Annals of Mathematics, **163** (2006), 723–741

The number of extensions of a number field with fixed degree and bounded discriminant

By Jordan S. Ellenberg and Akshay Venkatesh*

Abstract

We give an upper bound on the number of extensions of a fixed number field of prescribed degree and discriminant $\leq X$; these bounds improve on work of Schmidt. We also prove various related results, such as lower bounds for the number of extensions and upper bounds for Galois extensions.

1. Introduction

Let K be a number field, and let $N_{K,n}(X)$ be the number of number fields L (always considered up to K-isomorphism) such that $[L:K] = n$ and $\mathbf{N}_{\mathbb{Q}}^{K}\mathcal{D}_{L/K} < X$. Here $\mathcal{D}_{L/K}$ is the relative discriminant of L/K, and $\mathbf{N}_{\mathbb{Q}}^{K}$ is the norm on ideals of K, valued in positive integers. $\mathcal{D}_L = |\mathcal{D}_{L/\mathbb{Q}}|$ will refer to discriminant over $\mathbb{Q}$.

A folk conjecture, possibly due to Linnik, asserts that

$$N_{K,n}(X) \sim c_{K,n}X \qquad (n \text{ fixed}, X \to \infty).$$

This conjecture is trivial when $n = 2$; it has been proved for $n = 3$ by Davenport and Heilbronn [7] in case $K = \mathbb{Q}$, and by Datskovsky and Wright in general [6]; and for $n = 4, 5$ and $K = \mathbb{Q}$ by Bhargava [3], [2]. A weaker version of the conjecture for $n = 5$ was also recently established by Kable and Yukie [11]. These beautiful results are proved by methods which seem not to extend to higher n. The best upper bound for general n is due to Schmidt [18], who showed

$$N_{K,n}(X) \ll X^{(n+2)/4}$$

where the implied constant depends on K and n. We refer to [4] for a survey of results.

In many cases, it is easy to show that $N_{K,n}(X)$ is bounded below by a constant multiple of X; for instance, if n is even, simply consider the set of

*The first author was partially supported by NSA Young Investigator Grant MDA905-02-1-0097. The second author was partially supported by NSF Grant DMS-0245606.

quadratic extensions of a fixed L_0/K of degree $n/2$. For the study of lower bounds it is therefore more interesting to study the number of number fields L such that $[L:K]=n$, $\mathbf{N}_{\mathbb{Q}}^{K}\mathcal{D}_{L/K}<X$ and the Galois closure of L has Galois group S_n over K. Denote this number by $N'_{K,n}(X)$. Malle showed [14, Prop. 6.2] that

$$N'_{\mathbb{Q},n}(X) > c'_n X^{1/n}$$

for some constant c'_n.

The main result of this paper is to improve these bounds, with particular attention to the "large n limit." The upper bound lies much deeper than the lower bound.

Throughout this paper we will use $\ll$ and $\gg$ where the implicit constant depends on n; we will not make this n-dependency explicit (but see our appendix to [1] for results in this direction).

THEOREM 1.1. *For all $n>2$ and all number fields K, we have*

$$N_{K,n}(X) \ll (X\mathcal{D}_K^n A_n^{[K:\mathbb{Q}]})^{\exp(C\sqrt{\log n})}$$

where A_n is a constant depending only on n, and C is an absolute constant. Further,

$$X^{1/2+1/n^2} \ll_K N'_{K,n}(X).$$

In particular, for all $\varepsilon>0$

$$\limsup_{X\to\infty}\frac{\log N_{K,n}(X)}{\log X} \ll_\varepsilon n^\varepsilon, \quad \liminf_{X\to\infty}\frac{\log N'_{K,n}(X)}{\log X} \geq \frac{1}{2}+\frac{1}{n^2}. \tag{1.1}$$

Linnik's conjecture claims that the limit in (1.1) is equal to 1; thus, despite its evident imprecision, the upper bound in Theorem 1.1 seems to offer the first serious evidence towards this conjecture for large n. It is also worth observing that de Jong and Katz [9] have studied a problem of a related nature where the number field K is replaced by the function field $\mathbb{F}_q(T)$; even here, where much stronger geometric techniques are available, they obtain an exponent of the nature $c\log(n)$; a proof of this bound can be found in [8, Lemma 2.4]. This suggests that replacing n^ε in (1.1) by a constant will be rather difficult.

We will also prove various related results on the number of number fields with certain Galois-theoretic properties. For instance, if $G\leq S_n$, let $N_{K,n}(X;G)$ be the number of number fields L such that $[L:K]=n$, $\mathbf{N}_{\mathbb{Q}}^{K}\mathcal{D}_{L/K}<X$, and the action of $\mathrm{Gal}(\bar{K}/K)$ on embeddings $K\hookrightarrow\mathbb{C}$ is conjugate to the G-action on $\{1,\dots,n\}$. We describe how one can obtain upper bounds on $N_{K,n}(X;G)$ using the invariant theory of G. A typical example is:

PROPOSITION 1.2. *Let $G\leq S_6$ be a permutation group whose action is conjugate to the $\mathrm{PSL}_2(\mathbb{F}_5)$-action on $\mathbb{P}^1(\mathbb{F}_5)$. Then $N_{\mathbb{Q},6}(X;G)\ll_\varepsilon X^{8/5+\varepsilon}$.*

Specializing further, let $N_{K,n}(X;\mathrm{Gal})$ be the number of *Galois* extensions among those counted by $N_{K,n}(X)$; we prove the following upper bound.

PROPOSITION 1.3. *For each $n > 4$, one has $N_{K,n}(X;\mathrm{Gal}) \ll_{K,n,\varepsilon} X^{3/8+\varepsilon}$.*

In combination with the lower bound in Theorem 1.1, this shows that if one orders the number fields of fixed degree over $\mathbb{Q}$ by discriminant, a random one is not Galois.

Although we will use certain *ad hoc* tools, the central idea will always be to count fields by counting integral points on certain associated varieties, which are related to the invariant theory of the Galois group. These varieties must be well-chosen to obtain good bounds. In fact, the varieties we use are birational to the Hilbert scheme of r points in $\mathbb{P}^n$, suggesting the importance of a closer study of the distribution of rational points on these Hilbert schemes.

The results can perhaps be improved using certain techniques from the study of integral points, such as the result of Bombieri-Pila [15]. However, the proof of Theorem 1.1 turns out, somewhat surprisingly, to require only elementary arguments from the geometry of numbers and linear algebra.

Acknowledgments. The authors are grateful for the hospitality of the American Institute of Mathematics, where the first phase of this work was undertaken. We also thank Hendrik Lenstra for useful comments on an earlier draft.

2. Proof of upper bound

The main idea of Schmidt's proof is as follows: by Minkowski's theorem, an extension L/K contains an integer α whose archimedean valuations are all bounded by a function of $\Delta_L = \mathbf{N}_{\mathbb{Q}}^{K}\mathcal{D}_{L/K}$. Since all the archimedean absolute values are bounded in terms of Δ_L, so are the symmetric functions of these absolute values; in other words, α is a root of a monic polynomial in $\mathbb{Z}[x]$ whose coefficients have (real) absolute value bounded in terms of Δ_L. There are only finitely many such polynomials, and counting them gives the theorem of [18].

The main idea of Theorem 1.1 is to count r-tuples of integers in L instead of single integers.

Let $\mathbb{A}^n = \mathrm{Spec}(\mathbb{Z}[x_1, x_2, \ldots, x_n])$ denote affine n-space, which we regard as being defined over $\mathbb{Z}$. We fix an algebraic closure $\bar{K}$ of K. Let $\rho_1, \ldots, \rho_n$ be the embeddings of L into $\bar{K}$. Then the map $\phi_L = \rho_1 \oplus \cdots \oplus \rho_n$ embeds $\mathcal{O}_L$ in $\bar{K}^n = \mathbb{A}^n(\bar{K})$, and the direct sum of r copies of this map is an embedding $\mathcal{O}_L^r \to (\bar{K}^n)^r = (\mathbb{A}^n)^r(\bar{K})$ (which map we also, by abuse of notation, call ϕ_L).

