14 The Minkowski bound and finiteness results

14.1 Lattices in real vector spaces

Recall that for an integral domain \( A \) with fraction field \( K \), an \( A \)-lattice in a finite dimensional \( K \)-vector space \( V \) is a finitely generated \( A \)-submodule of \( V \) that contains a \( K \)-basis for \( V \) (see Definition 5.9). We now want to specialize to the case \( A = \mathbb{Z} \), but rather than working with the fraction field \( K = \mathbb{Q} \) we will instead work with its archimedean completion \( \mathbb{R} \). Note that a finitely generated \( \mathbb{Z} \)-submodule of a vector space is necessarily a free module, since \( \mathbb{Z} \) is a PID and a submodule of a vector space must be torsion-free. Now \( V \) is an \( \mathbb{R} \)-vector space of some finite dimension \( n \), and has a canonical structure as a topological metric space isomorphic to \( \mathbb{R}^n \) (by Proposition 10.5, there is a unique topology on \( V \) compatible with the topology of \( \mathbb{R} \), because \( \mathbb{R} \) is complete). This topology makes \( V \) a locally compact Hausdorff space, thus \( V \) is a locally compact group and is thus equipped with a Haar measure \( \mu \) that is unique up to scaling, by Theorem 13.14.

**Definition 14.1.** A subgroup \( H \) of a topological group \( G \) is discrete if the subspace topology on \( H \) is the discrete topology (every point is open), and cocompact if \( H \) is a normal subgroup of \( G \) and the quotient \( G/H \) is compact (here \( G/H \) denotes the group \( G/H \) with the quotient topology given by identifying elements of \( G \) that lie in the same coset of \( H \)).

**Definition 14.2.** Let \( V \) be an \( \mathbb{R} \)-vector space of finite dimension. A (full) lattice in \( V \) is a \( \mathbb{Z} \)-submodule generated by an \( \mathbb{R} \)-basis for \( V \); equivalently, a discrete cocompact subgroup.

See Problem Set 7 for a proof that these two definitions are equivalent.

**Remark 14.3.** A discrete subgroup of a Hausdorff topological group is always closed; see \([1, \text{III.2.1.5}]\) for a proof. This implies that the quotient of a Hausdorff topological group by a normal discrete subgroup is Hausdorff (which is false for topological spaces in general); see \([1, \text{III.2.1.18}]\). It follows that the quotient of a Hausdorff topological group (including all locally compact groups) by a discrete cocompact subgroup is a compact group. These facts are easy to see in the case of lattices: \( \mathbb{Z} \) is closed in \( \mathbb{R} \) (as the complement of a union of open intervals), so \( \mathbb{Z}^n \) is closed in \( \mathbb{R}^n \). Given a lattice \( \Lambda \) in \( V \), each \( \mathbb{Z} \)-basis for \( \Lambda \) determines an isomorphism of topological groups \( \Lambda \cong \mathbb{Z}^n \) and \( V \cong \mathbb{R}^n \), and the quotient \( V/\Lambda \cong \mathbb{R}^n/\mathbb{Z}^n \cong (\mathbb{R}/\mathbb{Z})^n \) (an \( n \)-torus), is compact Hausdorff and thus a compact group.

**Remark 14.4.** You might ask why we are using the archimedean completion \( \mathbb{R} = \mathbb{Q}_\infty \) rather than some other completion \( \mathbb{Q}_p \). The reason is that \( \mathbb{Z} \) is not a discrete subgroup of \( \mathbb{Q}_p \) (elements of \( \mathbb{Z} \) can be arbitrarily close to 0 under the \( p \)-adic metric).

Any basis \( v_1, \ldots, v_n \) for \( V \) determines a parallelepiped

\[
F(v_1, \ldots, v_n) := \{ t_1 v_1 + \cdots + t_n v_n : t_1, \ldots, t_n \in [0,1) \}
\]

that we may view as the unit cube by fixing an isomorphism \( \varphi : V \cong \mathbb{R}^n \) that maps \( (v_1, \ldots, v_n) \) to the standard basis of unit vectors for \( \mathbb{R}^n \). It then makes sense to normalize the Haar measure \( \mu \) so that \( \mu(F(v_1, \ldots, v_n)) = 1 \), and we then have \( \mu(S) = \mu_{\mathbb{R}^n}(\varphi(S)) \) for every measurable set \( S \subseteq V \), where \( \mu_{\mathbb{R}^n} \) denotes the standard Lebesgue measure on \( \mathbb{R}^n \).

For any other basis \( e_1, \ldots, e_n \) of \( V \), if we let \( E = [e_{ij}]_{ij} \) be the matrix whose \( j \)th column expresses \( e_j = \sum_i e_{ij} v_i \), in terms of our normalized basis \( v_1, \ldots, v_n \), then

\[
\mu(F(e_1, \ldots, e_n)) = | \det E | = \sqrt{\det E^t \det E} = \sqrt{\det(E^t E)} = \sqrt{\det([e_i, e_j]_{ij}),}
\]

Andrew V. Sutherland
where \( \langle e_i, e_j \rangle \) is the canonical inner product (the dot product) on \( \mathbb{R}^n \). Here we have used the fact that the determinant of a matrix in \( \mathbb{R}^{n \times n} \) is the signed volume of the parallelepiped spanned by its columns (or rows). This is a consequence of the following more general result, which is independent of the choice of basis or the normalization of \( \mu \).

