UPPER BOUNDS ON NUMBER FIELDS OF GIVEN DEGREE AND BOUNDED DISCRIMINANT

ROBERT J. LEMKE OLIVER AND FRANK THORNE

ABSTRACT. Let $N_n(X)$ denote the number of degree n number fields with discriminant bounded by X. In this note, we improve the best known upper bounds on $N_n(X)$, finding that $N_n(X) = O(X^{c(\log n)^2})$ for an explicit constant c.

1. INTRODUCTION AND STATEMENT OF RESULTS

Let $N_n(X) := \#\{K/\mathbb{Q} : [K : \mathbb{Q}] = n, |\mathrm{Disc}(K)| \leq X\}$ be the number of degree n extensions of $\mathbb{Q}$ with bounded absolute discriminant $\mathrm{Disc}(K)$. It follows from the Hermite–Minkowski theorem that $N_n(X)$ is finite, and in fact bounded by $O_n(X^n)$. This was substantially improved by Schmidt [Sch95], who shows that $N_n(X) \ll X^{(n+2)/4}$, by Ellenberg and Venkatesh [EV06], who obtain an exponent that is $O(\exp(c\sqrt{\log n}))$ for some constant c, and Couveignes [Cou19], who shows that $N_n(X) \ll X^{c(\log n)^3}$ for some unspecified constant c.

Our main theorem improves on these results.

Theorem 1.1. *There is a constant $c > 0$ such that $N_n(X) \ll_n X^{c(\log n)^2}$ for every $n \geq 6$. Explicitly, we may take $c = 1.564$, and for every $c' > 1/(4(\log 2)^2) \approx 0.52$, there is some $N > 0$ such that $N_n(X) \ll_n X^{c'(\log n)^2}$ for every $n \geq N$.*

The proof of Theorem 1.1 follows the same general strategy employed by Ellenberg and Venkatesh and by Couveignes; see Section 2 for a loose discussion of the differences.

In fact, Theorem 1.1 follows straightforwardly from a numerical computation and by combining the Schmidt bound with the following theorem that is somewhat more flexible than Theorem 1.1.

Theorem 1.2. *Let $n \geq 2$.*
   1) Let d be the least integer for which $\binom{d+2}{2} \geq 2n + 1$. Then

$$(1.1) \qquad N_n(X) \ll_n X^{2d - \frac{d(d-1)(d+4)}{6n}} \ll X^{\frac{8\sqrt{n}}{3}}.$$

   2) Let $3 \leq r \leq n$ and let d be such that $\binom{d+r-1}{r-1} > rn$. Then $N_n(X) \ll_{n,r,d} X^{dr}$.

Optimizing the choice of d, the second case of Theorem 1.2 yields an exponent that is $O(r^2 n^{1/(r-1)})$ with an absolute implied constant. We note that, in applying Theorem 1.2 to deduce Theorem 1.1, we will choose r to be a suitable multiple of $\log n$.

This improves upon the Schmidt bound $N_n(X) \ll X^{\frac{n+2}{4}}$ for $n \geq 95$. For $n \leq 5$, asymptotic formulas of the form $N_n(X) \sim c_n X$ were proved by Davenport-Heilbronn [DH71], Cohen-Diaz y Diaz-Olivier [CDyDO02], and Bhargava [Bha05, Bha10]; although our method still applies in these cases, it yields a substantially weaker result. In general, for $6 \leq n \leq 94$, the Schmidt bound remains the best known, but improvements are available for fields with restricted Galois structure due to work of Dummit [Dum18].

2. Setup for the proof

We begin by recalling the central idea of Schmidt's proof in language that will be of use to us. Let K be a number field of degree n. The ring of integers $\mathcal{O}_K$ is a lattice inside Minkowski space $K_\infty := K \otimes \mathbb{R} \simeq \mathbb{R}^n$ with covolume $c_K \sqrt{\mathrm{Disc}(K)}$, where c_K is a constant depending only on the signature of K. Similarly, the set $\mathcal{O}_K^0$ of integers in $\mathcal{O}_K$ with trace 0 forms a lattice inside the trace 0 subspace of K_∞ with covolume $c_K' \sqrt{\mathrm{Disc}(K)}$ for some c_K', and it follows that there is some $\alpha \in \mathcal{O}_K^0$ all of whose embeddings are at most $O(\mathrm{Disc}(K)^{\frac{1}{2n-2}})$. We assume for convenience of exposition that $\mathbb{Q}(\alpha) = K$; if not, Schmidt proceeds by induction, counting both the possible extensions $F = \mathbb{Q}(\alpha)/\mathbb{Q}$ and the possible K/F. The minimal polynomial $p_\alpha(x)$ of α is given by

