Heuristics for the asymptotics of the number of $S_n$-number fields

Arul Shankar | Jacob Tsimerman

Department of Mathematics, Bahen Centre, University of Toronto, Toronto, Ontario, Canada

Correspondence
Arul Shankar and Jacob Tsimerman,
Department of Mathematics, University of Toronto, 40 St. George St, Toronto, ON MSS 2E4, Canada.
Email: ashankar@math.toronto.edu, jacobt@math.toronto.edu

Funding information
NSERC, Grant/Award Numbers: Sloan fellowship, Ontario Early Researcher Award

Abstract
We give a heuristic argument supporting conjectures of Bhargava on the asymptotics of the number of $S_n$-number fields having bounded discriminant. We then make our arguments rigorous in the case $n = 3$ giving a new elementary proof of the Davenport–Heilbronn theorem. Our basic method is to count elements of small height in $S_n$-fields while carefully keeping track of the index of the monogenic ring that they generate.

MSC 2020
11R04, 11R09, 11R16

1 | INTRODUCTION

A classical question in analytic number theory is to count the number of algebraic number fields of bounded discriminant. To make the question more precise, for a transitive subgroup $G < S_n$, we define $N(G,X)$ to be the number of degree-$n$ number fields, with discriminant bounded by $X$ whose Galois closure has Galois group $G$. There has been much work on the function $N(G,X)$, both conjectural and unconditional. It is conjectured by Malle [21] that $N(G,X) \asymp X^a \ln(X)^b$ with precise values for $a, b$. This conjecture was shown to be incorrect for certain cases by Klüners [15], and later corrected by Turkelli [23].

Unconditionally, a stronger version of Malle’s conjecture (recovering the asymptotics of $N(G,X)$, and not merely its growth) is known for abelian groups $G$ by work of Wright [25]. For nilpotent groups $G$, Klüners–Malle [17] prove a weak form of Malle’s conjecture, in which upper and lower bounds, differing only by a factor of $O_e(X^e)$, are proved for $N(G,X)$. Versions of Malle’s conjecture are also known for certain products of groups, and wreath product of groups (see, for example, [9, 16, 19, 24]).

For the important case $G = S_n$, only four cases are known. The case $n = 2$ is trivial. The case $n = 3$ is due to seminal work of Davenport–Heilbronn [11], while the cases $n = 4, 5$ are both landmark
results of Bhargava [2, 4]. In all these cases, the we have \( N(S_n, X) \sim c_n X \) for explicit constants \( c_n \). For general \( n \), the best known upper and lower bounds are due to work of Lemke Oliver and Thorne [18], and Bhargava, Wang, and the first named author [8], respectively. Both these results draw from the methods of previous foundational work by Ellenberg–Venkatesh [13].

In a different direction, Bhargava [5] gives a constant \( c_n \) for all integers \( n \geq 2 \), and conjectures that \( N(S_n, X) \sim c_n X \). The constant \( c_n \) is inspired by (and consistent with) the known results for \( n = 2, 3, 4, \) and \( 5 \). The value of this constant is given by

\[
c_n := \frac{1}{2} \left( \sum_{[K : \mathbb{R}] = n \text{ etale}} \frac{1}{\#\text{Aut}(K)} \right) \prod_p \left( 1 - \frac{1}{p} \right)^{-1} \sum_{[K : \mathbb{Q}_p] = n \text{ etale}} \frac{1}{\text{Disc}_p(K) \cdot \#\text{Aut}(K)} \right]. \tag{1}
\]

The justification for this conjecture follows from an assumption that degree-\( n \) extensions of local fields can be independently patched together to form \( S_n \)-number fields. The leading factor of \( \frac{1}{2} \) in the above equation corresponds to the global condition that discriminants of number fields must be congruent to 0 or 1 modulo 4. Moreover, \( c_n \) is computed explicitly in [5] using Serre’s mass formula [22].

In this paper, we give a different heuristic justification for the constant \( c_n \). In fact, we give a procedure to compute \( N(S_n, X) \). At a certain point our method requires executing a sieve which we are not able to do in general. In particular, we need to show the independence of a certain set of thin congruence conditions on lattice points within a region of Euclidean space. Our inability to show this is the reason why we do not provably compute \( N(S_n, X) \). Nonetheless, in Section 3 we execute our method rigorously in the case of \( n = 3 \), recovering the result of Davenport–Heilbronn.

The way that the results for \( n = 3, 4, 5 \) have been proven is by finding a parameterization of the space of rings of rank \( n \) over \( \mathbb{Z} \), along with some additional algebraic structure, in terms of orbits of the \( \mathbb{Z} \)-points of a reductive group acting on a lattice. These parameterizations are due to [12, 20], and [14] in the case \( n = 3 \), due to [1] when \( n = 4 \), and due to [3] when \( n = 5 \). Asymptotics for the number of these orbits having bounded discriminant are then computed using geometry-of-numbers methods. Finally, a sieve is performed to compute the asymptotics of maximal orders. The main difficulty in generalizing this approach to counting fields of degree \( n > 5 \) is the lack of a convenient parameterization for rank \( n \) rings.

Our method is to instead count algebraic integers \( \alpha \) of height \( \leq Y \) inside every degree-\( n \) field \( K \) with bounded discriminant for varying \( Y \). With one hiccup, this is fairly straightforward to do since it essentially amounts to counting the total number of algebraic numbers \( \alpha \) of degree \( n \) of height \( \leq Y \) which can be done by looking at the minimal polynomial of \( \alpha \). On the other hand, if the lattice given by \( \mathcal{O}_K \) is sufficiently regular then the number of such \( \alpha \) is given by counting points in a lattice and can therefore be well-approximated by the ratio \( \frac{C(Y)}{\text{Disc}(K)} \), for some explicit function \( C(Y) \). This then gives a family of identities parameterized by \( Y \) from which it is straightforward to recover the asymptotic behavior of \( N(S_n, X) \).

The hiccup alluded to above is that to recover the discriminant of \( K \) from the minimal polynomial of \( \alpha \) we need to know the index of \( \mathbb{Z}[\alpha] \) inside the maximal order \( \mathcal{O}_K \). This is given by independent congruence conditions for each prime, and ends up giving a thin family of congruence conditions. Proving this independence is the only part of our argument which remains conditional.

This paper is organized as follows. In Section 2, we present the general setup we use in order to count number fields via counting algebraic integers within them. We formulate a general heuristic assumption, stated as Equation (3), on counting integers polynomials with bounded height and
discriminant and fixed index. (Here, the index of a monic integer polynomial \( f(x) \) having nonzero discriminant is the index of the monogenic ring \( \mathbb{Z}[x]/f(x) \) in its maximal order.) We then prove that the heuristic assumption implies that \( N(S_n, X) \sim c_n X \).

In Section 3, we execute the argument unconditionally in the case of \( n = 3 \). That is, we prove that the main heuristic assumption holds in this case. Combined with our results in Section 2, we recover the Davenport–Heilbronn theorem as an immediate consequence.

## 2 HEURISTICS

We fix a positive integer \( d \) and consider the family \( \mathcal{F}_d \) of degree-\( d S_d \)-number fields \( K \). The purpose of this section is to prove the following result.

**Theorem 2.1.** Assume the Main Heuristic Assumption (3). Then we have

\[
\# \left\{ K \in \mathcal{F}_d : |\Delta(K)| < X \right\} \sim \left( \frac{1}{2} \sum_{[K : \mathbb{R}] = d} \frac{1}{\# \text{Aut}(K)} \right) \prod_p \left( 1 - \frac{1}{p} \right) \left( \sum_{[K : \mathbb{Q}_p] = d} \frac{|\text{Disc}(K)|^p}{\# \text{Aut}(K)} \right) X.
\]

This section is organized as follows. In Subsection 2.1, we define heights and establish a height preserving bijection between the set of degree-\( d \) fields number \( K \) that the normal closure of \( K \) has Galois group \( S_d \) over \( \mathbb{Q} \) (degree-\( d S_d \)-number fields) along with an element of \( \mathcal{O}_K \) with a certain subset of monic integer degree-\( d \) polynomials. Then in Subsection 2.2, we make our fundamental heuristic assumption regarding the asymptotics of the number of such monic integral polynomials \( f(x) \), such that the coefficients of \( f \) are bounded and satisfy certain congruence conditions. This asymptotic is expressed as a product of local densities. In Subsection 2.3, we compute these local \( p \)-adic densities using a Jacobian change of variables and in Subsection 2.4 we compute the local volume at infinity. Finally, in Subsection 2.5, we combine our results to prove Theorem 2.1, recovering Bhargava’s heuristics.

### 2.1 The main bijection and setup

Fix an integer \( d \geq 2 \). We choose a height function

\[
h : \bigcup K_\infty \to \mathbb{R}_{\geq 0},
\]

where the union is over all degree-\( d \) étalé algebras \( K_\infty \) over \( \mathbb{R} \), such that \( h \) satisfies the following three conditions: (a) the set of elements with \( h = 1 \) is compact and has measure 0, (b) \( h \) scales linearly, that is, \( h(\lambda x) = \lambda h(x) \) for \( \lambda \in \mathbb{R} \) and \( x \in \bigcup K_\infty \), and (c) the function \( h \) is nonzero away from elements \( 0 \in K_\infty \).

