University of
Cambridge

Mathematics Tripos

Part II

Number Fields

Lent, 2018

Lectures by
J. A. Thorne

Notes by
Qiangru Kuang

Contents

0 Motivation **2**

1 Ring of integers **3**

2 Complex Embeddings **7**

3 Discriminants and integral bases **11**

4 Unique factorisation in $\mathcal{O}_L$ **18**

5 Dedekind's criterion **24**

6 Geometry of numbers **28**

7 Dirichlet's unit theorem **37**

8 Cyclotomic fields and the Fermat equation **43**

0 Motivation

Recall the following example from IB Groups, Rings and Modules:

Theorem 0.1. *Let p be an odd prime, then $p = a^2 + b^2$ if and only if $p = 1 \pmod 4$.*

Proof. If $p = a^2 + b^2$ then $p = 0, 1$ or $2 \pmod 4$ so this condition is necessary.

Suppose instead $p = 1 \pmod 4$, then $\left(\frac{-1}{p}\right) = 1$ so there exists $a \in \mathbb{Z}$ such that $a^2 = 1 \pmod p$, or $p \mid a^2 + 1$. We can factor $a^2 + 1 = (a + i)(a - i) \in \mathbb{Z}[i]$. We know from IB Groups, Rings and Modules that $\mathbb{Z}[i]$ is a UFD. As $p \mid (a+i)(a-i)$, if p is irreducible in $\mathbb{Z}[i]$ then $p \mid a + i$ or $p \mid a - i$. Thus $p \in \mathbb{Z}[i]$ is reducible so $p = z_1 z_2$ with $z_1 z_2 \in \mathbb{Z}[i]$. If $z_1 = A + Bi$ where $A, B \in \mathbb{Z}$ then $A^2 + B^2 = p$. □

Notation. If $R \subseteq S$ are rings and $\alpha \in S$ then

$$R[\alpha] = \left\{ \sum i = 0^n a_i \alpha^i \in s : a_i \in R \right\}$$

which is the smallest subring of S containing both R and α.

Another example is given p an odd prime, does the equation

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

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

Theorem 0.2 (Kummer, 1850)**.** *If $\mathbb{Z}[e^{2\pi i/p}]$ is a UFD then there are no solutions.*

The strategy is to factor

$$x^p + y^p = \prod_{j=0}^{p-1} (x + e^{2\pi i j/p} y) \in \mathbb{Z}[e^{2\pi i/p}].$$

We now know that $\mathbb{Z}[e^{2\pi i/p}]$ is a UFD if and only if $p \leq 19$, so unfortunately this does not lead us very far. Instead, we have the more powerful theorem

Theorem 0.3 (Kummer, 1850)**.** *If p is a* regular *prime then there are no solutions.*

We will define regular prime later in this course. This theorem is more powerful that the previous one. To give an idea, if $p < 100$ then p is regular if and only if $p \neq 37, 59, 67$.

This course studies the ring of integers of a number field, which is a finite extension of $\mathbb{Q}$. In the end of the course we will come back to Kummer's theorem.

1 Ring of integers

Recall that a field extension L/K is an inclusion $K \subseteq L$ of fields. The degree of L/K is

$$[L : K] = \dim_K L.$$

We say L/K is finite if $[L : K] < \infty$.

Definition (Number field)**.** 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 K and $f(x) \in K[x]$ be irreducible. Then $L = K[x]/(f(x))$ is a number field. Recall Tower Law from IID Galois Theory:

$$[L : \mathbb{Q}] = [L : K][K : \mathbb{Q}] < \infty.$$

Note that the first one comes with an embedding in $\mathbb{C}$, but the second one doesn't (and in general there are more than one).

Definition (Algebraic integer)**.**

1. Let L/K be a field extension. We say $\alpha \in L$ is *algebraic* over K if there exists a monic polynomial $f(x) \in K[x]$ such that $f(\alpha) = 0$.
2. Let $L/\mathbb{Q}$ be a field extension. We say $\alpha \in L$ is an *algebraic integer* if there exists a monic polynomial $f(x) \in \mathbb{Z}[x]$ such that $f(\alpha) = 0$.

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

Note that $f_\alpha(x)$ is well-defined: firstly there exists some monic $f(x) \in K[x]$ such that $f(\alpha) = 0$ since α is algebraic. If $f_\alpha(x), f'_\alpha(x) \in K[x]$ both satisfy the definition of minimal plynomial then we apply the polynomial division algorithm to write

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

where $p(x), r(x) \in K[x]$ and $\deg r < \deg f'_\alpha$. Evaluate at α, we get

$$0 = f_\alpha(\alpha) = p(\alpha)f'_\alpha(\alpha) + r(\alpha) = r(\alpha)$$

so by minimality of $\deg f'_\alpha$, $r = 0$. Then $\deg f_\alpha = \deg f'_\alpha$ and they are both monic so $p = 1$. $f_\alpha = f'_\alpha$.

Lemma 1.1. *Let $L/\mathbb{Q}$ be a field extension and let $\alpha \in L$ to 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*
>
> $$\begin{aligned}\mathbb{Z}[x] &\to L \\ f(x) &\mapsto f(\alpha)\end{aligned}$$
>
> *equals to* $(f_\alpha(x))$.

Proof.

1. Recall from IB Groups, Rings and Modules that given $f(x) = \sum_{i=0}^{n} a_i x^i \in \mathbb{Z}[x]$, we define the content to be

   $$c(f) = \gcd(a_n, \dots, a_0).$$

   Gauss' Lemma says that 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 a monic $f(x) \in \mathbb{Z}[x]$ such that $f(\alpha) = 0$. Thus $c(f) = 1$. Apply polynomial division in $\mathbb{Q}[x]$ to get

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

   where $p(x), r(x) \in \mathbb{Q}[x]$. Same as before, we must have $r(x) = 0$ so $f(x) = p(x)f_\alpha(x)$. 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

   $$nmf(x) = np(x) \cdot mf_\alpha(x).$$

   Take contents,

   $$nm = c(nmf(x)) = c(np \cdot mf_\alpha) = c(np)c(mf_\alpha) = 1.$$

   Thus $n = m = 1$ so $f_\alpha(x) \in \mathbb{Z}[x]$.

2. This is similar to the previous one. Let $g(x) \in \mathbb{Z}[x]$ be such that $g(\alpha) = 0$. wlog $g(x) \neq 0$ and $c(g) = 1$. We deduce $g(x) = q(x)f_\alpha(x)$ where $q(x) \in \mathbb{Q}[x]$. Choose $k \geq 1$ such that $kq(x) \in \mathbb{Z}[x]$ and $c(kq) = 1$. Then

   $$k = c(kg) = c(kq \cdot f_\alpha) = c(kq)c(f_\alpha) = 1$$

   so $q(x) \in \mathbb{Z}[x]$.

3. Reformulation of (2).

□

> **Corollary 1.2.** *If* $a \in \mathbb{Q}$*, then* α *is an algebraic integer if and only if* $\alpha \in \mathbb{Z}$*.*

Proof. By the above lemma, α is an algebraic integer if and only if $f_\alpha(x) \in \mathbb{Z}[x]$. If $\alpha \in \mathbb{Q}$ then $f_\alpha(x) = x - \alpha$. □

Notation. If $L/\mathbb{Q}$ is a field extension, we write

$$\mathcal{O}_L = \{\alpha \in L : \alpha \text{ is an algebraic integer}\}.$$

Proposition 1.3. *If $L/\mathbb{Q}$ is a field extension, $\mathcal{O}_L$ is a ring.*

Proof. $0, 1 \in \mathcal{O}_L$. If $\alpha \in \mathcal{O}_L$ then

$$f_{-\alpha}(x) = (-1)^{\deg f_\alpha} f_\alpha(-x)$$

so $-\alpha \in \mathcal{O}_L$. Easy. Now given $\alpha, \beta \in \mathcal{O}_L$, we need to show $\alpha + \beta, \alpha\beta \in \mathcal{O}_L$. First notice the following characterisation of algebraic integers: 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 $1, \alpha, \alpha^2, \dots$. Let

$$f_\alpha(x) = x^d + a_1 x^{d-1} + \dots + a_d \in \mathbb{Z}[x],$$

then

$$\alpha^d = -(a_1\alpha^{d-1} + \dots + a_d) \in \sum_{i=0}^{d-1} \mathbb{Z}\alpha^i.$$

Thus by induction, $\alpha^n \in \sum_{i=0}^{d-1} \mathbb{Z}\alpha^i$ for all $n \geq d$.

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 an $\mathbb{Z}$-module by $\{\alpha^i\beta^j\}_{i,j\in\mathbb{N}_0}$. The same argument shows that in fact the ring is generated as a $\mathbb{Z}$-module by $\{\alpha^i\beta^j\}_{0\leq i<d, 0\leq j<e}$. Now use classification of finitely generated $\mathbb{Z}$-modules, there is an isomorphism

$$\mathbb{Z}[\alpha, \beta] \cong \mathbb{Z}^r \oplus T$$

for some $r \geq 1$ and finite abelian group T. In fact $T = 0$: if $\gamma \in T$ then $|T|\gamma = 0$ by Lagrange. But $\mathbb{Z}[\alpha, \beta] \subseteq L$, a $\mathbb{Q}$-vector space, so this forces $\gamma = 0$. We can therefore fix an isomorphism

$$\mathbb{Z}[\alpha, \beta] \cong \mathbb{Z}^r$$

for some $r \geq 1$. Now there is a $\mathbb{Z}$-module endomorphism

$$\begin{aligned} m_{\alpha\beta} : \mathbb{Z}[\alpha, \beta] &\to \mathbb{Z}[\alpha, \beta] \\ \gamma &\mapsto \alpha\beta\gamma \end{aligned}$$

$m_{\alpha\beta}$ can be represented by an $r \times r$ matrix $A_{\alpha\beta} \in \mathcal{M}_{r\times r}(\mathbb{Z})$. Let

$$F_{\alpha\beta}(x) = \det(x \cdot I_r - A_{\alpha\beta}) \in \mathbb{Z}[x]$$

be the characteristic polynomial. Then by Cayley-Hamilton Theorem,

$$F_{\alpha\beta}(m_{\alpha\beta}) = 0.$$

Write

$$F_{\alpha\beta}(x) = x^r + b_1 x^{r-1} + \dots + b_r \in \mathbb{Z}[x]$$

so

$$m_{\alpha\beta}^r + b_1 m_{\alpha\beta}^{r-1} + \dots + b_r \cdot \mathrm{id} = 0.$$

Apply the above endomorphism to $1 \in \mathbb{Z}[\alpha, \beta]$, we get

$$(\alpha\beta)^r + b_1(\alpha\beta)^{r-1} + \dots + b_r = F_{\alpha\beta}(\alpha\beta) = 0$$

so $\alpha\beta \in \mathcal{O}_L$.

The argument to show $\alpha + \beta \in \mathcal{O}_L$ is identical, replacing $m_{\alpha\beta}$ by

$$\begin{aligned} m_{\alpha+\beta} : \mathbb{Z}[\alpha, \beta] &\to \mathbb{Z}[\alpha, \beta] \\ \gamma &\mapsto (\alpha + \beta)\gamma \end{aligned}$$

□

Definition (Ring of integers)**.** $\mathcal{O}_L$ is the *ring of algebraic integers* of L.

Lemma 1.4. *Let $L/\mathbb{Q}$ be a number field and let $\alpha \in L$. Then there exists $n \in \mathbb{Z}, n \geq 1$ 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 there exists $n \in \mathbb{Z}, n \geq 1$ such that $g(x) = n^{\deg f} f(x/n) \in \mathbb{Z}[x]$ is monic. 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** (Complex embedding)**.** 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 unique embedding $\mathbb{Q} \to \mathbb{C}$.

> **Proposition 2.1.** *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 $\sigma : L \to \mathbb{C}$ such that $\sigma|_K = \sigma_0$.*

Proof. By induction on $[L : K]$. If $[L : K] = 1$ then $L = K$.

In general, choose $\alpha \in L \setminus K$ 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 suffices to show that 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. Notice that there is an isomorphism of fields

$$K[x]/(f_\alpha(x)) \to K(\alpha)$$
$$x \mapsto \alpha$$

To get a complex embedding $\sigma : K(\alpha) \to \mathbb{C}$ extending σ_0, it's equivalent to give a root β of $(\sigma_0 f_\alpha)(x)$ in $\mathbb{C}$. We have

$$[K(\alpha) : K] = \deg f_\alpha = \deg \sigma_0 f_\alpha$$

so it suffices to show that $\sigma_0 f_\alpha$ has distinct roots in $\mathbb{C}$. The polynomial $f_\alpha(x) \in K[x]$ is irreducible so is prime to its derivative $f'_\alpha(x)$. We can therefore find $A(x), B(x) \in K[x]$ such that

$$Af_\alpha + Bf'_\alpha = 1.$$

Hence
$$(\sigma_0 A)(\sigma_0 f_\alpha) + (\sigma_0 B)(\sigma_0 f'_\alpha) = 1.$$
Hence if $\beta \in \mathbb{C}$ and $(\sigma_0 f_\alpha)(\beta) = 0$, $(\sigma_0 f'_\alpha)(\beta) \neq 0$. □

Notation. If $\sigma : L \to \mathbb{C}$ is a complex embedding, then $\overline{\sigma}$ is also a complex embedding where $\overline{\sigma}(\alpha) = \overline{\sigma(\alpha)}$. In the other words, complex conjugation is an automorphism of $\mathbb{C}$ and we can post-compose it with any field embedding.

If $\sigma = \overline{\sigma}$ then $\sigma(L) \subseteq \mathbb{R}$. Otherwise $\sigma \neq \overline{\sigma}$ and $\sigma(L)$ is not contained in $\mathbb{R}$. We write r for the number of complex embeddings σ such that $\sigma = \overline{\sigma}$ and s for the number of pairs of embeddings $\{\sigma, \overline{\sigma}\}$ where $\sigma \neq \overline{\sigma}$. It then follows that

$$r + 2s = [L : \mathbb{Q}].$$

Example (Quadratic field)**.** Let $d \in \mathbb{Z}$ be square-free and $d \neq 0, 1$. Let

$$\mathbb{Q}(\sqrt{d}) = \mathbb{Q}[x]/(x^2 - d).$$

If $d > 0$ then $r = 2, s = 0$, which we call real quadratic field. If $d < 0$ then $r = 0, s = 1$, which we call imaginary quadratic field.

Example. Let $m \in \mathbb{Z}$ be cube-free and $m \neq -1, 0, 1$. Let

$$\mathbb{Q}(\sqrt[3]{m}) = \mathbb{Q}[x]/(x^3 - m).$$

Then $r = 1, s = 1$.

Definition (Trace & norm)**.** Let L/K be a extension of number fields and let $\alpha \in L$. Let m_α be the K-linear map

$$\begin{aligned} m_\alpha : L &\to L \\ \beta &\mapsto \alpha\beta \end{aligned}$$

Then we define the *trace* of α to be

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

and the *norm* of α to be

$$\mathrm{N}_{L/K}(\alpha) = \det m_\alpha \in K.$$

Lemma 2.2. *If L/K is an extension of number fields and $\alpha \in L$, then*

1. $\operatorname{tr}_{L/K}(\alpha) = [L : K(\alpha)] \operatorname{tr}_{K(\alpha)/K}(\alpha)$.
2. $\mathrm{N}_{L/K}(\alpha) = \mathrm{N}_{K(\alpha)/K}(\alpha)^{[L:K(\alpha)]}$.

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

Lemma 2.3. *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 $\sigma_1, \dots, \sigma_n : L \to \mathbb{C}$ be complex embeddings extending σ_0. Then*

$$\sigma_0(\operatorname{tr}_{L/K}(\alpha)) = \sum_{i=1}^{n} \sigma_i(\alpha)$$
$$\sigma_0(\mathrm{N}_{L/K}(\alpha)) = \prod_{i=1}^{n} \sigma_i(\alpha)$$

Proof. wlog $L = K(\alpha)$. Let $f_\alpha(x) \in K[x]$ be the minimal polynomial of α over K. Recall that

$$(\sigma_0 f_\alpha)(x) = \prod_{i=1}^{n} (x - \sigma_i(\alpha)).$$

Write $f_\alpha(x) = x^n + a_1 x^{n-1} + \dots + a_n$. Then

$$\sigma_0(a_1) = -\sum_{i=1}^n \sigma_i(\alpha)$$
$$\sigma_0(a_n) = (-1)^n \prod_{i=1}^n \sigma_i(\alpha)$$

Let $g(x) \in K[x]$ be the characteristic polynomial of m_α. If $g(x) = x^n + b_1 x^{n-1} + \dots + b_n$ then

$$b_1 = -\operatorname{tr} m_\alpha = -\operatorname{tr}_{L/K}(\alpha)$$
$$b_n = (-1)^n \det m_\alpha = (-1)^n \operatorname{N}_{L/K}(\alpha)$$

so done if we can show $f_\alpha(x) = g(x)$. By Cayley-Hamilton, $g(m_\alpha) = 0$ so $g(\alpha) = 0$. Thus $f_\alpha(x) = g(x)$. □

Corollary 2.4. *If $\alpha \in \mathcal{O}_L$ then $\operatorname{tr}_{L/K}(\alpha), \operatorname{N}_{L/K}(\alpha) \in \mathcal{O}_K$.*

Proof. We have the following characterisation of ring of integers: if $\beta \in K$ then $\beta \in \mathcal{O}_K$ if and only if $\sigma_0(\beta) \in \mathcal{O}_\mathbb{C}$ as for all $f(x) \in \mathbb{Z}[x]$, $f(\beta) = 0$ if and only if $f(\sigma_0(\beta)) = 0$.

