Learning Seminar on Bhargava's proof of van der Waerden's Conjecture

Contents

1. Introduction – Danny Neftin (March 25)

Let $f(x) = x^n + a_1 x^{n-1} + \cdots + a_n \in \mathbb{Z}[x]$, and let $\|f\|$ denote $\max_i |a_i|$. Let $E_n(H)$ denote the number of monic integral f of degree n with Galois group $\neq S_n$ and $\|f\| \leq H$. By Hilbert's Irreducibility Theorem, we have $E_n(H) = o(H^n)$; by taking $a_0 = 0$ we see that $E_n(H) \gg H^{n-1}$. The conjecture of van der Waerden — now a theorem of Bhargava — is that this lower bound is also an upper bound.

Conjecture 1 (van der Waerden [7])**.** $E_n(H) = O(H^{n-1})$.

Let $G \subset S_n$ and let $N_n(G, H)$ be the number of monic f with Galois group (abstractly) isomorphic to G as a permutation group. Some previous results on $N_n(G, H)$: a result of Chela [2],

$$\sum_{G \leq S_n \text{ intrans.}} N_n(G, H) = c_n H^{n-1} + O(H^{n-2})$$

and a result of Widmer [8],

$$\sum_{G \leq S_n \text{ trans. and imprim.}} N_n(G, H) = O(H^{n/2+2}).$$

Date: Spring 2022.

1.1. **Statement of results.** Recall the *index* $\mathrm{Ind}(G)$ of a permutation group G is defined as

$$\mathrm{Ind}(G) = \min_{g \neq 1} (n - \#\text{orbits of } g).$$

Theorem 1 (Bhargava [1])**.** *For any permutation group $G \subset S_n$, we have*

$$N_n(G,H) = O\big(\min\{H^{n+1-\mathrm{Ind}(G)} + H^{n-(n-1)\frac{1-1/\mathrm{Ind}(G)}{a(G)+1-1/\mathrm{Ind}(G)-u}} \log^n H, \\ H^{(2n-2)(a(G)-u)+1} \log^{n-1} H\}\big)$$

where $u = 1/(n(n-1))$ if G is primitive and $u = 0$ otherwise.

Let $F_n(G,X)$ denote the number of degree n number fields whose Galois closure has Galois group isomorphic to G as a permutation group with absolute discriminant $\leq X$. The proof of Theorem 1 requires nontrivial upper bounds on $F_n(G,X)$. Recall (the weak form of) Malle's conjecture:

$$X^{1/\mathrm{Ind}(G)} \ll F_n(G,X) \ll X^{1/\mathrm{Ind}(G)+\varepsilon}.$$

Let $a(G)$ denote any constant such that $F_n(G,X) = O(X^{a(G)})$ (known to be finite by, e.g., Ellenberg–Venkatesh [3]).

Corollary. *Let $G \subset S_n$ be a permutation group.*

- *If $G \neq A_n, S_n$ is primitive non-elemental*[1] *then $N_n(G,H) = O(H^{b\sqrt{n}\log^2 n})$ for some $b > 0$.*
- *If $G \neq A_n, S_n$ is primitive elemental then $N_n(G,H) = O(H^{n-bn/\log^2 n})$ for some $b > 0$.*
- *If $G = C_n$ for n prime, then $N_n(G,H) = O(H^2)$.*
- *If G is regular (i.e. $|G| = n$) then $N_n(G,H) = O(H^{3n/11+1.164})$.*
- *Let $G \subset S_n$ be an intransitive permutation group, and write the action of G as a subdirect product $G \subset G_1 \times \cdots \times G_k$ where $G_i \subset G$ is a transitive permutation group on n_i letters, so that $n = n_1 + \cdots + n_k$. If $N_{n_i}(G_i,H) = O(H^{\alpha_i}\log^{\beta_i} H)$ then*

$$N_n(G_1 \times \cdots \times G_k, H) = O(H^{\max\{\alpha_1,\ldots,\alpha_k\}} \log^{s-1} H)$$

where $s = \sum_{\alpha_i = \max\{\alpha_1,\ldots,\alpha_k\}} (1+\beta_i)$.

Sketch of proof of Theorem 1. Let f be integral monic irreducible of degree n. Write $K_f = \mathbb{Q}[x]/(f(x))$. Let C denote the product of primes which ramify in $K_f/\mathbb{Q}$ and let D denote the absolute discriminant of K_f. Let $P(H)$ be the set of all irreducible integer monic polynomials f of degree n and $\|f\| \leq H$ and primitive Galois group $\neq S_n$. We cover $P(H)$ with three sets, depending on the magnitudes of C and D, and bound each set separately:

[1] A primitive permutation group G acting on a set Ω is called *non-elemental* if there are positive integers r, m, k (with $k < m/2$ and ($r > 1$ or $k > 1$)) such that Ω can be identified with Δ^r where Δ consists of *all* k-subsets of $[m]$, and $G \subset S_m \wr S_r$.

- $P_1(H)$ is the subset where $C \leq H^{1+\delta}$ but $D > H^{2+2\delta}$ for some $\delta > 0$.
- $P_2(H)$ is the subset where $D \leq H^{2+2\delta}$.
- $P_3(H)$ is the subset where $C > H^{1+\delta}$ (which implies $D > H^{2+2\delta}$).

We'll show $\#P_1(H) = O(H^{n-1})$. If $C < H$, then the subset of f with $\|f\| \leq H$ satisfying a set of congruence conditions $\mathcal{C}$ modulo C is $O(H^n \mu(\mathcal{C}))$ where $\mu(\mathcal{C})$ is the density of $\mathcal{C}$ (e.g. there are $O(H^n \frac{1}{C})$ polynomials with coefficients all divisible by C). We will show that the condition of an f having K_f with absolute discriminant $\geq D$ can be defined by a set of congruence conditions modulo C with density $2^{\omega(D)}/D$. Thus when $C < H$, we have that

$$\#P_1(H) = O(H^n 2^{\omega(D)}/D).$$

Fourier analysis (for finite abelian groups) will be used to prove there exists a positive δ such that this bound still holds even if $C < H^{1+\delta}$. The discriminants D are all squarefull (this will follow from G being primitive and $\neq S_n$), and summing $O(H^n 2^{\omega(D)}/D)$ over all $C < H^{1+\delta}$ and squarefull $D \geq H^{2+2\delta}$ is $O_\varepsilon(H^{n-1-\delta+\varepsilon})$, which proves the desired estimate in this case.

