Arithmetic of cyclotomic fields

Tudor Ciurca

September 26, 2018

*

Contents

*Department of Mathematics, Imperial College London, London, SW7 6AZ, United Kingdom
E-mail address: teodor.ciurca@gmail.com

Abstract

For 300 years since its conception, Fermat's last theorem went unsolved. At first, only special cases were demonstrated. Fermat himself proved the exponent 4 case. Euler proved the exponent 3 case. Dirichlet proved the exponent 5 case. Lamé's proof for the exponent 7 case was quite complicated. Most of these early proofs went by descent.

It was Lamé who first noticed that cyclotomic fields are the right environment to work with in regards to Fermat's last theorem. He showed that Fermat's last theorem for exponent p would follow from unique prime factorization in the ring $\mathbb{Z}[\zeta_p]$, but he incorrectly assumed that unique prime factorization holds in all such rings. Here is where Kummer comes into the picture. He arguably had the most impact on Fermat's last theorem before Wiles.

Kummer introduced the concept of ideal numbers, a precursor to ideals, in order to fix unique prime factorization when it fails. He proved that ideal numbers have unique prime factorization, and fixed Lame's proof, even though it was limited to prime exponents satisfying a certain condition. These primes were called regular. Statistically, around 61% of primes are regular, although the infinitude of regular primes is still an unsolved problem.

This document is the result of a UROP project undertaken in the summer of 2018 at Imperial College London. The preliminaries assumed in this document include definitions from an undergraduate course in commutative algebra and results from Galois theory. In later chapters we make use of class field theory and Kummer theory, and refer the reader to some sources where this theory is developed.

Section 1 provides a comprehensive introduction to some topics in algebraic number theory at the level of an advanced undergraduate course. Section 2 uses the theory developed in Section 1 to study specific examples and families of number fields. Section 3 gives a crash course on the theory of adeles, assuming some knowledge of local fields. This will be used to prove Dirichlet's unit theorem and the finiteness of the ideal class group.

Section 4 introduces Dirichlet L-series and proves Dirichlet's theorem on primes in arithmetic progression. We also introduce Dedekind zeta functions and derive the analytic class number formula. We then study the relationship between Dirichlet characters of abelian Galois groups and their associated intermediate number fields. Section 5 is where we actually begin studying the arithmetic of cyclotomic fields. We apply our results to prove Fermat's last theorem for regular primes, following Kummer. We then build upon results in Section 4 and derive the relative class number formula for prime cyclotomic fields.

In Section 6, our main aim is to prove Kummer's criterion for the regularity of a prime. We give both an analytic and algebraic derivation for the preliminary criterion of whether p divides the relative class number. The analytic method uses p-adic L-functions whilst the algebraic method uses the Stickelberger relation to prove Herbrand's theorem. Along the way we study Gauss sums which will hep us prove Stickelberger's relation. Finally we use Kummer theory to relate the relative class number to the class number of a prime cyclotomic field.

1 Dedekind domains and their ideals

Number theory began as the study of the integers. They act as a scaffolding for the field of rational numbers. During the 19^{th} century mathematicians realized that number fields, that is, field extensions of $\mathbb{Q}$ of finite degree, have similar properties to $\mathbb{Q}$ and are the right context to study solutions of polynomial equations with integer coefficients in one variable. The "number theory" of a number field can be developed in a similar way to that of $\mathbb{Q}$. There is an analogue to the integers for every number field, called its ring of integers, which acts as the scaffolding of the number field and contains many of its number-theoretic properties. In this section we will study number fields and their rings of integers. The important properties satisfied by these rings of integers are summarized in the abstract object known as a Dedekind domain, which is defined below. In this document every ring is commutative and has a multiplicative identity.

The material in this section is based on a variety of sources. The main source is a course in algebraic number theory that the author has undertaken at the summer school PROMYS Europe 2017. This is also supplemented by [Cox13] and several expository papers by Keith Conrad and William Stein. At the beginning of every subsection, the specific sources used will be mentioned. If not, then the material of the subsection originates from the algebraic number theory course that the author took.

Definition 1.0.1. *A Dedekind domain is an integral domain R which is also*

- *Noetherian*
- *of Krull dimension 1*
- *integrally closed in* $\mathrm{Frac}(R)$*, the fraction field of R*

1.1 Rings of integers are Dedekind domains

Why are we interested in Dedekind domains? It is because the ring of integers $\mathcal{O}_K$ of a number field K is a Dedekind domain. Recall that the ring of integers is the integral closure of $\mathbb{Z}$ in K, and these are objects of central importance in algebraic number theory.

The first two conditions above can be thought of as some sort of restriction on the size of the ring. Note that one does not imply the other; there are examples of Krull dimension 1 rings which are not Noetherian. The next lemma proves that the rank of the ring of integers as a $\mathbb{Z}$-module equals the vector space dimension of the number field over $\mathbb{Q}$. This will be the crucial step in proving that rings of integers satisfy the first two conditions.

Lemma 1.1.1. *Let K be a number field. Then $\mathcal{O}_K$ is a free $\mathbb{Z}-$module of rank $[K:\mathbb{Q}]$.*

Proof. Let $K = \mathbb{Q}(a_1 \dots a_n)$ so that $\{a_1 \dots a_n\}$ is a $\mathbb{Q}$-basis for K. We will show that the a_i can be chosen to be elements of $\mathcal{O}_K$. Assume a_i is not such an element and let $f_i = \sum_{j=0}^{m} b_{i,j} x^j$ be its

minimal polynomial with $b_{i,j}$ integers and $b_{i,m}$ non-zero. Then we can multiply f_i by $b_{i,m}^{m-1}$ to get

$$b_{i,m}^{m-1} f_i = \sum_{j=0}^{m} b_{i,j} b_{i,m}^{m-j-1} (b_{i,m} a_i)^j = 0$$

As a result the polynomial $b_{i,m}^{m-1} f_i(b_{i,m}^{-1} x)$ is monic and has a root $b_{i,m} a_i$ which is an element of $\mathcal{O}_K$. Hence we can replace a_i with $b_{i,m} a_i$ since $b_{i,m}$ is a non-zero integer.

Now $\mathbb{Z}[a_1 \dots a_n] \subset \mathcal{O}_K$ is a free $\mathbb{Z}$-module of rank n, because the $a_1 \dots a_n$ being linearly independent over $\mathbb{Q}$ implies that they are linearly independent over $\mathbb{Z}$. To show that $\mathcal{O}_K$ is also a free $\mathbb{Z}$-module of rank n, we will consider the embedding

$$\psi : \mathcal{O}_K \to \mathbb{Z}^n$$

$$\psi : g \mapsto (Tr(g \cdot a_1) \dots Tr(g \cdot a_n))$$

where $Tr(\cdot)$ is the absolute trace in K. This is clearly a $\mathbb{Z}$-module homomorphism. If each $Tr(g \cdot a_i)$ is zero then $Tr(g \cdot h) = 0$ for any $h \in K$ since the trace is $\mathbb{Q}$-linear and $\{a_1 \dots a_n\}$ is $\mathbb{Q}$-basis for K. In particular that means $Tr(N(g)) = 0$ by selecting $h = \frac{N(g)}{g}$, where $N(\cdot)$ is the absolute norm in K. This implies that $g = 0$ because $N(g)$ is an integer and the trace of integers satisfies

$$Tr(N(g)) = n \cdot N(g)$$

This shows that ψ is injective, and so it is indeed an embedding. This means that $\mathcal{O}_K$ is a finitely generated $\mathbb{Z}$-module. It is obviously torsion-free, because K is, and therefore free by the fundamental theorem of finitely generated abelian groups. Its rank is therefore at most n, since we embedded it in $\mathbb{Z}^n$, but it must also be at least n as it contains $\mathbb{Z}[a_1 \dots a_n]$ as a submodule. This completes the proof of this lemma. □

Proposition 1.1.2. *Let K be a number field. Then $\mathcal{O}_K$ is Noetherian.*

Proof. Let $\mathfrak{a}$ be an ideal of $\mathcal{O}_K$. We claim that $\mathfrak{a}$ can be generated by $[K : \mathbb{Q}] = n$ elements. Suppose not, then we can find $n + 1$ elements in $\mathfrak{a}$ which are linearly independent over $\mathbb{Z}$, but this is impossible to do in $\mathcal{O}_K$, which has $\mathbb{Z}$-rank equal to n by Lemma 1.1.1. □

Lemma 1.1.3. *Let $\mathcal{O}_K$ be the ring of integers of a number field K. Then every non-zero ideal $\mathfrak{a}$ of $\mathcal{O}_K$ is a free $\mathbb{Z}$-module of rank $n = [K : \mathbb{Q}]$*

Proof. We've shown in Proposition 1.1.2 that every ideal of $\mathcal{O}_K$ is a finitely generated $\mathbb{Z}$-module. These ideals are torsion-free, because $\mathcal{O}_K$ is, and so they are free $\mathbb{Z}$-modules by the fundamental theorem of finitely generated abelian groups. In Proposition 1.1.2 we showed that these ideals can be generated by n elements, so they have $\mathbb{Z}$-rank at most n.

We claim that every $\mathcal{O}_K$-ideal $\mathfrak{a}$ contains some non-zero integer m. Just take the norm of any non-zero element in the ideal, which is a non-zero integer and must also be in the ideal. Let us suppose that $\{a_1 \dots a_n\}$ is a $\mathbb{Z}$-basis for $\mathcal{O}_K$. Then $\mathbb{Z}[ma_1 \dots ma_n]$ is also a free $\mathbb{Z}$-module of rank n, which embeds into $\mathfrak{a}$ by inclusion. Thus the $\mathbb{Z}$-rank of $\mathfrak{a}$ is at least n, so it must be exactly n. □

Proposition 1.1.4. *$\mathcal{O}_K$ has Krull dimension* 1.

Proof. Let $\mathfrak{a}$ be an ideal and $m \in \mathfrak{a}$ a non-zero integer, which we showed exists in Lemma 1.1.3. Then $(m) \subset \mathfrak{a}$ as ideals. By the order reversing inclusions of ideals and their quotient rings, we have $\mathcal{O}_K/\mathfrak{a} \subset \mathcal{O}_K/(m)$. By restriction of scalars, $\mathcal{O}_K/(m) = \mathcal{O}_K \otimes_{\mathbb{Z}} \mathbb{Z}/(m)$ is a free $\mathbb{Z}/(m)$-module of rank $[K : \mathbb{Q}]$. By looking at the sizes of these quotient rings we have

$$|\mathcal{O}_K/\mathfrak{a}| < |\mathcal{O}_K/(m)| = |\mathbb{Z}/(m)|^{[K:\mathbb{Q}]}$$

and so the quotient rings of all non-zero ideals $\mathfrak{a}$ are finite. Note that finite integral domains are fields, and as a result every prime ideal is maximal. Hence $\mathcal{O}_K$ has Krull dimension 1. □

We defined the ring of integers to be the integral closure of $\mathbb{Z}$ in K, and so the third condition is automatically satisfied.

Corollary 1.1.5. *$\mathcal{O}_K$ is a Dedekind domain.*

1.2 Unique prime factorization (UPF) of ideals in Dedekind domains

The next step is to show that failure of unique prime factorization in rings of integers of number fields can be remedied by working with ideals instead, which can be factored uniquely into prime ideals. This is a characteristic property of Dedekind domains, although we only prove one direction, that Dedekind domains as defined in Section 1.1 admit unique prime factorization of ideals.

Definition 1.2.1. *A fractional ideal of an integral domain R is an R-submodule $\mathfrak{i}$ of* $\mathrm{Frac}(R)$*, so that there is a non-zero element $r \in R$ so that $r\mathfrak{i} \subset R$.*

Definition 1.2.2. *Let R be a Dedekind domain with fraction field K. We denote by $\mathcal{I}_K$ the monoid of non-zero fractional ideals of R under multiplication, and by $\mathcal{P}_K$ the group of non-zero principal fractional ideals under multiplication.*

We will need to work with the more general notion of fractional ideals later on. For now we prove some general results about ideals which lead up to our desired result.

Theorem 1.2.3. *Let R be a commutative domain with a subring S. Let M be a finitely generated nonzero free S-submodule of R and let $b \in R$. Then $bM \subset M \implies b$ is integral over S.*

Proof. Let $\{m_1 \dots m_n\} \subset R$ be a basis for M over S. As $bM \subset M$, we can write

$$bm_i = \sum_{j=1}^{n} a_{i,j} m_j$$

for each $i = 1 \dots n$, where $a_{i,j} \in S$. In matrix form this comes out as

$$\begin{bmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,n} \\ a_{2,1} & a_{2,2} & \dots & a_{2,n} \\ \dots & \dots & \dots & \dots \\ a_{n,1} & a_{n,2} & \dots & a_{n,n} \end{bmatrix} \cdot \begin{bmatrix} m_1 \\ m_2 \\ \dots \\ m_n \end{bmatrix} = b \begin{bmatrix} m_1 \\ m_2 \\ \dots \\ m_n \end{bmatrix}$$

so b is an eigenvalue of the matrix A of coefficients $a_{i,j}$. Hence it satisfies its characteristic polynomial, which is monic and has coefficients in S. Therefore b is integral over S. □

Lemma 1.2.4. *Let R be a Noetherian ring. Then every ideal of R not equal to R itself contains a finite product of prime ideals.*

Proof. Let $\mathfrak{i} \neq R$ be an ideal. If it is prime, we are done, and if not, there exist elements $a, b \in R$ such that $ab \in \mathfrak{i}$ but $a, b \notin \mathfrak{i}$. We can therefore write $(\mathfrak{i} + a)(\mathfrak{i} + b) \subset \mathfrak{i}$ and repeat the procedure on the two ideals $(\mathfrak{i} + a)$ and $(\mathfrak{i} + b)$. This process leads to a potentially infinite chain of ideals under inclusion, but since R is Noetherian, it must terminate. Therefore $\mathfrak{i}$ contains a finite product of prime ideals. □

Lemma 1.2.5. *Let $\mathfrak{p}$ be a prime ideal of a ring R and let $\mathfrak{i}, \mathfrak{j}$ be ideals such that $\mathfrak{ij} \subset \mathfrak{p}$. Then $\mathfrak{i} \subset \mathfrak{p}$ or $\mathfrak{j} \subset \mathfrak{p}$.*

Proof. Let $a \in \mathfrak{i}, b \in \mathfrak{j}$. Then $ab \in \mathfrak{ij} \subset \mathfrak{p}$ and so either $a \in \mathfrak{p}$ or $b \in \mathfrak{p}$. This occurs for every such pair $(a, b) \in \mathfrak{i} \times \mathfrak{j}$. If $\mathfrak{i} \subset \mathfrak{p}$ then we are done. Otherwise there is an element $m \in \mathfrak{i}$ not in $\mathfrak{p}$. Then $mn \in \mathfrak{p}$ for all $n \in \mathfrak{j}$ and since $m \notin \mathfrak{p}$, we must have $n \in \mathfrak{p}$ for all $n \in \mathfrak{j}$. As a result $\mathfrak{j} \subset \mathfrak{p}$. □

Lemma 1.2.6. *Let R be a Dedekind domain. If $\mathfrak{p}$ is a prime ideal of R, then there is an element $q \in \mathrm{Frac}(R) \backslash R$ such that $q\mathfrak{p} \subset R$.*

Proof. Let $x \in \mathfrak{p}$ be a non-zero element. (x) contains some minimal product of prime ideals $\mathfrak{p}_1 \cdots \mathfrak{p}_n$ by Lemma 1.2.4, so that $\mathfrak{p}_1 \cdots \mathfrak{p}_{n-1} \not\subset (x)$. Since $\mathfrak{p}$ is prime and $\mathfrak{p}_1 \cdots \mathfrak{p}_n \subset (x) \subset \mathfrak{p}$, we can say W.L.O.G. that $\mathfrak{p}_n \subset \mathfrak{p}$ by Lemma 1.2.5. These ideals also happen to be maximal, because R has Krull dimension 1, and so $\mathfrak{p}_n = \mathfrak{p}$.

Let $y \in \mathfrak{p}_1 \cdots \mathfrak{p}_{n-1}$ so that $y \notin (x)$, which is possible because $\mathfrak{p}_1 \cdots \mathfrak{p}_{n-1} \not\subset (x)$. Then $y\mathfrak{p} \subset \mathfrak{p}_1 \cdots \mathfrak{p}_n \subset (x) \subset \mathfrak{p}$ and so $\frac{y}{x}\mathfrak{p} \subset \mathcal{O}_K$. However $\frac{y}{x} \notin \mathcal{O}_K$, so $q = \frac{y}{x}$ satisfies the conditions of the lemma. □

Proposition 1.2.7. *If $\mathfrak{p}$ is a prime ideal of a Dedekind domain R, then there is a fractional ideal $\mathfrak{p}^{-1}$ so that $\mathfrak{p}\mathfrak{p}^{-1} = \mathcal{O}_K$.*

Proof. Let $q \in K \backslash \mathcal{O}_K$ so that $q\mathfrak{p} \subset \mathcal{O}_K$, which exists by Lemma 1.2.6. Then we have $\mathfrak{p} \subset \mathfrak{p} + q\mathfrak{p} \subset \mathcal{O}_K$, But $\mathfrak{p}$ is maximal and so either $\mathfrak{p} + q\mathfrak{p} = \mathfrak{p}$ or $\mathfrak{p} + q\mathfrak{p} = \mathcal{O}_K$. The first case would imply that $(1, q)\mathfrak{p} \subset \mathfrak{p}$, which would mean that q is an algebraic integer by Theorem 1.2.3. So $q \in \mathcal{O}_K$ since

$\mathcal{O}_K$ is integrally closed in K, which contradicts the assumptions on q. We must have $(1,q)\mathfrak{p} = \mathcal{O}_K$ instead, which means that $(1,q) = \mathfrak{p}^{-1}$ is the inverse ideal of $\mathfrak{p}$. □

Example 1.2.8. *This example illustrates the above results, where we have actually given a method for finding inverses of prime ideals. Let's take* $K = \mathbb{Q}(\sqrt{-13})$. *In this case* $\mathcal{O}_K = \mathbb{Z}(\sqrt{-13})$. *Finding inverses of principal prime ideals would be easy, so we will instead find the inverse of the nonprincipal prime ideal* $\mathfrak{p} = (11, 3+\sqrt{-13})$.

From the notation above we set $x = 11 \in \mathfrak{p}$. *We factor* $(x) = (11, 3-\sqrt{-13})(11, 3+\sqrt{-13})$, *and set* $y = 3-\sqrt{-13} \in (11, 3-\sqrt{-13})\backslash(x)$. *We check that* $\frac{y}{x}\mathfrak{p} = \frac{3-\sqrt{-13}}{11}(11, 3+\sqrt{-13}) = (3-\sqrt{-13}, 2) \subset \mathcal{O}_K$ *and that* $\frac{y}{x}$ *is not an algebraic integer, since its trace is* $\frac{6}{11}$.

Then by our results above the fractional ideal $(\frac{3-\sqrt{-13}}{11}, 1)$ *is the inverse* $\mathfrak{p}^{-1}$. *To check we see that* $(\frac{3-\sqrt{-13}}{11}, 1)(11, 3+\sqrt{-13}) = (3-\sqrt{-13}, 2, 11, 3+\sqrt{-13}) = \mathcal{O}_K$.

Proposition 1.2.9. *Let* R *be a Dedekind domain and* $\mathfrak{i}$ *an ideal of* R *not equal to* R *itself. Then* $\mathfrak{i}$ *can be factored into a finite number of prime ideals.*

Proof. The ideal (0) is already prime. By Zorn's lemma, every non-zero ideal of R not equal to R itself is contained in a maximal ideal. Let us assume that $\mathfrak{i} \subset \mathfrak{p}$ for some maximal ideal $\mathfrak{p}$. Then we have $\mathfrak{p}^{-1}\mathfrak{i} \subset \mathcal{O}_K$ because the inverse $\mathfrak{p}^{-1}$ exists by Proposition 1.2.7. We call this new ideal $\mathfrak{i}_0$ and repeat the process, getting a potentially infinite chain of ideals $\mathfrak{i} \subset \mathfrak{i}_0 \subset \mathfrak{i}_1 \subset \dots$ with $\mathfrak{i}_{j+1} = \mathfrak{p}_j^{-1}\mathfrak{i}_j$ for all j as long as some maximal ideal $\mathfrak{p}_j$ exists by Zorn's lemma. Since R is Noetherian, this chain must terminate and so eventually $\mathfrak{i}_n = R$ for some n. Hence we can write $\mathfrak{i} = \mathfrak{p} \cdot \mathfrak{p}_0 \cdots \mathfrak{p}_{n-1}$. □

This proves that factorization of ideals into primes exists. Now we have to show uniqueness, but first we will see what the above result says from the viewpoint of fractional ideals.

