite words on A. We write $\mathcal{A} = \{a_1, a_2, a_3\}$ and $\mathcal{B} = \{b_1, b_2, b_3\}$.

Definition C.1. Let $g \in G$. The *a-left representation* of g is the unique geodesic word $w \in W$ which represents g such that all the $\mathcal{A}$-letters in w have been commuted as far to the left as possible. For example, if $g = b_1a_2a_3b_2a_1b_1$ then the a-left representation of g is $w = a_2a_3b_1a_1b_2b_1$. The *b-left representation* of g is defined analogously.

We can map G into T_W via a map $F_{\mathcal{A}} : G \to T_W$ as follows. Let $g \in G$ and let $w \in W$ be the a-left representation of g. We may write

$$w = u_1a_1u_2a_2u_3a_3 \dots u_ma_mu_{m+1}$$

where $a_i \in \mathcal{A}$ and u_i is a word on the alphabet $\mathcal{B}$. Then we define

$$F_{\mathcal{A}}(g) = \overline{u_1a_1}\ \overline{u_2a_2}\ \overline{u_3a_3}\dots\overline{u_ma_m}$$

So, for example $F_{\mathcal{A}}(b_1a_2a_3b_2a_1b_1) = \overline{a_2}\ \overline{a_3}\ \overline{b_1a_1}$.

We define $F_{\mathcal{B}} : G \to T_W$ analogously. Let $F : G \to T_W \times T_W$ be $F = F_1 \times F_2$.

Lemma C.2. *Suppose $g, g' \in G$ and suppose they have a-left representations (respectively)*

$$w = u_1a_1u_2a_2\dots u_pa_pu_{p+1}a_{p+1}\dots u_{p+m}a_{p+m}u_{p+m+1}$$

and

$$w' = u_1a_1u_2a_2\dots u_pa_pu'_{p+1}a'_{p+1}\dots u'_{p+n}a'_{p+n}u'_{p+n+1}$$

where $u_{p+1}a_{p+1} \neq u'_{p+1}a'_{p+1}$. Then one can prove that the letters $a_{p+1}, a_{p+2}, \dots, a_{p+m}$ and $a'_{p+1}, a'_{p+2}, \dots, a'_{p+n}$ are exactly the letters in $\mathcal{A}$ which survive when you fully cancel the word $w^{-1}w'$. A similar statement can be said for the b-left representations.

Proof. This is is left to the reader. ☐

Corollary C.3. *$F : G \to T_W \times T_W$ is an isometric embedding.*

Proof. Let $g, g' \in G$. Lemma C.2 tells us that $F_{\mathcal{A}}$ counts the numbers of letters in $\mathcal{A}$ that are in a reduced word representing $g^{-1}g'$ and $F_{\mathcal{B}}$ counts the numbers of letters in $\mathcal{B}$ that are in a reduced word representing $g^{-1}g'$. ☐

Now, as discussed in Remark 3.36, we can combine (♌) and (♍), analogously to how (♌) and (♉) are combined in Corollary 3.35, to create a $D : T_W \to T_\Omega$ associated to the data

- $\mathcal{S}_{\text{finite}} = \{\text{the final letter of } w_i\}$;
- $J_{\text{finite}} = 2$;
- $\mathcal{S}_{\text{linear}} = \left\{\begin{matrix}\text{the final } 36c \text{ letters of } w_i \\ \text{the final } 36c \text{ letters of } \mathrm{MT}(w_i)\end{matrix}\right\}$;
- $\delta = 0$;
- $J_{\text{linear}} = 2$;
- $N = 18$;
- $\epsilon = 1$.

We imagine the single finite statistic and the two linear statistics as being applied to an arbitrary sentence $\overline{w_1}\ \overline{w_2}\dots\overline{w_i} \in T_W$. Let $M \geq 1$ be the associated constant so that $d_{T_\Omega}(D\alpha, D\beta)) \geq d_{T_W}(\alpha,\beta)/M$ for all α,β satisfying $\leo(\mathcal{S}_{\text{finite}}, J_{\text{finite}})$ or $\virgo(\mathcal{S}_{\text{linear}}, \delta, J_{\text{linear}}, N, \epsilon)$.

Theorem C.4. *The composition*

$$G \xrightarrow{F} T_W \times T_W \xrightarrow{D\times D} T_\Omega \times T_\Omega$$

is a quasiisometric embedding.

