Number Fields

March 1, 2018

Contents

-1 Miscellaneous

Book: Number Fields, Marcus

Course notes: www.dpmms.ac.uk/ jat58/nfl2018

0 Motivation

Theorem. If p is an odd prime, then $p = a^2 + b^2$ for $a, b \in \mathbb{Z} \iff p \equiv 1 \pmod 4$.

Proof. If $p = a^2 + b^2$, then $p \equiv 0, 1, 2 \pmod 4$. So this condition on p is necessary.
Suppose instead $p \equiv 1 \pmod 4$. Then $\left(\frac{-1}{p}\right) = 1$. Thus $\exists a \in \mathbb{Z}$ such that $a^2 \equiv -1 \pmod p$, or $p | a^2 + 1$. We can factor $a^2 + 1 = (a + i)(a - i)$ in the ring $\mathbb{Z}[i]$. Here we introduce a notation: if $R \subseteq S$ are rings and $\alpha \in S$, then

$$R[\alpha] = \{\sum_{i=0}^{n} a_i \alpha^i \in S | a_i \in R\}$$

, the smallest subring of S containing both R and α.

We know from IB GRM that $\mathbb{Z}[i]$ is a UFD. Now $p|(a+i)(a-i)$. If p is irreducible in $\mathbb{Z}[i]$ then $p|a + i$ or $p|a - i$, contradiction. Thus p is reducible in $\mathbb{Z}[i]$, hence $p = z_1 z_2$ with $z_1, z_2 \in \mathbb{Z}[i]$. If $z_1 = A + Bi$, $A, B \in \mathbb{Z}$, then $A^2 + B^2 = p$. $\square$

Another example is when p is an odd prime. Does the equation

$$x^p + y^p = z^p$$

have solutions with $x, y, z \in \mathbb{Z}$ and $xyz \neq 0$?

Theorem. (Kummer, 1850)
If $\mathbb{Z}[e^{2\pi i/p}]$ is a UFD, then there are no solutions.
Strategy: factor $x^p + y^p = \prod_{j=0}^{p-1}(x + e^{2\pi ij/p}y)$ in $\mathbb{Z}[e^{2\pi i/p}]$.

However, we now know $\mathbb{Z}[e^{2\pi i/p}]$ is a UFD $\iff p \leq 19$.

Theorem. (Kummer, 1850)
If p is a *regular* prime, then there are no solutions.
If $p < 100$, then p is regular $\iff p \neq 37, 59, 67$.

We have seen various examples such as $\mathbb{Z} \subseteq \mathbb{Q}$, $\mathbb{Z}[i] \subseteq \mathbb{Q}[i]$, $\mathbb{Z}[e^{2\pi i/p}] \subseteq \mathbb{Q}[e^{2\pi i/p}]$, or in general, $\mathcal{O}_L \subseteq L$, where a ring of "integers" lies in a number field.

1 Ring of integers

Recall: A field extension L/K is an inclusion $K \leq L$ of fields. The degree of L/K is $[L : K] = \dim_K L$. We say L/K is finite if $[L : K] < \infty$.

Definition. (1.1)
A number field is a finite extension $L/\mathbb{Q}$. Here are two ways to construct number fields:
(1) Let $\alpha \in \mathbb{C}$ be an algebraic number. Then $L = \mathbb{Q}(\alpha)$ is a number field;
(2) Let K be a number field, and let $f(X) \in K[X]$ be an irreducible polynomial. Then $L = K[X]/(f(X))$ is a number field.
(Recall Tower Law: $[L : Q] = [L : K][K : Q] < \infty$).

Definition. (1.2)
(1) Let L/K be a field extension. Then we say $\alpha \in L$ is algebraic over K if there exists a monic $f(X) \in K[X]$ such that $f(\alpha) = 0$;
(2) Let $L/\mathbb{Q}$ be a field extension. Then we say $\alpha \in L$ is an algebraic integer if there exists a monic $f(X) \in Z[X]$ such that $f(\alpha) = 0$.

Definition. (1.3)
Let L/K be a field extension, and let $\alpha \in L$ be algebraic over K. We call the minimal polynomial of α over K the monic polynomial $f_\alpha(X) \in K[X]$ of least degree such that $f_\alpha(\alpha) = 0$.

We recall why $f_\alpha(X)$ is well-defined: there exists some monic $f(X) \in K[X]$ with $f(\alpha) = 0$ as α is algebraic. If $f_\alpha(\alpha), f'_\alpha(\alpha) \in K[X]$ both satisfy the definition of minimal polynomial, then we apply the polynomial division algorithm to write

$$f_\alpha(X) = p(X)f'_\alpha(X) + r(X)$$

where $p(X), r(X) \in K[X]$, and $\deg r < \deg f'_\alpha$. Evaluate at $X = \alpha$, we have $0 = f_\alpha(\alpha) = p(\alpha)f'_\alpha(\alpha) + r(\alpha) = r(\alpha)$. By minimality of $\deg f'_\alpha$, we must have $r = 0$. Then $\deg f_\alpha = \deg f'_\alpha$, and $f_\alpha(X), f'(\alpha)$ are both monic, i.e. $p(X) = 1$ and $f_\alpha(X) = f'_\alpha(X)$.

Lemma. (1.4)
Let $L/\mathbb{Q}$ be a field extension, and let $\alpha \in L$ be an algebraic integer. Then:
(1) The minimal polynomial $f_\alpha(X)$ of α over $\mathbb{Q}$ lies in $\mathbb{Z}[X]$;
(2) If $g(X) \in \mathbb{Z}[X]$ satisfies $g(\alpha) = 0$, then there exists $q(X) \in \mathbb{Z}[X]$ such that $g(X) = f_\alpha(X)q(X)$;
(3) The kernel of the ring homomorphism $\mathbb{Z}[X] \to L$ by $f(X) \to f(\alpha)$ equals $(f_\alpha(X))$, the ideal generated by $f_\alpha(X)$.

Proof. (1) Recall that if $f(X) = a_nX^n + ... + a_0 \in \mathbb{Z}[X]$, then we define from GRM, the content $c(f) = \gcd(a_n, ..., a_0)$. Recall Gauss' Lemma: If $f(X), g(X) \in \mathbb{Z}[X]$, then $c(fg) = c(f)c(g)$. Since $\alpha \in L$ is an algebraic integer, there exists monic $f(X) \in \mathbb{Z}[X]$ such that $f(\alpha) = 0$, i.e. $c(f) = 1$. Apply polynomial division in $\mathbb{Q}[X]$ to get $f(X) = p(X)f_\alpha(X) + r(X)$, where $p(X), r(X) \in \mathbb{Q}[X]$, $\deg r < \deg f_\alpha$. The definition of $f_\alpha(X)$ implies that $r(X) = 0$, hence $f(X) = p(X)f_\alpha(X)$. Now choose integers $n, m \geq 1$ such that $np(X) \in \mathbb{Z}[X]$, $c(np) = 1$, and $mf_\alpha(X) \in$

$\mathbb{Z}[x]$, $c(mf_\alpha) = 1$. Then $nmf(x) = (np(x))(mf_\alpha(x)) \implies c(nmf(x)) = nm = 1$. So $n = m = 1$, hence $f_\alpha(x) \in \mathbb{Z}[X]$.
(2) Let $g(X) \in \mathbb{Z}[X]$ be such that $g(\alpha) = 0$. WLOG $g(x) \neq 0$ and $c(g) = 1$. Now apply polynomial division to write $g(x) = q(x)f_\alpha(x) + s(x)$ where $q(x), s(x) \in \mathbb{Q}[x]$, $\deg s < \deg f_\alpha$. Again by definition we have $s(x) = 0$. Choose an integer $k \geq 1$ such that $kq(x) \in Z[x]$ and $c(kq) = 1$. Then $kg(x) = kq(x)f_\alpha(x) \implies k = c(kg) = c(kq)c(f_\alpha) = 1$. So $k = 1$, hence $q(x) \in \mathbb{Z}[x]$.
(3) is a reformulation of (2). $\square$

Let $L/\mathbb{Q}$ be a field extension. Last time we said $\alpha \in L$ is an algebraic integer if $\exists$ monic polynomial $f(x) \in \mathbb{Z}[x]$ such that $f(\alpha) = 0$. We proved that if $\alpha \in L$ is an algebraic integer and $f_\alpha(x) \in \mathbb{Q}[x]$ is the minimal polynomial of α over $\mathbb{Q}$, then $f_\alpha(x) \in \mathbb{Z}[x]$. However there is a small problem, so we'll prove again.

Proof. Choose $f(x) \in \mathbb{Z}[x]$ monic with $f(\alpha) = 0$, and write

$$f(x) = q(x)f_\alpha(x) + r(x)$$

where $q(x), r(x) \in \mathbb{Q}[x]$, $\deg r < \deg f_\alpha$. Then $r(\alpha) = 0 \implies r(x) = 0$, by minimality of $\deg f_\alpha$. I said that we can find integer $n, m \geq 1$ s.t. $nf\alpha(x) \in \mathbb{Z}[x]$, $c(nf\alpha) = 1$, $mq(x) \in \mathbb{Z}[x]$, $c(mq) = 1$. However we need to explain why do they exist. Note $f_\alpha(x)$ and $q(x)$ are both monic. Choose integers $N, M \geq 1$ such that $Nf_\alpha(x) \in \mathbb{Z}[x]$, $Mq(x) \in \mathbb{Z}[x]$. Then $c(Nf_\alpha)|N$, $c(Mq)|M$ as those are the leading term of the polynomial. Now let $N/c(Nf\alpha) = n \in \mathbb{Z}$, $M/c(Mq) = m \in \mathbb{Z}$. Now $nmf(x) = (nf\alpha(x))(mq(x))$, so $c(nmf(x)) = nm = 1 \implies n = m = 1$. $\square$

Corollary. (1.5)
If $\alpha \in \mathbb{Q}$, then α is an algebraic integer $\iff \alpha \in \mathbb{Z}$.

Proof. By lemma 1.4, α is an algebraic integer $\iff f_\alpha(x) \in \mathbb{Z}[x]$. But if $\alpha \in \mathbb{Q}$, then $f_\alpha(x) = x - \alpha$, and the first needs to divide the second polynomial. $\square$

Notation. If $L/\mathbb{Q}$ is any field extension, we write $\mathcal{O}_L = \{\alpha \in L | \alpha \text{ is an algebraic integer}\}$.

Now we proceed to the first non-trivial result of the course:

Proposition. (1.6)
If $L/\mathbb{Q}$ is a field extension, $\mathcal{O}_L$ is a ring.

Proof. Clearly $0, 1 \in \mathcal{O}_L$. Now if $\alpha \in \mathcal{O}_L$, then $f_{-\alpha}(x) = (-1)^{\deg f_\alpha} f_\alpha(-x) \implies -\alpha \in \mathcal{O}_L$.
The hard part is to show that if $\alpha, \beta \in \mathcal{O}_L$, then $\alpha + \beta \in \mathcal{O}_L$ and $\alpha\beta \in \mathcal{O}_L$.
Observe that if $\alpha \in \mathcal{O}_L$, then $\mathbb{Z}[\alpha] \subseteq L$ is a finitely generated $\mathbb{Z}$-module. By definition, $\mathbb{Z}[\alpha]$ is generated by $1, \alpha, \alpha^2, \alpha^3, \ldots$. Let $f_\alpha(x) = x^d + a_1x^{d-1} + \ldots + ad$, $a_i \in \mathbb{Z}$. Then $\alpha^d = -(a_1\alpha^{d-1} + \ldots + ad)$, so $\alpha^d \in \sum_{i=0}^{d-1} \mathbb{Z}\alpha^i$. By induction, we see that $\alpha^n \in \sum_{i=0}^{d-1} \mathbb{Z}\alpha^i$ for all $n \geq d$. Hence $\mathbb{Z}[\alpha] = \sum_{i=0}^{d-1} \mathbb{Z}\alpha^i$. Now take $\alpha, \beta \in \mathcal{O}_L$ and let $d = \deg f_\alpha$, $e = \deg f_\beta$.

By definition, $\mathbb{Z}[\alpha,\beta] = \mathbb{Z}[\alpha][\beta]$ is generated as a $\mathbb{Z}$-module by $\{\alpha^i\beta^j\}_{i,j\in\mathbb{N}}$. The same argument show that in fact this ring is generated as a $\mathbb{Z}$-module by $\{\alpha^i\beta^j\}$ for $0 \le i \le d-1, 0 \le j \le e-1$. So $\mathbb{Z}[\alpha,\beta]$ is finitely generated. From GRM we know the classification of finitely generated $\mathbb{Z}$-modules implies that there's an isomorphism $\mathbb{Z}[\alpha,\beta] \cong \mathbb{Z}^r \oplus T$ for some $r \ge 1$ and finite abelian group T. In fact, $T = 0$: if $\gamma \in T$, then $|T|\gamma = 0$, by Lagrange's theorem. But $\mathbb{Z}[\alpha,\beta] \subseteq L$, a $\mathbb{Q}$-vector space, so this forces $\gamma = 0$. Now we can therefore fix an isomorphism $\mathbb{Z}[\alpha,\beta] \cong \mathbb{Z}^r$ ($r \ge 1$. There's an endomorphism $m_{\alpha\beta} : \mathbb{Z}[\alpha,\beta] \to \mathbb{Z}[\alpha,\beta]$ by $\gamma \to \alpha\beta\gamma$ (as a $\mathbb{Z}$-module). $m_{\alpha\beta}$ corredponds to an $r\times r$ matrx $A_{\alpha\beta} \in M_{r\times r}(\mathbb{Z})$. Let $F_{\alpha\beta}(x) = \det(x\cdot 1_r - A_{\alpha\beta}) \in \mathbb{Z}[x]$, a monic polynomial. By the Cayley-Hamilton theorem, $F_{\alpha\beta}(m_{\alpha\beta}) = 0$ as endomorphisms of $\mathbb{Z}[\alpha,\beta]$. Write $F_{\alpha\beta}(x) = x^r + b_1x^{r-1} + ... + b_r$ for $b_i \in \mathbb{Z}$. Thus $m_{\alpha\beta}^r + b_1 m_{\alpha\beta}^{r-1} + ... + b_r \cdot 1_r = 0$ as endomorphisms of $\mathbb{Z}[\alpha,\beta]$.
Now the image of 1 is $(\alpha\beta)^r + b_1(\alpha\beta)^{r-1} + ... + b_r = F_{\alpha\beta}(\alpha\beta) = 0$. So $\alpha\beta \in \mathcal{O}_L$. The argument to show $\alpha+\beta \in \mathcal{O}_L$ is identical, replacing $m_{\alpha\beta}$ by $m_{\alpha+\beta} : \mathbb{Z}[\alpha,\beta] \to \mathbb{Z}[\alpha,\beta]$ by $\gamma \to (\alpha+\beta)\gamma$. The detail is omitted here. □

We call $\mathcal{O}_L$ the ring of algebraic integers of L.

