14 The Minkowski bound and finiteness results

14.1 Lattices in real vector spaces

In previous lectures we defined, for an integral domain A, the notion of an A-*lattice* in a finite dimensional vector space V over its fraction field K as a finitely generated A-submodule of V that spans K. We now want to specialize to the case $A = \mathbb{Z}$, in which case every A-lattice is free as a $\mathbb{Z}$-module (because $\mathbb{Z}$ is a PID and a submodule of a vector space is torsion-free). Rather than working with the fraction field $K = \mathbb{Q}$ we will instead work with its archimedean completion $\mathbb{R}$. We now take V to be a vector space over $\mathbb{R}$ and may regard it as a topological space isomorphic to $\mathbb{R}^n$ (by Proposition 10.6, there is a unique topology on V compatible with the topology on $\mathbb{R}$).

Recall that a subset S of a topological group is *discrete* if every $s \in S$ has an open neighborhood U for which $S \cap U = \{s\}$; equivalently, the subspace topology on S is the discrete topology. A subgroup H of a topological group G is said to be *cocompact* if it is normal and the quotient G/H is compact.

Definition 14.1. Let V be a real vector space of finite dimension. A (full) *lattice* in V is a free $\mathbb{Z}$-module $\Lambda \subseteq V$ that spans V as a real vector space. Equivalently, Λ is a discrete cocompact subgroup of V (see Problem Set 7).

Remark 14.2. A discrete subgroup of a Hausdorff topological group is necessarily closed; see [1, III.2.1.5] for a proof. This is easy to see for lattices: $\mathbb{Z}$ is closed in $\mathbb{R}$ (it is the complement of a union of open intervals), so $\mathbb{Z}^n$ is closed in $\mathbb{R}^n$. Given a lattice Λ in V, each $\mathbb{Z}$-basis for Λ determines an isomorphism of topological groups $\Lambda \simeq \mathbb{Z}^n$ and $V \simeq \mathbb{R}^n$.

Remark 14.3. You might ask why we are using the archimedean completion $\mathbb{R}$ of $\mathbb{Q}$ rather than some other completion $\mathbb{Q}_p$ of $\mathbb{Q}$. The reason is that $\mathbb{Z}$ is not a discrete subset of $\mathbb{Q}_p$ (elements of $\mathbb{Z}$ can be arbitrarily close to 0 under the p-adic metric).

As a locally compact group, $V \simeq \mathbb{R}^n$ has a Haar measure μ that is unique up to a scaling. Any basis $u_1, \ldots, u_n$ for V determines a parallelepiped

$$F(u_1, \ldots, u_n) := \{a_1u_1 + \cdots + a_nu_n : a_1, \ldots, a_n \in [0, 1)\}$$

that we may view as the unit cube by taking $\varphi\colon V \xrightarrow{\sim} \mathbb{R}^n$ to be the isomorphism that maps $(u_1, \ldots, u_n)$ to the standard basis for $\mathbb{R}^n$ and normalizing the Haar measure μ so that $\mu(F(u_1, \ldots, u_n)) = 1$. For any measurable set $S \subseteq \mathbb{R}^n$ we then have $\mu_{\mathbb{R}^n}(S) = \mu(\varphi(S))$, 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 $E = [e_{ij}]$ be the matrix whose jth column expresses $e_j = \sum_i e_{ij}u_i$, in terms of our standard basis $u_1, \ldots, u_n$, then

$$\mu(F(e_1, \ldots, e_n)) = |\det E| = \sqrt{\det E^t \det E} = \sqrt{\det(E^tE)} = \sqrt{\det[\langle e_i, e_j\rangle]_{ij}}, \qquad (1)$$

where $\langle e_, e_j\rangle$ 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 μ.

