Extremality of the Rotation Quasimorphism on the Modular Group

Thesis by

Joel Ryan Louwsma

In Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy

California Institute of Technology

Pasadena, California

2011

(Defended May 26, 2010)

Acknowledgments

First and foremost, I thank my advisor, Danny Calegari, for his invaluable guidance and suggestions over the past several years. Much of the material in Chapter 9 of this thesis is taken from our joint paper [CL11], and I thank him for allowing me to include it here.

I also thank Nathan Dunfield for initially getting me interested in low-dimensional topology at Caltech and for answering many questions as I studied background material.

I thank Michael Aschbacher, Danny Calegari, and Matthew Day for serving on my candidacy committee, and Michael Aschbacher, Danny Calegari, Matthew Day, and Matilde Marcolli for serving on my thesis committee.

I thank several fellow graduate students at Caltech with whom I have had interesting discussions about my work, especially Steven Frankel, Claire Levaillant, Chris Lyons, Paul Nelson, Bob Pelayo, Rupert Venzke, Nahid Walji, Alden Walker, and Dongping Zhuang.

I thank the staff members of the math department at Caltech, Kristy Aubry, Kathy Carreon, Stacey Croomes, Pam Fong, Cherie Galvez, and Seraj Muhammed, for their assistance in many matters over the years.

I thank several people for giving me opportunities to speak about my work in seminars and for interesting conversations during my visits. Specifically, I thank Francis Bonahon at USC, Erica Flapan and Jim Hoste in Claremont, Vestislav Apostolov and Steve Boyer at CIRGET in Montréal, and Genevieve Walsh and Kim Ruane at Tufts.

I thank the Centre de Recherches Mathématiques in Montréal for allowing me to spend Fall 2010 there as part of their thematic semester on Geometric, Combinatorial and Computational Group Theory. In particular, I thank Olga Kharlampovich for organizing most of the activities of the semester and Louis Pelletier for handling many logistical details behind the scenes.

I also thank the Ohio State University for allowing me to spend Spring 2011 in Columbus as part of their Special Year in Topology and Geometric Group Theory. I especially thank Indira Chatterji for making many arrangements for my visit.

Last but not least, I thank my friends and colleagues, not only at Caltech, but at many universities around the world, who have helped to make being a graduate student an enjoyable experience.

Abstract

For any element A of the modular group $\mathrm{PSL}(2,\mathbb{Z})$, it follows from work of Bavard [Bav91] that $\mathrm{scl}(A) \geq \mathrm{rot}(A)/2$, where scl denotes *stable commutator length* and rot denotes the *rotation quasimorphism*. Sometimes this bound is sharp, and sometimes it is not. We study for which elements $A \in \mathrm{PSL}(2,\mathbb{Z})$ the rotation quasimorphism is *extremal* in the sense that $\mathrm{scl}(A) = \mathrm{rot}(A)/2$. First, we explain how to compute stable commutator length in the modular group, which allows us to experimentally determine whether the rotation quasimorphism is extremal for a given A. Then we describe some experimental results based on data from these computations.

Our main theorem is the following: for any element of the modular group, the product of this element with a sufficiently large power of a parabolic element is an element for which the rotation quasimorphism is extremal. We prove this theorem using a geometric approach. It follows from work of Calegari [Cal09a] that the rotation quasimorphism is extremal for a hyperbolic element of the modular group if and only if the corresponding geodesic on the modular surface virtually bounds an immersed surface. We explicitly construct immersed orbifolds in the modular surface, thereby verifying this geometric condition for appropriate geodesics. Our results generalize to the 3-strand braid group and to arbitrary Hecke triangle groups.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

This thesis studies the relationship between two topics of recent interest, stable commutator length [Cal08] and quasimorphisms [Kot04]. A fundamental connection between these concepts was first discovered by Bavard [Bav91]; his result implies that every homogeneous quasimorphism gives a lower bound on stable commutator length. We are interested in determining when these bounds are sharp, in which case the quasimorphism is said to be extremal. Although extremal quasimorphisms are known to exist, only a few examples of them have been found, such as those quasimorphisms recently constructed by Calegari–Walker [CWa] on free groups. The difficulty in finding extremal quasimorphisms is due both to the fact that stable commutator length can presently be computed in relatively few groups and to the fact that the space of all homogeneous quasimorphisms on a group is often poorly understood.

We focus on the modular group $\mathrm{PSL}(2, \mathbb{Z})$, a group that is important in many areas of mathematics, including algebra, geometry, and number theory. We study a particular quasimorphism, the rotation quasimorphism, that arises in many contexts. It may be regarded as the homogenization of the classical Rademacher function, which has numerous interpretations and has been extensively studied by number theorists.

We explain how to compute stable commutator length in the modular group $\mathrm{PSL}(2, \mathbb{Z})$ by giving an algorithm for reducing the problem to that of computing stable commutator length in the free group F_2, which can be done using the computer program `scallop` [CWb]. We have written a program that implements our algorithm and have used it to generate a significant amount of data about when the rotation quasimorphism is and is not extremal. We present some experimental results based on this data.

We also present some theoretical results. Our main result is the following stability theorem.

Theorem. *For every element of the modular group* $\mathrm{PSL}(2,\mathbb{Z})$*, the product of this element with a sufficiently large power of a parabolic element is an element for which the rotation quasimorphism is extremal.*

The proof of this theorem is independent of our algorithm for computing stable commutator length and thus gives an alternate way to determine stable commutator length for certain families of elements of the modular group. Such an approach is a promising way to study stable commutator length in groups in which it is difficult to compute directly. Our proof is primarily geometric, and we obtain the above result as a consequence of another result about corresponding geodesics on the modular surface $\mathbb{H}^2/\,\mathrm{PSL}(2,\mathbb{Z})$ bounding immersed orbifolds. Specifically, we show the following.

Theorem. *For every hyperbolic element of the modular group* $\mathrm{PSL}(2,\mathbb{Z})$*, the product of this element with a sufficiently large power of a parabolic element corresponds to a geodesic on the modular surface* $\mathbb{H}^2/\,\mathrm{PSL}(2,\mathbb{Z})$ *that bounds an immersed orbifold.*

This second result is also of independent interest. Recently there has been significant interest in immersing surfaces in various spaces, for example in the celebrated work of Kahn–Markovic [KM] on the Surface Subgroup Theorem.

We now briefly describe the organization of this thesis. Chapters 2 and 3 introduce the general topics of study. In Chapter 2, we define stable commutator length and discuss some of its properties. In Chapter 3, we define the notion of quasimorphism and give some examples. We also state the fundamental relationship between stable commutator length and quasimorphisms and discuss several general questions arising from this.

Chapters 4 and 5 introduce the particular objects of study in this thesis. In Chapter 4, we introduce the modular group and discuss several fundamental properties of its action on the hyperbolic plane. In Chapter 5, we define the rotation quasimorphism, explain how to compute it, and discuss its relationship with other well-known functions.

Chapters 6 and 7 are devoted to the computational analysis of the extremality of the rotation quasimorphism on the modular group. In Chapter 6, we explain how to compute stable commutator length in the modular group. This gives us a way to experimentally test when the rotation quasimorphism is extremal. In Chapter 7, we discuss several experimental observations based on our data. Some of the data referenced in Chapter 7 is presented in Appendix A.

Chapters 8 and 9 form the theoretical heart of this thesis. Chapter 8 presents the main stability theorem stated above and discusses how it follows from the second theorem about immersed orbifolds in the modular surface. In Chapter 9, we explicitly describe how to construct such immersed orbifolds, thus proving the second theorem.

Finally, Chapter 10 explains how to generalize our results to the 3-strand braid group and to Hecke triangle groups. We conclude with a brief discussion of possible further generalizations.

Chapter 2

Stable commutator length

Stable commutator length, a kind of relative Gromov–Thurston norm, has been the subject of much recent interest, especially by Calegari and his collaborators. In this chapter, we define stable commutator length and discuss some of its properties. For much more information about this concept, we refer the reader to the excellent monograph [Cal09b].

2.1 Commutator length

Let G be any group. The *commutator subgroup* of G, denoted $[G, G]$, is the subgroup generated by *commutators*, i.e. elements of the form $[b, c] := bcb^{-1}c^{-1}$ for $b, c \in G$. The *commutator length* of an element $a \in [G, G]$, denoted $\mathrm{cl}(a)$, is the word length of a with respect to this generating set. More explicitly, this means $\mathrm{cl}(a)$ is the smallest integer g such that $a = \prod_{i=1}^{g}[b_i, c_i]$ for some elements $b_i, c_i \in G$. If $a \notin [G, G]$, we use the convention that $\mathrm{cl}(a) = \infty$. Define the *commutator width* of G to be $\mathrm{cw}(G) := \sup_{a \in [G,G]} \mathrm{cl}(a)$. If this supremum does not exist, we say that $\mathrm{cw}(G) = \infty$.

The commutator length of an element is sometimes referred to as the *genus* of that element, for the following topological reason. Suppose X is a topological space with $\pi_1(X) = G$ and α is a loop representing an element $a \in G$. Let $\Sigma_{g,1}$ be an oriented surface of genus g with one boundary component. One can construct a continuous map $\Sigma_{g,1} \to X$ taking $\partial\Sigma_{g,1}$ to α exactly when a can be written as a product of g commutators. This can be most easily seen by thinking of $\Sigma_{g,1}$ as formed from an identification space with $4g + 1$ edges, where $4g$ edges are labeled according to the letters seen in a product of g commutators.

This means the commutator length of a is the smallest g for which there exists a continuous map $\Sigma_{g,1} \to X$ taking $\partial\Sigma_{g,1}$ to α. Note that, if α bounds an oriented surface in X, then it must be trivial when regarded as an element of $H_1(X;\mathbb{Z}) = G/[G,G]$. One may think of $\mathrm{cl}(a)$ as measuring the complexity of this triviality in homology.

From this topological perspective, it is reasonable to also consider finite collections of loops in X and to allow surfaces with multiple boundary components. Let $\alpha_1, \ldots, \alpha_m$ be loops in X, and let $a_1, \ldots, a_m$ be the corresponding elements of G. Then the commutator length of the formal sum of the a_i is the smallest genus of a surface $\Sigma_{g,m}$ with m boundary components such that there is a map $\Sigma_{g,m} \to X$ taking the boundary components of $\Sigma_{g,m}$ to the α_i. This can also be stated algebraically, by defining

$$\mathrm{cl}\left(\sum_{i=1}^{m} a_i\right) = \min_{t_i \in G} \mathrm{cl}\left(\prod_{i=1}^{m} t_i a_i t_i^{-1}\right).$$

If such a surface does not exist, i.e. if the product of the a_i is not in $[G,G]$, we say that $\mathrm{cl}\left(\sum_{i=1}^{m} a_i\right) = \infty$.

Commutator length is a notoriously difficult quantity to compute in general, even for finite groups. For example, Ore [Ore51] conjectured that every element of a finite nonabelian simple group is a commutator, a result that was only proven recently [LOST10] after receiving much attention. Free groups are one of the few classes of groups in which commutator length can be computed. Edmunds [Edm75, Edm79] found an effective procedure for computing commutator length in free groups using cancellation arguments. Culler [Cul81] later showed how to compute commutator length in free groups using a geometric approach.

Even in free groups, understanding the general behavior of commutator length is difficult. There are many identities, such as $[b,c]^3 = [bcb^{-1}, c^{-1}bcb^{-2}][c^{-1}bc, c^2]$, which may be surprising at first, and these contribute to the complexity of the study of commutator length. There are, however, a few families of elements on which commutator length is well understood. For example, Culler [Cul81] shows that, in the free group F_2 on two generators b and c, $\mathrm{cl}([b,c]^n) = \left\lfloor \frac{n}{2} \right\rfloor + 1$.

2.2 Stable commutator length

Because of the difficulty of computing commutator length, we focus instead on the related notion of stable commutator length, which, roughly speaking, measures the part of commutator length that is preserved under taking powers. Stable commutator length turns out to be both more tractable and a richer notion than commutator length, due largely to connections with subjects such as hyperbolic geometry and bounded cohomology. One of these connections, with the theory of quasimorphisms, will be discussed in Section 3.4.

For $a \in [G,G]$, define the *stable commutator length* of a to be

$$\mathrm{scl}(a) = \lim_{n\to\infty} \frac{cl(a^n)}{n}.$$

The sequence $\mathrm{cl}(a^n)$ is subadditive, which implies, by Fekete's lemma, that this limit always exists. If $a \notin [G,G]$, but $a^n \in [G,G]$ for some $n \in \mathbb{N}$, instead define $\mathrm{scl}(a) = \mathrm{scl}(a^n)/n$. If $a^n \notin [G,G]$ for any $n \in \mathbb{N}$, we use the convention that $\mathrm{scl}(a) = \infty$.

Stable commutator length was extended to finite formal sums of elements by Calegari (see [Cal09b]). For a finite collection of elements $a_i \in G$ whose product is in $[G,G]$, define

$$\mathrm{scl}\left(\sum_{i=1}^{m} a_i\right) = \lim_{n\to\infty} \frac{\mathrm{cl}\left(\sum_{i=1}^{m} a_i^n\right)}{n}.$$

If $a_1 \cdots a_m \notin [G,G]$, but $a_1^n \cdots a_m^n \in [G,G]$ for some $n \in \mathbb{N}$, define

$$\mathrm{scl}\left(\sum_{i=1}^{m} a_i\right) = \frac{\mathrm{scl}\left(\sum_{i=1}^{m} a_i^n\right)}{n}.$$

If $a_1^n \cdots a_m^n \notin [G,G]$ for any $n \in \mathbb{N}$, we use the convention that $\mathrm{scl}\left(\sum_{i=1}^{m} a_i\right) = \infty$.

As an example, in a finite group, or any group with finite commutator width, stable commutator length is identically 0 on the commutator subgroup since there is a uniform bound on commutator length. In the free group F_2 generated by b and c, the result of Culler stated at the end of Section 2.1 implies that

$$\mathrm{scl}([b,c]) = \lim_{n\to\infty} \frac{\left\lfloor \frac{n}{2} \right\rfloor + 1}{n} = \frac{1}{2}.$$

Using the geometric interpretation of commutator length given in Section 2.1, one can understand stable commutator length in terms of the genus of surfaces Σ with one boundary component whose boundary wraps multiple times around a loop $\alpha\colon S^1 \to X$ corresponding to a. More precisely, $\mathrm{scl}(a)$ is the infimum of $\mathrm{genus}(\Sigma)/n(\Sigma)$ over all maps $\Sigma \to X$ taking $\partial\Sigma$ to a degree $n(\Sigma)$ cover of α. A deficiency in this approach is that this infimum will never be achieved, because $\mathrm{genus}(\Sigma)/n(\Sigma)$ can always be made arbitrarily close to $-\chi(\Sigma)/2n(\Sigma)$ by passing to finite covers, where $\chi(\Sigma)$ denotes the Euler characteristic of Σ.

Since Euler characteristic is multiplicative under taking covers, however, it is more natural to consider $\chi(\Sigma)$ rather than $\mathrm{genus}(\Sigma)$. Call a map $f\colon \Sigma \to X$ *admissible* if there is a commutative diagram

$$\begin{array}{ccc} \partial\Sigma & \xrightarrow{\iota} & \Sigma \\ {\scriptstyle \partial f}\big\downarrow & & \big\downarrow{\scriptstyle f} \\ S^1 & \xrightarrow[\alpha]{} & X \end{array}$$

Let $n(\Sigma)$ denote the degree of the map $\partial f\colon \partial\Sigma \to S^1$. Define $\chi^-(\Sigma) = \min(\chi(\Sigma), 0)$. Using the relationship $\chi(\Sigma) = 2 - 2\,\mathrm{genus}(\Sigma)$, one obtains that

$$\mathrm{scl}(a) = \inf_{\Sigma} \frac{-\chi^-(\Sigma)}{2n(\Sigma)},$$

where the infimum is taken over all admissible maps. This infimum may or may not be achieved, but when it is achieved surfaces achieving it are of particular interest. An admissible map $f\colon \Sigma \to X$ that achieves the above infimum is said to give an *extremal surface* for α. We remark that extremal surfaces are always π_1-injective, meaning the induced map $f_*\colon \pi_1(\Sigma) \to G$ is injective whenever f gives an extremal surface.

For a finite collection of loops $\alpha_i\colon S^1 \to X$, say that a map $f\colon \Sigma \to X$ is *admissible* if there is a commutative diagram

$$\begin{array}{ccc} \partial\Sigma & \xrightarrow{\iota} & \Sigma \\ {\scriptstyle \partial f}\big\downarrow & & \big\downarrow{\scriptstyle f} \\ \coprod_{i=1}^m S^1 & \xrightarrow[\coprod_{i=1}^m \alpha_i]{} & X \end{array}$$

such that the map $\partial f\colon \partial\Sigma \to \coprod_{i=1}^m S^1$ is a covering map. Let $n(\Sigma)$ denote the degree of this cover. Then, if a_i is the element of G corresponding to the loop α_i, one has that

$$\mathrm{scl}\left(\sum_{i=1}^{m} a_i\right) = \inf_{\Sigma} \frac{-\chi^-(\Sigma)}{2n(\Sigma)},$$

where the infimum is again taken over all admissible maps. Surfaces realizing this infimum are also called *extremal*.