By the lemma, $\sigma_0 \operatorname{tr}_{L/K}(\alpha) = \sigma_1(\alpha) + \dots + \sigma_n(\alpha)$. If $\alpha \in \mathcal{O}_L$ then $\sigma_i(\alpha) \in \mathcal{O}_\mathbb{C}$ for all i. But $\mathcal{O}_\mathbb{C}$ is a ring so $\sigma_0 \operatorname{tr}_{L/K}(\alpha) \in \mathcal{O}_\mathbb{C}$. Thus $\operatorname{tr}_{L/K}(\alpha) \in \mathcal{O}_K$. Similar for norm. □

Proposition 2.5 (Classification of ring of integers of quadratic fields)**.** *Let $d \in \mathbb{Z}$ be square-free and $d \neq 0, 1$. Let $L = \mathbb{Q}(\sqrt{d})$. Then*

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

Proof. We have a nice characterisation of algebraic integers in quadratic fields: if $\alpha \in L$, then $\alpha \in \mathcal{O}_L$ if and only if $\operatorname{tr}_{L/\mathbb{Q}}(\alpha), \operatorname{N}_{L/\mathbb{Q}}(\alpha) \in \mathbb{Z}$. (Why?)

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

$$\operatorname{tr}_{L/\mathbb{Q}}(\alpha) = u \in \mathbb{Z}$$
$$\operatorname{N}_{L/\mathbb{Q}}(\alpha) = \frac{1}{4}(u + v\sqrt{d})(u - v\sqrt{d}) = \frac{1}{4}(u^2 - dv^2) \in \mathbb{Z}$$

so $u^2 - dv^2 \in 4\mathbb{Z}$, $dv^2 \in \mathbb{Z}$. Write $v = \frac{r}{s}$ where $r, s \in \mathbb{Z}$ and are coprime. Then $dr^2 \in s^2\mathbb{Z}$ so $s^2 \mid dr^2$. If p is a prime and $p \mid s$ then $p^2 \mid d$. But this is absurd as d is square-free. Thus $v \in \mathbb{Z}$.

We have shown that if $\alpha \in \mathcal{O}_L$ then $\alpha = \frac{u}{2} + \frac{v}{2}\sqrt{d}$ where $u, v \in \mathbb{Z}$ and $u^2 = dv^2 \pmod 4$. Split into cases:

1. $d = 2, 3 \pmod 4$: $u^2 = 0, 1 \pmod 4$, $v^2 = 0, 1 \pmod 4$. Consider the congruence $u^2 = dv^2 \pmod 4$ shows that $u, v \in 2\mathbb{Z}$. Hence $\alpha \in \mathbb{Z}[\sqrt{d}]$. Thus $\mathcal{O}_L = \mathbb{Z}[\sqrt{d}]$.

2. $d = 1 \pmod 4$: $u^2 = v^2 \pmod 4$ so $u = v \pmod 2$. Hence

$$\mathcal{O}_L \subseteq \left\{ \frac{u}{2} + \frac{v}{2}\sqrt{d} : u, v \in \mathbb{Z}, u = v \pmod 2 \right\} \cong \mathbb{Z} \oplus \mathbb{Z} \cdot \left(\frac{1+\sqrt{d}}{2} \right).$$

It thus remains to show that $\frac{1+\sqrt{d}}{2}$ is an algebraic integer. But we know

$$\begin{aligned} \operatorname{tr}_{L/\mathbb{Q}} \frac{1+\sqrt{d}}{2} &= 1 \\ \mathrm{N}_{L/\mathbb{Q}} \frac{1+\sqrt{d}}{2} &= \frac{1-d}{4} \in \mathbb{Z} \end{aligned}$$

so done.

□

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

$$R^\times = \{u \in R : u \text{ is a unit}\}$$

form a group under multiplication.

Lemma 2.6. *If L is a number field then*

$$\mathcal{O}_L^\times = \{\alpha \in \mathcal{O}_L : \mathrm{N}_{L/\mathbb{Q}}(\alpha) = \pm 1\}.$$

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

Proof. Norm is multiplicative so

$$\mathrm{N}_{L/\mathbb{Q}}(\alpha\beta) = \mathrm{N}_{L/\mathbb{Q}}(\alpha)\,\mathrm{N}_{L/\mathbb{Q}}(\beta)$$

for all $\alpha, \beta \in L$. If $\alpha \in \mathcal{O}_L^\times$ then there exists $\beta \in \mathcal{O}_L$ such that $\alpha\beta = 1$. Thus $\mathrm{N}_{L/\mathbb{Q}}(\alpha)\,\mathrm{N}_{L/\mathbb{Q}}(\beta) = 1$. As they are both integers,

$$\mathrm{N}_{L/\mathbb{Q}}(\alpha) \in \mathbb{Z}^\times = \{\pm 1\}.$$

Conversely, suppose $\alpha \in \mathcal{O}_L$ and $\mathrm{N}_{L/\mathbb{Q}}(\alpha) = \pm 1$. Then $\alpha^{-1} \in L$. Let $\sigma_1, \dots, \sigma_n : L \to \mathbb{C}$ be distinct complex embeddings of L. Then

$$\mathrm{N}_{L/\mathbb{Q}}(\alpha) = \prod_{i=1}^n \sigma_i(\alpha) = \pm 1$$

so

$$\sigma_1(\alpha^{-1}) = \pm \prod_{i=2}^n \sigma_i(\alpha) \in \mathcal{O}_\mathbb{C}$$

so $\alpha^{-1} \in \mathcal{O}_L$. □

3 Discriminants and integral bases

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

Definition (Discriminant)**.** Let $\alpha_1, \dots, \alpha_n \in L$. Then their *discriminant* is

$$\operatorname{disc}(\alpha_1, \dots, \alpha_n) = \det D^2$$

where $D \in \mathcal{M}_{n\times n}(\mathbb{C})$ is $D_{ij} = \sigma_i(\alpha_j)$.

Notation. Sometimes we use the alternative notation

$$\Delta(\alpha_1, \dots, \alpha_n) = \operatorname{disc}(\alpha_1, \dots, \alpha_n).$$

Note. This is independent of the choice of ordering of σ_i's and α_j's, as changing them amounts to permuting the rows and columns, which changes $\det D$ by a sign.

Lemma 3.1. *Let* $\alpha_1, \dots, \alpha_n \in L$. *Then*

$$\operatorname{disc}(\alpha_1, \dots, \alpha_n) = \det T$$

where $T \in \mathcal{M}_{n\times n}(\mathbb{Q})$ *is* $T_{ij} = \operatorname{tr}_{L/\mathbb{Q}}(\alpha_i\alpha_j)$.

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 3.2. $\operatorname{disc}(\alpha_1, \dots, \alpha_n) \in \mathbb{Q}$ *and if further* $\alpha_1, \dots, \alpha_n \in \mathcal{O}_L$ *then* $\operatorname{disc}(\alpha_1, \dots, \alpha_n) \in \mathbb{Z}$.

Proof. $\operatorname{disc}(\alpha_1, \dots, \alpha_n) = \det T \in \mathbb{Q}$.

If α_i's are in $\mathcal{O}_L$, then $D_{ij} \in \mathcal{O}_\mathbb{C}$ for all i, j. As det is a polynomial, $\operatorname{disc}(\alpha_1, \dots, \alpha_n) \in \mathcal{O}_\mathbb{C} \cap \mathbb{Q} = \mathbb{Z}$. □

Proposition 3.3. *Let* $\alpha_1, \dots, \alpha_n \in L$. *Then* $\operatorname{disc}(\alpha_1, \dots, \alpha_n) \neq 0$ *if and only if* α_i*'s form a* $\mathbb{Q}$*-basis of* L.

Proof. Suppose α_i's do not form a basis, i.e. they satisfy a non-trivial relation. Then the columns of the matrix $D_{ij} = \sigma_i(\alpha_j)$ are linearly dependent so $\operatorname{disc}(\alpha_1, \dots, \alpha_n) = 0$.

Conversely, suppose $\alpha_1, \dots, \alpha_n$ are linearly independent. Then $\operatorname{disc}(\alpha_1, \dots, \alpha_n) \neq 0$ if and only if $\det T \neq 0$, if and only if the symmetric bilinear form

$$\begin{aligned} \phi : L \times L &\to \mathbb{Q} \\ (\alpha, \beta) &\mapsto \operatorname{tr}_{L/\mathbb{Q}}(\alpha\beta) \end{aligned}$$

is non-degenerate. In other words, for all $\alpha \in L^\times$, there exists $\beta \in L$ such that $\phi(\alpha, \beta) \neq 0$. But if $\alpha \in L^\times$ then $\phi(\alpha, \alpha^{-1}) = \operatorname{tr}_{L/\mathbb{Q}}(1) = n \neq 0$. □

Definition (Integral basis)**.** We say $\alpha_1, \dots, \alpha_n \in L$ form an *integral basis* for $\mathcal{O}_L$ if

1. $\alpha_1, \dots, \alpha_n \in \mathcal{O}_L$,
2. $\alpha_1, \dots, \alpha_n$ generate $\mathcal{O}_L$ as a $\mathbb{Z}$-module.

Lemma 3.4. *If $\alpha_1, \dots, \alpha_n$ form an integral basis for $\mathcal{O}_L$ then the function*

$$\begin{aligned} f : \mathbb{Z}^n &\to \mathcal{O}_L \\ (m_1, \dots, m_n) &\mapsto \sum_{i=1}^n m_i \alpha_i \end{aligned}$$

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

Proof. f is clearly a surjective homomorphism so remains to show it is injective. Observe that $\alpha_1, \dots, \alpha_n$ form a $\mathbb{Q}$-basis of L: we know that if $\beta \in L$ then there exists $N \geq 1, N \in \mathbb{Z}$ such that $N\beta \in \mathcal{O}_L$. Write

$$N\beta = \sum_{i=1}^n m_i \alpha_i$$

for some $m_i \in \mathbb{Z}$. Thus $\beta = \sum_{i=1}^m \frac{m_1}{N} \alpha_i$. Thus α_i's span L and thus form a basis of L.

If $f(m_1, \dots, m_n) = 0$ then $\sum_{i=1}^n m_i \alpha_i = 0$ so $m_i = 0$ by linear independence of α_i's. ☐

We will soon prove that every number field has an integral basis.

Lemma 3.5 (Sandwich lemma)**.**

1. *If $H \leq G$ are abelian groups and $G \cong \mathbb{Z}^a$ for some integer $a \geq 0$, then $H \cong \mathbb{Z}^b$ for some $b \leq a$.*
2. *If $K \leq H \leq G$ are abelian groups and $K \cong \mathbb{Z}^a, G \cong \mathbb{Z}^a$ for some $a \geq 0$, then $H \cong \mathbb{Z}^a$.*
3. *If $H \leq G$ are abelian groups and $H \cong \mathbb{Z}^a, G \cong \mathbb{Z}^a$ for some $a \geq 0$ then G/H is finite.*

This is a generalisation of results about finite dimensional vector spaces (i.e. finitely generated free modules over fields) to finitely generated free $\mathbb{Z}$-modules.

Proof.

1. G/H is a finitely generated abelian group. By the classification, $G/H \cong \mathbb{Z}^n \oplus A$ where A is a finite abelian group. Choose p prime such that $p \nmid |A|$. Then the map

$$\begin{aligned} f : G/H &\to G/H \\ x + H &\mapsto px + H \end{aligned}$$

is injective. Consider the map

$$\begin{aligned} f' : H/pH &\to G/pG \\ x + pH &\mapsto x + pG \end{aligned}$$

Claim this map is also injective: if $x \in H, x \in pG$ then $x = py$ for some $y \in G$. Then $y + H \in \ker f = H$. Thus $x \in pH$.

By classification $H \cong \mathbb{Z}^b$. As f' is injective, $|H/pH| \leq |G/pG|$, i.e. $p^b \leq p^a$ so $b \leq a$.

2. Apply 1 to $K \leq H$ and $H \leq G$ to get $H \cong \mathbb{Z}^b$ where $a \leq b \leq a$ so $a = b$.
3. Again G/H is finitely generated so by classification $G/H \cong \mathbb{Z}^N \oplus A$ where A is a finite abelian group. Let p be a prime such that $p \nmid |A|$. The same proof as in 1 shows that $f' : H/pH \to G/pG$ is injective. Since $|H/pH| = |G/pG| = p^a$, f' is an isomorphism. Thus

$$G/H + pG \cong (\mathbb{Z}/p\mathbb{Z})^N$$

There is a surjective homomorphism $G/pG \to G/H + pG$ which has kernel containing the image of f'. Hence the map is surjective with kernel G/pG. This forces $N = 0$.

□

Proposition 3.6. *There exists an integral basis for $\mathcal{O}_L$.*

Proof. Let $\beta_1, \dots, \beta_n \in L$ be a $\mathbb{Q}$-basis for L. wlog $\beta_1, \dots, \beta_n \in \mathcal{O}_L$. Then $\mathcal{O}_L \supseteq \bigoplus_{i=1}^n \beta_i \mathbb{Z}$.

Recall that

$$\begin{aligned} \phi : L \times L &\to \mathbb{Q} \\ (\alpha, \beta) &\mapsto \mathrm{tr}_{L/\mathbb{Q}}(\alpha\beta) \end{aligned}$$

is a non-degenerate symmetric bilinear form. Let $\beta_1^*, \dots, \beta_n^*$ be the dual basis, i.e. $\mathrm{tr}_{L/\mathbb{Q}}(\beta_i \beta_j^*) = \delta_{ij}$. 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

$$\mathrm{tr}_{L/\mathbb{Q}}(\alpha\beta_i) = \sum_{j=1}^n \mathrm{tr}_{L/\mathbb{Q}}(a_j \beta_j^* \beta_i) = \sum_{j=1}^n a_j \, \mathrm{tr}_{L/\mathbb{Q}}(\beta_j^* \beta_i) = a_i \in \mathbb{Z}$$

so $\mathcal{O}_L \subseteq \bigoplus_{i=1}^n \beta_i^* \mathbb{Z}$. Thus by Sandwich lemma there is an isomorphism $\mathcal{O}_L \cong \mathbb{Z}^n$. □

If $\alpha_1, \dots, \alpha_n$ and $\beta_1, \dots, \beta_n$ are both integral basis for $\mathcal{O}_L$, then there exists $A \in \mathcal{M}_{n \times n}(\mathbb{Z})$ such that

$$\beta_j = \sum_{i=1}^n A_{ij} \alpha_i$$

for each $1 \leq j \leq n$. Moreover, we must have $\det A = \pm 1$ and thus $A \in \mathrm{GL}_n(\mathbb{Z})$. Let $D_{ij} = \sigma_i(\alpha_j), D'_{ij} = \sigma_i(\beta_j)$ and then $\operatorname{disc}(\beta_1, \dots, \beta_n) = \det(D')^2$. 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 thus conclude that

$$\operatorname{disc}(\beta_1, \dots, \beta_n) = \det(D')^2 = \det(DA)^2 = \det D^2 = \operatorname{disc}(\alpha_1, \dots, \alpha_n).$$

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

Proposition 3.7. *Let $L = \mathbb{Q}(\alpha)$ and let $f(x) \in \mathbb{Q}[x]$ be the minimal polynomial of α over $\mathbb{Q}$. Then*

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

Note. In IID Galois Theory, we defined

$$\operatorname{disc} f = \prod_{i<j}(\sigma_i(\alpha) - \sigma_j(\alpha))^2.$$

Proof. Let $D_{ij} = \sigma_i(\alpha^{j-1})$. Then $D \in \mathcal{M}_{n \times n}(\mathbb{C})$ and $\operatorname{disc}(1, \alpha, \dots, \alpha^{n-1}) = \det D^2$. D is a Vandermonde matrix with

$$\det D = \prod_{i<j}(\sigma_j(\alpha) - \sigma_i(\alpha)).$$

For the second equality, note that

$$\mathrm{N}_{L/\mathbb{Q}}(f'(\alpha)) = \prod_{i=1}^{n} \sigma_i(f'(\alpha)) = \prod_{i=1}^{n} f'(\sigma_i(\alpha)).$$

Also since $f(x) = \prod_{i=1}^{n}(x - \sigma_i(\alpha))$, $f'(x) = \sum_{i=1}^{n} \prod_{j \neq i}(x - \sigma_j(\alpha))$. Substitute into the above formula to get

$$\mathrm{N}_{L/\mathbb{Q}}(f'(\alpha)) = \prod_{i=1}^{n} \prod_{j \neq i}(\sigma_i(\alpha) - \sigma_j(\alpha)) = (-1)^{\binom{n}{2}} \prod_{i<j}(\sigma_i(\alpha) - \sigma_j(\alpha))^2.$$

□

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

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

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

If $d = 2, 3 \pmod 4$ then $\mathcal{O}_L = \mathbb{Z}[\sqrt{d}]$. Apply the proposition to $f(t) = t^2 - d$ to get

$$D_L = \operatorname{disc}(1, \sqrt{d}) = -\operatorname{N}_{L/\mathbb{Q}}(2\sqrt{d}) = 4d.$$

On the other hand if $d = 1 \pmod 4$ then $\mathcal{O}_L = \mathbb{Z}[\alpha]$ where $\alpha = \frac{1+\sqrt{d}}{2}$. Apply the proposition to $f(t) = t^2 - t + \frac{1-d}{4}$, $f'(t) = 2t - 1$, $f'(\alpha) = \sqrt{d}$. Thus

$$D_L = -\operatorname{N}_{L/\mathbb{Q}}(\sqrt{d}) = d.$$

Proposition 3.8. *If $\alpha_1, \dots, \alpha_n \in \mathcal{O}_L$ are such that $\operatorname{disc}(\alpha_1, \dots, \alpha_n)$ is a non-zero square-free integer then $\alpha_1, \dots, \alpha_n$ form an integral basis for $\mathcal{O}_L$.*

Proof. Let $\beta_1, \dots, \beta_n$ be an integral basis for $\mathcal{O}_L$. Then there exists $A \in \mathcal{M}_{n\times n}(\mathbb{Z})$ such that

$$\alpha_j = \sum_{i=1}^{n} A_{ij}\beta_i$$

for $1 \le j \le n$. Then

$$\operatorname{disc}(\alpha_1, \dots, \alpha_n) = \det A^2 \operatorname{disc}(\beta_1, \dots, \beta_n)$$

using a previous argument. If $\operatorname{disc}(\alpha_1, \dots, \alpha_n)$ is square-free and non-zero then $\det A = \pm 1$ so $A \in \operatorname{GL}_n(\mathbb{Z})$. Thus $\alpha_1, \dots, \alpha_n$ must generate $\mathcal{O}_L$ and thus form an integral basis. ☐

Example. Let $f(t) = t^3 - t - 1$. Use the formula

$$\operatorname{disc}(t^3 + at + b) = -4a^3 - 27b^2$$

to get $\operatorname{disc}(f) = -23$, which is square-free (and non-zero of course). If $L = \mathbb{Q}(\alpha)$ where α is a root of $f(t)$ then $\mathcal{O}_L = \mathbb{Z}[\alpha]$.

