REDUCTION OF DYNATOMIC CURVES

JOHN R. DOYLE, HOLLY KRIEGER, ANDREW OBUS, RACHEL PRIES,
SIMON RUBINSTEIN-SALZEDO, AND LLOYD WEST

ABSTRACT. In this paper, we make partial progress on a function field version of the dynamical uniform boundedness conjecture for certain 1-dimensional families $\mathcal{F}$ of polynomial maps, such as the family $f_c(x) = x^m + c$ where $m \geq 2$. We do this by making use of the dynatomic modular curves $Y_1(n)$ (resp. $Y_0(n)$) which parametrize maps f in $\mathcal{F}$ together with a point (resp. orbit) of period n for f. The key point in our strategy is to study the set of primes p for which the reduction of $Y_1(n)$ modulo p fails to be smooth or irreducible.

Morton gave an algorithm to construct, for each n, a discriminant D_n whose list of prime factors contains all the primes of bad reduction for $Y_1(n)$. In this paper, we refine and strengthen Morton's results. Specifically, we exhibit two criteria on a prime p dividing D_n: one guarantees that p is in fact a prime of bad reduction for $Y_1(n)$, yet this same criterion implies that $Y_0(n)$ is geometrically irreducible. The other guarantees that the reduction of $Y_1(n)$ modulo p is actually smooth. As an application of the second criterion, we extend results of Morton, Flynn, Poonen, Schaefer, and Stoll by giving new examples of good reduction of $Y_1(n)$ for several primes dividing D_n when $n = 7, 8, 11$ and $f_c(x) = x^2 + c$.

The proofs involve a blend of arithmetic and complex dynamics, reduction theory for curves, ramification theory, and the combinatorics of the Mandelbrot set.

1. INTRODUCTION

We study the reduction modulo p of the dynatomic modular curves, which parametrize polynomial maps together with a point of period n.

For a self-morphism $f : \mathbb{P}^1_K \to \mathbb{P}^1_K$, defined over a field K, denote the n-fold iterate by f^n. A *K-point of period n* for f is a point $P \in \mathbb{P}^1(K)$ such that $f^n(P) = P$; the set of such is denoted $\mathrm{Per}_n(f)(K)$.

Consider a polynomial $f(x, c) \in K[x, c]$. This defines a one dimensional family $\mathcal{F} = \{f_c\}$ of self-morphisms of the projective line, $f_c : \mathbb{P}^1_K \to \mathbb{P}^1_K; x \mapsto f_c(x) = f(x, c)$. The most famous of these is the family of quadratic polynomial maps $f_c(x) = x^2 + c$. Given such a family $\mathcal{F}$, there is an algebraic curve $Y_{1,\mathcal{F}}^{\mathrm{dyn}}(n)$, defined over K, whose K-points correspond to pairs (f_c, P) consisting of a map f_c with $c \in K$ together with a point $P \in \mathrm{Per}_n^*(f_c)(K)$ of (formal) period n, as defined in Definition 3.5. Moreover, $Y_{1,\mathcal{F}}^{\mathrm{dyn}}(n)$ admits a branched cover

$$\pi_1 : Y_{1,\mathcal{F}}^{\mathrm{dyn}}(n) \;\to\; \mathbb{A}^1$$

given by the map $(f_c, P) \mapsto c$.

The curves $Y_{1,\mathcal{F}}^{\mathrm{dyn}}(n)$ are analogous to the classical modular curves $Y_1^{\mathrm{ell}}(N)$ in the case of torsion points on elliptic curves; accordingly they are called *dynatomic modular curves*. For ease of notation, we drop the superscript "dyn" and omit the reference to the family $\mathcal{F}$ where

this causes no ambiguity. The curve classifying *orbits* of (formal) period n is denoted $Y_0(n)$; it is a $\mathbb{Z}/n\mathbb{Z}$-quotient of $Y_1(n)$, giving a tower of covers $Y_1(n) \to Y_0(n) \to \mathbb{A}^1$.

The curves $Y_1(n)$ are smooth and irreducible in characteristic 0 for families of the form $f_c(x) = x^m + c$ where $m \geq 2$ ([Bou92], [Mor96]). In [Mor96], Morton defines a discriminant D_n (see Definition 3.15), whose list of prime factors contains all the primes of bad reduction for $Y_1(n)$; our computational data indicates, however, that $Y_1(n)$ has good reduction at many of the primes dividing D_n. For example, for the family $f_c(x) = x^2 + c$, the discriminant D_5 has 7 odd prime factors (3, 5, 11, 31, 3701, 4217 and 86131), but only two of them (5 and 3701) are primes of bad reduction for $Y_1(5)$. The chief purpose of the present paper is to work toward a finer characterization of the primes p for which the reduction modulo p of $Y_1(n)$ remains smooth and/or irreducible.

One motivation for this project is the conjecture of Morton and Silverman ([MS94]) on uniform bounds for the number of periodic points; see Conjecture 4.1. In Proposition 4.6, we prove that the function field case of the conjecture can be deduced from sufficient information about the set of primes of bad reduction of $Y_1(n)$ as $n \to \infty$. The results in §4 implicitly relate the function field uniform boundedness conjecture to a question about the connectivity of the monodromy graphs for the covers $Y_1(n) \to \mathbb{A}^1$, using results from complex dynamics.

Here is a brief summary of the paper. A more complete summary of the results is contained in §2. The necessary dynamical background is contained in §3. After reaching Remark 3.10, the reader may wish to skip ahead to §4 for the link between our results and the uniform boundedness conjecture, returning to the rest of §3 afterward.

In §5, we give some useful lemmas for detecting good and bad reduction. In §6, we reprove and generalize Morton's necessary criterion for bad primes; specifically, by relating the reduction of $Y_1(n)$ to that of $Y_0(n)$, we give a criterion that guarantees that $Y_1(n)$ actually has good reduction modulo a prime p dividing D_n. In §7, for the family $f_c(x) = x^m + c$ and a prime $p \nmid m$, we prove that $Y_1(n)$ does, in fact, have bad reduction modulo p under the condition that either $p \mid n$ or p divides D_n exactly once. When $n = 5$, the primes 5 and 3701 in the above example fall into precisely these two cases.

In §8, we consider the quadratic polynomial family $f_c(x) = x^2 + c$ and a prime p that divides D_n exactly once. Although the reduction of $Y_0(n)$ may be singular in this case, we prove that it is nonetheless geometrically irreducible. This result is obtained from a study of the connectedness of the "monodromy graph" associated to $Y_0(n)$ in characteristic 0, which we defer to §9; this is a result of independent interest in complex dynamics because of its connection to combinatorics of the Mandelbrot set. Geometrically irreducible reduction is sufficient for the applications to the uniform boundedness conjecture.

In §10, we provide a conceptual explanation why $Y_1(n)$ can have good reduction modulo certain primes that divide D_n. We use this to give new examples of good reduction of $Y_1(n)$ for several primes dividing the ramification discriminant when $n = 7, 8, 11$, thus extending results for $n = 4, 5, 6$ in [Mor98a, page 92], [FPS97, §7], and [Sto08, §2].

Data is included in Appendices A and B.

We note that while this paper was under review, the first author together with Bjorn Poonen used different techniques to make partial but significant progress on Conjecture 4.1; see [DP17, Theorem 1.6].

2. Summary of main results about reduction

In this section, for simplicity, we assume that $f_c(x) = x^m + c$. In §3, we recall the definition and key properties of the dynatomic curves $Y_0(n)$ and $Y_1(n)$. For now, note that $Y_1(n)$ has a natural model over $\mathbb{Z}$ whose generic fiber is smooth; see Definition 3.7 for the case of $Y_0(n)$. The material in §4 builds a connection between the uniform boundedness conjecture for this family of polynomials and the reduction of $Y_0(n)$ modulo p. The term *good reduction modulo p* means that the corresponding curve is smooth in characteristic p.

Following Morton [Mor96], there is a natural choice of polynomial $\delta_n(1, c)$ whose roots are the branch points of the cover $\pi_1 : Y_1(n) \to \mathbb{A}^1$. This polynomial has a factorization over $\mathbb{Z}$ of the form $\delta_n(1, c) = \prod_{d|n} \Delta_{n,d}(c)$ such that: for $d \mid n$ and $d \neq n$, the roots of $\Delta_{n,d}(c)$ are parabolic parameters of formal period n and exact period d, i.e. parameters for which an orbit of primitive period n degenerates to an orbit of primitive period d; and the roots of $\Delta_{n,n}(c)$ are primitive parabolic parameters, i.e. parameters for which two orbits of exact period n collide. See Definition 3.16 for details. Writing $D_{n,d} := \operatorname{disc} \Delta_{n,d}(c)$ and $R_{n,d,e} := \operatorname{res}(\Delta_{n,d}(c), \Delta_{n,e}(c))$, then the discriminant $D_n := \operatorname{disc} \delta_n(1, c)$ has a factorization

$$D_n = \prod_{d|n} D_{n,d} \cdot \prod_{\substack{d|n, e|n \\ d<e}} R_{n,d,e}^2.$$

2.1. **Bad reduction.** We prove several results characterizing primes of bad reduction. First, if $n > 3$, $p \mid n$, and $p \nmid m$, we prove that $Y_1(n)$ has bad reduction at p in Proposition 7.1. Here is the main result about bad reduction:

Theorem A. *(Theorem 7.2) Let $p \nmid m$ be prime. If $\operatorname{ord}_p D_{n,n} = 1$, then the curve $Y_0(n)$ has bad reduction. If $\operatorname{ord}_p D_n = 1$, then the curve $Y_1(n)$ has bad reduction.*

Remark 2.1. In computational examples with $n \leq 6$ and $f_c(x) = x^2 + c$ the only odd primes of bad reduction for $Y_0(n)$ are precisely those accounted for in Theorem A. Further computations with $n = 7$ and $n = 8$ confirm this observation for many small prime factors of $D_{n,n}$; namely for the primes appearing in Tables 3–7. This explains the bad reduction of $Y_0(n)$ at "mysterious large primes," such as $p = 3701$ for $n = 5$ and $p = 8029187$ for $n = 6$, though it does *not* give a straightforward recipe for computing these primes!

Despite the fact that $Y_0(n)$ has singular reduction at primes that exactly divide the discriminant $D_{n,n}$, we prove the following theorem.

Theorem B. *(Corollary 8.4) If $f_c(x) = x^2 + c$, and if an odd prime p divides $D_{n,n}$ exactly once, then $Y_0(n)$ is geometrically irreducible modulo p.*

This theorem is deduced – using the theory of the étale fundamental group in mixed characteristic – from a result on the connectedness of the "monodromy graph" in characteristic 0. The result on the monodromy graph is proved in §9 using the combinatorics of the Mandelbrot set.

In fact, if one can show that all odd primes of bad reduction of $Y_0(n)$ are explained by Theorem A, then one can use Theorem B to obtain uniform bounds on periodic points of $x^2 + c$ over function fields (see Propositions 4.3 and 4.6). In particular, we show that one could deduce the dynamical uniform boundedness conjecture over function fields (see 4.1) from the existence of a sufficiently slow-growing sequence of primes $\{p_n\}$ with the property that $X_0(n)$ has geometrically irreducible reduction modulo p_n for each sufficiently large n (see Proposition 4.6). For $n \leq 8$, we have computed that Theorem A explains all of these primes.

2.2. **Good reduction.** If $p \nmid m$ is a prime of bad reduction for $Y_1(n)$, Morton proved that $p \mid D_n$ [Mor96, Theorem 15]. We reprove and strengthen this result using general techniques relating reduction of branched covers to the geometry of the branch locus. This gives the new criterion under which $Y_1(n)$ actually has good reduction for a prime $p \mid D_n$.

Theorem C. *(Propositions 6.5, 6.8, and 6.17) The primes $p \nmid m$ of bad reduction for $Y_0(n)$ divide $D_{n,n}$. There is a divisor D'_n of D_n with the property that $Y_1(n)$ actually has good reduction modulo p if $p \nmid D'_n$.*

We refer to Remarks 6.10 and 6.19 for the definition of D'_n.
Theorem C depends on the next result connecting the geometry of $Y_1(n)$ and $Y_0(n)$.

Theorem D. *(Propositions 6.12 and 6.15) Suppose $p \nmid nm(m-1)$.*
   (1) If n is prime, then $Y_1(n)$ has good reduction modulo p if and only if $Y_0(n)$ has good reduction modulo p.
   (2) The curve $Y_1(n)$ is geometrically irreducible modulo p if and only if $Y_0(n)$ is geometrically irreducible modulo p.

Remark 2.2. In Theorem D:
   (1) Part (1) is false when n is composite; see, for example, $n = 6$ and $p = 67$ when $m = 2$ in Remark 6.14.
   (2) Part (2) is a strengthening of [Mor96, Proposition 17].
   (3) Combining part (2) with Theorem C shows that the reduction of $Y_1(n)$ modulo p is geometrically irreducible modulo p if $p \nmid nm(m-1)D_{n,n}$.

2.3. **New applications for good reduction.** In §10, we restrict to the case $f_c(x) = x^2 + c$ and analyze the reduction of $Y_0(n)$ above the points $c = 0$ and $c = -2$. These values of c are special because they yield maps that come from algebraic groups, namely power and Chebyshev maps, respectively. We develop a geometric strategy to show that, for particular n and p:
   (1) $Y_0(n)$ does not obtain a singularity above $c = 0$ or $c = -2$ when reduced modulo p,
   (2) and the power of p dividing $D_{n,n}$ is fully explained by branch points colliding with $c = 0$ or $c = -2$ modulo p.

This allows us to verify certain primes of good reduction computed in [Mor98a], [FPS97], [Sto08] when $n = 4, 5, 6$. Furthermore, in Theorem 10.1, we prove that $Y_0(n)$ has good reduction modulo p, even though p divides $D_{n,n}$ in the following new cases:

$Y_0(7)$ for $p = 3, 43, 127$,
$Y_0(8)$ for $p = 3, 5, 17, 257$, and
$Y_0(11)$ for $p = 3, 23, 89, 683$.

3. Dynatomic modular curves

In this section, we review facts about the moduli spaces for polynomial maps with periodic points. Let K be a field.

3.1. Periodic points, dynatomic polynomials and dynatomic curves.

We begin by recalling the following standard definitions from dynamics ([Sil07, Chapter 4]):

Definition 3.1. Let $f : \mathbb{P}^1 \to \mathbb{P}^1$ be a rational map defined over a field K.

(1) Say $P \in \mathbb{P}^1(K)$ is a *point of period n* for f if $f^n(P) = P$. Denote the set of all K-points of period n for f by $\mathrm{Per}_n(f)(K)$. Also write $\mathrm{Per}(f)(K) = \bigcup_{n=1}^{\infty} \mathrm{Per}_n(f)(K)$; a point $P \in \mathrm{Per}(f)(K)$ is said to be a *periodic point*.

(2) For $P \in \mathrm{Per}_n(f)(K)$ the *n-th multiplier* is defined to be $\lambda_n(P) = (f^n)'(P)$.

(3) Say $P \in \mathbb{P}^1(K)$ is a *point of primitive period n* for f if $f^n(P) = P$ and $f^m(P) \neq P$ for $1 \leq m < n$.

Remark 3.2. By the chain rule, any two periodic points in the same orbit have the same multiplier, so it also makes sense to define the multiplier $\lambda_n(\mathcal{O})$, where $\mathcal{O}$ is an *orbit* of periodic points. If $K = \mathbb{C}$, one calls an n-cycle $\mathcal{O}$ *attracting*, *indifferent*, or *repelling* according to whether $|\lambda_n(\mathcal{O})|$ is less than 1, equal to 1, or greater than 1, respectively.

We now turn to the construction of the dynatomic curves $Y_1(n)$. Let $\mathcal{F}$ be a family of polynomials of the form $f_c(x) = f(x, c) \in \mathbb{Z}[x, c]$. Then, for any positive integer n, let $\Psi_{\mathcal{F},n}(x, c) = f_c^n(x) - x \in \mathbb{Z}[x, c]$. A point (z, b) on the variety defined by $\Psi_{\mathcal{F},n}(x, c)$ in $\mathbb{A}^2$ corresponds to a polynomial $f_b(x)$ together with a point z of period n for f_b. However, this variety is reducible, since $\Psi_{\mathcal{F},d}(x, c)$ is a factor of $\Psi_{\mathcal{F},n}(x, c)$ whenever $d \mid n$.

Therefore one defines

$$\Phi_{\mathcal{F},n}(x, c) = \prod_{d|n} \Psi_{\mathcal{F},d}(x, c)^{\mu(n/d)}.$$

By Möbius inversion, there is a factorization

$$\Psi_{\mathcal{F},n}(x, c) = \prod_{d|n} \Phi_{\mathcal{F},d}(x, c).$$

It can be shown that $\Phi_{\mathcal{F},n}(x, c)$ is a polynomial (see [Sil07, Theorem 4.5]); it is called the *n-th dynatomic polynomial for the family $\mathcal{F}$*.

Definition 3.3. The *dynatomic modular curve* is the subscheme $Y_{\mathcal{F},1}(n)$ of $\mathbb{A}^2_{\mathbb{Z}}$ defined by

$$\Phi_{\mathcal{F},n}(x, c) = 0.$$

We shall generally omit the reference to the family $\mathcal{F}$ where unambiguous, writing $Y_1(n)$ for $Y_{\mathcal{F},1}(n)$. We also use the shorter term "dynatomic curve" synonymously with "dynatomic modular curve". Given a scheme S, write $Y_1(n)_S$ for $Y_1(n) \times_{\mathbb{Z}} S$. If $S = \mathrm{Spec}\, R$, we write $Y_1(n)_R$ for short.

Definition 3.4. The *reduction modulo p* of $Y_1(n)$ is the fiber $Y_1(n)_{\mathbb{F}_p}$ of $Y_1(n)$ over the prime p of $\mathbb{Z}$.

In general, a point $(w, b) \in Y_1(n)(K)$ corresponds to a polynomial f_b and a point $w \in K$ of primitive period n for f_b. However, there are finitely many points where orbits of period n degenerate to orbits of lower primitive period. One therefore adopts the following definition:

Definition 3.5. Let $f : \mathbb{P}^1 \to \mathbb{P}^1$ be a polynomial map defined over K. Say $P \in \mathbb{P}^1(K)$ is a *point of formal period n* for f if P is a point of primitive period d where either

- $n = d$;
- $n = ds$ and $\lambda_d(P)$ is a primitive s-th root of unity; or
- $n = dsp^e$ and $\lambda_d(P)$ is a primitive s-th root of unity, $e \geq 1$, and $p = \mathrm{char}\, K$.

Denote the set of all K-points of formal period n for f by $\mathrm{Per}_n^*(f)(K)$.

Proposition 3.6 ([Sil07, Theorem 4.5]). *For a field K, the K-points of $Y_1(n)_K$ are in bijection with the set*

$$\{(f_c, P) \ : \ c \in K, \ P \in \mathrm{Per}_n^*(f_c)(K)\}$$

3.2. **The moduli of periodic orbits.** There is a morphism $\pi_1 : Y_1(n) \to \mathbb{A}^1$ defined by projection onto the second coordinate c. The cover $\pi_1 : Y_1(n) \to \mathbb{A}^1$ has a cyclic subgroup of automorphisms $\langle \sigma \rangle \cong \mathbb{Z}/n\mathbb{Z}$, where σ acts by $(x, c) \mapsto (f_c(x), c)$. Hence, $\langle \sigma \rangle$ acts on $Y_1(n)_S$ for any scheme S. The quotient under this action is a moduli space for pairs $(f_c, \mathcal{O})$ consisting of a polynomial f_c together with an orbit $\mathcal{O}$ of formal period n. Note that $\mathcal{O}$ may be defined over K even if it contains points that are not.

Definition 3.7. For any ring R, let $Y_0(n)_R = Y_1(n)_R/\langle \sigma \rangle$.

There are morphisms φ and π_0 as in the diagram:

3.3. **Completions of the dynatomic curves.** Let K be a field. It is sometimes convenient in later proofs to work with a completion, $X_1(n)_K$, of $Y_1(n)_K$, which we now define. The completion that we use is the one that is regular at infinity.

A family of polynomials $\mathcal{F}$ is *homogenous* if it is of the form $\mathcal{F} = \{f(x, c)\}$, such that $f(x, u^r) \in \mathbb{Z}[x, u]$ is a homogenous polynomial of degree m in u and x, for some positive integer r, and such that $f(x, 0) = x^m$.

Definition 3.8. Let $\mathcal{F}$ be a homogenous family of polynomials. Let K be a field such that $\mathrm{disc}\, f(x, 1) \neq 0$ in K. Let V be the singular locus of $Y_i(n)_K$, and let W be the image of V under $\pi_i : Y_i(n)_K \to \mathbb{A}^1_K$ for $i \in \{0, 1\}$. Embed $\mathbb{A}^1_K$ in $\mathbb{P}^1_K$ and let $Z_i(n)_K$ be the normalization of $\mathbb{P}^1 \setminus W$ in the total ring of fractions of $Y_i(n)_K$. Then we define $X_i(n)_K$ to be the result of gluing $Z_i(n)_K$ to $Y_i(n)_K$ along $Y_i(n)_K \setminus V$.

Remark 3.9. By [Mor96, Proposition 10], under the hypotheses of the definition, the points at infinity in $X_1(n)_{\overline{K}} \setminus Y_1(n)_{\overline{K}}$ correspond to the periodic sequences $\{\zeta_j : j \geq 0\}$ that are of exact period n, where each ζ_j is a root of $f(x, 1) = 0$ in $\overline{K}$.

Remark 3.10. If $Y_i(n)_K$ is smooth, then $X_i(n)_K$ is simply the smooth projective completion of $Y_i(n)_K$.

It is occasionally convenient to have a model of the curves $X_i(n)$ over a mixed characteristic dvr. We define one as follows:

Definition 3.11. Let R be a characteristic $(0, p)$ discrete valuation ring with field of fractions K and residue field k. Suppose that $\mathcal{F} = \{f(x, c)\}$ is a homogenous family of polynomials with $p \nmid \operatorname{disc} f(x, 1)$ and such that $Y_i(n)$ is smooth and irreducible in characteristic zero. Then we define $\mathfrak{X}_i(n)_R$ to be the normalization of $\mathbb{P}^1_R$ in the function field extension $K(Y_i(n)_K)/K(c)$. The special and generic fibers of $\mathfrak{X}_i(n)_R$ are written $\mathfrak{X}_i(n)_k$ and $\mathfrak{X}_i(n)_K$, respectively.

Remark 3.12. Since $\mathfrak{X}_i(n)$ is normal, it is smooth in codimension 1. In particular, the generic fiber $\mathfrak{X}_i(n)_K$ is the smooth curve $X_i(n)_K$. The special fiber $\mathfrak{X}_0(n)_k$ is a quotient of $\mathfrak{X}_1(n)_k$ by $\mathbb{Z}/n\mathbb{Z}$; see Remark 6.6.

Remark 3.13. Under the assumptions of the definition, the affine part of $\mathfrak{X}_1(n)_R$ is none other than the scheme $Y_1(n)_R$, and thus its special fiber is just $Y_1(n)_k$; see Proposition 6.4.

3.4. **Parabolic parameters.** The branched cover of curves $\pi_1 : Y_1(n)_K \to \mathbb{A}^1_K$ has degree $\nu(n) = \sum_{d|n} m^d \mu(n/d)$, where m is the degree in x of $f_c(x)$. For generic c, the map f_c has $\nu(n)$ points of primitive period n that fall into $r(n) = \nu(n)/n$ orbits; these constitute the fiber of π above c. For finitely many c, the fiber of π_1 above c consists of fewer than $\nu(n)$ points; such parameters are called *parabolic parameters*. The points of multiplicity greater than 1 in the fiber of π_1 above a parabolic parameter c are called *parabolic points* and the orbit of such a point under f_c is called a *parabolic orbit*. In characteristic zero, there is at most one parabolic orbit in any fiber of π_1; see Proposition 3.25(3).

Remark 3.14. For example, if f_c has an orbit $\mathcal{O}$ whose primitive period d is strictly less than its formal period n, then $\mathcal{O}$ is a parabolic orbit. In this case, $\lambda_d(\mathcal{O})$ is a primitive n/d-th root of unity (assuming char $K \nmid n$) and $\lambda_n(\mathcal{O}) = \lambda_d(\mathcal{O})^{n/d} = 1$.

Definition 3.15. For the generic polynomial f_c, let $\alpha_1, \ldots, \alpha_r$ be representatives for the $r = \nu(n)/n$ orbits of formal period n, i.e. α_i is a choice of point in the i-th orbit. Define

$$\delta_n(x, c) = \prod_{i=1}^{r} (x - \lambda_n(\alpha_i)) \text{ and } D_n = \operatorname{disc}(\delta_n(1, c)).$$

By [MV95, Theorem A], $\delta_n(x, c) \in \mathbb{Z}[x, c]$, and so $D_n \in \mathbb{Z}$.

Recall the following definitions from [MV95]:

Definition 3.16. Let $C_m(x)$ denote the m-th cyclotomic polynomial and res the resultant with respect to x. For $d \mid n$ and $d \neq n$, define

$$\Delta_{n,d}(c) = \operatorname{res}_x(C_{n/d}(x), \delta_d(x, c)).$$

By Remark 3.14, $\delta_n(1, c)$ is divisible by $\Delta_{n,d}(c)$. Define $\Delta_{n,n}(c)$ by

(3.1) $$\delta_n(1, c) = \Delta_{n,n}(c) \prod_{\substack{d|n \\ d \neq n}} \Delta_{n,d}(c).$$

There is therefore a factorization

$$(3.2) \qquad D_n = \prod_{d|n} \operatorname{disc}(\Delta_{n,d}) \cdot \prod_{\substack{d|n, e|n \\ d<e}} \operatorname{res}(\Delta_{n,d}, \Delta_{n,e})^2.$$

Let $D_{n,d} = \operatorname{disc}(\Delta_{n,d})$ and $R_{n,d,e} = \operatorname{res}(\Delta_{n,d}, \Delta_{n,e})$.