We'll show $\#P_2(H) = O(H^{n-1})$. A well-known bound of Schmidt shows that the number of possible K_f up to isomorphism (with C and D satisfying the assumed bounds) is at most $O(H^{(2+2\delta)(n+2)/4})$. A result of Lemke Oliver and Thorne implies that the number of possibilities of f for a given K_f is at most $O(H \log^{n-1} H / H^{1/(n(n-1))})$. Multiplying these two upper bounds together yields the desired upper bound for $\#P_2(H)$ when $n \geq 6$. For $n = 4, 5$ there are better bounds than Schmidt's for the number of K_f's, which yields the desired results in these cases also.

We'll show $\#P_3(H) = O(H^{n-1})$. In this case, a polynomial $f \in P_3(H)$ has the property that if $f \equiv g \pmod{C}$ then $\mathrm{disc}(g)$ is also a multiple of C^2 (we say that $\mathrm{disc}(f)$ *is a multiple of* C^2 *for mod* C *reasons*). Thinking of $f(x)$ as an element of $\mathbb{Q}(a_1, \ldots, a_n)[x]$, we consider the 'double discriminant' $\mathrm{disc}_{\mathbb{Q}[a_n]}\mathrm{disc}_{\mathbb{Q}[x]} f$, which is a polynomial only in $a_1, \ldots, a_{n-1}$. One can show that if $\mathrm{disc}(f)$ is a multiple of C^2 for mod C reasons then C divides $\mathrm{disc}_{\mathbb{Q}[a_n]}\mathrm{disc}_{\mathbb{Q}[x]} f$. There are

- $O(H^{n-1})$ choices for $a_1, \ldots, a_{n-1} \in [-H, H] \cap \mathbb{Z}$,
- $O_\varepsilon(H^\varepsilon)$ choices for C (for given $a_1, \ldots, a_{n-1}$) since C divides $\mathrm{disc}_{\mathbb{Q}[a_n]}\mathrm{disc}_{\mathbb{Q}[x]} f$ (now thought of as an integer where we've plugged in $a_1, \ldots, a_{n-1}$), which will not be identically zero for generic $a_1, \ldots, a_{n-1}$, and the number of divisors of any positive integer $\leq H$ is $O_\varepsilon(H^\varepsilon)$,
- $O_\varepsilon(H^\varepsilon)$ choices for a_n (for given $a_1, \ldots, a_{n-1}, C$) since $\mathrm{disc}_{\mathbb{Q}[x]} f(x) \xmapsto{a_1,\ldots,a_n}$ $\mathrm{disc} f$ (where '$\xmapsto{a_1,\ldots,a_n}$' means plug in the given $a_1, \ldots, a_n \in \mathbb{Z}$), and applying $\xmapsto{a_1,\ldots,a_{n-1}}$ to $\mathrm{disc}_{\mathbb{Q}[x]} f(x)$ results in a polynomial in a_n which vanishes modulo C when we plug in a_n, since C divides $\mathrm{disc} f \in \mathbb{Z}$ by assumption. The number of roots of this polynomial in a_n (which is nonzero for generic $a_1, \ldots, a_{n-1}$) is bounded by its degree, which is $n(n-1)/2$.

In total, we get $O_\varepsilon(H^{n-1}H^\varepsilon H^\varepsilon) = O_\varepsilon(H^{n-1+\varepsilon})$ possibilities for f in this case. Finally, the ε will be removed (and thereby give the desired estimate for $\#P_3(H)$) by a further splitting of $P_3(H)$ into two sets depending on the sizes of the primes dividing C.

2. Fourier analysis – Artane Siad (April 5)

2.1. **Recap.** Let f be a monic irreducible integral polynomial of degree n. Let $K_f = \mathbb{Q}[x]/(f(x))$ be the field obtained by adjoining a root of f and let D (resp. C) be the discriminant of K_f (resp. the product of the primes dividing D). We want to show the size of the set

$$P(H) = \{f : H(f) \leq H \text{ and } \operatorname{Gal} f \neq S_n \text{ and primitive}\}$$

is $O(H^{n-1})$. (We have already reduced to the case $\operatorname{Gal} f$ transitive and primitive by Chela and Widmer's results, respectively.) Recall our strategy is to cover $P(H) = P_1(H) \cup P_2(H) \cup P_3(H)$ with three sets and bound each subset: $P_1(H)$ is the subset with smooth "large" discriminant, $P_2(H)$ is the subset with "small" discriminant, and $P_3(H)$ we'll come back to. More precisely, $P_1(H)$ is the subset of f such that

$$C \leq H^{1+\delta} \text{ but } D > H^{2+2\delta} \text{ for some } \delta > 0$$

and we'll show that

$$\#P_1(H) = O(H^n 2^{\omega(D)}/D)$$

2.2. **Fourier analysis.** Let V_R (resp. V_R^1) be the free R-module of polynomials (resp. monic polynomials) over R of degree n. Let p be a prime. The Fourier transform of a function $\Psi_p \colon V^1_{\mathbb{F}_p} \to \mathbb{C}$ is the function $\widehat{\Psi_p} \colon V^{1*}_{\mathbb{F}_p} \to \mathbb{C}$ defined by the usual formula

$$\widehat{\Psi_p}(g) = \frac{1}{p^n} \sum_{f \in V^1_{\mathbb{F}_p}} \Psi_p(g) \exp\left(\frac{2\pi i[f,g]}{p}\right)$$

where $[f,g] = g(f)$. One can make the same definition with p replaced by a squarefree number $p_1 \cdots p_r$. Then

$$\widehat{\Psi_{p_1 \cdots p_r}}(g) = \prod_i \widehat{\Psi_{p_i}}(g \pmod{p_i})$$

which follows from the Chinese remainder theorem.

