Zeitschrift: Commentarii Mathematici Helvetici
Herausgeber: Schweizerische Mathematische Gesellschaft
Band: 89 (2014)

Artikel: Galois theory of quadratic rational functions
Autor: Jones, Rafe / Manes, Michelle
DOI: https://doi.org/10.5169/seals-515669

Comment. Math. Helv. 89 (2014), 173–213
DOI 10.4171/CMH/316

Commentarii Mathematici Helvetici

Galois theory of quadratic rational functions

Rafe Jones and Michelle Manes*

Abstract. For a number field K with absolute Galois group G_K, we consider the action of G_K on the infinite tree of preimages of $\alpha \in K$ under a degree-two rational function $\phi \in K(x)$, with particular attention to the case when ϕ commutes with a non-trivial Möbius transformation. In a sense this is a dynamical systems analogue to the ℓ-adic Galois representation attached to an elliptic curve, with particular attention to the CM case. Using a result about the discriminants of numerators of iterates of ϕ, we give a criterion for the image of the action to be as large as possible. This criterion is in terms of the arithmetic of the forward orbits of the two critical points of ϕ. In the case where ϕ commutes with a non-trivial Möbius transformation, there is in effect only one critical orbit, and we give a modified version of our maximality criterion. We prove a Serre-type finite-index result in many cases of this latter setting.

Mathematics Subject Classification (2010). 37P15, 11R32.

Keywords. Galois representations, arboreal Galois representations, quadratic rational maps, arithmetic dynamics, iteration of rational functions, ramification in iterated towers.

1. Introduction

Let K be a number field, and $\phi \in K(z)$ a rational function of degree $d \geq 2$. Put

$$\phi^n = \underbrace{\phi \circ \phi \circ \cdots \circ \phi}_{n},$$

and denote by $\phi^{-n}(\alpha)$ the set of preimages of the point α in $\mathbb{P}^1(\bar{K})$ under the map ϕ^n. To the pair (ϕ, α), where $\alpha \in K$, we associate a tree of preimages: let $V_n = \phi^{-n}(\alpha)$, and give the set $T_\alpha = \bigsqcup_{n \geq 1} V_n$ the structure of a tree with root α by assigning edges according to the action of ϕ. See Figure 2 for examples. Because elements of $\mathrm{Gal}(\bar{K}/K)$ commute with ϕ, we obtain a map

$$\rho\colon \mathrm{Gal}(\bar{K}/K) \to \mathrm{Aut}(T_\alpha),$$

where $\mathrm{Aut}(T_\alpha)$ denotes the group of tree automorphisms of T_α. We call ρ the *arboreal Galois representation* attached to (ϕ, α), and the main goal of the present work is to study the image of ρ in the case where ϕ is a degree-two rational function.

*The first author's research was partially supported by NSF grant DMS-0852826, and the second author's by NSF grant DMS-1102858.

Given a prime ℓ and an elliptic curve E defined over a number field K, we obtain the ℓ-adic Galois representation $\rho_E : \mathrm{Gal}(\bar{K}/K) \to \mathrm{GL}_2(\mathbb{Z}_\ell)$ in much the same manner as the previous paragraph. The ℓ-adic Tate module $T_\ell(E)$ is the inverse limit of the sets $[\ell]^{-n}(O)$, and the action of $\mathrm{Gal}(\bar{K}/K)$ on $T_\ell(E)$ gives ρ_E. In this context Serre [16] proved that for a given elliptic curve E without complex multiplication the image of ρ_E has finite index in $\mathrm{GL}_2(\mathbb{Z}_\ell)$ for all ℓ, and ρ_E is surjective for all but finitely many ℓ. In the case where E has complex multiplication and the full endomorphism ring is contained in the ground field K, similar statements hold (see e.g. [16], p. 302), provided that $\mathrm{GL}_2(\mathbb{Z}_\ell)$ is replaced by the largest subgroup of $\mathrm{GL}_2(\mathbb{Z}_\ell)$ that commutes with the action on the Tate module induced by the extra endomorphisms of E. In general this is a Cartan subgroup.

In this paper we formulate a similar conjecture for quadratic rational functions and prove it in certain cases. To do so, we develop a general theory of arboreal representations associated to quadratic rational functions. In contrast to the situation for ρ_E, there appears to be no finite quotient G of the target group such that surjectivity of the induced representation into G implies surjectivity of ρ (see [6] for details). Rather, infinitely many conditions must be checked, and by studying the ramification of ρ we give a formulation of these in terms of the critical orbits of ϕ (Corollary 3.8 and Theorem 4.6).

When ϕ commutes with a non-trivial $f \in \mathrm{PGL}_2(K)$ such that $f(\alpha) = \alpha$, the Galois action on T_α must commute with the action of f. Define the automorphism group of ϕ to be

$$A_\phi = \{f \in \mathrm{PGL}_2(\bar{K}) : \phi \circ f = f \circ \phi\},$$

and define

$$A_{\phi,\alpha} = \{f \in A_\phi : f(\alpha) = \alpha\}.$$

We know that A_ϕ is finite by Proposition 4.65 of [17].

Let G_∞ denote the image of $\rho : \mathrm{Gal}(\bar{K}/K) \to \mathrm{Aut}(T_\alpha)$, and let C_∞ denote the centralizer of the action of $A_{\phi,\alpha}$ on $\mathrm{Aut}(T_\alpha)$. Recall that $\phi(x) \in K(x)$ is said to be *post-critically finite* if the orbit of each of the critical points of ϕ is finite. Such maps have the property that the extension $K(T_\alpha)/K$ is ramified above only finitely many primes of K (see [2] and the remark following Theorem 3.2), and thus G_∞ is topologically finitely generated. In this work, we focus on the general case, and we do not consider post-critically finite maps.

Conjecture 1.1. *Let $\phi \in K(x)$ have degree $d = 2$ and let $\alpha \in K$. Suppose that ϕ is not post-critically finite. Then $[C_\infty : G_\infty]$ is finite.*

When $A_{\phi,\alpha}$ is trivial, Conjecture 1.1 has been proven only in the case of two families of quadratic polynomials ([8], Theorem 1.1 and first remark on p. 534), namely

$$f(x) = x^2 - kx + k, k \in \mathbb{Z} \quad \text{and} \quad f(x) = x^2 + kx - 1, k \in \mathbb{Z} \smallsetminus \{0, 2\},$$

for $\alpha = 0$. The key feature of these families is that the orbit of 0 is finite but not periodic, a property they share with the family in Conjecture 1.3 below. We establish here the first similar result for a rational function that is not conjugate to a polynomial.

Theorem 1.2. *Let* $\phi(x) = \dfrac{1+3x^2}{1-4x-x^2}$ *and* $\alpha = 0$. *Then for* $K = \mathbb{Q}$, $G_\infty \cong \mathrm{Aut}(T_\alpha)$.

The function in Theorem 1.2 is polynomial-like in that it has a periodic critical point, though here it is in a 2-cycle rather than being a fixed point. Moreover, the orbit of 0 under ϕ is finite but not periodic. See the discussion following Corollary 3.8 for a one-parameter family of such maps.

If f is a non-identity element of $A_{\phi,\alpha}$, then $f(\alpha) = \alpha$, and so $f(\phi^n(\alpha)) = \phi^n(\alpha)$ for all $n \geq 1$. Because f has exactly two fixed points, α is thus either in a cycle of ϕ of length at most two, or maps after one iteration onto a fixed point. In fact, under the hypotheses of Conjecture 1.1, α is either fixed by ϕ or maps to a fixed point of ϕ (see Section 2). Figure 1 shows possible pre-image trees under the hypotheses of Conjecture 1.1. Here,

$$V_1 = \phi^{-1}(\alpha) \setminus \{\alpha\}, \quad V_n = \phi^{-1}(V_{n-1}) \text{ for } n > 1, \quad \text{and} \quad T_\alpha = \bigsqcup_{n\geq 1} V_n. \tag{1}$$

Note that $K(\phi^{-n}(\alpha)) = K(V_n)$.

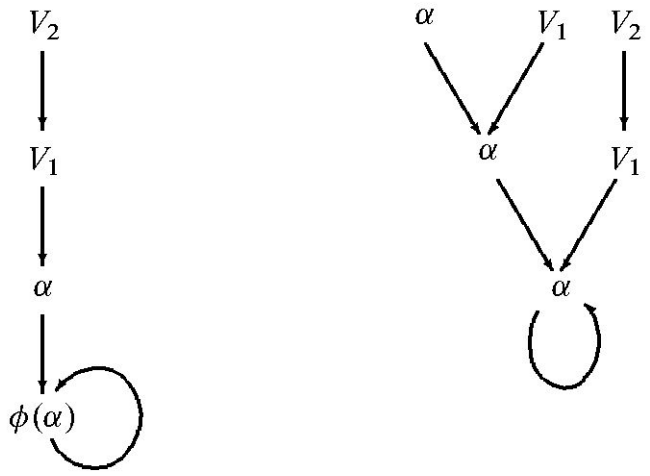

Figure 1. The first few levels of typical preimage trees when $\#A_{\phi,\alpha} > 1$ and $\deg \phi = 2$.

In Section 3, we give a criterion for a given pair (ϕ, α) to satisfy $[C_\infty : G_\infty] < \infty$ in the case $d = 2$. In Section 2 we show that in the case $\#A_{\phi,\alpha} > 1$, Conjecture 1.1 is equivalent to:

Conjecture 1.3. *Let $\phi(x) = k(x^2+1)/x$ and $\alpha = 0$, and suppose that ϕ is not post-critically finite. Then $[C_\infty : G_\infty] < \infty$.*

The conjectures above suggest a more general question.

Question 1.4. Let $\phi \in K(z)$ have degree $d \geq 2$ and $\alpha \in K$. Under what conditions is $[C_\infty : G_\infty]$ finite?

The post-critically finite maps of the form $\phi(x) = k(x^2+1)/x$ must have k with multiplicative height at most 2, and k can only be divisible by primes of K lying over (2) (see Proposition 5.1 for details). The structure of C_∞ in the setting of Conjecture 1.3 is described at the beginning of Section 4. In particular, C_∞ is an infinite-index subgroup of $\mathrm{Aut}(T_\alpha)$ with Hausdorff dimension 1/2 (see p. 192 for the definition); however, in contrast to Cartan subgroups of $\mathrm{GL}_2(\mathbb{Z}_\ell)$, it is highly non-abelian. Indeed, C_∞ has an index-two subgroup isomorphic to $\mathrm{Aut}(T_\alpha)$.

One of our main results is the following.

Theorem 1.5. *If K is a number field of odd degree over $\mathbb{Q}$, then Conjecture 1.3 is true for all k in a congruence class.*

We prove that Conjecture 1.3 is true in many other circumstances (see Corollary 5.11). For simplicity, we state here the result for the case $K = \mathbb{Q}$.

Theorem 1.6. *Conjecture 1.3 is true for $K = \mathbb{Q}$ provided k satisfies one of the following conditions (we write v_p for the p-adic valuation):*

- $v_2(k) = 0$ *or* $v_3(k) = 0$,
- $k \equiv 2, 3 \bmod 5$ *or* $k \equiv 1, 2, 5, 6 \bmod 7$,
- $v_p(2k \pm 1) > 0$ *for some* $p \equiv 3, 5 \bmod 8$,
- $v_p(2k^2 - k + 1) > 0$ *for some prime p with $-k$ not a square mod p, or*
- $v_p(2k^2 + k + 1) > 0$ *for some prime p with k not a square mod p.*

In the case where $k \in \mathbb{Z}$, we use Theorem 1.6 plus other results to verify Conjecture 1.3 for all k with $|k| \leq 10\,000$ (see the remark following Corollary 5.11). We also give the following sufficient conditions on k to ensure that the index in Conjecture 1.3 is one (see Theorem 5.13).

Theorem 1.7. *Let ϕ and α be as in Conjecture 1.3, and suppose that $K = \mathbb{Q}$. There exists an effectively computable set Σ of primes of $\mathbb{Z}$ of natural density zero, such that if $v_p(k) = 0$ for all primes belonging to Σ then $G_\infty \cong C_\infty$.*

For more on Σ, see Corollary 5.14 and the remark following. We note that all odd primes in Σ are congruent to 1 modulo 4, so that if k is an integer divisible only by primes congruent to 3 modulo 4, then Theorem 1.7 applies. We note that Theorem 1.7 may be far from best possible; indeed, we have been unable to find a single $k \in \mathbb{Z}$ for which $[G_\infty : C_\infty] > 1$. The fact that Σ has zero density is a consequence of our analysis of C_∞ and a result relating the structure of G_∞ to the density of prime divisors of orbits of a large class of rational functions (Theorem 6.1).

In order to prove our main results, we generalize techniques from [8], [14], [18] that treat case where ϕ is a polynomial. In particular, in Theorem 3.2 we obtain a formula for the discriminant of the numerator of ϕ^n, where ϕ is a rational function of degree $d \geq 2$, a problem that has interest in its own right (see [2]). In the case $d = 2$, we examine the irreducibility of the numerators of ϕ^n, both in the general quadratic case (Theorem 3.5) and in the case $\phi(x) = k(x^2 + b)/x$ (Theorem 4.5). We also analyze the extensions K_n/K_{n-1}, where $K_i = K(\phi^{-i}(\alpha))$. We give a criterion for $[K_n : K_{n-1}]$ to be as large as possible, both in the general quadratic case (Corollary 3.8) and in the case where $\phi(x) = k(x^2 + b)/x$ (Theorem 4.6). We use the former criterion to prove Theorem 1.2 (see the discussion following Corollary 3.8). The criteria for both irreducibility and maximality of the field extensions are arithmetic, and depend on knowledge of primes dividing elements of the form $\phi^n(\gamma)$, where γ is a critical point of ϕ. We assemble these pieces to prove the following result, which is the main engine behind Theorems 1.5, 1.6, and 1.7.

Theorem 1.8. *Let $\phi(x) = k(x^2 + 1)/x$ and $\alpha = 0$, and suppose that ϕ is not post-critically finite. Put*

$$\delta_n = \begin{cases} 2k^2 & \text{if } n = 1, \\ \delta_{n-1}^2 + \epsilon_{n-1}^2 & \text{if } n \geq 2, \end{cases} \quad \text{and} \quad \epsilon_n = \begin{cases} k & \text{if } n = 1, \\ \frac{\delta_{n-1}\epsilon_{n-1}}{k} & \text{if } n \geq 2. \end{cases}$$

If none of -1, δ_n, or $-\delta_n$ is a square for $n \geq 2$, then $[G_\infty : C_\infty]$ is finite.

See Theorem 5.3 for a slightly more general statement. The sequence (δ_n, ϵ_n) is related to the orbit of the critical point 1 of ϕ in that $\phi^n(1) = \delta_n/\epsilon_n$. What makes Theorem 1.8 possible is the fact that 0 is a pre-periodic point for ϕ, which ensures that the set of common prime ideal divisors of δ_i and δ_j is very restricted (see Lemma 5.2). This allows one to show that except in very special circumstances there must be a primitive prime divisor $\mathfrak{p}$ of δ_n (that is, $\mathfrak{p}$ does not divide δ_i for $i < n$) that divides δ_n to odd multiplicity. This is the key hypothesis of Theorem 4.6. The condition that 0 is pre-periodic has also been used to study primitive prime divisors in other dynamical sequences (see [3], [15]). A natural hope is that similar techniques might be used to tackle Conjecture 1.1 and Question 1.4, even in the case where A_ϕ is trivial.

2. Preliminaries and notation

In this section we fix some notation, and we show how to reduce Conjecture 1.1 to Conjecture 1.3 when $\#A_{\phi,\alpha} > 1$. For any $g \in \mathrm{PGL}_2(\bar{K})$, define the conjugate map

$$\phi^g = g\phi g^{-1}.$$

Proposition 2.1. *Suppose $\phi \in K(z)$ and basepoint $\alpha \in K$ satisfy $[C_\infty : G_\infty] < \infty$. Let $g \in \mathrm{PGL}_2(\bar{K})$ such that $\phi^g \in K(z)$. Then the finite index result also holds for ϕ^g with the basepoint $g(\alpha)$.*

Proof. To simplify notation, we let $\psi = \phi^g$. A computation reveals that

$$A_{\psi,g(\alpha)} = g \circ A_{\phi,\alpha} \circ g^{-1}. \tag{2}$$

First suppose $g \in \mathrm{PGL}_2(K)$. Then since

$$\psi^{-n}(g(\alpha)) = \{g(\beta) : \beta \in \phi^{-n}(\alpha)\}, \tag{3}$$

$\psi^{-n}(g(\alpha))$ and $\phi^{-n}(\alpha)$ generate the same extension of K.

Now suppose that ψ is a nontrivial twist of ϕ, meaning $g \in \mathrm{PGL}_2(L)$ for some finite extension L/K (here we take L minimal). From Lemma 2.6 of [10], there is an absolute bound B depending only on ϕ so that $[L : K] \leq B$. (In fact, this bound B can be chosen to depend only on the degree of ϕ and not on the specific map.)

