13 The Minkowski bound, finiteness results

13.1 Lattices in real vector spaces

In Lecture 6 we defined the notion of an A-*lattice* in a finite dimensional K-vector space V as a finitely generated A-submodule of V that spans V as a K-vector space, where K is the fraction field of the domain A. In our usual $AKLB$ setup, A is a Dedekind domain, L is a finite separable extension of K, and the integral closure B of A in L is an A-lattice in the K-vector space $V = L$. When B is a free A-module, its rank is equal to the dimension of L as a K-vector space and it has an A-module basis that is also a K-basis for L.

We now want to specialize to the case $A = \mathbb{Z}$, and rather than taking $K = \mathbb{Q}$, we will instead use the archimedean completion $\mathbb{R}$ of $\mathbb{Q}$. Since $\mathbb{Z}$ is a PID, every finitely generated $\mathbb{Z}$-module in an $\mathbb{R}$-vector space V is a free $\mathbb{Z}$-module (since it is necessarily torsion free). We will restrict our attention to free $\mathbb{Z}$-modules with rank equal to the dimension of V (sometimes called a *full lattice*).

Definition 13.1. Let V be a real vector space of dimension n. A (full) *lattice* in V is a free $\mathbb{Z}$-module of the form $\Lambda := e_1\mathbb{Z} + \cdots + e_n\mathbb{Z}$, where $(e_1, \ldots, e_n)$ is a basis for V.

Any real vector space V of dimension n is isomorphic to $\mathbb{R}^n$. By fixing an isomorphism, equivalently, choosing a basis for V that we identify with the standard basis for $\mathbb{R}^n$, we can equip V with an inner product $\langle \cdot, \cdot \rangle$ corresponding to the canonical inner product on $\mathbb{R}^n$ (the standard dot product). This makes V into a normed vector space with the norm

$$\|x\| := \sqrt{\langle x, x\rangle} \in \mathbb{R}_{\geq 0},$$

and also a metric space with distance metric

$$d(x, y) := \|x - y\|.$$

While the inner product $\langle \cdot, \cdot \rangle$ and distance metric $d(\cdot, \cdot)$ on V depend on our choice of basis (equivalently, the isomorphism $V \simeq \mathbb{R}^n$), the induced metric space topology does not; it is the same as the standard Euclidean topology on $\mathbb{R}^n$. The standard Lebesgue measure on $\mathbb{R}^n$ is the unique Haar measure that assigns measure 1 to the unit cube $[0, 1]^n$. This is consistent with Euclidean norm on $\mathbb{R}^n$, which assigns length 1 to the standard unit vectors. Having fixed an inner product $\langle \cdot, \cdot \rangle$ on $V \simeq \mathbb{R}^n$, we normalize the Haar measure on V so that the volume of a unit cube defined by any basis for V that is orthonormal with respect to $\langle \cdot, \cdot \rangle$ has measure 1.

Recall that a subset S of a topological space X is *discrete* if every $s \in S$ lies in an open neighborhood $U \subseteq X$ that intersects S only at s.

Proposition 13.2. *Let Λ be a subgroup of a real vector space V of finite dimension. Then Λ is a lattice if and only if Λ is discrete and V/Λ is compact (Λ is cocompact).*

Proof. Suppose $\Lambda = e_1\mathbb{Z} + \cdots e_n\mathbb{Z}$ is a lattice; then $e_1, \ldots, e_n$ is a basis for V. This basis determines an isomorphism $V \xrightarrow{\sim} \mathbb{R}^n$ of topological groups that sends Λ to $\mathbb{Z}^n \subseteq \mathbb{R}^n$. The subgroup $\mathbb{Z}^n \subseteq \mathbb{R}^n$ is clearly discrete and the quotient $\mathbb{R}^n/\mathbb{Z}^n \simeq \mathrm{U}(1)^n$ is clearly compact (here $\mathrm{U}(1)$ is the circle group).

For the converse, assume Λ is discrete and V/Λ is compact. Let W be the subspace of V spanned by Λ; the $\mathbb{R}$-vector space V/W cannot have positive dimension, since it is contained in the compact space V/Λ, thus $W=\{0\}$ and Λ spans V. By picking an $\mathbb{R}$-basis for V in Λ we obtain an isomorphism $V \xrightarrow{\sim} \mathbb{R}^n$ that allows us to identify Λ with a subgroup of $\mathbb{R}^n$ containing $\mathbb{Z}^n$. We claim that the index $[\Lambda:\mathbb{Z}^n]$ must be finite.

