Number Fields IID, Lent 2020*

Comments/corrections to `a.j.scholl@dpmms.cam.ac.uk`

Lecture 1

1 Algebraic numbers and integers. Number fields.

$\mathbb{Q}$ = rational field, F = any field containing $\mathbb{Q}$ (e.g. $F = \mathbb{C}$).

Recall: $\alpha \in F$ is *algebraic* (over $\mathbb{Q}$) if there exists nonzero $f \in \mathbb{Q}[T]$ with $f(\alpha) = 0$.

If so, there exists a unique monic polynomial $m_\alpha \in \mathbb{Q}[T]$ of minimal degree with $m_\alpha(\alpha) = 0$, called the *minimal polynomial* of α. The *degree* of α is the degree of m_α.

Proposition 1.1. *If $\alpha \in F$ is algebraic, then m_α is irreducible, and if $f \in \mathbb{Q}[T]$ then $f(\alpha) = 0 \iff m_\alpha | f$.*

Proof. (Proved in GRM hopefully) If $m_\alpha = fg$ in $\mathbb{Q}[T]$ then $f(\alpha)g(\alpha) = 0$, hence (F is a field!) $f(\alpha) = 0$ or $g(\alpha) = 0$, hence one of f, g is constant (by minimality of $\deg m_\alpha$).

If $f(\alpha) = 0$ then writing $f = gm_\alpha + h$, g, $h \in \mathbb{Q}[T]$, $\deg h < \deg m_\alpha$ we have $h(\alpha) = f(\alpha) - g(\alpha)m_\alpha(\alpha) = 0$, so (minimality again) $h = 0$ and $m_\alpha | f$. (Other direction is obvious) □

If $\alpha \in F$, then $\mathbb{Q}(\alpha)$ denotes the smallest subfield of F containing ($\mathbb{Q}$ and) α. So

$$\mathbb{Q}(\alpha) = \left\{ \frac{f(\alpha)}{g(\alpha)} \;\middle|\; f,\ g \in \mathbb{Q}[T],\ g(\alpha) \neq 0 \right\}.$$

Proposition 1.2. *If α is algebraic of degree n, then $1, \alpha, \dots, \alpha^{n-1}$ is a $\mathbb{Q}$-basis for $\mathbb{Q}(\alpha)$. Conversely, if $[\mathbb{Q}(\alpha) : \mathbb{Q}] = n < \infty$ then α is algebraic of degree n.*

Proof. Consider the homomorphism $\phi\colon \mathbb{Q}[T] \to F$, $\phi(f) = f(\alpha)$. Its kernel is the ideal (m_α) which is maximal, so its image is a subfield of F, so equals $\mathbb{Q}(\alpha)$. Now $1, T, \dots, T^{n-1}$ is obviously a basis for $\mathbb{Q}[T]/(m_\alpha)$. Hence the first part. In

*Parts in red were not done in lectures, but included here for completeness.

the other direction, $1, \alpha, \dots, \alpha^n \in \mathbb{Q}(\alpha)$ are linearly dependent over $\mathbb{Q}$, hence α is algebraic. ☐

Proposition 1.3. *$\{\alpha \in F \mid \alpha$ is algebraic$\}$ is a subfield of F.*

Proof. See Galois theory. Alternatively: enough to prove that the set is closed under $+$ and $\times$ — see 1.6 below for a stronger statement — and that if $\alpha \neq 0$ is algebraic then so is $1/\alpha$ — which follows from Proposition 1.2 or simply by

$$\sum_{j=0}^{n} b_j \alpha^j = 0 \implies \sum_{j=0}^{n} b_{n-j} (1/\alpha)^j = 0. \tag{1.1}$$

☐

Key Definition 1.4. $x \in F$ is an *algebraic integer* if there exists a monic $f \in \mathbb{Z}[T]$ with $f(\alpha) = 0$.

Lemma 1.5. *(1) Let $\alpha \in F$. TFAE:*

- *(i) α is an algebraic integer.*
- *(ii) α is algebraic and $m_\alpha \in \mathbb{Z}[T]$.*
- *(iii) $\mathbb{Z}[\alpha]$ is a finitely-generated $\mathbb{Z}$-module.*

If these hold, then $1, \alpha, \dots, \alpha^{d-1}$ is a $\mathbb{Z}$-basis for $\mathbb{Z}[\alpha]$, where $d = \deg \alpha$.

(2) $\alpha \in \mathbb{Q}$ is an algebraic integer iff $\alpha \in \mathbb{Z}$.

Notation for (iii): let $\alpha_1, \dots, \alpha_n \in F$. Define $\mathbb{Z}[\alpha_1, \dots, \alpha_n]$ to be the smallest subring of F containing $\{\alpha_i\}$, which is the set of all finite sums

$$\sum_{i_1, \dots, i_n \geq 0} c_{\underline{i}} \alpha_1^{i_1} \cdots \alpha_n^{i_n}, \qquad c_{\underline{i}} \in \mathbb{Z}.$$

So $\mathbb{Z}[\alpha] = \{g(\alpha) \mid g \in \mathbb{Z}[T]\}$.

Proof. (1) (i) $\implies$ (ii): suppose $f(\alpha) = 0$, $f \in \mathbb{Z}[T]$ monic. Then m_α divides f in $\mathbb{Q}[T]$, so $f = g m_\alpha$ with $g \in \mathbb{Q}[T]$ monic. By Gauss's Lemma, m_α (and also g) are in $\mathbb{Z}[T]$.

(ii) $\implies$ (iii): write $m_\alpha = T^d + \sum_{j=0}^{d} a_j T^j$, $a_j \in \mathbb{Z}$. Then $\alpha^d = -\sum_{j=0}^{d-1} a_j \alpha^j$. So for any $n \geq 0$, α^n is a $\mathbb{Z}$-linear combination of $1, \alpha, \dots, \alpha^{d-1}$. Therefore $\mathbb{Z}[\alpha]$ is generated by $1, \alpha, \dots, \alpha^{d-1}$ (and freely generated, since $\deg \alpha = d$).

(iii) $\implies$ (i): Let $\mathbb{Z}[\alpha]$ be generated by $g_1(\alpha), \dots, g_r(\alpha)$, $g_i \in \mathbb{Z}[T]$, and let $k = \max\{\deg g_i\}$. Then $\mathbb{Z}[\alpha]$ is generated by $1, \alpha, \dots, \alpha^k$, hence $\alpha^{k+1} = \sum_{j=0}^{k} b_j \alpha^j$ with $b_j \in \mathbb{Z}$, hence α is an algebraic integer.

(2) follows from (1)(ii). ☐

Theorem 1.6. *If α, $\beta \in F$ are algebraic integers, then so are $\alpha \pm \beta$ and $\alpha\beta$.*

Proof. The $\mathbb{Z}$-module $\mathbb{Z}[\alpha, \beta]$ is generated by $\{\alpha^i \beta^j \mid 0 \leq i < \deg \alpha,\ 0 \leq j < \deg \beta\}$. (This need not be a $\mathbb{Z}$-basis!) Then the submodule $\mathbb{Z}[\alpha\beta] \subset \mathbb{Z}[\alpha, \beta]$ is finitely generated, so by Lemma 1.5, $\alpha\beta$ is an algebraic integer. The same for $\alpha \pm \beta$. □

Proposition 1.7. *Let $\alpha \in F$ be algebraic. Then for some integer $b \geq 1$, $b\alpha$ is an algebraic integer.*

Proof. Exercise: check that for suitable b the minimal polynomial of $b\alpha$ has integer coefficients.

OR

Let $m_\alpha = T^d + \sum_{j=0}^{d-1} a_j T^j$, $a_j \in \mathbb{Q}$. Then $m_{b\alpha} = T^d + \sum_{j=0}^{d-1} b^{d-j} a_j T^j$ which for suitable $b \geq 1$ has integer coefficients. □

Now for the things this course is all about:

Definition. An *algebraic number field* (or simply *number field*) is a field K containing $\mathbb{Q}$ which is a finite extension; i.e. K is finite-dimensional as a $\mathbb{Q}$-vector space. The *degree* $[K : \mathbb{Q}]$ of K is the dimension $\dim_{\mathbb{Q}} K$.

The *ring of integers* $\mathfrak{o}_K$ of K is the set of algebraic integers in K (it is a subring of K by Theorem 1.6).

By Proposition 1.2, if α is algebraic then $\mathbb{Q}(\alpha)$ is a number field. The converse holds:

Theorem 1.8 (Theorem of the primitive element)**.** *If K is a number field, then $K = \mathbb{Q}(\alpha)$ for some $\alpha \in K$.*

See Galois theory for the proof. We will use this occasionally as it simplifies some proofs (though it's not essential).

——————— *Lecture 2* ———————

2 Example: quadratic fields

K is quadratic if $[K : \mathbb{Q}] = 2$. Suppose so, and let $\alpha \in K \setminus \mathbb{Q}$. Then m_α is quadratic, and solving we get $\alpha = x + \sqrt{y}$ (where $\sqrt{y}$ here simply denotes an element of K whose square is y), with x, $y \in \mathbb{Q}$ and y not a rational square. We can uniquely write $y = z^2 d$ where $z \in \mathbb{Q}$ and $d \in \mathbb{Z} \setminus \{0, 1\}$ is squarefree. Therefore $K = \mathbb{Q}(\sqrt{d})$, and is isomorphic to the quotient $\mathbb{Q}[T]/(T^2 - d)$ by GRM. If $d \neq d' \in \mathbb{Z} \setminus \{0, 1\}$ then $\mathbb{Q}(\sqrt{d})$ and $\mathbb{Q}(\sqrt{d'})$ are not isomorphic (exercise).

Let's compute $\mathfrak{o}_K$ for $K = \mathbb{Q}(\sqrt{d})$. Let $\alpha = u + v\sqrt{d} \in K$, $u, v \in \mathbb{Q}$. Then if $v = 0$, $\alpha \in \mathfrak{o}_K \iff \alpha \in \mathbb{Z}$ (Lemma 1.5(2)). Otherwise, $\alpha \notin \mathbb{Q}$ and $m_\alpha = T^2 - 2uT + u^2 - dv^2$. So $\alpha \in \mathfrak{o}_K$ iff $2u \in \mathbb{Z}$ and $u^2 - dv^2 \in \mathbb{Z}$.

If $u \in \mathbb{Z}$, then we require $dv^2 \in \mathbb{Z}$, and since d is squarefree, this holds iff $v \in \mathbb{Z}$.

Otherwise $u = (2a+1)/2$ for $a \in \mathbb{Z}$. Then $u^2 - dv^2 \in \mathbb{Z}$ iff $4dv^2 - (2a+1)^2 \in 4\mathbb{Z}$ iff (since d is squarefree) $v = k/2$ with $k \in \mathbb{Z}$ and $dk^2 \equiv 1 \pmod 4$. If $d \equiv 1 \pmod 4$ this congruence holds iff $k = 2b + 1$ is odd. Otherwise it holds for no integer k. Summing up, we have shown:

Theorem 2.1. *Suppose that $d \in \mathbb{Z} \setminus \{0, 1\}$ is squarefree, $K = \mathbb{Q}(\sqrt{d})$. Then:*
(i) If $d \not\equiv 1 \pmod 4$ then $\mathfrak{o}_K = \{u + v\sqrt{d} \mid u,\ v \in \mathbb{Z}\} = \mathbb{Z}[\sqrt{d}]$.
(ii) If $d \equiv 1 \pmod 4$ then

$$\mathfrak{o}_K = \{u + v\sqrt{d} \mid u,\ v \in \frac{1}{2}\mathbb{Z},\ u - v \in \mathbb{Z}\} = \mathbb{Z}\left[\frac{1+\sqrt{d}}{2}\right].$$

(It is an easy exercise to check the 2nd equality in (ii).)

Remarks. (1) For a general number field K it need not be the case that there exists $\alpha \in \mathfrak{o}_K$ such that $\mathfrak{o}_K = \mathbb{Z}[\alpha]$.

(2) For more complicated fields it will be impractical to try to compute $\mathfrak{o}_K$ in the naive way just used. We need some more tools.

3 Embeddings

K a number field, $[K : \mathbb{Q}] = n$.

Theorem 3.1. *There are exactly n field homomorphisms $\sigma_i \colon K \hookrightarrow \mathbb{C}$ $(1 \le i \le n)$, called the* complex embeddings *of K. More generally, if $\mathbb{Q} \subset F \subset K$ are number fields, then each of the $[F : \mathbb{Q}]$ complex embeddings of F has exactly $[K : F]$ extensions to K.*

Proof. (Galois theory) Assume $K = \mathbb{Q}(\theta) = \mathbb{Q}[T]/(m_\theta)$. Then to give a homomorphism $\sigma \colon K \hookrightarrow \mathbb{C}$ is the same as to give a ring homomorphism $\phi \colon \mathbb{Q}[T] \to \mathbb{C}$ satisfying $\phi(m_\theta) = 0$. If $z = \phi(T)$ then $\phi(m_\theta) = m_\theta(z)$, so the map $\phi \mapsto z$ gives a bijection between the set of such homomorphisms σ and the set of roots $z \in \mathbb{C}$ of m_θ, which has n elements. The second part is proved the same way, since θ has degree $[K : F]$ over F. □

Remarks. (i) If $K \subset \mathbb{C}$ then we can require, if we like, that σ_1 is the inclusion map.

(ii) Suppose that exactly r of the embeddings σ_i are real, i.e. $\sigma_i(K) \subset \mathbb{R}$. Then the remaining s_i fall into s complex conjugate pairs $(\sigma_i, \overline{\sigma_i})$, with $n = r + 2s$. (Some people, including me, prefer to use (r_1, r_2) instead of (r, s) but the Tripos has a history of using (r, s), so I'll stick with that.)

Proposition 3.2. *If $\alpha \in K$ then the complex numbers $\sigma_i(\alpha)$ are the complex roots of m_α, each taken $n/\deg\alpha$ times. (The* conjugates *of α.)*

Proof. Follows from second statement of Theorem 3.1 $\square$

4 Norm and trace

Let $\alpha \in K$. Define the map $u_\alpha \colon K \to K$ by $u_\alpha(x) = \alpha x$. It is a $\mathbb{Q}$-linear transformation of K. Define

$$f_\alpha = \text{characteristic polynomial of } u_\alpha = \det(T - u_\alpha) \in \mathbb{Q}[T]$$
$$\mathrm{N}_{K/\mathbb{Q}}(\alpha) = \det u_\alpha \text{ (norm)}, \qquad \mathrm{Tr}_{K/\mathbb{Q}}(\alpha) = \mathrm{tr}\, u_\alpha \text{ (trace)}.$$

Explicitly, let $(\beta_i)_{1\le i\le n}$ be a $\mathbb{Q}$-basis for K. Then we can write uniquely

$$\alpha\beta_i = \sum_{j=1}^{n} A_{ji}\beta_j, \qquad A \in \mathrm{Mat}_{n,n}(\mathbb{Q})$$

and then $f_\alpha = \det(T.I_n - A)$, and so on.