We have defined integral basis for rings of integers. In fact, we can generalise it to ideals of the ring:

Definition (Integral basis)**.** Let $I \subseteq \mathcal{O}_L$ be a non-zero ideal. Then elements $\alpha_1, \dots, \alpha_n \in L$ form an *integral basis* for I if

1. $\alpha_1, \dots, \alpha_n \in I$,
2. $\alpha_1, \dots, \alpha_n$ generate I as a $\mathbb{Z}$-module.

Proposition 3.9. *Let $I \subseteq \mathcal{O}_L$ be a non-zero ideal. Then there exists an integral basis for I.*

Proof. By definition $I \subseteq \mathcal{O}_L \cong \mathbb{Z}^n$. Let $\alpha_1, \dots, \alpha_n \in \mathcal{O}_L$ be an integral basis for $\mathcal{O}_L$. Let $a \in I$ be non-zero. Then $(a) \subseteq I$ and thus

$$\bigoplus_{i=1}^{n} a\alpha_i\mathbb{Z} \subseteq I \subseteq \mathcal{O}_L.$$

By Sandwich lemma $I \cong \mathbb{Z}^n$ as a $\mathbb{Z}$-module. Thus there exists an integral basis for I. ☐

Definition (Norm)**.** If $I \subseteq \mathcal{O}_L$ is a non-zero ideal then its *norm* is

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

Note that norm is finite by Sandwich lemma.

Definition (Discriminant)**.** If $I \subseteq \mathcal{O}_L$ is a non-zero ideal then its *discriminant* is

$$\mathrm{disc}(I) = \mathrm{disc}(\alpha_1, \dots, \alpha_n)$$

where $\alpha_1, \dots, \alpha_n$ is any integral basis for I.

Note that the same argument for discriminant of ring of integers shows that this is well-defined.

Lemma 3.10. *If $I \subseteq \mathcal{O}_L$ is a non-zero ideal then*

$$\mathrm{disc}(I) = \mathrm{disc}(\mathcal{O}_L) \cdot \mathrm{N}(I)^2.$$

Proof. Let $\alpha_1, \dots, \alpha_n$ be an integral basis for $\mathcal{O}_L$ and $\beta_1, \dots, \beta_n$ be an integral basis for I. Then there exists $A \in \mathcal{M}_{n\times n}(\mathbb{Z})$ such that

$$\beta_j = \sum_{i=1}^n A_{ij}\alpha_i$$

for $1 \leq j \leq n$ and

$$\mathrm{disc}(\beta_1, \dots, \beta_n) = \mathrm{disc}(\alpha_1, \dots, \alpha_n) \det A^2.$$

It thus suffices to show that $\det A^2 = [\mathcal{O}_L : I]^2$. In fact we'll show that if $B \in \mathcal{M}_{n\times n}(\mathbb{Z})$ and $\det B \neq 0$ then

$$|\mathbb{Z}^n / B\mathbb{Z}^n| = |\det B|.$$

Then the result follows from $\mathcal{O}_L \cong \mathbb{Z}^n$.

Proof. Recall from IB Groups, Rings and Modules that there exist $P, Q \in \mathrm{GL}_n(\mathbb{Z})$ such that

$$PBQ = D = \mathrm{diag}(d_1, \dots, d_n)$$

where $d_i \in \mathbb{Z}$ (Smith normal form). Thus

$$\mathbb{Z}^n / B\mathbb{Z}^n \cong \mathbb{Z}^n / D\mathbb{Z}^n \cong \bigoplus_{i=1}^n \mathbb{Z}/d_i\mathbb{Z}$$

so

$$|\mathbb{Z}^n / B\mathbb{Z}^n| = |\mathbb{Z}^n / D\mathbb{Z}^n| = \prod_{i=1}^n |d_i|.$$

On the other hand $|\det B| = |\det D| = \prod_{i=1}^n |d_i|$. ∎

∎

Lemma 3.11. *Let $\alpha \in \mathcal{O}_L \setminus \{0\}$. Then*

$$\mathrm{N}((\alpha)) = |\,\mathrm{N}_{L/\mathbb{Q}}(\alpha)|.$$

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

$$\begin{aligned} \operatorname{disc}(I) &= \operatorname{disc}(\alpha\alpha_1, \dots, \alpha\alpha_n) \\ &= \det(\sigma_i(\alpha\alpha_j))^2 \\ &= \det(\sigma_i(\alpha)\sigma_i(\alpha_j))^2 \\ &= \left(\prod_{i=1}^{n} \sigma_i(\alpha)\right)^2 \det(\sigma_i(\alpha_j))^2 \\ &= \mathrm{N}_{L/\mathbb{Q}}(\alpha)^2 \operatorname{disc}(\mathcal{O}_L) \end{aligned}$$

On the other hand, we showed last time that for any non-zero ideal $J \subseteq \mathcal{O}_L$,

$$\operatorname{disc}(J) = \mathrm{N}(J)^2 \operatorname{disc}(\mathcal{O}_L)$$

and the result follows. □

Notation. If $\alpha \in \mathcal{O}_L \setminus \{0\}$, we write

$$\mathrm{N}(\alpha) = \mathrm{N}((\alpha)).$$

Also define $\mathrm{N}(0) = 0$. Then for all $\alpha, \beta \in \mathcal{O}_L$, $\mathrm{N}(\alpha\beta) = \mathrm{N}(\alpha)\,\mathrm{N}(\beta)$.

In fact later we will show N is multiplicative for all ideals.

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

Recall that a ring R is a *unique factorisation domain* (UFD) if

1. R is an integral domain,
2. if $x \in R$ is non-zero and not a unit, then there exists an expression
$$x = p_1 \cdots p_r$$
where $p_i \in R$ are irreducibles. This expression is unique in the sense that if
$$x = q_1 \cdots q_s$$
is another such expressions then $r = s$ and after reordering each q_i is an associate of p_i, i.e. $q_i \in R^\times p_i$.

We know that $\mathbb{Z}$ is a UFD. However, if L is a number field then $\mathcal{O}_L$ need not be a UFD. Let's see an example where uniqueness fails.

Example. Let $L = \mathbb{Q}(\sqrt{-5})$. Then $\mathcal{O}_L = \mathbb{Z}[\sqrt{-5}]$. From example sheet we know $\mathcal{O}_L^\times = \{\pm 1\}$. In $\mathcal{O}_L$ we have
$$6 = 2 \times 3 = (1 + \sqrt{-5})(1 - \sqrt{-5}).$$
We can check that $2, 3, 1 \pm \sqrt{-5}$ are irreducibles and no two are associates. For example, suppose $2 = xy$ where $\mathrm{N}(x) > 1, \mathrm{N}(y) > 1$. As $\mathrm{N}(2) = 4$, $\mathrm{N}(x) = \mathrm{N}(y) = 2$. But $\mathrm{N}(a + b\sqrt{-5}) = a^2 + 5b^2$ which is never 2. Contradiction.

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

Proof. If $x \in \mathcal{O}_L$ then x is a non-zero non-unit if and only if $\mathrm{N}(x) > 1$. Suppose $x \in \mathcal{O}_L$ is a non-zero non-unit which cannot be written as a product of irreducibles, and with $\mathrm{N}(x)$ minimal among such elements. Then $x = yz$ with $\mathrm{N}(y), \mathrm{N}(z) > 1$, hence $\mathrm{N}(y), \mathrm{N}(z) < \mathrm{N}(x)$. By minimality of $\mathrm{N}(x)$, both y and z can be written as products of irreducibles. □

The way to get around this is to consider multiplication of ideals insteads of elements. Recall that if R is a ring and I, J are ideal of R, we can define
$$IJ = \left\{ \sum_{i=1}^k a_i b_i : a_i \in I, b_i \in J \right\}$$
$$I + J = \{a + b : a \in I, b \in J\}$$

Definition (Irreducible ideal)**.** A proper ideal $I \subsetneq R$ is *irreducible* if it does not admit an expression $I = JK$ where J, K are proper ideals of R.

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

The aim of this chapter is to prove that factorisation of ideals into prime ideals is unique. Recall from IB Groups, Rings and Modules

Definition (Prime ideal)**.** Let R is a ring. We say that a proper ideal $\mathfrak{p} \subseteq R$ is *prime* if for all $x, y \in R$, $xy \in \mathfrak{p} \implies x \in \mathfrak{p}$ or $y \in \mathfrak{p}$.

The following lemma gives us a way to characterise prime ideals:

Lemma 4.1. *Let R be a ring and $I, J, \mathfrak{p} \subseteq R$ be ideals. Suppose $\mathfrak{p}$ is prime and $IJ \subseteq \mathfrak{p}$ then $I \subseteq \mathfrak{p}$ or $J \subseteq \mathfrak{p}$.*

Proof. Wlog $I \not\subseteq \mathfrak{p}$. Choose $x \in I \setminus \mathfrak{p}$. For all $y \in J$, $xy \in IJ \subseteq \mathfrak{p}$ so $y \in \mathfrak{p}$. □

Note that the converse is trivially true, so we can think about a prime ideal as a "prime element" among all ideals, instead of breaking the ideal apart and talking about properties of elements in the ideal.

From now on let L be a number field.

Lemma 4.2. *Any non-zero prime ideal $\mathfrak{p} \subseteq \mathcal{O}_L$ is a maximal ideal.*

Proof. Recall that if R is a ring and I is a proper ideal of R, then I is prime if and only if R/I is an integral domain and I is maximal if and only if R/I is a field.

If $\mathfrak{p} \subseteq \mathcal{O}_L$ is a non-zero prime ideal, then $\mathcal{O}_L$ is a finite integral domain as its cardinality is $\mathrm{N}(\mathfrak{p})$. Any finite integral domain is a field. □

Lemma 4.3. *If $I \subseteq \mathcal{O}_L$ is a non-zero proper ideal then there exists non-zero prime ideals $\mathfrak{p}_1, \dots, \mathfrak{p}_r \subseteq \mathcal{O}_L$ such that $\mathfrak{p}_1 \cdots \mathfrak{p}_r \subseteq I$.*

Proof. For contradiction, let $I \subsetneq \mathcal{O}_L$ be an ideal which does not have this property with $\mathrm{N}(I)$ minimal among all such ideals. Clearly I is not prime so there exists $x, y \in \mathcal{O}_L$ such that $xy \in I$ but $x, y \notin I$. It follows that

$$I \subsetneq I + (x)$$
$$I \subsetneq I + (y)$$

and therefore

$$\mathrm{N}(I + (x)) < \mathrm{N}(I)$$
$$\mathrm{N}(I + (y)) < \mathrm{N}(I)$$

By minimality of $\mathrm{N}(I)$, we can find non-zero prime ideals $\mathfrak{p}_1, \dots, \mathfrak{p}_r, \mathfrak{q}_1, \dots, \mathfrak{q}_s$ such that

$$\mathfrak{p}_1 \cdots \mathfrak{p}_r \subseteq I + (x)$$
$$\mathfrak{q}_1 \cdots \mathfrak{q}_s \subseteq I + (y)$$

so

$$\mathfrak{p}_1 \cdots \mathfrak{p}_r \mathfrak{q}_1 \cdots \mathfrak{q}_s \subseteq (I + (x))(I + (y)) \subseteq I^2 + xI + yI + (xy) \subseteq I.$$

Absurd. □

Lemma 4.4. *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 $\mathfrak{p}_1, \dots, \mathfrak{p}_r \subseteq \mathcal{O}_L$ be non-zero prime ideals such that $\mathfrak{p}_1 \dots \mathfrak{p}_r \subseteq (\alpha)$. wlog r is minimal among all such expressions. Let $\mathfrak{p}$ be a maximal ideal containing I. Then

$$\mathfrak{p} \supseteq I \supseteq (\alpha) \supseteq \mathfrak{p}_1 \cdots \mathfrak{p}_r$$

so $\mathfrak{p} \supseteq \mathfrak{p}_i$ for some i. After reordering, assume $\mathfrak{p} \supseteq \mathfrak{p}_1$. Since non-zero prime ideals are maximal, we have $\mathfrak{p} = \mathfrak{p}_1$. Since r is minimal, we have $\mathfrak{p}_2 \cdots \mathfrak{p}_r \not\subseteq (\alpha)$. Choose $\beta \in \mathfrak{p}_2 \cdots \mathfrak{p}_r \setminus (\alpha)$. Claim that the element $\gamma = \frac{\beta}{\alpha}$ has the desired property: if $\gamma \in \mathcal{O}_L$ then $\beta = \alpha\gamma \in (\alpha)$. Absurd. In addition

$$\gamma I = \frac{\beta}{\alpha} I \subseteq \frac{1}{\alpha} \mathfrak{p}_2 \cdots \mathfrak{p}_r I \subseteq \frac{1}{\alpha} \mathfrak{p}_1 \cdots \mathfrak{p}_r \subseteq \mathcal{O}_L.$$

□

Proposition 4.5. *If $I \subseteq \mathcal{O}_L$ is a non-zero ideal then 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 $IJ \subseteq (\alpha)$. Suffices to show equality.

Let $K = \frac{1}{\alpha} IJ \subseteq \mathcal{O}_L$. We will show in fact that $K = \mathcal{O}_L$: if $K \neq \mathcal{O}_L$, there 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 $K = \frac{1}{\alpha} IJ \supseteq J$. Hence $\gamma J \subseteq \gamma K \subseteq \mathcal{O}_L$. We also have

$$\gamma IJ = \gamma\alpha K \subseteq (\alpha).$$

If $\beta \in \gamma J$ then $\beta \in \mathcal{O}_L$ and $\beta I \subseteq (\alpha)$ so $\gamma J \subseteq J$.

Recall that J admits an integral basis so there is an isomorphism $J \cong \mathbb{Z}^n$. Let $A \in \mathcal{M}_{n \times n}(\mathbb{Z})$ be the matrix representing multiplication by γ, with $f(x) \in \mathbb{Z}[x]$ its characteristic polynomial. Then by Cayley-Hamilton $f(\gamma) = 0$ so $\gamma \in \mathcal{O}_L$. Absurd.

This shows that $IJ = (\alpha)$. □

Now we have the machinery to define "division" of ideals:

Corollary 4.6. *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

$$\alpha J = AIJ = AIK = \alpha K$$

so $J = K$. □

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

Corollary 4.7. *If $I, J \subseteq \mathcal{O}_L$ are non-zero ideals, then $I \mid J$ if and only if $I \supseteq J$.*

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

Finally, the theorem we have promised:

Theorem 4.8. *If $I \subseteq \mathcal{O}_L$ is a non-zero ideal, then there exist prime ideals $\mathfrak{p}_1, \dots, \mathfrak{p}_r \subseteq \mathcal{O}_L$ such that*

$$I = \mathfrak{p}_1 \cdots \mathfrak{p}_r.$$

Moreover, the expression is unique up to reordering.

Proof. We show existence by contradiction. Suppose I is an ideal which cannot be written as a product of primes, and with $\mathrm{N}(I)$ minimal subject to this condition. We can find a maximal ideal $\mathfrak{p} \supseteq I$, which is also prime. Then $\mathfrak{p} \mid I$ so we can write $I = \mathfrak{p}J$ for some $J \subseteq \mathcal{O}_L$. Then $J \mid I$, hence $J \supseteq I$. If $J = I$ then we get $I = I\mathfrak{p}$ and hence $\mathcal{O}_L = \mathfrak{p}$, contradicting the maximality of $\mathfrak{p}$. Therefore $J \supsetneq I$, hence $N(J) < N(I)$. By minimality of $N(I)$, we can write $J = \mathfrak{p}_2 \cdots \mathfrak{p}_r$ where $\mathfrak{p}_i \subseteq \mathcal{O}_L$ are prime ideals. Hence

$$I = \mathfrak{p}J = \mathfrak{p}\mathfrak{p}_2 \cdots \mathfrak{p}_r.$$

Absurd.

