Chapter 4

Ideal Class Group and Units

We are now interested in understanding two aspects of ring of integers of number fields: "how principal they are" (that is, what is the proportion of principal ideals among all the ideals), and what is the structure of their group of units. For the former task, we will introduce the notion of class number (as the measure of how principal a ring of integers is), and prove that the class number is finite. We will then prove Dirichlet's Theorem for the structure of groups of units. Both results will be derived in the spirit of "geometry of numbers", that is as a consequence of Minkowski's theorem, where algebraic results are proved thanks to a suitable geometrical interpretation (mainly the fact that a ring of integers can be seen as a lattice in $\mathbb{R}^n$ via the n embeddings of its number field).

4.1 Ideal class group

Let K be a number field, and $\mathcal{O}_K$ be its ring of integers. We have seen in Chapter 2 that we can extend the notion of ideal to fractional ideal, and that with this new notion, we have a group structure (Theorem 2.5). Let I_K denote the group of fractional ideals of K. Let P_K denote the subgroup of I_K formed by the principal ideals, that is ideals of the form $\alpha\mathcal{O}_K$, $\alpha \in K^\times$.

Definition 4.1. The ideal class group, denoted by $\mathrm{Cl}(K)$, is

$$\mathrm{Cl}(K) = I_K/P_K.$$

Definition 4.2. We denote by h_K the cardinality $|\mathrm{Cl}_K|$, called the class number.

In particular, if $\mathcal{O}_K$ is a principal ideal domain, then $\mathrm{Cl}(K) = 0$, and $h_K = 1$.

Our goal is now to prove that the class number is finite for ring of integers of number fields. The lemma below is a version of Minkowski's theorem.

Lemma 4.1. *Let Λ be a lattice of $\mathbb{R}^n$. Let $X \subset \mathbb{R}^n$ be a convex, compact set (that is a closed and bounded set since we are in $\mathbb{R}^n$), which is symmetric with respect to 0 (that is, $x \in X \iff -x \in X$). If*

$$\mathrm{Vol}(X) \geq 2^n \mathrm{Vol}(\mathbb{R}^n/\Lambda),$$

then there exists $0 \neq \lambda \in \Lambda$ such that $\lambda \in X$.

Proof. Let us first assume that the inequality is strict: $\mathrm{Vol}(X) > 2^n\mathrm{Vol}(\mathbb{R}^n/\Lambda)$. Let us consider the map

$$\psi : \frac{1}{2}X = \{\frac{x}{2} \in \mathbb{R}^n \mid x \in X\} \to \mathbb{R}^n/\Lambda.$$

If ψ were injective, then

$$\mathrm{Vol}\left(\frac{1}{2}X\right) = \frac{1}{2^n}\mathrm{Vol}(X) \leq \mathrm{Vol}(\mathbb{R}^n/\Lambda)$$

that is $\mathrm{Vol}(X) \leq 2^n\mathrm{Vol}(\mathbb{R}^n/\Lambda)$, which contradicts our assumption. Thus ψ cannot be injective, which means that there exist $x_1 \neq x_2 \in \frac{1}{2}X$ such that $\psi(x_1) = \psi(x_2)$. By symmetry, we have that $-x_2 \in \frac{1}{2}X$, and by convexity of X (that is $(1-t)x + ty \in X$ for $t \in [0,1]$), we have that

$$\left(1 - \frac{1}{2}\right)x_1 + \frac{1}{2}(-x_2) = \frac{x_1 - x_2}{2} \in \frac{1}{2}X.$$

Thus $0 \neq \lambda = x_1 - x_2 \in X$, and $\lambda \in \Lambda$ (since $\psi(x_1 - x_2) = 0$).

