Number Fields *

Zhiyuan Bai

Compiled on July 29, 2022

This document serves as a set of revision materials for the Cambridge Mathematical Tripos Part II course *Number Fields* in Lent 2022. However, despite its primary focus, readers should note that it is NOT a verbatim recall of the lectures, since the author might have made further amendments in the content. Therefore, there should always be provisions for errors and typos while this material is being used.

Contents

0 Introduction **2**

1 Algebraic Numbers, Algebraic Integers and Number Fields **2**

2 Quadratic Fields **4**

3 Embeddings **5**

4 Norm and Trace **5**

5 Some Algebra **7**

6 Discriminants and Integral Bases **7**

7 Ideals **9**

8 Unique Factorisation **11**

9 Factorisation of Rational Primes **14**

10 Geometry of Numbers **16**

11 Units **20**

12 Diophantine Equations **23**

13 Analytic Class Number Formula **23**

*Based on the lectures under the same name taught by Prof. A. Scholl in Lent 2022.

0 Introduction

In IB Groups, Rings and Modules, we've seen the ring of Gaussian integers $\mathbb{Z}[i] = \{a + bi : a, b \in \mathbb{Z}\}$ is a PID, hence a UFD. A consequence of this is Fermat's theorem of two squares: A prime is a sum of two perfect squares if and only if it is congruent to 1 (mod 4).
A more general ring of algebraic integers, however, doesn't have to behave this well. For example, $\mathbb{Z}[\sqrt{-5}]$ is not a UFD as $6 = 2 \times 3 = (1 + \sqrt{-5})(1 - \sqrt{-5})$ (and it's not difficult to check that $2, 3, 1 \pm \sqrt{-5}$ are all irreducible in $\mathbb{Z}[\sqrt{-5}]$). A natural question to ask is how good or how bad can number rings be.
Another family of examples that sparks interests are the cyclotomic rings of integers. Take a prime p, let $\zeta_p = e^{2\pi i/p}$ and consider the ring $\mathbb{Z}[\zeta_p] = \{\sum_{j=0}^{p-2} a_j \zeta_p^j : a_j \in \mathbb{Z}\}$. This ring has its historical significance in the first attempts to prove Fermat's Last Theorem, since we have

$$x^p + y^p = z^p \iff x^p = \prod_{j=0}^{p-1} (z - \zeta_p^j y)$$

So studying the factorisations in these rings could give insights to some cases of Fermat's Last Theorem, although they're not quite enough to prove it.

1 Algebraic Numbers, Algebraic Integers and Number Fields

Let F be a field extension of $\mathbb{Q}$. Recall the followings from Galois theory:

Definition 1.1. $\alpha \in F$ is algebraic (over $\mathbb{Q}$) if $f(\alpha) = 0$ for some nonzero $f(X) \in \mathbb{Q}[X]$.

Proposition 1.1. *If $\alpha \in F$ is algebraic, then there is a unique monic irreducible $m_\alpha \in \mathbb{Q}[X]$ vanishing at α such that $\forall f \in \mathbb{Q}[T], f(\alpha) = 0 \iff m_\alpha \mid f$.*

Proof. $P_\alpha = \{p \in \mathbb{Q}[X] : p(\alpha) = 0\}$ is a prime ideal of $\mathbb{Q}[X]$, which is a PID. □

Remark. P_α can also be viewed as the kernel of the homomorphism $\varphi_\alpha : \mathbb{Q}[X] \to F, f \mapsto f_\alpha$.

Definition 1.2. m_α is called the minimal polynomial of α (over $\mathbb{Q}$). The degree of α (over $\mathbb{Q}$) is $\deg_{\mathbb{Q}}(\alpha) = \deg \alpha = \deg m_\alpha$.

Example 1.1. $\deg \alpha = 1 \iff \alpha \in \mathbb{Q}$.

For $\alpha \in F$, we write $\mathbb{Q}(\alpha) \subset F$ to denote the "smallest" subfield of F containing both $\mathbb{Q}$ and α. More precisely,

$$\mathbb{Q}(\alpha) = \left\{ \frac{f(\alpha)}{g(\alpha)} : f, g \in \mathbb{Q}[X] \right\} = \bigcap_{\mathbb{Q} \le K \le F, \alpha \in K} K$$

This is a disgusting expression. Let's simplify it when things are nice.

Proposition 1.2. *If α is algebraic and $\deg \alpha = n$, then $1, \alpha, \ldots, \alpha^{n-1}$ is a basis of $\mathbb{Q}(\alpha)$ as a $\mathbb{Q}$-vector space. Conversely, if $\dim_{\mathbb{Q}} \mathbb{Q}(\alpha) = n < \infty$, then α is algebraic and $\deg \alpha = n$.*

Proof. Suppose α is algebraic with minimal polynomial m_α. Then φ_α has image isomorphic to $\mathbb{Q}[X]/P_\alpha = \mathbb{Q}[X]/(m_\alpha)$. But P_α is maximal, so this quotient has to be a field. At the same time, any subfield of F containing $\mathbb{Q}, \alpha$ must also contain the image of φ_α, so indeed $1, \alpha, \ldots, \alpha^{n-1}$ span $\mathbb{Q}(\alpha)$ over $\mathbb{Q}$. They are linearly independent since $\deg m_\alpha = n$.
On the other hand, suppose $\dim_{\mathbb{Q}} \mathbb{Q}(\alpha) = n < \infty$, then $1, \alpha, \ldots, \alpha^n$ are $\mathbb{Q}$-linearly independent over $\mathbb{Q}$, which means that α is algebraic over $\mathbb{Q}$ of degree at most n. We then know its degree is exact n by the first part of the proposition. □

Proposition 1.3. *The set of algebraic numbers in F is a subfield of F.*

We have already seen this (and frankly most of the stuff we've talked about so far) in Galois theory, so let's try to produce a different proof here.

Proof. For $\alpha \neq 0$ and $\sum_{j=0}^n b_j \alpha^j = 0$ (where $b_n \neq 0$), then $\sum_{j=0}^n b_{n-j}(\alpha^{-1})^j = 0$. We will show the closure of it under addition and multiplication later, when we prove a stronger statement. □

Definition 1.3. $\alpha \in F$ is an algebraic integer if there is some monic $f \in \mathbb{Z}[T]$ such that $f(\alpha) = 0$.

Lemma 1.4. *Let $\alpha \in F$, then the followings are equivalent:*
(i) α is an algebraic integer.
(ii) α is algebraic and its minimal polynomial has integer coefficients.
(iii) $\mathbb{Z}[\alpha]$ is a finitely-generated $\mathbb{Z}$-module.
If any of the above is true, then $1, \ldots, \alpha^{d-1}$ is a $\mathbb{Z}$-basis for $\mathbb{Z}[\alpha]$ where $d = \deg \alpha$.

In case you are confused, for $\alpha_1, \ldots, \alpha_m \in F$, we write $\mathbb{Z}[\alpha_1, \ldots, \alpha_m]$ to be the "smallest" subring of F containing $\alpha_1, \ldots, \alpha_m$. That is,

$$\mathbb{Z}[\alpha_1, \ldots, \alpha_m] = \left\{ \sum_{\underline{k}} c_{\underline{k}} \alpha_1^{k_1} \cdots \alpha_m^{k_m} : k_i \geq 0, c_{\underline{k}} \in \mathbb{Z} \right\}$$

Proof. (i) $\implies$ (ii): Assume $f(\alpha) = 0$ for some monic $f \in \mathbb{Z}[T]$, then $f = g m_\alpha$ for some $g \in \mathbb{Q}[X]$. g is monic as f, m_α both are. Gauss's Lemma then forces m_α (and g) to have integer coefficients.
(ii) $\implies$ (iii): We have $\mathbb{Z}[\alpha] = \{g(\alpha) : \alpha \in \mathbb{Z}[T]\}$. Write $m_\alpha = T^d + \sum_{0 \leq j < d} a_j T^j$ for $a_j \in \mathbb{Z}$, then $\alpha^d = -\sum_{0 \leq j < d} a_j \alpha^j$, and therefore $g(\alpha)$ is a $\mathbb{Z}$-linear combination of $1, \alpha, \ldots, \alpha^{d-1}$ for all $g \in \mathbb{Z}[T]$. They are also $\mathbb{Z}$-linearly independent, so they form a $\mathbb{Z}$-basis for $\mathbb{Z}[\alpha]$.
(iii) $\implies$ (i): Suppose $\mathbb{Z}[\alpha]$ is finitely generated by $g_1(\alpha), \ldots, g_r(\alpha)$ for some $g_i \in \mathbb{Z}[T]$. Let $k = \max_i \deg g_i$, then $\mathbb{Z}[\alpha]$ is generated by $1, \alpha, \ldots, \alpha^k$ as a $\mathbb{Z}$-module, so $\alpha^{k+1} = \sum_{0 \leq j \leq k} b_j \alpha^j$ for some $b_j \in \mathbb{Z}$ and hence α is an algebraic integer. □

Corollary 1.5. *$\alpha \in \mathbb{Q}$ is an algebraic integer iff $\alpha \in \mathbb{Z}$.*

Proof. $m_\alpha(T) = T - \alpha$. □

Theorem 1.6. *If $\alpha, \beta \in F$ are algebraic integers, so are $\alpha\beta, \alpha \pm \beta$.*

Proof. The $\mathbb{Z}$-module $\mathbb{Z}[\alpha, \beta]$ is generated by $\{\alpha^i \beta^j : 0 \leq i < \deg \alpha, 0 \leq j < \deg \beta\}$. $\mathbb{Z}[\alpha \pm \beta]$ and $\mathbb{Z}[\alpha\beta]$ are $\mathbb{Z}$-submodules of $\mathbb{Z}[\alpha, \beta]$, hence are also finitely generated (as $\mathbb{Z}$ is an Euclidean domain). □

To finish the proof of Proposition 1.3, the only ingredient left is the following:

Proposition 1.7. *If $\alpha \in F$ is algebraic, then $b\alpha$ is an algebraic integer for some $b \in \mathbb{Z}, b \geq 1$.*

Proof. Look at the minimal polynomial of $b\alpha$. □

Definition 1.4. A number field (or an algebraic number field) K is a finite field extension of $\mathbb{Q}$.

If you haven't done Galois theory, to say that the field extension $K/\mathbb{Q}$ is finite is just to say that the dimension of K as a $\mathbb{Q}$-vector space is finite. This dimension is also called the degree $[K : \mathbb{Q}]$ of the field extension $K/\mathbb{Q}$.

Definition 1.5. The ring of integers $\mathcal{O}_K$ of a number field K is the set of algebraic integers contained in K, which is a subring of K by Theorem 1.6.

If α is algebraic, then $\mathbb{Q}(\alpha)$ is a number field by Proposition 1.2. Conversely,

Theorem 1.8. *Any number field K has the form $K = \mathbb{Q}(\alpha)$ for some $\alpha \in K$.*

Proof. Primitive element theorem from Galois theory. □

2 Quadratic Fields

A number field K is quadratic if $[K : \mathbb{Q}] = 2$. Suppose K is quadratic, then for any $\alpha \in K \setminus \mathbb{Q}$ we have $K = \mathbb{Q} \oplus \mathbb{Q}\alpha$ as a vector space and $\deg m_\alpha = 2$. By quadratic formula we can write $\alpha = x + \sqrt{y}$ for some $x, y \in \mathbb{Q}$ (to avoid confusion, $\sqrt{y}$ just means an element in K whose square is y). As $\alpha \notin \mathbb{Q}$, y is not a square in $\mathbb{Q}$ and therefore $y = z^2 d$ where $z, d \in \mathbb{Q}$ and $d \neq 0, 1$ is a square-free integer. Then $K = \mathbb{Q}(\alpha) = \mathbb{Q}(\sqrt{d})$ and $K \cong \mathbb{Q}[T]/(T^2 - d)$. If $d' \in \mathbb{Z} \setminus \{0, 1\}$ is also square-free and $\mathbb{Q}(\sqrt{d}) \cong \mathbb{Q}(\sqrt{d'})$ are isomorphic, then $d = d'$ (exercise).
What is $\mathcal{O}_K$? Let $\alpha = u + v\sqrt{d} \in K = \mathbb{Q}(\sqrt{d})$. If $v = 0$, then $\alpha = u \in \mathcal{O}_K$ iff $u \in \mathbb{Z}$. Otherwise, $m_\alpha(T) = T^2 - 2uT + (u^2 - dv^2)$, so $\alpha \in \mathcal{O}_K \iff 2u, u^2 - dv^2 \in \mathbb{Z}$.
Suppose $u \in \mathbb{Z}$, then $2u \in \mathbb{Z}$. We have $u^2 - dv^2 \in \mathbb{Z} \iff dv^2 \in \mathbb{Z} \iff v \in \mathbb{Z}$ as d is square-free. Otherwise, if $u \notin \mathbb{Z}$, then $2u \in \mathbb{Z}$ forces $u = (2a+1)/2$ for some $a \in \mathbb{Z}$. Then $\alpha \in \mathcal{O}_K \iff ((2a+1)/2)^2 - dv^2 \in \mathbb{Z} \iff (4a^2 + 4a + 1) - 4dv^2 \in 4\mathbb{Z} \iff v = k/2$ for some $k \in \mathbb{Z}$ and $dk^2 \equiv 1 \pmod 4$. The latter holds iff k is odd and $d \equiv 1 \pmod 4$.
To conclude,

Theorem 2.1. *Let $d \in \mathbb{Z} \setminus \{0, 1\}$ be square-free and $K = \mathbb{Q}(\sqrt{d})$.*
(i) If $d \equiv 2, 3 \pmod 4$, then $\mathcal{O}_K = \{u + v\sqrt{d} : u, v \in \mathbb{Z}\} = \mathbb{Z}[\sqrt{d}]$.
(ii) If $d \equiv 1 \pmod 4$, then $\mathcal{O}_K = \{u + v\sqrt{d} : u, v \in (1/2)\mathbb{Z}, u - v \in \mathbb{Z}\} = \mathbb{Z}[(1 + \sqrt{d})/2]$.