Proposition 4.1. $\mathrm{N}_{K/\mathbb{Q}}(\alpha\beta) = \mathrm{N}_{K/\mathbb{Q}}(\alpha)\mathrm{N}_{K/\mathbb{Q}}(\beta)$ *and* $\mathrm{Tr}_{K/\mathbb{Q}}(\alpha+\beta) = \mathrm{Tr}_{K/\mathbb{Q}}(\alpha) + \mathrm{Tr}_{K/\mathbb{Q}}(\beta)$.

Proof. From the definition, $u_{\alpha\beta} = u_\alpha u_\beta$ and $u_{\alpha+\beta} = u_\alpha + u_\beta$, so result follows from the same properties for determinant/trace of linear transformations. $\square$

Theorem 4.2. *(i)* $f_\alpha = \prod_{i=1}^{n}(T - \sigma_i(\alpha)) = m_\alpha^{n/d}$, $d = \deg(\alpha)$.
(ii) $f_\alpha = \prod_{i=1}^{n}(T - \sigma_i(\alpha))$, $\mathrm{N}_{K/\mathbb{Q}}(\alpha) = \prod_i \sigma_i(\alpha)$ *and* $\mathrm{Tr}_{K/\mathbb{Q}}(\alpha) = \sum_i \sigma_i(a)$.

——— *Lecture 3* ———

Proof. (i) Suppose first that $K = \mathbb{Q}(\alpha)$, so $\deg(\alpha) = n = \deg m_\alpha$. Then for every $\beta \in K$, $f_\alpha(\alpha)\beta = f_\alpha(u_\alpha)\beta = 0$ by Cayley-Hamilton, and $f_\alpha \in \mathbb{Q}[T]$, hence $m_\alpha | f_\alpha$ and therefore $f_\alpha = m_\alpha = \prod(T - \sigma_i(\alpha))$ by Proposition 3.2.

In the general case, consider $\mathbb{Q}(\alpha) \subset K$. By what we just proved, the characteristic polynomial of u_α on $\mathbb{Q}(\alpha)$ equals m_α, and its roots are the distinct conjugates of α, by 3.2. Then we have $K \simeq \mathbb{Q}(\alpha)^{n/d}$ as $\mathbb{Q}(\alpha)$-vector spaces, so

$$f_\alpha = (\text{characteristic polynomial of } u_\alpha \text{ on } \mathbb{Q}(\alpha))^{n/d} = m_\alpha^{n/d} = \prod_{i=1}^{n}(T - \sigma_i(\alpha)).$$

(ii) Follows from (i), since $\det u_\alpha = (-1)^n f_\alpha(0)$ and $\mathrm{tr}\, u_\alpha$ equals minus the coefficient of T^{n-1} in f_α.

$\square$

Remark. Some people take (ii) as the definition of norm and trace.

Corollary 4.3. *(i) Let $\alpha \in K$. Then $\alpha = 0 \iff \mathrm{N}_{K/\mathbb{Q}}(\alpha) = 0$.*

(ii) Let $\alpha \in \mathfrak{o}_K$. Then $f_\alpha \in \mathbb{Z}[T]$ and $\mathrm{N}_{K/\mathbb{Q}}(\alpha)$, $\mathrm{Tr}_{K/\mathbb{Q}}(\alpha) \in \mathbb{Z}$. Moreover, $\alpha \in \mathfrak{o}_K^\times \iff \mathrm{N}_{K/\mathbb{Q}}(\alpha) \in \{\pm 1\}$.

Proof. (i) $\alpha = 0$ iff every $\sigma_i(\alpha) = 0$.

(ii) $f_\alpha = m_\alpha^{n/d} \in \mathbb{Z}[T]$ as $\alpha \in \mathfrak{o}_K$, and $\mathrm{N}_{K/\mathbb{Q}}(\alpha)$, $\mathrm{Tr}_{K/\mathbb{Q}}(\alpha)$ are (up to sign) the coefficients of 1 and T^{n-1} in f_α. For the last part, if α, $\beta \in \mathfrak{o}_K$, then $\mathrm{N}_{K/\mathbb{Q}}(\alpha)\mathrm{N}_{K/\mathbb{Q}}(\beta) = \mathrm{N}_{K/\mathbb{Q}}(\alpha\beta)$, so if $\alpha \in \mathfrak{o}_K^\times$, $\mathrm{N}_{K/\mathbb{Q}}(\alpha)\mathrm{N}_{K/\mathbb{Q}}(\alpha^{-1}) = 1$ and $\mathrm{N}_{K/\mathbb{Q}}(\alpha) \in \{\pm 1\}$. Conversely, if $\mathrm{N}_{K/\mathbb{Q}}(\alpha) \in \{\pm 1\}$ then $f_\alpha = T^n + \sum_{j=1}^{n-1} b_j T^j$, with $b_j \in \mathbb{Z}$, $b_0 = \pm 1$, hence

$$\alpha^{-1} = b_0\big(\alpha^{n-1} + \sum_{j=1}^{n-1} b_j \alpha^{j-1}\big) \in \mathfrak{o}_K. \qquad \square$$

5 Some (GR)M

Proposition 5.1. *Let G be a finitely-generated abelian group with no torsion of rank n, generators $x_1, \ldots, x_n$. Let $H \subset G$ be the subgroup generated by some $y_1, \ldots, y_n$, where $y_i = \sum_j A_{ji} x_j$, $A \in \mathrm{Mat}_{n,n}(\mathbb{Z})$. Then if $\det A \neq 0$, H is a subgroup of finite index and $(G : H) = |\det A|$.*

Proof. Smith normal form: $A = PDQ$ with P, Q, $D \in \mathrm{Mat}_{n,n}(\mathbb{Z})$, $\det P$, $\det Q \in \{\pm 1\}$ and $D = \mathrm{diag}(d_1, \ldots, d_n)$ diagonal with $d_i > 0$. Then $G/H \simeq \mathbb{Z}/d_1\mathbb{Z} \times \cdots \times \mathbb{Z}/d_n\mathbb{Z}$ so $(G : H) = \prod d_i = |\det A|$. $\square$

Let V be a $\mathbb{Q}$-vector space of finite dimension n, $H \subset V$ a subgroup ($\mathbb{Z}$-submodule). Define

$$\mathrm{rank}(H) = \dim \mathrm{span}(H) \in \{0, 1, \cdots, n\}.$$

Proposition 5.2. *Let $H \subset V$ be a finitely-generated subgroup of rank r. Then $H = \bigoplus_{i=1}^r \mathbb{Z}x_i$ with (x_i) linearly independent in V.*

Proof. H has no torsion, so by classification, H is free and $H = \bigoplus_{i=1}^d \mathbb{Z}x_i$ for some d and $x_i \in V$. If x_i are linearly dependent, then there exist $m_i \in \mathbb{Q}$, not all 0, with $\sum m_i x_i = 0$. Clearing denominators, may assume $m_i \in \mathbb{Z}$. This contradicts freeness of H, so (x_i) are linearly independent, hence $d = r$. $\square$

6 Discriminants. Integral bases

Let $\alpha_1, \ldots, \alpha_n \in K$. Define

$$\mathrm{Disc}(\alpha_i) = \mathrm{Disc}(\alpha_1, \ldots, \alpha_n) = \det(\mathrm{Tr}_{K/\mathbb{Q}}\alpha_i\alpha_j) \in \mathbb{Q}.$$

Theorem 6.1. *(i)* $\mathrm{Disc}(\alpha_i) = \det(\sigma_i(\alpha_j))^2$.

(ii) $\mathrm{Disc}(\alpha_i) \neq 0 \iff (\alpha_i)$ *is a* $\mathbb{Q}$*-basis for* K.

(iii) If $\beta_i = \sum_{j=1}^{n} A_{ji}\alpha_j$ $(1 \leq i \leq n)$ *with* $A \in \mathrm{Mat}_{n,n}(\mathbb{Q})$*, then*

$$\mathrm{Disc}(\beta_i) = (\det A)^2 \,\mathrm{Disc}(\alpha_i).$$

(iv) Suppose (α_i) *is a* $\mathbb{Q}$*-basis for* K*. Then* $\mathrm{Disc}(\alpha_i)$ *depends only on the subgroup* $\mathbb{Z}\alpha_1 + \cdots + \mathbb{Z}\alpha_n$ *of* K *generated by* $\{\alpha_i\}$.

Proof. (i) Let $\Delta = (\sigma_i(\alpha_j))$. Then

$$({}^t\Delta\,\Delta)_{ij} = \sum_{k=1}^{n} \sigma_k(\alpha_i)\sigma_k(\alpha_j) = \mathrm{Tr}_{K/\mathbb{Q}}\alpha_i\alpha_j$$

hence $\mathrm{Disc}(\alpha_i) = \det({}^t\Delta\,\Delta) = (\det\Delta)^2$.

(ii) If $\sum b_j\alpha_j = 0$, $b_j \in \mathbb{Q}$, not all zero, then $\sum b_j\sigma_i(\alpha_j) = 0$ for all i, so $\det(\sigma_i(\alpha_j)) = 0$.

——————— *Lecture 4* ———————

Conversely, suppose that (α_j) is a basis. Let $T = (\mathrm{Tr}_{K/\mathbb{Q}}\alpha_i\alpha_j)_{ij}$. STP that if $\underline{0} \neq \underline{b} \in \mathbb{Q}^n$ then $T\underline{b} \neq \underline{0}$, or equivalently there exists $\underline{c} \in \mathbb{Q}^n$ such that ${}^t\underline{c}T\underline{b} \neq 0$. But if $\beta = \sum b_j\alpha_j$, $\gamma = \sum c_j\alpha_j$ then ${}^t\underline{c}T\underline{b} = \mathrm{Tr}_{K/\mathbb{Q}}(\beta\gamma)$, so taking $\gamma = \beta^{-1}$ will do.

(iii) Let $\Delta' = (\sigma_i(\beta_j))$. Then

$$\Delta'_{ij} = \sum_k \sigma_i(A_{kj}\alpha_k) = \sum_k \sigma_i(\alpha_k)A_{kj} = (\Delta\,A)_{ij}.$$

So $\det\Delta' = \det A \det\Delta$, and result follows from (i).

(iv) Let (α_i), (β_i) be $\mathbb{Q}$-bases for K generating the same subgroup of K. Then for some $A \in GL_n(\mathbb{Z}) = \{A \in \mathrm{Mat}_{n,n}(\mathbb{Z}) \mid \det A = \pm 1\}$, $\beta_i = \sum_j A_{ji}\alpha_j$. So by (iii), $\mathrm{Disc}(\beta_i) = (\det A)^2\,\mathrm{Disc}(\alpha_i) = \mathrm{Disc}(\alpha_i)$. □

If $H \subset K$ is a finitely generated subgroup of rank n, and $(x_1, \ldots, x_n)$ is a $\mathbb{Z}$-basis for H, (ii) and (iv) imply that $\mathrm{Disc}(x_1, \ldots, x_n)$ is a nonzero integer which does not depend on the choice of basis. We write it as $\mathrm{Disc}\,H$.

Lemma 6.2. *If* $H \subset H' \subset K$ *are finitely generated subgroups of rank* n*, then* $\mathrm{Disc}\,H = (H' : H)^2\,\mathrm{Disc}\,H'$.

Proof. Let H, H' have $\mathbb{Z}$-bases (α_i), (α'_i). Then $a_i = \sum_j B_{ji}\alpha'_j$, $B \in \mathrm{Mat}_{n,n}(\mathbb{Z})$. Then Proposition 5.1 and (ii) above give $(H' : H)^2 = (\det B)^2 = \mathrm{Disc}\,H / \mathrm{Disc}\,H'$. □

Theorem 6.3. *There exist* $\omega_1, \dots, \omega_n \in \mathfrak{o}_K$ *such that* $\mathfrak{o}_K = \mathbb{Z}\omega_1 \oplus \cdots \oplus \mathbb{Z}\omega_n$. *(We say that* (ω_i) *is an* integral basis *for* K*.)*

Proof. There exists a $\mathbb{Q}$-basis (ω_i) for K with $\omega_i \in \mathfrak{o}_K$ (take any basis and apply Proposition 1.7), and for such a basis, $0 \neq \operatorname{Disc} H \in \mathbb{Z}$ where $H \subset K$ is the subgroup generated by $\{\omega_i\}$. Choose such a basis with $|\operatorname{Disc} H|$ minimal. Claim that $H = \mathfrak{o}_K$. Indeed, let $\alpha \in \mathfrak{o}_K$, $H' = H + \mathbb{Z}\alpha$. Then H' is a finitely generated subgroup of K of rank n, so by the previous lemma, $\operatorname{Disc} H = (H' : H)^2 \operatorname{Disc} H'$. By minimality of $|\operatorname{Disc} H|$ this implies that $H = H'$ and $\alpha \in H$. □

Definition. The *discriminant* of K is $d_K := \operatorname{Disc} \mathfrak{o}_K$ (which equals $\operatorname{Disc}(\omega_i)$ for any integral basis (ω_i)).

Suppose $K = \mathbb{Q}(\sqrt{d})$ is quadratic, with d a squarefree integer. If $d \not\equiv 1 \pmod 4$ then an integral basis is $\{1, \sqrt{d}\}$, and

$$\Delta = (\sigma_i(\alpha_j)) = \begin{pmatrix} 1 & \sqrt{d} \\ 1 & -\sqrt{d} \end{pmatrix}, \qquad d_K = (\det \Delta)^2 = 4d.$$

If $d \equiv 1 \pmod 4$ then an integral basis is $\{1, (1+\sqrt{d})/2\}$ and so

$$d_K = (\det \Delta)^2 = \begin{vmatrix} 1 & (1+\sqrt{d})/2 \\ 1 & (1-\sqrt{d})/2 \end{vmatrix}^2 = d.$$

For calculating discriminants, the following results can be useful.

Proposition 6.4. *Suppose that* $K = \mathbb{Q}(\theta)$, *and* $f = m_\theta$ *is the minimal polynomial of* θ. *Then*

$$\operatorname{Disc}(1, \theta, \dots, \theta^{n-1}) = \prod_{i<j} (\sigma_i(\theta) - \sigma_j(\theta))^2 = (-1)^{n(n-1)/2} \mathrm{N}_{K/\mathbb{Q}}(f'(\theta)).$$

Proof. Recall the *Vandermonde determinant*:

$$VDM(X_1, \dots, X_n) := \begin{vmatrix} X_1^{n-1} & \cdots & X_n^{n-1} \\ \vdots & & \vdots \\ X_1 & \cdots & X_n \\ 1 & \cdots & 1 \end{vmatrix} = \prod_{1 \le i < j \le n} (X_i - X_j).$$

Then $\operatorname{Disc}(1, \dots, \theta^{n-1}) = VDM(\sigma_1(\theta), \dots, \sigma_n(\theta))^2$ by Theorem 6.1(i), hence the first equality. For the second, see example sheet 1, Q7. □

Proposition 6.5. *Let* $\omega_1, \dots, \omega_n \in \mathfrak{o}_K$ *with* $\operatorname{Disc}(\omega_i)$ *a* squarefree *integer. Then* (ω_i) *is an integral basis for* K.