Consider a degree-\( d S_d \)-number field \( K \in \mathcal{F}_d \). Then we use the natural embedding \( \iota : K \to K \otimes \mathbb{R} \) to define a height function on \( K \). Namely, we set \( h(\alpha) := h(\iota(\alpha)) \). Let \( S_K \) denote the set of elements \( \alpha \in \mathcal{O}_K \) that are reduced, that is, have trace in \([0, 1, \ldots, d - 1]\). For a real number \( Y > 0 \), we let \( S_K(Y) \) denote the elements \( \alpha \in S_K \) such that \( h(\alpha) < Y \).

Let \( V \) denote the space of monic polynomials \( f(x) \) of degree \( d \), and let \( V_0 \) (respectively, \( (V_d) \)) denote the subspace of \( V \) consisting of elements \( f \) with trace 0 (respectively, with trace in
\{0, 1, ..., d − 1\}. We let \( \Delta(f) \) denote the discriminant of \( f \). Given a polynomial \( f(x) \in V(\mathbb{R}) \) with nonzero discriminant, we obtain a pair \( (\mathbb{R}[x]/f(x), x) \) of a degree-\( d \) étale algebra \( K_{\infty} \) over \( \mathbb{R} \), along with an element in \( K_{\infty} \). We define the height of a polynomial \( f \in V(\mathbb{R}) \) with nonzero discriminant via \( h(f) := h(\alpha) \), where \( f \) corresponds to the pair \( (K_{\infty}, \alpha) \).

Let \( V(\mathbb{Z})^{\text{gen}} \) denote the subset of \( V(\mathbb{Z}) \) consisting of polynomials \( f \) such that \( R_{f} := \mathbb{Z}[x]/f(x) \) is an order in an \( S_{d} \)-number field, and for subsets \( L \) of \( V(\mathbb{Z}) \) define \( L^{\text{gen}} := L \cap V(\mathbb{Z})^{\text{gen}} \). For \( f \in V(\mathbb{Z})^{\text{gen}} \), we denote by \( \text{ind}(f) \) the index of the order \( R_{f} \) in the maximal order of its fraction field. By sending an element \( \alpha \in K \) to its minimal polynomial, we obtain a bijection between the set of pairs \( (K, \alpha \in S_{K}) \) and the set \( V_{d}(\mathbb{Z})^{\text{gen}} \). Throughout this section, we fix a constant \( \delta > 1 \). Keeping track of discriminants and the index, we obtain the following equality:

\[
\sum_{K \in F_{d}} |S_{K}(Y)| = \sum_{n=1}^{\infty} \# \{ f \in V_{d}(\mathbb{Z})^{\text{gen}} : \text{ind}(f) = n, \ h(f) < Y, \ n^{2}X < |\Delta(f)| < \delta n^{2}X \}. \tag{2}
\]

Now, by Minkowski’s theorem, for any such field \( K \), \( S_{K}(Y) \) is nonempty as long as \( Y \gg X^{1/(2d−2)} \). From now on we thus restrict to \( Y = X^{1/(2d−2)+\kappa} \), for some sufficiently small positive \( \kappa \). Note that for an element \( \alpha \in K \) with \( h(\alpha) = H \), the discriminant of \( \mathbb{Z}[\alpha] \) is of size at most \( O(H^{d(d−1)}) \). Therefore, the index \( n \) is at most \( O \left( H^{d(d−1)/2}/|\Delta_{K}|^{1/2} \right) \), and so the sum in (2) goes up to \( n = O(X^{d/4−1/2+\kappa d(d−1)/2}) \).

### 2.2 Local densities and big heuristic assumption

Let \( \sigma(n) \) denote the density in \( V_{d}(\hat{\mathbb{Z}}) \) of those \( f \in V_{d}(\hat{\mathbb{Z}}) \) such that the index of \( \hat{\mathbb{Z}}[x]/f \) in the corresponding maximal order is exactly \( n \). Fix a constant \( \delta > 1 \). We make the following assumption.

**Main Heuristic Assumption:** On average over \( n \), we have

\[
\# \left\{ f \in V_{d}(\mathbb{Z})^{\text{gen}} : \text{ind}(f) = n, \ h(f) < Y, \ n^{2}X < |\Delta(f)| < \delta n^{2}X \right\} \sim \sigma(n) \text{Vol} \left( \{ f \in V_{d}(\mathbb{R}) : h(f) < Y, \ n^{2}X < |\Delta(f)| < \delta n^{2}X \} \right), \tag{3}
\]

for \( Y = X^{1/(2d−2)+\kappa} \) for \( 0 \leq \kappa < \kappa' \) for some sufficiently small \( \kappa' > 0 \). More precisely, the sums of the two terms in the above equation over positive \( n \ll X^{d/4−1/2+\kappa d(d−1)/2} \) should have the same asymptotics as \( X \to \infty \).

Next, we apply the transformation \( \vartheta \) on \( V(\mathbb{R}) \), which acts on \( f(x) \in V(\mathbb{R}) \) by dividing all the roots of \( f(x) \) by \( Y \). Equivalently, for every \( k \), the map \( \vartheta \) scales the \( x^{k} \)-coefficient of \( f(x) \) by \( 1/Y^{k} \). It is easy to see that we have \( h(\vartheta \cdot f) = h(f)/Y \) and \( \Delta(\vartheta \cdot f) = Y^{-d(d−1)}\Delta(f) \). We consider \( V_{d}(\mathbb{R}) \) as a subset of \( V(\mathbb{R}) \). Applying \( \vartheta \) will map \( V_{d} \) into a union of hyperplanes in \( V(\mathbb{R}) \), namely those having traces in \{0, 1/Y, ..., d/Y\}. We may thus write

\[
\text{Vol} \left( \{ f \in V_{d}(\mathbb{R}) : h(f) < Y, \ n^{2}X < |\Delta(f)| < \delta n^{2}X \} \right) \sim dY^{\binom{d+1}{2}−1} \cdot \text{Vol} \left( \{ f \in V_{d}(\mathbb{R}) : h(f) < 1, \ n^{2}XY^{−d(d−1)} < |\Delta(f)| < \delta n^{2}XY^{−d(d−1)} \} \right).
\]
We make the following definitions. For complex numbers $s$ where the sum converges, and for real numbers $t > 0$, we define

$$L(s) = \sum_{n \geq 1} \frac{\sigma(n)}{n^s};$$

$$g(t) = d \cdot \text{Vol}\left(\{ f \in V_0(\mathbb{R}) : h(f) < 1, \ t^2 < |\Delta(f)| < \delta t^2 \}\right).$$

In the next subsection, we will see that $L(s)$ converges absolutely to the right of $\Re(s) = -1$, and has an analytic continuation to the right of $\Re(s) = -2$ with a simple pole at $s = -1$. It is easy to see that $g(t) = O(t^2)$ as $t \to 0$, and has bounded support. As a consequence, the Mellin transform $\tilde{g}$ of $g$ is analytic with no poles to the right of $\Re(s) = -2$.

Next, we set $R = Y(\frac{d}{2})X^{-\frac{1}{2}}$. Assuming the Main Heuristic Assumption (3), we see from (2) and the above discussion that we have for every $\varepsilon > 0$

$$\sum_{K \in \mathcal{T}_d \atop X < |\Delta(K)| < \delta X} |S_K(Y)| = Y^{(d+1)/2} \sum_{n \geq 1} \sigma(n)g\left(\frac{n}{R}\right).$$

(4)

Now if $\phi(x)$ is smooth of compact support on $\mathbb{R}_+$, we would have

$$\sum_{n \geq 1} \sigma(n)\phi\left(\frac{n}{R}\right) = \int_{\Re(s)=2} L(s)\tilde{\phi}(s)R^s ds \sim R^{-1}\tilde{\phi}(-1)\text{Res}_{s=-1}L(s),$$

where the final estimate follows by pulling back the line of integration from $s = 2$ to $s = \sigma$ for some $\sigma \in (-2, -1)$. This will pick up the pole at $s = -1$, and the remaining integral can be bounded using the super-polynomial decay of $\tilde{\phi}$. We will show that even though $g(x)$ is not smooth of compact support, the same asymptotic holds.

First, we fix the issue of compact support. Set $g_\varepsilon$ to be equal to $g$ on the interval $[\varepsilon, \infty]$, and 0 on $(0, \varepsilon)$. Then

$$Y^{(d+1)/2} \sum_{n \geq 1} \sigma(n)g\left(\frac{n}{R}\right) = Y^{(d+1)/2} \sum_{n \geq 1} \sigma(n)g_\varepsilon\left(\frac{n}{R}\right) + O(\varepsilon Y^{d-1}X^{1/2}).$$

To see that the final equality holds, note simply that

$$\sum_{n \leq R} \sigma(n)g\left(\frac{n}{R}\right) \ll \frac{1}{R^2} \sum_{n \leq R} n^2\sigma(n) \ll O(\varepsilon R^{-1})$$

from $g(t) = O(t^2)$ and the analytic properties of $L(s)$ by the Wiener–Ikehara theorem.