The affine variety $(\mathbb{A}^n)^r$ is naturally coordinatized by functions $\{x_{j,k}\}_{1\le j\le n, 1\le k\le r}$. The symmetric group S_n acts on $(\mathbb{A}^n)^r$ by permuting $x_{1,k}, \ldots, x_{n,k}$ for each k. The S_n-invariants in the coordinate ring of $(\mathbb{A}^n)^r$

are called *multisymmetric functions*. If f is a multisymmetric function, the composition $f \circ \phi_L : \mathcal{O}_L^r \to \bar{K}$ takes image in $\mathcal{O}_K$. It follows that if $R \subset \mathbb{Z}[\{x_{j,k}\}_{1\leq j\leq n, 1\leq k\leq r}]^{S_n}$ is a subring of the ring of multisymmetric functions and $A = \mathrm{Spec}(R)$, there is a map of sets

$$\mathcal{F} : \bigcup_L \mathcal{O}_L^r \to A(\mathcal{O}_K)$$

where the union is over all number fields L with $[L : K] = n$.

Our overall strategy can now be outlined as follows. If x is algebraic over K, write $||x||$ for the maximum of the archimedean absolute values of x. For a positive real number Y, let $B(Y)$ be the set of algebraic integers x in $\overline{K}$ with degree n over K and $||x|| < Y$. Let $f_1, \dots, f_s \in \mathbb{Z}[\{x_{j,k}\}_{1\leq j\leq n, 1\leq k\leq r}]^{S_n}$ be multisymmetric functions with degrees $d_1, \dots, d_s$. Put $R = \mathbb{Z}[f_1, \dots, f_s]$, and set $A = \mathrm{Spec}(R)$. Then there is a constant c such that (for any Y) one has $||f_i(\phi_L(\alpha_1, \alpha_2, \dots, \alpha_r))|| < cY^{d_i}$ whenever $\alpha_j \in B(Y)\,(1 \leq j \leq r)$ and the α_j all belong to some subextension $L \subset \bar{K}, [L : K] = n$. Let $A(\mathcal{O}_K)_Y$ be the subset of $A(\mathcal{O}_K)$ consisting of points P such that $||f_i(P)|| < cY^{d_i}$. Then for any subset S_Y of $B(Y)^r$, we have a diagram of sets

(2.1)

$$\begin{array}{ccc} \{(L, \alpha_1, \alpha_2, \dots, \alpha_r) : [L : K] = n, \Delta_L < X, (\alpha_1, \dots, \alpha_r) \in (\mathcal{O}_L)^r \cap S_Y\} & \xrightarrow{\mathcal{F}} & A(\mathcal{O}_K)_Y \\ \Big\downarrow & & \\ \{L : [L : K] = n, \Delta_L < X\}. & & \end{array}$$

The cardinality of the lower set is precisely $N_{K,n}(X)$. Our goal is to choose $A = \mathrm{Spec}(R), Y$, and S_Y in such a way that that the vertical map in (2.1) is surjective (by Minkowski's theorem), while the horizontal map $\mathcal{F}$ has finite fibers whose cardinality we can bound. This will yield the desired bound on $N_{K,n}(X)$. Since $|A(\mathcal{O}_K)_Y| \ll_K (c^s Y^{\sum_i d_i})^{[K:\mathbb{Q}]}$, it should be our aim to choose $f_1, \dots, f_s$ whose total degree is as low as possible.

We begin with a series of lemmas about polynomials over an arbitrary characteristic-0 field F.

Let S be any test ring. We give $\mathbb{A}^n$ the structure of a ring scheme so that the ring structure on $\mathbb{A}^n(S) = S^n$ is the natural one. Let Tr be the map $\mathbb{A}^n \to \mathbb{A}^1$ which, on S-points, induces the map $(z_1, \dots, z_n) \in S^n \mapsto z_1 + \cdots + z_n \in S$. Given an element $\mathbf{x} = (x_{j,k})_{1\leq j\leq n, 1\leq k\leq r} \in (S^n)^r$, we denote by $\mathbf{x}_k \in S^n$ the k-th "row" $(x_{1,k}, x_{2,k}, \dots, x_{n,k})$, and by $\mathbf{x}^{(j)} \in S^r$ the j-th "column" $(x_{j,1}, x_{j,2}, \dots, x_{j,r})$. These correspond to maps $\mathbf{x} \mapsto \mathbf{x}_k : (\mathbb{A}^n)^r \to \mathbb{A}^n, \mathbf{x} \mapsto \mathbf{x}^{(j)} : (\mathbb{A}^n)^r \to \mathbb{A}^r$.

Let $\sigma = (i_1, \dots, i_r)$ be an element of $\mathbb{Z}^r_{\geq 0}$; we will think of $\mathbb{Z}^r_{\geq 0}$ as an additive semigroup, operations being defined pointwise. Then σ defines a S_n-equivariant map $\chi_\sigma : (\mathbb{A}^n)^r \to \mathbb{A}^n$ by the rule

$$\chi_\sigma(\mathbf{x}) = \mathbf{x}_1^{i_1} \mathbf{x}_2^{i_2} \dots \mathbf{x}_r^{i_r}.$$

Here $\mathbf{x}_1^{i_1}$ denotes $\mathbf{x}_1$ raised to the i_1-th power, i.e. $\mathbf{x}_1^{i_1} = \mathbf{x}_1 \times \mathbf{x}_1 \cdots \times \mathbf{x}_1$ (i times), the "multiplication" being taken with respect to ring-scheme structure on $\mathbb{A}^n$.

In particular, $F^n = \mathbb{A}^n(F)$ has a ring structure, and Tr, χ_σ induce maps on F-points, namely $\mathrm{Tr} : F^n \to F, \chi_\sigma : (F^n)^r \to F^n$; we abuse notation and use the same symbols for these maps. The map $(x, y) \mapsto \mathrm{Tr}(xy)$ is a nondegenerate pairing on F^n, with respect to which we can speak of "orthogonal complement".

LEMMA 2.1. *Let $\mathbf{x} \in (F^n)^r$, and let Σ_0 be a subset of $\mathbb{Z}^r_{\geq 0}$ such that the $|\Sigma_0|$ vectors $\chi_\sigma(\mathbf{x})_{\sigma\in\Sigma_0}$ generate a subspace of F^n (considered as an F-vector space) of dimension greater than $n/2$.*

Denote by $\Sigma_1 = \Sigma_0 + \Sigma_0$ the set of sums of two elements of Σ_0. Let $W \subset F^n$ be the subspace of F^n spanned by $\chi_\sigma(\mathbf{x})_{\sigma\in\Sigma_1}$. Then the orthogonal complement of W is contained in a coordinate hyperplane $x_j = 0$ for some j.

Proof. Write m for $|\Sigma_0|$ and let $v_1, \ldots, v_m$ be the vectors $\chi_\sigma(\mathbf{x})$ as σ ranges over Σ_0. Then W is the space spanned by the products $v_a v_b$ (the algebra structure on F^n being as noted above). Suppose w is orthogonal to W. Then

$$\mathrm{Tr}(v_a v_b w) = 0 \tag{2.2}$$

for all a, b; if V is the space spanned by the $\{v_a\}$, then (2.2) implies that wV and V are orthogonal. This implies in turn that $\dim wV \leq n - \dim V < \dim V$, so multiplication by w is not an automorphism of F^n; in other words, w lies on a coordinate hyperplane. A subspace of F^n contained in a union of coordinate hyperplanes is contained in a single coordinate hyperplane; this completes the proof. □

For each $\sigma \in \mathbb{Z}^r_{\geq 0}$, let $f_\sigma : (\mathbb{A}^n)^r \to \mathbb{A}^1$ be the composition $\mathrm{Tr} \circ \chi_\sigma$. Then f_σ is a multisymmetric function. When Σ is a subset of $\mathbb{Z}^r_{\geq 0}$, we denote by R_Σ the subring of functions on $(\mathbb{A}^n)^r$ generated by $\{f_\sigma\}_{\sigma\in\Sigma}$. One has a natural map of affine schemes

$$F_\Sigma : (\mathbb{A}^n)^r \to \mathrm{Spec}\, R_\Sigma. \tag{2.3}$$

The goal of the algebro-geometric part of our argument is to show that, by choosing Σ large enough, we can guarantee that F_Σ is generically finite, and even place some restrictions on the locus in $(\mathbb{A}^n)^r$ where F_Σ has positive-dimensional fibers.