Proof. Above discussion. □

Example 2.1. For $d = -1$, we have $K = \mathbb{Q}(i)$ and $\mathcal{O}_K = \mathbb{Z}[i]$.
For $d = -3$, we have $\mathcal{O}_K = \mathbb{Z}[\omega]$ where $\omega = (1 + \sqrt{-3})/2$.

Remark. 1. For quadratic field, it's easy to work determine the ring of integers explicitly by hand. As might be expected, this is hard for other kinds of number fields. We will develop some algebraic tools to study the ring of integers for general K without having to compute it directly.
2. $\mathcal{O}_K$ does not have to have the form $\mathbb{Z}[\alpha]$ for some algebraic integer α.

3 Embeddings

For $d > 0$, there are two homomorphisms σ_1, σ_2 from $K = \mathbb{Q}(\sqrt{d})$ to $\mathbb{R}$ by mapping $\sqrt{d}$ to $\pm\sqrt{d}$ respectively. When $d < 0$, there are two homomorphisms σ_1, σ_2 from K to $\mathbb{C}$ by taking $\sqrt{d}$ to $\pm i\sqrt{-d}$. What about in a general number field?
Let K be a number field of degree $n = [K : \mathbb{Q}]$.

Theorem 3.1. *There are exactly n distinct field homomorphisms $\sigma_1, \ldots, \sigma_n : K \to \mathbb{C}$ (the "complex embeddings of K"). Moreover, if $\mathbb{Q} \subset F \subset K$ are number fields, then each of the $[F : \mathbb{Q}]$ complex embeddings of F extends to exactly $[K : F]$ embeddings of K.*

Proof. Say $K = \mathbb{Q}(\alpha) \cong \mathbb{Q}[T]/(m_\alpha)$. A field homomorphism $\sigma : K \to \mathbb{C}$ is the same as a ring homomorphism $\tilde{\sigma} : \mathbb{Q}[T] \to \mathbb{C}$ with $\tilde{\sigma}(m_\alpha(T)) = m_\alpha(\tilde{\sigma}(T)) = 0$, i.e. the information of a root of m_α in $\mathbb{C}$. This gives a bijection between embeddings $K \to \mathbb{C}$ and the roots of m_α in $\mathbb{C}$. Since m_α is irreducible (hence separable as we are in characteristic 0), the latter set has exactly n elements.
The exact same argument will work with the second part of the theorem. □

Remark. 1. If K is given as a subfield of $\mathbb{C}$, it is often convenient to take σ_1 to be this inclusion.
2. Suppose $\sigma_1, \ldots, \sigma_r$ are the embeddings whose images are contained in $\mathbb{R}$, then the remaining embeddings come in conjugate pairs since the nnoreal roots of a real polynomial come in conjugate pairs. So we can write $n = r + 2s$ where s is the number of such conjugate pairs $(\sigma, \bar{\sigma} \neq \sigma)$.

Proposition 3.2. *For $\alpha \in K$ (not necessarily a primitive element), the complex numbers $\sigma_1(\alpha), \ldots, \sigma_n(\alpha)$ (the "conjugates" of α) are the complex roots of m_α each with multiplicity $n/\deg \alpha$.*

Proof. Follows directly from the preceding theorem by taking $F = \mathbb{Q}(\alpha)$. □

4 Norm and Trace

Let K be a number field and $\alpha \in K$. Suppose $u_\alpha : K \to K$ is the map $u_\alpha(x) = \alpha x$. Then u_α is a $\mathbb{Q}$-linear map $K \to K$ with K now viewed as a $\mathbb{Q}$-vector space.

Definition 4.1. The charactertistic polynomial of α is $f_\alpha = \det(T - u_\alpha) \in \mathbb{Q}[T]$. The norm of α is $N_{K/\mathbb{Q}}(\alpha) = \det u_\alpha$ and the trace of α is $\mathrm{Tr}_{K/\mathbb{Q}}(\alpha) = \mathrm{tr}\, u_\alpha$.

Explicitly, if $\beta_1, \ldots, \beta_n$ is a $\mathbb{Q}$-basis for K and $\alpha\beta_i = \sum_j A_{ji}\beta_j$, then the rational matrix $A = (A_{ij})$ is the matrix for u_α and $f_\alpha, N_{K/\mathbb{Q}}(\alpha), \mathrm{Tr}_{K/\mathbb{Q}}(\alpha)$ are the charactertistic polynomial, determinant and trace of A respectively.

Example 4.1. Consider the case of a quadratic field $K = \mathbb{Q}(\sqrt{d})$. Under the basis $1, \sqrt{d}$, $u_{x+y\sqrt{d}}$ has matrix

$$\begin{pmatrix} x & dy \\ y & x \end{pmatrix}$$

So $f_\alpha(T) = T^2 - 2xT + (x^2 - dy^2), N_{K/\mathbb{Q}}(\alpha) = x^2 - dy^2$ and $\mathrm{Tr}_{K/\mathbb{Q}}(\alpha) = 2x$.

Remark. Clearly $u_{\alpha+\beta} = u_\alpha + u_\beta$ and $u_{\alpha\beta} = u_\alpha u_\beta = u_\beta u_\alpha$, and in particular $u_{m\alpha} = m u_\alpha$ for any $m \in \mathbb{Q}$. Combining them gives $u_{f(\alpha)} = f(u_\alpha)$ for any $f \in \mathbb{Q}[X]$.

Proposition 4.1. $N_{K/\mathbb{Q}}(\alpha\beta) = N_{K/\mathbb{Q}}(\alpha)N_{K/\mathbb{Q}}(\beta)$; $\mathrm{Tr}_{K/\mathbb{Q}}(\alpha+\beta) = \mathrm{Tr}_{K/\mathbb{Q}}(\alpha) + \mathrm{Tr}_{K/\mathbb{Q}}(\beta)$.

Proof. Immediate from the preceding remark. □

One also observe from the example of a quadratic field that there seem to be a relation between the conjugates of α and the values of $N_{K/\mathbb{Q}}(\alpha), \mathrm{Tr}_{K/\mathbb{Q}}(\alpha)$.

Theorem 4.2. *(i) Suppose K has complex embeddings $\{\sigma_i\}_{i=1}^n$ and $\alpha \in K$ has degree d, then*

$$f_\alpha(T) = \prod_{i=1}^{n}(T - \sigma_i(\alpha)) = m_\alpha(T)^{n/d}$$

(ii) $N_{K/\mathbb{Q}}(\alpha) = \prod_i \sigma_i(\alpha), \mathrm{Tr}_{K/\mathbb{Q}}(\alpha) = \sum_i \sigma_i(\alpha)$.

Proof. (i) Suppose first that $n = d$, then $f_\alpha(\alpha)\beta = f_\alpha(u_\alpha)\beta = 0$ for any $\beta \in K$ by Cayley-Hamilton. Consequently $f_\alpha(\alpha) = 0$, so $m_\alpha \mid f_\alpha$. But $\deg m_\alpha = \deg f_\alpha$ and both of them are monic, so $f_\alpha = m_\alpha$. We get the middle term by recalling that $\sigma_i(\alpha)$ are the complex roots of m_α.
The general case follows from the fact that $K \cong \mathbb{Q}(\alpha)^{n/d}$ as a $\mathbb{Q}(\alpha)$-vector space.
(ii) Expanding (i). □

Remark. Some people take the formulae in the theorem as the definitions of norm and trace. We are better than that.

Corollary 4.3. *(i) For $\alpha \in K$, $N_{K/\mathbb{Q}}(\alpha) = 0$ iff $\alpha = 0$.*
(ii) For $\alpha \in \mathcal{O}_K$, $f_\alpha \in \mathbb{Z}[T]$ and $N_{K/\mathbb{Q}}(\alpha), \mathrm{Tr}_{K/\mathbb{Q}}(\alpha)$ are integers. Moreover, $\alpha \in \mathcal{O}_K^\times$ iff $N_{K/\mathbb{Q}}(\alpha) = \pm 1$.

Proof. (i) Clear from the preceding theorem.
(ii) if $\alpha \in \mathcal{O}_K$, $m_\alpha \in \mathbb{Z}[T]$ and hence $f_\alpha \in \mathbb{Z}[T]$ since it's a power of m_α. $N_{K/\mathbb{Q}}(\alpha), \mathrm{Tr}_{K/\mathbb{Q}}(\alpha)$ are coefficients of f_α, hence also integers.
If $\alpha \in \mathcal{O}_K^\times$, there is some $\beta \in \mathcal{O}_K$ such that $\alpha\beta = 1$, hence $N_{K/\mathbb{Q}}(\alpha)N_{K/\mathbb{Q}}(\beta) = N_{K/\mathbb{Q}}(1) = 1$. But both $N_{K/\mathbb{Q}}(\alpha), N_{K/\mathbb{Q}}(\beta)$ are integers, so they can only be ± 1. Conversely, suppose $N_{K/\mathbb{Q}}(\alpha) = \pm 1$, then $f_\alpha(T) = T^n + \sum_{j=0}^{n-1} b_j T^j$ has $b_0 \in \{\pm 1\}$. Consequently $\alpha\beta = 1$ where $\beta = \mp\left(\alpha^{n-1} + \sum_{j=0}^{n-2} b_{j+1}\alpha^j\right)$ □

5 Some Algebra

Proposition 5.1. *Suppose G is a torsion-free finitely-generated abelian group with rank $n \geq 1$. Let $x_1, \ldots, x_n$ be a family of generators of it and $H \leq G$ be the subgroup generated by $y_1, \ldots, y_n$ where $y_i = \sum_j A_{ji} x_j$. If $\det A \neq 0$, then H is a subgroup of finite index and $[G : H] = |\det A|$.*

Proof. Find invertible matrices $P, Q \in \mathrm{Mat}_{n\times n}(\mathbb{Z})$ such that $\det P, \det Q \in \{\pm 1\}$ and $A = PDQ$ with D in Smith normal form, i.e. $D = \mathrm{diag}(d_1, \ldots, d_n)$ with $d_i \in \mathbb{Z}_{>0}$ and $d_1 \mid d_2 \mid \cdots \mid d_n$ (although we don't really need the divisibility part). Then $G/H = (\mathbb{Z}/d_1\mathbb{Z}) \times \cdots \times (\mathbb{Z}/d_n\mathbb{Z})$ and hence $[G : H] = \prod_i d_i = \det D = |\det A|$. □

Definition 5.1. Let V be a $\mathbb{Q}$-vector space with finite dimension $n \geq 1$ and $H \leq V$ be a subgroup (in other words a $\mathbb{Z}$-submodule). We define the rank of H to be $\dim \mathrm{Span}\, H$.

Proposition 5.2. *Let $H \leq V$ be a finitely generated subgroup of rank r, then $H = \bigoplus_{i=1}^{r} \mathbb{Z} x_i$ for some linearly independent $x_1, \ldots, x_r$ in V.*

Proof. H is torsion-free since V is. So $H = \bigoplus_{i=1}^{d} \mathbb{Z} x_i$ for some $x_i \in H \leq V$.
If the x_i's were linearly dependent, then there are some $m_i \in \mathbb{Q}$ not all zero with $\sum_i x_i m_i = 0$. Clearing denominators then yield a nontrivial $\mathbb{Z}$-linear relation, contradiction.
We also have $r = d$ since $\mathrm{Span}\{x_i\}_{i=1}^{d} = \mathrm{Span}\, H$. □

6 Discriminants and Integral Bases

For a number field K, we aim to show that $\mathcal{O}_K = \bigoplus_{i=1}^{n} \mathbb{Z}\omega_i$ for some algebraic integers $\omega_1, \ldots, \omega_n \in \mathcal{O}_K$ and $n = [K : \mathbb{Q}]$.

Definition 6.1. Let $\alpha_1, \ldots, \alpha_n \in K$. The discriminant of them is

$$\mathrm{Disc}(\alpha_1, \ldots, \alpha_n) = \det((\mathrm{Tr}_{K/\mathbb{Q}}(\alpha_i \alpha_j))_{1\leq i,j\leq n}) \in \mathbb{Q}$$

Theorem 6.1. *(i)* $\mathrm{Disc}(\alpha_1, \ldots, \alpha_n) = \det((\sigma_i(\alpha_j))_{1\leq i,j\leq n})^2$.
(ii) $\mathrm{Disc}(\alpha_1, \ldots, \alpha_n) \neq 0$ *iff* $\{\alpha_i\}_{i=1}^{n}$ *is a $\mathbb{Q}$-basis for K.*
(iii) If $\beta_i = \sum_j A_{ji}\alpha_j$ with $A_{ij} \in \mathbb{Q}$, then

$$\mathrm{Disc}(\beta_1, \ldots, \beta_n) = \mathrm{Disc}(\alpha_1, \ldots, \alpha_n)(\det A)^2$$

(iv) Suppose $\{\alpha_i\}_{i=1}^{n}$ is a $\mathbb{Q}$-basis for K, then $\mathrm{Disc}(\alpha_1, \ldots, \alpha_n)$ depends only on the subgroup $\mathbb{Z}\alpha_1 + \cdots + \mathbb{Z}\alpha_n \leq K$.