Lemma. (1.7)
Let $L/\mathbb{Q}$ be a number field, and let $\alpha \in L$. Then $\exists n \ge 1$ an integer such that $n\alpha \in \mathcal{O}_L$.

Proof. Let $f(x) \in \mathbb{Q}[x]$ be a monic polynomial such that $f(\alpha) = 0$. Then $\exists n \in \mathbb{Z}, n \ge 1$ such that $g(x) = n^{\deg f} f(x/n) \in \mathbb{Z}[x]$ is monic. But then $g(n\alpha) = n^{\deg f} f(\alpha) = 0$. So $n\alpha \in \mathcal{O}_L$. □

2 Complex embeddings

Let L be a number field.

Definition. (2.1)
A *complex embedding* of L is a field homomorphism $\sigma : L \to \mathbb{C}$. Note: in this case, σ is injective, and $\sigma|_{\mathbb{Q}}$ is the usual embedding $\mathbb{Q} \to \mathbb{C}$.

Proposition. (2.2)
Let L/K be an extension of number fields, and let $\sigma_0 : K \to \mathbb{C}$ be a complex embedding. Then there exist exactly $[L : K]$ embeddings $\sigma : L \to \mathbb{C}$ which extends σ_0 $(\sigma|_K = \sigma_0)$.

Proof. Induction on $[L : K]$. If $[L : K] = 1$, then $L = K$, so σ_0 determines σ. In general, choose $\alpha \in L - K$ and consider $L/K(\alpha)/K$. By the Tower law, $[L : K] = [L : K(\alpha)][K(\alpha) : K]$ and $[K(\alpha) : K] > 1$. By induction, it's enough to show there are exactly $[K(\alpha) : K]$ embeddings $\sigma : K(\alpha) \to \mathbb{C}$ extending σ_0. Let $f_\alpha(x) \in K[x]$ be the minimal polynomial of α over K. Observe there's an isomorphism $K[x]/(f_\alpha(x)) \to K(\alpha)$ by sending $x \to \alpha$. To give a complex embedding $\sigma : K(\alpha) \to \mathbb{C}$ extending σ_0, it's equivalent to give a root β of $(\sigma_0 f)(x)$ in $\mathbb{C}$ ($\sigma_0 f(x) \in \mathbb{C}[x]$ means apply σ_0 to the coefficients of $f(x)$). Dictionary: $\sigma \to \beta = \sigma(\alpha)$. We have $[K(\alpha) : K] = \deg f_\alpha = \deg \sigma_0 f_\alpha$. It's enough to show $\sigma_0 f_\alpha$ has distinct roots in $\mathbb{C}$. The polynomial $f_\alpha(x) \in K[x]$ is irreducible, so is prime to its derivative $f'_\alpha(x)$ (*char* $K = 0$). So α is separable over K. $\square$

Recall from last lecture, let L be a number field, a complex embedding is a field homomorphism $\sigma : L \to \mathbb{C}$. The number of such embeddings is $[L : \mathbb{Q}]$. If $L = \mathbb{Q}(\alpha)$, and $f_\alpha(x) \in \mathbb{Q}[x]$ is the minimal polynomial, then there is a bijection $\{\sigma : L \to \mathbb{C}\} \leftrightarrow \{$ roots $\beta \in \mathbb{C}$ of $f_\alpha(x)\}$ by sending $\sigma \to \beta = \sigma(alpha)$.

Notation: if $\sigma : L \to \mathbb{C}$ is a complex embedding, then $\bar{\sigma} : L \to \mathbb{C}$ is also a complex embedding, where $\bar{\sigma}(\alpha) = \overline{\sigma(\alpha)}$ (complex conjugation). If $\sigma = \bar{\sigma}$, then $\sigma(L) \subseteq \mathbb{R}$. Otherwise $\sigma \neq \bar{\sigma}$ and $\sigma(L) \nsubseteq \mathbb{R}$.

We write r for the number of complex embedding σ such that $\sigma = \bar{\sigma}$, s for the number of pairs of embeddings $\{\sigma, \bar{\sigma}\}$ where $\sigma \neq \bar{\sigma}$. Then $r + 2s = [L : \mathbb{Q}]$.

Example. Let $d \in \mathbb{Z}$ be square-free, $d \neq 0, 1$. Let $\mathbb{Q}(\sqrt{d}) = \mathbb{Q}[x]/(x^2 - d)$. If $d > 0$, then $r = 2, s = 0$ (real quadratic field).
If $d < 0$, then $r = 0, s = 1$ (imaginary quadratic field).

Example. Let $m \in \mathbb{Z}$ cube-free, $m \neq 0, 1, -1$. Let $\mathbb{Q}(\sqrt[3]{m}) = \mathbb{Q}[x]/(x^3 - m)$. Then $r = 1, s = 1$, since $x^3 - m$ has one real and two complex roots.

Definition. (2.3)
Let L/K be an extension of number fields, and let $\alpha \in L$. Let $m_\alpha : L \to L$ be the K-linear map defined by $m_\alpha(\beta) = \alpha\beta$. Then we define

$$\operatorname{tr}_{L/K}(\alpha) = \operatorname{tr} m_\alpha \in K$$
$$N_{L/K}(\alpha) = \det m_\alpha \in K$$

the trace and norm of α respectively.

Lemma. (2.4)
If L/K is an extension of number fields and $\alpha \in L$, then

$$\begin{aligned} \operatorname{tr}_{L/K}(\alpha) &= [L : K(\alpha)] \operatorname{tr}_{K(\alpha)/K}(\alpha) \\ N_{L/K}(\alpha) &= N_{K(\alpha)/K}(\alpha)^{[L:K(\alpha)]} \end{aligned}$$

Proof. There's an isomorphism $L \cong K(\alpha)^{[L:K(\alpha)]}$ of $K(\alpha)$-vector spaces(?). □

Lemma. (2.5)
Let L/K be an extension of number fields and let $\alpha \in L$. Let $\sigma_0 : K \to \mathbb{C}$ be a complex embedding, and let $\sigma_1, ..., \sigma_n : L \to \mathbb{C}$ be the embeddings of L extending σ_0.
Then

$$\begin{aligned} \sigma_0(\operatorname{tr}_{L/K}(\alpha)) &= \sigma_1(\alpha) + ... + \sigma_n(\alpha) \\ \sigma_0(N_{L/K}(\alpha)) &= \sigma_1(\alpha)...\sigma_n(\alpha). \end{aligned}$$

Proof. WLOG let $L = K(\alpha)$. Let $f_\alpha(x) \in K[x]$ be the minimal polynomial of α over K. Then

$$(\sigma_0 f_\alpha)(x) = (x - \sigma_1(\alpha))(x - \sigma_2(\alpha))...(x - \sigma_n(\alpha))$$

If $f(\alpha) = x^n + a_1 x^{n-1} + ... + a_n$, then $\sigma_0(a_1) = -(\sigma_1(\alpha) + ... + \sigma_n(\alpha))$, $\sigma_0(a_n) = (-1)^n \sigma_1(\alpha)...\sigma_n(\alpha)$.
Let $g(x) \in K[x]$ be the characteristic polynomial of m_α. If $g(x) = x^n + b_1 x^{n-1} + ... + b_n$, then $b_1 = -\operatorname{tr} m_\alpha = -\operatorname{tr}_{L/K}(\alpha)$, $b_n = (-1)^n \det m_\alpha = (-1)^n N_{L/K}(\alpha)$.By Cayley-Hamilton, $g(m_\alpha) = 0 \implies g(\alpha) = 0 \implies f_\alpha(x) = g(x)$. □

Corollary. (2.6)
If $\alpha \in \mathcal{O}_L$, then $\operatorname{tr}_{L/K}(\alpha)$, $N_{L/K}(\alpha) \in \mathcal{O}_K$.

Proof. If $\beta \in K$ then $\beta \in \mathcal{O}_K \iff \sigma_0(\beta) \in \mathcal{O}_\mathbb{C}$ (as $\forall f(x) \in \mathbb{Z}[x], f(\beta) = 0 \iff f(\sigma_0(\beta)) = 0$).
By the lemma, $\sigma_0 \operatorname{tr}_{L/K}(\alpha) = \sigma_1(\alpha) + ... + \sigma_n(\alpha)$. If $\alpha \in \mathcal{O}_L$, then $\sigma_1(\alpha), ..., \sigma_n(\alpha) \in \mathcal{O}_\mathbb{C} \implies \sigma_1(\alpha) + ... + \sigma_n(\alpha) \in \mathcal{O}_\mathbb{C} \implies \sigma_0 \operatorname{tr}_{L/K}(\alpha) \in \mathcal{O}_\mathbb{C} \implies \operatorname{tr}_{L/K}(\alpha) \in \mathcal{O}_K$.

The same argument works for the norm. □

Proposition. (2.7)
Let $d \in \mathbb{Z}$ be squarefree, $d \neq 0, 1$, and let $L = \mathbb{Q}(\sqrt{d})$. Then

$$\mathcal{O}_L = \begin{cases} \mathbb{Z}[\sqrt{d}] & d \equiv 2, 3 \pmod 4 \\ \mathbb{Z}[\frac{1+\sqrt{d}}{2}] & d \equiv 1 \pmod 4 \end{cases}$$

Proof. If $\alpha \in L$, then $\alpha \in \mathcal{O}_L$ if and only if both trace and norm (over $L/\mathbb{Q}$) of α is in $\mathbb{Z}$. Why? Forward direction is the previous corollary; if $\alpha \in L$, then $f(\alpha) = 0$, where $f(x) = (x - \sigma_1(\alpha))(x - \sigma_2(\alpha)) = x^2 - \operatorname{tr}_{L/\mathbb{Q}}(\alpha)x + N_{L/\mathbb{Q}}(\alpha) \in \mathbb{Q}[x]$, where σ_1, σ_2 are complex embeddings of L. So backward holds too.

Let $\alpha \in L$. Write $\alpha = \frac{u}{2} + \frac{v}{2}\sqrt{d}$ where $u, v \in \mathbb{Q}$. If $\alpha \in \mathcal{O}_L$, then $\mathrm{tr}_{L/\mathbb{Q}}(\alpha) = u \in \mathbb{Z}$, and $N_{L/\mathbb{Q}}(\alpha) = \frac{1}{4}(u + \sqrt{d}v)(u - \sqrt{d}v) = \frac{1}{4}(u^2 - dv^2) \in \mathbb{Z} \implies u^2 - dv^2 \in 4\mathbb{Z}$ $\implies dv^2 \in \mathbb{Z}$.
Write $v = \frac{r}{s}$ where $r, s \in \mathbb{Z}, s \neq 0, (r, s) = 1$. Then we get $dr^2 \in s^2\mathbb{Z} \implies s^2 | dr^2$. If p is a prime and $p|s$ then $p^2|d$. But we assumed d is square-free. So $s = 1$, so $v \in \mathbb{Z}$.

We've shown if $\alpha \in \mathcal{O}_L$, then $\alpha = \frac{u}{2} + \frac{v}{2}\sqrt{d}$ where $u, v \in \mathbb{Z}$ and $u^2 \equiv d^2 \pmod 4$.

Case 1: $d \equiv 2, 3 \pmod 4$. Then $u^2, v^2 \equiv 0, 1 \pmod 4$. Considering the congruence $u^2 \equiv dv^2 \pmod 4$ shows that both $u, v \in 2\mathbb{Z}$. Hence $\alpha \in \mathbb{Z}[\sqrt{d}] = \{a + b\sqrt{d} | a, b \in \mathbb{Z}\}$, and $\mathcal{O}_L = \mathbb{Z}[\sqrt{d}]$.

Case 2: $d \equiv 1 \pmod 4$. Hence $u^2 \equiv v^2 \pmod 4$, so $u \equiv v \pmod 2$. Hence $\mathcal{O}_L \subseteq \{\frac{u}{2} + \frac{v}{2}\sqrt{d} | u, v \in \mathbb{Z}, u \equiv 1 \pmod 2\} = \mathbb{Z} \oplus \mathbb{Z}(\frac{1+\sqrt{d}}{2})$. It remains to show that $\frac{1+\sqrt{d}}{2}$ is an algebraic integer.
We have $\mathrm{tr}_{L/\mathbb{Q}}(\frac{1+\sqrt{d}}{2}) = 1$, $N_{L/\mathbb{Q}}(\frac{1+\sqrt{d}}{2}) = \frac{1-d}{4} \in \mathbb{Z}$. □

Recall that if R is a ring, then a unit in R is an element $u \in R$ such that there exists $v \in R$ such that $uv = 1$.

The set $\mathbb{R}^* = \{u \in R | u \text{ is a unit}\}$ forms a group under multiplication.

Lemma. (2.8)
If L is a number field, then the units in $\mathcal{O}_L$ are $\mathcal{O}_L^* = \{\alpha \in \mathcal{O}_L | N_{L/\mathbb{Q}}(\alpha) = \pm 1\}$.

Proof. next time.

It's next time now! Let's prove this lemma.
$N_{L/\mathbb{Q}}(\alpha\beta) = N_{L/\mathbb{Q}}(\alpha)N_{L/\mathbb{Q}}(\beta)$ for any $\alpha, \beta \in L$.
If $\alpha \in \mathcal{O}_L^*$, then $\exists \beta \in \mathcal{O}_L$ such that $\alpha\beta = 1 \implies N_{L/\mathbb{Q}}(\alpha)N_{L/\mathbb{Q}}(\beta) = 1$. Since $N_{L/\mathbb{Q}}(\alpha), N_{L/\mathbb{Q}}(\beta) \in \mathbb{Z}$, we get $N_{L/\mathbb{Q}}(\alpha) \in \{\pm 1\}$.
Conversely, suppose $\alpha \in \mathcal{O}_L$ and $N_{L/\mathbb{Q}}(\alpha) = \pm 1$. Then $\alpha^{-1} \in L$. Let $\sigma_1, ..., \sigma_n : L \to \mathbb{C}$ be the distinct complex embeddings of L. Then

$$\begin{aligned} N_{L/\mathbb{Q}}(\alpha) = \sigma_1(\alpha)...\sigma_n(\alpha) = \pm 1 \\ \implies \sigma_1(\alpha^{-1}) = \pm\sigma_2(\alpha)...\sigma_n(\alpha) \in \mathcal{O}_{\mathbb{C}} \\ \implies \alpha^{-1} \in \mathcal{O}_L \end{aligned}$$

□

Remark. We'll prove later in the course that $\mathcal{O}_L^*$ is a finite group $\iff$ either $L = \mathbb{Q}$ or L is an imaginary quadratic field.

