Journal de Théorie des Nombres de Bordeaux

anciennement Séminaire de Théorie des Nombres de Bordeaux

Andrew OBUS

Fields of moduli of three-point G-covers with cyclic p-Sylow, II

Tome 25, n° 3 (2013), p. 579-633.

<http://jtnb.cedram.org/item?id=JTNB_2013__25_3_579_0>

Journal de Théorie des Nombres
de Bordeaux **25** (2013), 579–633

Fields of moduli of three-point G-covers with cyclic p-Sylow, II

par Andrew OBUS

Résumé. Nous poursuivons l'étude de la réduction stable et des corps de modules des G-revêtements galoisiens de la droite projective sur un corps discrètement valué de caractéristique mixte $(0, p)$, dans le cas où G a un p-sous-groupe de Sylow cyclique d'ordre p^n. Supposons de plus que le normalisateur de P agit sur lui-même via une involution. Sous des hypothèses assez légères, nous montrons que si $f : Y \to \mathbb{P}^1$ est un G-revêtement galoisien ramifié au-dessus de 3 points, défini sur $\overline{\mathbb{Q}}$, alors les n-ièmes groupes de ramification supérieure au-dessus de p, en numérotation supérieure, de (la clôture galoisienne de) l'extension $K/\mathbb{Q}$ sont triviaux, où K est le corps des modules de f.

Abstract. We continue the examination of the stable reduction and fields of moduli of G-Galois covers of the projective line over a complete discrete valuation field of mixed characteristic $(0, p)$, where G has a *cyclic* p-Sylow subgroup P of order p^n. Suppose further that the normalizer of P acts on P via an involution. Under mild assumptions, if $f : Y \to \mathbb{P}^1$ is a three-point G-Galois cover defined over $\overline{\mathbb{Q}}$, then the nth higher ramification groups above p for the upper numbering of the (Galois closure of the) extension $K/\mathbb{Q}$ vanish, where K is the field of moduli of f.

1. Introduction

1.1. Overview. This paper continues the work of the author in [12] about ramification of primes of $\mathbb{Q}$ in fields of moduli of three-point G-Galois covers of the Riemann sphere. We place bounds on the ramification of the prime p when a p-Sylow subgroup P of G is cyclic of arbitrary order (Theorem 1.1). This was done in [1] for P trivial, and in [23] for $|P| = p$. In [23], Wewers used a detailed analysis of the stable reduction of the cover to characteristic

Manuscrit reçu le 20 juin 2010.
The author was supported by a NDSEG Graduate Research Fellowship and an NSF Postdoctoral Research Fellowship in the Mathematical Sciences. Some editing was done while a guest at the Max-Planck-Institut Für Mathematik in Bonn. Final submission occurred while the author was supported by NSF grant DMS-1265290.
Mots clefs. field of moduli, stable reduction, Galois cover.
Classification math. 14G20, 11G22, 14H30, 14H25, 14G25, 11G20, 11S20.

p, inspired by results of Raynaud ([19]). However, many of the results on stable reduction in the literature (particularly, those in the second half of [19]) are only applicable when $|P| = p$.

In [15], the author generalized much of [19] to the case where P is cyclic of any order. In [12], these results were applied under the additional assumption that G is *p-solvable* (i.e., has no nonabelian simple composition factors with order divisible by p) to place bounds on ramification of p in the field of moduli of a three-point G-cover. In this paper, we drop the assumption of p-solvability, but we assume that the normalizer of a p-Sylow subgroup P acts on P via an involution. This hypothesis is satisfied for many non-p-solvable groups (Remark 1.3 (1)), and was used as a simplifying assumption in [2] (in the case where $|P| = p$) to examine the reduction of four-point G-covers of $\mathbb{P}^1$. It will simplify matters in our situation as well.

Let $f : Y \to X \cong \mathbb{P}^1_{\mathbb{C}}$ be a finite, connected, G-Galois branched cover of Riemann surfaces, branched only at $\mathbb{Q}$-rational points. Such a cover can always be defined over $\overline{\mathbb{Q}}$. If there are exactly three branch points (without loss of generality, 0, 1, and ∞), such a cover is called a *three-point cover*. The fixed field in $\overline{\mathbb{Q}}$ of all elements of $\mathrm{Aut}(\overline{\mathbb{Q}}/\mathbb{Q})$ fixing the isomorphism class of f as a G-cover (i.e., taking into account the G-action) is a number field called the *field of moduli* of f (as a G-cover). By [3, Proposition 2.7], it is also the intersection of all fields of definition of f (along with the G-action). For more details, see, e.g., [3] or the introduction to [12].

Since a branched G-Galois cover $f : Y \to X$ of the Riemann sphere is given entirely in terms of algebraic data (the branch locus C, the Galois group G, and an element $g_i \in G$ for each $c_i \in C$ such that $\prod_i g_i = 1$ and the g_i generate G), it is reasonable to try to draw inferences about the field of moduli of f based on these data. But this is a deep question, as the relation between topology of covers and their defining equations is given by [20], where the methods are non-constructive.

1.2. Main result. Let $f : Y \to X = \mathbb{P}^1$ be a three-point G-cover defined over $\overline{\mathbb{Q}}$. If $p \nmid |G|$, then p is unramified in the field of moduli of f ([1]). If G has a p-Sylow group of order p (thus cyclic), then p is tamely ramified in the field of moduli of f ([23]). Furthermore, if G has a *cyclic* p-Sylow subgroup of order p^n and is p-solvable (i.e., has no nonabelian simple composition factors with order divisible by p), then the nth higher ramification groups above p for the upper numbering of (the Galois closure of) $K/\mathbb{Q}$ vanish, where K is the field of moduli of f ([12]). Our main result extends this to many non-p-solvable groups.

If H is a subgroup of G, we write $N_G(H)$ for the normalizer of H in G and $Z_G(H)$ for the centralizer of H in G.

Theorem 1.1. *Let $f : Y \to X$ be a three-point G-Galois cover of the Riemann sphere, and suppose that a p-Sylow subgroup $P \leq G$ is cyclic of order p^n. Suppose $|N_G(P)/Z_G(P)| = 2$. Lastly, suppose that at least one of the three branch points has prime-to-p branching index (if exactly one, then we require $p \neq 3$). If $K/\mathbb{Q}$ is the field of moduli of f, then the nth higher ramification groups for the upper numbering of the Galois closure of $K/\mathbb{Q}$ vanish.*

Remark 1.2. By [12, Proposition 7.1], it suffices to prove Theorem 1.1 for covers over $\overline{\mathbb{Q}_p^{ur}}$, rather than $\overline{\mathbb{Q}}$, and for higher ramification groups over $\mathbb{Q}_p^{ur}$. Here, and throughout, $\mathbb{Q}_p^{ur}$ is the *completion* of the maximal unramified extension of $\mathbb{Q}_p$. This is the version we prove in §8. In fact, we prove even more, i.e., that the *stable model* of f can be defined over an extension $K/\mathbb{Q}_p^{ur}$ whose nth higher ramification groups for the upper numbering vanish above p.

Remark 1.3.
(1) Many simple groups G satisfy the hypotheses of Theorem 1.1 (for instance, any $PSL_2(\ell)$ where $p \neq 2$ and $v_p(\ell^2-1) = n$).
(2) If $N_G(P) = Z_G(P)$, then G is p-solvable by a theorem of Burnside ([25, Theorem 4, p. 169]), and thus falls within the scope of [12, Theorem 1.3], which treats p-solvable groups. Theorem 1.1 seems to be the next easiest case.
(3) I expect Theorem 1.1 to hold even if $p = 3$ or if all branching indices are divisible by p. See Question 9.1.

Remark 1.4. In the case that all three branch points of f as in Theorem 1.1 have prime-to-p branching index, we show (Proposition 8.1) that f actually has good reduction. We use this in Corollary 8.2 to give a proof that, when N is prime, the modular curve $X(N)$ has good reduction at all primes not dividing $6N$. In fact, Corollary 8.2 is more general, and works (in a slightly weakened sense) for all N. Our proof does not use the modular interpretation of $X(N)$, only that $X(N) \to X(1)$ can be given as a three-point $PSL_2(\mathbb{Z}/N)$-cover.

As in [12], our main technique for proving Theorem 1.1 will be an analysis of the *stable reduction* of the G-cover f to characteristic p. The major difference between the methods of this paper and those of [12] is this paper's use of the *auxiliary cover*. This is a construction, introduced by Raynaud ([19]), to simplify the group-theoretical structure of a Galois cover of curves. In particular, by replacing f with its auxiliary cover f^{aux}, we obtain a Galois cover whose field of moduli is related to that of f, but which now has a p-solvable Galois group.

Unfortunately, the cover f^{aux} in general has extra branch points, and it is not obvious where these extra branch points arise. The crux of the proof of Theorem 1.1 is understanding where these extra branch points are located; this is the content of §8.3.4, and is the reason why Theorem 1.1 is significantly more difficult than the analogous theorem where G is assumed to be p-solvable. Our assumption that $|N_G(P)/Z_G(P)| = 2$ alleviates this difficulty somewhat, as it allows us to work with more explicit equations (see §8.3 and Question 9.2).

1.3. Section-by-section summary and walkthrough. In §2, we give some explicit results on the reduction of μ_{p^n}-torsors. The purpose of §3, §4, and §5 is to recall the relevant material from [12]. In §3 and §4, we introduce stable reduction of G-covers, and state some of the basic properties. We also recall the *vanishing cycles formula* from [12], which is indispensable for the proof of Theorem 1.1. In §5, we recall the properties of *deformation data* (constructed in [15]), which give extra structure to the stable model.

The new part of the paper begins with §6, where we discuss *monotonicity* of stable reduction (a property that becomes relevant when $v_p(|G|) > 1$), and show that it is satisfied for all covers in this paper. In §7, we introduce the *auxiliary cover* and the *strong auxiliary cover*, and discuss why they are useful in calculating the field of moduli.

In §8, we prove our main result, Theorem 1.1. The proof is divided into §8.1, §8.2, and §8.3, corresponding to the case of 3, 2, and 1 branch point(s) with prime-to-p branching index, respectively. The proof for 1 branch point with prime-to-p index is by far the most difficult (as it involves the appearance of an extra branch point in the auxiliary cover), and toward the beginning of §8.3, we give an outline of the proof and of how it is split up over §8.3.3–§8.3.7.

In §9, we consider some questions arising from this work. In Appendix A we give an example of a three-point cover with nontrivial wild monodromy (see the appendix for more details). Appendix A is not needed for the rest of the paper.

For reasons that will become clear in §8.3, the case $p = 5$ presents some complications. The reader who is willing to assume $p > 5$ may skip Lemma 2.1 (2), Remark 8.18, Proposition 8.19 (2), Remark 8.25, Proposition 8.26 (2), Proposition 8.31 (2), (3), Lemma 8.33, and Proposition 8.35 (2c), which are among the more technical parts of the paper.

1.4. Notation and conventions. The letter k will always represent an algebraically closed field of characteristic $p > 0$.

If H is a subgroup of a finite group G, then $N_G(H)$ is the normalizer of H in G and $Z_G(H)$ is the centralizer of H in G. If G has a cyclic p-Sylow subgroup P, and p is understood, we write $m_G = |N_G(P)/Z_G(P)|$.

If K is a field, $\overline{K}$ is its algebraic closure. We write G_K for the absolute Galois group of K. If $H \leq G_K$, we write $\overline{K}^H$ for the fixed field of H in $\overline{K}$.

Similarly, if Γ is a group of automorphisms of a ring A, we write A^Γ for the fixed ring under Γ.

We use the standard theory of higher ramification groups for the upper and lower numbering from [21, IV]. However, unlike in [21], if L/K is a nontrivial G-Galois extension of complete discrete valuation rings with algebraically closed residue fields, then the *conductor* of L/K, written $h_{L/K}$, will be for us the greatest upper jump (i.e., the greatest i such that $G^i \neq \{id\}$). This is consistent with [12].

If R is any local ring, then $\hat{R}$ is the completion of R with respect to its maximal ideal. If R is any ring with a non-archimedean absolute value $|\cdot|$, then $R\{T\}$ is the ring of power series $\sum_{i=0}^{\infty} c_i T^i$ such that $\lim_{i\to\infty} |c_i| = 0$. If R is a discrete valuation ring with fraction field K of characteristic 0 and residue field k of characteristic p, we normalize the absolute value on K and on any subring of K so that $|p| = 1/p$. We always normalize the valuation on K so that p has valuation 1.

A *branched cover* $f : Y \to X$ is a finite, surjective, generically étale morphism of geometrically connected, smooth, proper curves. If f is of degree d and we choose an isomorphism $i : G \to \mathrm{Aut}(Y/X)$, then the datum (f, i) is called a *G-Galois cover* (or just a *G-cover*, for short). We will usually suppress the isomorphism i, and speak of f as a G-cover.

Suppose $f : Y \to X$ is a G-cover of smooth curves, and K is a field of definition for X. Then the *field of moduli of f relative to K (as a G-cover)* is $\overline{K}^{\Gamma^{in}}$, where $\Gamma^{in} = \{\sigma \in G_K | f^\sigma \cong f \text{ (as } G\text{-covers)}\}$ (see, e.g., [12, §1.1]). If X is $\mathbb{P}^1$, then the *field of moduli of f* means the field of moduli of f (as a G-cover) relative to $\mathbb{Q}$.

Let $f : Y \to X$ be any morphism of schemes and assume H is a finite group with $H \hookrightarrow \mathrm{Aut}(Y/X)$. If G is a finite group containing H, then there is a map $\mathrm{Ind}_H^G f : \mathrm{Ind}_H^G Y \to X$, where $\mathrm{Ind}_H^G Y$ is a disjoint union of $[G : H]$ copies of Y, indexed by the left cosets of H in G. The group G acts on $\mathrm{Ind}_H^G Y$, and the stabilizer of each copy of Y in $\mathrm{Ind}_H^G Y$ is a conjugate of H.

The set $\mathbb{N}$ is equal to $\{1, 2, 3, \ldots\}$.

Acknowledgements

This material is mostly adapted from my PhD thesis, and I thank my advisor, David Harbater, for much help related to this work. I also thank the referee for useful suggestions.

2. Reduction of μ_{p^n}-torsors

Let R be a mixed characteristic $(0, p)$ complete discrete valuation ring with residue field k and fraction field K. Let π be a uniformizer of R. Recall that we normalize the valuation of p (not π) to be 1. For any scheme or algebra S over R, write S_K and S_k for its base changes to K and k, respectively.

We state a partial converse of [12, Lemma 3.1], which will be used repeatedly in analyzing the stable reduction of covers (see §8.3):

Lemma 2.1. *Suppose R contains the p^nth roots of unity. Let $X = Spec\ A$, where $A = R\{T\}$. Let $f : Y_K \to X_K$ be a μ_{p^n}-torsor given by the equation $y^{p^n} = g$, where $g = 1 + \sum_{i=1}^{\infty} c_i T^i$. Suppose that $v(c_i) > n + \frac{1}{p-1}$ for all $i > p$ divisible by p. Suppose further that (at least) one of the following two conditions holds:*

1. *There exists i such that $v(c_i) < \min(v(c_p), n + \frac{1}{p-1})$.*
2. *$v(c_p) > n - \frac{p-2}{2(p-1)}$ and there exists $c'_p \in R$ with $v(c'_p - c_p) > n + \frac{1}{p-1}$ and $v\left(c_1 - \sqrt[p]{c'_p p^{(p-1)n+1}}\right) < n + \frac{1}{p-1}$.*

Then, even after a possible finite extension of K, the map $f : Y_K \to X_K$ does not *split into a union of p^{n-1} connected disjoint μ_p-torsors, such that if Y is the normalization of X in the total ring of fractions of Y_K, then $Y_k \to X_k$ is étale.*

Proof. Suppose we are in case (1). Pick $b \in R$ such that $v(b) = \min_i(v(c_i))$. Then $v(b) < n + \frac{1}{p-1}$, and $g = 1 + bw$ with $w \in A \backslash \pi A$. Let $a < n$ be the greatest integer such that $a + \frac{1}{p-1} < v(b)$. Then g has a p^ath root in A, given by the binomal expansion

$$\sqrt[p^a]{g} = 1 + \frac{1/p^a}{1!} bw + \frac{(1/p^a)((1/p^a) - 1)}{2!}(bw)^2 + \cdots .$$

Since $v(b) > a + \frac{1}{p-1}$, this series converges, and is in A. Furthermore, since the coefficients of all terms in this series of degree ≥ 2 have valuation greater than $v(b) - a$, the series can be written as $\sqrt[p^a]{g} = 1 + \frac{b}{p^a}u$, where u is congruent to $w \pmod{\pi}$.

Now, $v(b) - a = v(\frac{b}{p^a}) \leq 1 + \frac{1}{p-1}$. Furthermore, by assumption (1), the reduction $\overline{u}$ of u is not a pth power in A/π. Then [8, Proposition 1.6] shows that $\sqrt[p^a]{g}$ is not a pth power in A (nor in K). If $a < n - 1$, this proves that f does not split into a disjoint union of $n - 1$ torsors. If $a = n - 1$, then $v(\frac{b}{p^a}) < 1 + \frac{1}{p-1}$, and [8, Proposition 1.6] shows that the torsor given by $y^p = \sqrt[p^{n-1}]{g}$ does not have étale reduction. This proves the lemma in case (1).

Suppose we are in case (2) and not in case (1). It then suffices to show that there exists $h \in A$ such that $h^{p^n} g$ satisfies (1). Let $\eta = -\sqrt[p]{\frac{c'_p}{p^{n-1}}}$ (any pth root will do). Now, by assumption, $v(c'_p) - (n-1) > \frac{p}{2(p-1)}$, so

$v(\eta) > \frac{1}{2(p-1)}$. Then there exists $\epsilon > 0$ such that

$$(1+\eta T)^{p^n} \equiv 1 - p^n \sqrt[p]{\frac{c_p'}{p^{n-1}}}T - \binom{p^n}{p}\frac{c_p'}{p^{n-1}}T^p \pmod{p^{n+\frac{1}{p-1}+\epsilon}}.$$

It is easy to show that $\binom{p^n}{p} \equiv p^{n-1} \pmod{p^n}$ for all $n \geq 1$. So there exists $\epsilon > 0$ such that

$$(1+\eta T)^{p^n} \equiv 1 - \sqrt[p]{c_p' p^{(p-1)n+1}}T - c_p' T^p \pmod{p^{n+\frac{1}{p-1}+\epsilon}}.$$

Using the assumption that $v(c_1 - \sqrt[p]{c_p' p^{(p-1)n+1}}) < n + \frac{1}{p-1}$, we leave it to the reader to verify that $(1+\eta T)^{p^n} g$ satisfies (1) (with $i = 1$). $\square$

3. Semistable models of $\mathbb{P}^1$

Let R be a mixed characteristic $(0,p)$ complete discrete valuation ring with residue field k and fraction field K. If X is a smooth curve over K, then a *semistable* model for X is a relative flat curve $X_R \to \operatorname{Spec} R$ with $X_R \times_R K \cong X$ and semistable special fiber (i.e., the special fiber is reduced with only ordinary double points for singularities). If X_R is smooth, it is called a *smooth model*.

3.1. Models. Let $X \cong \mathbb{P}^1_K$. Write v for the valuation on K. Let X_R be a smooth model of X over R. Then there is an element $T \in K(X)$ such that $K(T) \cong K(X)$ and the local ring at the generic point of the special fiber of X_R is the valuation ring of $K(T)$ corresponding to the Gauss valuation (which restricts to v on K). We say that our model corresponds to the Gauss valuation on $K(T)$, and we call T a *coordinate* of X_R. Conversely, if T is any rational function on X such that $K(T) \cong K(X)$, there is a smooth model X_R of X such that T is a coordinate of X_R. In simple terms, T is a coordinate of X_R iff, for all $a, b \in R$, the subvarieties of X_R cut out by $T - a$ and $T - b$ intersect exactly when $v(a-b) > 0$.

Now, let X_R' be a semistable model of X over R. The special fiber of X_R' is a tree-like configuration of $\mathbb{P}^1_k$'s. Each irreducible component $\overline{W}$ of the special fiber $\overline{X}$ of X_R' yields a smooth model of X by blowing down all other irreducible components of $\overline{X}$. If T is a coordinate on the smooth model of X with $\overline{W}$ as special fiber, we will say that T corresponds to $\overline{W}$.

3.2. Disks and annuli. We give a brief overview here. For more details, see [7].

Let X_R' be a semistable model for $X = \mathbb{P}^1_K$. Suppose x is a smooth point of the special fiber $\overline{X}$ of X_R' on the irreducible component $\overline{W}$. Let T be a coordinate corresponding to $\overline{W}$ such that $T = 0$ specializes to x. Then the set of points of $X(\overline{K})$ which specialize to x is the *open p-adic disk* D given by $v(T) > 0$. The ring of functions on the formal disk corresponding to D is $\hat{\mathcal{O}}_{X,x} \cong R\{T\}$.

Now, let x be an ordinary double point of $\overline{X}$, at the intersection of components $\overline{W}$ and $\overline{W}'$. Then the set of points of $X(\overline{K})$ which specialize to x is an *open annulus* A. If T is a coordinate corresponding to $\overline{W}$ such that $T = 0$ specializes to $\overline{W}' \backslash \overline{W}$, then A is given by $0 < v(T) < e$ for some $e \in v(K^\times)$. The ring of functions on the formal annulus corresponding to A is $\hat{\mathcal{O}}_{X,x} \cong R[[T, U]]/(TU - p^e)$. Observe that e is independent of the coordinate. It is called the *épaisseur* of the annulus.

Suppose we have a preferred coordinate T on X and a semistable model X'_R of X whose special fiber $\overline{X}$ contains an irreducible component $\overline{X}_0$ corresponding to the coordinate T. If $\overline{W}$ is any irreducible component of $\overline{X}$ other than $\overline{X}_0$, then since $\overline{X}$ is a tree of $\mathbb{P}^1$'s, there is a unique non-repeating sequence of consecutive, intersecting components $\overline{X}_0, \ldots, \overline{W}$. Let $\overline{W}'$ be the component in this sequence that intersects $\overline{W}$. Then the set of points in $X(\overline{K})$ that specialize to the connected component of $\overline{W}$ in $\overline{X} \backslash \overline{W}'$ is a closed p-adic disk D. If the established preferred coordinate (equivalently, the preferred component $\overline{X}_0$) is clear, we will abuse language and refer to the component $\overline{W}$ as *corresponding to the disk* D, and vice versa. If U is a coordinate corresponding to $\overline{W}$, and if $U = \infty$ does not specialize to the connected component of $\overline{W}$ in $\overline{X} \backslash \overline{W}'$, then the ring of functions on the formal disk corresponding to D is $R\{U\}$.

4. Stable reduction

In §4, R is a mixed characteristic $(0, p)$ complete discrete valuation ring with residue field k and fraction field K. We set $X \cong \mathbb{P}^1_K$, and we fix a *smooth* model X_R of X. Let $f : Y \to X$ be a G-Galois cover defined over K, with G any finite group, such that the branch points of f are defined over K and their specializations do not collide on the special fiber of X_R. Assume that f is branched at at least three points. Using the stable reduction theorem for curves ([4, Corollary 2.7]), one can show that there is a unique minimal finite extension K^{st}/K with ring of integers R^{st} such that $f_{K^{st}} := f \times_K K^{st}$ has a *stable model* $f^{st} : Y^{st} \to X^{st}$ (which we will simply call the stable model of f). This model has the properties that:

- The special fiber $\overline{Y}$ of Y^{st} is semistable.
- The ramification points of $f_{K^{st}}$ specialize to *distinct* smooth points of $\overline{Y}$.
- Any genus zero irreducible component of $\overline{Y}$ contains at least three marked points (i.e., ramification points or points of intersection with the rest of $\overline{Y}$).
- G acts on Y^{st}, and $X^{st} = Y^{st}/G$.

