LTCC Advanced Course: Arithmetic Statistics

Christopher Keyes

Updated February 15, 2025

Course description

Arithmetic statistics encompasses a broad range of quantitative problems involving the distribution of number theoretic objects. This course is an introduction to the area through the problem of counting number fields by discriminant. After reviewing some preliminaries from algebraic number theory, we will explore how to count extensions of small degree by discriminant with prescribed Galois group, with the goal of covering in detail Davenport and Heilbronn's famous results on S_3-cubic extensions. We will then touch on some of Bhargava's more recent innovations in this area and their applications, and get a sense of the current state of the art for higher degrees.

Related reading

The central topic of counting cubic extensions follow Davenport and Heilbronn's original paper [DH71] and Bhargava, Shankar, and Tsimerman's more recent paper [BST13]. Other authors have written and published good notes on these topics, including Achenjang [Ach] and Landesman [Lan]. There are excellent notes from the Arizona Winter School on related topics, including (but not limited to) Bhargava's 2009 notes on parametrizing rings [Bha09] and Wood's 2014 notes on arithmetic statistics [Woo14]. (Better still, you can watch these lecture recordings online.)

1 Introduction

Arithmetic statistics is composed of two words, which loosely have the following meanings.

- **Arithmetic** (adj): relating to number theory.
- **Statistics** (noun): the study of collecting and analyzing quantitative data.

Put together, **arithmetic statistics** encompasses questions about number theoretic objects — think primes, number fields, elliptic curves — that are quantitative in nature — think probability, averages, or distributions. Let's start with some examples.

1.1 Primes

Several classical questions about the primes can framed as statistical.

Question 1.1. *How many primes are there?*

Answer 1.2 (Easy)**.** There are infinitely many primes. The proof is an exercise. :)

To give a more precise answer that sheds some light on how the primes are distributed, we need a way to count primes.

Definition 1.3. Define the prime counting function

$$\pi(X) = \#\{1 \leq p \leq X : p \text{ is prime}\}.$$

Here, the choice of counting function is rather obvious. (Can you think of a different one?) Still, we have made a choice, and it's worth thinking about why it's a good one, and what other options are out there.

- For all positive real X, $\pi(X)$ is finite. This is a key property of a counting function.
- It is easy to count all the natural numbers in $[1, X]$, to then compare with $\pi(X)$. This will not be true for all such problems; indeed, a great deal of current research goes into just counting things like number fields or elliptic curves, before counting those which have certain properties.
- We could have chosen $\pi'(X)$ to count the primes in $[-X, X]$. This satisfies both bullet points above. Of course, we have $\pi'(X) = 2\pi(X)$, so we lose nothing by specializing to positive numbers. (One could argue that sticking to positive is better; this hints at how one might want to count primes in a number field...)

Let us briefly recall some asymptotic notation which we will need throughout the course.

Definition 1.4 (Asymptotic notation)**.** Assume all functions are real valued.

- $f(X) = O(g(X))$ means there is a constant c and a number N such that
$$f(X) \leq cg(X) \text{ for all } X \geq N.$$
- $f(X) \ll g(X)$ is another common notation for $f(X) = O(g(X))$.
- $f(X) \sim g(X)$ means
$$\lim_{X\to\infty} \frac{f(X)}{g(X)} = 1.$$
This is quite a strong statement.
- $f(X) = o(g(X))$ means
$$\lim_{X\to\infty} \frac{f(X)}{g(X)} = 0,$$
i.e. that f grows slower than g asymptotically. In particular, if $f(X) = o(1)$ then $f(X) \to 0$.

With a counting function in hand, we can give a better answer to Question 1.1.

Theorem 1.5 (Prime number theorem, ∼1896)**.** *We have*

$$\pi(X) \sim \frac{X}{\log X}.$$

This long sought after theorem answers other statistical questions about the prime numbers. In particular, since $\pi(X) = o(X)$, we can see that the primes make up only "0%" of the natural numbers.

Example 1.6 (Naive attempt at proving PNT)**.** Let $\pi(X, X')$ denote the number of primes in $[X, X']$. To prove the prime number theorem, it is good enough to show $\pi(\sqrt{X}, X) \sim \frac{X}{\log X}$.

A number $n \in [\sqrt{X}, X]$ is prime if and only if for all primes $p \leq \sqrt{X}$ we have $p \nmid n$. One might be led to *guess* that the "probability" of $n \in [\sqrt{X}, X]$ being prime is equal to $\prod_{p \leq \sqrt{X}}(1 - \frac{1}{p})$. This relies on a naive heuristic assumption that these divisibility conditions are independent, so the product agrees with the density of primes in $[\sqrt{X}, X]$.

However, Mertens' product theorem (which *predates* the prime number theorem) states that

$$\prod_{p \leq X} \left(1 - \frac{1}{p}\right) \sim \frac{e^{-\gamma}}{\log X},$$

where $\gamma \approx 0.577$ is the Euler–Mascheroni constant. Thus our product grows like

$$\prod_{p \leq \sqrt{X}} \left(1 - \frac{1}{p}\right) \sim \frac{e^{-\gamma}}{\log \sqrt{X}} = \frac{2e^{-\gamma}}{\log X},$$

and $2e^{-\gamma} \approx 1.123$.

Thus our naive independence heurstic suggests $\pi(\sqrt{X}, X) \sim 1.123X/\log X$, which of course is the wrong answer! This tells us that the distribution of primes in intervals like $[\sqrt{X}, X]$ is not quite so nice, so we should be careful about making such independence assumptions.

With something of a handle on counting primes, we'd like to understand how they are distributed, in various ways. One such way is to look at arithmetic progressions. Let a, b be coprime positive integers and define

$$\pi(X; a, b) = \{1 \leq p \leq X : p \text{ is prime and } p \equiv a \pmod{b}\}.$$

Theorem 1.7. *We have*

$$\pi(X; a, b) \sim \frac{\pi(X)}{\varphi(b)}$$

where φ is Euler's phi function.

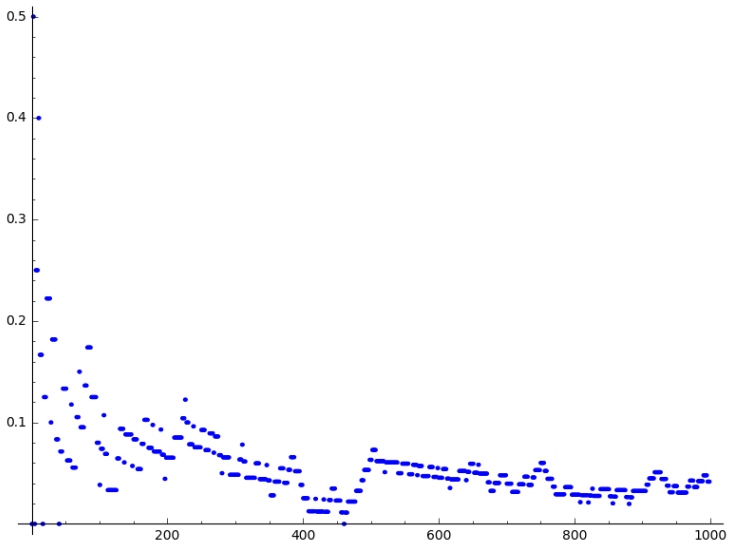
While we won't make this rigorous, this theorem may be interpreted as saying that the primes are *equidistributed* among residue classes of units modulo b. As a concrete example, if $b = 4$, then a(n odd) prime p is either 1 or 3 modulo 4, and each of these happens 50% of the time.

