An Introduction to

Algebraic Number Theory

Alistair Pattison

Carleton College Math Department

November 2023

Contents

1 Introduction

Imagine that you want to find all integer solutions of the equation

$$xy - x + 3\,y - z - 3 = 0. \tag{1}$$

This is an example of a Diophantine equation, and a classic way to solve them is via factoring. By rewriting Equation 1 as

$$\begin{aligned} z &= xy - x + 3\,y - 3 \\ &= (x+3)\,(y-1) \end{aligned} \tag{2}$$

and enumerating all factors of $z = ab$, one can recover $x = a - 3$ and $y = b + 1$. The first few solutions are

$$\begin{aligned} z &= 0, \quad x = -3, \quad y = 1; \\ z &= 1, \quad x = -2, \quad y = 2; \\ z &= 1, \quad x = -4, \quad y = 0. \end{aligned} \tag{3}$$

1.1 Primitive Pythagorean Triples

This technique is great when it applies, but it often doesn't work for more complicated Diophantine equation. For example, imagine we want to find all primitive Pythagorean triples, i.e., solutions to

$$z^2 = x^2 + y^2 \tag{4}$$

with $x, y, z \in \mathbb{Z}$ all relatively prime. The polynomial $x^2 + y^2$ is irreducible over $\mathbb{Z}[x]$, so we can't use the same factoring approach from the previous problem.

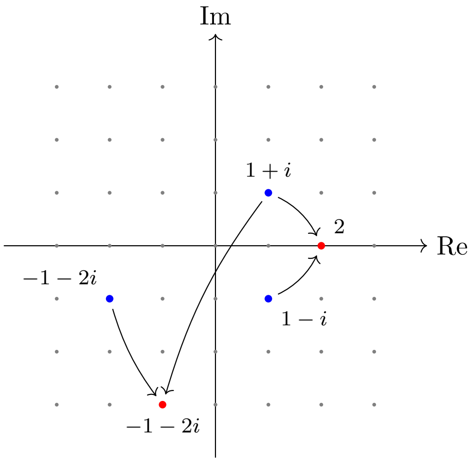

Figure 1: The prime factorizations of 2 and $-1-2i$ in the lattice of Gaussian integers $\mathbb{Z}[i]$.

a	b		
2	1	$3^2 + 4^2 = 5^2$	
3	2	$5^2 + 12^2 = 13^2$	
3	1	$8^2 + 6^2 = 10^2$	(not primitive)
4	3	$7^2 + 24^2 = 25^2$	
4	2	$12^2 + 16^2 = 20^2$	(not primitive)
4	1	$15^2 + 8^2 = 17^2$	
5	4	$9^2 + 40^2 = 41^2$	
5	3	$16^2 + 30^2 = 34^2$	(not primitive)

Table 1: The first few Pythagorean triples, primitive when the parity of a and b are different.

An unintuitive but fruitful idea is to take a leap of faith and reframe this question about the (traditional) integers as a factoring problem over the Gaussian integers $\mathbb{Z}[i]$ (Figure 1). We start by noting that Equation 4 factors as

$$\begin{aligned} z^2 &= x^2 + y^2 \\ &= \underbrace{(x + iy)}_{\alpha}\underbrace{(x - iy)}_{\beta} \qquad \in \mathbb{Z}[i][x], \end{aligned} \tag{5}$$

and one can show that $\alpha, \beta \in \mathbb{Z}[i]$ are relatively prime because $x, y \in \mathbb{Z}$ are relatively prime by assumption. Because $\mathbb{Z}[i]$ has unique factorization, α must be a square, i.e., $\alpha = u\gamma^2$ where $\gamma \in \mathbb{Z}[i]$ and $u \in \mathbb{Z}[i]^\times = \{\pm 1, \pm i\}$. (This is analogous to how in the traditional integers, $rs = n^2$ with $(r, s) = 1$ implies that $r = \pm k^2$ for some k. In $\mathbb{Z}[i]$, the units $\{\pm 1, \pm i\}$ take the place of the optional minus sign in front of k.)

For the sake of brevity, we ignore the pesky units and consider only the case where $u = 1$, i.e.,

$$\begin{aligned} \alpha &= \gamma^2 \\ &= (a + bi)^2 \\ &= (a^2 - b^2) + 2ab\,i \end{aligned} \tag{6}$$

for some $a, b \in \mathbb{Z}$. Referring back to our original definition of $\alpha = x + iy$, we have that

$$\begin{aligned} x &= \mathrm{Re}(\alpha) = a^2 - b^2, \\ y &= \mathrm{Im}(\alpha) = 2ab, \end{aligned} \tag{7}$$

and a bit of algebra gets us

$$z = a^2 + b^2. \tag{8}$$

Whenever a and b are relatively prime, the generated triple will be primitive. We enumerate the first few positive solutions (given by $a > b > 0$) in Table 1.

1.2 Overview

This is admittedly a toy example, but it illustrates an important point. We started with a problem strictly in the integers, and only by reframing it as a problem about a larger ring $\mathbb{Z}[i] \subset \mathbb{C}$ were we able to solve it. This is the essence of algebraic number theory: using tools from algebra like rings and field extensions to deepen our understanding about the integers and the primes.

While this is a nice, romantic vision of the field, it would be nice to have something more concrete. Thankfully, Daniel Marcus provides a much more pragmatic definition in the very first sentence of his book [3, p. 1]:

> "Algebraic number theory is essentially the study of number fields."

The goal of this paper is in its title: to provide an introduction to algebraic number theory. We'll start with a crash course in algebra, including some things omitted or very briefly covered in Carleton's intro algebra class. Then, we'll talk about algebraic number fields (the things Marcus speaks so highly of) and their corresponding integer rings before defining and unpacking the class group and the class number. To close, we'll talk a bit about the current state of the field and place this paper's results into the larger context of math. Unique factorization (and its failure) will provide a thread through the whole piece.

My hope is that anyone who has taken Math 342 at Carleton or the equivalent could conceivably pick up this paper and read it from top to bottom with no consultation of outside material. That's not to say that it'll be a light read–at times we'll move very quickly and leave gaps in the exposition for the sake of brevity. To quote Richard Feynman [2, p. 148]:

> "I am going to give what I will call an elementary demonstration. But elementary does not mean easy to understand. Elementary means that very little is required to know ahead of time in order to understand it, except to have an infinite amount of intelligence."