The field K^{st} is called the minimal field of definition of the stable model of f. If we are working over a finite extension K'/K^{st} with ring of integers R', we will sometimes abuse language and call $f^{st} \times_{R^{st}} R'$ the stable model of f.

Remark 4.1. Our definition of the stable model is the definition used in [23]. This differs from the definition in [19] in that [19] allows the ramification points to coalesce on the special fiber.

Remark 4.2. Note that X^{st} can be naturally identified with a blowup of $X \times_R R^{st}$ centered at closed points. Furthermore, the nodes of $\overline{Y}$ lie above nodes of the special fiber $\overline{X}$ of X^{st} ([18, Lemme 6.3.5]), and Y^{st} is the normalization of X^{st} in $K^{st}(Y)$.

If $\overline{Y}$ is smooth, the cover $f : Y \to X$ is said to have *potentially good reduction*. If f does not have potentially good reduction, it is said to have *bad reduction*. In any case, the special fiber $\overline{f} : \overline{Y} \to \overline{X}$ of the stable model is called the *stable reduction* of f. The strict transform of the special fiber of $X_{R^{st}}$ in $\overline{X}$ (Remark 4.2) is called the *original component*, and will be denoted $\overline{X}_0$.

Each $\sigma \in G_K$ acts on $\overline{Y}$ (via its action on Y). This action commutes with that of G and is called the *monodromy action*. Then it is known (see, for instance, [15, Proposition 2.9]) that the extension K^{st}/K is the fixed field of the group $\Gamma^{st} \leq G_K$ consisting of those $\sigma \in G_K$ such that σ acts trivially on $\overline{Y}$. Thus K^{st} is clearly Galois over K. Since k is algebraically closed, the action of G_K fixes $\overline{X}_0$ pointwise.

4.1. The graph of the stable reduction. As in [23], we construct the (unordered) dual graph $\mathcal{G}$ of the stable reduction of $\overline{X}$. An *unordered graph* $\mathcal{G}$ consists of a set of *vertices* $V(\mathcal{G})$ and a nonempty set of *edges* $E(\mathcal{G})$. Each edge has a *source vertex* $s(e)$ and a *target vertex* $t(e)$. Each edge has an *opposite edge* $\overline{e}$, such that $s(e) = t(\overline{e})$ and $t(e) = s(\overline{e})$. Also, $\overline{\overline{e}} = e$.

Given f, $\overline{f}$, $\overline{Y}$, and $\overline{X}$ as above, we construct two unordered graphs $\mathcal{G}$ and $\mathcal{G}'$. In our construction, $\mathcal{G}$ has a vertex v for each irreducible component of $\overline{X}$ and an edge e for each ordered triple $(\overline{x}, \overline{W}', \overline{W}'')$, where $\overline{W}'$ and $\overline{W}''$ are irreducible components of $\overline{X}$ whose intersection is $\overline{x}$. If e corresponds to $(\overline{x}, \overline{W}', \overline{W}'')$, then $s(e)$ is the vertex corresponding to $\overline{W}'$ and $t(e)$ is the vertex corresponding to $\overline{W}''$. The opposite edge of e corresponds to $(\overline{x}, \overline{W}'', \overline{W}')$. We denote by $\mathcal{G}'$ the *augmented* graph of $\mathcal{G}$ constructed as follows: consider the set B_{wild} of branch points of f with branching index divisible by p. For each $x \in B_{\text{wild}}$, we know that x specializes to a unique irreducible component $\overline{W}_x$ of $\overline{X}$, corresponding to a vertex A_x of $\mathcal{G}$. Then $V(\mathcal{G}')$ consists of the elements of $V(\mathcal{G})$ with an additional vertex V_x for each $x \in B_{\text{wild}}$. Also, $E(\mathcal{G}')$ consists of the elements of $E(\mathcal{G})$ with two additional opposite edges for each $x \in B_{\text{wild}}$, one with source V_x and target A_x, and one with source A_x and target V_x. We write v_0 for the vertex corresponding to the original component $\overline{X}_0$.

If $v, w \in V(\mathcal{G}')$, then a *path* from v to w is a sequence of nonrepeating vertices $\{v_i\}_{i=0}^n$ and edges $\{e_i\}_{i=0}^{n-1}$ such that $v = v_0$, $w = v_n$, $s(e_i) = v_i$, and $t(e_i) = v_{i+1}$. Let $u \in V(\mathcal{G}')$ correspond to the original component $\overline{X}_0$. We partially order the vertices of $\mathcal{G}'$ such that $v \preceq w$ if there is a path from u to w passing through v. The set of irreducible components of $\overline{X}$ inherits the partial order $\preceq$. Furthermore, if $\overline{x}_1$ and $\overline{x}_2$ are points of $\overline{X}$, we say that $\overline{x}_2$ *lies outward from* $\overline{x}_1$ if $\overline{x}_1 \neq \overline{x}_2$ and there are irreducible components $\overline{X}_1 \prec \overline{X}_2$ of $\overline{X}$ such that $\overline{x}_1 \in \overline{X}_1$ and $\overline{x}_2 \in \overline{X}_2$. Lastly, an irreducible component $\overline{W}$ of $\overline{X}$ *lies outward from* a point $\overline{x} \in \overline{X}$ if there is an irreducible component $\overline{W}' \prec \overline{W}$ such that $\overline{x} \in \overline{W}'$.

4.2. Inertia Groups of the Stable Reduction. Recall that G acts on $\overline{Y}$. By [18, Lemme 6.3.3], we know that the inertia groups of the action of G on $\overline{Y}$ at generic points of $\overline{Y}$ are p-groups. Also, at each node of $\overline{Y}$, the inertia group is an extension of a cyclic, prime-to-p order group by a p-group generated by the inertia groups of the generic points of the crossing components. If $\overline{V}$ is an irreducible component of $\overline{Y}$, we will always write $I_{\overline{V}} \leq G$ for the inertia group of the generic point of $\overline{V}$, and $D_{\overline{V}} \leq G$ for the decomposition group.

For the rest of this subsection, assume G has a *cyclic* p-Sylow subgroup. In this case, the inertia groups above a generic point of an irreducible component $\overline{W} \subset \overline{X}$ are conjugate cyclic groups of p-power order. If they are of order p^i, we call $\overline{W}$ a *p^i-component*. If $i = 0$, we call $\overline{W}$ an *étale component*, and if $i > 0$, we call $\overline{W}$ an *inseparable component*.

As in [19], we call irreducible component $\overline{W} \subseteq \overline{X}$ a *tail* if it is not the original component and intersects exactly one other irreducible component of $\overline{X}$. Otherwise, it is called an *interior component*. A tail of $\overline{X}$ is called *primitive* if it contains a branch point other than the point at which it intersects the rest of $\overline{X}$. Otherwise it is called *new*. This follows [23]. An inseparable tail that is a p^i-component will also be called a *p^i-tail*. Thus one can speak of, for instance, "new p^i-tails" or "primitive étale tails."

Lemma 4.3 ([15], Proposition 2.13)**.** *If $x \in X$ is branched of index $p^a s$, where $p \nmid s$, then x specializes to a p^a-component.*

Lemma 4.4 ([19], Proposition 2.4.8)**.** *If f has bad reduction and $\overline{W}$ is an étale component of $\overline{X}$, then $\overline{W}$ is a tail.*

Lemma 4.5 ([15], Proposition 2.16, see also [19], Remarque 3.1.8)**.** *If f has bad reduction and $\overline{W}$ is a p^i-tail of $\overline{X}$, then the component $\overline{W}'$ that intersects $\overline{W}$ is a p^j-component with $j > i$.*

Definition 4.6. (cf. [15], Definition 2.18) Consider a component $\overline{X}_b \neq \overline{X}_0$ of $\overline{X}$. Let $\overline{W}$ be the unique component of $\overline{X}$ such that $\overline{W} \prec \overline{X}_b$ and $\overline{W}$ intersects $\overline{X}_b$, say at $\overline{x}_b$. Suppose that $\overline{W}$ is a p^j-component and $\overline{X}_b$ is a p^i-component, $i < j$. Let $\overline{Y}_b$ be a component of $\overline{Y}$ lying above $\overline{X}_b$, and let $\overline{y}_b$ be a point lying above $\overline{x}_b$. Then the *effective ramification invariant* σ_b of $\overline{X}_b$ is defined as follows: If $\overline{X}_b$ is an étale component, then σ_b is the conductor of higher ramification for the extension $\hat{\mathcal{O}}_{\overline{Y}_b,\overline{y}_b}/\hat{\mathcal{O}}_{\overline{X}_b,\overline{x}_b}$. If $\overline{X}_b$ is a p^i-component ($i > 0$), then the extension $\hat{\mathcal{O}}_{\overline{Y}_b,\overline{y}_b}/\hat{\mathcal{O}}_{\overline{X}_b,\overline{x}_b}$ can be factored as $\hat{\mathcal{O}}_{\overline{X}_b,\overline{x}_b} \overset{\alpha}{\hookrightarrow} S \overset{\beta}{\hookrightarrow} \hat{\mathcal{O}}_{\overline{Y}_b,\overline{y}_b}$, where α is Galois and β is purely inseparable of degree p^i. Then σ_b is the conductor of higher ramification for the extension $S/\hat{\mathcal{O}}_{\overline{X}_b,\overline{x}_b}$.

Remark 4.7. By Lemma 4.5, the effective ramification invariant is defined for every tail.

Lemma 4.8 ([15], Lemma 2.20)**.** *The effective ramification invariants σ_b lie in $\frac{1}{m_G}\mathbb{Z}$.*

4.3. Vanishing cycles formula. Assume the notation of §4. The *vanishing cycles formula* stated below will be used repeatedly:

Theorem 4.9 (Vanishing cycles formula, cf., [15], Theorem 3.14, Corollary 3.15)**.** *Let $f : Y \to X \cong \mathbb{P}^1$ be a three-point G-Galois cover with bad reduction, where G has a cyclic p-Sylow subgroup. Let B_{new} be an indexing set for the new étale tails and let B_{prim} be an indexing set for the primitive étale tails. Let σ_b be the ramification invariant in Definition 4.6. Then we have the formula*

$$1 = \sum_{b \in B_{new}} (\sigma_b - 1) + \sum_{b \in B_{prim}} \sigma_b. \tag{4.1}$$

Corollary 4.10. *If $m_G = 2$, and if f has bad reduction, then there are at most two étale tails. Furthermore, for any étale tail $\overline{X}_b$, $\sigma_b \in \frac{1}{2}\mathbb{Z}$ (see §4).*

Proof. By Lemma [15, Lemma 4.2 (i)], each term on the right hand side of (4.1) is at least $\frac{1}{m_G} = \frac{1}{2}$, so both parts of the corollary follow immediately. ∎

5. Deformation data

Deformation data arise naturally from the stable reduction of covers. Much information is lost when we pass from the stable model of a cover to its stable reduction, and deformation data provide a way to retain some of this information. This process is described in detail in [15, §3.2], and we recall some facts here.

5.1. Generalities. Let $\overline{W}$ be any connected smooth proper curve over k. Let H be a finite group and χ a 1-dimensional character $H \to \mathbb{F}_p^\times$. A *deformation datum* over $\overline{W}$ of type (H, χ) is an ordered pair $(\overline{V}, \omega)$ such that: $\overline{V} \to \overline{W}$ is an H-cover; ω is a meromorphic differential form on $\overline{V}$ that is either logarithmic or exact (i.e., $\omega = du/u$ or du for $u \in k(\overline{V})$); and $\eta^*\omega = \chi(\eta)\omega$ for all $\eta \in H$. If ω is logarithmic (resp. exact), the deformation datum is called multiplicative (resp. additive). When $\overline{V}$ is understood, we will sometimes speak of the deformation datum ω.

If $(\overline{V}, \omega)$ is a deformation datum, and $w \in \overline{W}$ is a closed point, we define m_w to be the order of the prime-to-p part of the ramification index of $\overline{V} \to \overline{W}$ at w. Define h_w to be $\mathrm{ord}_v(\omega) + 1$, where $v \in \overline{V}$ is any point which maps to $w \in \overline{W}$. This is well-defined because $\eta^*\omega$ is a nonzero scalar multiple of ω for $\eta \in H$.

Lastly, define $\sigma_w = h_w/m_w$. We call w a *critical point* of the deformation datum $(\overline{V}, \omega)$ if $(h_w, m_w) \neq (1, 1)$. Note that every deformation datum contains only a finite number of critical points. The ordered pair (h_w, m_w) is called the *signature* of $(\overline{V}, \omega)$ (or of ω, if $\overline{V}$ is understood) at w, and σ_w is called the *invariant* of the deformation datum at w.

5.2. Deformation data arising from stable reduction. We use the notation of §4. Assume that a p-Sylow subgroup of G is cyclic. For each irreducible component of $\overline{Y}$ lying above a p^r-component of $\overline{X}$ with $r > 0$, we construct r different deformation data. The details of this construction are given in [15, Construction 3.4], and we do not give them here. Rather, we recall the important properties.

Suppose $\overline{V}$ is an irreducible component of $\overline{Y}$ with nontrivial generic inertia group $I_{\overline{V}} \cong \mathbb{Z}/p^r \subset G$. If $\overline{V}'$ is the smooth projective model of the function field of $k(\overline{V})^{p^r}$, then [15, Construction 3.4] constructs r meromorphic differential forms $\omega_1, \ldots, \omega_r$ on $\overline{V}'$ (well defined either up to scalar multiplication by $k^\times$ or by $\mathbb{F}_p^\times$, depending on whether the differential form is exact or logarithmic). Furthermore, if $H = D_{\overline{V}}/I_{\overline{V}}$, then H acts faithfully on $\overline{V}'$, and $\overline{W} \cong \overline{V}/H$. It is shown in [15, Construction 3.4] that $(\overline{V}', \omega_i)$ is in fact a deformation datum of type (H, χ) over $\overline{W}$ for $1 \leq i \leq r$, where χ is given by the conjugation action of H on $I_{\overline{V}}$. The invariant of σ_i at a point $w \in \overline{W}$ will be denoted $\sigma_{i,w}$. Since f is Galois, these invariants do not depend on which component $\overline{V}$ above $\overline{W}$ is chosen. The differential forms $\omega_1, \ldots, \omega_r$ correspond, in some sense, to the successive degree p extensions building a tower between $\overline{V}'$ and $\overline{V}$. For this reason, we will sometimes call the deformation datum $(\overline{V}', \omega_1)$ the *bottom deformation datum* for $\overline{V}$.

Furthermore, for $1 \leq i \leq r$, we associate a rational number δ_i (see [12, §5.2]). If ω_i is multiplicative, then $\delta_i = 1$. Otherwise, $0 < \delta_i < 1$. The *effective different* $\delta^{\text{eff}}_{\overline{W}}$ above $\overline{W}$ is defined by

$$\delta^{\text{eff}}_{\overline{W}} = \left(\sum_{i=1}^{r-1} \delta_i\right) + \frac{p}{p-1}\delta_r.$$

By convention, if $\overline{W}$ is an étale component, we set $\delta^{\text{eff}}_{\overline{W}} = 0$.

The following lemma will be very important in the main proof.

Lemma 5.1 ([15], Lemma 3.5, cf. [23], Proposition 1.7)**.** *Say* $(\overline{V}', \omega)$ *is a deformation datum arising from the stable reduction of a cover, and let* $\overline{W}$ *be the component of* $\overline{X}$ *lying under* $\overline{V}'$*. Then a critical point* w *of the deformation datum on* $\overline{W}$ *is either a singular point of* $\overline{X}$ *or the specialization of a branch point of* $Y \to X$ *with ramification index divisible by* p*. In the first case,* $\sigma_w \neq 0$*, and in the second case,* $\sigma_w = 0$ *and* ω *is logarithmic.*

Recall that $\mathcal{G}'$ is the augmented dual graph of $\overline{X}$ (§4.1). To each $e \in E(\mathcal{G}')$ we will associate the *effective invariant* σ^{eff}_b, and to each vertex of $\mathcal{G}$ we will associate the *effective different* δ^{eff}_v.

Definition 5.2 (cf. [15], Definition 3.10)**.** Let $e \in E(\mathcal{G}')$.

(1) Suppose e corresponds to the triplet $(w, \overline{W}, \overline{W}')$, where $\overline{W}$ is a p^r-component and $\overline{W}'$ is a $p^{r'}$-component with $r \geq r'$. Then $r \geq 1$ by Lemma 4.4. Let ω_i, $1 \leq i \leq r$, be the deformation data above $\overline{W}$. Then

$$\sigma^{\text{eff}}_e := \left(\sum_{i=1}^{r-1} \frac{p-1}{p^i}\sigma_{i,w}\right) + \frac{1}{p^{r-1}}\sigma_{r,w}.$$

Note that this is a weighted average of the $\sigma_{i,w}$'s.

(2) If $s(e)$ corresponds to a p^r-component and $t(e)$ corresponds to a $p^{r'}$-component with $r < r'$, then $\sigma^{\text{eff}}_e := -\sigma^{\text{eff}}_{\overline{e}}$.

(3) If either $s(e)$ or $t(e)$ is in $V(\mathcal{G}')\backslash V(\mathcal{G})$, then $\sigma^{\text{eff}}_e := 0$.

(4) For all $v \in V(\mathcal{G})$, define $\delta^{\text{eff}}_v = \delta^{\text{eff}}_{\overline{W}}$, where v corresponds to $\overline{W}$.

(5) For all $e \in E(\mathcal{G})$, define ϵ_e to be the épaisseur of the formal annulus corresponding to e (§3).

Remark 5.3. In the paper [16], similar ideas of deformation data are used, but the notation and method of calculation is somewhat different. Suppose $\overline{V}$ is an irreducible component of $\overline{Y}$ as in this section such that $D_{\overline{V}} = I_{\overline{V}} = \mathbb{Z}/p^r$, and let $\overline{W}$ be the component of $\overline{X}$ lying below it. If $\eta_{\overline{V}}$, $\eta_{\overline{W}}$ are the generic points of $\overline{V}$, $\overline{W}$, then $\hat{\mathcal{O}}_{Y^{st},\overline{V}}/\hat{\mathcal{O}}_{X^{st},\overline{W}}$ is a $\mathbb{Z}/p^r$-extension of complete discrete valuation rings, corresponding to a character χ in the language of [16, §5]. Our $\delta^{\text{eff}}_{\overline{W}}$ is equal to $\text{sw}(\chi)$ in [16, §5.3], and if $e \in E(\mathcal{G}')$ corresponds to $(w, \overline{W}, \overline{W}')$ as in Definition 5.2, then our σ^{eff}_e is equal to $\text{ord}_w(\text{dsw}(\chi)) + 1$ in [16, §5.3].

Lemma 5.4 ([15], Lemma 3.11 (i), (iii), [12], Lemma 5.10)**.**

(1) *For any* $e \in E(\mathcal{G}')$, *we have* $\sigma_e^{\mathrm{eff}} = -\sigma_{\overline{e}}^{\mathrm{eff}}$.

(2) *If* $t(e)$ *corresponds to an étale tail* $\overline{X}_b$, *then* $\sigma_e^{\mathrm{eff}} = \sigma_b$.

(3) *If* $e \in E(\mathcal{G})$, *then* $\delta_{s(e)}^{\mathrm{eff}} - \delta_{t(e)}^{\mathrm{eff}} = \sigma_e^{\mathrm{eff}} \epsilon_e$.

The following lemma is very important for §8.3:

Lemma 5.5 ([12], Lemma 5.7)**.** *Let* $e \in E(\mathcal{G})$ *such that* $s(e) \prec t(e)$. *Let* $\overline{w} \in \overline{X}$ *be the point corresponding to* e. *Let* Π_e *be the set of branch points of* f *with branching index divisible by* p *that specialize outward from* $\overline{w}$. *Let* B_e *index the set of étale tails* $\overline{X}_b$ *lying outward from* $\overline{w}$. *Then*

$$\sigma_e^{\mathrm{eff}} - 1 = \sum_{b \in B_e} (\sigma_b - 1) - |\Pi_e|.$$

Corollary 5.6 (Monotonicity of the effective different)**.** *If* $v, v' \in V(\mathcal{G})$, *and* $v \prec v'$, *then* $\delta_v^{\mathrm{eff}} \geq \delta_{v'}^{\mathrm{eff}}$.

Proof. Clearly we may assume that v and v' are adjacent, i.e., there is an edge e such that $s(e) = v$ and $t(e) = v'$. Since the branch points of f are assumed not to collide on the special fiber of our original smooth model X_R, there is at most one branch point of f specializing outward from the node $\overline{x}_e$ corresponding to e. That is, there is at most either one primitive tail or one branch point of index divisible by p lying outward from $\overline{x}_e$. Since $\sigma_b > 1$ for all new tails $\overline{X}_b$ ([15, Lemma 4.2 (i)]), we see by Lemma 5.5 that $\sigma_e^{\mathrm{eff}} \geq 0$. We conclude using Lemma 5.4 (3). ☐

6. Monotonicity

We maintain the assumptions and notation of §4, along with the assumption that a p-Sylow subgroup of G is cyclic of order p^n.

Lemma 6.1. *Let* x *be a branch point of* f *with branching index exactly divisible by* p^r. *If* x *specializes to an irreducible component* $\overline{W}$ *of* $\overline{X}$, *then either* $\overline{W}$ *is the original component, or the unique component* $\overline{W}'$ *such that* $\overline{W}' \prec \overline{W}$ *and* $\overline{W}'$ *intersects* $\overline{W}$ *is a* p^s*-component, for some* $s > r$.

Proof. By Lemma 5.1 the deformation data above $\overline{W}$ are all multiplicative, and by [12, Proposition 5.2] they are all identical. Thus $\delta_{\overline{W}}^{\mathrm{eff}} = r + \frac{1}{p-1}$, as $\delta_{\omega_i} = 1$ for all ω_i above $\overline{W}$. Assume $\overline{W}$ is not the original component. Then, by Corollary 5.6, $\delta_{\overline{W}'}^{\mathrm{eff}} \geq r + \frac{1}{p-1}$. This is impossible unless $\overline{W}$ is a p^s-component with $s \geq r$. If $s = r$, we must have $\delta_{\overline{W}'}^{\mathrm{eff}} = \delta_{\overline{W}}^{\mathrm{eff}} = r + \frac{1}{p-1}$. By Corollary 5.6, $\sigma_e^{\mathrm{eff}} = 0$ for e either edge corresponding to $\{w\} = \overline{W} \cap \overline{W}'$. Since the deformation data above $\overline{W}$ are identical and σ_e^{eff} is a weighted average of invariants, we have $\sigma_{i,w} = 0$ for all ω_i above $\overline{W}$. But this contradicts Lemma 5.1. So $s > r$. ☐

Lemma 6.2. *Let $\overline{x}$ be a singular point of $\overline{X}$ such that there are no étale tails $\overline{X}_b$ lying outward from $\overline{x}$. If $I_{\overline{x}} \leq G$ is an inertia group above $\overline{x}$, then $m_{I_{\overline{x}}} = 1$.*

Proof. We first claim that, given an inseparable component $\overline{W} \subseteq \overline{X}$, there cannot be exactly one point $\overline{w} \in \overline{W}$ such that $m_{I_{\overline{w}}} > 1$. To prove the claim, let $\overline{V} \in \overline{Y}$ be a component above $\overline{W}$. Since $\overline{W}$ is inseparable, it follows that $D_{\overline{V}}$ has a normal subgroup of order p (namely, the order p subgroup of $I_{\overline{V}}$). By [15, Corollary 2.4], $D_{\overline{V}}$ has a quotient of the form $\mathbb{Z}/p^{\nu} \rtimes \mathbb{Z}/m_{D_{\overline{V}}}$, where the action of $\mathbb{Z}/m_{D_{\overline{V}}}$ on $\mathbb{Z}/p^{\nu}$ is faithful, and $\mathbb{Z}/p^{\nu}$ is a p-Sylow subgroup of $D_{\overline{V}}$. If $m_{D_{\overline{V}}} = 1$, then $m_{I_{\overline{w}}} = 1$ for all $\overline{w} \in \overline{W}$, so assume $m_{D_{\overline{V}}} > 1$. Then $\overline{V} \to \overline{W}$ has a quotient $\mathbb{Z}/m_{D_{\overline{V}}}$-cover $\overline{V}' \to \overline{W}$, which must be branched at at least two points, say $\overline{w}_1$ and $\overline{w}_2$. Then $I_{\overline{w}_1}$ and $I_{\overline{w}_2}$ are non-abelian subgroups of $\mathbb{Z}/p^{\nu} \rtimes \mathbb{Z}/m_{D_{\overline{V}}}$, meaning that $m_{I_{\overline{w}_1}}$ and $m_{I_{\overline{w}_2}}$ are greater than 1. This proves the claim.