Proof. Since D is 1-Lipschitz, the composition is coarsely Lipschitz. So we only need to worry about the lower bound of the quasiisometric inequality.

Let $g, g' \in G$. Since we only care about the coarse geometry of G, we may assume that $d_G(g,g') \geq 12$. We may write $F(g) = (\alpha_1,\alpha_2)$ and $F(g') = (\beta_1,\beta_2)$ so that $d(F(g),F(g')) = d_{T_W}(\alpha_1,\beta_1) + d_{T_W}(\alpha_2,\beta_2)$. Without loss of generality, we may assume that $d_{T_W}(\alpha_1,\beta_1) \geq d_{T_W}(\alpha_2,\beta_2)$. So $d_{T_W}(\alpha_1,\beta_1) \geq \frac{1}{2}d(F(g),F(g'))$. We may write

$$\alpha_1 = \overline{u_1a_1}\ \overline{u_2a_2}\dots\overline{u_pa_p}\ \overline{u_{p+1}a_{p+1}}\dots\overline{u_{p+m}a_{p+m}}$$

and

$$\beta_1 = \overline{u_1a_1}\ \overline{u_2a_2}\dots\overline{u_pa_p}\ \overline{u'_{p+1}a'_{p+1}}\dots\overline{u'_{p+n}a'_{p+n}}$$

where $u_{p+1}a_{p+1} \neq u'_{p+1}a'_{p+1}$ and where the a_i and a'_i are letters in $\mathcal{A}$ and the u_i and u'_i are words on $\mathcal{B}$. Note that $d_{T_W}(\alpha_1,\beta_1) = m+n$.

Claim 1. *We may assume that $m,n \geq (m+n)/3$.*

Proof of Claim 1. Otherwise, we have $|m-n| \geq (m+n)/3$ and so

$$d_{T_\Omega}(D\alpha_1, D\beta_1) \geq (m+n)/3 = d_{T_W}(\alpha_1,\beta_1)/3 \geq \frac{1}{6}d(F(g),F(g')) = \frac{1}{6}d_G(g,g')$$

where the first inequality follows from the fact that D is height-preserving. So we are done in this case. ☐

It follows that

$$m,n \geq (m+n)/3 \geq \frac{1}{6}d(F(g),F(g')) = \frac{1}{6}d_G(g,g') \tag{C.1}$$

In particular, since $d_G(g,g') \geq 12$, we know that $m,n \geq 2$.

Claim 2. *We are done if the sentences $\alpha_1,\ \beta_1 \in T_W$ satisfy $\leo(\mathcal{S}_{\mathit{finite}}, J_{\mathit{finite}})$ or $\virgo(\mathcal{S}_{\mathit{linear}}, \delta, J_{\mathit{linear}}, N, \epsilon)$.*

Proof of Claim 2. If this occurs then we have

$$d_{T_\Omega}(D\alpha_1, D\beta_1) \geq d_{T_W}(\alpha_1, \beta_1)/M \geq \frac{1}{2M} d(F(g), F(g')) = \frac{1}{2M} d_G(g, g')$$

and we are done. ☐

If $a_{p+1} \neq a'_{p+1}$ or $a_{p+2} \neq a'_{p+2}$ then α_1 and β_1 satisfy $♌(\mathcal{S}_{\text{finite}}, J_{\text{finite}})$ and we are done. So we may assume that $a_{p+1} = a'_{p+1}$ and $a_{p+2} = a'_{p+2}$.

A crucial observation is that when we reduce (i.e. cancel as much as possible by commuting elements and deleting pairs of the form a_i^2 or b_i^2)

$$a_{p+m}^{-1} u_{p+m}^{-1} \dots a_{p+1}^{-1} u_{p+1}^{-1} u'_{p+1} a'_{p+1} \dots u'_{p+n} a'_{p+n} \tag{C.2}$$

we are left with a geodesic from g to g'. So any collection of letters in (C.2) which do not cancel have cardinality bounded above by $d_G(g, g')$.