Remark 1.8 (Prime number races)**.** With that said, there are *biases* in the distribution of primes in arithmetic progressions. Let's continue looking at primes modulo 4.

If we plot

$$\frac{\pi(X; 3, 4) - \pi(X; 1, 4)}{\pi(X)}$$

for $X \leq 1000$ we obtain the following.



Notice that this function is nonnegative; up to $X = 1000$, we never have *more* primes that are 1 modulo 4 than 3 modulo 4.

We have to go all the way to $X = 26861$ to find the first instance of this function dipping negative, and in fact it does so infinitely many times. However, it can be made precise that for 100% of X, the function is nonnegative. If you like this kind of thing, mess around with the computer (see the exercises), then search for "prime number races" or see this survey article [GM04].

1.2 Number fields

Let $\mathbb{Q}$ denote the field of rational numbers.

Definition 1.9. A **number field** is a finite extension $K/\mathbb{Q}$. Its **degree**, $[K : \mathbb{Q}]$, is equal to its dimension as a $\mathbb{Q}$-vector space.

Question 1.10. *For a given degree d, how many number fields are there?*

Like with the prime numbers, it's not so difficult to show that there are infinitely many number fields $K/\mathbb{Q}$ of degree $d > 1$. Again, we need to agree on a way to count them.

Definition 1.11 (Discriminant)**.** Let $\mathcal{O}_K$ be the ring of integers of K. Write $\mathcal{O}_K = \mathbb{Z}[\beta_i]$ as a free $\mathbb{Z}$-module, and let $\iota_j \colon K \to \overline{\mathbb{Q}}$ be the d embeddings of K into the algebraic closure. The **discriminant** of $K/\mathbb{Q}$ is

$$\mathrm{Disc}(K/\mathbb{Q}) = \det\left(\iota_j(\beta_i)\right)^2.$$

Another common equivalent definition uses the trace:

$$\mathrm{Disc}(K/\mathbb{Q}) = \det\left(\mathrm{Tr}(\beta_i\beta_j)\right).$$

The discriminant is a reasonable way to count number fields:

- It is actually a measure of size! Minkowski theory relates the discriminant to the volume of a fundamental domain for $\mathcal{O}_K$, viewed as a lattice in Minkowski space; see e.g. [Neu99, I.5].
- The number of $K/\mathbb{Q}$ of degree d with discriminant $|\mathrm{Disc}(K/\mathbb{Q})| \leq X$ is finite. This is known as Hermite's theorem, and its proof also uses Minkowski theory; see e.g. [Neu99, III.2].
- It is *isomorphism invariant*. This means that if we want to count *isomorphism classes* of number fields (which we will), we can still use it.
- The discriminant controls ramification: a prime p is ramified in K if and only if $p \mid \mathrm{Disc}(K/\mathbb{Q})$.

Remark 1.12. There are other invariants used to count number fields. One is called the **conductor**, a different measure of ramification for number fields. We won't talk about it further, but there is significant debate among researchers over what is the "right" invariant for these counting problems.

Definition 1.13. Let $N_n(X)$ denote the number of isomorphism classes of degree n extensions $K/\mathbb{Q}$ with discriminant $\mathrm{Disc}(K/\mathbb{Q}) \leq X$.

We can now state some results for low degrees. We will return to the $n = 2, 3$ cases in particular later in this course.

Theorem 1.14. *For $2 \leq n \leq 5$ there exists a constant c_n such that*

$$N_n(X) \sim c_n X.$$

Degree 2 is classical, closely related to the problem of counting squarefree numbers. The degree 3 case is due to Davenport and Heilbronn [DH71], and is much more involved. We will spend a good chunk of this course setting up and working through their original paper. The $n = 4, 5$ cases are celebrated work of Bhargava [Bha05, Bha10]. It is conjectured that similar results hold for all n.

Conjecture 1.15. *For all $n \geq 2$, there exists a constant c_n such that*

$$N_n(X) \sim c_n X.$$

We can also count number fields subject to certain conditions, much in the same way that we counted primes in arithmetic progressions earlier. A natural choice of condition is to specify the *Galois group*.

Definition 1.16 (G-extensions)**.** For $K/\mathbb{Q}$ of degree d, let $\widetilde{K}$ denote its Galois closure, i.e. the smallest Galois extension of $\mathbb{Q}$ containing K. If $\mathrm{Gal}(\widetilde{K}/\mathbb{Q}) \simeq G$, then we say $K/\mathbb{Q}$ is a **degree n G-extension** of $\mathbb{Q}$.

We can also define a counting function $N_n(X; G)$, which counts the number of such extensions with discriminant at most X.

Warning. When people say "count number fields with Galois group G," they (typically) don't actually mean that they are counting Galois extensions $K/\mathbb{Q}$ with $\mathrm{Gal}(K/\mathbb{Q}) = G$. They almost certainly mean counting G-extensions in the sense defined above.

We can now expand on Theorem 1.14.

Theorem 1.17. *In low degrees n and for certain G, we have asymptotics for $N_n(X;G)$.*

$$N_2(X;S_2) \sim \frac{X}{\zeta(2)} \tag{1.1}$$

$$N_3(X,C_3) \sim o(X) \tag{1.2}$$

$$N_3(X,S_3) \sim \frac{X}{3\zeta(3)} \qquad \textit{(Davenport–Heilbronn)} \tag{1.3}$$

$$N_4(X,D_4) \sim c_4(D_4)X \qquad \textit{(Cohen–Diaz y Diaz–Olivier)} \tag{1.4}$$

$$N_4(X,S_4) \sim c_4(S_4)X \qquad \textit{(Bhargava)} \tag{1.5}$$

$$N_5(X,S_5) \sim c_5(S_5)X \qquad \textit{(Bhargava)} \tag{1.6}$$

$$\tag{1.7}$$

More precise conjectures have been made for $N_n(X;G)$ in general.

Conjecture 1.18 (Malle, 2002)**.** *Fix a degree $n \geq 2$ and a transitive subgroup $G \subseteq S_n$. We have*

$$N_n(X;G) \sim c_n(G) X^{a(G)} (\log X)^{b(G)-1},$$

where $a(G)$ and $b(G)$ are invariants related to the group G, viewed as a permutation group on n elements.