Proof. Let $H = \sum \mathbb{Z}\omega_i$. Then $\operatorname{Disc}(H) = (\mathfrak{o}_K : H)^2 \operatorname{Disc}(\mathfrak{o}_K)$ by Lemma 6.2, hence $H = \mathfrak{o}_K$. □

7 Ideals I

Example: $K = \mathbb{Q}(\sqrt{-5})$, $\mathfrak{o}_K = \mathbb{Z}[\sqrt{-5}]$. The norm is $\mathrm{N}_{K/\mathbb{Q}}(x+y\sqrt{-5}) = x^2+5y^2$. Consider the equation:

$$6 = 2 \times 3 = (1+\sqrt{-5})(1-\sqrt{-5}).$$

It is not hard to see that these are two distinct factorisations of 6 into irreducibles. Indeed, we have $\mathrm{N}_{K/\mathbb{Q}}(2) = 4$, $\mathrm{N}_{K/\mathbb{Q}}(3) = 9$ and $\mathrm{N}_{K/\mathbb{Q}}(1 \pm \sqrt{-5}) = 6$. If any of these factors were reducible, then there would exist $\alpha \in \mathfrak{o}_K$ with $\mathrm{N}_{K/\mathbb{Q}}(\alpha) = 2$ or 3, and the equations $x^2+5y^2 = 2, = 3$ have no integer solutions. So $\mathfrak{o}_K$ is not a UFD.

We will show that we can restore unique factorisation if we replace elements by ideals.

Recall that an *ideal* I in a ring R is a subgroup (under addition) $I \subset R$ such that $\alpha \in R$, $\beta \in I \implies \alpha\beta \in I$. (Equivalently, I is an R-submodule of R.)

Lecture 5

In $\mathfrak{o}_K$, every nonzero ideal has finite index. More precisely:

Proposition 7.1. *(i) Let $I \subset \mathfrak{o}_K$ be a nonzero ideal. Then $I = \bigoplus_{i=1}^n \mathbb{Z}\alpha_i$ for some linearly independent (α_i), and*

$$(\mathfrak{o}_K : I)^2 = \frac{\operatorname{Disc} I}{d_K}.$$

(ii) Let $0 \neq \alpha \in \mathfrak{o}_K$. Then $(\mathfrak{o}_K : \alpha\mathfrak{o}_K) = \left|\mathrm{N}_{K/\mathbb{Q}}(\alpha)\right|$.

Proof. (i) Since $\mathfrak{o}_K$ is finitely generated, so is I. Let $0 \neq \alpha \in I$, and (ω_i) and integral basis for K. Then $\alpha\omega_1, \dots, \alpha\omega_n$ are linearly independent elements of I. So I has rank n. The first statement then follows from Proposition 5.2, and the second from Lemma 6.2.

(ii) If $I = \alpha\mathfrak{o}_K$ then we may take $\alpha_i = \alpha\omega_i$, and then

$$\operatorname{Disc}(\alpha_i) = \det(\sigma_i(\alpha\omega_j))^2 = \prod_i \sigma_i(\alpha)^2 \det(\sigma_i(\omega_j))^2 = \mathrm{N}_{K/\mathbb{Q}}(\alpha)^2 \operatorname{Disc}(\omega_i).$$

So by (i), $(\mathfrak{o}_K : I)^2 = \mathrm{N}_{K/\mathbb{Q}}(\alpha)^2$. □

We define for a nonzero ideal $N(I) = (\mathfrak{o}_K : I) \in \mathbb{Z}_{>0}$, the *norm* of I.

Corollary 7.2. *(i) $I \neq \{0\} \implies I \cap \mathbb{Z} \neq \{0\}$.*

(ii) There are only finitely many ideals of $\mathfrak{o}_K$ of given norm.

Proof. (i) For all $x \in \mathfrak{o}_K/I$, $N(I)x = 0$ by Lagrange, hence $N(I) \in I$.

(ii) Let $N(I) = M$. Then $M\mathfrak{o}_K \subset I$¿ By the isomorphism theorems for rings, there is a bijection

$$\{\text{ideals of } \mathfrak{o}_K \text{ containing } M\mathfrak{o}_K\} \overset{\sim}{\leftrightarrow} \{\text{ideals of } \mathfrak{o}_K/M\mathfrak{o}_K\}$$

and the set on the right is finite (since $\mathfrak{o}_K/M\mathfrak{o}_K$ is finite). □

Recall that an ideal $P \subset \mathfrak{o}_K$ is *prime* if: $P \neq \mathfrak{o}_K$ and if $\alpha, \beta \in \mathfrak{o}_K$ and $\alpha\beta \in P$ then $\alpha \in P$ or $\beta \in P$. (Equivalently, $\mathfrak{o}_K/P$ is an integral domain.)

Lemma 7.3. *Let $P \subset \mathfrak{o}_K$ be a prime ideal.*

(i) Either $P = \{0\}$ or P is a maximal ideal.

(ii) If $P \neq \{0\}$ then $P \cap \mathbb{Z} = p\mathbb{Z}$ for a prime number p, and $N(P)$ is a power p^f of p, for some $1 \leq f \leq n$.

Proof. (i) If $P \neq \{0\}$ then $\mathfrak{o}_K/P$ is a finite integral domain, which is therefore a field. So P is maximal.

(ii) By (i), $P \neq \{0\}$ implies that $P \cap \mathbb{Z}$ is nonempty, say $m \in P$, $m \geq 1$. As P is prime, some prime factor of m, call it p, must also belong to P. Therefore $(p) \subset P \subsetneqq \mathfrak{o}_K$, and since $N((p)) = \left|\mathrm{N}_{K/\mathbb{Q}}(p)\right| = p^n$, we must have $N(P) = p^f$ with $1 \leq f \leq n$. □

Henceforth, unless we say otherwise, by "prime ideal" we shall always mean "nonzero prime ideal". This is traditional terminology although it would be more natural just to say "maximal ideal".

Arithmetic of ideals: define sum and product of ideals to be

$$I + J = \{\alpha + \beta \mid \alpha \in I,\ \beta \in J\}$$
$$IJ = \{\text{finite sums } \sum_i \alpha_i\beta_i \mid \alpha_i \in I,\ \beta_i \in J\}.$$

These are also ideals (trivial check).

Notation: if $\alpha_1, \ldots, \alpha_k \in \mathfrak{o}_K$, we write $(\alpha_1, \ldots, \alpha_k)$ for the ideal generated by $\{\alpha_i\}$. So for $\alpha \in \mathfrak{o}_K$, $(\alpha) = \alpha\mathfrak{o}_K$. In this notation is it easy to see that

$$(\alpha_1, \ldots, \alpha_k) + (\beta_1, \ldots, \beta_\ell) = (\alpha_1, \ldots, \alpha_k, \beta_1, \ldots, \beta_\ell)$$
$$(\alpha_1, \ldots, \alpha_k)(\beta_1, \ldots, \beta_\ell) = (\alpha_1\beta_1, \alpha_1\beta_2, \ldots, \alpha_2\beta_1, \ldots, \alpha_i\beta_j, \ldots, \alpha_k\beta_\ell)$$

8 Ideals II: unique factorisation

Example: Let $K = \mathbb{Q}(\sqrt{-5})$. We saw earlier that $\mathfrak{o}_K = \mathbb{Z}[\sqrt{-5}]$ is not a UFD, as $6 = 2 \times 3 = (1 + \sqrt{-5}) \times (1 - \sqrt{-5})$.

Consider the ideal (2). Although 2 is irreducible in $\mathfrak{o}_K$ we have, from the formula for the product of ideals,

$$(2, 1+\sqrt{-5})^2 = (4, 2(1+\sqrt{-5}), (1+\sqrt{-5})^2) = (4, 2+2\sqrt{-5}, 4+2\sqrt{-5}) = (2)$$

the last equality since $2 = (4+2\sqrt{-5}) - (2+2\sqrt{-5})$. Likewise

$$(3) = (3, 1+\sqrt{-5})(3, 1-\sqrt{-5}), \qquad (1\pm\sqrt{-5}) = (2, 1+\sqrt{-5})(3, 1\pm\sqrt{-5})$$

so the ideal (6) can be written as a product of ideals

$$(6) = (2, 1+\sqrt{-5})^2(3, 1+\sqrt{-5})(3, 1-\sqrt{-5}).$$

All the ideals on the right hand side are prime ideals (their norms are 2 and 3, hence they are maximal), and it is not too hard to check that this is the only representation of (6) as a product of prime ideals. In this chapter we show that this is a general phenomenon.

Lemma 8.1. *$I \subset \mathfrak{o}_K$ a nonzero ideal, $\alpha \in K$ such that $\alpha I \subset I$. Then $\alpha \in \mathfrak{o}_K$.*

Proof. For every $k \geq$, $\alpha^k I \subset I$. Let $0 \neq \beta \in I$. Then $\mathbb{Z}[\alpha]\beta \subset I$, so $\mathbb{Z}[\alpha] \subset \beta^{-1} I$ is a finitely-generated $\mathbb{Z}$-module, so $\alpha \in \mathfrak{o}_K$. □

——— *Lecture 6* ———

Lemma 8.2. *(i) Let I be a nonzero ideal. Then there exists prime ideals $P_1, \ldots, P_r$ (not necessarily distinct) with $I \supset P_1 \cdots P_r$.*

(ii) Let P, $P_1, \ldots, P_r$ be prime ideals with $P \supset P_1 \cdots P_r$. Then for some i, $P = P_i$.

Proof. (i) Induction on $N(I)$. If $I = \mathfrak{o}_K$ or I is prime, trivial. Otherwise there exist α, $\beta \in \mathfrak{o}_K \setminus I$ with $\alpha\beta \in I$. Then $I + (\alpha)$, $I + (\beta)$ properly contain I, so by induction, $I + (\alpha) \supset P_1 \cdots P_r$ and $I + (\beta) \supset Q_1 \cdots Q_s$ for prime ideals P_i, Q_j. Then $P_1 \cdots Q_s \subset (I + (\alpha))(I + (\beta)) = I^2 + \alpha I + \beta I + (\alpha\beta) \subset I$. [OR: $I \supset (I + (\alpha))(I + (\beta)) \supset P_1 \cdots Q_s$.]

(ii) Suppose $P \neq P_1$. Let $x \in P_1 \setminus P$. Then for every $y \in P_2 \cdots P_r$, $xy \in P_1 \cdots P_r \subset P$, and so as P is prime and $x \notin P$, $y \in P$. Therefore $P \supset P_2 \cdots P_r$, so done by induction. □

Corollary 8.3. *Let $I \subset \mathfrak{o}_K$ be a nonzero proper ideal, $0 \neq \alpha \in I$. Then there exists $\beta \in \mathfrak{o}_K \setminus (\alpha)$ such that $\beta I \subset (\alpha)$.*

Proof. Let P be a prime ideal containing I. It's enough to find $\beta \notin (\alpha)$ with $\beta P \subset (\alpha)$. By Lemma 8.2, there exists a family of prime ideals $P_1, \ldots, P_r$ with $(\alpha) \supset P_1 \cdots P_r$, and WLOG $P = P_1$. Choose such a family with r minimal. Then $(\alpha) \not\supset P_2 \cdots P_r$, so let $\beta \in P_2 \cdots P_r \setminus (\alpha)$. Then $\beta P \subset P P_2 \cdots P_r \subset (\alpha)$ as required. □

Theorem 8.4. *Let $I \subset \mathfrak{o}_K$ be a nonzero ideal. Then there exists a nonzero ideal J such that IJ is principal.*

Proof. If $I = \mathfrak{o}_K$ then $J = \mathfrak{o}_K$ will do. So assume that $I \subsetneqq \mathfrak{o}_K$, and that the lemma holds for every ideal $I' \supsetneqq I$. Let $0 \neq \alpha \in I$, and choose β as in Corollary 8.3. Then $\alpha^{-1}\beta \notin \mathfrak{o}_K$ and $\alpha^{-1}\beta I \subset \mathfrak{o}_K$. Hence by Lemma 8.1, $\alpha^{-1}\beta I \not\subset I$. Therefore $I \subsetneqq I' := I + \alpha^{-1}\beta I \subset \mathfrak{o}_K$. So by induction, there exists $J' \subset \mathfrak{o}_K$ with $I'J' = (\gamma)$ principal. Let $J = \alpha J' + \beta J' = (\alpha, \beta)J'$. Then $IJ = (\alpha, \beta)IJ' = \alpha I'J' = (\alpha\gamma)$ is principal. □

Remark. The key point, which is somewhat obscured in this proof, is that if α, P and β are as in Corollary 8.3 then $(\alpha, \beta)P = (\alpha)$.

Alternatively: We use induction on $N(I)$. If $I = \mathfrak{o}_K$ then $J = \mathfrak{o}_K$ will do. Otherwise, pick a prime ideal P containing I. Let $0 \neq \alpha \in P$, and choose $\beta \notin (\alpha)$ as in the corollary. Write $H = (\alpha, \beta)$. Then $\alpha^{-1}\beta P \subset \mathfrak{o}_K$ but $\alpha^{-1}\beta P \not\subset P$. So as P is maximal, $P + \alpha^{-1}\beta P = \mathfrak{o}_K$. Therefore $PH = \alpha P + \beta P = (\alpha)$.

Now $IH \subset PH = (\alpha)$, so $I' := \alpha^{-1}IH \subset \mathfrak{o}_K$ is an ideal. Claim: $I' \supsetneqq I$. Indeed, as $H \supset (\alpha)$, $I' \supset I$. If $I' = I$ then $(\alpha^{-1}H)I = I$, and $\alpha^{-1}\beta \in \alpha^{-1}H \setminus \mathfrak{o}_K$, contradicting (8.1).

So by induction on norm, there exists J' with $I'J'$ principal, and then $I(J'H) = \alpha I'J'$ is also principal. □

Now we can prove the main properties of ideals in rings of integers.

Theorem 8.5. *Let I, I', J be nonzero ideals of $\mathfrak{o}_K$.*

(i) (Cancellation.) If $IJ = I'J$, then $I = I'$.

(ii) ("To divide is to contain") $I \supset J$ iff there exists an ideal H with $IH = J$.

(iii) There are unique distinct prime ideals $P_1, \dots P_r$ of $\mathfrak{o}_K$ and integers $a_i \geq 1$ such that $I = P_1^{a_1} \cdots P_r^{a_r}$.

Proof. (i) Choose J' with $JJ' = (\alpha)$ principal by Theorem 8.4. Then $\alpha I = IJJ' = I'JJ' = \alpha I'$, so $I = I'$.

(ii) "If" is clear. So suppose $I \supset J$. Let $II' = (\alpha)$ as in Theorem 8.4. Then $JI' \subset (\alpha)$, so $H := \alpha^{-1}JI' \subset \mathfrak{o}_K$ is an ideal, and $IH = \alpha^{-1}II'J = J$.

(iii) Existence: induction on $N(I)$. If $I \neq \mathfrak{o}_K$, let $P \supset I$, P prime. By (ii), $I = PJ$ for some $J \supset I$, and by (i), $J \neq I$. So by induction J is a product of prime ideals, hence so is I.

Uniqueness: suppose that $I = P_1 \cdots P_k = Q_1 \cdots Q_\ell$. We have to show that $k = \ell$ and that after reordering, $P_i = Q_i$. If $k = 0$ then $I = \mathfrak{o}_K$, hence $\ell = 0$ and nothing to prove. Otherwise, as $I \subset P_1$, by Lemma 8.2(ii), after possibly reordering the Q_i, $P_1 = Q_1$. Then by cancellation $P_2 \cdots P_k = Q_2 \cdots Q_\ell$. Repeating gives the result. □

Definition. Ideals I, J in $\mathfrak{o}_K$ are *equivalent* if there exist non-zero α, β with $\alpha I = \beta J$.