Theorem 1.2.10. *Let* R *be a Dedekind domain with fraction field* K. *Then* $\mathcal{I}_K$ *is an abelian group.*

Proof. Let $\mathfrak{i}$ be a non-zero ideal of R not equal to R itself. Then $\mathfrak{i}$ can be factored into a finite number of prime ideals by Proposition 1.2.9, each of which has an inverse by Proposition 1.2.9. Therefore, the ideal $\mathfrak{i}$ itself has an inverse.

Now let $\mathfrak{i}$ be a non-zero fractional ideal of R. By the definition of fractional ideals, there is some non-zero element $r \in R$ so that $r\mathfrak{i} \subset R$. $r\mathfrak{i}$ is an ideal too, which has an inverse $\mathfrak{j}$. The inverse of $\mathfrak{i}$ is therefore $r\mathfrak{j}$. This proves the statement. □

Definition 1.2.11. *Let* R *be a Dedekind domain with fraction field* K. *The ideal class group of* R *is defined to be the quotient* $\mathcal{I}_K/\mathcal{P}_K$, *which is an abelian group.*

Theorem 1.2.12. *Let* R *be a Dedekind domain. If* $\mathfrak{i} \subset \mathfrak{j}$ *as non-zero ideals of* R, *then there is some ideal* $\mathfrak{h}$ *of* R *such that* $\mathfrak{i} = \mathfrak{j}\mathfrak{h}$.

Proof. Inverses of arbitrary non-zero ideals exist by Theorem 1.2.10. As a result we can write $\mathfrak{i} \subset \mathfrak{j} \implies \mathfrak{j}^{-1}\mathfrak{i} \subset R$ and so $\mathfrak{h} = \mathfrak{i}\mathfrak{j}^{-1}$ is an ideal of R, which satisfies $\mathfrak{i} = \mathfrak{j}\mathfrak{h}$. □

This theorem is sometimes remembered as "to contain is to divide". Unique prime factorization of ideals is almost a direct consequence of this theorem.

Theorem 1.2.13. *Let R be a Dedekind domain. Then every ideal of R not equal to R itself can be factored uniquely into a finite product of prime ideals.*

Proof. Proposition 1.2.9 demonstrates existence. To show uniqueness, let's take an ideal $\mathfrak{i} \neq R$ and suppose we have two distinct factorizations into prime ideals $\mathfrak{i} = \mathfrak{p}_1^{e_1} \cdots \mathfrak{p}_n^{e_n} = \mathfrak{q}_1^{f_1} \cdots \mathfrak{q}_m^{f_m}$.

By Theorem 1.2.10, a cancellation law exists and so we can assume that the two factorizations have been reduced to $\mathfrak{p}_1^{g_1} \cdots \mathfrak{p}_n^{g_n} = \mathfrak{q}_1^{h_1} \cdots \mathfrak{q}_m^{h_m}$ with no factors in common and $h_1 \neq 0$ by rearrangement. This means we have $\mathfrak{p}_1^{g_1} \cdots \mathfrak{p}_n^{g_n} \subset \mathfrak{q}_1$ and so $\mathfrak{p}_i \subset \mathfrak{q}_1$ for some i. In particular $\mathfrak{p}_i = \mathfrak{q}_1$ since R has Krull dimension 1. This is a contradiction and so the two initial factorizations must have been identical. □

Example 1.2.14. *This is the example all undergraduate number theorists see. The ring $\mathbb{Z}[\sqrt{-5}]$ is not a UFD. Indeed, we have*

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

where all elements involved are prime. However, when passing to ideals, we see that the relevant ideals are not prime. In fact we can factorize each of them as follows:

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

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

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

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

It is then clear that (6) can be factored uniquely into prime ideals.

1.3 Ideal factorization

In this subsection we state some results that will allow us to factorize ideals into products of prime ideals and determine whether an ideal is prime or not. We will first introduce the notion of ideal norm, which is a way to measure the "size" of the ideal. In fact, the norm of an ideal is exactly the size of the corresponding quotient ring. We will later make use of the Chinese remainder theorem, which we shall prove now.

Theorem 1.3.1 (Chinese remainder theorem)**.** *Let R be a commutative ring and let $\mathfrak{i}_1 \dots \mathfrak{i}_n$ be a set of pairwise coprime ideals of R. That is, $\mathfrak{i}_i + \mathfrak{i}_j = R$ for all $i \neq j$. Then $R/(\prod_{i=1}^n \mathfrak{i}_i) \cong \bigoplus_{i=1}^n R/\mathfrak{i}_i$.*

Proof. It suffices to prove this in the case of 2 coprime ideals $\mathfrak{i}_1, \mathfrak{i}_2$. Let $a_1 \in \mathfrak{i}_1$, $a_2 \in \mathfrak{i}_2$ so that $a_1 + a_2 = 1$. We define the map $R/\mathfrak{i}_1\mathfrak{i}_2 \to R/\mathfrak{i}_1 \oplus R/\mathfrak{i}_2$ by sending x to the componentwise reduction $(x \pmod{\mathfrak{i}_1}, x \pmod{\mathfrak{i}_2})$. This is a ring homomorphism, because we are simply reducing modulo ideals in each component.

To show surjectivity, let $(m, n) \in R/\mathfrak{i}_1 \oplus R/\mathfrak{i}_2$ and choose representatives m', n' for m, n respectively in $R/\mathfrak{i}_1\mathfrak{i}_2$. Let $\bar{a_1}$ and $\bar{a_2}$ denote the reduction of a_1 and a_2 respectively modulo $\mathfrak{i}_1\mathfrak{i}_2$. Then $m'\bar{a_2} + n'\bar{a_1}$ will certainly map to (m, n).

To show injectivity, let $x \in R/\mathfrak{i}_1\mathfrak{i}_2$ be congruent to 0 modulo both $\mathfrak{i}_1$ and $\mathfrak{i}_2$. Choose a representative $x' \in R$ of x. Then x' is contained in both $\mathfrak{i}_1$ and $\mathfrak{i}_2$. Now a_1x' and a_2x' are both contained in $\mathfrak{i}_1\mathfrak{i}_2$, and thus $(a_1 + a_2)x' = x'$ is contained in $\mathfrak{i}_1\mathfrak{i}_2$. As a result x' must be congruent to 0 modulo $\mathfrak{i}_1\mathfrak{i}_2$ and so $x = 0$. This completes the proof. □

Definition 1.3.2. *Let K be a number field and let $\mathfrak{i}$ be an ideal of $\mathcal{O}_K$. Then the absolute norm of $\mathfrak{i}$ is defined as*

$$N(\mathfrak{i}) := |\mathcal{O}_K/\mathfrak{i}|$$

It is important to note that the absolute ideal norm exists. This follows from Section 1.1, where we showed that quotient rings of non-zero ideals in $\mathcal{O}_K$ are finite. We define the norm of the zero ideal to be zero. In this way the absolute ideal norm is multiplicative, and we will prove this fact soon. We will need to set up a linear algebra viewpoint of ideals in order to do so.

Let K be a number field and let $\mathfrak{i}$ be an ideal of $\mathcal{O}_K$. Let $n = [K : \mathbb{Q}]$ and fix a $\mathbb{Z}$-basis $\{a_1 \dots a_n\}$ for $\mathcal{O}_K$. Let $\{b_1 \dots b_n\}$ be a $\mathbb{Z}$-basis for $\mathfrak{i}$ and write

$$b_k = \sum_{g=1}^{n} s_{g,k} a_g$$

for each $k = 1 \dots n$ and some integers $s_{g,k}$. By using the $\mathbb{Z}$-basis $\{a_1 \dots a_n\}$ we may write any element $e = \sum_{k=1}^{n} s_k a_k$ of $\mathcal{O}_K$, where s_k are integers, as a column vector

$$l(e) = \begin{bmatrix} s_1 \\ s_2 \\ \dots \\ s_n \end{bmatrix}$$

Now consider the matrix

$$\begin{bmatrix} s_{1,1} & s_{1,2} & \dots & s_{1,n} \\ s_{2,1} & s_{2,2} & \dots & s_{2,n} \\ \dots & \dots & \dots & \dots \\ s_{n,1} & s_{n,2} & \dots & s_{n,n} \end{bmatrix}$$

The $\mathbb{Z}^n$-image of this matrix generates the set $\{l(e) : e \in \mathfrak{i}\}$. To see this, let $e = \sum_{k=1}^{n} r_k b_k$ be an

element of $\mathfrak{i}$ for integers r_i. Then we have

$$(\begin{bmatrix} s_{1,1} & s_{1,2} & \dots & s_{1,n} \\ s_{2,1} & s_{2,2} & \dots & s_{2,n} \\ \dots & \dots & \dots & \dots \\ s_{n,1} & s_{n,2} & \dots & s_{n,n} \end{bmatrix} \begin{bmatrix} r_1 \\ r_2 \\ \dots \\ r_n \end{bmatrix}) \cdot \begin{bmatrix} a_1 \\ a_2 \\ \dots \\ a_n \end{bmatrix} = \sum_{k=1}^{n} r_k b_K = e$$

so this matrix can be thought of as a manifestation of the ideal $\mathfrak{i}$. The *co-volume* of $\mathfrak{i}$, denoted covol($\mathfrak{i}$), is then defined as

$$\text{covol}(\mathfrak{i}) = |\det(\begin{bmatrix} s_{1,1} & s_{1,2} & \dots & s_{1,n} \\ s_{2,1} & s_{2,2} & \dots & s_{2,n} \\ \dots & \dots & \dots & \dots \\ s_{n,1} & s_{n,2} & \dots & s_{n,n} \end{bmatrix})|$$

This is a well-defined quantity because changing the $\mathbb{Z}$-basis of $\mathfrak{i}$ will not change the determinant of the matrix involved, since its $\mathbb{Z}^n$-image will be the same.

Lemma 1.3.3. *Let K be a number field with $[K : \mathbb{Q}] = n$ and let $\mathfrak{i}$ be an ideal of $\mathcal{O}_K$. Then $N(\mathfrak{i}) = \text{covol}(\mathfrak{i})$.*

Proof. Under the embedding $l : \mathcal{O}_K \to \mathbb{Z}^n$ defined above, ideals can be realized as n-dimensional lattices. Then the co-volume of an ideal is just the volume of the fundamental parallelepiped of the lattice corresponding to the ideal, by linear algebra.

By a simple geometrical argument, $N(\mathfrak{i}) = |\mathcal{O}_K/\mathfrak{i}|$ is simply the number of translates of $l(\mathfrak{i})$ required to cover $l(\mathcal{O}_K)$. How does this relate to the co-volumes of $\mathcal{O}_K$ and $\mathfrak{i}$? It means we require $N(\mathfrak{i})$ fundamental parallelepipeds of $l(\mathcal{O}_K)$ to cover $l(\mathfrak{i})$. As a result $N(\mathfrak{i})\text{covol}(\mathcal{O}_K) = \text{covol}(\mathfrak{i})$. It is easy to see that the embedding l is surjective, and so the volume of the fundamental parallelepiped of $\mathcal{O}_K$ under this embedding is 1. Then $N(\mathfrak{i}) = \text{covol}(\mathfrak{i})$. □

Proposition 1.3.4. *Let K be a number field with $[K : \mathbb{Q}] = n$, and let $\mathfrak{i}, \mathfrak{j}$ be ideals of $\mathcal{O}_K$. Then the absolute ideal norm is multiplicative. In other words*

$$N(\mathfrak{ij}) = N(\mathfrak{i}) \cdot N(\mathfrak{j})$$

Proof. We can assume that neither of the ideals $\mathfrak{j}$ or $\mathfrak{i}$ is the zero ideal or the whole ring, as then the statement would follow immediately. Fix a $\mathbb{Z}$-basis $\{a_1 \dots a_n\}$ for $\mathcal{O}_K$. By Lemma 1.3.3 it is sufficient to prove that

$$\text{covol}(\mathfrak{i})\text{covol}(\mathfrak{j}) = \text{covol}(\mathfrak{ij})$$

Let $[\mathfrak{ij} : \mathfrak{i}]$ denote the size of the additive group quotient quotient $\mathfrak{i}/\mathfrak{ij}$. Geometrically, this is the number of translates of the lattice $l(\mathfrak{ij})$ needed to cover $l(\mathfrak{i})$. As a result we have

$$[\mathfrak{ij} : \mathfrak{i}]\text{covol}(\mathfrak{i}) = \text{covol}(\mathfrak{ij})$$

We already know that $[\mathfrak{j} : \mathcal{O}_K] = \text{covol}(\mathfrak{j})$ by definition. Therefore it is sufficient to show that $[\mathfrak{ij} : \mathfrak{i}] = [\mathfrak{j} : \mathcal{O}_K]$. It is also sufficient to do so when $\mathfrak{j}$ is a prime ideal, as then we can factorize $\mathfrak{j}$ into prime ideals and then apply the result to each prime consequentially. We now follow [Cond, Theorem 7.5] to finish the proof.

Now $\mathcal{O}_K/\mathfrak{j}$ becomes a field and $\mathfrak{i}/\mathfrak{ij}$ becomes a vector space over $\mathcal{O}_K/\mathfrak{j}$, since $\mathfrak{j}$ annihilates $\mathfrak{i}/\mathfrak{ij}$. We have to show that $\dim(\mathfrak{i}/\mathfrak{ij}) = 1$. We have $\dim(\mathfrak{i}/\mathfrak{ij}) \geq 1$ because there exists some nonzero element $a \in \mathfrak{i} \backslash \mathfrak{ij}$ since $\mathfrak{i} \neq \mathfrak{ij}$ by UPF of ideals.

Since $\mathfrak{j}$ is prime and $(a) \subset \mathfrak{j}$ but $(a) \not\subset \mathfrak{ij}$ we must have $(a) + \mathfrak{ij} = \mathfrak{i}$. Therefore every element x in $\mathfrak{i}$ can be written as $x = y + ab$ for some $y \in \mathfrak{ij}$ and $b \in \mathcal{O}_K$. Then $x \equiv ab \pmod{\mathfrak{ij}}$ and it follows that a spans $\mathfrak{i}/\mathfrak{ij}$ over $\mathcal{O}_K$.

Let us write $b = mj + k$ for some $j \in \mathfrak{j}$ and $m, k \in \mathcal{O}_K$. Then $ab \equiv amj + ak \equiv ak \pmod{\mathfrak{ij}}$ since $aj \in \mathfrak{ij}$. Therefore a spans $\mathfrak{i}/\mathfrak{ij}$ as a vector space over $\mathcal{O}_K/\mathfrak{j}$. Hence $\dim(\mathfrak{i}/\mathfrak{ij}) = 1$ and we are done. □

Proposition 1.3.5. *Let K be a number field and let $\mathfrak{i}$ be an ideal of $\mathcal{O}_K$. Then*

- *$N(\mathfrak{i})$ is prime $\implies$ $\mathfrak{i}$ is a prime ideal.*
- *$\mathfrak{i}$ is a prime ideal $\implies$ $N(\mathfrak{i})$ is a prime power.*

Proof. The first statement follows from the fact that the only ideal with unit absolute norm is $\mathcal{O}_K$ itself, and the absolute norm is multiplicative by Proposition 1.3.4. Let $\mathfrak{i}$ be an ideal of $\mathcal{O}_K$ with $N(\mathfrak{i})$ prime. Suppose $\mathfrak{i}$ is not prime. Then it could be factored as a product of ideals $\mathfrak{i} = \mathfrak{ab}$ where neither of the ideals on the right are $\mathcal{O}_K$. Then $N(\mathfrak{i}) = N(\mathfrak{a})N(\mathfrak{b})$ where neither of the integers on the right are units, a contradiction.

To show the second statement, recall that quotient rings of prime ideals in the ring of integers of a number field are finite fields. It is a well known fact in field theory that the size of finite fields must be prime powers. □

Example 1.3.6. *Consider the ideal $\mathfrak{i} = (x + 126, x - 5)$ in the ring $R = \mathbb{Z}[x]/(x^3 + x + 1)$. The results above apply to this ring as well, since the only assumptions we used are that the ring involved is finitely generated over $\mathbb{Z}$. $\{1, x, x^2\}$ is clearly a $\mathbb{Z}$-basis for R. Now the set $\{x+126, x^2+126x, 126x^2 - x - 1, x - 5, x^2 - 5x, -5x^2 - x - 1\}$ must span $\mathfrak{i}$ over $\mathbb{Z}$. After doing some linear algebra, we get a $\mathbb{Z}$-basis for $\mathfrak{i}$. Here is the calculation:*

$$\begin{bmatrix} 126 & 1 & 0 \\ 0 & 126 & 1 \\ -1 & -1 & 126 \\ -5 & 1 & 0 \\ 0 & -5 & 1 \\ -1 & -1 & -5 \end{bmatrix} \to \begin{bmatrix} 0 & -125 & 15876 \\ 0 & 126 & 1 \\ 1 & 1 & -126 \\ 0 & 6 & -630 \\ 0 & -5 & 1 \\ 0 & 0 & -131 \end{bmatrix} \to \begin{bmatrix} 0 & 0 & -62749 \\ 0 & 0 & 79255 \\ 1 & 0 & 503 \\ 0 & 1 & -629 \\ 0 & 0 & -3144 \\ 0 & 0 & -131 \end{bmatrix} \to \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 114 \\ 0 & 1 & 21 \\ 0 & 0 & 0 \\ 0 & 0 & 131 \end{bmatrix}$$

The determinant of the rightmost matrix minus the zero rows is 131*, which is the norm of the ideal* $\mathfrak{i}$ *by previous results. This is prime and so the ideal* $\mathfrak{i}$ *is a prime ideal.*

Suppose we have a number field extension $L : K$ and an ideal $\mathfrak{i}$ of $\mathcal{O}_K$. Then we can realize $\mathfrak{i}$ as an ideal of $\mathcal{O}_L$ as follows. Write $\mathfrak{i} = (a_1 \dots a_n)$ for some $a_i \in \mathcal{O}_K$. Then $\mathfrak{i}$ as an ideal of $\mathcal{O}_L$ is simply the $\mathcal{O}_L$-module generated by $\{a_1 \dots a_n\}$. This is called the *extension* of $\mathfrak{i}$ to $\mathcal{O}_L$ and is denoted $\mathfrak{i}\mathcal{O}_L$.

Conversely, suppose we have an ideal $\mathfrak{i}$ of $\mathcal{O}_L$. Then its *contraction* to $\mathcal{O}_K$ is simply defined as the ideal $\mathfrak{i} \cap \mathcal{O}_K$. We leave it up to the reader to show that extension and contraction respects multiplication of ideals.

Definition 1.3.7. *Let* $L : K$ *be a number field. We say* L *is monogenic over* K *if* $\mathcal{O}_L = \mathcal{O}_K[\theta]$ *for some algebraic integer* θ*.*

We now prove a result called the Dedekind–Kummer theorem. This is the bread and butter of ideal factorization, even though it only applies to monogenic extensions of number fields. Given a monogenic number field extension $L : K$ and an element θ as above, we can factor the extensions of primes of K into primes of L using the minimal polynomial of θ. This theorem will be generalized to all number field extensions in Section 1.5.

Theorem 1.3.8. *Let* $L : K$ *be a monogenic number field extension, so that* $\mathcal{O}_L = \mathcal{O}_K[\theta]$ *for some algebraic integer* θ*. Let* f *be the minimal polynomial of* θ *over* K *and let* $\mathfrak{p}$ *be a prime ideal of* $\mathcal{O}_K$*. Let's say we can factorize* f *into irreducible factors modulo* $\mathfrak{p}\mathcal{O}_K$ *as*

$$\overline{f} \equiv \prod_{j=1}^{m} \overline{f_j}^{e_j} \pmod{\mathfrak{p}\mathcal{O}_K}$$

Then we get a corresponding factorization of the ideal $\mathfrak{p}$ *extended to* $\mathcal{O}_L$ *into prime ideals as*

$$\mathfrak{p}\mathcal{O}_L = \prod_{j=1}^{m} (\mathfrak{p}\mathcal{O}_L + (f_j(\theta)))^{e_j}$$

Proof. There is a canonical isomorphism $\mathcal{O}_L \cong \mathcal{O}_K[x]/(f)$ by sending θ to x. Let $\mathfrak{p}_x$ be the image of the prime ideal $\mathfrak{p}\mathcal{O}_L$ under this isomorphism. Then we have $\mathcal{O}_L/\mathfrak{p}\mathcal{O}_L \cong \mathcal{O}_K[x]/((f) + \mathfrak{p}_x)$. We use a bar to denote reduction modulo $\mathfrak{p}\mathcal{O}_L$ or $((f) + \mathfrak{p}_x)$. From the assumptions, there is a factorization of ideals in $\mathcal{O}_L/\mathfrak{p}\mathcal{O}_L$ as

$$(\overline{f(\theta)}) = (0) = \prod_{j=1}^{m} (\overline{f_j(\theta)})^{e_j}$$

Applying the Chinese remainder theorem then yields

$$(\mathcal{O}_L/\mathfrak{p}\mathcal{O}_L)/(0) \cong \mathcal{O}_L/\mathfrak{p}\mathcal{O}_L \cong \prod_{j=1}^{m} (\mathcal{O}_L/\mathfrak{p}\mathcal{O}_L)/(\overline{f_j(\theta)})^{e_j} \cong \prod_{j=1}^{m} \mathcal{O}_L/(\mathfrak{p}\mathcal{O}_L + (f_j(\theta))^{e_j})$$

Furthermore, $\mathcal{O}_K/\mathfrak{p}$ is a field, and as a result $\mathcal{O}_K[x]/\mathfrak{p}_x$ is a PID because it is a polynomial ring over a field with transcendence degree 1. It follows that $\mathcal{O}_L/\mathfrak{p}\mathcal{O}_L \cong \mathcal{O}_K[x]/((f)+\mathfrak{p}_x)$ is also a PID. Therefore the ideals $(\overline{f_j(\theta)})$, being generated by irreducible elements, are prime in $\mathcal{O}_L/\mathfrak{p}\mathcal{O}_L$. Hence $\mathcal{O}_L/(\mathfrak{p}\mathcal{O}_L + (f_j(\theta)))$ are also fields and so $\mathfrak{p}\mathcal{O}_L + (f_j(\theta))$ are prime ideals of $\mathcal{O}_L$.