With this notation, we can state the following key result:

Theorem 3.17 ([MV95, Theorem A]). *The $\Delta_{n,d}$ lie in $\mathbb{Z}[c]$ for all $d \mid n$, and satisfy the following identities:*

$$\operatorname{res}_x(\Phi_n(x,c), \Phi_d(x,c)) = \Delta_{n,d}(c)^d \qquad for\ d \mid n\ and\ d \neq n$$

and

$$(3.3) \qquad \operatorname{disc}_x(\Phi_n(x,c)) = \pm \Delta_{n,n}(c)^n \prod_{\substack{d|n \\ d\neq n}} \Delta_{n,d}(c)^{n-d}.$$

Moreover, if $\alpha_1, \ldots, \alpha_r$ are the representatives for the $r = \nu(n)/n$ orbits of formal period n for f_c and if

$$L_i(x) = \prod_{q=0}^{n-1}(x - f_c^q(\alpha_i)),$$

then, for some unit η in $\mathbb{Z}[c, \alpha_1, \ldots, \alpha_r]$,

$$(3.4) \qquad \Delta_{n,n}(c) = \eta \prod_{i\neq j} L_j(\alpha_i).$$

Theorem 3.17 shows that the parabolic parameters are precisely the roots of $\delta_n(1,c) = 0$. Moreover the factorization (3.3) can be interpreted as follows. For $d \mid n$ and $d \neq n$, the roots of $\Delta_{n,d}(c)$ are parabolic parameters for which an orbit of primitive period n degenerates to an orbit of primitive period d; in this case, $\lambda_d(x)$ is a primitive s-th root of unity for any point z in the parabolic orbit, where $s = n/(dp^e)$. For $d = n$, the identity (3.4) shows that the roots of $\Delta_{n,n}(c)$ are parameters for which two orbits of formal period n collide.

Definition 3.18. Roots of $\Delta_{n,d}(c)$, for $d \neq n$, are called *satellite parabolic parameters*. Roots of $\Delta_{n,n}(c)$ are called *primitive parabolic parameters*.

3.5. **Smoothness and irreducibility.** The next hypothesis was introduced in [Mor96].

Definition 3.19. Let $f(x,c) \in \mathbb{Z}[x,c]$ be a polynomial of degree α in x. The family $f_c(x) = f(x,c)$ satisfies *hypothesis (H)* if:
   (1) $f(x,0) = x^\alpha$ and $f(x,u^m)$ is a homogeneous polynomial in $\mathbb{Z}[x,u]$ for some m;
   (2) $\gcd(\operatorname{disc} f(x,1), \alpha^n - 1) = 1$;
   (3) $\delta_n(1,c)$ has no multiple roots.

Note that the family $f_c(x) = x^m + c$ satisfies parts (1) and (2) of hypothesis (H) for any $m \geq 2$, since $\operatorname{disc}(x^m + 1)$ is a power of m up to ± 1.

Theorem 3.20 ([Mor96, Theorem B]). *Let $f(x,c)$ be a family satisfying hypothesis (H). Suppose D_n and $\operatorname{disc} f(x,1)$ are nonzero in K. Then $Y_1(n)_K$ is geometrically irreducible.*

Remark 3.21. In [Mor98b] (Theorem B), Morton is able to prove Theorem 3.20 above with part (ii) of hypothesis (H) replaced by "$f(x,1)$ has distinct roots".

Theorem 3.22 ([Bou92, Chapter 3, Theorem 1, $m = 2$ case],[LS94, Theorem 4.1, $m \geq 2$ case]). *For the families of the form $f(x, c) = x^m + c$, the polynomial $\delta_n(1, c)$ has no multiple roots in characteristic 0. In particular, by Theorem 3.20, $Y_1(n)_{\mathbb{Q}}$ is geometrically irreducible.*

Theorem 3.23 ([DH84, Exposé XIV, $m = 2$ case], [GO14, Theorem 1.1, $m \geq 2$ case]). *For the families of the form $f(x, c) = x^m + c$, the curve $Y_1(n)_{\mathbb{C}}$ is smooth.*

Remark 3.24. The proofs of Theorems 3.22 and 3.23 rely on complex analysis and properties of the Mandelbrot set. Another proof is in [BL14, Theorems 1.1–1.2].

3.6. **Dynatomic curves as branched covers.** Consider the tower of covers

$$
\begin{array}{c}
& & X_1(n)_K \\
{\scriptstyle \mathbb{Z}/n\mathbb{Z}} \swarrow & {\scriptstyle \varphi} & \Big\downarrow {\scriptstyle \pi_1 \; \deg = \nu(n)} \\
X_0(n)_K & {\scriptstyle \pi_1} & \\
& {\scriptstyle \pi_0} \searrow & \\
{\scriptstyle \deg = \nu(n)/n} & & \mathbb{P}^1_K.
\end{array}
$$

Proposition 3.25. *Let $\mathcal{F}$ be a family of polynomials satisfying Hypothesis (H). Let K be an algebraically closed field over which $Y_1(n)_K$ is irreducible and $D_n \neq 0$.*

(1) *There are $\nu(n)/m$ distinct points of $X_1(n)_K$ in the fiber of $\pi_1 = \pi_0 \circ \varphi$ above ∞ in $\mathbb{P}^1$, each with ramification index m under π_1. Moreover, these points are all defined over the splitting field of $f(x, 1)$. Note: this holds without part (3) of Hypothesis (H).*

(2) *The cover $\pi_1 : Y_1(n)_K \to \mathbb{A}^1_K$ (resp. $\pi_0 : Y_0(n)_K \to \mathbb{A}^1_K$) is branched over the subset of $\mathbb{A}^1_K$ consisting of the roots of $\delta_n(1, c)$ (resp. $\Delta_{n,n}(c)$).*

(3) *If c is a root of $\Delta_{n,d}(c)$ for $d \leq n$, then there are d ramification points in $\pi_1^{-1}(c)$, which form a single orbit for the action of $\mathbb{Z}/n\mathbb{Z}$. At each such point z:*

(a) *If $d = n$, φ is unramified at z and π_1 has ramification index 2 at z.*

(b) *If $d < n$, φ has ramification index n/d at z and $\varphi(z)$ is unramified under π_0.*

Proof. See [Mor96, Proposition 9 and 10]. $\square$

3.7. **Mandelbrot set.** In this section, we consider the well-studied family $f_c(x) = x^2 + c$; we refer the interested reader to [Bra89], for example, for a survey. For $K = \mathbb{C}$, the base curve $\mathbb{A}^1_{\mathbb{C}}$ in the cover π_1 is known as the "parameter space." The Mandelbrot set $\mathcal{M}$ is the subset of parameters c for which the orbit of the critical point 0 under iteration of f_c is bounded. The Mandelbrot set is connected ([DH84, Corollary 8.3]). A parameter $c \in \mathcal{M}$ for which 0 converges to an attracting cycle of f_c is called *hyperbolic*. Each hyperbolic map f_c has exactly one attracting cycle. On a given connected component of the set of hyperbolic parameters, the period of the attracting cycle is a constant, say n; such a component is called a *hyperbolic component of period n*. The various hyperbolic components form the distinctive "bulbs" of the Mandelbrot set. The following is a well-known result in complex dynamics:

Proposition 3.26. *Let H be a hyperbolic component of period n. Define a map*

$$\lambda_n : H \to \mathbb{D}$$

which takes a parameter $c \in H$ to the n-th multiplier of the unique attracting cycle of f_c. Then λ_n is a conformal isomorphism and it extends to a homeomorphism $\bar{\lambda}_n : \overline{H} \to \overline{\mathbb{D}}$.

Definition 3.27. For a hyperbolic component H of period n, the unique point $c_0 \in H$ with $\lambda_n = 1$ is called the *root of the hyperbolic component H*.

Remark 3.28. The roots of hyperbolic components are precisely the parabolic parameters, i.e. the roots of $\delta_n(1, c)$. Satellite parabolic parameters, i.e. the roots of $\Delta_{n,d}(c)$ with $d \neq n$, are precisely the parameters c such that c lies on the boundary of some hyperbolic component H of period d and also is a root of a hyperbolic component of period n. Such parameters on the boundary of a given component H biject under λ_d with the primitive roots of unity in $\overline{\mathbb{D}}$. Primitive parabolic parameters, i.e. roots of $\Delta_{n,n}(c)$, occur at the cusps of the cardioid shaped hyperbolic components.

We discuss more properties of the Mandelbrot set in §9.

4. Remarks on uniform boundedness and gonality

We state a version of the uniform boundedness conjecture of Morton and Silverman for one dimensional families of polynomials ([MS94]; see [Poo12] for a generalized form).

Conjecture 4.1 (Uniform boundedness for one dimensional families of polynomials)**.** Fix positive integers N and $d \geq 2$. Let k be a field.

(1) Let $\mathcal{F}/\mathbb{Q}$ be a one dimensional family of polynomials of degree d. There exists $B = B_{N,\mathcal{F}} > 0$ such that

$$|\operatorname{Per}(f)(L)| < B$$

for all number fields L such that $[L : \mathbb{Q}] \leq N$ and for all rational maps $f \in \mathcal{F}(L)$.

(2) Let $\mathcal{F}/k$ be a one dimensional family of polynomials of degree d. There exists $B = B_{N,k,\mathcal{F}} > 0$ such that

$$|\operatorname{Per}(f)(L)| < B$$

for all function fields $L = k(C)$, where C/k is a curve with $\gamma_k(C) \leq N$, and for all non-isotrivial maps $f \in \mathcal{F}(L)$. Here $\gamma_k(C)$ denotes the k-gonality of the curve C/k.

Recall that f is *isotrivial* if there is a linear map $\mu \in \bar{k}(z)$ so that $\mu \circ f \circ \mu^{-1}$ is defined over a finite extension of k. In this section, we consider strategies for proving the function field version (2). For the quadratic family $f_c(x) = x^2 + c$ over $\mathbb{Q}$, the rational points of $X_0(n)$ have been computed for $n \leq 6$ ([Mor98a], [FPS97], and, conditional on BSD for the jacobian of a particular curve of genus 4 [Sto08]); but even in this simplest case, little is known about the uniform statement. This difficulty is due in part to lack of understanding about primes of bad reduction (see [Sto08, §5]); for the function field case this is the main difficulty, as we now explain.

Remark 4.2. Let $\mathcal{F}/k$ be a 1-dimensional family of polynomial maps of degree d. Denote by $M(L, \mathcal{F})$ the supremum of periods of any L-rational periodic point for a non-isotrivial map $f \in \mathcal{F}(L)$. Then a bound on $M(L, \mathcal{F})$ gives a bound on $\operatorname{Per}(f)(L)$, since for each period n there are at most d^n points of period n for f. Denote by $N(L, \mathcal{F})$ the maximal period of any L-rational periodic *orbit* for a non-isotrivial map $f \in \mathcal{F}(L)$. Then $M(L, \mathcal{F}) \leq N(L, \mathcal{F})$. An orbit of exact period n for a non-isotrivial map $f \in \mathcal{F}(L)$ corresponds to a non-constant point of $Y_0(n)(L)$. Therefore, to bound $|\operatorname{Per}(f)(L)|$, it suffices to find an n_0 such that $Y_0(n)(L)$ consists only of constant points (i.e., points defined over k) for all $n > n_0$.

Proposition 4.3. *Let k be a field of characteristic p with $p \nmid m$. Let $\mathcal{F}/k$ be a family satisfying parts (1)–(2) of hypothesis (H) (Definition 3.19). Then part (2) of Conjecture 4.1 holds for $\mathcal{F}/k$ if there exists n_1 such that reduction mod p of $X_0(n)$ is geometrically irreducible for all $n > n_1$.*

First recall the following two lemmas on gonality. Let X^{ns} denote the non-singular points of a curve X.

Lemma 4.4. *If $X/\mathbb{F}_q$ is a geometrically irreducible curve, then $\gamma_{\mathbb{F}_q}(X) \geq |X^{ns}(\mathbb{F}_q)|/(q+1)$.*

Proof. If $\varphi : X \dashrightarrow \mathbb{P}^1$ is a nonconstant rational map of degree d, then we have a morphism $\varphi : X^{ns} \to \mathbb{P}^1$ and each of the $q + 1$ points of $\mathbb{P}^1(\mathbb{F}_q)$ can have at most d preimages under φ, so $d(q+1) \geq |X^{ns}(\mathbb{F}_q)|$. $\square$

Lemma 4.5. *Consider an extension $k/\mathbb{F}$ of a perfect field $\mathbb{F}$. Suppose X_n is a sequence of curves defined over $\mathbb{F}$ with $\gamma_{\mathbb{F}}(X_n) \to \infty$ as $n \to \infty$ and $X_n(\mathbb{F}) \neq \varnothing$ for all n. Then $\gamma_k(X_n) \to \infty$ as $n \to \infty$.*

Proof. Under an algebraic extension of fields $L/\mathbb{F}$ we have $\gamma_L(X_n) \geq \sqrt{\gamma_{\mathbb{F}}(X_n)}$ ([Poo07, Theorem 2.5]). In particular, $\gamma_{\overline{\mathbb{F}}}(X_n) \geq \sqrt{\gamma_{\mathbb{F}}(X_n)}$. On the other hand, for any extension $M/\overline{\mathbb{F}}$, we have $\gamma_{\overline{\mathbb{F}}}(X_n) = \gamma_M(X_n)$ (by specialization; see [Poo07, Proposition 1(ii)]). Since k is contained in such an extension $M/\overline{\mathbb{F}}$, we have $\gamma_k(X_n) \geq \gamma_M(X_n) = \gamma_{\overline{\mathbb{F}}}(X_n) \geq \sqrt{\gamma_{\mathbb{F}}(X_n)}$, whence the claim follows. $\square$

Proof of Proposition 4.3. Let L be a function field, $L = k(C)$, where C/k is a curve with $\gamma_k(C) \leq N$. A nonconstant point of $X_0(n)(L)$ corresponds to a dominant rational map $C \dashrightarrow X_0(n)_k$. If $X_0(n)_k$ is geometrically irreducible, the existence of such a map implies that $\gamma_k(C) \geq \gamma_k(X_0(n))$ (see [Poo07, Proposition 1.1]). Therefore, by Remark 4.2, it is enough to find n_0 such that $\gamma_k(X_0(n)) > N$ for all $n > n_0$. That is, we need to show $\gamma_k(X_0(n)) \to \infty$ with n.

Suppose $X_0(n)$ has irreducible reduction mod p for all $n > n_1$. Let $\mathbb{F} = \mathbb{F}_{p^d}$ be the splitting field of $f(x, 1)$ over $\mathbb{F}_p$. By Proposition 3.25(1) and Proposition 6.4 (note that we do not use the results of this section below, so there is no circularity in this forward reference), there are at least $\nu(n)/mn$ smooth points at infinity in $X_0(n)_{\mathbb{F}_p}(\mathbb{F})$. So, by Lemma 4.4, $\gamma_{\mathbb{F}_p}(X_0(n)_{\mathbb{F}_p}) \geq \gamma_{\mathbb{F}}(X_0(n)_{\mathbb{F}}) \geq \nu(n)/(mn(p^d + 1))$ for all $n > n_1$. Since $\nu(n)/n \to \infty$ as $n \to \infty$, we have $\gamma_{\mathbb{F}_p}(X_0(n)_k) \to \infty$. Now by Lemma 4.5, $\gamma_k(X_0(n)) \to \infty$. $\square$

For the function field part of Conjecture 4.1 in characteristic 0, it suffices to find, for each n, one small prime of geometrically irreducible reduction, as in the next result.

Proposition 4.6. *Let k be a field of characteristic zero. Let $\mathcal{F}/k$ be a family satisfying parts (1) and (2) of hypothesis (H) (Definition 3.19), corresponding to a polynomial $f(x, c)$. Then Part (2) of Conjecture 4.1 holds for $\mathcal{F}/k$ if there exists n_0 and a sequence $\{p_n\}_{n>n_0}$ of primes such that: (1) $\nu(n)/n(p_n^{d_n} + 1) \to \infty$ as $n \to \infty$, where d_n is the degree of the splitting field of $f(x, 1)$ over $\mathbb{F}_{p_n}$; (2) $p_n \nmid m$ and (3) for $n > n_0$, the curve $X_0(n)$ has geometrically irreducible reduction mod p_n.*

Lemma 4.7. *Let X be a smooth curve over a number field K, and let $\mathfrak{p}$ be a prime of K with residue field $\mathbb{F}_q$, such that the reduction X' of X modulo $\mathfrak{p}$ is geometrically irreducible. Then $\gamma_K(X) \geq \gamma_{\mathbb{F}_q}(X')$.*

Proof. This is a slight generalization of [Xar12, Proposition 5]. Suppose $f : X \to Y$ is a morphism of degree d where $Y = \mathbb{P}^1_K$. To prove the lemma, it is enough to show that this implies the existence of a morphism $f' : X' \to \mathbb{P}^1_{\mathbb{F}_q}$ of degree $d' \leq d$. Applying [LL99, Proposition 4.14]), there are models $\mathcal{X}$ and $\mathcal{Y}$ of X and Y over $\mathcal{O}_{\mathfrak{p}}$ together with a rational map $\phi : \mathcal{X} \to \mathcal{Y}$ that restricts to a quasifinite map on X'. Let Z be the image of X' under ϕ. Consider $f' : \phi|_{X'} : X' \to Z$. Then f' is finite, since X' is proper. The curve Z smooth, since it is both geometrically irreducible, because X' is, and of arithmetic genus zero, because it is in the special fiber of $\mathcal{Y}$. Moreover, since Z is defined over a finite field, it has a rational point and is therefore isomorphic to $\mathbb{P}^1$. $\square$

Proof of Proposition 4.6. By the same reasoning as in the second paragraph of the proof of Proposition 4.3, the existence of a sequence of primes p_n satisfying (1) and (2) implies that $\gamma_{\mathbb{F}_{p_n}}(X_0(n)_{\mathbb{F}_{p_n}}) \to \infty$ as $n \to \infty$. By Lemma 4.7, this implies $\gamma_{\mathbb{Q}}(X_0(n)_{\mathbb{Q}}) \to \infty$. By Lemma 4.5, $\gamma_K(X_0(n)_K) \to \infty$ for any field K of characteristic 0. By the argument of the first paragraph of the proof of Proposition 4.3, this is sufficient. $\square$

We note that Benedetto [Ben05] and Canci [Can15] have proved a version of Conjecture 4.1 (1) for polynomials and rational functions, respectively, over function fields in characteristic 0; however, their bounds depend on the total number of poles of the coefficients. This is analogous to results given in [MS94] for $\mathbb{Q}$.

5. Preliminaries for good/bad reduction

We use a test of Fulton that detects singularities by comparing the cardinality of fibers of branched covers of curves with the degree of the local ramification divisor. If A is a Dedekind domain, whose maximal ideals all have the same residue field, and $I \subseteq A$ is a nonzero ideal, then the *degree* of I is $\sum_{x \in \mathrm{MaxSpec}\, A} \mathrm{ord}_x I$; this is generally applied to the *discriminant* ideal of an extension of A.

Theorem 5.1 ([Ful69, Theorem 2.3]). *Let $X = \mathrm{Spec}\, R$, where R is a discrete valuation ring with algebraically closed residue field k. Let $x \in X$ be the closed point. Let $f : Y = \mathrm{Spec}\, S \to \mathrm{Spec}\, R$ be a finite, flat, generically separable morphism of degree α. Let δ be the degree of the discriminant of S/R. Then $\delta \geq \alpha - |f^{-1}(x)|$, with equality holding if and only if Y is smooth and f is tamely ramified.*

Proof. That f is generically separable ensures that it is a *covering* in the sense of Fulton. The proposition then follows from [Ful69, Theorem 2.3]. $\square$

Corollary 5.2. *In the situation of Theorem 5.1, there are at least $\alpha - 2\delta$ points in $f^{-1}(x)$ of multiplicity one.*

Proof. By Theorem 5.1, $|f^{-1}(x)| \geq \alpha - \delta$ and $\sum_{y \in f^{-1}(x)} m_y = \alpha$, where m_y is the multiplicity of y. If β is the number of m_y equal to 1, then

$$\sum_{y \in f^{-1}(x)} m_y \geq \beta + 2(\alpha - \delta - \beta) = 2\alpha - 2\delta - \beta,$$

which is greater than α if $\beta < \alpha - 2\delta$. $\square$

For the rest of §5, let R be a complete discrete valuation ring with fraction field K and algebraically closed residue field k. Let $\overline{K}$ be the algebraic closure of K. Let $A = R[\![T]\!]$, and let B be a normal, finite, flat A-algebra of degree α. Write $A_{\overline{K}}, B_{\overline{K}}, A_k, B_k$ respectively for

$A \otimes_R \overline{K}$, $B \otimes_R \overline{K}$, $A \otimes_R k$, and $B \otimes_R k$. Assume B_k is reduced and B_k/A_k is generically étale. Write d_s (resp. d_η) for the degree of the discriminant of B_k/A_k (resp. $B_{\overline{K}}/A_{\overline{K}}$).

Proposition 5.3 ([GM98, I, 3.4])**.** *With notation as above, $d_s = d_\eta$.*

Proof. It suffices to consider the case that B is local, by localizing B at its (finitely many) maximal ideals and summing the equalities. Then the result follows from [GM98, I, 3.4]. $\square$

The following corollaries are important criteria for good and bad reduction, respectively. See [Ray94, §6] for similar material in the Galois case.

Corollary 5.4. *Suppose B is local and $\mathrm{Spec}(B_{\overline{K}}) \to \mathrm{Spec}(A_{\overline{K}})$ is ramified above at most one closed point z, for which the ramification is tame. Then $\mathrm{Spec}(B_{\overline{K}}) \to \mathrm{Spec}(A_{\overline{K}})$ is totally ramified (of index α) above z, $\mathrm{char}(k)$ does not divide α, and B_k is normal; in particular, $\mathrm{Spec}\, B_k$ is smooth.*

Proof. Let $e_1, \ldots, e_r$ be the ramification indices of the r points above z (or above some arbitrary $\overline{K}$-point, if there is no ramified point). Then $\alpha = \sum_{i=1}^r e_i$, and $d_\eta = \sum_{i=1}^r (e_i - 1)$. By Theorem 5.1 and Proposition 5.3, $d_\eta = d_s \geq \alpha - 1$. This is only possible if $r = 1$ and $e_1 = \alpha$, so that $d_\eta = d_s = \alpha - 1$. By Theorem 5.1, B_k is normal and $\mathrm{char}(k) \nmid \alpha$. $\square$

Corollary 5.5. *For any element a in the maximal ideal of R, let d_a be the degree of the discriminant of the extension S_a/R resulting from setting $T = a$ in B/A.*

    (1) If $d_a < d_\eta$, then either B_k is not normal or B_k/A_k is wildly ramified.

    (2) In addition, if there exists a closed point of $\mathrm{Spec}\, A_{\overline{K}}$ whose preimage under the map $\pi : \mathrm{Spec}\, B_{\overline{K}} \to \mathrm{Spec}\, A_{\overline{K}}$ consists of $\alpha - d_a$ points, then B_k is not normal.

Proof. Let ν be the number of maximal ideals in B (which is also the number of maximal ideals in B_k and in S_a, since A is local). Suppose B_k is normal and S_a/R is tamely ramified. Theorem 5.1 applied to B_k/A_k shows that $d_s = \alpha - \nu$. On the other hand, Theorem 5.1 applied to S_a/R shows that $d_a \geq \alpha - \nu = d_s$. Since $d_\eta = d_s$ by Proposition 5.3, this contradicts $d_a < d_\eta$. This proves (1).

Under the additional assumptions of (2), $\nu \leq \alpha - d_a$, and thus $\nu = \alpha - d_a$. By Theorem 5.1, the extension S_a/R is tamely ramified and S_a is normal. The ramification indices of S_a/R are the multiplicities of the points on the fiber of $\mathrm{Spec}\, B \to \mathrm{Spec}\, A$ above the maximal ideal of A, which are also equal to the ramification indices of B_k/A_k. Since S_a/R is tamely ramified, these are not divisible by $\mathrm{char}(k)$. Thus is it not possible for B_k/A_k to be wildly ramified with B_k normal. We conclude that B_k is not normal, proving (2). $\square$

6. Good reduction

We assume throughout this section that the family $\mathcal{F} = \{f_c\} = \{f(x, c)\}$ satisfies Morton's hypothesis (H) from Definition 3.19 for $c = u^m$. If $f_c(x) = x^m + c$, recall that this property is satisfied. In this section, let $Y_1(n)$ denote $Y_{1,\mathcal{F}}(n)$ (and similarly for $Y_0(n)$).

Under hypothesis (H), we prove that the dynatomic curves which are smooth in characteristic zero have good reduction modulo all primes not dividing $D_n = \mathrm{disc}\, \delta_n(1, c)$. We also show a similar statement about irreducibility that was already proven in [Mor96, Theorem 15], but the proof given here avoids working explicitly with factorizations of polynomials.

If R is a ring, recall that $Y_1(n)_R$ is the closed subscheme of $\mathbb{A}_R^2$ cut out by the dynatomic polynomial $\Phi_n \in \mathbb{Z}[x, c]$ and that $Y_0(n)_R := Y_1(n)_R/\langle\sigma\rangle$, where $\sigma(x, c) := (f_c(x), c)$.

Throughout this section, let k be a field of characteristic p. We assume that R is a characteristic zero complete discrete valuation ring with field of fractions K and residue field k. The generic and special fibers of $Y_0(n)_R$ are $Y_0(n)_K$ and $Y_0(n)_k$, respectively. Let ℓ be the splitting field of $f(x,1)$ over k. Note that ℓ/k is separable if $p \nmid \operatorname{disc} f(x,1)$.

Remark 6.1. If $f(x,c) = x^m + c$, then the following Propositions 6.2, 6.3, and 6.5 apply to all primes not dividing m, since $\operatorname{disc} f(x,1)$ is a power of m.

