Number Fields[*]

April 4, 2019

A number field L is a finite extension of $\mathbb{Q}$. The goal of this course is to define a subring $\mathcal{O}_L \subset L$, that we call the ring of algebraic integers of L, and to study the extent to which its properties generalize those of the usual ring of integers $\mathbb{Z} \subset \mathbb{Q}$. This is the beginning of the subject of algebraic number theory, which was invented by Kummer in the 1840's in order to attack the problem of supplying a proof of Fermat's Last Theorem. We will return to this problem at the end of the course.

The schedules for this course recommend several textbooks, but my personal favourite is Marcus' textbook *Number Fields*.

Non-examinable material, when it appears, is relegated to a subsection labelled with *asterisks*.

Contents

[*] These are Jack Thorne's notes for the Part II Number Fields course, lectured in Lent term 2019 at the University of Cambridge. Please email me with any comments or corrections – you can find my email address on my homepage.

1 The ring of algebraic integers

Recall that a field extension L/K is a pair of fields $K \subset L$. In this case L is a K-vector space and the number $[L : K] = \dim_K L$ is called the degree of the field extension L/K. We say that the extension L/K is finite if its degree is finite.

Definition 1.1. *A number field is a finite extension $L/\mathbb{Q}$.*

Here are two ways to construct number fields. If $\alpha \in \mathbb{C}$ is an algebraic number, then $\mathbb{Q}(\alpha)$ (the smallest field extension of $\mathbb{Q}$ containing α) is a number field. For example, $\mathbb{Q}(i)$ is a number field (an example of an imaginary quadratic field).

Alternatively, if K is a number field and and $f(x) \in K[x]$ is an irreducible monic polynomial, then $L = K[x]/(f(x))$ is a number field, with $[L : K] = \deg f$. Observe that in the first case our number field comes with a distinguished embedding in the complex numbers, while in the second case it does not. An important theme throughout this lecture course will be the role played by the set of all embeddings of a number field in the complex numbers.

In this course we will associate to any number field L a subring $\mathcal{O}_L \subset L$ called the ring of integers of L. This will generalize the inclusion $\mathbb{Z} \subset \mathbb{Q}$.

Definition 1.2.

1. *Let L/K be a field extension. An element $\alpha \in L$ is said to be algebraic over K if there exists a monic polynomial $f(x) \in K[x]$ such that $f(\alpha) = 0$.*
2. *Let $L/\mathbb{Q}$ be a field extension. An element $\alpha \in L$ is called an algebraic integer if there exists a monic polynomial $f(x) \in \mathbb{Z}[x]$ such that $f(\alpha) = 0$. We write $\mathcal{O}_L \subset L$ for the subset of algebraic integers.*

Definition 1.3. *If L/K is a field extension and $\alpha \in L$ is algebraic over K, then the minimal polynomial of α over K is by definition the monic polynomial $f_\alpha(x) \in K[x]$ of least degree such that $f_\alpha(\alpha) = 0$.*

The minimal polynomial is well-defined because we can apply the Euclidean algorithm in the ring $K[x]$: if $f_\alpha(x)$, $g_\alpha(x)$ are two polynomials with this property, then write $f_\alpha(x) = q(x)g_\alpha(x) + r(x)$ for some $q(x), r(x) \in K[x]$ with $\deg r < \deg g_\alpha$. Then setting $x = \alpha$ gives $r(\alpha) = 0$, hence $r = 0$, hence $q = 1$ (as $\deg f_\alpha = \deg g_\alpha$).

Lemma 1.4. *Let $L/\mathbb{Q}$ be a field extension, and let $\alpha \in L$ be an algebraic integer.*

1. *The minimal polynomial $f_\alpha(x)$ of α over $\mathbb{Q}$ is contained in $\mathbb{Z}[x]$. In other words, $f_\alpha(x)$ has integer coefficients.*
2. *If $g(x) \in \mathbb{Z}[x]$ is any polynomial such that $g(\alpha) = 0$, then we can find $q(x) \in \mathbb{Z}[x]$ such that $g(x) = q(x)f_\alpha(x)$.*
3. *Let $F : \mathbb{Z}[x] \to L$ be the ring homomorphism defined by $F(f(x)) = f(\alpha)$. Then the kernel of F equals the principal ideal $(f_\alpha(x))$ of $\mathbb{Z}[x]$.*

Proof. We recall (from IB Groups, Rings and Modules) that if $f(x) = a_0 + a_1 x + \cdots + a_n x^n \in \mathbb{Z}[x]$, then the content $c(f)$ is defined to be the greatest common divisor of $a_0, \ldots, a_n$. Moreover, Gauss' lemma states that if $f(x), g(x) \in \mathbb{Z}[x]$, then $c(fg) = c(f)c(g)$.

Let $f(x) \in \mathbb{Z}[x]$ be a monic polynomial such that $f(\alpha) = 0$. Applying the Euclidean algorithm in $\mathbb{Q}[x]$, we can find $q(x), r(x) \in \mathbb{Q}[x]$ such that $f(x) = q(x) f_\alpha(x) + r(x)$ and $\deg r < \deg f_\alpha$. Specializing to $x = \alpha$, we see $r(\alpha) = 0$, hence $r = 0$. Let n, m be positive integers such that $nq(x)$, $mf_\alpha(x) \in \mathbb{Z}[x]$. Then Gauss' lemma says $nm = c(nmf) = c(nqmf_\alpha) = c(nq)c(mf_\alpha)$. Since q and f_α are both monic, we have $c(nq)|n$ and $c(mf_\alpha)|m$. This is possible only if $c(nq) = n$ and $c(mf_\alpha) = m$, implying that in fact $f_\alpha(x) \in \mathbb{Z}[x]$.

Let $g(x) \in \mathbb{Z}[x]$ be a non-zero polynomial such that $g(\alpha) = 0$. Then we can write $g(x) = q(x) f_\alpha(x) + r(x)$ in $\mathbb{Q}[x]$ with $\deg r < \deg f_\alpha$. Again we see $r = 0$. Let $n \geq 1$ be an integer such that $nq(x) \in \mathbb{Z}[x]$. Then we get $c(ng) = nc(g) = c(nqf_\alpha) = c(nq)$, hence $n|c(nq)$, hence $q(x) \in \mathbb{Z}[x]$, as desired.

The third part of the lemma is simply a reformulation of the second part. □

Corollary 1.5. $\mathcal{O}_\mathbb{Q} = \mathbb{Z}$.

Proof. If $\alpha \in \mathbb{Q}$, its minimal polynomial is $f_\alpha(x) = x - \alpha$. Using the lemma, we see that α is an algebraic integer if and only if $\alpha \in \mathbb{Z}$. □

Proposition 1.6. *Let $L/\mathbb{Q}$ be a field extension. Then $\mathcal{O}_L$ is a ring.*

Proof. It is clear that $0, 1 \in \mathcal{O}_L$ and that if $\alpha \in \mathcal{O}_L$, then $-\alpha \in \mathcal{O}_L$. The hard part is to show that if $\alpha, \beta \in \mathcal{O}_L$, then $\alpha\beta \in \mathcal{O}_L$ and $\alpha + \beta \in \mathcal{O}_L$.

We first observe that the subring $\mathbb{Z}[\alpha] \subset L$ is a finitely generated $\mathbb{Z}$-module. Indeed, by definition, it is generated by the infinitely many elements $1, \alpha, \alpha^2, \ldots$. Let $d = \deg f_\alpha$. Since $f_\alpha(x)$ is monic and $f_\alpha(\alpha) = 0$, we see that $\alpha^d \in \sum_{i=0}^{d-1} \mathbb{Z}\alpha^i$. By induction this implies $\alpha^n \in \sum_{i=0}^{d-1} \mathbb{Z}\alpha^i$ for all $n \geq d$, showing that $\mathbb{Z}[\alpha]$ is in fact generated by the finitely many elements $1, \alpha, \ldots, \alpha^{d-1}$.

Now choose $\alpha, \beta \in \mathcal{O}_L$ and let $d = \deg f_\alpha$, $e = \deg f_\beta$. In the same way, we see that $\mathbb{Z}[\alpha, \beta]$ is generated as a $\mathbb{Z}$-module by the elements $\alpha^i \beta^j$ where $0 \leq i \leq d-1$, $0 \leq j \leq e-1$. Since $\mathbb{Z}[\alpha\beta] \subset \mathbb{Z}[\alpha, \beta]$, we see that $\mathbb{Z}[\alpha\beta]$ is a finitely generated $\mathbb{Z}$-module. It can therefore be generated by some sequence $1, \alpha\beta, (\alpha\beta)^2, \ldots, (\alpha\beta)^{n-1}$, implying the existence of a relation $(\alpha\beta)^n = a_{n-1}(\alpha\beta)^{n-1} + \cdots + a_1\alpha\beta + a_0$ for some integers $a_0, \ldots, a_{n-1}$. Equivalently, $\alpha\beta$ is a zero of the polynomial $x^n - a_{n-1}x^{n-1} - \cdots - a_0 \in \mathbb{Z}[x]$. This shows that $\alpha\beta$ is an algebraic integer. The same argument, using the fact that $\mathbb{Z}[\alpha + \beta] \subset \mathbb{Z}[\alpha, \beta]$ is a finitely generated $\mathbb{Z}$-module, shows that $\alpha + \beta$ is an algebraic integer. □

Accordingly, we call $\mathcal{O}_L$ the ring of algebraic integers of L.

Lemma 1.7. *Let L be a number field and let $\alpha \in L$. Then there exists an integer $n \geq 1$ such that $n\alpha \in \mathcal{O}_L$.*

Proof. Let $f(x) \in \mathbb{Q}[x]$ be a monic polynomial of degree d with α as a root. Then we can find an integer $n \geq 1$ such that the polynomial $g(x) = n^d f(x/n)$ has integer coefficients. The polynomial $g(x)$ is monic, and satisfies $g(n\alpha) = 0$, so $n\alpha \in \mathcal{O}_L$. □

2 Embeddings in $\mathbb{C}$

Let L be a number field.

Definition 2.1. *A complex embedding of L is a field homomorphism $\sigma : L \to \mathbb{C}$.*

Note that if σ is a complex embedding, then σ is injective and $\sigma|_{\mathbb{Q}}$ is the usual embedding $\mathbb{Q} \to \mathbb{C}$.

Proposition 2.2. *Let L/K be an extension of number fields, and let $\sigma_0 : K \to \mathbb{C}$ be a complex embedding. Then the number of distinct embeddings $\sigma : L \to \mathbb{C}$ such that $\sigma|_K = \sigma_0$ is equal to the degree $[L : K]$.*

Proof. We use induction on $[L : K]$. The case $[L : K] = 1$ is trivial since $L = K$ in this case. In general, choose $\alpha \in L - K$, giving a tower of extension $L/K(\alpha)/K$. The tower law from Part II Galois theory states that $[L : K] = [L : K(\alpha)][K(\alpha) : K]$. In particular we have $[L : K(\alpha)] < [L : K]$. If $[K(\alpha) : K] < [L : K]$ then we're done by induction. So we can suppose without loss of generality that $L = K(\alpha)$.

Let $f(x) \in K[x]$ denote the minimal polynomial of α over K. Then there is an isomorphism $K[x]/(f(x)) \to L$, $x \mapsto \alpha$. It follows that to give a homomorphism $\sigma : L \to \mathbb{C}$ extending σ_0 is to give a root of $\sigma_0 f(x)$ in $\mathbb{C}$. We therefore just need to explain why the polynomial $\sigma_0 f(x)$ has distinct roots in $\mathbb{C}$. Equivalently, we must explain why $\sigma_0 f(x)$ and $\sigma_0 f'(x)$ have no roots in common. However, this follows from the fact that $f(x)$ and $f'(x)$ together generate the unit ideal in $K[x]$ (as $f(x)$ is irreducible). □

An important special case of the proposition is where $L = \mathbb{Q}(\alpha)$ is generated by a single element, with minimal polynomial $f_\alpha(x) \in \mathbb{Q}[x]$. In this case the embeddings $L \to \mathbb{C}$ are in bijection with the roots of $f_\alpha(x)$ in $\mathbb{C}$.

If L is a number field and $\sigma : L \to \mathbb{C}$ is a complex embedding, then we write $\overline{\sigma} : L \to \mathbb{C}$ for the complex embedding given by the formula $\overline{\sigma}(\alpha) = \overline{\sigma(\alpha)}$ (i.e. complex conjugation of $\sigma(\alpha)$). There are two possibilities: either $\sigma = \overline{\sigma}$, in which case σ takes values in $\mathbb{R}$, or $\sigma \neq \overline{\sigma}$. We write r for the number of embedding $\sigma : L \to \mathbb{R}$, and s for the number of pairs $\sigma, \overline{\sigma} : L \to \mathbb{C}$ of embeddings with $\sigma \neq \overline{\sigma}$. Then $r + 2s = [L : \mathbb{Q}]$.

Example. If $d \in \mathbb{Z} - \{0, 1\}$ is a square-free integer, then the polynomial $x^2 - d \in \mathbb{Q}[x]$ is irreducible. The corresponding quadratic field is $L = \mathbb{Q}[x]/(x^2 - d) = \mathbb{Q}(\sqrt{d})$. We call it a real quadratic field if $d > 0$ (in which case $r = 2$, $s = 0$) or an imaginary quadratic field if $d < 0$ (in which case $r = 0$, $s = 1$).

Example. Let $m \in \mathbb{Z}$ be a cube-free integer, $m \neq 0, 1$. Let $L = \mathbb{Q}[x]/(x^3 - m)$. Then $r = 1, s = 1$, as follows from the fact that m has 1 real cube root and two complex cube roots in $\mathbb{C}$.

One application of complex embeddings is to understand the trace and norm.

Definition 2.3. *Let L/K be an extension of number fields. Let $\alpha \in L$, and think of $m_\alpha : L \to L$, $m_\alpha(\beta) = \alpha\beta$ as a K-linear endomorphism. Then we define the trace* $\mathrm{tr}_{L/K}(\alpha) = \mathrm{tr}\, m_\alpha$ *and the norm* $N_{L/K}(\alpha) = \det m_\alpha$.

Lemma 2.4. *We have* $\operatorname{tr}_{L/K}(\alpha) = [L : K(\alpha)] \operatorname{tr}_{K(\alpha)/K}(\alpha)$ *and* $N_{L/K}(\alpha) = N_{K(\alpha)/K}(\alpha)^{[L:K(\alpha)]}$.