Claim 3. *At most one $\mathcal{B}$-letter in $u_{p+3}a_{p+3} \dots u_{p+m}a_{p+m}$, and at most one $\mathcal{B}$-letter in $u'_{p+3}a'_{p+3} \dots u'_{p+n}a'_{p+n}$, can cancel when we reduce (C.2).*

Proof of Claim 3. This is left to the reader but the idea is that it is difficult for a $\mathcal{B}$-letter in $u_{p+3}a_{p+3} \dots u_{p+m}a_{p+m}$ to commute with both a_{p+1} and a_{p+2} (yet it needs to). ☐

Claim 4. *When we reduce (C.2), there can only be cancellation in at most one of the words u_{p+2} and u'_{p+2}.*

Proof of Claim 4. For there to be cancellation in u_{p+2}, we must have that u_{p+1} is a subword of $u'_{p+1}u'_{p+2}u'_{p+3} \dots u'_{p+n}$. If the final letter of u_{p+1} is within $u'_{p+2}u'_{p+3} \dots u'_{p+n}$, then, for there to be any cancellation within u_{p+2}, this final letter must commute with a'_{p+1}. But then the final letter of u_{p+1} would commute with $a_{p+1} = a'_{p+1}$ which contradicts the a-left representation. Thus u_{p+1} is actually a subword of u'_{p+1}. In exactly the same way for there to be any cancellation in u'_{p+2}, we must have that u'_{p+1} is a subword of u_{p+1}. Both these conditions cannot happen, and hence there can only be cancellation in at most one of u_{p+2} and u'_{p+2}. ☐

Claim 5. *We are done when $u_{p+2}a_{p+2} \neq u'_{p+2}a'_{p+2}$.*

Proof of Claim 5. We claim that $♍(\mathcal{S}_{\text{linear}}, \delta, J_{\text{linear}}, N, \epsilon)$ holds with $j = 2$ as the choice of index. Obviously (♍2) holds.

Let us prove that (♍1) holds. Since, by Claim 3, at most one $\mathcal{B}$-letter in $u_{p+3}a_{p+3} \dots u_{p+m}a_{p+m}$ can reduce when we cancel (C.2), it follows that

$$\text{length}(u_{p+3}a_{p+3} \dots u_{p+m}a_{p+m}) - 1 \leq d_G(g, g')$$

In particular, $\operatorname{length}(u_{p+3}a_{p+3}\dots u_{p+m}a_{p+m}) \le 2d_G(g,g')$. Hence, by (C.1), we have

$$\operatorname{AWL}(\overline{u_{p+3}a_{p+3}}\dots\overline{u_{p+m}a_{p+m}}) \le \frac{2d_G(g,g')}{d_G(g,g')/6} = 12 \le N$$

Similarly, we have $\operatorname{AWL}(\overline{u'_{p+3}a'_{p+3}}\dots\overline{u'_{p+n}a'_{p+n}}) \le N$.

Finally, we need to prove that ($\Virgo$3) holds. It follows from Claim 4 and Lemma C.2 that $\operatorname{length}(u_{p+2}a_{p+2}) \le d_G(g,g')$ or $\operatorname{length}(u'_{p+2}a'_{p+2}) \le d_G(g,g')$. Since $d_G(g,g') \le 2(m+n)$, it follows that the linear statistic "the final $36c$ letters of w_i" with $c = m+n$ can distinguish between the sentences $\overline{u_1a_1}\dots\overline{u_pa_p}\ \overline{u_{p+1}a_{p+1}}\ \overline{u_{p+2}a_{p+2}}$ and $\overline{u_1a_1}\dots\overline{u_pa_p}\ \overline{u'_{p+1}a'_{p+1}}\ \overline{u'_{p+2}a'_{p+2}}$. ☐

So we may assume that $u_{p+2}a_{p+2} = u'_{p+2}a'_{p+2}$. It follows from Claim 4 and Lemma C.2 that $\operatorname{length}(u_{p+2}a_{p+2}) \le d_G(g,g')$ and $\operatorname{length}(u'_{p+2}a'_{p+2}) \le d_G(g,g')$.

Claim 6. *We are done when* $u_{p+2}a_{p+2} = u'_{p+2}a'_{p+2}$.

Proof of Claim 6. We claim that $\Virgo(\mathcal{S}_{\text{linear}}, \delta, J_{\text{linear}}, N, \epsilon)$ holds with $j = 1$ as the choice of index. Obviously ($\Virgo$2) holds.

