The square sieve and the number of A_5-quintic extensions of bounded discriminant*

Manjul Bhargava, Alina Cojocaru, and Frank Thorne

November 11, 2014

Abstract

We prove that the number of quintic fields having associated Galois group A_5 and discriminant less than X is $O(X^{1-\delta})$ for $\delta = \frac{1}{40}$.

1 Introduction

A central problem in arithmetic statistics is that of understanding the asymptotic number of number fields having a given associated Galois group and bounded discriminant. For a permutation group $G \subset S_n$, let $N_n(X, G)$ denote the number of isomorphism classes of number fields K of degree n such that the Galois closure of K over $\mathbb{Q}$ has Galois group isomorphic to G (as a permutation group on the embeddings of K into $\bar{\mathbb{Q}}$). Then the aim is to understand the behavior of $N_n(X, G)$ as $X \to \infty$ for various groups G.

Due to its being the smallest noncyclic finite simple group and the largest finite subgroup of $\mathrm{PGL}_2(\mathbb{C})$, the alternating group $G = A_5$ on 5 letters (also called the icosahedral group) plays a rather special role. For example, A_5-quintic fields are naturally associated to certain weight one holomorphic cuspidal newforms via the work of Deligne and Serre (such forms are thus said to be of "icosahedral type"); information on $N_5(X, A_5)$ thus also gives information about the existence and number of weight one icosahedral cuspforms.

In [2], it was shown that the total number of quintic fields having bounded discriminant is $cX + o(X)$, for an explicit constant $c > 0$; furthermore, 100% of these quintic fields were shown to have associated Galois group S_5. Thus $N_n(X, S_5) = cX + o(X)$, and the best that can be deduced for A_5-quintic fields from the work in [2] is that

$$N_n(X, A_5) = o(X).$$

However, in the applications (e.g., the parallel work of Rohrlich [7] on self-dual two-dimensional characters of the absolute Galois group; see below), it becomes important to have a "power-saving" estimate on $N_5(X, A_5)$, i.e., an estimate of the form $O(X^{1-\delta})$ for some positive δ.

The purpose of this article is precisely to prove such an estimate:

Theorem 1 *We have $N_5(X, A_5) = O(X^{1-\delta})$ for any $\delta < \frac{1}{40}$.*

[*] 2000 AMS Subject Classification: 11R21, 11R45

To illustrate the (mildly unusual) nature of this problem we give not one but three short proofs of this result for varying values of δ. Each uses the results of [2] in combination with a sieve:

- In Section 3, we obtain $\delta = \frac{1}{200} - \epsilon$ using the Selberg sieve, closely following work of Shankar and Tsimerman [8] establishing a power saving error term for S_5-extensions. The same proof

- In Section 4, we obtain $\delta = \frac{1}{120}$ using Heath-Brown's square sieve [5], which was developed precisely to bound the number of squares in arithmetic sequences.

- In Section 5, we obtain $\delta = \frac{1}{40} - \epsilon$ using the *geometric* sieve of the first author, by describing discriminants divisible by q^2 in terms of geometric conditions $\pmod q$.

It seems that $\delta = \frac{1}{40}$ is a natural bottleneck, as all of the proofs use the first author's parametrization [1] of quintic rings by lattice points in a 40-dimensional vector space, and error terms of order $X^{39/40}$ naturally appear in various counts for these lattice points. This does not rule out the possibility of further improving our error terms, but for the moment this appears rather difficult.

In each case, we use the fact that A_5-extensions are precisely those of square discriminant, and in some cases we further apply the fact that A_5-extensions are distinguished by their splitting types at all primes. We expect that $N_5(X, A_5) = c'X^{1/2} \log X$ for some constant $c' > 0$; see, e.g., [6] for general conjectures of this type. It does not appear that any of our methods are capable of proving such a result, but in each section we discuss the limitations and possibilities inherent in each sieve.

Rohrlich has recently used Theorem 1 to prove that self-dual Artin representations of dimension two have density zero among all two-dimensional Artin representations; see [7].