Stable commutator length may be extended by linearity to $B_1(G;\mathbb{Q})$, the space of rational chains of elements of G that are trivial in $H_1(G;\mathbb{Q})$. It is subadditive on this space, and therefore extends continuously to $B_1(G;\mathbb{R})$, the space of real chains of elements of G that are trivial in $H_1(G,\mathbb{R})$. Stable commutator length gives a pseudo-norm on $B_1(G;\mathbb{R})$. Since it always vanishes on elements of the form $a - bab^{-1}$ and $a^n - na$ for $a, b \in G$ and $n \in \mathbb{Z}$, we form the subspace $H(G) \subseteq B_1(G;\mathbb{R})$ generated by elements of the form $a - bab^{-1}$ and $a^n - na$ and consider the quotient $B_1^H(G;\mathbb{R}) := B_1(G;\mathbb{R})/H(G)$. Stable commutator length descends to again give a pseudo-norm on $B_1^H(G;\mathbb{R})$. The advantage of doing this is that in some groups, such as hyperbolic groups, stable commutator length is a genuine norm on $B_1^H(G;\mathbb{R})$; see [CF10b].

2.3 Stable commutator length in free groups

Stable commutator length can be computed in relatively few groups, but Calegari [Cal09c] has an algorithm for computing it in free groups. Specifically, he considers the unit ball of $B_1^H(F_n;\mathbb{R})$ with respect to the stable commutator length norm, namely

$$\{C \in B_1^H(F_n;\mathbb{R}) : \mathrm{scl}(C) = 1\}.$$

Calegari shows that this ball is a rational polyhedron, meaning that its vertices are elements of $B_1^H(F_n;\mathbb{Q})$. He proves this by showing how to explicitly construct extremal surfaces bounding chains in $B_1(F_n;\mathbb{Z})$. A consequence of this result is that stable commutator length is a piecewise linear rational function on finite dimensional rational subspaces of $B_1^H(F_n;\mathbb{R})$, which implies that stable commutator length takes only rational values on $B_1^H(F_n;\mathbb{Q})$. In the course of the proof, Calegari obtains a polynomial-time algorithm for computing stable commutator length in free groups. This algorithm has been implemented in the computer program `scallop` [CWb].

Chapter 3

Quasimorphisms

In this chapter, we introduce the notion of quasimorphism and give some examples. Then we discuss a relationship between the theory of homogeneous quasimorphisms and stable commutator length, first discovered by Bavard [Bav91], that is fundamental to our work.

3.1 Definitions

A *quasimorphism* is a real-valued function on a group that fails to be a homomorphism by a bounded amount. More precisely, a function $\phi\colon G \to \mathbb{R}$ is a quasimorphism if it satisfies the property that

$$|\phi(ab) - \phi(a) - \phi(b)| \leq D$$

for some constant D that is independent of the choice of $a, b \in G$. Choose the smallest such D and denote it by $D(\phi)$, referred to as the *defect* of ϕ. One may think of $D(\phi)$ as measuring the amount by which ϕ fails to be a homomorphism.

We will primarily be concerned with homogeneous quasimorphisms. A quasimorphism ϕ is called *homogeneous* if it is a homomorphism on cyclic subgroups, i.e. if $\phi(a^n) = n\phi(a)$ for all $a \in G, n \in \mathbb{Z}$. As many naturally occurring quasimorphisms are not homogeneous, it is useful to note that any quasimorphism ϕ may be *homogenized* to give a homogeneous quasimorphism $\widetilde{\phi}$ defined by

$$\widetilde{\phi}(a) = \lim_{n\to\infty} \frac{\phi(a^n)}{n}.$$

A homogeneous quasimorphism is a class function, i.e. is constant on conjugacy classes. Homogeneous quasimorphisms also have the property that their defect is equal to the largest value taken on a commutator. More precisely, whenever ϕ is a homogeneous quasimorphism,

$$\sup_{b,c\in G} |\phi([b,c])| = D(\phi).$$

Let $Q(G)$ denote the vector space of homogeneous quasimorphisms on G. The defect gives a pseudo-norm on $Q(G)$, vanishing on the space $H^1(G;\mathbb{R})$ of homomorphisms $G \to \mathbb{R}$. It gives a norm on the quotient space $Q(G)/H^1(G;\mathbb{R})$. Observe that homomorphisms $G \to \mathbb{R}$ always take the value 0 on commutators. This means that, if $\phi\colon G \to \mathbb{R}$ is a homomorphism, we must have $\phi(a) = 0$ for all a satisfying $a^n \in [G,G]$ for some $n \in \mathbb{Z}$. As a result, an equivalence class in $Q(G)/H^1(G;\mathbb{R})$ takes a well-defined value on a whenever $a^n \in [G,G]$ for some $n \in \mathbb{Z}$.

Note that a quasimorphism ϕ on G can be extended by linearity to $B_1(G;\mathbb{R})$. In other words, one simply defines

$$\phi\left(\sum_{i=1}^m t_i a_i\right) := \sum_{i=1}^m t_i \phi(a_i).$$

Since (extensions of) elements of $H^1(G;\mathbb{R})$ take the value 0 on chains in $B_1(G;\mathbb{R})$, we also have that equivalence classes in $Q(G)/H^1(G;\mathbb{R})$ take well-defined values on chains in $B_1(G;\mathbb{R})$.

3.2 Rotation number

Poincaré [Poi81, Poi82] defined perhaps the first example of a quasimorphism in his study of homeomorphisms of the circle. Given an element of $\mathrm{Homeo}^+(S^1)$, the group of orientation-preserving homeomorphisms of the circle, one can try to define how much this homeomorphism rotates the circle, though this can only be defined up to adding an integer value. In order to avoid this ambiguity, instead consider

$$\widetilde{\mathrm{Homeo}}^+(S^1) := \{f \in \mathrm{Homeo}^+(\mathbb{R})\colon f(x+1) = f(x) \text{ for all } x \in \mathbb{R}\}.$$

Regarding S^1 as the unit interval $[0,1]$ with endpoints identified, the natural projection $\mathbb{R} \to S^1$ makes $\widetilde{\mathrm{Homeo}}^+(S^1)$ into a central extension of $\mathrm{Homeo}^+(S^1)$, i.e. there is an exact sequence

$$0 \to \mathbb{Z} \to \widetilde{\mathrm{Homeo}}^+(S^1) \to \mathrm{Homeo}^+(S^1) \to 1,$$

where $\mathbb{Z}$ is generated by unit translation of $\mathbb{R}$.

We define $\mathrm{rot}\colon \widetilde{\mathrm{Homeo}}^+(S^1) \to \mathbb{R}$ by setting

$$\mathrm{rot}(f) = \lim_{n\to\infty} \frac{f^n(0)}{n}.$$

This is a homogeneous quasimorphism with defect 1. Given an element $h \in \mathrm{Homeo}^+(S^1)$, choose an arbitrary lift $\widetilde{h} \in \widetilde{\mathrm{Homeo}}^+(S^1)$. Then the rotation number of h is $\mathrm{rot}\left(\widetilde{h}\right) \in \mathbb{R}/\mathbb{Z}$. This construction can also be used to define quasimorphisms on several other groups, and some of these generalizations are discussed in Chapter 5.

3.3 Counting quasimorphisms

Other basic examples of quasimorphisms are the counting quasimorphisms introduced by Rhemtulla [Rhe68] and Brooks [Bro81]. Consider the free group F_n on n letters, and fix a reduced word w in these letters (and their inverses). Define a function $C_w\colon F_n \to \mathbb{Z}$, called a *big counting function*, by setting $C_w(a)$ equal to the number of (possibly overlapping) copies of w in the reduced representative of a. While C_w is not a quasimorphism, one can make it into a quasimorphism by also taking into account appearances of w^{-1} in a. Define another function $H_w\colon F_n \to \mathbb{Z}$ by setting $H_w(a) = C_w(a) - C_{w^{-1}}(a)$. Then H_w defines a quasimorphism on F_n, called a *big counting quasimorphism*.

A slight variant of this construction was introduced by Epstein–Fujiwara [EF97]. Define another function $c_w\colon F_n \to \mathbb{Z}$, called a *small counting function*, by setting $c_w(a)$ equal to the maximal number of *disjoint* copies of w in the reduced representative of a. Then define another quasimorphism $h_w\colon F_n \to \mathbb{Z}$, known as a *small counting quasimorphism*, by the formula $h_w(a) = c_w(a) - c_{w^{-1}}(a)$. Epstein–Fujiwara also generalized little counting quasimorphisms to arbitrary hyperbolic groups. This construction was further generalized to mapping class groups by Bestvina–Feighn [BF02, BF07], and more recently to outer automorphism groups of free groups by Hamenstädt [Ham].

3.4 Bavard duality

A fundamental connection between the theory of quasimorphisms and stable commutator length is due to Bavard [Bav91]. He showed that, for all $a \in [G, G]$,

$$\mathrm{scl}(a) = \sup_{\substack{\phi \in Q(G)/H^1(G;\mathbb{R}) \\ D(\phi) \neq 0}} \frac{\phi(a)}{2D(\phi)},$$

a result we refer to as *Bavard duality.* Moreover, it is known that this supremum is always achieved (see [Cal09b]). Therefore it is of interest to find homogeneous quasimorphisms that achieve this supremum. A homogeneous quasimorphism ϕ that achieves this supremum for some a, i.e. satisfies $\mathrm{scl}(a) = \phi(a)/2D(\phi)$, is said to be *extremal* for a.

Bavard duality was extended to $B_1(G;\mathbb{R})$ by Calegari (see [Cal09b]). Given an element $\sum_{i=1}^m t_i a_i \in B_1(G;\mathbb{R})$, he shows that

$$\mathrm{scl}\left(\sum_{i=1}^m t_i a_i\right) = \sup_{\substack{\phi \in Q(G)/H^1(G;\mathbb{R}) \\ D(\phi) \neq 0}} \frac{\sum_{i=1}^m t_i \phi(a_i)}{2D(\phi)}.$$

A homogeneous quasimorphism that achieves this supremum is said to be *extremal* for the chain $\sum_{i=1}^m t_i a_i$, and extremal quasimorphisms are known to exist for all chains in $B_1(G;\mathbb{R})$.

3.5 Questions

In attempting to study stable commutator length and quasimorphisms in a particular group, one can ask the following two complementary questions.

Question 1. *Given a chain in $B_1(G;\mathbb{R})$, which homogeneous quasimorphisms are extremal for it?*

As mentioned in Section 3.4, extremal quasimorphisms always exist, and therefore the space of homogeneous quasimorphisms extremal for a given chain in $B_1(G;\mathbb{R})$ is always nonempty. However, the size of this space is not well understood. When is there a unique extremal quasimorphism in $Q(G)/H^1(G;\mathbb{R})$? When are there infinitely many linearly independent extremal quasimorphisms?

Question 2. *Given a homogeneous quasimorphism, for which chains in $B_1(G;\mathbb{R})$ is it extremal?*

In general, a homogeneous quasimorphism need not be extremal for any chains in $B_1(G;\mathbb{R})$. If, however, a homogeneous quasimorphism ϕ realizes its defect, i.e. there are elements $a,b \in G$ such that $\phi(ab)-\phi(a)-\phi(b) = D(\phi)$, then ϕ is extremal for an integral chain. This is because the quantity

$$\frac{\phi(ab)-\phi(a)-\phi(b)}{2D(\phi)}$$

can never exceed $1/2$, and therefore when it equals $1/2$ one has that $\mathrm{scl}(ab-a-b) = 1/2$ and ϕ is extremal for the integral chain $ab-a-b$. This happens, for example, when ϕ takes discrete values, which is the case for many known constructions of quasimorphisms.

Note that in this case ϕ is extremal for an *integral* chain, rather than simply for a chain in $B_1(G;\mathbb{R})$. Whenever a homogeneous quasimorphism is extremal for a chain in $B_1(G;\mathbb{Q})$, it must be extremal for a integral chain, simply by multiplying through by a constant to clear denominators. However, it is an interesting question to ask when a homogeneous quasimorphism is extremal for a rational/integral chain.

Question 3. *If a homogeneous quasimorphism is extremal for some chain in $B_1(G;\mathbb{R})$, must it be extremal for a chain in $B_1(G;\mathbb{Q})$?*

As explained above, the answer to this question is yes for a homogeneous quasimorphism that achieves its defect. The answer is also yes if G is a virtually free group and ϕ is a homogeneous quasimorphism that is extremal for a chain $C \in B_1(G;\mathbb{R})$. This is because C is contained in a finite dimensional rational subspace $V \subseteq B_1^H(G;\mathbb{R})$. The restriction of stable commutator length to V is a piecewise linear rational function (see [Cal09c]), and ϕ restricted to V is a linear function. The condition that ϕ is extremal for C implies that some level set $\phi = 2D(\phi)$ contains a face of the unit stable commutator length ball in V containing $C/\mathrm{scl}(C)$, and this face also contains rational chains.

In this thesis, we study Question 2 for the modular group $\mathrm{PSL}(2,\mathbb{Z})$ and the rotation quasimorphism.

Chapter 4

The modular group

In the following chapters, we restrict attention to a particular group of interest in many areas of mathematics, the modular group $\mathrm{PSL}(2,\mathbb{Z})$. In this chapter, we discuss several fundamental properties of the modular group and its action on the hyperbolic plane.

4.1 Definitions

Let $\mathbb{H}^2$ be the *hyperbolic plane*. We will typically think of $\mathbb{H}^2$ in terms of the Poincaré upper half-plane model, consisting of the points $\{z \in \mathbb{C} : \mathrm{Im}(z) > 0\}$ together with the metric $ds^2 = (dx^2 + dy^2)/y^2$, where $z = x + iy$. Geodesics in $\mathbb{H}^2$ consist of circular arcs perpendicular to the real axis and of vertical lines. Orientation-preserving isometries of $\mathbb{H}^2$ are Möbius transformations of the form

$$z \mapsto \frac{az+b}{cz+d},$$

where $a, b, c, d \in \mathbb{R}$ and $ad - bc = 1$. Such a transformation can also be thought of as a pair of matrices $\pm\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$, and composition of Möbius transformations corresponds to matrix multiplication. Therefore the group of orientation-preserving isometries of $\mathbb{H}^2$ is

$$\mathrm{PSL}(2,\mathbb{R}) \cong \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} : a, b, c, d \in \mathbb{R}, ad - bc = 1 \right\} / \{\pm I\}.$$

We will interchangeably use $A \in \mathrm{PSL}(2,\mathbb{R})$ to refer either to a Möbius transformation or to a matrix, with the understanding that as a matrix A is equivalent to $-A$.

The *modular group* $\mathrm{PSL}(2,\mathbb{Z})$ is the discrete subgroup of $\mathrm{PSL}(2,\mathbb{R})$ consisting of those elements of $\mathrm{PSL}(2,\mathbb{R})$ with $a,b,c,d \in \mathbb{Z}$. In other words,

$$\mathrm{PSL}(2,\mathbb{Z}) \cong \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} : a,b,c,d \in \mathbb{Z}, ad - bc = 1 \right\} / \{\pm I\}.$$

There are many possible generating sets for $\mathrm{PSL}(2,\mathbb{Z})$, but we prefer the generators $S = \left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right)$ and $U = \left(\begin{smallmatrix} 0 & -1 \\ 1 & 1 \end{smallmatrix}\right)$. The element S acts on $\mathbb{H}^2$ by a rotation of angle π about the point corresponding to the complex number i, and the element U acts on $\mathbb{H}^2$ by a rotation of angle $-2\pi/3$ about the point $(-1+\sqrt{3}i)/2$. The relations $S^2 = I$ and $U^3 = I$ generate all relations between S and U, meaning $\mathrm{PSL}(2,\mathbb{Z})$ has presentation

$$\langle S, U \mid S^2 = U^3 = 1 \rangle.$$

Therefore $\mathrm{PSL}(2,\mathbb{Z})$ is isomorphic to $\mathbb{Z}/2\mathbb{Z} * \mathbb{Z}/3\mathbb{Z}$, the free product of the group of order 2 and the group of order 3.

4.2 Classification of elements of the modular group

Elements of the modular group may be classified into various types, based either on algebraic information (their trace) or geometric information (fixed points in their action on $\mathbb{H}^2$). There are three distinct types of nonidentity elements $A \in \mathrm{PSL}(2,\mathbb{Z})$: finite order, parabolic, and hyperbolic. The type of an element may be determined as follows:

1. Finite order elements are those with $|\mathrm{tr}(A)| < 2$. These elements are characterized by the property that they fix a point in the interior of $\mathbb{H}^2$.

2. Parabolic elements are those with $|\mathrm{tr}(A)| = 2$. These elements are characterized by the property that they fix exactly one point on the boundary of $\mathbb{H}^2$.

3. Hyperbolic elements are those with $|\mathrm{tr}(A)| > 2$. These elements are characterized by the property that they fix exactly two points on the boundary of $\mathbb{H}^2$.

