JOURNAL de Théorie des Nombres de BORDEAUX

anciennement Séminaire de Théorie des Nombres de Bordeaux

Christopher KEYES

Growth of points on hyperelliptic curves over number fields

Tome 34, nº 1 (2022), p. 271-294.

https://doi.org/10.5802/jtnb.1201

Journal de Théorie des Nombres
de Bordeaux **34** (2022), 271–294

Growth of points on hyperelliptic curves over number fields

par Christopher KEYES

Résumé. Choisissons une courbe hyperelliptique $C/\mathbb{Q}$ de genre g et considérons les corps de nombres $K/\mathbb{Q}$ engendrés par les points algébriques de C. Dans cet article, nous étudions le nombre de telles extensions de degré fixé n et de discriminant inférieur ou égal à X. Nous montrons que lorsque $g \geq 1$ et n est suffisamment grand par rapport au degré de C (en supposant que n est pair si $\deg C$ est un nombre pair), ce nombre est $\gg X^{c_n}$, où c_n est une constante positive dépendant de g, qui tend vers $1/4$ lorsque $n \to \infty$. Ce résultat s'appuie sur le travail de Lemke Oliver et Thorne qui, dans le cas où C est une courbe elliptique, donnent une minoration pour le nombre d'extensions de degré fixé et de discriminant borné sur lesquelles le rang de C augmente avec une constante locale spécifiée.

Abstract. Fix a hyperelliptic curve $C/\mathbb{Q}$ of genus g, and consider the number fields $K/\mathbb{Q}$ generated by the algebraic points of C. In this paper, we study the number of such extensions with fixed degree n and discriminant bounded by X. We show that when $g \geq 1$ and n is sufficiently large relative to the degree of C, with n even if $\deg C$ is even, there are $\gg X^{c_n}$ such extensions, where c_n is a positive constant depending on g which tends to $1/4$ as $n \to \infty$. This result builds on work of Lemke Oliver and Thorne who, in the case where C is an elliptic curve, put lower bounds on the number of extensions with fixed degree and bounded discriminant over which the rank of C grows with specified root number.

1. Introduction

Let C be a smooth projective curve over $\mathbb{Q}$ and fix an algebraic closure $\overline{\mathbb{Q}}$. We say a field $K/\mathbb{Q}$ is *generated by a point of* C if $K = \mathbb{Q}(P)$ for some $P \in C(\overline{\mathbb{Q}})$. That is, K is the minimal field of definition for an algebraic point on C. For $n \geq 1$ an integer and X a positive real number, we define the quantity $N_{n,C}(X)$ to be the number of such extensions with degree $[K : \mathbb{Q}] = n$ and bounded absolute discriminant $|\mathrm{Disc}(K)| \leq X$. We also take $N_{n,C}(X, G)$ to be the number of those extensions with $\mathrm{Gal}(\widetilde{K}/\mathbb{Q}) \simeq G$, where $\widetilde{K}$ denotes the Galois closure of K.

Manuscrit reçu le 25 septembre 2020, révisé le 7 février 2021, accepté le 20 mai 2022.
Mathematics Subject Classification. 11G30, 12F05, 12E05.
Mots-clefs. Arithmetic statistics, hyperelliptic curves, Diophantine stability.

In their paper on Diophantine Stability, Mazur and Rubin [9] ask to what extent the set of fields generated by algebraic points determines the identity of the curve C. Motivated by this question, we want to understand how $N_{n,C}(X)$ grows as $X \to \infty$, and how this asymptotic depends on both the geometry of C and the degree n. When C is an elliptic curve, Lemke Oliver and Thorne [8] show there are $\gg X^{c_n - \epsilon}$ number fields $K/\mathbb{Q}$ of degree $n \geq 2$ and discriminant at most X, such that the Mordell–Weil rank of $C(K)$ is greater than that of $C(\mathbb{Q})$, and C/K has specified root number. Here c_n is a positive constant and tends to $1/4$ from below as $n \to \infty$.

In this paper, we consider the case where C is a hyperelliptic curve. Recall a hyperelliptic curve $C/\mathbb{Q}$ is given by an affine equation

$$C\colon y^2 = f(x),$$

where $f(x) \in \mathbb{Q}[x]$. If $f(x)$ is separable then C is nonsingular, and its genus g is related to its degree $d = \deg f$ by

$$d = \begin{cases} 2g+1 & d \text{ is odd} \\ 2g+2 & d \text{ is even.} \end{cases}$$

Our main result is an asymptotic lower bound for $N_{n,C}(X, S_n)$ when n is large relative to d, which generalizes that of Lemke Oliver and Thorne and recovers their bound when $g = 1$. We treat the cases of d odd and even separately in Theorems 1.1 and 1.2. In both cases, the implied constants depend on the degree n of the extension and the model f, and we are able to improve our results slightly when n is allowed to be sufficiently large.

Theorem 1.1. *Let C be a hyperelliptic curve with genus $g \geq 1$ and degree $d = 2g+1$. If $n \geq d$, then*

$$N_{n,C}(X, S_n) \gg X^{c_n}$$

where

$$c_n = \frac{1}{4} - \frac{gn^2 - (g^2 - 2g - 3)n - 2g^2}{2n^2(n-1)}.$$

Moreover, if n is sufficiently large, we have the improvement

$$c_n = \frac{1}{4} - \frac{gn + g^2 - 2g}{2n(n-1)}.$$

Theorem 1.1 applies whenever C has a rational Weierstrass point, as we can choose an equation for C with d odd. In the general case, when d is even, we restrict our attention to even n. This turns out to be a necessary restriction because a positive proportion of hyperelliptic curves over $\mathbb{Q}$ will have no points over any odd degree extensions, a result due to Bhargava, Gross, and Wang [1]. After making this restriction, we obtain a similar asymptotic bound to Theorem 1.1.

Theorem 1.2. *Let C be a hyperelliptic curve with genus $g \geq 1$ and degree $d = 2g + 2$. If $n \geq d + 2$ is even, then*

$$N_{n,C}(X, S_n) \gg X^{c_n}$$

where

$$c_n = \frac{1}{4} - \frac{(1 + 2g)n^2 - (2g^2 - 2g - 8)n - (4g^2 + 4g)}{4n^2(n-1)}.$$

Moreover, when n is sufficiently large, we have the improvement

$$c_n = \frac{1}{4} - \frac{(1 + 2g)n - 2g^2 + 2g + 2}{4n(n-1)}.$$

Remark 1.3. In both cases, the exponent c_n tends to $1/4$ from below as $n \to \infty$. If $d > 7$ is odd then c_n is positive for all $n \geq d$. Similarly, if $d \geq 4$ is even then c_n is positive for all $n \geq d + 2$. We discuss how to find the threshold where the improved exponent applies and give examples in Section 5.4.

We contrast Theorems 1.1 and 1.2 with a result of Granville [5] for quadratic twists of hyperelliptic curves, which tells a very different story for quadratic extensions. Granville proved, assuming the *abc*-conjecture, that when $g \geq 2$, the number of squarefree d such that $|d| \leq D$ and the quadratic twist

$$C_d \colon dy^2 = f(x)$$

has a nontrivial rational point is $\ll D^{1/(g-1)+o(1)}$. Here, nontrivial refers to points which don't arise from roots of $f(x)$ or points at infinity. Such points on twists give rise to points in $C(\mathbb{Q}(\sqrt{d}))$, suggesting an *upper* bound on $N_{2,C}(X)$ with vanishing exponent as $g \to \infty$.

The proofs of Theorems 1.1 and 1.2 employ a similar strategy as that used by Lemke Oliver and Thorne for elliptic curves and large degree fields. The approach is to produce a family of polynomials whose roots give rise to points on C. We will contrive this family to consist almost entirely of irreducible polynomials of the desired degree n and Galois group S_n. Then we count the family, adjusting for multiplicity of the fields generated, to give a lower bound for $N_{n,C}(X)$.

In Section 2, we recall the necessary Galois theory to justify using specializations to study irreducibility and Galois groups in polynomial families. We then introduce Newton polygons as a tool to determine Galois groups of polynomials. We apply these results in Section 3 to specific families to show that they are populated by irreducible polynomials with Galois group S_n. In Section 4, we state and prove a useful lemma relating the size of a polynomial's roots to its coefficients.

These ingredients are assembled in Section 5 into the proofs of Theorems 1.1 and 1.2. Here, we count specializations of our polynomial families while controlling multiplicity. We show that the contribution by fields with small discriminant is negligible, which improves our final lower bounds slightly. We make further improvements when n is sufficiently large by applying the best known upper bounds on the number of fixed degree number fields with bounded discriminant due to Lemke Oliver and Thorne [7].

Acknowledgments. The author would like to thank Robert Lemke Oliver for suggesting the topic and for his invaluable guidance, as well as Lea Beneish, Jackson Morrow, and David Zureick-Brown for their helpful comments on previous drafts. The author also thanks the referee for carefully reading the manuscript and making suggestions which improved the paper.

2. Irreducibility and Galois groups in families of polynomials

2.1. Hilbert's irreducibility theorem. A parameterized family of polynomials over $\mathbb{Q}$ is given by a polynomial $f(\boldsymbol{t}, x) \in \mathbb{Q}(\boldsymbol{t})[x]$, where $\boldsymbol{t} = (t_1, \dots t_k)$. If $\boldsymbol{t}_0 \in \mathbb{Q}^k$, then $f_{\boldsymbol{t}_0} = f(\boldsymbol{t}_0, x) \in \mathbb{Q}[x]$ is a *specialization* of f. We would like to understand how the irreducibility of f over $\mathbb{Q}(\boldsymbol{t})$ is related to that of its specializations $f_{\boldsymbol{t}_0}$ over $\mathbb{Q}$. Moreover, when $f_{\boldsymbol{t}_0}$ is irreducible, we would like to relate its Galois group $G_{\boldsymbol{t}_0}$ to that of f.