Proof. This follows from the fact that $L \cong K(\alpha)^{[L:K(\alpha)]}$ as $K(\alpha)$-vector spaces. □

Lemma 2.5. *Let L/K be an extension of number fields of degree $[L : K] = n$, and let $\sigma_0 : K \to \mathbb{C}$ be a complex embedding. Let $\sigma_1, \dots, \sigma_n : L \to \mathbb{C}$ be the distinct complex embeddings such that $\sigma_i|_K = \sigma_0$. Then for each $\alpha \in L$, we have $\sigma_0 \operatorname{tr}_{L/K}(\alpha) = \sigma_1(\alpha) + \dots + \sigma_n(\alpha)$ and $\sigma_0 N_{L/K}(\alpha) = \sigma_1(\alpha) \dots \sigma_n(\alpha)$.*

Proof. By the previous lemma, we can assume that $L = K(\alpha)$, in which case the minimal polynomial $f_\alpha(x)$ of α over K is equal to the characteristic polynomial of the multiplication-by-α endomorphism $m_\alpha : L \to L$. It follows that if $f_\alpha(x) = x^n + a_1 x^{n-1} + \dots + a_n$, then $\operatorname{tr}_{L/K}(\alpha) = -a_1$ and $N_{L/K}(\alpha) = (-1)^n a_n$. On the other hand, we know that there is a factorization

$$\sigma_0 f_\alpha(x) = (x - \sigma_1(\alpha)) \dots (x - \sigma_n(\alpha))$$

in $\mathbb{C}[x]$, showing that $-\sigma_0(a_1) = \sigma_1(\alpha) + \dots + \sigma_n(\alpha)$ and $(-1)^n \sigma_0(a_n) = \sigma_1(\alpha) \dots \sigma_n(\alpha)$. This completes the proof. □

Corollary 2.6. *If $\alpha \in \mathcal{O}_L$, then $\operatorname{tr}_{L/K}(\alpha), N_{L/K}(\alpha) \in \mathcal{O}_K$.*

Proof. Note that for $\beta \in K$, we have $\beta \in \mathcal{O}_K$ if and only if $\sigma_0(\beta) \in \mathcal{O}_\mathbb{C}$ (as for $f(x) \in \mathbb{Z}[x]$, $f(\beta) = 0$ if and only if $f(\sigma_0(\beta)) = 0$).

Let $\alpha \in \mathcal{O}_L$. The previous lemma shows that $\sigma_0(\operatorname{tr}_{L/K}(\alpha))$ is a sum of algebraic integers, hence an algebraic integer, hence $\operatorname{tr}_{L/K}(\alpha)$ is an algebraic integer. The same argument applies to the norm. □

Proposition 2.7. *Let $d \in \mathbb{Z} - \{0, 1\}$ be a square-free integer, and let L be the corresponding quadratic field.*

1. *If $d \equiv 2, 3$ mod 4, then $\mathcal{O}_L = \mathbb{Z}[\sqrt{d}]$.*

2. *If $d \equiv 1$ mod 4, then $\mathcal{O}_L = \mathbb{Z}[\frac{1}{2}(1 + \sqrt{d})]$.*

Proof. In either case we have $L = \{a + b\sqrt{d} \mid a, b \in \mathbb{Q}\}$. Let $\alpha = a + b\sqrt{d} \in L$. We claim that $\alpha \in \mathcal{O}_L$ if and only if $\operatorname{tr}_{L/\mathbb{Q}}(\alpha) = 2a \in \mathbb{Z}$ and $N_{L/\mathbb{Q}}(\alpha) = a^2 - db^2 \in \mathbb{Z}$. We have already seen that these conditions are necessary. To see that they are sufficient, note that the polynomial $x^2 - \operatorname{tr}_{L/\mathbb{Q}}(\alpha)x + N_{L/\mathbb{Q}}(\alpha) \in \mathbb{Q}[x]$ has α as a zero. This shows that the claimed rings are at least subrings of $\mathcal{O}_L$.

Let $\alpha = a + b\sqrt{d} \in \mathcal{O}_L$. Then $a = u/2$ for some $u \in \mathbb{Z}$, hence $4db^2 \in \mathbb{Z}$. Writing $b = r/s$ with $r, s \in \mathbb{Z}$ coprime and $s > 0$, we find $4dr^2 \in s^2\mathbb{Z}$, hence $s^2 | 4d$, hence $s = 1$ or 2 (as d is square-free). Thus we can write $b = v/2$ for some $v \in \mathbb{Z}$ and hence $\alpha = \frac{u}{2} + \frac{v}{2}\sqrt{d}$. We have $u^2 - dv^2 \in 4\mathbb{Z}$, hence $u^2 \equiv dv^2$ mod 4.

If $d \equiv 2, 3$ mod 4 then this forces u and v to both be even, hence $\alpha \in \mathbb{Z}[\sqrt{d}]$. On the other hand if $d \equiv 1$ mod 4 then this forces $u \equiv v$ mod 2, hence $\alpha \in \mathbb{Z}[\frac{1}{2}(1 + \sqrt{d})]$. This completes the proof. □

Another application of the norm is to characterize units. Recall that if R is a ring, then an element $u \in R$ is called a unit if there exists $v \in R$ such that $uv = 1$. The set $R^\times \subset R$ of units forms a group under multiplication.

Lemma 2.8. *Let L be a number field. Then $\mathcal{O}_L^\times = \{\alpha \in \mathcal{O}_L \mid N_{L/\mathbb{Q}}(\alpha) = \pm 1\}$.*

Proof. The norm is multiplicative. If $\alpha \in \mathcal{O}_L^\times$, then there exists $\beta \in \mathcal{O}_L$ such that $\alpha\beta = 1$, hence $N_{L/\mathbb{Q}}(\alpha\beta) = N_{L/\mathbb{Q}}(\alpha)N_{L/\mathbb{Q}}(\beta) = 1$. Since $\mathbb{Z}^\times = \{\pm 1\}$, we must have $N_{L/\mathbb{Q}}(\alpha) = \pm 1$. Suppose conversely that $\alpha \in \mathcal{O}_L$ and $N_{L/\mathbb{Q}}(\alpha) = \pm 1$. We must show that α^{-1} is an algebraic integer. Let $\sigma_1, \dots, \sigma_n : L \to \mathbb{C}$ denote the distinct complex embeddings of L. Then we have $\sigma_1(\alpha) \dots \sigma_n(\alpha) = \pm 1$, hence $\sigma_1(\alpha^{-1}) = \pm\sigma_2(\alpha) \dots \sigma_n(\alpha)$. This shows that $\sigma_1(\alpha^{-1})$ is an algebraic integer, hence that α^{-1} is an algebraic integer. □

3 Discriminant and integral bases

Let L be a number field of degree $n = [L : \mathbb{Q}]$. Let $\sigma_1, \dots, \sigma_n : L \to \mathbb{C}$ be the distinct complex embeddings.

Definition 3.1. *If $\alpha_1, \dots, \alpha_n$ are elements of L, then we define their discriminant*

$$\operatorname{disc}(\alpha_1, \dots, \alpha_n) = \det(\sigma_i \alpha_j)^2.$$

We note that this does not depend on the choice of ordering of the embeddings, or on the choice of ordering of the elements $\alpha_1, \dots, \alpha_n$. Indeed, a change in ordering changes the determinant by a sign, which disappears when we square.

Lemma 3.2. *We have $\operatorname{disc}(\alpha_1, \dots, \alpha_n) = \det(\operatorname{tr}_{L/\mathbb{Q}} \alpha_i \alpha_j)$.*

Proof. Define $n \times n$ matrices $T_{ij} = \operatorname{tr}_{L/\mathbb{Q}} \alpha_i \alpha_j$ and $D_{ij} = \sigma_i \alpha_j$. Then we have

$$T_{ij} = \sum_{k=1}^n \sigma_k(\alpha_i)\sigma_k(\alpha_j) = \sum_{k=1}^n D_{ki} D_{kj} = ({}^t D D)_{ij},$$

hence $T = {}^t D D$, hence $\det T = \det(D)^2 = \operatorname{disc}(\alpha_1, \dots, \alpha_n)$. □

Corollary 3.3. *We have $\operatorname{disc}(\alpha_1, \dots, \alpha_n) \in \mathbb{Q}$. If in fact $\alpha_1, \dots, \alpha_n \in \mathcal{O}_L$, then $\operatorname{disc}(\alpha_1, \dots, \alpha_n) \in \mathbb{Z}$.*

Proof. The previous lemma presents the discriminant as the determinant of a matrix with coefficients in $\mathbb{Q}$. On the other hand if $\alpha_1, \dots, \alpha_n$ are algebraic integers, then $\operatorname{disc}(\alpha_1, \dots, \alpha_n)$ is an algebraic integer in $\mathbb{C}$. Since it is also a rational number, it lies in $\mathcal{O}_\mathbb{Q} = \mathbb{Z}$. □

Proposition 3.4. *We have $\operatorname{disc}(\alpha_1, \dots, \alpha_n) \neq 0$ if and only if the elements $\alpha_1, \dots, \alpha_n$ form a basis for L as $\mathbb{Q}$-vector space.*

Proof. If the elements $\alpha_1, \ldots, \alpha_n$ are linearly dependent over $\mathbb{Q}$, then the columns of the matrix $\sigma_i \alpha_j$ are linearly dependent and the discriminant is zero. Suppose instead that the elements $\alpha_1, \ldots, \alpha_n$ form a basis for L. By the lemma, the discriminant is then non-zero if and only if the symmetric bilinear form $T : L \times L \to \mathbb{Q}$, $T(x, y) = \mathrm{tr}_{L/\mathbb{Q}}\, xy$, is non-degenerate. This is true: for any non-zero element $\beta \in L$, we have $T(\beta, \beta^{-1}) = n \neq 0$. □

The first application of the discriminant is to understanding integral bases.

Definition 3.5. *An integral basis for $\mathcal{O}_L$ is a tuple $(\alpha_1, \ldots, \alpha_n)$ of elements $\alpha_i \in \mathcal{O}_L$ which generate $\mathcal{O}_L$ as a $\mathbb{Z}$-module.*

Lemma 3.6. *If $(\alpha_1, \ldots, \alpha_n)$ is an integral basis, then the map $\mathbb{Z}^n \to \mathcal{O}_L$, $(m_1, \ldots, m_n) \mapsto m_1\alpha_1 + \cdots + m_n\alpha_n$ is an isomorphism.*

Proof. The map is surjective, by definition. Since any element of L admits an integer multiple which lies in $\mathcal{O}_L$, we see that $\alpha_1, \ldots, \alpha_n$ span L; they therefore form a basis of L as $\mathbb{Q}$-vector space and are linearly independent over $\mathbb{Q}$. This implies that the map is injective. □

Lemma 3.7 (Sandwich Lemma).

1. *Let $H \subset G$ be abelian groups such that $G \cong \mathbb{Z}^n$ for some $n \geq 1$. Then $H \cong \mathbb{Z}^m$ for some $m \leq n$.*

2. *Let $K \subset H \subset G$ be abelian groups such that $K \cong \mathbb{Z}^n$ and $G \cong \mathbb{Z}^n$ for some $n \geq 1$. Then $H \cong \mathbb{Z}^n$.*

3. *Let $H \subset G$ be abelian groups such that $H \cong G \cong \mathbb{Z}^n$ for some $n \geq 1$. Then G/H is finite.*

Proof. By the classification of finitely generated abelian groups, there is an isomorphism $H \cong \mathbb{Z}^m$ for some $m \geq 1$. We just need to explain why $m \leq n$. There is an isomorphism $G/H \cong \mathbb{Z}^k \oplus \mathbb{Z}/(d_1) \oplus \cdots \oplus \mathbb{Z}/(d_m)$ for some non-zero integers $d_1, \ldots, d_m$. Let p be a prime not dividing any of $d_1, \ldots, d_m$. Then multiplication by p is injective on G/H. It follows that the map $H/pH \to G/pG$ is injective: if $h \in H$ and $h = pg$ for some $g \in G$, then $g + H$ is an element of G/H which is in the kernel of multiplication by p, hence $g + H = H$ and $g \in H$. This in turn implies that $p^m \leq p^n$, hence $m \leq n$.

This proves the first part of the lemma. The first part implies the second. It remains to prove the third. There is an isomorphism $G/H \cong \mathbb{Z}^a \oplus T$, where $a \geq 0$ and T is a finite abelian group. We must show that $a = 0$. Let p be a prime not dividing the order of T, so that multiplication by p is again injective on G/H. Then the homomorphism $G/pG \to G/H + pG$ is surjective, and contains H/pH in its kernel. Since $|H/pH| = |G/pG| = p^n$ and $|G/H + pG| = p^a$, this can happen only if $a = 0$, as desired. □

Proposition 3.8. *There exists an integral basis for $\mathcal{O}_L$.*

Proof. Let $\beta_1, \ldots, \beta_n$ be a basis of L as $\mathbb{Q}$-vector space. After clearing denominators, we can assume that $\beta_i \in \mathcal{O}_L$ for each $i = 1, \ldots, n$. Thus there is an inclusion

$$\oplus_{i=1}^n \mathbb{Z}\beta_i \subset \mathcal{O}_L.$$

Let $\beta_1^*, \ldots, \beta_n^*$ denote the dual basis of L with respect to the trace form $T(x,y) = \mathrm{tr}_{L/\mathbb{Q}}(xy)$. Then there is an inclusion

$$\mathcal{O}_L \subset \oplus_{i=1}^n \mathbb{Z}\beta_i^*.$$

Indeed, if $\alpha = \sum_{i=1}^n a_i\beta_i^*$ is an element of $\mathcal{O}_L$ (with $a_i \in \mathbb{Q}$), then $T(\alpha, \beta_j) = \mathrm{tr}_{L/\mathbb{Q}}\, \alpha\beta_j \in \mathbb{Z}$ (as the trace of an algebraic integer is an integer). The sandwich lemma implies that $\mathcal{O}_L$ admits an integral basis. □

Definition 3.9. *We call the discriminant of the number field L the number $D_L = \mathrm{disc}(\alpha_1, \ldots, \alpha_n)$, where $(\alpha_1, \ldots, \alpha_n)$ is any integral basis.*

Note that this is independent of the choice of integral basis. Indeed, if $\alpha_1, \ldots, \alpha_n$ and $\beta_1, \ldots, \beta_n$ are two integral bases, then we can find a matrix $A \in \mathrm{GL}_n(\mathbb{Z})$ with $\beta_j = \sum_{k=1}^n A_{kj}\alpha_k$. Then we have $\mathrm{disc}(\beta_1, \ldots, \beta_n) = \mathrm{disc}(\alpha_1, \ldots, \alpha_n)\det(A)^2 = \mathrm{disc}(\alpha_1, \ldots, \alpha_n)$.

The following is a useful tool for calculating the discriminant in some situations.

Proposition 3.10. *Let $L = \mathbb{Q}(\alpha)$ be a number field of degree $[L : \mathbb{Q}] = n$, and let $f(x) \in \mathbb{Q}[x]$ be the minimal polynomial of α. Let $\sigma_1, \ldots, \sigma_n : L \to \mathbb{C}$ be the distinct complex embeddings of L. Then*

$$\mathrm{disc}(1, \alpha, \alpha^2, \ldots, \alpha^{n-1}) = \prod_{i<j}(\sigma_i(\alpha) - \sigma_j(\alpha))^2 = (-1)^{n(n-1)/2} N_{L/\mathbb{Q}}(f'(\alpha)).$$

We note that the term in the middle is what, in Part II Galois Theory, we call the discriminant $\mathrm{disc}\, f$ of the polynomial f.