Proof. (i) Let $\Delta = (\sigma_i(\alpha_j))_{1\leq i,j\leq n} \in \mathrm{Mat}_{n\times n}(\mathbb{C})$, then we have

$$(\Delta^\top \Delta)_{ij} = \sum_{k=1}^{n} \sigma_k(\alpha_i)\sigma_k(\alpha_j) = \sum_{k=1}^{n} \sigma_k(\alpha_i\alpha_j) = \mathrm{Tr}_{K/\mathbb{Q}}(\alpha_i\alpha_j)$$

(ii) Suppose $\alpha_1, \ldots, \alpha_n$ is not a basis, then $\sum_j b_j\alpha_j = 0$ for some $b_j \in \mathbb{Q}$ not all zero. Then $\sum_j \sigma_i(\alpha_j)b_j = 0$ for all i, which means that Δ is singular and hence $\mathrm{Disc}(\alpha_1, \ldots, \alpha_n) = (\det \Delta)^2 = 0$ by (i).

Conversely, suppose $\{\alpha_i\}_{i=1}^n$ is a basis. Let $T = (\mathrm{Tr}_{K/\mathbb{Q}}(\alpha_i\alpha_j))_{1\le i,j\le n}$. For any $b \in \mathbb{Q}^n \setminus \{0\}$, we want to show that $Tb \neq 0$. Equivalently, there is some $c \in \mathbb{Q}^n$ with $c^\top Tb \neq 0$.
For $\beta = \sum_i b_i\alpha_i \in K, \gamma = \sum_i c_i\alpha_i \in K$, we have

$$c^\top Tb = \sum_{1\le i,j\le n} b_i c_j \,\mathrm{Tr}_{K/\mathbb{Q}}(\alpha_i\alpha_j) = \mathrm{Tr}_{K/\mathbb{Q}}(\beta\gamma)$$

So we just need some c with $\beta\gamma = 1$, which is possible since $\{\alpha_i\}_{i=1}^n$ is a basis.
(iii) Let $\Delta' = (\sigma_i(\beta_j))_{1\le i,j\le n}$, then

$$\Delta'_{ij} = \sum_k \sigma_i(A_{kj}\alpha_k) = \sum_k A_{kj}\sigma_i(\alpha_k) = (\Delta A)_{ij}$$

Applying (i) gives the conclusion.
(iv) Suppose $(\alpha_i)_i, (\beta_i)_i$ generate the same subgroup of K, then $\beta_i = \sum_j A_{ji}\alpha_j$ for $A_{ij} \in \mathbb{Z}, \det A = \pm 1$. We conclude by (iii). □

Suppose $H \subset K$ is a finitely-generated subgroup of rank $n = [K : \mathbb{Q}]$, then $H = \mathbb{Z}\alpha_1 + \cdots + \mathbb{Z}\alpha_n$ for some linearly independent $(\alpha_i)_i$ by Proposition 5.2. We know that $\mathrm{Disc}(\alpha_1, \ldots, \alpha_n) \in \mathbb{Q} \setminus \{0\}$ depensds only on H by the preceding theorem, so it makes sense to write $\mathrm{Disc}(H) = \mathrm{Disc}(\alpha_1, \ldots, \alpha_n)$.

Lemma 6.2. *If $H \le H' \le K$ and both H, H' are finitely generated of rank $n = [K : \mathbb{Q}]$, then $\mathrm{Disc}(H) = \mathrm{Disc}(H')[H' : H]^2$.*

Proof. Say $H = \mathbb{Z}\alpha_1 + \cdots + \mathbb{Z}\alpha_n$ and $H = \mathbb{Z}\alpha'_1 + \cdots + \mathbb{Z}\alpha'_n$. As $H \le H'$, $\alpha_i = \sum_j B_{ji}\alpha'_j$ for some $B_{ij} \in \mathbb{Z}$. Then by Proposition 5.1 we have $[H' : H]^2 = \det B = \mathrm{Disc}(H)/\mathrm{Disc}(H')$. □

Theorem 6.3. *Let $n = [K : \mathbb{Q}]$. There exist $\omega_1, \ldots, \omega_n \in \mathcal{O}_K$ such that $\mathcal{O}_K = \mathbb{Z}\omega_1 \oplus \cdots \oplus \mathbb{Z}\omega_n$.*

Definition 6.2. We call such a collection $\omega_1, \ldots, \omega_n$ an integral basis for K.

Proof. There exists a $\mathbb{Q}$-basis $\omega_1, \ldots, \omega_n$ for K with $\omega_i \in \mathcal{O}_K$ for all i by clearing denominators. Let $H = \mathbb{Z}\omega_1 + \cdots + \mathbb{Z}\omega_n$, then $\mathrm{Disc}(H) = \mathrm{Disc}(\omega_1, \ldots, \omega_n) \in \mathbb{Z} \setminus \{0\}$ (as $\omega_i\omega_j \in \mathcal{O}_K$). This means that we can choose H (i.e. choose $\omega_1, \ldots, \omega_n$) such that $|\mathrm{Disc}\, H|$ is minimal.
Certainly $H \le \mathcal{O}_K$. Also, for any $\alpha \in \mathcal{O}_K$, $H' = H + \mathbb{Z}\alpha \supset H$ has rank n and, unless $H' = H$ (i.e. $\alpha \in H$), has strictly smaller absolute discriminant than H by the preceding lemma. But H already has minimal absolute discriminant, so necessarily $\alpha \in H$. This means that $H \ge \mathcal{O}_K$, hence $H = \mathcal{O}_K$.
Since $\omega_1, \ldots, \omega_n$ is a $\mathbb{Q}$-basis, they must freely generate $H = \mathcal{O}_K$ and we are done. □

Definition 6.3. $d_K = \mathrm{Disc}\, \mathcal{O}_K$ is the discriminant of the number field K.

Example 6.1. 1. $d_{\mathbb{Q}} = 1$.
2. Consider $K = \mathbb{Q}(\sqrt{d})$ for square-free $d \in \mathbb{Z} \setminus \{0, 1\}$. If $d \not\equiv 1 \pmod 4$, then $1, \sqrt{d}$ is an integral basis, in which case $\Delta = \begin{pmatrix} 1 & \sqrt{d} \\ 1 & -\sqrt{d} \end{pmatrix}$ and $d_K = (\det \Delta)^2 = 4d$. If $d \equiv 1 \pmod 4$, then $1, (1 + \sqrt{d})/2$ is an integral basis, thus a similar calculation gives $d_K = d$.

Proposition 6.4. *Suppose $K = \mathbb{Q}(\theta)$ and let $f = m_\theta, n = \deg f$, then*

$$\mathrm{Disc}(1, \theta, \ldots, \theta^{n-1}) = \prod_{i<j} (\sigma_i(\theta) - \sigma_j(\theta))^2 = (-1)^{n(n-1)/2} N_{K/\mathbb{Q}}(f'(\theta))$$

Proof. Recall the Vandermonde determinant

$$\mathrm{VDM}(X_1, \ldots, X_n) = \det \begin{pmatrix} X_1^{n-1} & \cdots & X_n^{n-1} \\ \vdots & \ddots & \vdots \\ X_1 & \cdots & X_n \\ 1 & \cdots & 1 \end{pmatrix} = \prod_{1 \le i < j \le n} (X_i - X_j)$$

from which we immediately have the first equality as $\mathrm{Disc}(1, \theta, \cdots, \theta^{n-1}) = \mathrm{VDM}(\sigma_1(\theta), \ldots, \sigma_n(\theta))^2$. The second equality is left as an exercise. □

Proposition 6.5. *Suppose $\omega_1, \ldots, \omega_n \in \mathcal{O}_K$ are such that $\mathrm{Disc}(\omega_1, \ldots, \omega_n) \neq 0$ is a square-free integer, then $\omega_1, \ldots, \omega_n$ is an integral basis.*

Proof. $H = \mathbb{Z}\omega_1 + \cdots + \mathbb{Z}\omega_n$ has rank n as $\mathrm{Disc}(\omega_1, \ldots, \omega_n) \neq 0$. Then $\mathrm{Disc}(H) = [\mathcal{O}_K : H]^2 \mathrm{Disc}(\mathcal{O}_K)$ forces $[\mathcal{O}_K : H] = 1$ as $\mathrm{Disc}(H) = \mathrm{Disc}(\omega_1, \ldots, \omega_n)$ is square-free. □

Remark. As shown in the quadratic field case, d_K is not always square-free, so one cannot always rely on this proposition to verify something is an integral basis.
But if $\alpha_1, \ldots, \alpha_n \in \mathcal{O}_K$ are such that $\mathrm{Disc}(\alpha_1, \ldots, \alpha_n) \neq 0$, then the idea of the proof shows that $[\mathcal{O}_K : \mathbb{Z}\alpha_1 + \cdots + \mathbb{Z}\alpha_n]^2 \mid \mathrm{Disc}(\alpha_1, \ldots, \alpha_n)$. Consequently, any $\theta \in \mathcal{O}_K$ has $m\theta \in \mathbb{Z}\alpha_1 + \cdots + \mathbb{Z}\alpha_n$ for some m with $m^2 \mid \mathrm{Disc}(\alpha_1, \ldots, \alpha_n)$. This gives an algorithm to compute $\mathcal{O}_K$.

7 Ideals

Definition 7.1. Let R be a (commutative) ring. An ideal I of R is an additive subgroup of it such that $\forall r \in R, x \in I$, we have $rx \in I$.

Equivalently, I is an R-submodule of R.
We are intersted in (nonzero) ideals of $\mathcal{O}_K$. Every such ideal has finite index. We actually have a more precise characterisation of this index.

Proposition 7.1. *Let K be a number field with $n = [K : \mathbb{Q}]$ and suppose $I \le \mathcal{O}_K$ is a nonzero ideal.*
(i) $I = \bigoplus_{i=1}^n \mathbb{Z}\alpha_i$ for some linearly independent $\alpha_1, \ldots, \alpha_n \in I$. Furthermore, $[\mathcal{O}_K : I]^2 = \mathrm{Disc}(I)/d_K$.
(ii) If $I = \alpha\mathcal{O}_K, \alpha \in \mathcal{O}_K \setminus \{0\}$ is a principal ideal, then $[\mathcal{O}_K : I] = |N_{K/\mathbb{Q}}(\alpha)|$.

Definition 7.2. Let I be an ideal of $\mathcal{O}_K$. We call $[\mathcal{O}_K : I]$ the norm $N(I)$ of I.

Proof. (i) As $\mathcal{O}_K$ is a finitely generated $\mathbb{Z}$-module, so is I. Take some $\alpha \in I \setminus \{0\}$ and let $\omega_1, \ldots, \omega_n$ be an integral basis for $\mathcal{O}_K$. Then $\alpha\omega_1, \ldots, \alpha\omega_n \in I$ are $\mathbb{Q}$-linearly independent, thus I has rank n. The claim then follows from Proposition

5.2 and Lemma 6.2.
(ii) $I = \alpha\mathcal{O}_K = \bigoplus_{i=1}^n \mathbb{Z}\alpha\omega_i$, so

$$\begin{aligned}\mathrm{Disc}(I) &= \mathrm{Disc}(\alpha\omega_1, \ldots, \alpha\omega_n) = \det((\sigma_i(\alpha\omega_j))_{ij})^2 = \det((\sigma_i(\alpha)\sigma_i(\omega_j))_{ij})^2 \\ &= \left(\prod_{i=1}^n \sigma_i(\alpha)\right)^2 \det((\sigma_i(\omega_j))_{ij})^2 = (N_{K/\mathbb{Q}}(\alpha))^2 d_K\end{aligned}$$

Applying (i) gives the result. □

Corollary 7.2. *(i) Every nonzero ideal contains a nonzero integer, namely its norm.*
(ii) There are only finitely many ideals of any given norm.

Proof. (i) $N(I)x = 0$ for all $x \in \mathcal{O}_K/I$ since $|\mathcal{O}_K/I| = N(I)$.
(ii) Let $M = N(I) \in \mathbb{Z}_{>0}$, then $M\mathcal{O}_K \leq I \leq \mathcal{O}_K$. But the set of ideals containing $M\mathcal{O}_K$ is in bijection with the ideals of $\mathcal{O}_K/M\mathcal{O}_K$, which is finite hence has only finitely many ideals. □

Definition 7.3. An ideal $P \leq R$ is prime if $P \neq R$ and $\forall \alpha, \beta \in R$, $\alpha\beta \in R$ implies that either $\alpha \in R$ or $\beta \in R$. Equivalently, R/P is an integral domain.
An ideal $I \leq R$ is maximal if $I \neq R$ and any $J \leq R$ with $I \leq J \leq R$ must either have $I = J$ or $J = R$. Equivalently, R/I is a field.

For our purpose, we will always mean a nonzero prime ideal when we mention a prime ideal.

Lemma 7.3. *Let $P \subset \mathcal{O}_K$ be a prime ideal, then:*
(i) P is maximal.
(ii) $P \cap \mathbb{Z} = p\mathbb{Z}$ for some (necessarily unique) prime $p \in \mathbb{Z}$ and $N(P) = p^f$ for some $1 \leq f \leq n = [K : \mathbb{Q}]$.

Proof. (i) Finite integral domains are fields.
(ii) Let $M = N(I)$. As $P \neq \mathcal{O}_K$, $M \neq 1$. Since P is prime, $p \in P$ for some $p \mid M$, thus $p\mathcal{O}_K \leq P \leq \mathcal{O}_K$. Hence $p\mathbb{Z} \leq P \cap \mathbb{Z} \leq \mathbb{Z}$. But $P \cap \mathbb{Z} \neq \mathbb{Z}$ as $P \neq \mathcal{O}_K$, so $P \cap \mathbb{Z} = p\mathbb{Z}$. Furthermore, $[\mathcal{O}_K : p\mathcal{O}_K] = p^n$, so $N(P) = [\mathcal{O}_K : P] \mid p^n$. □

As usual we write (S) to denote the ideal generated by $S \subset \mathcal{O}_K$.