In terms of this classification, one can describe the way an element of the modular group acts on $\mathbb{H}^2$:

1. Finite order elements act by rotation about their fixed point.

2. Parabolic elements act by a "limit rotation" about the fixed point on the boundary.

3. Hyperbolic elements act by translation along the geodesic determined by the two fixed points on the boundary.

4.3 The modular surface

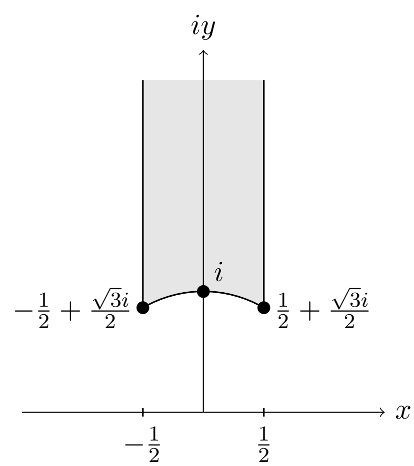

Figure 4.1: A fundamental domain for the action of $\mathrm{PSL}(2, \mathbb{Z})$ on $\mathbb{H}^2$

A fundamental domain of the action of $\mathrm{PSL}(2, \mathbb{Z})$ on $\mathbb{H}^2$ consists of the region

$$\left\{ z \in \mathbb{C} : |z| \geq 1, |\mathrm{Re}\, z| \leq \frac{1}{2} \right\},$$

shown in Figure 4.1. Under the action of $\mathrm{PSL}(2, \mathbb{Z})$, the vertical lines on the left and right of this region are identified with each other under the translation $z \mapsto z + 1$, and the circular arc at the bottom of the region is identified with itself under the transformation $z \mapsto -1/z$. The quotient of $\mathbb{H}^2$ by the action of $\mathrm{PSL}(2, \mathbb{Z})$ is a triangle orbifold of type $(2, 3, \infty)$, traditionally referred to as the *modular surface*, and shown in Figure 4.2. The cone point of order 2 corresponds to the point i in the fundamental domain, and the cone point

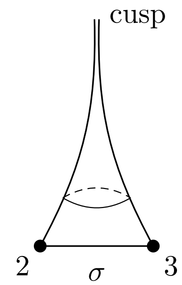

Figure 4.2: The modular surface $\mathbb{H}^2/\operatorname{PSL}(2,\mathbb{Z})$

of order 3 corresponds to the points $(\pm 1+\sqrt{3}i)/2$ in the fundamental domain, which are identified in the quotient.

This allows one to realize $\operatorname{PSL}(2,\mathbb{Z})$ as the *orbifold* fundamental group of the modular surface in a natural way. Here we think of the orbifold fundamental group as the ordinary fundamental group of the thrice punctured sphere with the specification that a loop around the order 2 cone point has order 2 and a loop around the order 3 cone point has order 3. Examining the action of the elements S and U on $\mathbb{H}^2$, one sees that S corresponds to a loop around the order 2 cone point and U corresponds to a clockwise loop around the order 3 cone point. Since S and U generate $\operatorname{PSL}(2,\mathbb{Z})$, one can appropriately concatenate these loops to obtain a loop corresponding to any element of the modular group. For a hyperbolic element $A \in \operatorname{PSL}(2,\mathbb{Z})$, one can also obtain a representative of the corresponding homotopy class by projecting the translation axis of A to the modular surface. We typically prefer this second construction because it gives a geodesic representative of the homotopy class. Closed geodesics on the modular surface have been extensively studied; see Series [Ser85] and Katok–Ugarcovici [KU07].

4.4 The Farey graph

The action of the modular group $\operatorname{PSL}(2,\mathbb{Z})$ on the hyperbolic plane can be understood combinatorially in terms of the induced action on the Farey graph. Let σ be the geodesic in the modular surface between the order 2 and order 3 cone points, as labeled in Figure 4.2. The *Farey graph* is $\widetilde{\sigma}$, the total preimage of σ in $\mathbb{H}^2$. It consists of the arc between the points $(-1+\sqrt{3}i)/2$ and $(1+\sqrt{3}i)/2$ along the boundary of the fundamental domain shown

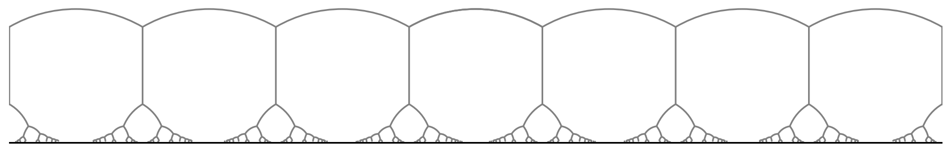

Figure 4.3: The Farey graph

in Figure 4.1 as well as all its translates under the action of $\mathrm{PSL}(2,\mathbb{Z})$. The Farey graph is a regular 3-valent tree, shown in Figure 4.3.

The Farey graph is dual to the *Farey tessellation* of the hyperbolic plane. Consider the ideal triangle in $\mathbb{H}^2$ with vertices 0, 1, and ∞. The tiles of the Farey tessellation consist of all translates of this triangle under the action of elements of $\mathrm{PSL}(2,\mathbb{Z})$. More explicitly, the ideal vertices of the Farey tessellation are all points $\mathbb{Q} \cup \{\infty\}$. Two points p/q and p'/q' (in lowest terms) are joined by a geodesic edge if and only if $|pq' - qp'| = 1$, where we regard $1/0$ as the representative of the point ∞ and consider an edge from a point to infinity to be a vertical line from that point.

Chapter 5

The rotation quasimorphism

In the following chapters, we restrict attention to a particular homogeneous quasimorphism on the modular group, the rotation quasimorphism. This quasimorphism can be defined in several equivalent ways, including as the homogenization of the Rademacher function, which has been widely studied in number theory and itself has numerous interpretations.

5.1 Definitions

The rotation number on $\widetilde{\mathrm{Homeo}}^+(S^1)$ defined in Section 3.2 may be extended to (central extensions of) other groups that act on S^1. The action of an element of $\mathrm{PSL}(2,\mathbb{R})$ on the hyperbolic plane $\mathbb{H}^2$ induces an orientation-preserving homeomorphism of the boundary S^1_∞ of $\mathbb{H}^2$. This defines an injective homomorphism $\mathrm{PSL}(2,\mathbb{R}) \to \mathrm{Homeo}^+(S^1)$, and so we regard $\mathrm{PSL}(2,\mathbb{R})$ as a subgroup of $\mathrm{Homeo}^+(S^1)$. Let $\widetilde{\mathrm{PSL}}(2,\mathbb{R})$ be the preimage of $\mathrm{PSL}(2,\mathbb{R})$ under the projection $\widetilde{\mathrm{Homeo}}^+(S^1) \to \mathrm{Homeo}^+(S^1)$. We get the following commutative diagram.

$$\begin{tikzcd}
1 \arrow[r] & \mathbb{Z} \arrow[r] \arrow[d, "\cong"'] & \widetilde{\mathrm{PSL}}(2,\mathbb{R}) \arrow[r] \arrow[d] & \mathrm{PSL}(2,\mathbb{R}) \arrow[r] \arrow[d] & 1 \\
1 \arrow[r] & \mathbb{Z} \arrow[r] & \widetilde{\mathrm{Homeo}}^+(S^1) \arrow[r] & \mathrm{Homeo}^+(S^1) \arrow[r] & 1
\end{tikzcd}$$

By precomposing with the map $\widetilde{\mathrm{PSL}}(2,\mathbb{R}) \to \widetilde{\mathrm{Homeo}}^+(S^1)$, Poincaré's rotation number extends to the group $\widetilde{\mathrm{PSL}}(2,\mathbb{R})$, and we call the resulting quasimorphism a rotation quasimorphism. It restricts to give a quasimorphism on (central extensions of) subgroups of $\mathrm{PSL}(2,\mathbb{R})$, such as $\widetilde{\mathrm{PSL}}(2,\mathbb{Z})$.

In view of Bavard's duality theorem, discussed in Section 3.4, Calegari [Cal09a] gives a far more general definition of rotation quasimorphism. For (virtually) free groups, he studies the unit ball of the stable commutator length norm on the space $B_1^H(G;\mathbb{R})$. He shows that codimension 1 faces of this ball correspond to realizations of G as the fundamental group of a surface (or orbifold). Bavard's duality theorem then says that there is a unique homogeneous quasimorphism of defect 1 in $Q(G)/H^1(G;\mathbb{R})$ dual to each such realization. The modular group $\mathrm{PSL}(2,\mathbb{Z})$ may be naturally identified with the orbifold fundamental group of the modular surface, and we call the unique homogeneous quasimorphism of defect 1 dual to this realization the *rotation quasimorphism*.

5.2 Significance of the rotation quasimorphism

The rotation quasimorphism is a natural one to study, for the following reason. Let G be a (virtually) free group, and let X be a space with $\pi_1(X) = G$. Let $C \in B_1(G;\mathbb{Q})$ be a rational chain. Suppose ϕ is a homogeneous quasimorphism that is extremal for C, and suppose C admits an extremal surface Σ. Then the map $\Sigma \to X$ induces a map $\pi_*\colon \pi_1(\Sigma) \to G$. The composition $\phi \circ \pi_*\colon \pi_1(\Sigma) \to \mathbb{R}$ defines a homogeneous quasimorphism on $\pi_1(\Sigma)$, and such a quasimorphism is always an extension of a rotation quasimorphism. This special role of rotation quasimorphisms is why we choose to study them. The rotation quasimorphism on the modular group $\mathrm{PSL}(2,\mathbb{Z})$ is also important because of its connection with several other functions, and some of these connections will be discussed in Section 5.4.

5.3 Computing the rotation quasimorphism

Despite the apparent difficulty of computing values of the rotation quasimorphism from the definition given in Section 5.1, it has a simple formula in the modular group $\mathrm{PSL}(2,\mathbb{Z})$, as we explain in this section. We use both the generators S and U given in Chapter 4 and the generators $L = SU = \left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right)$ and $R = SU^{-1} = \left(\begin{smallmatrix} 1 & 0 \\ 1 & 1 \end{smallmatrix}\right)$. The elements L and R are so denoted because they correspond to left and right turns in the action of $\mathrm{PSL}(2,\mathbb{Z})$ on the Farey graph. Specifically, if W is a positive word in L and R, then the path from the complex number i to $W(i)$ in the Farey graph turns left and right according to the appearances of L

and R. For example, the path from i to $L^2RL(i)$ in the Farey graph turns left, left, right, and left.

Using the presentation of $\mathrm{PSL}(2,\mathbb{Z})$ given in Chapter 4, any element $A \in \mathrm{PSL}(2,\mathbb{Z})$ may be written uniquely in the form

$$S^{\delta_i}U^{\epsilon_1}SU^{\epsilon_2}\cdots SU^{\epsilon_m}S^{\delta_2},$$

where each δ_i is either 0 or 1 and each ϵ_i is either -1 or 1. Given an element of $\mathrm{PSL}(2,\mathbb{Z})$, there is a standard way to obtain its expression in this form, as we now explain. Let $A = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathrm{PSL}(2,\mathbb{Z})$. Observe that left multiplication by L, L^{-1}, R, and R^{-1} corresponds to doing row operations on A and that left multiplication by S interchanges the rows of A (and multiplies one of them by -1). If $c = 0$, then $ad = 1$, and so $a = d = \pm 1$. This means $A = L^{ba/|a|}$. If $c \neq 0$, do row operations so as to perform the Euclidean algorithm on a and c until $c = 0$, keeping track of the corresponding matrices. In terms of the entries of the reduced matrix, the result of performing the Euclidean algorithm is $L^{ba/|a|}$. Solving for A then gives an expression for it in terms of L, L^{-1}, R, R^{-1}, and S. Using $L = SU$, $L^{-1} = U^{-1}S$, $R = SU^{-1}$, and $R^{-1} = US$ expresses A in terms of S, U, and U^{-1}.

Define a function $\phi\colon\ \mathrm{PSL}(2,\mathbb{Z}) \to \mathbb{Z}$ by setting

$$\phi(A) = \sum_{n=1}^{m} \epsilon_i.$$

This defines a quasimorphism on $\mathrm{PSL}(2,\mathbb{Z})$, and its homogenization is (up to a scalar multiple) the rotation quasimorphism.

The homogenization of this function can be understood more easily by conjugating elements of $\mathrm{PSL}(2,\mathbb{Z})$ to a standard form. All homogeneous quasimorphisms take the value 0 on finite order elements, so we consider only infinite order elements. Every infinite order element of $\mathrm{PSL}(2,\mathbb{Z})$ is conjugate to a positive word in L and R, and this is unique up to cyclic permutation, as we now explain. Suppose we have an infinite order element $A = S^{\delta_1}U^{\epsilon_1}SU^{\epsilon_2}\cdots SU^{\epsilon_m}S^{\delta_2}$. We show that A is conjugate to a word of this form with $\delta_1 = 1$ and $\delta_2 = 0$. If $\delta_2 = 1$, conjugate by S and simplify using $SS = I$. If $\delta_1 = \delta_2 = 0$ and

$\epsilon_1 = 1$, conjugate by U^{-1}. If $\delta_1 = \delta_2 = 0$ and $\epsilon_1 = -1$, conjugate by U. Then simplify as much as possible using $UU^{-1} = U^{-1}U = I$, $UU = U^{-1}$, and $U^{-1}U^{-1} = U$. Repeat these steps until $\delta_1 = 1$ and $\delta_2 = 0$. For every word of length at least 2 that is not already of the desired form, each pass through this algorithm either puts it in the correct form or shortens the length of the word, and therefore this process must terminate. We cannot be left with a single letter S, U, or U^{-1} since in this case A would have had finite order, which we assumed was not the case. Once we have $\delta_1 = 1$ and $\delta_2 = 0$, replace occurrences of SU with L and occurrences of SU^{-1} with R to get a positive word in L and R.

Let $\widetilde{\phi}$ denote the homogenization of ϕ. If W is a positive word in L and R, then $\phi(W^n) = n\phi(W)$ for all $n \in \mathbb{N}$, and hence $\widetilde{\phi}(W) = \phi(W)$. Since homogeneous quasimorphisms are constant on conjugacy classes, we may choose representatives of this form on which to compute $\widetilde{\phi}$. Given an arbitrary element $A \in \mathrm{PSL}(2, \mathbb{Z})$ of infinite order, conjugate A to get a word W in L and R. Then we have $\widetilde{\phi}(A) = \phi(W)$. It turns out $\widetilde{\phi}$ has defect 6, and $\mathrm{rot}(A) = -\widetilde{\phi}(A)/6$. Thus one may think of rot as counting the number of right turns minus the number of left turns in the action of an element on the Farey graph, appropriately scaled to make the defect 1.

5.4 The Rademacher function

The rotation quasimorphism is (up to a constant) the homogenization of the classical Rademacher function (see [KM94]). In number theory, the Rademacher function is often defined in terms of the Dedekind η-function, as follows. The Dedekind η-function is defined on $\mathbb{H}^2$ by

$$\eta(\tau) = e^{\pi i \tau/12} \prod_{n=1}^{\infty} (1 - e^{2\pi i n \tau}).$$

The 24th power of η is a modular form of weight 12, meaning that

$$\eta^{24}\left(\frac{a\tau + b}{c\tau + d}\right) = \eta^{24}(\tau)(c\tau + d)^{12}$$

for all matrices $A = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathrm{PSL}(2, \mathbb{Z})$.

Taking logarithms, we obtain

$$24(\log \eta)\left(\frac{a\tau + b}{c\tau + d}\right) = 24(\log \eta)(\tau) + 6\log(-(c\tau + d)^2) + 2\pi i \psi(A)$$

for some integer $\psi(A)$. Thus we obtain a function $\psi\colon \mathrm{PSL}(2, \mathbb{Z}) \to \mathbb{Z}$, and this is the *Rademacher function.* This function also arises in many other contexts, and Atiyah [Ati87] proves the equivalence of seven different definitions of it.

Yet another equivalent definition of the Rademacher function is given by Ghys [Ghy07]. Consider the quotient $\mathrm{PSL}(2, \mathbb{R})/\mathrm{PSL}(2, \mathbb{Z})$, which is homeomorphic to the complement of the trefoil knot in the 3-sphere. It may also be seen as the unit tangent bundle of the modular surface. There is a bijection between hyperbolic conjugacy classes of $\mathrm{PSL}(2, \mathbb{Z})$ and periodic orbits of the geodesic flow of the modular surface. Such a periodic orbit defines a knot in the complement of the trefoil knot, and the linking number of this knot with the trefoil knot is the value of the Rademacher function.

Chapter 6

Stable commutator length in the modular group

In this chapter, we explain how to compute stable commutator length in the modular group $\mathrm{PSL}(2,\mathbb{Z})$. We use a relationship between stable commutator length in a group and a finite index subgroup to reduce the problem to that of computing stable commutator length in the free group of rank 2, which allows us to use the algorithm mentioned in Section 2.3.

6.1 Stable commutator length in finite index subgroups

There is a relationship between stable commutator length in a group and a finite index subgroup, explained in [Cal09b]. Suppose H is a finite index subgroup of G. Let X be a topological space with $\pi_1(X) = G$. Then H corresponds to a finite degree cover $p\colon \widetilde{X} \to X$. Let $a_1, \ldots, a_n$ be elements of G whose formal sum is in $B_1(G, \mathbb{Q})$. Regard a_i as an element of $\pi_1(X)$, and let α_i be a loop in X representing this element. Let $\beta_1, \ldots, \beta_k$ be the total preimage of $\alpha_1, \ldots, \alpha_n$ in $\widetilde{X}$, and let h_i be the element of $H = \pi_1(\widetilde{X})$ corresponding to β_i. Then

$$\mathrm{scl}_G\left(\sum_{i=1}^{n} a_i\right) = \frac{1}{[G:H]}\,\mathrm{scl}_H\left(\sum_{i=1}^{k} h_i\right).$$

This can be shown using the definition of stable commutator length in terms of admissible maps. If there is an admissible map $d\colon \Sigma \to \widetilde{X}$ taking $\partial\Sigma$ to $\cup_i \beta_i$, then the composition $p \circ d$ maps Σ to X, taking $\partial\Sigma$ to $\cup_i \alpha_i$ with degree $[G : H]$. This shows that the left-hand side is less than or equal to the right-hand side. In the other direction, if there is an admissible map $f\colon \Sigma \to X$ taking $\partial\Sigma$ to $\cup_i \alpha_i$, we construct an appropriate admissible map to $\widetilde{X}$ as follows. Let K be a finite index subgroup of H that is normal in G. The map f induces a map $f_*\colon \pi_1(\Sigma) \to \pi_1(X) = G$, and composing this with the quotient map gives a map $\pi_1(\Sigma) \to G/K$. Let $\pi\colon \widetilde{\Sigma} \to \Sigma$ be the regular cover corresponding to the kernel of this map. Then the composition $f \circ \pi\colon \widetilde{\Sigma} \to X$ lifts to an admissible map $\widetilde{f}\colon \widetilde{\Sigma} \to \widetilde{X}$ taking $\partial\widetilde{\Sigma}$ to $\cup_i \beta_i$ and satisfying $p \circ \widetilde{f} = f \circ \pi$, as desired.