Trivially this is an equivalence relation, and an ideal is equivalent to $\mathfrak{o}_K$ iff it is principal.

Theorem 8.6. *The set of ideal classes is an abelian group under multiplication, the* class group $\mathrm{Cl}(K)$ *of* K*. The unit element is the the class of principal ideals.*

Proof. All the axioms are trivial apart from the existence of inverses, which is Theorem 8.4. ☐

——— *Lecture 7* ———

Variant: define a *fractional ideal* of K to be any subset of $\mathfrak{o}_K$ of K of the form αI, where $I \subset \mathfrak{o}_K$ is a nonzero ideal and $0 \neq \alpha \in K$. We multiply fractional ideals in the same way as ideals. A principal fractional ideal is one of the form $\alpha\mathfrak{o}_K$. We can restate most of the above as:

Theorem 8.7. *The set of fractional ideals of K is a group, freely generated by the prime ideals. The principal fractional ideals form a subgroup, and the quotient by it is isomorphic to* $\mathrm{Cl}(K)$*.* ☐

Remark. If I is a (usual) ideal then its inverse in the group of fractional ideals is $\alpha^{-1}J$, where $IJ = (\alpha)$.

Proposition 8.8. $\mathfrak{o}_K$ *is a PID iff* $\mathfrak{o}_K$ *is a UFD iff* $\mathrm{Cl}(K)$ *is trivial.*

Proof. By definition, $\mathrm{Cl}(K)$ is trivial iff $\mathfrak{o}_K$ is a PID.

PID $\implies$ UFD: see GRM.

UFD $\implies$ PID: enough to show that every prime ideal is principal. Let P be a prime ideal, $0 \neq \alpha \in P$. Factor $\alpha = \prod \pi_i$ where π_i are irreducible. Then some $\pi_j \in P$ as P is prime. As $\mathfrak{o}_K$ is a UFD, (π_j) is a (nonzero) prime ideal, hence maximal, and $(\pi_j) \subset P$, so $P = (\pi_j)$. ☐

Theorem 8.9. *Let I, J be nonzero ideals of $\mathfrak{o}_K$. Then $N(IJ) = N(I)N(J)$.*

Proof. STP (by previous theorem) that if P is prime then $N(IP) = N(I)N(P)$. Obviously $N(IP) = (\mathfrak{o}_K : I)(I : IP)$ so it's enough to show that $(I : IP) = N(P)$. By cancellation, $I \neq IP$. Claim: if $IP \subset J \subset I$ then $J = I$ or $J = IP$. Indeed, $J = IJ'$ for some J' by (ii), and then by cancellation, $P \subset J' \subset \mathfrak{o}_K$, and P is prime.

Let $\alpha \in I \setminus IP$. Then by the claim, $\alpha\mathfrak{o}_K + IP = I$. Consider the $\mathfrak{o}_K$-module homomorphism $\tilde{\alpha}\colon \mathfrak{o}_K/P \to I/IP$ given by multiplication by α. It is surjective, since $\mathrm{im}(\tilde{\alpha}) = (\alpha\mathfrak{o}_K + IP)/IP = I/IP$. It is also a homomorphism of $\mathfrak{o}_K/P$-vector spaces, and as $I \neq IP$, $\dim_{\mathfrak{o}_K/P}(I/IP) > 0$, so as it is surjective, $\dim = 1$ and therefore $\#(IP/I) = \#(\mathfrak{o}_K/P)$. ☐

Remark. This fails for the ring $R = \mathbb{Z}[2\sqrt{2}]$ and the prime ideal $P = (2, 2\sqrt{2})$, since $N(P) = 2$, whereas $P^2 = (4, 4\sqrt{2})$ and $N(P^2) = 8$.

9 Factorisation of rational primes

Recall that every prime ideal of $\mathfrak{o}_K$ contains a unique rational prime.

Theorem 9.1. *Let p be a rational prime, and let $\{P_i \mid 1 \leq i \leq k\}$ be the set of prime ideals of $\mathfrak{o}_K$ which contain p. Let $N(P_i) = p^{f_i}$. Then $(p) = P_1^{e_1} \cdots P_k^{e_k}$ for integers $e_i \geq 1$, satisfying $\sum e_i f_i = n$.*

Proof. The factorisation exists with $e_i \geq 1$ since the prime ideal factors of (p) are just the prime ideals containing p. Since $N((p)) = \left|\mathrm{N}_{K/\mathbb{Q}}(p)\right| = p^n$, the multiplicativity of the norm gives the stated equality. □

We say

- p *is ramified* in K if some $e_i > 1$; *unramified* if all $e_i = 1$.
- p *is inert* in K if (p) is prime (i.e. $r = 1 = e_1$, and so $f_1 = n$)
- p *splits completely* in K if $r = n$ (which implies that all $e_i = f_i = 1$)
- p is *totally ramified* in K if $e_1 = n$ (so $r = 1 = f_1$).

Later we'll show that only finitely many primes are ramified.

For computation of the factorisation: the following theorem is often (although not universally) applicable.

Theorem 9.2. *(Dedekind's criterion) Suppose that $K = \mathbb{Q}(\theta)$ for some $\theta \in \mathfrak{o}_K$, with minimal polynomial $g \in \mathbb{Z}[T]$. Assume that p is a prime not dividing $(\mathfrak{o}_K : \mathbb{Z}[\theta])$. Let the reduction $\bar{g} \in \mathbb{F}_p[T]$ of g factor as $\bar{g} = \prod \bar{g}_i^{e_i}$, where $\bar{g}_i \in \mathbb{F}_p[T]$ are monic, irreducible and distinct, and $e_i \geq 1$. Let $g_i \in \mathbb{Z}[T]$ be any monic polynomial whose reduction mod p is $\bar{g}_i$. Then $(p) = \prod P_i^{e_i}$, where $P_i = (p, g_i(\theta))$ are distinct prime ideals. Moreover $N(P_i) = p^{f_i}$ where $f_i = \deg(g_i)$.*

In the proof that follows, we make liberal use of the 3rd isomorphism theorem: if $J \subset I \subset R$ are ideals then $(R/J)/(I/J) \simeq R/I$.

Proof. First prove under the assumption $\mathfrak{o}_K = \mathbb{Z}[\theta]$.

Step 1. Since $\bar{g}_i \in \mathbb{F}_p[T]$ is irreducible,

$$\mathfrak{o}_K/P_i = \mathbb{Z}[\theta]/(p, g_i(\theta)) \simeq \mathbb{Z}[T]/(g, p, g_i) \simeq \mathbb{F}_p[T]/(\bar{g}, \bar{g}_i) = \mathbb{F}_p[T]/(\bar{g}_i)$$

is a field with $p^{\deg \bar{g}_i}$ elements. So P_i is prime, with norm p^{f_i}.

Step 2: Since $g = \prod g_i^{e_i} + ph$ for some $h \in \mathbb{Z}[T]$,

$$\prod P_i^{e_i} = \prod (p, g_i(\theta))^{e_i} \subset \prod (p, g_i(\theta)_i^{e}) \subset (p, \prod g_i(\theta)^{e_i}) = (p, -ph(\theta)) = (p).$$

Now compare norms: $N((p)) = \left|\mathrm{N}_{K/\mathbb{Q}}(p)\right| = p^n$, and $N(\prod P_i^{e_i}) = \prod p^{e_i f_i}$ by the multiplicativity of the norm and step 1. As $n = \deg g = \sum e_i \deg g_i$, the norms are equal, hence the inclusion $\prod P_i^{e_I} \subset (p)$ is an equality.

Lecture 8

In the general case, let $Q_i = p\mathbb{Z}[\theta] + g_i(\theta)\mathbb{Z}[\theta] \subset \mathbb{Z}[\theta]$. Step 1 shows that $\mathbb{Z}[\theta]/Q_i \simeq \mathbb{F}_p[T]/(\bar{g}_i)$ is a field with p^{f_i} elements. Consider the ring homomorphism $\phi\colon \mathbb{Z}[\theta]/Q_i \to \mathfrak{o}_K/P_i$, $\alpha + Q_i \mapsto \alpha + P_i$. As $\mathbb{Z}[\theta]/Q_i$ is a field, ϕ is injective. Its image is a subgroup of $\mathfrak{o}_K/P_i$ whose index divides both $(\mathfrak{o}_K : \mathbb{Z}[\theta])$ and the order of $\mathfrak{o}_K/P_i$, which is a power of p. As p doesn't divide $(\mathfrak{o}_K : H)$, ϕ is therefore an isomorphism. Now apply step 2.

□

Application: quadratic fields. Let $K = \mathbb{Q}(\sqrt{d})$, $d \notin \{0, 1\}$ a squarefree integer. We know that $\mathfrak{o}_K = \mathbb{Z}[\sqrt{d}]$ if $d \not\equiv 1 \pmod 4$, and $= \mathbb{Z}[(1 + \sqrt{d})/2]$ otherwise (in which case $(\mathfrak{o}_K : \mathbb{Z}[\sqrt{d}]) = 2$). The minimal polynomial $g = T^2 - d$ of $\sqrt{d}$ factors mod p as:

$$\bar{g} = \begin{cases} (T - \bar{a})(T + \bar{a}) & \text{if } p \neq 2, \left(\dfrac{d}{p}\right) = 1 \text{ and } a^2 \equiv d \pmod p \\ (T - \bar{d})^2 & \text{if } p = 2 \text{ or } p|d \\ \text{(irreducible)} & \text{if } p \neq 2, \left(\dfrac{d}{p}\right) = -1 \end{cases}$$

Dedekind's criterion therefore shows that if $p > 2$ then

- (inert) if d is not a square mod p, then (p) is prime (of norm p^2);
- (splits) if $(p, d) = 1$ and $d \equiv a^2$ mod p, then $(p) = PP'$ with $P = (p, a + \sqrt{d})$, $P' = (p, a - \sqrt{d})$ distinct prime ideals of norm p;
- (ramifies) if $p|d$ then $(p) = P^2$, where $P = (p, \sqrt{d})$ is prime (of norm p)

and that if $d \not\equiv 1 \pmod 4$ then $(2) = P^2$ with $P = (2, d - \sqrt{d})$ prime (of norm 2).

There remains the case $p = 2$, $d \equiv 1 \pmod 4$. Then $\mathfrak{o}_K = \mathbb{Z}[\theta]$ with $\theta = (1 + \sqrt{d})/2$, and $g = m_\theta = T^2 - T - (d - 1)/4$, and:

- (2 splits) if $d \equiv 1 \pmod 8$ then $g \equiv T(T - 1) \pmod 2$, and so $(2) = PP'$ where $P = (2, (1 + \sqrt{d})/2)$, $P' = (2, (1 - \sqrt{d})/2)$ are distinct of norm $= 2$;

- (2 is inert) if $d \equiv 5 \pmod 8$ then $g \equiv T^2+T+1 \pmod 2$, which is irreducible mod 2, hence (2) is prime.

Remark. Let $K = \mathbb{Q}(\theta)$ and suppose that $\mathfrak{o}_K = \mathbb{Z}[\theta]$. Then p splits completely in K iff $\bar{g}$ splits into distinct linear factors in $\mathbb{F}_p[T]$. This obviously implies that $n = [K : \mathbb{Q}] \leq p$. So if K is a number field in which some prime $p < [K : \mathbb{Q}]$ splits completely, $\mathfrak{o}_K$ cannot be of the form $\mathbb{Z}[\theta]$.

Theorem 9.3. *If p ramifies in K then $p|d_K$. In particular, only finitely many primes ramify in K.*

The converse is also true ($p|d_k \implies p$ ramifies), but the proof requires some Galois Theory.

Lemma 9.4. *If $\alpha \in \mathfrak{o}_K$ then $\mathrm{Tr}_{K/\mathbb{Q}}(\alpha^p) \equiv \mathrm{Tr}_{K/\mathbb{Q}}(\alpha)$ (mod p).*

Proof. By little Fermat it's enough to prove that $\mathrm{Tr}_{K/\mathbb{Q}}(\alpha^p) \equiv \mathrm{Tr}_{K/\mathbb{Q}}(\alpha)^p \pmod p$. But by the (generalised) binomial theorem

$$\begin{aligned}\mathrm{Tr}_{K/\mathbb{Q}}(\alpha)^p - \mathrm{Tr}_{K/\mathbb{Q}}(\alpha)^p &= \left(\sum \sigma_i(\alpha)\right)^p - \left(\sum \sigma_i(\alpha)^p\right) \\ &= \sum_{\substack{0 \leq k_1,\ldots,k_n < p \\ k_1+\cdots+k_n = p}} \frac{p!}{k_1! \cdots k_n!} \sigma_1(\alpha)^{k_1} \cdots \sigma_n(\alpha)^{k_n}\end{aligned}$$

and each of the coefficients is divisible by p. □

Proof of Theorem. Assume that $e_1 > 1$. Let $\alpha \in P_1^{e_1-1} P_e^{e_2} \cdots P_r^{e_r} \setminus (p)$. Then for any $\beta \in \mathfrak{o}_K$, $(\alpha\beta)^p \in (p)$, so by the Lemma, $\mathrm{Tr}_{K/\mathbb{Q}}(\alpha\beta) \equiv 0 \pmod p$.

Let (θ_i) be an integral basis for K, and write $\alpha = \sum b_i \theta_i$¡ where $b_i \in \mathbb{Z}$. Then

$$\sum c_i \mathrm{Tr}_{K/\mathbb{Q}}(\theta_i \theta_j) = \mathrm{Tr}_{K/\mathbb{Q}}(\alpha \theta_j) \equiv 0 \pmod p$$

and as $\alpha \notin (p)$, not all b_i are divisible by p. So the rows of the matrix $\mathrm{Tr}_{K/\mathbb{Q}}(\theta_i\theta_j)$ are linearly dependent mod p, hence $p|d_K$. □

Remarks. (1) With a bit more care (see example sheet) one can improve this to show that $\prod p^{(e_i-1)f_i}$ divides d_K.

(2) This can be useful in computing rings of integers. For example, let $K = \mathbb{Q}(\theta)$, $\theta = \sqrt[3]{p}$, $p \neq 3$ prime. Clearly $\mathbb{Z}[\theta] \subset \mathfrak{o}_K$ and $(p) = (\theta)^3$. So by the Theorem, $p|d_K$. Also $\mathrm{Disc}(\mathbb{Z}[\theta]) = (\mathfrak{o}_K : \mathbb{Z}[\theta])^2 d_K$. We have

$$\mathrm{Disc}(\mathbb{Z}[\theta]) = \det \mathrm{Tr}_{K/\mathbb{Q}} \begin{pmatrix} 1 & \theta & \theta^2 \\ \theta & \theta^2 & p \\ \theta^2 & p & p\theta \end{pmatrix} = \det \begin{pmatrix} 3 & 0 & 0 \\ 0 & 0 & 3p \\ 0 & 3p & 0 \end{pmatrix} = -27p^2$$

and since $p|d_K$ this means that $(\mathfrak{o}_K : \mathbb{Z}[\theta]) = 1$ or 3. (To distinguish which requires a further argument.)