Proof. The determinant $\det(\sigma_i\alpha^{j-1})$ is a Vandermonde determinant, equal to $\prod_{i<j}(\sigma_j\alpha - \sigma_i\alpha)$. This shows the first equality. For the second, we observe that

$$N_{L/\mathbb{Q}}(f'(\alpha)) = \prod_{i=1}^n \sigma_i f'(\alpha) = \prod_{i=1}^n \prod_{j \neq i}(\sigma_i(\alpha) - \sigma_j(\alpha)) = (-1)^{n(n-1)/2} \prod_{i<j}(\sigma_i(\alpha) - \sigma_j(\alpha))^2,$$

as required. □

Example. We can use the preceding proposition to calculate the discriminant of the quadratic field $L = \mathbb{Q}(\sqrt{d})$, where $d \neq 0, 1$ is a square-free integer. If $d \equiv 2, 3 \bmod 4$ then $\mathcal{O}_L = \mathbb{Z}[\sqrt{d}]$, so we take $f(x) = x^2 - d$ and get $D_L = -N_{L/\mathbb{Q}}(2\sqrt{d}) = 4d$. If $d \equiv 1 \bmod 4$, then $\mathcal{O}_L = \mathbb{Z}[\frac{1}{2}(1 + \sqrt{d})]$, so we take $f(x) = x^2 - x + (1-d)/4$ and get $D_L = -N_{L/\mathbb{Q}}(\sqrt{d}) = d$.

A useful sufficient (but not necessary!) criterion for elements $\alpha_1, \ldots, \alpha_n$ to form an integral basis for $\mathcal{O}_L$ is the following.

Proposition 3.11. *Suppose that $\alpha_1, \ldots, \alpha_n \in \mathcal{O}_L$ and $\mathrm{disc}(\alpha_1, \ldots, \alpha_n)$ is a non-zero square-free integer. Then $(\alpha_1, \ldots, \alpha_n)$ is an integral basis of L.*

Proof. The elements $\alpha_1, \dots, \alpha_n$ are independent over $\mathbb{Q}$, because their discriminant is non-zero. Let $M = \oplus_{i=1}^n \mathbb{Z}\alpha_i \subset \mathcal{O}_L$. Then the index $[\mathcal{O}_L : M]$ is finite, by the sandwich lemma. Let $\beta_1, \dots, \beta_n$ be an integral basis for $\mathcal{O}_L$, and choose a matrix B_{ij} in $M_n(\mathbb{Z})$ such that $\alpha_j = \sum_{k=1}^n B_{kj}\beta_k$. Then we have $\sigma_i(\alpha_j) = \sum_{k=1}^n \sigma_i(\beta_k)B_{kj}$, hence $D(\alpha) = D(\beta)B$ in the obvious notation. This shows that $\operatorname{disc}(\alpha_1, \dots, \alpha_n) = \operatorname{disc}(\beta_1, \dots, \beta_n)\det(B)^2$. If $\operatorname{disc}(\alpha_1, \dots, \alpha_n)$ is square-free, then we must have $\det(B) = \pm 1$, showing that $B \in \mathrm{GL}_n(\mathbb{Z})$ and hence that there exists a matrix $C \in M_n(\mathbb{Z})$ such that $\sum_{l=1}^n B_{kl}C_{lj} = \delta_{kj}$. It follows that $\beta_i \in M$ for each $i = 1, \dots, n$, and hence that $\mathcal{O}_L = M$ and $\alpha_1, \dots, \alpha_n$ is an integral basis for $\mathcal{O}_L$. □

Example. The discriminant of a cubic polynomial $f(x) = x^3 + ax + b$ is $-4a^3 - 27b^2$. Let $f(x) = x^3 - x - 1 \in \mathbb{Z}[x]$. Then $f(x)$ is irreducible over $\mathbb{Q}$ and $\operatorname{disc} f = -23$ is square-free. Let $L = \mathbb{Q}[x]/(f(x))$, and let $\alpha \in L$ denote the image of x mod $(f(x))$. Then α is an algebraic integer, and $\operatorname{disc} \mathbb{Z}[\alpha] = \operatorname{disc} f = -23$. The proposition shows that in fact $\mathbb{Z}[\alpha] = \mathcal{O}_L$.

Finally we introduce some useful definitions for ideals of $\mathcal{O}_L$.

Definition 3.12. *Let L be a number field, and let $I \subset \mathcal{O}_L$ be a non-zero ideal. An integral basis for I is, by definition, a tuple $\alpha_1, \dots, \alpha_n$ of elements of I which generate I as a $\mathbb{Z}$-module.*

Lemma 3.13. *Any non-zero ideal $I \subset \mathcal{O}_L$ admits an integral basis.*

Proof. We first suppose that $I = (\alpha)$ is principal. Then if $\alpha_1, \dots, \alpha_n$ is an integral basis for $\mathcal{O}_L$, then $\alpha\alpha_1, \dots, \alpha\alpha_n$ is an integral basis for I. In general, we can choose a non-zero element $\alpha \in I$, and then $(\alpha) \subset I \subset \mathcal{O}_L$. The result then follows from the sandwich lemma. □

Note that the proof of the lemma implies that if $\alpha_1, \dots, \alpha_n$ is an integral basis of I, then $\alpha_1, \dots, \alpha_n$ forms a basis of L as $\mathbb{Q}$-vector space and there is an isomorphism $I \cong \oplus_{i=1}^n \mathbb{Z}\alpha_i$ of $\mathbb{Z}$-modules.

Definition 3.14. *If $I \subset \mathcal{O}_L$ is a non-zero ideal, then we define $N(I) = [\mathcal{O}_L : I]$. (This index is finite, by the sandwich lemma.)*

Definition 3.15. *We define* $\operatorname{disc}(I) = \operatorname{disc}(\alpha_1, \dots, \alpha_n)$, *where $\alpha_1, \dots, \alpha_n$ is any integral basis for I. The same argument as for $\mathcal{O}_L$ shows that* $\operatorname{disc}(I)$ *is independent of the integral basis chosen.*

Lemma 3.16. *If $I \subset \mathcal{O}_L$ is a non-zero ideal, then* $\operatorname{disc}(I) = \operatorname{disc}(\mathcal{O}_L)N(I)^2$.

Proof. Let $\alpha_1, \dots, \alpha_n$ be an integral basis for $\mathcal{O}_L$, and let $\beta_1, \dots, \beta_n$ be an integral basis for I. Let $B \in M_n(\mathbb{Z})$ be the matrix such that $\beta_j = \sum_{k=1}^n B_{kj}\alpha_k$. We have seen that $\operatorname{disc}(I) = \operatorname{disc}(\mathcal{O}_L)\det(B)^2$, so we just need to show that $|\det(B)| = [\mathcal{O}_L : I]$. By one of the main results from IB Groups, Rings, and Modules, we can find choose $\alpha_1, \dots, \alpha_n$ and $\beta_1, \dots, \beta_n$ such that the matrix B is in Smith normal form, i.e. so that it is diagonal with diagonal entries $d_1|d_2|\dots|d_n$ for some non-zero integers $d_1, \dots, d_n$. Then we have $\det(B) = d_1 \dots d_n$, while $\mathcal{O}_L/I \cong \oplus_{i=1}^n \mathbb{Z}/d_i\mathbb{Z}$. The result follows. □

Lemma 3.17. *Let $\alpha \in \mathcal{O}_L$ be non-zero, and let $I = (\alpha)$. Then $N(I) = |N_{L/\mathbb{Q}}(\alpha)|$.*

Proof. Let $\alpha_1, \dots, \alpha_n$ be an integral basis for $\mathcal{O}_L$. Then $\alpha\alpha_1, \dots, \alpha\alpha_n$ is an integral basis for I. Let $\sigma_1, \dots, \sigma_n : L \to \mathbb{C}$ be the distinct complex embeddings of L. Then we have

$$\operatorname{disc}(I) = \det(\sigma_i(\alpha\alpha_j))^2 = \det(\sigma_i(\alpha)\sigma_i(\alpha_j))^2 = N_{L/\mathbb{Q}}(\alpha)^2 \operatorname{disc}(\mathcal{O}_L),$$

since $\sigma_1(\alpha) \dots \sigma_n(\alpha) = N_{L/\mathbb{Q}}(\alpha)$. Using Lemma 3.16 we find that $N_{L/\mathbb{Q}}(\alpha)^2 = N(I)^2$, hence $|N_{L/\mathbb{Q}}(\alpha)| = N(I)$. □

We will abuse notation slightly by writing $N(\alpha)$ for the norm of the ideal (α). We thus have $N(\alpha) = |N_{L/\mathbb{Q}}(\alpha)|$. It is convenient to define the norm of the zero ideal to be $N(0) = 0$; then this identity holds for $\alpha = 0$ also.

4 Ideals and unique factorization

Let L be a number field. A key difference between $\mathbb{Z}$ and the ring $\mathcal{O}_L$ is that $\mathcal{O}_L$ need not be a unique factorization domain in general. Recall that in a ring R, an element $x \in R$ is said to be irreducible if it is not zero, not a unit, and cannot be expressed as a product $x = yz$ with both y, z non-units. In fact every element of $\mathcal{O}_L$ can be written as a product of irreducibles. The argument, as in the case of $\mathbb{Z}$, is by induction on $N(x)$: we have $N(x) = 1$ if and only if x is a unit. If $N(x) > 1$, then either x is irreducible, or $x = yz$ with both y, z non-units. Then $N(y) < N(x)$ and $N(y) < N(x)$, so by induction both y, z can be written as products of irreducibles.

However, this expression as a product of irreducibles is not unique. Consider, for example, $L = \mathbb{Q}(\sqrt{-5})$, so $\mathcal{O}_L = \mathbb{Z}[\sqrt{-5}]$. The elements $2, 3, 1 + \sqrt{-5}$, $1 - \sqrt{-5}$ are all irreducible. For example, if $yz = 2$, then $N(y)N(z) = N(2) = 4$. If $y = a + b\sqrt{-5}$ and $N(y) = 2$, then $a^2 + 5b^2 = 2$. This equation has no solutions in integers a, b, so we see $N(y) = 1$ or $N(y) = 4$, showing that y or z is a unit. The other cases can be dealt with similarly. Moreover, it is easy to show that none of these elements are associates (as $\mathcal{O}_L^\times = \{\pm 1\}$).

However, we have $6 = 2 \times 3 = (1 + \sqrt{-5}) \times (1 - \sqrt{-5})$. This is an example of an element having two distinct factorizations as a product of irreducibles. How can we get around this? The point is that if we consider ideals, instead of ring elements, then we have the possibility of reducing elements further.

Let R be an integral domain, and let $x, y \in R$. Then $(x) \subset (y)$ if and only if y divides x. In particular, the ideals $(x), (y)$ are equal if and only if x, y are associates (i.e. one is a unit multiple of the other). We recall that if I, J are ideals of R, then we can define their sum and product

$$I + J = \{z + w \mid z \in I, w \in J\}$$

and

$$IJ = \{\sum_i z_i w_i \mid z_i \in I, w_i \in J\}.$$

These are again ideals of R. If $I = (x)$ and $J = (y)$, then $I + J = (x, y)$ and $IJ = (xy)$.

Let us say that an ideal $I \subset R$ is irreducible if it cannot be written as a product $I = JK$, where J, K are proper ideals of R. In the above example, the difficulty is resolved by the fact that the ideals $(2), (3), (1 + \sqrt{-5})$, $(1 - \sqrt{-5})$ are not irreducible, even though they are generated by irreducible elements. In fact, we have $(2) = (2, 1 + \sqrt{-5})^2$ and $(3) = (3, 1 + \sqrt{-5})(3, 1 - \sqrt{-5})$, and you can check that all of these ideals are proper ideals of $\mathcal{O}_L$.

In this chapter, we will show that any non-zero ideal of $\mathcal{O}_L$ can be written uniquely as a product of irreducible ideals of $\mathcal{O}_L$. In fact, it is more convenient to with prime ideals, although these we will eventually see that the two definitions are equivalent.

Definition 4.1. *Let R be a ring. A prime ideal $P \subset R$ is a proper ideal satisfying $xy \in P \Rightarrow x \in P$ or $y \in P$ for any elements $x, y \in R$.*

Lemma 4.2. *Let R be a ring, and let $P \subset R$ be a prime ideal. If $I, J \subset R$ are ideals such that $IJ \subset P$, then $I \subset P$ or $J \subset P$.*

Proof. Suppose that $I \not\subset P$, and let $x \in I \setminus P$. For any $y \in J$, we have $xy \in IJ \subset P$, hence $xy \in P$. Since P is prime, this forces $y \in P$. Since y was arbitrary, we find $J \subset P$. □

From now on, we assume that L is a number field.

Lemma 4.3. *Any non-zero prime ideal of $\mathcal{O}_L$ is maximal.*

Proof. We recall from IB Groups, Rings and Modules that an ideal $I \subset R$ of a ring is prime if and only if R/I is an integral domain, and is maximal if and only if R/I is a field. If $P \subset \mathcal{O}_L$ is a non-zero prime ideal, then R/P is a finite integral domain (of cardinality $N(P)$). Any finite integral domain is a field (as if x is non-zero, then the sequence $x, x^2, x^3, \dots$ must eventually be periodic, by the pigeonhole principle). It follows that P is maximal. □

Lemma 4.4. *Every non-zero ideal $I \subset \mathcal{O}_L$ contains a product of non-zero prime ideals.*

Proof. Suppose not, and let I be a counterexample with $N(I)$ minimal. Then $N(I) > 1$ as $\mathcal{O}_L$ contains prime ideals. (Take any maximal ideal.) Moreover I cannot be prime, so there exist elements $x, y \in \mathcal{O}_L$ such that $xy \in I$ but $x \notin I$ and $y \notin I$. Thus we have $N(I + (x)) < N(I)$ and $N(I + (y)) < N(I)$, so the ideals $I + (x)$ and $I + (y)$ contain products $P_1 \dots P_r$ and $Q_1 \dots Q_s$ of prime ideals. But then we have

$$P_1 \dots P_r Q_1 \dots Q_s \subset (I + (x))(I + (y)) \subset I + (xy) = I.$$

This contradiction concludes the proof. □

Lemma 4.5. *Let $I \subset \mathcal{O}_L$ be a proper non-zero ideal. Then there exists $\gamma \in L \setminus \mathcal{O}_L$ such that $\gamma I \subset \mathcal{O}_L$.*

Proof. Let $\alpha \in I$ be a non-zero element. Then we can find a product of non-zero prime ideals $P_1 \dots P_r \subset (\alpha)$, by the previous lemma. We assume that r is minimal with this property. On the other hand we can find a prime ideal P containing I. Thus P contains some P_i, hence equals P_i. After relabelling, we can assume that $P = P_1$. Since r is minimal, we have

$P_2 \dots P_r \not\subset (\alpha)$, we can choose an element $\beta \in P_2 \dots P_r - (\alpha)$. We claim that the element $\gamma = \beta/\alpha$ satisfies the requirements of the lemma.