Since stable commutator length takes only rational values on free groups (see Section 2.3), this relationship implies that stable commutator length also takes only rational values on groups that are *virtually* free, i.e. those with a finite index free subgroup.

6.2 Principal congruence subgroups

The modular group $\mathrm{PSL}(2,\mathbb{Z})$ has many finite index subgroups, and perhaps first among these are the principal congruence subgroups. Let n be a positive integer, and consider the map

$$\mathrm{PSL}(2,\mathbb{Z}) \to \mathrm{PSL}(2,\mathbb{Z}/n\mathbb{Z})$$

given by reducing each matrix entry modulo n. The kernel of this map is denoted $\Gamma(n)$ and called the *principal congruence subgroup of level* n. These subgroups of $\mathrm{PSL}(2,\mathbb{Z})$ have been much studied by number theorists. We are interested in $\Gamma(2)$, the principal congruence subgroup of level 2. Since $\mathrm{PSL}(2,\mathbb{Z}/2\mathbb{Z})$ has order 6, $\Gamma(2)$ is an index 6 subgroup of $\mathrm{PSL}(2,\mathbb{Z})$. A fundamental domain for its action on $\mathbb{H}^2$ consists of six copies of a fundamental domain for the action of $\mathrm{PSL}(2,\mathbb{Z})$ on $\mathbb{H}^2$, as shown in Figure 6.1.

Pairs of edges of this fundamental domain going to the same point at infinity are identified under the action of $\Gamma(2)$, and the quotient $\mathbb{H}^2/\Gamma(2)$ is a thrice-punctured sphere. Since the fundamental group of a thrice-punctured sphere is F_2, the free group of rank 2, we have that $\Gamma(2) \cong F_2$. By computing the elements of $\mathrm{PSL}(2,\mathbb{Z})$ needed to identify corresponding edges in the fundamental domain of the action of $\Gamma(2)$ on $\mathbb{H}^2$, one finds that free generators of $\Gamma(2)$ are $\left(\begin{smallmatrix}1 & 2\\ 0 & 1\end{smallmatrix}\right)$ and $\left(\begin{smallmatrix}1 & 0\\ 2 & 1\end{smallmatrix}\right)$.

We use the relationship between stable commutator length in a group and a finite index subgroup given in the previous section to turn the problem of computing stable commutator

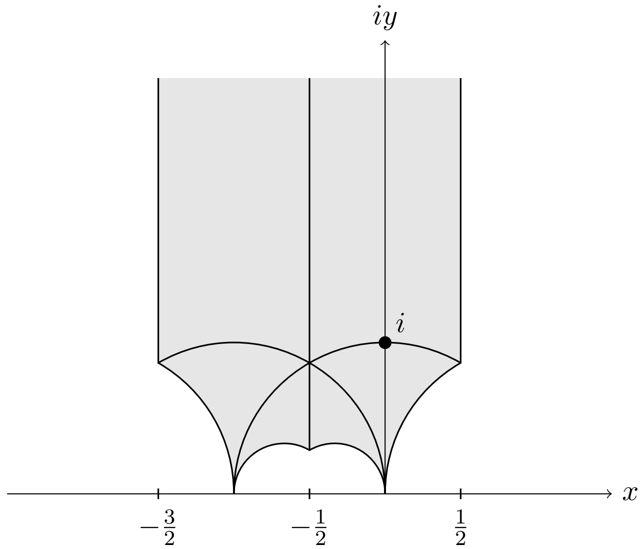

Figure 6.1: A fundamental domain for the action of $\Gamma(2)$ on $\mathbb{H}^2$

length in $\mathrm{PSL}(2,\mathbb{Z})$ into one of computing it in $\Gamma(2) \cong F_2$. Here, this relationship says that

$$\mathrm{scl}_{\mathrm{PSL}(2,\mathbb{Z})}(A) = \frac{1}{6}\,\mathrm{scl}_{\Gamma(2)}\left(\sum_{i=1}^{k} h_i\right),$$

where the elements h_i come from taking the total preimage of a loop corresponding to A in a degree 6 cover of the modular surface $\mathbb{H}^2/\,\mathrm{PSL}(2,\mathbb{Z})$ by the thrice-punctured sphere $\mathbb{H}^2/\Gamma(2)$. Since stable commutator length in F_2 can be computed using the program `scallop` [CWb], the problem of computing stable commutator length in $\mathrm{PSL}(2,\mathbb{Z})$ is thus reduced to finding a systematic way to determine the h_i.

6.3 The algorithm

In this section, we give an algorithm for explicitly determining the elements $h_i \in F_2$ corresponding to a given $A \in \mathrm{PSL}(2,\mathbb{Z})$. Recall that this involves finding the total preimage of a loop in the modular surface under a degree 6 cover by the thrice-punctured sphere.

Fix a basepoint p on the geodesic σ between the order 2 and order 3 cone points of the modular orbifold. In the fundamental domain for the action of $\mathrm{PSL}(2,\mathbb{Z})$ on $\mathbb{H}^2$, this corresponds to a pair of points on the unit circle that are identified by the order 2 element S, as shown in Figure 6.2, where the points in the pair are labeled p and p'.

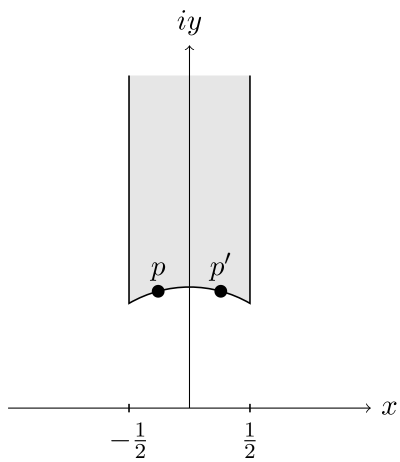

Figure 6.2: A basepoint in the fundamental domain for the action of $\mathrm{PSL}(2,\mathbb{Z})$

First consider how loops based at p corresponding to the elements S and U of $\mathrm{PSL}(2,\mathbb{Z})$ look in this fundamental domain. The element S acts on $\mathbb{H}^2$ by a rotation of angle π about i, and therefore the corresponding loop is formed by taking a simple arc from p' to p in the fundamental domain, where p and p' are identified. The element U acts by a rotation of angle $-2\pi/3$ about the point $(-1+\sqrt{3}i)/2$. Therefore the loop corresponding to U is formed by taking an arc that starts at p', crosses the right vertical edge of the fundamental domain (thereby moving to the left vertical edge of the fundamental domain), and ends at p.

Now consider how these loops lift to the thrice-punctured sphere under the degree 6 covering described in the previous section. Observe that the basepoint p on the modular surface has 6 preimages in the thrice-punctured sphere. These points are labeled $p_1,\ldots,p_6$ in the fundamental domain for the action of F_2 shown in Figure 6.3. Let A be an element of $\mathrm{PSL}(2,\mathbb{Z})$, and let γ be a corresponding curve on the modular surface, based at p. Choosing p_i as basepoint, γ lifts to a curve beginning at p_i and ending at a (likely different) p_j. In this way, a loop A induces a permutation of the points p_i. The length of a cycle in this permutation corresponds to a power to which γ needs to be raised in order for this lift to be a closed loop β_i in the thrice-punctured sphere. The number of cycles in this permutation corresponds to the number of closed loops β_i in the total preimage of γ.

To get an explicit algorithm for finding the total preimage of a loop on the modular surface, we first consider what happens for the elements S and U. The loop corresponding

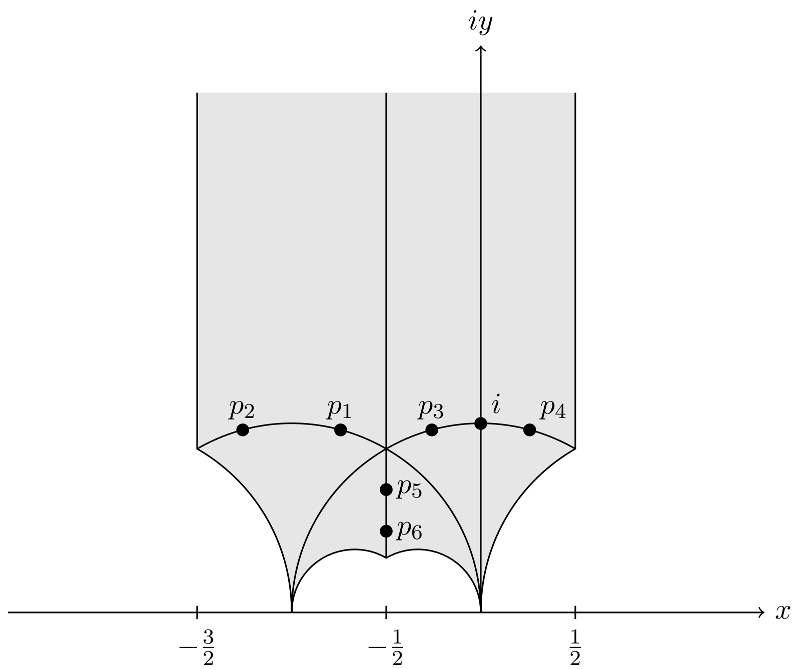

Figure 6.3: A fundamental domain for F_2 with lifts of a basepoint

to S lifts to curves from p_1 to p_2, from p_2 to p_1, from p_3 to p_4, from p_4 to p_3, from p_5 to p_6, and from p_6 to p_5. Hence S induces the permutation (12)(34)(56) on the lifts of p. The loop corresponding to U lifts to curves from p_1 to p_3, from p_3 to p_5, from p_5 to p_1, from p_2 to p_6, from p_4 to p_2, and from p_6 to p_4. Hence U induces the permutation (135)(264) on the lifts of p. Since S and U generate $\mathrm{PSL}(2, \mathbb{Z})$, this is enough information to determine the permutation on the p_i induced by any element of $\mathrm{PSL}(2, \mathbb{Z})$.

To determine the elements $h_i \in F_2$ corresponding to the β_i, we identify $F_2 = \langle b, c \rangle$ with the fundamental group of the thrice-punctured sphere by denoting by b the curve that loops once in the counterclockwise direction around the puncture corresponding to the point -1 in Figure 6.3 and denoting by c the curve that loops once in the counterclockwise direction around the puncture corresponding to the point 0. The loop in the counterclockwise direction around the puncture corresponding to the point ∞ in Figure 6.3 is then equal to $c^{-1}b^{-1}$.

To determine the element of F_2 corresponding to a closed loop in the thrice-punctured sphere based at a point p_i, one only needs to keep track of the punctures that are encircled, which can be done by recording the times one of the edges going to -1, 0, or ∞ in Figure 6.3 is crossed. For a path that has been decomposed into paths corresponding to S, U, and

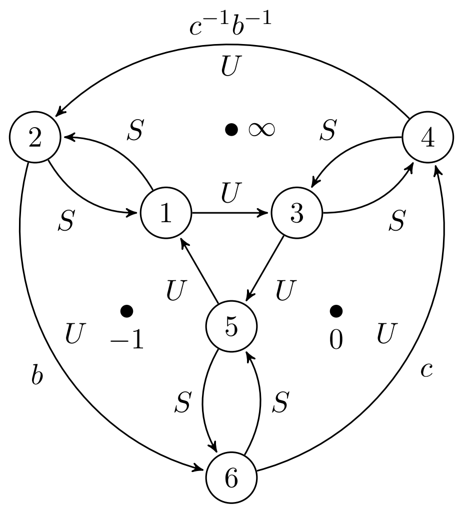

Figure 6.4: A diagram for computing elements of F_2

U^{-1}, these edges are crossed exactly when traveling between p_2 and p_4, between p_4 and p_6, or between p_6 and p_2. This is depicted in Figure 6.4, where a point p_i is replaced by a node i and lifts corresponding to S and U are shown as edges. By following the edges corresponding to the letters in an appropriate power of a word and recording the labels in b and c when traversing an outside edge, one obtains an element $h_i \in F_2$.

We have written a program that implements this algorithm to compute the elements $h_i \in F_2$ corresponding to a given element of $\mathrm{PSL}(2, \mathbb{Z})$. The stable commutator length of the formal sum of these elements can then be computed using `scallop` [CWb], allowing us to compute stable commutator length in $\mathrm{PSL}(2, \mathbb{Z})$. In the next chapter, we discuss some experimental observations based on our calculations.

Chapter 7

Experimental results

We use the algorithm described in the previous chapter to compute the stable commutator length of many elements of the modular group $\mathrm{PSL}(2,\mathbb{Z})$. Comparing these results with values of the rotation quasimorphism (computed as described in Chapter 5), we obtain experimental data about when the rotation quasimorphism is and is not extremal. In this chapter, we discuss several observations based on this data. Some graphs related to the material of Section 7.3 are presented in Appendix A.

7.1 How often is the rotation quasimorphism extremal?

Since the rotation quasimorphism has defect 1, it is extremal for an element $A \in \mathrm{PSL}(2,\mathbb{Z})$ exactly when $\mathrm{scl}(A) = \mathrm{rot}(A)/2$. Therefore, by computing values of rot as described in Chapter 5 and computing values of scl as described in Chapter 6, one can test when rot is extremal. If A has finite order, we have $\mathrm{scl}(A) = \mathrm{rot}(A) = 0$, and so we focus on infinite order elements. Since scl and rot are both *class functions*, i.e. are constant on conjugacy classes, it is only necessary to check one element of each conjugacy class of $\mathrm{PSL}(2,\mathbb{Z})$. As explained in Chapter 5, every infinite order element of $\mathrm{PSL}(2,\mathbb{Z})$ is conjugate to a positive word in the letters L and R, and therefore we test only elements of this form.

Recall from Chapter 5 that the value of the rotation quasimorphism on a word of the form $R^{a_1}L^{b_1}\cdots R^{a_n}L^{b_n}$, $a_i, b_i \geq 0$, is

$$\frac{1}{6}\left(\sum_{i=1}^{n} a_i - \sum_{i=1}^{n} b_i\right).$$

		Extremal	
Length	Number	Number	Fraction
1	2	2	1.000
2	4	4	1.000
3	8	8	1.000
4	16	16	1.000
5	32	32	1.000
6	64	64	1.000
7	128	128	1.000
8	256	208	0.813
9	512	458	0.895
10	1024	744	0.727
11	2048	1718	0.839
12	4096	2560	0.625
13	8192	6216	0.759
14	16384	8908	0.544
15	32768	22537	0.688
16	65536	33968	0.518

Table 7.1: Fraction of words for which either rot or $-\,\mathrm{rot}$ is extremal

As stable commutator length is always nonnegative, this immediately shows that the rotation quasimorphism cannot be extremal for words for which the total exponent of L is greater than the total exponent of R, since in this case rot would be negative. However, it is possible that $-\,\mathrm{rot}$ could instead be extremal for such an element. Therefore, in attempting to determine how frequently the rotation quasimorphism is extremal, we count the number of words for which either rot or $-\,\mathrm{rot}$ is extremal.

We find that, for short words in L and R, it is almost always the case that either rot or $-\,\mathrm{rot}$ is extremal, but that, for longer words in L and R, rot and $-\,\mathrm{rot}$ are less frequently extremal. More specifically, for all words of length at most 7 in L and R, either rot or $-\,\mathrm{rot}$ is extremal. However, for longer words, there are many instances when neither rot nor $-\,\mathrm{rot}$ is extremal. For example, the element R^4L^2RL has stable commutator length $5/24$, whereas the bound given by rot is $1/6$. Since the bound given by rot is always a multiple of $1/12$, it is clear that neither rot nor $-\,\mathrm{rot}$ can be extremal for any element whose stable commutator length is not a multiple of $1/12$. Having stable commutator length a multiple of $1/12$ is not enough to ensure that either rot or $-\,\mathrm{rot}$ is extremal, however. For example, the element $R^3L^2RL^2$ has stable commutator length $1/6$, whereas the bound given by rot is 0. Table 7.1

shows how frequently either rot or $-\mathrm{rot}$ is extremal for words of length less than or equal to 16 in L and R.

It is interesting to observe from this data that either rot or $-\mathrm{rot}$ seems to be extremal more frequently for words of odd length in L and R than for words of even length. We do not know the reason for this phenomenon, but suspect it is the result of a parity issue when trying to construct surfaces with a prescribed boundary. In Chapter 8, we discuss a geometric characterization of when rot is extremal in terms of curves virtually bounding an immersed surface. Apparently such immersed surfaces are easier to construct for words of odd length in L and R than for words of even length, though we do not know the reason for this.

