THE DENSITY OF PRIME DIVISORS IN THE ARITHMETIC DYNAMICS OF QUADRATIC POLYNOMIALS

RAFE JONES

ABSTRACT. Let $f \in \mathbb{Z}[x]$, and consider the recurrence given by $a_n = f(a_{n-1})$, with $a_0 \in \mathbb{Z}$. Denote by $P(f, a_0)$ the set of prime divisors of this recurrence, i.e., the set of primes p dividing some non-zero a_n, and denote the natural density of this set by $D(P(f, a_0))$. The problem of determining $D(P(f, a_0))$ when f is linear has attracted significant study, although it remains unresolved in full generality. In this paper we consider the case of f quadratic, where previously $D(P(f, a_0))$ was known only in a few cases. We show $D(P(f, a_0)) = 0$ regardless of a_0 for four infinite families of f, including $f = x^2 + k$, $k \in \mathbb{Z} \backslash \{-1\}$. The proof relies on tools from group theory and probability theory to formulate a sufficient condition for $D(P(f, a_0)) = 0$ in terms of arithmetic properties of the forward orbit of the critical point of f. This provides an analogy to results in real and complex dynamics, where analytic properties of the forward orbit of the critical point have been shown to determine many global dynamical properties of a quadratic polynomial. The article also includes apparently new work on the irreducibility of iterates of quadratic polynomials.

1. INTRODUCTION

[1] Let $a_n = f(a_{n-1}, \ldots, a_{n-k})$ be a recursively defined sequence of order k, where f is a polynomial in k variables. Many well-known sequences belong to this class, such as the Fibonacci numbers ($F_0 = F_1 = 1, F_n = F_{n-1} + F_{n-2}$), the Lucas numbers ($L_0 = 1, L_1 = 3, L_n = L_{n-1} + L_{n-2}$), the Fermat numbers $t_n = 2^{2^n} + 1$ ($t_0 = 3, t_n = (t_{n-1} - 1)^2 + 1$), and the Mersenne numbers $m_n = 2^n - 1$ ($m_0 = 0, m_n = 2m_{n-1} + 1$). The set of prime divisors of a_n, namely

$$(1) \qquad P(a_n) = \{p \text{ prime} : p \text{ divides } a_i \text{ for some } i \geq 0 \text{ with } a_i \neq 0\}$$

has attracted great interest, although it often is too much to hope for a complete understanding of this set. For instance, since the Fermat numbers t_n are relatively prime, a full understanding of $P(t_n)$ would permit one to resolve the well-known conjecture that only finitely many of the t_n are prime. In this paper we consider the natural density $D(P(a_n))$ for certain a_n. Recall that the natural density of a set S of primes is

$$(2) \qquad D(S) = \lim_{x \to \infty} \frac{\#\{p \in S : p \leq x\}}{\#\{p : p \leq x\}},$$

provided that this limit exists.

When a_n is a linear recurrence, the density of $P(a_n)$ has been much studied, e.g. [4], [10], [13], [21], [15]. Note that $P(a_n)$ is infinite, excluding a few degenerate cases [7], [23]. It is easy to see that $P(m_n)$ contains all $p \neq 2$ and $P(F_n)$ contains all p (see e.g. [24] for the latter). However, the same is not true of L_n: Lagarias [13] shows

---

[1] 11R32 (primary), 11B37, 11C08, 60G42 (secondary)

$D(P(L_n)) = 2/3$. Moreover, Hasse [10] showed that the sequence $a_n = 2^n + 1$, a relative of the Mersenne numbers, satisfies $D(P(a_n)) = 17/24$. More generally, Moree and Stevenhagen [15], drawing on and correcting work of Stephens [21], have studied a class of recurrences including the case $a_n = (b+1)a_{n-1} - ba_{n-2}$. Assuming the Riemann Hypothesis for certain Kummer extensions, and with mild conditions on a_0 and a_1, they show that $D(P(a_n))$ is positive and depends in a delicate way on b [15, Theorem 4]. Lagarias [13] remarks that under the Generalized Riemann Hypothesis it is possible to show that $D(P(a_n))$ exists and is positive for any non-degenerate second-order linear recurrence a_n.

We restrict our attention in this work to recurrences with first-order representations, that is, sequences $a_n = f(a_{n-1})$ with $f \in \mathbb{Z}[x]$ and $a_0 \in \mathbb{Z}$ (such a sequence also has a kth-order representation for any $k \geq 1$). These sequences are orbits of integers in the arithmetic dynamical system defined by iteration of f. We write $P(f, a_0)$ instead of $P(a_n)$, and our main concern is to determine $D(P(f, a_0))$ for various f and a_0. If f is linear, then writing $f = bx + c$ we note that our sequence satisfies $a_n = (b+1)a_{n-1} - ba_{n-2}$, with $a_1 = ba_0 + c$. This is the case treated in [15].

By contrast, the case where $\deg(f) \geq 2$ has received relatively little attention. The main result is of Odoni [17], who proves that "most" polynomials of degree d have the property that all of their integer orbits contain a "very small" proportion of the primes. More specifically, he shows the following: let $A(d, N)$ denote the set of monic f with $\deg(f) = d \geq 2$ and coefficients of absolute value at most N. Then for any $\epsilon > 0$, as $N \to \infty$ the proportion of $f \in A(d, N)$ with $D(P(f, a_0)) < \epsilon$ for all $a_0 \in \mathbb{Z}$ approaches 1. Odoni obtains his result by considering the generic monic of degree d and applying a version of the Hilbert irreducibility theorem. Thus the nature of the exceptional set is not explicit, and the result does not determine $D(P(f, a_0))$ for any specific f and a_0.

On the other hand, in [18] Odoni gives a thorough analysis of Sylvester's sequence, namely $w_0 = 2$, $w_n = 1 + w_0 w_1 \cdots w_{n-1}$. This sequence has a long and rich history and many interesting properties. Among those that attracted Odoni's attention are its reminiscence of Euclid's proof of the infinitude of the primes and the fact that the sequence is also given by $w_0 = 2$, $w_n = f(w_{n-1})$, where $f = x^2 - x + 1$. Odoni shows that $D(P(f, a_0)) = 0$ for all $a_0 \in \mathbb{Z}$.

In this article we consider the case $\deg(f) = 2$. We note that $P(f, a_0)$ is infinite as long as $f \neq x^2$ and $\{f^n(a_0) : n \in \mathbb{N}\}$ is infinite; we prove this in Section 6. For $n \geq 1$ we denote the nth iterate of f by f^n, and we take $f^0 = x$. To give our primary results, we need a few preliminaries. First, to deal with the case when some iterate of f is reducible, it is desirable to have results for *translated* iterated sequences, that is, those of the form $g(f^n(a_0))$, where $g, f \in \mathbb{Z}[x]$. We denote by $P(g, f, a_0)$ the set of primes dividing at least one non-zero term of this sequence. Second, by a *rigid divisibility sequence*, we mean a sequence b_n of S-integers in $\mathbb{Q}$ with two properties: first, for all $p \notin S$, $v_p(b_n) > 0$ implies $v_p(b_{mn}) = v_p(b_n)$ for all $m \geq 1$; second, for all $p \notin S$, $p^e \mid c_n$ and $p^e \mid c_m$ implies $p^e \mid c_{\gcd(m,n)}$. This is a strengthening of the notion of divisibility sequence (i.e., $b_n \mid b_{mn}$ in the case $S = \emptyset$). The cases of interest to us always have $S = \{2\}$ or $S = \emptyset$.

Theorem 1.1. *Let $f, g \in \mathbb{Z}[x]$ be monic with f quadratic, and let γ be the critical point of f. Suppose that $g \circ f^n$ is irreducible for all $n \geq 0$ and the set $\{g(f^n(\gamma))\}$ is infinite. If either*

(1) g divides f^r for some $r \geq 0$ and the set $\{f^n(0) : n = 1, 2, \ldots\}$ is finite and does not contain 0, or
(2) the sequence $(g(f^n(\gamma)) : n = 1, 2, \ldots)$ is a rigid divisibility sequence,

then $D(P(g, f, a_0)) = 0$ for all $a_0 \in \mathbb{Z}$.

Our most frequent application of 1 is when $g(x) = x$, in which case g divides $f^0 = x$. The polynomials f meeting the hypothesis of Theorem 1.1 fall into several families, and these are the subjects of our second main result. We remark that the proof of Theorem 1.1, part 1 generalizes easily to the setting where $\mathbb{Z}$ is replaced by the ring of integers in a number field, and there it may apply to more polynomials (see p. 13).

Theorem 1.2. *Suppose one of the following holds:*

(1) $f = x^2 - kx + k$ for some $k \in \mathbb{Z}$
(2) $f = x^2 + kx - 1$ for some $k \in \mathbb{Z}\backslash\{0, 2\}$
(3) $f = x^2 + k$ for some $k \in \mathbb{Z}\backslash\{-1\}$
(4) $f = x^2 - 2kx + k$ for some $k \in \mathbb{Z}\backslash\{\pm 1\}$.

Then $D(P(f, a_0)) = 0$ for all $a_0 \in \mathbb{Z}$.

We note that the recurrence given by $w_n = k + w_0 w_1 \cdots w_{n-1}$ satisfies $w_n = f(w_{n-1})$, where $f = x^2 - kx + k$, whence part 1 of Theorem 1.2 generalizes the sequence considered in [18]. Note also that it is not true that $D(P(f, a_0)) = 0$ for all monic, quadratic $f \in \mathbb{Z}[x]$. Indeed, if $f = (x - k)^2 + k$ with $k = \pm p$ for any prime $p \geq 3$, then one can show $D(P(f, 0)) = 1/3$. However, it seems likely that $D(P(f, a_0)) = 0$ for nearly all f whose critical point has infinite forward orbit (see Conjectures 1 and 2). The ineffectivity of Theorems 1.1 and 1.2 is due to the use of Siegel's theorem on finiteness of integral points on elliptic curves.

The method of proof of Theorems 1.1 and 1.2 revolves around a study of the Galois tower generated by the polynomials $g \circ f^n$ for $n = 1, 2, \ldots$. We focus on the Galois groups $H_n(f, g)$ of the extensions $\mathbb{Q}(\text{roots of } g \circ f^n)/\mathbb{Q}(\text{roots of } g \circ f^{n-1})$. To accomplish this, we amplify the techniques introduced in [12] from the theory of stochastic processes. This is the primary focus of sections 2 and 3, which build up to a proof of Theorem 1.1. One important step is Lemma 3.1, which shows that to establish $D(P(f, a_0)) = 0$, it is generally enough to prove that $H_n(f, g)$ is as large as possible for infinitely many n. By way of comparison, note that in [17] and [18] the results are achieved by showing $H_n(f, g)$ is as large as possible for all n. To apply Theorem 1.1 to specific families, we require results on the irreducibility of iterates of quadratic polynomials over $\mathbb{Q}$. This is the subject of section 4, and these results are apparently new. In section 5 we draw on the previous four sections to prove Theorem 1.2.

A key step in the proof of Theorem 1.1 is Theorem 3.3, which shows that $H_n(f, g)$ is as large as possible only when the set $\{g(f^i(\gamma)) : i = 1, 2, \ldots, n\}$ satisfies certain arithmetic properties, where γ is the critical point of f. Thus the density of $P(f, a_0)$ depends on arithmetic properties of the critical orbit $\{g(f^i(\gamma)) : i = 1, 2, \ldots\}$ of f. This makes for a striking analogy with complex and real dynamics, where analytic properties of the critical orbit of a quadratic polynomial have been shown to determine fundamental dynamical behavior of the polynomial. For instance, if $f \in \mathbb{C}[z]$ is quadratic, then membership in the Mandelbrot set – and equivalently the connectedness of the filled Julia set of f – is determined by whether the critical orbit remains bounded [5, Section 3.8]. If $f \in \mathbb{R}[x]$ is of the form $a - x^2$ for

$-1/4 \le a \le 2$ the Collet-Eckmann condition states that if $|f'|$ grows exponentially along the critical orbit, then f exhibits stochastic behavior on an invariant interval. An important question is whether almost all maps in this family satisfy the Collet-Eckmann condition; this question was recently resolved in the affirmative by Avila and Moreira [2].

2. Galois processes and eventual martingales

Our main technique is to relate the density of $P(g, f, a_0)$ to properties of the Galois groups of $g \circ f^n$ over $\mathbb{Q}$. In order to do this, we make use of the notion of *Galois processes* [12].

