# Chapter 5 : Units in number rings

There are just two units in $\mathbb{Z}$, namely $\pm 1$. In a general number ring there may be many more units. For example, all the units in the ring $\mathbb{Z}[\mathrm{i}]$ of Gaussian integers are $\pm 1, \pm \mathrm{i}$. There may even be an infinite number of units in a number ring. It will be shown in due course that $\pm(1+\sqrt{2})^n$, $n \in \mathbb{Z}$, are all the units of $\mathbb{Z}[\sqrt{2}]$. For all $n \neq 0$ the absolute values of $\pm(1+\sqrt{2})^n$ are different from 1. $\mathbb{Z}[\sqrt{2}]$ is a PID. So we can think of factorizations in $\mathfrak{O} := \mathfrak{O}_{\mathbb{Z}[\sqrt{2}]}$ as element-wise factorizations. To start with we fix a set of pairwise non-associate prime *elements* of $\mathfrak{O}$. Every non-zero element of $\mathfrak{O}$ admits a factorization $up_1^{e_1} \cdots p_r^{e_r}$ for prime 'representatives' $p_i$ and for a unit $u$ of the form $\pm(1+\sqrt{2})^n$. Thus in order to complete the picture of factorization we need machineries to handle the units in a number ring.

The biggest theorem in this chapter is *Dirichlet's unit theorem* that explains the group structure of $\mathfrak{O}_K^*$ for a number field $K$. In the process I will introduce a branch of mathematics (now more or less dead) known as the *Geometry of numbers*. The proof of Dirichlet's unit theorem is not very illuminating, but its implications are profound. Let's go ahead!

## 5.1 Some basic properties of units

As done in the last few chapters, let us fix the letter $K$ to denote an arbitrary number field of degree $d$ and signature $(r_1, r_2)$. We have $d = r_1 + 2r_2$. Let us name the real embeddings of $K$ as $\sigma_1, \ldots, \sigma_{r_1}$ and the properly complex embeddings of $K$ as $\sigma_{r_1+1}, \ldots, \sigma_{r_1+r_2}, \bar{\sigma}_{r_1+1}, \ldots, \bar{\sigma}_{r_1+r_2}$, where bar denotes complex conjugate. We will assume that $\sigma_1$ is the identity embedding of $K$. The set of units in $\mathfrak{O} := \mathfrak{O}_K$ will be denoted by $\mathfrak{U} := \mathfrak{U}_K := \mathfrak{O}^*$. We know that $\mathfrak{U}$ is an (Abelian) group under (complex) multiplication. Our basic aim in this chapter is to reveal the structure of the group $\mathfrak{U}$.

Every Abelian group is a $\mathbb{Z}$-module and, if finitely generated and not free, contains torsion elements, i.e., (non-identity) elements of finite order $> 1$.[1] $\mathfrak{U}$ always contains the element $-1$ of order 2. The torsion subgroup of $\mathfrak{U}$ is denoted by $\mathfrak{R} := \mathfrak{R}_K := \mathfrak{U}_{\mathrm{tors}}$. Thus $\mathfrak{U} \cong \mathfrak{R} \times \mathfrak{G}$, where $\mathfrak{G}$ is a torsion-free group. I will show that $\mathfrak{R}$ is a finite group and that $\mathfrak{G}$ is finitely generated and hence free, i.e., $\mathfrak{G} \cong \mathbb{Z}^\rho$ for some $\rho \in \mathbb{Z}_+$. From Dirichlet's unit theorem (that I will prove in Section 5.3) it follows that $\rho = r_1 + r_2 - 1$. Thus $\mathfrak{G}$ has a $\mathbb{Z}$-basis consisting of $\rho$ elements, say $\xi_1, \ldots, \xi_\rho$, and every unit of $\mathfrak{U}$ can be uniquely expressed as $\omega \xi_1^{e_1} \cdots \xi_\rho^{e_\rho}$, where $\omega$ is a root of unity and $e_i \in \mathbb{Z}$.

The following characterization of units of $\mathfrak{O}$ is very important.

**5.1 Proposition** Let $\alpha \in \mathfrak{O}$. Then $\alpha \in \mathfrak{U}$, if and only if $\mathrm{N}(\alpha) = \pm 1$.

*Proof* [if] Let $\beta := \prod_{i=2}^{d} \sigma_i(\alpha)$. Then $\alpha\beta = \mathrm{N}(\alpha) = \pm 1$. Now $\beta$, being a product of algebraic integers, is in $\mathbb{A}$ and, being equal to $\pm\alpha^{-1}$, is in $K$, i.e., $\beta \in \mathfrak{O} = \mathbb{A} \cap K$. Therefore $\alpha \in \mathfrak{U}$.

[only if] $\alpha \mid 1$ in $\mathfrak{O}$, so that $\mathrm{N}(\alpha) \mid \mathrm{N}(1) = 1$ in $\mathbb{Z}$. Thus we must have $\mathrm{N}(\alpha) = \pm 1$. ◀

Note that this proposition does not prevent an element of $K \setminus \mathfrak{O}$ from having norm $\pm 1$ (Exercise 5.1.1). Also for $\alpha \in \mathfrak{U}$ the condition $\mathrm{N}(\alpha) = \pm 1$ does not necessarily imply that the absolute value $|\alpha|$ is 1. I will now prove that the number of units of $\mathfrak{O}$ all the conjugates of which have bounded absolute values is finite. More generally:

**5.2 Proposition** Let $B$ be a given positive integer (or real number). The number of elements $\alpha \in \mathfrak{O}$ for which $|\sigma(\alpha)| \leqslant B$ for *every* complex embedding $\sigma$ of $K$ is finite.

[1] Every finitely generated torsion-free module over a PID is free.

*Proof* For such an element $\alpha$ the minimal polynomial $\mathrm{minpoly}_{\alpha,\mathbb{Q}}(X) \in \mathbb{Z}[X]$ has degree dividing $d$ and has all coefficients with absolute values $\leqslant \max(1, dB, \binom{d}{2}B^2, \binom{d}{3}B^3, \ldots, B^d)$. Since there are only finitely many such polynomials and each such polynomial has $\leqslant d$ roots, the result follows. ◀

**5.3 Proposition** The torsion subgroup $\mathfrak{R}$ of $\mathfrak{U}$ is finite. In particular, $\mathfrak{R}$ is cyclic.

*Proof* Let $\alpha \in \mathfrak{R}$ and let $m$ be the smallest positive integer for which $\alpha^m = 1$, i.e., $\alpha$ is a primitive $m$-th root of unity. Since $\mathbb{Q}(\alpha)$ is a subfield of $K$, it follows that $\phi(m) \mid d$. This means that $m$ can assume only finitely many values. Since for each such $m$ the number of primitive $m$-th roots of unity is finite (namely $\phi(m)$), $\mathfrak{R}$ is finite. The last statement of the proposition follows immediately from Theorem 1.77. ◀

Thus $\mathfrak{R} \cong \mu_m$ for some $m \in \mathbb{N}$ with $\phi(m) \mid d$, where $\mu_m$ denotes the (multiplicative) group of $m$-th roots of unity. Since for each $u \in \mathfrak{R}$ we have $-u \in \mathfrak{R}$, it is evident that $m = |\mathfrak{R}|$ is even. Note that we may have $\phi(m) < d$ (Exercise 5.1.2).

**Exercises for Section 5.1**

**1.** **(a)** Give an example of a number field $K$ and an element $\alpha \in K \setminus \mathfrak{O}_K$ with $\mathrm{N}(\alpha) = 1$.
**(b)** Give an example of a number field $K$ and an element $\alpha \in K \setminus \mathfrak{O}_K$ with $\mathrm{N}(\alpha) = -1$.
**(c)** Give an example of a number field $K$ and an element $\alpha \in \mathfrak{U}_K$ with $|\alpha| > 1$.
**(d)** Give an example of a number field $K$ and an element $\alpha \in \mathfrak{U}_K$ with $|\alpha| < 1$.

**2.** **(a)** Find all the even integers $m$ satisfying $\phi(m) \mid 2$.
**(b)** For each $m$ of Part (a) give an example of a quadratic number field $K$ for which $\mathfrak{R}_K \cong \mu_m$.

**3.** **(a)** Let $K$ be a real quadratic number field. Show that the only units of $\mathfrak{O}_K$ of finite order are $\pm 1$.
**(b)** Let $K$ be a number field of odd degree $d$. Show that the only units of $\mathfrak{O}_K$ of finite order are $\pm 1$.

**4.** Let $K$ be a number field, $\mathfrak{O} := \mathfrak{O}_K$, $\mathfrak{U} := \mathfrak{U}_K$, $\mathfrak{R} := \mathfrak{R}_K$ and $\alpha \in \mathfrak{O}$. Show that:
**(a)** $\alpha \in \mathfrak{U}$, if and only if $\mathrm{minpoly}_{\alpha,\mathbb{Q}}(0) = \pm 1$.
**(b)** $\alpha \in \mathfrak{R}$, if and only if $|\sigma(\alpha)| = 1$ for every complex embedding $\sigma$ of $K$.

**5.** Let $p$ be an odd prime, $\omega_p$ a primitive $p$-th root of unity and $K := \mathbb{Q}(\omega_p)$ (so that $\mathfrak{O}_K = \mathbb{Z}[\omega_p]$). Show that the torsion subgroup of $\mathfrak{U}_K$ is $\mathfrak{R}_K = \{\pm\omega_p^i \mid i = 0, 1, \ldots, p-1\}$. In particular, $\mathfrak{R}_K \cong \mu_{2p}$.

## 5.2 Lattices and Minkowski's theorem

As mentioned earlier, the main goal of this chapter is to prove that the torsion-free part of $\mathfrak{U} = \mathfrak{U}_K$ is of finite rank equal to $r_1 + r_2 - 1$, where $(r_1, r_2)$ is the signature of the field $K$. A popular approach to achieve that is by using the Geometry of Numbers, a field of mathematics introduced by Minkowski. One starts with the following basic definition.