Regrettably, the conjecture is false in general, with counterexamples given by Klüners [Klü05]. However, it is known to hold in several cases, including for small degrees as above, for abelian groups G, and for certain G arising from semidirect products of other groups where the conjecture is known. Very recent work of Loughran and Santens suggests an explicit description of $c_n(G)$ and a way to modify the conjecture to account for the known counterexamples; the introduction also lists many references if you are interested in further reading [LS24].

1.2.1 Local behavior

Let $K/\mathbb{Q}$ be *Galois* extension with ring of integers $\mathcal{O}_K$. Recall that $\mathcal{O}_K$ is a Dedekind domain, i.e. its ideals factor uniquely into prime ideals. If p is an integer prime, then

$$p\mathcal{O}_K = \left(\prod_{i=1}^{g} \mathfrak{p}_i \right)^e,$$

where the $\mathfrak{p}_i$ are distinct primes of $\mathcal{O}_K$. Let $\mathbb{F}_\mathfrak{p} = \mathcal{O}_K/\mathfrak{p}$ denote the residue field, which is finite. In the Galois case, we have $\mathbb{F}_{\mathfrak{p}_i} \simeq \mathbb{F}_{\mathfrak{p}_j}$, so we can safely denote them all by $\mathbb{F}_\mathfrak{p}$.

Let's quickly review some terms to describe prime splitting in number fields.

Definition 1.19 (Inertia and ramification)**.** The **inertia degree** of $\mathfrak{p} \mid p$ is the degree of $\mathbb{F}_\mathfrak{p}/\mathbb{F}_p$, denoted f.

The **ramification degree** of $\mathfrak{p} \mid p$ is the exponent e in the factorization. An important fact is

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

There are some important special cases:

- if $e = 1$ we say p is **unramified**;
- conversely, if $e > 1$, we say p is **ramified**, and if $g = 1$ then p is **totally ramified**;
- when $p\mathcal{O}_K$ itself is prime, we say p is **inert**, which coincides with $g = 1$, $e = 1$, and $f = [K : \mathbb{Q}]$;
- if p is unramified and $f = 1$ for all i, then we say p is **totally split**.

Definition 1.20 (Decomposition and inertia groups)**.** Let the **decomposition group**, $D_{\mathfrak{p}} \subset \mathrm{Gal}(K/\mathbb{Q})$ denote the subgroup of $\mathrm{Gal}(K/\mathbb{Q})$ consisting of automorphisms that fix $\mathfrak{p}$.

There is a natural exact sequence

$$0 \to I_{\mathfrak{p}} \to D_{\mathfrak{p}} \to \mathrm{Gal}(\mathbb{F}_{\mathfrak{p}}/\mathbb{F}_p) \to 0$$

where the kernel, $I_{\mathfrak{p}}$ is called the **inertia group**.

When $\mathfrak{p} \mid p$ is unramified, the inertia group vanishes, so $D_{\mathfrak{p}}$ is cyclic. In particular, the Frobenius map, which generates $\mathrm{Gal}(\mathbb{F}_{\mathfrak{p}}/\mathbb{F}_p)$ has a unique preimage in $D_{\mathfrak{p}}$. Moreover, since each $D_{\mathfrak{p}}$ is conjugate to $D_{\mathfrak{p}'}$ for $\mathfrak{p}, \mathfrak{p}' \mid p$, the respective Frobenius elements form a *conjugacy class* in $\mathrm{Gal}(K/\mathbb{Q})$.

Definition 1.21 (Frobenius conjugacy class)**.** For a prime p, the **Frobenius class** $\mathrm{Frob}_p \subset \mathrm{Gal}(K/\mathbb{Q})$ is this conjugacy class.

Example 1.22 (Quadratic fields)**.** Suppose $K/\mathbb{Q}$ is a quadratic field. Since $\mathrm{Gal}(K/\mathbb{Q}) \simeq C_2$ is abelian, a conjugacy class corresponds to a group element. If $p\mathcal{O}_K = \mathfrak{p}\overline{\mathfrak{p}}$ is split, Then $D_{\mathfrak{p}}, D_{\overline{\mathfrak{p}}}$ are both trivial, and Frob_p is the identity of C_2.

In the inert case, $\mathfrak{p} = p\mathcal{O}_K$, we have $D_{\mathfrak{p}} \simeq \mathrm{Gal}(\mathbb{F}_{\mathfrak{p}}/\mathbb{F}_p) \simeq C_2$, so its generator must also generate the Galois group.

Let's revisit the prime number theorem, but this time count only those primes with prescribed Frobenius class. Let

$$\pi(X; K, C) = \#\{1 \le p \le X : \mathrm{Frob}_p = C\},$$

for a fixed conjugacy class $C \subset \mathrm{Gal}(K/\mathbb{Q})$. Chebotarev's theorem is powerful generalization of Theorem 1.7.

Theorem 1.23 (Chebotarev Density Theorem)**.** *Let $K/\mathbb{Q}$ be a finite Galois extension. We have*

$$\pi(X; K, C) \sim \frac{\#C}{\#\,\mathrm{Gal}(K/\mathbb{Q})} \cdot \frac{X}{\log X}.$$

That is, the primes with a prescribed Frobenius class have density proportional to the size of Frob_p as a conjugacy class. This is in fact stronger than Theorem 1.7, and proving so is a good exercise.

1.2.2 Class groups

Let's get back to counting number fields, rather than primes, but this time count by their class groups.

Definition 1.24 (Class group)**.** Let $K/\mathbb{Q}$ be a number field. There is an exact sequence

$$1 \to K^\times/\mathcal{O}_K^\times \to \oplus_{\mathfrak{p} \text{ prime}} \mathbb{Z} \to \mathrm{Cl}(K) \to 0,$$

where the left hand map is div, which sends an element $x \in K^\times$ to the primes dividing its numerator minus those dividing its denominator (with multiplicity). The **ideal class group** is the quotient $\mathrm{Cl}(K)$.[1]

The class group measure the failure of $\mathcal{O}_K$ to be a unique factorization domain. Even in the simplest case of $K/\mathbb{Q}$ a quadratic field, they remain somewhat mysterious. Here are some things we know and don't know:

- there are finitely many *imaginary* quadratic fields with trivial $\mathrm{Cl}(K)$;
- in fact, as $\mathrm{Disc}(K/\mathbb{Q})$ grows, so does $\mathrm{Cl}(K)$ for an imaginary quadratic;
- it is predicted, but not known, that there are infinitely many *real* quadratic fields with trivial $\mathrm{Cl}(K)$.