$$p_\alpha(x) = \prod_\sigma (x - \sigma(\alpha)) = x^n + a_2(\alpha)x^{n-2} + \cdots + a_n(\alpha),$$

where the coefficients a_i are integers satisfying $|a_i(\alpha)| \ll \mathrm{Disc}(K)^{\frac{i}{2n-2}}$ and the product runs over the n embeddings $\sigma \colon K \hookrightarrow \mathbb{C}$. It follows that, ignoring the issue of subfields, the number of degree n fields with discriminant at most X may be bounded by the number of integral polynomials $f(x) = x^n + a_2 x^{n-2} + \cdots + a_n$ where each a_i is bounded by $O(X^{\frac{i}{2n-2}})$. The number of such polynomials is $O(X^{\frac{n+2}{4}})$, and this is Schmidt's bound.

The key idea of Ellenberg and Venkatesh's improvement, on which our work as well as Couveignes's is based, is that there are more invariants of small height attached to tuples of integers inside $\mathcal{O}_K^0$. For example, suppose α and β are elements of $\mathcal{O}_K^0$ whose maximum embedding is bounded by some Y. Then α and β (and hence K) are determined by their minimal polynomials, which in turn are determined by the traces $\mathrm{Tr}(\alpha^i)$ and $\mathrm{Tr}(\beta^i)$ for $1 \le i \le n$. These traces are integers of size $O(Y^i)$, and it follows that there are most $O(Y^{2\sum_{i=2}^n i}) = O(Y^{n^2+n-2})$ possible pairs (α, β). It is possible to do much better by exploiting mixed traces $\mathrm{Tr}(\alpha^i \beta^j)$, however.

By regarding the n embeddings $\alpha^{(1)}, \ldots, \alpha^{(n)}$ and $\beta^{(1)}, \ldots, \beta^{(n)}$ of α and β as variables, we might hope that once $2n$ of these mixed traces are specified, it is possible to recover the values $\alpha^{(1)}, \ldots, \alpha^{(n)}, \beta^{(1)}, \ldots, \beta^{(n)}$, and hence the pair (α, β) and the field K. There are $\binom{d+2}{2} - 1$ mixed traces with $i+j \le d$, so in particular we might hope that the traces $\mathrm{Tr}(\alpha^i \beta^j)$ with $i+j \le d \approx 2\sqrt{n}$ suffice to determine α and β. Since $\mathrm{Tr}(\alpha^i \beta^j) \ll Y^{i+j}$, this would yield that there are at most $Y^{O(n^{3/2})}$ such pairs.

In Lemma 3.1, we prove that the set of $2n$ traces $\mathrm{Tr}(\alpha^i \beta^j)$ with smallest possible $i+j$ suffices to determine the pair for a "generic" choice of α and β. We show that every field has such a choice of α and β with small height, and this leads to the first case of Theorem 1.2. This also provides the first insight into our improvement over the work of Ellenberg and Venkatesh, who only prove in this context that a set of roughly $8n$ traces suffices.

Consider now an r-tuple $\alpha_1, \ldots, \alpha_r \in \mathcal{O}_K$, with $r \ge 3$. By a similar heuristic as above, we might hope that once rn different traces $\mathrm{Tr}(\alpha_1^{i_1} \ldots \alpha_r^{i_r})$ are specified, the tuple $\alpha_1, \ldots, \alpha_r$ is determined. We show that this is the case in Lemma 3.3 for a set of rn traces with $i_1 + \cdots + i_r$ nearly as small as possible. By contrast, Ellenberg and Venkatesh only show that a set of roughly $2^{2r-1}n$ traces suffices. Couveignes's work is morally similar but structurally a little different; instead of working with mixed traces, he constructs a set of r polynomials, each with about rn coefficients, that determines each number field. Thus, his approach relies on taking roughly $r^2 n$ invariants of a number field.