2 Preliminaries

Algebraic number theory (unsurprisingly) requires a lot of algebra—most texts require a graduate course or two, but I've attempted to condense the necessary background into the following few pages. See e.g. Dummit and Foote [1] for more complete exposition.

2.1 Groups

We assume that the reader is familiar with the definition of a group and provide it here only for completeness.

Definition 2.1 (Group)**.** *A **group** is a set G along with a binary operation $\cdot : G \times G \to G$ such that,* *(i)* $a \cdot b \in G$ *for all* $a, b \in G$ *(closure),*

(ii) *there exists an* identity element e *such that* $a \cdot e = a$ *for all* $a \in G$ *(identity),*

(iii) *for every* $a \in G$*, there exists some element* a^{-1} *such that* $a \cdot a^{-1} = e$ *(inverses).*

If the $\cdot$ *operation is commutative, we call* G *an* ***Abelian group****.*

Definition 2.2 (Subgroup)**.** *We say that* $H \subset G$ *is a* subgroup of G *(written* $H \leq G$*) if* H *is a group with respect to the group operation of* G*.*

2.2 Rings

Much of algebraic number theory is concerned with generalizations of integers in the complex plane. This is done through the theory of rings.

Definition 2.3 (Commutative ring)**.** *A* ***commutative ring*** *is a set* R *with two commutative operations* $+$ *and* $\cdot$ *such that*

(i) *R is an Abelian group with respect to addition,*

(ii) *R is closed under multiplication, and*

(iii) *multiplication distributes over addition.*

Definition 2.4 (Ideal)**.** *An (additive) subgroup* $I \leq R$ *is an* ***ideal*** *if* $ra \in I$ *for all* $r \in R$*,* $a \in I$*.*

We define $A = (a_1, a_2, \ldots)$ *to be the smallest ideal containing every* a_i *and say that* A *is* ***generated*** *by* $\{a_1, a_2, \ldots\}$*.*

We define multiplication of ideals as follows: If A and B are ideals, then their product is the ideal generated by the set

$$\{ab : a \in A, b \in B\}. \tag{9}$$

For principal ideals, $(a)(b) = (ab)$, and the product of two non-principal, finitely generated ideals is

$$(a_1, \ldots, a_n)(b_1, \ldots, b_n) = (a_i b_j : i \leq n, j \leq m). \tag{10}$$

We leave it to the reader to show that if A and B are ideals, that AB is also an ideal and that $AB \subset A \cap B$.

Much of number theory concerns prime numbers, so we extend the notion of primality to ideals in the following way:

Definition 2.5 (Prime Ideal)**.** *An ideal* P *is* ***prime*** *if* $ab \in P$ *implies* $a \in P$ *or* $b \in P$*.*

This is a generalization of Euclid's Lemma which states that p is prime if and only if $p \mid ab$ implies $p \mid a$ or $p \mid b$. In the integers, prime ideals are those ideals generated by prime elements.

Definition 2.6. *A commutative ring is an **integral domain** if $ab = 0$ implies $a = 0$ or $b = 0$. (No zero divisors.)*

Definition 2.7 (Fraction field)**.** *The **fraction field** of an integral domain R is the smallest field containing R. It is equal to*

$$\operatorname{Frac} R = \left\{ \frac{a}{b} : a, b \in R, b \neq 0 \right\} / \sim \tag{11}$$

where $\sim$ is the equivalence relation $^a/_b = {}^p/_q$ if $aq = bp$. (This is just your familiar cross-multiplication.)

The classic example is that $\operatorname{Frac} \mathbb{Z} = \mathbb{Q}$; in other rings, the definition behaves very intuitively. One can almost always forget about the equivalence relation and just cancel like terms in the numerator and denominator. For example, the fraction field of the polynomial ring $\mathbb{C}[x]$ is the field of rational functions $\mathbb{C}(x)$.

Definition 2.8 (Integral elements)**.** *Let A be a ring a subring B. An element $a \in A$ is **integral over** B if a is the root of some monic polynomial $f \in B[x]$. If $A \geq B = \mathbb{Z}$, we often drop the "over B" part and say that f is **integral**.*

Any integer a is trivially integral because of the polynomial $f_a(x) = x - a$, but there are more complicated examples too. For example, $\frac{1}{2} + \frac{1}{2}\sqrt{5} \in \mathbb{Z}[\sqrt{5}]$ is integral by the polynomial $x^2 - x - 1$.

Definition 2.9 (Integrally Closed Domain)**.** *Let R be an integral domain. We say that R is **integrally closed** if the fact that $a \in \operatorname{Frac} R$ is integral over R implies that $a \in R$.*

Equivalently, R is integrally closed if there are no elements of $\operatorname{Frac} R \setminus R$ that are integral over R.

The integers are an integrally closed domain, which translates to the statement that monic polynomials with integer coefficients don't have fractional roots. This notion of integral closure will become important later in the paper in the context of Dedekind domains.

Definition 2.10 (Unique factorization domain)**.** *A commutative ring R is a **unique factorization domain** if every element factors uniquely into irreducible elements.*

The integers are famously a UFD by the fundamental theorem of arithmetic which states that every integer factors uniquely into a product of primes. But this isn't always the case, for example, in the ring $\mathbb{Z}[i\sqrt{5}]$,

$$\begin{aligned} 6 &= 2 \cdot 3 \\ &= (1 + i\sqrt{5})(1 - i\sqrt{5}) \end{aligned} \tag{12}$$

where 2, 3, $1 + i\sqrt{5}$ and $1 - \sqrt{5}$ are all irreducible (meaning their only divisors are themselves and units).

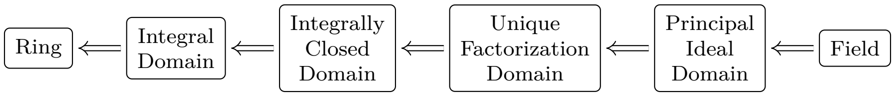

Figure 2: The chain of successively stronger ring definitions. Each definition in the chain implies the previous.