Sketch proof of converse. Suppose that $(p) = P_1 \cdots P_r$ is unramified in K. Write $k_i = \mathfrak{o}_K/P_i$, a finite field extension of $\mathbb{F}_p$. Recall from Galois theory:

Theorem. *Let $k/\mathbb{F}_p$ be a finite extension. The trace map $\mathrm{Tr}_{k/\mathbb{F}_p}\colon k \to \mathbb{F}_p$ is surjective.*

[Proof: if $[k : \mathbb{F}_p] = n$ then $\mathrm{Tr}_{k/\mathbb{F}_p}(x) = x + x^p + \cdots + x^{p^{n-1}}$. As this is a polynomial of degree $< p^n = \#k$, there exists $x \in k$ with $\mathrm{Tr}_{k/\mathbb{F}_p}(x) = a \neq 0$. Then for any $b \in \mathbb{F}_p$, $\mathrm{Tr}_{k/\mathbb{F}_p}(a^{-1}bx) = b$.]

The Chinese Remainder Theorem (CRT) says that $\mathfrak{o}_K/(p) \simeq k_1 \times \cdots \times k_r$, and then linear algebra gives:

Lemma. *Let $\alpha \in \mathfrak{o}_K$. Then the reduction mod p of $\mathrm{Tr}_{K/\mathbb{Q}}(\alpha)$ is $\sum_i \mathrm{Tr}_{k_i/\mathbb{F}_p}(\alpha + P_i)$.*

Let's show that d_K is prime to p. Let $\alpha \in \mathfrak{o}_K \smallsetminus (p)$. It's enough to show that there exists $\beta \in \mathfrak{o}_K$ such that $\mathrm{Tr}_{K/\mathbb{Q}}(\alpha\beta) \not\equiv 0 \pmod p$.

As $\alpha \notin (p) = P_1 \cdots P_r$, there exists j with $\alpha \notin P_j$. Let $x \in k_j$ with $\mathrm{Tr}_{k_j/\mathbb{F}_p}(x) = 1$. By CRT, there exists $\beta \in \mathfrak{o}_K$ with $\beta \in P_i$ for $i \neq j$ and $\beta + P_j = (\alpha + P_j)^{-1}x$. Then by the Lemma, $\mathrm{Tr}_{K/\mathbb{Q}}(\alpha\beta) \equiv 1 \pmod p$. ◻

Lecture 9

10 Geometry of numbers

We'll next prove some things which are not "pure algebra", using the embeddings $\sigma_i\colon K \hookrightarrow \mathbb{C}$ and real geometry. The main results of the next 5 lectures are:

Theorem. *(i) The class group $\mathrm{Cl}(K)$ is finite.*

(ii) The group of units $\mathfrak{o}_K^*$ is finitely generated, of rank $r + s - 1$

(Recall that $n = r + 2s$ where r is the number of real embeddings and s the number of conjugate pairs of complex embeddings.) Neither of these can be proved by algebra alone.

Definition. A *lattice* in $\mathbb{R}^n$ is a subgroup $\Lambda \subset \mathbb{R}^n$ generated by n $\mathbb{R}$-linearly independent elements.

E.g. $\mathbb{Z}^n \subset \mathbb{R}^n$. Let $e_1, \ldots, e_n$ be linearly independent, $\Lambda \subset \mathbb{R}^n$ the subgroup they generate. The *fundamental parallelopiped* attached to (e_i) is the subset $\mathcal{P} = \{\sum x_i e_i \mid 0 \leq x_i < 1\}$. [Draw a picture.] The *covolume* of Λ is $\mathrm{covol}(\Lambda) = \mathrm{vol}\,\mathcal{P}$. Equivalently,

$$\mathrm{covol}(\Lambda) = |\det(e_{ij})|$$

where $e_i = (e_{i1}, \ldots, e_{in})$. The covolume doesn't depend on the choice of basis (by Proposition 5.1, for example).

Example: let $K = \mathbb{Q}(\sqrt{-d})$ be imaginary quadratic. Consider the embedding $\sigma\colon K \hookrightarrow \mathbb{C}$, $\sqrt{-d} \mapsto i\sqrt{d}$. Then a basis for $\sigma(\mathfrak{o}_K)$ is $\{1, \sigma(\theta)\}$ where $\theta = \sqrt{-d}$ if $d \not\equiv 3 \pmod 4$, $\theta = (1+\sqrt{-d})/2$ otherwise. [Pictures.]

We have $\operatorname{covol}\sigma(\mathfrak{o}_K) = \sqrt{d}$ in the first case, $\sqrt{d}/2$ in the second, so in both cases, $\operatorname{covol}\sigma(\mathfrak{o}_K) = (1/2)\,|d_K|^{1/2}$.

Now state a special case of Minkowski's Theorem, which we'll prove next time (in greater generality):

Theorem 10.1. *Let $\Lambda \subset \mathbb{C}$ be a lattice, and $X = \{z \in \mathbb{C} \mid |z|^2 \leq R\}$. If $\pi R \geq 4\operatorname{covol}(\Lambda)$ then $X \cap \Lambda \neq \{0\}$.*

Remark. For $\Lambda = \mathbb{Z}[i]$ the theorem is rather easy. The point is that it holds for any "shape" of lattice.

Assuming this we can prove:

Theorem 10.2. *Let $K = \mathbb{Q}(\sqrt{-d})$, $I \subset \mathfrak{o}_K$ a nonzero ideal. Then there exists nonzero $\alpha \in I$ with $\mathrm{N}_{K/\mathbb{Q}}(\alpha) \leq c_K N(I)$ where $c_K = (2/\pi)\,|d_K|^{1/2}$.*

Proof. Consider $\sigma(I) \subset \sigma(\mathfrak{o}_K) \subset \mathbb{C}$, which is a lattice of covolume $N(I)\operatorname{covol}\sigma(\mathfrak{o}_K) = (1/2)\,|d_K|^{1/2} N(I)$. Intersect it with $X = \{z \in \mathbb{C} \mid |z| \leq R\}$ with $R = (2/\pi)\,|d_K|^{1/2} N(I)$. Since $\pi R \geq 4\operatorname{covol}\sigma(I)$ (in fact $=$), by the Theorem there exists $\alpha = u + v\sqrt{-d} \in I \setminus \{0\}$ with $\sigma(\alpha) \in X$. But then $\mathrm{N}_{K/\mathbb{Q}}(\alpha) = u^2 + dv^2 = |\sigma(\alpha)|^2 \leq R$ as required. □

Corollary 10.3. *Let $K = \mathbb{Q}(\sqrt{-d})$. Then:*

(i) $\mathrm{Cl}(K)$ *is finite.*

(ii) Every ideal class of K contains an ideal of norm $\leq c_K = (2/\pi)\,|d_K|^{1/2}$.

(iii) $\mathrm{Cl}(K)$ *is generated by the classes of prime ideals of norm $\leq c_K$.*

Proof. (ii) Let I be a nonzero ideal, and choose J with $IJ = (\beta)$ principal. Let $0 \neq \alpha \in J$ with $\left|\mathrm{N}_{K/\mathbb{Q}}(\alpha)\right| \leq c$. Then since $\alpha \in J$, for some ideal I' we have $JI' = (\alpha)$ and $N(I') = N((\alpha))/N(J) = \mathrm{N}_{K/\mathbb{Q}}(\alpha)/N(J) \leq c_K$. Also $(\alpha\beta) = \alpha IJ = \beta JI'$, so $\alpha I = \beta I'$ and thus I' is equivalent to I.

(i) By Corollary 7.2(ii), the set of ideals of norm $\leq c_K$ is finite, so by (ii), $\mathrm{Cl}(K)$ is finite.

For (iii) it's enough to write I' from (ii) as a product $P_1 \cdots P_r$ of primes, as then $N(P_i) \leq N(I') \leq c_K$. □

Some examples. $K = \mathbb{Q}(i)$. Then $d_K = -4$, so every ideal is equivalent to an ideal I with $N(I) \leq c_K = 4/\pi < 2$, which implies $N(I) = 1$, i.e. $I = \mathfrak{o}_K$. So every ideal is principal, and we have another (different) proof that $\mathbb{Z}[i]$ is a PID.

$K = \mathbb{Q}(\sqrt{-5})$. We have seen that $\mathfrak{o}_K$ is not a PID, so $\mathrm{Cl}(K) \neq \{1\}$. What is it?

We have $d_K = -20$ so can take $c = \frac{\sqrt{80}}{\pi} < \frac{9}{\pi} < 3$. So every ideal class contains an ideal of norm ≤ 2. Now $(2) = P^2$ with $P = (2, 1+\sqrt{5})$ non-principal. So P is the only ideal of norm 2, hence $\mathrm{Cl}(K) = \{[\mathfrak{o}_K], [P]\}$ has order 2.

——————————— *Lecture 10* ———————————

Theorem 10.4 (Minkowski's Theorem)**.** *Let $\Lambda \subset \mathbb{R}^n$ be a lattice, $X \subset \mathbb{R}^n$ a [measurable] convex subset, symmetric about* 0*. If* $\mathrm{vol}(X) > 2^n \,\mathrm{covol}(\Lambda)$*, or if X is compact and* $\mathrm{vol}(X) \geq 2^n \,\mathrm{covol}(\Lambda)$*, then $X \cap \Lambda \neq \{0\}$.*

Here $X \subset \mathbb{R}^n$ is:

- *convex* if $x, y \in X$ and $t \in [0,1]$ implies that $tx + (1-t)y \in X$
- *symmetric about* 0 if $x \in X$ implies $(-x) \in X$
- *compact* if it is closed and bounded.

(Measurable means that the volume of X is defined — it follows from the other conditions.)

Lemma 10.5 (Blichfeldt's Lemma)**.** *Let $\Lambda \subset \mathbb{R}^n$ be a lattice, $Y \subset \mathbb{R}^n$ a measurable subset. If* $\mathrm{vol}(Y) > \mathrm{covol}(\Lambda)$ *then there exist distinct $x, y \in Y$ with $x - y \in Y$.*

Proof. Let $\mathcal{P}$ be the fundamental parallelopiped with respect to some basis of Λ. For $\lambda \in \Lambda$ let $Y_\lambda = \{x \in Y \mid x - \lambda \in \mathcal{P}\} = Y \cap (\lambda + \mathcal{P})$. Then Y is the disjoint union of the (countable) family of (measurable) subsets Y_λ, so $\mathrm{vol}(Y) = \sum_\lambda \mathrm{vol}(Y_\lambda)$. Now $(-\lambda) + Y_\lambda \subset \mathcal{P}$, so as $\mathrm{vol}(Y) > \mathrm{vol}(\mathcal{P})$ there exist $\lambda \neq \mu \in \Lambda$ for which $(-\lambda) + Y_\lambda$ and $(-\mu) + Y_\mu$ are not disjoint, so contain some common v. Then $x = v + \lambda$, $y = v + \mu$ will do. □

Remark. Morally the proof is: consider $\pi\colon \mathbb{R}^n \to \mathbb{R}^n/\Lambda$. Then $\mathrm{vol}(Y) > \mathrm{covol}(\Lambda) = \mathrm{vol}(\mathbb{R}^n/\Lambda) \geq \mathrm{vol}(\pi(Y))$ so $Y \to \pi(Y)$ is not a bijection.

Proof of Minkowski's Theorem. First suppose $\mathrm{vol}(X) > 2^n \,\mathrm{covol}(\Lambda) = \mathrm{covol}(2\Lambda)$. Then by Blichfeldt, there exist $x, y \in X$ with $0 \neq x - y \in 2\Lambda$. By symmetry, $-y \in X$ and by convexity, $\lambda = (x + (-y))/2 \in X$. But $0 \neq \lambda \in \Lambda$.

Suppose that X is compact and $\mathrm{vol}(X) = 2^n \,\mathrm{covol}(\Lambda)$. For $\delta > 0$, let $X_\delta = \{(1+\delta)x \mid x \in X\}$. Then X_δ is convex and symmetric about 0, and $\mathrm{vol}(X_\delta) = (1+\delta)\,\mathrm{vol}(X) > 2^n \,\mathrm{covol}(\Lambda)$, so $X_\delta \cap \Lambda \neq \{0\}$. But also $X_\delta \cap \Lambda$ is finite, since X_δ is bounded. As X is closed, $X \cap \Lambda = \bigcap_{\delta > 0} X_\delta \cap \Lambda$ equals $X_{\delta'} \cap \Lambda$ for some $\delta' > 0$, so by the first part is $\neq \{0\}$. □

Consider the embeddings $\sigma_i \colon K \hookrightarrow \mathbb{R}$ $(1 \le i \le r)$ and $\sigma_i \colon K \hookrightarrow \mathbb{C}$ $(r < i \le r+s)$. Their product is an embedding

$$\sigma = (\sigma_1, \dots, \sigma_{r+s}) \colon K \hookrightarrow \mathbb{R}^r \times \mathbb{C}^s \simeq \mathbb{R}^n$$

(we identify $\mathbb{C}$ with $\mathbb{R}^2$ using the basis $\{1, i\}$).

Proposition 10.6. *$\sigma(\mathfrak{o}_K)$ is a lattice in $\mathbb{R}^n$ of covolume $2^{-s} |d_K|^{1/2}$.*

Proof. Let $\omega_1, \dots, \omega_n$ be an integral basis. Then $e_i = \sigma(\omega_i) \in \mathbb{R}^n$ is the vector

$$e_i = (\sigma_1(\omega_i), \dots, \sigma_r(\omega_i), \operatorname{Re}\sigma_{r+1}(\omega_i), \operatorname{Im}\sigma_{r+1}(\omega_i), \dots)$$

and so $\det(e_{ij}) = \pm(-1/2i)^s \det \sigma_j(\omega_i)_{1 \le i,j \le n}$, since if $r < j \le r+s$,

$$(\sigma_j(\alpha), \overline{\sigma}_j(\alpha)) = (\operatorname{Re}\sigma_j(\alpha), \operatorname{Im}\sigma_j(\alpha)) \begin{pmatrix} 1 & i \\ 1 & -i \end{pmatrix}.$$

So as $\det(\sigma_j(\omega_i))^2 = d_K \ne 0$, (e_i) is a basis for $\mathbb{R}^n$ and so $\sigma(\mathfrak{o}_K)$ is a lattice of covolume $|\det(e_{ij})| = 2^{-s} |d_K|^{1/2}$. □

Applying Proposition 7.1 then gives:

Corollary 10.7. *If $I \subset \mathfrak{o}_K$ is a nonzero ideal then $\sigma(I)$ is a lattice of covolume $2^{-s} |\operatorname{disc}(I)|^{1/2} = 2^{-s} N(I) |d_K|^{1/2}$.*

Theorem 10.8. *For any nonzero ideal $I \subset \mathfrak{o}_K$, there exists $0 \ne \alpha \in I$ with $|\mathrm{N}_{K/\mathbb{Q}}(\alpha)| \le c_K N(I)$, where*

$$c_K = \left(\frac{4}{\pi}\right)^s \frac{n!}{n^n} |d_K|^{1/2} \qquad \textit{(Minkowski's constant)}.$$

Corollary 10.9. *Every ideal class of K contains an ideal of norm $\le c$. In particular, $\operatorname{Cl}(K)$ is finite, and is generated by the classes of prime ideals of norm $\le cK$.*

Proof. Same as for imaginary quadratic fields — Corollary 10.3. □

Lecture 11

Proof of 10.8 — real quadratic case. Let $K = \mathbb{Q}(\sqrt{d})$ be real quadratic. Then

$$\sigma \colon K \hookrightarrow \mathbb{R}^2, \qquad \sigma(u + v\sqrt{d}) = (u + v\sqrt{d}, u - v\sqrt{d}).$$

As we saw at end of last lecture, if $\alpha = u + v\sqrt{d} \in K$, $\mathrm{N}_{K/\mathbb{Q}}(\alpha) = \sigma_1(\alpha)\sigma_2(\alpha) = u^2 - dv^2$, so $|\mathrm{N}_{K/\mathbb{Q}}(\alpha)| \le R$ iff $\sigma(\alpha)$ lies in the area bounded by the hyperbolae $x_1 x_2 = \pm R$. (Picture)

So apply Minkowski's theorem, we need to choose a convex symmetric set contained in this region. The largest such is the square

$$X = \left\{(x_1, x_2) \in \mathbb{R}^2 \;\middle|\; |x_1| + |x_2| \le 2R^{1/2}\right\}$$

whose area is $8R$. Then by Minkowski's theorem, $X \cap \Lambda \ne \{0\}$ as long as $8R \ge 4\,\mathrm{covol}\,\sigma(I) = 4\,|d_K|^{1/2}\,N(I)$. So there exists $0 \ne \alpha \in I$ with $\left|\mathrm{N}_{K/\mathbb{Q}}(\alpha)\right| \le R = (1/2)\,|d_K|^{1/2}\,N(I)$, which is $c_K N(I)$ in the special case $n = r = 2$, $s = 0$.