By equation (3), for each $n \geq 1$ an extension of $\phi^{-n}(\alpha)$ of degree at most B contains $K(\psi^{-n}(g(\alpha)))$. The finite index result follows. □

When $\deg \phi = 2$, A_ϕ is either trivial, cyclic of order two, or isomorphic to S_3 [12]. The third option occurs if and only if ϕ is conjugate over $\bar{K}$ to $1/z^2$, in which case ϕ is post-critically finite. Hence, if we assume $\#A_{\phi,\alpha} > 1$ in Conjecture 1.1, then necessarily $\#A_{\phi,\alpha} = 2$ and $A_{\phi,\alpha} = A_\phi$. In this case we have from Lemma 1 of [11] that ϕ is conjugate over $\bar{K}$ to

$$\psi(z) = k(z^2 + 1)/z, \quad \text{with } k \in K^* \smallsetminus \{0, -1/2\}.$$

Let g denote the conjugacy such that ϕ^g has the form above. By equation (2), we conclude that $\#A_{\psi,g(\alpha)} = 2$. Therefore $A_{\psi,g(\alpha)} = \{\mathrm{id}, z \mapsto -z\}$, and since $g(\alpha)$ must be fixed by elements of this set, it follows that $g(\alpha) = 0$ or $g(\alpha) = \infty$. Hence $g(\alpha)$ is either fixed by ψ or maps to a fixed point, so α is either fixed by ϕ or maps to a fixed point of ϕ, as shown in Figure 1. These two cases are illustrated in the specific case $k = 1$ in Figure 2.

When $g(\alpha) = \infty$, using the notation of Figure 1, we have $V_1 = \{0\}$, and thus $V_n = \phi^{-(n-1)}(0)$. It follows that the arboreal representation is the same as the case $g(\alpha) = 0$. Thus to prove Conjecture 1.1, we need only consider pairs of the form

$$(\phi, \alpha) = \big(k(z^2 + 1)/z, 0\big).$$

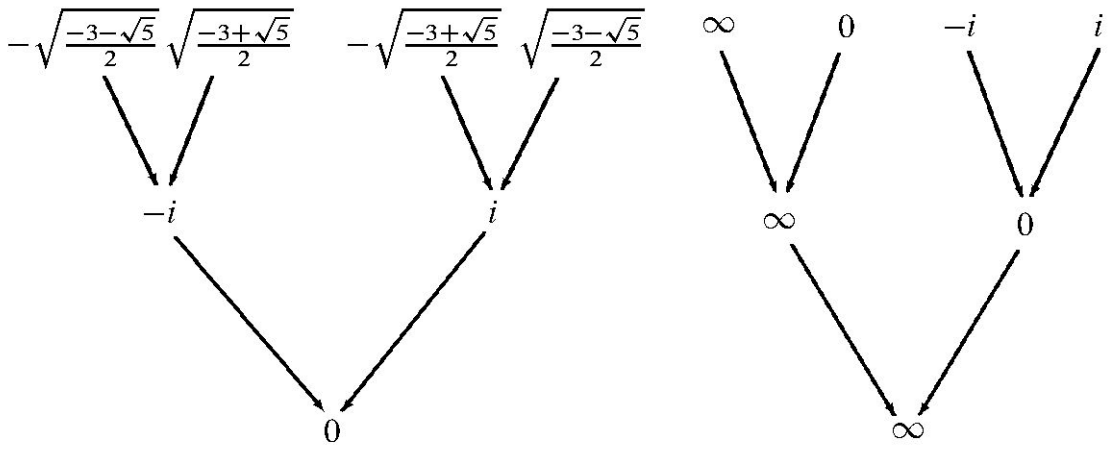

Figure 2. First few levels of the preimage trees of 0 and ∞ under $\phi(x) = (x^2 + 1)/x$.

Therefore to establish Conjecture 1.1, it is enough to prove Conjecture 1.3.

We return now to the general case and establish some notation. Let K be a number field, and $\phi\colon \mathbb{P}^1 \to \mathbb{P}^1$ a rational function defined over K. Suppose that $\phi([X, Y]) = [P(X, Y), Q(X, Y)]$ in homogeneous coordinates, with $P(X, Y)$ and $Q(X, Y)$ having no common factors of positive degree. Fix particular choices of P and Q, and let

$$\Phi(X, Y)\colon \mathbb{A}^2 \to \mathbb{A}^2, \quad (X, Y) \mapsto (P(X, Y), Q(X, Y)),$$

be a lift of ϕ. Define $P_n, Q_n \in K[X, Y]$ by

$$\Phi^n(X, Y) := (P_n(X, Y), Q_n(X, Y)),$$

where

$$P_n(X, Y) = P_{n-1}(P(X, Y), Q(X, Y))$$

and

$$Q_n(X, Y) = Q_{n-1}(P(X, Y), Q(X, Y)),$$

or equivalently

$$P_n(X, Y) = P(P_{n-1}(X, Y), Q_{n-1}(X, Y))$$

and

$$Q_n(X, Y) = Q(P_{n-1}(X, Y), Q_{n-1}(X, Y)).$$

Note that $\phi^n([X, Y]) = [P_n(X, Y), Q_n(X, Y)]$, though using homogeneous coordinates may involve cancellation of some common constant factors.

We use lower-case letters to denote de-homogenizations, and summarize our notation:

$\phi([X,Y]) = [P(X,Y), Q(X,Y)]$	a rational map on $\mathbb{P}^1$ of degree d
$\phi(x) = p(x)/q(x)$	the dehomogenization of ϕ
$\Phi(X,Y) = (P(X,Y), Q(X,Y))$	natural lift of ϕ to a map on $\mathbb{A}^2$
$\Phi^n(X,Y) = (P_n(X,Y), Q_n(X,Y))$	nth iterate of Φ
$p_n(x) = P_n(x,1), q_n(x) = Q_n(x,1)$	dehomogenized versions of P_n, Q_n
$A_\phi = \{f \in \mathrm{PGL}_2(\bar{K}) : \phi \circ f = f \circ \phi\}$	the automorphism group of ϕ
$A_{\phi,\alpha} = \{f \in A_\phi : f(\alpha) = \alpha\}$	
$\ell(R)$	leading coefficient of the polynomial R
d_R	the degree of R
V_n	$\phi^{-n}(0)$, unless 0 is periodic; see (1)
$T = \bigcup_{i \geq 1} V_n$	preimage tree of ϕ with root 0
$T_n = \bigcup_{1 \leq i \leq n} V_n$	truncation of T to level n
$K_n = K(V_n) = K(T_n)$	
$K_\infty = \bigcup_n K_n = K(T)$	
$G_n = \mathrm{Gal}\,(K_n/K)$	
$G_\infty = \mathrm{Gal}\,(K_\infty/K) = \varprojlim G_n$	

When we refer to a "separable polynomial," we mean that the polynomial has distinct roots. We also adopt the convention that $\alpha = 0$ and that ∞ does not appear in the pre-image tree of $\alpha = 0$; that is, we assume $\phi^n(\infty) \neq 0$. These assumptions make the statements and proofs of our results in Section 3 much simpler, and come at no cost. Indeed, as noted earlier in this section, the representations associated to (ϕ, α) and $(\phi^g, g(\alpha))$ are the same for $g \in \mathrm{PGL}_2(K)$. Choosing g with $g(\alpha) = 0$ then reduces to the case $\alpha = 0$. We may similarly require that $g(\beta) = \infty$, where $\beta \in K$ is any point disjoint from the preimage tree of α.

Because $\alpha = 0$, G_n is the Galois group of the de-homogenized polynomial $p_n(x) = P_n(x,1) \in K[x]$. We frequently move back and forth between $P_n(X,Y)$ and $p_n(x)$. The de-homogenized version of the recursion for P_n is

$$p_n(x) = q(x)^{\deg P_{n-1}}\, p_{n-1}(p(x)/q(x)) = q(x)^{d^{n-1}}\, p_{n-1}(p(x)/q(x)), \quad (4)$$

or equivalently

$$p_n(x) = q_{n-1}(x)^{\deg P}\, p(p_{n-1}(x)/q_{n-1}(x)) = q_{n-1}(x)^d\, p(p_{n-1}(x)/q_{n-1}(x)), \quad (5)$$

where in both cases $d = \deg P$. These hold for all x with $q(x) \neq 0$ and $q_{n-1}(x) \neq 0$, respectively.

3. Discriminants, irreducibility, and Galois theory of rational functions

We begin with results concerning the discriminant of the numerator of an iterate of a rational function. We then consider the case $d = 2$, prove results on the irreducibility of such polynomials, and then apply these results to the question of under what conditions the degree of the extension $[K_n : K_{n-1}]$ is as large as possible. These results hold without consideration of the automorphism group A_ϕ of ϕ. In Section 4 we examine the case where $\#A_\phi = 2$.

Throughout, we denote the degree of a polynomial $s(x) \in K[x]$ by d_s and its leading coefficient by $\ell(s)$. We recall the *resultant* of two polynomials $g_1, g_2 \in K[x]$ may be defined as

$$\operatorname{Res}(g_1, g_2) = \ell(g_1)^{d_{g_2}} \prod_{g_1(\alpha)=0} g_2(\alpha),$$

and is a homogeneous polynomial in the coefficients of g_1 and g_2 that vanishes if and only if g_1 and g_2 have a common root in $\bar{K}$. We will also make use of the following basic equality:

$$\prod_{g_1(\beta)=0} g_2(\beta) = (-1)^{d_{g_1} d_{g_2}} \ell(g_2)^{d_{g_1}} \ell(g_1)^{-d_{g_2}} \prod_{g_2(\alpha)=0} g_1(\alpha). \tag{6}$$

The discriminant of the polynomial p_n will prove to be a fundamental tool in what follows. However, p_n is constructed as a de-homogenized polynomial P_n, and P_n is given by a double-recursion. Standard results on calculating discriminants do not apply in this more complicated situation; we need a new tool. We begin with the somewhat simpler case of calculating the discriminant of a single polynomial that is the de-homogenization of a function of two other (homogeneous) polynomials. In Theorem 3.2, we apply this result recursively to find a discriminant formula for p_n.

Lemma 3.1. *Let $F, P, Q \in K[X, Y]$ be non-constant homogeneous polynomials with $\deg P = \deg Q = d$ and P and Q having no common roots in $\mathbb{P}^1(\bar{K})$. Let $H(X, Y) = F(P(X, Y), Q(X, Y))$, and let $h, f, p, q \in K[x]$ denote the de-homogenizations of H, F, P, and Q, respectively. Let $c = qp' - pq'$. Finally, assume that $H(1, 0) \neq 0$ and $F(1, 0) \neq 0$. Then*

$$\operatorname{Disc} h = \pm \ell(h)^{k_1} \ell(q)^{k_2} \ell(f)^{k_3} \ell(c)^{k_4} (\operatorname{Disc} f)^d (\operatorname{Res}(q, p))^{d_f(d_f-2)} \prod_{c(\gamma)=0} h(\gamma),$$

where

$$k_1 = d_f d - 2 - d_c - d_q(d_f - 2), \quad k_2 = d_f(d - d_p)(d_f - 2),$$
$$k_3 = (d_q - d)(d_f - 2), \quad \text{and} \quad k_4 = d_f d.$$

Proof. By definition,

$$\operatorname{Disc} h = \pm \ell(h)^{-1} \operatorname{Res}(h, h') = \pm \ell(h)^{d_h - 2} \prod_{h(\alpha)=0} h'(\alpha). \tag{7}$$

De-homogenizing H gives $h(x) = q(x)^{d_F} f\,(p(x)/q(x))$ provided $q(x) \neq 0$, and thus

$$h' = d_F q^{d_F - 1} q' \cdot f(p/q) + q^{d_F} f'(p/q) \cdot (p/q)', \tag{8}$$

assuming $q(x) \neq 0$. Since $F(1,0) \neq 0$, no root α of h can satisfy $q(\alpha) = 0$. This implies that for each α with $h(\alpha) = 0$, we have $f(p(\alpha)/q(\alpha)) = 0$, so the first summand in (8) vanishes when $x = \alpha$. We may thus rewrite the right side of (7) as

$$\pm\, \ell(h)^{d_h - 2} \Big(\prod_{h(\alpha)=0} q(\alpha) \Big)^{d_F - 2} \prod_{h(\alpha)=0} (f' \circ p/q)(\alpha) \prod_{h(\alpha)=0} (p'q - q'p)(\alpha). \tag{9}$$

Using (6), the first product in (9) is equal to $\big(\ell(q)^{d_h} \ell(h)^{-d_q} \prod_{q(\pi)=0} h(\pi)\big)^{d_F - 2}$. Moreover, for each root π of q we have

$$h(\pi) = H(\pi, 1) = F\,(P(\pi, 1), Q(\pi, 1)) = F\,(P(\pi, 1), 0) = \ell(f)\, p(\pi)^{d_F}.$$

Hence the first product in (9) becomes

$$\begin{aligned} &\Big(\ell(q)^{d_h} \ell(h)^{-d_q} \ell(f)^{d_q} \prod_{q(\pi)=0} p(\pi)^{d_f} \Big)^{d_F - 2} \\ &\qquad = \big(\ell(q)^{d_h} \ell(h)^{-d_q} \ell(f)^{d_q} \big(\ell(q)^{-d_p} \operatorname{Res}(q, p) \big)^{d_f} \big)^{d_F - 2}. \end{aligned}$$

Turning to the second product in (9), we have already noted that $h(\alpha) = 0$ implies $p(\alpha)/q(\alpha)$ is a root of f. Moreover, for each root β of f, there are with multiplicity precisely d elements α with $p(\alpha)/q(\alpha) = \beta$ (this is ensured by the assumption that $H(1,0) \neq 0$). Thus as α runs over all roots of h, $(p/q)(\alpha)$ runs over all roots of f, hitting each one d times. Hence the second product in (9) equals

$$\Big(\prod_{f(\beta)=0} f'(\beta) \Big)^{d} = \Big(\ell(f)^{-(d_f - 2)} \operatorname{Disc}(f) \Big)^{d}.$$

From (6), the third product in (9) equals

$$\pm \ell(c)^{d_h} \ell(h)^{-d_c} \prod_{c(\gamma)=0} h(\gamma).$$

Gathering the terms containing $\ell(h)$ and $\ell(f)$, and using the fact that $d_f = d_F$ (since $F(1,0) \neq 0$) and $d_h = d d_f$ (since $H(1,0) \neq 0$) completes the proof. □

Remark. Note the conditions $H(1,0) = 0$ and $F(1,0) = 0$ correspond to our assumption that ∞ is not in the pre-image tree of $\alpha = 0$. These assumptions greatly ease an (already complicated) calculation, but it is certainly possible to obtain similar formulas in the case that $H(1,0) = 0$ or $F(1,0) = 0$. In these cases F factors as a product F_1F_2 with $F_2(1,0) \neq 0$ and $H_2(1,0) \neq 0$, where $H_2 := F_2(P,Q)$. One then splits the product on the right side of (7) into the product over the roots of h_2 and the product over the remaining roots of h.

Theorem 3.2. *Let $\phi = p(x)/q(x) \in K(x)$ be a rational function of degree $d \geq 2$, let $n \geq 2$, and define p_n and q_n recursively so that $\phi^n = p_n(x)/q_n(x)$. Let $c = qp' - pq'$. Assume that $\phi^n(\infty) \neq 0$ and $\phi^{n-1}(\infty) \neq 0$. If $\phi(\infty) \neq \infty$, then*

$$\begin{aligned} \text{Disc } p_n = \pm \ell(p_n)^{k_1} \ell(q)^{k_2} \ell(c)^{k_3} (\text{Disc } p_{n-1})^d \\ (\text{Res}\,(q,p))^{d^{n-1}(d^{n-1}-2)} \prod_{c(\gamma)=0} p_n(\gamma), \end{aligned} \tag{10}$$

where

$$k_1 = 2d - 2 - d_c, \quad k_2 = d^{n-1}(d - d_p)(d^{n-1} - 2), \quad \text{and} \quad k_3 = d^n.$$

If $\phi(\infty) = \infty$, then

$$\text{Disc } p_n = \pm \ell(p)^{k_1} \ell(c)^{k_2} (\text{Disc } p_{n-1})^d \, (\text{Res}\,(q,p))^{d^{n-1}(d^{n-1}-2)} \prod_{c(\gamma)=0} p_n(\gamma), \tag{11}$$

where

$$k_1 = d^{2n-1} - d_q(d^{2n-2} - 2d^{n-1}) - d_c(1-d^n)/(1-d) - 2 \quad \text{and} \quad k_2 = d^n. \tag{12}$$

Remark. Suppose that ϕ is post-critically finite, so the forward orbit of each γ with $c(\gamma) = 0$ is finite. An induction on equation (4) shows that the set of primes dividing $p_n(\gamma)$ for all critical points γ and all $n \geq 1$ is finite. If $\phi(\infty) = \infty$, induction on equation (11) then shows that the set of primes dividing Disc(p_n) for any n is likewise finite.

When $\phi(\infty) \neq \infty$ and $d_c = 2d - 2$ (in other words, when ∞ is not a critical point), the term $\ell(p_n)$ does not contribute to the product in equation (10), so the set of primes dividing Disc(p_n) for any n is finite in this case as well. When $\phi(\infty) \neq \infty$ and ∞ is a critical point, then from (10), we have that Disc p_n is divisible by $\ell(p_n)$. However, $\ell(p_n) = P_n(1,0)$, and since by assumption ϕ is post-critically finite and ∞ is a critical point, $P_n(1,0)$ can take on only finitely many values as n varies.

Hence, Theorem 3.2 shows that if ϕ is post-critically finite, then there is a finite set of primes S such that for every $n \geq 1$, Disc p_n is divisible only by primes in S. In this case, then, the field K_∞ is ramified over only finitely many primes of K. (For another version of this result, see [2].)