Proof of claim: choose an integer $r\ge 1$ so that the ball of radius $\epsilon=\sqrt{n}/r$ about 0 intersects Λ only at 0; this is possible because Λ is discrete. We now subdivide the 1-cube in $\mathbb{R}^n$ into $\frac{1}{2r}$-cubes of which there are finitely many. If $[\Lambda:\mathbb{Z}^n]$ is infinite, then one of these $\frac{1}{2r}$-cubes contains at least two (in fact, infinitely many) distinct elements $v,w\in\Lambda$, which must be separated by a distance that is strictly less than ϵ. But then $0<\|v-w\|<\epsilon$, which contradicts our choice of ϵ.

The claim implies that Λ is a finitely generated $\mathbb{Z}$-module, hence a free $\mathbb{Z}$-module (it is torsion free and $\mathbb{Z}$ is a PID). It contains $\mathbb{Z}^n$ with finite index so its rank is n. $\square$

Remark 13.3. One might ask why we are using the archimedean completion $\mathbb{R}$ of $\mathbb{Q}$ rather than some nonarchimedean 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 (see Definition 12.11). Any basis $u_1,\ldots,u_n$ for V determines a parallelepiped

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

If we fix $u_1,\ldots,u_n$ as our basis for $V\simeq\mathbb{R}^n$, we then normalize the Haar measure μ so that it agrees with the standard normalization on $\mathbb{R}^n$ by defining $\mu(F(u_1,\ldots,u_n))=1$.

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

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

This is precisely the factor by which we rescale μ if we switch to the basis $e_1,\ldots,e_n$.

Remark 13.4. If $T\colon V\to V$ is a linear transformation on a real vector space $V\simeq\mathbb{R}^n$ with Haar measures μ, then for any measurable set S we have

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

This identity does not depend on a choice of basis; $\det T$ is the same regardless of which basis we use to compute it. It implies, in particular, that the absolute value of the determinant of any matrix in $\mathbb{R}^{n\times n}$ is equal to the volume of the parallelepiped spanned by its rows (or columns), a fact that we used above.

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 which we may identify with the parallelepiped $F(u_1,\ldots,u_n)\subset V$, which forms a set of unique coset representatives. More generally, we make the following definition.

Definition 13.5. 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 unique 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 13.6. *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. For $\lambda \in \Lambda$, the set $(F + \lambda) \cap G$ is the λ-translate of $F \cap (G - \lambda)$; these sets have the same measure since μ is translation-invariant. Partitioning F over translates of G yields

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

where we have used the countable additivity of μ and the fact that $\Lambda \simeq \mathbb{Z}^n$ is countable. $\square$

Definition 13.7. 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.

Remark 13.8. Note that volumes and covolumes depend on the normalization of the Haar measure μ, but ratios of them do not. In situations where we have a canonical way to choose an isomorphism $V \to \mathbb{R}^n$ (or $V \to \mathbb{C}^n$), such as when V is a number field (which is our main application), we normalize the Haar measure μ on V so that the inverse image of the unit cube in $\mathbb{R}^n$ has unit volume in V.

Proposition 13.9. *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. Let F be a fundamental domain for Λ and let L be a set of unique coset representatives for Λ/Λ'. Then L is finite (because Λ and Λ' are both cocompact) and

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

is a fundamental domain for Λ'. Thus

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

Definition 13.10. Let S be a subset of a real vector space. The set S is *symmetric* if it is closed under negation, and it is *convex* if for every pair of points $x, y \in S$ the line segment $\{tx + (1-t)y : t \in [0,1]\}$ between them is contained in S.

Lemma 13.11. *If $S \subseteq \mathbb{R}^n$ is a symmetric convex set of volume $\mu(S) > 2^n$ then S contains a nonzero element of $\mathbb{Z}^n$.*

Proof. See Problem Set 6. $\square$

Theorem 13.12 (Minkowski 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. $\square$

Example 13.13. As an application of the Minkowski lattice point theorem, let us prove FERMAT'S CHRISTMAS THEOREM: an odd prime p is a sum of two integer squares a^2+b^2 if and only if $p \equiv 1 \bmod 4$.[1] The "only if" direction is easy: a^2 and b^2 must be congruent to 0 or 1 mod 4, which implies that a^2+b^2 cannot be congruent to 3 mod 4.

