Algebra & Number Theory

Volume 20
2026
No. 3

Galois groups of reciprocal polynomials and the van der Waerden–Bhargava theorem

Theresa C. Anderson, Adam Bertelli and Evan M. O'Dorney

msp

ALGEBRA AND NUMBER THEORY 20:3 (2026)
https://doi.org/10.2140/ant.2026.20.603

msp

Galois groups of reciprocal polynomials and the van der Waerden–Bhargava theorem

Theresa C. Anderson, Adam Bertelli and Evan M. O'Dorney

We study the Galois groups G_f of degree $2n$ reciprocal (a.k.a. palindromic) polynomials f of height at most H, finding that G_f falls short of the maximal possible group $S_2 \wr S_n$ for a proportion of all f bounded above and below by constant multiples of $H^{-1}\log H$, whether or not f is required to be monic. This answers a 1998 question of Davis, Duke and Sun and extends Bhargava's 2023 resolution of van der Waerden's 1936 conjecture on the corresponding question for general polynomials. Unlike in that setting, the dominant contribution comes not from reducible polynomials but from those f for which $(-1)^n f(1) f(-1)$ is a square, causing G_f to lie in an index-2 subgroup.

1. Introduction

For a positive integer n, let $E_n(H)$ denote the number of degree n monic separable polynomials $f(x) = x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$ with integer coefficients $a_i \in [-H, H]$ whose Galois group is *not* S_n. Hilbert's irreducibility theorem implies that Galois groups not equal to S_n occur 0% of the time, in other words,

$$E_n(H) = o(H^n).$$

In 1936, van der Waerden [23] gave a quantitative upper bound and conjectured that the true order of growth is

$$E_n(H) \asymp H^{n-1}, \tag{1}$$

the lower bound coming from counting multiples of x (or any fixed monic linear polynomial). For the next several decades, progress continued, with many authors making improvements on the bound, including Knobloch (1955) [17], Gallagher (1972) [12], Chow and Dietmann (2020) [10] who proved it for $n \le 4$, and a group including the first author (2023) [1]. The conjecture (1) was finally resolved by Bhargava (2023) ([8]; an abridged version of the paper has appeared in print [7]). Bhargava's method harnesses sophisticated known results (classification of subgroups of S_n, distribution of discriminants of number fields) in combination with innovative recent methods, including Fourier equidistribution and the use of the *double discriminant* $\operatorname{disc}_{a_n} \operatorname{disc}_x f$. For transitive subgroups such as A_n, the frequency of appearance is expected to be much lower [5].

MSC2020: 11R32, 11R45, 11C08, 11N35, 20E22.
Keywords: arithmetic statistics, van der Waerden's conjecture, reciprocal polynomial, geometric sieve.

In this paper, we study the analogous question for the subspace of reciprocal polynomials, which was previously posed by Davis, Duke and Sun [11]. A polynomial f is called *reciprocal* if

$$f(1/x) = \frac{1}{x^{\deg f}} f(x),$$

in other words, if the coefficient list of f is palindromic. We focus on the case where f is of even degree $2n$ (in the odd-degree case, f must be reducible, equal to $x+1$ times an even-degree reciprocal polynomial). The roots of f come in reciprocal pairs $\{\alpha, 1/\alpha\}$; the Galois group must preserve the partition into pairs and thus is a subgroup of the wreath product $S_2 \wr S_n$, the subgroup of S_{2n} of order $2^n n!$ preserving this partition. In this paper, we provide a precise estimate for how often the group is strictly smaller:

Theorem 1.1. *Let $\mathcal{E}_n^{\text{monic}}(H)$ be the number of separable monic reciprocal polynomials f of degree $2n$ with coefficients in $[-H, H]$ whose Galois group is not $S_2 \wr S_n$. Then for each $n \geq 2$,*

$$\mathcal{E}_n^{\text{monic}}(H) \asymp H^{n-1} \log H.$$

Remark 1.2. Here and throughout the paper, if $f(H)$, $g(H)$ are real-valued functions of a sufficiently large real number H, then the usual notations

$$f(H) \ll g(H), \quad g(H) \gg f(H), \quad f(H) = O(g(H))$$

mean that $|f(H)| < c \cdot |g(H)|$ for sufficiently large H and some constant c, while the notations

$$f(H) \asymp g(H), \quad f(H) = \Theta(g(H))$$

mean that $f(H) \ll g(H)$ and $f(H) \gg g(H)$. Finally, $f(H) = o(g(H))$ means that $\lim_{H\to\infty} f(H)/g(H) = 0$. The implied constants may depend on n but not on H.

Theorem 1.1 is significant for several reasons:

- This is a sharp improvement on the work of Davis, Duke and Sun [11], who showed that $\mathcal{E}_n^{\text{monic}}(H) \ll H^{n-1/2} \log H$.
- The correct order of growth is *not* H^{n-1}, as one might expect by counting the reducible polynomials by analogy with van der Waerden's conjecture. Instead, the dominant contribution comes from reciprocal polynomials f such that $(-1)^n f(1) f(-1)$ is a square. This makes disc f a square and causes the Galois group to lie in an index-2 subgroup, called G_1 in the classification below. Reciprocal f such that $(-1)^n f(1) f(-1)$ is a square (among other conditions) show up as characteristic polynomials of automorphisms of lattices in [13].
- Since the characteristic polynomial of a symplectic matrix is reciprocal, a potential application of this work is to understand the characteristic polynomials of random elements of $\mathrm{Sp}_{2n}(\mathbb{Z})$, extending the work of Rivin [21] for $\mathrm{SL}_n(\mathbb{Z})$ and that of the first author and Lemke Oliver [16] for the symplectic case. Since the Weyl group of Sp_{2n} is $S_2 \wr S_n$, this fits into a wider principle that a random element of a connected

split reductive group $G/\mathbb{Q}$, faithfully embedded, should be almost surely equal to the Weyl group of G [15; 18]. Additionally, there are interesting connections to results in quantum chaos (see below).

Throughout this paper, n is fixed and the height $H \to \infty$ (the "large box model"). Work has also been fruitful for the complementary question on the "restricted coefficient model" where the coefficients are drawn from a fixed set and $n \to \infty$ [9; 3; 4]. In particular, Hokken [14] has recently proved an asymptotic count of the number of ± 1-coefficient reciprocal polynomials where the discriminant is a square, that is, $G_f \subseteq G_1$ in the notation below.

As with Bhargava [8] and many other papers on van der Waerden's conjecture, the methods are largely indifferent to whether we look at monic polynomials or general nonmonic polynomials f with all coefficients ranging through a box $[-H, H]^{n+1}$. The analogue of Theorem 1.1 for the nonmonic setting is:

Theorem 1.3. *Let $\mathcal{E}_n(H)$ be the number of separable reciprocal polynomials f of degree $2n$ with coefficients in $[-H, H]$ whose Galois group is not $S_2 \wr S_n$. Then for all $n \geq 1$,*

$$\mathcal{E}_n(H) \asymp H^n \log H.$$

Remark 1.4. Here the range of applicability is $n \geq 1$. In Theorem 1.1, we must exclude $n = 1$, because a monic reciprocal quadratic polynomial $f(x) = x^2 + a_1 x + 1$ has full Galois group S_2 for all $a_1 \neq \pm 2$.

Note that a degree $2n$ polynomial f is reciprocal if and only if the Cayley-transformed polynomial

$$\widetilde{f}(x) = (1+x)^{2n} f\left(\frac{1-x}{1+x}\right)$$

is even, that is, $\widetilde{f}(x) = \widetilde{g}(x^2)$ for a polynomial $\widetilde{g}$. (This is a straightforward computation and well known, see for instance [19, p. 275].) For an even polynomial, the roots again come in pairs $\{\alpha, -\alpha\}$. So we also have the following corollary:

Corollary 1.5. *The number of degree $2n$ even polynomials $\widetilde{f}$ with coefficients in $[-H, H]$ whose Galois group is not $S_2 \wr S_n$ is $\Theta(H^n \log H)$.*

Note that in this setting, the condition for the Galois group to lie in G_1 is that the product $(-1)^n a_{2n} a_0$ of the first and last coefficients of $\widetilde{f}$ be a square. If we impose $a_{2n} = 1$, the likelihood of this rises to $O(H^{-1/2})$, so the naïve analogue of Corollary 1.5 in the monic setting is false.

As alluded to earlier, there are potential applications of our work to the area of quantum chaos. In particular, studying the mass of eigenfunctions in quantum chaos is an area of great mathematical interest. This area shares exciting connections with reciprocal polynomials via quantum cat maps, a toy model of study in the area, that are given by symplectic matrices A. In particular, an important object to study the mass of eigenfunctions is the semiclassical measure μ associated to A. It turns out that μ has nice support properties if the characteristic polynomial of A^m for all $m \in \mathbb{N}$ (including A itself) is irreducible over the integers. In an appendix to a paper of Elena Kim, the first author and Lemke Oliver show that this irreducibility happens 100% of the time and moreover that the generic Galois group of such polynomials is the wreath product $S_2 \wr S_n$ (see [16] for more precise information and connections).

Based on this recent result, we conjecture that if M is a random $n \times n$ matrix, then the characteristic polynomials not only of M but of its powers M^i, $i \geq 1$, are almost surely S_n, where "almost surely" entails error bounds of van der Waerden type. This can be interpreted as saying that, if f is the characteristic polynomial, then not only its root α but each of its powers α^i, $i \geq 1$, generates an S_n-extension of $\mathbb{Q}$. The present work can be regarded as an extension of this to the fractional power $\sqrt{\alpha}$. It is natural to hope that the higher-order roots $\sqrt[m]{\alpha}$ of polynomials $g(x^m)$ can be handled in a similar way, the generic Galois group now being the semidirect product

$$(\mu_m \wr S_n) \rtimes (\mathbb{Z}/m\mathbb{Z})^\times,$$

where $(\mathbb{Z}/m\mathbb{Z})^\times$ acts naturally on the group μ_m of mth roots of unity.

Our methods are parallel to Bhargava's in [8], with judicious modifications when necessary. In Section 2, we lay out preliminary facts about reciprocal polynomials. In Section 3, we classify maximal subgroups of $S_2 \wr S_n$, and in particular, we show that we can narrow our focus to three groups, which we denote G_1, G_2, and G_3. In Sections 4, 5, and 6, we count the number $\mathcal{E}_n(G_i; H)$ of polynomials having each of those Galois groups (or a subgroup thereof). The G_1-polynomials we can count by direct parametrization. To count the G_3-polynomials f, we take advantage of the fact that f is reducible over a quadratic extension of $\mathbb{Q}$. The heart of the argument is to handle G_2, which plays a challenging role like that of the alternating group A_n in the study of general polynomials. Following Bhargava, we divide up the polynomials into three cases based on the size of the discriminant and its prime divisors. While attacking each case in turn, we are led to apply Fourier analysis for equidistribution and to construct a suitably modified double discriminant. The Fourier analysis is more involved than Bhargava's and involves breaking up the desired count into terms supported on different sublattices of the lattice $V(\mathbb{Z})$ of reciprocal polynomials. This type of decomposition is used in harmonic analysis to split up a function into pieces with different Fourier properties, but our use is perhaps novel in this setting.

Because Theorems 1.1 and 1.3 are so similar, we focus on Theorem 1.3, the nonmonic setting, which is the technically simpler of the two. At the end of each of Sections 4, 5, and 6, we explain how the proof must be adapted to the monic case to prove Theorem 1.1.