**5.4 Definition** Let $m, n \in \mathbb{Z}_+$, $m \leqslant n$ and $\{\mathbf{b}_1, \ldots, \mathbf{b}_m\}$ a set of $m$ $\mathbb{R}$-linearly independent vectors of $\mathbb{R}^n$.[2] The set $\Lambda$ of all $\mathbb{Z}$-linear combinations of $\mathbf{b}_1, \ldots, \mathbf{b}_m$, i.e., the set of elements of $\mathbb{R}^n$ of the form

$$u_1\mathbf{b}_1 + \cdots + u_m\mathbf{b}_m, \quad u_i \in \mathbb{Z},$$

[2]Let us plan to denote vectors of $\mathbb{R}^n$ by lower-case bold-face Roman letters. A vector $\mathbf{b}$ of $\mathbb{R}^n$ is an $n$-tuple $(b_1, b_2, \ldots, b_n)$ with each $b_i \in \mathbb{R}$. It is often convenient to treat $\mathbf{b}$ as the $1 \times n$ matrix $\begin{pmatrix} b_1 & b_2 & \cdots & b_n \end{pmatrix}$, i.e., as an $n$-dimensional row vector.

is called an $m$-dimensional lattice in $\mathbb{R}^n$ with a basis $(\mathbf{b}_1, \ldots, \mathbf{b}_m)$. If $m = n$, then $\Lambda$ is called a full lattice of $\mathbb{R}^n$.

**5.5 Example** $\mathbb{Z}^n = \{(u_1, \ldots, u_n) \mid u_i \in \mathbb{Z}\}$ is a full lattice of $\mathbb{R}^n$ with the canonical basis $(\mathbf{e}_1, \ldots, \mathbf{e}_n)$, where $\mathbf{e}_i := (0, \ldots, 0, 1, 0 \ldots, 0)$ (1 in the $i$-th position). A lattice need not have a unique basis. For example, $\mathbb{Z}^2$ has the canonical basis $((1,0),(1,1))$. $((2,3),(1,2))$ is also a basis of $\mathbb{Z}^2$, since $(1,0) = 2(2,3) - 3(1,2)$ and $(0,1) = -(2,3) + 2(1,2)$.

**5.6 Proposition** Let $\Lambda$ be a full lattice of $\mathbb{R}^n$ with basis $(\mathbf{b}_1, \ldots, \mathbf{b}_n)$. Then the vectors $\mathbf{c}_1, \ldots, \mathbf{c}_n \in \mathbb{R}^n$ constitute a basis of $\Lambda$, if and only if $(\, c_{ij} \,) = T\,(\, b_{ij} \,)$ for some $n \times n$ matrix $T$ with integer entries and of determinant $\pm 1$.[3] (Here $b_{ij}$ (resp. $c_{ij}$) refers to the $j$-th component of $\mathbf{b}_i$ (resp. $\mathbf{c}_i$).)

*Proof* [if] Since $T$ has integer entries, each $\mathbf{c}_i \in \Lambda$. Furthermore since $\det T = \pm 1$, $T$ is invertible and $T^{-1}$ has integer entries. Choose any $\mathbf{x} \in \Lambda$. Since $(\mathbf{b}_1, \ldots, \mathbf{b}_n)$ is a basis for $\Lambda$, we have integers $u_1, \ldots, u_n$ with $\mathbf{x} = u_1\mathbf{b}_1 + u_2\mathbf{b}_2 + \cdots + u_n\mathbf{b}_n = (\, u_1 \quad u_2 \quad \cdots \quad u_n \,)\,(\, b_{ij} \,) = (\, u_1 \quad u_2 \quad \cdots \quad u_n \,)\, T^{-1}\,(\, c_{ij} \,) = (\, v_1 \quad v_2 \quad \cdots \quad v_n \,)\,(\, c_{ij} \,) = v_1\mathbf{c}_1 + v_2\mathbf{c}_2 + \cdots + v_n\mathbf{c}_n$, where $v_1, \ldots, v_n$ are integers. Thus $\mathbf{x}$ is in the $\mathbb{Z}$-linear span of $\mathbf{c}_1, \ldots, \mathbf{c}_n$. Also $\det(\, c_{ij} \,) = \pm \det(\, b_{ij} \,) \neq 0$, so that $\mathbf{c}_1, \ldots, \mathbf{c}_n$ are $\mathbb{R}$-linearly independent.

[only if] Clearly each $\mathbf{c}_i$ belongs to $\Lambda$. Since $\mathbf{b}_1, \ldots, \mathbf{b}_n$ form a basis of $\Lambda$, every $\mathbf{c}_i$ can be written as a $\mathbb{Z}$-linear combination of $\mathbf{b}_1, \ldots, \mathbf{b}_n$. Therefore there exists a matrix $T$ with integer entries satisfying $(\, c_{ij} \,) = T\,(\, b_{ij} \,)$. Now $(\mathbf{c}_1, \ldots, \mathbf{c}_n)$ is also a basis of $\Lambda$ and $\mathbf{b}_i$ are in $\Lambda$. Hence there exists another matrix $S$ with integer entries satisfying $(\, b_{ij} \,) = S\,(\, c_{ij} \,)$. Thus $(\, b_{ij} \,) = ST\,(\, b_{ij} \,)$. Taking determinants and observing the facts that $(\, b_{ij} \,)$ is nonsingular ($\mathbf{b}_1, \ldots, \mathbf{b}_d$ being linearly independent) and that $\det S, \det T \in \mathbb{Z}$ complete the proof. ◀

**5.7 Corollary** Let $\Lambda$ be a full lattice of $\mathbb{R}^n$ and let $(\mathbf{b}_1, \ldots, \mathbf{b}_n)$ and $(\mathbf{c}_1, \ldots, \mathbf{c}_n)$ be two bases of $\Lambda$. Then $|\det(\, b_{ij} \,)| = |\det(\, c_{ij} \,)|$. Thus the value $|\det(\, b_{ij} \,)|$ is independent of the choice of the basis $(\mathbf{b}_1, \ldots, \mathbf{b}_n)$ of $\Lambda$ and is an invariant of $\Lambda$. ◀

**5.8 Definition** Let $\Lambda$ be a full lattice of $\mathbb{R}^n$ with basis $(\mathbf{b}_1, \ldots, \mathbf{b}_n)$. The value $|\det(\, b_{ij} \,)| \in \mathbb{R}$ is called the determinant of $\Lambda$ and is denoted by $\Delta(\Lambda)$. It is also common to call $\Delta(\Lambda)$ the discriminant of $\Lambda$.

It follows from elementary calculus that $\Delta(\Lambda)$ is equal to the volume[4] of each of the fundamental parallelepiped

$$\mathcal{P}_{\mathbf{x}} := \left\{ \mathbf{x} + \sum_{i=1}^{n} u_i\mathbf{b}_i \mid 0 \leqslant u_i < 1 \right\}, \quad \mathbf{x} \in \Lambda.$$

The fundamental parallelepipeds $\mathcal{P}_{\mathbf{x}}$ are pairwise disjoint and cover $\mathbb{R}^n$ (i.e., $\mathbb{R}^n = \bigcup_{\mathbf{x}\in\Lambda} \mathcal{P}_{\mathbf{x}}$).

I will now prove a theorem due to Minkowski, which states that a subset $\Omega \subseteq \mathbb{R}^n$ meeting certain restrictions (that I will state later) will have to contain a non-zero point of a full lattice $\Lambda$ of $\mathbb{R}^n$, whenever the volume of $\Omega$ is 'sufficiently large' compared to $\Delta(\Lambda)$. I start with the following lemma:

[3] A square matrix with integer entries and with determinant $\pm 1$ is often called a unimodular matrix.

[4] Here I use the term 'volume' rather intuitively without making an attempt to define it rigorously. A cynical reader may think of volume as a measure, for example, the Lebesgue measure, on $\mathbb{R}^n$, provided that this new term continues to make sense to him/her. For the time being it is sufficient to concentrate on regions $\Omega \subseteq \mathbb{R}^n$ whose 'volumes' can be calculated using the standard techniques of integral calculus.

**5.9 Lemma** [Blichfeldt's theorem][5] Let $n \in \mathbb{N}$, $\Lambda$ a full lattice in $\mathbb{R}^n$, and let $\Omega \subseteq \mathbb{R}^n$ have volume $> \Delta(\Lambda)$. Then there exist two distinct points $\mathbf{x}_1$ and $\mathbf{x}_2$ in $\Omega$ such that $\mathbf{x}_1 - \mathbf{x}_2 \in \Lambda$.

*Proof* For each $\mathbf{y} \in \Lambda$ define $\Omega_\mathbf{y} := \Omega \cap \mathcal{P}_\mathbf{y}$. Then $\Omega_\mathbf{y}$ are pairwise disjoint and cover $\Omega$. Thus the volume of $\Omega$ is the sum $\sum_{\mathbf{y}\in\Lambda} \mathrm{vol}(\Omega_\mathbf{y})$. The translation of each $\Omega_\mathbf{y}$ by $-\mathbf{y}$ shifts the region $\Omega_\mathbf{y}$ in the parallelepiped $\mathcal{P}_\mathbf{y}$ to the region $\Omega_\mathbf{y} - \mathbf{y}$ inside the parallelepiped $\mathcal{P} := \mathcal{P}_\mathbf{0}$. We clearly see that these translations do not change the volumes of $\Omega_\mathbf{y}$ and thus the volume of $\Omega$ equals $\sum_{\mathbf{y}\in\Lambda} \mathrm{vol}(\Omega_\mathbf{y} - \mathbf{y})$. Since $\Delta(\Lambda) = \mathrm{vol}(\mathcal{P}) < \mathrm{vol}(\Omega)$, it follows that we have some $\mathbf{y}_1 \neq \mathbf{y}_2 \in \Lambda$ for which there is an overlap of $\Omega_{\mathbf{y}_1} - \mathbf{y}_1$ and $\Omega_{\mathbf{y}_2} - \mathbf{y}_2$, i.e., there are points $\mathbf{x}_1 \in \Omega_{\mathbf{y}_1}$ and $\mathbf{x}_2 \in \Omega_{\mathbf{y}_2}$ with $\mathbf{x}_1 - \mathbf{y}_1 = \mathbf{x}_2 - \mathbf{y}_2$. But then $\mathbf{x}_1 - \mathbf{x}_2 = \mathbf{y}_2 - \mathbf{y}_1 \in \Lambda$. ◀

The main theorem of this section is an easy corollary to the last lemma and the following definitions.