Proposition 14.4. *If $T: V \to V$ is a linear transformation on a real vector space $V \simeq \mathbb{R}^n$ with Haar measure μ, then for every measurable set S we have*

$$\mu(T(S)) = |\det T|\, \mu(S). \tag{2}$$

Proof. See [8, Ex. 1.2.21]. □

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

Definition 14.5. Let Λ be a lattice in $V \simeq \mathbb{R}^n$. A *fundamental domain* for Λ is a measurable set $F \subseteq V$ such that

$$V = \bigsqcup_{\lambda \in \Lambda} (F + \lambda).$$

In other words, F is a measurable set of coset representatives for V/Λ. 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 14.6. *Let Λ be a lattice in $V \simeq \mathbb{R}^n$ with Haar measure μ. Then $\mu(F) = \mu(G)$ for all fundamental domains F and G for Λ.*

Proof. Using the translation invariance and countable additivity of μ (note that $\Lambda \simeq \mathbb{Z}^n$ is a countable set) along with the fact that Λ is closed under negation, we obtain

$$\begin{aligned}\mu(F) = \mu(F \cap V) &= \mu\left(F \cap \bigsqcup_{\lambda \in \Lambda} (G + \lambda)\right) = \mu\left(\bigsqcup_{\lambda \in \Lambda} (F \cap (G + \lambda))\right)\\ &= \sum_{\lambda \in \Lambda} \mu(F \cap (G + \lambda)) = \sum_{\lambda \in \Lambda} \mu((F - \lambda) \cap G) = \sum_{\lambda \in \Lambda} \mu((G + \lambda) \cap F).\end{aligned}$$

The proposition then follows by symmetry (swap F and G in the derivation above). □

Definition 14.7. Let Λ be a lattice in $V \simeq \mathbb{R}^n$ with Haar measure μ. The *covolume* $\operatorname{covol}(\Lambda)$ of Λ is the volume $\mu(F)$ of any fundamental domain F for Λ.

Note that volumes and covolumes depend on the normalization of the Haar measure μ, but ratios of them do not. Regardless of the normalization, the covolume of a lattice Λ is finite (because Λ is cocompact) and nonzero (because Λ is discrete).

Proposition 14.8. *If $\Lambda' \subseteq \Lambda$ are lattices in a real vector space V of finite dimension then*

$$\operatorname{covol}(\Lambda') = [\Lambda : \Lambda'] \operatorname{covol}(\Lambda)$$

Proof. Fix a fundamental domain F for Λ and a set of coset representatives L for Λ/Λ'. Then

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

is a fundamental domain for Λ', and $\#L = [\Lambda : \Lambda'] = \mu(F')/\mu(F)$ is finite, since F' and F both have finite nonzero measure. We then have

$$\operatorname{covol}(\Lambda') = \mu(F') = (\#L)\mu(F) = [\Lambda : \Lambda'] \operatorname{covol}(\Lambda).$$ □

Definition 14.9. 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 every pair of points $x, y \in S$ the line segment $\{tx + (1-t)y : t \in [0,1]\}$ between them lies in S.

Theorem 14.10 (Minkowski's Lattice Point Theorem)**.** *Let Λ be a lattice in a real vector space $V \simeq \mathbb{R}^n$ with Haar measure μ. If $S \subseteq V$ is a symmetric convex set such that*

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

then S contains a nonzero element of Λ.

Proof. See Problem Set 6. □

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, so that $n = r + 2s$. We then have

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

(the first isomorphism was proved in Lecture 13 and the second follows from the fact that every étale algebra over a separably closed field splits (see Example 4.30). We have a sequence of injective homomorphisms of topological groups

$$\mathcal{O}_K \hookrightarrow K \hookrightarrow K_\mathbb{R} \hookrightarrow K_\mathbb{C}, \tag{3}$$

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 $K \hookrightarrow K_\mathbb{C}$ is $\alpha \mapsto (\sigma_1(\alpha), \ldots, \sigma_n(\alpha))$, where $\operatorname{Hom}_\mathbb{Q}(K, \mathbb{C}) = \{\sigma_1, \ldots, \sigma_n\}$, which factors through the map $K_\mathbb{R} \hookrightarrow K_\mathbb{C}$ defined below;
- 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 $\mathbb{C}[x]/(x-\alpha) \times \mathbb{C}[x]/(x-\bar{\alpha}) \simeq \mathbb{C} \times \mathbb{C}$.