— *general case.* The two quadratic cases suggest that we should define, for any K,

$$X = X_R = \left\{(x_1, \ldots, x_r, z_1, \ldots, z_s) \in \mathbb{R}^r \times \mathbb{C}^s \;\middle|\; \sum_{1 \le j \le r} |x_j| + 2 \sum_{r < j \le r+s} |z_j| \le nR^{1/n}\right\}.$$

The *Arithmetic-Geometric Mean* inequality (AGM) says that for reals $a_i \ge 0$,

$$(a_1 \cdots a_n)^{1/n} \le \frac{a_1 + \cdots + a_n}{n}$$

and applying this to $a_j = |x_j|$ $(1 \le j \le r)$, $a_{r+2j-1} = a_{r+2j} = |z_j|$ $(1 \le j \le s)$ gives

$$(\underline{x}, \underline{z}) \in X_R \implies \prod_{j=1}^{r} |x_j| \prod_{j=1}^{s} |z_j|^2 \le R.$$

and therefore $\sigma(\alpha) \in X_R \implies \left|\mathrm{N}_{K/\mathbb{Q}}(\alpha)\right| \le R$.

Lemma 10.10.

$$\mathrm{vol}(X_R) = 2^r \left(\frac{\pi}{2}\right)^s \frac{n^n}{n!} R.$$

(Proof: tedious calculation by induction on r, s.) By Minkowski's theorem, there exists $0 \ne \alpha \in I$ with $\left|\mathrm{N}_{K/\mathbb{Q}}(\alpha)\right| \le R$ if R satisfies $\mathrm{vol}(X_R) \ge 2^n 2^{-s}\,|d_K|^{1/2}\,N(I)$, which by the Lemma is equivalent to $R \ge c_K N(I)$. □

Special cases worth remembering:

- K real quadratic $\implies c_K = (1/2)\,|d_K|^{1/2}$.
- K imaginary quadratic $\implies c_K = (2/\pi)\,|d_K|^{1/2}$.

For large n, the factor $n!/n^n$ is a significant gain (e.g. $10!/10^{10} = 0.00036...$)

A more complicated example. Let $K = \mathbb{Q}(\sqrt{-17})$. Then $d_K = -68$, so $c_K = 2\frac{\sqrt{68}}{\pi} < 2 \times \frac{9}{3} < 6$. So $\mathrm{Cl}(K)$ is generated by the classes of primes of norm ≤ 5 — i.e. by primes of norm 2, 3, 5, since a prime of norm p^2 equals (p), so is principal.

- $p=5$: $-17 \equiv -2$ not a square (mod 5) so 5 is inert.
- $p=3$: $-17 \equiv 1^2 \pmod 3$ so $(3)=P_3P_3'$ with
$$P_3 = (3, 1+\sqrt{-17}), \qquad P_3' = (3, 1-\sqrt{-17}).$$
- $p=2$: as $-17 \not\equiv 1 \pmod 4$, $(2) = P_2^2 = (2, 1+\sqrt{-17})^2$ is ramified.

In the class group we have the relations $[P_2]^2 = 1 = [P_3][P_3']$. Let's compute

$$\begin{aligned} P_3^2 &= (3, 1+\sqrt{-17})^2 = (9, 3+3\sqrt{-17}, -16+2\sqrt{-17}) \\ &= (9, 3+3\sqrt{-17}, 2+2\sqrt{-17}) = (9, 1+\sqrt{-17}) \end{aligned}$$

which has norm 9.

Now the norm of $1+\sqrt{-17}$ is 18, and $(1+\sqrt{-17}) \subset P_3^2$, so $(1+\sqrt{-17})$ equals P_3^2 times an ideal of norm 2, i.e. equals $P_2P_3^2$. So in $\mathrm{Cl}(K)$, $[P_3]^2 = [P_2]^{-1} = [P_2]$. As P_2 is not principal (otherwise could solve $u^2+17v^2=2$ in integers), this means that $\mathrm{Cl}(K) \simeq \mathbb{Z}/4\mathbb{Z}$, generated by $[P_3]$.

A quintic example. Let $K = \mathbb{Q}(\theta)$, where θ is a root of $g = T^5 - T + 1$. (This is irreducible mod 5 hence irreducible.) Easily see that $(r,s) = (1,2)$, and the discriminant of g is $2689 = 19 \times 151$ which is squarefree. So $\mathfrak{o}_K = \mathbb{Z}[\theta]$, $d_K = 2689$. Then
$$c_K = \left(\frac{4}{\pi}\right)^2 \frac{5!}{5^5}\sqrt{2869} = 3.3...$$
so $\mathrm{Cl}(K)$ is generated by prime ideals of norm 2 or 3. By Dedekind's criterion, a prime of norm p exists iff g has a root mod p. But if $p=2$ or 3, it doesn't, so $\mathrm{Cl}(K)$ is trivial.

Remark: one might think that the class group grows with the field, in some sense. This does not appear to be the case. Contrast:

- If $K = \mathbb{Q}(\sqrt{-d})$ is imaginary quadratic, it is known that $\#\,\mathrm{Cl}(K) \to \infty$ as $d \to \infty$. In particular, we know that $\mathrm{Cl}(K) \neq \{1\}$ if $d > 163$.
- For $K = \mathbb{Q}(\sqrt{d})$ is real quadratic, it appears that there are infinitely many d for which $\mathrm{Cl}(K)$ is trivial (but not proved).

Lecture 12

Example. Let $K = \mathbb{Q}(\sqrt{10})$. Then $c_K = (1/2)\sqrt{40} = \sqrt{10} < 4$, so $\mathrm{Cl}(K)$ is generated by the classes of the primes of norm 2 and 3.

- $(2) = P_2^2 = (2, \sqrt{10})^2$ is ramified;

- $(3) = P_3P_3' = (3, 1+\sqrt{10})(3, 1-\sqrt{10})$ splits.

So $[P_2]^2 = 1 = [P_3][P_3']$ in $\mathrm{Cl}(K)$. To find more relations, compute norms of small elements of $\mathfrak{o}_K$ (since relations between $[P_2]$ and $[P_3]$ in $\mathrm{Cl}(K)$ will be of the form $P_2^aP_3^b = (\alpha)$ for some $\alpha \in \mathfrak{o}_K$ of norm $\pm 2^a3^b$). For example, $\mathrm{N}_{K/\mathbb{Q}}(1+\sqrt{10}) = -9$, and P_3 divides $(1+\sqrt{10})$. But $1+\sqrt{10} \notin P_3'$ so we must have $(1+\sqrt{10}) = P_3^2$. Likewise, $\mathrm{N}_{K/\mathbb{Q}}(2-\sqrt{10}) = -6$ and $2-\sqrt{10} = 3-(1+\sqrt{10}) \in P_3$, so $(2-\sqrt{10}) = P_2P_3$. So $[P_2] = [P_3] = [P_3']$, and either $\mathrm{Cl}(K) = \{1\}$ or $\mathrm{Cl}(K) \simeq \mathbb{Z}/2\mathbb{Z}$ with P_2 a generator. Is P_2 principal? If so then it is generated by $\alpha = u + v\sqrt{10}$ and

$$\mathrm{N}_{K/\mathbb{Q}}(\alpha) = u^2 - 10v^2 = \pm 2.$$

But then $u^2 \equiv \pm 2 \pmod 5$ which is impossible. So $[P_2] \neq 1$ and $\mathrm{Cl}(K) \simeq \mathbb{Z}/2\mathbb{Z}$.

(Note that unlike in imaginary quadratic fields, a bit more work is needed to decide if an ideal is principal.)

Definition. The *class number* of $K$is the order of $\mathrm{Cl}(K)$, denoted h_K.

11 Units

Theorem 11.1 (Dirichlet's Unit Theorem)**.** *The groups $\mathfrak{o}_K^*$ of units of K is finitely generated, of rank $r+s-1$.*

The torsion subgroup of $\mathfrak{o}_K^*$ is the group of roots of unity in K (which is always cyclic, see Galois Theory). So equivalent statement is: there are units $\varepsilon_1, \ldots, \varepsilon_{r+s-1}$ such that every unit can be uniquely written as $\zeta\varepsilon_1^{m_1}\cdots\varepsilon_{r+s-1}^{m_{r+s-1}}$ for integers m_i and ζ a root of unity in K. Let's first do a couple of special cases.
Quadratic fields. Let $K = \mathbb{Q}(\sqrt{d})$ be quadratic. Then $\mathfrak{o}_K = \{\alpha = u + v\sqrt{d}\}$ with u, v integers or (if $d \equiv 1 \pmod 4$) halves of odd integers. As $\alpha \in \mathfrak{o}_K^* \iff \mathrm{N}_{K/\mathbb{Q}}(\alpha) = \pm 1$ (Corollary 4.3(ii)), for any unit α we have $u^2 - dv^2 = \pm 1$. If $d < 0$ this has only finitely many solutions, so $\mathfrak{o}_K^*$ is finite for K imaginary quadratic, in agreement with the Theorem ($r+s-1 = 0+1-1 = 0$).

Consider the case $d > 0$. Then every solution of *Pell's equation* $u^2 - dv^2 = 1$ gives a unit. From Number Theory IIC you know that Pell's equation has infinitely many solutions, so $\mathfrak{o}_K^*$ is infinite for K real quadratic. We can be more precise:

Theorem 11.2. *Let $K = \mathbb{Q}(\sqrt{d}) \subset \mathbb{R}$, $d > 0$. Then:*
(i) $\mathfrak{o}_K^*$ is infinite.
(ii) There exists a unique smallest unit $\varepsilon > 1$ (the fundamental unit *of K), and $\mathfrak{o}_K^* = \{\pm\varepsilon^m \mid m \in \mathbb{Z}\}$.*
(iii) If $d \neq 5$, then $\varepsilon = u + v\sqrt{d} \in \mathfrak{o}_K^*$ is the fundamental unit iff u, $v > 0$ and v is minimal. The fundamental unit of $\mathbb{Q}(\sqrt{5})$ is $(1+\sqrt{5})/2$.

Proof. (i) We give an alternative proof below.

(ii) As $K \subset \mathbb{R}$ the only roots of unity in K are ± 1. By (i) there exists a unit $\varepsilon = u + v\sqrt{d} \in \mathfrak{o}_K^* \setminus \{\pm 1\}$, $\mathrm{N}_{K/\mathbb{Q}}(\varepsilon) = \pm 1$. Claim: $\varepsilon > 1$ iff both u, $v > 0$. Indeed, all of the 4 numbers $\{\pm u \pm v\sqrt{d}\} = \{\pm\varepsilon, \pm 1/\varepsilon\}$ are units, and exactly one of them lies in each of the intervals $(-\infty, -1)$, $(-1, 0)$, $(0, 1)$, $(1, \infty)$. So $\varepsilon > 1$ iff ε is the largest of the four, which holds iff both u, $v > 0$.

So choose $\varepsilon > 1$ minimal (which exists since u, $v > 0$). If $\varepsilon' = u' + v'\sqrt{d} \in \mathfrak{o}_K^*$ with $\varepsilon' > 1$ then there exists a unique $m \geq 1$ with $\varepsilon^m \leq \varepsilon' < \varepsilon^{m+1}$. Then $1 \leq \varepsilon'/\varepsilon^m < \varepsilon$, so by minimality $\varepsilon' = \varepsilon^m$. So the units > 1 are just $\{\varepsilon^m \mid m \geq 1\}$, and so (by the division into four) $\mathfrak{o}_K^* = \{\pm\varepsilon^m \mid m \in \mathbb{Z}\}$.

(iii) The units with u, $v > 0$ are the powers $\varepsilon^m = u_m + v_m\sqrt{d}$ with $m \geq 1$. By binomial expansion we have $v_m \geq mu^{m-1}v$, and $2u \in \mathbb{Z}$, so if $m > 1$ and $u > 1/2$, $v_m > v$. If $u = 1/2$ then $dv^2 = \pm 1 + 1/4$ and so $d = 5$, $v = 1/2$ and $\varepsilon = (1+\sqrt{5})/2$. □

Another proof of (i). Rather than directly constructing a unit, we use Minkowski's Theorem to prove:

Proposition 11.3. *If $R \geq |d_K|^{1/2}$, there are infinitely many elements of $\mathfrak{o}_K$ with $\left|\mathrm{N}_{K/\mathbb{Q}}(\alpha)\right| \leq R$.*

Let's assume this. We know there are only finitely many ideals of norm $\leq R$, so there exists $\alpha \neq \beta \in \mathfrak{o}_K$ with $(\alpha) = (\beta)$. Then $\alpha/\beta \in \mathfrak{o}_K^*$.