Keeping the notation above, suppose f is irreducible over $\mathbb{Q}(\boldsymbol{t})$. Then the field $K = \mathbb{Q}(\boldsymbol{t})[x]/f(\boldsymbol{t}, x)$ is a finite extension of $\mathbb{Q}(\boldsymbol{t})$ of degree $n = \deg f$. The Galois closure of $K/\mathbb{Q}(\boldsymbol{t})$ is denoted $\widetilde{K}$, allowing us to define the Galois group $G = \mathrm{Gal}(f/\mathbb{Q}(\boldsymbol{t})) = \mathrm{Gal}(\widetilde{K}/\mathbb{Q}(\boldsymbol{t}))$. Let $g(\boldsymbol{t}, x) \in \mathbb{Q}(\boldsymbol{t}, x)$ generate the extension $\widetilde{K}$, that is $\widetilde{K} \simeq \mathbb{Q}(\boldsymbol{t})[x]/g(\boldsymbol{t}, x)$. Again, we use $g_{\boldsymbol{t}_0}$ to denote the specialization $g(\boldsymbol{t}_0, x) \in \mathbb{Q}[x]$.

Theorem 2.1 (Hilbert irreducibility)**.** *Using the notation above, suppose $\boldsymbol{t}_0 \in \mathbb{Q}^k$ is such that $g_{\boldsymbol{t}_0}$ is irreducible over $\mathbb{Q}$. Then the permutation representations of G and $G_{\boldsymbol{t}_0}$ acting on the roots of f and $f_{\boldsymbol{t}_0}$ are isomorphic.*

Moreover, the proportion of specializations $g_{\boldsymbol{t}_0}$ which are irreducible is $1 - o_H(1)$ for $\boldsymbol{t}_0$ in any rectangular region in $\mathbb{Z}^k$ having shortest side length H.

The fact that almost all specializations of an irreducible polynomial are irreducible is classical. For a proof of the isomorphism of permutation representations, we refer the reader to [8, Theorem 4.1]. Theorem 2.1 tells us that once we know f, and hence g, is irreducible over $\mathbb{Q}(\boldsymbol{t})$, then 100% of its integral specializations $f_{\boldsymbol{t}_0}$ are irreducible over $\mathbb{Q}$ with Galois groups $G \simeq G_{\boldsymbol{t}_0}$ isomorphic as permutation groups.

From Theorem 2.1 we obtain the following corollary, also appearing in [8, Corollary 4.2].

Corollary 2.2. *Suppose $f(\boldsymbol{t}, x)$ is irreducible over $\mathbb{Q}(\boldsymbol{t})$. If the permutation representation of $G_{\boldsymbol{t}_0}$ contains an element of a given cycle type for a positive proportion of $\boldsymbol{t}_0 \in \mathbb{Z}^k$, then the permutation representation of G must also contain an element of that type.*

Proof. By Theorem 2.1, choose a sufficiently large rectangular region in $\mathbb{Z}^k$ such that the proportion of $\boldsymbol{t}_0$ for which G and $G_{\boldsymbol{t}_0}$ do not have isomorphic permutation representations is smaller than the proportion of $\boldsymbol{t}_0$ for which $G_{\boldsymbol{t}_0}$ contains an element of the given cycle type. Then there is some $\boldsymbol{t}_0$ in the region such that both $G \simeq G_{\boldsymbol{t}_0}$ and $G_{\boldsymbol{t}_0}$ contains the given cycle type, thus G must contain that cycle type. □

The upshot of Corollary 2.2 is that we need only prove that a positive proportion of integral specializations $f_{\boldsymbol{t}_0}$ have Galois group $G_{\boldsymbol{t}_0}$ containing a given cycle type to see that G does. Then by another application of Theorem 2.1, 100% of specializations contain an element with the given type. In particular, if the presence of certain cycle types in G implies that G is the full symmetric group S_n, then we need only find that positive proportions of specializations $f_{\boldsymbol{t}_0}$ have each of these types to see that 100% of specializations have Galois group S_n.

2.2. Recognizing the symmetric group. Let S_n denote the symmetric group on the set $\{1, \ldots, n\}$, and let $G \subseteq S_n$ be a permutation subgroup. Recall that G is a *transitive* subgroup if for all pairs i, j with $1 \leq i, j \leq n$, there exists an element $\sigma \in G$ for which $\sigma(i) = j$. We will describe several ways to detect that a transitive subgroup G is isomorphic to the full symmetric group S_n using the presence of certain cycle types.

Lemma 2.3 (Lifting transitive subgroups)**.** *Let $G \subseteq S_n$ be a transitive permutation subgroup on the set $\{1, \ldots, n\}$. Assume G contains a subgroup H which is isomorphic as a permutation subgroup to S_k for some $k > n/2$. Then $G \simeq S_n$.*

Proof. The result is clearly true if $k = n$, so assume $k < n$. After renumbering if necessary, we may assume H acts nontrivially on $\{1, \ldots, k\}$ and acts trivially on $\{k+1, \ldots, n\}$. In particular, G contains the transpositions $(1\ a)$ for $2 \leq a \leq k$. Let $\sigma_j \in G$ be an element such that $\sigma_j(1) = j$, which exists by the transitivity of G. Then $\sigma_j(a)$ takes on $k-1$ different values for $2 \leq a \leq k$, none of which are equal to j.

Set $j = k+1$, so we have that $\sigma_{k+1}(a)$ takes $k-1$ distinct values when $2 \leq a \leq k$, none of which are equal to $k+1$. The hypothesis that $k > n/2$ is equivalent to $k - 1 > n - k - 1$, so by the pigeonhole principle, there exists at least one such a for which $\sigma_{k+1}(a) \leq k$. Conjugating $(1\ a)$ by σ_{k+1}, we see that G contains the transposition $(\sigma_{k+1}(a)\ k+1)$. Together with the subgroup H, this transposition generates a subgroup of G isomorphic to S_{k+1}. Finitely many applications of this procedure show $G \simeq S_n$. □

Proposition 2.4. *Suppose $G \subseteq S_n$ is a transitive permutation subgroup containing a transposition, τ, and a cycle, σ, of length $n-1$ or length $p > n/2$ for p a prime. Then $G \simeq S_n$.*

Proof. Suppose first that σ has length $n-1$ and renumber so that σ is given by $(1 \ \dots \ n-1)$ in cycle notation. Write $\tau = (a\ b)$. Since G is transitive, we can conjugate τ by some element of G to produce a transposition $(n\ c)$ where $1 \leq c \leq n-1$. Conjugation of $(n\ c)$ by powers of σ produces $\{(n\ d) \mid 1 \leq d \leq n-1\} \subseteq G$, which is a generating set for S_n. Hence $G \simeq S_n$.

Now suppose that σ has length p for some prime $p > n/2$, and again renumber so $\sigma = (1 \ \dots \ p)$. Conjugating τ produces a transposition $(1\ b) \in G$ for some b. Suppose that $b > p$. Then conjugation of $(1\ b)$ by powers of σ produces the subset $\{(a\ b) \mid 1 \leq a \leq p\} \subseteq G$, which generates a subgroup $H \subseteq G$ isomorphic to S_{p+1} acting on $\{1, \dots, p, b\}$.

If instead $1 < b \leq p$, then $\sigma^i(1) = (1\ b\ \dots)$ is a p-cycle for some $1 \leq i < p$. We may renumber again such that $b = 2$, making our transposition $(1\ 2)$, and $\sigma^i = (1\ 2\ \dots\ p)$. Conjugating $(1\ 2)$ by powers of σ^i, we obtain $\{(c\ c+1) \mid 1 \leq c \leq p-1\} \subseteq G$, which is a generating set for S_p acting on $\{1, \dots, p\}$.

In either case, we have shown that there exists a permutation subgroup $H \subseteq G$ such that $H \simeq S_k$ for $k = p$ or $k = p+1$. Since $p > n/2$, the hypothesis of Lemma 2.3 applies, so we may conclude $G \simeq S_n$. □

2.3. Newton polygons.

We now introduce the Newton polygon, our tool for showing that a positive proportion of integral specializations f_{t_0} have certain cycle types in their Galois group. Let p be a prime, $\mathbb{Q}_p$ the field of p-adic numbers, and $f(x) \in \mathbb{Q}_p[x]$ a polynomial.

Definition 2.5 (Newton polygon)**.** With the notation above, let $f(x)$ be given by $f(x) = \sum_{i=0}^{n} k_i x^i$. The *Newton polygon of f* is the lower convex hull of the set

$$\left\{ (i, v_p(k_i)) \in \mathbb{R}^2 \,\middle|\, 0 \leq i \leq n \right\},$$

where v_p denotes the p-adic valuation, and we set $v_p(0) = \infty$ by convention. We will denote the Newton polygon of f by $\mathrm{NP}_{\mathbb{Q}_p}(f)$, and simply $\mathrm{NP}(f)$ when it will not create confusion.

The Newton polygon $\mathrm{NP}(f)$ can be split up into segments of distinct increasing slopes. The number and slopes of segments in the Newton polygon determine the valuations of the roots of $f(x)$ in $\mathbb{Q}_p$. More precisely, if a segment of the Newton polygon with length l has slope s then $f(x)$ has l roots each with valuation $-s$ in $\overline{\mathbb{Q}_p}$. For a proof, see [10, II.6]. This key fact allows us to prove two lemmas.

Lemma 2.6. *Suppose* $\mathrm{NP}(f)$ *has a segment of length l and slope s, and no other segments of this slope (i.e. consider the* entire *segment of slope s). Then f has a factorization $f = f_0 f_1$ over $\mathbb{Q}_p$, such that* $\deg f_0 = l$ *and the roots of f_0 have valuation $-s$.*

Moreover, if $s = r/l$ has reduced fraction form r'/l' then all irreducible factors of f_0 over $\mathbb{Q}_p$ have degree divisible by l'. In particular, if $\gcd(r, l) = 1$ *then the f_0 produced above is irreducible over $\mathbb{Q}_p$.*

Proof. Since the action of the Galois group $\mathrm{Gal}(f/\mathbb{Q}_p)$ on the roots of f preserves their valuations, we see that for an irreducible polynomial over $\mathbb{Q}_p$, all roots have the same valuation. Therefore, we can decompose f into irreducible factors and group together those whose roots have valuation $-s$ into f_0. This must have degree l, since f has exactly l roots with valuation $-s$.