Now, if $\overline{W}$ is an inseparable tail, then there is only one possible point $\overline{w} \in \overline{W}$ where $m_{I_{\overline{w}}}$ might not be 1 (the intersection point with the rest of $\overline{X}$), and the claim shows that we do, in fact, have $m_{I_{\overline{w}}} = 1$. The lemma then follows by inward induction. □

Definition 6.3. We call the stable reduction $\overline{f}$ of f *monotonic* if for every $\overline{W} \preceq \overline{W}'$, the inertia group of $\overline{W}'$ is contained in the inertia group of $\overline{W}$. In other words, the stable reduction is monotonic if the generic inertia does not increase as we move outward from $\overline{X}_0$ along $\overline{X}$.

In the situation of Theorem 1.1, all covers are monotonic:

Proposition 6.4. *If f is a three-point G-cover of $\mathbb{P}^1$, where G has a cyclic p-Sylow subgroup of order p^n, and $m_G = 2$, then $\overline{f}$ is monotonic.*

Proof. Suppose f is not monotonic. Then there exist $j \leq n$ and a set Σ of p^j-components of $\overline{X}$ with the following properties: $\overline{X}_0 \notin \Sigma$; the union $\overline{U}$ of the components in Σ (viewed as a closed subset of $\overline{X}$) is connected; and each irreducible component of $\overline{X}$ that intersects $\overline{U}$ but is not in Σ is a p^i-component, $i < j$ (think of $\overline{U}$ as being a "plateau" for inertia). In particular, $j > 0$. Note that, by Lemmas 4.3 and 6.1, no branch point of f specializes to $\overline{U}$ (this is the only place where we use $\overline{X}_0 \notin \Sigma$).

Recall that $\mathcal{G}$ (resp. $\mathcal{G}'$) is the dual graph (resp. augmented dual graph) of $\overline{X}$ (§4.1). Let Φ be the set of all edges $e \in E(\mathcal{G}')$ such that $s(e)$ corresponds to a component in Σ, and let Φ' be the set of those $e \in \Phi$ such that $t(e)$ does not correspond to a component in Σ (in particular, the inertia of $t(e)$ has order smaller than p^j). If $e \in \Phi$ corresponds to a point $\overline{x} \in \overline{X}$, then we write σ_e^{bot} to mean the invariant of any bottom deformation datum above the component corresponding to $s(e)$ at $\overline{x}$ (this is equivalent to $\sigma_e^{\mathrm{eff},j-1}$

in the language of [15, Definition 3.10]). By [15, Lemma 3.11 (i)] we have $\sigma_e^{\mathrm{bot}} = -\sigma_{\overline{e}}^{\mathrm{bot}}$ for all $e \in \Phi\backslash\Phi'$. By [15, Lemma 3.12] (and the fact that no branch point of f specializes to $\overline{U}$) we have, for any vertex v representing a component in Σ, that

$$\sum_{\substack{e\in E(\mathcal{G}')\\ s(e)=v}} (\sigma_e^{\mathrm{bot}} - 1) = \sum_{\substack{e\in E(\mathcal{G})\\ s(e)=v}} (\sigma_e^{\mathrm{bot}} - 1) = -2.$$

A simple induction argument (cf. the proof of [23, Corollary 1.11] or [15, Theorem 3.14]) shows that

$$\sum_{e\in\Phi'} (\sigma_e^{\mathrm{bot}} - 1) = -2. \tag{6.1}$$

Suppose $e \in \Phi'$ corresponds to a point $\overline{x} \in \overline{X}$, and let $m_e = m_{I_{\overline{x}}}$, where $I_{\overline{x}}$ is an inertia group of $\overline{f}$ above $\overline{X}$. By Lemma 4.8, we have $\sigma_e^{\mathrm{bot}} \in \frac{1}{m_e}\mathbb{Z}$. Since $t(e)$ has smaller inertia order than $s(e)$, [15, Lemma 3.11 (ii)] shows that $\sigma_e^{\mathrm{bot}} > 0$. Clearly, $m_e \in \{1, 2\}$, and $m_e = 1$ if no étale tails of $\overline{X}$ lie outward from $\overline{x}$ (Lemma 6.2).

In particular, $\sigma_e^{\mathrm{bot}} - 1 \geq -\frac{1}{2}$, and $\sigma_e^{\mathrm{bot}} - 1 \geq 0$ if there are no étale tails lying outward from $\overline{x}$. Since $m_G = 2$, Corollary 4.10 shows that there can be at most two étale tails. Thus there are at most three $e \in \Phi'$ such that an étale tail lies outward from the node corresponding to e (at worst, the outermost e preceding each of the étale tails and the innermost e). So $\sigma_e^{\mathrm{bot}} - 1 \geq 0$ for all but at most three edges $e \in \Phi'$. This contradicts (6.1), proving the proposition. □

7. The auxiliary cover

We maintain the notation of §4. Assume that $f : Y \to X$ is a G-cover defined over K as in §4 with bad reduction, so that $\overline{X}$ is not just the original component (G need not have a cyclic p-Sylow group). Following [19, §3.2], we can construct an *auxiliary cover* $f^{aux} : Y^{aux} \to X$ with (modified) stable model $(f^{aux})^{st} : (Y^{aux})^{st} \to X^{st}$ and (modified) stable reduction $\overline{f}^{aux} : \overline{Y}^{aux} \to \overline{X}$, defined over some finite extension R' of R. We will explain what “modified" means in a remark following the construction. The construction is almost entirely the same as in [19, §3.2], and we will not repeat the details. Instead, we give an overview, and we mention where our construction differs from Raynaud's.

Let $B_{\text{ét}}$ index the étale tails of $\overline{X}$. Subdividing $B_{\text{ét}}$, we index the set of primitive tails by B_{prim} and the set of new tails by B_{new}. We will write $\overline{X}_b$ for the tail indexed by $b \in B_{\text{ét}}$.

The construction proceeds as follows: From $\overline{Y}$ remove all of the components that lie above the *étale* tails of $\overline{X}$ (as opposed to *all* the tails—this is the only thing that needs to be done differently than in [19], where all tails are étale). Now, what remains of $\overline{Y}$ is possibly disconnected. We choose one connected component, and call it $\overline{V}$.

For each $b \in B_{\text{prim}}$, let a_b be the branch point of f specializing to $\overline{X}_b$, let $\overline{x}_b$ be the point where $\overline{X}_b$ intersects the rest of $\overline{X}$, and let $p^r m_b$ be the index of ramification above $\overline{X}_b$ at $\overline{x}_b$, with m_b prime-to-p. Then $\overline{X}_b$ intersects a p^r-component. At each point $\overline{v}_b$ of $\overline{V}$ above $\overline{x}_b$, we attach to $\overline{V}$ a *Katz-Gabber* cover of $\overline{X}_b$ (cf. [10, Theorem 1.4.1], [19, Théorème 3.2.1]), branched of order m_b (with inertia groups isomorphic to $\mathbb{Z}/m_b$) at the specialization $\overline{a}_b$ of a_b and of order $p^r m_b$ (with inertia groups isomorphic to $\mathbb{Z}/p^r \rtimes \mathbb{Z}/m_b$) at $\overline{x}_b$. We choose our Katz-Gabber cover so that above the complete local ring of $\overline{x}_b$ on $\overline{X}_b$, it is isomorphic to the original cover. It is the composition of a cyclic cover of order m_b branched at $\overline{x}_b$ and $\overline{a}_b$ with a cyclic cover of order p^r branched at one point. Note that if $m_b = 1$, we have eliminated the branch point $\overline{a}_b$ of the original cover.

For each $b \in B_{\text{new}}$, we carry out the same procedure, except that we *introduce* an (arbitrary) branch point $\overline{a}_b \neq \overline{x}_b$ of ramification index m_b on the new tail $\overline{X}_b$.

Let $\overline{f}^{aux} : \overline{Y}^{aux} \to \overline{X}$ be the cover of k-schemes we have just constructed. Let $G^{aux} \leq G$ be the decomposition group of $\overline{V}$. As in [19, §3.2], one shows that, after a possible finite extension R' of R, we can lift $\overline{f}^{aux}$ to a map $(f^{aux})^{st} : (Y^{aux})^{st} \to X^{st}$ over R', satisfying the following properties:

(1) Above an étale neighborhood of the union of those components of $\overline{X}$ other than étale tails, the cover $f^{st} : Y^{st} \to X^{st}$ is $\text{Ind}_{G^{aux}}^{G}(f^{aux})^{st}$ (see §1.4).
(2) The generic fiber $f^{aux} : Y^{aux} \to X$ is a G^{aux}-cover branched exactly at the branch points of f and at a new point a_b of index m_b for each new tail $b \in B_{\text{new}}$ (unless $m_b = 1$, as noted above). Each a_b specializes to the corresponding branch point $\overline{a}_b$ introduced above.

Keep in mind that there is some choice here in how to pick the new branch points—for a new tail $\overline{X}_b$, depending on the choice of $\overline{a}_b$, we can choose a_b to be *any* point of X that specializes to $\overline{X}_b \backslash \overline{x}_b$. The set of such points forms a closed p-adic disk (§3.2)

The generic fiber f^{aux} of $(f^{aux})^{st}$ is called the *auxiliary cover*, and $(f^{aux})^{st}$ is called the *modified stable model* of the auxiliary cover. The special fiber $\overline{f}^{aux}$ is called the *modified stable reduction* of the auxiliary cover.

Remark 7.1. Usually, the stable model of f^{aux} is same as the modified stable model $(f^{aux})^{st}$. However, it may happen that the stable model of f^{aux} is a contraction of $(f^{aux})^{st}$ (or that it is not even defined, as we may have eliminated a branch point by passing to the auxiliary cover). This happens only if $\overline{X}$ has a primitive tail $\overline{X}_b$ for which $m_b = 1$, and for which the Katz-Gabber cover inserted above $\overline{X}_b$ has genus zero. Then this tail, and possibly some components inward, would be contracted in the stable model of f^{aux}. We use the term *modified* stable model to mean that we do not perform this contraction. Alternatively, we can think of $(f^{aux})^{st}$ as the stable model of f^{aux}, if we count specializations of all points of Y^{aux} above branch points of f (as opposed to f^{aux}) as marked points.

If we are interested in understanding the field of moduli of a G-cover (or more generally, the minimal field of definition of the stable model), it is in some sense good enough to understand the auxiliary cover, as the following lemma shows.

Lemma 7.2. *If the modified stable model $(f^{aux})^{st} : (Y^{aux})^{st} \to X^{st}$ of the auxiliary cover f^{aux} is defined over a Galois extension K^{aux}/K_0, then the stable model f^{st} of f can also be defined over K^{aux}.*

Proof. (cf. [23], Theorem 4.5) Take $\sigma \in \Gamma^{aux}$, the absolute Galois group of K^{aux}. We must show that $f^\sigma \cong f$ and that σ acts trivially on the stable reduction $\overline{f} : \overline{Y} \to \overline{X}$ of f. Let $\hat{f} : \hat{Y} \to \hat{X}$ be the formal completion of f^{st} at the special fiber and let $\hat{f}^{aux} : \hat{Y}^{aux} \to \hat{X}$ be the formal completion of $(f^{aux})^{st}$ at the special fiber. For each étale tail $\overline{X}_b$ of $\overline{X}$, let $\overline{x}_b$ be the intersection of $\overline{X}_b$ with the rest of $\overline{X}$. Write $\mathcal{D}_b$ for the formal completion of $\overline{X}_b \backslash \{\overline{x}_b\}$ in $X_{R^{st}}$. Then $\mathcal{D}_b$ is a closed formal disk, which is certainly preserved by σ. Also, let $\mathcal{U}$ be the disjoint union of the formal completion of $\overline{X} \setminus \bigcup_b \overline{X}_b$ with the formal completions of the $\overline{x}_b$ in $X_{R^{st}}$.

Write $\mathcal{V} = \hat{Y} \times_{\hat{X}} \mathcal{U}$. We know from the construction of the auxiliary cover that

$$\mathcal{V} = \mathrm{Ind}_{G^{aux}}^{G}\ \hat{Y}^{aux} \times_{\hat{X}} \mathcal{U}.$$

Since σ preserves the auxiliary cover and acts trivially on its special fiber, σ acts as an automorphism on $\mathcal{V}$ and acts trivially on its special fiber. By uniqueness of tame lifting, $\mathcal{E}_b := \hat{Y} \times_{\hat{X}} \mathcal{D}_b$ is the unique lift of $\overline{Y} \times_{\overline{X}} (\overline{X}_b \backslash \{\overline{x}_b\})$ to a cover of $\mathcal{D}_b$ (where the branching is compatible with that of f, if $\overline{X}_b$ is primitive). This means that σ acts as an automorphism on $\hat{Y} \times_{\hat{X}} \mathcal{D}_b$ as well.

Define $\mathcal{B}_b := \mathcal{U} \times_{\hat{X}} \mathcal{D}_b$, the boundary of the disk $\mathcal{D}_b$. A G-cover of formal schemes $\hat{Y} \to \hat{X}$ such that $\hat{Y} \times_{\hat{X}} \mathcal{U} \cong \mathcal{V}$ and $\hat{Y} \times_{\hat{X}} \mathcal{D}_b \cong \mathcal{E}_b$ is determined by a patching isomorphism

$$\varphi_b : \mathcal{V} \times_{\mathcal{U}} \mathcal{B}_b \xrightarrow{\sim} \mathcal{E}_b \times_{\mathcal{D}_b} \mathcal{B}_b$$

for each b. The isomorphism φ_b is determined by its restriction $\overline{\varphi}_b$ to the special fiber.

Let $\overline{X}_{b,\infty}$ be the generic point of Spec $\hat{\mathcal{O}}_{\overline{X}_b,\overline{x}_b}$, and define $\overline{Y}_{b,\infty}$ (resp. $\overline{Y}^{aux}_{b,\infty}$) to be $\overline{Y} \times_{\overline{X}} \overline{X}_{b,\infty}$ (resp. $\overline{Y}^{aux} \times_{\overline{X}} \overline{X}_{b,\infty}$). Then $\overline{Y}_{b,\infty} = \mathrm{Ind}_{G^{aux}}^{G} \overline{Y}^{aux}_{b,\infty}$. Since σ acts trivially on $\overline{Y}^{aux}_{b,\infty}$, it acts trivially on $\overline{Y}_{b,\infty}$, which is the special fiber of both $\mathcal{V} \times_{\mathcal{U}} \mathcal{B}_b$ and $\mathcal{E}_b \times_{\mathcal{D}_b} \mathcal{B}_b$. So σ acts trivially on $\overline{\varphi_b}$, and thus on φ_b. Thus $\hat{f}^\sigma \cong \hat{f}$, and by Grothendieck's Existence Theorem, $f^\sigma \cong f$.

Lastly, we must check that σ acts trivially on $\overline{f}$. This is clear away from the étale tails. Now, for each étale tail $\overline{X}_b$, we know σ acts trivially on $\overline{X}_b$, so it must act vertically on $\overline{Y}_b := \overline{Y} \times_{\overline{X}} \overline{X}_b$. But σ also acts trivially on $\overline{Y}_{b,\infty}^{aux}$. Since $\overline{Y}_{b,\infty}$ is induced from $\overline{Y}_{b,\infty}^{aux}$, σ acts trivially on $\overline{Y}_{b,\infty}$. Therefore, σ acts trivially on $\overline{Y}_b$. $\square$

The auxiliary cover f^{aux} is often simpler to work with than f due to the following:

Proposition 7.3. *If we assume that a p-Sylow subgroup of G is* cyclic, *then the group G^{aux} has a normal subgroup of order p.*

Proof. Let $\overline{S}$ be the union of all inseparable components of $\overline{X}$. By construction, the inverse image $\overline{V}$ of $\overline{S}$ in $\overline{Y}^{aux}$ is connected, and its decomposition group is G^{aux}. By [15, Corollary 2.12], G^{aux} has a normal subgroup of order p. $\square$

Lastly, in the case that a p-Sylow subgroup of G is cyclic, we make a further simplification of the auxiliary cover, as in [19, Remarque 3.1.8]. Since G^{aux} has a normal subgroup of order p, [15, Corollary 2.4 (i)] shows that the quotient G^{str} of G^{aux} by its maximal normal prime-to-p subgroup N is isomorphic to $\mathbb{Z}/p^n \rtimes \mathbb{Z}/m_{G_{aux}}$, where the action of $\mathbb{Z}/m_{G_{aux}}$ on $\mathbb{Z}/p^n$ is faithful. Note that $m_{G_{aux}} | m_G$. Then $Y^{str} := Y^{aux}/N$ is a branched G^{str}-cover of X, called the *strong auxiliary cover*. Constructing the strong auxiliary cover is one of the key places where it is essential to assume that a p-Sylow subgroup of G is cyclic, as otherwise G^{aux} does not necessarily have such nice group-theoretical properties.

The branching on the generic fiber of the strong auxiliary cover is as follows: At each point of X where the branching index of f was divisible by p, the branching index of f^{str} is a power of p (as G^{str} has only elements of p-power order and of prime-to-p order). At each branch point specializing to an étale tail $b \in B_{\text{ét}}$, the ramification index is m_b, where $m_b | (p-1)$ (cf. [19, §3.3.2]).

The following lemma shows that it will generally suffice to look at the strong auxiliary cover instead of the auxiliary cover.

Lemma 7.4. *Let $f : Y \to X$ be as in §4. Let L be a field over which both the stable model of f^{str} and all the branch points of $Y^{aux} \to Y^{str}$ are defined. Then the stable model f^{st} of f can be defined over a tame extension of L.*

Proof. Since $Y^{str} = Y^{aux}/N$, and $p \nmid |N|$, it follows from [12, Proposition 6.2] that $(f^{aux})^{st}$ can be defined over a tame extension of L. By Lemma 7.2, so can f^{st}. $\square$

While the Galois group of the (strong) auxiliary cover is simpler than the original Galois group of f, we generally are made to pay for this with the introduction of new branch points. Understanding where these branch points appear is key to understanding the minimal field of definition of the stable reduction of the auxiliary cover.

8. Proof of the Main Theorem

In this section, we will prove Theorem 1.1. Let k be an algebraically closed field of characteristic p, let $R_0 = W(k)$, and let $K_0 = \mathrm{Frac}(R_0)$. Note that if $k \cong \overline{\mathbb{F}}_p$, then $K_0 \cong \mathbb{Q}_p^{ur}$. Also, for all $i > 0$, we set $K_i = K_0(\zeta_{p^i})$, where ζ_{p^i} is a primitive p^ith root of unity.

Let $f : Y \to X \cong \mathbb{P}^1$ be a three-point G-cover defined over $\overline{K_0}$, where G has a cyclic p-Sylow subgroup of order p^n and $m_G = 2$. Since $m_G | (p-1)$ ([15, Lemma 2.1]), we may (and do) assume throughout this section that $p \neq 2$. We break this section up into the cases where the number τ of branch points of $f : Y \to X = \mathbb{P}^1$ with prime-to-p branching index is 1, 2, or 3. By Lemma 4.3, if f has bad reduction, then τ is the number of primitive tails of the stable reduction. The cases $\tau = 2$ and $\tau = 3$ are quite easy, whereas the case $\tau = 1$ is much more involved. This stems from the appearance of new tails in the stable reduction of f in the case $\tau = 1$. The ideas in the proof of the $\tau = 1$ case should work as well in the $\tau = 0$ case, but the computations will be more difficult. See Question 9.1.

We mention that, because any finite extension of K/K_0 has cohomological dimension 1, then if K is the field of moduli of a G-Galois cover relative to K_0, it is also a field of definition ([3, Proposition 2.5]).

8.1. The case $\tau = 3$.

Proposition 8.1. *Assume $f : Y \to X$ is a three-point G-cover defined over $\overline{K_0}$ where G has a cyclic p-Sylow subgroup P with $m_G = |N_G(P)/Z_G(P)| = 2$. Suppose that all three branch points of f have prime-to-p branching index. Then f has potentially good reduction. Additionally, f has a model defined over K_0, and thus the field of moduli of f relative to K_0 is K_0.*

Proof. Suppose f has bad reduction. We know that the stable reduction must have three primitive tails. But this contradicts Corollary 4.10. So f has potentially good reduction.

Let $\overline{f} : \overline{Y} \to \overline{X}$ be the reduction of f over k. Then f is tamely ramified. By [6, Theorem 4.10], if R is the ring of integers of any finite extension K/K_0, then there exists a unique deformation f_R of $\overline{f}$ to a cover defined over R. It follows that f_{R_0} exists, and $f_{R_0} \otimes_{R_0} R \cong f_R$. Thus $f_{R_0} \otimes_{R_0} K_0$ is the model we seek. $\square$

Proposition 8.1, while an easy consequence of the vanishing cycles formula, gives a proof that the modular curve $X(N)$ has good reduction to characteristic p for many p, without relying on its modular interpretation (see, for instance, [5]).

Corollary 8.2. *Let $N \in \mathbb{N}$ have prime factorization $N = \prod_{i=1}^{r} p_i^{a_i}$. Let M be the product of all primes that divide $p_i^2 - 1$ for more than one i. Then the modular curve $X(N)$ has good reduction at all primes not dividing $6NM$.*

Proof. The modular curve $X(N)$ can be realized (via the j-function) as a $PSL_2(\mathbb{Z}/N)$-cover $f : X(N) \to X(1) \cong \mathbb{P}^1$, branched at three points of index 2, 3, and N, respectively. Let $G = PSL_2(\mathbb{Z}/N)$, and let p be a prime dividing $|G|$ but not dividing $6NM$. By the Chinese remainder theorem, one sees that $G \cong \prod_i PSL_2(\mathbb{Z}/p_i^{a_i})$. Furthermore, for each i, we have an exact sequence

$$1 \to P_i \to PSL_2(\mathbb{Z}/p_i^{a_i}) \to PSL_2(p_i) \to 1,$$

where P_i is a p_i-group and the third map is the modulo p_i projection on matrix entries. The order of $PSL_2(p_i)$ is $p_i(p_i^2-1)/2$, and it is well known that $PSL_2(p_i)$ contains cyclic subgroups of order $\frac{p_i+1}{2}$ and $\frac{p_i-1}{2}$. By the Schur-Zassenhaus theorem, these subgroups lift to $PSL_2(\mathbb{Z}/p_i^{a_i})$. In particular, the p-Sylow subgroup of $PSL_2(\mathbb{Z}/p_i^{a_i})$ is cyclic.