If $\gamma \in \mathcal{O}_L$ then $\beta = \alpha\gamma \in (\alpha)$, which contradicts the definition of β. So $\gamma \in L \setminus \mathcal{O}_L$. On the other hand, we have

$$\gamma I = \frac{\beta}{\alpha} I \subset \frac{1}{\alpha} P_2 \dots P_r I \subset P_1 \frac{1}{\alpha} P_2 \dots P_r \subset \mathcal{O}_L,$$

as required. □

Proposition 4.6. *Let $I \subset \mathcal{O}_L$ be a non-zero ideal. Then there exists a non-zero ideal $J \subset \mathcal{O}_L$ such that IJ is principal.*

Proof. Let $\alpha \in I$ be a non-zero element, and set $J = \{\beta \in \mathcal{O}_L \mid \beta I \subset (\alpha)\}$. Then J is a non-zero ideal of $\mathcal{O}_L$, and $IJ \subset (\alpha)$. We will show that in fact $IJ = (\alpha)$.

Let $K = \frac{1}{\alpha} IJ$. Reformulating slightly, we see that $K \subset \mathcal{O}_L$ is a non-zero ideal, and we must show that $K = \mathcal{O}_L$. If $K \neq \mathcal{O}_L$, then we can find $\gamma \in L \setminus \mathcal{O}_L$ such that $\gamma K \subset \mathcal{O}_L$. Observe that this implies $\gamma IJ = \gamma \alpha K \subset (\alpha)$, hence $\gamma J \cap \mathcal{O}_L \subset J$. On the other hand we have $(\alpha) \subset I$, hence $J \subset K$, hence $\gamma J \subset \gamma K \subset \mathcal{O}_L$. We therefore have $\gamma J \subset J$. Since $J \cong \mathbb{Z}^n$ as abelian groups, this implies that γ satisfies a monic polynomial in $\mathbb{Z}[x]$, contradicting $\gamma \notin \mathcal{O}_L$. This contradiction completes the proof. □

Corollary 4.7. *If $I, J, K \subset \mathcal{O}_L$ are non-zero ideals and $IJ = IK$, then $J = K$.*

Proof. We can find an ideal $A \subset \mathcal{O}_L$ such that $AI = (\alpha)$ is principal. Then we find $\alpha J = \alpha K$, hence $J = K$. □

Let $I, J \subset \mathcal{O}_L$ be non-zero ideals. We say that I divides J, and write $I|J$, if there is an ideal $K \subset \mathcal{O}_L$ such that $IK = J$.

Corollary 4.8. *If $I, J \subset \mathcal{O}_L$ are non-zero ideals, then $I|J$ if and only if $I \supset J$.*

Proof. Clearly if there exists K such that $IK = J$, then $I \supset J$. Suppose conversely that $I \supset J$, and fix a non-zero ideal $K \subset \mathcal{O}_L$ such that $IK = (\alpha)$ is principal. Then $\frac{1}{\alpha} JK \subset \mathcal{O}_L$ is an ideal and $\frac{1}{\alpha} JKI = J$, showing that $I|J$. □

Theorem 4.9. *Any non-zero ideal $I \subset \mathcal{O}_L$ admits an expression $I = P_1 \dots P_r$ as a product of prime ideals of $\mathcal{O}_L$. This expression is unique up to re-ordering of terms.*

Proof. Suppose for contradiction that there exists an ideal I which does not admit such an expression. We can assume that $N(I)$ is minimal with this property. Let P be a prime ideal containing I. By the second corollary, we can write $I = PJ$ for some ideal J. Then $J \supset I$. If $J = I$ then by the first corollary we can divide to get $\mathcal{O}_L = P$, a contradiction. Therefore $J \neq I$ and by minimality, we can write J, and hence I, as a product of prime ideals: this is the desired contradiction.

If $I = P_1 \dots P_r = Q_1 \dots Q_s$ admits two expressions as a product of primes, then P_1 divides $Q_1 \dots Q_s$, so contains some Q_i, say Q_1 (after relabelling of terms). Then we must have $P_1 = Q_1$, so we can divide to find $P_2 \dots P_r = Q_2 \dots Q_s$. Continuing in this way we find that $r = s$ and the two expressions are the same. □

Now that we have shown that ideals cancel, we can define the ideal class group of a number field L.

Definition 4.10. *The ideal class group* $\mathrm{Cl}(\mathcal{O}_L)$ *is the set of equivalence classes of non-zero ideals* $I \subset \mathcal{O}_L$ *under the equivalence relation:* $I \sim J$ *if there exists* $\alpha \in L^\times$ *such that* $I = \alpha J$.

It is easy to see that this is an equivalence relation. We write $[I]$ for the equivalence class containing an ideal I.

Lemma 4.11. $\mathrm{Cl}(\mathcal{O}_L)$ *is a group under the operation* $[I][J] = [IJ]$.

Proof. The operation is well-defined on equivalence classes. An identity is $[\mathcal{O}_L]$. The existence of inverses is precisely the statement of Proposition 4.6. □

Proposition 4.12. *The following are equivalent:*

1. $\mathcal{O}_L$ *is a principal ideal domain (PID).*
2. $\mathcal{O}_L$ *is a unique factorization domain (UFD).*
3. *The group* $\mathrm{Cl}(\mathcal{O}_L)$ *is trivial.*

Proof. $(i) \Rightarrow (ii)$: proved in IB Groups, Rings, and Modules.

$(ii) \Rightarrow (iii)$: by unique factorization of ideals, it suffices to show that all non-zero prime ideals of $\mathcal{O}_L$ are principal. If $P \subset \mathcal{O}_L$ is a non-zero prime ideal, choose a non-zero element $\alpha \in P$, and factor $\alpha = \alpha_1 \dots \alpha_r$ as a product of irreducible elements of $\mathcal{O}_L$. Since P is prime, we have $\alpha_i \in P$ for some i; after relabelling, we can assume $\alpha_1 \in P$.

Since $\mathcal{O}_L$ is a UFD, the ideal (α_1) is prime. Since $(\alpha_1) \subset P$ and all non-zero prime ideals of $\mathcal{O}_L$ are maximal, we must have $(\alpha_1) = P$, showing that P is indeed principal.

$(iii) \Rightarrow (i)$: let $I \subset \mathcal{O}_L$ be a non-zero ideal. Then $[I] = [\mathcal{O}_L]$, so there exists $\alpha \in L^\times$ such that $I = \alpha\mathcal{O}_L$, hence $\alpha \in \mathcal{O}_L$ and $I = (\alpha)$. Therefore I is principal. □

Thus the ideal class group $\mathrm{Cl}(\mathcal{O}_L)$ measures the failure of the ring $\mathcal{O}_L$ to be a unique factorization domain.

We will end this section by using unique factorization to show that the ideal norm is multiplicative:

Proposition 4.13. *Let* $I, J \subset \mathcal{O}_L$ *be ideals. Then* $N(IJ) = N(I)N(J)$.

Proof. By convention, we have $N(0) = 0$, and the proposition is trivial if either I or J is the zero ideal. We can therefore assume that they are both non-zero. Let $I = P_1^{e_1} \dots P_r^{e_r}$ denote the factorization of I as a product of powers of distinct primes of $\mathcal{O}_L$. It is clearly enough to show that in fact $N(I) = \prod_{i=1}^r N(P)^{e_i}$. On the example sheet, you will prove the Chinese Remainder Theorem: this says that if $A, B \subset \mathcal{O}_L$ are non-zero ideals with no prime ideal factors in common, then there is an isomorphism $\mathcal{O}_L/(AB) \cong \mathcal{O}_L/A \times \mathcal{O}_L/B$. This clearly implies that $N(AB) = N(A)N(B)$, so reduces us to showing that $N(P^e) = N(P)^e$ for any non-zero prime ideal $P \subset \mathcal{O}_L$ and integer $e \geq 1$.

We have a chain of ideals $\mathcal{O}_L \supset P \supset P^2 \supset \cdots \supset P^e$, and each successive quotient P^i/P^{i+1} is a module over the field $\mathcal{O}_L/P$, which has cardinality $N(P)$. Since we have

$$N(P^e) = |\mathcal{O}_L/P^e| = \prod_{i=0}^{e-1} |P^i/P^{i+1}|,$$

it will suffice to show that each P^i/P^{i+1} is in fact a 1-dimensional vector space over the field $\mathcal{O}_L/P$. Choose an element $\alpha \in P \setminus P^2$. Then we can define a map $\mathcal{O}_L/P \to P^i/P^{i+1}$ by the formula $\beta + P \mapsto \alpha^i \beta + P^{i+1}$. It will suffice to show that this map is surjective. By definition of the map, this is equivalent to showing that $(\alpha^i) + P^{i+1} = P^i$.

We can factor $(\alpha) = PQ$, where Q is not divisible by P, hence $(\alpha^i) = P^i Q^i$. We recall (Corollary 4.8) that for non-zero ideals A, B of $\mathcal{O}_L$, A divides B if and only if A contains B. In paticular, A divides $(\alpha^i) + P^{i+1}$ if and only if A contains (α^i) and P^{i+1}, if and only if A divides $(\alpha^i) = P^i Q^i$ and P^{i+1}. It follows that we must have $(\alpha^i) + P^{i+1} = P^i$, as desired. □

5 Dedekind's criterion

Let L be a number field. We now discuss how to actually construct prime ideals of $\mathcal{O}_L$. If $P \subset \mathcal{O}_L$ is a prime ideal, then $\mathcal{O}_L/P$ is a finite field, so there is exactly one prime number p such that $P|(p)$. This reduces us to factorizing the ideal (p) for each prime number p. We begin with a preliminary observation:

Lemma 5.1. *Suppose that p is a prime, and factor $(p) = P_1^{e_1} \dots P_r^{e_r}$, where $P_1, \dots, P_r$ are distinct prime ideals of $\mathcal{O}_L$ and $e_i \geq 1$ for each $i = 1, \dots, r$. Then $n = \sum_{i=1}^r e_i f_i$, where $N(P_i) = p^{f_i}$.*

Proof. This follows immediately from the fact that the ideal norm $N(I)$ is multiplicative. □

Note in particular that this implies $r \leq n$. It is helpful to introduce some terminology:

Definition 5.2. *Let p be a prime number, and factor $(p) = P_1^{e_1} \dots P_r^{e_r}$, where $P_1, \dots, P_r$ are distinct prime ideals of $\mathcal{O}_L$.*

1. *We say that p ramifies in $\mathcal{O}_L$ if $e_i > 1$ for some $i = 1, \dots, r$.*
2. *We say that p is inert in $\mathcal{O}_L$ if (p) is prime (i.e. $r = 1$ and $e_1 = 1$).*
3. *We say that p splits completely in $\mathcal{O}_L$ if $r = n$ (equivalently, $e_i = f_i = 1$ for each $i = 1, \dots, r$).*

It is often possible to factor (p) using the following theorem:

Theorem 5.3 (Dedekind's theorem)**.** *Suppose that $\alpha \in \mathcal{O}_L$ is such that $L = \mathbb{Q}(\alpha)$. Let $f_\alpha(x) \in \mathbb{Z}[x]$ be the minimal polynomial of α over $\mathbb{Q}$, and let p be a prime number not* *dividing the index* $[\mathcal{O}_L : \mathbb{Z}[\alpha]]$. *Let* $\overline{f}_\alpha(x) \in \mathbb{F}_p[x]$ *denote the reduction of* $f_\alpha(x)$ *modulo* p, *and suppose that there is a factorization*

$$\overline{f}_\alpha(x) = \prod_{i=1}^{r} \overline{g}_i(x)^{e_i}$$

in $\mathbb{F}_p[x]$, *where the* $e_i \geq 1$ *are integers and the* $\overline{g}_i(x)$ *are pairwise distinct irreducibles.*

For each $i = 1, \dots, r$, *choose a polynomial* $g_i(x) \in \mathbb{Z}[x]$ *with reduction modulo* p *equal to* $\overline{g}_i$, *and let* $Q_i = (p, g_i(\alpha))$. *Then for each* $i = 1, \dots, r$, Q_i *is a prime ideal of* $\mathcal{O}_L$ *which does not depend on the choice of* g_i, *and*

$$(p) = \prod_{i=1}^{r} Q_i^{e_i}$$

is the factorization of (p) *as a product of powers of distinct prime ideals of* $\mathcal{O}_L$.

Proof. Consider the ring $A = \mathbb{Z}[\alpha]/p\mathbb{Z}[\alpha]$. Then there are isomorphisms

$$A \cong \mathbb{Z}[x]/(p, f_\alpha(x)) \cong \mathbb{F}_p[x]/(\overline{f}_\alpha(x)) \cong \prod_{i=1}^{r} \mathbb{F}_p[x]/(\overline{g}_i(x))^{e_i},$$

the last isomorphism by the Chinese Remainder Theorem applied to the Euclidean domain $\mathbb{F}_p[x]$. Each ring $\mathbb{F}_p[x]/(\overline{g}_i(x))^{e_i}$ has a quotient $\mathbb{F}_p[x]/(\overline{g}_i(x))$, which is a finite field of cardinality p^{f_i}, where $f_i = \deg \overline{g}_i(x)$. We define $\overline{P}_i \subset A$ to be the kernel of the composite map

$$A \to \mathbb{F}_p[x]/(\overline{g}_i(x))^{e_i} \to \mathbb{F}_p[x]/(\overline{g}_i(x)).$$

Then $\overline{P}_1, \dots, \overline{P}_r$ are pairwise distinct prime ideals, and in fact $\overline{P}_i = (\overline{g}_i(\alpha))$.

The map $\mathbb{Z}[\alpha] \to \mathcal{O}_L$ induces, by passage to quotient, a map $\mathbb{Z}[\alpha]/p\mathbb{Z}[\alpha] \to \mathcal{O}_L/(p)$. We claim that this map is an isomorphism. Observe that both source and target have cardinality p^n. It therefore suffices to show that the map is surjective. Let $N = [\mathcal{O}_L : \mathbb{Z}[\alpha]]$. By hypothesis, p does not divide N, so we can find integers $a, b \in \mathbb{Z}$ such that $aN + bp = 1$. If $x \in \mathcal{O}_L$ then $Nx \in \mathbb{Z}[\alpha]$, hence $x = aNx + bpx \in \mathbb{Z}[\alpha] + p\mathcal{O}_L$; hence $x \bmod (p)$ is equal to the image of the element $aNx \in \mathbb{Z}[\alpha]$, and our map is surjective.

The prime ideals of $\mathcal{O}_L$ containing p are in bijection with the prime ideals of $\mathcal{O}_L/(p)$. These in turn are in bijection with the ideals of $A = \mathbb{Z}[\alpha]/p\mathbb{Z}[\alpha]$. If $\overline{P} \subset A$ is a prime, then the corresponding ideal of $\mathcal{O}_L$ is (P, p), where P denotes the pre-image of $\overline{P}$ in $\mathbb{Z}[\alpha]$. Let $Q_i = (p, g_i(\alpha))$ be the prime ideal of $\mathcal{O}_L$ corresponding to $\overline{P}_i$. Then $Q_1, \dots, Q_r$ are pairwise distinct prime ideals of $\mathcal{O}_L$ containing p, and $N(Q_i) = p^{f_i}$ for each $i = 1, \dots, r$.