**5.10 Definition** Let $\Omega$ be a subset of $\mathbb{R}^n$. We call $\Omega$ symmetric (about the origin), if $-\mathbf{x} \in \Omega$ whenever $\mathbf{x} \in \Omega$. $\Omega$ is called convex, if $\mathbf{x}_1, \mathbf{x}_2 \in \Omega$ implies that the line segment joining $\mathbf{x}_1$ and $\mathbf{x}_2$ is completely contained in $\Omega$, i.e., $t\mathbf{x}_1 + (1-t)\mathbf{x}_2 \in \Omega$ for all $t$, $0 \leqslant t \leqslant 1$.

**5.11 Theorem** [Minkowski's convex-body theorem] Let $n \in \mathbb{N}$, $\Lambda$ a full lattice in $\mathbb{R}^n$ and $\Omega$ a symmetric and convex subset of $\mathbb{R}^n$ with $\mathrm{vol}(\Omega) > 2^n\Delta(\Lambda)$. Then $\Omega$ contains a non-zero point of $\Lambda$.

*Proof* Define the region $\frac{1}{2}\Omega := \{\mathbf{x} \in \mathbb{R}^n \mid 2\mathbf{x} \in \Omega\}$. Then $\mathrm{vol}(\frac{1}{2}\Omega) = \frac{1}{2^n}\mathrm{vol}(\Omega) > \Delta(\Lambda)$. By Lemma 5.9 there exist distinct points $\mathbf{y}_1, \mathbf{y}_2 \in \frac{1}{2}\Omega$ such that $\mathbf{0} \neq \mathbf{x} := \mathbf{y}_1 - \mathbf{y}_2 \in \Lambda$. Now $\mathbf{y}_1 = \frac{1}{2}\mathbf{x}_1$ and $\mathbf{y}_2 = \frac{1}{2}\mathbf{x}_2$ for some $\mathbf{x}_1, \mathbf{x}_2 \in \Omega$. Thus $\mathbf{x} = \frac{1}{2}\mathbf{x}_1 - \frac{1}{2}\mathbf{x}_2 = \frac{1}{2}\mathbf{x}_1 + (1-\frac{1}{2})(-\mathbf{x}_2)$. Since $\Omega$ is symmetric, $-\mathbf{x}_2 \in \Omega$. Moreover since $\Omega$ is convex, $\mathbf{x} \in \Omega$. ◀

It can be shown that if the convex and symmetric region $\Omega$ of Theorem 5.11 is also compact (i.e., closed and bounded), then the condition $\mathrm{vol}(\Omega) \geqslant 2^n\Delta(\Lambda)$ is sufficient for $\Omega$ to contain a non-zero point of $\Lambda$. If $\Omega$ is not compact, we need the strict inequality $\mathrm{vol}(\Omega) > 2^n\Delta(\Lambda)$.

Though Minkowski's theorem is fairly intuitive and straightforward to prove, it has important consequences. In the next section we will prove Dirichlet's unit theorem using Theorem 5.11. The Minkowski bound (Equation 4.4) can also be derived following this line of thought (Exercise 5.3.8). For the time being let us concentrate on an equivalent characterization of a lattice.

**5.12 Definition** The $n$-dimensional closed ball (or sphere) $\mathrm{B}_{n,r}$ of radius $r$ and with center at the origin is defined as

$$\mathrm{B}_{n,r} := \left\{ \mathbf{x} = (x_1, \ldots, x_n) \in \mathbb{R}^n \;\middle|\; |\mathbf{x}| = \sqrt{x_1^2 + \cdots + x_n^2} \leqslant r \right\}.$$

If $n$ is understood from the context, one can abbreviate $\mathrm{B}_{n,r}$ as $\mathrm{B}_r$. A subset $\Lambda$ of $\mathbb{R}^n$ is called discrete, if $\Lambda \cap \mathrm{B}_{n,r}$ is a finite set for every positive real number $r$.

**5.13 Proposition** A subset $\Lambda$ of $\mathbb{R}^n$ is a lattice, if and only if $\Lambda$ is a discrete (additive) subgroup of $\mathbb{R}^n$.

*Proof* [if] Let $\Lambda$ be a discrete additive subgroup of $\mathbb{R}^n$ and let $\{\mathbf{c}_1, \ldots, \mathbf{c}_m\} \subseteq \Lambda$, $0 \leqslant m \leqslant n$, be a maximal set of $\mathbb{R}$-linearly independent elements of $\Lambda$. If $m = 0$, then $\Lambda = \{(0, 0, \ldots, 0)\}$ is the 0-dimensional lattice. So let us assume that $m \geqslant 1$ and prove by induction on $m$ that $\Lambda$ is an $m$-dimensional lattice of $\mathbb{R}^n$.

[5] The original version of Blichfeldt's theorem (1914) states that if a bounded region $\Omega \subseteq \mathbb{R}^2$ is of area $> m$ (for some $m \in \mathbb{N}$), then $\Omega$ can be so translated as to contain at least $m+1$ points of the integer lattice $\mathbb{Z}^2$. This theorem can be generalized for any lattice and for any dimension $n$. We are here interested in the special case: $m = 1$.

If $m = 1$, every $\mathbf{x} \in \Lambda$ is a *real* multiple of $\mathbf{c}_1$. Since $\Lambda$ is discrete by hypothesis, it is easy to see that the set $\{a \in \mathbb{R} \mid a > 0, a\mathbf{c}_1 \in \Lambda\}$ contains a minimal element, say $b$. Define $\mathbf{b}_1 := b\mathbf{c}_1 \in \Lambda$. I claim that every $\mathbf{x} \in \Lambda$ is an *integral* multiple of $\mathbf{b}_1$. Clearly the claim holds for $\mathbf{x} = \mathbf{0}$. So assume $\mathbf{x} \neq \mathbf{0}$ and write $\mathbf{x} = a\mathbf{b}_1$ for some $a \in \mathbb{R} \setminus \{0\}$. Writing $a = \lfloor a \rfloor + r$ for some $r$, $0 \leqslant r < 1$, we get $rb\mathbf{c}_1 = r\mathbf{b}_1 = (a - \lfloor a \rfloor)\mathbf{b}_1 = \mathbf{x} - \lfloor a \rfloor \mathbf{b}_1 \in \Lambda$. The choice of $b$ then forces $rb = 0$, i.e., $r = 0$, i.e., $\mathbf{x} = \lfloor a \rfloor \mathbf{b}_1$. It follows that $\Lambda = \mathbb{Z}\mathbf{b}_1$, i.e., $\Lambda$ is a one-dimensional lattice of $\mathbb{R}^n$.

Now assume that $m > 1$ and that the result holds for $m-1$. Since $\Lambda' := \Lambda \cap V'$, $V' := \mathbb{R}\mathbf{c}_1 + \cdots + \mathbb{R}\mathbf{c}_{m-1}$, is a discrete subgroup of $\mathbb{R}^n$ containing $\{\mathbf{c}_1, \ldots, \mathbf{c}_{m-1}\}$ as a maximal subset of $\mathbb{R}$-linearly independent elements, by the induction hypothesis $\Lambda'$ is an $(m-1)$-dimensional lattice with some basis $(\mathbf{b}_1, \ldots, \mathbf{b}_{m-1})$. But then $(\mathbf{b}_1, \ldots, \mathbf{b}_{m-1}, \mathbf{c}_m)$ is an $\mathbb{R}$-basis of $V := \mathbb{R}\mathbf{c}_1 + \cdots + \mathbb{R}\mathbf{c}_m$. Therefore every $\mathbf{x} \in \Lambda \subseteq V$ can be written as $\mathbf{x} = a_1\mathbf{b}_1 + \cdots + a_{m-1}\mathbf{b}_{m-1} + a_m\mathbf{c}_m$ with $a_i \in \mathbb{R}$. The set $\{a_m \in \mathbb{R} \mid a_m > 0, a_1\mathbf{b}_1 + \cdots + a_{m-1}\mathbf{b}_{m-1} + a_m\mathbf{c}_m \in \Lambda \text{ with } |a_i| < 1 \text{ for } 1 \leqslant i < m\}$ is non-empty (since $\mathbf{c}_m \in \Lambda$, so that 1 belongs to this set) and by the discreteness property of $\Lambda$ contains a minimal element, say $b$. Choose any $\mathbf{b}_m := b_1\mathbf{b}_1 + \cdots + b_{m-1}\mathbf{b}_{m-1} + b\mathbf{c}_m \in \Lambda$. Clearly $\mathbf{b}_1, \ldots, \mathbf{b}_m$ are $\mathbb{R}$-linearly independent and span $V$.