Definition 2.11 (Principal Ideal Domain)**.** *A commutative ring R is a* ***principal ideal domain*** *if every ideal is generated by a single element.*

The integers are principal because every ideal is of the form $(a) = \{1, \pm 1, \pm 2a, \ldots\}$. The rings $\mathbb{Z}[x]$ and $\mathbb{Q}[x, y]$ are not because of the ideals $(2, x)$ and (x, y) respectively.

Definition 2.12 (Noetherian domain)**.** *A ring R is* ***Noetherian*** *if every ideal is finitely generated.*

We provide (without proof) the following equivalent definitions of Noetherian:

Theorem 2.13. *The following are equivalent*

(i) R is Noetherian

(ii) Every increasing sequence of ideals is eventually constant, i.e., $I_1 \subset I_2 \subset \cdots$ implies that there exists an M such that $I_n = I_m$ for $n, m > M$.

(iii) Every non-empty set of ideals S has a "maximal" element M such that $M \subset I$ implies $M = I$. There may be multiple such maximal elements.

2.3 Fields

Definition 2.14 (Field)**.** *A* ***field*** *is a commutative ring F where $0 \neq 1$ and every element $a \in F$ has a multiplicative inverse $a^{-1} \in F$ such that $aa^{-1} = e$.*

Fields are often touted as the end of the chain of implications shown in Figure 2. But there's always a bigger fish. For example, the insufficiency of the field $\mathbb{Q}$ becomes very apparent when we consider the polynomial $x^2 + 5$. It wants to factor as

$$\begin{aligned} f(x) &= x^2 + 5 \\ &= (x + i\sqrt{5})(x - i\sqrt{5}), \end{aligned} \tag{13}$$

but it can't because $i\sqrt{5} \notin \mathbb{Q}$. Over $\mathbb{C}$, f does factor completely, but there's a lot of extra stuff in $\mathbb{C}$ that f doesn't care about:

$$\pi, \quad e, \quad \sqrt{17}, \quad 4 + 3\sqrt[6]{5}, \quad \text{and} \quad e^{2\pi i/5} \tag{14}$$

to name a few. In some sense, f would be "equally happy" living in a much smaller field without these extraneous elements. This motivates the following definition.

Definition 2.15 (Finite field extensions)**.** *The* field extension $K(\alpha_1, \ldots, \alpha_n)$ *is the smallest field containing both K and each of $\alpha_1, \ldots, \alpha_n$.*

For f, the field of concern is

$$\mathbb{Q}(i\sqrt{5}) = \{a + bi\sqrt{5} : a, b \in \mathbb{Q}\}. \tag{15}$$

This is a number field.

3 Number Fields

We begin by defining the object of interest:

Definition 3.1 (Number Field)**.** *A* ***number field*** *$K \subset \mathbb{C}$ is a finite extension of $\mathbb{Q}$.*

For those comfortable with fields (e.g. through a course on Galois Theory), this definition seems completely natural. But it's often helpful to instead think about number fields $\mathbb{Q}(\alpha)$ as finite-degree vector spaces over $\mathbb{Q}$. For example, the number field introduced at the end of the previous section can be rewritten as

$$\mathbb{Q}(i\sqrt{5}) = \operatorname{span}_{\mathbb{Q}}\{1,\ i\sqrt{5}\}, \tag{16}$$

a degree-two rational vector space with basis $\{1,\ i\sqrt{5}\}$. We claim (without proof) that we can do the same thing for any number field.

Theorem 3.2. *Any number field can be written in the form*

$$K = \mathbb{Q}[\alpha] = \operatorname{span}_{\mathbb{Q}}\{1,\ \alpha,\ \alpha^2,\ \ldots,\ \alpha^{n-1}\}$$

where $\alpha \in \mathbb{C}$ is the root of some degree-n polynomial in $\mathbb{Q}[x]$. We call n the ***degree*** *of K.*

Numbers $\alpha \in \mathbb{C}$ that are roots of polynomials over $\mathbb{Q}$ are so special that mathematicians came up with a name for them:

Definition 3.3 (Algebraic number)**.** *A complex number $\alpha \in \mathbb{C}$ is an* ***algebraic number*** *if it is the root of some irreducible monic polynomial $f \in \mathbb{Q}[x]$. We call f the* ***minimal polynomial*** *of α and say that α has* ***degree*** *$\deg \alpha = \deg f$.*

If α and β are algebraic numbers with the same minimal polynomial, we say that they are *conjugate*—for example i and $-i$ are both roots of the polynomial $x^2 + 1$. In general, if $\alpha \in \mathbb{C}$ has degree $\deg \alpha = n$, then α has $n - 1$ conjugates by the fact that $\mathbb{C}$ is algebraically closed.

Most numbers we're familiar with are algebraic, for example any combination of radicals and rational numbers. Some famous non-examples are transcendental numbers like π and e.

Although any algebraic number α produces a valid number field $\mathbb{Q}(\alpha)$, for the purposes of this paper we'll focus on two "classic" examples: cyclotomic and quadratic fields.

Definition 3.4 (Cyclotomic field)**.** *The kth **cyclotomic field** is the number field $\mathbb{Q}(\zeta_k)$ where $\zeta_k = e^{2\pi i/k}$ is a kth root of unity.*

Definition 3.5 (Quadratic field)**.** *The mth **quadratic field** is the number field $K = \mathbb{Q}(\sqrt{m})$. When $m > 0$, we call K a **real quadratic field**. When $m < 0$, we call K an **imaginary quadratic field**.*

In quadratic fields, we typically assume that m is squarefree. If not, $m = nk^2$ for some $n, k \in \mathbb{Z}$ and

$$\mathbb{Q}(\sqrt{m}) = \mathbb{Q}(k\sqrt{n}) = \mathbb{Q}(n). \tag{17}$$

One might expect that $m > 0$ produces be the simpler case, but it turns out that the opposite is true: we know very little about real quadratic fields (more on this later).

3.1 Complex Embeddings

Definition 3.6 (Embedding)**.** *Let K be a number field. An **embedding** of K in $\mathbb{C}$ is a ring homomorphism $\sigma : K \hookrightarrow \mathbb{C}$.*

Embeddings are always injective. (This follows from the fact that the kernel of a homomorphism must be an ideal and the only ideals of a field are the zero ideal and the field itself).