It remains to check that $(p) = Q_1^{e_1} \dots Q_r^{e_r}$. However, we have

$$Q_i^{e_i} = (p, g_i(\alpha))^{e_i} \subset (p, g_i(\alpha)^{e_i}),$$

hence

$$Q_1^{e_1} \dots Q_r^{e_r} \subset (p, g_1(\alpha)^{e_1} \dots g_r(\alpha)^{e_i}) = (p),$$

as $g_1(x)^{e_1} \dots g_r(x)^{e_r} = f_\alpha(x) + pg(x)$ for some $g(x) \in \mathbb{Z}[x]$. Taking norms, we have

$$N(Q_1^{e_1} \dots Q_r^{e_r}) = N(Q_1)^{e_1} \dots N(Q_r)^{e_r} = p^{e_1 f_1 + \dots + e_r f_r} = p^n = N(p).$$

It follows that we must have $Q_1^{e_1} \dots Q_r^{e_r} = (p)$, which is what we needed to prove. □

As an example, let $L = \mathbb{Q}(\sqrt{-11})$ and consider the factorization of (5). Since $-11 \equiv 1 \bmod 4$, we have $\mathcal{O}_L = \mathbb{Z}[\frac{1}{2}(1+\sqrt{-11})]$; this ring contains $\mathbb{Z}[\sqrt{-11}]$ with index 2. Since 5 and 2 are coprime, we can apply Dedekind's criterion to the polynomial $f_\alpha(x) = x^2 + 11$. The reduction modulo 5 is $x^2 + 1$; since $-1 \equiv 2^2 \bmod 5$, this factors as $(x-2)(x-3)$ in $\mathbb{F}_5[x]$. We find that the prime factorization of (5) is

$$(5) = (5, \sqrt{-11} - 2)(5, \sqrt{-11} - 3),$$

and that these factors are distinct prime ideals of $\mathcal{O}_L$. In particular, 5 splits completely in $\mathcal{O}_L$.

In fact, the same argument works for any quadratic field:

Proposition 5.4. *Let d be a square-free integer, $d \neq 0, 1$. Let $L = \mathbb{Q}(\sqrt{d})$. Let p be a prime number.*

1. *Suppose that p is odd. Then:*

   (a) *If p divides d, then p ramifies in L; there exists a unique prime ideal $P \subset \mathcal{O}_L$ dividing (p), and $(p) = P^2$.*

   (b) *If p does not divide d and d is a quadratic residue modulo p, then p splits completely in L; there exist distinct prime ideals $P, Q \subset \mathcal{O}_L$ such that $(p) = PQ$.*

   (c) *If p does not divide d and d is not a quadratic residue modulo p, then p is inert in L; the ideal (p) is prime.*

2. *Suppose instead that $p = 2$. Then:*

   (a) *If $d \equiv 2, 3 \bmod 4$ then 2 is ramified in L.*

   (b) *If $d \equiv 1 \bmod 8$ then 2 splits completely in L.*

   (c) *If $d \equiv 5 \bmod 8$ then 2 is inert in L.*

Proof. We just treat the case $p = 2$. If $d \equiv 2, 3 \bmod 4$, then $\mathcal{O}_L = \mathbb{Z}[\sqrt{d}]$ so we can apply Dedekind's criterion with $f_\alpha(x) = x^2 - d$. Modulo 2, this polynomial is x^2 or $x^2 + 1 = (x+1)^2$; in either case we see that 2 is ramified.

If $d \equiv 1 \bmod 4$, then $\mathcal{O}_L = \mathbb{Z}[\frac{1}{2}(1+\sqrt{d})] = \mathbb{Z}[\alpha]$, say. We can apply Dedekind's criterion with $f_\alpha(x) = x^2 - x + \frac{1}{4}(1-d)$. If $d \equiv 1 \bmod 8$ then modulo 2 this polynomial is $x^2 + x = x(x+1)$, so 2 splits in $\mathcal{O}_L$. If $d \equiv 5 \bmod 8$ then modulo 2 this polynomial is $x^2 + x + 1$, which is irreducible in $\mathbb{F}_2[x]$; and 2 is inert. □

6 The geometry of numbers

The following theorem is the most important in the entire course.

Theorem 6.1. *Let L be a number field. Then there exists a constant C_L such that every class $[I] \in \mathrm{Cl}(\mathcal{O}_L)$ contains a representative ideal I of norm $N(I) \leq C_L$.*

We will calculate C_L explicitly later on. This will give an effective way to compute ideal class groups in practice.

Corollary 6.2. *$\mathrm{Cl}(\mathcal{O}_L)$ is a finite abelian group. Moreover, it is generated by the classes $[P]$ of the non-zero prime ideals $P \subset \mathcal{O}_L$ of norm $N(P) \leq C_L$.*

We call the cardinality of $\mathrm{Cl}(\mathcal{O}_L)$ the class number of L.

Proof. Every class in $\mathrm{Cl}(\mathcal{O}_L)$ is represented by an ideal $I \subset \mathrm{Cl}(\mathcal{O}_L)$ of norm $N(I) \leq C_L$. To prove the finiteness of $\mathrm{Cl}(\mathcal{O}_L)$, we need to show that there are only finitely many such ideals I. It suffices to show that for each integer $N \geq 1$, there are only finitely many ideals of norm $N(I) = N$. If $N(I) = N$ then, by Lagrange's theorem, N annihilates the finite abelian group $\mathcal{O}_L/I$, hence $N \in I$. Since ideals of $\mathcal{O}_L$ containing N correspond to ideals of the finite ring $\mathcal{O}_L/(N)$, there are only finitely many possibilities.

If $N(I) \leq C_L$ and $I = \prod_{i=1}^r P_i^{e_i}$ is the prime factorisation of I, then $[I] = \prod_{i=1}^r [P_i]^{e_i}$. This shows that $[I]$ is contained in the subgroup generated by the classes of the prime factors of $[I]$. On the other hand, since the norm is multiplicative we must have $N(P_i) \leq C_L$ for each $i = 1, \ldots, r$. This shows that $\mathrm{Cl}(\mathcal{O}_L)$ is generated by the classes of prime ideals of norm at most C_L. □

To prove Theorem 6.1, we will actually prove a slightly different statement:

Theorem 6.3. *There exists a constant C_L with the following property: let $I \subset \mathcal{O}_L$ be a non-zero ideal. Then there exists a non-zero element $\alpha \in I$ such that $N(\alpha) \leq C_L N(I)$.*

Proof of Theorem 6.1 using Theorem 6.3. Let $I \subset \mathcal{O}_L$ be a non-zero ideal. Choose an element $\alpha \in I$ of norm $N(\alpha) \leq C_L N(I)$. Then $I | (\alpha)$, so there exists an ideal $J \subset \mathcal{O}_L$ such that $IJ = (\alpha)$. It follows that $[I] = [J]^{-1}$ and $N(J) = N(\alpha)/N(I) \leq C_L$. In other words, the inverse of every ideal class contains a representative of norm at most C_L. Since $\mathrm{Cl}(\mathcal{O}_L)$ is a group, every ideal class is the inverse of another ideal class, so this completes the proof. □

We can therefore focus on proving Theorem 6.3. We first give a sketch of the proof in the case that L is an imaginary quadratic field. We will then go on to treat the general case.

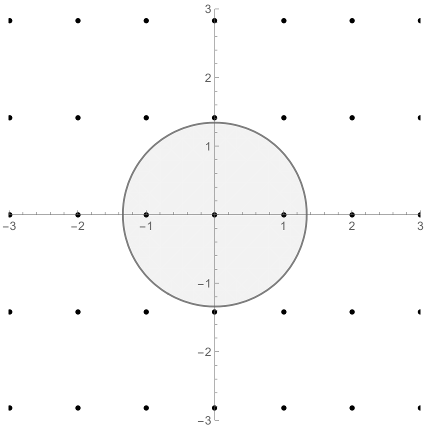
We will use the "geometry of numbers". The key idea is that of embedding the ring of integers of a number field in a Euclidean vector space. For this reason, we start with the following definitions:

Definition 6.4. *Let V be a finite-dimensional $\mathbb{R}$-vector space. A lattice $\Lambda \subset V$ is the $\mathbb{Z}$-span of a basis for V as $\mathbb{R}$-vector space.*

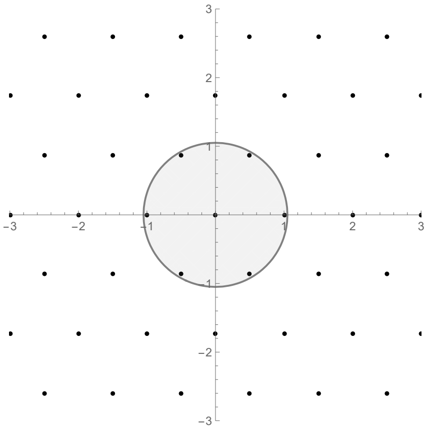
Definition 6.5. *Let V be a finite-dimensional inner product space over $\mathbb{R}$, and let $\Lambda \subset V$ be a lattice. The covolume $A(\Lambda)$ is the volume of a fundamental parallelotope, i.e. the set $\{\sum_{i=1}^n t_i v_i \mid t_i \in [0,1)\}$, where $v_1, \dots, v_n$ is a $\mathbb{Z}$-basis for Λ.*

We note that the covolume is independent of the choice of $\mathbb{Z}$-basis of Λ. Indeed, it equals the absolute value of the determinant of the matrix with columns equal to $v_1, \dots, v_n$. If $v_1', \dots, v_n'$ is a different choice of basis, then we chance the value of the covolume by a factor $|\det(B)|$ for some $B \in \mathrm{GL}_n(\mathbb{Z})$; and this factor equals 1.

Now let $d \in \mathbb{Z}$ be a negative square-free integer, and let $L = \mathbb{Q}(\sqrt{d})$. We recall that $\mathcal{O}_L = \mathbb{Z}[\sqrt{d}]$ (if $d \equiv 2, 3 \bmod 4$) or $\mathcal{O}_L = \mathbb{Z}[\frac{1}{2}(1+\sqrt{d})]$ (if $d \equiv 1 \bmod 4$). Let $\sigma : L \to \mathbb{C}$ be a complex embedding. Then $\sigma(\mathcal{O}_L)$ is a lattice in $\mathbb{C}$, viewed as a real vector space of dimension 2:



$\sigma(\mathbb{Z}[\sqrt{-2}])$



$\sigma(\mathbb{Z}[\frac{1}{2}(1+\sqrt{-3})])$

More generally, if $I \subset \mathcal{O}_L$ is a non-zero ideal then $\sigma(I)$ is a lattice in $\mathbb{C}$. Its covolume is given by the following lemma.

Lemma 6.6. *Let $I \subset \mathcal{O}_L$ be a non-zero ideal. Then $A(\sigma(I)) = \frac{1}{2}\sqrt{|\operatorname{disc}(I)|} = \frac{N(I)}{2}\sqrt{|D_L|}$.*

Proof. Let α_1, α_2 be an integral basis for I. Writing $\sigma(\alpha_1) = x_1 + iy_1$, $\sigma(\alpha_2) = x_2 + iy_2$, we get

$$A(I) = \left| \det \begin{pmatrix} x_1 & y_1 \\ x_2 & y_2 \end{pmatrix} \right|.$$

On the other hand, we have

$$\operatorname{disc}(I) = \det \begin{pmatrix} x_1 + iy_1 & x_2 + iy_2 \\ x_1 - iy_1 & x_2 - iy_2 \end{pmatrix}^2 = (2i)^2 A(I)^2.$$

This implies the desired result. □

The key geometric input is the following result about lattices in a 2-dimensional Euclidean vector space.

Theorem 6.7 (Minkowski's theorem in dimension 2)**.** *Let $\Lambda \subset \mathbb{R}^2$ be a lattice. Then there exists $\lambda \in \Lambda \setminus \{0\}$ such that $|\lambda|^2 \leq \frac{4}{\pi} A(\Lambda)$.*

The important point is that the existence of a non-zero point of Λ in the closed disk of given radius depends only on the area of $A(\Lambda)$, and not on the shape of Λ itself.

Corollary 6.8. *Let $C_L = \frac{2}{\pi}\sqrt{|D_L|}$. For any non-zero ideal $I \subset \mathcal{O}_L$, there exists a non-zero element $\alpha \in I$ such that $N(\alpha) \leq C_L N(I)$.*

Proof. Apply Minkowski's theorem (Theorem 6.7) to $\sigma(I) \subset \mathbb{C}$. □

Thus the class group of an imaginary quadratic field L is generated by the classes of non-zero prime ideals $P \subset \mathcal{O}_L$ of norm $N(P) \leq C_L$. This allows us to effectively calculate the ideal class group:

Example. Let $L = \mathbb{Q}(\sqrt{-7})$. Then $|D_L| = 7$, so $C_L = 2\sqrt{7}/\pi < 2$. There are no primes $p < 2$, so $\mathrm{Cl}(\mathcal{O}_L)$ must be the trivial group, and $\mathbb{Z}[\frac{1+\sqrt{-7}}{2}]$ must be a UFD.

Example. Let $L = \mathbb{Q}(\sqrt{-5})$. Then $|D_L| = 20$, so $C_L = 4\sqrt{5}/\pi < 3$. Using Dedekind's criterion, we find that $(2) = (2, 1+\sqrt{5})^2 = P^2$, say. If P is not principal, then we will have $\mathrm{Cl}(\mathcal{O}_L) \cong \mathbb{Z}/2\mathbb{Z}$. However, if $P = (a + b\sqrt{-5})$ then $N(P) = 2 = a^2 + 5b^2$. This equation has no solutions with $a, b \in \mathbb{Z}$, so P is indeed not a principal ideal.

We now go beyond the case of imaginary quadratic fields and treat the case of a general number field L. We first state and prove the more general version of Minkowski's theorem that we use.

Theorem 6.9 (Minkowski's theorem)**.** *Let Λ be a lattice in $\mathbb{R}^n$, and let $E \subset \mathbb{R}^n$ be a subset satisfying the following conditions:*

1. *The boundary ∂E has volume 0.*

2. *E is convex.*

3. *E is centrally symmetric: i.e. $E = -E$.*

Then if $\mathrm{vol}(E) > 2^n A(\Lambda)$, *then E contains a non-zero point of Λ. If E is moreover compact, this this conclusion holds under the weaker assumption that* $\mathrm{vol}(E) \geq 2^n A(\Lambda)$.

Proof. We first treat the case of strict inequality. Let $v_1, \dots, v_n$ be a $\mathbb{Z}$-basis for Λ, and let P denote the corresponding fundamental parallelotope. Then $\mathbb{R}^n$ is a disjoint union of translates $\lambda + P$, $\lambda \in \Lambda$. It follows that

$$\frac{1}{2}E = \sqcup_{\lambda \in \Lambda}((\frac{1}{2}E) \cap \lambda + P).$$

We thus have

$$\mathrm{vol}(P) < \mathrm{vol}(\frac{1}{2}E) = \sum_{\lambda \in \Lambda} \mathrm{vol}((\frac{1}{2}E) \cap \lambda + P) = \sum_{\lambda \in \Lambda} \mathrm{vol}((\frac{1}{2}E - \lambda) \cap P).$$

If the sets $\frac{1}{2}E - \lambda$ were pairwise disjoint, then the right-hand side here would be bounded above by $\mathrm{vol}(P)$: a contradiction. Therefore there must exist distinct elements $\lambda, \mu \in \Lambda$ such that the intersection

$$(\frac{1}{2}E - \lambda) \cap (\frac{1}{2}E - \mu)$$

is non-empty. Using that E is centrally symmetric and convex, this implies that $\lambda - \mu$ is a non-zero element of $\Lambda \cap E$.