What remains is to show that every $\mathbf{x} \in \Lambda$ is a $\mathbb{Z}$-linear combination of $\mathbf{b}_1, \ldots, \mathbf{b}_m$. This being obvious for $\mathbf{x} = \mathbf{0}$, take $\mathbf{x} \neq \mathbf{0}$ and write $\mathbf{x} = a_1\mathbf{b}_1 + \cdots + a_m\mathbf{b}_m$ for some real $a_i = \lfloor a_i \rfloor + r_i$, $0 \leqslant r_i < 1$. Consider $\mathbf{y} := \mathbf{x} - (\lfloor a_1 \rfloor \mathbf{b}_1 + \cdots + \lfloor a_m \rfloor \mathbf{b}_m) = r_1\mathbf{b}_1 + \cdots + r_m\mathbf{b}_m = s_1\mathbf{b}_1 + \cdots + s_{m-1}\mathbf{b}_{m-1} + r_m b\mathbf{c}_m \in \Lambda$, where $s_i := r_i + r_m b_i$ for $1 \leqslant i \leqslant m-1$. But then $(s_1 - \lfloor s_1 \rfloor)\mathbf{b}_1 + \cdots + (s_{m-1} - \lfloor s_{m-1} \rfloor)\mathbf{b}_{m-1} + r_m b\mathbf{c}_m \in \Lambda$ and so the choice of $b$ forces $r_m = 0$, i.e., $\mathbf{y} = r_1\mathbf{b}_1 + \cdots + r_{m-1}\mathbf{b}_{m-1} \in \Lambda \cap V' = \Lambda'$. Therefore $\mathbf{y} = c_1\mathbf{b}_1 + \cdots + c_{m-1}\mathbf{b}_{m-1}$ for some *integers* $c_1, \ldots, c_{m-1}$ and consequently $\mathbf{x} = (c_1 + \lfloor a_1 \rfloor)\mathbf{b}_1 + \cdots + (c_{m-1} + \lfloor a_{m-1} \rfloor)\mathbf{b}_{m-1} + \lfloor a_m \rfloor \mathbf{b}_m \in \mathbb{Z}\mathbf{b}_1 + \cdots + \mathbb{Z}\mathbf{b}_m$.

[only if] A lattice $\Lambda$ of $\mathbb{R}^n$ is clearly an additive subgroup of $\mathbb{R}^n$. Let $m$ be the dimension of $\Lambda$. If $m = 0$, then $\Lambda = \{(0, 0, \ldots, 0)\}$ is evidently discrete. So assume that $m \geqslant 1$ and let $(\mathbf{b}_1, \ldots, \mathbf{b}_m)$ be a basis of $\Lambda$. Then $\mathbf{b}_1, \ldots, \mathbf{b}_m$ span an $m$-dimensional $\mathbb{R}$-subspace $V$ of $\mathbb{R}^n$. The bijective linear transformation $\tau : V \to \mathbb{R}^m$, $\mathbf{b}_i \mapsto \mathbf{e}_i$, (where $(\mathbf{e}_1, \ldots, \mathbf{e}_m)$ is the *canonical basis* for $\mathbb{R}^m$) maps $\Lambda$ to the full lattice $\Lambda' := \tau(\Lambda) = \mathbb{Z}^m = \{(u_1, \ldots, u_m) \mid u_i \in \mathbb{Z}\}$ of $\mathbb{R}^m$. For every positive real number $r'$ the condition $(u_1, \ldots, u_m) \in \mathrm{B}_{m,r'}$ implies that each $|u_i| \leqslant r'$, i.e., $\Lambda' \cap \mathrm{B}_{m,r'}$ is finite. Now choose a real number $r > 0$ and look at $|\Lambda \cap \mathrm{B}_{n,r}| = |\Lambda \cap V \cap \mathrm{B}_{n,r}| = |\tau(\Lambda \cap V \cap \mathrm{B}_{n,r})| = |\Lambda' \cap \tau(V \cap \mathrm{B}_{n,r})|$. Since the region $\tau(V \cap \mathrm{B}_{n,r})$ is bounded[6] in $\mathbb{R}^m$, we can choose a real $r' > 0$ such that $\tau(V \cap \mathrm{B}_{n,r}) \subseteq \mathrm{B}_{m,r'}$. But then $|\Lambda \cap \mathrm{B}_{n,r}| \leqslant |\Lambda' \cap \mathrm{B}_{m,r'}| < \infty$. ◀

**5.14 Example** For any irrational (real number) $\xi$ the elements 1 and $\xi$ are linearly independent over $\mathbb{Z}$ and hence $\Lambda := \mathbb{Z} + \mathbb{Z}\xi$ is a free Abelian group of rank 2. But $\Lambda$ is not a lattice of $\mathbb{R}$, since 1 and $\xi$ are linearly dependent over $\mathbb{R}$. The set $\{u + v\xi \mid u, v \in \mathbb{Z}\}$ is dense in $\mathbb{R}$ and so $\Lambda$ is not a discrete set.

**Exercises for Section 5.2**

**1.** **(a)** Show that $2^n\Delta(\Lambda)$ is the best possible bound on the volume of $\Omega$ in Minkowski's convex-body theorem.

**(b)** Show that symmetry and convexity are both individually necessary for Minkowski's convex-body theorem.

**2.** Let $\Lambda$ be a full lattice of $\mathbb{R}^n$ and $\Omega$ a symmetric convex subset of $\mathbb{R}^n$ with $\mathrm{vol}(\Omega) > 2^n\Delta(\Lambda)$. Show that $\Omega$ contains at least three points of $\Lambda$.

**3.** Let $\Lambda$ be a full lattice of $\mathbb{R}^n$. Show that there exists a non-zero point $\mathbf{x} = (x_1, \ldots, x_n) \in \Lambda$ with $|x_i| \leqslant \sqrt[n]{\Delta(\Lambda)}$ for all $i = 1, \ldots, n$. Conclude that the shortest non-zero vector of $\Lambda$ is of length $\leqslant \sqrt{n}\sqrt[n]{\Delta(\Lambda)}$.

[6] The map $\tau$ is a homeomorphism of $V$ onto $\mathbb{R}^m$ and hence preserves compactness. In $\mathbb{R}^k$, $k \in \mathbb{N}$, the term 'compact' is synonymous with 'closed and bounded'.

**4.** [Minkowski's linear forms theorem] Let $A := (a_{ij})$ be an $n \times n$ matrix with real entries and with $\det A \neq 0$. Further let $\lambda_1, \ldots, \lambda_n$ be positive real numbers with $\lambda_1 \cdots \lambda_n > |\det A|$. Show that there exist integers $x_1, \ldots, x_n$, not all zero, such that $|\sum_{j=1}^{n} a_{ij}x_j| < \lambda_i$ for all $i = 1, \ldots, n$.

**5.** As a sample application of Minkowski's convex-body theorem this exercise demonstrates a proof of the fact that every prime $p$ is a sum of four squares of integers. Since $2 = 1^2 + 1^2 + 0^2 + 0^2$, let us concentrate on $p > 2$.

**(a)** Show that the congruence $x^2 + y^2 + 1 \equiv 0 \pmod{p}$ has a solution for $(x, y)$.

**(b)** Let $(x, y)$ be a solution of the congruence of Part (a). Define

$$\Lambda := \{(a, b, c, d) \in \mathbb{Z}^4 \mid c \equiv xa + yb \pmod{p}, d \equiv ya - xb \pmod{p}\}.$$

Show that $\Lambda$ is a full lattice of $\mathbb{R}^4$ with discriminant $\Delta(\Lambda) = p^2$. (**Hint:** $(1, 0, x, y)$, $(0, 1, y, -x)$, $(0, 0, p, 0)$ and $(0, 0, 0, p)$ constitute a basis of $\Lambda$.)

**(c)** Show that the ball $\mathrm{B}_{4,r}$ contains a non-zero point of $\Lambda$, if $r^2 \geqslant \frac{4\sqrt{2}}{\pi}p = (1.8006326323\ldots)p$. (**Hint:** You may assume that $\mathrm{vol}(\mathrm{B}_{4,r}) = \pi^2 r^4/2$.)

**(d)** If $\frac{4\sqrt{2}}{\pi}p \leqslant r^2 < 2p$, conclude that every non-zero point $(a, b, c, d) \in \Lambda \cap \mathrm{B}_{4,r}$ satisfies $a^2 + b^2 + c^2 + d^2 = p$.

## 5.3 Dirichlet's unit theorem

Now back to the business – the units in $K$. We continue to use the notations introduced at the beginning of Section 5.1. In order to bring lattices in the scene one uses the complex embeddings of $K$.

**5.15 Definition** The map

$$L : K^* \to \mathbb{R}^{r_1+r_2}, \quad \alpha \mapsto \Big(\ln|\sigma_1(\alpha)|, \ldots, \ln|\sigma_{r_1}(\alpha)|, 2\ln|\sigma_{r_1+1}(\alpha)|, \ldots, 2\ln|\sigma_{r_1+r_2}(\alpha)|\Big),$$

is called the logarithmic representation or the logarithmic map of $K$ (or $K^*$). The restriction of $L$ to $\mathfrak{U}$ will be denoted by $\mathfrak{L} := L_{|\mathfrak{U}}$. The space $\mathbb{R}^{r_1+r_2}$ is called the logarithmic space of $K$. Let us denote the coordinates of the logarithmic space by $x_1, \ldots, x_{r_1+r_2}$, where $x_i$ represents $\ln|\sigma_i(\alpha)|$. The hyperplane $\mathfrak{H} : x_1 + \cdots + x_{r_1} + 2x_{r_1+1} + \cdots + 2x_{r_1+r_2} = 0$ of the logarithmic space is of importance in what follows. We have $\dim \mathfrak{H} = r_1 + r_2 - 1$.

First let us look at the immediate properties of $L$ (and $\mathfrak{L}$).

**5.16 Lemma** (1) $L$ is a homomorphism of the multiplicative group $K^*$ to the additive group $\mathbb{R}^{r_1+r_2}$. In particular, $\mathfrak{L}$ is a group homomorphism from $\mathfrak{U}$ to $\mathbb{R}^{r_1+r_2}$.

(2) $\mathrm{Ker}\,\mathfrak{L} = \mathfrak{R}$. In particular, $\mathrm{Ker}\,\mathfrak{L}$ is finite.

(3) $\mathrm{Im}\,\mathfrak{L}$ is contained in the hyperplane $\mathfrak{H}$ and is isomorphic to $\mathfrak{G}$.

