From the holdings of the Department of Special Collections , Princeton University Library, Princeton, NJ

These documents are either in the public domain, or can only be used for educational and research purposes (Fair use) as per U.S. Copyright law (text below). By accessing this file, all users agree that their use falls in the public domain or within fair use as defined by the copyrightlaw. We do not charge any permission or usage fees for the publication of images of material inour collections, including those provided by Princeton University Library, via our website, catalog, or directly from staff.For more information regarding copyright, credit, and citation please review the information onthe Copyright, Credit and Citation Guidelines page of our website.
https://library.princeton.edu/special-collections/policies/copyright-credit-and-citation-guidelines. If you have any questions, including how to order higher-quality images from the collection, contact Special Collections Public Services via our Ask Us! Form. https://library.princeton.edu/special-collections/ask-us

---

Doctoral Dissertations, 2011-2026 Mathematics

2015

3-Manifold Topology, with Groups and Randomne

Wu, Shuyun Conan

Princeton, NJ : Princeton University

---

http://arks.princeton.edu/ark:/88435/dsp010g354h522

U.S. Copyright law: The copyright law of the United States (Title 17, §108, United States Code) governs the making of photocopies or other reproductions of copyrighted material. Under certain conditions specified in the law, libraries and archives are authorized to furnish a photocopy or other reproduction. One of these specified conditions is that the photocopy or reproduction is not to be used for any purpose other than private study, scholarship, or research. If a user makes a request for, or later uses, a photocopy or reproduction for purposes in excess of fair use that user may be liable for copyright infringement. This institution reserves the right to refuse to accept a copying order if, in its judgment, fulfillment of the order would involve a violation of copyright law. https://www.copyright.gov/title17/title17.pdf

3-manifold Topology, with Groups and Randomness

Conan Wu

A Dissertation
Presented to the Faculty
of Princeton University
in Candidacy for the Degree
of Doctor of Philosophy

Recommended for Acceptance
by the Department of
Mathematics
Adviser: David Gabai

June 2015

© Copyright by Conan Wu, 2015.

All Rights Reserved

Abstract

In recent years there has been an increasing trend of interactions between different fields of mathematics, in particular, as a few major problems in low-dimensional topology gets solved this way, the attention has been shifting towards exploring the interactions between topology and various other fields. Note that such interaction can happen two-ways: one may apply a result or technique from another field to solve a problem in topology, or one could apply theorems or methods in topology to gain insights and solve problems in seemingly unrelated fields. In this thesis we will present results in both of the above mentioned directions. The first part of the thesis being an application of probability and Teichmuller theory to proving a purely topological theorem while the second part being applying classical topology to gain insights towards solving an open problem in group theory.

More precisely, the first result gave exact exponential growth rates for primitive elements in free groups with any number of generators by taking a topological viewpoint. The second result applies probability, among other things to prove the existence of hyperbolic homology 3-spheres with any given Heegaard genus g and Casson invariant n.

In light of the collaborative nature of the topic, I am extremely fortunate to have talked and worked with outstanding mathematicians from a variety of fields. The work in the first part is done in collaboration with Doron Puder while the second part with Alex Lubotzky and Joseph Maher.

Acknowledgements

First of all, no amount of credits can justify the gratitude owed to my advisor, David Gabai, who went far beyond the call of duty and took constant care of my interests. I have been a singular student with many unusual thoughts, among which are my doubts about graduate school and life in academia. Early in my tenure as a graduate student, Dave and I both discovered that unlike him I cannot concentrate on any single problem for long periods of time, which was frustrating and discouraging to me. Dave encouraged me by pointing out that great mathematicians work in many different styles and provided me with the privilege to travel around and find my voice by collaborating, which eventually led to this thesis.

I would also like to thank individuals who exposed me to this fascinating world of mathematics: William Neal, Charles Pugh, Larry Guth, Amie Wilkinson, my mentors at Canada/USA Mathcamp and countless mathematicians who engaged me in insightful discussions and generously shared their ideas and experience with me. Although I decided to not stay in mathematics, I sincerely believe what I learned in math and in grad school would be invaluable for the upcoming chapters in my life.

I would like to acknowledge the flexibility and generosity of the Princeton mathematics department, to me it is unlike any other mathematics department in the world. I shall thank Jill LeClair for catering to each students's individual needs and making the graduate school experience much more personal.

Lastly I wish to sincerely thank my co-authors of the two papers presented in this thesis: Alex Lubotzky, Joseph Maher and Doron Puder. They were very generous with their time, constantly holding discussions and sharing their expertise in fields that I am not familiar.

As a side remark, parts of this thesis were written during my travels in the Philippines, Thailand, Germany, China and, of course, Princeton. The completion of this thesis is made possible by many individuals whom inspired, influenced or encouraged

me to finish, long after I wandered out of the mathematical world; Along those lines, special credits are due to Igor Rivin, Peter Sarnak, Alex Lubotzky, Gil Kalai, Emmanuel Breuillard, Alessandra Iozzi, Christian Bonatti, Nets Katz, Danny Calegari and various subscribers of my blog.

Contents

Abstract iii

Acknowledgements iv

1 Primitive Elements in Free Groups **4**

1.1 Statement and Background 4

1.2 Whitehead Graphs 9

1.3 Proof of Theorems 14

1.3.1 Primitives and cyclic primitives 15

1.3.2 Ingredients for bounding cyclic primitives 16

1.3.3 Proof of Theorem 3 19

1.4 Most Triplets are Negligible 25

1.5 Generalizations and Further Questions 30

2 Random Methods in 3-Manifold Theory **32**

2.1 Motivation and Outline 32

2.2 Applying Random Method 34

2.3 Heegaard splitting distance 40

2.4 Coarse geometry 46

2.5 Local estimates far from the disc set 57

2.6 Train tracks and shadows 64

2.7 Exponential decay for splitting distance 73

Introduction

This thesis is written at an exciting time in the development of topology and mathematics in general: the landscape has been changing since Thurston's geometrization conjecture was proved in 2003 by Perelman using methods from PDEs. During my time as a graduate student I have witnessed a few major break throughs, including the proof of virtual Haken conjecture in 2013, many of which, just like Perelman's result, utilizes methods outside of classical low-dimensional topology. In fact, I have came to realize this as a general trend in modern mathematics: we are gradually arriving at a more unified view in mathematics by finding more and more connections between different fields.

In light of the above, I set out to become somewhat a 'Renaissance man' in mathematics, studying bits and pieces of anything that caught my imagination, participating in a board range of conferences and seminars and talking to mathematicians both in and outside of low dimensional topology.

It was summer 2012, on a hard-core hiking trip in some remote canyons in Utah, when I first heard about the problem of estimating the asymptotic growth of primitive elements in free groups from Tsachik Gelander:

Problem 1. Let F_k be the free group of k generators, a word $w \in F_k$ is called *primitive* if it belongs to some free generating set. What is the exponential growth rate of primitive words as the length grow to infinity?

I could not take my mind off the question for the rest of the hike, at the time it was

surprising to me that something so basic could be open; it turns out that not much is know about asymptotic behavior of free groups. A few months later I discussed the problem with fellow graduate student Doron Puder who specializes in combinatorial group theory, and we answered the question (see [**?**]):

Theorem 1. *For all $k \geq 3$, the exponential growth rate of primitive elements in F_k is*

$$\rho_k = \lim_{N\to\infty} \sqrt[N]{|P_{k,N}|} = 2k - 3.$$

My input in this is mostly some insights from classical topology, i.e. Whitehead moves and viewing free group and primitive elements as manifolds and simple closed curves inside manifolds, which provides a way of partitioning and counting the elements.

During my time in Jerusalem in my third year, Alex Lubotzky introduced me to his fascinating vision of applying random methods in solving existence problems in topology: instead of proving some manifold exists by constructing one, one proves that they occupied a positive proportion of all manifolds in some reasonable measure, hence such objects must exist. In short, one proves something exist by showing there is a large amount of them without knowing a single example. This might seem counterintuitive at first glance, but such methods has been long used in fields such as graph theory and combinatorics, most famously by Paul Erdős in showing the existence of expander graphs.

By the end of my stay in Israel, with lots of guidance from Alex, we arrived at an outline of proof of

Theorem 2. *For any genus $g \geq 2$ and any integer k, there exist infinitely many closed hyperbolic homology 3-spheres with Casson invariant k and Heegaard genus g.*

The paper [LMW14] is later completed with Joesph Maher who specializes in the theory of Teichmuller spaces.

In light of the above, this thesis is naturally divided into two parts: the first part deals with Problem 1 and Theorem 1 while the second about Theorem 2.

Notes on Collaboration

As there are multiple authors involved in both papers, I would like to make my humble attempt to attribute credits whenever they are due. However, the collaborative nature of this work has made it almost impossible to completely separate the parts.

I introduced the use of the Whitehead algorithm to help address the problem of partitioning primitives. Crucial counting arguments were worked out (in coffee shops) directly with Doron Puder where I supplied various ideas such as negligible triples and he supplied combinatorial expertise. The results are published in [PW14].

In the second paper, as mentioned above, Alex Lubotzky presented me with the problem and guided me through the initial part of the thinking process until I was able to combine topology, Johnson homomorphism and random walks to show that hyperbolic homology 3-spheres with any given Casson invariant decays sub-exponentially when measured by the random walk in $MCG(S_g)$. After some discussions on the topic, both online and in-person, the proof of exponential decay of splittings with less than full Heegaard genus is due to Joseph Maher as an improved version of his pervious results. The results are presented at the Yale Geometry Seminar and appear in the preprint [LMW14].

Chapter 1

Primitive Elements in Free Groups

1.1 Statement and Background

Let F_k be the free group of k generators $F_k = \langle x_1, \ldots, x_k \rangle$ $(k \geq 2)$. Recall that elements in F_k are represented by words in the alphabet $\left\{x_{1,}^{\pm 1} \cdots, x_k^{\pm 1}\right\}$ and uniquely represented by reduced words, that is those that does not contain '$x_i^{\pm} x_i^{\mp}$' for any i.

Definition. $x \in F_k$ is called **primitive** if there exists $y_1, \cdots, y_{k-1}$ such that $F_k = \langle x, y_1, \cdots, y_{k-1} \rangle$. I.e. w belongs to a k-element generating set. Such set is called a **free generating set.**

As a simple example, consider $F_2 = \langle a, b \rangle$. By default we have a, b are primitive. Since $a = ab \cdot b^{-1}$, $\langle ab, b \rangle$ generates all of F_2, hence ab is primitive. Along the same line it is easy to see ab^n is primitive for all $n \in \mathbb{Z}$ and a^n is not primitive unless $n = \pm 1$. In some sense one can think of primitive elements as prime numbers in free groups.

Like the study of prime numbers, a natural question to ask would be how many primitive elements are there in F_k?

Recall the standard norm in F_k is given by

$$\| x \| = \text{word length of } w$$

where w is the reduced word representing $x \in F_k$.

Consider the sphere $\partial B_{N,k} = \{x \in F_k \mid\mid\mid x \mid\mid = N\}$, let

$$P_{N,k} = \{x \in \partial B_{N,k} \mid x \text{ is primitive}\}.$$

Estimating the size of $P_{N,k}$ or $P_{\leq N,k}$ would give an answer to the above question. As we shall see later, in our case there is little distinction between the two.

Let w be any reduced word of length n in F_k, there are exactly $2k-1$ reduced words of length $n+1$ that starts with w (i.e. w can be followed by any character in the alphabet except for the inverse of the last character of w); furthermore all elements of $\partial B_{n+1,k}$ are obtained that way. Hence we have

$$\mid \partial B_{N,k} \mid = 2k(2k-1)^{N-1}.$$

Similar to what we did to show $ab^n \in F_2$ is primitive, we can see that for any element in $x \in F_{k-1}$, $x_k x \in F_k$ belongs to the generating set $\langle x_1, \cdots, x_{k-1}, x_k x \rangle$ and $\| x_k x \| = \| x \| + 1$. This gives raise to a simple lower bound

$$\mid P_{N,k} \mid \geq \mid \partial B_{N-1,k-1} \mid = (2k-2)(2k-3)^{N-2} \geq (2k-3)^{N-1}.$$

i.e. $\mid P_{N,k} \mid$ grows exponentially as $N \to \infty$ with exponential growth rate at least

$$\lim_{N\to\infty} \sqrt[N]{(2k-3)^{N-1}} = 2k-3.$$

As a consequence of this exponential growth, we know for all $k > 2$ the size of $P_{N,k}$ and $P_{\leq N,k}$ will only differ by a multiplicative constant, hence having the same growth rate.

The asymptotic growth of primitive elements has attracted a considerable amount

of interest over the past years. The open question is listed, for example, by M. Wicks, in [BMS02a, Problem F17] and the active website [BGV, Problem F19].

It is known (see, for example, [BMS02b]) that the set of primitive elements is exponentially small in F_k. Namely

$$\rho_k := \limsup_{N\to\infty} \sqrt[N]{|P_{N,k}|} < \lim_{N\to\infty} \sqrt[N]{|\ \partial B_{N,k}\ |} = 2k-1.$$

In 2004 Igor Rivin proved ([Riv04]) that for F_2, the exponential growth rate

$$\lim_{N\to\infty} \sqrt[N]{|\ P_{N,2}\ |} = \sqrt{3}$$

which remained the only case where the exact growth rate is known.

For $k \geq 3$, various upper bounds on ρ_k have been established [BV02, BMS02b, Shp05]. The best to date is due to Shpilrain [Shp05] who showed $\rho_k < \lambda_k$, where λ_k is the greatest real root of $\lambda\left(\lambda^2-1\right)\left(\lambda-(2k-2)\right)+1$. We know that $\lambda_k < 2k-2$ for each k, but λ_k approaches $2k-2$ as $k \to \infty$.

Our work completely answered the above question by giving the following tight result:

Theorem. *(Main Theorem) For all $k \geq 3$,*

$$\rho_k = \lim_{N\to\infty} \sqrt[N]{|P_{k,N}|} = 2k-3.$$

Moreover, there are positive constants c_k and C_k such That

$$c_k \cdot N \cdot (2k-3)^N \leq |P_{k,N}| \leq C_k \cdot N \cdot (2k-3)^N .$$

To obtain the result we first reduce the problem to estimating the number of conjugacy classes of primitive elements rather than counting the words themselves.

A word $w = a_1 a_2 \cdots a_N$ is called **cyclically reduced** if $a_1 \neq a_N^{-1}$; a *cyclic word* is the equivalence class of cyclically reduced words up to cyclic permutation. One can easily observe that conjugacy classes in F_k are uniquely represented by cyclic words. We denote *the* **cyclic word** associated with w by $[w]$. Let the **cyclic length** of w be the length of cyclically reduced representatives of $[w]$, denoted it by $|w|_c$.

Since conjugates of primitive elements are primitive, we can break down the problem into estimating the growth of each conjugacy class and counting the number of primitive conjugacy classes with respect to cyclic length.

As mentioned above, the growth rate of primitive elements in F_2 is known to be $\sqrt{3}$, in fact the number of primitive conjugacy classes has exponential growth rate 1, the major contribution comes from the growth of individual conjugacy classes. But as we shall see in Proposition 1, things are very different in F_k with $k \geq 3$ where the generic primitive word is nearly cyclically reduced and the exponential growth rate would be the same for primitive elements and conjugacy classes of them.

Let $C_{k,N}$ be the set of primitive cyclic words of cyclic length N in F_k. As mentioned above, any word where some letter appears exactly once (either as itself or its inverse) is primitive by forming a generating set with the rest of the alphabet. This subset will be crucial in our analysis, let us denote it by $L_{k,N}$ i.e.

$$C_{k,N} \coloneqq \{[w] \,|\, w \in F_k \text{ is primitive and } |w|_c = N\}$$

$$L_{k,N} \coloneqq \{[w] \in C_{k,N} \,|\, \exists\, x \in X \text{ s.t. } x \text{ appears in } w \text{ exactly once}\}\,.$$

Since all elements of $L_{k,N}$ have cyclically reduced representative starting with a single appearance letter, we can easily count:

$$|L_{k,N}| \leq 2k\,(2k-2)\,(2k-3)^{N-2}\,.$$

The inequality comes from multiple counting where the word contains multiple letters appearing only once, since words cannot be counted more than $2k$ times we also have an easy lower bound:

$$|L_{k,N}| \geq (2k-2)(2k-3)^{N-2},$$

so the exponential growth rate of $L_{k,N}$ is $2k-3$.

How large is $L_{k,N}$ compare to $C_{k,N}$? It turned out, quite surprisingly that except for an exponentially small set, all primitive cyclic-words contain one of the letters exactly once, namely:

Theorem 3. *For $k \geq 3$,*

$$\limsup_{N\to\infty} \sqrt[N]{|C_{k,N} \setminus L_{k,N}|} < 2k-3.$$

Hence for $k \geq 3$ the growth rate of $C_{k,N}$ is equal to that of $L_{k,N}$. Combining that with a result in[MS03, Prop 1.4] which states that $|C_{k,N}| = 4\varphi(N)$ where $\varphi$is the Euler function, we have:

Corollary. *For $k \geq 2$*

$$\limsup_{N\to\infty} \sqrt[N]{|C_{k,N}|} = 2k-3.$$

Moreover, for $k \geq 3$ there are constants $d_k, D_k > 0$ such that

$$d_k \cdot (2k-3)^N \leq |C_{k,N}| \leq D_k \cdot (2k-3)^N.$$

Then we make use of this fact and bound the number of actual primitive elements in those conjugacy classes, namely:

Proposition 1. *For $k \geq 3$, if $|C_{k,N}| \leq D_k \cdot (2k-3)^N$, then there exists $D > 0$ s.t.*

$$|P_{k,N}| \leq D \cdot N \cdot (2k-3)^N.$$

Hence

$$\rho(k) = \limsup_{N\to\infty} \sqrt[N]{|P_{k,N}|} = 2k - 3.$$

The above proposition gave a complete answer to the growth question. The major component here is the surprising result in 3, the proof is an thorough analysis of the Whitehead algorithm and Whitehead graphs which arise in topology. We divide the set of primitives into finitely many classes according to certain properties of their Whitehead graphs.

The structure of this part of the thesis is as follows: In section 1.2 we introduce Whitehead graphs and prove key lemmas used in establishing 3. Then we apply them in Section 1.3 to establish our main theorem except for one technical counting lemma which is left to be proved in Section 1.4. Section 1.5 discusses some generalizations and further questions.

1.2 Whitehead Graphs

Let M_k be the connected sum of k copies of $\mathbb{S}^1 \times \mathbb{S}^2$, or equivalently, double of a genus k handlebody, we can easily see $\pi_1(M_k) = F_k$. Fix a set of k disjoint 2-spheres $S_1, S_2, \cdots, S_k$, one corresponding to each summand, so that $\widehat{M_k} = M_k \setminus \bigcup_{i=1}^{k} S_i$ is simply connected with $2k$ boundary components $S_1^+, S_1^-, S_2^+, \cdots, S_k^-$. i.e. $\{S_i\}$ is the double of a cut system of H_k (a cut system is a set of disjoint discs that cuts the handlebody into a ball).

For every $w \in F_k = \pi_1(M_k)$, the cyclic word $[w]$ can be realized as a simple curve in M_k. Conversely, given any oriented curve in M_k we can write down a cyclic word in F_k by reading off the sequence of spheres the curve intersects, with signs. Hence we get a bijective correspondence between cyclic words $[w]$ and homotopy classes of oriented simple curves in M_k.

Given any proper non-empty subset $\mathcal{U} \subset \left\{S_1^+, S_1^-, \cdots, S_k^-\right\}$, there is an embedded

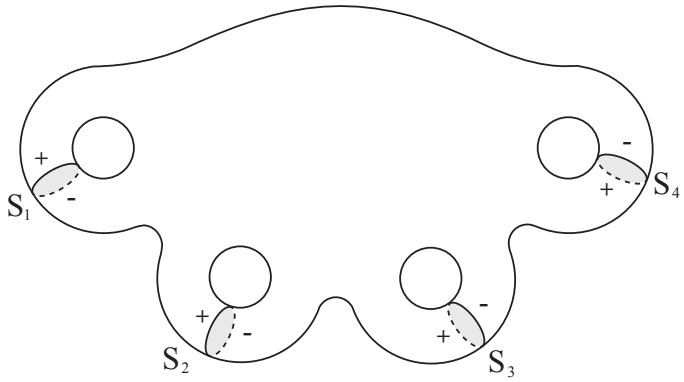


Figure 1.1: Spheres S_i in H_k

2-sphere $S_{\mathcal{U}}$ in $\widehat{M}$ separating the boundary components in $\mathcal{U}$ from those not in $\mathcal{U}$. For every $v \in X^{\pm 1}$ denote by S_v the corresponding boundary component of $\widehat{M}$ (so $S_{x_i} = S_i^+$ and $S_{x_i^{-1}} = S_i^-$). If there exists some $v = x_j^{\varepsilon} \in X^{\pm 1}$ such that $S_v \notin \mathcal{U}$ and $S_{v^{-1}} \in \mathcal{U}$ then $S_{\mathcal{U}}$ is an essential non-separating sphere in $\widehat{M} \cup S_j$, i.e. a non-contractible embedding of a sphere which does not separate $\widehat{M} \cup S_j$ into two connected components.