Since $p \nmid 6NM$, we have that p divides the order of exactly one $PSL_2(p_i)$. In particular, the p-Sylow subgroup of G is cyclic. It is well known that $m_{PSL_2(p_i)} = 2$ for this p_i (relative to the prime p), and thus the same is true for m_G. Since $p \nmid 6N$, Proposition 8.1 shows that the cover f, and thus $X(N)$, has good reduction to characteristic p. $\square$

8.2. The case $\tau = 2$. If there are exactly two branch points with prime-to-p branching index, then f has bad reduction (f cannot have good reduction because it will have a branch point with p dividing the branching index). We use the notation of §4. In particular, $f^{st} : Y^{st} \to X^{st}$ is the stable model of f and $\overline{f} : \overline{Y} \to \overline{X}$ is the stable reduction.

Proposition 8.3. *Assume $f : Y \to X$ is a three-point G-cover defined over $\overline{K_0}$ where G has a cyclic p-Sylow subgroup P of order p^n with $m_G = |N_G(P)/Z_G(P)| = 2$. Suppose that two of the three branch points of f have prime-to-p branching index. Then the stable model of f can be defined over a tame extension K of K_n. In particular, f can be defined over K. Thus the field of moduli of f relative to K_0 is contained in a tame extension of K_n.*

Proof. We know that $\overline{X}$ must have two primitive tails, and Corollary 4.10 shows that there are no new tails. The effective ramification invariant for each of the primitive tails is $\frac{1}{2}$ by the vanishing cycles formula (4.1). Then the strong auxiliary cover $f^{str} : Y^{str} \to X$ is a three-point $\mathbb{Z}/p^{\nu} \rtimes \mathbb{Z}/2$-cover, for some $\nu \leq n$.

By [12, Proposition 7.6], the stable model of f^{str} is defined over a tame extension K^{str} of $K_{\nu} \subseteq K_n$. Since the branch loci of f, f^{aux}, and f^{str} are each the same three points, all branch points of the canonical map $Y^{aux} \to Y^{str}$ are ramification points of f^{str}. The ramification points of f^{str} specialize to distinct points on $\overline{Y}^{str}$, so $G_{K^{str}}$ permutes them trivially. Thus they are defined over K^{str}. By Lemma 7.4, the stable model of f is defined over a tame extension K of K^{str}. □

8.3. The case $\tau = 1$. Now we consider the case where only one point, say 0, has prime-to-p branching index. As in the case $\tau = 2$, the cover f has bad reduction. The goal of this (rather lengthy) section is to prove the following proposition:

Proposition 8.4. *Assume $f : Y \to X$ is a three-point G-cover defined over $\overline{K_0}$ where G has a cyclic p-Sylow subgroup P with $m_G = |N_G(P)/Z_G(P)| = 2$ and $p \neq 3$. Suppose that exactly one of the three branch points of f has prime-to-p branching index. Then the stable model of f can be defined over a finite extension K/K_0 such that the nth higher ramification groups for the upper numbering for (the Galois closure of) K/K_0 vanish. In particular, f can be defined over such a K. Thus the nth higher ramification group for the upper numbering for the field of moduli of f relative to K_0 vanishes.*

We mention that, because $m_G = 2$, the stable reduction of f is monotonic (Proposition 6.4).

8.3.1. We first deal with the case where there is one primitive tail $\overline{X}_b$, but no new étale tails. Then the vanishing cycles formula (4.1) shows that $\sigma_b = 1$. Furthermore, we claim that $m_{G^{aux}} = 1$. If this were not the case, then the strong auxiliary cover would have Galois group $G^{str} \cong \mathbb{Z}/p^{\nu} \rtimes \mathbb{Z}/2$, for some $\nu \leq n$, but only one branch point with prime-to-p branching index. Then taking the quotient by $\mathbb{Z}/p^{\nu}$ would yield a contradiction.

Since we are assuming that the stable reduction of f has no new tails, the auxiliary cover $f^{aux} : Y^{aux} \to X$ is branched at either two or three points. If it is branched at three points, we conclude using [12, Proposition 7.15] that the stable model of f^{aux} (which is the modified stable model) is defined over some K such that the nth higher ramification groups of the extension K/K_0 vanish. If f^{aux} is branched at two points (without loss of generality, 0 and ∞), it is a cyclic cover, and thus clearly defined over K_n. Furthermore, the points in the fiber above 1 are defined over K_n. We conclude that the modified stable model of f^{aux} is defined over K_n (Remark 7.1). By [21, IV, Corollary to Proposition 18], the nth higher ramification groups of K_n/K_0 vanish. By Lemma 7.2, Proposition 8.4 is true in this case.

8.3.2. We now come to the main case, where there is a new étale tail $\overline{X}_b$ and a primitive tail $\overline{X}_{b'}$. We will assume for the remainder of §8.3 that $p \neq 3$ (although it is likely that the main result should hold in the case $p = 3$, see Question 9.1).

Fix, once and for all, a coordinate x corresponding to the smooth model X_{R_0} with special fiber $\overline{X}_0$, so that f is branched at $x = 0$, $x = 1$, and $x = \infty$. By the vanishing cycles formula (4.1), the new tail $\overline{X}_b$ has $\sigma_b = 3/2$ and the primitive tail $\overline{X}_{b'}$ has $\sigma_{b'} = 1/2$. It is then clear that the auxiliary cover has four branch points: at $x = 0$, $x = 1$, $x = \infty$, and $x = a$, where a is in the disk corresponding to $\overline{X}_b$. Keep in mind that, by the construction of the auxiliary cover (§7), we may always replace a by any other point in the disk corresponding to $\overline{X}_b$. Also, the modified stable model of the auxiliary cover is, in fact, the stable model. The strong auxiliary cover then has Galois group $G^{str} \cong \mathbb{Z}/p^\nu \rtimes \mathbb{Z}/2$ for some $\nu \leq n$. Without loss of generality, we can assume that 0 and a are branched of index 2, and 1 and ∞ are branched of p-power index. After a possible application of the transformation $x \to \frac{x}{x-1}$ of $\mathbb{P}^1$, which interchanges 1 and ∞ while fixing 0, we may and do further assume that a does not collide with ∞ on the smooth model of X corresponding to the coordinate x (i.e., $|a| \leq 1$).

Lemma 8.5. *At least one point of f^{str} is branched of index p^ν. Such a point specializes to the original component.*

Proof. Consider the $\mathbb{Z}/p \rtimes \mathbb{Z}/2$-cover $f' := Y^{str}/Q \to X$, where Q has order $p^{\nu-1}$. This must be branched at at least three points, thus at 1 or ∞. If 1 or ∞ is a branch point of f', then its branching index in f is p^ν. By Lemma 4.3, any branch point of index p^ν specializes to a p^ν-component of $\overline{X}$. By Lemma 6.1, it specializes to $\overline{X}_0$. □

Let us fix some additional notation for §8 by writing down the equations of the cover $f^{str} : Y^{str} \to X^{str}$. Let $Z^{str} = Y^{str}/(\mathbb{Z}/p^\nu)$. Then $Z^{str} \to X^{str}$ is a degree 2 cover of $\mathbb{P}^1$'s, branched at 0 and a. Therefore, Z^{str} can be given (birationally) over $\overline{K_0}$ by the equation

$$z^2 = \frac{x-a}{x}. \tag{8.1}$$

Fix a choice of $\sqrt{1-a}$. Since $z = \pm 1$ (resp. $\pm\sqrt{1-a}$) corresponds to $x = \infty$ (resp. $x = 1$), then $Y^{str} \to Z^{str}$ can be given (birationally) over $\overline{K_0}$ by the equation

$$y^{p^\nu} = g(z) := \left(\frac{z+1}{z-1}\right)^r \left(\frac{z+\sqrt{1-a}}{z-\sqrt{1-a}}\right)^s \tag{8.2}$$

for some integers r and s, which are well-defined modulo p^ν. Without loss of generality, we take $0 < r, s < p^\nu$. The branching index of f^{str} at ∞ is $p^{\nu - v(r)}$, and at 1 it is $p^{\nu - v(s)}$.

Write $\overline{Z}^{str}$ for $\overline{Y}^{str}/(\mathbb{Z}/p^\nu)$, and let $\overline{Z}_b$ (resp. $\overline{Z}_{b'}$) be the unique irreducible component of $\overline{Z}^{str}$ above the new tail $\overline{X}_b$ (resp. the primitive tail $\overline{X}_{b'}$).

We will work over a large enough finite extension K/K_0 (i.e., we assume the stable model of f^{str} is defined over K and we replace K by a finite extension whenever convenient). Let $e \in K$ be such that $|e|$ is the radius of the disk $\mathcal{D}$ corresponding to $\overline{Z}_b$. Since $x = a$ corresponds to $z = 0$, we can choose a coordinate t on the disk $\mathcal{D}$ such that $z = et$. If $\hat{Y}$, $\hat{Z}$ are the formal completions of $(Y^{str})^{st}$ and $(Z^{str})^{st}$ along their special fibers, then the torsor $\hat{Y} \times_{\hat{Z}} \mathcal{D} \to \mathcal{D}$ can be given generically, after a possible finite extension of K, by the equation

$$y^{p^\nu} = 1 + \frac{g'(0)}{1!}(et) + \frac{g''(0)}{2!}(et)^2 + \cdots . \tag{8.3}$$

Now, since $\sigma_b = \frac{3}{2}$, and since $\overline{X}_b$ intersects a p-component (see Lemma 8.6 below), we know that the generic fiber of this torsor must split into $p^{\nu-1}$ connected components, each of which has étale reduction and is birationally equivalent to a $\mathbb{Z}/p$-cover branched at one point with conductor 3. Let $c_i = \frac{g^{(i)}(0)}{i!}e^i$. Then

$$y^{p^\nu} = 1 + c_1 t + c_2 t^2 + \cdots . \tag{8.4}$$

Note that we have fixed the meaning of the symbols f, p, k, a, r, s, n, ν, x, y, z, e, t, c_i, $g(z)$, G, σ_b, $\sigma_{b'}$, $\overline{X}_b$, $\overline{X}_{b'}$, $\overline{X}_0$, $\overline{Z}_b$, and $\overline{Z}_{b'}$. We will also use the notation $\mathcal{G}$ and $\mathcal{G}'$ for the dual graph and augmented dual graph of $\overline{X}$ (§4.1).

The idea of the proof is as follows: In order to place bounds on the higher ramification filtration of the field of moduli of f, it suffices by the results of §7 to understand the minimal field of definition of the stable model of f^{str}. In order to do this, we must first calculate the disk corresponding to the new tail $\overline{X}_b$. Since we can take a to be any value in this disk, we choose the value defined over the "smallest" field possible to be our a, and then $K_n(a, \sqrt{1-a})$ will be a field of definition of f^{str}. This is done in the first large subsection, §8.3.4.

Since understanding the monodromy action is enough to pin down a field of definition of the stable model of f^{str}, the goal is then to determine the monodromy action of $\mathrm{Gal}(\overline{K_0}/K_0)$ on $\overline{f}^{str}$. We use the criterion of [12, Proposition 4.9], which essentially says that if this action fixes the tails of $\overline{Y}^{str}$, then it fixes all of $\overline{Y}^{str}$. To this end, in §8.3.5 (our second large subsection), we show exactly which disks correspond to inseparable tails of $\overline{X}$ (if there are any).

In §8.3.6, we put all of the information from §8.3.4 and §8.3.5 together to determine a field of definition for the stable model of f (up to tame extension). Lastly, in §8.3.7, we show that the appropriate higher ramification groups vanish.

We start with §8.3.3, where we show some basic properties of $(f^{str})^{st}$ and prove a couple of algebraic results that will be used later.

8.3.3. *Preliminary lemmas.*

Lemma 8.6. *Every étale tail $\overline{X}_c$ of $\overline{X}$ intersects a p-component.*

Proof. By Lemma 4.5, $\overline{X}_c$ intersects an inseparable component. If $\overline{X}_c$ intersects a p^{α}-component, with $\alpha > 1$, then [15, Lemma 4.2] shows that $\sigma_b \geq p/2$. Since $p/2 > 2$, this contradicts the vanishing cycles formula (4.1). ☐

Lemma 8.7. *The map f^{str} is branched at $x = \infty$ of index p^{ν}, and $x = \infty$ specializes to the original component $\overline{X}_0$. If $v(a-1) = 0$, then f^{str} is branched at $x = 1$ of index p^{ν}, and $x = 1$ also specializes to $\overline{X}_0$.*

Proof. Assume for a contradiction that ∞ is not branched of index p^{ν}. Then 1 is branched of index p^{ν}, and specializes to $\overline{X}_0$ by Lemma 6.1. Thus, the deformation data above $\overline{X}_0$ are multiplicative by Lemma 5.1, and identical by [12, Proposition 5.2]. By assumption and by Lemma 4.3, ∞ does not specialize to the original component. Then consider the unique point $\overline{x} \in \overline{X}_0$ such that the specialization $\overline{\infty}$ of ∞ lies outward from $\overline{x}$. Since $|a| \leq 1$, there is no étale tail lying outward from $\overline{x}$. If $\varepsilon \in E(\mathcal{G}')$ corresponds to $(\overline{x}, \overline{X}_0, \overline{W})$, for some $\overline{W}$, then Lemma 5.5 shows that $\sigma_{\varepsilon}^{\text{eff}} = 0$. But this means that $\sigma_{\varepsilon} = 0$ for each deformation datum above $\overline{X}_0$, which contradicts Lemma 5.1. We have thus shown that ∞ is branched of index p^{ν}. By Lemma 6.1, ∞ specializes to $\overline{X}_0$.

Now suppose $v(a-1) = 0$. Assume for a contradiction that 1 does not specialize to the original component. Consider the unique point $\overline{x} \in \overline{X}_0$ such that the specialization $\overline{1}$ of 1 lies outward from $\overline{x}$, and let $\varepsilon \in E(\mathcal{G}')$ correspond to $(\overline{x}, \overline{X}_0, \overline{W})$ for some $\overline{W}$. As in the previous paragraph, $\sigma_{\varepsilon} = 0$ for each deformation datum above $\overline{X}_0$, and we get a contradiction. ☐

Corollary 8.8. *All deformation data above the original component are multiplicative. In particular, $\delta_{\overline{X}_0}^{\text{eff}} = \nu + \frac{1}{p-1}$.*

Proof. By Lemma 8.7, $x = \infty$ specializes to the original component $\overline{X}_0$. By Lemma 5.1, all deformation data above $\overline{X}_0$ are multiplicative. ☐

Lemma 8.9. *Let $c = \alpha + \frac{\beta}{\sqrt{1-a}}$, where $\alpha, \beta, a \in K$ and $\sqrt{1-a}$ means either square root. Let $a_0 = 1 - \left(\frac{\beta}{\alpha}\right)^2$. If $v(c) > 0$ and $v(\alpha) = 0$, then $v(a - a_0) = v(c) + 2v(\beta)$. Note that $a_0 \in K_0(\alpha, \beta)$.*

Proof. Solving for a, we find that $a = 1 - \left(\frac{\beta}{c-\alpha}\right)^2$. Choose $a_0 = 1 - \left(\frac{\beta}{\alpha}\right)^2$. Then

$$a - a_0 = \beta^2 \left(\frac{2c\alpha - c^2}{\alpha^2(\alpha - c)^2}\right).$$

Clearly, $v(a - a_0) = 2v(\beta) + v(c)$. □

For positive integers q and $r_1, \ldots, r_n$ such that $\sum_i r_i = q$, define

$$\binom{q}{r_1, \ldots, r_n} = \frac{q!}{r_1! \cdots r_n!}.$$

We leave the proof of the following lemma to the reader:

Lemma 8.10. *For any prime p,*

$$v_p\left(\binom{q}{r_1, \ldots, r_n}\right) \geq \max_i \left(v_p\left(\binom{q}{r_i}\right)\right) = v_p(q) - \min_i(v_p(r_i))$$

(here $\binom{q}{r_i}$ is the standard binomial coefficient).

8.3.4. ***The new (étale) tail.*** An open p-adic disk is determined by its radius and any point inside. The disk corresponding to the new tail $\overline{X}_b$ is centered at a. The following lemma determines its radius.

Lemma 8.11. *Let ρ (resp. e) be an element of K such that $|\rho|$ (resp. $|e|$) is the radius of the disk centered at $x = a$ corresponding to $\overline{X}_b$ (resp. the disk centered at $z = 0$ corresponding to $\overline{Z}_b$).*

(1) *If $v(a) = v(a-1) = 0$, then $v(\rho) = \frac{2}{3}(\nu + \frac{1}{p-1})$ and $v(e) = \frac{1}{3}(\nu + \frac{1}{p-1})$.*
(2) *If $v(a) > 0$, then $v(\rho) = \frac{2}{3}(\nu + \frac{1}{p-1}) + \frac{1}{3}v(a)$ and $v(e) = \frac{1}{3}(\nu + \frac{1}{p-1} - v(a))$.*
(3) *If $v(a-1) > 0$, then $v(\rho) = \frac{2}{3}(\nu + \frac{1}{p-1} + v(1-a))$ and $v(e) = \frac{1}{3}(\nu + \frac{1}{p-1} + v(1-a))$.*

Proof. Since $z^2 = \frac{x-a}{x}$, then for any z, $v(z) = \frac{1}{2}(v(x-a) - v(x))$. Since $\overline{X}_b$ is a new tail, $x = 0$ does not specialize to the corresponding disk. So for any x in this disk, $v(x-a) > v(a)$, thus $v(x) = v(a)$. This shows that $v(z) = \frac{1}{2}(v(x-a) - v(a))$ in this disk, and thus $v(e) = \frac{1}{2}(v(\rho) - v(a))$. Therefore, it suffices to prove the statements about $v(\rho)$.

Consider the path $\{v_i\}_{i=0}^{j}$, $\{e_i\}_{i=0}^{j-1}$, where v_0 corresponds to $\overline{X}_0$ and v_j corresponds to $\overline{X}_b$ (§4.1). Write ϵ_i (resp. σ_i^{eff}, δ_i^{eff}) for ϵ_{e_i} (resp. $\sigma_{e_i}^{\mathrm{eff}}$, $\delta_{v_i}^{\mathrm{eff}}$) (see Definition 5.2). Then $\delta_0 = \nu + \frac{1}{p-1}$, whereas $\delta_j = 0$.

To (1): Suppose $v(a) = v(a-1) = 0$. Then $\overline{X}_b$ is the only étale tail lying outward from the point $\overline{x}_0$ corresponding to e_0. No branch points with branching index divisible by p lie outward from $\overline{x}_0$, either. By Lemma

5.5, $\sigma_i^{\mathrm{eff}} = \frac{3}{2}$ for all $0 \leq i < j$. By applying Lemma 5.4 (3) to each e_i, $0 \leq i < j$, we obtain $v(\rho) = \frac{2}{3}(\nu + \frac{1}{p-1})$.

To (2): Suppose $v(a) > 0$. In order to separate the specializations of $x = a$ and $x = 0$ on the special fiber, there must be a component $\overline{W}$ of $\overline{X}$ corresponding to the closed disk of radius $|a|$ and center 0 (or equivalently, center a). Suppose $\overline{W}$ corresponds to v_{i_0}. Then, for $i < i_0$, Lemma 5.5 shows that $\sigma_i^{\mathrm{eff}} = 1$. For $i \geq i_0$, Lemma 5.5 shows that $\sigma_i^{\mathrm{eff}} = \frac{3}{2}$. By construction, we have $\sum_{i=0}^{i_0-1} \epsilon_i = v(a)$. Applying Lemma 5.4 (3) to each of the edges $e_0, \ldots, e_{i_0-1}$, we see that $\delta_{\overline{W}}^{\mathrm{eff}} = \nu + \frac{1}{p-1} - v(a)$. Then, applying Lemma 5.4 (3) to each of the edges $e_{i_0}, \ldots, e_{j-1}$, we see that $\sum_{i=i_0}^{j-1} \epsilon_i = \frac{2}{3}(\nu + \frac{1}{p-1} - v(a))$. So $v(\rho) = \sum_{i=0}^{j-1} \epsilon_i = \frac{2}{3}(\nu + \frac{1}{p-1}) + \frac{1}{3}v(a)$.

To (3): Suppose $v(a-1) > 0$. In order to separate the specializations of $x = a$ and $x = 1$ on the special fiber, there must be a component $\overline{W}$ of $\overline{X}$ corresponding to the closed disk of radius $|1-a|$ and center 1 (or equivalently, center a). Suppose $\overline{W}$ corresponds to v_{i_0}. Then, for $i < i_0$, Lemma 5.5 shows that $\sigma_i^{\mathrm{eff}} = \frac{1}{2}$. For $i \geq i_0$, Lemma 5.5 shows that $\sigma_i^{\mathrm{eff}} = \frac{3}{2}$. By construction, we have $\sum_{i=0}^{i_0-1} \epsilon_i = v(1-a)$. Applying Lemma 5.4 (3) to each of the edges $e_0, \ldots, e_{i_0-1}$, we see that $\delta_{\overline{W}}^{\mathrm{eff}} = \nu + \frac{1}{p-1} - \frac{1}{2}v(1-a)$. Then, applying Lemma 5.4 to each of the edges $e_{i_0}, \ldots, e_{j-1}$, we see that $\sum_{i=i_0}^{j-1} \epsilon_i = \frac{2}{3}(\nu + \frac{1}{p-1} - \frac{1}{2}v(1-a))$. So $v(\rho) = \sum_{i=0}^{j-1} \epsilon_i = \frac{2}{3}(\nu + \frac{1}{p-1} + v(1-a))$. □

We now determine a point a_0 inside the disk corresponding to $\overline{X}_b$. It will turn out that a_0 (thus the disk) is uniquely determined by p, r, and s. We choose a_0 so that it is defined over as small an extension of K_0 as possible. Our strategy will be to look at equations (8.3) and (8.4), understand the dependence of the coefficients $c_i = \frac{g^{(i)}(0)}{i!}e^i$ on a, and use Lemma 2.1 to show that only for certain choices of a can the torsor given by (8.4) split into $p^{\nu-1}$ copies of an Artin-Schreier cover.

The main idea of the argument is completely present when $v(a) = v(a-1) = 0$ (Proposition 8.12, the simplest case). Unfortunately, calculational difficulties make this idea more difficult to implement when $v(a) > 0$ or $v(a-1) > 0$ (Propositions 8.19 and 8.26), and the arguments are much longer. Furthermore, when $p = 5$, we will have to use Lemma 2.1 (2) instead of Lemma 2.1 (1), which obscures the main idea even more. Thus, on a first reading, the reader might choose to read only §8.3.4.1, as well as the statements of Propositions 8.19 and 8.26, before moving on to §8.3.5.

Note that our choice of a does not affect which case we are in, as the radius of the disk corresponding to $\overline{X}_b$ is less than 1.

8.3.4.1. ***The case*** $\mathbf{v(a) = v(a-1) = 0}$.