LEMMA 2.2. *Let $\mathbf{x}$ be a point of $(\mathbb{A}^n)^r(F)$, and let Σ_1 be a subset of $\mathbb{Z}^r_{\geq 0}$ such that the $|\Sigma_1|$ vectors $\chi_\sigma(\mathbf{x})_{\sigma\in\Sigma_1}$ span F^n as an F-vector space. For each k between 1 and r let $e_k \in \mathbb{Z}^r_{\geq 0}$ be the vector with a 1 in the k-th coordinate and 0's elsewhere. Let Σ be a set which contains $\Sigma_1 + \Sigma_1$, and $\Sigma_1 + e_k$ for all k.*

Then the preimage $F_\Sigma^{-1}(F_\Sigma(\mathbf{x})) \subset (\mathbb{A}^n)^r(F)$ is finite, of cardinality at most $(n!)^r$.

Proof. Let $\mathbf{y}$ be $F_\Sigma(\mathbf{x})$. Let $m = |\Sigma_1|$. As in the proof of Lemma 2.1, let $v'_1, \dots, v'_m$ be the image of $\mathbf{x}$ under the $\{\chi_\sigma\}_{\sigma\in\Sigma_1}$. We may suppose by relabeling that $v'_1, \dots, v'_n$ form a basis for F^n (as an F-vector space). Since Σ contains $\Sigma_1 + \Sigma_1$, the determination of $\mathbf{y}$ fixes $\mathrm{Tr}(v'_a v'_b)$ for all a, b; and since Σ contains $\Sigma_1 + e_k$, we also know the traces $\mathrm{Tr}(v'_a \mathbf{x}_k)$ for all a and k. It follows that, for each k, we can represent the action of multiplication by $\mathbf{x}_k$ on the F-vector space spanned by $v'_1, \dots, v'_n$ by a matrix whose coefficients are determined by $\mathbf{y}$. But such a matrix evidently determines $\mathbf{x}_k$ up to permutation of coordinates; this proves the desired result. □

In the proof of Proposition 2.5 below, we will need to show that, by allowing $\mathbf{x}$ to vary over certain subspaces of $(F^n)^r$, we can ensure that $\mathbf{x}$ can be chosen in order to verify the hypothesis of Lemma 2.1.

LEMMA 2.3. *Let V be a F-subspace of F^n of dimension m, and let $\Sigma_0 \subset \mathbb{Z}^r_{\geq 0}$ be a subset of size m. Let $Z \subset V^r$ be the subset of points $\mathbf{x} \in V^r$ such that the m vectors $\chi_\sigma(\mathbf{x})_{\sigma\in\Sigma_0}$ are not linearly independent (over F) in F^n. Then Z is not the whole of V^r. If one identifies V^r with F^{mr}, Z is contained in the F-points of a hypersurface, defined over F, whose degree is bounded by a constant depending only on n and Σ_0.*

Proof. We may assume (by permuting coordinates) that the map "projection onto the first m coordinates," which we denote $\pi : F^n \to F^m$, induces an isomorphism $V \cong F^m$. Suppose there is a nontrivial linear relation

$$\sum_{\sigma\in\Sigma} c_\sigma \chi_\sigma(\mathbf{x}) = \mathbf{0} \in F^n, \tag{2.4}$$

that is, suppose $\mathbf{x} \in Z$. Each $\sigma \in \Sigma_0$ also defines a map $F^r \to F$ (derived from the map $\chi_\sigma : (\mathbb{A}^n)^r \to \mathbb{A}^n$ with $n = 1$) so we may speak of $\chi_\sigma(\mathbf{x}^{(j)}) \in F$ for $1 \leq j \leq n$. By abuse of notation we also use π to denote the projection of $(F^n)^r$ onto $(F^m)^r$. Then the restriction of π to V^r is an isomorphism $V^r \cong F^{mr}$.

Any nontrivial linear relation between the $\chi_\sigma(\mathbf{x})$ yields a nontrivial relation between the m vectors $\chi_\sigma(\pi(\mathbf{x}))$ in F^m. This in turn implies vanishing of the determinant

$$D = \begin{vmatrix} \chi_{\sigma_1}(\mathbf{x}^{(1)}) & \cdots & \chi_{\sigma_1}(\mathbf{x}^{(m)}) \\ \vdots & & \vdots \\ \chi_{\sigma_m}(\mathbf{x}^{(1)}) & \cdots & \chi_{\sigma_m}(\mathbf{x}^{(m)}) \end{vmatrix}.$$

The contribution of each $m \times m$ permutation matrix to D is a distinct monomial in the mr variables, so D is not identically 0 in $F[x_{1,1}, \dots, x_{m,r}]$. Evidently the degree of D is bounded in terms of n and Σ_0. Let $V(D)$ be the vanishing locus of D in $(F^m)^r$. Now the locus in Z is contained in $\pi^{-1}(V(D))$, which yields the desired result. □

Finally, we need a straightforward fact about points of low height on the complements of hypersurfaces.

LEMMA 2.4. *Let f be a polynomial of degree d in variables $x_1, \dots, x_n$. Then there exist integers $a_1, \dots, a_n$ such that $\max_{1 \le i \le n} |a_i| \le (1/2)(d+1)$ and $f(a_1, \dots, a_n) \neq 0$.*

Proof. There are at most d hyperplanes on which f vanishes, which means that the function $g(x_2, \dots, x_n) := f(a_1, x_2, \dots, x_n)$ is not identically 0 for some a_1 with absolute value at most $(1/2)(d+1)$. Now proceed by induction on n. □

Now we are ready for the key point in the proof of Theorem 1.1. The point is to use the lemmas above to construct Σ which is small enough that $\operatorname{Spec} R_\Sigma$ has few rational points of small height, but which is large enough so that F_Σ does not have too many positive-dimensional fibers.

PROPOSITION 2.5. *Let Σ_0 be a subset of $\mathbb{Z}^r_{\ge 0}$ of size $m > n/2$; let $\Sigma_1 \subset \mathbb{Z}^r_{\ge 0}$ contain $\Sigma_0 + \Sigma_0$; and let $\Sigma \subset \mathbb{Z}^r_{\ge 0}$ contain $\Sigma_1 + \Sigma_1$ and $\Sigma_1 + e_k$ for all k.*

Let L be a finite extension of K with $[L : K] = n$. Then there is an r-tuple $(\alpha_1, \dots, \alpha_r) \in \mathcal{O}_L^r$ such that

- *For every k,*

$$||\alpha_k|| \ll_\Sigma \mathcal{D}_L^{1/d(n-2)},$$

  where $d = [K : \mathbb{Q}]$.

- *The set $F_\Sigma^{-1}(F_\Sigma((\phi_L(\alpha_1, \dots, \alpha_r)))) \subset (\mathbb{A}^n)^r(\bar{K})$ has cardinality at most $(n!)^r$.*

- *The elements $\alpha_1, \dots, \alpha_r$ generate the field extension L/K.*

Proof. First of all, note that if $(\alpha_1, \dots, \alpha_r), \Sigma_0, \Sigma_1, \Sigma$ satisfy the conditions above, then so do $(\alpha_1, \dots, \alpha_r), \Sigma_0', \Sigma_1, \Sigma$ for any subset $\Sigma_0' \subset \Sigma_0$ with $|\Sigma_0'| > n/2$. So it suffices to prove the theorem in case $n/2 < m \le (n/2 + 1)$.

Let $1 = \beta_1, \dots, \beta_{nd}$ be a $\mathbb{Q}$-linearly independent set of integers in $\mathcal{O}_L$ such that $||\beta_i||$ is the i-th successive minimum of $||\cdot||$ on $\mathcal{O}_L$, in the sense of Minkowski's second theorem [20, III, §3]. The K-vector space spanned by $\beta_1, \dots, \beta_{md}$ has K-dimension at least m, so we may choose $\gamma_1, \dots, \gamma_m$ among the β_i which are linearly independent over K.