Let us prove that ($\Virgo$1) holds. As in the proof of Claim 5, we have

$$\operatorname{length}(u_{p+3}a_{p+3}\dots u_{p+m}a_{p+m}) \le 2d_G(g,g')$$

Further, we have $\operatorname{length}(u_{p+2}a_{p+2}) \le d_G(g,g')$ and $\operatorname{length}(u'_{p+2}a'_{p+2}) \le d_G(g,g')$. Hence, by (C.1), we have

$$\operatorname{AWL}(\overline{u_{p+2}a_{p+2}}\dots\overline{u_{p+m}a_{p+m}}) \le \frac{3d_G(g,g')}{d_G(g,g')/6} = 18 = N$$

Similarly, we have $\operatorname{AWL}(\overline{u'_{p+2}a'_{p+2}}\dots\overline{u'_{p+n}a'_{p+n}}) \le N$.

Finally, we need to prove that ($\Virgo$3) holds. Let u be the largest common initial part of $u_{p+1}a_{p+1}$ and $u'_{p+1}a'_{p+1}$ and write $u_{p+1}a_{p+1} = uv$ and $u'_{p+1}a'_{p+1} = uv'$.

We claim that $\operatorname{length}(v) \le 3d_G(g,g')$ and $\operatorname{length}(v') \le 3d_G(g,g')$. We have three cases: v, v' can both be non-empty, v can be non-empty and v' can be empty, or v' can be non-empty and v can be empty. In the first case, it is not hard to see that both v and v' survive in the cancellation of (C.2). So $\operatorname{length}(v) \le d_G(g,g')$ and $\operatorname{length}(v') \le d_G(g,g')$. In the second case, when we reduce (C.2), we know that v can only cancel with letters in $u'_{p+2}a'_{p+2}$ and $u'_{p+3}a'_{p+3}\dots u'_{p+n}a'_{p+n}$. So at most $d_G(g,g') + 1$ letters in v can cancel in (C.2). So $\operatorname{length}(v) \le 2d_G(g,g') + 1 \le 3d_G(g,g')$. We are done in the third case for similar reasons.

Hence $d_{T_A}(u_{p+1}a_{p+1}, u_{p+1}a_{p+1}) \leq 6d_G(g, g')$. It follows from Corollary 4.11 that either the final $18d_G(g, g')$ letters of $u_{p+1}a_{p+1}$ and $u'_{p+1}a'_{p+1}$ are distinct or the final $18d_G(g, g')$ letters of $\mathrm{MT}(u_{p+1}a_{p+1})$ and $\mathrm{MT}(u'_{p+1}a'_{p+1})$ are distinct. So the two linear statistics in $\mathcal{S}_{\text{linear}}$ (choosing $c = m + n$) successfully distinguish $\overline{u_1a_1} \ldots \overline{u_pa_p}\ \overline{u_{p+1}a_{p+1}}$ from $\overline{u_1a_1} \ldots \overline{u_pa_p}\ \overline{u'_{p+1}a'_{p+1}}$ since $36(m + n) \geq 18d_G(g, g')$. ☐

The proof of Theorem C.4 is complete. ☐

Bibliography

[AC16] Yago Antolín and Laura Ciobanu. Finite generating sets of relatively hyperbolic groups and applications to geodesic languages. *Trans. Amer. Math. Soc.*, 368(11):7965–8010, 2016.

[BBF15] Mladen Bestvina, Ken Bromberg, and Koji Fujiwara. Constructing group actions on quasi-trees and applications to mapping class groups. *Publ. Math. Inst. Hautes Études Sci.*, 122:1–64, 2015.

[BBF21] Mladen Bestvina, Ken Bromberg, and Koji Fujiwara. Proper actions on finite products of quasi-trees. *Ann. H. Lebesgue*, 4:685–709, 2021.

[BBFS19] Mladen Bestvina, Ken Bromberg, Koji Fujiwara, and Alessandro Sisto. Acylindrical actions on projection complexes. *Enseign. Math.*, 65(1-2):1–32, 2019.

[BDS07] Sergei Buyalo, Alexander Dranishnikov, and Viktor Schroeder. Embedding of hyperbolic groups into products of binary trees. *Invent. Math.*, 169(1):153–192, 2007.

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

[BL07] S. V. Buyalo and N. D. Lebedeva. Dimensions of locally and asymptotically self-similar spaces. *Algebra i Analiz*, 19(1):60–92, 2007.

[Bow12] B. H. Bowditch. Relatively hyperbolic groups. *Internat. J. Algebra Comput.*, 22(3):1250016, 66, 2012.

