Algebraic Number Theory

Alex Rutar*
University of Waterloo

Winter 2019†

*arutar@uwaterloo.ca
†Last updated: June 2, 2019

Contents

Preface

These are course notes for PMath 441/641 (algebraic number theory) at the University of Waterloo. The notes are not original work: they are derived from lectures taught by Matthew Satriano during the Winter 2019 semester. If you find mistakes in the document, you can send me an email at *arutar@uwaterloo.ca* or raise an issue in the repository at *https://github.com/alexrutar/UWaterloo-Course-Notes/*.

I. Field Theory in $\mathbb{C}$

1 Fields over $\mathbb{Q}$

Algebraic Numbers

Definition. An **algebraic integer** is a root of a monic polynomial in $\mathbb{Z}[x]$. An **algebraic number** is the root of any non-zero polynomial in $\mathbb{Z}[x]$. A **number field** is a finite extension of $\mathbb{Q}$. If K, L are fields and $K \subseteq L$, we say that L is an **extension field** of K and K is a **subfield** of L. We write $[L : K] = \dim_K L$, the dimension of L over K.

Example. Equivalently, algebraic numbers are the roots of polynomials in $\mathbb{Q}[x]$. $\sqrt{-5}$ is a root of $x^2 + 5 \in \mathbb{Z}[x]$ is an algebraic integer, and $[\mathbb{Q}(\sqrt{-5}) : \mathbb{Q}] = 2$. A basis for $\mathbb{Q}(\sqrt{-5})$ over $\mathbb{Q}$ is given by $\{1, \sqrt{-5}\}$.

Definition. If K is a field, then $f \in K[x]$ is **irreducible** if whenever $f = gh$, $g, h \in K[x]$, then g or h is constant.

> ***1.1 Proposition.*** *Let $K \subseteq \mathbb{C}$ is a subfield and suppose $f \in K[x]$ is irreducible. Then, f has distinct roots in $\mathbb{C}$.*

Proof Suppose not and write $f(x) = a_n(x-\alpha)^2 g(x)$ in $\mathbb{C}[x]$. Then $f'(x) = 2a_n(x-\alpha)g(x) + a_n(x-\alpha)^2 g(x)$, and $f'(\alpha) = 0$. Let p be the minimal polynomial of α. Then $p|f$ so $p = f$ up to a constant. As well, $f = p|f'$, a contradiction. ∎

Field Extensions

Definition. If $K \subseteq L$ are fields, then we write L/K and say that L is a **extension** of K. If $K \subseteq \mathbb{C}$ is a field $\theta \in \mathbb{C}$, then the field K **adjoin** θ, denoted $K(\theta)$, is defined to be the smallest subfield of $\mathbb{C}$ containing K and θ.

Example. Set $L := \{a + b\sqrt{-5} : a, b \in \mathbb{Q}\}$; why is it that $\mathbb{Q}(\sqrt{-5}) = L$? Certainly L is a field: the inverse of $a + b\sqrt{-5}$ is given by $\frac{a-b\sqrt{-5}}{a^2+5b^2}$, which always since $a^2 + 5b^2$ is not zero whenever $\alpha \neq 0$. To see equality, let M be any field containing $\mathbb{Q}$ and $\sqrt{-5}$. Then if a, b are both rational, then $a \in M$ and $b\sqrt{-5} \in M$ so $a + b\sqrt{-5} \in M$. Thus L is the smallest field containing $\mathbb{Q}$ and $\sqrt{-5}$.

Example. Consider $\zeta = e^{2\pi i/3}$. Then one can verify that $\mathbb{Q}(\zeta) = \{a + b\zeta + c\zeta^2 : a, b, c \in \mathbb{Q}\}$.

Definition. Let $K \subseteq \mathbb{C}$ be a subfield. Then we say $\theta \in \mathbb{C}$ is **algebraic over K** if there exists a polynomial $f \in K[x]$ such that $f(\theta) = 0$. We say $p \in K[x]$ is the **minimal polynomial** of θ if it is monic, has θ as a root, and if it has minimal degree. The **degree of θ over K** is $\deg p(x)$.

Example. $\sqrt{-5}$ has minimal polynomial $x^2 + 5$, and ζ has minimal polynomial $x^2 + x + 1$.

> ***1.2 Proposition. (Properties of the Minimal Polynomial)*** *Let $K \subseteq \mathbb{C}$ be a subfield, $\theta \in \mathbb{C}$ algebraic over K. Then there exists a unique minimal polynomial $p(x)$ of θ over K. In particular, the following hold:*

1. *If $f(\theta) = 0$, $p|f$.*
2. *p is irreducible in $K[x]$*

Proof If $p, q \in K[x]$ are both minimal polynomials, then $r = p - q$ has lower degree and $r(\theta) = 0$. If r is non-zero, let it have leading coefficient c so that $r(x)/c$ is monic. But then $\deg(r/c) < \deg p$ and $r(\theta)/c = 0$, contradicting minimality of p.

1. By the division algorithm, write $f = pq + r$. If $r \neq 0$, then $\deg r < \deg p$ and $r(\theta) = 0$, a contradiction by the same reasoning above.
2. If p is reducible, write $p = fg$ where f, g are not constant. Since $F[x]$ is a UFD, $0 = p(\theta) = f(\theta)g(\theta)$ so θ is a root of f or g, contradicting minimality.

Thus the result holds. ∎

Remark. Since p is irreducible, p has $n = \deg p$ distinct roots in $\mathbb{C}$.

Definition. Suppose θ has minimal polynomial $p(x)$. The roots $\theta_1, \ldots, \theta_n \in \mathbb{C}$ of p are called the **conjugates** of θ.

1.3 Proposition. *Let $K \subseteq \mathbb{C}$, $\theta \in \mathbb{C}$ algebraic over K, and let $n = \deg p$ be the degree of the minimal polynomial. Then every element $\alpha \in K(\theta)$ has a unique representation in the form*

$$\alpha = a_0 + a_1\theta + \cdots + a_{n-1}\theta^{n-1}$$

where $a_i \in K$.

Proof First note that

$$K(\theta) = \left\{ \frac{f(\theta)}{g(\theta)} : f, g \in K[x], g(\theta) \neq 0 \right\}$$

Set $\alpha = f(\theta)/g(\theta) \in K(\theta)$. Let's first see that p and g are coprime. Suppose not; then there exists non-constant $h \in K[x]$ such that $h|p$ and $h|g$. Since p is irreducible, $h = cp$ for some $c \in K^\times$. Then since $h|g$, $p|g$ as well and $g(\theta) = 0$, a contradiction.

Since $K[x]$ is a PID and p, g are coprime, there exist polynomials $s, t \in K[x]$ so that $sp + tg = 1$. Evaluating at g, we must have $t(\theta)g(\theta) = 1$ and

$$\alpha = \frac{f(\theta)}{g(\theta)} = f(\theta)t(\theta)$$

so α is a polynomial in θ. By the division algorithm, $ft = pq + r$ where $\deg r \leq n - 1$ and $\alpha = r(\theta)$ is a polynomial expression in θ with degree $n - 1$.

It remains to see uniqueness. Suppose $\alpha = r_1(\theta) = r_2(\theta)$ where $r_1, r_2 \in K[x]$ and $\deg r_i < n$. If $r_1(x) - r_2(x) \neq 0$, then $\deg(r_1 - r_2) < n$. But then $r_1 - r_2$ has θ as a root and $\deg(r_1 - r_2) < n$, contradicting minimality of p. ∎

Remark. This says that $\{1, \theta, \ldots, \theta^{n-1}\}$ is a basis for $K(\theta)$ over K. In general, when θ is algebraic over K, $K(\theta) = K[\theta]$.

1.4 Corollary. *Suppose $M/L/K$. Then $[M : K] = [M : L][L : K]$.*

Proof Exercise. ∎

2 Finite Extensions and Embeddings

Definition. An injective ring homomorphism $\phi : R \to S$ is called an **embedding**. We write $R \hookrightarrow S$ is the inclusion map.

t:embed

2.1 Theorem. *Let $K \subseteq \mathbb{C}$ is a subfield, L/K is a finite extension field. If $\sigma : K \hookrightarrow \mathbb{C}$ is an embedding, then σ extends to an embedding $L \hookrightarrow \mathbb{C}$ in exactly $[L : K]$ ways.*

Proof First, let's prove the theorem for extensions of the form $K(\alpha)/K$. Let $p(x) = a_0 + \cdots + a_m x^m \in K[x]$ be the minimal polynomial of α over K. Since σ is injective, $K \cong \sigma(K) \subseteq \mathbb{C}$. Let $g(x) = \sigma(a_0) + \cdots + \sigma(a_{m-1})x^{m-1} + x^m$, which is irreducible over $\sigma(K)$. To see this, if $(c_0 + c_1 x + \cdots + c_u x^u)(d_0 + d_1 x + \cdots + d_v x^v)$ is any factorization (with $c_i, d_i \in \sigma(K)$), then $(\sigma^{-1}(c_0) + \sigma^{-1}(c_1)x + \cdots + x^u)(\sigma^{-1}(d_0) + \sigma^{-1}(d_1)x + \cdots + x^v)$ is a factorization of $p(x)$, so it must be trivial. Now, let $\beta_1, \ldots, \beta_m \in \mathbb{C}$ be the distinct roots of $g(x)$, and let $\beta := \beta_i$ be arbitrary. Given an element $\gamma = b_0 + b_1\alpha + b_2\alpha^2 + \cdots + b_{m-1}\alpha^{m-1}$ in $K(\alpha)$, let

$$\lambda_\beta(\gamma) = \sigma(b_0) + \sigma(b_1)\beta + \cdots + \sigma(b_{m-1})\beta^{m-1}$$

One can verify that this is a ring homomorphism which respects σ. Furthermore, there no other embeddings λ since $0 = \lambda(0) = \lambda(p(\alpha)) = g(\lambda(\alpha))$. Thus, $\lambda(\alpha)$ is a root of g, so $\lambda(\alpha) = \beta_i$ for some i. Since λ is a homomorphism, if $\lambda_1(\alpha) = \lambda_2(\alpha)$, then $\lambda_1 = \lambda_2$, so there are at most $[K(\alpha) : K]$ embeddings.

Now, the proof follows by induction. If $[L : K] = 1$, we are done; if $[L : K] > 1$, get $\alpha \in L \setminus K$. From above, σ extends to $[K(\alpha) : K]$ embeddings $\lambda : K(\alpha) \to \mathbb{C}$, and by induction, any such embedding extends to $[L : K(\alpha)]$ embeddings $\lambda : L \to \mathbb{C}$. Thus there are $[L : K(\alpha)][K(\alpha) : K] = [L : K]$ embeddings extending σ, as desired. ∎

Remark. Our most common use case will be when σ is the identity map on K.

Example. Consider the embedding of $\mathbb{Q} \hookrightarrow \mathbb{C}$. If $\mathbb{Q}(\sqrt{d})/\mathbb{Q}$, then the two embeddings are given by $\sqrt{d} \mapsto \sqrt{d}$ or $\sqrt{d} \mapsto -\sqrt{d}$. Note that $\pm\sqrt{d}$ are conjugates: both are roots of the minimal polynomial $x^2 - d$.

Example. Suppose $K = \mathbb{Q}$, and $L = \mathbb{Q}(\sqrt[3]{2})$. Since $x^3 - 2$ is the minimal polynoial of $\sqrt[3]{2}$, its conjugates are $\sqrt[3]{2}, \sqrt[3]{2}\omega, \sqrt[3]{2}\omega^2$ where $\omega = e^{2\pi i/3}$. All embedings extend $\mathbb{Q} \subseteq \mathbb{C}$ are given by $\sqrt[3]{2} \mapsto \sqrt[3]{2}\omega^k$ for $k = 0, 1, 2$.

2.2 Theorem. (Primitive Element) *Let $K \subseteq L \subseteq \mathbb{C}$ with $[L : K] < \infty$. Then there $\theta \in L$ such that $L = K(\theta)$.*

Proof Since $[L : K] < \infty$, we have $L = K(\alpha_1, \alpha_2, \ldots, \alpha_m)$ for some m. By induction on m, it suffices to handle the case $L = K(\alpha, \beta)$.

Let $\{\alpha_1, \ldots, \alpha_n\}$ be the conjugates of α and $\{\beta_1, \ldots, \beta_m\}$ are conjugates of β (over K). Let $c \in K^\times$ be such that $\alpha + c\beta \neq \alpha_i + c\beta_j$ for any $(i, j) \neq (1, 1)$ (K is an infinite field, so such a c certainly exists), and set $\theta := \alpha + c\beta$. Certainly $K(\theta) \subseteq K(\alpha, \beta)$; for the reverse inclusion, it suffices to show that $\beta \in K(\theta)$. Let $f(x)$ be the minimal polynomial of α over K, and $g(x)$ the minimal polynomial of β over K. Note that β is a root of both $f(\theta - cx)$ and $g(x)$; and by choice of c, there are no others in common.

Let $h(x)$ be the minimal polynomial of β over $K(\theta)$. Since β is a root of both $f(\theta - cx)$ and $g(x) \in K[x] \subseteq K(\theta)[x]$, we must have $h | f(\theta - cx)$ and $h | g(x)$, so $\deg h = 1$ and $\beta \in K(\theta)$.∎

Normal Extensions

Definition. Let $K \subseteq L \subseteq \mathbb{C}$, $[L:K] < \infty$. We say L is a **normal extension** of K if it is closed under taking conjugates over K.

Example. For example, $\mathbb{Q}(\sqrt{d})/\mathbb{Q}$ is a normal extension. If $\alpha \in L$, then $\alpha = a + b\sqrt{d}$. The conjugate of α is $a - b\sqrt{d}$, which is also an element of L. On the other hand, a classic non-example is $L = \mathbb{Q}(\sqrt[3]{2})/\mathbb{Q}$. Then $\sqrt[3]{2} \in L$ but $\omega\sqrt[3]{2} \notin L$, since $\omega\sqrt[3]{2} \notin \mathbb{R}$.

2.3 Proposition. *Let $K \subseteq L \subseteq \mathbb{C}$, $[L:K] < \infty$. Then L/K is normal if and only if for all $\sigma : L \hookrightarrow \mathbb{C}$ such that $\sigma|_K = \mathrm{id}_K$, σ is an automorphism of L.*

Proof Note that σ is an automorphism of L if and only if $\sigma(L) = L$.

If L/K is normal, let $\alpha \in L$ be such that $L = K(\alpha)$. Then $\sigma : L \hookrightarrow \mathbb{C}$ is specified fully by $\sigma(\alpha) = \alpha_i$, where α_i is a conjugate of α. But then $\sigma : K(\alpha) \to K(\alpha_i)$ is an isomorphism, and since L/K is normal, $K(\alpha) = K(\alpha_i)$ and σ is an automorphism of L.

Conversely, let's show that L/K is normal. Let $\alpha \in L$, let α_i be the conjugates of α over K: we need to show that $\alpha_i \in L$. Let $\sigma(\alpha) = \alpha_i$ extend id_K, and by hypothesis, σ is an automorphism so $\alpha_i \in K(\alpha) = L$. ∎

Remark. Recall that there are $[L:K]$ embeddings that fix K; in other words, $\sigma : L \hookrightarrow \mathbb{C}$ such that $\sigma|_K = \mathrm{id}_K$. The corollary says that L/K is normal if and only if all of these embeddings are automorphisms. Thus L/K is normal if and only if exactly $[L:K]$ automorphisms of L fixing K.

2.4 Corollary. *Let $K \subseteq \mathbb{C}$, $\alpha_i \in \mathbb{C}$ algebraic over K. Then $L = K(\alpha_1, \ldots, \alpha_n)$ is normal over K if all the conjugates of α_i are in L.*

Proof Let $\sigma : L \hookrightarrow \mathbb{C}$ be an embedding extending id_K. If $\theta \in L$, then $\theta = f(\alpha_1, \ldots, \alpha_n)$ for $f(x) \in K[x_1, x_2, \ldots, x_n]$. Then $\sigma(\theta) = f(\sigma(\alpha_1), \ldots, \sigma(\alpha_n))$ where $\sigma(\alpha_i)$ is some conjugate of α_i, an element of L by hypothesis. Thus $\sigma(\theta) \in L$ so $\theta \in L$ as well. ∎

2.5 Corollary. *$K \subseteq L \subseteq \mathbb{C}$, $[L:K] < \infty$. Then there exists a finite extension M/L such that M/K is normal.*

Proof Get $\alpha \in L$ so that $L = K(\alpha)$. Let $\alpha_1, \ldots, \alpha_n$ be the conjugates of α over K. Set $M = K(\alpha_1, \ldots, \alpha_n)$, and by the previous corollary, M/K is normal. ∎

Example. Let $L = \mathbb{Q}(\sqrt[3]{2})$, $K = \mathbb{Q}$. L/K is not normal, but $M = \mathbb{Q}(\sqrt[3]{2}, \sqrt[3]{2}\omega, \sqrt[3]{2}\omega^2)/\mathbb{Q}$ is normal.

3 Galois Theory over $\mathbb{Q}$

Definition. Let L/K be any finite extension. The **Galois group** of L/K is defined

$$\mathrm{Gal}(L/K) = \left\{\sigma \in \mathrm{Aut}(L) \;\middle|\; \sigma|_K = \mathrm{id}_K\right\}$$

Now if $H \leq \mathrm{Gal}(L/K)$, $L^H = \{\alpha \in L : \sigma(\alpha) = \alpha \forall \sigma \in H\}$ is called the **fixed field of H**.

Remark. Recall that $|\mathrm{Gal}(L/K)| \leq [L:K]$, with equality if and only if L/K is normal. As well, one can verify that L^H is indeed a field, so L/L^H is an extension. In particular, this extension has certain properties:

t:galfix

3.1 Theorem. *Given $K \subseteq L \subseteq \mathbb{C}$, L/K a finite normal extension. Let $G = \mathrm{Gal}(L/K)$. Then*
- $L^G = K$
- *If $H \leq G$ and $L^H = K$, then $H = G$*

Proof We first see that $K = L^G$. Let $\sigma : L \hookrightarrow \mathbb{C}$ be an embedding fixing K. Since L is normal, $\sigma \in \mathrm{Gal}(L/K)$, so by definition of L^G, σ fixes L^G. But then $[L:L^G][L^G:K] \leq [L:L^G]$, so $[L^G:K] \leq 1$ and $L^G = K$.

Suppose now that $L^H = K$. Set $L = K(\alpha)$ and consider the polynomial

$$f(x) = \prod_{\sigma \in H} (x - \sigma(\alpha)) = x^{|H|} - e_1 x^{|H|-1} + \cdots + e_{|H|}(-1)^{|H|}$$

where the e_i are elementary symmetric functions in the $\sigma(\alpha)$. If $\tau \in H$, then

$$\tau(e_1) = \sum_{\sigma \in H} \tau\sigma(\alpha) = \sum_{\sigma \in H} \sigma(\alpha) = e_1$$

since $\tau \in H$ permutes the $\sigma(\alpha)$ and e_1 is a symmetric polynomial in $\sigma(\alpha)$. The same argument holds for any e_i, so $e_i \in L^H = K$ for all i; thus, $f(x) \in K[x]$. Since $\mathrm{id} \in H$, $f(\alpha) = 0$; and $\deg f = |H|$. Since the minimal polynomial of α over K has degree $\leq |H|$,

$$[L:K] = [K(\alpha):K] \leq |H| \leq |G| = [L:K]$$

so $H = G$. ∎

Remark. Suppose L/K is normal, and $L \supseteq F \supseteq K$ where F is a field. Then L/F is also normal since conjugates of $\alpha \in L$ over F are a subset of conjugates of α over K.

t:ftgt

3.2 Theorem. (Fundamental Theorem of Galois Theory) *Let $K \subseteq L \subseteq \mathbb{C}$, L/K normal, with $L/F/K$.*
- *(i)* $L^{\mathrm{Gal}(L/F)} = F$
- *(ii) If $H \leq G = \mathrm{Gal}(L/K)$, then $\mathrm{Gal}(L/L^H) = H$.*
- *(iii) F/K is normal if and only if $\mathrm{Gal}(L/F) \trianglelefteq \mathrm{Gal}(L/K)$. In this case,*

$$\mathrm{Gal}(F/K) \cong \mathrm{Gal}(L/K) \Big/ \mathrm{Gal}(L/F)$$

Proof (i) Since L/K is normal, L/F is normal and (Thm. 3.1) states that $F = L^{\mathrm{Gal}(L/F)}$.

(ii) Let $H' = \mathrm{Gal}(L/L^H)$. By definition, H fixes L^H, so $H \leq H' = \mathrm{Gal}(L/L^H)$. Since L/L^H is normal and $H \leq \mathrm{Gal}(L/L^H)$ has L^H as its fixed field, by the previous theorem, $H = \mathrm{Gal}(L/L^H) = H'$.

(iii) Let $H = \mathrm{Gal}(L/F)$. If $\sigma \in \mathrm{Gal}(L/K)$, then $\sigma : F \longrightarrow \sigma(F)$ is an isomorphism and $\sigma\,\mathrm{Gal}(L/F)\sigma^{-1} = \mathrm{Gal}(L/\sigma(F))$. Thus,

$$\begin{aligned} \mathrm{Gal}(L/F) \trianglelefteq \mathrm{Gal}(L/K) &\iff \mathrm{Gal}(L/\sigma(F)) = \sigma\,\mathrm{Gal}(L/F)\sigma^{-1} = \mathrm{Gal}(L/F) \\ &\iff \sigma(F) = F \text{ for all } \sigma \\ &\iff F/K \text{ is normal} \end{aligned}$$

since a field is normal if and only if it is fixed by all its automorphisms.

When this holds, we can compute $\mathrm{Gal}(F/K)$. Since $\sigma(F) = F$, we have a well-defined map $\mathrm{Gal}(L/K) \to \mathrm{Gal}(F/K)$ given by $\sigma \mapsto \sigma|_F$. The kernel is $\{\sigma \in \mathrm{Gal}(L/K) : \sigma|_F = \mathrm{id}_F\} = \mathrm{Gal}(L/F)$. Then by first isomorphism theorem,

$$\mathrm{Gal}(F/K) \cong \mathrm{Gal}(L/K) \Big/ \mathrm{Gal}(L/F)$$

as required. ∎

II. The Ring of Algebraic Integers

4 Number Fields

We now focus our attention on extensions, in particular finite extensions, of $\mathbb{Q}$ in $\mathbb{C}$. A major example throughout this section are the cyclotomic extensions of $\mathbb{Q}$; many of the theorems we will prove will provide tools to better understand extensions $\mathbb{Q}(\zeta_n)/\mathbb{Q}$. Recall that α is an algebraic integer if it has a minimal polynomial in $\mathbb{Z}[x]$.

Definition. K is a **number field** if K is a finite extension of $\mathbb{Q}$. The set of **algebraic numbers** over $\mathbb{Q}$ is denoted $\overline{\mathbb{Q}}$. The set of **algebraic integers** is denoted $\mathcal{O}_{\overline{\mathbb{Q}}}$. We write $\mathcal{O}_K \subseteq K = \mathcal{O}_{\overline{\mathbb{Q}}} \cap K$ to denote the subset of **algebraic integers of K**.

