UCLA
UCLA Electronic Theses and Dissertations

Title

Thurston theory and polymorphic maps

Permalink

https://escholarship.org/uc/item/8nv2361q

Author

Smith, Zachary Bruce

Publication Date

2024

Peer reviewed|Thesis/dissertation

eScholarship.org Powered by the California Digital Library
University of California

UNIVERSITY OF CALIFORNIA

Los Angeles

Thurston theory and polymorphic maps

A dissertation submitted in partial satisfaction of the
requirements for the degree Doctor of Philosophy
in Mathematics

by

Zachary Bruce Smith

2024

© Copyright by

Zachary Bruce Smith

2024

ABSTRACT OF THE DISSERTATION

Thurston theory and polymorphic maps

by

Zachary Bruce Smith
Doctor of Philosophy in Mathematics
University of California, Los Angeles, 2024
Professor Mario Bonk, Chair

Let $A \subseteq S^2$ be a finite set of points on the 2-sphere. A Thurston map $f\colon (S^2, A) \to (S^2, A)$ induces an associated holomorphic pullback map $\sigma_f\colon \mathcal{T}_A \to \mathcal{T}_A$, where $\mathcal{T}_A$ is the Teichmüller space of the marked sphere (S^2, A). By [KPS16] this pullback map is known to satisfy a family of functional identities of the form $\sigma_f \circ g = \widetilde{g} \circ \sigma_f$ where $g, \widetilde{g} \in \mathrm{Aut}(\mathcal{T}_A)$. In the case of four marked points we have $\mathcal{T}_A = \mathbb{H}$, and such maps are known as *polymorphic maps*. In this dissertation we develop a general framework for studying the cusp dynamics of polymorphic maps of the upper half-plane $\mathbb{H}$. Applying this framework to the Thurston pullback map, we obtain a new proof of Thurston's characterization theorem in the special case of four marked points. We also obtain new progress on the finite curve attractor conjecture of [Pil22]. Specifically, we prove that if a Thurston map $f\colon (S^2, A) \to (S^2, A)$ is totally unobstructed in the sense that all of its Thurston multipliers are strictly less than one, then (f, A) has a finite curve attractor.

The dissertation of Zachary Bruce Smith is approved.

Ko Honda

Mikhail Hlushchanka

Mikhail Khitrik

Christina Kim

Mario Bonk, Committee Chair

University of California, Los Angeles

2024

To my parents.

Contents

List of Figures

Notation

We record here the most important notation used in this thesis.

The sets $\mathbb{N}, \mathbb{Z}, \mathbb{Q}, \mathbb{R}$ and $\mathbb{C}$ will carry their usual meanings: natural numbers (starting from 1), integers, rationals, real numbers, and complex numbers respectively. We will denote the nonnegative numbers $\mathbb{R}_{\geq 0} := \mathbb{R} \cap \{x \in \mathbb{R} : x \geq 0\}$. Similar notation will be used for positive numbers $\mathbb{R}_{>0}$.

The cardinality of a set X is denoted by $|X|$ and the identity map on X by id_X.

If X is a topological space and $A \subseteq X$, then $\overline{A}$ denotes the closure, $\mathrm{int}(A)$ denotes the interior, and ∂A the boundary of A in X.

If $f \colon X \to X$ is a map and $n \in \mathbb{N}$, then

$$f^n := \underbrace{f \circ \cdots \circ f}_{n \text{ factors}}.$$

We denote the open unit disk in $\mathbb{C}$ as $\mathbb{D} := \{z \in \mathbb{C} : |z| < 1\}$. We denote the upper half-plane as $\mathbb{H} := \{\tau \in \mathbb{C} : \operatorname{Im} \tau > 0\}$.

We denote the Riemann sphere as $\widehat{\mathbb{C}} := \mathbb{C} \cup \{\infty\}$. Similarly, we denote the extended real line as $\widehat{\mathbb{R}} := \mathbb{R} \cup \{\infty\}$ and the extended rationals as $\widehat{\mathbb{Q}} := \mathbb{Q} \cup \{\infty\}$. Typically we will treat $\mathbb{H}$ as a subspace of the Riemann sphere, so $\partial\mathbb{H} = \widehat{\mathbb{R}}$.

ACKNOWLEDGMENTS

I would like to thank Mario Bonk, my PhD advisor, for both introducing me to the topic of Thurston maps and for his constant support and guidance. I would have neither a job nor a degree without his generous help. More importantly, I would like to thank him for the deep insights and truly beautiful mathematics he has shared with me.

I also give thanks to Mikhail Hlushchanka for the time he spent advising me when I first started this project and for the numerous helpful comments and remarks he has given me since then.

I thank the rest of my committee members: Michael Hitrik and Ko Honda, as well as last-minute addition Inwon Kim, who saved me from a bureaucratic snafu.

Daniel Meyer and Kevin Pilgrim have both supported my career and given me helpful comments about this and related work.

I frequently used Darragh Glynn's Python program for visualizing curve pullbacks of flapped Lattès maps during my initial studies, and it was used to help make Figure 2. Walter Parry's `Pf4` program also greatly aided my investigations. I thank him for both making it and for taking the time to explain how it works to me.

I would not have made it through graduate school without my friends, especially Harris Khan, Timothy Smits, Nicholas Boschert, Matthew Stone, and Abhishek Chaturvedi.

Thank you, Ana Saechao, for the happiness you have given me the past three years.

And lastly, I would like to thank my family, who have always supported and encouraged me. I cannot even begin to describe the love and gratitude I have for my parents, Mark and Kelly Smith. Since the English language fails me, I present this thesis instead. I dedicate it to them.

This work was partially supported by NSF grant DMS-2054987.

VITA

2019	M.A. (Mathematics), UCLA
2017	B.S. (Mathematics), University of Chicago

CHAPTER 1

Introduction

1.1. Thurston theory

The motivation of this thesis is to answer certain dynamical questions that arise in context of Thurston maps. A *Thurston map* is an orientation-preserving branched cover $f\colon S^2 \to S^2$ that is not a homeomorphism and such that each of its critical points has a finite forward orbit; it is considered together with the data of a finite forward-invariant set of marked points A that contains said forward orbits. These maps are named after William Thurston, who introduced them as combinatorial models of postcritically-finite rational maps.

In 1982 Thurston presented his celebrated characterization theorem, which gives necessary and sufficient conditions for when a Thurston map is "realized" by a rational map in a suitable sense. The proof, as explicated by Douady and Hubbard [DH93], uses an analytic map on a suitable Teichmüller space $\sigma_f\colon \mathcal{T}_A \to \mathcal{T}_A$ that is induced by f by pulling back complex structures. The question of whether f is realized by a rational map then reduces to whether σ_f has a fixed point in $\mathcal{T}_A$.

The main requirement in Thurston's theorem is the (non)existence of Jordan curves in $S^2 \setminus A$ with special invariance properties. More generally, every Thurston map induces a *pullback relation* on isotopy classes of Jordan curves in $S^2 \setminus A$. The restriction $f\colon S^2 \setminus f^{-1}(A) \to S^2 \setminus A$ is a covering map, so if $\gamma \subseteq S^2 \setminus A$ is a Jordan curve, then a component $\widetilde{\gamma}$ of $f^{-1}(\gamma)$ will also be a Jordan curve in $S^2 \setminus A$. We say that $\widetilde{\gamma}$ is a *pullback* of γ by f. Lifting isotopies shows that the set of isotopy classes of $f^{-1}(\gamma)$ rel. A depends only on the isotopy class of γ rel. A. See Figure 2.

With this in mind, understanding the dynamics of Thurston maps f, finding suitable invariants by which to classify such maps, and understanding the dynamics of the associated

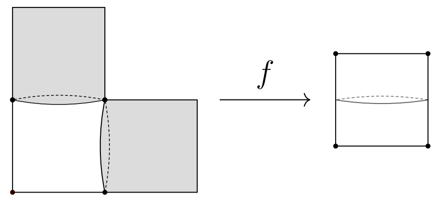


FIGURE 1. A combinatorial depiction of a Thurston map with four marked points. White tiles are mapped to the front face of the "pillow" on the right (which is a copy of S^2), and gray tiles are mapped to the back face.

pullback map σ_f are all closely related to the curve pullback relation described above. See, for example, [BEKP09],[Sel12], [Koc13], [KPS16], and [Pil22].

One of the major open problems in the study of the pullback relation on curves is the following (see [Lod13] and [Pil22]):

Conjecture 1.1 (Finite curve attractor (FCA) conjecture)**.** *Let $f\colon (S^2, A) \to (S^2, A)$ be a Thurston map that is realized by a rational map and that is not of type $(2,2,2,2)$. Then there is a finite set $\mathcal{A}(f)$ of Jordan curves in $S^2 \setminus A$ with the following property: for every Jordan curve γ in $S^2 \setminus A$ there is a positive integer $N(\gamma)$ such that, for $n \geq N(\gamma)$, all pullbacks $\widetilde{\gamma}$ of γ under f^n are contained in $\mathcal{A}(f)$ up to isotopy* rel. A*.*

The assumption that f is "not of type $(2,2,2,2)$" is a technical requirement that we will make precise later. For now, let it suffice to say that most Thurston maps meet this requirement. The FCA conjecture thus claims that, for most rational Thurston maps, if one iterates the pullback relation on curves, then one eventually lands in a finite set of isotopy classes.

The FCA conjecture has been the subject of much study. Koch, Pilgrim, and Selinger have shown it holds when the associated virtual endomorphism is contracting [KPS16]. Belk, Lanier, Margalit, and Winarski have proven it for all postcritically-finite polynomials [BLMW22]. Hlushchanka has proven it for rational maps where every critical point is a fixed point [Hlu19]. In the case of four postcritical points it is also known for certain NET maps [Lod13, FKK+17], for all quadratic non-Lattès maps [KL19], and for maps obtained

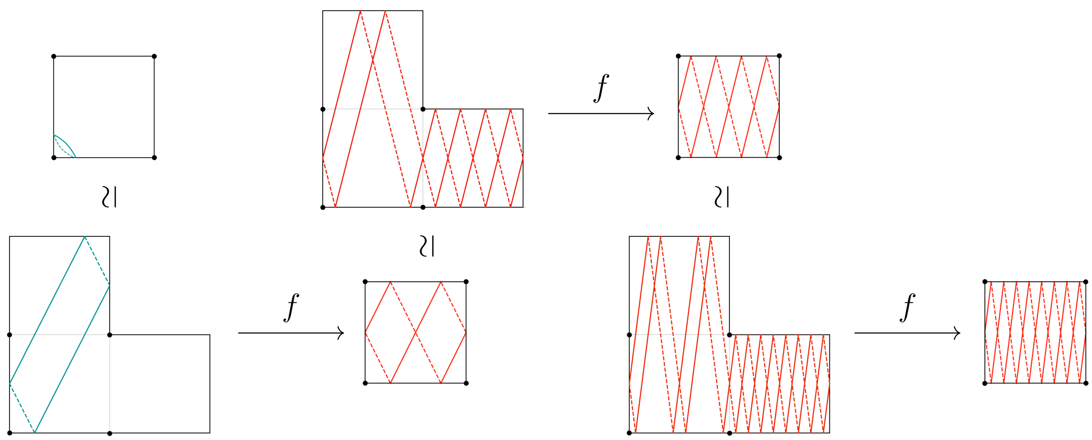


FIGURE 2. Some iterates of the pullback relation on curves for the map from Figure 1, where the isotopy class relative to the marked points is redrawn in the square pillow after each pullback. Each essential curve class has a representative that is a geodesic with extended rational slope in the square pillow picture, so there is a bijection between essential curve classes and the extended rationals in this case. Using this identification, the above picture represents $\mu_f\colon 8/1 \mapsto 4/1 \mapsto 2/1 \mapsto o$, where o denotes the "trivial" curve classes.

from a certain blowup of 2×2-Lattès maps [BHI21]. The general conjecture—even in the simplest nontrivial case of four marked points—still remains open.

It turns out that the Thurston pullback map and the pullback relation on curves are related to each other in a significant way: the map $\sigma_f\colon \mathcal{T}_A \to \mathcal{T}_A$ admits an extension to the *Weil–Petersson boundary* of Teichmüller space, and the pullback relation on curves is then encoded in the boundary behavior of this extension. This correspondence was demonstrated by Selinger in [Sel12]. In the case where $|A| = 4$, there is the pleasingly simple description of all these objects. The Teichmüller space is just the upper half-plane $\mathbb{H}$ and the Weil–Petersson boundary is just the extended rationals $\widehat{\mathbb{Q}} = \mathbb{Q} \cup \{\infty\}$. On the other hand, the isotopy classes $[\gamma]$ of (essential) Jordan curves of $S^2 \setminus A$ are in bijection with $\widehat{\mathbb{Q}}$; roughly speaking, one can think of each curve as equivalent to a geodesic with extended rational slope in the square pillow picture as in Figure 2. Thus the pullback relation on curves is just a function $\mu_f\colon \widehat{\mathbb{Q}} \cup \{o\} \to \widehat{\mathbb{Q}} \cup \{o\}$, where o represents isotopy classes of "trivial curves."

As we shall see later, these two pullback maps are essentially the same on the extended rationals (up to conjugation by negative inversion).

In this special setting, where $|A| = 4$ and $\sigma_f\colon \mathbb{H} \to \mathbb{H}$, it is natural to ask whether (and to what extent) classical function theory can be used to understand the pullback map σ_f and its boundary dynamics. After all, the boundary behavior of holomorphic self-maps of hyperbolic 2-space is well-understood via the notion of nontangential limits, angular derivatives, and the Julia–Wolff–Carathéodory theory.

The answer to the above question is *yes*, but only if we make use of a special additional structure that underlies the Thurston pullback map. Indeed, let G be the pure mapping class group of the marked sphere (S^2, A), which we identify with its action on $\mathcal{T}_A$ by isometries. Then a general Thurston pullback map $\sigma_f\colon \mathcal{T}_A \to \mathcal{T}_A$ satisfies a family of functional identities of the form

$$\sigma_f \circ g = \varphi_f(g) \circ \sigma_f$$

where g ranges over a finite-index subgroup H of G, and $\varphi_f : H \to G$ is a homomorphism (see [KPS16]). In the case where $|A| = 4$, we have $\mathcal{T}_A = \mathbb{H}$ and we can think of H and G as discrete subgroups of $\mathrm{Aut}(\mathbb{H})$, i.e., Fuchsian groups. Analytic maps on the upper half-plane $\mathbb{H}$ satisfying such a collection of functional identities are called *polymorphic maps* in the classical literature. We are henceforth interested in the dynamics of these polymorphic maps.

1.2. Polymorphic maps

The study of polymorphic maps has a venerable history going all the way back to Fricke and Klein, who originally studied them in relation to elliptic modular forms (see, e.g., [Fri12]). The term "polymorphic" was coined by Fricke (see [BWFH21, pp. 432]) so as to contrast these functions with automorphic functions. More recently they have been the subject of investigations by Hejhal [Hej75, Hej76], Mejía and Pommerenke [MP12a, MP12b, MP08], and many others.

In general, a polymorphic map is a meromorphic function $\sigma\colon \mathbb{H} \to \widehat{\mathbb{C}}$ that satisfies the intertwining relation

$$\sigma \circ g = \varphi(g) \circ \sigma$$

for all g in a Fuchsian group G, where $\varphi\colon G \to \mathrm{PSL}(2, \mathbb{C})$ is a homomorphism into the group of Möbius transformations. This setting is inappropriate for dynamics, so throughout this thesis we will use the following restricted definition: let G and G' be finite coarea Fuchsian groups, and let $\varphi\colon G \to G'$ be a homomorphism. We will say a nonconstant holomorphic function $\sigma\colon \mathbb{H} \to \mathbb{H}$ is *φ-polymorphic* (or just *polymorphic*) if σ satsifies the intertwining relation

$$\sigma \circ g = \varphi(g) \circ \sigma$$

for all $g \in G$. We shall call the homomorphism φ an *intertwining homomorphism.* In the special case where G and G' are finite-index subgroups of the modular group $\mathrm{PSL}(2, \mathbb{Z})$, we will call the map σ *modularly polymorphic.*

The polymorphicity property gives the map $\sigma\colon \mathbb{H} \to \mathbb{H}$ just enough additional structure to make the study of its boundary dynamics tractable. The boundary points can be divided into two classes: cusps and conical limit points of G. At *contact cusps* (those which themselves map to cusps) the angular derivative of σ both exists and can be computed from data encoded by the intertwining homomorphism. For simplicity we only formulate these results for modularly polymorphic maps, but similar results can be formulated even when G and G' are not subgroups of the modular group so long as $\mathrm{cusp}(G') \subseteq \mathrm{cusp}(G)$.

Since a Thurston map (f, A) with $|A| = 4$ induces a φ_f-polymorphic Thurston pullback map $\sigma_f\colon \mathbb{H} \to \mathbb{H}$ where φ_f is a homomorphism of finite-index subgroups of the modular group, the above considerations give us access to angular derivatives as a new tool to understand the boundary dynamics of σ_f. It also suggests that some problems, such as the FCA conjecture, have an analogous formulation for polymorphic maps beyond those which arise as Thurston pullbacks. In paricular, we conjecture the following:

Conjecture 1.2 (Finite cusp attractor (FCA) conjecture)**.** *Suppose $\sigma\colon \mathbb{H} \to \mathbb{H}$ is modularly polymorphic, has an interior fixed point, and is not an automorphism. Then σ has a finite cusp attractor in the following sense: there is a finite set $\mathcal{A} \subseteq \widehat{\mathbb{Q}}$ such that, for every $r \in \widehat{\mathbb{Q}}$, either $\sigma^N(r) \in \mathbb{H}$ for some positive integer N, or $\sigma^n(r) \in \mathcal{A}$ for all n sufficiently large.*

Since the finite cusp attractor conjecture for polymorphic maps would imply the finite curve attractor conjecture for Thurston maps with four marked points, we shall not distinguish the acronyms and instead use FCA to refer to both problems.

With all of this in mind, the central aims of this thesis are as follows: (i) to collect and codify the machinery of polymorphic maps in the dynamical setting (i.e., modularly polymorphic maps); (ii) to recast known facts about Thurston maps with four marked points in terms of this machinery; and (iii) to leverage this machinery to prove new results about Thurston maps.

1.3. Outline and main results

We will now outline the structure of the thesis and describe our results.

In Chapter 2 we will present the key definitions and constructions of Thurston theory. We provide a careful description of the problems we are most interested in—namely, Thurston's characterization theorem and the FCA conjecture. We also discuss how to phrase these problems in terms of the Thurston pullback map $\sigma_f\colon \mathcal{T}_A \to \mathcal{T}_A$ and prove important properties of this map. In Section 2.4 we describe the set of homeomorphisms which lift under a Thurston map (f, A). The mapping classes of these liftable homeomorphisms $\mathrm{LMod}(f, A)$ form a finite-index subgroup of the pure mapping class group $\mathrm{PMod}(S^2, A)$. This result was established in [Pil12, KPS16] and we present a full proof in Theorem 2.19. Using this we establish the polymorphic nature of the Thurston pullback σ_f in Theorem 2.30. In the last section of the chapter we specialize the considerations to the case where $|A| = 4$ and elaborate on the niceties of this setting, some of which we described earlier in the introduction.

In Chapter 3 we provide the additional background necessary for understanding polymorphic self-maps of hyperbolic 2-space. In the first three sections we review Fuchsian groups, their limit sets, and the special properties of the modular group $\mathrm{PSL}(2,\mathbb{Z})$. The remainder of the chapter is devoted to the classical theory of boundary behavior for holomorphic self-maps of hyperbolic 2-space. Nontangential limits and angular derivatives are defined and we establish an important estimate for the growth of horoballs in terms of the angular derivative in Corollary 3.18.

In Chapter 4 we define and study polymorphic maps. We show that a φ-polymorphic map with $\varphi\colon G \to G'$ admits a continuous extension to a map $\mathbb{H}\cup\mathrm{cusp}(G) \to \mathbb{H}\cup\mathrm{cusp}(G')$. We also present a rigidity result due to Pommerenke that says φ-polymorphic maps are uniquely determined by their intertwining homomorphism φ. We then restrict the setting to modularly polymorphic maps, for which we can compute the angular derivatives. The angular derivative is determined by a ratio of integer *cusp widths*. These ratios give us the notion of *cusp multipliers* $\lambda(r)$ and the reciprocal notion of *cusp dilation factors* $\delta(r) = 1/\lambda(r)$. In particular, we prove

Theorem 4.5. *Suppose $\sigma\colon \mathbb{H} \to \mathbb{H}$ is modularly polymorphic. Then we have the following properties:*

(i) *the map σ extends to a continuous map $\sigma\colon \mathbb{H}\cup\widehat{\mathbb{Q}} \to \mathbb{H}\cup\widehat{\mathbb{Q}}$*

(ii) *if $r \in \widehat{\mathbb{Q}}$ is a contact cusp in the sense that $r' = \sigma(r) \in \widehat{\mathbb{Q}}$, then the angular derivative $\sigma'(r)$ is finite. Moreover, if $r = p/q$ for coprime integers p and $q > 0$, and likewise $r' = p'/q'$ for coprime integers p' and $q' > 0$, then*

$$\sigma'(r) = \frac{1}{\lambda(r)}\left(\frac{q}{q'}\right)^2 = \delta(r)\left(\frac{q}{q'}\right)^2,$$

where $\lambda(r)$ is the cusp multiplier and $\delta(r) = 1/\lambda(r)$ is the cusp dilation factor. If r or r' is ∞, then we respectively take q or q' to be 1 in the above formula.

Since finite-index subgroups of the modular group have finitely many cusp classes, the set of cusp multipliers attained by σ, denoted $S_\lambda(\sigma)$, is a finite list. The same is also true for for the set of cusp dilation factors $S_\delta(\sigma)$.

