Algebraic Number Theory

George D. Torres

Math 390C - Fall 2017

These are lecture notes from Mirela Çiperiani's Algebraic Number Theory course M390C given Fall 2017. The reader should be comfortable with the essential notions of commutative algebra, general field theory, and Galois theory. Please forward any typos to `gdavtor@math.utexas.edu`.

Last updated: February 13, 2019

1. Number Fields

❖

A number field is a finite extension of $\mathbb{Q}$. To be more precise, we'll begin with some review of field extensions and algebraic elements.

Definition 1.1. A complex number $\alpha \in \mathbb{C}$ is *algebraic* if there exists $p(x) \in \mathbb{Q}[x]$ such that $p(\alpha) = 0$.

Any algebraic element α defines an ideal $I_\alpha = \{p \in \mathbb{Q}[x] \mid p(\alpha) = 0\}$. Since $\mathbb{Q}$ is a field, $\mathbb{Q}[x]$ is a PID and so $I_\alpha = (p)$ for some $p \in I_\alpha$ (which must be irreducible). The unique monic polynomial that generates I_α is called the minimal polynomial for α, which we'll denote $p_{\min,\alpha}$. We call the degree of an algebraic number $\deg(\alpha) := \deg(p_{\min,\alpha})$.

Proposition 1.2. *The following are equivalent:*

1. *α is algebraic.*
2. *$[\mathbb{Q}[\alpha] : \mathbb{Q}] < \infty$.*
3. *$\mathbb{Q}[\alpha] = \mathbb{Q}(\alpha)$.*

Proof:

($1 \Rightarrow 2$): If α is algebraic and $d = \deg(\alpha)$, then $\mathbb{Q}[\alpha] \cong \mathbb{Q}[x]/I_\alpha$ is a finite dimensional $\mathbb{Q}$ vector space with basis $\{1, \alpha, ..., \alpha^{d-1}\}$. Therefore $[\mathbb{Q}[\alpha] : \mathbb{Q}] = d < \infty$.

($1 \Rightarrow 3$): Since $p_{\min,\alpha}$ is irreducible, I_α is a maximal ideal and so $\mathbb{Q}[\alpha]$ is a field. Since it contains α, $\mathbb{Q}[\alpha] = \mathbb{Q}(\alpha)$.

($3 \Rightarrow 1$): The element $\frac{1}{\alpha} \in \mathbb{Q}(\alpha)$ is also contained in $\mathbb{Q}[\alpha]$, so $\frac{1}{\alpha} = q(\alpha)$ for $q \in \mathbb{Q}[\alpha]$. Then $\alpha q(\alpha) - 1 = 0$. Thus α satisfies the polynomial $xq(x) - 1$.

($2 \Rightarrow 1$): $\mathbb{Q}[\alpha]$ being a finite vector space over $\mathbb{Q}$ means that the elements $\{\alpha^i\}$ for i sufficiently large must be mutually dependent. This induces a polynomial relation in α with $\mathbb{Q}$ coefficients, hence α is algebraic.

□

Remark 1.3. If α, β are algebraic, then it is easy to check that $\alpha + \beta, \alpha\beta$, and α/β (the latter only when $\beta \neq 0$) are also algebraic. The algebraic elements of $\mathbb{C}$ therefore form a subfield of $\mathbb{C}$.

Definition 1.4. If F is a field containing $\mathbb{Q}$ and $[F : \mathbb{Q}] < \infty$, then F is called a *number field* of degree $n = [F : \mathbb{Q}]$.

Remark 1.5. For any number field F, the primitive element theorem guarantees us a (not necessariy unique) element $\alpha \in F$ such that $F = \mathbb{Q}(\alpha)$.

Definition 1.6. For a number field F, an *embedding* of F into $\mathbb{C}$ is a homomorphism $\phi : F \to \mathbb{C}$. An embedding is called real if $\mathrm{im}(\phi) \subset \mathbb{R}$ and complex otherwise.

Given the finiteness of F as an extension, a natural question to ask about embeddings is how many there are. We know that, if ϕ is a complex embedding, then so is its complex conjugate $\overline{\phi}$. Thus the number of embeddings must be of the form $r_1 + 2r_2$, where r_1 is the number of real embedings and $2r_2$ is the number of complex embeddings.

Corollary 1.7. *For a number field F of degree n, there are exactly n distinct embeddings.*

Proof:

Fix a primitive element α, so that $F = \mathbb{Q}(\alpha)$. Any homomorphism $\phi : F \to \mathbb{C}$ is determined by the image of α. Since $F \cong \phi(F) \cong \mathbb{Q}(\phi(\alpha)) \cong \mathbb{Q}[x]/(p_{\min,\phi(\alpha)})$ and $F \cong \mathbb{Q}[x]/(p_{\min,\alpha})$, the minimal polynomials of α and $\phi(\alpha)$ are the same. Therefore $\phi(\alpha)$ must be another root of $p_{\min,\alpha}$. Since it has no repeated roots, there are exactly n choices of roots in $\mathbb{C}$.

□

1.1 Norm, Trace, and Discriminant ❖

For $\alpha \in F$, consider the map $m_\alpha : F \to F$ given by $x \mapsto \alpha x$. This is a $\mathbb{Q}$ linear map, so it has a trace and a determinant. These are used to define the norm and trace of an element:

Definition 1.8. Given $\alpha \in F$, the *trace* of α is $\mathrm{tr}(m_\alpha)$ and is denoted $\mathrm{tr}_{F/\mathbb{Q}}(\alpha)$. The *norm* of α is the determinant $\det(m_\alpha)$ and is denoted $N_{F/\mathbb{Q}}(\alpha)$. We sometimes drop the "$F/\mathbb{Q}$" subscript when the number field is implicit.

Remark 1.9. Notice that basic properties of trace and determinant extend to the field norm and trace: $\mathrm{tr}(\alpha+\beta) = \mathrm{tr}(\alpha) + \mathrm{tr}(\beta)$ and $N(\alpha\beta) = N(\alpha)N(\beta)$.

Proposition 1.10. *Let* $\deg(F) = n$*, and* $\{\phi_1, ..., \phi_n\}$ *be the distinct embeddings of* F*. Then for any* $\alpha \in F$*:*

$$\mathrm{tr}_{F/\mathbb{Q}}(\alpha) = \sum_i \phi_i(\alpha)$$

$$N_{F/\mathbb{Q}}(\alpha) = \prod_i \phi_i(\alpha)$$

Proof:

Consider the field $\mathbb{Q}(\alpha) \subset F$ and the map $m_\alpha : \mathbb{Q}(\alpha) \to \mathbb{Q}(\alpha)$. Since the minimal polynomial p_{m_α} of m_α is $p_{\min,\alpha}$, which has no multiple roots, m_α is diagonalizable. The number of eigenvalues of m_α is the degree $m = [\mathbb{Q}(\alpha) : \mathbb{Q}]$. Then, if $\{\alpha_j\}$ are the roots of $p_{\min,\alpha}$:

$$\mathrm{tr}_{\mathbb{Q}(\alpha)/\mathbb{Q}}(\alpha) = \sum_{j=1}^{m} \alpha_j$$

Note that every embedding $\mathbb{Q}(\alpha) \to \mathbb{C}$ can be extended to $e = \deg(F/\mathbb{Q}(\alpha))$ embeddings $F \to \mathbb{C}$. Then, by the same reasoning as above, we have:

$$\mathrm{tr}_{F/\mathbb{Q}}(\alpha) = \sum_{j=1}^{m} e\alpha_j$$

But since the image of ϕ_i must be a root of $p_{\min,\alpha}$, we get:

$$\mathrm{tr}_{F/\mathbb{Q}}(\alpha) = \sum_{i=1}^{n} \phi(\alpha)$$

The same can be done to show that the product formula.

□

The trace as a map $F \to \mathbb{Q}$ induces a pairing:

$$\langle -, - \rangle_F : F \times F \to \mathbb{Q}$$

$$(x, y) \mapsto \mathrm{tr}_{F/\mathbb{Q}}(xy)$$

We sometimes drop the F subscript on the brackets unless it is necessary.

Exercise 1.11. Show that $\langle -, - \rangle_F$ is a nondegenerate bilinear pairing.

Definition 1.12. Given n elements $\alpha_1, ..., \alpha_n \in F$, the *discriminant* $\operatorname{disc}(\alpha_1, ..., \alpha_n)$ is defined to be the determinant of the matrix $\langle \alpha_i, \alpha_j \rangle$.

There are several observations about the discriminant that will be useful:

1. The discriminant can be evaluated using the embeddings ϕ_k:

$$\operatorname{disc}(\alpha_1, ..., \alpha_n) = \det(tr(\alpha_i \alpha_j)) = \det((\phi_k(\alpha_j)^T \phi_k(\alpha_j)) = \det(\phi_k(\alpha_j))^2$$

2. If $(\beta_1, ..., \beta_n) = M(\alpha_1, ..., \alpha_n)$ for some matrix M, then:

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

   This follows from the previous observation.

3. For $F = \mathbb{Q}(\alpha)$ of degree d, the discriminant of the power basis $\{1, \alpha, ..., \alpha^{d-1}\}$ is:

$$\begin{aligned} \operatorname{disc}(1, \alpha, ..., \alpha^{d-1}) &= \det \begin{pmatrix} 1 & \phi_1(\alpha) & \cdots & \phi_1(\alpha)^{d-1} \\ \vdots & \vdots & & \vdots \\ 1 & \phi_d(\alpha) & \cdots & \phi_d(\alpha)^{d-1} \end{pmatrix}^2 \\ &= \prod_{i<j} (\phi_i(\alpha) - \phi_j(\alpha))^2 \\ &= N_{F/\mathbb{Q}}(p'_{\min,\alpha}(\alpha)) \end{aligned}$$

   where we used the Vandermonde determinant formula.

Proposition 1.13. *For $F/\mathbb{Q}$ a degree d number field and $B = \{\alpha_1, ..., \alpha_d\}$ be a collection of elements in F. Then B forms a basis of $F/\mathbb{Q}$ if and only if* $\operatorname{disc}(\alpha_1, ..., \alpha_d) \neq 0$.

Proof:

> One direction is straight-forward: if they don't form a basis, there is a linear relation on the α_i. This extends to a linear relation on the columns of $\phi_i(\alpha_j)$, hence making the discriminant zero by the first observation above. For the other direction, it sufices to prove that the discriminant is nonzero for one basis (and then use observation 2). We choose the power basis $\{1, \alpha, ..., \alpha^{d-1}\}$, which has discriminant:
>
> $$\operatorname{disc}(1, \alpha, ..., \alpha^{d-1}) = \prod_{i<j} (\phi_i(\alpha) - \phi_j(\alpha))^2$$
>
> This is nonzero because otherwise $\phi_i(\alpha) = \phi_j(\alpha)$ for $j \neq i$, which would imply that ϕ_i and ϕ_j sent α to the same thing.

□

As an application, we construct the total embedding of F into $\mathbb{R}^{r_1} \times \mathbb{C}^{r_2}$, where r_1 is the number of real embeddings and r_2 is the number of pairs of complex embeddings. For each pair of complex embeddings $(\phi_i, \overline{\phi_i})$, we chose one representative to define a map:

$$\begin{aligned} \Phi : F &\to \mathbb{R}^{r_1} \times \mathbb{C}^{r_2} \\ \alpha &\mapsto (\underbrace{\phi_1(\alpha), ..., \phi_{r_1}(\alpha)}_{\text{real}}, \underbrace{\phi_{r_1+1}(\alpha), ..., \phi_{r_1+r_2}(\alpha)}_{\text{complex rep.'s}}) \end{aligned}$$

This is non-canonical, since we made arbitrary choices of representatives. The question we will answer is: what does the image of Φ look like?

Example 1.14. Consider $F = \mathbb{Q}(\alpha)$ for $\alpha^3 = 2$. Then there are two complex embeddings and one real. If ω is a (primitive) cube root of unity, then one example of Φ would be:

$$\Phi(\alpha) = \left(\sqrt[3]{2}, \omega\sqrt[3]{2}\right)$$

Theorem 1.15. *For a number field F of degree d:*

- Φ *maps a basis of $F/\mathbb{Q}$ to an $\mathbb{R}$ basis of $\mathbb{R}^{r_1} \times \mathbb{C}^{r_2}$.*
- Φ *is injective.*
- $\mathrm{im}(\Phi)$ *is dense.*

Proof:

We identify $\mathbb{C}$ with $\mathbb{R}^2$ in the usual way. This induces a map $\tilde{\Phi} : F \to \mathbb{R}^d$.

1. If $\omega_1, ..., \omega_d$ is a basis of $F/\mathbb{Q}$, then by multilinearity of the determinant:

$$\det \begin{pmatrix} \tilde{\Phi}(\omega_1) \\ \vdots \\ \tilde{\Phi}(\omega_d) \end{pmatrix} = (2i)^{-r_2} \underbrace{\det(\phi_i(\omega_j))}_{\neq 0}$$

   Therefore $\{\tilde{\Phi}(\omega_i)\}$ form a basis of $\mathbb{R}^d$, and hence so do $\{\Phi(\omega_i)\}$.
2. Suppose $\Phi(\beta) = 0$. Extend β to a basis B of $F/\mathbb{Q}$. Then $\Phi(B)$ cannot be a basis because it has at most $d - 1$ linearly independent elements. This contradicts 1).
3. We have shown that $\Phi(F)$ contains a $\mathbb{R}^{r_1} \times \mathbb{C}^{r_2}$ over $\mathbb{Q}$. Since $\mathbb{Q}$ is dense in $\mathbb{R}$, so too must $\Phi(F)$ be dense in $\mathbb{R}^{r_1} \times \mathbb{C}^{r_2}$.

□

1.2 Algebraic Integers ❖

Definition 1.16. For a number field F, an element $\alpha \in F$ is an *algebraic integer* (also called integral) if $p_{\min,\alpha}$ has integer coefficients.

Proposition 1.17. *Let α be an element of a number field F. The following are equivalent:*

1. *α is integral.*
2. *There exists a monic $f(x) \in \mathbb{Z}[x]$ such that $f(\alpha) = 0$.*
3. *$\mathbb{Z}[\alpha]$ is a finitely generated $\mathbb{Z}$ module.*
4. *There exists a finitely generated, nonzero $\mathbb{Z}$ module $M \subset F$ such that $\alpha M \subset M$.*

Proof:

($1 \Rightarrow 2$): Take $f(x)$ to be the minimal polynomial of α.
($2 \Rightarrow 1$): Write $f(x) = p_{\min,\alpha}(x)q(x)$ for $q(x) \in \mathbb{Q}[x]$ monic. If $r = \deg(p_{\min,\alpha})$ and $p_{\min,\alpha} = \sum_{j=1}^{r} \beta_j x^j$ then the coefficients of the r highest degree terms $f(x)$ are $\{\beta_1, ..., \beta_r\}$. These are integers by assumption, so in fact $p_{\min,\alpha} \in \mathbb{Z}[x]$.
($1 \Rightarrow 3$): A basis for $\mathbb{Z}[\alpha] \cong \mathbb{Z}[x]/(p_{\min,\alpha})$ is $\{1, \alpha, ..., \alpha^{r-1}\}$, so it is finitely generated.
($3 \Rightarrow 4$): Take $M = \mathbb{Z}[\alpha]$.

($4 \Rightarrow 2$): Take $x_1, ..., x_n$ as as generators for M. Then since $\alpha x_i \in M$ for all i, we have a matrix relation:

$$\begin{pmatrix} \alpha x_1 \\ \vdots \\ \alpha x_n \end{pmatrix} = N \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

for N a matrix with integer coefficients. Then $\det(\alpha I_n - N) = 0$, and so the polynomial $\det(xI_n - N) \in \mathbb{Z}[x]$ satisfies α.

□

Remark 1.18. If α and β are algebraic integers, then $\mathbb{Z}[\alpha + \beta] \subset \mathbb{Z}[\alpha] \oplus \mathbb{Z}[\beta]$ must be finitely generated as a $\mathbb{Z}$ module because each of $\mathbb{Z}[\alpha], \mathbb{Z}[\beta]$ is. Therefore, by above, $\alpha + \beta$ is an algebraic integer. For similar reasons, $\alpha\beta$ is also an algebraic integer. This proves:

Corollary 1.19. *The set of all algebraic integers in F (denoted O_F) is a subring.*

Exercise 1.20. Show that $\mathrm{Frac}(O_F) = F$.

Theorem 1.21. *For a degree number field F of degree d, the ring of integers O_F is a free $\mathbb{Z}$ module of rank d.*

Proof:

Let $\alpha_1, ..., \alpha_n$ be a basis of $F/\mathbb{Q}$. For each α_i, there exists an integer m such that $m\alpha_i$ is integral (just take m to be the lcm of the denominators of the coefficients of $p_{\min,\alpha_i}$). Thus, we can assume $\{\alpha_i\}$ are integral. This directly means that the rank of O_F is at least d. Now define the map $\varphi : F \to \mathbb{Q}^d$ by:

$$\varphi(\beta) = (\langle \alpha_1, \beta \rangle, \langle \alpha_2, \beta \rangle, ..., \langle \alpha_d, \beta \rangle)$$

Note that, for $x, y \in O_F$, $\langle x, y \rangle \in \mathbb{Z}$ because the trace of an algebraic integer is again an algebraic integers in $\mathbb{Q}$ (i.e. they are in $\mathbb{Z}$). Therefore φ sends O_F to $\mathbb{Z}^d$. Further, this is an injection because the $\langle -, - \rangle$ is nondegenerate. Therefore the rank of O_F is at most d. Since it clearly has no torsion, it must be free of rank exactly d.

□

Example 1.22. Let F be a quadratic number field (degree 2). Then there is a squarefree integer d such that $F = \mathbb{Q}(\sqrt{d})$. It is a standard exercise to show that the ring of integers O_F depends on d mod 4 in the following way:

$$O_F = \begin{cases} \mathbb{Z}[\sqrt{d}] & d \equiv 2, 3 \bmod 4 \\ \mathbb{Z}[(1 + \sqrt{d})/2] & d \equiv 1 \bmod 4 \end{cases}$$

Remark 1.23. Suppose $\alpha = (\alpha_1, ..., \alpha_n)$ and $\beta = (\beta_1, ..., \beta_n)$ are two bases of O_F. Then there is an integer matrix M such that $\alpha = M\beta$. Since both M and M^{-1} are integer matrices, $\det(M) \in \mathbb{Z}^\times = \{\pm 1\}$. In particualar, the discriminants $\mathrm{disc}(\alpha_1, ..., \alpha_n)$ and $\mathrm{disc}(\beta_1, ..., \beta_n)$ are equal and they define an invariant of O_F. This is known as the discriminant of the number field, denoted $\mathrm{disc}\, F$.

Lemma 1.24. *Any set of elements $\{\beta_1, ..., \beta_d\}$ in O_F with squarefree discriminant form a basis of O_F.*

Proof:

If $\{\beta_i\}$ is not a basis, then writing $\beta = M\alpha$ for a basis $\alpha = \{\alpha_i\}$ requires $\det(M) \notin \mathbb{Z}^*$; in particular, $\mathrm{disc}(\beta_1, ..., \beta_n) = \det(M)^2 \,\mathrm{disc}(\alpha_1, ..., \alpha_n)$ is not squarefree.

□

One should keep in mind that this is not an if and only if statement.

2. Dedekind Rings

❖

In this section, we will define Dedekind rings in order to give a theory of ideal facorization in O_F. Recall these equivalent definitions of a ring R being Noetherian:

1. Every ideal $I \subset R$ is finitely generated.
2. Every nonempty collection of ideals in R has a maximal element.
3. Every ascending chain of ideals $I_0 \subset I_1 \subset I_2 \subset ...$ eventually stabilizes.

Exercise 2.1. Show that these are equivalent.

An important fact about Noetherian rings is the Hilbert Basis Theorem:

Proposition 2.2. *If R is a Noetherian ring, then so is $R[x]$ (and hence so is $R[x_1, ...x_n]$).*

Definition 2.3. Suppose $R \subset S$ are rings. We say $\alpha \in S$ is *integral* over R if there exists a monic polynomial $f \in R[x]$ such that $f(\alpha) = 0$.

Definition 2.4. A domain R is *integrally closed* if every $\alpha \in \text{Frac}(R)$ that is integral is in R itself.

Example 2.5. $\mathbb{Z}[\sqrt{5}]$ is not integrally closed because $\frac{\sqrt{5}+1}{2} \in \mathbb{Q}(\sqrt{5})$ is integral over $\mathbb{Z}$ (and hence over $\mathbb{Z}[\sqrt{5}]$) but is not in $\mathbb{Z}[\sqrt{5}]$.

Definition 2.6. The *Krull dimension* of a commutative ring R is the length of the longest chain of nested prime ideals $\mathfrak{p}_0 \subset \mathfrak{p}_1 \subset \mathfrak{p}_2 \subset ... \subset \mathfrak{p}_n$ (length is determined by the number of inclusions, not the number of primes).

Example 2.7. The Krull dimension of $\mathbb{Z}$ is 1 and the Krull dimension of $\mathbb{Z}[x_1, ..., x_n]$ is $n+1$. More generally, the Krull dimension of $R[x]$ is one more than the Krull dimension of R.

Definition 2.8. A commutative ring R is *Dedekind* if:

1. R is a Noetherian domain.
2. R is integrally closed.
3. The Krull dimension of R is ≤ 1 (i.e. non-zero primes are maximal).

Proposition 2.9. *The ring of integers O_F of a number field F is a Dedekind ring.*

Proof:

Since O_F is finitely generated as a $\mathbb{Z}$ module, any submodule (ideal) is also finitely generated. To show that it is integrally closed, let $\alpha \in F$ be integral over O_F. Then $O_F[\alpha]$ is a finite O_F module, and hence also a finite $\mathbb{Z}$ module. Therefore the submodule $\mathbb{Z}[\alpha]$ is also finitely generated, which is true if and only if α is integral over $\mathbb{Z}$.

Finally, let $\mathfrak{p} \subset O_F$ be nonzero. Then $O_F/\mathfrak{p}$ is an integral domain. We claim that $\mathfrak{p}$ contains a nonzero integer m; inded, for any nonzero $\alpha \in \mathfrak{p}$, the degree 0 coefficient of $p_{\min,\alpha}$ is in $\mathfrak{p}$. Thus $(m) \subset \mathfrak{p}$ and we have a ring surjection:

$$O_F/(m) \twoheadrightarrow O_F/\mathfrak{p}$$

We note that, if d is the rank of O_F, then $O_F/(m) \cong (\mathbb{Z}/m\mathbb{Z})^{\oplus d}$. In particular, it is finite. Therefore $O_F/\mathfrak{p}$ is finite. Any finite integral domain is a field, so $\mathfrak{p}$ must be maximal. Therefore the Krull dimension is 1.

□

2.1 Fractional Ideals and Unique Factorization ❖

Definition 2.10. For a Dedekind ring R, let F denote the fraction field of R. A *fractional ideal* I of R is an additive subgroup of F such that $\exists \alpha \in F^*$ such that αI is a nonzero ideal of R.

Proposition 2.11. *The following are true for a Dedekind ring R with $F = Frac(R)$:*

1. *Every nonzero ideal of R is fractional.*
2. *Every fractional ideal that is contained in R is an ideal of R.*
3. *If I, J are fractional ideals, then $IJ = \{\sum \alpha_i \beta_i \mid \alpha_i \in I, \beta_i \in J\}$ is a fractional ideal.*
4. *αR is a fractional ideal of R for any nonzero α.*
5. *For all fractional ideals I, the set $I^{-1} = \{\alpha \in F \mid \alpha I \subset R\}$ is a fractional ideal of R.*

Proof:

1. Clearly $1 \cdot I \subset R$.
2. Let $I \subset R$ be fractional. Then it is easy to check that I is an additive subgroup of R, so we need only verify that $rI \subset I$ for any $r \in R$. Let $\alpha \in F^*$ be such that αI is an ideal of R. Then $r\alpha I \subset \alpha I$. For any $a \in I$, this inclusion means that there exists $b \in I$ such that $ra\alpha = b\alpha$, and therefore:
$$\alpha(ra - b) = 0 \Rightarrow ra - b = 0 \Rightarrow ra = b.$$
Therefore $ra \in I$, so $rI \subset I$.
3. If α, β are such that $\alpha I \subset R$ and $\beta J \subset R$, then clearly $\alpha\beta IJ \subset R$.
4. If $\alpha \neq 0$, then αR is an additive subgroup of F and $\alpha\alpha^{-1}R \subset R$, so it is also a fractional ideal.
5. Exercise.

□

The main fact that we will use about Dedekind rings is that they admit prime factorization, just like in the integers. In fact, we can think of Dedekind rings as a generalization of the integers (just as O_F was the generalization of integers as a subring $\mathbb{Q}$).