To prove it, embed $\sigma\colon K \hookrightarrow \mathbb{R}^2$ as before, and consider the rectangle $Y_\delta = [-\delta, \delta] \times [-R/\delta, R/\delta]$ which lies within the hyperbolae $x_1 x_2 = \pm R$. [DRAW PICTURE]

Let $\delta = \delta_0 = 1$ say. As $\operatorname{vol} Y_\delta = 4R \geq 4 \operatorname{covol} \sigma(\mathfrak{o}_K) = 4\,|d_K|^{1/2}$, by Minkowski there exists $\alpha_0 \in \mathfrak{o}_K$ nonzero with $\left|\mathrm{N}_{K/\mathbb{Q}}(\alpha_0)\right| \leq R$. Now choose $\delta_1 < |\sigma_1(\alpha_0)|$. By Minkowski again, we get $0 \neq \alpha_1 \in \mathfrak{o}_K$ with $\left|\mathrm{N}_{K/\mathbb{Q}}(\alpha_1)\right| \leq R$, and also $|\sigma_1(\alpha_1)| \leq \delta_1 < |\sigma_1(\alpha_0)|$. Continuing in this way we get a sequence of nonzero elements $\alpha_k \in \mathfrak{o}_K$ with $\left|\mathrm{N}_{K/\mathbb{Q}}(\alpha_k)\right| \leq R$ as required. □

——— *Lecture 13* ———

We'll need:

Lemma 11.4. *A subgroup $\Lambda \subset \mathbb{R}^n$ is a lattice iff (i) it contains a basis of $\mathbb{R}^n$ and (ii) for every bounded $X \subset \mathbb{R}^n$, $X \cap \Lambda$ is finite.*

Remark. A subgroup $\Lambda \subset \mathbb{R}^n$ satisfying (ii) is said to be a *discrete* subgroup of $\mathbb{R}^n$. If $V \subset \mathbb{R}^n$ is the subspace spanned by Λ, then the lemma also shows that Λ is a lattice in V, so is freely generated by $m \leq n$ $\mathbb{R}$-linearly independent elements of V.

Proof. Let $\Lambda \subset \mathbb{R}^n$ be a lattice. By definition (i) holds. There is an invertible linear transformation $u\colon \mathbb{R}^n \to \mathbb{R}^n$ such that $u(\Lambda) = \mathbb{Z}^n$. Then X is bounded iff $u(X)$ is bounded, so we may assume $\Lambda = \mathbb{Z}^n$, in which case (ii) is obvious.

Conversely, suppose that $\Lambda \subset \mathbb{R}^n$ satisfies (i) and (ii). Again by change of basis we may assume by (i) that $\Lambda \supset \mathbb{Z}^n$. Let $S = \{x \in \Lambda \mid 0 \leq x_i < 1 \text{ for every } i\}$. Then by (ii), S is finite, and every element of Λ can be (uniquely) written as $x + \lambda$ with $x \in S$, $\lambda \in \mathbb{Z}^n$. So $(\Lambda : \mathbb{Z}^n)$ is finite of index d say, and then $\Lambda \subset d^{-1}\mathbb{Z}^n$. Therefore by GRM $\Lambda = \sum_{i=1}^n \mathbb{Z} e_i$ for some e_i. As $\mathbb{Z}^n$ spans $\mathbb{R}^n$ over $\mathbb{R}$, so does Λ, hence (e_i) is a basis for $\mathbb{R}^n$ and Λ is a lattice. □

To deal with the general case, we first bound the torsion in $\mathfrak{o}_K^*$.

Lemma 11.5. *Let $C > 0$. Then $\{\alpha \in \mathfrak{o}_K \mid$ for every i, $|\sigma_i(\alpha)| \leq C\}$ is finite.*

Proof. The characteristic polynomial of α is

$$\prod_i (T - \sigma_i(\alpha)) = T^n + \sum_{r=1}^n c_r T^{n-r} = T^n + \sum_{r=1}^n (-1)^r \sum_{i_1 < \cdots < i_r} \sigma_{i_1}(\alpha) \cdots \sigma_{i_r}(\alpha) T^{n-r},$$

$$|c_r| \leq \binom{n}{r} C^r.$$

As $c_r \in \mathbb{Z}$ there are only finitely many such polynomials. $\square$

Corollary 11.6. *The group of roots of unity of K is finite (hence cyclic).*

Proof. If α is a root of unity then for every i, $|\sigma_i(\alpha)| = 1$. So the result follows from the Lemma. $\square$

To show that $\mathfrak{o}_K^*$ is finitely generated it is convenient to pass to an additive group, so we take logarithms!

Key Definition 11.7. The *logarithmic embedding* is the map $\mathcal{L}\colon K^* \to \mathbb{R}^{r+s}$, given by

$$\mathcal{L}(\alpha) = (\mathcal{L}(\alpha)_i) = (\log|\sigma_1(\alpha)|, \ldots, \log|\sigma_r(\alpha)|, 2\log|\sigma_{r+1}(\alpha)|, \ldots, 2\log|s_{r+s}(\alpha)|).$$

Note from the definition that:

(i) $\mathcal{L}$ is a homomorphism.

(ii) If $\alpha \in K^*$ then $\sum \mathcal{L}(\alpha)_i = \log\left|\mathrm{N}_{K/\mathbb{Q}}(\alpha)\right|$; indeed,

$$\left|\mathrm{N}_{K/\mathbb{Q}}(\alpha)\right| = \prod_{i=1}^{n} |\sigma_i(\alpha)| = \prod_{i=1}^{r} |\sigma_i(\alpha)| \prod_{i=r+1}^{n} |\sigma_i(\alpha)| \left|\overline{\sigma_i(\alpha)}\right| = \prod_{i=1}^{r} |\sigma_i(\alpha)| \prod_{i=r+1}^{n} |\sigma_i(\alpha)|^2 .$$

In particular, this means that

$$\mathcal{L}(\mathfrak{o}_K^*) \subset \mathbb{R}^{r+s,0} := \Big\{(y_i) \in \mathbb{R}^{r+s} \;\Big|\; \sum_i y_i = 0\Big\}.$$

and that $\mathcal{L}(\zeta) = 0$ if ζ is a root of unity.

Proposition 11.8. *(i)* $\ker \mathcal{L} \cap \mathfrak{o}_K^*$ *is the group of roots of unity of K.*
(ii) $\mathcal{L}(\mathfrak{o}_K^*)$ *is a discrete subgroup of* $\mathbb{R}^{r+s,0}$ *(cf. remark after Lemma 11.4).*

Proof. Let $M > 0$, and consider the set $Z = \{|y_i| \le M\} \subset \mathbb{R}^{r+s}$. Then $\mathcal{L}(\alpha) \in Z$ iff $e^{-M} \le |\sigma_i(\alpha)| \le e^M$ $(1 \le i \le r)$, $e^{-M/2} \le |\sigma_i(\alpha)| \le e^{M/2}$ $(i > r)$. In particular, $S = \{\alpha \in \mathfrak{o}_K^* \mid \mathcal{L}(\alpha) \in Z\}$ is finite, by Lemma 11.5. As $0 \in Z$, $S \supset \ker \mathcal{L} \cap \mathfrak{o}_K^*$, hence (i). Also S finite implies that $\mathcal{L}(\mathfrak{o}_K^*) \cap Z$ is finite for every M, hence (ii). $\square$

From this it follows that $\mathfrak{o}_K^*$ is finitely generated of rank $\le r+s-1 = \dim \mathbb{R}^{r+s,0}$. The unit theorem follows from:

Theorem 11.9. $\mathcal{L}(\mathfrak{o}_K^*)$ *is a lattice in* $\mathbb{R}^{r+s,0}$.

——— *Lecture 14* ———

By (ii) it's enough to show that there exists $(r+s-1)$ linearly independent units. This is **not examinable**, and follows from a slightly more complicated version of the proof of Theorem 11.2(i):

Proposition 11.10. *There exists a constant C with the following property: let $1 \le j \le r+s$. Then for every $\delta > 0$ there exists $0 \ne \alpha \in \mathfrak{o}_K$ such that*

(i) $\mathrm{N}_{K/\mathbb{Q}}(\alpha) \le C$, *and*

(ii) for every $i \ne j$, $|\sigma_i(\alpha)| \le \delta$.

Sketch proof. Consider the set $Y \subset \mathbb{R}^r \times \mathbb{C}^s$ given by the inequalities

$$|z_i| \le \begin{cases} A\delta^{1-n} & \text{if } i = j \\ \delta & \text{if } 1 \le i \le r+s \text{ and } i \ne j \end{cases}$$

for some $A > 0$. Then $\operatorname{vol} Y = 2^r \pi^s C$, where $C = A$ if $j \le r$, $C = A^2$ if $j \ge r+1$. Applying Minkowski for sufficiently large A, there exists $0 \ne \alpha \in \mathfrak{o}_K$ with $\sigma(\alpha) \in Y$, and (i) and (ii) are easily seen to hold. □

Corollary 11.11. *Let $1 \le j \le r+s$. Then there exists $\varepsilon = \varepsilon_j \in \mathfrak{o}_K^*$ such that $|\sigma_j(\varepsilon)| > 1$ and for every $i \ne j$, $|\sigma_i(\varepsilon)| < 1$.*

Proof. By the Proposition, we may inductively find a sequence of nonzero elements $\alpha_1, \alpha_2, \ldots$ of $\mathfrak{o}_K$ such that $\left|\mathrm{N}_{K/\mathbb{Q}}(\alpha_k)\right| \le C$ and for every $i \ne j$, $|\sigma_i(\alpha_{k+1})| < |\sigma_i(\alpha_k)|$. Then there exist k and ℓ with $k < \ell$ and $(\alpha_k) = (\alpha_\ell)$, and then $\varepsilon = \alpha_\ell/\alpha_k$ is a unit with the desired property. □

Final step in the proof: the elements $\mathcal{L}(\varepsilon_j) \in \mathbb{R}^{r+s,0}$, $1 \le j \le r+s$, span $\mathbb{R}^{r+s,0}$. For this, consider the $(r+s) \times (r+s)$ matrix $(\mathcal{L}(\varepsilon_j)_k)$. The sum of its columns is zero, since $\mathcal{L}(\varepsilon_j) \in \mathbb{R}^{r+s,0}$. So it satisfies the conditions of the following:

Lemma 11.12. *Let $A \in \mathrm{Mat}_{m,m}(\mathbb{R})$ be a matrix such that*

(i) for all $j \ne k$, $A_{jk} < 0$

(ii) for all j, $\sum_j A_{jk} = 0$.

Then A has rank $m-1$.

Proof. Let $\underline{x} \in \mathbb{R}^n$. We show that $A\underline{x} = \underline{0}$ iff $\underline{x}$ is a multiple of $(1, \ldots, 1)$. Then $\mathrm{nullity}(A) = 1$ as required.

By (ii), "if" holds. So suppose $\underline{x} \in \mathbb{R}^n$, and let x_k be its largest coordinate. Then if $A\underline{x} = \underline{0}$,

$$\sum_{j \ne k} A_{kj}(x_k - x_j) = \sum_{j=1}^{m} A_{kj} x_k - \sum_{j=1}^{m} A_{kj} x_j = 0$$

by (ii). But as $x_k \ge x_j$ and $A_{kj} < 0$ for every j, this forces $x_k = x_j$ for every j. □

12 Application to Diophantine equations

A *Diophantine equation* is a polynomial equation in ≥ 2 variables, for which we ask for solutions in $\mathbb{Z}$ (or sometimes in $\mathbb{Q}$). Example: the Fermat equation $x^n+y^n=z^n$.

Algebraic number theory can often help to solve such equations.

Example: Consider the equation $y^2+5=x^3$. Let's try to find all solutions (x,y) in integers.

Elementary consideration tell us that if $(x,y)\in\mathbb{Z}^2$ is a solution, then:

- x is odd, since if not $y^3+5\equiv 0 \pmod 4$ which is impossible
- $(5,x)=1$, since if not, $5|y$ and therefore $25|x^3-y^2=5$, contradiction.

So $(x,10)=1$. To go further, factor both sides in $\mathbb{Z}[\sqrt{-5}]=\mathfrak{o}_K$ where $K=\mathbb{Q}(\sqrt{-5})$.

Careful! $\mathfrak{o}_K$ is not a UFD. So we need to factor in ideals instead. We have:

$$(x)^3=(y+\sqrt{-5})(y-\sqrt{-5}).$$

Suppose P is a prime ideal dividing both $(y+\sqrt{-5})$ and $(y-\sqrt{-5})$. Then P divides $(y+\sqrt{-5})+(y-\sqrt{-5})\supset(2\sqrt{-5})\supset(10)$. But also P divides $(x)^3$ so P divides (x). As $(x,10)=(1)$, no such P exists. Therefore by unique factorisation of ideals, there exist ideals I and J such that

$$(y+\sqrt{-5})=I^3,\quad (y-\sqrt{-5})=J^3,\quad (x)=IJ.$$

Now $\mathrm{Cl}(K)$ has order 2, so as I^3 is principal, so is I, say $I=(a+b\sqrt{-5})$ with a, $b\in\mathbb{Z}$. Therefore $y+\sqrt{-5}=(\text{unit})\times(a+b\sqrt{-5})^3$, and as $\mathfrak{o}_K^*=\{\pm1\}$, this implies (replacing (a,b) by $(-a,-b)$ if necessary)

$$y+\sqrt{-5}=(a+b\sqrt{-5})^3=(a^3-15ab^2)+(3a^2b-5b^3)\sqrt{-5}.$$

Equating coefficients of $\sqrt{-5}$ gives $3a^2b-5b^3=1$, so $b=\pm1$ and $3a^2=5\pm1$ which is impossible. So the original equation has no integer solutions.

Lecture 15

13 Analytic class number formula

Recall the Riemann zeta function:

$$\zeta(s)=\sum_{n=1}^{\infty}\frac{1}{n^s}$$

defined for $s>1$ (in fact converges for $s\in\mathbb{C}$ with $\mathrm{Re}(s)>1$, but we'll only use real values of s here). For $s=1$ it diverges, and in fact:

Proposition 13.1.

$$\lim_{s\to 1+}(s-1)\zeta(s)=1$$

(More precisely, $\zeta(s)$ can be shown to have an analytic continuation with a pole at $s=1$).

Proof. The function x^{-s} is monotone decreasing, hence for any $s>1$

$$\int_n^{n+1}\frac{dx}{x^s}<\frac{1}{n^s}<\int_{n-1}^{n}\frac{dx}{x^s}$$

and summing gives

$$\frac{1}{s-1}=\int_1^\infty\frac{dx}{x^s}<\sum_{n=1}^\infty\frac{1}{n^s}<1+\int_1^\infty\frac{dx}{x^s}=1+\frac{1}{s-1}.$$

Multiplying by $(s-1)$ and taking the limit gives the result. □

Now let K be a number field. Define the *Dedekind zeta function* of K to be the series

$$\zeta_K(s)=\sum_{0\neq I\subset \mathfrak{o}_K}\frac{1}{N(I)^s}.$$

So if $K=\mathbb{Q}$, $\zeta_K(s)=\zeta(s)$ is the Riemann zeta function.

Here's the analytic clas number formula fo quadratic fields.

Theorem 13.2. *Let $K=\mathbb{Q}(\sqrt{d})$ be quadratic. The series $\zeta_K(s)$ converges for $s>1$, and*

$$\lim_{s\to 1+}(s-1)\zeta_K(s)=\begin{cases}\dfrac{2\pi\, h_K}{|d_K|^{1/2}\,w_K} & \text{if } d<0\\[2ex] \dfrac{4\,h_K\log\varepsilon}{|d_K|^{1/2}\,w_K} & \text{if } d>0\end{cases}$$

where $h_K=\#\,\mathrm{Cl}(K)$ is the class number, $w_K=\#\mu(K)$ is the number of roots of unity and (for $d>0$, $K\subset\mathbb{R}$) $\varepsilon>1$ is the fundamental unit of K.

(Note that $w_K=2$ except if $K=\mathbb{Q}(i)$ or $\mathbb{Q}(\sqrt{-3})$.)