Proposition 8.12. *If $v(a) = v(a-1) = 0$, then the disk Δ corresponding to $\overline{X}_b$ contains the K_0-rational point $x = a_0$, where $a_0 = 1 - \frac{s^2}{r^2}$.*

Proof. Recall that we use the notation of (8.3) and (8.4). We know $v(e) = \frac{1}{3}(\nu + \frac{1}{p-1})$ by Lemma 8.11. Since $g^{(i)}(0)/i!$ is the coefficient of z^i in the Maclaurin series expansion of g, and since $v(\sqrt{1-a}) = 0$, we obtain that $v(\frac{g^{(i)}(0)}{i!}) \geq 0$. Since $c_i = \frac{g^{(i)}(0)}{i!}e^i$, we have $v(c_i) \geq iv(e) = \frac{i}{3}(\nu + \frac{1}{p-1})$. It follows that for $p|i$ (and $i \neq 0$), $v(c_i) > \nu + \frac{1}{p-1}$. So for the torsor given by (8.4) to split into $p^{\nu-1}$ disjoint μ_p-torsors with étale reduction, Lemma 2.1 (1) says that we must have $v(c_1) \geq \nu + \frac{1}{p-1}$. In particular, we must have $v(g'(0)) \geq \frac{2}{3}(\nu + \frac{1}{p-1})$. A calculation shows that

$$g'(0) = 2r + \frac{2s}{\sqrt{1-a}}.$$

Since the branching index of 1 is p^ν, s is a unit modulo p and $v(2s) = 0$. Since $v(g'(0)) \geq \frac{2}{3}(\nu + \frac{1}{p-1})$, Lemma 8.9 (with $c = g'(0)$, $\alpha = 2r$, $\beta = 2s$) shows that $a_0 = 1 - \frac{s^2}{r^2} \in K_0$ satisfies $v(a - a_0) \geq \frac{2}{3}(\nu + \frac{1}{p-1})$. But by Lemma 8.11, $v(\rho) = \frac{2}{3}(\nu + \frac{1}{p-1})$, where $|\rho|$ is the radius of Δ. So $a_0 \in \Delta$. $\square$

Remark 8.13. In fact, if a is as in Proposition 8.12, we have $v(c_3) = \nu + \frac{1}{p-1}$ and $v(c_i) > \nu + \frac{1}{p-1}$ for $i \neq 3$. By [12, Lemma 3.1 (i)], the torsor given by (8.4) indeed splits into $p^{\nu-1}$ disjoint μ_p-torsors with étale reduction birational to an Artin-Schreier cover branched at one point of conductor 3.

8.3.4.2. ***The case*** $\mathbf{v(a) > 0}$. When $v(a) > 0$, the required calculations are somewhat more involved. By Lemma 8.11, we have $v(e) = \frac{1}{3}(\nu + \frac{1}{p-1} - v(a))$ in this case.

Lemma 8.14. *We have $v(a) = v(r+s) \leq \nu - 1$. In particular, $v(a) \in \mathbb{Z}$.*

Proof. The cover $\overline{Y}^{str} \to \overline{Z}^{str}$ splits completely above the specialization $\overline{z}$ of $z = 0$. Recall that t is a coordinate on the disk corresponding to $\overline{Z}_b$, so that $z = et$. Then $\overline{z}$ corresponds to the open disk $|t| < 1$, and [18, Proposition 3.2.3 (2)] shows that this disk splits into p^ν disjoint copies in $\overline{Y}^{str}$. In particular, $g(et)$ is a p^νth power in $R[[t]]$. If $\sum \alpha_i t^i$ is a power series in $R[[t]]$ that is a p^νth power, the coefficient of t must be divisible by p^ν. So the coefficient c_1 of t in $g(et)$, which is $g'(0)e$, has valuation at least ν, and thus $v(g'(0)) \geq \nu - v(e) = \frac{2}{3}\nu + \frac{1}{3}v(a) - \frac{1}{3(p-1)}$.

On the other hand, $g'(0) = 2r + \frac{2s}{\sqrt{1-a}}$, which can be written as

$$g'(0) = 2r + 2s(1 + \frac{a}{2} + O(a^2)) = 2(r+s) + sa + s(O(a^2)), \tag{8.5}$$

where $O(a^2)$ represents terms whose valuation is at least $2v(a)$. If we assume for the moment that $v(a) < \nu - \frac{1}{2(p-1)}$, then we must have

$$v(g'(0)) \geq \frac{2}{3}\nu + \frac{1}{3}v(a) - \frac{1}{3(p-1)} > v(a).$$

Since $v(a^2) > v(a)$, this means $v(2(r+s) + sa) > v(a)$, so

$$v(r+s) = v(sa) = v(a)$$

($v(s) = 0$, by Lemma 8.7). Since $v(r+s) \in \mathbb{Z}$, we have

$$v(a) = v(r+s) \leq \nu - 1.$$

If instead, we assume that $v(a) \geq \nu - \frac{1}{2(p-1)}$, then

$$v(g'(0)) \geq \frac{2}{3}\nu + \frac{1}{3}v(a) - \frac{1}{3(p-1)} > \nu - 1.$$

So $v(2(r+s)) = v(r+s) \geq \nu$ by (8.5).

It remains to show that we cannot have both $v(a) \geq \nu - \frac{1}{2(p-1)}$ and $v(r+s) \geq \nu$. Suppose, for a contradiction, that this is the case. Then, multiplying $g(z)$ by $\left(\frac{z-\sqrt{1-a}}{z+1}\right)^{r+s}$, which is a p^νth power, we obtain the alternative equation

$$y^{p^\nu} = \left(\frac{z+\sqrt{1-a}}{z+1}\right)^s \left(\frac{z-\sqrt{1-a}}{z-1}\right)^r$$

to (8.2). Consider the unique component $\overline{V}$ of $\overline{Z}^{str}$ above $\overline{W}$, the component of $\overline{X}$ corresponding to the disk of radius $|a|$ around $x = 0$. Then $\overline{V}$ corresponds to the coordinate z. The formal completion of $\overline{V} \backslash \{z = \pm 1\}$ in Z^{str} is isomorphic to Spec C where

$$C := R\{(z-1)^{-1}, (z+1)^{-1}\}.$$

We have

$$\frac{z+\sqrt{1-a}}{z+1} = 1 + (\sqrt{1-a} - 1)(z+1)^{-1}.$$

Since

$$v(\sqrt{1-a} - 1) = v(a) > \nu - 1 + \frac{1}{p-1},$$

this is a $p^{\nu-1}$st power in C (which follows from the binomial expansion). Likewise, $\left(\frac{z-\sqrt{1-a}}{z-1}\right)$ is a $p^{\nu-1}$st power in C. So

$$\left(\frac{z+\sqrt{1-a}}{z+1}\right)^s \left(\frac{z-\sqrt{1-a}}{z-1}\right)^r$$

is a $p^{\nu-1}$st power in C. But this means that there are at least $p^{\nu-1}$ irreducible components in the inverse image of $\overline{V}\backslash\{z=\pm 1\}$ in $\overline{Y}^{str}$, and thus that many irreducible components of $\overline{Y}^{str}$ above $\overline{V}$.

Now, $\overline{W}$ is not a tail, so it is not an étale component by Lemma 4.4. So $\overline{W}$ must be a p-component. Furthermore, it cannot intersect a p^2-component, because the inertia groups above the intersection point would have order divisible by p^2, and then there could not be $p^{\nu-1}$ irreducible components above $\overline{V}$. So $\overline{Y}$ has $p^{\nu-1}$ irreducible components above $\overline{V}$, each a radicial extension of degree p. Associated to each is one deformation datum. It has three critical points: two at the intersection of $\overline{W}$ with outward-lying components, and one at the intersection of $\overline{W}$ with inward lying components. By Lemma 5.5, the first two critical points have invariants $3/2$ and $1/2$, and [23, p. 998, (2)] shows that the third has invariant -1. Since no multiplicative deformation datum can have -1 for an invariant, the deformation datum must be additive. But this contradicts [22, Proposition 2.8], proving the lemma. □

Remark 8.15. Armed with the knowledge that $v(a)=v(r+s)$, we can run the argument of the second-to-last paragraph of the proof again to see that $g(z)$ is a $p^{v(a)-1}$st power in C, and thus there are at least $p^{v(a)-1}$ irreducible components of $\overline{Y}^{str}$ above $\overline{V}$. In particular, $\overline{W}$ is a p^i-component for $i\leq \nu-v(a)+1$.

Lemma 8.16. *The valuation $v\left(\frac{g^{(i)}(0)}{i!}\right)$ is at least $v(a)-v(i)$.*

Proof. Note that $\frac{g^{(i)}(0)}{i!}$ is the coefficient of z^i in the Maclaurin series expansion of $g(z)$. Since $v(a)\geq v(a)-v(i)$, it suffices to look modulo a. By (8.2), $g(z)$ is congruent (mod a) to

$$\left(\frac{z+1}{z-1}\right)^{r+s}=\left(1+\frac{2}{z-1}\right)^{r+s}=(-1-2z-2z^2-2z^3-\cdots)^{r+s}. \tag{8.6}$$

Expanding out the above expression gives

$$(-1)^{r+s}\left(1+\sum_{i=1}^{\infty}\sum_{\substack{I=\{i_1,\dots,i_q\}\subset\mathbb{N}\\ A=(a_1,\dots,a_q)\in\mathbb{N}^q\\ \sum_{j=1}^{q}a_j\le r+s\\ \sum_{j=1}^{q}a_ji_j=i}} 2^{|\sum_{j=1}^{q}a_j|}\binom{r+s}{a_1,\dots,a_q,r+s-\sum_{j=1}^{q}a_j}z^i\right) \tag{8.7}$$

(the contributions to the z^i term come from taking a_j different z^{i_j} terms for $j=1$ to q).

Now, if $\sum_{j=1}^{q}a_ji_j=i$, then there exists j such that $v(a_j)\le v(i)$. By Lemma 8.10, the coefficient of z^i has valuation at least $v(r+s)-v(i)$, which is $v(a)-v(i)$, by Lemma 8.14. □

Recall that $c_i := \frac{g^{(i)}(0)}{i!}e^i$.

Corollary 8.17. *For $i>3$, we have $v(c_i)>\nu+\frac{1}{p-1}$, unless $p=i=5$ and $v(a)=\nu-1$.*

Proof. By Lemma 8.16, $v(c_i)\ge iv(e)+v(a)-v(i)$. By Lemma 8.11 (ii), $v(e)=\frac{1}{3}(\nu+\frac{1}{p-1}-v(a))$. So

$$v(c_i)\ge\frac{i}{3}\left(\nu+\frac{1}{p-1}-v(a)\right)+v(a)-v(i) \tag{8.8}$$

$$=\nu+\frac{1}{p-1}+\frac{i-3}{3}\left(\nu+\frac{1}{p-1}-v(a)\right)-v(i). \tag{8.9}$$

Therefore, $v(c_i)>\nu+\frac{1}{p-1}$ whenever $\frac{(i-3)}{3}(\nu+\frac{1}{p-1}-v(a))>v(i)$. By Lemma 8.14, $\nu-v(a)\ge1$. One checks that $\frac{(i-3)}{3}(\nu+\frac{1}{p-1}-v(a))>v(i)$ always holds, except when $p=i=5$ and $v(a)=\nu-1$. This proves the corollary. □

Remark 8.18. In the case $p=5$ and $v(a)=\nu-1$, one uses (8.7) to see that $\frac{g^{(5)}(0)}{5!}$ is congruent to $32\binom{r+s}{5}\pmod{a}$ (this is the term where $I=\{1\}$ and $A=(5)$; all other terms are 0 modulo a). Thus c_5 is equal to $32\binom{r+s}{5}e^5$ plus terms of valuation greater than $v(ae^5)=\nu+\frac{13}{12}>\nu+\frac{1}{4}=\nu+\frac{1}{p-1}$. Also, $v(c_5)=\nu+\frac{1}{12}$.

Proposition 8.19. *Suppose $v(a)>0$.*

(1) *If $p>5$ or $v(a)<\nu-1$, then the disk Δ corresponding to $\overline{X}_b$ contains the K_0-rational point $x=a_0$, where $a_0=1-\frac{s^2}{r^2}$.*

(2) *If* $p = 5$ *and* $v(a) = \nu - 1$, *then* Δ *contains the point* $x = a_0$, *where* $a_0 = 1 - \left(\frac{s - \sqrt[5]{5^{4\nu+1}\binom{r+s}{5}}}{r}\right)^2$ *(for all choices of 5th root).*

Proof. To (1): By Corollary 8.17, $v(c_i) > \nu + \frac{1}{p-1}$ for all i such that $p|i$. By Lemma 2.1 (1), the torsor given by (8.4) can split into $p^{\nu-1}$ disjoint μ_p-torsors with étale reduction only if $v(c_1) \geq \nu + \frac{1}{p-1}$. In particular, we must have

$$v(g'(0)) = v(2r + \frac{2s}{\sqrt{1-a}}) \geq \nu + \frac{1}{p-1} - v(e) = \frac{2}{3}(\nu + \frac{1}{p-1}) + \frac{1}{3}v(a).$$

By Lemma 8.9, $a_0 = 1 - \frac{s^2}{r^2} \in K_0$ satisfies $v(a - a_0) \geq \frac{2}{3}(\nu + \frac{1}{p-1}) + \frac{1}{3}v(a)$. But by Lemma 8.11, $v(\rho) = \frac{2}{3}(\nu + \frac{1}{p-1}) + \frac{1}{3}v(a)$, where $|\rho|$ is the radius of Δ. So $a_0 \in \Delta$.

To (2): Let $c_5' = 32\binom{r+s}{5}e^5$. By Remark 8.18, $v(c_5' - c_5) > \nu + \frac{1}{4}$ and $v(c_5) > n$. Also, Corollary 8.17 shows that $v(c_i) > \nu + \frac{1}{4}$ for all $i > 5$ such that $5|i$. By Lemma 2.1 (2), the torsor given by (8.4) can split into $5^{\nu-1}$ disjoint μ_5-torsors with étale reduction only if

$$v\left(c_1 - \sqrt[5]{32 \cdot 5^{4\nu+1}\binom{r+s}{5}e^5}\right) \geq \nu + \frac{1}{4}.$$

In particular, we must have

$$v\left(g'(0) - \sqrt[5]{32 \cdot 5^{4\nu+1}\binom{r+s}{5}}\right) \geq \nu + \frac{1}{4} - v(e) = \nu - \frac{1}{6}. \tag{8.10}$$

Recall that $g'(0) = 2r + \frac{2s}{\sqrt{1-a}}$. Since

$$v\left((1 - \frac{1}{\sqrt{1-a}})\sqrt[5]{32 \cdot 5^{4\nu+1}\binom{r+s}{5}}\right) = v(a) + \nu - \frac{1}{5} > \nu - \frac{1}{6}$$

(as $v(a) \geq 1$), Equation (8.10) is equivalent to

$$v\left(2r + \frac{2s - \sqrt[5]{32 \cdot 5^{4\nu+1}\binom{r+s}{5}}}{\sqrt{1-a}}\right) \geq \nu - \frac{1}{6}.$$

By Lemma 8.9, we have $v(a - a_0) \geq \nu - \frac{1}{6}$, where $a_0 = 1 - \left(\frac{s - \sqrt[5]{5^{4\nu+1}\binom{r+s}{5}}}{r}\right)^2$. But $v(\rho) = \nu - \frac{1}{6}$, so a_0 specializes to $\overline{X}_b$, and we can take $a = a_0$. □

Remark 8.20. As in Remark 8.13, one shows that if a is as in Proposition 8.19, then the torsor given by (8.4) indeed splits into $p^{\nu-1}$ disjoint μ_p-torsors with étale reduction birational to an Artin-Schreier cover branched at one point of conductor 3. Also, if we are in case (1) of Proposition 8.19 and $a = 1 - \frac{s^2}{r^2}$, then since $v(g'(0)) > 0$, we must choose $\sqrt{1-a} = -\frac{s}{r}$.

8.3.4.3. ***The case*** $\mathbf{v(a-1) > 0}$. This case will be quite parallel to the $v(a) > 0$ case. We have $v(e) = \frac{1}{3}(\nu + \frac{1}{p-1} + v(1-a))$ by Lemma 8.11. We claim that $x = 1$ is branched of index strictly less than p^ν. Indeed, if $x = 1$ is branched of index p^ν, its specialization $\overline{1}$ would lie on $\overline{X}_0$ by Lemma 6.1. But $\overline{1}$ must be a smooth point of $\overline{X}$. So $\overline{1}$ corresponds to an open disk of radius 1. Since $v(a-1) > 0$, then a would specialize to $\overline{1}$ as well, contradicting the fact that a specializes to an étale tail. Let $p^{\nu_1} < p^\nu$ be the branching index of $x = 1$. Then we know $v(s) = \nu - \nu_1$.

Since the specializations of 1 and a cannot collide on $\overline{X}$, we must have a component $\overline{W}$ of $\overline{X}$ corresponding to the disk of radius $|1-a|$ centered at 1 (or equivalently, at a).

The next two lemmas play the role of Lemma 8.14 for the case $v(a-1) > 0$.

Lemma 8.21. *We have* $v(1-a) \leq 2(\nu - 1 + \frac{1}{p-1})$.

Proof. Let $Q \leq G^{str}$ be the unique subgroup of order p^{ν_1}. Consider the cover $(f^{st})' : (Y^{str})^{st}/Q \to X^{st}$, with generic fiber f'. Consider the path $\{v_i\}_{i=0}^{j}$, $\{e_i\}_{i=0}^{j-1}$ in $\mathcal{G}$, where v_0 corresponds to $\overline{X}_0$ and v_j corresponds to $\overline{W}$. Write $(\sigma_i^{\mathrm{eff}})'$ (resp. $(\delta_i^{\mathrm{eff}})'$) for the effective invariant at e_i for $(f^{st})'$ (resp. the effective different at v_i for $(f^{st})'$). Then $(\delta_0^{\mathrm{eff}})' = \nu - \nu_1 + \frac{1}{p-1}$.

Since no branch point of f' with index divisible by p, and only one branch point with index 2, specializes outward from the point corresponding to any e_i, Lemma 5.5 shows that $(\sigma_i^{\mathrm{eff}})' - 1$ is a sum of elements of the form $\sigma - 1$, where $\sigma > 0$, $\sigma \in \frac{1}{2}\mathbb{Z}$, and $\sigma \in \mathbb{Z}$ for all but one term in the sum. Therefore, $(\sigma_i^{\mathrm{eff}})' - 1 \geq -\frac{1}{2}$, so $(\sigma_i^{\mathrm{eff}})' \geq \frac{1}{2}$.

By Lemma 6.1 and monotonicity, $x = 1$ specializes to a component which intersects a component which is inseparable for $(f^{st})'$. In particular, since $x = 1$ specializes on or outward from $\overline{W}$, it must be the case that any component of $\overline{X}$ lying inward from $\overline{W}$ is inseparable for $(f^{st})'$. Then, we can apply Lemma 5.4 to each e_i, $0 \leq i < j$, to show that

$$\nu - \nu_1 + \frac{1}{p-1} - (\delta_j^{\mathrm{eff}})' \geq \frac{1}{2}v(1-a).$$

Since $(\delta_j^{\mathrm{eff}})' \geq 0$ and $\nu_1 \geq 1$, this yields $v(1-a) \leq 2(\nu - 1 + \frac{1}{p-1})$. ☐

Lemma 8.22. *We have* $v(\sqrt{1-a}) = v(s)$. *In particular,* $v(\sqrt{1-a}) \in \mathbb{Z}$.

Proof. As in the case $v(a) > 0$, we must have that $v(g'(0)e) \geq \nu$ (see beginning of proof of Lemma 8.14), so

$$v(g'(0)) \geq \frac{2}{3}\nu - \frac{1}{3}(\frac{1}{p-1} + v(1-a)).$$

Since $v(1-a) \leq 2(\nu - 1 + \frac{1}{p-1})$ (Lemma 8.21), we see that

$$v(g'(0)) \geq \frac{2}{3} - \frac{1}{p-1} > 0.$$

Recall that $g'(0) = 2r + \frac{2s}{\sqrt{1-a}}$. Since $v(2r) = 0$ (f is totally ramified above $x = \infty$), it follows that $v(\frac{2s}{\sqrt{1-a}}) = 0$. Therefore $v(\sqrt{1-a}) = v(2s) = v(s)$. □

The following lemma plays the role of Lemma 8.16 when $v(a-1) > 0$.

Lemma 8.23. *The valuation* $v\left(\frac{g^{(i)}(0)}{i!}\right) \geq (1-i)v(\sqrt{1-a}) - v(i)$.

Proof. Again, $\frac{g^{(i)}(0)}{i!}$ is the coefficient of z^i in the Maclaurin series expansion of $g(z)$. Since $(1-i)v(\sqrt{1-a}) - v(i) \leq 0$, it suffices to look at coefficients modulo R (as an R-submodule of K). Then, $g(z)$ is congruent (mod R) to

$$\left(\frac{z+\sqrt{1-a}}{z-\sqrt{1-a}}\right)^s = \left(1 + \frac{2}{\frac{z}{\sqrt{1-a}} - 1}\right)^s = (-1 - 2w - 2w^2 - 2w^3 - \cdots)^s, \tag{8.11}$$

where $w = \frac{z}{\sqrt{1-a}}$. This is analogous to (8.6), and we conclude as we do in Lemma 8.16 that the coefficient of w^i has valuation at least $v(s) - v(i)$. Thus the coefficient of z^i has valuation at least $v(s) - v(i) - iv(\sqrt{1-a})$. By Lemma 8.22, this is $(1-i)v(\sqrt{1-a}) - v(i)$. □

Recall that $c_i := \frac{g^{(i)}(0)}{i!}e^i$. Parallel to Corollary 8.17, we have:

Corollary 8.24. *For $i > 3$, we have $v(c_i) > \nu + \frac{1}{p-1}$, unless $p = i = 5$ and $v(\sqrt{1-a}) = \nu - 1$.*

Proof. By Lemma 8.23, $v(c_i) \geq iv(e) + (1-i)v(\sqrt{1-a}) - v(i)$. By Lemma 8.11 (3), $v(e) = \frac{1}{3}(\nu + \frac{1}{p-1} + v(1-a))$. So

$$v(c_i) \geq \frac{i}{3}\left(\nu + \frac{1}{p-1} + v(1-a)\right) + (1-i)v(\sqrt{1-a}) - v(i) \tag{8.12}$$

$$= \nu + \frac{1}{p-1} + \frac{i-3}{3}\left(\nu + \frac{1}{p-1} - v(\sqrt{1-a})\right) - v(i). \tag{8.13}$$

Therefore, $v(c_i) > \nu + \frac{1}{p-1}$ whenever $\frac{(i-3)}{3}(\nu + \frac{1}{p-1} - v(\sqrt{1-a})) > v(i)$. By Lemma 8.21, $\nu - v(\sqrt{1-a}) \geq 1$. One checks that $\frac{(i-3)}{3}(\nu + \frac{1}{p-1} - v(\sqrt{1-a})) \geq v(i)$ always holds, except when $p = i = 5$ and $v(\sqrt{1-a}) = \nu - 1$. This proves the corollary. □

Parallel to Remark 8.18, we have:

Remark 8.25. In the case $p = 5$ and $v(s) = v(\sqrt{1-a}) = \nu - 1$, one sees (using (8.7) and (8.11)) that $\frac{g^{(5)}(0)}{5!}$ is congruent to $\left(\sqrt{1-a}\right)^{-5} 32\binom{s}{5}$ modulo $\left(\sqrt{1-a}\right)^{-5} sR$ (viewed as an R-submodule of K). That is, $\frac{g^{(5)}(0)}{5!}$ is equal to $\left(\sqrt{1-a}\right)^{-5} 32\binom{s}{5}$ plus terms of valuation more than $-4v(\sqrt{1-a})$, which is $4(1-\nu)$. So c_5 is equal to $\left(\sqrt{1-a}\right)^{-5} 32\binom{s}{5}e^5$ plus terms of valuation greater than $4 - 4\nu + 5v(e) = \nu + \frac{13}{12} > \nu + \frac{1}{4} = \nu + \frac{1}{p-1}$. Also, $v(c_5) = \nu + \frac{1}{12}$.