By expressing the trace as a sum over embeddings, we may regard the mixed traces $\text{Tr}(\alpha_1^{i_1} \ldots \alpha_r^{i_r})$ as being governed by polynomial maps from $(\mathbb{A}^n)^r$ to $\mathbb{A}^1$. The key lemma we use to determine the α_i from these traces is then the following.

Lemma 2.1. *For $N \geq 1$, let $f_1, \ldots, f_N$ be polynomials from $\mathbb{A}^N$ to $\mathbb{A}^1$. Suppose that the determinant of the matrix $(\frac{\partial f_i}{\partial x_j})_{1 \leq i,j \leq N}$ is not identically 0. Then there is a nonzero polynomial $P \colon \mathbb{A}^N \to \mathbb{A}^1$ such that whenever $P(\mathbf{x}_0) \neq 0$ for some $\mathbf{x}_0 \in \mathbb{A}^N$, the variety $V := V_{\mathbf{x}_0}$ cut out by the equations $f_1(\mathbf{x}) = f_1(\mathbf{x}_0), \ldots, f_N(\mathbf{x}) = f_N(\mathbf{x}_0)$ consists of at most $\prod_i(\deg f_i)$ points.*

Proof. Let $F \colon \mathbb{A}^N \to \mathbb{A}^N$ be defined by $F(\mathbf{x}) = (f_1(\mathbf{x}), \ldots, f_N(\mathbf{x}))$, so that $V_{\mathbf{x}_0} = F^{-1}(F(\mathbf{x}_0))$. F is dominant, since for example the image of a Euclidean neighborhood of any $\mathbf{x}$ for which $\det\left(\frac{\partial f_i}{\partial x_j}\right)(\mathbf{x}) \neq 0$ is a neighborhood of $F(\mathbf{x})$. By [Mum99, Chapter 1, §8, Theorem 3], there is a Zariski open set $U \subseteq \text{im}(F)$ such that $F^{-1}(\mathbf{y})$ is of dimension zero for $\mathbf{y} \in U$. Choosing a polynomial Q vanishing on the complement of U, $P = Q \circ F$ is our desired polynomial.

Therefore, when $P(\mathbf{x}_0) \neq 0$, the variety $V := V_{\mathbf{x}_0}$ has dimension 0 and consists of a finite number of points. The quantitative bound follows by using Bézout's theorem [Har77, Theorem I.7.7] to iteratively bound the number of affine components of the projectivization of $V\big(f_1 - f_1(\mathbf{x}_0), \ldots, f_m - f_m(\mathbf{x}_0)\big)$ for each $m \leq N$. $\square$

3. The dimension of mixed trace varieties

In this section, we show that the varieties associated to fixed values of the "mixed traces" introduced in the previous section typically have dimension 0. We do so over $\mathbb{C}$. Thus, let $n \geq 2$ be an integer, corresponding to the degree of the extensions we wish to count, and let $r \geq 2$ be an integer corresponding to the number of elements of which we wish to take the mixed trace. To an r-tuple $\mathbf{a} = (a_1, \ldots, a_r) \in \mathbb{Z}_{\geq 0}^r$, we associate the function

$$\text{Tr}_{n,\mathbf{a}} \colon (\mathbb{A}^n)^r \to \mathbb{A}^1$$

given by

$$\text{Tr}_{n,\mathbf{a}}(\mathbf{x}_1, \ldots, \mathbf{x}_r) := \sum_{i=1}^{n} x_{1,i}^{a_1} \ldots x_{r,i}^{a_r}.$$

We let $|\mathbf{a}| = a_1 + \cdots + a_r$ denote the total degree of $\text{Tr}_{n,\mathbf{a}}$. Motivated by Lemma 2.1, let

$$D\text{Tr}_{n,\mathbf{a}} := \left(\frac{\partial}{\partial x_{k,i}} \text{Tr}_{n,\mathbf{a}}\right)_{\substack{1 \leq k \leq r \\ 1 \leq i \leq n}}$$

denote the (row) vector of partial derivatives of $\text{Tr}_{n,\mathbf{a}}$ (i.e. its gradient). Our goal, then, is to find a set A of rn different vectors $\mathbf{a}$ with small combined total degree for which the determinant of the matrix

$$(D\text{Tr}_{n,\mathbf{a}})_{\mathbf{a} \in A}$$

is not identically 0. We begin by considering the case $r = 2$, both to clarify ideas and because we obtain an essentially optimal result in this case.