Proposition (Twisted Poisson summation formula)**.** *For p a prime and ϕ a Schwartz function on $V^1_{\mathbb{R}}$,*

$$\sum_{f \in V^1_{\mathbb{Z}}} \Psi_p(f)\phi(\frac{f}{H}) = H^n \sum_{g \in V^{1*}_{\mathbb{Z}}} \widehat{\Psi_p}(g)\widehat{\phi}(\frac{gH}{p}).$$

For our application, ϕ will be a Schwartz function approximating the indicator function of the unit box $[-1,1]^n$ and Ψ_p will be the indicator function of a subset

$S \subset V^1_{\mathbb{F}_p}$. In this case, the left-hand side of the twisted Poisson formula becomes

$$\#\{v \in V^1_{\mathbb{Z}} : H(v) \le H \text{ and } v \pmod p \in S\}.$$

The right-hand side will be

$$\sim H^n(\widehat{\Psi_p}(0) + \text{error}).$$

where the error is $O(M(\Psi_p)(p^{1+\varepsilon}/H)^n)$ and $M(\Psi_p) = \max_{g\neq 0} |\widehat{\Psi_p(g)}|$. Roughly speaking, we'll want to apply this to S which is the set of polynomials modulo primes dividing C which have large indices modulo these primes.

Definition 2.1. A *splitting type* is a multiset of tuples of positive integers

$$\{(f_1, e_1), (f_2, e_2), \dots, (f_r, e_r)\}$$

(denoted $\sigma = (f_1^{e_1} \cdots f_r^{e_r})$). We define $\deg \sigma = \sum_i f_i e_i$, $\text{ind}\sigma = \sum_i (e_i - 1)f$, $\text{len}\sigma = \deg\sigma - \text{ind}\sigma = \sum_i f$. An automorphism of σ is a degree-and-multiplicity-preserving bijection of σ as a multiset.

The factorization of a polynomial into irreducibles naturally defines a splitting type which records the degrees f_i of irreducible factors and their multiplicities e_i.

Lemma 2.2. *For any splitting type σ let $w_{p,\sigma} \colon V^1_{\mathbb{F}_p} \to \mathbb{C}$ be defined to be the number of times σ appears as a sub-splitting type of the splitting type of f. If σ has degree $< n$ we define $m_\sigma(g)$ be the minimum length of a sub-splitting type $\tilde\sigma$ of σ having the property there exists a fixed realization f_σ of $\tilde\sigma$ such that $[f, g]$ is constant over the set of f divisible by f_σ. If σ has degree n we set $m_\sigma(g) = 0$. Then*

$$\widehat{w_{p,\sigma}}(0) = \frac{p^{-k}}{\#\operatorname{Aut}(\sigma)} + O(p^{-k-1})$$

and

$$M(w_{p,\sigma}) = \max_{g\neq 0} |\widehat{w_{p,\sigma}}(g)| = \begin{cases} O(p^{-k-\min_{i=1}^r f_i}) & \text{if } \deg\sigma < n, \\ O(p^{-k-1/2}) & \text{if } \deg\sigma = n. \end{cases}$$

Fix a positive integer D, let $C = \text{rad}(D)$, and let $\delta \in (0, 2n)$. Our goal is to show that $\#\{f \text{ with } \|f\| < H \text{ and } \text{disc}(K_f) = D\} = O_\varepsilon(H^{n+\varepsilon}/D)$ when $C \le H^{1+\delta}$ and $D \ge H^{2+2\delta}$ and squarefull. Our strategy will be to write $D = p_1^{k_1} \cdots p_m^{k_m}$ and impose congruence conditions on $f \pmod{p_i}$ so that the index of $f \pmod{p_i}$ is at least k_i.

Recall the Twisted Poisson formula strategy: for p a prime and ϕ a Schwartz function on $V^1_{\mathbb{R}}$,

$$\sum_{f\in V^1_{\mathbb{Z}}} \Psi_p(f)\phi(\frac{f}{H}) = H^n \sum_{g\in V^{1*}_{\mathbb{Z}}} \widehat{\Psi_p}(g)\widehat{\phi}(\frac{gH}{p}).$$

We'll take ϕ to be a smooth approximation to the indicator function on $[-1, 1]^n$ and Ψ_p to be the indicator function of a subset of $V(\mathbb{F}_p)$ of polynomials with

large indices. The right-hand side will be roughly $\sim H^n(\widehat{\Psi_p}(0)+\text{error})$ where the error is $O(M(\Psi_p)(p^{1+\varepsilon}/H)^n)$ and $M(\Psi_p)=\max_{g\neq 0}|\widehat{\Psi_p(g)}|$. The above lemma will help us bound this error. Let $\sigma=(f_1^{e_1}\cdots f_r^{e_r}$ be a type of degree d, index k, and length $\ell=d-k$. Recall that $w_{p,\sigma}\colon V^1_{\mathbb{F}_p}\to\mathbb{C}$ is defined by letting $w_{p,\sigma}(f)$ be the number of tuples $(P_1,\dots,P_r)$ with P_i distinct monic polynomials such that $\deg P_i=f_i$ and $P_1^{e_1}\cdots P_r^{e_r}$ divides f.

Lemma 2.3.

$$M(w_{p,\sigma})=\begin{cases}O(p^{-k-\min_i f_i}) & \text{if } d<n,\\ O(p^{-k-1/2}) & \text{if } d=n.\end{cases}$$

$$\widehat{w_{p,\sigma}}(0)=\frac{p^{-k}}{\#\mathrm{Aut}\sigma}+O(p^{-k-1}).$$

Corollaries of the lemma:

(1) For p a prime, the number of $f\in V^1_{\mathbb{Z}}\cap[-H,H]^n$ with mod p index at least k is
$$O_\varepsilon(\frac{H^n}{p^k}+p^{n-k-1/2+\varepsilon}).$$

(2) Let $D=p_1^{k_1}\cdots p_m^{k_m}$. Then the number of $f\in V^1_{\mathbb{Z}}\cap[-H,H]^n$ with mod p_i index at least k for all i is
$$O_\varepsilon(\frac{H^{n+\varepsilon}}{D}+H^\varepsilon\prod_i p_i^{n-k_i-1/2+\varepsilon}).$$
(Note that $100^{\omega(D)}=O_\varepsilon(D^\varepsilon)=O_\varepsilon(H^\varepsilon)$ for D is at most $O(H^{2n-2})$.)