If we fix a $\mathbb{Z}$-basis for $\mathcal{O}_K$, the image of this basis is a $\mathbb{Q}$-basis for K, an $\mathbb{R}$-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

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

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

$$\langle (a_1, \ldots, a_n), (b_1, \ldots, b_n) \rangle := \sum_{i=1}^n a_i \bar{b}_i \in \mathbb{C}.$$

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

$$\langle x, y\rangle := \sum_{\sigma \in \operatorname{Hom}_{\mathbb{Q}}(K,\mathbb{C})} \sigma(x)\overline{\sigma(y)} \in \mathbb{R}, \tag{4}$$

which is a real number because the non-real embeddings in $\operatorname{Hom}_{\mathbb{Q}}(K,\mathbb{C})$ come in complex conjugate pairs. The inner product defined in (4) is the *canonical inner product* on $K_{\mathbb{R}}$ (it applies to all of $K_{\mathbb{R}}$, not just the image of K in $K_{\mathbb{R}}$). The topology it induces on $K_{\mathbb{R}}$ is the same as the Euclidean topology on $\mathbb{R}^r \times \mathbb{C}^s$, but the corresponding norm $\|\ \|$ has a different normalization, as we now explain.

If we write the elements of $K_{\mathbb{C}} \simeq \mathbb{C}^n$ as vectors (z_σ) indexed by $\sigma \in \operatorname{Hom}_{\mathbb{Q}}(K,\mathbb{C})$, we may identify $K_{\mathbb{R}}$ with its image in $K_{\mathbb{C}}$ as the set

$$K_{\mathbb{R}} = \{(z_\sigma) \in K_{\mathbb{C}} : \bar{z}_\sigma = z_{\bar{\sigma}}\}.$$

When $\sigma = \bar{\sigma}$ is a real embedding we have $z \mapsto z_\sigma \in \mathbb{R} \subseteq \mathbb{C}$, while for pairs of conjugate complex embeddings $(\sigma, \bar{\sigma})$ we get the embedding $z \mapsto (z_\sigma, z_{\bar{\sigma}}) = (z_\sigma, \bar{z}_\sigma)$ of $\mathbb{C}$ into $\mathbb{C} \times \mathbb{C}$ noted above. Each vector $(z_\sigma) \in K_{\mathbb{R}}$ can be written uniquely in the form

$$(w_1, \ldots, w_r,\ x_1 + iy_1,\ x_1 - iy_1,\ \ldots,\ x_s + iy_s,\ x_s - iy_s), \tag{5}$$

with $w_i, y_j, z_j \in \mathbb{R}$, where each z_i corresponds to a z_σ 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 then becomes

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

Thus if we take the w_i, x_j, y_j as coordinates for $R^n \simeq \mathbb{R}^r \times \mathbb{C}^s \simeq K_{\mathbb{R}}$ (as $\mathbb{R}$-vector spaces), in order to normalize the Haar measure μ on $K_{\mathbb{R}}$ so that it is consistent with the Lebesgue measure $\mu_{\mathbb{R}^n}$ on $\mathbb{R}^n$ we define

$$\mu(S) \coloneqq 2^s \mu_{\mathbb{R}^n}(S),$$

for any measurable set S in $K_{\mathbb{R}}$ that we view as a subset of $\mathbb{R}^n$ by expressing it in w_i, x_j, y_j coordinates via the canonical embedding $z \mapsto (z_\sigma)$ as explained above.

Having fixed a normalized Haar measure μ for $K_{\mathbb{R}}$, we can now compute the covolume of the lattice $\mathcal{O}_K$ in $K_{\mathbb{R}}$.