For the uniqueness part, suppose $\mathfrak{p}_1, \dots, \mathfrak{p}_r$ and $\mathfrak{q}_1, \dots, \mathfrak{q}_s$ are non-zero ideals in $\mathcal{O}_L$ such that

$$\mathfrak{p}_1 \cdots \mathfrak{p}_r = \mathfrak{q}_1 \cdots \mathfrak{q}_s.$$

Then $\mathfrak{p}_1 \mid \mathfrak{q}_1 \cdots \mathfrak{q}_s$ so $\mathfrak{p}_1 \supseteq \mathfrak{q}_i$ for some $1 \leq i \leq s$. wlog $\mathfrak{p}_1 \supseteq \mathfrak{q}_1$. But both $\mathfrak{p}_1$ and $\mathfrak{q}_1$ are maximal so $\mathfrak{p}_1 = \mathfrak{q}_1$. Cancel to get

$$\mathfrak{p}_2 \cdots \mathfrak{p}_r = \mathfrak{q}_2 \cdots \mathfrak{q}_s.$$

Continue in this way to obtain $r = s$ and $\mathfrak{p}_i = \mathfrak{q}_i$ after reordering. □

Before going to construct prime ideals and do arithmetics on them, we first define

Definition (Ideal class group)**.** The *ideal class group* is defined to be

$$\mathrm{Cl}(\mathcal{O}_L) = \{I \subseteq \mathcal{O}_L \text{ non-zero ideal}\}/\sim$$

where $I \sim J$ if there exists $\alpha \in L^\times$ such that $\alpha I = J$.

Write $[I]$ for the equivalence class containing I.

Lemma 4.9. $\mathrm{Cl}(\mathcal{O}_L)$ *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 $\alpha, \beta \in L^\times$ are such that $\alpha I, \beta J \subseteq \mathcal{O}_L$ then

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

so the operation is well-defined.

For any $I \subseteq \mathcal{O}_L$, $\mathcal{O}_L I = I$ so $[\mathcal{O}_L]$ is the 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

$$[IJ] = [I][J] = [(\alpha)] = [\mathcal{O}_L]$$

so $[I]^{-1} = [J]$. Associativity follows from associativity of ideal multiplication. □

Proposition 4.10. *TFAE:*

1. $\mathcal{O}_L$ *is a PID.*
2. $\mathcal{O}_L$ *is a UFD.*
3. $\mathrm{Cl}(\mathcal{O}_L)$ *is trivial.*

Proof.

- $1 \implies 2$: See IB Groups, Rings and Modules.
- $2 \implies 3$: Suffices to show every ideal $I \subseteq \mathcal{O}_L$ is principal. We know that we can write

  $$I = \mathfrak{p}_1 \cdots \mathfrak{p}_r$$

  as a product of prime ideals. As products of principal ideals are principal, it suffices to show that every prime ideal of $\mathcal{O}_L$ is principal. Let $\mathfrak{p} \subseteq \mathcal{O}_L$ be a prime ideal and $\alpha \in \mathfrak{p}$ non-zero, and let

  $$\alpha = \alpha_1 \cdots \alpha_r$$

  be a factorisation of α into irreducibles. Recall that if a ring is a UFD then irreducible elements are prime. Since

  $$\mathfrak{p} \supseteq (\alpha) = (\alpha_1) \cdots (\alpha_r)$$

  so $\mathfrak{p} \mid \mathfrak{p}_1 \cdots \mathfrak{p}_r$ where $\mathfrak{p}_i = (\alpha_i)$. Since α_i's are prime, $\mathfrak{p}_i$ is a prime ideal. Hence we must have $\mathfrak{p} = \mathfrak{p}_i = (\alpha_i)$ for some i. Thus $\mathfrak{p}$ is principal.
- $3 \implies 1$: Let $I \subseteq \mathcal{O}_L$ be a non-zero ideal. Since $\mathrm{Cl}(\mathcal{O}_L)$ is trivial, we have $[I] = [\mathcal{O}_L]$, so there exists $\alpha \in L^\times$ such that $\alpha\mathcal{O}_L = I$. We have $\alpha \cdot 1 = \alpha \in I \subseteq \mathcal{O}_L$ so $\alpha \in \mathcal{O}_L$. Then $I = (\alpha)$ is principal.

□

Thus $\mathrm{Cl}(\mathcal{O}_L)$ can be seen as the obstruction to $\mathcal{O}_L$ being a UFD.

Lemma 4.11. *If $I, J \subseteq \mathcal{O}_L$ are non-zero ideals then*

$$\mathrm{N}(IJ) = \mathrm{N}(I)\,\mathrm{N}(J).$$

Proof. Example sheet. □

5 Dedekind's criterion

If $\mathfrak{p} \subseteq \mathcal{O}_L$ is a non-zero prime ideal, then there is a unique prime number $p \in \mathbb{Z}_{\geq 0}$ such that $p \in \mathfrak{p}$ since

$$(p) = \ker(\mathbb{Z} \to \mathcal{O}_L/\mathfrak{p}).$$

Then $\mathfrak{p} \mid p\mathcal{O}_L$ and $\mathrm{N}(\mathfrak{p}) = p^f$ for some $f \geq 1$.

Lemma 5.1. *Let p be a prime number and factor*

$$p\mathcal{O}_L = \prod_{i=1}^{r} \mathfrak{p}_i^{e_i}$$

where $\mathfrak{p}_1, \dots, \mathfrak{p}_r$ are distinct prime ideals of $\mathcal{O}_L$ and $e_i \geq 1$. Define $f_i \geq 1$ by $\mathrm{N}(\mathfrak{p}_i) = p^{f_i}$. Then

$$\sum_{i=1}^{r} e_i f_i = [L : \mathbb{Q}].$$

In particular, $r \leq [L : \mathbb{Q}]$.

Proof. Apply norm to get

$$p^{[L:\mathbb{Q}]} = \mathrm{N}(p\mathcal{O}_L) = \prod_{i=1}^{r} \mathrm{N}(\mathfrak{p}_i)^{e_1} = p^{\sum_{i=1}^{r} e_i f_i}.$$

□

Definition (Ramification)**.** Let p be a prime number and let $p\mathcal{O}_L = \prod_{i=1}^{r} \mathfrak{p}_i^{e_i}$ be the factorisation as above.

1. We say p *ramifies* in L if $e_i > 1$ for some i. We say p is *totally ramified* if $r = 1$ and $e_1 = [L : \mathbb{Q}]$, i.e. $p\mathcal{O}_L = \mathfrak{p}_1^{[L:\mathbb{Q}]}$.
2. We say p is *inert* in L if $r = 1$ and $e_1 = 1$, i.e. $p\mathcal{O}_L$ is prime.
3. We say p *splits completely* in L if $r = [L : \mathbb{Q}]$ and $e_i = f_i = 1$ for all i.

Theorem 5.2 (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 integer such that $p \nmid [\mathcal{O}_L : \mathbb{Z}[\alpha]]$. Let $\overline{f}(x) = f(x) \pmod{p}$ and factor*

$$\overline{f}(x) = \prod_{i=1}^{r} \overline{g}_i(x)^{e_i} \in \mathbb{F}_p[x]$$

where $\overline{g}_1(x), \dots, \overline{g}_r(x) \in \mathbb{F}_p[x]$ are distinct monic irreducible polynomials. Let $g_i(x) \in \mathbb{Z}[x]$ be any polynomial with $g_i(x) \pmod{p} = \overline{g}_i(x)$, and define

$$\mathfrak{p}_i = (p, g_i(\alpha)) \subseteq \mathcal{O}_L,$$

an ideal of $\mathcal{O}_L$. Let $f_i = \deg \overline{g}_i(x)$.

Then $\mathfrak{p}_1, \dots, \mathfrak{p}_r$ are disjoint prime ideals of $\mathcal{O}_L$ and

$$p\mathcal{O}_L = \prod_{i=1}^{r} \mathfrak{p}^{e_i}$$
$$\mathrm{N}(\mathfrak{p}_i) = p^{f_i}$$

Example. Let $L = \mathbb{Q}(\sqrt{-11})$ and $p = 5$. As $-11 = 1 \pmod 4$, $\mathcal{O}_L = \mathbb{Z}[\frac{1+\sqrt{-11}}{2}]$. Thus $\mathbb{Z}[\sqrt{-11}] \subseteq \mathcal{O}_L$ has index 2 as an additive subgroup. Therefore we can apply Dedekind's criterion to $\alpha = \sqrt{-11}$. $f(x) = x^2 + 11$.

$$\overline{f}(x) = f(x) \pmod 5 = x^2 + 1 = (x+2)(x+3) \in \mathbb{F}_5[x]$$

so $5\mathcal{O}_L = \mathfrak{p}\mathfrak{q}$ where

$$\mathfrak{p} = (5, \sqrt{-11} + 2)$$
$$\mathfrak{q} = (5, \sqrt{-11} + 3)$$

and $\mathfrak{p}, \mathfrak{q}$ are distinct prime ideals of $\mathcal{O}_L$. Thus 5 splits completely in $\mathbb{Q}(\sqrt{-11})$.

Proof. Recall that if R is a ring and $I \subseteq R$ is an ideal, then there is a bijection

$$\{\text{ideals } J \subseteq R \text{ containing } I\} \leftrightarrow \{\text{ideals } K \text{ of } R/I\}$$
$$J \mapsto J/I \subseteq R/I$$

Furthermore there is an isomorphism

$$R/J \cong (R/I)/(J/I).$$

We have $\mathbb{Z}[\alpha] \subseteq \mathcal{O}_L$ of finite index. Let $A = \mathbb{Z}[\alpha], \varphi : A \hookrightarrow \mathcal{O}_L$. By reduction mod p, get a ring homomorphism

$$\overline{\varphi} : A/pA \to \mathcal{O}_L/p\mathcal{O}_L$$
$$\beta + pA \mapsto \beta + p\mathcal{O}_L$$

Claim that $\overline{\varphi}$ is an isomorphism. Since both the domain and the codomain have the same cardinality $p^{[L:\mathbb{Q}]}$, it suffices to show $\overline{\varphi}$ 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 Lagrange, and $\beta = aN\beta + bp\beta$ so $\overline{\varphi}(aN\beta + pA) = \beta + p\mathcal{O}_L$.

Therefore

$$\{\text{ideals in } \mathcal{O}_L \text{ containing } p\} \leftrightarrow \{\text{ideals of } A/pA\}$$
$$(p) \subseteq I \leftrightarrow I \ni p$$

We have

$$A = \mathbb{Z}[\alpha] \cong \mathbb{Z}[x]/(f(x))$$
$$\alpha \leftarrow\!\shortmid x$$

Reduction mod p gives an isomorphism

$$A/pA \cong \mathbb{Z}[x]/(p, f(x)) \cong \mathbb{F}_p[x]/(\overline{f}(x)).$$

We have $\overline{f}(x) = \prod_{i=1}^r \overline{g}_i(x)^{e_i}$, so there are homomorphisms

$$\mathbb{F}_p[x]/(\overline{f}(x)) \to \mathbb{F}_p[x]/(\overline{g}_i(x))$$

given by quotienting by the ideal $(\overline{g}_i(x)) \supseteq (\overline{f}(x))$.

Define $\mathfrak{p}_i \subseteq \mathcal{O}_L$ to be the ideal containing p such that $\mathfrak{p}_i/(p)$ is the kernel of the ring homomorphism

$$\mathcal{O}_L/p\mathcal{O}_L \xrightarrow{\overline{\varphi}^{-1}} A/pA \xrightarrow{\cong} \mathbb{F}_p[x]/(\overline{f}(x)) \longrightarrow \mathbb{F}_p[x]/(\overline{g}_i(x))$$

This ring homomorphism is surjective and its image is a field of cardinality p^{f_i}. Hence $\mathcal{O}_L/\mathfrak{p}_i$ is a finite field of cardinality p^{f_i} so $\mathfrak{p}_i$ is a prime ideal of norm

$$\mathrm{N}(\mathfrak{p}_i) = p^{f_i}.$$

The $\mathfrak{p}_i$'s are distinct because their images in $\mathcal{O}_L/p\mathcal{O}_L$ are distinct, as if $i \neq j$ then $(\overline{g}_i(x), \overline{g}_j(x))$ is the unit ideal of $\mathbb{F}_p[x]$.

To show $\mathfrak{p}_i = (p, g_i(x))$, it suffices to show $\mathfrak{p}_i/(p) \subseteq \mathcal{O}_L/p\mathcal{O}_L$ is generated by $\overline{g}_i(\alpha)$. This is equivalent to showing

$$\ker(\mathbb{F}_p[x]/(\overline{f}(x)) \to \mathbb{F}_p[x]/(\overline{g}_i(x))$$

is generated by $\overline{g}_i(x)$, which is true by definition.

It remains to show

$$\mathfrak{p}_1^{e_1} \cdots \mathfrak{p}_r^{e_r} = p\mathcal{O}_L.$$

$$\begin{aligned}
\mathfrak{p}_1^{e_1} \cdots \mathfrak{p}_r^{e_r} &= (p, g_1(\alpha))^{e_1} \cdots (p, g_r(\alpha))^{e_r} \\
&\subseteq (p, g_1(\alpha)^{e_1}) \cdots (p, g_r(\alpha)^{e_r}) \\
&\subseteq (p, g_1(\alpha)^{e_1} \cdots g_r(\alpha)^{e_r}) \\
&= (p, f(\alpha)) \\
&= (p)
\end{aligned}$$

by noting that

$$(x, y)^n = (x^n, x^{n-1}y, \dots, y^n) \subseteq (x, y^n).$$

Take norm,

$$\begin{aligned}
\mathrm{N}(\mathfrak{p}_1^{e_q} \cdots \mathfrak{p}_r^{e_r}) &= \prod_{i=1}^r \mathrm{N}(\mathfrak{p}_i)^{e_i} \\
&= p^{\sum_{i=1}^r e_i f_i} \\
&= p^{\deg f} \\
&= p^{[L:\mathbb{Q}]} \\
&= \mathrm{N}(p)
\end{aligned}$$

so equality holds. ☐

One application is the classification of prime ideals in ring of integers of quadratic fields:

Proposition 5.3. *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*
   (a) *if $p \mid d$, then $(p) = \mathfrak{p}^2$ so p ramifies in L.*
   (b) *if $p \nmid d$ and $\left(\frac{d}{p}\right) = 1$ then $(p) = \mathfrak{p}\mathfrak{q}$ so p splits completely in L.*
   (c) *if $p \nmid d$ and $\left(\frac{d}{p}\right) = -1$ then (p) is prime and thus inert in L.*
2. *if $p = 2$ then*
   (a) *if $d = 2, 3 \pmod 4$ then 2 ramifies in L.*
   (b) *if $d = 1 \pmod 8$ then 2 splits completely in L.*
   (c) *if $d = 5 \pmod 8$ then 2 is inert in L.*

Proof. The case for p odd is similar to the worked example above and is left as an exercise. We just do the case for $p = 2$. If $d = 2, 3 \pmod 4$ then $\mathcal{O}_L = \mathbb{Z}[\sqrt{d}]$ so by Dedekind's criterion, we must factor $x^2 - d \pmod 2$. But

$$x^2 - d = (x - d)^2 \pmod 2.$$

If $d = 1 \pmod 4$ 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 $d = 1 \pmod 8$ then

$$x^2 + x = x(x + 1) \pmod 2.$$

If $d = 5 \pmod 8$ then the polynomial is irreducible. □

6 Geometry of numbers

Definition (Lattice)**.** If V is a finite-dimensional $\mathbb{R}$-vector space, then a *lattice* in V is a subgroup of the form

$$\Lambda = \bigoplus_{i=1}^{n} \mathbb{Z} v_i$$

where $v_1, \dots, v_n$ is a basis of V as an $\mathbb{R}$-vector space.

This is a generalisation of the ususal lattice $\mathbb{Z}^n \subseteq \mathbb{R}^n$.

Definition (Covolume)**.** If V is a finite-dimensional real inner product space and $\Lambda \subseteq V$ is a lattice, then the *covolume* of Λ is

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

where $\Lambda = \bigoplus_{i=1}^{n} \mathbb{Z} v_i$.

It is an exercise to check that it is independent of the choice of basis (that generate Λ).

We first consider only a fixed imaginary quadratic field $L = \mathbb{Q}(\sqrt{d})$ where $d < 0$ is square-free. Let $\sigma : L \to \mathbb{C}$ be a complex embedding. Our first observation is that $\sigma(\mathcal{O}_L)$ is a lattice in $\mathbb{C}$:

1. if $d = 2, 3 \pmod 4$, then $\sigma(\mathcal{O}_L) = \mathbb{Z} \oplus \mathbb{Z}\sqrt{d}$.
2. if $d = 1 \pmod 4$, then $\sigma(\mathcal{O}_L) = \mathbb{Z} \oplus \mathbb{Z}\frac{1+\sqrt{d}}{2}$.

More generally, if $I \subseteq \mathcal{O}_L$ is a non-zero ideal then $\sigma(I)$ is a lattice in $\mathbb{C}$.