The following proposition plays the role of Proposition 8.19 when $v(a-1) > 0$.

Proposition 8.26. *Suppose $v(a-1) > 0$.*

(1) *If $p > 5$ or $v(\sqrt{1-a}) < \nu - 1$, then the disk Δ corresponding to $\overline{X}_b$ contains the K_0-rational point $x = a_0$, where $a_0 = 1 - \frac{s^2}{r^2}$.*
(2) *If $p = 5$ and $v(\sqrt{1-a}) = \nu - 1$, then Δ contains the point $x = a_0$, where $a_0 = 1 - \left(\frac{s - \sqrt[5]{5^{4\nu+1}\binom{s}{5}}}{r}\right)^2$.*

Proof. To (1): By Corollary 8.24, $v(c_i) > \nu + \frac{1}{p-1}$ for all i such that $p|i$. By Lemma 2.1 (1), the torsor given by (8.4) can split into $p^{\nu-1}$ disjoint μ_p-torsors with étale reduction only if $v(c_1) \geq \nu + \frac{1}{p-1}$. In particular, we must have

$$v(g'(0)) = v(2r + \frac{2s}{\sqrt{1-a}}) \geq \nu + \frac{1}{p-1} - v(e) = \frac{2}{3}(\nu + \frac{1}{p-1} - v(\sqrt{1-a})).$$

By Lemma 8.9, $a_0 = 1 - \frac{s^2}{r^2} \in K_0$ satisfies $v(a - a_0) \geq \frac{2}{3}(\nu + \frac{1}{p-1} + v(1-a))$. But by Lemma 8.11, $v(\rho) = \frac{2}{3}(\nu + \frac{1}{p-1} + v(1-a))$, where $|\rho|$ is the radius of Δ. So $a_0 \in \Delta$.

To (2): Let $c_5' = \left(\sqrt{1-a}\right)^{-5} 32\binom{s}{5}e^5$. By Remark 8.25, $v(c_5' - c_5) > \nu + \frac{1}{4}$ and $v(c_5) > n$. Also, Corollary 8.24 shows that $v(c_i) > \nu + \frac{1}{4}$ for all $i > 5$ such that $5|i$. By Lemma 2.1 (2), the torsor given by (8.4) can split into $5^{\nu-1}$ disjoint μ_5-torsors with étale reduction only if

$$v\left(c_1 - \sqrt[5]{\left(\sqrt{1-a}\right)^{-5} 32 \cdot 5^{4\nu+1}\binom{s}{5}e^5}\right) \geq \nu + \frac{1}{4}.$$

In particular, we must have

$$v\left(g'(0) - \sqrt[5]{\left(\sqrt{1-a}\right)^{-5} 32 \cdot 5^{4\nu+1}\binom{s}{5}}\right) \geq \nu + \frac{1}{4} - v(e) = \frac{5}{6},$$

which can be rewritten as

$$v\left(2r+\frac{2s-\sqrt[5]{32\cdot 5^{4\nu+1}\binom{s}{5}}}{\sqrt{1-a}}\right)\geq\frac{5}{6}.$$

By Lemma 8.9, we have $v(a-a_0)\geq 2\nu-\frac{7}{6}$, where $a_0=1-\left(\frac{s-\sqrt[5]{5^{4\nu+1}\binom{s}{5}}}{r}\right)^2$. But $v(\rho)=2\nu-\frac{7}{6}$, so $a_0\in\Delta$. □

Remark 8.27. As in Remarks 8.13 and 8.20, one shows that if a is as in Proposition 8.26, then the torsor given by (8.4) indeed splits into $p^{\nu-1}$ disjoint μ_p-torsors with étale reduction birational to an Artin-Schreier cover branched at one point of conductor 3. Also, if we are in case (1) of Proposition 8.26 and $a=1-\frac{s^2}{r^2}$, then since $v(g'(0))>0$, we must choose $\sqrt{1-a}=-\frac{s}{r}$.

8.3.5. ***The inseparable tails.*** For the rest of the proof, we replace a with the a_0 calculated in Proposition 8.12, 8.19, or 8.26 of §8.3.4, depending on the congruency class of a (see the paragraph preceding Lemma 8.5). Maintain the notations of §8.3.4. Throughout §8.3.5, we assume that $\nu>1$ (if $\nu=1$, there can be no inseparable tails by Lemma 4.5). Our goal is to calculate exactly where the inseparable tails of $\overline{X}$ lie. We first place restrictions on what kinds of inseparable tails we can have and where they lie (Lemma 8.28, Proposition 8.30), and then explicitly exhibit an inseparable tail in each allowable situation (Proposition 8.31).

The next lemma shows us that we are not looking for too many tails.

Lemma 8.28. *Suppose $\overline{X}$ has a new inseparable p^j-tail $\overline{X}_c$.*

(1) *The only new inseparable p^j-tail is $\overline{X}_c$. Its effective ramification invariant is $\sigma_c=2$.*
(2) *There are two p^j-components $\overline{X}_\beta$ and $\overline{X}_{\beta'}$ of $\overline{X}$ that intersect p^{j+1}-components and have non-integral effective ramification invariant (see Definition 4.6). We have $\sigma_\beta=\sigma_{\beta'}=\frac{1}{2}$. Also, up to switching indices β and β', we have $\overline{X}_0\prec\overline{X}_\beta\prec\overline{X}_b$ and $\overline{X}_0\prec\overline{X}_{\beta'}\prec\overline{X}_{b'}$.*
(3) *If $v(a)>0$, then $j\leq\nu-v(a)$. If $v(a-1)>0$, then $j\leq\nu-v(\sqrt{1-a})-1$.*

Proof. To (1) and (2): Let $\{\overline{X}_\alpha,\ \alpha\in A\}$ be the set of all p^j-components that intersect $p^{j'}$-components with $j'>j$. By Lemma 4.5, this includes all the p^j-tails. Also, for $i>0$, let Π_i be the set of all branch points of f (equivalently f^{str}) with branching index divisible by p^i. As a consequence of [15, Proposition 3.17] (setting the "α" of [15] equal to our j) and monotonicity, we have the following equation relating the effective ramification invariants of the $\overline{X}_\alpha$ (see Definition 4.6):

$$|\Pi_{j+1}| - 2 = \sum_{\alpha \in A} (\sigma_\alpha - 1). \tag{8.14}$$

By definition, $\sigma_\alpha > 0$ for any $\overline{X}_\alpha$. Furthermore, by [15, Lemmas 4.1, 4.2 (i)], any new inseparable *tail* $\overline{X}_\alpha$ has invariant $\sigma_\alpha \geq 2$. By Lemmas 6.2 and 4.8, any $\overline{X}_\alpha$ that has non-integral effective ramification invariant and is not a tail must lie between $\overline{X}_0$ and either $\overline{X}_b$ or $\overline{X}_{b'}$, and must have $\sigma_\alpha \geq \frac{1}{2}$. By monotonicity, there are at most two such components: Call them $\overline{X}_\beta$ and possibly $\overline{X}_{\beta'}$. If there is a new inseparable p^j-tail $\overline{X}_c$, then the only way that (8.14) can be satisfied is if $|\Pi_{j+1}| = 2$, if $\overline{X}_c$ is the only new p^j-tail (and $\sigma_c = 2$), and if $\overline{X}_\beta$ and $\overline{X}_{\beta'}$ both exist (with $\sigma_\beta = \sigma_{\beta'} = \frac{1}{2}$). Since the conductor of a $\mathbb{Z}/p^a$-extension of $k[[t]]$ is always at least p^{a-1} (see, e.g., [17, Lemma 19]), we see that $\overline{X}_\beta$ (resp. $\overline{X}_{\beta'}$) intersects a p^{j+1}-component, because otherwise the invariant σ_β (resp. $\sigma_{\beta'}$) would be at least $\frac{p}{2}$.

To (3): If $v(a) > 0$, recall that $\overline{W}$ is the component of $\overline{X}$ separating 0 and a. Assume that β and β' are as in the statement of (2). Then $\overline{W} \prec \overline{X}_\beta$ (and $\overline{W} \prec \overline{X}_{\beta'}$). Now, by Remark 8.15, the order of generic inertia above $\overline{W}$ is at most $p^{\nu - v(a)+1}$. By monotonicity, $\overline{X}_\beta$ has less inertia than $\overline{W}$. Since $\overline{X}_\beta$ is a p^j-component, we see that $j \leq \nu - v(a)$.

If $v(a-1) > 0$, then Lemma 8.22 shows that f^{str} is branched above 1 of index $p^{\nu - v(\sqrt{1-a})}$. From the proof of (2), $|\Pi_{j+1}| = 2$, which means that 1 is branched of index at least p^{j+1}. It follows that $j \leq \nu - v(\sqrt{1-a}) - 1$. $\square$

Not only does Lemma 8.28 show us that we are not looking for too many tails, but it (in particular, part (2)) also gives information on the inseparable *interior* components $\overline{X}_\beta$ and $\overline{X}_{\beta'}$ of $\overline{X}$. The next lemma gives further information on the corresponding disks.

Lemma 8.29 (cf. Lemma 8.11)**.** *Suppose $\overline{X}$ has a new inseparable p^j-tail $\overline{X}_c$, and maintain the notation $\overline{X}_\beta$ from Lemma 8.28. Let $\overline{Z}_\beta$ be the unique component of $\overline{Z}^{str}$ above $\overline{X}_\beta$. Let e' (resp. ρ') $\in \overline{K}$ be such that the radius of the disk corresponding to $\overline{Z}_\beta$ (resp. $\overline{X}_\beta$) is $|e'|$ (resp. $|\rho'|$).*

(1) *If $v(a) = v(a-1) = 0$, then $v(\rho') > \frac{2}{3}(\nu - j + \frac{1}{p-1})$ and $v(e') > \frac{1}{3}(\nu - j + \frac{1}{p-1})$.*
(2) *If $v(a) > 0$, then $v(\rho') > \frac{2}{3}(\nu - j + \frac{1}{p-1}) + \frac{1}{3}v(a)$ and $v(e') > \frac{1}{3}(\nu - j + \frac{1}{p-1} - v(a))$.*
(3) *If $v(a-1) > 0$, then $v(\rho') > \frac{2}{3}(\nu - j + \frac{1}{p-1} - v(1-a))$ and $v(e') > \frac{1}{3}(\nu - j + \frac{1}{p-1} + v(1-a))$.*

Proof. We give only a sketch. As in Lemma 8.11, we have $v(e') = \frac{1}{2}(v(\rho') - v(a))$, so it suffices to prove the statements about ρ'.

Let $Q < G^{str}$ be the unique subgroup of order p^j, and let

$$(f^{st})' : (Y^{str})^{st}/Q \to X^{st}$$

be the canonical map.

Consider the path $\{v_i\}_{i=0}^{\ell}$, $\{e_i\}_{i=0}^{\ell-1}$, where v_0 corresponds to $\overline{X}_0$ and v_ℓ corresponds to $\overline{X}_\beta$. Write $(\sigma_i^{\text{eff}})'$ for the effective invariant for $(f^{st})'$ at e_i. The effective different for $(f^{st})'$ at v_0 is $\nu - j + \frac{1}{p-1}$, whereas at v_ℓ it is 0. Lastly, write $\Delta_i = 1$ (resp. 0) if $\overline{X}_c$ lies (resp. does not lie) outward from the point corresponding to e_i. In particular, $\Delta_{\ell-1} = 0$, so Δ_i is not always 1.

To (1) (cf. Proof of Lemma 8.11 (1)): By Lemmas 5.5 and 8.28, $(\sigma_i^{\text{eff}})' = \frac{1}{2} + \Delta_i$. In particular, $\Delta_{\ell-1} = 0$, so $(\sigma_i^{\text{eff}})' \leq \frac{3}{2}$ with equality not holding for some i. By applying Lemma 5.4 (3) to each e_i, $0 \leq i < \ell$, we obtain $v(\rho') > \frac{2}{3}(\nu - j + \frac{1}{p-1})$.

To (2): Recall that $\overline{W}$ is the component of $\overline{X}$ corresponding to the closed disk of radius $|a|$ and center a. Suppose $\overline{W}$ corresponds to v_{i_0}. Then, for $i < i_0$, Lemma 5.5 shows that $(\sigma_i^{\text{eff}})' = \Delta_i$. For $i > i_0$, Lemma 5.5 shows that $(\sigma_i^{\text{eff}})' = \frac{1}{2} + \Delta_i$. If Δ_i were equal to 1 for all i, we would have $v(\rho') = \frac{2}{3}(\nu - j + \frac{1}{p-1}) + \frac{1}{3}v(a)$ (cf. Proof of Lemma 8.11 (2)). But since, for large enough i, we have $\Delta_i = 0$, we can see from Lemma 5.4 (3) that $v(\rho') > \frac{2}{3}(\nu - j + \frac{1}{p-1}) + \frac{1}{3}v(a)$.

To (3): In this case, $\overline{W}$ is the component of $\overline{X}$ corresponding to the closed disk of radius $|1-a|$ and center a. Suppose $\overline{W}$ corresponds to v_{i_0}. Then, for $i < i_0$, Lemma 5.5 shows that $(\sigma_i^{\text{eff}})' = -\frac{1}{2} + \Delta_i$. For $i > i_0$, Lemma 5.5 shows that $(\sigma_i^{\text{eff}})' = \frac{1}{2} + \Delta_i$. If Δ_i were equal to 1 for all i, we would have $v(\rho') = \frac{2}{3}(\nu - j + \frac{1}{p-1} + v(1-a))$ (cf. Proof of Lemma 8.11 (3)). But since, for large enough i, we have $\Delta_i = 0$, we can see from Lemma 5.4 (3) that $v(\rho') > \frac{2}{3}(\nu - j + \frac{1}{p-1} + v(1-a))$. ☐

Knowledge of the inertia groups above $\overline{X}_\beta$, $\overline{X}_{\beta'}$, as well as the disks to which they correspond, gives further restrictions on the inseparable tails, as we see below.

Proposition 8.30. *Suppose that $\overline{X}$ has a new inseparable p^j-tail. Then $v(a) > 0$ or $v(a-1) > 0$. Furthermore, if $v(a) > 0$, then either $j = \nu - v(a)$, or both $p = 5$ and $j = \nu - v(a) - 1$. If $v(a-1) > 0$, then $p = 5$ and $j = \nu - v(\sqrt{1-a}) - 1$.*

Proof. Maintain the notation $\overline{X}_\beta$ and $\overline{Z}_\beta$ from Lemma 8.29. Let e' be such that the radius of the disk corresponding to $\overline{Z}_\beta$ is $|e'|$, and let t' be a coordinate on this disk. Let $Q < G^{str}$ be the unique subgroup of order p^j. Write $q : (Y^{str})' := Y/Q \to Z^{str}$ for the canonical map. We can write the equation of q in terms of t' as

$$y^{p^{\nu-j}} = g(e't') = 1 + \frac{g'(0)}{1!}(e't') + \frac{g''(0)}{2!}(e't')^2 + \cdots .$$

We claim that, unless j and a are as in the statement of the proposition, the right-hand side is a $p^{\nu-j}$th power in $\mathrm{Frac}(R\{t'\})$. This means that $(Y^{str})^{st}/Q$ splits into $p^{\nu-j}$ irreducible components above $\overline{Z}_\beta$, each mapping isomorphically to $\overline{Z}_\beta$. This implies that $\overline{X}_\beta$ is a p^j-component for f that does *not* intersect a p^{j+1}-component, which contradicts the definition of $\overline{X}_\beta$.

Let c_i' be the coefficient of $(t')^i$ in $g(e't')$. Then $c_i' = c_i(\frac{e'}{e})^i$, so $v(c_i') = v(c_i) - i(v(e) - v(e'))$, which is greater than $v(c_i) - \frac{i}{3}j$, by Lemmas 8.11 and 8.29.

In the case $v(a) = v(a-1) = 0$, it is clear from Lemma 8.29 that $v(c_i') > \nu - j + \frac{1}{p-1}$ for $i \geq 3$. Also, by Lemma 2.1 (1), we have $v(c_i) \geq \nu + \frac{1}{p-1}$ for $i = 1, 2$. Then

$$v(c_i') > \nu + \frac{1}{p-1} - \frac{i}{3}j > \nu - j + \frac{1}{p-1}$$

for $i = 1, 2$. Since $v(c_i') > \nu - j + \frac{1}{p-1}$ for all i, the binomial theorem shows that $g(e't')$ is a $p^{\nu-j}$th power in $\mathrm{Frac}(R\{t'\})$, finishing this case.

Now, suppose $v(a) > 0$. Since $v(c_i') > v(c_i) - \frac{i}{3}j$, Equation (8.9) shows that

$$v(c_i') > \nu - j + \frac{1}{p-1} + \frac{i-3}{3}(\nu - j + \frac{1}{p-1} - v(a)) - v(i).$$

By Lemma 8.28 (3), $\nu - j - v(a) \geq 0$. We obtain that $v(c_i') > \nu - j + \frac{1}{p-1}$ for $i \geq 3$, unless $j = \nu - v(a)$ or both $j = \nu - v(a) - 1$ and $p = i = 5$. Barring these possibilities, we conclude as in the case $v(a) = v(a-1) = 0$ that $v(c_i') > \nu - j + \frac{1}{p-1}$ for $i = 1, 2$. Since $v(c_i') > \nu - j + \frac{1}{p-1}$ for all i, the binomial theorem shows that $g(e't')$ is a $p^{\nu-j}$th power in $\mathrm{Frac}(R\{t'\})$, finishing this case.

Lastly, if $v(a-1) > 0$, then since $v(c_i') > v(c_i) - \frac{i}{3}j$, Equation (8.13) shows that

$$v(c_i') > \nu - j + \frac{1}{p-1} + \frac{i-3}{3}(\nu - j + \frac{1}{p-1} - v(\sqrt{1-a})) - v(i).$$

By Lemma 8.28 (3), $\nu - j - v(\sqrt{1-a}) \geq 1$. We obtain that $v(c_i') > \nu - j + \frac{1}{p-1}$ for $i \geq 3$, unless $p = i = 5$ and $j = \nu - v(\sqrt{1-a}) - 1$. We conclude as in the case $v(a) > 0$.

□

Proposition 8.30 has narrowed down the possibilities for inseparable tails to the point that we can now explicitly exhibit every possible inseparable tail.

Proposition 8.31. *A new inseparable tail in fact exists in all cases allowed by Proposition 8.30. In particular:*

(1) *If $v(a) > 0$, then $\overline{X}$ has a new inseparable $p^{\nu - v(a)}$-tail corresponding to the disk of radius $p^{-(v(a)+\frac{1}{(p-1)})}$ around $x = \frac{a}{2}$. The two components of $\overline{Z}^{str}$ lying above correspond to the disks of radius $p^{-(\frac{1}{2(p-1)})}$ around $z = \pm\sqrt{-1}$.*

(2) *If $v(a) > 0$, $p = 5$, and $v(a) < \nu - 1$, then $\overline{X}$ has a new inseparable $p^{\nu - v(a)-1}$-tail corresponding to the disk of radius $p^{-(v(a)+\frac{17}{20})}$ around $x = \frac{a}{1-d^2}$, where $d := \pm\left(\frac{5^{v(a)+1}}{r+s}\right)^{2/5}$ (for any choice of 5th root). The two components of $\overline{Z}^{str}$ lying above correspond to the disks of radius $p^{-\frac{17}{40}}$ around the two possible choices of d.*

(3) *If $v(a-1) > 0$, $p = 5$, and $v(\sqrt{1-a}) < \nu - 1$, then $\overline{X}$ has a new inseparable $p^{\nu - v(\sqrt{1-a})-1}$-tail corresponding to the disk of radius $p^{-(v(1-a)+\frac{17}{20})}$ around $x = \frac{a}{1-d^2}$, where $d := \pm 2\frac{s}{r}\left(\frac{5^{v(\sqrt{1-a})+1}}{s}\right)^{2/5}$ (for any choice of 5th root). The two components of $\overline{Z}^{str}$ lying above correspond to the disks of radius $p^{-(v(\sqrt{1-a})+\frac{17}{40})}$ around the two possible choices of d.*

Combining this with Lemma 8.28 (1), we obtain:

Corollary 8.32. *The new inseparable tails mentioned in Proposition 8.31 are all the new inseparable tails of $\overline{X}$.*

Before proving Proposition 8.31 we prove a lemma. Recall that

$$g(z) = \left(\frac{z+1}{z-1}\right)^r \left(\frac{z+\sqrt{1-a}}{z-\sqrt{1-a}}\right)^s.$$

Lemma 8.33. *Suppose $p = 5$.*

(1) *Suppose $v(a) > 0$, $v(a) < \nu - 1$, and $a = 1 - \frac{s^2}{r^2}$. Let $d, e'' \in R$ with $v(d) = \frac{2}{5}$ and $v(e'') = \frac{17}{40}$. Then, if we expand $g(d + e''t'')$ as a power series in $R\{t''\}$, we will have $g(d + e''t'') \equiv$*

$$g(d) + (-1)^{r+s}\left(2(r+s)d^2e''t'' + 2(r+s)d(e''t'')^2 + \frac{32(r+s)}{5}(e''t'')^5\right)$$

modulo $5^{v(a)+\frac{5}{4}+\epsilon}R\{t''\}$ for some $\epsilon > 0$. Furthermore,

$$v(g(d) - (-1)^{r+s}) = v(a) + 1.$$

(2) *Suppose $v(a-1)>0$, $v(\sqrt{1-a})<\nu-1$, and $a=1-\frac{s^2}{r^2}$. Let $d,e''\in R$ with $v(d)=v(\sqrt{1-a})+\frac{2}{5}$ and $v(e'')=v(\sqrt{1-a})+\frac{17}{40}$. Then, if we expand $g(d+e''t'')$ as a power series in $R\{t''\}$, we will have $g(d+e''t'')\equiv$*

$$g(d)+(-1)^{r+s}\left(-8\frac{r^3}{s^2}d^2e''t''-8\frac{r^3}{s^2}d(e''t'')^2-\frac{32r^5}{5s^4}(e''t'')^5\right)$$

modulo $5^{v(\sqrt{1-a})+\frac{5}{4}+\epsilon}R\{t''\}$ for some $\epsilon>0$. Furthermore,

$$v(g(d)-(-1)^{r+s})=v(\sqrt{1-a})+1.$$

Proof. Write $\sqrt{1-a}=-\frac{s}{r}$ (Remarks 8.20, 8.27).

To (1): We can write

$$g(z)=\left(\frac{z+1}{z-1}\right)^{r+s}\left(\left(\frac{z+\sqrt{1-a}}{z+1}\right)\left(\frac{z-1}{z-\sqrt{1-a}}\right)\right)^{s}. \tag{8.15}$$

Recall (Lemma 8.14) that $v(r+s)=v(a)$. We will calculate $g(d+e''t'')$ by first expanding $g(z)$ as a power series, and then plugging in $z=d+e''t''$. Note that, since $\min(v(d),v(e''))=\frac{2}{5}$, we can essentially think of z as having valuation $\frac{2}{5}$, and thus can ignore terms of the form cz^i such that $v(c)+\frac{2}{5}i>v(a)+\frac{5}{4}$. Throughout, we use the notation $\sigma\approx\tau$ if, thinking of z as a generic element of R having valuation $\frac{2}{5}$, we have $\sigma=\tau$ modulo terms with valuation $>v(a)+\frac{5}{4}$. Since $v(a)\geq 1$, note that $a^2z\approx 0$. This, as well as the fact that $v(a)=v(r+s)$ (Lemma 8.14), will be used repeatedly (without mention) below.