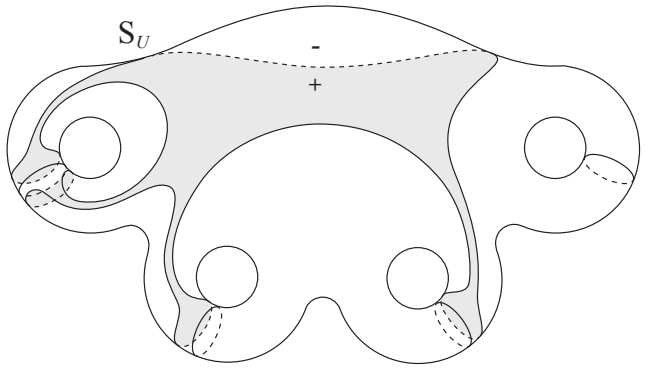


Figure 1.2: $S_{\mathcal{U}}$ with $\mathcal{U} = \left\{S_1^+, S_1^-, S_2^+, S_3^-\right\}$

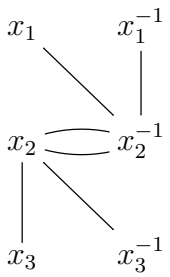
The **Whitehead automorphism** $\varphi_{(\mathcal{U},v)}$ of F_k is then defined by replacing the sphere S_j by $S_{\mathcal{U}}$ and write each cyclic word as the intersection pattern of the corresponding curve with the new set of spheres. In the example illustrated above, $S_3^+ \notin \mathcal{U}$ and $S_3^- \in \mathcal{U}$ hence we may replace S_3 with $S_{\mathcal{U}}$. Writing down $\varphi_{(\mathcal{U},v)}$ formally one gets:

- $\varphi_{(\mathcal{U},v)}(v) = v$; $\varphi_{(\mathcal{U},v)}(v^{-1}) = v^{-1}$;
- For $u \neq v$,
  - $\varphi_{(\mathcal{U},v)}(u) = u$ if $S_u, S_{u^{-1}} \notin \mathcal{U}$;
  - $\varphi_{(\mathcal{U},v)}(u) = vuv^{-1}$ if $S_u, S_{u^{-1}} \in \mathcal{U}$;
  - $\varphi_{(\mathcal{U},v)}(u) = vu$ and $\varphi_{(\mathcal{U},v)}(u^{-1}) = uv^{-1}$ if $S_u \in \mathcal{U}$, $S_{u^{-1}} \notin \mathcal{U}$;

Note that there is clearly no automorphism that replaces a generator-repersenting-sphere by either a contractible or a separating sphere, hence this is as good as one could hope for in terms of replacing individual generators.

The Whitehead automorphism is a classical tool dating back to 1936, see [Whi36a, Whi36b], it was originally invented to answer the question: Given two words $w_1, w_2 \in F_k$, when does there exist an automorphism $\phi \in \mathrm{Aut}\,(F_k)$ mapping w_1 to w_2? We shall observe that w_1 is primitive if and only if there is an automorphism mapping it to a single-letter. Hence we can test the primitivity of a given word by finding a sequence of length-reducing Whitehead automorphisms, we will soon see such sequence always exists when w is primitive.

By forgetting the order in which the spheres are intersected and looking only at the arcs connecting boundary components of $\widehat{M}$ one gets a finite graph with $2k$ vertices labeled $X^{\pm 1} = \left\{x_1^{\pm 1}, \cdots, x_k^{\pm 1}\right\}$. This is called the **Whitehead graph** of the cyclic word $[w]$, denoted by $\Gamma(w)$. For example, $\Gamma\,(w)$ for $w = x_1 x_2^2 x_3^{-1} x_2^{-2} \in F_3$ is:



Going from manifolds to graphs, every $(\mathcal{U}, v)$ defined as above corresponds to a partition of the vertices by $Z = \{x_u \,|\, S_u \in \mathcal{U}\}$ and $Y = X^{\pm 1} \backslash (Z \cup \{v\})$. Denote $\phi_{Y,Z,v} = \varphi_{(\mathcal{U},v)}$, and notice that $X^{\pm 1} = Y \sqcup Z \sqcup \{v\}$.

We now recall some elementary definitions in group and graph theory:

Definition. A **free factor** of F_k is a free normal subgroup N of F_k such that the quotient of F_k/N is free.

A **cut vertex** of a finite graph G is a vertex such that $G \backslash \{v\}$ is disconnected (this includes the case where G is itself disconnected).

The following theorem, part of the basis for Whitehead's algorithm, plays a central role in our argument:

Fact 1. *[Sta99, Thm 2.4] If w is contained in a proper free factor of F_k, then $\Gamma(w)$ has a cut vertex.*

In particular, all primitive elements are contained in a rank one free factor, hence have Whitehead graphs with cut vertices.

The cyclic length $|w|_c$ of w is the number of edges in the Whitehead graph corresponding to a cyclically reduced representative. A natural candidate for a length reducing Whitehead automorphism is therefore to replace the sphere corresponding to the cut vertex v by one that separates a connected component of $\Gamma(w) \backslash \{v\}$. Indeed,

Proposition 2. *[Sta99, Prop 2.3] Let v be a cut-vertex of $\Gamma(w)$, and let Y and Z be a non-trivial partition of the remaining vertices so that there are no edges between Y and Z, and $v^{-1} \in Z$. Then*

$$|\phi_{Y,Z,v}(w)|_c = |w|_c - E(Y, v).$$

Here $E(Y, v)$ is the number of edges connecting v to Y. For instance, for $w = x_1 x_2^2 x_3^{-1} x_2^{-2}$ as above, there are two possible cut-vertices: x_2 and x_2^{-1}. If one chooses

$v = x_2$ and $Z = \left\{x_1, x_1^{-1}, x_2^{-1}\right\}$, then $[\phi_{Y,Z,v}(w)] = [x_2x_1x_2x_3^{-1}x_2^{-2}] = [x_1x_2x_3^{-1}x_2^{-1}]$ has cyclic length 4.

Moreover, following Fact 1, if w is contained in a proper free factor and $|w|_c > 1$, then we always have a cut vertex; Proposition says we can always shorten the cyclic length of w through applying a Whitehead automorphism if there is a partition (Y, Z, v) with $E(Y, v) > 0$. Nonexistence of such triple would imply that $\Gamma(w)$ is disconnected and w is a word in a proper subset of the letters, say $x_1, \ldots, x_j$ $(j < k)$ but does not belong to a proper free factor in $F(x_1, \ldots, x_j)$. Luckily this cannot occur for primitive w with $|w|_c > 1$. Hence when w is primitive we can always apply a sequence of Whitehead automorphisms according to cut vertices in the Whitehead graphs, until it becomes a (conjugate of) single-letter word.

Now we may consider all possible triples (Y, Z, v) in order give an upper bound on the number of primitive cyclic words. We say thaT a triplet (Y, Z, v) is *valid* for the cyclic word $[w]$ if it satisfies the statement in Proposition 2 (namely, if v is a cut-vertex of $\Gamma(w)$, Y and Z are a non-trivial partition of the remaining vertices with $E(Y, Z) = 0$, and $v^{-1} \in Z$).

Let $A_{Y,Z,v}$ denote the set of all cyclic words having (Y, Z, v) as a valid triplet and $A^N_{Y,Z,v}$ be its subset where the words has cyclic length N; namely

$$A_{Y,Z,v} = \{[w] \mid (Y, Z, v) \text{ is a valid triplet for } w\}$$

and

$$A^N_{Y,Z,v} = \{[w] \in A_{Y,Z,v} \,|\, |w|_c = N\}.$$

By Fact 1,

$$C_{k,N} \subseteq \bigcup_{(Y,Z,v)} A^N_{Y,Z,v}$$

taking the union over all possible triplets partitioning $X^{\pm}$ (with $Y, Z \neq \emptyset$ and $v^{-1} \in$

Z).

We proceed by bounding the growth of primitives in $A_{Y,Z,v}$ for each of the finitely many triplets (Y, Z, v). Intuitively, the cut vertex and partition will restrict the number of possible ways to connect vertices, hence result in a smaller growth rate. In the extreme case, if both sets Y and Z contain roughly half of the elements of $X^{\pm 1}$, namely $|Y| \approx k$ and $|Z| \approx k$; then from any vertex in Yone can only connect to another vertex in Y, resulting in k choices. If we ignore the possibility of going through v, then the possible number of such cyclic words would be roughly only k^N, which makes the exponential growth rate k (note that this applies to the whole set $A_{Y,Z,v}$ and not only to the primitives in it). Hence we should expect $A_{Y,Z,v}$ to grow faster for triples (Y, Z, v) where one of Y, Z is almost all of $X^{\pm 1}$.

Indeed it turns out that $A_{Y,Z,v}$ is negligible unless one of Y, Z is very small:

Proposition 3. *Every triplet* (Y, Z, v) *satisfies*

$$\limsup_{N\to\infty} \sqrt[N]{\left|A^N_{Y,Z,v}\right|} < 2k - 3,$$

unless $min(|Y|, |Z|) = 1$ *or* $Y = \{x, x^{-1}\}$ *for some letter* x.

The proof of this proposition involves some careful analysis in various cases, we postpone it to Section 1.4. In Section 1.3 we assume this proposition, give the precise growth rates for the remaining essential partitions and obtain our theorems.

1.3 Proof of Theorems

In this section we complete the proofs of Theorem 3 and the main theorem.

1.3.1 Primitives and cyclic primitives

Here we present the observation that, unlike in F_2 , for $k \geq 3$ it suffices to count the conjugacy classes containing primitive words.

Proposition 4. *For $k \geq 3$, if $|C_{k,N}| \leq C \cdot (2k-3)^N$ for some $C > 0$ as in the conclusion of Theorem 3, then*

$$|P_{k,N}| \leq D \cdot N \cdot (2k-3)^N$$

for some $D > 0$.

Proof. Each $w \in P_{k,N}$ is of the form

$$uw'u^{-1}$$

where $w' \in F_k$ is cyclically reduced and primitive. Let ℓ be the word length of u, we have $\ell \leq \frac{N-1}{2}$ and $|w|_c = |w'|_c = N - 2\ell$.

Primitivity of w' implies that, in particular, it is not a proper power; hence all its cyclic shifts are different. Namely, the cyclic word $[w']$ is represented by exactly $N - 2\ell$ distinct cyclically reduced words. On the other hand, u can be any word of length ℓ as long as the first letter of u^{-1} and the last letter of u do not cancel out their adjacent letters in w'. There are $(2k-1)^{\ell-1}(2k-2) \leq (2k-1)^{\ell}$ such words.

Therefore,

$$\begin{aligned}
|P_{k,N}| &= \sum_{\ell=0}^{\lfloor \frac{N-1}{2} \rfloor} (N-2\ell) \left|C_{k,N-2\ell}\right| (2k-2) (2k-1)^{\ell-1} \\
&\leq N \cdot \sum_{\ell=0}^{\lfloor \frac{N-1}{2} \rfloor} \left|C_{k,N-2\ell}\right| (2k-1)^{\ell} \\
&\leq N \cdot \sum_{\ell=0}^{\lfloor \frac{N-1}{2} \rfloor} C \cdot (2k-3)^{N-2\ell} (2k-1)^{\ell} \\
&= N \cdot C \cdot (2k-3)^{N} \sum_{\ell=0}^{\lfloor \frac{N-1}{2} \rfloor} \left(\frac{2k-1}{(2k-3)^2}\right)^{\ell}.
\end{aligned}$$

For $k \geq 3$, $\left(\frac{2k-1}{(2k-3)^2}\right) \leq \frac{5}{9} < 1$. Bounding the geometric series we deduce $|P_{k,N}| \leq C \cdot N \cdot (2k-3)^N$. □

The proposition shows that the main theorem follows from Theorem 3. For the lower bound in the main theorem recall that the number of cyclically reduced primitive words of length N with one of the letters appearing exactly once is

$$N \cdot |L_{k,N}| = N \cdot 2k \cdot (2k-2) \cdot (2k-3)^{N-2}.$$

1.3.2 Ingredients for bounding cyclic primitives

It remains to bound the growth of cyclic primitives, i.e. prove Theorem 3. Before starting the proof we present a couple of useful facts which will be used in our proof. Prior to this project, my co-author, Doron Puder, has investigated extensively in the combinatorial and probability aspects of free groups and permutation groups. We will make use of one of his results (in collaboration with O. Parzanchevski) in [PP12]:

Theorem 4. *[PP12, Thm 1.7] Let $w = x_{i_1} x_{i_2} \cdots x_{i_N} \in F_k$. For each $n \in \mathbb{N}$we*

consider the symmetric group S_n and define a function $\psi_w : S_n^k \to S_n$ where

$$\psi_w(\sigma_1, \sigma_2, \cdots, \sigma_k) \longmapsto \sigma_{i_1}\sigma_{i_2}\cdots\sigma_{i_N}$$

Let $\sigma_1, \sigma_2, \cdots, \sigma_k$ be random permutations in S_n chosen independently with uniform distribution. Then w is non-primitive if and only if there exists some M such that for all $n > M$ we have

$$\mathbb{E}(|Fix(\sigma_w(\sigma_1, \sigma_2, \cdots, \sigma_k))|) > 1.$$

Where Fix(σ) denotes the set of fixed points of σ.

The moral here being when w is primitive, the push forward of the uniform measure from S_n^k to S_n by ψ_w would be the uniform measure on S_n giving expected number of fixed points exactly 1. This theorem says the converse is true: whenever w is non-primitive it will skew up the push forward measure in such a way that the expected number of fixed points would be strictly larger than 1 for large enough n.

This theorem serves as a useful tool in proving statements about primitive words which may otherwise be tedious to derive, for example:

Proposition 5. *Let $w_1, w_2 \in F_k$ be letter disjoint words (i.e. they do not share any common letters). If the concatenation w_1w_2 is primitive, then at least one of w_1 or w_2 is primitive.*

Proof. Since the push-forward measure by w is a class function, the probability $\mathbb{P}(w_{\sigma_1,\sigma_2,\cdots,\sigma_k}(i) = i)$ is independent of i (here $i \in \{1, \ldots, n\}$), and likewise, $\mathbb{P}(w_{\sigma_1,\sigma_2,\cdots,\sigma_k}(i) = j)$ is independent of i and j as long as $i \neq j$. Thus, $\mathbb{E}(|\mathrm{Fix}(w_{\sigma_1,\sigma_2,\cdots,\sigma_k})|) > 1$ if and only if $\mathbb{P}(w_{\sigma_1,\sigma_2,\cdots,\sigma_k}(1) = 1) > \frac{1}{n}$. Let $p(w, n) = \mathbb{P}(w_{\sigma_1,\sigma_2,\cdots,\sigma_k}(1) = 1)$.

Since w_1 and w_2 are letter disjoint, they induce independent push-forward measures on S_n. If both words are non-primitive then for large enough n, both $p(w_1, n) >$

$\frac{1}{n}$ and $p(w_2, n) > \frac{1}{n}$, which implies

$$\begin{aligned} p(w_1 w_2, n) &= \mathbb{P}((w_1 w_2)_{\sigma_1,\sigma_2,\cdots,\sigma_k}(1) = 1) \\ &= \sum_{j=1}^{n} \mathbb{P}((w_1)_{\sigma_1,\sigma_2,\cdots,\sigma_k}(1) = j) \cdot \mathbb{P}((w_2)_{\sigma_1,\sigma_2,\cdots,\sigma_k}(j) = 1) \\ &= p(w_1, n)p(w_2, n) + (n-1)\frac{1-p(w_1,n)}{n-1} \cdot \frac{1-p(w_2,n)}{n-1} \\ &= \frac{1}{n} + \frac{n}{n-1} \cdot \left(p(w_1, n) - \frac{1}{n}\right) \cdot \left(p(w_2, n) - \frac{1}{n}\right) > \frac{1}{n}. \end{aligned}$$

This contradicts the assumption that $w_1 w_2$ is primitive. □

Recall that by Proposition 3 which we will prove in Section 1.4, primitives from $A_{Y,Z,v}$ for most triplets (Y, Z, v) are negligible in our calculation for exponential growth. We make some simple observations about the remaining three types of triplets:

- $|Y| = 1$, say $Y = \{a\}$, and $w \in A_{Y,Z,v}$ then each appearance of a is followed by v^{-1} and each appearance of a^{-1} is preceded by v. It is not hard to see that the growth rate here is at least $\sqrt{(2k-3)^2 + 1} > 2k - 3$. Indeed, consider the $(2k-2)(2k-3)$ ordered reduced pairs of letters not containing $a^{\pm 1}$ and, in addition, the pair av^{-1}. Each one of these pairs can be followed by one of $(2k-3)(2k-3)+1$ of these pairs, which shows the lower bound. Since every possible pair of letters is followed by one of less than $(2k-1)^2$ possible pairs, the growth rate is strictly less than $2k-1$. In fact, the exact growth rate is the largest (real) root of $\lambda^5 - (2k-3)\lambda^4 - 3\lambda^3 + (2k-3)\lambda^2 + 3\lambda + (2k-3)$, which tends to $2k-3$ as $k \to \infty$.

- $Y = \{a, a^{-1}\}$ then every instance of $a^{\pm 1}$ in a word from $A_{Y,Z,v}$ is in the form $\ldots v a^m v^{-1} \ldots$ for some $0 \neq m \in \mathbb{Z}$. The exponential growth rate here is the largest (real) root of $\lambda^4 - (2k-2)\lambda^3 + (2k-4)\lambda^2 + (2k-2)\lambda - 6k + 11$, which

again approaches $2k-3$ from above as $k \to \infty$. Again, looking at pairs of letters one can easily infer the growth rate is strictly less than $2k-1$.

- $|Z| = 1$, namely $Z = \{v^{-1}\}$, then v^{-1} is followed only by v^{-1}, and v is preceded only by v. So $A_{Y,Z,v}$ consists of all cyclic words not containing $v^{\pm 1}$ (together with $\{v^m \mid m \in \mathbb{Z}\}$). Hence the growth rate is exactly $(2k-3)$.

We obtain the following naive bound from the observation above:

Corollary 1. *Let $k \geq 2$. Every triplet (Y, Z, v) satisfies*

$$\limsup_{N\to\infty} \sqrt[N]{\left|A^N_{Y,Z,v}\right|} < 2k-1.$$

1.3.3 Proof of Theorem 3

Recall that the second statement of Theorem 3 is that for $k \geq 3$,

$$\limsup_{N\to\infty} \sqrt[N]{|C_{k,N} \setminus L_{k,N}|} < 2k-3. \tag{1.1}$$

As mentioned in Section 1.1, when $k \geq 3$ the other two statements of the theorem follow from equation (1.1), and for $k = 2$, the relevant statement ($\limsup_{N\to\infty} \sqrt[N]{|C_{2,N}|} = 1$) is already known [MS03, Prop. 1.4]. We now prove equation (1.1) by induction on k. The moral of the proof will be that even when the set $A_{Y,Z,v}$ has exponential growth rate larger than $2k-3$, every cyclic *primitive* $[w] \in A_{Y,Z,v}$ can be shortened 'fast enough' by the corresponding Whitehead automorphism $\phi_{Y,Z,v}$.

Suppose $k \geq 3$ and

$$\limsup_{N\to\infty} \sqrt[N]{|C_{k-1,N}|} = 2k-5.$$

Recall from Proposition 2 we know the cyclic length reduction induced by Whitehead automorphism is the number of edges connecting the set Y to vertex v, we partition

$C_{k,N} \backslash L_{k,N}$ into a 'fast reducing' set and its complement:

Let $M_{k,N} \subseteq C_{k,N} \backslash L_{k,N}$be the set of cyclic primitive words w such that either

- $[w] \in A^N_{Y,Z,v}$ with $Y = \{x\}$ and x appearing at least 4 times in $[w]$, or
- $[w] \in A^N_{Y,Z,v}$ with $Y = \{x, x^{-1}\}$ and $[w]$ containing at least 2 instances of $vx^m v^{-1}$ (for any $0 \neq m \in \mathbb{Z}$).