3 Discriminants and integral bases

Let L be a number field, $n = [L : \mathbb{Q}]$, $\sigma_1, ..., \sigma_n : L \to \mathbb{C}$ be distinct complex embeddings.

Definition. (3.1)
Let $\alpha_1, ..., \alpha_n \in L$. Then their discriminant is $disc(\alpha_1, ..., \alpha_n) = \det(D)^2$, where $D = M_{n\times n}(F)$ is $D_{ij} = \sigma_i(\alpha_j)$. Note: this is independent of the choice of ordering of $\sigma_1, ..., \sigma_n$ and $\alpha_1, ..., \alpha_n$, as that's just permuting the rows or columns, hence changing only possibly signs; but we took a square in the definition.

Lemma. (3.2)
Let $\alpha_1, ..., \alpha_n \in L$. Then $disc(\alpha_1, ..., \alpha_n) = \det(T)$, where $T \in M_{n\times n}(\mathbb{Q})$ is $T_{ij} = \mathrm{tr}_{L/\mathbb{Q}}(\alpha_i\alpha_j)$.

Proof. $T_{ij} = \sum_{k=1}^n \sigma_k(\alpha_i\alpha_j) = \sum_{k=1}^n D_{ki}D_{kj} = (D^T D)_{ij}$. □

Corollary. (3.3)
$disc(\alpha_1, ..., \alpha_n) \in \mathbb{Q}$. If $\alpha_1, ..., \alpha_n \in \mathcal{O}_L$, then $disc(\alpha_1, ..., \alpha_n) \in \mathbb{Z}$.

Proof. $disc(\alpha_1, ..., \alpha_n) = \det(T)$, and entries of T is trace of some elements of L (over $\mathbb{Q}$) so is in the base field $\mathbb{Q}$ (think a bit). So this must be rational. If $\alpha_1, ..., \alpha_n \in \mathcal{O}_L$, then $\forall i, j$, $D_{ij} \in \mathcal{O}_\mathbb{C} \implies disc(\alpha_1, ..., \alpha_n) \in \mathcal{O}_\mathbb{C} \cap \mathbb{Q} = \mathbb{Z}$. □

Proposition. (3.4)
Let $\alpha_1, ..., \alpha_n \in L$. Then $disc(\alpha_1, ..., \alpha_n) \neq 0 \iff \alpha_1, ..., \alpha_n$ form a basis of L as $\mathbb{Q}$-vector space.

Proof. First suppose $\alpha_1, ..., \alpha_n$ are linearly dependent. Then the columns of the matrix $D_{ij} = \sigma_i(\alpha_j)$ are linearly depnedent $\implies disc(\alpha_1, ..., \alpha_n) = 0$ (determinant is 0).
Now suppose $\alpha_1, ..., \alpha_n$ are linearly independent. Then $disc(\alpha_1, ..., \alpha_n) \neq 0$ $\iff \det(T) \neq 0 \iff$ the symmetric bilinear form $\phi : L \times L \to \mathbb{Q}$ by $\phi(\alpha, \beta) = \mathrm{tr}_{L/\mathbb{Q}}(\alpha\beta)$ is non-degenerate, i.e. $\forall \alpha \in L^*, \exists \beta \in L$ such that $\phi(\alpha, \beta) \neq 0$.
If $\alpha \in L^*$, then $\phi(\alpha, \alpha^{-1}) = \mathrm{tr}_{L/\mathbb{Q}}(1) = n \neq 0$. □

Definition. (3.5)
We say elements $\alpha_1, ..., \alpha_n \in L$ form an *integral basis for* $\mathcal{O}_L$, if:
(i) $\alpha_1, ..., \alpha_n \in \mathcal{O}_L$;
(ii) $\alpha_1, ..., \alpha_n$ generate $\mathcal{O}_L$ as a $\mathbb{Z}$-module.

Lemma. (3.6)
If $\alpha_1, ..., \alpha_n$ form an integral basis for $\mathcal{O}_L$, then the function

$$f : \mathbb{Z}^n \to \mathcal{O}_L$$
$$(m_1, ..., m_n) \to \sum_{i=1}^n m_i\alpha_i$$

is an isomorphism of $\mathbb{Z}$-module.

Proof. f is a homomorphism, we must show it's bijective. Observe that $\alpha_1, ..., \alpha_n$ form a basis of L as $\mathbb{Q}$-vector space. We know that if $\beta \in L$, then $\exists N \in \mathbb{Z}^+$ such that $N\beta \in \mathcal{O}_L$ (I think (1.7)). So we can write $N\beta = \sum_{i=1}^n m_i\alpha_i$ for some $m_1 \in \mathbb{Z} \implies \beta = \sum_{i=1}^n \frac{m_i}{N}\alpha_i$. Hence $\alpha_1, ..., \alpha_n$ span L, so they form a basis of L.
If $f(m_1, ..., m_n) = 0$, then $\sum_{i=1}^n m_i\alpha_i = 0 \implies (m_1, ..., m_n) = (0, ..., 0)$, as $\alpha_1, ..., \alpha_n$ are independent over $\mathbb{Q}$. This shows f is injective. It's surjecitve by definition. $\square$

Lemma. (3.7, sandwich lemma)
(i) If $H \leq G$ are groups and $G \cong \mathbb{Z}^a$ for some $a \geq 0$, then $H \cong \mathbb{Z}^b$ for some $b \leq a$.
(ii) If $K \leq H \leq G$ are groups and $K \cong \mathbb{Z}^a$, $G \cong \mathbb{Z}^a$ for some $a \geq 0$, then $H \cong \mathbb{Z}^a$.
(iii) If $H \leq G$ are groups and $H \cong \mathbb{Z}^a$, $G \cong \mathbb{Z}^a$ for some $a \geq 0$, then G/H is finite.

Proof. (i) $H \leq G$, $G \cong \mathbb{Z}^a$. Then G/H is f.g abelian group. By the classification, there's an isomorphism $G/H \cong \mathbb{Z}^N \oplus A$, A finite abelian group. Choose p prime, $p \not| |A|$. Then the map $f : G/H \to G/H$ by $x + H \to px + H$ is injective, so $f' : H/pH \to G/pG$ by $x + pH \to x + pG$ is injecitve − why? If $x \in H, x \in pG$, then $x = py$ for some $y \in G$; then $y + H \in \ker(f) = H$. Hence $x \in pH$. So indeed f' is injective. By the classification, $H \cong \mathbb{Z}^b$. f' injective $\implies |H/pH| \leq |G/pG|$, i.e. $p^b \leq p^a$ so $b \leq a$.
(ii) Apply (i) to $K \leq H$ and $H \leq G$ to get $H \cong \mathbb{Z}^b$ where $a \leq b \leq a$.
(iii) $H \leq G$, $H \cong \mathbb{Z}^a, G \cong \mathbb{Z}^a$. Again G/H is finitely generated, so by the classification $G/H \cong \mathbb{Z}^N \oplus A$ where A is a finite abelian group.
Let p be a prime, $p \not| |A|$. same proof as in (i) shows that $f' : H/pH \to G/pG$ is injecitve. Since $|H/pH| = |G/pG| = p^a$, f' is a group isomorphism $G/H + pG \cong (\mathbb{Z}/p\mathbb{Z})^N$. There's a surjective homomorphism $G/pG \to G/H + pG$ which has kernel containing the image of f'. Hence $G/pG \to G/H + pG$ is surjective with kernel G/pG. This forces $N = 0$. $\square$

Let L be a number field, $n = [L : \mathbb{Q}]$, $\sigma_1, ..., \sigma_n : L \to \mathbb{C}$ be distinct complex embeddings; $\alpha_1, ..., \alpha_n \in L$, we defined $disc(\alpha_1, ..., \alpha_n) = \det(\sigma_i(\alpha_j))^2$. An alternative notation is $\Delta(\alpha_1, ..., \alpha_n)$. We also said $\alpha_1, ..., \alpha_n$ form an integral basis for $\mathcal{O}_L$ if they generate $\mathcal{O}_L$ as a $\mathbb{Z}$-module.

Proposition. (3.8)
There exists an integral basis for $\mathcal{O}_L$.

Proof. Let $\beta_1, ..., \beta_n \in L$ be a basis for L as $\mathbb{Q}$-vector space. WLOG, $\beta_1, ..., \beta_n \in \mathcal{O}_L$. Then $\mathcal{O}_L \supset \oplus_{i=1}^n \mathbb{Z}\beta_i$.
Recall $\phi : L \times L \to \mathbb{Q}$ by sending $(\alpha, \beta) \to \mathrm{tr}_{L/\mathbb{Q}}(\alpha\beta)$ is a non-degenerate symmetric bilinear form (we showed that last time). Let $\beta_1^*, ..., \beta_n^*$ be the dual basis. Then $\mathrm{tr}_{L/\mathbb{Q}(\beta_i\beta_j^*)} = \delta_{ij}$ (why?).
If $\alpha \in \mathcal{O}_L$, then we can write $\alpha = \sum_{i=1}^n a_i\beta_i^*$ where $a_i \in \mathbb{Q}$. We know $\alpha\beta_i \in \mathcal{O}_L$, hence $\mathrm{tr}_{L/\mathbb{Q}}(\alpha\beta) \in \mathbb{Z}$. However LHS $= \sum_{j=1}^n \mathrm{tr}_{L/\mathbb{Q}}(a_j\beta_j^*\beta_i) =$

$\sum_{j=1}^{n} a_j \operatorname{tr}_{L/\mathbb{Q}}(\beta_j^* \beta_i) = a_j$. So $\mathcal{O}_L \subseteq \oplus_{i=1}^{n} \mathbb{Z}\beta_i^*$. By sandwich lemma there is an isomorphism between $\mathbb{Z}^n$ and $\mathcal{O}_L$. $\square$

If $\alpha_1, ..., \alpha_n$, $\beta_1, ..., \beta_n$ are both integral bases for $\mathcal{O}_L$, then there exists $A \in M_{n\times n}(\mathbb{Z})$ such that $\beta_j = \sum_{i=1}^{n} A_{ij}\alpha_i$ for each $j = 1, ..., n$. Moreover, we must have $\det(A) \in \{\pm 1\}$, and $A \in GL_n(\mathbb{Z})$. Then $disc(\beta_1, ..., \beta_n) = \det(D')^2$, where $D'_{ij} = \sigma_i(\beta_j), D_{ij} = \sigma_i(\alpha_j)$. We have $D'_{ij} = \sum_{k=1}^{n} \sigma_i(A_{kj}\alpha_k) = \sum_{k=1}^{n} \sigma_i(\alpha_k)A_{kj} = (DA)_{ij}$.

We find $disc(\beta_1, ..., \beta_n) = \det(D')^2 = \det(DA)^2 = \det(D)^2 = disc(\alpha_1, ..., \alpha_n)$. Therefore we could define:

Definition. (3.9)
The discriminant D_L of the number field L is $disc(\alpha_1, ..., \alpha_n)$, where $\alpha_1, ..., \alpha_n$ is any integral basis for $\mathcal{O}_L$.

Proposition. (3.10)
Let $L = \mathbb{Q}(\alpha)$, and let $f(x) \in \mathbb{Q}[x]$ be the minimal polynomial of α over $\mathbb{Q}$. Then

$$disc(1, \alpha, \alpha^2, ..., \alpha^{n-1}) = \prod_{i<j}(\sigma_i(\alpha) - \sigma_j(\alpha))^2 = (-1)^{n(n-1)/2} N_{L/\mathbb{Q}}(f'(\alpha))$$

In part II Galois theory, we defined the discrimant of a polynomial, $discf = \prod_{i<j}(\sigma_i(\alpha) - \sigma_j(\alpha))^2$ where α_i's are the roots of f.

Proof. If $D_{ij} = \sigma_i(\alpha^{j-1})$, $D \in M_{n\times n}(\mathbb{C})$, then $disc(1, \alpha, ..., \alpha^{n-1}) = \det(D)^2$. D is a Vandermonde matrix, so we know $\det(D) = \prod_{i<j}(\sigma_j(\alpha) - \sigma_i(\alpha))$.
On the other hand, $N_{L/\mathbb{Q}}(f'(\alpha)) = \prod_{i=1}^{n} \sigma_i(f'(\alpha)) = \prod_{i=1}^{n} f'(\sigma_i(\alpha))$.
Using $f(x) = \prod_{j=1}^{n}(x - \sigma_j(\alpha))$, we get RHS $= \prod_{i=1}^{n}\prod_{j\neq i}(\sigma_i(\alpha) - \sigma_j(\alpha)) = (-1)^{\binom{n}{2}} \prod_{i<j}(\sigma_i(\alpha) - \sigma_j(\alpha))^2$. $\square$

Note: if $\alpha \in \mathcal{O}_L$ and $\mathbb{Z}[\alpha] = \mathcal{O}_L$, then $1, \alpha, ..., \alpha^{n-1}$ is an integral basi for $\mathcal{O}_L$. We can then use proposition to calculate D_L.

Example. Let $d \in \mathbb{Z}$ square-free, $d \neq 0, 1$, $L = \mathbb{Q}(\sqrt{d})$. Then

$$D_L = \begin{cases} 4d & d \equiv 2, 3 \pmod 4 \\ d & d \equiv 1 \pmod 4 \end{cases}$$

To see this, if $d \equiv 2, 3 \pmod 4$, then $\mathcal{O}_L = \mathbb{Z}[\sqrt{d}]$ (shown previously). Apply proposition to $x^2 - d = f(x)$, we get $D_L = disc(1, \sqrt{d}) = -N_{L/\mathbb{Q}}(2\sqrt{d}) = 4d$.
On the other hand, if $d \equiv 1 \pmod 4$, then $\mathcal{O}_L = \mathbb{Z}[\frac{1+\sqrt{d}}{2}]$. Apply proposition to the minimal polynomial of this element, $f(x) = x^2 - x + \frac{1-d}{4}$, so $f'(x) = 2x - 1$, so $f'(\alpha) = \sqrt{d}$. Therefore $D_L = -N_{L/\mathbb{Q}}(\sqrt{d}) = \sqrt{d}$.

Proposition. If $\alpha_1, ..., \alpha_n \in \mathcal{O}_L$ are such that $disc(\alpha_1, ..., \alpha_n)$ is a non-zero square-free integer, then $\alpha_1, ..., \alpha_n$ form an integral basis for $\mathcal{O}_L$.
Note: this is a sufficient condition, but is not necessary (the previous example).

Proof. Let $\beta_1, ..., \beta_n$ be an integral basis for $\mathcal{O}_L$. There exists $A \in M_{n\times n}(\mathbb{Z})$ such that $\alpha_j = \sum_{i=1}^n A_{ij}\beta_i \; \forall j = 1, ..., n$. Then $disc(\alpha_1, ..., \alpha_n) = \det(A)^2 disc(\beta_1, ..., \beta_n)$ (we proved this in the beginning of lecture: $D' = DA$). In particular, if this is square-free and non-zero, then $\det(A)$ must be $\{\pm 1\}$. So $A \in GL_n(\mathbb{Z})$. Hence $\alpha_1, ..., \alpha_n$ generate $\mathcal{O}_L$ (as they can generate β_i) and form an integral basis. □

This could save a lot of calculation if we are lucky.