Algebraic number fields $K = \mathbb{Q}(\alpha)$ are subsets of $\mathbb{C}$, so there is a trivial embedding given by $x \mapsto x$. But this isn't the only way to map number fields onto the complex plane. For quadratic fields $Q(\sqrt{m})$, the map

$$a + b\sqrt{m} \;\mapsto\; a - b\sqrt{m} \tag{18}$$

is an embedding (exercise: verify this) because $a - b\sqrt{m}$ and $a + \sqrt{m}$ are conjugate, i.e., roots of the same minimal polynomial in $\mathbb{Q}[x]$. In general for an algebraic number field $K = \mathbb{Q}(\alpha)$, there are $n = \deg \alpha$ embeddings: one trivial identity map and $n - 1$ others given by sending α to any of its conjugates.

For cyclotomic fields, this amounts to sending ζ_k to any of the other kth primitive roots of unity, of which there are $\varphi(k)$. Figure 3 shows the possible destinations of ζ for $k = 5$.

Before we move on, we define the norm which—while a useful object in its own right—we introduce mostly for its utility later on in proving Theorem 5.4.

Definition 3.7 (Norm)**.** *Let $K = \mathbb{Q}(\alpha)$ be a number field and let $\sigma_1, \ldots, \sigma_n : K \to \mathbb{C}$ denote the complex embeddings of K. The **norm** of $\alpha \in K$ is*

$$\mathrm{N}^K(\alpha) = \sigma_1(\alpha)\,\sigma_2(\alpha)\;\cdots\;\sigma_n(\alpha). \tag{19}$$

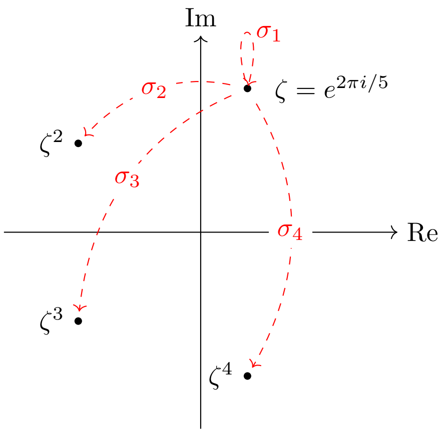

Figure 3: Where $\zeta = e^{2\pi i/5}$ gets sent by the four embeddings of the fifth cyclotomic field $\mathbb{Q}(\zeta)$.

4 Number Rings

In the previous section, we extend the notion of rational number into the complex plane. Now, we develop a complex analogue to the integers.

Definition 4.1 (Algebraic Integer)**.** *An **algebraic integer** is a complex number α that is the root of some monic polynomial $f \in \mathbb{Z}[x]$. We use $\mathbb{A}$ to denote the set of all algebraic integers.*

Using the idea of integrality from Definition 2.8, one can equivalently define

$$\mathbb{A} = \{\alpha \in \mathbb{C} : \alpha \textit{ is integral}\}. \tag{20}$$

As one would hope, our familiar integers $\mathbb{Z}$ remain integers under this more general definition: $a \in \mathbb{Z}$ is the root of $f_a(x) = x - a \in \mathbb{Z}[x]$. But $\mathbb{A}$ also contains more complicated elements that don't at all obey our intuition of integrality:

- $i\sqrt{5} \in \mathbb{A}$ because of $f(x) = x^2 + 5$.
- $2 + \sqrt[3]{17} \in \mathbb{A}$ because of $f(x) = x^3 - 6x^2 + 12x - 25$.
- $\frac{1}{2} + \frac{1}{2}\sqrt{5} \in \mathbb{A}$ because of $f(x) = x^2 - x - 1$.

Not all algebraic numbers are algebraic integers, but non-integral elements tend to look a little bit more complicated. For example, $-\frac{3}{4} + \frac{i}{4}\sqrt{7} \notin \mathbb{A}$ because its minimal polynomial $x^2 + \frac{3}{2}x + 1$ has a non-integer coefficient.

It turns out (although we won't prove it) that $\mathbb{A}$ is a subring of $\mathbb{C}$, i.e., closed under addition and subtraction. Given $\alpha, \beta \in \mathbb{A}$, the procedure for generating minimal polynomials for $\alpha + \beta$ and $\alpha\beta$ is given in Marcus [3, p. 12].

With this new understanding of integrality in $\mathbb{C}$, we propose the following natural definition for a number field's canonical "ring of integers".

Definition 4.2 (Number Ring)**.** *The **number ring** of a number field $K = \mathbb{Q}(\alpha)$ is the set of algebraic integers contained within K, denoted*

$$\mathcal{O}_K = \mathbb{A} \cap K.$$

We use $\mathcal{O}$ because it looks like a ring.

Sometimes, calculating the number ring for a number field is a simple as moving α from next to the $\mathbb{Q}$ to next to the $\mathbb{Z}$. For example, our friend $\mathbb{Z}[i\sqrt{5}]$ is the number ring of the imaginary quadratic field $\mathbb{Q}(i\sqrt{5})$. This happens to also be true for all cyclotomic fields:

$$\mathcal{O}_{\mathbb{Q}(\zeta_k)} = \mathbb{Z}[\zeta_k]. \tag{21}$$

Proving this fact is not trivial [3, p. 22].

Unfortunately, it's not always this simple. In fact, we've already seen evidence of a more complicated example when we noted above that $\frac{1}{2} + \frac{1}{2}\sqrt{5} \in \mathbb{Q}(\sqrt{5})$ is an integer. But it's not in $\mathbb{Z}[\sqrt{5}]$. In general, the number ring of a quadratic field $\mathbb{Q}(\sqrt{m})$ is

$$\mathcal{O}_{\mathbb{Q}(\sqrt{m})} = \begin{cases} \mathbb{Z}[1, \frac{1}{2} + \frac{1}{2}\sqrt{d}\,] & \text{if } d \equiv 1 \mod 4 \\ \mathbb{Z}[1, \sqrt{d}\,] & \text{otherwise.} \end{cases} \tag{22}$$

Thankfully, things are never too bad: we are always guaranteed an integral basis for $\mathcal{O}_K$, meaning that the ring of integers for any number field K can be written as

$$\mathcal{O}_K = \operatorname{span}_{\mathbb{Z}}\{b_1, \dots, b_n\} \tag{23}$$

for some $b_1, \dots, b_n \in \mathcal{O}_K$. Another way of saying this is that $\mathcal{O}_K$ is always a free Abelian group of rank n [3, p. 20].