We can prove Gauss' Lemma under the observation that $\mathbb{Z}_p[x]$ is an integral domain.

l:gauss

4.1 Lemma. (Gauss) *If $f, g \in \mathbb{Z}[x]$ are primitive (their coefficients have no non-trivial common factor), then fg is also primitive.*

Proof Suppose $f, g \in \mathbb{Z}[x]$ are primitive. If fg is not primitive, then some prime p divides all coefficients of fg. Consider modulo p, so $\overline{f}\,\overline{g} = 0$. Then $\overline{f} = 0$ or $\overline{g} = 0$, so p divides all coefficients of f or g and f, g are not primitive. ∎

p:min-int

4.2 Proposition. *Let α be an algebraic integer. Then the minimal polynomial of α over $\mathbb{Q}$ is in $\mathbb{Z}[x]$.*

Proof Let α be an algebraic integer, so there exists $h \in \mathbb{Z}[x]$ monic such that $h(\alpha) = 0$. Let $f \in \mathbb{Q}[x]$ be the minimal polynomial of α over $\mathbb{Q}$. Then $h = fg$ in $\mathbb{Q}[x]$. Since h, f are monic, g is also monic. Let $a, b \in \mathbb{Z}$ so that $af, bg \in \mathbb{Z}[x]$ and af, bg are primitive polynomials. Then by Gauss's Lemma (Lem. 4.1), $abh = (af)(bg) \in \mathbb{Z}[x]$ is primitive, so $ab = \pm 1$, so $a, b = \pm 1$ and $f, g \in \mathbb{Z}[x]$ to begin with. ∎

A simple observation following from this fact is that $\mathcal{O}_{\mathbb{Q}} = \mathbb{Z}$.

Example. (Quadratic Extensions) Let d be a squarefree integer. Then

$$\mathcal{O}_{\mathbb{Q}(\sqrt{d})} = \begin{cases} \mathbb{Z}[d] & : d \equiv 2, 3 \pmod 4 \\ \left\{ \frac{a+b\sqrt{d}}{2} : a \equiv b \pmod 2 \right\} & : d \equiv 1 \pmod 4 \end{cases}$$

Let $\alpha = r + s\sqrt{d}$, $r, s \in \mathbb{Q}$. If $s = 0$, then $\alpha = r \in \mathbb{Q}$, so $\alpha \in \mathcal{O}_{\mathbb{Q}(\sqrt{d})}$. Now consider $s \neq 0$. The minimal polynomial of α over $\mathbb{Q}$ is

$$\left(x - (r + s\sqrt{d})\right)\left(x - (r - s\sqrt{d})\right) = x^2 - 2rx + (r^2 - ds^2)$$

By $\alpha \in \mathcal{O}_{\mathbb{Q}(\sqrt{d})}$ if and only if $2r \in \mathbb{Z}$, $r^2 - ds^2 \in \mathbb{Z}$.

First, if $r \in \mathbb{Z}$, so $ds^2 \in \mathbb{Z}$ and since d is squarefree, $s \in \mathbb{Z}$.

The other case is $r = \frac{a}{2}$, where a is an odd integer. Then $ds^2 = \text{integer} + a^2/4$, so $s = b/2$ where b is an odd integer. Since $r^2 - ds^2 \in \mathbb{Z}$, we need $4 \mid (a^2 - db^2)$. Modulo 4, $a^2 \equiv db^2$, and since a, b are odd, $a^2 = b^2 \equiv 1 \pmod 4$ and $d \equiv 1 \pmod 4$.

Remark. Notice in the preceding examples, we got

$$\mathcal{O}_{\mathbb{Q}} = \mathbb{Z} \qquad\qquad \mathcal{O}_{\mathbb{Q}(\sqrt{d})} = \begin{cases} \mathbb{Z}[\sqrt{d}] \\ \mathbb{Z}\left[\frac{1+\sqrt{d}}{2}\right] \end{cases}$$

which are all rings!

t:alg-int

4.3 Theorem. *Let $\alpha \in \mathbb{C}$. Then the following are equivalent:*

(i) α is an algebraic integer
(ii) $\mathbb{Z}[\alpha]$ is finitely generated as an additive group
(iii) α is an element of some subring of $\mathbb{C}$ having finitely generated additive group.
(iv) $\alpha A \subseteq A$ for some finitely generated additive subgroup $A \subseteq \mathbb{C}$.

Proof $(i \Rightarrow ii)$ We know $\mathbb{Z}[\alpha] = \{a_0 + a_1\alpha + \cdots + a_{n-1}\alpha^{n-1} : a_i \in \mathbb{Z}\}$ where n is the degree of α over $\mathbb{Q}$. Then it is generated over $\mathbb{Z}$ by $\{1, \alpha, \alpha^2, \ldots, \alpha^{n-1}\}$.

$(ii \Rightarrow iii)$ $\alpha \in \mathbb{Z}[\alpha]$ and $\mathbb{Z}[\alpha]$ is a subring of $\mathbb{C}$ with finitely generated additive group.

$(iii \Rightarrow iv)$ Let $A \subseteq \mathbb{C}$ denote the subring with $\alpha \in A$; then $\alpha A \subseteq A$.

$(iv \Rightarrow i)$ Let $\{a_1, \ldots, a_n\}$ generate A as an additive group with $\alpha A \subseteq A$. In particular, $\alpha a_i \in A$, so there exists $\{m_{ij} : j = 1, \ldots, n\} \subset \mathbb{Z}$ such that $\alpha a_i = \sum_{j=1}^{n} m_{ij} a_j$. Let $M = (m_{ij})$ in $\mathbb{Z}$, so that

$$(\alpha I_n - M)\begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix} = 0$$

Thus, α is a root of $\det(xI_n - M) \in \mathbb{Z}[x]$. ∎

Remark. The proof of $(iv \Rightarrow i)$ gives a general method for computing polynomials which have a specific algebraic integer as a root.

4.4 Corollary. *$\mathcal{O}_{\overline{\mathbb{Q}}}$ is a ring. In particular, $\mathcal{O}_K$ is a ring for any number field K.*

Proof Say α has degree n and β has degree m over $\mathbb{Q}$. Then $\mathbb{Z}[\alpha, \beta] \subseteq \mathbb{C}$ is a subring with a finitely generated additive group because it is generated by $\alpha^i \beta^j$ where $0 \le i < n$, $0 \le j < m$. Since $\alpha\beta, \alpha + \beta \in \mathbb{Z}[\alpha, \beta]$ we are done by (iii) in (Thm. 4.3). Finally, $\mathcal{O}_K = K \cap \mathcal{O}_{\overline{\mathbb{Q}}}$ is an intersection of rings and thus also a ring. ∎

4.5 Proposition. *Let α be an algebraic number. Then there exists $r \in \mathbb{Z}^+$ such that $r\alpha$ is an algebraic integer.*

This essentially says that if $\alpha \in K$ and K is a number field, then there exists $r \in \mathbb{Z}^+$ such that $r\alpha \in \mathcal{O}_K$.

Proof Since α is an algebraic number, α satisfies a polynomial in $\mathbb{Q}[x]$. Clear denominators to get $h \in \mathbb{Z}[x]$ so $h(\alpha) = 0$. Write $h(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_0$. Then

$$\begin{aligned} a_n^{n-1} h(x) &= a_n^n x^n + a_n^{n-1} a_{n-1} x^{n-1} + \cdots + a_n^{n-1} a_0 \\ &= (a_n x)^n + a_{n-1}(a_n x)^{n-1} + \cdots + a_n^{n-1} a_0 \end{aligned}$$

Let $g(x) = x^n + a_{n-1}x^{n-1} + \cdots + a_n^{n-1}a_0$, so $g(a_n\alpha) = 0$ and $a_n\alpha$ is an algebraic integer. If a_n is negative, take $-a_n\alpha$ instead. ∎

Cyclotomic Extensions I: Introduction

Definition. We say ζ_n is a **primitive n^{th} root of unity** if $\zeta_n^n = 1$ and $\zeta_n^k \neq 1$ for any $k < n$. We call the extension $\mathbb{Q}(\zeta_n)$ a **cyclotomic field**.

Example. The 4^{th} roots of unity are $1, i, -1, -i$, so i and $-i$ are the primitive 4^{th} roots of unity.

The cyclotomic fields play a fundamental role in number theory. For example, in class field theory, we have the following theorem:

> ***4.6 Theorem. (Kronecker-Weber)*** *If $K/\mathbb{Q}$ is a finite normal extension and $\mathrm{Gal}(K/\mathbb{Q})$ is abelian, then $K \subseteq \mathbb{Q}(\zeta_n)$ for some n.*

We will not prove this theorem in full generality, but we will see partial results on assignments.

> ***4.7 Theorem.*** *ζ_n is an algebraic integer with minimal polynomial*
>
> $$\Phi_n(x) := \prod_{j \in (\mathbb{Z}_n)^\times} (x - \zeta_n^j)$$

Proof Note that ζ_n is a root of $x^n - 1$, so ζ_n is an algebraic integer. As in (Prop. 4.2), let $f(x) \in \mathbb{Z}[x]$ be the minimal polynomial of ζ_n over $\mathbb{Q}$ so that $f(x) \mid (x^n - 1)$ over $\mathbb{Z}[x]$. Recall that

$$x^n - 1 = \prod_{j \in \mathbb{Z}_n} (x - \zeta_n^j)$$

If $j \notin (\mathbb{Z}_n)^\times$, then ζ_n^j satisfies $x^{\frac{n}{\gcd(n,j)}} - 1$ but ζ_n does not, so ζ and ζ_n^j are not conjugates. Thus the only possible conjugates for ζ_n are the ζ_n^j where $j \in (\mathbb{Z}_n)^\times$; it suffices to show that these are precisely the conjugates. In particular, let's show that if $\theta = \zeta_n^t$ and p is prime with $p \nmid n$, then θ^p is conjugate to θ. With this, the result follows: if j is coprime to n, write $j = p_1^{e_1} \cdots p_m^{e_m}$ with $p_i \nmid n$ and repeatedly apply the above result to ζ_n for each p_i, e_i times.

Thus let's prove the claim. Write $x^n - 1 = f(x)g(x)$ with $f, g \in \mathbb{Z}[x]$; since θ^p is a root of $x^n - 1$, either it is a root of $f(x)$ - in which case we're done - or it is a root of $g(x)$. Suppose $g(\theta^p) = 0$, so θ is a root of $g(x^p) \in \mathbb{Z}[x]$ so $f(x) \mid g(x^p)$ over $\mathbb{Z}[x]$. Modulo p, $\overline{f}(x) \mid \overline{g}(x^p) = \overline{g}(x)^p$ in $\mathbb{Z}_p[x]$. Since $\mathbb{Z}_p[x]$ is a UFD, let $s(x)$ be an irreducible factor of $f(x)$ so that $s|\overline{f}$ and thus $s|\overline{g}$. But then $x^n - \overline{1} = \overline{f}\,\overline{g}$, so $s^2 \mid (x^n - 1)$ and $s \mid \overline{n}x^{n-1}$. Since n is coprime to p, this implies $s = cx$ for some $c \in \mathbb{Z}_p$. But then $cx \mid x^n - \overline{1}$, a contradiction. ∎

Remark.

1. For p prime, we have
$$\Phi_p(x) = \prod_{j=1}^{p-1} (x - \zeta_p^j) = \frac{x^p - 1}{x - 1} = x^{p-1} + x^{p-2} + \cdots + 1$$
2. $\mathbb{Q}(\zeta_n)/\mathbb{Q}$ is a normal extension since conjugates of ζ_n are $\zeta_n^j \in \mathbb{Q}(\zeta_n)$. As well, $[Q(\zeta_n) : \mathbb{Q}] = |(\mathbb{Z}_n)^\times| = \phi(n)$.

4.8 Proposition. $\mathrm{Gal}(\mathbb{Q}(\zeta_n)/\mathbb{Q}) \cong (\mathbb{Z}_n)^\times$.

Proof Set $G = \mathrm{Gal}(\mathbb{Q}(\zeta_n)/\mathbb{Q})$, which consists of automorphisms $\sigma : \mathbb{Q}(\zeta_n) \to \mathbb{Q}(\zeta_n)$ fixing $\mathbb{Q}$. For such a σ, we must have $\sigma(\zeta_n) = \zeta_n^j$ for some $\gcd(j,n) = 1$. Thus to every $\sigma \in G$, we can associate the index $j \in (Z_n)^\times$ so that $\sigma_j(\zeta_n) = \zeta_n^j$. This gives us a map $G \to (\mathbb{Z}_n)^\times$ by $\sigma_j \mapsto j$. This map is a homomorphism:

$$\sigma_k\sigma_j(\zeta_n) = \sigma_k(\zeta_n^j) = \sigma_k(\zeta_n)^j = \zeta_n^{jk} = \sigma_{jk}(\zeta_n)$$

and bijectivity is left as a straightforward exercise. ∎

5 Traces, Norms, and Units

Definition. Suppose K is a number field with $[K : \mathbb{Q}] = n$, and let $\sigma_1, \ldots, \sigma_n : K \hookrightarrow \mathbb{C}$ be the usual embeddings extending $\mathbb{Q} \subseteq \mathbb{C}$. Given $\alpha \in K$, we say its **trace** is

$$\mathrm{Tr}_{\mathbb{Q}}^K = Tr_{\mathbb{Q}}^K(\alpha) = \sum_{i=1}^n \sigma_i(\alpha)$$

and its **norm**

$$N_{\mathbb{Q}}^K(\alpha) = \prod_{i=1}^n \sigma_i(\alpha)$$

5.1 Proposition. *Let $r \in \mathbb{Q}$, $\alpha, \beta \in K$ as above. Then*

(i) $\mathrm{Tr}_{\mathbb{Q}}^K(r\alpha) = r\,\mathrm{Tr}_{\mathbb{Q}}^k(\alpha)$
(ii) $\mathrm{Tr}_{\mathbb{Q}}^K(\alpha+\beta) = \mathrm{Tr}_{\mathbb{Q}}^K(\alpha) + \mathrm{Tr}_{\mathbb{Q}}^K(\beta)$
(iii) $N_{\mathbb{Q}}^K(\alpha\beta) = N_{\mathbb{Q}}^K(\alpha)N_{\mathbb{Q}}^K(\beta)$
(iv) $N_{\mathbb{Q}}^K(r\alpha) = r^n N_{\mathbb{Q}}^K(\alpha)$

Proof Exercise. ∎

Example. Consider $\sqrt{2} \in \mathbb{Q}(\sqrt{2},\sqrt{3}) = K$. The minimal polynomial of $\sqrt{2}$ is $x^2 - 2$. The 4 embeddings $K \hookrightarrow \mathbb{C}$ are given by $\sqrt{2}, \sqrt{3} \mapsto \pm\sqrt{2}, \pm\sqrt{3}$, so $N_{\mathbb{Q}}^K(\sqrt{2}) = \sqrt{2}\sqrt{2}(-\sqrt{2})(-\sqrt{2}) = 4$.

t:relntr

5.2 Theorem. *If $[K : \mathbb{Q}] = n$, $\alpha \in K$, then*

$$\frac{1}{[K:\mathbb{Q}]}\mathrm{Tr}_{\mathbb{Q}}^K(\alpha) = \frac{1}{[\mathbb{Q}(\alpha):\mathbb{Q}]}\mathrm{Tr}_{\mathbb{Q}}^{\mathbb{Q}(\alpha)}(\alpha)$$

and

$$N_{\mathbb{Q}}^K(\alpha)^{\frac{1}{[K:\mathbb{Q}]}} = N_{\mathbb{Q}}^{\mathbb{Q}(\alpha)}(\alpha)^{\frac{1}{[\mathbb{Q}(\alpha):\mathbb{Q}]}}$$

Proof Each of the $[\mathbb{Q}(\alpha) : \mathbb{Q}]$ embeddings $\mathbb{Q} \hookrightarrow \mathbb{Q}(\alpha)$ extend to $[K : \mathbb{Q}(\alpha)]$ embeddings $\mathbb{Q} \hookrightarrow K$. So, letting σ_i be the embeddings $\mathbb{Q}(\alpha) \hookrightarrow \mathbb{C}$, let σ_{ij} be the $[K : \mathbb{Q}(\alpha)]$ extensions. Then

$$\mathrm{Tr}_{\mathbb{Q}}^K(\alpha) = \sum_{j=1}^{[K:\mathbb{Q}(\alpha)]}\sum_{i=1}^n \sigma_{ij}(\alpha) = \sum_{j=1}^{[K:\mathbb{Q}(\alpha)]}\left(\sum_{i=1}^n \sigma_i(\alpha)\right) = [K : \mathbb{Q}(\alpha)]\,\mathrm{Tr}_{\mathbb{Q}}^{\mathbb{Q}(\alpha)}\sigma_i(\alpha)$$

and the proof is identical for $N_{\mathbb{Q}}^K(\alpha)$. ∎

Remark. Given α an algberaic integer, the value $\mathrm{Tr}(\alpha)$ does not necessarily make sense since you need to choose the number field K containing α. However, the previous proposition (Thm. 5.2) says that this distinction is not too important: if we divide by $1/[K:\mathbb{Q}]$, the trace does not depend on K containing α.

5.3 Corollary. *If $\alpha \in K$, K is a number field, then $\mathrm{Tr}_{\mathbb{Q}}^{K}(\alpha)$, $N_{\mathbb{Q}}^{K}(\alpha) \in \mathbb{Q}$. In particular, if $\alpha \in \mathcal{O}_K$, then $\mathrm{Tr}_{\mathbb{Q}}^{K}(\alpha)$, $N_{\mathbb{Q}}^{K}(\alpha) \in \mathbb{Z}$.*

Proof Let α have minimal polynomial $x^n + a_{n-1}x^{n-1} + \cdots + a_0$. Note that $\mathrm{Tr}_{\mathbb{Q}}^{\mathbb{Q}(\alpha)}$ is the $-a_{n-1}$ coefficient and $N_{\mathbb{Q}}^{\mathbb{Q}(\alpha)}$ is the $\pm a_0$ coefficient of the minimal polynomial. These are both rationals, and if α is an algebraic integer, then they are both integers. Then by (Thm. 5.2), $\mathrm{Tr}_{\mathbb{Q}}^{K}$ and $\mathbb{N}_{\mathbb{Q}}^{K}$ are integer multples / powers, and are thus still rational or integer. ■

Since $\mathcal{O}_K$ is a ring, it is natural to ask what the units are.

p:unit

5.4 Proposition. *Let K be a number field and $\alpha \in \mathcal{O}_K$. Then $\alpha \in \mathcal{O}_K^{\times}$ if and only if $N_{\mathbb{Q}}^{K}(\alpha) = \pm 1$.*

Proof If $\alpha \in \mathcal{O}_K^{\times}$, then $\alpha\beta = 1$ for some $\beta \in \mathcal{O}_K$. Then $1 = N_{\mathbb{Q}}^{K}(1) = N_{\mathbb{Q}}^{K}(\alpha\beta) = N_{\mathbb{Q}}^{K}(\alpha)N_{\mathbb{Q}}^{K}(\beta)$ is a product of integers, so they must be ± 1.

Otherwise, suppose $\alpha \in \mathcal{O}_K$ and $N_{\mathbb{Q}}^{K}(\alpha) = 1$, so that $N_{\mathbb{Q}}^{\mathbb{Q}(\alpha)}(\alpha) = \pm 1$. Then if σ_i are the embeddings $\mathbb{Q}(\alpha) \hookrightarrow \mathbb{C}$ fixing $\mathbb{Q}$, $\sigma_1 = \mathrm{id}$,

$$\pm 1 = \prod_{i=1}^{n} \sigma_i(\alpha) = \alpha \prod_{i=2}^{n} \sigma_i(\alpha)$$

Note that each $\sigma_i(\alpha) \in \mathcal{O}_{\overline{\mathbb{Q}}}$, but since $\mathbb{Q}(\alpha)$ may not be normal, $\sigma_i(\alpha)$ may not be in $\mathbb{Q}(\alpha)$. However $\prod_{i=2}^{n} \sigma_i(\alpha) = \pm\alpha^{-1} \in \mathbb{Q}(\alpha)$ is an algebraic integeer and thus in $\mathcal{O}_K$, so α is a unit.■

Example. In $K = \mathbb{Q}(i)$, $\mathcal{O}_K^{\times} = \mathbb{Z}[i]^{\times}$ and $N(a+bi) = a^2 + b^2$. Thus the units are given by $\{\pm 1, \pm i\}$ More generally, if ζ is a root of unity and $\zeta \in K$, then $\zeta \in \mathcal{O}_K^{\times}$. This follows since $N_{\mathbb{Q}}^{\mathbb{Q}(\zeta)}(\zeta) = 1$ and we can apply (Thm. 5.2).

Units in Quadratic Extensions

5.5 Proposition. *Let d be a square-free negative integer. Then $\mathcal{O}_{\mathbb{Q}(\sqrt{d})}^{\times} = \{\pm 1\}$ unless*

- *$d = -1$, in which case the units are $\{\pm 1, \pm i\}$.*
- *$d = -3$, in which case the units are $\left\{\pm 1, \frac{\pm 1 \pm \sqrt{-3}}{2}\right\}$.*

Proof First suppose $\alpha \in \mathcal{O}_{\mathbb{Q}(\sqrt{d})}^{\times}$, where d is square-free. If $d \not\equiv 1 \pmod 4$, then $\alpha = a + b\sqrt{d}$, so $\alpha \in \mathbb{Z}[\sqrt{d}]^{\times}$ if and only if $N(\alpha) = a^2 - db^2 = \pm 1$. So $a + b\sqrt{d}$ is a unit if and only if (a, b) is a solution to the diophantine equation $x^2 - dy^2 = \pm 1$. Similarly, $\frac{a+b\sqrt{d}}{2} \in \mathcal{O}_{\mathbb{Q}(\sqrt{d})}$ for $d \equiv 1 \pmod 4$ is a unit if and only if $a^2 - db^2 = \pm 4$. Now suppose additionally that $d < 1$.

Case 1: $d \not\equiv 1 \pmod 4$. If $d < -1$, then the only solution to $x^2 - dy^2 = \pm 1$ is $(\pm 1, 0)$. If $d = -1$, then solutions to $x^2 + y^2 = \pm 1$ are $(\pm 1, 0)$ and $(0, \pm 1)$.

Case 2: $d \equiv 1 \pmod 4$. We want solutions to $x^2 - dy^2 = \pm 4$. If $d < -3$, then the only solutions are $(\pm 2, 0)$, which correspond to $\{\pm 1\} \in \mathcal{O}_K$. If $d = -3$, then the solutions are $(\pm 1, 0)$ and $(0, \pm 1)$. ∎

Remark. When $d < 0$, the graph of $x^2 - dy^2 = \{\pm 1, \pm 4\}$ is an ellipse so there are only a finite number of integer pair solutions. On the other hand, consider $d = 2$, so the graph is a hyperbola with asymptotes $\pm\sqrt{2}$. If we want integer solutions, we want solutions b/a that are close to $\sqrt{2}$, so we're looking for (good) rational approximations to $\sqrt{2}$. In a precise sense, one can define the "best" rational approximation to $\sqrt{2}$. One intuition about "best" is to bound the denominator and be close to $\sqrt{2}$. Given α, its continued fraction approximation of α is

$$\alpha = a_0 + \cfrac{1}{a_1 + \frac{1}{a_2 + \cdots}}$$

The first few convergents to the continued fraction expansion of $\sqrt{2}$ are $1, 3/2, 7/5$.