*Proof* (1) For each $\alpha, \beta \in K^*$ and $i \in \{1, \ldots, r_1 + r_2\}$ we have $\sigma_i(\alpha\beta) = \sigma_i(\alpha)\sigma_i(\beta)$ ($\sigma_i$ being a field homomorphism), so that $L(\alpha\beta) = L(\alpha) + L(\beta)$.

(2) Since $\mathfrak{R} = \mu_m$ for some $m \in \mathbb{N}$, for each $\alpha \in \mathfrak{R}$ and $i \in \{1, \ldots, r_1 + r_2\}$ we have $\sigma_i(\alpha)^m = \sigma_i(\alpha^m) = \sigma_i(1) = 1$, i.e., $|\sigma_i(\alpha)| = 1$. Thus $\mathfrak{R} \subseteq \mathrm{Ker}\,\mathfrak{L}$. Conversely if $\alpha \in \mathrm{Ker}\,\mathfrak{L}$, we have $|\sigma_i(\alpha)| = 1$ for all $i = 1, \ldots, d$. This implies that $|\sigma_i(\alpha^n)| = 1$ for all $i$ and for all $n \in \mathbb{N}$. By Proposition 5.2 $\alpha$ can not be of infinite order, i.e., $\alpha \in \mathfrak{R}$.

(3) Let $\alpha \in \mathfrak{U}$. By Proposition 5.1 we have $1 = |\mathrm{N}(\alpha)| = \prod_{i=1}^{r_1+2r_2} |\sigma_i(\alpha)|$. Now $\sigma_{r_1+r_2+j}(\alpha) = \overline{\sigma_{r_1+j}(\alpha)}$ for each $j = 1, \ldots, r_2$, so that $1 = \Big(\prod_{i=1}^{r_1} |\sigma_i(\alpha)|\Big)\Big(\prod_{i=r_1+1}^{r_1+r_2} |\sigma_i(\alpha)|^2\Big)$. Taking logarithm we conclude that $\mathfrak{L}(\alpha) \in \mathfrak{H}$, i.e., $\mathrm{Im}\,\mathfrak{L} \subseteq \mathfrak{H}$. Finally $\mathfrak{U} \cong \mathfrak{R} \times \mathfrak{G}$ and hence by the isomorphism theorem we have $\mathrm{Im}\,\mathfrak{L} \cong \mathfrak{U}/\mathrm{Ker}\,\mathfrak{L} = \mathfrak{U}/\mathfrak{R} \cong \mathfrak{G}$. ◀

**5.17 Theorem** $\operatorname{Im}\mathfrak{L}$ is a lattice of $\mathbb{R}^{r_1+r_2}$ of dimension $\leqslant r_1+r_2-1$. In particular, $\mathfrak{G}$ is finitely generated with rank $\leqslant r_1+r_2-1$.

*Proof* I will first show that $\operatorname{Im}\mathfrak{L}$ is a discrete subgroup of $\mathbb{R}^{r_1+r_2}$. Choose any real $r>0$. The ball $\mathrm{B}_{r_1+r_2,r}$ is contained in the cube $\mathrm{C}_{r_1+r_2,r} := \{\mathbf{x}\in\mathbb{R}^{r_1+r_2} \mid |x_i|\leqslant r \text{ for all } i\}$. Let $\alpha\in\mathfrak{U}$ be such that $\mathfrak{L}(\alpha)\in \mathrm{B}_{r_1+r_2,r}$, i.e., $\mathfrak{L}(\alpha)\in \mathrm{C}_{r_1+r_2,r}$, i.e., $|\sigma_i(\alpha)|\leqslant \exp r$ for all $i=1,\dots,d$. By Proposition 5.2 the number of such $\alpha$ is finite, i.e., $(\operatorname{Im}\mathfrak{L})\cap \mathrm{B}_{r_1+r_2,r}$ is a finite set. Hence $\operatorname{Im}\mathfrak{L}$ is a discrete subgroup and thereby a lattice of $\mathbb{R}^{r_1+r_2}$ (Proposition 5.13). Moreover since $\operatorname{Im}\mathfrak{L}\subseteq\mathfrak{H}$, the dimension of $\operatorname{Im}\mathfrak{L}$ is $\leqslant \dim\mathfrak{H}=r_1+r_2-1$. The last assertion in the theorem follows from the fact that $\mathfrak{G}\cong\operatorname{Im}\mathfrak{L}$. ◀

So far so good! It is at least proved that the free part $\mathfrak{G}$ of $\mathfrak{U}$ is of finite rank, i.e., is finitely generated. The rank of $\mathfrak{G}$ is no more than $r_1+r_2-1$. It turns out that the rank of $\mathfrak{G}$ is exactly *equal to* $r_1+r_2-1$. It is by no means an easy task to prove this. If $r_1+r_2-1=0$, then there is nothing to prove. So let us assume that $r_1+r_2-1\geqslant 1$ and proceed by proving a series of auxiliary results. First we have to look at a different kind of lattices. We again consider the complex embeddings of $K$.

**5.18 Definition** The map

$$\sigma : K\to\mathbb{R}^{r_1}\times\mathbb{C}^{r_2},\quad \alpha\mapsto\Big(\sigma_1(\alpha),\dots,\sigma_{r_1}(\alpha),\sigma_{r_1+1}(\alpha),\dots,\sigma_{r_1+r_2}(\alpha)\Big),$$

is called the canonical embedding of $K$ in $\mathbb{R}^{r_1}\times\mathbb{C}^{r_2}$. Here we focus on the additive groups of $K$ and $\mathbb{R}^{r_1}\times\mathbb{C}^{r_2}$. Using the standard $\mathbb{R}$-basis $(1,\mathrm{i})$ of $\mathbb{C}$ one can identify $\mathbb{R}^{r_1}\times\mathbb{C}^{r_2}$ with $\mathbb{R}^{r_1}\times\mathbb{R}^{2r_2}=\mathbb{R}^d$ and so $\sigma$ can be thought of as a group homomorphism from $(K,+)$ to $(\mathbb{R}^d,+)$.

**5.19 Proposition** Let $\mathfrak{a}$ be a non-zero ideal of $\mathfrak{O}$. Then $\sigma(\mathfrak{a})$ is a full lattice of $\mathbb{R}^d$ with discriminant $\Delta(\sigma(\mathfrak{a}))=2^{-r_2}\,\mathrm{N}(\mathfrak{a})\sqrt{|\Delta_K|}$. In particular, $\sigma(\mathfrak{O})$ is a full lattice of $\mathbb{R}^d$ with $\Delta(\sigma(\mathfrak{O}))=2^{-r_2}\sqrt{|\Delta_K|}$.

*Proof* Let $(\gamma_1,\dots,\gamma_d)$ be an integral basis of $\mathfrak{a}$. It suffices to show that $\sigma(\gamma_1),\dots,\sigma(\gamma_d)$ are linearly independent over $\mathbb{R}$, i.e., to show that the matrix

$$A:=\begin{pmatrix}
\sigma_1(\gamma_1) & \cdots & \sigma_{r_1}(\gamma_1) & \operatorname{Re}\sigma_{r_1+1}(\gamma_1) & \operatorname{Im}\sigma_{r_1+1}(\gamma_1) & \cdots & \operatorname{Re}\sigma_{r_1+r_2}(\gamma_1) & \operatorname{Im}\sigma_{r_1+r_2}(\gamma_1)\\
\sigma_1(\gamma_2) & \cdots & \sigma_{r_1}(\gamma_2) & \operatorname{Re}\sigma_{r_1+1}(\gamma_2) & \operatorname{Im}\sigma_{r_1+1}(\gamma_2) & \cdots & \operatorname{Re}\sigma_{r_1+r_2}(\gamma_2) & \operatorname{Im}\sigma_{r_1+r_2}(\gamma_2)\\
\vdots & \cdots & \vdots & \vdots & \vdots & \cdots & \vdots & \vdots\\
\sigma_1(\gamma_d) & \cdots & \sigma_{r_1}(\gamma_d) & \operatorname{Re}\sigma_{r_1+1}(\gamma_d) & \operatorname{Im}\sigma_{r_1+1}(\gamma_d) & \cdots & \operatorname{Re}\sigma_{r_1+r_2}(\gamma_d) & \operatorname{Im}\sigma_{r_1+r_2}(\gamma_d)
\end{pmatrix}$$

has non-zero determinant. Consider the matrix

$$B:=(\,\sigma_j(\gamma_i)\,)=\begin{pmatrix}
\sigma_1(\gamma_1) & \cdots & \sigma_{r_1}(\gamma_1) & \sigma_{r_1+1}(\gamma_1) & \cdots & \sigma_{r_1+r_2}(\gamma_1) & \overline{\sigma_{r_1+1}(\gamma_1)} & \cdots & \overline{\sigma_{r_1+r_2}(\gamma_1)}\\
\sigma_1(\gamma_2) & \cdots & \sigma_{r_1}(\gamma_2) & \sigma_{r_1+1}(\gamma_2) & \cdots & \sigma_{r_1+r_2}(\gamma_2) & \overline{\sigma_{r_1+1}(\gamma_2)} & \cdots & \overline{\sigma_{r_1+r_2}(\gamma_2)}\\
\vdots & \cdots & \vdots & \vdots & \cdots & \vdots & \vdots & \cdots & \vdots\\
\sigma_1(\gamma_d) & \cdots & \sigma_{r_1}(\gamma_d) & \sigma_{r_1+1}(\gamma_d) & \cdots & \sigma_{r_1+r_2}(\gamma_d) & \overline{\sigma_{r_1+1}(\gamma_d)} & \cdots & \overline{\sigma_{r_1+r_2}(\gamma_d)}
\end{pmatrix}.$$

Adding the $(r_1+r_2+1)$-st column to the $(r_1+1)$-st column of $B$ gives twice the $(r_1+1)$-st column of $A$. Subsequently subtracting half the $(r_1+1)$-st column from the $(r_1+r_2+1)$-st column of (the modified) $B$ gives $-\mathrm{i}$ times the $(r_1+2)$-nd column of $A$. Proceeding in this way for all $j=2,\dots,r_2$ shows that $|\det B|=2^{r_2}|\det A|$. But $0\neq\Delta(\mathfrak{a})=\Delta(\gamma_1,\dots,\gamma_d)=(\det B)^2$, which shows that $\det A\neq 0$. Moreover $\Delta(\sigma(\mathfrak{a}))=|\det A|=2^{-r_2}|\det B|=2^{-r_2}\sqrt{|\Delta(\mathfrak{a})|}=2^{-r_2}\,\mathrm{N}(\mathfrak{a})\sqrt{|\Delta_K|}$, where the last equality follows from Corollary 4.22. ◀

The canonical embedding of $K$ in $\mathbb{R}^{r_1}\times\mathbb{C}^{r_2}$ leads to a generalization of the concept of norms.