Lemma 3.1. *Let $n \geq 1$, and let $A_n = (\mathbf{a}_1, \cdots, \mathbf{a}_{2n})$ consist of the first $2n$ elements of the ordered set $\{(1,0), (0,1), (2,0), (1,1), (0,2), \ldots\}$, that is, the set of ordered pairs (i,j) ordered first by total degree $i + j$, then by j.*

Then with notation as above, $\det(D\text{Tr}_{n,\mathbf{a}})_{\mathbf{a} \in A_n} \neq 0$.

Proof. Induction on n, with $\det(D\mathrm{Tr}_{1,\mathbf{a}})_{\mathbf{a}\in A_1} = 1$.

Every entry in the matrix $\mathbf{D} = (D\mathrm{Tr}_{n,\mathbf{a}})_{\mathbf{a}\in A}$ is a monomial; the (k,n)-entry $\mathbf{D}_{k,n}$ equals $a_{k,1} x_{1,n}^{a_{k,1}-1} x_{2,n}^{a_{k,2}}$, and $\mathbf{D}_{k,2n} = a_{k,2} x_{1,n}^{a_{k,1}} x_{2,n}^{a_{k,2}-1}$. We write $\det(\mathbf{D})$ an alternating sum of products of these monomials. For each k and ℓ, the contribution of those terms involving either $\mathbf{D}_{k,n}$ and $\mathbf{D}_{\ell,2n}$, or alternatively $\mathbf{D}_{k,2n}$ and $\mathbf{D}_{\ell,n}$, is

$$(3.1) \qquad \pm \det \begin{bmatrix} a_{k,1} x_{1,n}^{a_{k,1}-1} x_{2,n}^{a_{k,2}} & a_{k,2} x_{1,n}^{a_{k,1}} x_{2,n}^{a_{k,2}-1} \\ a_{\ell,1} x_{1,n}^{a_{\ell,1}-1} x_{2,n}^{a_{\ell,2}} & a_{\ell,2} x_{1,n}^{a_{\ell,1}} x_{2,n}^{a_{\ell,2}-1} \end{bmatrix} \cdot \delta_{k,\ell}$$

$$(3.2) \qquad = \pm \det \begin{bmatrix} a_{k,1} & a_{k,2} \\ a_{\ell,1} & a_{\ell,2} \end{bmatrix} \cdot x_{1,n}^{a_{k,1}+a_{\ell,1}-1} x_{2,n}^{a_{k,2}+a_{\ell,2}-1} \cdot \delta_{k,\ell},$$

where $\delta_{k,\ell}$ is the relevant $(2n-2)\times(2n-2)$ matrix minor, which doesn't involve $x_{1,n}$ or $x_{2,n}$.

By construction, if $\mathbf{a}_k + \mathbf{a}_\ell = \mathbf{a}_{2n-1} + \mathbf{a}_{2n}$ in $\mathbb{Z}^2$, then $\{k,\ell\} = \{2n-1, 2n\}$. Since the exponents of $x_{1,n}$ and $x_{2,n}$ in (3.2) are given by $\mathbf{a}_k + \mathbf{a}_\ell - (1,1)$, this implies that the contribution (3.2) from $(l,\ell) = (2n-1, 2n)$ is not cancelled by any other contribution. It therefore suffices to prove that this contribution is not zero.