Let $M^c_{k,N}$denote the complement of $M_{k,N}$ inside $C_{k,N} \backslash L_{k,N}$. We proceed by showing that the exponential growth rates of $|M_{k,N}|$ and $|M^c_{k,N}|$ are both strictly less than $2k-3$, for different reasons.

Lemma 1.

$$\limsup_{N\to\infty} \sqrt[N]{\left|M^c_{k,N}\right|} < 2k-3.$$

Proof. In light of Proposition 3 we only need to consider cyclic primitive $[w] \in M^c_{k,N}$ with a Whitehead partition (Y, Z, v) where $\min |Y|, |Z| = 1$ or $Y = \{x, x^{-1}\}$.

If $|Z| = 1$ then by definition $Z = \{v^{-1}\}$, and every cyclic word in $A_{Y,Z,v}$ is either a power of v or a word in the alphabet $(X \setminus \{v\})^{\pm 1}$. The set $\{v^n\}$ grows linearly hence is clearly negligible. In the latter case, $[w]$ is primitive in F_{k-1} with letters $X \setminus \{v\}$. It follows from the induction hypothesis that the exponential growth rate of this set of cyclic primitives is $2k-5 < 2k-3$.

Assume next that $Y = \{x, x^{-1}\}$ and $[w]$ has exactly one instance of $vx^m v^{-1}$ for some $|m| \geq 2$ (the case $|m| = 1$ is impossible as $M^c_{k,N} \cap L_{k,N} = \emptyset$). Pick a representative of the form $[x^m w']$, where w' has length $< N$ and does not contain the letter x. By Proposition 5 $[w']$ is primitive in F_{k-1} hence this set has exponential growth rate $2k-5$.

The remaining case is $Y = \{x\}$ and $x^{\pm 1}$ appearing exactly twice or thrice in $[w]$ (again, if it $x^{\pm 1}$ appears only once, then $[w]$ belongs to $L_{k,N}$). Consider first the case where $x^{\pm 1}$ appears exactly **twice**:

We can write $[w]$ in the form

$$\left[(u_1 x u_2)^{\pm 1} w_1 (u_1 x u_2)^{\pm 1} w_2\right],$$

where u_1, u_2 are maximal sequence of letters preceding and following both instances of $x^{\pm 1}$, respectively.

Let ℓ_i be the length of u_i. Up to a factor of N^3, which is negligible in terms of exponential growth rates, we know ℓ_1, ℓ_2, $|w_1|$ and $|w_2|$. There are about $(2k-3)^{\ell_1+\ell_2}$ options for the values of u_1 and u_2. The automorphism $\psi \in \mathrm{Aut}(F_k)$ which maps $x \mapsto u_1^{-1} x u_2^{-1}$ and leaves unchanged the remaining letters, maps $[w]$ to the primitive cyclic word

$$[w'] = \left[x^{\pm 1} w_1 x^{\pm 1} w_2\right].$$

Let $N' = N - 2(\ell_1 + \ell_2)$ be the length of w'. We claim that the number of possible w' is bounded above by $C \cdot (2k-3-\varepsilon)^{N'}$ for some $C, \varepsilon > 0$, ε small. This will suffice as the number of possible $[w]$ is then bounded by some polynomial in N times

$$\begin{aligned} & (2k-3-\varepsilon)^{N'} \cdot (2k-3)^{\ell_1+\ell_2} \\ = \; & (2k-3-\varepsilon)^{N-2\ell_1-2\ell_2} \cdot \sqrt{2k-3}^{2\ell_1+2\ell_2} \\ < \; & (2k-3-\varepsilon)^{N}. \end{aligned}$$

First, if one of w_1 or w_2 is trivial, then as in the preceding case, the number of possible $[w']$s is at most some constant times $(2k-5)^{N'}$. So assume $w_1, w_2 \neq 1$ and the word $[w']$ belongs to some $A^{N'}_{Y',Z',v'}$. By the maximality of u_1 and u_2, each of the vertices x and x^{-1} in the Whitehead graph $\Gamma(w')$ has at least two neighbors in $X^{\pm 1} \setminus \{x, x^{-1}\}$. Hence Y' cannot be $\{x\}, \{x^{-1}\}$ or $\{x, x^{-1}\}$. Also, it is not possible that $v' = x^{\pm 1}$ and $Z' = \left\{(v')^{-1}\right\}$, since cyclic words corresponding to this triplet are

either words in $X \setminus \{v'\}$ or powers of v'. Hence the triplet (Y', Z', v') induces some non-trivial partition of $(X \setminus \{x\})^{\pm 1}$ (both Y' and Z' intersect $(X \setminus \{x\})^{\pm 1}$). Hence w_1 and w_2 are words (albeit not cyclic) corresponding to some non-trivial triplet partitioning $X \setminus \{x\}$. But by Corollary 1, the exponential growth of such subsets is strictly less than $2(k-1) - 1 = 2k - 3$.

Finally, consider the case where $Y = \{x\}$ and $x^{\pm 1}$ appears exactly **three times** in $[w]$:
The proof that this subset grows slower than $(2k-3)^N$is very similar to the previous case. This time, each such word is of the form

$$\left[(u_1 x u_2)^{\pm 1} w_1 (u_1 x u_2)^{\pm 1} w_2 (u_1 x u_2)^{\pm 1} w_3\right]$$

with u_1, u_2 maximal. It can be shortened via an automorphism to

$$[w'] = \left[x^{\pm 1} w_1 x^{\pm 1} w_2 x^{\pm 1} w_3\right]$$

of length N'. Again, up to a polynomial factor of N^4 we know $\ell_1, \ell_2, |w_1|, |w_2|$ and $|w_3|$, and we claim that the number of options for $[w']$ is bounded by some constant times $(2k-3-\varepsilon)^{N'}$ with $\varepsilon > 0$. Hence the total number of options for $[w]$ is bounded by some polynomial in N times

$$(2k-3-\varepsilon)^{N-3\ell_1-3\ell_2} \cdot (2k-3)^{\ell_1+\ell_2} < (2k-3-\varepsilon)^N$$

for ε small enough. Indeed, if two of w_1, w_2 and w_3 are trivial we are again in the same situation as in the case $Y = \{x, x^{-1}\}$. Otherwise, the exact same argument as before shows $[w'] \in A^{N'}_{Y',Z',v'}$ for some triplet (Y', Z', v') partitioning $(X \setminus \{x\})^{\pm 1}$ non-trivially.

This covered all the cases of $M^c_{k,N}$ and hence the lemma is established. □

Now we move on to the remaining set $M_{k,N}$. The idea is to shorten such words by applying appropriate Whitehead automorphisms until the result falls outside of $M_{k,*}$ (i.e. outside of $M_{k,n}$ for all n). To achieve this goal we first consider cyclic words $[w] \in M_{k,N}$ such that the corresponding automorphism $\phi_{Y,Z,v}$ maps them into $L_{k,*}$: $[\phi_{Y,Z,v}(w)] \in L_{k,*}$. Denote this subset by $\widetilde{L}_{k,N} \subseteq M_{k,N}$. We claim that

Lemma 2.

$$\limsup_{N\to\infty} \sqrt[N]{\left|\widetilde{L}_{k,N}\right|} < 2k-3.$$

Proof. Consider first the words $[w] \in \widetilde{L}_{k,N}$ with triplet (Y,Z,v) where $Y=\{x\}$. The effect of $\phi = \phi_{Y,Z,v}$ on $[w]$ is precisely that: each instance of the form $\dots xv^{-1}\dots$ becomes simply $\dots x\dots$, and each instance of $\dots vx^{-1}\dots$ turns into $\dots x^{-1}\dots$. In particular, the only letter whose number of appearances in $[w]$ is changed by ϕ is v. By definition, $[w] \in \widetilde{L}_{k,N} \subseteq M_{k,N} \subseteq C_{k,N} \setminus L_{k,N}$, so the letter which $[\phi(w)]$ contains exactly once is necessarily $v^{\pm 1}$. Hence, aside for one, all occurrences of $v^{\pm 1}$ in $[w]$ are as part of either xv^{-1} or vx^{-1}. We deduce that $[w]$ is of the form

$$[vw']$$

with w' being a word of length $N-1$ in $2(k-2)$ building blocks of length 1: $(X\setminus\{x,v\})^{\pm 1}$ and 2 building blocks of length 2: $(xv^{-1})^{\pm 1}$. (In other words, w' is any word in F_{k-1} but where one of the letters is of length 2). This kind of words clearly has exponential growth rate strictly less than $2k-3$. (To be precise, the growth rate is the larger root of $\lambda^2-(2k-4)\lambda-2$.)

The complement of this latter subset inside $\widetilde{L}_{k,N}$ consists of primitive cyclic words belonging to $A^N_{Y,Z,v}$ with $Y=\{x,x^{-1}\}$. This time, ϕ turns each instance of $\dots vx^m v^{-1}\dots$ into $\dots x^m\dots$. The same arguments as before show that if $[w]$ is such a word, then

$$[w] = [vw']$$

where w' is composed of building blocks from $(X \setminus \{x, v\})^{\pm 1}$ together with $\{vx^m v^{-1} \,|\, 0 \neq m \in \mathbb{Z}\}$. Every letter in w' is followed by one of at most $(2k-4)$ possible letters, showing this type of words also has exponential growth rate $< 2k-3$, and thus completing the proof. □

Lemma 3.

$$\limsup_{N\to\infty} \sqrt[N]{|M_{k,N}|} < 2k-3.$$

Proof. Every $[w] \in M_{k,N}$ is equipped with some triplet (Y, Z, v) from the definition of $M_{k,N}$. First we observe the corresponding Whitehead automorphism $\phi_{Y,Z,v}$ shortens $[w]$ by at least 4. There are in total $2k\,(2k-2)$ triplets with $Y = \{x\}$ and $k\,(2k-2)$ triplets with $Y = \{x, x^{-1}\}$. Let $\mathcal{W}$ denote the set of these $2k\,(2k-2) + k\,(2k-2) = 6k\,(k-1)$ possible Whitehead automorphisms.

In other words, for every $[w] \in M_{k,N}$ there exists $\phi_1 \in \mathcal{W}$ such that $|\phi\,(w)\,|_c = N' \leq N-4$. If $[\phi\,(w)] \in M_{k,N'}$ we apply the corresponding automorphism $\phi_2 \in \mathcal{W}$ and obtain a cyclic word of length $\leq N-8$. Since for all n we have $C_{k,n} = L_{k,n} \cup M_{k,n} \cup M_{k,n}^c$, one can continue this process until the resulting cyclic word is either in $M_{k,n}^c$ or in $\widetilde{L}_{k,n}$ for some $2 \leq n \leq N$. Let $[\widehat{w}] \in M_{k,n}^c \cup \widetilde{L}_{k,n}$ be the cyclic-word we obtain this way, i.e.

$$[w] \overset{\phi_1}{\to} [w_1] \overset{\phi_2}{\to} [w_2] \to \ldots \overset{\phi_r}{\to} [\widehat{w}]\,.$$

Each element in $M_{k,N}$ is uniquely determined by $n = |\widehat{w}|_c$, $[\widehat{w}] \in M_{k,n}^c \cup \widetilde{L}_{k,n}$ and $(\phi_1, \phi_2, \cdots, \phi_r) \in \mathcal{W}^r$, where $r \leq R = \left\lfloor \frac{N-n}{4} \right\rfloor$. For each n, the number of possible tuples $(\phi_1, \phi_2, \cdots, \phi_r)$ is

$$\sum_{i=0}^{R} |\mathcal{W}|^i = \sum_{i=0}^{R} [6k\,(k-1)]^i \leq \frac{[6k\,(k-1)]^{R+1} - 1}{6k\,(k-1) - 1} \leq 2\,[6k\,(k-1)]^R\,.$$

By Lemmas 1 and 2, the possible number of $[\widehat{w}]$ of length n is bounded from above by

$C \cdot (2k-3-\varepsilon)^n$ for some $C, \varepsilon > 0$. Summing over all possible values of n we obtain:

$$\begin{aligned} |M_{k,t}| &\leq 2C \cdot \sum_{n=2}^{N} (2k-3-\varepsilon)^n \cdot [6k\,(k-1)]^R \\ &\leq 2C \cdot \sum_{n=2}^{N} (2k-3-\varepsilon)^n \cdot (\sqrt[4]{6k(k-1)})^{N-n}. \end{aligned}$$

For $k \geq 3$ we can pick ε small enough so that

$$\sqrt[4]{6k\,(k-1)} < 2k-3-\varepsilon,$$

hence

$$|M_{k,t}| < 2C \cdot N \cdot (2k-3-\varepsilon)^N .$$

□

This completes the proof of 1.1, hence also Theorem 3. Our main theorem now follows by Proposition 4.

1.4 Most Triplets are Negligible

As discussed in Section 1.2, so far we have proved the theorem except for Proposition 3, the proof of which we have postponed to this section. We begin by recalling the statement:

Proposition. *Every triplet (Y, Z, v) satisfies*

$$\limsup_{N\to\infty} \sqrt[N]{\left|A^N_{Y,Z,v}\right|} < 2k-3,$$

unless $min(|Y|, |Z|) = 1$ or $Y = \{x, x^{-1}\}$ for some letter x.

To simplify the notation we denote $|Y|$, $|Z|$ by y and z, respectively. Note that $y + z = 2k-1$. The assumptions are hence $y, z \geq 2$ and $Y \neq \{x, x^{-1}\}$. The main

technique here is to show that each letter from $X^{\pm 1}$ appearing in a word from $A_{Y,Z,v}$ is followed by one of less than $(2k-3)^r$ possible r-tuples of letters, for some $r \geq 1$.

Case 1: $|Y|, |Z| \geq 4$: Let $[w] \in A_{Y,Z,v}$ and let $x \in X^{\pm 1}$ appear in w. If $x \in Y$ then either the inverse of the following letter is v or it belongs to $Y \setminus \{x\}$, so there are y options, and $y = 2k - 1 - z \leq 2k - 5$. The same argument applies for $x \in Z$. Finally, the only letter that cannot follow v is v^{-1}, so v is followed by:

- one of y letters from Y, which are followed in turn by one of y letters, or
- one of $z-1$ letters from $Z \setminus \{v^{-1}\}$, which are followed in turn by one of z letters, or
- v, which is followed by one of $2k-1$ letter.

Overall, there are $y^2 + z\,(z-1) + (2k-1)$ possibilities for the two letters following v. It is easy to see that under the assumptions of the lemma, this expression is largest when $y = 2k-5$ and $z = 4$. But even in this case,

$$y^2 + z\,(z-1) + (2k-1) = (2k-5)^2 + 12 + (2k-1) < (2k-3)^2\,,$$

(note that if $y, z \geq 4$ then $k \geq 5$). This completes the proof.

Case 2: $|Y| = 3$: Assume first that $k \geq 5$, so $2k - 3 \geq 7$. As in the previous proof, let $x \in X^{\pm 1}$ be a letter in a word from $A_{Y,Z,v}$. If $x \in Y$, the following letter is one of $y = 3$ possibilities. If $x \in Z$ there are at most $z = 2k-4$ possibilities. Finally, if $x = v$, then v is followed by

- one of 3 letters from Y, which are followed in turn by one of 3 letters, or
- one of $2k-5$ letters from $Z \setminus \{v^{-1}\}$, which are followed in turn by one of $2k-4$ letters, or

- v, which is followed by one of $2k-1$ letter.

Overall, there are $3^2 + (2k-5)(2k-4) + (2k-1)$ possibilities for the two letters following v. This is strictly less than $(2k-3)^2$ for $k \geq 5$.

Suppose next that $k = 4$, so now $2k - 3 = 5$. Any $x \in X^{\pm 1} \setminus \{v\}$ is followed by at most 4 possible letters. As for v itself, we need to distinguish between two cases: either Y does not contain a letter and its inverse, in which case w.l.o.g. $Y = \{a, b, c\}$ and $Z = \{a^{-1}, b^{-1}, c^{-1}, v^{-1}\}$; or w.l.o.g. $Y = \{a, a^{-1}, b\}$ and $Z = \{b^{-1}, v^{-1}, c^{\pm 1}\}$. In the first case an easy computation shows that v is followed by one of 115 possible triplets of letters, and we are done as $115 < 5^3$. In the second case, v is followed by one of 617 possible quadruplets, and $617 < 5^4$.

Finally, if $y = 3$ and $k = 3$ (k cannot be smaller than 3 if $y = 3$), then $2k - 3 = 3$. The partition is, up to name changes, $Y = \{a, a^{-1}, b\}$ and $Z = \{b^{-1}, v^{-1}\}$. Through computation, we have any letter $x \neq v$ is followed by at most 8 possible pairs of letters, and v is followed by at most 17,883 possible 9-tuples (and $17{,}833 < 3^9 = 19{,}683$), alternatively, one can show that exponential growth rate here equals the largest real root of $\lambda^5 - 3\lambda^4 + \lambda^3 - \lambda^2 - \lambda + 7$, which is about 2.68.

Case 3: $|Z| = 3$ Assume first that $k \geq 4$, hence $2k - 3 \geq 5$. If $x \in Y$, it is followed by the inverse of one of $2k-5$ letters from $Y \setminus \{x\}$ or by v^{-1}, or a total of $2k-4$ possibilities. A letter from Z is followed by one of $3 < 2k-3$ letters (v^{-1} and two letters whose inverse belongs to Z).

To analyze the number of possibilities after v, we distinguish between two cases:

1. Assume that Z contains a letter and its inverse, i.e. $Z = \{v^{-1}, a, a^{-1}\}$. We note:

   - Every $x \in Y$, is followed either by one of $2k-5$ letters from $Y \setminus \{x^{-1}\}$ which are then followed by one of $(2k-4)$ letters, or by v^{-1} which is followed by one of 3 possible letters: a total of $(2k-5)(2k-4) + 3 = 4k^2 - 18k + 23$

possible pairs.

- The letter a is followed either by one of two letters from Z (a or v^{-1}), then followed by one of 3 possible letters, or by v which is followed by one of $2k-1$ letters: a total of $2 \cdot 3 + (2k-1) = 2k+5$ possible pairs. The same computation holds for a^{-1}.
- Every pair of letters following v starts either with some $x \in Z$ ($2 \cdot 3$ options), by v ($2k-1$ options) or by some $x \in Y$ ($(2k-4)^2$ options): a total of $6 + 2k - 1 + (2k-4)^2 = 4k^2 - 14k + 21$ possible pairs.

We can now count the number of possible triplets of letters following v:

- $2\,(2k+5)$ triplets begin with a or a^{-1}.
- $(2k-4)\,(4k^2-18k+23)$ triplets begin with some $x \in Y$.
- $4k^2 - 14k + 21$ triplets begin with v.

The total number of possible triplets following v is, therefore, $8k^3 - 48k^2 + 108k - 61$ which is strictly less than $(2k-3)^3$ when $k \geq 4$.

2. The other possibility is that Z does not contain a letter and its inverse, hence $Z = \{v^{-1}, a, b\}$. A similar computation shows that v is followed in this case by one of $8k^3 - 56k^2 + 160k - 145$ which is again strictly less than $(2k-3)^3$ when $k \geq 4$.

Finally, if $k = 3$ then $Z = \{v^{-1}, a, b\}$ and $Y = \{a^{-1}, b^{-1}\}$ (for otherwise $Y = \{x, x^{-1}\}$). Another technical computation shows that a^{-1} and b^{-1} are followed by at most 2 possible letters, v^{-1} by at most 7 possible pairs of letters, a and b by at most $237 < 3^5$ 5-tuples, and v by at most $41{,}372{,}449 < 3^{16}$ 16-tuples of letters, alternatively, the exponential growth rate here equals the largest real root of $\lambda^4 - 2\lambda^3 - 4\lambda^2 + 2\lambda + 7$, which is about 2.85.

Case 4: $|Y| = 2$ Since $Y \neq \{x, x^{-1}\}$, assume w.l.o.g. that $Y = \{a, b\}$. The case $k = 3$ was already dealt with in the previous case, so assume $k \geq 4$. Similar calculations to those above show that:

- a and b are followed by one of $2 < (2k-3)$ letters.
- a^{-1} and b^{-1} are followed by one of $(4k^2 - 14k + 16) < (2k-3)^2$ pairs of letters.
- Any letter x such that $x, x^{-1} \in Z$ is followed by one of $(4k^2 - 16k + 21) < (2k-3)^2$ pairs of letters.
- v^{-1} is followed by one of $(4k^2 - 16k + 19) < (2k-3)^2$ pairs of letters.
- v is followed by one of $(8k^3 - 44k^2 + 106k - 91)$ triplets of letters. This is strictly less than $(2k-3)^3$ for $k \geq 5$. For $k = 4$, a concrete computation shows v is followed by one of $613 < 5^4$ possible 4-tuples of letters.