Let $V \subset \bar{K}^n$ be the $\bar{K}$-vector space spanned by $\{\phi_L(\gamma_i)\}_{1 \le i \le m}$. Then by Lemma 2.3 there is a constant C_{n,Σ_0} and a hypersurface $Z \subset V^r$ of degree C_{n,Σ_0} such that, for all $\mathbf{x}$ not in $Z(\bar{K})$, the m vectors $\chi_\sigma(\mathbf{x})_{\sigma \in \Sigma_0}$ are $\bar{K}$-linearly independent in $\bar{K}^n$.

For every field M strictly intermediate between K and L, we let $V_M \subset \bar{K}^n$ be the $\bar{K}$-vector subspace $\phi_L(M) \subset \bar{K}^n$. Each $(V \cap V_M)^r$ is a certain linear subspace of V^r; note that, since $m > n/2$, no subspace V_M contains V. Let Z' be the union of $Z(\bar{K})$ with $(V \cap V_M)^r$, as M ranges over all fields between K and L.

Now let Y be a hypersurface of V^r so that $Y(\bar{K})$ contains Z'; one may choose Y so that the degree of Y is bounded in terms of n and Σ_0. By Lemma 2.4, there is a constant H, depending only on n and Σ_0, so that, for any lattice $\iota : \mathbb{Z}^{mr} \hookrightarrow V^r$ (i.e. we require $\iota(\mathbb{Z}^{mr})$ spans V^r over $\bar{K}$) there is a point $p \in \mathbb{Z}^{mr}$, with $\iota(p) \notin Y(\bar{K})$, whose coordinates have absolute value at most H.

It follows that there exists a set of mr integers $c_{1,1}, \dots, c_{m,r}$ with $|c_{j,k}| \leq H$, such that

$$\mathbf{x} = (\phi_L(c_{1,1}\gamma_1 + \cdots + c_{m,1}\gamma_m), \dots, \phi_L(c_{1,r}\gamma_1 + \cdots + c_{m,r}\gamma_m))$$

is not in $Y(\bar{K})$. For each k between 1 and r define $\alpha_k \in \mathcal{O}_L$ via

$$\alpha_k = c_{1,k}\gamma_1 + \cdots + c_{m,k}\gamma_m.$$

Let $W \subset L$ be the K-subspace spanned by $\chi_\sigma(\alpha_1, \dots, \alpha_r)$ as σ ranges over Σ_1 (here we regard χ_σ as a map $L^r \to L$, cf. remarks after (2.4)). Suppose W is not the whole of L. Then there is a nonzero element $t \in L$ such that $\mathrm{Tr}_K^L tw = 0$ for all $w \in W$. It follows that $\phi_L(t) \in \bar{K}$ lies in the orthogonal complement (w.r.t. the form Tr on $\bar{K}^n$) of $\phi_L(W) \subset \bar{K}^n$. But the orthogonal complement to the $\bar{K}$-span of $\phi_L(W)$ is contained in a coordinate hyperplane by Lemma 2.1. Since $\rho_j(t)$ cannot be 0 for any j and any nonzero t, this is a contradiction; we conclude that $W = L$, and thus that the vectors $\{\chi_\sigma(\mathbf{x})\}_{\sigma \in \Sigma_1}$ span L as a K-vector space.

The bound on the size of the fiber $F_\Sigma^{-1}(F_\Sigma(\mathbf{x}))$ follows from Lemma 2.2, and the fact that $\mathbf{x} \notin V_M^r$ for any M implies that $\alpha_1, \dots, \alpha_r$ generate the extension L/K.

It remains to bound the archimedean absolute values of the α_i. The image of $\mathcal{O}_L$ in $\mathcal{O}_L \otimes_{\mathbb{Z}} \mathbb{R}$ is a lattice of covolume $\mathcal{D}_L^{1/2}$, so by Minkowski's second theorem [20, Th. 16],

$$\prod_{i=1}^{nd} ||\beta_i|| \leq \mathcal{D}_L^{1/2}.$$

The $||\beta_i||$ form a nondecreasing sequence, so for $m < n$, we have

$$||\beta_{md}||^{(n-m)d} \leq \prod_{i=md+1}^{nd} ||\beta_i|| \leq \mathcal{D}_L^{1/2}.$$

Since $m \leq (1/2)n + 1$, we get

$$||\beta_i|| < (\mathcal{D}_L)^{1/d(n-2)}$$

for all $i \le m$. It follows that all archimedean absolute values of γ_i for $i \le m$ are bounded by a constant multiple of $\mathcal{D}_L^{1/d(n-2)}$, the implicit constant being absolute. The result follows, since each α_k is an integral linear combination of the γ_i with coefficients bounded by H. □

We are now ready to prove the upper bound in Theorem 1.1; what remains is merely to make a good choice of Σ and apply Proposition 2.5. Let r and c be positive integers such that $\binom{r+c}{r} > n/2$, and let Σ_0 be the set of all r-tuples of nonnegative integers with sum at most c. We shall choose r, c in the end; but r, c, Σ_0, Σ will all depend only on n, so that all constants that depend on them in fact depend only on n.

Now take Σ to be the set of all r-tuples of nonnegative integers with sum at most $4c$, and consider the map

$$F_\Sigma : (\mathbb{A}^n)^r \to \operatorname{Spec} R_\Sigma.$$

By Proposition 2.5, to every field L with $[L : K] = n$ we can associate an r-tuple $(\alpha_1, \dots, \alpha_r)$ of integers satisfying the three conditions in the statement of the proposition. Define $Q_L \in (\mathbb{A}^n)^r(\bar{K})$ to be $\phi_L(\alpha_1, \dots, \alpha_r)$, and let $P_L \in \operatorname{Spec} R_\Sigma(\mathcal{O}_K)$ be the point $F_\Sigma(Q_L)$.

By the second condition on $\alpha_1, \dots, \alpha_r$, there are at most $(n!)^r$ points in $F_\Sigma^{-1}(P_L)$. By the third condition, $Q_L = Q_{L'}$ only if L and L' are isomorphic over K. We conclude that at most $(n!)^r$ fields L are sent to the same point in $\operatorname{Spec} R_\Sigma(\mathcal{O}_K)$.

We now restrict our attention to those fields L satisfying

$$\mathbf{N}_{\mathbb{Q}}^K \mathcal{D}_{L/K} < X.$$

In this case, for every archimedean valuation $|\cdot|$ of L and every $k \le r$ we have the bound

$$|\alpha_k| \ll \mathcal{D}_L^{1/d(n-2)} \ll (X\mathcal{D}_K^n)^{1/d(n-2)}. \tag{2.5}$$

Now, f_σ being as defined prior to (2.3), $f_\sigma(Q_L)$ is an element of $\mathcal{O}_K$, which (by choice of Σ) we can express as a polynomial of degree at most $4c$ (and absolutely bounded coefficients) in the numbers $\rho_j(\alpha_k) \in \bar{K}$. If $|\cdot|$ is any archimedean absolute value on K, we can extend $|\cdot|$ to a archimedean absolute value on L, and by (2.5) we have

$$|f_\sigma(Q_L)| \ll (X\mathcal{D}_K^n)^{4c/d(n-2)}.$$

The number of elements of $\mathcal{O}_K$ with archimedean absolute values at most B is $\le (2B+1)^d$. (For large enough B, one can save an extra factor of $\mathcal{D}_K^{1/2}$; this is not necessary for our purpose.) In view of the above equation, the number of possibilities for $f_\sigma(Q_L)$ is $\ll (X\mathcal{D}_K^n A_n^d)^{4c/(n-2)}$ where A_n is a constant depending only on n.

Now the point $P_L \in \operatorname{Spec} R_\Sigma(\mathcal{O}_K)$ is determined by $f_\sigma(Q_L)$ $(\sigma \in \Sigma)$ and we have $|\Sigma| = \binom{r+4c}{r}$. The number of possibilities for P_L is therefore $\ll (X\mathcal{D}_K^n A_n^d)^{(4c/(n-2))\binom{r+4c}{r}}$.