2 Parametrization of quintic rings and fields

We briefly recall the results of [1] and [2] that we will need. For any ring T (commutative, with unit), let V_T denote the space $T^4 \otimes \wedge^2 T^5$ of quadruples of 5×5 skew-symmetric matrices with entries in T. The group $G_T = \mathrm{GL}_4(T) \times \mathrm{SL}_5(T)$ naturally acts on V_T, and there is a natural invariant polynomial of degree 40 for the action of $G_{\mathbb{Z}}$ on $V_{\mathbb{Z}}$, called the *discriminant*, which in fact generates the ring of polynomial invariants. We say that an orbit of G_T on V_T is *nondegenerate* if the discriminant of any element in that orbit is nonzero.

The nondegenerate orbits of G_T on V_T in the case of fields T were first classified by Wright and Yukie [9], and were shown to be in natural correspondence with étale degree 5 extensions of T. The orbits of $G_{\mathbb{Z}}$ on $V_{\mathbb{Z}}$ were classified in [1] in terms of quintic rings and their sextic resolvent rings, as follows.

Theorem 2 *The nondegenerate $G_{\mathbb{Z}}$-orbits on $V_{\mathbb{Z}}$ are in canonical bijection with isomorphism classes of pairs (R, S), where R is a quintic ring and R' is a sextic resolvent ring of R. In this bijection, the discriminant of an element $v \in V_{\mathbb{Z}}$ equals the discriminant of the corresponding ring R of rank n. Furthermore, every isomorphism class of ring R of rank n occurs in this bijection, and every isomorphism class of maximal ring occurs exactly once.*

Recall that a *quintic ring* is a ring R (commutative, with unit) such that R is free of rank 5 as a $\mathbb{Z}$-module. A *sextic resolvent ring* S of R is a ring that is free of rank 6 as a $\mathbb{Z}$-module and which satisfies certain properties which will not be needed here (see [1] for details).

Let $\mathcal{Q}(X)$ denote the set of nondegenerate $G_{\mathbb{Z}}$-orbits on $V_{\mathbb{Z}}$ having absolute discriminant less than X, and let $N(X)$ denote the cardinality of $\mathcal{Q}(X)$. More generally, for any subset $U \subset V_{\mathbb{Z}}$ defined by congruence conditions, let $\mathcal{Q}(U;X)$ denote the set of nondegenerate $G_{\mathbb{Z}}$-orbits on U having absolute discriminant less than X, and let $N(U;X)$ denote the cardinality of $\mathcal{Q}(U;X)$. Then the following theorem was proven in (27) and (28) of [2]:

Theorem 3 *Suppose $U \subset V_{\mathbb{Z}}$ is the union of k distinct translates $L_1, \ldots, L_k$ of the lattice $m \cdot V_{\mathbb{Z}}$. Then, for $m < X^{1/40}$ we have*

$$N(U;X) = km^{-40}\,\frac{\zeta(2)^2\zeta(3)^2\zeta(4)^2\zeta(5)}{2n_i}X + O(km^{-39}X^{39/40}), \tag{1}$$

where the implied constant is independent of k and m.

It is interesting to consider potential improvements to Theorem 3. For example, if U consists of all m^{40} translates of $m \cdot V_{\mathbb{Z}}$, then we obtain an error term of $mX^{39/40}$, even though we are simply counting all of $V_{\mathbb{Z}}$.

Improvements, as well as a relaxation of the condition $m < X^{1/40}$, should be possible if U can be proved to be equidistributed within boxes modulo m. An example of such an improvement, proved by Fourier analytic means, is Lemma 36 of [4], and the techniques applied there can be generalized and possibly improved.

For example, square root cancellation in the relevant exponential sums should lead, very roughly speaking, to being able to regard congruence conditions modulo m as being modulo $m^{1/2}$ instead, with correspondingly improved error terms. This would presumably improve our error terms in our square sieve and Selberg sieve arguments, but probably not past $X^{39/40}$. In contrast, the geometric sieve does not use Theorem 3, and at present it is unclear if this argument could be improved.