To prove the "if" direction, let $p \equiv 1 \bmod 4$ be prime. The cyclic group $\mathbb{F}_p^\times$ has order $p-1$ divisible by 4, so it contains an element α of order 4 whose square must be -1, the unique element of order 2 in $\mathbb{F}_p^\times$. Let $i \in [1, p-1]$ be a lift of $\alpha \in \mathbb{F}_p \simeq \mathbb{Z}/p\mathbb{Z}$ to $\mathbb{Z}$ and define

$$\Lambda := \{(x,y) \in \mathbb{Z}^2 : y \equiv ix \bmod p\},$$

so that $x^2+y^2 \equiv (x+iy)(x-iy) \equiv 0 \bmod p$ for all $x, y \in \Lambda$. Then $\Lambda = (1,i)\mathbb{Z} + (0,p)\mathbb{Z}$ is a lattice in $\mathbb{R}^2$ with covolume

$$\operatorname{covol}(\Lambda) = \left|\det \begin{bmatrix} 1 & i \\ 0 & p \end{bmatrix}\right| = p.$$

The set

$$S := \{v \in \mathbb{R}^2 : \|v\| < \sqrt{2p}\},$$

is a symmetric convex set in $\mathbb{R}^2$ with measure $\mu(S) = 2\pi p > 4p = 2^2 \operatorname{covol}(\Lambda)$. By Corollary 13.12, S contains a nonzero $(a,b) \in \Lambda$. Then $a^2+b^2 \equiv 0 \bmod p$, since $(a,b) \in \Lambda$ and $0 < a^2+b^2 < 2p$, since (a,b) is a nonzero element of S; therefore $a^2+b^2 = p$.

13.2 The canonical inner product

Let $K/\mathbb{Q}$ be a number field with $K_\mathbb{R} := K \otimes_\mathbb{Q} \mathbb{R} \simeq \mathbb{R}^r \times \mathbb{C}^s \simeq \mathbb{R}^n$ and $K_\mathbb{C} := K \otimes_\mathbb{Q} \mathbb{C} \simeq \mathbb{C}^n$ and $r + 2s = n$. 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 an 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

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

[1] In a letter from Fermat to Mersenne dated December 25, 1640 (whence the name) Fermat claimed a proof of this theorem; as usual, he did not actually supply one, but in this case he almost certainly had one.

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 embeddings in $\operatorname{Hom}_{\mathbb{Q}}(K, \mathbb{C})$ are either real or 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 $K \hookrightarrow K_{\mathbb{R}}$). 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, $\bar{z}_\sigma = z_{\bar{\sigma}} \in \mathbb{R}$, while for pairs of conjugate complex embeddings $(\sigma, \bar{\sigma})$ we get the embedding $z \mapsto (z, \bar{z})$ 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 $w_i, y_j, z_i \in \mathbb{R}$, 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

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

and if we normalize the Haar measure μ on $K_{\mathbb{R}}$ consistently we will have

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

where $\mu_{\mathbb{R}_n}$ denotes the standard Lebesgue measure on $\mathbb{R}^n$. Having fixed a normalization of the Haar measure on $K_{\mathbb{R}}$, we can compute the covolume of the lattice $\mathcal{O}_K$ in $K_{\mathbb{R}}$.

13.3 Covolumes of ideals

Proposition 13.14. *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 $\mathcal{O}_K$, and let $\operatorname{Hom}_{\mathbb{Q}}(K, \mathbb{C}) = \{\sigma_1, \ldots, \sigma_n\}$. Let $A := [\sigma_i(e_j)]_{ij} \in \mathbb{C}^{n \times n}$. Viewing $\mathcal{O}_K \hookrightarrow K_{\mathbb{R}}$ as a lattice in $K_{\mathbb{R}}$ with basis $e_1, \ldots, e_n$, using

(1) to compute $\operatorname{covol}(\mathcal{O}_K)^2 = \mu(F(e_1,\ldots,e_n))^2$ yields

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

and by Proposition 11.13, $|\operatorname{disc}\mathcal{O}_K| = |\det A|^2 = \operatorname{covol}(\mathcal{O}_K)^2$. $\square$

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

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

as having image in $\mathbb{Q}_{>0}$ by identifying $N(I) = (x) \in \mathcal{I}_{\mathbb{Z}}$ with $|x| \in \mathbb{Q}_{>0}$. For ideals $I \subseteq \mathcal{O}_K$ this is just the positive integer $[\mathcal{O}_K : I]$; by definition, the norm $N(I)$ is the module index $(\mathcal{O}_K : I)_{\mathbb{Z}}$, and for $I \subseteq \mathcal{O}_K$ this is simply the $\mathbb{Z}$-ideal generated by $[\mathcal{O}_K : I]$.