2. Reciprocal polynomials

We define the *height* of an integer polynomial

$$P(x) = c_n x^n + c_{n-1} x^{n-1} + \cdots + c_1 x + c_0 \in \mathbb{Z}[x]$$

to be the maximum of the coefficients:

$$\mathrm{Ht}\, P = \max\bigl\{|c_n|, |c_{n-1}|, \ldots, |c_0|\bigr\} \in \mathbb{Z}_{\geq 0}.$$

Let

$$f(x) = a_0 x^{2n} + a_1 x^{2n-1} + \cdots + a_{n-1} x^{n+1} + a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0$$

be a reciprocal polynomial of degree $2n$. Note that there is a unique degree n integer-coefficient polynomial

$$g(u) = b_n u^n + b_{n-1} u^{n-1} + \cdots + b_1 u + b_0$$

such that

$$f(x) = x^n g\left(x + \frac{1}{x}\right).$$

The passage between f and g is bijective and linear, and it increases or decreases heights by at most a bounded factor, depending only on n. Hence it is immaterial whether we count f or g of height at most H. For most purposes, it is more convenient to count g.

We denote the roots of f by

$$\alpha_1, \frac{1}{\alpha_1}, \alpha_2, \frac{1}{\alpha_2}, \ldots, \alpha_n, \frac{1}{\alpha_n}.$$

Then the roots of g are $\beta_1, \ldots, \beta_n$, where

$$\beta_i = \alpha_i + \frac{1}{\alpha_i}.$$

We will sometimes write $\alpha = \alpha_1$ and $\beta = \beta_1$ when the choice of root is irrelevant.

In view of the main theorem we would like to prove, we can assume any statement that occurs for all but $O(H^n)$ of the $O(H^{n+1})$ polynomials g of height at most H. For example, we may assume that g is irreducible and that $g(2)$ and $g(-2)$ are nonzero. We have the tower of number fields

$$K_f = \mathbb{Q}(\alpha) \quad \supseteq \quad K_g = \mathbb{Q}(\beta) \quad \supset \quad \mathbb{Q},$$

where K_g is an S_n-extension of degree n, while K_f/K_g is of degree at most 2, given by $K_f = K_g(\sqrt{\beta^2 - 4})$. Let $\widetilde{K}_g$ and $\widetilde{K}_f$, respectively, be the splitting fields of g and f, and let G_g and G_f be their respective Galois groups, which are subgroups of S_n and $S_2 \wr S_n$, with $G_f \twoheadrightarrow G_g$ under the natural projection $S_2 \wr S_n \twoheadrightarrow S_n$. By the main result of Bhargava [8, Theorem 1], we can assume that G_g is the whole S_n. Our aim in this paper is to understand when G_f is not the whole $S_2 \wr S_n$.

By the usual formula for the discriminant of a number field tower, we have

$$(\operatorname{Disc} K_f) = (\operatorname{Disc} K_g)^2 \cdot N_{K_g/\mathbb{Q}} \operatorname{disc}_{K_g} K_f$$

as ideals in $\mathbb{Z}$. The following is a closely related result on the discriminants of the associated polynomials.

Lemma 2.1. $\operatorname{disc} f = (-1)^n g(2) g(-2) (\operatorname{disc} g)^2$.

Proof. We have

$$\begin{aligned}
\operatorname{disc} f &= a_0^{4n-2} \prod_{i=1}^{n} (\alpha_i - \alpha_i^{-1})^2 \prod_{i<j} (\alpha_i - \alpha_j)^2 (\alpha_i - \alpha_j^{-1})^2 (\alpha_i^{-1} - \alpha_j)^2 (\alpha_i^{-1} - \alpha_j^{-1})^2 \\
&= b_n^{4n-2} \prod_{i=1}^{n} (\beta_i^2 - 4) \cdot \prod_{i<j} \alpha_i^{-4} \alpha_j^{-4} (\alpha_i - \alpha_j)^4 (\alpha_i \alpha_j - 1)^4 \\
&= (-1)^n \cdot b_n \prod_{i=1}^{n} (2 - \beta_i) \cdot b_n \prod_{i=1}^{n} (-2 - \beta_i) \cdot b_n^{4n-4} \prod_{i<j} (\beta_i - \beta_j)^4 \\
&= (-1)^n \cdot g(2) \cdot g(-2) \cdot (\operatorname{disc} g)^2.
\end{aligned}$$

□

3. Maximal subgroups of $S_2 \wr S_n$

If G_f is not the full $S_2 \wr S_n$, it is contained in a maximal subgroup. We have $S_2 \wr S_n = \mathbb{F}_2^n \rtimes S_n$. The following elementary lemma classifies the maximal subgroups of a semidirect product whose normal factor is abelian.

Lemma 3.1. *Let $X \rtimes S$ be a semidirect product of groups, with X abelian. The maximal subgroups of $X \rtimes S$ are of two types*:

(a) *$X \rtimes S'$, for $S' < S$ a maximal subgroup, and*

(b) *those groups $G < X \rtimes S$ for which $Y = G \cap X$ is a maximal S-invariant subgroup of X such that the projection $G \to S$ is surjective.*

Proof. Let $G < X \rtimes S$ be a maximal subgroup, and let S' be the projection of G onto S. If $S' \neq S$, then $G \leq X \rtimes S'$, and we must have equality so we get a group of type (a).

So we assume $S' = S$. Then $Y = G \cap X$ is a subgroup of X, not the whole of X since $G \neq X \rtimes S$. The conjugation action of $X \rtimes S$ on X is the S-action. Since G surjects onto S, the fact that G is closed under conjugation by itself implies that Y is closed under the S-action. Suppose that Y is not maximal as an S-invariant subspace, $Y < Y' < X$. Let

$$GY' = \{gy : g \in G,\ y \in Y'\} \subset X \rtimes S.$$

We claim that GY' is a subgroup. Since G and Y' are subgroups, it suffices to show that any product yg, $y \in Y'$, $g \in G$, belongs to GY'. Write $g = sx$, $x \in X$, $s \in S$. Since X is abelian,

$$yg = ysx = sy^s x = sxy^s = gy^s,$$

where y^s denotes the conjugate $s^{-1}ys$. Since Y' is S-invariant, the last product belongs to GY'. So GY' is a subgroup.

It is evident that $GY' \cap X = Y'$. So $G < GY' < X \rtimes S$, contradicting maximality of G. So Y is maximal, and G is a group of type (b). □

By Bhargava's result [8, Theorem 1], the extension $\widetilde{K}_g$ has full Galois group S_n for all but $O(H^n)$ polynomials g (or $O(H^{n-1})$ in the monic case), so Theorems 1.1 and 1.3 are already known for subgroups of type (a). To classify the subgroups of type (b), we must find the maximal S_n-invariant subspaces of $X = \mathbb{F}_2^n$, a vector space equipped with an S_n-action permuting the n basis vectors freely. For convenience we let

$$\mathbf{0} = (0, \dots, 0), \quad \mathbf{1} = (1, \dots, 1) \in X.$$

We use a superscript $\perp$ to denote the orthogonal complement of a space under the S_n-invariant inner product

$$(x_1, \dots, x_n) \bullet (y_1, \dots, y_n) = \sum_{i=1}^{n} x_i y_i.$$

Lemma 3.2. *Let $X = \mathbb{F}_2^n$, with the permutation action of S_n. The S_n-invariant subspaces of X are*

- X,
- $\langle \mathbf{1} \rangle^\perp = \{\mathbf{v} = (v_1, \dots, v_n) : \sum_i v_i = 0\}$,
- $\langle \mathbf{1} \rangle = \{\mathbf{0}, \mathbf{1}\}$,
- $\{\mathbf{0}\}$.

Proof. Let $W \subseteq X$ be an invariant subspace. If W contains a vector $\mathbf{v} = (v_1, \dots, v_n)$ besides $\mathbf{0}$ and $\mathbf{1}$, then upon applying some element of S_n we can assume $v_1 = 1$, $v_2 = 0$. Then $W \ni \mathbf{v} + (12)\mathbf{v} = \langle 1, 1, 0, \dots, 0 \rangle$. Applying further permutations from S_n, we get that W contains every vector with exactly two nonzero coordinates. Then W contains their span, which is $\langle \mathbf{1} \rangle^\perp$. Hence $W = \langle \mathbf{1} \rangle^\perp$ or $W = X$. □

Among these, the maximal subgroups are

- $\langle \mathbf{1} \rangle^\perp$, and
- $\langle \mathbf{1} \rangle$, for n odd (since for n even, $\langle \mathbf{1} \rangle \subseteq \langle \mathbf{1} \rangle^\perp$).

Note that for n odd, we have a direct sum decomposition $X = \langle \mathbf{1} \rangle \oplus \langle \mathbf{1} \rangle^\perp$. We now classify the groups G using group cohomology as follows:

Lemma 3.3. *Let $X \rtimes S$ be a semidirect product of groups, with X abelian, and let $K \subseteq X$ be a subgroup fixed by S. The subgroups of $G \subseteq X \rtimes S$ such that $G \cap X = K$ are parametrized by 1-cocycles $\varepsilon \in Z^1(S, X/K)$, in other words, maps $\varepsilon \colon S \to X/K$ satisfying the cocycle condition*

$$\varepsilon(\sigma\tau) = \varepsilon(\sigma) + \sigma(\varepsilon(\tau)). \tag{2}$$

The map sends each ε to the group

$$G = \{(x, \sigma) : x \equiv \varepsilon(\sigma) \bmod K\}. \tag{3}$$

Moreover, two such subgroups G, G' are conjugate if and only if the corresponding ε, ε' are in the same cohomology class in $H^1(S, X/K)$; in other words, if there is a $y \in X$ such that for all $\sigma \in S$,

$$\varepsilon'(\sigma) = \varepsilon(\sigma) + \sigma(y) - y. \tag{4}$$

Proof. This is a fairly standard use of group cohomology. Indeed, it is easy to check that G must have the form (3), that closure under multiplication enforces (2), and that conjugation by $y \in X$ induces (4). (Since G surjects onto S, conjugation by X is sufficient to produce all the conjugates of G in $X \rtimes S$.) □

Theorem 3.4. *The maximal subgroups of $S_2 \wr S_n$ whose projection onto S_n is the whole group are*

$$G_1 = \Big\{(\mathbf{v}, \sigma) \in \mathbb{F}_2^n \rtimes S_n : \sum_i v_i = 0\Big\} = \langle \mathbf{1} \rangle^\perp \rtimes S_n \quad \text{for } n \geq 1,$$

$$G_2 = \Big\{(\mathbf{v}, \sigma) \in \mathbb{F}_2^n \rtimes S_n : \sum_i v_i = \operatorname{sgn} \sigma\Big\} \quad \text{for } n \geq 2,$$

$$G_3 = \langle \mathbf{1} \rangle \times S_n \quad \text{for } n \geq 3 \text{ odd}.$$

Proof. By the preceding lemmas, we are left with computing $H^1(S_n, X/Y)$ for each of the maximal S_n-invariant subspaces Y in Lemma 3.2.

If $Y = \langle \mathbf{1} \rangle^\perp$, then

$$H^1(S_n, X/\langle \mathbf{1} \rangle^\perp) = H^1(S_n, C_2) = \mathrm{Hom}(S_n, C_2) = C_2,$$

the two maps ε being the zero map and the sign map, giving the subgroups G_1 and G_2 claimed.