Consider $\epsilon = 1 + \sqrt{2} \in \mathbb{Z}[\sqrt{2}]^\times$, so $N(\epsilon) = -1$. As well, ϵ^n is also a unit for any n. For example, $\epsilon^2 = 3 + 2\sqrt{2}$, $\epsilon^3 = 7 + 5\sqrt{2}$. It turns out that $\epsilon^n = p_n + q_n\sqrt{2}$, where p_n/q_n is the n^{th} convergent of the continued fraction expansion of $\sqrt{2}$.

t:dir-approx

5.6 Theorem. (Dirichlet Approximation) *Let $\alpha \in \mathbb{R} \setminus \mathbb{Q}$, let $Q > 1$, $Q \in \mathbb{Z}$. Then there exists $p, q \in \mathbb{Z}$ such that $1 \le q \le Q$ and $|q\alpha - p| < \frac{1}{Q}$. In particular, there are infinitely many pairs $(p, q) \in \mathbb{Z}^2$ for which $|\alpha - p/q| < 1/q^2$.*

Proof The "in particular" statement follows from the first statement because $|\alpha - p/q| < \frac{1}{Qq} \le \frac{1}{q^2}$. Since Q can be chosen arbitrarily, there are infinitely many such solutions.

Let's now prove the main statement. For any $x \in \mathbb{R}$, let $\{x\} = x - \lfloor x \rfloor$ denote the integer part of x. Consider the Q intervals

$$\left\{\left(0, \frac{1}{Q}\right), \left(\frac{1}{Q}, \frac{2}{Q}\right), \ldots, \left(\frac{Q-1}{Q}, 1\right)\right\}$$

and consider the $Q+1$ numbers $\{\{\alpha\}, \{2\alpha\}, \ldots, \{(Q+1)\alpha\}\}$. Since α is irrational, each of these numbers lies in one of the above intervals. By the pidgeonhole principle, get $1 \le m < n \le Q$ such that $|\{n\alpha\} - \{m\alpha\}| < 1/Q$ so that

$$\left|n\alpha - \lfloor n\alpha \rfloor - m\alpha + \lfloor m\alpha \rfloor\right| = \left|(n-m)\alpha - (\lfloor n\alpha \rfloor - \lfloor m\alpha \rfloor)\right| < \frac{1}{Q}$$

Take $q = n - m$, $p = \lfloor n\alpha \rfloor - \lfloor m\alpha \rfloor$, and we are done. ∎

t:dir-unit-q

5.7 Theorem. (Dirichlet Unit, Quadratic Extensions) *If $d > 1$ be squarefree and set $K = \mathbb{Q}(\sqrt{d})$. Then, there exists a smallest unit $\epsilon > 1$ such that*

$$\mathcal{O}_K^\times = \{\pm\epsilon^n : n \in \mathbb{Z}\} \cong \mathbb{Z}_2 \times \mathbb{Z}$$

Proof We treat the case where $d \not\equiv 1 \pmod 4$; the proof when $d \equiv 1 \pmod 4$ follows identically.

Let $\theta = p + q\sqrt{d}$, $p, q \in \mathbb{Z}$, $q > 0$. Then,

$$|N(\theta)| = |p + q\sqrt{d}||p - q\sqrt{d}| = \left|\frac{p}{q} + \sqrt{d}\right|\left|\frac{p}{q} - \sqrt{d}\right| q^2$$

By Dirichlet approximation (Thm. 5.6), there are infinitely many pairs $(p,q) \in \mathbb{Z}^2$ such that $|p/q - \sqrt{d}| < 1/q^2$. For such (p,q), $\left|\frac{p}{q} + \sqrt{d}\right| < 2\sqrt{d} + 1$. Since $2\sqrt{d}+1$ is independent of the value of (p,q), by the pidgeonhole principle, there exists $m \in \mathbb{Z}^+$ such that there are infinintely many $\theta = p + q\sqrt{d}$ with $|N(\theta)| = m$. Enumerate these by $\theta_i = p_i + q_i\sqrt{d}$ for $i \in \mathbb{N}$.

Let's show that $\mathcal{O}_K^\times$ is an infinite set. We might take θ_i/θ_1 for infinitely many θ_i (which certainly has norm 1), but θ_i/θ_1 might not be an algebraic integer. We can, however, amend this as follows. By the pidgeonhole principle, there exists some $\theta_0 := \theta_j$ such there are infinitely many θ_i with $p_i \equiv p_0 \pmod m$ and $q_i \equiv q_0 \pmod m$. Let θ_0' be the conjugate of θ_0, so that

$$\begin{aligned}\frac{\theta_i}{\theta_0} &= 1 + \frac{\theta_i - \theta_0}{\theta_0} = 1 + \frac{\theta_i - \theta_0}{\theta_0\theta_0'}\theta_0' \\ &= 1 + \frac{(p_i - p_0) + (q_i - q_0)\sqrt{d}}{m}\theta_0' \in \mathcal{O}_K\end{aligned}$$

Thus, we have infinitely many $\beta \in \mathcal{O}_K^\times$.

Now, let $S = \{\gamma \in \mathcal{O}_K^\times : \gamma > 0\}$, so $|S| = \infty$; let's show that S has a minimal element. Assuming this, let $\epsilon \in S$ be minimal and set $\lambda \in \mathcal{O}_K^\times$: taking $-\lambda$ if necessary, we may assume $\lambda > 0$. Then there exists $n \in \mathbb{Z}$ so that $\epsilon^n \le \lambda < \epsilon^{n+1}$. Then $1 \le \lambda\epsilon^n < \epsilon$, and since $\epsilon > 1$ is minimal, we must have $\lambda/\epsilon^n = 1$; i.e. $\lambda = \epsilon^n$.

Note the following: if $1 < \gamma = x + y\sqrt{d}$ is a unit, then $x, y \ge 1$. To see this, consider the four values $\gamma, -\gamma, \gamma^{-1}, -\gamma^{-1}$, which are $\frac{\pm x \pm y\sqrt{d}}{2}$. Since x and x^{-1} cannot both be greater than 1, exactly one of the four values are greater than 1, so it must be the largest one; i.e. the one with $x, y \ge 1$. But now let $\gamma > 1$ be arbitrary; by positivity, there are only finitely many $\gamma_0 < \gamma$, so there must be some minimal element. ∎

Remark. A natural question is to ask this question for a general number field. For example, if K is cubic, then $\mathcal{O}_K^\times$ may or may not have a smallest unit $\epsilon > 1$; see (Thm. 14.9) for the general case.

6 Discriminants and Integral Bases

Definition. Let K be a number field, and let $\sigma_1, \dots, \sigma_n : K \hookrightarrow \mathbb{C}$ be embeddings extending $\mathbb{Q} \subseteq \mathbb{C}$. Given $\alpha_1, \dots, \alpha_n \in K$, we define the **discriminant of $\alpha_1, \dots, \alpha_n$** to be

$$\operatorname{disc}(\alpha_1, \dots, \alpha_n) = \det\begin{pmatrix} \sigma_1(\alpha_1) & \dots & \sigma_1(\alpha_n) \\ \vdots & \ddots & \vdots \\ \sigma_n(\alpha_1) & \dots & \sigma_n(\alpha_n) \end{pmatrix}^2$$

If $K = \mathbb{Q}(\alpha)$, for notational simplicity, we say $\operatorname{disc}(\alpha) = \operatorname{disc}(1, \alpha, \dots, \alpha^{n-1})$.

Remark. The value of $\operatorname{disc}(\alpha_1, \dots, \alpha_n)$ is independent of the ordering of the α_i: swapping rows or columns only changes sign in the determinant.

Example. If $K = \mathbb{Q}(\sqrt{d})$, then

$$\operatorname{disc}(1, \sqrt{d}) = \det\begin{pmatrix} 1 & 1 \\ \sqrt{d} & -\sqrt{d} \end{pmatrix}^2 = 4d$$

6.1 Proposition. *Let K be a number field of degree n. Then*

$$\operatorname{disc}(\alpha_1,\dots,\alpha_n) = \det\begin{pmatrix} \operatorname{Tr}_{\mathbb{Q}}^K(\alpha_1\alpha_1) & \dots & \operatorname{Tr}_{\mathbb{Q}}^K(\alpha_1\alpha_n) \\ \vdots & \ddots & \vdots \\ \operatorname{Tr}_{\mathbb{Q}}^K(\alpha_n\alpha_1) & \dots & \operatorname{Tr}_{\mathbb{Q}}^K(\alpha_n\alpha_n) \end{pmatrix}$$

Proof Let $M = (\sigma_i(\alpha_j))_{ij}$. Then

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

where

$$\left(M^tM\right)_{ij} = \sum_{k=1}^n M_{ik}M_{jk} = \sum_{k=1}^n \sigma_k(\alpha_i)\sigma_k(\alpha_j) = \sum_{k=1}^n \sigma_k(\alpha_i\alpha_j) = \operatorname{Tr}(\alpha_i\alpha_j)$$ ■

Example. Again, take $K = \mathbb{Q}(\sqrt{d})$. Then

$$\operatorname{disc}(1,\sqrt{d}) = \det\begin{pmatrix} \operatorname{Tr}(1) & \operatorname{Tr}(\sqrt{d}) \\ \operatorname{Tr}(\sqrt{d}) & \operatorname{Tr}(d) \end{pmatrix} = \det\begin{pmatrix} 2 & 0 \\ 0 & 2d \end{pmatrix} = 4d$$

which is the same as the previous example.

6.2 Corollary. *We have* $\operatorname{disc}(\alpha_1,\dots,\alpha_n) \in \mathbb{Q}$*, and if* $\alpha_i \in \mathcal{O}_K$*, then* $\operatorname{disc}(\alpha_1,\dots,\alpha_n) \in \mathbb{Z}$*.*

Proof $\operatorname{Tr}(\alpha_i\alpha_j) \in \mathbb{Q}$ and if the $\alpha_i \in \mathcal{O}_K$, then $\operatorname{Tr}(\alpha_i\alpha_j) \in \mathbb{Z}$. ■

Change of Basis

Let's now understand how discriminants change under change of basis. Suppose $\alpha_1,\dots,\alpha_n$ is a basis for $K/\mathbb{Q}$, and let $\beta_1,\dots,\beta_n \in K$ are arbitrary (possibly not a basis). Since $\sigma_i(\beta_k) \in K$, there exists c_{kj} such that $\sigma_i(\beta_k) = \sum_{j=1}^n c_{kj}\sigma_i(\alpha_j)$. Then

$$\begin{pmatrix} \sigma_1(\beta_1) & \cdots & \sigma_n(\beta_1) \\ \vdots & & \vdots \\ \sigma_1(\beta_n) & \cdots & \sigma_n(\beta_n) \end{pmatrix} = \begin{pmatrix} c_{11} & \cdots & c_{1n} \\ \vdots & & \vdots \\ c_{n1} & \cdots & c_{nn} \end{pmatrix}\begin{pmatrix} \sigma_1(\alpha_1) & \cdots & \sigma_n(\alpha_1) \\ \vdots & & \vdots \\ \sigma_1(\alpha_n) & \cdots & \sigma_n(\alpha_n) \end{pmatrix}$$

Let $C = \left(c_{ij}\right)$ denote the above transition matrix; then,

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

Now, if $K/\mathbb{Q}$ is a finite extension, then we know there exists $\theta \in K$ such that $K = \mathbb{Q}(\theta)$. Thus, $\{1,\theta,\dots,\theta^{n-1}\}$ is a basis for $K/\mathbb{Q}$. In particular,

$$\begin{aligned} \operatorname{disc}(1,\theta,\dots,\theta^{n-1}) &= \det\begin{pmatrix} \sigma_1(1) & \sigma_1(\theta) & \cdots & \sigma_1(\theta^{n-1}) \\ \vdots & \vdots & & \vdots \\ \sigma_n(1) & \sigma_n(\theta) & \cdots & \sigma_n(\theta^{n-1}) \end{pmatrix}^2 \\ &= \det\begin{pmatrix} \sigma_1(1) & \sigma_1(\theta) & \cdots & \sigma_1(\theta)^{n-1} \\ \vdots & \vdots & & \vdots \\ \sigma_n(1) & \sigma_n(\theta) & \cdots & \sigma_n(\theta)^{n-1} \end{pmatrix}^2 \\ &= \prod_{i<j}(\sigma_i(\theta) - \sigma_j(\theta))^2 \end{aligned}$$

since it is the square of the determinant of a Vandermonde matrix. In particular, this value is non-zero since the $\sigma_i(\theta)$ are distinct. Now the following proposition follows from this discussion:

6.3 Proposition. *Let $\alpha_1,\dots,\alpha_n \in K$ where $n = [K:\mathbb{Q}]$. Then $\operatorname{disc}(\alpha_1,\dots,\alpha_n) \neq 0$ if and only if $\alpha_1,\dots,\alpha_n$ is a basis for $K/\mathbb{Q}$.*

Proof Let C denote the transition matrix for $\{\alpha_1,\dots,\alpha_n\}$ in terms of the (θ^j). Then

$$\operatorname{disc}(\alpha_1,\dots,\alpha_n) = \det(C)^2 \operatorname{disc}(1,\theta,\dots,\theta^{n-1})$$

so that $\operatorname{disc}(\alpha_1,\dots,\alpha_n) = 0$ if and only if $\det(C) = 0$ if and only if $\alpha_1,\dots,\alpha_n$ are linearly dependent. ∎

t:disc-pb

6.4 Theorem. *Let $K = \mathbb{Q}(\theta)$, $[K:\mathbb{Q}] = n$. Then*

$$\operatorname{disc}(\theta) := \operatorname{disc}(1,\theta,\dots,\theta^{n-1}) = (-1)^{\binom{n}{2}} N_{\mathbb{Q}}^{K}(f'(\theta))$$

where $f(x) \in \mathbb{Q}[x]$ is the minimal polynomial of θ over $\mathbb{Q}$.

Proof Let $\theta_1,\dots,\theta_n$ be the conjugates of θ. Then $f(x) = \prod_{i=1}^{n}(x-\theta_i)$, so $f'(x) = \sum_{j=1}^{n}\prod_{i\neq j}(x-\theta_i)$. Thus

$$\begin{aligned} N_{\mathbb{Q}}^{K}(f'(\theta)) &= \prod_{k=1}^{n}\sigma_n(f'(\theta)) = \prod_{k=1}^{n} f'(\theta_k) \\ &= \prod_{k=1}^{n}\prod_{i\neq k}(\theta_k-\theta_i) = \prod_{i<k}(\theta_k-\theta_i)(\theta_i-\theta_k) \\ &= (-1)^{\binom{n}{2}}\prod_{i<k}(\theta_i-\theta_k)^2 = \operatorname{disc}(1,\theta,\dots,\theta^{n-1}) \end{aligned}$$

∎

Cyclotomic Extensions II: Discriminants

t:disc-cycl

6.5 Theorem. *Let $\zeta_n = e^{2\pi i/n}$ and set $d = \operatorname{disc}\left(1,\zeta_n,\dots,\zeta_n^{\phi(n)-1}\right)$. Then $d \mid n^{phi(n)}$, and if p is an odd prime, $d = (-1)^{\binom{p}{2}}p^{p-2}$.*

Proof Let $\Phi_n(x)$ be the minimal polynomial of ζ_n, and write $x^n - 1 = \Phi_n(x)g(x)$ where $g(x) \in \mathbb{Z}[x]$. Then $nx^{n-1} = \Phi_n'(x)g(x) + \Phi_n(x)g'(x)$, so $n\zeta_n^{n-1} = \Phi'(\zeta_n)g(\zeta_n)$. Thus

$$N\left(n\zeta_n^{n-1}\right) = N\left(\Phi'(\zeta_n)\right)\cdot N\left(g(\zeta_n)\right)$$

Since $\zeta_n \in \mathcal{O}_{\mathbb{Q}(\zeta_n)}^{\times}$, $N(\zeta_n) = \pm 1$. Thus

$$\pm n^{\phi(n)} = (-1)^{\binom{\phi(n)}{2}} N\left(\Phi_n'(\zeta_n)\right)\cdot N(g(\zeta_n))$$

so $\pm\operatorname{disc}(\zeta_n)N(g(\zeta_n)) = n^{\phi(n)}$. Since $g \in \mathbb{Z}[x]$, $g(\zeta_n) \in \mathcal{O}_{\mathbb{Q}(\zeta_n)}$ and $N(g(\zeta_n)) \in \mathbb{Z}$. Thus $\operatorname{disc}(\zeta_n) \mid n^{\phi(n)}$, as required.

Now, if p is an odd prime, $x^p - 1 = \Phi_p(x)(x-1)$, so $px^{p-1} = \Phi_p'(x)(x-1) + \Phi_p(x)$. Thus $p\zeta_p^{p-1} = \Phi_p(\zeta_p)(\zeta_p - 1)$. Note that $N(\zeta_p^{p-1}) = N(\zeta_p)^{p-1} = 1$ and since $p-1$ is even. We can also compute

$$N(\zeta_p - 1) = (-1)^{p-1}\prod_{i=1}^{p-1}(1-\zeta_p^i) = \Phi_p(1) = p$$

so that

$$\begin{aligned} p\zeta_p^{p-1} = \Phi_p'(\zeta_p)(\zeta_p - 1) &\Rightarrow p^{p-1} = N(\Phi_p'(\zeta_p))p \\ &\Rightarrow (-1)^{\binom{p}{2}}p^{p-2} = \operatorname{disc}(\zeta_p) \end{aligned}$$

as required. ∎

Remark. In general, we have

$$\operatorname{disc}\left(1, \zeta_n, \ldots, \zeta_n^{\phi(n)-1}\right) = (-1)^{\phi(n)/2}\frac{n^{\phi(n)}}{\prod_{p|n} p^{\phi(n)/(p-1)}}$$

which we state here without proof.

Integral Bases

Definition. Let K be a number field, $[K:\mathbb{Q}] = n$. We say $A = \{\alpha_1, \ldots, \alpha_n\}$ is an **integral basis** for K if $\mathcal{O}_K = \operatorname{span}_{\mathbb{Z}}(A)$. When it exists, a **power basis** for $\mathcal{O}_K$ is an integral basis of the form $\{1, \alpha, \cdots, \alpha^{n-1}\}$; i.e. $\mathcal{O}_K = \mathbb{Z}[\alpha]$.

Remark. Clearly we must have $\alpha_i \in \mathcal{O}_K$. As well, $\alpha_1, \ldots, \alpha_n$ is a basis for $K/\mathbb{Q}$: given $\theta \in K$, there exists $r \in \mathbb{Z}^+$ such that $r\theta \in \mathcal{O}_K$, so $r\theta \in \operatorname{span}_{\mathbb{Z}}(A)$ and $\theta \in \operatorname{span}_{\mathbb{Q}}(A)$. Since $[K:\mathbb{Q}] = n$, $\alpha_1, \ldots, \alpha_n$ is a basis. In particular, this means that A is in fact a $\mathbb{Z}$–basis for $\mathcal{O}_K$ (justifying the terminology).

6.6 Theorem. *If K is a number field, then K has an integral basis.*

Proof Write $K = \mathbb{Q}(\theta)$ where $\theta \in \mathcal{O}_K$. Consider the set of all bases $\{\beta_1, \ldots, \beta_n\}$ for $K/\mathbb{Q}$ such that $\beta_i \in \mathcal{O}_K$. Such a basis certainly exists; given any basis, we can clear denominators such that they are in $\mathcal{O}_K$. Let A have $|\operatorname{disc}(A)|$ minimal (the discriminant is an integer, so such an A exists); let's show that A is in fact an integral basis.

Suppose not. Then there exists $\gamma \in \mathcal{O}_K$ where $\gamma = a_1\alpha_1 + \cdots + a_n\alpha_n$ and $a_1 \notin \mathbb{Z}$. Let $a_1 = a + r$ with $a \in \mathbb{Z}$, $0 < r < 1$; consider the basis $\{\alpha_1', \ldots, \alpha_n'\}$ where $\alpha_i' = \alpha_i$ for $i > 1$, and $\alpha_1' = \gamma - a\alpha_1$. Then

$$\begin{aligned} \operatorname{disc}(\alpha_1', \alpha_2', \ldots, \alpha_n') &= \det\begin{pmatrix} a_1 - a & a_2 & a_3 & \cdots & a_n \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{pmatrix}^2 \operatorname{disc}(\alpha_1, \ldots, \alpha_n) \\ &= r^2 \operatorname{disc}(\alpha_1, \ldots, \alpha_n) \end{aligned}$$

Since $0 < r < 1$, $|\operatorname{disc}(\alpha_1', \ldots, \alpha_n')| < |\operatorname{disc}(\alpha_1, \ldots, \alpha_n)|$, contradicting minimality. ∎

6.7 Proposition. *If K is a number field, then all integral bases have the same discriminant.*

Proof Let $\{\alpha_1,\dots,\alpha_n\}$ and $\{\beta_1,\dots,\beta_n\}$ be two integral bases; then

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

for $\alpha_j \in \mathcal{O}_K$ and $c_{ij} \in \mathbb{Z}$. Let $C = (c_{ij})$. Since $\{\alpha_1,\dots,\alpha_n\}$ is also an integral basis, $\left(C^{-1}\right)_{ij} \in \mathbb{Z}$ as well. Thus $C \in \mathrm{GL}_n(\mathbb{Z})$ so $\det(C)^2 = 1$ and

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

indeed have the same discriminant. ∎

Definition. If K is a number field, we say its **discriminant** $\mathrm{disc}(K)$ is the discriminant of any integral basis.

Example. (Quadratic Extensions) Consider $\mathbb{Q}(\sqrt{d})$. If $d \not\equiv 1 \pmod 4$, then $\{1, \sqrt{d}\}$ is an integral basis; if $d \equiv 1 \pmod 4$, then $\left\{1, \frac{1+\sqrt{d}}{2}\right\}$ is an integral basis. Thus

$$\mathrm{disc}(\mathbb{Q}(\sqrt{d})) = \begin{cases} 4d & d \not\equiv 1 \pmod 4 \\ d & d \equiv 1 \pmod 4 \end{cases}$$

p:ext

6.8 Proposition. *Let K be a number field, $\{\alpha_1,\dots,\alpha_n\}$ a basis for $K/\mathbb{Q}$ with $\alpha_i \in \mathcal{O}_K$. If $d = \mathrm{disc}(\alpha_1,\dots,\alpha_n)$, then for all $\alpha \in \mathcal{O}_K$, there exists $m_i \in \mathbb{Z}$ such that*

$$\alpha = \frac{\sum_{i=1}^{n} m_i\alpha_i}{d} \qquad\qquad d \,\Big|\, m_i^2$$

Example. Consider $\mathbb{Q}(\sqrt{d})$, where $d \equiv 1 \pmod 4$. Then $\{1, \sqrt{d}\}$ is a $\mathbb{Q}$–basis, $\sqrt{d} \in \mathcal{O}_K$, and $\mathrm{disc}(1, \sqrt{d}) = 4d$. Since d is squarefree, if $4d \,\Big|\, m_i^2$, then $d \,\Big|\, m_i$. Thus, the proposition states that any $\gamma \in \mathcal{O}_K$ can be expressed in the form $\frac{m_1+m_2\sqrt{d}}{2}$ for some $m_1, m_2 \in \mathbb{Z}$. Note that the converse is not necessarily true: not all such expressions are in $\mathcal{O}_K$ (indeed, we need $m_1 \equiv m_2 \pmod 2$).