Corollary 13.15. *Let K be a number field and let I be a nonzero fractional ideal of $\mathcal{O}_K$. Then*

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

Proof. Let $n = [K:\mathbb{Q}]$. Since $\operatorname{covol}(bI) = b^n\operatorname{covol}(I)$ and $N(bI) = b^nN(I)$ 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 13.9 and 13.14, we have

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

as claimed. $\square$

13.4 The Minkowski bound

Theorem 13.16 (Minkowski bound)**.** *Let K be a number field of degree $n = r + 2s$ with s complex embeddings. 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

$$|N_{K/\mathbb{Q}}(a)| \leq m_K N(I).$$

Before proving the theorem we first prove a lemma.

Lemma 13.17. *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

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

It follows that

$$\mu(S_t) = 2^s \mu_{\mathbb{R}^n}(V) \tag{7}$$

where $V \subseteq \mathbb{R}^n$ is the region defined by (6) and $\mu_{\mathbb{R}^n}$ is the standard Lebesgue measure on $\mathbb{R}^n$. We now show that the volume of V is a scalar multiple of the volume of the set

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

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 $d \in [0, t]$ determined by (6). 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 all but w_r fixed and replace w_r ranging over $[-d, d]$ 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

$$\mu_{\mathbb{R}^n}(V) = 2^{r-s}\pi^s \mu_{\mathbb{R}^n}(U).$$

Plugging this into (7) and applying $\mu_{\mathbb{R}^n}(U) = t^n/n!$ yields

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

as desired. This completes the proof of the lemma. $\square$

Proof of Theorem 13.16. Let I be a nonzero fractional ideal of $\mathcal{O}_K$. By Minkowski's Lattice Point Theorem (Corollary 13.12) and Corollary 13.15, if we choose t so that

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

then S_t will contain a nonzero element $a \in I$ which must satisfy

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

where σ ranges over the n elements of $\operatorname{Hom}_\mathbb{Q}(K, \mathbb{C})$.

By Lemma 13.17, we want to choose t so that

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

equivalently,

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

Let us now pick t so that $\left(\frac{t}{n}\right)^n > m_K N(I)$. Recalling that the geometric mean is bounded above by the arithmetic mean, we have

$$\sqrt[n]{|N_{K/\mathbb{Q}}(a)|} = \sqrt[n]{\prod |\sigma(a)|} \le \frac{1}{n}\sum |\sigma(a)| < \frac{t}{n},$$

Thus $|N_{K/\mathbb{Q}}(a)| < \left(\frac{t}{n}\right)^n$. If we now take the limit as $\left(\frac{t}{n}\right)^n \to m_K N(I)$ from above, we obtain $|N_{K/\mathbb{Q}}(a)| \le m_K N(I)$ as desired. $\square$

13.5 Finiteness of the ideal class group

Recall that the ideal class group $\operatorname{Pic}\mathcal{O}_K = \operatorname{cl}\mathcal{O}_K = \mathcal{I}_K/\mathcal{P}_K$ is the quotient of the ideal group $\mathcal{I}_K$ of $\mathcal{O}_K$ by its subgroup of principal fractional ideals $\mathcal{P}_K$.

We now use the Minkowski bound to prove that every ideal class contains a representative ideal of small norm. It will then follow that the ideal class group is finite.

Theorem 13.18. *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 $N(I) \le m_K$, where m_K is the Minkowski constant.*

Proof. Let $[J]$ be an ideal class of $\mathcal{O}_K$ represented by the nonzero fractional ideal J. By Theorem 13.16, the ideal J^{-1} contains a nonzero element a for which

$$|N_{K/\mathbb{Q}}(a)| \le m_K N(J^{-1}) = m_K/N(J),$$

and therefore $N(aJ) = |N_{K/\mathbb{Q}}(a)|N(J) \le m_K$. We have $a \in J^{-1}$, thus $aJ \subseteq J^{-1}J = \mathcal{O}_K$ and aJ is an $\mathcal{O}_K$-ideal as desired. $\square$

Lemma 13.19. *Let K be a number field and let M be a real number. The set of ideals $I \subseteq \mathcal{O}_K$ with $N(I) \le M$ 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 $m \le M$, since