7.2 Statistical expectations

Another observation based on the above data is that the proportion of words for which rot or $-\mathrm{rot}$ is extremal generally decreases as word length increases. This is not surprising in light of other results, and in fact we believe that, for a generic element of $\mathrm{PSL}(2,\mathbb{Z})$, neither rot nor $-\mathrm{rot}$ should be extremal. This expectation implies that the numbers in the rightmost column of Table 7.1 should go to 0 as word length increases.

Our expectation is based on comparing results about the behavior of scl and rot on generic words of long length. Calegari–Maher [CM] have shown that, in any hyperbolic group, generic rationally nullhomologous words of length m have scl $\sim \log m/m$. On the other hand, Calegari–Fujiwara [CF10a] have shown that any bicombable quasimorphism on a word-hyperbolic group satisfies a central limit theorem. (Björklund–Hartnick [BH] have also recently shown that arbitrary quasimorphisms along random walks on countable groups satisfy a central limit theorem.) This result implies that rot $\sim \sqrt{m}$ for generic words of length m. These differing rates of growth of scl and rot show that rot should not be extremal for a generic element of $\mathrm{PSL}(2,\mathbb{Z})$.

7.3 The function $n(W)$

Despite the fact that rot and $-$rot are typically not extremal, we have been able to prove a stability theorem about elements for which rot is extremal. Our main theorem (see Chapter 8) shows that multiplying any word W in L and R by a sufficiently large power of R ensures that the rotation quasimorphism will be extremal for the resulting element. In light of this theorem, if rot is not extremal for W, it is natural to wonder how large a natural number n is needed to make rot extremal for R^nW. One might think of this as giving a measure of "how far" rot is from being extremal for W. On the other hand, if rot is extremal for W, it will fail to be extremal for some L^mW simply because $\mathrm{rot}(L^mW)$ will be negative for sufficiently large m. The exponent needed here gives some measure of how "strongly" extremal rot is for W.

To make this more precise, for every word W in L and R we define an integer $n(W)$ as follows. If rot is not extremal for W, then $n(W)$ is the smallest natural number for which rot is extremal for R^nW. If rot is extremal for W, then $n(W)$ is $-m$, where m is the smallest natural number for which the rotation quasimorphism is extremal for L^mW but not for $L^{m+1}W$.

If rot were extremal exactly when it is nonnegative, $n(W)$ would follow a binomial distribution for words of a fixed length. This is true for words of length 1, 2, 3, and 4, but is not true in general since rot is often not extremal even when it is nonnegative. For example, $n(L^2RL^2)$ could be as small as 3, but it is actually 6. One begins to see irregularity in the distribution of $n(W)$ for words W of length 5, as shown in Figure A.1, and this becomes more pronounced for longer words. The distributions of $n(W)$ for words of length 6, 7, 8, and 9 are shown in Figures A.2, A.3, A.4, and A.5. Especially in the last figure, one has the impression that the distribution of $n(W)$ is not centered about 0 but rather is skewed toward positive values. It would be interesting to study the behavior of the distribution of $n(W)$ for long words.

7.4 Stuttering

Word
RRRLRLL
RRRLLRL
RRLRLLR
RRLLRLR
RLRLLRR
RLLRLRR
LRRRRLL
LRLLRRR
LLRRRRL
LLRLRRR

Table 7.2: All examples of 1-stuttering for words of length 7

Word
RRLRLLRL
RLRLRRLL
RLRLLRLR
RLLRRLRL
LRRRLRLL
LRRLLRRL
LRLRRLLR
LRLLRLRR
LLRRRRLL
LLRRLRLR
LLRLRRRL

Table 7.3: All examples of 1-stuttering for words of length 8

Given that we will show in Chapter 8 that rot is extremal for R^nW for sufficiently large n, one might expect that if rot is extremal for R^nW then it is also extremal for $R^{n+1}W$. However, this is not necessarily the case. Our theorem only shows that there is some N such that, whenever $n \geq N$, rot is extremal for R^nW. We do not have any control over what happens when $n < N$, and indeed we have found examples of various types of behavior. If rot is extremal for W but not for RW,...,R^mW, we say that W is an example of *m-stuttering*.

The first examples of 1-stuttering occur for words of length 7. There are exactly ten such examples among the words of length 7 in L and R, and these are listed in Table 7.2. For words of length 8, there are exactly eleven such examples, and these are listed in Table 7.3.

The number of examples of 1-stuttering seems to continue to grow as longer words are considered.

The first examples of 2-stuttering occur for words of length 8, where there are exactly four such examples: R^2LRL^2RL, $RLRL^2RLR$, LRL^2RLR^2, and $L^2R^4L^2$. The first examples of 3-stuttering occur among words of length 14, and $RLR^2L^2RLRL^2R^2L$ is such an example. Although we have not found examples of m-stuttering for $m \geq 4$, we have no reason to believe that m-stuttering cannot happen for larger m. It would be interesting to study whether m-stuttering can occur for arbitrarily large m.

Chapter 8

Main theorem

In this chapter, we state our main stability theorem about elements of the modular group $\mathrm{PSL}(2,\mathbb{Z})$ for which the rotation quasimorphism is extremal. We separately consider the cases of finite order, parabolic, and hyperbolic elements, reducing to consideration of elements of a particular form. Our theorem then follows from another theorem about when closed geodesics on the modular surface bound immersed orbifolds, and we prove this second theorem in Chapter 9.

Theorem. *For every element $A \in \mathrm{PSL}(2,\mathbb{Z})$, there exists a parabolic element $P \in \mathrm{PSL}(2,\mathbb{Z})$ and an integer $N \in \mathbb{Z}$ such that, whenever $n \geq N$, the rotation quasimorphism is extremal for the element P^nA.*

8.1 Parabolic elements

As explained in Chapter 5, every infinite order element $A \in \mathrm{PSL}(2,\mathbb{Z})$ is conjugate to a positive word in L and R. We first consider the case when A is conjugate to R^a, $a > 0$, or to L^b, $b > 0$. (This corresponds to the case when A is parabolic.) Observe that $SL^bS = S(SU)^bS = (US)^b = R^{-b}$. This shows that every element conjugate to R^a or L^b is in fact conjugate to R^a for some (possibly negative) a. Suppose $A = BR^aB^{-1}$, and let $P = BRB^{-1}$. Then $P^nA = BR^nB^{-1}BR^aB^{-1} = BR^{a+n}B^{-1}$. We compute that $\mathrm{scl}(R) = 1/12$ and $\mathrm{rot}(R) = 1/6$, which implies that $\mathrm{scl}(R^m) = \mathrm{rot}(R^m)/2 = m/12$ for all $m \in \mathbb{N}$. Choosing $N \geq -a$, we get that rot is extremal for R^{a+n}. Since scl and rot are both class functions, it follows that rot is also extremal for $BR^{a+n}B^{-1} = P^nA$.

8.2 Hyperbolic elements

We next consider the case when A is conjugate to a positive word in L and R but not to R^a or L^b. (This corresponds to the case when A is hyperbolic.) In this case, A is conjugate to an element of the form $W = R^{a_1}L^{b_1}\cdots R^{a_k}L^{b_k}$, where $a_i, b_i > 0$. Suppose $A = BWB^{-1}$, and let $P = BRB^{-1}$. Then $P^nA = BR^nB^{-1}BWB^{-1} = BR^nWB^{-1}$. This means that, in order to show rot is extremal for P^nA, it suffices to show it is extremal for R^nW. Note that $R^nW = R^{a_1+n}L^{b_1}\cdots R^{a_n}L^{b_n}$, which is again a positive word in L and R. Therefore it suffices to only consider words of this standard form. In Section 8.4 and following, we will show that, for every element of the form $R^{a_1}L^{b_1}\cdots R^{a_k}L^{b_k}$, $a_i, b_i > 0$, making a_1 sufficiently large is enough to ensure that rot is extremal for this element.

8.3 Finite order elements

If A is of finite order, it is conjugate to S, U, or U^{-1}. We consider these three cases separately.

1. Suppose $A = BSB^{-1}$. Let $P = BRB^{-1}$. Then, when $n \geq 2$,

$$\begin{aligned}
P^nA &= BR^nSB^{-1} \\
&= B(SU^{-1})^nSB^{-1} \\
&= (BSU^{-1})(SU^{-1})^{n-2}SU^{-1}SSU^{-1}(USB) \\
&= (BSU^{-1})(SU^{-1})^{n-2}SU(BSU^{-1})^{-1} \\
&= (BSU^{-1})R^{n-2}L(BSU^{-1})^{-1}.
\end{aligned}$$

   Therefore it suffices to show that rot is extremal for $R^{n-2}L$ for sufficiently large n. This fits the standard form of hyperbolic elements to be considered in Section 8.4 and following.

2. Suppose $A = BUB^{-1}$. Let $P = BRB^{-1}$. Then, when $n \geq 3$,

$$\begin{aligned} P^n A &= BR^n UB^{-1} \\ &= B(SU^{-1})^n UB^{-1} \\ &= B(SU^{-1})^{n-1} SB^{-1} \\ &= (BSU^{-1})(SU^{-1})^{n-3} SU^{-1} SSU^{-1}(USB) \\ &= (BSU^{-1})(SU^{-1})^{n-3} SU(BSU^{-1})^{-1} \\ &= (BSU^{-1}) R^{n-3} L (BSU^{-1})^{-1}. \end{aligned}$$

   Therefore it suffices to show that rot is extremal for $R^{n-3}L$ for sufficiently large n. This fits the standard form of hyperbolic elements to be considered in Section 8.4 and following.

3. Suppose $A = BU^{-1}B^{-1}$. Let $P = BRB^{-1}$. Then, when $n \geq 1$,

$$\begin{aligned} P^n A &= BR^n U^{-1} B^{-1} \\ &= B(SU^{-1})^n U^{-1} B^{-1} \\ &= B(SU^{-1})^{n-1} SUB^{-1} \\ &= BR^{n-1} LB^{-1}. \end{aligned}$$

   Therefore it suffices to show that rot is extremal for $R^{n-1}L$ for sufficiently large n. This fits the standard form of hyperbolic elements to be considered in Section 8.4 and following.

8.4 The geometric approach

It remains to show that the rotation quasimorphism is extremal for a hyperbolic element of the form $R^{a_1} L^{b_1} \cdots R^{a_k} L^{b_k}$, $a_i, b_i > 0$, whenever a_1 is sufficiently large. It follows from work of Calegari [Cal09a] that rot is extremal for a hyperbolic element of $\mathrm{PSL}(2, \mathbb{Z})$ exactly when the corresponding geodesic on the modular surface virtually bounds an immersed surface.

We use this geometric condition to finish the proof of our main theorem. First, we explain what it means for a curve to virtually bound an immersed surface.

Recall that a differentiable map between differentiable manifolds is an *immersion* if its derivative at every point is injective. When immersing surfaces in 2-dimensional orbifolds, we require that, around any point on the surface that maps to an order n cone point, the map factors through the quotient by a rotation of angle $2\pi/n$. We say that an immersed curve c in a differentiable manifold X *bounds an immersed surface* if there is an immersion $\Sigma \to X$ mapping $\partial\Sigma$ to c in an orientation-preserving way. An immersed curve c in X *virtually* bounds an immersed surface if there is an immersion $\Sigma \to X$ mapping $\partial\Sigma$ to a cover of c in an orientation-preserving way.

We want to show that, for a hyperbolic element of the form $R^{a_1}L^{b_1}\cdots R^{a_k}L^{b_k}$, $a_i, b_i > 0$, the corresponding geodesic on the modular surface virtually bounds an immersed surface whenever a_1 is sufficiently large. The proof of this result occupies Section 8.5 and Chapter 9.

It was observed experimentally that a similar result seems to hold in the free group F_2, namely that for any word w the curve on the once-punctured torus corresponding to $w[a,b]^n$ virtually bounds an immersed surface for all sufficiently large n. Thus our result may be regarded as an analogue of Conjecture 3.16 from [Cal09a].

8.5 Cone points

In Chapter 9, we show that a hyperbolic element $R^{a_1}L^{b_1}\cdots R^{a_k}L^{b_k}$, $a_i, b_i > 0$, corresponds to a geodesic that bounds an immersed orbifold whenever a_1 is sufficiently large. In this section, we explain how this result implies that such a geodesic also virtually bounds an immersed surface.

We say that an immersed curve c in a 2-dimensional orbifold X *bounds an immersed orbifold* if there is an immersion $\Sigma \to X$ of a 2-dimensional orbifold Σ that takes $\partial\Sigma$ to c in an orientation-preserving way. If an order m cone point maps to an order n cone point, we require that $m \mid n$ and that around the order m cone point the map factors through the quotient by an order n/m rotation.

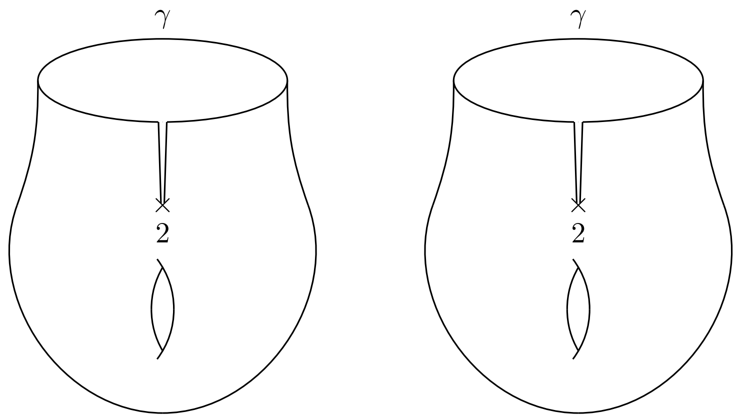

Figure 8.1: Eliminating cone points

If a curve γ bounds an immersed orbifold, we construct an immersed surface that it virtually bounds as follows. Suppose such an immersed orbifold has k cone points, and let the orders of these cone points be $n_1, \ldots, n_k$. For each cone point of the immersed orbifold, cut along an arc from the cone point to the boundary (which maps to γ), as depicted in Figure 8.1. The k arcs along which cuts are made should be disjoint. Now take $\operatorname{lcm}(n_1, \ldots, n_k)$ copies of the resulting orbifold. For the ith cone point, divide the $\operatorname{lcm}(n_1, \ldots, n_k)$ copies of the orbifold into $\operatorname{lcm}(n_1, \ldots, n_k)/n_i$ groups of n_i each. Within each group of n_i copies, glue up the cuts to the ith cone point by gluing the left side of the cut on the jth copy to the right side of the cut on the $(j+1)$st copy for all $1 \leq j \leq n_i - 1$, as well as the left side of the cut on the n_ith copy to the right side of the cut on the first copy. Doing this for all cone points, we create an immersed surface with no cone points whose boundary maps to a degree $\operatorname{lcm}(n_1, \ldots, n_k)$ cover of γ. Thus γ virtually bounds an immersed surface.

To prove our main theorem, it therefore remains only to show that appropriate geodesics on the modular surface bound immersed orbifolds. This is shown in the next chapter.

Chapter 9

Immersed orbifolds in the modular surface

As explained in Chapter 8, our main theorem follows from a stability theorem about when certain geodesics on the modular surface bound immersed orbifolds. In this chapter, we state and prove this theorem. The material of this chapter is largely taken from [CL11].

Theorem. *Consider a hyperbolic conjugacy class in* $\mathrm{PSL}(2,\mathbb{Z})$*, represented by a word* W *of the form* $R^{a_1}L^{b_1}\cdots R^{a_k}L^{b_k}$ *for* $a_i, b_i > 0$*. If* a_1 *is sufficiently large, the corresponding geodesic on the modular surface bounds an immersed orbifold.*

The requirement that a_1 be sufficiently large means explicitly that $a_1 \geq \sum_{i=2}^{k} a_i + \sum_{i=1}^{k} b_i + 11k + 7$. This is simply what is needed to ensure that our proof will work; it is certainly not a necessary condition. Note that if a_1 is sufficiently large for a word W, it will also be sufficiently large for all R^mW, $m \in \mathbb{N}$, which is why we call this a stability theorem.

The remainder of this chapter is devoted to the proof of this theorem.

9.1 Decomposing the geodesic

Fix a hyperbolic conjugacy class in $\mathrm{PSL}(2,\mathbb{Z})$, represented by a word $W = R^{a_1}L^{b_1}\cdots R^{a_k}L^{b_k}$, $a_i, b_i > 0$, and let γ be the corresponding closed geodesic on the modular surface. We first show how to divide γ into arcs corresponding roughly to the terms R^{a_i} and L^{b_i}. As in Section 4.4, let σ be the geodesic on the modular surface between the order 2 and order 3 cone points. Recall that the total preimage of σ in $\mathbb{H}^2$, denoted $\widetilde{\sigma}$, is the Farey graph, shown in Figure 4.3. Let V denote the region of $\mathbb{H}^2$ above $\widetilde{\sigma}$, as shown in Figure 9.1.