Definition 7.4. For ideals $I, J \leq R$, their sum is $I + J = \{\alpha + \beta : \alpha \in I, \beta \in J\} \leq \mathcal{O}_K$ and their product is $IJ = (\{\alpha\beta : \alpha \in I, \beta \in J\}) = \{\sum_{i=1}^n \alpha_i\beta_i : \alpha_i \in I, \beta_i \in J\}$.

It is clear that these operations are associative and commutative. Also, in terms of generators, we have $(\alpha_1, \ldots, \alpha_k) + (\beta_1, \ldots, \beta_l) = (\alpha_1, \ldots, \alpha_k, \beta_1, \ldots, \beta_l)$ and $(\alpha_1, \ldots, \alpha_k)(\beta_1, \ldots, \beta_l) = (\alpha_1\beta_1, \ldots, \alpha_k\beta_1, \alpha_1\beta_2, \ldots, \alpha_k\beta_l)$. Note that although we have $(\alpha)(\beta) = (\alpha\beta)$, we don't in general have $(\alpha) + (\beta) = (\alpha + \beta)$. Arithmetic of ideals turns out to be the correct way to rectify unique factorisation, which fails in general rings of integers.

Example 7.1. $\mathbb{Z}[\sqrt{-5}] = \mathcal{O}_K$ for $K = \mathbb{Q}(\sqrt{5})$. This is not a UFD: $(1 + \sqrt{-5}) \times (1 - \sqrt{-5}) = 2 \times 3$. But the news aren't all bad. Turns out, we can rectify this by looking at a different kind of unique factorisation formulated using ideals. We have $2\mathcal{O}_K = P_2^2$ where $P_2 = 2\mathcal{O}_K + (1 + \sqrt{-5})\mathcal{O}_K$, $3\mathcal{O}_K = P_{3,+}P_{3,-}$

where $P_{3,\pm} = 3\mathcal{O}_K + (1 \pm \sqrt{-5})\mathcal{O}_K$, and $(1 \pm \sqrt{-5})\mathcal{O}_K = P_2 P_{3,\pm}$. The non-unique factorisation we've seen can then be explained with $(2\mathcal{O}_K)(3\mathcal{O}_K) = (P_2^2)(P_{3,+}P_{3,-}) = (P_2 P_{3,+})(P_2 P_{3,-}) = ((1+\sqrt{-5})\mathcal{O}_K)((1-\sqrt{-5})\mathcal{O}_K)$.

It turns out that such unique factorisation of ideals is a general phenomenon, even when $\mathcal{O}_K$ fails to be a UFD. What really makes $\mathcal{O}_K$ not a UFD is the fact that P_2 is not a principal ideal. To measure the failure of unique factorisation, we will study the obstructions to ideals being principal.

8 Unique Factorisation

Our goal is to prove that every nonzero ideal is uniquely expressible as a product of prime ideals.
For an ideal $I \leq \mathcal{O}_K$ and $\alpha \in \mathcal{O}_K$, we certainly have $\alpha I \leq I$. Conversely,

Lemma 8.1. *Let $I \leq \mathcal{O}_K$ be a nonzero ideal and $\alpha \in K$. Suppose $\alpha I \leq I$, then $\alpha \in \mathcal{O}_K$.*

Proof. $\alpha^k I \leq I$ for every $k \geq 0$. Pick any $\beta \in I \setminus \{0\}$, then $\mathbb{Z}[\alpha]\beta \leq I$, so $\mathbb{Z}[\alpha] \leq \beta^{-1} I$. But $\beta^{-1} I$ is a finitely generated $\mathbb{Z}$-module, so $\mathbb{Z}[\alpha]$ has to be as well, which precisely means that $\alpha \in \mathcal{O}_K$. □

Our first aim is to show that every ideal is a product of prime ideals. To start with,

Lemma 8.2. *(i) Let $I \leq \mathcal{O}_K$ be a nonzero ideal, then there exists (nonzero) prime ideals $P_1, \ldots, P_r \leq \mathcal{O}_K$, not necessarily distinct, such that $I \supset P_1 \cdots P_r$. (ii) Suppose $P, P_1, \ldots, P_r \leq \mathcal{O}_K$ are (nonzero) prime ideals with $P \supset P_1 \cdots P_r$, then $P = P_i$ for some $1 \leq i \leq r$.*

The second statement is somewhat an analogy to the fact that if $p, p_1, \ldots, p_r$ are primes in $\mathbb{Z}$ and $p \mid p_1, \ldots, p_r$, then $p = p_i$ for some i. The analogy can be strengthened by rephrasing $m \mid n$ as $(m) \supset (n)$.

Proof. (i) Induction on $N(I)$. There is nothing to prove if $I = \mathcal{O}_K$ or if I is prime. Otherwise, I is not prime and hence we can find $\alpha, \beta \notin I$ with $\alpha\beta \in I$. Then $I + (\alpha), I + (\beta)$ are ideals properly containing I, so each of them has strictly smaller norm. We are then done by induction hypothesis since $(I + (\alpha))(I + (\beta)) = I^2 + \alpha I + \beta I + (\alpha\beta) \subset I$.
(ii) Suppose $P \neq P_1$, then we can find some $\alpha \in P_1 \setminus P$ (noting that P_1 is maximal since it's prime). For any $\beta \in P_2 \cdots P_r$, we have $\alpha\beta \in P_1 \cdots P_r \subset P$, so $\beta \in P$. We therefore have $P \supset P_2 \cdots P_r$. An induction argument then gives the result. □

Corollary 8.3. *Suppose $I \leq \mathcal{O}_K$ is a nonzero proper ideal and $0 \neq \alpha \in I$, then there exists $\beta \in \mathcal{O}_K$ such that $\beta \notin (\alpha)$ with $\beta I \subset (\alpha)$.*

Proof. As I is proper, there exists some prime ideal P such that $I \subset P$ (e.g. by taking the ideal containing I with minimal norm). It suffices to find $\beta \in \mathcal{O}_K \setminus (\alpha)$ with $\beta P \subset (\alpha)$. By the preceding lemma, we can find $P_1, \ldots, P_r$ with r minimal such that $(\alpha) \supset P_1 \cdots P_r$. But $P \supset (\alpha)$, so WLOG $P = P_1$. The minimality of r means that we can find $\beta \in P_2 \cdots P_r \setminus (\alpha)$ which would have $\beta P \subset (\alpha)$. □

Theorem 8.4. *Let $I \leq \mathcal{O}_K$ be a nonzero ideal, then there exists a nonzero ideal J such that IJ is principal.*

Proof. Again induction on $N(I)$. The case $N(I) = 1$ is clear. For $N(I) > 1$, let $I \subset P$ for P prime and pick $0 \neq \alpha \in P$. By the preceding corollary, we can find $\beta \in \mathcal{O}_K \setminus (\alpha)$ with $\beta P \subset \alpha$. As $\beta P \subset (\alpha)$, $\alpha^{-1}\beta P \subset \mathcal{O}_K$, so $\alpha^{-1}\beta P \nsubseteq P$ as $\alpha^{-1}\beta \notin \mathcal{O}_K$ (Lemma 8.1). The maximality of P then shows that $\alpha^{-1}\beta P + P = \mathcal{O}_K$.
Consider $H = (\alpha, \beta) \leq \mathcal{O}_K$, then $PH = \beta P + \alpha P = (\alpha)$ is principal. What about I? We have $IH \subset PH = (\alpha)$, so $I' = \alpha^{-1}IH \leq \mathcal{O}_K$. If we can show that $I' \supsetneq I$ then we are done by induction hypothesis, for if $I'J'$ is principal then $IJ'H = \alpha I'J'$ too is principal. Indeed, we have $I' \supset I$ since $H \supset (\alpha)$. If $I' = I$, then $(\alpha^{-1}H)I = I$, which then means that $\alpha^{-1}\beta I \subset I$. But $\alpha^{-1}\beta \notin \mathcal{O}_K$ which contradicts Lemma 8.1. □

Theorem 8.5. *Suppose $I, I', J \leq \mathcal{O}_K$ are nonzero ideals, then:*
(i) (Cancellation Law) $IJ = I'J \iff I = I'$.
(ii) (To Divide is to Contain) $I \supset J \iff \exists H \leq \mathcal{O}_K, IH = J$.
(iii) (Unique Factorisation) There exists distinct prime ideals $P_1, \ldots, P_r$ and positive integers $a_1, \ldots, a_r$ with $I = P_1^{a_1} \cdots P_r^{a_r}$, and they are unique up to reordering.

Proof. (i) Choose J' such that $JJ' = (\alpha) \neq (0)$ is principal. Then $\alpha I = \alpha I'$, so $I = I'$.
(ii) Suppose $I \supset J$. Take I' such that $II' = (\alpha) \neq (0)$. Then $JI' \subset II' = (\alpha)$, so $H = \alpha^{-1}JI' \leq \mathcal{O}_K$ and we have $IH = J$.
(iii) For existence, we do induction on the norm $N(I)$ (again). The case $N(I) = 1$ is clear. For $N(I) > 1$, choose a prime $P \supset I$. Then $I = PJ$ for some $J \leq \mathcal{O}_K$ by (ii) and $J \supsetneq I$ by (i). We are then done by induction hypothesis.
As for uniqueness, suppose $I = P_1 \cdots P_r = Q_1 \cdots Q_s$ (but now P_i, Q_j don't have to be nonrepeating). Since $P_1 \supset I = Q_1 \cdots Q_s$, $P_1 = Q_j$ for some j by Lemma 8.2(i). WLOG $j = 1$. We can then use (i) to conclude that $P_2 \cdots P_r = Q_2 \cdots Q_s$. Induction finishes the proof. □

Definition 8.1. Suppose $I, J \leq \mathcal{O}_K$ are nonzero. We say I, J are equivalent ($I \sim J$) if $\alpha I = \beta J$ for some $\alpha, \beta \in \mathcal{O}_K \setminus \{0\}$.

It's clear that $\sim$ is an equivalence relation and I is principal iff $I \sim \mathcal{O}_K$.

Definition 8.2. The equivalence classes of $\sim$ are called ideal classes.

Theorem 8.6. *The set of ideal classes in $\mathcal{O}_K$ forms an abelian group under multiplication whose identity is the class of principal ideals.*

Definition 8.3. This group is called the ideal class group $\mathrm{Cl}(K)$ of K.

Proof. It's clear that multiplication is well-defined on $\mathrm{Cl}(K)$, is associatiev and commutative, and admits $[\mathcal{O}_K]$ as identity. Theorem 8.4 shows the existence of inverse. □

We will show later that $\mathrm{Cl}(K)$ is always finite.
In order to talk about $\mathrm{Cl}(K)$ more conveniently, it often helps to introduce the notion of a fractional ideal.

Definition 8.4. A fractional ideal in K is any subset of K of the form αI where $I \leq \mathcal{O}_K$ is a nonzero ideal and $\alpha \in K^\times$.

We can of course define addition and multiplication of fractional ideals in exactly the same way we did for (ordinary) ideals. Theorem 8.4 then gives

Theorem 8.7. *The set of fractional ideals is an abelian group under multiplication, and it's freely generated by the (nonzero) prime ideals.*
Furthermore, the principal fractional ideals (i.e. $\alpha\mathcal{O}_K, \alpha \in K^\times$) form a subgroup and its quotient is $\mathrm{Cl}(K)$.

Proposition 8.8. *The followings are equivalent:*
(i) $\mathcal{O}_K$ is a PID.
(ii) $\mathrm{Cl}(K)$ *is trivial.*
(iii) $\mathcal{O}_K$ is a UFD.

Remark. In general there are UFDs that are not PIDs, e.g. $\mathbb{Z}[T]$ is a UFD but $(2, T) \leq \mathbb{Z}[T]$ is not principal.

Proof. (i) $\iff$ (ii) and (i) $\implies$ (iii) are clear.
Suppose now that $\mathcal{O}_K$ is a UFD. It suffices to show that every prime ideal is principal. Let $\alpha \in P \setminus \{0\}$, then α factors uniquely (up to units) as $\alpha = \pi_1 \cdots \pi_R$ where each π_i is irreducible. We know that $\pi_i \in P$ for some i, but (π_i) is prime and $(0) \neq (\pi_i) \subset P \subset \mathcal{O}_K$, which means that $(\pi_i) = P$ since every prime ideal in $\mathcal{O}_K$ is maximal. □

Theorem 8.9. *If I, J are nonzero ideals, then $N(IJ) = N(I)N(J)$.*

This is highly nontrivial, as it is in fact equivalent to Theorem 8.5.

Remark. 1. If I, J are principal, then the theorem follows from the fact that $N((\alpha)) = |N_{K/\mathbb{Q}}(\alpha)|$.
2. If I, J are coprime, i.e. $I + J = \mathcal{O}_K$, then the theorem follows from Chinese Remainder Theorem ($R/(IJ) = R/(I \cap J) \cong (R/I) \times (R/J)$).