The containment $(\mathfrak{p}\mathcal{O}_L + (f_j(\theta)))^{e_j} \subset \mathfrak{p}\mathcal{O}_L + (f_j(\theta))^{e_j}$ is clear and thus $\mathfrak{p}\mathcal{O}_L + (f_j(\theta))^{e_j} = (\mathfrak{p}\mathcal{O}_L+(f_j(\theta)))^{r_j}$ for some $r_j \leq e_j$ by UPF of ideals. However there is a chain of proper containments

$$\mathfrak{p}\mathcal{O}_L + (f_j(\theta)) \subsetneq \mathfrak{p}\mathcal{O}_L + (f_j(\theta))^2 \subsetneq \cdots \subsetneq \mathfrak{p}\mathcal{O}_L + (f_j(\theta))^{e_j}$$

so finally we must have $r_j = e_j$ and $(\mathfrak{p}\mathcal{O}_L + (f_j(\theta)))^{e_j} = \mathfrak{p}\mathcal{O}_L + (f_j(\theta))^{e_j}$. This makes

$$\mathcal{O}_L/\mathfrak{p}\mathcal{O}_L \cong \prod_{j=1}^{m} \mathcal{O}_L/(\mathfrak{p}\mathcal{O}_L + (f_j(\theta)))^{e_j}$$

Finally $\mathfrak{p}\mathcal{O}_L = \prod_{j=1}^{m}(\mathfrak{p}\mathcal{O}_L + (f_j(\theta)))^{e_j}$ by the Chinese remainder theorem. □

Example 1.3.9. *Let us see an example of the above when the base field is $\mathbb{Q}$. Let $L = \mathbb{Q}(\sqrt{67})$ in which case $\mathcal{O}_L = \mathbb{Z}[\sqrt{67}]$ and so this is a monogenic extension. The minimal polynomial of $\theta = \sqrt{67}$ is $f = x^2 - 67$. Let's factorize the ideal $\mathfrak{i} = (10 + \sqrt{67})$ in $\mathcal{O}_L$ using what we learned so far.*

The norm of this ideal can be calculated as 33. In particular $(33) \subset \mathfrak{i}$ so we can factor $\mathfrak{i}$ by factoring the primes $(3), (11)$ using the above result. We have

$$x^2 - 67 \equiv (x-1)(x+1) \pmod 3$$

$$x^2 - 67 \equiv (x-1)(x+1) \pmod{11}$$

The Dedekind–Kummer theorem then gives $(3) = (3, \sqrt{67} - 1)(3, \sqrt{67} + 1)$ and $(11) = (11, \sqrt{67} - 1)(11, \sqrt{67} + 1)$. After a finite number of checks we arrive at

$$(3, \sqrt{67} + 1)(11, \sqrt{67} - 1) = (33, 66, 3\sqrt{67} - 3, 11\sqrt{67} + 11) = (10 + \sqrt{67})$$

1.4 Decomposition of primes in field extensions.

Let $L : K$ be a number field extension. This automatically implies that $\mathcal{O}_K \subset \mathcal{O}_L$. Prime ideals of $\mathcal{O}_K$ may no longer be prime when extended to $\mathcal{O}_L$, as we saw in the previous subsection. The Dedekind–Kummer theorem gave a method to factorize the extension of prime ideals in monogenic extensions. In this subsection we develop further results about the prime factorization of a prime ideal of $\mathcal{O}_K$ extended to $\mathcal{O}_L$. This is called decomposition of primes. The material in this section is explained in [Cox13, 5.A] without proof.

Definition 1.4.1. *Let $L : K$ be a number field extension and let $\mathfrak{p}$ be a prime ideal of $\mathcal{O}_K$. Let $\mathfrak{P}$ be a prime ideal of $\mathcal{O}_L$ containing $\mathfrak{p}\mathcal{O}_L$. Then $\mathfrak{P}$ is called a prime above $\mathfrak{p}$ in the extension $L : K$, and $\mathfrak{p}$ is called a prime below $\mathfrak{P}$ in the extension $L : K$.*

Proposition 1.4.2. *Let $L : K$ be a number field extension. Then for any prime ideal $\mathfrak{P}$ of $\mathcal{O}_L$, there is a unique prime $\mathfrak{p}$ of $\mathcal{O}_K$ below $\mathfrak{P}$.*

Proof. We claim that $\mathfrak{p} = \mathfrak{P} \cap \mathcal{O}_K$. Assume it is not prime. Then we can write $\mathfrak{p} = \mathfrak{ij}$ for some ideals $\mathfrak{i}, \mathfrak{j}$ of $\mathcal{O}_K$ not equal to $\mathcal{O}_K$. When we extend the ideals $\mathfrak{p}, \mathfrak{i}, \mathfrak{j}$ to $\mathcal{O}_L$ we have

$$\mathfrak{ij}\mathcal{O}_L \subset \mathfrak{p}\mathcal{O}_L \subset \mathfrak{P}$$

$\mathfrak{P}$ is a prime ideal so $\mathfrak{i}\mathcal{O}_L \subset \mathfrak{P}$ or $\mathfrak{j}\mathcal{O}_L \subset \mathfrak{P}$. Let's say $\mathfrak{i}\mathcal{O}_L \subset \mathfrak{P}$ and so $\mathfrak{i}\mathcal{O}_L \cap \mathcal{O}_K \subset \mathfrak{P} \cap \mathcal{O}_K = \mathfrak{p}$. Since $\mathfrak{i}\mathcal{O}_L \cap \mathcal{O}_K = \mathfrak{i}$, we have the containment $\mathfrak{i} \subset \mathfrak{p}$ and so $\mathfrak{p} = \mathfrak{i}$. However, this contradicts $\mathfrak{j} \neq \mathcal{O}_K$. It follows that $\mathfrak{p}$ is a prime. In particular, it is a prime below $\mathfrak{P}$.

Suppose $\mathfrak{q}$ is another prime below $\mathfrak{P}$. Then $\mathfrak{p} = \mathfrak{P} \cap \mathcal{O}_K \subset \mathfrak{q}$. But this implies $\mathfrak{q} = \mathfrak{p}$ since we are in a Dedekind domain. This proves uniqueness. □

Definition 1.4.3. *Let $L : K$ be a number field extension. Let $\mathfrak{P}_1$ be a prime ideal of $\mathcal{O}_L$ and let $\mathfrak{p}$ be the unique prime below $\mathfrak{P}_1$. Let $\mathfrak{p}\mathcal{O}_L = \prod_{i=1}^{n} \mathfrak{P}_i^{e_i}$ be the unique prime factorization of $\mathfrak{p}\mathcal{O}_L$ into prime ideals in $\mathcal{O}_L$. The ramification degree of $\mathfrak{P}_1$ in the extension $L : K$ is defined to be*

$$e_{L:K}(\mathfrak{P}_1) := e_1$$

and the inertia degree of $\mathfrak{P}_1$ in the extension $L : K$ is defined to be

$$f_{L:K}(\mathfrak{P}_1) := [\mathcal{O}_L/\mathfrak{P}_1 : \mathcal{O}_K/\mathfrak{p}]$$

Proposition 1.4.4. *Let $M : L : K$ be number field extensions and let $\mathfrak{p}$ be a prime ideal of $\mathcal{O}_K$. Let $\mathfrak{P}$ be a prime above $\mathfrak{p}$ in L and let $\mathcal{P}$ be a prime above $\mathfrak{P}$ in M. Then the inertia and ramification degrees are multiplicative in the sense that*

$$f_{M:K}(\mathcal{P}) = f_{M:L}(\mathcal{P}) \cdot f_{L:K}(\mathfrak{P})$$

$$e_{M:K}(\mathcal{P}) = e_{M:L}(\mathcal{P}) \cdot e_{L:K}(\mathfrak{P})$$

Proof. The multiplicativity of inertia degrees is equivalent to

$$[\mathcal{O}_M/\mathcal{P} : \mathcal{O}_K/\mathfrak{p}] = [\mathcal{O}_M/\mathcal{P} : \mathcal{O}_L/\mathfrak{P}][\mathcal{O}_L/\mathfrak{P} : \mathcal{O}_K/\mathfrak{p}]$$

which is a result in field theory. By Proposition 1.4.2, $\mathfrak{P}\mathcal{O}_M$ is the only prime ideal in $\mathcal{O}_L$ that $\mathcal{P}$ divides, and it does so with multiplicity $e_{M:L}(\mathcal{P})$. Meanwhile, $\mathfrak{P}$ divides $\mathfrak{p}\mathcal{O}_L$ with multiplicity $e_{L:K}(\mathfrak{P})$. Multiplicities are multiplicative, so $\mathcal{P}$ divides $\mathfrak{p}\mathcal{O}_M$ with multiplicity $e_{M:L}(\mathcal{P})e_{L:K}(\mathfrak{P})$. But this also equals the ramification degree $e_{M:K}(\mathcal{P})$. □

Proposition 1.4.5. *Let $L : K$ be a number field extension. Let $\mathfrak{p}$ be a prime ideal of $\mathcal{O}_K$. Let $\mathfrak{p}\mathcal{O}_L = \prod_{i=1}^{n} \mathfrak{P}_i^{e_i}$ be the unique prime factorization of $\mathfrak{p}\mathcal{O}_L$ into prime ideals in $\mathcal{O}_L$. Then*

$$\sum_{i=1}^{n} e_{L:K}(\mathfrak{P}_i) \cdot f_{L:K}(\mathfrak{P}_i) = [L : K]$$

Proof. Firstly, by restriction of scalars we have

$$N(\mathfrak{p}\mathcal{O}_L) = |\mathcal{O}_L/\mathfrak{p}\mathcal{O}_L| = |\mathcal{O}_K/\mathfrak{p}|^{[L:K]} = N(\mathfrak{p})^{[L:K]}$$

Then using our definition of inertia degree we get

$$N(\mathfrak{P}_i) = |\mathcal{O}_L/\mathfrak{P}_i| = |\mathcal{O}_K/\mathfrak{p}|^{[\mathcal{O}_L/\mathfrak{P}_1:\mathcal{O}_K/\mathfrak{p}]} = N(\mathfrak{p})^{[\mathcal{O}_L/\mathfrak{P}_1:\mathcal{O}_K/\mathfrak{p}]}$$

Applying the absolute ideal norm to the prime factorization of $\mathfrak{p}\mathcal{O}_L$ yields

$$N(\prod_{i=1}^{n}\mathfrak{P}_i^{e_i}) = \prod_{i=1}^{n}\cdot N(\mathfrak{P}_i)^{e_{L:K}(\mathfrak{p}_i)} = \prod_{i=1}^{n} N(\mathfrak{p})^{e_{L:K}(\mathfrak{p}_i)\cdot f_{L:K}(\mathfrak{p}_i)}$$

Putting everything together gives us

$$N(\mathfrak{p})^{[L:K]} = N(\mathfrak{p}\mathcal{O}_L) = N(\prod_{i=1}^{n}\mathfrak{P}_i^{e_i}) = \prod_{i=1}^{n} N(\mathfrak{p})^{e_{L:K}(\mathfrak{p}_i)\cdot f_{L:K}(\mathfrak{p}_i)}$$

Staring at the exponents yields $[L:K] = \sum_{i=1}^{n} e_{L:K}(\mathfrak{P}_i)\cdot f_{L:K}(\mathfrak{P}_i)$. □

We will now draw our attention to Galois extensions, where we can say more about the decomposition of primes in terms of Galois actions. Let $L:K$ be a Galois extension and let $\mathfrak{P}$ be a prime ideal of $\mathcal{O}_L$. Let $\mathfrak{p} = \mathfrak{P}\cap\mathcal{O}_K$ be the unique prime below $\mathfrak{P}$ and let $\sigma\in\mathrm{Gal}(L:K)$. Then

$$\mathfrak{p}\mathcal{O}_L \subset \mathfrak{P} \implies \sigma(\mathfrak{p}\mathcal{O}_L) = \mathfrak{p}\mathcal{O}_L \subset \sigma(\mathfrak{P})$$

so $\sigma(\mathfrak{P})$ is another prime ideal above $\mathfrak{p}$. To show that it is prime, suppose $\sigma(\mathfrak{P}) = \mathfrak{ij}$. Then we have $\mathfrak{P} = \sigma^{-1}(\mathfrak{i})\sigma^{-1}(\mathfrak{j})$ and so either $\sigma^{-1}(\mathfrak{i})$ or $\sigma^{-1}(\mathfrak{j})$ must be the entire ring $\mathcal{O}_L$. It follows that $\mathfrak{i} = \mathcal{O}_L$ or $\mathfrak{j} = \mathcal{O}_L$ and we conclude that $\sigma(\mathfrak{P})$ is a prime ideal. Therefore $\mathrm{Gal}(L:K)$ acts on the prime ideals above $\mathfrak{p}$. The next result shows that this action is transitive.

Proposition 1.4.6. *Let $L:K$ be a Galois extension and let $\mathfrak{p}$ be a prime ideal of $\mathcal{O}_K$. Let $\mathfrak{p}\mathcal{O}_L = \prod_{i=1}^{n}\mathfrak{P}_i^{e_i}$ be its prime factorization into prime ideals of $\mathcal{O}_L$. Then for any pair of prime ideals $\mathfrak{P}_i, \mathfrak{P}_j$ there is some element $\sigma\in\mathrm{Gal}(L:K)$ such that $\sigma(\mathfrak{P}_i) = \mathfrak{P}_j$.*

Proof. Suppose that is not the case for a proof by contradiction. Then by the Chinese remainder theorem there is an element $a\in\mathcal{O}_L$ satisfying

- $a \equiv 0 \pmod{\mathfrak{P}_i}$
- $a \equiv 1 \pmod{\sigma(\mathfrak{P}_j)}\ \forall\sigma\in\mathrm{Gal}(L:K)$

since none of the prime ideals $\sigma(\mathfrak{P}_j)$ are equal to $\mathfrak{P}_i$. However

$$N_{L:K}(a) = \prod_{\sigma\in\mathrm{Gal}(L:K)} \sigma(a) \in \mathfrak{P}_i\cap\mathcal{O}_k = \mathfrak{p}$$

so in particular $N_{L:K}(a) \subset \mathfrak{p}\mathcal{O}_L \subset \mathfrak{P}_j$. Therefore $N_{L:K}(a) \equiv 0 \pmod{\mathfrak{P}_j}$ which means that $\sigma(a) \equiv 0 \pmod{\mathfrak{P}_j}$ for some $\sigma \in \mathrm{Gal}(L:K)$, because $\mathcal{O}_L/\mathfrak{P}_j$ is an integral domain. As a result $a \equiv 0 \pmod{\sigma^{-1}(\mathfrak{P}_j)}$, which contradicts the Chinese remainder theorem. We must therefore have $\sigma(\mathfrak{P}_i) = \mathfrak{P}_j$ for some σ. □

Proposition 1.4.7. *Let $L:K$ be a Galois extension and let $\mathfrak{p}$ be a prime ideal of $\mathcal{O}_K$. Let $\mathfrak{P}_i$ and $\mathfrak{P}_j$ be two primes above $\mathfrak{p}$. Then we have*

$$e_{L:K}(\mathfrak{P}_i) = e_{L:K}(\mathfrak{P}_j)$$

$$f_{L:K}(\mathfrak{P}_i) = f_{L:K}(\mathfrak{P}_j)$$

Proof. By Proposition 1.4.6, there is some $\sigma \in \mathrm{Gal}(L:K)$ so that $\sigma(\mathfrak{P}_i) = \mathfrak{P}_j$. Then $\mathfrak{p}\mathcal{O}_L = \sigma(\mathfrak{p}\mathcal{O}_L) = \prod_{k=1}^{n} \sigma(\mathfrak{P}_k)^{e_k}$ and therefore $\sigma(\mathfrak{P}_i)^{e_i} = \mathfrak{P}_j^{e_i} = \mathfrak{P}_j^{e_j}$ by UPF of ideals. It follows that $e_{L:K}(\mathfrak{P}_i) = e_{L:K}(\mathfrak{P}_j)$.

Galois actions on ideals induce homomorphisms of their quotient rings as

$$\sigma : \mathcal{O}_L/\mathfrak{P}_i \to \mathcal{O}_L/\mathfrak{P}_j$$

Since $\sigma(\mathfrak{P}_i) \subset \mathfrak{P}_j$. However, this homomorphism is invertible since $\mathrm{Gal}(L:K)$ is a group, and so the homomorphism above is actually an isomorphism. It follows that $N(\mathfrak{P}_i) = N(\mathfrak{P}_j)$ since the size of their quotient rings must be the same. Hence $|\mathcal{O}_K/\mathfrak{p}|^{f_{L:K}(\mathfrak{P}_j)} = |\mathcal{O}_K/\mathfrak{p}|^{f_{L:K}(\mathfrak{P}_i)}$ so it follows that $f_{L:K}(\mathfrak{P}_i) = f_{L:K}(\mathfrak{P}_j)$. □

From the above proposition, we see that in a Galois extension $L:K$, the inertia and ramification degrees of a prime $\mathfrak{P}_i$ in L is determined solely by their unique prime $\mathfrak{p}$ below. We can therefore denote $e_{L:K}(\mathfrak{p}) := e_{L:K}(\mathfrak{P}_i)$ and $f_{L:K}(\mathfrak{p}) = f_{L:K}(\mathfrak{P}_i)$ when we work with Galois extensions.

Corollary 1.4.8. *Let $L:K$ be a Galois extension and let $\mathfrak{p}$ be a prime ideal of $\mathcal{O}_K$. Let $\mathfrak{p}\mathcal{O}_L = \prod_{k=1}^{n} \mathfrak{P}_k^{e_k}$ be its prime factorization into prime ideals of $\mathcal{O}_L$. Then we have*

$$[L:K] = n \cdot e_{L:K}(\mathfrak{p}) \cdot f_{L:K}(\mathfrak{p})$$

Example 1.4.9. *Let $f = x^3 - 3x + 1$. Then $K = \mathbb{Q}[x]/(f)$ turns out to be a Galois extension of $\mathbb{Q}$. The formula from the corollary above tells us that for every prime ideal $\mathfrak{p}$ of $\mathbb{Z}$, one of the following cases occurs with the same notation:*

$$e_{K:\mathbb{Q}}(\mathfrak{p}) = 3, f_{K:\mathbb{Q}}(\mathfrak{p}) = 1, n = 1$$

$$e_{K:\mathbb{Q}}(\mathfrak{p}) = 1, f_{K:\mathbb{Q}}(\mathfrak{p}) = 3, n = 1$$

$$e_{K:\mathbb{Q}}(\mathfrak{p}) = 1, f_{K:\mathbb{Q}}(\mathfrak{p}) = 1, n = 3$$

We will look at a prime ideal which splits completely (case 3) when extended to K. It turns out that we can apply the Dedekind–Kummer theorem to this extension. We factorize the ideal $(19)\mathcal{O}_K$ as

$$x^3 - 3x + 1 \equiv (x-3)(x-7)(x-10) \pmod{19}$$

$$(19)\mathcal{O}_K = (19, x-3)(19, x-7)(19, x-10)$$

Then the Galois group should act transitively on the ideals on the LHS.

1.5 Orders of number fields in general

Here we will describe some tools to find the rings of integers of number fields. The idea is to start with an order, which we will define shortly, and then gradually extend the order until we get the entire ring of integers. We will introduce the discriminant, which will act as a measure of the size of an order, and allow us to find the ring of integers in a finite number of steps, since each algebraic integer we add to our order will decrease the discriminant of the order by some factor. We will also study orders for their own sake and apply their theory to produce a generalization of the Dedekind–Kummer theorem.