Since each number field L contributes a point to this count, and since no point is counted more than $(n!)^r$ times, we have

$$N_{K,n}(X) \ll (X\mathcal{D}_K^n A_n^d)^{(4c/(n-2))\binom{r+4c}{r}}. \tag{2.6}$$

Now is a suitable time to optimize r and c. We may assume $n \geq 3$. Take r to be the greatest integer $\leq \sqrt{\log(n)}$, and choose c to be the least integer $\geq (nr!)^{1/r}$. Note that $c \geq n^{1/r} \geq e^{\sqrt{\log(n)}} \geq e^r \geq r$ and $c \leq 2(nr!)^{1/r}$. Then $\binom{r+c}{r} > c^r/r! \geq n$ whereas $\binom{r+4c}{r} \leq \binom{5c}{r} \leq \frac{(5c)^r}{r!} \leq 10^r n$. Substituting these values of r, c into (2.6) yields the upper bound of Theorem 1.1.

In the language of the beginning of this section, we have taken A to be $\operatorname{Spec} R_\Sigma$ and the map $\mathcal{F}$ to be F_Σ. The set S_Y can be taken to be the set of r-tuples of integers $\alpha_1, \ldots, \alpha_r$ so that $\alpha_j \in B(Y)(1 \leq j \leq r)$, and so that there exists a subextension $L \subset \overline{K}, [L : K] = n$ with $\alpha_j \in L$ and such that $\phi_L(\alpha_1, \ldots, \alpha_r) \in V^r - Z'$ (notation of proof of Proposition 2.5). Minkowski's theorem guarantees that each number field L contains an r-tuple of integers in S_Y for some reasonably small Y, while the lemmas leading up to Proposition 2.5 show that the fibers of F containing a point of S_Y have cardinality at most $(n!)^r$.

Another way to think of the method is as follows: we can factor F_Σ as

$$(\mathbb{A}^n)^r \to X = (\mathbb{A}^n)^r/S_n \to A = \operatorname{Spec} R_\Sigma$$

where the intervening quotient is just the affine scheme associated to the ring of multisymmetric functions. Every r-tuple of integers in $\mathcal{O}_L$ corresponds to an integral point of X; however, the fact that X fails to embed naturally in a low-dimensional affine space makes it difficult to count points of $X(\mathbb{Z})$ with bounded height. The method used here identifies a locus $W \subset X$ which is contracted in the map to $\operatorname{Spec} R_\Sigma$, and shows that the map $X(\mathbb{Z}) \to A(\mathbb{Z})$ has fibers of bounded size away from W; this gives an upper bound on the number of integral points on $X \backslash W$ of bounded height. One might ask whether the estimates on rational points of bounded height predicted by the Batyrev-Manin conjecture could be applied to X. Any such prediction would lead to a refinement of our upper bound on the number of number fields.

2.1. *Improvements, invariant theory, and the large sieve.*

Remark 2.6. The method we have used above may be optimized in various ways: by utilizing more of the invariant theory of S_n, and by using results about counting integral points on varieties. These techniques may be used, for any fixed n, to improve the exponent in the upper bound of Theorem 1.1.

(The invariant theory, however, becomes more computationally demanding as n increases.) However, they do not change the limiting behavior as $n \to \infty$. We have therefore chosen to present a different example of this optimization: giving good bounds on $N_{K,n}(X; G)$ for $G \neq S_n$. For simplicity of exposition we take $K = \mathbb{Q}$.

Example 2.7. Let $G = \langle (1,6,2)(3,4,5), (5,6)(3,4) \rangle$; it is a primitive permutation group on $\{1,2,3,4,5,6\}$ whose action is conjugate to the action of $\mathrm{PSL}_2(\mathbb{F}_5)$ on $\mathbb{P}^1(\mathbb{F}_5)$.

We will show $N_{\mathbb{Q},6}(X; G) \ll_\varepsilon X^{8/5+\varepsilon}$, a considerable improvement over Schmidt's bound of X^2 (over which, in turn, Theorem 1.1 presents no improvement for $n = 6$).

Let G act on monomials $x_1, x_2, \ldots, x_6$ by permutation of the indices. Set $f_i = \sum_{j=1}^6 x_j^i$ for $1 \leq i \leq 5$, and $f_6 = x_1x_2(x_3+x_4)+x_1x_3x_5+x_1x_4x_6+x_1x_5x_6+x_2x_3x_6 + x_2x_4x_5 + x_2x_5x_6 + x_3x_4(x_5 + x_6)$. Set $A = \mathbb{C}[f_1, f_2, \ldots, f_6]$. Then $R = \mathbb{C}[x_1, \ldots, x_6]^G$ is a free A-module of degree 6; indeed $R = \oplus_{i=1}^6 A \cdot g_i$, where $g_1 = 1$ and $g_2, g_3, \ldots, g_6$ can be chosen to be homogeneous of degree $5, 6, 6, 7, 12$. (This data was obtained with the commands `InvariantRing`, `PrimaryInvariants`, and `SecondaryInvariants` in Magma.) One checks that $\overline{R} = R/f_1R$ is an integral domain.

Let S be the subring of $\overline{R}$ generated by $\overline{f_2}, \ldots, \overline{f_6}$ and $\overline{g_2}$, and let $Z = \mathrm{Spec}(S)$. S is an integral domain since $\overline{R}$ is; thus Z is irreducible. The map $\mathbb{C}[f_2, f_3, f_4, f_5, f_6] \to S$ induces a finite projection $Z \xrightarrow{\Pi} \mathbb{A}^5$ (it is finite since R is finite over A, so $\overline{R}$ is finite over $\mathbb{C}[f_2, f_3, \ldots, f_6]$). Also $\overline{g_2} \notin \mathbb{C}[\overline{f_2}, \ldots, \overline{f_6}]$, as follows from the fact that $R = \oplus_{i=1}^6 Ag_i$; thus the degree of Π is at least 2.

Suppose L is a number field with $[L : \mathbb{Q}] = 6$ with Galois group G and $\mathcal{D}_L < X$. Minkowski's theorem implies there exists $x \in \mathcal{O}_L$ with $\mathrm{Tr}_{\mathbb{Q}}^L(x) = 0$ and $||x|| \ll X^{1/10}$; here $||x||$ is defined as in the proof of Proposition 2.5. The element $x \in \mathcal{O}_L$ gives rise to a point $\mathbf{x} \in Z(\mathbb{Z})$ whose projection $\Pi(\mathbf{x}) = (y_1, y_2, y_3, y_4, y_5) \in \mathbb{Z}^5$ satisfies:

$$|y_1| \ll X^{2/10}, |y_2| \ll X^{3/10}, |y_3| \ll X^{4/10}, |y_4| \ll X^{5/10}, |y_5| \ll X^{3/10}. \tag{2.7}$$

We must count integral points on Z whose projection to $\mathbb{A}^5$ belong to the skew-shaped box defined by (2.7). It is clear that the number of points on $Z(\mathbb{Z})$ projecting to the box (2.7) is at most $X^{17/10}$, but applying the large sieve to the map $Z \xrightarrow{\Pi} \mathbb{A}^5$ (cf. [5] or [19]) one obtains the improved bound $X^{8/5+\varepsilon}$. (Note that the results, for example in [19], are stated only for a "square" box (all sides equal) around the origin — but indeed they apply, with uniform implicit constant, to a square box centered at *any* point. Now we tile the skew box (2.7) by square boxes of side length $X^{2/10}$ to obtain the claimed result.)

One expects that one can quite considerably improve this bound given more explicit understanding of the variety Z; ideally speaking one would like

to slice it, show that most slices are geometrically irreducible, and apply the Bombieri-Pila bound [15]. It is the intermediate step — showing that very few slices have irreducible components of low degree — which is difficult. This seems like an interesting computational question.

We remark that this particular example can also be analyzed by constructing an associated *quintic* extension (using the isomorphism of $\mathrm{PSL}_2(\mathbb{F}_5)$ with A_5) and counting these quintic extensions. This is close in spirit to the idea of the next section; in any case the method outlined above should work more generally.

2.2. *Counting Galois extensions.* In this section, we give bounds on the number of Galois extensions of $\mathbb{Q}$ with bounded discriminant. In combination with the lower bound in Theorem 1.1 for the total number of extensions, this yields the fact that "most number fields, counted by discriminant, are not Galois."