Proposition 6.2. *Let p be a prime not dividing* $\operatorname{disc} f(x,1)$. *Then $Y_i(n)_R$ has reduced special fiber. If $Y_i(n)_K$ is smooth, then $Y_i(n)_R$ is isomorphic to the normalization of $\mathbb{A}_R^1$ in the function field extension $K(Y_i(n)_K)/K(c)$ coming from the generic fiber of $\pi_i : Y_i(n)_R \to \mathbb{A}_R^1$.*

Proof. Write c_k for the reduction of c to $\mathbb{A}_k^1$. The fiber product $Y_1(n)_k \times_{\mathbb{A}_k^1} \operatorname{Spec} \ell((1/c_k))$ is reduced by [Mor96, Proof of Proposition 10], which means that $Y_1(n)_k$ is reduced; (in [Mor96], it is assumed that $Y_1(n)_k$ is irreducible, but this is not used). Since $Y_0(n)_k$ is a quotient of $Y_1(n)_k$, it is reduced as well.

To show the second statement about $Y_i(n)_R$, it suffices to show that $Y_i(n)_K$ smooth implies $Y_i(n)_R$ is normal. Since $Y_i(n)_k$ is reduced, this follows from [Liu02, Lemma 4.1.18]. $\qquad\square$

Proposition 6.3. *Let p be a prime not dividing $m \operatorname{disc} f(x,1)$, where m is the parameter in hypothesis (H) for f. The natural projection $\pi_{1,k} : Y_1(n)_k \to \mathbb{A}_k^1$ given by $(x,c) \mapsto c$ is generically separable of degree $\nu(n)$ (the degree of $Y_1(n)_k \to \mathbb{A}_k^1$).*

Proof. By [Mor96, Proposition 10], taking the base change of $\pi_{1,k}$ by the map $\operatorname{Spec} \ell((1/c)) \to \mathbb{A}_k^1$ yields a separable map π of degree $\nu(n)$; indeed, each point has degree m over $\operatorname{Spec} \ell((1/c))$. Then $\pi_{1,k}$ is generically separable since $\ell((1/c))$ is separable over $k(c)$. $\qquad\square$

Recall that $\mathfrak{X}_i(n)_R$ is the normalization of $\mathbb{P}_R^1$ in $K(Y_i(n)_K)$ for $i \in \{0,1\}$, assuming $Y_i(n)_K$ is smooth.

Proposition 6.4. *Let p be a prime not dividing $m \operatorname{disc} f(x,1)$, let $i \in \{0,1\}$, and assume that $Y_i(n)_K$ is smooth.*

 (1) *The curve $Y_i(n)_k$ is isomorphic to the preimage of $\mathbb{A}_k^1$ in the special fiber $\mathfrak{X}_i(n)_k \to \mathbb{P}_k^1$ of $\mathfrak{X}_i(n)_R \to \mathbb{P}_R^1$.*
 (2) *If k is perfect, then $\mathfrak{X}_i(n)_k \cong X_i(n)_k$. In particular, $\mathfrak{X}_i(n)_k$ is smooth above ∞.*

Proof. By Proposition 6.2, whose hypotheses are satisfied by Remark 6.1, $Y_i(n)_k$ is the special fiber of the normalization of $\mathbb{A}_R^1$ in $K(Y_K)$. The definition of $\mathfrak{X}_i(n)_R$ implies part (1).

To prove part (2), it suffices by part (1) and the definition of $X_i(n)_k$ (Definition 3.8) to show that $\mathfrak{X}_i(n)_k$ is smooth above ∞. By Corollary 5.4, it suffices to show that no branch point of $\mathfrak{X}_i(n)_K = X_i(n)_K \to \mathbb{P}_K^1$ specializes to ∞ other than ∞. But any such finite branch point is a root of $\delta_n(1,c)$ (Proposition 3.25(2)). These roots are integral over R by [Mor96, top of p. 331], so they do not specialize to ∞. $\qquad\square$

Proposition 6.5. *Let p be a prime not dividing $m \operatorname{disc} f(x,1)$.*

 (1) *If $p \nmid D_n$, then $Y_1(n)_k$ is irreducible. If, in addition, $Y_1(n)_K$ is smooth, then so is $Y_1(n)_k$ (cf. [Mor96, Theorem 15]).*
 (2) *The same holds when $Y_1(n)_k$ is replaced by $Y_0(n)_k$ and D_n is replaced by $D_{n,n}$.*

Proof. We may assume k algebraically closed. For (1), first assume that $Y_1(n)_K$ is smooth. By Proposition 6.2, $Y_1(n)_R$ is a normal scheme, and the special fiber $Y_1(n)_k$ of $Y_1(n)_R$ is reduced. By Proposition 6.3, the special fiber of the map $Y_1(n)_R \to \mathbb{P}_R^1$ is generically separable.

Now, let x be a closed point of $\mathbb{A}^1_R$. Then $\widehat{\mathcal{O}}_{\mathbb{A}^1_R,x} \cong R[\![T]\!]$. Since $p \nmid D_n$, Proposition 3.25(2) shows that *at most one* branch point of $Y_1(n)_K \to \mathbb{A}^1_K$ specializes to x. Let $y \in Y_1(n)_k$ be a closed point above x. The extension $\widehat{\mathcal{O}}_{Y_1(n)_R,y}/\widehat{\mathcal{O}}_{\mathbb{P}^1_R,x}$ is of the form considered in §5. By Corollary 5.4, y is a smooth point of $Y_1(n)_R$. Since y is an arbitrary closed point and k is algebraically closed and thus perfect, $Y_1(n)_k$ is smooth. Since $Y_1(n)_K$ is connected by Theorem 3.20, so is the generic fiber $\mathfrak{X}_1(n)_K$ of $\mathfrak{X}_1(n)_R$. By Zariski's Connectedness Theorem, $\mathfrak{X}_1(n)_k$ is also connected. Identify $Y_1(n)_k$ with an open subscheme of $\mathfrak{X}_1(n)_k$ using Proposition 6.4(1). Since all points of $\mathfrak{X}_1(n)_k \setminus Y_1(n)_k$ are smooth by Proposition 6.4(2), $Y_1(n)_k$ is connected. Since $Y_1(n)_k$ is smooth and connected, it is irreducible.

If $Y_1(n)_K$ is not smooth; it is still irreducible by Theorem 3.20. If $\widetilde{Y}_K$ is the normalization of $Y_1(n)_K$, then the induced map $\widetilde{Y}_K \to \mathbb{A}^1_K$ is a branched cover with branch locus contained inside that of $Y_1(n)_K \to \mathbb{A}^1_K$. If $\widetilde{Y}_R$ is the normalization of $\mathbb{A}^1_R$ in $K(\widetilde{Y}_K)$, and $\widetilde{Y}_k$ is its special fiber, then the same proof as in the smooth case shows that $\widetilde{Y}_k$ is smooth and irreducible. Since $Y_1(n)_k$ is birational to $\widetilde{Y}_k$, this finishes the proof of (1).

The same proof works for (2), using Proposition 3.25(3) in place of Proposition 3.25(2). $\square$

Remark 6.6. If $Y_1(n)_K$ is smooth, then $Y_1(n)_R$ and $Y_0(n)_R$ are normal. Since the σ-action is "vertical" and $Y_1(n)_k \to \mathbb{A}^1_k$ is separable, the induced σ-action on $Y_1(n)_k$ is *faithful*.

Good reduction of $Y_1(n)$ can sometimes be characterized in terms of the non-collision of parabolic orbits, as in the following lemma.

Lemma 6.7. *Suppose $p \nmid m \operatorname{disc} f(x,1)$ and $Y_1(n)$ is smooth in characteristic zero. If $Y_0(n)$ has good reduction modulo p, and no branch points of $Y_1(n) \to Y_0(n)$ collide modulo p, then $Y_1(n)$ has good reduction modulo p.*

Proof. Let y be a closed point on the special fiber of $Y_1(n)_R$, and x be its image in $Y_0(n)_R$. By Propositions 6.2 and 6.3, the extension $\widehat{\mathcal{O}}_{Y_1(n)_R,y}/\widehat{\mathcal{O}}_{Y_0(n),x}$ is of the form considered in §5. By Corollary 5.4, y is smooth. The lemma follows. $\square$

Recall the factorization of the discriminant $D_n = \operatorname{disc}(\delta_n(1,c))$ given by equation 3.1:

$$(6.1) \qquad D_n = \prod_{d|n} D_{n,d} \cdot \prod_{\substack{d<e \\ d|n,e|n}} R^2_{n,d,e},$$

where $D_{n,d} = \operatorname{disc}(\Delta_{n,d})$ and $R_{n,d,e} = \operatorname{res}(\Delta_{n,d}, \Delta_{n,e})$.

We can now precisely state a new criterion for good reduction in terms of these factors of D_n. Proposition 6.17 contains an additional criterion for quadratic maps.

Proposition 6.8. *Let $p \nmid m \operatorname{disc} f(x,1)$ and suppose $Y_1(n)$ is smooth in characteristic zero. Suppose that p divides $\prod_{e|n,e<n} R_{n,e,n}$ but does not divide $D_n / \prod_{e|n,e<n} R^2_{n,e,n}$. Then $Y_1(n)$ has good reduction modulo p.*

Proof. Since $p \nmid D_{n,n}$, Proposition 6.5(2) shows that $Y_0(n)$ has good reduction modulo p. The finite branch points of $Y_1(n) \to Y_0(n)$ lie above the roots of $F(c) := \prod_{\substack{d|n \\ d<n}} \Delta_{n,d}(c)$ in $\mathbb{A}^1$. By assumption,

$$p \nmid \operatorname{disc}(F(c)) = \prod_{d<n} D_{n,d} \prod_{\substack{d|n,e|n \\ d<e<n}} R^2_{n,d,e}.$$

So the images of these branch points in $\mathbb{A}^1$ do not collide modulo p. Thus the branch points themselves do not collide on $Y_0(n)_k$. The result follows from Lemma 6.7. $\square$

Remark 6.9. Proposition 6.8 states that the set of primes where $Y_1(n)$ has bad reduction does not contain any prime p which divides D_n only because it divides $\prod_{e<n} R_{n,e,n}$. In Tables 3–7 for $f(x,c) = x^2 + c$, this guarantees good reduction if p appears in the list of (e,n) rows for $e < n$ but not in any of the other rows. When $5 \leq n \leq 8$, this verifies good reduction for $Y_1(n)$ for most primes dividing D_n but not $nD_{n,n}$. For example, this includes $n = 5$ and $p = 86131$; $n = 6$ and $p = 79, 211, 239, 409, 3331$ or 68700493; $n = 7$ and $p = 987211$ or $\scriptstyle 6170196067469432327626567036587732898819798393402690697954898729590090140927$; and $n = 8$ for all $p \mid D_8$ with $p \nmid D_{8,8}$ other than 1321 and 18967471 (there are at least 11 such primes including $p = 53, 593, 12073, 22129, 248117, \ldots$).

Remark 6.10. In light of Proposition 6.8, the quantity D'_n in Theorem C can be taken to be

$$D'_n = \frac{D_n}{\prod_{\substack{e|n \\ e<n}} R^2_{n,e,n}}.$$

6.1. **The case of $f(x,c) = x^m + c$.** In §6.1, we obtain additional results on good reduction when $f(x,c) = x^m + c$. By Theorem 3.23, $X_1(n)$ is smooth for such f in characteristic zero.

Lemma 6.11. *In characteristic zero, for $n > 2$ and $f(x) = x^m + c$, the totally ramified points of $Y_1(n) \to Y_0(n)$ are exactly the points*

$$(6.2) \qquad z = (x_0, c_0) := \left(\frac{\zeta_{(m-1)n}}{m^{1/(m-1)}}, x_0 - x_0^m \right),$$

as $\zeta_{(m-1)n}$ ranges through the $(m-1)$-st roots of the primitive n-th roots of unity. Furthermore, over any $\mathbb{Z}[\zeta_{(m-1)n}, m^{1/(m-1)}]$-algebra, we have

$$(6.3) \qquad \frac{\partial \Phi_n}{\partial c}(z) \prod_{\substack{d|n \\ d\neq 1,n}} \Phi_d(z) = -\frac{(m-1)m^{1/(m-1)}n}{\zeta_n(\zeta_n - 1)^2}$$

for z as in (6.2). Here we take $\zeta_n = \zeta_{(m-1)n}^{m-1}$.

Proof. The points z in (6.2) are exactly those points in $\mathbb{A}^2$ such that x_0 is a fixed point of $x \mapsto x^m + c_0$ with multiplier $\lambda_1(x_0)$ equal to a primitive nth root of unity. This proves the first statement. Next, fix $\zeta_{(m-1)n}$. By [BL14, Lemma 3.7] (using $m = 1$ and $s = n$),

$$\frac{\partial \Phi_n}{\partial c}(z) \prod_{\substack{d|n \\ d\neq 1,n}} \Phi_d(z) = \frac{n \, \partial\rho/\partial c}{\rho(\rho - 1)}(c_0)$$

over $\mathbb{C}$, where ρ is the multiplier function on $Y_1(1)$, thought of as a function of c (note that [BL14] assumes $f(x,c) = x^2 + c$, but the proof of this statement is the same for $x^m + c$ — see also [GO14, Lemma 2.6]). Now, $\rho(c_0) = \zeta_n$. Since $\rho(x,c) = mx^{m-1}$,

$$(6.4) \qquad \frac{\partial \rho}{\partial c} = m(m-1)x^{m-2}\frac{\partial x}{\partial c}.$$

The variables x and c are related by $x^m + c = x$, so $\partial x/\partial c$ at $c = c_0$ equals $1/(1 - mx_0^{m-1})$, or $1/(1 - \zeta_n)$. Plugging in x_0 to (6.4) and simplifying proves (6.3) over $\mathbb{C}$. Since Φ_n is defined over $\mathbb{Z}$, the formula holds as well for all $\mathbb{Z}[\zeta_{(m-1)n}]$-algebras. $\square$

Proposition 6.12. *Assume $f(x,c) = x^m + c$. If p and n are distinct primes with $p \nmid m(m-1)$, then $Y_1(n)$ has good reduction modulo p if and only if $Y_0(n)$ does.*

Proof. Since $Y_1(n)$ is a finite cover of $Y_0(n)$, it has bad reduction if $Y_0(n)$ does — indeed, its minimal semistable model has nodes at the preimages of the nodes of the minimal semistable model of $Y_0(n)$.

Suppose $Y_0(n)$ has good reduction. By Lemma 6.7, it suffices to show that the branch points of $\phi : Y_1(n) \to Y_0(n)$ do not collide on the special fiber. Since ϕ is cyclic of prime order, this is the same as checking that the ramification points do not collide on the special fiber. These points are given in Lemma 6.11. By the assumptions on p, none of them collide on the special fiber. □

Example 6.13. Assume $f(x,c) = x^2 + c$. Let $n = 5$ and $p = 11$. Then $Y_0(5)$ has good reduction modulo 11 since $11 \nmid D_{5,5}$. So $Y_1(5)$ has good reduction modulo 11 by Proposition 6.12 (even though $11 \mid D_{1,5}$ and $11 \mid D_{1,1}$.)

Remark 6.14. Proposition 6.12 is false when n is composite. For example, take the family $f_c(z) = z^2 + c$ with $n = 6$ and $p = 67$. The prime 67 does not divide $D_{6,6}$, so $Y_0(6)$ has good reduction at 67 by Proposition 6.5(2). On the other hand, 67 divides D_n exactly once (it divides $D_{6,3}$); so Theorem 7.2 below shows that $Y_1(6)$ has bad reduction at 67. Nonetheless, $Y_1(6)$ is still *irreducible* modulo 67, as Proposition 6.15 below shows.

Proposition 6.15. *Assume $f(x,c) = x^m + c$. If $p \nmid (m-1)mn$, then $Y_1(n)$ is geometrically irreducible modulo p if and only if $Y_0(n)$ is.*

Proof. Since $Y_1(n)$ is a finite cover of $Y_0(n)$, it is geometrically reducible modulo p if $Y_0(n)$ is. Now, suppose $Y_0(n)$ is geometrically irreducible modulo p. Then the $\mathbb{Z}/n\mathbb{Z}$-action on $Y_1(n)$ in characteristic p permutes the irreducible components transitively. To show that $Y_1(n)$ is geometrically irreducible modulo p, it suffices to show that $Y_1(n) \to Y_0(n)$ has a smooth, totally ramified point in characteristic p, since such a point can lie on only one irreducible component but must simultaneously lie on all irreducible components. Since $p \nmid (m-1)mn$, the right hand side of (6.3) is a p-adic unit. Lemma 6.11 then implies that $\partial \Phi_n / \partial c$ (at the point z) does not vanish modulo p, which shows that z is the desired point. □

6.2. **Quadratic maps.** In this section, we restrict to the quadratic maps $f_c(x) = x^2 + c$. In this case, we can rule out prime factors of D_n that only divide the factors $D_{n,1}$ and $D_{n,2}$.

Lemma 6.16. *Let $d < n$ be positive integers with $d \mid n$. Let p be a prime not dividing $2n/d$. Suppose that there are distinct roots c_1 and c_2 of $\Delta_{n,d}$ and ramification points (x_1, c_1) and (x_2, c_2) of $Y_1(n) \to \mathbb{A}^1$ that collide on the special fiber $Y_1(n)_{\mathbb{F}_p}$. Then c_1 and c_2 lie on the boundaries of distinct period d hyperbolic components of the Mandelbrot set.*

Proof. By Proposition 3.25(4), for $i = 1, 2$, the orbit of x_i is a parabolic orbit for f_{c_i}; hence each c_i lies on the boundary of a unique period d hyperbolic component H_i. The d-th multipliers $\lambda_i = \lambda_d(x_i)$ are primitive (n/d)-th roots of unity (by the comment after Proposition 3.17). Both x_i and c_i are p-adic integers at all odd primes p by [MV95, Corollary 3.5]. Since $x_1 \equiv x_2$ and $c_1 \equiv c_2 \pmod{p}$, it follows that $\lambda_1 \equiv \lambda_2 \pmod{p}$. If $c_1 \neq c_2$ and $H_1 = H_2$, then the multipliers λ_i would be *distinct* primitive (n/d)-th roots of unity (by Remark 3.28), hence distinct mod p for $p \nmid n/d$. Thus $H_1 \neq H_2$. □

Recall that Proposition 6.8 contains a criterion under which $Y_1(n)$ actually has good reduction modulo a prime dividing D_n. We can now precisely state an additional criterion for good reduction in the case of quadratic maps.

Proposition 6.17. *Suppose $p \nmid 2n$ and $n > 2$. Let $S = \{1\}$ if n is odd and $S = \{1, 2\}$ if n is even. If p divides $\prod_{d \in S} D_{n,d}$ but not $D_n / \prod_{d \in S} D_{n,d}$, then $Y_1(n)$ has good reduction at p.*

Proof. Since $p \nmid D_{n,n}$, the curve $Y_0(n)$ has good reduction modulo p by Proposition 6.5(2). By Lemma 6.7, $Y_1(n)$ has good reduction modulo p unless there are two ramification points (x_1, c_1) and (x_2, c_2) of $Y_1(n) \to Y_0(n)$ that collide on the special fiber. In this case, c_1 and c_2 are roots of $\delta_n(1, c)$ that collide modulo p. By the hypotheses, c_1 and c_2 are roots of $\Delta_{n,d}$ for some $d \in S$. By Lemma 6.16, c_1 and c_2 lie on the boundary of distinct hyperbolic components of period d. But there is only one hyperbolic component of period d, since $d \in \{1, 2\}$. $\square$

Remark 6.18. Proposition 6.17 states that the set of primes of bad reduction for $Y_1(n)$ does not contain any prime p which divides D_n only because it divides $D_{n,1}$ (resp. $D_{n,1}D_{n,2}$) when n is odd (resp. n is even). In Tables 3–7, this means bad reduction cannot occur if p appears in the $(1, 1)$ (resp. and/or $(2, 2)$) row but not in other rows. For example, $Y_1(7)$ has good reduction modulo $p = 29$.

Remark 6.19. In light of Proposition 6.17, if $f_c(x) = x^2 + c$ then we can further divide the quantity D'_n from Remark 6.10 by the prime-to-$2n$ part of $\prod_{d \in S} D_{n,d}$.

7. BAD REDUCTION

As in §6, we assume in this section that the family $\mathcal{F} = \{f_c\} = \{f(x, c)\}$ satisfies Morton's hypothesis (H) from Definition 3.19 for $c = u^m$. By $Y_1(n)$ we mean $Y_{1,\mathcal{F}}(n)$ (and the same for $Y_0(n)$). We further assume that $Y_1(n)$ is *smooth* in characteristic zero, so that results about bad reduction are not vacuous: this holds when $f(x, c) = x^m + c$, by Theorem 3.23. In this section we give two criteria for the dynatomic curve to have bad reduction at a prime.

Proposition 7.1. *Assume $f_c(x) = x^m + c$. If p is a prime not dividing m but dividing $n > 3$, then $Y_1(n)$ has bad reduction modulo p.*

Proof. We show that the reduction of (any) point $z = (x_0, c_0)$ from Lemma 6.11 is a singular point of $Y_1(n)_{\overline{\mathbb{F}}_p}$. Since z is a ramification point of $Y_1(n) \to \mathbb{A}^1$, it follows that $\partial \Phi_n / \partial x = 0$ at z, even in characteristic zero. To compute $\partial \Phi_n / \partial c$ at z, we use (6.3) and compute $\prod_{\substack{d|n \\ d \neq 1, n}} \Phi_d(z)$. Writing Ψ_n for $f_c^n(x) - x$, then

$$\prod_{\substack{d|n \\ d \neq 1,n}} \Phi_d = \frac{\Psi_n}{\Phi_1 \Phi_n} = \frac{\Psi_n}{\Psi_1} \prod_{d|n} \Psi_d^{-\mu(n/d)}.$$

The right hand side has an equal number of factors with positive and negative exponents, and all factors (other than the Ψ_n's, which cancel out) have simple zeroes at x_0, so the product is equal to the product of the x-derivatives of the non-Ψ_n terms at x_0. Since $\partial \Psi_d / \partial x = \zeta_n^d - 1$ at z, the right hand side equals

$$\frac{1}{\zeta_n - 1} \prod_{\substack{d|n \\ d \neq n}} (\zeta_n^d - 1)^{-\mu(n/d)}.$$

Since $\zeta_n^d - 1$ is a p-adic unit whenever n/d is not a power of p, the expression simplifies to $(\zeta_n^{n/p} - 1)/(\zeta_n - 1)$ up to p-adic units. Now, using (6.3), there is an equality, up to p-adic units,

$$\frac{\partial \Phi_n}{\partial c}(z) = \frac{(m-1)m^{1/(m-1)}n}{(\zeta_n^{n/p} - 1)(\zeta_n - 1)}.$$

If $p \mid n$ and $n > 3$, then the right hand side has positive p-valuation, and thus reduces to 0 in characteristic p. So the reduction of z modulo p is not smooth. $\square$

Our second criterion for bad reduction is in terms of the valuation of the discriminant:

Theorem 7.2. *Suppose p is a prime not dividing m disc $f(x,1)$. If p divides the order of $D_{n,n}$ (resp. D_n) exactly once, then $Y_0(n)$ (resp. $Y_1(n)$) has bad reduction modulo p.*

Remark 7.3. For $n \le 6$ and $f(x,c) = x^2 + c$, we verified that the *only* odd primes p of bad reduction for $Y_0(n)$ are those with $v_p(D_{n,n}) = 1$.

Remark 7.4. If p divides $D_{n,n}$ (resp. D_n), then the branch points of $Y_0(n) \to \mathbb{P}^1$ (resp. $Y_1(n) \to \mathbb{P}^1$) collide modulo p. The proof of Theorem 7.2 shows that when p divides $D_{n,n}$ or D_n exactly once, then exactly two *ramification* points must also collide modulo p, which is what causes the bad reduction.

We start with two lemmas from algebraic number theory.

Lemma 7.5. *Suppose the leading coefficient of $\phi \in \mathbb{Z}_p[x]$ is a unit. If $v_p(\mathrm{disc}(\phi)) = 1$, then there exists $a \in \mathbb{Z}_p$ such that $v_p(\phi(a)) = 1$. Furthermore, there exist exactly two roots r_1 and r_2 of ϕ whose residues in $\overline{\mathbb{F}}_p$ are the same as that of a, and these are the only two roots that have the same residue in $\overline{\mathbb{F}}_p$.*

Proof. Adjoining a root x of ϕ gives an integrally closed extension of $\mathbb{Z}_p$, as otherwise p^{2r} would divide disc ϕ, where p^r is the index of the ring $\mathbb{Z}_p[x]$ in its integral closure. Suppose the prime (p) splits as $\mathfrak{p}_1^{e_1} \cdots \mathfrak{p}_s^{e_s}$ in $\mathbb{Z}_p[x]$. Since $p \mid \mathrm{disc}(\phi)$ exactly once, $e_i = 1$ for all but one i, for which $e_i = 2$ and $\mathfrak{p}_i$ has residue field of cardinality p. Assume without loss of generality that $i = 1$. Then the reduction $\overline{\phi}$ of $\phi \pmod{p}$ has irreducible factorization $\overline{\phi}_1^2 \overline{\phi}_2 \cdots \overline{\phi}_s$, where $\overline{\phi}_1$ has degree 1. By Hensel's lemma, this lifts to a splitting $\phi = \beta\phi_2 \cdots \phi_s$, where $\beta \in \mathbb{Z}_p[x]$ is a lift of $\overline{\phi}_1^2$. In particular, ϕ has exactly two roots r_1 and r_2 having the same residue in $\overline{\mathbb{F}}_p$, and these roots are conjugate over $\mathbb{Z}_p$ since they are the roots of β. By the definition of the discriminant, $v_p(r_1 - r_2) = 1/2$. Thus $r_1, r_2 = a \pm b\sqrt{p}$, where b is integral over $\mathbb{Z}_p$ with valuation 0 and $a = (r_1 + r_2)/2 \in \mathbb{Z}_p$ is the element we seek. $\square$

Notation. Write $\mathbb{Q}_p^{ur}$ for the completion of the maximal unramified extension of $\mathbb{Q}_p$, and $\mathbb{Z}_p^{ur}$ for its valuation ring.