Let K be a field and $g, f \in K[x]$ such that $g \circ f^n$ is separable for all $n \ge 0$ (we take f^0 to be the identity map). Denote the splitting field of $g \circ f^n$ over K by $K(g \circ f^n)$. Clearly $K(g \circ f^n) \supseteq K(g \circ f^{n-1})$. We now let $GP(f, g)$ be the Galois process of this tower, as defined in [12, page 13]; for completeness we give the construction in the present case. Define $G_n(f, g)$ to be $\mathrm{Gal}\,(K(g \circ f^n)/K)$ and $H_n(f, g)$ to be $\mathrm{Gal}\,(K(g \circ f^n)/K(g \circ f^{n-1}))$. Let $G(f, g) = \varprojlim G_n(f, g)$, and take $\mathcal{P}$ to be a Harr measure on $G(f, g)$ with $\mathcal{P}(G(f, g)) = 1$. Letting $\mathcal{B}$ be the Borel sigma algebra, the triple $(G(f, g), \mathcal{P}, \mathcal{B})$ is then a probability space. Denote by ψ_n the natural projection $G(f, g) \to G_n(f, g)$, and define random variables X_n on $G(f, g)$ as follows:
$$X_n(\sigma) = \text{number of roots of } g \circ f^n \text{ fixed by } \psi_n(\sigma).$$
The data $(G(f, g), \mathcal{P}, \mathcal{B}, \{X_n\}_{n \ge 0})$ by definition give a stochastic process, which we call the Galois process of $g \circ f^n$, and denote $GP(f, g)$. Intuitively, this process resembles a random walk through successively larger quotients of $G(f, g)$ that is compatible with the natural restriction maps. Positions at each level are assigned a value based on the number of roots of $g \circ f^n$ left fixed. Note that it follows from basic properties of Haar measure that

$$(3) \quad \mathcal{P}(X_0 = t_0, \ldots, X_n = t_n) =$$
$$\frac{1}{\#G_n(f, g)} \# \left\{ \sigma \in G_n(f, g) : \sigma \text{ fixes } t_i \text{ roots of } g \circ f^i \text{ for } i = 0, 1, \ldots, n \right\}.$$

The following result links $D(P(f, a_0))$ to Galois processes. We wish to consider f that may have reducible iterates, and thus it becomes necessary to deal with translates of recurrences. Given a recurrence a_n with terms in K and $g \in K[x]$, we refer to the sequence $g(a_n)$ as a *polynomial translate* of a_n.

Theorem 2.1. *Let $f, g \in \mathbb{Z}[x]$ be polynomials with $g \circ f^n$ separable for all n. Let $a_n = f^n(a_0)$ with $a_0 \in \mathbb{Z}$. The density of prime divisors of $g(a_n) : n = 0, 1, 2, \ldots$ is bounded above by*
$$\lim_{n \to \infty} \mathcal{P}(X_n > 0),$$
where X_n is the nth random variable in $GP(f, g)$.

Remark. *It is also true that if $g = \mathrm{id}$, then $\lim_{n \to \infty} \mathcal{P}(X_n > 0)$ furnishes an upper bound for the density of primes p such that 0 is periodic in $\mathbb{Z}/p\mathbb{Z}$ under iteration of f. This follows from the fact that 0 is periodic in $\mathbb{Z}/p\mathbb{Z}$ if and only if $f^{-n}(0) \cap \mathbb{Z}$ is non-empty for all $n \ge 1$; cf [12, Proposition 3.1]. Thus Theorem 1.2 shows that for the families in question, the density of p with 0 periodic in $\mathbb{Z}/p\mathbb{Z}$ under f is zero.*

Remark. *With trivial modifications to the proof the Theorem can be made to apply with $\mathbb{Z}$ replaced by the ring of integers $\mathcal{O}_K$ in a number field K.*

Proof. Let

$$\Omega_N = \{p : p \nmid \operatorname{Disc}(g \circ f^N) \text{ and } g(f^N(x)) \equiv 0 \bmod p \text{ has no solution in } \mathbb{Z}\}.$$

If $p \in \Omega_N$, then clearly $g(f^n(a_0)) \not\equiv 0 \pmod{p}$ for all $n \geq N$. Thus $p \nmid g(a_n)$ for all $n \geq N$. There are only finitely many p with $p \mid g(a_n)$ for some $n < N$, and also only finitely many dividing $\operatorname{Disc}(g \circ f^N)$. Therefore

$$(4) \qquad D(\Omega_N) \leq D(\{p : p \nmid g(a_n) \text{ for all } n \geq 1\}).$$

If $p \nmid \operatorname{Disc}(g \circ f^N)$, then p cannot divide the discriminant of the splitting field $\mathbb{Q}(g \circ f^N)$ of $g \circ f^N$ [16, Corollary 2, p. 157], so p is unramified in $\mathbb{Q}(g \circ f^N)$. Now $g(f^N(x)) \equiv 0 \pmod{p}$ having a solution in $\mathbb{Z}$ is equivalent to $g(f^N(x))$ having at least one linear factor in $(\mathbb{Z}/p\mathbb{Z})[x]$. Except for possibly finitely many p, this implies that $p\mathcal{O}_K = \mathfrak{P}_1 \cdots \mathfrak{P}_r$, where $K/\mathbb{Q}$ is obtained by adjoining a root of $g \circ f^N$, $\mathcal{O}_K$ is the ring of integers, and at least one of the $\mathfrak{P}_i$ has residue class degree one [16, Theorem 4.12]. This is equivalent to the disjoint cycle decomposition of the Frobenius conjugacy class at p having a fixed point (in the natural permutation representation of $G_N(f, g) = \operatorname{Gal}(\mathbb{Q}(g \circ f^N)/\mathbb{Q})$ on the roots of $g \circ f^N$). From the Chebotarev Density Theorem it follows [16, Proposition 7.15] that the density of p with $p\mathcal{O}_K$ having such a decomposition is

$$\frac{1}{\#G_N(f,g)} \#\{\sigma \in G_N(f,g) : \sigma \text{ fixes at least one root of } g \circ f^N\},$$

and by (3) this is simply $\mathcal{P}(X_N > 0)$. Thus $\mathcal{P}(X_N > 0) = 1 - D(\Omega_N)$. From (4), we now have

$$\mathcal{P}(X_N > 0) \quad \geq \quad 1 - D(\{p : p \nmid g(a_n) \text{ for all } n \geq 1\}),$$

and this last expression is the same as $D(\{p : p \mid g(a_n) \text{ for some } n \geq 1\})$. $\qquad\square$

In light of this result, we study the long-term behavior of $GP(f, g)$; this occupies the remainder of this section and also the next section. Our main results show that $GP(f, g)$ has certain convergence properties, and require the following definition:

Definition. *A stochastic process $X_0, X_1, X_2, \ldots$ taking values in $\mathbb{Z}$ is a* martingale *if for all $n \geq 1$ and any $t_i \in \mathbb{Z}$,*

$$E(X_n \mid X_0 = t_0, X_1 = t_1, \ldots, X_{n-1} = t_{n-1}) = t_{n-1}.$$

We call $X_0, X_1, X_2, \ldots$ an eventual martingale *if for some N the process $X_N, X_{N+1}, X_{N+2}, \ldots$ is a martingale.*

Martingales are important chiefly because they often converge in the following sense:

Definition. *Let $X_1, X_2, \ldots$ be a stochastic process defined on the probability space $(\Omega, \mathcal{P}, \mathcal{F})$. The process* converges *if*

$$\mathcal{P}\left(\omega \in \Omega : \lim_{n \to \infty} X_n(\omega) \text{ exists}\right) = 1.$$

Note that convergence does not depend on any finite collection of X_n. We give one martingale convergence theorem (see e.g. [9, Section 12.3] for a proof)

Theorem 2.2. *Let $M = (X_1, X_2, \ldots)$ be a martingale whose random variables take nonnegative real values. Then M converges.*

Since the random variables in $GP(f, g)$ take nonnegative integer values, we immediately have the following:

Corollary 2.3. *Suppose that $GP(f, g)$ is an eventual martingale. Then*

$$\mathcal{P}(\{\sigma \in G(f, g) : X_1(\sigma), X_2(\sigma), \ldots \text{ is eventually constant}\}) = 1.$$

A fruitful point of view of the groups $G_n(f, g)$ is as automorphism groups of rooted forests. A rooted forest is a disjoint union of trees, each with a distinguished vertex called its root. There is a natural filtration of the forest provided by the distance to a root. We say a rooted forest is *complete of degree d* if each vertex is connected to precisely d others, and that the forest has height h if the length of the longest path is h. We may form a rooted forest from the polynomials $g \circ f^n$ as follows: take the roots (in the polynomial sense) of g to be V_0, the set of roots (in the graph sense) of the forest, and for each $n \geq 1$, let V_n be the set of roots of $g \circ f^n$. Two vertices $v_{n-1} \in V_{n-1}$ and $v_n \in V_n$ are connected if $f(v_n) = v_{n-1}$. For each root α of g, the set $\{v : f^i(v) = \alpha \text{ for some } i\}$ forms a tree with root α. It is easy to check that this defines a rooted forest.

A group G acts on a rooted forest with vertices V if there is an action of G on the set V such that for all $\sigma \in G$, v is connected to v' if and only if $\sigma(v)$ is connected to $\sigma(v')$.

Lemma 2.4. *Let G be a group acting on a rooted forest of height n, and denote the vertices of distance i from a root by V_i. Let $v_0 \in V_{n-1}$, and suppose the subgroup $H = \{\sigma \in G : \sigma(v) = v \text{ for all } v \in V_{n-1}\}$ acts transitively on $E_{v_0} = \{v \in V_n : v \text{ is connected to } v_0\}$. Finally, let $\sigma \in G$ satisfy $\sigma(v_0) = v_0$. Then*

$$(5) \qquad \frac{1}{\#H} \sum_{\tau \in H} \#\{v \in E_{v_0} : \sigma\tau(v) = v\} = 1.$$

Proof. Since $\sigma(v_0) = v_0$, each element of σH permutes E_{v_0}. Note that

$$\sum_{\tau \in H} \#\{v \in E_{v_0} : \sigma\tau(v) = v\} = \sum_{\tau \in H} \sum_{v \in E_{v_0}} \epsilon(v, \tau),$$

where $\epsilon(v, \tau) = 1$ if $\sigma\tau(v) = v$ and $\epsilon(v, \tau) = 0$ otherwise. Interchanging the sums yields

$$(6) \qquad \sum_{v \in E_{v_0}} \#\mathrm{Stab}_{\sigma H}(v).$$

Now given $v \in E_{v_0}$, by the transitivity of H we can choose $\tau' \in H$ with $\tau'(v) = \sigma^{-1}(v)$. We then have $\sigma\tau(v) = v$ if and only if $\tau'^{-1}\tau(v) = v$. Thus

$$\#\mathrm{Stab}_{\sigma H}(v) = \#\{\tau \in H : \tau'^{-1}\tau(v) = v\} = \#\mathrm{Stab}_H(v).$$

By standard group theory and the transitivity of H, we have $\#\mathrm{Stab}_H(v) = \#H/\#E_{v_0}$. Hence the expression in (6) is $\#H$. $\qquad\qquad \square$

Theorem 2.5. *Let $f, g \in K[x]$ be such that $g \circ f^n$ is separable for all $n \geq 0$. Suppose that there exists n_0 such that for all $n \geq n_0$ and every root α of $g \circ f^{n-1}$, the polynomial $f(x) - \alpha$ is irreducible over the splitting field $K(g \circ f^{n-1})$ of $g \circ f^{n-1}$. Then $GP(f, g)$ is an eventual martingale.*

Proof. By the definition of eventual martingale and conditional expectation, we must show that there is an N such that for $n > N$, the expression

(7) $$\sum_k k \cdot \frac{\mathcal{P}(X_N = t_N, \ldots, X_{n-1} = t_{n-1}, X_n = k)}{\mathcal{P}(X_N = t_N, \ldots, X_{n-1} = t_{n-1})}$$

is equal to t_{n-1}, for any fixed values of the t_i. We take N to be n_0. Put

$$
\begin{aligned}
S &= \{\sigma \in G_n(f, g) : \sigma \text{ fixes } t_i \text{ roots of } g \circ f^i \text{ for } N \le i \le n-1\} \\
S_k &= \{\sigma \in S : \sigma \text{ fixes } k \text{ roots of } g \circ f^n\}
\end{aligned}
$$

From the basic property of $\mathrm{GP}(f, g)$ given in (3), the expression in (7) is equal to

$$\sum_k k \cdot \frac{\#S_k}{\#S}$$

This in turn may be rewritten as

(8) $$\frac{1}{\#S} \sum_{\sigma \in S} (\text{number of roots of } g \circ f^n \text{ fixed by } \sigma).$$

We denote $H_n(f, g)$ by H. Note that if $\tau \in H$, then τ fixes the roots of $g \circ f^i$ for $N \le i \le n-1$. Thus S is invariant under multiplication by H, whence S is a union of cosets of H. Let $\sigma H \subseteq S$, and let T be the complete binary infinite rooted forest corresponding to the translated iterates $g \circ f^n$. Note that if T_n denotes the height n rooted forest consisting of the first n levels of T, then $G_n(f, g)$ acts as automorphisms of T_n and H is the subgroup stabilizing the vertices of height $n-1$.

