

FIXED-POINT-FREE ELEMENTS OF ITERATED MONODROMY GROUPS

RAFE JONES

ABSTRACT. The iterated monodromy group of a post-critically finite complex polynomial of degree $d \geq 2$ acts naturally on the complete d -ary rooted tree T of preimages of a generic point. This group, as well as its profinite completion, acts on the boundary of T , which is given by extending the branches to their “ends” at infinity. We show that, in most cases, elements that have fixed points on the boundary are rare, in that they belong to a set of Haar measure 0. The exceptions are those polynomials linearly conjugate to multiples of Chebyshev polynomials and a case that remains unresolved, where the polynomial has a non-critical fixed point with many critical preimages. The proof involves a study of the finite automaton giving generators of the iterated monodromy group, and an application of a martingale convergence theorem. Our result is motivated in part by applications to arithmetic dynamics, where iterated monodromy groups furnish the “geometric part” of certain Galois extensions encoding information about densities of dynamically interesting sets of prime ideals.

1. INTRODUCTION

Suppose that $f \in \mathbb{C}[x]$ is a polynomial of degree $d \geq 2$, and let $\text{Crit } f$ be the set of critical points of f . Define the *post-critical set* $P_f := \bigcup_{n \geq 1} f^n(\text{Crit } f)$, where f^n denotes the n th iterate of f . Note that P_f consists of all points over which the map $f^n : \mathbb{C} \rightarrow \mathbb{C}$ is ramified for at least one n . When P_f is finite, we call f *post-critically finite*. Choose a point $\beta \in \mathbb{C} \setminus P_f$, and denote by T_β the disjoint union of the sets $f^{-n}(\beta)$ of preimages of β under f^n , for $n \geq 0$ (we take f^0 to be the identity map). Assigning edges according to the action of f , T_β becomes a complete rooted d -ary tree whose n th level is $f^{-n}(\beta)$. The fundamental group $\pi_1(\mathbb{C} \setminus P_f, \beta)$ acts on each set $f^{-n}(\beta)$ by monodromy, and thus gives a subgroup of $\text{Aut}(T_\beta)$ that we call the *iterated monodromy group* of f , and write $\text{IMG}(f)$.

In the decade since their introduction, iterated monodromy groups have become a powerful tool used in a variety of settings (see the introduction in [16] for a summary of the history of the subject). They are both computable and have deep connections to the dynamics of the underlying polynomial. Indeed, the action of a set of generators of $\text{IMG}(f)$ on T_β can be given by a simple finite automaton that depends largely on the structure of the set P_f (see Section 5 for more details). To illustrate the connections to dynamics, one can associate to $\text{IMG}(f)$ a *limit dynamical system* whose points are equivalence classes of left-infinite paths in T_β , and whose map is the shift map. This dynamical system is topologically conjugate to the action of f on its Julia set [15, Section 3.6 and Theorem 6.4.4]. For this

Received by the editors June 9, 2012 and, in revised form, May 1, 2013.

2010 *Mathematics Subject Classification*. Primary 37F10, 20E08; Secondary 37P05, 37F20.

The author’s research was partially supported by NSF grant DMS-0852826.

reason the group $\text{IMG}(z^2 - 1)$ has become known as the *Basilica group*, since the top half of the Julia set of $z^2 - 1$ bears a striking resemblance to the profile of the Basilica di San Marco in Venice.

Applications of iterated monodromy groups abound. Defined in a more general setting, they have been used by Bartholdi and Nekrashevych to resolve the well-known “twisted rabbit” problem of J. Hubbard [1]. They have also attracted interest for their purely group-theoretic properties; for instance, the Basilica group is the first known example separating the classes of amenable groups and subexponentially amenable groups [3]. The more general class of groups generated by finite automata includes the renowned Grigorchuk group [6], the first example of a group of intermediate growth. The survey [16] and the monograph [15] give an overview of iterated monodromy groups and their applications, as well as extensive bibliographies.

Our interest in iterated monodromy groups comes from arithmetic, where properties of the action of $\text{IMG}(f)$ on the boundary of T_β yield information about interesting sets of prime ideals (See Section 2 for more details). In particular, we are interested in elements of $\text{IMG}(f)$ that fix at least one point on the boundary of T_β , or equivalently, at least one infinite branch of T_β . Our main result is that such elements are rare for a large class of f .

Let us introduce some notation. We may identify T_β with the set X^* of all finite words (including the empty word) over an alphabet X containing d letters. The root of T_β corresponds to the empty word, and $f^{-n}(\beta)$ corresponds to X^n , the set of all words of length n . The boundary of X^* is the set $X^\omega = \{x_1x_2x_3 \cdots : x_i \in X\}$ of *ends* of X^* . Let G be a group of automorphisms of X^* , and denote by G_n the image of G under the restriction map $\text{Aut}(X^*) \rightarrow \text{Aut}(X^n)$. Define

$$(1.1) \quad \mathcal{F}(G) = \lim_{n \rightarrow \infty} \frac{\#\{g \in G_n : g \text{ fixes at least one element of } X^n\}}{\#G_n}.$$

Note that the fraction above is non-increasing, since all lifts to G_{n+1} of an element of G_n with no fixed points again have no fixed points. Thus the limit in (1.1) exists. We may also describe $\mathcal{F}(G)$ by considering the closure G_∞ of G in $\text{Aut}(X^*)$, which is a compact topological group and thus comes with a natural probability measure μ (the normalized Haar measure). It is straightforward to show that

$$(1.2) \quad \mathcal{F}(G) = \mu(\{g \in G_\infty : g \text{ fixes at least one element of } X^\omega\}).$$

We use the notation G_∞ for the closure of G because the latter coincides with the inverse limit of the groups G_n . In the case where $G = \text{IMG}(f)$ we denote G_∞ by $\overline{\text{IMG}}(f)$; it is also the profinite completion of $\text{IMG}(f)$, and hence known as the *profinite iterated monodromy group* of f .

Following the terminology of [13, section 1.3], define $f \in \mathbb{C}[z]$ to be *exceptional* if there exists a finite, non-empty set Σ with $f^{-1}(\Sigma) \setminus \text{Crit } f = \Sigma$. Our main result is the following.

Theorem 1.1. *Let $f \in \mathbb{C}[z]$ be a post-critically finite polynomial of degree at least 2, with iterated monodromy group G . If f is not exceptional, then $\mathcal{F}(G) = 0$.*

Exceptional polynomials have appeared as a distinguished class in a variety of settings; for instance, the affine orbifold lamination attached to certain exceptional polynomials has an isolated leaf [10, Section 2] (see also [11–13] for special properties of such polynomials). It is not difficult to show that if f is exceptional, then

$\#\Sigma \leq 2$ (see p. 2046). If $\#\Sigma = 2$, then f is linearly conjugate to $\pm T_d$, where T_d is the Chebyshev polynomial of degree d (see Proposition 8.4). Note that linear conjugacy preserves the conjugacy class in $\text{Aut}(X^*)$ of $\text{IMG}(f)$. If $\#\Sigma = 1$, then f has a fixed point z_0 all of whose preimages are critical except z_0 itself, so f is conjugate to a polynomial of the form

$$(1.3) \quad z(z - a_1)^{k_1} \cdots (z - a_m)^{k_m},$$

where $a_i \in \mathbb{C} \setminus \{0\}$ and $k_i \geq 2$.

We also compute $\mathcal{F}(G)$ for exceptional polynomials with $\#\Sigma = 2$. Note that T_d is conjugate to $-T_d$ when d is even.

Proposition 1.2. *If $f(z) \in \mathbb{C}[z]$ is conjugate to T_d for d even, then $\mathcal{F}(\text{IMG}(f)) = 1/4$. If f is conjugate to $\pm T_d$ for d odd, then $\mathcal{F}(\text{IMG}(f)) = 1/2$.*

Thus the only post-critically finite polynomials f for which $\mathcal{F}(\text{IMG}(f))$ remains unknown are non-Chebyshev maps conjugate to a map of the form (1.3). We remark that the power maps $f(z) = z^d$ are not exceptional, and hence have $\mathcal{F}(\text{IMG}(f)) = 0$, in contrast to Chebyshev polynomials.

When f is quadratic, it must be conjugate to $z^2 + c$ for $c \in \mathbb{C}$. The only exceptional polynomial of this form is $f(z) = z^2 - 2$, since those of the form (1.3) have degree at least 3. Theorem 1.1 and Proposition 1.2 thus give the result that furnished the original motivation for this project:

Corollary 1.3. *Let $f(z) = z^2 + c \in \mathbb{C}[z]$ be post-critically finite, with iterated monodromy group G . Then $\mathcal{F}(G) = 0$ unless $f(z)$ is the Chebyshev polynomial $z^2 - 2$, in which case $\mathcal{F}(G) = 1/4$.*

To prove Theorem 1.1, we study groups of automorphisms of rooted trees, and draw heavily on a characterization due to V. Nekrashevych [15, Theorem 6.10.8] of which such groups are $\text{IMG}(f)$ for some post-critically finite f . Along the way we derive some results that apply more generally. For instance, define $g \in \text{Aut}(X^*)$ to be *spherically transitive* if it acts transitively on X^n for each $n \geq 1$. Every iterated monodromy group of a polynomial contains a spherically transitive element (see Theorem 5.8 and Lemma 6.2), which is furnished by monodromy at infinity. This element plays a crucial role in our analysis.

Theorem 1.4. *Suppose that $G \leq \text{Aut}(X^*)$ has a spherically transitive element. Then $\mu(\{g \in G_\infty : g \text{ fixes infinitely many elements of } X^\omega\}) = 0$.*

We prove Theorem 1.4 in Section 4, where we define a stochastic process encoding information about fixed-point-free elements of every G_n . The presence of a spherically transitive element implies this process is a martingale (Theorem 4.2), and we establish Theorem 1.4 using a basic martingale convergence theorem.

We give two other results that lead up to the proof of Theorem 1.1. A salient feature of X^* is its self-similarity, and we use this to describe elements of $\text{Aut}(X^*)$ recursively. Let $g \in \text{Aut}(X^*)$, and for a vertex $v \in X^*$ consider the subtrees vX^* and $g(v)X^*$ with root v and $g(v)$, respectively. Both are naturally isomorphic to X^* , and identifying them gives an automorphism $g|_v \in \text{Aut}(X^*)$, called the *restriction of g at v* . See Section 3 for examples and further definitions. We call G *contracting* if there is a finite set $\mathcal{N} \subset G$ such that for each $g \in G$, there is $n_g \geq 1$ such that all restrictions of g at words of length at least n_g belong to \mathcal{N} . Roughly, this property means that the action of g is relatively restrained, at least close to the

boundary of X^* . In particular, many computations in G can be reduced to finite considerations; see Section 3 for more details. It is known that iterated monodromy groups of post-critically finite polynomials are contracting [15, Theorems 3.9.12 and 6.10.8]. Let

$$\mathcal{N}_1 = \{g \in G : g|_v = g \text{ and } g(v) = v \text{ for some non-empty } v \in X^*\}.$$

When G is contracting, \mathcal{N}_1 is finite; see Proposition 3.5.

Theorem 1.5. *Suppose that $G \leq \text{Aut}(X^*)$ is contracting and has a spherically transitive element. If every $g \in \mathcal{N}_1$ fixes infinitely many ends of X^* , then $\mathcal{F}(G) = 0$.*

It is not hard to show that when G is contracting, \mathcal{N}_1 is torsion (see the end of Section 4), and this gives the following corollary.

Corollary 1.6. *Suppose that $G \leq \text{Aut}(X^*)$ is contracting and has a spherically transitive element. If G is torsion-free, then $\mathcal{F}(G) = 0$.*

The Basilica group B is known to be torsion-free [2, Proposition 3.11], and so Corollary 1.6 proves that $\mathcal{F}(B) = 0$.

Theorem 1.5 and Corollary 1.6 are proven using only the tools from Section 4, which do not use specific facts about iterated monodromy groups. On the other hand, in order to use Theorem 1.5 to prove Theorem 1.1, we apply a characterization of iterated monodromy groups of post-critically finite polynomials due to Nekrashevych [15, Theorem 6.10.8] (we give a restatement in Theorem 5.8). This gives a natural finite generating set A for $\text{IMG}(f)$. We introduce a *stability graph* associated to A , and use it to show that every element of \mathcal{N}_1 is conjugate to a power of an element of $\mathcal{N}_1 \cap A$ (Theorem 7.5). Thus we reduce questions about fixed points of the action of elements of \mathcal{N}_1 on X^* to the study of the action of elements of A on X^* , and these are directly related to the orbits of the critical points of f (Theorem 5.1). In Section 8 we use a strong property of A given in Theorem 5.8 to show that only very special configurations of A allow for elements of \mathcal{N}_1 to fix a finite number of ends of X^* .