In the last section of Chapter 4 we apply the Denjoy–Wolff fixed point theorem to our setting. Recall that the Denjoy-Wolff theorem says a nonidentity holomorphic self-map of $\mathbb{H}$ has a unique nonrepelling fixed point in $\mathbb{H}$, which is called the Denjoy–Wolff (DW) point. We present a second rigidity theorem that shows the DW point of a polymorphic map σ can only occur at a boundary conical limit point when σ is a Möbius transformation. Presuming G admits at least one cusp class, then we can further say that σ is an automorphism of $\mathbb{H}$. This together with Theorem 4.5 will imply

Theorem 4.9. *Suppose $\sigma\colon \mathbb{H} \to \mathbb{H}$ is modularly polymorphic and not an automorphism. Then σ has an interior fixed point $\tau_0 \in \mathbb{H}$ if and only if there is no fixed cusp $r \in \widehat{\mathbb{Q}}$ with dilation factor $\delta(r) \leq 1$.*

Since the cusp multiplier $\lambda(r)$ of the Thurston pullback map σ_f coincides with the Thurston multiplier of the corresponding curve (which has slope class $s = -1/r$), applying the above theorem to the Thurston pullback map σ_f produces a new proof of Thurston's characterization theorem in the case of four marked points:

Theorem 4.10 (Thurston's criterion for four points)**.** *Let (f, A) be a Thurston map with $|A| = 4$ and suppose f is not of type $(2, 2, 2, 2)$. Then f is combinatorially equivalent to a rational map if and only if f has no Thurston obstruction, i.e., there is no f-invariant essential Jordan curve γ with $\lambda_f(\gamma) \geq 1$.*

Note that an essential Jordan curve γ in (S^2, A) is said to *f-invariant* if each essential pullback of γ under f is isotopic to γ rel. A. The quantity λ_f is the *Thurston multiplier* of γ; see Section 2.5 for the precise definition.

Our proof of the above theorem is remarkably shorter and different in flavor from Thurston's original proof as presented in [DH93], which mostly works on the level of moduli space and us a nontrivial amount of hyperbolic geometry.

There are other facts about the pullback relation that can be proven using these calculations. For example, we can apply an inequality due to Cowen and Pommerenke [CP82] to obtain

Corollary 4.12. *Suppose $\sigma\colon \mathbb{H} \to \mathbb{H}$ is a modularly polymorphic map with interior DW point $\tau_0 \in \mathbb{H}$. Define*

$$C := \min\{\sigma'(r) : r \in \widehat{\mathbb{Q}} \text{ and } \sigma(r) = r\}.$$

Then $C > 1$ and

$$|\operatorname{Fix}(\sigma \cap \widehat{\mathbb{Q}})| \leq \frac{1}{C-1}\left(\frac{1 - |\sigma'(\tau_0)|^2}{|1 - \sigma'(\tau_0)|^2}\right).$$

In particular, σ has finitely many fixed cusps.

Applying this to the Thurston pullback map gives

Corollary 4.13. *Suppose (f, A) is a rational Thurston map with $|A| = 4$ that is not of type* $(2, 2, 2, 2)$. *Then f has finitely many f-invariant essential curve classes.*

We also obtain a new estimate for the number of fixed curves classes. The finiteness portion of this corollary was previously known even in the general case (see [Pil12, Theorem 1.5]). [Par18, Theorem 10.1] has a similar estimate for the $|A| = 4$ case, but it is given in terms of group data rather than the Thurston pullback map itself.

In Chapter 5, we return our attention to the FCA conjecture for polymorphic maps. Combining our machinery with techniques previously used by the author in [Smi24] we prove

Theorem 5.1. *Let $\sigma\colon \mathbb{H} \to \mathbb{H}$ be a modularly polymorphic map with spectrum of cusp dilation factors satisfying* $\min S_\delta(\sigma) > 1$. *Then σ has a finite cusp attractor.*

Of course, there is also a corresponding result for Thurston maps. We say a Thurston map (f, A) is *totally unobstructed* if all of its Thurston multipliers satisfy $\lambda_f(r) < 1$. Then the above theorem implies

Corollary 5.12. *If (f, A) is a totally unobstructed Thurston map, then (f, A) has a finite global curve attractor.*

Totally unobstructed Thurston maps is a nonempty class and we provide several examples. Thus Corollary 5.12 represents new progress on the FCA conjecture.

Finally, in Chapter 6, we formulate some conjectures and discuss possible directions for future study.

CHAPTER 2

Thurston theory

2.1. Basic definitions

In this section we summarize the basic definitions and constructions of Thurston theory. We omit proofs of the stated propositions, but they can be found in any standard reference on the subject (e.g., [DH93], [BCT14], [BM17, Chapter 2]).

Let S^2 be the 2-sphere, which we take to be oriented. We say $f\colon S^2 \to S^2$ is a *branched covering map* if it is: (i) continuous and surjective; and (ii) expressible as a power map $z \mapsto z^d$ in local homeomorphic coordinates, where d is some positive integer. In other words, given any $x \in S^2$, we require that there are open topological disks $U, V \subseteq S^2$ with $x \in U$ and $y := f(x) \in V$ as well as orientation-preserving homeomorphisms $\varphi\colon U \to \mathbb{D}$ and $\psi\colon V \to \mathbb{D}$ with $\varphi(x) = 0$ and $\psi(y) = 0$ such that

$$(\psi \circ f \circ \varphi^{-1})(z) = z^d$$

for all $z \in \mathbb{D}$. This is equivalent to the commutativity of the following diagram:

$$\begin{array}{ccc} x \in U \subseteq S^2 & \xrightarrow{\quad f \quad} & y \in V \subseteq S^2 \\ \Big\downarrow{\scriptstyle \varphi} & & \Big\downarrow{\scriptstyle \psi} \\ 0 \in \mathbb{D} & \xrightarrow[z \mapsto z^d]{} & 0 \in \mathbb{D}. \end{array}$$

The positive integer d in the above definition only depends on f and the point $x \in S^2$; it is called the *local degree* of f at x, and is denoted by $\deg(f, x)$. The topological degree of $f\colon S^2 \to S^2$ is denoted by $\deg(f)$. The two quantities are related by the formula

$$\sum_{x \in f^{-1}(y)} \deg(f, x) = \deg(f)$$

where $y \in S^2$ is any point.

A point $c \in S^2$ where the local mapping degree $\deg(f, c)$ is at least 2 is a *critical point*, and we denote the set of critical points by C_f. The immediate images of critical points are called *critical values*, and the set of all critical values is $V_f := f(C_f)$. If $f\colon S^2 \to S^2$ is a branched covering map, then it restricts to an ordinary covering map from $S^2 \setminus f^{-1}(V_f)$ to $S^2 \setminus V_f$.

A *postcritical point* of f is a point $p \in S^2$ of the form $p = f^n(c)$ where c is a critical point of f and f^n denotes the nth iterate of f for a nonnegative integer n. We denote the set of all postcritical points P_f, so

$$P_f = \bigcup_{n \geq 1} \{f^n(c) : c \in C_f\}.$$

Denote the cardinality of the above set by $|P_f|$. If $|P_f|$ is finite, then f is said to be *postcritically-finite*.

Definition 2.1. Let $f\colon S^2 \to S^2$ be an orientation-preserving postcritically-finite branched covering map with $\deg(f) \geq 2$. Let $A \subseteq S^2$ be a finite set of marked points with the properties $P_f \subseteq A$ and $f(A) \subseteq A$. We call the map of pairs $f\colon (S^2, A) \to (S^2, A)$ a *Thurston map*.

Remarks 2.2.

(1) In the above definition it may be the case that P_f is a proper subset of A. This slightly extends the standard definition of Thurston maps, where $A = P_f$. We will sometimes use the phrase *marked Thurston map* to emphasize the more general setting. We refer the reader to [BCT14] for a general overview of Thurston theory with markings and how it differs from the original setting of [DH93].

(2) The $(\cdot, A)$ portion of the notation will be suppressed when the marked set is clear.

(3) Conversely, when we wish to emphasize that a map f is being considered relative to a particular set of marked points, we may write (f, A).

(4) We will assume throughout our discussions that $|A| \geq 3$.

It is clear from the definitions that any rational map $f\colon \widehat{\mathbb{C}} \to \widehat{\mathbb{C}}$ is a branched covering map, and so any postcritically-finite rational map is Thurston map. Accordingly, we shall refer to such maps as *rational Thurston maps* from now on. The class of rational Thurston maps is quite large and it is not difficult to produce examples.

Example 2.3. Consider the quadratic map $f(z) = (1-2z)^2$. This map has the following dynamical portrait:

$$\tfrac{1}{2} \xrightarrow{2:1} 0 \longrightarrow 1 \circlearrowleft \qquad\qquad \infty \circlearrowleft 2:1$$

The critical set is $C_f = \{1/2, \infty\}$ and the postcritical set is $P_f = \{0, 1, \infty\}$. Since this is a postcritically-finite rational map, (f, P_f) is a rational Thurston map. We might also consider this map relative to a larger set of marked points. For example, there is a nonpostcritical fixed point at $1/4$, so the set $A = \{0, 1, \infty, 1/4\} \supsetneq P_f$ is a valid marking set and (f, A) is a Thurston map with four marked points.

The dynamics on the postcritical set determine some of the geometry of the underlying Thurston map. This geometry is naturally described by a type of object called an *orbifold*. An orbifold is a space that is locally the quotient of a model space by some group action. For Thurston maps the model space is 2-dimensional and the group actions will be given by cyclic groups. To this end, consider the *ramification function* of a Thurston map $f\colon S^2 \to S^2$, which is the function $\nu_f\colon S^2 \to \mathbb{N} \cup \{\infty\}$ defined by

$$\nu_f(p) = \operatorname{lcm}\{\deg(f^n, x) : x \in S^2, n \in \mathbb{N}, \text{ and } f^n(x) = p\}.$$

This function has the property $\nu_f(p) = 1$ for $p \in S^2 \setminus P_f$ and $\nu_f(p) \geq 2$ for $p \in P_f$. It can be thought of as encoding the orders of the cyclic group actions generating the orbifold. Accordingly, the *orbifold* $\mathcal{O}_f$ associated to a Thurston map f is just the pair (S^2, ν_f). If we label the finitely many points of P_f as $p_1, \ldots, p_k$ in such a way that $2 \leq \nu_f(p_1) \leq \cdots \leq \nu_f(p_k)$, then the k-tuple

$$(\nu_f(p_1), \ldots, \nu_f(p_k))$$

is the *signature* of the orbifold $\mathcal{O}_f$. We will sometimes refer to the orbifold signature of $\mathcal{O}_f$ as the *type* of the Thurston map f.

The orbifold $\mathcal{O}_f$ has Euler characteristic

$$\chi(\mathcal{O}_f) := 2 - \sum_{p \in P_f} \left(1 - \frac{1}{\nu_f(p)}\right).$$

We use the convention $1/\infty := 0$ in the formula above. If $\chi(\mathcal{O}_f) = 0$ we say the orbifold is *parabolic* and if $\chi(\mathcal{O}_f) < 0$ we say the orbifold is *hyperbolic*. The orbifold $\mathcal{O}_f$ of a Thurston map f is always parabolic or hyperbolic (see, e.g., [BM17, Proposition 2.12]).

For more background on orbifolds and ramification functions for Thurston maps, we refer the reader to [BM17, Section 2.5 and Appendix A.10] and [Mil06, Appendix E].

Example 2.4. Consider again the map $f(z) = (1 - 2z)^2$ of Example 2.3. It is not hard to see that $\nu_f(\infty) = \infty$ since this point is a fixed critical point. Meanwhile, $\nu_f(0) = \nu_f(1) = 2$. Thus the signature of the Thurston map (f, P_f) is $(2, 2, \infty)$. The ramification funtion is supported on the postcritical set, so adding additional points to the marked set does not change the orbifold signature. Since

$$\chi(\mathcal{O}_f) = 2 - \left(1 + \frac{1}{2} + \frac{1}{2}\right) = 0,$$

this map has parabolic orbifold.

Example 2.5. Consider a Thurston map with the following combinatorial description: glue two copies of the unit square together along their boundary to obtain a "pillow" as depicted on the right side of Figure 3. This is a topological 2-sphere. Choose the four corners $\{a, b, c, d\}$ as a marking set A. On the left, we have a tiled surface which is also a topological 2-sphere. We define a Thurston map L from the left sphere to the right sphere by scaling a white tile on the left to the front face of the pillow on the right, and then extending to a map defined on the whole sphere by a successive Schwarz reflection process. Thus the gray tiles map to the back face of the pillow while the white tiles map to the front face of the pillow.

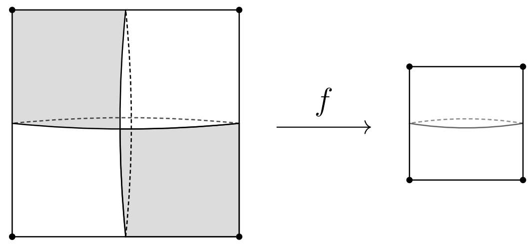


FIGURE 3. A combinatorial depiction of the 2×2 Lattès map.

The map L is a degree 4 Thurston map with $P_L = \{a, b, c, d\}$. The preimage of any of the points $p \in P_L$ consists of two critical points, both of which map to p with degree 2. Thus this map is of type $(2, 2, 2, 2)$. It is an example of what is known as a *Lattès map*. We refer the reader to [BM17, Chapter 3] for more information about Lattès maps.

2.2. Isotopies and combinatorial equivalence

Definition 2.6. Let X and Y be topological spaces. A continuous map $H\colon X \times [0,1] \to Y$ is called an *isotopy* if $H_t := H(\cdot, t)\colon X \to Y$ is a homeomorphism for each $t \in [0,1]$.

The map H_t is called the *time-t* map of the isotopy. Given a distinguished subset $A \subseteq X$, the map H is said to be an *isotopy relative to A* if $H_t|_A = H_0|_A$ for each $t \in [0,1]$. Two homeomorphisms $h_0, h_1\colon X \to Y$ are called *isotopic* (*relative to A*) if there exists an isotopy $H\colon X \times [0,1] \to Y$ (relative to A) with $H_0 = h_0$ and $H_1 = h_1$. If h_0 and h_1 are isotopic relative to A, we will write $h_0 \simeq h_1$ rel. A.

We can now introduce a notion of equivalence for Thurston maps.

Definition 2.7. Two marked Thurston maps (f, A) and (g, A') are said to be *combinatorially equivalent* or *Thurston equivalent* if there exist orientation-preserving homeomorphisms

$h_0, h_1 \colon (S^2, A) \to (S^2, A')$ such that the diagram

$$\begin{CD}
(S^2, A) @>{h_1}>> (S^2, A') \\
@V{f}VV @VV{g}V \\
(S^2, A) @>>{h_0}> (S^2, A')
\end{CD}$$

commutes, and also $h_0 \simeq h_1$ rel. A.

Combinatorial equivalence is, in general, weaker than topological conjugacy (though there are some special cases where the two notions coincide). A basic problem in Thurston theory is how to determine when a given Thurston map (f, A) is combinatorially equivalent to a rational Thurston map. When this occurs, we say that the Thurston map (f, A) is *realized* by a rational map.

The answer to the above problem is the content of Thurston's characterization theorem. We review some important relevant definitions before describing the theorem and the key idea behind its proof.

Definition 2.8. We will say two orientation-preserving homeomorphisms $\varphi_0, \varphi_1 \colon S^2 \to \widehat{\mathbb{C}}$ are *marking equivalent* and write $\varphi_0 \sim \varphi_1$ if there is a Möbius transformation $M \colon \widehat{\mathbb{C}} \to \widehat{\mathbb{C}}$ such that $\varphi_1 \simeq M \circ \varphi_0$ rel. A and also the diagram

$$\begin{array}{ccc}
 & & \widehat{\mathbb{C}} \\
 & \overset{\varphi_0}{\nearrow} & \\
(S^2, A) & & \Big\downarrow M \\
 & \underset{\varphi_1}{\searrow} & \\
 & & \widehat{\mathbb{C}}
\end{array}$$

commutes on A. The equivalence class of a marking φ under the relation $\sim$ is denoted by $[\varphi]$. The *Teichmüller space* of a marked sphere (S^2, A), denoted $\mathcal{T}_A$ or $\operatorname{Teich}(S^2, A)$, is the space of all such marking equivalence classes.

To each Thurston map $f\colon (S^2, A) \to (S^2, A)$ there is an associated pullback map on Teichmüller $\sigma_{f,A}\colon \mathcal{T}_A \to \mathcal{T}_A$. The construction of this map is based on the following proposition:

Proposition 2.9. *Let (f, A) be a Thurston map, and let $\varphi\colon (S^2, A) \to (\widehat{\mathbb{C}}, \varphi(A))$ be a marking homeomorphism. Then there is another marking homeomorphism $\psi\colon (S^2, A) \to (\widehat{\mathbb{C}}, \psi(A))$ and a rational map $R\colon (\widehat{\mathbb{C}}, \psi(A)) \to (\widehat{\mathbb{C}}, \varphi(A))$ so that the following diagram commutes:*

$$\begin{array}{ccc} (S^2, A) & \overset{\psi}{\dashrightarrow} & (\widehat{\mathbb{C}}, \psi(A)) \\ {\scriptstyle f}\downarrow & & \downarrow{\scriptstyle R} \\ (S^2, A) & \underset{\varphi}{\longrightarrow} & (\widehat{\mathbb{C}}, \varphi(A)). \end{array}$$

Moreover, the map ψ is unique up to equivalence with respect to the relation $\sim$ defining $\mathcal{T}_A$.

Definition 2.10. Given a Thurston map $f\colon (S^2, A) \to (S^2, A)$ the associated *Thurston pullback map* $\sigma_{f,A}\colon \mathcal{T}_A \to \mathcal{T}_A$ is the map $[\varphi] \mapsto [\psi]$, where ψ is the map provided by the previous proposition.

Remarks 2.11.

(1) The "$\cdot, A$" portion of the subscript is to emphasize the more general setting of a marked Thurston map (f, A), since σ_f typically denotes the pullback of (f, P_f). Since we work almost exclusively in the marked setting in this paper, we will suppress the subscript and simply write $\sigma_f := \sigma_{f,A}$.

(2) The map σ_f can be defined even when the requirement $\deg(f) \geq 2$ is dropped from the definition of a Thurston map, i.e., when f is simply a homeomorphism with $f(A) = A$.

(3) The map σ_f is holomorphic with respect to the complex structure of Teichmüller space $\mathcal{T}_A$ (see [DH93, Proposition 2.2]).

(4) We have the iteration property $\sigma_{f^n} = \sigma_f^n$ for any $n \in \mathbb{N}$.

The main utility of the Thurston pullback map σ_f is that it reduces the problem of finding of finding a rational representative of a given combinatorial equivalence class to a fixed point problem:

Proposition 2.12. *The Thurston map (f, A) is combinatorially equivalent to a rational Thurston map if and only if $\sigma_f \colon \mathcal{T}_A \to \mathcal{T}_A$ has a fixed point.*

We omit the proof, but it follows from the definitions of combinatorial equivalence and the Thurston pullback map in a relatively straight-forward manner.

An immediate consequence of the above proposition is that every Thurston map (f, A) where $|A| = 3$ is realized by a rational map as $\mathcal{T}_A$ is just a single point in this case.

The proposition also hints at why type $(2, 2, 2, 2)$ maps are usually excluded from our considerations. As a self-map of the hyperbolic space $\mathcal{T}_A$ the map σ_f is always weakly contracting (see, e.g., [Roy71]), but the additional assumption that (f, A) is not of type $(2, 2, 2, 2)$ allows us to further say that $\sigma_f^2 = \sigma_{f^2}$ is strictly contracting on $\mathcal{T}_A$ (see [DH93, Proposition 3.3]). Type $(2, 2, 2, 2)$ maps, on the other hand, admit Thurston pullbacks σ_f that are automorphisms of $\mathcal{T}_A$. The Lattès map presented in Example 2.5 has $\mathcal{T}_A = \mathbb{H}$ and $\sigma_f = \mathrm{id}_{\mathbb{H}}$, for instance. Some of our later considerations do apply to even these cases, but we will generally focus on the setting where some iterate of σ_f (and later, a generic polymorphic map σ) is strictly contracting.

2.3. Pullback relation on curves

In this section we will finally give a precise statement of Thurston's characterization theorem. We will also provide some additional details for our description of the the curve pullback relation that appears in the FCA conjecture.

We will call a curve $\gamma \subseteq S^2$ a *Jordan curve* if it is simple and closed. For a finite set $A \subseteq S^2$, we will call a curve γ a *Jordan curve in the marked sphere* (S^2, A) if $\gamma \subseteq S^2 \setminus A$. A Jordan curve γ in (S^2, A) is *essential* if each of the two connected components of $S^2 \setminus \gamma$ contain at least two points of A. We call a Jordan curve *peripheral* if it is not essential (so γ either encircles a single marked point or is nullhomotopic in $S^2 \setminus A$). Denote by $\mathcal{C}(S^2, A)$ the set of isotopy classes of essential unoriented Jordan curves in $S^2 \setminus A$. A *multicurve* Γ is a nonempty set of distinct elements of $\mathcal{C}(S^2, A)$ represented by pairwise nonintersecting curves.

Consider a Thurston map $f\colon S^2 \to S^2$ with marked set A. As stated in the introduction, there is a pullback relation on Jordan curves in (S^2, A) defined in the following manner: if γ is a Jordan curve in $S^2 \setminus A$, then a connected component $\widetilde{\gamma}$ of $f^{-1}(\gamma)$ is a *pullback* of γ by f. The restriction of $f\colon S^2 \setminus f^{-1}(A) \to S^2 \setminus A$ is an ordinary covering map, so each pullback $\widetilde{\gamma}$ of γ is itself a Jordan curve in (S^2, A). The restriction $f|_{\widetilde{\gamma}} : \widetilde{\gamma} \to \gamma$ is also a covering map, and so has some finite degree, which we denote $\deg(f\colon \widetilde{\gamma} \to \gamma)$.

By lifting isotopies one can see that the isotopy classes of curves in $f^{-1}(\gamma)$ rel. A depends only on the isotopy class of γ rel. A, and not on the specific choice of γ. More precisely, we have the following proposition:

Proposition 2.13. *Let (f, A) be a Thurston map and let α and β be Jordan curves in (S^2, A) with $\alpha \simeq \beta$ rel. A. Then there is a bijection $\widetilde{\alpha} \leftrightarrow \widetilde{\beta}$ between the pullbacks $\widetilde{\alpha}$ of α and the pullbacks $\widetilde{\beta}$ of β under f such that, for all pullbacks corresponding under this bijection, we have $\widetilde{\alpha} \simeq \widetilde{\beta}$ rel. A and $\deg(f\colon \widetilde{\alpha} \to \alpha) = \deg(f\colon \widetilde{\beta} \to \beta)$.*

See [BM17, Lemma 6.9] for a proof.