(3) Fix $\delta\in(0,\frac{1}{2n-1}]$ and $C=p_1\cdots p_m<H^{1+\delta}$ squarefree and $D=\prod_i p_i^{k_i}$. Then the number of $f\in V^1_{\mathbb{Z}}\cap[-H,H]^n$ with mod p_i index at least k for all i is
$$O_\varepsilon(\frac{H^{n+\varepsilon}}{\prod_i p_i^{k_i}}+\frac{H^\varepsilon\prod_i p_i^{n-1/2+\varepsilon}}{\prod_i p_i^{k_i}}) \tag{1}$$
Observe that
$$\frac{H^\varepsilon\prod_i p_i^{n-1/2+\varepsilon}}{\prod_i p_i^{k_i}}=O_\varepsilon(\frac{H^\varepsilon H^{(1+\delta)(\frac{2n-1}{2}+\varepsilon)}}{\prod_i p_i^{k_i}})=O_\varepsilon(\frac{H^{\varepsilon+n-1/2+1/2+\varepsilon}}{\prod_i p_i^{k_i}})=O_\varepsilon(\frac{H^{n+\varepsilon}}{\prod_i p_i^{k_i}})$$
which shows that the second summand in (1) can be absorbed into the first summand.

3. Case I of Proof of Theorem 1 – Andrew O'Desky (April 5)

Let V denote the affine space of degree n monic polynomials.

3.1. The density of irreducible polynomials with given discriminant.

Recall we regard splitting types of polynomials as unordered tuples σ of pairs (f_i,e_i) of positive integers, written $\sigma=(f_1^{e_1}f_2^{e_2}\cdots)=\vec{f}^{\vec{e}}$.

Lemma 3.1. *The mod p density of $\{f \in V(\mathbb{Z}) : \sigma(f \pmod p) = \sigma\}$ is $O(p^{-\mathrm{ind}(\sigma)})$.*

Proof. Say $\sigma = (f_1^{e_1} \cdots f_r^{e_r})$. There is a surjective function

$$\prod_{i=1}^{r} \{g_i \in \mathbb{F}_p[x] \text{ monic irred. of degree } f_i\} \to \{f \in \mathbb{F}_p[x] : \sigma(f) = \sigma\}$$
$$(g_1, \ldots, g_r) \mapsto g_1^{e_1} \cdots g_r^{e_r}$$

whose fibers have cardinalities $\leq r!$ so are bounded independently of p. By Gauss's formula, $\#\{g \text{ irred. of degree } f\} = \frac{1}{f}p^f + O(p^{f-1})$. Therefore

$$\#\{f \in \mathbb{F}_p[x] : f \text{ monic has type } \sigma\} = O(\prod_{i=1}^{r} \tfrac{1}{f_i} p^{f_i}) = O(p^{f_1+\cdots+f_r}).$$

□

Let K be a number field and let σ_p denote the splitting type of a prime p in the extension $K/\mathbb{Q}$. We will need an upper bound on the p-adic valuation of the discriminant.

Lemma 3.2. $v_p(d_{K/\mathbb{Q}}) \leq \mathrm{ind}(\sigma_p) + n[\log_p n]$.

All we need from this is that $v_p(d_{K/\mathbb{Q}}) = \mathrm{ind}(\sigma_p)$ for all $p > n$ and that $|v_p(d_{K/\mathbb{Q}}) - \mathrm{ind}(\sigma_p)|$ is bounded uniformly in p and K.

Proof. Let $P \subset K$ be a prime ideal containing p. Then $v_P(\mathcal{D}_{K/\mathbb{Q}}) \leq e_P - 1 + v_P(e_P) = e_P - 1 + e_P v_p(e_P)$ (Serre's *Local Fields*, Ch. III, §7, p. 58). Thus $v_p(d_{K/\mathbb{Q}}) \leq \sum_{P|p} f_P(e_P - 1 + e_P v_p(e_P)) \leq \mathrm{ind}(\sigma_p) + n[\log_p n]$ since $e_P \leq n$ and $v_p(e_P)$ is at most the highest power of p dividing $\mathrm{lcm}(1, \ldots, n)$ which is $[\log_p n]$. □

Corollary 3.3. *For positive integers C and D with $C = \mathrm{rad}(D)$, the mod C density of*

$$V_D = \{f \in V(\mathbb{Z}) : f \text{ irred.}, \mathrm{disc}(K_f) = D\}$$

is $O(c^{\omega(D)}/D)$ for some $c > 0$.

Proof. Let $f \in V_D$, let $a \in K_f$ be a root of f, let $O_f = \mathbb{Z}[a]$ and let $O \subset K_f$ be the maximal order. For any prime p, $O_f \otimes \mathbb{F}_p$ is isomorphic to

$$\frac{\mathbb{F}_{p^{f_1}}[x]}{(x)^{e_1}} \times \cdots \times \frac{\mathbb{F}_{p^{f_r}}[x]}{(x)^{e_r}}$$

for a uniquely determined type $\vec{f^e}$ (namely the splitting type of $f \pmod p$). In the same manner, let σ_p be the type corresponding to $O \otimes \mathbb{F}_p$. One checks using the map $O_f \otimes \mathbb{F}_p \to O \otimes \mathbb{F}_p$ that $\mathrm{ind}(\sigma(f \pmod p)) \geq \mathrm{ind}(\sigma_p)$ (same number of maximal ideals for both algebras and the residue degrees only go up).

Thus $\mathrm{ind}(\sigma(f \pmod p)) \geq \mathrm{ind}(\sigma_p) \geq v_p(D) - n[\log_p n]$ (Lemma 3.2), which implies the mod p reduction of the set V_D is contained in

$$\bigcup_{\substack{\sigma \text{ such that} \\ \mathrm{ind}(\sigma) \geq v_p(D) - n[\log_p n]}} \{f \in V(\mathbb{F}_p) : \sigma(f) = \sigma\}.$$

By Lemma 3.1, the cardinality of this set is

$$\sum_{\substack{\sigma \text{ such that} \\ \mathrm{ind}(\sigma) \geq v_p(D) - n[\log_p n]}} \#\{f \in V(\mathbb{F}_p) : \sigma(f) = \sigma\} = O(p^{n-(v_p(D)-n[\log_p n])})$$

$$= O(p^{n-v_p(D)} n^n) = O(p^{n-v_p(D)}).$$

By the Chinese remainder theorem, the density mod C is the product of the density mod primes dividing C, so the mod C density of V_D is

$$\prod_{p|C} O(p^{-v_p(D)}) = O(c^{\omega(D)}/D).$$

□

3.2. **Summing over discriminants.** Let G be the Galois group of the normal closure of $K/\mathbb{Q}$. We can compute the index of σ_p using the action of an(y) inertia subgroup at p on $H\backslash G$, where $H \subset G$ is the subgroup fixing K.