Case 5: $|Z| = 2$ Assume w.l.o.g. that $Z = \{a, v^{-1}\}$. The case $k = 3$ was handled in the case $|Y| = 3$, so assume $k \geq 4$. Again, the following formulas can be easily computed:

- a and v^{-1} are each followed by one of $2 < (2k-3)$ letters.
- a^{-1} is followed by one of $(4k^2 - 14k + 14) < (2k-3)^2$ pairs of letters.
- Any letter x such that $x, x^{-1} \in Y$ is followed by one of $(4k^2 - 16k + 19) < (2k-3)^2$ pairs of letters.
- v is followed by one of $(8k^3 - 40k^2 + 80k - 51)$ triplets of letters. This is strictly less than $(2k-3)^3$ for $k \geq 6$. For $k = 5$, a concrete computation shows v is followed by one of $1{,}951 < 7^4$ possible 4 tuples of letters, and for $k = 4$, v is followed by one of $557 < 5^4$ possible 4-tuples of letters.

This finishes the proof of Proposition 3. $\square$

1.5 Generalizations and Further Questions

Recall $S_{k,N}$ is the set of all words of length N in F_k which are contained in a proper free factor. Since the primitives are contained in a free factor of rank 1, $S_{k,N} \supseteq P_{k,N}$ is a strictly larger set. A natural question along the same vein would be to estimate growth of $S_{k,N}$. Recall the statement of Fact 1 in Section 1.2:

Fact. *The Whitehead graph of any element in $S_{k,N}$ has a cut vertex.*

Our proof of the main theorem also applies to this case and yields the stronger result that $S_{k,N}$ has the same exponential growth rate: $2k-3$.

By the main theorem, most words in any size $k-1$ subset of the letters are non-primitive. We conclude that the number of non-primitive words in $S_{k,N}$ grows at least as fast as $(2k-3)^N$. Hence we know that $S_{k,N}$ is indeed larger than $P_{k,N}$ in a non-negligible way. More precisely, we prove:

Corollary 2. *For $k \geq 3$ we have*

$$\lim_{N\to\infty} \sqrt[N]{|S_{k,N}|} = \lim_{N\to\infty} \sqrt[N]{|S_{k,N} \setminus P_{k,N}|} = 2k-3.$$

Proof. As already mentioned in Section 1.1 that $(2k-3)$ is a lower bound. To show it is also an upper bound, we repeat similar arguments in Section 1.3.1: □

Firstly, the exponential growth of $S_{k,N}$ is the same as the exponential growth of $\overline{S_{k,N}}$, the set of cyclic-words of length N belonging to a proper free factor. By Fact 1, each $[w] \in \overline{S_{k,N}}$ belongs to some $A^N_{Y,Z,v}$. For most triplets, Proposition 3 shows they grow slower than $(2k-3)^N$. When $Z = \{v^{-1}\}$, $A^N_{Y,Z,v} \subseteq \overline{S_{k,N}}$ and grows exponentially with base $(2k-3)$.

Consider next words in $A^N_{Y,Z,v}$ such that either

- $Y = \{x, x^{-1}\}$ and there is exactly one instance of $vx^m v^{-1}$, or

- $Y = \{x\}$ and there are up to three instances of $x^{\pm 1}$.

It is evident that this set of words has exponential growth rate $(2k-3)$.

The remaining words from $\overline{S_{k,N}}$, which we denote by $Q_{k,N}$, can be described in a similar fashion to the cyclic words from $M_{k,N}$. In a similar argument as in Lemma 3, we can shorten each word from $Q_{k,N}$ by the corresponding Whitehead automorphisms until we get a word outside $Q_{k,N}$. Since we have already seen that $\left|\overline{S_{k,N}} \setminus Q_{k,N}\right|$ has exponential growth rate $(2k-3)$, we can complete the proof of Corollary 2 in the same manner we proved Lemma 3. $\square$

In fact, the proof of Corollary 2 alone could be shorter than the proof of Theorem 3. The same shorter proof would show that $\limsup_{N\to\infty} \sqrt[N]{C_{k,N}} = 2k-3$. In other words, much of the complexity of the analysis in Sections 1.3.2 and 1.3.3 is required only for showing the stronger result that the growth rate of $|C_{k,t} \setminus L_{k,t}|$ is strictly smaller than $2k-3$.

Finally, we mention the following closely related questions which are still open:

1. What can be said about the growth of $\mathrm{Aut}\,(F_k)$ with respect to standard generating sets such as
   - Nielsen moves?
   - Whitehead automorphisms?
2. What is the smallest possible exponential growth rate of $P_{k,N}$ (or $C_{k,N}$) with respect to arbitrary finite generating sets of F_k (not necessarily bases)?

Chapter 2

Random Methods in 3-Manifold Theory

2.1 Motivation and Outline

Over the years, random methods have evolved into powerful techniques in several areas of mathematics. Most notably, as pioneered by Paul Erdős [Erd59], the study of random graphs has become an important branch of contemporary graph theory. What is especially fascinating about this development is the fact that such techniques solved many problems which have nothing to do with probability: one can use random constructions to show the existence of a graph satisfying particular properties without constructing explicit examples. For example, graphs with both arbitrarily large girth and arbitrarily large chromatic number were shown to exist by random methods by Erdős [Erd59] long before explicit examples were found by Lovász [Lov68] and Lubotzky, Phillips and Sarnak [LPS88].

The goal of this paper is to present similar methods within the world of 3-manifolds. In 2006, Dunfield and W. Thurston [DT06a] presented a model of 'random' 3-manifolds by considering random walks on the mapping class group, and a theory

of random 3-manifolds is starting to emerge ([DT06b], [Riv08], [Kow08], [Mah10], [DW11], [Mah11], [Riv14]). Here we will use this theory to prove the following existence result, which a priori has nothing to do with randomness.

Theorem. *For any integers k and g with $g \geq 2$, there exist infinitely many closed hyperbolic 3-manifolds which are integral homology 3-spheres with Casson invariant k and Heegaard genus equal to g.*

In [Riv14]*Theorem 8.9, Rivin investigates the Casson invariant for a different model of random 3-manifolds, random mapping tori, where the random walk on the mapping class group is used as the monodromy map for a 3-manifold fibering over the circle. In particular, by doing Dehn fillings on cusped random mapping tori, he produces sets of manifolds with Casson invariants $\{N, N+1, N+2, \dots\}$, for some N depending on the cusped manifold. As the manifolds are produced by filling a cusped manifold, their volume is bounded. In contrast, results announced by Brock and Souto [BS09] would show that the volume of the 3-manifolds produced in the proof of Theorem 2.1 tend to infinity.

There is however a difference between our methodology and the common practice of random methods in graph theory. In graph theory, as said, existence is usually shown by proving that "most" objects satisfy the desired property, even though no explicit examples are given. Here our considerations will be somewhat more delicate: we will have to compare rates of decay of various properties along random walks, and the difference between them will ensure the existence of the desired manifolds. We hope that our results will be an initial example of applying random methods to 3-manifolds.

We now give a brief outline of the proof. From well-known results about random walks on $\mathbb{Z}$, one deduces that for any fixed n the probability of hitting n decays polynomially with respect to the length of the walk. Since the non-hyperbolic and Heegaard genus $< g$ decay exponentially, we conclude that there must exist hyperbolic

elements in H corresponding to Heegaard genus g homology 3-spheres with Casson invariant n. This finishes the proof.

The paper is organized as follows: Section 2.2 reduces the main theorem to Theorem 5 which claims that the set of $h \in H$ inducing 3-manifolds with Heegaard genus strictly less than g is exponentially small. Section 2.3 then gives background on the curve complex and reformulates Theorem 5 in terms of Heegaard splitting distances resulting in Theorem 6. The section then went on to divide the proof of Theorem 6 into two cases: Far from disc set and close to the disc set. Sections 2.4, 2.5 and 2.6 provides tools and treats the two cases separately. The two propositions are then put together in Section 2.7 into a complete proof of Theorem 6

2.2 Applying Random Method

Before starting the proof we give some background and fix notation.

Let S_g be an embedded copy of the closed genus g surface in the standard 3-sphere (such embedding is unique up to homotopy); separating the sphere into two genus g handlebodies H_1 and H_2.

Since all handlebodies with the same genus are homeomorphic, we find homeomorphism

$$\varphi : H_2 \to H_1$$

which induces an automorphism

$$\partial\varphi : \partial H_2 \to \partial H_1$$

by taking the extension and restriction to the boundary.

Composing with the restriction of the identity map in S^3

$$i : \partial H_1 \to \partial H_2,$$

we obtain automorphism

$$\partial\varphi^{-1} \circ i : \partial H_1 \to \partial H_1$$

where $\partial H_1 = S_g$.

Let $MCG(S_g)$ be the mapping class group of S_g, let $h_{S^3} \in MCG(S_g)$ be the isotopy class of $\partial\varphi^{-1} \circ i$.

For each element ϕ in $MCG(S_g)$, we may consider the map $\phi \circ h_{S^3}$, it defines a gluing of the two handlebodies and induces a 3-manifold M_ϕ with a genus g Heegaard splitting. Note that we compose with h_{S^3} so that the identity element in $MCG(S_g)$ would correspond to the standard 3-sphere.

It is easy to see that M_h is a homology sphere if and only if h is in the Torelli subgroup $\mathcal{T}$ of the mapping class group.

For any finitely generated subgroup $G < MCG(S_g)$ with a finite set S of generators, consider the random walk on the Cayley graph with equal probability of going along each element in $S \cup S^{-1}$. Let w_N denote the group element corresponding to Nth step of the random walk. Taking the probability of the random walk landing in the set gives rise to a way of measuring the size of subsets of our group.

Given $h \in MCG(S_g)$ pseudo Anosov, let $\mathcal{L}^s(h)$ $\mathcal{L}^u(h)$ denote the stable and unstable laminations of h.

Definition 1. A finitely generated subgroup G of $MCG(S_g)$ is said to be **sufficiently large** if it contains two pseudo Anosov elements $\varphi,\ \psi$ with distinct stable and unstable laminations, namely

$$\{\mathcal{L}^s(\varphi),\ \mathcal{L}^u(\varphi)\} \cap \{\mathcal{L}^s(\psi),\ \mathcal{L}^u(\psi)\} = \phi.$$

Fixing a finite generating set, we say that a subset $E \subseteq G$ is **exponentially small** if there exists constants $C, c > 0$ such that for every $N \in \mathbb{N}$,

$$P(w_N \in E) \leq Ce^{-cN}.$$

We will call a subset of G **exponentially large** if its complement is exponentially small.

Let $\mathcal{K}$ be the subgroup of $\mathcal{T}$ generated by Dehn twists along separating curves.

In fact, as discussed in Chapter 6 of [FM12], the group $\mathcal{K}$ coincides with the second Torelli group, $\mathcal{T}^2(S_g)$, defined as the kernel of the action of $\mathcal{T}(S)$ on the quotient $\Gamma/[\Gamma, \Gamma']$ where $\Gamma = \pi_1(S_g)$ and $\Gamma' = [\Gamma, \Gamma]$ is the commutator subgroup.

We have:

Fact 2. *(Morita [?]) The Casson invariant $\lambda : \mathcal{K} \to \mathbb{Z}$ is a homomorphism.*

Claim 1. $\mathcal{K}$ is sufficiently large.

All we need to do is exhibit a pair of pseudo Anosov elements in $\mathcal{K}$ with distinct stable and unstable laminations. To this end let us define:

Definition 2. A collection of curves is said to **fill** the surface if the complementary region of their union is a set of discs.

By Theorem 3.1 in [PH92], for any pair of simple closed geodesics that fill the surface, the composition ϕ of Dehn twists along them is pseudo Anosov.

Hence we need a pair of separating simple curves which fill the surface. It is not hard to find such pair, for example, the curves in Figure 2.2 below works for any odd $g \geq 2$.

For g even we replace the second curve by the one in Figure 2.2.

Thus, with appropriate choice of orientation, $\varphi \in \mathcal{K}$ - the composition of Dehn twists τ_α, τ_β along these curves, is pseudo Anosov.

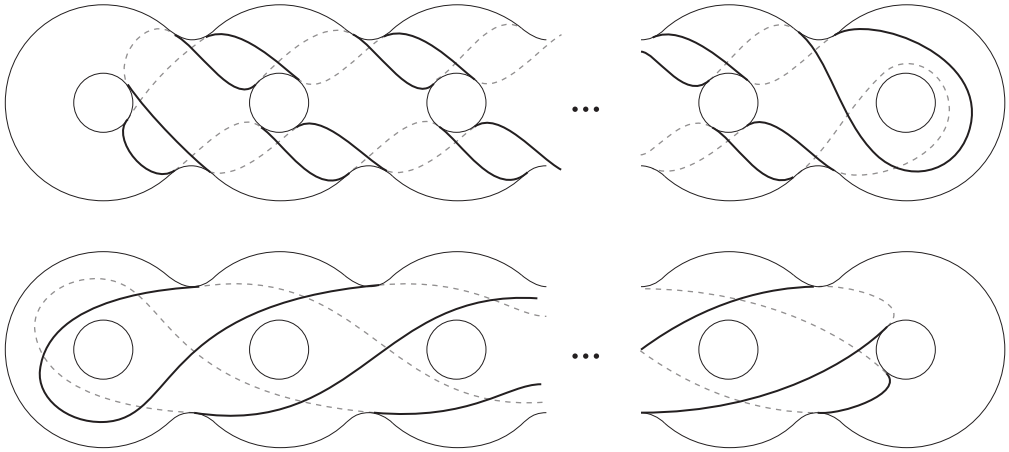


Figure 2.1: Filling pair of separating curves

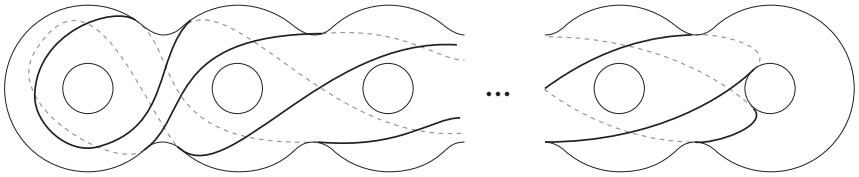


Figure 2.2: Second curve for even genus

Furthermore, following the steps of producing stable and unstable laminations for Dehn twists along filling curves as presented in section 13.3 of [FLP12], one can easily see that there exist integers a, b such that $\psi = \tau_\alpha^a \circ \tau_\beta^b$ is again pseudo Anosov and has stable and unstable laminations different from that of φ. Establishes the claim.

Remark 1. One can prove Claim 1 more abstractly: As $\mathcal{K}$ is normal in $\mathcal{T}$ and $\mathcal{T}/\mathcal{K}$ is abelian, it can be deduced that $\mathcal{K}$ is sufficiently large just like $\mathcal{T}$.

Lemma 4. $\lambda : \mathcal{K} \to \mathbb{Z}$ *is surjective.*

Proof. By Section 3.2.8 of [Sav02], for Dehn surgeries on the trefoil knot κ we have:

$$|\lambda(S^3 + \frac{1}{m+1}\kappa) - \lambda(S^3 + \frac{1}{m}\kappa)| = 1$$

independent of the integer m.

Now we can embed κ as a separating curve C in the closed genus two surface as follows:

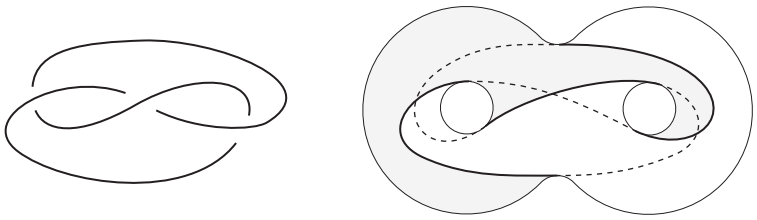


Figure 2.3: Embedded separating trefoil

It is a standard fact (see, for example, 3.2.5 of [Sav02] that the $S^3 + \frac{1}{m}\kappa$ is the same as taking the standard embedding of S_2 in S^3 and gluing the two handlebodies with m Dehn twists along the embedded copy of κ.

Let $\tau \in \mathcal{K}$ be the Dehn twist along C, so we have

$$|\lambda(\tau^2) - \lambda(\tau)| = 1.$$

This finishes the proof when $g = 2$.

Also, Casson invariant is additive under connected sums, hence for any $g > 2$ we may simply add handles on the standard embedding of S_2 away from the embedding of k and obtain the same manifold by applying m Dehn twists along κ.

Since $\mathcal{K}$ is the group generated by Dehn twists along separating curves in S_g, we deduce that for all m, $S^3 + \frac{1}{m}\kappa$ are manifolds obtained as M_g for some $g \in \mathcal{K}$. We conclude there are always consecutive integers in the image $\lambda(\mathcal{K})$, i.e. $\lambda|_{\mathcal{K}}$ is surjective. □

Corollary 3. *$\lambda : H \to \mathbb{Z}$ is surjective for any subgroup $H < \mathcal{K}$ with $\tau \in H$.*

Now let H be any finitely generated sufficiently large subgroup of $MCG(S_g)$, the following two results in [Mah10] apply:

Fact 3. *The set of elements of H inducing hyperbolic 3-manifolds is exponentially large in H.*

and

Fact 4. *The set of pseudo Anosov elements is exponentially large in H.*

We denote the above sets by H_{hyp} and H_{pA}, respectively. We also need such an estimate on the set H_g of elements in H that induce 3-manifolds with Heegaard genus g. For this we will prove:

Theorem 5. *There exists finitely generated sufficiently large subgroup $H < \mathcal{K}$ such that the set H_g of elements in H that induce 3-manifolds with Heegaard genus g is exponentially large in H.*

In [Mah10], Maher proved that $\lim_{N\to\infty} P(w_N \in H_g) = 1$, but we need the stronger version claiming that H_g is in fact exponentially large. We will spend the rest of our paper on the proof of Theorem 5.

Now we are ready to put the parts together and obtain the main theorem.

Proof. (of main theorem) Let H be the subgroup produced by Theorem 5, without loss of generality suppose $\tau \in H$.

Suppose some $n \in \mathbb{Z}$ is not achieved as the Casson invariant of a hyperbolic homology sphere with Heegaard genus g, then in particular we have

$$\lambda^{-1}(n) \cap H \subseteq (H_{\text{hyp}})^c \cup (H_g)^c.$$

Since H satisfies the conditions of Fact 3, we know H_{hyp} is exponentially small. H_g is exponentially small by construction.

Hence we deduce that the level set $\lambda^{-1}(n)$ is exponentially small in H.

Now since λ is a homomorphism on H, it projects the random walk on H to a Markov process on $\mathbb{Z}$, in particular, at step L,

$$P(\pi_\lambda(w_L) = n) > c/\sqrt{L}$$

with some small constant c, (see for example section 2.1 of [LL10]. This contradicts the assumption of $\lambda^{-1}(n)$ being exponentially small. $\square$

2.3 Heegaard splitting distance

Let S_g be the complete orientable surface of genus g. The **curve complex**, $C(S_g)$ is a simplicial complex defined as follows:

1. The vertex set $V(C(S_g)) := \{\text{Simple closed curves on } S_g\} /\text{isotopy}$
2. There is a simplex spanning $v_1, \cdots, v_k$ iff there exists mutually disjoint curves representing isotopy classes $v_1, \cdots, v_k$.

Fix $S_g = \partial H$, The **disc set** $\mathcal{D}$ is defined as set of all vertices in $C(S_g)$ corresponding to isotopy classes of simple closed curves which bounds discs in H.

The **Handlebody group** is a subgroup of $MCG(S_g)$ consisting of isotopy classes of all automorphisms of S_g which extends to automorphisms of H.

Given sets $A, B \subseteq C(S_g)$, let the distance

$$d(A, B) = \min \{d(a, b \mid a \in A, b \in B\}.$$

For any Heegaard splitting $M = H_1 \sqcup S_g \sqcup H_2$, since H_1 and H_2 are homeomorphic, we have $h \in MCG\{S_g\}$ with $h(\mathcal{D}_1) = \mathcal{D}_2$, h is unique up to composition with elements of the handlebody group.