Lemma 6.1. *If $I \subseteq \mathcal{O}_L$ is a non-zero ideal then*

$$A(I) = \frac{1}{2}\sqrt{|\operatorname{disc}(I)|} = \frac{\mathrm{N}(I)}{2}\sqrt{|D_L|}.$$

Proof. Let α_1, α_2 be an integral basis for I. Then

$$\sigma(I) = \mathbb{Z}\sigma(\alpha_1) \oplus \mathbb{Z}\sigma(\alpha_2).$$

If $\sigma\alpha_1 = x_1 + iy_1, \sigma\alpha_2 = x_2 + iy_2$ where x_i, y_i's are real, then

$$A(\sigma(I)) = \left| \det \begin{pmatrix} x_1 & x_2 \\ y_1 & y_2 \end{pmatrix} \right|$$

which is the area of the parallelogram spanned by the two vectors. Also by definition,

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

□

To demonstrate how to actually compute and use covolume, we state a theorem whose general version will be proved later:

Theorem 6.2 (Special case of Minkowski'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* $\operatorname{area}(S) \geq 4A(\Lambda)$ *then there exists* $\lambda \in \Lambda \setminus \{0\}$ *such that* $\lambda \in S$*.*

The surprising thing about this theorem is that it is independent of the shape of the lattice.

In particular, there exists $\lambda \in \Lambda \setminus \{0\}$ such that

$$|\lambda|^2 \leq \frac{4}{\pi} A(\Lambda).$$

Corollary 6.3. *If* $I \subseteq \mathcal{O}_L$ *is a non-zero ideal then there exists* $\alpha \in I \setminus \{0\}$ *such that*

$$\mathrm{N}(\alpha) \leq c_L \, \mathrm{N}(I)$$

where $c_L = \frac{2}{\pi}\sqrt{|D_L|}$*.*

Proof. We apply the theorem to $\sigma(I) \subseteq \mathbb{C}$ to get there exists $\lambda \in \sigma(I) \setminus \{0\}$ such that

$$|\lambda|^2 \leq \frac{4}{\pi} \frac{\mathrm{N}(I)}{2} \sqrt{|D_L|} = c_L \, \mathrm{N}(I).$$

If $\alpha \in I$ is such that $\sigma(\alpha) = \lambda$ then

$$\mathrm{N}(\alpha) = \sigma(\alpha)\overline{\sigma(\alpha)} = |\sigma(\alpha)|^2 = |\lambda|^2.$$

□

Corollary 6.4. *If* $[I] \in \mathrm{Cl}(\mathcal{O}_L)$ *then there exist* $J \in [I]$ *such that*

$$\mathrm{N}(J) \leq c_L.$$

Proof. Choose $K \in [I]^{-1}$ such that IK is principal. Apply the previous corollary to find $\alpha \in K \setminus \{0\}$ such that

$$\mathrm{N}(\alpha) \leq c_L \, \mathrm{N}(K).$$

As $(\alpha) \subseteq K$, $K \mid (\alpha)$ so there exist $J \subseteq \mathcal{O}_L$ non-zero such that $JK = (\alpha)$. Then done since $[J] = [K]^{-1} = [I]$ and

$$\mathrm{N}(J) = \frac{\mathrm{N}(\alpha)}{\mathrm{N}(K)} \leq c_L.$$

□

Finally we can prove our first result in algebraic number theory:

Theorem 6.5. *The group* $\mathrm{Cl}(\mathcal{O}_L)$ *is finite.*

In fact, we will later prove that this is true for any number field L.

Proof. We've shown that every class $[I] \in \mathrm{Cl}(\mathcal{O}_L)$ has a representative of norm $\leq c_L$. Thus suffices to show that for every $m \in \mathbb{Z}, m \geq 1$, the number of ideals $I \subseteq \mathcal{O}_L$ of norm $\mathrm{N}(I) = m$ is finite.

If $\mathrm{N}(I) = m$ then $[\mathcal{O}_L : I] = m$ so by Lagrange $m \in I$. Thus I comes from an ideal of the finite ring $\mathcal{O}_L/m\mathcal{O}_L$. □

Note. We see $\mathrm{Cl}(\mathcal{O}_L)$ is generated by ideal classes $[\mathfrak{p}]$ where $\mathfrak{p} \subseteq \mathcal{O}_L$ is a non-zero prime ideal of norm $\mathrm{N}(\mathfrak{p}) \leq c_L$. To see this, any class has the form $[I]$ where $\mathrm{N}(I) \leq c_L$. If $I = \prod_{i=1}^r \mathfrak{p}_i^{e_i}$ then

$$\begin{aligned} [I] &= \prod_{i=1}^r [\mathfrak{p}_i]^{e_i} \\ \mathrm{N}(I) &= \prod_{i=1}^r \mathrm{N}(\mathfrak{p}_i)^{e_i} \end{aligned}$$

Thus $\mathrm{N}(\mathfrak{p}_i) \leq \mathrm{N}(I) \leq c_L$.

Example.

1. $d = -7$. As $d = 1 \pmod 4$, $D_L = d$ based on our results in previous chapters. Thus
$$c_L = \frac{2}{\pi}\sqrt{7} < \frac{2}{3}\sqrt{7} < 2$$
so $\mathrm{Cl}(\mathcal{O}_L)$ is generated by ideals of norm < 2. There are none except $\mathcal{O}_L$. Thus $\mathrm{Cl}(\mathcal{O}_L)$ is trivial. Hence $\mathcal{O}_L = \mathbb{Z}[\frac{1+\sqrt{-7}}{2}]$ is a UFD.

2. $d = -5$. We already knew this is not a UFD. $D_L = 4d$ so
$$c_L = \frac{2}{\pi}\sqrt{20} = \frac{4}{\pi}\sqrt{5} < \frac{4}{3}\sqrt{5} < 3$$
so $\mathrm{Cl}(\mathcal{O}_L)$ is generated by prime ideals $\mathfrak{p} \subseteq \mathcal{O}_L$ of norm $\mathrm{N}(\mathfrak{p}) = 2$. We know by Dedekind's criterion that $2\mathcal{O}_L = \mathfrak{p}^2$. Thus $\mathrm{Cl}(\mathcal{O}_L)$ is generated by $[\mathfrak{p}]$ and $[\mathfrak{p}]^2 = [2\mathcal{O}_L] = [\mathcal{O}_L]$ is the trivial class. Hence there are two possibilities:

   (a) if $\mathfrak{p}$ is principal then $\mathrm{Cl}(\mathcal{O}_L)$ is trivial.

   (b) if $\mathfrak{p}$ is not principal then $\mathrm{Cl}(\mathcal{O}_L) \cong \mathbb{Z}/2\mathbb{Z}$.

But we already knew that $\mathcal{O}_L$ is not a UFD so $\mathrm{Cl}(\mathcal{O}_L)$ is not trivial so must have

$$\mathrm{Cl}(\mathcal{O}_L) \cong \mathbb{Z}/2\mathbb{Z}.$$

Having a grasp of the tools we have, we will now move on to a general number field L.

First we have a theorem that does not necessarily have any relation with number fields:

Theorem 6.6 (Minkowski)**.** *Let $\Lambda \subseteq \mathbb{R}^n$ be a lattice and let $E \subseteq \mathbb{R}^n$ be a measurable subset which is convex and centrally symmetric, i.e. $E = -E = \{x \in \mathbb{R}^n : -x \in E\}$. Then*

1. *if $\mathrm{vol}(E) > 2^n A(\Lambda)$, then there exists $\lambda \in \Lambda \setminus \{0\}$ such that $\lambda \in E$.*
2. *if $\mathrm{vol}(E) \geq 2^n A(\Lambda)$ and E is compact, then there exists $\lambda \in \Lambda \setminus \{0\}$ such that $\lambda \in E$.*

Note that the special case we used above corresponds to $n = 2$ and E closed disk.

Proof. Let $\Lambda = \bigoplus_{i=1}^n \mathbb{Z}v_i, P = \left\{\sum_{i=1}^n t_i v_i : t_i \in [0,1)\right\}$. Then $\mathrm{vol}(P) = A(\Lambda)$ and $\mathbb{R}^n = \bigcup_{\lambda\in\Lambda}(P + \lambda)$. Then

1. $$\begin{aligned} \mathrm{vol}(P) &< \frac{1}{2^n}\mathrm{vol}(E) \\ &= \mathrm{vol}(\frac{1}{2}E) \\ &= \sum_{\lambda\in\Lambda} \mathrm{vol}(\frac{1}{2}E \cap (\lambda + P)) \\ &= \sum_{\lambda\in\Lambda} \mathrm{vol}((\frac{1}{2}E - \lambda) \cap P) \end{aligned}$$

   Claim that there exists $\lambda, \mu \in \Lambda$ distinct such that $(\frac{1}{2}E - \lambda) \cap (\frac{1}{2}E - \mu)$ is non-empty: if not, the sets $\frac{1}{2}E - \lambda$ are pairwise disjoint so

   $$\mathrm{vol}(P) < \sum_{\lambda\in\Lambda} \mathrm{vol}((\frac{1}{2}E - \lambda) \cap P) \leq \mathrm{vol}(P),$$

   absurd. Hence there exists $z, w \in E$ such that $\frac{z}{2} - \lambda = \frac{w}{2} - \mu$. Thus

   $$\lambda - \mu = \frac{z}{2} - \frac{w}{2} = \frac{z}{2} + \frac{-w}{2}.$$

   As E is centrally symmetric, $-w \in E$. Finally as E is convex, $\frac{z}{2} + \frac{-w}{2} \in E$. Thus $\lambda - \mu \in (\Lambda \setminus \{0\}) \cap E$.

2. Given the further assumption that E is compact, E is closed and bounded. $\mathrm{vol}(E) \geq 2^n A(\Lambda)$ implies that for $m \geq 1$

   $$\mathrm{vol}((1 + \frac{1}{m})E) > 2^n A(\Lambda).$$

By 1, for all $m \in \mathbb{N}$ there exists $\lambda_m \in (\Lambda \setminus \{0\}) \cap ((1+\frac{1}{m})E)$. $(1+\frac{1}{m})E \subseteq 2E$ and $2E \cap \Lambda$ is finite as $2E$ is bounded. By pigeonhole principle, we can assume there exists $\lambda \in \Lambda \setminus \{0\}$ such that $\lambda_m = \lambda$ for all $m \geq 1$. E is closed and $\lambda \in (1+\frac{1}{m})E$ for all $m \geq 1$. Thus $\lambda \in E$.

□

Now let L be a number field. Let $n = [L : \mathbb{Q}]$ and $\tau_1, \dots, \tau_r : L \to \mathbb{R}$ be real embeddings of L and $\sigma_1, \overline{\sigma}_1, \dots, \sigma_s, \overline{\sigma}_s : L \to \mathbb{R}$ be complex embeddings. We have $n = r + 2s$.

Define a map

$$\begin{aligned} S : L &\to \mathbb{R}^r \times \mathbb{C}^s \\ \alpha &\mapsto (\tau_1(\alpha), \dots, \tau_r(\alpha), \sigma_1(\alpha), \dots, \sigma_s(\alpha)) \end{aligned}$$

This is a homomorphism of additive groups.

Lemma 6.7. *If $I \subseteq \mathcal{O}_L$ is a non-zero ideal then $S(I)$ is a lattice.*

Proof. Let $\alpha_1, \dots, \alpha_n$ be an integral basis of I. Then

$$S(I) = \bigoplus_{i=1}^{n} \mathbb{Z} S(\alpha_i)$$

and $\mathbb{R}^r \times \mathbb{C}^s$ is an n-dimensional $\mathbb{R}$-vector space. So we must show that $S(\alpha_1), \dots, S(\alpha_n)$ are independent, or equivalently that

$$\det \begin{pmatrix} \tau_1(\alpha_1) & \cdots & \tau_1(\alpha_n) \\ \vdots & & \vdots \\ \tau_r(\alpha_1) & \cdots & \tau_r(\alpha_n) \\ \operatorname{Re}\sigma_1(\alpha_1) & \cdots & \operatorname{Re}\sigma_1(\alpha_n) \\ \operatorname{Im}\sigma_1(\alpha_1) & \cdots & \operatorname{Im}\sigma_1(\alpha_n) \\ \vdots & & \vdots \\ \operatorname{Re}\sigma_s(\alpha_1) & \cdots & \operatorname{Re}\sigma_s(\alpha_n) \\ \operatorname{Re}\sigma_s(\alpha_1) & \cdots & \operatorname{Re}\sigma_s(\alpha_n) \end{pmatrix} \neq 0.$$

Note that for $z \in \mathbb{C}$, we have

$$\begin{pmatrix} z \\ \overline{z} \end{pmatrix} = \begin{pmatrix} 1 & i \\ 1 & -i \end{pmatrix} \begin{pmatrix} \operatorname{Re} z \\ \operatorname{Im} z \end{pmatrix}.$$

So this determinant equals to

$$\left(\frac{1}{-2i}\right)^s \det \begin{pmatrix} \tau_1(\alpha_1) & \cdots & \tau_1(\alpha_n) \\ \vdots & & \vdots \\ \tau_r(\alpha_1) & \cdots & \tau_r(\alpha_n) \\ \sigma_1(\alpha_1) & \cdots & \sigma_1(\alpha_n) \\ \overline{\sigma}_1(\alpha_1) & \cdots & \overline{\sigma}_1(\alpha_n) \\ \vdots & & \vdots \\ \sigma_s(\alpha_1) & \cdots & \sigma_s(\alpha_n) \\ \overline{\sigma}_s(\alpha_1) & \cdots & \overline{\sigma}_s(\alpha_n) \end{pmatrix} \neq 0$$

as $\operatorname{disc}(I) \neq 0$. □

Lemma 6.8. *If $I \subseteq \mathcal{O}_L$ is a non-zero ideal, then*

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

Proof. Same calculation with determinants as before. □

Proposition 6.9. *If $I \subseteq \mathcal{O}_L$ is a non-zero ideal then there exists $\alpha \in I \setminus \{0\}$ such that*

$$\mathrm{N}(\alpha) \leq c_L \, \mathrm{N}(I)$$

where

$$c_L = \left(\frac{4}{\pi}\right)^s \frac{n!}{n^n}\sqrt{|D_L|}.$$

Definition (Minkowski constant)**.** c_L above is called the *Minkowski constant* of L.

Proof. Apply Minkowski to the lattice $S(I)$ and the region, which might not be the most intuitive choice,

$$B_{r,s}(t) = \left\{ (\mathbf{x}, \mathbf{z}) \in \mathbb{R}^r \times \mathbb{C}^s : \sum_{i=1}^r |x_i| + 2\sum_{i=1}^s |z_i| \leq t \right\}.$$

Check that it is convex, centrally symmetric and compact. If

$$\mathrm{vol}(B_{r,s}(t)) \geq 2^n A(S(I))$$

then there exists $\alpha \in I \setminus \{0\}$ such that $S(\alpha) \in B_{r,s}(t)$.

Now we use AM-GM inequality to bound $\mathrm{N}(\alpha)$:

$$\begin{aligned}
\mathrm{N}(\alpha)^{1/n} &= \left(\prod_{i=1}^r |\tau_i(\alpha)| \prod_{i=1}^s |\sigma_i(\alpha)|^2 \right)^{1/n} \\
&\leq \frac{1}{n}\left(\sum_{i=1}^r |\tau_i(\alpha)| + 2\sum_{i=1}^s |\sigma_i(\alpha)| \right) \\
&\leq \frac{t}{n}
\end{aligned}$$

and therefore $\mathrm{N}(\alpha) \leq \frac{t^n}{n^n}$. To get the optimal bound, choose t so that

$$\mathrm{vol}(B_{r,s}(t)) = 2^n A(S(I)).$$

It is an elementary exercise to show that

$$\mathrm{vol}(B_{r,s}(t)) = 2^r \left(\frac{\pi}{2}\right)^s \frac{t^n}{n!}$$

by induction on r and s. Thus

$$2^r \left(\frac{\pi}{2}\right)^s \frac{t^n}{n!} = 2^n A(S(I)) = 2^{r+s}\,\mathrm{N}(I)\sqrt{|D_L|}.$$

Rearrange,

$$\mathrm{N}(\alpha) \leq \frac{t^n}{n^n} = c_L \, \mathrm{N}(I).$$

□

Similar corollaries:

Corollary 6.10. *For any class $[I] \in \mathrm{Cl}(\mathcal{O}_L)$, there exists $J \in [I]$ such that*

$$\mathrm{N}(J) \leq c_L.$$

Corollary 6.11. *The group $\mathrm{Cl}(\mathcal{O}_L)$ is finite and generated by $[\mathfrak{p}]$ where $\mathfrak{p}$ is a prime ideal of norm $\mathrm{N}(\mathfrak{p}) \leq c_L$.*

Proof. Exactly the same as before. □

Remark. In practice, the Minkowski constant is a very effective bound.

Example. Let $f(x) = x^5 - x + 1$. This is irreducible modulo 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 discrimiant is

$$\operatorname{disc} f = 2869 = 19 \cdot 151$$

which is square-free. Thus $\mathcal{O}_L = \mathbb{Z}[\alpha]$ and $D_L = \operatorname{disc} f$. Thus

$$c_L = \left(\frac{4}{\pi}\right)^2 \frac{5!}{5^5} \sqrt{2869} < 4.$$

Hence $\mathrm{Cl}(\mathcal{O}_L)$ is generated by prime ideals $\mathfrak{p}$ of norm $\mathrm{N}(\mathfrak{p}) = 2$ or 3. But by Dedekind's criterion, such primes exists if and only if $f(x)$ has a root in $\mathbb{F}_2$ or $\mathbb{F}_3$. In this case there are not such roots so $\mathrm{Cl}(\mathcal{O}_L)$ is trivial so $\mathbb{Z}[\alpha]$ is a UFD.