**Proposition 14.5.** Let \( T: V \to V \) be a linear transformation of \( V \simeq \mathbb{R}^n \). For any Haar measure \( \mu \) on \( V \) and every measurable set \( S \subseteq V \) we have

\[
\mu(T(S)) = |\det T| \mu(S).
\]

**(2)**

*Proof.* See [8, Ex. 1.2.21].

If \( \Lambda \) is a lattice \( e_1 \mathbb{Z} + \cdots + e_n \mathbb{Z} \) in \( V \), the quotient \( V/\Lambda \) is a compact group that we may identify with the parallelepiped \( F(e_1, \ldots, e_n) \subseteq V \), which forms a set of unique coset representatives. More generally, we make the following definition.

**Definition 14.6.** Let \( \Lambda \) be a lattice in \( V \simeq \mathbb{R}^n \). A *fundamental domain* for \( \Lambda \) is a measurable set \( F \subseteq V \) such that

\[
V = \bigcup_{\lambda \in \Lambda} (F + \lambda).
\]

In other words, \( F \) is a measurable set of coset representatives for \( V/\Lambda \). Fundamental domains exist: if \( \Lambda = e_1 \mathbb{Z} + \cdots + e_n \mathbb{Z} \) we may take the parallelepiped \( F(e_1, \ldots, e_n) \).

**Proposition 14.7.** Let \( \Lambda \) be a lattice in \( V \simeq \mathbb{R}^n \) and let \( \mu \) be a Haar measure on \( V \). Every fundamental domain for \( \Lambda \) has the same measure, and this measure is finite and nonzero.

*Proof.* Let \( F \) and \( G \) be two fundamental domains for \( \Lambda \). Using the translation invariance and countable additivity of \( \mu \) (note that \( \Lambda \simeq \mathbb{Z}^n \) is a countable set) along with the fact that \( \Lambda \) is closed under negation, we obtain

\[
\mu(F) = \mu(F \cap V) = \mu\left(F \cap \bigcup_{\lambda \in \Lambda} (G + \lambda)\right) = \mu\left(\bigcup_{\lambda \in \Lambda} (F \cap (G + \lambda))\right)
\]

\[
= \sum_{\lambda \in \Lambda} \mu(F \cap (G + \lambda)) = \sum_{\lambda \in \Lambda} \mu((F \cap G) + \lambda) = \sum_{\lambda \in \Lambda} \mu(G \cap (F + \lambda)) = \mu(G),
\]

where the last equality follows from the first four (swap \( F \) and \( G \)). If we fix a \( \mathbb{Z} \)-basis \( e_1, \ldots, e_n \) for \( \Lambda \), the parallelepiped \( F(e_1, \ldots, e_n) \) is a fundamental domain for \( \Lambda \), and its closure is compact, so \( \mu(F(e_1, \ldots, e_n)) \) is finite, and it is nonzero because there is an isomorphism \( V \simeq \mathbb{R}^n \) that maps the closure of \( F(e_1, \ldots, e_n) \) to the unit cube in \( \mathbb{R}^n \) whose Lebesgue measure is nonzero (whether a set has zero measure or not does not depend on the normalization of the Haar measure and is therefore preserved by isomorphisms of locally compact groups).

**Definition 14.8.** Let \( \Lambda \) be a lattice in \( V \simeq \mathbb{R}^n \) and fix a Haar measure \( \mu \) on \( V \). The *covolume* \( \text{covol}(\Lambda) \in \mathbb{R}_{>0} \) of \( \Lambda \) is the measure \( \mu(F) \) of any fundamental domain \( F \) for \( \Lambda \).

Note that covolumes depend on the normalization of \( \mu \), but ratios of covolumes do not.

**Proposition 14.9.** If \( \Lambda' \subseteq \Lambda \) are lattices in \( V \simeq \mathbb{R}^n \), then \( \text{covol}(\Lambda') = [\Lambda : \Lambda'] \text{covol}(\Lambda) \).
Proof. Fix a fundamental domain $F$ for $\Lambda$ and a set of coset representatives $S$ for $\Lambda / \Lambda'$. Then

$$F' := \bigsqcup_{\lambda \in S} (F + \lambda)$$

is a fundamental domain for $\Lambda'$, and $\#S = [\Lambda : \Lambda'] = \mu(F')/\mu(F)$ is finite. We then have

$$\text{covol}(\Lambda') = \mu(F') = (\#S)\mu(F) = [\Lambda : \Lambda'] \text{covol}(\Lambda).$$

Definition 14.10. Let $S$ be a subset of a real vector space. The set $S$ is symmetric if it is closed under negation, and convex if for all $x, y \in S$ we have $\{tx + (1-t)y : t \in [0,1]\} \subseteq S$.

Theorem 14.11 (Minkowski’s Lattice Point Theorem). Let $\Lambda$ be a lattice in $V \simeq \mathbb{R}^n$ and $\mu$ a Haar measure on $V$. If $S \subseteq V$ is a symmetric convex measurable set that satisfies

$$\mu(S) > 2^n \text{covol}(\Lambda),$$

then $S$ contains a nonzero element of $\Lambda$.

Proof. See Problem Set 6.