2.1 The square sieve

We prove Theorem 1 by bounding the number of squares in $\mathcal{A}(X)$, using the *square sieve*, which Heath-Brown introduced in [5] to estimate the number of consecutive squarefree integers. Heath-Brown's result is the following:

Proposition 4 (Heath-Brown [5]) *Let $\mathcal{A}$ denote any multiset of nonzero integers of absolute value at most X, and let $\mathcal{P}$ be any set of primes with $e^{|\mathcal{P}|} > X$. If $S(\mathcal{A})$ denotes the number of squares in $\mathcal{A}$, then we have*

$$S(\mathcal{A}) \ll \frac{|\mathcal{A}|}{|\mathcal{P}|} + \frac{1}{|\mathcal{P}|^2}\sum_{p \neq q \in \mathcal{P}}\left|\sum_{n \in \mathcal{A}}\left(\frac{n}{pq}\right)\right|. \tag{2}$$

Proof: The proof is simple and so we recall it here, following [5]. Denote

$$R := \sum_{n \in \mathcal{A}}\left(\sum_{p \in \mathcal{P}}\left(\frac{n}{p}\right)\right)^2. \tag{3}$$

The inner sum is nonnegative for all n, and for square n we observe that

$$\sum_{p \in \mathcal{P}}\left(\frac{n}{p}\right) = \sum_{p \in \mathcal{P},\, p \nmid n} 1 \gg |\mathcal{P}|, \tag{4}$$

so long as X is sufficiently large. (Here we have applied our lower bound on $|\mathcal{P}|$ relative to X.) Now expanding the square in (3) and switching the order of summation, we obtain

$$R = \sum_{p \in \mathcal{P}} \sum_{n \colon p \nmid n} 1 + \sum_{p \neq q \in \mathcal{P}} \sum_{n} \left(\frac{n}{pq} \right) \leq |\mathcal{P}||\mathcal{A}| + \sum_{p \neq q \in \mathcal{P}} \left| \sum_{n} \left(\frac{n}{pq} \right) \right|. \tag{5}$$

Comparing (4) and (5) and dividing through by $|\mathcal{P}|^2$, we obtain the result. □

We can now prove our main theorem. As before, let $\mathcal{A} = \mathcal{A}(X)$ be the multiset of discriminants of quintic rings (counted once for each sextic resolvent ring), and let $\mathcal{P}$ consist of all primes $< X^\alpha$, where $\alpha > 0$. Applying Proposition 5 in Proposition 4, we see that the number of squares in $\mathcal{A}(X)$ satisfies the bound

$$S(\mathcal{A}) \ll \frac{|\mathcal{A}|}{|\mathcal{P}|} + \frac{1}{|\mathcal{P}|^2} \sum_{p \neq q \in \mathcal{P}} \left| \sum_{n \in \mathcal{A}} \left(\frac{n}{pq} \right) \right| \ll X^{1-\alpha} \log X + X^{39/40 + 2\alpha}, \tag{6}$$

and we conclude by taking $\alpha = 1/120$.

3 Proof via the Selberg sieve

In a beautiful short paper [8], Shankar and Tsimerman obtained a power-saving error term in the counting function for S_5-quintc fields, via the Selberg sieve. Their proof is quite simple and flexible, and indeed our application here will illustrate the versatility of their methods.

The key improvement of [8] was an improved bound for points $x \in V_{\mathbb{Z}}$ in an average of fundamental domains with $a_{12} \neq 0$ (i.e., not in the cusp), which don't correspond to S_5-quintic rings. (By Lemma 11 – of [2], the number of irreducible x with $a_{12} = 0$ is $\ll X^{39/40}$, so that these may be disregarded.) Because S_5 is generated by a 5-cycle and a transposition, any such x must either not have splitting type (5) at any prime, or must not have splitting type (1112) at any prime. Writing $S_p(5)$ and $S_p(1112)$ for the sets of $x \in V_{\mathbb{Z}}$ with $a_{12} \neq 0$ whose splitting type is respectively not (5) or (1112), Shankar and Tsimerman proceed by giving good bounds for $\cap_p S_p(5)$ and $\cap_p S_p(1112)$.

At this point we may simply quote their results. Because any $x \in V_{\mathbb{Z}}$ corresponding to an A_5-quintic ring cannot have splitting type (1112) for any prime, we obtain

$$N_5(X, A_5) \ll_\epsilon X^{\frac{199}{200} + \epsilon} \tag{7}$$

by their bound on $\cap_p S_p(1112)$.