Let α_0 be a vertex of height $n-1$, and suppose that $\sigma(\alpha_0) = \alpha_0$. We have the following factorization of $g \circ f^n$ over $K(g \circ f^{n-1})$:

(9) $$g \circ f^n = \prod_{\alpha \text{ root of } g \circ f^{n-1}} f(x) - \alpha.$$

Thus the irreducibility of $f(x) - \alpha_0$ implies that H acts transitively on the set $E_{\alpha_0} = \{\beta \text{ root of } g \circ f^n : f(\beta) = \alpha_0\}$, which is the set of vertices of height n connected to α_0. By Lemma 2.4, we have

$$\frac{1}{\#H} \sum_{\tau \in H} \#\{\beta \in E_{\alpha_0} : \sigma\tau(\beta) = \beta\} = 1.$$

On the other hand, if α is a vertex of height $n-1$ and $\sigma(\alpha) = \alpha'$ for some $\alpha' \ne \alpha$ then σ maps the roots of $f(x) - \alpha$ to the roots of $f(x) - \alpha'$, and hence fixes no roots of $f(x) - \alpha$. Since $\sigma \in S$, we have that σ fixes t_{n-1} roots of $g \circ f^{n-1}$, so it follows that

$$\frac{1}{\#H} \sum_{\tau \in H} \#\{\text{roots of } g \circ f^n \text{ fixed by } \sigma\tau\} = t_{n-1}.$$

Since S is a union of cosets of H, the expression in (8) equals t_{n-1}. $\square$

Lemma 2.6. *Let $f, g \in K[x]$ where f and g have respectively degrees d_f, d_g and leading coefficients a, b, and let $\gamma_1, \gamma_2, \ldots, \gamma_{d-1}$ be the critical points of f. Put $\Delta_n = \mathrm{Disc}\,(g \circ f^n)$. Then for all $n \ge 1$ we have*

$$\Delta_n = \pm \Delta_{n-1}^{d_f} d_f^{k_1} a^{k_2} b^{k_3} \prod_{i=1}^{d-1} g(f^n(\gamma_i)),$$

where $k_1 = d_g d_f^n$, $k_2 = d_g^2 d_f^{2n-1} - d_g$, and $k_3 = d_f - 1$.

Proof. We use the definition of discriminant via resultants. Throughout, we denote by $\ell(h)$ the leading coefficient of the polynomial h. Let $h_1, h_2 \in \mathbb{Z}[x]$ be two polynomials with degrees δ_1, δ_2, and let α_i, $i = 1, \ldots, \delta_1$, denote the roots of h_1 and β_j, $j = 1, \ldots, \delta_2$, denote the roots of h_2. The resultant may be defined as $R(h_1, h_2) = \ell(h_1)^{\delta_2}\ell(h_2)^{\delta_1} \prod(\alpha_i - \beta_j)$, or equivalently (and more usefully for us) as

$$\ell(h_1)^{\delta_2} \prod_{h_1(\alpha)=0} h_2(\alpha).$$

Note that the product is taken over roots with multiplicity. The discriminant $\Delta(h_1)$ of h_1 then satisfies

$$(10) \qquad\qquad R(h_1, h_1') = (-1)^{\delta_1(\delta_1-1)/2}\ell(h_1)\Delta(h_1);$$

see for instance [14, p.204]. From the definition it follows immediately that $R(h_1, h_2) = (-1)^{\delta_1\delta_2}R(h_2, h_1)$ and $R(h_1 h_2, h_3) = R(h_1, h_3)R(h_2, h_3)$. With slightly more work, one obtains

$$(11) \qquad\qquad R(h_1 \circ h_2, h_1' \circ h_2) = \ell(h_2)^{(\delta_1^2-\delta_1)\delta_2}R(h_1, h_1')^{\delta_2}.$$

From (10) we have

$$(12) \qquad \left| \frac{\Delta_n}{(\Delta_{n-1})^{d_f}} \right| = \left| \frac{\ell(g \circ f^{n-1})^{d_f}}{\ell(g \circ f^n)} \cdot \frac{R(g \circ f^n, (g \circ f^n)')}{R(g \circ f^{n-1}, (g \circ f^{n-1})')^{d_f}} \right|,$$

Writing $g \circ f^n$ as $(g \circ f^{n-1}) \circ f$, we have $(g \circ f^n)' = f' \cdot (g \circ f^{n-1})'(f)$. Thus

$$\begin{aligned} |R(g \circ f^n, (g \circ f^n)')| &= |R((g \circ f^n)', g \circ f^n)| \\ &= \left|R(f', g \circ f^n) \cdot R((g \circ f^{n-1})'(f), g \circ f^n)\right|. \end{aligned}$$

Using (11) and the fact that the degree of $g \circ f^{n-1}$ is $d_g d_f^{n-1}$, this last expression becomes

$$|a|^{(d_g d_f^{n-1}-1)d_g d_f^n} \left|R(f', g \circ f^n) \cdot (R(g \circ f^{n-1}, (g \circ f^{n-1})')^{d_f}\right|.$$

The right side of (12) then becomes

$$\left| \frac{\ell(g \circ f^{n-1})^{d_f}}{\ell(g \circ f^n)} \right| \cdot |a|^{(d_g d_f^{n-1}-1)d_g d_f^n} \cdot |R(f', g \circ f^n)|.$$

From the definition of resultant we have $R(f', g \circ f^n) = (ad_f)^{d_g d_f^n} \prod_{i=1}^{d-1} g(f^n(\gamma_i))$. Hence the powers of a, b, and d_f in the formula of the Lemma are those in the expression

$$a^{(d_g d_f^{n-1}-1)d_g d_f^n}(ad_f)^{d_g d_f^n} \frac{\ell(g \circ f^{n-1})^{d_f}}{\ell(g \circ f^n)}.$$

Now $\ell(g \circ f^n) = ba^{c_1}$, where $c_1 = d_g(d_f^{n-1} + d_f^{n-2} + \cdots + 1)$, while $\ell(g \circ f^{n-1}) = ba^{c_2}$, where $c_2 = d_g(d_f^{n-2} + d_f^{n-3} + \cdots + 1)$. Hence

$$\frac{\ell(g \circ f^{n-1})^{d_f}}{\ell(g \circ f^n)} = b^{d_f-1}a^{-d_g}.$$

The formulas in the Lemma follow immediately. $\qquad\qquad\qquad\qquad\qquad\qquad\square$

One consequence of Lemma 2.6 is that $g \circ f^n$ is separable for all n if and only if g is separable and 0 is not contained in the forward orbit of any critical point of f.

We also remark that in the case $\deg(f) = 2$ and g is monic and linear, Lemma 2.6 gives

$$\Delta_n = \pm\Delta_{n-1}^2 2^{2^n} a^{2^{2n-1}-1} g(f^n(\gamma)),$$

where γ is the unique critical point of f.

We call a polynomial $f \in K[x]$ *critically finite* if all of its critical points have finite forward orbits, or equivalently are eventually periodic. In particular, a quadratic polynomial $f = ax^2 + bx + c$ is critically finite if and only if the forward orbit of $-b/2a$ is finite. We remark that it follows from Lemma 2.6 and the appendix to this article that the compositum of $\mathbb{Q}(g \circ f^n), n = 1, 2, \ldots$ is finitely ramified if and only if f is critically finite (cf [1]).

We now specialize to the case where f is quadratic and K is a number field.

Theorem 2.7. *Let K be a number field, $f, g \in K[x]$ with f quadratic, and suppose $G_0(f, g) = \text{Gal}\,(K(g)/K)$ is a 2-group. Suppose further that f is critically infinite, and $g \circ f^n$ is irreducible for all $n \geq 0$. Then $GP(f, g)$ is an eventual martingale.*

Remark. *For the conclusion of Theorem 2.7 to hold, we only need the sequence $\{g(f^n(\gamma)) : n = 1, 2, \ldots\}$ to contain only finitely many squares. This happens for many critically finite polynomials, e.g. $x^2 - 2$ and $(x - k)^2 + k$ for all nonsquare $k \in \mathbb{Z}$.*

Proof. Let K_n be the splitting field of $g \circ f^n$ over K, let V_n be the roots of $g \circ f^n$, and let $\mathfrak{C}$ be the partition of V_n into the sets $\{$roots of $f(x) - \alpha\}$, where α varies over roots of $g \circ f^{n-1}$. Let $\sigma_{\mathfrak{C}} \in \text{Sym}(V_n)$ be the permutation associated to $\mathfrak{C}$, i.e. the unique permutation whose orbits are precisely the sets belonging to $\mathfrak{C}$. We wish to show that $\sigma_{\mathfrak{C}} \in G_n(f, g)$ for n sufficiently large.

For all $n \geq 1$, it follows from (9) that the field K_n is obtained from K_{n-1} by adjoining square roots. Since $G_0(f, g)$ is a 2-group, it follows that $G_n(f, g)$ is a 2-group as well, and thus has non-trivial center. Let δ be a nontrivial element of $Z(G_n(f, g))$, and $\mathfrak{D}$ the corresponding central fiber system (see [12, Proposition-Definition 4.10]). Then by [12, Theorem 4.9] we have either $\sigma_{\mathfrak{C}} = \delta$ or $G_n(f, g)$ is composed entirely of alternating permutations.

In the latter case, $\Delta_n = \text{Disc}\,(g \circ f^n)$ is a square in K. By Lemma 2.6 this implies that $sg(f^n(\gamma))$ is a square in K, where γ is the critical point of f and s satisfies $sb \in K^{*2}$ if g has even degree and $sab \in K^{*2}$ if g has odd degree (as in Lemma 2.6, a and b are the leading coefficients of f and g, respectively). Now for $n \geq 2$, this implies the curve $C : y^2 = sg(f^2(x))$ has a pair of points that are S-integers, indeed a pair with $x = f^{n-2}(\gamma)$ (we may take S to be the primes in $\mathcal{O}_K$ dividing either 2 or the denominator of one of the coefficients of f). Since f is critically infinite, $f^k(\gamma) \neq f^j(\gamma)$ for all j, k, so that the pairs obtained from different n are all distinct. Since $\deg(g \circ f^2) \geq 4$, and $g \circ f^2$ is irreducible and thus separable, the genus of C is at least 1, whence by Siegel's Theorem [11, p.353] C has only finitely many S-integral points. Thus $g(f^n(\gamma))$ is a square for only finitely many n. Denote the largest such n by n', and take $n_0 = n' + 1$. Thus for $n \geq n_0$ we have shown that $\sigma_{\mathfrak{C}} = \delta$, and it follows that $\sigma_{\mathfrak{C}} \in G_n(f, g)$. Indeed, $\sigma_{\mathfrak{C}} \in H_n(f, g) \stackrel{\text{def}}{=} \text{Gal}\,(K_n/K_{n-1})$, whence $H_n(f, g)$ is nontrivial.

Therefore by (9) there is some root α of $g \circ f^{n-1}$ such that $f(x) - \alpha$ is irreducible over K_{n-1}. Since $g \circ f^{n-1}$ is irreducible over $\mathbb{Q}$, the action of $G_{n-1}(f, g)$ is transitive on the roots of $g \circ f^{n-1}$, and it follows that $f(x) - \alpha$ is irreducible over K_{n-1} for all

α. The irreducibility of $g \circ f^n$ over K implies separability, and we now have from Theorem 2.5 that $GP(f, g)$ is an eventual martingale. $\square$

3. Density results

In this section we prove Theorem 1.1. Note that by Theorem 2.7, the hypotheses of Theorem 1.1 imply that $GP(f, g)$ is an eventual martingale. In order to parlay this into a zero-density result, we make use of the following lemma. We remark that it follows from the factorization in (9) that the extension $\mathbb{Q}(g \circ f^n)/\mathbb{Q}(g \circ f^{n-1})$ is the compositum of at most $\deg(g \circ f^{n-1})$ quadratic extensions. Thus $H_n(f, g) \cong (\mathbb{Z}/2\mathbb{Z})^m$ for some $0 \le m \le \deg(g \circ f^{n-1})$. If $m = \deg(g \circ f^{n-1})$, then we call $H_n(f, g)$ *maximal*.