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

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

Proof 2. Let I be an ideal of absolute norm $N(I) \le M$ and let $I = \mathfrak{p}_1 \cdots \mathfrak{p}_k$ be its factorization into (not necessarily distinct) prime ideals. Then $M \ge N(I) = N(\mathfrak{p}_1)\cdots N(\mathfrak{p}_k) \ge 2^k$, since the norm of each $\mathfrak{p}_i$ is a prime power, and in particular at least 2. It follows that $k \le \log_2 M$ is bounded, independent of I. Each prime ideal $\mathfrak{p}$ lies above some prime $p \le M$, of which there are $\pi(M) \approx M/\log M$ (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 $(n\pi(M))^{\log_2 M}$ ideals of norm at most M, a finite number. $\square$

Theorem 13.20. *Let K be a number field. The ideal class group of $\mathcal{O}_K$ is finite.*

Proof. By Theorem 13.18, each ideal class is represented by an ideal of norm at most m_K, and clearly distinct ideal classes must be represented by distinct ideals. By Lemma 13.19, the number of such ideals is finite. $\square$

Remark 13.21. 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* [2] and proved by Heilbronn [4] in 1934; the first fully explicit lower bound was obtained by Oesterlé in 1988 [5].

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 [6] and Heegner [3].[2] Complete lists of imaginary quadratic fields for each class number $h_K \le 100$ are now available [7].

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.[3]

Corollary 13.22. *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, thus Theorem 13.18 implies $m_K \ge 1$. Therefore

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

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

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

for all $n \ge 1$, by 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.

	$n=2$	$n=3$	$n=4$	$n=5$	$n=6$	$n=7$
lower bound from Corollary 13.22	3	11	46	210	1014	5014
minimum value of $\lvert\operatorname{disc}\mathcal{O}_K\rvert$	3	23	275	4511	92799	2306599

Corollary 13.23. *If K is a number field other than $\mathbb{Q}$ then $|\operatorname{disc}\mathcal{O}_K| > 1$. In particular, there is no non-trivial unramified extension of $\mathbb{Q}$.*

Proposition 13.24. *For $M \in \mathbb{R}_{>0}$ 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

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

[2] Heegner's 1952 result [3] was essentially correct but contained some gaps that prevented it from being generally accepted until 1967 when Stark gave a complete proof in [6].

[3] 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 [1].

Then

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

and by the Minkowski lattice point theorem (Corollary 13.12), 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

$$|N_{K/\mathbb{Q}}(a)| = \left|\prod_{i=1}\sigma_i(a)\right| \in \mathbb{Z}_{>0},$$

which must be at least 1, and $|a_2|, \ldots, |a_n| < 1$ so $|a_1| > 1 > |a_i|$ for $i = 2, \ldots, n$.

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 $\operatorname{Hom}_{\mathbb{Q}}(K, \mathbb{C})$ that fix $\mathbb{Q}(a)$, namely, those lying in the kernel of the map $\operatorname{Hom}_{\mathbb{Q}}(K, \mathbb{C}) \to \operatorname{Hom}_{\mathbb{Q}}(\mathbb{Q}(a), \mathbb{C})$ induced by restriction. But this is impossible since $|a_1| > |a_i|$ for $i \neq 1$.

Now $a \in \mathcal{O}_K$, so its minimal polynomial is a monic irreducible polynomial $f \in \mathbb{Z}[x]$ of degree n. The roots of $f(x)$ correspond to the $a_i = \sigma_i(a) \in \mathbb{R}$ which are all bounded in absolute value; and the coefficients of $f(x)$ are the elementary symmetric functions of the roots, hence also bounded in absolute value. The coefficients of f 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, where $n = r + 2s$ and $K_{\mathbb{R}} \simeq \mathbb{R}^r \times \mathbb{C}^s$. Now let

$$S := \{(w_1, \ldots, w_r,\ x_1{+}iy_1,\ \ldots,\ x_s{+}iy_s) \in K_{\mathbb{R}} : |x_1| < c\sqrt{M} \text{ and } |w_i|, |x_j|, |y_k| < 1\ (j \neq 1)\}$$

with c chosen so that $\mu(S) > 2^n\operatorname{covol}(\mathcal{O}_K)$ (the exact value of c depends on n but clearly this can be done). 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. $\square$

Lemma 13.25. *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 12.8 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 $v_p(e_v)$ cannot exceed $\log_p(n)$, so

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

as claimed. $\square$

Remark 13.26. The bound in Lemma 13.25 is tight. It is achieved by $K = \mathbb{Q}[x]/(x^{p^e} - p)$, for example.

Theorem 13.27 (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 $p \in S$, so $|\operatorname{disc}\mathcal{O}_K|$ is bounded. The theorem then follows from Proposition 13.24. $\square$

References

[1] 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.

[2] 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.

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

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

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

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

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

18.785 Number Theory I
Fall 2015