Example. Let $L = \mathbb{Q}(\sqrt{10})$. Then

$$c_L = \frac{1}{2}\sqrt{4 \cdot 10} = \sqrt{10} < 4.$$

Thus $\mathrm{Cl}(\mathcal{O}_L)$ is generated by $[\mathfrak{p}]$ where $\mathrm{N}(\mathfrak{p}) = 2$ or 3. By Dedekind's criterion,

$$(2) = \mathfrak{p}_2^2$$
$$(3) = \mathfrak{p}_3 \mathfrak{p}_3'$$

where

$$\mathfrak{p}_2 = (2, \sqrt{10})$$
$$\mathfrak{p}_3 = (3, 1 + \sqrt{10})$$
$$\mathfrak{p}_3' = (3, 1 - \sqrt{10})$$

To find relations in $\mathrm{Cl}(\mathcal{O}_L)$, we can calculate the norm. For example,

$$\mathrm{N}(2 + \sqrt{10}) = |4 - 10| = 6$$

so

$$(2+\sqrt{10}) = \mathfrak{p}_2\mathfrak{p}_3 \text{ or } \mathfrak{p}_2\mathfrak{p}_3'.$$

In either case we see that $[\mathfrak{p}_2]$ generates $\mathrm{Cl}(\mathcal{O}_L)$ so $\mathrm{Cl}(\mathcal{O}_L)$ is either trivial or isomorphic to $\mathbb{Z}/2\mathbb{Z}$, with the second case occurring if and only if $\mathfrak{p}_2$ is not principal. $\mathfrak{p}_2$ is principal if and only if there exists $a+b\sqrt{10} \in \mathcal{O}_L$ such that $(a+b\sqrt{10}) = \mathfrak{p}_2$. Taking norm,

$$a^2 - 10b^2 = \pm 2.$$

Reduce modulo 5, neither 2 or -2 is a quadratic residue. Absurd. Thus $\mathfrak{p}_2$ is not principal and

$$\mathrm{Cl}(\mathcal{O}_L) \cong \mathbb{Z}/2\mathbb{Z}.$$

Example. Let $L = \mathbb{Q}(\sqrt{-17})$. Then

$$c_L = \frac{4}{\pi} \cdot \frac{1}{2} \cdot \sqrt{4 \cdot 17} = \frac{4}{\pi}\sqrt{17} < \frac{4}{3}\sqrt{17} < 6$$

So $\mathrm{Cl}(\mathcal{O}_L)$ is generated by primes of norm $2, 3$ or 5.

By Dedekind's criterion,

- $x^2 + 17 = x^2 + 2 \pmod 5$ so (5) is prime of norm 25.
- $x^2 + 17 = x^2 - 1 \pmod 3$ so

$$(3) = \mathfrak{q}_3\mathfrak{q}_3'$$

  where

$$\begin{aligned} \mathfrak{q}_3 &= (3, 1+\sqrt{-17}) \\ \mathfrak{q}_3' &= (3, 1-\sqrt{-17}) \end{aligned}$$

- $x^2 + 17 = (x+1)^2 \pmod 2$ so

$$(2) = \mathfrak{q}_2^2$$

  where

$$\mathfrak{q}_2 = (2, 1+\sqrt{-17}).$$

Now compute, for example, the norm

$$\mathrm{N}(1+\sqrt{-17}) = 18 = 2 \cdot 3^2.$$

Note that $1+\sqrt{-17} \in \mathfrak{q}_3$ so $\mathfrak{q}_3 \mid (1+\sqrt{-17})$. So we must have one of

$$\begin{aligned} (1+\sqrt{-17}) &= \mathfrak{q}_2\mathfrak{q}_3\mathfrak{q}_3' \\ (1+\sqrt{-17}) &= \mathfrak{q}_2\mathfrak{q}_3^2 \end{aligned}$$

Note that they result in different structures of $\mathrm{Cl}(\mathcal{O}_L)$. To decide between these, we compute

$$\begin{aligned} \mathfrak{q}_3^2 &= (9, 3+3\sqrt{-17}, (1+\sqrt{-17})^2) \\ &= (9, 3+3\sqrt{-17}, -16+2\sqrt{-17}) \\ &= (9, 3+3\sqrt{-17}, 2+2\sqrt{-17}) \\ &= (9, 1+\sqrt{-17}) \end{aligned}$$

We see $1+\sqrt{-17} \in \mathfrak{q}_3^2$ so we have $(1+\sqrt{-17}) = \mathfrak{q}_2\mathfrak{q}_3^2$.[1]

We see $[\mathfrak{q}_3]$ generates $\mathrm{Cl}(\mathcal{O}_L)$ and if $\mathfrak{q}_2$ is not principal then $\mathrm{Cl}(\mathcal{O}_L) \cong \mathbb{Z}/4\mathbb{Z}$. But $\mathfrak{q}_2$ is principal if and only if we can solve

$$a^2 + 17b^2 = 2$$

in $\mathbb{Z}$. This is impossible so

$$\mathrm{Cl}(\mathcal{O}_L) \cong \mathbb{Z}/4\mathbb{Z}.$$

Remark. There are many open questions about ideal class groups, even for quadratic fields.

- We know: $|\mathrm{Cl}(\mathcal{O}_{\mathbb{Q}(\sqrt{d})})| \to \infty$ as $d \to -\infty$ through square-free integers. In particular, there are only finitely many imaginary quadratic fields of given cardinality. For example, there are exactly 9 imaginary quadratic fields with trivial ideal class group. See example sheet for the existence (the uniqueness part is much more difficult).
- 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 A be a finite abelian group of p-power order. Then for $d < 0$ square-free, conjecture that

  $$\mathbb{P}(\mathrm{Cl}(\mathcal{O}_{\mathbb{Q}(\sqrt{d})})_p \cong A) = \frac{\prod_{i=1}^{\infty}(1-\frac{1}{p^i})}{|\mathrm{Aut}(A)|}.$$

  where for a finite abelian group M, M_p is the (unique) p-Sylow subgroup and the probability on LHS is defined to be

  $$\lim_{x\to\infty} \frac{|\{d<0 : |d| < x, d \text{ square-free}, \mathrm{Cl}(\mathcal{O}_{\mathbb{Q}(\sqrt{d})})_p \cong A\}|}{|\{d<0 : |d|<x, d \text{ square-free}\}|}.$$

[1] In this specific case, one can take a shortcut by noting that $\mathfrak{q}_3\mathfrak{q}_3' = (3)$ which does not divide $(1+\sqrt{-17})$.

7 Dirichlet's unit theorem

Let L be a number field of degree n and Let $\tau_1, \dots, \tau_r : L \to \mathbb{R}$ be real embeddings, $\sigma_1, \dots, \sigma_s, \overline{\sigma}_1, \dots, \overline{\sigma}_s : L \to \mathbb{C}$ be distinct complex embeddings.

Theorem 7.1 (Dirichlet's unit theorem)**.** *There is an isomorphism*

$$\mathcal{O}_L^\times \cong \mu_L \times \mathbb{Z}^{r+s-1}$$

where $\mu_L \subseteq \mathcal{O}_L^\times$ is the finite cyclic group of roots of unity in $\mathcal{O}_L^\times$.

In fact, the proof shows more: define a map $\ell : \mathcal{O}_L^\times \to \mathbb{R}^{r+s}$ by

$$\alpha \mapsto (\log|\tau_1(\alpha)|, \dots, \log|\tau_r(\alpha)|, 2\log|\sigma_1(\alpha)|, \dots, 2\log|\sigma_s(\alpha)|).$$

Then this is a homomorphism of abelian groups, and $\ell(\mathcal{O}_L^\times)$ 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 implies that if $\alpha \in \mathcal{O}_L^\times$ then

$$\log \mathrm{N}(\alpha) = \sum_{i=1}^{r} \log|\tau_i(\alpha)| + 2\sum_{i=1}^{s} \log|\sigma_i(\alpha)| = 0.$$

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

Example. $\mathcal{O}_L^\times$ is finite if and only if $r + s = 1$, i.e.

- $r = 1, s = 0$, so $L = \mathbb{Q}$.
- $r = 0, s = 1$, so $L = \mathbb{Q}(\sqrt{d})$ where $d < 0$ is square-free.

The first case where $\mathcal{O}_L^\times$ is infinite is $L = \mathbb{Q}(\sqrt{d})$, $d > 0$ square-free. Then $r + s - 1 = 1$, so $\ell(\mathcal{O}_L^\times)$ is infinite cyclic. Fix $\sigma : \mathbb{Q}(\sqrt{d}) \to \mathbb{R}$ to be the real embedding with $\sigma(\sqrt{d}) > 0$. $\sigma(\mu_L) \subseteq \mathbb{R}^\times$ so $\mu_L = \{\pm 1\}$. In this case we can consider the map

$$\begin{aligned} \ell' : \mathcal{O}_L^\times &\to \mathbb{R} \\ \alpha &\mapsto \log|\sigma(\alpha)| \end{aligned}$$

We know that $\ell'(\mathcal{O}_L^\times) \subseteq \mathbb{R}$ is a lattice. In particular, there is a unique characterised unit $\alpha \in \mathcal{O}_L^\times$ satisfying $\sigma(\alpha) > 0, \log|\sigma(\alpha)| > 0$ and as small as possible. In other words, $\alpha \in \mathcal{O}_L^\times$ 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

$$\mathcal{O}_L^\times = \{\pm\alpha^n : n \in \mathbb{Z}\}.$$

How to find fundamental units?

Lemma 7.2.

1. *If $d = 2, 3 \pmod 4$ and $v \in \mathcal{O}_L^\times$ satisfies $v > 1$, then*

$$v = a + b\sqrt{d}$$

where $a \geq b \geq 1$.

2. *If* $d = 1 \pmod 4$ *and* $v \in \mathcal{O}_L^\times$ *satisfies* $v > 1$, *then*

$$v = \frac{1}{2}(a + b\sqrt{d})$$

where $a \geq b \geq 1$.

Proof.

1. Let $v' = a - b\sqrt{d}$. Then

$$vv' = a^2 - db^2 = \mathrm{N}_{L/\mathbb{Q}}(v) = \pm 1$$

so $v > 1$ implies that $|v'| < 1$. Hence

$$\begin{aligned} v + v' &= 2a > 0 \\ v - v' &= 2b\sqrt{d} > 0 \end{aligned}$$

As a, b are integers, we must have $a \geq 1, b \geq 1$. Also

$$\left(\frac{a}{b}\right)^2 = d \pm \frac{1}{b^2} \geq 1$$

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

$$\begin{aligned} v + v' &= a > 0 \\ v - v' &= b\sqrt{d} > 0 \end{aligned}$$

so $a \geq 1, b \geq 1$. Also

$$\left(\frac{a}{b}\right)^2 = d \pm \frac{4}{b^2} \geq 1$$

as $d \geq 5$.

□

We can use this to find the fundamental unit in a quadratic field $\mathbb{Q}(\sqrt{d})$ where $d \in \mathbb{Z}$ is positive square-free.

1. $d = 2, 3 \pmod 4$: let $u = a + b\sqrt{d}$. Let $u^k = a_k + b_k\sqrt{d}$. Then we have the relation

$$\begin{aligned} u^{k+1} &= u \cdot u^k \\ &= (a_1 + b_1\sqrt{d})(a_k + b_k\sqrt{d}) \\ &= (a_1a_k + db_1b_k) + (b_1a_k + a_1b_k)\sqrt{d} \end{aligned}$$

Hence

$$b_{k+1} = b_1a_k + a_1b_k > b_k$$

so the sequence $(b_k)_{k\in\mathbb{N}}$ is strictly increasing.

We can therefore characterise u as follow: 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 $u = a + b\sqrt{d}$.

2. $d = 1 \pmod 4$: let $u = \frac{1}{2}(a + b\sqrt{d})$. 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 if and only if $a_k = b_k$ and $a_1 = b_1 = 1$. Note that if $a_1 = b_1 = 1$ then

$$\mathrm{N}(u) = \left| \frac{1-d}{4} \right| = 1$$

so $d = 5$.

   (a) $d > 5$: the sequence $(b_k)_{k \in \mathbb{N}}$ is strictly increasing. The fundamental unit can therefore be found as follow: let $b \in \mathbb{N}$ be the least integer such that $db^2 + 4$ or $db^2 - 4$ is of the form a^2 for some $a \in \mathbb{N}$. Then $\frac{1}{2}(a + b\sqrt{d})$ is the fundamental unit.

   (b) $d = 5$: the sequence $(b_k)_{k \in \mathbb{N}}$ is non-decreasing and each value b_i can appear at most twice (as occurrence corresponds to solutions to $db_i^2 \pm 4 = a_i^2$). We can therefore characterise the fundamental unit u as follow: let $b \in \mathbb{N}$ be the least positive integer for which $db^2 + 4 = a^2$ or $db^2 - 4 = a'^2$ for $a, a' \in \mathbb{N}$. 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 $d + 4 = 3^2, d - 4 = 1^2$ so $\frac{1}{2}(1 + \sqrt{5})$ is the fundamental unit.

Example.

1. $d = 2$. Then $b = 1$ works since $2 - 1 = 1^2$ so $1 + \sqrt{2}$ is a fundamental unit.
2. $d = 7$.

$$\begin{aligned} b &= 1 : 7 \pm 1 \text{ not a square} \\ b &= 2 : 4 \cdot 7 \pm 1 \text{ not a square} \\ b &= 3 : 9 \cdot 7 + 1 = 8^2 \end{aligned}$$

   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 fraction, but it is not discussed in this course.

Now we start the proof of Dirichlet's unit theorem, which in non-examinable.

Proof of Dirichlet's unit theorem. Recall the setup: let L be a number field, $\tau_1, \dots, \tau_r : L \to \mathbb{R}$ are the real embeddings and $\sigma_1, \overline{\sigma}_1, \dots, \sigma_s, \overline{\sigma}_s : L \to \mathbb{C}$ are the complex embeddings of L. Define a map $\ell : \mathcal{O}_L^\times \to \mathbb{R}^{r+s}$ by

$$\alpha \mapsto (\log|\tau_1(\alpha)|, \dots, \log|\tau_r(\alpha)|, 2\log|\sigma_1(\alpha)|, \dots, 2\log|\sigma_s(\alpha)|).$$

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. *Extend ℓ to $\mathcal{O}_L \setminus \{0\}$. Let $\alpha \in \mathcal{O}_L \setminus \{0\}$ be such that $\ell(\alpha) = (a_1, \dots, a_{r+s})$. Fix an integer $1 \leq k \leq r+s$. Then there exists $\beta \in \mathcal{O}_L \setminus \{0\}$ such that if $\ell(\beta) = (b_1, \dots, b_{r+s}) \in \mathbb{R}^{r+s}$ then $b_i < a_i$ if $i \neq k$. Moreover,*

$$\mathrm{N}(\beta) \leq \left(\frac{2}{\pi}\right)^s \sqrt{|D_L|}.$$

Proof. This proof is similar to the derivation of Minkowski constant but using a slightly different convex body. Let $c_1, \dots, c_{r+s} \in \mathbb{R}_{>0}$ and let

$$E = \{(\mathbf{x}, \mathbf{z}) \in \mathbb{R}^r \times \mathbb{C}^s : |x_i| \leq c_i, |z_i|^2 \leq c_{r+i}\}.$$

Then if $\mathrm{vol}(E) \geq 2^{r+2s} A(S(\mathcal{O}_L)) = 2^{r+s}\sqrt{|D_L|}$, then by Minkowski there exists $\beta \in \mathcal{O}_L \setminus \{0\}$ such that $S(\beta) \in E$. In particular,

$$\mathrm{N}(\beta) = \prod_{i=1}^{r} |\tau_i(\beta)| \prod_{i=1}^{s} |\sigma_i(\beta)|^2 \leq \prod_{i=1}^{r+s} c_i.$$

We choose c_i so that $0 < c_i < e^{a_i}$ if $i \neq k$ and

$$\mathrm{vol}(E) = \pi^s 2^r \prod_{i=1}^{r+s} c_i = 2^{r+s}\sqrt{|D_L|}.$$

The first property gives $b_i < a_i$ if $i \neq k$ while the second gives

$$\mathrm{N}(\beta) \leq \prod_{i=1}^{r+s} c_i = \left(\frac{2}{\pi}\right)^s \sqrt{|D_L|}.$$

□

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

Proof. By the lemma we can find elements $\alpha_1, \alpha_2, \dots \in \mathcal{O}_L \setminus \{0\}$ such that

$$\mathrm{N}(\alpha_i) \leq \left(\frac{2}{\pi}\right)^s \sqrt{|D_L|}$$

for all $i \in \mathbb{N}$ and if $\ell(\alpha_i) = (b_{i,1}, \dots, b_{i,r+s})$ then $b_{i+1,j} < b_{i,j}$ if $j \neq k$ for all $i \geq 1$. The ideals (α_i) have bounded norm, so are finite in number. So there exist $N < M$ such that $(\alpha_N) = (\alpha_M)$. Then the element

$$\varepsilon = \frac{\alpha_N}{\alpha_M}$$

has the desired property. □

Lemma 7.5. *Let $N \geq 1$ and $A \in \mathcal{M}_{N \times N}(\mathbb{R})$ be such that*

1. *$\sum_{i=1}^N A_{ij} = 0$ for $1 \leq j \leq N$,*
2. *$A_{ij} > 0$ if $i = j$ and $A_{ij} < 0$ if $i \neq j$,*

then A has rank $N-1$.