Next, we fix the issue of smoothness. For this it suffices to note that we may bound $g_\varepsilon$ from above and below by smooth functions. Hence, combining (4) with the above displayed equation yields

$$\sum_{K \in \mathcal{T}_d \atop X < |\Delta(K)| < \delta X} |S_K(Y)| \sim Y^{d-1}X^{1/2}\tilde{g}(-1)\text{Res}_{s=-1}L(s).$$

(5)
In the next two subsections, we compute the residue of $L(s)$ at $s = -1$, and the value of $\tilde{g}(-1)$, respectively.

### 2.3 Computing the residue of $L(s)$

Let $K_p$ (respectively, $K_\infty$) be a degree $d$ étale extension of $\mathbb{Q}_p$ (respectively, $\mathbb{R}$). For $v = p$ or $\infty$, we have a map

$$
\phi : K_v \to V(K_v) \\
\alpha \mapsto \text{Nm}(x - \alpha),
$$

which is $\# \text{Aut}(K_v) \to 1$. Moreover, for $p$ a prime number, the image of $\phi(O_{K_p})$ is contained in $V(\mathbb{Z}_p)$. We fix the Haar-measure $\nu$ on $V(\mathbb{Z}_p) \cong \mathbb{Z}_p^d$ (respectively, $V(\mathbb{R}) \cong \mathbb{R}^d$) normalized so that $\nu(V(\mathbb{Z}_p)) = 1$ (respectively, $\nu(V(\mathbb{R})/V(\mathbb{Z})) = 1$). We also fix the Haar-measure $\mu$ on $K_v$ normalized so that $\mu(O_{K_p}) = 1$ when $v = p$ is prime, and normalized to be standard Euclidean measure, after identifying $\mathbb{C} \equiv \mathbb{R}^2$ via the basis $\{1, i\}$, when $v = \infty$. The following lemma relates the measures $\phi^*\nu$ and $\mu$.

**Lemma 2.2.** With the measures $\nu$ and $\mu$ normalized as above, we have

$$
\phi^*\nu = |\text{Disc}(K_p)|^{1/2}_p |\text{Disc}(x)|^{1/2}_p \mu(x) \quad \text{when } v = p; \\
\phi^*\nu = |\text{Disc}(x)|^{1/2} \mu(x) \quad \text{when } v = \infty.
$$

The above lemma is proved by Serre [22, Lemma 3].

The Dirichlet coefficients $\sigma(n)$ of $L(s)$ are clearly multiplicative. Therefore, we have $L(s) = \prod_p L_p(s)$, where the Euler factors $L_p$ are given by

$$
L_p(s) = \sum_{k=0}^{\infty} \frac{\sigma_p(p^k)}{p^{ks}}; \quad \sigma_p(p^k) = \text{Vol}(\{ f \in V(\mathbb{Z}_p) : \text{ind}(f) = p^k \}).
$$

Note that the discriminant-0 locus of $V(\mathbb{Z}_p)$ has measure 0. Therefore, we obtain

$$
L_p(s) = \int_{V(\mathbb{Z}_p)} |\text{ind}(f)|_p^s \nu(f) \\
= \sum_{[K_p : \mathbb{Q}_p] = d} \frac{|\text{Disc}(K_p)|_p^{1/2}}{|\text{Aut}(K_p)|} \int_{O_{K_p}} |\text{ind}(\alpha)|_p^s |\text{Disc}(\alpha)|_p^{1/2} \mu(\alpha) \\
= \sum_{[K_p : \mathbb{Q}_p] = d} \frac{|\text{Disc}(K_p)|_p}{|\text{Aut}(K_p)|} \int_{O_{K_p}} |\text{ind}(\alpha)|_p^{s+1} \mu(\alpha).
$$

Taking $s = -1$ now yields

$$
L_p(-1) = \int_{V(\mathbb{Z}_p)} |\text{ind}(f)|_p^{-1} \nu(f) = \sum_{[K_p : \mathbb{Q}_p] = d} \frac{|\text{Disc}(K_p)|_p}{|\text{Aut}(K_p)|} = 1 + 1/p + O(1/p^2),
$$
where the third equality is a consequence of [5, Theorem 1.1]. It thus follows that $L(s)$ has its rightmost pole at $s = -1$ and that this pole is simple. Moreover, we clearly have

$$\text{Res}_{s=-1} L(s) = \prod_p \left(1 - \frac{1}{p}\right) \left(\frac{\sum_{[K_p:Q_p]=d} |\text{Disc}(K_p)_p|}{|\text{Aut}(K_p)|}\right); \quad (8)$$

and that $L(s)$ has analytic continuation to the right of $s = -2$.

### 2.4 Computing $\tilde{g}(-1)$

We start by writing

$$\tilde{g}(-1) = d \int_0^\infty \text{Vol}\left(\{ f \in V_0(\mathbb{R}) : h(f) < 1, \ t^2 < |\Delta(f)| < \delta t^2 \}\right) \frac{dt}{t^2}$$

$$= d \int_0^\infty \text{Vol}\left(\{ f \in V_0(\mathbb{R}) : h(f) < 1, s < |\Delta(f)| < \delta s \}\right) \frac{ds}{2s^{3/2}}$$

$$= d \int_{s=0}^\infty \int_{f \in V_0(\mathbb{R})} \frac{ds}{2s^{3/2}} \frac{d_0 f}{s < |\Delta(f)| < \delta s}$$

$$= d \int_{f \in V_0(\mathbb{R})} \frac{d_0 f}{h(f) < 1} \int_{(\Delta(f)/\delta) < s < \Delta(f)} \frac{ds}{2s^{3/2}}$$

$$= d(\sqrt{\delta} - 1) \int_{f \in V_0(\mathbb{R})} \frac{1}{|\Delta(f)|} d_0(f), \quad (9)$$

where we denote the Haar measure on $V_0(\mathbb{R})$ by $d_0(f)$.

Let $K_\infty$ be a fixed degree-$d$ etale algebra over $\mathbb{R}$. In Lemma 2.2, we computed a Jacobian change of variables which applies to the map $\phi : K_\infty \to V(\mathbb{R})$. This yields the equality $df = |\text{Disc}(x)|^{1/2} dx$, where $df$ and $dx$ denote the previously normalized Haar-measures on $V(\mathbb{R})$ and $K_\infty$, respectively.

The additive group $\mathbb{R}$ acts on $V(\mathbb{R})$ via linear change of variables: an element $\lambda$ of $\mathbb{R}$ sends $f(x)$ to $f(x + \lambda)$. This action clearly preserves the discriminant. Furthermore, $\mathbb{R}$ acts on $K_\infty$ by addition.

It is easy to see that the map $\phi$ of (6) respects the action of $\mathbb{R}$ on $K_\infty$ and $V(\mathbb{R})$, which is to say that $\phi(\alpha + \lambda) = \lambda \cdot \phi(\alpha)$. This action of $\mathbb{R}$ allows us to write

$$V(\mathbb{R}) = \mathbb{R} \times V_0(\mathbb{R});$$

$$K_\infty = \mathbb{R} \times K_\infty^{(\text{tr}=0)}.$$
and $K_{\infty} \to \mathbb{R} \times K_{\infty}^{(tr=0)}$ are easily computed to be $d$ and $1$, respectively. Denoting the Haar-measure on $K_{\infty}^{(tr=0)}$ by $d_0 \alpha$, we obtain from Lemma 2.2 that

$$d_0(f) = \frac{1}{d} |\text{Disc}(\alpha)|^{1/2} d_0 \alpha.$$ 

Therefore, computing the integral in the final line of (9) using the above change of variables formula, we obtain the following Archimedean analog of (7):

$$\tilde{g}(-1) = (\sqrt{\delta} - 1) \sum_{K_{\infty}} \frac{1}{\#\text{Aut}(K_{\infty})} \int_{\alpha \in K_{\infty}^{(tr=0)}} \frac{d_0 \alpha}{h(\alpha) < 1}.$$  

(10)

Combining (5), (8), and (10), we obtain the following result.

**Theorem 2.3.** Let $X > 0$ be a real number, eventually going to infinity. Fix a constant $\delta > 1$ and sufficiently small $\kappa > 0$. Let $Y = X^{1/(2d-2)+\kappa}$. Conditional on the Main Heuristic Assumption (3), we have

$$\sum_{K \in \mathcal{P}_d} |S_K(Y)| 
\sum_{X < |\Delta(K)| < \delta X} \frac{1}{\#\text{Aut}(K_{\infty})} \int_{\alpha \in K_{\infty}^{(tr=0)}} d_0 \alpha \prod_{p} \left(1 - \frac{1}{p}\right) \left(\sum_{[K_p : \mathbb{Q}_p] = d} \frac{|\text{Disc}(K_p)|_p}{|\text{Aut}(K_p)|_p}\right).$$

Remark 2.4. The theorem above is easily modified to accommodate finitely many local conditions on the fields $K$. Indeed, if we restrict the sum on the left to only those fields $K$ whose completion $K_{\nu}$ is in a particular finite set, then on the right we restrict the sum accordingly. The analysis is identical.

2.5 | Cutting off the cusp

The purpose of this subsection is to deduce Theorem 2.1 from Theorem 2.3. For this, we will need to do two things. First, we must bound the number of fields $K$ for which $\mathcal{O}_K$ is ‘skewed’, and as a result the size of $S_K(Y)$ is anomalous. Second, for the remainder of fields $K$, we must precisely estimate the average size of $S_K(Y)$ for the relevant ranges of $Y$. To accomplish these two goals, we analyze how $S_K(Y)$ behaves using results purely from lattice theory.