Proof Let $\alpha \in \mathcal{O}_K$ be arbitrary so that $\alpha = a_1\alpha_1 + \cdots + a_n\alpha_n$ for some $a_i \in \mathbb{Q}$. Let $\sigma_1,\dots,\sigma_n : K \hookrightarrow \mathbb{C}$ extend $\mathbb{Q} \subseteq \mathbb{C}$. For each $j = 1,\dots,n$, we have $\sigma_j(\alpha) = a_1\sigma_j(\alpha_1) + \cdots + a_n\sigma_j(\alpha_n)$ so that

$$\begin{pmatrix} \sigma_1(\alpha_1) & \dots & \sigma_1(\alpha_n) \\ \vdots & & \vdots \\ \sigma_n(\alpha_1) & \dots & \sigma_n(\alpha_n) \end{pmatrix} \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} = \begin{pmatrix} \sigma_1(\alpha) \\ \vdots \\ \sigma_n(\alpha) \end{pmatrix}$$

Define

$$\gamma_j := \det \begin{pmatrix} \sigma_1(\alpha_1) & \cdots & \sigma_1(\alpha) & \dots & \sigma_1(\alpha_n) \\ \vdots & & \vdots & & \vdots \\ \sigma_n(\alpha_1) & \dots & \sigma_n(\alpha) & \dots & \sigma_n(\alpha_n) \end{pmatrix}$$

where the j^{th} column is replaced, and $\delta = \det(\sigma_j(\alpha_i))$. Since $\alpha_i \in \mathcal{O}_K$, $\sigma_j(\alpha_i) \in \mathcal{O}_K$ for any j, so $\gamma_j, \delta \in \mathcal{O}_K$. Note that $d := \mathrm{disc}(K) = \delta^2$. By Cramer's rule, $a_j = \frac{\gamma_j}{\delta}$. Take $m_j := da_j \in \mathbb{Q}$; but then $da_j = \delta\gamma_j \in \mathcal{O}_K$, so $m_j \in \mathbb{Q} \cap \mathcal{O}_K = \mathbb{Z}$.

For the second part, we have

$$\frac{m_j^2}{d} = da_j^2 = d\left(\frac{\gamma_j}{\delta}\right)^2 = \frac{d\gamma_j^2}{\delta^2} = \gamma_j^2 \in \mathcal{O}_K$$

so $m_j^2/d \in \mathbb{Z}$ as well. ∎

Real and Complex Embeddings

Let K be a number field and let $\sigma_1, \dots, \sigma_n : K \hookrightarrow \mathbb{C}$ be the embeddings extending $\mathbb{Q} \subseteq \mathbb{C}$. Let r_1 denote the number of embeddings where $K \hookrightarrow \mathbb{R}$; then, the other embeddings come in pairs: if $\sigma : K \hookrightarrow \mathbb{C}$, then $\overline{\sigma} : K \hookrightarrow \mathbb{C}$ is a (distinct) embedding.

We say that r_1 is the number of real embeddings, and $2r_2$ is the number of complex embeddings; in this case, $n = r_1 + 2r_2$.

Example. Let d be squarefree. Then $\mathbb{Q}(\sqrt{d})$ for $d > 0$ has $r_1 = 2$, $r_2 = 0$, while $\mathbb{Q}(\sqrt{d})$ for $d < 0$ has $r_1 = 0$, $r_2 = 1$.

6.9 Proposition. *Let $[K : \mathbb{Q}] = n$; then, the sign of* $\operatorname{disc}(K)$ *is* $(-1)^{r_2}$.

Proof Let $\alpha_1, \dots, \alpha_n$ be an integral basis for $K/\mathbb{Q}$. Consider

$$\delta = \det\begin{pmatrix} \sigma_1(\alpha_1) & \dots & \sigma_1(\alpha_n) \\ \vdots & & \vdots \\ \sigma_n(\alpha_1) & \dots & \sigma_n(\alpha_n) \end{pmatrix} \qquad \overline{\delta} = \det\begin{pmatrix} \overline{\sigma_1}(\alpha_1) & \dots & \overline{\sigma_1}(\alpha_n) \\ \vdots & & \vdots \\ \overline{\sigma_n}(\alpha_1) & \dots & \overline{\sigma_n}(\alpha_n) \end{pmatrix}$$

where $\operatorname{disc}(K) = \delta^2$. If σ_i is real, then $\overline{\sigma_i} = \sigma_i$. If (σ_i, σ_j) are complex conjugate pairs, then in $\overline{\delta}$ we swap column i with column j. Thus $\overline{\delta} = (-1)^{r_2}\delta$, so δ is purely imaginary if r_2 is odd, and real if r_2 is even. This proves the claim. ∎

Cyclotomic Extensions III: Algebraic Integers in $\mathbb{Q}(\zeta_{p^r})$

t:cepr

6.10 Theorem. *If p is prime, $r \in \mathbb{Z}^+$, then $\mathcal{O}_{\mathbb{Q}(\zeta_{p^r})} = \mathbb{Z}[\zeta_{p^r}]$.*

Proof For notation $\zeta = \zeta_{p^r}$ and we take $\mathbb{Q}(\zeta) = \mathbb{Q}(1-\zeta)$.

Let $s = \phi(p^r)$, so $\{1, 1-\zeta, \dots, (1-\zeta)^{s-1}\}$ is a $\mathbb{Q}$–basis for $\mathbb{Q}(\zeta)$. Let's show that it is an integral basis. By (Prop. 6.8), we know if $\alpha \in \mathcal{O}_K$, there exist $m_i \in \mathbb{Z}$ such that $\alpha = \frac{\sum_{i=1}^n m_j(1-\zeta)^j}{d}$ where

$$\begin{aligned} d = \operatorname{disc}(1-\zeta) &= \prod_{\substack{i<j \\ i,j \in (\mathbb{Z}/p^r)^\times}} ((1-\zeta^i) - (1-\zeta^j))^2 \\ &= \prod_{\substack{i<j \\ i,j \in (\mathbb{Z}/p^r)^\times}} (\zeta^i - \zeta^j)^2 = \operatorname{disc}(\zeta) = \pm p^{p-2} \end{aligned}$$

Let's first treat the case $\beta = \frac{l_1 + l_2(1-\zeta) + \cdots + l_s(1-\zeta)^{s-1}}{p}$. Let i be minimal so that $p \nmid l_i$. Set

$$\gamma = \frac{l_i(1-\zeta)^{i-1} + \cdots + l_s(1-\zeta)^{s-1}}{p} \in \mathcal{O}_K$$

Since $(1-x) \mid (1-x^j)$ in $\mathbb{Z}[x]$, $(1-\zeta) \mid (1-\zeta^j)$ in $\mathcal{O}_K$ so that

$$(1-\zeta)^s \mid \prod_{p\nmid j}(1-\zeta^j) = \Phi_{p^r}(1) = p$$

over $\mathcal{O}_K$. Thus $p = (1-\zeta)^s\lambda$ for some $\lambda \in \mathcal{O}_K$. Since $\lambda, \gamma, 1-\zeta \in \mathcal{O}_K$, $(1-\zeta)^{s-i}\lambda\gamma \in \mathcal{O}_K$. However,

$$(1-\zeta)^{s-i}\lambda\gamma = \lambda\frac{l_i(1-\zeta)^{s-1}}{p} + \lambda\frac{l_{i+1}(1-\zeta)^s}{p} + \cdots$$

where the tail terms are all algebraic integers, so

$$\theta := \frac{l_i}{1-\zeta} = \lambda\frac{l_i(1-\zeta)^{s-1}}{p} \in \mathcal{O}_K$$

Then $(1-\zeta)\theta = l_i$ and, taking norms, $N(1-\zeta)N(\theta) = N(l_i)$ so that $pN(\theta) = l_i^s$ and $p \mid l_i$ and no such l_i exists. But now since $d = \pm p^{p-2}$, we may repeat the above argument for each factor of p, and we are done. ∎

Remark. This proof demonstrates a general tool for verifying that a given basis of algebraic integers is indeed integral. One need simply check each prime p such that $p^2 \mid d$; if there are no algebraic integers of the form $\alpha = \frac{m_1\beta_1+\cdots+m_n\beta_n}{p}$ where $|m_i| < p$ for every such p, then β is indeed an integral basis.

If there is some α of this form, then update $\{\beta_1, \ldots, \beta_n\}$ with the new algebraic integer α; the new discriminant is d/p^2, and we may repeat the above process. This process will terminate after a finite number of steps (though it may take a while), giving a general procedure to compute integral bases for arbitrary number fields.

7 Composita and Resultants

Composita

Definition. If K, L are number fields, then the **compositum** of K and L is the smallest field containing $K \cup L$. We denote it by $KL = LK$.

Our goal in this section is to relate $\mathcal{O}_K$, $\mathcal{O}_L$, and $\mathcal{O}_{KL}$.

l:ext-comp

7.1 Lemma. *Suppose* $[K:\mathbb{Q}] = m$, $[L:\mathbb{Q}] = n$.

(i) $[KL:\mathbb{Q}] \le mn$.

(ii) $[KL:\mathbb{Q}] = mn$ *if and only if for any embeddings* $\sigma : K \hookrightarrow \mathbb{C}$, $\tau : L \hookrightarrow \mathbb{C}$ *embeddings, there exists a unique embedding* $\epsilon : KL \hookrightarrow \mathbb{C}$ *such that* $\epsilon|_K = \sigma$, $\epsilon|_L = \tau$.

Proof (i) We have $[KL:\mathbb{Q}] = [KL:K]\cdot[K:\mathbb{Q}] \le [L:\mathbb{Q}]\cdot[K:\mathbb{Q}] = mn$.

(ii) Consider $\epsilon|_K : K \hookrightarrow \mathbb{C}$ and $\epsilon|_L : L \hookrightarrow \mathbb{C}$; since the products $\alpha\beta$ for $\alpha \in K$ and $\beta \in L$ generate KL, ϵ is determined uniquely by $\epsilon|_K$ and $\epsilon|_L$. Since $[KL:\mathbb{Q}] = mn$, then there are mn embeddings $\epsilon : KL \hookrightarrow \mathbb{C}$, so this must be all of them, and the equivalence follows. ∎

t:comp

7.2 Theorem. *Let $[K:\mathbb{Q}]=n$, $[L:\mathbb{Q}]=m$, $[KL:\mathbb{Q}]=mn$, and $d=\gcd(\operatorname{disc}(K),\operatorname{disc}(L))$. Then $\mathcal{O}_{KL}\subseteq\frac{1}{d}\mathcal{O}_K\mathcal{O}_L$.*

Proof Let $\{\alpha_1,\dots,\alpha_n\}$ be an integral basis for $K/\mathbb{Q}$ and $\{\beta_1,\dots,\beta_m\}$ an integral basis for $L/\mathbb{Q}$. Then $KL=\operatorname{span}_{\mathbb{Q}}\{\alpha_i\beta_j:(i,j)\in[n]\times[m]\}$. Since $[KL:\mathbb{Q}]=mn$, the $\alpha_i\beta_j$ are a $\mathbb{Q}$–basis of algebraic integers. Then $\alpha\in KL$ can be represented as

$$\alpha=\sum_{i=1}^{m}\sum_{j=1}^{n}\frac{\alpha_i\beta_j a_{ij}}{r}$$

with $a_{ij},r\in\mathbb{Z}$ and $\gcd(a_{11},\dots,a_{nm},r)=1$. If $\alpha\in\mathcal{O}_{KL}$ we want to show that $r\mid\operatorname{disc}(K)$ and $r\mid\operatorname{disc}(L)$ so that $r|d$ and $\alpha\in\frac{1}{d}\mathcal{O}_K\mathcal{O}_L$.

By symmetry, let's show that $r\mid\operatorname{disc}(K)$. Given $\sigma_1,\dots,\sigma_n:K\hookrightarrow\mathbb{C}$, by (Lem. 7.1) there exists $\sigma_i':KL\hookrightarrow\mathbb{C}$ so that $\sigma_i'|_K=\sigma_i$ and $\sigma_i'|_L=\mathrm{id}_L$. Then

$$\sigma_i'(\alpha)=\sum_{i=1}^{m}x_i\sigma(\alpha_i)\qquad\qquad x_i=\sum_{j=1}^{n}\frac{a_{ij}\beta_j}{r}$$

since $x_i\in L$. Equivalently,

$$\begin{pmatrix}\sigma_1(\alpha_1)&\dots&\sigma_1(\alpha_n)\\ \vdots&\ddots&\vdots\\ \sigma_n(\alpha_1)&\dots&\sigma_n(\alpha_n)\end{pmatrix}\begin{pmatrix}x_1\\ \vdots\\ x_n\end{pmatrix}=\begin{pmatrix}\sigma_1'(\alpha)\\ \vdots\\ \sigma_n'(\alpha)\end{pmatrix}$$

Let

$$\gamma_i=\begin{pmatrix}\sigma_1(\alpha_1)&\cdots&\sigma_1'(\alpha_1)&\cdots&\sigma_1(\alpha_n)\\ \vdots&&\vdots&&\vdots\\ \sigma_n(\alpha_1)&\dots&\sigma_n'(\alpha_n)&\cdots&\sigma_n(\alpha_n)\end{pmatrix}$$

where the i^{th} column is replaced. Then by Cramer's rule, $x_i=\frac{\gamma_i}{\delta}$ where $\gamma_i,\delta\in\mathcal{O}_K$ and $\delta^2=\operatorname{disc}(K)$. Thus $\operatorname{disc}(K)x_i=\delta\gamma_i\in\mathcal{O}_K$ is an algebraic integer, but also $\operatorname{disc}(K)\in\mathbb{Z}$ so $\operatorname{disc}(K)x_i\in L$. Thus $\operatorname{disc}(K)x_i\in\mathcal{O}_L$; but then, since

$$\operatorname{disc}(K)x_i=\sum_{j=1}^{m}\left(\frac{\operatorname{disc}(K)a_{ij}}{r}\right)\beta_j$$

and the β_j form an integral basis for $\mathcal{O}_L$, we have $\frac{\operatorname{disc}(K)a_{ij}}{r}\in\mathbb{Z}$. Since $\gcd(a_{11},\dots,a_{mn},r)=1$, this forces $r\mid\operatorname{disc}(K)$. ∎

Cyclotomic Extensions IV: Algebraic Integers in $\mathbb{Q}(\zeta_n)$

7.3 Theorem. *$\mathcal{O}_{\mathbb{Q}(\zeta_n)}=\mathbb{Z}[\zeta_n]$.*

Proof Let's do this by induction on the number of prime factors of n; we already did the base case $n=p^r$ in (Thm. 6.10). For $k\geq 2$ let

$$n=p_1^{e_1}\cdots p_k^{e_k}\qquad m:=p_1^{e_1}\cdots p_{k-1}^{e_{k-1}}\qquad K=\mathbb{Q}(\zeta_m)\qquad L=\mathbb{Q}\left(\zeta_{p_k^{e_k}}\right)$$

First, let's see that $KL = \mathbb{Q}(\zeta_n)$. Note that $\zeta_n \in KL$ since m and $p_k^{e_k}$ are coprime; thus, there exists $x, y \in \mathbb{Z}$ so that $xm + yp_k^{e_k} = 1$. Then $\zeta_m^y \zeta_{p_k^{e_k}}^x = e^{2\pi i/n}$, so $\mathbb{Q}(\zeta_n) \subseteq KL$. As well,

$$\phi(n) = \phi(m)\phi\left(p_k^{e_k}\right) = [K:\mathbb{Q}]\cdot[L:\mathbb{Q}] \geq [KL:\mathbb{Q}] \geq [\mathbb{Q}(\zeta_n):\mathbb{Q}] = \phi(n)$$

so $\mathbb{Q}(\zeta_n) = KL$ and $[KL:\mathbb{Q}] = [K:\mathbb{Q}][L:\mathbb{Q}]$. Thus, $\mathbb{Q}(\zeta_n) = KL$ and $[KL:\mathbb{Q}] = [K:\mathbb{Q}][L:\mathbb{Q}]$ and by (Thm. 7.2), we have

$$\mathbb{Z}[\zeta_n] \subseteq \mathcal{O}_{\mathbb{Q}(\zeta_n)} \subseteq \frac{1}{d}\mathcal{O}_{\mathbb{Q}(\zeta_m)}\mathcal{O}_{\mathbb{Q}(\zeta_{p_k^{e_k}})} = \frac{1}{d}\mathbb{Z}[\zeta_m]\mathbb{Z}[\zeta_{p_k^{e_k}}] = \frac{1}{d}\mathbb{Z}[\zeta_n]$$

where $d = \gcd(\operatorname{disc} K, \operatorname{disc} L)$. Recall by (Thm. 6.5), we have $\operatorname{disc}(\mathbb{Q}(\zeta_n)) \mid n^{\phi(n)}$. Thus $\operatorname{disc}(K) \mid m^{\phi(m)}$ and $\operatorname{disc}(L) \mid (p_k^{e_k})^{\phi(p_k^{e_k})}$ so that $d = 1$ and $\mathcal{O}_{\mathbb{Q}(\zeta_n)} = \mathbb{Z}[\zeta_n]$. ∎

Resultants

Definition. Let $f(x), g(x) \in \mathbb{C}[x]$ with $f(x) = a_n x^n + \cdots + a_1 x + a_0$, $g(x) = b_m x^m + \cdots + b_1 x + b_0$. The **resultant** of f and g is

$$R(f,g) = \det \begin{pmatrix} a_n & a_{n-1} & \cdots & a_1 & a_0 & 0 & \cdots & 0 \\ 0 & a_n & a_{n-1} & \cdots & a_1 & a_0 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & a_n & a_{n-1} & \cdots & a_1 & a_0 \\ b_m & b_{m-1} & \cdots & b_0 & 0 & 0 & \cdots & 0 \\ 0 & b_m & \cdots & b_1 & b_0 & 0 & \cdots & 0 \\ 0 & 0 & b_m & \cdots & b_1 & b_0 & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & 0 \\ 0 & 0 & \cdots & 0 & b_m & \cdots & b_1 & b_0 \end{pmatrix}$$

Remark. $R(f,g)$ is homoeneous of degree m in the a_i and degree n in the b_j.

We want to show that if $f, g \in \mathbb{Q}[x]$, then $R(f,g) = 0$ if and only if f and g have a common factor in $\mathbb{Q}[x]$. In particular, we have the following proposition:

p:res

7.4 Proposition. *Let $f, g \in \mathbb{C}[x]$. The following are equivalent:*

(i) f and g have a common root in $\mathbb{C}$

(ii) There exists $h, k \in \mathbb{C}[x]$ such that $hf = kg$ and $\deg(h) \leq m-1$, $\deg(k) \leq n-1$.

(iii) $R(f,g) = 0$

Proof $(i \Rightarrow ii)$ If f, g have a common root $\alpha \in \mathbb{C}$, then $(x-\alpha) \mid f$ and $(x-\alpha) \mid g$. Then $f = (x-\alpha)k$, $g = (x-\alpha)h$ and $hf = (x-\alpha)kh = kg$.

$(ii \Rightarrow i)$ If $hf = kg$ with $\deg h \leq m-1$, $\deg k \leq n-1$, then by Pigeonhole principle, the roots of k cannot contain all the roots of f, so one root must be a root of g.

$(ii \Leftrightarrow iii)$ We can now turn our question into one of linear algebra. Given f, g, we want to compute h, k such that $hf = kg$ where $\deg h = \deg g - 1$, and $\deg k = \deg f - 1$. Let

$$h = c_{m-1}x^{m-1} + \cdots + c_1 x + c_0$$
$$k = d_{n-1}x^{n-1} + \cdots + d_0$$

Treate c_i, d_j as indeterminants so that the statement $hf = kg$ encodes $n+m$ equations by comparing coefficients of the same degree. For example, the x^{n+m-2} equation $a_n c_{m-2} + a_{n-1}c_{m-1} = b_m d_{n-2} + b_{m-1}d_{n-1}$. In particular, $(c_0,\dots,c_{m-1};-d_0,\dots,-d_{n-1})$ is a solution if and only if it is in the kernel of the matrix

$$A = \begin{pmatrix} a_n & 0 & \cdots & 0 & b_m & 0 & 0 & \cdots & 0 \\ a_{n-1} & a_n & \ddots & \vdots & b_{m-1} & b_m & 0 & \cdots & 0 \\ \vdots & a_{n-1} & \ddots & 0 & \vdots & \vdots & b_m & \ddots & \vdots \\ a_1 & \vdots & \ddots & a_n & b_0 & b_1 & \vdots & \ddots & 0 \\ a_0 & a_1 & \ddots & a_{n-1} & 0 & b_0 & b_1 & \ddots & b_m \\ 0 & a_0 & \ddots & \vdots & 0 & 0 & b_0 & \ddots & \vdots \\ \vdots & \ddots & \ddots & a_1 & \vdots & \vdots & \ddots & \ddots & b_1 \\ 0 & \cdots & 0 & a_0 & 0 & 0 & \cdots & 0 & b_0 \end{pmatrix}$$

and this matrix has non-trivial kernel if and only if $0 = \det(A) = \det(A^t) = R(f,g)$. ∎

Let $x_1,\dots,x_n$ denote the roots of f and $y_1,\dots,y_m$ denote roots of g. Then $a_1,\dots,a_n$ are a_n times an elementary symmetric function in the x_i, $b_1,\dots,b_{m-1}$ are b_m times an elementary symmetric function in the y_i. Thus $R(f,g) \in \mathbb{C}[x_1,\dots,x_n,y_1,\dots,y_m] =: P$ is a symmetric polynomial times $a_n^m b_m^n$. By (Prop. 7.4), if $x_i = y_j$, then $R(f,g) = 0$. In other words, every $(x_i - y_j) \mid R(f,g)$ (as polynomials). Since each $(x_i - y_j)$ is an irreducible coprime factor of P, $\prod_{i,j}(x_i - y_j) \,\Big|\, R(f,g)$. Set $S := a_n^m b_m^n \prod_{i,j}(x_i - y_j)$.