Definition 1.5.1. *Let K be a number field. An order $\mathcal{O}$ of K is defined to be a subring of K that*

- *has fraction field equal to K*
- *is a finitely generated $\mathbb{Z}$-module*

Proposition 1.5.2. *Let K be a number field with $[K : \mathbb{Q}] = n$. Then $\mathcal{O}$ is an order of K if and only if it is a subring of $\mathcal{O}_K$ whose rank as a free $\mathbb{Z}$-module is n.*

Proof. We prove the forward direction. Let $a \in \mathcal{O}$. Then $\mathbb{Z}[a]$ is a finitely generated $\mathbb{Z}$-module because it is a subring of $\mathcal{O}$. Clearly $a\mathbb{Z}[a] \subset \mathbb{Z}[a]$ and $\mathbb{Z} \subset \mathbb{Z}[a]$. By applying Theorem 1.2.3 we get that a is an algebraic integer. This occurs for every $a \in \mathcal{O}$ and so $\mathcal{O} \subset \mathcal{O}_K$. A submodule of a free $\mathbb{Z}$-module is free so $\mathcal{O}$ is also free. Let's say $\{a_1 \dots a_k\}$ is a $\mathbb{Z}$-basis for $\mathcal{O}$. Then $\mathrm{Frac}(\mathcal{O}) = K$ and so $\{a_1 \dots a_k\}$ is in fact also a $\mathbb{Q}$-basis for K. It follows that $k = n = rank(\mathcal{O})$.

Now we prove the other direction. Let $\mathcal{O}$ be a subring of $\mathcal{O}_K$ whose rank as a free $\mathbb{Z}$-module is n. A submodule of a finitely generated $\mathbb{Z}$-module is also finitely generated, and so $\mathcal{O}$ is a finitely generated $\mathbb{Z}$-module. It remains to show that $\mathrm{Frac}(\mathcal{O}) = K$. We have $rank(\mathcal{O}) = n$ and so there is a set $\{a_1 \dots a_n\} \subset \mathcal{O}$ which is linearly independent over $\mathbb{Z}$. This implies that $\{a_1 \dots a_n\}$ is linearly independent over $\mathbb{Q}$. By linear algebra $\{a_1 \dots a_n\}$ is in fact a $\mathbb{Q}$-basis for K and so indeed $\mathrm{Frac}(\mathcal{O}) = K$. □

The ring of integers of a number field is sometimes referred to as the maximal order. The above proposition makes sense of this fact, because every order is contained in the ring of integers.

Example 1.5.3. *Let us see a simple example of an order. Let* $K = \mathbb{Q}(\sqrt{5})$. *Then* $\mathbb{Z}[\sqrt{5}]$ *is an order of* K, *since* $\mathrm{Frac}(\mathbb{Z}[\sqrt{5}]) = K$ *and* $\{1, \sqrt{5}\}$ *is a* $\mathbb{Z}$*-basis for* $\mathbb{Z}[\sqrt{5}]$ *so it is finitely generated. However, as we shall see in Section* 2, *the maximal order of* K *is* $\mathbb{Z}[\frac{1+\sqrt{5}}{2}]$.

Note that $\mathbb{Z}$ *is not an order of* K, *since* $\mathrm{Frac}(\mathbb{Z}) = \mathbb{Q} \neq K$.

Definition 1.5.4. *Let* K *be a number field with* $[K : \mathbb{Q}] = n$ *and* $\mathcal{O}$ *an order of* K. *Let* $\sigma_1 \dots \sigma_n$ *be the embeddings of* K *into* $\mathbb{C}$. *Let* $\omega_1 \dots \omega_n$ *be a* $\mathbb{Z}$*-basis for* $\mathcal{O}$. *We define the discriminant of* $\mathcal{O}$ *as*

$$\Delta(\mathcal{O}) := \det \begin{bmatrix} \sigma_1(\omega_1) & \sigma_1(\omega_2) & \dots & \sigma_1(\omega_n) \\ \sigma_2(\omega_1) & \sigma_2(\omega_2) & \dots & \sigma_2(\omega_n) \\ \dots & \dots & \dots & \dots \\ \sigma_n(\omega_1) & \sigma_n(\omega_2) & \dots & \sigma_n(\omega_n) \end{bmatrix}^2$$

Proposition 1.5.5. *Let* K *be a number field and let* $\mathcal{O}$ *be an order of* K. *The discriminant of* $\mathcal{O}$ *is well-defined and it is an integer.*

Proof. Suppose we have a different $\mathbb{Z}$-basis for the order in the definition. Let it be $\{r_1 \dots r_n\}$. Let the integer change of basis matrix from $\{r_1 \dots r_n\}$ to $\{\omega_1 \dots \omega_n\}$ be M so that

$$\begin{bmatrix} \omega_1 \\ \omega_2 \\ \dots \\ \omega_n \end{bmatrix} = M \begin{bmatrix} r_1 \\ r_2 \\ \dots \\ r_n \end{bmatrix}$$

M must be invertible in $GL_n(\mathbb{Z})$ since we can also specify an integer change of basis matrix from $\{\omega_1 \dots \omega_n\}$ to $\{r_1 \dots r_n\}$ which is the inverse of M. Hence the determinant of M must be a unit of $\mathbb{Z}$, so we must have $\det(M)^2 = 1$. We also have

$$\begin{bmatrix} \sigma_1(\omega_1) & \sigma_1(\omega_2) & \dots & \sigma_1(\omega_n) \\ \sigma_2(\omega_1) & \sigma_2(\omega_2) & \dots & \sigma_2(\omega_n) \\ \dots & \dots & \dots & \dots \\ \sigma_n(\omega_1) & \sigma_n(\omega_2) & \dots & \sigma_n(\omega_n) \end{bmatrix} = \begin{bmatrix} \sigma_1(r_1) & \sigma_1(r_2) & \dots & \sigma_1(r_n) \\ \sigma_2(r_1) & \sigma_2(r_2) & \dots & \sigma_2(r_n) \\ \dots & \dots & \dots & \dots \\ \sigma_n(r_1) & \sigma_n(r_2) & \dots & \sigma_n(r_n) \end{bmatrix} M^T$$

and so it is clear upon taking the square determinant of both sides, that the discriminant is the same when calculated using different integral bases. Hence it is well-defined.

Now we prove the discriminant is an integer. Let's apply an embedding σ_i to the matrix in question. We get

$$\sigma_i\left(\begin{bmatrix} \sigma_1(\omega_1) & \sigma_1(\omega_2) & \dots & \sigma_1(\omega_n) \\ \sigma_2(\omega_1) & \sigma_2(\omega_2) & \dots & \sigma_2(\omega_n) \\ \dots & \dots & \dots & \dots \\ \sigma_n(\omega_1) & \sigma_n(\omega_2) & \dots & \sigma_n(\omega_n) \end{bmatrix} \right) = \begin{bmatrix} \sigma_i\sigma_1(\omega_1) & \sigma_i\sigma_1(\omega_2) & \dots & \sigma_i\sigma_1(\omega_n) \\ \sigma_i\sigma_2(\omega_1) & \sigma_i\sigma_2(\omega_2) & \dots & \sigma_i\sigma_2(\omega_n) \\ \dots & \dots & \dots & \dots \\ \sigma_i\sigma_n(\omega_1) & \sigma_i\sigma_n(\omega_2) & \dots & \sigma_i\sigma_n(\omega_n) \end{bmatrix}$$

The embeddings form a groupoid so applying σ_i just permutes the rows of this matrix. The determinant of the matrix is changed by a factor of ± 1. Since we are taking the square of the determinant, the discriminant of $\mathcal{O}$ is unaffected. Therefore the discriminant must be a rational number, because it is fixed by all embeddings. However, the matrix entries are all algebraic integers, and so is the discriminant. Hence the discriminant is an integer. $\square$

Proposition 1.5.6. *Let K be a number field and let $\mathcal{O}, \mathcal{O}'$ be orders of K so that $\mathcal{O} \subset \mathcal{O}'$. Then $\frac{\Delta(\mathcal{O})}{\Delta(\mathcal{O}')}$ is a perfect integer square.*

Proof. Let $\{\omega_1 \dots \omega_n\}$ be a $\mathbb{Z}$-basis for $\mathcal{O}$ and $\{r_1 \dots r_n\}$ a $\mathbb{Z}$-basis for $\mathcal{O}'$. Since $\mathcal{O} \subset \mathcal{O}'$, there is an integer matrix M so that

$$\begin{bmatrix} \omega_1 \\ \omega_2 \\ \dots \\ \omega_n \end{bmatrix} = M \begin{bmatrix} r_1 \\ r_2 \\ \dots \\ r_n \end{bmatrix}$$

In terms of discriminants this says that

$$\begin{bmatrix} \sigma_1(\omega_1) & \sigma_1(\omega_2) & \dots & \sigma_1(\omega_n) \\ \sigma_2(\omega_1) & \sigma_2(\omega_2) & \dots & \sigma_2(\omega_n) \\ \dots & \dots & \dots & \dots \\ \sigma_n(\omega_1) & \sigma_n(\omega_2) & \dots & \sigma_n(\omega_n) \end{bmatrix} = \begin{bmatrix} \sigma_1(r_1) & \sigma_1(r_2) & \dots & \sigma_1(r_n) \\ \sigma_2(r_1) & \sigma_2(r_2) & \dots & \sigma_2(r_n) \\ \dots & \dots & \dots & \dots \\ \sigma_n(r_1) & \sigma_n(r_2) & \dots & \sigma_n(r_n) \end{bmatrix} M^T$$

so that after taking determinants, $\Delta(\mathcal{O}) = \Delta(\mathcal{O}')\det(M)^2$ and the result follows. $\square$

Remark 1.5.7. *If $\mathcal{O}, \mathcal{O}'$ are orders of K so that $\mathcal{O} \subset \mathcal{O}'$ and $\Delta(\mathcal{O}) = \Delta(\mathcal{O}')$, then $\mathcal{O} = \mathcal{O}'$. This is because in this case the integer matrix M has determinant 1 or -1, and is therefore invertible. It follows that it is a change of basis matrix.*

Example 1.5.8. *Let $K = \mathbb{Q}(\sqrt{5})$. Recall that the order $\mathbb{Z}[\sqrt{5}]$ has a $\mathbb{Z}$-basis $\{1, \sqrt{5}\}$ and so it has discriminant*

$$\Delta(\mathbb{Z}[\sqrt{5}]) = \det(\begin{bmatrix} 1 & \sqrt{5} \\ 1 & -\sqrt{5} \end{bmatrix})^2 = (-2\sqrt{5})^2 = 20$$

The only square factor of 20 is 4, and so the maximal order could potentially have discriminant 5. This turns out to be the case, since the ring of integers of K is $\mathbb{Z}[\frac{1+\sqrt{5}}{2}]$ and so

$$\Delta(\mathcal{O}_K) = \det(\begin{bmatrix} 1 & \frac{1+\sqrt{5}}{2} \\ 1 & \frac{1-\sqrt{5}}{2} \end{bmatrix})^2 = (-\sqrt{5})^2 = 5$$

We will now describe a process for finding the ring of integers of a number field K, inspired by the example above. This will take a finite number of steps given a starting order $\mathcal{O}$ because $\Delta(\mathcal{O})$ has a finite number of square prime factors. Let $\{r_1 \dots r_n\}$ be an integral basis for $\mathcal{O}$.

Since the fraction field of $\mathcal{O}$ is K, every algebraic integer in K will take the form $x = \frac{\sum_{i=1}^{n} a_i r_i}{m}$ for a_i and m integers. We can suppose $(a_j, m) = 1$ for some j. By adding integer multiples of r_i to x, we can assume that the coefficients a_i lie in $\{0 \dots m-1\}$. Then, by multiplying x by some integer, we can assume additionally that $a_j = 1$. In this case x can replace r_j to form the integral basis $\{r_1 \dots x \dots r_n\}$ of an extended order $\mathcal{O}'$.

Note then that $\{r_1 \dots mx \dots r_n\}$ is an integral basis for $\mathcal{O}$. Then $\Delta(\mathcal{O}) = m^2 \cdot \Delta(\mathcal{O}')$ because

$$\det(\begin{bmatrix} \sigma_1(r_1) & \dots & \sigma_1(mx) & \dots & \sigma_1(r_n) \\ \sigma_2(r_1) & \dots & \sigma_2(mx) & \dots & \sigma_2(r_n) \\ \dots & \dots & \dots & \dots & \dots \\ \sigma_n(r_1) & \dots & \sigma_n(mx) & \dots & \sigma_n(r_n) \end{bmatrix}) = m \cdot \det(\begin{bmatrix} \sigma_1(r_1) & \dots & \sigma_1(x) & \dots & \sigma_1(r_n) \\ \sigma_2(r_1) & \dots & \sigma_2(x) & \dots & \sigma_2(r_n) \\ \dots & \dots & \dots & \dots & \dots \\ \sigma_n(r_1) & \dots & \sigma_n(x) & \dots & \sigma_n(r_n) \end{bmatrix})$$

We can now describe an algorithm to finding the ring of integers of K.

1. Begin with an order $\mathcal{O}$ with integral basis $\{r_1 \dots r_n\}$. Calculate its discriminant $\Delta(\mathcal{O})$ and choose a prime factor m so that m^2 divides $\Delta(\mathcal{O})$.

2. An algebraic integer in K would have to take the form $x = \frac{\sum_{i=1}^{n} a_i r_i}{m}$ where each coefficient a_i lies in $\{0 \dots m-1\}$. If any such algebraic integers are found, extend the order by adding them in, and go back to step 1 with the new order. This will decrease the discriminant by a factor of m^2.

3. If no algebraic integers are found, go back to step 1 and choose a different prime factor m.

4. If no square prime factors remain, then the current order must be the maximal one.

There is a very nice way of determining whether an algebraic number is an algebraic integer using linear algebra. Combining this with our recipe for finding the ring of integers is particularly nice, and lends itself well to a computer program.

Remark 1.5.9. *Let K be a number field and fix a vector space basis $\{r_1 \dots r_n\}$ for K over $\mathbb{Q}$. Let $a = \sum_{i=1}^{n} a_i r_i \in K$. Then a can be viewed as the column vector*

$$\begin{bmatrix} a_1 \\ a_2 \\ \dots \\ a_n \end{bmatrix}$$

in $\mathbb{Q}^n$. Multiplication by a fixed element becomes linear map on $\mathbb{Q}^n$. Let M_a be the matrix corresponding to multiplication by a. By Cayley-Hamilton, M_a will satisfy its characteristic equation which has degree at most n. Therefore a will be a root of this equation, and so the minimal polynomial of a will divide the characteristic equation. Hence a is an algebraic integer if and only if the characteristic equation of M_a is monic with integer coefficients.

Example 1.5.10. *The above method allows us to find minimal polynomials of elements in abstract number fields. Let $K = \mathbb{Q}[x]/(x^3 + 7x - 5)$. We will compute the minimal polynomial of x^2, whose degree must be 3. We have $x^3 = 5 - 7x$ and $x^4 = 5x - 7x^2$. Therefore*

$$M_{x^2} = \begin{bmatrix} 0 & 5 & 0 \\ 0 & -7 & 5 \\ 1 & 0 & -7 \end{bmatrix}$$

The characteristic polynomial of this matrix is

$$x^3 + 14x^2 + 49x - 25$$

which has degree 3 so it must be the minimal polynomial of x^2.

See Section 2 for more examples of computations of the ring of integers.

The important thing to remember now that we move on to general orders is that the situation is not so drastically different. What we actually proved in Section 1.2 is that all *good* ideals are invertible. In Dedekind domains, which are integrally closed, all ideals are good. In general orders, it turns out that most ideals are good, and have an inverse. The conductor ideal will measure those ideals of the order which are not good. We follow [Cona] for the remainder of this subsection.

Definition 1.5.11. *Let K be a number field and $\mathcal{O}$ an order of K. The conductor ideal of $\mathcal{O}$ is defined as*

$$\mathfrak{C}_{\mathcal{O}} := \{x \in \mathcal{O}_K : x\mathcal{O}_K \subset \mathcal{O}\}$$

Note that the conductor ideal is an ideal in both $\mathcal{O}_K$ and $\mathcal{O}$. To see this, for any element $x \in \mathfrak{C}_{\mathcal{O}}$ we have $x \cdot 1 \in x\mathcal{O}_K \subset \mathcal{O}$ and so $x \in \mathcal{O}$. In fact, the conductor ideal is the largest ideal of $\mathcal{O}_K$ which is also contained in $\mathcal{O}$, as the following proposition shows.

Proposition 1.5.12. *Let K be a number field and $\mathcal{O}$ an order of K. Any ideal $\mathfrak{i}$ of $\mathcal{O}_K$ which is contained in $\mathcal{O}$ is also contained in $\mathfrak{C}_{\mathcal{O}}$.*

Proof. If $\mathfrak{i} \subset \mathcal{O}$ then for every element $x \in \mathfrak{i}$ we have $x\mathcal{O}_K \subset \mathfrak{i} \subset \mathcal{O}$. As a result $x \in \mathfrak{C}_{\mathcal{O}}$ for every element $x \in \mathfrak{i}$ and so $\mathfrak{i} \subset \mathfrak{C}_{\mathcal{O}}$. □

Definition 1.5.13. *Let K be a number field and $\mathcal{O}$ an order of K. An ideal $\mathfrak{i}$ of $\mathcal{O}$ is good if*

$$\{x \in K : x\mathfrak{i} \subset \mathfrak{i}\} = \mathcal{O}$$

and otherwise it is called bad.

Proposition 1.5.14. *Let K be a number field and $\mathcal{O}$ an order of K. If an ideal $\mathfrak{i}$ of $\mathcal{O}$ is coprime to the conductor ideal $\mathfrak{C}_{\mathcal{O}}$, then it is good.*

Proof. Assume $\mathfrak{i} \subset \mathcal{O}$ is coprime to the conductor ideal so that $\mathfrak{i} + \mathfrak{C}_\mathcal{O} = \mathcal{O}$. Then there are elements $b \in \mathfrak{i}, c \in \mathfrak{C}_\mathcal{O}$ so that $b + c = 1$. Let $x \in K$ so that $x\mathfrak{i} \subset \mathfrak{i}$. We have to show that $x \in \mathcal{O}$. We have $xb \in \mathfrak{i} \subset \mathcal{O}$ and $xc \in \mathfrak{C}_\mathcal{O} \subset \mathcal{O}$. Since $x = xb + xc$ we have that $x \in \mathcal{O}$. Therefore $\{x \in K : x\mathfrak{i} \subset \mathfrak{i}\} = \mathcal{O}$ and $\mathfrak{i}$ is a good ideal. □

We prove next that good prime ideals of any order are invertible, and their inverse is a good fractional ideal. Notice the parallels with Lemma 1.2.6 and Proposition 1.2.7.

Lemma 1.5.15. *Let K be a number field and $\mathcal{O}$ and order of K. Let $\mathfrak{p}$ be a prime ideal of $\mathcal{O}$. Then there is some element $q \in K\backslash\mathcal{O}$ so that $q\mathfrak{p} \subset \mathcal{O}$.*

Proof. Let $\mathfrak{p}$ be a prime ideal. Let $x \in \mathfrak{p}$ be a nonzero element element. (x) contains a minimal product of prime ideals $\mathfrak{p}_1 \cdots \mathfrak{p}_n$ by Lemma 1.2.4, such that $\mathfrak{p}_1 \cdots \mathfrak{p}_{n-1} \not\subset (x)$. Since $\mathfrak{p}$ is prime and $\mathfrak{p}_1 \cdots \mathfrak{p}_n \subset (x) \subset \mathfrak{p}$, we can say W.L.O.G. that $\mathfrak{p}_n \subset \mathfrak{p}$ by Lemma 1.2.5. $\mathcal{O}$ has Krull dimension 1 and so in fact we must have $\mathfrak{p}_n = \mathfrak{p}$.

Let $y \in \mathfrak{p}_1 \cdots \mathfrak{p}_{n-1}$ such that $y \notin (x)$, which exists because $\mathfrak{p}_1 \cdots \mathfrak{p}_{n-1} \not\subset (x)$. Then $y\mathfrak{p} \subset \mathfrak{p}_1 \cdots \mathfrak{p}_n \subset (x) \subset \mathfrak{p}$. As a result $\frac{y}{x}\mathfrak{p} \subset \mathcal{O}$, but $\frac{y}{x} \notin \mathcal{O}$ because $y \notin (x)$. Thus $q = \frac{y}{x}$ is an element of $K\backslash\mathcal{O}$ so that $q\mathfrak{p} \subset \mathcal{O}$. □

Proposition 1.5.16. *Let K be a number field and $\mathcal{O}$ and order of K. Every good prime ideal of $\mathcal{O}$ is invertible, and their inverse is a good fractional ideal.*

Proof. By Lemma 1.5.15 there is some element $q \in K\backslash\mathcal{O}$ so that $q\mathfrak{p} \subset \mathcal{O}$. Now $\mathfrak{p} \subset (1, q)\mathfrak{p} \subset \mathcal{O}$ as before and so either $(1, q)\mathfrak{p} = \mathfrak{p}$ or $(1, q)\mathfrak{p} = \mathcal{O}$. In the first case $q\mathfrak{p} \subset \mathfrak{p}$, but $\mathfrak{p}$ is a good ideal and so $q \in \mathcal{O}$ which is a contradiction. Therefore we must have $(1, q)\mathfrak{p} = \mathcal{O}$ and so the inverse is $\mathfrak{p}^{-1} = (1, q)$. It remains to show that $(1, q)$ is a good fractional ideal.