To treat the case of non-strict inequality, we recall (theorem of Heine–Borel) that E is compact if and only if it is closed and bounded. Applying the first part of the theorem to the sets $(1 + \frac{1}{m})E$, we find non-zero elements $\lambda_m \in (1 + \frac{1}{m})E \cap \Lambda$. These points are all contained in $2E \cap \Lambda$, which is a finite set; we can therefore find a non-zero element $\lambda \in \Lambda$ which is contained in $(1 + \frac{1}{m})E$ for all $m \geq 1$. □

Now let $n = [L : \mathbb{Q}]$, let $\sigma_1, \dots, \sigma_r$ denote the real embeddings of L, and let $\tau_1, \overline{\tau}_1, \dots, \tau_s, \overline{\tau}_s$ denote the conjugate pairs of complex embeddings. We identify $\mathbb{C} = \mathbb{R}^2$ and define a map $S : L \to \mathbb{R}^n = \mathbb{R}^r \times \mathbb{C}^s$ by the formula

$$S : \alpha \mapsto (\sigma_1(\alpha), \dots, \sigma_r(\alpha), \tau_1(\alpha), \dots, \tau_s(\alpha)).$$

This is an injective homomorphism.

Lemma 6.10. *$S(\mathcal{O}_L)$ is a lattice. More generally, if $I \subset \mathcal{O}_L$ is a non-zero ideal, then $S(I)$ is a lattice.*

Proof. Fix an integral basis $\alpha_1, \dots, \alpha_n$ of I. We must show that the vectors $S(\alpha_1), \dots, S(\alpha_n)$ are linearly independent; in other words, that the matrix A with columns

$$(\sigma_1(\alpha_i), \dots, \sigma_r(\alpha_i), \mathrm{Re}\tau_1(\alpha_i), \mathrm{Im}\tau_1(\alpha_i), \dots, \mathrm{Re}\tau_s(\alpha_i), \mathrm{Im}\tau_s(\alpha_i))$$

is non-singular. After performing column operations on S, we can transform this matrix to the matrix with columns

$$(\sigma_1(\alpha_i), \dots, \sigma_r(\alpha_i), \tau_1(\alpha_i), \overline{\tau_1(\alpha_i)}, \dots, \tau_s(\alpha_i), \overline{\tau_s(\alpha_i)}).$$

We find $(-2i)^{2s} \det A^2 = \mathrm{disc}(I)$. In particular, $\det A \neq 0$ and $S(I)$ is a lattice. □

Lemma 6.11. *Let $I \subset \mathcal{O}_L$ be a non-zero ideal. Then $A(S(I)) = \frac{1}{2^s}\sqrt{|\,\mathrm{disc}(I)|} = \frac{N(I)}{2^s}\sqrt{|D_L|}$.*

Proof. The proof is the same computation with determinants as in the proof in the previous lemma. □

The following is a restatement of Theorem 6.3.

Theorem 6.12. *For any non-zero ideal $I \subset \mathcal{O}_L$, there exists an element $\alpha \in I$ such that $N(\alpha) \leq C_L N(I)$, where $C_L = (\frac{4}{\pi})^s \frac{n!}{n^n}\sqrt{|D_L|}$.*

We call the value $C_L = (\frac{4}{\pi})^s \frac{n!}{n^n}\sqrt{|D_L|}$ the Minkowski constant.

Proof. We will apply Minkowski's theorem to the lattice $S(I) \subset \mathbb{R}^n = \mathbb{R}^r \times \mathbb{C}^s$. For any $t \geq 0$, we define a region

$$B_{r,s}(t) = \{(y_1, \dots, y_r, z_1, \dots, z_s) \in \mathbb{R}^r \times \mathbb{C}^s \mid \sum_{i=1}^{r} |y_i| + 2\sum_{j=1}^{s} |z_j| \leq t\}.$$

This is a compact subset of $\mathbb{R}^n$. It is clearly centrally symmetric. It is convex, as follows immediately from the triangle inequality. We claim that $\operatorname{vol}(B_{r,s}(t)) = 2^r \left(\frac{\pi}{2}\right)^s \frac{t^n}{n!}$. We assume this for now, and prove the proposition.

We choose t so that $B_{r,s}(t)$ has volume equal to $2^n A(S(I))$, or in other words so that

$$t^n = n! \left(\frac{4}{\pi}\right)^s N(I)\sqrt{|D_L|}.$$

Then Minkowski's lemma implies that there exists a non-zero element $\alpha \in I$ such that $S(\alpha) \in B_{r,s}(t)$. Let $S(\alpha) = (y_1, \dots, y_r, z_1, \dots, z_s)$. Using the AM-GM inequality, we have

$$N(\alpha)^{1/n} = (|y_1 \dots y_r z_1 \overline{z}_1 \dots z_s \overline{z}_s|)^{1/n} \leq \frac{|y_1| + \cdots + |y_r| + 2|z_1| + \cdots + 2|z_r|}{n} \leq \frac{t}{n},$$

hence $N(\alpha) \leq t^n/n^n = N(I)C_L$, as required.

It remains to prove that $\operatorname{vol}(B_{r,s}(t)) = 2^r(\frac{\pi}{2})^s\frac{t^n}{n!}$. This can be proved by induction on r, s, the base cases being $B_{1,0}(t) = [-t, t]$ and $B_{0,1}(t) = D(0, \frac{t}{2})$ (we leave it as an exercise). □

We note that other regions could have been used in the proof of Theorem 6.12; however, there is no simple choice that gives a better value for the constant C_L. We repeat the following consequences of Theorem 6.12:

Theorem 6.13. *The ideal class group* $\mathrm{Cl}(\mathcal{O}_L)$ *is finite. It is generated by the classes of non-zero prime ideals* $P \subset \mathcal{O}_L$ *of norm* $N(P) \leq C_L$.

This bound is amazingly sharp! The following example illustrates this.

Example. Let $f(x) = x^5 - x + 1$, and let $L = \mathbb{Q}[x]/(f(x))$. (One can check that $f(x)$ is irreducible modulo 5, hence irreducible over $\mathbb{Q}$.) Let $\alpha = x \bmod (f(x)) \in \mathcal{O}_L$. The discriminant of the polynomial $x^5 + ax + b$ is $5^5b^4 + 2^8a^5$. In the present case we have $a = -1$, $b = 1$, giving

$$\operatorname{disc} \mathbb{Z}[\alpha] = 2869 = 19 \times 151.$$

In particular, this discriminant is square-free, showing that $\mathbb{Z}[\alpha] = \mathcal{O}_L$. The polynomial $f(x)$ has exactly one real root, so $r = 1$, $s = 2$. We have

$$C_L = \left(\frac{4}{\pi}\right)^2 \frac{5!}{5^5}\sqrt{2869} < 4,$$

so every ideal class of $\mathrm{Cl}(\mathcal{O}_L)$ contains a representative I of norm $N(I) < 4$. We will show that there are no ideals of norm $N(I) = 2$ or 3; this will imply that $\mathrm{Cl}(\mathcal{O}_L)$ is trivial, and hence that $\mathbb{Z}[\alpha]$ is a PID. If $N(I) = 2$, then I is a prime ideal dividing 2. By Dedekind's criterion there exists such an ideal I if and only if $f(x)$ mod 2 has a linear factor. This is not the case. Similarly $f(x)$ mod 3 does not have a linear factor, showing that the field L has class number 1.

We conclude with some more examples.

Example. Let $L = \mathbb{Q}(\sqrt{10})$. Then $C_L = \sqrt{10} < 4$, so $\mathrm{Cl}(\mathcal{O}_L)$ is generated by the primes dividing 2 and 3. Applying Dedekind's criterion, we find that $2 = P_2^2$, where $P_2 = (2, \sqrt{10})$, and $3 = P_3 P_3'$, where $P_3 = (3, 1 + \sqrt{10})$ and $P_3' = (3, 1 - \sqrt{10})$. Thus $\mathrm{Cl}(\mathcal{O}_L)$ is generated by $[P_2]$ and $[P_3]$. The principal ideal $(2 + \sqrt{10})$ has norm 6, so in fact $\mathrm{Cl}(\mathcal{O}_L)$ is generated by P_2. If P_2 is not principal, then $\mathrm{Cl}(\mathcal{O}_L) \cong \mathbb{Z}/2\mathbb{Z}$. However, if $P_2 = (a + b\sqrt{10})$ then $a^2 - 10b^2 = \pm 2$, implying that one of $2, -2$ is a quadratic residue modulo 5. This is not the case, so we find that P is not principal.

Example. Let $L = \mathbb{Q}(\sqrt{-17})$. Then $|D_L| = 4 \times 17$, so $C_L = 4\sqrt{17}/\pi < 6$. Using Dedekind's criterion, we find that $(2) = P_2^2$, where $P_2 = (2, 1 + \sqrt{-17})$. We have $(3) = P_3 P_3'$, where $P_3 = (3, 1 + \sqrt{-17})$ and $P_3' = (3, 1 - \sqrt{-17})$ (and these ideals are distinct). The ideal (5) is prime. Thus $\mathrm{Cl}(\mathcal{O}_L)$ is generated by $[P_2]$ and $[P_3]$.

To understand the structure of this group, we look for relations between these ideals. For example, we have $N(1 + \sqrt{-17}) = 1 + 17 = 18 = 2 \cdot 3 \cdot 3$. A calculation shows that $P_2 P_3^2 = (1 + \sqrt{-17})$, hence $[P_2] = [P_3]^{-2}$ and $\mathrm{Cl}(\mathcal{O}_L)$ is generated by P_3. If P_2 is not principal, then we will have $\mathrm{Cl}(\mathcal{O}_L) \cong \mathbb{Z}/4\mathbb{Z}$. The equation $a^2 + 17b^2 = 2$ is not soluble in integers a, b, so indeed P_2 is not principal.

7 Unit group and Dirichlet's theorem

Let L be a number field of degree $[L : \mathbb{Q}] = n$. We have proved the finiteness of the class number of L. The second fundamental theorem that we will prove concerns the group $\mathcal{O}_L^\times$ of units in $\mathcal{O}_L$. Recall that these are precisely the elements $\alpha \in \mathcal{O}_L$ with $N(\alpha) = 1$.

Theorem 7.1 (Dirichlet's unit theorem). *Let $\mu_L \subset \mathcal{O}_L^\times$ denote the group of roots of unity in $\mathcal{O}_L^\times$. Then μ_L is a finite cyclic group and there is an isomorphism $\mathcal{O}_L^\times \cong \mu_L \times \mathbb{Z}^{r+s-1}$.*

In fact, the proof will show more. Let $\tau_1, \dots, \tau_r : L \to \mathbb{R}$ denote the distinct real embeddings of L, and let $\sigma_1, \overline{\sigma}_1, \dots, \sigma_s, \overline{\sigma}_s : L \to \mathbb{C}$ be the distinct complex embeddings. Define a map

$$\ell : \mathcal{O}_L^\times \to \mathbb{R}^{r+s}$$

by the formula

$$\ell(\alpha) = (\log(|\tau_1(\alpha)|), \dots, \log(|\tau_r(\alpha)|), 2\log(|\sigma_1(\alpha)|), \dots, 2\log(|\sigma_s(\alpha)|)).$$

Then ℓ is group homomorphism of abelian groups, and its image is contained inside the hyperplane

$$H = \{(x_1, \dots, x_{r+s}) \in \mathbb{R}^{r+s} \mid \sum_{i=1}^{r+s} x_i = 0\}.$$

This inclusion expresses the identity (for $\alpha \in \mathcal{O}_L^\times$):

$$0 = \log N(\alpha) = \sum_{i=1}^{r} \log(|\tau_i(\alpha)|) + 2\sum_{i=1}^{s} \log(|\sigma_i(\alpha)|).$$

The proof of Theorem 7.1 will in fact show that $\ker \ell = \mu_L$ and that $\ell(\mathcal{O}_L^\times) \subset H$ is a lattice; in particular, there is an isomorphism $\ell(\mathcal{O}_L^\times) \cong \mathbb{Z}^{r+s-1}$.

The proof of Theorem 7.1 is non-examinable, so we first focus on giving examples. First, we see that the unit group is finite if and only if $r + s - 1 = 0$. There are two possibilities: either $r = 1, s = 0$ (in which case $L = \mathbb{Q}$) or $r = 0, s = 1$ (in which case L is an imaginary quadratic field).

The first interesting case is when $r = 2$ and $s = 0$, which is that of real quadratic fields. Let $d > 1$ be a squarefree integer and let $L = \mathbb{Q}(\sqrt{d})$. Let $\sigma : L \to \mathbb{R}$ be the embedding which sends $\sqrt{d}$ to the positive square-root of d. The existence of this embedding shows that $\mu_L = \{\pm 1\}$. Consider the map $\ell' : \mathcal{O}_L^\times \to \mathbb{R}$, $\alpha \mapsto \log|\sigma(\alpha)|$. From the above discussion, we see that ℓ' is a homomorphism with kernel μ_L, and its image is a lattice in $\mathbb{R}$. In particular, there is exactly one unit $\alpha \in \mathcal{O}_L^\times$ such that $\sigma(\alpha) > 0$, $\log(|\sigma(\alpha)|) > 0$, and $\mathcal{O}_L^\times = \{\pm\alpha^n \mid n \in \mathbb{Z}\}$. We call α the fundamental unit of L. Equivalently, $\alpha \in \mathcal{O}_L^\times$ is the unique unit such that $\sigma(\alpha) > 1$ and $\mathcal{O}_L^\times = \{\pm\alpha^n \mid n \in \mathbb{Z}\}$.

How can we find the fundamental unit α? We first prove a simple lemma. In order to simplify notation, we now identify L with its image in $\mathbb{R}$ (so σ is the identity embedding).

Lemma 7.2. *1. If $d \equiv 2, 3$ mod 4, let $u = a + b\sqrt{d} \in \mathcal{O}_L^\times$ satisfy $u > 1$. Then $a \geq b \geq 1$.*

2. If $d \equiv 1$ mod 4, let $u = \frac{1}{2}(a + b\sqrt{d}) \in \mathcal{O}_L^\times$ satisfy $u > 1$. Then $a \geq b \geq 1$.

Proof. We treat each case in turn. First suppose that $d \equiv 2, 3 \bmod 4$. Let $\overline{u} = a - \sqrt{d}b$. Then $u\overline{u} = \pm 1$, hence $u^{-1} = \pm\overline{u}$ and $|\overline{u}| < 1$. Since $2a = u + \overline{u}$ and $2b\sqrt{d} = u - \overline{u}$, we see that $a > 0$ and $b > 0$, hence $a \geq 1$ and $b \geq 1$ (as a, b are integers). The equation $(a/b)^2 = d \pm 1/b^2$ shows that $a \geq b$ (as $d \geq 2$).