Let K be a number field of degree d over $\mathbb{Q}$ and G a finite group; we denote by $N_K(X, G)$ the number of Galois extensions of K with Galois group G such that

$$\mathbf{N}_{\mathbb{Q}}^{K} \mathcal{D}_{L/K} < X.$$

PROPOSITION 2.8. *If* $|G| > 4$, *then* $N_K(X, G) \ll_{K,G,\varepsilon} X^{3/8+\varepsilon}$.

Remark 2.9. Proposition 2.8 is not meant to be sharp; our aim here is merely to show that most fields are not Galois, so we satisfy ourselves with giving a bound smaller than $X^{1/2}$. In fact, according to a conjecture of Malle [14], $N_K(X, G)$ should be bounded between $X^{\frac{\ell}{(\ell-1)|G|}}$ and $X^{\frac{\ell}{(\ell-1)|G|}+\varepsilon}$, where ℓ is the smallest prime divisor of $|G|$. This conjecture is true for all abelian groups G by a theorem of Wright [21], and is proved for all nilpotent groups in a paper of Klüners and Malle [13].

Remark 2.10. The proof of the proposition depends on the fact that any finite simple group S has a proper subgroup H with $|H| > \sqrt{|S|}$. This may be verified directly from the classification of finite simple groups. A much weaker result would also suffice if we used Theorem 1.1 in place of Schmidt's result in the argument.

Proof. We proceed by induction on $|G|$. In this proof, all implicit constants in $\ll, \gg$ depend on K, ε and G, although we do not always explicitly note this.

Write an exact sequence

$$1 \to H \to G \to Q \to 1$$

where H is a minimal normal subgroup of G. Then H is a direct sum of copies of some simple group [16, 3.3.15].

Suppose that L/K is a Galois extension with $G_{L/K} \cong G$ and $\mathbf{N}_{\mathbb{Q}}^{K}\mathcal{D}_{L/K} < X$. Fixing an isomorphism of $G_{L/K}$ with G, let M be the subfield of L fixed by H. Then M/K is a Galois extension with Galois group Q and $N_{\mathbb{Q}}^{K}\mathcal{D}_{M/K} < X^{1/|H|}$. The number of such extensions M/K is $N_K(X^{1/|H|}, Q)$, which by the induction hypothesis is $\ll_Q X^{3/8|H|+\varepsilon}$ in case $|Q| > 4$. If $|Q| \leq 4$, then Q is abelian and by Wright's result [21] $N_K(X^{1/|H|}, Q) \ll X^{1/|H|+\varepsilon}$.

Now take M/K to be fixed; then the number of choices for L is bounded above by $N_M(X, H)$.

First, suppose H is not abelian. Let H_0 be a proper subgroup of H that does not contain any normal subgroups of H and is of maximal cardinality subject to this restriction.

If L is any H-extension of M, fix an isomorphism of $G_{L/M}$ with H and set $L' = L^{H_0}$; then L is the normal closure of L' over M. Further, $\mathcal{D}_{L'} \leq (\mathcal{D}_L)^{1/|H_0|} \ll_K (\mathbf{N}_{\mathbb{Q}}^{K}\mathcal{D}_{L/K})^{1/|H_0|}$. It follows from the main theorem of [18] that the number of possibilities for L' (and hence the number of possibilities for L), *given* M, is $\ll X^{\frac{(|H|/|H_0|+2)}{4|H_0|}}$, where the implicit constant is independent of M.

The group H_0 can be chosen to have size at least $\sqrt{|H|}$ (cf. [12], comments after 5.2.7) so summing over all choices of M, we find

$$N_K(X, G) \ll_G X^{1/4+1/(2\sqrt{|H|})+1/|H|+\varepsilon}$$

which, since $|H| \geq 60$, proves Proposition 2.8 in case H is non-abelian.

Now, suppose H is abelian; so $H = (\mathbb{Z}/p\mathbb{Z})^r$ for some prime p and some positive integer r. By [21] we may assume $|Q| \geq 2$.

Let $b_M(Y)$ be the number of H-extensions of M such that $\mathbf{N}_{\mathbb{Q}}^{M}\mathcal{D}_{L/M} = Y$. Let S be the set of primes of $\mathbb{Q}$ dividing Y, let $G_S(M)$ be the Galois group of the maximal extension of M unramified away from primes dividing S, and for each prime λ of M let I_λ be the inertia group at λ. Then $b_M(Y) \leq |\operatorname{Hom}(G_S(M), H)|$. Moreover, the kernel of the map

$$\operatorname{Hom}(G_S(M), H) \to \bigoplus_{\lambda|S} \operatorname{Hom}(I_\lambda, H)$$

is isomorphic to a subgroup of the r-th power of the class group of M, and as such has cardinality $\ll_\varepsilon \mathcal{D}_{M/\mathbb{Q}}^{r/2+\varepsilon}$, by the easy part of the Brauer-Siegel theorem. On the other hand, the number of primes λ is $\ll |S|$, and $|\operatorname{Hom}(I_\lambda, H)|$ is bounded by some constant C depending only on $[M : \mathbb{Q}]$; so the image of the map above has cardinality at most

$$(C')^{|S|} \ll_{\varepsilon,K,G} Y^{\varepsilon}.$$

We conclude that

$$b_M(Y) \ll \mathcal{D}_{M/\mathbb{Q}}^{r/2+\varepsilon} Y^{\varepsilon}. \tag{2.8}$$

Let μ be a prime of K such that μ does not divide $|G|\mathcal{D}_{M/K}$ and primes of M above μ ramify in L. Then the image of $I_\mu \subset \mathrm{Gal}(\bar{K}/K)$ in G is a cyclic subgroup whose order is a multiple of p; it follows that $(p-1)|G|/p$ divides $\mathrm{ord}_\mu \mathcal{D}_{L/K}$. So $\mathbf{N}_{\mathbb{Q}}^M \mathcal{D}_{L/M}$ lies in one of a finite set of cosets of $\mathbb{Q}^*/(\mathbb{Q}^*)^{(p-1)|G|/p}$. Let Σ be this union of cosets. Since the valuation of $\mathbf{N}_{\mathbb{Q}}^M \mathcal{D}_{L/M}$ is divisible by $\frac{(p-1)|G|}{p}$ at primes not dividing $|G|\mathbf{N}_{\mathbb{Q}}^K \mathcal{D}_{M/K}$, it follows that we may take Σ so that the number of cosets in Σ is $\ll_{\varepsilon,G} (\mathbf{N}_{\mathbb{Q}}^K \mathcal{D}_{M/K})^{\varepsilon}$.

When M is a Q-extension of K, we write N_1 for $\mathbf{N}_{\mathbb{Q}}^K \mathcal{D}_{M/K}$. Then

$$N_K(X,G) \leq \sum_{M: N_1 \leq X^{1/|H|}} \sum_{N_2 < X N_1^{-|H|}, N_2 \in \Sigma} b_M(N_2).$$

The inner sum has length $\ll_\varepsilon N_1^{\varepsilon}(XN_1^{-|H|})^{p/(p-1)|G|}$, which, combined with (2.8), gives

$$\begin{aligned} N_K(X,G) &\ll_\varepsilon \sum_{M: N_1 \leq X^{1/|H|}} X^{p/(p-1)|G|+\varepsilon} N_1^{r/2-p/(p-1)|Q|+\varepsilon} \\ &\leq N_K(X^{1/|H|},Q) X^{p/(p-1)|G|+\varepsilon} \max_{N_1 < X^{1/|H|}} N_1^{r/2-p/(p-1)|Q|+\varepsilon} \\ &= N_K(X^{1/|H|},Q) X^{\alpha+\varepsilon} \end{aligned}$$

where $\alpha = \max(\frac{r}{2|H|}, \frac{p}{(p-1)|G|})$.

By the induction hypothesis, $N_K(X^{1/|H|},Q) \ll X^{3/8|H|+\varepsilon}$ when $|Q| \geq 5$, while $N_K(X^{1/|H|},Q)$ is asymptotic to $X^{1/2|H|}$ if $|Q| = 3,4$ and to $X^{1/|H|}$ when $|Q| = 2$. Define

$$\beta(Q) = \begin{cases} 3/8 & |Q| \geq 5; \\ 1/2 & |Q| = 3,4; \\ 1 & |Q| = 2. \end{cases}$$

Then

$$N_K(X^{1/|H|},Q) X^{r/2|H|} \ll X^{(r/2+\beta)/|H|+\varepsilon}$$

and the exponent $\frac{r/2+\beta}{|H|}$ is at most $3/8$ unless either $|H| = 2$, or $|Q| = 2$ and $|H| = 3,4$. In case $|Q| = 2, |H| = 4$, the group G is nilpotent and Proposition 2.8 is proved by Klüners and Malle.