Using (8.6), (8.7), and simplifying, we can write

$$\left(\frac{z+1}{z-1}\right)^{r+s}\approx(-1)^{r+s}\left(1+2(r+s)z+\frac{8}{3}(r+s)z^3+\frac{32}{5}(r+s)z^5\right). \tag{8.16}$$

On the other hand, since $\sqrt{1-a}=-\frac{s}{r}$, if we let $\mu=\sqrt{1-a}-1=-\frac{r+s}{r}$, then $v(\mu)=v(a)$. We can write

$$\begin{aligned}\left(\left(\frac{z+\sqrt{1-a}}{z+1}\right)\left(\frac{z-1}{z-\sqrt{1-a}}\right)\right)^{s} &= \left((1+\frac{\mu}{z+1})(1+\frac{\mu}{z-1-\mu})\right)^{s}\\ &\approx 1+2s\mu\frac{z}{z^2-1},\end{aligned}$$

and thus

$$\begin{aligned}&\left(\left(\frac{z+\sqrt{1-a}}{z+1}\right)\left(\frac{z-1}{z-\sqrt{1-a}}\right)\right)^{s}\approx\\ &\qquad 1+2\frac{s(r+s)}{r}(z+z^3)\approx 1-2(r+s)(z+z^3).\end{aligned} \tag{8.17}$$

Combining (8.15), (8.16), and (8.17) yields

$$g(z) \approx (-1)^{r+s}\left(1 + \frac{2}{3}(r+s)z^3 + \frac{32}{5}(r+s)z^5\right).$$

Now, if we plug in $z = d + e''t''$ (and ignore all t'' terms where the coefficient has valuation $> v(a) + \frac{5}{4}$), then z^3 becomes $d^3 + 3d^2e''t'' + 3d(e''t'')^2$ and z^5 becomes $d^5 + (e''t'')^5$. We obtain the expression in the lemma.

Since $v(\frac{32}{5}(r+s)d^5) = v(a) + 1$ and $v(\frac{2}{3}(r+s)d^3) = v(a) + \frac{6}{5}$, we see that $v(g(d) - (-1)^{r+s}) = v(a) + 1$.

To (2): Recall by Lemma 8.22 that $v(s) = v(\sqrt{1-a})$. Let $w = \frac{z}{\sqrt{1-a}} = -\frac{r}{s}z$. Again, we calculate $g(d + e''t'')$ by first expanding $g(z)$, and then plugging in $z = d + e''t''$. Since $\min(v(d), v(e'')) = v(\sqrt{1-a}) + \frac{2}{5}$, we can essentially think of z as having valuation $v(\sqrt{1-a}) + \frac{2}{5}$ and w as having valuation $\frac{2}{5}$. We will write $\sigma \approx \tau$ if, thinking of z (resp. w) as a generic element of R having valuation $v(\sqrt{1-a}) + \frac{2}{5}$ (resp. $\frac{2}{5}$), we have $\sigma = \tau$ modulo terms with valuation $> v(\sqrt{1-a}) + \frac{5}{4}$. Note also that $v(s) = v(\sqrt{1-a})$ (Lemma 8.22).

Using (8.7), we have

$$\left(\frac{z+1}{z-1}\right)^r \approx (-1)^r(1 + 2rz) \tag{8.18}$$

(note that $z^2 \approx 0$).

On the other hand, using (8.11), (8.7), and simplifying, we can write

$$\left(\frac{z+\sqrt{1-a}}{z-\sqrt{1-a}}\right)^s \approx (-1)^s\left(1 + 2sw + \frac{8}{3}sw^3 + \frac{32}{5}sw^5\right). \tag{8.19}$$

Combining (8.18), (8.19), and substituting $w = -\frac{r}{s}z$ yields

$$g(z) \approx (-1)^{r+s}\left(1 - \frac{8r^3}{3s^2}z^3 - \frac{32r^5}{5s^4}z^5\right). \tag{8.20}$$

If we plug in $z = d + e''t''$ (and ignore all t'' terms where the coefficient has valuation $> v(\sqrt{1-a}) + \frac{5}{4}$), then z^3 becomes $d^3 + 3d^2e''t'' + 3d(e''t'')^2$ and z^5 becomes $d^5 + (e''t'')^5$. We obtain the expression in the lemma.

Since $v(\frac{32r^5}{5s^4}d^5) = v(\sqrt{1-a}) + 1$ and $v(\frac{8r^3}{3s^2}d^3) = v(\sqrt{1-a}) + \frac{6}{5}$, we see that $v(g(d) - (-1)^{r+s}) = v(\sqrt{1-a}) + 1$. □

We now prove Proposition 8.31:

Proof. Since $z^2 = \frac{x-a}{x}$, it is easy to check (as in Lemma 8.11) that it suffices to prove the statements about $\overline{Z}^{str}$.

Say that we wish to show that a disk $\mathcal{D}$ in $\mathbb{P}^1_z$ corresponds to a component of $\overline{Z}^{str}$ lying above a new inseparable p^j-tail $\overline{X}_c$ of $\overline{X}$. Let $Q < G^{str}$ be the unique subgroup of order p^j. If $\hat{Y}$, $\hat{Z}$ are the formal completions of

$(Y^{str})^{st}/Q$ and $(Z^{str})^{st}$ along their special fibers, then we claim that it suffices to show that the generic fiber of the torsor $\hat{f}: \hat{Y} \times_{\hat{Z}} \mathcal{D} \to \mathcal{D}$ splits into $p^{\nu-j-1}$ μ_p-torsors, each of which has étale reduction with conductor 2.

We prove the claim. Suppose $\mathcal{D}$ is such that $\hat{f}$ splits as desired. Since an Artin-Schreier cover of conductor 2 has genus $\frac{p-1}{2} > 0$, then [12, Lemma 4.3] shows that $\overline{X}_c$ is contained in the stable reduction of $\overline{X}$. Since the proposed disks corresponding to $\overline{X}_c$ do not contain $x = 0$, 1, a or ∞, it follows that no branch point specializes to or outward from $\overline{X}_c$. So either $\overline{X}_c$ is a new inseparable tail, or there exists a new inseparable tail lying outward from $\overline{X}_c$. In cases (2) and (3), no new inseparable tail can lie outward from $\overline{X}_c$, by Proposition 8.30, Lemma 4.5, and monotonicity. In case (1), any new inseparable tail lying outward from $\overline{X}_c$ would have to be one of the tails in (2) or (3) (again using Proposition 8.30, Lemma 4.5, and monotonicity), and inspection shows that this is not the case. Thus $\overline{X}_c$ is an inseparable tail, proving the claim.

It remains to show that, for the disks $\mathcal{D}$ in the proposition, $\hat{f}$ splits as desired. Let $z = d$ be a center of $\mathcal{D}$, and enlarge K (if necessary) so that K contains an element e'' such that $|e''|$ is the radius of $\mathcal{D}$. Then we can choose a coordinate t'' on $\mathcal{D}$ such that $z = d + e''t''$. Enlarge K again (if necessary) so that $g(d) \in (K^\times)^{p^\nu}$. The generic fiber of $\hat{f}$ can be given by the equation

$$y^{p^{\nu-j}} = \frac{g(d + e''t'')}{g(d)}.$$

To (1): Here $d = \sqrt{-1}$ (either square root) and $e'' \in K$ with $v(e'') = \frac{1}{2(p-1)}$. Also, $j = \nu - v(a)$, so $\nu - j = v(a) = v(r+s)$, by Lemma 8.14. So we may multiply $\frac{g(z)}{g(d)}$ by $p^{v(r+s)}$th powers without changing the generic fiber of $\hat{f}$. By (8.15), we may assume that the generic fiber of $\hat{f}$ is given by

$$y^{p^{v(a)}} = \left(\left(\frac{z+\sqrt{1-a}}{z+1}\right)\left(\frac{z-1}{z-\sqrt{1-a}}\right)\right)^s.$$

Let $\mu = -\frac{r+s}{r} = \sqrt{1-a} - 1$. Then $v(\mu) = v(a)$. Thus we can write

(8.21)
$$y^{p^{v(a)}} = h(z) = \left((1+\frac{\mu}{z+1})(1+\frac{\mu}{z-1-\mu})\right)^s = 1 + 2s\mu\frac{z}{z^2-1} + O(\mu^2),$$

where $O(\mu^2)$ signifies terms in z whose coefficients have valuation at least $v(\mu^2)$.

After a possible finite extension, assume $h(d) \in (K^\times)^{p^{v(a)}}$. Thus we may replace $h(z)$ by $\frac{h(z)}{h(d)}$, without changing $\hat{f}$. Expanding $\frac{h(z)}{h(d)}$ out in terms of

t'' gives
$$1+\frac{h'(d)}{h(d)1!}e''t''+\frac{h''(d)}{h(d)2!}(e''t'')^2+\cdots.$$
For all $i>0$, Equation (8.21) shows that
$$v(\frac{h^{(i)}(d)}{i!h(d)})\geq v(\mu)=v(a)=\nu-j.$$
So for $i>2$,
$$v(\frac{h^{(i)}(d)}{i!h(d)}(e'')^i)>\nu-j+\frac{1}{p-1}.$$
For $i=1$, a direct calculation shows $\frac{h'(d)}{h(d)}=O(\mu^2)$, so
$$v(\frac{h'(d)}{h(d)}e'')>2v(\mu)=2v(a)>\nu-j+\frac{1}{p-1}.$$
For $i=2$, a direct calculation shows
$$h''(z)=2s\mu\frac{2z}{(z^2-1)^3}\left(-4(1+z^2)+2z(z^2-1)\right)+O(\mu^2).$$
Then $v(h''(d))=v(\mu)=\nu-j$. Since $v(h(d))=0$, we have that
$$v(\frac{h''(d)}{h(d)2!}(e'')^2)=\nu-j+\frac{1}{p-1}.$$
By [12, Lemma 3.1 (i)], the generic fiber of $\hat{f}$ splits into $p^{v(a)-1}$ μ_p-torsors, each of which has étale reduction with conductor 2.

To (2): Here $d=\pm\left(\frac{5^{v(a)+1}}{r+s}\right)^{2/5}$ (any 5th root) and $e''\in K$ with $v(e'')=\frac{17}{40}$. Since $v(r+s)=v(a)$, then $v(d)=\frac{2}{5}$. Also, $\nu-j=v(a)+1$. Then the generic fiber of $\hat{f}$ can be given by $y^{p^{v(a)+1}}=\frac{g(d+e''t'')}{g(d)}$. By Lemma 8.33 (1), this is equivalent to
$$y^{p^{v(a)+1}}\approx 1+\left(2(r+s)d^2e''t''+2(r+s)d(e''t'')^2+\frac{32(r+s)}{5}(e''t'')^5\right),$$
where "$\approx$" means we have equality up to terms with coefficients of valuation $>v(a)+\frac{5}{4}$. Note that, since $v(g(d)-(-1)^{r+s})=v(a)+1$, dividing out by $g(d)$ is the same as dividing the coefficients of positive powers of t'' by $(-1)^{r+s}$, up to $\approx$.

The valuation of the coefficient of $(t'')^2$ is $v(a)+\frac{5}{4}$. If c_1'' and c_5'' are the coefficients of t'' and $(t'')^5$, respectively, then plugging in d shows $c_5''-\frac{(c_1'')^5}{5^{4v(a)+5}}=0$. By [12, Lemma 3.1 (ii)], the special fiber of $\hat{f}$ splits into $p^{v(a)}$ μ_p-torsors, each of which has étale reduction with conductor 2.

To (3): Here $d=\pm 2\frac{s}{r}\left(\frac{5^{v(\sqrt{1-a})+1}}{s}\right)^{2/5}$ (any 5th root) and $e''\in K$ with $v(e'')=v(\sqrt{1-a})+\frac{17}{40}$. Since $v(s)=v(\sqrt{1-a})$ by Lemma 8.22, then

$v(d) = v(\sqrt{1-a}) + \frac{2}{5}$. Also, $\nu - j = v(a) + 1$. Then the generic fiber of $\hat{f}$ can be given by $y^{p^{v(a)+1}} = \frac{g(d+e''t'')}{g(d)}$. By Lemma 8.33 (2), this is equivalent to

$$y^{p^{v(a)+1}} \approx 1 + \left(-8\frac{r^3}{s^2}d^2e''t'' - 8\frac{r^3}{s^2}d(e''t'')^2 - \frac{32r^5}{5s^4}(e''t'')^5 \right),$$

where "$\approx$" is as in (2). As in (2), dividing out by $g(d)$ is the same as dividing the coefficients of positive powers of t'' by $(-1)^{r+s}$, up to $\approx$.

The valuation of the coefficient of $(t'')^2$ is $v(\sqrt{1-a}) + \frac{5}{4}$. If c_1'' and c_5'' are the coefficients of t'' and $(t'')^5$, respectively, then plugging in d shows

$$c_5'' - \frac{(c_1'')^5}{5^{4v(\sqrt{1-a})+5}} = (2^{25} - 2^5)\frac{r^5}{5s^4}(e'')^5,$$

which has valuation $v(\sqrt{1-a}) + \frac{25}{8} > v(\sqrt{1-a}) + \frac{5}{4}$. By [12, Lemma 3.1 (i)], the generic fiber of $\hat{f}$ splits into $p^{v(a)-1}$ μ_p-torsors, each of which has étale reduction with conductor 2. $\square$

8.3.6. *A field of definition of the stable model.* We first determine a field of definition of f^{str}. Recall that f^{str} is branched at 0, 1, ∞, and a, and that $K_\nu = K_0(\zeta_{p^\nu})$.

Proposition 8.34. *The cover f^{str} is defined (as a G^{str}-cover) over $K^{str} := K_\nu(a, \sqrt{1-a}) = K_\nu(a)$.*

Proof. The explicit equations (8.1) and (8.2) give the cover f^{str}. So it is immediate that f^{str} is defined as a mere cover (i.e., without the $G^{str}-$ *action*) over $K_0(a, \sqrt{1-a})$.

Let α be a generator of $\mathbb{Z}/p^\nu \leq G^{str}$, let β be an element of order 2 in G^{str}, and let ζ_{p^ν} be a p^νth root of unity. Since α^* fixes z, Equation (8.2) shows that $\alpha^*(y) = \zeta_{p^\nu}^i y$ for some $i \in \mathbb{Z}$. Also, Equation (8.1) shows that $\beta^*(z) = -z$. Then we see that $\beta^*(g(z)) = g(z)^{-1}$. Thus $\beta^*(y) = \zeta_{p^\nu}^\ell y^{-1}$ for some $\ell \in \mathbb{Z}$. This shows that the action of G^{str} is defined over $K_0(\zeta_{p^\nu}) = K_\nu$. So f^{str} is defined over $K_\nu(a, \sqrt{1-a})$ as a G^{str}-cover.

To conclude the proof, note that either $v(1-a) = 0$ or $v(1-a) = 2v(s) \in 2\mathbb{Z}$ (Lemma 8.22). Since $v(K_\nu(a)^\times) \supset \mathbb{Z}$ and $p \neq 2$, it follows that $\sqrt{1-a} \in K_\nu(a)$. $\square$

Recall that we have fixed $a = 1 - \frac{s^2}{r^2}$, unless we are in the situation of Propositions 8.19 (2) or 8.26 (2). In these cases, $a = 1 - \left(\frac{s - \sqrt[5]{5^{4\nu+1}\binom{r+s}{5}}}{r}\right)^2$ or $a = 1 - \left(\frac{s - \sqrt[5]{5^{4\nu+1}\binom{s}{5}}}{r}\right)^2$, respectively (see Propositions 8.12, 8.19, and 8.26).

Proposition 8.35. (1) *If $v(a) = v(a-1) = 0$, then $(f^{str})^{st}$ can be defined over a tame extension $(K^{str})^{st}$ of K_ν.*

(2) *Suppose $v(a) > 0$ or $v(a-1) > 0$.*
  (a) *We have $\nu > 1$.*
  (b) *If $p > 5$, $v(a) = \nu - 1$, or $v(\sqrt{1-a}) = \nu - 1$, then $(f^{str})^{st}$ can be defined over a tame extension $(K^{str})^{st}$ of $K_\nu(a)(\sqrt[p]{1+u})$, where $u \in K_0(a)$ has valuation 1.*
  (c) *If $p = 5$ and both $v(a)$ and $v(\sqrt{1-a})$ are less than $\nu - 1$, then $(f^{str})^{st}$ can be defined over a tame extension $(K^{str})^{st}$ of $K_\nu(\sqrt[p]{\eta})(\sqrt[p]{1+u}, \sqrt[p]{1+u'})$, where $\eta \in \mathbb{Q}$ has prime-to-p valuation, $u \in K_0$ has valuation 1, and $u' \in K_0(\sqrt[p]{\eta})$ has valuation 1.*

Proof. Let $K^{str} = K_\nu(a)$, as in Proposition 8.34.

We will use the criterion of [12, Proposition 4.9], which states that if f has monotonic stable reduction, if L/K^{str} is such that G_L fixes a smooth point of $\overline{X}$ on each tail of $\overline{X}$, and if G_L fixes a smooth point of $\overline{Y}^{str}$ above each tail of $\overline{X}$, then $(f^{str})^{st}$ can be defined over a tame extension of L.

To (1): It is clear that there is a K^{str}-rational point specializing to each étale tail (namely, $x = 0$ and $x = a$). By (8.1) and (8.2), the fibers of f^{str} above $x = 0$ and $x = a$ consist of K^{str}-rational points. These points are fixed by $G_{K^{str}}$. Now, if $v(a) = v(a-1) = 0$, there are no inseparable tails (Proposition 8.30) and $K^{str} = K_\nu$. This, together with the criterion of [12, Proposition 4.9], proves (1).

To (2a): If $v(a) > 0$, then Lemma 8.14 shows that $\nu > 1$. If $v(a-1) > 0$, then Lemmas 8.21 and 8.22 show that $\nu > 1$.

To (2b): Suppose we are in the situation of (2) and $v(a) > 0$. By Proposition 8.31 (1), there is a $p^{\nu - v(a)}$-inseparable tail $\overline{X}_c$ to which the K^{str}-rational point $x = \frac{a}{2}$ specializes. Then each component of $\overline{Z}^{str}$ above $\overline{X}_c$ contains the specialization of one of $z = \pm\sqrt{-1}$. Consider the cover $(Y^{str})' \to Z^{str}$, where $(Y^{str})' = Y^{str}/Q$ with Q the unique subgroup of G^{str} of order $p^{\nu - v(a)}$. Equation (8.21) shows that this cover can be given by the equation

$$y^{p^{v(a)}} = 1 + 2s\mu\frac{z}{z^2-1} + O(\mu^2),$$

where $v(\mu) = v(a)$, and where the terms on the right-hand side are all in $K_0(a)$. Plugging in $z = \pm\sqrt{-1}$, we get that $y^{p^{v(a)}} = 1 + \alpha$, with $v(\alpha) = v(a)$. By a binomial expansion (cf. proof of Lemma 2.1), $1+\alpha$ has a $p^{v(a)-1}$st root in $K_0(a)$, which is of the form $1+u$ with $v(u) = 1$. So $G_{K^{str}(\sqrt[p]{1+u})}$ fixes the fiber above the specialization of $x = \frac{a}{2} \in \overline{X}_c$ for this cover. Since the quotient by Q is radicial above $\overline{X}_c$, it follows that $G_{K^{str}(\sqrt[p]{1+u})}$ fixes the fiber of $\overline{Y}^{str}$ above the specialization of $x = \frac{a}{2} \in \overline{X}_c$ pointwise. In particular, it fixes a point above $\overline{X}_c$.

If, instead, $v(a-1) > 0$, then there is a (not new) $p^{\nu - v(s)}$-inseparable tail $\overline{X}_c$ containing the specialization of $x = 1$. Then each component of $\overline{Z}^{str}$ above $\overline{X}_c$ contains the specialization of one of $z = \pm\sqrt{1-a}$. Consider the cover $(Y^{str})' \to Z^{str}$, where $(Y^{str})' = Y^{str}/Q$ and Q is the unique subgroup of G^{str} of order $p^{\nu - v(s)}$. After multiplying by $p^{v(s)}$th powers, this cover can be given by the equation

$$y^{p^{v(s)}} = \left(\frac{z+1}{z-1}\right)^r = \left(\frac{2z}{z-1} - 1\right)^r.$$

Recall that, by Lemma 8.22, we have $v(s) = v(\sqrt{1-a})$. Plugging in $z = \pm\sqrt{1-a}$ and multiplying by $(-1)^r$, which is a $p^{v(s)}$th power in K^{str}, we get that $y^{p^{v(s)}} = 1 + \alpha$, with $v(\alpha) = v(s) = v(\sqrt{1-a})$. As in the previous paragraph, we conclude that there exists $u \in K_0(a)$ with $v(u) = 1$ such that $G_{K^{str}(\sqrt[p]{1+u})}$ fixes a point above $\overline{X}_c$.

By Proposition 8.31 (1), these are the only inseparable tails in the situation of (2b). Applying the criterion of [12, Proposition 4.9] finishes the proof of (2b).

To (2c): Assume we are in the situation of (2c). By Propositions 8.19 and 8.26, we have $a \in \mathbb{Q} \subset K_\nu$, so $K^{str} = K_\nu$ and u (from (2b)—the inseparable tail in that case still exists in this case) is in K_0.

Suppose $v(a) > 0$. Then there is a new inseparable $p^{\nu - v(a) - 1}$-tail $\overline{X}_{c'}$ containing the specialization of a $K' := K_0(\sqrt[p]{\eta})$-rational point, where $\eta \in \mathbb{Q}$ and $p \nmid v(\eta)$ (Proposition 8.31 (2)). Each component of $\overline{Z}^{str}$ above $\overline{X}_{c'}$ contains the specialization of the K'-rational point $z = d$ of Proposition 8.31 (2). Consider the cover $(Y^{str})' \to Z^{str}$, where $(Y^{str})' = Y^{str}/Q$ and Q is the unique subgroup of G^{str} of order $p^{\nu - v(a) - 1}$. This cover is given by the equation $y^{p^{v(a)+1}} = g(z)$. Now, by Lemma 8.33 (1), $g(d) = (-1)^{r+s} + \alpha'$, with $v(\alpha') = v(a) + 1$. Since $(-1)^{r+s}$ is a $p^{v(a)+1}$st power in K_ν, we may assume $g(d) = 1 \pm \alpha'$. By the binomial expansion, $1 \pm \alpha'$ has a $p^{v(a)}$th root in K', which is of the form $1 + u'$ with $v(u') = 1$. So $G_{K'(\sqrt[p]{1+u'})}$ fixes a point above $\overline{X}_c$ for this cover. Since the quotient by Q is radicial above $\overline{X}_c$, it follows that $G_{K'(\sqrt[p]{1+u'})}$ fixes a point above $\overline{X}_{c'}$.

If, instead, $v(a-1) > 0$, the exact same proof (using Proposition 8.31 (3) instead of (2) and Lemma 8.33 (2) instead of (1)) shows that we can find $\eta \in \mathbb{Q}$ with $p \nmid v(\eta)$ and $u' \in K' := K_0(\sqrt[p]{\eta})$ such that $G_{K'(\sqrt[p]{1+u'})}$ fixes a point above the new inseparable tail $\overline{X}_{c'}$. We have now addressed all the inseparable tails (Proposition 8.31), so we can apply the criterion of [12, Lemma 4.9] to complete the proof of (2c). □

Proposition 8.36. *In all cases of Proposition 8.35, the stable model f^{st} of f can be defined over $(K^{str})^{st}$.*

Proof. Since the branch loci of f^{aux} and f^{str} are the same, all branch points of $Y^{aux} \to Y^{str}$ are ramification points of f^{str}. Thus their specializations do not coalesce on $\overline{f}^{str}$, and $G_{(K^{str})^{st}}$ must permute them trivially. So they are each defined over $(K^{str})^{st}$. By Lemma 7.4, f^{st} is defined over $(K^{str})^{st}$. □

8.3.7. *Higher ramification groups.* In this section, we calculate the bounds on the conductors of the fields in Proposition 8.35, considered as extensions of K_0. Recall that, for any finite Galois extension L/K (K a finite extension of K_0), we write $h_{L/K}$ for the conductor of L/K.