While most of the proofs here are group-theoretic, the consequences are of interest to number theorists, and hence we have made an effort to make the exposition relatively self-contained. After giving more details on links between our results and number theory (Section 2), we give in Section 3 just the background necessary to prove Theorem 1.4, Theorem 1.5, and Corollary 1.6. These proofs are in Section 4. For the remainder of the paper, more background is required, which we describe in Section 5. Sections 6, 7, and 8 contain the rest of the proofs.

2. CONNECTIONS TO NUMBER THEORY AND ARITHMETIC DYNAMICS

We give here some links between iterated monodromy groups and density questions for sets of dynamical interest in arithmetic contexts.

2.1. Iterated monodromy groups and Galois groups. Let $f \in \mathbb{C}[z]$ be post-critically finite. The action of $\text{IMG}(f)$ on the set $f^{-n}(\beta)$, where β is outside the post-critical set, is given by monodromy, and we refer to this quotient as G_n . Recall that the inverse limit of the G_n is by definition the profinite iterated monodromy group $\overline{\text{IMG}}(f)$. On the other hand, the Galois group $\tilde{G}_n := \text{Gal}(K_n/\mathbb{C}(t))$, where K_n is the splitting field of the polynomial $f^n(z) - t \in \mathbb{C}(t)[z]$, has a natural action on the d^n roots of $f^n(z) - t$ over $\mathbb{C}(t)$. It is well known (see e.g. [5, Theorem 8.12])

that $G_n \cong \tilde{G}_n$, and the corresponding actions on $f^{-n}(\beta) \subset \mathbb{C}$ and $f^{-n}(t) \subset \overline{\mathbb{C}(t)}$ are conjugate subgroups of S_{d^n} . We summarize this in the following proposition, which is essentially [15, Proposition 6.4.2].

Proposition 2.1. *The profinite iterated monodromy group $\overline{\text{IMG}}(f)$ is isomorphic to the Galois group of K_∞ over $\mathbb{C}(t)$, where $K_\infty = \bigcup_{n \geq 1} K_n$. Moreover, the corresponding actions on the preimage trees $T_\beta \subset \mathbb{C}$ and $T_t \subset \overline{\mathbb{C}(t)}$ are conjugate.*

Proposition 2.1 prompted the introduction of iterated monodromy groups, as a tool for computing the group \tilde{G}_∞ [15, p. 174]. In the remainder of this section, we give connections of iterated monodromy groups to arithmetic problems, which proceed via the link to Galois theory in Proposition 2.1.

2.2. Density problems: Global fields. Let k be a perfect field of characteristic $p \geq 0$, let \bar{k} be an algebraically closed field containing k , and suppose that $f \in k[x]$ has degree $d \geq 2$ that is prime to p . For the moment we do not assume that f is post-critically finite. Let K_n be the splitting field of $f^n(x) - t$ over $k(t)$, and put $K_\infty = \bigcup_n K_n$. For $n \leq \infty$, define the *arithmetic monodromy group* A_n to be the Galois group of K_n over $k(t)$. The field of constants of K_n is $\bar{k} \cap K_n$, which we denote by k'_n . The *geometric monodromy group* G_n is the normal subgroup of A_n whose elements restrict to the identity on k'_n . Clearly A_n/G_n is isomorphic to the Galois group of k'_n/k , and hence we have an exact sequence

$$(2.1) \quad 1 \rightarrow G_n \rightarrow A_n \rightarrow \text{Gal}(k'_n/k) \rightarrow 1$$

for each $n \leq \infty$. We may also take a specialization $t = t_0 \in k$, thereby obtaining a specialized form of (2.1):

$$(2.2) \quad 1 \rightarrow G_n(t_0) \rightarrow A_n(t_0) \rightarrow \text{Gal}(k'_n/k) \rightarrow 1.$$

Note that the extension k'_n/k of constants is independent of specialization. For $n \leq \infty$, the groups A_n and G_n both act naturally on the set R_n of roots of $f^n(x) - t$ over $k(t)$, while $A_n(t_0)$ and $G_n(t_0)$ act on the set $R_n(t_0)$ of roots of $f^n(x) - t_0$ over k . We thus define

$$(2.3) \quad \mathcal{F}(A_\infty) = \lim_{n \rightarrow \infty} \frac{\#\{g \in A_n : g \text{ fixes at least one root of } f^n(x) - t\}}{\#A_n},$$

with similar definitions for $\mathcal{F}(G_\infty)$, $\mathcal{F}(A_\infty(t_0))$, and $\mathcal{F}(G_\infty(t_0))$. We remark that the set $\bigcup_{n \geq 0} R_n$ has a natural structure as a complete $(\deg f)$ -ary rooted tree, while the same is true of $\bigcup_{n \geq 0} R_n(t_0)$ provided that $f^n(x) - t_0$ is separable for all $n \geq 1$ (or equivalently, there are no critical points of f mapping to t_0 under any iterate of f). We can thus identify $\bigcup_{n \geq 0} R_n$ with X^* , and $A_n, G_n, A_n(t_0), G_n(t_0)$ with subgroups of $\text{Aut}(X^*)$. With this identification, (2.3) is the same as (1.1).

If $f \in \mathbb{C}[x]$ is a post-critically finite polynomial, its coefficients must satisfy algebraic relations imposed by the self-intersections of the orbits of the critical points, and hence f is defined over a finite extension k of \mathbb{Q} . We may take $\bar{k} = \mathbb{C}$, and then (2.1) with $n = \infty$ becomes

$$(2.4) \quad 1 \rightarrow \overline{\text{IMG}}(f) \rightarrow A_\infty \rightarrow \text{Gal}(k'_\infty/k) \rightarrow 1.$$

Let k be a global field, that is, a finite extension of \mathbb{Q} or a finite extension of the function field $\mathbb{F}_q(t)$ of \mathbb{P}^1 over the finite field with q elements, and we take the ring of integers O_k to be the integral closure in k of \mathbb{Q} or $\mathbb{F}_q[t]$. We wish to have a notion of size for a set of prime ideals in O_k .

Definition 2.2. Let k be a global field and \mathcal{P} be a set of primes in O_k . The Dirichlet density of \mathcal{P} is

$$\delta(\mathcal{P}) = \limsup_{s \rightarrow 1^+} \frac{\sum_{\mathfrak{p} \in \mathcal{P}} N\mathfrak{p}^{-s}}{\sum_{\mathfrak{p} \subset O_k} N\mathfrak{p}^{-s}},$$

where $N\mathfrak{p}$ is the number of elements in the field $O_k/\mathfrak{p}O_k$.

The Chebotarev density theorem allows one to relate the Dirichlet density of various naturally-occurring sets of primes in O_k to group-theoretic properties of the Galois groups of certain extensions of k . The following theorem is an instance of this.

Theorem 2.3 (Theorem 2.1, [9]). *Let k be a number field with ring of integers O_k , and let $f \in O_k[x]$ with $a_0, t_0 \in O_k$. Let $\mathcal{P}(f, t_0)$ be the set of primes dividing at least one element of the sequence $f^n(a_0) - t_0, n \geq 1$. Suppose that $f^n(x) - t_0$ is separable for all $n \geq 1$, and let $A_\infty(t_0)$ be as in (2.2). Then*

$$\delta(\mathcal{P}(f, t_0)) \leq \mathcal{F}(A_\infty(t_0)).$$

Note that the conclusion is independent of the choice of a_0 . In the case of number fields as above, we may also replace δ with natural density, namely

$$D(\mathcal{P}) = \lim_{x \rightarrow \infty} \frac{\#\{\mathfrak{p} \in \mathcal{P} : N\mathfrak{p} \leq x\}}{\#\{\mathfrak{p} \in O_k : N\mathfrak{p} \leq x\}},$$

and part of the conclusion of the theorem is that this limit exists.

Theorem 2.3 says that $\mathcal{F}(A_\infty)$ gives the “generic” value of the density of prime divisors of an orbit of f translated by a constant t_0 . Indeed, if $\mathcal{F}(A_\infty) = 0$, then one can use the Hilbert irreducibility theorem to show that for any $\epsilon > 0$, $\mathcal{P}(f, t_0) < \epsilon$ for all but a thin set of t_0 . In the case where f is post-critically finite, we have the exact sequence (2.4), and in light of Theorem 1.1, one needs to study the extension of constant fields k'_∞/k and understand how it interacts with $\overline{\text{IMG}}(f)$. Indeed, if k'_∞ is a finite extension of k , one could replace the ground field k by k'_∞ and obtain the desired result. However, it seems unlikely that this is the case in most circumstances. For instance, when $f(x) = x^2$ and $k = \mathbb{Q}$, we have that $k'_\infty = \mathbb{Q}_{\zeta_{2^\infty}}$.

2.3. Density problems over finite fields. Let \mathbb{F}_q be the finite field with q elements, let $f \in \mathbb{F}_q[x]$, and let $\alpha \in \overline{\mathbb{F}_q}$. Clearly the forward orbit $\{f^n(\alpha) : n \geq 0\}$ of any such α is contained in a finite extension of \mathbb{F}_q , whence it must be finite. We thus have two fundamental behaviors: if there is a $j \geq 1$ with $f^j(\alpha) = \alpha$ we call α *purely periodic* under f , while if there is no such j , then we call α *pre-periodic* under f . Let $\text{Per}(f)$ be the set of purely periodic points for f in $\overline{\mathbb{F}_q}$. Note that by construction f must be post-critically finite, since all its orbits are finite. Define the Dirichlet density of a set $S \subseteq \overline{\mathbb{F}_q}$ to be

$$(2.5) \quad \delta(S) = \limsup_{s \rightarrow 1^+} \frac{\sum_{\alpha \in S} (\deg \alpha)^{-1} N(\alpha)^{-s}}{\sum_{\alpha \in \overline{\mathbb{F}_q}} (\deg \alpha)^{-1} N(\alpha)^{-s}},$$

where $\deg \alpha = [\mathbb{F}_q(\alpha) : \mathbb{F}_q]$, and $N(\alpha) = q^{\deg \alpha}$. This is essentially identical to Definition 2.2; the $(\deg \alpha)$ term is necessary because there are $\deg \alpha$ conjugates of α corresponding to the prime of $\mathbb{F}_q[t]$ with root α .

We sketch an argument showing how $\delta(\text{Per}(f))$ is given by statistics of an arithmetic iterated monodromy group as in (2.1), where $k = \mathbb{F}_q$. Note that $\alpha \in \text{Per}(f)$ if

and only if some infinite branch of the tree of preimages $\bigsqcup_{n \geq 0} f^{-n}(\alpha)$ is contained in the base field \mathbb{F}_q . Let \mathfrak{p} be the prime ideal generated by the minimal polynomial of α over \mathbb{F}_q . Then a branch of $\bigsqcup_{n \geq 0} f^{-n}(\alpha)$ is contained in \mathbb{F}_q if and only if $\text{Frob}_{\mathfrak{p}} \in A_n$ fixes a root of $f^n(x) - t$ for each $n \geq 1$ (denote by \mathcal{P} the set of such \mathfrak{p}). Here $\text{Frob}_{\mathfrak{p}}$ is the conjugacy class of elements of A_n that act on the residue class field $O_{K_n}/\mathfrak{p}O_{K_n}$ as $x \mapsto x^q$. The Chebotarev density theorem for function fields [17, Theorem 9.13A] then gives that the Dirichlet density of \mathcal{P} is bounded above, for each $n \geq 1$, by the proportion of $g \in A_n$ that fix at least one root of $f^n(x) - t$. Thus this density is bounded above by $\mathcal{F}(A_\infty)$. It is then straightforward to show that this implies $\delta(\text{Per}(f)) \leq \mathcal{F}(A_\infty)$.

3. BACKGROUND AND EXAMPLES, PART I

3.1. Wreath recursion and spherically transitive elements. In this section we give the background required to prove the results in Section 4. We draw on the exposition in [15, Chapter 1], including following the convention there of writing group actions on the left. From now on we suppose that our alphabet X is given by $\{0, \dots, d - 1\}$, and we let S_d denote the symmetric group on d letters. Then there is a natural isomorphism

$$\psi : \text{Aut}(X^*) \rightarrow S_d \wr \text{Aut}(X^*),$$

where \wr denotes the wreath product, that takes g to $(\sigma, (g|_0, \dots, g|_{d-1}))$, where $\sigma \in S_d$ is the action of g on X (i.e., on the first level of X^*). In other words, we may describe g by specifying its restriction at each element of X and its action on X . We call this the *wreath recursion* describing g . We generally drop the parentheses and equate g with its image under ψ , writing

$$(3.1) \quad g = \sigma(g|_0, \dots, g|_{d-1}).$$

We write the identity element as 1, and when the permutation σ is the identity, we omit it. Hence the identity element of $\text{Aut}(X^*)$ is given in wreath recursion by $(1, 1, \dots, 1)$. Note that the element $a = (a, 1, 1, \dots, 1)$ is also the identity, since by induction it acts trivially on X^n for all n , and thus acts trivially on X^* . Given $g = \sigma(g|_0, \dots, g|_{d-1})$, we can make explicit its action on any X^n thanks to the following formulas, which are straightforward to prove:

$$(3.2) \quad g|_{vw} = g|_v|_w \quad g(vw) = g(v)g|_v(w),$$

for any $v, w \in X^*$.