14.3 Covolumes of fractional ideals

Let K be a number field. Recall that a $\mathbb{Z}$-lattice in the $\mathbb{Q}$-vector space K is a finitely generated $\mathbb{Z}$ module with $\mathbb{Q}$-span K. Every $\mathbb{Z}$-lattice M in K corresponds to a lattice in the $\mathbb{R}$-vector space $K_{\mathbb{R}}$ under the canonical embedding $K \hookrightarrow K \otimes_{\mathbb{Q}} \mathbb{R} = K_{\mathbb{R}}$: the image of M is still a finitely generated $\mathbb{Z}$-module, and any $\mathbb{Q}$-basis for K that lies in M gets mapped to an $\mathbb{R}$-basis for $K_{\mathbb{R}}$ that lies in the image of M. We may thus view any fractional ideal of $\mathcal{O}_K$ (including $\mathcal{O}_K$ itself) as a lattice in $K_{\mathbb{R}}$. We now determine the covolume of these lattices.

Proposition 14.11. *Let K be a number field with ring of integers $\mathcal{O}_K$. Then*

$$\operatorname{covol}(\mathcal{O}_K) = \sqrt{|\operatorname{disc}\mathcal{O}_K|}.$$

Proof. Let $e_1, \ldots, e_n \in \mathcal{O}_K$ be a $\mathbb{Z}$-basis for $\mathcal{O}_K$, let $\operatorname{Hom}_{\mathbb{Q}}(K, \mathbb{C}) = \{\sigma_1, \ldots, \sigma_n\}$, and define $A := [\sigma_i(e_j)]_{ij} \in \mathbb{C}^{n \times n}$. 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 (1) to compute $\operatorname{covol}(\mathcal{O}_K)^2 = \mu(F(e_1, \ldots, e_n))^2$ as

$$\begin{aligned}\operatorname{covol}(\mathcal{O}_K)^2 &= \det[\langle e_i, e_j \rangle]_{ij} = \det\left[\sum_k \sigma_k(e_i)\overline{\sigma_k(e_j)}\right]_{ij} \\ &= \det(A^{\mathrm{t}}\overline{A}) = (\det A)(\overline{\det A}) \\ &= |\det A|^2 = |\operatorname{disc}\mathcal{O}_K|^2,\end{aligned}$$

where the last line follows from Proposition 12.6. □

Recall from Remark 6.12 that for number fields K we view the absolute norm

$$\begin{aligned}\mathrm{N}\colon \mathcal{I}_{\mathcal{O}_K} &\to \mathcal{I}_{\mathbb{Z}} \\ I &\mapsto [\mathcal{O}_K : I]_{\mathbb{Z}}\end{aligned}$$

as having image in $\mathbb{Q}_{>0}$ by identifying $\mathrm{N}(I) = (t) \in \mathcal{I}_{\mathbb{Z}}$ with $t \in \mathbb{Q}_{>0}$ (here $[\mathcal{O}_K : I]_{\mathbb{Z}}$ is a module index of $\mathbb{Z}$-lattices in the $\mathbb{Q}$-vector space K, see Definitions 6.1 and 6.4). 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 $\mathrm{N}(a) := \mathrm{N}((a)) = |\mathrm{N}_{K/\mathbb{Q}}(a)|$

Corollary 14.12. *Let K be a number field and let I be a nonzero fractional ideal of $\mathcal{O}_K$. Then*

$$\operatorname{covol}(I) = \mathrm{N}(I)\sqrt{|\operatorname{disc}\mathcal{O}_K|}$$