But this is immediate: the 2×2 determinant in (3.2) is nonzero because consecutive elements of A_n are never scalar multiples of one another, and $\delta_{2n-1,2n} = \det(D\mathrm{Tr}_{n-1,\mathbf{a}})_{\mathbf{a}\in A_{n-1}}$. $\qquad\square$

Example. Let $n = 3$. Then $A = \{(1,0),(0,1),(2,0),(1,1),(0,2),(3,0)\}$. The associated matrix is

$$\mathbf{D} = \begin{bmatrix} \boxed{1} & \boxed{1} & 1 & \boxed{0} & \boxed{0} & 0 \\ \boxed{0} & \boxed{0} & 0 & \boxed{1} & \boxed{1} & 1 \\ \boxed{2x_{1,1}} & \boxed{2x_{1,2}} & 2x_{1,3} & \boxed{0} & \boxed{0} & 0 \\ \boxed{x_{2,1}} & \boxed{x_{2,2}} & x_{2,3} & \boxed{x_{1,1}} & \boxed{x_{1,2}} & x_{1,3} \\ 0 & 0 & \boxed{0} & 2x_{2,1} & 2x_{2,2} & \boxed{2x_{2,3}} \\ 3x_{1,1}^2 & 3x_{1,2}^2 & \boxed{3x_{1,3}^2} & 0 & 0 & \boxed{0} \end{bmatrix}.$$

The boxed entries comprise the 4×4 and 2×2 matrices in (3.2) for $(k,\ell) = (2n-1, 2n)$, whose determinants are nonzero and multiply to a summand of $\det(\mathbf{D})$.

For larger r, we apply a theorem due to Alexander and Hirschowitz [AH95], that we state in the following manner to be consistent with our notation. (See also [BO08].) This theorem is also an important ingredient in Couveignes's work.

Theorem 3.2 (Alexander–Hirschowitz). *Let V denote the complex vector space of homogeneous degree d polynomials in r variables. Given n general points in $\mathbb{P}^{r-1}$, let $W \subseteq V$ denote the subspace of polynomials whose first order partial derivatives all vanish at each of the n points. Then W has the "expected" codimension in V, namely*

$$\mathrm{codim}\, W = \min\{\dim V, rn\},$$

except for the following cases:

- *$d = 2$, $2 \le n \le r - 1$;*
- *$d = 3$, $r = 5$, $n = 7$;*
- *$d = 4$, $(r,n) \in \{(3,5),(4,9),(5,14)\}$.*

With Theorem 3.2, we are able to find a good choice of the set A in general.

Lemma 3.3. *Let $n \geq 6$, $3 \leq r \leq n$, and suppose d is such that $\binom{d+r-1}{r-1} > rn$. Then there is a set A of rn vectors $\mathbf{a} \in \mathbb{Z}_{\geq 0}^r$ of total degree d for which the determinant $\det(D\mathrm{Tr}_{n,\mathbf{a}})_{\mathbf{a} \in A} \neq 0$.*

Proof. As in Theorem 3.2, let V denote the vector space of complex homogeneous polynomials in r variables with degree d. Choose an arbitrary set of n general points in $\mathbb{P}^{r-1}$, and let W denote the subspace with vanishing first order partials at each. Under the hypotheses of Lemma 3.3, W has codimension rn, since $\dim V = \binom{d+r-1}{r-1}$ and none of the exceptional cases of Theorem 3.2 apply.

Let $\mathfrak{M}_d$ denote the set of monomials of degree d, which naturally forms a basis for V. For each $m \in \mathfrak{M}_d$, form an rn-dimensional column vector v_m by evaluating each of the first order partials of m at the n general points. Let M denote the $rn \times \binom{d+r-1}{r-1}$ matrix whose columns are the vectors v_m for $m \in \mathfrak{M}_d$. The subspace W may be identified with the kernel of M, and since W has the expected codimension, it follows that M has full rank, namely $\mathrm{rank}(M) = rn$.

Choose an $rn \times rn$ minor M' of M of full rank. The columns of M' are indexed by monomials of total degree d that may be identified with elements of $\mathbb{Z}_{\geq 0}^r$. Let A consist of those associated elements in $\mathbb{Z}_{\geq 0}^r$. Then M' is the transpose of the matrix $(D\mathrm{Tr}_{n,\mathbf{a}})_{\mathbf{a} \in A}$, evaluated at our set of n points, and since $\det(M') \neq 0$ we have $\det(D\mathrm{Tr}_{n,\mathbf{a}})_{\mathbf{a} \in A} \neq 0$ as well. $\qquad\square$

Remark. In Lemma 3.3 (and hence also Theorem 1.2) we may also allow $\binom{d+r-1}{r-1} = rn$, provided that (d, r, n) is not $(3, 5, 7)$ or $(4, 5, 14)$.