We pick a small constant $\kappa > 0$ and a constant $0 < \kappa_1 < \kappa$. For $\delta > 1$, let $B_{X}^{(\delta)}$ denote the set of fields $K$ with $X < |\text{Disc}(K)| < \delta X$, such that the largest vector in a Minkowski basis for $\mathcal{O}_K$ has length bounded by $X^{1/(2d-2)+\kappa_1}$. We let $C_{X}^{(\delta)}$ denote the set of fields $K$ with $X < |\text{Disc}(K)| < \delta X$, and such that $K \not\in B_{X}^{(\delta)}$. Then, we have the following facts, which follow immediately from the theory of Minkowski bases and Minkowski’s theorem.

(a) If $|\text{Disc}(K)| < 2X$ then, for some absolute constant $c > 0$, we have $S_K(cX^{1/(2d-2)}) \geq 1$.
(b) For $K$ as above, if $Y > cX^{1/(2d-2)}$ and $Z > 1$, then we have $S_K(YZ) \ll S_K(Y)Z^{d-1}$. 
(c) For $K \in C_X^{(δ)}$, we have $S_K(cX^{1/(2d-2)+κ_1}) \ll X^{(d-2)κ_1}S_K(cX^{1/(2d-2)})$. It thus follows that

$$S_K(cX^{1/(2d-2)+κ}) \ll X^{(d-1)κ-κ_1}S_K(cX^{1/(2d-2)}).$$

(d) For $K \in B_X^{(δ)}$, we have

$$S_K(X^{1/(2d-2)+κ}) \sim \frac{X^{1/2+(d-1)κ}}{|Disc(K)|^{1/2}} \int_{α \in K_{tr=0}^∞} d_0α.$$
Estimating the size of $B^{(\delta)}_X$

We fix a signature $\sigma$ at infinity corresponding to the algebra $K_\infty$ over $\mathbb{R}$. Given a set $F$ of degree-$d$ fields $K$, we let $F^{(\sigma)}$ denote the subset of fields $K \in F$ such that $K \otimes \mathbb{R} \cong K_\infty$. For ease of notation, we define

$$M = \frac{1}{\#\text{Aut}(K_\infty)} \prod_p \left(1 - \frac{1}{p}\right) \left(\sum_{\left[K_p: \mathbb{Q}_p\right] = d} \frac{|\text{Disc}(K_p)|_p}{|\text{Aut}(K_p)|}\right).$$

From Theorem 2.3†, we have

$$\sum_{K \in F^{(\sigma)}_d, \Delta(K) X < \delta X} |S_K(X^{1/(2d-2)+\kappa})| \sim (\sqrt{\delta} - 1)M \int_{h(\alpha) < 1} \int_{\alpha \in K^{(\sigma)}_\infty} d_0 \alpha \cdot X^{1+(d-1)\kappa}.$$  

Thus, using Fact (d) in conjunction with Lemma 2.5, we obtain

$$\sum_{K \in B^{(\delta)}_X, \Delta(K) X < \delta X} \sqrt{\frac{X}{|\text{Disc}(K)|}} \sim (\sqrt{\delta} - 1)M \cdot X.$$

Conclusion

Set $\delta = 1 + \epsilon$, where $\epsilon$ will eventually tend to 0. From (11) and Lemma 2.5, we have

$$\sum_{K \in F^{(\sigma)}_d, \Delta(K) X < \delta X} \sqrt{\frac{X}{|\text{Disc}(K)|}} \sim \frac{\epsilon}{2} MX.$$

Summing over the $\epsilon$-adic ranges in $[1, X]$, we obtain

$$\sum_{K \in F^{(\sigma)}_d, 1 < \Delta(K) < X} t_K \sim \frac{1}{1 + \frac{M}{2}} X,$$

where $1/\sqrt{\delta} \leq t_K \leq 1$. Now letting $\epsilon \to 0$ yields Theorem 2.1.

2.6 Remarks

(1) The heuristic above is flexible enough to accommodate finitely many local conditions on the fields $K$. Indeed, the Archimedean places conditions are already accommodated by the height

† Technically we are applying the strengthening of this theorem where we restrict to those fields with Archimedean completion $K_\infty$, as discussed in Remark 2.4.
function $h$. If we want to impose conditions on the étale algebra $K_p$ at a finite $p$, we may simply record that condition on $f$ when making the main bijection, and it will only affect the density computation in Subsection 2.3.

(2) One may ask a more precise equidistribution question by asking about the shape of the lattice $\mathcal{O}_K^{\text{tr}=0}$, or even better by asking about the distribution of the co-volume 1 lattice $|\text{Disc}K|^{1/(2d-2)}\mathcal{O}_K^{\text{tr}=0}$ inside the space of all co-volume 1 lattices in $K_\infty^{\text{tr}=0}$. The natural guess is that it is equidistributed with respect to Haar measure on $\text{SL}_d(K_\infty^{\text{tr}=0})$, and this is proven modulo an $\text{SO}$-action by Bhargava–Harron [6] in the cases $d = 3, 4, 5$. By varying the height functions given in our heuristic one obtains a family of test functions for the resulting measure, but it appears to the authors to be insufficient to determine the measure completely.

However, the heuristic does recover one interesting case: Recall that the theta function of a lattice $(L, Q)$ is

$$\Theta_{L, Q}(z) := \sum_{v \in L} e^{2\pi i z Q(v)},$$

viewed as a function on the upper half-plane. We may recover the distribution on the theta functions associated to $|\text{Disc}K|^{1/(2d-2)}\mathcal{O}_K^{\text{tr}=0}$ by varying the function $h$ as an appropriate function only of $|\alpha|$.

(3) For the case of $d \geq 6$, our main heuristic really requires an average over $n$, since there is $O(1)$ expected points for each $n$ when $d = 6$ and fewer than 1 expected points when $d > 6$. To see this, note that

$$\sigma(n)\text{Vol}\left(\{ f \in V_d(\mathbb{R}) : h(f) < Y, n^2X < |\Delta(f)| < \delta n^2X \}\right),$$

at least for squarefree numbers $n$ and our key ranges $n \asymp X^{\frac{d}{2} - \frac{1}{2}}$ and $Y \asymp X^{\frac{1}{2d-2}}$, is roughly asymptotic to $n^{-2}Y^{(d+1)/2} \asymp X^{\frac{3}{2} - \frac{d}{4}}$.

### 3.1 The Number of Cubic Fields Having Bounded Discriminant

Consider a cubic field $K$ over $\mathbb{Q}$ with ring of integers $\mathcal{O}_K$, and discriminant $\Delta(K)$. We say that an element $\alpha \in \mathcal{O}_K$ is reduced if the trace of $\alpha$ is 0, 1, or 2. Define $|\alpha|_\infty = \max_{v|\infty} |\alpha|_v$, and for a real number $Y > 0$, let $S_K(Y)$ to be the set of reduced elements $\alpha \in \mathcal{O}_K \setminus \mathbb{Z}$ satisfying $|\alpha|_\infty < Y$.

**Definition 1.** For a ring $R$, let $V(R)$ denote the set of monic cubic polynomials $f(x) = x^3 + tx^2 + Ax + B$, where $t \in \{0, 1, 2\}$ and $A, B \in R$. **Note that this is in contrast to our notation in Section 2.**

We denote the discriminant of $f(x)$ by $\Delta(f)$. Define the height function

$$h : V(\mathbb{R}) \to \mathbb{R}$$

$$h(f) := \max |\alpha|,$$
where the maximum is taken over the roots of \( f \). We then have the following lemma whose proof is immediate.

**Lemma 3.1.** There is a bijection between the following two sets.

1. The set of pairs \((K, \alpha)\), where \( K \) is a cubic field (up to isomorphism), \( \alpha \in S_K(Y) \), and \( \Delta(K) < X \).
2. The set of irreducible polynomials \( f(x) \in V(\mathbb{Z}) \) such that \( h(f) < Y \) and \( \Delta(\mathbb{Q}[x]/f(x)) < X \).

For a subset \( L \) of \( V(\mathbb{Z}) \), we denote the set of irreducible elements in \( L \) by \( L^{irr} \). Given \( f \in V(\mathbb{Z}) \) (respectively, \( V(\mathbb{Z}_p) \)) with \( \Delta(f) \neq 0 \), we define \( \text{ind}(f) \) to be the index of \( \mathbb{Z}[x]/f(x) \) (respectively, \( \mathbb{Z}_p[x]/f(x) \)) in the ring of integral elements in \( \mathbb{Q}[x]/f(x) \) (respectively, \( \mathbb{Q}_p[x]/f(x) \)). We then have the following consequence of Lemma 3.1.

\[
\sum_{[K: \mathbb{Q}] = 3} |S_K(Y)| = \sum_{n \geq 1} \# \{ f \in V(\mathbb{Z})^{irr} : \text{ind}(f) = n, h(f) < Y, n^2X < |\Delta(f)| < \delta n^2X \} \quad (12)
\]

for a fixed constant \( \delta > 1 \).

Recall from Section 2 that we denote the density of the set of elements \( f \in V(\mathbb{Z}) \) with index \( n \) by \( \sigma(n) \). Let \( V(\mathbb{R})_{n^2X,Y} \) denote the set of elements \( f(x) \in V(\mathbb{R}) \) such that \( h(f) < Y \) and \( n^2X < |\Delta(f)| < \delta n^2X \). Then the main result of this section is as follows.