If $Y = \langle \mathbf{1} \rangle$ for n odd, we must compute

$$H^1(S_n, X/\langle \mathbf{1} \rangle).$$

Since n is odd, we have that $X = \langle \mathbf{1} \rangle \oplus \langle \mathbf{1} \rangle^\perp$ is a direct sum and $X/\langle \mathbf{1} \rangle \cong \langle \mathbf{1} \rangle^\perp$. First consider

$$H^1(S_n, X).$$

With this action, $X = \mathrm{Ind}_{S_{n-1}}^{S_n} C_2$ is an induced module, so by Shapiro's lemma,

$$H^1(S_n, X) = H^1(S_{n-1}, C_2) = \mathrm{Hom}(S_{n-1}, C_2) = C_2.$$

Hence

$$H^1(S_n, X/\langle \mathbf{1} \rangle) = \frac{H^1(S_n, X)}{H^1(S_n, C_2)} = 0.$$

Therefore there is only the trivial extension G_3.

Note that G_1, G_2, and G_3 are normal subgroups of X, so we get no further maximal subgroups via conjugation.

The restrictions on n are provided due to the fact that, for some n, the G_i are nonmaximal or coincident:

- For $n = 1$, $G_2 = G_1$ and $G_3 = S_2 \wr S_n$.
- For n even, $G_3 \subseteq G_1$. □

Remark 3.5. A proof of Theorem 3.4 without group cohomology is possible, but it involves some in-depth case analysis with many computations of products in S_n and $S_2 \wr S_n$. After this paper was circulated in preprint form, we learned of the useful reference [2] where these computations are carried out.

Remark 3.6. With a few more cohomological computations, we can classify *all* the subgroups of $S_2 \wr S_n$ that surject onto S_n. They are the whole $S_2 \wr S_n$, the groups G_1, G_2, and G_3 (without parity restrictions on n), and the following additional groups:

- $\{0\} \times S_n \cong S_n$
- $\big\{\big((\mathrm{sgn}\,\sigma)\mathbf{1}, \sigma\big) : \sigma \in S_n\big\}$, a twisted copy of S_n
- and, for $n = 4$ only, the group $\mathrm{GL}_2\mathbb{F}_3$, which acts on the 8-element set $\mathbb{F}_3^2 \setminus \{0\}$, preserving the partition into opposite pairs of vectors. In terms of its map to S_4, this is the double cover denoted $2 \cdot S_4^+$ in [24; 20]. For instance, this is the generic Galois group of the even octic minimal polynomial of the y-coordinate of a 3-torsion point on an elliptic curve over $\mathbb{Q}$. It is a proper subgroup of the maximal subgroup G_2.

To prove Theorems 1.1 and 1.3, we must bound the number of polynomials f, equivalently g, for which G_f is (conjugate to) a subgroup of G_1, G_2, or G_3 for the values of n listed in Theorem 3.4.

Definition 3.7. For $G \subseteq S_2 \wr S_n$ and $H \geq 2$, let $\mathcal{E}_n(G; H)$ be the number of separable reciprocal polynomials f of degree $2n$ with coefficients in $[-H, H]$ such that $G_g = S_n$ and $G_f \subseteq G$. Let $\mathcal{E}_n^{\text{monic}}(G; H)$ count the subset of these that are monic.

The following lemma reduces computing $\mathcal{E}_n(G_1; H)$ and $\mathcal{E}_n(G_2; H)$ to number-theoretic conditions on g that must hold in order to fulfill these conditions. (For G_3, we will use a different technique: see Section 6.)

Lemma 3.8. *Assume that G_g is the whole of S_n. Then*:

(a) *$G_f \subseteq G_1$ if and only if $(-1)^n g(2)g(-2)$ is a square.*

(b) *$G_f \subseteq G_2$ if and only if $(-1)^n g(2)g(-2)\operatorname{disc} g$ is a square.*

Proof. For (a), note that G_1 is the preimage of A_{2n} under the usual inclusion $S_2 \wr S_n \hookrightarrow S_{2n}$. Thus $G_f \subseteq G_1$ if and only if $\operatorname{disc} f = (-1)^n g(2)g(-2)(\operatorname{disc} g)^2$ is a square, which happens exactly when $(-1)^n g(2)g(-2)$ is a square.

For (b), consider the embedding of $S_2 \wr S_n$ into S_{3n} given by its action on the disjoint union of the roots of f and of g. Note that G_2 is the preimage of A_{3n} under this embedding. Hence $G_f \subseteq G_2$ if and only if $\operatorname{disc} f \cdot \operatorname{disc} g$ is a square, which happens exactly when $(-1)^n g(2)g(-2)\operatorname{disc} g$ is a square. □

4. Counting G_1-polynomials

We first deal with the case G_1, which yields the main term of Theorems 1.1 and 1.3.

Theorem 4.1. *For $n \geq 1$,*

$$\mathcal{E}_n(G_1; H) \asymp H^n \log H \tag{5}$$

and for $n \geq 2$,

$$\mathcal{E}_n^{\text{monic}}(G_1; H) \asymp H^{n-1} \log H. \tag{6}$$

By Lemma 3.8(a), it suffices to count g such that $(-1)^n g(2)g(-2)$ is a square z^2.

Lemma 4.2. *The number of solutions to the equation $xy = z^2$, $1 \leq x, y, z \leq H$ $(H \geq 2)$ is $\Theta(H \log H)$.*

Proof. A parametrization of the solutions is given by

$$x = ku^2, \quad y = kv^2, \quad z = kuv$$

where k, u, v are positive integers and $\gcd(u, v) = 1$. For each k, $1 \leq k \leq H$, the pair (u, v) is chosen from the box $1 \leq u, v \leq \sqrt{H/k}$, and the number of coprime pairs in this box is $\Theta(H/k)$ (the lower bound comes from citing the limiting proportion $6/\pi^2 > 0$ of coprime pairs when H/k is large, and noting that there is always at least one solution $u = v = 1$). So the total number N of solutions satisfies $N \asymp \sum_{k=1}^{H} H/k \asymp H \log H$, as desired. □

Proof of Theorem 4.1. For simplicity, we prove the nonmonic case (5). As $g(2), g(-2) \ll H$ and each pair $(x, y) = (|g(2)|, |g(-2)|)$ appears $O(H^{n-1})$ times, we get that $G_f \subseteq G_1$ at most $O(H^n \log H)$ times. Conversely, if we take $|x|, |y| \le cH$ for an appropriate constant c, and with $x \equiv y \bmod 4$ (which can be arranged, for instance by taking $4|k$), we find that there are $\Theta(H^{n-1})$ polynomials g with $g(2) = x$ and $g(2) = y$, and thus $\Theta(H^{n-1} \log H)$ polynomials overall with Galois group $G_f \subseteq G_1$. □

4.1. *Remarks on the monic case.* For the monic case, the argument is identical, replacing n by $n-1$. It is only necessary to have at least two free coefficients so that $g(2)$ and $g(-2)$ can be adjusted independently, requiring $n \ge 2$.

5. Counting G_2-polynomials

For G_2, we prove the following bounds, which are stronger than those for G_1 by a factor of $\log H$:

Theorem 5.1. *For $n \ge 2$,*

$$\mathcal{E}_n(G_2; H) \ll H^n, \tag{7}$$

$$\mathcal{E}_n^{\text{monic}}(G_2; H) \ll H^{n-1}. \tag{8}$$

By Lemma 3.8(b), we wish to count g such that $(-1)^n g(2)g(-2)\operatorname{disc} g$ is a square. We use a sieve method adapted from Bhargava [8]. We begin with some analytic preliminaries.

5.1. *Twisted Poisson summation.* Let $\Phi\colon \mathbb{R}^n \to \mathbb{C}$ be a Schwartz function. We normalize the Fourier transform by

$$\widehat{\Phi}(y) = \int_{\mathbb{R}^n} e^{-2\pi\sqrt{-1}x\bullet y}\,\Phi(x)\,dx.$$

The usual Poisson summation formula

$$\sum_{x\in\mathbb{Z}^n} \Phi(x) = \sum_{y\in\mathbb{Z}^n} \widehat{\Phi}(y)$$

can be extended in various ways. If $L \subseteq \mathbb{Z}^n$ is a lattice (a subgroup of finite index), then

$$\sum_{x\in L} \Phi(x) = \frac{1}{[\mathbb{Z}^n : L]} \sum_{y\in L^*} \widehat{\Phi}(y), \tag{9}$$

where $L^* \supseteq \mathbb{Z}^n$ is the dual lattice. More generally:

Proposition 5.2. *Let $L \subseteq \mathbb{Z}^n$ be a lattice, and let $\Psi\colon (\mathbb{Z}/M\mathbb{Z})^n \to \mathbb{C}$ be any function, where the modulus M is coprime to $[\mathbb{Z}^n : L]$. Let $\widehat{\Psi}\colon (\mathbb{Z}/M\mathbb{Z})^n \to \mathbb{C}$ be the Fourier transform*

$$\widehat{\Psi}(y) = \frac{1}{M^n} \sum_{x\in(\mathbb{Z}/M\mathbb{Z})^n} e^{-2\pi\sqrt{-1}x\bullet y/M}\,\Psi(x).$$

For any Schwartz function Φ,

$$\sum_{x \in L} \Psi(x)\Phi(x) = \frac{1}{[\mathbb{Z}^n : L]} \sum_{y \in L^*} \widehat{\Psi}(y)\widehat{\Phi}\left(\frac{y}{M}\right).$$

Observe that $\widehat{\Psi}$ is well defined on L^* since M is coprime to $[\mathbb{Z}^n : L]$. (In fact, one can do without this hypothesis, but then $\widehat{\Psi}$ becomes a function on L^*/ML^* instead of being independent of L.) To prove this proposition, one may assume by linearity that Ψ is supported on a single point, and the result reduces to Poisson summation.

5.2. *The index of a polynomial over a field.* Following Bhargava [8, Proposition 26], we make the following definition. If $P \in \Bbbk[x, y]$ is a nonzero homogeneous polynomial over a field, we factor $P = \prod_i P_i^{e_i}$ into powers of distinct irreducibles and define the *index* of P to be

$$\operatorname{ind}(P) = \sum_i (e_i - 1) \deg P_i.$$

The index is significant for bounding the power of a prime p dividing the discriminant of the polynomial and the extension field it defines. The following lemma is used implicitly in [8]; for completeness, we offer a statement and proof.