4. BOUNDS ON THE NUMBER OF NUMBER FIELDS

Let $K/\mathbb{Q}$ be a number field of degree n. Then K may be embedded into $\mathbb{C}^n$. Lemmas 3.1 and 3.3 produce, for any $r \geq 2$, a set $A \in \mathbb{Z}_{\geq 0}^r$ for which $\det(D\mathrm{Tr}_{n,\mathbf{a}})_{\mathbf{a} \in A} \neq 0$. By Lemma 2.1, there is a hypersurface outside of which the variety cut out by specifying the mixed traces $\mathrm{Tr}_{n,\mathbf{a}}$ for $\mathbf{a} \in A$ consists of a bounded number of points. The following lemma, due to Ellenberg and Venkatesh [EV06, Lemma 2.4], will be used to show that there is an r-tuple of integers in K, at least one of which cuts out K, of small height that avoids this hypersurface.

Lemma 4.1 (Ellenberg–Venkatesh). *Let $P \colon \mathbb{A}^N \to \mathbb{A}^1$ be a polynomial of degree d. Then there are integers $a_1, \ldots, a_N$ with $|a_i| \leq (d+1)/2$ for which $P(a_1, \ldots, a_N) \neq 0$.*

Lastly, we shall make use of the following upper bound on the height of the largest Minkowski minimum of a number field that follows from work of Bhargava, Shankar, Taniguchi, Thorne, Tsimerman, Zhao [BST$^+$].

Lemma 4.2. *Given a number field K of degree n, there is an integral basis $\{\beta_1, \ldots, \beta_n\}$ of its ring of integers for which $|\beta_i| \ll_n \mathrm{Disc}(K)^{1/n}$ in each archimedean embedding.*

Proof of Theorem 1.2. Consider first the case $r \geq 3$ and let d be as in the statement of the theorem. Then by Lemma 3.3, there is a set A of $\mathbf{a} \in \mathbb{Z}_{\geq 0}^r$ of size rn and degree d for which $\det(D\mathrm{Tr}_{n,\mathbf{a}})_{\mathbf{a} \in A} \neq 0$. By Lemma 2.1, there is a polynomial $P \colon (\mathbb{A}^n)^r \to \mathbb{A}^1$ such that whenever $P(\mathbf{x}_0) \neq 0$, the variety

$$\{\mathbf{x} \in (\mathbb{A}^n)^r : \mathrm{Tr}_{n,\mathbf{a}}(\mathbf{x}) = \mathrm{Tr}_{n,\mathbf{a}}(\mathbf{x}_0) \text{ for all } \mathbf{a} \in A\}$$

consists of $O_{d,r,n}(1)$ points. Given $\mathbf{x} \in (\mathbb{A}^n)^r$, we also define $\mathrm{Disc}^{(1)}(\mathbf{x})$ to denote its discriminant in the first copy of $\mathbb{A}^n$, i.e.

$$\mathrm{Disc}^{(1)}(\mathbf{x}) := \prod_{\substack{i,j \leq n \\ i \neq j}} (x_{1,i} - x_{1,j}).$$

Now, let K be a number field of degree n and discriminant at most X. By fixing an embedding $K \hookrightarrow \mathbb{C}^n$, we may regard an r-tuple of integers $\alpha_1, \ldots, \alpha_r \in \mathcal{O}_K$ as giving rise to a point $\mathbf{x}_\alpha \in (\mathbb{A}^n)^r$. Combining Lemmas 4.1 and 4.2, we find there are $\alpha_1, \ldots, \alpha_r \in \mathcal{O}_K$ with each $|\alpha_i| \ll_{n,d,r} X^{1/n}$ in each archimedean embedding for which the associated point $\mathbf{x}_\alpha$ satisfies both $P(\mathbf{x}_\alpha) \neq 0$ and $\mathrm{Disc}^{(1)}(\mathbf{x}_\alpha) \neq 0$. Since $\mathrm{Disc}^{(1)}(\mathbf{x}_\alpha) \neq 0$, α_1 cuts out a degree n extension of $\mathbb{Q}$, which must therefore be equal to K. It follows that any such K is determined, up to $O_{n,r,d}(1)$ choices, by the rn values $\mathrm{Tr}_{n,\mathbf{a}}(\mathbf{x}_\alpha)$ for $\mathbf{a} \in A$. Each of these quantities is an integer of size $O(X^{d/n})$, whence there are $O(X^{rd})$ choices in total, and hence $O(X^{rd})$ number fields K.