Let's focus on imaginary quadratics $K/\mathbb{Q}$. One sensible question might be "how often is the order of $\mathrm{Cl}(K)$ divisible by a prime p?" Or, "what is the average size of $\mathrm{Cl}(K)[p]$?" The case of $p = 2$ is rather classical; see the exercises.

For odd p, we can make sense of the latter question via the Cohen–Lenstra heuristics.

Conjecture 1.25 (Cohen–Lenstra)**.** *Let p be an odd prime. Then the average p-torsion in the class group of an imaginary quadratic field is*

$$\lim_{X\to\infty} \frac{\sum_{K/\mathbb{Q} \text{ im. quad.}} \#\,\mathrm{Cl}(K)[p]}{\sum_{K/\mathbb{Q} \text{ im. quad.}} 1} = 2.$$

This conjecture, and many other statements about average class group behavior, are based on the following heuristic assumption: *class groups behave like random finite abelian groups A, weighted by $\frac{1}{\#\,\mathrm{Aut}(A)}$*. Conjecture 1.25 boils down to equating the average on the left-hand-side with

$$\frac{\sum_A \frac{\#A[p]}{\#\,\mathrm{Aut}(A)}}{\sum_A \frac{1}{\#\,\mathrm{Aut}(A)}},$$

where A runs over all finite abelian p-groups. In their seminal paper, Cohen and Lenstra actually computed this ratio to be 2, along with a number of other averages and probability results for finite abelian groups assuming this distribution.

Remark 1.26. Conjecture 1.25 is only known to hold for $p = 3$, due to Davenport and Heilbronn [DH71, Theorem 3]. We will come back to this later.

1.3 Elliptic curves

Definition 1.27 (Elliptic curve)**.** Let K be a field. An **elliptic curve** E/K is a smooth projective geometrically integral curve of genus 1.

Alternatively, and more explicitly, if $\mathrm{char}(K) \neq 2, 3$, an elliptic curve is isomorphic to a plane cubic curve given by a **short Weierstrass equation**

$$E\colon y^2 = x^3 + ax + b,$$

where $4a^3 + 27b^2 \neq 0$ (for smoothness).

[1] The middle group, $\oplus_{\mathfrak{p}}\mathbb{Z}$ should really be thought of as the group of fractional ideals in K.

A characteristic feature of an elliptic curve is that the K-points satisfy a *group law*. We write $E(K)$ for this group of points; if E is in short Weierstrass form, the identity is the point at infinity, $[0:1:0]$ in projective coordinates.

Theorem 1.28 (Mordell–Weil)**.** *Let K be a number field and E/K an elliptic curve. Then $E(K)$ is a finitely generated abelian group,*

$$E(K) \simeq \mathbb{Z}^r + T,$$

for $r \geq 0$ and T a finite abelian group.

The integer r is known as the **rank** of E/K, and T is known as the **torsion subgroup** of $E(K)$.

There are lots of statistical questions we can ask about elliptic curves and their Mordell–Weil groups.

- As $E/\mathbb{Q}$ varies, what groups T arise as torsion subgroups? In the 1970s, Mazur proved a longstanding conjecture that exactly 16 groups T can show up, each one infinitely often. Similar questions have been studied for different base fields.
- As $E/\mathbb{Q}$ varies, what ranks r arise? This is not only an open question, but one that divides our community: some believe that the ranks of $E/\mathbb{Q}$ is bounded, while others believe they are not. The largest known rank is at least 29, for a curve $E/\mathbb{Q}$ discovered in 2024 (last year!) by Elkies and Klagsbrun.
- For a fixed $E/\mathbb{Q}$, how does $E(K)$ vary as K varies? In particular, when is $E(K) \supsetneq E(\mathbb{Q})$? Or, how often does the rank change when going from $E(\mathbb{Q})$ to $E(K)$? These are areas of ongoing research, and can be interpreted as counting number fields $K/\mathbb{Q}$ coming from E in particular ways.

Let's think a little more about ranks. Even though the maximum is not known, we could try to access other statistical statements about them, such as the *average*. As always, we'll need a reasonable way to count elliptic curves.

Definition 1.29 (Naive height)**.** Let $E/\mathbb{Q}$ be an elliptic curve given by

$$E\colon y^2 = x^3 + ax + b.$$

The **naive height** is given by

$$\mathrm{ht}(E) = \max(4|a|^3, 27b^2).$$

This height satisfies some of the counting properties we like, namely that there are finitely many E with $\mathrm{ht}(E) \leq X$. However, notice that it fails to be an isomorphism invariant.

Example 1.30. Let

$$E\colon y^2 = x^3 + x + 1$$
$$E'\colon y^2 = x^3 + 16x + 64.$$

There is an isomorphism $E \to E'$ defined on points by $(x, y) \mapsto (4x, 8y)$. The naive heights of E, E', respectively, are 27 and 110592.

While this is a little inconvenient, we can get around it by recognizing that the problem here is that we have a prime p such that $p^4 \mid a$ and $p^6 \mid b$. As long as we impose that there is no such p, then the short Weierstrass form of $E/\mathbb{Q}$, and hence its naive height, are unique.

Remark 1.31 (Other heights)**.** As always, there are other notions of height. Namely, we could've taken the discriminant of x^3+ax+b or the conductor, an invariant which measures how degenerate the curve becomes when reduced modulo primes p.

There are also theories for counting points on *weighted* spaces, which captures what's going on in the example above. In some sense, these height functions might be considered more natural, since they from the geometry of the space of elliptic curves themselves.

Theorem 1.32 (Bhargava–Shankar [BS15])**.** *We have*

$$\limsup_{X\to\infty} \frac{\sum_{\mathrm{ht}(E)\le X} \mathrm{rk}(E)}{\sum_{\mathrm{ht}(E)\le X} 1} \le 1.5.$$

In other words, the average rank of an elliptic curve over $\mathbb{Q}$, when counted by naive height, is at most 1.5.

This remarkable result both improved upon previous results and removed conditions on the Generalized Riemann Hypothesis and Birch and Swinnerton-Dyer Conjecture. The proof actually involves bounding the average size of the 2-Selmer group of $E/\mathbb{Q}$. To do this, they associate to each element of the 2-Selmer group a binary quartic form (up to some group action), and then count orbits quartic forms. We won't have time to cover this further, but the basic philosophy has a lot in common with our approach to cubic fields.

2 Quadratic fields and squarefree numbers

As something of a warmup, let's carefully prove (1.1),

$$N_2(X) \sim \frac{X}{\zeta(2)}.$$

We will need to borrow some standard facts from analytic number theory, which we will try to keep to a minimum.

2.1 Möbius inversion

We will only need two tools: the Möbius function and Dirichlet convolution.

Definition 2.1 (Möbius function)**.** The **Möbius function**, denoted $\mu(n)$ is defined on $n \in \mathbb{Z}$ by

$$\mu(n) = \begin{cases} (-1)^{\omega(n)} & n \text{ squarefree,} \\ 0 & n \text{ not squarefree,} \end{cases}$$

where $\omega(n)$ is the number of prime factors of n.