One can multiply elements in wreath recursion form using the normal multiplication in a semi-direct product:

$$(3.3) \quad \sigma(g_0, \dots, g_{d-1}) \cdot \tau(h_0, \dots, h_{d-1}) = \sigma\tau(g_{\tau(0)}h_0, \dots, g_{\tau(d-1)}h_{d-1}),$$

where $g_i = g|_i$ and $h_i = h|_i$. If we take $v \in X^*$ of length n , we may consider (3.3) as giving the wreath recursion of $g, h \in \text{Aut}(X^*)$ acting on X^n . This gives

$$(3.4) \quad (gh)(v) = g(h(v)) \quad \text{and} \quad (gh)|_v = g|_{h(v)} \cdot h|_v.$$

Example 3.1. Let $d = 2$ and take σ to be the non-trivial element of S_2 . Let $a = (a, b)$, $b = \sigma(1, 1)$ and $G = \langle a, b \rangle$. From (3.3), we have $b^2 = \sigma^2(1, 1) = 1$ and $a^2 = (a^2, b^2) = (a^2, 1)$. By induction this gives $a^2 = 1$. However, the element $ba = \sigma(b, a)$ is spherically transitive, i.e., acts on each X^n as a 2^n -cycle, and in particular has infinite order. This is a consequence of Proposition 3.3. In Section 5

we show that G is isomorphic to the iterated monodromy group of the polynomial $z^2 - 2$, which is conjugate to the Chebyshev polynomial $2z^2 - 1$.

Example 3.2. Let $d = 2$ and take σ to be the non-trivial element of S_2 . Let $a = \sigma(1, b)$, $b = (1, a)$ and $G = \langle a, b \rangle$. This is the Basilica group, mentioned on page 2024. If we write $X^2 = \{00, 01, 10, 11\}$, then from (3.2), the wreath recursion for a acting on X^2 is $\tau(1, 1, 1, a)$, where $\tau = (00, 10)(01, 11)$. Hence from (3.3), a^2 acts on X^2 as $(1, a, 1, a)$. It follows that the restrictions of a^n to words of length 2 are all of the form a^k for $k < n$. If a is torsion of order n , then all restrictions of a^n are trivial, and so $a^k = 1$ for some $k < n$, a contradiction. Hence a has infinite order, though it is not spherically transitive.

As an illustration of the preceding ideas, we give a characterization of spherically transitive elements of $\text{Aut}(X^*)$. The proof is left as an exercise.

Proposition 3.3. *Let X have d elements and $g \in \text{Aut}(X^*)$. For each $v \in X^{n-1}$, let $\gamma_v \in S_d$ denote the action of $g|_v$ on X , suppose that γ_v commutes with $\gamma_{v'}$ for all $v, v' \in X^{n-1}$, and let $\rho_n = \prod_{v \in X^{n-1}} \gamma_v$. Then g is spherically transitive if and only if ρ_n is a d -cycle for every $n \geq 1$.*

Remark 3.4. Note that by convention $X^0 = \emptyset$, and $\gamma_\emptyset = \rho_1$ is the action of g on X . The commutativity hypothesis is automatically satisfied in the case $d = 2$, and in this setting ρ_n is the identity precisely when the number of $v \in X^{n-1}$ with $\gamma_v \neq 1$ is even. Thus Proposition 3.3 says that for $d = 2$, g is spherically transitive when $\gamma_v \neq 1$ for an odd number of $v \in X^{n-1}$, for all $n \geq 1$. For the element $ba = \sigma(b, a)$ in Example 3.1, it is easy to see that $\gamma_v \neq 1$ for only one v in each X^{n-1} .

3.2. Self-similar and contracting groups. A group $G \leq \text{Aut}(X^*)$ is *self-similar* if $g|_x \in G$ for all $g \in G$ and $x \in X$. We call G *contracting* if there is a finite set $\mathcal{N} \subset G$ such that for every $g \in G$, $g|_v \in \mathcal{N}$ for all $v \in X^*$ sufficiently long. The smallest set satisfying this condition is called the *nucleus* of the group. In a contracting group, one can reduce many computations to considerations involving only a finite set. For instance, as pointed out in [1], solving the so-called word problem (determining whether a given product of n generators is trivial) can be done in polynomial time in a contracting group.

We now consider the set of *stable* elements of G ,

$$\mathcal{N}_0 = \{g \in G : g|_v = g \text{ for some non-empty } v \in X^*\}.$$

Proposition 3.5. *If $G \leq \text{Aut}(X^*)$ is contracting, then \mathcal{N}_0 is finite and the nucleus of G is equal to*

$$(3.5) \quad \{h \in G : h = g|_w \text{ for some } g \in \mathcal{N}_0, w \in X^*\}.$$

Proof. By definition, the nucleus of G consists of the elements $g \in G$ for which there exists $r \in G$ with $r|_w = g$ for arbitrarily long words w . If $g|_v = g$ for some non-empty v and v^n is the n -fold concatenation of v with itself, then from (3.2) we have $g|_{v^n} = g$ for all $n \geq 1$. Moreover, any h with $g|_w = h$ for some $w \in X^*$ must also occur as the restriction of g at arbitrarily long words. Hence the set in (3.5) is contained in the nucleus, and in particular \mathcal{N}_0 is finite. On the other hand, if h is in the nucleus, let $r \in G$ with $r|_w = h$ for arbitrarily long words w . Let n_1 be the size of the nucleus and n_2 be such that $r|_u$ is in the nucleus when u has length at least n_2 . We may take the length of w to exceed $n_1 + n_2$. Hence if w_k is the length- k initial word of w , then $r|_{w_k}$ is in the nucleus for more than n_1 values of k ,

and hence $r|_{w_k} = r|_{w_j}$ for some $k < j$. Therefore $r|_{w_k} \in \mathcal{N}_0$ and there is a word w' with $r|_{w_k|_{w'}} = r|_w = h$. \square

It is known that standard actions on X^* of iterated monodromy groups of post-critically finite polynomials are always contracting [15, Theorem 6.4.4], and Proposition 7.1 gives a method for computing \mathcal{N}_0 for a class of groups including iterated monodromy groups. For the group from Example 3.1, we have that $\mathcal{N}_0 = \{1, a\}$, and hence G has nucleus $\{1, a, b\}$. For the Basilica group (Example 3.2), we have $\mathcal{N}_0 = \{1, a, b, a^{-1}, b^{-1}, ba^{-1}, ab^{-1}\}$ (see the remark following Proposition 7.1), and in this case \mathcal{N}_0 coincides with the nucleus.

4. THE FIXED-POINT PROCESS

As noted in the introduction, the profinite completion G_∞ of G with respect to the G_n comes equipped with a natural probability measure that projects to the discrete uniform measure on each G_n . In this section we define a stochastic process – that is, an infinite collection of random variables defined on a common probability space – that encodes information about the number of fixed points in X^n of elements of G_n . We then adapt techniques of [8] to show that this process is a martingale provided that G contains a spherically transitive element. Finally, we apply a martingale convergence theorem that leads to the proofs of Theorem 1.4, Theorem 1.5, and Corollary 1.6.

Given $g \in G$ where the group G acts naturally on a set S , we denote by $\text{Fix}(g)$ the set of $s \in S$ with $g(s) = s$. Define a stochastic process Y_1, Y_2, \dots on G_∞ by taking $Y_n(g) = \#\text{Fix}(\pi_n(g))$, where π_n is the natural projection $G_\infty \rightarrow G_n$ and G_n acts on X^n . We call this the *fixed-point process* of G , and write it $FP(G)$. Because $\mu(\pi_n^{-1}(T)) = \#T/\#G_n$ for any $T \subseteq G_n$, we have that $\mu(Y_1 = t_1, \dots, Y_n = t_n)$ is given by

$$(4.1) \quad \frac{1}{\#G_n} \# \{g \in G_n : g \text{ fixes } t_i \text{ elements of } X^i \text{ for } i = 1, 2, \dots, n\}.$$

We denote by $E(Y)$ the expected value of the random variable Y .

Definition 4.1. A stochastic process with probability measure μ and random variables Y_1, Y_2, \dots taking values in \mathbb{R} is a *martingale* if for all $n \geq 2$ and any $t_i \in \mathbb{R}$,

$$E(Y_n \mid Y_1 = t_1, Y_2 = t_2, \dots, Y_{n-1} = t_{n-1}) = t_{n-1},$$

provided $\mu(Y_1 = t_1, Y_2 = t_2, \dots, Y_{n-1} = t_{n-1}) > 0$.

Theorem 4.2. *Let $G \leq \text{Aut}(X^*)$ have a spherically transitive element. Then $FP(G)$ is a martingale.*

Proof. We must show that

$$(4.2) \quad E(Y_n \mid Y_1 = t_1, \dots, Y_{n-1} = t_{n-1}) = t_{n-1},$$

where t_1, \dots, t_{n-1} satisfy $\mu(Y_1 = t_1, \dots, Y_{n-1} = t_{n-1}) > 0$. Because the Y_i take integer values, each t_i must be an integer. By definition, the left-hand side of (4.2) is

$$(4.3) \quad \sum_k k \cdot \frac{\mu(Y_1 = t_1, \dots, Y_{n-1} = t_{n-1}, Y_n = k)}{\mu(Y_1 = t_1, \dots, Y_{n-1} = t_{n-1})}.$$

Put

$$\begin{aligned} S &= \{g \in G_n : g \text{ fixes } t_i \text{ elements of } X^i \text{ for } 1 \leq i \leq n-1\}, \\ S_k &= \{g \in S : g \text{ fixes } k \text{ elements of } X^n\}. \end{aligned}$$

By (4.1), the expression in (4.3) is equal to $\sum_k k \cdot (\#S_k/\#S)$. This in turn may be rewritten

$$(4.4) \quad \frac{1}{\#S} \sum_{g \in S} \#\text{Fix}(g).$$

Let $\sigma \in G_n$ be the image under π_n of the spherically transitive element of G_∞ assumed to exist. Then $\tau = \sigma^{d^{n-1}}$ acts trivially on X^{n-1} , and hence S is invariant under multiplication by powers of τ , and therefore is a disjoint union of cosets of $\langle \tau \rangle$. Note that because $\langle \sigma \rangle$ acts transitively on X^n , $\langle \tau \rangle$ must act transitively on each set $v* = \{vx : x \in X\}$ for $v \in X^{n-1}$.

Now take $g\langle \tau \rangle \subseteq S$, and let $R = \{vx : v \in X^{n-1}, g(v) = v, x \in X\}$. Note that because $g \in S$, we have $\#R = dt_{n-1}$. If $vx \in R$, then $g(vy) = vx$ for some unique $y \in X$. There is a unique $i \in \{0, \dots, d-1\}$ such that $\tau^i(vx) = vy$, and thus $g\tau^i(vx) = vx$. If $I(g, s)$ is the function that takes the value 1 when $g(s) = s$ and 0 otherwise, we have shown that $\sum_{i=0}^{d-1} I(g\tau^i, vx) = 1$ and hence

$$\sum_{vx \in R} \sum_{i=0}^{d-1} I(g\tau^i, vx) = dt_{n-1}.$$

Inverting the order of summation and using that $g(w) \neq w$ for $w \notin R$, we have

$$\sum_{i=0}^{d-1} \#\text{Fix}(g\tau^i) = dt_{n-1}.$$

But S is the disjoint union of cosets of $\langle \tau \rangle$, and hence

$$\sum_{g \in S} \#\text{Fix}(g) = \#S \cdot t_{n-1}.$$

Therefore the expression in (4.4) equals t_{n-1} . □

Martingales are useful tools because they often converge in the following sense:

Definition 4.3. Let Y_1, Y_2, \dots be a stochastic process defined on the probability space Ω with probability measure μ . The process *converges* if

$$\mu \left(\omega \in \Omega : \lim_{n \rightarrow \infty} Y_n(\omega) \text{ exists} \right) = 1.$$

We give one standard martingale convergence theorem (see e.g. [7, Section 12.3] for a proof).

Theorem 4.4. *Let $M = (Y_1, Y_2, \dots)$ be a martingale whose random variables take non-negative real values. Then M converges.*

Since the random variables in $FP(G)$ take non-negative integer values, we immediately have the following:

Corollary 4.5. *Let $G \leq \text{Aut}(X^*)$ contain a spherically transitive element. Then*

$$\mu(\{g \in G_\infty : Y_1(g), Y_2(g), \dots \text{ is eventually constant}\}) = 1.$$

In particular, any $g \in G_\infty$ fixing infinitely many ends of X^* must have $Y_i(g) \rightarrow \infty$, and hence lie in a set of measure zero. This proves Theorem 1.4.