**5.20 Definition** Let $\mathbf{x} := (x_1, \dots, x_{r_1}, z_1, \dots, z_{r_2}) \in \mathbb{R}^{r_1} \times \mathbb{C}^{r_2}$, where $x_i$ are real numbers and $z_j$ are complex numbers. ($\mathbf{x}$ can be viewed as the $d$-tuple $(x_1, \dots, x_{r_1}, \operatorname{Re} z_1, \operatorname{Im} z_1, \dots, \operatorname{Re} z_{r_2}, \operatorname{Im} z_{r_2})$ under the identification of $\mathbb{R}^{r_1} \times \mathbb{C}^{r_2}$ with $\mathbb{R}^d$.) We define the norm of $\mathbf{x}$ as

$$\mathrm{N}(\mathbf{x}) := x_1 \cdots x_{r_1} z_1 \bar{z}_1 \cdots z_{r_2} \bar{z}_{r_2}.$$

Clearly $\mathrm{N}(\alpha) = \mathrm{N}(\sigma(\alpha))$ for $\alpha \in K$, where the norm of the left hand side corresponds to Definition 3.8.

Since $\mathbb{R}^{r_1} \times \mathbb{C}^{r_2}$ is also a ring (It is in fact an $\mathbb{R}$-algebra.), one can define the product of two elements $\mathbf{x} := (x_1, \dots, x_{r_1}, z_1, \dots, z_{r_2})$ and $\mathbf{x}' := (x'_1, \dots, x'_{r_1}, z'_1, \dots, z'_{r_2})$ of $\mathbb{R}^{r_1} \times \mathbb{C}^{r_2}$ as the coordinate-wise product $\mathbf{x}\mathbf{x}' := (x_1x'_1, \dots, x_{r_1}x'_{r_1}, z_1z'_1, \dots, z_{r_2}z'_{r_2})$, which viewed as an element of $\mathbb{R}^d$ is to be identified with $(x_1x'_1, \dots, x_{r_1}x'_{r_1}, \operatorname{Re}(z_1z'_1), \operatorname{Im}(z_1z'_1), \dots, \operatorname{Re}(z_{r_2}z'_{r_2}), \operatorname{Im}(z_{r_2}z'_{r_2}))$.[7] Clearly this new norm function is multiplicative, i.e., $\mathrm{N}(\mathbf{x}\mathbf{x}') = \mathrm{N}(\mathbf{x})\,\mathrm{N}(\mathbf{x}')$ for every $\mathbf{x}, \mathbf{x}' \in \mathbb{R}^{r_1} \times \mathbb{C}^{r_2}$. Let us also define the set

$$\mathbf{x}\sigma(\mathfrak{O}) := \{\mathbf{x}\sigma(\alpha) \mid \alpha \in \mathfrak{O}\} \subseteq \mathbb{R}^{r_1} \times \mathbb{C}^{r_2}.$$

**5.21 Proposition** For $\mathbf{x} \in \mathbb{R}^{r_1} \times \mathbb{C}^{r_2}$ with $|\mathrm{N}(\mathbf{x})| \neq 0$ the set $\mathbf{x}\sigma(\mathfrak{O})$ is a full lattice of $\mathbb{R}^d$ with discriminant $\Delta(\mathbf{x}\sigma(\mathfrak{O})) = 2^{-r_2}|\mathrm{N}(\mathbf{x})|\sqrt{|\Delta_K|}$.

*Proof* The proof is similar to that of Proposition 5.19. If $(\gamma_1, \dots, \gamma_d)$ is an integral basis of $K$, then it suffices to show that $\mathbf{x}\sigma(\gamma_1), \dots, \mathbf{x}\sigma(\gamma_d)$ are $\mathbb{R}$-linearly independent. The details are left to the reader. ◀

Now we have the requisite machineries to prove the unit theorem. In what follows let us denote by $\Lambda_{\mathbf{x}}$ the lattice $\mathbf{x}\sigma(\mathfrak{O})$ of $\mathbb{R}^d$.

**5.22 Theorem** [Dirichlet's unit theorem] $\operatorname{Im}\mathfrak{L}$ is a lattice of $\mathbb{R}^{r_1+r_2}$ of dimension $r_1 + r_2 - 1$, i.e., $\mathfrak{G}$ is a free Abelian group of rank $r_1 + r_2 - 1$.

*Proof* Let $X := \left\{\mathbf{x} \in \mathbb{R}^{r_1} \times \mathbb{C}^{r_2} \mid \tfrac{1}{2} \leqslant |\mathrm{N}(\mathbf{x})| \leqslant 1\right\}$. We have $2^{-r_2-1}\sqrt{|\Delta_K|} \leqslant \Delta(\Lambda_{\mathbf{x}}) \leqslant 2^{-r_2}\sqrt{|\Delta_K|}$ for every $\mathbf{x} \in X$. Choose any convex closed bounded symmetric region $\Omega$ of $\mathbb{R}^{r_1} \times \mathbb{C}^{r_2}$ having $\operatorname{vol}(\Omega) \geqslant 2^{r_1+r_2}\sqrt{|\Delta_K|}$. By Minkowski's convex-body theorem (5.11) for each $\mathbf{x} \in X$ there exists a non-zero $\alpha \in \mathfrak{O}$ such that $\mathbf{x}\sigma(\alpha) \in \Omega$. Since points in $\Omega$ have bounded coordinates, we have a positive real constant $B$ depending on $\Omega$ such that $|\mathrm{N}(\mathbf{y})| \leqslant B/2$ for every $\mathbf{y} \in \Omega$. Then for each $\mathbf{x} \in X$ and $\alpha \in \mathfrak{O}$ with $\mathbf{x}\sigma(\alpha) \in \Omega$ we have $|\mathrm{N}(\mathbf{x}\sigma(\alpha))| = |\mathrm{N}(\mathbf{x})||\mathrm{N}(\sigma(\alpha))| = |\mathrm{N}(\mathbf{x})||\mathrm{N}(\alpha)| \leqslant B/2$, so that $|\mathrm{N}(\alpha)| \leqslant B$.

Now consider the set $S$ consisting of the non-zero principal ideals $\alpha\mathfrak{O}$ for which $\mathbf{x}\sigma(\alpha) \in \Omega$ for some $\mathbf{x} \in X$. For each $\alpha\mathfrak{O} \in S$ we have $\mathrm{N}(\alpha\mathfrak{O}) = |\mathrm{N}(\alpha)| \leqslant B$. By Proposition 4.29 $S$ is a finite set, say $S = \{\alpha_1\mathfrak{O}, \dots, \alpha_k\mathfrak{O}\}$. Then for every $\mathbf{x} \in X$ and $0 \neq \alpha \in \mathfrak{O}$ with $\mathbf{x}\sigma(\alpha) \in \Omega$ we have $\alpha\mathfrak{O} = \alpha_i\mathfrak{O}$ for some $i \in \{1, \dots, k\}$, i.e., $\alpha = \xi\alpha_i$ for some unit $\xi \in \mathfrak{U}$, i.e., $\mathbf{x}\sigma(\xi) = \mathbf{x}\sigma(\alpha\alpha_i^{-1}) = \sigma(\alpha_i^{-1})\mathbf{x}\sigma(\alpha) \in \sigma(\alpha_i^{-1})\Omega$. The region $\Omega' := \bigcup_{i=1}^{k} \sigma(\alpha_i^{-1})\Omega \subseteq \mathbb{R}^{r_1} \times \mathbb{C}^{r_2}$ is bounded (since $\Omega$ is so), i.e., for every $\mathbf{x} \in X$ we have a unit $\xi$ of $\mathfrak{O}$ such that the absolute value of each coordinate of $\mathbf{x}\sigma(\xi)$ is $< M$ for some finite bound $M > 0$ that depends on $\Omega'$ (and not on $\mathbf{x}$).

For each $i \in \{1, \dots, r_1 + r_2\}$ choose a point $\mathbf{x}_i = (x_{i1}, \dots, x_{i,r_1+r_2}) \in \mathbb{R}^{r_1} \times \mathbb{C}^{r_2}$ with $|x_{ij}| = M$ for $j \neq i$ and with $|x_{ii}|$ so adjusted that $|\mathrm{N}(\mathbf{x}_i)| = 1$. Then each $\mathbf{x}_i \in X$ and hence we have a unit $\xi_i \in \mathfrak{U}$ such that $\mathbf{x}_i\sigma(\xi_i) = \Big(x_{i1}\sigma_1(\xi_i), \dots, x_{ir_1}\sigma_{r_1}(\xi_i), x_{i,r_1+1}\sigma_{r_1+1}(\xi_i), \dots, x_{i,r_1+r_2}\sigma_{r_1+r_2}(\xi_i)\Big) \in \Omega'$. The absolute values of the coordinates of $\mathbf{x}_i\sigma(\xi_i)$ are $< M$ and so $|\sigma_j(\xi_i)| < 1$, i.e., $\ln|\sigma_j(\xi_i)| < 0$ for $j \neq i$.