Theorem 2.12. *Let R be a Dedekind ring. Then the set $\mathcal{I}(R) = \{$fractional ideals $I \subset Frac(R)\}$ is abelian group under multiplication. Moreover, $\mathcal{I}(R) \cong \bigoplus_{primes} \mathbb{Z}$. In other words, every fractional ideal factors uniquely into a finite product of powers of prime ideals.*

To prove this, we will need three lemmata:

Lemma 2.13. *For R a Noetherian ring, every ideal $I \subset R$ contains a product of prime ideals.*

Proof:

Define the set:
$$S = \{I \subset R \mid I \text{ does not contain a product of prime ideals}\}$$
Assume that S is nonempty; then there is a maximal element I_0 of S. Since I_0 cannot be a prime ideal, there exists $r, s \in R$ such that $rs \in I_0$ but $r, s \notin I_0$. Then (I_0, r) and (I_0, s) are ideals containing I_0 properly. Since I_0 was maximal in S, then (I_0, r) and (I_0, s) both contain a product of prime ideals. Then $(I_0, r)(I_0, s)$ also contains a product of prime ideals. However, by construction of r and s, $(I_0, r)(I_0, s) \subset I_0$, and so I_0 contains a product of primes. This is a contradiction, so $S = \emptyset$.

□

Lemma 2.14. *Let $R \subset F$ be a Dedekind ring contained in a field. For $I \subset R$ a proper ideal, then $\exists \lambda \in F \setminus R$ such that $\lambda I \subset R$.*

Proof:

Let $\mathfrak{p}$ be a prime ideal containing I and fix a nonzero element $a \in I$. Then by the previous Lemma, the ideal it generates contains a product of primes: $\mathfrak{p}_1\mathfrak{p}_2\cdots\mathfrak{p}_n \subset (a)$. Further, we can take n to be the minimal such length. Then we have $\mathfrak{p}_1\mathfrak{p}_2\cdots\mathfrak{p}_n \subset \mathfrak{p}$, which means $\mathfrak{p} = \mathfrak{p}_i$ for some i. Without loss of generality, let $i = 1$. Since n was minimal, $\mathfrak{p}_2\cdots\mathfrak{p}_n$ is no longer a subset of (a), and so we may take $b \in \mathfrak{p}_2\cdots\mathfrak{p}_n \setminus (a)$. Then $b/a \in F \setminus R$. Further, any $x \in I$ is also in $\mathfrak{p}_1 = \mathfrak{p}$, so $bx \in \mathfrak{p}_1\mathfrak{p}_2\cdots\mathfrak{p}_n \subset (a)$. Therefore $bx/a \in R$ and $\frac{b}{a}I \subset R$.

□

Lemma 2.15. *If I is a fractional ideal of R, then $II^{-1} = R$.*

Proof:

By definition of I^{-1}, the fractional ideal II^{-1} is contained in R, so it must be an ideal of R (part 2. of Proposition 2.11). Assume that $II^{-1} \neq R$; then by the previous Lemma, there exists some $\lambda \in F \setminus R$ such that $\lambda II^{-1} \subset R$. If we rewrite λII^{-1} as $I \cdot \lambda I^{-1}$, we see that:

$$\lambda I^{-1} \subset I^{-1} \Rightarrow \lambda^n I^{-1} \subset I^{-1}\ \forall n$$

Now pick $y \in I$ nonzero, and notice that $\lambda^n y I^{-1} \subset yI^{-1} \subset R$. Fixing $x \in I^{-1}$ nonzero, this says that $\lambda^n xy \in R$ and therefore $\lambda^n \in Rx^{-1}y^{-1}$. This means that $R[\lambda] \subset Rx^{-1}y^{-1}$. Since R is Noetherian, we must conclude that $R[\lambda]$ is a finitely generated submodule, which is true if and only if λ is integral. But $\lambda \notin R$ by choice, and R is integrally closed. This is a contradiction.

□

The proof of Theorem 2.12 will follow from the next few claims.

Claim 1: Every nonzero prime ideal of R is a finite product of prime ideals.

Proof:

Let $S = \{J \subset R \mid J \text{ is not a product of primes}\}$ and assume it is nonempty. Then it has a maximal element I. Since I is not prime, it is not maximal so $I \subset \mathfrak{p}$ properly for some prime $\mathfrak{p}$. Consider $I\mathfrak{p}^{-1} \subset \mathfrak{p}\mathfrak{p}^{-1} = R$. This is a proper ideal of R. By Lemma 2.14, there exists some $\lambda \in F \setminus R$ such that $\lambda \in \mathfrak{p}^{-1}$. Therefore $\mathfrak{p}^{-1} \not\subset R$, which makes I a strict subset of $I\mathfrak{p}^{-1}$. Since I was maximal in S, we must be able to write $I\mathfrak{p}^{-1} = \mathfrak{p}_1\cdots\mathfrak{p}_n$ for $\mathfrak{p}_i$ prime. However, that would mean $I = \mathfrak{p}\mathfrak{p}_1\cdots\mathfrak{p}_n$, which is a contradiction. Therefore S is empty.

□

Claim 2: The expression of an ideal as a product of primes is unique (up to reordering).

Proof:

Suppose $\mathfrak{p}_1\cdots\mathfrak{p}_r = \mathfrak{q}_1\cdots\mathfrak{q}_s$ are two expressions of I (we can assume $s \geq r$). Then $\mathfrak{q}_1\cdots\mathfrak{q}_s \subset \mathfrak{p}_1$. Assume $\mathfrak{q}_i \neq \mathfrak{p}_1$ for all i. Then choose $x_i \in \mathfrak{q}_i$ with $x_i \notin \mathfrak{p}_1$. Then the product $x_1\cdots x_s$ is in $\mathfrak{q}_1\cdots\mathfrak{q}_s$ and hence is in $\mathfrak{p}_1$. This is a contradiction because $\mathfrak{p}_1$ is prime, so we must have $\mathfrak{q}_i = \mathfrak{p}_1$ for some i. Without loss of generality $i = 1$, so:

$$\mathfrak{p}_1\cdots\mathfrak{p}_r = \mathfrak{p}_1\mathfrak{q}_2\cdots\mathfrak{q}_s$$

We can multiply both sides by $\mathfrak{p}_1^{-1}$ to eliminate $\mathfrak{p}_1$ from the product. Repeating this, we end up with $R = \mathfrak{q}_{s-r}\cdots\mathfrak{q}_s$. Since prime ideals are proper, so we must have $r = s$. Further, each $\mathfrak{q}_i$ was equal to some $\mathfrak{p}_j$.

□

Claim 3: The previous two claims apply to all fractional ideals (except possibly with negative powers).

Proof:

If I is fractional, then let $a, b \in R$ be such that $\frac{a}{b} I \subset R$. Then we can write:

$$\frac{a}{b} I = (a)(b)^{-1} I = \mathfrak{p}_1 \cdots \mathfrak{p}_n$$

uniquely. Rearranging:

$$I = \mathfrak{p}_1 \cdots \mathfrak{p}_n \cdot (a)^{-1}(b)$$

Further, we can also write (a) and (b) as a product of primes, which in turn expresses I as a product of (possibly negative) powers of primes. We leave verifying uniqueness as an exercise.

□

The following useful fact comes as a corollary to unique factorzation:

Corollary 2.16. *For ideals $I \subset J$ ideals of a Dedekind ring R, there exists an ideal $J' \subset R$ such that $I = JJ'$.*

Proof:

Write $J = \mathfrak{p}_1^{e_1} \cdots \mathfrak{p}_n^{e_n}$. Since $I \subset J$, we have $I \subset \mathfrak{p}_1$ and therefore $I\mathfrak{p}_1^{-e_1} \subset J\mathfrak{p}_1^{-e_1} = \mathfrak{p}_2^{e_2} \cdots \mathfrak{p}_n^{e_n}$. Repeating this until we exhaust prime factors of J, we get $IJ^{-1} \subset R$. Since it is contained in R, it is an ideal so we can set $J' = IJ^{-1}$.

□

As an application, let's briefly return to the case of $R = O_F$ for a number field F. Fix a prime ideal $\mathfrak{p} \subset O_F$ and consider the diagram:

$$\begin{array}{ccccc} F & \supset & O_F & \supset & \mathfrak{p} \\ | & & | & & | \\ Q & \supset & \mathbb{Z} & \supset & \mathfrak{p} \cap \mathbb{Z} \end{array}$$

The ideal $\mathfrak{p} \cap \mathbb{Z}$ has a generator which must be prime, so we can write $(p) = \mathfrak{p} \cap \mathbb{Z}$ for a prime $p \in \mathbb{Z}$. In this situation, we say that "$\mathfrak{p}$ lies above p." It is not hard to show that $O_F/\mathfrak{p}$ is a finite integral domain[1] , and contains a copy of $\mathbb{Z}/p\mathbb{Z}$. Therefore $O_F/\mathfrak{p} \cong \mathbb{F}_{p^n}$ for some n. Moreover, since $p \in \mathfrak{p}$, we have $pO_F \subset \mathfrak{p}$. Corollary 2.16 then says that there is some J such that $J\mathfrak{p} = pO_F$. In particular, $\mathfrak{p}$ divisdes pO_F. Since pO_F has only finitely many prime factors, it follows that there can only be finitely many primes $\mathfrak{p}$ lying over p.

2.2 The Ideal Class Group ❖

For a Dedekind ring R, we how have a well-defined group structure on the set of fractional ideals $\mathcal{I}(R)$. We can then define the group homomorphism:

$$\varphi : F^* \to \mathcal{I}(R)$$
$$\alpha \mapsto (\alpha)$$

The kernel of φ is the group of units R^*. The cokernel of φ is $\mathcal{I}(R)/\operatorname{im}(\varphi)$. This is the *ideal class group* of R, denoted $\mathrm{Cl}(R)$. It is the group of fractional ideals modulo principal fractional ideals. Thus, for any Dedekind ring R, we have the exact sequence:

$$0 \longrightarrow R^* \longrightarrow F^* \longrightarrow \mathcal{I}(R) \longrightarrow \mathrm{Cl}(R) \longrightarrow 0$$

The first fact we would like to show is that $\mathrm{Cl}(R)$ is finite for any R. To start, it is easy to characterize when it is trivial:

Proposition 2.17. *Let R be a Dedekind ring. Then the following are equivalent:*

1. $\mathrm{Cl}(R) = \{1\}$.

[1] Hint: show that $\mathfrak{p}$ contains a prime number p, hence $|O_F/\mathfrak{p}| \leq |O_F/p| < \infty$

2. *Every fractional ideal is principal.*
3. *R is a PID.*
4. *R is a UFD.*

Proof:

$(1 \iff 2)$ This is by definition of $\mathrm{Cl}(R)$.
$(2 \Rightarrow 3 \Rightarrow 4)$ Exercise.
$(4 \Rightarrow 2)$ It suffices to show that every prime ideal is principal. Let $\mathfrak{p}$ be prime and fix $x \in \mathfrak{p}$ nonzero. Since R is a UFD, we can write $x = \alpha_1 \alpha_2 \cdots \alpha_n$ for α_i irreducible. Since $\mathfrak{p}$ is prime, $\alpha_i \in \mathfrak{p}$ for some i. Therefore $(\alpha_i) \subset \mathfrak{p}$; but since α_i is irreducible, (α_i) is prime. Therefore Krull dimension ≤ 1 implies that $\mathfrak{p} = (\alpha_i)$.

□

Theorem 2.18. *Let F be a number field with ring of integers O_F. Then $\mathrm{Cl}(O_F)$ is a finite group.*

Remark 2.19. Sometimes we write $\mathrm{Cl}(F)$ instead of $\mathrm{Cl}(O_F)$. The order of the class group $h_F := |\mathrm{Cl}(O_F)|$ is known as the *class number* of F.

Our strategy for proving this will be to define a norm on fractional ideals and show that this norm is bounded on $\mathrm{Cl}(O_F)$. We can define the norm of an ideal using the following fact: every ideal $I \subset O_F$ is a product of prime ideals and $O_F/\mathfrak{p}$ is finite for any prime, so the ring O_F/I is finite.

Definition 2.20. Let $I \subset O_F$ be an ideal. Then the norm of I is $N_{F/\mathbb{Q}}(I) := |O_F/I|$.

Remark 2.21. Recall that for any $x \in O_F$, the determinant of the multiplication map $m_x : O_F \to O_F$ defines the norm of x. However, $|\det(m_x)| = |\mathrm{coker}(m_x)| = |O_F/mO_F| = N_{F/\mathbb{Q}}(mO_F)$. Thus the norm of an element coincides with the norm of the ideal that it generates (in absolute value).

Lemma 2.22. *Let I, J be ideals of O_F. Then $N(IJ) = N(I)N(J)$.*

Proof:

Since we have ideal factorization, it suffices to prove this when I is prime. Notice that:

$$N(IJ) = |O_F/IJ| = |O_F/J||J/IJ|$$

Therefore we must show $|I/IJ| = |O_F/I|$. We showed earlier that O_F/I is a finite field, and hence J/IJ is a vector space over that field. Pick $\alpha \in J \setminus IJ$, and consider the linear map $m_\alpha : O_F/I \to J/IJ$ given by $x \mapsto \alpha x$. This is an injection and $\mathrm{im}(m_\alpha) = \alpha O_F + IJ$. Since $\alpha O_F \subset J$, by Lemma 2.16 there exists an ideal J' such that $\alpha O_F = JJ'$. Further:

$$\begin{aligned} JJ' + IJ = J &\iff J' + I = O_F \\ &\iff I \nmid J' \\ &\iff J' \not\subset I \end{aligned}$$

If we suppose $J' \subset I$, then $\alpha O_F = JJ' \subset IJ$, and hence $\alpha \in IJ$. This is a contradiction, so $J' \not\subset I$. By above, this then means $JJ' + IJ = J$. Therefore $\mathrm{im}(m_\alpha) = \alpha O_F + IJ = J$, which makes m_α a surjection. We have thus demonstrated an isomorphism $I/IJ \cong O_F/I$ as (finite) vector spaces, which gives us the desired result.

□

We can now use the multiplicativity of the norm to define norms of fractional ideals:

$$1 = N(II^{-1}) = N(I)N(I^{-1}) \Rightarrow N(I^{-1}) = \frac{1}{N(I)}.$$

In other words, if $I = \prod_i \mathfrak{p}_i^{e_i}$, then:

$$N(I) = \prod_i N(\mathfrak{p}_i)^{e_i}$$

Proposition 2.23. *For every number field $F/\mathbb{Q}$, there exists an integer G such that for every ideal I there exists $\alpha \in I$ with $N(\alpha) \leq GN(I)$.*

Proof:

Let $\alpha_1, ..., \alpha_d$ be a basis of O_F and consider the set:

$$S_m = \left\{ \sum_i x_i \alpha_i \mid x_i \in \mathbb{Z}, 0 \leq x_i \leq m \right\} \subseteq O_F$$

This has size $(m+1)^d$. Choose m such that $m^d \leq N(I) < (m+1)^d$ and let $\pi : O_F \to O_F/I$ be the natural projection. We have chosen m such that $\pi(S_m) = O_F/I$, so there must exist $s_1, s_2 \in S_m$ such that $\pi(s_1) = \pi(s_2)$. Let $\alpha = s_1 - s_2 \in I$. Since the ocoefficients of s_1 and s_2 were positive and less than m, we can write:

$$\alpha = \sum_i x_i \alpha_i$$

for $|x_i| < m$. Now for each i define:

$$G_i = \sum_{j=1}^{d} |\sigma_i(\alpha_j)|$$

where $\{\sigma_j\}$ are the distinct embeddings $F \hookrightarrow \mathbb{C}$. We note that:

$$|\sigma_j(\alpha)| \leq \sum_{j=1}^{d} |x_i||\sigma_j(\alpha_i)| \leq G_j m$$

Therefore:

$$N(\alpha) = \left| \prod_j \sigma_j(\alpha) \right| \leq \underbrace{\left(\prod_j^d G_j \right)}_{G} m^d$$

Since $N(I) \geq m^d$, this inequality becomes $N(\alpha) \leq GN(I)$ as desired.

□

This bound is the necessary tool in proving that $\mathrm{Cl}(O_F)$ is finite.

Proof (of Theorem 2.18):

We will first show that every ideal class $c \in \mathrm{Cl}(O_F)$ has a representing ideal with norm less than G, then we will show that the set of ideals with norm less than G is finite. From this, it will follow that $\mathrm{Cl}(O_F)$ is finite.

Given $c \in \mathrm{Cl}(O_F)$, let I be an ideal representative of c^{-1}. Then, by Proposition 2.23, there exists $\alpha \in I$ such that $N(\alpha) \leq GN(I)$. If we consider the ideal $\alpha O_F \subset I$, by Corollary 2.16 there exists J such that $\alpha O_F = IJ$. Further, J is a representative of c and:

$$N(J) = N(\alpha)N(I)^{-1} \leq GN(I)N(I)^{-1} = G$$

Therefore every ideal class has a representative with norm bounded by G.

Now consider the set:

$$B = \{I \subset O_F \mid N(I) \leq G\}$$

If we write any element of this set as $I = \prod_{i=1}^{t} \mathfrak{p}_i^{e_i}$, we have:

$$N(I) = \prod_{i=1}^{t} N(\mathfrak{p}_i)^{e_i} \leq G$$

We showed at the end of Section 2.1 that $N(\mathfrak{p}_i) = p_i^{m_i}$, where $\mathfrak{p}_i$ lies over p_i and $m_i \geq 1$. Since there are finitely many primes less than or equal to G and finitely many primes lying over a given prime ideal, there can only be finitely many products shown above. Since ideal factorization is unique, it then follows that B is a finite set.

□

3. Geometry of Numbers

❖

While the preceeding discussion could provide a bound on the size of $\mathrm{Cl}(O_F)$, it would be very crude and wouldn't be very sensitive to the structure of the number field F. Introducing the language of geometry and lattices will, among other things, give us a better picture of the class group and better bounds on it size. To start, here a few characteristics of finitely generated abelian groups that we will use:

Theorem 3.1. *Let G be a free abelian group of rank n and H be a subgroup. Then:*

1. *H is free of rank $m \leq n$.*
2. *There is a basis $e_1, ..., e_n$ of G and a set of positive integers $a_1|a_2|...|a_m$ such that $\{a_1e_1, ..., a_me_m\}$ is a basis of H.*

Using this, one can show that if A is a finitely generated abelian group, it is isomorphic to $\mathbb{Z}^r \oplus \mathbb{Z}/m_1\mathbb{Z} \oplus ... \oplus \mathbb{Z}/m_t\mathbb{Z}$ for some r and m_i. In particular, A is finite if and only if its rank is 0.

Corollary 3.2. *For an n by n matrix M with entries in $\mathbb{Z}$ acting on $G = \mathbb{Z}^n$ with image $H = M(G)$, we have:*

1. *G/H is finite if and only if $\det(M) \neq 0$.*
2. *If $\det(M) \neq 0$, then $|\det(M)| = [G : H]$.*

3.1 Lattices ❖

Lattices discrete subgroups of the Euclidean plane. Here we will develop some of the necessary tools associated to lattices to prove two of Minkowski's important theorems about lattices and number fields. One of these says that given a lattice, a convex symmetric body of a certain volume must intersect it somewhere. The other uses this fact to place a tighter bound than the one in Proposition 2.23.

Definition 3.3. For a real vector space V of finite dimension, a subset $L \subset V$ is a *lattice* if there exists $e_1, ..., e_n \in L$ such that:

1. $L = \mathbb{Z}e_1 + ... + \mathbb{Z}e_n$.
2. $e_1, ..., e_n$ is a basis of V over $\mathbb{R}$.

Remark 3.4. Some sources relax the second condition, allowing for lattices whose rank is less than the dimension of V. However, these are clearly just full rank lattices in a subspace of V.

Example 3.5. The standard example of lattice is $\mathbb{Z}^n \subset \mathbb{R}^n$. So too is $M(\mathbb{Z}^n)$ for a matrix M of full rank. In fact, these are all of the lattices in $\mathbb{R}^n$.

Proposition 3.6. *For a number field F, the ring of integers O_F is a lattice in $F \otimes_{\mathbb{Q}} \mathbb{R} \cong \mathbb{R}^{r_1} \times \mathbb{C}^{r_2}$ (where r_1 is the number of real embeddings of F and r_2 is half the number of complex embeddings).*

Proof:

> Since O_F free of rank $d = \deg(F/\mathbb{Q})$, we can write $O_F = \mathbb{Z}e_1 + ... + \mathbb{Z}e_d$ for $\{e_i\}$ a basis of $F/\mathbb{Q}$. By Theorem 1.15, the map $\Phi : F \to F \otimes_{\mathbb{Q}} \mathbb{R}$ sends $\{e_i\}$ to a basis of $F \otimes_{\mathbb{Q}} \mathbb{R}$.

□

Example 3.7. For d squarefree, the ring of integers of $\mathbb{Q}(\sqrt{d})$, which we characterized in Example 1.22, is a lattice in $\mathbb{R}^2$ (or $\mathbb{C}$). The two possibilities of this lattice are showin in Figure 3.1.

Definition 3.8. A subgroup $G \subset V$ of a vector space V is *discrete* if $\forall x \in G$, there exists $\epsilon > 0$ such that $B_\epsilon(x) \cap G = x$. It is *co-compact* if V/G is compact.

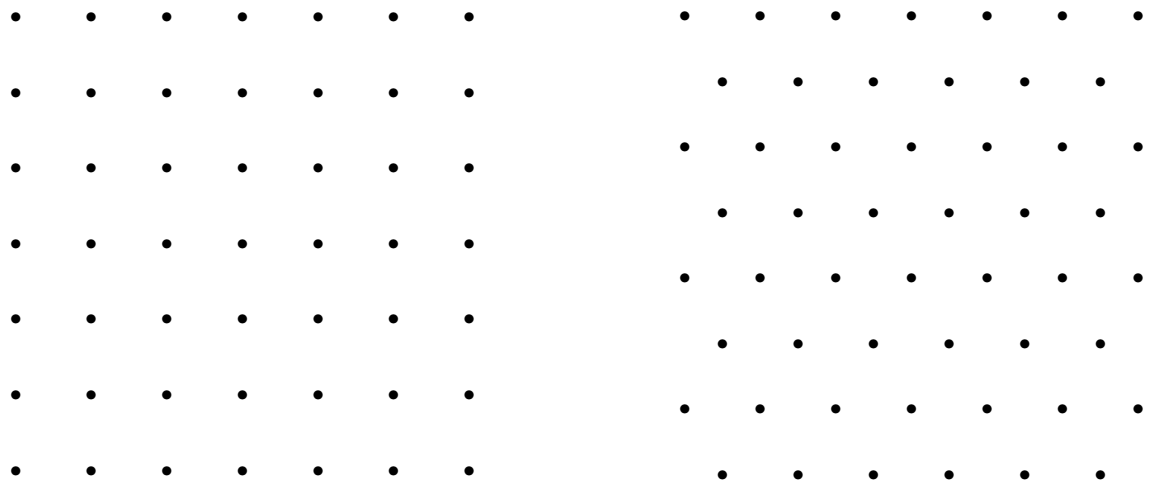

Figure 3.1: Lattices representing the ring of integers of $\mathbb{Q}(i)$ and $\mathbb{Q}(\sqrt{-3})$.

Proposition 3.9. *For V a finite dimensional vector space over $\mathbb{R}$, $L \subset V$ a subgroup, the following are equivalent:*

1. *L is a lattice.*
2. *L is a discrete subgroup and is co-compact.*
3. *L generates V over $\mathbb{R}$ (i.e. $V = L \otimes_{\mathbb{Z}} \mathbb{R}$) and for every bounded set $B \subset V$, the intersection $B \cap L$ is finite.*

Proof:

(1 $\Rightarrow$ 2) Since $L = \mathbb{Z}e_1 + ... + \mathbb{Z}e_n$, we have a surjection:

$$\left\{ \sum x_i e_i \mid 0 \leq x_i < 1 \right\} \to V/L$$

Since the left is compact, the right must also be compact.
FINISH

□

Corollary 3.10. *If I is a fractional ideal of a number field F, then I is a lattice in $F \otimes \mathbb{R}$.*

Proof:

There exists $\alpha \in F$ such that $\alpha I \subset O_F$, and we can assume $\alpha \in \mathbb{Z}^+$ (exercise). Let $m \in \mathbb{Z} \cap \alpha I$, which we have shown is nonempty in the proof of Proposition 2.9. Then we have:

$$\frac{m}{\alpha} O_F \subset I \subset \frac{1}{\alpha} O_F.$$

Since O_F is a lattice, so is qO_F for any $q \in \mathbb{Q}^*$. Therefore the $\frac{m}{\alpha} O_F$ and $\frac{1}{\alpha} O_F$ are both lattices. Then since $\frac{m}{\alpha}$ is co-compact, so is I. Additionally, since $\frac{1}{\alpha} O_F$ is discrete, so is I. Therefore I is a lattice.