Proof. The rank is at most $N-1$. We show that the first $N-1$ rows of A are independent. Suppose there exists $t_i \in \mathbb{R}$ for $1 \le i < N$, not all zero, such that

$$\sum_{i=1}^{N-1} t_i A_{ij} = 0$$

for $1 \le j \le N$. wlog after rescaling, there exists $1 \le k < N$ such that $t_k = 1$ and $t_i \le 1$ if $i \ne k$. Then

$$0 = \sum_{i=1}^{N-1} t_i A_{ik} \ge \sum_{i=1}^{N-1} A_{ik} > \sum_{i=1}^{N} A_{ik} = 0$$

Absurd. □

Lemma 7.6. *Fix $B > 0$. Let*

$$X_B = \{\alpha \in \mathcal{O}_L : \forall \sigma : L \to \mathbb{C}, |\sigma(\alpha)| \le B\},$$

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 $\mathbb{R}^r \times \mathbb{C}^s$ so must be finite. □

Finally we get something we promised earlier:

Proposition 7.7. *$\ell(\mathcal{O}_L^\times)$ form a lattice in $H \le \mathbb{R}^{r+s}$.*

Proof. We must show that there exist units $v_1, \dots, v_{r+s-1} \in \mathcal{O}_L^\times$ such that their images under ℓ span H as an $\mathbb{R}$-vector space and generate $\ell(\mathcal{O}_L^\times)$ as an abelian group.

By Corollary 7.4, we can find $\varepsilon_1, \dots, \varepsilon_{r+s} \in \mathcal{O}_L^\times$ such that if $\ell(\varepsilon_j) = (A_{1,j}, \dots, A_{r+s,j})$ then $A_{i,j} < 0$ if $i \ne j$ and $A_{i,j} > 0$ if $i = j$. By Lemma 7.5, the matrix A has rank $r+s-1$ so we can find $v_1, \dots, v_{r+s-1} \in \mathcal{O}_L^\times$ such that $\ell(v_1), \dots, \ell(v_{r+s-1})$ span H as an $\mathbb{R}$-vector space.

Let $\Lambda = \bigoplus_{i=1}^{r+s-1} \mathbb{Z}\ell(v_i) \le H$ which is a lattice. Then $\Lambda \le \ell(\mathcal{O}_L^\times)$ and if $u \in \mathcal{O}_L^\times$ then there exists $\lambda \in \Lambda$ such that

$$\ell(u) - \lambda \in \left\{ \sum_{i=1}^{r+s-1} t_i \ell(v_i) : t_i \in [0,1) \text{ for all } 1 \le i \le r+s-1 \right\} = P.$$

But the set of units in $\ell^{-1}(P)$ is finite by Lemma 7.6. Hence the quotient $\ell(\mathcal{O}_L^\times)/\Lambda$ is finite. By Lagrange, there exists $N \in \mathbb{Z}, N \ge 1$ such that $N \cdot \ell(\mathcal{O}_L^\times) \le \Lambda$. Hence

$$\Lambda \le \ell(\mathcal{O}_L^\times) \le \frac{1}{N}\Lambda.$$

By sandwich lemma, $\ell(\mathcal{O}_L^\times)$ is a free abelian group of rank $r+s-1$. In particular, it is a lattice in H. □

Let's now finish the proof the unit theorem, i.e. show there is an isomorphism

$$\mathcal{O}_L^\times \cong \mu_L \times \mathbb{Z}^{r+s-1}$$

where μ_L is the (finite) group of roots of unity in L. Note that $\mu_L = \ker \ell$: if $\zeta \in \mu_L$ then $\zeta^N = 1$ for some $N \geq 1$. Hence $\ell(\zeta^N) = N \cdot \ell(\zeta) = 0$. As $\ell(\zeta) \in \mathbb{R}^{r+s}$, a vector space, we have $\ell(\zeta) = 0$. Conversely, if $\alpha \in \mathcal{O}_L^\times$ and $\ell(\alpha) = 0$ then for all $\sigma : L \to \mathbb{C}, |\sigma(\alpha)| = 1$. By Lemma 7.6 $\ker \ell$ is finite. By Lagrange it consists of roots of unity.

Choose $v_1, \dots, v_{r+s-1} \in \mathcal{O}_L^\times$ such that $\ell(v_1), \dots, \ell(v_{r+s-1})$ is a $\mathbb{Z}$-basis of $\ell(\mathcal{O}_L^\times)$. Define a map

$$\begin{aligned} \mu_L \times \mathbb{Z}^{r+s-1} &\to \mathcal{O}_L^\times \\ (\zeta, n_1, \dots, n_{r+s-1}) &\mapsto \zeta v_1^{n_1} \cdots v_{r+s-1}^{n_{r+s-1}} \end{aligned}$$

It is an exercise to check this is an isomorphism. □

8 Cyclotomic fields and the Fermat equation

An warm-up exercise:

Question. Find all Pythagorean triples $x^2 + y^2 = z^2$ where $x, y, z \in \mathbb{Z}$ not all zero.

wlog $\gcd(x, y, z) = 1$. Consider the parity: if 2 divides both x and y then 2 divides z, so assume x is odd, y is even. The idea is to factor the equation in $\mathbb{Z}[i]$ to get

$$(x + iy)(x - iy) = z^2.$$

Claim that the ideals $(x + iy)$ and $(x - iy)$ of $\mathbb{Z}[i]$ are coprime, i.e. there is no prime ideal $\mathfrak{p} \in \mathbb{Z}[i]$ which divides both of them: if $\mathfrak{p}$ divides both then $\mathfrak{p} \mid (2x), \mathfrak{p} \mid (2y)$. If ℓ is an odd prime such that $\ell \mid \mathrm{N}(\mathfrak{p})$ then this implies $\ell \mid 2x, \ell \mid 2y$ so $\ell \mid x$ and $\ell \mid y$, impossible. Thus $\mathfrak{p} \mid (2)$, hence $\mathfrak{p} \mid (z^2)$, so $2 \mid z$, absurd. Thus there is no such prime $\mathfrak{p}$.

Using the identity $(x + iy)(x - iy) = (z)^2$, we see that $(x + iy)$ must be the square of another ideal. Using the fact that $\mathbb{Z}[i]$ is a UFD, we get

$$(x + iy) = (a + ib)^2 = (a^2 - b^2 + 2abi)$$

where $a, b \in \mathbb{Z}$. Hence $x + iy = u(a^2 - b^2 + 2abi)$ for some $u \in \mathbb{Z}[i]^\times = \{\pm 1, \pm i\}$. It is left as an exercise to show that there exists $A, B \in \mathbb{Z}$ such that

$$\begin{aligned} x &= A^2 - B^2 \\ y &= 2AB \\ z &= A^2 + B^2 \end{aligned}$$

The aim of this section is to do something similar for

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

where p is an odd prime.

From now on p is an odd prime.

Definition (Cyclotomic field)**.** The *pth cyclotomic field* is $K = \mathbb{Q}(\zeta_p)$ where $\zeta_p = e^{2\pi i/p} \in \mathbb{C}$.

Lemma 8.1.

1. *$(1 - \zeta_p)^{p-1} = (p)$ in $\mathcal{O}_K$, $\mathrm{N}(1 - \zeta_p) = p$ and $(1 - \zeta_p) \subseteq \mathcal{O}_K$ is a prime ideal.*
2. *Let $f_p(x) = \frac{x^p - 1}{x - 1} \in \mathbb{Z}[x]$. Then $f_p(x)$ is irreducible and $[K : \mathbb{Q}] = p - 1$.*

Proof. We can factorise

$$f_p(x) = \prod_{j=1}^{p-1} (x - \zeta_p^i).$$

In particular, $f_p(\zeta_p) = 0$ and $[K : \mathbb{Q}] \leq p - 1$. We also have

$$f_p(1) = p = \prod_{j=1}^{p-1}(1 - \zeta_p^j).$$

Claim that for $1 \leq j < p$, we have $(1 - \zeta_p^j) = (1 - \zeta_p)$ as ideals of $\mathcal{O}_K$. We show this by exhibiting inclusion both ways:

$$\frac{1 - \zeta_p^j}{1 - \zeta_p} = 1 + \zeta_p + \dots + \zeta_p^{j-1} \in \mathcal{O}_K$$

so $1 - \zeta_p^j \in (1 - \zeta_p)$. Choose $k \in \mathbb{Z}, k \geq 1$ such that $jk = 1 \pmod p$, then

$$\frac{1 - \zeta_p}{1 - \zeta_p^j} = \frac{1 - \zeta_p^{jk}}{1 - \zeta_p^j} = 1 + \zeta_p^j + \dots + \zeta_p^{j(k-1)} \in \mathcal{O}_K$$

so $1 - \zeta_p \in (1 - \zeta_p^j)$.

Thus $(p) = (1-\zeta_p)^{p-1}$ is an ideal in $\mathcal{O}_K$. It follows that $p^{[K:\mathbb{Q}]} = \mathrm{N}(1-\zeta_p)^{p-1}$. But since we already know $[K : \mathbb{Q}] \leq p - 1$, we must have $\mathrm{N}(1 - \zeta_p) = p$ and $[K : \mathbb{Q}] = p - 1$. □

Lemma 8.2.

$$\mathrm{disc}(1, \zeta_p, \dots, \zeta_p^{p-2}) = (-1)^{\frac{p-1}{2}} p^{p-2}.$$

Proof. We know

$$\mathrm{disc}(1, \zeta_p, \dots, \zeta_p^{p-2}) = (-1)^{\binom{p-1}{2}} \mathrm{N}_{K/\mathbb{Q}}(f_p'(\zeta_p)),$$

but

$$f_p'(x) = \frac{(x-1)px^{p-1} - (x^p - 1)}{(x-1)^2}$$

so $f_p'(\zeta_p) = \frac{p\zeta_p^{-1}}{\zeta_p - 1}$ and its norm is

$$\mathrm{N}_{K/\mathbb{Q}}(f_p'(\zeta_p)) = \frac{p^{p-1}\,\mathrm{N}_{K/\mathbb{Q}}(\zeta_p)^{-1}}{\mathrm{N}_{K/\mathbb{Q}}(\zeta_p - 1)}.$$

Now notice that every embedding $\sigma : K \to \mathbb{C}$ is purely complex so they appear in conjugate pairs. Thus for any $\alpha \in K^\times$, $\mathrm{N}_{K/\mathbb{Q}}(\alpha) \in \mathbb{Q}$ is positive. We already know

$$|\,\mathrm{N}_{K/\mathbb{Q}}(\zeta_p - 1)| = \mathrm{N}(1 - \zeta_p) = p$$
$$|\,\mathrm{N}_{K/\mathbb{Q}}(\zeta_p)| = \mathrm{N}(\zeta_p) = 1$$

so putting everything together,

$$\mathrm{disc}(1, \zeta_p, \dots, \zeta_p^{p-2}) = (-1)^{\frac{p-1}{2}} \frac{p^{p-1}}{p} = (-1)^{\frac{p-1}{2}} p^{p-2}.$$

□

Proposition 8.3.

$$\mathcal{O}_K = \mathbb{Z}[\zeta_p].$$

Proof. We already know $[\mathcal{O}_K : \mathbb{Z}[\zeta_p]] < \infty$ and

$$\mathrm{disc}(\mathcal{O}_K)[\mathcal{O}_K : \mathbb{Z}[\zeta_p]]^2 = \mathrm{disc}(1, \zeta_p, \dots, \zeta_p^{p-2}) = \pm p^{p-2}.$$

Hence $\mathbb{Z}[\zeta_p] \subseteq \mathcal{O}_K$ is of p-power index, which we are going to prove to be 1. Look at the quotient ring $\mathcal{O}_K/(1-\zeta_p)$, which has order $\mathrm{N}(1-\zeta_p) = p$ so is just the finite field of p elements. Thus the characteristic homomorphism $\mathbb{Z} \to \mathcal{O}_K/(1-\zeta_p)$ is surjective. Hence for any $z_0 \in \mathcal{O}_K$, there exists $a_0 \in \mathbb{Z}, z_1 \in \mathcal{O}_K$ such that

$$z_0 = a_0 + (1-\zeta_p)z_1.$$

Repeat for z_1, there exists $a_1 \in \mathbb{Z}, z_2 \in \mathcal{O}_K$ such that

$$\begin{aligned} z_0 &= a_0 + (1-\zeta_p)(a_1 + (1-\zeta_p)z_2) \\ &= a_0 + (1-\zeta_p)a_1 + (1-\zeta_p)^2 z_2 \end{aligned}$$

By induction, we see we can write

$$z_0 = \underbrace{a_0 + (1-\zeta_p)a_1 + \dots + (1-\zeta_p)^{n-1}a_{n-1}}_{\in \mathbb{Z}[1-\zeta_p]} + (1-\zeta_p)^n z_n$$

where $a_1, \dots, a_{n-1} \in \mathbb{Z}, z_n \in \mathcal{O}_K$ for any $n \geq 1$, i.e.

$$\mathcal{O}_K = \mathbb{Z}[1-\zeta_p] + (1-\zeta_p)^n \mathcal{O}_K$$

for any $n \geq 1$.

Observe that $\mathbb{Z}[1-\zeta_p] = \mathbb{Z}[\zeta_p]$ and $(1-\zeta_p)^{(p-1)N}\mathcal{O}_K = p^N \mathcal{O}_K$ for any $N \geq 1$. Thus

$$\mathcal{O}_K = \mathbb{Z}[\zeta_p] + p^N \mathcal{O}_K.$$

We know $\mathbb{Z}[\zeta_p] \subseteq \mathcal{O}_K$ has p-power index, so by Lagrange there exists $N \geq 1$ such that $p^N \mathcal{O}_K \subseteq \mathbb{Z}[\zeta_p]$. Hence

$$\mathcal{O}_K = \mathbb{Z}[\zeta_p] + p^N \mathcal{O}_K = \mathbb{Z}[\zeta_p].$$

□

What are the roots of unity in this ring? The ζ_p^i's certainly are. Stare at it a bit longer and you will find their negatives are as well. We use the following lemma to show that's all of them.

Lemma 8.4. *If ℓ is a prime number, then ℓ ramifies in K if and only if $\ell = p$.*

Proof. Recall that by definition, ℓ ramifies in K if and only if there exists $\mathfrak{p} \subseteq \mathcal{O}_K$ prime such that $\mathfrak{p}^2 \mid \ell\mathcal{O}_K$.

We've seen that $(1-\zeta_p)^{p-1} = p\mathcal{O}_K$, so p is ramified in K. Let $\ell \neq p$ be a prime. Since $\mathbb{Z}[\zeta_p] = \mathcal{O}_K$, Dedekind's criterion tells us that ℓ is ramified in K if and only if $f_p(x) \pmod{\ell}$ has a repeated root. We know $\mathrm{disc}\, f_p = \pm p^{p-2}$, hence $\mathrm{disc}(f_p \pmod{\ell}) \neq 0$ so $f_p(x) \pmod{\ell}$ does not have any repeated roots. □

Proposition 8.5. *Let $\mu_K \subseteq \mathcal{O}_K^\times$ be the group of roots of unity in K. Then*

$$\mu_K = \{\pm\zeta_p^i : 0 \le i < p\}.$$

Proof. $\{\pm\zeta_p^i : 0 \le i < p\} \subseteq \mu_K$ is a subgroup of order $2p$ so it suffices to show $|\mu_K| = 2p$. If $\ell \ne p$ is an odd prime and $\ell \mid |\mu_K|$ then since μ_K is cyclic, $\zeta_\ell \in K$ so $\mathbb{Q}(\zeta_\ell) \subseteq K$. As $(1-\zeta_\ell)^{\ell-1}\mathcal{O}_{\mathbb{Q}(\zeta_\ell)} = \ell\mathcal{O}_{\mathbb{Q}(\zeta_\ell)}$, we get $(1-\zeta_\ell)^{\ell-1}\mathcal{O}_K = \ell\mathcal{O}_K$, contradicting the fact that ℓ is unramified in K.

Similarly if $4 \mid |\mu_K|$ then $i \in K$ and hence $(1+i)^2\mathcal{O}_K = 2\mathcal{O}_K$, contradicting the fact that 2 is unramified in K.