Lemma 7.6. *Let $a \in \mathbb{Z}_p^{ur}$, and let $S_a/\mathbb{Z}_p^{ur}$ be the extension given by restricting $Y_0(n)_{\mathbb{Z}_p^{ur}} \to \mathbb{A}^1$ above $c = a$. Then the degree of the discriminant of $S_a/\mathbb{Z}_p^{ur}$ is $v_p(\Delta_{n,n}(a))$.*

Proof. Write $R = \mathbb{Z}_p^{ur}$, $k = \overline{\mathbb{F}}_p$, and $K = \mathbb{Q}_p^{ur}$ for simplicity. By [Mor96, Proposition 9d], $\Delta_{n,n}(c)$ is the discriminant of the field extension for $Y_0(n) \to \mathbb{A}^1$ over any field of characteristic zero; (the proposition is only stated for algebraically closed fields, but it holds over all fields since it is a polynomial identity and everything is defined over $\mathbb{Q}$ by Theorem 3.17).

Identifying a with a point in $\mathbb{A}_R^1$ and $\overline{a}$ with its specialization to $\mathbb{A}_k^1$, the map $Y_0(n)_R \times_{\mathbb{A}_R^1}$ $\mathrm{Spec}\,\widehat{\mathcal{O}}_{\mathbb{A}_R^1,\overline{a}} \to \mathrm{Spec}\,\widehat{\mathcal{O}}_{\mathbb{A}_R^1,\overline{a}}$ is finite and free, and thus has discriminant given by a *principal* ideal (δ) of $\widehat{\mathcal{O}}_{\mathbb{A}_R^1,\overline{a}} \cong R[\![T]\!]$, where T is a coordinate vanishing at a, e.g., $T = c - a$. Since the discriminant respects the unramified base change $\mathrm{Spec}\,\widehat{\mathcal{O}}_{\mathbb{A}_R^1,\overline{a}} \to \mathbb{A}_R^1$ and the map $Y_0(n)_R \to$ $\mathbb{A}_R^1$ is not ramified along the special fiber (Propositions 6.2 and 6.3), it follows that $(\delta) =$ $(\Delta_{n,n})$, where $\Delta_{n,n}$ is considered as a function of T. Then the discriminant ideal of $S_a/\mathbb{Z}_p^{ur}$ is just the result of setting $T = 0$ in $(\Delta_{n,n})$, that is, $(\Delta_{n,n}(a))$. The lemma follows. $\square$

Proof of Theorem 7.2. Let $R = \mathbb{Z}_p^{ur}$ and $K = \mathbb{Q}_p^{ur}$. We show that the special fiber of $Y_0(n)_R$ (resp. $Y_1(n)_R$) is singular when p divides $D_{n,n}$ (resp. D_n) exactly once. First note that, by Proposition 6.2, $Y_i(n)_R$ is the normalization of $\mathbb{A}_R^1$ in $K(Y_i(n)_K)$ for $i \in \{0,1\}$. Also recall that the polynomials $\Delta_{n,d}(c)$ are all defined over $\mathbb{Z}$ (Theorem 3.17) and have leading coefficients that are units in $\mathbb{Z}_p$ ([Mor96, top of p. 331]).

If $v_p(D_n) = 1$, then by (6.1) p does not divide any $R_{n,d,e}$ with $d \neq e$ and there exists a unique $d \mid n$ such that p divides $D_{n,d}$. Then Lemma 7.5 guarantees an $a \in \mathbb{Z}_p$ (in particular, in $\mathbb{Z}_p^{ur}$) for which $v_p(\Delta_{n,d}(a)) = 1$, and for which exactly two roots r_1 and r_2 of $\delta_n(1,c)$ have the same reduction as a in an appropriate extension of $\mathbb{Z}_p^{ur}$. Likewise, if $v_p(D_{n,n}) = 1$, then Lemma 7.5 guarantees $a \in \mathbb{Z}_p^{ur}$ for which $v_p(\Delta_{n,n}(a)) = 1$ and two roots r_1 and r_2 of $\Delta_{n,n}(c)$ have the same reduction as a. In either case, let $\overline{a}$ be the specialization of a to $\mathbb{A}_{\overline{\mathbb{F}}_p}^1$.

Let $\mathcal{Z} = \mathrm{Spec}\,\widehat{\mathcal{O}}_{\mathbb{P}_R^1,\overline{a}} \cong \mathrm{Spec}\,R[\![T]\!]$, and let

$$\mathcal{Y}_i := Y_i(n)_R \times_{\mathbb{A}_R^1} \mathcal{Z} \text{ for } i \in \{0,1\}.$$

Now, $\mathcal{Y}_0$ and $\mathcal{Y}_1$ are flat over $\mathcal{Z}$, and the maps $\mathcal{Y}_0 \to \mathcal{Z}$ and $\mathcal{Y}_1 \to \mathcal{Z}$ correspond to extensions $B_0/R[\![T]\!]$ and $B_1/R[\![T]\!]$ as in §5. Write $B_{i,\overline{K}}$, $R[\![T]\!]_{\overline{K}}$ for $B_i \otimes_R \overline{K}$, $R[\![T]\!] \otimes_R \overline{K}$, respectively (with $i \in \{0,1\}$).

Suppose $v_p(D_{n,n}) = 1$. The only two (geometric) branch points on the generic fiber of $\mathcal{Y}_0 \to \mathcal{Z}$ come from r_1 and r_2. The map $Y_0(n) \to \mathbb{P}^1$ has simple ramification above each of them (Proposition 3.25(4a)), which means that the degree of the discriminant of $B_{0,\overline{K}}/R[\![T]\!]_{\overline{K}}$ is 2. By Lemma 7.6, the degree of the discriminant of $\mathcal{Y}_0 \to \mathcal{Z}$ restricted above $c = a$ is $v_p(\Delta_{n,n}(a)) = 1$. The cardinality of a fiber of $\mathcal{Y}_0 \times_R \overline{K} \to \mathcal{Z} \times_R \overline{K}$ above a branch point is $\alpha - 1$, where α is the degree of $\mathcal{Y}_0 \to \mathcal{Z}$. Thus the assumptions of Corollary 5.5(2) hold, and $\mathcal{Y}_0$ has a singular point. Thus $Y_0(n)$ has bad reduction modulo p.

Now, suppose $v_p(D_n) = v_p(D_{n,d}) = 1$ for some $d < n$. The (geometric) branch points of the generic fiber of $\mathcal{Y} \to \mathcal{Z}$ are again r_1 and r_2, but this time there are d ramification points above each of them, each with ramification index n/d (Proposition 3.25(4b)). So the degree of the discriminant of $B_{1,\overline{K}}/R[\![T]\!]_{\overline{K}}$ is $2d(n/d - 1) = 2(n - d)$. On the other hand,

$$v_p(\mathrm{disc}_x(\Phi_n(x,a))) = (n - d)v_p(\Delta_{n,d}) = n - d$$

by Theorem 3.17. Also, the cardinality of a fiber of $\mathcal{Y}_1 \times_R \overline{K} \to \mathcal{Z} \times_R \overline{K}$ above a branch point is $\alpha - (n - d)$, where α is the degree of $\mathcal{Y}_1 \to \mathcal{Z}$. The assumptions of Corollary 5.5(2) hold, and $\mathcal{Y}_1$ has a singular point. Thus $Y_1(n)$ has bad reduction modulo p. $\square$

Remark 7.7. In the proof above, it is not really important to work over $\mathbb{Z}_p^{ur}$ rather than $\mathbb{Z}_p$. We only use a ring with algebraically closed residue field in order to quote results from §5.

8. Irreducibility of the dynatomic curve at primes exactly dividing the discriminant

We assume throughout this section that $f_c(x) = x^2 + c$. Our goal is to show that the reduction of $Y_0(n)$ modulo p is geometrically irreducible whenever $v_p(D_{n,n}) = 1$. This will depend on a result about the so-called *monodromy graph* (Definition 8.2), whose proof we defer until the next section.

First, we recall a few definitions related to étale fundamental groups. If S is a normal scheme and s is a geometric point of S, then $\pi_1(S, s)$ is defined to be the automorphism group of the fiber functor F_s (see, e.g., [SGA03, V] or [Méz00]). Note that different choices s and s' of geometric points lead to isomorphic fiber functors as long as S is connected. A *path* (really, a homotopy class of paths) from s to t is an isomorphism of functors $F_s \to F_t$. If $h : S \to S'$ is a morphism such that $h(s) = s'$, then functoriality gives rise to a homomorphism $\pi_1(S, s) \to \pi_1(S', s')$.

Suppose S is a curve over a perfect field K with smooth projective completion $\widetilde{S}$, and t is a K-point of $\widetilde{S} \setminus S$. Fix a geometric point $\bar{t}$ of $\operatorname{Spec} \widehat{\mathcal{O}}_{\widetilde{S},t} \setminus \{t\} \subseteq S$ corresponding to an algebraic closure $\mathcal{K}$ of $\operatorname{Frac} \widehat{\mathcal{O}}_{\widetilde{S},t}$. The inertia group of $\pi_1(\operatorname{Spec} \widehat{\mathcal{O}}_{\widetilde{S},t} \setminus \{t\}, \bar{t})$ is simply the inertia group of $\mathcal{K} / \operatorname{Frac} \widehat{\mathcal{O}}_{\widetilde{S},t}$. This can be identified with a subgroup I of $\pi_1(S, \bar{t})$ via the inclusion $\operatorname{Spec} \widehat{\mathcal{O}}_{\widetilde{S},t} \setminus \{t\} \subseteq S$. If s is a geometric point of S and we additionally fix a path from $\bar{t}$ to s in S, then the *inertia subgroup* of $\pi_1(S, s)$ at t is the image of I under this path. If the path and the geometric point $\bar{t}$ are not fixed, then the inertia subgroup of $\pi_1(S, s)$ at t is only well-defined up to conjugacy.

Proposition 8.1. *Let R be a mixed characteristic complete discrete valuation ring with fraction field K and residue field k. Let S be a smooth, projective, geometrically connected relative R-curve. Let $x_1, \ldots, x_n$ be distinct R-sections of S. Let $h : X \to S$ be a finite flat R-morphism, such that the special fiber of X is reduced, the special fiber h_k of h is generically separable, and the generic fiber h_K of h is étale outside $\{x_{1,K}, \ldots, x_{n,K}\}$, where $x_{i,K}$ is the intersection of x_i with the generic fiber S_K of S. Pick a $\overline{K}$-point $s_{\overline{K}} \in S_K \setminus \{x_{1,K}, \ldots, x_{n,K}\}$ whose specialization does not coincide with that of any $x_{i,K}$.*

Let $m \leq n$ be such that $x_1, x_2, \ldots, x_m$ are pairwise non-overlapping. For each $1 \leq i \leq m$, fix a geometric point $\eta_{i,K}$ corresponding to the algebraic closure of $\operatorname{Frac} \widehat{\mathcal{O}}_{S_K, x_{i,K}}$ and a path from $\eta_{i,K}$ to $s_{\overline{K}}$. If the action of the subgroup of $\pi_1(S_K \setminus \{x_{1,K}, \ldots, x_{n,K}\}, s_{\overline{K}})$ generated by the inertia subgroups at $x_{1,K}, \ldots, x_{m,K}$ on a geometric fiber of h_K above $s_{\overline{K}}$ is transitive, then the special fiber of X is geometrically irreducible.

Proof. We may assume that k is algebraically closed. We use the notation h_k (resp. X_k, S_k, $x_{i,k}$) to represent the special fiber of h (resp. X, S, x_i), and similarly for K and the generic fiber. Write T for $x_1 \cup \cdots \cup x_n$, and define T_k and T_K similarly. Let $U \subseteq X$ be equal to $h^{-1}(S \setminus T)$, with U_k and U_K defined accordingly. By assumption, $s_{\overline{K}}$ specializes to a point $s_k \in S_k \setminus T_k$. Since X has reduced special fiber and h_k is generically separable, purity of the branch locus ([Sza09, Theorem 5.2.13]) shows that $h_k|_{U_k}$ is étale, and thus U_k is smooth. So connectedness of U_k implies irreducibility of U_k and consequently irreducibility of X_k. We are reduced to showing that U_k is connected, or equivalently, that the action of $\pi_1(S_k \setminus T_k, s_k)$ on the fiber $h_k^{-1}(s_k)$ is transitive.

Let $\eta_{i,k}$ be a geometric point of $S_k \setminus T_k$ corresponding to an algebraic closure of $\operatorname{Frac} \widehat{\mathcal{O}}_{S_k, x_{i,k}}$. Since $x_1, \ldots, x_m$ do not overlap on the special fiber, Corollary 5.4 implies that h_k is tamely

ramified above $x_{1,k}, \ldots, x_{m,k}$ with the same ramification indices as h_K, and all points lying above $x_{1,k}, \ldots, x_{m,k}$ are smooth. Thus, for $1 \leq i \leq m$, specialization gives a bijection from $h_K^{-1}(\eta_{i,K})$ to $h_k^{-1}(\eta_{i,k})$ and the action of the inertia subgroup of $\pi_1(S_K \setminus T_K, \eta_{i,K})$ at $x_{i,K}$ on $h_K^{-1}(\eta_{i,K})$ is identified with that of the inertia subgroup of $\pi_1(S_k \setminus T_k, \eta_{i,k})$ at $x_{i,k}$ on $h_k^{-1}(\eta_{i,k})$.

Choose a path on $S_k \setminus T_k$ from $\eta_{i,k}$ to s_k such that the induced bijection from $h_k^{-1}(\eta_{i,k})$ to $h_k^{-1}(s_k)$ is the specialization of the bijection from $h_K^{-1}(\eta_{i,K})$ to $h_K^{-1}(s_{\overline{K}})$ given by the path assumed in the proposition. Then specialization identifies the action of the inertia subgroup of $\pi_1(S_k \setminus T_k, s_k)$ at $x_{i,k}$ on $h_k^{-1}(s_k)$ with that of the inertia subgroup of $\pi_1(S_K \setminus T_K, s_{\overline{K}})$ at $x_{i,K}$ on $h_K^{-1}(s_{\overline{K}})$. Since the combined action of the inertia groups of $\pi_1(S_K \setminus T_K, s_{\overline{K}})$ at $x_{1,K}, \ldots, x_{m,K}$ on $h_K^{-1}(s_{\overline{K}})$ is transitive by assumption, the same is true with k replacing K. This proves the desired result. $\square$

We now return briefly to the characteristic zero setting. Define the *sheets* of the topological cover $\pi_0 : X_0(n)(\mathbb{C}) \to \mathbb{P}^1(\mathbb{C})$ to be the connected components of $\pi_0^{-1}(A)$, where $A := \mathbb{P}^1(\mathbb{C}) \setminus (\mathcal{M} \cup [0, \infty])$. Note that A is simply connected and π_0 is unramified above A, as the Mandelbrot set $\mathcal{M}$ contains all finite branch points of π_0. Fix a base point $a \in A$. For each branch point c of π_0, we fix a loop γ_c around c based at a as follows: take a path within A from a to a point very close to c. Then take a small loop around c (the direction does not matter), and then retrace the path back to a. The homotopy class of such a loop in $\mathbb{P}^1(\mathbb{C}) \setminus \{\text{branch locus of } \pi_0\}$ is well-defined up to inversion.

By Proposition 3.25, all ramification points of the topological cover $\pi_0 : X_0(n)(\mathbb{C}) \to \mathbb{P}^1(\mathbb{C})$ have ramification index two. This means that γ_c induces a transposition of the sheets of the cover for each branch point c. Note that this transposition does not depend on the base point a. Thus we make the following definition.

Definition 8.2. The *monodromy graph* for $X_0(n)$ is the graph $\Gamma(n)$ whose vertices correspond to the sheets of the topological branched cover $\pi_0 : X_0(n)(\mathbb{C}) \to \mathbb{P}^1(\mathbb{C})$ and whose edges correspond to branch points of π_0. The edge associated to a branch point c connects the two vertices corresponding to the sheets that are transposed by γ_c. An edge is *finite* if c is finite and is *infinite* otherwise.

For example, Figures 1 and 2 from Appendix A are the graphs of $\Gamma(5)$ and $\Gamma(7)$, respectively (the labeling of the vertices with 0's and 1's is explained in §9).

Corollary 8.3. Let p be an odd prime. Consider the cover $h : \mathfrak{X}_0(n)_R \to \mathbb{P}^1_R$, where R is a finite extension of $\mathbb{Z}_p^{ur}$ with fraction field K. Let T be the set of branch points of the generic fiber h_K of h, let $T' \subseteq T$ be those branch points whose specializations do not collide with that of some other branch point on the special fiber, and assume that all points in T' are K-rational. Identify the geometric generic fiber $h_K \times_K \overline{K}$ with an isomorphic map $h_{\mathbb{C}} : X_0(n)_{\mathbb{C}} \to \mathbb{P}^1_{\mathbb{C}}$ via a fixed isomorphism $\iota : \overline{K} \cong \mathbb{C}$, and correspondingly identify the set of branch points of $h_{\mathbb{C}}$ with T. If the monodromy graph $\Gamma(n)$ remains connected when all edges corresponding to points in $T \setminus T'$ are removed, then the special fiber of $\mathfrak{X}_0(n)_R$ is irreducible.

Proof. Pick a base point $a \in \mathbb{P}^1_{\mathbb{C}} \setminus (\mathcal{M} \cap [0, \infty]) \subseteq \mathbb{P}^1_{\mathbb{C}} \setminus T$, and pick loops γ_c based at a for all $c \in T$ as in the discussion before Definition 8.2. Since the monodromy graph is connected even when the edges corresponding to points of $T \setminus T'$ are removed, the monodromy action of the subgroup $G \subseteq \pi_1^{\text{top}}(\mathbb{P}^1(\mathbb{C}) \setminus T, a)$ generated by the loops γ_c for $c \in T'$ on $h_{\mathbb{C}}^{-1}(a)$ is transitive. The monodromy action, along with GAGA, gives a canonical injection $\pi_1^{\text{top}}(\mathbb{P}^1(\mathbb{C}) \setminus T, a) \to \pi_1(\mathbb{P}^1_{\mathbb{C}} \setminus T, a)$, which is the profinite completion. Under this injection, the

closure $\hat{G}$ of the image of G is topologically generated by the inertia subgroups of $\pi_1(\mathbb{P}^1_\mathbb{C} \setminus T, a)$ at the points of T', when paths to a are chosen appropriately. The action of $\hat{G}$ on $h_\mathbb{C}^{-1}(a)$ is transitive.

The same is therefore true for the action of $\hat{G} \subseteq \pi_1(\mathbb{P}^1_{\overline{K}} \setminus T, a)$ on $h_{\overline{K}}^{-1}(a)$, where we identify $\mathbb{C}$ with $\overline{K}$ via ι. In particular, $\hat{G}$ is topologically generated by the inertia groups at the points of T' for the appropriate paths. Since $h_{\overline{K}}$ and the branch points are defined over K, the same is true if we replace $\overline{K}$ with K. The corollary now follows from Proposition 8.1, which applies to $\mathfrak{X}_0(n)_R \to \mathbb{P}^1_R$ by Remark 6.1 and Propositions 6.2, 6.3, and 6.4. $\qquad\square$

Corollary 8.4. *If p is an odd prime such that $v_p(D_{n,n}) = 1$, then $Y_0(n)$ is geometrically irreducible modulo p. If $p \nmid n$, then $Y_1(n)$ is also geometrically irreducible modulo p.*

Proof. Let R be a finite extension of $\mathbb{Z}_p^{ur}$ with fraction field K such that all branch points of $\mathfrak{X}_0(n)_K = X_0(n)_K \to \mathbb{P}^1_K$ are K-rational. Recall that $\Delta_{n,n}(c)$ is defined over $\mathbb{Z}$ and its leading coefficient is a unit in $\mathbb{Z}_p$ (Theorem 3.17 and [MV95, Proposition 3.4]). By Lemma 7.5, there are exactly 2 roots of $\Delta_{n,n}(c)$ with the same residue in $\overline{\mathbb{F}}_p$, and these roots lie in R. Thus exactly 2 branch points of the generic fiber of $\mathfrak{X}_0(n)_R \to \mathbb{P}^1_R$ collide on the special fiber, and neither of these specializes to ∞. By Theorem 9.1 below (whose proof does not depend on anything from this section), the monodromy graph $\Gamma(n)$ is connected when the edges corresponding to the two colliding branch points are removed (under some identification of $\overline{K}$ with $\mathbb{C}$). By Corollary 8.3, the special fiber $\mathfrak{X}_0(n)_k$ of $\mathfrak{X}_0(n)_R$ is irreducible. Since $Y_0(n)_k$ is a dense open subset of $\mathfrak{X}_0(n)_k$ (Proposition 6.4(1)), the first statement of the corollary follows. The second statement follows from Proposition 6.15. $\qquad\square$

Example 8.5. Corollary 8.4 shows that $Y_0(5)$ is geometrically irreducible modulo every odd prime p (including $p = 3701$, the only prime of bad reduction, with $v_p(D_{5,5}) = 1$).

We conclude by noting that with additional knowledge of the characteristic 0 monodromy graph, one could deduce geometric irreducibility modulo p:

Proposition 8.6. *Write $v_p(D_{n,n}) = d$, and suppose the monodromy graph $\Gamma(n)$ is connected even when any $2d$ finite edges are removed. Then $X_0(n)$ is geometrically irreducible in characteristic p.*

Proof. Applying Corollary 5.2 to $\operatorname{Spec} R[c]/\Delta_{n,n}(c) \to \operatorname{Spec} R$, where R is as in Corollary 8.4, shows that there are at least $\deg \Delta_{n,n} - 2d$ roots of $\Delta_{n,n}(c)$ that are unique in their residue class in $\overline{\mathbb{F}}_p$. So at most $2d$ branch points of $\mathfrak{X}_0(n)_R \to \mathbb{P}^1_R$ can collide on the special fiber. If we could show that the monodromy graph $\Gamma(n)$ is connected, even when any $2d$ finite edges are removed, then the proof of Corollary 8.4 carries through to show that $X_0(n)$ is geometrically irreducible in characteristic p. $\qquad\square$

In other words, if we know that $\Gamma(n)$ remains connected after removing E finite edges, then $X_0(n)$ is geometrically irreducible in characteristic p whenever $v_p(D_{n,n}) \leq E/2$. A preliminary investigation for $n \leq 18$ indicates that excluding small n, one can take $E = n$ when n is odd and $E = n/2$ when n is even.

9. Connectedness of the monodromy graph

In this section, we consider only the family of quadratic polynomials $f_c(x) = x^2 + c$. The purpose of this section is to show that the monodromy associated to the cover

$$\pi_0 : X_0(n) \to \mathbb{P}^1$$

is highly transitive for all $n \in \mathbb{N}$. More precisely, the main result of the section is about the connectedness of the monodromy graph $\Gamma(n)$ from Definition 8.2.

Theorem 9.1. *For each positive integer n, the monodromy graph $\Gamma(n)$ remains connected after removing any two finite edges.*

We rely on the combinatorial description of the monodromy action for complex quadratic polynomials as developed by Lau and Schleicher [LS94, §3]; we refer the reader to their paper for more details. Our argument refines and expands the techniques of [BL14] by combining the combinatorial and the complex dynamical descriptions of quadratic polynomial dynamics. The methods of this section are independent of those used in the rest of the paper; the reader interested only in the result is directed to Theorem 9.1, proved at the end of the section.

9.1. **Itineraries.** Let $\Sigma := \{0, 1\}^{\mathbb{N}}$ be the space of binary sequences. This space comes equipped with a metric: if we write $v = v_1 v_2 v_3 \ldots$ and $w = w_1 w_2 w_3 \ldots$, then

$$d(v, w) = \sum_{i=1}^{\infty} \frac{|v_i - w_i|}{2^i}.$$

Note that $d(v, w) \leq 1/2^n$ if and only if v and w agree for the first n terms or agree for all but the n-th term. We define the *shift* map from Σ to itself to be the continuous function

$$\sigma(v_1 v_2 v_3 \ldots) = v_2 v_3 v_4 \ldots.$$

Now let $c \in \mathbb{C}$, and suppose that c does not lie in the Mandelbrot set $\mathcal{M}$. We may assign to each point α in the Julia set $\mathcal{J}_c$ of f_c an *itinerary*, which is an element $\iota(\alpha) \in \Sigma$ determined by the relative positions of the iterates of α in the Julia set $\mathcal{J}_c$. We direct the reader to [LS94, p. 8] for the definition of the itinerary in this setting. We do not use the definition explicitly in the current paper; however, we note that the itinerary is defined in such a way that the map $\alpha \mapsto \iota(\alpha)$ is a homeomorphism from $\mathcal{J}_c$ to Σ, and the map f_c restricted to $\mathcal{J}_c$ is conjugate to the shift map on Σ; that is, $\iota \circ f_c = \sigma \circ \iota$. (See [LS94, Theorem 3.1].) Therefore, much of the dynamics of f_c on $\mathcal{J}_c$ may be understood by studying the dynamics of the shift map on Σ.