**Theorem 3.2.** For some sufficiently small \( \kappa > 0 \), set \( Y = X^{1/4+\kappa} \). Then we have

\[
\sum_{[K: \mathbb{Q}] = 3} |S_K(Y)| = \sum_{n \geq 1} \sigma(n) \text{Vol}(V(\mathbb{R})_{n^2X,Y}) + o(X). \quad (13)
\]

Note that the above result, along with (12), is essentially a restatement of the Main Heuristic Assumption of Section 2. Therefore, in conjunction with the results of Section 2, Theorem 3.2 immediately recovers the Davenport–Heilbronn result on the density of discriminants of cubic fields.

**Theorem 3.3** [11]. Let \( N_3^\pm(X) \) denote the number of cubic fields \( K \) such that \( 0 < \pm \Delta(K) < X \). Then

\[
N_3^+(X) \sim \frac{1}{12\zeta(3)}X; \quad N_3^-(X) \sim \frac{1}{4\zeta(3)}X.
\]

The values \( 1/(12\zeta(3)) \) and \( 1/(4\zeta(3)) \) arising in the above theorem can be recovered from the constant \( c_3 \) in (1) using Serre’s mass formula. See, for example, [5], where \( c_n \) is computed for all \( n \).

This section is organized as follows. First, in Subsection 3.1, we prove a variety of estimates and bounds on sets of elements in \( V(\mathbb{Z}) \) satisfying various height, discriminant, and index conditions. Then in Subsection 3.2, we provide an upper bound for the left-hand side of (13), which is optimal up to a factor of \( O(\zeta(X)) \). Finally, in Subsection 3.3, we execute an inclusion exclusion sieve in order to prove Theorem 3.2 using the counting results of the previous two subsections.
### 3.1 Counting non-maximal integer monic cubic polynomials

To estimate the number of lattice points in the bounded subsets of $V(\mathbb{R})$, we need the following proposition due to Davenport [10].

**Proposition 3.4.** Let $\mathcal{R}$ be a bounded, semi-algebraic multiset in $\mathbb{R}^n$ having maximum multiplicity $m$, and that is defined by at most $k$ polynomial inequalities each having degree at most $\ell$. Then the number of integral lattice points (counted with multiplicity) contained in the region $\mathcal{R}$ is

$$\text{Vol}(\mathcal{R}) + O(\max\{\text{Vol}(\tilde{\mathcal{R}}), 1\}),$$

where $\text{Vol}(\tilde{\mathcal{R}})$ denotes the greatest $d$-dimensional volume of any projection of $\mathcal{R}$ onto a coordinate subspace obtained by equating $n - d$ coordinates to zero, where $d$ takes all values from 1 to $n - 1$. The implied constant in the second summand depends only on $n$, $m$, $k$, and $\ell$.

**Congruence conditions on polynomials with index divisible by an integer $n$**

Let $p$ be a prime. The following criterion for $f \in V(\mathbb{Z})$ with $\Delta(f) \neq 0$ to have index divisible by a prime $p$ follows immediately from [7, Theorem 14] (originally due to work of Davenport–Heilbronn [11]).

**Lemma 3.5.** An element $f(x) \in V(\mathbb{Z})$ has index divisible by a prime $p$ if and only if there exists $\bar{r} \in \mathbb{Z}/p\mathbb{Z}$ such that for every lift $r \in \mathbb{Z}$ of $\bar{r}$, we have $p^2 \mid f(r)$ and $p \mid f'(r)$, where $f'(x)$ is the derivative of $f(x)$.

Let $f(x) = x^3 + kx^2 + Ax + B$ be an element of $V(\mathbb{Z})$. It follows from Lemma 3.5 that the residue classes of $A$ and $k$ modulo $p$ and the residue class of $B$ modulo $p^2$ determine whether or not $p \mid \text{ind}(f)$. More precisely, we have the following immediate consequence of Lemma 3.5.

**Corollary 3.6.** Let $p$ be a fixed prime and let $k$ and $A$ be fixed integers. The number of $\bar{B} \in \mathbb{Z}/p^2\mathbb{Z}$ such that $p \mid \text{ind}(x^3 + kx^2 + Ax + B)$, for lifts $B \in \mathbb{Z}$ of $\bar{B}$, is determined by the residue classes of $k$ and $A$ in $\mathbb{Z}/p\mathbb{Z}$. In fact, the number of such $B \in \mathbb{Z}/p^2\mathbb{Z}$ is equal to the number of roots modulo $p$ of $3x^2 + 2kx + A$.

For a positive integer $n$, let $\Sigma_n$ denote the set of polynomials $f(x)$ in $V(\mathbb{Z})$ with nonzero discriminant such that $n \mid \text{ind}(f)$. We will need a version of Corollary 3.6 for arbitrary integers $n$. For $f(x)$ having non-zero discriminant we define $R_f := \mathbb{Z}[x]/f(x)$ and $\mathcal{O}_f \supset R_f$ the ring of integers of $R_f \otimes \mathbb{Q}$. To analyze the case when $n$ is divisible by a prime power $p^\ell$ with $\ell \geq 2$, we write the set $\Sigma_{p^\ell}$ as the disjoint union

$$\Sigma_{p^\ell} = \Sigma_{p^\ell}^{(1)} \cup \Sigma_{p^\ell}^{(2)},$$

where $\Sigma_{p^\ell}^{(1)}$ denotes the set of elements $f(x) \in \Sigma_{p^\ell}$ such that the image of $x$ in $\mathcal{O}_f/\mathbb{Z}$ is not a multiple of $p$. Then we have the following lemma:
Lemma 3.7. If an element $f(x) \in V(\mathbb{Z})^{irr}$ belongs to $\Sigma_{p^\ell}^{(1)}$, then there exists $\tilde{r} \in \mathbb{Z} / p^\ell \mathbb{Z}$ such that for every lift $r \in \mathbb{Z}$ of $\tilde{r}$, we have $\frac{1}{4} p^{2\ell} \mid f(r)$ and $\frac{1}{2} p^\ell \mid f'(r)$.

Proof. We start with the case when the splitting type of $f$ at $p$ is $(1^21)$. Let $\sigma_1$, $\sigma_2$, and $\theta$ denote the roots of $f$ in $\overline{Q}_p$, with $\sigma_1 \equiv \sigma_2 \pmod{p}$. Then $\theta$ belongs to $\mathbb{Z}_p$ and $\sigma_1$ and $\sigma_2$ either belong to $\mathbb{Z}_p$ and are congruent elements in the ring of integers of a ramified extension of $Q_p$. For $r = (\sigma_1 + \sigma_2)/2$, the roots of $f(x + r)$ are $\theta' = \theta - r$, $\sigma_1' = \sigma_1 - r$, and $\sigma_2' = \sigma_2 - r = -\sigma_1'$. The $p$-part of the discriminant $\Delta(f) = \Delta(f(x - r))$ is then computed to be equal to the $p$-part of $(\sigma_1' - \sigma_2')^2 = 4\sigma_1'^2$. It follows that $\frac{1}{2} p^\ell$ divides $\sigma_1'$ and $\sigma_2'$, and therefore that $f(x + r)$ is of the form $x^3 + ax^2 + \frac{1}{2} p^{\ell} bx + \frac{1}{4} p^{2\ell} c$. Clearly, this remains true for all $r_1$ that are congruent to $r$ modulo $p^\ell$. Evaluating $f(x)$ and $f'(x)$ at $r$ yields the lemma for this case.

Next assume that the splitting type of $f$ is $(1^3)$. For an appropriate integer $r$, replace $f(x)$ by $f(x + r)$ such that the triple zero of $f$ modulo $p$ is at 0, that is, $p \mid a$, $p \mid b$, and $p \mid c$. Since $p \mid \text{ind}(f)$ we must have $p^2 \mid c$. Denote the image of $x$ in $R_f = \mathbb{Z}[x]/f(x)$ by $\alpha$. Since $\alpha$ is not a multiple of $p$, it cannot simultaneously happen that $p^2 \mid b$ and $p^3 \mid c$. Therefore the only possibilities are that $p \mid b$, $p^2 \mid c$ or that $p^2 \mid b$ and $p^2 \mid c$. We claim in both cases that the $p$-part of the index of $f$ is $p$ and that therefore $\ell = 1$. Indeed, in the former case, we have $p^3 \mid \Delta(f)$ and hence $p^2 \nmid \text{ind}(f)$.

In the latter case, since $p$ divides $a$, $b$, and $c$, it follows that $\alpha$, considered as an element of the ring of integers of $\overline{Q}_p$, belongs to the maximal ideal. If $\mathcal{O}_f$ is unramified at $p$, then it would follow that $\alpha$ is a multiple of $p$, a contradiction. Hence $p \mid \text{Disc}\mathcal{O}_f$. Combined with the fact that in this case we have $p^3 \mid \Delta(f)$, we conclude that $p^3 \nmid \text{ind}(f)$ as needed.

The result now follows from Lemma 3.5.