Note that the inequality in Theorem 14.11 bounds the ratio of the measures of two sets ($S$ and a fundamental domain for $\Lambda$), and is thus independent of the choice of $\mu$.

14.2 The canonical inner product

Let $K/\mathbb{Q}$ be a number field of degree $n$ with $r$ real places and $s$ complex places; then $n = r + 2s$, by Corollary 13.9. We now want to consider the base change of $K$ to $\mathbb{R}$ and $\mathbb{C}$:

$$K_{\mathbb{R}} := K \otimes_{\mathbb{Q}} \mathbb{R} \simeq \mathbb{R}^r \times \mathbb{C}^s,$$

$$K_{\mathbb{C}} := K \otimes_{\mathbb{Q}} \mathbb{C} \simeq \mathbb{C}^n.$$  

The isomorphism $K_{\mathbb{R}} \simeq \mathbb{R}^r \times \mathbb{C}^s$ follows from Theorem 13.5 and the isomorphism $K_{\mathbb{C}} \simeq \mathbb{C}^n$ follows from the fact that $\mathbb{C}$ is separably closed; see Example 4.31. We note that $K_{\mathbb{R}}$ is an $\mathbb{R}$-vector space of dimension $n$, thus $K_{\mathbb{R}} \simeq \mathbb{R}^n$, but this is an isomorphism of $\mathbb{R}$-vector spaces and is not an $\mathbb{R}$-algebra isomorphism unless $s = 0$.

We have a sequence of injective homomorphisms of topological rings

$$\mathcal{O}_K \hookrightarrow K \hookrightarrow K_{\mathbb{R}} \hookrightarrow K_{\mathbb{C}},$$

which are defined as follows:

- the map $\mathcal{O}_K \hookrightarrow K$ is inclusion;
- the map $K \hookrightarrow K_{\mathbb{R}} = K \otimes_{\mathbb{Q}} \mathbb{R}$ is the canonical embedding $\alpha \mapsto \alpha \otimes 1$;
- the map $K_{\mathbb{R}} \simeq \mathbb{R}^r \times \mathbb{C}^s \hookrightarrow \mathbb{C}^r \times \mathbb{C}^{2s} \simeq K_{\mathbb{C}}$ embeds each factor of $\mathbb{R}^r$ in a corresponding factor of $\mathbb{C}^r$ via inclusion and each $\mathbb{C}$ in $\mathbb{C}^s$ is mapped to $\mathbb{C} \times \mathbb{C}$ in $\mathbb{C}^{2s}$ via $z \mapsto (z, \bar{z})$.

To better understand the last map, note that each $\mathbb{C}$ in $\mathbb{C}^s$ arises as $\mathbb{R}[\alpha] = \mathbb{R}[x]/(f) \simeq \mathbb{C}$ for some monic irreducible $f \in \mathbb{R}[x]$ of degree 2, but when we base-change to $\mathbb{C}$ the field $\mathbb{R}[\alpha]$ splits into the étale algebra $\mathbb{C}[x]/(x - \alpha) \times \mathbb{C}[x]/(x - \bar{\alpha}) \simeq \mathbb{C} \times \mathbb{C}$. The composition $K \hookrightarrow K_{\mathbb{R}} \hookrightarrow K_{\mathbb{C}}$ is given by the map

$$x \mapsto (\sigma_1(x), \ldots, \sigma_n(x)),$$
where \( \text{Hom}_\mathbb{Q}(K, \mathbb{C}) = \{\sigma_1, \ldots, \sigma_n\} \). If we put \( K = \mathbb{Q}(\alpha) := K[x]/(f) \) and let \( \alpha_1, \ldots, \alpha_n \in \mathbb{C} \) be the roots of \( f \) in \( \mathbb{C} \), each \( \sigma_i \) is the \( \mathbb{Q} \)-algebra homomorphism \( K \to \mathbb{C} \) defined by \( \alpha \mapsto \alpha_i \).

If we fix a \( \mathbb{Z} \)-basis for \( \mathcal{O}_K \), its image under the maps in (3) is a \( \mathbb{Q} \)-basis for \( K \), a \( \mathbb{R} \)-basis for \( K_\mathbb{R} \), and a \( \mathbb{C} \)-basis for \( K_\mathbb{C} \), all of which are vector spaces of dimension \( n = [K : \mathbb{Q}] \). We may thus view the injections in (3) as inclusions of topological groups (but not rings!)

\[
\mathbb{Z}^n \hookrightarrow \mathbb{Q}^n \hookrightarrow \mathbb{R}^n \hookrightarrow \mathbb{C}^n.
\]

The ring of integers \( \mathcal{O}_K \) is a lattice in the real vector space \( K_\mathbb{R} \simeq \mathbb{R}^n \), which inherits an inner product from the canonical Hermitian inner product on \( K_\mathbb{C} \simeq \mathbb{C}^n \) defined by

\[
\langle z, z' \rangle := \sum_{i=1}^{n} z_i \bar{z}_i' \in \mathbb{C}.
\]

For elements \( x, y \in K \hookrightarrow K_\mathbb{R} \hookrightarrow K_\mathbb{C} \) the Hermitian inner product can be computed as

\[
\langle x, y \rangle := \sum_{\sigma \in \text{Hom}_\mathbb{Q}(K, \mathbb{C})} \sigma(x) \overline{\sigma(y)} \in \mathbb{R},
\]

which is a real number because the non-real embeddings in \( \text{Hom}_\mathbb{Q}(K, \mathbb{C}) \) come in complex conjugate pairs. The inner product defined in (4) agrees with the restriction of the Hermitian inner product on \( K_\mathbb{R} \hookrightarrow K_\mathbb{C} \). The metric space topology it induces on \( K_\mathbb{R} \) is the same as the Euclidean topology on \( K_\mathbb{R} \simeq \mathbb{R}^n \) induced by the usual dot product on \( \mathbb{R}^n \), but the corresponding norm \( \|x\| := \langle x, x \rangle \) has a different normalization, as we now explain.