Why is this function useful? It tends to come up in inclusion-exclusion arguments involving primes. It also appears in several handy identities that relate arithmetic functions.

The first such formula states that if f, g are arithmetic functions,

$$g(n) = \sum_{d\mid n} f(d) \implies f(n) = \sum_{d\mid n} \mu(d) g\left(\frac{n}{d}\right).$$

Definition 2.2. Let f, g be arithmetic functions. The **Dirichlet convolution**, denoted $f * g$, is an arithmetic function given by

$$(f * g)(n) = \sum_{ab=n} f(a)g(b) = \sum_{d|n} f(d)g(n/d).$$

Analytic number theorists love Dirichlet series. Given this fact, you can see why they also love convolution by observing the following identity[2]

$$\left(\sum_n \frac{f(n)}{n^s}\right)\left(\sum_n \frac{g(n)}{n^s}\right) = \left(\sum_n \frac{(f*g)(n)}{n^s}\right). \tag{2.1}$$

Simply expanding the product on the left reveals the identity.

Let's define two auxiliary arithmetic functions.

$$\mathbb{1}(n) = 1 \text{ for all } n, \text{and}$$
$$\varepsilon(n) = \begin{cases} 1 & n = 1, \\ 0 & \text{otherwise.} \end{cases}$$

We claim that

$$\mathbb{1} * \mu = \varepsilon.$$

This is easy to verify by unwinding the definitions. The key point is that $\sum_{d|n} \mu(d) = 0$ unless $n = 1$.

We'd like to apply (2.1) to this identity. Indeed, the Dirichlet series for $\mathbb{1}$ is the Riemann ζ function! This provides that

$$\sum_n \frac{\mu(n)}{n^s} = \frac{1}{\zeta(s)}. \tag{2.2}$$

This is the bare minimum we need to prove things about squarefree numbers and quadratic fields.

2.2 Squarefree numbers

Let $\mathrm{SF}(X)$ denote the counting function

$$\mathrm{SF}(X) = \#\{1 \le n \le X : n \text{ squarefree}\}.$$

Our immediate goal is to prove the following classical result.

Proposition 2.3. *We have*

$$\mathrm{SF}(X) \sim \frac{X}{\zeta(2)} = \frac{6}{\pi^2} X.$$

Proof. Observe that

$$\mathrm{SF}(X) = \sum_{n \le X} \sum_{d^2|n} \mu(d),$$

[2] Let's not worry about convergence, though in general one probably should. Just think of this as a formal identification for now.

since if n is squarefree, the inner sum is just 1 (coming from $d = 1$), and if not, the inner sum is over all factors of the square part of n, and it vanishes. Switching order of summation and collecting terms, we have

$$\begin{aligned}\sum_{n\le X}\sum_{d^2|n}\mu(d) &= \sum_{d\le\sqrt{X}}\mu(d)\sum_{\substack{n\le X\\ d^2|n}}1\\ &= \sum_{d\le\sqrt{X}}\mu(d)\left\lfloor\frac{X}{d^2}\right\rfloor\\ &= X\sum_{d\le\sqrt{X}}\frac{\mu(d)}{d^2}+O(\sqrt{X}).\end{aligned}$$

We now compare $\sum_{d\le X}\frac{\mu(d)}{d^2}$ to the Dirichlet series:

$$\sum_{d\le\sqrt{X}}\frac{\mu(d)}{d^2}=\sum_{d}\frac{\mu(d)}{d^2}-\sum_{d>\sqrt{X}}\frac{\mu(d)}{d^2}.$$

The first term on the right-hand-side is $\frac{1}{\zeta(2)}$ by (2.2), while the second term is seen to be $O(1/\sqrt{X})$, by estimating $|\mu(d)|\le 1$ and taking the integral. This gives the result. □

Notice that we actually showed $\mathrm{SF}(X) = X/\zeta(2) + O(\sqrt{X})$, which is stronger than the statement! If you're the sort of person that really cares about sharpening error terms, then arithmetic statistics would make a wonderful playground. In this course, we'll be focusing on the main terms for the most part.

We actually will want to know something stronger than Proposition 2.3. How are the squarefree numbers distributed in residue classes modulo n? For this we define

$$\mathrm{SF}(X;a,b) = \#\{1\le n\le X : n \text{ squarefree and } n\equiv a \pmod b\}.$$

Example 2.4 (Squarefrees modulo 4)**.** If we compute the first million integers, we count 607926 are squarefree. Of these, the counts of those which are 1, 2, or 3 modulo 4 are 202636, 202640, and 202650, respectively. These are all pretty close, lending credence to the possibility that $\mathrm{SF}(X;a,4)\sim X/3\zeta(2)$ for $a\in\{1,2,3\}$.

Let's try to prove this for $b = p^2$ by modifying our previous proof of Proposition 2.3.

Proposition 2.5. *Let p be a prime and a such that $a\not\equiv 0 \pmod{p^2}$. Then we have*

$$\mathrm{SF}(X;a,p^2)=\frac{X}{(p^2-1)\zeta(2)}+O(\sqrt{X}).$$

Proof. Following our previous argument, we have

$$\begin{aligned}\mathrm{SF}(X;a,p^2) &= \sum_{\substack{n\le X\\ n\equiv a\,(p^2)}}\sum_{d^2\mid n}\mu(d)\\ &= \sum_{d\le\sqrt{X}}\mu(d)\sum_{\substack{n\le X\\ d^2\mid n\\ n\equiv a\,(p^2)}}1\\ &= \frac{X}{p^2}\sum_{\substack{d\le\sqrt{X}\\ \gcd(d,p)=1}}\frac{\mu(d)}{d^2}+O(\sqrt{X}) \qquad (2.3)\end{aligned}$$

Note the factor of $1/p^2$ that has appeared. We also see that the sum over d is restricted to d coprime to p; since $a \not\equiv 0 \pmod{p^2}$, those terms with $p \mid d$ contribute nothing.