Let us now assume that $\mathrm{Vol}(X) = 2^n\mathrm{Vol}(\mathbb{R}^n/\Lambda)$. By what we have just proved, there exists $0 \neq \lambda_\epsilon \in \Lambda$ such that $\lambda_\epsilon \in (1+\epsilon)X$ for all $\epsilon > 0$, since

$$\begin{aligned}\mathrm{Vol}((1+\epsilon)X) &= (1+\epsilon)^n\mathrm{Vol}(X)\\ &= (1+\epsilon)^n 2^n\mathrm{Vol}(\mathbb{R}^n/\Lambda)\\ &> 2^n\mathrm{Vol}(\mathbb{R}^n/\Lambda), \text{ for all } \epsilon > 0.\end{aligned}$$

In particular, if $\epsilon < 1$, then $\lambda_\epsilon \in 2X \cap \Lambda$. The set $2X \cap \Lambda$ is compact and discrete (since Λ is discrete), it is thus finite. Let us now understand what is happening here. On the one hand, we have a sequence λ_ϵ with infinitely many terms since there is one for every $0 < \epsilon < 1$, while on the other hand, those infinitely many terms are all lattice points in $2X$, which only contains finitely many of them. This means that this sequence must converge to a point $0 \neq \lambda \in \Lambda$ which belongs to $(1+\epsilon)X$ for infinitely many $\epsilon > 0$. Thus $\lambda \in \Lambda \cap (\cap_{\epsilon \to 0}(1+\epsilon)X - 0)$. Since X is closed, we have that $\lambda \in X$. □

Let $n = [K : \mathbb{Q}]$ be the degree of K and let (r_1, r_2) be the signature of K. Let $\sigma_1, \ldots, \sigma_{r_1}$ be the r_1 real embeddings of K into $\mathbb{R}$. We choose one of the two embeddings in each pair of complex embeddings, which we denote by

$\sigma_{r_1+1},\dots,\sigma_{r_1+r_2}$. We consider the following map, called canonical embedding of K:

$$\begin{aligned} \sigma: \quad K &\to \quad \mathbb{R}^{r_1}\oplus\mathbb{C}^{r_2}\simeq\mathbb{R}^n \\ \alpha &\mapsto \quad (\sigma_1(\alpha),\dots,\sigma_{r_1}(\alpha),\sigma_{r_1+1}(\alpha),\dots,\sigma_{r_1+r_2}(\alpha)). \end{aligned} \tag{4.1}$$

We have that the image of $\mathcal{O}_K$ by σ is a lattice $\sigma(\mathcal{O}_K)$ in $\mathbb{R}^n$ (we have that $\sigma(\mathcal{O}_K)$ is a free abelian group, which contains a basis of $\mathbb{R}^n$). Let $\alpha_1,\dots,\alpha_n$ be a $\mathbb{Z}$-basis of $\mathcal{O}_K$. Let M be the generator matrix of the lattice $\sigma(\mathcal{O}_K)$, given by

$$\begin{pmatrix} \sigma_1(\alpha_1) & \dots & \sigma_{r_1}(\alpha_1) & \mathrm{Re}(\sigma_{r_1+1}(\alpha_1)) & \mathrm{Im}(\sigma_{r_1+1}(\alpha_1)) & \dots & \mathrm{Re}(\sigma_{r_1+r_2}(\alpha_1)) & \mathrm{Im}(\sigma_{r_1+r_2}(\alpha_1)) \\ \vdots & & & & & & \vdots & \\ \sigma_1(\alpha_n) & \dots & \sigma_{r_1}(\alpha_n) & \mathrm{Re}(\sigma_{r_1+1}(\alpha_n)) & \mathrm{Im}(\sigma_{r_1+1}(\alpha_n)) & \dots & \mathrm{Re}(\sigma_{r_1+r_2}(\alpha_n)) & \mathrm{Im}(\sigma_{r_1+r_2}(\alpha_n)) \end{pmatrix}$$

whose determinant is given by

$$\mathrm{Vol}(\mathbb{R}^n/\sigma(\mathcal{O}_K)) = |\det(M)| = \frac{\sqrt{|\Delta_K|}}{2^{r_2}}.$$