Proof. Let $n = [K : \mathbb{Q}]$. Since $\operatorname{covol}(bI) = b^n \operatorname{covol}(I)$ and $\mathrm{N}(bI) = b^n \mathrm{N}(I)$ for any $b \in \mathbb{Z}_{\geq 0}$, without loss of generality we may assume $I \subseteq \mathcal{O}_K$ (replace I with a suitable bI if not). Applying Propositions 14.8 and 14.11, we have

$$\operatorname{covol}(I) = [\mathcal{O}_K : I]\operatorname{covol}(\mathcal{O}_K) = \mathrm{N}(I)\operatorname{covol}(\mathcal{O}_K) = \mathrm{N}(I)\sqrt{|\operatorname{disc}\mathcal{O}_K|}$$

as claimed. □

14.4 The Minkowski bound

Theorem 14.13 (Minkowski bound)**.** *Let K be a number field of degree $n = r + 2s$ with s complex places. Define the Minkowski constant m_K for K as the positive real number*

$$m_K := \frac{n!}{n^n}\left(\frac{4}{\pi}\right)^s \sqrt{|\operatorname{disc}\mathcal{O}_K|}.$$

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

$$\mathrm{N}(a) \leq m_K \mathrm{N}(I).$$

Before proving the theorem we first prove a lemma.

Lemma 14.14. *Let K be a number field of degree $n = r + 2s$ with r real and s complex places. For each $t \in \mathbb{R}_{>0}$, the volume of the convex symmetric set*

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

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

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

Proof. As in (5), we may uniquely write each $(z_\sigma) \in \mathcal{K}_\mathbb{R}$ 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 $\sum_\sigma |z_\sigma| \le t$ if and only if

$$\sum_{i=1}^{r} |w_i| + \sum_{j=1}^{s} 2\sqrt{|x_j|^2 + |y_j|^2} \le t. \tag{6}$$

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

$$U := \{(u_1, \ldots, u_n) \in \mathbb{R}^n \colon \sum u_i \le t \text{ and } u_i \ge 0\} \subseteq \mathbb{R}^n,$$

which is $\mu_{\mathbb{R}^n}(U) = t^n/n!$ (the 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 (6) (the value of d depends on the fixed values of w_i, x_j, y_j for $1 \le i \le r$ and $1 \le j \le s-1$). If we replace (x_s, y_s) with (u_{n-1}, u_n) ranging over the triangular region bounded by $u_{n-1} + u_n \le 2d$ and $u_{n-1}, u_n \ge 0$, we need to incorporate a factor of $\pi/2$ to account for the difference between $(2d^2)/2 = 2d^2$ and πd^2; repeat this s times. Similarly, we now 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 $[0, d]$, and incorporate a factor of 2 to account for this change of variable; repeat r times. We then have

$$\mu(S_t) = 2^s \mu_{\mathbb{R}^n}(S_t) = 2^s \left(\frac{\pi}{2}\right)^s 2^r \mu_{\mathbb{R}^n}(U) = 2^r \pi^s \frac{t^n}{n!}$$

as desired. This completes the proof of the lemma. □

Proof of Theorem 14.13. Let I be a nonzero fractional ideal of $\mathcal{O}_K$. By Theorem 14.10 and Corollary 14.12, if we choose t so that

$$\mu(S_t) > 2^n \operatorname{covol}(I) = 2^n \mathrm{N}(I)\sqrt{|\operatorname{disc}\mathcal{O}_K|},$$

then S_t will contain a nonzero element $a \in I$ satisfying

$$\sum_\sigma |\sigma(a)| \le t,$$

where σ ranges over the n elements of $\operatorname{Hom}_\mathbb{Q}(K, \mathbb{C})$. By Lemma 14.14, we want t to satisfy

$$2^r \pi^s \frac{t^n}{n!} = \mu(S_t) > 2^n \mathrm{N}(I)\sqrt{|\operatorname{disc}\mathcal{O}_K|},$$

equivalently,

$$t^n > \frac{2^{n-r} n!}{\pi^s} N(I)\sqrt{|\operatorname{disc}\mathcal{O}_K|} = n! \left(\frac{4}{\pi}\right)^s \sqrt{|\operatorname{disc}\mathcal{O}_K|}\mathrm{N}(I) = n^n m_K \mathrm{N}(I).$$