Example. Let $f(x) = x^3 - x - 1$. Then $discf = -4a^3 - 27b^2 = -23$. This is square-free! If $L = \mathbb{Q}(\alpha)$, α a root of $f(x)$, then $\mathcal{O}_L = \mathbb{Z}[\alpha]$.

Definition. (3.12)
Let $I \subseteq \mathcal{O}_L$ be a no-zero ideal. Then elements $\alpha_1, ..., \alpha_n \in L$ form an integral basis for I if:
(i) $\alpha_1, ..., \alpha_n \in I$;
(ii) $\alpha_1, ..., \alpha_n$ generate I as a $\mathbb{Z}$-module.

Proposition. (3.13)
Let $I \subseteq \mathcal{O}_L$ be a non-zero ideal. Then there exists an integral basis for I.

Definition. By definition, $I \subseteq \mathcal{O}_L \cong \mathbb{Z}^n$. Let $\alpha_1, ..., \alpha_n \in \mathcal{O}_L$ be an integral basis for $\mathcal{O}_L$. Let $\alpha \in I$ be non-zero. Then $(\alpha) \subseteq I$, hence $\oplus_{i=1}^n \mathbb{Z}\alpha\alpha_i \subseteq I \subseteq \mathcal{O}_L$. So by sandwich lemma, there is an isomorphism between I and $\mathbb{Z}^n$ as $\mathbb{Z}$-module. Hence there exists an integral basis for I.

An interesting consequence of the proof:

Definition. (3.14)
If $I \subseteq \mathcal{O}_L$ is a non-zero ideal, then we define its norm

$$N(I) = [\mathcal{O}_L : I]$$

which is finite by the sandwich lemma.

Definition. (3.15)
If $I \subset \mathcal{O}_L$ is a non-zero ideal then we define $disc(I) = disc(\alpha_1, ..., \alpha_n)$ where $\alpha_1, ..., \alpha_n$ is an integral basis for I. (same argument shows $disc(I)$ depends only on I).

Lemma. (3.16)
If $I \subseteq \mathcal{O}_L$ is a non-zero ideal, then $disc(I) = disc(\mathcal{O}_L)N(I)^2$.

Proof. Let $\alpha_1, ..., \alpha_n$, $\beta_1, ..., \beta_n$ be integral bases for $\mathcal{O}_L$ and I respectively. Then $\exists A \in M_{n\times n}(\mathbb{Z})$ such that $\beta_j = \sum_{i=1}^n A_{ij}\alpha_i \; \forall j = 1, ...n$, and $disc(\alpha_1, ..., \alpha_n)\det(A)^2 = disc(\beta_1, ..., \beta_n)$. We must show $\det(A)^2 = [\mathcal{O}_L : I]^2$.

In fact, we'll show if $B \in M_{n\times n}(\mathbb{Z})$ and $\det(B) \neq 0$, then $|\mathbb{Z}^n/B\mathbb{Z}^n| = |\det(B)|$. This suffices after identify $\mathcal{O}_L \cong \mathbb{Z}^n$.

Recall: $\exists P, Q \in GL_n(\mathbb{Z})$ such that $PBQ = D = Diag(d_1, ..., d_n)$, $d_i \in \mathbb{Z}$ (Smith normal form). Hence we have $\mathbb{Z}^n/B\mathbb{Z}^n \cong \mathbb{Z}^n/D\mathbb{Z}^n \cong \oplus_{i=1}^n \mathbb{Z}/d_i\mathbb{Z} \implies |\mathbb{Z}^n/B\mathbb{Z}^n| = |\mathbb{Z}^n/D\mathbb{Z}^n| = \prod_{i=1}^n |d_i|$.
On the other hand, $|\det(B)| = |\det(D)| = \prod_{i=1}^n |d_i|$. □

Remember we have L a number field, $n = [L : \mathbb{Q}]$, $\sigma_1, ..., \sigma_n : L \to \mathbb{C}$ are distinct complex embeddings of L.

Lemma. (3.17)
Let $\alpha \in \mathcal{O}_L \setminus \{0\}$. Then $N((\alpha)) = |N_{L/\mathbb{Q}}(\alpha)|$ (Note that's an ideal).

Proof. Let $\alpha_1, ..., \alpha_n$ be an integral basis for $\mathcal{O}_L$. Then $\alpha\alpha_1, ..., \alpha\alpha_n$ is an integral basis for $I = (\alpha)$. So

$$\begin{aligned} disc(I) &= disc(\alpha\alpha_1, ..., \alpha\alpha_n) \\ &= \det(\sigma_i(\alpha\alpha_j))^2 \\ &= \det(\sigma_i(\alpha)\sigma_i(\alpha_j))^2 \\ &= (\prod_{i=1}^{n} \sigma_i(\alpha))^2 \det(\sigma_i(\alpha_j))^2 \\ &= N_{L/\mathbb{Q}}(\alpha)^2 disc(\mathcal{O}_L) \end{aligned}$$

And we showed last time that for any non-zero ideal $J \subseteq \mathcal{O}_L$, $disc(J) = N(J)^2 disc(\mathcal{O}_L)$. $\square$

Notation: If $\alpha \in \mathcal{L} - \{0\}$, we let $N(\alpha) = N((\alpha))N(0) = 0$.
Then $\forall \alpha, \beta \in \mathcal{O}_L$, $N(\alpha\beta) = N(\alpha)N(\beta)$.

4 Unique factorisation in $\mathcal{O}_L$

Recall: we say a ring R is a unique factorisation domain (UFD) if
(i) R is an integral domain;
(ii) if $x \in R$ is non-zero and not a unit, then there exists an expression $x = p_1 ... p_r$ where $p_i \in R$ are irreducible elements. This expression is unique in the sense that if $x = q_1 ... q_s$ is another such expression, then $r = s$ and after re-ordering, each q_i is an associate of p_i (i.e. $q_i \in R^* p_i$, where R^* is the field of units).

After 2 years of Cambridge Maths we certainly know $\mathbb{Z}$ is a UFD. However, if L is a number field, $\mathcal{O}_L$ need not be a UFD.

In fact, any non-zero $x \in \mathcal{O}_L$ which is not a unit can be expressed as a product of irreducible elements.

If $x \in \mathcal{O}_L$, then x is a no-zero non-unit $\iff N(x) > 1$. Suppose $x \in \mathcal{O}_L$ is a non-zero non-unit which cannot be written as a product of irreducible elements, and with $N(x)$ minimal among elements with this property. Then $x = yz$ with $N(y) > 1$, $N(z) > 1$, hence $N(y) < N(x)$, $N(z) < N(x)$. By minimality of $N(x)$, both y, z can be written as products of irreducible; contradiction.

Example. Consider $L = \mathbb{Q}(\sqrt{-5}$, $\mathcal{O}_L = \mathbb{Z}[\sqrt{-5}]$, and $\mathcal{O}_L^* = \{\pm 1\}$. In $\mathcal{O}_L$ we have $6 = 2 \times 3 = (1 + \sqrt{-5})(1 - \sqrt{-5})$, and all of the four are irreducibles, and no two are associates (norms). So $\mathcal{O}_L$ is not a UFD (famous example).

Idea: introduce ideal multiplication in order to reduce elements further.

Recall that if R is a ring and I, J are ideals of R, then we define

$$IJ = \{\sum_{i=1}^k a_i b_i | a_i \in I, b_i \in J\},$$
$$I + J = \{a + b | a \in I, b \in J\}$$

We can define an ideal $I \subsetneq R$ to be irreducible if it does not admit an expression $I = JK$ where J, K are proper ideals of R.

Key point: even if $\alpha \in \mathcal{O}_L$ is irreducible, the ideal (α) need not be irreducible. For example in $\mathbb{Z}[\sqrt{-5}]$, we have $(2) = (2, 1+\sqrt{-5})^2$, $(3) = (3, 1+\sqrt{-5})(3, 1-\sqrt{-5})$.

Definition. (4.1)
If R is a ring, we say that an ideal $P \subsetneq R$ is prime if $\forall x, y \in R$, $xy \in P \implies x \in P$ or $y \in P$.

Lemma. (4.2)
Let R be a ring, and let $I, J, P \subseteq R$ be ideals, and suppose P is prime and $IJ \subseteq P$. Then $I \subseteq P$ or $J \subseteq P$.

Proof. WLOG $I \not\subseteq P$. Choose some $x \in I \setminus P$. If $y \in J$, is any element, then $xy \in IJ \subseteq P$. So $y \in P$. So $J \subseteq P$. $\square$

From now on, L is a number field.

Lemma. (4.3)
Any non-zero prime ideal $P \subseteq \mathcal{O}_L$ is a maximal ideal.

Proof. Recall: if R is a ring and $I \subsetneq R$ is an ideal, then I is prime $\iff R/I$ is an integral domain, and I is maximal $\iff R/I$ is a field. If you don't remember these statements then I strongly encourage you to review GRM. If $p \subseteq \mathcal{O}_L$ is a non-zero prime ideal, then $\mathcal{O}_L/P$ is a finite integral domain (of cardinality $N(P)$); any such ring is a field, so P is also maximal. □

Lemma. (4.4)
If $I \subsetneq \mathcal{O}_L$ is a non-zero ideal, then there exist non-zero prime ideals $P_1, ..., P_r \subseteq \mathcal{O}_L$ such that $P_1...P_r \subseteq I$.

Proof. For contradiction, let $I \subsetneq \mathcal{O}_L$ be an ideal whicih does not have this property, and such that $N(I)$ is minimal among ideals not having this property. Then I is not prime, so there exist elements $x, y \in \mathcal{O}_L$ such that $xy \in I$ but $x \notin I$, $y \notin I$. But then it follows that $I \subsetneq I + (x)$ and $I \subsetneq I + (y)$. So $N(I + (x)), N(I + (y)) < N(I)$. By minimality of $N(I)$, we can find non-zero prime ideals $P_1...P_r \subseteq I + (x)$ and $Q_1...Q_r \subseteq I + (y)$. Then $P_1...P_rQ_1...Q_r \subseteq (I + (x))(I + (y)) \subseteq I^2 + xI + yI + (xy) \subseteq I$. Contradiction. □

Lemma. (4.5)
If $I \subsetneq \mathcal{O}_L$ is a non-zero ideal, then there exists $\gamma \in L \setminus \mathcal{O}_L$ such that $\gamma I \subseteq \mathcal{O}_L$.

Proof. Let $\alpha \in I \setminus \{0\}$. Let $P_1, ..., P_r \subseteq \mathcal{O}_L$ be non-zero prime ideals such that $P_1...P_r \subseteq (\alpha)$. WLOG r is minimal with this property. Let P be a minimal ideal containing I. Then $P \supseteq I \supseteq (\alpha) \supseteq P_1...P_r$, hence $P \supset P_i$ for some i. After relabelling assume $P \supset P_1$. Since non-zero prime ideals are maximla, we have $P = P_1$. Since r is minimal, we have $P_2...P_r \not\subseteq (\alpha)$. Choose $\beta \in P_2...P_r \setminus (\alpha)$.
Claim: the element $\gamma = \beta/\alpha$ has the desired property.
If $\gamma \in \mathcal{O}_L$, then $\beta = \alpha\gamma \in (\alpha)$, contradiction;
$\gamma I = \frac{\beta}{\alpha} I \subseteq \frac{1}{\alpha} P_2...P_r \cdot I \subseteq \frac{1}{\alpha} P_1P_2...P_r \subseteq \mathcal{O}_L$. □

Let L be a number field. Last lecture we proved that if $I \subsetneq \mathcal{O}_L$ is a non-zero ideal, then there exist $\gamma \in L \setminus \mathcal{O}_L$ such that $\gamma I \subseteq \mathcal{O}_L$.

Proposition. (4.6)
If $I \subseteq \mathcal{O}_L$ is a non-zero ideal, there exists a non-zero ideal $J \subseteq \mathcal{O}_L$, such that IJ is principal.

Proof. Choose $\alpha \in I \setminus \{0\}$. Define $J = \{\beta \in \mathcal{O}_L | \beta I \subseteq (\alpha)\}$. J is a non-zero ideal, as $\alpha \in J$. We have $IJ \subseteq (\alpha)$. We will show $IJ = (\alpha$.
Let $K = \frac{1}{\alpha} IJ \subseteq \mathcal{O}_L$. We will show in fact that $K = \mathcal{O}_L$. Suppose otherwise, that $K \neq \mathcal{O}_L$, then $\exists \gamma \in L \setminus \mathcal{O}_L$ such that $\gamma K \subseteq \mathcal{O}_L$.
We have $(\alpha) \subseteq I$, hence $\frac{1}{\alpha} I \supseteq \mathcal{O}_L$, hence $underbrace\frac{1}{\alpha}IJ_K \supset J$. Hence $\gamma J \subseteq \gamma K \subseteq \mathcal{O}_L$.
Another observation is that, we also have $\gamma IJ = \gamma\alpha K \subseteq (\alpha)$.

If we have $\beta \in \gamma J$, on one hand $\beta \in \mathcal{O}_L$; on the other hand, $\beta I \subseteq (\alpha)$. So $\beta \in J$, hence $\gamma J \subseteq J$.
Recall that J admits an integral basis, so ther's an isomorphism $J \cong \mathbb{Z}^n$. If $A \in M_{n\times n}(\mathbb{Z})$ is the matrix representing multiplication by γ, and if $f(x) \in \mathbb{Z}[x]$ is the characteristic polynomial of A, then $f(\gamma) = 0$.
Hence $\gamma \in \mathcal{O}_L$. Contradiction. So $K = \mathcal{O}_L$. □

Corollary. (4.7)
If $I, J, K \subseteq \mathcal{O}_L$ are non-zero ideals and $IJ = IK$, then $J = K$.

Proof. Choose a non-zero ideal $A \subseteq \mathcal{O}_L$ such that $AI = (\alpha)$ is principal. Then $AIJ = \alpha J = AIK = \alpha K \implies J = K$. □

If $I, J \subseteq \mathcal{O}_L$ are non-zero ideals, say I divides J (or $I|J$) if there exists an ideal $K \subseteq \mathcal{O}_L$ such that $IK = J$.

Corollary. (4.8)
If $I, J \subseteq \mathcal{O}_L$ are non-zero ideals, then $I|J \iff I \supseteq J$.