For the convenience of the reader, we recall their argument. For any subset $S \subseteq V_{\mathbb{Z}}$, write $N^*_{12}(S, X)$ for the averaged number of $x \in S$ with $a_{12} \neq 0$ and $|\operatorname{Disc}(x)| < X$, where this averaging is defined in (1) of [8] and originally in [2].

Fix a parameter $z < X$, write $P(z) = \prod_{p<z} p$, and for each squarefree $d \mid P(z)$ write

$$a_d = N^*_{12}\left(\bigcap_{p \mid d} T_p(1112) \bigcap_{p \mid \frac{P(z)}{d}} S_p(1112), X \right), \tag{8}$$

where $T_p(1112)$ is the complement of $S_p(1112)$, and $a_d = 0$ for $d \nmid P(z)$.

By Theorem 3, equivalently (4) of [8][1], we have

$$\sum_{n\equiv 0 \pmod{d}} a_n = N_{12}^*(\cap_{p|d} T_p(1112), X) = c_i g_d(1112)X + r_d, \tag{9}$$

where the density $g_d(1112)$ is multiplicative in d and satisfies $g_p(1112) = \frac{1}{12} + O(p^{-1})$, and $r_d = O(dg_d(5)X^{39/40})$. Defining

$$h_d(1112) = \prod_{p|d} \frac{g_p(5)}{1 - g_p(5)}, \quad H = \sum_{\substack{d<\sqrt{D} \\ d|P(z)}} h_d(5), \tag{10}$$

the Selberg sieve yields that

$$a_1 \leq c_i X H^{-1} + O\Bigg(\sum_{d<D,\ d|P(z)} \tau_3(d) r_d \Bigg). \tag{11}$$

Choosing $z = \sqrt{X}$ and $D = X^{1/100}$, and using that

$$d^{-\epsilon} \ll_\epsilon g_d(1112), h_d(1112) \ll_\epsilon d^\epsilon, \tag{12}$$

a quick computation shows that $a_1 = N_{12}^*(\cap_{p<z} S_p(1112)) \ll_\epsilon X^{199/200+\epsilon}$, as desired.

One naturally asks if we can obtain a better bound by sieving out more, as (1112) is not the only splitting type forbidden for A_5-quintic fields. The answer seems to be no; by [1], the splitting types forbidden for A_5-quintic fields have total density $\frac{1}{2} + O(p^{-1})$ as opposed to $\frac{1}{12} + O(p^{-1})$, but all that is used in [8] is that this density is bounded away from 0 and 1.

4 Proof via the square sieve

4.1 Equidistribution of Jacobi symbols

Now let $\mathcal{A}(X)$ denote the multiset of absolute discriminants of elements in $\mathcal{Q}(X)$. We use Theorem 3 to deduce the following equidistribution result:

Proposition 5 *Let m be a squarefree integer, and let $\mathcal{A}(X)$ be as above. Then*

$$\Bigg| \sum_{n\in\mathcal{A}(X)} \left(\frac{n}{m}\right) \Bigg| = O(mX^{39/40}). \tag{13}$$

Proof: Let $U_\pm$ denote the multisubset of elements $n \in \mathcal{A}(X)$ such that $\left(\frac{n}{m}\right) = \pm 1$. Then $U_\pm$ is the union of $k_\pm = k_{m,\pm}$ translates of $m \cdot V_{\mathbb{Z}}$ for some positive integers $k_\pm$. By Theorem 3, we have

$$\Bigg| \sum_{n\in\mathcal{A}(X)} \left(\frac{n}{m}\right) \Bigg| = (k_+ - k_-)m^{-40}\, \frac{\zeta(2)^2\zeta(3)^2\zeta(4)^2\zeta(5)}{2n_i} X + O((k_+ + k_-)m^{-39}X^{39/40}). \tag{14}$$

[1] The statements of Theorem 3 and (4) of [8] are not identical, but are equivalent in the cases of interest. The condition that $m < X^{1/40}$ is not stated in [8], but this result is never applied with $m > X^{1/40}$. Similarly, the condition that $a_{12} \neq 0$ is included in (4) of [8] and not in Theorem 3; in all cases where we apply this result, the error term is larger than the error term $O(X^{39/40})$ of Lemma 11 of [2] discussed earlier.