Lemma 3.4. *We have* $\mathrm{ind}(\sigma_p) = [K : \mathbb{Q}] - i_p$ *where* i_p *is the number of orbits of an(y) inertia subgroup* I_p *at* p *on* $H\backslash G$. *If* p *is ramified then* $\mathrm{ind}(\sigma_p) \geq \mathrm{ind}(G)$.

Proof. Let D (resp. I) denote the decomposition group (resp. inertia group) of a fixed prime Q of the normal closure of K lying over p. Recall that $H\sigma D \mapsto P_\sigma = \sigma Q \cap K$ defines a bijection $H\backslash G/D \to \{\text{primes of } K \text{ dividing } p\}$. Since the (lower-numbering) ramification filtration of H is determined by restriction from G, we have that $|H\sigma D| = e_{P_\sigma} f_{P_\sigma} |H|$ and $|H\sigma I| = e_{P_\sigma}|H|$. Let $\{H\sigma_i D\}_i$ be representatives for $H\backslash G/D$. Then under the action of I, the set $H\backslash G$ decomposes into $f_{P_{\sigma_1}}$ sets of size $e_{P_{\sigma_1}}$, $f_{P_{\sigma_2}}$ sets of size $e_{P_{\sigma_2}}$, and so on, showing that $\mathrm{ind}(\sigma_p) = |H\backslash G| - f_{P_{\sigma_1}} - f_{P_{\sigma_2}} - \cdots = |H\backslash G| - i_p$.

If p is ramified, then I contains a nontrivial element g. Thus $\mathrm{ind}(G) \leq \mathrm{ind}(g) \leq [K : \mathbb{Q}] - i_p = \mathrm{ind}(\sigma_p)$. □

Lemma 3.5.

$$\sum_{D > X \text{ sq.full}} 1/D = O(X^{-1/2}).$$

Proof. Note that $\frac{1-p^{-s}+p^{-2s}}{1-p^{-s}} = 1 + p^{-2s} + p^{-3s} + \cdots$ so the Dirichlet series

$$\sum_{D \text{ sq.full}} 1/D^s = \prod_p \frac{1 - p^{-s} + p^{-2s}}{1 - p^{-s}} = \frac{\zeta(2s)\zeta(3s)}{\zeta(6s)}$$

has simple poles at $s = 1/3, 1/2$. By the theorem of Wiener–Ikehara this means $A(X) = \#\{\text{sq.full } D \leq X\} \sim \frac{\zeta(3/2)}{\zeta(6/2)} X^{1/2}$. Now applying Abel's summation formula to $\sum_{X<n\leq Y} a_n \phi(n)$ for $\phi(x) = x^{-s}$ (for any $s > 0$) and a_n the nth Dirichlet coefficient of $\frac{\zeta(2s)\zeta(3s)}{\zeta(6s)}$ gets that

$$\begin{aligned}\sum_{X<n\leq Y,\ \text{sq.full}} n^{-s} &= A(Y)\phi(Y) - A(X)\phi(X) + s\int_X^Y A(Z)Z^{-s-1}\,dZ \\ &= O(Y^{1/2-s}) + O(X^{1/2-s}) + s\int_X^Y O(Z^{1/2-2s-1})\,dZ \\ &= O(Y^{1/2-s}) + O(X^{1/2-s}) + O(Y^{1/2-2s} + X^{1/2-2s}) \\ &= O(Y^{1/2-s}) + O(X^{1/2-s}).\end{aligned}$$

Taking $s = 1$ and Y to infinity gets

$$\sum_{n>X,\ \text{sq.full}} n^{-1} = O(X^{-1/2}).$$

□

Proposition 3.6 (Jordan)**.** *If a degree n permutation group is transitive and primitive and not S_n, then* $\operatorname{ind}(G) = \min_{g\neq 1} \operatorname{ind}(g)$ *is at least two.*

Proof. The relation $i \sim j$ if $(ij) \in G$ on $[n]$ is preserved by G, so it is discrete or indiscrete; if the former, G contains no transpositions, and otherwise $G = S_n$. □

For $H > 0$ and $\delta \in (0, 1/(2n))$ let $P_1(H)$ be the set of $f \in V(\mathbb{Z})$ with $\|f\| \leq H$, primitive Galois group $G_f \neq S_n$, where K_f has discriminant D and $\operatorname{rad}(D) = C$ satisfying $C \leq H^{1+\delta}$ and $D > H^{2(1+\delta)}$.

Proposition 3.7 (Case I of Theorem 1)**.** $\#P_1(H) = O(H^{n-1})$.

Proof. Let $V_D(H)$ denote the subset of $f \in P_1(H)$ for which $\operatorname{disc}(K_f) = D$. Supposing $P_1(H)$ were defined instead with $\delta = 0$, we would have $C \leq H$, in which case $\#V_D(H) = O(H^n \delta_C(V_D(H)))$ where $\delta_C(V_D(H))$ is the mod C density of $V_D(H)$. The results of §2 show however that we may take these same bounds for any $\delta \in (0, 1/(2n))$. Using Corollary 3.3 shows that $\#V_D(H) = O(H^n c^{\omega(D)}/D)$ for some $c > 0$. By the index formula (Lemma 3.4) and Jordan's classical result we know that D must be squarefull, so with the help of Lemma 3.5,

$$\begin{aligned}\#P_1(H) = \sum_{D>H^{2(1+\delta)}\ \text{sq.full}} O(H^n c^{\omega(D)}/D) &= \sum_{D>H^{2(1+\delta)}\ \text{sq.full}} O_\varepsilon(H^{n+\varepsilon}/D) \\ &= O_\varepsilon(H^{n+\varepsilon-(1+\delta)}) = O(H^{n-1}).\end{aligned}$$