Finally let us come back to our old lattice $\operatorname{Im}\mathfrak{L} = \mathfrak{L}(\mathfrak{U}) = \mathfrak{L}(\mathfrak{G})$ of the logarithmic space $\mathbb{R}^{r_1+r_2}$ of $K$. We have seen (Theorem 5.17) that $\operatorname{Im}\mathfrak{L}$ has dimension $\leqslant r_1 + r_2 - 1$. I now claim that the vectors

[7] $\mathbf{x}\mathbf{x}'$ is *not* to be identified with $(x_1x'_1, \dots, x_{r_1}x'_{r_1}, \operatorname{Re}(z_1)\operatorname{Re}(z'_1), \operatorname{Im}(z_1)\operatorname{Im}(z'_1), \dots, \operatorname{Re}(z_{r_2})\operatorname{Re}(z'_{r_2}), \operatorname{Im}(z_{r_2})\operatorname{Im}(z'_{r_2}))$. We identify $\mathbb{R}^{r_1} \times \mathbb{C}^{r_2}$ with $\mathbb{R}^d$ as an $\mathbb{R}$-vector space, not as a ring.

$\mathfrak{L}(\xi_1), \ldots, \mathfrak{L}(\xi_{r_1+r_2-1}) \in \operatorname{Im}\mathfrak{L}$ are $\mathbb{R}$-linearly independent, so that $\operatorname{Im}\mathfrak{L}$ has dimension $\geqslant r_1 + r_2 - 1$, thereby proving Dirichlet's unit theorem.

For the proof of the claim just mentioned let us project the $(r_1 + r_2)$-dimensional vectors $\mathfrak{L}(\xi_i)$ to $\mathbb{R}^{r_1+r_2-1}$ by dropping the last coordinate[8] and show that the matrix $A := (\, l_{ij}\,)_{1\leqslant i,j\leqslant r_1+r_2-1}$ is non-singular, where $l_{ij} := \begin{cases} \ln|\sigma_j(\xi_i)| & \text{for } 1 \leqslant j \leqslant r_1, \\ 2\ln|\sigma_j(\xi_i)| & \text{for } r_1+1 \leqslant j \leqslant r_1+r_2-1. \end{cases}$ We have seen that $l_{ij} < 0$ for $i \neq j$. Furthermore since $\mathrm{N}(\xi_i) = \pm 1$, we have $\sum_{j=1}^{r_1+r_2-1} l_{ij} = -\epsilon \ln|\sigma_{r_1+r_2}(\xi_i)| > 0$ for $i < r_1 + r_2$, where $\epsilon$ is 1 or 2 depending on whether $r_2 = 0$ or $r_2 > 0$ respectively. It is left to the reader as an easy exercise to show that the above two conditions on the coefficients $l_{ij}$ of $A$ are sufficient to prove the non-singularity of $A$ (Exercise 5.3.2). ◀

That's it! And it's already too much for a theorem with such a simple statement! The above proof for the unit theorem is not constructive, i.e., it does not specify how to compute a set of generators of $\mathfrak{G}$ (often called a set of fundamental units of $\mathfrak{O}$). For real quadratic number fields one can use the theory of continued fractions[9] to compute the fundamental unit.

**5.23 Example** Let $D \neq 0, 1$ be a square-free integer, $K := \mathbb{Q}(\sqrt{D})$ and $\mathfrak{O} := \mathfrak{O}_K$. If $D < 0$, then the signature of $K$ is $(0, 1)$ and the rank of $\mathfrak{G}$ is $0 + 1 - 1 = 0$, i.e., $\mathfrak{U} = \mathfrak{R} = \mu_m$ for some even $m$ with $\phi(m) = 1$ or 2. That is, $\mathfrak{U}$ is finite in this case. (See Exercises 5.1.2 and 5.3.3.)

For $D > 0$ the signature of $K$ is $(2, 0)$, i.e., the rank of $\mathfrak{G}$ is $2 + 0 - 1 = 1$. This means that $\mathfrak{O}$ contains an infinite number of units. Let $\xi$ be a fundamental unit of $\mathfrak{O}$. Then every unit of $\mathfrak{O}$ is of the form $\pm\xi^n$, $n \in \mathbb{Z}$ (Exercise 5.1.3). The question is how to compute $\xi$ for such an $\mathfrak{O}$.

First consider the easier case $D \equiv 2, 3 \pmod 4$. In this case every element of $\mathfrak{O}$ is of the form $a + b\sqrt{D}$ with $a, b \in \mathbb{Z}$. For $a + b\sqrt{D} \in \mathfrak{U}$ the condition $\mathrm{N}(a + b\sqrt{D}) = \pm 1$ implies that $a^2 - Db^2 = \pm 1$.[10] Since a solution $(a, b)$ of these equations gives four solutions $(\pm a, \pm b)$, we need to concentrate only on the solutions with $a \geqslant 0$ and $b \geqslant 0$. Furthermore the only solutions with $ab = 0$ are $(\pm 1, 0)$, so that we may consider $a > 0$ and $b > 0$ only.

The solutions of the equations $a^2 - Db^2 = \pm 1$ are related to the the (simple) continued fraction expansion of $\sqrt{D}$, which is periodic and is of the form $\langle a_0, \overline{a_1, \ldots, a_{r-1}, 2a_0}\rangle$.[11] The (least) period of this expansion is $r \geqslant 1$. Let $h_n/k_n, n \in \mathbb{Z}_+$, be the convergents of the expansion of $\sqrt{D}$. Then each positive solution $(a, b)$ of $a^2 - Db^2 = \pm 1$ is from the set $\{(h_n, k_n) \mid n \in \mathbb{Z}_+\}$. If $r$ is even, there are no solutions of $a^2 - Db^2 = -1$, whereas the positive solutions of $a^2 - Db^2 = 1$ are given by $(h_{nr-1}, k_{nr-1})$ for all $n \in \mathbb{N}$. If $r$ is odd, then the positive solutions of $a^2 - Db^2 = -1$ are $(h_{nr-1}, k_{nr-1})$ for all odd $n \in \mathbb{N}$, and the positive solutions of $a^2 - Db^2 = 1$ are $(h_{nr-1}, k_{nr-1})$ for all even $n \in \mathbb{N}$. In either case we can name the positive solutions of $a^2 - Db^2 = \pm 1$ as $(a_n, b_n) := (h_{nr-1}, k_{nr-1})$, $n \in \mathbb{N}$. Then $a_n + b_n\sqrt{D} = (a_1 + b_1\sqrt{D})^n$, that is, $\xi := a_1 + b_1\sqrt{D} = h_{r-1} + k_{r-1}\sqrt{D}$ is a fundamental unit of $\mathfrak{O}$.

The following table gives the (positive) fundamental unit $\xi > 1$ of $\mathfrak{O}$ for some small positive square-free values of $D \equiv 2, 3 \pmod 4$.

[8] The last unit $\xi_{r_1+r_2}$ and the last coordinate were dropped on aesthetic grounds. Dropping the $i$-th vector and (hence) the $i$-th coordinate would have served our purpose equally brilliantly for any $i \in \{1, \ldots, r_1 + r_2\}$.

[9] See any book on elementary number theory, say, the book by Niven, Zuckerman and Montgomery.

[10] The Diophantine equation $x^2 - Dy^2 = 1$ is known as Pell's equation after John Pell. However John Pell contributed very little to solving or even noticing this equation. The equation bears his name owing to a mistake of Euler who supposedly wanted to give the credit to Brouncker, the founder of continued fractions. What would better be called 'Brouncker's equation' has been since then (and till now) erroneously dubbed as Pell's equation.

[11] This property of the continued fraction expansion of $\sqrt{D}$ is true even when $D \equiv 1 \pmod 4$. What is demanded is that $D$ should be square-free.

| $D$ | $\sqrt{D}$ | $r$ | $h_{r-1}/k_{r-1}$ | $\xi$ |
|---|---|---|---|---|
| 2 | $\langle 1, \overline{2}\rangle$ | 1 | $1/1$ | $1+\sqrt{2}$ |
| 3 | $\langle 1, \overline{1,2}\rangle$ | 2 | $2/1$ | $2+\sqrt{3}$ |
| 6 | $\langle 2, \overline{2,4}\rangle$ | 2 | $5/2$ | $5+2\sqrt{6}$ |
| 7 | $\langle 2, \overline{1,1,1,4}\rangle$ | 4 | $8/3$ | $8+3\sqrt{7}$ |
| 10 | $\langle 3, \overline{6}\rangle$ | 1 | $3/1$ | $3+\sqrt{10}$ |
| 11 | $\langle 3, \overline{3,6}\rangle$ | 2 | $10/3$ | $10+3\sqrt{11}$ |
| 14 | $\langle 3, \overline{1,2,1,6}\rangle$ | 4 | $15/4$ | $15+4\sqrt{14}$ |
| 15 | $\langle 3, \overline{1,6}\rangle$ | 2 | $4/1$ | $4+\sqrt{15}$ |
| 19 | $\langle 4, \overline{2,1,3,1,2,8}\rangle$ | 6 | $170/39$ | $170+39\sqrt{19}$ |

Note that the fundamental unit $\xi$ may be quite large even for small values of $D$. For example, the continued fraction expansion of $\sqrt{991}$ is

$$\begin{aligned}\sqrt{991} &= \langle 31, \overline{2,12,10,2,2,2,1,1,2,6,1,1,1,1,3,1,8,4,1,2,1,2,3,1,4,1,20,6,4,31,}\\ &\qquad \overline{4,6,20,1,4,1,3,2,1,2,1,4,8,1,3,1,1,1,1,6,2,1,1,2,2,2,10,12,2,62}\rangle,\end{aligned}$$

which corresponds to a period of $r = 60$, so that

$$\begin{aligned}\xi &= h_{59} + k_{59}\sqrt{991}\\ &= 379516400906811930638014896080 + 12055735790331359447442538767\sqrt{991}.\end{aligned}$$

Now let us consider a positive square-free integer $D \equiv 1 \pmod 4$. In this case the elements of $\mathfrak{O}$ are of the form $a + b\left(\frac{1+\sqrt{D}}{2}\right)$ with $a, b \in \mathbb{Z}$. Putting $x := 2a + b$ and $y := b$ allows us to rewrite this element as $(x + y\sqrt{D})/2$, where $x$ and $y$ are integers of the same parity. Applying the norm condition on a unit of this form implies that $x^2 - Dy^2 = \pm 4$. Since $D$ is odd, any integer solution of these equations corresponds to $x$ and $y$ of the same parity and therefore to a unit of $\mathfrak{O}$. As in the previous case we need to concentrate only on the positive solutions of $x^2 - Dy^2 = \pm 4$.