Thus there is a well-defined *pullback relation* on the set of isotopy classes of Jordan curves in $S^2 \setminus A$. Under this relation, a multicurve pulls back to another multicurve after discarding peripheral pullbacks. We say a multicurve Γ is *f-invariant* if $f^{-1}(\Gamma) \subseteq \Gamma$.

Denote by $\mathbb{Z}[\mathcal{C}(S^2, A)]$ and $\mathbb{R}[\mathcal{C}(S^2, A)]$ the free $\mathbb{Z}$- and $\mathbb{R}$-modules generated by $\mathcal{C}(S^2, A)$, so that an element $w \in \mathbb{Z}[\mathcal{C}(S^2, A)]$ is given by the formal finite linear combination

$$w = \sum_{i=1}^{k} a_i[\gamma_i],$$

where each $a_i \in \mathbb{Z}$. The free submodules generated by a multicurve Γ will be denoted by $\mathbb{Z}^\Gamma$ and $\mathbb{R}^\Gamma$.

The *Thurston linear transformation* $\mathcal{L}_f\colon \mathbb{R}[\mathcal{C}(S^2, A)] \to \mathbb{R}[\mathcal{C}(S^2, A)]$ is defined on basis vectors $[\gamma]$ in the following way: suppose a representative γ has essential pullbacks $\widetilde{\gamma}_1 \dots, \widetilde{\gamma}_n$. Then

$$\mathcal{L}_f([\gamma]) = \sum_{i=1}^{n} \lambda_i[\widetilde{\gamma}_i] \tag{2.1}$$

where

$$\lambda_i = \sum_{f^{-1}(\gamma)\supseteq\delta\simeq\widetilde{\gamma}_i} \frac{1}{\deg(f\colon \delta \to \gamma)};$$

the sum ranges over preimages δ of γ isotopic to $\widetilde{\gamma}_i$. If a multicurve Γ happens to be f-invariant, then $\mathcal{L}_f$ restricts to a linear transformation of the submodule $\mathbb{R}^\Gamma$; we shall denote this restriction by $\mathcal{L}_{f,\Gamma}$.

Definition 2.14. A *Thurston obstruction* of a Thurston map (f, A) is an f-invariant multicurve such that the spectral radius of $\mathcal{L}_{f,\Gamma}$ satisfies $\rho(\mathcal{L}_{f,\Gamma}) \geq 1$.

Theorem 2.15 (Thurston's characterization)**.** *Let (f, A) be a Thurston map that is not of type $(2, 2, 2, 2)$. Then f is combinatorially equivalent to a rational map if and only if f has no Thurston obstruction.*

2.4. Liftables and polymorphicity

Definition 2.16. For a finite set of points $A \subseteq S^2$, let $\operatorname{Homeo}^+(S^2, A)$ be the group of orientation-preserving homeomorphisms $\varphi\colon S^2 \to S^2$ such that $\varphi|_A = \mathrm{id}_A$; we endow the space with the compact-open topology.

Let $\operatorname{Homeo}_0^+(S^2, A)$ denote the path-component of $\operatorname{Homeo}^+(S^2, A)$ that contains the identity map, meaning this set consists precisely of those elements that are isotopic to the identity relative to A. This is easily seen to be a normal subgroup of $\operatorname{Homeo}^+(S^2, A)$; define the *pure mapping class group* of (S^2, A) to be the quotient

$$\operatorname{PMod}(S^2, A) := \operatorname{Homeo}^+(S^2, A)/\operatorname{Homeo}_0^+(S^2, A).$$

Definition 2.17. Let (f, A) be a Thurston map. We say that $\varphi \in \operatorname{Homeo}^+(S^2, A)$ is *liftable* by f if there is a $\widetilde{\varphi} \in \operatorname{Homeo}^+(S^2, A)$ such that

$$\varphi \circ f = f \circ \widetilde{\varphi}. \tag{2.2}$$

In this situation, the following diagram commutes:

$$\begin{array}{ccc} (S^2, A) & \xrightarrow{\widetilde{\varphi}} & (S^2, A) \\ {\scriptstyle f}\downarrow & & \downarrow{\scriptstyle f} \\ (S^2, A) & \xrightarrow[\varphi]{} & (S^2, A). \end{array}$$

If we weaken the definition so that $\varphi, \widetilde{\varphi}\colon S^2 \to S^2$ are just orientiation-preserving homeomorphisms not necessarily fixing A pointwise, but we retain (2.2), then we will say that φ is *weakly liftable*. When we wish to emphasize the difference between the two concepts, we will refer to the former property as *strong liftability.*

Denote the set of weakly liftable and strongly liftable homeomorphisms by $\mathrm{Lift}(f)$ and $\mathrm{Lift}(f, A)$ respectively. By lifting isotopies rel. A by the Thurston map (f, A), we get

Proposition 2.18. *If $\varphi_0 \simeq \varphi_1$ rel. A and φ_0 is weakly or strongly liftable by (f, A), then φ_1 is respectively weakly or strongly liftable.*

In light of the above proposition, let $\mathrm{LMod}(f)$ and $\mathrm{LMod}(f, A)$ be the subgroups of $\mathrm{PMod}(S^2, A)$ consisting of weakly and strongly liftable mapping classes respectively. One of the deepest results in Thurston theory is that the pure mapping class group is virtually strongly liftable, in the sense that there is a subgroup of finite index which is strongly liftable:

Theorem 2.19 (Virtual liftability)**.** *For any Thurston map (f, A), $\mathrm{LMod}(f, A)$ has finite index in $\mathrm{PMod}(S^2, A)$. Furthermore, the assignment $[\varphi] \mapsto [\widetilde{\varphi}]$ is a well-defined homomorphism from $\mathrm{LMod}(f, A)$ to $\mathrm{PMod}(S^2, A)$.*

A version of this result first appeared in [Pil12] and was generalized to the nondynamical setting of admissible covers $f\colon (S^2, A) \to (S^2, B)$, where A and B are potentially different, in [KPS16]. The proof is quite complicated, but since the result is essential for establishing the polymorphicity of the Thurston pullback map, we endeavor to present it in full. Our argument mostly follows that of [KPS16] with some modifications.

Algebraic preliminaries. We first fix some notation. Let G be a group and suppose $G \curvearrowright X$ is a group action on a set X. We define the *orbit* of an element $x \in X$ under G to be $Gx := \{gx : g \in G\}$. We define the *stabilizer* of an element $x \in X$ to be $G_x := \{g \in G : gx = x\}$. Clearly G_x is a subgroup of G. Denote the index of G_x in G by $[G : G_x]$. Then we have

Lemma 2.20 (Orbit-Stabilizer theorem)**.** *Suppose $G \curvearrowright X$ is a group action on a set X. Then, for each $x \in X$, we have*

$$|Gx| = [G : G_x].$$

In particular, if Gx is finite, then G_x is a finite index subgroup of G.

Proof. Fix some $x \in X$, and put $H := G_x$. For $g_1, g_2 \in G$, note that $g_1 x = g_2 x$ if and only if $g_1^{-1} g_2 \in H$, which is true if and only if $g_1 H = g_2 H$. Thus there is a well-defined bijection between points in the orbit Gx and left cosets of H in G. □

Lemma 2.21. *Let $\alpha \colon G \to G'$ be a surjective group homomorphism. If H is a finite-index subgroup of G, then $\alpha(H)$ is a finite-index subgroup of G'. Moreover, $[G' : \alpha(H)] \leq [G : H]$*

Proof. Suppose $[G : H] = n$. Then there are distinct elements $g_1, \ldots, g_n \in G$ such that $G = g_1 H \cup \cdots \cup g_n H$. Let $g_i' = \alpha(g_i)$ for $i = 1, \ldots, n$. Since α is surjective, we have $K = \alpha(G) = g_1' \alpha(H) \cup \cdots \cup g_n' \alpha(H)$. This shows that $\alpha(H)$ has finite index in K. The index inequality follows from the fact that the list $g_1', \ldots, g_n'$ has at most n distinct elements. □

Lemma 2.22. *If G is a finitely-generated group, then it has at most finitely many subgroups of a given index $n \in \mathbb{N}$.*

Proof. Let H be a subgroup of index n, so there are n distinct (left) cosets of H. If X is the set of these cosets, then there is a group action $G \curvearrowright X$ given by left-multiplication: for $g \in G$, we have $g_0 H \mapsto g g_0 H$. If we label the cosets $\{1, \ldots, n\}$ (where H itself is labeled 1), then the group action gives a homomorphism $\alpha \colon G \to S_n$. We also have $H = G_1 = \{g \in G : \alpha(g)(1) = 1\}$. Thus every index n subgroup of G arises from a homomorphism $\alpha \colon G \to S_n$.

Now if $g_1, \ldots, g_m$ is a set of generators for G, then α is completely determined by the images $\alpha(g_1), \ldots, \alpha(g_m) \in S_n$. There are $|S_n| = n!$ choices for each image, and thus there are at most $(n!)^m$ homomorphisms $\alpha\colon G \to S_n$. The previous paragraph shows that every index n subgroup arises from such a homomorphism, so there are finitely many index n subgroups. $\square$

Liftability. Our goal in this section to develop a characterization for when a homeomorphism is weakly liftable. By Proposition 2.18, the weak liftability of $\varphi \in \operatorname{Homeo}^+(S^2, A)$ depends only on the isotopy class of φ relative to A. Fix a point $x_0 \in S^2 \setminus A$. Standard arguments show that each isotopy class in the punctured sphere contains an element that fixes x_0. We will accordingly focus our analysis on this case.

In the following, $\pi_1(X, x_0)$ will denote the fundamental group of a topological space X relative to the basepoint x_0. Recall that any continuous map of pointed topological spaces $f\colon (X, x_0) \to (Y, y_0)$ induces a homomorphism on fundamental groups $f_*\colon \pi_1(X, x_0) \to \pi_1(Y, y_0)$. We refer the reader to [Hat02, Chapter 1] for additional details about fundamental groups.

Lemma 2.23. *Let $u_0, u_1 \in f^{-1}(x_0)$ and put $H = f_*(\pi_1(S^2 \setminus f^{-1}(A), u_0))$ and $H' = f_*(\pi_1(S^2 \setminus f^{-1}(A), u_1))$. Suppose $\varphi \in \operatorname{Homeo}^+(S^2, A)$ and $\varphi(x_0) = x_0$. Then φ is weakly liftable to a homeomorphism $\widetilde{\varphi}\colon S^2 \to S^2$ with $\widetilde{\varphi}(u_0) = u_1$ if and only if $\varphi_*(H) = H'$.*

Proof. If φ is weakly liftable, then we have the following diagram of pointed topological spaces:

$$\begin{array}{ccc}
(S^2 \setminus f^{-1}(A), u_0) & \xrightarrow{\ \widetilde{\varphi}\ } & (S^2 \setminus f^{-1}(A), u_1) \\
\Big\downarrow{\scriptstyle f} & & \Big\downarrow{\scriptstyle f} \\
(S^2 \setminus A, x_0) & \xrightarrow[\ \varphi\ ]{} & (S^2 \setminus A, x_0).
\end{array}$$

Applying the π_1 functor (and keeping in mind that φ and $\widetilde{\varphi}$ are homeomorphisms and thus induce group isomorphisms), we get $\varphi_*(H) = H'$.

Now suppose that $\varphi_*(H) = H'$, and consider the following redrawn diagram:

$$\begin{array}{ccccc} & & & & (S^2 \setminus f^{-1}(A), u_1) \\ & & & \overset{\widetilde{\varphi}}{\nearrow} & \downarrow f \\ (S^2 \setminus f^{-1}(A), u_0) & \xrightarrow{f} & (S^2 \setminus A, x_0) & \xrightarrow{\varphi} & (S^2 \setminus A, x_0). \end{array}$$

Since

$$(\varphi \circ f)_*(\pi_1(S^2 \setminus f^{-1}(A), u_0)) = \varphi_*(H) \subseteq H' = f_*(\pi_1(S^2 \setminus f^{-1}(A), u_1))$$

by assumption, the lifting criterion for covering spaces (see, e.g., [Hat02, Proposition 1.33]) implies the existence of a continuous lift $\widetilde{\varphi}$ making the above diagram commute. Since $\varphi_*^{-1}(H') = (\varphi^{-1})_*(H') \subseteq H$, we can also apply the lifting criterion to $\psi = \varphi^{-1}$ as well; call the continuous lift so-obtained $\widetilde{\psi}$.

We claim that $\widetilde{\psi} = \widetilde{\varphi}^{-1}$, which will show that $\widetilde{\varphi}$ is a homeomorphism. To see this, observe that

$$f \circ \widetilde{\psi} \circ \widetilde{\varphi} = \psi \circ f \circ \widetilde{\varphi} = \psi \circ \varphi \circ f = f,$$

so $\widetilde{\psi} \circ \widetilde{\varphi}$ is a deck transformation of f. Also note that $(\widetilde{\psi} \circ \widetilde{\varphi})(u_0) = u_0$. Since the only deck transformation that fixes a point is the identity, we conclude $\widetilde{\psi} \circ \widetilde{\varphi} = \mathrm{id}_{S^2 \setminus f^{-1}(A)}$. A similar argument works for the other composition order, so the claim is proven.

We have shown that there exists a lift homeomorphism $\widetilde{\varphi}$ of the punctured sphere $S^2 \setminus f^{-1}(A)$ to itself. Taking the end compactification (also known as the Freudenthal compactification; see, e.g., [AN93, Section VI.3] for details) uniquely extends this map to a homeomorphism of the whole sphere. This completes the proof. □

We can say even more about the subgroups H and H' considered in the previous lemma. Since they are the images of fundamental groups based at points in the same fiber of f, they will in fact be conjugate to each other in $\pi_1(S^2 \setminus A, x_0)$ (see, e.g., [Hat02, Theorem 1.38]). Thus the lemma might be rephrased as follows:

Lemma 2.24. *Suppose* $\varphi \in \mathrm{Homeo}^+(S^2, A)$ *with* $\varphi(x_0) = x_0$*. Let* $u_0 \in f^{-1}(x_0)$ *and* $H = f_*(\pi_1(S^2 \setminus f^{-1}(A), u_0))$*. Then* φ *is weakly liftable if and only if* $\varphi(H) = gHg^{-1}$ *for some* $g \in \pi_1(S^2 \setminus A, x_0)$*.*

Virtuality. To prove the finite-index claim of Theorem 2.19, we will need the next two lemmas.

Lemma 2.25. $\mathrm{LMod}(f, A)$ *has finite index in* $\mathrm{LMod}(f)$*.*

Proof. The main issue to contend with is that a given weak liftable φ might have multiple lifts, each fixing $f^{-1}(A)$ setwise. Following [KPS16], let $\mathcal{H}_A := \mathrm{Homeo}^+(S^2, A)$ and let $\mathcal{H}_{\{f^{-1}(A)\}}$ denote the group of homeomorphisms that fix $f^{-1}(A)$ setwise. Let $\mathcal{Q}$ be the subgroup of $\mathcal{H}_A \times \mathcal{H}_{\{f^{-1}(A)\}}$ consisting of ordered pairs $(\varphi, \widetilde{\varphi})$ satisfying (2.2) where φ is weakly liftable. This group acts on $f^{-1}(A)$ since $\widetilde{\varphi}$ induces a permutation on this set. By the orbit-stabilizer theorem, for each $a \in A \subseteq f^{-1}(A)$, the subgroup $\mathcal{Q}_a$ consisting of pairs $(\varphi, \widetilde{\varphi})$ with $\widetilde{\varphi}(a) = a$ has finite index in $\mathcal{Q}$. Since the intersection of finitely many finite-index subgroups also has finite index, we see that

$$\mathcal{Q}_A = \bigcup_{a \in A} \mathcal{Q}_a$$

has finite index in $\mathcal{Q}$ as well. Moreover, we have the following commutative diagram:

$$\begin{array}{ccc}
\mathcal{Q}_A & \hookrightarrow & \mathcal{Q} \\
\downarrow & & \downarrow \\
\mathrm{Lift}(f, A) & \hookrightarrow & \mathrm{Lift}(f) \\
\downarrow & & \downarrow \\
\mathrm{LMod}(f, A) & \longrightarrow & \mathrm{LMod}(f).
\end{array}$$

The first row of vertical arrows are the surjective homomorphisms given by projecting to the first coordinate, i.e., $(\varphi, \widetilde{\varphi}) \mapsto \varphi$. The second row of vertical arrows are the surjective homomorphisms given by taking the isotopy class relative to A, i.e., $\varphi \mapsto [\varphi]$. The lemma now follows from Lemma 2.21 and the fact $[\mathcal{Q} : \mathcal{Q}_A] < \infty$. □

Lemma 2.26. $\mathrm{LMod}(f)$ *has finite index in* $\mathrm{PMod}(S^2, A)$.

Proof. Let $d = \deg f \geq 2$. Consider the set of conjugacy classes of index d subgroups in $\pi_1(S^2 \setminus A, x_0)$:

$$X := \Big\{ \mathrm{conj}(H) : [\pi_1(S^2 \setminus A, x_0) : H] = d \Big\}.$$

The fundamental group $\pi_1(S^2 \setminus A, x_0)$ is finitely generated, so it has a finite number of index d subgroups by Lemma 2.22, and hence a finite number of such conjugacy classes.

Let $G := \mathrm{PMod}(S^2, A)$. We can define a group action $G \curvearrowright X$ by $\mathrm{conj}(H) \mapsto \mathrm{conj}(\varphi_*(H))$ where φ is a representative of $[\varphi] \in G$ with $\varphi(x_0) = x_0$. This action is well-defined because isotopic maps both fixing x_0 will induce the same map on fundamental groups.

Now for $u_0 \in f^{-1}(x_0)$, the image $f_*(\pi_1(S^2 \setminus f^{-1}(A), u_0))$ is an index d subgroup of $\pi_1(S^2 \setminus A, x_0)$, and so determines a conjugacy class $\xi \in X$. By Lemma 2.24 we see that $[\varphi] \in \mathrm{LMod}(f)$ if and only if $[\varphi]$ stabilizes ξ under the group action defined above. Since X is a finite set, $G_\xi = \mathrm{LMod}(f)$ has finite index in G by the orbit-stabilizer theorem. This completes the proof. $\square$

Proof of Theorem 2.19. That $\mathrm{LMod}(f, A)$ has finite index in $\mathrm{PMod}(S^2, A)$ immediately follows from the previous two lemmas. All that remains to show is that the assignment $[\varphi] \mapsto [\widetilde{\varphi}]$ is a well-defined homomorphism from $\mathrm{LMod}(f, A)$ to $\mathrm{PMod}(S^2, A)$, which we will do in two steps.

First we claim given a representative φ of an element in $\mathrm{LMod}(f, A)$, there is a unique lift $\widetilde{\varphi}$ that fixes A pointwise. This statement is equivalent to the following claim: if $f = f \circ \widetilde{\varphi}$ with $\widetilde{\varphi}|_A = \mathrm{id}_A$, then $\widetilde{\varphi} = \mathrm{id}_{S^2}$. There are several ways to show this, but perhaps the simplest is to equip the spheres with complex structures such that f is holomorphic, which we may do by Proposition 2.9. Then $\widetilde{\varphi}$ will be topologically conjugate to a Möbius transformation which also fixes the points of A. Since $|A| \geq 3$, this Möbius transformation will be the identity, and thus $\widetilde{\varphi}$ must be the identity as well. This proves the claim.

Finally, we must show that given two strong liftables with $\varphi_0 \simeq \varphi_1$ rel. A, the unique lifts fixing A described above have $\widetilde{\varphi}_0 \simeq \widetilde{\varphi}_1$ rel. A. This follows by lifting the isotopy by the Thurston map f, so we are done. □

The homomorphism described by Theorem 2.19 is traditionally called the *virtual endomorphism* for the Thurston map (f, A), and it is denoted by $\varphi_f \colon H_f \to G$. Here, $H_f = \mathrm{LMod}(f, A)$ and $G = \mathrm{PMod}(S^2, A)$. We remark that this homomorphism φ_f should not be confused with the liftable homeomorphisms just discussed.

Evaluation. Determining which homeomorphisms are liftables, and if so, what their lifts are, can be computationally difficult. There is one class of homeomorphisms for which this problem is easy: Dehn twists about essential curves in (S^2, A). We refer the reader to [FM12, Chapter 3] for general background on Dehn twists.

We denote the (left) Dehn twist about such a curve γ by T_γ. It is easy to see that if Jordan curves $\gamma_0, \gamma_1 \subseteq S^2 \setminus A$ are isotopic relative to A, then T_{γ_0} and T_{γ_1} belong to the same isotopy class in $\mathrm{PMod}(S^2, A)$. In what follows, we will simplify notation by identifying T_γ with its mapping class.

Let $\overline{\mathcal{C}}(S^2, A)$ denote the union of $\mathcal{C}(S^2, A)$ together with all of the peripheral isotopy classes in (S^2, A). We can extend the definition of the Thurston linear transformation $\mathcal{L}_f$ to a transformation $\overline{\mathcal{L}_f} \colon \mathbb{R}[\mathcal{C}(S^2, A)] \to \mathbb{R}[\overline{\mathcal{C}}(S^2, A)]$ by including peripheral pullbacks in the defining equation (2.1).