□

4. Case II of Proof of Theorem 1 – Danny Neftin (April 19)

Case II means we assume $C \leq H^{1+\delta}$ and $D \leq H^{2+2\delta}$ where δ is some fixed number in $(0, 1/(2n))$. Schmidt proved that

$$\#\{K_f/\mathbb{Q} : [K_f : \mathbb{Q}] = n, d_{K_f/\mathbb{Q}} \leq X\} = O(X^{(n+2)/4}).$$

By a result of Lemke Oliver and Thorne, a given number field K of degree n with primitive Galois group is equal to K_f for at most

$$O(\frac{H(\log H)^{n-1}}{|d_{K/\mathbb{Q}}|^{\frac{1}{n(n-1)}}}) = O(H^{n-1}).$$

So $\#P_2(H)$ is

$$\sum_{K: d_{K/\mathbb{Q}} < H^{2+2\delta}} O(\frac{H(\log H)^{n-1}}{|d_{K/\mathbb{Q}}|^{\frac{1}{n(n-1)}}})) = O(H^{(2+2\delta)\frac{n+2}{4}})O(H(\log H)^{n-1}) = O(H^{n-1}).$$

5. Case III of Proof of Theorem 1 – Andrew O'Desky (April 26)

Let $f \in V(\mathbb{Z})$ with primitive Galois group $G_f \neq S_n$, $d_{K_f/\mathbb{Q}} = D$, and suppose that (for fixed $\delta \in (0, 1/(2n))$) $C = \prod_{p|D} p > H^{1+\delta}$. Let $a_1, \ldots, a_{n-1}$ be the first $n-1$ coefficients of f and consider the one-parameter family of polynomials $\mathbb{A}^1 \to V : a \mapsto f_a = x^n + a_1 x^{n-1} + \cdots + a_{n-1}x + a$. Let $DD = DD(a_1, \ldots, a_{n-1})$ denote the following "double discriminant":

$$DD = \operatorname{disc}(\mathbb{A}^1 \xrightarrow{f_a} V \xrightarrow{\operatorname{disc}} \mathbb{A}^1)$$

(For example, $DD_{n=3} = (18a_1a_2 - 4a_1^3)^2 + 108(a_1^2a_2^2 - 4a_2^3)$.)

Proposition (Proposition 5.2, [1])**.** *Let $h(a) \in \mathbb{Z}[a]$. Suppose that p^2 divides $h(a)$ as well as $h(c + pa)$ for some $a \not\equiv 0 \pmod p$. Then $h'(c)$ is divisible by p.*

Proof. We can write $h(a) = h(c) + h'(c)(x - c) + (x - c)^2 r(x)$ for an integral polynomial $r(x)$. Setting $x = c + pa$ and reducing mod p^2 shows $0 = h'(c)pa$. □

Lemma 5.1. $\operatorname{disc}(f) \equiv 0 \pmod{p^2}$ *if $f_{\overline{\mathbb{F}_p}}$ has at least a triple root or two double roots.*

Proof. We may assume f has distinct roots. Suppose that f has roots

$$r_1, r_2, \ldots, r_a, s_1^{(1)}, \ldots, s_{e_1}^{(1)}, \ldots, s_1^{(\ell)}, \ldots, s_{e_\ell}^{(\ell)}$$

which all lie over the same root of $f_{\overline{\mathbb{F}_p}}$, where $r_i \in \mathbb{Q}_p$ and $\{s_j^{(i)}\}_j$ consists of $G(\mathbb{C}_p/\mathbb{Q}_p^{\mathrm{ur}})$-conjugates for all i. Then the ramification index of $\mathbb{Q}_p(s_j^{(i)})$ is e_i for

all i, j. Then

$$\begin{aligned} v_p(\mathrm{disc}(f)) &\geq \sum_{i\neq j} v_p(r_i - r_j) + 2\sum_{k=1}^{a}\sum_{i=1}^{\ell}\sum_{j=1}^{e_i} v_p(r_k - s_j^{(i)}) + \sum_{i=1}^{\ell}\sum_{j\neq j'} v_p(s_j^{(i)} - s_{j'}^{(i)}) \\ &\geq a(a-1) + 2\sum_{k=1}^{a}\sum_{i=1}^{\ell}\sum_{j=1}^{e_i}\frac{1}{e_i} + \sum_{i=1}^{\ell}\frac{1}{e_i}e_i(e_i-1) \\ &= a^2 - 2a + n + \ell(2a-1). \end{aligned}$$

If $f_{\mathbb{Q}_p^{\mathrm{ur}}}$ has an irreducible factor of degree $d > 1$, then the inequality shows that $v_p(\mathrm{disc}(f)) \geq 0^2 - 0 + n + 1(0-1) = n - 1$. If $f_{\overline{\mathbb{F}_p}}$ has at least two double roots or a root of multiplicity at least three, then we see that $v_p(\mathrm{disc}(f))$ is at least 2 in either case. □

In particular, the condition that $p^2|\mathrm{disc}(f)$ follows from splitting conditions on f mod p. This shows that $p^2|\mathrm{disc}(f + pg)$ for any integral g with $\deg g < \deg f$, as $f + pg$ will have the same splitting type mod p as f. This means the polynomial $a \mapsto h(a) = \mathrm{disc}(f_a)$ satisfies the hypothesis of the last proposition with $c = a_n$ for any $p|C$, which means $h'(a_n)$ is divisible by C. This means that $DD(a_1, \ldots, a_{n-1}) = \mathrm{Res}(h, h')$ is divisible by C because modulo every $p|C$ the polynomials h and h' share a common root $a = a_n$. The number of prime divisors of n satisfies the bound $\omega(n) = O(\frac{\log n}{\log\log n})$, which means the number of squarefree-divisors function satisfies $d^{sf}(n) = 2^{\omega(n)} = n^{O(\frac{1}{\log\log n})} = O_\varepsilon(n^\varepsilon)$.