Proposition 8.37. (1) *If L is a tame extension of K_ν ($\nu \geq 1$), then L/K_0 is Galois and $h_{L/K_0} = \nu - 1$.*

(2) *Let $\eta \in K_0$ be such that $p \nmid v(\eta)$, let $u \in K_0$ such that $v(u) = 1$, and let $u' \in K_0(\sqrt[p]{\eta})$ such that $v(u') = 1$. Let $\nu > 1$, and let K' be a tame extension of $K := K_\nu(\sqrt[p]{\eta})(\sqrt[p]{1+u}, \sqrt[p]{1+u'})$. If L is the Galois closure of K' over K_0, then $h_{L/K_0} = \max(\nu - 1, \frac{p}{p-1})$.*

Proof. To (1): By [21, IV, Corollary to Proposition 18], $h_{K_\nu/K_0} = \nu - 1$. Note that any tame extension of a Galois extension of K_0 is, in fact, Galois over K_0. So L/K_0 is Galois. By [12, Lemma 2.2], its conductor is also $\nu - 1$.

To (2): Fix an algebraic closure $\overline{K_0}$ of K_0. Let K'' be the Galois closure of K over K_0. Then, because any tame extension of a Galois extension of K_0 is Galois over K_0, we see that the compositum $K'K''$ is Galois over K_0, thus $L = K'K''$. In particular, L/K'' is tame. By [12, Lemma 2.2], it suffices to show that $h_{K''/K_0} = \max(\nu - 1, \frac{p}{p-1})$.

Let u'_i, $1 \leq i \leq c$, be the distinct Galois conjugates of u' in an algebraic closure of K_0. Then K'' is the compositum of K_ν with

$$M := K_1(\sqrt[p]{\eta})(\sqrt[p]{1+u}, \sqrt[p]{1+u'_1}, \ldots \sqrt[p]{1+u'_c}).$$

Note that M/K_0 is Galois. The conductor of a compositum is the maximum of the conductors ([12, Lemma 2.3]), so $h_{K''/K_0} = \max(\nu - 1, h_{M/K_0})$. It will suffice to show that $h_{M/K_0} = \frac{p}{p-1}$.

Note that the absolute ramification index of K_1 is $p-1$. By [14, Lemma 3.2(ii)], $h_{K_1(\sqrt[p]{\eta})/K_1} = p$. Since the lower numbering is invariant under subgroups, the greatest lower jump for the higher ramification filtration of $G(K_1(\sqrt[p]{\eta})/K_0)$ is p. Then $h_{K_1(\sqrt[p]{\eta})/K_0} = \frac{p}{p-1}$ by the definition of the upper numbering.

Now, by [14, Lemma 3.2, Remark 3.4],

$$h_{K_1(\sqrt[p]{\eta})(\sqrt[p]{1+u'_i})/K_1(\sqrt[p]{\eta})} < \frac{p}{p-1}(p(p-1)) - p(p-1) = p.$$

The same holds for $K_1(\sqrt[p]{\eta})(\sqrt[p]{1+u})/K_1(\sqrt[p]{\eta})$. Thus, again using [12, Lemma 2.3], we see that $h_{M/K_1(\sqrt[p]{\eta})} < p$. By [14, Lemma 2.1] (with our K_0, $K_1(\sqrt[p]{\eta})$, and M playing the roles of K, L, and M in [14]), either $h_{M/K_0} = \frac{p}{p-1}$ or $h_{M/K_0} > \frac{p}{p-1}$ and

$$\frac{1}{p(p-1)}(h_{M/K_1(\sqrt[p]{\eta})} - p) = h_{M/K_0} - \frac{p}{p-1}.$$

Since the left hand side is negative whereas the right hand side is positive, the second option cannot hold. So $h_{M/K_0} = \frac{p}{p-1}$. ☐

Proof of Proposition 8.4. Note that in Proposition 8.35 (2b), either $a \in K_0$ or $a \in K_0(\sqrt[p]{\eta})$, for some $\eta \in \mathbb{Q}$ with $p \nmid v_p(\eta)$. Thus Proposition 8.37 shows that the Galois closures (over K_0) of all of the extensions $(K^{str})^{st}$ in Proposition 8.35 have conductor $< \nu$ over K_0. This, along with Proposition 8.36 and the fact that $\nu \leq n$, completes the proof of Proposition 8.4. ☐

Proof of Theorem 1.1. By Proposition 8.37 (1), the extensions in Propositions 8.1 and 8.3 have conductor $< n$ over K_0. This fact, combined with Proposition 8.4 and Remark 1.2, proves Theorem 1.1. ☐

9. Further questions

Question 9.1. Does Theorem 1.1 hold even if we allow $p = 3$ or no prime-to-p branch points?

If there are no prime-to-p branch points ($\tau = 0$, in the language of §8), then $\overline{X}$ can have up to two new (étale) tails. This allows for a greater proliferation of subcases (for instance, the two new tails could branch out from the same point of the original component). Techniques similar to those used in §8.3 should work, and we have worked out some of the easier cases (unpublished). However, keeping track of all the possibilities will be quite tedious, and we do not pursue the calculation here.

If we allow $p = 3$ (even if we assume that there is a branch point of f with prime-to-p branching index), then the stable reduction of f^{str} looks very different than what we determine in §8.3. In particular, Lemma 8.6 does not hold. In fact, if $\overline{X}_b$ is an étale tail with ramification invariant $\sigma_b = \frac{3}{2}$, then it *must* intersect a p^2-component (because the conductor of a $\mathbb{Z}/3$-extension in residue characteristic 3 cannot be 3). To pursue this case using the techniques of §8.3 would require good, explicit conditions characterizing the following extensions, in analogy with Lemma 2.1: Say R is a complete discrete valuation ring with fraction field K of characteristic 0, residue field k algebraically closed of characteristic 3, and uniformizer π. Suppose further that R contains the 9th roots of unity. Write $A = R\{t\}$ and $L = \mathrm{Frac}(A)$. We wish to characterize $\mathbb{Z}/9$-extensions M of L such that the normalization B of A in M satisfies the following conditions:

- Spec $B/\pi \to$ Spec A/π is an étale extension with conductor 3.
- Spec B/π is integral.

For a similar statement in residue characteristic 2, see [12, Proposition C.1].

Question 9.2. What if the condition $m_G = 2$ is relaxed in Theorem 1.1?

Allowing arbitrary m_G will require some new techniques, as the strong auxiliary cover is in general no longer a p^ν-cyclic extension of $\mathbb{P}^1$, making it difficult to perform explicit computations. However, it turns out to be true that, if the cover has bad reduction, the deformation data above the original component of $\overline{X}$ are all *multiplicative*, even for arbitrary m_G (this holds for $m_G = 2$, by adapting Lemma 8.5 to the $\tau = 0$ and $\tau = 2$ cases). One might hope to obtain results using variations on the deformation theory of torsors under multiplicative group schemes that Wewers develops in [24]. In the case where $v_p(|G|) = 1$, such deformation theory leads to a conceptual understanding of where the disks corresponding to the étale tails of the stable reduction (as in §8.3.4) are located, in particular showing that such disks must contain rational points over a small field. It also does not rely at all on having $m_G = 2$. If one considers a more generalized version of Theorem 1.1 where no restrictions on m_G or the branching indices are required, then generalizing [24] to the case of larger cyclic p-Sylow subgroups might provide a more conceptual proof (in particular, with regards to the analog of §8.3.4).

Appendix A. An Example of Wild Monodromy

In [15, Theorem 1.1], it is proven that if $f : Y \to X$ is a three-point G-cover defined over a complete discrete valuation field K of mixed characteristic $(0, p)$, where G has a cyclic p-Sylow subgroup of order p^n and p does not divide the order of the center of G, then the wild monodromy group Γ_w of f (i.e., the p-Sylow subgroup of $\mathrm{Gal}(K^{st}/K)$, notation of §4) has exponent dividing p^{n-1}. In particular, if $n = 1$, then Γ_w is trivial (this is [19, Théorème 4.2.10]). In this appendix, we exhibit an example showing that the wild monodromy can be nontrivial when $n > 1$. This is surprisingly difficult (see Remark A.4). Our example is based on the calculations of §8.3.

Throughout the appendix, let $G = SL_2(251)$, and let k be an algebraically closed field of characteristic $p = 5$. Let $R_0 = W(k)$ and $K_0 = \mathrm{Frac}(R_0)$. Lastly, let $K = K_0(\mu_{5^\infty})$ (that is, we adjoin all 5th-power roots of unity to K), and let R be the valuation ring of K. Note that G has a cyclic 5-Sylow subgroup of order $5^3 = 125$ and $m_G = 2$. We normalize all valuations on R_0, K_0, or any extensions thereof so that $v(5) = 1$.

Proposition A.1. *There exists a three-point G-cover $f : Y \to X = \mathbb{P}^1_K$, defined over K, such that the branching indices of the three branch points are e_1, e_2, and e_3, with $v_5(e_1) = 0$, $v_5(e_2) = 2$, and $v_5(e_3) = 3$.*

Proof. We show that such a cover can be defined over $\mathbb{Q}^{ab}$. Since $\mathbb{Q}^{ab} \hookrightarrow K$, this will prove the proposition.

Let $\alpha = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \in G$. This has order 251. We claim there exists $\beta = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in G$ satisfying the following properties:

- The order of β is 250.
- The order of $\alpha\beta$ is 50.
- The matrices α and β generate $SL_2(251)$.

To prove the claim, first note that any $GL_2(251)$-conjugacy class in G is determined by the trace of the matrices it contains, unless the trace is ± 2. In particular, the trace of a matrix determines its order if it is not ± 2. Let τ be the trace of the matrices in some conjugacy class of order 250, and let ρ be the trace of the matrices in some conjugacy class of order 50. Then τ, ρ, 2, and -2 are pairwise distinct. Choose a, b, c, and d in $\mathbb{F}_{251}$ solving the (clearly solvable) system of equations:

$$
\begin{aligned}
a + d &= \tau \\
a + c + d &= \rho \\
ad - bc &= 1
\end{aligned}
$$

Since the trace of $\alpha\beta$ is $a + c + d$, these equations ensure that β and $\alpha\beta$ have the desired orders. Let $\overline{\alpha}$ and $\overline{\beta}$ be the images of α and β in $H := PSL_2(251)$. Since $c \neq 0$, one checks that $\overline{\beta}$ does not normalize the subgroup generated by $\overline{\alpha}$. Then, by [9, II, Hauptsatz 8.27], we have that $\overline{\alpha}$ and $\overline{\beta}$ generate H. Furthermore, since β is diagonalizable over $GL_2(251)$ and has eigenvalues of order 250, we have $\beta^{125} = -I_2$. Since $\overline{\alpha}$ and $\overline{\beta}$ generate H, and β generates $\ker(G \to H)$, then α and β generate G.

Consider the triple $([\alpha], [\beta], [\alpha\beta]^{-1})$ of conjugacy classes of G. By [11, I, Theorem 5.10 and Remark afterward], this triple is rigid. By [11, I, Theorem 4.8], there exists a three-point G-cover of $\mathbb{P}^1$, defined over $\mathbb{Q}^{ab}$, with branching indices $e_1 = \mathrm{ord}(\alpha) = 251$, $e_3 = \mathrm{ord}(\beta) = 250$, and $e_2 = \mathrm{ord}((\alpha\beta)^{-1}) = 50$. This completes the proof of the proposition. $\square$

Proposition A.2. *If $f : Y \to X = \mathbb{P}^1_K$ is a cover satisfying the properties of Proposition A.1, then f has nontrivial wild monodromy Γ_w.*

Proof. To fix notation, we assume f is branched at $x = 0$, $x = 1$, and $x = \infty$ of index e_1, e_2, and e_3, respectively, with $v_5(e_1) = 0$, $v_5(e_2) = 2$, and $v_5(e_3) = 3$. By [13, Lemma 3.2], the stable reduction of f has both a primitive étale tail and a new étale tail. Construct the strong auxiliary cover $f^{str} : Y^{str} \to X$ of f (§7). This is a four-point G^{str}-cover, with $G^{str} \cong \mathbb{Z}/125 \rtimes \mathbb{Z}/2$ such that the action of $\mathbb{Z}/2$ is faithful. As in (8.1) and (8.2), this cover is given by

$$z^2 = \frac{x-a}{x} \tag{A.1}$$

$$y^{125} = g(z) := \left(\frac{z+1}{z-1}\right)^r \left(\frac{z+\sqrt{1-a}}{z-\sqrt{1-a}}\right)^s, \tag{A.2}$$

where r and s are integers satisfying $v_5(r) = 0$ and $v_5(s) = 1$. Replacing y with a prime-to-5 power, we can assume $s = 5$. By Lemma 8.7, we have $v(1-a) > 0$ in $K(a)/K$, and then Lemma 8.22 and Proposition 8.26 (1) show that we can take $a = 1 - \frac{25}{r^2}$. In particular, f^{str} is defined over K. By Proposition 8.31, the stable model $(f^{str})^{st} : (Y^{str})^{st} \to X^{st}$ of f^{str} has a new inseparable 5-tail $\overline{X}_c$. We claim that there is an extension L/K such that $\mathrm{Gal}(L/K)$ fixes $\overline{X}_c$ pointwise, and acts nontrivially of order 5 on the stable reduction $\overline{f}^{str} : \overline{Y}^{str} \to \overline{X}$ of f^{str} above $\overline{X}_c$. Since $(f^{str})^{st}$ is a quotient of the stable model $(f^{aux})^{st}$ of the (standard) auxiliary cover f^{aux} (§7), then $\mathrm{Gal}(L/K)$ will act nontrivially of order divisible by 5 above $\overline{X}_c$ in $(f^{aux})^{st}$ as well. Lastly, since, above an étale neighborhood of $\overline{X}_c$, the stable model f^{st} of f is isomorphic to a set of disconnected copies of $(f^{aux})^{st}$, the same holds true over a formal neighborhood $\hat{X} \subseteq X^{st}$ of $\overline{X}_c$. That is,

$$Y^{st} \times_{X^{st}} \hat{X} \cong \mathrm{Ind}_{G^{aux}}^{G} (Y^{aux})^{st} \times_{X^{st}} \hat{X}.$$

Since f is defined over K, the $\mathrm{Gal}(L/K)$-action on $Y^{st} \times_{X^{st}} \hat{X}$ is determined by the action on $(Y^{aux})^{st} \times_{X^{st}} \hat{X}$ and the fact that it commutes with the G-action. So the action of $\mathrm{Gal}(L/K)$ on $\overline{Y} \times_{\overline{X}} \overline{X}_c$, thus on $\overline{Y}$, is nontrivial of order divisible by 5. This means that f has nontrivial wild monodromy.

It remains to prove the claim. Let $Z^{str} = Y^{str}/(\mathbb{Z}/125)$, with stable model $(Z^{str})^{st}$ and stable reduction $\overline{Z}^{str}$. Then z is a coordinate on Z^{str}, and by Proposition 8.31 (3), there is a component of $\overline{Z}^{str}$ above $\overline{X}_c$ containing the specialization $\overline{d}$ of

$$z = d := \frac{2 \cdot 5^{7/5}}{r},$$

where we can use any choice of 5th root. Since $\overline{X}_c$ is a p-component, there are 25 points of $\overline{Y}^{str}$ above $\overline{d}$. If $g(d)$ is a 5th power, but not a 25th power, in $K(d) = K(\sqrt[5]{5})$, and if $L = K(d, \sqrt[25]{g(d)})$, then the action of $\mathrm{Gal}(L/K(d))$ will permute these 25 points in orbits of order 5, and we will be done (it turns out that $\mathrm{Gal}(K(d)/K)$ fixes $\overline{d}$, even though it clearly does not fix d). This follows from Lemma A.3 below. $\square$

Lemma A.3. *Let $d = \frac{2 \cdot 5^{7/5}}{r}$, where r is a prime-to-p integer and we choose any pth root of* 5*. Let g be the rational function in (A.2), with $s = 5$ and $a = 1 - \frac{25}{r^2}$. Then $g(d)$ is a* 5*th power, but not a* 25*th power, in $K(\sqrt[5]{5})$.*

Proof. Fix a 5th root of 5 in $\overline{K}$, which we will denote by either $\sqrt[5]{5}$ or $5^{1/5}$. We first note that $g(d) \in K_0(\sqrt[5]{5})$. By (8.20), we have

$$g(d) = \pm\left(1 - \frac{8r^3}{75}d^3 - \frac{32r^5}{5^5}d^5\right) + o(5^{9/4}),$$

where the o represents terms of valuation greater than $\frac{9}{4}$. Upon plugging in d and simplifying, this gives

$$g(d) = \pm(1 - 3 \cdot 5^{11/5} - 4 \cdot 5^2) + o(5^{9/4}).$$

Using the binomal theorem, we see that $g(d)$ has a 5th root η in $K_0(\sqrt[5]{5})$, and

$$\eta = \pm(1 - 3 \cdot 5^{6/5} - 20) + o(5^{5/4}).$$

We wish to show that η is not a 5th power in $K(\sqrt[5]{5})$.

Now, since $K(\sqrt[5]{5})/K_0(\sqrt[5]{5})$ is abelian, any subextension is Galois. So if η is a 5th power in $K(\sqrt[5]{5})$, then taking a 5th root of η must generate a Galois extension of $K_0(\sqrt[5]{5})$. This is clearly not the case unless η is already a 5th power in $K_0(\sqrt[5]{5})$, so it suffices to show that η is not a 5th power in $K_0(\sqrt[5]{5})$. Since -1 is a 5th power in K_0, we may assume that $\eta = 19 + 3 \cdot 5^{6/5} + o(5^{5/4})$.

Suppose that $\theta \in K_0(\sqrt[5]{5})$ such that $\theta^5 = \eta$, and write

$$\theta = \alpha + \beta \cdot 5^{1/5} + \gamma \cdot 5^{2/5} + \delta \cdot 5^{3/5} + \epsilon \cdot 5^{4/5},$$

where α, β, γ, δ, and ϵ are in K_0. Comparing valuations, we see that $\theta \in R_0[\sqrt[5]{5}]$. Equating coefficients of 1 and $5^{6/5}$ gives the equations

$$\alpha^5 + 5\beta^5 \equiv 19 \pmod{25}$$
$$\alpha^4 \beta \equiv 3 \pmod{5}$$

The second equation yields $\alpha \equiv 4 \pmod 5$, and then the first equation yields $\beta \equiv 3 \pmod 5$. But then $\alpha^5 + 5\beta^5 \equiv 24 + 5 \cdot 18 \equiv 14 \not\equiv 19 \pmod{25}$. So ϵ cannot exist, and we are done. ☐

Remark A.4. The example above is quite complicated, and is not generalizable in any meaningful way (for instance, it depends critically on having $p = 5$). One hopes for easier examples, but they are difficult to come by. For instance, results of [12, §7.1] show that no examples of three-point G-covers with nontrivial wild monodromy can exist when G is p-solvable and $m_G > 1$. So if one wants to find an easier example where p does not divide the order of the center of G, one needs to look either at a group that is not p-solvable, or at a group where $m_G = 1$. By Burnside's theorem (see, e.g., [15, Lemma 2.2]), having $m_G = 1$ implies that G is of the form $G \cong H \rtimes \mathbb{Z}/p^n$, where the action of $\mathbb{Z}/p^n$ on H is faithful.

References

[1] S. Beckmann, *Ramified primes in the field of moduli of branched coverings of curves.* J. Algebra **125** (1989), 236–255.

[2] I. I. Bouw and S. Wewers, *Reduction of covers and Hurwitz spaces.* J. Reine Angew. Math. **574** (2004), 1–49.

[3] K. Coombes and D. Harbater, *Hurwitz families and arithmetic Galois groups.* Duke Math. J. **52** (1985), 821–839.

[4] P. Deligne and D. Mumford, *The irreducibility of the space of curves of given genus.* Inst. Hautes Études Sci. Publ. Math. **36** (1969), 75–109.

[5] P. Deligne and M. Rapoport, *Les schémas de modules de courbes élliptiques.* Modular functions of one variable II, LNM **349**, Springer-Verlag (1972), 143–316.

[6] W. Fulton, *Hurwitz schemes and irreducibility of moduli of algebraic curves.* Ann. of Math. **90** (1969), no. 2, 542–575.

[7] Y. Henrio, *Disques et couronnes ultramétriques.* Courbes semi-stables et groupe fondamental en géométrie algébrique, Progr. Math., **187**, Birkhäuser Verlag, Basel (1998), 21–32.

[8] Y. Henrio, *Arbres de Hurwitz et automorphismes d'ordre p des disques et des couronnes p-adiques formels.* arXiv:math/0011098

[9] B. Huppert, *Endliche gruppen.* Springer-Verlag, Berlin, 1987.

[10] N. Katz, *Local-to-global extensions of fundamental groups.* Ann. Inst. Fourier, Grenoble **36** (1986), 69–106.

[11] G. Malle and B. H. Matzat, *Inverse Galois theory.* Springer-Verlag, Berlin, 1999.

[12] A. Obus, *Fields of moduli of three-point G-covers with cyclic p-Sylow, I.* Algebra Number Theory **6** (2012), no. 5, 833–883.

[13] A. Obus, *Toward Abhyankar's inertia conjecture for $PSL_2(\ell)$.* Groupes de Galois géométriques et différentiels, Séminaires et Congrès, **27**, Société Mathématique de France (2013), 191–202.

[14] A. Obus, *Conductors of extensions of local fields, especially in characteristic* $(0, 2)$. To appear in Proc. Amer. Math. Soc.

[15] A. Obus, *Vanishing cycles and wild monodromy.* Int. Math. Res. Notices (2012), 299–338.

[16] A. Obus and S. Wewers, *Cyclic extensions and the local lifting problem.* To appear in Ann. of Math.

[17] R. Pries, *Wildly ramified covers with large genus.* J. Number Theory **119** (2006), 194–209.

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

[19] M. Raynaud, *Specialization des revêtements en caractéristique* $p > 0$. Ann. Sci. École Norm. Sup. **32** (1999), 87–126.

[20] J.-P. Serre, *Géométrie algébrique et géométrie analytique*, Ann. Inst. Fourier, Grenoble **6** (1955–1956), 1–42.

[21] J.-P. Serre, *Local fields.* Springer-Verlag, New York, 1979.

[22] S. Wewers, *Reduction and lifting of special metacyclic covers.* Ann. Sci. École Norm. Sup. (4) **36** (2003), 113–138.

[23] S. Wewers, *Three point covers with bad reduction.* J. Amer. Math. Soc. **16** (2003), 991–1032.

[24] S. Wewers, *Formal deformation of curves with group scheme action.* Ann. Inst. Fourier **55** (2005), 1105-1165.

[25] H. J. Zassenhaus, *The theory of groups, 2nd ed..* Chelsea Publishing Company, New York, 1956.

Andrew Obus
University of Virginia
141 Cabell Drive
Charlottesville, VA 22904
E-mail: andrewobus@gmail.com
URL: people.virginia.edu/~aso9t/