If $p^2 \mid |\mu_K|$, then $\omega = e^{2\pi i/p^2} \in K$. Let $f(x) = \frac{x^{p^2}-1}{x^p-1} \in \mathbb{Z}[x]$, then

$$f(x) = \prod_{\substack{1 \le a \le p^2 \\ p \nmid a}} (x - \omega^a).$$

Then

$$f(1) = p = \prod_{\substack{1 \le a \le p^2 \\ p \nmid a}} (1 - \omega^a).$$

By the same argument as for ζ_p, $(1-\omega^a)\mathcal{O}_K = (1-\omega)\mathcal{O}_K$ if $(a,p) = 1$. Hence

$$p\mathcal{O}_K = (1-\omega)^{\phi(p^2)} = (1-\omega)^{p(p-1)}.$$

Taking norm, get $p^{p-1} = \mathrm{N}(1-\omega)^{p(p-1)}$, absurd. Thus $|\mu_K| = 2p$. □

Lemma 8.6 (Kummer)**.** *If $u \in \mathcal{O}_K^\times$, there exists $g \in \mathbb{Z}$ such that*

$$\zeta_p^g u \in K \cap \mathbb{R}.$$

For those familiar with Galois theory, we have the tower of fields

$$\begin{array}{c} K = \mathbb{Q}(\zeta_p) \\ \Big| \, 2 \\ K \cap \mathbb{R} = \mathbb{Q}(\zeta_p + \zeta_p^{-1}) \\ \Big| \, \frac{p-1}{2} \\ \mathbb{Q} \end{array}$$

Proof. Claim that if $\sigma : K \to \mathbb{C}$ is a complex embedding, then for all $\alpha \in K$, $\sigma(\overline{\alpha}) = \overline{\sigma(\alpha)}$: suffices to check this for $\alpha = \zeta_p$. If $\sigma(\zeta_p) = \zeta_p^a$ then

$$\sigma(\overline{\zeta_p}) = \sigma(\zeta_p^{-1}) = \zeta_p^{-a} = \overline{\zeta_p^a} = \overline{\sigma(\zeta_p)}.$$

If $u \in \mathcal{O}_K^\times$, then for any embedding $\sigma : K \to \mathbb{C}$,

$$|\sigma(u/\overline{u})| = |\sigma(u)\overline{\sigma(u)}^{-1}| = 1.$$

Hence $u/\overline{u} \in \mu_K$, so we can write $u/\overline{u} = (-1)^b\zeta_p^k$ for some $b \in \{0,1\}, k \in \mathbb{Z}$. After replacing k by $k+p$, wlog $k = 2g$. Then $u = \overline{u}(-1)^b\zeta_p^{2g}$.

Now look at the residue ring $\mathcal{O}_K/(1-\zeta_p) \cong \mathbb{Z}/p\mathbb{Z}$. The ideal $(1-\zeta_p)\mathcal{O}_K$ is stable under complex conjugation, so complex conjugation induces an automorphism of $\mathcal{O}_K/(1-\zeta_p)$. As $\mathbb{Z} \to \mathcal{O}_K/(1-\zeta_p)$ is surjective, this automorphism is trivial, so for all $\alpha \in \mathcal{O}_K$, $\alpha = \overline{\alpha} \mod (1-\zeta_p)\mathcal{O}_K$. Hence for all $u \in \mathcal{O}_K^\times$,

$$\begin{aligned} u &= \overline{u} \mod (1-\zeta_p) \\ &= \overline{u}(-1)^b\zeta_p^{2g} = \overline{u}(-1)^b \mod (1-\zeta_p) \end{aligned}$$

Since $u \in \mathcal{O}_K^\times$, $u \neq 0 \pmod{(1-\zeta_p)}$ so must have $b = 0$. Hence $u = \overline{u}\zeta_p^{2g}$, so $\zeta_p^{-g}u = \overline{\zeta_p^{-g}u} \in K \cap \mathbb{R}$. □

Lemma 8.7. *If $\alpha \in \mathcal{O}_K$, then there exists $a \in \mathbb{Z}$ such that*

$$\alpha^p = a \mod p\mathcal{O}_K.$$

Proof. For all $\alpha \in \mathcal{O}_K$, there exists $b \in \mathbb{Z}$ such that $\alpha = b \mod (1-\zeta_p)$. Note the identity

$$\alpha^p - b^p = \prod_{i=0}^{p-1}(\alpha - \zeta_p^i b).$$

For any $i \geq 0$,

$$\alpha - \zeta_p^i b = \alpha - b = 0 \mod (1-\zeta_p).$$

Hence

$$\alpha^p - b^p \in (1-\zeta_p)^p \subseteq (1-\zeta_p)^{p-1} = p\mathcal{O}_K.$$

□

We now discuss Fermat's Last Theorem:

Theorem 8.8 (Wiles, 1994)**.** *Let $n \geq 3$ be an integer, and let $x, y, z \in \mathbb{Z}$ be such that*

$$x^n + y^n = z^n$$

then $xyz = 0$.

A little history: in early 19th century, there are many false proofs of this theorem relying on the false assumption that $\mathbb{Z}[e^{2\pi i/n}]$ is a UFD. In 1840s, Kummer invented the theory of ideal factorisation in number fields in order to try to give a correct proof, which worked for a large class of primes. The complete proof was announced by Wiles in 1993 in a room less than 100 yards from where we are now. Despite the geographical proximity, we are NOT going to prove it in this course!

Definition (Regular prime)**.** An prime p is *regular* if

$$p \nmid |\operatorname{Cl}(\mathbb{Z}[\zeta_p])|.$$

Theorem 8.9 (Kummer)**.** *Let p be an regular prime, then Fermat's Last Theorem holds in exponent $n = p$.*

Again we will not prove this. Instead we will prove

Theorem 8.10. *Let p be an odd regular prime. Let $x, y, z \in \mathbb{Z}$ be such that $p \nmid xyz$. Then*
$$x^p + y^p \neq z^p.$$

Kummer called this the "first case" of Fermat's Last Theorem. He dealt with the "second case" (where $p \mid xyz$) using similar techniques.

Proof. Let $x, y, z \in \mathbb{Z}$ such that $x^p + y^p + z^p = 0$, $p \nmid xyz$. wlog $\gcd(x, y, z) = 1$. Then we factor
$$x^p + y^p = \prod_{i=0}^{p-1}(x + \zeta_p^i y) = -z^p \in \mathbb{Z}[\zeta_p].$$
Claim that the ideals $(x + \zeta_p^i y)$ are pairwise coprime:

Proof. Suppose $\mathfrak{q} \subseteq \mathcal{O}_K$ is a prime ideal dividing $(x + \zeta_p^i y)$ and $(x + \zeta_p^j y)$ where $0 \leq i < j < p$. Then
$$\mathfrak{q} \mid ((\zeta_p^i - \zeta_p^j)y) = (1 - \zeta_p)(y).$$
If $\mathfrak{q} \mid (y)$ then $\mathfrak{q} \mid (z)$ so $\mathfrak{q} \mid (x)$. Taking norm, we get $\ell \mid \gcd(x, y, z)$ where $\mathrm{N}(\mathfrak{q}) = \ell^f$, absurd.

If $\mathfrak{q} \mid (1 - \zeta_p)$ then $\mathfrak{q} = (1 - \zeta_p)$ and $(1 - \zeta_p) \mid (z)$ so $p \mid z$, absurd. □

Thus by
$$\prod_{i=0}^{p-1}(x + \zeta_p^i y) = (z)^p$$
there exists an ideal $I \subseteq \mathcal{O}_K$ such that $(x + \zeta_p y) = I^p$. Since p is regular, i.e. $p \nmid |\mathrm{Cl}(\mathcal{O}_K)|$, this implies that I is principal. If $I = (\delta)$ then $(x + \zeta_p y) = (\delta^p)$, so there exists $u \in \mathcal{O}_K^\times$ such that $x + \zeta_p y = u\delta^p$. By previous lemmas, there exists $v \in \mathcal{O}_K^\times \cap \mathbb{R}, g \in \mathbb{Z}, a \in \mathbb{Z}$ such that
$$u\delta^p = \zeta_p^g va \mod p\mathcal{O}_K.$$
Hence $\zeta_p^{-g}(x + \zeta_p y) = va \mod p\mathcal{O}_K$.

Observe that $va \in \mathcal{O}_K \cap \mathbb{R}$ so is invariant under complex conugation. Hence
$$\zeta_p^{-g}(x + \zeta_p y) = \zeta_p^g(x + \zeta_p^{-1} y) \mod p\mathcal{O}_K$$
so
$$\zeta_p^{-g}x + \zeta_p^{1-g}y - \zeta_p^g x - \zeta_p^{g-1}y = 0 \mod p\mathcal{O}_K.$$
What is g? First note that $g \neq 0, 1 \pmod p$: if $g = 0 \pmod p$, then
$$x + \zeta_p y - x - \zeta_p^{-1}y = \zeta_p(1 - \zeta_p^{-2})y = 0 \mod (1 - \zeta_p)^{p-1}$$
and hence $y \in (1 - \zeta_p)^{p-2}$, so $p \mid y$. Similar if $g = 1 \pmod p$.

Now observe that two of $-g, 1 - g, g, g - 1$ must be congruent module p, as otherwise the identity
$$\zeta_p^{-g}x + \zeta_p^{1-g}y - \zeta_p^g x - \zeta_p^{g-1}y = 0 \mod p\mathcal{O}_K$$
together with the fact that $\{\zeta_p^i\}_{i=1}^{p-1}$ is an integral basis of $\mathbb{Z}[\zeta_p]$, forces $p \mid x, p \mid y$, absurd.

Since $g \neq 0, 1 \pmod p$, the only possibility is $-g = g - 1 \pmod p$, i.e. $2g = 1 \pmod p$. Hence

$$\begin{aligned} &\zeta_p^{-g}(x + \zeta_p y - \zeta_p^{2g} x - \zeta_p^{2g-1} y) \\ &= \zeta_p^{-g}(x + \zeta_p y - \zeta_p x - y) \\ &= \zeta_p^{-g}(x - y)(1 - \zeta_p) \\ &= 0 \mod (1 - \zeta_p)^{p-1} \end{aligned}$$

Thus $x - y = 0 \mod (1 - \zeta_p)^{p-2}$, hence $x = y \pmod p$. Recall the equation

$$x^p + y^p + z^p = 0$$

is symmetric in x, y, z so the same argument also gives $y = z \pmod p$, hence

$$3x^p = 0 \pmod p.$$

If $p \neq 3$ then $p \mid x$, absurd. If $p = 3$ then reducing modulo 9 shows there are no solutions, which is left as an exercise. □

The rest of the course is non-examinable.

The question now is how to decide if p is regular. Unfortunately Minkowski's bound is not very effective. To give an idea let $h_p = |\operatorname{Cl}(\mathbb{Z}[\zeta_p])|$ be the class number. The table of class number of cyclotomic fields begins with

p	h_p	p	h_p
3	1	37	37
5	1	41	121
7	1	43	211
11	1	47	695
13	1	53	4889
17	1	59	41241
19	1	61	76301
23	3	67	853513
29	8	71	3882809
31	9	73	11957417

We observe that h_p seems to grow quickly with p. Also most primes seem to regular: of those in the table, all but $p = 37, 59, 67$ are regular.

Kummer gave a criterion to decide whether or not p is regular in terms of the Bernoulli numbers B_n.

Definition (Bernoulli number)**.** For $n \geq 0$, the nth *Bernoulli number* is defined by the formula

$$\frac{t}{1 - e^{-t}} = \sum_{n \geq 0} B_n \frac{t^n}{n!}.$$

Note that $B_n \in \mathbb{Q}$. The first few Bernoulli numbers are

n	B_n	n	B_n
0	1	6	$\frac{1}{42}$
1	$\frac{1}{2}$	7	0
2	$\frac{1}{6}$	8	$-\frac{1}{30}$
3	0	9	0
4	$-\frac{1}{30}$	10	$\frac{5}{66}$
5	0	11	0
		12	$-\frac{691}{2730}$

Theorem 8.11 (Kummer's criterion)**.** *If p is an odd prime, then p is regular if and only if p does not divide the numerator of B_n for any $n = 2, 4, \dots, p-3$.*

Example. $p = 691$ is prime. 691 divides the numerator of B_{12} so by Kummer's criterion $691 \mid h_{691}$.

Alternatively we may define B_n as

$$B_n = -n\zeta(1-n)$$

where $\zeta(s)$ is the Riemann zeta function. This is not coincidental. In fact the Riemann zeta function and its generalisation are closely related to the arithmetics of number fields.

Definition (Dedekind zeta function)**.** Let L be a number field. Its *Dedekind zeta function* is

$$\zeta_L(s) = \sum_{I \subseteq \mathcal{O}_L} \mathrm{N}(I)^{-s}$$

where the sum is over all non-zero ideals.

Note.

1. One can show that the sum is absolutely convergent in the region $\operatorname{Re} s > 1$ and it defines a holomorphic function there. This boils down to bound the number of ideals with certain norm.

2. If $L = \mathbb{Q}$ then

$$\zeta_L(s) = \zeta(s) = \sum_{n \geq 1} n^{-s}$$

is the usual Riemann zeta function. Other properties such as Euler product generalises as well.

In general, unique factorisation of ideals gives an identity

$$\zeta_L(s) = \prod_{\mathfrak{p} \subseteq \mathcal{O}_K} (1 - \mathrm{N}(\mathfrak{p})^{-s})^{-1}$$

where the sum is over all non-zero prime ideals.

Definition (Regulator)**.** Let L be a number field. The *regulator* of L, R_L, is defined as follow: let $v_1, \dots, v_{r_1+r_2-1} \in \mathcal{O}_L^\times$ generate the free abelian group $\mathcal{O}_L^\times / \mu_L$. Let A be the matrix with columns $\ell(v_1), \dots, \ell(v_{r_1+r_2-1})$ where $\ell : \mathcal{O}_L^\times \to \mathbb{R}^{r_1+r_2}$ is the logarithmic map from the proof of Dedekind's unit theorem. Then R_L is the absolute value of any $(r_1 + r_2 - 1) \times (r_1 + r_2 - 1)$

minor of the matrix A.

The idea is that $\ell(\mathcal{O}_L^\times)$ is a lattice in the hyperplane H and R_L is essentially the covolume of the lattice. This is a generalisation of fundamental unit.

Theorem 8.12.

1. *$\zeta_L(s)$ has a meromorphic continuation to all $s \in \mathbb{C}$ with no poles except a simple pole at 1. It satisfies the functional equation*

$$\Lambda_L(s) = \Lambda_L(1-s)$$

where by definition

$$\Lambda_L(s) = \zeta_L(s)|D_L|^{s/2}(\pi^{-s/2}\Gamma(s/2))^{r_1} \cdot (2(2\pi)^{-s}\Gamma(s))^{r_1}$$

where r_1 is the number of real embeddings of L and r_2 is the number of pairs of complex embeddings.

2. *Analytic class number formula: the residue of $\zeta_L(s)$ at $s = 1$ is*

$$\frac{2^{r_1}(2\pi)^{r_2}h_L R_L}{w_L\sqrt{|D_L|}}$$

where by definition $h_L = |\operatorname{Cl}(\mathcal{O}_L)|$ and $w_L = |\mu_L|$.

What does this have to do with cyclotomic fields? Let p be an odd prime and $K = \mathbb{Q}(\zeta_p)$ and $E = K \cap \mathbb{R} = \mathbb{Q}(\zeta_p + \zeta_p^{-1})$.

There exists factorisation

$$\zeta_K(s) = \prod_{\chi:(\mathbb{Z}/p\mathbb{Z})^\times \to \mathbb{C}^\times} L(\chi, s)$$

where χ is a group character and $L(\chi, s)$ is the Dirichlet L-function

$$L(\chi, s) = \prod_{\ell \neq p}(1 - \chi(\ell \mod p)\ell^{-s})^{-1}.$$

There is a similar factorisation

$$\zeta_E(s) = \prod_{\chi:\chi(-1)=1} L(\chi, s).$$

In fact, if $M/\mathbb{Q}$ is any finite Galois extension, then there is a factorisation

$$\zeta_M(s) = \prod_\rho L(\rho, s)^{\dim \rho}$$

induced by irreducible representations ρ of $\operatorname{Gal}(M/\mathbb{Q})$.

Taking the quotient of these two factorisations gives

$$\frac{\zeta_K(s)}{\zeta_E(s)} = \prod_{\chi:\chi(-1)=-1} L(\chi, s).$$

Note that both sides are holomorphic at $s = 1$. Kummer's criterion for the regularity of the prime p is proved by evaluating either side at $s = 1$.

On LHS, we can apply the analytic class number formula for K and E together to get

$$\frac{h_K R_K}{h_E R_E} \cdot \text{explicit factor.}$$

Note that $r_1 + r_2$ is the same for K and E:

	r_1	r_2
K	0	$\frac{p-1}{2}$
E	$\frac{p-1}{2}$	0
$\mathbb{Q}$	1	0

$\mathcal{O}_L^\times$ is a subgroup of $\mathcal{O}_K^\times$ of finite index and thus R_k/R_E is an integer which can be explicitly evaluated. Thus LHS is $h_K/h_E \cdot$ explicit factor.

On RHS, each $L(\chi, s)$ is holomorphic at $s = 1$ and $L(\chi, 1)$ can be evaluated explicitly in terms of (generalised) Bernoulli numbers using purely analytic techniques.

With more work, this leads to Kummer's criterion for the p-divisibility of h_K.

Index

algebraic integer, 3

Bernoulli number, 49

complex embedding, 7
covolume, 28
cyclotomic field, 43

Dedekind zeta function, 50
Dedekind's criterion, 24
Dirichlet's unit theorem, 37
discriminant, 11, 14, 16

fundamental unit, 37

ideal class group, 21
ideal divisibility, 21
inert, 24
integral basis, 12, 15
irreducible ideal, 18

Kummer's criterion, 50
Kummer's lemma, 46

lattice, 28

minimal polynomial, 3
Minkowski constant, 33
Minkowski's theorem, 31

norm, 8, 16
number field, 3

prime ideal, 19

quadratic field, 8

ramification, 24
regular prime, 47
regulator, 50
ring of integers, 6

split completely, 24

trace, 8