In this case also the continued fraction expansion of $\sqrt{D}$ helps. Let $\sqrt{D} = \langle a_0, \overline{a_1, \ldots, a_{r-1}, 2a_0}\rangle$ with least period $r$. We compute the least positive solution $> 1$ of $a^2 - Db^2 = \pm 1$ as $\eta := h_{r-1} + k_{r-1}\sqrt{D}$. Clearly $2\eta$ (i.e., $(2h_{r-1}, 2k_{r-1})$) satisfies $x^2 - Dy^2 = \pm 4$, i.e., $\eta$ is a unit in $\mathfrak{O}$. But the fundamental unit $\xi := (u + v\sqrt{D})/2 > 1$ of $\mathfrak{O}$ may be smaller than $\xi$. One has:

$$\eta = \begin{cases} \xi & \text{if } D \equiv 1 \pmod 8,\\ \xi & \text{if } D \equiv 5 \pmod 8 \text{ and } u \equiv v \equiv 0 \pmod 2,\\ \xi^3 & \text{if } D \equiv 5 \pmod 8 \text{ and } u \equiv v \equiv 1 \pmod 2.\end{cases}$$

Since $u$ and $v$ are not known in advance, one should try to compute the cube root of $\eta$ in $\mathfrak{O}$ (for the case $D \equiv 5 \pmod 8$). If the attempt is successful, one gets $\xi = \eta^{1/3}$ (the real cube root), otherwise one takes $\xi = \eta$. The following table supplements the previous one by listing the fundamental unit $\xi > 1$ of $\mathfrak{O}$ for some small square-free $D \equiv 1 \pmod 4$.

| $D$ | $\sqrt{D}$ | $r$ | $h_{r-1}/k_{r-1}$ | $\eta$ | $\xi$ |
|---|---|---|---|---|---|
| 5 | $\langle 2, \overline{4}\rangle$ | 1 | $2/1$ | $2+\sqrt{5}$ | $\eta^{1/3} = (1+\sqrt{5})/2 = 0 + 1\left(\frac{1+\sqrt{5}}{2}\right)$ |
| 13 | $\langle 3, \overline{1,1,1,1,6}\rangle$ | 5 | $18/5$ | $18+5\sqrt{13}$ | $\eta^{1/3} = (3+\sqrt{13})/2 = 1 + \left(\frac{1+\sqrt{13}}{2}\right)$ |
| 17 | $\langle 4, \overline{8}\rangle$ | 1 | $4/1$ | $4+\sqrt{17}$ | $\eta = 4 + \sqrt{17} = 3 + 2\left(\frac{1+\sqrt{17}}{2}\right)$ |

For a general number field computing a set of fundamental units can be carried out by discovering pairs $(\alpha, \beta) \in \mathfrak{O}^2$ with $\mathrm{N}(\alpha) = \mathrm{N}(\beta) = m$ for some small integer $m$. Since there are only finitely many

(principal) ideals of $\mathfrak{O}$ of norm $m$, there is a finite (i.e., non-zero) chance that $\alpha\mathfrak{O} = \beta\mathfrak{O}$. If so, $\alpha/\beta$ is a unit of $\mathfrak{O}$. When many units are available in this way, we have to find out a set of $r_1 + r_2 - 1$ linearly independent units. Suitable roots of these linearly independent units will give us the desired set of fundamental units of $\mathfrak{O}$.

**Exercises for Section 5.3**

**1.** Complete the details in the proof of Proposition 5.21.

**2.** Let $A := (l_{ij})$ be an $n \times n$ matrix with real entries having the properties that $l_{ij} < 0$ for $i \neq j$ and that $\sum_{j=1}^{n} l_{ij} > 0$ for all $i \in \{1, \ldots, n\}$. Show that $A$ is non-singular (i.e., invertible). (**Hint:** If $A$ is singular, the system $A\mathbf{x} = \mathbf{0}$ has a non-zero solution $\mathbf{x}$. Look at the $i$-th equation, where $|x_i| = \max\limits_{1 \leqslant j \leqslant n} |x_j|$.)

**3.** Compute $\mathfrak{U}_K$ for $K := \mathbb{Q}(\sqrt{D})$, where $D < 0$ is a square-free integer.

**4.** Let $\alpha$ be a root of the polynomial $f(X) := X^3 + 10X + 1$ (Show that $f(X)$ is irreducible in $\mathbb{Q}[X]$.), $K := \mathbb{Q}(\alpha)$ and $\mathfrak{O} := \mathfrak{O}_K$.

**(a)** Compute that $\Delta(\alpha) = -4027$ and hence deduce that $\mathfrak{O} = \mathbb{Z}[\alpha]$.

**(b)** Show that the signature of $K$ is $(1, 1)$.

**(c)** Show that $\alpha$ is a unit of $K$. (**Remark:** $\alpha$ is indeed a fundamental unit of $\mathfrak{O}$, but proving that requires some boring calculations.)

**(d)** Conclude that $\mathfrak{U}_K = \{\pm\alpha^n \mid n \in \mathbb{Z}\}$.

**5.** **(a)** Let $\rho \in \mathbb{N}$. Find all possible values of $d$ for which there may exist a number field $K$ of degree $d$ with $\mathfrak{G}_K$ having rank $\rho$.

**(b)** For which values of $d$ found in Part (a) will a complex embedding of $K$ have a real fundamental root $> 1$?

**6.** Find the fundamental unit $> 1$ of $\mathfrak{O}_K$ for $K := \mathbb{Q}(\sqrt{21})$, $K := \mathbb{Q}(\sqrt{22})$, $K := \mathbb{Q}(\sqrt{23})$, $K := \mathbb{Q}(\sqrt{33})$ and for $K := \mathbb{Q}(\sqrt{37})$.

**7.** Let $n \in \mathbb{N}$.

**(a)** Show that $\sqrt{n^2+1} = \langle n, \overline{2n} \rangle$. Conclude that if $n \not\equiv 2, 6 \pmod 8$ and $n^2 + 1$ is square-free, then a fundamental unit of $\mathfrak{O}_{\mathbb{Q}(\sqrt{n^2+1})}$ is $n + \sqrt{n^2+1}$.

**(b)** Show that $\sqrt{n^2-1} = \langle n-1, \overline{1, 2(n-1)} \rangle$. Conclude that if $n \geqslant 2$ and $n^2 - 1$ is square-free, then a fundamental unit of $\mathfrak{O}_{\mathbb{Q}(\sqrt{n^2-1})}$ is $n + \sqrt{n^2-1}$.

**(c)** Show that $\sqrt{n^2+2} = \langle n, \overline{n, 2n} \rangle$. Conclude that if $n^2 + 2$ is square-free, then a fundamental unit of $\mathfrak{O}_{\mathbb{Q}(\sqrt{n^2+2})}$ is $(n^2+1) + n\sqrt{n^2+2}$.

**8.** [The Minkowski bound] Let $K$ be a number field of degree $d$ and signature $(r_1, r_2)$ and let $\mathfrak{O} := \mathfrak{O}_K$. Show that every non-zero ideal $\mathfrak{a}$ of $\mathfrak{O}$ contains a non-zero element $\alpha$ with $|\mathrm{N}(\alpha)| \leqslant M_K \, \mathrm{N}(\mathfrak{a})$, where

$$M_K := \frac{d!}{d^d} \left(\frac{4}{\pi}\right)^{r_2} \sqrt{|\Delta_K|}$$

is a constant that depends only on $K$. (**Hint:** Apply Minkowski's convex-body theorem on the lattice $\sigma(\mathfrak{a})$ of $\mathbb{R}^{r_1} \times \mathbb{C}^{r_2}$ and on the compact region $\Omega_t := \left\{(x_1, \ldots, x_{r_1}, z_1, \ldots, z_{r_2}) \in \mathbb{R}^{r_1} \times \mathbb{C}^{r_2} \;\middle|\; \sum_{i=1}^{r_1} |x_i| + 2\sum_{j=1}^{r_2} |z_j| \leqslant t\right\}$ for a suitable $t$. You may use the fact that $\mathrm{vol}(\Omega_t) = 2^{r_1}(\pi/2)^{r_2} t^d/d!$.)

**9.** Let $K$ be a number field of degree $d$. Show that:

**(a)** $|\Delta_K| \geqslant \left(\frac{\pi}{4}\right)^d \left(\frac{d^d}{d!}\right)^2$.

**(b)** $|\Delta_K| > 1$ for $K \neq \mathbb{Q}$.

**(c)** $|\Delta_K| \to \infty$ as $d \to \infty$. (**Hint:** $\ln n! = n \ln n - n + \mathrm{O}(\ln n)$.)

*** 10.** [Hermite's theorem] Let $\Delta \in \mathbb{N}$. Show that there are only finitely many number fields $K$ with $|\Delta_K| = \Delta$. (**Hint:** By Minkowski's theorem on the lattice $\sigma(\mathfrak{O})$ there exists a non-zero $\alpha \in \mathfrak{O}_K$ with $|\sigma_1(\alpha)| < \sqrt{\Delta + 1}$ and $|\sigma_i(\alpha)| < 1$ for $i = 2, \ldots, r_1 + r_2$. Argue that $K = \mathbb{Q}(\alpha)$. Also use Exercise 5.3.9(c).)