Proof. If $IK = J$, then $J \subseteq I$.
Suppose instead that $I \supseteq J$. Choose a non-zero ideal $A\mathcal{O}_L$ such that $AI = (\alpha)$ is principal (by 4.6). Then $AI = (\alpha) \supseteq AJ$, hence $\mathcal{O}_L \supseteq \frac{1}{\alpha}AJ$. So $K = \frac{1}{\alpha}AJ$ is a non-zero ideal of $\mathcal{O}_L$, and $IK = \frac{1}{\alpha}AIJ = J$. □

Theorem. (4.9)
If $I \subseteq \mathcal{O}_L$ is a non-zero ideal, then there exist prime ideals $P_1, ..., P_r \subseteq \mathcal{O}_L$ such that $I = P_1P_2...P_r$. Moreover, this expression is unique up to re-ordering of terms.

Proof. We show existence by contradiction. Suppose I is an ideal which cannot be written as product of primes, and with $N(I)$ minimal subject to this condition. We can find a maximal ideal $P \supset I$. P is also prime. Then $P|I$, so we can write $I = PJ$ for some ideal $J \subseteq \mathcal{O}_L$. Then $J|I$, hence $J \supset I$. If $J = I$, then we get $I = IP$, hence $\mathcal{O}_L = P$ as we can cancel, but that's a contradiction as prime ideals by definition cannot be $\mathcal{O}_L$.
Therefore $J \supsetneq I$, hence $N(J) < N(I)$. By minimality, we can write J as $J = P_2...P_r$ where each $P_i \subseteq \mathcal{O}_L$ are prime ideals. Then we have $I = PJ$. Contradiction. This shows existence.

For uniqueness, suppose $P_1, ..., P_r$, $Q_1, ..., Q_s$ are non-zero prime ideals in $\mathcal{O}_L$ such that $P_1...P_r = Q_1...Q_s$. Then $P_1|Q_1...Q_r$, so $P_1 \supseteq Q_i$ for some $i = 1, ..., s$. WLOG $P_1 \supset Q_1$. Since both P_1, Q_1 are maximal, $P_1 = Q_1$. Then we cancel to obtain $P_2...P_r = Q_2...Q_s$; continue this to get $r = s$ and $P_i = Q_i$ after re-ordering. □

Definition. (4.10)
The ideal class group $Cl(\mathcal{O}_L) = \{I \subseteq \mathcal{O}_L \text{ non-zero ideal}\}$. $I \sim J$ if $\exists \alpha \in L^*$ such that $\alpha I = J$.
We write $[I]$ for the equivalence class containing I.

Lemma. (4.11)
$Cl(\mathcal{O}_L$ is a group under the operation

$$[I][J] = [IJ]$$

with identity $[\mathcal{O}_L]$.

Proof. If $I, J \subseteq \mathcal{O}_L$ are non-zero ideals and $\alpha, \beta \in L^*$ are such that $\alpha I \subseteq \mathcal{O}_L$ and $\beta J \subseteq \mathcal{O}_L$. Then

$$(\alpha I)(\beta J) = \alpha\beta IJ$$

so ideal multiplication is well-defined on equivalent classes.
For any $I \subseteq \mathcal{O}_L$, $\mathcal{O}_L I = I$, so $[\mathcal{O}_L]$ is an identity.
We showed that if $I \subseteq \mathcal{O}_L$ is any non-zero ideal, then there exists a non-zero ideal $J \subseteq \mathcal{O}_L$ such that $IJ = (\alpha)$ is principal. Then $[I][J] = [IJ] = [(\alpha)] = [\mathcal{O}_L]$. Hence $[I]^{-1} = [J]$. □

Proposition. (4.12)
The following are equivalent:
(i) $\mathcal{O}_L$ is a PID;
(ii) $\mathcal{O}_L$ is a UFD;
(iii) The ideal class group, $Cl(\mathcal{O}_L)$, is trivial.

Proof. (i) implies (ii): In IB GRM.
(ii) implies (iii): We must show any ideal $I \subseteq \mathcal{O}_L$ is principal. We know that we can write $I = P_1 ... P_r$ as a product of prime ideals.
It's therefore enough to show that every prime ideal of $\mathcal{O}_L$ is principal. Let $P \subseteq \mathcal{O}_L$ be a non-zero prime ideal, let $\alpha \in P$ be non-zero, and let $\alpha = \alpha_1 ... \alpha_r$ be an expression of α as a product of irreducibles.
Recall: if R is a ring, then we say $x \in R$ is prime if $\forall y, z \in R, x|yz \implies x|y$ or $x|z$. Also we learned from GRM that if R is a UFD then irreducible elements of R are prime.
We find $P \supset \alpha = (\alpha_1)...(\alpha_r) \implies P|P_1...P_r$ where $P_i = (\alpha_i)$. Since α_i is prime, P_i is a prime ideal. Hence we must have $P = P_i = (\alpha_i)$ for some i, and hence P is principal.
(iii) implies (i): Let $I \subseteq \mathcal{O}_L$ be a non-zero ideal. Since $Cl(\mathcal{O}_L$ is trivial, we have $[I] = [\mathcal{O}_L]$, so there exists $\alpha \in L^*$ such that $\alpha\mathcal{O}_L = I$. We have $\alpha \cdot 1 = \alpha \in I \subseteq \mathcal{O}_L$, so $\alpha \in \mathcal{O}_L$, hence $I = (\alpha)$ is principal. □

Lemma. (4.13)
If $I, J \subseteq \mathcal{O}_L$ are non-zero ideals, then $N(IJ) = N(I)N(J)$.

Proof. Example sheet 2. □

Example sheet 2 now available!

Last time we learned that, if L is a number field, then we know any non-zero ideal $I \subseteq \mathcal{O}_L$ canbe written uniquely as $I = \prod_{i=1}^r P_i^{e_i}$, wher the p_i are distinct prime ideals, and $e_i \geq 1$. We also defined $Cl(\mathcal{O}_L)$ as the obstruction to $\mathcal{O}_L$ being a UFD.

5 Dedekind's criteion

If $P \subseteq \mathcal{O}_L$ is a non-zero prime ideal, then there's a unique prime number $p \in \mathbb{Z}_{\geq 0}$ such that $p \in P$. $(p) = \ker(\mathbb{Z} \to \mathcal{O}_L/P)$. Then $P|p\mathcal{O}_L$, and $N(P) = p^f$ for some $f \geq 1$.

Lemma. (5.1)
Let p be a prime number, and factor $p\mathcal{O}_L = \prod_{i=1}^r P_i^{e_i}$ where $P_1, ..., P_r$ are distinct prime ideals of $\mathcal{O}_L$, $e_i \geq 1$. Define $f_i \geq 1$ by $N(P_i) = p^{f_i}$. Then $\sum_{i=1}^r e_i f_i = [L : \mathbb{Q}]$. In particular, $r \leq [L : \mathbb{Q}]$.

Proof. Apply norm to get $N(p\mathcal{O}_L)(= p^{[L:\mathbb{Q}]}) = \prod_{i=1}^r N(P_i)^{e_i}(= p^{\prod_{i=1}^r e_i f_i})$. □

Definition. (5.2)
Let p be a prime number, and let $p\mathcal{O}_L = \prod_{i=1}^r P_i^{e_i}$ be the factorization as above.
(i) We say p *ramifies* in L if $e_i > 1$ for some i. We say p is totally *ramified* if $r = 1$ and $e_1 = [L : \mathbb{Q}]$. In other words, $p\mathcal{O}_L = P_i^{[L:\mathbb{Q}]}$.
(ii) We say p is *inert* in L if $r = 1$ and $e_1 = 1$, i.e. $p\mathcal{O}_L$ is prime.
(iii) We say p *splits completely* in L if $r = [L : \mathbb{Q}]$ and $e_i = f_i = 1$ for all i.

Note that these don't cover all the possible cases.

Theorem. (5.3, Dedekind's criterion)
Let $\alpha \in \mathcal{O}_L$ be such that $L = \mathbb{Q}(\alpha)$. Let $f(x) \in \mathbb{Z}[x]$ be its minimal polynomial and let p be a prime such that $p \nmid [\mathcal{O}_L : \mathbb{Z}[\alpha]]$.
Let $\bar{f}(x) = f(x) \pmod p$, and factor $\bar{f}(x) = \prod_{i=1}^r \bar{g}_i(x)^{e_i}$ in $F_p[x]$, where $\bar{g}_1(x), ..., \bar{g}_r(x) \in F_p[x]$ are distinct monic irreducible polynomials. Let $g_i(x) \in \mathbb{Z}[x]$ be any polynomial with $g_i(x) \pmod p = \bar{g}_i(x)$, and define $Q_i = (p, g_i(\alpha)) \subseteq \mathcal{O}_L$, an ideal of $\mathcal{O}_L$. Let $f_i = \deg \bar{g}_i(x)$.
Then $Q_1, ..., Q_r$ are distinct prime ideals of $\mathcal{O}_L$, and $p\mathcal{O}_L = \prod_{i=1}^r Q - i^{e_i}$, and $N(Q_i) = p^{f_i}$.

For example, let's take $L = \mathbb{Q}(\sqrt{-11})$, $p = 5$. We see $-11 \equiv 1 \pmod 4$, so $\mathcal{O}_L = \mathbb{Z}[\frac{1+\sqrt{-11}}{2}]$. Thus $\mathbb{Z}[\sqrt{-11}] \subseteq \mathcal{O}_L$ has index 2 as an additive subgroup. Therefore we can apply Dedekind's criterion to $\alpha = \sqrt{-11}$, with $f(x) = x^2 + 11$ in order to factorize $5\mathcal{O}_L$. We see $\bar{f}(x) = f(x) \pmod 5 = x^2 + 1 = (x+2)(x+3)$ in $F_5[x]$. So $t\mathcal{O}_L = PQ$ where $P = (5, \sqrt{-11} + 2), Q = (5, \sqrt{-11}, 3)$, and hence P, Q are the same prime ideals (of $\mathcal{O}_L$). Thus $5\mathcal{O}_L$ splits completely in $\mathcal{Q}\sqrt{-11}$.

Proof. (of 5.3)
Recall: if R is a ring and $I \subseteq R$ is an ideal, then there's a bijection between ideals containing I and idealks of R/I. 3rd isomorphism theorem gives $R/J \cong (R/I)/(J/I)$. We have $\mathbb{Z}[\alpha] \subseteq \mathcal{O}_L$ of finite index. Let $A = \mathbb{Z}[\alpha]$, $\phi : A \to \mathcal{O}_L$. By reduction mod p, we get another ring homomorphism $\bar{\phi} : A/pA \to \mathcal{O}_L/p\mathcal{O}_L$ by $\bar{\phi}(\beta + pA) = \beta + p\mathcal{O}_L$.
We claim that this is actually an isomorphism. Both source and targe have cardinality $p^{[L:\mathbb{Q}]}$, so it's enough to show $\bar{\phi}$ is surjective. Let $N = [\mathcal{O}_L : \mathbb{Z}[\alpha]]$. We can find $a, b \in \mathbb{Z}$ such that $aN + bp = 1$. If $\beta \in \mathcal{O}_L$, then $N\beta \in \mathbb{Z}[\alpha]$ (by

Lagrance), and $\beta = aN\beta + bp\beta \implies \bar{\phi}(aN\beta + pA) = \beta + p\mathcal{O}_L$. Therefore there is a bijection between ideals in $\mathcal{O}_L$ containing p and ideals of A/pA.
We have $A = \mathbb{Z}[\alpha] \cong \mathbb{Z}[x]/(f(x))$ by sending α to x. Reduction mod p gives an isomorphism $A/pA \cong \mathbb{Z}[x]/(p, f(x)) \cong F_p[x]/(\bar{f}(x))$. We have $\bar{f}(x) = \prod_{i=1}^r \bar{g}_i(x)^{e_i}$, so there are homomorphisms $F_p[x]/(\bar{f}(x)) \to \mathbb{F}_p[x]/(\bar{g}_i(x))$, given by quotient by the ideal $(\bar{g}_i(x)) \supseteq (\bar{f}(x))$. Define $\mathbb{Q}_i \subseteq \mathcal{O}_L$ to be the ideal containing p such that $\mathbb{Q}_i/(p)$ is the kernel of the ring homomorphism $\mathcal{O}_L/p\mathcal{O}_L \xrightarrow{\bar{\phi}^{-1}} A/pA \xrightarrow{\cong} F_p[x]/(\bar{f}(x)) \to F_p[x]/(\bar{g}_i(x))$. This ring homomorphism is surjective, and its image is a field of cardinality p^{f_i}. Hence $\mathcal{O}_L/\mathbb{Q}_i$ is a finite field of cardinality p^{f_i}, hence $\mathbb{Q}_i$ is a prime ideal of norm $N(\mathbb{Q}_i) = p^{f_i}$.
Also, the $\mathbb{Q}_i$ are distinct, because their images in $\mathcal{O}_L/p\mathcal{O}_L$ are distinct, as if $i \neq j$ then $(\bar{g}_i(x), \bar{g}_j(x))$ is the unit ideal of $F_p[x]$. To show $\mathbb{Q}_i = (p, g_i(\alpha))$, it's enough to show $\mathbb{Q}_i/(p) \subseteq \mathcal{O}_L/p\mathcal{O}_L$ is generated by $\bar{g}_i(\alpha)$. This is equivalent to showing that $\ker(F_p[x]/(\bar{f}(x)) \to F_p[x]/(\bar{g}_i(x)))$ is generated by $\bar{g}_i(x)$. This is true by definition.
It remains to show $Q_1^{e_1}...Q_r^{e_r} = p\mathcal{O}_L$. We have

$$\begin{aligned} Q_1^{e_1}...Q_r^{e_r} &= (p_1 g_1(\alpha))^{e_1}...(p_r g_r(\alpha))^{e_r} \\ &= (p_1 g_1(\alpha)^{e_1})...(p_1 g_r(\alpha)^{e_r}) \\ &\leq (p, g_1(\alpha)^{e_1})...(g_r(\alpha)^{e_r}) = (p, f(\alpha)) = (p) \end{aligned}$$

Take norms, $N(LHS) = \prod_{i=1}^r N(Q_i)^{e_i} = p^{\sum_{i=1}^r e_i f_i} = p^{\deg f} = p^{[L:\mathbb{Q}]} = N(p) = N(RHS)$. This forces $Q_1^{e_1}...Q_r^{e_r} = p\mathcal{O}_L$. □

Let L be a number field. Last time we had that if $\alpha \in \mathcal{O}_L$, $\mathbb{Q}(\alpha) = L$, $p \nmid [\mathcal{O}_L : \mathbb{Z}[\alpha]]$. Dedekind's criterion: can factor $p\mathcal{O}_L$ by factoring $f_\alpha(x) \pmod p$.