[BS00] M. Bonk and O. Schramm. Embeddings of Gromov hyperbolic spaces. *Geom. Funct. Anal.*, 10(2):266–306, 2000.

[BSCSW11] Alexander Bendikov, Laurent Saloff-Coste, Maura Salvatori, and Wolfgang Woess. The heat semigroup and Brownian motion on strip complexes. *Adv. Math.*, 226(1):992–1055, 2011.

[BST12] Itai Benjamini, Oded Schramm, and Ádám Timár. On the separation profile of infinite graphs. *Groups Geom. Dyn.*, 6(4):639–658, 2012.

[Can84] James W. Cannon. The combinatorial structure of cocompact discrete hyperbolic groups. *Geom. Dedicata*, 16(2):123–148, 1984.

[CM04] Ruth Charney and John Meier. The language of geodesics for Garside groups. *Math. Z.*, 248(3):495–509, 2004.

[Coz21] Corentin Le Coz. Poincaré profiles of lamplighter diagonal products, 2021.

[DL01] Reinhard Diestel and Imre Leader. A conjecture concerning a limit of non-Cayley graphs. *J. Algebraic Combin.*, 14(1):17–25, 2001.

[Dra03] A. N. Dranishnikov. On hypersphericity of manifolds with finite asymptotic dimension. *Trans. Amer. Math. Soc.*, 355(1):155–167, 2003.

[DS05] Cornelia Druţu and Mark Sapir. Tree-graded spaces and asymptotic cones of groups. *Topology*, 44(5):959–1058, 2005. With an appendix by Denis Osin and Mark Sapir.

[DWXX11] Juan Deng, Zhi-Ying Wen, Ying Xiong, and Li-Feng Xi. Bilipschitz embedding of self-similar sets. *J. Anal. Math.*, 114:63–97, 2011.

[DY05] François Dahmani and AslıYaman. Bounded geometry in relatively hyperbolic groups. *New York J. Math.*, 11:89–95, 2005.

[Dyu00] Anna Dyubina. Instability of the virtual solvability and the property of being virtually torsion-free for quasi-isometric groups. *Internat. Math. Res. Notices*, (21):1097–1101, 2000.

[ECH$^+$92] David B. A. Epstein, James W. Cannon, Derek F. Holt, Silvio V. F. Levy, Michael S. Paterson, and William P. Thurston. *Word processing in groups.* Jones and Bartlett Publishers, Boston, MA, 1992.

[EFW12] Alex Eskin, David Fisher, and Kevin Whyte. Coarse differentiation of quasi-isometries I: Spaces not quasi-isometric to Cayley graphs. *Ann. of Math. (2)*, 176(1):221–260, 2012.

[EFW13] Alex Eskin, David Fisher, and Kevin Whyte. Coarse differentiation of quasi-isometries II: Rigidity for Sol and lamplighter groups. *Ann. of Math. (2)*, 177(3):869–910, 2013.

[Far98] B. Farb. Relatively hyperbolic groups. *Geom. Funct. Anal.*, 8(5):810–840, 1998.

[FM92] K. J. Falconer and D. T. Marsh. On the Lipschitz equivalence of Cantor sets. *Mathematika*, 39(2):223–233, 1992.

[FM98] Benson Farb and Lee Mosher. A rigidity theorem for the solvable Baumslag-Solitar groups. *Invent. Math.*, 131(2):419–451, 1998. With an appendix by Daryl Cooper.

[GM08] Daniel Groves and Jason Fox Manning. Dehn filling in relatively hyperbolic groups. *Israel J. Math.*, 168:317–429, 2008.

[Gro81] Mikhael Gromov. Groups of polynomial growth and expanding maps. *Inst. Hautes Études Sci. Publ. Math.*, (53):53–73, 1981.

[Gro93] M. Gromov. Asymptotic invariants of infinite groups. In *Geometric group theory, Vol. 2 (Sussex, 1991)*, volume 182 of *London Math. Soc. Lecture Note Ser.*, pages 1–295. Cambridge Univ. Press, Cambridge, 1993.

[Hed67] G. A. Hedlund. Remarks on the work of Axel Thue on sequences. *Nordisk Mat. Tidskr.*, 15:148–150, 1967.

[HM20] David Hume and John M. Mackay. Poorly connected groups. *Proc. Amer. Math. Soc.*, 148(11):4653–4664, 2020.