On the other hand,

$$N_K(X^{1/|H|},Q) X^{p/(p-1)|G|} \ll X^{(p/(p-1)|Q|+\beta)/|H|+\varepsilon}.$$

Here, the exponent is once again at most $3/8$ unless either $|H| = 2$, or $|Q| = 2$ and $|H| = 3,4$.

We have thus proven Proposition 2.8 unless $G = S_3$ or $H = \mathbb{Z}/2\mathbb{Z}$. In the former case, the proposition follows from the theorem of Datskovsky and Wright [6] on the number of cubic extensions of number fields. (More precisely,

one may count Galois S_3-extensions by controlling the 3-class numbers of the quadratic subextensions; this can be done using the results of [6].) In the latter case, we can refine the argument above; let $b'_M(Y)$ be the number of quadratic extensions L/M which are preserved by the action of Q and so that $\mathbf{N}_{\mathbb{Q}}^M \mathcal{D}_{L/M} = Y$. Choosing S to consist of all divisors of $Y\mathbf{N}_{\mathbb{Q}}^K \mathcal{D}_{M/K}$ and utilising the inflation-restriction sequence

$$\mathrm{Hom}(G_S(K), \mathbb{Z}/2\mathbb{Z}) \to \mathrm{Hom}(G_S(M), \mathbb{Z}/2\mathbb{Z})^Q \to H^2(Q, \mathbb{Z}/2\mathbb{Z})$$

we see that $b'_M(Y) \ll (Y \cdot \mathbf{N}_{\mathbb{Q}}^K \mathcal{D}_{M/K})^{\varepsilon}$. (Here $G_S(K)$ is defined analogously to $G_S(M)$.) This saves a factor of $N_1^{r/2}$ throughout the rest of the argument, and in particular we have

$$N_K(X, G) \ll_{\varepsilon} N_K(X^{1/2}, Q) X^{2/|G|+\varepsilon}.$$

Since we may assume G non-nilpotent, we can take $|Q| \geq 6$, which yields

$$N_K(X, G) \ll_{\varepsilon} X^{3/16} X^{1/6+\varepsilon}$$

which again yields the desired result. □

3. Proof of lower bound for S_n extensions

We now turn to the (easier) question of proving the lower bounds for $N'_{K,n}(X)$ asserted in Theorem 1.1, and finish with a brief discussion of some related issues.

We make some preliminary remarks. Firstly, as was discussed in the introduction, this question is often much easier if one is counting extensions for G a *proper* subgroup of S_n (see Malle [14] for some examples). On the other hand, the general question of lower bounds subsumes the inverse Galois problem over $\mathbb{Q}$. The method we give can be generalized to G-extensions, so long as one can construct a family of polynomials with generic Galois group G (i.e. an element $p \in K[t_1, \ldots, t_k, X]$ such that $K(t_1, \ldots, t_n)[X]/(p)$ is a G-extension of $K(t_1, \ldots, t_n)$; of course the bound will depend on p).

As before, let K be a fixed extension of $\mathbb{Q}$ of degree d. We also set $\Delta_L = \mathbf{N}_{\mathbb{Q}}^K(\mathcal{D}_{L/K})$ and $\mathcal{O}_L^0 = \{x \in \mathcal{O}_L : \mathrm{Tr}_K^L(x) = 0\}$. In this section, we will not aim for any uniformity in K; the implicit constants in this section will always depend on K and n. As before, for x an algebraic number, we denote by $||x||$ the largest archimedean valuation of x.

LEMMA 3.1. *Let $[L : K] = n$ be so that L/K has no proper subextensions. Then $||x|| \gg \Delta_L^{\frac{1}{n(n-1)d}}$ for all $x \in \mathcal{O}_L^0$, $x \neq 0$.*

Proof. If $x \in \mathcal{O}_L^0$, then $O_K[x]$ is a subring of O_L which generates O_L as a K-vector-space, since L/K has no proper subextensions and x is not in

K; in particular, the discriminant $\mathcal{D}(\mathcal{O}_K[x])$ of $\mathcal{O}_K[x]$ over $\mathcal{O}_K$ is divisible by $\mathcal{D}_{L/K}$. In particular, $\mathbf{N}_{\mathbb{Q}}^K(\mathcal{D}_{L/K}) \leq \mathbf{N}_{\mathbb{Q}}^K \mathcal{D}(\mathcal{O}_K[x])$. $\mathcal{D}(\mathcal{O}_K[x])$ is the same as the discriminant of the characteristic polynomial of x; from this, one deduces that $\mathcal{D}(\mathcal{O}_K[x])$ is a principal ideal of $\mathcal{O}_K$, generated by a polynomial of degree $n(n-1)$ in the Galois conjugates of x. In particular, one deduces $\mathbf{N}_{\mathbb{Q}}^K(\mathcal{D}(\mathcal{O}_K[x])) \ll ||x||^{n(n-1)d}$, whence the assertion. □

See Remark 3.2 for generalizations.

In the lower bound proved below, we have not aimed to optimize the exponent $1/2+1/n^2$. It will be obvious from the proof that it can be improved somewhat, both by replacing Schmidt's upper bound with that of Theorem 1.1, and by utilizing successive maxima and Remark 3.2 rather than just Lemma 3.1. This seems like an interesting optimization question; the gain for small n can be significant although one does not obtain an exponent near 1.

Proof (*of lower bound* $N'_{K,n}(X) \gg_{K,n} X^{1/2+1/n^2}$ *in Theorem* 1.1). We fix as before an algebraic closure $\bar{K}$. Consider the set $S(Y)$ of algebraic integers $x \in \bar{K}$ so that $[K(x) : K] = n$, $\mathrm{Tr}_K^{K(x)}(x) = 0$ and $||x|| \leq Y$. Let $S(Y; S_n)$ be the subset of those x so that the Galois closure of $K(x)$ over K has Galois group S_n.

Then, by considering the characteristic polynomial, we see that $|S(Y)| \gg Y^{d(n(n+1)/2-1)}$. Considering (the proof of) Hilbert's irreducibility theorem, we see that the same bound holds for $S(Y; S_n) \subset S(Y)$:

$$|S(Y; S_n)| \gg Y^{d(\frac{n(n+1)}{2}-1)} = Y^{\frac{(n-1)(n+2)d}{2}}. \tag{3.1}$$

Indeed one may put a congruence constraint on the characteristic polynomial to guarantee that the Galois closure has group S_n (cf. [17]).

Suppose L is an S_n-extension of K (i.e. $[L : K] = n$ and the Galois closure of L/K has Galois group S_n). $\mathcal{O}_L^0$ is a free $\mathbb{Z}$-module of rank $(n-1)d$; then Lemma 3.1 guarantees that the number of $x \in S(Y; S_n)$ such that $K(x) \cong L$ is $\ll (\frac{Y}{\Delta_L^{1/n(n-1)d}})^{(n-1)d}$; in particular if there is at least one such x, one must have $\Delta_L \ll Y^{n(n-1)d}$. Combining these comments with (3.1) we find that for some constant c:

$$\sum_{\substack{L:\Delta_L \leq cY^{n(n-1)d} \\ L/K\ S_n-\text{extension}}} \left(\frac{1}{\Delta_L}\right)^{\frac{1}{n}} \gg Y^{dn(n-1)/2}. \tag{3.2}$$

However, Schmidt's upper bound $N_{K,n}(X) \ll X^{(n+2)/4}$ easily shows that

$$\sum_{\substack{L:\Delta_L < Y^{d(n-1)} \\ [L:K]=n}} \left(\frac{1}{\Delta_L}\right)^{\frac{1}{n}} \ll Y^{\frac{dn(n-1)}{2}-\delta}$$

for some $\delta > 0$; thus one can replace the range of summation in (3.2) by $Y^{d(n-1)} < \Delta_L \leq cY^{dn(n-1)}$ without changing the result. In particular $N'_{K,n}(cY^{dn(n-1)}) \gg Y^{dn(n-1)(\frac{1}{2}+\frac{1}{n^2})}$, which implies the result. □

Remark 3.2 (Shape of number field lattices). Lemma 3.1 emphasizes the importance of understanding the shape of number field lattices. For clarity, fix attention on totally real number fields of degree $n \geq 3$ over $\mathbb{Q}$ with no proper subfields; one can formulate similar ideas in the general case.