Proof. By unique factorisation, it suffices to show that $N(IP) = N(I)N(P)$ if P is prime. We also have $N(IP) = [\mathcal{O}_K : IP] = [\mathcal{O}_K : I][I : IP]$, so we need only to show that $N(P) = [I : IP]$.
We know $[I : IP] \neq 1$ by Theorem 8.5(i). For any ideal J with $I \supset J \supset IP$, Theorem 8.5(ii) yields $J' \leq \mathcal{O}_K$ with $J = IJ'$. $I \supset IJ' \supset IP$ then gives $\mathcal{O}_K \supset J' \supset P$ by Theorem 8.4. Thus $J' \in \{\mathcal{O}_K, P\}$ as P is maximal, i.e. either $J = I$ or $J = IP$.
Pick $\alpha \in I \setminus IP$, then the above means that $(\alpha) + IP = I$. Consider $A : \mathcal{O}_K/P \to I/IP$ via $x + P \mapsto \alpha x + IP$ which is a well-defined homomorphism of $\mathcal{O}_K/P$-vector spaces. A is surjective as $(\alpha) + IP = I$. But $\dim_{\mathcal{O}_K/P}(\mathcal{O}_K/P) = 1$ and $|I/IP| \neq 1$, so A would have to be bijective. Hence the result. □

Remark. If we consider $R = \mathbb{Z}[2\sqrt{2}]$ (which is not a ring of integers) and $P = (2, 2\sqrt{2})$, then $N(P) = 2$ but $P^2 = (4, 4\sqrt{2})$ has norm $8 \neq 2^2$.

9 Factorisation of Rational Primes

Recall that if p is an odd prime, then p is irreducible in $\mathbb{Z}[i]$ if $p \equiv 3 \pmod 4$ and we can factor $p = \pi\bar{\pi} = (x+iy)(x-iy)$ into irreducibles if $p \equiv 1 \pmod 4$. We want to generalise this to more general rings of integers.

Theorem 9.1. *Suppose $p \in \mathbb{Z}$ is a prime number ("rational prime"). Let $\{P_i : 1 \le i \le k\}$ be the prime ideals of $\mathcal{O}_K$ containing p. Write $N(P_i) = p^{f_i}$, then $(p) = P_1^{e_1} \cdots P_k^{e_k}$ for some $e_i \ge 1, \sum_i e_i f_i = n = [K : \mathbb{Q}]$.*

Proof. The existence of the factorisation is clear by Theorem 8.5. Theorem 8.9 then shows that $\sum_i e_i f_i = n$. □

The numbers $e_1, \ldots, e_k$ are called the ramification indices of p in K, and we'll soon prove that they are "usually" 1.

Definition 9.1. We say p is ramified (or p ramifies) if there is some i with $e_i > 1$.
p is inert if (p) is prime (i.e. $k = 1 = e_1 = 1, f_1 = n$).
p splits completely if $k = n$ (so $e_i = f_i = 1$ for all i).
p is totally ramified if $k = 1, e_1 = n$ (so $f_1 = 1$).

The factorisation of (p) is usually calculated with

Theorem 9.2 (Kummer-Dedekind Criterion)**.** *Suppose $K = \mathbb{Q}(\theta)$ for some $\theta \in \mathcal{O}_K$ with minimal polynomial $f \in \mathbb{Z}[T]$. Let p be a prime with $p \nmid [\mathcal{O}_K : \mathbb{Z}[\theta]]$ (e.g. if $\mathcal{O}_K = \mathbb{Z}[\theta]$).*
Suppose the reduction $\bar{f} \in \mathbb{F}_p[T]$ of f modulo p factors as $\bar{f} = \prod_{i=1}^k \bar{g}_i^{e_i}, e_i \ge 1$ where $\bar{g}_i \in \mathbb{F}_p[T]$ are distinct, monic and irreducible. Let $g_i \in \mathbb{Z}[T]$ be monic with reduction $\bar{g}_i$ modulo p. Then $(p) = \prod_{i=1}^k P_i^{e_i}$ where $P_i = (p, g_i(\theta)) \le \mathcal{O}_K$ are distinct primes in $\mathcal{O}_K$. Moreover, $N(P_i) = p^{f_i}$ with $f_i = \deg g_i$.

We'll use the third isomorphism theorem of rings: If $I \le J \le R$ are ideals, then $(R/I)/(J/I) \cong R/J$.

Proof. First assume $\mathcal{O}_K = \mathbb{Z}[\theta]$. Since $\bar{g}_i$ is irreducible,

$$\mathcal{O}_K/(P_i) = \mathbb{Z}[\theta]/(p, g_i(\theta)) \cong \mathbb{Z}[T]/(f, p, g_i) \cong \mathbb{F}_p[T]/(\bar{f}, \bar{g}_i) = \mathbb{F}_p[T]/(\bar{g}_i)$$

is a finite field with p^{f_i} elements. So P_i is prime and $N(P_i) = p^{f_i}$.
As $f = \prod_{i=1}^k g_i^{e_i} + ph$ for some $h \in \mathbb{Z}[T]$, we have

$$\begin{aligned}\prod_{i=1}^k P_i^{e_i} &= \prod_{i=1}^k (p, g_i(\theta))^{e_i} \subset \prod_{i=1}^n (p, g_i(\theta)^{e_i}) \\ &\subset \left(p, \prod_{i=1}^k g_i(\theta)^{e_i}\right) = (p, (f - ph)(\theta)) = (p)\end{aligned}$$

Taking norms, we obtain $\prod_i p^{e_i f_i} = \prod_i N(P_i)^{e_i} \ge N((p)) = p^n$, so $\sum_i e_i f_i \ge n$ with equality iff $(p) = \prod_i P_i^{e_i}$. But $n = \deg f = \sum_i e_i \deg \bar{g}_i = \sum_i e_i f_i$, so indeed $(p) = \prod_i P_i^{e_i}$.
The only thing left to show is that $P_i \ne P_i$ if $i \ne j$. We have $P_i + P_j = (p, g_i(\theta), g_j(\theta))$. As $\bar{g}_i, \bar{g}_j$ are coprime, there are $a, b \in \mathbb{Z}[T]$ with $ag_i + bg_j \equiv 1$

(mod p). This then means that $P_i + P_j = (1)$ which in particular means that $P_i \neq P_j$.
How about the general case? Let $Q_i = p\mathbb{Z}[\theta] + g_i(\theta)\mathbb{Z}[\theta] \leq \mathbb{Z}[\theta]$, then Q_i is a prime ideal of $\mathbb{Z}[\theta]$ and $\mathbb{Z}[\theta]/Q_i$ is a field with p^{f_i} elements like we did before. As $Q_i \subset P_i$, we can consider the ring homomorphism $\phi_i : \mathbb{Z}[\theta]/Q_i \to \mathcal{O}_K/P_i$ induced by inclusion. We know that $[\mathcal{O}_K : P_i] \mid [\mathcal{O}_K : (p)] = p^n$. But as $p \nmid [\mathcal{O}_K : \mathbb{Z}[\theta]]$, the image of ϕ_i also has index (essentially $[\mathcal{O}_K : P_i + \mathbb{Z}[\theta]]$) coprime to p, hence must be 1. As $\mathbb{Z}[\theta]/Q_i$ is field, any surjective ring homomorphism going out of is an isomorphism, so P_i is prime.
Its certainly still true that $\prod_i P_i^{e_i} = (p)$, and $\prod_i p^{e_i f_i} = \prod_i [\mathbb{Z}[\theta] : Q_i]^{e_i} = \prod_i [\mathcal{O}_K : P_i]^{e_i} \geq N((p)) = p^n$. We must then have equality everywhere. □

Example 9.1. Consider the case of a quadratic field $K = \mathbb{Q}(\sqrt{d})$ where $d \neq 0, 1$ is a square-free integer. We know that $\mathcal{O}_K = \mathbb{Z}[\sqrt{d}]$ or $\mathbb{Z}[(1+\sqrt{d})/2]$ with the latter happening iff $d \equiv 1 \pmod 4$. The minimal polynomial $f = T^2 - d$ of $\sqrt{d}$ factors modulo p as

$$\bar{f} = \begin{cases} (T-\bar{a})(T+\bar{a}) & \text{if } p \neq 2 \text{ and there is some } a \text{ with } a^2 \equiv d \pmod p \\ (T-\bar{d})^2 & \text{if } p = 2 \\ \bar{f} & \text{otherwise} \end{cases}$$

If p is an odd prime, then $p \nmid [\mathcal{O}_K : \mathbb{Z}[\sqrt{d}]]$. So by the preceding theorem, if d is not a square modulo p, then (p) is prime (p is inert); If $p \nmid d$ and $d \equiv a^2 \pmod p$ for some a, then $(p) = PP'$ where P, P' are distinct and $P = (p, a + \sqrt{d}), P' = (p, a - \sqrt{d})$ (p splits completely); If $p \mid d$, then $(p) = P^2$ where $P = (p, \sqrt{d}) = (p, \sqrt{d} - d)$ (p is totally ramified).
When $p = 2$, if $d \not\equiv 1 \pmod 4$, then $\mathcal{O}_K = \mathbb{Z}[\sqrt{d}]$. We have $T^2 - d \equiv (T+d)^2 \pmod 2$, so the preceding theorem shows that $(2) = P^2, P = (2, d + \sqrt{d})$.
If $d \equiv 1 \pmod 4$, then $\mathcal{O}_K = \mathbb{Z}[\theta]$ where $\theta = (1+\sqrt{d})/2$ and we are in trouble: $[\mathcal{O}_K : \mathbb{Z}[\sqrt{d}]] = 2$, so we can't use the theorem directly. What do we do? Well, we still have $K = \mathbb{Q}(\theta)$. The minimal polynomial of θ is $f(T) = T^2 - T - (d-1)/4$. So if $d \equiv 1 \pmod 8$, then $f(T) \equiv T(T+1) \pmod 2$ and therefore $(2) = PP'$ where $P = (2, \theta)$ and $P' = (2, \theta + 1)$. Otherwise $d \equiv 5 \pmod 8$, in which case $f(T) \equiv T^2 + T + 1 \pmod 2$ is irreducible modulo 2, so (2) is prime.

Remark. Recall that for quadratic fields $K = \mathbb{Q}(\sqrt{d})$ we had

$$d_K = \begin{cases} d & \text{for } d \equiv 1 \pmod 4 \\ 4d & \text{otherwise} \end{cases}$$

So indeed the prime factors of d_K are exactly those who ramifies. This generalises.

Theorem 9.3. *Let K be a number field and p a rational prime. If p ramifies in K, then $p \mid d_K$. In particular, only finitely many primes are ramified in K.*

Remark. The converse is also true, but we'll not prove that nor use that here.

Lemma 9.4. *If $\alpha \in \mathcal{O}_K$, then $\mathrm{Tr}_{K/\mathbb{Q}}(\alpha^p) \equiv \mathrm{Tr}_{K/\mathbb{Q}}(\alpha) \pmod p$.*

Proof. We know that $\mathrm{Tr}_{K/\mathbb{Q}}(\alpha)^p \equiv \mathrm{Tr}_{K/\mathbb{Q}}(\alpha) \pmod{p}$. Now,

$$\begin{aligned}\mathrm{Tr}_{K/\mathbb{Q}}(\alpha)^p - \mathrm{Tr}_{K/\mathbb{Q}}(\alpha^p) &= \left(\sum_{i=1}^n \sigma_i(\alpha)\right)^p - \sum_{i=1}^n \sigma_i(\alpha)^p \\ &= \sum_{k_1+\cdots+k_n=p, k_i<p} \frac{p!}{k_1!\cdots k_n!}\sigma_1(\alpha)^{k_1}\cdots\sigma_n(\alpha)^{k_n}\end{aligned}$$

This is an integer which is the product of p and an algebraic integer, hence an integer multiple of p. □

Proof of Theorem 9.3. Suppose $(p) = P_1^{e_1}\cdots P_r^{e_r}$ with, say, $e_1 > 1$. Let $\alpha \in P_1^{e_1-1}P_2^{e_2}\cdots P_r^{e_r} \setminus (p)$. Then for any $\beta \in \mathcal{O}_K$ we have $(\alpha\beta)^p \in (p)$ as $e_i > 1$ (so $p(e_1-1) \geq e_1$). By the preceding lemma, $\mathrm{Tr}_{K/\mathbb{Q}}(\alpha\beta) \equiv \mathrm{Tr}_{K/\mathbb{Q}}((\alpha\beta)^p) \equiv 0 \pmod{p}$.
Suppose $(\theta_i)_i$ is an integral basis and $\alpha = \sum_i a_i\theta_i, a_i \in \mathbb{Z}$. Since $\alpha \notin (p)$, not all a_i's are divisible by p. Then for any j,

$$\sum_{i=1}^n a_i \,\mathrm{Tr}_{K/\mathbb{Q}}(\theta_i\theta_j) = \mathrm{Tr}_{K/\mathbb{Q}}(\alpha\theta_j) \equiv 0 \pmod{p}$$

So the rows of $(\mathrm{Tr}_{K/\mathbb{Q}}(\theta_i\theta_j))_{i,j}$ are linearly dependent modulo p, which in particular means that $p \mid d_K$. □

This can sometimes help in determining the ring of integers.

Example 9.2. Suppose $K = \mathbb{Q}(\theta)$ where $\theta = \sqrt[3]{p}$ for a prime $p \neq 3$. Clearly $\mathbb{Z}[\theta] \subset \mathcal{O}_K$ and $(p) = (\theta)^3$. So $P = (\theta)$ must be prime and $e_1 = 3$ (p is totally ramified). By Theorem 9.3, we have $p \mid d_K = \mathrm{Disc}(\mathbb{Z}[\theta])/[\mathcal{O}_K : \mathbb{Z}[\theta]]^2$. But

$$\mathrm{Disc}(\mathbb{Z}[\theta]) = \det \mathrm{Tr}_{K/\mathbb{Q}} \begin{pmatrix} 1 & \theta & \theta^2 \\ \theta & p & p\theta \\ \theta^2 & p\theta & p\theta^2 \end{pmatrix} = \det \begin{pmatrix} 3 & 0 & 0 \\ 0 & 0 & 3p \\ 0 & 3p & 0 \end{pmatrix} = -27p^2$$

which means that we must have $[\mathcal{O}_K : \mathbb{Z}[\theta]] \in \{1, 3\}$.