Lemma 3.1. *If $GP(f, g)$ is an eventual martingale and $H_n(f, g)$ is maximal for infinitely many n, then $GP(f, g)$ converges to 0, i.e.*

$$\lim_{n \to \infty} \mathcal{P}(X_n = 0) = 1.$$

Proof. The proof of [12, Lemma 5.2] requires only straightforward adaptation to apply to $GP(f, g)$. The proof of [12, Theorem 1.3] only uses [12, Lemma 5.2] and that the stochastic process is eventually constant with probability 1, which follows from Corollary 2.3. $\square$

We now give two results that relate the maximality of $H_n(f, g)$ to the translated critical orbit $\{g(f^n(\gamma)) : n \ge 1\}$.

Lemma 3.2. *Let K be a number field and $f, g \in K[x]$ with $f = ax^2 + bx + c$, and let γ be the critical point of f. Suppose that $g \circ f^n$ is irreducible over K for all $n \ge 1$. Then for $n \ge 2$, $H_n(f, g)$ is maximal if and only if $g(f^n(\gamma))$ is not a square in $K(g \circ f^{n-1})$.*

Proof. We require only minor modifications of [22, Lemma 1.6]. Let $d = \deg(g \circ f^{n-1})$, $K_n = K(g \circ f^n)$, $K_{n-1} = K(g \circ f^{n-1})$, and denote the roots of $g \circ f^{n-1}$ by $\alpha_j, j = 1, \ldots, d$. From (9), we have that K_n/K_{n-1} is a 2-Kummer extension obtained by adjoining to K_{n-1} the square roots of $b^2 - 4ac + 4a\alpha_j$ for $j = 1, \ldots, d$. Thus $[K_n : K_{n-1}] = 2^{d - \dim V}$, where V is the $\mathbb{F}_2$-vector space of all $(\epsilon_1, \ldots, \epsilon_d) \in \mathbb{F}_2^d$ such that $\prod_{j=1}^d (b^2 - 4ac + 4a\alpha_j)^{\epsilon_j}$ is a square in K_{n-1}. By the irreducibility of $g \circ f^n$, we have that $\mathrm{Gal}(K_n/K)$ acts transitively on the α_j. We now follow precisely the reasoning in the proof of [22, Lemma 1.6] to show that $V \ne 0$ if and only if $(1, \ldots, 1) \in V$. Therefore $H_n(f, g)$ is maximal if and only if $\prod_{j=1}^d (b^2 - 4ac + 4a\alpha_j)$ is a square in K_{n-1}. Since $n \ge 2$, we have d even, and thus $H_n(f, g)$ is maximal if and only if

$$\prod_{j=1}^d \left(\alpha_j - \frac{b^2}{4a} + c \right)$$

is a square in K_{n-1}. The same argument as in the proof of Proposition 4.2 (p. 13) shows that this is equivalent to $g(f^n(\gamma))$ being a square in K_{n-1}. $\square$

Theorem 3.3. *Let K be a number field with ring of integers $\mathcal{O}_K$, and let $f, g \in \mathcal{O}_K[x]$ with $f = ax^2 + bx + c$, and let γ be the critical point of f. Suppose that $g \circ f^n$ is irreducible over K for all $n \ge 1$, and denote by $v_\mathfrak{p}$ the $\mathfrak{p}$-adic valuation for $\mathfrak{p}$ a prime in $\mathcal{O}_K$. If $n \ge 2$ and there exists $\mathfrak{p}$ with $v_\mathfrak{p}(g(f^n(\gamma)))$ odd, $v_\mathfrak{p}(g(f^m(\gamma))) = 0$ for all $1 \le m \le n - 1$, and $v_\mathfrak{p}(2a) = 0$ then $H_n(f, g)$ is maximal.*

Proof. It follows from Lemma 2.6 and [16, Corollary 2, p. 157] that the only primes ramifying in $K(g \circ f^{n-1})$ are a subset of those dividing $2a$ and those dividing the numerator of $g(f^m(\gamma))$ for some $0 \le m \le n-1$. Thus if $v_{\mathfrak{p}}(g(f^m(\gamma))) = 0$ for all $1 \le m \le n-1$, and $v_{\mathfrak{p}}(2a) = 0$ then $\mathfrak{p}$ is unramified in $K(g \circ f^{n-1})$. Hence if $v_{\mathfrak{p}}(g(f^n(\gamma)))$ is odd, then $g(f^n(\gamma))$ cannot be a square in $K(g \circ f^{n-1})$. By Lemma 3.2 we then have $H_n(f,g)$ maximal. $\qquad\square$

As an illustration of Theorem 3.3, consider the case of $K = \mathbb{Q}$, $f(x) = x^2 + 5$, $g(x) = x$. We have $f(\gamma) = 5$, $f^2(\gamma) = 2 \cdot 3 \cdot 5$, $f^3(\gamma) = 5 \cdot 181$, and $f^4(\gamma) = 2 \cdot 3 \cdot 5 \cdot 23 \cdot 1187$. In each case there is a $p \neq 2$ with $v_p(f^n(\gamma)) = 1$ but $v_p(f^m(\gamma)) = 0$ for all $1 \le m \le n-1$, and therefore $H_n(f)$ is maximal for $1 \le n \le 4$.

Theorem 3.3 says that the maximality of $H_n(f,g)$ can be ascertained by examining arithmetic properties of the critical orbit. This makes for a striking analogy with complex and real dynamics, where analytic properties of the critical orbit of a quadratic polynomial have been shown to determine fundamental dynamical behavior of the polynomial. See the last paragraph of the Introduction for more details.

We make one final remark before proving Theorem 1.1. Consider $f \in \mathbb{Z}[x]$ quadratic, and let $g(x) = x$. To analyze the arithmetic of the critical orbit of f, we split the primes dividing some $f^n(\gamma)$ into two classes. We say that p is *isolated* if $v_p(f^n(\gamma)) > 0$ for some n and there is no other n' with $v_p(g(f^{n'}(\gamma))) > 0$. When we reduce the coefficients of f modulo p and iterate, this behavior corresponds to 0 being in the orbit of γ in $\mathbb{Z}/p\mathbb{Z}$ but not in the periodic part. If p is not isolated, then there are infinitely many n' with $v_p(f^{n'}(\gamma)) > 0$, and we call p *recurrent*. This corresponds to 0 being in the periodic part of the orbit of γ in $\mathbb{Z}/p\mathbb{Z}$. Heuristics and computation suggest that most primes p dividing some element of the critical orbit are isolated, but in general there appear to be infinitely many recurrent primes, and their behavior seems difficult to control. There is one exception, namely when all primes dividing an element of the translated critical orbit are recurrent. This is the case when $\gamma = 0$, e.g. for $f(x) = x^2 + k$, $k \in \mathbb{Z}$. At the other extreme are f such that all but finitely many primes appearing in the critical orbit are isolated. This is the case when $\{f^n(0) : n \ge 1\}$ is finite and does not contain 0, as in the case of $f = x^2 - kx + k$, for which 0 maps to the fixed point k. These two extremes correspond to parts (2) and (1), respectively, of Theorem 1.1.

We now give the proof of Theorem 1.1:

Proof. We apply Theorem 2.7 to get that $GP(f)$ is an eventual martingale. Then by Lemma 3.1 and Theorem 2.1 we need only show that $H_n(f,g)$ is maximal for infinitely many n. Suppose first that we are in case 1, and let $\{f^n(0) : n = 1, 2, \ldots\} = \{b_1, \ldots, b_k\}$ and $\{g(f^n(0)) : n = 1, 2, \ldots\} = \{\beta_1, \ldots, \beta_{k'}\}$. We remark that by hypothesis $g \mid f^r$, whence $g(f^n(0)) \mid f^{n+r}(0)$ for each n. This means β_i divides some b_j, and in particular the assumption that the b_i are all nonzero implies the β_i are as well. The outline of the argument is this: it follows from Siegel's theorem on S-integral points that for infinitely many n, the rational number $|g(f^n(\gamma))/d|$ is not a square for any $d \mid \prod_i b_i$. We use this fact to show that there must exist p with $v_p(g(f^n(\gamma)))$ odd and $v_p(g(f^j(\gamma))) = 0$ for all $j < n$. It follows from Theorem 3.3 that H_n is maximal, and this proves the theorem.

To apply Siegel's theorem, we assume there are infinitely many n with $g(f^n(\gamma)) > 0$; the argument is similar if there are infinitely many n with $g(f^n(\gamma)) < 0$. If

$g(f^n(\gamma))/|d|$ is a square for $n \geq 2$, then the curve $C_d : |d|y^2 = g(f^2(x))$ has a pair of S-integral points ($S = \{2\}$) with $x = f^{n-2}(\gamma)$. By assumption the set $\{g(f^n(\gamma))\}$ is infinite, so all points obtained in this way are distinct. Since $g(f^2(x))$ is irreducible and hence separable, it follows that C_d has genus at least 1. From Siegel's Theorem on S-integral points [11, p. 353] and the fact that there are only finitely many values for d (recall d varies over the divisors of $\prod_i b_i$), we have that for all n large enough, $g(f^n(\gamma))/|d|$ is not a square for any d. Since $g(f^n(\gamma)) > 0$ for infinitely many n, we have for these n that $|g(f^n(\gamma))/d|$ is not a square for any d, as desired.

Suppose for a moment that $\gamma \in \mathbb{Z}$. We now show that for all n large enough, there exists $d \mid \prod_i b_i$ such that $g(f^n(\gamma))/d$ is an integer not divisible by any primes $p \mid \prod_i b_i$. This is equivalent to showing that $v_p(g(f^n(\gamma))) \leq v_p(\prod_i b_i)$ for all $p \mid \prod_i b_i$. Momentarily fix a prime p dividing $\prod_i b_i$, and let $m = \max_i v_p(b_i)$. Suppose there is some n_0 with $v_p(g(f^{n_0}(\gamma))) > m$. By assumption $g \mid f^r$, say $gh = f^r$. Then multiplying through by $h(f^{n_0}(\gamma))$ gives $f^{n_0+r}(\gamma) \equiv 0 \bmod p^{m+1}$. Thus for any $j > n_0 + r$ we have

$$f^j(\gamma) \equiv f^{j-(n_0+r)}(f^{n_0+r}(\gamma)) \equiv b_i \not\equiv 0 \bmod p^{m+1}$$

for some i. Since each β_i divides some b_i, none of the β_i can vanish modulo p^{m+1}. Thus we have $g(f^j(\gamma)) \equiv \beta_{i'} \not\equiv 0 \bmod p^{m+1}$ for some i'. We obtain in this way an n_0 for each $p \mid \prod_i b_i$, and for any n larger than their max we have $v_p(g(f^n(\gamma))) \leq m \leq v_p(\prod_i b_i)$ for all $p \mid \prod_i b_i$. If $\gamma \notin \mathbb{Z}$, then we wish to show that for all n large enough there exists an *odd* $d \mid \prod_i b_i$ such that the numerator of $g(f^n(\gamma))/d$ is an integer not divisible by any primes $p \mid \prod_i b_i$ (the denominator of $g(f^n(\gamma))/d$ is an even power of 2). We proceed exactly as before, noting that the numerator of $g(f^n(\gamma))$ is odd, and thus d can be taken to be odd as well.