For a set $S \subset V(\mathbb{Z})$, let $\nu(S)$ denote the volume of the closure of $S$ in $V(\hat{\mathbb{Z}})$. For integers $k$ and $A$, let $S(k, A)$ denote the set of integers $B$ such that $x^3 + kx^2 + Ax + B$ belongs to $S$. Let $\nu(k, A; S)$ denote the volume of the closure of $S(k, A)$ in $\mathbb{Z}_p$. Here, we compute volumes in $V(\mathbb{Z}_p)$ and $\mathbb{Z}_p$ in terms of Euclidean measure normalized so that $V(\mathbb{Z}_p)$ and $\mathbb{Z}_p$, respectively, have volume 1. Then we have the following result.

Proposition 3.8. Let $n$ be a positive integer and write $n = q^3 m$, where $m$ is cube-free. Then

(a) the set $\Sigma_n$ is defined via congruence conditions modulo $n^2$;
(b) for $k, A \in \mathbb{Z}$, the density of $\Sigma_n(k, A)$ depends only on the congruence classes of $k$ and $A$ modulo $n$;
(c) we have the bound $\nu(\Sigma_n) \ll \frac{n^c}{q^3 m^2}$.

Proof. By the Chinese Remainder Theorem, we may assume that $n = p^\ell$ is a prime power. We proceed by induction on $\ell$. For $\ell = 1$ all three claims follow from Lemma 3.5. Furthermore, for all $\ell \geq 1$, the claims of the proposition with $\Sigma_{p^\ell}$ replaced by $\Sigma_{p^{\ell-1}}^{(1)}$, follow from Lemma 3.7. This also yields the required results for $\ell = 2$, since $\Sigma_{p^2} = \Sigma_{p^{2-1}}^{(1)}$.

We now assume that $\ell \geq 3$ and prove the claims of the proposition with $\Sigma_{p^\ell}$ replaced with $\Sigma_{p^{\ell-2}}^{(2)}$. To this end, let $f(x) = x^3 + kx^2 + Ax + B$ be an element of $\Sigma_{p^{\ell-2}}^{(2)}$ and let $\alpha$ denote the image of
In $\mathbb{Z}[x]/f(x)$, then there exists $r \in \mathbb{Z}$, defined uniquely modulo $p$, such that $\alpha - r$ is a multiple of $p$. The polynomial $f(x + r) = x^3 + k'x^2 + A'x + B'$ satisfies $p \mid k'$, $p^2 \mid A'$ and $p^3 \mid B'$. Furthermore, the unique $\mathbb{Z}$-translate of

$$x^3 + (k'/p)x^2 + (A'/p^2)x + (B'/p^3)$$

which belongs to $V(\mathbb{Z})$ also belongs to $\Sigma_{p^{r-1}}$. Parts (a) and (b) of the proposition follow immediately from induction. Part (c) also follows since for each fixed $r$, the volume of the corresponding subset of cubic polynomials with index divisible by $p^r$ is bounded by $O(p^{-6} \cdot \nu(\Sigma_{p^{r-1}}))$. Since $r$ is defined modulo $p$, Part (c), and the proposition, follows.

**Bounds on reducible elements**

For squarefree integers $n$, we obtain a bound on the number of reducible polynomials in $\Sigma_n$ having bounded height. For a set $S \subset V(\mathbb{Z})$, let $S^{\text{red}}$ denote the set of reducible polynomials in $S$. We have the following result.

**Proposition 3.9.** Let $n$ be a positive squarefree integer. Then

$$\# \{f \in \Sigma_n^{\text{red}} : h(f) < Y\} \ll Y^{3/n^{1-\varepsilon} + n^{\varepsilon}Y};$$

$$\# \{f \in V(\mathbb{Z}) : h(f) < Y, \Delta(f) = 0\} \ll Y^2.$$  \hspace{1cm} (14)

**Proof.** If $f \in V(\mathbb{Z})$ is reducible, then there exists $r \in \mathbb{Z}$ such that $f(r) = 0$. Hence, we have

$$f(x) = (x - r)(x^2 + (r + k)x + b) = x^3 + kx^2 + (b - r(r + k))x - br$$

with $k \in \{0, 1, 2\}$. For such an element $f$ with $h(f) < Y$, it follows that $|br| \ll Y^3$ and $|b - r(r + k)| \ll Y^2$. The latter inequality implies that if $|b| > CY^2$ for some $C$, then $|r| \gg Y$. Taking $C$ sufficiently large contradicts the inequality $|br| \ll Y^3$, and hence we obtain $|r| \ll Y$ and $|b| \ll Y^2$. Note that for fixed $r$ and $k$, the polynomial $\Delta_{r,k}(b) := \Delta(f)$ is a cubic polynomial in $b$ with leading coefficient 4, and in particular, that it is nonzero.

Let $f \in \Sigma_n$ be a polynomial with fixed $r$ and $k$. It follows that $n^2 \mid \Delta(f)$ and that therefore the residue of $b$ modulo $n$ has $O(n^\varepsilon)$ choices. The first claim of the proposition now follows from the bounds on $|r|$ and $|b|$.

The second claim is immediate since given values for the $x$- and $x^2$-coefficients of $f$, the constant coefficient is determined by the $\Delta(f) = 0$ condition. \hfill \Box

**Estimates and bounds on irreducible elements**

Let $V(\mathbb{R})_{X,Y}^{\pm}$ denote the set of elements $f \in V(\mathbb{R})^\pm$ such that $|\Delta(f)| < X$ and $h(f) < Y$. We start by estimating the number of elements in $\Sigma_n$ with bounded height and discriminant, for squarefree integers $n$. 
Theorem 3.10. Let \( m \) be a positive integer and let \( n \) be a positive squarefree integer relatively prime to \( m \). Let \( \mathcal{L} \subset V(\mathbb{Z}) \) be a set defined by congruence conditions modulo \( m \). Then

\[
\# \{ f \in \mathcal{L} \cap \Sigma_n : 0 < \pm \Delta(f) < X, \ h(f) < Y \} = \nu(\mathcal{L})\nu(\Sigma_n)\text{Vol}(V(\mathbb{R})_{X,Y}^\pm) + O \left( Y^3 m/n^{1-\epsilon} + Y^2 mn^\epsilon \right).
\]

Proof. Given \( A \in \mathbb{Z} \), let \( R_A \) denote the set of polynomials \( f(x) \in V(\mathbb{R})_{X,Y}^\pm \) with \( x \)-coefficient equal to \( A \). Let \( \nu_{m,n,A} \) denote the density of the set of polynomials \( f(x) \in \mathcal{L} \cap \Sigma_n \) whose \( x \)-coefficient is \( A \) within the set of polynomials \( f \in V(\mathbb{Z}) \) whose \( x \)-coefficient is \( A \). From Proposition 3.8, it follows that \( \nu_{m,n,A} \) depends only on the residue of \( A \) modulo \( mn \), and that \( \nu_{m,n,A} \ll 1/n^{2-\epsilon} \). Fibering by \( A \), we obtain

\[
\# \{ f \in \mathcal{L} \cap \Sigma_n : |\Delta(f)| < X, \ h(f) < Y \} = \sum_{|A| \leq 3Y^2} \nu_{m,n,A} |R_A| + O(Y^2 mn^\epsilon),
\]

where \( |R_A| \) denotes the length of \( R_A \). (The second line in the above equation follows by splitting the integer points in \( R_A \) into congruence classes modulo \( mn^2 \). For fixed \( A \), the number of \( B \) mod \( n^2 \) with \( x^3 + kx^2 + Ax + B \) having index divisible by \( n \) is bounded by \( O(n^6) \). Thus, the number of such congruence classes is bounded by \( O(mn^6) \), yielding the error term of \( O(Y^2 mn^\epsilon) \).

Note that we have \( |R_A| \ll Y^3 \) from the height bound. Now the average value of \( \nu_{m,n,A} \), as \( A \) varies over a complete residue system modulo \( mn \), is clearly equal to \( \nu(\mathcal{L})\nu(\Sigma_n) \). Consider the main term of the second line of (15). We break up the sum over arithmetic progressions modulo \( mn \). From Proposition 3.4, we obtain

\[
\sum_{|A| \leq 3Y^2} \nu_{m,n,A} |R_A| = \sum_{d \in \mathbb{Z} / (mn)} \nu_{m,n,d} \sum_{|A| \leq 3Y^2} |R_A|
\]

\[
= \sum_{d \in \mathbb{Z} / (mn)} \nu_{m,n,d} \left( \frac{\nu(\mathcal{L})\nu(\Sigma_n)\text{Vol}(V(\mathbb{R})_{X,Y}^\pm)}{mn} + O(Y^3) \right)
\]

\[
= \nu_{m}(L)\nu_{n}(\Sigma_n)\text{Vol}(V(\mathbb{R})_{X,Y}^\pm) + O \left( \frac{mY^3}{n^{1-\epsilon}} \right),
\]

where we use the fact that \( \nu_{m,n,A} \ll 1/n^{2-\epsilon} \). This concludes the proof of the theorem. \( \square \)

Finally, we prove a bound on the number of polynomials \( f \) such that \( \text{ind}(f) \) is divisible by arbitrary positive integers \( n \).