The **Heegaard splitting distance** of splitting $H_1 \sqcup S_g \sqcup H_2$ is then simply the distance between two disc sets:

$$d_{sp}(H_1 \sqcup S_g \sqcup H_2) := d(\mathcal{D}_1, \mathcal{D}_2) = d(\mathcal{D}, h(\mathcal{D})).$$

With the identification above, any element $\phi \in MCG(S_g)$ gives raise to Heegaard splitting with gluing map $\phi \circ h_{S^3}$, resulting in 3-manifold M_ϕ. The Heegaard splitting distance is then

$$d_{sp}(M_\phi) = d(\mathcal{D}, \phi \circ h_{S^3}(\mathcal{D})).$$

One can derive many interesting properties of the 3-manifold by consider the Heegaard splitting distance, we shall state the following facts:

Fact 5. *(Kobayashi[Kob88] and Hempel [Hem76]) If the Heegaard splitting distance is > 2 then the corresponding 3-manifold is irreducible, atoroidal and not Seifert-fibered.*

Furthermore, it follows from Thurston's geometrization conjecture that

Fact 6. *(Perelman [MT07]) If the Heegaard splitting distance is > 2 then the corresponding 3-manifold is hyperbolic.*

Most relevant to our following proof is:

Fact 7. *(Scharleman-Tomova [ST06]) If the Heegaard splitting distance is $> 2g$ for some genus gHeegaard splitting, then the Heegaard genus of the 3-manifold is equal to g.*

We now consider Heegaard splitting distance of splittings induced by random elements in the mapping class group, in [Mah10], Maher proved that it grows at least linearly with respect to the number of steps of the random walk, namely:

Fact 8. *(Maher) For any finitely generated complete subgroup of $MCG(S_g)$ with random walk w_n, there exists $l > 0$ such that*

$$\lim_{N\to\infty} \mathbb{P}(d_{sp}(M(w_N)) \geq l \cdot N) = 1$$

However, in order to utilize Fact 7 and conclude that the set of manifolds with Heegaard genus equal to $g$is exponentially large, we need a stronger version of Fact 8, namely:

Theorem 6. *For any sufficiently large complete subgroup $G < MCG(S_g)$, there exists sufficiently large finitely generated subgroup $H < G$ with random walk w_n and numbers $K, l > 0$ and $c < 1$ such that*

$$\mathbb{P}(d_{sp}(M(w_n)) \leqslant l \cdot n) \leqslant Kc^n.$$

We call such random variables $(w_n)_{n\in\mathbb{N}}$ **grow linearly with exponential decay**.

Theorem 5 now easily follows from Theorem 6:

Proof. Since $\mathcal{K}$ is normal and sufficiently large, it is complete.

Apply Theorem 6 we obtain finitely generated H where the Heegaard splitting distance grow linearly with exponential decay.

Pick $N > \frac{2g}{l}$, for all $n > N$ we have

$$\mathbb{P}(d_{\mathrm{sp}}(M(w_n)) \leqslant 2g) \leqslant \mathbb{P}(d_{\mathrm{sp}}(M(w_n)) \leqslant l \cdot n) \leqslant Kc^n$$

is exponentially small.

By Fact 7 we conclude H_g^c is exponentially small in H. □

Let us now sketch the proof of Theorem 6. Recall that the splitting distance is merely the distance in $C(S_g)$ between the disc set $\mathcal{D}$ and its 'image under translation'

of the mapping class group element composed with some standard gluing map h_{S^3} resulting in the 3-sphere. Since h_{S^3} does not depend on the random walk, the key of estimating the growth in Heegaard splitting distance would be to analyze how does a random walk translate $\mathcal{D}$.

A key intermediate step would be to look at the trejecture of a single point under the random walk and estimate the growth of its distance to $\mathcal{D}$. We want the point to leave $\mathcal{D}$ fast enough. More precisely:

Fix a vertex $x_0 \in C(S_g)$ as basepoint, let $H < \mathcal{K}$ be a finitely generated subgroup, consider image of the random walk on H under the orbit map $w_n \mapsto w_n x_0$. i.e. we let the point x_0 travel in $C(S_g)$ by following path generated by the random walk in H.

Proposition 6. *For a suitably chosen H, the distance from $w_n x_0$ to the disc set, $d(\mathcal{D}, w_n x_0)$, grows linearly with exponential decay. i.e.*

$$\mathbb{P}(d(\mathcal{D}, w_n x_0) < l \cdot n) < K c^n$$

for some $K, l > 0$ and $c < 1$.

We now give an outline of the argument for this result, omitting certain technical details.

Let $\mathbb{N}_0$ be the set of non-negative integers.

Consider the sequence of random variables $X_n = d(\mathcal{D}, w_n x_0)$ with values in $\mathbb{N}_0$. This is not a Markov chain, but the conditional probabilities

$$\mathbb{P}(X_{n+1} = j \mid w_n = h)$$

are well-defined for any $h \in MCG(S_g)$.

We will break the analysis into two cases:

There exists some large M such that, for all $n \in \mathbb{N}_0$:

1. If hx_0 lies inside the disc set at the n-th step, then there is a definite chance $\varepsilon > 0$ where after M further steps x will be at least R away from the disc set with probability ε, namely

$$\mathbb{P}(d(w_{n+M}x_0, \mathcal{D}) \geq R \mid w_n = h \text{ with } hx_0 \in \mathcal{D}) \geqslant \varepsilon.$$

2. If hx_0 is some distance $T > 0$ away from $\mathcal{D}$, then the probability that after M steps x_0 moved exactly distance $r$closer to the disc set decays exponentially in r, namely there exists $q < 1$ where

$$\mathbb{P}(d(w_{n+M}x_0, \mathcal{D}) = T - R \mid w_n = h \text{ with } d(\mathcal{D}, hx_0) = T) \leqslant q^{R+1}.$$

The moral of the above breakdown being that, if x_0 is inside $\mathcal{D}$ at time n, it will leave $\mathcal{D}$ and be at a certain distance with fixed probability; once x_0 is far from $\mathcal{D}$ the distance stays large.

Now we briefly indicate how those estimates imply X_n grows linear with exponential decay:

The sequence of random variables $\{X_n\}$ does not arise from a Markov chain, but we can compare the distributions of the $\{X_n\}$ with the distributions $\{Y_n\}$ arising from a Markov chain on $\mathbb{N}_0$, which never increases by more than one unit per step, and has transition probabilities given by $p(0,0) = 1 - \varepsilon$, and $p(t, t-r) = q^{r+1}$, for $r \geqslant 0$. The first few vertices of this Markov chain are illustrated in Figure 2.4.

Intuitively, it is always more likely that the random variables $\{X_n\}$ move further to the right than the random variables $\{Y_n\}$ arising from the Markov chain.

This means that the distribution of each Y_n will have greater weight on small values than the distribution for the corresponding X_n, or more precisely $F_{X_n}(t) \leqslant F_{Y_n}(t)$ for all t, where F_{X_n} and F_{Y_n} are the cumulative probability functions for X_n and Y_n

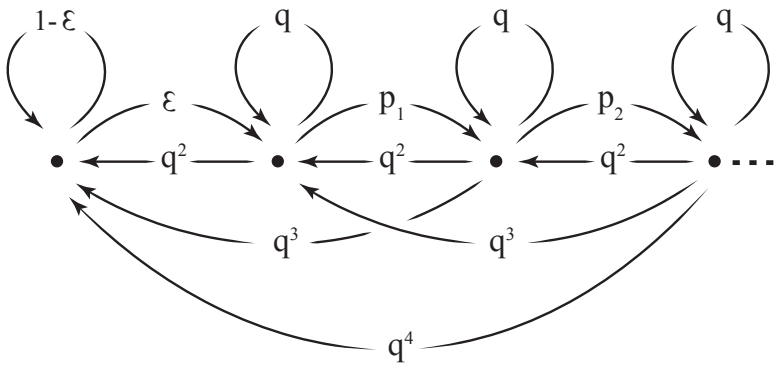


Figure 2.4: The Markov chain $(\mathbb{N}_0, P)$, where $p_i = 1 - q - q^2 - \cdots - q^{i+1}$.

respectively. This property is usually described by saying that the random variables X_n *stochastically dominate* the random variables Y_n, written as $Y_n \prec X_n$.

A standard argument from the theory of Markov chains shows that the random variables Y_n arising from the Markov chain satisfies the linear progress with exponential decay property that we require, and so this implies the linear progress with exponential decay property for the X_n.

Note that the above two properties are approximate versions of what we are going to prove, for technical reasons it is more convenient to work with a coarse version of distance, in which we choose X_n to be the integer part of $d(\mathcal{D}, w_n x_0)/K$, for some reasonably large K, rather than $d(\mathcal{D}, w_n x_0)$ itself.

The two properties are shown using different sets of tools, we first address present some backgrounds and results in corse geometry in Section 2.4 which will prepare us for addressing the second property in Section 2.5. Then we present some additional background on train trucks and prove the first property in Section 2.6. Finally, Section 2.7 applies the two estimates to show that $d(\mathcal{D}, w_n x_0)$ grows linearly, with exponential decay and then passing from individual points to splitting distances and establish Theorem 6.

2.4 Coarse geometry

Recall that a geodesic metric space (X, d) is **δ-hyperbolic** if for any triple of points (x, y, z), the geodesic connecting y and z is contained in the δ-neighbourhood of the other two geodesics.

Fix δ-hyperbolic space (X, d), the **Gromov product** of two points y and z in X, with respect to a basepoint $x \in X$, is defined to be

$$(y \cdot z)_x = \tfrac{1}{2}(d(x, y) + d(x, z) - d(y, z)).$$

We can think of the Gromov product as roughly equal to the distance from x to a geodesic from y to z. Infact this is true up to an additive constant.

We start with an elementary observation concerning distances in δ-hyperbolic spaces. We say two geodesics starting at the same point **fellow-travel** for a distance T if corresponding points on the geodesics are less than 2δ away for any $t \in [0, T]$. By definition of δ-hyperbolicity, any two geodesics γ_{xy}, γ_{yz} would fellow-travel for some distance near y. The total distance $d(x, z)$ is equal to $d(x, y) + d(y, z)$ minus approximately twice this overlap. Hence the fellow-travel distance is approximately the Gromov product $(x \cdot z)_y$.

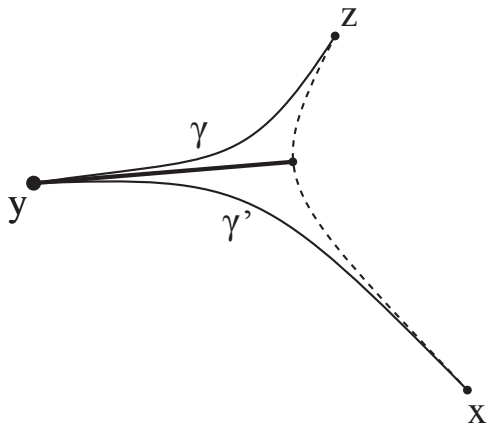


Figure 2.5: Overlapping geodesics

Re-stating the above we have

$$d(x,z) \approx d(x,y) + d(y,z) - 2(x \cdot z)_y,$$

where approximate means up to a universal additive constant involving δ, for simplicity we will write '=' instead from now on.

A set $K \subset X$ is said to be q-**quasiconvex** if any geodesic connecting two points of K is contained in the q-neighbourhood of K.

We now produce similar estimates for the case in which the point x is replaced by a quasiconvex set, and also for the case in which both the points x and z are replaced with quasiconvex sets, as illustrated in Figure 2.6, where x is the closest point on K_x to y and z is the closest point on K_z to y.

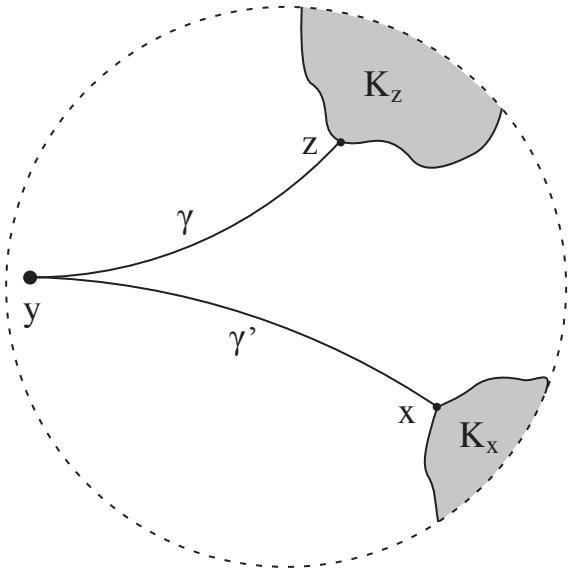


Figure 2.6: Estimating distances between quasiconvex sets and points.

First, when we replace x with a quasiconvex set K, a version of the above works in producing a lower bound on $d(K,z)$, with an additional hypothesis on $(x \cdot z)_y$. Furthermore the Gromov product based at y of z with any point in K is equal to $(x \cdot z)_y$ up to bounded additive error depending only on δ and q. More precisely,

Proposition 7. *For (X, d) δ-hyperbolic, given a number q, there are numbers C_1, C_2 and C_3, depending on δ and q, such that if K is a q-quasiconvex set, and x is a closest point in K to y, then for any point z with*

$$(x \cdot z)_y \leqslant d(x, y) - C_1, \tag{2.1}$$

then

$$d(K, z) \geqslant d(x, y) + d(y, z) - 2(x \cdot z)_y - C_2. \tag{2.2}$$

Furthermore, for any point $x' \in K$,

$$|(x \cdot z)_y - (x' \cdot z)_y| \leqslant C_3. \tag{2.3}$$

Next, we replace both x and z with quasiconvex sets K_x and K_z, similar to the result above, we require the follow-travel distance $(x \cdot z)_y$ to be a bounded constant shorter than both geodesics. We show that the Gromov product based at y of any point in K_x with any point in K_z, is equal to $(x \cdot z)_y$, again up to bounded additive error depending only on δ and q, namely:

Proposition 8. *For (X, d) δ-hyperbolic, given a number q, there exists C_1, C_2 and C_3 depending only on δ and q, such that if K_x and K_z are q-quasiconvex sets and x is a closest point in K_x to y; z is a closest point in K_z to y, then if*

$$(x \cdot z)_y \leqslant \min\{d(x, y), d(y, z)\} - C_1, \tag{2.4}$$

then

$$d(K_x, K_z) \geqslant d(x, y) + d(y, z) - 2(x \cdot z)_y - C_2. \tag{2.5}$$

Furthermore, for any points $x' \in K_x$ and $z' \in K_z$,

$$|(x \cdot z)_y - (x' \cdot z')_y| \leqslant C_3. \tag{2.6}$$

Recall the following result regarding approximate trees in δ-hyperbolic spaces, see for example Ghys and de la Harpe [**?**]Section 2.2.

Fact 9. *[Approximate tree] Let (X, d) be a δ-hyperbolic space, there is a number C_T depending only on δ, such that for any finite collection of points $x_1, \ldots, x_n$, there is a geodesic tree T in X containing all x_is, such that*

$$d_T(x, y) - C_T \cdot n \leqslant d(x, y) \leqslant d_T(x, y) \tag{2.7}$$

for all points x and y in T. (Here d_T is the distance inside the tree T.) We call T the approximate tree *determined by the x_is.*

Note that any tree is automatically a 0-hyperbolic space. As the Gromov product is defined in terms of distances, if we write $(y \cdot z)_x^T$ for the Gromov product in the tree T with vertices $\{x_1, \ldots x_n\}$, then we have

$$\left|(y \cdot z)_x^T - (y \cdot z)_x\right| \leq 3C_T \cdot n/2,$$

which only depends on δ and n, where n is the number of vertices in T.

Recall that any three distinct points in a tree determine a unique center, namely the unique point that lies in the intersection of the three geodesics connecting the three possible pairs of points. It follows from the definition of the Gromov product that for any points x, y and z in a tree T, if v is the center for x, y and z, then

$$(x \cdot y)_z^T = d_T(z, v).$$

Let I be a connected subinterval of $\mathbb{R}$. We say a path $\gamma\colon I \to X$ is a (K, c)-**quasigeodesic** if

$$\tfrac{1}{K}\,|x-y| - c \leqslant d(\gamma(x), \gamma(y)) \leqslant K\,|x-y| + c,$$

for all x and y in $\mathbb{R}$.

In a δ-hyperbolic space, quasigeodesics are contained in bounded neighborhoods of geodesics. This is often known as the Morse Lemma, see for example Bridson and Haefligger [BH99]*Theorem III.H.1.7.

Fact 10. *Let γ be a (K, c)-quasigeodesic in a δ-hyperbolic space. Then there is a constant L, depending only on K, c and δ, such that γ is contained in an L-neighbourhood of the geodesic connecting its endpoints.*

As geodesics in the approximate tree T are $(1, C_T n)$-quasigeodesics, any geodesic in T is contained in an L_n-neighbourhood of the geodesic in X connecting its endpoints, for some number L_n, depending only on δ and n.

Proposition 9. *Let (X, d) be a δ-hyperbolic space, let K be a q-quasiconvex set. $x, y, z \in X$ with x being the closest point in K to y, and let x' be an arbitrary point in K. Let T be an approximate tree determined by a set of n points, which include x, x' and y, and let v be the center of x, x' and y in T. Then there is a number C, which only depends on δ, q and n, such that*

$$d(x, v) \leqslant C. \tag{2.8}$$

Although the constant C depends on the number of points determining the approximate tree T, we will only consider approximate trees with $n \leq 5$.

Proof. (of Proposition 9.) Let v be the center of x, x' and y in T. In particular, the minimal subtree in T containing x, x' and y also contains v, and has the configuration

illustrated in Figure 2.7, possibly with edges of zero length.

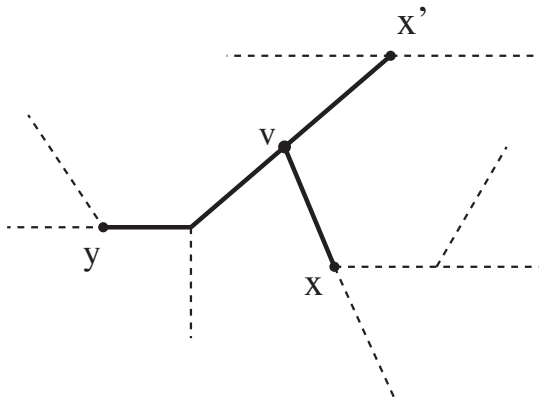


Figure 2.7: A minimal subtree containing x, x' and y, with center v.

By Fact 10, since v lies on the geodesic in T connecting x and x', v is within a bounded distance L_n of the geodesic in X connecting x and x'.

Since K is q-quasiconvex, the geodesic in X from x to x' is contained in a q-neighbourhood of K, hence

$$d(K, v) \leqslant q + L_n. \tag{2.9}$$

As x is a closest point in K to y, $d(x, y) = d(K, y)$, the triangle inequality implies

$$d(x, y) \leqslant d(K, v) + d(v, y).$$

By (2.9) , this implies

$$d(x, y) \leqslant q + L_n + d(v, y).$$

We know metric in X is coarsely equivalent to the metric in T, (2.7) implies

$$d_T(x, y) - C_T \cdot n \leqslant q + L_n + d_T(v, y).$$

Since T is a tree,

$$d_T(x, y) = d_T(x, v) + d_T(v, y).$$

Let $C = q + L_n + C_T \cdot n$, we have

$$d(x, v) \leqslant d_T(x, v) \leqslant C, \tag{2.10}$$

i.e. the distance from x to v in X is bounded by a number which only depends on δ, q and n as required. □

We now prove Proposition 7, recall the proposition states that we can find C_1, C_2 and C_3, depending on δ and q such that for all q-quasiconvex set K with $x \in K$ being a closest point to y, for any z with $(x \cdot z)_y \leqslant d(x, y) - C_1$ we have $d(K, z) \geqslant d(x, y) + d(y, z) - 2(x \cdot z)_y - C_2$. Moreover, for any $x' \in K$ we have $|(x \cdot z)_y - (x' \cdot z')_y| \leqslant C_3$.

Proof. Let T be an approximate tree containing the points x, x', y and z, and let v be the center in T for x, x' and y, let w be the center in T for x, y and z. The set K and the points x, x' and y satisfy the conditions of Proposition 9 hence $d(x, v) \leqslant C$, where C depends only on δ, q and n, in this case $n = 4$.

We choose

$$C_1 = 6C_T + C.$$

In order to obtain (2.1) we need the centers v and w occur in that order along the geodesic from x to y, so the approximate tree T has the configuration illustrated in Figure 2.8.