For the second statement, we use the same observation above to recognize that the Newton polygon of an irreducible polynomial has one segment. Let g be an irreducible polynomial over $\mathbb{Q}_p$ dividing f_0. Then $\mathrm{NP}(g)$ has one segment of slope $s = r_g/\deg g$. Since reducing this fraction also produces r'/l', we must have $l' \mid \deg g$. □

Lemma 2.7. *Suppose $f(x) \in \mathbb{Q}[x]$, $p > \deg f$, and* $\mathrm{NP}_{\mathbb{Q}_p}(f)$ *has a segment of length l and slope r/l such that* $\gcd(r, l) = 1$. *If $f = f_0 f_1$ is the factorization from Lemma 2.6 and l is pairwise coprime to the degrees of the irreducible factors of f_1 over $\mathbb{Q}_p$, then* $\mathrm{Gal}(f/\mathbb{Q})$ *contains an l-cycle.*

Proof. We begin by factoring $f = f_0 f_1$ as in Lemma 2.6, noting that f_0 must be irreducible of degree l. Let $E_0, E_1, E = E_0 E_1$ be the splitting fields of f_0, f_1, f respectively, obtained by adjoining roots. Let T_0, T_1, T denote the maximal unramified subextensions of E_0, E_1, E over $\mathbb{Q}_p$. We will find an l-cycle in $\mathrm{Gal}(E_0/T_0)$ and make identifications

$$\mathrm{Gal}(E_0/T_0) \simeq \mathrm{Gal}(E_0 T_0 T_1/T_0 T_1) \simeq \mathrm{Gal}(E/E_1 T_0 T_1) \subseteq \mathrm{Gal}(E/\mathbb{Q}_p). \tag{2.1}$$

Identifying $\mathrm{Gal}(E/\mathbb{Q}_p)$ with $\mathrm{Gal}(f/\mathbb{Q}_p)$ and taking the natural inclusion into $\mathrm{Gal}(f/\mathbb{Q})$ gives the result.

Let $L = \mathbb{Q}_p[x]/(f_0(x))$, which is a degree l extension of $\mathbb{Q}_p$. By hypothesis, the set of p-adic valuations $v_p(L^\times)$ contains $\frac{r}{l}$, and $\gcd(r, l) = 1$ implies that $\frac{1}{l}\mathbb{Z} \subseteq v_p(L^\times)$. Hence, the ramification index of $L/\mathbb{Q}_p$ is equal to l, making $L/\mathbb{Q}_p$ totally ramified.

The hypothesis $p > \deg f$ implies $p \nmid l$, meaning $L/\mathbb{Q}_p$ is totally tamely ramified, so there exists a uniformizer $\pi \in L$ which satisfies $x^l - p = 0$ [6, II.5, Proposition 12], and thus $L = \mathbb{Q}_p(\pi)$. Whether or not $L/\mathbb{Q}_p$ is Galois, the Galois closure E_0 has an cyclic automorphism of order l, coming from a primitive l-th root of unity. This automorphism fixes T_0 and necessarily acts nontrivially on L, permuting the roots of f_0 in a cyclic fashion.

The first identification in (2.1) follows from elementary Galois theory, since $E_0 \cap T_0T_1 = T_0$ and E_0/T is a Galois extension. See e.g. [2, §14.4, Proposition 19].

For the second identification in (2.1), we remark that the ramification index $[E_1 : T_1]$ divides the product of the degrees of the irreducible factors of f_1 over $\mathbb{Q}_p$, which is coprime to l by our hypotheses. Moreover, we have $[E_1 : T_1] = [E_1T_0T_1 : T_0T_1]$ since $E_1 \cap T_0T_1 = T_1$. Thus $l = [E_0 : T_0] = [E_0T_0T_1 : T_0T_1]$ is coprime to $[E_1T_0T_1 : T_0T_1]$, so $E_0T_0T_1 \cap E_1T_0T_1 = T_0T_1$, and the identification of Galois groups follows from the same argument as the first one. □

To connect this result to families of polynomials, suppose we have an irreducible $f(\boldsymbol{t}, x) \in \mathbb{Q}(\boldsymbol{t})[x]$, which is equivalent to its Galois group G being a transitive subgroup of S_n. We look for integral specializations $\boldsymbol{t}_0 \in \mathbb{Z}^k$ such that for some prime p, the Newton polygon of $f_{\boldsymbol{t}_0}$ has a segment satisfying the hypothesis of Lemma 2.7, and hence $G_{\boldsymbol{t}_0}$ contains an l-cycle. If we find that we need only specify the p-adic valuations of the specialization $\boldsymbol{t}_0$ to obtain such a cycle, then $G_{\boldsymbol{t}_0}$ contains an l-cycle for a positive proportion of $\boldsymbol{t}_0 \in \mathbb{Z}^k$, and Corollary 2.2 implies that G contains an l-cycle as well. Repeating this procedure to find cycles of different lengths, we can hope to satisfy the hypotheses of Proposition 2.4 to see that G is in fact the full symmetric group, in which case Theorem 2.1 implies 100% of specializations have $G_{\boldsymbol{t}_0} = S_n$. We realize this procedure in the following section for specific polynomial families.

3. Polynomial families arising from hyperelliptic curves

Suppose $C/\mathbb{Q}$ is a nonsingular hyperelliptic curve given by

$$C\colon y^2 = f(x) = \sum_{i=0}^{d} c_i x^i$$

for a squarefree polynomial $f(x) \in \mathbb{Z}[x]$. In this section, we construct families of polynomials whose specializations give rise to number fields generated by algebraic points on C.

Let $g(x) = \sum_{i=0}^{d_g} a_i x^i \in \mathbb{Q}(\boldsymbol{a})[x]$ and $h(x) = \sum_{i=0}^{d_h} b_i x^i \in \mathbb{Q}(\boldsymbol{b})[x]$, where $\boldsymbol{a} = (a_0, \dots a_{d_g})$ and $\boldsymbol{b} = (b_0, \dots, b_{d_h})$. Then consider the polynomial in $\mathbb{Q}(\boldsymbol{a}, \boldsymbol{b})[x]$ given by

$$F_f(\boldsymbol{a}, \boldsymbol{b}, x) = g(x)^2 - f(x)h(x)^2, \tag{3.1}$$

which has degree $n = \max(2d_g, d + 2d_h)$. We will use $F_{f,\boldsymbol{a}_0,\boldsymbol{b}_0}(x)$ to denote a specialization with $\boldsymbol{a}_0 \in \mathbb{Q}^{d_g+1}$ and $\boldsymbol{b}_0 \in \mathbb{Q}^{d_h+1}$.

Let $\boldsymbol{a}_0, \boldsymbol{b}_0$ be rational specializations such that $F_{f,\boldsymbol{a}_0,\boldsymbol{b}_0}(x)$ has degree n over $\mathbb{Q}$, and suppose further that it is irreducible. Take α to be a root of

$F_{f,\boldsymbol{a}_0,\boldsymbol{b}_0}(x)$, which by rearranging (3.1) satisfies

$$\left(\frac{g_{\boldsymbol{a}_0}(\alpha)}{h_{\boldsymbol{b}_0}(\alpha)}\right)^2 = f(\alpha).$$

Thus we have $P = \left(\alpha, \frac{g_{\boldsymbol{a}_0}(\alpha)}{h_{\boldsymbol{b}_0}(\alpha)}\right)$ is an algebraic point on C and $\mathbb{Q}(P)$ is precisely the degree n field $\mathbb{Q}(\alpha)$.

Given $f(x)$ and a degree n, our goal is now to describe a polynomial family $F_f(\boldsymbol{a}, \boldsymbol{b}, x)$, and use the methods of the previous section to prove that it is irreducible over $\mathbb{Q}(\boldsymbol{a}, \boldsymbol{b})$ with Galois group $G \simeq S_n$. This will give us a means of producing many degree n number fields which are generated by algebraic points of C, which we can count later.

3.1. Odd degree curves. Fix f with odd degree $d \geq 3$. Fix a degree $n \geq d$. We take the degrees d_g and d_h to be as large as possible so (3.1) has degree n,

$$d_g = \begin{cases} (n-1)/2, & n \text{ odd}, \\ n/2, & n \text{ even}, \end{cases}$$

$$d_h = \begin{cases} (n-d)/2, & n \text{ odd}, \\ (n-d-1)/2, & n \text{ even}. \end{cases}$$

For simplicity, we denote the polynomial family (3.1) by $F(x) \in \mathbb{Q}(\boldsymbol{a}, \boldsymbol{b})[x]$ and a specialization by $F_{\boldsymbol{a}_0,\boldsymbol{b}_0}(x) \in \mathbb{Q}[x]$, leaving both f and n implicit when it will not create confusion.

Proposition 3.1. *Fix a polynomial f and integers n, d_g, d_h as above. Then F_f is irreducible in $\mathbb{Q}(\boldsymbol{a}, \boldsymbol{b})[x]$ and* $\mathrm{Gal}(F_f/\mathbb{Q}(\boldsymbol{a}, \boldsymbol{b})) \simeq S_n$.

Proof. The irreducibility and Galois group of $F_f(x)$ over $\mathbb{Q}(\boldsymbol{a}, \boldsymbol{b})$ are invariant under a linear change of variables in x. It will be convenient to assume that the constant term of f, c_0, is nonzero, which is always possible after a linear change of variables. We treat the cases of n even and odd separately.