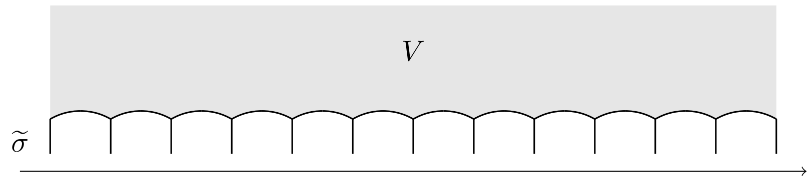

Figure 9.1: The region V

For each term R^{a_i}, let F_i be a bi-infinite path in $\widetilde{\sigma}$ that makes left and right turns according to the appearances of the letters L and R in W, chosen so it makes a_i consecutive right turns along ∂V. An example of such a path is shown in Figure 9.2. The endpoints of this path are fixed points for the action of a conjugate of W on $\partial\mathbb{H}^2$. Let $\widetilde{\gamma}$ be the lift of γ with these fixed points as endpoints. Then let α_i be the projection of $\widetilde{\gamma} \cap V$ to the modular surface. (If $a_i = 1$, it is possible this intersection could be empty, in which case α_i does not exist.) Similarly, for each term L^{b_i}, let G_i be a bi-infinite path corresponding to W that makes b_i consecutive left turns along ∂V. The endpoints of this path determine a lift $\widetilde{\gamma}$ of γ, and we denote the projection of $\widetilde{\gamma} \cap V$ to the modular surface by β_i (which might not exist if $b_i = 1$). Thus we have decomposed the closed geodesic γ into arcs α_i and β_i, each of which travel between consecutive intersection points of γ with σ.

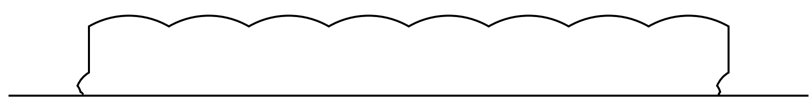

Figure 9.2: A bi-infinite path in $\widetilde{\sigma}$ corresponding to the term R^7 of the word R^7L^2RL

The lengths of the arcs α_i and β_i correspond roughly to the exponents a_i and b_i, as we now explain. Consider the path in $\widetilde{\sigma}$ corresponding to the bi-infinite word $\dot{L}R^{a_i}\dot{L}$ that makes a_i right turns in the same places F_i does, where $\dot{L}$ indicates an infinite sequence of Ls. The endpoints of this path are inside the endpoints of F_i, and hence the geodesic connecting them is below $\widetilde{\gamma}$. Also consider the path corresponding to $\dot{R}LR^{a_i}L\dot{R}$ that makes a_i right turns in the same places F_i does. The endpoints of this path are outside the endpoints of F_i, and hence the geodesic connecting them is above $\widetilde{\gamma}$. These two geodesics constrain the length of $\widetilde{\gamma} \cap V$, as shown in Figure 9.3. For terms L^{b_i}, we similarly consider the paths in $\widetilde{\sigma}$ corresponding to $\dot{R}L^{b_i}\dot{R}$ and $\dot{L}RL^{b_i}R\dot{L}$ that make b_i left turns in the same places G_i does.

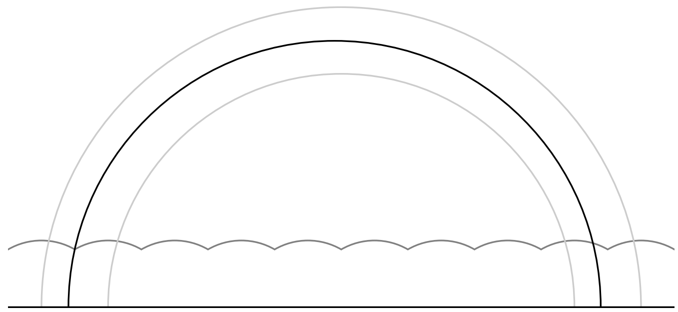

Figure 9.3: An example of the constraints on $\widetilde{\gamma} \cap V$

The geodesics connecting the endpoints of these paths are below and above the original geodesic, and hence constrain the length of $\widetilde{\gamma} \cap V$.

These paths in $\widetilde{\sigma}$ beginning and ending in $\dot{L}$ or $\dot{R}$ have endpoints on integer values along the real axis, and hence the geodesics bounding $\widetilde{\gamma}$ have centers at half-integer points and half-integer radii. One can thus compute how long these geodesics stay in V. If there are m unintersected segments of ∂V between those segments intersected by a geodesic, we say that this geodesic *uses* $m + 2$ segments of ∂V. One finds that the geodesic corresponding to $\dot{L}R^{a_i}\dot{L}$ uses at least $a_i - 1$ segments of ∂V and the geodesic corresponding to $\dot{R}LR^{a_i}L\dot{R}$ uses at most $a_i + 1$ segments of ∂V. Thus we conclude that the $\widetilde{\gamma}$ corresponding to R^{a_i} uses between $a_i - 1$ and $a_i + 1$ segments of ∂V. In particular, this shows α_i is nonempty whenever $a_i \geq 2$. One also finds that the geodesic corresponding to $\dot{R}L^{b_i}\dot{R}$ involves at least $b_i - 1$ segments of ∂V and the geodesic corresponding to $\dot{L}RL^{b_i}R\dot{L}$ involves at most $b_i + 1$ segments of ∂V, meaning the $\widetilde{\gamma}$ corresponding to L^{b_i} uses between $b_i - 1$ and $b_i + 1$ segments of ∂V. The most important point is that the numbers a_i and b_i control the number of segments of ∂V used by lifts of α_i and β_i to V. Since we assumed that a_1 is very large relative to the b_i and the other a_i, we know that lifts of α_1 to V will use many more segments of ∂V than lifts of any of the β_i or the other α_i.

9.2 Choosing appropriate lifts

We choose lifts $\widetilde{\alpha}_i$ of α_i and $\widetilde{\beta}_i$ of β_i to V in a particular way. Roughly speaking, we want the lifts $\widetilde{\alpha}_i$ and $\widetilde{\beta}_i$ to be as shown in Figure 9.4. Such an arrangement is needed to ensure

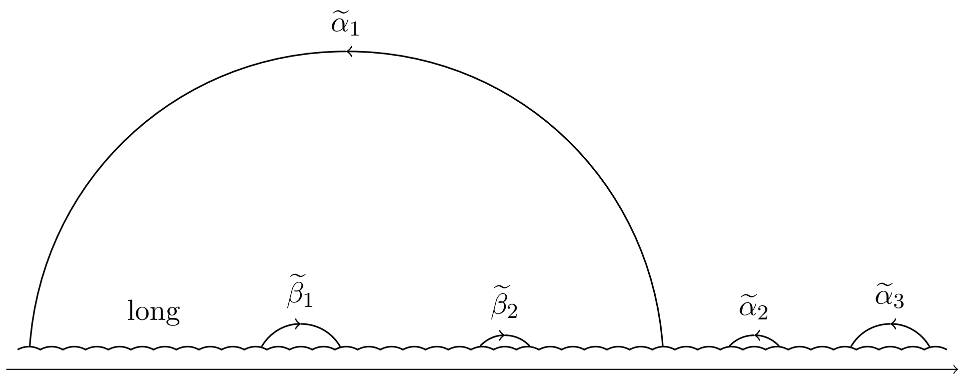

Figure 9.4: An example of how to lift to $\widetilde{\alpha}_i$ and $\widetilde{\beta}_i$

that the subsequent steps of the proof will work. More specifically, we choose $\widetilde{\alpha}_i$ and $\widetilde{\beta}_i$ satisfying the following conditions.

1. None of the $\widetilde{\alpha}_i$ or $\widetilde{\beta}_i$ intersect each other.
2. No segment of ∂V is intersected by more than one of the $\widetilde{\alpha}_i$ or $\widetilde{\beta}_i$.
3. None of the $\widetilde{\alpha}_i$ are nested inside each other.
4. None of the $\widetilde{\beta}_i$ are nested inside each other.
5. All the $\widetilde{\beta}_i$ are under $\widetilde{\alpha}_1$.
6. There are exactly five segments of ∂V between segments where the $\widetilde{\beta}_i$ intersect ∂V.
7. There are exactly five segments of ∂V between the rightmost segment of ∂V intersected by a $\widetilde{\beta}_i$ and the rightmost segment of ∂V intersected by $\widetilde{\alpha}_1$.

The most important consequence of this arrangement is that there is a long sequence of unintersected segments of ∂V between the leftmost segment ∂V intersected by of $\widetilde{\alpha}_1$ and the leftmost segment of ∂V intersected by $\widetilde{\beta}_i$.

9.3 Arranging the surface to be glued

We have arranged the curves $\widetilde{\alpha}_i$ and $\widetilde{\beta}_i$ in this way so that we can construct an immersed orbifold from the pieces of V they bound. Consider the portion of V to the left of the $\widetilde{\alpha}_i$ and $\widetilde{\beta}_i$, as shown in Figure 9.5. Observe that, when these pieces are projected to the modular

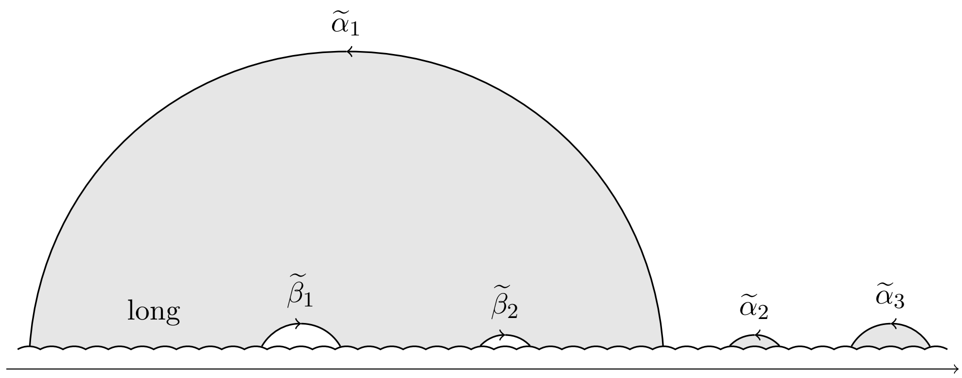

Figure 9.5: The pieces of V that will be glued together

surface, their boundary consists of all of the curves α_i and β_i, plus a number of whole or partial segments of ∂V. We attempt to glue together these segments of ∂V to obtain an orbifold whose boundary maps to only of the α_i and β_i.

Observe that γ traverses the arcs α_i and β_i in a certain order. Therefore, in gluing together pieces of V to construct an orbifold whose boundary maps to γ, we want the curves $\widetilde{\alpha}_i$ and $\widetilde{\beta}_i$ to be arranged in the corresponding way on the boundary. This gives a natural way to glue the (partial) segments of ∂V that are intersected by one of the $\widetilde{\alpha}_i$ or $\widetilde{\beta}_i$, and we do this gluing. If γ passes through the order 2 cone point of the modular surface, the corresponding segments of ∂V can be glued to each other as usual. If γ passes through the order 3 cone point, we cannot glue the corresponding segments of ∂V to each other directly as this would create a cone point on the boundary of the orbifold we are constructing. Rather, we need to glue another piece with angle $2\pi/3$ between these segments. There are many suitable pieces along the long sequence of unintersected segments of ∂V. We use a pair of segments near the end of this sequence for this purpose, resolving the issue of γ passing through the order 3 cone point. For each α_i that intersects adjacent segments of ∂V, there is a point about which there are three corners of angle $2\pi/3$, and this forces us to glue an additional pair of segments, which we do. There are then no additional points with three corners of angle $2\pi/3$, due to the space that was left between the $\widetilde{\beta}_i$.

We are left with a surface with several boundary components, such as the one shown in Figure 9.6. One of these components consists only of the curves $\widetilde{\alpha}_i$ and $\widetilde{\beta}_i$, traversed in the order corresponding to γ. This component maps to γ under the map to the modular surface

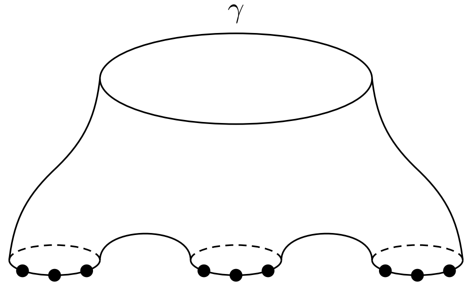

Figure 9.6: The result after gluing several segments of ∂V

induced by the projection $\mathbb{H}^2 \to \mathbb{H}^2/\operatorname{PSL}(2,\mathbb{Z})$, and so we label it by γ in Figure 9.6. The other boundary components contain several points that are preimages of the order 3 cone point in the modular surface; these are indicated by dots in Figure 9.6. The angle between segments meeting at these points is either $2\pi/3$, if no gluing has yet occurred at this point, or $4\pi/3$, if a pair of edges has already been glued at this point.

9.4 Reducing to the case of one component

In trying to construct an orbifold that immerses in the modular surface and whose boundary maps to γ, we want to glue up these other boundary components entirely, so that the boundary segment labeled by γ is the only one remaining. In doing this, edges must either be glued in pairs or folded in half about a preimage of the order 2 cone point. About points labeled by dots the total angle must be either 2π or $2\pi/3$, since these points map to the order 3 cone point. This is a completely combinatorial problem, and we explain how to solve it under the hypotheses of our theorem. To simplify the diagrams, we label dots about which the angle is $2\pi/3$ by 1 and dots about which the angle is $4\pi/3$ by 2.

We need to glue several components, such as those shown in Figure 9.7. Based on the way the curves $\widetilde{\alpha}_i$ and $\widetilde{\beta}_i$ were chosen in Section 9.2, we can say several things about points on these components. First, every component contains at least part of one of the sequences of five segments of ∂V left between two of the $\widetilde{\beta}_i$ or between a $\widetilde{\beta}_i$ and $\widetilde{\alpha}_1$. The first and last of these five segments might have already been identified with other segments, but the middle three segments remain unglued. Therefore the angle about the two points in the middle of the sequence of five segments is still $2\pi/3$, which means each component to be

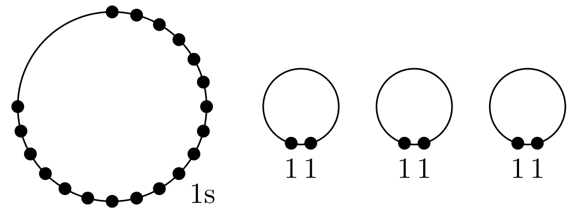

Figure 9.7: The components to be glued, one with a long sequence of 1s

glued must contain a pair of adjacent 1s. We also know that the long sequence of segments of ∂V between the leftmost endpoint of $\widetilde{\alpha}_1$ and the leftmost endpoint of a $\widetilde{\beta}_i$ must have a long sequence of unidentified edges, and therefore one of the components in Figure 9.7 must have a long sequence of 1s. We call this component the primary component and all other components secondary components. We now show how to use this long sequence of 1s to glue up all the components.

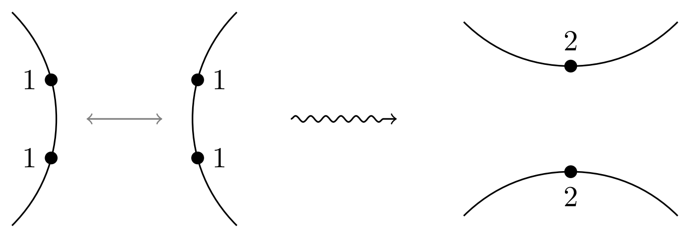

Figure 9.8: Gluing two 11 segments

Observe that we can glue a pair of 11 segments as shown in Figure 9.8, joining pairs of points labeled by 1 to give points labeled by 2. For each secondary component, we use this move to identify a 11 sequence with a 11 sequence near the end of the long string of 1s on the primary component. In this way, all secondary components are joined to the primary component, and the primary component still has a long sequence of 1s. Therefore we have reduced to the problem of showing that a single component with a long sequence of 1s can be glued up.

9.5 Reducing to the case of one 2

We have a single component with a long sequence of 1s and some other sequence of numbers v, as indicated in Figure 9.9. The main use of the long sequence of 1s is to create the

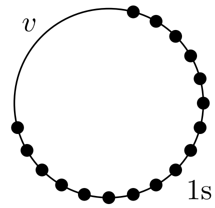

Figure 9.9: The remaining component to be glued, with a long sequence of 1s

complimentary sequence v^c to v, in order that v can be glued up. By complementary sequence, we mean that if v consists of points labeled by a, b, c we create the sequence labeled by $3 - c, 3 - b, 3 - a$. (We reverse the order so that v^c can be glued to v.) To create a complementary sequence, we must be able to create points labeled by 2 wherever we wish in the long sequence of 1s. This can be done simply by gluing an edge to itself, folding about the point that maps to the order 2 cone point in the modular surface, as shown in Figure 9.10.

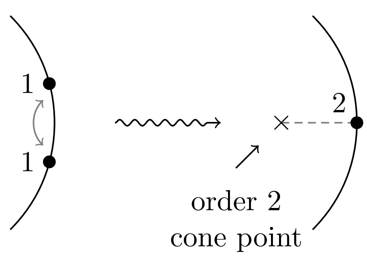

Figure 9.10: Folding a 11 segment

Now we use this to create the compliment v^c of v somewhere in the middle of the sequence of 1s. We then glue v^c to v, as shown in Figure 9.11.

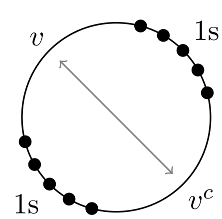

Figure 9.11: Gluing a sequence to its complement

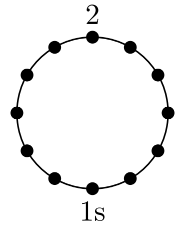

Figure 9.12: A component with a single 2

After doing this gluing, there are two remaining boundary components, each consisting of part of the long sequence of 1s not used in creating v^c. A 1 next to v is identified with a 1 next to v^c, so therefore these remaining components have a single point labeled 2 and all other points labeled 1. (We choose the placement of v^c to ensure that each resulting component has at least one 1.) It turns out each of these components can be glued up separately, and therefore we reduce to the case of considering a single component with exactly one 2, as depicted in Figure 9.12.