We may now give a short proof of Theorem 1.5. Assume the hypotheses of that theorem, and let \mathcal{N} be the nucleus of G . Suppose that $g \in G_\infty$ fixes some end $w = x_1x_2 \cdots$ of X^* . Let $v_n = x_1x_2 \cdots x_n$ for each $n \geq 1$, and consider the sequence of restrictions $g|_{v_1}, g|_{v_2}, \dots$. For n large enough, we have $g|_{v_n} \in \mathcal{N}$, and $g|_{v_n}$ fixes the end $x_{n+1}x_{n+2} \cdots$ since g fixes w . Because \mathcal{N} is finite, there must be $i < j$ with $g|_{v_i} = g|_{v_j}$. Let $h = g|_{v_i}$, and note that for $w = x_{i+1}x_{i+2} \cdots x_j$ we have $h(w) = w$ and $h|_w = h$. Hence $h \in \mathcal{N}_1$, and by hypothesis fixes infinitely many ends of X^* . Inserting v_i on the beginning of each of these ends, we obtain infinitely many ends of X^* fixed by g . Hence by Theorem 1.4, g lies in a set of measure zero, proving the theorem.

To derive Corollary 1.6, note that if $g \in \mathcal{N}_1$, then $g(v) = v$ and $g|_v = g$ for some non-empty $v \in X^*$. From (3.4) it follows that $g^n(v) = v$ and $g^n|_v = g^n$ for all $n \geq 1$, and hence $g^n \in \mathcal{N}_1 \subseteq \mathcal{N}_0$ for all $n \geq 1$. Because G is contracting, \mathcal{N}_0 is finite by Proposition 3.5, and thus two distinct powers of g are equal, implying that g is torsion. Therefore if G is torsion-free, then \mathcal{N}_1 is trivial, and Corollary 1.6 follows from Theorem 1.5.

5. BACKGROUND AND EXAMPLES, PART II

5.1. Computations of iterated monodromy groups. Recall from Section 2.1 that if f is a post-critically finite polynomial with post-critical set P_f , then $\text{IMG}(f)$ acts naturally on the tree $T_\beta = \bigsqcup_{n \geq 1} f^{-n}(\beta)$ of preimages of any $\beta \in \mathbb{C} \setminus P_f$. If f has degree d , then we may take $X = \{0, 1, \dots, d - 1\}$, choose a bijection $\Lambda : X \rightarrow f^{-1}(\beta)$, and for each $x \in X$, fix a path starting at β and ending in $\Lambda(x)$. This data allows one to extend Λ to an isomorphism $\Lambda : X^* \rightarrow T$ [15, Proposition 5.2.1] that conjugates the action of $\text{IMG}(f)$ to that of some $G \leq \text{Aut}(X^*)$ on X^* [15, Proposition 5.2.2]. We call this a *standard action* of $\text{IMG}(f)$ on X^* . Any two standard actions yield conjugate subgroups of $\text{Aut}(X^*)$. A standard action has the advantage of furnishing a method for computing the action on X^* of elements of $\text{IMG}(f)$ via a wreath recursion [15, Proposition 5.2.2] (see also [1, Proposition 2.2]). In [15, Definition 6.8.1, Theorem 6.8.3] a specific standard action is constructed, in which G is generated by $\{g_z : z \in P_f\}$. For given $z \in P_f$, the action of g_z on X is the monodromy action of a small loop going around z in the positive direction, and the restrictions of g_z are either trivial or g_y , where $y \in P_f$ and $f(y) = z$. A more specific description is given in the following theorem.

Theorem 5.1. *Let $f \in \mathbb{C}[z]$ be a post-critically finite polynomial, with post-critical set P_f . Any standard action of $\text{IMG}(f)$ on X^* is generated by a set $\{g_z : z \in P_f\}$ with the following properties:*

- (1) *The action of g_z on X contains one m -cycle for each $y \in f^{-1}(z)$ with $\text{ord}_f(y) = m$, where $\text{ord}_f(y)$ denotes the order of vanishing of f at y .*
- (2) *Let (x_1, x_2, \dots, x_m) be the cycle corresponding to y in part (1), and let $g = g_z$. If $y \notin P_f$, then $g|_{x_i} = 1$ for each $i = 1, \dots, m$. If $y \in P_f$, then there is a unique $i \in \{1, \dots, m\}$ such that $g|_{x_i} = g_y$, and $g|_{x_i} = 1$ otherwise.*

Remark 5.2. While [15, Theorem 6.8.3] gives a specific standard action of $\text{IMG}(f)$ on X^* , Theorem 5.1 is stated for any standard action. This is because the statements in Theorem 5.1 are all invariant when G is conjugated by an element of $\text{Aut}(X^*)$, and any two standard actions are conjugate.

Remark 5.3. Although we don't regard ∞ as being in P_f , the conclusion of Theorem 5.1 nonetheless applies to it. Because $f^{-1}(\infty) = \{\infty\}$, it is a point of multiplicity d , and we have that g_∞ acts as a d -cycle on X , with restriction to some $x \in X$ giving g_∞ and the other restrictions being trivial. Hence for each $n \geq 1$, there is a unique $v \in X^n$ with $g_\infty|_v = g_\infty$ and $g_\infty|_{v'} = 1$ for all $v' \in X^n$ with $v' \neq v$. Proposition 3.3 then shows that g_∞ is spherically transitive. The fact that G contains a spherically transitive element is also a consequence of Theorem 5.8 and Lemma 6.2.

As an illustration of this result, we show that the group in Example 3.1 is a standard action of the iterated monodromy group of $f(z) = z^2 - 2$ on X^* , where $X = \{0, 1\}$. We have $P_f = \{-2, 2\}$, so that $G = \langle g_{-2}, g_2 \rangle$. Now $f^{-1}(-2) = \{0\}$ and $\text{ord}_f(0) = 2$, implying that g_{-2} acts on X as a 2-cycle. Because $0 \notin P_f$, the restrictions of g_{-2} are trivial. On the other hand $f^{-1}(2) = \{-2, 2\}$, so g_2 acts trivially on X . Because $-2 \in P_f$ but $2 \notin P_f$, the restriction of g_2 to one element of X is trivial, while the other one is g_{-2} . Either choice gives the same group, since conjugating by g_{-2} exchanges the restrictions of g_2 .

5.2. Automata and Moore diagrams. A very useful description of $g \in \text{Aut}(X^*)$ in terms of its wreath recursion comes via automata theory. The set $Q(g) = \{g|_v : v \in X^*\}$ of all restrictions of g may be viewed as the set of states of an automaton. Being in a state $g|_w$ for some $w \in X^n$ and receiving an input letter $x \in X$, the automaton types on the output tape $g|_w(x)$ and proceeds to the state $(g|_w)|_x$, which by (3.2) is just $g|_{wx}$. In this way the action of g on any $v \in X^*$ may be determined. We formalize this in the following definition:

Definition 5.4. An *automaton* A over the set X is given by

- the set of states, which we denote also by A ;
- a map $\tau : A \times X \rightarrow X \times A$.

If $\tau(a, x) = (y, b)$, then y and b as functions of (a, x) are called the *output* and *transition function*, respectively. We say that A is *invertible* if each $a \in A$ acts on X as a permutation.

The *Moore diagram* of an automaton A provides a good method of visualization. It is a directed labeled graph whose vertex set is the set A of states of the automaton. If $\tau(a, x) = (y, b)$, then there is an arrow from a to b labeled by (x, y) . If A is invertible, the Moore diagram of the inverse automaton is given by formally replacing each state a by a^{-1} and changing each arrow labeling from (x, y) to (y, x) . Given an automaton A over a set X , it is easy to see that the states of A define elements of $\text{Aut}(X^*)$. Indeed, we can recover the wreath recursion for $a \in A$ by noting that if $\tau(a, x) = (y, b)$, then $a(x) = y$ and $a|_x = b$. In this case we say that $G = \langle A \rangle$ is generated by the automaton A .

By Theorem 5.1, a standard action of the iterated monodromy group of a post-critically finite polynomial is generated by a set that is closed under restrictions. Hence the automaton generating such a group is finite. See Figure 1 for an example.

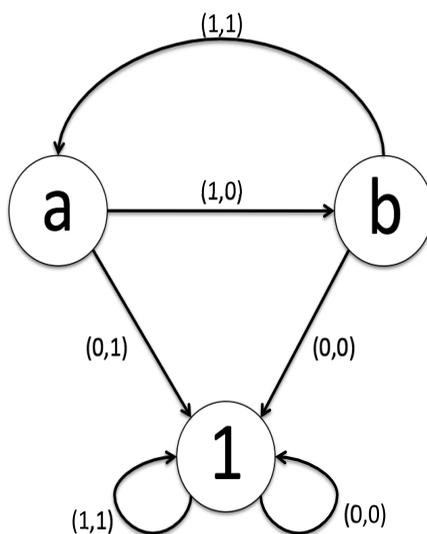


FIGURE 1. Moore diagram of the automaton generating the Basilica group; see Example 3.2. The state labeled 1 is the identity state, and is distinct from the element $1 \in X$.

5.3. Bounded and finitary automorphisms.

Definition 5.5. We say that $g \in \text{Aut}(X^*)$ is *finite-state* if it is defined by a finite automaton, or equivalently if $\{g|_v : v \in X^*\}$ is a finite set. We call g *bounded* if it is finite-state and the sequence

$$q_n = \#\{v \in X^n : g|_v \neq 1\}$$

is bounded. We call g *finitary* if $q_n = 0$ for all n sufficiently large, or equivalently if there exists n_0 such that $g|_v$ is trivial for all words of length at least n_0 .

Finitary automorphisms will play a major role in Sections 7 and 8. The main fact we will use about the more general notion of bounded automorphisms is the following special case of a theorem of Nekrashevych and Bondarenko:

Theorem 5.6 ([4], [15, Theorem 3.9.12]). *Let $G \leq \text{Aut}(X^*)$ be generated by a finite automaton whose states define bounded automorphisms of X^* . Then G is contracting.*

5.4. Kneading automata and theorem of Nekrashevych. We require a strong result of Nekrashevych that characterizes the $G \leq \text{Aut}(X^*)$ that are isomorphic to a standard action of the iterated monodromy group of a post-critically finite polynomial. This characterization is purely in terms of a finite automaton that generates G . To state this result, we require the notion of a *tree-like multi-set of permutations*. Recall that a multi-set of permutations of a set X is a map $i \mapsto \pi_i$ from a set I of indices to the set $\text{Sym}(X)$ of permutations of X . Thus for instance distinct indices may give the same permutation. We denote the set $\{\pi_i : i \in I\}$ by T . The *cycle diagram* associated to T is an oriented 2-dimensional CW-complex whose set of 0-cells is X . For each cycle (x_1, x_2, \dots, x_n) of each $\pi_i \in T$, there is a 2-cell whose boundary passes through x_1, x_2, \dots, x_n and no other elements of

X , and whose order on the boundary corresponds to the order in the cycle. Two different 2-cells can only intersect at 0-cells. We call the *reduced cycle diagram* of T the diagram obtained by deleting the 2-cells corresponding to fixed points of the π_i .

Definition 5.7. A multi-set T of permutations of a set X is said to be *tree-like* if the cycle diagram of T is contractible.

For an example of a tree-like multi-set, see Figure 2. Note that we could add the identity to this multi-set any number of times and it would still be tree-like. However, adding any non-trivial element of S_6 would yield a non-tree-like multi-set.

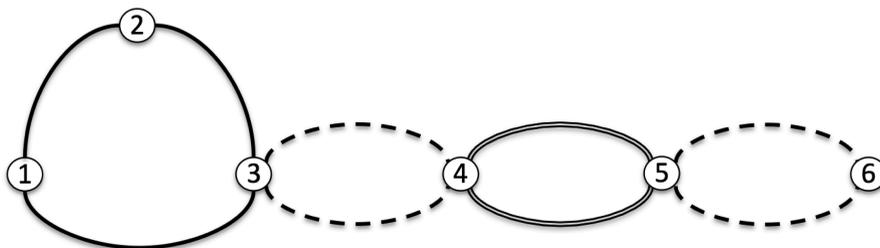


FIGURE 2. The reduced cycle diagram of the multi-set $\{(1, 2, 3), (3, 4)(5, 6), (4, 5)\}$ of elements of S_6 . The arrows are omitted, and the action of the elements are given by the solid, dashed, and doubled lines, respectively. This multi-set is tree-like.

Another way to visualize the action of a multi-set of permutations T on a set X is via its *cycle graph*. We define it to be a bipartite graph obtained from the reduced cycle diagram by coloring each vertex of the former white, and replacing each 2-cell by a black vertex connected to the white vertices on the boundary of the 2-cell. See Figure 3 for the cycle graph corresponding to the multi-set from Figure 2. Note that our definition differs slightly from that of [15, p. 186], where the cycle graph is not defined to be bipartite, but is otherwise identical.