[HMT22] David Hume, John M. Mackay, and Romain Tessera. Poincaré profiles of Lie groups and a coarse geometric dichotomy. *Geom. Funct. Anal.*, 32(5):1063–1133, 2022.

[HNS22] Sam Hughes, Patrick S. Nairne, and Davide Spriano. Regularity of quasi-geodesics characterises hyperbolicity, 2022.

[HR03] Derek F. Holt and Sarah Rees. Regularity of quasigeodesics in a hyperbolic group. *Internat. J. Algebra Comput.*, 13(5):585–596, 2003.

[HS13] David Hume and Alessandro Sisto. Embedding universal covers of graph manifolds in products of trees. *Proc. Amer. Math. Soc.*, 141(10):3337–3340, 2013.

[Hum17] David Hume. Embedding mapping class groups into a finite product of trees. *Groups Geom. Dyn.*, 11(2):613–647, 2017.

[Kas22] Daniel Kasprowski. Coarse embeddings into products of trees. *Kyoto J. Math.*, 62(1):225–229, 2022.

[Ker23] Alice Kerr. Tree approximation in quasi-trees. *Groups Geom. Dyn.*, 17(4):1193–1233, 2023.

[MM99] Howard A. Masur and Yair N. Minsky. Geometry of the complex of curves. I. Hyperbolicity. *Invent. Math.*, 138(1):103–149, 1999.

[MM00] H. A. Masur and Y. N. Minsky. Geometry of the complex of curves. II. Hierarchical structure. *Geom. Funct. Anal.*, 10(4):902–974, 2000.

[MS83] David E. Muller and Paul E. Schupp. Groups, the theory of ends, and context-free languages. *J. Comput. System Sci.*, 26(3):295–310, 1983.

[MS13] John M. Mackay and Alessandro Sisto. Embedding relatively hyperbolic groups in products of trees. *Algebr. Geom. Topol.*, 13(4):2261–2282, 2013.

[MS20] John M. Mackay and Alessandro Sisto. Quasi-hyperbolic planes in relatively hyperbolic groups. *Ann. Acad. Sci. Fenn. Math.*, 45(1):139–174, 2020.

[Nai23] Patrick S. Nairne. Embeddings of trees, Cantor sets and solvable Baumslag-Solitar groups. *Geom. Dedicata*, 217(2):Paper No. 13, 26, 2023.

[NS95] Walter D. Neumann and Michael Shapiro. Automatic structures, rational growth, and geometrically finite hyperbolic groups. *Invent. Math.*, 120(2):259–287, 1995.

[Osi05] D. Osin. Asymptotic dimension of relatively hyperbolic groups. *Int. Math. Res. Not.*, (35):2143–2161, 2005.

[Osi06a] D. V. Osin. Relative Dehn functions of amalgamated products and HNN-extensions. In *Topological and asymptotic aspects of group theory*, volume 394 of *Contemp. Math.*, pages 209–220. Amer. Math. Soc., Providence, RI, 2006.

[Osi06b] Denis V. Osin. Relatively hyperbolic groups: intrinsic geometry, algebraic properties, and algorithmic problems. *Mem. Amer. Math. Soc.*, 179(843):vi+100, 2006.

[Pap05] Panos Papasoglu. Quasi-isometry invariance of group splittings. *Ann. of Math. (2)*, 161(2):759–830, 2005.

[Pau01] Scott D. Pauls. The large scale geometry of nilpotent Lie groups. *Comm. Anal. Geom.*, 9(5):951–982, 2001.

[Pet22] Harry Petyt. Mapping class groups are quasicubical, 2022.

[Rul08] Alina Rull. *Embedding of right-angled Artin and Coxeter groups into products of trees.* PhD thesis, University of Zurich, Zürich, 2008.

[Sis13] Alessandro Sisto. Projections and relative hyperbolicity. *Enseign. Math. (2)*, 59(1-2):165–181, 2013.

[Why01] K. Whyte. The large scale geometry of the higher Baumslag-Solitar groups. *Geom. Funct. Anal.*, 11(6):1327–1343, 2001.

[Woe13] Wolfgang Woess. What is a horocyclic product, and how is it related to lamplighters ? *Internationale Mathematische Nachrichten*, 224:1–27, 2013.

[Yu98] Guoliang Yu. The Novikov conjecture for groups with finite asymptotic dimension. *Ann. of Math. (2)*, 147(2):325–355, 1998.