Proposition. (5.4)
Let d be a square-free integer, $d \neq 0, 1$, $L = \mathbb{Q}(\sqrt{d})$, and let p be a prime number. Then
(1) If p is odd, then:
- if $p|d$, then $(p) = P^2$, so p ramifies in L;
- if $p \nmid d$ and $(\frac{d}{p}) = 1$, then $(p) = PQ$, so p splits completely in L;
- if $p \nmid d$ and $(\frac{d}{p} = -1$, then (p) is prime and p is inert in L.

(2) If $p = 2$, then:
- if $d \equiv 2, 3 \pmod 4$, then 2 ramifies in L;
- if $d \equiv 1 \pmod 8$, then 2 splits completely in L;
- if $d \equiv 5 \pmod 8$, then 2 is inert in L.

Proof. We just do the case where $p = 2$. If $d \equiv 2, 3 \pmod 4$, then $\mathcal{O}_L = \mathbb{Z}[\sqrt{d}]$, so by Dedekind's criterion, we must factor $x^2 - d \pmod 2$. But $x^2 - d \equiv (x - d)^2 \pmod 2$. If $d \equiv 1 \pmod 4$, then $\mathcal{O}_L = \mathbb{Z}[\frac{1+\sqrt{d}}{2}]$, so we must factor $x^2 + x + \frac{1-d}{4} \pmod 2$. If $d \equiv 1 \pmod 8$, this is $x^2 + x = x(x+1) \pmod 2$. If $d \equiv 5 \pmod 8$, this is $x^2 + x + 1 \pmod 2$ which is irreducible. □

6 Geometry of numbers

Definition. (6.1)
If V is a finite dimensional $\mathbb{R}$-vector space, then a lattice in V is a subgroup of the form $\Lambda = \oplus_{i=1}^m \mathbb{Z}v_i$, where $v_1, ..., v_n$ is a basis of V as $\mathbb{R}$-vector space (for example, $\mathbb{Z}^n \subseteq \mathbb{R}^n$).

Definition. (6.2)
If V is a finite-dimensional inner product space over $\mathbb{R}$, and $\Lambda \subseteq V$ is a lattice, then the covolume of Λ is

$$A(\Lambda) = vol(\{\sum_{i=1}^n t_i v_i | t_i \in [0,1)\})$$

where $\Lambda = \oplus_{i=1}^n \mathbb{Z}v_i$.
Check: this is independent of the choice of basis $v_1, ..., v_n$.

For today, let's consider only a fixed imaginary quadratic field $L = \mathbb{Q}(\sqrt{d})$ where $d < 0$ is a square-free integer. Let's take $\sigma : L \to \mathbb{Q}$ be a complex embedding. Then $\sigma(\mathcal{O}_L)$ is a lattice in ϕ. If $d \equiv 2, 3 \pmod 4$, then $\sigma(\mathcal{O}_L) = \mathbb{Z} \oplus \mathbb{Z}[\sqrt{d}]$; if $d \equiv 1 \pmod 4$ then $\sigma(\mathcal{O}_L) = \mathbb{Z} \oplus \mathbb{Z}(\frac{1+\sqrt{d}}{2})$
If $I \leq \mathcal{O}_L$ is a non-zero ideal, then $\sigma(I)$ is a lattice in $\mathbb{C}$.

Lemma. (6.3)
If $I \subseteq \mathcal{O}_L$ is a non-zero ideal, then $A(I) = \frac{1}{2}\sqrt{|disc(I)|} = \frac{N(I)}{2}\sqrt{|D_L|}$.

Proof. Let α_1, α_2 be an integral basis for I. Then $\sigma(I) = \mathbb{Z}\sigma(\alpha_1) \oplus \mathbb{Z}\sigma(\alpha_2)$. Write $\alpha_1 = x_1 + iy_1, \alpha_2 = x_2 + iy_2$, then $A(\sigma(I)) = |\det \left(\begin{smallmatrix} x_1 & x_2 \\ y_1 & y_2 \end{smallmatrix}\right)|$ (area of a parallelogram).
Then

$$disc(I) = \det \begin{pmatrix} x_1 + iy_1 & x_2 + iy_2 \\ x_1 - iy_1 & x_2 - iy^2 \end{pmatrix} = (2i)^2 \det \begin{pmatrix} y_1 & y_2 \\ x_1 & x_2 \end{pmatrix}$$

□

Theorem. (6.4, special case of Minkovski's theorem)
Let $\Lambda \subseteq \mathbb{R}^2$ be a lattice, and let $S = D(0, r) \subseteq \mathbb{R}^2$ be the closed disk of radius r. Then if $area(S) \geq 4A(\Lambda)$, then $\exists \lambda \in \Lambda - \{0\}$ such that $\lambda \in S$.
In particular, there exists $\lambda \in \Lambda - \{0\}$ such that $|\lambda|^2 \leq \frac{4}{\pi}A(\Lambda)$.

Corollary. (6.5)
If $I \subseteq \mathcal{O}_L$ is a non-zero ideal, then there exists $\alpha \in I - \{0\}$ s.t. $N(\alpha) \leq c_L N(I)$, where $c_L := \frac{2}{\pi}\sqrt{|D_L|}$.

Proof. We apply the theorem to $\sigma(I) \subseteq \mathbb{C}$ to get $\lambda \in \sigma(I) - \{0\}$, such that $|\lambda|^2 \leq \frac{4}{\pi} \cdot \frac{N(I)}{2}\sqrt{|D_L|} = c_l N(I)$. If $\alpha \in I$ is such that $\sigma(\alpha) = \lambda$, then $N(\alpha) = \sigma(\alpha)\overline{\sigma(\alpha)} = |\sigma(\alpha)|^2 = |\lambda|^2$. □

Corollary. (6.6)
If $[I] \in Cl(\mathcal{O}_L)$, then there exists $J \in [I]$ such that $N(J) \leq c_L$.

Proof. Choose $k \in [I]^{-1}$ so that IK is principal. Apply the corollary to find $\alpha \in K - \{0\}$, such that $N(\alpha) \leq c_L N(K)$. Then $(\alpha) \subseteq K \implies K|(\alpha) \implies \exists J \subseteq \mathcal{O}_L$ non-zero ideal such that $JK = (\alpha)$. We have $[J] = [K]^{-1} = [I]$, so $J \in [I]$. Also, $N(J) = N(\alpha)/N(K) \leq c_L$. $\square$

Theorem. (6.7)
The group $Cl(\mathcal{O}_L)$ is finite. (we'll prove this for any L next time).

Proof. We've shown every class $[I] \in Cl(\mathcal{O}_L)$ has a representative of norm $\leq c_L$. It therefore suffices to show that $\forall m \in \mathbb{Z}, m \geq 1$, the number of ideals $I \subseteq \mathcal{O}_L$ of norm $N(I) = m$ is finite. If $N(I) = m$, then $[\mathcal{O}_L : I] = m$, so by Lagrance, $m \in I$. Thus I comes from an ideal of the finite ring $\mathcal{O}_L/m\mathcal{O}_L$. $\square$

Note: we see $CL(\mathcal{O}_L)$ is generated by ideal classes $[P]$, where $P \subseteq \mathcal{O}_L$ is a non-zerp prime ideal of norm $N(P) \leq c_L$. Why? Any class has the form $[I]$, where $N(I) \leq c_L$. If $I = \prod_{i=1}^r p_i^{e_i}$, then $[I] = ()i = 1^r [P_i]^{e_i}$ and $N(I) = \prod_{i=1}^r N(P_i)^{e_i}$, so $N(P_i) \leq N(I) \leq c_L$ for each $i = 1, ..., r$.

Example. Consider $d = -7$. $d \equiv 1 \pmod 4$, so $D_L = -d$, $c_l = \frac{2}{\pi}\sqrt{7} < \frac{2}{3}\sqrt{7} < 2$.
$Cl(\mathcal{O}_L)$ is generated by ideals of norm < 2. There are none except $\mathcal{O}_L$, so $Cl(\mathcal{O}_L)$ is the trivial group. Hence $\mathcal{O}_L = \mathbb{Z}[\frac{1+\sqrt{-7}}{2}]$ is a UFD.
$d = -5$: $D_L = -4d$, $c_L = \frac{2}{\pi}\sqrt{70} = \frac{4}{\pi}\sqrt{5} < \frac{4}{3}\sqrt{5} < 3$. Hence $Cl(\mathcal{O}_L)$ is generated by prime ideals $P \subseteq \mathcal{O}_L$ of norm $N(P) = 2$. We know by Dedekind's criterion that $2\mathcal{O}_L = P^2$. Hence $Cl(\mathcal{O}_L)$ is generated by $[P]$, and $[P]^2 = [2\mathcal{O}_L]$ is the trivial class.
Hence there are two possibilities: if P is principal, then $Cl(\mathcal{O}_L)$ is trivial; if P is not principal, then $Cl(\mathcal{O}_L) \cong \mathbb{Z}/2\mathbb{Z}$. We know $\mathcal{O}_L$ is not a UFD, so we must have $Cl(\mathcal{O}_L) \cong \mathbb{Z}/2\mathbb{Z}$.

Last time we see that if L is an imaginary quadratic field, then $Cl(\mathcal{O}_L)$ is finite, generated by $[P]$ where P is a prime ideal of norm $N(P) \leq C_L$, where $C_L = \frac{2}{\pi}\sqrt{|D_L|}$.

This time we will show the case of a general number field L.

Theorem. (6.8, Minkowski's theorem)
Let $\Lambda \subseteq \mathbb{R}^n$ be a lattice, and let $E \subseteq \mathbb{R}^n$ be a measurable subset which is conve, and centrally symmetric ($E = -E = \{x \in \mathbb{R}^n | -x \in E\}$). Then:
(i) If $vol(E) > 2^n A(\Lambda)$, then $\exists \lambda \in \Lambda \setminus \{0\}$ such that $\lambda \in E$;
(i) If $vol(E) \geq 2^n A(\Lambda)$ and E is compact, then $\exists \lambda \in \Lambda \setminus \{0\}$ such that $\lambda \in E$.
(we used this last time in the special case $n = 2$, E=closed disk).

Proof. Let $\Lambda = \oplus_{i=1}^n \mathbb{Z}v_i$, $P = \{\sum_{i=1}^n t_i v_i | t_i \in [0,1)\}$. Then $vol(P) = A(\Lambda)$, and $\mathbb{R}^n = \sqcup_{\lambda \in \Lambda}(P + \lambda)$.
(i) $vol(P) < \frac{1}{2^n} vol(E) = vol(\frac{1}{2}E) = \sum_{\lambda \in \Lambda} vol([\frac{1}{2}E] \cap [\lambda + P]) = \sum_{\lambda \in \Lambda} vol([\frac{1}{2}E - \lambda] \cap P)$.
We claim that there exists $\lambda \neq \mu \in \Lambda$ such that $(\frac{1}{2}E - \lambda) \cap (\frac{1}{2}E - \mu)$ is non-empty. Why? If not, sets $\frac{1}{2}E - \lambda$ are pairwise disjoint, so $vol(P) <$

$\sum_{\lambda\in\Lambda} vol([\frac{1}{2}E-\lambda]\cap P)\leq vol(P)$, contradiction.
Hence $\exists z, w\in E$ such that $\frac{z}{2}-\lambda=\frac{w}{2}-\mu$, where $\lambda\neq\mu\in\Lambda$, so $\lambda-\mu=\frac{z}{2}-\frac{w}{2}=\frac{z}{2}+\frac{(-w)}{2}$. Since E is centrally symmetric, $-w\in E$, and E is convex implies that $\frac{z}{2}+\frac{(-w)}{2}\in E$, so $\lambda-\ mu\in(\Lambda\setminus\{0\})\cap E$.
(ii) E compact implies that E is closed and bounded. $vol(E)\geq 2^nA(\Lambda)$ so $\forall m\geq 1$, $vol((1+\frac{1}{m})E)>2^nA(\Lambda)$. By (i), $\forall m\in\mathbb{N}\exists s\lambda_m\in(\Lambda\setminus\{0\})\cap((1+\frac{1}{m})E)$, and $(1+\frac{1}{m})E\subseteq 2E$, and $2E\cap\Lambda$ is finite as $2E$ is bounde. By pigeonhole principle we can assume $\exists\lambda\in\Lambda\setminus\{0\}$ such that $\lambda_m=\lambda\forall m\geq 1$. E closed and $\lambda\in(1+\frac{1}{m})E\forall m\geq 1\implies\lambda\in E$. Now let L be a number field. Let $n=[L:\mathbb{Q}]$, let $\tau_1,...,\tau_r:L\to\mathbb{R}$ be the real embeddings of L, and let $\sigma_1,\bar{\sigma}_1,...,\sigma_s,\bar{\sigma}_s:L\to\mathbb{C}$ be the remaining distinct complex embeddings of L. Then $r+2s=n$.

Define a map $S:l\to\mathbb{R}^r\times\mathbb{C}^s$ by $\alpha\to(\tau_1(\alpha),...,\tau_r(\alpha),\sigma_1(\alpha),...,\sigma_s(\alpha))$. This is a homomorphism of additive groups. □

Lemma. If $I\subseteq\mathcal{O}_L$ is a non-zero ideal, then $S(I)$ is a lattice.

Proof. Let $\alpha_1,...,\alpha_n$ be an integral basis of I. Then $S(I)=\oplus_{i=1}^n\mathbb{Z}s(\alpha_i)$ and $\mathbb{R}^r\times\mathbb{C}^3$ has dimension n as $\mathbb{R}$-vector space. So we must show that $S(\alpha_1),...,S(\alpha_n)$ are independent or equivalently that

$$\det\begin{pmatrix}\tau_1(\alpha)1)...\tau_1(\alpha_n)\\ ...\\ \tau_r(\alpha_1)...\tau_r(\alpha_n)\\ Re\sigma_1(\alpha_1)...Re\sigma_1(\alpha_n)\\ Im\sigma_1(\alpha_1)...Im\sigma_1(\alpha_n)\\ ...\\ Im\sigma_n(\alpha_1)...Im\sigma_s(\alpha_n)\end{pmatrix}\neq 0$$

Note: for $z\in\mathbb{C}$,

$$\begin{pmatrix}z\\ z\end{pmatrix}=\begin{pmatrix}1 & i\\ 1 & -i\end{pmatrix}\begin{pmatrix}Rez\\ Imz\end{pmatrix}$$

So this determinant equals

$$(\frac{1}{-2i})^s\det\begin{pmatrix}\tau_1(\alpha)1)...\tau_1(\alpha_n)\\ ...\\ \tau_r(\alpha_1)...\tau_r(\alpha_n)\\ \sigma_1(\alpha_1)...\sigma_1(\alpha_n)\\ ...\\ sigma_n(\alpha_1)...\sigma_s(\alpha_n)\end{pmatrix}\neq 0$$

as $disc(I)\neq 0$. □

Lemma. (6.10)
If $I\subseteq\mathcal{O}_L$ is a non-zero ideal, then

$$A(S(I))=\frac{1}{2^s}\sqrt{|disc(I)|}=\frac{N(I)}{2^s}\sqrt{|D_L|}$$

Proposition. (6.11)
If $I\subseteq\mathcal{O}_L$ is a non-zero ideal, then there exists $\alpha\in I\setminus\{0\}$ such that $N(\alpha)\leq C_LN(I)$, where $C_L=(\frac{4}{\pi})^s\frac{n!}{n^n}\sqrt{|D_L|}$.
Here C_L is called the Minkowski constant of L.