Lemma 5.3. *Let $g \in R[x, y]$ be a separable homogeneous binary form of degree n over a Dedekind domain R. Let $F = \operatorname{Frac} R$, and let $E = F[\beta]/g(\beta, 1)$ be the étale algebra (product of separable finite field extensions) defined by g. If $\mathfrak{p}$ is a prime ideal of R such that E is tamely ramified at $\mathfrak{p}$ (e.g., $\operatorname{char}(R/\mathfrak{p}) > n$) and g is not identically zero modulo $\mathfrak{p}$, then*

$$v_{\mathfrak{p}}(\operatorname{Disc} E) \le \operatorname{ind}(g \bmod p) \le v_{\mathfrak{p}}(\operatorname{disc} g). \tag{10}$$

Proof. We have deliberately stated the lemma in greater generality than needed to allow for making some reductions. First, we may replace R by its completion at $\mathfrak{p}$. Let $\Bbbk$ be the residue field of R. Now, if $g \equiv g_1 g_2 \bmod \mathfrak{p}$ with $g_i \in \Bbbk[x, y]$ homogeneous and coprime, then by Hensel's lemma, the factorization lifts to R and induces a splitting $E = E_1 \times E_2$. The inequality (10) can then be deduced by summing the corresponding inequalities for g_1 and g_2. Thus we may assume that

$$g \equiv c \cdot g_1^e \bmod \mathfrak{p},$$

where c is a constant and $g_1 \in \Bbbk[x, y]$ is irreducible of some degree f, with $ef = n$. We may change coordinates so that $g_1 \neq y$ is monic in x. Now

$$E = \prod_{j=1}^{r} E_j$$

is a product of fields with the same inertia index f and possibly different ramification indices e_j. We compute, using the usual formula for the discriminant of a tamely ramified extension:

- First,
$$v_{\mathfrak{p}}(\operatorname{Disc} E) = \sum_{j=1}^{r} v_{\mathfrak{p}}(\operatorname{Disc} E_i) = \sum_{j=1}^{r} (e_j - 1) f = n - rf.$$
- $\operatorname{ind}(g \bmod p) = (e-1)f = n - f$.
- Finally, we need to understand the ring $S = R[\beta]/g(\beta, 1) \subset E$. Note that S is contained in
$$S' = \{(x_1, \dots, x_r) \in \mathcal{O}_E : x_1 \equiv \cdots \equiv x_r \bmod \mathfrak{p}\},$$
a subring of $\mathcal{O}_E$ of index $(r-1)f$. Thus
$$\begin{aligned} v_{\mathfrak{p}}(\operatorname{disc} g) &= v_{\mathfrak{p}}(\operatorname{Disc} S) \\ &\geq v_{\mathfrak{p}}(\operatorname{Disc} S') = v_{\mathfrak{p}}(\operatorname{Disc} E) + 2 v_{\mathfrak{p}}([\mathcal{O}_E : S']) \\ &= n - rf + 2(r-1)f = n - f + (r-1)f. \end{aligned}$$

The desired inequality follows immediately. Equality for both parts holds exactly when $r = 1$, or in the original setup, when $\mathfrak{p} \nmid [\mathcal{O}_E : S]$. □

Following Bhargava [8, §5], we define
$$D = \operatorname{Disc} K_g \quad \text{and} \quad C = \prod_{p \mid D} D$$
and divide the counting into three cases based on the sizes of C and D relative to H. Unlike in [8], we do not have that D is squareful; but by Lemma 3.8(b) we have that $(-1)^n D g(2) g(-2)$ is a square, which limits the cases.

Let δ be a small constant (such as $1/4n$). For R a ring, denote by $V^{\text{hom}}(R)$ the $(n+1)$-dimensional space of binary n-ic forms $P(x, y)$ over R, and denote by $V(R)$ the space of polynomials $g(u)$ of degree at most n over R. The two R-modules are isomorphic, but we will need to identify them in multiple ways.

5.3. ***Case I:*** $C \leq H^{1+\delta}$, $D \geq H^{2+2\delta}$. In this case, we need to estimate the number of g for which $D = \operatorname{Disc}(K_g)$, $g(2)$, and $g(-2)$ have certain factors. We begin with a short argument that yields the result up to ε.

Lemma 5.4. *Let p be a prime, and let k be an integer. The number of binary forms $g \in V^{\text{hom}}(\mathbb{F}_p)$ such that*

- *$g(1, 0) = 0$, that is, $y \mid g$, and*
- *$\operatorname{ind}(g) \geq k$*

is $O(p^{n-k})$.

Proof. We can immediately dispose of the case $p \leq n$, for here both the number of g and the desired bound are $O(1)$. In [8, Corollary 27], it is shown that the number of degree-n binary forms g such that $\operatorname{ind}(g) \geq k$ is $O(p^{n+1-k})$. To impose the condition $g(1, 0) = 0$, we can consider each g as lying in the family of p translates

$$g(x, y+ax), \quad a \in \mathbb{F}_p.$$

The translates all have the same index, and if $p > n$, the translates are all distinct. Moreover, since g has at most n roots over $\mathbb{F}_p$, at most $n = O(1)$ of the translates satisfy the added condition $g(1, 0) = 0$. Hence the total number of such g is $O(p^{n-k})$, as desired. □

Let $p > n$ be a prime dividing C, and suppose $p^k \parallel D$. By the first inequality in (10), we have $\operatorname{ind}(g) \geq k$, and this occurs for a proportion p^{-k} of g. If k is odd, then we additionally have $p \mid g(2)$ or $p \mid g(-2)$, and altogether there is a proportion p^{-k-1} of g satisfying these conditions. Multiplying over p, the proportion of G_2-polynomials with Disc $K_g = D$ is bounded by

$$\prod_{p \mid D} O\big(p^{-2\lceil k/2\rceil}\big) = \frac{O(c^{\omega(D_1)})}{D_1^2},$$

where $D_1^2 = \prod_{p \mid D} p^{2\lceil k/2\rceil}$ is the least square divisible by D. Observe that $D_1 \gg H$ and that each D_1 occurs for at most $2^{\omega(D_1)}$ values of D. Moreover, these g are cut out by congruence conditions mod C. If $C \leq H$, then we can estimate the number of lattice points very precisely because our modulus is lower than the size of the box. We get that the number of polynomials g is

$$\ll H^{n+1} \sum_{D_1 \geq H} \frac{c^{\omega(D_1)}}{D_1^2} \ll_\varepsilon H^{n+\varepsilon}.$$

Using Fourier analysis we can remove the ε and also extend the validity of this case from $C \leq H$ to $C \leq H^{1+\delta}$.

Recall some definitions from [8, §4.1]. If a binary n-ic form f (over $\mathbb{Z}$, or over $\mathbb{F}_p$) factors modulo p as $\prod_{i=1}^r P_i^{e_i}$, with P_i irreducible and $\deg(P_i) = f_i$, then the *splitting type* (f, p) of f is defined as $(f_1^{e_1} \cdots f_r^{e_r})$, and the *index* $\operatorname{ind}(f)$ of f modulo p (or the *index* of the splitting type (f, p) of f) is defined to be $\sum_{i=1}^r (e_i - 1) f_i$. If $p \mid f$, we put $\operatorname{ind}(f) = \infty$. More abstractly, a *splitting type* is an expression σ of the form $(f_1^{e_1} \cdots f_r^{e_r})$, where the f_i and e_i are positive integers. The *degree* $\deg(\sigma)$ is $\sum_{i=1}^r e_i f_i$, and the *index* $\operatorname{ind}(\sigma)$ is $\sum_{i=1}^r (e_i - 1) f_i$. Finally, $\#\operatorname{Aut}(\sigma)$ is defined to be $\prod_i f_i$ times the number of permutations of the factors $f_i^{e_i}$ that preserve σ.

Later on we will need to deal with splitting types with a distinguished factor of degree 1. If σ has $f_1 = 1$, we let $\#\operatorname{Aut}'(\sigma)$ be $\prod_i f_i$ times the number of permutations of the factors $f_i^{e_i}$, $i \geq 2$, that preserve σ.

We first recall the following lemma of Bhargava:

Lemma 5.5 [8, Proposition 26]. *Let $\sigma = (f_1^{e_1} \cdots f_r^{e_r})$ be a splitting type with* $\deg(\sigma) = d \leq n$ *and* $\operatorname{ind}(\sigma) = k$. *Let $w_{p,\sigma} : V^{\mathrm{hom}}(\mathbb{F}_p) \to \mathbb{C}$ be defined by*

$w_{p,\sigma}(f) :=$ *the number of r-tuples $(P_1, \ldots, P_r)$, up to the action of the group of permutations of $\{1, \ldots, r\}$ preserving σ, such that the P_i are distinct irreducible binary forms where, for each i, we have $P_i(x, y)$ is y or is monic as a polynomial in x,* $\deg P_i = f_i$, *and $P_1^{e_1} \cdots P_r^{e_r} \mid f$.*

Then

$$\widehat{w}_{p,\sigma}(g) = \begin{cases} \dfrac{p^{-k}}{\#\mathrm{Aut}(\sigma)} + O(p^{-(k+1)}) & \text{if } g = 0; \\ O(p^{-(k+1)}) & \text{if } g \neq 0. \end{cases}$$

The significance of this function $w_{p,\sigma}$ is twofold: First, $w_{p,\sigma}(f)$ is nonnegative, and is equal to 1 when f has splitting type σ; second, the definition is arranged so that $w_{p,\sigma}$ is a sum of characteristic functions of subspaces, which makes the Fourier transform nonnegative, small, and easily computable.

Bhargava uses this lemma to bound the number of integer polynomials having high index at a set of primes. In our setting, we also need to bound the number of integer polynomials having high index *and* which vanish at a given point mod p; hence we modify the lemma as follows:

Lemma 5.6. *Let* $\sigma = (f_1^{e_1} \cdots f_r^{e_r})$ *be a splitting type with* $f_1 = 1$. *Let* $\deg(\sigma) = d$ *and* $\mathrm{ind}(\sigma) = k$. *Let* $w'_{p,\sigma} : V^{\mathrm{hom}}(\mathbb{F}_p) \to \mathbb{C}$ *be defined by*

$w'_{p,\sigma}(f) :=$ *the number of r-tuples* $(P_2, \dots, P_r)$, *up to the action of the group of permutations of* $\{2, \dots, r\}$ *preserving* σ, *such that the* P_i *are distinct irreducible binary forms where, for each i, we have* $P_i(x, y)$ *is monic as a polynomial in x,* $\deg P_i = f_i$, *and* $y^{e_1} P_2^{e_2} \cdots P_r^{e_r} \mid f$.

Let $V^{\mathrm{hom}}_{e_1,\mathbb{F}_p}$ *denote the subspace of* $V^{\mathrm{hom}}(\mathbb{F}_p)$ *comprising polynomials divisible by* y^{e_1}, *and let* $(V^{\mathrm{hom}}_{e_1,\mathbb{F}_p})^{\perp} \subseteq V^{\mathrm{hom}}(\mathbb{F}_p)^*$ *be its dual. Then*

$$\hat{w}'_{p,\sigma}(g) = \begin{cases} \dfrac{p^{-(k+1)}}{\#\mathrm{Aut}'(\sigma)} + O(p^{-(k+2)}) & \text{if } g \in (V^{\mathrm{hom}}_{e_1,\mathbb{F}_p})^{\perp}; \\ O(p^{-(k+2)}) & \text{if } g \notin (V^{\mathrm{hom}}_{e_1,\mathbb{F}_p})^{\perp}. \end{cases}$$

Proof. Rather than starting from scratch, we derive this lemma from the preceding one. Indeed, let σ' be the splitting type obtained by deleting the first factor $f_1^{e_1} = 1^{e_1}$ from σ. Then $\mathrm{Aut}(\sigma') \cong \mathrm{Aut}'(\sigma)$, and $\mathrm{ind}(\sigma') = k - e_1 + 1$. We have that $w'_{p,\sigma}(f)$ vanishes unless $f \in V^{\mathrm{hom}}_{e_1,\mathbb{F}_p}$ (implying in particular that $\hat{w}'_{p,\sigma}$ is constant on cosets of $(V^{\mathrm{hom}}_{e_1,\mathbb{F}_p})^{\perp}$), and

$$w'_{p,\sigma}(f) = \breve{w}_{p,\sigma'}(f/y^{e_1})$$

is *almost* $w_{p,\sigma}(f/y^{e_1})$. We say "almost" because we need to exclude the case that one of the P_i is y; so we define $\breve{w}_{p,\sigma}$ to be just like $w_{p,\sigma}$, except that the $P_i(x, y)$ in the definition are not allowed to equal y. Since $w_{p,\sigma'}$ is a sum of characteristic functions of subspaces corresponding to the various choices of the P_i and $\breve{w}_{p,\sigma}$ is obtained by deleting some of the terms from this sum, we have, by Lemma 5.5, $\breve{w}_{p,\sigma'}(f) \leq w_{p,\sigma'}(f)$ and

$$\widehat{\breve{w}}_{p,\sigma'}(g) \leq \widehat{w}_{p,\sigma'}(g) = \begin{cases} \dfrac{p^{-(k-e_1+1)}}{\#\mathrm{Aut}'(\sigma) + O(p^{-(k-e_1+2)})} & \text{if } g \in V^{\perp}_{e_1,\mathbb{F}_p}; \\ O(p^{-(k-e_1+2)}) & \text{if } g \notin V^{\perp}_{e_1,\mathbb{F}_p}. \end{cases}$$

Re-embedding $V^{\mathrm{hom}}_{e_1,\mathbb{F}_p}$ into $V^{\mathrm{hom}}(\mathbb{F}_p)$, the Fourier transform drops by a factor of p^{e_1}, yielding upper bounds of the claimed order of magnitude.