From the previous two paragraphs it follows that for infinitely many n, there exists a prime p, not dividing any of the b_i, such that $v_p(g(f^n(\gamma)))$ is odd. Therefore writing $gh = f^r$ and multiplying by $h(f^n(\gamma))$ we have $f^{n+r}(\gamma) \equiv 0 \bmod p$, so for each $j > n+r$ we have $f^j(\gamma) \equiv b_i \not\equiv 0 \bmod p$ for some i. If $v_p(g(f^l(\gamma))) > 0$ for some $l > n$, then multiplying by $h(f^l(\gamma))$ gives $f^{l+r}(\gamma) \equiv 0 \bmod p$, a contradiction. It follows that $v_p(g(f^l(\gamma))) = 0$ for all $l < n$ as well, since otherwise $v_p(g(f^n(\gamma))) = 0$, a contradiction (hence p is isolated in the terminology of the discussion preceding this proof). It remains to show we can take n large enough to ensure $p \neq 2$. If $\gamma \notin \mathbb{Z}$, then the numerator of $f^n(\gamma)$ is odd and for all n, whence the same is true for $g(f^n(\gamma))$, since g divides f^r. In this situation $v_2(g(f^n(\gamma))$ is never odd. If $\gamma \in \mathbb{Z}$ then either $2 \mid \prod_i b_i$ or there is at most one n with $v_2(g(f^n(\gamma))) > 0$. Hence $p \neq 2$ for n large enough, and the theorem now follows from Theorem 3.3.

Now suppose that we are in case 2, and let $c_n = g(f^n(\gamma))$ for $n \geq 1$. Recall that we are assuming (c_n) to be a rigid divisibility sequence with $S = \{2\}$ if $\gamma \notin \mathbb{Z}$ and $S = \emptyset$ if $\gamma \in \mathbb{Z}$ (see p. 2 for definitions). Consider c_q, where q is a prime. If $v_p(c_q) > 0$ for some prime $p \neq 2$, then we have either $v_p(c_1) > 0$ or $v_p(c_n) = 0$ for $1 \leq n < q$. If $v_p(c_1) > 0$, then $v_p(c_1) = v_p(c_q)$, whence $v_p(c_q/c_1) = 0$.

Thus we need only show that for infinitely many q there exists $p \neq 2$ with $v_p(c_q/c_1)$ odd: by Theorem 3.3 this implies that $H_q(f, g)$ maximal for infinitely many q, which completes the proof.

Note that $v_2(c_q/c_1)$ cannot be odd for $q > 2$. Indeed, if $\gamma \notin \mathbb{Z}$ then since g and f are monic, $v_2(c_n) = -2^n$ for all n. If $\gamma \in \mathbb{Z}$ then either $v_2(c_1) > 0$, implying $v_2(c_n/c_1) = 0$ for all n, or $v_2(c_1) = 0$ and $v_2(c_2) > 0$, implying $v_2(c_n) = 0$ for n

odd, or $v_2(c_1) = 0$ and $v_2(c_2) = 0$, implying that $v_2(c_n) = 0$ for all n. The first two implications follow from the fact that c_n is a rigid divisibility sequence, and the last since $g(f(\gamma)) \equiv 1 \bmod 2$ and $g(f^2(\gamma)) \equiv 1 \bmod 2$ imply either $g(0) \equiv g(1) \equiv 1 \bmod 2$ or $f^n(\gamma) \equiv f^m(\gamma) \bmod 2$ for all m, n.

Let $q > 2$. To show there exists $p \neq 2$ with $v_p(c_q/c_1)$ odd, it is thus enough to show that $|c_q/c_1|$ is not a square in $\mathbb{Q}$. Suppose that infinitely many of the c_q are positive; the proof is similar in the case that infinitely many are negative. If $q \geq 2$ and $c_q/|c_1|$ is a square, then the curve $C : |c_1| y^2 = f^2(x)$ has a pair of integral points with $x = f^{q-2}(\gamma) = c_{q-2}$. Since $\{c_n : n \geq 1\}$ is infinite by assumption, each pair obtained in this way is distinct. Since f^2 is separable, it follows that C has genus 1, whence by Siegel's Theorem [11, p. 353] there can be only finitely many such pairs, and thus we have $c_q/|c_1|$ not a square for q sufficiently large. By assumption $c_q = |c_q|$ for infinitely many q, and this completes the proof. □

Remark. *For essentially all quadratic $f \in \mathbb{Z}[x]$, either $f^n(\gamma) > 0$ for all n large enough or $f^n(\gamma) < 0$ for all n large enough. In this case, the proof of Theorem 1.1, part 1, shows that for f satisfying the hypotheses of the theorem, $G_n(f)$ has finite index in the full automorphism group of the complete infinite binary rooted tree.*

Remark. *With only minor modifications, the proof of Theorem 1.1 applies to the case of $f \in \mathcal{O}_K[x]$, where $\mathcal{O}_K$ is the ring of integers in a number field K. In this case, there can be many more f that satisfy the hypotheses of the theorem. For instance, if we let α be a root of $x^3 + 3x^2 + 2x + 1$ and $K = \mathbb{Q}(\alpha)$, then not only do the families in Theorem 1.2 have $\{f^n(0)\}$ finite and not containing 0 (taking $k \in \mathcal{O}_K$ rather than $k \in \mathbb{Z}$), but so does $f(x) = x^2 + x + \alpha$. Indeed, we have $\{f^n(0)\} = \{\alpha, \alpha^2 + 2\alpha, -1\}$.*

4. Irreducibility and stability of polynomial iterates

In order to deduce Theorem 1.2 from Theorem 1.1, we must show that for f belonging to the four families in question, f is critically infinite and $g \circ f^n$ is irreducible for all n and some appropriate g. This section examines the irreducibility question, which has been studied in its own right, e.g. in [8], [3].

Definition. *Let K be a field and f and g be polynomials in $K[x]$. We say g is f-stable if $g \circ f^n$ is irreducible over K for $n = 0, 1, 2, \ldots$ (note that in particular g is irreducible over K). We say f is* stable *if f is f-stable, i.e., all iterates of f are irreducible. We say f is* eventually stable *if some iterate of f is a product of f-stable polynomials.*

Lemma 4.1 (Capelli's Lemma). *Let K be a field, $f(x), g(x) \in K[x]$, and let $\beta \in \overline{K}$ be any root of $g(x)$. Then $g(f(x))$ is irreducible over K if and only if both g is irreducible over K and $f(x) - \beta$ is irreducible over $K(\beta)$.*

For a proof, see [8, Lemma 0.1]. The following proposition is a remarkably useful tool for determining f-stability when f is quadratic.

Proposition 4.2. *Let $f(x) = ax^2 + bx + c \in K[x]$ and let $\gamma = -b/2a$ be the critical point of f. Suppose that $g \in K[x]$ is such that $g \circ f^{n-1}$ is irreducible over K for some $n \geq 2$. Then $g \circ f^n$ is irreducible over K if $g(f^n(\gamma))$ is not a square in K.*

Proof. (cf [12, Lemma 4.13]) Denote by d the degree of $g \circ f^{n-1}$. By Capelli's Lemma and the irreducibility of $g \circ f^{n-1}$, we have $g \circ f^n$ irreducible if for any root

β of $g \circ f^{n-1}$, Disc $(f(x) - \beta) = b^2 - 4ac + 4a\beta$ is not a square in $K(\beta)$. This must hold if $N_{K(\beta)/K}(b^2 - 4ac + 4a\beta)$ is not a square in K. But

$$
\begin{aligned}
N_{K(\alpha)/K}(b^2 - 4ac + 4a\beta) &= (-4a)^d \prod_{\beta \text{ root of } g \circ f^{n-1}} \left[\left(-\frac{b^2}{4a} + c \right) - \beta \right] \\
&= (-4a)^d \cdot g(f^{n-1}(-b^2/4a + c)) = (-4a)^d \cdot g(f^{n-1}(f(\gamma))).
\end{aligned}
$$

Since $n \geq 2$, we have d even, which proves the proposition. $\qquad\square$

From Proposition 4.2, it follows that stability is a generic property for quadratic polynomials with integral coefficients. As an example of the kind of result that Proposition 4.2 allows, we give the following lemma and theorem, which show that for quadratic polynomials with f^2 irreducible and given critical point γ, all but finitely many are stable.

Lemma 4.3. *Let $f \in \mathbb{Z}[x]$ be monic, quadratic, and nonsquare, and write $f = (x - \gamma)^2 + \gamma + m$. Suppose that $|m| > 6 + 3\sqrt{|\gamma| + 1}$. Then for every $n \geq 3$, $f^n(\gamma)$ is not a square in $\mathbb{Q}$. Furthermore, if $\gamma \in \mathbb{Z}$ then the same conclusion holds if $|m| > 1 + \sqrt{|\gamma| + 1}$.*

Proof. Note that either 1) $\gamma \in \mathbb{Z}$ or 2) $\gamma \notin \mathbb{Z}$ but $2\gamma \in \mathbb{Z}$. In either case we have by hypothesis $|m| > 2$, and since the denominator of m cannot exceed 4 we in fact have $|m| \geq 9/4$. For $r/s \in \mathbb{Q}$ we have

$$
f\left(\frac{r}{s}\right) = \frac{1}{s^2}\left[(r - s\gamma)^2 + s^2(\gamma + m) \right],
$$

and if we assume that $s\gamma$ and $s^2(\gamma + m)$ are integers then $f(r/s)$ is not a square if

$$
\tag{13} \left| s^2(\gamma + m) \right| < 2\left| r - s\gamma \right| - 1.
$$

This follows because $\gamma + m \neq 0$ (since f is nonsquare) and

$$
\min \left\{ \left| x^2 - (x-1)^2 \right|, \left| x^2 - (x+1)^2 \right| \right\} \geq 2|x| - 1
$$

for $x \in \mathbb{Z}$.

Put $\frac{r}{s} = f^{n-1}(\gamma)$, and note that $s = j^{2^{n-1}}$, where j is 1 if $\gamma \in \mathbb{Z}$ and 2 otherwise. Dividing (13) by s^2 then yields that $f^n(\gamma)$ is not a square if

$$
\tag{14} |\gamma + m| < j^{-2^{n-1}}\left(2\left| f^{n-1}(\gamma) - \gamma \right| - 1 \right).
$$

A quick induction shows that $f^{n-1}(\gamma) - \gamma = f_m^{n-1}(0)$, where $f_m(x) = x^2 + m$.

We now show that (14) holds for $n \geq 3$. First we show that it holds for $n = 3$; in this case (14) becomes

$$
\tag{15} j^4|\gamma + m| < 2\left| m^2 + m \right| - 1.
$$

Suppose $j = 2$, or equivalently that $\gamma \notin \mathbb{Z}$. We have $|m| - 6 > 3\sqrt{|\gamma| + 1}$, whence by squaring both sides and subtracting we have $|m|^2 - |m| > 11|m| + 9|\gamma| + 9 - 36$. Since $\gamma \notin \mathbb{Z}$, we have $\gamma \geq 1/2$. Thus $|m| \geq 9 + 3/4$ (since the denominator of m is 4), and we have $11|m| + 9|\gamma| - 27 > 8|m| + 8|\gamma| + 1$. Now

$$
2\left| m^2 + m \right| - 1 \geq 2(|m|^2 - |m|) - 1 > 16|m| + 16|\gamma| + 1 > 16|\gamma + m|,
$$

and thus (15) is satisfied. In the case $j = 1$, we have $(|m| - 1)^2 > |\gamma| + 1$, and thus $|m|^2 - |m| > |\gamma| + |m|$. Therefore

$$
2\left| m^2 + m \right| - 1 \geq 2(|m|^2 - |m|) - 1 > 2(|\gamma| + |m|) - 1 > |\gamma| + |m|,
$$

where the last inequality follows since $|m| \geq 9/4$. Hence (15) is satisfied.

To show that (14) holds for $n > 3$, we must show the function $j^{-2^n} \left(2 \left|f_m^n(0)\right| - 1\right)$ increases with n, i.e.

$$(16) \qquad j^{2^{n-1}} < \frac{2 \left|f_m^n(0)\right| - 1}{2 \left|f_m^{n-1}(0)\right| - 1}$$

Noting that $f_m^n(0) = (f_m^{n-1}(0))^2 + m$, it is enough to prove

$$j^{2^{n-1}} < \left|f_m^{n-1}(0)\right| - \frac{2|m| + 1}{2 \left|f_m^{n-1}(0)\right|}.$$

Since $|m| \geq 9/4$, it follows that $2|m| + 1 < 2 \left|f_m^{n-1}(0)\right|$ for all $n \geq 3$. Thus it suffices to show

$$(17) \qquad \left|f_m^{n-1}(0)\right| > 1 + j^{2^{n-1}}$$

for all $n \geq 4$. One can easily use $|m| > 9$ to show that this holds for $n = 4$ (in the case $j = 1$, $|m| > 2$ suffices). Now assume that (17) holds for $n - 1$, with $n \geq 5$. We show that it holds for n.