Proof. As $\phi^n(\infty) \neq 0$, and $\phi^{n-1}(\infty) \neq 0$, we have $P_n(1,0) \neq 0$ and $P_{n-1}(1,0) \neq 0$, respectively. We may thus apply Lemma 3.1 with $F = P_{n-1}$ and $H = P_n$. Note also that $\phi^{n-1}(\infty) \neq 0$ implies that $\deg p_{n-1} = d^{n-1}$. Moreover, if we assume that $\phi(\infty) \neq \infty$, then $d_q = d$. Lemma 3.1 then immediately gives formula (10).

Assuming now that $\phi(\infty) = \infty$, we have $d_p = d$, which kills the $\ell(q)$ term in Lemma 3.1. Further, $\phi(\infty) = \infty$ also implies that for all k, $\ell(p_k) = \ell(p)\ell(p_{k-1})^d$, and an induction gives $\ell(p_k) = \ell(p)^{(1-d^k)/(1-d)}$. The power of $\ell(p)$ in the expression in Lemma 3.1 is thus

$$\frac{(1-d^n)(d^n-2-d_c-d_q(d^{n-1}-2))}{1-d} + \frac{(1-d^{n-1})(d_q-d)(d^{n-1}-2)}{1-d}.$$

This simplifies to the value of k_1 given in equation (12). □

We now consider the irreducibility of the p_n in the case of quadratic rational functions.

Lemma 3.3. *Let $\phi(x) \in K(x)$ have degree 2, let t be a parameter, and let $\gamma_1, \gamma_2 \in \mathbb{P}^1(\bar{K})$ be the critical points of ϕ. Then there exists $C \in K$ such that*

$$\operatorname{Disc}_x (p(x) - tq(x)) = C \prod_{\phi(\gamma_i)\neq\infty} (\phi(\gamma_i) - t)\,. \tag{13}$$

Moreover, if $p(x)$ is separable, then $C = \operatorname{Disc} p(x) \cdot \prod_{\phi(\gamma_i)\neq\infty} (\phi(\gamma_i)^{-1})$.

Remark. This is a special case of a more general phenomenon; see [2], Proposition 1.

Proof. Note that both sides of (13) are in $K[t]$ (the right side because the $\phi(\gamma_i)$ are either rational or Galois-conjugate). We show that the roots of $\operatorname{Disc}_x (p(x) - tq(x))$ in $\bar{K}$ are precisely the $\phi(\gamma_i)$ with $\phi(\gamma_i) \neq \infty$, and this is enough to establish the lemma.

Note first that since ϕ is quadratic, by the Riemann–Hurwitz formula we must have $\gamma_1 \neq \gamma_2$ and $\phi(\gamma_1) \neq \phi(\gamma_2)$. For a given $t \in \bar{K}$, the degree-two homogeneous polynomial $P(X,Y) - tQ(X,Y)$ has a single root in $\mathbb{P}^1(\bar{K})$ if and only if $\phi^{-1}(t)$ has a single element, which occurs precisely when $t = \phi(\gamma_i)$. Thus $\operatorname{Disc}(P(X,Y) - tQ(X,Y))$ vanishes if and only if $t = \phi(\gamma_i)$, and since $t \in \bar{K}$ we cannot have $t = \infty$. Finally, both $P(X,Y) - tQ(X,Y)$ and $p(x) - tq(x)$ have degree 2, and thus $\operatorname{Disc}(P(X,Y) - tQ(X,Y))$ is the same as $\operatorname{Disc}_x (p(x) - tq(x))$.

The last statement of the proposition comes from setting $t = 0$ in (13) and noting that the separability of $p(x)$ implies $\operatorname{Disc} p(x)$ does not vanish, and therefore $\prod_{\phi(\gamma_i)\neq\infty} (\phi(\gamma_i))$ cannot vanish either. □

Lemma 3.4. *Let $\phi(x) \in K(x)$ have degree 2, and for each $i \geq 0$, denote by K_i the splitting field of p_i. Assume that p_{n-1} is irreducible in $K[x]$, let α be a root of p_{n-1}, and let $\gamma_1, \gamma_2 \in \mathbb{P}^1(\bar{K})$ be the critical points of ϕ. Then there exists $C \in K$ such that p_n is irreducible in $K[x]$ if and only if*

$$C \prod_{\phi(\gamma_i)\neq\infty} (\phi(\gamma_i) - \alpha) \notin K_{n-1}^{*2}.$$

If $p(x)$ is separable, then C has the same value as in Lemma 3.3.

Proof. Denote the roots of p_{n-1} by $\alpha_1, \ldots, \alpha_r$, and take $\alpha_1 = \alpha$. Since p_{n-1} is irreducible, the α_i are Galois conjugates, and hence the action of Galois on the roots of p_n is either as a single orbit of $2r$ elements (and thus p_n is irreducible) or two orbits of r elements. The latter case holds if and only if each orbit contains exactly one element in each fiber $\phi^{-1}(\alpha_i)$, or equivalently if and only if the roots of $p(x)-\alpha_i q(x)$ are not conjugate for any i. This holds if and only if the $p(x)-\alpha_i q(x)$ are all reducible over K_{n-1}. Because the α_i are all conjugate, this is equivalent to the reducibility of $p(x) - \alpha q(x)$ over K_{n-1}, which occurs precisely when $\mathrm{Disc}\,(p(x) - \alpha q(x))$ is a square in K_{n-1}. The result now follows from Lemma 3.3. □

Recall that G_n is the Galois group of the splitting field of the polynomials p_n. To understand when G_n is as large as possible, it will be necessary to have conditions under which the polynomials p_n are irreducible. We now give a criterion for the irreducibility of p_n assuming that p_{n-1} is irreducible and has even degree. Note that the criterion here is sufficient but not necessary. The result is useful in that it applies to all degree 2 rational maps, but unfortunately the hypotheses are not satisfied in the case of quadratic maps with a nontrivial automorphism. We will need a refinement of this result in that case, which we provide in Theorem 4.5.

Theorem 3.5. *Let $\phi(x) \in K(x)$ have degree 2. Suppose that $n \geq 2$, and that p_{n-1} is irreducible in $K[x]$ and has even degree. Let $\ell(p_{n-1})$ be the leading coefficient of p_{n-1}, let $\gamma_1, \gamma_2 \in \mathbb{P}^1(\bar{K})$ be the critical points of ϕ, and without loss say $\phi(\gamma_1) \neq \infty$. If $\phi(\gamma_2)$ is not (resp. is) ∞, then p_n is irreducible in $K[x]$ provided*

$$p_{n-1}(\phi(\gamma_1)) \cdot p_{n-1}(\phi(\gamma_2)) \notin K^{*2} \quad (\textit{resp. } \ell(p_{n-1}) \cdot p_{n-1}(\phi(\gamma_1)) \notin K^{*2}). \quad (14)$$

Remark. The condition that p_{n-1} have even degree is implied by $\phi^{n-1}(\infty) \neq 0$. Moreover, if $\gamma_i \neq \infty$, then from (4) and the assumption $n \geq 2$, $p_{n-1}(\phi(\gamma_i)) = p_n(\gamma_i)$ up to squares. Thus if both γ_1 and γ_2 are finite, then (14) becomes

$$p_n(\gamma_1)\, p_n(\gamma_2) \notin K^{*2} \quad (\text{resp. } \ell(p_{n-1})\, p_n(\gamma_1) \notin K^{*2}).$$

Proof. By Lemma 3.4, we must show

$$C \prod_{\phi(\gamma_i)\neq\infty} (\phi(\gamma_i) - \alpha) \notin K_{n-1}^{*2},$$

for some root α of p_{n-1}, and for this it is sufficient to show that the norm of the left side as an element of K_{n-1}/K is not a square in K. This norm equals

$$C^{\deg p_{n-1}} \prod_{i=1}^{2} \prod_{j=1}^{r} \big(\phi(\gamma_i) - \alpha_j\big) \quad \text{(resp. } C^{\deg p_{n-1}} \prod_{j=1}^{r} \big(\phi(\gamma_1) - \alpha_j\big)\text{)},$$

where $\alpha_1, \ldots, \alpha_r$ denote the roots of p_{n-1}. This is the same as

$$C^{\deg p_{n-1}} \ell(p_{n-1})^{-2} \prod_{i=1}^{2} p_{n-1}(\phi(\gamma_i)) \quad \text{(resp. } C^{\deg p_{n-1}} \ell(p_{n-1})^{-1} p_{n-1}(\phi(\gamma_i))\text{)}.$$

Since $\deg p_{n-1}$ is even, $C^{\deg p_{n-1}}$ is a square. □

For the remainder of this section, we let n be fixed, and assume that $\phi^n(\infty) \neq 0$ and $\phi^{n-1}(\infty) \neq 0$. We also assume that p_n is separable, which by Theorem 3.2 is equivalent to $\phi^i(\gamma) \neq 0$, $i = 1, \ldots, n$ for all critical points γ of ϕ. Together, these assumptions imply that there are d^n distinct roots $\alpha_1, \ldots, \alpha_{d^n}$ of p_n, and d^{n-1} distinct roots $\beta_1, \ldots, \beta_{d^{n-1}}$ of p_{n-1}. Moreover, the α_i are precisely the roots of $p(x) - \beta_j q(x)$, for $j = 1, \ldots, d^{n-1}$. Recall that $K_n = K(\alpha_1, \ldots, \alpha_{d^n})$ and $K_{n-1} = K(\beta_1, \ldots, \beta_{d^{n-1}})$.

We examine the extension K_n/K_{n-1} in the case $d = 2$ and give conditions that ensure it is as large as possible. Recall that $G_n = \mathrm{Gal}\,(K_n/K)$. The assumptions of the previous paragraph imply an injection $G_n \hookrightarrow \mathrm{Aut}(T_n)$, where T_n is the complete binary rooted tree of height n. Restriction gives a homomorphism $\mathrm{Aut}(T_n) \to \mathrm{Aut}(T_{n-1})$, whose kernel is generated by the transpositions swapping a single pair of vertices at level n, both connected to a given vertex at level $n-1$. Thus the kernel is isomorphic to $(\mathbb{Z}/2\mathbb{Z})^{2^{n-1}}$. Hence $\mathrm{Gal}\,(K_n/K_{n-1})$ must inject into this group. We now show how one can see this directly from the way that K_n is constructed from K_{n-1}; this point of view will also be the most useful for establishing our maximality results.

Because $\deg \phi = 2$, $p(x) - \beta_j q(x) \in K_{n-1}(x)$ is a quadratic polynomial. Note that K_n is obtained from K_{n-1} by adjoining the roots of $p(x) - \beta_j q(x)$ for $j = 1, \ldots, 2^{n-1}$, so we have that K_n is a 2-Kummer extension of K_{n-1}, and indeed letting

$$\delta_j = \mathrm{Disc}(p(x) - \beta_j q(x)), \text{ we have } K_n = K_{n-1}\Big(\sqrt{\delta_j} : j = 1, \ldots, 2^{n-1}\Big). \quad (15)$$

It follows that $\mathrm{Gal}\,(K_n/K_{n-1}) \hookrightarrow (\mathbb{Z}/2\mathbb{Z})^{2^{n-1}}$. Using Kummer theory (e.g. [9], Section VI.8), $[K_n : K_{n-1}]$ is the order of the group D generated by the classes of the δ_j in K_{n-1}^*/K_{n-1}^{*2}. Now,

$$\#D = \frac{2^{2^{n-1}}}{\#V}, \quad \text{where } V = \Big\{(e_1, \ldots, e_{2^{n-1}}) \in \mathbb{F}_2^{2^{n-1}} : \prod_j \delta_j^{e_j} \in K_{n-1}^{*2}\Big\}.$$

That is, V is the group of relations among the δ_j. One sees easily that V is an $\mathbb{F}_2$-vector space, and that the action of $G_{n-1} := \mathrm{Gal}\,(K_{n-1}/K)$ on the δ_j gives an action of G_{n-1} on V as linear transformations. It follows that V is an $\mathbb{F}_2[G_{n-1}]$-module.

The following lemma is due to M. Stoll [18]. We give the proof here for the sake of completeness.

Lemma 3.6 (Stoll). *Let Γ be a 2-group and $M \neq 0$ a $\mathbb{F}_2[\Gamma]$-module. Then the submodule M^Γ of Γ-invariant elements is non-trivial.*

Proof. Induct on $\#\Gamma$. Suppose $\Gamma = \{e, \sigma\} \cong \mathbb{Z}/2\mathbb{Z}$, and take $m \in M$ with $m \neq 0$. Then either $\sigma(m) = m$ or $m + \sigma(m) \neq 0$ (since M is an $\mathbb{F}_2$-module). In the former case, m is a nontrivial element of M^Γ, while in the latter case $m + \sigma(m)$ is a nontrivial element of M^Γ.

If $\#\Gamma > 2$, then let N be a nontrivial normal subgroup of Γ (possible since Γ is a 2-group). Then M is an $\mathbb{F}_2[N]$-module also, so by induction $M^N \neq 0$. However, M^N is an $\mathbb{F}_2[\Gamma/N]$-module, so again by induction $0 \neq (M^N)^{\Gamma/N} = M^\Gamma$. □

We now give a condition that will guarantee the extension $[K_n : K_{n-1}]$ is as large as possible. *A priori*, this result depends on deciding whether an element of K_{n-1} is a square. However, we can actually give a condition ensuring $[K_n : K_{n-1}] = 2^{2^{n-1}}$ solely in terms of the arithmetic of K. We provide such a condition in Corollary 3.8.

Theorem 3.7. *Let $\phi \in K(x)$ have degree 2 with $\phi^n(\infty) \neq 0$ and $\phi^{n-1}(\infty) \neq 0$. Suppose that $n \geq 2$ and that p_{n-1} is irreducible in $K[x]$. Let $\ell(p_{n-1})$ be the leading coefficient of p_{n-1}, let $\gamma_1, \gamma_2 \in \mathbb{P}^1(\bar{K})$ be the critical points of ϕ, and without loss say $\phi(\gamma_1) \neq \infty$. If $\phi(\gamma_2)$ is not (resp. is) ∞, then $[K_n : K_{n-1}] = 2^{2^{n-1}}$ if and only if*

$$p_{n-1}(\phi(\gamma_1))\,p_{n-1}(\phi(\gamma_2)) \notin K_{n-1}^{*2} \quad (\textit{resp. } \ell(p_{n-1})\,p_{n-1}(\phi(\gamma_1)) \notin K_{n-1}^{*2}). \tag{16}$$

Remark. As in Theorem 3.5, if both γ_1 and γ_2 are finite, then it follows from (4) and $n \geq 2$ that (16) may be replaced by

$$p_n(\gamma_1)\,p_n(\gamma_2) \notin K_{n-1}^{*2} \quad (\text{resp. } \ell(p_{n-1})\,p_n(\gamma_1) \notin K_{n-1}^{*2}). \tag{17}$$

Proof. From the discussion immediately preceding Lemma 3.6, we have

$$[K_n : K_{n-1}] = 2^{2^{n-1}}/\#V, \quad \text{where } V = \left\{(e_1, \ldots, e_{2^{n-1}}) \in \mathbb{F}_2^{2^{n-1}} : \prod_j \delta_j^{e_j} \in K_{n-1}^{*2}\right\}$$

has a natural structure of a $\mathbb{F}_2[G_{n-1}]$-module. Thus $[K_n : K_{n-1}] < 2^{2^{n-1}}$ if and only if $V \neq 0$, which by Lemma 3.6 occurs if and only if $V^{G_{n-1}} \neq 0$. However, since p_{n-1} is irreducible, G_{n-1} acts transitively on the δ_j defined in (15), implying that the only

possible nontrivial element in $V^{G_{n-1}}$ is $(1, 1, \ldots, 1)$. Hence $[K_n : K_{n-1}] < 2^{2^{n-1}}$ if and only if

$$\prod_{j=1}^{2^{n-1}} \mathrm{Disc}(p(x) - \beta_j q(x)) \in K_{n-1}^{*2},$$

where as before the β_j are the 2^{n-1} distinct roots of p_{n-1}. By Lemma 3.3, this is equivalent to

$$\prod_{i=1}^{2} \prod_{j=1}^{2^{n-1}} \big(\phi(\gamma_i) - \beta_j\big) \in K_{n-1}^{*2} \quad (\text{resp. } \prod_{j=1}^{2^{n-1}} \big(\phi(\gamma_1) - \beta_j\big) \in K_{n-1}^{*2}).$$

The theorem now follows from the fact that $p_{n-1}(x) = \ell(p_{n-1}) \prod (x - \beta_j)$. □

Let $s = 2$ if $\phi(\gamma_2) \neq \infty$ and $s = 1$ otherwise. Then equation (17) can be summarized as

$$\ell(p_{n-1})^s \prod_{i=1}^{s} p_n(\gamma_i) \notin K_{n-1}^{*2}.$$

Since $\ell(p_{n-1})^s \prod_{i=1}^{s} p_n(\gamma_i)$ is not just an element of K_{n-1} but also an element of K, it cannot be a square in K_{n-1} unless all the primes of K_{n-1} that divide it lie over primes of K that ramify in K_{n-1}. Thanks to Theorem 3.2, we have a handle on the primes that can ramify in K_{n-1}, and thus we can use Theorem 3.7 to give a simpler condition ensuring that $[K_n : K_{n-1}] = 2^{2^{n-1}}$. In the corollary, we limit ourselves for simplicity of statement to the case when ∞ is distinct from the critical points and values of ϕ, and does not have 0 in its forward orbit. Denote by $v_{\mathfrak{p}}$ the $\mathfrak{p}$-adic valuation at a prime $\mathfrak{p}$ in the ring of integers of K.