In particular, note that $g(x) = b_m \prod_{j=1}^m (x - y_j)$ so that $a_n^m \prod_{i=1}^n g(x_i) = S$ and

$$S = a_n^m \prod_{i=1}^n g(x_i) \tag{7.1}$$ {eq:S1}

Similarly, $f(x) = a_n \prod_{i=1}^n (x - x_i) = (-1)^n a_n \prod_{i=1}^n (x_i - x)$ so that

$$S = (-1)^{mn} b_m^n \prod_{i=1}^n f(y_i) \tag{7.2}$$ {eq:S2}

(Eq. 7.1) tells us that S is homogeneous of degree n in the b_j's, and (Eq. 7.2) says S is homogenous of degree m in the a_i. Since $R(f,g)$ has the same property and $S \,\Big|\, R(f,g)$, $R = cS$ for some $c \in \mathbb{C}$. However, S has constant term $a_n^m b_m^n$, which is the same as $R(f,g)$; thus, $c = 1$. In particular, we've shown the following proposition:

p:res-alt

7.5 Proposition. *Let $f,g \in \mathbb{C}[x]$ with $\deg f = n$, $\deg g = m$, and f have roots $x_1,\dots,x_n$ and g have roots $y_1,\dots,y_m$ (perhaps with repetitions). Then*

$$R(f,g) = a_n^m b_m^n \prod_{i,j}(x_i - y_j)$$

This gives us an easy way to compute certain types of discriminants:

7.6 Corollary. *Let α be algebraic over $\mathbb{Q}$ and f the minimal polynomial of α. Then* $\operatorname{disc}(\alpha) = (-1)^{\binom{n}{2}} R(f,f')$.

Proof Let's apply (Prop. 7.5) in the case where $g = f'$. Let $f(x) = x^n + a_{n-1}x^{n-1} + \cdots + a_0 = (x-\alpha_1)\cdots(x-\alpha_n)$. Let $\sigma_1, \ldots, \sigma_n : \mathbb{Q}(\alpha) \hookrightarrow \mathbb{C}$ extend $\mathbb{Q} \subseteq \mathbb{C}$. Then applying (Eq. 7.1), we have

$$R(f, f') = \prod_{i=1}^{n} f'(\alpha_i) = \prod_{i=1}^{n} \sigma_i(f'(\alpha)) = N_{\mathbb{Q}}^{\mathbb{Q}(\alpha)}(f'(\alpha))$$

and the result follows by (Thm. 6.4). ∎

As a fun application of this result, let's prove the following proposition. Note that the result was not strictly necessary to do this, but we get to use it to do one of the computations.

7.7 Proposition. *Let θ be a root of $f(x) = x^3 + x^2 - 2x + 8$, and $K = \mathbb{Q}(\theta)$. Then $\mathcal{O}_K$ has no power basis.*

Proof Let's calculate $\mathcal{O}_K$. First, we have

$$\operatorname{disc}(\theta) = -R(f, f') = \det \begin{pmatrix} 1 & 1 & -2 & 8 & 0 \\ 0 & 1 & 1 & -2 & 8 \\ 3 & 2 & -2 & 0 & 0 \\ 0 & 3 & 2 & -2 & 0 \\ 0 & 0 & 3 & 2 & -2 \end{pmatrix}^2 = -4 \cdot 503$$

Thus $\operatorname{disc}(K) = -4 \cdot 503$ or $\operatorname{disc}(K) = -503$ since, under change of basis, the factor must change by a square of an integer. We know from a homework assignment (or by direct computation) that $\mathcal{O}_K \neq \mathbb{Z}[\theta]$ since $(\theta^2 - \theta)/2 \in \mathcal{O}_K$ and $\operatorname{disc}(K) = -503$. In particular, one has that $\operatorname{disc}(1, \theta, \frac{\theta^2-\theta}{2}) = -503$ by change of basis. Since 503 is squarefree, $\left\{1, \theta, \frac{\theta^2-\theta}{2}\right\}$ is an integral basis of $\mathcal{O}_K$.

Now, let $\lambda \in \mathcal{O}_K$. We'll show that $2 \mid \operatorname{disc}(\lambda)$ so that $\operatorname{disc}(\lambda) \neq -503$ and $\{1, \lambda, \lambda^2\}$ is not an integral basis. We can write $\lambda = a + b\theta + c\frac{\theta^2-\theta}{2}$ for $a, b, c \in \mathbb{Z}$. In particular, after some computation, one has $\lambda^2 = A_1 + A_2\theta + A_3\frac{\theta^2-\theta}{2}$, where

$$\begin{aligned} A_1 &= a^2 - 2c^2 - 8bc \\ A_2 &= -2c^2 + 2ab + 2bc - b^2 \\ A_3 &= 2b^2 + 2ac + c^2 \end{aligned}$$

Then by change of basis,

$$\begin{aligned} \operatorname{disc}(\lambda) &= -503 \cdot \det \begin{pmatrix} 1 & 0 & 0 \\ a & b & c \\ A_1 & A_2 & A_3 \end{pmatrix}^2 = -503 \cdot (bA_3 - cA_2)^2 \\ &= -503 \cdot (2b^3 - bc^2 + b^2c + 2c^3)^2 \end{aligned}$$

where $2b^3 - bc^2 + b^2c + 2c^3 \equiv bc(b-c) \equiv 0 \pmod 2$. ∎

III. Prime Ideals in Number Rings

8 Dedekind Domains

Definition. R is **Noetherian** if every ideal of R is finitely generated; that is, $I = (r_1, \dots, r_n)$.

t:noe

8.1 Proposition. *The following are equivalent:*

(i) Every ascending chain of ideals in R stabilizes.

(ii) Every non-empty set S of ideals of R has a maximal element in S.

(iii) R is Noetherian.

Proof $(i \Rightarrow ii)$ Let S be a non-empty set of ideals with no maximal element. Since S is non-empty, get $I_1 \in S$. Then for any $I_k \in S$, I_k is not maximal and get $I_{k+1} \supsetneq I_k$. This is an infinite chain of ideal which does not stabilize.

$(ii \Rightarrow i)$ Let $I_1 \subseteq I_2 \subseteq \cdots$ be an ascending chain of ideals, and let $S = \{I_k : k \in \mathbb{N}\}$. By assumption, S has a maximal element, I_N; but then for any $n \geq N$, $I_n = I_N$ and the chain stabilizes.

$(i \Rightarrow iii)$ Let I be an ideal of R not finitely generated. Then $I \neq (0)$, so get $a_1 \in I$. For any finite $a_1, \dots, a_k \in I$, since I is not finitely generated, there exists $a_{k+1} \in I \setminus (a_1, \dots, a_k)$. But then $\{(a_1, \dots, a_i) : i \in \mathbb{N}\}$ does not stabilize, a contradiction.

$(iii \Rightarrow i)$ Let $I_1 \subseteq I_2 \subseteq \cdots$ be an ascending chain of ideals, and set $I = \bigcup_{i=1}^{\infty} I_n$. By assumption, $I = (x_1, \dots, x_n)$. Since each $x_i \in I_j$ for some j, get k so that $x_1, \dots, x_n \in I_k$; but then $I_k = I_n$ for all $n \geq k$ and the chain stabilizes. ∎

8.2 Theorem. (Hilbert) *If R is Noetherian, then $R[x]$ is Noetherian.*

Proof See PMATH 446 notes. ∎

Remark. The most basic example of a Noetherian domain is a PID. It is also easy to see that if R is Noetherian, R/I is also Noetherian. This means that a lot of rings are Noetherian.

Definition. If $R \subseteq S$ subrings with R, S integral domains, we say $s \in S$ is **integral over R** if there exists $f(x) \in R[x]$, f monic, such that $f(s) = 0$. We say R is **integrally closed in S** if $s \in S$ is integral over R if and only if $s \in R$.

Example.

1. Let K be a number field so that $\mathbb{Z} \subseteq K$. Then $\alpha \in K$ is integral over $\mathbb{Z}$ if and only if $\alpha \in \mathcal{O}_K$.
2. If $R = \mathbb{Z}$, $\mathrm{Frac}(R) = \mathbb{Q}$ and $\alpha \in \mathbb{Q}$ is integral over R if and only if $\alpha \in \mathbb{Z}$, so $\mathbb{Z}$ is integrally closed in $\mathbb{Q}$.
3. If $R = \mathbb{Z}[\sqrt{5}]$, then $\mathrm{Frac}(R) = \mathbb{Q}(\sqrt{5})$. Note that $(1 + \sqrt{5})/2$ is integral over $\mathbb{Z}[\sqrt{5}]$ (in fact, it is not even integral over $\mathbb{Z}$), so $\mathbb{Z}[\sqrt{5}]$ is not integrally closed in $\mathbb{Q}(\sqrt{5})$.

Definition. A **Dedekind domain** is an integral domain R satisfying 3 properties:

1. R is Noetherian.
2. Every prime ideal is maximal.
3. R is integrally closed in its field of fractions.

p:id-scale

8.3 Proposition. *Let K be a number field, $0 \neq I \subseteq \mathcal{O}_K$ an ideal. Then there exists $a \in \mathbb{Z}\setminus\{0\}$ such that $a \in I$.*

Proof Say $\alpha \in I$, $\alpha \neq 0$. Let $\alpha_1, \dots, \alpha_n$ be conjugates of $\alpha = \alpha_1$ so that $a := N_{\mathbb{Q}}^{\mathbb{Q}(\alpha)}(\alpha) = \alpha_1 \cdots \alpha_n \in \mathbb{Z}\setminus\{0\}$. As in the proof of (Prop. 5.4), $\alpha_2 \cdots \alpha_n \in \mathcal{O}_K$ so that $a \in I$. ∎

Remark. If $0 \neq I \subseteq \mathcal{O}_K$, this proposition show that $I \cap \mathbb{Z} \subseteq \mathbb{Z}$ is a non-zero ideal.

Definition. Given $I \subseteq \mathcal{O}_K$ an ideal, then $\{\alpha_1, \dots, \alpha_n\}$ is called an **integral basis** of I if $\alpha_i \in I$ and every element of I has a unique representation as an integer linear combination of the α_i.

t:id-basis

8.4 Theorem. *Every non-zero ideal $I \subseteq \mathcal{O}_K$ has an integral basis. More specifically, if $\{\omega_1, \dots, \omega_n\}$ is an integral basis for $\mathcal{O}_K$, then there exists $a_{ij} \in \mathbb{Z}$, $a_{ii} \in \mathbb{Z}^+$ such that $\{\alpha_1, \dots, \alpha_n\}$ is an integral basis for I and*

$$\begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix} = \begin{pmatrix} a_{11} & 0 & \dots & 0 \\ a_{21} & a_{22} & \ddots & \vdots \\ \vdots & \vdots & \ddots & 0 \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix} \begin{pmatrix} \omega_1 \\ \vdots \\ \omega_n \end{pmatrix}$$

Proof From (Prop. 8.3) there exists $a \in I \cap \mathbb{Z}^+$; in particular, for any $\omega \in \mathcal{O}_K$, $a\omega \in I$. We shall use this fact throughout the proof. We thus inductively define a_{ij} as follows:

- Let $a_{11} \in \mathbb{Z}^+$ be minimal such that $a_{11}\omega_1 \in I$; set $\alpha_1 := a_{11}\omega_1$.
- Let $a_{21} \in \mathbb{Z}$ and $a_{22} \in \mathbb{Z}^+$ minimal such that $\alpha_2 := a_{21}\omega_1 + a_{22}\omega_2 \in I$. Again, such an α_2 exists since $a(\omega_1 + \omega_2) \in I$.
- In general, let $\alpha_i := a_{i1}\omega_1 + a_{i(i-1)}\omega_{i-1} + \cdots + a_{ii}\omega_i \in I$ with $a_{ii} \in \mathbb{Z}^+$ minimal.

Let $A = (a_{ij})$ which satisfies the requirements; it remains to show that $\{\alpha_1, \dots, \alpha_n\}$ is in fact an integral basis. Since $\{\omega_1, \dots, \omega_n\}$ is a basis for $K/\mathbb{Q}$ and $\det(A) \neq 0$, $\{\alpha_1, \dots, \alpha_n\}$ is as well.

Now, let $\beta \in I$ be arbitrary. Since $\{\omega_1, \dots, \omega_n\}$ is an integral basis for $\mathcal{O}_K$, get $b_i \in \mathbb{Z}$ such that $\beta = b_1\omega_1 + \cdots + b_n\omega_n$. Write $b_n = a_{nn}q + r$ with $0 \leq r < a_{nn}$ and $q, r \in \mathbb{Z}$. Then

$$b_1\omega_1 + \cdots + b_{n-1}\omega_{n-1} + r\omega_n = \beta - q\alpha_n \in I$$

so by minimaliy of a_{nn}, we must have $r = 0$. Thus

$$\beta = b_1\omega_1 + \cdots + b_{n-1}\omega_{n-1} + qa_{nn}\omega_n \qquad\qquad a_{nn}\omega_n = \alpha_n + \gamma$$

where $\gamma \in \operatorname{span}_{\mathbb{Z}}\{\omega_1, \dots, \omega_{n-1}\}$. Thus $b_1\omega_1 + \cdots + b_{n-1}\omega_{n-1} + q\gamma = \beta - q\alpha_n \in I$.

The proof follows by repeating the same argument with $\beta - q\alpha_n$. ∎

Example. Consider $I = (7) \subseteq \mathbb{Z}[\sqrt{2}]$. An integral basis for (7) is $\{7, 7\sqrt{2}\}$ since $7, 7\sqrt{2} \in I$ and every element of I is of the form $7(a + b\sqrt{2})$ for some $a, b \in \mathbb{Z}$.

8.5 Theorem. *If K is a number field, then $\mathcal{O}_K$ is a Dedekind domain.*

Proof We verify the three requirements:

1. *$\mathcal{O}_K$ is Noetherian.* Suppose $I \subseteq \mathcal{O}_K$. If $I = (0)$ we're done; otherwise, choose an integral basis $\{\alpha_1, \dots, \alpha_n\}$ for I and $I = (\alpha_1, \dots, \alpha_n)$.

2. *Every non-zero prime ideal is maximal.* Let $0 \neq P \subseteq \mathcal{O}_K$ be prime. It suffices to show that $|\mathcal{O}_K/P| < \infty$ since finite integral domains are fields*.

   Let $\{\omega_1, \ldots, \omega_n\}$ be an integral basis for $\mathcal{O}_K$. Then by (Prop. 8.3), there exists $a \in \mathbb{Z}^+ \cap P$ so that $a\omega_i \in P$. Thus there are at most a^n possible elements in $\mathcal{O}_K/P$.

3. *$\mathcal{O}_K$ is integrally closed in K.* Suppose $\gamma \in K$ is integral over $\mathcal{O}_K$, so that $\gamma^n + \alpha_{n-1}\gamma^{n-1} + \cdots + \alpha_1\gamma + \alpha_0 = 0$. Note that $\gamma \in \mathbb{Z}[\alpha_0, \alpha_1, \ldots, \alpha_{n-1}, \gamma] = A$; it suffices to show that A is finitely generated as an additive group by (Thm. 4.3). Since A is generated over $\mathbb{Z}$ by all $\alpha_0^{m_0} \cdots \alpha_{n-1}^{m_{n-1}} \gamma^{m_n}$, let's show that only finitely many such products are necessary. Since $\alpha_i \in \mathcal{O}_K$, we can take $m_i < [K : \mathbb{Q}]$; and, since γ^n is expressible as a product over the α_i from its minimal polynomial, we can take $m < n$.

Thus $\mathcal{O}_K$ is a Dedekind domain. ∎

9 Prime Factorization of Ideals

Unique Factorization

l:id-cont

9.1 Lemma. *Let Q be a prime ideal in a ring R such that $Q \supseteq J_1 \cdots J_r$. Then $Q \supseteq J_i$ for some i.*

Proof Suppose $J_1, \ldots, J_{r-1} \not\subset Q$. Thus get $j_i \in J_i$ for $i < r$ with $j_i \notin Q$. If $\alpha \in J_r$ arbitrary, then $j_1 \cdots j_{r-1}\alpha \in q$ so by primality, $\alpha \in Q$ and $J_r \subseteq \alpha$. ∎

l:prod-id

9.2 Lemma. *If R is a Dedekind domain, then every non-zero ideal contains a product of prime ideals.*

Proof Let S be the set of non-zero ideals that don't contain a product of primes; suppose $S \neq \emptyset$. Since R is Noetherian, by (Prop. 8.1), there exists $M \in S$ maximal. Since M is not prime, get $r, s \in R \setminus M$ with $rs \in M$. But then $M_1 := M + (r)$ and $M_2 := M + (s)$ properly contain M and are not in S, so M_1 and M_2 both contains products of primes. Furthermore, $M_1 M_2 \subseteq M$ so M contains a product of prime ideals, forcing $S = \emptyset$. ∎

l:id-frac

9.3 Lemma. *Let R be a Dedekind domain, $I \subsetneq R$, and $K = \mathrm{Frac}(R)$. Then there exists $\gamma \in K \setminus R$ such that $\gamma I \subseteq \mathbb{R}$.*

Proof This is obvious if $I = (0)$, so we may assume $I \neq (0)$ and let $0 \neq a \in I$. Since $I \subsetneq R$, a is not a unit, so $\frac{1}{a} \in K \setminus R$. By (Lem. 9.2), $(a) \supseteq P_1 \cdots P_r$ where P_i are prime ideals; we may take r to be minimal. Let M be a maximal ideal containing I, so $M \supseteq I \supseteq (a) \supseteq P_1 \cdots P_r$, and since M is maximal, M is prime.

Without loss of generality, $P_1 \subseteq M$ by (Lem. 9.1); since R is Dedekind, $M = P_1$. If $r = 1$, set $b = 1$; and if $r > 1$, let $b \in P_2 \cdots P_r$. Set $\gamma = \frac{b}{a}$ so that

$$\gamma I = \frac{b}{a} I \subseteq \frac{b}{a} P_1 \subseteq \frac{1}{a} P_1 \cdots P_r \subseteq \frac{1}{a}(a) = R$$ ∎

*Let R be an integral domain and consider $\alpha : R \to R$ by $\alpha(x) = xr$. This is injective since R is an integral domain and, since R is finite, it is surjective as well. Thus there exists $s \in R$ with $rs = 1$.

p:id-inv

9.4 Proposition. *Suppose R is a Dedekind domain and $(0) \neq I \subseteq R$ is an ideal. Then for any $0 \neq \alpha \in I$, there exists an ideal $J \subseteq R$ such that $IJ = (\alpha)$ is principal.*

Proof Set $J = \{\beta \in R : \beta I \subseteq (\alpha)\}$[†]. Certainly $IJ \subseteq (\alpha)$ by definition, so we need to show that $(\alpha) \subseteq IJ$.

Let $B = \frac{1}{\alpha}IJ$. We know $B \subseteq R$ is an ideal, so we want to show that $B = R$. Suppose not. Then by (Lem. 9.3), get $\gamma \in \mathrm{Frac}(R) \setminus R$ such that $\gamma B \subseteq R$. Since $\alpha \in I$, $J \subseteq B$ and $\gamma J \subseteq \gamma B \subseteq R$. Then, $\gamma \frac{1}{\alpha} IJ = \gamma B \subseteq R$ which can be rephrased as $(\gamma J)I \subseteq (\alpha)$; thus, $\gamma J \subseteq J$ by definition of J.

Now since J has an integral basis, it is a finitely generated additive group. But then since $\gamma \in K \setminus R$, we cannot have $J \supseteq \gamma J \supseteq \gamma J^2 \supseteq \cdots$. Thus in fact $B = R$ and the result follows. ∎

Definition. If A, B are ideals in R, we say that ***A* divides *B*** and write $A \mid B$ if there exists an ideal C such that $AC = B$.

cor:id-div

9.5 Corollary. *Let A, B, C be ideals in a Dedekind domain, with $C \neq 0$. Then*

(i) $A \supseteq B$ if and only if $A \mid B$.

(ii) If $AC = BC$, then $A = B$.

Proof (i) If $A|B$, then get C such that $B = AC \subseteq A$. Conversely, suppose $A \supseteq B$. This is clear if $A = (0)$; else, let $0 \neq \alpha \in A$. Then by (Prop. 9.4), get J such that $JA = (\alpha)$. Then $(\alpha) = JA \supseteq JB$, so $R \supseteq \frac{1}{\alpha}JB$. Let $C = \frac{1}{\alpha}JB$ so that $AC = \frac{1}{\alpha}AJB = B$.

(ii) By (Prop. 9.4), get I such that $CI = (\alpha)$. Then $(\alpha)A = ACI = BCI = (\alpha)B$, so $A = B$. ∎

9.6 Theorem. *In a Dedekind domain, every proper non-zero ideal factors uniquely into a product of prime ideals.*

Proof Let S be the set of non-zero proper ideals that cannot be written as a product of primes. If $S \neq \emptyset$, let $M \in S$ be a maximal. Since M is not prime, so M is not maximal: thus, let $P \supsetneq M$ be maximal. Since $M \subseteq P$, $P|M$ so there is an ideal C such that $M = PC$; since $M \neq \rho$, $C \neq R$. C cannot be a product of prime ideals, or M is a product of prime ideals, so $C \in S$. But then by maximality of M, $M = C$ and $M = PC = PM$, so $P = R$, a contradiction.

It remains to show unique factorization. Given $I \neq 0$, R, say $I = P_1 \cdots P_r = Q_1 \cdots Q_s$. Then $I \subseteq P_1$, so without loss of generality, $P_1 \supseteq Q_1$ by (Lem. 9.1). Since prime ideals are maximal, this forces $P_1 = Q_1$ so we can cancel to get $P_2 \cdots P_r = Q_2 \cdots Q_s$ and we are done by induction. ∎

Example. Consider the ring $\mathbb{Z}[\sqrt{-5}]$, which does not have unique factorization: $6 = 2 \cdot 4 = (1+\sqrt{-5})(1-\sqrt{-5})$. In particular, this means that (2) is not a prime ideal: $1+\sqrt{-5}, 1-\sqrt{-5} \notin (2)$, but $6 \in (2)$. We do, however, still have unique factorization:

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

[†]From commutative algebra, this is the **ideal quotient**, denoted $((\alpha) : I)$.

9.7 Theorem. *R is a PID if and only if R is a UFD and a Dedekind domain.*

Proof ($\Rightarrow$) Since every PID is a UFD; it suffices to show that R is also Dedekind. This is left as an assignment exercise.

($\Leftarrow$) Let R be a UFD and a Dedekind domain. By unique factorization, it suffices to show that every non-zero prime ideal P is principal. Let $0 \neq \alpha \in P$; since P is proper, α is not a unit. Since R is a UFD, write $\alpha = p_1^{a_1} \cdots p_k^{a_k}$ with $k > 0$. Since P is prime, without loss of generality, $p_1 \in P$. Then $(p_1) \subseteq P$ but (p_1) is prime and thus maximal (since R is Dedekind), so $P = (p_1)$ is principal. ∎

9.8 Corollary. *If K is a number field, then $\mathcal{O}_K$ has unique factorization into primes if and only if $\mathcal{O}_K$ is a PID.*

Proof This follows immediately since $\mathcal{O}_K$ is Dedekind. ∎

Example. Consider $K = \mathbb{Q}(\sqrt{-d})$, for $d > 0$. For $d = 1, 2$, K is in fact a Euclidean domain; more generally, we may ask when $\mathcal{O}_K$ is a PID. It was conjectured (correctly) by Gauss that this is true when $d \in \{1, 2, 3, 7, 11, 19, 43, 67, 163\}$.