The case $r = 2$ is similar, except appealing to Lemma 3.1 for the construction of the set A, and noting that $\mathrm{Tr}_{n,\mathbf{a}}(\mathbf{x}_\alpha)$ is an integer of size $O(X^{|\mathbf{a}|/n})$. For the set A produced by Lemma 3.1, we find

$$\sum_{\mathbf{a} \in A} |\mathbf{a}| = 2nd - \frac{d(d-1)(d+4)}{6},$$

where d is the least integer for which $\binom{d+2}{2} \geq 2n + 1$. This yields the remaining case of the theorem; for the second inequality in (1.1), take $d = \lfloor 2\sqrt{n} \rfloor$. $\square$

Proof of Theorem 1.1. We describe an asymptotically optimal choice of r and d as $n \to \infty$ in the second part of Theorem 1.2. This will show that the exponent in Theorem 1.1 may be taken to be $(1/(4(\log 2)^2) + o(1))(\log n)^2$ as $n \to \infty$.

Thus, let n be large. We choose $d = \alpha \log n$ and $r - 1 = \beta \log n$ for constants α and β. So doing, a computation with Stirling's formula reveals

$$(4.1) \qquad \log \binom{d+r-1}{r-1} = ((\alpha + \beta) \log(\alpha + \beta) - \alpha \log \alpha - \beta \log \beta) \log n + O(\log \log n).$$

On the other hand, $\log(rn) = \log n + O(\log \log n)$, so we find that asymptotically optimal choices of α and β will satisfy

$$(\alpha + \beta) \log(\alpha + \beta) - \alpha \log \alpha - \beta \log \beta = 1 + O\left(\frac{\log \log n}{\log n}\right).$$

This expression is symmetric in α and β, as is the exponent $dr = \alpha\beta(\log n)^2 + O(\log n)$ produced by Theorem 1.2. As a Lagrange multipliers computation shows, the exponent is minimized by choosing $\alpha = \beta = 1/2 \log 2$. This yields the second part of the theorem.

The first part of the theorem (with $c = 1.564$) follows for $n < e^{12}$ by a numerical computation, with $n = 805$ being the bottleneck. For $n > e^{12}$ we choose $d = r - 1 = \lceil \log(n) \rceil$, apply the bound $\binom{2d}{d} \geq \frac{4^d}{2\sqrt{d}}$, and verify that $\binom{2d}{d} \geq rn$ and $dr < 1.564(\log n)^2$ with this choice. $\square$

5. Scope for improvement

We list here three possible directions in which our results may be improved.

First, for $r \geq 3$, it would be desirable to incorporate mixed traces of total degree *at most d*, as opposed to restricting attention to mixed traces of total degree exactly d as is done in Lemma 3.3. An optimistic version of this improvement would imply the bound $N_n(X) \ll X^{dr}$, say, whenever $\binom{d+r}{r} \geq rn$. This would substantially improve the efficiency of the method for fixed r, as this exponent would be $O(r^2 n^{1/r})$, as opposed to that obtained from Theorem 1.2, which is $O(r^2 n^{1/(r-1)})$. It would not, however, yield an improved version of Theorem 1.1.

Second, it would be worthwhile to incorporate greater input from the geometry-of-numbers. For $r < n$, there are linearly independent r-tuples of integers $\alpha_1, \ldots, \alpha_r \in \mathcal{O}_K$ for which $|\alpha_i| \ll \mathrm{Disc}(K)^{\frac{1}{2(n-r+1)}}$, which for $r \asymp \log(n)$ is significantly smaller than the bound $\mathrm{Disc}(K)^{1/n}$ coming from Lemma 4.2. However, in applying Lemma 4.1, it is a priori necessary to work with the full ring of integers, and not a small rank sublattice, in which case Lemma 4.2 is essentially optimal. It would be interesting to exclude the possibility that the hypersurface outside of which a mixed trace variety is a complete intersection contains these small rank sublattices for varying K. This would lead to an improvement in the exponents in Theorems 1.1 and 1.2 essentially by a factor of 2. The authors hope to return to this question in future work.