If we write elements \( z \in K_\mathbb{C} \simeq \mathbb{C}^n \) as vectors \( (z_\sigma) \) indexed by the set \( \sigma \in \text{Hom}_\mathbb{Q}(K, \mathbb{C}) \) in some fixed order, we may identify \( K_\mathbb{R} \) with its image in \( K_\mathbb{C} \) as the set

\[
K_\mathbb{R} = \{z \in K_\mathbb{C} : \bar{z}_\sigma = z_\bar{\sigma} \text{ for all } \sigma \in \text{Hom}_\mathbb{Q}(K, \mathbb{C})\}.
\]

For real embeddings \( \sigma = \bar{\sigma} \) we have \( z_\sigma \in \mathbb{R} \subseteq \mathbb{C} \), and for pairs of conjugate complex embeddings \( (\sigma, \bar{\sigma}) \) we get the embedding \( z \mapsto (z_\sigma, z_{\bar{\sigma}}) \) of \( \mathbb{C} \) into \( \mathbb{C} \times \mathbb{C} \) used to defined the map \( K_\mathbb{R} \hookrightarrow K_\mathbb{C} \) above. Each \( z \in K_\mathbb{R} \) can be uniquely written in the form

\[
(w_1, \ldots, w_r, x_1 + iy_1, x_1 - iy_1, \ldots, x_s + iy_s, x_s - iy_s),
\]

with \( w_i, x_j, y_j \in \mathbb{R} \). Each \( w_i \) corresponds to a \( z_\sigma \) with \( \sigma = \bar{\sigma} \), and each \( (x_j + iy_j, x_j - iy_j) \) corresponds to a complex conjugate pair \( (z_\sigma, z_{\bar{\sigma}}) \) with \( \sigma \neq \bar{\sigma} \). The canonical inner product on \( K_\mathbb{R} \) can then be written as

\[
\langle z, z' \rangle = \sum_{i=1}^{r} w_i w_i' + 2 \sum_{j=1}^{s} (x_j x_j' + y_j y_j').
\]

Thus if we take \( w_1, \ldots, w_r, x_1, y_1, \ldots, x_s, y_s \) as coordinates for \( K_\mathbb{R} \simeq \mathbb{R}^n \) (as \( \mathbb{R} \)-vector spaces), in order to normalize the Haar measure \( \mu \) on \( K_\mathbb{R} \) so that it is consistent with the Lebesgue measure \( \mu_\mathbb{R}^n \) on \( \mathbb{R}^n \) we define

\[
\mu(S) := 2^s \mu_\mathbb{R}^n(S)
\]

for any measurable set \( S \subseteq K_\mathbb{R} \) that we may view as a subset of \( \mathbb{R}^n \) by expressing it in \( w_i, x_j, y_j \) coordinates as above. With this normalization, the identity (1) still holds when
we replace $\mu_{\mathbb{R}^n}$ with $\mu$ and the dot product on $\mathbb{R}^n$ with the Hermitian inner product on $K_{\mathbb{R}}$, that is, for any $\mathbb{R}$-basis $e_1, \ldots, e_n$ of $K_{\mathbb{R}}$ we still have

$$\mu(F(e_1, \ldots, e_n)) = \sqrt{|\det([e_i, e_j]_{ij})|}$$

(7)

Using the Hermitian inner product on $K_{\mathbb{R}} \subseteq K_{\mathbb{C}}$ rather than the dot product on $K_{\mathbb{R}} \simeq \mathbb{R}^n$ multiplies $2s$ of the columns in the matrix $[e_i, e_j]_{ij}$ by 2, and thus multiplies the RHS by $\sqrt{2^{2s}} = 2^s$; our normalization of $\mu = 2^s \mu_{\mathbb{R}^n}$ multiplies the LHS by $2^s$ so that (7) still holds.

### 14.3 Covolumes of fractional ideals

Having fixed a normalized Haar measure $\mu$ for $K_{\mathbb{R}}$, we can now compute covolumes of lattices in $K_{\mathbb{R}} \simeq \mathbb{R}^n$. This includes not only (the image of) the ring of integers $\mathcal{O}_K$, but also any nonzero fractional ideal $I$ of $\mathcal{O}_K$: every such $I$ contains a nonzero principal fraction ideal $a\mathcal{O}_K$, and if $e_1, \ldots, e_n$ is a $\mathbb{Z}$-basis for $\mathcal{O}_K$ then $ae_1, \ldots, ae_n$ is a $\mathbb{Z}$-basis for $a\mathcal{O}_K$ that is an $\mathbb{R}$-basis for $K_{\mathbb{R}}$ that lies in $I$.