We first prove the proposition in the case when m is a prime p. By [1, §12], an element $v \in V_{\mathbb{Z}}$ has a nonzero square discriminant modulo p if and only if the corresponding quintic ring R, when reduced modulo p, has square discriminant in $\mathbb{F}_p^\times$, i.e., if and only if R has splitting type (11111), (113), (22), or (5) at p. Similarly, an element $v \in V_{\mathbb{Z}}$ has a nonsquare discriminant modulo p if and only if the corresponding quintic ring R, when reduced modulo p, has splitting type (1112), (23), or (4) at p. For each of these 7 splitting types σ, there is a unique quintic ring $R_p(\sigma)$ over $\mathbb{Z}_p$ having the splitting type σ at p. By [2, (29)], we then have

$$(15)\quad k_+ = p^{40}\mu(U_+) = \sum_{\sigma\in\{(11111),(113),(22),(5)\}} \frac{|G_{\mathbb{F}_p}|}{|\operatorname{Aut}(R(\sigma))|} = |G_{\mathbb{F}_p}|\cdot\Big(\frac{1}{120}+\frac{1}{6}+\frac{1}{8}+\frac{1}{5}\Big) = \frac{1}{2}\cdot|G_{\mathbb{F}_p}|.$$

Similarly,

$$(16)\quad k_- = p^{40}\mu(U_-) = \sum_{\sigma\in\{(1112),(23),(4)\}} \frac{|G_{\mathbb{F}_p}|}{|\operatorname{Aut}(R(\sigma))|} = |G_{\mathbb{F}_p}|\cdot\Big(\frac{1}{12}+\frac{1}{6}+\frac{1}{4}\Big) = \frac{1}{2}\cdot|G_{\mathbb{F}_p}|.$$

Combining (14), (15), and (16) with the fact that

$$G(\mathbb{F}_p) = p^{40} + O(p^{39})$$

now yields (13) in the case that $m = p$.

If $m = p_1\cdots p_r$, the Chinese Remainder Theorem implies that $k_{m,+} = \sum k_{p_1,\pm}\cdots k_{p_r,\pm}$, where the sum is over the 2^{r-1} choices of signs for the p_i where an even number of signs are negative. By the above calculations, each summand is equal to $2^{-r}\prod_i |G_{\mathbb{F}_{p_i}}| = 2^{-r}\cdot|G_{\mathbb{Z}/m\mathbb{Z}}|$, so that $k_{m,+} = \frac{1}{2}|G_{\mathbb{Z}/m\mathbb{Z}}| < m^{40}$. For $k_{m,-}$, we sum over the 2^{r-1} choices of sign where an odd number of signs are negative, so that $k_{m,-} = \frac{1}{2}|G_{\mathbb{Z}/m\mathbb{Z}}|$ as well, yielding (13). □

5 Proof via the geometric sieve

Finally, we will obtain the strongest error terms using the *geometric* sieve of the first author [3].

The idea to be pursued is the following. A_5 fields K with discriminant in $[X^{1/2}, X]$ have the property that for some squarefree $q \gg X^{1/8}$ with $(q, 30) = 1$, we have $v_p(\mathrm{Disc}(K)) \in \{2, 4\}$ for each prime p dividing q. (We take q coprime to 30 so as to avoid complications due to wild ramification; also note that there are $\ll X^{1/2}$ A_5-fields K with discriminant $< X^{1/2}$.)

In most sieve applications, such K would constitute the 'tail' of elements to be sieved *out*; for example, in [3] bounds for this tail yield asymptotics for fields with square-*free* discriminant. Here the 'tail' is precisely what we want to count.

We apply the geometric sieve as follows. Suppose that, as above, $x \in V_{\mathbb{Z}}$ corresponds to a maximal quintic order and satisfies $v_p(\mathrm{Disc}(x)) \geq 2$ (and in particular $v_p(\mathrm{Disc}(x)) \in \{2, 4\}$) for each prime $p \mid q$, for some squarefree $q > X^{1/8}$. Then, $\mathrm{Disc}(x')$ will be divisible by p^2 for each $x' \equiv x \pmod{p^2}$; in the language of [3], the discriminant polynomial is strongly a multiple of p^2 at x. By Lemma 3.6 of [3] it follows that there is a subscheme Y_2 of $\mathbb{A}^{40}_{\mathbb{Z}}$ such that $x \pmod p \in Y_2(\mathbb{F}_p)$ for each $p \mid q$; indeed, we may choose one of the 40 coordinates x_i arbitrarily, and then Y_2 is defined by $\mathrm{Disc} = \frac{\partial\,\mathrm{Disc}}{\partial x_i} = 0$.