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

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

Taking the limit as $\left(\frac{t}{n}\right)^n \to m_K \mathrm{N}(I)$ from above yields $\mathrm{N}(a) \le m_K \mathrm{N}(I)$. □

14.5 Finiteness of the ideal class group

Recall that the ideal class group $\operatorname{cl}\mathcal{O}_K$ is the quotient of the ideal group of $\mathcal{O}_K$ by its subgroup of principal fractional ideals. We now use the Minkowski bound to prove that every ideal class $[I] \in \operatorname{cl}\mathcal{O}_K$ 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 14.15. *Let K be a number field. Every ideal class in $\operatorname{cl}\mathcal{O}_K$ contains an ideal $I \subseteq \mathcal{O}_K$ of absolute norm $\mathrm{N}(I) \leq m_K$, 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 14.13, the fractional ideal J^{-1} contains a nonzero element a for which

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

and therefore $\mathrm{N}(aJ) = \mathrm{N}(a)\mathrm{N}(J) \leq m_K$. 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 $\mathrm{N}(I) \leq m_K$ as desired. □

Lemma 14.16. *Let K be a number field and let $M > 1$ be a real number. The set of ideals $I \subseteq \mathcal{O}_K$ with $\mathrm{N}(I) \leq M$ 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 $m^n = [\mathbb{Z}^n : m\mathbb{Z}^n] = [\mathbb{Z} : m\mathbb{Z}]^n$ in $\mathbb{Z}^n$. □

The proof of Lemma 14.16 is effective: the number of ideals $I \subseteq \mathcal{O}_K$ with $\mathrm{N}(I) \leq M$ 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 $\mathrm{N}(I) \leq M$ and let $I = \mathfrak{p}_1 \cdots \mathfrak{p}_k$ be its factorization into (not necessarily distinct) prime ideals. Then $M \geq \mathrm{N}(I) = \mathrm{N}(\mathfrak{p}_1)\cdots\mathrm{N}(\mathfrak{p}_k) \geq 2^k$, since the norm of each $\mathfrak{p}_i$ 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 $\mathfrak{p}$ lies above some prime $p \leq M$, of which there are $\pi(M) \approx M/\log M \leq M$ (here $\pi(x)$ is the prime counting function), and for each prime p the number of primes $\mathfrak{p}|p$ is at most n. Thus there are at most $(n\pi(M))^{\log_2 M} \leq (nM)^{\log_2 M}$ ideals of norm at most M in $\mathcal{O}_K$. □

Corollary 14.17. *Let K be a number field. The ideal class group of $\mathcal{O}_K$ is finite.*

Proof. By Theorem 14.15, each ideal class is represented by an ideal of norm at most m_K, and distinct ideal classes must be represented by distinct ideals. By Lemma 14.16, the number of such ideals is finite. □

Remark 14.18. For imaginary quadratic fields $K = \mathbb{Q}(\sqrt{-d})$ it is known that the *class number* $h_K := \#\operatorname{cl}\mathcal{O}_K$ 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].[1] Complete lists of imaginary quadratic fields for each class number $h_K \le 100$ are now available [9].

The situation for real quadratic fields is quite different; it is generally believed that there are infinitely many real quadratic fields with class number 1.[2]

Corollary 14.19. *Let K be a number field of degree n with s complex places. Then*

$$|\operatorname{disc}\mathcal{O}_K| \ge \left(\frac{n^n}{n!}\right)^2\left(\frac{\pi}{4}\right)^{2s} > \frac{1}{2\pi n}\left(\frac{\pi e^2}{4}\right)^n.$$

Proof. The absolute norm of an integral ideal is a positive integer. By Theorem 14.15,

$$m_K = \frac{n!}{n^n}\left(\frac{4}{\pi}\right)^s\sqrt{|\operatorname{disc}\mathcal{O}_K|} \ge 1.$$

The first lower bound on $|\operatorname{disc}\mathcal{O}_K|$ follows from $s \le n/2$, and the second follows from

$$n! \ge \sqrt{2\pi n}\left(\frac{n}{e}\right)^n,$$

an explicit version of Stirling's approximation. □

We note that $\pi e^2/4 > 5.8$, so the minimum value of $|\operatorname{disc}\mathcal{O}_K|$ 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 $|\operatorname{disc}\mathcal{O}_K|$ that actually occurs. As can be seen in the table, $|\operatorname{disc}\mathcal{O}_K|$ appears to grow substantially faster than the corollary suggests. Better lower bounds can be proved using more advanced techniques, but a significant gap still remains.