If $j = 1$, then $|m| > 2$ guarantees that $|f_m^n(0)| > \left|f_m^{n-1}(0)\right|$, and we are done. If $j = 2$, suppose first that $|m| > 2j^{2^{n-1}}$. Since $|m| > 3$ (indeed we know $|m| > 9$), we have $|m|^2 - 3|m| > 0$, and this implies that $\left|m^2\right| - |m| > 2|m| > |m| + 1 + j^{2^{n-1}}$. Therefore $\left|f_m^2(0)\right| > |m| + 1 + j^{2^{n-1}}$. Squaring both sides leads to $\left|f_m^3(0)\right| > 1 + j^{2^n}$. Since $|f_m^n(0)| > \left|f_m^{n-1}(0)\right|$, we have $|f_m^n(0)| > 1 + j^{2^n}$.

Now suppose that $|m| < 2j^{2^{n-1}}$ (equality is impossible since $j = 2$ implies $\gamma \notin \mathbb{Z}$). Given $\left|f_m^{n-1}(0)\right| > 1 + j^{2^{n-1}}$, squaring both sides gives

$$|f_m^n(0) - m| > 1 + 2j^{2^{n-1}} + j^{2^n} > 1 + j^{2^n} + |m|.$$

Therefore $|f_m^n(0)| + |m| > 1 + j^{2^n} + |m|$, and we are done. $\qquad \square$

Theorem 4.4. *Let $f \in \mathbb{Z}[x]$ be monic, quadratic, and irreducible, and write $f = (x - \gamma)^2 + \gamma + m$. Suppose that $|m| > 6 + 3\sqrt{|\gamma| + 1}$ (if $\gamma \in \mathbb{Z}$ then $|m| > 1 + \sqrt{|\gamma| + 1}$ suffices), and that*

$$(18) \qquad \frac{-m \pm \sqrt{f^2(\gamma)}}{2} \notin \mathbb{Q}^{*2}.$$

Then f is stable.

Remark. *Condition (18) is always true if $\gamma \notin \mathbb{Z}$. Indeed, let $f = x^2 + ax + b$, with a odd. We have $\gamma = -a/2$, $m = (2a - a^2 + 4b)/4$, and*

$$f^2(\gamma) = \frac{a^4 - 4a^3 - 8a^2 b + 16ab + 16b^2 + 16b}{16}.$$

Reducing the numerator of $f^2(\gamma)$ modulo 8 gives $a^4 - 4a^3$. Since a is odd, this is equivalent to 5 modulo 8, showing that $f^2(\gamma)$ cannot be a square in $\mathbb{Q}$.

Proof. Given that f is irreducible, condition (18) is equivalent to f^2 being irreducible, as a quick calculation and application of Proposition 4.2 shows. Since f is irreducible, it is not a square, whence Lemma 4.3 shows that $f^n(\gamma)$ is not a square for all $n \geq 3$. By induction and Proposition 4.2 this proves that f is stable. $\qquad \square$

Proposition 4.5. *Let $k \in \mathbb{Z}\backslash\{0, -1\}$, and put $f = x^2 + k$. Then f is stable if $-k$ is not a square in $\mathbb{Z}$ and f factors as the product of two linear f-stable polynomials if $-k$ is a square in $\mathbb{Z}$.*

Proof. Suppose first that $-k$ is not a square in $\mathbb{Z}$, whence f is irreducible. The critical point of f is $\gamma = 0$, and $f^2(0) = k^2 + k$, which is not a square since $k \notin \{0, -1\}$. Thus the condition (18) is satisfied, as indeed are all the hypotheses of Theorem 4.4 as long as $k \neq 1$, $k \neq 2$, and $k \neq -2$. For these cases, we verify (14) directly for all $n \geq 2$. If $k = 1$, then (14) becomes $1 < 2|f^n(0)| - 1$, which clearly holds for $n \geq 2$. A similar argument applies to $k = 2$. If $k = -2$ the forward orbit of 0 lies in $\{\pm 2\}$, whence $f^n(0)$ is not a square for $n \geq 2$. In all three cases induction and Proposition 4.2 show f is stable.

If $-k = a^2$ for some $a \in \mathbb{Z}$ then we have $f = g_1 g_2$, where $g_1 = x + a$ and $g_2 = x - a$. Note that since $k \notin \{0, -1\}$, we have $|a| \geq 2$. To show that $g_i(f^n(0))$ is not a square for all $n \geq 2$, we note $g_i(f^n(0)) = f^{n-1}(0)^2 - a^2 \pm a$, so by the same reasoning used to obtain (14) we have $g_i(f^n(0))$ not a square if $a^2 \pm a < 2|f^{n-1}(0)| - 1$. Since $|a| \geq 2$ this holds for $n = 2$. Moreover, again because $|a| \geq 2$, it follows that $|f^{n+1}(0)| > |f^n(0)|$ for all $n \geq 2$. Induction and Proposition 4.2 now show g_1 and g_2 are f-stable. $\square$

Proposition 4.6. *Let $k \in \mathbb{Z}\backslash\{0\}$, and put $f = x^2 + kx - k$. Then f is stable if $k \neq 4$ and f is the square of an f-stable polynomial if $k = 4$.*

Proof. In the notation of Theorem 4.4, we have $\gamma = -k/2$ and $m = k(2 - k)/4 = \gamma - \gamma^2$. Thus f is irreducible as long as $2\gamma - \gamma^2 \neq 0$, which is the case when $k \neq 0, 4$. To show that f^2 is irreducible, we note that $f^2(\gamma) = (\gamma - \gamma^2)^2 + 2\gamma - \gamma^2$, so to ensure $f^2(\gamma)$ is not a square, it is enough to show $|2\gamma - \gamma^2| < 2|\gamma - \gamma^2| - 1$. This holds for $|\gamma| > 2$. Now to use Theorem 4.4 we must find the γ with

$$(19) \qquad\qquad |\gamma - \gamma^2| > 6 + 3\sqrt{|\gamma| + 1}.$$

A derivative shows that $|\gamma - \gamma^2| - 6 - 3\sqrt{|\gamma| + 1}$ increases with $|\gamma|$ provided $|\gamma| > 2$. One easily sees that (19) holds for all $\gamma \notin (-7/2, 9/2)$. Similarly, a derivative shows that $|\gamma - \gamma^2| - 1 - \sqrt{|\gamma| + 1}$ increases with $|\gamma|$ provided $|\gamma| > 1$. A quick calculation now shows $|\gamma - \gamma^2| > 1 + \sqrt{|\gamma| + 1}$ for all $\gamma \notin (-2, 3)$. Therefore by Theorem 4.4 we have that f is stable unless $k = -5, -3, -2, -1, 1, 2, 3, 5, 7$. For all of these save $k = \pm 1$ one can easily check that all iterates of f are Eisenstein and thus irreducible. The case $k = -1$ is proven in [18, Proposition 1.1]. The case $k = 1$ is handled in Proposition 4.7.

For $k = 4$, we clearly have $f = g^2$, with $g = (x - 2)$. Moreover, $f^n(\gamma) = 0$ for $n = 1$ and $f^n(\gamma) = 4$ for $n \geq 2$. Thus $g \circ f^n(\gamma)$ is not a square for all $n \geq 1$, and it follows that $g \circ f^n(\gamma)$ is irreducible for all $n \geq 1$. $\square$

Proposition 4.7. *Let $k \in \mathbb{Z}\backslash\{0\}$, and put $f = x^2 + kx - 1$. Then f is stable if $k \neq -1$ and f^3 is the product of two f-stable polynomials if $k = -1$.*

Proof. In the notation of Theorem 4.4, we have $\gamma = -k/2$ and $m = (-k^2 + 2k - 4)/4$. We have that f is irreducible as long as $k^2 + 4$ is not a square, which holds for all $k \neq 0$. To show $f^2(x)$ is irreducible, note that

$$f^2(\gamma) = \frac{1}{16}(k^4 - 4k^3 + 8k^2 + 16k) = \left(\frac{1}{4}(k - 1)^2\right)^2 + \frac{1}{16}(2k^2 - 12k - 1).$$

Thus $f^2(\gamma)$ is not a square as long as $\left|2k^2 - 12k - 1\right| < 8\left|k-1\right|^2 - 1$, which holds for $k \notin \{0,1,2,3\}$.

One checks that $\left|(-k^2 + 2k - 4)/4\right| > 6 + 3\sqrt{|k/2|+1}$ as long as $k \notin (-7,9)$ and also $\left|(-k^2 + 2k - 4)/4\right| > 1 + \sqrt{|k/2|+1}$ as long as $k \notin (-2,4)$. Using Theorem 4.4, we now have that f is stable as long as $k \notin \{-5,-3,-1,0,1,2,3,5,7\}$. Consider $k = 7$. We have $\gamma = -7/2$, $m = -39/4$. We note that in the proof of Lemma 4.3 we need only verify (14) for some n and (17) for all $i \geq n$. When $n = 4$ we have $f^{n-1}(\gamma) - \gamma > 7268$ and $2^{2^{n-1}} \cdot |\gamma + m| = 3392$, whence (14) holds. To verify (17) for $n \geq 4$, note that it holds for $n = 4$ since $f_m^3(0) > 7268 > 1 + 2^8$. Moreover, given that (17) holds for n, to show it holds for $n+1$ it is enough to have $|m| > 3$ (see proof of Lemma 4.3). Hence we have shown that $f^n(\gamma)$ not a square for $n \geq 4$. One checks the remaining cases by hand, proving that f is stable. In a similar manner, one verifies that f is stable for $k \in \{-5,-3,5,7\}$, although for $k = -3$ one must take $n = 6$ and for $k = 5$ one must take $n = 7$.

For $k \in \{1,2,3\}$, let $\alpha_1 < \alpha_2$ be the two (real) roots of f, and put $I = (\alpha_1, 0)$. Using the fact that $f(\gamma) \in I$, it follows that $f(I) \subseteq I$, whence $f^n(\gamma) < 0$ for all n. Thus $f^n(\gamma)$ cannot be a square, so f is stable. This leaves the case $k = -1$, where we see that although f and f^2 are irreducible,

$$f^3(x) = (x^4 - 3x^3 + 4x - 1)(x^4 - x^3 - 3x^2 + x + 1) := g_1(x)g_2(x),$$

where g_1 and g_2 are irreducible. To prove that g_1 is f-stable, it is enough to show that $g_1(f^n(\gamma))$ is not a square for all $n \geq 1$. Note that $g_1(-1) = -1$ and also $f^2(-1) = -1$. Working modulo 3, we have $\gamma \equiv -1$, whence $g_1(f^n(\gamma)) \equiv -1 \mod 3$ for n even. Working modulo 19, we have $f^5(\gamma) \equiv -1 \mod 19$, whence $g_1(f^n(\gamma)) \equiv -1 \mod 19$ for n odd, $n \geq 5$. One now checks that $g_1(f^n(\gamma))$ is not a square in $\mathbb{Q}$ for $n = 1,3$, showing that g_1 is f-stable. For g_2 we have $g_2(1) = g_2(-1) = -1$, and since $f^n(\gamma) \equiv \pm 1 \mod 3$ for all n, we have $g_2(f^n(\gamma)) \equiv -1 \mod 3$ for all n, proving that g_2 is f-stable. $\qquad \square$

We note that it is possible for a quadratic f to be irreducible over $\mathbb{Q}$ yet for f^2 to have a nontrivial factorization. One example is $x^2 + 10x + 17$, which is irreducible yet clearly fails to satisfy condition (18); see [3, p. 91] for an explicit description of all such polynomials. Moreover, as in the case of $f(x) = x^2 - x - 1$, it is also possible for f and f^2 to be irreducible yet for f^3 to have a nontrivial factorization. Computer searches have revealed no monic, quadratic $f \in \mathbb{Z}[x]$ whose first three iterates are irreducible yet whose fourth is not. Moreover, in all cases checked, the reducible iterate has factored into two f-stable polynomials. Generalizing this observation, we make the following conjecture:

Conjecture 1. *Let $f \in \mathbb{Z}[x]$ be monic and quadratic, and suppose that 0 is not periodic under f. Then f is eventually stable.*

We remark that if 0 is periodic under $f \in \mathbb{Z}[x]$, then f cannot be eventually stable. This is because the factors of f^n are the minimal polynomials over $\mathbb{Q}$ of the preimages of 0 (in $\overline{\mathbb{Q}}$) under f. Thus if 0 is periodic under f, x occurs as a factor of some f^m, and it follows that f^n has a linear factor for each $n \geq 1$. Thus for instance no member of the family $x^2 + kx - (k+1), k \in \mathbb{Z}$, is eventually stable.