Ramification

Suppose $P \subseteq \mathcal{O}_K$ is a prime ideal and let $0 \neq a \in P \cap \mathbb{Z}^+$. Write $a = p_1^{e_1} \cdots p_r^{e_r}$, so by primality, some $p := p_i \in P$. Then $(p) \subseteq P$, so $P \mid (p)$; thus, $PQ_1 \cdots Q_s = (p)$ for some prime ideals Q_i. In particular, every prime ideal of $\mathcal{O}_K$ is a factor of (p) for some $p \in \mathbb{Z}$.

Furthermore, P cannot be a factor of (p) and (q) for distinct primes $p, q \in \mathbb{Z}$: otherwise, $p, q \in P$ so $1 \in \rho$. Thus $p \in \mathbb{Z}^+$ is the unique prime number such that $(p) \subseteq P$; in this case, we say that ***P* lies over *p***.

Definition. Let K be a number field, $p \in \mathbb{Z}^+$ a prime. We say ***p* ramifies in *K*** if there exists some prime ideal $P \subseteq \mathcal{O}_K$ such that $P^2 \mid (p)$ in $\mathcal{O}_K$.

Remark. By unique factorization into prime ideals, we can write $(p) = P_1^{e_1} \cdots P_r^{e_r}$. Then p ramifies in K if $e_i > 1$ for some i. We say that e_i is the **ramification index** for the prime P_i.

We can interpret the idea of ramification in the sense of algebraic geometry. First, let's consider a well-known example: consider the map $f : \mathbb{C} \to \mathbb{C}$ be given by $x \mapsto x^n$. Then most points $z \in \mathbb{C}$ have n distinct preimages - in fact, all of them except $z = 0$.

In algebraic geometry, since $\mathcal{O}_K/\mathbb{Z}$ is an integral extension of rings, we consider $\pi : \mathrm{Spec}(\mathcal{O}_K) \to \mathrm{Spec}(\mathbb{Z})$ given by $\rho \mapsto \rho \cap \mathbb{Z}$. This is a surjective homeomorphism. In general, if $R \subseteq S$ is integral and R is Noetherian, then for any prime $P \subseteq R$, the set of primes $Q \subseteq S$ containing P are precisely those given by $\pi^{-1}(P)$. In the number field case, if K/L is a finite field extension so that $\mathcal{O}_K \subseteq \mathcal{O}_L$; then $Q \subseteq \mathcal{O}_L$ lies over $P \subseteq \mathcal{O}_K$ precisely when Q occurs in the prime factorization of $P\mathcal{O}_L$. If we take $K = \mathbb{Q}$, then $\mathcal{O}_K = \mathbb{Z}$ and we consider prime ideals $(p) \subseteq \mathbb{Z}$. As we will see shortly, for most points, $\pi^{-1}((p))$ consists of distinct prime factors; the question of checking if p is ramified is searching for points where this is not true.

t:id-ram

9.9 Theorem. *Let $D = \mathrm{disc}(K)$ and $p \in \mathbb{Z}^+$. Then p is ramified if and only if $p \mid D$.*

Proof ($\Rightarrow$) Get P such that $P^2 \mid (p)$; let $(p) = P^2 Q$. Note that $PQ \neq P^2 Q$ since P is proper; thus, let $\alpha \in PQ \setminus P^2 Q$. In particular, $\frac{\alpha}{p} \notin \mathcal{O}_K$ but $\alpha^2 \in P^2 Q^2 \subseteq (p)$ so $\frac{\alpha^2}{p} \in \mathcal{O}_K$. Thus

if $\beta \in \mathcal{O}_K$ is arbitrary, then $\frac{(\alpha\beta)^p}{p} \in \mathcal{O}_K$. Note that $\operatorname{Tr}((\alpha\beta)^p) = p \operatorname{Tr}\left(\frac{(\alpha\beta)^p}{p}\right)$, so $p \mid \operatorname{Tr}((\alpha\beta)^p)$ and

$$\operatorname{Tr}((\alpha\beta)^p) = \left(\sum_{i=1}^{n} \sigma_i(\alpha\beta)\right)^p = \sum_{i=1}^{n} \sigma_i(\alpha\beta)^p + p\gamma = \operatorname{Tr}((\alpha\beta)^p) + p\gamma$$

for some $\gamma \in \mathcal{O}_K$. Thus $p \mid \operatorname{Tr}(\alpha\beta)^p$, so $p \mid \operatorname{Tr}(\alpha\beta)$.

Let $\{\omega_1, \ldots, \omega_n\}$ be an integral basis for K so that $\alpha = a_1\omega_1 + \cdots + a_n\omega_n$ for $a_i \in \mathbb{Z}$. Since $\frac{\alpha}{p} \notin \mathcal{O}_K$, without loss of generality, $p \nmid a_1$. Note that $p \mid \operatorname{Tr}(\alpha\omega_i)$ for any i and

$$\operatorname{Tr}(\alpha\omega_i) = \operatorname{Tr}((a_1\omega_1 + \cdots + a_n\omega_n)\omega_i) = \sum_{j=1}^{n} a_j \operatorname{Tr}(\omega_j\omega_i)$$

Now,

$$D = \operatorname{disc}(K) = \det \begin{pmatrix} \operatorname{Tr}(\omega_1\omega_1) & \cdots & \operatorname{Tr}(\omega_1\omega_n) \\ \vdots & \ddots & \vdots \\ \operatorname{Tr}(\omega_n\omega_1) & \cdots & \operatorname{Tr}(\omega_n\omega_n) \end{pmatrix}$$

so that by standard matrix manipulation preserving the determinant,

$$a_1 D = \det \begin{pmatrix} a_1 \operatorname{Tr}(\omega_1\omega_1) & \cdots & a_1 \operatorname{Tr}(\omega_1\omega_n) \\ \operatorname{Tr}(\omega_2\omega_1) & \cdots & \operatorname{Tr}(\omega_2\omega_n) \\ \vdots & \ddots & \vdots \\ \operatorname{Tr}(\omega_n\omega_1) & \cdots & \operatorname{Tr}(\omega_n\omega_n) \end{pmatrix} = \det \begin{pmatrix} \sum_{i=1}^{n} a_i \operatorname{Tr}(\omega_i\omega_1) & \cdots & \sum_{i=1}^{n} \operatorname{Tr}(\omega_i\omega_n) \\ \operatorname{Tr}(\omega_2\omega_1) & \cdots & \operatorname{Tr}(\omega_2\omega_n) \\ \vdots & \ddots & \vdots \\ \operatorname{Tr}(\omega_n\omega_1) & \cdots & \operatorname{Tr}(\omega_n\omega_n) \end{pmatrix}$$

and $p \mid a_1 D$. Since $p \nmid a_i$, $p \mid D$.

($\Leftarrow$) This implication is beyond the scope of this course. ∎

Example. Consider $\mathbb{Z}[\sqrt{3}] = \mathcal{O}_K$. Since $\operatorname{disc}(K) = 12$, we see that $(3) = (\sqrt{3})^2$ indeed ramifies.

10 Norms of Ideals

Definition. Let $(0) \neq I \subseteq \mathcal{O}_K$. Then we define **norm** of I to be $|\mathcal{O}_K/I|$, and write $N_{\mathbb{Q}}^K(I)$ or $N(I)$ or $\|I\|$ when the context is clear.

Remark. Equivalently, $N(I) = [\mathcal{O}_K : I]$ since $I \subseteq \mathcal{O}_K$ is an additive subgroup.

t:id-norm

10.1 Theorem. *Let K be a number field, $I \subseteq \mathcal{O}_K$ and let $\{\alpha_1, \ldots, \alpha_n\}$ be an integral basis for I. Then*

$$N(I) = \left|\frac{\operatorname{disc}(\alpha_1, \ldots, \alpha_n)}{\operatorname{disc}(K)}\right|^{1/2}$$

Proof Let's first show that this quantity does not depend on the choice of integral basis. Suppose $\{\alpha_1, \ldots, \alpha_n\}$ and $\{\beta_1, \ldots, \beta_n\}$ are choices of integral bases for I. Then if P is a change of basis, $P \in \mathrm{GL}_n(\mathbb{Z})$, so $\det(P) = \pm 1$.

It thus suffices to do this with the integral basis as in (Thm. 8.4). Fix $\{\omega_1,\dots,\omega_n\}$ an integral basis for $\mathcal{O}_K$, and let

$$A = \begin{pmatrix} a_{11} & 0 & \dots & 0 \\ a_{21} & a_{22} & \ddots & \vdots \\ \vdots & \vdots & \ddots & 0 \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix} \qquad \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix} = A \begin{pmatrix} \omega_1 \\ \vdots \\ \omega_n \end{pmatrix}$$

so that

$$\operatorname{disc}(\alpha_1,\dots,\alpha_n) = \det(A)^2 \operatorname{disc}(\omega_1,\dots,\omega_n)$$

Thus,

$$\frac{\operatorname{disc}(\alpha_1,\dots,\alpha_n)}{\operatorname{disc}(K)} = (a_{11}\cdots a_{nn})^2$$

Thus we need to show that $N(I) = \prod_{i=1}^n a_{ii}$. Let's show that every element of $\mathcal{O}_K/I$ has a unique representation as $r_1\omega_1 + \cdots + r_n\omega_n$ where $0 \le r_i < a_{ii}$.

First let $\gamma \in \mathcal{O}_K$ be arbitrary and write $\gamma = \sum_{i=1}^n b_i\omega_i$.

- Write $b_n = q_n a_{nn} + r_n$, where $0 \le r_n < a_{nn}$ and set $\gamma_n = \gamma - q_n\alpha_n$.
- Let c_{n-1} denote the coefficient of ω_{n-1} in γ_n, and write $c_{n-1} = q_{n-1}a_{n-1,n-1} + r_{n-1}$, and set $\gamma_{n-1} = \gamma_n - q_{n-1}\alpha_{n-1}$.

Repeating the above process, we note that

$$r_1\omega_1 + \cdots + r_n\omega_n = \gamma - (q_1\alpha_1 + \cdots + q_n\alpha_n)$$

and since $\sum_{i=1}^n q_i\alpha_i \in I$, γ has a representation in the desired form.

Furthermore, such a representation is unique: suppose

$$r_1\omega_1 + \cdots + r_n\omega_n \equiv s_1\omega_1 + \cdots + s_n\omega_n \pmod{I}$$

where $r_i, s_i < a_{ii}$. Then since $\{\alpha_1,\dots,\alpha_n\}$ is an integral basis for I, we have

$$\sum_{i=1}^n (r_i - s_i)\omega_i = \sum_{i=1}^n t_i\alpha_i = \begin{pmatrix} t_1a_{11} & 0 & \dots & 0 \\ t_2a_{21} & t_2a_{22} & \ddots & \vdots \\ \vdots & \vdots & \ddots & 0 \\ t_na_{n1} & t_na_{n2} & \dots & t_na_{nn} \end{pmatrix} \cdot \begin{pmatrix} \omega_1 \\ \vdots \\ \omega_n \end{pmatrix}$$

But then comparing coefficients starting at the last row, since $|r_i - s_i| < a_{ii}$, we have $t_n = \cdots = t_1 = 0$ so that $r_i = s_i$ for all i. ∎

10.2 Corollary. *Let K be a number field, $I = (\alpha)$. Then $N(I) = |N(\alpha)|$.*

Proof Let $\{\omega_1,\dots,\omega_n\}$ be a basis for $\mathcal{O}_K$ so that $\{\alpha\omega_1,\dots,\alpha\omega_n\}$ is a basis for I. Let $\sigma_1,\dots,\sigma_n$ be the embeddings of $K \hookrightarrow \mathbb{C}$. Then

$$\begin{pmatrix} \sigma_1(\alpha\omega_1) & \dots & \sigma_1(\alpha\omega_n) \\ \vdots & \ddots & \vdots \\ \sigma_n(\alpha\omega_1) & \dots & \sigma_n(\alpha\omega_n) \end{pmatrix} = \begin{pmatrix} \sigma_1(\alpha) & & \\ & \ddots & \\ & & \sigma_n(\alpha) \end{pmatrix} \begin{pmatrix} \sigma_1(\omega_1) & \dots & \sigma_1(\omega_n) \\ \vdots & \ddots & \vdots \\ \sigma_n(\omega_1) & \dots & \sigma_n(\omega_n) \end{pmatrix}$$

so that $\operatorname{disc}(\alpha\omega_1,\dots,\alpha\omega_n) = N(\alpha)^2 \operatorname{disc}(\omega_1,\dots,\omega_n)$. Then $N(I) = |N(\alpha)|$ by (Thm. 10.1). ∎

Remark. In the sense of the previous two propositions, we see that $N(I)$ is a generalization of norms of elements.

10.3 Theorem. (Fermat) *Let K be a number field, $P \subseteq \mathcal{O}_K$ a prime ideal, $\alpha \in \mathcal{O}_K$, and $P \nmid (\alpha)$. Then $\alpha^{N(p)-1} \equiv 1 \pmod{P}$.*

Proof Since $\mathcal{O}_K/P$ is a field, so $(\mathcal{O}_K/P)^\times$ is a group with size $N(P)-1$. Then $P \nmid (\alpha)$ if and only if $\alpha \notin \rho$ and the result follows by Lagrange's theorem. ∎

p:norm-in-id

10.4 Proposition. *If $I \subseteq \mathcal{O}_K$ is an ideal, then $N(I) \in I$.*

Proof Consider the element $1 + I \in \mathcal{O}_K/I$. Since $|\mathcal{O}_K/I| = N(I)$, by Lagrange, $N(I)(1 + I) = N(I) + I = 0$, so $N(I) \in I$. ∎

cor:fin-norm

10.5 Corollary. *If K is a number field and $a \in \mathbb{Z}^+$, then there are only finitely many ideals $I \subseteq \mathcal{O}_K$ with $N(I) = a$.*

Proof If $I \subseteq \mathcal{O}_K$ is an ideal and $N(I) = a$, then $a \in I$. Thus $(a) \subseteq I$ so $I \mid (a)$; since we have unique factorization of ideals, there are only finitely many such I. ∎

Example. Which $I \subseteq \mathbb{Z}[i]$ have norm 5? Note that $5 = (1+2i)(1-2i)$ is a factorization into primes, since if $N(I) = 5$, then $I = (1+2i)^a(1-2i)^b$ for $a, b \in \{0, 1\}$. Thus one can verify $N(I) = 5$ if and only if $I = (1+2i)$ or $I = (1-2i)$.

The Ideal Norm is Multiplicative

Definition. If $B, C \subseteq \mathcal{O}_K$ are ideals, we say $D \subseteq \mathcal{O}_K$ is the **gcd of B, C** if $D \mid B$, $D \mid C$ and, whenever $E|B$ and $E|C$, we have $E \mid D$. We also say $L = \subseteq \mathcal{O}_K$ is the **lcm of B, C** if $B \mid L$, $C \mid L$, and L is minimal with this property.

Remark. One can see that the gcd and lcm both exist and are unique by factorization of ideals into primes. Alternatively, in light of (Cor. 9.5), $\gcd(I, J) = I + J$ is the smallest ideal containing both I and J, and $\operatorname{lcm}(I, J) = I \cap J$ is the smallest ideal contained in I and J.

l:gcd-pr

10.6 Lemma. *Suppose $B, C \subseteq \mathcal{O}_K$ are non-zero ideals. Then there exists $\alpha \in B$ such that $\gcd\left(\frac{(\alpha)}{B}, C\right) = 1$.*

Remark. This makes sense since if $\alpha \in B$, then $(\alpha) \subseteq B$ and $B \mid (\alpha)$.

Proof If $C = \mathcal{O}_K$, choose any $\alpha \in B$. Otherwise, write $C = \prod_{i=1}^r P_i^{e_i}$ as a product of prime ideals. If $r = 1$, then $C = P^e$. Let $\alpha \in B \setminus BP$ be arbitrary, so that

$$\gcd\left(\frac{(\alpha)}{B}, P^e\right) = P^m$$

Suppose for contradiction $m \geq 1$; then, $P \,\Big|\, \frac{(\alpha)}{B}$ so $(\alpha) = B\frac{(\alpha)}{B} = BPE \subseteq BP$ so that $\alpha \in BP$, a contradiction.

Now suppose $r > 1$; then by the $r = 1$ case for any $m \in \{1, \ldots, r\}$, set

$$B_m := B\frac{P_1 \cdots P_r}{P_m} \qquad \alpha_m \in B_m \quad \text{s.t.} \quad \gcd\left(\frac{(\alpha_m)}{B_m}, P_m\right) = 1 \qquad \alpha = \sum_{i=1}^r \alpha_i$$

In particular, since $B \supseteq B_i$, $\alpha_i \in B$ for any i so that $\alpha \in B$. Furthermore, for $i \neq m$, $\alpha_i \in B_i \subseteq BP_m$.

Let's show that $\alpha \notin BP_m$ for any m. If $\alpha \in BP_m$, then since $\alpha_i \in BP_m$ for $i \neq m$, we have $\alpha_m \in BP_m$. Thus $BP_m \mid (\alpha_m)$ and $B_m \mid (\alpha_m)$ by assumption so that $P_m \mid \frac{(\alpha_m)}{B}$ and $\frac{P_1 \cdots P_r}{P_m} \mid \frac{(\alpha_m)}{B}$. Thus $BP_1 \cdots P_r \mid (\alpha_m)$. But then since $\frac{BP_1 \cdots P_r}{B_m} = P_m$, $P_m \mid \frac{(\alpha)}{B_m}$, contradicting the choice of α_m.

Now suppose $\gcd\left(\frac{(\alpha)}{B}, C\right) \neq 1$. Then there exists m such that $P_m \mid \frac{(\alpha)}{B}$, so $BP_m \mid (\alpha)$ and $\alpha \in BP_m$, a contradiction. Thus the result follows. ∎

1:gcd-cp

10.7 Lemma. *Suppose $B, C \subseteq \mathcal{O}_K$ are non-zero ideals. If $\alpha\beta \equiv 0 \pmod{BC}$ and $\gcd\left(\frac{(\alpha)}{B}, C\right) = 1$, then $\beta \equiv 0 \pmod{C}$.*

Proof Since $\alpha\beta \in BC$, $BC \mid (\alpha)(\beta)$ so that $C \mid \frac{(\alpha)}{B}(\beta)$. Since $\gcd\left(\frac{(\alpha)}{B}, C\right) = 1$, we have $C \mid (\beta)$ so $\beta \in C$. ∎

10.8 Theorem. *If $B, C \subseteq \mathcal{O}_K$ are non-zero ideals, then $N(BC) = N(B)N(C)$.*

Proof By (Lem. 10.6), get $\alpha \in B$ such that $\gcd\left(\frac{(\alpha)}{B}, C\right) = 1$. Let $\beta_1, \ldots, \beta_{N(B)} \in \mathcal{O}_K$ and $\gamma_1, \ldots, \gamma_{N(C)} \in \mathcal{O}_K$ represent the distinct classes in $\mathcal{O}_K/B$ and $\mathcal{O}_K/C$ respectively. Let's show that $\beta_i + \alpha\gamma_j$ represent the distinct classes in $\mathcal{O}_K/BC$, which would give the result $N(BC) = N(B)N(C)$.

Let's first see that $\beta_i + \alpha\gamma_j$ are distinct mod BC and suppose

$$\beta_i + \alpha\gamma_j \equiv \beta_k + \alpha\gamma_l \pmod{BC} \iff \beta_i - \beta_k \equiv \alpha(\gamma_l - \gamma_i) \pmod{BC}$$

Thus since $BC \subseteq B$, the congruence also holds mod B. Then since $\alpha \in B$, $\beta_i - \beta_k \equiv 0 \pmod{B}$ and $i = k$. As a result, $0 \equiv \alpha(\gamma_l - \gamma_j) \pmod{BC}$, so by (Lem. 10.7), $\gamma_l - \gamma_j \equiv 0 \pmod{C}$ and $j = l$.

Next, we need to show if $\omega \in \mathcal{O}_K$, then there exists i, j such that $\omega \equiv \beta_i + \alpha\gamma_j \pmod{BC}$. Let i be such that $\omega \equiv \beta_i \pmod{B}$. Then $\omega - \beta_i \in B = \gcd((\alpha), BC) = (\alpha) + BC$ so that $\omega - \beta_i = \alpha a + b$ for some $a \in \mathcal{O}_K$ and $b \in BC$. Let j be such that $a \equiv \gamma_j \pmod{C}$. Then

$$\omega = \beta_i + \alpha\gamma_j + \alpha(a - \gamma_j) + b$$

with $\alpha \in B$, $a - \gamma_j \in C$. Then $\alpha(a - \gamma_j) \in BC$ and $b \in BC$, so $\omega = \beta_i + \alpha\gamma_j \pmod{BC}$. ∎

Remark. Note that $N(I) = 1$ if and only if $I = \mathcal{O}_K$. If $[K : \mathbb{Q}] = n$ and P is a prime ideal, we already showed $P \mid (p)$ for some prime $p \in \mathbb{Z}$. Factor $(p) = PJ$ and since the norm is multiplicative, $N((p)) = N(P)N(J)$ so $N(P) = p^f$ for some $1 \leq f \leq n$. We write $f(P|p) := f$. If $N(I)$ is prime, then I is a prime ideal (though the converse is not true!)

11 Class Group

Fix $\mathcal{O}_K$ and let $I, J \subseteq \mathcal{O}_K$ be ideals. Define an equivalence relation $\sim$ on the set of ideals of $\mathcal{O}_K$ by $I \sim J$ if there exists $\alpha, \beta \in \mathcal{O}_K$ such that $(\alpha)I = (\beta)J$. One can verify that this is indeed an equivalence relation.

Definition. The **class group of K** is $\mathrm{Cl}(\mathcal{O}_K)$ is the set of ideals modulo the above equivalence relation, where the group operation is multiplication of ideals. We define the **class number** $h_K := |\mathrm{Cl}(\mathcal{O}_K)|$.

t:ck

11.1 Theorem. *If K is a number field, then there exists a constant C_K such that for all $0 \neq A \subseteq \mathcal{O}_K$ ideal, there exists $\alpha \in A$ such that $|N(\alpha)| \leq C_K N(A)$.*

Proof Let $\{\omega_1, \dots, \omega_n\}$ be an integral basis for $\mathcal{O}_K$ and let $t := \lfloor N(A)^{1/m} \rfloor$. Consider all elements of $\mathcal{O}_K$ of the form $c_1\omega_1 + \cdots + c_n\omega_n$ where $0 \leq c_i \leq t$. There are $(t+1)^n$ such elements where $(t+1)^n > N(A)$; thus, by the pidgeonhole principle, there exists some $\beta_1 \neq \beta_2$ so that $\beta_1 \equiv \beta_2 \pmod{A}$. Set $\alpha = \beta_1 - \beta_1 \in A$; then, $\alpha = t_1\omega_1 + \cdots + t_n\omega_n$ with $|t_i| \leq t$. Then

$$\begin{aligned} |N(\alpha)| &= \left| \prod_{j=1}^{n} \sigma_j(\alpha) \right| = \left| \prod_{j=1}^{n} t_1\sigma_j(\omega_1) + \cdots + t_n\sigma_j(\omega_n) \right| \\ &\leq \prod_{j=1}^{n} \left(|t_1| \cdot |\sigma_j(\omega_1)| + \cdots + |t_n| \cdot |\sigma_j(\omega_n)| \right) \\ &\leq t^n \prod_{j=1}^{n} (|\sigma_j(\omega_1)| + \cdots + |\sigma_j(\omega_n)|) < N(A)C_K \end{aligned}$$

where $C_K = \prod_{j=1}^{n} \left(\sum_{j=1}^{n} |\sigma_j(\omega_i)| \right)$ depends only on K. ∎

Remark. This bound is not very good; we will later show that much better bounds are indeed possible.

t:cl-bound

11.2 Theorem. *If K is a number field, then the class number $h_K < \infty$.*

Proof By (Cor. 10.5), it suffices to show that that every ideal class contains an ideal with norm at most C_K.