Case 1: n is even. When n is even, we take $d_g = n/2$ and $d_h = (n-d-1)/2$. Let p be a prime that does not divide any nonzero coefficient of f. Consider an integral specialization $\boldsymbol{a}_0 = (a_0, \ldots, a_{n/2})$ and $\boldsymbol{b}_0 = (b_0, \ldots, b_{(n-d-1)/2})$ with the following p-adic valuations:

$$\begin{aligned} v_p(a_0) &= 1 \\ v_p(a_i) &\geq 1 \text{ for } 0 < i < n/2 \\ v_p(a_{n/2}) &= 0 \\ v_p(b_j) &\geq 2 \text{ for } 0 \leq j \leq (n-d-1)/2. \end{aligned} \tag{3.2}$$

These requirements on the valuations of b_j allow us to effectively ignore the $h_{\boldsymbol{b}_0}(x)^2 f(x)$ term of $F_{\boldsymbol{a}_0,\boldsymbol{b}_0}$ in constructing the Newton polygon. Inspecting the valuations of the coefficients of $g_{\boldsymbol{a}_0}(x)^2$ gives the resulting $\mathbb{Q}_p$-adic Newton polygon for $F_{\boldsymbol{a}_0,\boldsymbol{b}_0}$, shown in Figure 3.1.

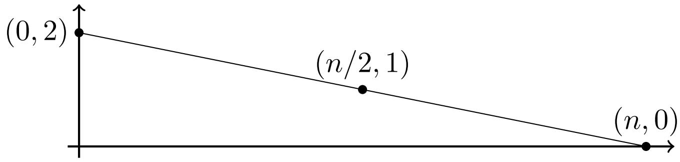

FIGURE 3.1. $\mathrm{NP}_{\mathbb{Q}_p}(F_{\boldsymbol{a}_0,\boldsymbol{b}_0})$ with one segment of slope $-2/n$

The Newton polygon $\mathrm{NP}_{\mathbb{Q}_p}(F_{\boldsymbol{a}_0,\boldsymbol{b}_0})$ has one segment of slope $-2/n$, so by Lemma 2.6, if $F_{\boldsymbol{a}_0,\boldsymbol{b}_0}$ is reducible over $\mathbb{Q}_p$ then it is the product of two degree $n/2$ irreducible factors. In particular, if F is reducible over $\mathbb{Q}(\boldsymbol{a},\boldsymbol{b})$, it must also be the product of two degree $n/2$ irreducible factors, as any other factorization would yield an incompatible factorization upon specializing by $\boldsymbol{a}_0,\boldsymbol{b}_0$ with the valuations given in (3.2).

Let us now consider a different integral specialization $\boldsymbol{a}_0,\boldsymbol{b}_0$ with the following p-adic valuations:

$$\begin{aligned} v_p(a_0) &= 0 \\ v_p(a_i) &\geq 2 \text{ for } 0 < i \leq n/2 \\ v_p(b_j) &\geq 2 \text{ for } 0 \leq j < (n-d-1)/2 \\ v_p(b_{(n-d-1)/2}) &= 1. \end{aligned} \tag{3.3}$$

The constant term of $F_{\boldsymbol{a}_0,\boldsymbol{b}_0}$ is $a_0^2 - b_0^2 c_0$ which has valuation 0. All other coefficients can be seen to have valuation at least 2, with the leading coefficient having valuation at least 4. The coefficient of x^{n-1} is given by $2a_{n/2-1}a_{n/2} - b_{(n-d-1)/2}^2 c_d$, which has valuation exactly 2. The resulting Newton polygon is shown below in Figure 3.2.

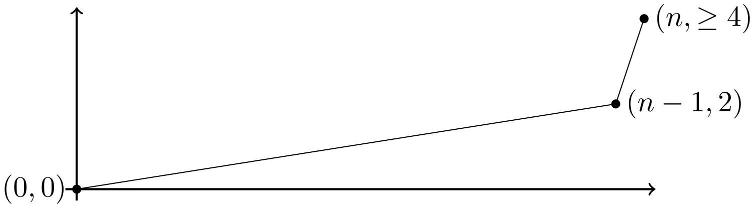

FIGURE 3.2. $\mathrm{NP}_{\mathbb{Q}_p}(F_{\boldsymbol{a}_0,\boldsymbol{b}_0})$ with $(n-1)$-cycle

This Newton polygon has a segment of length $n-1$ and slope equal to $2/(n-1)$, so by Lemma 2.6 whenever $\boldsymbol{a}_0,\boldsymbol{b}_0$ have the p-adic valuations given in (3.3), we have that $F_{\boldsymbol{a}_0,\boldsymbol{b}_0}$ factors as a degree $n-1$ irreducible polynomial times a linear polynomial over $\mathbb{Q}_p$. Such a factorization cannot occur if F has two irreducible degree $n/2$ factors over $\mathbb{Q}(\boldsymbol{a}, \boldsymbol{b})$, so we may conclude that F is irreducible, and hence G is a transitive permutation subgroup of S_n. Moreover, Lemma 2.7 implies that the Galois group of $F_{\boldsymbol{a}_0,\boldsymbol{b}_0}$ over $\mathbb{Q}$ contains a cycle of length $n-1$ whenever $\boldsymbol{a}_0$ and $\boldsymbol{b}_0$ satisfy the valuations in (3.3). These valuation criteria are satisfied for a positive proportion of integral specializations $\boldsymbol{a}_0$ and $\boldsymbol{b}_0$, so Corollary 2.2 implies that G contains an $(n-1)$-cycle.

To produce a transposition in G, we first argue that the set of primes p such that $f(k) \equiv 0 \pmod{p}$ for some $k \in \mathbb{Z}$, is infinite. This fact has an elementary proof, but is also seen to be a consequence of the Chebotarev density theorem (see e.g. [10, §13]), as this set contains the positive density set of primes which split completely in the splitting field of $f(x)$. In any case, we may fix a prime $p > n$ with $p \nmid \operatorname{Disc} f, c_d$ such that $p \mid f(k)$ for some integer k, which implies that p divides the constant coefficient of the translation $f(x+k)$, but not the linear term, as $p \nmid \operatorname{Disc} f$ implies that the reduction of $f(x)$ mod p is also squarefree. Using a Hensel's lemma lifting argument, we can further find an integral solution to $f(x) \equiv 0 \pmod{p}$ such that $f(x) \not\equiv 0 \pmod{p^2}$. Thus after possibly another change of variables, we may assume that $v_p(c_0) = 1$ and $v_p(c_1) = 0$.

We consider an integral specialization $\boldsymbol{a}_0, \boldsymbol{b}_0$ with the following p-adic valuations:

$$\begin{aligned} v_p(a_0) &= 2 \\ v_p(a_1) &= 0 \\ v_p(a_i) &\geq 2 \text{ for } 1 < i < n/2 \\ v_p(a_{n/2}) &= 3 \\ v_p(b_0) = v_p(b_{(n-d-1)/2}) &= 1 \\ v_p(b_j) &\geq 1 \text{ for } 0 < j < (n-d-1)/2. \end{aligned} \tag{3.4}$$

These requirements ensure that the constant term of $F_{\boldsymbol{a}_0,\boldsymbol{b}_0}$ has valuation exactly 3, the coefficient of x^2 has valuation exactly 0, the x^{n-1} coefficient $2a_{n/2}a_{n/2-1} - b^2_{(n-d-1)/2}c_d$ has valuation exactly 2, and the leading term has valuation exactly 6, with all other coefficients having valuation at least 2. The resulting Newton polygon is shown below in Figure 3.3.

That $n > d \geq 3$ ensures that $\frac{2}{n-3} < 4$, so the two rightmost segments are distinct. These, together with the segment of length 2 and slope $-3/2$ above, ensure that $F_{\boldsymbol{a}_0,\boldsymbol{b}_0}$ has factors of degree 2, $n-3$, and 1 over $\mathbb{Q}_p$, so Lemma 2.7 applies to reveal a transposition in $G_{\boldsymbol{a}_0,\boldsymbol{b}_0}$.

Since a positive proportion of integer tuples $\boldsymbol{a}_0, \boldsymbol{b}_0$ satisfy (3.4), Corollary 2.2 implies that G also contains a transposition. Thus G satisfies the hypotheses of Proposition 2.4 and we conclude that $G \simeq S_n$.

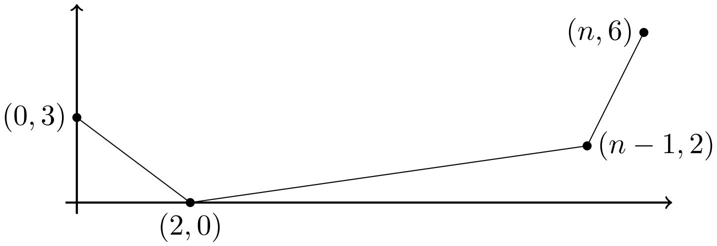

Figure 3.3. $\mathrm{NP}_{\mathbb{Q}_p}(F_{\boldsymbol{a}_0,\boldsymbol{b}_0})$ with transposition

Case 2: n is odd. Now we take $d_g = (n-1)/2$ and $d_h = (n-d)/2$. Fix a prime p not dividing any nonzero coefficient of f. Consider an integral specialization $\boldsymbol{a}_0 = (a_0, \dots, a_{(n-1)/2})$ and $\boldsymbol{b}_0 = (b_0, \dots b_{(n-d)/2})$ with the following p-adic valuations:

$$\begin{aligned} v_p(a_0) &= 0 \\ v_p(a_i) &\geq 2 \text{ for } i > 0 \\ v_p(b_j) &\geq 2 \text{ for } j < (n-d)/2 \\ v_p(b_{(n-d)/2}) &= 1. \end{aligned} \tag{3.5}$$

These requirements ensure that the constant term $a_0^2 - b_0^2 c_0$ has valuation exactly 0, the leading coefficient $b_{(n-d)/2}^2 c_d$ has valuation exactly 2, and all intermediate coefficients have valuation at least 2. This produces the p-adic Newton polygon for $F_{\boldsymbol{a}_0,\boldsymbol{b}_0}$ shown below in Figure 3.4.