□

Definition 3.11. The *fundamental domain* of a lattice $L \subset V$ is:

$$\mathcal{D}_L = \left\{ \sum x_i e_i \mid 0 \leq x_i < 1 \right\}$$

where $\{e_i\}$ are a basis of L.

Notice that the fundamental domain depends on the choice of basis. What doesn't depend on this choice, however, is the volume of the fundamental domain. This is known as the *covolume* of L. More precisely, if μ is a Haar measure on V, then $\operatorname{covol}(L) = \mu(\mathcal{D}_L)$. It is not hard to show that this is independent of basis choice of L, since the determinant of an integral change of basis matrix must be ± 1 (see Remark 1.23). Since we will only ever use lattices in $\mathbb{R}^n$, from here out the measure chosen is the Lebesgue measure.

Lemma 3.12. *Let M be a matrix with $\mathbb{R}$ coefficients and let $L = M(\mathbb{Z}^n)$. Then:*

1. *L is a lattie if and only if $\det(M) \neq 0$.*
2. *If L is a lattice, then $\operatorname{covol}(L) = |\det(M)|$.*

Proof:

1. L is a lattice if and only if M takes the standard basis to another basis of $\mathbb{R}^n$, which is true if and only if $\det(M) \neq 0$.
2. Let $\{e_i\}$ be the standard basis of $\mathbb{Z}^n$. Since $L = \mathbb{Z}M(e_1) + ... + \mathbb{Z}M(e_n)$, the fundamental domain of L is the image of the fundamental domain of $\mathbb{Z}^n$. Its volume is exactly the area spanned by the columns of M, which is the determinant.

□

Lemma 3.13. *If $S \subset \mathbb{R}^n$ and $L \subset \mathbb{R}^n$ is a lattice such that $\mu(S) > \operatorname{covol}(L)$, then there exist $s_1 \neq s_2 \in S$ such that $s_1 - s_2 \in L$.*

Proof:

Let $\pi : \mathbb{R}^n \to \mathbb{R}^n/L$ be the projection map. The volume of $\mathbb{R}^n/L$ as a set in $\mathbb{R}^n$ is $\operatorname{covol}(L)$. Note that if $\pi|_S$ was an injection, then it would be area preserving, and therefore $\mu(\pi(S)) = \mu(S) > \operatorname{covol}(L)$. But a set in $\mathbb{R}^n/L$ cannot have volume more than $\operatorname{covol}(L)$, so π must not have been injective on S. The result follows.

□

Definition 3.14. A set $C \subset \mathbb{R}^n$ is *symmetric* (with respect to the origin) if $x \in C \iff -x \in C$. It is *convex* if for all $x, y \in C$, $(t-1)x + ty \in C$ for $t \in [0, 1]$.

Theorem 3.15 (Minkowski)**.** *Let L be a lattice in $\mathbb{R}^n$ and let $C \subset \mathbb{R}^n$ be measurable which is convex and symmetric. Suppose also that $\mu(C) > 2^n \operatorname{covol}(L)$. Then C contains a nonzero element of L. Moreover, if C is compact, then $\mu(C) \geq 2^n \operatorname{covol}(L)$ is sufficient.*

Proof:

Let $S = \frac{1}{2}C$, so that $\mu(S) = 2^{-n}\mu(C) > \operatorname{covol}(L)$. By Lemma 3.13, there exist $s_1, s_2 \in S$ such that $s_1 - s_2 \in L$. Therefore there are $c_1, c_2 \in C$ such that $\frac{1}{2}(c_1 - c_2) \in L$. Since C is symmetric, $-c_2 \in C$, so by convexity we get $\frac{1}{2}c_1 + \frac{1}{2}(-c_2) \in C$. Therefore $\frac{1}{2}(c_1 - c_2) \in L \cap C$ and is nonzero.

In the case where C is compact, fix $\epsilon > 0$ and define:

$$C_\epsilon = \bigcup_{x \in C} B_\epsilon(x).$$

Since C is compact, we can refine this to a finite cover $C_\epsilon = \bigcup_i^N B_\epsilon(x_i)$. Then $\mu(C_\epsilon) \leq \mu(C) + \epsilon M$ for some M, so we can apply the previous case to C_ϵ. Letting ϵ be sufficiently small, we can ensure that the lattice point we found is contained in C.

□

3.2 Minkowski's Theorem ❖

This section we will state and prove the second Minkowski theorem, which is a stronger version of Proposition 2.23 with an explicit value of G depending on the lattice discriminant of F.

Proposition 3.16. *Let F be a number field, with r_1 the number of real embeddings and r_2 the number of complex embeddings. Let $\Phi : F \to F \otimes_{\mathbb{Q}} \mathbb{R}$ be the map defined in Theorem 1.15. Then:*

1. $\operatorname{covol}(O_F) = 2^{-r_2} |\operatorname{disc} F|^{1/2}$.
2. *For I a fractional ideal,* $\operatorname{covol}(I) = N(I) \operatorname{covol}(O_F)$.

Proof:

We identify $\mathbb{C}$ with $\mathbb{R}^2$ in the usual way. Let $\{\omega_1, ..., \omega_n\}$ be a basis of O_F and let $\phi_1, ..., \phi_{r_1}$ be the real embeddings of F and let $\phi_{r_1+1}, ..., \phi_{r_1+r_2}$ be distinct, non-conjugate embeddings of F. Define the matrix:

$$M = \begin{pmatrix} \phi_1(\omega_1) & \cdots & \phi_{r_1}(\omega_1) & \operatorname{Re}\phi_{r_1+1}(\omega_1) & \operatorname{Im}\phi_{r_1+1}(\omega_1) & \cdots & \operatorname{Im}\phi_{r_1+r_2}(\omega_1) \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ \phi_1(\omega_n) & \cdots & \phi_{r_1}(\omega_n) & \operatorname{Re}\phi_{r_1+1}(\omega_n) & \operatorname{Im}\phi_{r_1+1}(\omega_n) & \cdots & \operatorname{Im}\phi_{r_1+r_2}(\omega_n) \end{pmatrix}$$

Notice that $M(\mathbb{Z}^n) = \Phi(O_F)$. Applying Lemma 3.12, we get $\operatorname{covol}(O_F) = |\det(M)| = 2^{-r_2} |\det(\phi_i(\omega_j))| = 2^{-r_2} |\operatorname{disc} F|^{1/2}$. The last equality came from observation 1 about the discriminant in Section 1.1.

Now let I be a fractional ideal. Then there exists a nonzero integer m such that $mI \subset O_F$. Since O_F/mI is finite, mI is a subgroup of O_F whose rank is the same as that of O_F. Then there exists an n by n integer matrix A such that $A\omega$ is a basis for mI (here $\omega = (\omega_1, ..., \omega_n)$ is a basis for O_F). By Corollary 3.2, $[O_F : mI] = |\det(A)|$. However, we also know that $[O_F : mI] = N(mI) = m^n N(I)$. By Lemma 3.12, the covolume of mI is $|\det(AM)|$, where M is the same as above. Therefore:

$$\operatorname{covol}(mI) = |\det(AM)| = |\det(A)||\det(M)| = m^n N(I) \operatorname{covol}(O_F)$$

On the other hand, multiplying the lattice I by m has the effect of pulling out a factor m^n in its covolume, so $\operatorname{covol}(mI) = m^n \operatorname{covol}(I)$. Therefore we get $\operatorname{covol}(I) = N(I) \operatorname{covol}(O_F)$ as desired.

□

Lemma 3.17. *For $r_1, r_2 \in \mathbb{Z}^+$ and $R \geq 0$, define the set:*

$$W(r_1, r_2, R) = \left\{ (x_1, ..., x_{r_1}, y_1, ..., y_{r_2}) \in \mathbb{R}^{r_1} \times \mathbb{C}^{r_2} \;\middle|\; \sum_i^{r_1} |x_i| + 2 \sum_i^{r_2} |y_i| \leq R \right\}$$

Letting $n = r_1 + 2r_2$, then the volume of this set is $\mu(W(r_1, r_2, R)) = \frac{2^{r_1}}{n!} \left(\frac{\pi}{2}\right)^{r_2} R^n$.

Proof:

We will induct on n. There are two base cases: $r_1 = 1, r_2 = 0$ and $r_1 = 0, r_2 = 1$. In the former case, $W(1, 0, R) = \{x \in \mathbb{R} \mid |x| \leq R\}$ has volume $2R = \frac{2^1}{1!} \left(\frac{\pi}{2}\right)^0 R^1$. In the latter case, $W(0, 1, R) = \{y \in \mathbb{C} \mid 2|y| \leq R\}$ has volume $\pi(R/2)^2 = \frac{2^0}{2!} \left(\frac{\pi}{2}\right)^1 R^2$.

There are two inductive steps. The first is fixing r_2 and sending $r_1 \to r_1 + 1$. We are assuming the

formula is true for $n-1 = r_1 + 2r_2$. In this case the volume is:

$$\begin{aligned}
|W(r_1+1, r_2, R)| &= \int_{-R}^{R} W(r_1, r_2, R-|t|)dt \\
&= \int_{-R}^{0} W(r_1, r_2, R+t)dt + \int_{0}^{R} W(r_1, r_2, R-t)dt \\
&= 2^{r_1}\left(\frac{\pi}{2}\right)^{r_2} \frac{1}{(n-1)!}\left[\int_{-R}^{0}(R+t)^{n-1}dt + \int_{0}^{R}(R-t)^{n-1}dt\right] \\
&= \frac{2^{r_1+1}}{n!}\left(\frac{\pi}{2}\right)^{r_2} R^n.
\end{aligned}$$

The second case is $r_2 \mapsto r_2 + 1$ fixing r_1. Then:

$$\begin{aligned}
|W(r_1, r_2+1, R)| &= \int_{\{0\le|z|\le R/2\}} W(r_1, r_2, R-2|z|)dz \\
&= \frac{2^{r_1}}{(n-2)!}\left(\frac{\pi}{2}\right)^{r_2} \int_0^{R/2}\int_0^{2\pi} t(R-2t)^{n-1}d\theta dt \\
&= \frac{2^{r_1}}{(n-2)!}\left(\frac{\pi}{2}\right)^{r_2} \cdot 2\pi \int_0^{R/2} t(R-2t)^{n-2}dt \\
&= \frac{2^{r_1}}{(n-2)!}\left(\frac{\pi}{2}\right)^{r_2+1} \frac{R^n}{(n-1)n}
\end{aligned}$$

Therefore the formula also holds for n.

□

Theorem 3.18 (Minkowski)**.** *Let $F/\mathbb{Q}$ be a degree n number field, r_1 and r_2 be the number of real and non-conjugate embeddings of F into $\mathbb{C}$. Then every ideal I of O_F contains an element $x \in I$ such that:*

$$N(x) \le \frac{n!}{n^n}\left(\frac{4}{\pi}\right)^{r_2} |\operatorname{disc} F|^{1/2} N(I)$$

Proof:

Recall from Proposition 3.16 that $\operatorname{covol}(I) = 2^{-r_2}N(I)|\operatorname{disc} F|^{1/2}$. Let $X(R) := W(r_1, r_2, R)$, and choose R such that:

$$\mu(X(R)) = \frac{2^{r_1}}{n!}\left(\frac{\pi}{2}\right)^{r_2} R^n = 2^n \operatorname{covol}(I) \tag{3.2.1}$$

$$= 2^n 2^{-r_2} N(I)|\operatorname{disc} F|^{1/2} \tag{3.2.2}$$

Since $X(R)$ is symmetric, convex, and compact, by Theorem 3.15 there exists a nonzero $x \in I \cap X(R)$.[a] By virtue of being in $X(R)$, this element satisfies:

$$\sum_i^n |\phi_i(x)| \le R$$

Combining this with the inequality of geometric and arithmetic means, we have:

$$|N(x)| = \left|\prod_i \phi_i(x)\right| \le \left(\frac{\sum_i^n |\phi_i(x)|}{n}\right)^n \le \frac{R^n}{n^n}$$

Using equation 3.2.2, we get the desired inequality:

$$N(x) \leq \frac{n!}{n^n}\left(\frac{4}{\pi}\right)^{r_2} N(I)|\operatorname{disc} F|^{1/2}$$

[a]Were are slightly abusing notation and thinking of I as the lattice $\Phi(I) \subset \mathbb{R}^{r_1} \times \mathbb{C}^{r_2}$.

□

Remark 3.19. As n grows, the two constant terms in front of the discriminant decrease monotonically. As a result, this bound gets tigher as the degree of $F/\mathbb{Q}$ increases. The bounding coefficient of $N(I)$ is called the *Minkowski constant*.

Corollary 3.20. *Let $F/\mathbb{Q}$ be of degree n. Then:*

1. $|\operatorname{disc} F| \geq \left(\frac{n^n}{n!}\left(\frac{\pi}{2}\right)^{r_2}\right)^2$.
2. $|\operatorname{disc} F| \geq \frac{\pi^n}{4}$. *In particular, if* $|\operatorname{disc} F| = 1$, *then* $F = \mathbb{Q}$.
3. *Every ideal clas contains an ideal $I \subset O_F$ such that* $N(I) \leq \frac{n!}{n^n}\left(\frac{4}{\pi}\right)^{r_2}|\operatorname{disc} F|^{1/2}$.
4. $|\operatorname{Cl}(O_F)| < \infty$.

Proof:

We have shown that $3) \Rightarrow 4)$ in the proof of Theorem 2.18, so we will only prove the first three.

1. Let $I \subset O_F$ be any ideal. For any $\alpha \in I$, we must have $N(\alpha) \geq N(I)$. By Minkowski's theorem, there exists $x \in I$ such that $N(x) \geq \frac{n!}{n^n}\left(\frac{4}{\pi}\right)^{r_2}|\operatorname{disc} F|^{1/2}N(I)$. Combining these inequalities gives:

$$\frac{n!}{n^n}\left(\frac{4}{\pi}\right)^{r_2}|\operatorname{disc} F|^{1/2} \geq 1$$

   from which the desired inequality follows.

2. Since $n^n \geq 2^{n-1}n!$, we have:

$$\left(\frac{\pi}{4}\right)^{2r_2}\left(\frac{n^n}{n!}\right)^2 \geq 4^{n-1}\left(\frac{\pi}{4}\right)^{2r_2} \geq 4^{n-1}\left(\frac{\pi}{4}\right)^{2r_1}\left(\frac{\pi}{4}\right)^{r_1} = \frac{\pi^n}{4}$$

   Combining this with the first inequality, we find $|\operatorname{disc} F| \geq \frac{\pi^n}{4}$.

3. Let c be an ideal class, and let $I \subset O_F$ be an ideal representing c. Let $x \in I$ be such $N(x) \leq \frac{n!}{n^n}\left(\frac{4}{\pi}\right)^{r_2}|\operatorname{disc} F|^{1/2}N(I)$, by Minkowski's theorem. Then since $xO_F \subset I$, we have $xI^{-1} \subset O_F$ and so xI^{-1} is an ideal also representing c and:

$$N(xI^{-1}) = \frac{N(x)}{N(I)} \leq \frac{n!}{n^n}\left(\frac{4}{\pi}\right)^{r_2}|\operatorname{disc} F|^{1/2}.$$

□

Exercise 3.21. Show that $\operatorname{Cl}(O_F) = \{1\}$ for $F = \mathbb{Q}(\sqrt{14})$. In this case $\frac{n!}{n^n}\left(\frac{4}{\pi}\right)^{r_2}\sqrt{|\operatorname{disc} F|} \approx 3.74$. What ideals have norm less than or equal to 3? Clearly $N(I) = 1$ is trivial, so consider $N(I) = 2, 3$. Since 2 and 3 are prime numbers, I must be a prime ideal. We have shown that the norm of a prime ideal is p^n for some prime $p \in \mathbb{Z}$ lying over I and $n \geq 1$. In this case, our only choices are $p = 2, 3$ and $n = 1$. Therefore I lies over 2 or it lies over 3, which means I divides either (2) or (3). We note that (3) and (2) have prime factorizations (3) and $(4-\sqrt{14})(4+\sqrt{14})$, respectively.[2] In either case, all divisors of (2) and (3) are principal. Therefore I is principal, and so every ideal class can be represented by a principal ideal.

[2]The claim that (3) is prime we will prove later

Corollary 3.22. *For any positive integer Δ, there are finitely many number fields with discriminant less than or equal to Δ.*

Proof (sketch):

Use Minkowski's theorem to show that there exists $\alpha \in O_F$ such that $F = \mathbb{Q}(\alpha)$ and $|\phi_i(\alpha)| \leq 1 + \Delta$. Then the coefficients of the minimal polynomial of α, which are elementary symmetric polynomials in $\phi_i(\alpha)$, are bounded by a function of the Minkowski constant. There can be only finitely many such polynomials, hence finitely many such fields.

□

3.3 Dirichlet's Theorem ❖

The next important theorem is Dirichlet's Theorem. This characterizes the multiplicative group of O_F, which we'll denote by $O_F^\times$. Recall that an element $x \in O_F$ is called a root of unity if $x^n = 1$ for some n. We denote the roots of unity by μ_F. Our first step for proving this theorem will be showing that the roots of unity are the kernel of the map:

$$\text{Log} : O_F^\times \to \mathbb{R}^{r_1+r_2}$$

$$x \mapsto (\log|\phi_1(x)|, \log|\phi_2(x)|, ..., 2\log|\phi_{r_1+1}(x)|, ..., 2\log|\phi_{r_1+r_2}(x)|)$$

where $\{\phi_i\}_{i\leq r_1}$ are real embeddings the others are a choice of non-conjugate complex embeddings as usual. This is clearly a group homomorphism, so it makes sense to talk about its kernel.

Lemma 3.23. *Let F be a number field. If $\alpha \in O_F$ is nonzero such that $|\phi_i(\alpha)| \leq 1$ for all i, then α is a root of unity.*

Proof:

If $|\phi_i(\alpha)| \leq 1$, then $|\phi_i(\alpha^k)| \leq 1$ for all k. Let f_k be the characteristic polynomial of α^k over F:

$$f_k = \prod_{i=1}^{n}(x - \phi_i(\alpha^k))$$

We can expand this to something of the form $f_k = x^n + \sum a_j x^j$, where $\binom{|a_j|\leq n}{j\leq n^n}$ because $\phi_i(\alpha^k) \leq 1$. Since f_k is of degree n and its coefficients are bounded by n^n, the set of such f_k is finite. This means that the set of their roots is finite (i.e. $\{\phi_1(\alpha^k)\}_{k\in\mathbb{N}}$ is finite). Therefore $\phi_1(\alpha) = \phi_1(\alpha^N)$ for some N, which implies $\alpha - \alpha^N = 0 \Rightarrow \alpha^{N-1} = 1$. Therefore α is a root of unity.

□

Corollary 3.24. $\ker(\text{Log}) = \mu_F$.

Proof:

If $\alpha \in \mu_F$, then $\phi_i(\alpha^N) = \phi_i(1) = 1$. Therefore $|\phi_i(\alpha)| = 1$, which means $\log(|\phi_i(\alpha)|) = 0$, so $\alpha \in \ker(\text{Log})$. On the other hand, if $\log(|\phi_i(\alpha)|) = 0$, then it is a root of unity by the Lemma above.

□

Theorem 3.25 (Dirichlet)**.** *Let F be a number field and $O_F^\times$ be the multiplicative group of its ring of integers. Then $O_F^\times \cong \mu_F \oplus \mathbb{Z}^{r_1+r_2-1}$.*

This is saying that the torsion part of $O_F^\times$ is exactly μ_F and its free part has rank $r_1 + r_2 - 1$. It is clear that μ_F is the torsion part of $O_F^\times$ (by definition of torsion). Since $\text{im}(\text{Log}) \subset \mathbb{R}^{r_1+r_2}$, the Corollary shows that $\text{im}(\text{Log}) = \mathbb{Z}^m$ for some m. Therefore, to prove Dirichlet's theorem it suffices to show that $m = r_1 + r_2 - 1$.

Lemma 3.26. *Fix $1 \leq k \leq r_1 + r_2$, and $\alpha \in O_F$ nonzero. Then there exists $\beta \in O_F$ nonzero such that:*

$$\log|\phi_i(\beta)| < \log|\phi_i(\alpha)|$$

for all $i \neq k$ and $N(\beta) \leq \left(\frac{2}{\pi}\right)^{r_2} \sqrt{|\operatorname{disc} F|}$.

Proof:

This will be an application of Minkowski's theorem. Define the constant:

$$H = \frac{\left(\frac{2}{\pi}\right)^{r_2} \sqrt{|\operatorname{disc} F|}}{\prod_{i \neq k} \frac{|\phi_i(\alpha)|}{2}}$$

If $k \leq r_1$, define:

$$E = \left\{ z \in \mathbb{R}^{r_1} \times \mathbb{C}^{r_2} \;\middle|\; \begin{array}{c} |z_i| \leq |\phi_i(\alpha)|/2 \; \forall i \leq r_1, i \neq k \\ |z_i|^2 \leq |\phi_i(\alpha)|/2 \; \forall i > r_1, i \neq k \\ |z_k| \leq H \end{array} \right\}$$

otherwise, if $k > r_1$ define E by:

$$E = \left\{ z \in \mathbb{R}^{r_1} \times \mathbb{C}^{r_2} \;\middle|\; \begin{array}{c} |z_i| \leq |\phi_i(\alpha)|/2 \; \forall i \leq r_1, i \neq k \\ |z_i|^2 \leq |\phi_i(\alpha)|/2 \; \forall i > r_1, i \neq k \\ |z_k|^2 \leq H \end{array} \right\}$$

It is easy to see that in both cases E is compact, symmetric, and convex. Further, its volume is:

$$\mu(E) = \prod_{i \leq r_1, i \neq k} \frac{2|\phi_i(\alpha)|}{2} \cdot \prod_{r_1 < i, i \neq k} \frac{\pi|\phi_i(\alpha)|}{2} \cdot \delta \cdot \left(\frac{2}{\pi}\right)^{r_2} \frac{\sqrt{|\operatorname{disc} F|}}{\prod_{i \neq k} \frac{|\phi_i(\alpha)|}{2}}$$

where $\delta = 2$ if $k \leq r_1$ and $\delta = \pi$ if $k > r_1$. This simplifies to:

$$\begin{aligned} \mu(E) &= 2^{r_1} \pi^{r_2} \left(\frac{2}{\pi}\right)^{r_2} \sqrt{|\operatorname{disc} F|} \\ &= 2^n 2^{-r_2} \sqrt{|\operatorname{disc} F|} \\ &= 2^n \operatorname{covol}(O_F) \end{aligned}$$

Therefore by Minkowski's theorem E and O_F has a nonzero intersection, so let $\beta \in E \cap O_F$ be nonzero. Moreover, β manifestly satisfies the required conditions since it is contained in E.

□

Corollary 3.27. *If $1 \leq k \leq r_1 + r_2$, there exists a unit u_k such that $\log|\phi_i(u_k)| < 0$ for all $i \neq k$.*

Proof:

Pick $\alpha_0 \in O_F$ nonzero, and use the above lemma to construct a sequence $\alpha_0, \alpha_1, \alpha_2, \ldots$ such that $\log|\phi_i(\alpha_m)| < \log|\phi_i(\alpha_{m-1})|$ for all $i \neq k$ and $N(\alpha_m) \leq \left(\frac{2}{\pi}\right)^{r_2} \sqrt{|\operatorname{disc} F|}$. Since $\alpha_m O_F$ has norm bounded by a function independent of m, the set of ideals $\{\alpha_m O_F\}$ is finite (see the second part of our proof of Theorem 2.18). Therefore $\alpha_\ell O_F = \alpha_j O_F$ for integers $\ell > j$, which implies $\frac{\alpha_\ell}{\alpha_j}$ is a unit in O_F. By construction, $\log|\phi_i(\alpha_\ell/\alpha_j)| < 0$. Therefore $u_k = \alpha_\ell/\alpha_j$ is the desired unit.

□

Now we can prove Theorem 3.25. Observe that $N(\alpha) = 1$ for any $\alpha \in O_F^\times$, hence $\mathrm{im}(Log)$ is contained in a hyperplane:

$$\mathrm{im}(\mathrm{Log}) \subseteq \mathcal{H} = \left\{ (x_1, ..., x_{r_1+r_2}) \mid \sum_i x_i = 0 \right\}$$

Since $\dim(\mathcal{H}) = r_1 + r_2 - 1$, it suffices to show that $\mathrm{im}(\mathrm{Log})$ is a (full rank) lattice in $\mathcal{H}$.

Proposition 3.28. $\mathrm{Log}(O_F^\times)$ *is a lattice in* $\mathcal{H}$.