Corollary 3.8. *Let $\phi = p(x)/q(x) \in K(x)$ have degree 2, let $c = qp' - pq'$, let $\ell(p_{n-1})$ be the leading coefficient of p_{n-1}, and suppose that $\phi^n(\infty) \neq 0$ for all $n \geq 1$ and that ϕ has two finite critical points γ_1, γ_2 with $\phi(\gamma_i) \neq \infty$ for each i. Suppose further that there exists a prime $\mathfrak{p}$ of K with $v_{\mathfrak{p}}(p_n(\gamma_1)p_n(\gamma_2))$ odd and*

$$0 = v_{\mathfrak{p}}(\ell(p)) = v_{\mathfrak{p}}(\ell(c)) = v_{\mathfrak{p}}(\mathrm{Res}\,(q, p)) = v_{\mathfrak{p}}(\mathrm{Disc}\ p) = v_{\mathfrak{p}}(p_j(\gamma_i)) \qquad (18)$$

for $1 \leq i \leq 2$, $2 \leq j \leq n-1$. Then $[K_n : K_{n-1}] = 2^{2^{n-1}}$.

Proof. If $\phi(\infty) = \infty$, the conditions in (18), along with (11) and induction, imply $v_{\mathfrak{p}}(\mathrm{Disc}(p_{n-1})) = 0$. If $\phi(\infty) \neq \infty$, then the conditions in (18) and (10) give the same conclusion (note that ∞ not a critical point and $\phi(\infty) \neq 0$ imply that $k_1 = k_2 = 0$ in (10)). Hence $\mathfrak{p}$ does not ramify in K_{n-1}. Therefore there is a prime $\mathfrak{P}$ in the ring of integers of K_{n-1} with $v_{\mathfrak{P}}(\mathfrak{p})$ odd, and it follows that $v_{\mathfrak{P}}(p_n(\gamma_1)p_n(\gamma_2))$ is odd, so $p_n(\gamma_1)p_n(\gamma_2)$ cannot be a square in K_{n-1}^*. The corollary now follows from Theorem 3.7 and the remark preceding it. □

Corollary 3.8 provides a convenient method for checking the maximality of G_n for an arbitrary quadratic ϕ, at least for small n. In certain circumstances, it can even be used to determine G_∞, although the difficulties in disentangling possible interactions of the two critical orbits at various primes are considerable. We illustrate with the family of quadratic rational functions

$$\phi_a(x) = \frac{1 + ax + (3+a)x^2}{1 - (4+a)x - (a+1)x^2}, \quad a \in \mathbb{Q} \smallsetminus \{-2\}.$$

The critical points of ϕ_a are 1 and $-1/3$; in addition, ϕ_a has the two-cycle $1 \mapsto -1 \mapsto 1$, and ϕ_a sends 0 to 1. The behavior of one critical orbit is thus quite simple, and the fact that 0 is preperiodic ensures that elements of the other critical orbit are close to relatively prime. Define the polynomials p and q by $\phi = p/q$; then

$$\operatorname{Res}(q, p) = 16(a+2)^2.$$

For $k \geq 2$, we have the recursion

$$\begin{aligned} p_k &= q_{k-1}^2 + a q_{k-1} p_{k-1} + (3+a) p_{k-1}^2, \\ q_k &= q_{k-1}^2 - (4+a) q_{k-1} p_{k-1} - (a+1) p_{k-1}^2. \end{aligned} \tag{19}$$

Note that the only primes ℓ where we might have

$$p_k \equiv 0 \pmod{\ell} \quad \text{and} \quad q_k \equiv 0 \pmod{\ell}$$

are those dividing $2(a+2)$. The reason is that if $\ell \nmid \operatorname{Res}(q, p)$, then ϕ has *good reduction* at ℓ, and thus so do all iterates of ϕ [17], Theorem 2.18, implying that $\ell \nmid \operatorname{Res}(p_k, q_k)$. We remark that one can also apply Lemma 3.1 to the polynomials p_k, q_k, and $p_k q_k$ to obtain an exact formula for $\operatorname{Res}(p_k, q_k)$, which turns out to be a power of $\operatorname{Res}(p, q)$.

Now let $t \in \mathbb{Q}$, and suppose that for some $k \geq 1$, we have $p_k(t) = q_k(t)$. An induction shows that

$$\begin{aligned} p_{k+i}(t) &= (4+2a)(p_{k+i-1}(t))^2 \quad \text{if } i \text{ is odd, and} \\ p_{k+i}(t) &= 4(p_{k+i-1}(t))^2 \quad \text{if } i \text{ is even.} \end{aligned}$$

It follows that there are positive integers r_i, s_i with

$$p_{k+i}(t) = p_k(t)^{2^i} 2^{r_i} (2+a)^{s_i}, \tag{20}$$

where $r_i \equiv s_i \bmod 2$.

Suppose that $\ell \nmid 2(a+2)$ satisfies $\ell \mid p_k(t)$ for some $k \geq 1$, and take k minimal with this property. Then since $\ell \nmid q_k(t)$,

$$p_{k+1}(t) \equiv q_k(t)^2 \not\equiv 0 \pmod{\ell} \quad \text{and} \quad q_{k+1}(t) \equiv p_{k+1}(t) \pmod{\ell}.$$

It then follows from applying (20) with $k+1$ in place of k that $\ell \nmid p_{k+i}(t)$ for all $i \geq 1$.

Proof of Theorem 1.2. Consider the specialization $a = 0$, so

$$\phi(x) = \frac{1 + 3x^2}{1 - 4x - x^2}.$$

By the above analysis, any odd prime divides at most one term of the sequence $p_n(-1/3)$. Moreover, $p_1(1) = 4$ and $p_2(1) = q_2(1) = 2^6$, and thus it follows from (20) that $p_n(1)$ is an even power of 2 for all $n \geq 1$.

We also note that $(p_n(-1/3), q_n(-1/3)) \in (\mathbb{Z}/5\mathbb{Z})^2$ for $n \geq 1$ gives the orbit

$$(3,0) \mapsto (2,1) \mapsto (3,4) \mapsto (3,4) \mapsto \cdots$$

and thus neither of $\pm p_n(-1/3)$ is a square for all $n \geq 1$. Hence for each n there is a prime at which $p_n(-1/3)$ – and therefore $p_n(-1/3)p_n(1)$ – has odd valuation. To apply Corollary 3.8, we need to show that for each n, this prime is not 2 or 3, since the leading coefficient of c and Disc p are both -12. Note, however, that we don't need to consider $n = 1$, since clearly $[K_1 : K] = 2$.

Consider first the 3-adic behavior of $p_n(-1/3)$. We have $p(-1/3) = 4/3$ and $q(-1/3) = 20/9$. From (19), we see that for $k \geq 2$,

$$-1 \geq v_3(p_{k-1}(-1/3)) > v_3(q_{k-1}(-1/3)),$$

which implies that

$$v_3(p_k(-1/3)) = 2v_3(q_{k-1}(-1/3)) = v_3(q_k(-1/3)).$$

Hence $v_3(p_n(-1/3))$ is even for all $n \geq 2$.

Turning now to the 2-adic perspective, suppose that for some $k \geq 2$,

$$1 \leq e = v_2(p_{k-1}(-1/3)) = v_2(q_{k-1}(-1/3)),$$

and write

$$p_{k-1}(-1/3) = 2^e u \text{ and } q_{k-1}(-1/3) = 2^e w, \text{ where } v_2(u) = v_2(w) = 0.$$

We then have

$$p_k \equiv 2^{2e}(u^2 + 3v^2) \bmod 2^{2e+3} \equiv 2^{2e+2} \bmod 2^{2e+3},$$

and similarly for q_k. It follows that $v_2(p_k(-1/3)) = v_2(q_k(-1/3)) = 2e + 2$, and since $v_2(p(-1/3)) = v_2(q(-1/3)) = 2$, we thus have that $v_2(p_n(-1/3))$ is even for all $n \geq 1$.

Finally, we must show that $\phi_0^n(\infty) \neq 0$ for all $n \geq 1$. But $\phi_0(\infty) \equiv 0 \bmod 3$, implying that ∞ maps modulo 3 into the 2-cycle $1 \mapsto -1 \mapsto 1$. We have thus shown that when $a = 0$, $G_\infty \cong \mathrm{Aut}(T)$. □

4. Discriminants, irreducibility, and Galois theory of quadratic rational functions with an order-2 automorphism

In this section, we consider the setting of Conjecture 1.3, namely $\phi(x) = k(x^2+b)/x$. In the interest of describing exactly the arboreal Galois representation associated to such a map, we choose not to take $b = 1$, since doing so implies that conjugation by $x \to x/\sqrt{b}$ is defined over K, introducing a possible additional quadratic extension. Note that $\phi^n(\infty) \neq 0$ and $\phi^{n-1}(\infty) \neq 0$. Let $\iota(x) = -x$, and note that ι acts on the roots of p_n without fixed points, since 0 and ∞ are the only fixed points of ι and neither maps to 0 under any iterate of ϕ.

We wish to apply the same general program from Section 3 to this case. However, Theorem 3.5 and Theorem 3.7 do not apply, since the critical points satisfy $\gamma_1 = -\gamma_2$ and p_n is always an even function. Hence $p_n(\gamma_1)p_n(\gamma_2) = p_n(\gamma_1)^2$ is a square in K_{n-1}^* for all n. Indeed, we will show that $[K_n : K_{n-1}] \neq 2^{2^{n-1}}$ for all $n \geq 2$. As in Section 3, we have $G_n \hookrightarrow \mathrm{Aut}(T_n)$, and T_n is the complete binary rooted tree of height n, provided that p_n is separable. However, now the image of G_n must commute with the action of ι on T_n. We thus have $G_n \subseteq C_n$, where C_n denotes the centralizer in $\mathrm{Aut}(T_n)$ of the element corresponding to the action of ι. As in Section 1, $C_\infty := \varprojlim C_n$ plays roughly the role of a Cartan subgroup in the theory of Galois representations attached to elliptic curves with complex multiplication. We begin by describing the structure of C_n in purely group-theoretic terms; then we proceed to give discriminant, irreducibility, and Galois-maximality results for maps of the form $k(x^2 + b)/x$.

By slight abuse of notation, we write ι for the action induced by ι on T_n. Because ι acts on T_n without fixed points, its action on T_1 is non-trivial. Note that for any $j < n$ there is a natural epimorphism $C_n \to C_j$ obtained by restriction. For a vertex $v \in T_n$, we define the height of v to be $\min_i\{v \in T_i\}$.

Proposition 4.1. *Let $\iota \in \mathrm{Aut}(T_n)$ be any involution whose restriction to $\mathrm{Aut}(T_1)$ is non-trivial. Let C_j be the centralizer in $\mathrm{Aut}(T_j)$ of ι restricted to T_j, and let T_a be a subtree of T_n rooted at a height-one vertex of T_1. Then the map*

$$h\colon \ker(C_n \to C_1) \longrightarrow \mathrm{Aut}(T_a)$$

given by $h(\tau) = \tau|_{T_a}$ is an isomorphism.

Proof. Because there are exactly two branches from the root of T_n, there are exactly two subtrees of T_n rooted at a height-one vertex of T_1; call them T_a and T_b. The height-n vertices V of T_n may be decomposed into the union of the height-$(n-1)$ vertices $V_a \in T_a$ and $V_b \in T_b$. Because ι acts non-trivially on T_1 and is an automorphism of T_n, we have $\iota(T_a) = T_b$ and $\iota(T_b) = T_a$.

Clearly h is a homomorphism. To show h is surjective, let $\sigma \in \mathrm{Aut}(T_a)$ and define $\tau \in \mathrm{Aut}(T_n)$ by

$$\tau|_{T_1} = \mathrm{id}, \quad \tau|_{T_a} = \sigma, \quad \text{and} \quad \tau|_{T_b} = \iota\sigma\iota.$$

One then checks that $\iota\tau\iota = \tau$. Since τ acts trivially on T_1, $\tau \in \ker(C_n \to C_1)$ and $h(\tau) = \sigma$.

To show that h is injective, let $\tau \in \ker h$, so that $\tau(x) = x$ for all $x \in T_a$. Then since $\tau \in C_n$ we have $\tau(\iota(x)) = \iota(\tau(x)) = \iota(x)$; that is, τ acts trivially on all elements of T_b as well. Thus $\tau = \mathrm{id}$. □

In the next corollary, we describe the kernel of the restriction map $C_n \to C_{n-1}$. Recall that the kernel of the restriction map $\mathrm{Aut}(T_j) \to \mathrm{Aut}(T_{j-1})$ is generated by the transpositions swapping a single pair of vertices connected to a given vertex at level $j-1$, and thus is isomorphic to $(\mathbb{Z}/2\mathbb{Z})^{2^{j-1}}$. Recall also that the *Hausdorff dimension* of a subgroup H of $\mathrm{Aut}(T)$ is defined to be

$$\lim_{n\to\infty} \frac{\log_2 \# H_n}{\log_2 \# \mathrm{Aut}(T_n)},$$

where H_n is the restriction of the action of H to the tree T_n. This gives a rough measure of the size of H in $\mathrm{Aut}(T)$.

Corollary 4.2. *Assume the hypotheses of Proposition 4.1, and assume also that p_n is separable. Then there is an isomorphism between $\ker(C_n \to C_{n-1})$ and $\ker(\mathrm{Aut}(T_{n-1}) \to \mathrm{Aut}(T_{n-2}))$. In particular,*

$$\# \ker (C_n \to C_{n-1}) = 2^{2^{n-2}}$$

and the Hausdorff dimension of C_∞ is $1/2$.

Proof. Because we have assumed that p_n is separable, T_a is a complete binary rooted tree of height $n-1$, and we have $\mathrm{Aut}(T_a) \cong \mathrm{Aut}(T_{n-1})$. By Proposition 4.1 we then have a commutative diagram

$$\begin{array}{ccccccccc}
0 & \longrightarrow & \mathrm{Aut}(T_{n-1}) & \longrightarrow & C_n & \longrightarrow & C_1 & \longrightarrow & 0 \\
\Big\downarrow{\scriptstyle \mathrm{id}} & & \Big\downarrow{\scriptstyle r_1} & & \Big\downarrow{\scriptstyle r_2} & & \Big\downarrow{\scriptstyle \mathrm{id}} & & \\
0 & \longrightarrow & \mathrm{Aut}(T_{n-2}) & \longrightarrow & C_{n-1} & \longrightarrow & C_1 & \longrightarrow & 0,
\end{array}$$

where the rows are exact and the maps r_1 and r_2 are restriction. It is straightforward to check that this gives an exact sequence

$$\ker \mathrm{id} \to \ker r_1 \to \ker r_2 \to \ker \mathrm{id},$$

which completes the proof. The statement about Hausdorff dimension follows since

$$\# \ker (\mathrm{Aut}(T_n) \to \mathrm{Aut}(T_{n-1})) = 2^{2^{n-1}}.$$ □

The orbit of the critical point when $b = 1$ will play a special role in the sequel, so we introduce the following notation.

Definition 4.3. Let

$$\phi^*(x) = k(x^2 + 1)/x := p^*(x)/q^*(x),$$

and define p_n^* and q_n^* by the recursion in equation (4). Finally, let $\delta_n = kp_n^*(1)$.

We now turn to the discriminant of p_n. Note that one consequence of the following corollary is that $\mathrm{Disc}(p_n)$ is a square in $K_1 = K(\sqrt{-b})$ for all $n \geq 2$, so that the action of $\mathrm{Gal}\,(K_n/K_1)$ on the roots of p_n is contained in the alternating group on 2^n letters.