Continuing with our previous approach, we want to compare $\sum_{\substack{d\le\sqrt{X}\\ \gcd(d,p)=1}}\frac{\mu(d)}{d^2}$ to $1/\zeta(2)$. For this we see

$$\begin{aligned}\sum_d\frac{\mu(d)}{d^s} &= \sum_{p\nmid d}\frac{\mu(d)}{d^s}+\sum_{p\mid d}\frac{\mu(d)}{d^s}\\ &= \sum_{p\nmid d}\frac{\mu(d)}{d^s}+\sum_{d'}\frac{\mu(pd')}{pd'^s}\\ &= \sum_{p\nmid d}\frac{\mu(d)}{d^s}+\sum_{p\nmid d'}\frac{\mu(pd')}{pd'^s} && (\text{if } p\mid d' \text{ then } \mu(pd')=0)\\ &= \sum_{p\nmid d}\frac{\mu(d)}{d^s}-\frac{1}{p^s}\sum_{p\nmid d'}\frac{\mu(d')}{d'^s} && (\mu(pd')=\mu(p)\mu(d') \text{ for } p\nmid d')\\ &= \left(1-\frac{1}{p^s}\right)\sum_{p\nmid d}\frac{\mu(d)}{d^s}.\end{aligned}$$

This last line yields, with $s=2$,

$$\sum_{p\nmid d}\frac{\mu(d)}{d^2}=\frac{p^2}{p^2-1}\sum_d\frac{\mu(d)}{d^2}. \qquad (2.4)$$

Combining (2.3) and (2.4), we have

$$\mathrm{SF}(X;a,p^2)=\frac{X}{p^2}\left(\frac{p^2}{p^2-1}\sum_d\frac{\mu(d)}{d^2}-\sum_{\substack{d>\sqrt{X}\\ p\nmid d}}\frac{\mu(d)}{d^2}\right)+O(\sqrt{X}).$$

The sum over $d>\sqrt{X}$ term is estimated the same way as in Proposition 2.3. This leaves the main term of $X/(p^2-1)\zeta(2)$, as desired. □

Proposition 2.5 illustrates an idea that comes back again and again in these sort of counting problems. We often reduce problems we care about to counting (things like)

integers, but we need this count to still behave well when we impose extra *local conditions*. In the case of squarefree numbers, our count does indeed behave well: squarefree numbers are equidistributed among residue classes mod p^2.[3]

2.3 Quadratic fields

We now turn our attention to the $n = 2$ case of Theorem 1.17. For this, we make use of the characterization of quadratic fields. This is well known; see Cox's book [Cox22] for a great reference for all things quadratic fields.

Proposition 2.6 (Characterization of quadratic fields)**.** *There is a bijection*

$$\{\textit{Squarefree numbers } B \in \mathbb{Z}\} \to \{\textit{Quadratic extensions } K/\mathbb{Q}\}/\textit{iso}.$$

Given a squarefree B, its image is (the class of) the extension $K = \mathbb{Q}(\sqrt{B})$, which has:

- $\mathcal{O}_{\mathbb{Q}(\sqrt{B})} \simeq \begin{cases} \mathbb{Z}\left[\frac{1+\sqrt{B}}{2}\right] & B \equiv 1 \pmod 4, \\ \mathbb{Z}[\sqrt{B}] & B \equiv 2, 3 \pmod 4. \end{cases}$
- $\mathrm{Disc}(\mathbb{Q}(\sqrt{B})) = \begin{cases} B & B \equiv 1 \pmod 4, \\ 4B & B \equiv 2, 3 \pmod 4. \end{cases}$

Proof of Theorem 1.17 ($n = 2$ case). We have

$$N_2(X) = 2\left(\mathrm{SF}(X; 1, 4) + \mathrm{SF}(X/4; 2, 4) + \mathrm{SF}(X/4; 3, 4)\right),$$

with the factor of 2 accounting for the fact that $\mathrm{SF}(X)$ only counts positive squarefree integers. By Proposition 2.5, this becomes

$$\frac{2X}{\zeta(2)}\left(\frac{1}{3} + \frac{1}{12} + \frac{1}{12}\right) = \frac{X}{\zeta(2)}.$$

Thus despite some discrepancy in the discriminant, we actually get the same answer as counting squarefree numbers. □

3 Parametrizing rings ranks 2 and 3

A great reference for this material is [Bha09].

Definition 3.1 (rank n ring)**.** A **rank n ring** is a ring R which is isomorphic to $\mathbb{Z}^n$ as a $\mathbb{Z}$-module.

Example 3.2 (Quadratic rings)**.** Let $n = 2$. Here are some examples of quadratic rings:

- If $K/\mathbb{Q}$ is a quadratic field, then $\mathcal{O}_K$ is a quadratic ring. As a $\mathbb{Z}$-module, it is $\mathbb{Z} + \omega\mathbb{Z}$, where ω might be $\sqrt{B}$ or $\frac{1+\sqrt{B}}{2}$.
- The ring $\mathbb{Z} \oplus \mathbb{Z}$ is a quadratic ring. However, it is degenerate in some sense, being the direct product of degree 1 rings.

[3] At least, those for which it's possible to be squarefree!

- Fix $K/\mathbb{Q}$ a quadratic field and let $f > 1$ be a positive integer. Write $\mathcal{O}_K = \mathbb{Z} + \omega\mathbb{Z}$ as above and let $\mathcal{O} = \mathbb{Z} + f\omega\mathbb{Z}$ be the order of conductor f in $\mathcal{O}_K$. This is easily checked to be a ring.
- Quadratic rings can have nilpotents, e.g. $\mathbb{Z}[x]/(x^2)$.

Definition 3.3. A rank n ring is **maximal** if it is not contained in another rank n ring.

3.1 Quadratic rings

Let's set the stage by classifying quadratic rings. Suppose $R = \mathbb{Z} + \omega\mathbb{Z}$. What could its ring structure look like? We understand the addition structure, since R is a free $\mathbb{Z}$-module. For multiplication, we have

$$(a + b\omega)(c + d\omega) = ac + (ad + bc)\omega + bd\omega^2.$$

The ac and $(ad + bc)\omega$ terms make sense. What about the $bd\omega^2$ term? We know $\omega^2 \in R$, so it must be that

$$\omega^2 = \tau + \theta\omega.$$

The values of $\tau, \theta \in \mathbb{Z}$ control the multiplication in R, and hence the ring structure! Another way to put this is that the matrix representing multiplication by ω,

$$\begin{pmatrix} 0 & \tau \\ 1 & \theta \end{pmatrix},$$

controls the ring R.

Remark 3.4. Note that ω, τ, θ are not unique. For example, $\mathbb{Z}[\omega + 1]$ or $\mathbb{Z}[-\omega]$ give the same ring as $\mathbb{Z}[\omega]$. In fact, if $\omega' = \omega + 1$ for instance, then the associated invariants are $\tau' = \tau + 1$, $\theta' = \theta + 2$. In particular, this allows us to always take $\theta \in \{0, 1\}$.