Let $0 \neq I \subseteq \mathcal{O}_K$; then, get $0 \neq A$ such that IA is principal. By (Thm. 11.1), there exists $0 \neq \alpha \in A$ such that $|N((\alpha))| \leq C_K N(A)$. Since $\alpha \in A$, it follows that $(\alpha) \subseteq A$, so $(\alpha) = AB$ for some B. Thus, in the class group $\mathrm{Cl}(\mathcal{O}_K)$, we have $A = I^{-1}$ and $B = A^{-1}$, so $B = I$ in $\mathrm{Cl}(\mathcal{O}_K)$. Since $AB = (\alpha)$, $N(A)N(B) \leq C_K N(A)$, so $N(B) \leq C_K$. In other words, B and I are in the same class and $N(B) \leq C_K$. ∎

Computing an ideal class group

Last time, we showed every ideal class has a representation of norm at most C_K; this yields a general procedure for computing the class group, assuming we have a bound M of C_K.

1. Take a bound M for C_K; for example, we may take $M = \sqrt{|\mathrm{disc}(K)|}$.
2. From the proof of (Thm. 11.2), it suffices to consider ideals with norm at most M. Since $I = \prod P_i^{e_i}$ is a factorization, the primes P with $N(P) \leq M$ will generate $\mathrm{Cl}(\mathcal{O}_K)$.
3. Since $N(P) \in P$ by (Prop. 10.4), P lies over a prime p with $p \leq M$. Thus $P \subseteq (p)$, so every such prime will arise as a factor of (p) with $p \leq C_K$.

Example. Consider $\mathrm{Cl}(\mathbb{Q}(\sqrt{-23}))$, so $C_K = \sqrt{23} < 5$. Thus we need ideals with norm a most 4, so it suffices to consider $(2), (3)$. From a homework assignment, we know that

$$(2) = \underbrace{\left(2, \frac{1+\sqrt{-23}}{2}\right)}_{P}\underbrace{\left(2, \frac{1-\sqrt{-23}}{2}\right)}_{P'}$$
$$(3) = \underbrace{\left(3, \frac{1-\sqrt{-23}}{2}\right)}_{Q}\underbrace{\left(3, \frac{1+\sqrt{-23}}{2}\right)}_{Q'}$$

is a factorization into primes, so that all the ideals of norm at most 4 are products of the above primes. We will write $I \sim J$ if $I = J$ in $\mathrm{Cl}(K)$. Note that $PP' \sim (2)$ so that $P' = P^{-1}$; similarly, $Q' = Q^{-1}$. Furthermore, one can verify that

$$PQ = \left(6, 2\frac{1-\sqrt{-23}}{2}, 3\left(\frac{1-\sqrt{-23}}{2}\right), \left(\frac{1-\sqrt{-23}}{2}\right)^2\right) = (6)$$

so $P = Q^{-1}$ and $Q \sim P'$. An analoguous computation shows $P'Q' \sim (1)$ so $Q' \sim P$. We now have

$$N\left(\frac{3+\sqrt{-23}}{2}\right) = 8 = N\left(\frac{3-\sqrt{-23}}{2}\right)$$

and these two ideals are distinct. Since p is not principal b/c there is no principal ideal of norm 2. Similarly, p' is not principal. Since we know there are at last two distinct principal ideals of norm 8, we must have p, p' not principal. Thus p^3, p'^3 are not principal, so $p^3 \sim 1$ and p has order 3 in $\mathrm{Cl}(K)$.

IV. Topics in Number Theory

12 Quadratic Reciprocity

Suppose we wish to solve the equation $x^2 + bx + c$ in $\mathbb{F}_p$. After completing the square, this reduces to the question of solving $x^2 = a \pmod{p}$. Is there a nice way to determine when a is a square mod p?

Definition. We define the **Legendre symbol** by

$$\left(\frac{a}{p}\right) = \begin{cases} 1 & a \text{ is a square in } \mathbb{F}_p \\ -1 & : \text{otherwise} \end{cases}$$

Let H_p be the set of squares in $(\mathbb{Z}_p)^\times$. On a homework assignment, we showed that $[(\mathbb{Z}_p)^\times : H_p] = 2$; in particular, $(\mathbb{Z}_p)^\times \big/ H_p \cong \mathbb{Z}_2$. In particular, if a, b are not squares, then ab is a square; if a is not square and b is square, then ab is not square. This observation is the fact that

$$\left(\frac{ab}{p}\right) = \left(\frac{a}{p}\right)\left(\frac{b}{p}\right)$$

Said another way, the map $\phi : (\mathbb{Z}_p)^\times \to \mathbb{Z}_2$ with $a \mapsto \left(\frac{a}{p}\right)$ is a homomorphism with $\ker \phi = H_p$.

More generally, since $(\mathbb{Z}_p)^\times$ is cyclic, let $\alpha \in (\mathbb{Z}_p)^\times$ be a generator. Then the map $\theta(\mathbb{Z}_p)^\times \to \mathbb{Z}_{p-1}$ given by $\alpha^k \mapsto k$ is a group isomorphism. To summarize, the following diagram commutes

$$\begin{array}{ccc} (\mathbb{Z}_p)^\times & \xrightarrow{x \mapsto x^2} & H_p \\ \downarrow \theta & & \downarrow \theta|_{H_p} \\ \mathbb{Z}_{p-1} & \xrightarrow{x \mapsto 2x} & 2\mathbb{Z}_{p-1} \end{array}$$

Thus $2\mathbb{Z}_{p-1} = \ker(\psi)$ where $\psi : \mathbb{Z}_{p-1} \to \mathbb{Z}_2$ is the map $a \mapsto a \cdot (p-1)/2$; this map lifts to a map $\phi^* : H_p \to \{-1, 1\}$. In particular, taking $a = -1$, we have that

$$\left(\frac{-1}{p}\right) = \begin{cases} 1 & p \not\equiv 3 \pmod 4 \\ -1 & p \equiv 3 \pmod 4 \end{cases}$$

Let's record this as a proposition:

p:lgs-1

12.1 Proposition. *Let p be prime.*

(i) $\left(\frac{-1}{p}\right) = \begin{cases} 1 & p \not\equiv 3 \pmod 4 \\ -1 & p \equiv 3 \pmod 4 \end{cases}$

(ii) $\left(\frac{ab}{p}\right) = \left(\frac{a}{p}\right)\left(\frac{b}{p}\right)$

Let's view $(\mathbb{Z}_p)^\times \cong \mathrm{Gal}(\mathbb{Q}(\zeta_p)/\mathbb{Q})$. Let $p^* = (-1)^{\frac{p-1}{2}} p$, H_p denote the squares in $(\mathbb{Z}_p)^\times$, and $K_p = \mathbb{Q}(\sqrt{p^*})$. Then $\mathbb{Q}(\zeta_p)^{H_p} = K_p$ and $H_p = \mathrm{Gal}(\mathbb{Q}(\zeta_p)/K_p)$ by the fundamental theorem of Galois theory (Thm. 3.2). In particular,

$$\left(\frac{q}{p}\right) = 1 \iff q \in H_p \iff \sigma_q \text{ fixes } K_p \iff \mathcal{O}_K/P = \mathbb{F}_q \tag{12.1}$$ {eq:f}

Given $a \in (\mathbb{Z}_p)^\times$, let σ_a denote the Galois group element $\sigma_a : \zeta_p \mapsto \zeta_p^a$. Let Q denote any ideal of $\mathbb{Z}[\zeta_p]$ lying over q, so σ_q acts on $\mathbb{Z}[\zeta_p]\big/Q$ by

$$\sigma_q\left(\sum_{i=0}^{p-1} a_i \zeta_p^i\right) = \sum_{i=0}^{p-1} a_i \zeta_p^{qi} = \left(\sum_{i=0}^{p-1} a_i \zeta^i\right)^q$$

since $\mathbb{Z}[\zeta_p]\big/Q$ has characteristic q. Thus for all $\alpha \in \mathbb{Z}[\zeta_p]\big/Q$, $\sigma_q(\alpha) = \alpha^q$. We say that σ_q is the **Frobenius map** associated to q, and write $\sigma_q = \mathrm{Frob}_q \in \mathrm{Gal}(\mathbb{Q}(\zeta_p)/\mathbb{Q})$. The reasoning behind this name is that $\mathbb{Z}[\zeta_p]\big/Q \,/\, \mathbb{Z}/(p) = \mathbb{F}_p$, so $\mathbb{Z}[\zeta_p]\big/Q$ is a an extension of finite fields, and the map σ_q induces the Frobenius map on this extension.

In particular, since $K_p/\mathbb{Q}$ is also Galois, $\sigma_q = \mathrm{Frob}_q$ in $\mathrm{Gal}(K_p/\mathbb{Q})$ as well. One can show that Frob_q is the unique map up to conjugacy since the Galois group acts transitively on the primes lying over p.

Now let $P \subseteq \mathcal{O}_{K_p}$; note that $\sigma_q = \mathrm{id}|_{\mathrm{Gal}(K_p/\mathbb{Q})}$ if and only if $\mathcal{O}_K/P = \mathbb{F}_q$. From a homework assignment, this happens if and only if $(q) \subseteq K_p$ is not prime. Thus combining with (Eq. 12.1), we have that

$$\left(\frac{q}{p}\right) = 1 \iff (q) \text{ is not prime in } \mathbb{Q}(\sqrt{p^*}) \iff \left(\frac{p^*}{q}\right) = 1$$

Thus we have proven

12.2 Theorem. (Quadratic Reciprocity) *Let p, q be distinct primes. Then*

$$\left(\frac{q}{p}\right) = \left(\frac{p^*}{q}\right) \text{ or equivalently } \left(\frac{p^*}{q}\right) = (-1)^{\frac{p-1}{2}\cdot\frac{q-1}{2}}\left(\frac{p}{q}\right)$$

Example. A good way to illustrate the computational power of quadratic reciprocity is through an example:

$$\begin{aligned}\left(\frac{17}{113}\right) = \left(\frac{113}{17}\right) = \left(\frac{11}{17}\right) = \left(\frac{17}{11}\right)\\ = \left(\frac{6}{11}\right) = \left(\frac{2}{11}\right)\left(\frac{3}{11}\right)\\ = -\left(\frac{11}{3}\right) = -\left(\frac{-1}{3}\right) = 1\end{aligned}$$

Remark. Let $G = \mathrm{Gal}(\mathbb{Q}(\zeta_p)/\mathbb{Q}) \cong (\mathbb{Z}_p)^\times$. If q is prime, we say $\sigma \in G$ is Frob_q if, given Q lying over q, $\sigma(\alpha) = \alpha^q \pmod{Q}$. Which elements of G are of the form Frob_q for some q?

The perhaps surprising answer is that for every a, this happens infinitely often. In fact, this is a way of rephrasing Dirichlet's theorem for primes in arithmetic progressions.

This fact generalizes to all Galois groups, and is called the Chebotarev density theorem. One consequence: if K is a number field, then $\{p \in \mathbb{Z} : p \text{ splits completely in } \mathcal{O}_K\}$.

13 Fermat's Last Theorem

13.1 Theorem. (Fermat's Last) *If $p \geq 3$ is prime and $p \nmid x,y,z$ for $x,y,z, \in \mathbb{Z}\setminus\{0\}$, then $x^p + y^p \neq z^p$.*

Definition. We say that p is **regular** if $p \nmid h_{\mathbb{Q}(\zeta_p)}$.

The goal of this section is to prove Fermat's last theorem for regular primes. The key observation is that if p is a regular prime, if $I \subseteq \mathcal{O}_{\mathbb{Q}(\zeta_p)}$ is ideal with I^p is principal, then I is principal as well.

13.2 Theorem. (Hilbert Class Field) *Let K be a number field; then there exists a number field L such that L/K is normal and $\mathrm{Gal}(L/K) = \mathrm{Cl}(\mathcal{O}_K)$. Such an L is called the **Hilbert class field**.*

In particular, this has the property that for every ideal $I \subseteq \mathcal{O}_K$, $I\mathcal{O}_L$ is principal in $\mathcal{O}_L$.

Definition. We say p is a **regular prime** if $p \nmid h_{\mathbb{Q}(\zeta_p)}$.

1:km

13.3 Lemma. *Let $\zeta = \zeta_p$. Then in $\mathbb{Z}[\zeta]$,*

(i) the elements $1-\zeta, 1-\zeta^2, \dots, 1-\zeta^{p-1}$ are associates.
(ii) $1+\zeta$ is a unit
(iii) There exists $u \in \mathbb{Z}[\zeta]^\times$ so that $p = u(1-\zeta)^{p-1}$ so that $(1-\zeta)$ is the only prime ideal dividing (p).

Proof (i) Consider $\frac{1-\zeta^j}{1-\zeta} = 1+\zeta+\cdots+\zeta^{j-1} \in \mathbb{Z}[\zeta]$. As well, $\frac{1-\zeta}{1-\zeta^j} = \frac{1-\zeta^{jk}}{1-\zeta^j} \in \mathbb{Z}[\zeta]$ where $jk \equiv 1 \pmod{p}$. Thus $1-\zeta, \dots, 1-\zeta^p$ are associates.

(ii) We have $1+\zeta = \frac{1-\zeta^2}{1-\zeta}$, so $1+\zeta$ is a unit.

(iii) Recall that

$$1 + x + \cdots + x^{p-1} = \prod_{j=1}^{p-1} (x - \zeta^j)$$

so

$$p = \prod_{j=1}^{p-1} (1-\zeta^j) = (1-\zeta)^{p-1} \prod_{j=1}^{p-1} \frac{1-\zeta^j}{1-\zeta} = u(1-\zeta)^{p-1}$$

where $u = \prod_{j=1}^{p-1} \frac{1-\zeta^j}{1-\zeta} \in \mathbb{Z}[\zeta]^\times$ from (i). ∎

1:km2

13.4 Lemma. *If $u \in \mathbb{Z}[\zeta]^\times$, then $\frac{u}{\overline{u}}$ is a root of unity.*

Proof Let $\sigma \in \mathrm{Gal}(\mathbb{Q}(\zeta)/\mathbb{Q})$. Then $\sigma(\zeta) = \zeta^a$ for some a, so $\overline{\sigma(\zeta)} = \zeta^{-a} = \sigma(\overline{\zeta})$. Thus for any such σ,

$$\left|\sigma\left(\frac{u}{\overline{u}}\right)\right|^2 = \sigma\left(\frac{u}{\overline{u}}\right)\overline{\sigma\left(\frac{u}{\overline{u}}\right)} = 1$$

so all the conjugates of $\frac{u}{\overline{u}}$ have complex norm 1. One can show that if α is an algebraic integer and all its conjugates have norm 1, then α is a root of unity. ∎

13.5 Theorem. (Kummer) *If $p \geq 3$ is a regular prime and $p \nmid x,y,z$ for $x,y,z,\in \mathbb{Z}\setminus\{0\}$, then $x^p + y^p \neq z^p$.*

Proof Recall from a homework that the roots of unity in $\mathbb{Z}[\zeta_p]$ are $\pm\zeta^j$. In $\mathbb{Z}[\zeta]$, $z^p = x^p + y^p = \prod_{j=0}^{p-1}(x+\zeta^j y)$.

Let's show that the ideals $(x+\zeta^j y)$ are relatively prime. To the contrary, suppose ρ is a common prime factor of $(x+\zeta^j)$ and $(x+\zeta^{j'}y)$. In particular, ρ is a prime factor of

$$(x+\zeta^j y) - (x+\zeta^{j'}y) = \left(\zeta^j y(1-\zeta^{j'-j})\right) = (y(1-\zeta))$$

and $(y(1-\zeta)) \,\Big|\, (yp)$ from (iii) in (Lem. 13.3). Thus, $\rho \,\Big|\, (yp)$; but also, $\rho \,\Big|\, (z^p)$, a contradiction since (z^p), (yp) are coprime.

Since the $(x+y\zeta^j)$ are coprime and $\prod_{j=0}^{p-1}(x+y\zeta^j) = (z^p)$ is a p^{th} power, we must have each $(x+y\zeta^j) = I_j^p$. Since $p \nmid h_{\mathbb{Q}(\zeta)}$ and I_j^p is trivial in $\mathrm{Cl}(\mathbb{Q}(\zeta))$, we have that I_j is a principal.

Fix $j=1$, and we have $(x+\zeta y) = (t)^p$ so that for some $t \in \mathbb{Z}[\zeta]$ and $u \in \mathbb{Z}[\zeta]^\times$, $x+\zeta y = ut^p$. Write $t = b_0 + b_1\zeta + \cdots + b_{p-2}\zeta^{p-2}$; then modulo (p), we have $t^p \equiv b_0 + b_1 + \cdots + b_{p-2} \pmod{p}$. But then $\overline{t} = b_0 + \cdots + b_{p-2}\zeta^{-1}$, so $\overline{t}^p \equiv b_0 + \cdots + b_{p-2} \pmod{p}$ and $t^p \equiv \overline{t}^p \pmod{p}$. From (Lem. 13.4), we have that $\frac{u}{\overline{u}} = \pm\zeta^j$ for some j. Let's treat the case $\pm\zeta^j = \zeta^j$, so that

$$x + y\zeta = ut^p = \zeta^j \overline{u} t^p \equiv \zeta^j \overline{u}\overline{t}^p = \zeta^j(x+\overline{\zeta}y)$$

and

$$x + y\zeta - y\zeta^{j-1} - x\zeta^j \equiv 0 \pmod{p} \tag{13.1}$$ {e:km}

But now,

$$\begin{aligned}\mathbb{Z}[\zeta]\Big/(p) &= \mathbb{Z}[x]\Big/(p, x^{p-1}+\cdots+x+1) \\ &= F_p[x]\Big/(x^{p-1}+\cdots+x+1) = \mathbb{F}_p[x]\Big/(x-1)^{p-1}\end{aligned}$$

so, modulo p, $1, \zeta, \zeta^2, \ldots, \zeta^{p-2}$ form a basis. If $j \notin \{0,1,2,p-1\}$, then $(*)$ contradicts linear independence. (Incomplete?) ∎

14 Lattices and Minkowski's Theorem

Definition. A **lattice** is an abelian subgroup Λ of $\mathbb{R}^n$ such that $\Lambda \cong \mathbb{Z}^n$.

Example. If $[K:\mathbb{Q}] = n$, then $\mathcal{O}_K$ is a lattice in $K \cong \mathbb{Q}^n \subseteq \mathbb{R}^n$. $\mathcal{O}_K$ is a lattice since it has an integral basis.

Example. Consider $\mathbb{C} \cong \mathbb{R}^2$, and let τ be in the upper half plane. Then $\Lambda = \mathbb{Z} \oplus \mathbb{Z}\tau$, and $\mathbb{C}/\Lambda = \mathcal{T}$ is the torus. In a sense that can be made precise, $\mathcal{T}$ is an elliptic curve, and every elliptic curve arises like this.

Choose a basis $\{\alpha_1, \ldots, \alpha_n\}$ for Λ; this basis is also an $\mathbb{R}$–basis for $\mathbb{R}^n$. If $\{\alpha_1, \ldots, \alpha_n\}$ is a basis for Λ and $\{\alpha_1', \ldots, \alpha_n'\}$ is a basis for Λ, then we have a change of basis matrix

$$\begin{pmatrix}\alpha_1' \\ \vdots \\ \alpha_n'\end{pmatrix} = P \begin{pmatrix}\alpha_1 \\ \vdots \\ \alpha_n\end{pmatrix}$$

Since $P \in \mathrm{GL}_n(\mathbb{Z})$, $\det P = \pm 1$. Thus, we can define the **volume of Λ** to be

$$d(\Lambda) = |\det(\alpha_1, \ldots, \alpha_n)|$$

which is independent of the choice of matrix by the above observation.

t:blich

14.1 Theorem. (Blichfeldt) *Suppose $\Lambda \subseteq \mathbb{R}^n$ is a lattice, $m \in \mathbb{Z}^+$ and $S \subseteq \mathbb{R}^n$ with Lebesgue measure $\mu(S)$. Suppose $\mu(S) \geq md(\Lambda)$ S is compact if equality holds. Then there exist distinct $x_1, \ldots, x_{m+1} \in S$ such that $x_i - x_j \in \Lambda$.*

Proof Let $\alpha_1, \ldots, \alpha_n$ be a basis for Λ. Let $P = \left\{\sum_{i=1}^n \theta_i \alpha_i \mid 0 \leq \theta_i < 1\right\}$, so that $\mu(P) = d(\Lambda)$. For each $\lambda \in \Lambda$, let $R(\lambda) = \{\nu \in P \mid \lambda + \nu \in S\}$. Then

$$\sum_{\lambda \in \Lambda} \mu(R(\lambda)) = \mu(S) \geq m\mu(P)$$

If S is not compact, then there exists $\nu_0 \in P$ which occurs in at least $m+1$ of the $R(\lambda)$'s. If instead S is compact, for any $\epsilon_r > 0$, get $\nu_r \in P(1+\epsilon_r)$ which occurs in at least $m+1$ of $R(\lambda)$'s. This sequence has a convergent subsequence with limit ν_0 which has the same property.

Let $\lambda_1, \ldots, \lambda_{m+1}$ distinct such that if $\nu \in R(\lambda_i)$, then $x_i = \lambda_i + v_0 \in S$. Then $x_i - x_j = \lambda_i - \lambda_j \in \Lambda$. ∎

t:mink

14.2 Theorem. (Minkowski) *Let $\Lambda \subseteq \mathbb{R}^n$ be a lattice, $m \in \mathbb{Z}^+$, $S \subseteq \mathbb{R}^n$ convex and symmetric about the origin. Suppose $\mu(S) \geq m2^n d(\Lambda)$ with S compact if equality holds. Then there exist m pairs $(\lambda_1, -\lambda_1), \ldots, (\lambda_m, -\lambda_m)$ with $\lambda_j \in \Lambda \setminus \{0\}$, $\lambda_j \in S$.*

Proof Either $\mu(S/2) > md(\Lambda)$ or $\mu(S/2) = md(\Lambda)$ and $S/2$ is compact. Thus by (Thm. 14.1), there exist $x_1, \ldots, x_{m+1} \in S$ such that $x_i/2 - x_j/2 \in \Lambda$. Order these x_i such that $x_1 > x_2 > \cdots > x_{m+1}$ where we say $x_i > x_j$ if the first non-zero coordinate of $x_i - x_j$ is positive. Take $\lambda_j = x_j/2 - x_{m+1}/2$. By choice of ordering, the $\pm\lambda_j$ are distinct. Since S is symmetric, $-x_{m+1}/2 \in S$. Since S is convex, $\lambda_j = x_j/2 + (-x_{m+1})/2 \in S$. ∎

Remark. The bound is sharp: consider $S = \{(x_1, \ldots, x_n) \in \mathbb{R}^n : |x_1| < m, |x_j| < 1\}$. Then $\mu(S) = m2^n = m2^n d(\mathbb{Z}^n)$ and S contains exactly m lattice points.