Corollary 4.4. *Let $k, b \in K^*$ and $\phi(x) = k(x^2 + b)/x$. Then for all $n \geq 2$, we have*

$$\mathrm{Disc}(p_n) = \pm k^{2^n(2^n-1)} b^{2^{2n-2}} \mathrm{Disc}(p_{n-1})^2 p_n^*(1)^2.$$

Proof. Since $\phi(\infty) = \infty$, we have $\phi^n(\infty) = \infty \neq 0$ for all n. We thus may apply (11) with the following data:

$$d = 2, \quad d_p = 2, \quad d_q = 1, \quad \text{and} \quad c = k(x^2 - b), \text{ which gives } d_c = 2.$$

We can then compute the exponents given in (12):

$$k_1 = 2^{2n-2} - 2^n \quad \text{and} \quad k_2 = 2^n.$$

We also have

$$\ell(p) = \ell(c) = k, \quad \ell(q) = 1, \quad p_n = k(p_{n-1}^2 + bq_{n-1}^2), \quad q_n = p_{n-1}q_{n-1}.$$

Since $p(x) = k(x^2 + b)$ and $q(x) = x$, an induction shows that p_n is even for all n. Thus $p_n(\sqrt{b}) = p_n(-\sqrt{b})$. A double induction on both q_n and p_n gives

$$q_n(\sqrt{b}) = b^{(2^n-1)/2} q_n^*(1) \quad \text{and} \quad p_n(\sqrt{b}) = b^{2^{n-1}} p_n^*(1). \tag{21}$$

Finally, $\mathrm{Res}\,(q, p) = \ell(q)^2 p(0) = kb$. The corollary now follows from substituting the relevant values into (11) and simplifying. □

Theorem 4.5. *Let $k, b \in K$, $\phi(x) = k(x^2 + b)/x$. Then p_n is irreducible if none of $-b$, $-b\delta_i$, δ_i is a square in K for $2 \leq i \leq n$.*

Remark. It is necessary to assume that both $-b\delta_i$ and δ_i are not squares in K. Indeed, in the case $k = 1, b = -5$, one has

$$-b\delta_2 = 25, \quad \delta_2 = -5, \quad \text{and} \quad p_2 = (x^2 - 5x + 5)(x^2 + 5x + 5).$$

In the case $k = 2/3, b = 1$ one has

$$\delta_2 = 100/81, \quad -b\delta_2 = -100/81, \quad \text{and} \quad p_2 = (2/27)(4x^2 + 1)(x^2 + 4).$$

Proof. We begin by considering p_1 and p_2. Clearly p_1 is irreducible if and only if $-b$ is not a square in K. From (4) we have

$$p_2 = k(p_1^2 + bq_1^2) = k(p_1 - x\sqrt{-b})(p_1 + x\sqrt{-b}).$$

Assuming that $-b$ is not a square in K, we have that p_2 is irreducible if and only if $p_1 - x\sqrt{-b}$ is irreducible over $K(\sqrt{-b})$, which holds if $\mathrm{Disc}(p_1 - x\sqrt{-b}) = -b(1 + 4k^2)$ is not a square in $K(\sqrt{-b})$. A straightforward computation shows this holds if and only if $1 + 4k^2$ is a square in K or $-b$ times a square in K. Since $1 + 4k^2 = k^{-1}p_2^*(1)$ and neither $-b\delta_2 = -bkp_2^*(1)$ nor $\delta_2 = kp_2^*(1)$ is a square in K, we conclude p_2 is irreducible.

Now induct on n. The cases $n = 1, 2$ have been handled, so let $n \geq 3$ and assume that p_{n-1} is irreducible. By Lemma 3.4, it is enough to show that for some root α of p_{n-1},

$$C\big(\phi(\sqrt{b}) - \alpha\big)\big(\phi(-\sqrt{b}) - \alpha\big) \notin K_{n-1}^{*2}.$$

We do this by taking the norm of the left-hand side over $K_1 = K(\sqrt{-b})$:

$$\begin{aligned} N_{K_{n-1}/K_1}\big(-C\big(\phi(\sqrt{b}) - \alpha\big)\big(\phi(\sqrt{b}) + \alpha\big)\big) \\ = \prod_{\phi^{n-2}(\alpha) = \sqrt{-b}} -C\big(\phi(\sqrt{b}) - \alpha\big)\big(\phi(\sqrt{b}) + \alpha\big). \end{aligned} \tag{22}$$

Since $\phi(-x) = -\phi(x)$, $\phi^{n-2}(\alpha) = \sqrt{-b}$ implies $\phi^{n-2}(-\alpha) = -\sqrt{-b}$. Thus

$$\big\{\pm\alpha : \phi^{n-2}(\alpha) = \sqrt{-b}\big\} = \big\{\alpha : \phi^{n-1}(\alpha) = 0\big\}.$$

Hence the right-hand side of (22) is the same as

$$(-C)^{(\deg p_{n-1})/2} \prod_{\phi^{n-1}(\alpha) = 0} \big(\phi(\sqrt{b}) - \alpha\big).$$

Because $\phi(\infty) = \infty$, we have $\phi^n(\infty) \neq 0$ for all n. Hence $\deg p_{n-1} = 2^{n-1}$, and $(\deg p_{n-1})/2$ is even when $n \geq 3$. Furthermore, since $\{\alpha : \phi^{n-1}(\alpha) = 0\}$ is the same as the set of roots of p_{n-1} and since $p_{n-1}(\alpha) = \ell(p_{n-1})\prod(x - \alpha)$, the left-hand side of (22) is not a square in K_1 provided

$$\ell(p_{n-1})^{-1} p_{n-1}\big(\phi(\sqrt{b})\big) \notin K_1^{*2}.$$

Finally, the recursion in (4) applied in this case gives $p_n(\sqrt{b}) = b^{2^{n-2}} p_{n-1}(\phi(\sqrt{b}))$. Inductive arguments show that

$$\ell(p_{n-1}) = k^{2^{n-1}} \quad \text{and} \quad p_n(\sqrt{b}) = b^{2^{n-1}} p_n^*(1),$$

meaning we must show

$$k^{-(2^n-1)}b^{2^{n-1}-2^{n-2}}p_n^*(1) \notin K_1^{*2}.$$

But by assumption neither $\delta_n = kp_n^*(1)$ nor $-b\delta_n = -bkp_n^*(1)$ is a square in K, and thus δ_n is not a square in K_1. (To see this, suppose that $c \in K$ with $c = (a_1 + a_2\sqrt{-b})^2$. Then $c = a_1^2 - ba_2^2$ with either $a_1 = 0$ or $a_2 = 0$, meaning either c or $-bc$ is a square in K.) This completes the main induction. □

Recall from Corollary 4.2 that $[K_n : K_{n-1}] \le 2^{2^{n-2}}$, with equality occurring if and only if $\ker(G_n \to G_{n-1}) \cong \ker(C_n \to C_{n-1})$. Using the methods of Section 3, we give a criterion ensuring that $[K_n : K_{n-1}]$ is as large as possible.

Theorem 4.6. *Let $k, b \in K^*$ and define $\phi(x) = k(x^2 + b)/x$. Assume that p_{n-1} is irreducible and $n \ge 3$. Then we have $[K_n : K_{n-1}] = 2^{2^{n-2}}$ provided that there exists a prime $\mathfrak{p}$ of K with*

$$v_\mathfrak{p}(\delta_n) \text{ odd}, \quad v_\mathfrak{p}(\delta_j) = 0 \text{ for } 1 \le j \le n-1, \quad \text{and} \quad v_\mathfrak{p}(k) = v_\mathfrak{p}(b) = v_\mathfrak{p}(2) = 0. \tag{23}$$

Remark. Because $\#C_1 = 2 = \deg p_1$ and $\#C_2 = 4 = \deg p_2$, it follows that $[K_1 : K] \le 2$ and $[K_2 : K_1] \le 2$. We have $[K_1 : K] = 2$ if and only if p_1 is irreducible and $[K_1 : K] = 4$ if and only if p_2 is irreducible. Note that p_1 is irreducible if and only if $-b$ is not a square in K, and from the proof of Theorem 4.5 we have that p_2 is irreducible if and only if $-b$, $-b\delta_2$ and δ_2 are all not squares in K.

Proof. As in the discussion preceding Lemma 3.6, K_n is obtained from K_{n-1} by adjoining the square roots of Disc $p(x) - \beta q(x)$, as β varies over the roots of p_{n-1}. In the present case, Disc $p(x) - \beta q(x) = \beta^2 - 4bk^2$. Since $-\beta$ is also a root of p_{n-1}, half of the square roots are redundant, and we have

$$K_n = K_{n-1}\big(\sqrt{\beta^2 - 4bk^2} : p_{n-2}(\beta) = \sqrt{-b}\,\big).$$

In analogy with the discussion preceding Lemma 3.6, we have $[K_n : K_{n-1}] = 2^{2^{n-2}}/\#V$, where

$$\begin{aligned} V = \{(e_1, \ldots, e_{2^{n-2}}) \in \mathbb{F}_2^{2^{n-2}} : \prod_j (\beta_j^2 - 4bk^2)^{e_j} \in K_{n-1}^{*2}\}, \\ \text{and } \beta_1, \ldots, \beta_j \text{ are the } 2^{n-2} \text{ solutions to } p_{n-2}(\beta) = \sqrt{-b}. \end{aligned} \tag{24}$$

The action of $G := \mathrm{Gal}\left(K_{n-1}/K(\sqrt{-b})\right)$ on the β_j gives an action of G on V as linear transformations, thereby making V a $\mathbb{F}_2[G]$-module. Lemma 3.6 now applies to show that if $\#V > 1$, then V contains a G-invariant element. Since p_{n-1} is irreducible, $\mathrm{Gal}\,(K_{n-1}/K)$ acts transitively on the β_j. By the definition of the

β_j in (24), any σ mapping one β_j to another must fix $\sqrt{-b}$ and thus must lie in G. Hence $[K_n : K_{n-1}] = 2^{2^{n-2}}$ provided that

$$\prod_j (\beta_j^2 - 4bk^2) \notin K_{n-1}^{*2}.$$

The hypotheses ensure that none of b, k, or $p_i^*(1)$ can be zero, which by Theorem 4.5 shows that p_{n-1} is irreducible. Because $n \geq 3$ and p_{n-1} is separable (since K is perfect), there are an even number of the β_j, and we may replace $\prod_j (\beta_j^2 - 4bk^2)$ with $\prod_j (4bk^2 - \beta_j^2)$. Further, the roots of p_{n-1} consist of $\{\pm\beta_1, \ldots, \pm\beta_j\}$, so we have that $[K_n : K_{n-1}] = 2^{2^{n-2}}$ provided that

$$\prod_{p_{n-1}(\beta)=0} (2k\sqrt{b} - \beta) \notin K_{n-1}^{*2}. \tag{25}$$

This product equals $\ell(p_{n-1})^{-1} p_{n-1}(\phi(\sqrt{b}))$, which via equation (4) is the same as $\ell(p_{n-1})^{-1} p_n(\sqrt{b})$ up to squares, since $n \geq 3$. Because $\ell(p_{n-1}) = k^{2^n - 1}$, we have that $\ell(p_{n-1})^{-1} p_n(\sqrt{b})$ is a square in K_{n-1} if and only if $k p_n(\sqrt{b})$ is a square in K_{n-1}. As in equation (21), $p_n(\sqrt{b}) = b^{2^{n-1}} p_n^*(1)$, so (25) holds provided that $\delta_n = k p_n^*(1) \notin K_{n-1}^{*2}$.

By assumption in (23), there is a prime $\mathfrak{p}$ of K with $v_\mathfrak{p}(k) = v_\mathfrak{p}(b) = v_\mathfrak{p}(2) = 0$. We thus have

$$v_\mathfrak{p}(\text{Disc } p_1) = v_\mathfrak{p}(4bk^2) = 0.$$

Also from (23), we assume $v_\mathfrak{p}(\delta_j) = v_\mathfrak{p}(k p_j^*(1)) = 0$ for $1 \leq j \leq n-1$, and since $v_\mathfrak{p}(k) = 0$, we have

$$v_\mathfrak{p}(p_j^*(1)) = 0 \quad \text{for } 1 \leq j \leq n-1.$$

By induction, Corollary 4.4 implies that $v_\mathfrak{p}(\text{Disc } p_{n-1}) = 0$. Therefore $\mathfrak{p}$ does not ramify in K_{n-1}, whence there is a prime $\mathfrak{P}$ of K_{n-1} with $v_\mathfrak{P}(\mathfrak{p})$ odd. We then have $v_\mathfrak{P}(\delta_n)$ odd, which means that δ_n cannot be a square in K_{n-1}. □

By Theorem 4.5, to show that p_n is irreducible for all $n \geq 1$, it suffices to show that none of $-b$, $-b\delta_n$, or δ_n is a square in K for all $n \geq 2$. By Theorem 4.6, a relatively small amount of knowledge about the primes dividing the δ_n then allows one to show $G_n \cong C_n$. We also note that given Proposition 2.1, the role played by b is actually a minor one because it is simply a twist parameter.

In the next section, we prove Theorem 5.3, which implies Theorem 1.8. We then give several sufficient conditions to show $-b\delta_n$ and δ_n are not squares in K (Theorems 5.7–5.10), before giving a further sufficient condition on k that ensures that $G_n \cong C_n$ (Theorem 5.13).

5. Maximality and finite index results

In this section we apply the results of Section 4 to obtain results showing G_∞ is a large subgroup of C_∞ in many cases. The map $\phi(x) = k(x^2 + b)/x \in K(x)$ is given in homogeneous coordinates by

$$\phi([X, Y]) = [k(X^2 + bY^2), XY], \quad k, b \in K. \tag{26}$$

Recall that

$$P_n(X, Y) = \begin{cases} k\left(X^2 + bY^2\right) & \text{if } n = 1, \\ k\left(P_{n-1}(X, Y)^2 + bQ_{n-1}(X, Y)^2\right) & \text{if } n \geq 2, \end{cases}$$

and

$$Q_n(X, Y) = \begin{cases} XY & \text{if } n = 1, \\ P_{n-1}(X, Y)Q_{n-1}(X, Y) & \text{if } n \geq 2. \end{cases}$$

Thus $\phi^n([X, Y]) = [P_n(X, Y), Q_n(X, Y)]$. Recall also that

$$\begin{aligned} p_n(x) &= P_n(x, 1) & q_n(x) &= Q_n(x, 1), \\ P_n^*(X, Y) &= P_n(X, Y)|_{b=1}, & Q_n^*(X, Y) &= Q_n(X, Y)|_{b=1}, \\ p_n^*(x) &= P_n^*(x, 1), & q_n^*(x) &= Q_n^*(x, 1), \\ \phi^*(x) &= k(x^2 + 1)/x, \text{ and} & \delta_n &= kp_n^*(1). \end{aligned}$$

Many of our results in this section exclude the case where ϕ is post-critically finite, and so we begin by showing that this case is rare.

Proposition 5.1. *Let K be a number field and $\phi(x) = k(x^2 + b)/x$ with $k \in K^*$. If ϕ is post-critically finite, then the standard (absolute) multiplicative height of k is at most 2. In particular, there are only finitely many post-critically finite ϕ over any number field.*

Remark. This is best possible, since $k = \pm 1/2$ give post-critically finite maps. Using Proposition 5.1, one easily checks that these are the only $k \in \mathbb{Q}$ that give post-critically finite maps.

Proof. Begin by noting that ϕ is conjugate to $k(x^2 + 1)/x$ over $\bar{K}$, and a map is post-critically finite if and only if all its conjugates are. Therefore we may consider $\phi^*(x) = k(x^2 + 1)/x$. The critical points of ϕ^* are ± 1, and their orbits are interchanged by the involution $z \mapsto -z$. So ϕ^* is post-critically finite if and only if the orbit of $z = 1$ is finite.

Recall that the standard (absolute) multiplicative height of $k \in K$ is defined to be

$$\left(\prod_{v \in M_K} \max\{1, |k|_v\}^{n_v} \right)^{1/[K:\mathbb{Q}]},$$

where M_K is the set of places of K and $n_v = [K_v : \mathbb{Q}_v]$ is the local degree of v. Consider first an archimedean place v of K, and for simplicity denote $|\cdot|_v$ by $|\cdot|$. Suppose $|k| > 1$. One checks that ∞ is a fixed point of K with multiplier $1/k$, and hence is attracting. By Theorem 9.3.1 of [1], ∞ must attract a critical point monotonically (i.e. the critical point does not land on ∞), proving that ϕ is not post-critically finite.

Now let v be a non-archimedean place of K. If $|k| > 1$, then for $|x| > 1$ we get

$$|\phi(x)| = |k| \cdot \left|\frac{x^2+1}{x}\right| = |k| \cdot \frac{|x^2|}{|x|} > |x|.$$

It follows that x cannot have a finite orbit under ϕ. Now $|\phi(1)| = |2k|$, and this equals $|k|$ provided that v does not lie over 2. If v does divide 2, then we still have $|2k| > 1$ provided that $|k| > |1/2| = 2$. Hence if ϕ is post-critically finite, we must have $|k|_v \leq 1$ for each place $v \in M_K$ not over 2, and $|k|_v \leq 2$ for each place $v \in M_K$ over 2.

Suppose that ϕ is post-critically finite, and assume $[K : \mathbb{Q}] = d$. We have

$$\prod_{v \in M_K} \max\{1, |k|_v\}^{n_v} \leq \prod_{v|2} 2^{n_v} = 2^d.$$

Taking the d^{th} root gives the desired height bound of 2. □

Lemma 5.2. *Suppose that for some prime* $\mathfrak{p}$ *of* $\mathcal{O}_K$, $v_\mathfrak{p}(\delta_n) > 0$ *and* $v_\mathfrak{p}(\delta_m) > 0$ *for some* $m \neq n$. *Then* $v_\mathfrak{p}(k) > 0$.

Proof. Without loss of generality, we may assume $n < m$ and that n is the smallest positive integer satisfying $v_\mathfrak{p}(\delta_n) > 0$. The rough idea is that ϕ maps 0 to ∞, which is a fixed point. Thus if $\phi^n(1) \equiv 0 \bmod \mathfrak{p}$, then $\phi^m(1) \not\equiv 0 \bmod \mathfrak{p}$, and $\mathfrak{p}$ cannot divide both $p_n(1)$ and $p_m(1)$.

We are given that $v_\mathfrak{p}(\delta_n) = v_\mathfrak{p}(kp_n^*(1)) > 0$. From the recursion, we see that $p_n^*(1)$ is a polynomial in k with integral coefficients and no constant term; hence $v_\mathfrak{p}(k) < 0$ implies $v_\mathfrak{p}(p_n^*(1)) < 0$. It follows that either $v_\mathfrak{p}(k) > 0$ and we're done or $v_\mathfrak{p}(k) = 0$ and $v_\mathfrak{p}(p_n^*(1)) > 0$. We assume the latter scenario and derive a contradiction.