9.6 Reducing to special cases

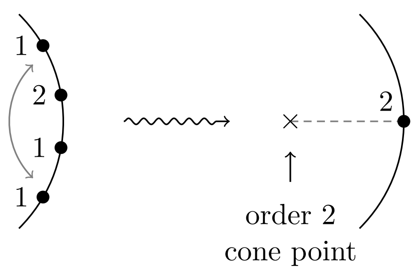

Figure 9.13: Folding a 1211 sequence

If the component we are considering has four or more 1s, we can fold a 1211 sequence as shown in Figure 9.13, identifying a pair of edges and folding the middle edge in half about a point mapping to the order 2 point in the modular surface. This has the effect of converting the 1211 sequence to a single 2 point, thus reducing the number of 1s by three. We do this repeatedly until we are left with only one, two, or three 1s.

9.7 Special cases

It remains to show how to glue up three special cases, a 21 circle, a 211 circle, and a 2111 circle. Each of these is explained below.

1. In the case of one 2 and one 1, fold along the line intersecting the midpoints of the edges, as shown in Figure 9.14. This produces two cone points of order 2, both of which map to the order 2 cone point of the modular surface.

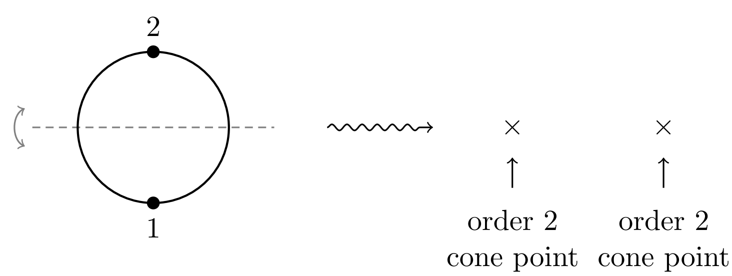

Figure 9.14: Folding up a 21 circle

2. In the case of one 2 and two 1s, fold along the line intersecting the midpoint of a 21 edge and the other 1 point, as shown in Figure 9.15. This produces one order 2 cone point, which maps to the order 2 cone point of the modular surface, and one order 3 cone point, which maps to the order 3 cone point of the modular surface.

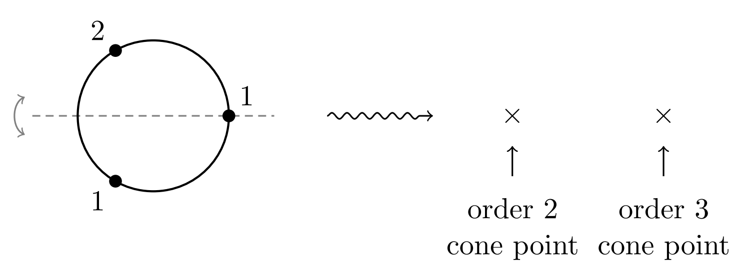

Figure 9.15: Folding up a 211 circle

3. In the case of one 2 and three 1s, fold along the line intersecting the two opposite 1 points, as shown in Figure 9.16. This produces two order 3 cone points, both of which map to the order 3 cone point of the modular surface.

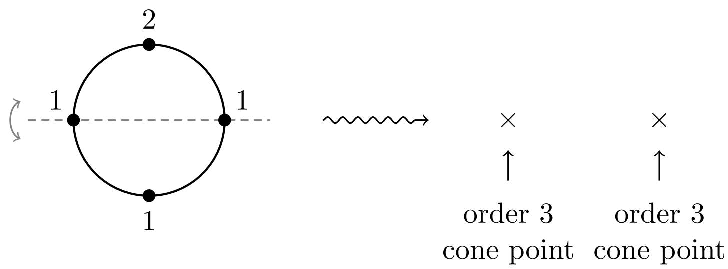

Figure 9.16: Folding up a 2111 circle

This completes the proof of the theorem stated at the beginning of this chapter. This result implies our main theorem, as explained in Chapter 8.

Chapter 10

Generalizations

One reason for studying the modular group $\mathrm{PSL}(2,\mathbb{Z})$ is that it is a prototype for several classes of groups. One can therefore hope to generalize results about the modular group to groups such as braid groups, mapping class groups, or other lattices in $\mathrm{PSL}(2,\mathbb{R})$. In this chapter, we discuss two generalizations of our results, one to the 3-strand braid group and the other to Hecke triangle groups.

10.1 The 3-strand braid group

The study of stable commutator length in the modular group is closely related to the study of stable commutator length in the 3-strand braid group. The *n-strand braid group* B_n has presentation

$$\left\langle \sigma_1, \ldots, \sigma_{n-1} \;\middle|\; \begin{array}{l} \sigma_i\sigma_{i+1}\sigma_i = \sigma_{i+1}\sigma_i\sigma_{i+1} \\ \sigma_i\sigma_j = \sigma_j\sigma_i \text{ when } |i-j| \geq 2 \end{array} \right\rangle,$$

where geometrically one thinks of σ_i as consisting of n vertical strands where the ith strand crosses under the $(i+1)$st strand. The abelianization of B_n is obtained by adding the commutativity relations $\sigma_i\sigma_{i+1} = \sigma_{i+1}\sigma_i$ to the above presentation. Combined with the braid relations, these show exactly that all generators σ_i are equal in the abelianization, so hence $B_n/[B_n, B_n] \cong \mathbb{Z}$.

Consider the surjective homomorphism $\mathrm{lk}_n \colon B_n \to \mathbb{Z}$ that maps each generator σ_i to 1. Thus lk_n counts the "total exponent" of a word in the braid generators, and is well defined since this quantity is preserved by the braid relations. Any element of the commutator subgroup must have total exponent 0, so it is clear that $[B_n, B_n] \trianglelefteq \ker \mathrm{lk}_n$. By the isomorphism theorems,

$$\mathbb{Z} \cong B_n / \ker \mathrm{lk}_n \cong (B_n/[B_n, B_n])/(\ker \mathrm{lk}_n /[B_n, B_n]) \cong \mathbb{Z}/(\ker \mathrm{lk}_n /[B_n, B_n]).$$

Since $\mathbb{Z}$ can only be a quotient of itself by the trivial subgroup, this shows that $[B_n, B_n] = \ker \mathrm{lk}_n$, i.e. that the commutator subgroup of B_n consists of exactly those braids with total exponent 0.

The center of B_n is the infinite cyclic subgroup generated by the "full twist" $(\sigma_1 \cdots \sigma_{n-1})^n$. In the case $n = 3$, $B_3/Z(B_3) \cong \mathrm{PSL}(2, \mathbb{Z})$, which allows the use of techniques not available for higher-strand braid groups. The quotient (anti-)homomorphism $\phi\colon B_3 \to \mathrm{PSL}(2, \mathbb{Z})$ takes σ_1 to L^{-1} and σ_2 to R. It turns out that commutator length and stable commutator length in B_3 are completely determined by commutator length and stable commutator length in the modular group, for the following reason. Let $\beta \in [B_3, B_3]$. If $\beta = \prod_{i=1}^{g} [\delta_i, \gamma_i]$, then $\phi(\beta) = \phi\big(\prod_{i=1}^{g} [\delta_i, \gamma_i]\big) = \prod_{i=1}^{g} [\phi(\delta_i), \phi(\gamma_i)]$. Conversely, if $\phi(\beta) = \prod_{i=1}^{g} [B_i, C_i]$, choose any preimages $\delta_i \in \phi^{-1}(B_i)$ and $\gamma_i \in \phi^{-1}(C_i)$. Since β and $\prod_{i=1}^{g} [\delta_i, \gamma_i]$ have the same image under ϕ, they are equal in B_3 up to a power of the full twist. But the full twist has total exponent 6, whereas β and $\prod_{i=1}^{g} [\delta_i, \gamma_i]$ both have total exponent 0. Thus we must in fact have $\beta = \prod_{i=1}^{g} [\delta_i, \gamma_i]$. This shows that, for all $\beta \in [B_3, B_3]$, we have $\mathrm{cl}_{B_3}(\beta) = \mathrm{cl}_{\mathrm{PSL}(2,\mathbb{Z})}(\phi(\beta))$, and hence $\mathrm{scl}_{B_3}(\beta) = \mathrm{scl}_{\mathrm{PSL}(2,\mathbb{Z})}(\phi(\beta))$.

The rotation quasimorphism may also be extended from $\mathrm{PSL}(2, \mathbb{Z})$ to B_3. In the following, we use rot to denote the rotation quasimorphism on $\mathrm{PSL}(2, \mathbb{Z})$ and rot_{B_3} to denote the corresponding quasimorphism that we define on the 3-strand braid group. Given $\beta \in B_3$, set $\mathrm{rot}_{B_3}(\beta) = \mathrm{rot}(\phi(\beta))$. Then

$$\begin{aligned}
|\mathrm{rot}_{B_3}(\beta\delta) - \mathrm{rot}_{B_3}(\beta) - \mathrm{rot}_{B_3}(\delta)| &= |\mathrm{rot}(\phi(\beta\delta)) - \mathrm{rot}(\phi(\beta)) - \mathrm{rot}(\phi(\delta))| \\
&= |\mathrm{rot}(\phi(\beta)\phi(\delta)) - \mathrm{rot}(\phi(\beta)) - \mathrm{rot}(\phi(\delta))| \\
&\leq D(\mathrm{rot}),
\end{aligned}$$

so rot_{B_3} is a quasimorphism. This also shows that $D(\mathrm{rot}_{B_3}) \leq D(\mathrm{rot})$. To show that

$D(\mathrm{rot}_{B_3}) \geq D(\mathrm{rot})$, choose elements $B, C \in \mathrm{PSL}(2,\mathbb{Z})$ such that

$$|\mathrm{rot}(BC) - \mathrm{rot}(B) - \mathrm{rot}(C)| + \epsilon > D(\mathrm{rot}).$$

Then for any $\beta \in \phi^{-1}(B)$ and $\delta \in \phi^{-1}(C)$, we have

$$\begin{aligned} |\mathrm{rot}_{B_3}(\beta\delta) - \mathrm{rot}_{B_3}(\beta) - \mathrm{rot}_{B_3}(\delta)| + \epsilon &= |\mathrm{rot}(BC) - \mathrm{rot}(B) - \mathrm{rot}(C)| + \epsilon \\ &> D(\mathrm{rot}). \end{aligned}$$

Thus $D(\mathrm{rot}_{B_3}) = D(\mathrm{rot})$.

Since for $\beta \in [B_3, B_3]$ we have that $\mathrm{scl}(\beta) = \mathrm{scl}(\phi(\beta))$ and $\mathrm{rot}_{B_3}(\beta) = \mathrm{rot}(\phi(\beta))$, it follows that rot_{B_3} is extremal for $\beta \in [B_3, B_3]$ exactly when rot is extremal for $\phi(\beta)$. As $[B_3, B_3]$ is not a finite index subgroup of B_3, the analogue of our main theorem must be stated in such a way as to ensure that the braids under consideration are trivial in rational homology. We accomplish this by multiplying by a power of the full twist $(\sigma_1\sigma_2)^3$ (which projects to the identity in $\mathrm{PSL}(2,\mathbb{Z})$) to ensure that the braids under consideration have total exponent 0. Then our main theorem implies the following result.

Theorem. *Let $\beta \in [B_3, B_3]$ be a positive word in σ_1^{-1} and σ_2. Then* rot_{B_3} *is extremal for $\sigma_2^{6n}\beta(\sigma_1\sigma_2)^{-3n}$ for all sufficiently large $n \in \mathbb{N}$.*

10.2 Hecke triangle groups

The modular group also generalizes to the Hecke triangle groups, originally introduced by Hecke [Hec36]. The *Hecke triangle group* G_q, $q \geq 3$, is the discrete subgroup of $\mathrm{PSL}(2,\mathbb{Z})$ generated by $S = \left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right)$ and $U = \left(\begin{smallmatrix} 0 & -1 \\ 1 & 2\cos(\pi/q) \end{smallmatrix}\right)$. The element S has order 2 and the element U has order q, and in fact $G_q \cong \langle S, U \mid S^2 = U^q = 1\rangle$. The quotient $\mathbb{H}^2/G_q$ is an orbifold with an order 2 cone point, an order q cone point, and a cusp, known as a *Hecke triangle surface.* When $q = 3$, this is just the ordinary modular group and modular surface.

As is the case for the modular group, the Hecke group G_q is naturally identified with the orbifold fundamental group of $\mathbb{H}^2/G_q$. Let P be the element of G_q corresponding to a negative loop around the cusp of $\mathbb{H}^2/G_q$. Then the theorem of Chapter 9 generalizes to the following result.

Theorem. *Let A be a hyperbolic element of the Hecke triangle group G_q. For all sufficiently large n, the geodesic on the Hecke triangle surface $\mathbb{H}^2/G_q$ corresponding to the element P^nA bounds an immersed orbifold.*

We now outline how the argument given in Chapter 9 can be used to prove this theorem. Let σ_q be the geodesic segment between the order 2 and order q cone points of $\mathbb{H}^2/G_q$. The total preimage of σ_q in $\mathbb{H}^2$, denoted $\widetilde{\sigma}_q$, is a regular q-valent tree. Let V_q be the component of $\mathbb{H}^2 \setminus \widetilde{\sigma}_q$ stabilized by the translation $\left(\begin{smallmatrix} 1 & 2\cos(\pi/q) \\ 0 & 1 \end{smallmatrix}\right)$. The boundary ∂V_q consists of circular arcs meeting at angles $2\pi/q$.

As in Section 9.1, a hyperbolic conjugacy class in G_q is represented by a geodesic γ in the Hecke triangle surface, and we again decompose γ into arcs between successive intersection points with σ_q. We denote by α_i those arcs whose lifts to V_q travel left and by β_i those arcs whose lifts to V_q travel right.

When n is sufficiently large, the geodesic γ corresponding to P^nA contains an arc α_1 that is very long compared to the other α_i and the β_i. As in Section 9.2, we choose lifts $\widetilde{\alpha}_i$ of α_i and $\widetilde{\beta}_i$ of β_i such that none of the $\widetilde{\alpha}_i$ are under $\widetilde{\alpha}_1$ and all of the $\widetilde{\beta}_i$ are under $\widetilde{\alpha}_1$ with five segments between them.

Partially gluing segments of V_q that are to the left of the $\widetilde{\alpha}_i$ and $\widetilde{\beta}_i$ as determined by the way γ was cut up, we obtain a surface with one boundary component mapping to γ and several other boundary components mapping onto σ. These other boundary components contain points around which the angle is some multiple of $2\pi/q$. Labeling these points with integers from 1 to $p-1$ (corresponding to the multiple of $2\pi/q$), we reduce to a combinatorial gluing problem, as in Section 9.3.

We again know that each component contains a 11 sequence and that one component contains a long sequence of 1s. Therefore we glue a 11 sequence from each component to the end of the long sequence of 1s to reduce to the case of one component to be glued. Let v denote the resulting sequence of numbers not in the long sequence of 1s. The complement v^c of v is a sequence that has the number $q-k$ across from where v has the number k. By folding in half the appropriate number of consecutive segments in the long sequence of 1s, we can form any number we wish, and we use this method to create V^c from the long sequence of 1s. Gluing v to v^c, we reduce to the case of all 1s except for one 2.

We may still assume the sequence of 1s is as long as necessary. We reduce its length using two moves: folding edges in half and gluing adjacent edges. Folding an edge in half turns a $1a$ edge into a point labeled by $a+1$, creating an order 2 cone point from the midpoint of the edge. Gluing adjacent edges is only allowed if the middle point is labeled by 1, and this move turns a $11a$ edge into a point labeled by $a+1$, creating an order q cone point from the middle 1. In the case $a = q-1$, these moves also force us to glue adjacent segments, creating a point labeled by 2. Notice that folding an edge does not change the sum of the labels on the vertices modulo q, whereas gluing adjacent edges reduces the sum of the labels on the vertices by 1. We first glue adjacent edges repeatedly, until the sum of the labels on the vertices is congruent to 0 modulo q. We then repeatedly fold edges involving the point not labeled by 1. Eventually we reach the case where there are only two points, one labeled by 1 and the other labeled by $q-1$. We then fold both edges in half, gluing the point labeled by 1 to the point labeled by $q-1$ and creating two order 2 cone points. This establishes our result for arbitrary Hecke triangle surfaces $\mathbb{H}^2/G_q$.

10.3 Further directions

One might try to generalize our results to other triangle orbifolds or to higher-strand braid groups. It seems plausible that the statement of Section 10.2 could be true for triangle orbifolds of type (p, q, ∞), $p, q \in \mathbb{N}$, rather than simply for triangle orbifolds of type $(2, q, \infty)$, and it would be interesting to try to modify our proof to work in this case.

Gambaudo–Ghys [GG05] used symplectic geometry to construct families of quasimorphisms on braid groups B_n that generalize the rotation quasimorphism, and it would also be interesting to study when these quasimorphisms are extremal. Since there is no known algorithm for computing stable commutator length in higher-strand braid groups, statements about the extremality of these quasimorphisms would give an indirect method for computing the stable commutator length of higher-strand braids.