5. Results for specific families

In this section we draw together the results from all the previous sections to prove Theorem 1.2, treating one family at a time. We also give Conjecture 2, which treats more general f than those meeting the hypotheses of Theorem 1.1.

Theorem 5.1. *Let $f = x^2 - kx + k$ for $k \in \mathbb{Z}$. Then $D(P(f, a_0)) = 0$ for all a_0.*

Proof. Clearly $P(f, a_0)$ is finite for $k = 0$, so this case is covered. To apply Theorem 1.1 we must show that f is stable and critically infinite. If $k \neq 4$, then by Proposition 4.6 we have that f is stable. We now show that $\{f^n(\gamma) : n \geq 1\}$ is infinite for $k \notin \{-2, 2, 4\}$. Recall that we denote the critical point of f by γ, which is $k/2$ for the family under consideration. Suppose first that k is odd. Letting a_{n-1} and a_n denote the numerators of $f^{n-1}(\gamma)$ and $f^n(\gamma)$ respectively, we have $a_n = a_{n-1}^2 - k2^{2^{n-1}} a_{n-1} + k2^{2^n}$. Moreover, $f^n(\gamma) = a_n/2^{2^n}$. Thus by induction a_n must be odd, and therefore f is critically infinite since the denominator of $f^n(\gamma)$ in lowest terms is increasing.

Now suppose that k is even. We note that f is increasing on (γ, ∞), and that the largest fixed point of f is $\rho = \max\{1, k\}$. If $f^n(\gamma) > \rho$ for some n, one easily sees that $f^{m+1}(\gamma) > f^m(\gamma)$ for all $m \geq n$, showing that f is critically infinite. We show $f^n(\gamma) > \rho$ for $n = 2$. Note that $f^2(\gamma) = \frac{1}{16}(k^4 - 4k^3) + k$. If $k > 0$ then $\rho = k$, and we have $f^2(\gamma) > \rho$ as long as $k > 4$. If $k < 0$ then $\rho = 1$, and we have $f^2(\gamma) > \rho$ as long as $k^4 - 4k^3 + 16k - 16 > 0$. This is equivalent to $(k - 2)^3(k + 2) > 0$, which holds when $k < -2$. Thus f is critically infinite as long as $k \notin \{-2, 2, 4\}$, and the theorem is proved in all but these cases.

For $k \in \{-2, 2, 4\}$, we first show that $\#H_n(f, g) = 2$ for all n, where $g = x - 2$ for $k = 4$ and $g = x$ in the other two cases. When $k = 4$, the roots of $g \circ f^n$ are the same as the roots of iterates of $x^2 - 2$, that is, $\zeta_{2^{n+2}} + \zeta_{2^{n+2}}^{-1}$, and the extensions thus generated are $\mathbb{Q}(\zeta_{2^{n+2}}) \cap \mathbb{R}$. For $k = -2$, i.e. $f = x^2 + 2x - 2$, iterates of f form extensions that are $\mathbb{Q}(\zeta_{3 \cdot 2^{n+1}}) \cap \mathbb{R}$. Finally, for $k = 2$, roots of iterates of f are roots of $x^{2^n} + 1$ for all $n \geq 1$. In all three cases one easily shows that $\#H_n(f, g) = 2$ for all n. We now show by induction that the identity is the only element of $G_n(f, g)$ that fixes a root of $g \circ f^n$, whence the proportion of such elements is 2^{-n} and from Theorem 2.1 we have $D(P(f, a_0)) = 0$. The case $n = 1$ is clear. Assuming that only the identity in $G_{n-1}(f, g)$ fixes any roots of $g \circ f^{n-1}$, we have that only elements of $H_n(f, g)$ can fix any roots of $g \circ f^n$. Since $H_n(f, g)$ is not trivial, $g \circ f^n$ factors over $\mathbb{Q}(g \circ f^{n-1})$ as a product of quadratics, and $H_n(f, g)$ must act transitively on the roots of each of these factors. Thus $H_n(f, g)$ consists of the identity and an element that exchanges the roots of each quadratic factor of $g \circ f^n$ over $\mathbb{Q}(g \circ f^{n-1})$, and thus fixes no roots of $g \circ f^n$. $\square$

Theorem 5.2. *Let $f = x^2 + kx - 1$ for $k \in \mathbb{Z} \backslash \{0, 2\}$. Then $D(P(f, a_0)) = 0$ for all a_0.*

Proof. Part (1) of Theorem 1.1 applies immediately (with $g(x) = x$) in the case that f is stable and critically infinite. If $k \neq -1$, then by Proposition 4.7 we have that f is stable. We now show that $\{f^n(\gamma) : n \geq 1\}$ is infinite for $k \notin \{0, 2\}$. As in the proof of Theorem 5.1, the case k odd is covered.

Now suppose that k is even, and note that $\gamma = -k/2$, $m = (-k^2 + 2k - 4)/4$. Following the proof of Theorem 5.1, we have here that the largest fixed point ρ of f is at most $\max\{1, -k + 2\}$. Again it is enough to show $f^n(\gamma) > \rho$ for some n,

and we take $n = 2$. We have $f^2(\gamma) = \frac{1}{16}(k^4 - 4k^3 + 8k^2 - 16k)$, which is greater than 1 for $k \geq 4$ and greater than $-k + 2$ for $k \leq -2$.

This leaves the case $k = -1$. Since from Proposition 4.7 we have $f^3 = g_1 g_2$ with g_1 and g_2 f-stable, Part (1) of Theorem 1.1 applies to show that both $P(g_1, f, a_0)$ and $P(g_2, f, a_0)$ have density 0 for every $a_0 \in \mathbb{Z}$. Since $P(f, a_0)$ is the union of these two sets, it too has density 0. The theorem is proved. $\square$

The cases of $x^2 - 1$ and $x^2 + 2x - 1 = (x+1)^2 - 2$ present peculiar difficulties, as both are critically finite, but the relevant Galois groups do not lend themselves to easy analysis. We remark that to show $D(P(x^2 - 1, a_0)) = 0$, put $f = x^2 - 1$ and note that $f^n(a_0) \equiv 0 \bmod p$ for some n implies either 1) $a_0 \equiv 0$ or $-1 \bmod p$ or 2) $f^m(a_0) \equiv 1 \bmod p$ for some m. Thus it suffices to show that the density of primes p with $f^n(a_0) \equiv 1 \bmod p$ is zero, and this may be accomplished by considering the Galois groups of $f^n(x) - 1$. These are the same as those generated by iterates of $h(x) = (x + 1)^2 - 2$. Iterates of h generate extensions that are finitely ramified (indeed, unramified outside $\{2\}$), yet are much larger than the extensions generated in the exceptional cases handled in Theorem 5.1. There appears to be no easy way to show that $\lim_{n \to \infty} \mathcal{P}(X_n) = 0$ in this case.

We now turn to the case where f is monic and quadratic and all primes dividing an element of the critical orbit are recurrent. This is the case, for instance, whenever 0 is the critical point, as the following lemma shows. See [19, Section 3] for a nice discussion of matters similar to the following lemma and proposition.

Lemma 5.3. *Let $f = x^2 + k$ for $k \in \mathbb{Z}\backslash\{0, -1\}$, and put $f^n(0) = t_n$ for $n \geq 1$. Then t_n is a rigid divisibility sequence with $S = \emptyset$.*

Proof. (cf [12, Lemma 6.1]) Since we have excluded $k \in \{0, -1\}$, we have $t_n \neq 0$ for all n. We use induction on j to show $v_p(t_n) = e > 0$ implies $v_p(t_{nj}) = e$ for all $j \geq 1$. The case $j = 1$ is trivial. Suppose inductively that $v_p(t_{n(j-1)}) = e$. Note that $t_{nj} = f^{n(j-1)}(t_n)$, and also $f^{n(j-1)}$ is a polynomial in x^2. Thus we can write

$$f^{n(j-1)}(x) = x^2 g(x) + f^{n(j-1)}(0) = x^2 g(x) + t_{n(j-1)},$$

for some $g \in \mathbb{Z}[x]$. Hence putting $x = t_n$ we have $t_{nj} = t_n^2 g(t_n) + t_{n(j-1)}$. Now $v_p\left[(t_n)^2(g(t_n))\right] \geq 2e$, and by our inductive hypothesis $v_p(t_{n(j-1)}) = e$. Since $e \geq 1$, the first summand vanishes to higher order at p than the second, so we conclude $v_p(p_{nj}) = e$.

We must also show that $p^e \mid t_n$ and $p^e \mid t_m$ implies $p^e \mid t_{\gcd(m,n)}$. Let $p^e \mid t_n$ and $p^e \mid t_m$, and suppose without loss of generality that $n > m$. Then $f^n(0) \equiv f^m(0) \equiv 0 \bmod p^e$, whence $0 \equiv f^{n-m}(f^m(0)) \equiv f^{n-m}(0) \bmod p^e$. Applying this repeatedly and using the Euclidean algorithm yields $p^e \mid t_{\gcd(m,n)}$. $\square$

If $f^n(0)$ is a rigid divisibility sequence and f is reducible, its factors give rise to other rigid divisibility sequences, as the following Proposition shows.

Proposition 5.4. *Suppose that $f \in \mathbb{Z}[x]$ has all iterates separable, and that 0 is a critical point of f and $(f^n(0))$ is a rigid divisibility sequence $(S = \emptyset)$ with infinitely many values. If g divides f then the sequence $g(0), g(f(0)), g(f^2(0)), \ldots$ is a rigid divisibility sequence $(S = \emptyset)$.*

Proof. Suppose that $gh = f$, and for $n \geq 1$ put $a_n = f^n(0)$ and b_n, c_n respectively for $g(f^{n-1}(0)), h(f^{n-1}(0))$. Recall that by convention $f^0(x) = x$, and note that we

have $b_n c_n = a_n$ for all $n \geq 1$. If $v_p(b_i) = e > 0$, then we have $a_i = f^i(0) \equiv 0 \bmod p^e$, whence 0 is periodic modp^e under f, with period dividing i. Hence in particular $f^{mi-1}(0) \equiv f^{i-1}(0) \bmod p^e$ for all $m \geq 1$. Now $v_p(b_i) = e > 0$ implies that $g(f^{i-1}(0)) \equiv 0 \bmod p^e$, and thus $f^{i-1}(0)$ is a root of $g \bmod p^e$. Therefore $f^{mi-1}(0)$ is also a root of $g \bmod p^e$, implying that $v_p(b_{mi}) \geq e$. Similarly, if $v_p(c_i) = e > 0$, then $v_p(c_{mi}) \geq e$ for all $m \geq 1$.

Now suppose that $v_p(b_i) = e_1 > 0$. Then $v_p(c_i) = e_2 \geq 0$, and $v_p(a_i) = e_1 + e_2$. Since (a_n) is a rigid divisibility sequence, we have $v_p(a_{mi}) = e_1 + e_2$ for all $m \geq 1$. But by the previous paragraph $v_p(b_{mi}) \geq e_1$ and $v_p(c_{mi}) \geq e_2$. Thus $a_{mi} = b_{mi}c_{mi}$ implies $v_p(b_{mi}) = e_1$ and $v_p(c_{mi}) = e_2$.

To show that $p^e \mid b_n$ and $p^e \mid b_m$ implies $p^e \mid b_{\gcd(m,n)}$, one may use the same argument as in the proof of Lemma 5.3. Hence (b_n) is a rigid divisibility sequence, as desired. $\qquad\square$

Theorem 5.5. *Let $f = x^2 + k \in \mathbb{Z}[x]$ with $k \neq -1$. Then $D(P(f, a_0)) = 0$ for all $a_0 \in \mathbb{Z}$.*

Proof. First suppose that $-k$ is not a square in $\mathbb{Z}$. By Proposition 4.5 f is stable, and provided that $k \neq -2$ we have f critically infinite, whence the theorem follows from Theorem 1.1, part 2, taking $g(x) = x$. The case $k = -2$ may be resolved directly, as has already been done in the proof of Theorem 5.1.