Lastly, unlike in the case of Schmidt's work, which considers minimal polynomials, it is not typically the case that algebraic points on a mixed trace variety cut out degree n extensions, even when the variety is a complete intersection. That is, if integers $m_{\mathbf{a}}$ are chosen for each $\mathbf{a} \in A$, as is done in the proof of Theorem 1.2, the points $\{\mathbf{x} \in (\mathbb{A}^n)^r : \mathrm{Tr}_{n,\mathbf{a}}(\mathbf{x}) = m_{\mathbf{a}} \text{ for all } \mathbf{a} \in A\}$ need not define a degree n field extension of the rationals. For example, let $n = 5$ and $r = 2$, and consider the set A produced by Lemma 3.1. Choosing integers $m_{\mathbf{a}} \in [-10^{|\mathbf{a}|}, 10^{|\mathbf{a}|}]$ randomly in `Magma` and performing a Groebner basis computation, the authors find that a "typical" choice of $\{m_{\mathbf{a}} : \mathbf{a} \in A\}$ gives rise to solutions that define a degree 30 extension of the rationals that may be realized as a degree 5 extension of a sextic field $F/\mathbb{Q}$. Thus, integers corresponding to the mixed traces of elements in quintic fields should be such that the polynomial typically defining this sextic field admits a rational root. More generally, the authors speculate that mixed traces of r-tuples of integers in degree n extensions lie on *thin* subsets of $\mathbb{Z}^{rn}$ in the sense of Serre, and that an understanding of these subsets could yield a substantial improvement to the resulting bounds on number fields.

References

[AH95]   J. Alexander and A. Hirschowitz. Polynomial interpolation in several variables. *J. Algebraic Geom.*, 4(2):201–222, 1995.

[Bha05]   Manjul Bhargava. The density of discriminants of quartic rings and fields. *Ann. of Math. (2)*,
           162(2):1031–1063, 2005.
[Bha10]   Manjul Bhargava. The density of discriminants of quintic rings and fields. *Ann. of Math. (2)*,
           172(3):1559–1591, 2010.
[BO08]   Maria Chiara Brambilla and Giorgio Ottaviani. On the Alexander-Hirschowitz theorem. *J. Pure
           Appl. Algebra*, 212(5):1229–1251, 2008.
[BST⁺]   M. Bhargava, A. Shankar, T. Taniguchi, F. Thorne, J. Tsimerman, and Y Zhao. Bounds on
           2-torsion in class groups of number fields and integral points on elliptic curves. *J. Amer. Math.
           Soc.*, to appear.
[CDyDO02] Henri Cohen, Francisco Diaz y Diaz, and Michel Olivier. Enumerating quartic dihedral exten-
           sions of $\mathbb{Q}$. *Compositio Math.*, 133(1):65–93, 2002.
[Cou19]   Jean-Marc Couveignes. Enumerating number fields. *arXiv e-prints*, page arXiv:1907.13617, July
           2019.
[DH71]   H. Davenport and H. Heilbronn. On the density of discriminants of cubic fields. II. *Proc. Roy.
           Soc. London Ser. A*, 322(1551):405–420, 1971.
[Dum18]   Evan P. Dummit. Counting G-extensions by discriminant. *Math. Res. Lett.*, 25(4):1151–1172,
           2018.
[EV06]   Jordan S. Ellenberg and Akshay Venkatesh. The number of extensions of a number field with
           fixed degree and bounded discriminant. *Ann. of Math. (2)*, 163(2):723–741, 2006.
[Har77]   Robin Hartshorne. *Algebraic geometry*. Springer-Verlag, New York-Heidelberg, 1977. Graduate
           Texts in Mathematics, No. 52.
[Mum99]   David Mumford. The red book of varieties and schemes. 1358:x+306, 1999. Includes the Michi-
           gan lectures (1974) on curves and their Jacobians, With contributions by Enrico Arbarello.
[Sch95]   Wolfgang 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).