We have $v_\mathfrak{p}(P_n^*(1,1)) > 0$, and thus $Q_{n+1}^*(1,1) = P_n^*(1,1)Q_n^*(1,1) \equiv 0 \bmod \mathfrak{p}$ and

$$P_{n+1}^*(1,1) = k\left(P_n^*(1,1)^2 + Q_n^*(1,1)^2\right) \equiv kQ_n^*(1,1)^2 \mod \mathfrak{p}.$$

Induction gives $Q_m^*(1,1) \equiv 0 \bmod \mathfrak{p}$ and

$$P_m^*(1,1) \equiv k^{2^{m-n}-1} Q_n^*(1,1)^{2^{m-n}} \mod \mathfrak{p}.$$

Now $Q_n^*(1,1) = P_1^*(1,1)P_2^*(1,1)\cdots P_{n-1}^*(1,1)$, so by the minimality of n we have $Q_n^*(1,1) \not\equiv 0 \bmod \mathfrak{p}$. By assumption we have $k \not\equiv 0 \bmod \mathfrak{p}$, and so it follows that

$p_m^*(1) = P_m^*(1,1) \not\equiv 0 \bmod \mathfrak{p}$. This contradicts our supposition that $v_\mathfrak{p}(\delta_m) > 0$. □

Theorem 5.3. *Let ϕ be defined as in (26). Suppose that none of $-b$, δ_n, and $-b\delta_n$ is a square in K for $n \geq 2$, and also assume that ϕ is not post-critically finite. Then G_∞ has finite index in C_∞.*

Remark. From the proof below, it follows that only finitely many of the numbers δ_n and $-b\delta_n$ can be squares. However, this is not enough to ensure that $p_n(x)$ is irreducible for all n, and failure of this irreducibility provides an obstacle to showing $[C_\infty : G_\infty] < \infty$.

Proof. We first claim that for any $c \in K^*$, $c\delta_n$ is a square in K for at most finitely many n. Note that $c\delta_n = ckp_n^*(1)$. Let $\phi^*(x) = k(x^2+1)/x$ and apply equation (5) to see that $ckp_n^*(1)$ is a square if and only if

$$ck \cdot p^*\left(\phi^{*n-1}(1)\right) \in K^{*2}.$$

Without loss of generality, we may take $n \geq 4$, and thus rewrite

$$ck \cdot p^*\left(\phi^{*n-1}(1)\right) = ck \cdot p^*\left(\phi^{*2} \circ \phi^{*n-3}(1)\right).$$

It follows that $c\delta_n$ is a square if and only if the curve $ck \cdot p^*(\phi^{*2}(x)) = y^2$ has a K-rational point (x, y) with $x = (\phi^*)^{n-3}(1)$. Since

$$p^*(\phi^{*2}(x)) = k\left(p_2^*(x)^2 + q_2^*(x)^2\right)/q_2^*(x)^2,$$

this is equivalent to the curve

$$C : y^2 = c\left(p_2^*(x)^2 + q_2^*(x)^2\right) \tag{27}$$

having a rational point with $x = (\phi^*)^{n-3}(1)$. The right-hand side of (27) is simply $ck^{-1}p_3^*(x)$, whose discriminant, by Corollary 4.4, is divisible only by c, k, b, $p_1^*(1)$, $p_2^*(1)$, and $p_3^*(1)$. None of these is zero because none of δ_1, δ_2, and δ_3 is a square in K by hypothesis. Thus the right-hand side of (27) has distinct roots, and hence the genus of C is 3. It follows from Faltings' Theorem, see [5], Part E, that C has only finitely many rational points, and thus $c\delta_n$ is a square for only finitely many n.

To prove the theorem, note that the hypotheses on $-b$, $-b\delta_n$ and δ_n allow us to apply Theorem 4.5 to show that $p_n(x)$ is irreducible for all n. We now wish to apply Theorem 4.6. Let S be a finite set of places of K, including all places dividing k, b, or 2, and all archimedean places. Expand S further, if necessary, so that the ring $\mathcal{O}_{K,S}$ of S-integers is a principal ideal domain. Let $U_{K,S}$ denote the multiplicative

group of S-units. Note that since $\delta_n \in \mathbb{Z}[k^2]$, there exists $a_n \in K$ with $a_n^2\delta_n \in \mathcal{O}_K$. Since $\mathcal{O}_{K,S}$ is a UFD, we may write for each n,

$$a_n^2\delta_n = u\beta^2 \prod_{i=1}^{j_n} \pi_i, \tag{28}$$

with $u \in U_{K,S}/U_{K,S}^2$, $\beta \in \mathcal{O}_{K,S}$, and $\pi_i \in \mathcal{O}_{K,S}$ irreducible, and this decomposition is unique. We permit the product on the right of (28) to be empty, and say $j_n = 0$ in this case.

By Dirichlet's Theorem for S-units ([4], p. 174), $U_{K,S}/U_{K,S}^2$ is a finite group, and we let Σ consist of a set of coset representatives. Suppose that there are infinitely many n for which $j_n = 0$. Because Σ is finite, there is a product c of elements in Σ with $ca_n^2\delta_n$ a square in $\mathcal{O}_{K,S}$ for infinitely many n. This contradicts the conclusion of the previous paragraph.

It follows that for all but finitely many n, there must be at least one π_i in (28). Thus setting $\mathfrak{p}_i = (\pi_i) \cap \mathcal{O}_K$ we have

$$v_{\mathfrak{p}_i}(\delta_n) \text{ odd}, \quad \text{and} \quad v_{\mathfrak{p}_i}(k) = v_{\mathfrak{p}_i}(b) = v_{\mathfrak{p}_i}(2) = 0.$$

Because $v_{\mathfrak{p}_i}(k) = 0$, Lemma 5.2 implies that $\mathfrak{p}_i$ divides at most one δ_n. Theorem 4.6 then applies to complete the proof. □

In light of Theorem 5.3, we now study the quantities $-b$, δ_n, and $-b\delta_n$. We begin with a fundamental result on the polynomials $P_n(X, Y)$ and $Q_n(X, Y)$.

Lemma 5.4. *Let $S_n, T_n \in \mathbb{Z}[k, X, Y]$ be the polynomials not divisible by k that satisfy $P_n = k^{s(n)}S_n$ and $Q_n = k^{t(n)}T_n$ respectively, for some $s(n), t(n) \in \mathbb{Z}$. Then we have*

$$S_n = \begin{cases} S_{n-1}^2 + bT_{n-1}^2 & \text{if } n \text{ is odd,} \\ k^2S_{n-1}^2 + bT_{n-1}^2 & \text{if } n \text{ is even,} \end{cases}$$

$$T_n = S_{n-1}T_{n-1},$$

$$s(n) = \frac{1}{3}\left(2^n - (-1)^n\right),$$

and

$$t(n) = \begin{cases} s(n) - 1 & \text{if } n \text{ is odd,} \\ s(n) & \text{if } n \text{ is even.} \end{cases}$$

Moreover, for any n, S_n and T_n are homogeneous in X and Y, and relatively prime as polynomials in X and Y with coefficients in $\mathbb{Z}[k]$.

Proof. We proceed by induction. In this case it will be convenient for us to start with $n = 0$, in which case we put $\phi^0([X, Y]) = [X, Y]$. This gives

$$S_0 = X, \quad T_0 = Y, \quad \text{and} \quad s(0) = t(0) = 0.$$

For $n = 1$, we have $\phi([X, Y]) = [k(X^2 + bY^2), XY]$, showing that

$$S_1 = S_0^2 + bT_0^2, \quad T_1 = S_0T_0, \quad s(1) = 1, \quad \text{and} \quad t(1) = 0,$$

which agree with the statements in the lemma.

For $n = 2$, we have

$$P_2(X, Y) = k\left(k^2(X^2 + bY^2)^2 + b(XY)^2\right)$$

and

$$Q_2(X, Y) = k(X^2 + bY^2)XY,$$

so that

$$S_2 = k^2S_1^2 + bT_1^2, \quad T_2 = S_1T_1, \quad \text{and} \quad s(2) = t(2) = 1,$$

which again agree with the statements in the lemma.

Now suppose that n is even and that the statement of the lemma holds for $n - 1$, so in particular $t(n - 1) = s(n - 1) - 1$. Then

$$\begin{aligned} P_n = k\left(P_{n-1}^2 + bQ_{n-1}^2\right) &= k\left(k^{2s(n-1)}S_{n-1}^2 + bk^{2t(n-1)}T_{n-1}^2\right) \\ &= k \cdot k^{2s(n-1)-2}\left(k^2S_{n-1}^2 + bT_{n-1}^2\right). \end{aligned}$$

Since k does not divide T_{n-1}, it also does not divide $k^2S_{n-1}^2 + bT_{n-1}^2$, showing that $S_n = k^2S_{n-1}^2 + bT_{n-1}^2$. Thus

$$\begin{aligned} s(n) = 2s(n - 1) - 1 &= \frac{2}{3}\left(2^{n-1} - (-1)^{n-1}\right) - 1 && \text{(by induction)} \\ &= \frac{1}{3}\left(2^n + 2 - 3\right) && \text{(since } n - 1 \text{ is odd)} \\ &= \frac{1}{3}\left(2^n - (-1)^n\right). \end{aligned}$$

We also have

$$Q_n = P_{n-1}Q_{n-1} = \left(k^{s(n-1)}S_{n-1}\right)\left(k^{t(n-1)}T_{n-1}\right) = k^{2s(n-1)-1}S_{n-1}T_{n-1}.$$

Since k divides neither S_{n-1} nor T_{n-1}, it also does not divide $S_{n-1}T_{n-1}$, showing that $T_n = S_{n-1}T_{n-1}$. Thus $t(n) = 2s(n - 1) - 1 = s(n)$.

Now suppose that n is odd and that the statement of the lemma holds for $n-1$, so in particular $t(n-1) = s(n-1)$. Then

$$\begin{aligned} P_n = k\left(P_{n-1}^2 + bQ_{n-1}^2\right) &= k\left(k^{2s(n-1)}S_{n-1}^2 + bk^{2t(n-1)}T_{n-1}^2\right) \\ &= k \cdot k^{2s(n-1)}\left(S_{n-1}^2 + bT_{n-1}^2\right). \end{aligned}$$

As in the case of n even, we have $S_n = k^2 S_{n-1}^2 + bT_{n-1}^2$. Thus

$$\begin{aligned} s(n) = 2s(n-1) + 1 &= \frac{2}{3}\left(2^{n-1} - (-1)^{n-1}\right) + 1 \quad \text{(by induction)} \\ &= \frac{1}{3}\left(2^n - 2 + 3\right) \quad \text{(since } n-1 \text{ is even)} \\ &= \frac{1}{3}\left(2^n - (-1)^n\right). \end{aligned}$$

We also have

$$Q_n = P_{n-1}Q_{n-1} = \left(k^{s(n-1)}S_{n-1}\right)\left(k^{t(n-1)}T_{n-1}\right) = k^{2s(n-1)}S_{n-1}T_{n-1}.$$

As in the case of n even, it follows that $T_n = S_{n-1}T_{n-1}$. Thus $t(n) = 2s(n-1) = s(n) - 1$.

It remains to show that S_n and T_n are relatively prime as homogeneous polynomials in X and Y with coefficients in $\mathbb{Z}[k]$. Assume inductively the same statements hold for S_{n-1} and T_{n-1}. The homogeneity of S_n and T_n follows immediately from the recursions in the lemma, which have already been established. Let F be an irreducible non-constant homogeneous polynomial in X and Y with coefficients in $\mathbb{Z}[k]$. If F divides T_n, then F must divide either S_{n-1} or T_{n-1}, but cannot divide both since S_{n-1} and T_{n-1} are relatively prime. From the formula for S_n in the lemma it follows that F cannot divide S_n, regardless of the parity of n. □

For the remainder of this section we assume that $b = 1$, as in Conjecture 1.3. We thus have $\phi^* = \phi$, $p_n^* = p_n$, and $q_n^* = q_n$. We put Lemma 5.4 to use to study the δ_n, which will allow us to apply Theorem 5.3 in the case where K is real, since $\delta_n > 0$ and thus $-\delta_n$ cannot be a square. Before proceeding, we note that

$$\delta_n = kP_n(1,1) = k^{s(n)+1}S_n(1,1)$$

by Lemma 5.4. Since $s(n)+1$ is always even, we have that δ_n is a square in K if and only if $S_n(1,1)$ is a square in K. To make this a bit more concrete, here are the first few $P_n(1,1)$ and $Q_n(1,1)$, with corresponding $S_n(1,1)$ and $T_n(1,1)$ easy to read off.

$$\begin{aligned} &P_1(1,1) = 2k, \quad P_2(1,1) = k(4k^2+1), \quad P_3(1,1) = k^3(16k^4 + 8k^2 + 5), \\ &Q_1(1,1) = 1, \quad Q_2(1,1) = 2k, \quad Q_3(1,1) = k^2(8k^2+2), \end{aligned}$$

$$\begin{aligned} P_4(1,1) &= k^5(256k^{10} + 256k^8 + 224k^6 + 144k^4 + 57k^2 + 4), \text{ and} \\ Q_4(1,1) &= k^5(128k^6 + 96k^4 + 56k^2 + 10). \end{aligned}$$

Lemma 5.5. *Suppose that $b = 1$ and there is a prime $\mathfrak{p}$ of $\mathcal{O}_K$ with $\mathfrak{p} \mid (5)$ and $[\mathcal{O}_K/\mathfrak{p} : \mathbb{Z}/5\mathbb{Z}]$ odd. If $k \equiv \pm 2 \bmod \mathfrak{p}$, then neither of $\pm\delta_n$ is a square for any n.*

Proof. From Lemma 5.4 and the fact that $\mathcal{O}_K/\mathfrak{p}$ has characteristic 5, we have that the sequence $(S_n(1,1), T_n(1,1))$ modulo $\mathfrak{p}$ is $(2,1), (2,2), (3,4), (2,2), (3,4), \ldots$ and repeats in the obvious way. Because $[\mathcal{O}_K/\mathfrak{p} : \mathbb{Z}/5\mathbb{Z}]$ is odd, $\mathcal{O}_K/\mathfrak{p}$ has no quadratic sub-extensions, and thus neither of ± 2 is a square in $\mathcal{O}_K/\mathfrak{p}$. Hence $\pm S_n(1,1)$ is not a square modulo $\mathfrak{p}$ for all n. □

Corollary 5.6. *Suppose that K is a number field of odd degree, and take $b = 1$. Then there is a congruence class of $k \in K$ with $[C_\infty : G_\infty]$ finite.*

Proof. Because K has odd degree, it has no quadratic sub-extenstions, and hence -1 cannot be a square in K. Moreover, the product of the residue class degrees of ideals of $\mathcal{O}_K$ dividing (5) must be odd, and thus there is some $\mathfrak{p} \mid (5)$ with $[\mathcal{O}_K/\mathfrak{p} : \mathbb{Z}/5\mathbb{Z}]$ odd. By Lemma 5.5, when $k \equiv \pm 2 \bmod \mathfrak{p}$, neither $\pm\delta_n$ is a square for any n. By Theorem 5.3, $[C_\infty : G_\infty]$ is finite for all $k \equiv \pm 2 \bmod \mathfrak{p}$. □

We now present several results that show δ_n is not a square for all n provided that there exist certain primes of $\mathcal{O}_K$ and k satisfies conditions relating to these primes. These lead into Corollary 5.11, which shows that Conjecture 1.3 is true for certain real number fields K, including $\mathbb{Q}$.

Theorem 5.7. *Suppose that $b = 1$ and $v_\mathfrak{p}(k) = 0$ for some prime $\mathfrak{p} \subset \mathcal{O}_K$ with $\#\mathcal{O}_K/\mathfrak{p} = 2$. Then δ_n is not a square for all $n \geq 2$.*

Proof. Note first that $\#\mathcal{O}_K/\mathfrak{p} = 2$ implies $\mathfrak{p} \mid (2)$, and let $e \geq 1$ be such that $\mathfrak{p}^e \| (2)$. We claim that $\mathfrak{p} \nmid (P_n(1,1))$ and $\mathfrak{p}^e \| (Q_n(1,1))$ for all $n \geq 2$. We have

$$(P_1(1,1)) = (2k) \text{ and } (Q_1(1,1)) = (1), \text{ whence } (P_2(1,1)) = \left(k(4k^2+1)\right),$$

which is not divisible by $\mathfrak{p}$ because $v_\mathfrak{p}(k) = 0$ and $\mathfrak{p} \mid (2)$. Also, $(Q_2(1,1)) = (2k)$, which is exactly divisible by $\mathfrak{p}^e$. Since

$$P_n(1,1) = k(P_{n-1}(1,1)^2 + Q_{n-1}(1,1)^2) \text{ and } Q_n(1,1) = P_{n-1}(1,1)Q_{n-1}(1,1),$$

the claim follows by induction.