Theorem 3.11. Let \( n \) be a positive integer. We have

\[
\# \{ f \in \Sigma_n : |\Delta(f)| < X, \ h(f) < Y \} \ll \nu(\Sigma_n) \left( Y^2 + n \right) \min \{ Y^3, X^{\frac{1}{2}} \} + nY^2.
\]

Proof. Let \( R_A \) be defined as in the proof Theorem 3.10, and note that \( |R_A| \ll \min(Y^3, X^{\frac{1}{2}}) \). We fiber over \( k \) and \( A \). For fixed \( k \) and \( A \), the condition \( n^2 \ | \Delta(x^3 + kx^2 + Ax + B) \) has at most \( O(n) \)
solutions $B$ mod $n^2$. Therefore, we obtain

$$
\{f \in \Sigma_n : |\Delta(f)| < X, h(f) < Y\} \ll \sum_{k \in \{0,1,2\}} \sum_{A \leq 3Y^2} (\nu(k, A; \Sigma_n(k, A)) \cdot |R_A| + O(n))
$$

$$
\ll nY^2 + \sum_{k \in \{0,1,2\}} \sum_{A \leq 3Y^2} \nu(k, A; \Sigma_n(k, A)) (\min\{Y^3, X^{\frac{1}{2}}\})
$$

$$
\ll nY^2 + (Y^2 + n) \min\{Y^3, X^{\frac{1}{2}}\} \operatorname{Avg}(\nu(k, A; \Sigma_n(k, A))),
$$

where the average is over $k \in \{0, 1, 2\}$ and $A$ modulo $n$. Since this average is equal to $\nu(\Sigma_n)$, the theorem follows.

3.2 An upper bound

We fix a constant $C > 1$ such that for every cubic field $K$ with $\Delta(K) \leq X$, the set $S_K(CX^{1/4})$ is nonempty. Let $X$ and $Y$ be positive real numbers such that $Y \geq CX^{1/4}$. Our goal in this section is to prove an upper bound for number of cubic fields $K$ with discriminant bounded by $X$, where each field $K$ is counted with weight $|S_K(Y)|$. We start with the following important lemma.

**Lemma 3.12.** Let $X$ and $Y$ be as above. Let $K$ be a cubic field such that $X/2 \leq |\Delta(K)| \leq X$. Then we have

$$
\#S_K(Y)/\#S_K(CX^{1/4}) \ll Y^2/X^{1/2},
$$

where the implied constant is independent of $X$, $Y$, and $K$.

**Proof.** We start by picking a Minkowski basis $(1, \alpha, \beta)$ for $O_K$. Let $\delta_1 \leq \delta_2$ be such that $|\alpha|_\infty = X^{\delta_1}$ and $|\beta|_\infty = X^{\delta_2}$. From our assumption on $C$, it follows that $|\alpha|_\infty < CX^{1/4}$. We have

$$
\#S_K(Y) \asymp \frac{Y}{X^{\delta_1}} \cdot \max\left\{ \frac{Y}{X^{\delta_2}}, 1 \right\};
$$

$$
\#S_K(CX^{1/4}) \asymp \frac{CX^{1/4}}{X^{\delta_1}} \cdot \max\left\{ \frac{CX^{1/4}}{X^{\delta_2}}, 1 \right\}.
$$

The proof now follows from the fact that $X^{\delta_2} \gg X^{1/4}$.

We now prove the following crucial upper bound:

**Theorem 3.13.** Let $X$ and $Y$ be as above. Then

$$
\sum_{|K : \mathbb{Q}| = 3 \atop |\Delta(K)| \leq X} |S_K(Y)| \ll \varepsilon X^{1/2 + \varepsilon} Y^2,
$$

where the implied constant only depends on $C$. 

Proof. We start by counting cubic fields whose discriminants are in a dyadic range of $M < X$. If $K$ is such a field and if $\alpha \in S_K(CM^{1/4})$, then any polynomial $f$ corresponding to $(K, \alpha)$ under Lemma 3.1 must satisfy $\text{ind}(f) \ll M^{1/4}$ (since $\Delta(f) \ll M^{3/2}$). From Lemma 3.12 and (12), we obtain

$$
\sum_{[K: \mathbb{Q}] = 3} |S_K(Y)| \ll \frac{Y^2}{M^{1/2}} \sum_{[K: \mathbb{Q}] = 3} |S_K(CM^{1/4})| 
\ll \frac{Y^2}{M^{1/2}} \sum_{n \ll M^{1/4}} \# \{ f \in \Sigma_n : h(f) < CM^{1/4}, |\Delta(f)| < n^2M \}
\ll \frac{Y^2}{M^{1/2}} \sum_{n \ll M^{1/4}} \left( nM^{1/2} + \nu(n)M^{1/2} \min\{M^{3/4}, nM^{1/2}\} \right)
\ll Y^2M^{1/2} + Y^2M^{1/2} \sum_{n \ll M^{1/4}} \nu(n)n
\ll M^{1/2+\varepsilon}Y^2,
$$

where the third estimate follows from Theorem 3.11, and the last estimate follows from Proposition 3.8. Summing $M < X$ over powers of 2 yields the theorem. \hfill \Box

3.3 The sieve

To translate our results regarding polynomials with index divisible by some integer $n$ to polynomials with index exactly $n$, it is necessary to employ a sieve. To this end, fix $C$ as in the previous subsection. Throughout this section, we set $Y = CX^{1/4+\varepsilon}$. We apply the inclusion exclusion sieve to obtain

$$
\sum_{[K: \mathbb{Q}] = 3} |S_K(Y)| = \sum_{n \geq 1} \# \{ f \in V(\mathbb{Z})^{\text{irr}} : \text{ind}(f) = n, h(f) < Y, |\Delta(f)| < n^2X \}
= \sum_{n \geq 1} \sum_{d \geq 1} \mu(d) \# \{ f \in \Sigma^{\text{irr}}_{dn} : h(f) < Y, |\Delta(f)| < n^2X \}. \tag{16}
$$

The next result bounds the tail of the above sum. Let $\kappa$, $\delta_1$, and $\delta_2$ be positive real numbers to be chosen later. Then we have

Lemma 3.14. We have

$$
\sum_{n, d \geq 1 \atop nd > X^{1/4+\delta_1}} \# \{ f \in \Sigma^{\text{irr}}_{dn} : h(f) < Y \} \ll \varepsilon X^{1+5\kappa-\delta_1+\varepsilon}.
$$

Proof. Let $f \in V(\mathbb{Z})^{\text{irr}}$ be such that $\text{ind}(f) > X^{1/4+\delta_1}$ and $h(f) < Y$. Denote $\mathbb{Q}[x]/f(x)$ by $K$. Then

$$
|\Delta(K)| \ll \frac{Y^6}{\text{ind}(f)^2} \ll X^{1+6\kappa-2\delta_1}.
$$
Therefore, denoting the number of divisors of \( m \) by \( \sigma_0(m) \), we have

\[
\sum_{n,d \geq 1 \atop nd > X^{1/4+\delta_1}} \# \{ f \in \Sigma_{dn}^{\text{irr}} : h(f) < Y \} = \sum_{m > X^{1/4+\delta_1}} \sigma_0(m) \# \{ f \in \Sigma_m^{\text{irr}} : h(f) < Y \}
\ll \sum_{[K : \mathbb{Q}] = 3 \atop |\Delta(K)| < X^{1+6\epsilon-2\delta_1}} X^\epsilon S_K(Y) \ll X^{1+5\epsilon-\delta_1+\epsilon},
\]

where the final estimate follows from Theorem 3.13.

For an integer \( m \), let \( sq(m) \) denote the product of the prime powers dividing \( m \) to exponent at least 2. Next, we bound the sum over the terms in the second line of (16), where \( sq(dn) \) is large.

**Lemma 3.15.** We have

\[
\sum_{n,d \geq 1 \atop nd \leq X^{1/4+\delta_1} \atop sq(dn) > X^{\delta_2}} \sum_{f \in \Sigma_{dn}^{\text{irr}} : h(f) < Y, |\Delta(f)| < n^2X} \ll \epsilon X^{1+2\epsilon+2\delta_1-\delta_2/2+\epsilon} + X^{1+2\epsilon-\delta_2/9+\epsilon}.
\]

**Proof.** Applying Theorem 3.11, we see that the left-hand side of the above equation is

\[
\ll \epsilon X^\epsilon \sum_{m \leq X^{1/4+\delta_1} \atop sq(m) > X^{\delta_2}} \left( mY^2 + \nu(\Sigma_m)(Y^2 + m) \min\{Y^3, mX^{1/2}\} \right)
\ll X^{1+2\epsilon+2\delta_1-\delta_2+\epsilon} + X^{1+2\epsilon} \sum_{m \leq X^{1/4+\delta_1} \atop sq(m) > X^{\delta_2}} \nu(\Sigma_m)m,
\]

where the first term in the second line above corresponds to the sum of \( mY^2 \) over the appropriate \( m \), and the second term corresponds to the sum over \( m \) of \( m\nu(\Sigma_m)Y^2X^{1/2} \). (Note that since \( \delta_1 \) will be taken to be small, we have \( Y^2 \geq m \) for all \( m \) in our range.) Writing \( m = rs \), where \( r \) is squarefull, \( s \) is squarefree, and \( (r, s) = 1 \), we get

\[
\sum_{m \leq X^{1/4+\delta_1} \atop sq(m) > X^{\delta_2}} \nu(m)m \ll \epsilon X^\epsilon \sum_{r > X^{\delta_2}} \sum_{s \leq X^{1/4+\delta_1}/r} \frac{rs}{r^{5/3}s^2}
\ll \epsilon X^\epsilon \sum_{r > X^{\delta_2}} r^{-2/3},
\]

where the first equality follows from Proposition 3.8.