To obtain that the main term is undiminished by the $w \mapsto \breve{w}$ replacement, we can argue that

$$\breve{w}_{p,\sigma} = w_{p,\sigma} - \sum_{i:f_i=1} \breve{w}_{p,\sigma_i}$$

where σ_i is obtained by deleting $f_i^{e_i}$ from σ. Thus

$$\begin{aligned}\widehat{\breve{w}}_{p,\sigma}(0) &= \widehat{w}_{p,\sigma}(0) - \sum_{i:f_i=1} \widehat{\breve{w}}_{p,\sigma_i}(0) \\ &= \frac{p^{-(k-e_1+1)}}{\#\mathrm{Aut}'(\sigma)} + O(p^{-(k-e_1+2)}) + \sum_{i:f_i=1} O(p^{-(k-e_1-e_i+1)}) \\ &= \frac{p^{-(k-e_1+1)}}{\#\mathrm{Aut}'(\sigma)} + O(p^{-(k-e_1+2)}),\end{aligned}$$

as desired. □

We can now estimate the number of polynomials in Case I.

Lemma 5.7. *Let D be a positive integer. Let $C = \prod_{p|D} p$ be its radical, and let $D'^2 = \prod_{p|D} p^{2\lceil v_p(D)/2\rceil}$ be its smallest square multiple. Assume that $C < H^{1+\delta}$. The number of reciprocal integer polynomials f of height $\le H$ for which $G_f \subseteq G_2$ and $D \mid \mathrm{Disc}\, K_g$ is*

$$\ll \frac{O(1)^{\omega(C)} H^{n+1}}{D'^2}.$$

Remark 5.8. Note that in contrast to the notation in the rest of the paper, we do *not* set $D = \mathrm{Disc}\, K_g$, but rather assume only that $D \mid \mathrm{Disc}\, K_g$. For Case I, we set $D = \mathrm{Disc}\, K_g$, but we will reuse this lemma in Case III, and there we will pick a general divisor D.

Note that Lemma 5.7 implies the bound of Theorem 5.1 on the number of g in Case I because each D' determines D up to $O(1)^{\omega(C)}$ possibilities, and the total number of g is thus

$$\ll O(1)^{\omega(C)} H^{n+1} \sum_{D' \ge H^{1+\delta}} \frac{1}{D'^2} \ll O(1)^{\omega(C)} H^{n-\delta} \ll H^{n-\delta+\varepsilon} \ll H^n.$$

Proof of Lemma 5.7. First of all, we divide out all primes $p \le n$ from D. If there is at least one K_g with $D \mid \mathrm{disc}\, K_g$, this change only affects D by a bounded factor, since $v_p(\mathrm{disc}\, K_g)$ is uniformly bounded. Thus we can assume that every prime $p \mid C$ is at least n.

For each prime $p \mid C$, excluding the degenerate case that $g \equiv 0 \bmod p$ (see below), let $\sigma_p = (f_1^{e_1} \cdots f_r^{e_r})$ be the splitting type of its homogenization $\tilde{g} = y^n g(x/y)$. By the second inequality in (10),

$$v_p(D) \le \mathrm{ind}(\tilde{g} \bmod p).$$

Also, if $v_p(D)$ is odd, the relations

$$1 \equiv v_p(D) \equiv v_p(\mathrm{disc}\, g) \equiv v_p g(2) g(-2) \bmod 2$$

imply that $g(u)$ mod p is divisible by either $u-2$ or $u+2$. Let $k_p = v_p(D') = \lceil v_p(D)/2 \rceil$. We thus get one of three cases restricting σ_p and $\tilde{g}$:

(0) $\operatorname{ind}(\sigma_p) \geq 2k_p$;

(2) $\operatorname{ind}(\sigma_p) \geq 2k_p - 1$ and $f_1 = 1$ with $P_1 = x - 2y$;

(−2) $\operatorname{ind}(\sigma_p) \geq 2k_p - 1$ and $f_1 = 1$ with $P_1 = x + 2y$.

We call an *annotated splitting type* a pair (σ, j) of a splitting type $\sigma = \sigma_p$ with a choice of case $j \in \{0, 2, -2\}$. Set

$$w_{p,\sigma,j}(h) = \begin{cases} w_{p,\sigma}\big(y^n h(x/y)\big) & \text{if } j = 0, \\ w'_{p,\sigma}\big(x^n h(y/x - j)\big) & \text{if } j = \pm 2, \end{cases}$$

the coordinate changes in the latter two cases being to send the distinguished linear factors $u - j$ in our situation to the distinguished linear factor y in Lemma 5.6. Likewise, let

$$\#\operatorname{Aut}(\sigma, j) = \begin{cases} \#\operatorname{Aut}(\sigma) & \text{if } j = 0, \\ \#\operatorname{Aut}'(\sigma) & \text{if } j = \pm 2. \end{cases}$$

For (σ_p, j_p) an annotated splitting type, let $\Psi_p = w_{p,\sigma_p,j_p} : V_{\mathbb{F}_p} \to \mathbb{R}$, and let $\Psi : V_{\mathbb{Z}/C\mathbb{Z}} \to \mathbb{R}$ be the product of the Ψ_p. Observe that there are only $O(1)^{\omega(C)}$ choices of the annotated splitting types for all p, and for each g as in the statement of the lemma, there is a choice for which $\Psi(g) \geq 1$. (In the degenerate case that $g \equiv 0$ mod p, we pick σ_p and j_p arbitrarily, because $w_{p,\sigma,j}(0) \geq 1$ for all σ and j.) Thus it suffices to prove, for a fixed choice of σ_p and case (a)–(c) for each $p \mid C$, that

$$\sum_{\operatorname{Ht} g \leq H} \Psi(g) \ll \frac{O(1)^{\omega(C)} H^{n+1}}{D'^2}.$$

We claim that there is a decomposition $\Psi_p = \Lambda_p + \Delta_p$ with the following properties:

- $\Lambda_p = a_p \mathbf{1}_{L_p}$ is a rescaled characteristic function of a sublattice $L_p \subseteq V(\mathbb{Z})$, with $0 < a_p \leq 1$, and $\widehat{\Lambda}_p = \hat{a}_p \mathbf{1}_{L_p^\perp}$ where $0 < \hat{a}_p \leq p^{-2k_p}$;
- Δ_p has uniformly small Fourier transform: $\widehat{\Delta}_p(h) \ll p^{-2k_p - 1}$.

Indeed, we take

$$L_p = \begin{cases} V(\mathbb{Z}) & \text{if } j_p = 0, \\ \big\{g : g(j_p) \equiv g'(j_p) \equiv \cdots \equiv g^{(e_{1,p}-1)}(j_p) \equiv 0 \text{ mod } p\big\} & \text{if } j_p = \pm 2. \end{cases}$$

Note that in the latter case, L_p is the image of $V^{\text{hom}}_{e_1, \mathbb{F}_p}$ under the relevant identification of V^{hom} with V sending $x^i y^{n-i}$ to $(u - j)^{n-i}$. Choose the constant

$$a_p = \frac{[V(\mathbb{Z}) : L_p] \cdot p^{-\operatorname{ind}\sigma - |j|/2}}{\#\operatorname{Aut}(\sigma, j)} = \frac{p^{-\sum_{i > |j|/2} (e_i - 1) f_i}}{\#\operatorname{Aut}(\sigma, j)} \leq 1$$

so that $\widehat{\Lambda}_p$ is the dominant term

$$\hat{a}_p \mathbf{1}_{L_p^\perp}, \quad \hat{a}_p = \frac{p^{-\operatorname{ind}\sigma - |j|/2}}{\#\mathrm{Aut}(\sigma, j)} \le p^{-2k_p}$$

of $\widehat{\Psi}_p$ as computed in Lemmas 5.5 and 5.6, which show the smallness of $\widehat{\Delta}_p$.

Now

$$\Psi = \prod_p (\Lambda_p + \Delta_p) = \sum_{q \mid C} \Lambda_q \Delta_{C/q},$$

where we set

$$\Lambda_q = \prod_{p \mid q} \Lambda_p = a_q \mathbf{1}_{L_q}, \quad a_q = \prod_{p \mid q} a_p, \quad L_q = \bigcap_{p \mid q} L_p, \quad \text{and} \quad \Delta_q = \prod_{p \mid q} \Delta_p.$$

We now use Fourier analysis, as in Case I of [8], to rewrite the desired count of g. Let $\phi \colon \mathbb{R} \to \mathbb{R}$ be a Schwartz function on $\mathbb{R}$ with the following properties:

(a) $\phi(u)$ is nonnegative, and $\phi(u) \ge 1$ for $|u| \le 1$;

(b) ϕ is compactly supported;

(c) The Fourier transform $\widehat{\phi}$ is real and nonnegative.