Suppose that $DD(a_1, \ldots, a_{n-1}) = 0$. In other words, the integer point $P = (a_1, \ldots, a_{n-1})$ lies on the affine variety $V(DD)$ where DD vanishes. Using the large sieve and Lang–Weil for the mod p estimates, one can show that for any thin set $T \subset \mathbb{Z}^d$

$$\{P \in T : \|P\| \leq H\} = O_\varepsilon(H^{d-1/2+\varepsilon})$$

(due to S. Cohen). If V is an affine variety of dimension d, then there is a finite map $V \to \mathbb{A}^d$ by Noether normalization, and applying this bound to the image of $V(\mathbb{Z})$ proves that

$$\{P \in V(\mathbb{Z}) : \|P\| \leq H\} = O_\varepsilon(H^{d-1/2+\varepsilon}).$$

In particular the set of integer points on $V(DD)$ of height $\leq H$ is $O_\varepsilon(H^{n-2-1/2+\varepsilon})$, so the set of $f = x^n + a_1x^{n-1} + \cdots + a_n$ of height $\leq H$ with $DD(a_1, \ldots, a_{n-1}) = 0$ is $O_\varepsilon(H^{n-2-1/2+\varepsilon+1}) = O(H^{n-1})$.

Suppose that $DD(a_1, \ldots, a_{n-1}) \neq 0$. Then C is determined up to

$$d^{sf}(DD(a_1, \ldots, a_{n-1})) = O_\varepsilon(H^\varepsilon)$$

possibilities. Once C is determined, we can determine the residue class of $a_n \pmod p$ up to $O(1)$ possibilities for any $p|C$, since it is a root of $h(a) = \mathrm{disc}(f_a)$ modulo p. (Note $h(a)$ is not identically zero modulo p once p does not divide any of the coefficients of $\mathrm{disc} f$, for the discriminant of $x^n + a$ is not identically zero,

which implies that $a^{\deg \operatorname{disc} f}$ appears as a monomial in $\operatorname{disc} f$, and this term will be nonzero modulo p.) So we can determine $a_n \pmod C$ up to $O(1)^{\omega(C)} = O_\varepsilon(H^\varepsilon)$ possibilities, but then this actually determines a_n since $C > H$. So the number of f such that $DD(a_1, \dots, a_{n-1}) \neq 0$ is $O_\varepsilon(H^{n-1+\varepsilon})$.

One needs an additional argument to remove the epsilson from this last bound.

6. Section 2 – Eilidh McKemmie (May 17)

6.1. **Recap.** Let $G \leq S_n$ and let $N_n(G, H)$ denote the number of monic integral polynomials with height $\leq H$ and Galois group G.

Theorem (Theorem 1)**.** *$N_n(G, H) = O(H^{n-1})$ if $G \neq S_n$.*

The case G intransitive was done by van der Waerden, and Chela gave better bounds. The case G imprimitive was proven by Widmer. We'll focus on the primitive case. For an element g of the degree n permutation group G, we define the index of g to be $\operatorname{ind}(g) = n - \#\{\text{orbits of } g\}$ and define $\operatorname{ind}(G) = \min_{g \neq 1} \operatorname{ind}(g)$. We also define $F_n(G, X)$ to be the number of number fields with Galois group G and (absolute) discriminant $< X$. We let a_G denote any constant satisfying $F_n(G, X) = O(X^{a_G})$. (Malle's conjecture is that for any $\varepsilon > 0$, we can take $a_G = \frac{1}{\operatorname{ind}(G)} + \varepsilon$.)

6.2. **Theorem 2 of Bhargava** [1]**.**

Theorem (Theorem 2, [1])**.**

$$N_n(G, H) = O(\min\{H^{n+1-\operatorname{ind}(G)} + ?, *\})$$

where

$$? = H^{n-(n-1)\frac{1-1/\operatorname{ind}(G)}{a_G+1-1/\operatorname{ind}(G)-u}} \log^n H,$$
$$* = H^{(2n-2)(a_G-u)+1} \log^{n-1} H,$$

and

$$u = \begin{cases} \frac{1}{n(n-1)} & \text{if } G \text{ primitive,} \\ 0 & \text{otherwise.} \end{cases}$$

If Malle's conjecture is true, then the ? term can be removed. To get unconditional results, we'll need to bound ? using worse upper bounds for a_G than what Malle's conjecture provides.

Jordan proved that if G is primitive and $\neq S_n$, then G contains no transposition, and therefore $\operatorname{ind}(G) \geq 2$. In this case, $N_n(G, H) = O(H^{n-1})$. The idea was to split up the set we're counting $P(H)$ into three sets $P_1(H), P_2(H)$, and $P_3(H)$. Theorem 2 is proved with an analogous approach but with modified subsets. The subset $P_3(H)$ will be unchanged, while the first two subsets are modified as follows: $P_1(H) \coloneqq \{C \leq H^{1+\delta}, D > Y\}$ and $P_2(H) \coloneqq \{D \leq Y\}$.

Corollary (Corollary 3, [1])**.**

(1) *if G is non-elemental, then $N_n(G,H) = O(H^{b\sqrt{n}\log^2 n})$ for an absolute constant $b > 0$,*
(2) *if G is elemental, then $N_n(G,H) = O(H^{n-\frac{bn}{\log^2 n}})$ for an absolute constant $b > 0$.*

It will be easier to find the index for the elemental groups, while it will be easier to find a_G for the non-elemental groups. The work is mostly done already for elemental groups, so we'll do those first. Let G be a permutation group acting on a set Ω. Let $\binom{[m]}{k}$ denote the *set* of all k-sets in $[m] = \{1, \dots, m\}$. We say G is *non-elemental* if there exist $r, m, k > 0$ positive integers satisfying $k < m/2$ and $rk > 1$ such that

$$\Omega = \binom{[m]}{k}^r$$

and G is contained in the subgroup

$$S_m \wr S_r \subset \mathrm{Aut}(\Omega).$$

Recall that $S_m \wr S_r$ is defined to be $(S_m)^r \rtimes S_r$, so the condition means that any $g \in G$ can be realized by a permutation of the form $g = (g_1, \dots, g_r, h)$ for $g_i \in S_m$ and $h \in S_r$. For elemental groups, we have the following lower bound for the index:

Theorem (Theorem 13, [1]; Guralnick–Magaard [4])**.** *If $G \leq S_n$ is elemental and primitive, then* $\mathrm{ind}(G) \geq \frac{3n}{14}$.