Appendix A

Distribution of $n(W)$

In Section 7.3, we defined a function $n(W)$ to measure how "strongly" extremal or not extremal rot is for W. In particular, if rot is not extremal for W then $n(W)$ is the smallest power of R by which W needs to be multiplied in order to make rot extremal for the resulting element. If rot is extremal for W then $n(W)$ is nonpositive and measures the power of L by which W needs to be multiplied in order to make rot not extremal for the resulting element. In this appendix, we provide graphs of the distribution of $n(W)$ for words W of length 5, 6, 7, 8, and 9.

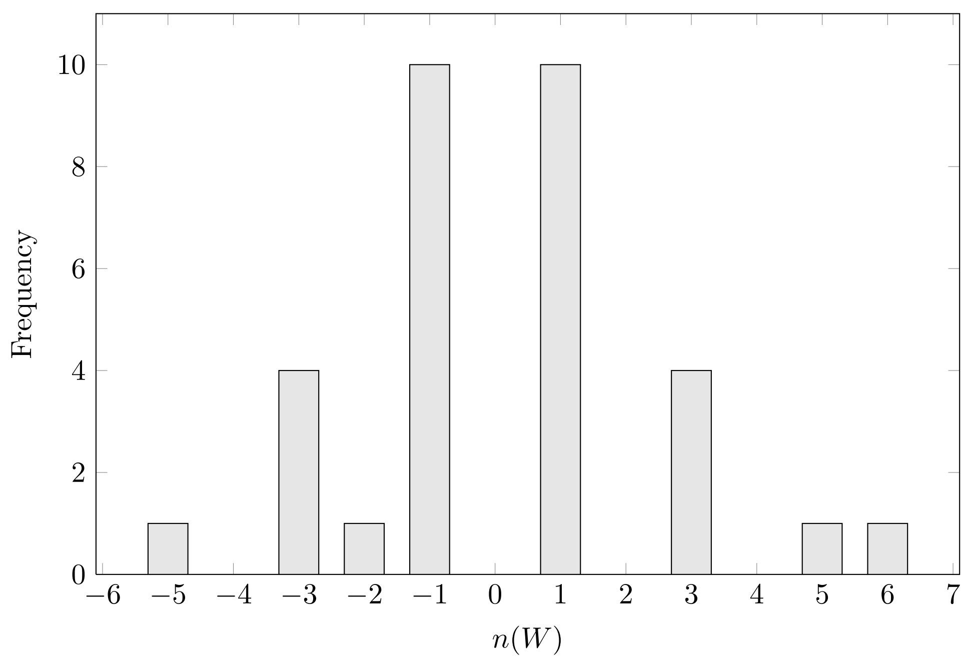

Figure A.1: Distribution of $n(W)$ for words of length 5

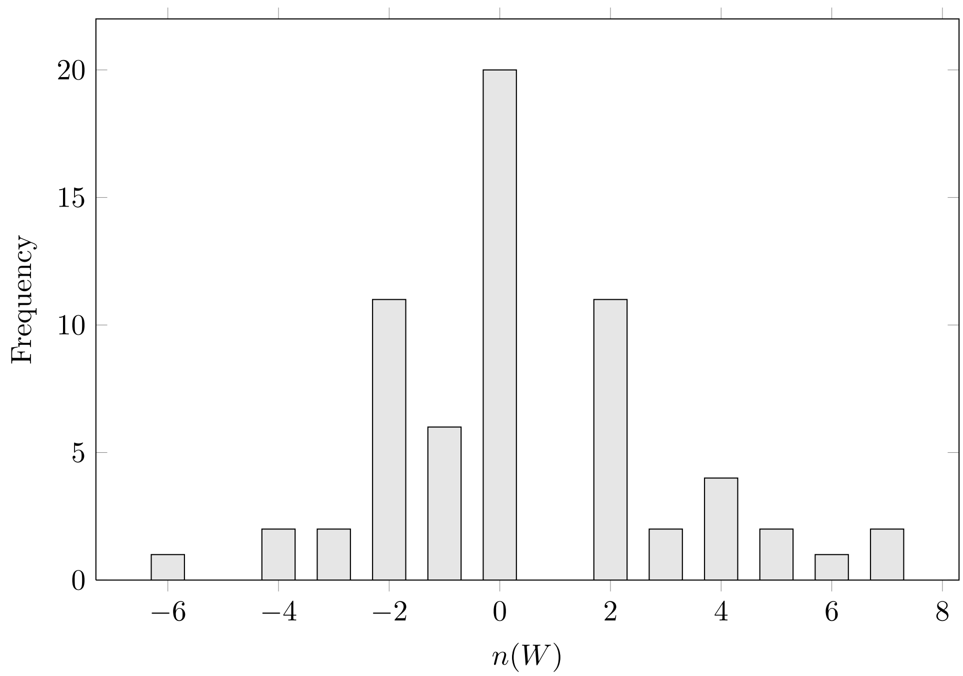

Figure A.2: Distribution of $n(W)$ for words of length 6

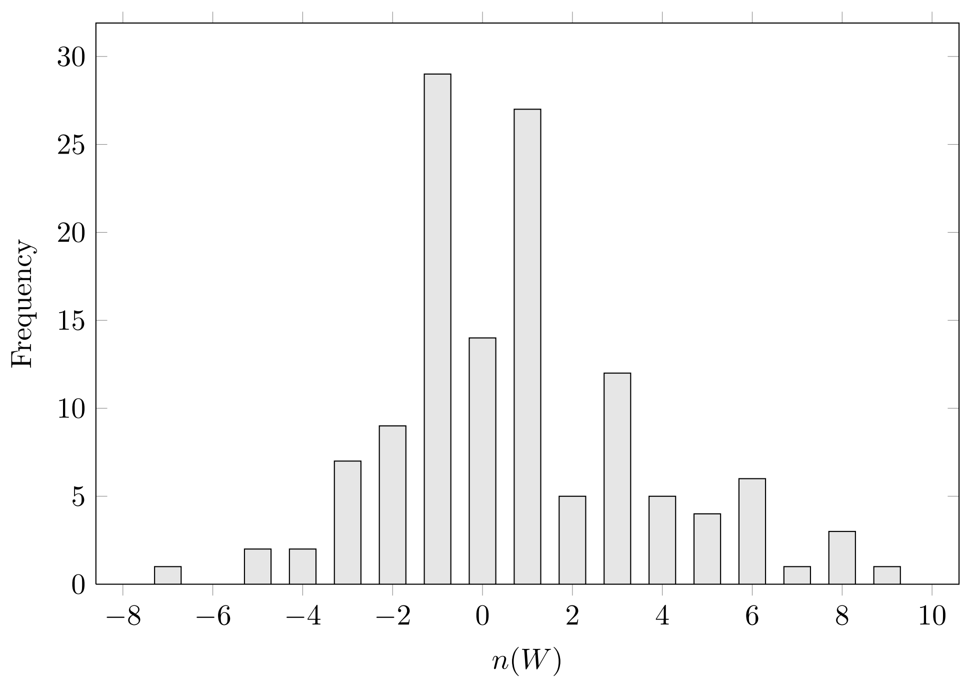

Figure A.3: Distribution of $n(W)$ for words of length 7

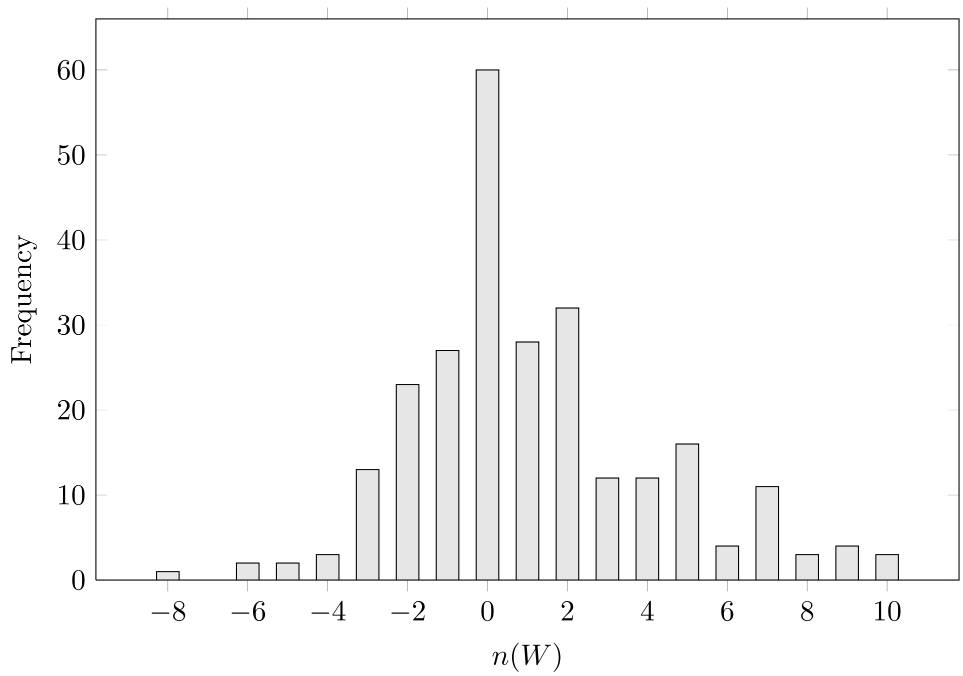

Figure A.4: Distribution of $n(W)$ for words of length 8

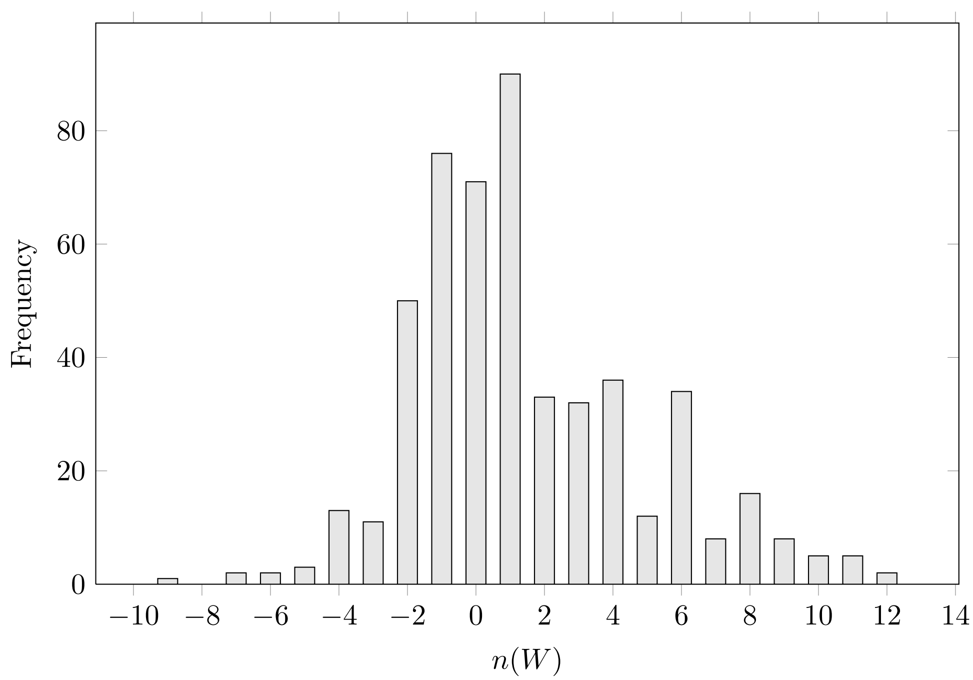

Figure A.5: Distribution of $n(W)$ for words of length 9

Bibliography

[Ati87] Michael Atiyah, *The logarithm of the Dedekind η-function*, Math. Ann. **278** (1987), no. 1-4, 335–380. MR 909232 (89h:58177)

[Bav91] Christophe Bavard, *Longueur stable des commutateurs*, Enseign. Math. (2) **37** (1991), no. 1-2, 109–150. MR 1115747 (92g:20051)

[BF02] Mladen Bestvina and Koji Fujiwara, *Bounded cohomology of subgroups of mapping class groups*, Geom. Topol. **6** (2002), 69–89 (electronic). MR 1914565 (2003f:57003)

[BF07] ———, *Quasi-homomorphisms on mapping class groups*, Glas. Mat. Ser. III **42(62)** (2007), no. 1, 213–236. MR 2332668 (2008k:57002)

[BH] Michael Björklund and Tobias Hartnick, *Biharmonic functions on groups and limit theorems for quasimorphisms along random walks*, arXiv:1005.0077.

[Bro81] Robert Brooks, *Some remarks on bounded cohomology*, Riemann surfaces and related topics: Proceedings of the 1978 Stony Brook Conference (State Univ. New York, Stony Brook, N.Y., 1978), Ann. of Math. Stud., vol. 97, Princeton Univ. Press, Princeton, N.J., 1981, pp. 53–63. MR 624804 (83a:57038)

[Cal08] Danny Calegari, *What is... stable commutator length?*, Notices Amer. Math. Soc. **55** (2008), no. 9, 1100–1101. MR 2451345

[Cal09a] ———, *Faces of the scl norm ball*, Geom. Topol. **13** (2009), no. 3, 1313–1336. MR 2496047

[Cal09b] ———, *scl*, MSJ Memoirs, vol. 20, Mathematical Society of Japan, Tokyo, 2009. MR 2527432

[Cal09c] ———, *Stable commutator length is rational in free groups*, J. Amer. Math. Soc. **22** (2009), no. 4, 941–961. MR 2525776 (2010k:57002)

[CF10a] Danny Calegari and Koji Fujiwara, *Combable functions, quasimorphisms, and the central limit theorem*, Ergodic Theory Dynam. Systems **30** (2010), no. 5, 1343–1369.

[CF10b] ———, *Stable commutator length in word-hyperbolic groups*, Groups Geom. Dyn. **4** (2010), no. 1, 59–90. MR 2566301 (2011a:20109)

[CL11] Danny Calegari and Joel Louwsma, *Immersed surfaces in the modular orbifold*, Proc. Amer. Math. Soc **139** (2011), 2295–2308.

[CM] Danny Calegari and Joseph Maher, *Statistics and compression of scl*, arXiv:1008.4952.

[Cul81] Marc Culler, *Using surfaces to solve equations in free groups*, Topology **20** (1981), no. 2, 133–145. MR 605653 (82c:20052)

[CWa] Danny Calegari and Alden Walker, *Isometric endomorphisms of free groups*, arXiv:1101.4055.

[CWb] ———, *scallop*, computer program, available from `http://www.its.caltech.edu/~dannyc/`.

[Edm75] Charles C. Edmunds, *On the endomorphism problem for free groups*, Comm. Algebra **3** (1975), 1–20. MR 0369530 (51 #5763)

[Edm79] ———, *On the endomorphism problem for free groups. II*, Proc. London Math. Soc. (3) **38** (1979), no. 1, 153–168. MR 520977 (80d:20025)

[EF97] David B. A. Epstein and Koji Fujiwara, *The second bounded cohomology of word-hyperbolic groups*, Topology **36** (1997), no. 6, 1275–1289. MR 1452851 (98k:20088)

[GG05] Jean-Marc Gambaudo and Étienne Ghys, *Braids and signatures*, Bull. Soc. Math. France **133** (2005), no. 4, 541–579. MR 2233695 (2007d:57014)

[Ghy07] Étienne Ghys, *Knots and dynamics*, International Congress of Mathematicians. Vol. I, Eur. Math. Soc., Zürich, 2007, pp. 247–277. MR 2334193 (2008k:37001)

[Ham] Ursula Hamenstädt, *Lines of minima in Outer space*, arXiv:0911.3620.

[Hec36] E. Hecke, *Über die Bestimmung Dirichletscher Reihen durch ihre Funktionalgleichung*, Math. Ann. **112** (1936), no. 1, 664–699. MR 1513069

[KM] Jeremy Kahn and Vladimir Markovic, *Immersing almost geodesic surfaces in a closed hyperbolic three manifold*, arXiv:0910.5501.

[KM94] Robion Kirby and Paul Melvin, *Dedekind sums, μ-invariants and the signature cocycle*, Math. Ann. **299** (1994), no. 2, 231–267. MR 1275766 (95h:11042)

[Kot04] D. Kotschick, *What is… a quasi-morphism?*, Notices Amer. Math. Soc. **51** (2004), no. 2, 208–209. MR 2026941

[KU07] Svetlana Katok and Ilie Ugarcovici, *Symbolic dynamics for the modular surface and beyond*, Bull. Amer. Math. Soc. (N.S.) **44** (2007), no. 1, 87–132. MR 2265011 (2007j:37050)

[LOST10] Martin W. Liebeck, E. A. O'Brien, Aner Shalev, and Pham Huu Tiep, *The Ore conjecture*, J. Eur. Math. Soc. (JEMS) **12** (2010), no. 4, 939–1008. MR 2654085 (2011e:20016)

[Ore51] Oystein Ore, *Some remarks on commutators*, Proc. Amer. Math. Soc. **2** (1951), 307–314. MR 0040298 (12,671e)

[Poi81] H. Poincaré, *Mémoire sur les courbes défines par une équation différentielle*, J. de Mathématiques **7** (1881), 375–422.

[Poi82] ———, *Mémoire sur les courbes défines par une équation différentielle*, J. de Mathématiques **8** (1882), 251–296.

[Rhe68] A. H. Rhemtulla, *A problem of bounded expressibility in free products*, Proc. Cambridge Philos. Soc. **64** (1968), 573–584. MR 0225889 (37 #1480)

[Ser85] Caroline Series, *The modular surface and continued fractions*, J. London Math. Soc. (2) **31** (1985), no. 1, 69–80. MR 810563 (87c:58094)