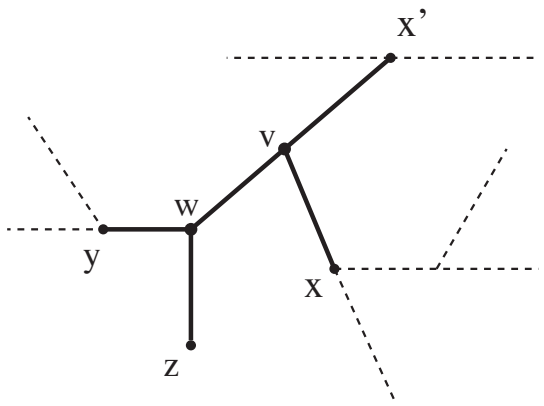


Figure 2.8: An approximate tree containing x, x', y and z.

As the metric in T is equal to the metric in X up to bounded additive error, condition (2.1) is equivalent to

$$(x \cdot z)_y^T \leqslant d_T(x, y) - C_1 + 6C_T.$$

In T, the Gromov product $(x \cdot z)_y^T$ is equal to $d_T(y, w)$, and the center v lies on the geodesic connecting x to y, so

$$d_T(w, y) \leqslant d_T(x, v) + d_T(v, y) - C_1 + 6C_T.$$

As $d(x, v) \leqslant C$, this implies

$$d_T(w, y) \leqslant d_T(v, y) - C_1 + 6C_T + C.$$

Substituting $C_1 = 6C_T + C$, we have

$$d_T(w, y) \leqslant d_T(v, y),$$

hence indeed the centers v and w must lie in that order along the geodesic from x to

y.

The constant C_T depends only on δ and C depends on δ, q and $n = 4$. Hence $C_1 = 6C_T + C$ only depends on δ and q, as required.

In order to show (2.2), let x' be the closest point in K to z. Then, given the configuration of the approximate tree shown in Figure 2.8,

$$d_T(x', z) \geqslant d_T(v, z).$$

Since $d_T(x, v) \leqslant C$,

$$d_T(x', z) \geqslant d_T(x, z) - C,$$

and now using the definition of the Gromov product,

$$d_T(x', z) \geqslant d_T(x, y) + d_T(y, z) - 2(x \cdot z)_y^T - C.$$

Choose

$$C_2 = C + 14C_T,$$

depending only on δ and q.

The metrics in T and X are equal up to additive constants, hence

$$d(K, z) \geqslant d(x, y) + d(y, z) - 2(x \cdot z)_y - C_2.$$

Finally, the configuration of the approximate tree shown in Figure 2.8, shows that the two Gromov products in metric space T, $(x \cdot z)_y^T$ and $(x' \cdot z)_y^T$, are both equal to $d_T(y, w)$. Since the metrics in T and X are equal up to additive error, this shows (2.3) for some constant C_3 which only depends on δ and q, as required. □

Finally, we prove Proposition 8 where both x and z are replaced by quasiconvex sets.

Proof. Consider an approximate tree T containing the points x, x', y, y' and z. Let v be the center in T for x, x' and y, let w be the center in T for z, z' and y, and let u be the center for x, y and z.

The set K_x and the points x, x' and y satisfy the conditions of Proposition 9, so $d(x, v) \leqslant C$ for some C depending on δ, q and n, in this case we have $n = 5$.

Similarly, by applying Proposition 9 to the set K_z and points z, z' and y', we get $d(z, w) \leqslant C$.

Set

$$C_1 = 15C_T/2 + C,$$

i.e. the centers v, u and w occur in that order along the geodesic from x to z and the approximate tree T has the configuration illustrated in Figure 2.9, possibly with some zero length edges.

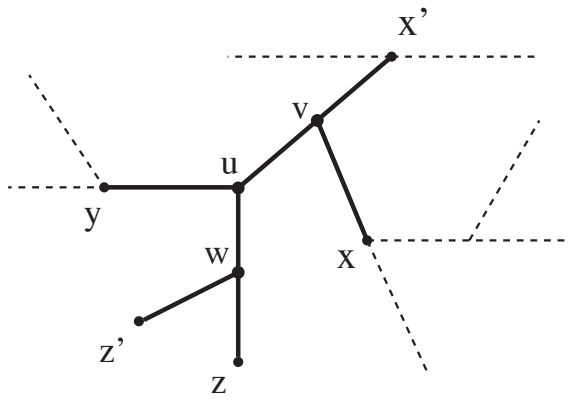


Figure 2.9: An approximate tree for x, x', y, z and z'.

As the metric in T is equal to the metric in X up to bounded additive error, condition (2.4) is equivalent to

$$(x \cdot z)_y^T \leqslant \min\{d_T(x, y), d_T(y, z)\} - C_1 + 15C_T/2.$$

In T, the Gromov product $(x \cdot z)_y^T$ is equal to $d_T(y, u)$, and the center v lies on the

geodesic in T connecting x to y, and the center w lies on the geodesic connecting z to y, so we have

$$d_T(y,u) \leqslant \min\{d_T(x,v) + d_T(v,y), d_T(z,w) + d_T(w,y)\} - C_1 + 15C_T/2.$$

Combining with the fact that $d(x,v) \leqslant C$ and $d(z,w) \leqslant C$, we have

$$d_T(y,u) \leqslant \min\{d_T(v,y), d_T(w,y)\} - C_1 + 15C_T/2 + C.$$

Substituting $C_1 = 15C_T/2 + C$, we get

$$d_T(y,u) \leqslant \min d_T(v,y), d_T(w,y),$$

i.e. the centers v and w must lie further from y than u.

The only way in which this can happen is if the center u lies between v and w, as illustrated in Figure 2.9. Establishing (2.4).

For inequality (2.5), the distance $d(K_x, K_z)$ between the two quasiconvex sets.

Suppose that x' and z' are closest points in K_x and K_z respectively, i.e. $d(x', z') = d(K_x, K_z)$.

In the approximate tree we have $d_T(x', z') \geqslant d_T(x, z)$.

The definition of the Gromov product gives

$$d_T(x', y') \geqslant d_T(x,y) + d_T(y,z) - 2(x \cdot z)_y^T - 2C.$$

Now using Fact 9 to estimate distances in X,

$$d(K_x, K_z) \geqslant d(x,y) + d(y,z) - 2(x \cdot z)_y - 2C - 5C_T.$$

By setting

$$C_2 = 2C + 5C_T$$

which again only depends on δ and q, we establish (2.5).

For the final estimate for the Gromov products, observe that in the approximate tree T we have $(x \cdot z)_y^T = (x' \cdot z')_y^T$. Again by Fact 9, we have

$$|(x \cdot z)_y - (x' \cdot z')_{y'}| \leqslant 15C_T/2.$$

Setting

$$C_3 = 15C_T/2$$

depending on δ and q, would then yield (2.6). □

2.5 Local estimates far from the disc set

In this section we prove the second estimate stated in Section 2.3, namely:

If hx_0 is some distance $T > 0$ away from $\mathcal{D}$, then the probability that after M steps x_0 moved exactly distance $r$closer to the disc set decays exponentially in r, namely there exists $q < 1$ where

$$\mathbb{P}(d(w_{n+M}x_0, \mathcal{D}) = T - R \mid w_n = h \text{ with } d(\mathcal{D}, hx_0) = T) \leqslant q^{R+1}.$$

Definition 3. Given a basepoint $x \in X$, and a number $R \geqslant 0$, the **shadow set** of a point $y \in X$, with parameter R, to be

$$S_x(y, R) = \{z \in X \mid (y \cdot z)_x \geqslant d(x, y) - R\}. \tag{2.11}$$

We are interested in estimating distances from points in the curve complex to a particular disc set $\mathcal{D}$.

Given a positive number R, we define the step function $\phi_R \colon \mathcal{C}(S_g) \to \mathbb{N}_0$ by

$$\phi_R(x) = \lfloor d(\mathcal{D}, x)/R \rfloor,$$

where $\lfloor r \rfloor$ is the largest integer less than or equal to r. We shall think of ϕ_R as a corse distance at scale R.

In particular, a random walk on the mapping class group gives rise to a sequence of random variables $X_n = \phi_R(w_n x_0)$, with values in $\mathbb{N}_0$.

A key proposition would be to show that for X_m sufficiently large, the probability that X_{m+n} is less than $X_m - r$ decays exponentially in r, for all large enough n. As the random walk is a Markov chain on G, it suffices to consider the case of a random walk of length n starting at some fixed point $x_m = g_m x_0$ in $\mathcal{C}(S_g)$. i.e. we reformulate the estimate in Section 2.3:

Proposition 10. *Let G be a finitely generated sufficiently large subgroup of $MCG(S_g)$ with the probability distribution generated by random walk with a fixed set of generators. For any number $q > 0$ there exists numbers R and N depending on q and μ, such that: For all $n \geqslant N$ and $r \geqslant 0$,*

$$\mathbb{P}(\phi_R(g_m w_n x_0) \leqslant t + 1 - r \mid \phi_R(g_m x_0) = t \geqslant 1) \leqslant q^r.$$

Recall that the probability distribution μ depends on the group G, any number which depends on μ implicitly depends on the group G and hence on the coarse geometry constants determined by G.

In order to show this, we make use of Proposition 7, which enables us to estimate the distance between $g_m w_n x_0$ and a quasiconvex set in terms of the distance travelled by the sample path together with an estimate on the size of a particular Gromov product. We will also use the fact that the distance travelled by the sample path in the curve complex $\mathcal{C}(S_g)$ grows linearly in n with exponential decay.

Fact 11. *[Mah12] With the above assumptions, there are numbers $K_\ell, L > 0$ and $c_\ell < 1$ depending on μ, such that:*

For all $n \in \mathbb{N}_0$

$$\mathbb{P}(d(w_n x_0, x_0) \leqslant Ln) \leqslant K_\ell \cdot c_\ell^n$$

In particular, this result holds for $n = 0$, hence we know that $K_\ell \geq 1$.

In order to estimate probabilities involving Gromov products, we will also use the following estimate for the probability a sample path lies in a shadow set, which we shall refer to as exponential decay for shadows.

Fact 12. *[Mah12] Then there are numbers $K_S > 0$ and $c_S < 1$ depending on μ, such that for any point y,*

For all $n \in \mathbb{N}_0$

$$\mathbb{P}(w_n x_0 \in S_{x_0}(y, R)) \leqslant K_S \cdot c_S^{d(x_0,y)-R}.$$

The following Gromov product estimate then follows immediately from the definition of shadow sets (2.11):

$$\mathbb{P}\left(\ (y \cdot w_n x_0)_{x_0} \geqslant d(x_0, y) - R \ \right) \leqslant K c^{d(x_0,y)-R}.$$

To simplify notation, let $K = \max\{K_\ell, K_S\}$ and $c = \max\{c_\ell, c_S\}$.

The proof for Proposition 10 goes roughly as follows: Consider a random walk of length n starting at a point x_0, reasonably far from the disc set $\mathcal{D}$, and so the endpoint of the random walk is $w_n x_0$. By linear progress, it is very likely that the random walk has gone a reasonable distance, and by exponential decay for shadows, it is very unlikely that the Gromov product $(x_0 \cdot g_m w_n x_0)_{g_m x_0}$ would be large, and this Gromov product measures how much the geodesic from $g_m x_0$ to $g_m w_n x_0$ fellow-travels, or backtracks, along the path from $g_m x_0$ to x_0. In this case, we may then apply Proposition 7 to estimate the distance from $\mathcal{D}$ to $g_m w_n x_0$. This is the main

case we consider in the proof below, but we also need to estimate separately the less likely cases in which the random walk does not go very far, or the Gromov product is so large that Proposition 7 does not apply. The final bounds arise from adding the bounds we get in each case.

Proof: (of Proposition 10)

Let $R \in \mathbb{N}$ such that

$$R \geqslant \max\{\log(q^2/3K)/\log(c) + C_1 + C_3, 2\log(q^2/3K)/\log(c), C_2 + 2C_3\}, \quad (2.12)$$

where C_1, C_2 and C_3 are constants from Proposition 7, which only depend on the coarse geometry constants δ and q; K and c are the exponential decay constants, which depend on μ.

We then choose

$$N \geqslant \max\{2R/L, \log(q^3/3K)/\log(c)\}, \quad (2.13)$$

with the same notation for constants as above where L is the linear progress constant, which only depends on μ.

We emphasize that our choice of R only depends on q and μ, and our choice of N depends on R, q and μ.

To estimate the probability that $\phi_R(gw_nx_0)$ takes certain values we need to estimate the distance from the disc set $\mathcal{D}$ to gw_nx_0. In the main case we consider we shall do this by using Proposition 7. We shall apply Proposition 7 with $y = g_mx_0$, $z = g_mw_nw_0$, and x the nearest point in the disc set to g_mx_0.

Proposition 7 then says that the following condition on the Gromov product

$$(x \cdot g_mw_nx_0)_{gx_0} \leqslant d(x, g_mx_0) - C_1, \quad (2.14)$$

implies the following bound on the distance from $g_m w_n x_0$ to the disc set $\mathcal{D}$,

$$d(\mathcal{D}, g_m w_n x_0) \geqslant d(x, g_m x_0) + d(g_m x_0, g_m w_n x_0) - 2(x \cdot g_m w_n x_0)_{g_m x_0} - C_2.$$

Recall that as $x_0 \in \mathcal{D}$, the final part of Proposition 7 shows that we may replace $(x \cdot g_m w_n x_0)_{g x_0}$ with $(x_0 \cdot g_m w_n x_0)_{g_m x_0}$ up to bounded error, so (2.14) is implied by the following condition

$$(x_0 \cdot g_m w_n x_0)_{g_m x_0} \leqslant d(x, g_m x_0) - C_1 - C_3.$$

As $d(x, g_m x_0) = d(\mathcal{D}, g_m x_0)$, and $d(\mathcal{D}, g_m x_0) \leqslant R\phi_R(g_m x_0)$, and we have assumed $\phi_R(g_m x_0) = t$, we can rewrite the condition above as

$$(x_0 \cdot g_m w_n x_0)_{g_m x_0} \leqslant Rt - C_1 - C_3. \tag{2.15}$$

We shall consider various cases, depending on whether some combination of condition (2.15), and the following condition (2.16), hold. The second condition is that the sample path has travelled a distance at least $2R$ from $g_m x_0$ to $g_m w_n x_0$, i.e.

$$d(g_m x_0, g_m w_n x_0) \geqslant 2R. \tag{2.16}$$

We shall consider the following three cases, defined in terms of the conditions above, which cover all possibilities. The table below summarizes the three cases, the various possibilities for $\phi_R(g_m w_n x_0)$ which may occur given the conditions, and the bounds on the probabilities that these values of $\phi_R(g_m w_n x_0)$ occur.

Case	Conditions	Value of $\phi_R(g_m w_n x_0)$	Probability
1	(2.15) and (2.16) hold	$\phi_R(g_m w_n x_0) \leqslant t - r$	$\leqslant q^{r+1}/3, r \geqslant 1$
			$\leqslant 1, r = 0$
2	(2.15) fails	$\phi_R(g_m w_n x_0) \geqslant 0$	$\leqslant q^{t+1}/3$
3	(2.16) fails	$\phi_R(g_m w_n x_0) \geqslant t - 2$	$\leqslant q^3/3$

We now consider each case in turn.

Case 1. If both conditions (2.15) and (2.16) hold, and so we may apply Proposition 7. In this case, consequences (2.2) and (2.3) imply

$$d(\mathcal{D}, g_m w_n x_0) \geqslant Rt + d(g_m x_0, g_m w_n x_0) - 2(x_0 \cdot g_m w_n x_0)_{g_m x_0} - 2C_3 - C_2,$$

where C_2 and C_3 only depend on δ and q. As condition (2.16) holds, we are in the case in which $d(g_m x_0, g_m w_n x_0) \geqslant 2R$, and as we have chosen $R \geqslant C_2 + 2C_3$ this implies that

$$d(\mathcal{D}, g_m w_n x_0) \geqslant Rt + R - 2(x_0 \cdot g_m w_n x_0)_{g_m x_0},$$

and so $\phi_R(g_m w_n x_0) \leqslant t + 1 - r$ may only occur if $(x_0 \cdot g_m w_n x_0)_{g x_0} \geqslant Rr/2$. By exponential decay for shadows this happens with probability at most $Kc^{Rr/2}$, which is at most $q^r/3$, for $r \geqslant 1$, and at most $K \geqslant 1$ for $r = 0$, by (2.12), as we have chosen $R \geqslant 2\log(q^2/3K)/\log(c)$. Therefore this case contributes an amount $q^r/3$ to the upper bound for the probability that $\phi_R(g_m w_n x_0) = t + 1 - r$, for all $1 \leqslant r \leqslant t + 1$, and an amount 1 to the upper bound for $\phi_R(g_m w_n x_0) = t + 1$.

Case 2. If condition (2.15) fails, i.e.

$$(x_0 \cdot g_m w_n x_0)_{g_m x_0} \geqslant Rt - C_3 - C_1. \tag{2.17}$$

By exponential decay for shadows, Fact 12, the probability that condition (2.17) does not hold is at most $Kc^{Rt-C_3-C_1}$, which is at most $q^{t+1}/3$, for $t \geqslant 1$, as by (2.12) we have chosen $R \geqslant \log(q^2/3K)/\log(c) + C_1 + C_3$. In this case there is no restriction on the possible value of $\phi_R(g_m w_n x_0)$, and so this case contributes an amount of $q^{t+1}/3$ to the upper bound for the probability that $\phi_R(g_m w_n x_0) = r$ for every possible value of $r \in \mathbb{N}_0$.

Case 3. If condition (2.16) fails, i.e. the sample path travels distance at most $2R$. By linear progress with exponential decay, Fact 11,

$$\mathbb{P}(d(g_m x_0, g_m w_n x_0) \leqslant L \cdot n) \leqslant Kc^n.$$

By (2.13) we have chosen $N \geqslant 2R/L$, and so this implies that

$$\mathbb{P}(d(g_m x_0, g_m w_n x_0) \leqslant 2R) \leqslant Kc^n,$$

for all $n \geqslant N$. Also by (2.13) we have chosen $N \geqslant \log(q^3/3K)/\log(c)$, and so $Kc^n \leqslant q^3/3$ for all $n \geqslant N$. As the sample path has not gone very far, then the distance from $\mathcal{D}$ to $g_m w_n x_0$ can not have decreased too much. In fact, by the triangle inequality, as

$$d(\mathcal{D}, g_m w_n x_0) \geqslant d(\mathcal{D}, g_m x_0) - d(g_m x_0, g_m w_n x_0),$$

this implies that

$$d(\mathcal{D}, g_m w_n x_0) \geqslant Rt - 2R,$$

i.e. $\phi_R(g_m w_n x_0) \geqslant \phi_R(g_m x_0) - 2$ in this case, which occurs with probability at most $q^3/3$. This case contributes an amount of $q^3/3$ to the upper bound for those values of $r \in \mathbb{N}_0$ satisfying $r \geqslant t - 2$.

At least one of the three cases above must occur, and so the desired upper bounds arise from summing the probabilities in each case, which we summarize in the table below.

		Upper bound			
r	Value of $\phi_R(gw_nx_0) = t+1-r$	Case 1	Case 2	Case 3	Total
$4 \leqslant r \leqslant t+1$	$0 \leqslant t+1-r \leqslant t-3$	$q^r/3$	$q^{t+1}/3$		$\leqslant q^r$
$1 \leqslant r \leqslant 3$	$t-2 \leqslant t+1-r \leqslant t$	$q^r/3$	$q^{t+1}/3$	$q^3/3$	$\leqslant q^r$
$r = 0$	$t+1 \leqslant t+1-r$	1	$q^{t+1}/3$	$q^3/3$	$\leqslant 1$

The Total column gives an upper bound on probability that any of the cases occur, which in the final row is the trivial upper bound of 1.

This completes the proof of Claim 10.

2.6 Train tracks and shadows

It remains to prove the first estimate stated at the end of Section 2.3:

If hx_0 lies inside the disc set at the n-th step, then there is a definite chance $\varepsilon > 0$ where after M further steps x will be at least R away from the disc set with probability ε, namely

$$\mathbb{P}(d(w_{n+M}x_0, \mathcal{D}) \geq R \mid w_n = h \text{ with } hx_0 \in \mathcal{D}) \geqslant \varepsilon.$$

The key observation leading to the proof is that a maximal recurrent train track determines a subset of the curve complex which contains a shadow set.

Let us first give some basic definitions of train tracks, see Penner and Harer [PH92] for more details.