More specifically, Guralnick–Magaard [4] gave upper bounds for the number of fixed points of any permutation in G. Fixed points relate to the index since one has the easy inequality $\mathrm{ind}(g) \geq \frac{1}{2}\#\{\text{moved points}\}$.

Theorem (Theorem 19, [1]; Schmidt [6] and Lemke-Oliver–Thorne [5])**.** *We may take $a_G = c\log^2 n$.*

These bounds for a_G and the index, plugged in to the ? term of Theorem 2, (unconditionally) deduce the second part of Corollary 3 for elemental groups.

Now for non-elemental groups. We need to bound a_G. Let G' denote an abstract copy of G (which may act differently on Ω). We will let G' act (imprimitively) on rm symbols $(x_{ij})_{1\leq i\leq m, 1\leq j\leq r}$, where the ith copy of S_m acts on the ith column of letters, and S_r permutes the columns. We claim that

$$F_n(G,X) \ll F_{rm}(G', X^{\frac{3rm}{n}}).$$

This claim is what we need to prove the second part of Corollary 3: combining the claim with the Schmidt bound [6], we have that (for $n < 95$)

$$F_n(G,X) \ll F_{rm}(G', X^{\frac{3rm}{n}}) = O(X^{\frac{3rm(rm+2)}{4n}}) = O(X^4).$$

Combining the claim with the bound from Lemke-Oliver–Thorne [5], we have that (for $n \geq 95$)

$$F_n(G, X) \leq F_{rm}(G', X^{\frac{3rm}{n}}) = O(X^{\frac{3brm\log^2(rm)}{n}}) = O(X^{\frac{b\log^2 n}{\sqrt{n}}})$$

for $r^2m^2 \ll n$. In these two ranges of n, these two inequalities (respectively) imply smaller values of a_G than were previously known. Plugging these bounds for a_G into Theorem 2 proves the Corollary. So now we turn to proving the claim.

Theorem (Theorem 16, [1])**.** *Let G be non-elemental and primitive. Let $g \in G$ correspond to $g' \in G'$. If $g \neq 1$, then $\frac{\mathrm{ind}(g)}{\mathrm{ind}(g')} > \frac{n}{3rm}$.*

Proof. Let $g = (g_1, \ldots, g_r, h)$. If $h \neq 1$, then g moves at least $\binom{m}{k}^r - \binom{m}{k}^{r-1}$ k-sets. Then

$$\mathrm{ind}(g) \geq \frac{\#\{\text{moved}\}}{2} \geq \frac{1}{2}\left(\binom{m}{k}^r - \binom{m}{k}^{r-1}\right)$$

From this we see that

$$\frac{\mathrm{ind}(g)}{\mathrm{ind}(g')} > \frac{1}{2rm}\left(\binom{m}{k}^r - \binom{m}{k}^{r-1}\right) > \frac{n}{3rm}$$

If $n = 1$, consider i such that $\mathrm{ind}(g_i)$ is maximal. We will associate an element $\overline{g_i}$ to g_i which satisfies $\mathrm{ind}(g) \geq \mathrm{ind}(\overline{g_i})$. For example, if $g_i = (123)(4567)(8)$ then $\overline{g_i}$ will be $(12)(3)(45)(67)(8)$. (This association will let us reduce the computation to the easy case when the permutation g_i is a product of disjoint transpositions.) Define $f(m, k, y)$ to be the number of k-sets S of $[m]$ such that S is moved by a $\sigma \in S_m$ where σ is a product of y disjoint transpositions. Bhargava proves that

$$f(m, k, y) \geq \frac{8}{5k}\binom{m-1}{k-1}y$$

by induction on k. (Proof skipped). The number of disjoint transpositions in $\overline{g_i}$ is at least $\frac{\mathrm{ind}(g_i)}{2}$, so $\overline{g_i}$ moves at least $\frac{8}{mk}\binom{m-1}{k-1}\frac{\mathrm{ind}(g_i)}{2}$ elements. (For the example above, the index of g_i is five, and $y = 3 \geq \frac{5}{2}$.) We have that

$$\mathrm{ind}(g) \geq \frac{\#\{\text{moved}\}}{2} \geq \frac{4}{5k}\binom{m-1}{k-1}\frac{\mathrm{ind}(g_i)}{2}\binom{m}{k}^{r-1}$$

and $\mathrm{ind}(g') \leq r\mathrm{ind}(g_i)$. □

Any prime $p > n$ will be tamely ramified in $K_f = K$ (for a given f with Galois group G). Let K' be the G'-resolvent of K (i.e. $K' = \Omega^{H'}$ where H' is the stabilizer of any point in the action of G'). Let $g \in G$ be any generator of the inertia group at p. Then $\mathrm{ind}(g) = v_p(\mathrm{Disc}(K)$ and $\mathrm{ind}(g') = v_p(\mathrm{Disc}(K')$.

Corollary (Corollary 18, [1])**.**

$$|\mathrm{Disc}(K)| \gg |\mathrm{Disc}(K')|^{\frac{n}{3rm}}.$$

References

[1] M. Bhargava. Galois groups of random integer polynomials and van der Waerden's Conjecture, 2021.
[2] R. Chela. Reducible polynomials. *J. London Math. Soc.*, 38:183–188, 1963.
[3] J. S. Ellenberg and A. Venkatesh. The number of extensions of a number field with fixed degree and bounded discriminant. *Ann. of Math. (2)*, 163(2):723–741, 2006.
[4] R. Guralnick and K. Magaard. On the minimal degree of a primitive permutation group. *J. Algebra*, 207(1):127–145, 1998.
[5] R. J. L. Oliver and F. Thorne. Upper bounds on number fields of given degree and bounded discriminant, 2020.
[6] W. M. Schmidt. Number fields of given degree and bounded discriminant. Number 228, pages 4, 189–195. 1995. Columbia University Number Theory Seminar (New York, 1992).
[7] B. L. van der Waerden. Die Seltenheit der reduziblen Gleichungen und der Gleichungen mit Affekt. *Monatsh. Math. Phys.*, 43(1):133–147, 1936.
[8] M. Widmer. On number fields with nontrivial subfields. *Int. J. Number Theory*, 7(3):695–720, 2011.