Now suppose that $d \equiv 1 \bmod 4$. Let $\overline{u} = \frac{1}{2}(a - b\sqrt{d})$. Then $u\overline{u} = \pm 1$, hence $u^{-1} = \pm\overline{u}$ and $|\overline{u}| < 1$. Since $a = u + \overline{u}$ and $\sqrt{d}b = u - \overline{u}$, we see that $a > 0$ and $b > 0$, hence $a \geq 1$ and $b \geq 1$ (as a, b are integers). The equation $(a/b)^2 = d \pm 4/b^2$ again shows that $a \geq b$ (as $d \geq 5$). □

We can now explain how to find the fundamental unit. First suppose that $d \equiv 2, 3 \bmod 4$, and let $\alpha = a_1 + b_1\sqrt{d}$ be the fundamental unit. Write $\alpha^k = a_k + b_k\sqrt{d}$. Then the relation $b_{k+1} = a_1 b_k + b_1 a_k$ shows that the sequence $b_1, b_2, \dots$ is strictly increasing. We can therefore characterize the fundamental unit as follows: let $b \geq 1$ be the least positive integer such that $db^2 \pm 1 = a^2$ is a perfect square, where $a > 0$. Then $\alpha = a + b\sqrt{d}$ is the fundamental unit.

Now suppose $d \equiv 1 \bmod 4$, and let $\alpha = \frac{1}{2}(a_1 + b_1\sqrt{d})$ be the fundamental unit. Write $\alpha^k = \frac{1}{2}(a_k + b_k\sqrt{d})$. Then $b_{k+1} = \frac{1}{2}(a_1 b_k + b_1 a_k)$. This shows that $b_{k+1} \geq b_k$, with equality if and only if $a_1 = b_1 = 1$ and $a_k = b_k$. If $a_1 = b_1 = 1$ then the identity $a_1^2 - db_1^2 = \pm 4$ shows that we must have $d = 5$. Excluding the case $d = 5$ for the moment, we find that

$b_{k+1} > b_k$, and hence this sequence is strictly increasing. We can therefore characterize the fundamental unit as follows: let $b \geq 1$ be the least positive integer such that $db^2 \pm 4 = a^2$ is a perfect square, where $a > 0$. Then $\alpha = \frac{1}{2}(a + b\sqrt{d})$ is the fundamental unit.

We now return to the case $d = 5$. In this case we see that the sequence b_k is at least non-decreasing. Note that for each positive integer b there are at most two positive integers a such that $\frac{1}{2}(a + b\sqrt{d})$ is a unit. Using the characterization of the fundamental unit as the least element $\beta \in \mathcal{O}_L^\times$ such that $\beta > 1$, we see that we can characterize the fundamental unit in this case as follows: let $b \geq 1$ be the least positive integer such that $db^2 + 4$ or $db^2 - 4$ has the form a^2, where $a \geq 1$ is a positive integer; among these, choose the smallest possible value of a. Then $\alpha = \frac{1}{2}(a + b\sqrt{d})$ is the fundamental unit.

Carrying this out in the case $d = 5$, we see that both $d+4$ and $d-4$ are squares, and that a fundamental unit is obtained by taking $a = 1$, hence $\alpha = \frac{1}{2}(1 + \sqrt{5})$.

Some examples of fundamental units are shown in the following table.

d	α
2	$1+\sqrt{2}$
3	$2+\sqrt{3}$
5	$\frac{1}{2}(1+\sqrt{5})$
7	$8+3\sqrt{7}$
10	$3+\sqrt{10}$
11	$10+3\sqrt{11}$
13	$\frac{1}{2}(3+\sqrt{13})$
14	$15+4\sqrt{14}$
15	$4+\sqrt{15}$
17	$4+\sqrt{17}$
19	$170+39\sqrt{19}$
21	$\frac{1}{2}(5+\sqrt{21})$
22	$197+42\sqrt{22}$

The examples show that even for relatively small values of d, the above procedure may not be very efficient. There is a more efficient algorithm which allows one to write down the fundamental unit in terms of the continued fraction expansion of $\sqrt{d}$. We do not discuss this in this course.

7.1 *Proof of Dirichlet's unit theorem*

The proof of Dirichlet's unit theorem is again based on the geometry of numbers. As usual, let $\sigma_1, \dots, \sigma_r, \tau_1, \overline{\tau}_1, \dots, \tau_s, \overline{\tau}_s$ denote the complex embeddings of L. We extend ℓ to a group homomorphism $\ell : L^\times \to \mathbb{R}^{r+s}$ given by the formula

$$\ell(\alpha) = (\log|\sigma_1(\alpha)|, \dots, \log|\sigma_r(\alpha)|, 2\log|\tau_1(\alpha)|, \dots, 2\log|\tau_s(\alpha)|).$$

Then $\ell(\mathcal{O}_L^\times)$ is contained in the codimension 1 subspace $H \subset \mathbb{R}^{r+s}$ where all of the coordinates sum to zero, because if $\alpha \in \mathcal{O}_L^\times$ then

$$\log|\sigma_1(\alpha)| + \cdots + \log|\sigma_r(\alpha)| + 2\log|\tau_1(\alpha)| + 2\log|\tau_s(\alpha)| = \log|N_{L/\mathbb{Q}}(\alpha)| = \log N(\alpha) = 0.$$

The geometry of numbers enters in the proof of the following lemma.

Lemma 7.3. *Fix an integer k with $1 \leq k \leq r+s$ and $\alpha \in \mathcal{O}_L \setminus \{0\}$. Let $\ell(\alpha) = (a_1, \ldots, a_{r+s})$. Then we can find $\beta \in \mathcal{O}_L \setminus \{0\}$ satisfying the following conditions:*

1. *$N(\beta) \leq (\frac{2}{\pi})^s \sqrt{|\operatorname{disc}(\mathcal{O}_L)|}$.*

2. *Let $\ell(\beta) = (b_1, \ldots, b_{r+s})$. Then $b_i < a_i$ if $i < k$.*

Proof. Let $E \subset \mathbb{R}^n = \mathbb{R}^r \times \mathbb{C}^s$ be the region defined by the inequalities

$$|y_1| \leq c_1, \ldots, |y_r| \leq c_r, |z_1|^2 \leq c_{r+1}, \ldots, |z_s|^2 \leq c_{r+s},$$

where the c_i are positive real numbers chosen to satisfy the conditions

$$0 < c_i < e^{a_i} \ (i \neq k)$$

and

$$c_1 \ldots c_{r+s} = (\frac{2}{\pi})^s \sqrt{|D_L|}.$$

We apply Minkowski's theorem to the lattice $S(\mathcal{O}_L)$ and the compact region E in $\mathbb{R}^r \times \mathbb{C}^s$. The volume of E equals $2^n A(S(\mathcal{O}_L))$, so Minkowski's theorem implies the existence of a non-zero element $\beta \in \mathcal{O}_L$ such that $S(\beta) \in E$. This element has the desired properties. □

Corollary 7.4. *Fix an integer k with $1 \leq k \leq r+s$. Then there exists an element $\epsilon \in \mathcal{O}_L^\times$ such that, writing $\ell(\epsilon) = (e_1, \ldots, e_{r+s})$, we have $e_i > 0$ if $i \neq k$ and $e_k < 0$.*

Proof. Choose an arbitrary element $\alpha \in \mathcal{O}_L \setminus \{0\}$. Applying Lemma 7.3 repeatedly, we can find elements $\alpha_1, \alpha_2, \ldots$ such that $N(\alpha_j) \leq (\frac{2}{\pi})^s \sqrt{|D_L|}$ for all j and the i-entry of $\ell(\alpha_j/\alpha_{j+1})$ is positive if $i \neq k$.

Since the elements α_j are infinite in number and bounded in norm, there must exist $j < j'$ such that $(\alpha_j) = (\alpha_{j'})$. This implies that $\alpha_j/\alpha_{j'} \in \mathcal{O}_L^\times$. It has the required property by construction. □

Lemma 7.5. *Let $N \geq 1$ be an integer and let $A \in M_N(\mathbb{R})$ be a matrix satisfying the following conditions:*

1. *For each $j = 1, \ldots, N$, $\sum_{i=1}^N A_{ij} = 0$.*

2. *For each $i, j = 1, \ldots, N$, we have $A_{ij} > 0$ if $i = j$ and $A_{ij} < 0$ if $i \neq j$.*

Then A has rank $N - 1$.

Proof. Since the rows of A sum to zero, the rank is at most $N-1$. We show that the first $N-1$ rows of A are in fact linearly independent. Suppose that there are real numbers $t_1, \dots, t_{N-1}$, not all zero, such that $\sum_{i=1}^{N-1} t_i A_{ij} = 0$ for each $j = 1, \dots, N$. After rescaling, we can assume that there is k such that $t_k = 1$ and $t_i \leq 1$ for all $i = 1, \dots, N-1$. Then we have

$$0 = \sum_{i=1}^{N-1} t_i A_{ik} \geq \sum_{i=1}^{N-1} A_{ik} > \sum_{i=1}^{N} A_{ik} = 0.$$

This is a contradiction. □

Lemma 7.6. *Let $B > 0$ be a real number, and let $X_B = \{\alpha \in \mathcal{O}_L \mid \forall \sigma : L \to \mathbb{C}, |\sigma(\alpha)| \leq B\}$. Then X_B is finite.*

Proof. $S(X_B) \subset \mathbb{R}^r \times \mathbb{C}^s$ is the intersection of a compact set with a lattice. It is therefore finite. □

Proposition 7.7. *$\ell(\mathcal{O}_L^\times)$ is a lattice in H.*

Proof. We first show that $\ell(\mathcal{O}_L^\times)$ spans H. By the corollary, we can find elements $v_1, \dots, v_{r+s} \in \ell(\mathcal{O}_L^\times)$ such that the i-entry of v_j is strictly positive if $i \neq j$ and strictly negative if $i = j$. We claim that these vectors span H. This follows from Lemma 7.5. Indeed, let $A \in M_{r+s}(\mathbb{R})$ be the matrix with column j given by v_j. Then A satisfies the hypotheses of that lemma, so has rank $r + s - 1$.

Let us therefore choose vectors $v_1, \dots, v_{r+s-1} \in L(\mathcal{O}_L^\times)$ which form a basis of H as $\mathbb{R}$-vector space. They span a lattice $\Lambda \subset H$. Let $P \subset H$ denote the set of combinations $\sum_{i=1}^{r+s-1} t_i v_i$, where $t_i \in [0, 1]$ for each $i = 1, \dots, r+s-1$ (in other words, P is the closure of a fundamental parallelotope). We observe that $P \cap \ell(\mathcal{O}_L^\times)$ is finite. Indeed, if $\ell(\alpha) \in P$ then $|\sigma(\alpha)|$ is bounded independently of the embedding $\sigma : L \to \mathbb{C}$, so we can apply Lemma 7.6.

For any element $x \in \ell(\mathcal{O}_L^\times)$, we can write x in the form $x = \lambda + p$ where $\lambda \in \Lambda$ and $p \in P \cap \ell(\mathcal{O}_L^\times)$. It follows that the index $N = [\ell(\mathcal{O}_L^\times) : \Lambda]$ is finite. Lagrange's theorem implies that $N\ell(\mathcal{O}_L^\times) \subset \Lambda$, hence $\Lambda \subset \ell(\mathcal{O}_L^\times) \subset \frac{1}{N}\Lambda$. By the sandwich lemma, we find that $\ell(\mathcal{O}_L^\times) \cong \mathbb{Z}^{r+s-1}$, and hence that $\ell(\mathcal{O}_L^\times)$ is a lattice in H. □

We can finally complete the proof of Dirichlet's unit theorem.

Theorem 7.8. *The group μ_L of roots of unity in $\mathcal{O}_L^\times$ is finite and cyclic, and there is an isomorphism $\mathcal{O}_L^\times \cong \mu_L \times \mathbb{Z}^{r+s-1}$.*

Proof. Note that $\ker \ell$ is contained in the set X_1 of Lemma 7.6, so it is finite. It must therefore be equal to μ_L, which is a cyclic group. Let $u_1, \dots, u_{r+s-1} \in \mathcal{O}_L^\times$ be elements which project to a $\mathbb{Z}$-basis of $\ell(\mathcal{O}_L^\times)$. We define a map $f : \mu_L \times \mathbb{Z}^{r+s-1} \to \mathcal{O}_L^\times$ by the formula $(w, a_1, \dots, a_{r+s-1}) \mapsto w u_1^{a_1} \dots u_{r+s-1}^{a_{r+s-1}}$. This is an isomorphism. □

8 Cyclotomic fields

We are going to use cyclotomic number fields to study the Fermat equation. As a warm-up, let's see how to use number fields to find all the Pythagorean triples $x^2 + y^2 = z^2$, where $x, y, z \in \mathbb{Z}$ satisfy $\gcd(x, y, z) = 1$. Note that x and y cannot both be even; and looking modulo 4 shows that (without loss of generality, after switching x and y) we can assume that x is odd and y is even.

Given such a triple, we can factor $x^2 + y^2 = (x + iy)(x - iy)$ in $\mathbb{Z}[i]$. I claim that the ideals $(x + iy)$, $(x - iy)$ of $\mathbb{Z}[i]$ are coprime (i.e. have no prime ideal factors in common). Indeed, if $P \subset \mathbb{Z}[i]$ is a prime ideal dividing both, then P divides $(2x)$ and $(2y)$. If $N(P) = p^f$ then taking norms, we see that this forces p to divide $2x$ and $2y$, hence to divide 2. This would imply that z is divisible by 2, a contradiction.

This shows that $(x + iy)$ and $(x - iy)$ are coprime ideals. Since their product $(z)^2$ is a square, $(x + iy)$ must be the square of an ideal of $\mathbb{Z}[i]$. Since $\mathbb{Z}[i]$ is a PID, we can write $(x + iy) = (a + ib)^2 = (a^2 - b^2 + 2abi)$ for some $a, b \in \mathbb{Z}$. In particular, $x + iy$ is a multiple of $a^2 - b^2 + 2abi$ by a unit of $\mathbb{Z}[i]$. Since we have assumed that x is odd and y is even, this unit must be ± 1, showing that $x = a^2 - b^2$, $y = 2ab$ or $x = b^2 - a^2$, $y = -2ab$.

We now begin our study of cyclotomic fields. Let p be an odd prime number, which will be fixed throughout this section.

Definition 8.1. *Let $\zeta_p = e^{2\pi i/p} \in \mathbb{C}$. The p^{th} cyclotomic field is $K = \mathbb{Q}(\zeta_p)$.*

Since ζ_p is a zero of the polynomial $X^p - 1$, $\mathbb{Q}(\zeta_p)$ is a number field.

Lemma 8.2.

1. *The element $1 - \zeta_p \in \mathcal{O}_K$ satisfies $N(1 - \zeta_p) = p$ and $(1 - \zeta_p)^{p-1} = (p)$. The ideal $(1 - \zeta_p) \subset \mathcal{O}_K$ is prime.*
2. *The polynomial $f_p(x) = (x^p - 1)/(x - 1) \in \mathbb{Z}[x]$ is irreducible, and $[K : \mathbb{Q}] = p - 1$.*

Proof. The polynomial $f_p(x)$ has ζ_p as a zero. In fact, we can factor $f_p(x) = \prod_{i=1}^{p-1}(x - \zeta_p^i)$. This shows that $[K : \mathbb{Q}] \leq p - 1$. We have $f_p(1) = p = \prod_{i=1}^{p-1}(1 - \zeta_p^i)$.