4.1 Dedekind Domains

Number rings have some nice properties, so mathematicians have done what they love best and given those properties a definition.

Definition 4.3 (Dedekind domain)**.** *A **Dedekind domain** is a ring R such that*

(i) R is integrally closed (Definition 2.9),

(ii) R is Noetherian (Definition 2.12), and

(iii) every nonzero prime ideal is maximal.

On its own, this definition is uninspiring, but over the next several pages it will allow us to establish the following: Even though elements of number rings don't factor uniquely, the *ideals* of number rings do decompose uniquely into primes. But before we do anything too impressive, we show that number rings satisfy this definition.

Theorem 4.4. *Number Rings are Dedekind domains.*

Proof. We will show part (i) of Definition 4.3. Sketches of proofs for (ii) and (iii) are available in [3, p. 40].

Let $K = \mathbb{Q}[\alpha]$ be a number field. We wish to show that $\mathcal{O}_K$ is integrally closed, i.e., that if $\alpha \in K$ is the root of some monic $f \in \mathcal{O}_K[x]$, then $\alpha \in \mathcal{O}_K$.

To start, let $f(x) = x^n + \alpha_{n-1}x^{n-1} + \cdots + \alpha_0$ and note that because α_i is an algebraic integer, it's the root of some monic $f_i \in \mathbb{Z}[x]$ of degree d_i. Using this fact, we are able to write $\alpha_i^{d_i}$ as an integral linear combination of lower-order powers of α_i:

$$\alpha_i^{d_i} = c_0 + c_1\alpha_i + \ldots + c_{d_i-1}\alpha_i^{d_i-1}, \quad c_i \in \mathbb{Z}. \tag{24}$$

By repeating this rewriting process, we can express *any* power of α_i as an integral linear combination of $1, \alpha_i, \alpha_i^2, \ldots, \alpha_i^{d_i-1}$.

We now show that $R = \mathbb{Z}[\alpha_0, \ldots, \alpha_{n-1}, \alpha]$ has a finite integral basis. We know that elements of R are linear combinations of terms of the form

$$x = \alpha_0^{b_0} \cdots \alpha_{n-1}^{b_{n-1}}\, \alpha^{b_n} \tag{25}$$

where $b_i \in \mathbb{Z}$. First, we rewrite α^{b_n} in terms of $\alpha, \ldots, \alpha_{n-1}$ and then use the result from the previous paragraph to rewrite high powers of $\alpha_0, \ldots, \alpha_{n-1}$. We are left with only low power terms—specifically an integral linear combination of elements in the set

$$B = \left\{\alpha_0^{k_0} \cdots \alpha_{n-1}^{k_{n-1}}\, \alpha^k : 0 \leq k_i < d_i, 0 \leq k < n\right\} \tag{26}$$

So B additively generates R. The generating set B set has size $d_0 \cdots d_{n-1} \cdot n$, which is finite. Because $\alpha \in R$ and R is finitely additively generated, it follows that $\alpha \in A$ by [3, Theorem 2.2]. We conclude that $\alpha \in \mathbb{A} \cap K = \mathcal{O}_K$, and therefore that $\mathcal{O}_K$ is integrally closed. □

Our ultimate goal for this section is to show that ideals in Dedekind domains factor uniquely into primes, but to get there, we need a much better sense for how the ideals themselves behave. The following theorems begin to build that understanding. For the sake of brevity, we defer proof of Theorem 4.5 to Marcus [3, p. 40] (the proof is long and not illuminating). However, we will use the result to prove two desirable properties of Dedekind ideals that we make use of in the remainder of the section.

Theorem 4.5 (Inverses exist)**.** *For any nonzero ideal $I \subseteq R$, there exists some ideal $J \subseteq R$ such that IJ is principal.*

***Corollary* 4.5.1** *(Cancellation law).* If A, B, and C are ideals in a Dedekind domain with $AB = AC$, then $B = C$.

Proof. Take $J \subset R$ such that JA is principal, i.e., $JA = (\alpha)$ for some $\alpha \in R$. Then

$$\alpha B = (\alpha)B = JAB = JAC = (\alpha)B = \alpha C, \tag{27}$$

so $B = C$. □

***Corollary* 4.5.2** *(Ideal divisibility).* If A and B are ideals of a Dedekind domain, then $A \mid B$ iff $A \supset B$.

Proof. The forward direction is easy: If $A \mid B$, there must exist some nontrivial ideal I such that $AI = B$, from which $B \subset A$ follows immediately.

To prove the reverse direction, assume $B \subset A$ and choose J such that $AJ = (\alpha)$ is principal. Define $C = \frac{1}{\alpha}JB$ and observe that

$$C = \frac{1}{\alpha} = \frac{1}{\alpha}JB \subset \frac{1}{\alpha}JA = \frac{1}{\alpha}(\alpha) \subseteq R. \tag{28}$$

Additionally, for any $x \in R$ and $c \in C$, we have

$$xr = x(\frac{1}{\alpha}jb) = \frac{1}{\alpha}(xj)b \in \frac{1}{\alpha}JB \tag{29}$$

by the fact that J is an ideal and $xj \in J$. So C is an ideal and

$$AC = \frac{1}{\alpha}AJB = \frac{1}{\alpha}(\alpha)B = RB = B. \tag{30}$$

by the fact that $RB = B$ for any ideal B. □

With these in hand, we move to the main result of this section.

Theorem 4.6. *Ideals in Dedekind domains uniquely factor into prime ideals.*

Proof. Let R be a Dedekind domain. We show that every ideal is representable as a product of primes. Marcus [3, p. 42] provides a proof that this representation is unique up to ordering.

Let A be all those ideals that are not representable as a product of primes and assume towards contradiction that A is non-empty. Because R is Noetherian A must contain some maximal ideal M. Note that $M \neq (R)$ by the convention that (R) "factors" as the empty product, so $(R) \notin A$.