Recall from Remark 12.14 that the discriminant of a number field $K$ is the integer

$$D_K := \text{disc } \mathcal{O}_K := \text{disc}(e_1, \ldots, e_n) \in \mathbb{Z}.$$ 

### Proposition 14.12

Let $K$ be a number field. Using the normalized Haar measure on $K_{\mathbb{R}}$ defined in (6),

$$\text{covol}(\mathcal{O}_K) = \sqrt{|D_K|}.$$ 

**Proof.** Let $e_1, \ldots, e_n \in \mathcal{O}_K$ be a $\mathbb{Z}$-basis for $\mathcal{O}_K$, let $\text{Hom}_\mathbb{Q}(K, \mathbb{C}) = \{\sigma_1, \ldots, \sigma_n\}$, and define $A := [\sigma_i(e_j)]_{ij} \in \mathbb{C}^{n \times n}$. Then $D_K = \text{disc}(e_1, \ldots, e_n) = (\det A)^2$, by Proposition 12.6.

Viewing $\mathcal{O}_K \hookrightarrow K_{\mathbb{R}}$ as a lattice in $K_{\mathbb{R}}$ with basis $e_1, \ldots, e_n$, we may use (7) to compute $\text{covol}(\mathcal{O}_K) = \mu(F(e_1, \ldots, e_n)) = \sqrt{|\det([e_i, e_j]_{ij})|}$. Applying (4) yields

$$\det([e_i, e_j]_{ij}) = \det \left[ \sum_k \sigma_k(e_i) \sigma_k(e_j) \right] = \det(A^tA) = (\det A)(\det A).$$

Noting that $\det A$ is the square root of an integer (hence either real or purely imaginary), we have $\text{covol}(\mathcal{O}_K)^2 = |(\det A)^2| = |D_K|$, and the proposition follows. \hfill \Box

Recall from Remark 6.13 that for number fields $K$ we view the absolute norm

$$N : \mathcal{O}_K \to \mathbb{Z}$$

$$I \mapsto [\mathcal{O}_K : I]_\mathbb{Z}$$

as having image in $\mathbb{Q}_{>0}$ by identifying $N(I) \in \mathbb{Z}$ with a positive generator for $N(I)$ (note that $\mathbb{Z}$ is a PID). Recall that $[\mathcal{O}_K : I]_\mathbb{Z}$ is a module index of $\mathbb{Z}$-lattices in the $\mathbb{Q}$-vector space $K$, see Definitions 6.1 and 6.5), and for ideals $I \subseteq \mathcal{O}_K$ this is just the positive integer $[\mathcal{O}_K : I]_\mathbb{Z} = [\mathcal{O}_K : I]$. When $I = (a)$ is a principal fractional ideal with $a \in K$, we may simply write $N(a) := N((a)) = [N_{K/\mathbb{Q}}(a)]$

### Corollary 14.13

Let $K$ be a number field and let $I$ be a nonzero fractional ideal of $\mathcal{O}_K$. Then

$$\text{covol}(I) = N(I) \sqrt{|D_K|}.$$
Proof. Let $n = [K : \mathbb{Q}]$. Since $\text{covol}(bI) = b^n \text{covol}(I)$ and $N(bI) = b^n N(I)$ for any $b \in \mathbb{Z}_{>0}$, without loss of generality we may assume $I \subseteq \mathcal{O}_K$ (replace $I$ with a suitable $bI$ if not). Applying Propositions 14.9 and 14.12, we have

$$\text{covol}(I) = [\mathcal{O}_K : I] \text{covol}(\mathcal{O}_K) = N(I) \text{covol}(\mathcal{O}_K) = N(I) \sqrt{|D_K|}$$

as claimed. \qed

14.4 The Minkowski bound

Theorem 14.14. Minkowski bound Let $K$ be a number field of degree $n$ with $s$ complex places. Define the Minkowski constant $m_K$ for $K$ as the positive real number

$$m_K := \frac{n!}{n^n} \left(\frac{4}{\pi}\right)^s \sqrt{|D_K|}.$$

For every nonzero fractional ideal $I$ of $\mathcal{O}_K$ there is a nonzero $a \in I$ for which

$$N(a) \leq m_K N(I).$$

To prove this theorem we need the following lemma.

Lemma 14.15. Let $K$ be a number field of degree $n$ with $r$ real and $s$ complex places. For each $t \in \mathbb{R}_{>0}$, the measure of the convex symmetric set

$$S_t := \{(z_{\sigma}) \in K_\mathbb{R} : \sum |z_{\sigma}| \leq t\} \subseteq K_\mathbb{R}$$

with respect to the normalized Haar measure $\mu$ on $K_\mathbb{R}$ is

$$\mu(S_t) = 2^r \pi^s t^n/n!.$$

Proof. As in (5), we may uniquely write each $z = (z_{\sigma}) \in K_\mathbb{R}$ in the form

$$(w_1, \ldots, w_r, x_1 + iy_1, x_1 - iy_1, \ldots, x_s + iy_s, x_s - iy_s)$$

with $w_i, x_j, y_j \in \mathbb{R}$. We will have $\sum_{\sigma} |z_{\sigma}| \leq t$ if and only if

$$\sum_{i=1}^r |w_i| + \sum_{j=1}^s 2\sqrt{|x_j|^2 + |y_j|^2} \leq t. \quad (8)$$

We now compute the volume of this region in $\mathbb{R}^n$ by relating it to the volume of the simplex

$$U_t := \{(u_1, \ldots, u_n) \in \mathbb{R}_{>0}^n : u_1 + \cdots + u_n \leq t\} \subseteq \mathbb{R}^n,$$

which is $\mu_{\mathbb{R}^n}(U_t) = t^n/n!$ (volume of the standard simplex in $\mathbb{R}^n$ scaled by a factor of $t$).