We note that for each $i = 1, \ldots, p - 1$, we have $(1 - \zeta_p^i) = (1 - \zeta_p)$ as ideals of $\mathcal{O}_K$. Indeed, we can write $1 - \zeta_p^a = (1 - \zeta_p)(1 + \zeta_p + \cdots + \zeta_p^{a-1})$, showing that $1 - \zeta_p$ divides $1 - \zeta_p^a$. Conversely, if $ab \equiv 1 \bmod p$ then we have $1 - \zeta_p = 1 - \zeta_p^{ab} = (1 - \zeta_p^a)(1 + \zeta_p^a + \cdots + \zeta_p^{a(b-1)})$, showing that $1 - \zeta_p^a$ divides $1 - \zeta_p$.

We therefore have $(1 - \zeta_p)^{p-1} = (p)$ as ideals of $\mathcal{O}_K$. Taking norms gives $N(1 - \zeta_p)^{p-1} = p^{[K:\mathbb{Q}]}$. This is only possible if $N(1 - \zeta_p) = p$ and $[K : \mathbb{Q}] = p - 1$. In particular, $f_p(x)$ is irreducible. Since the ideal $(1 - \zeta_p)$ has prime norm, it is prime. □

Observe that the roots of $f_p(x)$ in $\mathbb{C}$ are precisely the primitive p^{th} roots of unity $\zeta_p^a = e^{2\pi ia/p}$, $a = 1, \ldots, p - 1$. It follows that $K = \mathbb{Q}(\zeta_p)$ has $r = 0$, $s = (p - 1)/2$. Moreover, the elements $1, \zeta_p, \ldots, \zeta_p^{p-2}$ are linearly independent over $\mathbb{Q}$.

Lemma 8.3. *We have* $\mathrm{disc}(1, \zeta_p, \ldots, \zeta_p^{p-2}) = (-1)^{(p-1)/2} p^{p-2}$.

Proof. We use the formula $\mathrm{disc}(1, \zeta_p, \dots, \zeta_p^{p-2}) = (-1)^{(p-1)(p-2)/2} N_{L/\mathbb{Q}}(f_p'(\zeta_p))$. Let $\pi = 1 - \zeta_p$. We have already seen that $f_p(1) = \prod_{i=1}^{p-1}(1 - \zeta_p^i) = N_{K/\mathbb{Q}}(1 - \zeta_p) = p$. On the other hand, we have $f_p'(x) = ((x-1)px^{p-1} - (x^p - 1))/(x-1)^2$, hence $f_p'(\zeta_p) = p\zeta_p^{-1}/(\zeta_p - 1)$. All together this gives

$$\mathrm{disc}(1, \zeta_p, \dots, \zeta_p^{p-2}) = (-1)^{(p-1)/2} p^{p-2},$$

as required. □

It follows that $\mathbb{Z}[\zeta_p] \subset \mathcal{O}_K$ is a finite index subring, this index being divisible at most by powers of p.

Proposition 8.4. *The ring of integers of $\mathbb{Q}(\zeta_p)$ equals $\mathbb{Z}[\zeta_p]$.*

Proof. We have already seen that $[\mathcal{O}_K : \mathbb{Z}[\zeta_p]]$ is a power of p. Therefore there exists $N \geq 1$ such that $p^N \mathcal{O}_K \subset A$. On the other hand, the residue field $\mathcal{O}_K/(1 - \zeta_p)$ has cardinality p, so is isomorphic to $\mathbb{Z}/p\mathbb{Z}$. Therefore any element $z_0 \in \mathcal{O}_K$ can be written in the form $z_0 = a_0 + (1 - \zeta_p)z_1$, where $a_0 \in \mathbb{Z}$ and $z_1 \in \mathcal{O}_K$. Applying the same argument to z_1, we can write $z_0 = a_0 + (1 - \zeta_p)a_1 + (1 - \zeta_p)^2 z_2$, where $a_0, a_1 \in \mathbb{Z}$ and $z_2 \in \mathcal{O}_K$. Proceeding by induction, we see that for any $n \geq 1$, each element $z_0 \in \mathcal{O}_K$ admits an expression $x = a_0 + (1 - \zeta_p)a_1 + \cdots + (1 - \zeta_p)^{n-1}a_{n-1} + (1 - \zeta_p)^n z_n$, where $a_0, \dots, a_{n-1} \in \mathbb{Z}$ and $z_n \in \mathcal{O}_K$. Consequently, we have $\mathcal{O}_K = \mathbb{Z}[1 - \zeta_p] + (1 - \zeta_p)^n \mathcal{O}_K$. Taking $n = (p-1)N$ and using that $\mathbb{Z}[\zeta_p] = \mathbb{Z}[1 - \zeta_p]$, we get $\mathcal{O}_K = \mathbb{Z}[\zeta_p] + p^N \mathcal{O}_K = \mathbb{Z}[\zeta_p]$, as required. □

Corollary 8.5. *The only prime l which ramifies in L is $l = p$.*

Proof. We can apply Dedekind's criterion. Since the polynomial $x^p - 1$ has distinct roots modulo l for any prime $l \neq p$, we see that any such prime l is unramified in L. The prime p is ramified, since we have $(p) = (1 - \zeta_p)^{p-1}$ and $p - 1 > 1$. □

We now turn to the units of $\mathbb{Z}[\zeta_p]$.

Lemma 8.6. *The only roots of unity in $\mathbb{Q}(\zeta_p)$ are $\pm\zeta_p^a$ for $a = 0, \dots, p-1$.*

Proof. The group μ_K is cyclic; we wish to show it is cyclic of order $2p$. We first show that its order is not divisible by 4, i.e. that $i \notin L$. Otherwise, we would have $(2) = (1+i)^2$, implying that 2 is ramified in K: this would contradict the corollary. The same argument shows that the order of μ_K is not divisible by any odd prime $l \neq p$. It remains to check that the order of μ_K is not divisible by p^2. Suppose that there exists a primitive p^2th root of unity ω in L. The same argument as above then shows that the ideals $(1 - \omega^a)$, $a \in \mathbb{Z}/p^2\mathbb{Z}$, $(a, p) = 1$, are all equal. Evaluating the polynomial $(x^{p^2} - 1)/(x^p - 1)$ at $x = 1$ gives the identity $(1 - \omega)^{p(p-1)} = (p)$. However, this would imply that $N(1 - \omega)^{p(p-1)} = p^{(p-1)}$, hence $N(1 - \omega)^p = p$, a contradiction. □

Lemma 8.7 (Kummer's Lemma). *Let $u \in \mathcal{O}_K^\times$. Then there exists $a \in \mathbb{Z}$ such that $\zeta_p^a u \in K \cap \mathbb{R}$.*

Note that $[K : K \cap \mathbb{R}] = 2$, and in fact $K \cap \mathbb{R} = \mathbb{Q}(\zeta_p + \zeta_p^{-1})$.

Proof. If $\sigma : K \to \mathbb{C}$ is a complex embedding, then for all $y \in K$, we have $\sigma(\overline{y}) = \overline{\sigma(y)}$; in other words, σ commutes with the action of complex conjugation. It suffices to check this for $y = \zeta_p$, so $\sigma(\zeta_p) = \zeta_p^a$ for some $a = 1, \dots, p-1$. Then $\sigma(\overline{\zeta_p}) = \sigma(\zeta_p^{-1}) = \zeta_p^{-a} = \overline{\sigma(\zeta_p)}$, as required.

Let $y = u/\overline{u}$. Then y is a unit and for any $\sigma : K \to \mathbb{C}$, $\sigma(y) = \sigma(u)/\overline{\sigma(u)}$ is a complex number of absolute value 1. We have seen in the course of providing Dirichlet's unit theorem that this forces y to be a root of unity. We can therefore write $u = \pm\zeta_p^k\overline{u}$ for some $k \in \mathbb{Z}$. After possibly replacing k by $k+p$, we can assume that $k = 2g$ for some $g \in \mathbb{Z}$.

We claim that $u = \zeta_p^{2g}\overline{u}$ (i.e. the sign is $+$ and not $-$). Let $v \in \mathbb{Z}$ be an integer such that $u \equiv v \bmod (1-\zeta_p)$. Using the equality of ideals $(1-\zeta_p) = (1-\zeta_p^{-1})$, we see that $\overline{u} \equiv v \bmod (1-\zeta_p)$. If the sign is $-$ then we get $v \equiv -v \bmod (1-\zeta_p)$, implying that $2 \in (1-\zeta_p)$: a contradiction, since the norm of $(1-\zeta_p)$ is odd.

It follows that $\zeta_p^{-g}u = \overline{\zeta_p^{-g}u}$, hence that $\zeta_p^{-g}u$ is real. This is what we needed to prove. □

We prove one more technical lemma that will be useful in a moment.

Lemma 8.8. *For any $\alpha \in \mathbb{Z}[\zeta_p]$, there exists $a \in \mathbb{Z}$ such that $\alpha^p \equiv a$ mod (p).*

Proof. There exists $b \in \mathbb{Z}$ such that $\alpha \equiv b \bmod (1-\zeta_p)$. Then $\alpha^p - b^p = \prod_{i=0}^{p-1}(\alpha - \zeta_p^i b)$. We have $\alpha - \zeta_p^i b \equiv \alpha - b \equiv 0 \bmod (1-\zeta_p)$, so we find $\alpha^p \equiv b^p \bmod (1-\zeta_p)^p$, hence mod $(1-\zeta_p)^{p-1} = (p)$. □

We are now going to use everything we have done so far to prove a special case of the following theorem.

Theorem 8.9. *Let $n \geq 3$ be an integer. Suppose that there exist integers x, y, z such that $x^n + y^n = z^n$. Then $xyz = 0$.*

In fact it was the hope of proving Fermat's Last Theorem that motivated the development of the theory of number fields. One reduces easily to the case where $n = p$ is an odd prime. Several mathematicians in the 1800's had the idea of trying to prove Fermat's Last Theorem by factoring

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

and trying to use unique factorization in the ring $\mathbb{Z}[\zeta_p]$. This led, for example, to Gabriel Lamé's premature announcement to the Académie des Sciences in Paris of a proof of the theorem, which was based on the false assumption that unique factorization always holds. (Lamé did give a correct proof in the special case $p = 7$.) Around the same time, Kummer developed the theory of ideals in order to deal with precisely this difficulty.

The special case we are going to prove is the following.

Theorem 8.10 (Kummer, 1850). *Let p be an odd prime, and suppose that there exist integers x, y, z such that $x^p + y^p = z^p$. Suppose further that p does not divide $\#\,\mathrm{Cl}(\mathcal{O}_K)$, where $K = \mathbb{Q}(\zeta_p)$. Then xyz is divisible by p.*

If p does not divide the class number of $\mathbb{Q}(\zeta_p)$, then we say that p is a regular prime. For historical reasons, the case where $(p, xyz) = 1$ is referred to as the 'first case' of Fermat's Last Theorem. The 'second case' (where p divides xyz) can be treated using similar, but more sophisticated, methods. (Kummer also dealt with the second case under the assumption that p is a regular prime.)

Proof. After making a change of variable, we can assume that $x^p + y^p + z^p = 0$ and that x, y, z have no common factor. Factoring in $\mathbb{Z}[\zeta_p]$, we get the relation

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

hence the equality of ideals

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

We claim that the ideals appearing on the left-hand side of this equality are pairwise coprime. Suppose for contradiction that $P \subset \mathbb{Z}[\zeta_p]$ is a prime ideal dividing both $(x+\zeta_p^i y)$ and $(x+\zeta_p^j y)$ for some $i < j$. Then P contains the element $(\zeta_p^i - \zeta_p^j)y = \zeta_p^i(1-\zeta_p^{j-i})y$, so P divides $(1-\zeta_p)(y)$. If P divides (y), then it must also divide (x) and (z), implying that x, y, z have a common factor, namely the unique prime factor of $N(P)$ – a contradiction. If P divides $(1 - \zeta_p)$, then $P = (1 - \zeta_p)$ and p divides z – another contradiction. Therefore the ideals are pairwise coprime.

It follows that $(x + \zeta_p y) = I^p$ is the p^{th} power of another ideal of $\mathcal{O}_K$. In particular, $[I]^p = 1$ in $\mathrm{Cl}(\mathcal{O}_K)$, hence $[I] = 1$ and I is principal (since p does not divide the order of $\mathrm{Cl}(\mathcal{O}_K)$, by assumption). Let $I = (\delta)$. Then we get a relation $x + \zeta_p y = u\delta^p$ for some $u \in \mathcal{O}_K^\times$, $\delta \in \mathcal{O}_K$, hence $x + \zeta_p y = \zeta_p^g r \delta^p$ for some $g \in \mathbb{Z}$ and some real unit $r \in \mathcal{O}_K^\times \cap \mathbb{R}$. By Lemma 8.8, we can find $a \in \mathbb{Z}$ such that $\delta^p \equiv a \bmod (p)$. Altogether this gives us

$$\zeta_p^{-g}(x + \zeta_p y) \equiv ra \bmod (p)$$

and (taking complex conjugates)

$$\zeta_p^g(x + \zeta_p^{-1} y) \equiv ra \bmod (p),$$

hence

$$\zeta_p^{-g}x + \zeta_p^{1-g}y - \zeta_p^g x - \zeta_p^{g-1}y \equiv 0 \bmod (p).$$

We now try to decide the value of $g \bmod p$. If $g \equiv 0 \bmod p$, then we get $(\zeta_p - \zeta_p^{-1})y \equiv 0 \bmod p$, hence p divides y: a contradiction. Similarly if $g \equiv 1 \bmod p$, then we get $p|x$, another contradiction. Therefore none of $g, -g, g-1, 1-g$ are congruent to 0 modulo p.

It follows that two of $g, -g, g-1, 1-g$ must be congruent to each other modulo p. Indeed, otherwise we can find $\beta \in \mathcal{O}_K$ such that

$$\beta = \zeta_p^{-g}\frac{x}{p} + \zeta_p^{1-g}\frac{y}{p} - \zeta_p^g\frac{x}{p} - \zeta_p^{g-1}\frac{y}{p}.$$

The elements $\zeta_p, \zeta_p^2, \dots, \zeta_p^{p-1}$ form an integral basis for $\mathcal{O}_K$, so this would imply that $\frac{x}{p} \in \mathbb{Z}$, a contradiction.

If $g \equiv -g \bmod p$ or $g \equiv g - 1 \bmod p$, then $g \equiv 0 \bmod p$ or $1 \equiv 0 \bmod p$. Neither of these possibilities can occur. Treating the other cases in the same way, we see that we must have $2g \equiv 1 \bmod p$, leading to a relation

$$p\zeta_p^g\beta = x + \zeta_p y - \zeta_p x - y = (x - y)(1 - \zeta_p).$$

It follows that $x \equiv y \bmod p$. Since our original equation $x^p + y^p + z^p = 0$ was symmetric in x, y, z, the same argument applied to y, z gives $y \equiv z \bmod p$, hence $3x^p \equiv 0 \bmod p$. If $p \neq 3$ then we're done. However, the case $p = 3$ can be treated directly (look at congruences modulo 9 – we leave this as an exercise). □