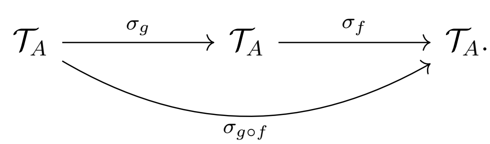
Theorem 2.27 ([Pil12, Corollary 3.4]). *Let $[\gamma] \in \mathcal{C}(S^2, A)$ and suppose a representative γ has pullbacks $\widetilde{\gamma}_1 \ldots, \widetilde{\gamma}_n$, some of which may be peripheral. Then $T_\gamma^a \in \mathrm{LMod}(f, A)$ for $a \in \mathbb{Z}$ if and only if $\overline{\mathcal{L}_f}(a[\gamma]) \in \mathbb{Z}[\overline{\mathcal{C}}(S^2, A)]$. In this case, suppose $\overline{\mathcal{L}_f}(a[\gamma]) = b_1[\widetilde{\gamma}_1] + \cdots + b_n[\widetilde{\gamma}_n]$ for $b_1, \ldots, b_n \in \mathbb{Z}$. Then*

$$\varphi_f(T_\gamma^a) = \prod_{k=1}^{n} T_{\widetilde{\gamma}_k}^{b_k}.$$

If a particular $\widetilde{\gamma}_i$ happens to be peripheral, then the corresponding twist $T_{\widetilde{\gamma}_i}$ is isotopic to the identity map rel. A. We remark that our formulation of the above theorem is a slight correction to the original in [Pil12], where the subtle issue of peripheral pullbacks was not considered.

Polymorphicity. We will now pass the results of the previous section through to the Thurston pullback. We first state two simple propositions, both of which directly follow from our construction of the the Thurston pullback map in Section 2.2. Call a map $f\colon (S^2, A) \to (S^2, A)$ *admissible* if it is either a Thurston map or an element of $\mathrm{Homeo}^+(S^2, A)$.

Proposition 2.28. *Suppose $f\colon (S^2, A) \to (S^2, A)$ and $g\colon (S^2, A) \to (S^2, A)$ are admissible. Then $g \circ f\colon (S^2, A) \to (S^2, A)$ is admissible, and $\sigma_{g\circ f} = \sigma_f \circ \sigma_g$. In other words, the following diagram commutes:*



Proposition 2.29. *If $[\varphi] \in \mathrm{PMod}(S^2, A)$, then $\sigma_\varphi\colon \mathcal{T}_A \to \mathcal{T}_A$ is an automorphism.*

Thus consider the subgroup $G \leq \mathrm{Aut}(\mathcal{T}_A)$ determined by the image of $\mathrm{PMod}(S^2, A)$ under the antihomomorphism $[\varphi] \mapsto \sigma_\varphi$. Let $H \leq G$ be the finite-index subgroup which corresponds to the image of $\mathrm{LMod}(f, A)$. Then Proposition 2.28 and Theorem 2.19 together imply

Theorem 2.30 (Polymorphicity of Thurston pullback)**.** *For any Thurston map (f, A), there is some subgroup G of $\mathrm{Aut}(\mathcal{T}_A)$, a finite-index subgroup $H \leq G$, and homomorphism $\varphi_f\colon H \to G$ such that*

$$\sigma_f \circ g = \varphi_f(g) \circ \sigma_f$$

for all $g \in H$.

2.5. Specialization to four marked points

When $|A| = 4$, multicurves have at most one essential curve class. Furthermore, these curve classes are in bijection with $\widehat{\mathbb{Q}}$. There are multiple constructions of this bijection; see, for example, [BHI21, Appendix] or [FM12, Proposition 2.6]. As we did in the introduction, we can describe the pullback relation on curves as a function $\mu_f\colon \widehat{\mathbb{Q}} \cup \{o\} \to \widehat{\mathbb{Q}} \cup \{o\}$ where o represents any of the peripheral curve classes.

The Thurston linear transformation reduces to a matrix with a single value, which we call the *Thurston multiplier*. In particular, we have

Definition 2.31. For a Thurston map (f, A) with $|A| = 4$, let γ be an essential Jordan curve of (S^2, A). Let $\widetilde{\gamma}_1, \ldots, \widetilde{\gamma}_n$ be the set of essential pullbacks of γ under f. The *Thurston multiplier* of γ is defined to be

$$\lambda_f(\gamma) := \sum_{j=1}^{n} \frac{1}{\deg(f\colon \widetilde{\gamma}_j \to \gamma)}.$$

If all pullbacks of γ are peripheral, then $\lambda_f(\gamma) = 0$.

Remarks 2.32.

(1) By Proposition 2.13, the multiplier may be regarded as being defined on isotopy classes.

(2) A simple but powerful fact is that a given Thurston map (f, A) has finitely many possible multiplier values.

Theorem 2.33 (Thurston's criterion for four points)**.** *Let (f, A) be a Thurston map with $|A| = 4$ and suppose f is not of type $(2,2,2,2)$. Then f is combinatorially equivalent to a rational map if and only if f has no Thurston obstruction, i.e., there is no f-invariant essential Jordan curve γ with $\lambda_f(\gamma) \geq 1$.*

As described in Selinger's thesis [Sel12], σ_f extends to the Weil–Petersson completion of Teichmüller space $\overline{\mathcal{T}}_A$, which for $|A| = 4$ is $\overline{\mathcal{T}}_A = \mathbb{H} \cup \widehat{\mathbb{Q}}$. The behavior of σ_f on $\widehat{\mathbb{Q}}$ encodes the pullback action on essential Jordan curves in (S^2, A). In particular, if $\mu_f(s) = s'$ for $s, s' \in \widehat{\mathbb{Q}}$, then $\sigma_f(r) = r'$ where $r = -1/s$ and $r' = -1/s'$. If $\mu_f(s) = o$, then $\sigma(r) \in \mathbb{H}$. For a detailed discussion of these facts, see [CFPP12, Section 6].

Since we will only every work on the level cusps rather than curves, we shall use the notation $\lambda_f(r)$ to denote the multiplier for the curve γ with slope class $s = -1/r$ throughout the remainder of this paper.

As for the pure mapping class group, there is a natural identification of $\mathrm{PMod}(S^2, A)$ with $\overline{\Gamma}(2)$, the projectivized level-2 congruence subgroup of $\mathrm{PSL}(2, \mathbb{Z})$. This identification is given by the correspondence

$$T_0 \mapsto \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix} \qquad \text{and} \qquad T_\infty \mapsto \begin{bmatrix} 1 & 0 \\ -2 & 1 \end{bmatrix}$$

where T_0 and T_∞ are the Dehn twists which fix the curves of slope 0 and ∞ respectively. We can also specialize the result of Theorem 2.27 to the case where $|A| = 4$ to get:

Theorem 2.34. *For a Thurston map (f, A) with $|A| = 4$, suppose a curve with slope class s pulls back to a curve with slope class s'. If $\lambda_f(s) = b/a$ for positive integers a, b, then there is some positive integer n such that the Dehn twist T_s^{na} is liftable and $\varphi_f(T_s^{na}) = T_{s'}^{nb}$.*

If the curve of slope class s has no peripheral pullbacks, or the peripheral pullbacks all have degree dividing a, then we can take $n = 1$ in the above theorem. Otherwise it is necessary to pass to some multiple in order to ensure that $\overline{\mathcal{L}_f}(na[s]) \in \mathbb{Z}[\overline{\mathcal{C}}(S^2, A)]$ in Theorem 2.27.

CHAPTER 3

Background for polymorphic maps

In this chapter we collect all the background material needed for our systematic study of polymorphic maps.

In Section 3.1 we define Fuchsian groups and give an overview of their basic properties.

In Section 3.2 we describe the structure of the limit set of a finite coarea Fuchsian group. The behavior of a φ-polymorphic map near the boundary depends largely only the limit set structure of the Fuchsian groups G and G'.

In Section 3.3 we define the modular group and discuss some of its basic properties. In particular, we define notion of cusp width, which is essential for our computation of angular derivatives of modularly polymorphic maps.

In Section 3.4 we review the classical theory of complex functions near a boundary. In particular, we define the notions of nontangential limits and angular derivatives. We also present the famous Julia–Wolff–Carathéodory theorem.

3.1. Fuchsian groups

By $\mathrm{Aut}(\mathbb{H})$ we shall denote the group of conformal automorphisms of the upper half-plane, which consists of orientation-preserving isometries with respect to the hyperbolic metric of $\mathbb{H}$. All such maps are Möbius transformations with real coefficients, i.e., maps g given by

$$g(\tau) = \frac{a\tau + b}{c\tau + d}$$

where $a, b, c, d \in \mathbb{R}$ and $ad - bc \neq 0$. Möbius transformations are unchanged by scaling the coefficients by a nonzero parameter, so we will normalize so that $ad - bc = 1$. This yields the standard identification of $\mathrm{Aut}(\mathbb{H})$ with the matrix group $\mathrm{PSL}(2, \mathbb{R})$:

$$\mathrm{PSL}(2,\mathbb{R}) = \left\{ \tau \mapsto \frac{a\tau+b}{c\tau+d} : \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in \mathrm{SL}(2,\mathbb{R}) \right\}. \tag{3.1}$$

Note that $\mathrm{PSL}(2,\mathbb{R}) = \mathrm{SL}(2,\mathbb{R})/\{\pm I\}$ where I is the identity matrix, so elements of $\mathrm{PSL}(2,\mathbb{R})$ are uniquely represented by $\mathrm{SL}(2,\mathbb{R})$ matrices up to sign.

The *absolute trace* of $g \in \mathrm{Aut}(\mathbb{H})$ is defined by $|\operatorname{tr}(g)| = |a+d|$, where

$$g(\tau) = \frac{a\tau+b}{c\tau+d} \quad \text{for} \quad \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in \mathrm{SL}(2,\mathbb{R}).$$

The absolute trace detects the fixed point structure of an automorphism of $\mathbb{H}$. Accordingly, there is a classification of elements of $g \in \mathrm{Aut}(\mathbb{H}) \setminus \{\mathrm{id}_{\mathbb{H}}\}$ into three types:

(1) g is *elliptic* if $|\operatorname{tr}(g)| < 2$, or equivalently, if g fixes a single point in $\mathbb{H}$,

(2) g is *parabolic* if $|\operatorname{tr}(g)| = 2$, or equivalently, if g fixes a single point on $\widehat{\mathbb{R}}$,

(3) g is *hyperbolic* if $|\operatorname{tr}(g)| > 2$, or equivalently, if g fixes two points on $\widehat{\mathbb{R}}$.

In what follows, it will sometimes be convenient to use the disk model $\mathbb{D}$ of hyperbolic plane. Nonidentity elements of $\mathrm{Aut}(\mathbb{D})$ fall into the same three categories above, distinguished by the conditions on the fixed points (i.e., having one fixed point in $\mathbb{D}$, one fixed point on $\partial\mathbb{D}$, or two fixed points on $\partial\mathbb{D}$).

Definition 3.1. A *Fuchsian group* G is a discrete subgroup of $\mathrm{Aut}(\mathbb{H})$.

Many authors alternatively define Fuchsian groups as subgroups of $\mathrm{Aut}(\mathbb{H})$ that act *properly discontinuously* in the following sense: given any compact set $K \subseteq \mathbb{H}$, there are only finitely many maps $g \in G$ for which $g(K) \cap K \neq \varnothing$. These two definitions are equivalent.

The quotient surface $\mathbb{H}/G$ is a hyperbolic orbifold with a metric induced by the standard Poincaré metric on $\mathbb{H}$. The *genus* g of a Fuchsian group is the genus of the underlying topological surface of $\mathbb{H}/G$.

A Fuchsian group G is said to be *cocompact* if $\mathbb{H}/G$ is compact, and it is *finite coarea* if $\mathbb{H}/G$ has finite area. If G is finite coarea, then let m be the number of conjugacy classes of maximal parabolic or maximal elliptic cyclic subgroups of G. Each maximal elliptic cyclic subgroup corresponds to a cone point, while each maximal parabolic cyclic subgroup corresponds to a puncture. Since a finite area orbifold has finitely many cone points and punctures, m is indeed a finite number. Furthermore, if G is finite coarea and has no elliptic elements then $\mathbb{H}/G$ is a compact Riemann surface with m punctures. See, e.g., [Kat92] or [Bea95] for more details.

A useful fact, which we shall employ several times in later chapters, is that every finitely generated Fuchsian group is virtually torsion-free:

Lemma 3.2 (Selberg's lemma)**.** *A finitely generated Fuchsian group has a finite-index subgroup that is torsion-free, i.e., has no elliptic elements.*

A modern proof of Selberg's lemma can be found in [Alp87].

Using this lemma we can obtain as a corollary the following:

Proposition 3.3. *Given any homomorphism $\varphi\colon G \to G'$ where G, G' are finitely generated Fuchsian groups, there are torsion-free finite-index subgroups $H \leq G$ and $H' \leq G'$ so that $\varphi(H) \subseteq H'$.*

Proof. Use Selberg's lemma to obtain a torsion-free finite-index subgroup H' of G'. Then $\varphi^{-1}(H')$ has finite index in G. Apply Selberg's lemma again to obtain a finite-index torsion-free subgroup $H \leq \varphi^{-1}(H')$. This subgroup H also has finite index in G, and it has the property $\varphi(H) \subseteq H'$ by construction. □

3.2. Limit sets of Fuchsian groups

Let G be a nonelementary Fuchsian group, that is, one which does not admit a finite orbit in $\overline{\mathbb{H}} := \mathbb{H} \cup \widehat{\mathbb{R}}$. The *limit set of* G, denoted by $L(G)$, is the set of limit points of an orbit $G\tau$ for some $\tau \in \mathbb{H}$:

$$L(G) = \{x \in \overline{\mathbb{H}} : \text{there exist } g_n \in G \text{ such that } g_n(\tau) \to x \text{ as } n \to \infty\}.$$

This set does not depend on the choice of $z \in \mathbb{H}$. Furthermore, $L(G)$ is closed in $\overline{\mathbb{H}}$, is G-invariant, and satisfies $L(G) \subseteq \widehat{\mathbb{R}}$.

We say that a Fuchsian group G is *of the first kind* if $L(G) = \widehat{\mathbb{R}}$. It is a basic fact that any finite coarea Fuchsian group is of the first kind.

We will divide $L(G)$ into points of two types:

Definition 3.4. A point $x \in L(G)$ is called a *cusp* of G if there exists a parabolic element $g \in G$ such that g fixes x. We denote the set of all cusps of G by $\mathrm{cusp}(G)$.

Definition 3.5. A point $x \in L(G)$ is called a *conical limit point* if there exists a Stolz angle S and a sequence of elements $g_n \in G$ such that $g_n(i) \to x$ in S. We denote the set of all conical limit points of G by $\mathrm{con}(G)$.

A *Stolz angle* within a particular domain with smooth boundary is an angular sector with vertex on the boundary and which, in a neighborhood of the vertex, intersects the boundary only at the vertex. For the upper half-plane $\mathbb{H} = \{x + iy : y > 0\}$, a Stolz angle based at $x_0 \in \mathbb{R}$ is given by the region $y \geq m|x - x_0|$ where $m > 0$. A Stolz angle with vertex ∞ is just the image of a Stolz angle with vertex 0 under the map $\tau \mapsto -1/\tau$.

The choice of distinguished baspoint i in the definition of conical limit point may be replaced with any $\tau \in \mathbb{H}$ (see [Bea95, Theorem 10.2.1]). In other words, $x \in L(G)$ is a conical point if there is a Stolz angle S and a sequence $g_n \in G$ such that, for each $\tau \in \mathbb{H}$, $g_n(\tau) \to x$ in S.

The set of cusps and conical limit points are disjoint. Moreover, for a finitely generated Fuchsian group G, these are the *only* points of $L(G)$, and this property characterizes finitely generated groups:

Theorem 3.6. *A Fuchsian group G is finitely generated if and only if each point of $L(G)$ is either a cusp or a conical limit point, i.e., $L(G) = \mathrm{cusp}(G) \cup \mathrm{con}(G)$.*

A proof can be found in [Bea95, Theorem 10.2.5].

It will sometimes be convenient in our calculations to replace G with a finite-index subgroup $H \leq G$. Accordingly, we must understand how the structure of the limit set changes when passing to finite subgroups.

Proposition 3.7. *If $H \leq G$ is a finite-index subgroup of a finitely-generated Fuchsian group G, then we have the following:*

(i) $L(H) = L(G)$

(ii) $\text{cusp}(H) = \text{cusp}(G)$

(iii) $\text{con}(H) = \text{con}(G)$.

Proof. In light of Theorem 3.6 it will suffice to establish any two items in the list.

We first prove (i). The inclusion $L(H) \subseteq L(G)$ is automatic, so we need only consider the other direction. We make the following claim: given any $\tau \in \mathbb{H}$ there is some $r > 0$ such that for any $g \in G$ there exists some $h \in H$ such that $d_{\mathbb{H}}(g(\tau), h(\tau)) \leq r$. To see this, let $n = [G : H]$ and consider a decomposition of G into n (right) cosets of H: $G = Hg_1 \cup \cdots \cup Hg_n$. Given any $g \in G$, there is some $i = 1, \ldots, n$ such that $gg_i^{-1} \in H$. Moreover, since G consists of hyperbolic isometries,

$$d_{\mathbb{H}}(g(\tau), gg_i^{-1}(\tau)) = d_{\mathbb{H}}(\tau, g_i^{-1}(\tau))$$

for any $\tau \in \mathbb{H}$. Thus, for fixed $\tau \in \mathbb{H}$, put

$$r := \max_{1 \leq i \leq n} d_{\mathbb{H}}(\tau, g_i^{-1}(\tau)).$$

The claim then follows.

Now let $x \in L(G)$, so there is some $\tau \in \mathbb{H}$ and sequence $g_n \in G$ such that $g_n(\tau) \to x$. The preceding claim implies the existence of an $r > 0$ and a sequence $h_n \in H$ such that $d_{\mathbb{H}}(g_n(\tau), h_n(\tau)) \leq r$ for each $n \in \mathbb{N}$. By the properties of the hyperbolic metric, since $g_n(\tau) \to x \in \widehat{\mathbb{R}}$ as $n \to \infty$, it must be the case that $h_n(\tau) \to x$ as $n \to \infty$ as well. This proves (i).

We now prove (ii). The inclusion $\text{cusp}(H) \subseteq \text{cusp}(G)$ is again automatic. For the other direction, suppose $x \in \widehat{\mathbb{R}}$ is a cusp of G, so there is some parabolic element $g \in G$ such that $g(x) = x$. We claim there is some positive integer N such that $g^N \in H$. This will immediately imply the remaining direction, since then $g^N(x) = x$ for $g^N \in H$.

To prove the claim, again let $n = [G : H]$, and consider the list of cosets

$$\{gH, \ldots, g^n H, g^{n+1} H\}.$$

Since there are only n distinct cosets of H, two cosets in the above list must coincide by the pigeonhole principle. If $g^i H = g^j H$ for $i \neq j$, then we see that $g^{i-j} H = H$, and thus $g^N \in H$ for $N = |i - j|$. This completes the proof. $\square$

3.3. The modular group

The *modular group* is the subgroup of matrices in $\mathrm{PSL}(2, \mathbb{R})$ with matrix representatives belonging to $\mathrm{SL}(2, \mathbb{Z})$. In other words,

$$\mathrm{PSL}(2, \mathbb{Z}) = \left\{ \tau \mapsto \frac{a\tau + b}{c\tau + d} : \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in \mathrm{SL}(2, \mathbb{Z}) \right\}. \tag{3.2}$$

This Fuchsian group has a few special properties that we shall make use of. For one, $\mathrm{PSL}(2, \mathbb{Z}) = \langle S, T \rangle$ where $S(\tau) = -\dfrac{1}{\tau}$ and $T(\tau) = \tau + 1$ for $\tau \in \mathbb{H}$, so it is finitely generated. It also has finite coarea (see [Kat92, Example 3.A]), so it is of the first kind.

Proposition 3.8. *If* $G = \mathrm{PSL}(2, \mathbb{Z})$*, then* $\mathrm{cusp}(G) = \widehat{\mathbb{Q}}$*. Moreover, there is a single cusp class and* $\mathrm{cusp}(G) = G\infty$*.*

Proof. First observe that ∞ is a cusp of G, since $T \in \mathrm{PSL}(2, \mathbb{Z})$ given by $T(\tau) = \tau + 1$ has $T(\infty) = \infty$ and T is parabolic.

Next, let $r \in \mathbb{Q}$ be a finite rational number, so we may write $r = p/q$ for coprime $p, q \in \mathbb{Z}$ with $q > 0$. Since p and q are coprime, there are integers $u, v \in \mathbb{Z}$ such that $up + vq = -1$. Then there is $g_r \in G$ such that $g_r(r) = \infty$ given by

$$g_r(\tau) = \frac{u\tau + v}{q\tau - p}.$$

The element $T_r := g_r^{-1} \circ T \circ g_r$ is a parabolic element which fixes r, so r is also a cusp. Thus $\widehat{\mathbb{Q}} \subseteq \mathrm{cusp}(G)$.

On the other hand, if $x \in \text{cusp}(G)$ then there must be some parabolic element $g \in \text{PSL}(2, \mathbb{Z})$ such that $g(x) = x$. Suppose g has form

$$g(\tau) = \frac{a\tau + b}{c\tau + d}$$

where $ad - bc = 1$ and $a, b, c, d \in \mathbb{Z}$. We get the equation

$$\frac{ax + b}{cx + d} = x.$$

If $c = 0$ then $ad = 1$ and so $a = d = \pm 1$. Thus the equation reduces to the statement $x \pm b = x$. We cannot have $b = 0$ since then g is the identity and not a parabolic element. If $b \neq 0$ though, then we must have $x = \infty \in \widehat{\mathbb{Q}}$.

Next suppose $c \neq 0$. Rearranging the equation and using the quadratic formula gives

$$x = \frac{(a - d) \pm \sqrt{(a + d)^2 - 4}}{2c}.$$

Since g is parabolic we must have $(a + d)^2 = |a + d|^2 = 4$. It follows that $x \in \mathbb{Q}$ in this case. This completes the proof. □

Since $\text{cusp}(G) = \widehat{\mathbb{Q}}$ and G is a finitely-generated Fuchsian group of the first kind, it follows that $\text{con}(G) = \widehat{\mathbb{R}} \setminus \widehat{\mathbb{Q}} = \mathbb{I}$, the set of irrational numbers.