Suppose not. Then there is some $x \in K\backslash\mathcal{O}$ so that $x(1, q) \subset (1, q)$. However, this means $x(1, q)\mathfrak{p} \subset (1, q)\mathfrak{p}$ and in particular $x \subset \mathcal{O}$, a contradiction. Hence $\mathfrak{p}^{-1}$ is a good fractional ideal. □

What follows is a weak version of unique prime factorization of ideals in arbitrary orders. It states that ideals coprime to the conductor ideal can be factored uniquely into prime ideals coprime to the conductor ideal.

Proposition 1.5.17. *Let K be a number field and $\mathcal{O}$ an order of K. Let $\mathfrak{i}$ be an ideal of $\mathcal{O}$ not equal to $\mathcal{O}$ itself, and coprime to $\mathfrak{C}_\mathcal{O}$. Then $\mathfrak{i}$ can be factored into a finite number of prime ideals coprime to $\mathfrak{C}_\mathcal{O}$.*

Proof. The ideal (0) is already prime. By Zorn's lemma, every non-zero ideal of $\mathcal{O}$ not equal to $\mathcal{O}$ itself is contained in a maximal ideal. Let us assume that $\mathfrak{i} \subset \mathfrak{p}$ for some maximal ideal $\mathfrak{p}$. Then $\mathfrak{p}$ must be coprime to $\mathfrak{C}_\mathcal{O}$, since $\mathcal{O} = \mathfrak{C}_\mathcal{O} + \mathfrak{i} \subset \mathfrak{C}_\mathcal{O} + \mathfrak{p}$.

Then we have $\mathfrak{i} \subset \mathfrak{p}^{-1}\mathfrak{i} \subset \mathcal{O}$, since the inverse $\mathfrak{p}^{-1}$ exists by Proposition 1.5.16. We denote the ideal $\mathfrak{p}^{-1}\mathfrak{i}$ by $\mathfrak{i}_0$ and note that it must also be coprime to $\mathfrak{C}_\mathcal{O}$ by the same reason that $\mathfrak{p}$ is coprime

to $\mathfrak{C}_\mathcal{O}$. We repeat the process on $\mathfrak{i}_0$, getting a potentially infinite chain of ideals $\mathfrak{i} \subset \mathfrak{i}_0 \subset \mathfrak{i}_1 \subset \dots$ with $\mathfrak{i}_{j+1} = \mathfrak{p}_j^{-1}\mathfrak{i}_j$ for all j as long as some maximal ideal $\mathfrak{p}_j$ exists by Zorn's lemma. All the ideals mentioned will be coprime to $\mathfrak{C}_\mathcal{O}$. Since $\mathcal{O}$ is Noetherian, this chain must terminate and so eventually $\mathfrak{i}_n = \mathcal{O}$ for some n. Hence we can write $\mathfrak{i} = \mathfrak{p} \cdot \mathfrak{p}_0 \cdots \mathfrak{p}_{n-1}$. □

This proves the existence of a prime factorization of ideals coprime to the conductor ideal. To prove uniqueness, we develop a cancellation law by showing that ideals coprime to the conductor ideal are invertible. Note that UPF of good ideals in general is not always true.

Proposition 1.5.18. *Let K be a number field and $\mathcal{O}$ an order of K. Let $\mathfrak{i}$ be an ideal of $\mathcal{O}$ coprime to $\mathfrak{C}_\mathcal{O}$. Then $\mathfrak{i}$ is invertible.*

Proof. $\mathfrak{i}$ can be written as a finite product of prime ideals coprime to $\mathfrak{C}_\mathcal{O}$ by Proposition 1.5.17, each of which is invertible by Proposition 1.5.16. Therefore $\mathfrak{i}$ itself is invertible. □

Corollary 1.5.19. *Let K be a number field and $\mathcal{O}$ an order of K. Then ideals of $\mathcal{O}$ coprime to $\mathfrak{C}_\mathcal{O}$ will factor uniquely into a product of prime ideals coprime to $\mathfrak{C}_\mathcal{O}$.*

We will now work towards a generalization of the Dedekind–Kummer theorem.

Lemma 1.5.20. *Let K be a number field and $\mathcal{O}$ and order of K. Let $\mathfrak{i}$ be an ideal of $\mathcal{O}_K$ which is coprime to $\mathfrak{C}_\mathcal{O}$. Then $\mathcal{O}/\mathfrak{i} \cap \mathcal{O} \cong \mathcal{O}_K/\mathfrak{i}$ and $\mathfrak{i} \cap \mathcal{O}$ is a good ideal.*

Proof. To show this, note that $\mathfrak{i} + \mathfrak{C}_\mathcal{O} = \mathcal{O}_K$ by assumption. Hence $\mathfrak{i} + \mathcal{O} = \mathcal{O}_K$ because $\mathfrak{C}_\mathcal{O} \subset \mathcal{O} \subset \mathcal{O}_K$. Therefore the quotient map composed with the inclusion $\mathcal{O} \to \mathcal{O}_K \to \mathcal{O}_K/\mathfrak{i}$ is surjective. The kernel is clearly $\mathcal{O} \cap \mathfrak{i}$, and so $\mathcal{O}/\mathfrak{i} \cap \mathcal{O} \cong \mathcal{O}_K/\mathfrak{i}$.

From $\mathfrak{i} + \mathfrak{C}_\mathcal{O} = \mathcal{O}_K$ we get $\mathfrak{i} \cap \mathcal{O} + \mathfrak{C}_\mathcal{O} = \mathcal{O}$, since $\mathfrak{C}_\mathcal{O} \subset \mathcal{O}$. As a result $\mathfrak{i} \cap \mathcal{O}$ is a good ideal by Proposition 1.5.14, since it is coprime to the conductor ideal. □

Theorem 1.5.21. *Let $L : K$ be a number field extension and let $\theta \in \mathcal{O}_L$ so that $K(\theta) = L$. Then $\mathcal{O} = \mathcal{O}_K[\theta]$ is an order of L. Let $\mathfrak{C}$ be its conductor ideal. Let $\mathfrak{p}$ be a prime ideal of $\mathcal{O}_K$ so that $\mathfrak{p}\mathcal{O}_L$ is relatively prime to $\mathfrak{C}$. Let f be the minimal monic polynomial of θ over K.*

We claim that under these conditions, we can apply the Dedekind–Kummer theorem. In other words, if we can factorize f modulo $\mathfrak{p}$ into irreducibles as

$$\bar{f} \equiv \prod_{i=1}^{n} \bar{f_i}^{e_i} \pmod{\mathfrak{p}}$$

then $\mathfrak{p}$ factorizes into prime ideals of $\mathcal{O}_L$ as

$$\mathfrak{p}\mathcal{O}_L = \prod_{i=1}^{n} (\mathfrak{p}\mathcal{O}_L + (f_i(\theta)))^{e_i}$$

Proof. By Lemma 1.5.20 we get that $\mathcal{O}/\mathfrak{p}\cap\mathcal{O} \cong \mathcal{O}_L/\mathfrak{p}\mathcal{O}_L$ since $\mathfrak{p}\mathcal{O}_L$ is coprime to $\mathfrak{C}$. We also know that $\mathcal{O} \cong \mathcal{O}_K[x]/(f)$. Let $\mathfrak{p}_x$ denote the image of the ideal $\mathfrak{p}$ under this isomorphism. Then

$$\mathcal{O}/\mathfrak{p}\cap\mathcal{O} \cong \mathcal{O}_K[x]/(\mathfrak{p}_x \cap \mathcal{O}_K[x] + (f)) \cong (\mathcal{O}_K/(\mathfrak{p}\cap\mathcal{O}_K))[x]/(\bar{f})$$

We get a factorization of the zero element as

$$\bar{f}(\theta) = (0) \equiv \prod_{i=1}^{n} \bar{f}_i(\theta)^{e_i} \pmod{\mathfrak{p}}$$

Then by applying Chinese remainder theorem gives us

$$\mathcal{O}/\mathfrak{p}\cap\mathcal{O} \cong \prod_{i=1}^{n}(\mathcal{O}_K/(\mathfrak{p}\cap\mathcal{O}_K))[x]/(f_i)^{e_i} \cong \prod_{i=1}^{n}(\mathcal{O}/(\mathfrak{p}\cap\mathcal{O}))/(f_i(\theta))^{e_i} \cong \prod_{i=1}^{n}\mathcal{O}/(\mathfrak{p}\cap\mathcal{O}+(f_i(\theta))^{e_i})$$

We will repeat the same argument as with Theorem 1.3.8. Firstly note that $(\mathcal{O}_K/(\mathfrak{p}\cap\mathcal{O}_K))[x]$ is a PID, because it is a polynomial ring over a field with transcendence degree 1. Therefore the ideals $(f_i(\theta))$ are prime in $\mathcal{O}/\mathfrak{p}\cap\mathcal{O}$, because they are generated by irreducible elements. Hence the quotients $\mathcal{O}/(\mathfrak{p}\cap\mathcal{O}+(f_i(\theta)))$ are fields, and so $(\mathfrak{p}\cap\mathcal{O}+(f_i(\theta)))$ are prime ideals of $\mathcal{O}$.

$\mathfrak{p}\cap\mathcal{O}$ is coprime to $\mathfrak{C}$ by Lemma 1.5.20. Furthermore, $(\mathfrak{p}\cap\mathcal{O}+(f_i(\theta)))$ is coprime to $\mathfrak{C}$, since it contains the ideal $\mathfrak{p}\cap\mathcal{O}$. $(\mathfrak{p}\cap\mathcal{O}+(f_i(\theta))^{e_i})$ is also coprime to $\mathfrak{C}$ because it is a factor of $\mathfrak{p}\cap\mathcal{O}$. Now we have the containment of ideals $(\mathfrak{p}\cap\mathcal{O}+(f_i(\theta)))^{e_i} \subset (\mathfrak{p}\cap\mathcal{O}+(f_i(\theta))^{e_i}) \subset (\mathfrak{p}\cap\mathcal{O}+(f_i(\theta)))$ as in Theorem 1.3.8, each of which is coprime to $\mathfrak{C}$. By unique prime factorization of ideals coprime to the conductor ideal, we have $(\mathfrak{p}\cap\mathcal{O}+(f_i(\theta))^{e_i}) = (\mathfrak{p}\cap\mathcal{O}+(f_i(\theta)))^r$ for some integer r. However, there are chains of proper containments

$$(\mathfrak{p}\cap\mathcal{O}+(f_i(\theta))) \subsetneq (\mathfrak{p}\cap\mathcal{O}+(f_i(\theta))^2) \subsetneq \cdots \subsetneq (\mathfrak{p}\cap\mathcal{O}+(f_i(\theta))^{e_i})$$

$$(\mathfrak{p}\cap\mathcal{O}+(f_i(\theta))) \subsetneq (\mathfrak{p}\cap\mathcal{O}+(f_i(\theta)))^2 \subsetneq \cdots \subsetneq (\mathfrak{p}\cap\mathcal{O}+(f_i(\theta)))^{e_i}$$

All of these ideals are coprime to $\mathfrak{C}$, so one can argue $(\mathfrak{p}\cap\mathcal{O}+(f_i(\theta))^{e_i}) = (\mathfrak{p}\cap\mathcal{O}+(f_i(\theta)))^{e_i}$ by UPF of ideals coprime to the conductor ideal. Therefore, by Lemma 1.5.20 we get

$$\mathcal{O}_L/\mathfrak{p}\mathcal{O}_L \cong \mathcal{O}/\mathfrak{p}\cap\mathcal{O} \cong \prod_{i=1}^{n}\mathcal{O}/(\mathfrak{p}\cap\mathcal{O}+(f_i(\theta))^{e_i}) \cong \prod_{i=1}^{n}(\mathcal{O}/\mathfrak{p}\cap\mathcal{O}+(f_i(\theta)))^{e_i} \cong \prod_{i=1}^{n}\mathcal{O}_L/(\mathfrak{p}\mathcal{O}_L+(f_i(\theta)))^{e_i}$$

It follows that $\mathfrak{p}\mathcal{O}_L = \prod_{i=1}^{n}(\mathfrak{p}\mathcal{O}_L + (f_i(\theta)))^{e_i}$. Each of the ideals on the RHS are prime because their quotient rings are fields. This completes the proof. □

Now we get some information on the conductor ideal, which will help us produce a slick corollary of the above theorem. Then we will see an example of its use.

We work in the absolute case: a number field extension $K : \mathbb{Q}$. Let $\theta \in K$ so that $K = \mathbb{Q}(\theta)$ and let f be its minimal polynomial. Then the absolute discriminant of the order $\mathbb{Z}[\theta]$ is just $|disc(f)|$, the discriminant of the polynomial. The algebraic integers in $\mathcal{O}_K\backslash\mathbb{Z}[\theta]$ will all have

reduced denominator dividing $|disc(f)|$, due to our algorithm for finding the ring of integers. Hence $(disc(f)) \subset \mathfrak{C}_{\mathbb{Z}[\theta]}$. To see this, let g be the least common multiple of the denominators. Then $(disc(f)) \subset (g)$ and $(g)\mathcal{O}_K = \mathbb{Z}[\theta]$ since it cancels out the denominators in a minimal way. Therefore $(g) \subset \mathfrak{C}_{\mathbb{Z}[\theta]}$ by Proposition 1.5.12.

Corollary 1.5.22. *Let $K = \mathbb{Q}[x]/(f)$ be a number field for f a monic irreducible polynomial with integer coefficients. Let p be an integer prime whose square does not divide the discriminant of f. Then (p) is coprime to the conductor ideal of $\mathbb{Z}[x]/(f)$ and so it can be factorized using the generalized Dedekind–Kummer theorem.*

Example 1.5.23. *The polynomial $f = x^5 + 3x + 3$ is irreducible by Eisenstein's criterion. Let $K = \mathbb{Q}[x]/(f)$. Then f has discriminant 315333. 5 does not divide this discriminant so we can factor $(5)\mathcal{O}_K$ using the above method. We have*

$$x^5 + 3x + 3 \equiv (x-3)(x^4 + 3x^3 + 4x^2 + 2x + 4) \equiv (x-3)((x+2)^4 + 3) \pmod 5$$

By doing the substitution $y = x + 2$, we see that the polynomial $y^4 + 3$ has no roots modulo 5, so it is either irreducible or factorizes into two quadratic factors. In the latter case we may write

$$y^4 + 3 = (y^2 + ay + b)(y^2 + my + n) = (y^4 + (a+m)y^3 + (b+n+am)y^2 + (an+bm)y + bn)$$

A quick check of the cases shows that this cannot occur, and so $x^4 + 3x^3 + 4x^2 + 2x + 4$ is irreducible modulo 5 and we may write

$$(5) = (5, x-3)(5, x^4 + 3x^3 + 4x^2 + 2x + 4)$$

229 divides the discriminant with multiplicity 1, and so $(229)\mathcal{O}_K$ can also be factored using the above method. After running a computer program we see that the only roots of $x^5 + 3x + 3$ modulo 229 are 180 and 56. By looking at the roots of the derivative of $x^5 + 3x + 3$, we see that 56 is a root with multiplicity 2 and 180 is a root with multiplicity 1. Therefore we can factorize

$$x^5 + 3x + 3 \equiv (x - 180)(x - 56)^2(x^2 + 63x + 138) \pmod{229}$$

Now $x^2 + 63x + 138$ is irreducible since there are no other roots. As a result we finally have

$$(229) = (229, x - 180)(229, x - 56)^2(229, x^2 + 63x + 138)$$

1.6 More on prime decomposition

Let $L : K$ be a Galois extension with Galois group G throughout this subsection. We will study more carefully the prime decomposition of Galois extensions. We follow [Stea] and [Steb].

Definition 1.6.1. *Let $\mathfrak{P}$ be a prime ideal in $\mathcal{O}_L$. $D_{L:K}(\mathfrak{P})$ is defined as the decomposition group of $\mathfrak{P}$. This is the subgroup of G that fixes $\mathfrak{P}$. That is,*

$$D_{L:K}(\mathfrak{P}) = \{\sigma \in G : \sigma(\mathfrak{P}) = \mathfrak{P}\}$$

Definition 1.6.2. *Let $\mathfrak{P}$ be a prime ideal in $\mathcal{O}_L$. $I_{L:K}(\mathfrak{P})$ is defined as the Inertia group of $\mathfrak{P}$. This is the subgroup of G that fixes $\mathcal{O}_L/\mathfrak{P}$. That is,*

$$I_{L:K}(\mathfrak{P}) = \{\sigma \in G : \sigma(a) \equiv a \pmod{\mathfrak{P}} \, \forall a \in \mathcal{O}_L/\mathfrak{P}\}$$

Note that $I_{L:K}(\mathfrak{P})$ necessarily fixes $\mathfrak{P}$, and so it a subgroup of $D_{L:K}(\mathfrak{P})$.

Lemma 1.6.3. *Let $\mathfrak{p}$ be a prime ideal of $\mathcal{O}_K$. Let $\mathfrak{P}$ be the prime above $\mathfrak{p}$ in the extension $L : K$. Then $|D_{L:K}(\mathfrak{P})| = f_{L:K}(\mathfrak{P})e_{L:K}(\mathfrak{P})$*

Proof. From Section 1.4, we have the formula $[L : K] = nf_{L:K}(\mathfrak{P})e_{L:K}(\mathfrak{P})$ where n is the number of distinct primes above $\mathfrak{p}$. Let's say $\mathfrak{P}_1 \dots \mathfrak{P}_n$ are those primes with $\mathfrak{P} = \mathfrak{P}_1$. Then G acts transitively on these primes, and $D_{L:K}(\mathfrak{P})$ is the stabilizer of the element $\mathfrak{P}$, whilst $\{\mathfrak{P}_1 \dots \mathfrak{P}_n\}$ is the orbit. By the orbit-stabilizer theorem, $|D_{L:K}(\mathfrak{P})| = \frac{|G|}{n} = f_{L:K}(\mathfrak{P})e_{L:K}(\mathfrak{P})$. □

Lemma 1.6.4. *Let $\mathfrak{p}$ be a prime ideal of $\mathcal{O}_K$. Then the decomposition groups of the primes above $\mathfrak{p}$ are conjugate in G.*

Proof. Let $\mathfrak{P}_1, \mathfrak{P}_2$ be two primes above $\mathfrak{p}$. From Section 1.4, we know that there is an element $\sigma \in G$ so that $\sigma(\mathfrak{P}_1) = \mathfrak{P}_2$. Then $\sigma^{-1}D_{L:K}(\mathfrak{P}_2)\sigma \subset D_{L:K}(\mathfrak{P}_1)$. However, by Lemma 1.6.3 we have $|D_{L:K}(\mathfrak{P}_2)| = |D_{L:K}(\mathfrak{P}_1)| = f_{L:K}(\mathfrak{P})e_{L:K}(\mathfrak{P})$ and so in fact $\sigma^{-1}D_{L:K}(\mathfrak{P}_2)\sigma = D_{L:K}(\mathfrak{P}_1)$. This completes the proof. □

Keep in mind that decomposition groups are not necessarily normal in G. However, we can still construct their fixed fields. Let $\mathfrak{p}$ be a prime ideal in $\mathcal{O}_K$. The fixed fields of the decomposition groups of the primes above $\mathfrak{p}$ will be isomorphic by the above lemma, which states that the corresponding groups are conjugate in G.

Lemma 1.6.5. *Let $\mathfrak{P}$ be a prime ideal in $\mathcal{O}_L$ and let $\mathfrak{p}$ be the prime below $\mathfrak{P}$ in K. Then $\mathfrak{P} \cap L^{D_{L:K}(\mathfrak{P})}$ has inertia and ramification degree 1 over the extension $L^{D_{L:K}(\mathfrak{P})} : K$ and it does not split in the extension $L : L^{D_{L:K}(\mathfrak{P})}$.*

Proof. By Lemma 1.6.3 and field theory we have

$$[L^{D_{L:K}(\mathfrak{P})} : K] = \frac{[L : K]}{|D_{L:K}(\mathfrak{P})|} = \frac{[L : K]}{e_{L:K}(\mathfrak{P})f_{L:K}(\mathfrak{P})} = n$$

Let $\mathfrak{q} = \mathfrak{P} \cap L^{D_{L:K}(\mathfrak{P})}$. By this definition, $\mathfrak{P}$ is a prime above $\mathfrak{q}$. Note that $L : L^{D_{L:K}(\mathfrak{P})}$ is a Galois extension with Galois group $D_{L:K}(\mathfrak{P})$ and so we can use the more specific tools we have developed. We will show that $\mathfrak{q}$ does not split in the extension $L : L^{D_{L:K}(\mathfrak{P})}$. $D_{L:K}(\mathfrak{P}) = \mathrm{Gal}(L : L^{D_{L:K}(\mathfrak{P})})$ fixes $\mathfrak{P}$, a prime above $\mathfrak{q}$ in the extension $L : L^{D_{L:K}(\mathfrak{P})}$. Since the Galois group acts transitively, $\mathfrak{P}$ must be the only prime above $\mathfrak{q}$, and so $\mathfrak{q}$ is inert (does not split) in $L : L^{D_{L:K}(\mathfrak{P})}$.