A **train track** is a smoothly embedded 1-complex τ on a surface such that edges (called **branches**) at each vertex (called a **switch**) are all tangent, and there is at

least one edge in both possible tangent directions at each vertex. Therefore, for each switch, the branches are divided into two non-empty sets of branches with the same signed tangent vector, which are called the incoming and outgoing branches.

The complementary regions of a train track are surfaces with piecewise smooth boundaries and vertices. None of the complementary regions may be annuli or discs with two or fewer vertices. A train track on a closed surface is **maximal** if every complementary region is a triangle.

A train route is a smooth path in τ on the traintrack, i.e. it crosses a switch by going from an incoming branch to an outgoing branch, or vice versa. A train track is **recurrent** if every branch is contained in a closed train route. A **transverse measure** on τ is a non-negative function on the branches which satisfies the switch condition, i.e. at each switch, the sum of the measures of the incoming branches is equal to the sum of the measures of the outgoing branches. Any closed train route induces a transverse measure on τ given by the counting measure.

We shall write $P(\tau)$ for the collection of transverse measures supported by τ, which is a subset of $\mathcal{ML}$, Thurston's space of measured laminations on the surface, and $P(\tau)$ is a cone on a compact polyhedron in $\mathcal{PML}$, the projectivization of $\mathcal{ML}$. If the train track is maximal, then $P(\tau)$ has the same dimension as $\mathcal{ML}$.

A simple closed curve x is said to be **carried** by τ if x is homotopic to a train route. A train track σ is **carried** by another train track τ if every train route on σ is homotopic to a train route on τ, and we denote this by $\sigma \prec \tau$.

Given a train track τ, we may produce new train tracks by **splitting** τ, by the following local modification, illustrated in Figure 2.10, in which a subset of the train track corresponding to the top configuration is replaced by one of the three lower configurations.

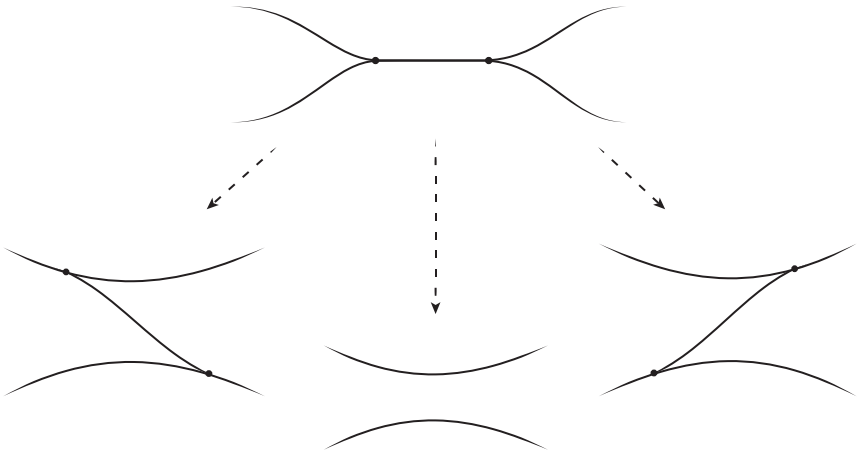


Figure 2.10: Splitting a train track.

The central configuration is called a **collision** or degenerate split. For our purposes, we will only need to consider non-degenerate splits.

Given a recurrent maximal train track τ, let $C(\tau)$ be the collection of simple closed curves carried by τ.

In order to prove the first property mentioned in Section 2.3 a key observation is that a maximal recurrent train track determines a subset of the curve complex which contains a shadow set.

To prove this, we start by showing if two shadow sets intersect, then we may increase the parameter of one of them by a bounded amount such that the new shadow set contains both of the original ones.

Proposition 11. *There is a number R depending only on δ, such that if*

$$S_{x_0}(x_1, R_1) \cap S_{x_0}(x_2, R_2) \neq \varnothing$$

then

$$S_{x_0}(x_1, R_1) \subset S_{x_0}(x_2, \min\{d(x_0, x_2) - d(x_0, x_1) + R_1, R_2\} + R).$$

Proof. Choose

$$y \in S_{x_0}(x_1, R_1) \cap S_{x_0}(x_2, R_2).$$

Since y lies in both shadow sets, we have

$$(x_i \cdot y)_{x_0} \geqslant d(x_i, y) - R_i$$

for $i = 1, 2$.

We know that (see, for example, Bridson and Haefliger [**?**]*III.H.1.20) for any points x, y and z,

$$(x \cdot z)_{x_0} \geqslant \min\{(x \cdot y)_{x_0}, (y \cdot z)_{x_0}\} - \delta, \tag{2.18}$$

Applying (2.18) to the three points x_1, x_2 and $y$yields

$$(x_1 \cdot x_2)_{x_0} \geqslant \min\{d(x_0, x_1) - R_1, d(x_0, x_2) - R_2\} - \delta.$$

Fix $z \in S_{x_0}(x_1, R_1)$, $(x_1 \cdot z)_{x_0} \geqslant d(x_0, z) - R_1$.

Similarly, applying (2.18) to the three points x_1, x_2 and $z$yields

$$(x_2 \cdot z)_{x_0} \geqslant \min\{d(x_0, x_1) - R_1, d(x_0, x_2) - R_2\} - 2\delta.$$

Setting $R = 2\delta$ would then finish the proof. □

It is easy to see that shadow set is always non-empty as long as $R \geqslant 0$. Furthermore, if the parameter R is sufficiently large, the shadow set would have non-empty limit sets in $\partial\mathcal{C}(S_g)$.

We say a group G acts **coarsely transitively** on X if there is a number K such that for any x and y in X, there is a group element g such that $d(gx, y) \leqslant K$.

Check that the action of $MCG(S_g)$ on $\mathcal{C}(S_g)$ is coarsely transitive.

Theorem 7. *[BHM11] Let X be a Gromov hyperbolic space whose isometry group*

acts coarsely transitively on X. There exists $R_0 \geqslant 0$ depending only on δ such that for all $R \geqslant R_0$ and $x, x_0 \in X$ the shadow set $S_{x_0}(x, R)$ has a limit set in ∂X which contains a non-empty open set.

For any $A > 0$ we have $S_{x_0}(x, R) \subseteq S_{x_0}(x, R + A)$ and if A is sufficiently large the limit sets of the shadow sets are strictly nested and the distance between them in $\mathcal{C}(S_g)$ is bounded below in terms of A.

Fact 13. *[Mah12] Let X be a δ-hyperbolic space. There exists K depending only on δ, such that for any $A, R > 0$ and any $x, y \in X$ with $d(x, y) \geqslant A + R + 2K$, the closure of the shadow $S_x(y, R)$ is disjoint from the closure of the complement of the shadow $S_x(y, R + A + K)$, i.e.*

$$\overline{S_x(y, R)} \cap \overline{X \setminus S_x(y, R + A + K)} = \varnothing.$$

Furthermore for any pair of points $a, b \in X$ with $a \in S_x(y, R)$ and $b \in X \setminus S_x(y, R + A + K)$, the distance between a and b is at least A.

We will use these two properties in the form of the following elementary corollary, which says that every shadow set contains shadow sets nested inside it by an arbitrarily large distance.

Corollary 4. *There exists $R_0 \geqslant 0$ such that for any $A \geqslant 0$, and for any shadow set $S_x(y, R_0)$, there is a shadow set $S_x(z, R_0) \subset S_x(y, R_0)$ with*

$$\overline{\mathcal{C}(S_g) \setminus S_x(y, R_0)} \cap \overline{S_x(z, R_0)} = \varnothing$$

and

$$d(\mathcal{C}(S_g) \setminus S_x(y, R_0), S_x(z, R_0)) \geqslant A.$$

We also observe that the complement of a shadow set is roughly a shadow set.

Fact 14. *[Mah12] There exists K depending on δ, such that for all $R \geqslant 2K$ and all $x, z \in \mathcal{C}(S_g)$ with $d(x,z) \geqslant R + K$,*

$$S_x(z, d(x,z) - R - K) \subset \mathcal{C}(S_g) \setminus S_z(x, R) \subset S_x(z, d(x,z) - R + K).$$

Finally, we recall the following 'change of basepoint' result for shadow sets.

Fact 15. *[Mah12] There exist numbers A and B depending on δ such that for any r and any three points x, y and z with $(x \cdot y)_z \leqslant r - A$, we have*

$$S_z(x, r) \subset S_y(x, s)$$

where $s = d(x,y) - d(x,z) + r - B$.

We now provide a link between subsets of the curve complex determined by train tracks and shadow sets, by showing that every maximal train track τ determines a subset $C(\tau)$ of the curve complex which contains a shadow set.

Definition 4. A measured lamination is **minimal** if it

- does not contain any simple closed curves; and
- is not disjoint from any simple closed curve.

Let $\mathcal{PML}_{\min}(S_g)$ be the set of all minimal laminations. We have $\mathcal{PML}_{\min}(S_g) \subseteq \mathcal{PML}(S_g)$ with the relative topology.

The set of **ending laminations** $\mathcal{EL}(S_g)$ is a quotient of $\mathcal{PML}_{\min}(S_g)$ by topological equivalence, i.e. two measured laminations are identified if they correspond to different measures on the same lamination.

The **Teichmüller space** of the surface S_g, denoted by $\mathcal{T}(S_g)$, is the space of hyperbolic metrics on the S_g.

Define map $\iota\colon \mathcal{T}(S_g) \to \mathcal{C}(S_g)$ where for any hyperbolic metric $d \in \mathcal{T}(S_g)$

$$\iota : d \mapsto \text{shortest curve on } S_g \text{ measured in } d.$$

It was shown independently by Klarreich [Kla] and Hamenstädt [Ham06] that the Gromov boundary of the curve complex is homeomorphic to the ending lamination space.

Theorem 8. *[Kla] The inclusion map $\iota\colon \mathcal{T}(S_g) \to \mathcal{C}(S_g)$ extends continuously to $\mathcal{PML}_{\min}(S_g)$ to $\pi\colon \mathcal{PML}_{\min}(S_g) \to \partial\mathcal{C}(S_g)$. π is surjective and $\pi((\mathcal{L}_1, \mu_1)) = \pi((\mathcal{L}_2, \mu_2))$ if and only if $\mathcal{L}_1 = \mathcal{L}_2$.*

Furthermore, π induces a homeomorphism between $\mathcal{EL}(S_g)$ and $\partial\mathcal{C}(S_g)$.

We now use this to show that every maximal recurrent train track τ determines a subset $C(\tau)$ of the complex of curves which contains a shadow set.

Proposition 12. *For any simple closed curve α and any maximal recurrent train track τ there is a simple closed curve β and a number $R \geqslant R_0$ such that the shadow set $S_x(y, R)$ is contained in $C(\tau)$.*

Proof. Let $S_\alpha(\beta, R)$ be any shadow set with $R \geqslant R_0$, with $d(\alpha, \beta)$ sufficiently large such that the closure of $S_\alpha(\beta, R)$ is not equal to all of $\mathcal{C}(S_g)$.

Given a pseudo-Anosov element ϕ in $MCG(S_g)$, we shall write $(\mathcal{L}^s(\phi), \mathcal{L}^u(\phi)) \in \mathcal{PML}(S_g) \times \mathcal{PML}(S_g)$ consisting of its stable and unstable laminations of ϕ.

Such pairs $(\mathcal{L}^s(\phi), \mathcal{L}^u(\phi))$ are dense in $\mathcal{PML}(S_g) \times \mathcal{PML}(S_g)$, as ϕ runs over all pseudo-Anosov elements in $MCG(S_g)$, see for example [Mah10]*Lemma 3.4.

Any pseudo-Anosov element ϕ acts with north-south dynamics on the Thurston compactification $\mathcal{T}(S_g) \cup \mathcal{PML}(S_g)$, with source and sink being the unstable and stable laminations, respectively. i.e. for any open sets U and closed set V with $\mathcal{L}^u(\phi) \in U$, $\mathcal{L}^s(\phi) \notin U$ and $\mathcal{L}^s(\phi) \notin V$; there exists N and g such that for all $n \geqslant N$, $\phi^n(V) \subset U$.

Therefore, it suffices to show that there is a shadow set $S_\alpha(\beta, R)$, such that the closure of $\iota^{-1}(S_\alpha(\beta, R))$ in the Thurston compactification $\mathcal{T}(S_g) \cup \mathcal{PML}(S_g)$ is disjoint from an open set in $\mathcal{PML}(S_g)$. Suppose not, then $\iota^{-1}(S_\alpha(\beta, R))$ is dense, i.e. $\overline{\iota^{-1}(S_\alpha(\beta, R))} = \mathcal{PML}(S_g)$.

Since $\mathcal{EL}$ is dense in $\mathcal{PML}$ and its image under ι is equal to $\partial\mathcal{C}(S_g)$, this implies $\overline{S_\alpha(\beta, R)} = \partial C(S_g)$, a contradiction. □

We now show that if the closure of a shadow set $S_x(y, R)$ contains a limit point of the subgroup supporting the random walks, then there is a slightly larger shadow set $S_x(y, R + R_1)$ whose closure has positive measure with respect to the hitting measure ν.

Proposition 13. *There exists $R_1 > 0$ such that if $S_{x_0}(y, R)$ contains a limit point of H, then $\nu(\overline{S_{x_0}(y, R + R_1)}) > 0$.*

Furthermore, for any number D, there is a shadow set $S_{x_0}(z, R_0) \subset S_{x_0}(y, R + R_1)$ which is at least distance D away from x_0, with $\nu(\overline{S_{x_0}(z, R_0)}) > 0$.

Proof. There is a sequence of group elements $(h_i)_{i\in\mathbb{N}}$ such that $(h_i x_0)_{i\in\mathbb{N}}$ converges to $\mathcal{L}$ in $S_{x_0}(y, R)$, so infinitely many of the $(h_i x_0)_{i\in\mathbb{N}}$ lie in $S_{x_0}(y, R + R_1)$.

Choose one with $\mu_n(h_i) > 0$ for some n. By exponential decay for shadows, a definite proportion of sample paths starting from $h_i x_0$ at time n converge into $S_{x_0}(y, R + R_2)$, so we have $\nu(\overline{S_{x_0}(y, R + R_2)}) > 0$, as required.

To obtain the second claim, consider the countable collection

$$\mathcal{C}_D = \{S_{x_0}(p, R_0) \mid p \in C(S_g) \text{ is a vertex}; d(x_0, p) \geq D\}.$$

$\mathcal{C}_D$ is a cover of $S_{x_0}(y, R + R_2)$, hence at least one of $S_{x_0}(p, R_0)$ has positive measure and is contained in $S_{x_0}(y, R + R_3)$, for some R_3. □

Now we are ready to show the desired estimate from Section 2.3, more precisely we prove:

Proposition 14. *For any complete subgroup $G < MCG(S_g)$ there is a finitely generated non-elementary subgroup $H < G$ with probability distribution generated by random walk with a fixed set of generators. For any $R > 0$ there exists numbers N and $\varepsilon > 0$ such that for any $g \in MCG(S_g)$ and $n \geq N$, the random walk of length n starting at g will be at least R away from the disc set $\mathcal{D}$ with probability ε, i.e.*

$$\mathbb{P}(\phi_R(gw_nx_0) = 1 \mid \phi_R(gx_0) = 0) \geqslant \varepsilon.$$

The following result of Kerckhoff [Ker90] plays an important role in our construction of H.

Theorem 9. *[Ker90]*Proposition on p36 There is a recurrent maximal train track τ on a closed orientable surface S_g such that for any identification of S_g with the boundary of a handlebody, τ can be split at most $-9\chi(S_g)$ times to a recurrent maximal train track τ' where $N(\tau')$ is disjoint from the disk set $\mathcal{D}$ of the handlebody.*

Here $\chi(S_g) = -2(g-1)$ is the Euler characteristic of the surface.

Proof. (of Proposition 14.) Let τ be a recurrent maximal train track on S_g satisfying Theorem 9, and let $\mathcal{T}$ be the finite collection of maximal train tracks obtained by splitting τ at most $-9\chi(S_g)$ times.

By Proposition 12, the subset of the curve complex corresponding to each maximal train track $\tau_i \in \mathcal{T}$ contains some shadow set $S_i = S_{x_0}(y_i, R_0)$.

Furthermore, for any $R \geqslant 0$, for each i we may choose a shadow set $S_i' = S_{x_0}(y_i', R_0)$ with $\overline{S_i'} \cap \overline{\mathcal{C}(S_g) \backslash S_i} = \varnothing$ and $d(\mathcal{C}(S_g) \setminus S_i, S_i') \geqslant R$.

Since the subgroup G is complete in $MCG(S_g)$, endpoints of pseudo-Anosov elements are dense in $\partial\mathcal{C}(S_g)$. In fact, the pairs $(\mathcal{L}^s, \mathcal{L}^u)$ of stable and unstable laminations are dense in $\partial\mathcal{C}(S_g) \times \partial\mathcal{C}(S_g)$.

Each shadow set S_i' contains a non-empty open set in $\partial\mathcal{C}(S_g)$, so for each S_i' we can choose a pseudo-Anosov element $g_i \in G$ with at least one limit points in S_i'.

Define

$$H \coloneqq \langle g_i \mid \tau_i \in \mathcal{T} \rangle.$$

By Proposition 13 and the constriction of S_i, we have the hitting measure $\nu(\overline{S_i}) > 0$ for each i.

Set

$$\varepsilon = \min \left\{ \nu(\overline{S_i}) \mid \tau_i \in \mathcal{T} \right\} /2,$$

as the measures induced by finite random walks weakly converge to ν, there is an N such that $\mu_n(\overline{S_i}) \geqslant \varepsilon$ for all $n \geqslant N$ and for all i.

Now consider a random walk of length n starting from gx_0. We wish to estimate the distance $d(\mathcal{D}, gw_nx_0)$, and by applying the isometry g^{-1}, this is equivalent to considering the distance $d(g^{-1}\mathcal{D}, w_nx_0)$. By Theorem 9, for any disc set, in particular the disc set $g^{-1}\mathcal{D}$, there is some maximal train track τ_i in T disjoint from $g^{-1}\mathcal{D}$, and so if w_nx_0 lies in the nested shadow set S_i' contained in $C(\tau_i)$, then $d(g^{-1}\mathcal{D}, w_nx_0) \geqslant R$, and so $\phi_R(gw_nx_0) \geqslant 1$.

As $\mu_n(S_i') \geqslant \varepsilon$ for all $n \geqslant N$, this implies that w_nx_0 lies in S_i' with probability at least ε, completing the proof of Proposition 14. □

2.7 Exponential decay for splitting distance

In this section we prove the exponential decay for Heegaard splitting distance and hence Theorem 5.

First we use the local estimates for distance from the disc set obtained in the previous sections to show that the distance from w_nx_0 to the disc set grows linearly with exponential decay.

Proposition 15. *For any complete subgroup G of $MCG(S_g)$, there is a finitely generated $H < G$ with probability distribution μ generated by random walk with a fixed*

set of generators; Such that there exists numbers $K, L > 0$ and $c < 1$ depending on μ, for any disc set $\mathcal{D}$ and any basepoint x_0, the distance from $w_n x_0$ to the disc set $\mathcal{D}$ grows linearly with exponential decay, i.e.

$$\mathbb{P}(d(\mathcal{D}, w_n x_0) \leqslant Ln) \leqslant Kc^n.$$

We show this by comparing the distribution of the random variables $d(\mathcal{D}, w_n x_0)$ with a Markov chain on $\mathbb{N}_0$ which gives an upper bound on the probability that the random variable takes small values. Ultimately, we show the Markov chain $(\mathbb{N}, P_0)$ has spectral radius $\rho(P)$ strictly less than 1, which gives the exponential decay result.

A useful observation is that if an exponential decay estimate holds for the iterated random walk w_{nN} generated by the N-fold convolution measure μ_N (i.e. if one take the walk N steps at a time), then a similar exponential decay estimate holds for w_n with different constants, i.e.:

Proposition 16. *Let μ be a finitely supported probability distribution such that for some $N > 0$ we have*

$$\mathbb{P}(d(\mathcal{D}, w_{nN} x_0) \leqslant Ln) \leqslant Kc^n$$

for all n, then there exists $L' > 0$ and $c' < 1$ such that

$$\mathbb{P}(d(\mathcal{D}, w_n x_0) \leqslant L'n) \leqslant Kc'^n,$$

for all n.