Proof:

We will show that $\mathrm{Log}(O_F^\times)$ satisfies condition 3) of Proposition 3.9, which is equivalent to it being a lattice. Let $B_c(0) \subset \mathbb{R}^{r_1+r_2}$ be the ball of radius c centered at 0. Then $\mathrm{Log}(x) \in H \cap B_c(0)$ satisfies $|\phi_i(x)| \leq e^c$ for all i. Since O_F is a lattice, the set of elements $x \in O_F$ satisfying $|\phi_i(x)| \leq e^c$ is finite, so therefore $\mathrm{Log}(O_F^\times) \cap B_c(0)$ must also be finite. Since every bounded set is contained in $B_c(0)$ for some c, we have $\mathrm{Log}(O_F^\times) \cap B$ being finite for any bounded set $B \subset \mathbb{R}^{r_1+r_2}$.

By the previous Corollary, there are units $\{u_k\} \in O_F^\times$ for all $1 \leq k \leq r_1 + r_2$ that are not in μ_F. Since their image under Log is contained in $\mathcal{H}$ and $\log|\phi_i(u_k)| < 0$, we must have $\log|\phi_k(u_k)| > 0$. Then the matrix $M = \log|\phi_i(u_j)|$ has positive entries on the diagonals and negative entries everywhere else. This property ensures that the rank of M is at least $r_1 + r_2 - 1$. Since the all ones vector is in the kernel of M, we conclude that the rank of M is $r_1 + r_2 - 1$. Therefore $\{u_k\}$ span $\mathcal{H}$ over $\mathbb{R}$.

□

4. Ramification and Decomposition of Primes

❖

Recall that, in working through Exercise 3.21, we claimed that (3) was prime in $O_{\mathbb{Q}(\sqrt{14})}$. In this section we will show why this was true in more generality. The problem we would like to answer is: given a prime number $p \in \mathbb{Z}$, how does $(p) = pO_F$ factorize? After we answer this, we will also be able to state some important facts about ramification of primes and generalize this discussion to extensions of number fields other than $\mathbb{Q}$.

Lemma 4.1. *Let $F/\mathbb{Q}$ be a number field and $\alpha \in O_F$ be such that $F = \mathbb{Q}(\alpha)$ and $[O_F : \mathbb{Z}[\alpha]] < \infty$. Then $\operatorname{disc}(p_{\min,\alpha}) = [O_F : \mathbb{Z}[\alpha]]^2 \operatorname{disc} F$.*

Proof:

Exercise. (*Hint:* Use observation 2. about the discriminant from §1.1)

□

Theorem 4.2 (Kummer's Lemma)**.** *Let $f(x) \in \mathbb{Z}[x]$ be irreducible and monic, and let α be a root of f in its splitting field. Form $F = \mathbb{Q}(\alpha)$ and let $p \in \mathbb{Z}$ be a prime such that $p \nmid [O_F : \mathbb{Z}[\alpha]]$. Suppose that $f(T) = h_1^{e_1}(T) \cdot \ldots \cdot h_r^{e_r}(T)$ in $\mathbb{F}_p(T)$, with $h_i(T)$ irreducible in $\mathbb{F}_p[T]$ and $h_i \neq h_j$. If $g_i(T) \in \mathbb{Z}[T]$ are such that $g_i \equiv h_i$ mod p, then $\mathfrak{p}_i = (g_i(\alpha), p)$ are distinct prime ideals of O_F and $p = \mathfrak{p}_1^{e_1} \cdot \ldots \cdot \mathfrak{p}_r^{e_r}$. Moreover, $N(\mathfrak{p}_i) = p^{\deg(h_i)}$.*

Proof:

First we claim that $\mathbb{Z}[\alpha]/(g_i(\alpha), p) \cong \mathbb{F}_p[T]/h_i(T) \cong \mathbb{F}_q$, where $q = p^{\deg(h_i)}$. This follows from the fact that $\mathbb{Z}[\alpha]/(g_i(\alpha), p) \cong \mathbb{Z}_p[\alpha]/g_i(\alpha) = \mathbb{Z}_p[\alpha]/h_i(\alpha)$. Through this isomorphism, we will show that $O_F/\mathfrak{p}_i \cong \mathbb{F}_q$, from which we will deduce the theorem.

Let $d = [O_F : \mathbb{Z}[\alpha]$; then $p \nmid d$ implies that there exist $a, b \in \mathbb{Z}$ such that $ap + bd = 1$. Define the map $m_{bd} : O_F/\mathfrak{p}_i \to \mathbb{Z}[\alpha]/(g_i(\alpha), p)$ by $x \mapsto bdx$. We must check a few things to verify that this is well-defined. First, $bdO_F \subset \mathbb{Z}[\alpha]$ because $\mathbb{Z}[\alpha]$ is an index d sublattice of O_F. Moreover, if $x \in p_i$ we can write it as $x = \beta g_i(\alpha) + \gamma p$ ($\beta, \gamma \in O_F$) and therefore:

$$bdx = bg_i(\alpha)(d\beta) + bp(d\gamma)$$

Since $d\beta, d\gamma \in \mathbb{Z}[\alpha]$, we have $bdx \in (g_i(\alpha), p)$. Finally, to verify that it is a homomorphism, keep in mind $bd - 1 = ap$ and consider:

$$\begin{aligned} m_{bd}(x) \cdot m_{bd}(y) - m_{bd}(xy) &= bdx \cdot bdy - bdxy \\ &= bdxy(bd - 1) \\ &= p \cdot (dxy)ab && (\in p\mathbb{Z}[\alpha]) \\ &= 0 \end{aligned}$$

Therefore m_{bd} is well-defined. To see that this map is injective, suppose $bdx \in (g_i(\alpha), p) \subset \mathbb{Z}[\alpha]$. Then:

$$x = (ap + bd)x = a(px) + bdx \in \mathfrak{p}_i$$

Similar reasoning shows that m_{bd} is surjective. Therefore it is an isomorphism and $O_F/\mathfrak{p}_i \cong \mathbb{Z}[\alpha]/(g_i(\alpha), p) \cong \mathbb{F}_q$. This means that $\mathfrak{p}_i$ is maximal, and therefore prime. and $N(\mathfrak{p}_i) = p^{\deg(h_i)}$.

Note that $\prod \mathfrak{p}_i^{e_i} = \prod (g_i(\alpha), p)^{e_i} \subseteq pO_F$ because $\prod g_i(\alpha) \equiv 0 \bmod p$. Additionally, since h_i are monic:

$$N\left(\prod \mathfrak{p}_i^{e_i}\right) = \prod N(\mathfrak{p}_i)^{e_i} = p^{\sum e_i \deg(h_i)} = p^{\deg(f)} = N(pO_F)$$

Therefore $pO_F = \prod \mathfrak{p}_i^{e_i}$. The final thing we must verify is that $\mathfrak{p}_i \neq \mathfrak{p}_j$ for $i \neq j$. Since h_i and h_j are coprime in $\mathbb{F}_p[T]$, there are $a(T), b(T) \in \mathbb{F}_p[T]$ such that $a(T)h_i(T) + b(T)h_j(T) = 1$ in $\mathbb{F}_p[T]$. Passing to

appropriate lifts to $\mathbb{Z}[x]$ and evaluating at α, we get:

$$a(x)g_i(x) + b(x)h_j(x) = 1 + pk \Rightarrow 1 = a(\alpha)g_i(\alpha) + b(\alpha)g_j(\alpha) - pk \in (g_i(\alpha), g_j(\alpha), p)$$

Note that $(g_i(\alpha), g_j(\alpha), p) = (\mathfrak{p}_i, \mathfrak{p}_j)$. Therefore $1 \in (\mathfrak{p}_i, \mathfrak{p}_j)$, so they are not the same ideal.

□

Definition 4.3. A prime $p \in \mathbb{Z}$ factoring as $pO_F = \prod \mathfrak{p}_i^{e_i}$ is called *ramified* (in F) if $e_i > 1$ for some i.

Corollary 4.4. *Let $F/\mathbb{Q}$ be a number field, let $p \in \mathbb{Z}$ be prime, and let $\alpha \in O_F$. Then if p ramifies in F, then either $p \mid [O_F : \mathbb{Z}[\alpha]]$ or $p \mid \operatorname{disc} F$.*

Proof:

Let $pO_F = \prod \mathfrak{p}_i^{e_i}$ be the factorization of p. By assumption, $e_i > 1$ for some i. Then by Kummer's Lemma, $f_{\min,\alpha}$ has a multiple root mod p, which is true if and only if p divides $\operatorname{disc}(f)$. By Lemma 4.1, we also have:

$$\operatorname{disc}(f) = [O_F : \mathbb{Z}[\alpha]]^2 \operatorname{disc} F$$

From here, we see that if $p \nmid [O_F : \mathbb{Z}[\alpha]]$, then $p \mid \operatorname{disc} F$ and vice versa.

□

Example 4.5. Returning to the case $F = \mathbb{Q}(\sqrt{14})$ and $O_F = \mathbb{Z}[\sqrt{14}]$, we can apply Kummer's lemma to the ideal (3) in O_F. The minimal polynomial of $\alpha = \sqrt{14}$ is $f(x) = x^2 - 14$, which is $T^2 + 1$ in $\mathbb{F}_3[T]$. It is easy to check that this is irreducible, so we $h_1(T) = f(T)$, which implies (3) is prime.

Example 4.6. Let α be a root of $f(x) = x^3 - x - 1$ and let $F = \mathbb{Q}(\alpha)$. The discriminant of f is 23, which is squarefree and so therefore $[O_F : \mathbb{Z}[\alpha]] = 1$. Consider the ideal $(5) \in O_F$. Since 5 does not divide $[O_F : \mathbb{Z}[\alpha]]$, we can apply Kummer's lemma as we did in the previous example. In this case $f(T) = (T-2)(T^2 + 2x - 2)$ is the irreducible representation of $f(T)$ in $\mathbb{F}_5[T]$. Therefore $(5) = \mathfrak{p}_1\mathfrak{p}_2$, where $\mathfrak{p}_1 = (\alpha - 2, 5)$ and $\mathfrak{p}_2 = (\alpha^2 + 2\alpha - 2, 5)$.

Proposition 4.7. *Let $p \in \mathbb{Z}$ be a prime, $f(x) \in \mathbb{Z}[x]$ which is Eisenstein with respect to p, and let α be a root of f. Then $p \nmid [O_F : Z[\alpha]]$, where $F = \mathbb{Q}(\alpha)$.*

Proof:

Write $f(x) = x^n + a_{n-1}x^{n-1} + \cdots + a_0$. We are assuming $p \mid a_i$ and $p^2 \nmid a_0$. Suppose that $p \mid [O_F : \mathbb{Z}[\alpha]]$. Consider the ideal:

$$I = \{h \in \mathbb{F}_p[x] \mid \exists g \in \mathbb{Z}[x] \text{ such that } h \equiv g \bmod p \text{ and } g(\alpha) \in pO_F\}$$

Since $\mathbb{F}_p[x]$ is a PID, we have $I = (h_0)$ for some $h_0 \in \mathbb{F}_p[x]$. Note that $x^n = f \bmod p$, so I is not empty (taking $g = f$). Since $p \mid [O_F : \mathbb{Z}[\alpha]]$, there exists $\beta \in O_F \backslash \mathbb{Z}[\alpha]$ such that $p\beta \in \mathbb{Z}[\alpha]$. Write $p\beta = \sum_{i=0}^{n-1} b_i\alpha^i$, which implies that $p \nmid b_i$ for some i. Therefore:

$$k(x) = \sum_{i=0}^{n-1} b_i x^i \notin p\mathbb{Z}[x]$$

Moreover, $\overline{k}(x) \neq 0$, $k(\alpha) = p\beta \in pO_F$. Therefore $\overline{k}(x) \in I$, which means $\deg(\overline{k}) \leq n - 1$. Since $x^n \in I$, this means $h_0 = x^m$ for $m \leq n - 1$. Thus $x^{n-1} \in I$ because it is an ideal. By the conditions on I, this means $\alpha^{n-1} \in pO_F$. Using the minimal polynomial of α:

$$-\alpha \underbrace{(\alpha^{n-1} + a_{n-1}\alpha^{n-2} + \cdots + a_1)}_{\in pO_F} = a_0$$

Therefore $a_0 \in p\alpha O_F \Rightarrow \frac{a_0}{p} \in \alpha O_F \Rightarrow (a_0/p)^{n-1} \in \alpha^{n-1} O_F \subseteq pO_F$. This implies $\frac{a_0}{p} \in p\mathbb{Z}$, which contradicts $p^2 \nmid a_0$.

□

Theorem 4.8. *Let F be a number field, $p \in \mathbb{Z}$ a prime. Then p ramifies in $F/\mathbb{Q}$ if and only if $p \mid \operatorname{disc} F$.*

Lemma 4.9. *Let A be a ring and $B_i \supset A$ be rings containing A (for $1 \leq i \leq n$). Suppose also that B_i are finitely generated free A modules. Write $B = B_1 \times ... \times B_n$. Then $\operatorname{disc} B/A = \prod_{i=1}^{n} \operatorname{disc} B_i/A$.*

Proof:

It is sufficient to prove for $n = 2$. Suppose that $\{x_1, ..., x_{d_1}\}$ is a basis of B_1/A, and $\{y_1, ..., y_{d_2}\}$ is a basis of B_2/A. Then $\{(x_i, 0), (0, y_j)\}$ are a basis of B/A. Then:

$$\operatorname{disc} B/A = \det \begin{pmatrix} \operatorname{tr}(m_{x_i x_j}) & 0 \\ 0 & \operatorname{tr}(m_{y_i y_j}) \end{pmatrix} = \det(\operatorname{tr}(m_{x_i x_j})) \cdot \det(\operatorname{tr}(m_{y_i y_j})) = \operatorname{disc} B_1/A \cdot \operatorname{disc} B_2/A$$

□

Similar reasoning can also prove:

Lemma 4.10. *If I is an ideal of a ring A, and B is a ring containing A that is free and finitely generated over A. Then:*

$$\operatorname{disc}(B/IB) = \operatorname{im}(\operatorname{disc}(B/A)) \in A/I$$

Where the discriminant of B/IB is over A/I.

Using these lemmata, we can prove the theorem:

Proof of Theorem 4.8:

Take $\omega_1, ..., \omega_n$ to be a basis of O_F. These reduce to a basis $\overline{\omega}_i$ of O_F/pO_F over $\mathbb{Z}/p\mathbb{Z}$. Write $p = \mathfrak{p}_1^{e_1} \cdots \mathfrak{p}_r^{e_r}$ for $\mathfrak{p}_i$ prime ideals. By the Chinese Remainder Theorem:

$$O_F/pO_F \cong O_F/\mathfrak{p}_1^{e_1} \times \cdots \times O_F/\mathfrak{p}_r^{e_r}$$

Let $\{\delta_{ik}\}$ be a basis of $O_F/\mathfrak{p}_i^{e_i}$ over $\mathbb{Z}/p\mathbb{Z}$, where $k = 1, ..., r$. Then $\operatorname{disc}(F/\mathbb{Q}) = \det(\operatorname{tr}(m_{\omega_i \omega_j}))$. Changing basis appropriately, this is $\operatorname{disc}(O_F/\mathbb{Z}) \equiv \operatorname{disc}(O_F/p/\mathbb{Z}/p\mathbb{Z})$ mod p. Then:

$$\begin{aligned} u \operatorname{disc}(F/\mathbb{Q}) &= \prod_{k=1}^{r} \det(\operatorname{tr}(m_{\delta_{ik}\delta_{jk}})) \\ &= \prod_{k=1}^{r} \operatorname{disc}(O_F/\mathfrak{p}_i^{e_i}/\mathbb{Z}/p\mathbb{Z}) \end{aligned}$$

Where u is a unit in $\mathbb{Z}/p\mathbb{Z}$. If $e_i = 1$, then $O_F/p = \mathbb{F}_q$ and $\operatorname{disc}(\mathbb{F}_q/\mathbb{F}_p) = \operatorname{disc}(x^{p^a} - x) \neq 0$ because $x^{p^a} - x$ has no multiple roots. Moreover, if for some i we have $e_i > 1$, we claim that $\operatorname{disc}(O_F/\mathfrak{p}^{e_i}/\mathbb{Z}/p\mathbb{Z}) = 0$. To see this, take $\pi \in \mathfrak{p}_i \setminus \mathfrak{p}_i^2$. Then $\pi \in O_F/\mathfrak{p}_i^{e_i}$ is nilpotent. Construct a basis of $O_F/\mathfrak{p}_i^{e_i}$ over $\mathbb{Z}/p\mathbb{Z}$ starting with π. Let $\{\pi, \gamma_1, ..., \gamma_f\}$ be this basis. Then $\pi\gamma_j$ is nilpotent, which implies $\operatorname{tr}(m_{\pi\gamma_j}) = 0$ for all j. Therefore $\operatorname{disc}(O_F/\mathfrak{p}_i^{e_i}/\mathbb{Z}/p\mathbb{Z}) = 0$ as claimed. Finally, this means $\operatorname{disc}(O_F/\mathbb{Z}) = \operatorname{disc}(F/\mathbb{Q}) = 0$ mod p.

□

Corollary 4.11. *Let F be a number field. Then there are only finitely many primes $p \in \mathbb{Z}$ that ramify in F.*

Corollary 4.12. *If F is a number field not equal to $\mathbb{Q}$, then there is a prime $p \in \mathbb{Z}$ that ramifies in F.*

4.1 Prime ideals in Galois Extensions ❖

In the above discussion, for any prime $p \in Z$ and number field F, we were able to characterize factorization $pO_F = \prod_i \mathfrak{p}_i^{e_i}$. In this case, if we define $f_i := [O_F/\mathfrak{p}_i : \mathbb{Z}/p\mathbb{Z}]$, then multiplicativity of the norm implies that:

$$N(pO_F) = p^{\sum e_i f_i} \Rightarrow \sum_i e_i f_i = [F : \mathbb{Q}]$$

We can strengthen this fact by considering the more general setting of field extensions $L \supset K \supset \mathbb{Q}$, where K, L are number fields. Fix a prime ideal $\mathfrak{p}$ of O_K. Then we can consider its saturation in $\mathfrak{p}O_L$ in O_L. As usual, it factorizes as $\mathfrak{p}O_L = \mathfrak{q}_1^{e_1}...\mathfrak{q}_r^{e_r}$ for primes $\mathfrak{q}_i \subset O_L$. If we define $f_i(L/K) := [O_L/\mathfrak{q}_i : O_K/\mathfrak{p}]$ (this is known as the *inertia degree* of $\mathfrak{q}_i$ over $\mathfrak{p}$), then the same reasoning shows that:

$$\sum_i e_i f_i = [L : K]$$

Remark 4.13. Just as in the case above where $K = \mathbb{Q}$, we say that $\mathfrak{p}$ is ramified when at least one of the exponents e_i is greater than 1.

Proposition 4.14. *If L/K is a Galois extension of number fields and $\mathfrak{p} \subset O_K$ is a prime ideal with factorization $\mathfrak{p}O_L = \prod_i^r \mathfrak{q}_i^{e_i}$, then $e_1 = e_2 = ... = e_r$ and $f_1(L/K) = f_2(L/K) = ... = f_r(L/K)$. In particular, $f \cdot r \cdot e = [L : K]$.*

Lemma 4.15. *For a Galois extension L/K and $\mathfrak{p} \subset O_K$ a prime ideal, the Galois group $G = \mathrm{Gal}(L/K)$ acts transitively on the set of prime ideals of O_L dividing $\mathfrak{p}O_L$.*

Proof:

Let $\{\mathfrak{q}_1, ..., \mathfrak{q}_r\}$ be the primes dividing $\mathfrak{p}O_L$ and suppose this is false; then without loss of generality $\mathfrak{q}_2 \neq g\mathfrak{q}_1$ for all $g \in G$. Then, since they are both prime, we have $\mathfrak{q}_2 + g\mathfrak{q}_1 = O_L$ for all g. Then we can write $1 = x_g + y_g$ for $x_g \in \mathfrak{q}_2, y_g \in g\mathfrak{q}_1$. Note that $x_g \equiv 0 \bmod \mathfrak{q}_2$ and $x_g \equiv 1 \bmod g\mathfrak{q}_1$ for all g. Then:

$$1 = \prod_{g \in G}(x_g + y_g) = x + \prod_{g \in G} y_g$$

where $x \equiv 0 \bmod \mathfrak{q}_2$ and $x \equiv 1 \bmod g\mathfrak{q}_1$ for all g (or equivalently $g^{-1}x \equiv 1 \bmod \mathfrak{q}_1$). Then taking the norm of x:

$$N_{L/K}(x) = \prod_{g \in G} g^{-1}x \Rightarrow N_{L/K}(x) \equiv 1 \bmod \mathfrak{q}_1$$

Since $x \in \mathfrak{q}_2$, the product $\prod_{g \in G} gx$ is also in $\mathfrak{q}_2$ because $\mathfrak{q}_2$ is an ideal. Therefore $N_{L/K}(x) \in \mathfrak{q}_2 \Rightarrow N_{L/K}(x) \in \mathfrak{q}_2 \cap O_K = \mathfrak{p}$. But we also showed that $N_{L/K}(x) \equiv 1 \bmod \mathfrak{q}_1$, which means $N_{L/K}(x) \equiv 1 \bmod \mathfrak{p}$, which is a contradiction.

□

Proof (Proposition 4.14):

Let $g \in \mathrm{Gal}(L/K)$ be such that $g(\mathfrak{q}_1) = \mathfrak{q}_2$. Then $g\mathfrak{p}O_L = \mathfrak{p}O_L$, which implies $g\mathfrak{q}_1^{e_1} g\mathfrak{q}_2^{e_2} \cdots g\mathfrak{q}_r^{e_r} = \mathfrak{q}_1^{e_1}\mathfrak{q}_2^{e_2} \cdots \mathfrak{q}_r^{e_r}$, and therefore $g\mathfrak{q}_1^{e_1} = \mathfrak{q}_2^{e_2}$ by unique factorization. Thus $e_1 = e_2$, and similarly for every other e_i.

Now let $f_1 = [O_L/\mathfrak{q}_1 : O_k/\mathfrak{p}]$ and $f_2 = [O_L/\mathfrak{q}_2 : O_k/\mathfrak{p}]$, and g as above. Then $g : O_L/\mathfrak{q}_1 \to O_L/\mathfrak{q}_2$ is a well-defined bijection. Further, $g|_{O_k/\mathfrak{p}}$ is the identity. Thus, if $\{v_1, ..., v_{f_1}\}$ is a basis of $O_L/\mathfrak{p}$ over $O_k/\mathfrak{p}$, then $\{gv_1, ..., gv_{f_1}\}$ is a basis for $O_L/\mathfrak{q}_2$ over $O_K/\mathfrak{p}$. Hence $f_1 = f_2$ and similarly for every other f_i.

□

4.1.1 Decomposition and Inertia Groups

In this section L/K is assumed to be a Galois extension of number fields with Galois group G. Fix $\mathfrak{p} \subset O_K$ and $\mathfrak{q} \subset O_L$ prime such that $\mathfrak{q}$ lies over $\mathfrak{p}$. Then let $\mathfrak{p}O_L = \prod_i^r \mathfrak{q}_i^e$ and $f = f_i(L/K)$.

Definition 4.16. The *decomposition group* associated to $\mathfrak{q}$ is:

$$D_\mathfrak{q} = \{g \in G \mid g\mathfrak{q} = \mathfrak{q}\} \subset G$$

A few properties about $D_\mathfrak{q}$ are immediate. One is that $|D_\mathfrak{q}| = ef$, by the orbit stabilizer theorem. The other is that $D_{g\mathfrak{q}} = gD_\mathfrak{q}g^{-1}$. If G is abelian, the latter shows that the decomposition group depends only on the choice of $\mathfrak{p}$ and not on the prime lying over it. Standard Galois theory says that $D_\mathfrak{q}$ uniquely corresponds to an intermediate field:

$$L^{D_\mathfrak{q}} = \{x \in L \mid g(x) = x \ \forall g \in D_\mathfrak{q}\}$$

where $L \supset L^{D_\mathfrak{q}} \supset K$. Denote $\mathfrak{q}_D := \mathfrak{q} \cap O_{L^{D_\mathfrak{q}}}$.