Indeed, we have that $\mathrm{Re}(x)=(x+\bar{x})/2$ and $\mathrm{Im}(x)=(x-\bar{x})/2i$, $x\in\mathbb{C}$, and

$$|\det(M)| = |\det(M')|$$

where M' is given by

$$\begin{pmatrix} \sigma_1(\alpha_1) & \dots & \sigma_{r_1}(\alpha_1) & \sigma_{r_1+1}(\alpha_1) & \frac{\sigma_{r_1+1}(\alpha_1)-\overline{\sigma_{r_1+1}(\alpha_1)}}{2i} & \dots & \sigma_{r_1+r_2}(\alpha_1) & \frac{\sigma_{r_1+r_2}(\alpha_1)-\overline{\sigma_{r_1+r_2}(\alpha_1)}}{2i} \\ \vdots & & & & & & \vdots & \\ \sigma_1(\alpha_n) & \dots & \sigma_{r_1}(\alpha_n) & \sigma_{r_1+1}(\alpha_n) & \frac{\sigma_{r_1+1}(\alpha_n)-\overline{\sigma_{r_1+1}(\alpha_n)}}{2i} & \dots & \sigma_{r_1+r_2}(\alpha_n) & \frac{\sigma_{r_1+r_2}(\alpha_n)-\overline{\sigma_{r_1+r_2}(\alpha_n)}}{2i} \end{pmatrix}.$$

Again, we have that $|\det(M')| = 2^{-r_2}|\det(M'')|$, with M'' given by this time

$$\begin{pmatrix} \sigma_1(\alpha_1) & \dots & \sigma_{r_1}(\alpha_1) & \sigma_{r_1+1}(\alpha_1) & \overline{\sigma_{r_1+1}(\alpha_1)} & \dots & \sigma_{r_1+r_2}(\alpha_1) & \overline{\sigma_{r_1+r_2}(\alpha_1)} \\ \vdots & & & & & & \vdots & \\ \sigma_1(\alpha_n) & \dots & \sigma_{r_1}(\alpha_n) & \sigma_{r_1+1}(\alpha_n) & \overline{\sigma_{r_1+1}(\alpha_n)} & \dots & \sigma_{r_1+r_2}(\alpha_n) & \overline{\sigma_{r_1+r_2}(\alpha_n)} \end{pmatrix},$$

which concludes the proof, since (recall that complex embeddings come by pairs of conjugates)

$$|\det(M)| = 2^{-r_2}|\det(M'')| = 2^{-r_2}\sqrt{\Delta_K}.$$

We are now ready to prove that $\mathrm{Cl}(K)=I_K/P_K$ is finite.

Theorem 4.2. *Let K be a number field with discriminant Δ_K.*

1. *There exists a constant $C = C_{r_1,r_2} > 0$ (which only depends on r_1 and r_2) such that every ideal class (that is every coset of $\mathrm{Cl}(K)$) contains an integral ideal whose norm is at most*

$$C\sqrt{|\Delta_K|}.$$

2. The group Cl(K) *is finite.*

Proof. Recall first that by definition, a non-zero fractional ideal J is a finitely generated $\mathcal{O}_K$-submodule of K, and there exists $\beta \in K^\times$ such that $\beta J \subset \mathcal{O}_K$ (if β_i span J as $\mathcal{O}_K$-module, write $\beta_i = \delta_i/\gamma_i$ and set $\beta = \prod \gamma_i$). The fact that $\beta J \subset \mathcal{O}_K$ exactly means that $\beta \in J^{-1}$ by definition of the inverse of a fractional ideal (see Chapter 2). The idea of the proof consists of, given a fractional ideal J, looking at the norm of a corresponding integral ideal βJ, which we will prove is bounded as claimed.