Suppose now that $-k = c^2$ for some $c \in \mathbb{Z}$ with $|c| \geq 2$, thereby excluding $k \in \{0, -1\}$. Then $\{f^n(0)\}$ is infinite. Letting $g(x) = x - c$, we have that $g \circ f^n$ is irreducible for all $n \geq 1$ by Proposition 4.5. Moreover, the critical point of f is $\gamma = 0$ and Lemma 5.3 and Proposition 5.4 imply that $g(f^n(0))$ is a rigid divisibility sequence. Thus part 2 of Theorem 5.2 shows that $D(P(g, f, a_0)) = 0$ for all $a_0 \in \mathbb{Z}$. Similarly, $D(P(h, f, a_0)) = 0$, where $h(x) = x + c$. Since $g(f^n(a_0)) \cdot h(f^n(a_0)) = f^{n+1}(a_0)$, we have that $P(f, a_0)$ is the union of two density zero sets, and thus has density zero.

Finally, if $k = 0$ it is clear that the set of divisors of a_n is finite. $\qquad\square$

For our final family, we ride the coattails of Theorem 5.5. This is very much along the lines of the sequence factorization described in [19, Section 3].

Theorem 5.6. *Let $f = x^2 - 2kx + k \in \mathbb{Z}[x]$ with $k \neq \pm 1$. Then $D(P(f, a_0)) = 0$ for all $a_0 \in \mathbb{Z}$.*

Proof. Note first that $f(x) = (x - k)^2 - k^2 + k$, so that $f = \phi \circ h \circ \phi^{-1}$, where $\phi(x) = x + k$ and $h(x) = x^2 - k^2$. Hence $f^n = \phi \circ h^n \circ \phi^{-1}$, implying that $f^n(a_0) = h^n(a_0 - k) + k$. But $(h^n(a_0 - k) + k)(h^n(a_0 - k) - k) = h^{n+1}(a_0 - k)$, and $D(P(h, a_0 - k)) = 0$ by Theorem 5.5 provided $k^2 \neq 1$. $\qquad\square$

We have now completed the proof of Theorem 1.2, in which Theorem 1.1 has been the driving force. Unfortunately, quadratic $f \in \mathbb{Z}[x]$ meeting the hypotheses of Theorem 1.1 are rare (and indeed remain rare despite growing more numerous if $\mathbb{Z}$ is replaced by the ring of integers of a number field). However, we conjecture that the conclusion of Theorem 1.1 holds for generic quadratic $f \in \mathbb{Z}[x]$:

Conjecture 2. *Let $f \in \mathbb{Z}[x]$ be quadratic, and suppose that f is stable and critically infinite. Then $D(P(f, a_0)) = 0$ for all $a_0 \in \mathbb{Z}$.*

The obstacle impeding a proof of Conjecture 2 is a knowledge of the arithmetic of the critical orbit of f. It is enough to show that given a stable, quadratic $f \in \mathbb{Z}[x]$

and $b \in \frac{1}{2}\mathbb{Z}$ with $\{f^n(b) : n \geq 1\}$ infinite, then there are infinitely many n such that there exists a prime p with $v_p(f^n(b))$ odd and $v_p(f^m(b)) = 0$ for all $m < n$. This assertion is plausible in light of work of [6] showing that for any nonzero $c \in \mathbb{Z}$, "most" n have the property that there exists p with $v_p(n^2 + c) = 1$ but $v_p(m^2 + c) = 0$ for all $m < n$. It seems plausible as well that a similar assertion is true for translated sequences $\{g(f^n(b)) : n \geq 1\}$. This would prove that Conjecture 2 holds for all quadratic f that are eventually stable and critically infinite. If Conjecture 1 is true as well, this would show that $D(P(f, a_0)) = 0$ for virtually all quadratic f.

6. The infinitude of prime divisors of first-order polynomial recurrences

Theorem 6.1. *Suppose that $f \in \mathbb{Z}[x]$ is of degree $d \geq 2$ and not of the form ax^d, and let $a_0 \in \mathbb{Z}$ be such that $\{f^n(a_0) : n \in \mathbb{N}\}$ is infinite. Then there is a constant c with $\#(P(f, a_0) \cap \{1, \ldots, n\}) \geq c \log \log n$ for n large enough. In particular, $P(f, a_0)$ is infinite.*

Proof. We adopt an approach suggested independently to the author by I. Shparlinski and V. Dimitrov. First note that by [20, Theorem 3.20] the limit

$$\lim_{n \to \infty} \frac{1}{d^n} \log |f^n(a_0)|$$

exists, and by [20, Theorem 3.22] is 0 if and only if $\{f^n(a_0) : n \in \mathbb{N}\}$ is finite. Since here $\{f^n(a_0) : n \in \mathbb{N}\}$ is infinite, it follows that there are constants $C, C' > 1$ with

$$C^{d^n} < |f^n(a_0)| < (C')^{d^n}$$

for n large enough.

Let $h(k) = \#\{p : p \mid a_i \text{ for some } i \leq k\}$. We show that $h(k) > k/3$ for k sufficiently large. Consider the $h(k) + 1$ numbers $a_{k-h(k)}, \ldots, a_k$, and suppose that there are infinitely many k with $h(k) < k$ (otherwise we are done). For a natural number m, call p a *champion* prime for m if $v_p(m) \geq v_q(m)$ for all $p \neq q$. Note that $p^{v_p(m)} \geq |m|^{1/s}$ for any champion prime of m, where s is the number of distinct prime factors of m. By the pigeonhole principle, there are i and j with $k - h(k) \leq i < j \leq k$ such that a_i and a_j share a champion prime. It follows that

$$(20) \qquad \gcd(a_i, a_j) \geq |a_i|^{1/h(k)} \geq \left|a_{k-h(k)}\right|^{1/h(k)} \geq C_0^{d^{k-h(k)}/h(k)}.$$

The inequality $|a_i| \geq \left|a_{k-h(k)}\right|$ follows by taking k sufficiently large and noting $|f^n(a_0)| \to \infty$ since $\{f^n(a_0) : n \in \mathbb{N}\}$ is infinite.

On the other hand, taking i and j as in the previous paragraph and writing $j = i + r$ with $r \leq h(k)$, we have $a_j = f^r(a_i) \equiv f^r(0) \bmod a_i$. It follows that $\gcd(a_i, a_j) \mid f^r(0)$. Suppose now that $f^r(0) \neq 0$ for all $r \geq 1$ (we deal with the other case in a moment). Then we now have $\gcd(a_i, a_j) \leq C_1^{d^{h(k)}}$. Combining this with (20), taking log twice and rearranging yields $h(k) + C_3 \log h(k) + C_4 > k/2$. For k sufficiently large this gives $h(k) > k/3$.

Now each of the primes contributing to $h(k)$ can be at most $|a_k| \leq C_4^{d^k}$. Hence $\#\{p \in P(f, a_0) : p \leq C_4^{d^k}\} \geq k/3$ for k sufficiently large. It now follows that $\#(P(f, a_0) \cap \{1, \ldots, n\}) \geq C_5 \log \log n$ for some $C_5 > 0$.

Finally, suppose that $f^r(0) = 0$ for some $r \geq 1$. Write $f^r(x) = x^l g(x)$ with $g(0) \neq 0$. Since f is not of the form ax^d we may take $\deg(g) \geq 1$. Consider

the sequence $b_n = a_{rn}$, so that $b_{n+1} = f^r(b_n)$. Note that $b_{n+1} = b_n^l \cdot g(b_n)$. If $v_p(b_n) = e \geq 1$ then $g(b_n) \equiv g(0) \bmod p^e$. It follows that if $p \mid g(0)$ then $v_p(b_n)$ goes to infinity with n, whence for n large enough we have $v_p(b_n) > v_p(g(0))$. Clearly the same conclusion holds if $p \nmid g(0)$. Therefore in either situation we have that for n large enough $v_p(b_n) = e \geq 1$ implies $v_p(g(b_n)) < e$. Thus if all primes dividing $g(b_n)$ also divide b_n, then $g(b_n) \leq b_n/2$. However, by hypothesis b_n takes on infinitely many values, and since $\deg(g) \geq 1$, we have $g(b_n) \geq b_n - C$ for some constant C. Hence for n large enough there must exist p with $p \mid g(b_n)$ but $p \nmid b_n$, whence $p \mid b_{n+1}$ but $p \nmid b_n$. Putting $h(k) = \#\{p : p \mid a_i \text{ for some } i \leq k\}$ as above, we have now shown that $h(k) \geq k/(r+\epsilon)$, and $\#(P(f, a_0) \cap \{1, \ldots, n\}) \geq c \log \log n$ follows as above. $\qquad \square$

References

[1] Wayne Aitken, Farshid Hajir, and Christian Maire. Finitely ramified iterated extensions. *Int. Math. Res. Not.*, (14):855–880, 2005.

[2] Artur Avila and Carlos Gustavo Moreira. Statistical properties of unimodal maps: the quadratic family. *Ann. of Math. (2)*, 161(2):831–881, 2005.

[3] Mohamed Ayad and Donald L. McQuillan. Irreducibility of the iterates of a quadratic polynomial over a field. *Acta Arith.*, 93(1):87–97, 2000.

[4] Christian Ballot. Density of prime divisors of linear recurrences. *Mem. Amer. Math. Soc.*, 115(551):viii+102, 1995.

[5] Robert L. Devaney. *An introduction to chaotic dynamical systems*. Addison-Wesley Studies in Nonlinearity. Addison-Wesley Publishing Company Advanced Book Program, Redwood City, CA, second edition, 1989.

[6] Graham Everest, Shaun Stevens, Duncan Tamsett, and Tom Ward. Primes generated by recurrence sequences. *Amer. Math. Monthly*, 114(5):417–431, 2007.

[7] Jan-Hendrik Evertse. On sums of S-units and linear recurrences. *Compositio Math.*, 53(2):225–244, 1984.

[8] Burton Fein and Murray Schacher. Properties of iterates and composites of polynomials. *J. London Math. Soc. (2)*, 54(3):489–497, 1996.

[9] Geoffrey R. Grimmett and David R. Stirzaker. *Probability and random processes*. Oxford University Press, New York, third edition, 2001.

[10] Helmut Hasse. Über die Dichte der Primzahlen p, für die eine vorgegebene ganzrationale Zahl $a \neq 0$ von gerader bzw.ungerader Ordnung mod. p ist. *Math. Ann.*, 166:19–23, 1966.

[11] Marc Hindry and Joseph H. Silverman. *Diophantine geometry*, volume 201 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 2000. An introduction.

[12] Rafe Jones. Iterated Galois towers, their associated martingales, and the p-adic Mandelbrot set. *Compositio Math.*, 43(5):1108–1126, 2007.

[13] J. C. Lagarias. The set of primes dividing the Lucas numbers has density 2/3. *Pacific J. Math.*, 118(2):449–461, 1985.

[14] Serge Lang. *Algebra*, volume 211 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, third edition, 2002.

[15] Pieter Moree and Peter Stevenhagen. A two-variable Artin conjecture. *J. Number Theory*, 85(2):291–304, 2000.

[16] Władysław Narkiewicz. *Elementary and analytic theory of algebraic numbers*. Springer Monographs in Mathematics. Springer-Verlag, Berlin, third edition, 2004.

[17] R. W. K. Odoni. The Galois theory of iterates and composites of polynomials. *Proc. London Math. Soc. (3)*, 51(3):385–414, 1985.

[18] R. W. K. Odoni. On the prime divisors of the sequence $w_{n+1} = 1 + w_1 \cdots w_n$. *J. London Math. Soc. (2)*, 32(1):1–11, 1985.

[19] Brian Rice. Primitive prime divisors in polynomial arithmetic dynamics. *Integers*, 7:A26, 16 pp. (electronic), 2007.

[20] Joseph H. Silverman. *The arithmetic of dynamical systems*, volume 241 of *Graduate Texts in Mathematics*. Springer, New York, 2007.

[21] P. J. Stephens. Prime divisors of second-order linear recurrences. I. *J. Number Theory*, 8(3):313–332, 1976.

[22] Michael Stoll. Galois groups over **Q** of some iterated polynomials. *Arch. Math. (Basel)*, 59(3):239–244, 1992.

[23] A. J. van der Poorten. Some problems of recurrent interest. In *Topics in classical number theory, Vol. I, II (Budapest, 1981)*, volume 34 of *Colloq. Math. Soc. János Bolyai*, pages 1265–1294. North-Holland, Amsterdam, 1984.

[24] Nicolai N. Vorobiev. *Fibonacci numbers*. Birkhäuser Verlag, Basel, 2002. Translated from the 6th (1992) Russian edition by Mircea Martin.

R. Jones
Department of Mathematics
University of Wisconsin-Madison
480 Lincoln Dr.
Madison, WI 53706
USA
jones@math.wisc.edu