Proof. We apply Minkowski's theorem to the lattice $S(I)$, and region $B_{r,s}(t) = \{(\mathbf{x},\mathbf{z}) \in \mathbb{R}^r \times \mathbb{C}^s | \sum_{i=1}^r |X_i| + 2\sum_{i=1}^s |z_i| \leq t\}$.
Note: $B_{r,s}(t)$ is convex, centrally symmetric and compact.
If $vol(B_{r,s}(t)) \geq 2^n A(S(I))$, then there exists $\alpha \in I \setminus \{0\}$ such that $S(\alpha) \in B_{r,s}(t)$.
We use a tuck with the AM-GM inequality to bound $N(\alpha)$:

$$N(\alpha)^{1/n} = (\prod_{i=1}^r |\tau_i(\alpha)) \prod_{i=1}^s |\sigma_i(\alpha)|^2)^{1/n} \leq \frac{(\sum_{i=1}^r |\tau_1(\alpha)| + 2\sum_{i=1}^s |\sigma_i(\alpha)|)}{n}$$

Hence $N(\alpha) \leq t^n/n^n$. To get optimal bound, choose t so that $vol(B_{r,s}(t)) = 2^n A(S(I))$.
Exercise: $vol(B_{r,s}(t)) = 2^r(\frac{\pi}{2})^s t^n/n!$ (Induction on r and s).
We have

$$\begin{aligned} 2^r(\pi/2)^s t^n/n! = 2^n A(S(I)) = 2^{r+s} N(I)\sqrt{|D_L|} \\ \implies t^n = (4/\pi)^s n! N(I)\sqrt{|D_L|} \\ \implies N(\alpha) \leq t^n/n^n = C_L N(I) \end{aligned}$$

□

Corollary. (6.12)
For any class $[I] \in Cl(\mathcal{O}_L)$, there exists $J \in [I]$ such that $N(J) \leq C_L$.

Corollary. (6.13)
The group $Cl(\mathcal{O}_L)$ is finite, generated by $[P]$ where P is a prime ideal of norm $N(P) \leq C_L$.

These corollaries are deduced from the proposition exactly as in the case $L = \mathbb{Q}(\sqrt{d})$, $d < 0$.

Remark. In practice this bound is very effective. For example consider $f(x) = x^5 - x + 1$, this is irreducible mod 5, so over $\mathbb{Q}$. Let $L = \mathbb{Q}(\alpha)$ where α is a root of $f(x)$. In this case $r = 1, s = 2$, the discriminant $discf = 2869 = 19 \cdot 151$ is square-free, so $\mathcal{O}_L = \mathbb{Z}[\alpha]$, and $D_L = discf$, so $c_L = (4/\pi)^2(5^!/5^5)\sqrt{2869} < 4$. Hence $Cl)\mathcal{O}_L$ is generated by P of norm $N(P) = 2$ or 3. By Dedekind's criterion, such primes exist iff $f(x)$ has a root in F_2 or F_3. But there are no such roots. Hence $Cl(\mathcal{O}_L)$ is trivial, hence $\mathbb{Z}[\alpha]$ is a UFD.

Last time we showed $CL(\mathcal{O}_L)$ is generated by $[P]$ where $[P]$ is a prime ideal of norm $N(P) \leq C_L = (4/\pi)^3 n!/n^n\sqrt{|D_L|}$. For example, if $L = \mathbb{Q}(\sqrt{10})$, $C_L = \frac{1}{2}\sqrt{4 \cdot 10} = \sqrt{10} < 4$. $Cl(\mathcal{O}_L$ is generated by $[P]$ where $N(P) = 2$ or 3.
Dedekind's criterion: $2\mathcal{O}_L = P_2^2$, where $P_2 = (2, \sqrt{10})$. $x^2 - 10 \equiv x^2 - 1 \pmod 3$ so $3\mathcal{O}_L = P_3 P_3'$, where $P_3 = (3, 1 + \sqrt{10})$. To find relatoins in $Cl(\mathcal{O}_L)$, we can calculate norms, e.g. $N(2 + \sqrt{10}) = |4 - 10| = 6$, so $(2 + \sqrt{10}) = P_2 P_3$ or $P_2 P_3'$. In either case we see that $[P_2]$ generates $Cl(\mathcal{O}_L$. So either $Cl(\mathcal{O}_L)$ is trivail, or $Cl(\mathcal{O}_L \cong \mathbb{Z}/2\mathbb{Z}$ with the second case occuring iff So P_2 is not principal. P_2 is principal $\iff \exists a + b\sqrt{10} \in \mathcal{O}_L$ such that $(a + b\sqrt{10}) = P_2 \iff \exists a, b \in \mathbb{Z}$ s.t. $a^2 - 10b^2 = \pm 2$.
If $a^2 - 10b^2 = \pm 2$, then either 2 or -2 is a quadratic residue $\pmod 5$. So in fact P_2 is not principal. So $Cl(\mathcal{O}_L) \cong \mathbb{Z}/2\mathbb{Z}$.

Now take $L = \mathbb{Q}(\sqrt{-17})$. $C_l = \frac{4}{\pi} \cdot \frac{1}{2}\sqrt{4 \cdot 17} = 4/\pi\sqrt{17} < \frac{4}{3}\sqrt{17} < 6$. So $Cl(\mathcal{O}_L)$ is generated by primes of norm $2, 3$ or 5. Dedekind's criterion: $x^2 + 17 \equiv x^2 + 2 \pmod 5$, so $5\mathcal{O}_L$ is prime of norm 25. $x^2 + 17 \equiv x^2 - 1 \pmod 3$, so $3\mathcal{O}_L = Q_3 Q_3'$ where $Q_3 = (3, 1 + \sqrt{-17})$, $Q_3' = (3, 1 - \sqrt{-17})$. $x^2 + 17 = (x + 1)^2 \pmod 2$, so $2\mathcal{O}_L = Q_2^2$ where $Q_2 = (2, 1 + \sqrt{-17})$.
Now $N(1 + \sqrt{-17}) = 18 = 2 \times 3^2$. Note $1 + \sqrt{-17} \in Q_3 \implies Q_3 | (1 + \sqrt{-17})$. So we must have either $(1 + \sqrt{-17}) = Q_2 Q_3 Q_3'$, or $(1 + \sqrt{-17}) = Q_2 Q_3^2$. To decide between these, we compute

$$\begin{aligned} Q_3^2 &= (0, 3 + 3\sqrt{-17}, (1 + \sqrt{-17})^2) \\ &= (9, 3 + 3\sqrt{-17}, -16 + 2\sqrt{-17}) \\ &= (9, 3 + 3\sqrt{-17}, 2 + 2\sqrt{-17}) \\ &= (9, 1 + \sqrt{-17}) \end{aligned}$$

We see $1 + \sqrt{-17} \in Q_3^2$ so $Q_3^2 | (1 + \sqrt{-17})$, hence $(1 + \sqrt{-17}) = Q_2 Q_3^2$. We see $[Q_3]$ generates $Cl(\mathcal{O}_L)$ and if Q_2 is not principal then $Cl(\mathcal{O}_L) \cong \mathbb{Z}/4\mathbb{Z}$. But Q_2 is principal iff we can solve $a^2 + 17b^2 = 2$ with $a, b \in \mathbb{Z}$. This is impossible, so $Cl(\mathcal{O}_L) \cong \mathbb{Z}/4\mathbb{Z}$.

Remark. Ther are many open questions about ideal class groups even for quadratic fields.
Things we know: Number of $Cl(\mathcal{O}_{\mathbb{Q}(\sqrt{d})} \to \infty$ as $d \to -\infty$ through squaree-free integers. There are exactly 9 imaginary quadratic fields with trivial ideal class group (hard).
Things we don't know: are there infinitely many real quadratic fields of trivial ideal class group?

Cohen-Lenstra heuristics: let p be an odd prime, and let A be a finite abelian group of p-power order. Then for $d < 0$ square-free, $\mathbb{P}(Cl(\mathcal{O}_{\mathbb{Q}(\sqrt{d})}) \cong A) = \frac{\prod_{i=1}^{\infty}(1 - 1/p^i)}{\text{Number of } Aut(A)}$.

For M a finite abelian group, M_p is the (unique) p-sylow subgroup.

By definition, The above probablity is the ratio between the number of $d < 0$ square-free, $Cl(\mathcal{O}_{\mathbb{Q}(\sqrt{d})})_p \cong A$, $|d| < X$ and the number of $d < 0$ square-free, $|d| < x$.

7 Dirichlet's unit theorem

Let L be a number field of degree $n = [L : \mathbb{Q}]$, $\tau_1, ..., \tau_r : L \to \mathbb{R}$ are real embeddings, $\sigma_1, ..., \sigma_s, \bar{\sigma}_1, ..., \bar{\sigma}_s : L \to \mathbb{C}$ are distinct complex embeddings.

Theorem. (7.1)
There is an isomorphism $\mathcal{O}_L^* \cong \mu_L \times \mathbb{Z}^{r+s-1}$, where $\mu_L \subseteq \mathcal{O}_L^*$ is the finite cyclic group of roots of unity in $\mathcal{O}_L^*$.
In fact the proof shows omre: define a map $l : \mathcal{O}_L^* \to \mathbb{R}^{r+s}$: $l(\alpha) = (\log|\tau_1(\alpha)|, ..., \log|\tau_r(\alpha)|, 2\log|\sigma_1(\alpha)|, ..., 2\,\mathrm{lo}$
then this is a homomorphism of abelian groups, and $l(\mathcal{O}_L^*)$ is contained in the hyperplane $H = \{\mathbf{x} \in \mathbb{R}^{r+s} | \sum_{i=1}^{r+s} x_i = 0\} \subseteq \mathbb{R}^{r+s}$. This expresses the condition $\alpha \in \mathcal{O}_L^* \implies \log N(\alpha) = \sum_{i=1}^{r} \log|\tau_i(\alpha)| + 2\sum_{i=1}^{s} |\sigma_i(\alpha)|$.

The proof of the theorem will show $l(\mathcal{O}_L^*)$ is a lattice in H.

Example: $\mathcal{O}_L^*$ is finite $\iff$ $r + s = 1$, i.e. $r = 1, s = 0$ ($L = \mathbb{Q}$), or $r = 0, s = 1$ ($L = \mathbb{Q}(\sqrt{d})$,$d < 0$ square-free). The first case where $\mathcal{O}_L^*$ is infinite is $L = \mathbb{Q}\sqrt{d}), d > 0$, square-free. Then $+s - 1 = 1$, so $l(\mathcal{O}_L^*)$ is infinite cyclic. Let's fix $\sigma : \mathbb{Q}(sqrtd) \to \mathbb{R}$ to be the real embedding with $\sigma(\sqrt{d}) \geq 0$. $\sigma(\mu_L) \subseteq \mathbb{R}^*$, so $\mu_L = \{\pm 1\}$ in this case. In this case, we can consider the map $l' : \mathcal{O}_L^* \to \mathbb{R}$ by $\alpha \to \log|\sigma(\alpha)|$. We know that $l'(\mathcal{O}_L^*) \subseteq \mathbb{R}$ is a lattice, in particular there is a uniquely characterised unit $\alpha \in \mathcal{O}_L^*$ satisfying $\sigma(\alpha) > 0$, $\log|\sigma(\alpha)| > 0$ and as small as possible. In other words, $\alpha \in \mathcal{O}_L^*$ is the unit for which $\sigma(\alpha) > 1$ and $\sigma(\alpha)$ is minimal with respect to this property. We call α the fundamental unit of $L = \mathbb{Q}(\sqrt{d})$. Then we have $\mathcal{O}_L^* = \{\pm\alpha^n | n \in \mathbb{Z}\}$.

Example sheet 3 is now online!

Last time we have: if L is a number field, then $\mathcal{O}_L^* \cong \mu_L \times \mathbb{Z}^{r+s-1}$, where μ_L are roots of unity.

Now suppose $L = \mathbb{Q}(\sqrt{d})$ where $d \in \mathbb{Z}$ is a square free integer, $d > 1$. We identify L with a subfield of $\mathbb{R}$, where $\sqrt{d}$ is the positive square root.

We saw that the Dirichlet's unit theorem implies $\exists u \in \mathcal{O}_L^*$ such that $u = \min\{v \in \mathcal{O}_L^* | v > 1\}$. u is called the fundamental unit, and $\mathcal{O}_L^* = \{\pm u^n | n \in \mathbb{Z}\}$.

Lemma. (7.2)
(1) If $d \equiv 2, 3 \pmod 4$ and $v \in \mathcal{O}_L^*$ satisfies $v > 1$, then $v = a + b\sqrt{d}$ where $a \geq b \geq 1$;
(2) If $d \equiv 1 \pmod 4$, and $v \in \mathcal{O}_L^*$ satisfies $v > 1$, then $v = \frac{1}{2}(a + b\sqrt{d})$ wher $a \geq b \geq 1$.

Proof. (1) Let $v' = a - b\sqrt{d}$. Then $vv' = a^2 - db^2 = N_{L/\mathbb{Q}}(v) = \pm 1$. So $v > 1 \implies |v'| < 1$. Hence $v + v' = 2a > 0$, $v - v' = sb\sqrt{d} > 0$. As a, b are integers, we must have $a \geq 1, b \geq 1$.
Also, $(a/b)^2 = d \pm 1/b^2 \geq 1$ as $d \geq 2$.
(2) Let $v' = \frac{1}{2}(a - b\sqrt{d})$. Then $vv' = \pm 1$ and $a^2 - db^2 = \pm 4$. Then $v + v' = a > 0$, and $v - v' = b\sqrt{d} > 0$. Hence $a \geq 1, b \geq 1$. Also, $(a/b)^2 = d \pm 4/b^2$ as $d \geq 5$ as $d \equiv 1 \pmod 4$. $\square$

We can use this to find the fundamental unit $u \in \mathcal{O}_L^*$. First suppose $d \equiv 2, 3 \pmod 4$ and let $u = a + b\sqrt{d}$. Let $u^k = a_k + b_k\sqrt{d}$. Then $u^{k+1} = (a_1 + b_1\sqrt{d})(a_k + b_k\sqrt{d}) = (a_1 a_k + d b_1 b_k) + (b_1 a_k + a_1 b_k)\sqrt{d}$. Hence $b_{k+1} = b_1 a_k + a_1 b_k > b_k$. Hence the sequence b_1, b_2, b_3 is strictly increasing.