If we view all the $w_i, x_j, y_j$ as fixed except the last pair $(x_s, y_s)$, then $(x_s, y_s)$ ranges over a disk of some radius $d \in [0, t/2]$ determined by (8). If we replace $(x_s, y_s)$ with $(u_{n-1}, u_n)$ ranging over the triangular region bounded by $u_{n-1} + u_n \leq 2d$ and $u_{n-1}, u_n \geq 0$, we need to incorporate a factor of $\pi/2$ to account for the difference between $(2d)^2/2 = 2d^2$ and $\pi d^2$; repeat this $s$ times. Similarly, if we hold everything but $w_r$ fixed and replace $w_r$ ranging over $[-d, d]$ for some $d \in [0, t]$ with $u_r$ ranging over $[0, d]$, we need to incorporate a factor of $2$ to account for this change of variable; repeat $r$ times. We then have

$$\mu(S_t) = 2^s \mu_{\mathbb{R}^n}(S_t) = 2^s \left(\frac{\pi}{2}\right)^s 2^r \mu_{\mathbb{R}^n}(U) = 2^r \pi^s t^n/n!.$$
Proof of Theorem 14.14. Let $I$ be a nonzero fractional ideal of $\mathcal{O}_K$. By Theorem 14.11, if we choose $t$ so that $\mu(S_t) > 2^n \cvol(I)$, then $S_t$ will contain a nonzero $a \in I$. By Lemma 14.15 and Corollary 14.13, it suffices to choose $t$ so that

$$\left(\frac{t}{n}\right)^n = \frac{n!\mu(S_t)}{n^n 2^n \pi^s} > \frac{n!2^n}{n^n 2^n \pi^s} \cvol(I) = \frac{n!}{n^n} \left(\frac{4}{n}\right)^s \sqrt{|D_K|N(I)} = m_K N(I).$$

Let us now pick $t$ so that $\left(\frac{t}{n}\right)^n > m_K N(I)$. Then $S_t$ contains $a \in I$ with $\sum_\sigma |\sigma(a)| \leq t$ Recalling that the geometric mean is bounded above by the arithmetic mean, we then have

$$N(a) = \left(N(a)^{1/n}\right)^n = \left(\prod_\sigma |\sigma(a)|^{1/n}\right)^n \leq \left(\frac{1}{n} \sum_\sigma |\sigma(a)|\right)^n \leq \left(\frac{t}{n}\right)^n,$$

Taking the limit as $\left(\frac{t}{n}\right)^n \rightarrow m_K N(I)$ from above yields $N(a) \leq m_K N(I)$. \hfill \Box

14.5 Finiteness of the ideal class group

Recall that the ideal class group $\text{cl}\mathcal{O}_K$ is the quotient of the ideal group $\mathcal{I}_K$ of $\mathcal{O}_K$ by its subgroup of principal fractional ideals. We now use the Minkowski bound to prove that every ideal class $[I] \in \text{cl}\mathcal{O}_K$ can be represented by an ideal $I \subseteq \mathcal{O}_K$ of small norm. It will then follow that the ideal class group is finite.

Theorem 14.16. Let $K$ be a number field. Every ideal class in $\text{cl}\mathcal{O}_K$ contains an ideal $I \subseteq \mathcal{O}_K$ of absolute norm $N(I) \leq m_K$, where $m_K$ is the Minkowski constant for $K$.

Proof. Let $[J]$ be an ideal class of $\mathcal{O}_K$ represented by the nonzero fractional ideal $J$. By Theorem 14.14, the fractional ideal $J^{-1}$ contains a nonzero element $a$ for which

$$N(a) \leq m_K N(J^{-1}) = m_K N(J)^{-1},$$

and therefore $N(aJ) = N(a)N(J) \leq m_K$. We have $a \in J^{-1}$, thus $aJ \subseteq J^{-1}J = \mathcal{O}_K$, so $I = aJ$ is an $\mathcal{O}_K$-ideal in the ideal class $[J]$ with $N(I) \leq m_K$ as desired. \hfill \Box

Lemma 14.17. Let $K$ be a number field and let $M$ be a real number. The set of ideals $I \subseteq \mathcal{O}_K$ with $N(I) \leq M$ is finite.

Proof 1. As a lattice in $K_{\mathbb{R}} \simeq \mathbb{R}^n$, the additive group $\mathcal{O}_K \simeq \mathbb{Z}^n$ has only finitely many subgroups $I$ of index $m$ for each positive integer $m \leq M$, since $[\mathbb{Z}^n : I] = m$ implies

$$(m\mathbb{Z})^n \subseteq I \subseteq \mathbb{Z}^n,$$

and $(m\mathbb{Z})^n$ has finite index $m^n = [\mathbb{Z}^n : m\mathbb{Z}] = [\mathbb{Z} : m\mathbb{Z}]^n$ in $\mathbb{Z}^n$. \hfill \Box

The proof of Lemma 14.17 is effective: the number of ideals $I \subseteq \mathcal{O}_K$ with $N(I) \leq M$ clearly cannot exceed $M^{n+1}$. But in fact we can give a much better bound than this.