Such a ϕ can be constructed, for instance, by taking a usual even "bump function" and convolving with itself, which squares the Fourier transform, ensuring nonnegativity. Then let $\Phi \colon V(\mathbb{R}) \to \mathbb{R}$ be the product $\Phi(g) = \phi(b_0)\phi(b_1) \cdots \phi(b_n)$. We use this as a smoothing factor:

$$\begin{aligned} X &:= \#\{g \in V(\mathbb{Z}) : G_f \subseteq G_2,\, D \mid \operatorname{disc} K_g,\, \operatorname{Ht} g \le H\} \\ &\le \sum_{\operatorname{Ht} g \le H} \Psi(g) \le \sum_{g \in V(\mathbb{Z})} \Psi(g) \Phi\left(\frac{g}{H}\right) = \sum_{q \mid C} \sum_{g \in V(\mathbb{Z})} \Lambda_q(g) \Delta_{C/q}(g) \Phi\left(\frac{g}{H}\right). \end{aligned}$$

Since C has $O(1)^{\omega(C)}$ divisors q, it suffices to prove that for all $q \mid C$,

$$X_q := \sum_{g \in V(\mathbb{Z})} \Lambda_q(g) \Delta_{C/q}(g) \Phi\left(\frac{g}{H}\right) \ll \frac{O(1)^{\omega(C)} H^{n+1}}{D'^2}.$$

We apply twisted Poisson summation (Proposition 5.2) with modulus $M = C/q$, which is coprime to $[V(\mathbb{Z}) : L_q] \mid q^n$, to get

$$\begin{aligned} X_q &= a_q \sum_{g \in L_q} \Delta_{C/q}(g) \Phi\left(\frac{g}{H}\right) = \frac{a_q H^{n+1}}{[V(\mathbb{Z}) : L_q]} \sum_{h \in L_q^*} \widehat{\Delta}_{C/q}(h) \widehat{\Phi}\left(\frac{qHh}{C}\right) \\ &\ll \frac{a_q H^{n+1}}{[V(\mathbb{Z}) : L_q]} \prod_{p \mid C/q} O(p^{-2k_p - 1}) \sum_{h \in L_q^*} \widehat{\Phi}\left(\frac{qHh}{C}\right). \end{aligned}$$

We apply Poisson summation again, now untwisted, to get

$$\begin{aligned} X_q &\ll \frac{a_q C^{n+1}}{q^{n+1}} \prod_{p \mid C/q} O(p^{-2k_p-1}) \sum_{g \in L_q} \Phi\Big(\frac{Cg}{qH}\Big) \\ &\le \frac{a_q C^{n+1}}{q^{n+1}} \prod_{p \mid C/q} O(p^{-2k_p-1}) \sum_{g \in V(\mathbb{Z})} \Phi\Big(\frac{Cg}{qH}\Big) \\ &\ll a_q \prod_{p \mid C/q} O(p^{-2k_p-1}) \left(\frac{C}{q} \max\left\{\frac{qH}{C}, 1\right\}\right)^{n+1} \\ &\ll O(1)^{\omega(C)} \prod_{p \mid q} p^{-2k_p} \prod_{p \mid C/q} O(p^{-2k_p-1}) \cdot \max\Big\{H, \frac{C}{q}\Big\}^{n+1} \\ &= \frac{O(1)^{\omega(C)} q}{C D'^2} \max\Big\{H, \frac{C}{q}\Big\}^{n+1}. \end{aligned}$$

If the first argument to the maximum dominates, we get a bound

$$X_q \ll \frac{O(1)^{\omega(C)} q}{C D'^2} H^{n+1} \le \frac{O(1)^{\omega(C)} H^{n+1}}{D'^2},$$

as desired. If instead the second argument dominates, we get a bound

$$X_q \ll \frac{O(1)^{\omega(C)} q}{C D'^2} \left(\frac{C}{q}\right)^{n+1} = \frac{O(1)^{\omega(C)} C^n}{q^n D'^2} \le \frac{O(1)^{\omega(C)} H^{n+n\delta}}{D'^2},$$

as desired, since $n\delta \le 1$. □

5.4. *Case II: $D \le H^{2+2\delta}$.* Here we simply invoke Case II of Bhargava's treatment [8, §5], which shows that the number of irreducible g of height $< H$ defining a number field K_g of primitive Galois group G_g and discriminant $D \le H^{2+2\delta}$ is $O(H^n)$ (or $O(H^{n-1})$ in the monic case). In our setting we are only concerned with the case $G_g = S_n$. We do not need to use the added knowledge that $G_f \subseteq G_2$.

5.5. *Case III: $C \ge H^{1+\delta}$.* Here we adapt the method of Bhargava using the double discriminant.

Lemma 5.9 [8, Proposition 34]. *Let $p > 2$ be a prime. If $h(x_1, \dots, x_n)$ is an integer polynomial, such that $h(c_1, \dots, c_n)$ is a multiple of p^2 for mod p reasons, that is, $h(c_1 + pd_1, \dots, c_n + pd_n)$ is a multiple of p^2 for all $(d_1, \dots, d_n) \in \mathbb{Z}^n$, then $\frac{\partial}{\partial x_n} h(c_1, \dots, c_n)$ is a multiple of p.*

Let

$$h = g(2)g(-2)\operatorname{disc} g,$$

considered as a polynomial in the coefficients $b_0, \dots, b_n$ of g. Let $p \mid C$, $p > n$. We claim that $p^2 \mid h$ for mod p reasons. We have that $p \mid D$, and p^2 divides the square $(-1)^n D g(2) g(-2)$. If $p^2 \mid D$, then by the first inequality in (10), the index of g modulo p is at least 2. So by the second inequality on that same line, we have $p^2 \mid \operatorname{disc} g$ for mod p reasons. Otherwise, we have $p \mid D$, so $p \mid \operatorname{disc} g$, and either $p \mid g(2)$ or $p \mid g(-2)$. Thus in all cases the product h is divisible by p^2 for mod p reasons. By Lemma 5.9, this implies that the derivative $\frac{\partial}{\partial b_0}h$ with respect to the constant term is divisible by p. Hence their resultant

$$R(b_1, \dots, b_n) = \mathrm{Res}_{b_0}\left(h, \frac{\partial}{\partial b_0}h\right) = \pm b_n \operatorname{disc}_{b_0} h$$

is a multiple of p.

The polynomial R is the analogue in our setting of the *double discriminant* DD of [8, Proposition 34]. For brevity, let $G = \operatorname{disc}_u g$. Then

$$\begin{aligned} R(b_1, \dots, b_n) &= \pm b_n \operatorname{disc}_{b_0}(g(2)g(-2)G) \\ &= \pm b_n \operatorname{disc}_{b_0} G \cdot \mathrm{Res}_{b_0}(g(2), g(-2))^2 \,\mathrm{Res}_{b_0}(g(2), G)^2 \,\mathrm{Res}_{b_0}(g(-2), G)^2 \\ &= \pm b_n \operatorname{disc}_{b_0} G \cdot (g(2) - g(-2))^2 \big(\operatorname{disc}_u(g - g(2))\big)^2 \big(\operatorname{disc}_u(g - g(-2))\big)^2, \end{aligned} \tag{11}$$

where, in the last step, we took advantage of the linearity of $g(\pm 2)$ in b_0 to use the standard formula

$$\mathrm{Res}_x(x - a, P(x)) = P(a).$$

Thus R is a product of which one factor is the double discriminant $\mathrm{DD}(g) = \operatorname{disc}_{b_0} G = \operatorname{disc}_{b_0} \operatorname{disc}_u g$. One easily sees from the factorization (11) that R is not identically zero as a function of $b_1, \dots, b_n$.

We now proceed as in [8]. The number of $b_1, \dots, b_n \in [-H, H]^n$ such that $R(b_1, \dots, b_n) = 0$ is $O(H^{n-1})$ (by, e.g., [6, Lemma 3.1]), and so the number of g with such $b_1, \dots, b_n$ is $O(H^n)$. We now fix $b_1, \dots, b_n$ such that $R(b_1, \dots, b_n) \neq 0$. Then $R(b_1, \dots, b_n)$ has at most $O_\varepsilon(H^\varepsilon)$ factors $C > H$. Once C is determined by $b_1, \dots, b_n$ (up to $O_\varepsilon(H^\varepsilon)$ possibilities), then the number of solutions for b_0 (mod C) to $h \equiv 0 \bmod C$ is $(\deg_{b_0} h)^{\omega(C)} = O_\varepsilon(H^\varepsilon)$, as the number of possibilities for b_0 modulo p such that $h \equiv 0 \bmod p$ for each $p \mid C$ is at most $\deg_{b_0}(h)$. Since $C > H$, the number of possibilities for $b_0 \in [-H, H]$ is also at most $O_\varepsilon(H^\varepsilon)$, and so the total number of g in this case is $O_\varepsilon(H^{n+\varepsilon})$.

To eliminate the ε, we divide into two subcases, the first of which is reduced to Case I, just as in [8]:

Subcase (i): $A = \prod_{\substack{p \mid C \\ p > H^{\delta/2}}} p \leq H$.

In this subcase, C has a factor C_{I} between $H^{1+\delta/2}$ and $H^{1+\delta}$, with $A \mid C_{\mathrm{I}} \mid C$. Pick C_{I} to be the largest such factor, and let $D_{\mathrm{I}} = \prod_{p \mid C_{\mathrm{I}}} p^{v_p(D)}$. We now appeal to Lemma 5.7 from Case I, with C_{I} in place of C, and D_{I} in place of D. (Note that D determines C, C_{I}, and D_{I}.) We get that the number of g with a given D is at most

$$\ll \frac{O(1)^{\omega(C_{\mathrm{I}})} H^{n+1}}{D_{\mathrm{I}}'^2},$$

where

$$D_{\mathrm{I}}' = \prod_{p \mid C_{\mathrm{I}}} p^{\lceil v_p(D)/2 \rceil} \geq C_{\mathrm{I}} \geq H^{1+\delta/2}.$$

For each D_{I}', there are at most $2^{\omega(D)}$ values of D_{I}, so the total number of reciprocal polynomials in this subcase is

$$\sum_{D_1' > H^{1+\delta/2}} \frac{O(1)^{\omega(C_1)} H^{n+1}}{D_1'^2} = O_\varepsilon(H^{n-\delta/2+\varepsilon}) = O(H^n).$$

Subcase (ii): $A = \prod_{\substack{p \mid C \\ p > H^{\delta/2}}} p > H.$

In this subcase, we carry out the original argument of Case III, with C replaced by A. We have $A \mid R(b_1, \dots, b_n)$.

Fix one of the $O(H^n)$ choices of $b_1, \dots, b_n$ such that $R(b_1, \dots, b_n) \neq 0$. Being bounded above by a fixed power of H, we see that $R(b_1, \dots, b_n)$ can have at most a bounded number of possibilities for the factor A (since all prime factors of A are bounded below by a fixed positive power of H). Once A is determined by $b_1, \dots, b_n$, then the number of solutions for $b_0 \pmod{A}$ to $\operatorname{Disc}(f) \equiv 0 \pmod{A}$ is $O(n^{\omega(A)}) = O(1)$. Since $A > H$, the total number of f in this subcase is also $O(H^n)$, completing the proof of Theorem 5.1 in the nonmonic case.