Now suppose that $\delta_n \in K^2$ for some $n \geq 2$. Then $kP_n(1,1) \in K^2$, and so

$$P_{n-1}(1,1)^2 + Q_{n-1}(1,1)^2 = z^2 \quad \text{for some } z \in K.$$

Rewrite this as

$$Q_{n-1}(1,1)^2 = (z + P_{n-1}(1,1))(z - P_{n-1}(1,1)),$$

and to ease notation set

$$z + P_{n-1}(1,1) = s \quad \text{and} \quad z - P_{n-1}(1,1) = t.$$

This gives

$$Q_{n-1}(1,1)^2 = st \quad \text{and} \quad 2P_{n-1}(1,1) = s - t.$$

From the previous paragraph, $v_\mathfrak{p}(Q_{n-1}(1,1)) = e = v_\mathfrak{p}(2)$ and $v_\mathfrak{p}(P_{n-1}(1,1)) = 0$, giving

$$2e = v_\mathfrak{p}(s) + v_\mathfrak{p}(t) \tag{29}$$

and

$$e = v_\mathfrak{p}(s - t); \tag{30}$$

hence

$$2v_\mathfrak{p}(s - t) = v_\mathfrak{p}(s) + v_\mathfrak{p}(t). \tag{31}$$

If $v_\mathfrak{p}(s) \neq v_\mathfrak{p}(t)$, then $2v_\mathfrak{p}(s - t) = 2\max\{v_\mathfrak{p}(s), v_\mathfrak{p}(t)\} > v_\mathfrak{p}(s) + v_\mathfrak{p}(t)$, contradicting (31). If $v_\mathfrak{p}(s) = v_\mathfrak{p}(t)$, they are both e by equation (29). So $s, t \in \mathfrak{p}^e$ and $s, t \notin \mathfrak{p}^{e+1}$, so neither is the identity in $\mathfrak{p}^e/\mathfrak{p}^{e+1}$. But $\mathfrak{p}^e/\mathfrak{p}^{e+1} \cong O_K/\mathfrak{p}$ (see e.g. [4], p. 43) and thus has only two elements, implying that $s - t \in \mathfrak{p}^{e+1}$. This contradicts (30), proving the theorem. □

Theorem 5.8. *Suppose that $b = 1$ and there is $\mathfrak{p} \subset \mathcal{O}_K$ with $\#\mathcal{O}_K/\mathfrak{p} = 3$. Then δ_n is not a square for all odd n. If $v_\mathfrak{p}(k) = 0$ then δ_n is not a square for all $n \geq 1$.*

Proof. Note that $\mathcal{O}_K/\mathfrak{p} \cong \mathbb{Z}/3\mathbb{Z}$. Because $S_1(1,1) = 2$ and $T_1(1,1) = 1$, and the sum of the squares of two non-zero elements of $\mathbb{Z}/3\mathbb{Z}$ cannot be zero, it follows by induction that $S_n(1,1)$ and $T_n(1,1)$ are not zero in $\mathcal{O}_K/\mathfrak{p}$. It then follows immediately from the recurrence for S_n in Lemma 5.4 that $S_n(1,1) \equiv 2 \bmod \mathfrak{p}$ for n odd, and thus cannot be a square in K for n odd. If $v_\mathfrak{p}(k) = 0$, then the same statement holds for n even. □

Theorem 5.9. *Let $b = 1$. If one of the following holds then δ_n is not a square for all $n \geq 1$:*

(1) *There is $\mathfrak{p} \subset \mathcal{O}_K$ with $\mathcal{O}_K/\mathfrak{p} \cong \mathbb{Z}/7\mathbb{Z}$ and $k \equiv 2, 5 \bmod \mathfrak{p}$.*

(2) *There is $\mathfrak{p} \subset \mathcal{O}_K$ with $\mathcal{O}_K/\mathfrak{p} \cong \mathbb{Z}/7\mathbb{Z}$, $k \equiv 1, 6 \bmod \mathfrak{p}$, and the hypotheses of Theorem 5.8 hold.*

Proof. For $k \equiv \pm 2 \bmod \mathfrak{p}$, the sequence $(S_n(1,1), T_n(1,1))$ modulo $\mathfrak{p}$ is $(2,1), (3,2)$ and then a repeating cycle of $(6,6), (5,1), (5,5), (6,4), (3,3), (3,2)$, so $S_n(1,1)$ is never a square modulo $\mathfrak{p}$. For $k \equiv \pm 1 \bmod \mathfrak{p}$, the sequence in question consists of the length-12 repeating cycle

$$(2,1),\ (5,2),\ (1,3),\ (3,3),\ (4,2),\ (6,1),\ (2,6),\ (5,5),\ (1,4),\ (3,4),\ (4,5),\ (6,6).$$

Hence $S_n(1,1)$ is not a square modulo $\mathfrak{p}$ for all even n, and combined with Theorem 5.8 this shows that δ_n is not a square for all $n \geq 1$. □

Theorem 5.10. *Let $b = 1$. Suppose that one of the following holds:*

(1) $v_{\mathfrak{p}}(2k-1) > 0$ *or* $v_{\mathfrak{p}}(2k+1) > 0$ *for a prime* $\mathfrak{p} \subset \mathcal{O}_K$ *such that 2 is not a square in* $\mathcal{O}_K/\mathfrak{p}$.

(2) $v_{\mathfrak{p}}(2k^2-k+1) > 0$ *for a prime* $\mathfrak{p} \subset \mathcal{O}_K$ *such that* $-k$ *is not a square in* $\mathcal{O}_K/\mathfrak{p}$.

(3) $v_{\mathfrak{p}}(2k^2+k+1) > 0$ *for a prime* $\mathfrak{p} \subset \mathcal{O}_K$ *such that* k *is not a square in* $\mathcal{O}_K/\mathfrak{p}$.

Then δ_n is not a square for all n.

Remark. When $K = \mathbb{Q}$, case (1) of Theorem 5.10 applies provided that there is a prime p with $v_p(2k \pm 1) > 0$ and p congruent to 3 or 5 modulo 8. In particular, if k is an integer and $2k \not\equiv 0 \bmod 8$, then one of $2k \pm 1$ must be equivalent to 3 or 5 modulo 8, implying that some divisor of $2k \pm 1$ is equivalent to 3 or 5 modulo 8. Hence if k is an integer not divisible by 4 then δ_n is not a square for all n.

Proof. It follows from the definitions of $P_n(X,Y)$ and $Q_n(X,Y)$ that if $\epsilon_n = kQ_n(1,1)$, then

$$\delta_n = \delta_{n-1}^2 + \epsilon_{n-1}^2, \quad \epsilon_n = (1/k)\delta_{n-1}\epsilon_{n-1}. \tag{32}$$

Suppose first that we are in case (1). One checks that $\delta_2 - \delta_1$, $\delta_3 - \delta_2$, and $\epsilon_3 - \epsilon_1$ are all divisible by both $2k-1$ and $2k+1$. Hence for $\mathfrak{p} \mid (2k-1)$ or $\mathfrak{p} \mid (2k+1)$, we have $\delta_3 \equiv \delta_2 \equiv \delta_1 \bmod \mathfrak{p}$ and $\epsilon_3 \equiv \epsilon_1 \bmod \mathfrak{p}$, which by (32) and induction ensures that $\delta_n \equiv \delta_1 \bmod \mathfrak{p}$ for all n. But $\delta_1 = 2k^2$, which is not a square modulo $\mathfrak{p}$ by assumption.

For the other cases, one checks that $2k^2 \pm k + 1$ divides $\delta_3 - \delta_2$, $\delta_4 - \delta_3$, and $\epsilon_4 - \epsilon_2$. As in the previous paragraph, this implies that $\delta_n \equiv \delta_2 \bmod \mathfrak{p}$ for all $n \geq 2$. Now $\delta_2 = k^2(4k^2+1)$. Suppose first that $\mathfrak{p} \mid (2k^2-k+1)$. Then $4k^2 \equiv 2k-2 \bmod \mathfrak{p}$, so δ_2 is a square modulo $\mathfrak{p}$ if and only if $2k-1$ is. But $-2k^2 + k \equiv 1 \bmod \mathfrak{p}$, so $-k(2k-1)$ is a square modulo $\mathfrak{p}$. Hence $2k-1$ is a square modulo $\mathfrak{p}$ if and only if $-k$ is.

If $\mathfrak{p} \mid (2k^2+k+1)$, then $4k^2 \equiv -2k-2 \bmod p$, so δ_2 is a square modulo p if and only if $-2k-1$ is. But $-2k^2 - k \equiv 1 \bmod p$, so $k(-2k-1)$ is a square modulo p. Hence $-2k-1$ is a square modulo p if and only if k is. □

Corollary 5.11. *Let K be a number field with a real embedding, and suppose $b = 1$ and the hypotheses of one of Theorems 5.8–5.10 hold. Then none of $\{\pm\delta_n : n = 2, 3, \ldots\}$ is a square in K, $p_n(x)$ is irreducible for all $n \geq 1$ and G_∞ has finite index in C_∞.*

Proof. Because K has a real embedding, -1 is not a square in K, and by hypothesis we have that δ_n is not a square for all $n \geq 2$. Moreover,

$$\delta_n = kP_n(1,1) = k^2(P_{n-1}(1,1)^2 + Q_{n-1}(1,1)^2) > 0 \quad \text{for all } n.$$

Hence $-\delta_n$ cannot be a square in K either. Thus by Theorem 4.5, $p_n(x)$ is irreducible for all $n \geq 1$. Theorem 5.3 applies as well, proving the corollary. □

Remark. In the case $K = \mathbb{Q}$, Corollary 5.11 applies to most values of k. For instance, when k is a positive integer, Theorem 5.10 alone applies to all $k \leq 10\,000$ except for 55 values of k. Of these, the third part of Theorem 5.9 eliminates 21 values.

The remaining 34 values can be ruled out with additional congruences. As in the argument of Theorem 5.9, one can compute the eventually periodic sequence of ordered pairs ($\delta_n \bmod p, \epsilon_n \bmod p$) and check whether δ_n is a square mod p for elements in the cycle and also in the tail, i.e., those elements before the cycle begins. For instance, when $p = 11$ and $k \equiv \pm 1 \bmod 11$, the sequence has a tail of length 2 and a cycle of length 4. The four elements of the cycle have δ_n not a square, and while the second element of the tail is a square, we need not worry about δ_2 being a square, since $4k^2 + 1$ is not a square for k a positive integer. Hence δ_n is not a square for all n when $k \equiv \pm 1 \bmod 11$. This eliminates 11 of the 34 remaining k values.

The 23 values of k that still remain may be handled with congruences involving higher moduli, as given in Table 1. The second column is the modulus p of the congruence, the third and fourth columns give the tail length and cycle length of the sequence ($\delta_n \bmod p, \epsilon_n \bmod p$), and the fifth column gives the n, if any, such that δ_n is a square modulo p. When there is a prime $p < 200$ such that both the tail and cycle contain no δ_n that are squares modulo p, we have listed that prime. Otherwise, we have chosen p to minimize the exceptional n, which are always at most 3. Note that by Theorem 5.8, δ_1 and δ_3 are not squares, and as noted in the previous paragraph δ_2 cannot be a square for k a positive integer. We remark that the δ_n need not be distinct for different values of n, which explains why it is reasonable to have long cycles not containing square δ_n, as in the case of $k = 840$, $p = 197$ or $k = 1620$, $p = 37$.

We now wish to apply Theorem 4.6 to show that $G_\infty \cong C_\infty$ for certain values of k, which demands finding a prime ideal dividing δ_n to odd multiplicity but not dividing any δ_m for $m \neq n$. In light of Lemma 5.2, it is sufficient to know that $v_\mathfrak{p}(k) = 0$ to show that $\mathfrak{p}$ must divide (δ_n) for at most one n. So if $\mathfrak{p} \mid P_n(1,1)$, it is advantageous to know that $\mathfrak{p} \nmid (k)$. We thus study the divisibility by k of the coefficients of $P_n(X,Y)$ and $Q_n(X,Y)$, given in the following data and lemma.

The special case $k = 1$ (corresponding to $\phi(x) = (x^2+1)/x$) plays an important role. We thus put

$$a_n = P_n(1,1)|_{k=1}, \quad b_n = Q_n(1,1)|_{k=1}.$$

Table 1. Congruences used to show $[G_\infty : C_\infty]$ is finite for given values of k.

k	p	tail length	cycle length	exceptional n
444	61	0	4	
840	197	1	84	
1620	37	0	36	
1764	83	0	60	
3000	13	1	12	
3336	37	2	6	$n = 2$
4176	13	1	12	
4224	19	0	6	
4620	41	4	4	$n = 1, 2$
4704	43	2	6	
5184	13	1	12	
5904	31	3	4	$n = 1, 2$
6240	17	4	4	$n = 1$
6384	37	4	2	$n = 2, 3$
6996	71	2	4	$n = 1$
7224	17	4	4	$n = 1$
7620	31	2	4	$n = 1$
7836	13	1	12	
7956	83	1	60	
8004	31	2	4	$n = 1$
8316	19	0	6	
9720	131	3	12	
9804	29	1	12	

Note that $a_1 = 2$ and $b_1 = 1$ and

$$a_n = a_{n-1}^2 + b_{n-1}^2, \quad b_n = a_{n-1}b_{n-1} \tag{33}$$

for $n \geq 2$. We have for instance $a_2 = 5, b_2 = 2, a_3 = 29$, and $b_3 = 10$.

Another important quantity is the constant term of $S_n(1, 1)$, regarded as a polynomial in k. Set

$$\sigma_n = S_n(1, 1)|_{k=0}, \quad \tau_n = T_n(1, 1)|_{k=0}.$$

We have $\sigma_1 = 2$, $\tau_1 = 1$, and it follows from Lemma 5.4 that for all $j \geq 1$,

$$\sigma_{2j+1} = \sigma_{2j}^2 + \tau_{2j}^2, \quad \sigma_{2j} = \tau_{2j-1}^2, \quad \tau_j = \sigma_{j-1}\tau_{j-1}. \tag{34}$$

For example, we have $\sigma_2 = 1$, $\tau_2 = 2$, $\sigma_3 = 5$, $\tau_3 = 2$, $\sigma_4 = 4$, and $\tau_4 = 10$.

The following lemma relates these quantities.

Lemma 5.12. *Let a_n and σ_n be defined as above, and let c be the smallest integer that is at least $n/2$. Then for each $n \geq 2$, $\sigma_n = \prod_{i=1}^{c} a_i^{e_i}$, with $e_c = 1$ when n is odd and $e_c = 0$ when n is even. In particular, σ_n is a product of powers of the a_i with $i < n/2 + 1$.*

Proof. The third part of (34) implies that $\tau_j = \sigma_{j-1} \cdots \sigma_1$ for all j. The second part of (34) then gives that

$$\sigma_{2j} = (\sigma_{2j-2} \cdots \sigma_1)^2. \tag{35}$$

We claim that $\sigma_{2j+1}/\tau_{2j+1} = a_{j+1}/b_{j+1}$. For $j = 1$ we have $\sigma_3/\tau_3 = 5/2 = a_2/b_2$, so the claim holds. Now

$$\begin{aligned} \frac{\sigma_{2j+1}}{\tau_{2j+1}} &= \frac{\sigma_{2j}^2 + \tau_{2j}^2}{\sigma_{2j}\tau_{2j}} = \frac{\tau_{2j-1}^4 + \sigma_{2j-1}^2\tau_{2j-1}^2}{\tau_{2j-1}^3\sigma_{2j-1}} \\ &= \frac{\tau_{2j-1}^2 + \sigma_{2j-1}}{\tau_{2j-1}\sigma_{2j-1}} = \frac{1 + (\sigma_{2j-1}/\tau_{2j-1})^2}{\sigma_{2j-1}/\tau_{2j-1}}. \end{aligned}$$

By inductive assumption the last expression becomes $(1 + a_j^2/b_j^2)/(a_j/b_j)$, and clearing denominators gives $(b_j^2 + a_j^2)/(b_j a_j)$. The claim now follows from (33). The claim, together with the fact that $b_{j+1} = a_j \cdots a_1$ and $\tau_{2j+1} = \sigma_{2j} \cdots \sigma_1$, gives

$$\sigma_{2j+1} a_j \cdots a_1 = \sigma_{2j} \cdots \sigma_1 a_{j+1}. \tag{36}$$

We now prove the lemma by induction. Since $\sigma_1 = 2 = a_1^1$ and $\sigma_2 = 1 = a_1^0$, the lemma holds in these cases. Suppose now that the lemma holds for all σ_i with $i < n$. If $n = 2j$ for some j, then $c = j$ and (35) yields that σ_{2j} is a product of powers of the a_i. The maximum index occurring in the right-hand side of (35) is due to σ_{2j-3}, for which the smallest integer that is at least $(2j-3)/2$ is $j-1$, which is the same as $c-1$. Hence σ_{2j} is a product of powers of the a_i with $i \leq c-1$, as desired.

If $n = 2j+1$, then $c = j+1$. On the right-hand side of (36), we have by inductive hypothesis

$$a_j \mid \sigma_{2j-1}, \quad a_{j-1} \mid \sigma_{2j-3}, \quad \ldots \quad a_1 \mid \sigma_1.$$

Hence we may cancel the $a_j \cdots a_1$ on the left-hand side of (36) to get that σ_{2j+1} is a product of powers of the a_i. The largest index occurring on the right-hand side of

(36) is $j+1$, which equals c. Moreover, by inductive hypothesis the largest index appearing in any of the other factors is j, showing that a_{j+1} appears to only the first power. □

Theorem 5.13. *Let $b = 1$ and assume that -1 is not a square in K and each of the fractional $\mathcal{O}_K$-ideals $(\delta_2), (\delta_3), \ldots, (\delta_m)$ is not the square of a fractional $\mathcal{O}_K$-ideal. Assume also that $v_{\mathfrak{p}}(k) = 0$ for all primes $\mathfrak{p}$ dividing $a_n := P_n(1,1)|_{k=1}$ for some $n < m/2 + 1$. Then $G_m \cong C_m$.*

Proof. We begin by noting that if $m = 1$ or $m = 2$, then $G_m \cong C_m$ is equivalent to $p_n(x)$ being irreducible, which is ensured by -1 and $\pm\delta_2$ not being squares in K (see first paragraph of the proof of Theorem 4.5).