Lemma 4.17. *For $\mathfrak{q}_D, \mathfrak{p}$ and $\mathfrak{q}$ as above, $\mathfrak{q}_D O_L = \mathfrak{q}^{e_D}$ and $\mathfrak{p}O_{L^{D_\mathfrak{q}}} = \prod_{i\in I} g_i\mathfrak{q}_D$, where $e_D = e$ and I is an index set of G with $|I| = r$.*

Proof:

Since $\mathrm{Gal}(L/L^{D_\mathfrak{q}}) = D_\mathfrak{q}$, Lemma 4.15 says that the only prime ideal that can lie over $\mathfrak{q}_D$ is $\mathfrak{q}$. Therefore $\mathfrak{q}_D O_L = \mathfrak{q}^{e_D}$ for some e_D. Since every prime in O_L lying over $\mathfrak{p}$ is of the form $g\mathfrak{q}$ for some $g \in G$, every prime in $O_{L^{D_\mathfrak{q}}}$ lying over $\mathfrak{p}$ must be of the form $g\mathfrak{q} \cap O_{L^{D_\mathfrak{q}}} = g\mathfrak{q}_D$. This means that $pO_{L^{D_\mathfrak{q}}} = \prod_{i\in I} g\mathfrak{q}_D$ for some index set I. All we now have to verify is that $e_D = e$ and $|I| = r$.

Since the left cosets of $D_\mathfrak{q}$ in G act uniquely on $\mathfrak{q}$ and hence $\mathfrak{q}_D$, the size of I must be equal to the number of left cosets of $D_\mathfrak{q}$. Since $|D_\mathfrak{q}| = ef$ and $|G| = fre$, we have $|I| = r$. Moreover, notice that:

$$\mathfrak{p}O_L = (\mathfrak{p}O_{L^{D_\mathfrak{q}}})O_L = \prod_{i\in I} g_i(\mathfrak{q}_D O_L) = \prod_{i\in I} g_i\mathfrak{q}^{e_D} = \mathfrak{q}_1^{e_D} \cdots \mathfrak{q}_r^{e_D}$$

Where $\mathfrak{q}_i = g_i\mathfrak{q}$. By unique factorization we then have $e_D = e$.

□

Remark 4.18. The field tower picture in this discussion is:

$$\begin{array}{ccccc}
L & \supset & O_L & \supset & \mathfrak{q}_1, ..., \mathfrak{q}_r \\
{\scriptstyle ef}\big| & & \big| & & \big| \\
L^{D_\mathfrak{q}} & \supset & O_{L^{D_\mathfrak{q}}} & \supset & \mathfrak{q}_{D,1}, ..., \mathfrak{q}_{D,r} \\
{\scriptstyle r}\big| & & \big| & & \big| \\
K & \supset & O_K & \supset & \mathfrak{p}
\end{array}$$

Where $\mathfrak{q}_{D,i} = \mathfrak{q}_i \cap O_{L^{D_{\mathfrak{q}_i}}}$.

Exercise 4.19. Let $\mathfrak{q} \subset O_L$ be a fixed prime lying over $\mathfrak{p}$, denote $k_\mathfrak{q} := O_L/\mathfrak{q}$ and $k_\mathfrak{p} := O_K/\mathfrak{p}$. Show that:

$$O_K/\mathfrak{p} = O_{L^{D_\mathfrak{q}}}/\mathfrak{q}_D$$

(*Hint:* show that $f(L^{D_\mathfrak{q}}/K) = [O_{L^{D_\mathfrak{q}}}/\mathfrak{q}_D : O_K/\mathfrak{p}] = 1$).

Definition 4.20. The dimension $[k_\mathfrak{q} : k_\mathfrak{p}]$ is called the *inertia degree* of $\mathfrak{p}$ over $\mathfrak{q}$.

Note that $f = f(L/K)$ defined above is exactly the inertia degree. Since any $g \in D_\mathfrak{q}$ fixes $\mathfrak{q}$, it restricts to an automorphism of $k_\mathfrak{q}$ fixing $k_\mathfrak{p}$. Therefore we have a well-defined map:

$$\phi_\mathfrak{q} : D_\mathfrak{q} \to \mathrm{Gal}(k_\mathfrak{q}/k_\mathfrak{p})$$

$$g \mapsto g|_{k_\mathfrak{q}}$$

Definition 4.21. The kernel of this homomorphism is called the *inertia group* of $\mathfrak{q}$, denoted $I_\mathfrak{q}$. It is manifestly a normal subgroup of the decomposition group $D_\mathfrak{q}$.

Proposition 4.22. *The map $\phi_\mathfrak{q}$ is a surjection.*

Proof:

Pick $\theta \in O_L$ such that $k_\mathfrak{p}(\overline{\theta}) = k_\mathfrak{q}$, where $\overline{\theta}$ is the image of θ in $k_\mathfrak{q} = O_L/\mathfrak{q}$. Since $k_\mathfrak{p} = O_{L^{D_\mathfrak{q}}}/\mathfrak{q}_D$, we can assume $\theta \in O_{L^{D_\mathfrak{q}}}$. Let $F(x) \in O_{L^{D_\mathfrak{q}}}[x]$ be the minimal polynomial of θ and $\overline{f}(x) \in k_\mathfrak{p}[x]$ be the minimal polynomial of $\overline{\theta}$. Certainly $\overline{F}(\overline{\theta}) = 0$, where $\overline{F}(x)$ is the reduction of F mod $\mathfrak{q}_D$, so we have $\overline{f}(x) \mid \overline{F}(x)$.

Now let $\sigma \in \mathrm{Gal}(k_\mathfrak{q}/k_\mathfrak{p})$. Then σ is determined by where it sends $\overline{\theta}$, so let $\overline{\gamma} = \sigma(\overline{\theta})$. Since $\overline{\gamma}$ is also a root of $\overline{f}$, it must further be a root of $\overline{F}(x)$ since $\overline{f}(x)$ is a factor of $\overline{F}(x)$. Therefore one of the roots of $F(x)$ reduces to $\overline{\gamma}$, call it γ. The Galois group $D_\mathfrak{q}$ acts transitively on roots of any minimal polynomial, so there is some $\tau \in D_\mathfrak{q}$ such that $\tau(\theta) = \gamma$. Restricting τ to $k_\mathfrak{q}$ gives us an automorphism of $k_\mathfrak{q}$ that sends $\overline{\theta}$ to $\overline{\gamma}$, so $\phi_\mathfrak{q}(\tau) = \tau|_{k_\mathfrak{q}} = \sigma$.

□

Since $|D_\mathfrak{q}| = ef$ and $|\mathrm{Gal}(k_\mathfrak{q}/k_\mathfrak{p})| = f$, we have $|I_\mathfrak{q}| = e$. Since the order of the inertia group of $\mathfrak{q}$ has order e, the exponent of the factorization of $\mathfrak{p}O_L$, in a sense it measures how ramified $\mathfrak{p}$ is. In particular, if $\mathfrak{p}$ doesn't ramify, then $I_\mathfrak{q}$ is trivial and $\phi_\mathfrak{q}$ is an isomorphism.

4.1.2 The Abelian, Unramified Case

From here on, assume $\mathfrak{p}$ doesn't ramify and that L/K is an abelian Galois extension. If $|k_\mathfrak{q}| = p^m$, then there is a canonical generator of $\mathrm{Gal}(k_\mathfrak{q}/k_\mathfrak{q})$, the Frobenius automorphism $x \mapsto x^{p^m}$. We denote $(\mathfrak{q}, L/K) := \phi_\mathfrak{q}^{-1}(x \mapsto x^{p^m})$ to be the unique preimage of the Frobenius automorphism, which is known as the *Frobenius of L/K*. We also note that $(g\mathfrak{q}, L/K) = g^{-1}(\mathfrak{q}, L/K)g$ for any $g \in \mathrm{Gal}(L/K)$ because $g^{-1}D_\mathfrak{q}g = D_{g\mathfrak{q}}$ and:

$$\phi_{g\mathfrak{q}}(g^{-1}(\mathfrak{q}, L/K)g)(\overline{x}) = g^{-1}(\mathfrak{q}, L/K)\overline{gx} = g^{-1}\overline{(gx)^{p^m}} = \overline{g^{-1}gx^{p^m}} = \overline{x}^{p^m}$$

Since $\mathrm{Gal}(L/K)$ is abelian, we have $g^{-1}(\mathfrak{q}, L/K)g = (g\mathfrak{q}, L/K)$ for any g. This means that the Frobenius of L/K only depends on the lower prime $\mathfrak{p} \subset O_K$. Therefore we denote it by $(\mathfrak{p}, L/K)$, keeping in mind that $\mathfrak{p}$ is not a prime of O_L, but a prime of O_K.

Definition 4.23. Let $I = \prod_i \mathfrak{p}_i^{e_i}$ be a fractional ideal of O_K. Then the *Frobenius* of I is:

$$(I, L/K) := \prod_i (\mathfrak{p}_i, L/K)^{e_i} \in \mathrm{Gal}(L/K)$$

Proposition 4.24. *Let $L \supset K \supset \mathbb{Q}$ be finite field extensions such that $L/\mathbb{Q}$ is abelian and Galois. Fix $\mathfrak{q}$ a prime of O_L, $\mathfrak{p}$ a prime of O_K lying below $\mathfrak{q}$, and $p \in \mathbb{Z}$ a prime below $\mathfrak{p}$. Suppose that p is unramified in $L/\mathbb{Q}$. Then:*

1. *$(p, K/\mathbb{Q})$ is the image of $(p, L/\mathbb{Q})$ in $\mathrm{Gal}(K/\mathbb{Q}) = \mathrm{Gal}(L/\mathbb{Q})/\mathrm{Gal}(L/K)$.*
2. *$(p, L/\mathbb{Q})^{f(K/\mathbb{Q})} = (\mathfrak{p}, L/K)$ in $\mathrm{Gal}(L/\mathbb{Q})$.*

Proof:

If p is unramified, then so is $\mathfrak{p}$, so $\phi_\mathfrak{q}$ and $\phi_\mathfrak{p}$ are isomorphisms. Since $k_p = \mathbb{F}_p$, we have $\phi_\mathfrak{q}(p, L/\mathbb{Q}) = (x \mapsto x^p) \in \mathrm{Gal}(k_\mathfrak{q}/\mathbb{F}_p)$. Moreover, since $k_\mathfrak{p}$ is also characteristic p, we have $\phi_\mathfrak{p}(p, K/\mathbb{Q}) = (x \mapsto x^p) \in \mathrm{Gal}(k_\mathfrak{p}/\mathbb{F}_p)$. Therefore the resriction of $\phi_\mathfrak{q}(p, L/\mathbb{Q})$ to $k_\mathfrak{p}$ is $\phi_\mathfrak{p}(p, L/\mathbb{Q})$, and hence the restriction of $(p, L/\mathbb{Q})$ to K is $(p, K/\mathbb{Q})$. This proves the first claim.

Suppose $|k_\mathfrak{p}| = p^m$, from whence $m = f(K/\mathbb{Q})$. Then $\phi_\mathfrak{q}(\mathfrak{p}, L/K) = (x \mapsto x^{p^m})$ and $\phi_\mathfrak{q}(p, L/\mathbb{Q}) = (x \mapsto x^p)$. Thus $\phi_\mathfrak{q}(\mathfrak{p}, L/K) = \phi_\mathfrak{q}(p, L/\mathbb{Q})^m \Rightarrow (\mathfrak{p}, L/K) = (p, L/\mathbb{Q})^m$ as desired.

□

5. Local Fields

❖

All fields we have dealt with so far ebedded into Euclidean space and consequently we had many geometric tools at our disposal. In this section, we will define another class of field that doesn't obey standard Euclidean properties by imposing a different norm, called the p-adic norm. After commpleting with respect to this norm, we will be able to state similar theorems about ramification, decomposition of primes, and Galois groups.

Definition 5.1. A *norm* on a field F is a function $|\cdot| : F \to \mathbb{R}_{\geq 0}$ such that:

1. $|x| = 0$ if and only if $x = 0$.
2. $|xy| = |x||y|$.
3. $|x + y| \leq |x| + |y|$.

A norm is called *Archimedean* if $|x + y| \leq \max(|x|, |y|)$.

It follows from these properties that for any field norm $|\cdot|$, we have $|1| = 1$ and $|-x| = |x|$. The most trivial example of a norm on any field is $|0| = 0$ and $|x| = 1$ for all $x \in F$ nonzero. This is known as the trivial norm. From here out, any time we suppose the existence of a norm it is assumed to be nonrivial.

Example 5.2. Let $F = \mathbb{Q}$. A norm is detetermined by how it acts the primes in $\mathbb{Z} \subset \mathbb{Q}$ by the multiplicative property. For a prime $p \in \mathbb{Z}$, the p-adic norm $|\cdot|_p$ is defined by $|0| = 0$, $|p| = p^{-1}$, $|m| = 1$, where $(m, p) = 1$. For example, if $p = 5$, the norms of $5, 10$, and $\frac{1}{12}$ are $\frac{1}{5}$, $\frac{1}{5}$, and 1, respectively.

Lemma 5.3. *If a norm is non-Archimedean, then $|x + y| = \max(|x|, |y|)$ for all x, y satisfying $|x| \neq |y|$.*

Proof:

Assume without loss of generality $|x| < |y|$, and suppose $|x + y| < |y|$; then since $y = x + y - x$, we have $|y| \leq \max(|x+y|, |x|) < |y|$, a contradiction. Therefore $|x+y| \geq |y|$. We also have $|x+y| \leq \max(|x|, |y|) = |y|$ by the non-Archimedean property, so therefore $|x + y| = |y|$.

□

Remark 5.4. A norm is non-Archimedean if and only if $|z| \leq 1 \Rightarrow |z+1| \leq 1$, which can be seen by taking $z = x/y$ satisfying $|z| \leq 1$.

Lemma 5.5. *A norm is non-Archimedean if and only if $|m| \leq 1$ for all $m \in \mathbb{Z}$.*

Proof:

For one direction, we use the fact that $|1| = 1 \Rightarrow |m| = |1 + (m-1)| \leq 1$. For the other direction, assume $|m| \leq 1$ for all $m \in \mathbb{Z}$. Let $x \in F$ be such that $|x| \leq 1$; then for any positive integer n:

$$\begin{aligned} |(1+x)^n| &\leq |1 + nx + ... + nx^{n-1} + x^n| \\ &\leq |1| + |x| + ... + |x|^n && (|n| \leq 1) \\ &\leq 1 + n|x| \end{aligned}$$

Therefore $|1 + x| \leq (1 + n|x|)^{1/n}$ for all n. Taking $n \to \infty$ gives $|x + 1| \leq 1$, so this norm is non-Archimedean.

□

Any norm on a field induces a topology on that field by considering balls around each point in the usual way. We say that two norms are *equivalent* if they induce the same topology on F.

Lemma 5.6. *Let $|\cdot|_1$ and $|\cdot|_2$ be two norms on a field F. Then they are equivalent if and only if there exists $s > 0$ such that $|x|_2 = |x|_1^s$ for all $x \in F$.*

Proof:

One direction is clear, since any $|\cdot|_1$ ball of radius r is a $|\cdot|_2$ ball of radius r^s. Now assume $|\cdot|_1$ and $|\cdot|_2$ are equivalent. Fix $y \in F$ such that $|y|_1 > 1$, and let $x \in F$ be arbitrary. Let α be such that $|x|_1 = |y|_1^\alpha$. Then choosing $n/m > \alpha$ and $n/m \to \alpha$, we have

$$|x^m y^{-n}|_1 = |x|_1^m |y|_1^{-n} = |y|_1^{\alpha m/n} < 1$$

Moreover, $|x^m y^{-n}|_1 \to 1$ as $n/m \to \alpha$. Therefore we also have $|x^m y^{-n}|_2 < 1$ and $|x^m y^{-n}|_2 \to 0$ as $n/m \to \alpha$ by equivalence of norms. This shows that $|x|_2 = |y|_2^\alpha$. Let s be such that $|y|_2 = |y|_1^s$, which is independent of x. Thus $|x|_2 = |x|_1^s$ for s independent of x.

□

Theorem 5.7. *If $|\cdot|$ is a non-trivial, non-Archimedean norm on $\mathbb{Q}$, then it is equivalent to the p-adic norm $|\cdot|_p$ for some prime p.*

Proof:

Using Lemma 5.5, we have $|m| \leq 1$ for all $m \in \mathbb{Z}$. Since this norm is nontrivial, there exists a nonzero $q \in \mathbb{Q}$ with $|q| \neq 1$. Without loss of generality assume $|q| < 1$, which means there is some prime p such that $|p| < 1$. For any other prime ℓ, we have $|\ell| \geq 1$ because otherwise we could write:

$$1 = pk + \ell m \Rightarrow 1 \leq \max(|pk|, |\ell m|) < 1$$

So indeed $|\ell| \geq 1$. But $\ell \leq 1$ as well, so $|\ell| = 1$. Then pick α such that $|p|^\alpha = \frac{1}{p}$. Therefore $|\cdot|$ is equivalent to $|\cdot|_p$.

□

5.1 Defining Local Fields ❖

Our definition of local field will require a few definitions. As a warm-up, we will define the most canonical such field: the p-adic numbers and then proceed to define them more generally. For any prime p, we call $\mathbb{Q}_p$, the p-adic numbers, the completion of $\mathbb{Q}$ with respect to the p-adic metric $|\cdot|_p$ (i.e. equivalence classes of Cauchy sequences). Because we have chosen a non-archimedean norm, this completion has a very different structure compared to completions with respect to archimedean norms, like $\mathbb{R}$. As an exercise, the reader can show that $\mathbb{Q}_p$ has a power series representation:

$$\mathbb{Q}_p = \left\{ \sum_{i=v}^{\infty} x_i p^i \;\middle|\; x_i \in \{0, ..., p-1\}, v \in \mathbb{Z}, x_v \neq 0 \right\}$$

Similarly, the p-adic integers are defined as $\mathbb{Z}_p = \{x \in \mathbb{Q}_p \mid |x| \leq 1\}$. Under this definition, is is immediate that $\mathbb{Z} \subset \mathbb{Z}_p$ from Lemma 5.5. The power series representation of $\mathbb{Z}_p$ is:

$$\mathbb{Z}_p = \left\{ \sum_{i=0}^{\infty} x_i p^i \;\middle|\; x_i \in \{0, ..., p-1\}, x_0 \neq 0 \right\}$$

Remark 5.8. For the purposes of this course, the power series representations of $\mathbb{Q}_p$ and $\mathbb{Z}_p$ are not particularly important. However, we encourage the reader to investigate the arithmetic of p-adic integers using the series representations if only for self-edification.

Definition 5.9. If F is a field with a non-archimedean norm, then define $\mathcal{O}_F = \{x \in F \mid |x| \leq 1\}$ and $\mathfrak{m}_F = \{x \in F \mid |x| < 1\}$.

Note that our choice of notation for $\mathcal{O}_F$ suggests that perhaps it is the ring of integers of F (provided it is a finite extension). This is true in the case of finite extensions of $\mathbb{Q}_p$ and is an exercise in the next section.

Exercise 5.10. For $F = \mathbb{Q}_p$, show that $\mathfrak{m}_F = p\mathbb{Z}_p$.

Proposition 5.11. *$\mathcal{O}_F$ is a local ring with maximal ideal $\mathfrak{m}_F$.*

Proof:

First we claim that the units of $\mathcal{O}_F$ are exactly those of norm 1. Given a unit $x \in O_F$, we have $x^{-1} \in \mathcal{O}_F \Rightarrow |x|^{-1} \geq 1$. Combining this with $|x| \leq 1$ gives $|x| = 1$. Conversely, if $|x| = 1$, then clearly $x^{-1} \in F$ has norm 1 as well, so $x^{-1} \in O_F$. If we suppose $I \supsetneq \mathfrak{m}_F$ is an ideal, then I has an element of norm 1 and thus a unit, so $I = O_F$. Thus $\mathfrak{m}_F$ is maximal. Moreover, it is unique because if there were another maximal ideal, it couldn't contain units so must be contained in $\mathfrak{m}_F$, which means it wasn't maximal to start with.

□

We define the *residue field* of F to be $k_F = \mathcal{O}_F/\mathfrak{m}_F$. Since k_F is compact and discrete, it must be finite. From the proof of Theorem 5.7, there exists p such that $|p| < 1$, which means $p \in \mathfrak{m}_F$. This means k_F is a field extension of $\mathbb{Z}/p\mathbb{Z}$.

Definition 5.12. Given a field F, an $\mathbb{R}$*-valuation* (or simply a *valuation*) is a map $v : F^* \to \mathbb{R}$ such that $v(xy) = v(x) + v(y)$ and $v(x + y) \leq \min(v(x), v(y))$.

If F is a characteristic zero field with non-archimedean norm, then we can define a valuation on F by $|x| = p^{-v(x)}$, where p is the prime contained in $\mathfrak{m}_F$. We call this the induced valuation map on F.

Definition 5.13. A *local field* (also called a *p-adic* field) is a field F of characteristic zero complete with respect to a non-trivial non-archimedean norm $|\cdot|$ such that the image of the induced valuation map is descrete (i.e. $v(F^*) \subset \mathbb{R}$ is a lattice).

For a local field F, the lattice $v(F^*)$ has a minimal element v_0. For any $\pi \in F$ such that $v(\pi) = v_0$, we have that π generates $\mathfrak{m}_F$ because v_0 generates the group $v(F^*)$. Such an element is called a *uniformizer* of F.

Exercise 5.14. Show that any $x \in F^*$ can be written uniquely as $x = \pi_F^m u$ for $u \in \mathcal{O}_F^\times, \pi_F$ a uniformizer of F, and m an integer.

Lemma 5.15. *Let F be a field with non-archimedean norm and let $R \subset \mathcal{O}_F$ be a set of representatives of k_F and π_F be a uniformizer of F. Then every $x \in \mathcal{O}_F$ (and every $x \in F$) can be written uniquely as a convergent sum:*

$$x = \sum_{i=0}^{\infty} x_i \pi_F^i,\ x_i \in R$$

Proof:

It suffices to show that for any n, we can uniquely write:

$$x = \sum_{i=0}^{n-1} x_i \pi^i + \pi^n b \tag{1}$$

where $b \in O_F$ and $x_i \in R$. We prove this by induction on n. The base case $n = 0$ amounts to finding $x_0 \in R$ such that $x - x_0 \in \mathfrak{m}_F$. Clearly we can take x_0 to be x mod $\mathfrak{m}_F$, and it is unique because we are assuming R doesn't contain duplicate representatives. Now assume we can uniquely write x as in (1). Then once again we can find a unique $x_n \in R$ such that $x_n - b \in \mathfrak{m}_F$. This means:

$$x = \sum_{i=0}^{n-1} x_i \pi^i + \pi^n (x_n + c\pi) = \sum_{i=0}^{n} x_i \pi^i + \pi^{n+1} c$$

for some $c \in O_F$. The result then follows by induction.

□

If $f \in \mathbb{Z}[x]$ is a polynomial with an integer root, then $\overline{f} \in \mathbb{F}_p[x]$ has a root in $\mathbb{F}_p$. However, the converse is doesnt generally hold. In the case of a p-adic field, there are conditions under which the converse does hold. This is the content of Hensel's lemma, which we will state but not prove.

Lemma 5.16 (Hensel)**.** *Let F be a p-adic field, $f(x) \in \mathcal{O}_F[x]$, and $\alpha_0 \in \mathcal{O}_F$ such that $|f(\alpha_0)| < 1$ and $|f'(\alpha_0)| = 1$. Then there exists $\alpha \in \mathcal{O}_F$ such that $\pi_F \mid (\alpha - \alpha_0)$ and $f(\alpha) = 0$.*

5.2 Finite extensions of $\mathbb{Q}_p$ ❖

Now we specialize to the case of finite extensions of $\mathbb{Q}_p$. We will first show that these are local fields. From there, we will try to tell the same story that we did with Number Fields, giving an account of ramification and Galois extensions. As before, out starting point will be the ring of integers. Let F be a finite extension of $\mathbb{Q}_P$ and consider:

$$O_F = \{\alpha \in F \mid p_{\min,\alpha} \in \mathbb{Z}_p[x]\}$$

The reader can verify that this is a ring and that its field of fractions is F as we did in the Number Field case.

Lemma 5.17. *A Dedekind domain R with finitely many prime ideals is a PID.*

Proof:

It suffices to prove that any prime ideal is prinicpal, as the result will follow for any other ideal by unique factorization. Let $\mathfrak{p}_1, ..., \mathfrak{p}_n$ be the distinct set of primes of R. Pick $x \in \mathfrak{p}_1 \setminus \mathfrak{p}_1^2$ be such that $\exists c \in \mathfrak{p}_1 \cdot ... \cdot \mathfrak{p}_n$ with $x_1 + c = 1$. We can do this because $O_F = \mathfrak{p}_1 + \mathfrak{p}_2 \cdot ... \cdot \mathfrak{p}_n$. We note that $x_1 \notin \mathfrak{p}_i$ for $i > 1$ because otherwise $1 = x_1 + c \in \mathfrak{p}_i$. Thus $x_1 \in \mathfrak{p}_1$, which means the prime factorization of (x_1) is $\mathfrak{p}_1^r$. This means $x_1 \in \mathfrak{p}_1^r$, so $r = 1$ by assumption on x_1. Thus $\mathfrak{p}_1 = (x_1)$.