Let us pick a non-zero fractional ideal I. Since I is a finitely generated $\mathcal{O}_K$-module, we have that $\sigma(I)$ is a lattice in $\mathbb{R}^n$, and so is $\sigma(I^{-1})$, with the property that

$$\mathrm{Vol}(\mathbb{R}^n/\sigma(I^{-1})) = \mathrm{Vol}(\mathbb{R}^n/\sigma(\mathcal{O}_K))\mathrm{N}(I^{-1}) = \frac{\sqrt{|\Delta_K|}}{2^{r_2}\mathrm{N}(I)},$$

where the first equality comes from the fact that the volume is given by the determinant of the generator matrix of the lattice. Now since we have two lattices, we can write the generator matrix of $\sigma(I^{-1})$ as being the generator matrix of $\sigma(\mathcal{O}_K)$ multiplied by a matrix whose determinant in absolute value is the index of the two lattices. Let X be a compact convex set, symmetrical with respect to 0. In order to get a set of volume big enough to use Minkowski theorem, we set a scaling factor

$$\lambda^n = 2^n \frac{\mathrm{Vol}(\mathbb{R}^n/\sigma(I^{-1}))}{\mathrm{Vol}(X)},$$

so that the volume of λX is

$$\mathrm{Vol}(\lambda X) = \lambda^n \mathrm{Vol}(X) = 2^n \mathrm{Vol}(\mathbb{R}^n/\sigma(I^{-1})).$$

By Lemma 4.1, there exists $0 \neq \sigma(\alpha) \in \sigma(I^{-1})$ and $\sigma(\alpha) \in \lambda X$. Since $\alpha \in I^{-1}$, we have that αI is an integral ideal in the same ideal class as I, and

$$\mathrm{N}(\alpha I) = |N_{K/\mathbb{Q}}(\alpha)|\mathrm{N}(I) = |\prod_{i=1}^n \sigma_i(\alpha)|\mathrm{N}(I) \leq M\lambda^n \mathrm{N}(I),$$

where $M = \max_{x\in X} \prod |x_i|$, $x = (x_1, \ldots, x_n)$, so that the maximum over λX gives $\lambda^n M$. Thus, by definition of λ^n, we have that

$$\begin{aligned}
\mathrm{N}(\alpha I) &\leq \frac{2^n \mathrm{Vol}(\mathbb{R}^n/\sigma(I^{-1}))}{\mathrm{Vol}(X)} M\mathrm{N}(I) \\
&= \frac{2^n M}{2^{r_2}\mathrm{Vol}(X)}\sqrt{\Delta_K} \\
&= \underbrace{\frac{2^{r_1+r_2} M}{\mathrm{Vol}(X)}}_{C}\sqrt{\Delta_K}.
\end{aligned}$$

This completes the first part of the proof.

Figure 4.1: Johann Peter Gustav Lejeune Dirichlet (1805-1859)

We are now left to prove that $\mathrm{Cl}(K)$ is a finite group. By what we have just proved, we can find a system of representatives J_i of I_K/P_K consisting of integral ideals J_i, of norm smaller than $C\sqrt{|\Delta_K|}$. In particular, the prime factors of J_i have a norm smaller than $C\sqrt{|\Delta_K|}$. Above the prime numbers $p < C\sqrt{|\Delta_K|}$, there are only finitely many prime ideals (or in other words, there are only finitely many integrals with a given norm). □

4.2 Dirichlet Units Theorem

By abuse of language, we call units of K the units of $\mathcal{O}_K$, that is the invertible elements of $\mathcal{O}_K$. We have seen early on (Corollary 1.11) that units are characterized by their norm, namely units are exactly the elements of $\mathcal{O}_K$ with norm ± 1.

Theorem 4.3. (Dirichlet Units Theorem.) *Let K be a number field of degree n, with signature (r_1, r_2). The group $\mathcal{O}_K^*$ of units of K is the product of the group $\mu(\mathcal{O}_K)$ of roots of unity in $\mathcal{O}_K$, which is cyclic and finite, and a free group on $r_1 + r_2 - 1$ generators. In formula, we have that*

$$\mathcal{O}_K^* \simeq \mathbb{Z}^{r_1+r_2-1} \times \mu(\mathcal{O}_K).$$

The most difficult part of this theorem is actually to prove that the free group has exactly $r_1 + r_2 - 1$ generators. This is nowadays usually proven using Minkowski's theorem. Dirichlet though did not have Minkowski's theorem available: he proved the unit theorem in 1846 while Minkowski developed the geometry of numbers only around the end of the 19th century. He used instead the pigeonhole principle. It is said that Dirichlet got the main idea for his proof while attending a concert in the Sistine Chapel.