From the proof of Theorem 4.5 it follows that $-1, \pm\delta_2, \pm\delta_3, \ldots, \pm\delta_m$ not being squares in K implies $p_n(x)$ is irreducible for all $n \leq m$. We may thus apply Theorem 4.6, provided that for each n with $3 \leq n \leq m$ we can find a prime $\mathfrak{p}$ with $v_{\mathfrak{p}}(k) = v_{\mathfrak{p}}(2) = 0$, $v_{\mathfrak{p}}(\delta_n)$ odd, and $v_{\mathfrak{p}}(\delta_t) = 0$ for each $t \leq n$.

Now,

$$\delta_n = kP_n(1,1) = k^{s(n)+1}S_n(1,1)$$

by Lemma 5.4. Note that $s(n)+1$ is always even, and thus the squarefree part of the fractional ideal factorization of (δ_n) (which is non-trivial by hypothesis) divides $S_n(1,1)$. Therefore there is a prime ideal $\mathfrak{q} \subset \mathcal{O}_K$ with $\mathfrak{q} \,||\, (S_n(1,1))$, so that $v_{\mathfrak{q}}(\delta_n) = 1$.

We first show that $\mathfrak{q} \nmid (k)$. If $\mathfrak{q} \mid (k)$, then $0 \equiv S_n(1,1) \equiv S_n(1,1)|_{k=0} \bmod \mathfrak{q}$. Thus $\mathfrak{q} \mid (\sigma_n)$ in the notation of Lemma 5.12, and that lemma shows that $\mathfrak{q} \mid (a_n)$ for some $n < m/2 + 1$, contradicting our hypotheses.

Because $\mathfrak{q} \nmid (k)$, we may apply Lemma 5.2 to show that $\mathfrak{q} \nmid (P_i(1,1))$ for all $i \neq n$. It then follows from $kP_i(1,1) = \delta_i$ that $v_{\mathfrak{q}}(\delta_i) = 0$ for all $i \neq n$. In particular, since $P_1(1,1) = 2k$, we have that $\mathfrak{q} \nmid (2)$, completing the proof. □

Corollary 5.14. *Suppose that $K = \mathbb{Q}$. Then there is a density zero set of primes Σ, consisting of 2 and primes congruent to 1 modulo 4, such that if $v_p(k) = 0$ for all $p \in \Sigma$, then $G_\infty \cong C_\infty$.*

Proof. Let Σ be the set of primes dividing $a_n := P_n(1,1)|_{k=1}$ for at least one $n \geq 1$. Note that $a_1 = 2$, and so by assumption $v_2(k) = 0$. Hence by Theorem 5.7 and the fact that $\delta_1 = 2k^2$, δ_n is not a square for all $n \geq 1$. Because $\delta_n > 0$ for all n, this shows that the fractional ideals (δ_n) are all not squares. We now apply Theorem 5.13, showing that $G_\infty \cong C_\infty$.

Note that for $k = 1$ it is certainly the case that $v_p(k) = 0$ for all $p \in \Sigma$, and so we have $G_\infty \cong C_\infty$ in this case. However, a_n is the numerator of $\phi^n(1)$ in the case where $k = 1$, and hence Theorem 6.2 applies to show that the natural density of Σ is zero.

Finally, a simple induction on the recurrence in (33), shows that a_n and b_n are relatively prime for all n. Since a_n is the sum of two relatively prime squares, no prime divisors of a_n can be congruent to 3 modulo 4. □

Remark. It is easy to see if a given prime $p \in \mathbb{Z}$ belongs to Σ. Indeed, letting $\phi(x) = (x^2+1)/x$, then we have $\phi^n(1) = a_n/b_n$, provided that $b_n \neq 0$. Because the preimages of ∞ under ϕ are ∞ and 0, we have that $b_n = 0$ only when $a_{n-1} = 0$; thus to see if $p \mid a_n$ for some n, we need only see if 0 occurs in the sequence $(\phi^n(1) \bmod p)_{n\geq 1}$. This is easily computable, since the sequence repeats after at most $(p-1)/2$ entries. For instance, the primes in Σ less than 2000 are 2, 5, 29, 41, 89, 101, 109, 269, 421, 509, 521, 709, 929, 941, 1549, 1861. Some of these do not divide a_n until n is rather large. For instance, 929 divides a_{42}, but not a_n for $n < 42$.

Remark. As noted in the Introduction, Corollary 5.14 may be far from best possible. Indeed, we have found no $k \in \mathbb{Z}$ for which $[G_\infty : C_\infty] \geq 2$. The only $k \in \mathbb{Q}$ where we can be sure this holds are those of the form a/b, where $4a^2 + b^2$ is a square, since in that case the numerator of $\phi^2(x)$ factors as two quadratic polynomials (see the remark following Theorem 4.5), and note that $kp_2^*(1) = 4a^2 + b^2$, so the image of the action of Gal $(\bar{\mathbb{Q}}/\mathbb{Q})$ on the second level of the tree T_0 of preimages of zero has order two. However, the image of C_∞ on the second level of T_0 has order four, implying that $[G_\infty : C_\infty] \geq 2$.

6. Density of prime divisors in orbits

In this section we use the group-theoretic description of C_n given at the beginning of Section 4 to show that if p_n is separable and $G_\infty \cong C_\infty$, then for any $a \in K$, the density of the set of primes of O_K dividing some element of the orbit $\{\phi^n(a) : n = 1, 2, \ldots\}$ is zero.

We begin with a version of Theorem 2.1 of [8] that applies to a large class of rational functions. By the *natural upper density* of a set of primes S in O_K, we mean

$$D(S) = \limsup_{x\to\infty} \frac{\#\{\mathfrak{p} \in S : N(\mathfrak{p}) \leq x\}}{\#\{\mathfrak{p} : N(\mathfrak{p}) \leq x\}}, \tag{37}$$

where $N(\mathfrak{p}) = N_{K/\mathbb{Q}}(\mathfrak{p})$ is the norm of $\mathfrak{p}$.

Theorem 6.1. *Let $\phi \in K(x)$ be a rational function with p_n separable for all n, and suppose that $\phi^n(\infty) \neq 0$ for all $n > n_0$. Let $a_n = \phi^n(a_0)$ with $a_0 \in K$. Then for any $N > n_0$, the density of primes $\mathfrak{p}$ of K with $v_\mathfrak{p}(a_n) > 0$ for at least one $n \geq 1$ is bounded above by*

$$\frac{1}{\#G_N}\#\{\sigma \in G_N : \sigma \text{ fixes at least one root of } p_N\}. \tag{38}$$

Remark. It is also true that (38) furnishes an upper bound for the density of primes $\mathfrak{p}$ such that 0 is periodic in $O_K/\mathfrak{p}O_K$ under iteration of ϕ. This follows from the fact that 0 is periodic in $O_K/\mathfrak{p}O_K$ if and only if $\phi^{-n}(0) \cap O_K$ is non-empty for all $n \geq 1$; cf. [7], Proposition 3.1.

Proof. We denote by $\mathbb{F}_\mathfrak{p}$ the field $O_K/\mathfrak{p}O_K$, which is the same as $O_{K,\mathfrak{p}}/\mathfrak{p}O_{K,\mathfrak{p}}$, where $O_{K,\mathfrak{p}}$ is the localization of O_K at $\mathfrak{p}$. Any $x \in K$ has a reduction $\bar{x} \in \mathbb{P}^1(\mathbb{F}_\mathfrak{p})$. Provided that $\mathfrak{p} \nmid \mathrm{Res}(p,q)$ (i.e. ϕ has *good reduction* modulo $\mathfrak{p}$), we may reduce the coefficients of ϕ modulo $\mathfrak{p}$ and obtain a morphism $\bar{\phi}\colon \mathbb{P}^1(\mathbb{F}_\mathfrak{p}) \to \mathbb{P}^1(\mathbb{F}_\mathfrak{p})$ such that $\bar{\phi}^N = \overline{\phi^N}$ [17], Theorem 2.15.

Fix $N > n_0$, and consider

$$\Omega_N = \{\mathfrak{p} : \mathfrak{p} \nmid \mathrm{Res}(p,q) \text{ and } \bar{\phi}^N(y) = 0 \text{ has no solution in } \mathbb{P}^1(\mathbb{F}_\mathfrak{p})\}.$$

If $\mathfrak{p} \in \Omega_N$, then we have $\phi^{N+m}(x) = \phi^N(\phi^m(x)) \not\equiv 0 \bmod \mathfrak{p}$ for all $x \in K$, since otherwise $\bar{\phi}^m(\bar{x}) \in \mathbb{P}^1(\mathbb{F}_\mathfrak{p})$ gives a solution to $\bar{\phi}^N(y) = 0$. Thus $v_\mathfrak{p}(a_{N+m}) \leq 0$ for all $m \geq 0$. There are only finitely many $\mathfrak{p}$ with $v_\mathfrak{p}(a_n) > 0$ for some $n < N$, and thus we have

$$D(\Omega_N) \leq D\left(\{\mathfrak{p} : v_\mathfrak{p}(a_n) \leq 0 \text{ for all } n \geq 1\}\right). \tag{39}$$

Because $N > n_0$, we have $\phi^N(\infty) \neq 0$, and hence there are only finitely many $\mathfrak{p}$ for which $\bar{\phi}^N(\infty) = 0$. But if $\bar{\phi}^N(y) = 0$ has a solution in $\mathbb{P}^1(\mathbb{F}_\mathfrak{p})$, then either $\bar{\phi}^N(\infty) = 0$ or $p_N(x) \equiv 0 \bmod \mathfrak{p}$ has a solution in O_K. It follows that

$$D(\Omega_N) = 1 - D(\{\mathfrak{p} : p_N(x) \equiv 0 \bmod \mathfrak{p} \text{ has a solution in } O_K\}). \tag{40}$$

We now use the Chebotarev Density Theorem to show that

$$D(\{\mathfrak{p} : p_N(x) \equiv 0 \bmod \mathfrak{p} \text{ has a solution in } O_K\})$$

is given by the expression in (38), which along with (39) and (40) completes the proof. Except for finitely many primes ramifying in $K(p_N)$, $p_N(x) \equiv 0 \pmod{\mathfrak{p}}$ having a solution in O_K is equivalent to $p_N(x)$ having at least one linear factor in $\mathbb{F}_\mathfrak{p}[x]$. Except for possibly finitely many $\mathfrak{p}$, this implies that $\mathfrak{p}O_L = \mathfrak{p}_1 \cdots \mathfrak{p}_r$, where L/K is obtained by adjoining a root of p_N and at least one of the $\mathfrak{p}_i$ has residue class degree one [13], Theorem 4.12. This is equivalent to the disjoint cycle decomposition of the Frobenius conjugacy class at $\mathfrak{p}$ having a fixed point (in the natural permutation representation of G_N acting on the roots of p_N). From the Chebotarev Density Theorem it follows ([13], Proposition 7.15) that the density of $\mathfrak{p}$ with $\mathfrak{p}O_L$ having such a decomposition is the expression in (38). □

Theorem 6.2. *Assume the hypotheses of Theorem 6.1. Moreover, let C_n be the centralizer in* $\mathrm{Aut}(T_n)$ *of an involution $\iota \in \mathrm{Aut}(T_n)$ acting non-trivially on T_1. Suppose that $\phi \in K(x)$ satisfies $G_n \cong C_n$ for all $n \geq 1$, and let $a_n = \phi^n(a_0)$ with $a_0 \in K$. Then*

$$D(\mathfrak{p} \in O_K : v_\mathfrak{p}(a_n) > 0 \text{ for at least one } n \geq 1) = 0. \tag{41}$$

Proof. For n large enough, we have from Theorem 6.1 that (41) is bounded above by

$$\frac{1}{\#C_n}\#\{\sigma \in C_n : \sigma \text{ fixes at least one top-level vertex in } T_n\}. \tag{42}$$

If $\sigma \in C_n$ satisfies $\sigma|_{T_1} \neq e$, then clearly σ can fix no top-level vertices of T_n. On the other hand, if $\sigma|_{T_1} = e$ then $\sigma \in \ker(C_n \to C_1)$. Therefore by Proposition 4.1, we have that (42) is the same as

$$b_n := \frac{\#\{\sigma \in \mathrm{Aut}(T_{n-1}) : \sigma \text{ fixes at least one top-level vertex in } T_{n-1}\}}{2\#\mathrm{Aut}(T_{n-1})}. \tag{43}$$

From Propositions 5.5, 5.6 of [7], it follows that $b_n = (1 - c_n)/2$, where c_n is given by the evaluation at $z = 0$ of the nth iterate of $f(z) = \frac{1}{2}z^2 + \frac{1}{2}$. This implies that $c_n \to 1$, and thus $b_n \to 0$; indeed, it is enough to note that f maps $I = (0, 1]$ to itself, $f(1) = 1$, and f is increasing on I. Moreover, from Proposition 5.6, part ii, in [7] we have $b_n = 1/n + O((\log n)/n^2)$. □

Acknowledgments. The authors thank the Institute for Computational and Experimental Research in Mathematics for an enjoyable semester, during which a revision of this paper was completed. We also thank the referee for numerous helpful comments.

References

[1] A. F. Beardon, *Iteration of rational functions*. Grad. Texts in Math. 132, Springer, New York 1991. Complex analytic dynamical systems. Zbl 1120.30300 MR 1128089

[2] J. Cullinan and F. Hajir, Ramification in iterated towers for rational functions. *Manuscripta Math.* **137** (2012), no. 3–4, 273–286. Zbl 1235.14023 MR 2875279

[3] X. Faber and A. Granville, Prime factors of dynamical sequences. *J. Reine Angew. Math.* **661** (2011), 189–214. Zbl 05995789 MR 2863906

[4] A. Fröhlich and M. J. Taylor, *Algebraic number theory*, Cambridge Stud. Adv. Math. 27, Cambridge University Press, Cambridge 1993. Zbl 0744.11001 MR 1215934

[5] M. Hindry and J. H. Silverman, *Diophantine geometry*. Grad. Texts in Math. 201, Springer, New York 2000. Zbl 0948.11023 MR 1745599

[6] R. Jones, An iterative construction of irreducible polynomials reducible modulo every prime. *J. Algebra* **369** (2012), 114–128. Zbl 06155226 MR 2959789

[7] R. Jones, Iterated Galois towers, their associated martingales, and the p-adic Mandelbrot set, *Compos. Math.* **143** (2007), no. 5, 1108–1126. Zbl 1166.11040 MR 2360312

[8] R. Jones, The density of prime divisors in the arithmetic dynamics of quadratic polynomials, *J. London Math. Soc.* (2) **78** (2008), no. 2, 523–544. Zbl 1193.37144 MR 2439638

[9] S. Lang, *Algebra*. Third edition, Grad. Texts in Math. 211, Springer, New York 2002. Zbl 0984.00001 MR 1878556

[10] A. Levy, M. Manes, and B. Thompson, Uniform bounds for preperiodic points in families of twists. Available at http://arxiv.org/abs/1204.4447, 2012.

[11] M. Manes, $\mathbb{Q}$-rational cycles for degree-2 rational maps having an automorphism. *Proc. London Math. Soc.* (3) **96** (2008), no. 3, 669–696. Zbl 1213.14048 MR 2407816

[12] J. Milnor, Geometry and dynamics of quadratic rational maps. With an appendix by the author and Lei Tan, *Experiment. Math.* **2** (1993), no. 1, 37–83. Zbl 0922.58062 MR 1246482

[13] W. Narkiewicz, *Elementary and analytic theory of algebraic numbers*. Third edition, Springer Monogr. Math., Springer, Berlin 2004. Zbl 1159.11039 MR 2078267

[14] R. W. K. Odoni, Realising wreath products of cyclic groups as Galois groups. *Mathematika* **35** (1988), no. 1, 101–113. Zbl 0662.12010 MR 0962740

[15] B. Rice, Primitive prime divisors in polynomial arithmetic dynamics. *Integers* **7** (2007), A26, 16 pp. (electronic). Zbl 1165.11028 MR 2312276

[16] J.-P. Serre, Propriétés galoisiennes des points d'ordre fini des courbes elliptiques. *Invent. Math.* **15** (1972), no. 4, 259–331. Zbl 0235.14012 MR 0387283

[17] J. H. Silverman, *The arithmetic of dynamical systems*. Grad. Texts in Math. 241, Springer, New York 2007. Zbl 1130.37001 MR 2316407

[18] M. Stoll, Galois groups over **Q** of some iterated polynomials. *Arch. Math. (Basel)* **59** (1992), no. 3, 239–244. Zbl 0758.11045 MR 1174401

Received August 10, 2011

Rafe Jones, Department of Mathematics, Carleton College, Northfield, MN 55057, U.S.A.
E-mail: rfjones@carleton.edu

Michelle Manes, Department of Mathematics, University of Hawaii, 2565 McCarthy Hall, Honolulu, HI 96813, U.S.A.
E-mail: mmanes@math.hawaii.edu