Let L be such a number field. Then $\mathcal{O}_L^0$ is a lattice endowed with a natural quadratic form, namely $x \mapsto \mathrm{tr}(x^2)$; as such, it defines an element $[L]$ of the moduli space $\mathfrak{S}$ of *homothety classes of positive definite quadratic forms*. $\mathfrak{S}$ can be identified with $\mathrm{PGL}_{n-1}(\mathbb{Z})\backslash \mathrm{PGL}_{n-1}(\mathbb{R})/PO_{n-1}(\mathbb{R})$. It is reasonable to ask about the distribution of $[L]$, as L varies, in the finite volume space $\mathfrak{S}$.

Hendrik Lenstra has informed us that David Terr has proven the equidistribution of a closely related set in the case $n = 3$ in his Ph.D. thesis.

General results in this direction seem out of reach; one can at least prove, however, mild constraints on $[L]$ that show it does not lie *too* far into the cusp. Let $a_1 \leq a_2 \leq \cdots \leq a_{n-1}$ be the successive minima (in the sense of Minkowski) of $\mathcal{O}_L^0$. Then one has automatically $a_1 a_2 \ldots a_{n-1} \asymp \sqrt{\mathcal{D}_L}$; however, on account of the assumption that K has no proper subfield, one further has for $1 \leq j \leq n-2$ that $a_1 a_j \gg a_{j+1}$ (indeed, were this not so, the lattice spanned by $a_1, a_2, \ldots, a_j$ would be stable under multiplication by a_1, and so $\mathbb{Q}(a_1)$ is a proper subfield of L). Finally one evidently has $a_j \gg 1$. Combining these constraints gives nontrivial constraints on the a_i; for example, one recovers Lemma 3.1, and one obtains $a_{n-1} \ll \mathcal{D}_L^{\frac{1}{2([\sqrt{2n}]-1)}}$, where $[\alpha]$ is the greatest integer $\leq \alpha$. One may use this type of result to further improve the exponents in Theorem 1.1 for specific n.

Remark 3.3 (Alternate ways of ordering number fields). There are many ways to order lattices of rank > 1; the ordering by volume is completely different than that by shortest vector.

We continue to work over the base field $\mathbb{Q}$. Given a number field L, we define $\mathbf{s}(L) = \inf(||x|| : x \in \mathcal{O}_L, \mathbb{Q}(x) = L)$. It is then immediate that, for any $C > 0$, the number of number fields L with $[L : \mathbb{Q}] = n$ and $\mathbf{s}(L) \leq C$ is finite; indeed one may verify that $\mathbf{s}(L)$ is "comparable" to the discriminant: $\mathcal{D}_L^{\frac{1}{n(n-1)}} \ll \mathbf{s}(L) \ll \mathcal{D}_L^{\frac{1}{2[(n-1)/2]}}$.

Let $N_{n,\mathbf{s}}(Y)$ be the number of L with $[L : \mathbb{Q}] = n$ and $\mathbf{s}(L) \leq Y$. Then one may show quite easily that $Y^{\frac{(n-1)n}{2}} \ll N_{n,\mathbf{s}}(Y) \ll Y^{\frac{(n-1)(n+2)}{2}}$; in particular, the discrepancy between upper and lower bounds is much better than when counting by discriminant. Further, the (approximate) asymptotic $N_{n,\mathbf{s}}(Y) \asymp Y^{\frac{(n-1)(n+2)}{2}}$ follows from Hypothesis 3.4 below, which seems very difficult (Granville [10] and Poonen have proved versions of this — too weak

for our purposes — using the ABC conjecture). The idea is to use Hypothesis 3.4 to construct many polynomials with square-free discriminant.

Hypothesis 3.4. Let $f \in \mathbb{Z}[x_1, \ldots, x_n]$. Then, if B_i is any sequence of boxes all of whose side lengths go to infinity, one has:

$$\lim_i \frac{\#\{x \in B_i : f(x) \text{ squarefree}\}}{\#\{x \in B_i\}} = C_f$$

where C_f is an appropriate product of local densities.

University of Wisconsin-Madison, Madison, WI
E-mail address: ellenber@math.wisc.edu

Courant Institute of Mathematical Sciences, New York University, New York, NY
E-mail address: venkatesh@cims.nyu.edu

References

[1] M. Belolipetsky, Counting maximal arithmetic subgroups, preprint; available on arXiV as math.GR/0501198.

[2] M. Bhargava, The density of discriminants of quartic rings and fields, *Ann. of Math.* **162** (2005), 1031–1063.

[3] ———, The density of of discriminants of quintic rings and fields, *Ann. of Math.*, to appear.

[4] H. Cohen, Constructing and counting number fields, *Proc. Internat. Congress of Mathematicians* (Beijing, 2002) **II** (2002), 129–138.

[5] S. D. Cohen, The distribution of Galois groups and Hilbert's irreducibility theorem, *Proc. London Math. Soc.* **43** (1981), 227–250.

[6] B. Datskovsky and D. J. Wright, Density of discriminants of cubic extensions, *J. reine angew. Math.* **386** (1988), 116–138.

[7] H. Davenport and H. Heilbronn, On the density of discriminants of cubic fields. II, *Proc. Royal Soc. London Ser. A* **322** (1971), 405–420.

[8] J. Ellenberg and A. Venkatesh, Counting extensions of function fields with specified Galois group and bounded discriminant, in *Geometric Methods in Algebra and Number Theory* (F, Bogomolov and Y. Tschinkel, eds.), *Progress in Math.* **235**, 151–168, Birkhäuser Boston, MA (2005).

[9] A. J. de Jong and N. M. Katz, personal communication.

[10] A. Granville, ABC allows us to count squarefrees, *Internat. Math. Research Notices* **19** (1998), 991–1009.

[11] A. Kable and A. Yukie, On the number of quintic fields, *Invent. Math.* **160** (2005), 217–259.

[12] P. Kleidman and M. Liebeck, *The Subgroup Structure of the Finite Classical Groups*, *London Math. Society Lecture Note Series*, Vol. 129, Cambridge Univ. Press, Cambridge, 1990.

[13] J. Klüners and G. Malle, Counting nilpotent Galois extensions, *J. reine angew. Math.* **572** (2004), 1–26.

[14] G. Malle, On the distribution of Galois groups, *J. Number Theory* **92** (2002), 315–329.

[15] J. Pila, Density of integer points on plane algebraic curves, *Internat. Math. Research Notices* **18** (1996), 903–912.

[16] D. J. S. Robinson, *A Course in the Theory of Groups*, second edition, *Graduate Texts in Mathematics* **80**, Springer-Verlag, New York, 1996.

[17] A. Schinzel, On Hilbert's irreducibility theorem, *Ann. Polon. Math.* **16** (1965), 333–340.

[18] W. M. Schmidt, Number fields of given degree and bounded discriminant, Columbia University Number Theory Seminar (New York, 1992), *Astérisque* **228** (1995), 189–195.

[19] J-P. Serre, *Lectures on the Mordell-Weil Theorem* (Translated from the French and edited by Martin Brown from notes by Micdhel Waldschmidt), *Aspects of Math.* **E15**, Friedr. Vieweg & Sohn, Braunschweig, 1989.

[20] C. L. Siegel, *Lectures on the Geometry of Numbers* (Notes by B. Friedman, Rewritten by Komaravolu Chandrasekharan with the assistance of Rudolf Suter, with a preface by Chandrasekharan), Springer-Verlag, New York, 1989.

[21] D. J. Wright, Distribution of discriminants of abelian extensions, *Proc. London Math. Soc.* **58** (1989), 17–50.

(Received March 11, 2004)