Remarks. (i) This result was used by Siegel in 1935 to show that for *quadratic* fields K, h_kR_k is of order $|d_K|^{1/2}$ as $|d_K|\to\infty$. So for imaginary quadratic fields ($R_K=1$), $h_K\to\infty$, whereas for real quadratic fields, $h_K\log\varepsilon\to\infty$. So real quadratic fields of class number 1 (of which there are believed to be infinitely many) will tend to have very large fundamental units. E.g. $K=\mathbb{Q}(\sqrt{3001})$ has class number 1. Its fundamental unit is $\varepsilon=u+v\sqrt{3001}$ where u, $v\in\mathbb{N}$ with $u>4\times 10^{36}$.

(ii) Another application is the following explicit formula for class numbers: let $p \equiv 7 \pmod 8$ be prime. Then $h_{\mathbb{Q}(\sqrt{-p})} = R - N$, where

$$R = \#\left\{ \text{quadratic residues mod } p \text{ in the interval } \left[1, \frac{p-1}{2}\right] \right\}$$
$$N = \#\left\{ \text{quadratic nonresidues mod } p \text{ in the interval } \left[1, \frac{p-1}{2}\right] \right\}$$

In particular this implies that $R > N$, but there seems to be no way to prove this with proving the much stronger statement above.

Proof for imaginary quadratic K. We start by choosing representatives $J_1, \dots J_h$ of the ideal classes of K, $h = h_K$. Then for any ideal I there exist a unique $i \in \{1, \dots h\}$ such that $IJ_i = (\alpha)$ is principal (and necessarily then $\alpha \in J_i$). Conversely, for any i and any $0 \neq \alpha \in J_i$ there exists I with $IJ_i = (\alpha)$, and elements α, $\alpha' \in J_i$ determine the same I iff $\alpha' = \varepsilon\alpha$ for $\varepsilon \in \mathfrak{o}_K^*$. By the multiplicativity of the norm, $N(I) = N(J_j)^{-1} \left|\mathrm{N}_{K/\mathbb{Q}}(\alpha)\right|$.

We now use the fact that K is imaginary quadratic. Then $\mathfrak{o}_K^*$ is finite, so each I corresponds to exactly w_K different α, and therefore

$$\zeta_K(s) = \sum_{j=1}^{h} N(J_i)^s \frac{1}{w_K} \sum_{0 \neq \alpha \in J_i} \frac{1}{\left|\mathrm{N}_{K/\mathbb{Q}}(\alpha)\right|^s}. \qquad (*)$$

Fix $K \subset \mathbb{C}$. Recall from §10 that if I is a nonzero ideal, then $I \subset \mathbb{C} \simeq \mathbb{R}^2$ is a lattice, of covolume $(1/2)N(I)\left|d_K\right|^{1/2}$, and that if $\alpha \in I$ then $\mathrm{N}_{K/\mathbb{Q}}(\alpha) = |\alpha|^2$. We'll show next time the following generalisation of Proposition 13.1:

Theorem 13.3. *Let $\Lambda \subset \mathbb{C}$ be a lattice and*

$$Z(s) = \sum_{0 \neq x \in \Lambda} \frac{1}{|x|^{2s}}.$$

Then $Z(s)$ converges for $s > 1$ and

$$\lim_{s \to 1+} (s-1) Z(s) = \frac{\pi}{\operatorname{covol} \Lambda}.$$

Given this:

$$\lim_{s \to 1+} (s-1) \sum_{0 \neq \alpha \in J_i} \frac{1}{\left|\mathrm{N}_{K/\mathbb{Q}}(\alpha)\right|^s} = \frac{\pi}{(1/2)N(J_i)\left|d_K\right|^{1/2}}$$

and so putting this into $(*)$ gives

$$\lim_{s \to 1+} (s-1)\zeta_K(s) = \frac{2\pi h_K}{w_K \left|d_K\right|^{1/2}}$$

as the norms cancel. □

We'll prove something a bit more general than Theorem 13.3. Let $\Lambda \subset \mathbb{R}^n$ be a lattice, and $C \subset \mathbb{R}^n$ a nonempty closed convex cone. This means that C is a closed convex subset of $\mathbb{R}^n$, and if $x \in C$ then for every $a \geq 0$, $ax \in C$. (So $0 \in C$). [Draw a picture.]

Let $F\colon \mathbb{R}^n \to \mathbb{R}$ be a continuous function, continuously differentiable [1] on $C \setminus \{0\}$. satisfying:

(i) $F(ax) = a^n F(x)$ for every $a \geq 0$, $x \in \mathbb{R}^n$.

(ii) $F(x) > 0$ for all $x \in C \setminus \{0\}$.

For example, $F(x) = \|x\|^n$ satisfies this for any C and $n > 1$.

For $t > 0$, define $C_t = \{x \in C \mid F(x) \leq t^n\}$. By (i), $C_t = tC_1$. The sets C_t are closed and bounded.

Closed is clear as C is closed. For bounded, consider $S = \{x \in C \mid \|x\| = 1\}$. Then S is compact and F is continuous and > 0 on S, hence there exists $c > 0$ with $F \geq c$ on S. But then for any $0 \neq x \in C$, $F(x) = \|x\|^n F(x/\|x\|) \geq c\|x\|^n$. So $x \in C_t \implies \|x\|^n \leq c^{-1}t$.

Let

$$Z(s) = \sum_{0 \neq x \in \Lambda \cap C} \frac{1}{F(x)^s}.$$

Theorem 13.4. *The series $Z(s)$ converges for $s > 1$, and*

$$\lim_{s \to 1+} (s-1)Z(s) = \mu := \frac{\operatorname{vol} C_1}{\operatorname{covol} \Lambda}.$$

Example: $n = 1$, $\Lambda = \mathbb{Z}$, $C = \mathbb{R}_{\geq 0}$. Then $Z(s) = \zeta(s)$ and the theorem reduces to Proposition 13.1. If $n = 2$, $F(x) = |x|^2$ and $C = \mathbb{R}^2$, this is Theorem 13.3.

Proof. First note: if we change the basis of $\mathbb{R}^n$, $Z(s)$ does not change, but $\operatorname{vol} C_1$, $\operatorname{covol} \Lambda$ are both multiplied by the same factor (the absolute value of the determinant of the change-of-basis matrix). So after a change of basis we may assume that $\Lambda = \mathbb{Z}^n$, and then $\mu = \operatorname{vol} C_1$.

Let $N(t) = \#(\mathbb{Z}^n \cap C_t)$. Then $N(t) < \infty$ as C_t is bounded, and $N(t) = \#((1/t)\mathbb{Z}^n \cap C_1)$. Therefore (covering C_1 with cubical boxes of side $1/t$)

$$\mu = \operatorname{vol} C_1 = \lim_{t \to \infty} \frac{N(t)}{t^n}. \tag{13.1}$$

[1] This condition is included to ensure that $\operatorname{vol}(C_1)$ is well-defined, and could be replaced by something weaker. For the F, C we'll use, this is not an issue.

Write $\mathbb{Z}^n \cap C = \{0, x_1, x_2, \dots\}$ where $0 < F(x_1) \le F(x_2) \le \dots$, and put $t_k = F(x_k)^{1/n}$. Then a similar argument shows that

$$\lim_{k\to\infty} \frac{k}{t_k^n} = \mu.$$

In detail: $\mathbb{Z}^n \cap C_{t_k} \supset \{0, t_1, \dots, t_k\}$ but for any $\delta > 0$, $x_k \notin \mathbb{Z}^n \cap C_{t_k - \delta}$. Therefore for every $k \ge 1$, $\delta > 0$,

$$N(t_k) > k \ge N(t_k - \delta)$$

and therefore if $0 < \delta < t_1$

$$\frac{N(t_k)}{t_k^n} > \frac{k}{t_k^n} \ge \frac{N(t_k - \delta)}{(t_k - \delta)^n} \left(\frac{t_k - \delta}{t_k}\right)^n$$

and letting $k \to \infty$ and using (13.1),

$$\lim_{k\to\infty} \frac{k}{t_k^n} = \mu.$$

Now $Z(s) = \sum_{k\ge 1} t_k^{-ns}$. By the previous equation, for every $\epsilon > 0$ there exists k_ϵ such that

$$(\mu - \epsilon)^s \frac{1}{k^s} < \frac{1}{t_k^{ns}} < (\mu + \epsilon)^s \frac{1}{k^s} \qquad \text{if } k \ge k_\epsilon$$

and therefore

$$(\mu - \epsilon)^s \sum_{k \ge k_\epsilon} \frac{1}{k^s} < \sum_{k \ge k_\epsilon} \frac{1}{t_k^{ns}} < (\mu + \epsilon)^s \sum_{k \ge k_\epsilon} \frac{1}{k^s}.$$

By Proposition 13.1 this shows that the series for $Z(s)$ converges for $s > 1$, and it's an exercise in real analysis to deduce (also by 13.1) that

$$\lim_{s\to 1+} (s-1) Z(s) = \mu.$$

In detail:

$$\limsup_{s\to 1+} (s-1) \sum_{k \ge k_\epsilon} \frac{1}{t_k^{ns}} \le \limsup_{s\to 1+} (\mu + \epsilon)^s (s-1) \sum_{k \ge k_\epsilon} \frac{1}{k^s} = \mu + \epsilon.$$

Likewise

$$\liminf_{s\to 1+} (s-1) \sum_{k \ge k_\epsilon} \frac{1}{t_k^{ns}} \ge \mu - \epsilon.$$

Therefore, since $(s-1) \sum_{k < k_\epsilon} t_k^{-ns} \to 0$ as $s \to 1$,

$$\mu - \epsilon \le \liminf_{s\to 1+} (s-1) Z(s) \le \limsup_{s\to 1+} (s-1) Z(s) \le \mu + \epsilon$$

and as this holds for every $\epsilon > 0$, we are done. $\square$

Proof of ACNF for real quadratic fields. We fix $K = \mathbb{Q}(\sqrt{d}) \subset \mathbb{R}$, and then have the embedding $\sigma\colon K \hookrightarrow \mathbb{R}^2$, $u + v\sqrt{d} \mapsto (u + v\sqrt{d}, u - v\sqrt{d})$. Let $\varepsilon > 1$ be a fundamental unit. Suppose $0 \neq \alpha \in \mathfrak{o}_K$, and let $\sigma(\alpha) = (\alpha, \alpha')$. Then if $m \in \mathbb{Z}$ and $\beta = \varepsilon^m \alpha$, $|\beta/\beta'| = |\varepsilon/\varepsilon'|^m |\alpha/\alpha'| = \varepsilon^{2m} |\alpha/\alpha'|$, since $\varepsilon\varepsilon' = \mathrm{N}_{K/\mathbb{Q}}(\varepsilon) = \pm 1$. This shows that if J is a principal ideal, then there exists α with $J = (\alpha)$ and $1/\varepsilon < |\alpha/\alpha'| \leq \varepsilon$, and α is then uniquely determined up to sign. So in place of $(*)$ we have:

$$\zeta_K(s) = \sum_{j=1}^{h} N(J_i)^s \frac{1}{2} \sum_{\substack{0 \neq \alpha \in J_i \\ \varepsilon^{-1} < |\alpha/\alpha'| \leq \varepsilon}} \frac{1}{|\mathrm{N}_{K/\mathbb{Q}}(\alpha)|^s} \qquad (**)$$

So let's consider the set

$$D = \{(x_1, x_2) \in \mathbb{R}^2 \mid \varepsilon^{-1} \leq x_1/x_2 \leq \varepsilon\}$$

which is the union of 4 cones: one is $C = D \cap \mathbb{R}^2_{\geq 0}$, and the other 3 are the reflections of C in the coordinate axes (draw a picture). The inner series in $(**)$ is almost the same as

$$\sum_{0 \neq \alpha \in J_i \cap D} \frac{1}{|\mathrm{N}_{K/\mathbb{Q}}(\alpha)|^s} = 4 \sum_{0 \neq \alpha \in J_i \cap C} \frac{1}{|\mathrm{N}_{K/\mathbb{Q}}(\alpha)|^s} \qquad (1)$$

In fact the two series differ by

$$\sum_{\substack{0 \neq \alpha \in J_i \\ \varepsilon^{-1} = |\alpha/\alpha'|}} \frac{1}{|\mathrm{N}_{K/\mathbb{Q}}(\alpha)|^s} \qquad (2)$$

and it is an exercise to show that this series converges for $s = 1$. In detail: $\alpha = \varepsilon\alpha'$ implies that α and α' generate the same ideal. If $(p) = PP'$ is split in K, then this means that if $P^a|(\alpha)$, then also $(P')^a|(\alpha)$. By Theorem 9.3, this implies that $\mathrm{N}_{K/\mathbb{Q}}(\alpha) = \pm M^2 L$, where M and L are integers with L divisible only by primes dividing d_K. So (2) is bounded by

$$\sum_{\substack{M \geq 1 \\ L | d_K^R,\ \text{some } R}} \frac{1}{M^{2s} L^s} = \sum_{M \geq 1} \frac{1}{M^2 s} \prod_{p | d_K} \frac{1}{1 - p^{-s}}$$

which converges for $s > 1/2$.

Now take $F(\underline{x}) = |x_1 x_2|$, so $F(\sigma(\alpha)) = |\mathrm{N}_{K/\mathbb{Q}}(\alpha)|$. Conditions (i) and (ii) are satisfied for the cone C, and calculus gives $\operatorname{vol} C_1 = \log \varepsilon$. Therefore the series (1) converges for $s > 1$, and

$$\lim_{s \to 1+} (s-1) \sum_{0 \neq \alpha \in J_i \cap D} \frac{1}{|\mathrm{N}_{K/\mathbb{Q}}(\alpha)|^s} = 4 \frac{\operatorname{vol} C_1}{\operatorname{covol} J_i} = 4 \frac{\log \varepsilon}{N(J_i) d_K^{1/2}}$$

so plugging into $(**)$ gives

$$\lim_{s\to 1+}(s-1)\zeta_K(s) = 2\frac{h_K \log\varepsilon}{d_K^{1/2}}$$

which is the analytic class number formula for $(r_1, r_2) = (2, 0)$ since $w_K = 2$. $\square$

Theorem 13.5. *The series for $\zeta_K(s)$ converges for $s > 1$, and*

$$\lim_{s\to 1+}(s-1)\zeta_K(s) = 2^{r_1+r_2}\pi^{r_2}\left|d_K\right|^{-1/2}\frac{h_K R_K}{w_K}.$$

Here the undefined terms are:

- (r_1, r_2) = what we called (r, s) up till now;
- $h_K = \#\,\mathrm{Cl}(K)$, the *class number* of K;
- w_K is the number of roots of unity in K; and
- R_K is the *regulator* of K, defined to be

  $$R_K = \left|\det(\mathcal{L}(\varepsilon_j)_k)_{1\le j,k\le r_1+r_2-1}\right|$$

  where $\varepsilon_1, \ldots \varepsilon_{r_1+r_2-1}$ are generators for the torsion free part of $\mathfrak{o}_K^*$. ($R_K = 1$ if $K = \mathbb{Q}$ or $\mathbb{Q}(\sqrt{-d})$.)

For example, if K is real quadratic, $R_K = \log\varepsilon$ where $\varepsilon > 1$ is a fundamental unit. For general K the same proof works, except the cone C becomes a bit more complicated.