Recall that the inertia and ramification degrees are multiplicative. In particular we have

$$|D_{L:K}(\mathfrak{P})| = e_{L:K}(\mathfrak{P})f_{L:K}(\mathfrak{P}) = e_{L:L^{D_{L:K}(\mathfrak{P})}}(\mathfrak{P})e_{L^{D_{L:K}(\mathfrak{q})}:K}(\mathfrak{P})f_{L:L^{D_{L:K}(\mathfrak{P})}}(\mathfrak{P})f_{L^{D_{L:K}(\mathfrak{P})}:K}(\mathfrak{q})$$

On the other hand, since $\mathfrak{q}$ is inert in $L : L^{D_{L:K}(\mathfrak{P})}$, we get

$$|D_{L:K}(\mathfrak{P})| = e_{L:L^{D_{L:K}(\mathfrak{P})}}(\mathfrak{P}) f_{L:L^{D_{L:K}(\mathfrak{P})}}(\mathfrak{P})$$

so we can conclude that

$$e_{L^{D_{L:K}(\mathfrak{P})}:K}(\mathfrak{q}) f_{L^{D_{L:K}(\mathfrak{P})}:K}(\mathfrak{q}) = 1$$

It follows that $\mathfrak{q}$ has inertia and ramification degrees 1 over the extension $L^{D_{L:K}(\mathfrak{P})} : K$. □

Proposition 1.6.6. *Let $\mathfrak{P}$ be a prime ideal in $\mathcal{O}_L$ and let $\mathfrak{p}$ be the prime below $\mathfrak{P}$ in K. Then $D_{L:K}(\mathfrak{P})/I_{L:K}(\mathfrak{P}) \cong \mathrm{Gal}(\mathcal{O}_L/\mathfrak{P} : \mathcal{O}_K/\mathfrak{p})$, the group of automorphisms of $\mathcal{O}_L/\mathfrak{P}$ that fixes the natural embedding of $\mathcal{O}_K/\mathfrak{p}$.*

Proof. $\mathcal{O}_L/\mathfrak{P} : \mathcal{O}_K/\mathfrak{p}$ is an extension of finite fields, and so its Galois group is cyclic of order $f_{L:K}(\mathfrak{P})$. Its generator is induced by the Frobenius automorphism $\mathrm{Frob}_{\mathcal{O}_K/\mathfrak{p}}$, that sends x to $x^{N(\mathfrak{p})}$ in $\mathcal{O}_L/\mathfrak{P}$. Each element of $D_{L:K}(\mathfrak{P})$ defines an automorphism of $\mathcal{O}_L/\mathfrak{P}$ that fixes $\mathcal{O}_K/\mathfrak{p}$. This gives a group homomorphism from $D_{L:K}(\mathfrak{P})$ to $\mathrm{Gal}(\mathcal{O}_L/\mathfrak{P} : \mathcal{O}_K/\mathfrak{p})$ whose kernel is $I_{L:K}(\mathfrak{P})$ by definition. It remains to show that this homomorphism is surjective.

Since we have a finite extension of fields, we can use the primitive element theorem to get some element $a \in \mathcal{O}_L/\mathfrak{P}$ so that $\mathcal{O}_L/\mathfrak{P} = (\mathcal{O}_K/\mathfrak{p})(a)$. Let a' be a representative of a in $\mathcal{O}_L$ and let f be the polynomial

$$f = \prod_{\sigma \in D_{L:K}(\mathfrak{P})} (x - \sigma(a'))$$

over $L^{D_{L:K}(\mathfrak{P})}$. Let $\mathfrak{q}$ be the prime below $\mathfrak{P}$ in $L^{D_{L:K}(\mathfrak{P})}$. Then $\mathcal{O}_{L^{D_{L:K}(\mathfrak{P})}}/\mathfrak{q} \cong \mathcal{O}_K/\mathfrak{p}$ since the corresponding inertia degree is 1, by Lemma 1.6.5. Hence the reduction of f modulo $\mathfrak{q}$ can be made to have coefficients in $\mathcal{O}_K/\mathfrak{p}$. Let this polynomial be $\bar{f}$. It splits completely in the extension $\mathcal{O}_L/\mathfrak{P} : \mathcal{O}_K/\mathfrak{p}$ because it has roots $\overline{\sigma(a')} = \sigma(a)$ for every $\sigma \in D_{L:K}(\mathfrak{P})$.

Now $\mathrm{Frob}_{\mathcal{O}_K/\mathfrak{p}}(a)$ is also a root of $\bar{f}$ since the Frobenius automorphism fixes the coefficients of $\bar{f}$, which are in $\mathcal{O}_K/\mathfrak{p}$. Hence $\mathrm{Frob}_{\mathcal{O}_K/\mathfrak{p}}(a) = \sigma(a)$ for some $\sigma \in D_{L:K}(\mathfrak{P})$. Since $\mathcal{O}_L/\mathfrak{P} = (\mathcal{O}_K/\mathfrak{p})(a)$, and both σ and the Frobenius automorphism fix $\mathcal{O}_K/\mathfrak{p}$, we must actually have $\mathrm{Frob}_{\mathcal{O}_K/\mathfrak{p}} = \sigma$. As a result σ is sent to $\mathrm{Frob}_{\mathcal{O}_K/\mathfrak{p}}$ which generates $\mathrm{Gal}(\mathcal{O}_L/\mathfrak{P} : \mathcal{O}_K/\mathfrak{p})$, and so the homomorphism $D_{L:K}(\mathfrak{P}) \to \mathrm{Gal}(\mathcal{O}_L/\mathfrak{P} : \mathcal{O}_K/\mathfrak{p})$ is surjective. This completes the proof. □

With this last proposition the entire picture can be painted. We shall see that for any Galois extension $L : K$ and prime ideal $\mathfrak{P}$ of $\mathcal{O}_L$, there are intermediate fields $L : A : B : K$ so that

- The prime below $\mathfrak{P}$ in B has inertia and ramification degrees 1 in the extension $B : K$.
- The prime below $\mathfrak{P}$ in B remains inert in $A : B$.
- The prime below $\mathfrak{P}$ in A totally ramifies in $L : A$.

Proposition 1.6.7. *Let $\mathfrak{P}$ be a prime ideal in $\mathcal{O}_L$ and let $\mathfrak{p}$ be the prime below $\mathfrak{P}$ in K. Then $|I_{L:K}(\mathfrak{P})| = e_{L:K}(\mathfrak{P})$ and $\mathfrak{P} \cap L^{I_{L:K}(\mathfrak{P})}$ ramifies completely with ramification degree $e_{L:K}(\mathfrak{P})$ in the extension $L : L^{I_{L:K}(\mathfrak{P})}$.*

Proof. From Galois theory we have the equality

$$\frac{|D_{L:K}(\mathfrak{p})|}{|I_{L:K}(\mathfrak{p})|} = |\mathrm{Gal}(\mathcal{O}_L/\mathfrak{P} : \mathcal{O}_K/\mathfrak{p})| = f_{L:K}(\mathfrak{P})$$

As a result $|I_{L:K}(\mathfrak{P})| = e_{L:K}(\mathfrak{P})$ by Lemma 1.6.3. By field theory the extension $L : L^{I_{L:K}(\mathfrak{P})}$ has degree $e_{L:K}(\mathfrak{P})$, and so the extension $L^{I_{L:K}(\mathfrak{P})} : L^{D_{L:K}(\mathfrak{P})}$ has degree $f_{L:K}(\mathfrak{P})$.

Let $\mathfrak{q} = \mathfrak{P} \cap L^{I_{L:K}(\mathfrak{P})}$. The extension $L : L^{I_{L:K}(\mathfrak{P})}$ is Galois with Galois group $I_{L:K}(\mathfrak{P})$. The group $I_{L:K}(\mathfrak{P})$ fixes the quotient ring of $\mathfrak{P}$, which is a prime above $\mathfrak{q}$. It follows that $[\mathcal{O}_L/\mathfrak{P} : \mathcal{O}_{L^{I_{L:K}(\mathfrak{P})}}/\mathfrak{q}] = 1$ because the entire Galois group fixes the quotient ring and so $|\mathrm{Gal}(\mathcal{O}_L/\mathfrak{P} : \mathcal{O}_{L^{I_{L:K}(\mathfrak{P})}}/\mathfrak{q})| = f_{L:L^{I_{L:K}(\mathfrak{P})}}(\mathfrak{P}) = 1$. Recall also that $f_{L^{D_{L:K}(\mathfrak{P})}:K}(\mathfrak{q} \cap L^{D_{L:K}(\mathfrak{P})}) = 1$ from Lemma 1.6.5. Since the inertia degree is multiplicative, we get

$$f_{L:K}(\mathfrak{P}) = f_{L:L^{I_{L:K}(\mathfrak{P})}}(\mathfrak{P}) \cdot f_{L^{I_{L:K}(\mathfrak{P})}:L^{D_{L:K}(\mathfrak{P})}}(\mathfrak{q}) \cdot f_{L^{D_{L:K}(\mathfrak{P})}:K}(\mathfrak{q} \cap L^{D_{L:K}(\mathfrak{P})}) = f_{L^{I_{L:K}(\mathfrak{P})}:L^{D_{L:K}(\mathfrak{P})}}(\mathfrak{q})$$

so all the inertia must occur in the extension $L^{I_{L:K}(\mathfrak{p})} : L^{D_{L:K}(\mathfrak{p})}$. By applying Proposition 1.4.5 we get

$$\begin{aligned} f_{L^{I_{L:K}(\mathfrak{P})}:L^{D_{L:K}(\mathfrak{P})}}(\mathfrak{q}) &= [L^{I_{L:K}(\mathfrak{p})} : L^{D_{L:K}(\mathfrak{p})}] = \\ &= \sum_{\mathfrak{r} | (\mathfrak{P} \cap L^{D_{L:K}(\mathfrak{P})}) \mathcal{O}_{L^{I_{L:K}(\mathfrak{P})}}} f_{L^{I_{L:K}(\mathfrak{P})}:L^{D_{L:K}(\mathfrak{P})}}(\mathfrak{r}) e_{L^{I_{L:K}(\mathfrak{P})}:L^{D_{L:K}(\mathfrak{P})}}(\mathfrak{r}) \end{aligned}$$

Since the sum on the RHS consists of integers, $\mathfrak{q}$ must be the only prime above $\mathfrak{P} \cap L^{D_{L:K}(\mathfrak{P})}$ in $L^{I_{L:K}(\mathfrak{P})}$, and it is unramified. In other words, $e_{L^{I_{L:K}(\mathfrak{P})}:L^{D_{L:K}(\mathfrak{P})}}(\mathfrak{q}) = 1$. We also have $e_{L^{D_{L:K}(\mathfrak{P})}:K}(\mathfrak{q} \cap L^{D_{L:K}(\mathfrak{P})}) = 1$ by Lemma 1.6.5. The ramification degree is multiplicative, so we get

$$e_{L:K}(\mathfrak{P}) = e_{L:L^{I_{L:K}(\mathfrak{P})}}(\mathfrak{P}) \cdot e_{L^{I_{L:K}(\mathfrak{P})}:L^{D_{L:K}(\mathfrak{P})}}(\mathfrak{q}) \cdot e_{L^{D_{L:K}(\mathfrak{P})}:K}(\mathfrak{q} \cap L^{D_{L:K}(\mathfrak{P})}) = e_{L:L^{I_{L:K}(\mathfrak{P})}}(\mathfrak{P})$$

As a result all ramification must occur in the extension $L : L^{I_{L:K}(\mathfrak{p})}$. □

Here is a table that illustrates the results thus far:

We will now define the Artin symbol for unramified ideals in Galois extensions, which relates them to elements of the Galois group. For a Galois extension $L : K$, the corresponding Artin symbol is a group homomorphism from the group of fractional ideals of L coprime to the ramified primes, to the Galois group $\mathrm{Gal}(L : K)$. This turns out to be very special in abelian extensions.

Number field extension	Degree	Prime below $\mathfrak{P}$	Ramification degree	Inertia degree
$L : L^{I_{L:K}(\mathfrak{P})}$	$e_{L:K}(\mathfrak{P})$	$\mathfrak{P} \cap L^{I_{L:K}(\mathfrak{P})}$	$e_{L:K}(\mathfrak{P})$	1
$L^{I_{L:K}(\mathfrak{P})} : L^{D_{L:K}(\mathfrak{P})}$	$f_{L:K}(\mathfrak{P})$	$\mathfrak{P} \cap L^{D_{L:K}(\mathfrak{P})}$	1	$f_{L:K}(\mathfrak{P})$
$L^{D_{L:K}(\mathfrak{P})} : K$	n	$\mathfrak{p}$	1	1

Definition 1.6.8. *Let $L : K$ be a Galois extension and $\mathfrak{P}$ an unramified prime ideal of $\mathcal{O}_L$. Let $\mathfrak{p}$ be the prime below $\mathfrak{P}$ in K. Then the inertia group $I_{L:K}(\mathfrak{P})$ is trivial because $\mathfrak{P}$ is unramified. Therefore $D_{L:K}(\mathfrak{P}) \cong \mathrm{Gal}(\mathcal{O}_L/\mathfrak{P} : \mathcal{O}_K/\mathfrak{p})$. As a result there is a unique $\sigma \in \mathrm{Gal}(L : K)$ that corresponds to the Frobenius automorphism of $\mathcal{O}_K/\mathfrak{p}$ in $\mathcal{O}_L/\mathfrak{P}$.*

We define the Artin symbol of the prime $\mathfrak{P}$ in the extension $L : K$ as

$$(\frac{L : K}{\mathfrak{P}}) = \sigma$$

Proposition 1.6.9. *Let $L : K$ be an abelian Galois extension and $\mathfrak{p}$ a prime ideal of $\mathcal{O}_K$ that does not ramify. Then the Artin symbol for any prime above $\mathfrak{p}$ in L is the same.*

Proof. Let $\mathfrak{P}, \mathfrak{P}'$ be two primes above $\mathfrak{p}$ in L. Let $\sigma \in \mathrm{Gal}(L : K)$ so that $\sigma(\mathfrak{P}) = \mathfrak{P}'$. Since we are in a Galois extension, we have $\mathcal{O}_L/\mathfrak{P} \cong \mathcal{O}_L/\mathfrak{P}'$. This isomorphism is given by σ. As a result, the corresponding Galois groups are isomorphic, and the isomorphism is given by conjugation with σ. Under this isomorphism, the Frobenius automorphisms are mapped to each other, and so

$$(\frac{L : K}{\mathfrak{P}'}) = \sigma(\frac{L : K}{\mathfrak{P}})\sigma^{-1}$$

Since $Gal(L : K)$ is abelian, we get $(\frac{L:K}{\mathfrak{P}'}) = (\frac{L:K}{\mathfrak{P}})$. □

As a result of the above proposition, the Artin symbol for unramified primes is solely determined by the prime below, in abelian Galois extensions. Let $L : K$ be an abelian Galois extension which is unramified, meaning that no prime ideal in $\mathcal{O}_L$ is ramified. Let $\mathfrak{P}$ be a prime ideal in $\mathcal{O}_L$ and let $\mathfrak{p}$ be the prime below $\mathfrak{P}$ in K. Then we can define without loss of generality

$$(\frac{L : K}{\mathfrak{p}}) := (\frac{L : K}{\mathfrak{P}})$$

Definition 1.6.10. *Let $L : K$ be an unramified abelian Galois extension. There is a map called the Artin map, which is the group homomorphism*

$$\phi_{L:K} : \mathcal{I}_K \to \mathrm{Gal}(L : K)$$

$$\prod_i \mathfrak{p}_i^{e_i} \mapsto \prod_i (\frac{L : K}{\mathfrak{p}_i})^{e_i}$$

1.7 More on discriminants

We've already described the absolute discriminant of a number field along with some of its basic properties. However there are much deeper facts which makes the discriminant a powerful computational tool. We also introduce the relative discriminant here.

Theorem 1.7.1 (Brill)**.** *Let K be number field with r_2 pairs of complex conjugate places. Then the sign of $\Delta(K)$ is $(-1)^{r_2}$*

Proof. Without taking the square in the definition of discriminant, $\sqrt{\Delta(K)} = b$ or bi for some integer b. Applying complex conjugation on the matrix will swap r_2 rows, and thus change the sign of $\sqrt{\Delta(K)}$ by $(-1)^{r_2}$. If the sign is changed, which occurs when r_2 is odd, then we must have the case $\sqrt{\Delta(K)} = bi$ since conjugation alters it. In this case $sgn(\Delta(K)) = -1 = (-1)^{r_2}$. If r_2 is even, conjugation does not change $\sqrt{\Delta(K)}$ and so we have the case $\sqrt{\Delta(K)} = b$ in which case $sgn(\Delta(K)) = 1 = (-1)^{r_2}$. This completes the proof. □

The following theorem was completed with the help of hints at the end of [BA].

Theorem 1.7.2 (Stickelberger)**.** *Let $\Delta(K)$ be the absolute discriminant of a number field K. Then $\Delta(K) \equiv 0$ or $1 \pmod 4$.*

Proof. Write the determinant of the discriminant matrix as $P + N$ where P is the sum of the terms given by even permutations and N is the sum of the terms given by an odd permutations. Let σ be an embedding of K. Applying σ to the discriminant matrix may swap the sign of the determinant, as discussed in the previous theorem. In particular, since σ permutes the rows, every even permutation becomes odd and vice versa. Hence $\sigma(P) = -N$ and $\sigma(N) = -P$.

As a result PN and $P - N$ will be fixed by every embedding. Thus they are rational integers by the same logic that $\Delta(K)$ is a rational integer. Then we have

$$\Delta(K) = (P+N)^2 = (P-N)^2 + 4PN \implies (P+N)^2 \equiv (P-N)^2 \pmod 4$$

so $\Delta(K)$ is congruent to a square integer $(P - N)^2$ modulo 4. The only integer squares modulo 4 are 0 and 1 and so we are done. □

We will now give an alternate description of the absolute discriminant. The remainder of this subsection is based on [Oss] and related lectures.

Proposition 1.7.3. *Let K be a number field and choose a $\mathbb{Z}$-basis $\{a_1 \dots a_n\}$ for $\mathcal{O}_K$. Then*

$$\Delta(K) = \det \begin{bmatrix} Tr(a_1a_1) & Tr(a_2a_1) & \dots & Tr(a_na_1) \\ Tr(a_1a_2) & Tr(a_2a_2) & \dots & Tr(a_na_2) \\ \dots & \dots & \dots & \dots \\ Tr(a_1a_n) & Tr(a_2a_n) & \dots & Tr(a_na_n) \end{bmatrix}$$

Proof. Let M be the original matrix used in the computation of the discriminant. That is,

$$M = \begin{bmatrix} \sigma_1(a_1) & \sigma_1(a_2) & \dots & \sigma_1(a_n) \\ \sigma_2(a_1) & \sigma_2(a_2) & \dots & \sigma_2(a_n) \\ \dots & \dots & \dots & \dots \\ \sigma_n(a_1) & \sigma_n(a_2) & \dots & \sigma_n(a_n) \end{bmatrix}$$

Let $R = M^T M$. Then we have

$$\Delta(K) = \det(M^2) = \det(M^T M) = \det(R)$$

The (i, j) entry in matrix R is

$$\sum_{k=1}^{n} \sigma_k(a_i)\sigma_k(a_j) = \sum_{k=1}^{n} \sigma_k(a_i a_j) = Tr(a_i a_j) \quad \square$$

This description is reminiscent of the trace trick we used to embed orders into $\mathbb{Z}^n$. Under this description we can view the discriminant as some sort of measure of size of an order, since it does actually correspond to the volume of the fundamental parallelepiped of the embedding of the order into $\mathbb{Z}^n$. Now we will discuss the notion of relative discriminant.