Proof. Let $\sigma : K \to \mathbb{R}^{r_1} \times \mathbb{C}^{r_2} \simeq \mathbb{R}^n$ be the canonical embedding of K (see (4.1)). The logarithmic embedding of K is the mapping

$$\begin{array}{rrl}\lambda: & K^* \to & \mathbb{R}^{r_1+r_2} \\ & \alpha \mapsto & (\log|\sigma_1(\alpha)|, \ldots, \log|\sigma_{r_1+r_2}(\alpha))|.\end{array}$$

Since $\lambda(\alpha\beta) = \lambda(\alpha) + \lambda(\beta)$, λ is a homomorphism from the multiplicative group K^* to the additive group of $\mathbb{R}^{r_1+r_2}$.

Step 1. We first prove that the kernel of λ restricted to $\mathcal{O}_K^*$ is a finite group. In order to do so, we prove that if C is a bounded subset of $\mathbb{R}^{r_1+r_2}$, then $C' = \{x \in \mathcal{O}_K^*,\ \lambda(x) \in C\}$ is a finite set. In words, we look at the preimage of a bounded set by the logarithmic embedding (more precisely, at the restriction of the preimage to the units of $\mathcal{O}_K$).

Proof. Since C is bounded, all $|\sigma_i(x)|$, $x \in \mathcal{O}_K^*$, $i = 1, \ldots, n$ belong to some interval say $[a^{-1}, a]$, $a > 1$. Thus the elementary polynomials in the $\sigma_i(x)$ will also belong to some interval of the same form. Now they are the coefficients of the characteristic polynomial of x, which has integer coefficients since $x \in \mathcal{O}_K^*$. Thus there are only finitely many possible characteristic polynomials of elements $x \in C'$, hence only finitely many possible roots of minimal polynomials of elements $x \in C'$, which shows that x can belong to C' for only finitely many x. Now if we set $C = \{0\}$, C' is the kernel $\ker(\lambda)|_{\mathcal{O}_K^*}$ of λ restricted to $\mathcal{O}_K^*$ and is thus finite.

Step 2. We now show that $\ker(\lambda)|_{\mathcal{O}_K^*}$ consists of exactly all the roots of unity $\mu(\mathcal{O}_K)$.

Proof. That it does consist of roots of unity (and is cyclic) is a known property of any subgroup of the multiplicative group of any field. Thus if $x \in \ker(\lambda)|_{\mathcal{O}_K^*}$ then x is a root of unity. Now conversely, suppose that $x^m = 1$. Then x is an algebraic integer, and

$$|\sigma_i(x)|^m = |\sigma_i(x^m)| = |1| = 1$$

so that $|\sigma_i(x)| = 1$, and thus $\log|\sigma_i(x)| = 0$ for all i, showing that $x \in \ker(\lambda)|_{\mathcal{O}_K^*}$.

Step 3. We are now ready to prove that $\mathcal{O}_K^*$ is a finite generated abelian group, isomorphic to $\mu(\mathcal{O}_K) \times \mathbb{Z}^s$, $s \le r_1 + r_2$.