10 Geometry of Numbers

The beauty of algebraic number theory is that not only does algebra help, but we also get to use all sort of geometric tools. In this section, we'll explore one of such tricks, namely to deduce results from the geometry of the real/complex numbers via the embeddings $K \hookrightarrow \mathbb{C}$. Using such techniques, we will be able to show that $\mathrm{Cl}(K)$ is finite and we'll also provide a way to compute it. Another related theorem is that the group of units $\mathcal{O}_K^\times$ is finitely generated and has rank $r + s - 1$ (where as usual $r + 2s = n$ and r is the number of real embeddings of K).
$(\sigma_i)_{i=1}^n$ embeds K into Euclidean space. We want to investigate the image of $\mathcal{O}_K$ under this embedding. For example, the embedding $\mathbb{Q}(i) \hookrightarrow \mathbb{C} \cong \mathbb{R}^n$ maps $\mathbb{Z}[i]$ to the lattice $\mathbb{Z}^2 \leq \mathbb{R}^2$.

Definition 10.1. A lattice in $\mathbb{R}^n$ is a subgroup of $(\mathbb{R}^n, +)$ generated by n linearly independent elements of $\mathbb{R}^n$

Example 10.1. $\mathbb{Z}^n \leq \mathbb{R}^n$ is a lattice.

In general, if $e_1, \ldots, e_n \in \mathbb{R}^n$ are linearly independent, we can form the lattice $\Lambda = \mathbb{Z}e_1 + \cdots + \mathbb{Z}e_n$. The fundamental parallelopiped attached to the basis $(e_i)_i$ is the set $\mathscr{P} = \{\sum_i x_i e_i : 0 \leq x_i < 1\}$.

Definition 10.2. The covolume $\operatorname{covol}(\Lambda)$ of Λ is defined as the volume $\operatorname{vol}(\mathscr{P})$ of the parallelopiped $\mathscr{P}$, which equals $|\det(e_{ij})|$ where $e_i = (e_{ij})_j$.

This is independent of the choice of basis $(e_i)_i$ generating Λ.

Remark. 1. Suppose we have a finite index subgroup $\Lambda' \leq \Lambda$, then Λ' is also a lattice and its covolume is $\operatorname{covol}(\Lambda') = [\Lambda : \Lambda'] \operatorname{covol}(\Lambda)$.
2. The same definition extends to complex vector spaces by identifying $\mathbb{C} \cong \mathbb{R}^2$.

Example 10.2. Suppose $K = \mathbb{Q}(\sqrt{-d})$ (an "imaginary quadratic field") where $d > 0$ is square-free. It has the embedding $\sigma(a + b\sqrt{-d}) = a + b\sqrt{d}i$ into $\mathbb{C}$. Then $1, \sigma(\theta)$ constitute a basis for $\sigma(\mathcal{O}_K)$ where $\theta = (1 + \sqrt{-d})/2$ if $-d \equiv 1 \pmod 4$ and $\theta = \sqrt{-d}$ otherwise. The former looks like a slanted lattice and the latter a rectangular lattice. We have $\operatorname{covol}(\sigma(\mathcal{O}_K)) = \sqrt{d}/2$ in the first case and $\operatorname{covol}(\sigma(\mathcal{O}_K)) = \sqrt{d}$ in the second case. Either way, we have $\operatorname{covol}(\sigma(\mathcal{O}_K)) = |d_K|^{1/2}/2$.

Theorem 10.1 (Special Case of Minkowski's Theorem)**.** *Let $\Lambda \subset \mathbb{C}$ be a lattice and $X = \{z \in \mathbb{C} : |z|^2 \leq R\}$. If $\pi R \geq 4 \operatorname{covol} \Lambda$, then $\Lambda \cap X \neq \{0\}$.*

This can be used to show a cool number theoretic result.

Theorem 10.2. *Let $K = \mathbb{Q}(\sqrt{-d})$ (where as usual $d > 0$ is square-free) and $I \subset \mathcal{O}_K$ a nonzero ideal. Then there exists some $\alpha \in I \setminus \{0\}$ such that $N_{K/\mathbb{Q}}(\alpha) \leq c_K N(I)$ where $c_K = (2/\pi)|d_K|^{1/2}$.*

Proof. Consider the lattice $\sigma(I) \subset \sigma(\mathcal{O}_K) \subset \mathbb{C}$ which is a lattice with covolume $\operatorname{covol}(I) = N(I) \operatorname{covol}(\sigma(\mathcal{O}_K)) = (N(I)/2)|d_K|^{1/2}$. Take X as in the preceding theorem. Then we know that $X \cap \sigma(I) \neq \{0\}$ whenever $\pi R \geq 4 \operatorname{covol}(\sigma(I)) = 2N(I)|d_K|^{1/2}$. $R = c_K N(I) = (2N(I)/\pi)|d_K|^{1/2}$ satisfies this inequality, so there is some $\alpha \in I \setminus \{0\}$ with $N_{K/\mathbb{Q}}(\alpha) = |\sigma(\alpha)|^2 \leq R = c_K N(I)$. □

Theorem 10.3. *Let $K = \mathbb{Q}(\sqrt{-d})$ ($d > 1$ square-free). Then:*
(i) $\operatorname{Cl}(K)$ is finite.
(ii) Every ideal class of K contains an ideal of norm at most c_K.
(iii) $\operatorname{Cl}(K)$ is generated by classes of prime ideals of norm at most c_K.

Proof. (i) Follows from (ii) by Corollary 7.2.
(ii) Take $I \leq \mathcal{O}_K$ nonzero. First choose J such that $IJ = (\beta), \beta \in \mathcal{O}_K \setminus \{0\}$ is principal. Take $\alpha \in J \setminus \{0\}$ such that $N_{K/\mathbb{Q}}(\alpha) \leq c_K N(J)$, which is possible by the preceding theorem. Take I' such that $I'J = (\alpha)$ (since $\alpha \in J$), then $\alpha IJ = (\alpha\beta) = \beta I'J$, so I, I' are in the same ideal class. But $N(I') = N((\alpha))/N(J) = N_{K/\mathbb{Q}}(\alpha)/N(J) \leq c_K$.
(iii) Follows from (ii) and Theorem 8.5. □

Example 10.3. 1. Take $K = \mathbb{Q}(i)$. Then $d_K = -4$ and so every ideal is equivalent to an ideal I with $N(I) \leq c_K = 4/\pi < 2$, so $I = \mathcal{O}_K$ and hence every ideal is principal.

2. Take $K = \mathbb{Q}(\sqrt{-5})$, then $\mathcal{O}_K = \mathbb{Z}[\sqrt{-5}]$ is not a PID, so the class group is nontrivial. $d_K = -20$, so $c_K = 4\sqrt{5}/\pi < 3$. That is, every ideal class contains an ideal of norm at most 2. If the ideal has norm 1 then we arrive at the principal ideal class. If the ideal has norm 2, we note that $(2) = P^2, P = (2, 1 + \sqrt{5})$ which forces P to be the only ideal of norm 2. Thus $\mathrm{Cl}(K) = \{[\mathcal{O}_K], [P]\}$. This computation works for imaginary quadratic fields in general.

Let's now turn to the full version of Minkowski's theorem.

Definition 10.3. A subset $X \subset \mathbb{R}^n$ is convex if $tx + (1-t)y \in X$ for any $x, y \in X, t \in [0, 1]$. It's symmetric if $x \in X \implies -x \in X$.

Theorem 10.4 (Minkowski's Theorem)**.** *Let $\Lambda \leq \mathbb{R}^n$ be a lattice and $X \subset \mathbb{R}^n$ a (measurable) convex symmetric subset. If either $\mathrm{vol}(X) > 2^n \,\mathrm{covol}(\Lambda)$ or $\mathrm{vol}(X) \geq 2^n \,\mathrm{covol}(\Lambda)$ and X is compact, then $X \cap \Lambda \neq \{0\}$.*

Lemma 10.5 (Blichfeldt's Lemma)**.** *Suppose $\Lambda \leq \mathbb{R}^n$ is a lattice and $Y \subset \mathbb{R}^n$ measurable. If $\mathrm{vol}(Y) > \mathrm{covol}(\Lambda)$, then there exists distinct $x, y \in Y$ with $x - y \in \Lambda$.*

Proof. Choose a basis for Λ and let $\mathscr{P}$ be the corresponding fundamental parallelopiped. Then $\mathbb{R}^n = \coprod_{\lambda \in \Lambda} (\lambda + \mathscr{P})$ and so $Y = \coprod_\lambda Y_\lambda$ where $Y_\lambda = Y \cap (\lambda + \mathscr{P})$ and $\mathrm{vol}(Y) = \sum_\lambda \mathrm{vol}(Y_\lambda)$. We have $\sum_\lambda \mathrm{vol}(-\lambda + Y_\lambda) = \sum_\lambda \mathrm{vol}(Y_\lambda) = \mathrm{vol}(Y) > \mathrm{vol}(\mathscr{P})$. However, $-\lambda + Y_\lambda \subset \mathscr{P}$ for all λ. So there must exist distinct λ, μ such that $(-\lambda + Y_\lambda) \cap (-\mu + Y_\mu) \neq \varnothing$. Take $z \in (-\lambda + Y_\lambda) \cap (-\mu + Y_\mu)$. Then $x = z + \lambda, y = z + \mu$ satisfies the claim. □

Proof of Theorem 10.4. Note that $2^n \,\mathrm{covol}(\Lambda) = \mathrm{covol}(2\Lambda)$.
Suppose $\mathrm{vol}(X) > 2^n \,\mathrm{covol}(\Lambda)$. By the preceding lemma, there are $x, y \in X, x \neq y$ with $x - y \in 2\Lambda$. As X is symmetric we have $-y \in X$; As X is convex we have $0 \neq (x-y)/2 \in X \cap \Lambda$, in particular $X \cap \Lambda \neq \{0\}$.
Now suppose that $\mathrm{vol}(X) = 2^n \,\mathrm{covol}(\Lambda)$ and X is assumed to be compact. For $\delta > 0$, let $X_\delta = (1+\delta)X$, then $\mathrm{vol}(X_\delta) > 2^n \,\mathrm{covol}(\Lambda)$. Then $X_\delta \neq \{0\}$ by above discussion. As X is compact, X_δ are bounded, so $X_\delta \cap \Lambda$ is finite. X is also closed, so $X \cap \Lambda = \bigcap_{\delta > 0} X_\delta \cap \Lambda$, which then has to be equal to $X_{\delta'} \cap \Lambda$ for some $\delta' > 0$, and we are done. □

We've seen that a special case of Theorem 10.4 implies the finiteness of $\mathrm{Cl}(K)$ for K an imaginary quadratic field. This is true for a general number field K. Let $n = [K : \mathbb{Q}] = r + 2s$ as usual. Write $\sigma_j : K \hookrightarrow \mathbb{R}, 1 \leq j \leq r$ and $\sigma_j, \bar{\sigma}_j : K \hookrightarrow \mathbb{R}, r < j \leq r + s$ to denote the embeddings. The product of them is a injective map $\sigma : K \hookrightarrow \mathbb{R}^r \times \mathbb{C}^s \cong \mathbb{R}^n$ by identifying $\mathbb{C} \cong \mathbb{R}^2$ using the usual basis $\{1, i\}$.

Proposition 10.6. *$\sigma(\mathcal{O}_K)$ is a lattice in $\mathbb{R}^n$ with covolume $2^{-s}|d_K|^{1/2}$.*

Proof. Let $\omega_1, \ldots, \omega_n$ be an integral basis for K, then $e_i = \sigma(\omega_i) \in \mathbb{R}^n$ is the vector

$$\begin{aligned} e_i = (&\sigma_1(\omega_i), \ldots, \sigma_r(\omega_i), \\ &\mathrm{Re}\,\sigma_{r+1}(\omega_i), \mathrm{Im}\,\sigma_{r+1}(\omega_i), \ldots, \mathrm{Re}\,\sigma_{r+s}(\omega_i), \mathrm{Im}\,\sigma_{r+s}(\omega_i)) \end{aligned}$$

For $r < j \leq r+s$, we have

$$(\operatorname{Re}\sigma_j(\alpha), \operatorname{Im}\sigma_j(\alpha))\begin{pmatrix}1 & 1\\ i & -i\end{pmatrix} = (\sigma_j(\alpha), \overline{\sigma_j(\alpha)})$$

So $\det e_{ij} = \pm(1/(2i))^s \det(\sigma_i(\omega_i))$ and $\det(\sigma_j(\omega_i))^2 = d_K \neq 0$. This means that $(e_i)_i$ is a basis for $\mathbb{R}^n$ and also a $\mathbb{Z}$-basis for $\sigma(\mathcal{O}_K)$, so $\sigma(\mathcal{O}_K)$ is a lattice with covolume $|\det e_{ij}| = 2^{-s}|d_K|^{1/2}$. □

Pretty much the same argument shows that

Corollary 10.7. *Suppose $I \leq \mathcal{O}_K$ is a nonzero ideal, then $\sigma(I) \subset \mathbb{R}^n$ is a lattice of covolume $2^{-s}|\operatorname{disc}(I)|^{1/2} = 2^{-s}N(I)|d_K|^{1/2}$.*

Theorem 10.8. *Suppose $I \leq \mathcal{O}_K$ is a nonzero ideal, then there is some $0 \neq \alpha \in I$ with $|N_{K/\mathbb{Q}}(\alpha)| \leq c_K N(I)$ where*

$$c_K = \left(\frac{4}{\pi}\right)^s \frac{n!}{n^n}|d_K|^{1/2}$$

is called the Minkowski constant for K.