Now, a squarefull number \( r \) can always be written as \( r = y^2r' \) where \( r' \) is square-free and \( y \geq r^{1/3} \). Moreover, the fibers of \( r \to y \) correspond to choices for \( r' \) and \( r' \mid y \), and thus they are of size
\[ \tau(y) = y^{o(1)}. \] Therefore, we have
\[
\sum_{m \in X^{1/4+\delta_1}} \nu(m)m \ll_{\varepsilon} X^{\varepsilon} \sum_{y > X^{2/3}} y^{-4/3+o(1)} \ll_{\varepsilon} X^{-\delta_2/9+\varepsilon},
\]
concluding the proof of the lemma.

We are now ready to prove the main result of this section.

**Proof of Theorem 3.2.** Equation (16) and Lemmas 3.14 and 3.15 imply that we have
\[
\sum_{[K: \mathbb{Q}=3]} |S_K(Y)| = \sum_{n,d \geq 1 \atop nd \leq X^{1/4+\delta_1} \atop \text{sq}(dn) \leq X^{\delta_2}} \mu(d) \# \{ f \in \Sigma_{dn} : h(f) < Y, |\Delta(f)| < n^2 X \} + O_{\varepsilon}(E_1)
\]
\[
= \sum_{n,d \geq 1 \atop nd \leq X^{1/4+\delta_1} \atop \text{sq}(dn) \leq X^{\delta_2}} \mu(d) \# \{ f \in \Sigma_{dn} : h(f) < Y, |\Delta(f)| < n^2 X \} + O_{\varepsilon}(E_1 + X^{3/4+3\varepsilon+\delta_2}),
\]
where we use Proposition 3.9 to prove that the number of reducible elements \( \Sigma_{dn} \) is negligible, and where the error term \( E_1 \) is defined to be
\[ E_1 := X^{1+5\varepsilon-\delta_1+\varepsilon} + X^{1+2\kappa+2\delta_1-\delta_2/2+\varepsilon} + X^{1+2\kappa-\delta_2/9+\varepsilon}. \]
Write \( nd = m\ell \), where \( m \) is squarefree, \( \ell \) is squarefull, and \( (m, \ell) = 1 \). Estimating the number of irreducible elements in \( \Sigma_{dn} \) having bounded height and discriminant using Theorem 3.10, we obtain
\[
\{ f \in \Sigma_{dn} : h(f) < Y, |\Delta(f)| < n^2 X \} = \nu(\Sigma_{dn}) \text{Vol}(V(\mathbb{R})_{n^2 XY}) + O_{\varepsilon}(Y^{3+\varepsilon} \ell^2 / m^{1-\varepsilon} + Y^2 \ell^2 m^{\varepsilon}).
\]
Adding the error term in the right-hand side of the above equation over \( m \leq X^{1/4+\delta_1} \) and \( \ell \leq X^{\delta_2} \), we obtain the following estimate.
\[
\sum_{[K: \mathbb{Q}=3]} |S_K(Y)| = \sum_{n,d \geq 1 \atop nd \leq X^{1/4+\delta_1} \atop \text{sq}(dn) \leq X^{\delta_2}} \mu(d) \nu(\Sigma_{dn}) \text{Vol}(V(\mathbb{R})_{n^2 XY}) + O(E(\kappa, \delta_1, \delta_2))
\]
\[
= \sum_{n \geq 1} \sigma(n) \text{Vol}(V(\mathbb{R})_{n^2 XY}) + O(E(\kappa, \delta_1, \delta_2)),
\]
where
\[ E(\kappa, \delta_1, \delta_2) := X^{1+5\kappa-\delta_1+\varepsilon} + X^{1+2\kappa+2\delta_1-\delta_2/2+\varepsilon} + X^{1+2\kappa-\delta_2/9+\varepsilon} + X^{3/4+3\kappa+3\delta_2+\varepsilon} + X^{3/4+2\kappa+\delta_1+3\delta_2+\varepsilon}. \]
Since it is clearly possible to pick positive constants \( \kappa, \delta_1, \) and \( \delta_2 \) such that \( E(\kappa, \delta_1, \delta_2) = o(X) \), we recover Theorem 3.2.
Finally, the Davenport–Heilbronn theorem [11], stated as Theorem 3.3, follows from Theorem 3.2 in conjunction with Theorem 2.1.

ACKNOWLEDGEMENTS
It is our pleasure to thank Manjul Bhargava, Peter Sarnak, and Xiaoheng Wang for many helpful conversations. We are very grateful to the referee for many helpful comments and suggestions. Arul Shankar is supported by an NSERC discovery grant and a Sloan fellowship. Jacob Tsimerman is supported by an NSERC discovery grant and an Ontario Early Researcher Award.

JOURNAL INFORMATION
The Journal of the London Mathematical Society is wholly owned and managed by the London Mathematical Society, a not-for-profit Charity registered with the UK Charity Commission. All surplus income from its publishing programme is used to support mathematicians and mathematics research in the form of research grants, conference grants, prizes, initiatives for early career researchers and the promotion of mathematics.

REFERENCES
1. M. Bhargava, Higher composition laws III: the parametrization of quartic rings, Ann. Math. (2) 159 (2004), no. 3, 1329–1360.
2. M. Bhargava, The density of discriminants of quartic rings and fields, Ann. Math. 162 (2005), 1031–1063.
3. M. Bhargava, Higher composition laws IV: the parametrization of quintic rings, Ann. Math. (2) 167 (2008), no. 1, 53–94.
4. M. Bhargava, The density of discriminants of quintic rings and fields, Ann. Math. (2) 172 (2010), no. 3, 1559–1591.
5. M. Bhargava, Mass formulae for extensions of local fields, and conjectures on the density of number field discriminants, Int. Math. Res. Not. 17 (2007), 20.
6. M. Bhargava and P. Harron, The equidistribution of lattice shapes of rings of integers in cubic, quartic, and quintic number fields, Compos. Math. 152 (2016), no. 6, 1111–1120.
7. M. Bhargava, A. Shankar, and J. Tsimerman, On the Davenport–Heilbronn theorems, and second order terms, Invent. Math. 193 (2013), no. 2, 439–499.
8. M. Bhargava, A. Shankar, and X. Wang, Squarefree values of polynomial discriminants I, Invent. Math. 228 (2022), no. 3, 1037–1073.
9. H. Cohen, F. Diaz, Y. Diaz, and M. Olivier, Enumerating quartic dihedral extensions of \(\mathbb{Q}\), Compositio Mathematica 133 (2002), no. 1, 65–93.
10. H. Davenport, On a principle of Lipschitz, J. Lond. Math. Soc. 26 (1951), 179–183. (Corrigendum: “On a principle of Lipschitz”, J. Lond. Math. Soc. 39 (1964), 580.)
11. H. Davenport and H. Heilbronn, On the density of discriminants of cubic fields II, Proc. Roy. Soc. London Ser. A 322 (1971), no. 1551, 405–420.
12. B. N. Delone and D. K. Faddeev, The theory of irrationalities of the third degree, Translations of Mathematical Monographs, vol. 10, American Mathematical Society, Providence, RI, 1964.
13. J. Ellenberg and A. Venkatesh, The number of extensions of a number field with fixed degree and bounded discriminant, Ann. of Math. (2) 163 (2006), no. 2, 723–741.
14. W. T. Gan, B. Gross, and G. Savin, Fourier coefficients of modular forms on \(G_2\), Duke Math. J. 115 (2002), no. 1, 105–169.
15. J. Klüners, A counter example to Malle’s conjecture on the asymptotics of discriminants, C. R. Math. 6 (2005), 411–414.
16. J. Klüners, The distribution of number fields with wreath products as Galois groups, Int. J. Number Theory 8 (2012), no. 3, 845–858.
17. J. Klüners and G. Malle, Counting nilpotent Galois extensions, J. für Reine Angew. Math. 572 (2004), 1–26.
18. R. Lemke Oliver and F. Thorne, Upper bounds on number fields of given degree and bounded discriminant, Duke Math. J. 171 (2015), 3077–3087.
19. R. Lemke Oliver, J. Wang, and M. M. Wood, *Upcoming preprint on extensions with wreath products as Galois groups*.

20. F. Levi, *Kubische Zahlkörper und binäre kubische Formenklassen*, Ber. Sächs. Akad. Wiss. Leipzig, Math.-Naturwiss. Kl 66 (1914), 26–37.

21. G. Malle, *On the distribution of Galois groups, II*, Exp. Math. 13 (2004), no. 2, 129–135.

22. J. P. Serre, *Une “formule de masse” pour les extensions totalement ramifiées de degré donné d’un corps local*, C. R. Acad. Sci. Paris Sér. A-B 286 (1978), no. 22, A1031–A1036.

23. S. Türkelli, *Connected components of Hurwitz schemes and Malle’s conjecture*, J. Number Theory 155 (2015), 163–201.

24. J. Wang, *Malle’s conjecture for $S_n \times A$ for $n = 3, 4, 5$*, Compos. Math. 157 (2021), no. 1, 83–121.

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