5.6. *Remarks on the monic case.* It would be impractical to replicate the full proof of Theorem 5.1 in the monic case, most of which consists of changing n to $n-1$ and correcting calculations appropriately. The following changes are worthy of note:

- Unlike in the nonmonic case, the space $V(R)$ of monic polynomials over a ring does not have a natural origin. We must fix one, such as $g(u) = u^n$, in order to carry out the Fourier analysis.
- In Case I, we replace Lemma 5.5 with the following lemma from Bhargava:

Lemma 5.10 [8, Proposition 30]. *Let $\sigma = (f_1^{e_1} \cdots f_r^{e_r})$ be a splitting type with $\deg(\sigma) = d \leq n$ and $\operatorname{ind}(\sigma) = k$. Let $w_{p,\sigma} : V(\mathbb{F}_p) \to \mathbb{C}$ be defined by*

$w_{p,\sigma}(f) :=$ *the number of r-tuples $(P_1, \dots, P_r)$, up to the action of the group of permutations of $\{1, \dots, r\}$ preserving σ, such that the P_i are distinct irreducible monic polynomials with $\deg P_i = f_i$ for each i and $P_1^{e_1} \cdots P_r^{e_r} \mid f$.*

Then

$$\widehat{w}_{p,\sigma}(g) = \begin{cases} \dfrac{p^{-k}}{\#\operatorname{Aut}(\sigma)} + O(p^{-(k+1)}) & \text{if } g = 0, \\ O(p^{-(k+1)}) & \text{if } g \neq 0 \text{ and } d < n, \\ O(p^{-(k+1/2)}) & \text{if } g \neq 0 \text{ and } d = n. \end{cases}$$

- We then replace Lemma 5.6 as follows:

Lemma 5.11. *Let $\sigma = (f_1^{e_1} \cdots f_r^{e_r})$ be a splitting type with $f_1 = 1$. Let $\deg(\sigma) = d$ and $\operatorname{ind}(\sigma) = k$. Let $w'_{p,\sigma} : V(\mathbb{F}_p) \to \mathbb{C}$ be defined by*

$w_{p,\sigma}(f) :=$ *the number of r-tuples $(P_1, \dots, P_r)$, up to the action of the group of permutations of $\{2, \dots, r\}$ preserving σ, such that the P_i are distinct irreducible monic polynomials with $\deg P_i = f_i$ for each i, $f_1 = x$, and $P_1^{e_1} \cdots P_r^{e_r} \mid f$.*

Let $V_{e_1,\mathbb{F}_p}$ denote the subspace of $V(\mathbb{F}_p)$ comprising polynomials divisible by x^{e_1}, and let $V_{e_1,\mathbb{F}_p}^{\perp} \subseteq V(\mathbb{F}_p)^*$ be its dual. Then

$$\widehat{w}_{p,\sigma}(g) = \begin{cases} \dfrac{p^{-(k+1)}}{\#\mathrm{Aut}'(\sigma)} + O(p^{-(k+2)}) & \text{if } g \in V_{e_1,\mathbb{F}_p}^{\perp}, \\ O(p^{-(k+2)}) & \text{if } g \notin V_{e_1,\mathbb{F}_p}^{\perp} \text{ and } d < n, \\ O(p^{-(k+3/2)}) & \text{if } g \notin V_{e_1,\mathbb{F}_p}^{\perp} \text{ and } d = n. \end{cases}$$

- Our Ψ_p is then supported on

$$\widetilde{L}_p = \begin{cases} V(\mathbb{Z}) & \text{in case (0) on page 618}, \\ \{g : g(2) \equiv g'(2) \equiv \cdots \equiv g^{(e_{1,p}-1)}(2) \equiv 0 \bmod p\} & \text{in case (2)}, \\ \{g : g(-2) \equiv g'(-2) \equiv \cdots \equiv g^{(e_{1,p}-1)}(-2) \equiv 0 \bmod p\} & \text{in case (−2)}, \end{cases}$$

which is no longer a lattice but a coset $g_p + L_p$ for some fixed monic polynomial g_p and some lattice L_p of index dividing p^n. The Fourier transform $\widehat{\Psi}$ is small away from $L_p^{\perp}$. The intersection

$$\widetilde{L}_q = \bigcap_{p \mid q} \widetilde{L}_p = g_q + L_q$$

is likewise a coset of a lattice (if nonempty). When we carry out the twisted Poisson summation, the translation g_q contributes a twist factor $e^{-2\pi\sqrt{-1}g_q \bullet h}$ to the values of $\widehat{\Lambda}_q$. Because we then immediately bound each term by its absolute value, this twist factor drops out.

- The final step of Case I, bounding X_q, proceeds as follows:

$$X_q \ll \prod_{p \mid q} O(p^{-2k_p}) \prod_{p \mid \frac{C}{q}} O(p^{-2k_p - 1/2}) \cdot \max\left\{H, \frac{C}{q}\right\}^n = \frac{O(1)^{\omega(C)}\sqrt{q}}{\sqrt{C}\,D'^2} \max\left\{H, \frac{C}{q}\right\}^n$$
$$= \frac{O(1)^{\omega(C)}}{D'^2} \max\left\{\frac{H^n\sqrt{q}}{\sqrt{C}}, \left(\frac{C}{q}\right)^{n-1/2}\right\} \leq \frac{O(1)^{\omega(C)}}{D'^2} \max\left\{H^n, H^{n-1/2+n\delta}\right\} = \frac{O(1)^{\omega(C)} H^n}{D'^2}.$$

6. Counting G_3-polynomials

Finally, we count the polynomials f having $G_f \subseteq G_3$ for each odd $n \geq 3$ (the even case is subsumed by G_1, as we noted in stating Theorem 3.4).

Theorem 6.1. *For $n \geq 3$ odd,*

$$\mathcal{E}_n(G_3; H) \ll \begin{cases} H^2 \log^2 H & \text{if } n = 3, \\ H^{(n+1)/2} & \text{if } n \geq 5, \end{cases} \tag{12}$$

$$\mathcal{E}_n^{\text{monic}}(G_3; H) \ll \begin{cases} H^2 & \text{if } n = 3, \\ H^2 \log H \log\log H & \text{if } n = 5, \\ H^{(n-1)/2} \log H & \text{if } n \geq 7. \end{cases} \tag{13}$$

6.1. *Heights.* We need a notion of height more general than the naïve height on integer polynomials used in Section 5.

If $P = [x_0 : \cdots : x_N] \in \mathbb{P}^N(K)$ is a point in projective space over a number field K, we define its (exponential, projective, Weil) *height* (denoted $H(P)$ in Silverman [22, §VIII.5]) as

$$\operatorname{Htp} P = \prod_v \max\left\{|x_0|_v, \ldots, |x_N|_v\right\}^{[K_v:\mathbb{Q}_v]/[K:\mathbb{Q}]},$$

where the product is taken over the places v of K, and the norm $|x|_v$ extends the usual v-adic norm on $\mathbb{Q}$. This normalization ensures that the height is unchanged if K is embedded into a larger field. There are two natural ways to define a height on a polynomial $f(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_0$ over a number field, and we distinguish the *projective* and *affine* heights

$$\operatorname{Htp} f = \operatorname{Htp}\,[a_n : \cdots : a_0], \tag{14}$$

$$\operatorname{Ht} f = \operatorname{Htp}\,[a_n : \cdots : a_0 : 1]. \tag{15}$$

For instance, if $f \in \mathbb{Z}[x]$ is nonzero, then $\operatorname{Ht} f = \max\{|a_0|, \ldots, |a_n|\}$ is the naïve height already introduced in Section 2, while $\operatorname{Htp} f = \operatorname{Ht} f / \operatorname{ct} f$ is smaller by a factor of the *content* $\operatorname{ct}(f) = \gcd(a_0, \ldots, a_n)$. In particular, we define heights of algebraic numbers α by $\operatorname{Ht}\alpha = \operatorname{Htp}\,[\alpha : 1]$. The following properties should be noted:

- If α is an algebraic integer with conjugates $\alpha = \alpha_1, \ldots, \alpha_n$, then

$$\operatorname{Htp}\alpha = \prod_i \max\{1, \alpha_i\}^{1/n}$$

  is none other than the *Mahler measure* of its minimal polynomial (suitably normalized).

- If f has roots $\alpha_1, \ldots, \alpha_n$, then

$$2^{-n}\prod_i \operatorname{Ht}\alpha_i \le \operatorname{Htp} f \le 2^{n-1}\prod_i \operatorname{Ht}\alpha_i \tag{16}$$

  [22, Theorem VIII.5.9]. In particular, for any factorization $f = gh$, we have

$$\operatorname{Htp} f \asymp_{\deg f} \operatorname{Htp} g \cdot \operatorname{Htp} h. \tag{17}$$

6.2. *Proof of Theorem 6.1.*

Proof. Under the conditions, there are only two possibilities for G_f: either $G_f \cong G_3 \cong S_2 \times S_n$, or $G_f \cong S_n$. In the latter case, f is reducible and factors as a product

$$f(x) = c \cdot h(x) \cdot x^n h(1/x), \quad c \in \mathbb{Z}, \quad h \in \mathbb{Z}[x].$$

By taking $|c| = \operatorname{ct} f$, we may assume that $\operatorname{ct} h = 1$. We have $\operatorname{Htp} f = \operatorname{Ht} f/|c| \le H/|c|$, so by (17),

$$\operatorname{Ht} h = \operatorname{Htp} h \ll \sqrt{\frac{H}{|c|}}.$$

As h has $n+1$ free coefficients, the number of polynomials f we get is

$$\ll \sum_{c=1}^{H} \left(\frac{H}{c}\right)^{(n+1)/2} \ll H^{(n+1)/2}.$$

We now assume that $G_f = G_3$. In this case there is a quadratic field $K_2 = \mathbb{Q}(\sqrt{k})$, where k is a squarefree integer, such that $K_f = K_g \cdot K_2$, so that over K_2, the polynomial f factorizes:

$$f(x) = c' \cdot h(x) \cdot x^n h(1/x), \quad c' \in K_2, \quad h \in K_2[x]. \tag{18}$$

We cannot assume that this factorization holds in $\mathcal{O}_{K_2}[x]$ because $\mathcal{O}_{K_2}$ need not be a UFD. However, we can rescale h so that $h(1) \in \mathbb{Z}$. Then $c' = f(1)/h(1)^2 \in \mathbb{Q}$, and the contents $c = \mathrm{ct}(f)$ and $\mathfrak{a} = \mathrm{ct}(h)$ satisfy

$$c\mathcal{O}_{K_2} = c' \cdot \mathfrak{a}^2.$$

Therefore $\mathfrak{a}^2 = (c/c')$ is principal, generated by a rational number; in particular, $\mathfrak{a} = \bar{\mathfrak{a}}$ is self-conjugate. The self-conjugate ideals in a quadratic field are simply the rational multiples of products of ramified primes. Thus, after rescaling by a rational scalar, we have that there is a positive divisor $d \mid k$ such that

$$\mathfrak{a} = \sqrt{(d)} = \begin{cases} \langle d, \sqrt{k} \rangle & \text{if } k \equiv 2, 3 \bmod 4 \\ \left\langle d, \frac{1}{2}(k + \sqrt{k}) \right\rangle & \text{if } k \equiv 1 \bmod 4, \end{cases}$$

and $c' = \pm c/d$. Note that $\bar{h}(x)$ and $x^n h(1/x)$ have the same roots, and comparing values at $x = 1$, they must be equal. Thus the coefficients $\theta_i \in K_2$ of $h(x) = \sum_{i=0}^n \theta_i x^i$ satisfy

$$\theta_i = \overline{\theta_{n-i}}. \tag{19}$$

We now bound the height of the θ_i. This requires us to control the projective-to-affine height ratio

$$\frac{\mathrm{Htp}\, h}{\mathrm{Ht}\, h} = \prod_v \left(\frac{\max\{|\theta_0|_v, \ldots, |\theta_n|_v\}}{\max\{|\theta_0|_v, \ldots, |\theta_n|_v, 1\}} \right)^{[(K_2)_v:\mathbb{Q}_v]/[K_2:\mathbb{Q}]}.$$

We consider the contributions of the different places v in turn:

- If $v = p \mid d$ is finite, then the numerator is $|\mathfrak{a}|_p = 1/\sqrt{p}$, while the denominator is 1. Since $[(K_2)_v : \mathbb{Q}_v] = [K_2 : \mathbb{Q}] = 2$, we get a contribution $1/\sqrt{p}$ in this case.
- For all other finite v, the numerator and denominator are both 1.
- If v is infinite, the parenthesized fraction is 1 because, by (19),

$$|\theta_0|_v \cdot |\theta_n|_v = |\theta_0|_v \cdot |\overline{\theta_0}|_v = d \cdot \frac{|f(0)|}{c} \geq 1.$$