For $c \in \mathbb{C} \setminus \mathcal{M}$, all periodic points for f_c lie in the Julia set $\mathcal{J}_c$. Therefore each periodic point α has a well-defined itinerary, and the period of α under f_c is precisely the period of $\iota(\alpha)$ under the shift map. If α has period n, we write

$$\iota(\alpha) = \overline{v_1 \ldots v_n}.$$

Let $W := \mathbb{C} \setminus (\mathcal{M} \cup [0, \infty))$; note that the map π_0 is unramified over W, so $\pi_0^{-1}(W) \to W$ is an étale cover. As c varies within W, the periodic points of f_c move continuously in $\mathbb{C}$. Buff and Lei use this fact to prove that the itinerary may be used to label the sheets of the cover $\pi_1^{-1}(W) \to W$, so there is a one-to-one correspondence between sequences in Σ of exact period n and sheets of the cover $X_1(n) \to \mathbb{P}^1$. Since $X_0(n)$ is the quotient of $X_1(n)$ by the cyclic group generated by the automorphism f_c, it follows that there is also a one-to-one correspondence between the sheets of $X_0(n)$ and the set of *orbits* of length n (under the shift map) in Σ. Two period-n elements $v, w \in \Sigma$ correspond to the same sheet of $X_0(n) \to \mathbb{P}^1$ if and only if $w = \sigma^k(v)$ for some integer k.

9.2. **Kneading sequences.** For an angle $\theta \in (0, 1)$, we define the *kneading sequence* of θ as follows: Partition $\mathbb{R}/\mathbb{Z}$ into three sets,

$$S_0 := \left(\frac{\theta + 1}{2}, \frac{\theta}{2} \right); \ \ S_1 := \left(\frac{\theta}{2}, \frac{\theta + 1}{2} \right); \ \ S_\star := \left\{ \frac{\theta}{2}, \frac{\theta + 1}{2} \right\}.$$

We assume $\mathbb{R}/\mathbb{Z}$ is oriented so that $0 \in S_0$. The *kneading sequence* of θ is the sequence of symbols $K(\theta) = v_1 v_2 v_3 \cdots \in \{0, 1, \star\}^{\mathbb{N}}$ defined so that $(2^{i-1}\theta \mod 1) \in S_{v_i}$ for all $i \geq 1$. The kneading sequence always begins with a 1, and $K(\theta)$ contains a $\star$ if and only if θ is periodic under doubling. If θ has exact period n under doubling, then

$$K(\theta) = \overline{v_1 \dots v_{n-1} \star},$$

where each of $v_1, \dots, v_{n-1}$ is either 0 or 1. In this case, we denote by $K_0(\theta)$ and $K_1(\theta)$ the sequences $\overline{v_1 \dots v_{n-1} 0}$ and $\overline{v_1 \dots v_{n-1} 1}$, respectively. For a periodic itinerary $v = \overline{v_1 \dots v_n}$, we let $v^\star$ denote the periodic sequence $v = \overline{v_1 \dots v_{n-1} \star}$. Thus $K(\theta) = K_0(\theta)^\star = K_1(\theta)^\star$.

Definition 9.2. We say an n-periodic itinerary is *maximal* if it is maximal (with respect to the lexicographic order) within its orbit under the shift map. Note that there is a unique shift (modulo n) which takes a given itinerary to a maximal itinerary.

Lemma 9.3. *Let* $v = \overline{v_1 \dots v_n}$ *be a maximal itinerary. Then* $v_n = 0$ *and the angle* θ *defined by the binary number* $.\overline{v_1 \dots v_n}$ *has kneading sequence* $K(\theta) = v^\star = \overline{v_1 \dots v_{n-1} \star}$.

Proof. See [BL14, Lemma 4.2]. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\quad\square$

Example 9.4. Consider the maximal itinerary $v = \overline{110}$, and set $\theta = 6/7$, which has binary expansion $\theta = .\overline{110}$. Then

$$\theta = 6/7 \in S_1; \quad 2\theta = 5/7 \in S_1; \quad 2^2\theta = 3/7 \in S_\star;$$

so the kneading sequence of θ is equal to $\overline{11\star}$.

9.3. **External rays of the Mandelbrot set.** Douady and Hubbard showed in [DH84] that there is a conformal isomorphism $\Phi : \mathbb{C} \setminus \mathcal{M} \to \mathbb{C} \setminus \overline{\mathbb{D}}$, where $\overline{\mathbb{D}}$ denotes the closed unit disk. An *external (parameter) ray* is the preimage $\mathcal{R}(\theta)$ of a ray $\{ re^{2\pi i\theta} : r > 1 \}$ under the map Φ. We say that the ray $\mathcal{R}(\theta)$ *lands* at the parameter $c \in \partial\mathcal{M}$ if

$$\lim_{r \to 1^+} \Phi^{-1}(re^{2\pi i\theta}) = c.$$

It is a well-known open conjecture that $\mathcal{R}(\theta)$ lands for all $\theta \in \mathbb{R}/\mathbb{Z}$. It was shown by Douady and Hubbard [DH84, Theorem 13.1] that $\mathcal{R}(\theta)$ lands if $\theta \in \mathbb{Q}/\mathbb{Z}$.

If $\theta \in \mathbb{Q}/\mathbb{Z}$ has odd denominator, then θ is periodic under the doubling map, and the ray $\mathcal{R}(\theta)$ lands at the root of a hyperbolic component of the same period. Conversely, if c is the root of a period-n hyperbolic component, then there are precisely two angles $\theta, \theta' \in \mathbb{Q}/\mathbb{Z}$ such that $\mathcal{R}(\theta)$ and $\mathcal{R}(\theta')$ land at c — unless $c = 1/4$, in which case only $\mathcal{R}(0) = \mathcal{R}(1)$ lands at c; see [DH84, §14.6]. The angles θ and θ' both have period n under the doubling map, and they have the same n-periodic kneading sequence by [Sch00, Lemma 3.9]. We may therefore assign a well-defined kneading sequence to any root c of a hyperbolic component of $\mathcal{M}$. As an example, we give in Appendix B the kneading sequences of all roots of period-5 hyperbolic components of $\mathcal{M}$ — or, equivalently, the kneading sequences of all angles $\theta \in \mathbb{Q}/\mathbb{Z}$ of exact period 5 under the shift map.

We now give an elementary lemma about kneading sequences that is used in §9.5.

Lemma 9.5. *Let $\theta \in (0,1)$. The kneading sequence $K(\theta)$ begins with 11 if and only if $\theta < 1/3$ or $\theta > 2/3$. In this case, if the ray $\mathcal{R}(\theta)$ lands, it does so at a non-real parameter.*

Proof. By the definition of the kneading sequence, the kneading sequences $K(\theta)$ and $K(-\theta)$ agree for any angle θ. To prove the first claim, therefore, it suffices to assume that $0 < \theta \leq 1/2$ and to show that $K(\theta)$ begins with 11 if and only if $\theta < 1/3$. Indeed, to say that the kneading sequence $K(\theta)$ begins with 11 is equivalent to saying that $\theta/2 < 2\theta < (\theta+1)/2$, which in turn is equivalent to saying that $0 < \theta < 1/3$. This proves the first claim.

To prove the second claim, we may once again assume that $0 < \theta < 1/3$ due to the symmetry of $\mathcal{M}$ across the real axis. Suppose $\mathcal{R}(\theta)$ lands at a parameter $c \in \partial\mathcal{M}$. For such θ, the rays $\mathcal{R}(0)$ and $\mathcal{R}(1/3)$, together with the main cardioid, separate c from the real line. Thus c is non-real, as claimed. $\qquad\square$

9.4. **The monodromy action of the cover $X_1(n) \to \mathbb{P}^1$.** Though our main result involves the monodromy associated to the cover $\pi_0 : X_0(n) \to \mathbb{P}^1$, we begin with the combinatorial description of the monodromy action for the cover $\pi_1 : X_1(n) \to \mathbb{P}^1$ appearing in [LS94, Lemmas 3.4 and 3.5]. First, we require a definition.

Definition 9.6. For $\varepsilon \in \{0,1\}$, let $\widehat{\varepsilon} = 1 - \varepsilon$. For an element $v = v_0 v_1 v_2 \ldots$, we define the *binary complement* of v to be the sequence

$$\widehat{v} = \widehat{v_0}\widehat{v_1}\widehat{v_2}\ldots.$$

The branch points of the cover $X_1(n) \to \mathbb{P}^1$ are the roots of the period-n hyperbolic components of $\mathcal{M}$ as well as the point at infinity, and the monodromy action associated to a loop around such a branch point c depends on whether $c = \infty$ or c is a *primitive* or *satellite* parabolic parameter. (Refer to Definition 3.18 for these terms.) Parts (1) and (2) of the following proposition are due to Lau and Schleicher, and part (3) is due to Blanchard, Devaney, and Keen; see Proposition 3.25 for other information about the branched cover π_1.

Proposition 9.7 ([LS94, Lemmas 3.4, 3.5]; [BDK91, Theorem 1.3])**.** *Let c be a branch point for the cover $\pi_1 : X_1(n) \to \mathbb{P}^1$. Making a small turn around c induces the following permutation of the sheets of the cover:*

(1) *Suppose c is the root of a primitive component of $\mathcal{M}$, with parameter ray θ landing at c. Then both $K_0(\theta)$ and $K_1(\theta)$ have exact period n, and a small loop around c interchanges the sheets labeled $K_0(\theta)$ and $K_1(\theta)$.*

(2) *Suppose c is a bifurcation point from a period-k component, with parameter ray θ landing at c. Write $n = qk$ for some integer $q \geq 2$, and let $e^{2\pi i p/q}$ be the multiplier of the parabolic orbit of f_c. (Note that p must be coprime to q.) Exactly one of $K_0(\theta)$ and $K_1(\theta)$ has period equal to n; let $K'(\theta)$ denote this period-n sequence. Then a small loop around c sends the sheet labeled $K'(\theta)$ to the sheet labeled with the kp'-fold shift of $K'(\theta)$, where $pp' \equiv 1 \pmod{q}$.*

(3) *Suppose $c = \infty$. Then a small turn around c transposes sheets whose representative itineraries are binary complements.*

Definition 9.8. An n-periodic kneading sequence K is *primitive* if both K_0 and K_1 have period n. Otherwise, K is *imprimitive*.

Recall from §9.1 that the sheets of $\pi_0 : X_0(n) \to \mathbb{P}^1$ are labeled by *orbits* of period-n binary sequences under the shift map; that is, two period-n binary sequences v and w represent the same sheet if and only if $w = \sigma^k(v)$ for some $k \in \mathbb{N}$. The only branch points for π_0

are the roots of *primitive* hyperbolic components of $\mathcal{M}$ and the point at infinity. Applying Proposition 9.7 immediately yields the following:

Corollary 9.9. *Let c be a branch point for $\pi_0 : X_0(n) \to \mathbb{P}^1$.*

(1) Suppose c is the root of a primitive component of $\mathcal{M}$, with parameter ray θ landing at c. Then both $K_0(\theta)$ and $K_1(\theta)$ have exact period n, and a small loop around c interchanges the sheets labeled $K_0(\theta)$ and $K_1(\theta)$.

(2) Suppose $c = \infty$. Then a small turn around c transposes sheets whose representative itineraries are binary complements.

In particular, a loop around a single branch point always induces a transposition.

9.5. **Structure of the monodromy graph.** Recall that $\Gamma(n)$ is the monodromy graph from Definition 8.2, associated to the cover $X_0(n) \to \mathbb{P}^1$.

As the vertices of $\Gamma(n)$ correspond to orbits of n-periodic itineraries, the number of vertices grows rapidly with n, with leading term $2^n/n$. The number of finite edges can also be computed, and it has leading term 2^{n-1}. However, the graph $\Gamma(n)$ is generally far from regular.

Recall that each sheet of $X_0(n) \to \mathbb{P}^1$ corresponds to an orbit of a period-n itinerary under the shift map. Thus it makes sense to label the vertices of $\Gamma(n)$ by period-n itineraries, with the understanding that two itineraries v and v' are associated to the same vertex $\mathfrak{v}$ if and only if $v' = \sigma^k(v)$ for some k. We call such an itinerary a *representative itinerary* for $\mathfrak{v}$. Note that each vertex of $\Gamma(n)$ has a unique maximal representative.

Remark 9.10. To distinguish vertices from their representatives, we use the typeface $\mathfrak{v}$ to denote a vertex of $\Gamma(n)$, reserving the standard typeface v for a representative itinerary. If $\mathfrak{v}$ is the vertex associated to the itinerary v, we denote by $\widehat{\mathfrak{v}}$ the binary complement of $\mathfrak{v}$; that is, the vertex associated to $\widehat{v}$.

As Figure 2 suggests, the monodromy graph has additional structure; we require some definitions to make that structure clear.

Definition 9.11. Let v be an n-periodic itinerary, and suppose that the first n letters of v contain A zeros and B ones. We define the *disparity* $D(v)$ of an n-periodic itinerary to be the difference $A - B$. The disparity $D(\mathfrak{v})$ of a vertex $\mathfrak{v}$ of the monodromy graph is defined to be the disparity of any representative itinerary.

Note the disparity of a vertex is well-defined independent of the choice of representative itinerary. As $\mathfrak{v}$ ranges over the vertices of $\Gamma(n)$, $D(\mathfrak{v})$ takes all values from $2 - n$ to $n - 2$ with the same parity as n. We also note that $D(\widehat{\mathfrak{v}}) = -D(\mathfrak{v})$.

Lemma 9.12. *The disparities of two vertices connected by a finite edge differ by 2.*

Proof. By Corollary 9.9, the edge corresponding to a given branch point c connects two vertices only if the kneading sequence for c has the form $K = \overline{v_1 \ldots v_{n-1}\star}$ and the associated itineraries $K_0 = \overline{v_1 \ldots v_{n-1}0}$ and $K_1 = \overline{v_1 \ldots v_{n-1}1}$ are representatives for the two vertices. The result is immediate. $\square$

We arrange the vertices into rows by decreasing disparity, with the unique vertex of maximal disparity $n - 2$ (represented, for example, by $\overline{11 \ldots 10}$) on the top row and the vertex of minimal disparity on the bottom row. The preceding lemma shows that there is only one possibility for any finite edge: it connects two vertices in adjacent rows.

Remark 9.13. There is one final structural feature of the graph worth describing. If $c \in \mathbb{C}$ is a branch point of π_0, then its complex conjugate $\bar{c}$ is also a branch point of π_0, and the angles of the rays landing at $\bar{c}$ are the negatives of those landing at c; in particular, c and $\bar{c}$ have the same kneading sequence. Therefore, edges corresponding to *non-real* branch points come in pairs.

Note that it is also possible for branch points which are not conjugate to have the same kneading sequence, accounting for the high multiplicity of some of the connections.

9.6. **Successor edges.** One may use Lemma 9.3 to show that, given any vertex $\mathfrak{v}$ in $\Gamma(n)$, there is at least one edge connected to $\mathfrak{v}$. This construction is highly useful for proving the existence of edges of the monodromy graph; for example, it is used by Buff and Lei [BL14] to prove that $Y_1(n)$ is connected for all n. Because of it's importance in this section, we give a name to this construction.

Definition 9.14. Let $v = \overline{v_1 \ldots v_n}$ be a maximal n-periodic itinerary, so that $v_n = 0$. Define the *successor* of v to be

$$s(v) := \overline{v_1 \ldots v_{n-1} 1}.$$

If v is non-maximal, define its successor $s(v)$ to be the successor of the shift of v which is maximal. Note that the successor increases the disparity by 2.

Lemma 9.15. *Let v be an itinerary of period $n \geq 3$ satisfying $D(v) < n - 2$. Then $s(v)$ also has period n, hence $v^\star = s(v)^\star$ is the kneading sequence of a primitive branch point of π_0. Consequently, the vertices $\mathfrak{v}$ and $s(\mathfrak{v})$ associated to v and $s(v)$, respectively, are connected by an edge in $\Gamma(n)$.*

Moreover, if $D(\mathfrak{v}) \geq 0$, then there are at least two edges of $\Gamma(n)$ which connect $\mathfrak{v}$ to $s(\mathfrak{v})$.

Because the edges constructed in Lemma 9.15 are so useful later in this section, we give them a name:

Definition 9.16. Let $\mathfrak{v}$ be a vertex in $\Gamma(n)$ with disparity $D(\mathfrak{v}) < n - 2$. If v is a representative itinerary for $\mathfrak{v}$, we define the *successor* of $\mathfrak{v}$ to be the vertex $s(\mathfrak{v})$ associated to the itinerary $s(v)$, and we call an edge $\mathfrak{v} \to s(\mathfrak{v})$ constructed as in Lemma 9.15 a *successor edge*.

Proof of Lemma 9.15. Without loss of generality, assume v is maximal, with

$$(9.1) \qquad v = \overline{\underbrace{1\ldots1}_{a_1} \underbrace{0\ldots0}_{b_1} \underbrace{1\ldots1}_{a_2} \cdots\cdots \underbrace{1\ldots1}_{a_r} \underbrace{0\ldots0}_{b_r}}.$$

By maximality, $a_1 \geq a_j$ for all $1 \leq j \leq r$. Since

$$(9.2) \qquad s(v) = \overline{\underbrace{1\ldots1}_{a_1} \underbrace{0\ldots0}_{b_1} \underbrace{1\ldots1}_{a_2} \cdots\cdots \underbrace{1\ldots1}_{a_r} \underbrace{0\ldots0}_{b_r-1} 1},$$

a right shift puts $s(v)$ into maximal form, with at least $a_1 + 1$ leading consecutive ones. Since $a_1 + 1 > a_j$ for all j, $s(v)$ has a unique block of 1's of length $a_1 + 1$. Combining this with the fact that $s(v)$ has at least one 0 (otherwise, v would have only one 0, making the disparity of v equal to $n - 2$), we conclude that $s(v)$ must also have period n, hence the kneading sequence $v^\star$ connecting v and $s(v)$ is primitive. By Lemma 9.3, it is the kneading sequence of a primitive branch point c, and the vertices $\mathfrak{v}$ and $s(\mathfrak{v})$ are connected by the edge corresponding to c.

If also $D(v) \geq 0$, then since $n > 2$ and v is maximal, v begins with at least two consecutive ones. By Lemma 9.5, the branch point c guaranteed by the previous paragraph is non-real, and its conjugate $\bar{c}$ provides a second edge connecting $\mathfrak{v}$ to $s(\mathfrak{v})$ in $\Gamma(n)$ by Remark 9.13.   □

Together with the edges provided by the branch point at infinity, we see that every vertex in the graph is connected to the vertex of maximal disparity via a path on which each *finite* edge occurs with multiplicity at least 2. In particular, $\Gamma(n)$ is connected (cf. [BL14, Lemma 4.2]).

9.7. Additional edges of $\Gamma(n)$.

Lemma 9.15 is insufficient to guarantee that $\Gamma(n)$ remains connected after two finite edges are removed. In this section, we construct additional edges of $\Gamma(n)$, which requires a more delicate analysis. We begin by describing two types of vertices that will require special attention later.

Definition 9.17. A vertex $\mathfrak{v}$ is *repetitive* if its maximal representative has the form

$$(9.3) \qquad v = \overline{\underbrace{1 \ldots 1}_{b+1} \underbrace{0 \ldots 0}_{a} \underbrace{1 \ldots 1}_{b} \cdots\cdots\cdots \underbrace{1 \ldots 1}_{b} \underbrace{0 \ldots 0}_{a}}$$

for some positive integers a and b.

Definition 9.18. A vertex $\mathfrak{v}$ of $\Gamma(n)$ is *self-complementary* if $\widehat{\mathfrak{v}} = \mathfrak{v}$. In other words, if v is a representative itinerary of $\mathfrak{v}$, then $\mathfrak{v}$ is self-complementary if there is an integer $1 \leq k < n$ such that $\sigma^k(v) = \widehat{v}$. Note that, in this case, n must be even and $D(\mathfrak{v})$ must be zero.

If a vertex $\mathfrak{v}$ is neither repetitive nor self-complementary, then it is connected to another vertex of equal disparity via a path that does not include a successor edge from $\mathfrak{v}$ to $s(\mathfrak{v})$.

Lemma 9.19. *Suppose $\mathfrak{v}$ is a vertex of $\Gamma(n)$ with disparity $2 - n < D(\mathfrak{v}) < n - 2$. If $\mathfrak{v}$ is not repetitive or self-complementary, then there exists a path in $\Gamma(n)$ connecting $\mathfrak{v}$ to a vertex $\mathfrak{w}$ with $D(\mathfrak{w}) = D(\mathfrak{v})$ and $\mathfrak{w} \neq \mathfrak{v}$. Furthermore, this path can be chosen disjoint from the edges connecting $\mathfrak{v}$ to its successor constructed in Lemma 9.15.*

Corollary 9.20. *If $n \geq 3$ and $\mathfrak{v}$ satisfies $D(\mathfrak{v}) < n - 2$ and is not repetitive or self-complementary, then $\mathfrak{v}$ is connected to a vertex of higher disparity via a path that does not include a successor edge $\mathfrak{v} \to s(\mathfrak{v})$.*

Proof. If $\mathfrak{v}$ has minimal disparity $2 - n < 0$, then $\mathfrak{v}$ is connected by an infinite edge to $\widehat{\mathfrak{v}}$, which has (maximal) disparity $n - 2 > 0$. Now suppose $2 - n < D(\mathfrak{v}) < n - 2$. By Lemma 9.19, $\mathfrak{v}$ is connected to a different vertex $\mathfrak{w}$ of the same disparity as $\mathfrak{v}$ via a path that does not include the successor edge from $\mathfrak{v}$. Following the successor edge $\mathfrak{w} \to s(\mathfrak{w})$ completes the desired path.   □

Proof of Lemma 9.19. To obtain the path, note that any vertex is connected to its binary complement by the branch point at infinity. Therefore, given a vertex $\mathfrak{v}$, the successor function together with the branch point at infinity provide a path

$$(9.4) \qquad \mathfrak{v} \xrightarrow{E_1} \widehat{\mathfrak{v}} \xrightarrow{E_2} s(\widehat{\mathfrak{v}}) \xrightarrow{E_3} \widehat{s(\widehat{\mathfrak{v}})} \xrightarrow{E_4} s(\widehat{s(\widehat{\mathfrak{v}})}),$$

which induces the following map on disparity:

$$D(\mathfrak{v}) \longrightarrow -D(\mathfrak{v}) \longrightarrow -D(\mathfrak{v}) + 2 \longrightarrow D(\mathfrak{v}) - 2 \longrightarrow D(\mathfrak{v}).$$

In this way we produce a path to a vertex of equal disparity.

We now claim that, for each $i \in \{1, 2, 3, 4\}$, the edge E_i is not a successor edge from $\mathfrak{v}$ (as constructed in Lemma 9.15). The edges E_1 and E_3 are infinite edges, so they are

not successor edges. If E_2 were a successor edge from $\mathfrak{v}$ to $s(\mathfrak{v})$, we would have $\mathfrak{v} = \widehat{\mathfrak{v}}$, contradicting our assumption that $\mathfrak{v}$ is not self-complementary. Finally, E_4 cannot be a successor edge $\mathfrak{v} \to s(\mathfrak{v})$ by considering the disparity. Thus the path given in (9.4) does not pass through a successor edge from $\mathfrak{v}$ to $s(\mathfrak{v})$.

To complete the proof, we must show that the final vertex $s(\widehat{s(\mathfrak{v})})$ in the path (9.4) is different from $\mathfrak{v}$. In other words, we must show that if v is a representative itinerary of $\mathfrak{v}$, then $s(\widehat{s(\widehat{v})})$ is not a shift of v. We choose a representative v of $\mathfrak{v}$ so that $\widehat{v}$ is maximal. Writing

$$v = \underbrace{0\ldots0}_{a_1} \underbrace{1\ldots1}_{b_1} \underbrace{0\ldots0}_{a_2} \cdots\cdots\cdots \underbrace{0\ldots0}_{a_r} \underbrace{1\ldots1}_{b_r},$$

we have

$$\widehat{v} = \underbrace{1\ldots1}_{a_1} \underbrace{0\ldots0}_{b_1} \underbrace{1\ldots1}_{a_2} \cdots\cdots\cdots \underbrace{1\ldots1}_{a_r} \underbrace{0\ldots0}_{b_r},$$

so that we have

$$\sigma\big(s(\widehat{v})\big) = \underbrace{1\ldots1}_{a_1+1} \underbrace{0\ldots0}_{b_1} \cdots\cdots\cdots \underbrace{1\ldots1}_{a_r} \underbrace{0\ldots0}_{b_r-1}.$$

Note that it is possible that $b_r = 1$ and $\sigma(s(\widehat{v}))$ ends in a 1. The complement of $\sigma(s(\widehat{v}))$ is

$$\widehat{\sigma\big(s(\widehat{v})\big)} = \underbrace{0\ldots0}_{a_1+1} \underbrace{1\ldots1}_{b_1} \underbrace{0\ldots0}_{a_2} \cdots\cdots\cdots \underbrace{0\ldots0}_{a_r} \underbrace{1\ldots1}_{b_r-1}.$$

Now let u be the maximal shift of $\widehat{\sigma\big(s(\widehat{v})\big)}$ and consider its successor $s(u)$; we claim that $s(u)$ does not represent $\mathfrak{v}$. As $\widehat{\sigma(s(\widehat{v}))}$ has at least $a_1 + 1$ consecutive zeros, so does u and its successor $s(u)$, unless the maximal shift u of $\widehat{\sigma(s(\widehat{v}))}$ begins with the length-b_1 block of ones. Since v has at most a_1 consecutive zeros by maximality of $\widehat{v}$, we need only deal with the case that the maximal shift of $\widehat{\sigma(s(\widehat{v}))}$ is

$$u = \underbrace{1\ldots1}_{b_1} \underbrace{0\ldots0}_{a_2} \cdots\cdots\cdots \underbrace{0\ldots0}_{a_r} \underbrace{1\ldots1}_{b_r-1} \underbrace{0\ldots0}_{a_1+1}.$$

In this case, $b_1 \geq b_j$ for $j \in \{1,\ldots,r-1\}$ and $b_1 \geq b_r - 1$, so the maximal shift of $s(u)$ is

$$\sigma^{n-1}\big(s(u)\big) = \underbrace{1\ldots1}_{b_1+1} \underbrace{0\ldots0}_{a_2} \cdots\cdots\cdots \underbrace{0\ldots0}_{a_r} \underbrace{1\ldots1}_{b_r-1} \underbrace{0\ldots0}_{a_1}.$$

As noted above, b_1 is maximal among b_j except possibly b_r, so v cannot contain $b_1 + 1$ consecutive ones unless $b_1 + 1 = b_r$, and the maximal shift of v is

$$\sigma^{n-b_r}(v) = \underbrace{1\ldots1}_{b_r} \underbrace{0\ldots0}_{a_1} \underbrace{1\ldots1}_{b_1} \underbrace{0\ldots0}_{a_2} \cdots\cdots\cdots \underbrace{0\ldots0}_{a_r},$$

If $s(u)$ represents $\mathfrak{v}$, then as both are maximal representatives, $\sigma^{n-1}(s(u)) = \sigma^{n-b_r}(v)$. Identifying the lengths of equal blocks, we see that $a_1 = a_2 = \cdots = a_r$ and $b_1 = b_2 = \cdots = b_{r-1} = b_r - 1$; since $D(v) > 2 - n$, the common value of the b_i is nonzero, so $\mathfrak{v}$ is repetitive, contradicting the hypotheses. $\qquad\square$

9.8. **Admissibility and repetitive vertices.** The path constructed in Lemma 9.19 may not succeed in producing a new endpoint for a repetitive vertex; for example, if $v = \overline{01011}$, then the path from (9.4)

$$\overline{01011} \to \overline{10100} \to \overline{11010} \to \overline{00101} = \overline{10100} \to \overline{11010},$$

and the final vertex is indeed a shift of the initial. (Note that we are representing the vertices in the path by itineraries.) Fortunately, the repetition in these vertices allows us to construct other edges. To do this, however, we will require the full machinery of admissibility for the first time. A more complete account should include the theory of Hubbard trees, but it suffices for us to use a combinatorial criterion for admissibility due to Bruin and Schleicher.