□

Lemma 5.18. *O_F is a Dedekind domain and a PID.*

Proof:

Since O_F is a finite $\mathbb{Z}_p$ module, any submodule (ideal) of O_F is also finitely generated, so O_F is Noetherian. Moreover, if $\alpha \in F$ is integral over O_F, then $O_F[\alpha]$ is a finite O_F module, and hece a finite $\mathbb{Z}_p$ module. Therefore $\mathbb{Z}_p[\alpha]$ is also finitely genrated, so by Proposition 1.17, $\alpha \in O_F$. Finally, if $I \subset O_F$ is a nonzero prime ideal, then $\mathbb{Z}_p \cap I$ is a prime ideal of $\mathbb{Z}_p$. The only such ideal is $p\mathbb{Z}_p$. We see that $[O_F/I : \mathbb{Z}_p/p\mathbb{Z}_p] < \infty$ because O_F is a finite $\mathbb{Z}_p$ module. Moreover, since $\mathbb{Z}_p/p\mathbb{Z}_p$ is finite, O_F/I must also be finite. Any finite integral domain is a field, so I is maximal. The proves that the Krull dimension of O_F is 1.

To show that O_F is a PID, we will show that it has finitely many prime ideals. Since all prime ideals lie over a prime of $\mathbb{Z}_p$, and there is only one prime ideal of $\mathbb{Z}_p$, the factorization $pO_F = \mathfrak{p}_1^{e_1} \cdot ... \cdot \mathfrak{p}_n^{e_n}$ gives us all possible ideals of O_F.

□

Exercise 5.19. If F is a finite extension of $\mathbb{Q}_p$, show that $O_F = \mathcal{O}_F$.

Remark 5.20. In light of the above exercise, we will drop the "$\mathcal{O}_F$" notation and just use O_F. We will also denote the maximal ideal of O_F as $\mathfrak{p}$ instead of $\mathfrak{m}_F$ to be consistent with previous sections. Just as before, we will also call $k_\mathfrak{p} = O_F/\mathfrak{p}$ to be the residue field.

Theorem 5.21. *If F is a finite extension of $\mathbb{Q}_p$, then there is a unique norm on F that extends the p-adic norm on $\mathbb{Q}_p$.*[3]

Proof:

[3] "Extends" here means $|\alpha v| = |\alpha|_p |v|$ for all $\alpha \in \mathbb{Q}_p, v \in F$

We will first prove uniqueness, assuming that it exists. Suppose $|\cdot|_1$ and $|\cdot|_2$ are two norms extending $\mathbb{Q}_p$. For every $v \in F$, let c_v denote the smallest real number such that $|v|_1 \leq c_v|v|_2$. If B_1 denotes the radius 1 ball around the origin (under $|\cdot|$) the function $|\cdot| : B_1 \to \mathbb{R}$ must achieve a maximum since B_1 is compact. Let c_1 be the maximum. Then for any $v \in F$, we claim that $|v|_1 \leq c_1|v|_2$. To see this, pick $\alpha \in \mathbb{Q}_p$ such that $|\alpha|_p = \frac{1}{|v|_1}$ (we can do this by completeness of $\mathbb{Q}_p$); then we have:

$$|v|_1 \leq c_1|v|_2 \iff |\alpha v|_1 \leq c_1|\alpha v|_2 \iff 1 \leq c_1|\alpha v|_2$$

The right most statement is true by construction of c_1. The same argument shows that there exists c_2 independent of v such that $|v|_2 \leq c_2|v|_1$. Therefore $|\cdot|_1$ and $|\cdot|_2$ induce the same topology on F, so by Lemma 5.6, $|\cdot|_1 = |\cdot|_2^\alpha$ for some α. But restricting to $\mathbb{Q}_p \subset F$ requires $\alpha = 1$ because both norms agree on $\mathbb{Q}_p$. Therefore $|\cdot|_1 = |\cdot|_2$.

To show existence, it suffices to define a norm on O_F. Fix $x \in O_F$ and let $(x) = \mathfrak{p}_1^{a_1} \cdots \mathfrak{p}_r^{a_r}$, where $\mathfrak{p}_i \subset O_F$ are all prime ideals of O_F. Then we can write $x = \mu x_1^{a_1} \cdots x_r^{a_r}$ for some unit μ. If we factor (p) as $\mathfrak{p}_1^{e_1} \cdots \mathfrak{p}_r^{e_r}$, then for every $j = 1, ..., r$, define the norm $|x|_j = p^{-a_j/e_j}$. We note that $|p|_j = p^{-1}$, so this norm extends the p-adic norm on $\mathbb{Q}_p$. The reader can check that it is non-archimedean.

□

Remark 5.22. If $L/\mathbb{Q}_p$ is Galois, then $|\sigma x| = |x|$ for all $x \in F, \sigma \in \mathrm{Gal}(F/\mathbb{Q}_p)$ because otherwise $|\sigma(\cdot)|$ would define another norm extending $|\cdot|_p$.

Lemma 5.23. *Let F be a complete field with respect to some norm $|\cdot|$ and V be a finite dimensional normed vector space over F with norm compatible with $|\cdot|$. Let $\{v_1, ..., v_n\}$ be a basis for V over F. Then the infinity norm $|x_1v_1 + ... + x_nv_n|_\infty := \max(\{|x_i|\})$ is equivalent to the norm on V. In particular, V is complete with respect to its norm.*

Proof:

It suffices to show that there exist c, c' such that $c|\cdot|_\infty \leq |\cdot| \leq c'|\cdot|_\infty$. The second constant c' is straightforward to find using the triangle inequality:

$$|x_1v_1 + ... + x_nv_n| \leq |x_1||v_1| + ... + |x_n||v_n| \leq (|v_1| + ... + |v_n|)|v|_\infty$$

So we take $c' = |v_1| + ... + |v_n|$. To find the other constant c, we proceed by induction on the dimension of V. For $n = 1$, $c = |v_1|$ suffices. Now for $\dim(V) = n$, consider the following subspaces:

$$V_j = Fv_1 + ... + \widehat{Fv_i} + ... + Fv_n$$

where the hat means omission. Our inductive hypothesis is that the infinity norm on V_j is equivalent to $|\cdot|$, which then means that V_j is complete with respect to this norm. Therefore V_j is closed in V, and hence so is $V_j + v_j$. Because $0 \notin V_j + v_j$, there exists $\epsilon_j > 0$ such that $|w_i + v_i| > \epsilon_j$ for all $w_j \in V_j$. Let $c = \min(\{\epsilon_i\})$, so that for every i we have $|w_i + v_i| > c$ for all $w_i \in V_i$. For any $v = x_1v_1 + ... + x_nv_n$, let r be the index such that $|v|_\infty = |x_r|$. Then:

$$|x_1v_1 + ... + x_nv_n| = |x_r| \left| v_r + \sum_{i \neq r} \frac{x_i}{x_r} v_i \right| > |v|_\infty c$$

□

Applying the Lemma above to F as a $\mathbb{Q}_p$ vector space, we obtain:

Corollary 5.24. *Any finite extension F of $\mathbb{Q}_p$ is complete with respect to the unique norm on F extending $|\cdot|_p$, and hence F is a local field.*

Proposition 5.25. *If $F/\mathbb{Q}_p$ is a finite extension, then $\exists \alpha \in O_F$ such that $O_F = \mathbb{Z}_p[\alpha]$.*

Proof:

Let $\overline{\alpha} \in O_F/\mathfrak{p}$ be such that $k_\mathfrak{q} = \mathbb{Z}/p\mathbb{Z}[\overline{\alpha}]$. If $|k_\mathfrak{q}| = p$, then $\overline{\alpha}^q - \overline{\alpha} = 0$. Then applying Hensel's lemma to $p(x) = x^q - x$, there exists $\alpha_0 \in O_F$ such that $p(\alpha_0) = 0$ and $\alpha_0 \equiv \overline{\alpha} \bmod \mathfrak{p}$. Let π_F be a uniformizer of $\mathfrak{p}$. Then we claim that $O_F = \mathbb{Z}_p[\alpha_0 + \pi_F]$. If $k_i = \alpha_0^i$, the collection $\{k_i\}$ for $i = 0, ..., q$ is a complete set of representatives of $k_\mathfrak{p}$. By Lemma 5.15, we have:

$$O_F = \sum_{i=0}^{\infty} k_i \pi_F^i \mathbb{Z}_p$$

We note that $(\alpha_0 + \pi_F)^{q^n} \to \alpha_0$ as $n \to \infty$. Since O_F is complete, this shows $\alpha_0 \in \mathbb{Z}_p[\alpha_0 + \pi_F]$. Therefore $\pi_F \in \mathbb{Z}_p[\alpha_0 + \pi_F]$ and so $O_F \subseteq \mathbb{Z}_p[\alpha_0 + \pi_F]$, which means $O_F = \mathbb{Z}_p[\alpha_0 + \pi_F]$.

□

5.2.1 Ramification in p-adic fields

Let $L \supset K \supset \mathbb{Q}_p$ be finite extensions. We have shown that each of these is a p-adic field whose ring of inters is a local ring. We denote $\mathfrak{p}_L \subset O_L$ and $\mathfrak{p}_K \subset O_K$ to be the associated maximal ideals, $e = e(F/K)$ to be the ramification index (so that $\mathfrak{p}_K O_L = \mathfrak{p}_L^e$), and $f = f(L/K) = [O_L/\mathfrak{p}_L : O_K/\mathfrak{p}_K]$ to be the inertia degree. Since the residue fields only depend on K and L, we denote them by $k_L = O_F/\mathfrak{p}_L$ and $k_K = O_K/\mathfrak{p}_K$. Since there is only one ideal lying over p in each number ring, we have $[L : K] = fe$.

Proposition 5.26. *Let $\beta = \{k_1, ..., k_f\} \subset O_L$ be a basis of $O_L/\mathfrak{p}_L$ over $O_K/\mathfrak{p}_K$, and let π_L be a uniformizer of $\mathfrak{p}_L$. Then:*

$$O_L = \bigoplus_{i=1}^{f} \bigoplus_{j=0}^{e-1} k_i \pi_L^j O_K$$

Proof:

Let $M = \sum_j \sum_i k_i \pi_L^j O_K$. We will show that $O_L = M$ and the remainder is left as an exercise. Fix $x \in O_L$. In $O_L/\mathfrak{p}_L$, we can write $\overline{x} = \sum_{i=0}^{f} k_i \overline{x}_i$ for $\overline{x}_i \in O_K/\mathfrak{p}_K$ since β is a basis. Now let $\{x_i\}$ be lifts of $\{\overline{x}_i\}$ and define $m_0 = \sum_i k_i x_i \in M$. Since x and m_0 represent the same coset in $O_L/\mathfrak{p}_L$, the difference $x - m_0$ is divisible by π_L. We then repeat this process on $\frac{x-m_0}{\pi_L}$ to get $m_1 \in M$ with $\frac{x-m_0}{\pi_l} - m_1 \in \pi_L O_L$. Repeating *ad infinitum*, we obtain a series $m_0 + \pi_L m_1 + \pi_L^2 m_2 + ... \in M$ which approximates x in the sense that:

$$\lim_{N \to \infty} \left| x - \sum_{i=1}^{N} m_i \pi_L^i \right| = 0$$

Since M is complete (as O_K is complete), $x \in M$.

□

Definition 5.27. A finite extension $K \subset L$ of p-adic fields is called *totally ramified* if $f(L/K) = 1$ and $e(L/K) = [L : K]$. It is called *unramified* if $e(L/K) = 1$ and $f(L/K) = [L : K]$.

Theorem 5.28. *With L and K as above, there exists a unique extension L_0 such that $K \subset L_0$ is unramified and $L_0 \subset L$ is totally ramified, from which it follows that $[L_0 : K] = f$ and $[L : L_0] = e$.*

Proof:

We will show existence, then uniqueness. Let $\alpha_0 \in k_L$ be such that $k_L = k_K(\alpha_0)$. If $p_{\min,\alpha_0}(x) \in k_K[x]$ is the minimal polynomial of α_0, we can pick a lift of this polynomial $p(x) \in O_K[x]$. Since $p_{\min,\alpha_0}$ has no multiple roots, we have $p'_{\min,\alpha_0}(\alpha_0) \neq 0$, and so we may apply Hensel's lemma to obtain $\alpha \in O_L$ such that $p(\alpha) = 0$. Let $L_0 = K(\alpha)$. Note that $[L_0 : K] \leq \deg(p)$. Since $[L_0 : K] = e(L_0/K) f(L_0/K)$, and $f(L_0/K) = [k_L : k_K] = \deg(p)$, we must have $[L_0 : K] = \deg(p)$. In particular, $e(L_0/K) = 1$, so L_0/K is

unramified. For any intermiediate extension, we always have:

$$f(L_0/K)f(L/L_0) = f(L/K)$$

Since we have shown $f(L_0/K) = \deg(p) = f(L/K)$, the above equation then implies that $f(L/L_0) = 1$. Therefore L/L_0 is totally ramified.

Now suppose that L_0' is another intermediate extension with L_0'/K unramified and L/L_0' totally ramified. We apply Hensel's lemma once again to the same polynomial p to get $\alpha' \in O_{L_0'}$ such that $\alpha' \equiv \alpha_0 \bmod \mathfrak{p}_K$ and $p(\alpha') = 0$. But the roots of p are distinct mod $\mathfrak{p}_K$, so we must have $\alpha = \alpha'$. This means $L_0 \subset L_0'$. But since both L_0 and L_0' are unramified over K, we have $[L_0 : K] = f(L/K) = [L_0' : K]$. This means $L_0 = L_0'$.

□

Remark 5.29. If $q = |k_K|$ and $f = f(L/K)$, then $|k_L| = q^f$ and $p_{\min,\alpha_0}$ as above divides $x^{q^f} - x$. Then we can pick our lift $p \in O_K[x]$ so that it divides $x^{q^f} - x$ and we can choose α_0 to be a generator of $k_L^* \cong C_{q^f}$. This implies that our Hensel lift α satisfies $x^{q^f} - x$ (i.e. it is a root of unity), and moreover it is primitive because α_0 was. Thus $L_0 = K(\alpha) = K(\zeta_{q^f-1})$ is a cyclotomic extension.

As a result of the above observation, we have the following corollary:

Corollary 5.30. *If $L/\mathbb{Q}_p$ is unramified, then $L = \mathbb{Q}_p(\zeta)$, where ζ is a primitive $p^f - 1$ root of unity. In particular, every unramified extension of $\mathbb{Q}_p$ is Galois.*

Another observation comes from analysing the case where L/K was totally ramified to begin with (i.e. $L_0 = K$). In this case, $f(L/K) = 1$, so $k_L = k_K$. Since there is a trivial basis $\beta = \{1\}$ of k_L over k_K, by Proposition 5.26 we have $O_L = \bigoplus_{j=0}^{e-1} \pi_L^j O_k$. Since $\pi_L^e = \pi_K$, this is equivalent to $O_L = O_K[\pi_L]$. Then:

$$\text{Frac}(O_L) = \text{Frac}(O_K[\pi_L]) \Rightarrow L = K(\pi_L)$$

The minimal polynomial of π_L is easily seen as $x^e - \pi_K$ because it satisfies Eisenstein's criterion. The converse is also true:

Proposition 5.31. *Let $L = K(\alpha)$ be an extension of local fields with $p_{\min,\alpha}(x)$ satisfying Eisenstiein's criterion. Then L/K is totally ramified.*

5.2.2 Krasner's Lemma

Krasner's Lemma provides a useful topological criterion for when one extension of a local field is contained in another. We will use it to show that there are only finitely many extensions of $\mathbb{Q}_p$ of a given degree.

Lemma 5.32 (Krasner)**.** *Let $F/\mathbb{Q}_p$ be a finite extension, and let $\alpha, \beta \in \overline{\mathbb{Q}}_p$, the algebraic closure of $\mathbb{Q}_p$. Suppose that $|\alpha - \beta| < |\alpha - \sigma(\alpha)|$ for all non-identity $\sigma \in \text{Gal}(\overline{\mathbb{Q}_p(\alpha)})$, where $\overline{\mathbb{Q}_p(\alpha)}$ is the Galois closure of $\mathbb{Q}_p(\alpha)$. Then $F(\alpha) \subset F(\beta)$.*

Proof:

Let K be a Galois extension of F containing α and β. We wish to show that any $\sigma \in \text{Gal}(K/F(\beta))$ fixes α. By the non-archimedean property, we have:

$$\begin{aligned} |\sigma(\alpha) - \alpha| &= |\sigma(\alpha) - \beta + \beta - \alpha| \\ &\leq \max(|\sigma(\alpha) - \beta|, |\beta - \alpha|) \\ &= \max(|\sigma(\alpha - \beta)|, |\beta - \alpha|) && (\sigma(\beta) = \beta) \\ &= |\beta - \alpha| \end{aligned}$$

This contradicts our hypothesis. Therefore $\sigma(\alpha) = \alpha$.

□

Definition 5.33. Let $F/\mathbb{Q}_p$ be finite and $P(x), Q(x) \in F[x]$ be degree n polynomials with coefficients $\{p_i\}, \{q_i\}$, respectivel. We define the difference of P and Q as $|P - Q| := \max(\{|p_i - q_i|\})$.

We will need a technical lemma in addition to Krasner's Lemma to prove that there are finitely many extensions of $\mathbb{Q}_p$ of a given degree.

Lemma 5.34. *Let $P(x) \in F[x]$ be monic of degree d with no double roots and fix $\epsilon > 0$. Then there exists $\delta > 0$ such that if $Q(x) \in F[x]$ is monic of degree d with $|P - Q| < \delta$, then for all roots α of P there is some root β of Q such that $|\alpha - \beta| < \epsilon$.*

The proof is left as an exercise.

Theorem 5.35. *For every $d \geq 1$, there are finitely many degree d extensions of $\mathbb{Q}_p$.*

Proof:

Let $F/\mathbb{Q}_p$ be finite. We will first show that there are finitely many totally ramified extensions K/F of degree e. By Proposition, these are in bijective correspondence to extensions $K = F(\alpha)$ with $p_{\min,\alpha}(x)$ Eisenstein. Write any such polynomial as:

$$p_{\mathbf{a}}(x) = x^e + \sum_0^{e-1} a_i x^i$$

where $\mathbf{a}$ denotes the coefficient vector (a_i). Eisentstein's criterion requires $\mathbf{a} \in \mathfrak{p}_F \times ... \times \mathfrak{p}_F \times (\mathfrak{p}_F \setminus \mathfrak{p}_F^2)$. Since $\mathfrak{p}_{\mathbf{a}}$ is irreducible over F, it has no multiple roots. Then there is some $\epsilon > 0$ such that $\epsilon < \min |\alpha - \alpha_i|$, where $\alpha_i \neq \alpha$ are the Galois conjugates of α. By Lemma 5.34, there exists δ such that for all $\mathbf{b} \in \mathfrak{p}_F \times ... \times \mathfrak{p}_F \times (\mathfrak{p}_F \setminus \mathfrak{p}_F^2)$, $p_{\mathbf{b}}(x)$ is Eisenstein of degree e and $|b_i - a_i| < \delta \Rightarrow |\beta_i - \alpha_i| < \epsilon$, where β_i are the roots of $p_{\mathbf{b}}(x)$. For any such $\mathbf{b}$, we then get $|\beta_i - \alpha_i| < \min |\alpha_i - \alpha_j|$, $i \neq j$. By Krasner's Lemma, we have $F(\alpha_i) \subset F(\beta_i)$. Both $F(\alpha_i), F(\beta_i)$ are of degree e, so $F(\alpha_i) = F(\beta_i)$. Thus, if any two coefficient vectors $\mathbf{a}, \mathbf{b}$ are close enough in the max norm, their roots induce the same extension K/F. The space $\mathfrak{p}_F \times ... \times \mathfrak{p}_F \times (\mathfrak{p}_F \setminus \mathfrak{p}_F^2)$ is compact, so there is a finite open cover by δ nieghborhoods. Coefficients in each open set of the cover induce the same extension, so there are finitely many totally ramified extensions K/F of degree e for fixed F.

Now we finish the proof by strong induction on the degree. The base case is is clear, as $\mathbb{Q}_p$ is the only extension of degree 1. Now consider all degree d extensions and assume that there are finitely many extensions of degree f for all $f < d$. Then there are finitely many extensions F_0 of degree less than d and finitely many totally ramified extensions of F_0 of degree $d/[F_0 : \mathbb{Q}_p]$. Since any extension of degree d is totally ramified over some F_0, there can only be finitely many such extensions.

□

5.3 The Approximation Theorem and Ostrowski's Theorem ❖

In this section we will prove a stronger version of Theorem 5.7. Before we do so, we will prove a related Theorem:

Theorem 5.36. *Let F be a field and $|\cdot|_1, ..., |\cdot|_n$ be non-equivalent norms on F. Then given $a_1, ..., a_n \in F$, there for all $\epsilon > 0$, there exists $x \in F$ such that $|x - a_i|_i < \epsilon$ for $i = 1, ..., n$.*

Proof:

Since $|\cdot|_1$ and $|\cdot|_n$ are not equivalent, there exist $\alpha_1, \alpha_2 \in F$ such that $|\alpha_1|_1 < 1, |\alpha_1|_n \geq 1, |\alpha_2|_1 \geq 1, |\alpha_2|_n < 1$. Let $y = \frac{\alpha_2}{\alpha_1}$. Notice that $|y|_1 > 1$ and $|y|_n < 1$.

We first claim that there exists $z_1 \in F$ such that $|z_1|_1 > 1, |z_1|_j < 1$ for all $j > 1$. To prove this claim, we proceed inductively on n. For $n = 2$, taking $z_1 = y$ as above suffices. Now assume that $n > 2$ and that z_1 exists for any collection of $n-1$ norms. If $|z_1|_n < 1$, we have already finished, so we consider the

two cases $|z_1|_n = 1$ and $|z_1|_n > 1$. In the former case, replace z_1 by $z_1^m y$ for some m. Then we note that:

$$\begin{aligned} |z_1^m y|_1 &= |z|_1^m |y|_1 > 1 \\ |z_1^m y|_n &= |y|_n < 1 \\ |z_j^m y|_j &= |z_1|_j^m |y|_j \qquad (1 < j < n) \end{aligned}$$

Since $|z_1|_j < 1$, choosing m sufficiently large makes the last equation be less than 1. Therefore in the case $|z_1|_n = 1$, we have proven the inductive hypothesis. Now consider the case $|z_1|_n > 1$. For every $m \in \mathbb{N}$, define $t_m = z_1^m/(1+z_1^m)$. Notice that $|t_m| \to 1$ under the norms $|\cdot|_1, |\cdot|_n$ and $|t_m| \to 0$ under the norms $|\cdot|_j$ for $1 < j < n$. Now we replace z_1 by $t_m y$. Then for m sufficiently large we have $|t_m y|_j < 1$. Moreover we have $|t_m y|_1 > 1$ and $|t_m y|_n < 1$ for m sufficiently large.

By this claim, for every j there exists $z_j \in F$ such that $|z_j|_j > 1$ and $|z_j|_i < 1$ for all $i \neq j$. Define $x_{j,m} = \frac{z_j^m}{1+z_j^m}$. Then for each j, $|x_{j,m}|_j \to 1$ and $|x_{j,m}|_i \to 0$ for $i \neq j$. Let $a = \max_{i,j} |a_i|_j$. Then there exists M such that $|x_{j,M} - 1|_j < \frac{\epsilon}{na}$ and $|x_{j,M}|_i < \frac{\epsilon}{na}$ for $i \neq j$. Let $x_j \equiv x_{j,M}$ and set:

$$x = \sum_{j=1}^{n} x_j a_j$$

Then:

$$\begin{aligned} |x - a_j|_j &= \left| \sum_{i \neq j} x_i a_i + a_j(x_j - 1) \right|_j \\ &\leq \sum_{i \neq j} |a_i|_j |a_i|_j + |a_j|_j |x_j - 1|_j \\ &< a \frac{\epsilon}{na} n \\ &= \epsilon \end{aligned}$$

□

Proposition 5.37. *Every nontrivial norm $|\cdot|$ on $\mathbb{Q}$ is equivalent to $|\cdot|_p$ or $|\cdot|_\infty$.*

Proof:

If $|\cdot|$ is non-archimedean, we have shown $|\cdot| \sim |\cdot|_p$ by Theorem 5.7, so suppose it is archimedean. Fix $m, n \in \mathbb{N}$ greater than 1 and write m in base n; that is, pick $0 \leq a_i \leq n-1$ such that:

$$m = a_0 + a_1 n + ... + a_r n^r$$

where $r = \text{floor}(\log m / \log n)$. Note that $|a_i| < n$ by the triangle inequality. Then:

$$\begin{aligned} |m| &\leq |a_0| + |a_1||n| + \cdots + |a_r||n|^r \\ &< n + n|n| + ... + n|n|^r \\ &\leq (r+1) n |n|^r \\ &\leq \left(1 + \frac{\log m}{\log n}\right) n |n|^{\log m / \log n} \end{aligned}$$

Doing the same for m^k for some $k \in \mathbb{N}$, we get:

$$|m^k| \leq \left(1 + k\frac{\log m}{\log n}\right) n|n|^{k \log m/ \log n}$$

$$\Rightarrow |m| \leq \left(1 + k\frac{\log m}{\log n}\right)^{1/k} n^{1/k}|n|^{\log m/ \log n}$$

Sending $k \to \infty$, we find $|m|^{1/\log m} \leq |n|^{1/\log n}$. Doing the same with m and n swapped, we obtain equality: $|m|^{1/\log m} = |n|^{1/\log n}$ for all $m, n \in \mathbb{N}$ greater than 1, and denote this constant by c. The multiplicative property of norms implies $|n| = c^{\log n}$ for all $n \in \mathbb{Q}$. Thus $|\cdot|$ is equivalent to $|\cdot|_\infty$.