Hence

$$\mathrm{Ht}\, h = \sqrt{d}\, \mathrm{Htp}\, h \asymp \sqrt{d\, \mathrm{Htp}\, f} = \sqrt{\frac{d\, \mathrm{Ht}\, f}{c}} \leq \sqrt{\frac{dH}{c}}.$$

Write

$$\theta_i = \frac{du_i + v_i\sqrt{k}}{2}, \quad u_i, v_i \in \mathbb{Z},$$

and observe that not *all* the u_i nor *all* the v_i can be zero, or f would be reducible. Now some coefficient du_i of $h + \bar{h}$ is nonzero and thus at least d in absolute value. So

$$d \le \operatorname{Ht}(h + \bar{h}) \ll \operatorname{Ht}(h) \ll \sqrt{\frac{dH}{c}},$$

which gives us the bound

$$d \ll \frac{H}{c}.$$

Analogously, analyzing the v_i by considering $h - \bar{h}$ gives us the bound

$$d' := \frac{|k|}{d} \ll \frac{H}{c}.$$

Since du_i and $v_i\sqrt{|k|}$ are bounded by $\operatorname{Ht} h$, the numbers of possibilities for each u_i and v_i are $\ll \sqrt{H/cd}$ and $\ll \sqrt{H/cd'}$ respectively, and so the total number of polynomials is

$$\begin{aligned} E_n(G_3; H) &\ll \sum_{c=1}^{H} \sum_{d \ll H/c} \sum_{d' \ll H/c} \left(\sqrt{\frac{H}{cd}}\sqrt{\frac{H}{cd'}}\right)^{\frac{n+1}{2}} \\ &= H^{\frac{n+1}{2}} \sum_{c=1}^{H} c^{-\frac{n+1}{2}} \biggl(\sum_{d \ll H/c} d^{-\frac{n+1}{4}}\biggr)^2. \end{aligned}$$

If $n \ge 5$, the two remaining sums are both $O(1)$, and we get a bound of $O(H^{\frac{n+1}{2}})$. If $n = 3$, we get

$$E_n(G_3; H) \ll H^{\frac{n+1}{2}} \sum_{c=1}^{H} c^{-\frac{n+1}{2}} \log^2 H \ll H^{\frac{n+1}{2}} \log^2 H = H^2 \log^2 H. \qquad \square$$

Remark 6.2. The "$\ll$" in Theorem 6.1 can be sharpened to "$\asymp$" in the nonmonic case by checking that a positive proportion of choices of the independent variables c, d, d', u_i, v_i satisfy the needed conditions:

- d and d' are squarefree and coprime;
- u_i and v_i satisfy the conditions mod 2 so that $\theta_i \in \mathfrak{a}$;
- and the θ_i do not all lie in an ideal strictly smaller than $\mathfrak{a}$.

All these can be managed using an appropriate form of the squarefree sieve. We omit the details.

6.3. *Remarks on the monic case.* The monic case is proved similarly, but the possibilities are more restricted because we can factor $f(x) = \pm h_1(x) \cdot x^n h_1(1/x)$ over $\mathcal{O}_{K_2}$, where $h_1(x)$ is monic. This scaling h_1 of h may or may not be compatible with the one used above, but since $\operatorname{ct}(h_1) = (1)$, we derive that $\mathfrak{a} = (\theta)$ is principal. Since $\mathfrak{a}$ is known to be self-conjugate, we have a relation

$$\theta = \varepsilon\bar{\theta}, \tag{20}$$

where $\varepsilon \in \mathcal{O}_{K_2}^\times$ is a unit (necessarily of norm 1). The relation (20) determines θ up to scaling by $\mathbb{Q}^\times$; namely

$$\theta = w(1+\varepsilon) \quad \text{and/or} \quad \theta = w\sqrt{k}(1-\varepsilon), \quad w \in \mathbb{Q}^\times,$$

the expression "and/or" being used because these formulas become invalid if $\varepsilon = -1$, respectively $\varepsilon = 1$. Thus ε determines d. Since θ can be rescaled by a unit, it follows that rescaling ε by a square of a unit does not affect d. Since $|\mathcal{O}_{K_2}^\times/(\mathcal{O}_{K_2}^\times)^2|$ is uniformly bounded (at most 4), there are $O(1)$ possible values of d for each k. Each of the coefficients $\theta_1, \dots, \theta_{(n-1)/2}$ has $O(H/\sqrt{k})$ values, as above. As to θ_0, we have that $\eta = \theta_0^2/d$ is a unit and determines θ_0 up to sign. We have $\operatorname{Ht}\theta_0 \ll \sqrt{H/d}$ so $\operatorname{Ht}\eta \ll H$. Up to the finite group of roots of unity, η is a power η_1^m of the fundamental unit η_1 of K_2. We have $|\eta_1| \gg \sqrt{k}$, so the number of possibilities for η is $\ll \log H/\log k$, for a grand total of

$$\mathcal{E}_n^{\text{monic}}(G_3;\, H) \ll \sum_{2\le k\ll H^2} \left(\frac{H}{\sqrt{k}}\right)^{\frac{n-1}{2}} \cdot \frac{\log H}{\log k}.$$

Estimating this sum yields the claimed bounds. In contrast to the nonmonic case, it is unclear whether our bounds are sharp, specifically in the $n = 3$ and $n = 5$ cases. A more accurate count of monic G_3-polynomials demands a delicate understanding of the distribution of sizes of fundamental units in real quadratic fields.

Acknowledgements

Anderson was partially supported by the National Science Foundation (grants 2231990, 2237937). We thank Hongyi (Brian) Hu and the rest of the participants in the Research Topics in Mathematical Sciences semester at CMU in spring 2024 for their contributions to the discussion. We thank Arno Fehm, Benedict Gross, Igor Shparlinski, and the referees for insightful comments. We thank Robert Lemke Oliver for a careful read of an earlier draft.

References

[1] T. C. Anderson, A. Gafni, R. J. Lemke Oliver, D. Lowry-Duda, G. Shakan, and R. Zhang, "Quantitative Hilbert irreducibility and almost prime values of polynomial discriminants", *Int. Math. Res. Not.* **2023**:3 (2023), 2188–2214. MR

[2] L. Bary-Soroker and A. Fehm, "Correlations of sums of two squares and other arithmetic functions in function fields", *Int. Math. Res. Not.* **2019**:14 (2019), 4469–4515. MR

[3] L. Bary-Soroker and G. Kozma, "Irreducible polynomials of bounded height", *Duke Math. J.* **169**:4 (2020), 579–598. MR

[4] L. Bary-Soroker, D. Koukoulopoulos, and G. Kozma, "Irreducibility of random polynomials: general measures", *Invent. Math.* **233**:3 (2023), 1041–1120. MR

[5] L. Bary-Soroker, O. Ben-Porath, and V. Matei, "Probabilistic Galois theory: the square discriminant case", *Bull. Lond. Math. Soc.* **56**:6 (2024), 2162–2177. MR

[6] M. Bhargava, "The geometric sieve and the density of squarefree values of invariant polynomials", preprint, 2014. arXiv 1402.0031

[7] M. Bhargava, "A proof of van der Waerden's conjecture on random Galois groups of polynomials", *Pure Appl. Math. Q.* **19**:1 (2023), 45–60. MR

[8] M. Bhargava, "Galois groups of random integer polynomials and van der Waerden's conjecture", preprint, 2024. To appear in *Ann. of Math.* arXiv 2111.06507

[9] E. Breuillard and P. P. Varjú, "Irreducibility of random polynomials of large degree", *Acta Math.* **223**:2 (2019), 195–249. MR

[10] S. Chow and R. Dietmann, "Enumerative Galois theory for cubics and quartics", *Adv. Math.* **372** (2020), 107282, 37. MR

[11] S. Davis, W. Duke, and X. Sun, "Probabilistic Galois theory of reciprocal polynomials", *Exposition. Math.* **16**:3 (1998), 263–270. MR

[12] P. X. Gallagher, "The large sieve and probabilistic Galois theory", pp. 91–101 in *Analytic number theory* (St. Louis, MO, 1972), Proc. Sympos. Pure Math. **24**, Amer. Math. Soc., 1973. MR

[13] B. H. Gross and C. T. McMullen, "Automorphisms of even unimodular lattices and unramified Salem numbers", *J. Algebra* **257**:2 (2002), 265–290. MR

[14] D. Hokken, "Counting (skew-)reciprocal Littlewood polynomials with square discriminant", preprint, 2023. arXiv 2301.05656

[15] F. Jouve, E. Kowalski, and D. Zywina, "Splitting fields of characteristic polynomials of random elements in arithmetic groups", *Israel J. Math.* **193**:1 (2013), 263–307. MR

[16] E. Kim, T. C. Anderson, and R. J. L. Oliver, "Characterizing the support of semiclassical measures for higher-dimensional cat maps", preprint, 2024. arXiv 2410.13449

[17] H.-W. Knobloch, "Zum Hilbertschen Irreduzibilitätssatz", *Abh. Math. Sem. Univ. Hamburg* **19** (1955), 176–190. MR

[18] A. Lubotzky and L. Rosenzweig, "The Galois group of random elements of linear groups", *Amer. J. Math.* **136**:5 (2014), 1347–1383. MR

[19] S. Mattarei and M. Pizzato, "Generalizations of self-reciprocal polynomials", *Finite Fields Appl.* **48** (2017), 271–288. MR

[20] V. Naik, "Double cover of symmetric group", Groupprops article, 2013, available at https://groupprops.subwiki.org/wiki/Double_cover_of_symmetric_group. Accessed 4 May 2024.

[21] I. Rivin, "Walks on groups, counting reducible matrices, polynomials, and surface and free group automorphisms", *Duke Math. J.* **142**:2 (2008), 353–379. MR

[22] J. H. Silverman, *The arithmetic of elliptic curves*, 2nd ed., Graduate Texts in Mathematics **106**, Springer, 2009. MR

[23] B. L. van der Waerden, "Die Seltenheit der reduziblen Gleichungen und der Gleichungen mit Affekt", *Monatsh. Math. Phys.* **43**:1 (1936), 133–147. MR

[24] "Covering groups of the alternating and symmetric groups", Wikipedia page, 2024, available at http://en.wikipedia.org/wiki/Covering_groups_of_the_alternating_and_symmetric_groups. Accessed 4 May 2024.

Communicated by Melanie Matchett Wood
Received 2024-07-17 Revised 2025-01-16 Accepted 2025-03-03

tanders2@andrew.cmu.edu *Carnegie Mellon University, Pittsburgh, PA, United States*

abertell@alumni.cmu.edu *Pennsylvania State University, University Park, PA, United States*

eodorney@andrew.cmu.edu *Carnegie Mellon University, Pittsburgh, PA, United States*

Algebra & Number Theory

msp.org/ant

See inside back cover or msp.org/ant for submission instructions.

The subscription price for 2026 is US $590/year for the electronic version, and $865/year (+$75, if shipping outside the US) for print and electronic. Subscriptions, requests for back issues and changes of subscriber address should be sent to MSP.

Algebra & Number Theory (ISSN 1944-7833 electronic, 1937-0652 printed) at Mathematical Sciences Publishers, 2000 Allston Way # 59, Berkeley, CA 94701-4004, is published continuously online.

ANT peer review and production are managed by EditFLOW® from MSP.

PUBLISHED BY

mathematical sciences publishers

nonprofit scientific publishing

http://msp.org/

Algebra & Number Theory

Volume 20 No. 3 2026