Proof 2. Let $I$ be an ideal of absolute norm $N(I) \leq M$ and let $I = p_1 \cdots p_k$ be its factorization into (not necessarily distinct) prime ideals. Then $M \geq N(I) = N(p_1) \cdots N(p_k) \geq 2^k$, since the norm of each $p_i$ is a prime power, and in particular, at least 2. It follows that $k \leq \log_2 M$ is bounded, independent of $I$. Each prime ideal $p$ lies above some prime $p \leq M$, of which there are fewer than $M$, and for each prime $p$ the number of primes $p|p$ is at most $n$. Thus there are fewer than $(nM)^{\log_2 M}$ ideals of norm at most $M$ in $\mathcal{O}_K$. \hfill \Box
Corollary 14.18. Let \( K \) be a number field. The ideal class group of \( \mathcal{O}_K \) is finite.

Proof. By Theorem 14.16, each ideal class is represented by an ideal of norm at most \( m_K \), and by Lemma 14.17, the number of such ideals is finite. \( \square \)

Remark 14.19. For imaginary quadratic fields \( K = \mathbb{Q}(\sqrt{-d}) \) it is known that the class number \( h_K := \# \text{cl} \mathcal{O}_K \) tends to infinity as \( d \to \infty \) ranges over square-free integers. This was conjectured by Gauss in his *Disquisitiones Arithmeticae* [3] and proved by Heilbronn [5] in 1934; the first fully explicit lower bound was obtained by Oesterlé in 1988 [6]. This implies that there are only a finite number of imaginary quadratic fields with any particular class number. It was conjectured by Gauss that there are exactly 9 imaginary quadratic fields with class number one, but this was not proved until the 20th century by Stark [7] and Heegner [4]. Complete lists of imaginary quadratic fields for each class number \( h_K \leq 100 \) are now available [9]. By contrast, Gauss predicted that infinitely many real quadratic fields should have class number 1, however this question remains completely open. \(^2\)

Corollary 14.20. Let \( K \) be a number field of degree \( n \) with \( s \) complex places. Then

\[
|D_K| \geq \left( \frac{n^n}{n!} \right)^2 \left( \frac{\pi}{4} \right)^{2s} > \frac{1}{e^{2n}} \left( \frac{\pi e^2}{4} \right)^n.
\]

Proof. If \( I \) is an ideal and \( a \in I \) is nonzero, then \( N(a) \geq N(I) \), so Theorem 14.16 implies

\[
m_K = \frac{n!}{n^n} \left( \frac{4}{\pi} \right)^s \sqrt{|D_K|} \geq 1,
\]

the first inequality follows. The second uses an explicit form of Stirling’s approximation,

\[
n! \leq e\sqrt{n} \left( \frac{n}{e} \right)^n,
\]

and the fact that \( 2s \leq n \). \( \square \)

We note that \( \pi e^2/4 \approx 5.8 > 1 \), so the minimum value of \( |D_K| \) increases exponentially with \( n = [K : \mathbb{Q}] \). The lower bounds for \( n \in [2, 7] \) given by the corollary are listed below, along with the least value of \( |D_K| \) that actually occurs. As can be seen in the table, \( |D_K| \) appears to grow much faster than the corollary suggests. Better lower bounds can be proved using more advanced techniques, but a significant gap still remains.

<table>
<thead>
<tr>
<th>( n )</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>lower bound from Corollary 14.20</td>
<td>3</td>
<td>13</td>
<td>44</td>
<td>259</td>
<td>986</td>
<td>6267</td>
</tr>
<tr>
<td>minimum value of (</td>
<td>D_K</td>
<td>)</td>
<td>3</td>
<td>23</td>
<td>275</td>
<td>4511</td>
</tr>
</tbody>
</table>

Corollary 14.21. If \( K \) is a number field other than \( \mathbb{Q} \) then \( |D_K| > 1 \); equivalently, there are no nontrivial unramified extensions of \( \mathbb{Q} \).

Theorem 14.22. For every real \( M \) the set of number fields \( K \) with \( |D_K| < M \) is finite.

\(^1\)Heegner’s 1952 result [4] was essentially correct but contained some gaps that prevented it from being generally accepted until 1967 when Stark gave a complete proof in [7].

\(^2\)In fact it is conjectured that \( h_K = 1 \) for approximately 75.446\% of real quadratic fields with prime discriminant; this follows from the Cohen-Lenstra heuristics [2].
Proof. It follows from Corollary 14.20 that it suffices to prove this for fixed \(n := [K : \mathbb{Q}]\), since for all sufficiently large \(n\) we will have \(|D_K| > M\) for all number fields \(K\) of degree \(n\).