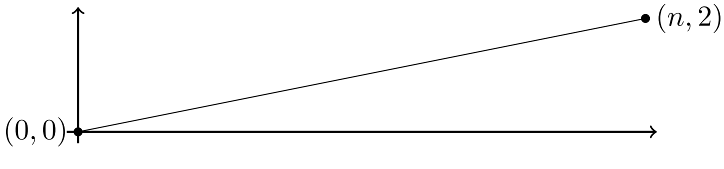

Figure 3.4. $\mathrm{NP}_{\mathbb{Q}_p}(F_{\boldsymbol{a}_0,\boldsymbol{b}_0})$ with n-cycle

This Newton polygon has one segment of slope $2/n$, and since n is odd we have $\gcd(2,n) = 1$. Thus Lemma 2.6 implies that the specialization $F_{\boldsymbol{a}_0,\boldsymbol{b}_0}$ is irreducible over $\mathbb{Q}_p$, hence over $\mathbb{Q}$, and we have that F must be irreducible over $\mathbb{Q}(\boldsymbol{a}, \boldsymbol{b})$, with its Galois group G a transitive subgroup of S_n.

Next, we aim to produce a q-cycle in G for a prime $q > n/2$. We will assume $n > 3$ for now, as the case of $n = d = 3$ will be handled by later arguments. Recalling Bertrand's postulate, there exists some prime q such that $\frac{n-1}{2} < q < n-1$, which is odd and satisfies $q > n/2$. Consider now a specialization $\boldsymbol{a}_0, \boldsymbol{b}_0$ satisfying

$$\begin{aligned} v_p(a_{(n-q)/2}) &= 0 \\ v_p(a_i) &\geq 2 \text{ for } i \neq (n-q)/2 \\ v_p(b_j) &\geq 2 \text{ for } j < (n-d)/2 \\ v_p(b_{(n-d)/2}) &= 1. \end{aligned} \tag{3.6}$$

These requirements ensure that the valuations of all coefficients of $F_{\boldsymbol{a}_0,\boldsymbol{b}_0}$ are at least 2, except for the degree $n-q$ term, whose coefficient has valuation zero coming from the presence of an $a_{(n-q)/2}^2$ term. The leading coefficient $b_{(n-d)/2}^2 c_d$ has valuation exactly 2. An example p-adic Newton polygon for such a specialization $F_{\boldsymbol{a}_0,\boldsymbol{b}_0}$ is shown below in Figure 3.5.

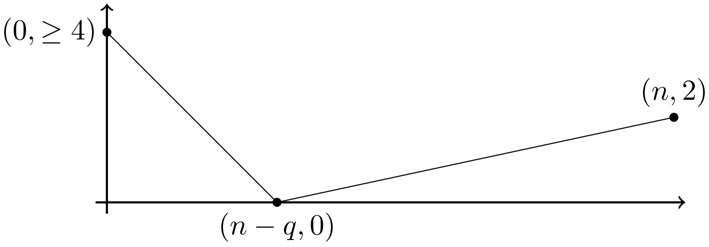

Figure 3.5. $\mathrm{NP}_{\mathbb{Q}_p}(F_{\boldsymbol{a}_0,\boldsymbol{b}_0})$ with q-cycle

Note that the left side of the Newton polygon in Figure 3.5 need not be a single segment. This is inconsequential however, because the right side is of interest to us, in particular the segment of slope $2/q$ and length q. Since $q > n/2$ is an odd prime, we have $\gcd(2, q) = 1$ and q is coprime to any integers less than or equal to $n - q$, so Lemma 2.7 applies, ensuring the existence of a q-cycle in $G_{\boldsymbol{a}_0,\boldsymbol{b}_0}$. Since a positive proportion of integral specializations satisfy (3.6), Corollary 2.2 implies that G contains a q-cycle as well.

Finally, we can produce a transposition in G using essentially the same argument as in the case of even n. After a possible change of variables, let $p > n$ be a prime such that $v_p(c_0) = 1$ and $p \nmid \operatorname{Disc} f, c_d$. We consider specializations with the following p-adic valuations.

$$\begin{aligned} v_p(a_0) &= 2 \\ v_p(a_1) &= 0 \\ v_p(a_i) &\geq 2 \text{ for } 1 < i \leq (n-1)/2 \\ v_p(b_0) = v_p(b_{(n-d)/2}) &= 1 \\ v_p(b_j) &\geq 1 \text{ for } 0 < j < (n-d)/2. \end{aligned} \tag{3.7}$$

These conditions produce the Newton polygon shown below in Figure 3.6.

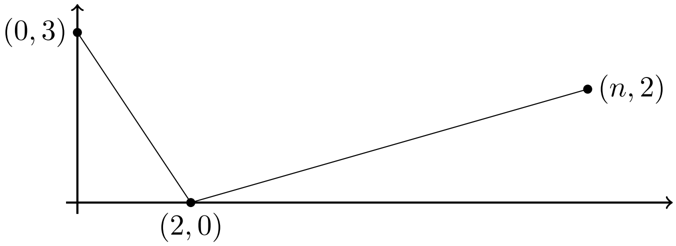

FIGURE 3.6. $\mathrm{NP}_{\mathbb{Q}_p}(F_{\boldsymbol{a}_0,\boldsymbol{b}_0})$ with transposition

Since $n-2$ is odd, Lemma 2.7 applied to the segment of slope $-3/2$ implies that $G_{\boldsymbol{a}_0,\boldsymbol{b}_0}$, and hence G by Corollary 2.2, contains a transposition. Therefore, G satisfies the hypotheses of Proposition 2.4, and we conclude $G \simeq S_n$. □

3.2. Even degree curves. We now present the analogous proposition for the case of d even. Let $f(x) \in \mathbb{Z}[x]$ be squarefree given by $f(x) = \sum_{i=0}^{d} c_i x^i$, with $d \geq 4$ even. Fix an even integer $n \geq d+2$ and take $d_g = n/2$ and $d_h = (n-d)/2 - 1$. Let $F_f(\boldsymbol{a}, \boldsymbol{b}, x) \in \mathbb{Q}(\boldsymbol{a}, \boldsymbol{b})[x]$ denote the polynomial family in (3.1), which is seen to have degree n. Again, for simplicity we denote this by $F(x)$ when it will not create confusion.

Proposition 3.2. *Fix a polynomial f, an even integer n, and degrees d_g, d_h as above. Then F_f is irreducible in $\mathbb{Q}(\boldsymbol{a}, \boldsymbol{b})[x]$ and $\mathrm{Gal}(F_f/\mathbb{Q}(\boldsymbol{a}, \boldsymbol{b})) \simeq S_n$.*

Proof. We will again need that the irreducibility of F_f and its Galois group G are invariant under linear change of coordinates in x, to allow us to assume certain conditions on the valuations of the c_i.

As in the proof of Proposition 3.1, there exists a prime $p > n$ not dividing both $\operatorname{Disc} f, c_d$ such that p divides $f(k)$ exactly once for some integer k. Thus after changing variables, we assume that $v_p(c_0) = 1$.

Consider now the change of variables by scaling x to be px. The constant term c_0 remains unchanged, but this allows us to assume that $p \mid c_i$ for $i \geq 1$. These assumptions are useful for finding long cycles in $G = \mathrm{Gal}(F_f/\mathbb{Q}(\boldsymbol{a}, \boldsymbol{b}))$. We consider an integral specialization $\boldsymbol{a}_0, \boldsymbol{b}_0$ with the following p-adic valuations:

$$\begin{aligned} v_p(a_i) &\geq 1 \text{ for } i < n/2 \\ v_p(a_{n/2}) &= 0 \\ v_p(b_0) &= 0, \end{aligned} \tag{3.8}$$

and no restrictions on b_j for $j > 0$. These restrictions, and assumptions on the coefficients c_i, ensure that every term of $F(x)$ is divisible by p, except for the leading coefficient $a_{n/2}^2$, which has valuation 0. Moreover, the valuation of the constant term $a_0^2 - b_0^2 c_0$ is exactly 1, so the Newton polygon of $F_{\boldsymbol{a}_0,\boldsymbol{b}_0}$ has exactly one segment of length n and slope $-1/n$, as shown in Figure 3.7.

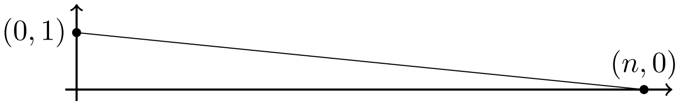

Figure 3.7. $\mathrm{NP}_{\mathbb{Q}_p}(F_{\boldsymbol{a}_0,\boldsymbol{b}_0})$ with n-cycle

Lemma 2.7 implies that $F_{\boldsymbol{a}_0,\boldsymbol{b}_0}$ is irreducible over $\mathbb{Q}_p$, and hence over $\mathbb{Q}$, so F is irreducible over $\mathbb{Q}(\boldsymbol{a},\boldsymbol{b})$ and G is transitive, containing an n-cycle by Corollary 2.2.

We use a variation of this argument to find an $(n-1)$-cycle in G. Fix another prime $p > n$ such that after a change of variables we have $v_p(c_0) = 1$ and $p \nmid c_1$. We consider an integral specialization $\boldsymbol{a}_0, \boldsymbol{b}_0$ with the following p-adic valuations:

$$\begin{aligned} v_p(a_i) &\geq 3 \text{ for } i < n/2 \\ v_p(a_{n/2}) &= 0 \\ v_p(b_0) &= 1 \\ v_p(b_j) &\geq 2 \text{ for } j > 0. \end{aligned} \tag{3.9}$$

These restrictions ensure that the constant term has valuation 3, while the linear coefficient, $2a_0a_1 - b_0^2c_1 - 2b_0b_1c_0$, has valuation exactly 2. All other terms have valuation at least 2 except for the leading term, which has valuation 0. This produces the Newton polygon below in Figure 3.8.