Denote by $\Sigma^\star$ the collection of sequences $v \in \{0,1\}^{\mathbb{N}}$, together with $\star$-periodic sequences, subject to the condition that $v_1 = 1$. (A sequence v is $\star$-*periodic* if it has the form $v = \overline{v_1 \ldots v_{n-1}\star}$ with $v_i \in \{0,1\}$).

Definition 9.21 ([BS08]). A sequence $v \in \Sigma^\star$ is *admissible* (resp., *real-admissible*) if it is the (necessarily $\star$-periodic) kneading sequence of a parabolic parameter $c \in \mathbb{C}$ (resp., $c \in \mathbb{R}$).

Definition 9.22. Given $v \in \Sigma^\star$, the ρ-*function* of v is the function $\rho_v : \mathbb{N} \to \mathbb{N} \cup \{\infty\}$ defined by

$$\rho_v(n) = \inf\{k > n : v_k \neq v_{k-n}\}.$$

The orbit of 1 under ρ_v is called the *internal address* of v. If the orbit of 1 contains ∞, it is standard convention to truncate the internal address just before the ∞ term.

As suggested by the terminology, if c is in the Mandelbrot set, the internal address gives a guide map to c in terms of a path through hyperbolic components of minimal period.

We now give a necessary and sufficient condition, due to Bruin and Schleicher, for an element of $\Sigma^\star$ to be admissible.

Theorem 9.23 ([BS08, Theorem 4.2]). *Given a $\star$-periodic sequence $v \in \Sigma^\star$ of period n, v is admissible if and only if m is an admissible period for all $1 \leq m < n$; that is, at least one of the following holds:*

(A1) *m is in the internal address of v.*
(A2) *there exists $k < m$ with $k \mid m$ and $\rho_v(k) > m$.*
(A3) *if $r \in \{1, \ldots, m\}$ is congruent to $\rho_v(m)$ modulo m, then m is not in the ρ_v-orbit of r.*

With this criterion for admissibility, we can now address the case of repetitive vertices.

Proposition 9.24. *Suppose $\mathfrak{v}$ is a repetitive vertex of $\Gamma(n)$. Then $\mathfrak{v}$ is connected to a vertex of higher disparity by a path which does not include a successor edge $\mathfrak{v} \to s(\mathfrak{v})$.*

Proof. The maximal representative of $\mathfrak{v}$ has the form of (9.3); that is,

$$v = \overline{\underbrace{1\ldots1}_{b+1} \underbrace{0\ldots0}_{a} \underbrace{1\ldots1}_{b} \cdots\cdots\cdots \underbrace{1\ldots1}_{b} \underbrace{0\ldots0}_{a}}$$

for some positive integers a and b.

First suppose that v has a single block of 1's. Consider the shift

$$\sigma(v) = \overline{\underbrace{1\ldots1}_{b} \underbrace{0\ldots0}_{a} 1}.$$

Replacing the final 1 of $\sigma(v)$ with a 0 yields the itinerary

$$u = \underbrace{1\ldots1}_{b} \; \underbrace{0\ldots0}_{a+1},$$

which is maximal in its orbit. If $\mathfrak{u}$ is the vertex corresponding to u, then this implies that $\mathfrak{v} = s(\mathfrak{u})$, so there is a disparity-*decreasing* edge from $\mathfrak{v}$ to $\mathfrak{u}$. Continuing in this way, one constructs a path from $\mathfrak{v}$ to the vertex of minimal disparity, represented by $\overline{10\ldots00}$, which is connected by an infinite edge to the vertex of maximal disparity.

We henceforth assume that v has at least two blocks of 1's, and we consider the shift w of v given by

$$w = \underbrace{1\ldots1}_{b} \; \underbrace{0\ldots0}_{a} \cdots\cdots\cdots \underbrace{1\ldots1}_{b} \; \underbrace{0\ldots0}_{a} \; \underbrace{1\ldots1}_{b+1} \; \underbrace{0\ldots0}_{a}.$$

If $w^\star$ is admissible, then there is an edge in $\Gamma(n)$ connecting the vertex $\mathfrak{v}$ associated to w (hence also to v) to the vertex $\mathfrak{u}$ associated to the itinerary

$$u = \underbrace{1\ldots1}_{b+1} \; \underbrace{0\ldots0}_{a} \cdots\cdots\cdots \underbrace{1\ldots1}_{b} \; \underbrace{0\ldots0}_{a} \; \underbrace{1\ldots1}_{b+1} \; \underbrace{0\ldots0}_{a-1}.$$

Note that u is primitive since it contains a single block of zeroes of length $(a-1)$, unless $a = 1$, in which case the maximal shift of u has a single block of 1's of length $2(b+1)$. Moreover, by comparing the sizes of the blocks of 0's and 1's, u and $s(v)$ do not represent the same vertex; hence there is a non-successor edge connecting $\mathfrak{v}$ to a vertex of higher disparity. It therefore remains to show that $w^\star$ is admissible, which we do by direct computation of the $\rho_{w^\star}$ function, abbreviated ρ here for clarity.

Let m be an integer with $1 < m < n$; we must show that m satisfies the three conditions of Theorem 9.23. First, because $w^\star$ is repetitive, it is relatively straightforward to show that the internal address of $w^\star$ is

$$1 \to (1+b) \to (2+b) \to \cdots \to (a+b) \to (n-a) \to (n-a+1) \to \cdots \to (n-1) \to n.$$

Suppose $1 \le m \le a+b$. If $m = 1$ or $m \ge b+1$, then m is in the internal address of $w^\star$, hence m satisfies (A1). On the other hand, if $2 \le m \le b$, then $\rho(1) = b+1 > m$, so m satisfies condition (A2).

Now suppose that $m \ge a+b+1$. If $m \ge n-a$, then m is in the internal address of $w^\star$, hence m again satisfies (A1). We now assume that $a+b+1 \le m < n-a$, in which case m is not in the internal address. If $(a+b)$ divides m, then $(a+b)$ is a *proper* divisor of m since $m > a+b$ by assumption. The fact that $\rho(a+b) = n-a > m$ then implies that m satisfies (A2). Finally, if $(a+b)$ does not divide m, then $r := \rho(m) - m < b < m$ satisfies $\rho(r) = b+1$; since m is not in the internal address, m cannot be in the orbit of r, so m satisfies condition (A3). $\qquad\square$

9.9. **Real admissibility and self-complementary vertices.** We use the Milnor-Thurston theory of real quadratic dynamics to construct additional edges for $\Gamma(n)$. Recalling the definition of real-admissibility from Definition 9.21, we have the following:

Theorem 9.25 ([MT88, Theorem 12.1]). *Let v be an n-periodic itinerary. Then there exists a shift w of v such that $w^\star$ is real-admissible.*

The reader should be aware that the notation and coordinates of [MT88] differ slightly from the current conventions (necessary for complex dynamics) that are used in this paper; in particular, their notion of a minimal shift, which corresponds to a real dynamical system,

is *not* the same as lexicographic minimality in our coordinates. However, the existence of a shift with an associated real-admissible kneading sequence is independent of these choices, and a conscientious reader can easily translate between the two notions if desired.

Unfortunately, Theorem 9.25 does not guarantee that every vertex of $\Gamma(n)$ has an adjacent edge associated to a real branch point. If $\mathfrak{v}$ is represented by an itinerary v for which $v^\star$ is real-admissible, it may be that $v^\star$ is imprimitive, in which case the real parabolic parameter with kneading sequence $v^\star$ is a *satellite* parabolic parameter. If this is the case, then we say v is *nearly imprimitive*, and write

$$(9.5) \qquad v = \overline{\underbrace{\mu \ldots \mu}_{k-1 \text{ times}} \mu'}$$

for some divisor $k > 1$ of n and some sequence μ of length $\ell = n/k$, where μ' changes the ℓth entry of μ. We now show that if v is self-complementary, then this problem cannot arise.

Proposition 9.26. *If $n > 2$ and v is self-complementary, then v is not nearly imprimitive.*

In order to prove the proposition, we require the following lemma.

Lemma 9.27. *Let v be a self-complementary itinerary of period n. Then n is even, and $v = \overline{w\widehat{w}}$ for some binary sequence w of length $n/2$.*

Proof. That n must be even follows immediately from the fact that a self-complementary itinerary has disparity zero.

Since v is self-complementary, there exists some integer k such that $\sigma^k(v) = \widehat{v}$; we may assume $0 < k < n$ since v has period n. Using the fact that the shift map σ commutes with binary complements, we therefore also have

$$v = \widehat{\widehat{v}} = \widehat{\sigma^k(v)} = \sigma^k(\widehat{v}) = \sigma^{2k}(v).$$

This implies that n divides $2k$, and since we assumed that $0 < k < n$, it follows that $k = n/2$.

Now write $v = \overline{wx}$, where w and x are binary sequences of length $n/2$. By the previous paragraph, $\overline{\widehat{w}\widehat{x}} = \widehat{v} = \sigma^{n/2}(v) = \overline{xw}$, so $x = \widehat{w}$, completing the proof. $\square$

Proof of Proposition 9.26. Suppose v is nearly imprimitive, written as in (9.5). If v is also self-complementary, then $D(v) = 0$, and as disparity is additive over concatenated strings,

$$0 = D(v) = (k-1)D(\mu) + D(\mu').$$

Since μ' and μ differ only at the ℓ-th digit, $D(\mu') = D(\mu) \pm 2$, hence

$$0 = kD(\mu) \pm 2.$$

It follows that $k = 2$ and $D(\mu) = \pm 1$. But if $k = 2$, then by Lemma 9.27 $\mu' = \widehat{\mu}$, and so $n/2 = \ell = 1$, contradicting our assumption that $n > 2$. $\square$

Corollary 9.28. *If $\mathfrak{v}$ is self-complementary, then there is a path connecting $\mathfrak{v}$ to a vertex of higher disparity which does not contain a successor edge from $\mathfrak{v}$.*

Proof. Let v be the representative itinerary for $\mathfrak{v}$ such that $v^\star$ is real-admissible, guaranteed by Theorem 9.25. As v is self-complementary, $D(v) = 0$ and Lemma 9.26 guarantees that v is not nearly imprimitive. The real parabolic parameter c with kneading sequence $v^\star$ therefore corresponds to an edge from $\mathfrak{v}$ to a vertex $\mathfrak{w}$.

If $D(\mathfrak{w}) > D(\mathfrak{v})$, then we are done, since the successor edges from $\mathfrak{v}$ correspond to complex branch points and this new edge corresponds to a real branch point. If $D(\mathfrak{w}) < D(\mathfrak{v})$, then

$D(\mathfrak{w}) = D(\mathfrak{v}) - 2 = -2$. Then $\mathfrak{w}$ is connected by the branch point at infinity to the vertex $\widehat{\mathfrak{w}}$ of disparity $2 > D(\mathfrak{v})$. $\qquad\square$

9.10. Connectivity of the monodromy graph. We may now prove the main result of this section.

Proof of Theorem 9.1. If $n \leq 4$, the graph $\Gamma(n)$ is easily computed and checked directly to satisfy the statement of the theorem, so assume $n > 4$. Let Γ' be a graph obtained from $\Gamma(n)$ by removing any two finite edges. We proceed by constructing a path from any vertex of Γ' to the vertex $\mathfrak{m}$ of maximal disparity, which is represented by

$$m = \underbrace{\overline{1 \ldots 1}\,0}_{n-1}.$$

If $\mathfrak{v}$ has negative disparity, then its complement $\widehat{\mathfrak{v}}$ has positive disparity, and there is an infinite edge connecting $\mathfrak{v}$ to $\widehat{\mathfrak{v}}$; we therefore assume $D(\mathfrak{v}) \geq 0$. By Lemma 9.15, repeated application of the successor function yields a path (in $\Gamma(n)$) from $\mathfrak{v}$ to $\mathfrak{m}$ where each edge has multiplicity at least two. Therefore, we may immediately conclude that every vertex is connected to $\mathfrak{m}$ in Γ' unless Γ' was obtained from $\Gamma(n)$ by the removal of a pair of successor edges coming from a single vertex $\mathfrak{v}_0$ of nonnegative disparity.

To complete the proof, it suffices to show that there is a path in $\Gamma(n)$ from $\mathfrak{v}_0$ to a vertex of higher disparity that avoids the successor edges from $\mathfrak{v}_0$ to $s(\mathfrak{v}_0)$. If $\mathfrak{v}_0$ is not repetitive of self-complementary, this is precisely Corollary 9.20; if $\mathfrak{v}_0$ is repetitive, this follows from Proposition 9.24; if $\mathfrak{v}_0$ is self-complementary, this is the content of Corollary 9.28. In all cases, therefore, we find that Γ' is connected. $\qquad\square$

10. Good reduction at primes dividing $2^n \pm 1$

In this section, we restrict to the case $f_c(x) = x^2 + c$. We provide a geometric explanation for good reduction of $Y_0(n)$ modulo a prime p such that p divides $2^n \pm 1$, even though p divides $D_{n,n}$. As an application, we verify certain examples of good reduction when $n = 5, 6$ and prove new cases of good reduction when $n = 7, 8, 11$.

Theorem 10.1. *The curve $Y_0(n)$ has good reduction modulo p in the following cases even though $p \mid D_{n,n}$:*
$n = 4$: $Y_0(4)$ for $p = 3, 5, 17$; [Mor98a, page 92]
$n = 5$: $Y_0(5)$ for $p = 3, 11, 31$; [FPS97, §7]
$n = 6$: $Y_0(6)$ for $p = 3, 5, 7, 13$; [Sto08, §2]
$n = 7$: $Y_0(7)$ for $p = 3, 43, 127$;
$n = 8$: $Y_0(8)$ for $p = 3, 5, 17, 257$;
$n = 11$: $Y_0(11)$ for $p = 3, 23, 89, 683$;

When n is prime and $p \nmid n$, then $Y_1(n)$ has good reduction modulo p if and only if $Y_0(n)$ does by Proposition 6.12. We computed that $Y_1(6)$ has bad reduction modulo $p = 3, 5$ and good reduction modulo $p = 7, 13$ and that $Y_1(8)$ has good reduction modulo $p = 3, 5, 17, 257$ (see Appendix A).

The strategy of the section is to study the reduction $\overline{Y}_0(n) := Y_0(n) \times_{\mathbb{Z}} \mathbb{F}_p$ above $c = 0$ and $c = -2$ modulo primes such that $p \mid (2^n \pm 1)$. (These values of c are special because they yield maps that come from algebraic groups, namely power and Chebyshev maps, respectively.) For particular n and p, we prove that:

(1) $\overline{Y}_0(n)$ has no singularity above $\bar{c} = 0$ and $\bar{c} = -2$, and
(2) the power of p dividing $D_{n,n}$ is fully explained by the points specializing to $\bar{c} = 0$ and $\bar{c} = -2$.

For part (1), we:

(i) determine the number of branch points which specialize to $c = 0$ and $c = -2$;
(ii) determine information about the degree of the discriminant above $\bar{c} = 0$ and $\bar{c} = -2$ in characteristic p; and
(iii) verify that the contribution to the degree of the discriminant above $c = 0$ and $c = -2$ is the same in characteristics zero and p.

For part (i), we need to compute $\Delta_{n,n}(c)$ modulo p. This step is the limiting factor for extending Theorem 10.1 to the case of larger n.

There are a few other cases that we expect can be solved using the techniques of this section, which we leave for an interested reader.

Question 10.2.

(1) *Does $Y_0(9)$ have good reduction modulo $p = 3, 7, 19, 73$?*
(2) *Does $Y_0(10)$ have good reduction modulo $p = 3, 5, 11, 33, 41$?*
(3) *Does $Y_0(13)$ have good reduction modulo $p = 3, 2731, 8191$?*

10.1. **Points of formal period n above $c = 0$ and $c = -2$.** We describe (the orbits of) the points of formal period n when $c = 0$ and $c = -2$ and their collision behavior modulo p.

Lemma 10.3.

(1) *The points x of formal period n above $c = 0$ occur at*

$$R_0 := \mu_{2^n - 1} \setminus \bigcup_{\substack{d|n \\ d \neq n}} \mu_{2^d - 1}.$$

(2) *The points x of formal period n above $c = -2$ occur at*

$$R_{-2} := \{\zeta + \zeta^{-1} : \zeta \in \mu_{2^n - 1} \cup \mu_{2^n + 1} \setminus \bigcup_{\substack{d|n \\ d \neq n}} \mu_{2^d - 1} \setminus \bigcup_{\substack{d|n \\ d \neq n}} \mu_{2^d + 1}\}.$$

Note that these $(2^n \pm 1)$-st roots of unity are not necessarily primitive. Note that $2^d + 1$ divides $2^n - 1$ if n/d even and divides $2^n + 1$ if n/d is odd.

Proof.

(1) If $c = 0$, then

$$\Phi_n(x, 0) = \prod_{d|n} (x^{2^d} - 1)^{\mu(n/d)}.$$

So $\Phi_n(x, 0)$ is the product of those cyclotomic polynomials $C_i(x)$ where i is a divisor of $2^n - 1$ but not of $2^d - 1$ for d a proper divisor of n.

(2) If $z \in \mathbb{C}^\times$, then $(z + z^{-1})^2 - 2 = z^2 + z^{-2}$. If $\zeta \in \mu_{2^n-1} \cup \mu_{2^n+1}$, then $\Psi_n(\zeta + \zeta^{-1}, -2) = 0$, i.e., $\zeta + \zeta^{-1}$ is an n-periodic point. If $\zeta + \zeta^{-1} = \eta + \eta^{-1}$, then $\eta = \zeta^{\pm 1}$. Thus there are 2^{n-1} numbers $\zeta + \zeta^{-1}$ with $\zeta \in \mu_{2^n-1}$ and $2^{n-1} + 1$ such numbers with $\zeta \in \mu_{2^n+1}$. Since $\mu_{2^n-1} \cap \mu_{2^n+1} = \{1\}$, $|\mu_{2^n-1} \cup \mu_{2^n+1}| = 2^n$, which is the degree of $\Psi_n(x, -2)$. Thus the roots of $\Psi_n(x, -2)$ are distinct and all arise in this way.

If $\zeta + \zeta^{-1}$ is also a d-periodic point, then $f_{-2}^d(\zeta + \zeta^{-1}) = \zeta + \zeta^{-1}$. The roots of $\Phi_n(x, -2)$ are the points of exact period n, which by induction are the values $\zeta + \zeta^{-1}$ such that ζ is not in any $\mu_{2^d-1} \cup \mu_{2^d+1}$ for $d \mid n$ and $d \neq n$.

$\square$

The points of $Y_1(n)$ above $c = 0$ and $c = -2$ are the points of R_0 and R_{-2}, respectively. The points of $Y_0(n)$ are the orbits of these under the C_n-action defined as follows: for $c = 0$, the action is $(1 \in C_n) \mapsto (\zeta \mapsto \zeta^2)$; for $c = -2$, the action is $(1 \in C_n) \mapsto (\zeta + \zeta^{-1} \mapsto \zeta^2 + \zeta^{-2})$. In other words, the action squares the root of unity in both cases.

If the reduction of $Y_1(n)$ modulo p has a singularity above $\bar{c} = 0$ (resp. $\bar{c} = -2$), then points of formal period n above $c = 0$ (resp. $c = -2$) collide modulo p. Let $\overline{R}_0$ and $\overline{R}_{-2}$ denote the reduction of R_0 and R_{-2}, respectively, modulo p.

Lemma 10.4. $|\overline{R}_0| < |R_0|$ (resp. $|\overline{R}_{-2}| < |R_{-2}|$) if and only if $p \mid (2^n - 1)$ (resp. $p \mid (2^n - 1)$ or $p \mid (2^n + 1)$).

Proof. The polynomial $x^k - 1$ is separable modulo p if and only if $p \nmid k$. Thus the k-th roots of unity are distinct in $\overline{\mathbb{F}}_p$ if and only if $p \nmid k$. It follows that if $p \nmid (2^n - 1)$, then $|\overline{R}_0| = |R_0|$.

If $p \mid (2^n - 1)$, then it suffices to show that there are two $(2^n - 1)$-st roots of unity *that are not $(2^d - 1)$-st roots for $d \mid n$ and $d < n$* that coalesce modulo p; that is, two elements of R_0 coalesce modulo p. To do this, let ζ be a primitive $(2^n - 1)$-st root of unity. Suppose that $2^n - 1 = p^k m$, where $p \nmid m$. Then $\zeta^i \equiv \zeta^j$ modulo some prime $\mathfrak{p}$ of $\mathbb{Z}[\zeta]$ above p if and only if $i \equiv j \pmod{m}$. Since $\nu(n) = |R_0| > m$, there must be two elements of R_0 coalescing in $\overline{\mathbb{F}}_p$ if $p \mid (2^n - 1)$.

For $c = -2$, then a similar argument applies, with the main difference being that $\zeta^i + \zeta^{-i} \equiv \zeta^j + \zeta^{-j} \pmod{\mathfrak{p}}$ if and only if $i \equiv \pm j \pmod{m}$, since $\zeta^i + \zeta^{-i} \equiv \zeta^j + \zeta^{-j}$ is equivalent to $\zeta^{-i}(\zeta^{i+j} - 1)(\zeta^{i-j} - 1) \equiv 0$. $\square$

10.2. **Degree of discriminant.** We now define four numbers which measure various contributions to the degree of the discriminant of the cover $\pi_0 : Y_0(n) \to \mathbb{A}^1$ and its reduction $\bar{\pi}_0$ modulo p. Let $\bar{\rho}_0$ (resp. $\bar{\rho}_{-2}$) be the contribution to the degree of the discriminant of the normalization of $\bar{\pi}_0$ over $\bar{c} = 0$ (resp. $\bar{c} = -2$). The number of ramification points of $\bar{\pi}_0$ above $\bar{c} = 0$ and $\bar{c} = -2$ is easy to compute combinatorially, but wild ramification often makes it harder to compute $\bar{\rho}_0$ and $\bar{\rho}_{-2}$.

Define ρ_0 (resp. ρ_{-2}) to be the degree of the discriminant of π_0 above the branch points which specialize to $\bar{c} = 0$ (resp. $\bar{c} = 2$) modulo p.

The values ρ_0 and $\bar{\rho}_0$ (resp. ρ_{-2} and $\bar{\rho}_{-2}$) determine whether or not $\overline{Y}_0(n)$ has a singularity above $\bar{c} = 0$ (resp. $\bar{c} = -2$).

Lemma 10.5. *With notation as above, $\bar{\rho}_0 \leq \rho_0$ and $\bar{\rho}_{-2} \leq \rho_{-2}$. The curve $\overline{Y}_0(n)$ is smooth above $\bar{c} = 0$ (resp. $\bar{c} = -2$) if and only if $\rho_0 = \bar{\rho}_0$ (resp. $\rho_{-2} = \bar{\rho}_{-2}$).*

Proof. Consider $\bar{c} = 0$; (the case $\bar{c} = -2$ is exactly the same). In the notation used in Proposition 5.3, $\rho_0 = d_\eta$ and $\bar{\rho}_0 \leq d_s$, with equality holding if and only if $\overline{Y}_0(n)$ is normal (that is, smooth) above $\bar{c} = 0$. Since $d_\eta = d_s$ (Proposition 5.3), it follows that $\rho_0 = d_\eta = d_s \geq \bar{\rho}_0$, with equality holding if and only if $\overline{Y}_0(n)$ is smooth above $\bar{c} = 0$. This proves the lemma. $\square$

We do not have a good method to compute ρ_0 and ρ_{-2} in general but the following lemma is useful to compute them in specific cases.

Lemma 10.6. *The values ρ_0 and ρ_{-2} are the integers such that there exists a polynomial $g(c) \in \mathbb{Z}[c]$ with $g(0)g(-2) \not\equiv 0 \pmod{p}$ and*

$$\Delta_{n,n}(c) \equiv c^{\rho_0}(c+2)^{\rho_{-2}} g(c) \bmod p.$$

Proof. By Proposition 3.25(4a), each branch point of π_0 has one ramification point η in its fiber, and η has ramification index 2. So ρ_0 equals the number of branch points which specialize to $\bar{c} = 0$. Thus ρ_0 is the integer such that $\Delta_{n,n}(c) \equiv c^{\rho_0} g_0(c) \bmod p$ with $g_0(0) \not\equiv 0 \bmod p$. Similarly, ρ_{-2} is the integer such that $\Delta_{n,n}(c) \equiv (c+2)^{\rho_{-2}} g_{-2}(c) \bmod p$ with $g_{-2}(-2) \not\equiv 0 \bmod p$. $\square$

10.3. **Other contributions to the discriminant.** We also use the polynomial $\Delta_{n,n}(c)$ to check that $\overline{Y}_0(n)$ is smooth everywhere, not just over the points $\bar{c} = 0, -2$. In order to guarantee that $Y_0(n)$ has good reduction at a prime $p \mid (2^n \pm 1)$, it is necessary to show that no points $\alpha \in \overline{\mathbb{F}}_p$ other than $0, -2$ contribute to the power of p in the discriminant $D_{n,n}$.