Definition 1.7.4. *Let $L : K$ be a number field extension. Let $n = [L : K]$ so that there are embeddings $\sigma_1 \dots \sigma_n$ of L into $\mathbb{C}$ that fix K. Let $\{a_1 \dots a_n\}$ be a vector space basis for L over K which is integral (inside $\mathcal{O}_L$). For this basis we can define the discriminant $\Delta(a_1 \dots a_n)$ as*

$$\Delta(a_1 \dots a_n) = \det \begin{bmatrix} \sigma_1(a_1) & \sigma_1(a_2) & \dots & \sigma_1(a_n) \\ \sigma_2(a_1) & \sigma_2(a_2) & \dots & \sigma_2(a_n) \\ \dots & \dots & \dots & \dots \\ \sigma_n(a_1) & \sigma_n(a_2) & \dots & \sigma_n(a_n) \end{bmatrix}^2$$

We define the relative discriminant of the extension $L : K$ as an ideal $\Delta(L : K)$ generated by the elements $\Delta(a_1 \dots a_n)$ as $\{a_1 \dots a_n\}$ runs through all integral vector space bases of L over K.

Note that for a number field extension $L : K$, $\mathcal{O}_L$ is not always a free $\mathcal{O}_K$-module. As a result the relative discriminant may not be a principal ideal. When $K = \mathbb{Q}$, the relative discriminant is principal and is generated by the absolute discriminant, so this definition extends that of the absolute discriminant. To see this, the discriminant of each vector space basis for K over $\mathbb{Q}$ which is integral is the same as the absolute discriminant of the corresponding order. We know that the discriminant of the maximal order divides the discriminant of the other orders and so the absolute discriminant generates the relative discriminant.

It is also easy to see that the ideal $\Delta(L : K)$ lies in $\mathcal{O}_K$. To see this, each individual element is fixed by the embeddings σ_i and lies in $\mathcal{O}_L$, and $\mathcal{O}_L \cap K = \mathcal{O}_K$. The definition of relative discriminant can also be extended to arbitrary orders in L, still giving an ideal in $\mathcal{O}_K$.

We can also define discriminants of extensions of quotient rings in number field extensions. Let $L : K$ be a number field extension and $\mathfrak{p}$ a prime ideal in $\mathcal{O}_K$. Then $\mathcal{O}_L/\mathfrak{p}\mathcal{O}_L$ is a vector space over $\mathcal{O}_K/\mathfrak{p}$ of degree equal to $n = [L : K]$, so we can choose a basis $\{a_1 \dots a_n\}$ in $\mathcal{O}_L/\mathfrak{p}\mathcal{O}_L$. The discriminant of this basis is defined as usual, and the discriminant of the quotient ring extension is the ideal $\Delta(\mathcal{O}_L/\mathfrak{p}\mathcal{O}_L : \mathcal{O}_K/\mathfrak{p})$ generated by the discriminants of all such bases. Since this ideal lies in $\mathcal{O}_K/\mathfrak{p}$, it will either be (0) or $\mathcal{O}_K/\mathfrak{p}$.

Example 1.7.5. *Let's look at the simple absolute case $\mathbb{Q}(\sqrt{-5}) : \mathbb{Q}$ where the ring of integers is $\mathbb{Z}[\sqrt{-5}]$. The discriminant of this extension is -20. Let's look at the ideal (2). Its quotient ring in $\mathbb{Z}[\sqrt{-5}]$ has representatives $\{0, 1, \sqrt{-5}, 1 + \sqrt{-5}\}$. The single Galois action does not alter this set of representatives, since $-\sqrt{-5} \equiv \sqrt{-5} \pmod 2$.*

Therefore no matter what basis we choose for this quotient ring over the field $\mathbb{Z}/(2)$, the two rows in the matrix will be equal and so the discriminant of this quotient ring extension is clearly 0. This is related to the fact that (2) ramifies, since the minimal polynomial of $\sqrt{-5}$ factors as a square modulo (2) and so the Galois action obviously fixes the quotient ring. It is also related to the fact that $-20 \equiv 0 \pmod 2$, as the next lemma tells us.

Lemma 1.7.6. *Let $L : K$ be a number field extension and $\mathfrak{p}$ a prime ideal in $\mathcal{O}_K$. Then*

$$\Delta(\mathcal{O}_L : \mathcal{O}_K) \equiv \Delta(\mathcal{O}_L/\mathfrak{p}\mathcal{O}_L : \mathcal{O}_K/\mathfrak{p}) \pmod{\mathfrak{p}}$$

Proof. Let $\{a_1 \dots a_n\}$ be an integral basis for L over K. It is clear that this is also an integral basis for $\mathcal{O}_L/\mathfrak{p}\mathcal{O}_L$ over $\mathcal{O}_K/\mathfrak{p}$ as long as $\Delta(a_1 \dots a_n)$ does not lie in $\mathfrak{p}$, since $\{a_1 \dots a_n\}$ is an integral basis for $\mathcal{O}_L/\mathfrak{p}\mathcal{O}_L$ over $\mathcal{O}_K/\mathfrak{p}$ if and only if it is linearly independent over $\mathcal{O}_K/\mathfrak{p}$.

Conversely, starting with an integral basis $\{a_1 \dots a_n\}$ for $\mathcal{O}_L/\mathfrak{p}\mathcal{O}_L$ over $\mathcal{O}_K/\mathfrak{p}$, we can find corresponding representatives $b_1 \dots b_n$ in $\mathcal{O}_L$. These are linearly independent over $\mathcal{O}_K/\mathfrak{p}$, in the sense that $\sum_{i=1}^n k_i b_i = 0$ for $k_i \in \mathcal{O}_K$ implies that each $k_i \in \mathfrak{p}$. Suppose this occurs and let $v = min_{i=1\dots n}(v_\mathfrak{p}(k_i))$. Choose some element $l \in \mathfrak{p}^v$. Then $\sum_{i=1}^n \frac{k_i}{l} b_i = 0$ with $\frac{k_i}{l} \not\equiv 0 \pmod{\mathfrak{p}}$ for at least some i which is a contradiction unless all the k_i were initially 0. Therefore $\{b_1 \dots b_n\}$ are linearly independent over $\mathcal{O}_K$, hence K and we have successfully lifted any basis for $\mathcal{O}_L/\mathfrak{p}\mathcal{O}_L$ over $\mathcal{O}_K/\mathfrak{p}$ to an integral basis for L over K. The calculation of the individual discriminants themselves obviously commute with reduction modulo $\mathfrak{p}$ and so we are done. □

Lemma 1.7.7. *Let $L : K$ be a number field extension. Let $\mathfrak{p}$ be a prime ideal in $\mathcal{O}_K$. Then $\mathfrak{p}$ ramifies if and only if $\mathcal{O}_L/\mathfrak{p}\mathcal{O}_L$ has nontrivial nilpotents.*

Proof. Suppose $\mathfrak{p}$ ramifies so we can factorize $\mathfrak{p}\mathcal{O}_L = \prod_{i=1}^m \mathfrak{P}_i^{e_i}$ with $e_1 > 1$. Choose a nonzero element $a \in (\mathfrak{P}_1^{e_1-1} \prod_{i=2}^m \mathfrak{P}_i^{e_i}) \backslash \mathfrak{p}\mathcal{O}_L$. Then $a^2 \in \mathfrak{p}\mathcal{O}_L$ so $a \pmod{\mathfrak{p}\mathcal{O}_L}$ is nilpotent.

Now suppose we have $x^k \equiv 0 \pmod{\mathfrak{p}\mathcal{O}_L}$ for some $k > 1$ and $x \not\equiv 0 \pmod{\mathfrak{p}\mathcal{O}_L}$. Letting x' be a representative for x in $\mathcal{O}_L$, we have $(\mathfrak{p}\mathcal{O}_L, x')^k \subset \mathfrak{p}\mathcal{O}_L$ yet $(\mathfrak{p}\mathcal{O}_L, x')^{k-1} \not\subset \mathfrak{p}\mathcal{O}_L$. In terms of division of ideals this says that $\mathfrak{p}\mathcal{O}_L | (\mathfrak{p}\mathcal{O}_L, x')^k$ whilst $\mathfrak{p}\mathcal{O}_L \nmid (\mathfrak{p}\mathcal{O}_L, x')^{k-1}$. It follows that in the factorization of $\mathfrak{p}\mathcal{O}_L$, some prime above $\mathfrak{p}$ must have ramification degree greater than 1. □

Theorem 1.7.8. *Let $L : K$ be a number field extension and $\mathfrak{p}$ a prime ideal in $\mathcal{O}_K$. Then $\mathfrak{p}$ ramifies if and only if it divides the relative discriminant $\Delta(L : K)$.*

Proof. $\mathfrak{p}$ divides the relative discriminant if and only if $\Delta(L : K) \equiv (0) \pmod{\mathfrak{p}}$. By Lemma 1.7.6 this occurs if and only if $\Delta(\mathcal{O}_L/\mathfrak{p}\mathcal{O}_L : \mathcal{O}_K/\mathfrak{p}) = (0)$.

Let $\{b_1 \dots b_n\}$ be a basis for $\mathcal{O}_L/\mathfrak{p}\mathcal{O}_L : \mathcal{O}_K/\mathfrak{p}$. Then let M be the corresponding discriminant matrix so that $\Delta(b_1 \dots b_n) = \det(M)^2$. If $\det(M)^2 = 0$ but $\det(M) \neq 0$ then $\mathcal{O}_L/\mathfrak{p}\mathcal{O}_L$ has nontrivial nilpotents and therefore $\mathfrak{p}$ ramifies by Lemma 1.7.7. Assume otherwise that $\det(M) = 0$. By linear algebra this occurs if and only if there are elements $a_i \in \mathcal{O}_K/\mathfrak{p}$ not all zero so that $\sum_{i=1}^n \sigma(a_i b_i) = 0$ for each embedding σ of $L : K$. In particular, the characteristic polynomial of the linear map corresponding to multiplication by $\sum_{i=1}^n a_i b_i$ is x^n, So either $\sum_{i=1}^n a_i b_i = 0$ or it is nilpotent. In the first case, we get a contradiction because $\{b_1 \dots b_n\}$ was chosen as a basis. in the second case, $\mathfrak{p}$ ramifies because of Lemma 1.7.7. This demonstrates the direction that $\Delta(\mathcal{O}_L/\mathfrak{p}\mathcal{O}_L : \mathcal{O}_K/\mathfrak{p}) = (0)$ implies $\mathfrak{p}$ ramifies.

The converse is much easier. If $\mathfrak{p}$ ramifies, then $\mathcal{O}_L/\mathfrak{p}\mathcal{O}_L$ has nontrivial nilpotents and so the extension $\mathcal{O}_L/\mathfrak{p}\mathcal{O}_L : \mathcal{O}_K/\mathfrak{p}$ is not separable. It follows that there exist a pair of embeddings σ_1, σ_2 which coincide in $\mathcal{O}_L/\mathfrak{p}\mathcal{O}_L$ which creates two identical rows in any discriminant matrix. Therefore $\Delta(\mathcal{O}_L/\mathfrak{p}\mathcal{O}_L : \mathcal{O}_K/\mathfrak{P}) = (0)$ and we are done. □

We prove some partial results about discriminants. These results are proven the author.

Lemma 1.7.9. *Let $M : L : K$ be a tower of number fields, with $[M : L] = m$ and $[L : K] = n$. Then by the tower law there are $[M : L][L : K] = mn$ embeddings of M into $\mathbb{C}$ that fix K. Let $\{\sigma_1 \dots \sigma_m\}$ be the set of embeddings of M fixing L and let $\{\tau_1 \dots \tau_n\}$ be the set of embeddings of L fixing K. Then each embedding of M fixing K is given by $\sigma_i \circ \tau_j$ for some i, j.*

Proof. The $\sigma_i \circ \tau_j$ gives us mn embeddings of M into $\mathbb{C}$ that fix K. To see this, let $\{a_1 \dots a_n\}$ be an integral basis for $L : K$ and let $\{c_1 \dots c_m\}$ be an integral basis for $M : L$. It follows that $\{a_1 c_1 \dots a_n c_m\}$ is an integral basis for $M : K$. Then $\sigma_i \circ \tau_j$ gives a well-defined embedding by sending $a_v c_w$ to $\tau_j(a_v)\sigma_i(c_w)$. Check that this is additive and multiplicative and hence an embedding.

It remains to show that these embeddings are distinct. Suppose $\sigma_i \circ \tau_j = \sigma_v \circ \tau_w$. Then restricting to the extension $L : K$ gives $\tau_j = \tau_w$. Since the embeddings form a groupoid, we also get $\sigma_i = \sigma_v$ □

Proposition 1.7.10. *Let $M : L : K$ be number field extensions. Then $\Delta(M : K)$ divides $\Delta(L : K)^{[M:L]}\Delta(M : L)^{[L:K]}$.*

Proof. We use the same setup as in Lemma 1.7.9. Pick integral bases $\{a_1 \dots a_n\}$ and $\{c_1 \dots c_m\}$ for $L : K$ and $M : L$ respectively. We can arrange the discriminant matrix for the integral basis

$\{a_1c_1 \dots a_nc_m\}$ so that it is a Kronecker product of the discriminant matrices of $\{a_1 \dots a_n\}$ and $\{c_1 \dots c_m\}$. We arrange it as

$$\begin{bmatrix} \sigma_1(a_1)\tau_1(c_1) & \dots & \sigma_1(a_1)\tau_1(c_m) & \dots \\ \dots & \dots & \dots & \dots \\ \sigma_1(a_1)\tau_m(c_1) & \dots & \sigma_1(a_1)\tau_m(c_m) & \dots \\ \dots & \dots & \dots & \dots \end{bmatrix}^2$$

so that it becomes the Kronecker product:

$$\left(\begin{bmatrix} \sigma_1(a_1) & \dots & \sigma_1(a_n) \\ \dots & \dots & \dots \\ \sigma_n(a_1) & \dots & \sigma_n(a_n) \end{bmatrix} \otimes \begin{bmatrix} \tau_1(c_1) & \dots & \tau_1(c_m) \\ \dots & \dots & \dots \\ \tau_m(c_1) & \dots & \tau_m(c_m) \end{bmatrix}\right)^2$$

whose determinant is $\Delta(a_1 \dots a_n)^m \Delta(c_1 \dots c_m)^n$. This shows that $\Delta(a_1 \dots a_n)^m \Delta(c_1 \dots c_m)^n$ is contained in the ideal $\Delta(M : K)$ and we are done. □

Proposition 1.7.11. *Let $L : K$ and $M : K$ be number field extensions with $L \cap M = K$. Let LM denote their compositum. Then $\Delta(LM : K)$ divides $\Delta(L : K)^{[M:K]} \Delta(M : K)^{[L:K]}$.*

Proof. By linear algebra, we get

$$[LM : K] = \frac{[L : K][M : K]}{[L \cap M : K]} = [L : K][M : K]$$

We consider the tower $LM : M : K$ and apply Proposition 1.7.10 to get that $\Delta(LM : K)$ divides $\Delta(LM : M)^{[M:K]} \Delta(M : K)^{[LM:M]}$. It remains to show that $\Delta(LM : M)$ divides $\Delta(L : K)$, which occurs if and only if $\Delta(L : K) \subset \Delta(LM : M)$. Every basis for $L : K$ in $\mathcal{O}_L$ will also be an basis for $LM : M$ in $\mathcal{O}_{LM}$ by dimensional considerations. Hence the inclusion $\Delta(L : K) \subset \Delta(LM : M)$ follows. □

1.8 The different ideal

Recall that the trace embedding of a ring of integers into $\mathbb{Z}^n$ is not always surjective. The different ideal is defined to be the inverse of an ideal called Dedekind's complementary module, which measures all elements of the number field which can be trace-embedded into $\mathbb{Z}^n$. This definition will easily extend to relative differents.

The different ideal has many useful applications. It gives us a method for computing the inverse of ideals. It is also heavily related to the discriminant ideal, in almost a dual nature. In fact, we shall see that the discriminant ideal is the relative ideal norm of the different ideal, which gives us a different way to compute the discriminant. The different ideal is easier to work with in proofs, and can be used to prove results about discriminants.

Just as discriminants tell us which primes below an extension will ramify, the different ideal will tell us which primes above an extension are ramified. Again, for monogenic extensions there is a particularly nice description of the different ideal in terms of the minimal polynomial of the primitive element. The different ideal will be $(f'(a))$, which gives some idea as to why it is called the different ideal. We will not be proving these results here. For proofs of these results, see [Conb]. We will follow this source to start with, and then move on to [Lan94].

Definition 1.8.1. *Let $L : K$ be a number field extension. Dedekind's complementary module for this extension is defined as*

$$\mathfrak{C}_{L:K} = \{x \in L : Tr_{L:K}(x\mathcal{O}_L) \subset \mathcal{O}_K\}$$

Definition 1.8.2. *Let $L : K$ be a number field extension. The relative different $\delta_{L:K}$ is defined as the ideal inverse of Dedekind's complementary module. Clearly $\mathcal{O}_L \subset \mathfrak{C}_{L:K}$ and so $\delta_{L:K} \subset \mathcal{O}_L$.*

The next proposition gives us an easy way to calculate the relative different ideal when the base ring is a PID.

Proposition 1.8.3. *Let $L : K$ be a number field extension of degree n so that $\mathcal{O}_K$ is a PID. Let $\{a_1 \dots a_n\}$ be an $\mathcal{O}_K$-basis for $\mathcal{O}_L$, which is possible since the base ring is a PID. Suppose we have found elements $b_1 \dots b_n \in L$ so that for all $i, j = 1 \dots n$,*

$$Tr_{L:K}(a_i b_j) = \delta(i, j)$$

where δ is the Kronecker delta. Then $\{b_1 \dots b_n\}$ is an $\mathcal{O}_K$-basis for $\mathfrak{C}_{L:K}$.

Proof. The fractional $\mathcal{O}_L$-ideal $(b_1 \dots b_n)$ will surject onto $\mathcal{O}_K^n$ under the trace map. We already know that the trace map is injective, and as a result $(b_1 \dots b_n)$ must be the complimentary ideal. In this situation $\{b_1 \dots b_n\}$ is called the dual basis of $\{a_1 \dots a_n\}$. □

Example 1.8.4. *We will compute the different ideal in quadratic fields and show that our hypotheses work in this case.*

In the case $d \not\equiv 1 \pmod 4$ the ring of integers of $\mathbb{Q}(\sqrt{d})$ is $\mathbb{Z}(\sqrt{d})$. We need to find elements b_1, b_2 so that

$$Tr(b_1) = 0, Tr(b_1\sqrt{d}) = 1$$

$$Tr(b_2) = 1, Tr(b_2\sqrt{d}) = 0$$

This is a linear algebra problem, for which there must be a unique solution. A quick computation shows that $b_1 = \frac{\sqrt{d}}{2d}$ and $b_2 = \frac{1}{2}$. As a result $\mathfrak{C}_{\mathbb{Q}(\sqrt{d})} = \frac{1}{2}(1, \frac{\sqrt{d}}{d}) = \frac{1}{2}(\frac{\sqrt{d}}{d})$ and so the different ideal is $(2\sqrt{d})$. In this monogenic case we do get $(2\sqrt{d}) = ((x^2 - d)'(\sqrt{d}))$. Also, $N(2\sqrt{d}) = 4d$ is the discriminant. Since we are in the Galois case, any ideal in the quadratic field dividing the different (hence the discriminant) will be ramified.

From now on we follow [Lan94, Chapter 3].

Proposition 1.8.5. *Let $M : L : K$ be a tower of number fields. Then the relative differents satisfy*

$$\delta_{M:K} = \delta_{M:L}\delta_{L:K}$$

Proof. Upon inverting the ideals, this is equivalent to showing that

$$\mathfrak{C}_{M:K} = \mathfrak{C}_{M:L}\mathfrak{C}_{L:K}$$

Suppose $x \in \mathfrak{C}_{M:L}$ and $y \in \mathfrak{C}_{L:K}$. If we can show that $xy \in \mathfrak{C}_{M:K}$, then it will follow that $\mathfrak{C}_{M:L}\mathfrak{C}_{L:K} \subset \mathfrak{C}_{M:K}$. Note that

$$Tr_{M:K}(xy\mathcal{O}_M) = Tr_{L:K}(Tr_{M:L}(xy\mathcal{O}_M)) = Tr_{L:K}(yTr_{M:L}(x\mathcal{O}_M)) \subset Tr_{L:K}(y\mathcal{O}_L)$$

so $Tr_{L:K}(y\mathcal{O}_L) \subset \mathcal{O}_K$ and indeed $xy \in \mathfrak{C}_{M:K}$. To get the other containment, let $x \in \mathfrak{C}_{M:K}$ and note that

$$Tr_{M:K}(x\mathcal{O}_M) = Tr_{L:K}(Tr_{M:L}(x\mathcal{O}_M)) = Tr_{L:K}(\mathcal{O}_L Tr_{M:L}(x\mathcal{O}_M)) \subset \mathcal{O}_K$$

where the second equality holds because $\mathcal{O}_L \subset \mathcal{O}_M$ and $Tr_{M:L}$ is L-linear. Then it follows that $Tr_{M:L}(x\mathcal{O}_M) \subset \mathfrak{C}_{L:K}$ by applying definitions. $\mathfrak{C}_{L:K}$ is just a fractional ideal of $\mathcal{O}_L$ and so

$$\mathfrak{C}_{L:K}^{-1}Tr_{M:L}(x\mathcal{O}_M) = Tr_{M:L}(x\mathfrak{C}_{L:K}^{-1}\mathcal{O}_M) \subset \mathcal{O}_L \implies x\mathfrak{C}_{L:K}^{-1} \subset \mathfrak{C}_{M:L} \implies x \in \mathfrak{C}_{M:L}\mathfrak{C}_{L:K}$$

It follows that $\mathfrak{C}_{M:K} \subset \mathfrak{C}_{M:L}\mathfrak{C}_{L:K}$ and we are done. □

We will need some basic results about localization.