Proof. By Step 1, we know that $\lambda(\mathcal{O}_K^*)$ is a discrete subgroup of $\mathbb{R}^{r_1+r_2}$, that is, any bounded subset of $\mathbb{R}^{r_1+r_2}$ contains only finitely many points of $\lambda(\mathcal{O}_K^*)$. Thus $\lambda(\mathcal{O}_K^*)$ is a lattice in $\mathbb{R}^s$, hence a free $\mathbb{Z}$-module of rank s, for some $s \le r_1 + r_2$. Now by the first isomorphism theorem, we have that

$$\lambda(\mathcal{O}_K^*) \simeq \mathcal{O}_K^*/\mu(\mathcal{O}_K)$$

with $\lambda(x)$ corresponding to the coset $x\mu(\mathcal{O}_K)$. If $x_1\mu(\mathcal{O}_K), \ldots, x_s\mu(\mathcal{O}_K)$ form a basis for $\mathcal{O}_K^*/\mu(\mathcal{O}_K)$ and $x \in \mathcal{O}_K^*$, then $x\mu(\mathcal{O}_K)$ is a finite produce of powers of the x_iG, so x is an element of $\mu(\mathcal{O}_K)$ times a finite product of powers of the x_i. Since the $\lambda(x_i)$ are linearly independent, so are the x_i (provided that the notion of linear independence is translated to a multiplicative setting: $x_1, \ldots, x_s$ are multiplicatively independent if $x_1^{m_1} \cdots x_s^{m_s} = 1$ implies that $m_i = 0$ for all i, from which it follows that $x_1^{m_1}\cdots x_s^{m_s} = x_1^{n_1}\cdots x_s^{n_s}$ implies $m_i = n_i$ for all i). The result follows.

Step 4. We now improve the estimate of s and show that $s \leq r_1 + r_2 - 1$. **Proof.** If x is a unit, then we know that its norm must be ± 1. Then

$$\pm 1 = N(x) = \prod_{i=1}^{n} \sigma_i(x) = \prod_{i=1}^{r_1} \sigma_i(x) \prod_{j=r_1+1}^{r_1+r_2} \sigma_j(x)\overline{\sigma_j(x)}.$$

By taking the absolute values and applying the logarithmic embedding, we get

$$0 = \sum_{i=1}^{r_1} \log|\sigma_i(x)| + \sum_{j=r_1+1}^{r_1+r_2} \log(|\sigma_j(x)||\overline{\sigma_j(x)}|)$$

and $\lambda(x) = (y_1, \ldots, y_{r_1+r_2})$ lies in the hyperplane W whose equation is

$$\sum_{i=1}^{r_1} y_i + 2\sum_{j=r_1+1}^{r_1+r_2} y_j = 0.$$

The hyperplane has dimension $r_1 + r_2 - 1$, so as above, $\lambda(\mathcal{O}_K^*)$ is a free $\mathbb{Z}$-module of rank $s \leq r_1 + r_2 - 1$.

Step 5. We are left with showing that $s = r_1 + r_2 - 1$, which is actually the hardest part of the proof. This uses Minkowski theorem. The proof may come later... one proof can be found in the online lecture of Robert Ash. □

Example 4.1. Consider K an imaginary quadratic field, that is of the form $K = \mathbb{Q}(\sqrt{-d})$, with d a positive square free integer. Its signature is $(r_1, r_2) = (0, 1)$. We thus have that its group of units is given by

$$\mathbb{Z}^{r_1+r_2-1} \times G = G,$$

that is only roots of unity. Actually, we have that the units are the 4rth roots of unity if $K = \mathbb{Q}(\sqrt{-1})$ (that is $\pm 1, \pm i$), the 6th roots of unity if $K = \mathbb{Q}(\sqrt{-3})$ (that is $\pm 1, \pm\zeta_3, \pm\zeta_3^2$), and only ± 1 otherwise.

Example 4.2. For $K = \mathbb{Q}(\sqrt{3})$, we have $(r_1, r_2) = (2, 0)$, thus $r_1 + r_2 - 1 = 1$, and $\mu(\mathcal{O}_K) = \pm 1$. The unit group is given by

$$\mathcal{O}_K^* \simeq \pm(2 + \sqrt{3})^{\mathbb{Z}}.$$

The main definitions and results of this chapter are

- Definition of ideal class group and class number.
- The fact that the class number of a number field is finite.
- The structure of units in a number field (the statement of Dirichlet's theorem)