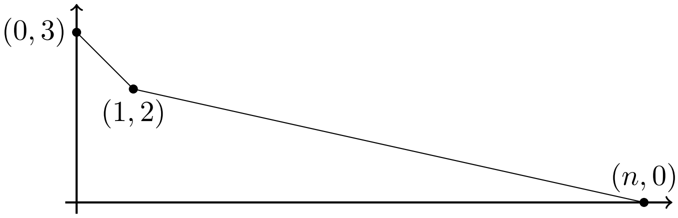

Figure 3.8. $\mathrm{NP}_{\mathbb{Q}_p}(F_{\boldsymbol{a}_0,\boldsymbol{b}_0})$ with $(n-1)$-cycle

Since $n \geq 4$, the two segments are distinct, with the rightmost one of length $n-1$ and slope $-2/(n-1)$. As n is even, Lemma 2.7 is satisfied, producing an $(n-1)$-cycle in $G_{\boldsymbol{a}_0,\boldsymbol{b}_0}$ and thus in G.

Finally, we produce a transposition in G, assuming that $n \geq 8$ for simplicity; nearly identical arguments suffice for the case of $d = 4$ and $n = 6$. More care is needed here to find a Newton polygon with exactly one segment of even length to satisfy the hypotheses of Lemma 2.7.

Fix a prime $p > n$ such that $p \nmid c_d, \operatorname{Disc} f$, c_d is a quadratic residue modulo p, and $p \mid f(k)$ for some integer k. Such a prime exists by our earlier Chebotarev argument, this time looking for primes splitting completely in the splitting field of $f(x)(x^2 - c_d)$. After a change of coordinates, we assume $v_p(c_0) = 1$ and $p \nmid c_1$. We consider an integral specialization $\boldsymbol{a}_0, \boldsymbol{b}_0$ with the following restrictions:

$$\begin{aligned} v_p(a_i) &\geq 4 \text{ for } i < \frac{n}{2} - 2 \\ v_p(a_{n/2-2}) &= 0 \\ v_p(a_{n/2-1}) &= 1 \\ v_p(a_{n/2}) &= 1 \text{ such that } \frac{a_{n/2}^2}{p^2} \equiv c_d \pmod{p^2} \\ v_p(b_0) &= 1 \\ v_p(b_j) &\geq 1 \\ v_p(b_{(n-d)/2}) &= 1 \text{ such that } \frac{b_{(n-d)/2}^2}{p^2} \equiv 1 \pmod{p^2}. \end{aligned} \tag{3.10}$$

Note that such $a_{n/2}$ exists, since c_d is a quadratic residue, and these assumptions ensure that $p^4 \mid a_{n/2}^2 - b_{(n-d)/2}^2 c_d$, the leading coefficient. Furthermore, we have that the constant coefficient has valuation 3, the linear coefficient has valuation 2, the x^{n-4} coefficient has valuation 0, and both the x^{n-3} and x^{n-2} coefficients have valuation 1, with all other terms having valuation at least 2.

Looking more closely at the coefficient of x^{n-1} given by

$$2a_{n/2-1}a_{n/2} - b_{(n-d)/2-1}b_{(n-d)/2}c_d - b_{(n-d)/2}^2 c_{d-1},$$

we see that its valuation at least 2. To ensure it has valuation exactly 2, we fix a residue class for $\frac{1}{p} b_{(n-d)/2-1}$ modulo p and ask that $a_{n/2-1}$ satisfy

$$\frac{a_{n/2-1}}{p} \not\equiv \left(2\frac{a_{n/2}}{p}\right)^{-1} \frac{1}{p^2}\left(b_{(n-d)/2-1}b_{(n-d)/2}c_d - b_{(n-d)/2}^2 c_{d-1}\right) \pmod{p}. \tag{3.11}$$

Thus combining (3.10) and (3.11), we produce the Newton polygon in Figure 3.9 below.

The segment of length 2 and slope 1/2, together with the fact that all other segments have odd length l' and slopes r'/l' with $\gcd(r', l') = 1$, allow us to apply Lemma 2.7 with $l = 2$ to produce a transposition in $G_{\boldsymbol{a}_0, \boldsymbol{b}_0}$. The requirements (3.10) and (3.11) are satisfied for a positive proportion of integral $\boldsymbol{a}_0, \boldsymbol{b}_0$, so Corollary 2.2 implies that G contains a transposition. Thus with its n-cycle, $(n-1)$-cycle, and transposition, Proposition 2.4 gives that $G \simeq S_n$. □

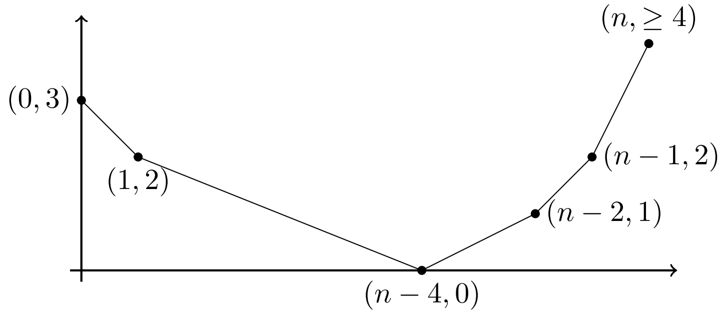

FIGURE 3.9. $\mathrm{NP}_{\mathbb{Q}_p}(F_{\boldsymbol{a}_0,\boldsymbol{b}_0})$ with transposition

4. Relating coefficients to roots

In this brief section we state a result which relates the absolute value of a polynomial's coefficients to that of its roots, which will be useful later when counting multiplicities of fields generated by a family of polynomials. To avoid confusion, we note that for the purposes of this section $f(x)$ denotes a general polynomial in $\mathbb{C}[x]$, rather than squarefree integral polynomial defining a nonsingular hyperelliptic curve, as in the previous section.

Lemma 4.1. *Let $f(x) = \sum_{i=0}^{n} c_i x^i \in \mathbb{C}[x]$ be monic and have degree n. There exist positive constants k_i such that for any $Y > 0$, if $|c_i| \leq k_i Y^{n-i}$ for $0 \leq i \leq n$ then $|\alpha| \leq Y$ for all roots α of $f(x)$.*

Proof. This result follows from classical upper bounds on the absolute value of complex roots given by Lagrange and Cauchy. A clean proof yielding explicit values of k_i follows from the following bound due to Fujiwara [4],

$$|\alpha| \leq 2\max\left\{\left|\frac{c_{n-1}}{c_n}\right|, \left|\frac{c_{n-2}}{c_n}\right|^{1/2}, \dots, \left|\frac{c_1}{c_n}\right|^{1/(n-1)}, \left|\frac{c_0}{2c_n}\right|^{1/n}\right\}. \tag{4.1}$$

Set $k_0 = \frac{1}{2^{n-1}}$ and $k_i = \frac{1}{2^{n-i}}$ for $1 \leq i \leq n$. In our case we have $c_n = 1$, so if $|c_0| \leq \frac{1}{2^{n-1}} Y^n$, then $|\frac{c_0}{2}| \leq \frac{1}{2^n} Y^n$. Taking n-th roots, we have $|\frac{c_0}{2}|^{1/n} \leq \frac{Y}{2}$. Similarly, for $1 \leq i \leq n-1$, we have $|c_i| \leq \left(\frac{Y}{2}\right)^{n-i}$, so taking $(n-i)$-th roots implies $|c_i|^{1/(n-i)} \leq \frac{Y}{2}$. Thus

$$\max\left\{|c_{n-1}|, |c_{n-2}|^{1/2}, \dots, |c_1|^{1/(n-1)}, \left|\frac{c_0}{2}\right|^{1/n}\right\} \leq \frac{Y}{2},$$

so applying (4.1) gives $|\alpha| \leq Y$ for any root α of $f(x)$. □

5. Proofs of main theorems

We begin with the proof of the first bound in Theorem 1.1, which covers Sections 5.1–5.3. In Section 5.4 we describe the modifications necessary to obtain the improved bound in Theorem 1.1 for sufficiently large n. The proof of Theorem 1.2 is nearly identical, and we highlight the differences in Section 5.5.

5.1. Parameterization. Let C be a nonsingular hyperelliptic curve over $\mathbb{Q}$ of odd degree $d = 2g + 1$. Then C has a model

$$C\colon y^2 = f(x) = \sum_{i=0}^{d} c_i x^i$$

where $c_i \in \mathbb{Z}$ for all i and $f(x)$ is squarefree. We may further assume that $c_0 \neq 0$ by translating x if needed. If necessary, we may also take f to be monic, by multiplying by c_d^{d-1} and changing variables again.

Let Y be a positive real number and $n \geq d$ an integer. We now construct a family of polynomials $P_{f,n}(Y)$ arising from certain specializations of (3.1). When n is even, take

$$\begin{aligned} g(x) &= x^{n/2} + a_{n/2-1}x^{n/2-1} + \cdots + a_0 \\ h(x) &= b_{(n-d-1)/2}x^{(n-d-1)/2} + b_{(n-d-1)/2-1}x^{(n-d-1)/2-1} + \cdots + b_0, \end{aligned} \tag{5.1}$$

with the restrictions that $a_i, b_j \in \mathbb{Z}$, $|a_{n/2-i}| \leq Y^i$, and $|b_{(n-d-1)/2-j}| \leq Y^{j+1/2}$. In the case of n odd we take

$$\begin{aligned} g(x) &= a_{(n-1)/2}x^{(n-1)/2} + a_{(n-1)/2-1}x^{(n-1)/2-1} + \cdots + a_0 \\ h(x) &= x^{(n-d)/2} + b_{(n-d)/2-1}x^{(n-d)/2-1} + \cdots + b_0, \end{aligned} \tag{5.2}$$

with $|a_{(n-1)/2-i}| \leq Y^{i+1/2}$ and $|b_{(n-d)/2-j}| \leq Y^j$.

Let $P_{f,n}(Y)$ be the set of polynomials $F(x) = g(x)^2 - f(x)h(x)^2$ for $g(x), h(x)$ of the form above. Note that any such $F(x)$ has degree n, and by Lemma 4.1 any root α of F satisfies $|\alpha| \ll_{n,f} Y$. Hence $\mathrm{Disc}(F) \leq kY^{n(n-1)}$ for a constant k depending on f and n.