Proposition 3.5. *The set of quadratic rings $\mathbb{Z}[\omega]$, up to isomorphism, is in bijection with the set of integers $\{D \equiv 0, 1 \pmod 4\}$.*

Proof. The map is the discriminant,

$$\mathrm{Disc}(\mathbb{Z}[\omega]) = \det \begin{pmatrix} \mathrm{Tr}(1) & \mathrm{Tr}(\omega) \\ \mathrm{Tr}(\omega) & \mathrm{Tr}(\omega^2) \end{pmatrix}.$$

By definition, $\mathrm{Tr}(\omega) = \theta$, and by linearity we have $\mathrm{Tr}(\omega^2) = \mathrm{Tr}(\tau) + \mathrm{Tr}(\theta\omega) = 2\tau + \theta^2$, so

$$\mathrm{Disc}(\mathbb{Z}[\omega]) = \det \begin{pmatrix} 2 & \theta \\ \theta & 2\tau + \theta^2 \end{pmatrix} = 4\tau + \theta^2.$$

We'll leave showing that the discriminant is an isomorphism class invariant to the exercises. □

A different approach to counting quadratic fields is to instead count quadratic *rings*, then sieve out those which are not integral domains and those which are not maximal. Indeed, this is the approach we will take for cubic fields. Here's what it might look like in the quadratic case.

Lemma 3.6. *Let R_D be the quadratic ring of discriminant D. In light of Proposition 3.5, we can write*

$$D = \begin{cases} 4D' & D \equiv 0 \pmod 4, \\ D' & D \equiv 1 \pmod 4. \end{cases}$$

Then we can infer the following about R_D from D'.

(i) R_D is an integral domain if and only if D' is not an integer square.

(ii) R_D is maximal if and only if D' is squarefree.

Proof. See the exercises. □

To count quadratic fields, we need to count *maximal* quadratic rings R_D which are *integral domains*.

- Proposition 3.5 tells us that there are X quadratic rings R_D with $|\operatorname{Disc}(R_D)| \leq X$.
- Lemma 3.6(i) tells us that $O(\sqrt{X})$ of those R_D are not integral domains.
- We can then sieve out those D' which are not squarefree — by repeating the arguments in the previous section — to obtain the count of maximal quadratic rings.

3.2 Cubic rings

We will parameterize cubic rings in terms of **integral binary cubic forms**,

$$f(x, y) = ax^3 + bx^2y + cxy^2 + dy^3$$

where $a, b, c, d \in \mathbb{Z}$. There is a natural twisted action of $\mathrm{GL}_2(\mathbb{Z})$ on the space of such forms, given by

$$\begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \cdot f(x, y) = \det \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}^{-1} f(\alpha x + \beta y, \gamma x + \delta y).$$

It turns out that cubic rings are parameterized by orbits of cubic forms under this action, in a way that preserves their discriminant.

Proposition 3.7. *There is a bijection*

$$\{\textit{cubic rings } R\}/\textit{iso} \to \{\textit{binary cubic forms } f\}/\operatorname{GL}_2(\mathbb{Z}).$$

If R_f denotes a cubic ring mapping to a cubic form $f(x, y)$, then moreover we have

$$\operatorname{Disc}(R_f) = \operatorname{Disc}(f) = b^2c^2 - 4ac^3 - 4b^3d - 27a^2d^2 + 18abcd.$$

Proof. We follow the exposition in [BST13]. Let $R = \mathbb{Z}[\omega, \theta]$. After possibly translating ω, θ by integers, we have $\omega\theta \in \mathbb{Z}$. Write

$$\begin{aligned} \omega\theta &= n, \\ \omega^2 &= m - b\omega + a\theta, \\ \theta^2 &= \ell - d\omega + c\theta, \end{aligned} \tag{3.1}$$

for integers n, m, ℓ, a, b, c, d. We define the correspondence by

$$R \mapsto f(x,y) = ax^3 + bx^2y + cxy^2 + dy^3.$$

To see this map is well defined, let's introduce a more conceptual perspective for where this map comes from. We have $R/\mathbb{Z} \simeq \mathbb{Z}^2$, with basis $\{\omega, \theta\}$. Define a map

$$\begin{aligned} R/\mathbb{Z} &\xrightarrow{\phi} \wedge^2(R/\mathbb{Z}) \\ r &\mapsto r \wedge r^2. \end{aligned}$$

Note that $\wedge^2 R/\mathbb{Z} \simeq \mathbb{Z}$, generated by the symbol $\omega \wedge \theta = -\theta \wedge \omega$. Using our formulae for ω^2 and θ^2 above, an explicit calculation shows

$$\phi(x\omega + y\theta) = f(x,y)(\omega \wedge \theta).$$

If $\{\omega', \theta'\}$ is another basis for $R/\mathbb{Z}$ (we should think of this as coming from an isomorphic cubic ring R') then in $R/\mathbb{Z}$ we have

$$\begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \begin{pmatrix} \omega \\ \theta \end{pmatrix} = \begin{pmatrix} \omega' \\ \theta' \end{pmatrix} \text{ for } \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \in \mathrm{GL}_2(\mathbb{Z}).$$

Moreover, we see that

$$\phi(x\omega' + y\theta') = \phi((x\alpha + y\beta)\omega + (x\gamma + y\delta)\theta) = f(x\alpha + y\beta, x\gamma + y\delta)(\omega \wedge \theta).$$

The extra factor of $\det \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}$ is added to make this compatible with the natural isomorphism $\wedge^2 R/\mathbb{Z} \to \wedge^2 R/\mathbb{Z}$ coming from the change of basis.

Finally, we construct an inverse to the correspondence. Given a binary cubic form $f(x,y) = ax^3+bx^2y+cxy^2+dy^3$, we can construct a candidate $R = \mathbb{Z}[\omega, \theta]$ using the relations above; we just need to come up with n, m, ℓ values. Since we want R to be associative, we need $\omega^2\theta = \omega^2 \cdot \theta = \omega \cdot \omega\theta$. Flushing this through the equations defining multiplication in R, we find $n = -ad$. A similar argument with $\omega\theta^2$ yields $m = -ac$, $\ell = -bd$. If we took f, f' in the same $\mathrm{GL}_2(\mathbb{Z})$-equivalence class, then we would find using the methods above that $R_f/\mathbb{Z} \simeq R_{f'}/\mathbb{Z}$.

The calculation that $\mathrm{Disc}(R_f) = \mathrm{Disc}(f)$ is left as an exercise. □

With this characterization, we would like to read off information about R_f from $f(x,y)$. Namely, we'd like to be able to tell when R_f is an integral domain, and when it is a maximal cubic ring.

Lemma 3.8 (See [BST13, Prop. 11])**.** *R_f is an integral domain if and only if f is irreducible.*

Proof. If f is reducible, then after a change of coordinates, we may assume $y \mid f$, so $a = 0$. Constructing R_f as in the proof of Proposition 3.7, we have $\omega\theta = n = -ad = 0$, so R_f is not an integral domain.