Proof. Since μ has finite support, let

$$R \coloneqq \operatorname{diam}(\operatorname{supp}(\mu)).$$

For any $0 \leq k \leqslant N$, we have

$$d(\mathcal{D}, w_{nN+k}x_0) \leqslant d(\mathcal{D}, w_{nN}x_0) + RN,$$

choose $L' = L + RN$ and $c' = c^{1/N}$, yielding

$$\mathbb{P}(d(\mathcal{D}, w_n x_0) \leqslant L'n) \leqslant Kc'^n.$$

□

We may now replace μ with μ_N and assume that the local estimates proved in pervious sections hold for $N = 1$, we shall do this for the remainder of this section to simplify notation.

Consider the sequence of random variables $X_n \colon \Omega \to \mathbb{N}_0$ defined by

$$X_n(\omega) = \phi_R(w_n x_0),$$

we wish to compare the distributions of the X_n with the distributions X'_n arising from the following Markov chain $(\mathbb{N}_0, P)$, starting with total mass 1 at $0 \in \mathbb{N}_0$ at time $n = 0$. We shall write $p(i, j)$ for the probability you go from i to j.

$$p(0,j) = \begin{cases} 1-\varepsilon & \text{if } j = 0 \\ \varepsilon & \text{if } j = 1 \\ 0 & \text{if } j \geqslant 2 \end{cases}$$

and for $i > 0$,

$$p(i,j) = \begin{cases} q^{i-j+1} & \text{if } j \leqslant i \\ 1 - q - q^2 - \cdots q^{i+1} & \text{if } j = i+1 \\ 0 & \text{if } j \geqslant i+2 \end{cases}$$

Figure 2.4 illustrates the first few vertices of this Markov chain.

Given a probability measure P on $\mathbb{N}_0$, let F_P be the cumulative distribution function of X:

$$F_P(n) = \sum_{i=0}^{n} P(i).$$

Given two probability measures P and P' on $\mathbb{N}_0$, we say that P' **stochastically dominates** P if $F_P(n) \geqslant F_{P'}(n)$ for all n, denoted by $P \preccurlyeq P'$.

Similarly, given a random variable $X\colon (I, \mathbb{P}) \to \mathbb{N}_0$, we write F_X for the cumulative distribution function of X. A random variable X **stochastically dominates** another random variable X', i.e. $X' \preccurlyeq X$, if $F_X(n) \leqslant F_{X'}(n)$ for all n.

Let $\{X_n\}$ be a sequence of random variables with values in $\mathbb{N}_0$. For any subset $A \subset \mathbb{N}_0$, the **transition kernels** for the sequence are the measures on $\mathbb{N}_0$ given by

$$K_n(i_1, \ldots, i_{n-1})(A) = \mathbb{P}(X_n \in A \mid (X_1, \ldots, X_{n-1}) = (i_1, \ldots i_{n-1})).$$

We shall see that if the transition kernels for two sequences of random variables $\{X_n\}$ and $\{X'_n\}$ satisfy $K'_n(i_1, \ldots i_{n-1}) \preccurlyeq K_n(i_1, \ldots i_{n-1})$ for all n and $i_1, \ldots i_{n-1}$, then $X'_n \preccurlyeq X_n$ for all n.

The precise statement of this result is often referred to as Strassen's Theorem, or the Stochastic Domination Theorem, see Lindvall [Lin92]*Chapter IV.

Fact 16. *[Lin92]*Theorem 5.8 Let $\{X_n\}$ and $\{X'_n\}$ be sequences of random variables with values in $\mathbb{N}_0$ such that $X'_0 \preccurlyeq X_0$, $K'_n(i_1, \ldots i_{n-1}) \preccurlyeq K_n(i_1, \ldots i_{n-1})$ for all n and $i_1, \ldots i_{n-1}$.*

Then $X'_n \preccurlyeq X_n$ for all n.

In our case we have $X'_0 = X_0$, in order to show that $X'_n \preccurlyeq X_n$ it suffices for the transition kernels to satisfy $K'_n \preccurlyeq K_n$. We now prove this using the two local estimates from Section 2.5 and 2.6, namely Proposition 10 and 14.

Proposition 17. *For all n and $i_1, \ldots, i_{n-1}$, the transition kernels for $\{X_n\}$ and $\{X'_n\}$*

satisfy

$$K'_n(i_1, \ldots, i_{n-1}) \preccurlyeq K_n(i_1, \ldots, i_{n-1}).$$

Proof. As the random variables $\{X'_n\}$ arise from a Markov chain, $K'_n(i_1, \ldots, i_{n-1})$ only depends on the value of i_{n-1} and may be computed from the defining transition probabilities $p(i, j)$ of the Markov chain.

We first consider the case $i_{n-1} = 0$:

Applying Proposition 14, we have for all n,

$$\mathbb{P}(X_n(\omega) = 0 \mid X_{n-1}(\omega) = 0) \leqslant 1 - \varepsilon, \tag{2.19}$$

writing this out in terms of the cumulative probability distributions gives

$$F_{K_n(i_1,\ldots,i_{n-2},0)}(0) \leqslant F_{K'_n(i_1,\ldots i_{n-2},0)}(0).$$

By definition of the Markov chain, the value of X'_n may not increase by more than one from the value of X'_{n-1}, i.e.

$$F_{K'_n(i_1,\ldots i_{n-2},0)}(1) = 1,$$

hence for all n and all $i_1, \ldots, i_{n-2}$ we have

$$K'_n(i_1, \ldots i_{n-2}, 0) \preccurlyeq K_n(i_1, \ldots i_{n-2}, 0).$$

We now consider the case $i_{n-1} > 0$:

Applying Proposition 10, for all $0 \leqslant i \leqslant k$ and all n, we have

$$\mathbb{P}(X_n(\omega) = k - i \mid X_{n-1}(\omega) = k) \leqslant q^{i+1}. \tag{2.20}$$

Summing over values of i with $l \leqslant i \leqslant k$ results that for all $l \leqslant k$,

$$F_{K_n(i_1,\dots i_{n-2},k)}(l) \leqslant F_{K'_n(i_1,\dots i_{n-2},k)}(l).$$

Again, by definition of the Markov chain, the value of X'_n may not increase by more than 1 at any step,

$$F_{K'_n(i_1,\dots i_{n-2},k)}(k+1) = 1$$

which implies for all n and all $i_1, \dots i_{n-2}$, and with $k > 0$,

$$K'_n(i_1, \dots i_{n-2}, k) \preccurlyeq K_n(i_1, \dots i_{n-2}, k).$$

□

We now show that the Markov chain $(\mathbb{N}_0, P)$ has spectral radius $\rho(P) < 1$.

For all functions f on $\mathbb{N}_0$, define

$$Pf(k) = \sum p(k,j) f(j).$$

Proposition 18. *If $q < 1/4$ then the Markov chain $(\mathbb{N}_0, P)$ has spectral radius $\rho < 1$.*

Proof. Recall from Woess [Woe00]*Section 7 that $\rho \leqslant t$ if there is a strictly positive t-superharmonic function f on $\mathbb{N}_0$, i.e the function f satisfies $f(k) > 0$ for each k and $Pf \leqslant tf$.

Let $t = \max\{1 - \varepsilon(1-2q), 4q\}$ we show $f(k) = (2q)^k$ is t-superharmonic for $t < 1$ as long as $q < 1/4$.

For $k = 0$, the inequality $Pf(k) \leqslant tf(k)$ can be written as

$$(1-\varepsilon)f(0) + \varepsilon f(1) \leqslant tf(0). \tag{2.21}$$

When $k \geq 1$ the inequalities are of the form

$$q^{k+1}f(0) + q^k f(1) + \cdots + qf(k) + p_k f(k+1) \leqslant tf(k), \tag{2.22}$$

where $p_k = 1 - q - q^2 - \cdots q^{k+1}$.

(2.21) gives

$$1 - \varepsilon(1-2q) \leqslant t,$$

which is satisfied for some $t < 1$ if $q < 1/2$.

For (2.22), we obtain

$$q^{k+1} + q^k(2q) + \cdots + q(2q)^k + p_k(2q)^{k+1} \leqslant t(2q)^k.$$

As $p_k \leqslant 1$, this inequality is satisfied if

$$q^{k+1}(1 + 2 + \cdots 2^k) + (2q)^{k+1} \leqslant t(2q)^k.$$

As the sum of the geometric series is less than 2^{k+1}, it suffices to choose q such that $4q \leqslant t$, which holds for some $t < 1$ if $q < 1/4$, as required. □

We shall write $p^{(n)}(i,j)$ for the probability that the Markov chain starting at i at time 0 is at location j on the nth step.

A Markov chain on a graph is **uniformly irreducible** if there are numbers N and ε_0 such that for any pair of neighboring vertices i and j, $p^{(n)}(x,y) \geqslant \varepsilon_0$ for some $n \leqslant N$.

If we consider $\mathbb{N}_0$ to have the standard graph structure in which i and j are connected by an edge if and only if $|i-j| \leqslant 1$, then the Markov chain $(\mathbb{N}_0, P)$ above is uniformly irreducible with $N = 1$ and $\varepsilon_0 = \min\{\varepsilon, q^2, 1 - q/(1-q)\}$.

Fact 17. *[Woe00]*Lemma 8.1 If the Markov chain (X,P) is uniformly irreducible,*

then there exists $A > 0$ *such that*

$$p^{(n)}(x, y) \leqslant A^{d(x,y)} \rho(P)^n.$$

Applying Fact 17, we get $A > 0$ such that $F_{X_n}(Ln) \leqslant LnA^{Ln}\rho^n$ which decays exponentially for some $L > 0$, chosen sufficiently close to zero. This completes our proof of Proposition 15.

Now we prove Theorem 5 using Proposition 15 and Proposition 8:

Proof. (of Theorem 5) Choose $m = \lfloor n/2 \rfloor$, consider the sample path w_n to be consisting of two segments: an initial segment w_m of length m and a final segment $w_m^{-1} w_n$, of length $n - m$.

These two sample paths w_m and $w_m^{-1} w_n$ are independently distributed.

By Proposition 15 we know that the distance from $\mathcal{D}$ to $w_m x_0$ decays exponentially in n, i.e.

$$\mathbb{P}(d(\mathcal{D}, w_m x_0) \leqslant Ln/2) \leqslant Kc^{n/2}. \tag{2.23}$$

Note that Proposition 15 also applies to the reflected random walk generated by the probability distribution $\check{\mu}(g) = \mu(g^{-1})$, with the same basepoint x_0, but with the disc set $h_{S^3}\mathcal{D}$, though possibly with different numbers $\check{K}, \check{L} > 0$ and $\check{c} < 1$.

To simplify notation, we shall write $\mathcal{D}'$ for $h_{S^3}\mathcal{D}$. Applying Proposition 15 to the reflected random walk of length $n - m$ yields

$$\mathbb{P}(d(\mathcal{D}', (w_m^{-1} w_n)^{-1} x_0) \leqslant \check{L}(n/2 + 1)) \leqslant \check{K}\check{c}^{\,n/2+1}.$$

Now apply the isometry w_n we obtain

$$\mathbb{P}(d(w_m x_0, w_n \mathcal{D}') \leqslant \check{L}(n/2 + 1)) \leqslant \check{K}\check{c}^{\,n/2+1}. \tag{2.24}$$

Choose $L_1 = \min\{L, \check{L}\}$, let y, y' be a closest point in $\mathcal{D}$ to $w_m x_0$ and y' be a closest point in $\mathcal{D}'$ to $w_m x_0$.

Recall the statement of Proposition 8:

Proposition: For (X, d) δ-hyperbolic, given a number q, there exists C_1, C_2 and C_3 depending only on δ and q, such that if K_x and K_z are q-quasiconvex sets and x is a closest point in K_x to y; z is a closest point in K_z to y, then if

$$(x \cdot z)_y \leqslant \min\{d(x, y), d(y, z)\} - C_1, (2.4)$$

then

$$d(K_x, K_z) \geqslant d(x, y) + d(y, z) - 2(x \cdot z)_y - C_2. (2.5)$$

Furthermore, for any points $x' \in K_x$ and $z' \in K_z$,

$$|(x \cdot z)_y - (x' \cdot z')_y| \leqslant C_3. (2.6)$$

We would like to apply this to the two discs sets as quasiconvex sets and $y = w_m x_0$, the proposition applied unless (2.4) fails, which means

$$(y \cdot y')_{w_m x_0} \leqslant L_1(n/2 + 1) - C_1, \tag{2.25}$$

where $C_1 = 15C_T/2 + C$ is as constructed in the proof of Proposition 8.

However, we shall see that the probability for (2.4) to fail decays exponentially in n:

Let $N = 4C_1/L_1$, for all $n \geq N$ we have

$$L_1(n/2 + 1) - C_1 \geqslant L_1 n/4 - C_3 - C_2/2,$$

where C_1, C_2 and C_3 are constructed in the proof of Proposition 8.

Using the Gromov product estimate (2.6) and exponential decay for shadows, there exists K' and c' such that

$$\mathbb{P}\left(\ (x_0 \cdot w_n x_0)_{w_m x_0} \geqslant L_1 n/4 - C_3 - C_2/2 \ \right) \leqslant K'(c')^{L_1 n/4 - C_3 - C_2/2}, \tag{2.26}$$

the latter term decays exponentially in n, as desired.

If $(y \cdot y')_{w_m x_0} > L_1(n/2+1) - C_1$ then Proposition 8 applies and we have the distance from $\mathcal{D}$ to $w_n \mathcal{D}'$ is at least

$$d(\mathcal{D}, w_n \mathcal{D}') \geqslant d(\mathcal{D}, w_m x_0) + d(w_m x_0, w_n \mathcal{D}') - 2(y \cdot y')_{w_m x_0} - C_2$$

and the difference between the Gromov products $(y \cdot y')_{w_m x_0}$ and $(x_0 \cdot w_n x_0)_{w_m x_0}$ are bounded.

Since $d(w_n x_0, w_n \mathcal{D}') = d(x_0, \mathcal{D}')$ and we have chosen a basepoint x_0 which lies in both $\mathcal{D}$ and $\mathcal{D}'$, we have

$$d(\mathcal{D}, w_n \mathcal{D}') \geqslant d(\mathcal{D}, w_m x_0) + d(w_m x_0, w_n \mathcal{D}') - 2(x_0 \cdot w_n x_0)_{w_m x_0} - 2C_3 - C_2.$$

The three events whose probabilities are estimated in lines (2.23), (2.24) and (2.26) need not be independent, but the probability that at least one of them occurs is at most the sum of the probabilities that each occurs.

Therefore for all $n \geq N$ we have

$$\mathbb{P}(d(\mathcal{D}, w_n \mathcal{D}') \leqslant L_1 n/4) \leqslant K c^{n/2} + \check{K}\check{c}^{\,n/2+1} + K'(c')^{L_1 n/4 - C_3 - C_2/2},$$

which decays exponentially in n, as required. □

Bibliography

[BGV] Myasnikov A Baumslag G. and Shpilrain V., *Open problems in combinatorial and geometric group theory*, http://http://www.sci.ccny.cuny.edu/ shpil/gworld/problems/oproblems.html.

[BH99] Martin R. Bridson and André Haefliger, *Metric spaces of non-positive curvature*, Springer-Verlag, 1999.

[BHM11] Sébastien Blachère, Peter Haïssinsky, and Pierre Mathieu, *Harmonic measures versus quasiconformal measures for hyperbolic groups*, Ann. Sci. Éc. Norm. Supér. (4) **44** (2011), no. 4, 863–721.

[BMS02a] G. Baumslag, A. Myasnikov, and V. Shpilrain, *Open problems in combinatorial group theory.*, Contemporary Mathematics **296** (2002), 1–38.

[BMS02b] A.V. Borovik, A. Myasnikov, and V. Shpilrain, *Measuring sets in infinite groups*, Computational and Statistical Group Theory (Las Vegas, NV/Hoboken, NJ), Contemp. Math., American Mathematical Society, 2002, pp. 21–42.

[BS09] J. Brock and J. Souto, *Volume and distances in the pants complex.*

[BV02] J. Burillo and E. Ventura, *Counting primitive elements in free groups*, Geometriae Dedicata **93** (2002), no. 1, 143–162.

[DT06a] Nathan M. Dunfield and William P. Thurston, *Finite covers of random 3-manifolds*, Invent. Math. **166** (2006), no. 3, 457–521.

[DT06b] ———, *A random tunnel number one 3-manifold does not fiber over the circle*, Geom. Topol. **10** (2006), 2431–2499.

[DW11] Nathan M. Dunfield and Helen Wong, *Quantum invariants of random 3-manifolds*, Algebr. Geom. Topol. **11** (2011), no. 4, 2191–2205.

[Erd59] P. Erdős, *Graph theory and probability*, Canad. J. Math. **11** (1959), 34–38.

[FLP12] A. Fathi, F. Landenbach, and V. Poenaru, *Thurston's work on surfaces*, Princeton University Press, 2012.

[FM12] B. Farb and D. Margalit, *A primer on mapping class groups*, Princeton Mathematical Series, vol. 49, Princeton University Press, 2012.

[Ham06] Ursula Hamenstädt, *Train tracks and the gromov boundary of the complex of curves*, London Math. Soc. Lecture Note Ser. **329** (2006), 187–207.

[Hem76] John Hempel, 3-*manifolds*, Ann. of Math. Studies, No. 86, Princeton University Press, 1976.

[Ker90] Steven P. Kerckhoff, *The measure of the limit set of the handlebody group*, Topology **29** (1990), no. 1, 27–40.

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

[Kob88] Tsuyoshi Kobayashi, *Heights of simple loops and pseudo-anosov homeomorphisms*, Contemp. Math. **78** (1988), no. 327–338.

[Kow08] E. Kowalski, *The large sieve and its applications*, Cambridge Tracts in Mathematics, Cambridge University Press, 2008.

[Lin92] Torgny Lindvall, *Lectures on the coupling method*, Wiley Series in Probability and Mathematical Statistics, John Wiley & Sons Inc., 1992.

[LL10] Gregory F. Lawler and Vlada Limic, *Random walk: a modern introduction*, Cambridge Studies in Advanced Mathematics, vol. 123, Cambridge University Press, 2010.

[LMW14] A. Lubotzky, J. Maher, and C. Wu, *Random methods in 3-manifold theory*, Preprint (2014).

[Lov68] L. Lovász, *On chromatic number of finite set-systems*, Acta Math. Acad. Sci. Hungar. **19** (1968), 59–67.

[LPS88] A. Lubotzky, R. Phillips, and P. Sarnak, *Ramanujan graphs*, Combinatorica **8** (1988), no. 3, 261–277.

[Mah10] Joseph Maher, *Random heegaard splittings*, J. Topol. **3** (2010), no. 4, 997–1025.

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

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

[MS03] A. Myasnikov and V. Shpilrain, *Automorphic orbits in free groups*, Journal of Algebra **269** (2003), no. 1, 18–27.

[MT07] John Morgan and Gang Tian, *Ricci flow and the poincaré conjecture*, Clay Mathematics Monographs, vol. 3, American Mathematical Society, 2007.

[PH92] R. C. Penner and J. L. Harer, *Combinatorics of train tracks*, Annals of Mathematics Studies, vol. 125, Princeton University Press, 1992.

[PP12] D. Puder and O. Parzanchevski, *Measure preserving words are primitive*, Arxiv preprint arXiv:1202.3269 (2012).

[PW14] D. Puder and C. Wu, *Primitive elements in free groups*, Journal of London Math. Soc. (2014).

[Riv04] I. Rivin, *A remark on'counting primitive elements in free groups'(by j. burillo and e. ventura)*, Geometriae Dedicata **107** (2004), no. 1, 99–100.

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

[Riv14] ———, *Statistics of random 3-manifolds occasionally fibering over the circle*, preprint (2014).

[Sav02] Nikolai Saveliev, *Invariants for homology* 3*-spheres*, Encyclopaedia of Mathematical Sciences, vol. 140, Springer-Verlag, 2002.

[Shp05] V. Shpilrain, *Counting primitive elements of a free group*, Contemporary Mathematics **372** (2005), 91–98.

[ST06] Martin Scharlemann and Maggy Tomova, *Alternate heegaard genus bounds distance*, Geom. Topol. **10** (2006), 593–617.

[Sta99] John R. Stallings, *Whitehead graphs on handlebodies*, Geometric group theory down under (J. Cossey, C. F. Miller, W.D. Neumann, and M. Shapiro, eds.), de Gruyter, Berlin, 1999, pp. 317–330.

[Whi36a] J.H.C. Whitehead, *On certain sets of elements in a free group*, Proc. London Math. Soc. **41** (1936), 48–56.

[Whi36b] ———, *On equivalent sets of elements in a free group*, Ann. of Math. **37** (1936), 768–800.

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