□

A very similar proof idea can show:

Theorem 5.38 (Ostrowski)**.** *For F a normed number field with archimedean norm, there exists an injection $\sigma : F \to \mathbb{C}$ such that $|x| = |\sigma(x)|^s$ for some s independent of x.*

6. Analytic Methods in Number Fields

❖

A central player in the analytic theory of number fields is Dedekind's Zeta function, which we will define here for $\mathbb{Q}$ (in which case it is known as the Riemann Zeta function) and subsequently generalize for an arbitrary number field. For $s \in \mathbb{C}$, the Zeta function is defined by:

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

which converges for $\mathrm{Re}(s) > 1$. Then ζ defines a holomorphic function on the subset of the plane $\{z \in \mathbb{C} \mid \mathrm{Re}(z) > 1\}$. Notice that:

$$\frac{1}{2^s}\zeta(s) = \frac{1}{2^s} + \frac{1}{4^s} + \frac{1}{6^s} + ...$$
$$\Rightarrow \zeta(s)\left(1 - \frac{1}{2^s}\right) = \frac{1}{1^s} + \frac{1}{3^s} + \frac{1}{5^s} + \frac{1}{7^s} + \cdots$$

Repeating again:

$$\zeta(s)\left(1 - \frac{1}{2^s}\right)\left(1 - \frac{1}{3^s}\right) = \frac{1}{1^s} + \frac{1}{5^s} + \frac{1}{7^s} + \cdots$$

Doing so *ad infinitum* for all primes yields the identity:

$$\zeta(s) \prod_{p \text{ prime}} \left(1 - p^{-s}\right) = 1 \text{ or } \zeta(s) = \prod_{p \text{ prime}} \frac{1}{1 - p^{-s}}$$

Thus if we are able to study analytic properties of ζ, we can perhaps use those to understand the primes. To do this, we need to extend the domain of ζ, which can be done through the use of the Gamma function. Defined on $\{z \in \mathbb{C} \mid \mathrm{Re}(s) > 0\}$, the Gamma function is:

$$\Gamma(s) = \int_0^{\infty} t^{s-1} e^{-t} dt$$

Some properties of this function are that $\Gamma(s+1) = s\Gamma(s)$ and $\Gamma(s) \neq 0$ for any $\mathrm{Re}(s) > 0$. Substituting $t \mapsto nt$, we find:

$$\Gamma(s) = \int_0^{\infty} n^s t^{s-1} e^{-nt} dt \Rightarrow \frac{1}{n^s} = \frac{1}{\Gamma(s)} \int_0^{\infty} t^{s-1} e^{-nt} dt$$

Then:

$$\begin{aligned}
\zeta(s)\Gamma(s) &= \sum_{n=1}^{\infty} \int_0^{\infty} t^{s-1} e^{-nt} dt \\
&= \int_0^{\infty} \frac{t^{s-1}}{e^t - 1} dt \\
&= \int_0^1 \frac{t^{s-1}}{e^t - 1} dt + \int_1^{\infty} \frac{t^{s-1}}{e^t - 1} dt \\
&= \frac{1}{1-s} + \int_0^1 t^{s-1} \left(\frac{1}{e^t - 1} - \frac{1}{t}\right) dt + \int_1^{\infty} \frac{t^{s-1}}{e^t + 1} dt
\end{aligned}$$

Since the second and third terms converge for all $\mathrm{Re}(s) > 0$, this shows that $\zeta(s)$ can be extended to a meromorphic function on $\{z \in \mathbb{C} \mid \mathrm{Re}(s) > 0\}$ with a pole of order 1 at $s = 1$. In fact, using the Monodromy Theorem from complex analysis, it is possible to extend $\zeta(s)$ to the entire complex plane. Moreover, the only singularity is at $s = 1$ just as before and this extension satisfies:

$$\zeta(1-s) = 2(2\pi)^{-s}\Gamma(s)\cos(\pi s/2)\zeta(s)$$

Exercise 6.1. Use the above functional relation to show that $\zeta(0) = -\frac{1}{2}$ and $\zeta(-1) = -\frac{1}{12}$.

6.1 The Zeta Function of a Number Field ❖

Throughout this section, let $K/\mathbb{Q}$ be a number field. We will define a generalized Zeta function for this number field and find its domain of convergence, which will depend on $[K:\mathbb{Q}]$. We will also find the residue of this function at its $s=1$ pole, which will provide a formula for the class number. This residue formula will be important in the next section.

Definition 6.2. The *Dedekind Zeta function* associated to K is:

$$\zeta_K(s) = \sum_{I \subset O_K} \frac{1}{N(I)^s}$$

If we define $f(n)$ to be the number of idelas of O_K of norm n, the Zeta function takes the form of a Dirichlet series:

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

Moreover, there is an analogous product formula which can be derived in a similar manner to the Riemann Zeta function:

$$\zeta_K(s) = \prod_{\mathfrak{p} \subset O_K} \frac{1}{(1 - N(\mathfrak{p})^{-s})}$$

Definition 6.3. The *regulator* R_K of a number field $K/\mathbb{Q}$ is the determinant of a minor of $\mathrm{Log}(u_i)$, where $u_1, ..., u_{r_1+r_2-1}$ are generators of the free part of $O_K^\times$ and $\mathrm{Log}(-)$ is the map defined in Section 3.3. If $r_1 + r_2 = 1$ (i.e. $O_K^\times$ has no free components), then the convention is to let $R_K = 1$.

Exercise 6.4. Recall that we showed $\mathrm{Log}(O_F^\times)$ is a lattice in the subspace $\mathcal{H} = \{x \in \mathbb{R}^{r_1+r_2} \mid \sum_i x_i = 0\}$. Show that the covolume of this lattice is $\sqrt{r_1 + r_2} R_K$. (*Hint:* Take the unit orthogonal vector u to $\mathcal{H}$ and show that the covolume of $\mathrm{Log}(O_F^\times) + \mathbb{Z}u$ is the same as the covolume of $\mathrm{Log}(O_F^\times)$. Then compute the former by re-arranging columns in the matrix whose rows are u and $\mathrm{Log}(u_i)$, where $\{u_i\}$ are generators of the free part of $O_F^\times$.)

Theorem 6.5. *The Dedekind Zeta function is a meromorphic function on* $\{z \in \mathbb{C} \mid \mathrm{Re}(s) > 1 - 1/[K:\mathbb{Q}]\}$ *with a simple pole at* $s = 1$*. Moreover the residue at this singularity is:*

$$\mathrm{Res}(\zeta_K, 1) = \frac{2^{r_1}(2\pi)^{r_2} |\mathrm{Cl}(K)| R_k}{w_k \sqrt{|\operatorname{disc} K|}}$$

where r_1 *is the number of real embeddings of* K*,* r_2 *is the number of non-conjugate complex embeddings of* K*,* w_K *is the number of roots of unity in* K*, and* R_K *is a regulator of* K*.*

The main takeaway from this result is that if we know the residue of ζ_K at 1, we have a formula for the class number $h_K = |\mathrm{Cl}(K)|$. One of the tools we will use to prove Theorem 6.5 is the following lemma about Dirichlet series:

Lemma 6.6. *Let* $Z(s) = \sum_n^\infty a_n/n^s$ *be a Dirichlet series. Suppose that the partial sums of the coefficients satisfy:*

$$S(t) := \sum_{n=1}^{t} a_n = \mathcal{O}(t^r)$$

for $r > 0$*. Then* $Z(s)$*converges in* $\{s \in \mathbb{C} \mid \mathrm{Re}(s) > r\}$ *and is analytic in that region. Morevoer:*

$$\lim_{s \to 1}(s-1)f(s) = \lim_{t \to \infty} \frac{S(t)}{t}$$

Where the limit in s *is taken inside the region of convergence of* Z*.*

Proof:

We will prove the first part and relegate the second claim to Proposition 2.1 in Chapter 4 of Janusz's *Algebraic Number Fields*. For any integers $v > u$, we have:

$$\begin{aligned}\left|\sum_{n=u}^{v} \frac{a_n}{n^s}\right| &= \left|\sum_{n=u}^{v} \frac{S(n) - S(n-1)}{n^s}\right| \\ &= \left|\sum_{u}^{v} \frac{S(n)}{n^s} - \sum_{u-1}^{v-1} \frac{S(n)}{(n+1)^s}\right| \\ &= \left|\frac{S(v)}{v^s} - \frac{S(u-1)}{u^s} + \sum_{u}^{v-1} S(n)\left(\frac{1}{n^s} - \frac{1}{(n+1)^s}\right)\right| \\ &\leq \left|\frac{S(v)}{v^s}\right| + \left|\frac{S(u-1)}{u^s}\right| + \sum_{u}^{v-1} |S(n)| \left|\frac{1}{n^s} - \frac{1}{(n+1)^s}\right|\end{aligned}$$

Since $|S(t)| \leq at^r$ for some constant $a > 0$ and t sufficiently large, we have:

$$\left|\frac{S(v)}{v^s}\right| + \left|\frac{S(u-1)}{u^s}\right| \leq \frac{a}{v^{\sigma-r}} + \frac{a}{u^{\sigma-r}} < \frac{2a}{u^{\sigma-r}} \qquad (u \gg 1)$$

where $\sigma = \mathrm{Re}(s)$ and we used that $v > u$. Using the same approximation to $S(t)$ and identity:

$$\frac{1}{n^s} - \frac{1}{(n+1)^s} = s\int_n^{n+1} \frac{dt}{t^{s+1}}$$

we find:

$$\begin{aligned}\sum_{u}^{v-1} |S(n)| \left|\frac{1}{n^s} - \frac{1}{(n+1)^s}\right| &\leq \sum_{u}^{v-1} at^r |s| \left|\int_n^{n+1} \frac{dt}{t^{s+1}}\right| \\ &< |s|a \int_u^{\infty} \frac{t^r dt}{|t^{s+1}|} \\ &\leq \frac{|s|a}{(\sigma - r)u^{\sigma-r}}\end{aligned}$$

Putting everything together, we get:

$$\left|\sum_{u}^{v} \frac{a_n}{n^s}\right| < \frac{2a}{u^{\sigma-r}} + \frac{|s|a}{(\sigma-r)u^{\sigma-r}} = \frac{a}{u^{\sigma-r}}\left(2 + \frac{|s|}{\sigma - r}\right)$$

for u sufficiently large. For $\epsilon, \delta > 0$, define the region:

$$D_{\epsilon,\delta}(r) = \{s \in \mathbb{C} \mid \mathrm{Re}(s) > r + \delta, |\arg(s-r)| \leq \pi/2 - \epsilon\}$$

In this region, $\frac{r}{\mathrm{Re}\, s - r}$ is bounded by a positive constant M and $\frac{1}{\cos \arg(s-r)}$ is bounded by a positive constant N. Using this and the triangle inequality:

$$\frac{|s|}{\sigma - r} \leq \frac{|s-r| + r}{\sigma - r} = \frac{1}{\cos \arg(s-r)} + \frac{r}{\sigma - r} \leq M + N$$

Thus:

$$\left|\sum_{u}^{v} \frac{a_n}{n^s}\right| < \frac{a(2 + M + N)}{u^{\sigma-r}}$$

This can be made as small as we wish by taking u sufficiently large. Therefore $Z(s)$ is Cauchy convergent in (and also uniformly convergent) for all $s \in D_{\epsilon,\delta}(r)$ for any $\epsilon, \delta > 0$. Taking $\epsilon, \delta \to 0$, we have the desired result.

Since we showed that $Z(s)$ is a uniformly convergent series of analytic functions in some region for every s, it follows also that $Z(s)$ is analytic for all $\mathrm{Re}(s) > r$.

□

Returning to ζ_K, we can rewrite the sum by splitting it over ideal classes in a way that lends itself to the above lemma. Let $c \in \mathrm{Cl}(K)$ be an ideal class and define:

$$f_c(t) = \#\{I \in c \mid I \subset O_K, N(I) = t\}$$

$$i_c(t) = \sum_{n=1}^{t} f_c(n) = \#\{I \in c \mid I \subset O_K, N(I) \leq t\}$$

Then:

$$\zeta_K(s) = \sum_{n=1}^{\infty} \sum_{c \in \mathrm{Cl}(K)} f_c(n) \frac{1}{n^s}$$

Lemma 6.7. *For any ideal class $c \in \mathrm{Cl}(K)$, we have:*

$$i_c(t) = \frac{2^{r_1}(2\pi)^{r_2} R_K}{w_K \sqrt{|\operatorname{disc} K|}} t + \mathcal{O}(t^{1-1/[K:\mathbb{Q}]})$$

Proof:

We will prove this for $[K : \mathbb{Q}] \leq 2$, since the full proof is more compliated than we have time to do and isn't much different from the quadratic case. If $K = \mathbb{Q}$, the claim is trivial since $i_c(t) = \#\{n \in \mathbb{Z} \mid n \leq t\} = \lfloor t \rfloor = t + \mathcal{O}(1)$. Thus we consider K quadratic, which has two subcases (real and complex).

Case 1 (K complex). Write $K = \mathbb{Q}(\sqrt{-d})$ for $d > 0$ squarefree. Let $\sigma : K \to \mathbb{R}^2$ be the embedding $a + b\sqrt{-d} \mapsto (a, b\sqrt{d})$. Fix an ideal $I \subset O_K$ in the class of c and a fractional ideal J in the ideal class c^{-1}. Since $IJ = (\alpha)$ for some $\alpha \in J$ (which is unique up to a unit), notice we have the following correspondence (up to units):

$$\{I \in c \mid I \subset O_K, N(I) \leq t\} \leftrightarrow \{\alpha \in J \mid N(\alpha J^{-1}) \leq t\} = \{\alpha \in J \mid N(\alpha) \leq tN(J)\} \tag{6.1.1}$$

Applying σ to the set on the right, we get:

$$\{I \in c \mid I \subset O_K, N(I) \leq t\} \leftrightarrow \{(x, y) \in \sigma(J) \mid x^2 + y^2 \leq tN(J)\}$$

Geometrically, the RHS is $\sigma(J) \cap B_0(\sqrt{tN(J)})$. Recall from Corollary 3.10 that $\sigma(J)$ is a lattice, so this intersection is a finite set as we'd expect. Now let δ be the diameter of the fundamental domain of $\sigma(J)$. Then for any $r \geq 0$, we have:

$$\pi(r - \delta)^2 \leq \mathrm{covol}(J) \cdot \#\{\sigma(J) \cap B_0(r)\} \leq \pi(r + \delta)^2$$

and therefore for $r = \sqrt{tN(J)}$:

$$\begin{aligned} \#\{\sigma(J) \cap B_0(\sqrt{tN(J)})\} &= \frac{\pi t N(J)}{\mathrm{covol}(J)} + \mathcal{O}(\sqrt{t}) \\ &= \frac{2\pi}{\sqrt{|\operatorname{disc} K|}} t + \mathcal{O}(\sqrt{t}) \end{aligned}$$

where we used Proposition 3.16 to compute $\mathrm{covol}(J)$. Finally, we notice that:

$$i_c(t) = \frac{1}{w_K} \cdot \#\{\sigma(J) \cap B_0(\sqrt{tN(J)})\}$$

which comes from accounting for units in K in the correspondence in (6.1.1). Since $r_1 = 0$, $r_2 = 1$, and $R_K = 1$, this proves the formula for a complex quadratic field.

Case 2 (K real). Write $K = \mathbb{Q}(\sqrt{d})$ for $d > 0$ squarefree. Let $\sigma : K \to \mathbb{R}^2$ be $a+b\sqrt{d} \mapsto (a+b\sqrt{d}, a-b\sqrt{d})$ (the two real embeddings of K). Note that $\sigma(\alpha) = (x, y)$ satisfies $N(\alpha) = xy$ by Proposition 1.10. Since $r_1+r_2 = 0+2 = 2$, the rank of $O_K^\times$ is 1; so let u_K be a positive generator (this is called a *fundamental unit*). We will use the same correspondence (6.1.1), but we have to account for more than just roots of unity, since $O_F^\times$ also contains the span of u_K. To account for this, we note that multiplying α by a sufficiently high power of u_K gives the inequality $1 < \left|\frac{y}{x}\right| < u_K^2$, where $(x, y) = \sigma(\alpha)$.[a] We then want to count pairs $(x, y) \in \sigma(J)$ that satisfy this inequality and $|xy| \leq tN(J)$; in other words, we want to compute:

$$\#\left\{(x, y) \in \sigma(J) \;\middle|\; |xy| \leq tN(J), 1 < \left|\frac{y}{x}\right| < u_K^2\right\} = w_K i_c(t)$$

Chapter 6 of Marcus's *Number Fields* computes the area of the region $R = \{(x, y) \in \mathbb{R}^2 \mid |xy| \leq A, 1 < |y/x| < B^2\}$ to be $4A\log(B)$. Therefore, by intersecting this region with the lattice $\sigma(J)$, we get:

$$\begin{aligned} i_c(t) &= \frac{1}{w_K} \frac{4tN(J)\log(u_K)}{\mathrm{covol}(J)} + \mathcal{O}(\sqrt{t}) \\ &= \frac{4\log(u_K)}{w_K\sqrt{|\operatorname{disc} K|}} t + \mathcal{O}(\sqrt{t}) \end{aligned}$$

This proves the formula, since $r_2 = 2$, $r_1 = 0$, and $R_K = \log(u_K)$.

[a] We are thinking of u_K as a real number in this inequality

□

Exercise 6.8. Verify Marcus's formula for the area of the region R that we used in the above proof.

For a complete proof of Lemma 6.7 for a general number field, consult Chapter 6 of Marcus. Finally, we can use this estimate on the size of $i_c(t)$ to prove Theorem 6.5:

Proof (of Theorem 6.5):

Recall we wrote:

$$\zeta_K(s) = \sum_{n=1}^{\infty} \sum_{c \in \mathrm{Cl}(K)} \frac{f_c(n)}{n^s}$$

Letting $a_n = \sum_{c \in \mathrm{Cl}(K)} f_c(n)$, the partial sums of a_n satisfy:

$$\begin{aligned} S(t) &= \sum_{n=1}^{t} \sum_{c \in \mathrm{Cl}(K)} f_c(n) \\ &= \sum_{c \in \mathrm{Cl}(K)} i_c(t) \\ &= \frac{2^{r_1}(2\pi)^{r_2} R_K}{w_K\sqrt{|\operatorname{disc} K|}} \cdot |\mathrm{Cl}(K)| \cdot t + \mathcal{O}(t^{1-1/[K:\mathbb{Q}]}) \end{aligned}$$

Therefore by Lemma 6.6, $\zeta_K(s)$ converges for all $\mathrm{Re}(s) > 1 - 1/[K : \mathbb{Q}]$ and is analytic in this domain.

Moreover, since the pole at $s = 1$ is simple, the residue at this point is:

$$\begin{aligned}\operatorname{Res}(\zeta_K, 1) &= \lim_{s\to 1}(s-1)\zeta_K(s) \\ &= \lim_{t\to\infty}\frac{S(t)}{t} \\ &= \frac{2^{r_1}(2\pi)^{r_2}R_K|\operatorname{Cl}(K)|}{w_K\sqrt{|\operatorname{disc} K|}}\end{aligned}$$

□

Just as with the Riemann Zeta function, the Dedekind Zeta function can be extended to a meromorphic function on all of $\mathbb{C}$ using the theory of analytic continuations.

6.2 Dirichlet Characters and L-Functions ❖

Using the theory of Dirichlet characters and L-functions will allow us to derive an explicit formula for the class number of a quadratic field.

Definition 6.9. A *Dirichlet character* is a multiplicative homomorphism $\chi : (\mathbb{Z}/n\mathbb{Z})^\times \to \mathbb{C}^*$.

Every Dirichlet character has an associated conjugate character $\overline{\chi}$, which is given by post-composing with complex cojugation. Given Dirichlet a character, we can consider the set of integers $m \mid n$ such that we have a factorization:

$$\chi : (\mathbb{Z}/n\mathbb{Z})^\times \longrightarrow (\mathbb{Z}/m\mathbb{Z})^\times \longrightarrow \mathbb{C}^*$$

where the first map is reduction mod m. The smallest such m where such a factorization occurs is called the *conductor* of χ, denoted f_χ. Given a character $\chi : (\mathbb{Z}/f_\chi\mathbb{Z})^\times \to \mathbb{C}^*$, we can extend it to a function $\mathbb{Z} \to \mathbb{C}$ by $x \mapsto \chi(x \mod f_\chi)$ if $(x, f_\chi) = 1$ and $x \mapsto 0$ otherwise. We will also call this function χ.

Definition 6.10. Given a character χ and an integer $m \in \mathbb{Z}$, the *Gauss sum* associated to χ and m is:

$$G(m, \chi) = \sum_{a\in\mathbb{Z}/f_\chi\mathbb{Z}} \chi(a)\xi_{f_\chi}^{am}$$

where ζ_{f_χ} is a primitive f_χ-root of unity. We will usually drop the "$a \in \mathbb{Z}/f_\chi\mathbb{Z}$" subscript unless it is necessary.

Proposition 6.11. *Given $m \in \mathbb{Z}$ and a character χ, the Gauss sum satisfies:*

1. $G(m, \chi) = \overline{\chi}(m)G(1, \chi)$.
2. $G(1, \chi)G(1, \overline{\chi}) = \overline{\chi}(-1)f_\chi$.

Proof:

To prove the first part, if $(m, f\chi) = 1$, then m has an inverse mod f_χ and:

$$\begin{aligned}G(m, \chi) &= \sum_a \chi(a)\zeta^{am} \\ &= \sum_a \chi(am^{-1})\zeta^a \\ &= \chi(m)^{-1}\sum_a \chi(a)\zeta^a \\ &= \overline{\chi}(m)G(1, \chi)\end{aligned}$$

Where we used that $\chi(m)^{-1} = \overline{\chi}(m)$ because any nonzero $\chi(x)$ has complex norm 1. On the other hand,

if $(m, f_\chi) = d > 1$, then $\chi(m) = 0$, so we wish to show that $G(m, \chi) = 0$.

$$\begin{aligned} G(m,\chi) &= \sum_a \chi(a)\zeta_{f_\chi}^{am} = \sum_a \chi(a)\zeta_{f_\chi/d}^{am/d} \\ &= \sum_a \zeta_{f_\chi/d}^{bm/d} \sum_{\substack{b\equiv a \\ \bmod f_\chi/d}} \chi(b) \end{aligned}$$

We will show that the inner sum is zero. We will first show that it is zero for $a = 1$. For all $c \equiv 1 \bmod f_\chi/d$, we have:

$$\sum_{\substack{b\equiv 1 \\ \bmod f_\chi/d}} \chi(b) = \sum_{\text{“ ”}} \chi(bc) = \chi(c) \sum_{\text{“ ”}} \chi(b)$$

We note that if $\chi(c) = 1$ for all $c \equiv 1 \bmod f_\chi/d$, then the conductor of χ would be f_χ/d, which is false. Therefore $\chi(c) \neq 1$ for any $c \equiv 1 \bmod f_\chi/d$. In particular, this forces the above sum to be zero. Moreover for any a:

$$\sum_{\substack{b\equiv a \\ \bmod f_\chi/d}} \chi(b) = \sum_{\substack{a'\equiv 1 \\ \bmod f_\chi/d}} \chi(ba') = \chi(b)\cdot 0 = 0$$

Therefore $G(m, \chi) = 0$. This completes the proof of the first statement.