Conversely, suppose R has zero divisors, $\alpha\beta = 0$ for $\alpha, \beta \neq 0$. Multiplying the characteristic cubic equation for α by β, we have

$$\beta(\alpha^3 + c_1\alpha^2 + c_2\alpha + c_3) = 0 \implies c_3 = 0.$$

Thus $\alpha(\alpha^2 + c_1\alpha + c_2) = 0$. If $\alpha^2 + c_1\alpha + c_2 = 0$, then set $\omega = \alpha$. Otherwise, set $\omega = \alpha^2 + c_1\alpha + c_2$ and observe $\omega^2 = c_2\omega$. In both cases, we can extend $\{1, \omega\}$ to a basis $\{1, \omega, \theta\}$ for R. From our computation of ω^2, we see that the associated form $f(x,y)$ has $a = 0$, and so it is reducible. □

For maximality, we recognize that a cubic ring R is maximal if and only if it is either

- $R = \mathcal{O}_K$ for a cubic number field K,
- $R = \mathcal{O}_K \times \mathbb{Z}$ for a quadratic number field K, or
- $R = \mathbb{Z}^3$.

That is, R is maximal if and only if it is the product of rings of integers of number fields.

Definition 3.9. Let p be a prime. A cubic ring R is **maximal at p** if $R_p = R \otimes \mathbb{Z}_p$ is maximal as a cubic $\mathbb{Z}_p$-algebra.

By the characterization above, R is maximal if and only if it is maximal at p for all primes p. Later, we will need to know how to see whether R_f is maximal at p by looking at f.

Lemma 3.10 (See [BST13, Lemma 13])**.** *Let R be a cubic ring. R is not maximal at p if there is a basis $\{1, \omega, \theta\}$ such that one of the following holds:*

- *$\mathbb{Z}[\omega/p, \theta]$ is a ring;*
- *$\mathbb{Z}[\omega/p, \theta/p]$ is a ring.*

Proof. Suppose $R \subset R'$ for a cubic ring R' and $p \mid [R' : R]$. After possibly replacing R' by $R' \cap (R \otimes_{\mathbb{Z}} \mathbb{Z}[1/p])$, we may assume that the index $[R' : R]$ is a (nonzero) power of p. This means that as abelian groups, we have $(R'/\mathbb{Z})/(R/\mathbb{Z}) \simeq \mathbb{Z}/p^i\mathbb{Z} \times \mathbb{Z}/p^j\mathbb{Z}$. Equivalently, $R' = \mathbb{Z}[\omega/p^i, \theta/p^j]$ for $R = \mathbb{Z}[\omega, \theta]$. We may also safely assume that $\{1, \omega/p^i, \theta/p^j\}$ is a normalized basis, i.e. $\omega\theta/p^{i+j} \in \mathbb{Z}$.

After possibly swapping basis elements, we may assume $i \geq j$ (and not both zero). If $i = 1$, then we are done. If both $i, j \geq 1$, then we argue that $\mathbb{Z}[\omega/p^{i-1}, \theta/p^{j-1}]$ is a ring. To see this, it suffices to check it contains $(\omega/p^{i-1})^2$ and $(\theta/p^{j-1})^2$. But

$$\left(\frac{\omega}{p^{i-1}}\right)^2 = p^2\left(\frac{\omega}{p^i}\right) = p^2 m - pb\frac{\omega}{p^{i-1}} + pa\frac{\theta}{p^{j-1}} \in \mathbb{Z}[\omega/p^{i-1}, \theta/p^{j-1}],$$

where m, a, b are integers coming from the multiplication structure of R'. A similar calculation shows $(\theta/p^{j-1})^2$ is in the ring. The same argument in the $j = 0$ case shows that $\mathbb{Z}[\omega/p^{i-1}, \theta]$ is a ring.

Iterating this process as necessary, we eventually produce R' of one of the forms desired in the statement of the lemma. □

What does this tell us about $f(x, y) = ax^3 + bx^2y + cxy^2 + dy^3$?

- If $R' = \mathbb{Z}[\omega/p, \theta]$ is a ring, with multiplication given by (3.1) with constants $n', m', \ell', a', b', c', d'$, then for R we have

$$\omega^2 = p^2(\omega/p)^2 = p^2m' - pb'\omega + p^2a'\theta,$$
$$\theta^2 = \ell - d\omega + c\theta.$$

  That is, we have $p^2 \mid a$ and $p \mid b$.

- If $R' = \mathbb{Z}[\omega/p, \theta/p]$ then a similar argument shows

$$\omega^2 = p^2(\omega/p)^2 = p^2m' - pb'\omega + pa'\theta,$$
$$\theta^2 = p^2(\theta/p)^2 = p^2\ell' - pd'\omega + pc'\theta.$$

  In this case, $p \mid a, b, c, d$.

References

[Ach] Niven Achenjang. Counting cubic number fields. Available at https://www.mit.edu/~NivenT/assets/pdf/Counting_Cubic_Number_Fields.pdf.

[Bha05] Manjul Bhargava. The density of discriminants of quartic rings and fields. *Ann. of Math. (2)*, 162(2):1031–1063, 2005.

[Bha09] Manjul Bhargava. On the classification of rings of "small" rank. Arizona Winter School notes, https://swc-math.github.io/aws/2009/09BhargavaNotes.pdf, 2009.

[Bha10] Manjul Bhargava. The density of discriminants of quintic rings and fields. *Ann. of Math. (2)*, 172(3):1559–1591, 2010.

[BS15] Manjul Bhargava and Arul Shankar. Binary quartic forms having bounded invariants, and the boundedness of the average rank of elliptic curves. *Ann. of Math. (2)*, 181(1):191–242, 2015.

[BST13] Manjul Bhargava, Arul Shankar, and Jacob Tsimerman. On the Davenport-Heilbronn theorems and second order terms. *Invent. Math.*, 193(2):439–499, 2013.

[Cox22] David Cox. *Primes of the Form x^2+ny^2: Fermat, Class Field Theory, and Complex Multiplication*, volume 387 of *American Mathematical Society Chelsea Publishing*. American Mathematical Society, Providence, RI, 2022.

[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.

[GM04] Andrew Granville and Greg Martin. Prime number races. Available at https://arxiv.org/abs/math/0408319, 2004.

[Klü05] Jürgen Klüners. A counterexample to Malle's conjecture on the asymptotics of discriminants. *C. R. Math. Acad. Sci. Paris*, 340(6):411–414, 2005.

[Lan] Aaron Landesman. Notes on counting extensions of degrees 2 and 3, following Bhargava. Available at https://people.math.harvard.edu/~landesman/assets/bhargavology-seminar-notes.pdf.

[LS24] Daniel Loughran and Tim Santens. Malle's conjecture and brauer groups of stacks, 2024.

[Neu99] Jürgen Neukirch. *Algebraic number theory*, volume 322 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 1999. Translated from the 1992 German original and with a note by Norbert Schappacher, With a foreword by G. Harder.

[Woo14] Melanie Wood. Asymptotics for number fields and class groups. Arizona Winter School notes, https://swc-math.github.io/aws/2014/2014WoodNotes.pdf, 2014.