5.2. Bounding multiplicities. We can count the number fields arising from specializations of (3.1) by counting elements of $P_{f,n}(Y)$, provided that we can control the multiplicity. This multiplicity arises from two sources. We may have different choices of $g(x)$ and $h(x)$ that produce the same element $F(x) \in P_{f,n}(Y)$, or we may find multiple elements of $P_{f,n}(Y)$ that produce isomorphic number fields. We deal with the former case in the following lemma.

Lemma 5.1. *Let $F(x) \in \mathbb{Z}[x]$ be a polynomial of degree n. The number of ways to choose $g(x), h(x) \in \mathbb{Z}[x]$ with at least one monic so that $F(x) = g(x)^2 - f(x)h(x)^2$ is $O_n(1)$.*

Proof. Note that f has no repeated roots, so the complex affine coordinate ring, given by $\mathbb{C}[x,y]/(y^2-f(x))$, is a Dedekind domain. With this, one follows the justification of [8, Lemma 7.4] to argue that (F) factors uniquely into a product of $2n$ prime ideals.

Given any such g, h with $F = g^2 - fh^2$, we associate the factorization $F = (g-\sqrt{f}h)(g+\sqrt{f}h)$. Thus the ideal $(g-\sqrt{f}h)$ factors as a product of the $2n$ primes dividing (F), giving at most 2^{2n} possibilities for the ideal $(g-\sqrt{f}h)$.

Since $\deg f$ is odd, the units in the coordinate ring consist of the constants, as an element $u+\sqrt{f}v$ has norm u^2-fv^2, which is a unit in $\mathbb{C}[x]$ if and only if $v=0$ and u is a nonzero constant. Thus the ideal $(g-\sqrt{f}h)$ determines g and h exactly by the monicity assumption. □

Now we can give a count for $\#P_{f,n}(Y)$, since Lemma 5.1 gives that each choice of a_i and b_j above coincides with at most a constant number of other choices. In the case of n even, we have $\#P_{f,n}(Y) \asymp Y^c$ where

$$\begin{aligned} c = \sum_{i=1}^{n/2} i + \sum_{j=0}^{(n-d-1)/2} (j+1/2) &= \frac{1}{4}\left(n^2 + (2-d)n + \frac{d^2-2d+1}{2}\right) \\ &= \frac{1}{4}\left(n^2 + (1-2g)n + 2g^2\right). \end{aligned} \tag{5.3}$$

The same approach yields the same count for n odd. Since the elements of $P_{f,n}(Y)$ arise as specializations of the family (3.1), Proposition 3.1 implies that $\#P_{f,n}(Y,S_n) \asymp Y^c$ where c is given in (5.3) and $P_{f,n}(Y,S_n)$ is the subset consisting of irreducible $F \in P_{f,n}(Y)$ with $\mathrm{Gal}(F/\mathbb{Q}) \simeq S_n$.

We now address the second source of multiplicity, namely that there may be many $F \in P_{f,n}(Y)$ for which $K \simeq \mathbb{Q}[x]/F(x)$. To deal with this, we employ machinery developed by Ellenberg and Venkatesh [3] for counting number fields, and the multiplicity counts of Lemke Oliver and Thorne [8].

Following their lead we define

$$S(Y) := \left\{F = x^n + c'_{n-1}x^{n-1} + \cdots + c'_0 \in \mathbb{Z}[x] \;\middle|\; |c'_{n-i}| \ll_{n,f} Y^i\right\}$$

with the condition that $F(x)$ is irreducible. Note that by this construction $P_{f,n}(Y,S_n) \subseteq S(Y,S_n)$, provided we choose the implied constant appropriately. We now define for a number field K its multiplicity within $S(Y,S_n)$,

$$M_K(Y) := \#\left\{F \in S(Y) \mid \mathbb{Q}[x]/F(x) \simeq K\right\}.$$

Lemma 5.2 (Lemke Oliver–Thorne, [8, Proposition 7.5])**.** *We have*

$$M_K(Y) \ll \max\left(Y^n \left|\mathrm{Disc}(K)\right|^{-1/2}, Y^{n/2}\right).$$

The proof uses the geometry of numbers, building on the strategy suggested in [3].

We now state an upper bound for the asymptotics of general number field counts, without respect to any curve. We use $N_n(X)$ to denote the number of degree n number fields K with $|\mathrm{Disc}(K)| \leq X$.

Theorem 5.3 (Schmidt, [11]). *For $n \geq 3$, we have*

$$N_n(X) \ll X^{\frac{n+2}{4}}. \tag{5.4}$$

We leverage Theorem 5.3 to show that the contribution to $N_{n,C}(X, S_n)$ by fields of low discriminant is negligible. This allows for some improvement to the eventual exponent.

Lemma 5.4. *Let $T \leq Y^n$. Then*

$$\sum_{|\mathrm{Disc}(K)| \leq T} M_K(Y) \ll Y^n T^{n/4},$$

where the sum runs over all degree n number fields K with $|\mathrm{Disc}(K)| \leq T$.

Proof. We begin by rewriting the sum as a Riemann–Stieljes integral,

$$\begin{aligned}\sum_{|\mathrm{Disc}(K)| \leq T} M_K(Y) &= \sum_{1 \leq t \leq T} \left(N_n(t) - N_n(t-1)\right) M_K(Y)(t) \\ &= \int_{1^-}^{T} M_K(Y)(t) \mathrm{d}N_n(t) \qquad (5.5) \\ &\ll Y^n \int_{1^-}^{T} \frac{\mathrm{d}N_n(t)}{t^{1/2}}, \qquad (5.6)\end{aligned}$$

where (5.6) follows from (5.5) by the multiplicity bound from Lemma 5.2. Integrating by parts in (5.6) produces

$$Y^n \int_{1^-}^{T} \frac{\mathrm{d}N_n(t)}{t^{1/2}} = Y^n \frac{N_n(T)}{T^{1/2}} + \frac{Y^n}{2} \int_{1-}^{T} \frac{N_n(t)}{t^{3/2}} \mathrm{d}t. \tag{5.7}$$

Recalling Schmidt's bound in (5.4), we estimate (5.7) by

$$\begin{aligned}Y^n \frac{N_n(T)}{T^{1/2}} + \frac{Y^n}{2} \int_{1^-}^{T} \frac{N_n(t)}{t^{3/2}} \mathrm{d}t &\ll Y^n T^{n/4} + \frac{Y^n}{2} \int_{1}^{T} t^{\frac{n}{4}-1} \mathrm{d}t \\ &= Y^n T^{n/4} + \frac{2Y^n}{n}(T^{n/4} - 1) \\ &= Y^n \left(\left(1 + \frac{2}{n}\right) T^{n/4} - \frac{2}{n} \right) \\ &\ll Y^n T^{n/4}. \qquad \square\end{aligned}$$

5.3. Final steps. We are now ready to assemble the proof of Theorem 1.1. By our construction, for any $F \in P_{f,n}(Y, S_n)$ and any root α of F, we have $\left(\alpha, \frac{g(\alpha)}{h(\alpha)}\right) \in C(K)$ where $K = \mathbb{Q}(\alpha)$ is a field of degree n with $\mathrm{Gal}(\widetilde{K}/\mathbb{Q}) \simeq S_n$. We then have that $|\mathrm{Disc}(K)| \leq kY^{n(n-1)}$ for a constant k depending on f, n.

Choose $T = \kappa Y^{n-(3+2g)+2g^2/n}$ for a positive constant κ to be determined shortly. By Lemma 5.4, we have

$$\sum_{|\mathrm{Disc}(K)|\leq T} M_K(Y) \ll \kappa^{n/4} Y^c, \tag{5.8}$$

and we recall from our earlier discussion that

$$\#P_{f,n}(Y, S_n) \asymp Y^c. \tag{5.9}$$

We then choose κ sufficiently small so that the quantity in (5.8) is at most $\#P_{f,n}(Y, S_n)/2$. Then, fields K with $T < |\mathrm{Disc}(K)| \leq kY^{n(n-1)}$ arise from a positive proportion of the polynomials in $P_{f,n}(Y, S_n)$. Counting just these fields and recognizing the bound for $M_K(Y)$ in Lemma 5.2 is decreasing with respect to $|\mathrm{Disc}(K)|$, we have $M_K(Y) \ll T^{-1/2}Y^n$ for all K with $T < |\mathrm{Disc}(K)| \leq kY^{n(n-1)}$. Thus we have

$$N_{n,C}(kY^{n(n-1)}, S_n) \gg Y^{c-n}T^{1/2} = Y^{\frac{1}{4}\left(n^2-(1+2g)n+2g^2-4g-6+4g^2/n\right)}. \tag{5.10}$$

Upon replacing Y in (5.10) by $(X/k)^{1/n(n-1)}$ and simplifying, we obtain as the exponent

$$c_n = \frac{1}{4} - \frac{gn^2 - (g^2-2g-3)n - 2g^2}{2n^2(n-1)}$$

and thus $N_{n,C}(X, S_n) \gg X^{c_n}$, which is the first statement of Theorem 1.1.

5.4. Improvements. To improve the exponent in the previous section, we seek to find when fields of discriminant less than Y^n contribute negligibly, allowing us to use the best possible multiplicity bound in Lemma 5.2, $M_K(Y) \ll Y^{n/2}$. If we assume this is true for some n, then we immediately have

$$N_{n,C}(kY^{n(n-1)}, S_n) \gg Y^{c-n/2},$$

and after simplifying and making the same substitutions as earlier, we obtain the improvement in Theorem 1.1.

It now remains to argue that this is possible. Suppose that $N_n(X) \ll X^{\alpha(n,g)}$ is valid for large enough n. With this assumption, we use the same procedure as the proof of Lemma 5.4 to show that

$$\sum_{|\mathrm{Disc}\, K|\leq Y^n} M_K(Y) \ll Y^{n/2+n\alpha(n,g)} \tag{5.11}$$

To make the right hand side of (5.11) be $o(Y^c)$, it suffices to take any $\alpha(n,g)$ satisfying

$$\alpha(n,g) < \frac{n}{4} - \frac{1+2g}{4} + \frac{g^2}{2n}. \tag{5.12}$$

Theorem 5.3 is insufficient for this purpose. We turn to the improved upper bounds for counting number fields of fixed degree by discriminant due to Lemke Oliver and Thorne [7].