The points $\alpha \in \overline{\mathbb{F}}_p$ that contribute to the power of p in $D_{n,n}$ are exactly those α for which $(c - \alpha)^2 \mid (\Delta_{n,n}(c) \bmod p)$. These α can be found by factoring $\Delta_{n,n}(c)$ modulo p directly, but it is easier computationally to compute

(10.1) $$\Gamma_{n,p}(c) := \gcd(\Delta_{n,n} \bmod p, \Delta'_{n,n} \bmod p) \in \mathbb{F}_p[c].$$

Then $(c - \alpha)^2 \mid (\Delta_{n,n}(c) \bmod p)$ if and only if $(c - \alpha) \mid \Gamma_{n,p}(c)$.

If $\bar{c} = 0, -2$ are the only roots of $\Gamma_{n,p}(c)$, then the curve $\overline{Y}_0$ is smooth except possibly above $\bar{c} = 0, -2$. This is the outcome in all but one of the cases we computed. The one exception is $n = 6$, $p = 3$, when $\Gamma_{6,3} = c^5(c+2)^2(c^2 + 2c + 2)$. Hence, in this case, we must check that $\overline{Y}_0$ is smooth above the roots of $c^2 + 2c + 2$ as well, which we do in Example 10.16. Coincidentally, this is the only case we computed for which $p^2 \mid (2^n - 1)$. The next such example is when $n = 9$, $p = 3$.

10.4. **Degree of discriminant in characteristic p.** This section contains an exact value (resp. lower bound) for $\bar{\rho}_0$ and $\bar{\rho}_{-2}$ when the ramification is tame (resp. wild).

Consider the reduction $\overline{R}_0$ of R_0 modulo p. Let t_0 be the number of orbits of $\overline{R}_0$ under the action of C_n. For each orbit O_i, its ramification index e_i is the number of orbits in R_0 specializing to it. Then $\nu(n)/n = \sum_{i=1}^{t_0} e_i$. Similarly, let t_{-2} be the number of orbits of $\overline{R}_{-2}$ under the action of C_n; for each orbit O'_i, let e'_i denote the ramification index; then $\nu(n)/n = \sum_{i=1}^{t_{-2}} e'_i$.

The cover $\bar{\pi}_0$ is tamely ramified above $\bar{c} = 0$ (resp. $\bar{c} = -2$) when $p \nmid e_i$ for any orbit O_i (resp. $p \nmid e'_i$ for any orbit O'_i).

Lemma 10.7. *In the tame case, the degree of the discriminant above $\bar{c} = 0$ (resp. $\bar{c} = -2$) is $\bar{\rho}_0 = \sum(e_i - 1) = \nu(n)/n - t_0$ (resp. $\bar{\rho}_{-2} = \sum(e_i - 1) = \nu(n)/n - t_{-2}$).*

In the wild case, let s_0 (resp. s_{-2}) be the number of orbits O_i of $\overline{R}_0$ (resp. orbits O'_i of $\overline{R}_{-2}$) for which p divides the ramification index e_i (resp. e'_i).

Lemma 10.8. *In the wild case, the degree of the discriminant above $\bar{c} = 0$ (resp. $\bar{c} = -2$) satisfies*

$$\bar{\rho}_0 \geq \sum(e_i - 1) + s_0 = \nu(n)/n - t_0 + s_0,$$

(resp.

$$\bar{\rho}_{-2} \geq \sum(e_i - 1) + s_{-2} = \nu(n)/n - t_{-2} + s_{-2}).$$

Proof. At each ramification point above $\bar{c} = 0$, the contribution to the degree of the discriminant of $\bar{\rho}_0$ is an integer. At each wildly ramified point, the contribution is strictly greater than $e_i - 1$. $\qquad\square$

10.5. **Ramification indices modulo p.** In general, the combinatorics of the collisions depends on the divisors of n and $2^n - 1$ and their p-adic valuations. When n is prime, the ramification indices at $\bar{c} = 0$ and $\bar{c} = -2$ for $\bar{\varphi} : \overline{Y}_1(n) \to \mathbb{P}^1$ and $\bar{\pi}_0 : \overline{Y}_0(n) \to \mathbb{P}^1$ can be computed precisely.

Lemma 10.9. *Suppose n is an odd prime and $2^n \pm 1 = p^k m$ with $p \nmid m$. Then $\bar{\pi}_1 : \overline{Y}_1(n) \to \mathbb{P}^1$ has the following ramification indices at $\bar{c} = 0$ and $\bar{c} = -2$:*

(1) *If $p \mid (2^n - 1)$, then the fiber of $\bar{\pi}_1$ above 0 has m points, one with ramification index $p^k - 1$ and the other $m - 1$ with ramification index p^k.*

(2) *If $p \mid (2^n - 1)$, then the fiber of $\bar{\pi}_1$ above -2 has one point with ramification index $e = \frac{p^k-1}{2}$ and $\frac{m-1}{2}$ points with $e = p^k$, and the remaining points are unramified.*

(3) *If $p \mid (2^n + 1)$ and $p \neq 3$, then the fiber of $\bar{\pi}_1$ above -2 has one point with ramification index $e = \frac{p^k-1}{2}$ and one with $e = p^k - 1$ and $\frac{m-3}{2}$ points with $e = p^k$, and the remaining points are unramified.*

(4) *Let $n > 3$. If $p \mid (2^n + 1)$ and $p = 3$, then $k = 1$ (since n is prime), and the fiber of $\bar{\pi}_1$ above -2 has $\frac{m-1}{2}$ points with $e = 3$, and the remaining points are unramified.*

Proof.

(1) By Lemma 10.3(1), the points of $Y_1(n)$ above $c = 0$ occur when $x \in R_0$. Since n is prime, the elements of $R_0 = \mu_{2^n-1} \setminus \{1\}$ can be identified with pairs $(a, b) \in \mathbb{Z}/p^k\mathbb{Z} \times \mathbb{Z}/m\mathbb{Z} \setminus \{(0,0)\}$. The pairs (a_1, b_1) and (a_2, b_2) coalesce modulo p if and only if $b_1 = b_2$. Thus the ramification index is $e = p^k - 1$ when $b = 0$ and is $e = p^k$ otherwise.

(2) By Lemma 10.3(2), the points of $Y_1(n)$ above $c = -2$ occur when $x \in R_{-2}$. Since n is an odd prime, the elements of R_{-2} are of the form $x = \zeta + \zeta^{-1}$, where $\zeta \in (\mu_{2^n-1} \setminus \{1\}) \cup (\mu_{2^n+1} \setminus \mu_3)$.

If $p \mid (2^n - 1)$, then the $2^{n-1} - 1$ values $x = \zeta + \zeta^{-1}$ for $\zeta \in \mu_{2^n+1} \setminus \mu_3$ remain distinct modulo p. Similarly to part (1), the remaining elements can be identified with pairs $(a, b) \in \mathbb{Z}/p^k\mathbb{Z} \times \mathbb{Z}/m\mathbb{Z} \setminus \{(0,0)\}$ up to equivalence $(a, b) \equiv (-a, -b)$. Modulo p, two pairs (a_1, b_1) and (a_2, b_2) coalesce if and only if $b_1 = \pm b_2$. Thus the ramification index is $e = (p^k - 1)/2$ when $b = 0$ and is $e = p^k$ for the $(m - 1)/2$ choices of $\pm b \neq 0$.

(3) If $p \mid (2^n + 1)$, then the $2^{n-1} - 1$ values $x = \zeta + \zeta^{-1}$ with $\zeta \in \mu_{2^n-1} \setminus \{1\}$ remain distinct modulo p. If $p \neq 3$, the other points of period n can be identified with pairs $(a, b) \in \mathbb{Z}/p^k\mathbb{Z} \times \mathbb{Z}/m\mathbb{Z}$ such that $a \neq 0$ if b has order 1 or 3, up to equivalence $(a, b) \equiv (-a, -b)$. Modulo p, two pairs (a_1, b_1) and (a_2, b_2) coalesce if and only if $b_1 = \pm b_2$. Thus the ramification index is $e = (p^k - 1)/2$ when $b = 0$ and is $e = p^k - 1$ when b has order 3 and is p^k otherwise.

(4) If $p = 3 \mid (2^n + 1)$, then the points of period n of the form $\zeta + \zeta^{-1}$ where $\zeta \in \mu_{2^n+1} \cong \mathbb{Z}/3\mathbb{Z} \times \mathbb{Z}/m\mathbb{Z}$ can be identified with pairs $(a, b) \in \mathbb{Z}/3\mathbb{Z} \times \mathbb{Z}/m\mathbb{Z}$ such that $b \neq 0$, up to equivalence $(a, b) \equiv (-a, -b)$. Thus the ramification index is 3. The other points of period n remain distinct modulo p.

$\qquad\square$

Example 10.10. When $n = 5$ and $c = -2$, then $\zeta \in (\mu_{31} \setminus \{1\}) \cup (\mu_{33} \setminus \mu_3)$. When $p = 11$, there is no collision of the roots of unity in μ_{31} but there is collision of the roots of unity in μ_{33}. Above $\bar{c} = -2$ in $\overline{Y}_1(5)$, there is one point of ramification index 5, one point of ramification index 10, and 15 unramified points. The next result shows that above $\bar{c} = -2$ in $\overline{Y}_0(5)$, there is one point with ramification index 2 and 4 unramified points.

Lemma 10.11. *Suppose n is prime and $2^n \pm 1 = p^k m$ with $p \nmid m$. Then $\bar{\pi}_0 : \overline{Y}_0(n) \to \mathbb{A}^1$ has the following ramification indices at $\bar{c} = 0$ and $\bar{c} = -2$:*

(1) *If $p \mid (2^n - 1)$, then there are $1 + \frac{m-1}{n}$ points of $\overline{Y}_0(n)$ above $\bar{c} = 0$, one with ramification index $\frac{p^k-1}{n}$ and the other $\frac{m-1}{n}$ points with ramification index p^k.*

(2) *If $p \mid (2^n - 1)$, then there is one point in $\overline{Y}_0(n)$ above $\bar{c} = -2$ with ramification index $\frac{p^k-1}{2n}$, there are $\frac{m-1}{2n}$ points of ramification index p^k, and the remaining points are unramified.*

(3) *If $p \mid (2^n + 1)$ and $p \neq 3$, then there is one point in $\overline{Y}_0(n)$ above $\bar{c} = -2$ with ramification index $\frac{p^k-1}{2n}$, there is one point with ramification index $\frac{p^k-1}{n}$, there are $\frac{m-3}{2n}$ points with ramification index p^k, and the remaining points are unramified.*

(4) *If $p \mid (2^n + 1)$ and $p = 3$, then there are $\frac{m-1}{2n}$ points in $\overline{Y}_0(n)$ above $\bar{c} = -2$ with ramification index 3, and the remaining points are unramified.*

Proof. Let η be a ramification point of π_1 with ramification index ε. Let η' be the image of η in $\overline{Y}_0(n)$ and let ε' be its ramification index. Since n is prime, either $\varepsilon' = \varepsilon$ or $\varepsilon' = \varepsilon/n$. If η is the unique point with ramification index ε in a fiber of π_1 then $\varepsilon' = \varepsilon/n$. If $n \nmid \varepsilon$, then $\varepsilon' = \varepsilon$. Parts (1)–(4) then follow from Lemma 10.9). $\qquad\square$

Example 10.12. Let $n = 11$, then $R_0 = \mu_{2047} \setminus \{1\}$. Note that $R_0 \simeq \mathbb{Z}/23 \times \mathbb{Z}/89 \setminus \{(0,0)\}$, which has order $186 \cdot 11$, so $\nu(11)/11 = 186$. The doubling map has 2 orbits of length 11 on $(\mathbb{Z}/23)^\times$ and 8 orbits of length 11 on $(\mathbb{Z}/89)^\times$. So, R_0 has 8 orbits in $0 \times (\mathbb{Z}/89)^\times$ and 2 orbits in $(\mathbb{Z}/23)^\times$ and $176 = 8 \cdot 22$ other orbits containing pairs (a, b) where $ab \neq 0$.

For $p = 23$, we index the $176 + 8$ orbits by (a, b_0) where $a \in \mathbb{Z}/23$ and b_0 is the smallest value in one of the 8 orbits. Modulo 23, $\overline{Y}_0(11)$ has 8 points with wild ramification of order 23 (from the collapsing of a) and 1 point with ramification of order 2.

For $p = 89$, we index the $176 + 2$ orbits by (a_0, b) where $b \in \mathbb{Z}/89$ and a_0 is the smallest value in one of the two orbits. Modulo 89, $\overline{Y}_0(11)$ has 2 points with wild ramification of order 89 (from the collapsing of b) and 1 point with ramification of order 8.

10.6. **The tame case.** If n is prime and $p \mid (2^n \pm 1)$, then the ramification is tame in Lemma 10.11 items 1,2,4 exactly when $m = 1$ and in item 3 when $m = 3$. By Catalan's Conjecture, the condition $m = 1$ occurs if and only if n is prime and $p = 2^n - 1$ is a Mersenne prime or when $n = p = 3$. There are currently 49 known Mersenne primes.

When $m = 1$ and $p = 2^n - 1$, then there is a unique p^kth root of unity in $\overline{\mathbb{F}}_p$. So all $p^k - 1 = 2^n - 2$ points of $Y_1(n)$ above $c = 0$ collide modulo p. Thus $t_0 = 1$ and $e_0 = p^k - 1$. Since $p \nmid e_0$, ramification at $\bar{c} = 0$ of $\bar{\pi}_1$ is tame and so ramification at $\bar{c} = 0$ of $\bar{\pi}_0$ is also tame. The contribution to the degree of the discriminant of $\bar{\pi}_0$ is $\bar{\rho}_0 = \frac{p^k-1}{n} - 1 = \frac{\nu(n)}{n} - 1$.

10.7. **The wild case.** In the wild case, we expect that the lower bound in Lemma 10.8 for $\bar{\rho}_0$ (resp. $\bar{\rho}_{-2}$) typically equals the upper bound ρ_0 (resp. ρ_{-2}). In the data we computed, this was true except when $n = 6$ and $p = 7$. When the lower bound and upper bound are equal, then $\overline{Y}_0(n)$ is smooth above $\bar{c} = 0$ (resp. $\bar{c} = -2$).

Table 1. Reduction data for n prime. Here e denotes the ramification indices, either above $\bar{c} = 0$ or $\bar{c} = -2$; points with ramification index 1 are omitted, and exponents denote multiplicity. So, for example, $e = 23^8, 2$ means that there are 8 points with ramification index 23 and 1 point with ramification index 2, and all other points have ramification index 1.

n	p	ρ_0	$\bar{\rho}_0$	e	0 tame?	ρ_{-2}	$\bar{\rho}_{-2}$	e	-2 tame?	Other sing?
5	3	0	0	—	yes	3	3	3	no	no
5	11	0	0	—	yes	1	1	2	yes	no
5	31	5	5	6	yes	2	2	3	yes	no
7	3	0	0	—	yes	9	9	3^3	no	no
7	43	0	0	—	yes	7	7	6, 3	yes	no
7	127	17	17	18	yes	8	8	9	yes	no
11	3	0	0	—	yes	93	93	3^{31}	no	no
11	23	185	185	$23^8, 2$	no	92	92	23^4	no	no
11	89	185	185	$89^2, 8$	no	92	92	89, 4	no	no
11	683	0	0	—	yes	91	91	62, 31	yes	no

We proved that this equality holds under certain algebraic and numerical conditions. First, under certain numerical conditions on n and p and when $k = 1$, we proved a generalization of [Bou11, Proposition 1.3, Lemma 3.2] to show that the contribution to the degree of the discriminant at a wildly ramified point η for the *non-Galois* cover $\bar{\pi}_0$ is determined by the upper jump σ of the ramification filtration of a pre-image of η in the Galois closure of $\bar{\pi}_0$. Next, under a certain algebraic condition on $\zeta \in R_0$ modulo a prime of $\mathbb{Z}[\zeta]$ above p, we determined the upper jump σ for a point x of formal period n specializing to a wildly ramified point of $\bar{\pi}_0$. Specifically, we proved that the upper jump is as small as possible, namely $\sigma = 1/(p^k - 1)$. This proof relied on an algebraic analysis of valuations of differences of roots of $\Psi_n(x, c)$.

This material about the wild case is not included here since it is lengthy and not used in the proof of Theorem 10.1. The limiting factor in extending Theorem 10.1 to larger n is the difficulty in computing $\Delta_{n,n}$ modulo p. In Example 10.18, we briefly describe the unusual case $n = 6$ and $p = 7$.

10.8. **Proof of Theorem 10.1.** The proof of Theorem 10.1 is now a long calculation. We calculate $\Delta_{n,n}(c)$ modulo p. (See the ancillary file `delta_nn.txt` for the explicit polynomials $\Delta_{n,n}(c)$ for $n = 5, 6, 7, 8$, as well as $\Delta_{11,11}(c)$ modulo p for $p = 3, 23, 89, 683$.) Using Lemma 10.6, this determines ρ_0 and ρ_{-2}. Using (10.1), we check there is no singularity away from $\bar{c} = 0, -2$.

If $n \in \{5, 7, 11\}$, we use Lemma 10.11 to compute the ramification indices for $\overline{Y}_0(n)$ above $\bar{c} = 0, -2$ and determine if the ramification is tame or wild (if $n \in \{6, 8\}$ we make a similar one-off calculation). When tame, Lemma 10.7 determines $\bar{\rho}_0$ and $\bar{\rho}_{-2}$. When wild, we use Lemma 10.8 to determine a lower bound for $\bar{\rho}_0$ and $\bar{\rho}_{-2}$. We summarize these calculations in Table 1. By Lemma 10.5, there is good reduction for all pairs (n, p) in Table 1.

We give some details about the proof in a few sample cases.

Table 2. Reduction data for n composite

n	p	ρ_0	$\bar{\rho}_0$	e	0 tame?	ρ_{-2}	$\bar{\rho}_{-2}$	e	-2 tame?	Other sing?
6	3	6	6	4^2	yes	3	3	4	yes	$c^2 + 2c + 2, \rho = 2$
6	5	0	0	—	yes	5	5	5	no	no
6	7	9	9	$7, 2$	no	2	2	3	yes	no
6	13	0	0	—	yes	3	3	4	yes	no
8	3	30	30	3^{10}	no	12	12	3^4	no	no
8	5	30	30	5^6	no	12	12	$5^2, 2^2$	no	no
8	17	25	25	$8^3, 4, 2$	yes	11	11	$8, 4, 2$	yes	no
8	257	0	0	—	yes	15	15	16	yes	no

Example 10.13. Let $n = 5$ and $p = 31$. Then $\Delta_{5,5}(c) \equiv c^5(c+2)^2 g(c) \bmod 31$, where $g(0), g(-2) \not\equiv 0 \bmod 31$. Thus $\rho_0 = 5$ and $\rho_{-2} = 2$. Furthermore, $\Gamma_{5,31}(c) = c^4(c+2)$, so the branch points remain distinct modulo 31 unless they specialize to $\bar{c} = 0, -2$.

By Lemma 10.11 (1), $\bar{\pi}_0 : \overline{Y}_0(5) \to \mathbb{P}^1$ is tamely ramified at $\bar{c} = 0$, with ramification index $\frac{p-1}{n} = 6$. Thus $\bar{\rho}_0 = 5$. If $\bar{c} = -2$, then $\bar{\rho}_{-2} = \frac{30}{2 \times 5} - 1 = 2$. By Lemma 10.5, $\overline{Y}_0(5)$ has no singularity above $\bar{c} = 0$ or $\bar{c} = -2$ modulo 31. Thus $Y_0(5)$ has good reduction modulo 31.

Example 10.14. When $n = 7$ and $p = 127$, then $\Delta_{7,7}(c) \equiv c^{17}(c+2)^8 g(c) \bmod 127$, where $g(c)$ has simple roots not congruent to 0 or -2 modulo 127. Thus $\rho_0 = 17$ and $\rho_{-2} = 8$. Also $\bar{\rho}_0 = \frac{127-1}{7} - 1 = 17$ and $\bar{\rho}_{-2} = \frac{127-1}{2 \times 7} - 1 = 8$. So $Y_0(7)$ has good reduction modulo 127.

Example 10.15. Let $n = 7$ and $p = 43$. Then $p \mid (2^n + 1)$ with $m = 3$. Note that $43^{47} \mid D_{7,7}$. Above $\bar{c} = -2$, there is one point with ramification index 3 and one with ramification index 6 so $\bar{\rho}_{-2} = 7$. Also $\rho_{-2} = 7$, so $\overline{Y}_0$ is smooth above $c = -2$.

10.9. **Composite n.** A similar analysis is possible when n is composite. When $n = 6, 8$, we compute the data in Table 2. When $n = 6$, this gives a new proof of a result in [Sto08, §2].

10.9.1. *The case $n = 6$.* The degree of $\pi_0 : Y_0(6) \to \mathbb{A}^1$ is 9. By Proposition 6.5 and Tables 3–7, there are 9 odd primes dividing $D_{n,n}$, which are possibly primes of bad reduction for $Y_0(6)$. The methods of this section are useful for analyzing four of these, namely $p = 3, 5, 7, 13$.

By Lemma 10.3, the points of $Y_1(6)$ above $c = 0$ are $R_0 = \mu_{63} - (\mu_7 \cup \mu_3)$, so that $|R_0| = 54$. We identify the elements of R_0 with pairs $(a, b) \in \mathbb{Z}/9\mathbb{Z} \times \mathbb{Z}/7\mathbb{Z}$, excluding $(0, b)$ for $b \in \mathbb{Z}/7\mathbb{Z}$ and $(3, 0)$ and $(6, 0)$. The squaring map has 9 orbits over $c = 0$, represented by:

$$A = (1, 0), \quad B_i = (1, i) \text{ for } 1 \le i \le 6, \quad C_1 = (3, 1), \quad C_2 = (3, 5).$$

For C_1 and C_2, note that 1 is a quadratic residue and 5 is a quadratic nonresidue modulo 7.

Example 10.16. Let $p = 3$. Then (10.1) is $\Gamma_{6,3}(c) = c^5(c+2)^2(c^2 + 2c + 2)$. Modulo 3, at $\bar{c} = 0$, orbit A stays disjoint from the other orbits, orbit C_1 collides with the B_i orbits with $\left(\frac{i}{7}\right) = +1$, and orbit C_2 collides with the B_i orbits with $\left(\frac{i}{7}\right) = -1$. Thus the ramification type is $(4, 4, 1)$. A similar analysis shows that the ramification type at $\bar{c} = -2$ is $(4, 1^5)$.

There is an additional complication here, namely that there are points other than $c = 0$ and $c = -2$ which contribute to $v_3(D_{6,6})$, coming from the additional factor of $\Gamma_{6,3}$. These are the roots of $c^2 + 2c + 2$, where $\rho = 2$ in characteristic 0. To compute $\bar{\rho}$ at these points,

we factor the dynatomic polynomial $\Phi_6(x)$ over $\mathbb{F}_3[c]$ when c is a root of $c^2 + 2c + 2$:

$$(x^6 + 2cx^5 + x^4 + (c + 1)x^3 + (2c + 1)x^2 + (2c + 2)x + c + 1)^2 \cdot$$
$$(x^6 + (2c + 1)x^5 + (2c + 1)x^4 + (2c + 2)x^3 + 2cx^2 + 2cx + 2c + 2)^2 \cdot g(x),$$

where g has simple roots which do not collide with those of the other two factors listed. If x is a root of the first square factor, then $x^2 + c$ is a root of the second square factor, and vice versa. It follows that $\bar{\pi}_0$ has two points of ramification index 2 above each of these values of c. This matches the degree in characteristic 0, so $Y_0(6)$ has good reduction modulo $p = 3$.

Example 10.17. When $p = 5$, then $\bar{c} = 0$ is not a branch point. There is one ramified point above $\bar{c} = -2$, coming from the orbits of $\zeta + \zeta^{-1}$ with $\zeta \in \mu_{65}$, with $e' = 5$. Thus $\bar{\rho}_{-2} \geq 5$, and since $\rho_{-2} = 5$ and $\bar{\rho}_{-2} \leq \rho_{-2}$, it follows that $\bar{\rho}_{-2} = 5$. Thus $Y_0(6)$ has good reduction at $p = 5$. (On the other hand, $Y_1(6)$ has bad reduction modulo $p = 5$.)

Example 10.18. When $p = 7$, then $\Delta_{6,6}(c) \equiv c^9(c + 2)^2 h(c) \pmod{7}$. One can check that $h(c)$ is squarefree modulo 7 by computing $\Gamma_{6,7}(c)$ from (10.1). Thus it suffices to check that there is no singularity above $\bar{c} = 0$ and $\bar{c} = -2$.

When $p = 7$ and $c = 0$, then orbit A and all the B orbits collapse to a point, and the C orbits collapse to a different point. The ramification type is $(7, 2)$, so ramification is wild.

At the wildly ramified point, the minimal possible contribution to $\bar{\rho}_0$ is 7. However, the contribution is actually 8 because the algebraic condition mentioned in § 10.7 is not satisfied for some $\zeta \in R_0$, namely ζ a primitive 21-st root of unity. We compute that $\sigma = \frac{1}{3}$ so the wild contribution to the degree of the discriminant is $(p - 1)(1 + \sigma) = 8$. The tame contribution is 1. Thus $\bar{\rho}_0 = 9$, so $Y_0(6)$ has no singularity over $\bar{c} = 0$ when $p = 7$.