	$n=2$	$n=3$	$n=4$	$n=5$	$n=6$	$n=7$
lower bound from Corollary 14.19	3	11	46	210	1014	5014
minimum value of $\lvert\operatorname{disc}\mathcal{O}_K\rvert$	3	23	275	4511	92799	2306599

Corollary 14.20. *If K is a number field other than $\mathbb{Q}$ then $|\operatorname{disc}\mathcal{O}_K| > 1$. Equivalently, there are no nontrivial unramified extensions of $\mathbb{Q}$.*

Theorem 14.21. *For $M \in \mathbb{R}$ the set of number fields K with $|\operatorname{disc}\mathcal{O}_K| < M$ is finite.*

Proof. Since we know that $|\operatorname{disc}\mathcal{O}_K| \to \infty$ as $n \to \infty$, it suffices to prove this for each fixed degree $n = [K:\mathbb{Q}]$.

Case 1: Let K be a totally real field (so every place $v|\infty$ is real) with $|\operatorname{disc}\mathcal{O}_K| < M$. Then $r = n$ and $s = 0$, so $K_\mathbb{R} \simeq \mathbb{R}^r \times \mathbb{C}^s = \mathbb{R}^n$. Consider the convex symmetric set

$$S := \{(x_1,\ldots,x_n) \in K_\mathbb{R} \simeq \mathbb{R}^n : |x_1| \le \sqrt{M} \text{ and } |x_i| < 1 \text{ for } i > 1\}.$$

Then

$$\mu(S) = 2\sqrt{M}2^{n-1} = 2^n\sqrt{M} > 2^n\sqrt{|\operatorname{disc}\mathcal{O}_K|} = 2^n\operatorname{covol}(\mathcal{O}_K),$$

[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].

so S contains a nonzero element $a \in \mathcal{O}_K \subseteq K \hookrightarrow K_{\mathbb{R}}$ that we may write as $a = (a_\sigma) = (\sigma_1(a), \ldots, \sigma_n(a))$, where the σ_i are the n embeddings of K into $\mathbb{C}$, all of which are real embeddings. We have

$$\mathrm{N}(a) = \left|\prod \sigma_i(a)\right| \geq 1$$

and $|a_2|, \ldots, |a_n| < 1$, so $|a_1| > 1 > |a_i|$ for $i = 2, \ldots, n$. In particular, $a_1 \neq a_i$ for any $i > 1$.

We now 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 $\mathrm{Hom}_{\mathbb{Q}}(K, \mathbb{C})$ that fix $\mathbb{Q}(a)$, namely, those lying in the kernel of the map $\mathrm{Hom}_{\mathbb{Q}}(K, \mathbb{C}) \to \mathrm{Hom}_{\mathbb{Q}}(\mathbb{Q}(a), \mathbb{C})$ 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 $|a_i| \leq \sqrt{M}$. The coefficients of $f(x)$ are elementary symmetric functions of its roots, hence also bounded in absolute value, and they are integers, so there are only finitely many possibilities for $f(x)$, given the bound M, hence only finitely many totally real number fields K of degree n.