By [**T**]heorem 15, Lemma 2], we have that $M \subset P$ for some prime ideal P. So there must exist an ideal I such that $M = PI$. This implies that $M \subseteq I$, and in fact, the containment must be strict: If $I = M$, then $RM = PM$ so $R = P$ by the cancellation law (Corollary 4.5.1). Therefore, $M \subset I$, so $I \notin A$ by the choice that M is maximal. This means that I factors uniquely as a product of primes, so M must also factor uniquely. This is a contradiction. □

Although we won't prove it, the converse of Theorem 4.6 is also true. In fact, prime factorization of ideals is sometimes taken as an alternative definition of Dedekind domains.

This theorem allows us to still have some notion of unique factorization even in number rings that aren't UFDs. For example, in the now-familiar $\mathbb{Z}[i\sqrt{5}]$, we observed earlier that 6 fails to uniquely factor because

$$\begin{aligned} 6 &= 2 \cdot 3 \\ &= (1 + i\sqrt{5})(1 - i\sqrt{5}). \end{aligned} \tag{31}$$

Thanks to Theorem 4.6, we now have some consolation: the ideal generated by 6 does uniquely factor into primes as

$$(6) = (2, 1 + i\sqrt{5})\,(2, 1 - i\sqrt{5})\,(3, 1 + i\sqrt{5})\,(3, 1 - i\sqrt{5}). \tag{32}$$

Notice that all four prime divisors of 6 (the ring element) appear in the prime factorization of (6) (the ring ideal). This is not a coincidence. In fact, we can recover the prime factors of the *element* 6 by taking the pairwise products of the prime factors of the *ideal* (6):

$$\begin{aligned} (3) &= (3, 1 + i\sqrt{5})\,(3, 1 - i\sqrt{5}) \\ (2) &= (2, 1 + i\sqrt{5})\,(2, 1 - i\sqrt{5}) \\ (1 + i\sqrt{5}) &= (2, 1 + i\sqrt{5})\,(3, 1 + i\sqrt{5}) \\ (1 - i\sqrt{5}) &= (2, 1 - i\sqrt{5})(3, 1 - i\sqrt{5}) \end{aligned} \tag{33}$$

It appears that in a hand-wavey, very non-technical way, the unique factorization of the ideal (6) completely captures the *failure* of unique factorization of the ring element 6. We'll pursue this idea further in the next section.

For now, we conclude by noting that this recovery of prime factors is only possible because the prime ideals of (6) are not principally generated. The following theorem formalizes that intuition.

Theorem 4.7. *Let R be a Dedekind domain. Then R is a unique factorization domain if and only if it's a principal ideal domain.*

First, a lemma.

***Lemma* 4.7.1.** Let $p \in R$ be prime. Then $(p) \subseteq R$ is prime.

Proof. Let $ab \in (p)$, i.e., $p \mid ab$ because $(p) = pR$ is principal. By Euclid's Lemma, $p \mid a$ or $p \mid b$, meaning either a or b is in (p). □

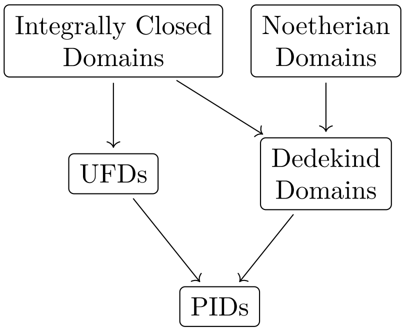

Figure 4: The two paths for an integrally closed domain with upward ambitions.

We now proceed to the proof of Theorem 4.7.

Proof. The converse (PID implies UFD) is a well known result outside the context of Dedekind domains. See e.g. Dummit and Foote [1].

To prove the forward direction, let I be an ideal of R; we show that I is principal.

By Theorem 4.5, $IJ = (\alpha)$ for some principal ideal (α) and some (not necessarily principal) ideal J. Because R is a UFD, we have that $\alpha = p_1 \cdots p_k$ for some primes $p_i \in R$, so

$$(\alpha) = (p_1 \cdots p_k) = (p_1) \cdots (p_k) \tag{34}$$

where each (p_k) is a prime ideal by Lemma 4.7.1.

Because R is a Dedekind domain, this factorization is unique, and because $IJ = (\alpha)$, it must be that I is the product of some subset of these (p_i). Let $A \subseteq \{p_i\}$ be the primes that generate those ideals. It follows that

$$I = \prod_{p_i \in A} (p_i) = \left(\prod_{p_i \in A} p_i \right). \tag{35}$$

So I is principal. □

One interpretation of Theorem 4.7 is that there are two choices for an integrally closed domain that wants to move up in the world: It can either take the traditional path and have unique factorization of its elements, or it can take the non-traditional path and have unique factorization of its ideals. But you can't do both without becoming a PID. This is illustrated in Figure 4.

We've seen examples of rings that fail to have unique factorization and compensate by taking the Dedekind domain path. Are there rings that take the other path, i.e., rings that are not Dedekind but do admit unique factorization? The answer is yes. For example, the ring $\mathbb{Q}[x, y]$ is not Dedekind because the ideal (x) is prime but not maximal: $(x) \subset (x, y)$. The infinite polynomial ring $\mathbb{Q}[x_1, x_2, \ldots]$ is not Dedekind because the ascending chain of ideals

$$(x_1) \subset (x_1, x_2) \subset \cdots \tag{36}$$

never stabilizes, so the ring isn't Noetherian. Yet both of these examples are UFDs.

5 The Ideal Class Group

We noticed in the previous section that the unique prime factorization of ideals revealed information about the *failure* of unique factorization of ring elements. This section gives a little more formality to that idea. We begin with the following construction:

Definition 5.1 (Ideal Class Group)**.** *Let $K = \mathbb{Q}[\alpha]$ be a number field. The class group of K is the set of ideals of $\mathcal{O}_K$, modulo the equivalence relation*

$$I \sim J \text{ iff } \alpha I = \beta J \text{ for some nonzero } \alpha, \beta \in R. \tag{37}$$

This happens to be the smallest equivalence relation that "kills" the principal ideals, i.e., collapses them into a single equivalence class. This follows from the fact that one can analogously define the ideal class group as the quotient of fractional ideals by principal ideals.