Proposition 3.9. *If $g \in \text{PSL}(2, \mathbb{Z})$ fixes ∞, then $g(\tau) = \tau + b$ for some $b \in \mathbb{Z}$.*

Proof. Suppose g fixes ∞ and has form

$$g(\tau) = \frac{a\tau + b}{c\tau + d}$$

where $ad - bc = 1$ and $a, b, c, d \in \mathbb{Z}$. Since $g(\infty) = a/c = \infty$, we must have $c = 0$ and thus $ad = 1$ by the determinant condition. Thus either $a = d = 1$ or $a = d = -1$. We may suppose without loss of generality that the former is true. Then $g(\tau) = \tau + b$ for $b \in \mathbb{Z}$, as claimed. □

We know by the proof of Proposition 3.7(ii) that if $G = \mathrm{PSL}(2, \mathbb{Z})$ and $H \leq G$ is a finite-index subgroup, then there is a unique minimal positive integer N such that $T^N \in H$ for $T(\tau) = \tau + 1$. In fact, given any $r \in \widehat{\mathbb{Q}}$ we see that there is a unique minimal positive integer $N(r)$ depending only on r such that $T_r^N \in H$, where $T_r = g_r^{-1} \circ T \circ g_r$ is as in the proof of Proposition 3.8. Any parabolic element of H fixing r is of this form. Indeed, if $h \in H$ and $h(r) = r$, then $g_r \circ h \circ g_r^{-1} \in G$ fixes ∞ and hence must be of the form T^b for some integer $b \in \mathbb{Z}$ by Proposition 3.9. Since $N(r)$ was chosen to be minimal we must have $h = T_r^b \in \langle T_r^{N(r)} \rangle$.

We shall call $N(r)$ the *width* of the cusp r in the subgroup H.

3.4. Julia–Wolff–Carathéodory theory

In this section we shall record some classical results about the boundary behavior of holomorphic self-maps of hyperbolic 2-space. See [Aba23, Chapter 2], [Sha93, Chapter 4], or [Bur79, Section VI.4] for modern expositions on this material.

In keeping with the literature, we shall initially present our results in the disk model $\mathbb{D}$ and then describe how to translate the results to the upper half-plane model $\mathbb{H}$ afterwards.

Definition 3.10. Given a holomorphic map $\sigma \colon \mathbb{D} \to \widehat{\mathbb{C}}$, we say σ has *angular* or *nontangential limit* at $\zeta \in \partial\mathbb{D}$ if there is $L \in \widehat{\mathbb{C}}$ so that, given any Stolz angle S with vertex ζ, and any sequence $z_n \to \zeta$ in S, we have $\sigma(z_n) \to L$ as $n \to \infty$.

In this case we write

$$\angle \lim_{z \to \zeta} \sigma(z) = L.$$

We are generally interested in the case where $\sigma \colon \mathbb{D} \to \mathbb{D}$. When such a map σ has an nontangential limit $\eta \in \overline{\mathbb{D}}$ at ζ we shall simply write $\sigma(\zeta) = \eta$. When $\eta \in \partial\mathbb{D}$, the limit

$$\sigma'(\zeta) := \angle \lim_{z \to \zeta} \frac{\sigma(z) - \eta}{z - \zeta}$$

exists (where we allow ∞), and we call this quantity the *angular derivative* of σ at ζ.

Theorem 3.11 (Julia–Wolff–Carathéodory (JWC) theorem)**.** *For a holomorphic function $\sigma \colon \mathbb{D} \to \mathbb{D}$ and $\zeta \in \partial\mathbb{D}$, the following statements are equivalent:*

(i) $\delta := \liminf_{z\to\zeta} \dfrac{1-|\sigma(z)|}{1-|z|}$ *is finite*

(ii) $\angle \lim_{z\to\zeta} \dfrac{\sigma(z)-\eta}{z-\zeta}$ *exists and is finite for some* $\eta \in \partial\mathbb{D}$

(iii) $\angle \lim_{z\to\zeta} \sigma(z) = \eta \in \partial\mathbb{D}$ *and* $\angle \lim_{z\to\zeta} \sigma'(z) \in \mathbb{C}$ *exists.*

Moreover, if any of the previous conditions hold, then

$$\sigma'(\zeta) = \angle \lim_{z\to\tau} \frac{\sigma(z)-\eta}{z-\zeta} = \angle \lim_{z\to\zeta} \sigma'(z) = \overline{\zeta}\eta\delta.$$

In particular, if σ has a fixed point at ζ and satisfies any of the above conditions, then $\sigma'(\zeta) = \delta$ is a finite positive real number.

We record one more classical result for the disk case, which will help us establish the existence of nontangential limits in later proofs. See [Aba23, Section 2.4] for a proof.

Theorem 3.12 (Lindelöf principle)**.** *Let* $\gamma\colon [0,1) \to \mathbb{D}$ *be a continuous curve such that*

$$\lim_{t\to 1^-} \gamma(t) = \zeta \in \partial\mathbb{D}.$$

If $\sigma : \mathbb{D} \to \mathbb{C}$ *is a bounded holomorphic function such that the limit*

$$\lim_{t\to 1^-} \sigma(\gamma(t)) = L$$

exists, then σ *has nontangential limit* L *at* ζ*.*

Let us now state the reformulation of these results for the upper half-plane model $\mathbb{H} = \{\tau \in \mathbb{C} : \operatorname{Im}\tau > 0\}$. We shall do this by means of the *Caley transformation*, which is the Möbius transformation $\kappa\colon \widehat{\mathbb{C}} \to \widehat{\mathbb{C}}$ defined by

$$\kappa(z) := i\frac{1+z}{1-z}$$

for all $z \in \widehat{\mathbb{C}}$. The Cayley transformation conformally maps $\mathbb{D}$ onto $\mathbb{H}$; when we write κ we will typically means its restriction to these sets. Note that a region $U \subseteq \mathbb{D}$ is contained in a Stolz angle in the disk if and only if $\kappa(U)$ is contained in a Stolz angle in $\mathbb{H}$.

Now given $\sigma\colon \mathbb{H} \to \mathbb{H}$, we can define a corresponding function $\sigma_\kappa\colon \mathbb{D} \to \mathbb{D}$ on the disk defined by

$$\sigma_\kappa := \kappa^{-1} \circ \sigma \circ \kappa.$$

Given $x \in \widehat{\mathbb{R}}$, define a corresponding point $\zeta_x := \kappa^{-1}(x)$.

Then the nontangential limit of σ at x exists if and only if the nontangential limit of σ_κ at ζ_x exists. Similarly, we say that σ has finite angular derivative at $x \in \widehat{\mathbb{R}}$ if and only if σ_κ has finite angular derivative at $\zeta_x \in \partial\mathbb{D}$. In view of the JWC theorem for disks, if $\sigma'(x)$ is finite then the corresponding limit, which we denote $\sigma(x)$, must have $\sigma(x) \in \widehat{\mathbb{R}}$. Moreover,

$$\sigma'_\kappa(\zeta_x) = \angle \lim_{\tau \to x} \frac{\dfrac{\sigma(\tau) - i}{\sigma(\tau) + i} - \dfrac{\sigma(x) - i}{\sigma(x) + i}}{\dfrac{\tau - i}{\tau + i} - \dfrac{x - i}{x + i}}.$$

We are now ready to give the precise formulas for the angular derivatives of σ at its boundary. These formulas must be stated in cases due to the distinguished boundary point ∞.

Definition 3.13. Suppose $\sigma\colon \mathbb{H} \to \mathbb{H}$ is holomorphic and has nontangential limit $\sigma(x) \in \widehat{\mathbb{R}}$ at $x \in \mathbb{H}$.

(1) If $x, \sigma(x) \in \mathbb{R}$, then

$$\sigma'(x) := \angle \lim_{\tau \to x} \frac{\sigma(\tau) - \sigma(x)}{\tau - x} = \sigma'_\kappa(\zeta_x) \left(\frac{\sigma(x) + i}{x + i} \right)^2.$$

(2) If $x \in \mathbb{R}$, $\sigma(x) = \infty$, then

$$\sigma'(x) := \angle \lim_{\tau \to x} -\frac{1}{(\tau - x)\sigma(\tau)} = -\sigma'_\kappa(\zeta_x) \left(\frac{1}{x + i} \right)^2.$$

(3) If $x = \infty$, $\sigma(x) \in \mathbb{R}$, then

$$\sigma'(\infty) := \angle \lim_{\tau \to \infty} -\tau(\sigma(\tau) - \sigma(\infty)) = -\sigma'_\kappa(1)(\sigma(\infty) + i)^2.$$

(4) If $x = \sigma(x) = \infty$, then

$$\sigma'(\infty) := \angle \lim_{\tau\to\infty} \frac{\tau}{\sigma(\tau)} = \sigma'_\kappa(1).$$

With this in hand, one can prove the following planar version of the Julia–Wolff–Carathéodory theorem (specialized to the case of $\sigma(\infty) = \infty$):

Theorem 3.14 (Planar JWC theorem)**.** *For a holomorphic function $\sigma\colon \mathbb{H} \to \mathbb{H}$ the following are equivalent:*

(i) $\sup_{\tau\in\mathbb{H}} \dfrac{\operatorname{Im}\tau}{\operatorname{Im}\sigma(\tau)} < \infty$

(ii) $\limsup_{\tau\to\infty} \dfrac{\operatorname{Im}\tau}{\operatorname{Im}\sigma(\tau)} < \infty$

(iii) $\sigma(\infty) = \infty$ *and* $\sigma'(\infty) = \angle \lim_{\sigma\to\infty} \dfrac{\tau}{\sigma(\tau)} < \infty$.

Moreover, if any of the previous conditions hold, then

$$\sigma'(\infty) = \limsup_{\tau\to\infty} \frac{\operatorname{Im}\tau}{\operatorname{Im}\sigma(\tau)} = \sup_{\tau\in\mathbb{H}} \frac{\operatorname{Im}\tau}{\operatorname{Im}\sigma(\tau)}.$$

We remark that for a fixed point at ∞, we obtain the (somewhat displeasing) property

$$\angle \lim_{\tau\to\infty} \sigma'(\tau) = \frac{1}{\sigma'(\infty)}.$$

We conclude this section with some example angular derivative calculations that will be useful for our later work.

Example 3.15. Consider $g \in \operatorname{Aut}(\mathbb{H})$ which transposes ∞ and a finite real number $x \in \mathbb{R}$. Then g is given by the formula

$$g(\tau) = \frac{x\tau - (x^2+1)}{\tau - x}.$$

Since $g(\infty) = x$, we use formula (3) from Definition 3.13 to find $g'(\infty)$. Note that

$$-\tau\left(g(\tau) - x\right) = -\tau\left(\frac{x\tau - (x^2+1)}{\tau - x} - x\right) = -\tau\left(\frac{-1}{\tau - x}\right) = \frac{\tau}{\tau - x}.$$

Taking the limit as $\tau \to \infty$ in a Stolz angle, we find

$$g'(\infty) = \angle \lim_{\tau \to \infty} \frac{\tau}{\tau - x} = 1.$$

Example 3.16. Let us now consider a map $g \in \mathrm{PSL}(2, \mathbb{Z})$ that sends a finite rational $r \in \mathbb{Q}$ to ∞. Write $r = p/q$ for coprime $p, q \in \mathbb{Z}$ with $q \neq 0$. Since p and q are coprime, there are integers $u, v \in \mathbb{Z}$ such that $up + vq = -1$. Then an example of $g \in \mathrm{PSL}(2, \mathbb{Z})$ with the desired property is

$$g(\tau) = \frac{u\tau + v}{q\tau - p}.$$

Since $g(r) = \infty$, we use formula (2) from Definition 3.13 to find $g'(r)$. Note that

$$-\frac{1}{(\tau - r)g(\tau)} = -\frac{1}{\tau - p/q} \cdot \frac{q\tau - p}{u\tau + v} = -\frac{q}{u\tau + v}.$$

Letting $\tau \to p/q$ in a Stolz angle, we find

$$g'(r) = \angle \lim_{\tau \to r} -\frac{q}{u\tau + v} = -\frac{q}{u(p/q) + v} = -\frac{q^2}{up + vq} = q^2.$$

3.5. Horoballs in $\mathbb{H}$

In this section we record some lemmas regarding horoballs in $\mathbb{H}$ for later use. Let $\widehat{\mathbb{R}} := \mathbb{R} \cup \{\infty\}$ denote the extended real line.

Definition 3.17. Let $t > 0$ and $x \in \widehat{\mathbb{R}}$. A *horoball* based at x with *hororadius* t is

$$H_t(x) := \left\{ \tau \in \mathbb{H} : \frac{\operatorname{Im} \tau}{|\tau - x|^2} > \frac{1}{t} \right\}$$

for $x \in \mathbb{R}$ and

$$H_t(\infty) := \left\{ \tau \in \mathbb{H} : \operatorname{Im} \tau > \frac{1}{t} \right\}$$

for $x = \infty$.

A horoball $H_t(x)$ based at a finite point $x \in \mathbb{R}$ is simply a ball in $\mathbb{H}$ that is tangent to the real line at x; a simple calculation shows the ball has Euclidean diameter t. The horoball $H_t(\infty)$, on the other hand, is a half-plane.

If H is a horoball, we call its boundary ∂H a *horocycle*. Thus a horocycle based at $x \in \mathbb{R}$ is a circle in $\mathbb{H}$ that is tangent to $\mathbb{R}$ at x, while a horocycle based at ∞ is a horizontal line.

From here we can use the Julia–Wolff–Carathéodory Theorem 3.14 to obtain the following reformulation of the Julia Lemma in terms of horoballs:

Corollary 3.18. *For a holomorphic function $\sigma\colon \mathbb{H} \to \mathbb{H}$ with $\sigma(x) = x'$ for $x, x' \in \widehat{\mathbb{R}}$, suppose $\delta := \sigma'(x)$ satisfies $\delta \in (0, \infty)$. Then $\sigma(H_t(x)) \subseteq H_{\delta t}(x')$.*

Proof. First suppose that $\sigma(\infty) = \infty$. In this case, $\delta = \sigma'(\infty) < \infty$, so Theorem 3.14 applies and

$$\delta = \sigma'(\infty) = \sup_{\tau \in \mathbb{H}} \frac{\operatorname{Im} \tau}{\operatorname{Im} \sigma(\tau)}.$$

In particular, $\operatorname{Im} \sigma(\tau) \geq (1/\delta) \operatorname{Im} \tau$ for all $\tau \in \mathbb{H}$. Now if $\tau \in H_t(\infty)$, then $\operatorname{Im} \tau > 1/t$, so it follows that $\operatorname{Im} \sigma(\tau) > 1/(\delta t)$. Hence $\sigma(\tau) \in H_{\delta t}(\infty)$ and the claim is proven.

For the general case, consider maps $g_1, g_2 \in \operatorname{Aut}(\mathbb{H})$ such that g_1 transposes x and ∞ and g_2 transposes x' and ∞. If $x = \infty$ then we take $g_1 = \mathrm{id}$, and likewise for x' and g_2. Otherwise, we use the map considered in Example 3.15:

$$g_1(\tau) = \frac{x\tau - (x^2 + 1)}{\tau - x} \quad \text{and} \quad g_2(\tau) = \frac{x'\tau - ((x')^2 + 1)}{\tau - x'}.$$

Note that g_1 and g_2 are both their own inverses. A simple calculation shows that

$$\operatorname{Im} g_1(\tau) = \frac{\operatorname{Im} \tau}{|\tau - x|^2}.$$

Thus $g_1(\tau) \in H_t(\infty)$ if and only if $\tau \in H_t(x)$. An analogous statement holds for g_2.

The map $g_2 \circ \sigma \circ g_1 \colon \mathbb{H} \to \mathbb{H}$ now fixes ∞, so the previous case applies with dilation factor $(g_2 \circ \sigma \circ g_1)'(\infty)$. The usual chain rule holds for angular derivatives at contact points (see [Aba23, Proposition 2.4.7]). Since $g_1'(\infty) = 1$ and $g_2'(x') = 1$, we see that $(g_2 \circ \sigma \circ g_1)'(\infty) = \sigma'(x) = \delta$. Thus

$$(g_2 \circ \sigma \circ g_1)(H_t(\infty)) \subseteq H_{\delta t}(\infty).$$

We have $g_1(H_t(\infty)) = H_t(x)$ and $g_2^{-1}(H_{\delta t}(\infty)) = g_2(H_{\delta t}(\infty)) = H_{\delta t}(x')$, so the above line is equivalent to the statement

$$\sigma(H_t(x)) \subseteq H_{\delta t}(x').$$

This completes the proof. □

We will now define a special family of normalized horoballs based at rational cusps that are paticularly well-behaved under the action by modular group elements. Suppose $r = \frac{p}{q} \in \mathbb{Q}$ is written in lowest terms. Then put

$$B_t(r) := H_{t/q^2}(r).$$

For $r = \infty = 1/0$, we define $B_t(\infty) := H_t(\infty)$. We shall sometimes call such a horoball $B_t(r)$ a *modular horoball.* Note the Euclidean radius of $B_t(r)$ for $r \neq \infty$ is given by $R = t/(2q^2)$. Since t no longer represents the true hororadius for these normalized horoballs, we shall instead call t the *hororadius parameter* in this context.

Lemma 3.19. *Let $g \in \mathrm{PSL}(2,\mathbb{Z})$ and suppose $g(r) = r'$ for $r, r' \in \widehat{\mathbb{Q}}$. Then*

$$g(B_t(r)) = B_t(r').$$

Proof. Note that if $r = p/q$ for coprime integers p, q with $q > 0$, then

$$B_t(r) = H_{t/q^2} = \left\{ \tau \in \mathbb{H} : \frac{\operatorname{Im}\tau}{|q\tau - p|^2} > \frac{1}{t} \right\}.$$

We show that if $g \in \mathrm{PSL}(2,\mathbb{Z})$ and $g(r) = r' = p'/q'$ for coprime integers p', q' with $q' > 0$, then

$$\frac{\operatorname{Im}\tau}{|q\tau - p|^2} = \frac{\operatorname{Im} g(\tau)}{|q' g(\tau) - p'|^2}.$$

It suffices to check the formula for the generators T and S of $\mathrm{PSL}(2,\mathbb{Z})$, where $T(\tau) = \tau + 1$ and $S(\tau) = -1/\tau$. In the first case, note $\operatorname{Im} T(\tau) = \operatorname{Im}\tau$ and $p'/q' = p/q + 1$, so after putting $p' = p + q$ and $q' = q$ the result follows for T. Similarly, in the second case we have $\operatorname{Im} S(\tau) = \operatorname{Im}\tau/|\tau|^2$ and $p'/q' = -q/p$. Putting $p' = -q$ and $q' = p$, the stated formula is true in this case as well. This completes the proof. □

An easy corollary of this lemma is the fact that, if $t < 1$, then the collection $\{B_t(r) : r \in \widehat{\mathbb{Q}}\}$ is disjoint. Suppose for contradiction two horoballs from the family intersect. By application of a suitably chosen $g \in \mathrm{PSL}(2, \mathbb{Z})$, the lemma allows us to assume one of the horoballs is $B_t(\infty) = \{\tau \in \mathbb{H} : \mathrm{Im}(\tau) > 1/t\}$. Any other horoball $B_t(r)$ for $r \in \mathbb{R}$ is tangent to the real line and has Euclidean diameter $t/q^2 \leq t < 1$, and so has $\mathrm{Im}(\tau) < 1$ for all $\tau \in B_t(r)$. Since $1 < 1/t$, the two horoballs must be disjoint and the claim follows.

CHAPTER 4

Polymorphic maps

4.1. Definition, cusps, and rigidity

We shall now finally turn our attention to polymorphic maps. Recall:

Definition 4.1. Let G, G' be finite coarea Fuchsian groups, and let $\varphi : G \to G'$ be a homomorphism. We will say a nonconstant holomorphic function $\sigma\colon \mathbb{H} \to \mathbb{H}$ is *φ-polymorphic* if σ satsifies the intertwining relation

$$\sigma \circ g = \varphi(g) \circ \sigma \tag{4.1}$$

for all $g \in G$.

An important property of φ-polymorphic maps is that they admit continuous extensions to $\operatorname{cusp}(G)$:

Proposition 4.2. *Suppose $\sigma\colon \mathbb{H} \to \mathbb{H}$ is φ-polymorphic. Then for any cusp $x \in \widehat{\mathbb{R}}$ of G, the nontangential limit $\sigma(x)$ exists. If $\sigma(x) \in \widehat{\mathbb{R}}$, then $\sigma(x)$ is a cusp of G'.*

This is result is at least implicitly contained in [MP08, Theorem 2]. Since our definition of polymorphic map is more restricted than the one used there, we can give a somewhat simplified proof.

Proof. If $x \in \widehat{\mathbb{R}}$ is a cusp of G, then there is a parabolic $g \in G$ such that $g(x) = x$. Let $\widetilde{g} = \varphi(g)$, so we have the intertwining relation

$$\sigma \circ g = \widetilde{g} \circ \sigma.$$

Our goal will be to show that the nontangential limit lies on the boundary only in the case where $\widetilde{g}$ is parabolic. To do this, we will divide the argument into cases based on the automorphism type of $\widetilde{g}$.

Case 1: $\widetilde{g}$ is the identity. Let $\langle g\rangle \le \mathrm{Aut}(\mathbb{H})$ be the subgroup generated by the parabolic element g. Then we have the conformal isomorphism $\mathbb{H}/\langle g\rangle \cong \mathbb{D}\setminus\{0\}$, and in this identification, $x\in\widehat{\mathbb{R}}$ on the left corresponds to the puncture at the origin on the right. Thus σ descends to an induced map $\overline{\sigma}\colon \mathbb{D}\setminus\{0\}\to\mathbb{H}$, and it is easy to see (say, by the Lindelöf principle) that σ has a nontangential limit at x if and only if $\overline{\sigma}$ extends to be continuous at the origin. This statement is also true if we replace $\overline{\sigma}$ with with $\kappa^{-1}\circ\overline{\sigma}\colon \mathbb{D}\setminus\{0\}\to\mathbb{D}$, where $\kappa^{-1}\colon\mathbb{H}\to\mathbb{D}$ is the inverse Cayley transformation introduced in the previous section. Since $\kappa^{-1}\circ\overline{\sigma}$ is bounded in punctured neighborhoods of the origin, Riemann's theorem on removable singularities guarantees $\kappa^{-1}\circ\overline{\sigma}$, and hence also $\overline{\sigma}$, extends to be continuous at the origin. This proves the existence of $\sigma(x)$ in the sense of nontangential limits. We note that $\sigma(x)\in\widehat{\mathbb{R}}$ cannot happen in this case (where $\widetilde{g}=\mathrm{id}_{\mathbb{H}}$). Indeed, if we supposed that $\kappa^{-1}\circ\overline{\sigma}(0)\in\partial\mathbb{D}$, then the maximum principle would imply $\kappa^{-1}\circ\overline{\sigma}$ is constant on the disk. This could only happen if $\sigma\colon\mathbb{H}\to\mathbb{H}$ itself were constant, thus contradicting our assumption that σ was nonconstant as a polymorphic function.