To prove the second, we use the identity that we just established to write:

$$\begin{aligned} G(1,\overline{\chi})G(1,\chi) &= G(1,\overline{\chi}) \sum_b \chi(b)\zeta^b = \sum_b \chi(b) G(1,\overline{\chi})\zeta^b \\ &= \sum_b G(b,\overline{\chi})\zeta^b = \sum_a \sum_b \overline{\chi}(a)\zeta^{ab}\zeta^b \\ &= \sum_a \overline{\chi}(a) \sum_b \zeta^{(a+1)b} \end{aligned}$$

We claim that the only contributions to the sum over a happen when $a+1 \equiv 0 \bmod f_\chi$. If $(a+1, f_\chi) = 1$, then:

$$\sum_b \zeta^{(a+1)b} = \sum_b \zeta^b = 0$$

because the sum of all f_χ powers of a f_χ root of unity is zero. The same reasoning holds for $(a+1, f_\chi) = d \neq 0$. Therefore the only contribution happens for $a \equiv -1$:

$$G(1,\overline{\chi})G(1,\chi) = \chi(-1) \sum_b \zeta^{0\cdot b} = \chi(-1) f_\chi$$

Since $\chi(-1) = \overline{\chi}(-1)$, the desired result holds.

□

Definition 6.12. Given a Dirichlet character χ, the *Dirichlet L-function* associated to χ is:

$$L(s,\chi) = \sum_{n=1}^{\infty} \frac{\chi(n)}{n^s}$$

Notice that if f is the conductor of χ, then:

$$\sum_{n=1}^{f} \chi(n) = G(0,\chi) = \chi(0)G(1,\chi) = 0$$

Therefore $\sum_{n=1}^{\infty} \chi(n) = \mathcal{O}(1)$. By Lemma 6.6, $L(s,\chi)$ converges for all $\mathrm{Re}(s) > 0$ and is holomorphic in that region. Moreover, one can show that L has a product formula:

$$L(s,\chi) = \prod_p \left(1 - \chi(p)p^{-s}\right)^{-1}$$

which we won't prove here. The remainder of this section will focus on a special class of L-functions that are associated to what are known as quadratic characters.

6.2.1 Quadratic Characters and Quadratic Fields

A character $\chi : (\mathbb{Z}/f\mathbb{Z})^* \to \mathbb{C}^*$ is quadratic if $\mathrm{im}(\chi) \subset \{\pm 1\}$. We say that χ is odd if $\chi(-1) = 1$, and even otherwise. The following are some examples of quadratic characters that we will use:

- For p an odd prime, define $\chi_p = \left(\frac{\cdot}{p}\right) : (\mathbb{Z}/p\mathbb{Z})^* \to \{\pm 1\}$, where $\left(\frac{\cdot}{p}\right)$ is the Legendre symbol. This is a quadratic character of conductor p.
- $\chi_4 : (\mathbb{Z}/4\mathbb{Z})^* \to \{\pm 1\}$ defined by sending -1 to -1 is a character of conductor 4.
- $\chi_8^+ : (\mathbb{Z}/8\mathbb{Z})^* \to \{\pm 1\}$ defined by sending -1 to 1 and 5 to -1 is a character of conductor 8.
- $\chi_8^- : (\mathbb{Z}/8\mathbb{Z})^* \to \{\pm 1\}$ defined by sending -1 to -1 and 5 to 1 is a character of conductor 8.

Quadratic characters are used to embed any quadratic number field in a cyclotomic extension of $\mathbb{Q}$, as claimed in the next Proposition:

Proposition 6.13. *Let K be a quadratic number field of discriminant D. Then $K \subset \mathbb{Q}(\zeta_{|D|})$ canonically.*

Proof:

First notice that, given any quadratic character χ of conductor f, we have $\chi = \overline{\chi}$ and $G(1,\chi)G(1,\chi) = \chi(-1)f$. Since $G(1,\chi)$ is a linear combination of elements in $\mathbb{Q}(\zeta_f)$ and its square is $\chi(-1)f$, we have $\mathbb{Q}(\sqrt{\chi(-1)f}) \subset \mathbb{Q}(\zeta_f)$, where ζ_f is a primitive f root of unity. Since K can we written as $\mathbb{Q}(\sqrt{d})$ for some squarefree d, it suffices to find a quadratic character χ such that $\chi(-1)f = D$. Recall that the discriminant D is equal to d if $d \equiv 1 \bmod 4$ and is equal to $4d$ otherwise. Let $d = \pm p_1 ... p_s$, with p_i prime and $p_i < p_j$ for $i < j$.

We claim that if $d \equiv 1 \bmod 4$, the character $\chi = \chi_{p_1} ... \chi_{p_s}$ satisfies $\chi(-1)f_\chi = D$. If $d \equiv 1 \bmod 4$ and d odd, we take $\chi = \chi_4 \chi_{p_1} ... \chi_{p_s}$. If d is even (i.e. $p_1 = 2$), then we take $\chi = \chi_8^{\pm} \chi_{p_2} ... \chi_{p_s}$. The reader can verify that all of these satisfy the desired criterion.

□

Remark 6.14. The character that we constructed is sometimes referred to as the character of K.

As a consequence, given K quadratic with discriminant D, we have a surjection $\phi : \mathrm{Gal}(\mathbb{Q}(\zeta_{|D|})/\mathbb{Q}) \to \mathrm{Gal}(K/\mathbb{Q})$ given by restriction. These groups are canonically identified with $(\mathbb{Z}/|D|\mathbb{Z})^*$ and $\{\pm 1\}$, respectively. Therefore ϕ defines a quadratic character.

Claim: ϕ is the character that we constructed in the proof of Proposition 6.13.

Proof:

An element $\sigma_n \in \mathrm{Gal}(\mathbb{Q}(\zeta_{|D|})/\mathbb{Q})$ is defined by $\zeta_{|D|} \mapsto \zeta_{|D|}^n$. Notice that $\phi(\sigma_n) = -1$ if and only if $\sigma_n|_K \neq \mathrm{id}$. Since $K = \mathbb{Q}(G(1,\chi))$, this happens if and only if $\sigma_n(G(1,\chi)) = -G(1,\chi)$. Then we can

compute:

$$\begin{aligned}\sigma_n(G(1,\chi)) &= \sigma_n \sum_a \chi(a)\zeta_{|D|}^a \\ &= \sum_a \chi(a)\zeta_{|D|}^{na} \\ &= G(n,\chi) = \chi(n)G(1,\chi)\end{aligned}$$

The last equality being by Proposition 6.11. We see that $\chi(n) = -1$ if and only if $\phi(\sigma_n) = -1$. Since χ and ϕ are quadratic, this is enough to show that they are the same.

□

Recall that, for any prime $p \in \mathbb{Z}$, there are three posibilities for how $(p) = pO_K$ factorizes. If $(p) = \mathfrak{p}$, it is called inert; if $(p) = \mathfrak{p}\bar{\mathfrak{p}}$ [4], it is called split; if $(p) = \mathfrak{p}^2$, it is called ramified. The character we constructed associated to K tells us which of these is the case for any p:

Proposition 6.15. *Let χ be the associated character of a quadratic field K. Then $\chi(p) = 0$ if p ramifies, $\chi(p) = 1$ if p splits, and $\chi(p) = -1$ if p is inert.*

Proof:

This follows from the two ways we can write the quotient O_K/pO_K. On one hand, we have:

$$O_K/pO_K = \begin{cases} \mathbb{Z}/p^2\mathbb{Z} & \text{iff } p \text{ ramifies} \\ (\mathbb{Z}/p\mathbb{Z}) \times (\mathbb{Z}/p\mathbb{Z}) & \text{iff } p \text{ splits} \\ \mathbb{F}_{p^2} & \text{iff } p \text{ is inert} \end{cases}$$

This follows from the Chinese remainder theorem and that O_K is a rank 2 free module. On the other hand, we know that $O_K = \mathbb{Z}[x]/Q(x)$ for some quadratic $Q(x)$. Then:

$$O_K/pO_K = (\mathbb{Z}[x]/Q(x))/(p \cdot \mathbb{Z}[x]/Q(x)) = \mathbb{F}_p[x]/q(x)$$

where $q(x) \in \mathbb{F}_p[x]$ is the reduction of $Q(x)$ mod p. There are three posibilities for how $q(x)$ can factor, depending on the discriminant D of K. Since the discriminant of Q is the discriminant of q mod p, we see that $q(x)$ has two roots if D is a square mod p, $q(x)$ has one root of multiplicity 2 if D is zero mod p, and $q(x)$ is irreducible otherwise. Therefore:

$$O_K/pO_K = \begin{cases} \mathbb{Z}/p^2\mathbb{Z} & \text{iff } D \equiv 0 \text{ mod } p \\ (\mathbb{Z}/p\mathbb{Z}) \times (\mathbb{Z}/p\mathbb{Z}) & \text{iff } D \text{ is a nonzero square mod } p \\ \mathbb{F}_{p^2} & \text{otherwise} \end{cases}$$

The reader can check that the right-hand condidtions are equivalent to $\chi(p) = 0$, $\chi(p) = 1$, and $\chi(p) = -1$, respectively.

□

Corollary 6.16. *If p is a prime and χ is the character of a quadratic field, then:*

$$\prod_{\mathfrak{p}|p}(1 - N(\mathfrak{p})^{-s}) = (1 - p^{-s})(1 - \chi(p)p^{-s}).$$

The proof follows directly from the previous proposition and is left as an exercise.

[4] $\bar{\mathfrak{p}}$ is the conjugate of $\mathfrak{p}$, where conjugation is $a + b\sqrt{d} \mapsto a - b\sqrt{d}$.

Theorem 6.17. *If K is a quadratic number field and χ is the associated character, then:*

$$L(1,\chi) = \begin{cases} \frac{2\pi|\operatorname{Cl}(K)|}{w_K\sqrt{|\operatorname{disc} K|}} & K \text{ imaginary} \\ \frac{2|\operatorname{Cl}(K)|\log(u_K)}{\sqrt{|\operatorname{disc} K|}} & K \text{ real} \end{cases}$$

where u_K is the fundamental unit of K if it is real.

Proof:

Using the previous Corollary, we can write:

$$\zeta_K(s) = \prod_{\mathfrak{p}\subset O_K} \frac{1}{1-N(\mathfrak{p})^{-s}} = \prod_p \frac{1}{(1-p^{-s})}\frac{1}{(1-\chi(p)p^{-s})} = \zeta(s)L(s,\chi)$$

Since ζ_K has a simple pole at $s=1$ and ζ has a simple pole at $s=1$, it follows that $L(1,\chi)$ is finite and is equal to $\frac{\operatorname{Res}(\zeta_K,1)}{\operatorname{Res}(\zeta,1)} = \operatorname{Res}(\zeta_K,1)$. The theorem then follows from the residue formula from Theorem 6.5.

□

Example 6.18. Let $K=\mathbb{Q}(i)$, where $i^2=-1$. Then $\chi=\chi_4$, which has the formula:

$$\chi(a) = \begin{cases} 0 & a \text{ even} \\ (-1)^{(a-1)/2} & a \text{ odd} \end{cases}$$

Then the L series at $s=1$ is:

$$L(1,\chi) = 1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \cdots = \frac{\pi}{4}$$

By the above theorem, we get:

$$\frac{\pi}{4} = \frac{\pi|\operatorname{Cl}(K)|}{4} \Rightarrow |\operatorname{Cl}(K)| = 1$$

Therefore K has trivial class group.

Proposition 6.19. *If K is a quadratic field with associated character χ whose conductor is $f>1$, then:*

$$L(1,\chi) = \begin{cases} \frac{\pi}{f\sqrt{f}}\left|\sum_a \chi(a)a\right| & \text{if } K \text{ is imaginary} \\ \frac{1}{\sqrt{f}}\left|\sum_a \chi(a)\log\sin(\pi a/f)\right| & \text{if } K \text{ is real} \end{cases}$$

where the sums are taken over $a\in\mathbb{Z}/f\mathbb{Z}$.

Proof:

Using the first property of the Gauss sum, we have:

$$\begin{aligned} G(1,\chi)L(1,\chi) &= \sum_n^\infty \frac{\chi(n)G(1,\chi)}{n} = \sum_n^\infty \frac{G(n,\chi)}{n} \\ &= \sum_{a\in\mathbb{Z}/f\mathbb{Z}} \chi(a)\sum_n^\infty \frac{\zeta_f^{an}}{n} \end{aligned}$$

Since $\sum_{n=1}^\infty z^n/n = -\log(1-z)$ for any $|z|\le 1, z\neq 1$, we get:

$$G(1,\chi)L(1,\chi) = -\sum_{a\in\mathbb{Z}/f\mathbb{Z}} \chi(a)\log(1-\zeta_f^a)$$

We can rewrite $\zeta_f^a - 1$ as:

$$\begin{aligned}\zeta_f^a - 1 &= e^{2\pi i a/f} - 1 \\ &= i\sin(2\pi a/f) + \cos(2\pi a/f) - 1 \\ &= 2i\sin(\pi a/f)\cos(\pi a/f) - 2\sin^2(\pi a/f) \\ &= e^{i\pi a/f} \cdot 2i\sin(\pi a/f)\end{aligned}$$

Then:

$$\begin{aligned}G(1,\chi)L(1,\chi) &= -\sum_a \chi(a)\log(-e^{i\pi a/f}\cdot 2i\sin(\pi a/f)) \\ &= -\sum_a \chi(a)\left(\frac{i\pi a}{f} - \frac{i\pi}{2} + \log 2 + \log\sin(\pi a/f)\right) \\ &= -\sum_a \chi(a)\cdot\frac{i\pi a}{f} - \sum_a \chi(a)\log\sin(\pi a/f)\end{aligned}$$

where we used that $\sum_a \chi(a) = 0$ because χ is nontrivial. We note that $|G(1,\chi)| = \sqrt{f}$ and that $L(1,\chi)$ is positive real, so that:

$$|G(1,\chi)L(1,\chi)| = \sqrt{f}L(1,\chi)$$

$$\implies L(1,\chi) = \frac{1}{\sqrt{f}}\left|\frac{i\pi}{f}\sum_a \chi(a)a + \sum_a \chi(a)\log\sin(\pi a/f)\right|$$

If χ is even, and hence K is real, the first sum is zero. If χ is odd, and hence K is complex, the second sum is zero. The result follows.

□

These formulas for $L(1,\chi)$ generalize the $\pi/4$ sum that we saw in the previous example. Just as with that example, as a consequence we can get a formula for the class number of a quadratic number field:

Corollary 6.20. *Let K be a quadratic number field with associated Dirichlet character χ whose conductor is $f > 1$. Then:*

$$|\operatorname{Cl}(K)| = \begin{cases} \frac{w_K}{2\sqrt{|\operatorname{disc} K|}}\left|\sum_a \chi(a)a\right| & K \text{ imaginary} \\ \frac{1}{2\log(u_K)}\left|\sum_a \chi(a)\log\sin(\pi a/f)\right| & K \text{ real}\end{cases}$$

where the sums are over $a \in \mathbb{Z}/f\mathbb{Z}$, u_K is a fundamental unit, and w_K is the number of roots of unity in K.

Proof:

This follows immediately from Proposition 6.19 and Theorem 6.17.

□

Example 6.21. Let $K = \mathbb{Q}(\sqrt{-5})$. Then $\chi = \chi_4\chi_5$ because $f = 20$. Since K is imaginary, we have:

$$\begin{aligned}|\operatorname{Cl}(K)| &= \frac{2}{2\cdot 20}\left|\chi(1) + 3\chi(3) + 7\chi(7) + 9\chi(9) + 11\chi(11) + 13\chi(13) + 17\chi(17) + 19\chi(19)\right| \\ &= \frac{1}{20}\left|1 + 3 + 7 + 9 - 11 - 13 - 17 - 19\right| \\ &= \frac{40}{20} = 2\end{aligned}$$

7. Introduction to Global Class Field Theory

❖

In this section, we will give a brief introduction to the class field theory of global fields. Roughly speaking, a global field is either a number field or a function field of an algebraic curve. Since we have not dealt with the latter, this section will only deal with number fields. Our goal will be to define the Ray Class group, which is a generalization of the Ideal Class Group, and state some of its properties. Then we will finish with the Artin symbol.

7.1 Moduli and the Ray Class Group ❖

Let K be a number field. Recall that for every prime ideal $\mathfrak{p} \subset O_K$, we have a $\mathfrak{p}$-adic valuation:

$$v_{\mathfrak{p}}(x/y) = e_x - e_y$$

where e_x and e_y are the unique exponents of $\mathfrak{p}$ in the factorization of x and y, respectively. This induces the $\mathfrak{p}$-adic norm $|\cdot|_{\mathfrak{p}} := \pi^{-v(\cdot)}$, where π is a generator of the prime ideal $\mathfrak{p}(O_K)_{\mathfrak{p}}$.[5] Moreover, for every real or complex embedding τ, there is an induced norm given by $|\tau(\cdot)|$, where $|\cdot|$ is the standard norm on $\mathbb{R}$ and $\mathbb{C}$. Ostrowski's theorem says that these two are the only norms that K admits.

Definition 7.1. A *place* of a number field K is an equivalence class of norm on K.

A place is distinguished by whether it corresponds to a prime ideal $\mathfrak{p} \subset O_K$ (called a *finite prime*) or an embedding τ (called an *infinite prime*). From here out, we will denote $O := O_K$, $U := O_K^\times$ and $\mathrm{Spec}(O_K)$ to be the set of all places of K.

Definition 7.2. Let K/F be a finite extension of number fields and let q be an infinite prime of K and P be an infinite prime of F such that $q \mid p$ (i.e. q agrees with p when restricted to F). Then we say that p is *ramified* in K/F if p is real and q is complex.

Definition 7.3. The *support* $\mathrm{Supp}(c_0)$ of an integral ideal $c_0 = \prod_i^k \mathfrak{p}_i^{e_i}$ is the set of prime ideals $\{\mathfrak{p}_1, ..., \mathfrak{p}_k\}$ dividing c_0.

Definition 7.4. A *modulus* of K is a formal product $c = c_0 c_\infty$, where $c_0 = \prod_i^k \mathfrak{p}_i^{e_i}$ is an integral ideal ($e_i \geq 0$) and c_∞ is a formal square-free product of real places of K. A modulus is also called a *cycle*. We say c is *finite* if $c_\infty = 1$.

Example 7.5. If $K = \mathbb{Q}$, there is only one infinite prime, denoted ∞, which corresponds to the absolute value on $\mathbb{Q}$. Examples of moduli are $c = 1, c = 7^2 \cdot 11$ and $c = 7^2 \cdot 5 \cdot \infty$.

Given $\alpha \in K^*$ and a modulus $c = c_0 c_\infty$, we say that $\alpha \equiv 1 \mod^* c$ if $v_{\mathfrak{p}_i}(\alpha - 1) \geq e_i$ for all $\mathfrak{p}_i \in \mathrm{Supp}(c_0)$ and if $\phi_i(\alpha) > 0$ for all ϕ_i real infinite places in c_∞. Then we define:

$$I(c) = \{\text{All fractional ideals of } O \text{ prime to } c_0\}$$

$$P_c = \{\text{All fractional ideals of } O \text{ that are principal and generated by } \alpha \in K, \alpha \equiv 1 \mod^* c\}$$

Remark 7.6. If $(\alpha) \in P_c$, then there exists $u \in U$ such that $\alpha u \equiv 1 \mod^* c$.

The reader can check that P_c is a subgroup of $I(c)$. The Ray Class group (otherwise called the Generalized Ideal Class group) is the quotient $I(c)/P_c$. As our notation suggests, this group depends on the modulus c. There are three special cases of c:

[5] Recall that $R_{\mathfrak{p}}$ is a PID when R is a Dedekind domain

- If $c = 1$, then I_c is the set of all fractional ideals and P_c is the set of principal fractional ideals. Therefore $I(1)/P_1 = \mathrm{Cl}(O)$.
- If $c = c_\infty$, then $I(c)/P_c$ is called the Narrow Class group.
- If $c = c_0$ (i.e. c is finite), then we have the following exact sequence:

$$1 \longrightarrow U_c \longrightarrow U \longrightarrow (O/cO)^\times \longrightarrow I(c)/P_c \longrightarrow \mathrm{Cl}(O) \longrightarrow 1$$

  where $U_c = \{u \in U \mid u \equiv 1 \mod^* c\}$. Since $(O/cO)^\times$ is finite and $\mathrm{Cl}(O)$ is finite, everything in the above sequence is finite. Moreover, if we define $\phi(c) := |(O/cO)^\times|$, then using exacness we can calculate:

$$h_c = \frac{h\phi(c)}{[U : U_c]}$$

  where $h_c = |I(c)/P_c|$ and $h = |\mathrm{Cl}(O)|$.

For a general modulus, define $\phi(c) = 2^{s(c_\infty)}|(O/c_0O)^\times|$, where $s(c_\infty)$ is the number of terms in c_∞.[6] Then the formula we found for h_c above actually holds for general c:

Proposition 7.7. *For any modulus c, we have $|I(c)/P_c| = h_c = \frac{h\phi(c)}{[U:U_c]}$.*

In particular, the Ray Class group is finite for any c. The finiteness of this group can be thought of as a generalization of Dirichlet's theorem of primes in arithmetic progressions, since every Ray class must contain infinitely many prime ideals.

Example 7.8. Let $K = \mathbb{Q}$ again. Let $c = m\infty$ be a modulus for some $m > 2$. Since $U = \{\pm 1\}$, we have $U_{c_0} = \{1\}$ and:

$$I(c) = \{r = p/q \in \mathbb{Q}^\times \mid (p, c_0) = (q, c_0) = 1\}$$

$$P_c = \{(r) \in \mathcal{I}(O) \mid r \equiv 1 \mod^* c\} = \{r \in \mathbb{Q}^+ \mid c_0 \mid (r-1)\}$$

Therefore $I(c)/P_c \cong (\mathbb{Z}/m\mathbb{Z})^\times$ and $I(c_0)/P_{c_0} \cong (\mathbb{Z}/m\mathbb{Z})^\times/\{\pm 1\}$.

7.2 The Artin Symbol ❖

Let L/K be a Galois extension of number fields, and fix $\mathfrak{p} \subset O_K$ and $\mathfrak{q} \subset O_L$ lying over $\mathfrak{p}$. Recall the decomposition group $D_\mathfrak{q} \subset \mathrm{Gal}(L/K)$ was defined as the set of elements in $\mathrm{Gal}(L/K)$ that fixed $\mathfrak{q}$. We saw in Section 4.1.1 that there is surjection $\phi_\mathfrak{q} : D_\mathfrak{q} \to \mathrm{Gal}(k_\mathfrak{q}/k_\mathfrak{p})$ given by restriction to $k_\mathfrak{q}$ whose kernel is the intertia group $I_\mathfrak{q}$. In the case where L/K was abelian, we defined the Frobeneus element $(\mathfrak{p}, L/K) \in \mathrm{Gal}(L/K)$ as the unique preimage of the Frobenius automorphism in $\mathrm{Gal}(k_\mathfrak{q}/k_\mathfrak{p})$.

Given a modulus c that is divisible by all ramified primes in O_K, we now have a well-defined map for any abelian extension L of K:

$$\phi_{c,L,K} : I(c) \to \mathrm{Gal}(L/K)$$

$$\prod \mathfrak{p}_i^{e_i} \mapsto \prod (\mathfrak{p}, L/K)^{e_i}$$

This is known as the *Artin symbol*. We sometimes denote $\phi_{c,L,K}(I)$ by $(I, L/K)$. Some properties of ϕ that we won't prove are:

1. If $\sigma : L \to F$ is a field isomorphism, then $\sigma L/\sigma K$ is another abelian extension. For any ideal $\mathfrak{a} \in I(c)$, we have $(\sigma\mathfrak{a}, \sigma L/\sigma K) = \sigma(\mathfrak{a}, L/K)\sigma^{-1}$.
2. If $K \subset L \subset M$ are abelian extensions, then $(\mathfrak{a}, L/K) = \mathrm{res}_K(\mathfrak{a}, M/K)$.

[6] This is a generalization of the totient function.

3. If L/K is abelian and E/K is finite, then $\mathrm{Gal}(LE/E)$ is abelian as well. Let $\mathfrak{p}$ be a prime of K that is unramified in L/K, and let $\mathfrak{q}$ be a prime of E with $\mathfrak{q} \mid \mathfrak{p}$. Then $\mathfrak{q}$ is unramified in LE/E. Moreover, $(\mathfrak{p}, L/K)^{f(E/K)} = \mathrm{res}_L(\mathfrak{q}, LE/E)$. Since we have:

$$(\mathfrak{p}, L/K)^e = (\mathfrak{p}^e, L/K)$$

for any e, this also means that $(N_{E/K}\mathfrak{q}, L/K) = (\mathfrak{p}^{f(E/K)}, L/K) = \mathrm{res}_L(\mathfrak{q}, LE/E)$.