Corollary 10.9. *Every ideal class of K contains an ideal of norm at most c_K. In particular, $\mathrm{Cl}(K)$ is finite and is generated by the classes of prime ideals of norm at most c_K.*

Before proving Theorem 10.8 in full generality, let's first look at the real quadratic case.

Proof of Theorem 10.8 for real quadratic fields. Suppose $K = \mathbb{Q}(\sqrt{d})$ for some square-free $d > 1$. Then $\sigma : K \hookrightarrow \mathbb{R}^2$ is given by the map $u + v\sqrt{d} \mapsto (u + v\sqrt{d}, u - v\sqrt{d})$. For $\alpha = u+v\sqrt{d}$, we have $N_{K/\mathbb{Q}}(\alpha) = u^2 - dv^2 = \sigma_1(\alpha)\sigma_2(\alpha)$. So $|N_{K/\mathbb{Q}}(\alpha)| \leq R$ iff $\sigma(\alpha)$ lies in the region bounded by the hyperbola $x_1x_2 = \pm R$. We need a convex symmetric $X \subset \mathbb{R}^n$ contained in this region that's as large as possible. Let's use the square with vertices $(\pm 2R^{1/2}, 0), (0, \pm 2R^{1/2})$, whose area is $8R$. Theorem 10.4 shows that if $8R \geq 4\operatorname{covol}(\sigma(I)) = 4|d_K|^{1/2}N(I)$, then there is a nonzero $\alpha \in I \setminus \{0\}$ with $\sigma(\alpha) \subset X$. In other words, there is a nonzero $\alpha \in I$ with $|N_{K/\mathbb{Q}}(\alpha)| \leq c_K N(I)$. □

Motivated by this, we are now ready to prove the theorem in full

Proof of Theorem 10.8. Consider

$$X_R = \left\{(x_1, \ldots, x_r, z_1, \ldots, z_s) \in \mathbb{R}^r \times \mathbb{C}^s : \sum_{j=1}^{r} |x_j| + 2\sum_{j=r+1}^{r+s} |z_j| \leq nR^{1/n}\right\}$$

which is clearly convex and symmetric. We have $|N_{K/\mathbb{Q}}(\alpha)| \leq R$ whenever $\sigma(\alpha) \in X_R$ by AM-GM.
Integration reveals that the volume of X_R is $2^n(\pi/2)^s(n!/n^n)R$ (example sheet). So by Theorem 10.4, there is $\alpha \in I \setminus \{0\}$ with $|N_{K/\mathbb{Q}}(\alpha)| \leq R$ when $\operatorname{vol}(X_R) \geq 2^n \operatorname{covol}(\sigma(I)) = 2^n 2^{-s}|d_K|^{1/2}N(I)$, which holds when $R \geq c_K N(I)$. □

In general, the factor $n!/n^n$ in c_K goes to 0 pretty quickly as n grows large, so it's a significant gain.

Example 10.4. For $K = \mathbb{Q}(\sqrt{-17}), d_K = -68$, so $c_K = 2\sqrt{68}/\pi < 6$. So the class group is generated by the classes of prime ideals of norm at most 5. One can check using Theorem 9.2 that (5) is prime, $(3) = P_3P_3', P_3 = (3, 1+\sqrt{-17}), P_3' = (3, 1-\sqrt{-17})$ and $(2) = P_2^2, P_2 = (2, 1+\sqrt{-17})$. We have $[P_2]^2 = 1 = [P_3][P_3']$ and $[P_3]^2 = [P_3^2] = [(9, 1+\sqrt{-17})]$. As $N_{K/\mathbb{Q}}(1+\sqrt{-17}) = 18, N((9, 1+\sqrt{-17})) = 9$, we have $(1+\sqrt{-17}) = P_3^2 I$ for some I with $N(I) = 2$. Then $I = P_2$, thus $[P_3]^2 = [P_2]^{-1} = [P_2]$. Now P_2 is clearly not principal ($x^2 + 17y^2 = 2$ has no integer solutions (x, y)). So $\mathrm{Cl}(K) \cong \mathbb{Z}/4\mathbb{Z}$ with $[P_3]$ as a generator.

Example 10.5. Consider the quintic field $K = \mathbb{Q}(\theta)$ where θ is a root of the irreducible quintic $g(T) = T^5 - T + 1$. As g has only one real root, $r = 1, s = 2$. Also, $\mathrm{disc}(g) = 2689 = 19 \times 151$ which is square-free. So $\mathcal{O}_K = \mathbb{Z}[\theta]$, and $c_K = (4/\pi)^2(5!/5^5)\sqrt{2689} < 4$. This means that $\mathrm{Cl}(K)$ is generated by classes of prime ideals of norm 2 or 3. But g has no root modulo 2 or 3, so by Theorem 9.2 there is no prime ideals of norm $2, 3$, forcing $\mathrm{Cl}(K)$ to be trivial.

How large can the class group get as K varies?
For imaginary quadratics $K = \mathbb{Q}(\sqrt{-d})$, it's known that $|\mathrm{Cl}(K)| \to \infty$ as $d \to \infty$. We also know that $\mathrm{Cl}(K)$ is nontrivial whenever $d > 163$.
On the other hand, if $K = \mathbb{Q}(\sqrt{d})$ is a real quadratic field, although $c_K \to \infty$ as $d \to \infty$, it was conjectured (by Gauss) that $\mathrm{Cl}(K)$ is trivial for infinitely many d. This is still open.

Example 10.6. Take $K = \mathbb{Q}(\sqrt{10})$ which has $c_K = (1/2)\sqrt{40} < 4$. So $\mathrm{Cl}(K)$ is generated by the classes of prime ideals of norm 2 or 3. We have $(2) = P_2^2, P_2 = (2, \sqrt{10}), (3) = P_3P_3', P_3 = (3, 1+\sqrt{10})(3, 1-\sqrt{10})$. Thus $[P_2]^2 = 1 = [P_3][P_3]'$. As usual, we can get other relations between them by factoring some small elements of $\mathcal{O}_K$. $N_{K/\mathbb{Q}}(1+\sqrt{10}) = -9, P_3 \mid (1+\sqrt{10})$, but $3 \nmid 1+\sqrt{10}$, so $(1+\sqrt{10}) = P_3^2$ and thus P_3^2 is principal. Also, $2-\sqrt{10} = 3-(1+\sqrt{10}) \in P_3, N_{K/\mathbb{Q}}(2-\sqrt{10}) = 4-10 = -6$. We also have $2-\sqrt{10} \in P_2$, so $(2-\sqrt{10}) = P_2P_3$ as $(2-\sqrt{10})$ has norm 6. Hence $[P_2] = [P_3] = [P_3']$ and it has order either 1 or 2.
If P_2 were principal, say $P_2 = u + v\sqrt{10}$, then $u^2 - 10v^2 = \pm 2$ which has no integer solution by taking mod 5. So $\mathrm{Cl}(K) \cong \mathbb{Z}/2\mathbb{Z}$ and is generated by $[P_2] = [P_3] = [P_3']$.

Definition 10.4. $h_K = |\mathrm{Cl}(K)|$ is called the class number of K.

11 Units

We will show the following theorem:

Theorem 11.1 (Dirichlet's Unit Theorem)**.** *The group $\mathcal{O}_K^\times$ is finitely generated and has rank $r+s-1$. Furthermore, the torsion subgroup of $\mathcal{O}_K^\times$ is the group of roots of unity in K.*
Equivalently, there exists $\epsilon_1, \ldots, \epsilon_{r+s-1} \in \mathcal{O}_K^\times$ such that every unit can be uniquely written in the form $\zeta\epsilon_1^{m_1} \cdots \epsilon_{r+s-1}^{m_{r+s-1}}$ with $m_i \in \mathbb{Z}$ and ζ a root of unity in K.

Example 11.1. $r+s-1 = 0$ iff $(r, s) = (1, 0)$ or $(r, s) = (0, 1)$. The former case is just $K = \mathbb{Q}, \mathcal{O}_K^\times = \{\pm 1\}$. The latter happens precisely when $K = \mathbb{Q}(\sqrt{-d})$ is

an imaginary quadratic, so $\mathcal{O}_K^\times$ consists of those $u+v\sqrt{d}$ with $u^2+dv^2=1$. This is clearly a finite collection. More precisely, when $d=-1$ we have $K=\mathbb{Q}(i)$ and $\mathcal{O}_K^\times=\{\pm 1,\pm i\}$; When $d=-3$ we have $K=\mathbb{Q}(\omega)$ where $\omega=(-1+\sqrt{-3})/2$ and $\mathcal{O}_K^\times=\{\pm 1,\pm\omega,\pm\omega^2\}$; Otherwise, $\mathcal{O}_K^\times=\{\pm 1\}$.

Theorem 11.2. *Let $K=\mathbb{Q}(\sqrt{d})\subset\mathbb{R}$ be a real quadratic field, then:*
(i) $\mathcal{O}_K^\times$ is infinite.
(ii) There exists a smallest unit $\epsilon>1$ (the "fundamental unit") and $\mathcal{O}_K^\times=\{\pm\epsilon^m : m\in\mathbb{Z}\}$.
(iii) A unit $u+v\sqrt{d}$ is the fundamental unit iff $u,v>0$ and v is minimal.

Consequently $\mathcal{O}_K\cong(\mathbb{Z}/2\mathbb{Z})\times\mathbb{Z}$.

Proof. (i) Recall that $\alpha\in\mathcal{O}_K$ is a unit iff $N_{K/\mathbb{Q}}(\alpha)\in\{\pm 1\}$, so $u+v\sqrt{d}\in\mathcal{O}_K$ is a unit iff $u^2-dv^2=\pm 1$ – this is just Pell's equation! So we know that this has infinitely many solution $(u,v)\in\mathbb{Z}^2$, which means that $\mathcal{O}_K$ is infinite.
(ii) As $K\subset\mathbb{R}$, the only roots of unity in K are $\{\pm 1\}$. By (i) there is some unit $\epsilon=u+v\sqrt{d}$ that's not ± 1. We claim that $\epsilon>1$ iff $u,v>0$. Indeed, the four numbers $\{\pm u\pm v\sqrt{d}\}=\{\pm\epsilon,\pm\epsilon^{-1}\}$ are all distinct units. It's then clear that no two of them can lie in the same interval amongst $(-\infty,-1),(-1,0),(0,1),(1,\infty)$, hence the claim.
This means that we can indeed choose a smallest $\epsilon>1$. If $\epsilon'=u'+v'\sqrt{d}\in\mathcal{O}_K^\times$ has $\epsilon'>1$, then there is a unique $m\geq 1$ such that $\epsilon^m\leq\epsilon'<\epsilon^{m+1}$, i.e. $1\leq\epsilon'/\epsilon^m<\epsilon$, so $\epsilon'=\epsilon^m$. So $\mathcal{O}_K^\times\cap(1,\infty)=\{\epsilon^m : m\geq 1\}$, which implies that $\mathcal{O}_K^\times=\{\pm\epsilon^m : m\in\mathbb{Z}\}$.
(iii) The units with $u,v>0$ have the form $\epsilon^m=u_m+v_m\sqrt{d}, m\geq 1$. Binomial expansion reveals that $v_m\geq mu_1^{m-1}v_1>v_1$ whenever $m>1$. □

There is another way to see (i), which can be more easily generalised.

Proposition 11.3. *Suppose K is a real quadratic field and $R\geq|d_K|^{1/2}$, then there are infinitely many $\alpha\in\mathcal{O}_K$ with $|N_{K/\mathbb{Q}}(\alpha)|\leq R$.*

Pick any such R. There are finitely many ideals with norm at most R, so there exists an ideal I with norm at most R and infinitely many $(\alpha_i)_{i\in\mathbb{N}}\in\mathcal{O}_K$ with $I=(\alpha_i)$ for every i. Then $\alpha_i/\alpha_0\in\mathcal{O}_K^\times$ for each i, thus $\mathcal{O}_K^\times$ is infinite.

Proof. Take the embedding $\sigma: K\hookrightarrow\mathbb{R}^2, x+y\sqrt{d}\mapsto(x+y\sqrt{d},x-y\sqrt{d})$. For $\delta>0$, let K_δ be the rectangle $[-\delta,\delta]\times[-R/\delta,R/\delta]$, which lies inside the region $\{(x_1,x_2)\in\mathbb{R}^2 : |x_1x_2|\leq R\}$. Let $\delta=\delta_0=1$, say. Then $\mathrm{vol}(K_\delta)=4R\geq 4|d_K|^{1/2}\geq 4\,\mathrm{covol}(\sigma\mathcal{O}_K)$. Theorem 10.4 then shows that there is $\alpha_0\in\mathcal{O}_K\setminus\{0\}$ with $\sigma\alpha_0\in K_\delta$, which means that $|N_{K/\mathbb{Q}}(\alpha_0)|\leq R$. But we can always choose δ_1 such that $0<\delta_1<|(\sigma\alpha_0)_1|$ and repeat the procedure. Since $\sigma\alpha_0\notin K_{\delta_1}$, this gives some $\alpha_1\neq\alpha_0$ with $|N_{K/\mathbb{Q}}(\alpha_1)|\leq R$. Doing this inductively gives the result. □

Now consider general K. Let's first show that $\mathcal{O}_K^\times$ is finitely generated with rank at most $r+s-1$.