Case 2: $\widetilde{g}$ is parabolic. In this case, σ descends to a holomorphic map from $\mathbb{H}/\langle g\rangle\cong\mathbb{D}\setminus\{0\}$ to $\mathbb{H}/\langle\widetilde{g}\rangle\cong\mathbb{D}\setminus\{0\}$, so we get an induced map $\overline{\sigma}\colon\mathbb{D}\setminus\{0\}\to\mathbb{D}\setminus\{0\}$. Arguing as before, $\overline{\sigma}$ extends to be continuous at the origin and so $\sigma(x)$ exists in the nontangential sense. The intertwining relation implies $\sigma(x)$ is a fixed point of $\widetilde{g}$, so $\sigma(x)\in\widehat{\mathbb{R}}$ since $\widetilde{g}$ is parabolic.

Case 3: $\widetilde{g}$ is hyperbolic. This case is impossible. Indeed, suppose otherwise. Then $\mathbb{H}/\langle\widetilde{g}\rangle\cong A$ for some open annulus $A=\{w : r_1<|w|<r_2\}$, where in this identification, the two fixed points of $\widetilde{g}$ in $\widehat{\mathbb{R}}$ on the left correspond to the inner and outer boundaries of A on the right. As in the previous cases the existence of $\sigma(x)$ follows from noting the removability of the singularity at the origin in the induced map $\overline{\sigma}\colon\mathbb{D}\setminus\{0\}\to A$. By the maximum principle it cannot be the case that $\overline{\sigma}(0)$ lies on the outer boundary $|w|=r_2$, and similarly by the minimum principle $\overline{\sigma}(0)$ cannot lie on the inner boundary $|w|=r_1$, so $\overline{\sigma}(0)\in A$. But then $\sigma(x)$ cannot be a fixed point of $\widetilde{g}$, since if it were then in the induced

map $\overline{\sigma}(0)$ would lie on one of the two boundaries of the annulus A. Since $\sigma(x)$ must be a fixed point of $\widetilde{g}$ by the intertwining relation, we have derived a contradiction.

Case 4: $\widetilde{g}$ is elliptic. Since G' is assumed to be Fuchsian, $\widetilde{g}$ must be finite-order (if it were infinite order, then G' would not act properly discontinuously on $\mathbb{H}$). Thus there is a positive integer n such that $\widetilde{g}^n$ is the identity. The result will then follow from Case 1 by replacing g with g^n, which is still parabolic with cusp x. □

In the remainder of this section we present the proof of an important rigidity result from [Pom81, Theorem 5].

Theorem 4.3. *Suppose $\varphi\colon G \to G'$ is a homomorphism of Fuchsian groups where G is finite coarea. If $\sigma_1\colon \mathbb{H} \to \mathbb{H}$ and $\sigma_2\colon \mathbb{H} \to \mathbb{H}$ are both φ-polymorphic, then $\sigma_1 = \sigma_2$.*

Proof. By Fatou's theorem (see, e.g., [Zyg02, Theorem 7.2.5]), σ_1 and σ_2 have nontangential limits for almost every $x \in \widehat{\mathbb{R}}$. Almost all of these points are conical limit points of G. Let $x \in \widehat{\mathbb{R}}$ be such a conical limit point. By definition there is a Stolz angle S and a sequence $g_n \in G$ such that, for any $\tau \in \mathbb{H}$, $g_n(\tau) \to x$ in S as $n \to \infty$. Put $\widetilde{g}_n = \varphi(g_n)$. Then

$$\sigma_1(g_n(\tau)) = \widetilde{g}_n(\sigma_1(\tau)) \tag{4.2}$$

for all $\tau \in \mathbb{H}$. We can give a precise description of the pointwise limit function of $\widetilde{g}_n$ by using a theorem due to Piranian and Thron ([PT57, Theorem 1]). After replacing g_n and $\widetilde{g}_n$ with suitable subsequences, there will be some function $\widetilde{g}$ so that $\widetilde{g}_n \to \widetilde{g}$ pointwise on either $\widehat{\mathbb{C}}$ or $\widehat{\mathbb{C}} \setminus \{z_0\}$, the complement of a single point. Moreover, this limit function $\widetilde{g}$ is either a constant or a Möbius transformation.

Taking the limit $n \to \infty$ in equation (4.2), we see that

$$\sigma_1(x) = \widetilde{g}(\sigma_1(\tau))$$

for either all $\tau \in \mathbb{H}$ or all $\tau \in \mathbb{H} \setminus \sigma_1^{-1}(\{z_0\})$ where $\sigma_1^{-1}(\{z_0\})$ is a countable set. In either case $\widetilde{g}$ is a constant and thus $\sigma_1(x)$ is determined solely by the sequence $\widetilde{g}_n$, which itself is determined by the intertwining homomorphism φ. Thus $\sigma_1(x) = \sigma_2(x)$ for almost every $x \in \widehat{\mathbb{R}}$, so $\sigma_1 = \sigma_2$ in $\mathbb{H}$ by Privalov's uniqueness theorem (see, e.g., [Zyg02, pp. 276]). $\square$

We remark that an alternate proof can be given by passing to quotient spaces (after first replacing G and G' with torsion-free finite-index subgroups as in Proposition 3.3). The induced maps on Riemann surfaces are uniquely determined by the intertwining homomorphism φ and are hence equal.

An important consequence of Theorem 4.3 is that we may always replace an intertwining homomorphism $\varphi\colon G \to G'$ with a restriction $\varphi|_H : H \to H'$, where H and H' are finite-index subgroups of G and G' respectively, and no information about the associated polymorphic map $\sigma\colon \mathbb{H} \to \mathbb{H}$ will be lost.

4.2. Modularly polymorphic maps

We now further restrict our setting to φ-polymorphic maps where $\varphi\colon G \to G'$ is a homomorphism between finite-index subgroups of the modular group $\mathrm{PSL}(2, \mathbb{Z})$. Such groups are still finite coarea and so the results of the previous section apply. The main advantage of working in this setting is that there is a natural notion of cusp width (see Section 3.3), which allows us to compute angular derivatives at the cusps of G.

Suppose $r \in \widehat{\mathbb{Q}}$ has cusp width $a \in \mathbb{N}$ within the group G, and that $r' = \sigma(r) \in \widehat{\mathbb{Q}}$ is also a cusp and has width $b \in \mathbb{N}$ in G'. Then the *cusp multiplier* at r is $\lambda(r) = b/a$. If $\sigma(r)$ is not a cusp of G' (i.e., $\sigma(r) \in \mathbb{H}$), then we shall declare $\lambda(r) = 0$. Since G and G' have finitely many cusp classes and hence only finitely many cusp widths, the *spectrum of cusp multipliers* for σ is a finite list:

$$S_\lambda(\sigma) := \{\lambda \in \mathbb{Q}_{\geq 0} : \lambda = \lambda(r) \text{ is a cusp multiplier for } \sigma\}.$$

It will sometimes be convenient to consider the reciprocals of these multipliers instead. We shall define the *cusp dilation factor* at a cusp $r \in \widehat{\mathbb{Q}}$ to be $\delta(r) = 1/\lambda(r)$. Thus there is also a *spectrum of cusp dilation factors* of σ defined by

$$S_\delta(\sigma) = \{\delta \in \widehat{\mathbb{Q}}_{>0} : \delta = \delta(r) \text{ is a cusp dilation factor for } \sigma\}.$$

Since these two sets are related by reciprocal, we have $\max S_\lambda(\sigma) = 1/\min S_\delta(\sigma)$.

We are now ready to begin the calculation of angular derivatives of modularly polymorphic maps. We first establish the following lemma:

Lemma 4.4. *Suppose $\sigma\colon \mathbb{H} \to \mathbb{H}$ is a holomorphic map with $\sigma(\infty) = \infty$ in the sense of nontangential limits. If there are real numbers $a, b \neq 0$ such that*

$$\sigma(\tau + a) = \sigma(\tau) + b$$

for all $\tau \in \mathbb{H}$, then $\sigma'(\infty) = a/b$.

Proof. By repeated applications of the relation $\sigma(\tau + a) = \sigma(\tau) + b$, we obtain

$$\sigma(\tau + na) = \sigma(\tau) + nb$$

for all $n \in \mathbb{Z}$. Putting $\tau = ni$ and dividing both sides by n, one can show that

$$(i + a)\frac{\sigma(ni + na)}{n(i + a)} = i\frac{\sigma(ni)}{ni} + b. \tag{4.3}$$

On the other hand,

$$\delta := \sigma'(\infty) = \angle \lim_{\tau \to \infty} \frac{\tau}{\sigma(\tau)}$$

necessarily exists in the sense that $\delta \in (0, \infty]$. Thus

$$\lim_{n \to \infty} \frac{\sigma(ni + na)}{n(i + a)} \text{ and } \lim_{n \to \infty} \frac{\sigma(ni)}{ni}$$

both exist and are equal to a finite nonnegative real number L, where $L = 1/\delta$. Taking the limit as $n \to \infty$ in equation (4.3) yields

$$(i + a)L = iL + b.$$

Taking real parts, we have $aL = b$ for $a, b \neq 0$, so $L = b/a$ and $\sigma'(\infty) = 1/L = a/b$ is a finite positive real number. This completes the proof. $\square$

Theorem 4.5. *Suppose $\sigma\colon \mathbb{H} \to \mathbb{H}$ is modularly polymorphic. Then we have the following properties:*

(i) *the map σ extends to a continuous map $\sigma\colon \mathbb{H} \cup \widehat{\mathbb{Q}} \to \mathbb{H} \cup \widehat{\mathbb{Q}}$*

(ii) *if $r \in \widehat{\mathbb{Q}}$ is a contact cusp in the sense that $r' = \sigma(r) \in \widehat{\mathbb{Q}}$, then the angular derivative $\sigma'(r)$ is finite. Moreover, if $r = p/q$ for coprime integers p and $q > 0$, and likewise $r' = p'/q'$ for coprime integers p' and $q' > 0$, then*

$$\sigma'(r) = \frac{1}{\lambda(r)} \left(\frac{q}{q'}\right)^2 = \delta(r) \left(\frac{q}{q'}\right)^2,$$

where $\lambda(r)$ is the cusp multiplier and $\delta(r) = 1/\lambda(r)$ is the cusp dilation factor. If r or r' is ∞, then we respectively take q or q' to be 1 in the above formula.

Proof. Since G and G' are both finite-index subgroups of the modular group, $\operatorname{cusp}(G) = \operatorname{cusp}(G') = \widehat{\mathbb{Q}}$ by Propositions 3.8 and 3.7. Part (i) then immediately follows from Proposition 4.2.

The first claim of part (ii), which is that $\sigma'(r)$ is finite when r is a contact cusp, is a consequence of the JWC theorem. For the derivative calculation, recall that every parabolic element of G that fixes r has form T_r^a for some integer $a \neq 0$, where $T_r = g_r^{-1} \circ T \circ g_r$ is as in Section 3.3 (note that we take $g_r = \mathrm{id}$ if $r = \infty$). Since $\varphi(T_r^a)$ is a parabolic element of G' that fixes r', we must have $\varphi(T_r^a) = T_{r'}^b$ for some integer $b \neq 0$. By putting $\widetilde{\sigma} := g_{r'} \circ \sigma \circ g_r^{-1}$ we obtain a map fixing ∞ such that $\widetilde{\sigma} \circ T^a = T^b \circ \widetilde{\sigma}$. Lemma 4.4 thus implies $\widetilde{\sigma}'(\infty) = a/b = 1/\lambda(r)$. On the other hand,

$$\widetilde{\sigma}'(\infty) = g_{r'}'(r')\sigma'(r)(g_r^{-1})'(\infty)$$

by the angular derivative chain rule. Assuming $r, r' \in \mathbb{Q}$, we have $g'_{r'}(r') = (q')^2$ and $g'_r(r) = q^2$ by Example 3.16. Rearranging the above equation then gives

$$\sigma'(r) = \frac{1}{\lambda(r)} \left(\frac{q}{q'}\right)^2.$$

If r or r' is ∞ then we respectively take $g_r = \mathrm{id}$ or $g_{r'} = \mathrm{id}$ in the above calculations. Since the identity has angular derivative 1 at every boundary point, this justifies the last claim of the theorem statement. □

Corollary 4.6. *If $\sigma(r) = r' \in \widehat{\mathbb{Q}}$ has cusp dilation factor $\delta := \delta(r)$, then for any modular horoball we have $\sigma(B_t(r)) \subseteq B_{\delta t}(r')$.*

Proof. This follows from Theorem 4.5 and Proposition 3.18. □

Compare this result to [CFPP12, Section 6], which establishes the same property for the Thurston pullback map σ_f using an argument based on the moduli of curve families.

4.3. Fixed points

In this section we present two results regarding the fixed points of modularly polymorphic maps. Theorem 4.9 characterizes the possible location of the DW point using another rigidity statement (Theorem 4.8). We obtain as a corollary a new proof of Thurston's theorem in the special case of four marked points.

In Corollary 4.12 we give an estimate on the number of fixed cusps for a modularly polymorphic with interior DW point. Appealing again to the polymorphicity of the Thurston pullback σ_f, this shows that a rational Thurston map (f, A) with four marked points has finitely many f-invariant curve classes.

Let us begin by recalling the following classical result:

Theorem 4.7 (Denjoy–Wolff theorem)**.** *If $\sigma\colon \mathbb{H} \to \mathbb{H}$ is holomorphic and not the identity, then it has a unique fixed point $\tau_0 \in \overline{\mathbb{H}}$ with the property $|\sigma'(\tau_0)| \leq 1$.*

In the above theorem statement, the derivative is interpreted in the usual sense when $\tau_0 \in \mathbb{H}$ and in the sense of angular derivatives when $\tau_0 \in \widehat{\mathbb{R}}$.

The unique fixed point $\tau_0 \in \overline{\mathbb{H}}$ is often referred to as the Denjoy–Wolff (DW) point. We would like to characterize the location of the DW point for our polymorphic maps. It turns out that the DW point cannot occur at a conical limit point on the boundary unless the map is an automorphism.

Theorem 4.8. *Let $\varphi\colon G \to G'$ be a homomorphism of finite coarea Fuchsian groups where G admits at least one cusp class. If $\sigma\colon \mathbb{H} \to \mathbb{H}$ is φ-polymorphic and not an automorphism, then its Denjoy–Wolff point is either a cusp of G or in $\mathbb{H}$.*

Note that we do not need σ to be modularly polymorphic here.

The author discovered a proof of this result independently before finding that it had, essentially, already been established within the proof of a similar statement in [Pom81, Theorem 3].

Proof. Suppose for contradiction that the Denjoy–Wolff point of σ is neither in $\mathbb{H}$ nor a cusp of G. By Theorem 3.6 it must be that the Denjoy–Wolff point is a conical limit point. Let $x \in \mathbb{R}$ be this conical limit point. It has the property $\sigma'(x)$ exists and $\sigma'(x) \in (0, 1]$. By definition there is a Stolz angle S and a sequence $g_n \in G$ such that, for any $\tau \in \mathbb{H}$, $g_n(\tau) \to x$ in S as $n \to \infty$. Put $\widetilde{g}_n = \varphi(g_n)$. Then

$$\sigma(g_n(\tau)) = \widetilde{g}_n(\sigma(\tau)).$$

Differentiating this expression gives

$$\sigma'(g_n(\tau))g_n'(\tau) = \widetilde{g}_n'(\sigma(\tau))\sigma'(\tau). \tag{4.4}$$

Suppose that

$$g_n(\tau) = \frac{a_n\tau + b_n}{c_n\tau + d_n} \quad \text{and} \quad \widetilde{g}_n(\tau) = \frac{A_n\tau + B_n}{C_n\tau + D_n}.$$

Then (4.4) can be rewritten as

$$\sigma'(g_n(\tau)) = \left(\frac{c_n\tau + d_n}{C_n\sigma(\tau) + D_n}\right)^2 \sigma'(\tau). \tag{4.5}$$

Put $M_n = |c_n| + |d_n| + |C_n| + |D_n|$, and multiply the right-hand side of the above by M_n^{-2} over itself. The sequences c_n/M_n, d_n/M_n, C_n/M_n and D_n/M_n are all bounded, so by passage to a subsequence we may assume they all converge to real numbers c, d, C and D respectively. Note that these numbers are not all zero, since by construction they satisfy $|c| + |d| + |C| + |D| = 1$. Combining this with the fact that $\sigma'(g_n(\tau)) \to \sigma'(x)$ for any $\tau \in \mathbb{H}$ as $n \to \infty$, taking the limit in (4.5) yields, for all $\tau \in \mathbb{H}$,

$$\sigma'(x) = \left(\frac{c\tau + d}{C\sigma(\tau) + D}\right)^2 \sigma'(\tau).$$

Note that, since $\sigma'(x)$ is nonzero and finite, it cannot be the case that C and D are both zero, and similarly it cannot be the case that c and d are both zero. We thus have the equation

$$\frac{\sigma'(\tau)}{(C\sigma(\tau) + D)^2} = \frac{\sigma'(x)}{(c\tau + d)^2}.$$

Integrating both sides shows that σ is a Möbius transformation.

Next we claim that, given that G admits a cusp, the Möbius transformation must be an automorphism of $\mathbb{H}$. Since $\sigma(\mathbb{H}) \subseteq \mathbb{H}$ and $\sigma(x) = x$, the image $\sigma(\widehat{\mathbb{R}})$ is either $\widehat{\mathbb{R}}$ itself or a horocycle based at x. If we have the former then we are done, so suppose we have the latter. Then any cusp $r \in \operatorname{cusp}(G)$ must have $\sigma(r) \in \mathbb{H}$. The proof of Proposition 4.2 implies there is some parabolic element g fixing r that satisfies the functional equation $\sigma \circ g = \sigma$. Since σ is a Möbius transformation it is invertible, and we find $g = \mathrm{id}$. This contradicts the assumption that g was parabolic, so this case cannot happen and the proof is complete. □

Theorems 4.5, 4.7, and 4.8 immediately imply the following:

Theorem 4.9. *Suppose $\sigma\colon \mathbb{H} \to \mathbb{H}$ is modularly polymorphic and not an automorphism. Then σ has an interior fixed point $\tau_0 \in \mathbb{H}$ if and only if there is no fixed cusp $r \in \widehat{\mathbb{Q}}$ with dilation factor $\delta(r) \leq 1$.*

Using this theorem and the polymorphicity of the Thurston pullback map, we can now prove Thurston's characterization theorem for four marked points.

Theorem 4.10 (Thurston's criterion for four points)**.** *Let (f, A) be a Thurston map with $|A| = 4$ and suppose f is not of type $(2, 2, 2, 2)$. Then f is combinatorially equivalent to a rational map if and only if f has no Thurston obstruction, i.e., there is no f-invariant essential Jordan curve γ with $\lambda_f(\gamma) \geq 1$.*

Proof of Theorem 4.10. Consider the Thurston pullback map $\sigma_f \colon \mathbb{H} \to \mathbb{H}$, which is φ_f-polymorphic where $\varphi_f : H \to G$ is the virtual endomorphism induced on the subgroup of liftables. We fix the identification $G = \mathrm{PMod}(S^2, A) = \overline{\Gamma}(2)$, so this map is actually modularly polymorphic. If (f, A) is not of type $(2, 2, 2, 2)$, then $\sigma_f^2 = \sigma_{f^2}$ is strictly contracting and so σ_f is not an automorphism.

Cusp multipliers exactly coincide with Thurston multipliers of the associated curve class by Corollary 2.34. The result then follows from Theorems 4.5 and 4.9. □

We can also say something about the number of fixed cusps on the boundary. This is accomplished with the following general result, due to Cowen and Pommerenke (see [CP82] or [Aba23, Theorem 4.8.4]):

Theorem 4.11 (Cowen-Pommerenke inequality)**.** *Suppose $\sigma \colon \mathbb{H} \to \mathbb{H}$ is a holomorphic map that is not the identity with interior DW point $\tau_0 \in \mathbb{H}$. Suppose that $x_1, \dots, x_n$ are boundary fixed points of σ. Then*

$$\sum_{i=1}^{n} \frac{1}{\sigma'(x_i) - 1} \leq \frac{1 - |\sigma'(\tau_0)|^2}{|1 - \sigma'(\tau_0)|^2}.$$

Note that the list of boundary fixed points in the above theorem is not assumed to be a complete list.

Corollary 4.12. *Suppose $\sigma \colon \mathbb{H} \to \mathbb{H}$ is a modularly polymorphic map with interior DW point $\tau_0 \in \mathbb{H}$. Define*

$$C := \min\{\sigma'(r) : r \in \widehat{\mathbb{Q}} \text{ and } \sigma(r) = r\}.$$

Then $C > 1$ and

$$|\operatorname{Fix}(\sigma \cap \widehat{\mathbb{Q}})| \leq \frac{1}{C - 1} \left(\frac{1 - |\sigma'(\tau_0)|^2}{|1 - \sigma'(\tau_0)|^2} \right).$$

In particular, σ has finitely many fixed cusps.

Using the polymorphicity of the Thurston pullback map σ_f, we immediately obtain

Corollary 4.13. *Suppose (f, A) is a rational Thurston map with $|A| = 4$ that is not of type $(2,2,2,2)$. Then f has finitely many f-invariant essential curve classes.*

Furthermore, we can explicitly estimate the number of f-invariant essential curve classes using the inequality of Corollary 4.12. The constant C only depends on the underlying Thurston map (f, A) and can be computed in terms of the Thurston multipliers as

$$C = \frac{1}{\max\{\lambda_f(r) : r \in \widehat{\mathbb{Q}} \text{ and } \sigma_f(r) = r\}}.$$

See [Par18, Theorem 10.1] for a similar estimate.