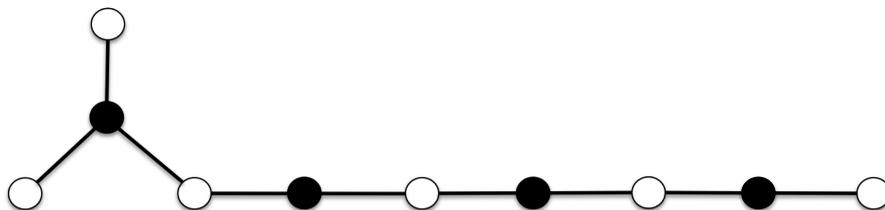


FIGURE 3. The cycle graph of the multi-set of permutations given in Figure 2.

The cycle graph and cycle diagram are clearly homotopically equivalent, and thus a multi-set of permutations is tree-like if and only if its cycle graph is a tree. In Section 6 we give several results on tree-like sets of permutations.

We may now state the characterization of iterated monodromy groups:

Theorem 5.8 ([15, Theorem 6.10.8]). *A subgroup $G \leq \text{Aut}(X^*)$ is isomorphic to a standard action of the iterated monodromy group of a post-critically finite polynomial if and only if G is conjugate in $\text{Aut}(X^*)$ to a group generated by a finite invertible automaton A with the following properties:*

- (1) *For each non-trivial $a \in A$, there is a unique arrow into the state a . In other words, there is a unique $b \in A$ and $x \in X$ with $b|_x = a$.*
- (2) *For each $a \in A$ and each cycle (x_1, x_2, \dots, x_n) of the action of a on X , the restriction $a|_{x_i}$ is non-trivial for at most one x_i .*
- (3) *The multi-set of permutations defined by the set of states of A acting on X is tree-like.*
- (4) *Let $a_1 \neq a_2$ be non-trivial states of A with $v_1, v_2 \in X^* \setminus \{\emptyset\}$ satisfying $a_i|_{v_i} = a_i$ and $a_i(v_i) = v_i$ for $i = 1, 2$. Then there is no $h \in G$ with $h(v_1) = v_2$ and $h|_{v_1} = h$.*

For example, the automaton given in Figure 1 satisfies all the conditions of Theorem 5.8. We do not use even close to the full strength of Theorem 5.8. Indeed, we require only the far easier direction, which is that if G is isomorphic to a standard action of an iterated monodromy group, then $G = \langle A \rangle$, where A satisfies conditions (1)-(4). Moreover, we do not use condition (2).

We introduce a definition following the terminology of [15]:

Definition 5.9. *A kneading automaton is a finite invertible automaton satisfying conditions (1)-(3) of Theorem 5.8.*

6. RESULTS ON TREE-LIKE SETS OF PERMUTATIONS

In this section we present several results that will play roles in the proofs of our main theorems. The first two appear in [15].

Lemma 6.1 ([15, Proposition 6.7.5]). *Let A be a kneading automaton. Then for any $n \geq 1$, the multi-set of permutations defined by the states of A acting on X^n is tree-like.*

Lemma 6.2 ([15, Corollary 6.7.7]). *If A is a kneading automaton, then the product of the states of A (taken in any order) is a spherically transitive element of $\text{Aut}(X^*)$.*

Lemma 6.3. *Let $T = \{\pi_1, \dots, \pi_n\}$ be a tree-like multi-set of permutations of a set X , and let $\sigma = \prod_{i \in I} \pi_i$ for some non-empty $I \subseteq \{1, \dots, n\}$. Suppose that $\sigma(x) = x$ for some $x \in X$. Then $\pi_i(x) = x$ for all $i \in I$.*

Proof. Induct on $\#I$. When $\#I = 1$ the statement is trivial. Suppose that $\#I \geq 2$ and $(\prod_{i \in I} \pi_i)(x) = x$, and let $k \in I$ and

$$\sigma_k = \left(\prod_{i \in I \setminus \{k\}} \pi_i \right).$$

If $\sigma_k(x) = y \neq x$, then necessarily $\pi_k(y) = x$. Thus in the cycle graph of T there is a path from the white vertex corresponding to x to the white vertex corresponding to y , given by the action of σ_k . There is a distinct path from the vertex corresponding to y back to the vertex corresponding to x , given by the action of π_k . This contradicts the hypothesis that the cycle graph is a tree.

Therefore $\sigma_k(x) = x$, and hence $\pi_k(x) = x$. Applying the inductive hypothesis to σ_k gives that $\pi_i(x) = x$ for all $i \in I$. \square

Lemma 6.4. *Let T be a tree-like multi-set of permutations acting on a set X with $\#X = d \geq 2$. Then the reduced cycle diagram of T has at most $(d - 1)$ 2-cells, with equality if and only if every element of T acts on X as a (possibly empty) disjoint product of transpositions.*

Proof. The condition that every element of T acts on X as a (possibly empty) disjoint product of transpositions is equivalent to the condition that each 2-cell in the reduced cycle diagram of T contains precisely two elements of X . In this case if we contract each 2-cell into a 1-cell joining the corresponding elements of X , we obtain a tree with d vertices. This tree has $d - 1$ edges by Euler’s formula, and hence there are $(d - 1)$ 2-cells in the reduced cycle diagram of T .

If there is a 2-cell C in the reduced cycle diagram of T containing c elements of X , with $c \geq 3$, then we may replace C by $(c - 1)$ 2-cells that each contain only two elements of X , in such a way that contractibility is preserved. Repeating this with every such C , we obtain a diagram with each 2-cell containing precisely two elements of X , which by the previous paragraph has $(d - 1)$ 2-cells. Hence the number of 2-cells in the reduced cycle diagram of T is at most $d - 1$, with equality only in the case where each 2-cell contains only two elements of X . \square

Lemma 6.5. *Let $T = \{\pi_1, \dots, \pi_n\}$ be a tree-like multi-set of permutations of a set X with $\#X = d$.*

- (1) *For any $i \neq j$, we have $\#\text{Fix}(\pi_i) + \#\text{Fix}(\pi_j) \geq 2$.*
- (2) *If $\#\text{Fix}(\pi_i) + \#\text{Fix}(\pi_j) \leq 3$ for some $i \neq j$, then π_k is the identity for all $k \notin \{i, j\}$.*

Proof. We begin by noting that by definition the cycle graph (and thus the reduced cycle diagram) of T is a contractible tree, and hence connected. If the cycle graph (equivalently, reduced cycle diagram) of some subset S of T is also connected, then the two cycle graphs must coincide, and hence all elements of the multi-set $S \setminus T$ are the identity.

Consider the reduced cycle diagram of the multi-set $\{\pi_i, \pi_j\}$, where $i \neq j$. It is a (possibly disconnected) planar graph, and hence by Euler’s formula satisfies

$$(6.1) \quad V - E + F = 2 + (c - 1),$$

where V, E , and F denote the numbers of vertices, edges, and faces (counting the face at infinity), respectively, and c denotes the number of connected components of the graph. Now the vertex set is just X , so $V = d$. There are r edges for each r -cycle of π_i or π_j , where $r > 1$ (recall that fixed points do not appear in the reduced cycle diagram). Thus $E = (d - \#\text{Fix}(\pi_i)) + (d - \#\text{Fix}(\pi_j))$. Finally, there is one face for each cycle of π_i or π_j , plus the face at infinity. Because the reduced cycle diagram of $\{\pi_i, \pi_j\}$ is a subset of the reduced cycle diagram for T , we have from Lemma 6.4 that π_i and π_j have at most $d - 1$ cycles between them, and hence $F \leq d - 1 + 1 = d$. Therefore (6.1) gives

$$(6.2) \quad 2d - (\#\text{Fix}(\pi_i) + \#\text{Fix}(\pi_j)) + 2 = E + 2 = V + F - (c - 1) \leq 2d - (c - 1),$$

and assertion (1) follows.

Note that in (6.2) we have equality if and only if $F = d$, which occurs precisely when the cycle diagram of $\{\pi_i, \pi_j\}$ has $(d - 1)$ 2-cells. When this happens, we

have by Lemma 6.4 that the number of 2-cells of the cycle diagram of $\{\pi_i, \pi_j\}$ is the same as the number of 2-cells of the cycle diagram of T , and hence the two diagrams coincide. It follows that $c = 1$. We have thus shown that either $c = 1$ or

$$(6.3) \quad 2d - (\#\text{Fix}(\pi_i) + \#\text{Fix}(\pi_j)) + 2 \leq 2d - c.$$

Now (6.3) implies that either $c = 1$ or $c \leq \#\text{Fix}(\pi_i) + \#\text{Fix}(\pi_j) - 2$. In particular, either $c = 1$ or $\#\text{Fix}(\pi_i) + \#\text{Fix}(\pi_j) \geq 4$. This together with the remarks at the beginning of the proof establish assertion (2). \square

7. STABILITY GRAPHS AND THE STRUCTURE OF \mathcal{N}_1

In this section we exploit condition (1) of Theorem 5.8 and the results of Section 6 to study the set

$$\mathcal{N}_1 = \{g \in G : g|_v = g \text{ and } g(v) = v \text{ for some non-empty } v \in X^*\}$$

first defined on p. 2026.

Condition (1) of Theorem 5.8 implies that if we delete the trivial state from the Moore diagram of a kneading automaton A (along with all the arrows originating at the trivial state), then the resulting graph is a disjoint union of cycles with trees attached to them. We call such a diagram the *reduced Moore diagram* of A . See Figure 4.

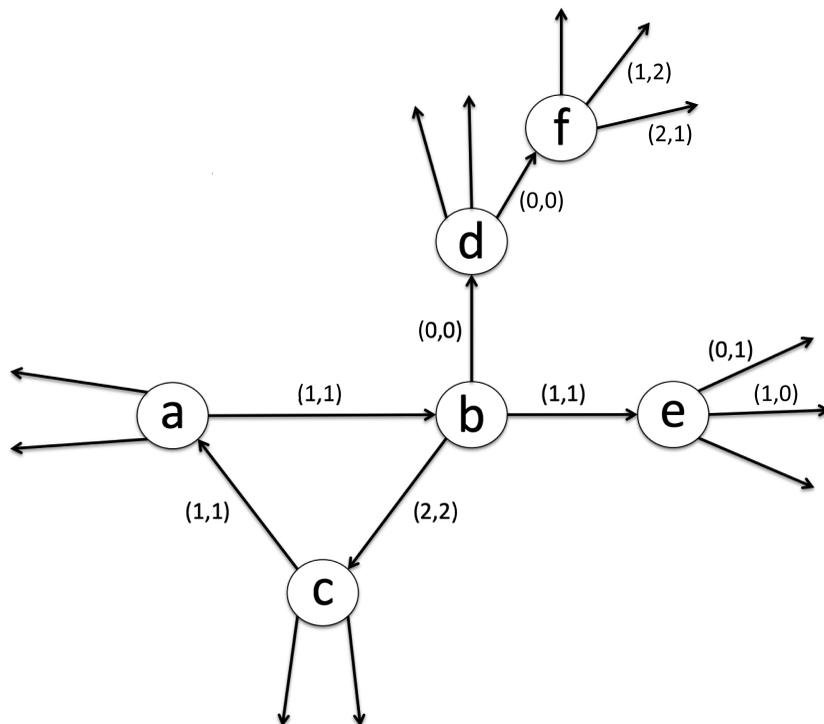


FIGURE 4. The reduced Moore diagram of a kneading automaton over $X = \{0, 1, 2\}$ satisfying all the conditions of Theorem 5.8. Labelings on arrows to the identity state have been deleted, except where the action on X is non-trivial.

In particular, the states not in cycles have the property that all restrictions to sufficiently long words are the identity, and hence they define finitary automorphisms of X^* . To each state a in a cycle of the Moore diagram we can associate its *stable sequence* $x_1x_2\cdots \in X^\omega$, which is the unique infinite word such that for each $v_n := x_1\cdots x_n$, $a|_{v_n}$ belongs to the cycle containing a . We refer to v_n as the *length- n stable sequence* of a . The (infinite) stable sequence of any given state is periodic, with period dividing the length of the cycle in which the element lies. For instance, for the automaton in Figure 4, the stable sequences of a , b , and c are $\overline{121}$, $\overline{211}$, and $\overline{112}$, respectively, where the bars denote repeating. By hypothesis A is invertible, and recall that the Moore diagram of the inverse automaton is given by replacing each state a by a^{-1} and changing each arrow labeling from (x, y) to (y, x) . Hence a^{-1} is in a cycle of the Moore diagram of the inverse automaton of A if and only if a is in a cycle of the Moore diagram of A . Each such a^{-1} has a stable sequence as before. Let C denote the collection of states of A that are in cycles of the Moore diagram, together with their inverses.

Let m be the least common multiple of the periods of the stable sequences of the elements of C . The *stability graph* of the automaton A is the directed graph whose vertex set is the set of length- m stable sequences of the states belonging to C . There is a directed edge from s_1 to s_2 if s_1 is the length- m stable sequence for some $a \in C$ and $a(s_1) = s_2$. We label such an edge with the element a . Two stability graphs are pictured in Figure 5.