Case 1: Let \(K\) be a totally real field (so every place \(v\) is real) with \(|D_K| < M\). Then \(r = n\) and \(s = 0\), so \(K_R \cong \mathbb{R}^r \times \mathbb{C}^s = \mathbb{R}^n\). Consider the convex symmetric set
\[
S := \{(x_1, \ldots, x_n) \in K_R \cong \mathbb{R}^n : |x_1| \leq \sqrt{M} \text{ and } |x_i| < 1 \text{ for } i > 1\}
\]
with measure
\[
\mu(S) = 2\sqrt{M}2^{n-1} = 2^n \sqrt{M} > 2^n \sqrt{|D_K|} = 2^n \text{ covol}(\mathcal{O}_K).
\]
By Theorem 14.11, the set \(S\) contains a nonzero \(a \in \mathcal{O}_K \subseteq K \hookrightarrow K_R\) that we may write as \(a = (a_1, \ldots, a_n) = (\sigma_1(a), \ldots, \sigma_n(a))\), where the \(\sigma_i\) are the \(n\) embeddings of \(K\) into \(\mathbb{C}\), all of which are real embeddings. We have
\[
N(a) = \left| \prod \sigma_i(a) \right| \geq 1,
\]
since \(N(a)\) must be a positive integer, and \(|a_2|, \ldots, |a_n| < 1\), so \(|a_1| > 1 > |a_i| \text{ for all } i \neq 1\).

We claim that \(K = \mathbb{Q}(a)\). If not, each \(a_i = \sigma_i(a)\) would be repeated \([K : \mathbb{Q}(a)] > 1\) times in the vector \((a_1, \ldots, a_n)\), since there must be \([K : \mathbb{Q}(a)]\) elements of \(\text{Hom}_{\mathbb{Q}}(K, \mathbb{C})\) that fix \(\mathbb{Q}(a)\), namely, those lying in the kernel of the map \(\text{Hom}_{\mathbb{Q}}(K, \mathbb{C}) \to \text{Hom}_{\mathbb{Q}}(\mathbb{Q}(a), \mathbb{C})\) induced by restriction. But this is impossible since \(a_i \neq a_1\) for \(i \neq 1\).

The minimal polynomial \(f \in \mathbb{Z}[x]\) of \(a\) is a monic irreducible polynomial of degree \(n\). The roots of \(f(x)\) in \(\mathbb{C}\) are precisely the \(a_i = \sigma_i(a) \in \mathbb{R}\), all of which are bounded by \(|a_i| \leq \sqrt{M}\). Each coefficient \(f_i\) of \(f(x)\) is an elementary symmetric functions of its roots, hence also bounded in absolute value (certainly \(|f_i| \leq 2^n M^{n/2}\) for all \(i\)). The \(f_i\) are integers, so there are only finitely many possibilities for \(f(x)\), hence only finitely many totally real number fields \(K\) of degree \(n\).

Case 2: \(K\) has \(r\) real and \(s > 0\) complex places, and \(K_R \cong \mathbb{R}^r \times \mathbb{C}^s\). Now let
\[
S := \{(w_1, \ldots, w_r, z_1, \ldots, z_s) \in K_R : |z_1|^2 < c\sqrt{M} \text{ and } |w_i|, |z_j| < 1 \text{ for } j > 1\}
\]
with \(c\) chosen so that \(\mu(S) > 2^n \text{ covol}(\mathcal{O}_K)\) (the exact value of \(c\) depends on \(s\) and \(n\)). The argument now proceeds as in case 1: we get a nonzero \(a \in \mathcal{O}_K \cap S\) for which \(K = \mathbb{Q}(a)\), and only a finite number of possible minimal polynomials \(f \in \mathbb{Z}[x]\) for \(a\).

Lemma 14.23. Let \(K\) be a number field of degree \(n\). For each prime number \(p\) we have
\[
v_p(D_K) \leq n[\log_p n] + n - 1.
\]
In particular, \(v_p(D_K) \leq n[\log_2 n] + n - 1\) for all \(p\).

Proof. We have
\[
v_p(D_K) = v_p(\mathcal{N}_{K/Q}(D_{K/Q})) = \sum_{q|p} f_q v_q(D_{K/Q})
\]
where \(\mathcal{D}_{K/Q}\) is the different ideal and \(f_q\) is the residue degree of \(q|p\). Using Theorem 12.26 to bound \(v_q(D_{K/Q})\) yields
\[
v_p(D_K) \leq \sum_{q|p} f_q(e_q - 1 + v_q(e_q)) = n - \sum_{q|p} f_q + \sum_{q|p} f_q e_q v_p(e_q) \leq n - 1 + n[\log_p n],
\]
where we have used \(-1\) as an upper bound on \(-\sum_{q|p} f_q\) and \([\log_p n]\) as an upper bound on each \(v_p(e_q)\) (since \(e_q \leq n\)), and the fact that \(\sum_{q|p} e_q f_q = n\) (by Theorem 5.34). \(\square\)
Remark 14.24. The bound in Lemma 14.23 is tight; it is achieved by $K = \mathbb{Q}[x]/(x^{p^e} - p)$, for example.

Theorem 14.25 (Hermite). Let $S$ be a finite set of places of $\mathbb{Q}$, and let $n$ be an integer. The number of extensions $K/\mathbb{Q}$ of degree $n$ unramified outside of $S$ is finite.

Proof. By Lemma 14.23, since $n$ is fixed, the valuation $v_p(D_K)$ is bounded for each $p \in S$ and must be zero for $p \notin S$. Thus $|D_K|$ is bounded, and the theorem then follows from Proposition 14.22. 

References