Embedding $\mathcal{O}_K$ in $\mathbb{R}^n$

Suppose $[K : \mathbb{Q}] = n$ so that $K = \mathbb{Q}(\theta)$. Let $\sigma_1, \ldots, \sigma_n \hookrightarrow \mathbb{C}$ be the embeddings, so r_1 is the number of real embeddings $\{\sigma_1, \ldots \sigma_{r_1}\}$ and r_2 pairs of complex embeddings $\{\sigma_{r_1+1}, \overline{\sigma}_{r_1+1}, \ldots, \sigma_{r_1+r_2}, \overline{\sigma}_{r_1+r_2}\}$. With these embeddings, we can define $\sigma : K \hookrightarrow \mathbb{R}^n$ by

$$\alpha \mapsto \left(\sigma_1(\alpha), \ldots, \sigma_{r_1}(\alpha), \mathrm{Re}\,\sigma_{r_1+1}(\alpha), \mathrm{Im}\,\sigma_{r_1+1}(\alpha), \ldots, \mathrm{Re}\,\sigma_{r_1+r_2}(\alpha), \mathrm{Im}\,\sigma_{r_1+r_2}(\alpha)\right)$$

Equivalently, we say $\sigma : K \hookrightarrow \mathbb{R}^{r_1} \times \mathbb{C}^{r_2}$ given by

$$\alpha \mapsto \left(\sigma_1(\alpha), \ldots, \sigma_{r_1}(\alpha), \sigma_{r_1+1}(\alpha), \ldots, \sigma_{r_1+r_2}(\alpha)\right)$$

and these are equivalent by identifying $\mathbb{R}^{2r_2} \cong \mathbb{C}^{r_2}$ as vector spaces. We call this embedding the **Minkowski embedding**.

l:lat-sz

14.3 Lemma. *Let $A \neq 0$ be an ideal of $\mathcal{O}_K$. Then $\sigma(A)$ is a lattice $\Lambda \subseteq \mathbb{R}^n$ and $d(\Lambda) = 2^{-r_2}\sqrt{|\operatorname{disc}(K)|}N(A)$.*

Proof Let $\alpha_1, \dots, \alpha_n$ be an integral basis for A. Let D_0 be the determinant of the matrix whose i^{th} row is

$$D_0 = \begin{pmatrix} \sigma_1(\alpha_1) & \dots & \sigma_{r_1}(\alpha_1) & \operatorname{Re}\sigma_{r_1+1}(\alpha_1) & \cdots & \operatorname{Im}\sigma_{r_1+r_2}(\alpha_1) \\ \vdots & & \vdots & \vdots & & \vdots \\ \sigma_1(\alpha_n) & \dots & \sigma_{r_1}(\alpha_n) & \operatorname{Re}\sigma_{r_1+1}(\alpha_n) & \cdots & \operatorname{Im}\sigma_{r_1+r_2}(\alpha_n) \end{pmatrix}$$

From (Thm. 10.1), we know $\det(\sigma_j(\alpha_i)) = \sqrt{|\operatorname{disc}(K)|} \cdot N(A)$. Using the fact that $\operatorname{Re}\sigma = \frac{\sigma+\overline{\sigma}}{2}$ and $\operatorname{Im}\sigma = \frac{\sigma-\overline{\sigma}}{2i}$ combined with elementary row operations, we have that $D_0 = \left(\frac{1}{-2i}\right)^{r_2} \det(\sigma_j(\alpha_i))$. In particular, since $D_0 \neq 0$, Λ is a lattice and $d(\Lambda) = D_0$. ∎

14.4 Theorem. *If A is a non-zero ideal in $\mathcal{O}_K$, then there exists $0 \neq \alpha \in A$ such that $|N(\alpha)| \leq \left(\frac{2}{\pi}\right)^{r_2}\sqrt{|\operatorname{disc}(K)|}$.*

Proof Given $t \in \mathbb{R}^+$, let

$$S_t = \left\{(x_1, \dots, x_n) \in \mathbb{R}^n : |x_i| \leq t, i = 1, \dots, r_1; x_{r_1+2j+1}^2 + x_{r_1+2j+2}^2 \leq t^2, j = 0, \dots, r_2 - 1\right\}$$

S_t is clearly convex and symmetric, and $\mu(S_t) = 2^{r_1}\pi^{r_2}t^n$. Define

$$t = \left(\left(\frac{2}{\pi}\right)^{r_2}\sqrt{|\operatorname{disc}(K)|}N(A)\right)^{1/n}$$

so that

$$2^{r_1}\pi^{r_2}t^n = 2^n\frac{1}{r^{r_2}}\sqrt{|\operatorname{disc}(K)|}N(A)$$

We now apply (Thm. 14.2) with $m = 1$ and (Lem. 14.3) to get $0 \neq \alpha \in A$ such that $\sigma(\alpha) \in S_t$. Then

$$|N(\alpha)| = \left|\prod_{i=1}^{r_1}\sigma(\alpha_i)\right| \cdot \left|\prod_{i=1}^{r_2}\sigma_{i+r_2}(\alpha)\overline{\sigma_{i+r_2}(\alpha)}\right| \leq t^{r_1+2r_2} = t^n$$

since $\sigma(\alpha) \in S_t$. Thus, $|N(\alpha)| \leq t^n = \left(\frac{2}{\pi}\right)^{r_2}\sqrt{|\operatorname{disc}(K)|} \cdot N(A)$. ∎

The following corollary is then immediate.

14.5 Corollary. *$C_K \leq \left(\frac{2}{\pi}\right)^{r_2}\sqrt{|\operatorname{disc}(K)|}$.*

Four Squares Theorem

14.6 Theorem. *Let p be an odd prime. Then the following are equivalent:*

(i) $\left(\frac{-1}{p}\right) = 1$

(ii) $p \equiv 1 \pmod 4$

(iii) *there exists $x, y \in \mathbb{Z}$ such that $p = x^2 + y^2$.*

Here's a proof of this theorem using the algebraic tools we have developed so far.

Proof $(i \Leftrightarrow ii)$ This is (Prop. 12.1).

$(ii \Rightarrow iii)$ Since $\left(\frac{-1}{p}\right) = 1$, (p) splits in $\mathbb{Z}[i]$ so that $(p) = PQ$ in $\mathbb{Z}[i]$, where P, Q are primes. Then $p^2 = N(p) = N(P)N(Q)$ and $N(P) = p$. Since $\mathbb{Z}[i]$ is a PID, $P = (a + bi)$ so that $p = N(P) = a^2 + b^2$.

$(iii \Rightarrow ii)$ The squares modulo 4 are 0 and 1, and since p is odd, we see that $p \equiv 1 \pmod 4$. ∎

We can also see the hard implication $(ii \Rightarrow iii)$ using Minkowski's Theorem:

Proof $(ii \Rightarrow iii)$ Get $l \in \mathbb{Z}$ such that $l^2 \equiv -1 \pmod p$. Let $\Lambda \subseteq \mathbb{R}^2$ be the lattice with $\mathbb{Z}$–basis $(1, l)$ and $(0, p)$, so that $d(\Lambda) = p$. Let S be a disc with radius r; then $\mu(S) = \pi r^2 > 2^2 p$. Set $r = 2\sqrt{p/\pi}$.

By Minkowski (Thm. 14.2), S contains a non-zero lattice point $m(1, l) + n(0, p) = (m, ml + np) \in S$; in particular, $0 < m^2 + (ml + np)^2 \leq r^2 < 2p$. Then,

$$m^2 + (ml + np)^2 \equiv m^2 + (ml)^2 \equiv m^2(1 + l^2) \equiv 0 \pmod p$$

so $m^2 + (ml + np)^2 = p$. ∎

Remark. If $a_i, b_i \in \mathbb{Z}$, then $(a_1^2 + b_1^2)(a_2^2 + b_2^2) = c_1^2 + c_2^2$. To see this, $a := a_1^2 + a_2^2 = N(a_1 + ia_2)$ and $b := b_1^2 + b_2^2 = N(b_1 + ib_2)$. Then $ab = N(z^2) = c_1^2 + c_2^2$ where $z = (a_1 + ia_2)(b_1 + ib_2)$. In particular, if $n = \prod p_i$ where $p_i \equiv 1 \pmod 4$, then $n = x^2 + y^2$. In fact, you can prove $n = x^2 + y^2$ if and only if the prime factors $p \equiv 3 \pmod 4$ occur to even exponents.

p:efsi

14.7 Proposition. (Euler's Four Squares Identity) *We have*

$$\left(\sum_{i=1}^{4} a_i^2\right) \cdot \left(\sum_{i=1}^{4} b_i^2\right) = \sum_{i=1}^{4} c_i^2$$

where

$$\begin{aligned} c_1 &= a_1 b_1 - a_2 b_2 - a_3 b_3 - a_4 b_4 \\ c_2 &= a_1 b_2 + a_2 b_1 + a_3 b_4 - a_4 b_3 \\ c_3 &= a_1 b_3 - a_2 b_4 + a_3 b_1 + a_4 b_2 \\ c_4 &= a_1 b_4 + a_2 b_3 - a_3 b_2 + a_4 b_1 \end{aligned}$$

Proof You can compute this directly or recall that the norm on quaternions is multiplicative and proceed as in the preceding remark. ∎

14.8 Theorem. (Four Squares) *If $n \in \mathbb{Z}^+$, then there exists $x, y, z, w \in \mathbb{Z}$ such that $n = x^2 + y^2 + z^2 + w^2$.*

In light of the four squares identity (Prop. 14.7), it suffices to show that primes are sums of four squares.

Claim I *If p is prime, then there exists $x, y \in \mathbb{Z}$ such that $x^2 + y^2 \equiv -1 \pmod p$.*

Proof If $p \equiv 1 \pmod 4$, then $\left(\frac{-1}{p}\right) = 1$ and we may take $y = 0$. Otherwise, we suppose that $\left(\frac{-1}{p}\right) = -1$. Equivalently, we want to solve $y^2 + 1 \equiv -x^2 \pmod p$.

Note that

$$|\{y^2+1 \mid y \in \mathbb{F}_p\}| = \frac{p+1}{2}$$

($(p-1)/2$ squares, plus 0)) and $y^2+1 \neq 0$ since $\left(\frac{-1}{p}\right) = -1$. Thus y^2+1 takes $(p+1)/2$ non-zero values; since there are only $(p-1)/2$ non-zero squares, so y^2+1 must be non-square for some $y = y_0$. Then $-(y_0^2+1)$ is a square modulo p; that is, there exists x such that $x^2 \equiv -(y_0^2+1) \pmod{p}$, as required.

We can now prove the theorem.

Proof Given p prime, choose $a, b \in \mathbb{Z}$ such that $a^2+b^2 \equiv -1 \pmod{p}$. Consider the lattice $\Lambda \subseteq \mathbb{R}^4$ with basis

$$\{(1,0,a,b),(0,1,b,-a),(0,0,p,0),(0,0,0,p)\}$$

such that $d(\Lambda) = p^2$. Let S be a ball of radius r, and $\mu(S) = \pi^2 r^4/2$. Let $r^2 = 4p/\pi\sqrt{2}$, so $\mu(S) = p^2$. By Minkowski (Thm. 14.2), S contains a non-zero lattice point (x,y,z,w) with $0 < x^2+y^2+z^2+w^2 \leq r^2 < 2p$. Note that $(x,y,z,w) = \alpha v_1 + \beta v_2 + \gamma v_3 + \delta v_4$. Then $x = \alpha$, $y = \beta$, $z = a\alpha + b\beta + p\gamma$, $w = b\alpha - a\beta + p\delta$. Modulo p, we see that

$$\begin{aligned} x^2+y^2+z^2+w^2 &\equiv x^2+y^2+(ax+by)^2+(bx-ay)^2 \\ &\equiv x^2+y^2+a^2x^2+b^2y^2+b^2x^2+a^2y^2 \\ &\equiv (1+a^2+b^2)x^2+(1+a^2+b^2)y^2 \equiv 0 \end{aligned}$$

since $a^2+b^2 \equiv -1 \pmod{p}$. Thus the result follows. ∎

Dirichlet Unit Theorem

Our final application of Minkowski's theorem is a generalization of the classification of units in a quadratic number field (Thm. 5.7) to any number field.

t:dir-unit

14.9 Theorem. (Dirichlet Unit) *If K is a number field with r_1 real embeddings and $2r_2$ complex embeddings, then $\mathcal{O}_K^\times \cong U_K \times \mathbb{Z}^{r_1+r_2-1}$, where U_K is the set of roots of unity in K.*

Let $\theta : K \to V := \mathbb{R}^{r_1} \times \mathbb{C}^{r_2}$ be given by

$$\alpha \mapsto \left(\sigma_1(\alpha), \ldots, \sigma_{r_1}(\alpha), \sigma_{r_1+1}(\alpha), \ldots, \sigma_{r_1+r_2}(\alpha)\right)$$

where $\sigma_1, \ldots, \sigma_r$ are the real emebddings and $\sigma_{r_1+1}, \ldots, \overline{\sigma}_{r_1+r_2}$ are the complex embeddings. Let $N : V \to \mathbb{R}$ be given by

$$N\left(x_1, \ldots, x_{r_1}, z_{r_1+1}, \ldots, z_{r_1+r_2}\right) = \prod_{i=1}^{r_1} x_i \cdot \prod_{i=r_1+1}^{r_1+r_2} |z_i|^2$$

N is chosen so that $N(\theta(\alpha)) = N_{\mathbb{Q}}^K(\alpha)$. Furthermore, V is a ring with coordinate-wise operations; in particular, it has units $V^\times = (\mathbb{R}^\times)^{r_1} \times (\mathbb{C}^\times)^{r_2}$. Set

$$G := \{v \in V^\times : |N(v)| = 1\}$$

so that G is a subgroup of $V^\times$. G is also closed as a topological space since it is the inverse image of 1 under the continuous map $v \mapsto |N(v)|$. Define

$$U := \theta(\mathcal{O}_K^\times) = \theta(\mathcal{O}_K) \cap G$$

In particular, $\theta(\mathcal{O}_K) \subseteq V$ is a lattice (identifying $\mathbb{C} \cong \mathbb{R}^2$), and $U \subseteq G$ is a discrete group. In order to understand $\mathcal{O}_K^\times$, we want to understand the group U. A natural way to do this is to transform the multiplicative structure on U to an additive structure via the "log map"

$$L : V^\times \to \mathbb{R}^{r_1+r_2} \text{ given } \mathrm{b}(x_i; z_j) \mapsto (\log|x_i|; 2\log|z_j|)$$

This is a surjective continuous group homomorphism; as a straightforward exercise, one can show that $L(G) \cong \mathbb{R}^{r_1+r_2-1}$. We have the following diagram summarizing these constructions:

$$\begin{array}{ccccc}
G/U & \xrightarrow{L\circ\pi} & L(G)/L(U) & & \\
\pi\uparrow & & \uparrow & & \\
G & \xrightarrow{L} & L(G) & \xrightarrow{\cong} & \mathbb{R}^{r_1+r_2-1} \\
\cup & & \cup & & \\
U & \longrightarrow & L(U) & &
\end{array}$$

In order to understand $U \subseteq G$, by the first isomorphism theorem, it suffices to understand $\ker L|_U$ (cl. III) and $L(U)$ (cl. II). In fact, we claim that $L(U) \cong \mathbb{R}^{r_1+r_2-1}$. We will prove that $L(U) \subseteq L(G)$ is a lattice. Assuming this, it suffices to show that $L(G)/L(U)$ is compact; and since L is surjective, it is in fact enough to show that G/U is compact (cl. I).

c:k1 Claim I *G/U is compact in the quotient topology of $V^\times/U$.*

Proof Let $v \in V^\times$ be arbitrary; then, multiplication by v is a linear map with matrix representation M; in particular, $\det(M) = |N(v)|$. If $R \subseteq V$ is any region, then $\mu(vR) = \lambda(R) \cdot |N(v)|$; if $v \in G$, then $\lambda(R) = \lambda(vR)$.

Let $C \subseteq G$ be any compact, symmetric, convex region with $\mu(C) \geq 2^n$. For all $g \in G$, $\mu(g^{-1}C) = \lambda(C) \geq 2^n\lambda(\theta(\mathcal{O}_K))$. As well, one can verify that $g^{-1}C$ is also symmetric, compact, and convex. Thus by (Thm. 14.2), there exists $0 \neq \alpha \in \mathcal{O}_K$ such that $\theta(\alpha) \in g^{-1}C$. In particular,

$$\left|N_{K/\mathbb{Q}}(\alpha)\right| = |N(\theta(\alpha))| \in \left|N(g^{-1}C)\right|$$

Since C is compact, $|N(C)| \subseteq \mathbb{R}$ is also compact and thus contains finitely many integers. If $\alpha_1, \ldots, \alpha_m \in \mathcal{O}_K$ represent all possible $|N(\alpha_i)| \in |N(C)|$, then $|N(\alpha)| = |N(\alpha_i)|$ for some i, so $\alpha \in \alpha_i\mathcal{O}_K^\times$. But then for any $g \in G$, there exists i such that $g^{-1}C \cap \theta(\alpha_i\mathcal{O}_K^\times) \neq \emptyset$, so

$$gU \cap \theta(\alpha^{-1})C \neq \emptyset$$

Thus G/U has representatives covered by $G \cap \bigcup_{i=1}^m \theta(\alpha_i)^{-1}C$ which is a finite union of compact sets. Since G/U is closed, we are done.

c:k2 Claim II *$L(U) \subseteq L(G)$ is a lattice, and $L(U) \cong \mathbb{R}^{r_1+r_2-1}$.*

Proof *Want to show that $L(U)$ is a full lattice in $\mathbb{R}^{r_1+r_2-1}$; this happens if and only if the quotient is compact.* Recall that $L(G) \cong \mathbb{R}^{r_1+r_2-1}$ let $B = \{(y_i) : |y_i| \le b\}$ be an arbitrary hypercube. Let's show that $L(U) \cap B$ is finite. If $L(\theta(\alpha)) \in B$, then $|\sigma(\alpha)| \le e^b$ for some σ real, and $|\sigma(\alpha)| \le e^{b/2}$ for some σ complex. Then

$$\prod_\sigma (t - \sigma(\alpha)) \in \mathbb{Z}[t]$$

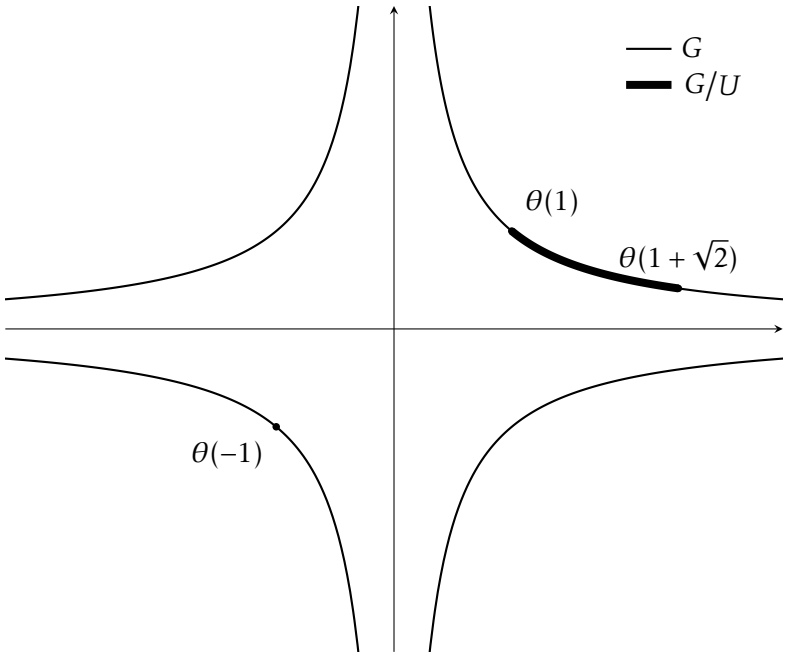
has bounded coefficients. There are only finitely many such polynomials, so there are only finitely many α and $L(U) \subseteq L(G)$ is a discrete subgruop.

Thus $L(U) \cong \mathbb{Z}^r$ for some $r \le r_1 + r_2 - 1$. Since $G/U \to L(G)/L(U)$ is a surjection, $L(G)/L(U) \cong (S^1)^r \times \mathbb{R}^{r_1+r_2-1-r}$ is compact, so $r = r_1 + r_2 - 1$.

c:k3 Claim III $\ker L|_U \cong U_K$.

Proof Clearly $\ker(L) = \{\pm 1\}^{r_1} \times (S^1)^{r_2}$ is a compact set and $\theta(U_K) \subseteq U \cap \ker(L)$. Since $U \subseteq V^\times$ is discrete, $U \cap \ker(L) \subseteq \ker L$ is compact and, in particular, a finite group. Thus by Lagrange's theorem, each $x \in U \cap \ker(L)$ has finite order, so in fact $U \cap \ker(L) \subseteq \theta(U_K)$, and equality holds. ∎

Example. To illustrate the definitions explicitly, consider the following case. Take $K = \mathbb{Q}(\sqrt{2})$ so that $\theta : K \to V = \mathbb{R}^2$ is given by $a + b\sqrt{2} \mapsto (a + b\sqrt{2}, a - b\sqrt{2})$. Then $N : \mathbb{R}^2 \to \mathbb{R}$ is given by $(x, y) \mapsto xy$, so $G = \{(x, y) : xy = \pm 1\}$ is the set of hyperbolas. Note that G is closed but not compact. In this case, $U = \theta(\pm(1 + \sqrt{2})^{\mathbb{Z}})$, and $U \subseteq G$ is discrete.



Here, we can see that G/U is compact.

Example. If $K = \mathbb{Q}(\sqrt{d})$, then $r_1 = 2$, $r_2 = 0$, $\mu_k = \{\pm 1\} \cong \mathbb{Z}/2$ and $\mathcal{O}_K^\times \cong \mathbb{Z}_2 \times \mathbb{Z}$ as in (Thm. 5.7). If $K = \mathbb{Q}(\sqrt{2}, \sqrt{3})$, then $r_2 = 0$, $r_1 = 4$, $\mathcal{O}_K^\times = \{\pm 1\} \times \mathbb{Z}^3$.