CHAPTER 5

Cusp attractors of polymorphic maps

In this chapter we return our attention to cusp attractors for modularly polymorphic maps. Specifically, we will prove the following partial solution to the FCA conjecture:

Theorem 5.1. *Let* $\sigma\colon \mathbb{H} \to \mathbb{H}$ *be a modularly polymorphic map with spectrum of cusp dilation factors satisfying* $\min S_\delta(\sigma) > 1$*. Then* σ *has a finite cusp attractor.*

The techniques we will employ were previously used by the author to verify the FCA conjecture for a special subclass of Thurston maps with four marked points in [Smi24]. We begin with an overview of the strategy.

In Section 5.1 we will find a truncated hyperbolic space $X \subseteq \mathbb{H}$ that is forward-invariant under the map $\sigma\colon \mathbb{H} \to \mathbb{H}$. A truncated hyperbolic space is the complement of a countable disjoint family of horoballs. Specifically, we will punch out horoballs based at every rational contact point $r \in \widehat{\mathbb{Q}}$. The space so-obtained will be cocompact in the sense that X has compact image under the projection $\pi\colon \mathbb{H} \to \mathbb{H}/G$. Accordingly, σ will be uniformly contracting on X.

In Section 5.2 we shall then attach a "leash" from the DW point of σ to each horoball in the complement of X. Iterating the map σ will "tighten" this leash. There will only be finitely many horoballs that we may land on after this procedure, and the base cusps of these horoballs will exactly be our desired attractor on cusps. This will prove Theorem 5.1.

In Section 5.3 we apply the theorem to the the setting of Thurston maps. We will call a Thurston map *totally unobstructed* if its set of Thurston multipliers has

$$\Lambda_f := \max S_\lambda(\sigma_f) = \max\{\lambda_f(r) : r \in \widehat{\mathbb{Q}}\} < 1.$$

Our results show that totally unobstructed Thurston maps with four marked points have finite curve attractors. We will also provide examples to show that such maps indeed exist.

5.1. Invariant truncated space construction

Much of our work in this section involves "capturing" the image of σ in a particular truncation of hyperbolic space. Our first task is to characterize sets with this property.

Definition 5.2. For G a finite coarea Fuchsian group, we will call a subset $X \subseteq \mathbb{H}$ *cocompact with respect to* G if the image of X under the natural projection map $\pi_G : \mathbb{H} \to \mathbb{H}/G$ is compact in $\mathbb{H}/G$.

Note that, if G is a cocompact Fuchsian group, then every closed subset of $\mathbb{H}$ is cocompact with respect to G.

Lemma 5.3. *Let G be a torsion-free finite-index subgroup of* $\mathrm{PSL}(2, \mathbb{Z})$*. A closed set $X \subseteq \mathbb{H}$ is cocompact with respect to G if and only if there is some $t > 0$ so that*

$$\bigcup_{r \in \widehat{\mathbb{Q}}} B_t(r) \subseteq X^c.$$

Proof. First observe that, since G is a torsion-free finite coarea Fuchsian group, $W = \mathbb{H}/G$ is a Riemann surface obtained by removing a finite number of points from a closed surface. Label these punctures $p_1, \ldots, p_n$.

If X is cocompact, then $\pi_G(X)$ is compact in W. About each puncture p_i is an open set U_i that both: (i) lies in the complement of $\pi_G(X)$, and (ii) is homeomorphic to a punctured topological disk. Moreoever, we can choose each U_i so that some component of $\pi_G^{-1}(U_i)$ is a normalized horoball B_{t_i} in $\mathbb{H}$ with hororadius parameter $t_i > 0$. It will follow from Lemma 3.19 that all components of $\pi_G^{-1}(U_i)$ are horoballs of this form. These horoballs lie in the complement of X by construction. Since there are finitely many punctures, we may put $t = \min(t_1, \ldots, t_n)$, and then

$$\bigcup_{r \in \widehat{\mathbb{Q}}} B_t(r) \subseteq X^c,$$

as desired.

Conversely, suppose there is some hororadius parameter $t > 0$ such that

$$\bigcup_{r \in \widehat{\mathbb{Q}}} B_t(r) \subseteq X^c.$$

For each puncture p_i, choose a cusp $r_i \in \widehat{\mathbb{Q}}$ that corresponds to it. Then $U_i = \pi_G(B_t(r_i))$ is a neighborhood of the puncture p_i in the surface W. Moreover, any other cusp s_i which is equivalent to r_i under the action of G has $\pi_G(B_t(r_i)) = \pi_G(B_t(s_i))$. Thus

$$\pi_G\left(\bigcup_{r \in \widehat{\mathbb{Q}}} B_t(r)\right) = U_1 \cup \cdots \cup U_n,$$

which implies

$$\pi_G(X) \subseteq W \setminus (U_1 \cup \cdots \cup U_n).$$

Since the latter set is topologically a closed surface with a finite number of open disks removed, it is compact. Hence $\pi_G(X)$ is compact in W as a closed subset of a compact set. $\square$

With the previous lemma in mind, let us define a family of truncated spaces $X_t \subseteq \mathbb{H}$ given by

$$X_t = \mathbb{H} \setminus \left(\bigcup_{r \in \widehat{\mathbb{Q}}} B_t(r)\right),$$

where $0 < t < 1$. The aim will be to cleverly select the parameter t to obtain our forward-invariant set. The next few lemmas will aid us in this selection.

Throughout the remainder of this section, $\sigma\colon \mathbb{H} \to \mathbb{H}$ will be used to denote a modularly polymoprhic map unless explicitly noted otherwise. We also remind the reader that a contact cusp $r \in \widehat{\mathbb{Q}}$ is one for which $\sigma(r) \in \widehat{\mathbb{Q}}$, while a noncontact cusp is one for which $\sigma(r) \in \mathbb{H}$.

Lemma 5.4. *For all $t > 0$ sufficiently small,*

$$\sigma(B_t(r)) \subseteq X_t$$

for all noncontact cusps $r \in \widehat{\mathbb{Q}}$.

Proof. We assume without loss of generality that the intertwining homomorphism $\varphi \colon G \to G'$ is such that G and $\varphi(G)$ are both torsion-free finite-index subgroups of the modular group, so that the quotients $W = \mathbb{H}/G$ and $Z = \mathbb{H}/\varphi(G)$ are Riemann surfaces and we have the following commutative diagram:

$$\begin{array}{ccc} \mathbb{H} & \xrightarrow{\ \sigma\ } & \mathbb{H} \\ \Big\downarrow{\scriptstyle \pi_G} & & \Big\downarrow{\scriptstyle \pi_{\varphi(G)}} \\ W & \xrightarrow[\ \overline{\sigma}\ ]{} & Z. \end{array}$$

For $t > 0$ define the set

$$N_t = \bigcup B_t(r),$$

where the union is taken over all noncontact cusps $r \in \widehat{\mathbb{Q}}$. Our goal is to show there exists $t > 0$ so that $\sigma(N_t) \subseteq X_t$. Once such a parameter t is found, then all smaller t will work as well and the lemma will follow.

For all $t > 0$ sufficiently small, $\pi_G(N_t)$ is a finite union of neighborhoods of punctures in W. None of the corresponding punctures themselves map to punctures in Z (for otherwise the corresponding cusps would be contact), so $\overline{\sigma}(\pi_G(N_t))$ is a precompact set in Z. It follows that $\sigma(N_t)$ lies inside a cocompact set. By Lemma 5.3 there is some $t' > 0$ such that $\sigma(N_t) \subseteq X_{t'}$. If $t \leq t'$, then $X_{t'} \subseteq X_t$ and we are done. If $t > t'$, then since $\sigma(N_{t'}) \subseteq \sigma(N_t)$ we can just take t' as our desired hororadius parameter instead. This completes the proof. □

Lemma 5.5. *For all $t > 0$ sufficiently small,*

$$\sigma(B_t(r)) \cap B_t(s) = \varnothing$$

for all contact cusps $r \in \widehat{\mathbb{Q}}$ with $\sigma(r) \neq s$.

Proof. Modularly polymorphic maps have finitely many cusp dilation factors; let $\delta_0 = \max S_\delta(\sigma)$ denote the largest. For all $t > 0$ sufficiently small, $\delta_0 t < 1$. Corollary 4.6 implies $\sigma(B_t(r)) \subseteq B_{\delta_0 t}(\sigma(r))$. Since both $\delta_0 t < 1$ and $t < 1$, this cannot intersect the modular horoball $B_t(s)$. □

Lemma 5.6. *Suppose $\sigma\colon \mathbb{H} \to \mathbb{H}$ is a holomorphic map that fixes ∞ and satisfies the functional equation*

$$\sigma(\tau + a) = \sigma(\tau) + b$$

where $a, b > 0$ are real numbers and $\sigma'(\infty) = a/b > 1$. Then for all $t > 0$ sufficiently small, σ is injective on $H_t(\infty)$ and

$$\sigma(H_t(\infty)) \supseteq H_t(\infty).$$

Proof. We first show that σ is injective on $H_t := H_t(\infty)$ for all $t > 0$ sufficiently small. By the JWC theorem,

$$\angle \lim_{\tau\to\infty} \sigma'(\tau) = \frac{1}{\sigma'(\infty)} > 0.$$

Thus there is some Stolz angle S based at ∞ such that, for all $t > 0$ sufficiently small, $\operatorname{Re}\sigma'$ is positive on $S \cap H_t$. Using the polymorphic relation

$$\sigma(\tau + a) = \sigma(\tau) + b$$

we can show that $\operatorname{Re}\sigma'$ is positive on the whole horoball H_t. Since H_t is convex, it will follow that σ is injective on this set. Indeed, suppose to the contrary that there are points $\tau_1, \tau_2 \in H_t$ such that $\sigma(\tau_1) = \sigma(\tau_2)$. Let $\alpha \subseteq H_t$ be the line segment connecting these two points with parameterization $\alpha(t) = \tau_1 + t(\tau_2 - \tau_1)$ for $t \in [0, 1]$. Then the fundamental theorem of calculus yields

$$0 = \sigma(\tau_2) - \sigma(\tau_1) = \int_\alpha \sigma'\, d\tau = (\tau_2 - \tau_1)\int_0^1 \sigma'(\tau_1 + t(\tau_2 - \tau_1))\, dt.$$

Dividing both sides of the equation by $\tau_2 - \tau_1 \neq 0$ and taking the real part gives

$$0 = \int_0^1 \operatorname{Re}\big(\sigma'(\tau_1 + t(\tau_2 - \tau_1))\big)\, dt,$$

which is impossible since $\operatorname{Re}\sigma' > 0$ on $H_t \supseteq \alpha$. This proves the first claim.

Next we claim that, for all $t > 0$ sufficiently small, if $\operatorname{Im}\tau = 1/t$ then $\operatorname{Im}\sigma(\tau) < 1/t$. Suppose for contradiction the claim is false. Then there is a sequence of $t_n \to 0$ and $\tau_n \in \mathbb{H}$ such that $\operatorname{Im}\tau_n = 1/t_n$ but $\operatorname{Im}\sigma(\tau_n) \geq 1/t_n$. By repeated application of the polymorphic relation

$$\sigma(\tau + a) = \sigma(\tau) + b$$

where $a, b \neq 0$ are real, we may assume that each τ_n lies in the strip $S = \{\tau : |\operatorname{Re}\tau| \leq a\}$. It follows that $\tau_n \to \infty$ in a Stolz angle, so we should have

$$\lim_{n\to\infty} \frac{\operatorname{Im}\tau_n}{\operatorname{Im}\sigma(\tau_n)} = \sigma'(\infty) > 1$$

by the JWC theorem. Yet

$$\frac{\operatorname{Im}\tau_n}{\operatorname{Im}\sigma(\tau_n)} \leq \frac{1/t_n}{1/t_n} = 1$$

by construction. This proves the second claim. Combining this with the injectivity of σ now implies the lemma. □

Lemma 5.7. *Let $\sigma\colon \mathbb{H} \to \mathbb{H}$ be modularly polymorphic with* $\min S_\delta(\sigma) > 1$*. Then for all $t > 0$ sufficiently small, σ is injective on $B_t(r)$ for all contact cusps $r \in \widehat{\mathbb{Q}}$ and*

$$\sigma(B_t(r)) \supseteq B_t(\sigma(r)).$$

Proof. This follows from Lemma 5.6 and the derivative calculation of Theorem 4.5. □

In the following, by $\|d\sigma_\tau\|$ we shall mean the norm of the linear map $d\sigma_\tau\colon T_\tau\mathbb{H} \to T_{\sigma(\tau)}\mathbb{H}$ with respect to the hyperbolic metric.

Lemma 5.8. *Let $\sigma\colon \mathbb{H} \to \mathbb{H}$ be modularly polymorphic with* $\min S_\delta(\sigma) > 1$*. Then there is a $t > 0$ such that the set $X := X_t$ is forward-invariant under σ. Furthermore, there is some $\alpha < 1$ such that $\|d\sigma_\tau\| \leq \alpha$ for all $\tau \in X$.*

Proof. As we did in the proof of Lemma 5.4, we assume without loss of generality that the intertwining homomorphism $\varphi\colon G \to G'$ is such that G and $\varphi(G)$ are are torsion-free finite-index subgroups of the modular group, so that the quotients $W = \mathbb{H}/G$ and $Z = \mathbb{H}/\varphi(G)$

are Riemann surfaces and we have the following commutative diagram:

$$\begin{array}{ccc} \mathbb{H} & \xrightarrow{\sigma} & \mathbb{H} \\ {\scriptstyle \pi_G}\downarrow & & \downarrow{\scriptstyle \pi_{\varphi(G)}} \\ W & \xrightarrow[\overline{\sigma}]{} & Z. \end{array}$$

Choose $t > 0$ so small that Lemmas 5.4, 5.5, and 5.7 all hold.

The set $\pi_G(X_t)$ is compact in W, so $\overline{\sigma}(\pi_G(X_t))$ must be compact in Z. It follows that $\sigma(X_t)$ must be cocompact with respect to $\varphi(G)$. By Lemma 5.3, this implies there is some hororadius parameter $t' > 0$ such that $\sigma(X_t) \subseteq X_{t'}$. Now put $X := X_t \cup X_{t'}$. We will show this is the desired forward-invariant subspace.

There are several cases to consider.

If $t \leq t'$, then $X_{t'} \subseteq X_t$ and $X = X_t$. The paragraph above shows that $\sigma(X) \subseteq X_{t'} \subseteq X$, so we are done with this case.

Now suppose that $t' < t$. Then $X_t \subseteq X_{t'}$ and $X = X_{t'}$. We already know that X_t is mapped into $X_{t'}$ by σ, so we must show that $X_{t'} \setminus X_t$ maps into $X_{t'}$. Suppose otherwise. Then there is some $r \in \widehat{\mathbb{Q}}$ and some $\tau \in B_t(r) \setminus B_{t'}(r)$ such that $\sigma(\tau) \notin X_{t'}$, or equivalently, $\sigma(\tau) \in B_{t'}(s)$ for some $s \in \widehat{\mathbb{Q}}$. We further divide into two subcases:

If $r \in \widehat{\mathbb{Q}}$ is a contact cusp, then Lemma 5.5 implies $s = r'$. But then $\sigma(\tau) \in B_{t'}(r') \subseteq \sigma(B_{t'}(r))$. Thus there is some $\tau_1 \in B_{t'}(r)$ such that $\sigma(\tau_1) = \sigma(\tau)$, and $\tau_1 \neq \tau$ since $\tau \in B_t(r) \setminus B_{t'}(r)$. This contradicts the assumption that σ is injective on $B_t(r)$, so we are done in this case.

Now suppose $r \in \widehat{\mathbb{Q}}$ is not a contact cusp, i.e., $\sigma(r) \in \mathbb{H}$. In this case there is nothing to do, since t was chosen such that $\sigma(B_t(r)) \subseteq X_t$ by Lemma 5.4.

All cases have been exhausted, so the first claim of the lemma is proven.

For the second claim, consider the function $K\colon \mathbb{H} \to \mathbb{R}$ defined by $K(\tau) = \|d\sigma_\tau\|$. This map is continuous. Furthermore, if $\tau_2 = g(\tau_1)$ for some $\tau_1 \in \mathbb{H}$ and $g \in G$, then by polymorphicity and the chain rule we have

$$d\sigma_{\tau_2} = d(\varphi(g) \circ \sigma)_{\tau_1} = d(\varphi(g))_{\sigma(\tau_1)} d\sigma_{\tau_1}.$$

Since $\varphi(g)$ is an isometry of $\mathbb{H}$, we have $\|d(\varphi(g))_\tau\| = 1$ for all $\tau \in \mathbb{H}$. Thus $K(\tau_2) = K(\tau_1)$, and K descends to a well-defined continuous map $\overline{K}\colon W \to \mathbb{R}$ defined by $\overline{K}(G\tau) = K(\tau)$.

Since $\pi_G(X)$ is compact and $\overline{K}$ is a continuous function with $\overline{K}(w) < 1$ for all $w \in W$, there is some $\alpha < 1$ such that $\overline{K}(w) \leq \alpha$ for all $w \in \pi_G(X)$. But then $\|d\sigma_\tau\| = K_f(\tau) \leq \alpha$ for all $\tau \in X$. This completes the proof. □

5.2. Leashing and proof of Theorem 5.1

Fix X as in the previous section. If γ is a smooth path in X and $\gamma' = \sigma(\gamma)$, then $l(\gamma') \leq \alpha l(\gamma)$ where $l(\cdot)$ denotes the length of the curve with respect to the hyperbolic metric. We also know that $\gamma' \subseteq X$ since X was chosen to be σ-invariant. With these nice properties in mind, consider the following path metric induced on X:

$$d_X(\tau_1, \tau_2) = \inf\{l(\gamma) : \gamma\colon [0,1] \to X \text{ is a smooth path connecting } \tau_1 \text{ to } \tau_2\}.$$

This metric is well-defined since X is path-connected, and by construction σ is uniformly contracting on (X, d_X). To avoid notational clutter, we shall simply write $d(\cdot,\cdot) = d_X(\cdot,\cdot)$.

Lemma 5.9. *There is some constant $C \geq 0$ depending only on σ such that, for all contact cusps $r \in \widehat{\mathbb{Q}}$ with $r' = \sigma(r) \in \widehat{\mathbb{Q}}$,*

$$d\big(\tau_0, \partial B_t(r')\big) \leq \alpha d\big(\tau_0, \partial B_t(r)\big) + C.$$

Proof. Let $r \in \widehat{\mathbb{Q}}$ be a contact cusp so that $r' = \sigma(r) \in \widehat{\mathbb{Q}}$. Let $\epsilon > 0$, and take a path γ from τ_0 to some point on $\partial B_t(r)$ such that $l(\gamma) < d(\tau_0, \partial B_t(r)) + \epsilon$.

Note that for $\delta := \delta(r)$, we have

$$\sigma(B_t(r)) \subseteq B_{\delta t}(r').$$

Put $\gamma' = \sigma(\gamma)$. Since $\delta > 1$, we obtain the configuration depicted in Figure 4. This yields

$$\begin{aligned} d(\tau_0, \partial B_t(r') &\leq l(\gamma') + d(\partial B_{\delta t}(r'), \partial B_t(r')) \\ &\leq \alpha l(\gamma) + d(\partial B_{\delta t}(r'), \partial B_t(r')). \end{aligned}$$

On the other hand, a simple calculation shows that $d(\partial B_{\delta t}(r'), \partial B_t(r')) = |\ln(\delta)|$.

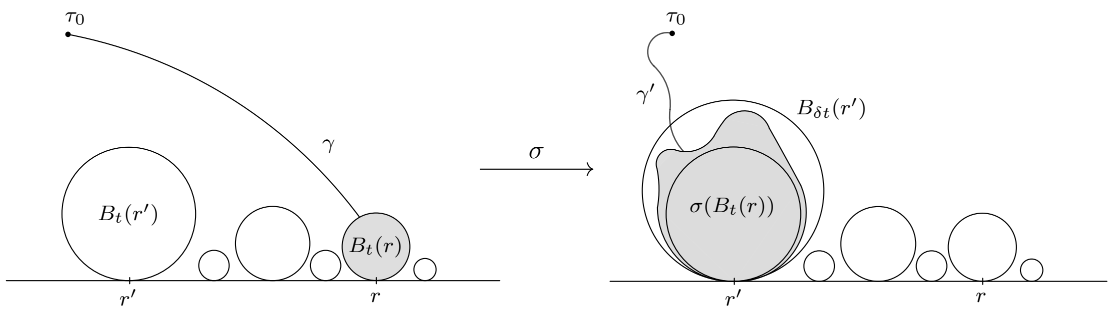


FIGURE 4. Generic picture of leash image.

The quantity $|\ln(\delta)|$ where δ is a contact cusp dilation factor takes on finitely many positive values; choose $C > 0$ so that $|\ln(\delta)| \leq C$. It follows that

$$d\big(\tau_0, \partial B_t(r')\big) \leq \alpha(d\big(\tau_0, \partial B_t(r)\big) + \epsilon) + C$$

for all $\epsilon > 0$. Since $\epsilon > 0$ was arbitrary, we obtain the desired inequality. $\square$

The next two lemmas show that our "leash" tightens in such a way that there are only finitely many cusps we may land on. They are reproduced from [Smi24, Section 4.5].

Lemma 5.10. *Let $(x_n) = (x_0, x_1, x_2, \dots)$ be a sequence of nonnegative real numbers with the property that there is some $\alpha \in (0, 1)$ and $C \geq 0$ for which*

$$x_n \leq \alpha x_{n-1} + C$$

for all $n \geq 1$. Then for each $\epsilon > 0$ there is an N so that

$$x_n \leq \frac{C}{1-\alpha} + \epsilon$$

for all $n \geq N$.