Theorem 5.5 (Lemke Oliver–Thorne, [7, Theorems 1.1, 1.2])**.** *For $n \geq 6$ we have*

$$N_n(X) \ll X^{1.564(\log n)^2}. \tag{5.13}$$

Moreover, for $n \geq 2$ we have the following.

(1) *Let m be the least integer for which $\binom{m+2}{2} \geq 2n+1$. Then $N_n(X) \ll X^{2m-\frac{m(m-1)(m+4)}{6n}}$.*

(2) *Let $3 \leq r \leq n$ and let m be an integer such that $\binom{m+r-1}{r-1} > rn$. Then $N_n(X) \ll X^{mr}$.*

By taking $\alpha(n,g) = 1.564(\log n)^2$, as in Theorem 5.5, we see that (5.12) is satisfied for n sufficiently large, since $(\log n)^2$ grows more slowly than the right hand side for any fixed g. This completes our justification of the improved exponent in Theorem 1.1.

For any fixed g, one can compute the n at which (5.12) takes effect for $\alpha(n,g) = 1.564(\log n)^2$. Then, one can use (1) and (2) of Theorem 5.5 to search by computer for the least n for which there exists $\alpha(n,g)$ such that $N_n(X) \ll X^{\alpha(n,g)}$ and (5.12) is satisfied, by checking all appropriate pairs of integers (m,r). When $g = 1$, the improved exponent is valid for all $n \geq 106$. When $g = 10$, this approach shows the improved exponent is valid for $n \geq 138$. For $g = 100$, this increases to $n \geq 324$.

Since Theorem 1.1 is only valid for degrees $n \geq d = 2g+1$, when g is sufficiently large, the improved exponent will be valid for all $n \geq d$. We computed this to be true for all $g \geq 238$.

5.5. Even degree curves. The proof of Theorem 1.2 follows the approach of the previous subsection. We begin with a hyperelliptic curve $C\colon y^2 = f(x)$ with $f(x) = \sum_{i=0}^{d} c_i x^i \in \mathbb{Z}[x]$ squarefree for $d \geq 4$ even.

For $Y > 0$ and an even integer $n \geq d+2$ we define a family $P_{f,n}(Y)$ by polynomials of the form $F(x) = g(x)^2 - f(x)h(x)^2$, where

$$\begin{aligned} g(x) &= x^{n/2} + a_{n/2-1}x^{n/2-1} + \cdots + a_0 \\ h(x) &= b_{(n-d)/2-1}x^{(n-d)/2-1} + b_{(n-d)/2-2}x^{(n-d)/2-2} + \cdots + b_0, \end{aligned}$$

satisfy $|a_{n/2-i}| \leq Y^i$ and $|b_{(n-d)/2-j}| \leq Y^j$. We address the possibility that multiple choices of $g(x)$ and $h(x)$ produce coinciding $F(x)$ in the following analogue to Lemma 5.1.

Lemma 5.6. *Let $F(x) \in \mathbb{Z}[x]$ be a polynomial of even degree n. The number of ways to choose $g(x) \in \mathbb{Z}[x]$ monic of degree $n/2$ and $h(x) \in \mathbb{Z}[x]$ of degree $(n-d)/2-1$ such that $F(x) = g(x)^2 - f(x)h(x)^2$ is $O_n(1)$.*

Proof. By the first paragraph of the proof of Lemma 5.1, we have that the (F) in the ring $\mathbb{C}[x,y]/(y^2 = f(x))$ factors uniquely into a product of $2n$

primes. Again, any such g, h give us a factorization $F = (g - \sqrt{f}h)(g + \sqrt{f}h)$ with at most 2^{2n} possibilities for the ideal $(g - \sqrt{f}h)$.

Unlike in Lemma 5.1, there may be nontrivial units in the coordinate ring. Suppose $F = g^2 - fh^2 = (g')^2 - f(h')^2$ and $(g - \sqrt{f}h) = (g' - \sqrt{f}h')$ for some g', h' of degrees $n/2, (n-d)/2 - 1$ respectively. Thus for some unit in the coordinate ring of the form $u + \sqrt{f}v$, i.e. we have

$$\begin{aligned} gu - fhv &= g' \\ hu - gv &= h'. \end{aligned}$$

The above implies that

$$g'h - gh' = h(gu - fhv) - g(hu - gv) = v(g^2 - fh^2) = vF.$$

However, the degree of the left hand side is at most $\deg g + \deg h < n$, while the degree of the right hand side is at least n if v is nonzero, a contradiction. Therefore, only nonzero constants preserve both the ideal $(g - \sqrt{f}h)$ and the desired degrees of g and h, so the monicity assumption ensures that the ideal determines g and h precisely. □

By Lemma 5.6 and the same argument as for the odd degree case, we have $\#P_{f,n}(Y) \asymp Y^c$ for

$$c = \frac{1}{4}\left(n^2 - 2gn + 2g^2 + 2g\right). \tag{5.14}$$

Proposition 3.2 guarantees that a positive proportion of the elements of $P_{f,n}(Y)$ are irreducible of degree n and have Galois group S_n.

We define $S(Y)$ and $M_K(Y)$ as in the odd degree case. Taking $T = \kappa Y^{n-(4+2g)+\frac{2g^2+2g}{n}}$ and applying Lemma 5.4, we obtain the analogue of (5.8),

$$\sum_{|\mathrm{Disc}(K)| \le T} M_K(Y) \ll \kappa^{n/4} Y^c, \tag{5.15}$$

with c as in (5.14). As before, we choose κ sufficiently small so the left hand side of (5.15) is at most $\#P_{f,n}(Y, S_n)/2$, allowing us to only count the contribution of fields K with $T < \mathrm{Disc}(K)kY^{n(n-1)}$. Proceeding as in (5.10), we have

$$N_{n,C}(kY^{n(n-1)}, S_n) \gg Y^{c-n}T^{1/2} = Y^{\frac{1}{4}\left(n^2 - (2+2g)n + 2g^2 - 2g - 8 + \frac{4g^2+4g}{n}\right)}.$$

Replacing Y above by $(X/k)^{1/n(n-1)}$ we obtain $N_{n,C}(X, S_n) \gg X^{c_n}$, with

$$c_n = \frac{1}{4} - \frac{(1+2g)n^2 - (2g^2 - 2g - 8)n - (4g^2 + 4g)}{4n^2(n-1)},$$

as in Theorem 1.2.

To obtain the improved lower bound when n is sufficiently large, the procedure is identical to that of Section 5.4. The improved bound is again $N_{n,C}(kY^{n(n-1)}, S_n) \gg Y^{c-n/2}$, but with c given by (5.14) instead, leading to the exponent in the second statement of Theorem 1.2. For this to be valid, we need $N_n(X) \ll X^{\alpha(n,g)}$ with

$$\alpha(n,g) < \frac{n}{4} - \frac{1+g}{2} + \frac{g^2+g}{2n}. \tag{5.16}$$

This is satisfied for n sufficiently large by (5.13) of Theorem 5.5. A computer search using (1) and (2) of Theorem 5.5 can be used to explicitly find when the improved exponent takes effect. These come out to be quite similar to the odd degree case; for example, when $g = 1$, the improved exponent is valid for all $n \geq 108$. When $g = 10$, it is valid for all $n \geq 139$. For $g = 100$, this increases to $n \geq 325$. As in the previous case, for g sufficiently large, the improved exponent of Theorem 1.2 will be valid for all degrees $n \geq d + 2$. We computed this to be true for all $g \geq 237$.

References

[1] M. Bhargava, B. H. Gross & X. Wang, “A positive proportion of locally soluble hyperelliptic curves over $\mathbb{Q}$ have no point over any odd degree extension”, *J. Am. Math. Soc.* **30** (2017), no. 2, p. 451-493.

[2] D. S. Dummit & R. M. Foote, *Abstract algebra*, 3rd ed., Wiley, 2004.

[3] J. S. Ellenberg & A. Venkatesh, “The number of extensions of a number field with fixed degree and bounded discriminant”, *Ann. Math.* **163** (2006), no. 2, p. 723-741.

[4] M. Fujiwara, “Über die obere Schranke des absoluten Betrages der Wurzeln einer algebraischen Gleichung”, *Tôhoku Math. J.* **10** (1916), p. 167-171.

[5] A. Granville, “Rational and integral points on quadratic twists of a given hyperelliptic curve”, *Int. Math. Res. Not.* **2007** (2007), no. 8, article no. rnm027 (25 pages).

[6] S. Lang, *Algebraic number theory*, Graduate Texts in Mathematics, vol. 110, Springer, 1994.

[7] R. J. Lemke Oliver & F. Thorne, “Upper bounds on number fields of given degree and bounded discriminant”, `https://arxiv.org/abs/2005.14110`, to appear in *Duke Math. J.*, 2020.

[8] ———, “Rank growth of elliptic curves in non-abelian extensions”, *Int. Math. Res. Not.* **2021** (2021), no. 24, p. 18411-48441.

[9] B. Mazur, K. Rubin & M. Larsen, “Diophantine stability”, *Am. J. Math.* **140** (2018), no. 3, p. 571-616.

[10] J. Neukirch, *Algebraic number theory*, Graduate Texts in Mathematics, vol. 322, Springer, 1999.

[11] W. M. Schmidt, “Number fields of given degree and bounded discriminant”, in *Columbia University number theory seminar (New York, 1992)*, Astérisque, vol. 228, Société Mathématique de France, 1995, p. 189-195.

Christopher Keyes
Emory University
Department of Mathematics
400 Dowman Dr.,
Atlanta, GA 30223, USA
E-mail: `christopher.keyes@emory.edu`
URL: `http://www.math.emory.edu/~ckeyes3/index.html`