If $p = 7$ and $c = -2$ then $\rho_{-2} = 2$. The orbits from the (real parts of the) 65-th roots of unity do not collide modulo 7. The orbits above $c = -2$ coming from the 63-rd roots of unity are closely related to the orbits above $c = 0$, but the orbit of A above $c = 0$ is not included in R_{-2}. The six B orbits turn into three orbits in pairs, which collide. The two C orbits turn into one orbit. Thus there is one point with ramification index 3, yielding $\bar{\rho}_{-2} = 2$, as desired. Thus $Y_0(6)$ has good reduction modulo $p = 7$.

Example 10.19. Let $n = 6$ and $p = 13$. Then $\bar{c} = 0$ is unbranched. There is one ramified point above $\bar{c} = -2$, coming from $\zeta + \zeta^{-1}$ with $\zeta \in \mu_{65}$, with $e = 4$. Thus $\bar{\rho}_{-2} = 3$ and $Y_0(6)$ has good reduction modulo $p = 13$.

10.9.2. *The case $n = 8$.* The degree of $\pi_0 : Y_0(8) \to \mathbb{A}^1$ is $\frac{2^8 - 2^4}{8} = 30$. By Proposition 6.5 and Tables 3–7, there are many possible primes of bad reduction for $Y_0(8)$. We use the methods of this section to show that $Y_0(8)$ has good reduction modulo $p = 3, 5, 17, 257$ below.

When $n = 8$ and $p = 3$, there are ten points of $\overline{Y}_0(8)$ above $\bar{c} = 0$, each of which has $e = 3$. Thus all ramification points are wild, so $\bar{\rho}_0 \geq 3 \times 10 = 30$. Equality occurs because $\rho_0 = 30$, and $\bar{\rho}_0 \leq \rho_0$ by Lemma 10.5. Above $\bar{c} = -2$, there are six points coming from orbits of points of the form $\zeta + \zeta^{-1}$ with $\zeta \in \mu_{255}$. (There are others coming from $\zeta \in \mu_{257}$, but these are unramified and thus do not contribute to $\bar{\rho}_{-2}$.) Four of these points have $e = 3$, and the other two have $e = 1$. Thus $\bar{\rho}_{-2} \geq 4 \times 3 = 12$, and since $\rho_{-2} = 12$, equality occurs.

The case $p = 5$ is similar. Above $\bar{c} = 0$, there are six points, each of which has ramification index 5, so $\bar{\rho}_0 \geq 30$, and again equality occurs by comparison with ρ_0. Above $\bar{c} = -2$, there are four points coming from orbits of points of the form $\zeta + \zeta^{-1}$ with $\zeta \in \mu_{255}$, two with $e = 5$ and two with $e = 2$. Thus $\bar{\rho}_{-2} \geq 5 \times 2 + 1 \times 2 = 12$, and again there is equality.

Let $n = 8$ and $p = 17$. Then $\Delta_{8,8} \equiv c^{25}(c+2)^{11}g(c) \bmod 17$, where $g(c)$ has simple roots and $g(0), g(-2) \not\equiv 0 \pmod{17}$. Note that $R_0 = \mu_{2^8-1} \backslash \mu_{2^4-1} = (\mu_{17} \times \mu_5 \times \mu_3) \backslash (\{1\} \times \mu_5 \times \mu_3)$. The orbits of the squaring map on R_0 have lengths 8, 8, 8, 4, 2. Thus $\bar{\rho}_0 = 25$ and there is no singularity above $\bar{c} = 0$. When $\bar{c} = -2$, then all ramification again comes from those $\zeta + \zeta^{-1}$ with $\zeta \in \mu_{255}$. There are three points of this form above $\bar{c} = -2$, with ramification index 8, 4, and 2, for a contribution to $\bar{\rho}_{-2}$ of $7 + 3 + 1 = 11$.

When $n = 8$ and $p = 257$, then $\bar{c} = 0$ is unramified, but $\bar{c} = -2$ is ramified, with all contribution coming from $\zeta + \zeta^{-1}$ with $\zeta \in \mu_{257}$. There is only one ramification point above $\bar{c} = -2$, with ramification index 16, so $e = 16$ and $\bar{\rho}_{-2} = 15$.

References

[BDK91] Paul Blanchard, Robert L. Devaney, and Linda Keen. The dynamics of complex polynomials and automorphisms of the shift. *Invent. Math.*, 104(3):545–580, 1991.

[Ben05] Robert L. Benedetto. Heights and preperiodic points of polynomials over function fields. *Int. Math. Res. Not.*, (62):3855–3866, 2005.

[BL14] Xavier Buff and Tan Lei. The quadratic dynatomic curves are smooth and irreducible. In *Frontiers in complex dynamics*, volume 51 of *Princeton Math. Ser.*, pages 49–72. Princeton Univ. Press, Princeton, NJ, 2014.

[Bou92] T. Bousch. *Sur quelques problèmes de dynamique holomorphe*. PhD thesis, Université de Paris-Sud, Centre d'Orsay, 1992.

[Bou11] Irene I. Bouw. Covers of the affine line in positive characteristic with prescribed ramification. In *WIN—women in numbers*, volume 60 of *Fields Inst. Commun.*, pages 193–200. Amer. Math. Soc., Providence, RI, 2011.

[Bra89] Bodil Branner. The Mandelbrot set. In *Chaos and fractals (Providence, RI, 1988)*, volume 39 of *Proc. Sympos. Appl. Math.*, pages 75–105. Amer. Math. Soc., Providence, RI, 1989.

[BS08] Henk Bruin and Dierk Schleicher. Admissibility of kneading sequences and structure of Hubbard trees for quadratic polynomials. *J. Lond. Math. Soc. (2)*, 78(2):502–522, 2008.

[Can15] Jung Kyu Canci. Preperiodic points for rational functions defined over a rational function field of characteristic zero. *New York J. Math.*, 21:1295–1310, 2015.

[DH84] A. Douady and J. H. Hubbard. *Étude dynamique des polynômes complexes. Partie I*, volume 84 of *Publications Mathématiques d'Orsay [Mathematical Publications of Orsay]*. Université de Paris-Sud, Département de Mathématiques, Orsay, 1984.

[DP17] John R. Doyle and Bjorn Poonen. Gonality of dynatomic curves and strong uniform boundedness of preperiodic points. preprint. `arXiv:1711.04233`, 2017.

[FPS97] E. V. Flynn, Bjorn Poonen, and Edward F. Schaefer. Cycles of quadratic polynomials and rational points on a genus-2 curve. *Duke Math. J.*, 90(3):435–463, 1997.

[Ful69] William Fulton. Hurwitz schemes and irreducibility of moduli of algebraic curves. *Ann. of Math. (2)*, 90:542–575, 1969.

[GM98] Barry Green and Michel Matignon. Liftings of Galois covers of smooth curves. *Compositio Math.*, 113(3):237–272, 1998.

[GO14] Yan Gao and YaFei Ou. The dynatomic periodic curves for polynomial $z \mapsto z^d + c$ are smooth and irreducible. *Sci. China Math.*, 57(6):1175–1192, 2014.

[Liu02] Qing Liu. *Algebraic geometry and arithmetic curves*, volume 6 of *Oxford Graduate Texts in Mathematics*. Oxford University Press, Oxford, 2002. Translated from the French by Reinie Erné, Oxford Science Publications.

[LL99] Qing Liu and Dino Lorenzini. Models of curves and finite covers. *Compositio Math.*, 118(1):61–102, 1999.

[LS94] Eike Lau and Dierk Schleicher. Internal addresses in the mandelbrot set and irreducibility of polynomials. *SUNY Stony Brook Preprint*, 19, 1994.

[Méz00] Ariane Mézard. Fundamental group. In *Courbes semi-stables et groupe fondamental en géométrie algébrique (Luminy, 1998)*, volume 187 of *Progr. Math.*, pages 141–155. Birkhäuser, Basel, 2000.

[Mor96]   Patrick Morton. On certain algebraic curves related to polynomial maps. *Compositio Math.*, 103(3):319–350, 1996.

[Mor98a] Patrick Morton. Arithmetic properties of periodic points of quadratic maps. II. *Acta Arith.*, 87(2):89–102, 1998.

[Mor98b] Patrick Morton. Galois groups of periodic points. *J. Algebra*, 201(2):401–428, 1998.

[MS94]   Patrick Morton and Joseph H. Silverman. Rational periodic points of rational functions. *Internat. Math. Res. Notices*, (2):97–110, 1994.

[MT88]   John Milnor and William Thurston. On iterated maps of the interval. In *Dynamical systems (College Park, MD, 1986–87)*, volume 1342 of *Lecture Notes in Math.*, pages 465–563. Springer, Berlin, 1988.

[MV95]   Patrick Morton and Franco Vivaldi. Bifurcations and discriminants for polynomial maps. *Nonlinearity*, 8(4):571–584, 1995.

[Poo07]   Bjorn Poonen. Gonality of modular curves in characteristic p. *Math. Res. Lett.*, 14(4):691–701, 2007.

[Poo12]   Bjorn Poonen. Uniform boundedness of rational points and preperiodic points, 2012.

[Ray94]   M. Raynaud. Revêtements de la droite affine en caractéristique $p > 0$ et conjecture d'Abhyankar. *Invent. Math.*, 116(1-3):425–462, 1994.

[Sch00]   Dierk Schleicher. Rational parameter rays of the Mandelbrot set. *Astérisque*, (261):xiv–xv, 405–443, 2000. Géométrie complexe et systèmes dynamiques (Orsay, 1995).

[SGA03]   *Revêtements étales et groupe fondamental (SGA 1)*. Documents Mathématiques (Paris) [Mathematical Documents (Paris)], 3. Société Mathématique de France, Paris, 2003. Séminaire de géométrie algébrique du Bois Marie 1960–61. [Algebraic Geometry Seminar of Bois Marie 1960-61], Directed by A. Grothendieck, With two papers by M. Raynaud, Updated and annotated reprint of the 1971 original [Lecture Notes in Math., 224, Springer, Berlin; MR0354651 (50 #7129)].

[Sil07]   Joseph H. Silverman. *The arithmetic of dynamical systems*, volume 241 of *Graduate Texts in Mathematics*. Springer, New York, 2007.

[Sto08]   Michael Stoll. Rational 6-cycles under iteration of quadratic polynomials. *LMS J. Comput. Math.*, 11:367–380, 2008.

[Sza09]   Tamás Szamuely. *Galois groups and fundamental groups*, volume 117 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 2009.

[Xar12]   Xavier Xarles. Squares in arithmetic progression over number fields. *J. Number Theory*, 132(3):379–389, 2012.

Appendix A. Data

The following is a complete list of odd primes of bad reduction of $Y_1(n)$ with $f(x,c) = x^2 + c$ for $6 \leq n \leq 8$:

- $Y_1(6) : 3, 5, 67, 8029187$. (Note that $Y_0(6)$ has good reduction modulo 3, 5, and 67).
- $Y_1(7) : 7, 8456262122135977535818884167 2549561$.
- $Y_1(8) : 1567, 18967471, 120664513, 268015967, 17510504527986299347 84579, 48452315131278500437597584581$.

A.1. **Computation.** We would like to thank Michael Stoll for suggesting the following method to find candidates for primes of bad reduction. Let $F(x,c) = \Phi_n(x,c)$ and write F_x and F_c for its two partial derivatives. The primes of bad reduction are those primes p such that F, F_x, F_c vanish at a common point mod p. A necessary condition for this is that the three resultants

$$R_1 = \mathrm{res}_x(F, F_x), \quad R_2 = \mathrm{res}_x(F, F_c), \quad R_3 = \mathrm{res}_x(F_x, F_c)$$

have a common root mod p.

By Theorem 3.17, R_1 factors as $\pm\Delta_{n,n}(c)^n \prod_{\substack{d|n \\ d\neq n}} \Delta_{n,d}(c)^{n-d}$. In the cases that we computed, R_2 was irreducible. To find a list of candidate primes of bad reduction, we computed

$$\gcd(\operatorname{res}(f,g), \operatorname{res}(f,R_2))$$

for each irreducible factor f of R_3 and each irreducible factor g of R_1. Any bad prime $p > n$ must divide one of these greatest common divisors. For $n \leq 8$, the computations terminate in a reasonable time, giving a set L_n of candidate primes of bad reduction for $Y_1(n)$. One can then check for each prime $p \in L_n$ whether $Y_1(n)$ is singular mod p. For $5 \leq n \leq 8$, the set L_n consists of precisely the bad primes greater than n; on the other hand, the set L_4 contains 107, which is not a bad prime for $Y_1(4)$.

A.2. **Tables.** In Tables 3–7, we describe the factorizations of $D_{n,d} = \operatorname{disc}(\Delta_{n,d})$ and of $\operatorname{res}(\Delta_{n,d}, \Delta_{n,e})$. The notation (e,d) stands for $\operatorname{disc}(\Delta_{n,d})$ when $d = e$, and for $\operatorname{res}(\Delta_{n,d}, \Delta_{n,e})^2$ when $d \neq e$. The discriminants and resultants for $n \geq 7$ are large and could not be completely factored within a reasonable amount of time. All prime factors less than 10^{10} are listed, as well as some larger factors we were able to find. The odd bad primes for $X_0(n)$ are underlined, and the odd bad primes for $X_1(n)$ are in boldface.

Table 3. $n = 4$

(e,d)	Factorization
$(1,1)$	2^8
$(1,2)$	$2^{14} \cdot 5^2$
$(1,4)$	$2^{32} \cdot 5^2 \cdot 17^4$
$(2,2)$	1
$(2,4)$	$2^{16} \cdot 5^4$
$(4,4)$	$2^{16} \cdot 3^9$

Table 4. $n = 5$

(e,d)	Factorization
$(1,1)$	$2^{24} \cdot \mathbf{5}^7 \cdot 11^2$
$(1,5)$	$2^{232} \cdot 11^6 \cdot 31^{18} \cdot 86131^2$
$(5,5)$	$2^{274} \cdot 3^{12} \cdot 31^{27} \cdot \underline{\mathbf{3701}}^1 \cdot 4217^3$

Table 5. $n = 6$

(e,d)	Factorization
$(1,1)$	$2^4 \cdot \mathbf{3}^1$
$(1,2)$	$2^{20} \cdot \mathbf{3}^6 \cdot 13^2$
$(1,3)$	$2^{40} \cdot \mathbf{3}^4 \cdot 7^2$
$(1,6)$	$2^{192} \cdot \mathbf{3}^{12} \cdot 7^4 \cdot 13^6 \cdot 211^4 \cdot 68700493^2$
$(2,2)$	$2^4 \cdot \mathbf{3}^1$
$(2,3)$	$2^{28} \cdot \mathbf{3}^4 \cdot 157^2$
$(2,6)$	$2^{204} \cdot \mathbf{3}^{12} \cdot 7^{18} \cdot 13^{12} \cdot 79^2$
$(3,3)$	$2^{12} \cdot \mathbf{3}^4 \cdot \mathbf{5}^2 \cdot \mathbf{67}^1$
$(3,6)$	$2^{296} \cdot \mathbf{3}^{66} \cdot 7^6 \cdot 239^4 \cdot 409^2 \cdot 3331^2$
$(6,6)$	$2^{956} \cdot \mathbf{3}^{91} \cdot \mathbf{5}^{25} \cdot 7^{66} \cdot 13^8 \cdot 29^3 \cdot 61^2 \cdot \underline{\mathbf{8029187}}^1 \cdot 55218797^3 \cdot 47548578843011867^2$

Table 6. $n = 7$

(e, d)	Factorization
$(1, 1)$	$2^{60} \cdot \mathbf{7}^{11} \cdot 29^2 \cdot 43^2$
$(1, 7)$	$2^{1668} \cdot 43^{32} \cdot 127^{54} \cdot 987211^4$
	$\cdot \, 61701960674694323276265567036587732898819798393402690697954898729590901409 27^2$
$(7, 7)$	$2^{7712} \cdot 3^{81} \cdot 43^{47} \cdot 127^{351} \cdot 10273^2 \cdot 194003^2$
	$\cdot \underline{\mathbf{845626212213597753581888841672549561}}^1$
	$\cdots$ other factors (good primes)

Table 7. $n = 8$

| (e, d)	Factorization
$(1, 1)$	$2^{40} \cdot 3^2$
$(1, 2)$	$2^{100} \cdot 17^8$
$(1, 4)$	$2^{264} \cdot 7^8 \cdot 17^{16} \cdot 1321^4$
$(1, 8)$	$2^{4096} \cdot 17^{156} \cdot 257^{34} \cdot 593^4 \cdot 12073^8 \cdot 158091133929843713^4$
	$\cdot (\text{102-digit prime})^4$
$(2, 2)$	2^6
$(2, 4)$	$2^{132} \cdot 17^{16}$
$(2, 8)$	$2^{1028} \cdot 17^{152} \cdot 53^8 \cdot 248117^4$
	$\cdot 18205929282889572644368241959974766492471262272780397^4$
$(4, 4)$	$2^{80} \cdot 3^2 \cdot 5^2 \cdot 17^6 \cdot \mathbf{18967471}^1$
$(4, 8)$	$2^{3072} \cdot 3^{48} \cdot 5^{52} \cdot 17^{180} \cdot 22129^2$
	$\cdot 6200609^2 \cdot 7408189^2 \cdot 27756089^4 \ldots \text{other factors}$
$(8, 8)$	$2^{27528} \cdot 3^{1044} \cdot 5^{1052} \cdot 17^{738} \cdot 257^{224} \cdot \underline{\mathbf{1567}}^1 \cdot 17863^2 \cdot 1733999^2 \cdot 8711621^3$
	$\cdot 141715459^2 \cdot \underline{\mathbf{120664513}}^1 \cdot \underline{\mathbf{268015967}}^1 \cdot \underline{\mathbf{17510504527986299347 84579}}^1$
	$\cdot \underline{\mathbf{4845231513127850043759 7584581}}^1 \cdots \text{other factors (good primes)}$

Appendix B. Monodromy data for $n = 5$

In Table 9, we describe the monodromy data at the branch points of π_1 for $Y_1(5)$. As mentioned in §9, sheets of the cover π_1 may be identified with sequences in $\{0, 1\}^{\mathbb{N}}$ of period 5; for convenience, we order the sheets 1 through 30 according to Table 8. The monodromy action at each branch point is then obtained from its kneading sequences via Proposition 9.7.

There are fifteen branch points for π_1, and there are exactly two parameter rays that land at each branch point. We therefore use the landing rays to differentiate the branch points in Table 9. The "bifurcation type" of a branch point c is $(5, 1)$ or $(5, 5)$ according to whether c is a satellite parabolic parameter (in which case a 5-cycle collapses to a 1-cycle) or a primitive parabolic parameter (in which case two 5-cycles collide). The $(5, 5)$-branch points are precisely the finite branch points of $\pi_0 : Y_1(5) \to \mathbb{A}^1$, and the data in Table 9 associated to such branch points is used to construct the monodromy graph $\Gamma(5)$ in Figure 1.

Though we do not give the analogous data for $n = 7$, we illustrate $\Gamma(7)$ in Figure 2.

Table 8. Ordering of sheets for $n = 5$

| 1	2	3	4	5
$\overline{00001}$	$\overline{00010}$	$\overline{00100}$	$\overline{01000}$	$\overline{10000}$
6	7	8	9	10
$\overline{00011}$	$\overline{00110}$	$\overline{01100}$	$\overline{11000}$	$\overline{10001}$
11	12	13	14	15
$\overline{00101}$	$\overline{01010}$	$\overline{10100}$	$\overline{01001}$	$\overline{10010}$
16	17	18	19	20
$\overline{00111}$	$\overline{01110}$	$\overline{11100}$	$\overline{11001}$	$\overline{10011}$
21	22	23	24	25
$\overline{01011}$	$\overline{10110}$	$\overline{01101}$	$\overline{11010}$	$\overline{10101}$
26	27	28	29	30
$\overline{01111}$	$\overline{11110}$	$\overline{11101}$	$\overline{11011}$	$\overline{10111}$
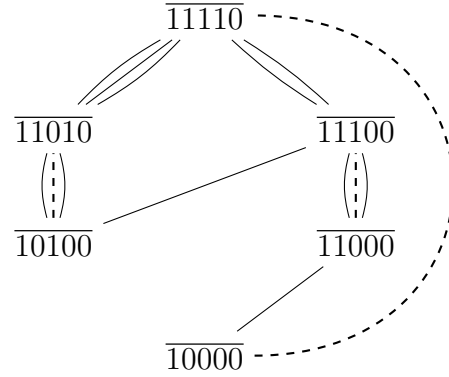
Figure 1. The monodromy graph $\Gamma(5)$. Finite edges are solid, and infinite edges are dashed.

Table 9. Branch points for $n = 5$

Landing rays ($\theta = \square/31$)	$K(\theta)$	Bifurcation type	Multiplier ($\lambda = e^{2\pi i \square}$)	Cycle structure
1, 2	1111∗	(5,1)	1/5	(26 27 28 29 30)
3, 4	1110∗	(5,5)	1	(16 26)(17 27)(18 28)(19 29)(20 30)
5, 6	1101∗	(5,5)	1	(21 26)(22 27)(23 28)(24 29)(25 30)
7, 8	1100∗	(5,5)	1	(6 16)(7 17)(8 18)(9 19)(10 20)
9, 10	1111∗	(5,1)	2/5	(26 29 27 30 28)
11, 12	1010∗	(5,5)	1	(11 23)(12 24)(13 25)(14 21)(15 22)
13, 18	1011∗	(5,5)	1	(21 29)(22 30)(23 26)(24 27)(25 28)
14, 17	1001∗	(5,5)	1	(11 16)(12 17)(13 18)(14 19)(15 20)
15, 16	1000∗	(5,5)	1	(1 6)(2 7)(3 8)(4 9)(5 10)
19, 20	1010∗	(5,5)	1	(11 23)(12 24)(13 25)(14 21)(15 22)
21, 22	1111∗	(5,1)	3/5	(26 28 30 27 29)
23, 24	1100∗	(5,5)	1	(6 16)(7 17)(8 18)(9 19)(10 20)
25, 26	1101∗	(5,5)	1	(21 26)(22 27)(23 28)(24 29)(25 30)
27, 28	1110∗	(5,5)	1	(16 26)(17 27)(18 28)(19 29)(20 30)
29, 30	1111∗	(5,1)	4/5	(26 30 29 28 27)
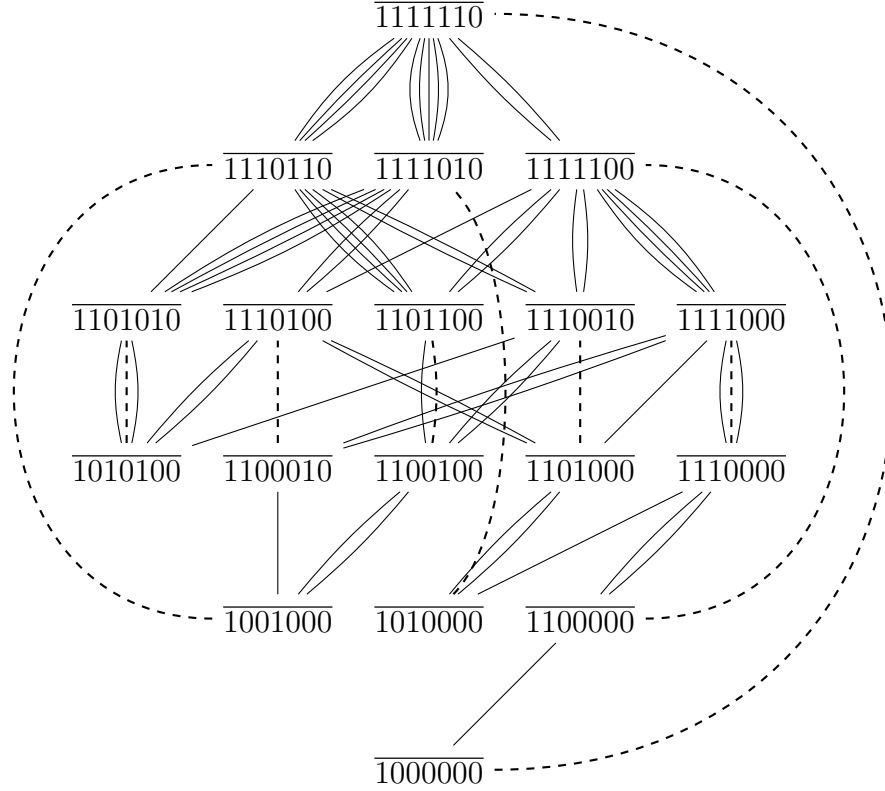
Figure 2. The monodromy graph $\Gamma(7)$. Finite edges are solid, and infinite edges are dashed.

JOHN R. DOYLE: UNIVERSITY OF ROCHESTER, ROCHESTER, NY 14627, USA
Current address: Louisiana Tech University, Ruston, LA 71272
E-mail address: jdoyle@latech.edu

HOLLY KRIEGER: UNIVERSITY OF CAMBRIDGE, CAMBRIDGE, CB3 0WB, ENGLAND
E-mail address: hkrieger@dpmms.cam.ac.uk

ANDREW OBUS: UNIVERSITY OF VIRGINIA, CHARLOTTESVILLE, VA 22904, USA
E-mail address: obus@virginia.edu

RACHEL PRIES: COLORADO STATE UNIVERSITY, FORT COLLINS, CO 80523, USA
E-mail address: pries@math.colostate.edu

SIMON RUBINSTEIN-SALZEDO: EULER CIRCLE, PALO ALTO, CA 94306, USA
E-mail address: simon@eulercircle.com

LLOYD WEST: UNIVERSITY OF VIRGINIA, CHARLOTTESVILLE, VA 22904, USA
E-mail address: lww8k@virginia.edu