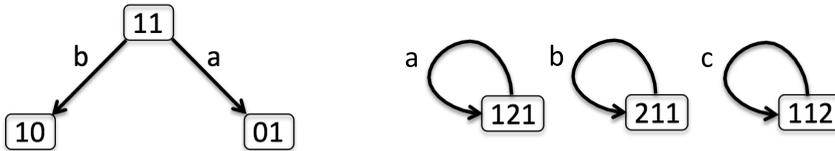


FIGURE 5. The stability graphs of the Basilica group (left) and the group generated by the automaton shown in Figure 4.

Recall that the set of *stable* elements of $G \leq \text{Aut}(X^*)$ is

$$\mathcal{N}_0 = \{g \in G : g|_v = g \text{ for some non-empty } v \in X^*\}.$$

Note that an element is stable if and only if it has a stable sequence. When G is generated by an automaton A satisfying the hypotheses of Theorem 5.8, the stability graph of A provides an algorithm for determining \mathcal{N}_0 and \mathcal{N}_1 . This idea first appeared in [2, Lemma 3.2], which deals with certain automata in the case $d = 2$.

We require some terminology relating to the stability graph. By a *path* we mean any sequence s_1, s_2, \dots, s_n of vertices such that there is a directed edge from s_{i-1} to s_i or a directed edge from s_i to s_{i-1} , for all $i = 2, \dots, n$. Note that this is more general than the usual notion of a path in a directed graph, since we permit paths to traverse edges against their direction. We further stipulate that our paths have *no back-tracking*, that is, each edge traversed is either distinct from the previous edge, or is the same as the previous edge and also in the same direction (i.e. consists of going again around a cycle of length one). By a *circuit*, we mean a path with a common starting and ending vertex; we allow repeats of vertices and edges. A

cycle is a circuit that repeats only its common starting and ending vertex. Recall that a kneading automaton is one satisfying conditions (1)-(3) of Theorem 5.8.

Proposition 7.1. *Let G be generated by a kneading automaton A . Then \mathcal{N}_0 consists of words in $A \cup A^{-1}$ obtained from the labels of paths in the stability graph of A , where one reads the inverse of the labeled element if one follows an arrow backwards. To assemble the word corresponding to a given path in the stability graph, one copies the letters down from right to left.*

In addition, \mathcal{N}_1 consists of the words obtained from labels of circuits in the stability graph of A .

Remark 7.2. For instance, the path of length 2 going from 01 to 10 in the stability graph of the Basilica group (Figure 5, left) gives ba^{-1} . The other path of length 2 gives ab^{-1} , while the paths of length 1 yield a, a^{-1}, b , and b^{-1} . Thus \mathcal{N}_0 consists of these six elements plus the identity. Since there are no circuits in the stability graph, \mathcal{N}_1 is trivial.

Proof. Recall that condition (1) of Theorem 5.8 ensures that each $a \in A$ not in a cycle of the reduced Moore diagram is finitary, that is, has trivial restriction on all sufficiently long words in X^* . If a is in a cycle, the length- n stable sequence v_n of a is the unique word of length n such that $a|_{v_n}$ is not finitary. We often simply call v_n the stable sequence of a when the length n is clear from context. If $a \in A$ has stable sequence v_n , then from (3.4) we have

$$(7.1) \quad 1 = (a^{-1}a)|_{v_n} = a^{-1}|_{a(v_n)}a|_{v_n}.$$

If $h \in G$ is finitary and $g \in G$ is not, then for sufficiently large k and any $v \in X^k$, we have $(hg)|_v = h|_{g(v)}g|_v = g|_v$. Hence hg cannot be finitary since g is not finitary. By hypothesis $a|_{v_n}$ is not finitary, and thus from (7.1) we have that $a^{-1}|_{a(v_n)}$ is not finitary, so that $a(v_n)$ is the stable sequence for a^{-1} . Because $a^{-1}a(v_n) = v_n$, multiplication by a^{-1} sends the stable sequence of a^{-1} to the stable sequence of a .

For $g \in \langle A \rangle$, let

$$(7.2) \quad g = a_1^{\epsilon_1} a_2^{\epsilon_2} \cdots a_k^{\epsilon_k},$$

where the a_i are (not necessarily distinct) elements of A , $\epsilon_i \in \{\pm 1\}$, and this expression is minimal length among all words in $A \cup A^{-1}$ giving g . We denote k by $\ell(g)$, and call it the length of g . From (3.3) it follows that $\ell(g|_v) \leq \ell(g)$ for any $g \in \langle A \rangle$ and any $v \in X^*$.

Suppose that $g \in \mathcal{N}_0$, so that there exists a non-empty $v \in X^*$ with $g|_v = g$. If v_n is the length- n initial word of $vv \cdots$, then $\ell(g|_{v_n}) = \ell(g)$. Hence each a_i lies in a cycle of the Moore diagram of A , since otherwise at least one would be finitary, implying $\ell(g|_w) < \ell(g)$ for w sufficiently long. Let m be the least common multiple of the periods of the (infinite) stable sequences of the a_i . Then

$$(7.3) \quad g|_{v_m} = (a_1^{\epsilon_1} a_2^{\epsilon_2} \cdots a_{k-1}^{\epsilon_{k-1}} a_k^{\epsilon_k})|_{v_m} \\ = a_1^{\epsilon_1}|_{a_2^{\epsilon_2} \cdots a_{k-1}^{\epsilon_{k-1}} a_k^{\epsilon_k}(v_m)} \cdot a_2^{\epsilon_2}|_{a_3^{\epsilon_3} \cdots a_{k-1}^{\epsilon_{k-1}} a_k^{\epsilon_k}(v_m)} \cdots a_{k-1}^{\epsilon_{k-1}}|_{a_k^{\epsilon_k}(v_m)} \cdot a_k^{\epsilon_k}|_{v_m}.$$

None of the elements in the right-hand side of (7.3) can be finitary, for otherwise $\ell(g|_w) < \ell(g)$ for sufficiently long w . Hence v_m is the (length- m) stable sequence for $a_k^{\epsilon_k}$, $a_k^{\epsilon_k}(v_m)$ is the stable sequence for $a_{k-1}^{\epsilon_{k-1}}$, $a_{k-1}^{\epsilon_{k-1}}a_k^{\epsilon_k}(v_m)$ is the stable sequence for $a_{k-2}^{\epsilon_{k-2}}$, and so on. Thus g determines a path in the stability graph of A , beginning at v_m , proceeding to $a_k^{\epsilon_k}(v_m)$, then to $a_{k-1}^{\epsilon_{k-1}}a_k^{\epsilon_k}(v_m)$, and so forth, ending at

$a_1^{\epsilon_1} a_2^{\epsilon_2} \cdots a_{k-1}^{\epsilon_{k-1}} a_k^{\epsilon_k}(v_m)$. There can be no back-tracking because of the minimality of (7.2). The path from v_m to $a_k^{\epsilon_k}(v_m)$ follows the arrow labeled a_k if $\epsilon_k = 1$, and runs against the arrow labeled a_k if $\epsilon_k = -1$. Assembling the labels along this path from right to left as indicated in the statement of the proposition then yields g .

Conversely, any path in the stability graph beginning at a vertex v_m yields a word $g = a_1^{\epsilon_1} a_2^{\epsilon_2} \cdots a_k^{\epsilon_k}$. Now v_m is the (length- m) stable sequence of $a_k^{\epsilon_k}$, and because v_m consists of some number of full cycles of the periodic part of the infinite stable sequence of $a_k^{\epsilon_k}$, we have $a_k^{\epsilon_k}|_{v_m} = a_k^{\epsilon_k}$. Similarly, $a_{k-1}^{\epsilon_{k-1}}|_{a_k^{\epsilon_k}(v_m)} = a_{k-1}^{\epsilon_{k-1}}$. Continuing in this manner we obtain $g|_{v_m} = g$, and hence $g \in \mathcal{N}_0$.

Now take $g \in \mathcal{N}_1$, so that there is some non-empty $v \in X^*$ with $g|_v = v$ and $g(v) = v$. Then for any length- n initial word v_n of $vv \cdots$, we have $\ell(g|_{v_n}) = \ell(g)$ and $g(v_n) = v_n$. Hence if m is as above, we have $a_1^{\epsilon_1} a_2^{\epsilon_2} \cdots a_{k-1}^{\epsilon_{k-1}} a_k^{\epsilon_k}(v_m) = v_m$, and so the path in the stability graph corresponding to g is a circuit. Conversely, any circuit yields g with $g|_{v_m} = g$ and $g(v_m) = v_m$. □

Theorem 7.3. *Let A be a kneading automaton. Then every cycle in the stability graph of A has length one.*

Proof. Let the cycle in question consist of the vertices s_0, s_1, \dots, s_n , with $s_n = s_0$ and s_0, \dots, s_{n-1} distinct. Suppose that $n \geq 2$, so that $s_0 \neq s_1$. Then assembling the labelings along this path as in Proposition 7.1 gives an element $g = a_1^{\epsilon_1} a_2^{\epsilon_2} \cdots a_n^{\epsilon_n}$ with the a_i distinct. Note that s_0 is the (length- m) stable sequence of $a_n^{\epsilon_n}$, and $a_i^{\epsilon_i} \cdots a_n^{\epsilon_n}(s_0) = s_{n-i+1}$ for each $i \geq 1$. In particular, $i = 1$ gives

$$(7.4) \quad a_1^{\epsilon_1} a_2^{\epsilon_2} \cdots a_n^{\epsilon_n}(s_0) = s_0.$$

By Lemma 6.1, the set of permutations given by the states of A acting on X^m is tree-like, and hence the cycle diagram of its action is contractible. If we replace some elements of A by their inverses, the cycle diagram is only altered by changing the directions of some arrows; in particular it is still contractible. Hence the multi-set of permutations given by $\{a_i^{\epsilon_i} : i = 1, \dots, n\}$ acting on X^m is a subset of a tree-like multi-set. From Lemma 6.3 and (7.4) we then have $a_i^{\epsilon_i}(s_0) = s_0$ for all i , which contradicts the fact that $a_n^{\epsilon_n}(s_0) = s_1 \neq s_0$. □

Theorem 7.4. *Let $G = \langle A \rangle$ satisfy all the conditions of Theorem 5.8. Then each component of the stability graph of A contains at most one cycle.*

Proof. By Theorem 7.3, each cycle of the stability graph of A has length one. Suppose that there are two such one-cycles at vertices s_1 and s_2 lying in a connected component of the stability graph of A , and let a_1 and a_2 be the elements labeling them. Then $a(s_i) = s_i$ for $i = 1, 2$. Moreover, s_1 and s_2 have length a multiple of the period of the stable sequences of a_1 and a_2 , and so $a_i|_{s_i} = a_i$ for $i = 1, 2$. Now s_1 and s_2 are in the same component of the stability graph, and so there is a path connecting them. Assembling the labelings along this path as in Proposition 7.1 gives $h \in G$ with $h(s_1) = s_2$ and $h|_{s_1} = h$ (see the construction in the converse portion of the proof of Proposition 7.1). But this contradicts condition (4) of Theorem 5.8. □

Theorem 7.5. *Let $G = \langle A \rangle$ satisfy all the conditions of Theorem 5.8. Then every element of \mathcal{N}_1 is conjugate to a power of an element of $A \cap \mathcal{N}_1$.*

Proof. By Proposition 7.1, elements of \mathcal{N}_1 correspond to circuits in the stability graph of A , which by definition have no back-tracking. By Theorems 7.3 and 7.4,

each such circuit belongs to a component C having at most a single cycle, which must have length one. Thus C is either a tree, or becomes a tree when we delete the edge forming the one-cycle. A tree cannot contain a non-trivial circuit, and so if γ is a non-trivial circuit in C , then C must contain a one-cycle at a vertex s , labeled by $a \in A$. Clearly $a \in \mathcal{N}_1$. Let s_0 be the starting point of γ , and note that if γ does not contain the one-cycle at s , then it lies entirely within a tree, which is impossible. Thus γ must proceed along the unique path to s , go around the one-cycle at s a non-zero number of times in the same direction each time, and return to s_0 the same way it came. If g is the element labeling the path from s_0 to s (assembled as in Proposition 7.1), then g^{-1} is the element labeling the reverse path. Hence the element labeling γ is conjugate to a power of a . \square

Corollary 7.6. *Let $G = \langle A \rangle$ satisfy all the conditions of Theorem 5.8, and suppose that every element of A that is in a cycle of the reduced Moore diagram either fixes no ends of X^* or fixes infinitely many. Then $\mathcal{F}(G) = 0$.*

Proof. Condition (1) of Theorem 5.8 implies that each $a \in A$ is bounded (see Definition 5.5). Indeed, for each $a \in A$, the restrictions of a to words of length n consist of the endpoints of all paths of length n in the reduced Moore diagram (following the arrows) starting at a . Because every non-trivial state has a unique incoming arrow, there can be at most one such path ending in each state. Hence $\#\{v \in X^n : g|_v \neq 1\}$ is bounded by $\#A$. By Theorem 5.6, G is therefore contracting. By Lemma 6.2, G contains a spherically transitive element. We may thus apply Theorem 1.5, and so to show $\mathcal{F}(G) = 0$ it is enough to show that every $g \in \mathcal{N}_1$ fixes infinitely many ends of X^* . By Theorem 7.5, each $g \in \mathcal{N}_1$ is conjugate to a^n for some $a \in A \cap \mathcal{N}_1$ and $n \in \mathbb{Z}$. Because $a \in \mathcal{N}_1$, a lies in a cycle of the reduced Moore diagram of A and also fixes at least one end of X^* . Thus by hypothesis a fixes infinitely many ends. But a^n fixes at least as many elements of each X^n as a does, and hence a^n fixes infinitely many ends of X^* . \square

8. THE FINAL STEPS

Recall that $g \in \text{Aut}(X^*)$ is finitary if all its restrictions at sufficiently long words are the identity. If g is finitary, then it either fixes no ends of X^* or fixes infinitely many such ends. Indeed, if w is an end fixed by g and w_n is the length- n initial word of w , then $g(w_n) = w_n$ for all n and we may take n large enough so that $g|_{w_n} = 1$. Thus g fixes all ends with initial word w_n , which is an infinite set.