Returning to our old friend $\mathbb{Z}[i\sqrt{5}]$, one can verify that $(2) \sim (3)$ by the fact that $3(2) = (6) = 2(3)$. In fact, for any number field K, any two principal ideals $(a), (b) \subseteq \mathcal{O}_K$ are equivalent by the fact that $a(b) = (ab) = b(a)$. This equivalence class of principal ideals plays the role of the identity element in the class group which is—as the name suggests—a group. The group operation given by ideal multiplication of some representative element from each ideal class. We leave it to the reader to verify that this operation is well-defined. Inverses are guaranteed to exist by Theorem 4.5 which guarantees that for any ideal I, there exists some J such that IJ is principal.

What about the non-principal ideals of $\mathbb{Z}[i\sqrt{5}]$? We saw that the ideal (6) factored into four prime ideals; under $\sim$, they're all equivalent:

$$\begin{aligned} (6, 3 + 3i\sqrt{5}) &= 3\,(2, 1 + i\sqrt{5}) \\ &= 3\,(2, 1 - i\sqrt{5}) \\ &= (1 - i\sqrt{5})\,(3, 1 + i\sqrt{5}) \\ &= (1 + i\sqrt{5})\,(3, 1 - i\sqrt{5}). \end{aligned} \tag{38}$$

(In the last two lines, the left-hand terms of the multiplication are elements of $\mathbb{Z}[i\sqrt{5}]$, not ideals.)

It turns out that these are the only two equivalence classes, so we say that the class group of $\mathbb{Q}[i\sqrt{5}]$ is (isomorphic to) the cyclic group $\mathcal{C}_2$. The principal ideals play the role of the identity element 1; the non-principal ideals play the role of -1. We saw evidence of this earlier in Equation 33 when we were able to multiply the prime factors of (6) to recover the irreducible factors of 6. We didn't know it at the time, but what we really were observing is that $(-1)^2 = 1$.

5.1 The Class Number

Definition 5.2 (Class Number)**.** *The class number of a number field $K = \mathbb{Q}(\alpha)$, denoted $h(K)$, is the size of its ideal class group.*

Before we do anything too complicated, note that this new definition allows us to rephrase Theorem 4.7:

Theorem 5.3. *The number ring $\mathcal{O}_K$ has unique factorization if and only if $h(K) = 1$.*

Proof. The class number $h(K) = 1$ iff all ideals in $\mathcal{O}_K$ are principal, i.e., iff $\mathcal{O}_K$ is a PID. Number rings are Dedekind domains, so the result follows from Theorem 4.7. □

We now proceed to the main goal of our paper. With so much machinery now built up, stating the theorem is quite simple:

Theorem 5.4. *Class numbers are always finite.*

Proof. We prove Theorem 5.4 through a sequence of lemmas.

***Lemma* 5.4.1.** Let K be a number field and $\mathcal{O}_K = \mathbb{A} \cap K$ its field of integers. Then, there exists some $\lambda > 0$ such that every non-trivial ideal of R contains a nonzero α such that

$$\left|\mathrm{N}^K(\alpha)\right| \leq \lambda ||I||. \tag{39}$$

Take $\alpha_1, \ldots, \alpha_n$ to be an integral basis for $\mathcal{O}_K = \operatorname{span}_{\mathbb{Z}}\{\alpha_1, \ldots, \alpha_n\}$ and let $\sigma_1, \ldots, \sigma_n$ be the embeddings of K in $\mathbb{C}$. We claim that Equation 39 holds for

$$\lambda = \prod_{i \leq n} \sum_{j \leq n} |\sigma_i(\alpha_j)|. \tag{40}$$

To find α, fix some ideal $I \subseteq R$, take m to be the largest integer such that $m^n \leq ||I|| < (m+1)^n$ and consider the following set:

$$M = \left\{ \sum_{j \leq n} m_j \alpha_j : m_j = 0, 1, \ldots, m \right\}. \tag{41}$$

To generate an element of M, one must make n choices for m_j, each of which has $m + 1$ options (the integers $0, \ldots, m$. So $|M| = (m+1)^n > ||I||$, and the pigeonhole principle guarantees that two elements of M must be congruent modulo I—call them a and b. We take

$$\alpha = b - a = \sum_{j \leq n} m_j \alpha_j \tag{42}$$

for some integer m_j with $m_j \leq m$. It follows that

$$\begin{aligned}\left|\mathrm{N}^K(\alpha)\right| &= \prod_{i \leq n} |\sigma_i(\alpha)| \\ &\leq \prod_{i \leq n} \sum_{j \leq n} m_j |\sigma_i(\alpha_j)| \\ &\leq \lambda m^n \\ &\leq \lambda ||I||.\end{aligned} \tag{43}$$

***Lemma* 5.4.2.** Every equivalence class of $\mathcal{O}_R$ contains an ideal J with $||J|| \leq \lambda$.

Let C be an equivalence class of O_R with respect to $\sim$ (from Definition 5.1). Take some $I \in \mathbb{C}^{-1}$ and define α as in Equation 42. By construction, $\alpha \in I$, so $(\alpha) \subseteq I$ and therefore there must exist some ideal J such that $IJ = (\alpha)$ by Corollary 4.5.2. Because IJ is principal and $I \in C^{-1}$, it must be that $J \in C$. Finally,

$$\lambda ||I|| \geq |N^K(\alpha)| = ||(\alpha)|| = ||I||\,||J|| \tag{44}$$

where the last two equalities follow from [3, Theorem 22]. So $||J|| \leq \lambda$.

***Lemma* 5.4.3.** There are only finitely many ideals satisfying $||J|| \leq \lambda$.

Let P_i be the prime divisors of J with corresponding multiplicities b_i. Then,

$$\lambda \geq ||J|| = ||\prod P_i^{n_i}|| = \prod ||P_i||^{n_i}. \tag{45}$$

There are only finitely many primes P_i and multiplicities b_i satisfying this inequality, so there are only finitely many possible J.

Each ideal class must contain some ideal satisfying $||J|| \leq \lambda$, of which there are only finitely many. Theorem 5.4 follows. □

We won't give any rigorous meaning to this claim, but it seems that the class number of a number field K captures "how far away" a group is from achieving unique factorization. In the previous section, we saw that the class number of $\mathbb{Q}[i\sqrt{5}]$ was 2, so in some sense it couldn't be any closer to being a UFD.

There's an encouraging undertone here: If every number ring yearns for unique factorization, it's never infinitely far from its dreams.