We can therefore characterise u as follows: let $b \in \mathbb{N}$ be the least positive integer such that $db^2 + 1$ or $db^2 - 1$ is of the form a^2 for some $a \in \mathbb{N}$. Then $u = a + b\sqrt{d}$. Now suppose $d \equiv 1 \pmod 4$, and let $u = \frac{1}{2}(a + b\sqrt{d})$, $a, b \in \mathbb{Z}$. Let $u^k = \frac{1}{2}(a_k + b_k\sqrt{d})$. Then $b_{k+1} = \frac{1}{2}(a_1 b_k + b_1 a_k)$. Using lemma 7.2, we see $b_{k+1} \geq b_k$. If (??)
This is wrong. Let's correct this next time. Sorry!

Example. $d = 2.L = \mathbb{Q}(\sqrt{2})$. $b = 1$ works: $2 - 1 = 1^2$. So $1 + \sqrt{2}$ is a fundamental unit.
$d = 7$. Try $b = 1$: 7 ± 1 is not a square; $b = 2$, doesn't work either; $b = 3$: $9 \cdot 7 \pm 1 = 8^2$. So $8 + 3\sqrt{7}$ is a fundamental unit.

Note: This procedure is not always efficient. For example, the fundamental unit in $\mathbb{Q}(\sqrt{22})$ is $197 + 42\sqrt{22}$.

There is a more efficient algorithm which uses continued fractions, but it is not discussed in this course (see number theory).

We now prove the unit theorem (this is non-examinable).

We recall the setup: L is a number field, $\tau_1, ..., \tau_r : L \to \mathbb{R}$, $\sigma_1, \bar{\sigma}_1, ..., \sigma_s, \bar{\sigma}_s : L \to \mathbb{C}$ are real and complex embeddings of L respectively.
Last time we defined a map: $l : \mathcal{O}_L^* \to \mathbb{R}^{r+s}$ by $\alpha \to (\log(\tau_1(\alpha)), ..., \log(\tau_r(\alpha)), 2\log(\sigma_1(\alpha)), ..., 2\log(\sigma_s(\alpha)))$.
The image is contained inside the subspace $H = \{\mathbf{x} \in \mathbb{R}^{r+s} | \sum_{i=1}^{r+s} x_i = 0\}$.

Lemma. (7.3)
Let $\alpha \in \mathcal{O}_L \setminus \{0\}$ be such that the above image vector is $(a_1, ..., a_{r+s}) \in \mathbb{R}^{r+s}$. Fix an integer $1 \leq k \leq r + s$. Then ther exists $\beta \in \mathcal{O}_L \setminus \{0\}$ such that if $l(\beta) = (b_1, ..., b_{r+s}) \in \mathbb{R}^{r+s}$, then $b_i < a_i$ if $i \neq k$. Moreover, $N(\beta) \leq (\frac{2}{\pi})^s \sqrt{|D_L|}$.

Proof. Let $c_1, ..., c_{r+s} \in \mathbb{R}_{>0}$, and let

$$E = \{(\mathbf{x}, \mathbf{z}) \in \mathbb{R}^r \times \mathbb{C}^s | |x_1| \leq c_1, ..., |x_r| \leq c_r, |z_1|^2 \leq c_{r+1}, ..., |z_r|^2 \leq c_{r+s}\}$$

Then if $vol(E) \geq 2^{r+2s} A(S(\mathcal{O}_L)) = 2^{r+s}\sqrt{|D_L|}$, then $(S : \mathcal{O}_L \to \mathbb{R}^r \times \mathbb{C}^s)$.

There exists $\beta \in \mathcal{O}_L \setminus \{0\}$ such that $S(\beta) \in E$ (by Minkovski's theorem). In particular, $N(\beta) = \prod_{i=1}^r |\tau_1(\beta)| \prod_{i=1}^s |\sigma_i(\beta)|^2 \leq c_1 ... c_{r+s}$ (by defiintion of E).

We choose c_i so that $0 < c_i < e^{a_i}$ if $i \neq k$, and $vol(E) = \pi^s 2^r c_1 ... c_{r+s} = 2^{r+s}\sqrt{|D_L|}$.

The first property gives $b_i < a_i$ if $i \neq k$, and the second property gives $N(\beta) \geq c_1 ... c_{r+s} = (\frac{2}{\pi})^s \sqrt{|D_L|}$. $\square$

Corollary. (7.4)
Fix an integer $1 \leq k \leq r + s$. Then there exists $\varepsilon \in \mathcal{O}_L^*$ such that if $l(\varepsilon) = (a_1, ..., a_{r+s})$ then $a_i < 0$ if $i \neq k$, and $a_k > 0$.

Proof. By the lemma, we can find elements $\alpha_1, \alpha_2, ...$ of $\mathcal{O}_L \setminus \{0\}$ such that $N(\alpha_1) \leq (\frac{2}{\pi})^s \sqrt{|D_L|}$ $\forall i \in \mathbb{N}$, and if $l(\alpha_i) = (b_{i_1}, ..., b_{i,r+s})$, then $b_{ij} < b_{i-1,j}$ if $j \neq k$ $\forall i = 2, 3, ...$ The ideals (α_i) have bounded norm, so are finite in number, so there exist elements α_N, α_M with $(\alpha_N) = (\alpha_M)$. Then the element $\varepsilon = \alpha_N/\alpha_M \in \mathcal{O}_L^*$ has the desired property. □

We continue with the non-examinable proof of Dirichlet's unit theorem.

We proved propoition: let $\alpha \in \mathcal{O}_L \setminus \{0\}$ be such that $l(\alpha)$ fix $1 \leq k \leq r+s$. Then $\exists \beta \in \mathcal{O}_L \setminus \{0\}$ such that $N(\beta) \leq (\frac{2}{\pi})^s \sqrt{|D_L|}$, and if $l(\beta) = (b_1, ..., b_{r+s})$ then $b_i < a_i$ if $i \neq k$.

We deduced Corllary 7.4: fix $1 \leq k \leq r+s$. Then there exists $\varepsilon \in \mathcal{O}_L^*$ such that if $l(\varepsilon) = (a_1, ..., a_{r+s})$, then $a_i < 0$ if $i \neq k$.

Proof. Choose $\alpha \in \mathcal{O}_L \setminus \{0\}$. By the proposition, we can find elements $\alpha_1, ...$ such that $N(\alpha_i) \leq (2/\pi)^s \sqrt{|D_L|}$, and if $l(i) = (b_{i1,...,ir+s})$ then $b_{ij} > b_{i+1j}$ if $j \neq k$ for all $i \geq 1$.

We now look at the ideals $(\alpha_1), (\alpha_2), ...$ These have norm at most $(2/\pi)\sqrt{|D_L|}$. We know there are only finitely many ideals of $\mathcal{O}_L$ of norm at monst that, so there must exist $N < M$ such that $(\alpha_N) = (\alpha_M)$. Hence $\exists u \in \mathcal{O}_L^*$ such that $\alpha_M = u\alpha_N$. Also, $u = \alpha_M/\alpha_N \implies l(u) = (b_{m1} - b_{N1}, ..., b_{mr+s} - b_{Nr+s})$. But $N < M$, so $b_{Nj} > b_{Mj}$ if $j \neq k$. So $B_{Mj} - b_{Nj} < 0$ if $j \neq k$. □

Lemma. (7.5)
Let $N \geq 1$, and let $A \in M_{N\times N}(\mathbb{R})$ be such that:
- $\sum_{i=1}^{N} A_{ij} = 0$ for all $j = 1, ..., N$;
- $A_{ij} > 0$ if $i = j$, and < 0 if $i \neq j$.

Then A has rank $N - 1$.

Proof. The rank is at most $N-1$. We show the first $N-1$ rows of A are LI. Suppose there exist $t_i \in \mathbb{R}, i = 1, ..., N-1$ not all zero s.t. $\sum_{i=1}^{N-1} t_i A_{ij} = 0$ for each $j = 1, ..., N$. WLOG after rescaling ther exists k that $t_k = 1$ and $t_i \leq 1$ if $i \neq k$. Then $0 = \sum_{i=1}^{N-1} t_i A_{ik} \geq \sum_{i=1}^{N-1} A_{ik} > \sum_{i=1}^{N} A_{ik} = 0$, contradiction. □

Lemma. (7.6)
Fix $B > 0$. Let $X_B = \{\alpha \in \mathcal{O}_L | \forall \sigma : L \to \mathbb{C}, |\sigma(\alpha)| \leq B\}$. THen X_B is finite.

Proof. Recall the map $S : \mathcal{O}_L \to \mathbb{R}^r \times \mathbb{C}^s$. $S(\mathcal{O}_L)$ is a lattice in $\mathbb{R}^r \times \mathbb{C}^s$. $S(X_B)$ is the intersection of the lattice $S(\mathcal{O}_L)$ with a compact subset of $\mathbb{R}^r \times \mathbb{C}^s$. Therefore it must be finite. □

Proposition. (7.7)
$l(\mathcal{O}_L^*)$ is a lattice in $H \leq \mathbb{R}^{r+s}$.

Proof. We must show there exist units $v_1, ..., v_{r+s-1} \in \mathcal{O}_L^*$ such that $l(v_1), ..., l(v_{r+s-1})$ span H as an $\mathbb{R}$-vector space and generate $l(\mathcal{O}_L^*)$ as an abelian group.
By corollary 7.4, we can find $\varepsilon_1, ..., \varepsilon_{r+s} \in \mathcal{O}_L^*$ such htat if $l(\varepsilon_j) = (A_{ij}, ..., A_{r+sj})$,

then $A_{ij} < 0$ if $i \neq j$ and $A_{ij} > 0$ if $i = j$. By lemma 7.5, the matrix A has rank $r + s - 1$, so we can find $v_1, ..., v_{r+s-1} \in \mathcal{O}_L^*$ such that $l(v_1), ..., l(v_{r+s-1})$ span $\mathcal{O}_L^*$ as an $\mathbb{R}$-vector space.
Let $\Lambda = \oplus_{i=1}^{r+s-1} \mathbb{Z} l(v_i) \leq H$. This is a lattice in H. Then $\Lambda \leq l(\mathcal{O}_L^*)$ and if $u \in \mathcal{O}_L^*$, then $\exists \lambda \in \Lambda$ such that $l(u) - \lambda \in \{\sum_{i=1}^{r+s-1} t_i l(v_i) | t_1, ..., t_{r+s-1} \in [0, 1]\} = P$. But the set of units $l(P)$ is finite by Lemma 7.6. Hence the quotien $l(\mathcal{O}_L^*)/\Lambda$ is finite. By Lagrange's theorem, $\exists N \in \mathbb{Z}, N > 1$ such that $Nl(\mathcal{O}_L^*) \leq \Lambda$. Hence $\Lambda \leq l(\mathcal{O}_L^*) \leq \frac{1}{N}\Lambda$. By the sandwich lemma, $l(\mathcal{O}_L^*)$ is a free abelian group of rank $r + s - 1$. In particular, it is a lattice in H. □

Let's now finish the proof of the unit theorem, i.e. show there's an isomorphism $\mathcal{O}_L^* \cong \mu_L \times \mathbb{Z}^{r+s-1}$, where μ_L is the (finite) group of roots of unity in $\mathcal{O}_L$.

Proof. We have $\mu_L = \ker l$. If $\xi \in \mu_L$, then $\xi^N = 1$ for some $N \geq 1$, hence $l(\xi^N) = 0 = Nl(\xi) \implies l(\xi) = 0$ as $l(\xi \in \mathbb{R}^{r+s}$. If $\alpha \in \mathcal{O}_L^*$ and $l(\alpha) = 0$ then $\forall \sigma : L \to \mathbb{C}$, $|\sigma(\alpha)| = 1$. By lemma 7.6, $\ker l$ is finite. By Lagrange's theorem, it consists of roots of unity.

Choose $v_1, ..., v_{r=s-1} \in \mathcal{O}_L^*$ such that $l(v_1), ..., l(v_{r+s-1})$ is a $\mathbb{Z}$-basis of $l(\mathcal{O}_L^*)$. Define a map $f : \mu_L \times \mathbb{Z}^{r+s-1} \to \mathcal{O}_L^*$ by $(\xi, n_1, ..., n_{r+s-1}) \to \xi v_1^{n_1} ... v_{r+s-1}^{n_{r+s-1}}$. □

Exercise: this is an isomorphism.

Return to the examinable parts:

We now show how to find the fundamental unit in $\mathbb{Q}(\sqrt{d})$, where $\sqrt{d} \in \mathbb{R}_{>0}$ and $d \in \mathbb{Z}$ is a positive square-free integer.

$d > 1, d \equiv 1 \pmod 4$:

Recall: the fundamental unit $u \in \mathcal{O}_L^*$ is the least unit $u > 1$. We saw last time that if $v = \frac{1}{2}(a + b\sqrt{d}) \in \mathcal{O}_L^*$ is any unit with $v > 1$, then $a \geq b \geq 1$.

Let $u^k = \frac{1}{2}(a_k + b_k\sqrt{d})$. Then $b_{k+1} = \frac{1}{2}(a_1 b_k + b_1 a_k) \geq \frac{1}{2}(a_1 + b_1)b_k \geq b_k$. We see $b_{k+1} \geq b_k$, with equality iff $a_k = b_k$ and $a_1 = b_1 = 1$. Note: if $a_1 = b_1 = 1$, then $N(u) = |\frac{1-d}{4}| = 1 \implies d = 5$. Assume first that $d > 5$. Then the sequence $b_1 < b_2 < b_3 < ...$ is strictly increasing. The fundamental unit u can therefore be found as following: let $b \in \mathbb{N}$ be the least positive integer such that $db^2 + 4 = a^2$ or $db^2 - 4 = a^2$, where $a \in \mathbb{N}$. Then $\frac{1}{2}(a + b\sqrt{d})$ is the fundamental unit.

Now suppose $d = 5$. Then at least $b_1 \leq b_2 \leq ...$ is non-decreasing, and each value b_i can appear at most twice: this is because occurrences correspond to solutions to $b_i^2 d \pm 4 = a_i^2$. We can therefore characterize the fundamental unit u as follows: let $b\ in\mathbb{N}$ be the least positive integer for which $db^2 + 4 = a^2$ or $db^2 - 4 = a^2$ for $a, a' \in \mathbb{N}$ (units $\frac{1}{2}(a + b\sqrt{d}) and \frac{1}{2}(a' + b\sqrt{d})$). Recall that the fundamental unit is the least unit with $u > 1$. Of these two possibilities, choose the unit with the smaller value of a or a'. In this case, $b = 1$ gives $d + 4 = 3^2, d - 4 = 1$. So $\frac{1}{2}(1 + \sqrt{5})$ is the fundamental unit in this case.