Throughout this section, when we write $\text{Fix}(g)$ for $g \in \text{Aut}(X^*)$, we mean the set of fixed points of the action of g on X (not on the ends of X^*).

Lemma 8.1. *Let A be a kneading automaton, and let $a, b \in A$ be finitary with $a \neq b$. Then at least one of a, b fixes infinitely many ends of X^* .*

Proof. By part (1) of Lemma 6.5, $\#\text{Fix}(a) + \#\text{Fix}(b) \geq 2$, and hence there exist elements $x_0, y_0 \in X$ that are fixed by either a or b . Renaming if necessary, assume that a fixes x_0 , and let $a_1 = a|_{x_0}$. Let $b_1 = a|_{y_0}$ if $a(y_0) = y_0$ and $b_1 = b|_{y_0}$ if $b(y_0) = y_0$. Applying part (1) of Lemma 6.5 again, there exist $x_1, y_1 \in X$ that are fixed by either a_1 or b_1 . Without loss of generality, assume $a_1(x_1) = x_1$, and let $a_2 = a_1|_{x_1}$. Let $b_2 = a_1|_{y_1}$ if $a_1(y_1) = y_1$ and $b_2 = b_1|_{y_1}$ if $b_1(y_1) = y_1$. Proceeding in this manner yields a sequence a_1, a_2, \dots of elements of A and words $w_n = x_0 x_1 \cdots x_n$ such that $a(w_n) = w_n$ and $a|_{w_n} = a_{n+1}$.

Because a is finitary, there exists $j > 0$ such that all a_n are trivial for $n \geq j$. Hence a fixes the word $w_j = x_0x_1 \cdots x_j$ and $a|_{w_j} = 1$, implying that a fixes all ends of X^* with initial word w_j , which is an infinite set. \square

Theorem 8.2. *Let $G = \langle A \rangle$ satisfy the conditions of Theorem 5.8, where A is defined over an alphabet X with $\#X = d$. Let C be a cycle of the reduced Moore diagram of A . If C contains at least two elements and one of them fixes a non-empty, finite set of ends of X^* , then d is odd, C is the only cycle, and up to conjugation in $\text{Aut}(X^*)$ we have $A = \{1, a, b\}$ with*

$$(8.1) \quad a = \sigma(b, 1, 1, \dots, 1) \quad b = \tau(1, 1, \dots, 1, a),$$

where σ and τ are products of disjoint transpositions, σ fixes only $0 \in X$, and τ fixes only $d - 1 \in X$. In particular, $\langle A \rangle$ is infinite dihedral.

Proof. Let $C = \{c_1, c_2, \dots, c_n\}$, so that there is an arrow in the reduced Moore diagram from c_i to c_{i+1} for $i = 1, \dots, n - 1$ and also from c_n to c_1 . Let $v_1 = x_1x_2 \cdots x_n$ be the length- n stable sequence of c_1 , implying that $v_i = x_ix_{i+1} \cdots x_nx_1 \cdots x_{i-1}$ is the stable sequence of c_i . Then $c_i|_{v_i} = c_i$ and from (3.2),

$$c_i(v_i) = c_i(x_i)c_{i+1}(x_{i+1}) \cdots c_n(x_n)c_1(x_1) \cdots c_{i-1}(x_{i-1}).$$

Hence c_i fixes v_i if and only if each c_j fixes x_j .

Assume that C contains at least two elements and one of them fixes a non-empty, finite set of ends of X^* . Suppose first that c_i does not fix v_i for some i . If c_i fixes an end w of X^* , then w cannot be the (infinite) stable sequence of c_i . Letting w_j be the length- j initial word of w , we can thus take j large enough so that $c_i|_{w_j}$ is finitary. Let $b = c_i|_{w_j}$, and note that for $k > j$, $c_i|_{w_k}$ is a restriction of b at a word of length $k - j$. Because b is finitary, we may take k large enough so that $c_i|_{w_k} = 1$. But c_i fixes w , and thus $c_i(w_k) = w_k$, ensuring that c_i fixes all ends of X^* with initial word w_k . Hence all c_i either fix no ends of X^* or infinitely many, a contradiction.

Suppose now that c_i fixes v_i for some i (equivalently, all i). Then for each $j \neq i$ there is some word w_j with $c_i(w_j) = w_j$ and $c_i|_{w_j} = c_j$. Thus if any c_j fixes infinitely many ends of X^* , then the same conclusion holds for all the c_j . If $n \geq 3$, then we claim

$$(8.2) \quad \#\text{Fix}(c_1) + \#\text{Fix}(c_2) + \#\text{Fix}(c_3) \geq 5.$$

To see why, note that $\#\text{Fix}(c_i) \geq 1$ for all i by assumption, and if $\#\text{Fix}(c_i) + \#\text{Fix}(c_j) \leq 3$ for some $i \neq j$, then part (2) of Lemma 6.5 gives $\#\text{Fix}(c_k) = d$ for the remaining k . If $d = 2$, then we must have $\#\text{Fix}(c_i) = 2$ for all i , and (8.2) holds. If $d \geq 3$ and $\#\text{Fix}(c_i) = \#\text{Fix}(c_j) = 1$, then $\#\text{Fix}(c_k) = d$ and again (8.2) holds.

Thus by (8.2) there are $y_1, y_2 \in X \setminus \{x_1, x_2, x_3\}$ with $c_i(y_1) = y_1$ and $c_j(y_2) = y_2$ (here i is allowed to equal j). Now $c_i|_{y_1}$ and $c_j|_{y_2}$ are distinct and finitary, and by Lemma 8.1 at least one of them fixes infinitely many ends of X^* . Thus at least one of the c_i fixes infinitely many ends of X^* , and hence all do. This gives a contradiction.

We have therefore shown that $\#C = 2$, and we write $C = \{a, b\}$. Let x_1x_2 be the stable sequence of a , implying that x_2x_1 is the stable sequence of b , and recall that both a and b must fix their stable sequences. If $\#\text{Fix}(a) + \#\text{Fix}(b) \geq 4$, then there exist $y_1, y_2 \in X \setminus \{x_1, x_2\}$ fixed by either a or b . As in the previous paragraph,

we conclude that both a and b fix infinitely many ends of X^* , a contradiction. Hence $\#\text{Fix}(a) + \#\text{Fix}(b) \leq 3$, and from part (2) of Lemma 6.5 we have that every $a' \in A \setminus \{a, b\}$ must act as the identity on X . If a' is in the component of the reduced Moore diagram containing C , then it cannot be part of C , and neither can any of its restrictions. Thus a' acts trivially on X^* . If a' is in a component of the reduced Moore diagram of A that does not contain C , then every element of this component must act trivially on X . Since components are closed under restriction, it follows that every element of this component is trivial. Thus $A = \{1, a, b\}$.

Now if $\#\text{Fix}(a) + \#\text{Fix}(b) = 3$, then there is some $y \in X \setminus \{x_1, x_2\}$ that we may assume without loss is fixed by a . Then $a|_y$ is not in the cycle $\{a, b\}$, and so $a|_y = 1$, showing that a fixes infinitely many ends of X^* . Thus b must as well, which is a contradiction. Therefore $\#\text{Fix}(a) = \#\text{Fix}(b) = 1$, and we must have $\text{Fix}(a) \cap \text{Fix}(b) = \emptyset$. Otherwise a and b have the same stable sequence, and thus give two one-cycles in the same component of the stability graph of A , violating Theorem 7.4. Hence conjugating by an appropriate $\gamma(1, 1, \dots, 1) \in \text{Aut}(X^*)$ we may assume that $\text{Fix}(a) = \{0\}$ and $\text{Fix}(b) = \{d - 1\}$, giving the forms in (8.1). Moreover, because $\#\text{Fix}(a) + \#\text{Fix}(b) = 2$ we must have equality in (6.1), which implies that the reduced cycle diagram of $\{\sigma, \tau\}$ has $(d - 1)$ 2-cells. By Lemma 6.4 it follows that σ and τ are products of disjoint transpositions, and hence d must be odd.

Now $a^2 = (a^2, 1, 1, \dots, 1)$, and so $a^2 = 1$, and similarly $b^2 = 1$. By Lemma 6.2 we have that ab has infinite order, and conjugating ab by a gives ba , which is $(ab)^{-1}$. Hence $\langle A \rangle$ is infinite dihedral. □

Theorem 8.3. *Let $G = \langle A \rangle$ satisfy the conditions of Theorem 5.8, where A is defined over an alphabet X with $\#X = d$. Let $C = \{c\}$ be a 1-cycle in the reduced Moore diagram of A , and suppose that c fixes at least two points of X . If c fixes a non-empty, finite set of ends of X^* , then d is even, C is the only cycle, and up to conjugation in $\text{Aut}(X^*)$ we have $A = \{c, a, 1\}$ with*

$$(8.3) \quad c = \sigma(c, a, 1, 1, \dots, 1) \quad a = \tau(1, 1, \dots, 1),$$

where σ and τ are products of disjoint transpositions, $\text{Fix}(\sigma) = \{0, 1\}$, and $\text{Fix}(\tau) = \emptyset$. In particular, $\langle A \rangle$ is infinite dihedral.

Proof. Suppose that c fixes a non-empty, finite set of ends of X^* . Let $x \in X$ be such that $c|_x = x$. If $\#\text{Fix}(c) \geq 3$, then there are $y_1, y_2 \in X \setminus \{x\}$ with $y_1 \neq y_2$ and $c(y_i) = y_i$. Then $c|_{y_i}$ is finitary for $i = 1, 2$, and by Lemma 8.1 one of them fixes infinitely many ends of X^* . It follows that c fixes infinitely many ends of X^* , a contradiction.

Hence $\#\text{Fix}(c) = 2$. If $c(x) \neq x$, then $\text{Fix}(c) = \{y_1, y_2\} \subseteq X \setminus \{x\}$, and as before we have a contradiction. Thus $\text{Fix}(c) = \{x, y\}$ for some $y \neq x$. Let $a = c|_y$. If $\#\text{Fix}(a) \geq 2$, then there are $z_1, z_2 \in X \setminus \{y\}$ with $z_1 \neq z_2$ and $a(z_i) = z_i$. But $a|_{z_i}$ is finitary for $i = 1, 2$, and as before this gives a contradiction. Hence $\#\text{Fix}(a) \leq 1$, so that $\#\text{Fix}(c) + \#\text{Fix}(a) \leq 3$, and we may apply part (2) of Theorem 6.5 as in the proof of Theorem 8.2 to get that $A = \{c, a, 1\}$. If $\#\text{Fix}(a) = 1$, then the restriction of a at this fixed point must be trivial, which implies that c fixes infinitely many ends of X^* . Thus $\text{Fix}(a) = \emptyset$. Conjugating by an appropriate element $\gamma(1, 1, \dots, 1) \in \text{Aut}(X^*)$ allows us to move x to 0 and y to 1, so that c and a have the forms in (8.3). Because $\#\text{Fix}(c) + \#\text{Fix}(a) = 2$ we must have equality in (6.1), which implies that the reduced cycle diagram of $\{\sigma, \tau\}$ has $(d - 1)$ 2-cells.

By Lemma 6.4 it follows that σ and τ are products of disjoint transpositions, and hence d must be even. That $\langle A \rangle$ is dihedral follows just as in the proof of Theorem 8.2. □

We now have all the tools in place to finish proving our main result.

Proof of Theorem 1.1. Recall that $f \in \mathbb{C}[z]$ is exceptional if there exists a finite, non-empty set Σ such that $f^{-1}(\Sigma) \setminus \text{Crit } f = \Sigma$. Suppose f is non-exceptional, put $d = \deg f$, and let $G \leq \text{Aut}(X^*)$ be a standard action of $\text{IMG}(f)$ on X^* . There is an automaton A that generates G and satisfies the conditions of Theorem 5.8.