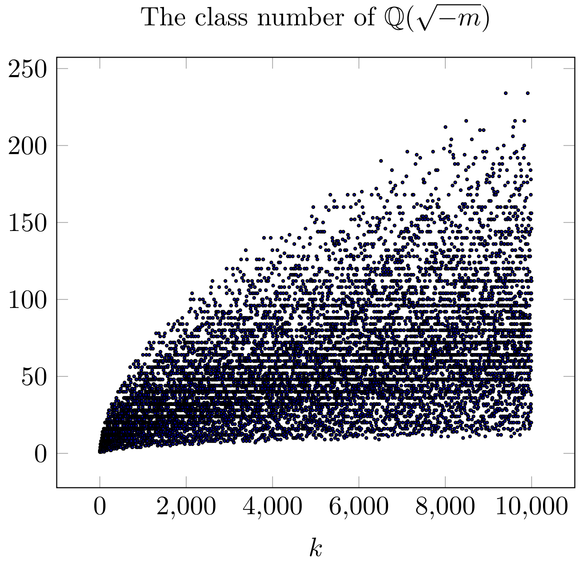

Figure 5: The long-term behavior of $h(\mathbb{Q}(\sqrt{-m}))$, taken from the OEIS [4, A000924].

6 Conclusion

To conclude, we briefly describe the behavior of the class number in the special cases of quadratic and cyclotomic fields.

In the imaginary quadratic case, it was recently proven that there are only 9 values for m for which $\mathbb{Q}[\sqrt{-m}]$ is a UFD, i.e., has class number 1. They are

$$m = 1, 2, 3, 7, 11, 19, 43, 67, 163 \tag{46}$$

and together known as the Heegner numbers [4, A000924] after Kurt Heegner who proved[1] in 1952 that this list was complete. In the long term, the class numbers for imaginary quadratic fields tend towards infinity as is shown in Figure 5.

One might reasonably hope that the purely-real quadratic case is simpler: it's not. We know comparatively very little about $\mathbb{Q}(\sqrt{m})$ when $m > 0$. It's conjectured that there are infinitely many m such that $\mathbb{Q}(\sqrt{m})$ is a UFD, but the problem is open. The first few m that produce quadratic fields with unique factorization are

$$m = 2, 3, 5, 6, 7, 11, 13, 14, 17, 19, 21, 22, 23, 29, \ldots \tag{47}$$

More are available on the OEIS [4, A003172]. Unlike the imaginary case, the there isn't clear asymptotic behavior for h as m grows (Figure 6).

[1] with minor flaws

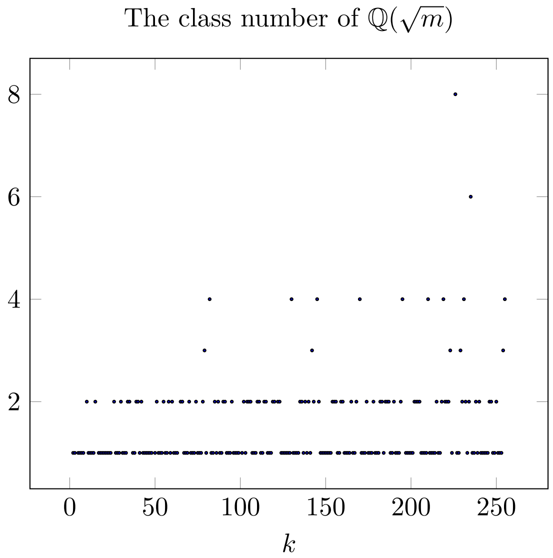

Figure 6: The long-term behavior of $h(\mathbb{Q}(\sqrt{m}))$, calculated with `sage`.

For cyclotomic fields, things are quite strange: the class number of $\mathbb{Q}[\zeta_k]$ is 1 for the first 22 integers, but when $k = 23$, suddenly it jumps to 3. At $k = 43$, the class number is already up to 211, and by the time you reach $k = 211$ the class number is the enormous value

$$49238446584179914120276706365116286443831$$

But for the preceding field ($k = 210$), the class number is a measly 13. In the long term, h becomes arbitrarily large: Figure 7.

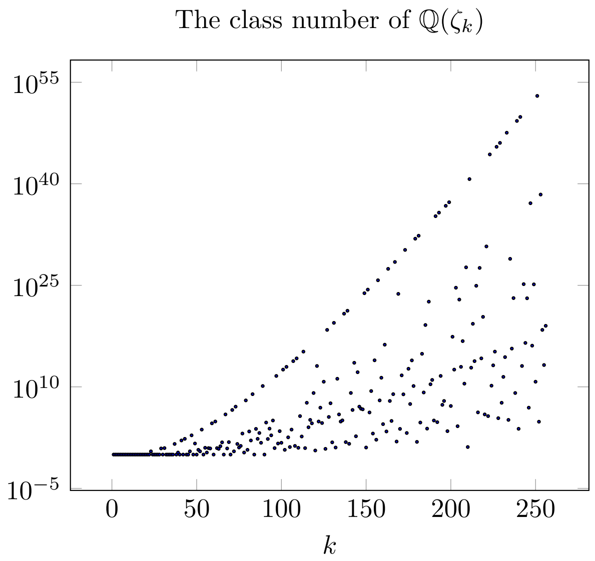

Figure 7: The long-term behavior of $h(\mathbb{Q}(\zeta_k))$, taken from the OEIS [4, A061653]

References

[1] D.S. Dummit and R.M. Foote. *Abstract Algebra*. Wiley, 2004. ISBN: 9780471452348. URL: https://books.google.com/books?id=QkAxJgAACAAJ.

[2] D.L. Goodstein and J.R. Goodstein. *Feynman's Lost Lecture: The Motion of Planets Around the Sun*. Norton, 1996. ISBN: 9780393039184. URL: https://books.google.com/books?id=o2VLdx2td0cC.

[3] D.A. Marcus. *Number Fields*. Universitext. Springer International Publishing, 2018. ISBN: 9783319902333. URL: https://link.springer.com/book/10.1007/978-3-319-90233-3.

[4] OEIS Foundation Inc. *The On-Line Encyclopedia of Integer Sequences*. 2023. URL: http://oeis.org.