Lemma 11.4. *For any $C>1$, $\{\alpha\in\mathcal{O}_K : \forall i, |\sigma_i\alpha|\leq C\}$ is finite.*

Proof. The characteristic polynomial of α has the form $\prod_i (T - \sigma_i(\alpha)) = T^n + \sum_{r=1}^n c_r T^{n-r}$ where we have

$$|c_r| = \left| (-1)^r \sum_{i_1 < i_2 < \cdots < i_r} \sigma_{i_1}(\alpha) \cdots \sigma_{i_r}(\alpha) \right| \leq \binom{n}{r} C^r$$

As $c_r \in \mathbb{Z}$, there are only finitely many possibilities for them, hence for this polynomial. Consequently there are only finitely many possible choices of α. □

Proposition 11.5. *The group of roots of unity in K is finite (and hence cyclic).*

Proof. Follows immediately from the preceding lemma with $C = 1$. □

To show that $\mathcal{O}_K^\times$ is finitely generated, we use the logarithm trick.

Definition 11.1. The logarithmic map of K is $\mathcal{L} : K^\times \to \mathbb{R}^{r+s}$ given by

$$\mathcal{L}\alpha = (\log|\sigma_1(\alpha)|, \ldots, \log|\sigma_r(\alpha)|, 2\log|\sigma_{r+1}(\alpha)|, \ldots, 2\log|\sigma_{r+s}(\alpha)|)$$

Then $\mathcal{L}$ is a group homomorphism and $\sum_i (\mathcal{L}\alpha)_i = \log|N_{K/\mathbb{Q}}(\alpha)|$. In particular, $\epsilon \in \mathcal{O}_K^\times$ iff $N_{K/\mathbb{Q}}(\alpha) \in \{\pm 1\}$ iff $\mathcal{L}\epsilon$ lies in the hyperplane $\mathbb{R}^{r+s,0} = \{y \in \mathbb{R}^{r+s} : \sum_i y_i = 0\}$. If ϵ is a root of unity, then $|\sigma_i \epsilon| = 1$ for all i, so $\mathcal{L}\epsilon = 0$. The converse is also true.

Proposition 11.6. *(i)* $(\ker \mathcal{L}) \cap \mathcal{O}_K^\times = \{\zeta \in \mathcal{O}_K : \exists n \in \mathbb{Z}_{>0}, \zeta^n = 1\}$.
(ii) $\mathcal{L}\mathcal{O}_K^\times$ is a free abelian group of rank at most $r + s - 1$.

Proof. (i) Take $M > 0$ and let $Z = Z_M = \{|y_i| \leq M\} \subset \mathbb{R}^{r+s}$. Whenever $\mathcal{L}\alpha \in Z$ we have $|\sigma_i \alpha| \leq e^M$ if $1 \leq i \leq r$ and $|\sigma_i \alpha| \leq e^{M/2}$ if $i > r$. Lemma 11.4 implies that $\{\alpha \in \mathcal{O}_K^\times : \mathcal{L}\alpha \in Z\}$ is finite. As this contains $(\ker \mathcal{L}) \cap \mathcal{O}_K^\times$, we conclude the result.
(ii) Follows from above discussion and the next lemma. □

Lemma 11.7. *Let $L \subset \mathbb{R}^m$ be a subgroup such that $L \cap Z$ is finite for any bounded $Z \subset \mathbb{R}^m$, then L is free abelian of rank $d = \dim(\mathrm{Span}_{\mathbb{R}} L) \leq m$.*

Proof. Let $d = \dim(\mathrm{Span}_{\mathbb{R}} L)$. By replacing $\mathbb{R}^m$ by subspace spanned by L and changing basis (which preserve boundedness), we can assume WLOG that $m = d$ and $L \supset \mathbb{Z}^d$. Proposition 5.2 tells us that it suffices to show that L is finitely generated.
Every $x \in L$ can be written in the form $x = a + z$ where $a \in \mathbb{Z}^d$ and $z \in [0,1)^d \cap L$. But $[0,1)^d \cap L$ is bounded, so there are only finitely many possibilities for z, so L is finitely generated. □

Remark. To complete the proof of Theorem 11.1, we just need to find $r + s - 1$ linearly independent unit. The way of doing it is to mimic the proof in the case of real quadratic fields.

12 Diophantine Equations

An diophantine equation is a polynomial equation, usually in multiple variables, for which we look for integer solutions. One famous example of this is Fermat's equation $x^d + y^d = z^d$. Algebraic number theory was basically invented to deal with these kind of problems, and are often helpful.

Example 12.1. Say you're interested in the integer solutions to the equation $y^2 + 5 = x^3$. One can spot immediately that x is odd (by taking modulo 4) and not divisible by 5 (by looking at the 5-adic valuation).
Let's do some elementary number theory. The natural thing to do is to write down the factorisation $(y+\sqrt{-5})(y-\sqrt{-5}) = x^3$ in $\mathbb{Z}[\sqrt{-5}]$, the ring of integers of $\mathbb{Q}(\sqrt{-5})]$.
We'd quite like to use a unique factorisation argument, which albeit tempting does not work directly, as $\mathbb{Z}[\sqrt{-5}]$ is not a UFD. We can, however, deduce the equality for ideals $(x)^3 = (y + \sqrt{-5})(y - \sqrt{-5})$.
Suppose P is a prime ideal of $\mathbb{Z}[\sqrt{-5}]$ dividing both $(y + \sqrt{-5})$ and $(y - \sqrt{-5})$, then P divides $(y + \sqrt{-5}) + (y - \sqrt{-5}) \supset (2\sqrt{-5}) \supset 10$ and also (x). But x is coprime to 10, so such P does not exist. Consequently, Theorem 8.5 tells us that there are ideals I, J such that $(y + \sqrt{-5}) = I^3, (y - \sqrt{-5}) = J^3$ and $IJ = (x)$.
But $\mathrm{Cl}(K)$ has order 2, so I is principal. Suppose $I = (a+b\sqrt{-5})$, then $y+\sqrt{-5}$ is $u(a + b\sqrt{-5})^3$ for some unit u. The only units of $\mathbb{Z}[\sqrt{-5}]$ are ± 1, so after possibly replacing a, b by $-a, -b$ we can write down $y + \sqrt{-5} = (a + b\sqrt{-5})^3 = (a^3 - 15ab^2) + (3a^2b - 5b^3)\sqrt{-5}$. In particular, $3a^2b - 5b^3 = 1$, so $b \in \{\pm 1\}$ and hence $3a^2 - 5 = \pm 1$, impossible.

13 Analytic Class Number Formula

For a number field K, the class number $h_K = |\mathrm{Cl}(K)|$ is not at all trivial to compute. One of the formulas to compute this is called the analytic class number formula. Recall that the Riemann ζ functions $\zeta(s) = \sum_n n^{-s}$ tells you a lot about the distribution of primes. This series converges for $\mathrm{Re}\, s > 1$, diverges at $s = 1$, and in fact

Proposition 13.1. *$(s - 1)\zeta(s) \to 1$ as $s \to 1^+$.*

Proof. x^{-s} is monotone, so for every $s > 1$ we get to write down

$$\int_n^{n+1} \frac{\mathrm{d}x}{x^s} \leq \frac{1}{n^s} \leq \int_{n-1}^{n} \frac{\mathrm{d}x}{x^s}$$

Summing everything gives $(s - 1)^{-1} \leq \zeta(s) \leq (s - 1)^{-1} + 1$. □

We are not gonna use this, but we are gonna generalise this to another L-function.

Definition 13.1. Let $K/\mathbb{Q}$ be a number field. Its Dedekind ζ function is

$$\zeta_K(s) = \sum_{0 \neq I \leq \mathcal{O}_K} \frac{1}{N(I)^s}$$

Example 13.1. $\zeta_{\mathbb{Q}} = \zeta$ is the ordinary Riemann ζ function.

Theorem 13.2. *Suppose K is a quadratic field, then*

$$\lim_{s\to 1^+}(s-1)\zeta_K(s) = \begin{cases} 2\pi h_K/(|d_K|^{1/2}w_K) & \text{if } K \text{ is imaginary} \\ 4h_K(\log\epsilon)/(|d_K|^{1/2}w_K) & \text{if } K \text{ is real} \end{cases}$$

where h_K is the class number of K, w_K the number of roots of unity in K, and ϵ the fundamental unit.

Note that

$$w_K = \begin{cases} 4 & \text{if } K = \mathbb{Q}(i) \\ 6 & \text{if } K = \mathbb{Q}(\sqrt{-3}) \\ 2 & \text{otherwise} \end{cases}$$

Remark. Siegel (1935) showed that the right hand side has order 1 as $|d_K| \to \infty$ by an estimation on the left hand side. For K imaginary quadratic, we have $h_K \to \infty$ as $|d_K| \to \infty$; For K real quadratic, we have $h_K \log\epsilon \to \infty$ as $|d_K| \to \infty$. It was conjectured that there are infinitely many real quadratic fields with $h_K = 1$. So we know from this estimation that these fields must have huge fundamental unit. For example, for $K = \mathbb{Q}(\sqrt{3001})$ has $h_K = 1$ and $\epsilon = u + v\sqrt{3001}$ with $u > 4 \times 10^{36}$.

Proof in the imaginary case. For simplicity, we assume K is not $\mathbb{Q}(i), \mathbb{Q}(\sqrt{-3})$ (both can be verified similarly and separately). So $\mathcal{O}_K^\times = \{\pm 1\}$
Let $J_1, \dots, J_h$ be the ideals representing all of $\mathrm{Cl}(K)$, where $h = h_K$. Take any ideal $I \neq (0)$, there is a unique $i \in \{1, \dots, h\}$ with $IJ_i = (\alpha)$ for some $\alpha \in \mathcal{O}_K$. We have $\alpha \in J_i \setminus \{0\}$, unique to I up to units.
Conversely, if $0 \neq \alpha \in J_i$, then there is a unique ideal I with $IJ_i = (\alpha)$, and we have $N(I)N(J) = N_{K/\mathbb{Q}}(\alpha) > 0$ as K is imaginary. So

$$\zeta_K(s) = \frac{1}{2}\sum_{i=1}^{h} N(J_i)^s \sum_{0\neq\alpha\in J_i} \frac{1}{N_{K/\mathbb{Q}}(\alpha)^s} = \frac{1}{2}\sum_{i=1}^{h} N(J_i)^s \sum_{0\neq\alpha\in J_i} \frac{1}{|\alpha|^{2s}}$$

If we embed $K \subset \mathbb{C}$, then $J_i \subset K \subset \mathbb{C}$ is a lattice of covolume $(1/2)N(J_i)|d_K|^{1/2}$. The result then follows from the next theorem. □

Theorem 13.3. *Let $\Lambda \subset \mathbb{C}$ be a lattice and*

$$Z(s) = Z_\Lambda(s) = \sum_{0\neq x\in\Lambda} \frac{1}{|x|^{2s}}$$

Then $Z(s)$ converges for $s > 1$ and $(s-1)Z(s) \to \pi/\operatorname{covol}(\Lambda)$ as $s \to 1^+$.

Proof. Let w_1, w_2 be a basis for Λ. Take the change of basis $f : \mathbb{R}^2 \to \mathbb{C}, x = (x_1, x_2) \mapsto x_1w_1 + x_2w_2$. Then f takes $\mathbb{Z}^2$ to Λ and $\|x\| = |f(x)|$ is a norm on $\mathbb{R}^2$, whose unit ball $B = \{x \in \mathbb{R}^2 : \|x\| \leq 1\} =$ has area $\mu = \pi/\operatorname{covol}(\Lambda)$.
Let $N(t) = |\{m \in \mathbb{Z}^2 : \|m\| \leq t\}| = |B \cap (1/t)\Lambda|$. Covering B with squares of side length $1/t$, we get $N(t)/t^2 \to \mu$ as $t \to \infty$.
Choose an ordering $\mathbb{Z}^2 = \{x_0 = 0, x_1, x_2, \dots\}$ with $0 < \|x_1\| \leq \|x_2\| \leq \cdots$ and

let $t_k = \|x_k\|$. The same trick as above gives $k/t_k^2 \to \mu$ as $k \to \infty$.
Then

$$Z(s) = \sum_{0 \neq m \in \mathbb{Z}^2} \frac{1}{\|m\|^{2s}} = \sum_{k \geq 1} \frac{1}{t_k^{2s}}$$

For every $\epsilon > 0$, there is some k_ϵ with $\forall k \geq k_\epsilon, (\mu - \epsilon)k^{-1} < t_k^{-2} < (\mu + \epsilon)k^{-1}$, which means that

$$(\mu - \epsilon)^s \sum_{k \geq k_\epsilon} \frac{1}{k^s} < \sum_{k \geq k_\epsilon} \frac{1}{t_k^{2s}} < (\mu + \epsilon)^s \sum_{k \geq k_\epsilon} \frac{1}{k^s}$$

For $s > 1$, we conclude from this that $Z(s)$ converges. Multiplying everything by $s - 1$ and sending $s \to 1^+$, we deduce by Proposition 13.1 that

$$\mu - \epsilon \leq \liminf_{s \to 1^+} (s-1)Z(s) \leq \limsup_{s \to 1^+} (s-1)Z(s) \leq \mu + \epsilon$$

But $\epsilon > 0$ is arbitrary, so $(s-1)Z(s) \to \mu$ as $s \to 1^+$. □

Remark. 1. We can do the same for real quadratic K, but in this case $N_{K/\mathbb{Q}}(\alpha)$ is no longer a norm and one have to keep track of the (infinity family of) units.
2. There is a similar, but more complicated, formula for number fields in general.
3. If you have some other way to compute the left hand side limit, then this does give an explicit formula for the class number.

Theorem 13.4. *Suppose $p \equiv 7 \pmod 8$ is prime and $K = \mathbb{Q}(\sqrt{-p})$. Then $h_K = R - N$ where R is the number of quadratic residues modulo p in the interval $[1, (p-1)/2]$ and N is the number of quadratic non-residues in the same interval.*

Corollary 13.5. *$R > N$ (seems hard without the theorem!).*