Suppose that there is a cycle C in the reduced Moore diagram of A such that some element of C fixes a non-empty, finite set of ends of X^* . If $\#C \geq 2$, then by Theorem 8.2, G is generated by two elements of the form (8.1). From Theorem 5.1 it follows that the post-critical set of f consists of $\{z_0, z_1\} \subset \mathbb{C}$ where $a = g_{z_0}, b = g_{z_1}, z_0 \neq z_1$, and neither z_0 nor z_1 is critical. Moreover, $f^{-1}(\{z_0, z_1\}) \setminus \{z_0, z_1\}$ is contained in the set of critical points of f , and hence f is exceptional (indeed, by Proposition 8.4 it is conjugate to $-T_d$ for some odd d). If $C = \{c\}$, then let $x \in X$ be such that $c|_x = c$, and let $\text{Fix}(c)$ be the set of fixed points in the action of c on X . By assumption $\text{Fix}(c)$ is non-empty. If $\#\text{Fix}(c) \geq 2$, then by Theorem 8.3, G is generated by two elements of the form (8.3). As before Theorem 5.1 implies that f is exceptional (it is conjugate to T_d for some even d). If $\text{Fix}(c) = \{y\}$ with $y \neq x$, then $c|_y$ must be finitary and hence fix either no ends or infinitely many ends by the remarks at the beginning of this section. Thus c either fixes no ends or infinitely many ends, contrary to our assumption. If $\text{Fix}(c) = \{x\}$, then by Theorem 5.1 we have $x = g_{z_0}$ where z_0 is non-critical and $f^{-1}(\{z_0\}) \setminus \{z_0\}$ is contained in the set of critical points of f . Once again f must be exceptional (it is conjugate to a polynomial of the form (1.3) or to T_d for some odd d).

Therefore the reduced Moore diagram of A contains no such cycle C . By Corollary 7.6 we have $\mathcal{F}(G) = 0$. □

In order to prove Proposition 1.2 and the remarks preceding it, we make a brief study of exceptional polynomials (see the discussion following Lemma 2.3 in [10] for similar remarks in the case of rational functions). Suppose that f has degree d and is exceptional, with Σ a finite set such that $f^{-1}(\Sigma) \setminus \text{Crit } f = \Sigma$. Following [10], we note that each preimage of a point in Σ is either a critical point or in Σ . Hence, letting $\Gamma = \text{Crit } f \cap f^{-1}(\Sigma)$, we have

$$d \cdot \#\Sigma - \sum_{c \in \Gamma} (\deg(c) - 1) = \#f^{-1}(\Sigma) \leq \#\Sigma + \#\Gamma,$$

where $\deg(c)$ is the local degree of f at c . Because f has $d - 1$ critical points up to multiplicity, this gives

$$(d - 1)\#\Sigma = \sum_{c \in \Gamma} (\deg(c) - 1) + \#\Gamma \leq 2(d - 1),$$

with equality holding if and only if f has $d - 1$ distinct critical points (necessarily each having multiplicity 2), all of which are contained in $f^{-1}(\Sigma)$.

We now give a complete characterization of the case $\#\Sigma = 2$. While our result is not new, the statement and proof differ in form from the standard treatments in the literature (e.g. [14, Theorem 19.9]). Recall that the Chebyshev polynomial of degree d is given by $T_d(z) = \cos(d \arccos z)$. Its critical set Γ consists of $d - 1$ distinct

points and satisfies $\Gamma \cap \{\pm 1\} = \emptyset$, $T_d(\Gamma) = \{\pm 1\}$, $T_d(1) = 1$, $T_d(-1) = (-1)^d$, and $T_d^{-1}\{\pm 1\} = \{\pm 1\} \cup \Gamma$.

Proposition 8.4. *Let $f \in \mathbb{C}[z]$ be exceptional, with $\Sigma = \{z_0, z_1\}$ and $z_0 \neq z_1$. Then one of the following holds:*

- (1) $f(z_0) = z_0$, $f(z_1) = z_1$, and f is conjugate to T_d for some odd d .
- (2) $f(z_0) = z_1$, $f(z_1) = z_0$, and f is conjugate to $-T_d$ for some odd d .
- (3) $f(z_0) = f(z_1) \in \{z_0, z_1\}$, and f is conjugate to T_d for some even d .

Proof. We outline an algebraic approach, which differs somewhat from the well-known geometric arguments (see e.g. [14, Theorem 19.9]). First note that by the definition of exceptional polynomial, $f(\Sigma) \subseteq \Sigma$ and Σ contains no critical points. By the discussion preceding the proposition, f has $d - 1$ critical points, all of which have multiplicity 2. Applying an appropriate affine conjugation, we may assume that $z_0 = -1$ and $z_1 = 1$. In case (1) of the proposition, we then have

$$(8.4) \quad f(z) + 1 = r(z + 1)(g(z))^2 \quad f(z) - 1 = r(z - 1)(h(z))^2,$$

where $r \in \mathbb{C}$, and $g, h \in \mathbb{C}[z]$ are monic and relatively prime. Differentiating gives $f' = r[2(z + 1)g'g + g^2] = r[2(z - 1)h'h + h^2]$, but the roots of f' are the same as the roots of gh , whence $f' = drgh$, where d is the degree of f . Because g and h are relatively prime, we obtain $dh = 2(z + 1)g' + g$ and $dg = 2(z - 1)h' + h$. Differentiating again and substituting yields

$$d^2g = 4(z + 1)(z - 1)g'' + (8z - 4)g' + g.$$

This differential equation gives a recurrence relation on the coefficients of g ; with the assumption that g is monic, this uniquely determines all coefficients of g . Because $f(1) = 1$, we have by (8.4) that $2 = 2r(g(1))^2$, thereby determining r , and thus also f . However, T_d clearly satisfies the same conditions as f , and thus $f = T_d$. Part (1) of the proposition follows. The other parts proceed similarly. □

To prove Proposition 1.2, we require the following result.

Proposition 8.5. *Let $\#X = d$ and suppose that $G \leq \text{Aut}(X^*)$ is generated by $\{a, b\}$ where a and b are distinct and non-trivial, $a^2 = b^2 = 1$, and ab is spherically transitive. Then $\mathcal{F}(G) = r/4$, where r is the number of elements in $\{a, b\}$ fixing at least one end of X^* .*

Proof. Because ab is spherically transitive, we have that for any $n \geq 1$ its action on X^n is a d^n -cycle. Clearly conjugation of ab by a gives $(ab)^{-1}$, and it follows that if G_n is the action of G on X^n , then G_n is dihedral of order $2d^n$ and a complete list of its elements is given by the actions of $(ab)^k$ and $a(ab)^k$ for $k = 0, \dots, d^n$. None of the elements of the form $(ab)^k$ for $k = 1, \dots, d^n - 1$ can have fixed points in X^n . An element of the form $a(ab)^k$ is conjugate to b if k is odd and to a if k is even. Now the action of a either has a fixed point in X^n for all n (if a fixes an end of X^*) or has no fixed points in X^n for n large enough, and similar statements hold for b . Thus for n sufficiently large, the number of elements of G_n fixing at least one point of X^n is $1 + r(d^n/2)$. Dividing by $\#G_n = 2d^n$ and letting $n \rightarrow \infty$ gives $\mathcal{F}(G) = r/4$. □

Proof of Proposition 1.2. Let $f(z)$ have degree d , and let $G \leq \text{Aut}(X^*)$ be a standard action of $\text{IMG}(f)$ on X^* . If $f(z)$ is conjugate to T_d for d even, then it follows

from Theorem 5.1 that G is generated by two elements of the form (8.3). Proposition 8.5 then applies to show $\mathcal{F}(G) = 1/4$. If $f(z)$ is conjugate to $-T_d$ for d odd, then it follows from Theorem 5.1 that G is generated by two elements of the form (8.1). If $f(z)$ is conjugate to T_d for d odd, then G is generated by two elements of the form

$$a = \sigma(a, 1, 1, \dots, 1) \quad b = \tau(1, 1, \dots, 1, b),$$

with assumptions as in (8.1). In either case Proposition 8.5 applies to show $\mathcal{F}(G) = 1/2$. \square

ACKNOWLEDGEMENTS

The author would like to thank Lasse Rempe, Juan Rivera-Letelier, and Mikhail Lyubich for helpful comments and references to the literature on exceptional maps. The author also extends his thanks to the Institute for Computational and Experimental Research in Mathematics, where he presented and received valuable feedback on some of these results as part of the semester on complex and arithmetic dynamics.

REFERENCES

- [1] Laurent Bartholdi and Volodymyr Nekrashevych, *Thurston equivalence of topological polynomials*, Acta Math. **197** (2006), no. 1, 1–51, DOI 10.1007/s11511-006-0007-3. MR2285317 (2008c:37072)
- [2] Laurent Bartholdi and Volodymyr Nekrashevych, *Iterated monodromy groups of quadratic polynomials. I*, Groups Geom. Dyn. **2** (2008), no. 3, 309–336, DOI 10.4171/GGD/42. MR2415302 (2009d:37074)
- [3] Laurent Bartholdi and Bálint Virág, *Amenability via random walks*, Duke Math. J. **130** (2005), no. 1, 39–56, DOI 10.1215/S0012-7094-05-13012-5. MR2176547 (2006h:43001)
- [4] E. Bondarenko and V. Nekrashevych, *Post-critically finite self-similar groups*, Algebra Discrete Math. **4** (2003), 21–32. MR2070400 (2005d:20041)
- [5] Otto Forster. *Lectures on Riemann surfaces*, volume 81 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1991. Translated from the 1977 German original by Bruce Gilligan, Reprint of the 1981 English translation.
- [6] Rostislav Grigorchuk, Dmytro Savchuk, and Zoran Šunić, *The spectral problem, substitutions and iterated monodromy*, Probability and mathematical physics, CRM Proc. Lecture Notes, vol. 42, Amer. Math. Soc., Providence, RI, 2007, pp. 225–248. MR2352271 (2008h:20040)
- [7] Geoffrey R. Grimmett and David R. Stirzaker, *Probability and random processes*, 3rd ed., Oxford University Press, New York, 2001. MR2059709 (2004m:60002)
- [8] Rafe Jones, *Iterated Galois towers, their associated martingales, and the p -adic Mandelbrot set*, Compos. Math. **143** (2007), no. 5, 1108–1126, DOI 10.1112/S0010437X07002667. MR2360312 (2008i:11131)
- [9] Rafe Jones, *The density of prime divisors in the arithmetic dynamics of quadratic polynomials*, J. Lond. Math. Soc. (2) **78** (2008), no. 2, 523–544, DOI 10.1112/jlms/jdn034. MR2439638 (2010b:37239)
- [10] Jeremy Kahn, Mikhail Lyubich, and Lasse Rempe, *A note on hyperbolic leaves and wild laminations of rational functions*, J. Difference Equ. Appl. **16** (2010), no. 5-6, 655–665, DOI 10.1080/10236190903257867. MR2675597 (2011j:37080)
- [11] Mikhail Lyubich and Yair Minsky, *Laminations in holomorphic dynamics*, J. Differential Geom. **47** (1997), no. 1, 17–94. MR1601430 (98k:58191)
- [12] N. Makarov and S. Smirnov, *Phase transition in subhyperbolic Julia sets*, Ergodic Theory Dynam. Systems **16** (1996), no. 1, 125–157, DOI 10.1017/S0143385700008749. MR1375130 (98g:58148)
- [13] N. Makarov and S. Smirnov, *On “thermodynamics” of rational maps. I. Negative spectrum*, Comm. Math. Phys. **211** (2000), no. 3, 705–743, DOI 10.1007/s002200050833. MR1773815 (2001m:37088)

- [14] John Milnor, *Dynamics in one complex variable*, 3rd ed., Annals of Mathematics Studies, vol. 160, Princeton University Press, Princeton, NJ, 2006. MR2193309 (2006g:37070)
- [15] Volodymyr Nekrashevych, *Self-similar groups*, Mathematical Surveys and Monographs, vol. 117, American Mathematical Society, Providence, RI, 2005. MR2162164 (2006e:20047)
- [16] Volodymyr Nekrashevych, *Iterated monodromy groups*, Groups St Andrews 2009 in Bath, Volume 1, London Math. Soc. Lecture Note Ser., vol. 387, Cambridge Univ. Press, Cambridge, 2011, pp. 41–93. MR2858850
- [17] Michael Rosen, *Number theory in function fields*, Graduate Texts in Mathematics, vol. 210, Springer-Verlag, New York, 2002. MR1876657 (2003d:11171)

DEPARTMENT OF MATHEMATICS, CARLETON COLLEGE, ONE NORTH COLLEGE STREET, NORTH-FIELD, MINNESOTA 55057

E-mail address: rfjones@carleton.edu