Proof. Observe that

$$x_2 \leq \alpha x_1 + C \leq \alpha(\alpha x_0 + C) + C = \alpha^2 x_0 + \alpha C + C.$$

Continuing inductively, we get

$$x_n \leq \alpha^n x_0 + C \sum_{k=0}^{n-1} \alpha^k = \alpha^n x_0 + C \frac{1-\alpha^n}{1-\alpha} = \frac{C}{1-\alpha} + \alpha^n \left(x_0 - \frac{C}{1-\alpha} \right)$$

for $n \geq 1$. Thus, after taking absolute values we have the estimate

$$x_n \leq \frac{C}{1-\alpha} + \alpha^n \left| x_0 - \frac{C}{1-\alpha} \right|.$$

Since $\alpha^n \to 0$ as $n \to \infty$, the claimed result follows. □

Lemma 5.11. *A disjoint collection of nonempty open horoballs in $\mathbb{H}$ is locally finite, meaning any compact subset of $\mathbb{H}$ intersects only finitely many horoballs from the collection.*

Proof. This result is easier to see in the Poincaré disk model of the hyperbolic plane. In this setting, a compact set K in $\mathbb{D}$ must be contained inside some Euclidean disk $D(0, r)$ with $r < 1$. Suppose that a horoball intersects K, and hence also $D(0, r)$. Then the diameter of the horoball is at least $1 - r$, its Euclidean radius is at least $(1 - r)/2$, and its Euclidean area is at least $\pi(1 - r)^2/4$. Since $\mathbb{D}$ has finite Euclidean area π, there can be at most finitely many disjoint horoballs in the collection which intersect K. This completes the proof. □

Proof of Theorem 5.1. Let $r \in \widehat{\mathbb{Q}}$. If r or any of its images is noncontact then we are done, so suppose that $\sigma^n(r) \in \widehat{\mathbb{Q}}$ for all $n \in \mathbb{N}$. Define the sequence

$$x_n = d(\tau_0, \partial B_t(\sigma^n(r)))$$

for integers $n \geq 1$, and we put $x_0 = d(\tau_0, r)$. By Lemma 5.9, this sequence satisfies the hypotheses of Lemma 5.10. Thus there is some $N(r)$ so that

$$x_n := d(\tau_0, \partial B_t(\sigma^n(r))) \leq \frac{C}{1-\alpha} + 1$$

for all $n \geq N(r)$. Yet the set

$$K = \left\{ \tau \in \mathbb{H} : d(\tau_0, \tau) \leq \frac{C}{1-\alpha} + 1 \right\}$$

is a compact subset of X and hence also of $\mathbb{H}$. Thus K intersects only finitely many horoballs from the collection $\{B_t(r) : r \in \widehat{\mathbb{Q}}\}$ by Lemma 5.11. Let $\mathcal{A} = \{r_1, \ldots, r_k\}$ be the base cusps of this finite collection of horoballs. Then it must be the case $\sigma_f^n(r) \in \mathcal{A}$ for $n \geq N(r)$. Since C, α and hence $\mathcal{A}$ depend only on the map σ and not on the initial cusp r, the set $\mathcal{A}$ is our desired finite cusp attractor. $\square$

5.3. Totally unobstructed Thurston maps

As mentioned in the introduction to this chapter, given a Thurston map (f, A) with $|A| = 4$, we will say that (f, A) is *totally unobstructed* if all of its Thurston multipliers have $\lambda_f(r) < 1$, i.e.,

$$\Lambda_f := \max\{\lambda_f(r) : r \in \widehat{\mathbb{Q}}\} < 1.$$

We shall sometimes refer to Λ_f as the *multiplier radius* of the Thurston map (f, A). Note that $\Lambda_f = \max S_\lambda(\sigma_f) = 1/\min S_\delta(\sigma_f)$ since Thurston multipliers and cusp multipliers coincide for the Thurston pullback map, and cusp multipliers and cusp dilation factors are reciprocally related. Thus, if (f, A) is totally unobstructed, then the associated Thurston pullback map σ_f is a modularly polymorphic map satisfying the hypotheses of Theorem 5.1. This proves

Corollary 5.12. *If (f, A) is a totally unobstructed Thurston map, then (f, A) has a finite global curve attractor.*

The class of totally unobstructed Thurston maps is nonempty, as shall see in the following examples.

Example 5.13. Consider the rational Thurston map $f\colon \widehat{\mathbb{C}} \to \widehat{\mathbb{C}}$ defined by

$$f(z) = \frac{2z^3+1}{3z^2}.$$

Taking the derivative gives

$$f'(z) = -\frac{2(1-z^3)}{3z^3}.$$

If we put $\omega = e^{2\pi i/3} = (-1+i\sqrt{3})/2$, then $C_f = \{0, 1, \omega, \omega^2\}$ and $P_f = \{1, \omega, \omega^2, \infty\}$. We have the following dynamical portrait:

$$0 \xrightarrow{2:1} \infty \circlearrowright \qquad 1 \circlearrowleft 2{:}1 \qquad \omega \circlearrowleft 2{:}1 \qquad \omega^2 \circlearrowleft 2{:}1$$

This map has orbifold signature $(2, \infty, \infty, \infty)$, and so is not of type $(2, 2, 2, 2)$. It is an example of a *nearly Euclidean Thurston (NET) map*, which are Thurston maps with exactly four postcritical points and the local degree at each critical point is 2. These maps have been extensively in [CFPP12] and [FKK$^+$17]. The program `NETmap` written by Walter Parry with assistance from Bill Floyd [PF16] is able to generate a complete list of Thurston multipliers for such maps. For our particular example, the set of Thurston multipliers is calculated by `NETmap` to be $S_\lambda = \{0, 1/3\}$. Thus (f, P_f) is a totally unobstructed Thurston map and hence has a finite curve attractor.

Example 5.14. Consider the Thurston map (f, P_f) with four poscritical points which is determined by the combinatorial square-tile picture in the figure below (compare Example 2.5):

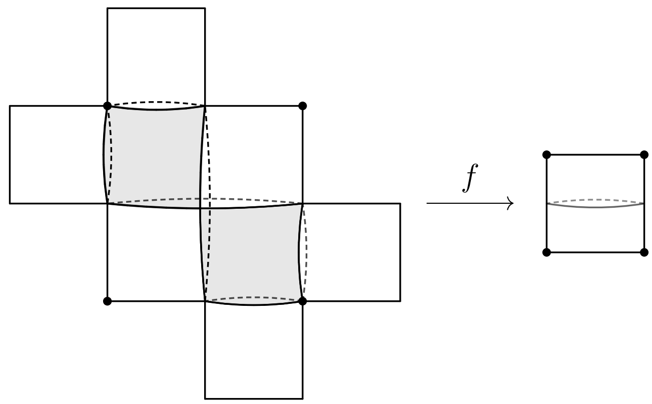


FIGURE 5. A combinatorial depiction of the map (f, P_f) obtained from blowing up the 2×2 Lattès map along four arcs.

This map is easily seen to not be of type $(2, 2, 2, 2)$. Walter Parry has written a program `Pf4` which is also able to produce a full set of Thurston multipliers for maps described by such data (at least for maps of sufficiently small degree); see [Par22]. In this case, $S_\lambda = \{0, 2/3\}$. Thus this map is also totally unobstructed and has a finite curve attractor.

CHAPTER 6

Further discussion

In this final chapter we discuss some conjectures and possible future directions of study.

The most obvious remaining question is whether the FCA conjecture for four marked points can be proven using our analysis. In Sections 6.1 and 6.2 we present some initial attempts at this. We introduce quantities that capture the long-term behavior of the Thurston pullback at cusps. We speculate that further studying these quantities might allow us to extend the results of the previous chapter to the general case of rational Thurston maps that are not of type $(2,2,2,2)$. These discussions also apply more generally to modularly polymorphic maps, but we do not formulate them as such.

In Section 6.3 we discuss what changes in the setting of a Thurston map (f, A) where $|A| \geq 5$. There is no easy translation of the polymorphic machinery to this case and we point out some difficulties.

6.1. Dynamical multiplier radius

Recall the multiplier radius for a Thurston map (f, A) with $|A| = 4$ is given by

$$\Lambda_f := \max\{\lambda_f(r) : r \in \widehat{\mathbb{Q}}\}.$$

This quantity is positive unless all pullbacks of essential Jordan curves in (S^2, A) are peripherial.

We conjecture the following:

Conjecture 6.1. *If (f, A) is a rational Thurston map that is not of type $(2,2,2,2)$, then there is some iterate (f^N, A) which is totally unobstructed.*

To study this conjecture, we shall consider a quantity Λ which captures the long term dynamical behavior of the multiplier radius.

For nonnegative integers n, m, we have

$$\Lambda_{f^{n+m}} \leq \Lambda_{f^n} \Lambda_{f^m}$$

since a maximal multiplier $\lambda_{f^{n+m}}(r)$ on the left can be decomposed as

$$\lambda_{f^{n+m}}(r) = \lambda_{f^n}(r) \lambda_{f^m}(\sigma_f^n(r)) \leq \Lambda_{f^n} \Lambda_{f^m}.$$

Note if $\Lambda_{f^N} = 0$ for some integer N, then the pullback action of (f^N, A) is trivial. We are not interested in this case, so we shall discard it henceforth. Thus in all interesting cases, the sequence Λ_{f^n} is a submultiplicative sequence of positive numbers, and so

$$\Lambda = \lim_{n \to \infty} (\Lambda_{f^n})^{1/n}$$

exists by Fekete's lemma. We shall call this quantity Λ the *dynamical multiplier radius* of f.

An immediate consequence of the definition is

Proposition 6.2. *If $\Lambda < 1$, then some iterate of (f, A) is totally unobstructed.*

We make the complementary conjecture:

Conjecture 6.3. *If $\Lambda \geq 1$, all iterates of (f, A) are obstructed.*

6.2. Asymptotic multiplier growth

For each $r \in \widehat{\mathbb{Q}}$ with $\sigma_f^n(r)$ always contact, consider the sequence $s_n(r) := \lambda_{f^n}(r)$. Note that if r_1 and r_2 belong to the same grand orbit under iteration by σ_f, then the tails of $s_n(r_1)$ and $s_n(r_2)$ match. Thus if r_1 and r_2 belong to the same grand orbit, the orbits of the multipliers have Θ-equivalent asymptotic growth rates. In other words, $s_n(r_1) = O(s_n(r_2))$ and $s_n(r_2) = O(s_n(r_1))$. The property of being in the same grand orbit is an equivalence relation. Since there are countably many extended rational cusps, there are at most countably many grand orbit equivalence classes. To each class $[r]$ we may assign an asymptotic growth Θ-equivalence class, which we may represent as $g\colon \mathbb{N}_0 \to [0, \infty)$.

Proposition 6.4. *If there is a grand orbit cusp class whose asymptotic growth function has* $\limsup g(n) > 0$*, then* (f, A) *is obstructed.*

Proof. Suppose for contradiction that (f, A) is not obstructed, meaning σ_f has a unique fixed point $\tau_0 \in \mathbb{H}$. The hypotheses say there is some $r \in \widehat{\mathbb{Q}}$ such that $\limsup \lambda_{f^n}(r) > 0$. Let $r_n := \sigma_f^n(r) \in \widehat{\mathbb{Q}}$ for $n \geq 0$, and write $r_n = p_n/q_n$ for coprime integers p_n and $q_n \geq 0$. We have the inequality

$$|p_n\tau_0 + q_n|^2 \leq \frac{1}{\lambda_{f^n}(r)}|p_0\tau_0 + q_0|^2. \tag{6.1}$$

This is essentially just the inclusion $\sigma_f(B_t(r)) \subseteq B_{t/\lambda_f(r)}(r')$ for $\sigma_f(r) = r'$ iterated n times.

Put $M = \limsup \lambda_{f^n}(r) > 0$, which may be infinite. There is a subsequence such that $\lambda_{f^{n_k}}(r) \to M$ as $k \to \infty$. If $M = \infty$, then for this subsequence (6.1) gives us

$$\lim_{k\to\infty} |p_{n_k}\tau_0 + q_{n_k}|^2 = 0.$$

If M is finite, we have

$$\limsup_{k\to\infty} |p_{n_k}\tau_0 + q_{n_k}|^2 \leq \frac{1}{M}|p_0\tau_0 + q_0|^2.$$

In either case, we see that the set $|p_{n_k}\tau_0 + q_{n_k}|$ is bounded in k. Since a bounded set intersects at most finitely many points in the lattice $\tau_0\mathbb{Z} \oplus \mathbb{Z}$, the pigeonhole principle implies there are numbers k and j such that $r_{n_k} = r_{n_j}$. Yet this implies r is preperiodic and thus has finite orbit.

Next we claim that $\lambda_{f^n}(r) = \Theta(\alpha^n)$ for some $\alpha \geq 1$. Indeed, since r is preperiodic there are minimal integers $N \geq 0$ and $m \geq 1$ such that $\sigma_f^{N+m}(r) = \sigma_f^m(r)$. Put

$$\alpha = \left(\prod_{k=0}^{m} \lambda_f(\sigma_f^{N+k}(r))\right)^{1/m}.$$

From here it is not hard to see that $\lambda_{f^n}(r) = \Theta(\alpha^n)$, as desired.

If $\alpha < 1$, then the original assumption $\limsup \lambda_{f^n}(r) > 0$ is false, so we must have $\alpha \geq 1$. Yet if this is the case, then $\alpha^m \geq 1$ as well. Since $\sigma_f^m = \sigma_{f^m}$ fixes every point of the cycle $\{\sigma_f^N(r), \ldots, \sigma_f^{N+m}(r)\}$ and each of these cusps have multiplier α^m under f^m, the Thurston map (f^m, A) is obstructed. This is impossible though, since if (f, A) is unobstructed then all iterates of this map are unobstructed as well. This completes the proof. $\square$

It follows that for an unobstructed map (f, A), we always have $\limsup g(n) = 0$ for every multiplier growth function associated to f.

In the above proof, we proved the following proposition along the way:

Proposition 6.5. *If r is a preperiodic cusp, then $\lambda_{f^n}(r) = \Theta(\alpha_r^n)$ for a unique $\alpha_r > 0$. If (f, A) is unobstructed, then $\alpha_r < 1$ for each preperiodic r.*

Conjecture 6.6. *For (f, A) unobstructed, $\sup \alpha_r < 1$, where the supremum is taken over all preperiodic cusps.*

6.3. Generalization to more marked points

A natural question to ask is whether any of our considerations extend to Thurston maps (f, A) with $|A| \geq 5$. After all, Theorem 2.30 concerning the polymorphicity of the Thurston pullback holds even in this case, so long as you interpret the pure mapping class group and the subgroup of liftables as subgroups of $\mathrm{Aut}(\mathcal{T}_A)$.

There are many difficulties with this idea. Although $\dim \mathcal{T}_A = |A| - 3$ and $\mathcal{T}_A = \mathbb{H}$ for $|A| = 4$, it is not the case that $\mathcal{T}_A = \mathbb{H}^{|A|-3}$. The structure of the higher dimensional Teichmüller spaces is significantly more complicated than this. The WP-completion of $\mathcal{T}_A$ in these cases is a *stratified space* with strata corresponding to multicurves in (S^2, A) (see [Sel12]). Classical function theory, which was the main tool four our analysis in this thesis, does not apply to this setting. Nevertheless, one wonders if some appropriate analogue of angular derivatives and the Denjoy–Wolff theorems might hold for self-maps of $\mathcal{T}_A$.

Question 6.7. *Is there a formulation of the Denjoy–Wolff theorem for holomorphic self-maps of $\mathcal{T}_A$? For polymorphic self-maps? And if so, what is the correct analogue of angular derivatives in these cases?*

Bibliography

[Aba23] Marco Abate, *Holomorphic Dynamics on Hyperbolic Riemann Surfaces*, De Gruyter Studies in Mathematics, vol. 89, De Gruyter, Berlin, 2023.

[Alp87] Roger C. Alperin, *An elementary account of Selberg's lemma*, Enseign. Math. (2) **33** (1987), no. 3-4, 269–273.

[AN93] J. M. Aarts and T. Nishiura, *Dimension and Extensions*, North-Holland Mathematical Library, vol. 48, North-Holland Publishing Co., Amsterdam, 1993.

[BCT14] Xavier Buff, Guizhen Cui, and Lei Tan, *Teichmüller spaces and holomorphic dynamics*, Handbook of Teichmüller theory. Vol. IV, 2014, pp. 717–756.

[Bea95] Alan F. Beardon, *The Geometry of Discrete Groups*, Graduate Texts in Mathematics, vol. 91, Springer-Verlag, New York, 1995.

[BEKP09] Xavier Buff, Adam Epstein, Sarah Koch, and Kevin Pilgrim, *On Thurston's pullback map*, Complex dynamics, 2009, pp. 561–583.

[BHI21] Mario Bonk, Mikhail Hlushchanka, and Annina Iseli, *Eliminating Thurston obstructions and controlling dynamics on curves*, Preprint: arXiv:2105.06938, 2021.

[BLMW22] James Belk, Justin Lanier, Dan Margalit, and Rebecca R. Winarski, *Recognizing topological polynomials by lifting trees*, Duke Math. J. **171** (2022), no. 17, 3401–3480.

[BM17] Mario Bonk and Daniel Meyer, *Expanding Thurston maps*, Mathematical Surveys and Monographs, vol. 225, American Mathematical Society, Providence, RI, 2017.

[Bur79] Robert B. Burckel, *An Introduction to Classical Complex Analysis. Vol. 1*, Pure and Applied Mathematics, vol. 82, 1979.

[BWFH21] Heinrich Burkhardt, Wilhelm Wirtinger, Robert Fricke, and Emil Hilb, *Encyklopadie der mathematischen Wissenschaften mit Einschluss Ihrer Anwendungen. bd. 2: Analysis. 2. Teil*, Leipzig, 1921.

[CFPP12] J. W. Cannon, W. J. Floyd, W. R. Parry, and K. M. Pilgrim, *Nearly Euclidean Thurston maps*, Conform. Geom. Dyn. **16** (2012), 209–255.

[CP82] Carl C. Cowen and Christian Pommerenke, *Inequalities for the angular derivative of an analytic function in the unit disk*, J. London Math. Soc. (2) **26** (1982), no. 2, 271–289.

[DH93] Adrien Douady and John H. Hubbard, *A proof of Thurston's topological characterization of rational functions*, Acta Math. **171** (1993), no. 2, 263–297.

[FKK+17] William Floyd, Gregory Kelsey, Sarah Koch, Russell Lodge, Walter Parry, Kevin M. Pilgrim, and Edgar Saenz, *Origami, affine maps, and complex dynamics*, Arnold Math. J. **3** (2017), no. 3, 365–395.

[FM12] Benson Farb and Dan Margalit, *A Primer on Mapping Class Groups*, Princeton Mathematical Series, vol. 49, Princeton University Press, Princeton, NJ, 2012.

[Fri12] Robert Fricke, *Die elliptischen Funktionen und ihre Anwendungen. Erster Teil. Die funktionentheoretischen und analytischen Grundlagen*, Springer, Heidelberg, 2012. Original work published 1916.

[Hat02] Allen Hatcher, *Algebraic Topology*, Cambridge University Press, Cambridge, 2002.

[Hej75] Dennis A. Hejhal, *Monodromy groups and linearly polymorphic functions*, Acta Math. **135** (1975), no. 1, 1–55.

[Hej76] ———, *The variational theory of linearly polymorphic functions*, J. Analyse Math. **30** (1976), 215–264.

[Hlu19] Mikhail Hlushchanka, *Tischler graphs of critically fixed rational maps and their applications*, Preprint: arXiv:1904.04759, 2019.

[Kat92] Svetlana Katok, *Fuchsian Groups*, Chicago Lectures in Mathematics, University of Chicago Press, Chicago, IL, 1992.

[KL19] Gregory Kelsey and Russell Lodge, *Quadratic Thurston maps with few postcritical points*, Geom. Dedicata **201** (2019), 33–55.

[Koc13] Sarah Koch, *Teichmüller theory and critically finite endomorphisms*, Adv. Math. **248** (2013), 573–617.

[KPS16] Sarah Koch, Kevin M. Pilgrim, and Nikita Selinger, *Pullback invariants of Thurston maps*, Trans. Amer. Math. Soc. **368** (2016), no. 7, 4621–4655.

[Lod13] Russell Lodge, *Boundary values of the Thurston pullback map*, Conform. Geom. Dyn. **17** (2013), 77–118.

[Mil06] John Milnor, *Dynamics in One Complex Variable*, Third, Annals of Mathematics Studies, vol. 160, Princeton University Press, Princeton, NJ, 2006.

[MP08] Diego Mejía and Christian Pommerenke, *On groups and normal polymorphic functions*, Rev. Colombiana Mat. **42** (2008), no. 2, 167–181.

[MP12a] ———, *On the boundary behaviour of polymorphic functions*, Comput. Methods Funct. Theory **12** (2012), no. 1, 201–212.

[MP12b] ———, *On the construction of polymorphic functions*, 60 years of analytic functions in Lublin, 2012, pp. 157–167.

[Par18] Walter Parry, *NET map slope functions*, Preprint: arXiv:1811.01274, 2018.

[Par22] ———, *Pf4*, 2022. https://sourceforge.net/projects/pf4/.

[PF16] Walter Parry and William Floyd, *NETmap*, 2016. https://intranet.math.vt.edu/netmaps/.

[Pil12] Kevin M. Pilgrim, *An algebraic formulation of Thurston's characterization of rational functions*, Ann. Fac. Sci. Toulouse Math. (6) **21** (2012), no. 5, 1033–1068.

[Pil22] ———, *On the pullback relation on curves induced by a Thurston map*, In the tradition of Thurston II. Geometry and groups, 2022, pp. 385–399.

[Pom81] Ch. Pommerenke, *Polymorphic functions for groups of divergence type*, Math. Ann. **258** (1981/82), no. 4, 353–366.

[PT57] G. Piranian and W. J. Thron, *Convergence properties of sequences of linear fractional transformations*, Michigan Math. J. **4** (1957), 129–135.

[Roy71] H. L. Royden, *Automorphisms and isometries of Teichmüller space*, Advances in the Theory of Riemann Surfaces (Proc. Conf., Stony Brook, N.Y., 1969), 1971, pp. 369–383.

[Sel12] Nikita Selinger, *Thurston's pullback map on the augmented Teichmüller space and applications*, Invent. Math. **189** (2012), no. 1, 111–142.

[Sha93] Joel H. Shapiro, *Composition Operators and Classical Function Theory*, Universitext: Tracts in Mathematics, Springer-Verlag, New York, 1993.

[Smi24] Zachary Smith, *Curve attractors for marked rational maps*, Preprint: arXiv:2401.16636, 2024.

[Zyg02] A. Zygmund, *Trigonometric Series. Vol. I, II*, Third, Cambridge Mathematical Library, Cambridge University Press, Cambridge, 2002.