Let B be the intersection of the set $\{x \in V_{\mathbb{R}} : 0 < |\operatorname{Disc}(x)| < 1\}$ and a fundamental domain for the action of $G_{\mathbb{Z}}$ on $V_{\mathbb{R}}$. Then, with $r = X^{1/40}$, $rB \cap \mathbb{Z}^{40}$ corresponds precisely to the set of $G_{\mathbb{Z}} \backslash V_{\mathbb{Z}}$ with $0 < |\operatorname{Disc}(x)| < X$. By Theorem 3.3 of [3], we have

$$\#\{a \in rB \cap \mathbb{Z}^{40} \ \mid \ a \pmod{p} \in Y(\mathbb{F}_p) \text{ for some prime } p > X^{1/8}\} = O(X^{39/40+\epsilon}); \tag{17}$$

the idea of the proof is to project onto one of the 40 coordinates x_i and prove that for each value a_i of the coordinate x_i, only finitely many $a \in rB \cap \mathbb{Z}^{40}$ counted above have x_i-coordinate equal to a_i.

Indeed, (17) is also valid for general squarefree q, as the proof of Theorem 3.3 of [3] remains equally valid for squarefree q, except that we obtain an additional factor of X^{ϵ} in the error term at three locations:

- In the verification of (17) of [3] for $k = 1$, $f_i(a)$ may have $\ll_{\epsilon} X^{\epsilon}$ squarefree divisors $>$, as opposed to simply $O(1)$ prime factors $p > r$.
- Similarly, in the verification of (19) of [3], $f_k(b, a_n)$ may have $\ll_{\epsilon} X^{\epsilon}$ squarefree divisors $> r$, as opposed to $O(1)$ prime factors $p > r$.
- In the verification of (21) of [3], the number of values of a_n such that $f_k(b, a_n) \equiv 0 \pmod{p}$ and $a = (b, a_n) \in rB \cap \mathbb{Z}^n$ is now $d^{\nu(q)} \cdot O(1)$, where $\nu(q)$ denotes the number of prime divisors of q. This quantity is once again $\ll_{\epsilon,d} X^{\epsilon}$.

We are therefore done, except for one technical point: Theorem 3.3 of [3] assumes that B is compact, which in our case it is not. This is dealt with in [3] (just before (27)) by removing a region of B of volume ϵ; however, this doesn't suffice for a power-saving error term. (**To do.** This can be dealt with following Remark 4.2; check and write up the details.)

Remark 6 *The proof of* (17) *illustrates that* $X^{39/40}$ *is a natural bottleneck with this method, as it is the volume of our projection. To go further, we would have to prove that for most* a_i*,* no $a \in rB \cap \mathbb{Z}^{40}$ *have* x_i*-coordinate equal to* a_i*.*

This completes the proof.

References

[1] M. Bhargava, Higher composition laws IV: The parametrization of quintic rings, *Ann. Math.* **167** (2008), no. 1, 53–94.

[2] M. Bhargava, The density of discriminants of quintic rings and fields, *Ann. of Math.* **72** (2010), 1559–1591.

[3] M. Bhargava, The geometric sieve and the density of squarefree values of invariant polynomials, preprint.

[4] M. Bhargava, A. Shankar and J. Tsimerman, On the Davenport–Heilbronn theorems and second order terms, *Invent. Math.* **193**, 439-499.

[5] D. R. Heath-Brown, *The square sieve and consecutive square-free numbers*, *Math. Ann.* **266** (1984), no. 3, 251–259.

[6] G. Malle, On the distribution of Galois groups II, *Experiment. Math.* **13** (2004), 129–135,

[7] D. Rohrlich, Self-dual Artin representations, preprint.

[8] A. Shankar and J. Tsimerman, Counting S_5-fields with a power saving error term, preprint.

[9] D. J. Wright and A. Yukie, Prehomogeneous vector spaces and field extensions, *Invent. Math.* **110** (1992), 283–314.