Case 2: K has r real and $s > 0$ complex places, and $K_{\mathbb{R}} \simeq \mathbb{R}^r \times \mathbb{C}^s$. Now let

$$S := \{(w_1, \ldots, w_r,\ z_1,\ \ldots,\ z_s) \in K_{\mathbb{R}} : |z_1|^2 < c\sqrt{M} \text{ and } |w_i|, |z_j| < 1\ (j > 1)\}$$

with c chosen so that $\mu(S) > 2^n \operatorname{covol}(\mathcal{O}_K)$ (the exact value of c depends on s and n). The argument now proceeds as in case 1: we get a nonzero $a \in \mathcal{O}_K \cap S$ with $K = \mathbb{Q}(a)$, and only a finite number of possible minimal polynomials $f \in \mathbb{Z}[x]$ for a. □

Lemma 14.22. *Let K be a number field of degree n. For each prime $p \in \mathbb{Z}$ we have*

$$v_p(\operatorname{disc} \mathcal{O}_K) \leq n(\log_p n + 1) - 1.$$

In particular, $v_p(\operatorname{disc} \mathcal{O}_K) \leq n(\log_2 n + 1) - 1$ for all primes $p \in \mathbb{Z}$.

Proof. We have

$$|\operatorname{disc} \mathcal{O}_K|_p = |N_{K/\mathbb{Q}}(\mathcal{D}_{K/\mathbb{Q}})|_p = \prod_{v|p} |\mathcal{D}_{K_v/\mathbb{Q}_p}|_v,$$

where $\mathcal{D}_{K_v/\mathbb{Q}_p}$ denotes the different ideal. It follows from Theorem 12.26 that

$$v_p(\operatorname{disc} \mathcal{O}_K) \leq \sum_{v|p} (e_v - 1 + e_v v_p(e_v)),$$

where e_v is the ramification index of $K_v/\mathbb{Q}_p$. We have $\sum_{v|p} e_v \leq n$ and $v_p(e_v) \leq \log_p(n)$, so

$$v_p(\operatorname{disc} \mathcal{O}_K) \leq n(\log_p n + 1) - 1.$$ □

Remark 14.23. The bound in Lemma 14.22 is tight; it is achieved by $K = \mathbb{Q}[x]/(x^{p^e} - p)$, for example.

Theorem 14.24 (Hermite)**.** *Let S be a finite set of places of $\mathbb{Q}$, and let $n \in \mathbb{Z}_{>1}$. The number of extensions $K/\mathbb{Q}$ of degree n unramified outside of S is finite.*

Proof. By the lemma, since n is fixed, the valuation $v_p(\operatorname{disc} \mathcal{O}_K)$ is bounded for each $p \in S$ and most be zero for $p \notin S$. Thus $|\operatorname{disc} \mathcal{O}_K|$ is bounded and the theorem then follows from Proposition 14.21. □

References

[1] Nicolas Bourbaki, *General Topology: Chapters 1-4*, Springer, 1995.

[2] Henri Cohen and Hendrik W. Lenstra Jr., *Heuristics on class groups of number fields*, in *Number Theory (Noordwijkerhout 1983)*, Lecture Notes in Mathematics **1068**, Springer, 1984, 33–62.

[3] Carl F. Gauss, *Disquisitiones Arithmeticae*, Göttingen (1801), English translation by Arthur A. Clark, revised by William C. Waterhouse, Spring-Verlag 1986 reprint of Yale University Press 1966 edition.

[4] Kurt Heegner, *Diophantische Analysis und Modulfunktionen*, Math. Z. **56** (1952), 227–253.

[5] Hans Heilbronn, *On the class number in imaginary quadratic fields*, Quart. J. of Math. Oxford **5** (1934), 150–160.

[6] Joseph Oesterlé, *La probléme de Gauss sur le nombre de classes*, Enseign. Math. **34** (1988), 43–67.

[7] Harold Stark, *A complete determination of the complex quadratic fields of class–number one*, Mich. Math. J. **14** (1967), 1–27.

[8] Terence Tao, *An introduction to measure theory*, Graduate Studies in Mathematics **126**, AMS, 2010.

[9] Mark Watkins, *Class numbers of imaginary quadratic fields*, Math. Comp. **73** (2004), 907–938.

18.785 Number Theory I
Fall 2016