Definition 1.8.6. *A semilocal ring is a ring with a finite number of maximal ideals.*

A local ring is a ring with a single maximal ideal.

Proposition 1.8.7. *Every semilocal Dedekind domain R is a PID.*

Proof. We apply the Chinese remainder theorem to the finite number of prime ideals $\mathfrak{p}_1 \dots \mathfrak{p}_n$, to show that each one is principal. Since R is Dedekind, there is some nonzero element $a_i \in \mathfrak{p}_i \backslash \mathfrak{p}_i^2$ for each $i = 1 \dots n$. For a fixed i the ideals $\mathfrak{p}_1 \dots \mathfrak{p}_i^2 \dots \mathfrak{p}_n$ are coprime so we can apply the Chinese remainder theorem to get an element $b_i \in R$ so that $b_i \equiv a_i \pmod{\mathfrak{p}_i^2}$ and $b_i \equiv 1 \pmod{\mathfrak{p}_j}$ for all $j \neq i$. As a result the only prime ideal containing b_i is $\mathfrak{p}_i$, and $\mathfrak{p}_i^2$ does not contain it. Hence $\mathfrak{p}_i = (b_i)$ by UPF of ideals. □

Proposition 1.8.8. *Let $\mathfrak{i}, \mathfrak{j}$ be ideals of a Dedekind domain R. If $\mathfrak{i}R_\mathfrak{p} = \mathfrak{j}R_\mathfrak{p}$ for every prime ideal $\mathfrak{p}$ of R, then $\mathfrak{i} = \mathfrak{j}$.*

Proof. This will follow from unique factorization of ideals in Dedekind domains. The localization of R at each prime $\mathfrak{p}$ is a PID with unique maximal ideal $\mathfrak{p}$ and so $\mathfrak{i}R_\mathfrak{p} = \mathfrak{j}R_\mathfrak{p} = \mathfrak{p}^n$ for some positive integer n. This will tell us that the multiplicities of each prime dividing $\mathfrak{i}$ and $\mathfrak{j}$ are the same, so the ideals themselves must be the same by UPF of ideals. □

Definition 1.8.9. *Let $L : K$ be a number field extension of degree n and $\mathfrak{p}$ a prime ideal of $\mathcal{O}_K$. Denote by $\Delta(\mathcal{O}_{L,\mathfrak{p}} : \mathcal{O}_{K,\mathfrak{p}})$ the ideal generated by the discriminants $\Delta(a_1 \dots a_n)$ where $\{a_1 \dots a_n\}$ is a basis for L over K lying in $\mathcal{O}_{L,\mathfrak{p}}$.*

Denote by $\mathfrak{C}_{\mathcal{O}_{L,\mathfrak{p}}:\mathcal{O}_{K,\mathfrak{p}}}$ the ideal

$$\{x \in L : Tr(x)\mathcal{O}_{L,\mathfrak{p}} \subset \mathcal{O}_{K,\mathfrak{p}}\}$$

and then define $\delta_{\mathcal{O}_{L,\mathfrak{p}}:\mathcal{O}_{K,\mathfrak{p}}} = \mathfrak{C}^{-1}_{\mathcal{O}_{L,\mathfrak{p}}:\mathcal{O}_{K,\mathfrak{p}}}$.

Lemma 1.8.10. *Let $L : K$ be an extension of number fields, and $\mathfrak{p}$ a prime ideal of $\mathcal{O}_K$. Then*

$$\Delta(\mathcal{O}_{L,\mathfrak{p}} : \mathcal{O}_{K,\mathfrak{p}}) = \Delta(L : K)\mathcal{O}_{K,\mathfrak{p}}$$

$$\delta_{\mathcal{O}_{L,\mathfrak{p}}:\mathcal{O}_{K,\mathfrak{p}}} = \delta_{L:K}\mathcal{O}_{L,\mathfrak{p}}$$

Proof. Let $\{a_1 \dots a_n\}$ be basis for $L : K$ lying in $\mathcal{O}_{L,\mathfrak{p}}$. Then there is some element $b \in \mathbb{Z}\backslash\mathfrak{p} \cap \mathbb{Z}$ so that $\{ba_1 \dots ba_n\}$ is a basis for $L : K$ lying in $\mathcal{O}_L$, simply by canceling denominators. Note then that b is invertible in $\mathcal{O}_{K,\mathfrak{p}}$ and so $\Delta(a_1 \dots a_n) \in \Delta(L : K)\mathcal{O}_{K,\mathfrak{p}}$, and one containment follows. The reverse containment is obvious, as every basis for $L : K$ lying in $\mathcal{O}_L$ also lies in $\mathcal{O}_{L,\mathfrak{p}}$.

The second statement will follow from $\mathfrak{C}_{\mathcal{O}_{L,\mathfrak{p}}:\mathcal{O}_{K,\mathfrak{p}}} = \mathfrak{C}_{L:K}\mathcal{O}_{L,\mathfrak{p}}$ by inversion. The reverse containment is clear. Let $x \in \mathfrak{C}_{\mathcal{O}_{L,\mathfrak{p}}:\mathcal{O}_{K,\mathfrak{p}}}$. Then $Tr_{L:K}(x\mathcal{O}_{L,\mathfrak{p}}) \subset \mathcal{O}_{K,\mathfrak{p}}$. Again by canceling denominators, there is some $y \in \mathcal{O}_{L,\mathfrak{p}}$ so that $Tr_{L:K}(xy\mathcal{O}_L) \subset \mathcal{O}_K$. Then $xy \in \mathfrak{C}_{L:K}$, in which case $x \in \mathfrak{C}_{L:K}\mathcal{O}_{L,\mathfrak{p}}$.This proves the forward containment. □

Proposition 1.8.11. *Let $L : K$ be a number field extension. Then we have a relation between the relative discriminant and relative different that takes the form*

$$\Delta(L : K) = N_{L:K}(\delta_{L:K})$$

Proof. We will prove the local version for each prime ideal $\mathfrak{p}$ of $\mathcal{O}_K$, that

$$\Delta(\mathcal{O}_{L,\mathfrak{p}} : \mathcal{O}_{K,\mathfrak{p}}) = N_{L:K}(\delta_{\mathcal{O}_{L,\mathfrak{p}}:\mathcal{O}_{K,\mathfrak{p}}})$$

We know that $\mathcal{O}_{K,\mathfrak{p}}$ is local and $\mathcal{O}_{L,\mathfrak{p}}$ is semilocal, hence both are PID's. Therefore there must be some basis $\{a_1 \dots a_n\}$ for L over K lying in $\mathcal{O}_{L,\mathfrak{p}}$ so that

$$\Delta(\mathcal{O}_{L,\mathfrak{p}} : \mathcal{O}_{K,\mathfrak{p}}) = (\Delta(a_1 \dots a_n))$$

Now take the dual basis $\{b_1 \dots b_n\}$ of $\{a_1 \dots a_n\}$ as described in Proposition 1.8.3. This will generate the complementary module $\mathfrak{C}_{\mathcal{O}_{L,\mathfrak{p}}:\mathcal{O}_{K,\mathfrak{p}}}$, which is itself a principal fractional ideal generated by some element $c \in L$. We have $\mathcal{O}_{L,\mathfrak{p}} = \bigoplus_{i=1}^{n} a_i \mathcal{O}_{K,\mathfrak{p}}$ as additive groups and so

$$\mathfrak{C}_{\mathcal{O}_{L,\mathfrak{p}}:\mathcal{O}_{K,\mathfrak{p}}} = (c) = \bigoplus_{i=1}^{n} ca_i \mathcal{O}_{K,\mathfrak{p}}$$

We know that $\Delta(ca_1 \dots ca_n) = N_{L:K}(c)^2 \Delta(a_1 \dots a_n)$ and also

$$(N_{L:K}(c)) = N_{L:K}(\mathfrak{C}_{\mathcal{O}_{L,\mathfrak{p}}:\mathcal{O}_{K,\mathfrak{p}}}) = N_{L:K}(\delta^{-1}_{\mathcal{O}_{L,\mathfrak{p}}:\mathcal{O}_{K,\mathfrak{p}}}) = N_{L:K}(\delta_{\mathcal{O}_{L,\mathfrak{p}}:\mathcal{O}_{K,\mathfrak{p}}})^{-1}$$

The fact $\Delta(a_1 \dots a_n)\Delta(b_1 \dots b_n) = 1$ is easy to spot by using the trace definition of the discriminant. Also $(\Delta(ca_1 \dots ca_n)) = (\Delta(b_1 \dots b_n))$ since both $\{b_1 \dots b_n\}$ and $\{ca_1 \dots ca_n\}$ are $\mathcal{O}_{K,\mathfrak{p}}$-bases for the complementary module and so their discriminants differ by a unit. Putting it all together gives

$$N_{L:K}(\delta_{\mathcal{O}_{L,\mathfrak{p}}:\mathcal{O}_{K,\mathfrak{p}}})^{-2}(\Delta(a_1 \dots a_n)) = (\Delta(ca_1 \dots ca_n)) = (\Delta(b_1 \dots b_n)) = (\Delta(a_1 \dots a_n))^{-1}$$

so that indeed $(\Delta(a_1 \dots a_n)) = \Delta(\mathcal{O}_{L,\mathfrak{p}} : \mathcal{O}_{K,\mathfrak{p}}) = N_{L:K}(\delta_{\mathcal{O}_{L,\mathfrak{p}}:\mathcal{O}_{K,\mathfrak{p}}})$. Now Lemma 1.8.10 gives

$$\Delta(L:K)\mathcal{O}_{K,\mathfrak{p}} = \Delta(\mathcal{O}_{L,\mathfrak{p}} : \mathcal{O}_{K,\mathfrak{p}}) = N_{L:K}(\delta_{\mathcal{O}_{L,\mathfrak{p}}:\mathcal{O}_{K,\mathfrak{p}}}) = N_{L:K}(\delta_{L:K}\mathcal{O}_{K,\mathfrak{p}}) = N_{L:K}(\delta_{L:K})\mathcal{O}_{K,\mathfrak{p}}$$

where the last equality follows since the relative ideal norm commutes with localization. Since $\Delta(L : K)$ and $N_{L:K}(\delta_{L:K})$ agree locally for every prime $\mathfrak{p}$, they must be the same by Proposition 1.8.8. □

Theorem 1.8.12. *Let $M : L : K$ be a tower of fields. Then*

$$\Delta(M:K) = \Delta(L:K)^{[M:L]} N_{L:K}(\Delta(M:L))$$

Proof. We start with the multiplicativity of the different; Proposition 1.8.5 gives $\delta_{M:K} = \delta_{M:L}\delta_{L:K}$. We then apply the relative ideal norm $N_{M:K}$ to both sides and use Proposition 1.8.11 to get

$$\Delta(M:K) = N_{M:K}(\delta_{M:K}) = N_{M:K}(\delta_{M:L}\delta_{L:K}) = N_{L:K}(N_{M:L}(\delta_{M:L}))N_{L:K}(N_{M:L}(\delta_{L:K}))$$

Note that $\delta_{L:K}$ is an ideal in $\mathcal{O}_L$ and the relative norm $N_{M:L}$ has the effect of raising every element in $\mathcal{O}_L$ to the power of $[M : L]$. Then applying Proposition 1.8.11 again gives

$$\Delta(M:K) = N_{L:K}(\Delta(M:L))N_{L:K}(\delta_{L:K}^{[M:L]}) = \Delta(L:K)^{[M:L]} N_{L:K}(\Delta(M:L))$$

□

We have an obvious but crucial corollary.

Corollary 1.8.13. *Let $M : L : K$ be a tower of fields. Then $\Delta(L : K)$ divides $\Delta(M : K)$. In fact, $\Delta(L : K)$ divides $\Delta(M : K)$ with multiplicity at least $[M : L]$.*

Theorem 1.8.14. *Let $L : K$ and $M : K$ be two field extensions so that $M \cap L = K$ and the relative discriminants $\Delta(L : K)$ and $\Delta(M : K)$ are relatively prime. Let LM denote the compositum of L and M. Then the relative discriminant of $LM : K$ is $\Delta(L : K)^{[M:K]}\Delta(M : K)^{[L:K]}$ and the ring of integers of LM is $\mathcal{O}_L\mathcal{O}_M$.*

Proof. From the above corollary, we have $\Delta(L : K)^{[M:K]}|\Delta(LM : K)$ and $\Delta(M : K)^{[L:K]}|\Delta(LM : K)$. By assumption $\Delta(L : K)$ and $\Delta(M : K)$ are relatively prime and so we get $\Delta(L : K)^{[M:K]}\Delta(M : K)^{[L:K]}|\Delta(LM : K)$. Conversely, $\Delta(LM : K)|\Delta(L : K)^{[M:K]}\Delta(M : K)^{[L:K]}$ by Proposition 1.7.11. Therefore as ideals we get $\Delta(LM : K) = \Delta(L : K)^{[M:K]}\Delta(M : K)^{[L:K]}$.

Moreover, the relative discriminant $\Delta(L : K)^{[M:K]}\Delta(M : K)^{[L:K]}$ corresponds to the order $\mathcal{O}_L\mathcal{O}_M$ and so this must be the ring of integers of LM. □

2 Examples of prime decomposition in number fields

In this section we apply the theory of Section 1 to certain families of number fields. We will compute a lot of their invariants, namely the absolute discriminant and the ring of integers. Ultimately we are looking for a classification of the decomposition of integer primes in these extensions. Sections 2.1, 2.3 and 2.5 are based on material from the algebraic number theory course at the summer school PROMYS Europe 2017. Sections 2.2 and 2.4 are original.

2.1 Prime decomposition in quadratic fields

A quadratic field will always be of the form $\mathbb{Q}(\sqrt{d})$ for some square-free integer d. The first thing we need to do is find the ring of integers. We start with the order $\mathbb{Z}(\sqrt{d})$ which has an integral basis $\{1, \sqrt{d}\}$ and we compute its discriminant as

$$\Delta(\mathbb{Z}(\sqrt{d})) = \det \begin{bmatrix} 1 & \sqrt{d} \\ 1 & -\sqrt{d} \end{bmatrix}^2 = (-2\sqrt{d})^2 = 4d$$

Since d is square-free, the only possible prime square factor of the discriminant is 4. Recall our algorithm for finding the ring of integers from Section 1.5. We only have to check whether $\frac{a\cdot 1+b\cdot\sqrt{d}}{2}$ is an algebraic integer for $a, b \in \{0, 1\}$. Now $\frac{1}{2}$ and $\frac{\sqrt{d}}{2}$ are clearly not algebraic integers, and 0 is already in $\mathbb{Z}[\sqrt{d}]$, so we only need to check $\frac{1+\sqrt{d}}{2}$.

$(\frac{1+\sqrt{d}}{2})^2 = \frac{1+d+2\sqrt{d}}{4}$ and so $(\frac{1+\sqrt{d}}{2})^2 - \frac{1+\sqrt{d}}{2} = \frac{d-1}{4}$. As a result the minimal monic polynomial of $\frac{1+\sqrt{d}}{2}$ is $x^2 - x - \frac{d-1}{4}$, and it is clear that $\frac{1+\sqrt{d}}{2}$ is an algebraic integer if and only if $d \equiv 1 \pmod 4$. In these cases the ring of integers is $\mathbb{Z}[\frac{1+\sqrt{d}}{2}]$ since the discriminant now becomes d which is square-free. In all other cases our original order is the ring of integers.

Theorem 2.1.1. *Let $K = \mathbb{Q}(\sqrt{d})$ be a quadratic number field with d a square-free nonzero integer. Then we have the following classification of the ring of integers and discriminant of K:*

- $d \equiv 1 \pmod 4 \implies \mathcal{O}_K = \mathbb{Z}[\frac{1+\sqrt{d}}{2}]$, $\Delta(K) = d$
- $d \not\equiv 1 \pmod 4 \implies \mathcal{O}_K = \mathbb{Z}[\sqrt{d}]$, $\Delta(K) = 4d$

It is convenient that the rings of integers have a primitive element in both cases, and so we can use the original Dedekind–Kummer theorem to decompose primes. Let p be an integer prime. We are in a Galois extension so symbolically, the ways in which p can decompose are

1. $(p) = \mathfrak{p}_1\mathfrak{p}_2$, in which case we say p is split completely.
2. $(p) = \mathfrak{p}^2$, in which case we say p is totally ramified.
3. $(p) = (p)$, in which case we say p is inert.

- First we look at the case $d \equiv 1 \pmod 4$ where the minimal polynomial of the primitive element $\theta = \frac{1+\sqrt{d}}{2}$ is $f = x^2 - x - \frac{d-1}{4}$. We are interested when f has a root modulo p. If $p \neq 2$ then we can write the roots in terms of the quadratic formula as

$$x = \frac{1 \pm \sqrt{d}}{2}$$

  These roots exist and are distinct when d is a quadratic residue modulo p. In that case p is split by Dedekind–Kummer. If $p|d$, then the roots exist but are the same, and so p is ramified. Otherwise, when d is a quadratic non-residue, there is no root and so p is inert. Now we need to make sure 2 behaves nicely. We have

$$f \equiv x^2 - x \equiv x(x-1) \pmod 2$$

  so 2 is split, which is nice because d is a quadratic residue modulo 2, and so it fits into the pattern given by the other primes.

- Now we look at the case $d \not\equiv 1 \pmod 4$ where the minimal polynomial of the primitive element $\theta = \sqrt{d}$ is $f = x^2 - d$. Here it is obvious that f has 2 distinct roots modulo p if and only if d is a quadratic residue modulo p. As long as $p \neq 2$, this occurs if and only if $4d$ is a quadratic residue modulo p, in which case p is split. If $p|d$ then $f \equiv x^2 \pmod p$, giving us a repeat root and so p is ramified. In the case $p \neq 2$, there is no root if and only if $4d$ is a quadratic non-residue modulo p, in which case p is inert. Now in the case $p = 2$ we get

$$d \equiv 0 \pmod 2 \implies x^2 - d \equiv x^2 \pmod 2$$

$$d \equiv 1 \pmod 2 \implies x^2 - d \equiv (x-1)^2 \pmod 2$$

  so 2 always ramifies. Therefore we can say in general that p ramifies if and only if $p|4d$.

Putting all this together gives us a nice classification of prime decomposition in quadratic extensions only in terms of their discriminant:

Theorem 2.1.2. *Let $(\frac{\cdot}{\cdot})$ denote the Legendre symbol. Let p be an integer prime and K a quadratic number field with discriminant Δ. Then*

- *p is inert if and only if $(\frac{\Delta}{p}) = -1$*
- *p is ramified if and only if $(\frac{\Delta}{p}) = 0$*
- *p is split if and only if $(\frac{\Delta}{p}) = 1$*

2.2 Prime decomposition in pure cubic fields

Pure cubic fields are those of the form $\mathbb{Q}(\sqrt[3]{d})$ for a cube-free integer d. In particular we will take d to be positive here since -1 is a cube. Pure cubic fields are never Galois over $\mathbb{Q}$ because the minimal polynomial of $\sqrt[3]{d}$ is $x^3 - d$ which has 1 real root and 2 complex roots. Hence $\mathbb{Q}(\sqrt[3]{d})$ has 1 real embedding and 2 complex embeddings. In this subsection we will study pure cubic fields.

Let $K = \mathbb{Q}(\sqrt[3]{d})$. We will find the ring of integers of K. We start with the order $\mathbb{Z}[\sqrt[3]{d}]$ which has an integral basis $\{1, \sqrt[3]{d}, \sqrt[3]{d^2}\}$. We compute its discriminant as

$$\Delta(\mathbb{Z}[\sqrt[3]{d}]) = \det \begin{bmatrix} 1 & \sqrt[3]{d} & \sqrt[3]{d^2} \\ 1 & \zeta_3 \sqrt[3]{d} & \zeta_3^2 \sqrt[3]{d^2} \\ 1 & \zeta_3^2 \sqrt[3]{d} & \zeta_3 \sqrt[3]{d^2} \end{bmatrix}^2 =$$

$$= (1 \cdot (\zeta_3^2 d - \zeta_3 d) - \sqrt[3]{d} \cdot (\zeta_3 \sqrt[3]{d^2} - \zeta_3^2 \sqrt[3]{d^2}) + \sqrt[3]{d^2} \cdot (\zeta_3^2 \sqrt[3]{d} - \zeta_3 \sqrt[3]{d}))^2 =$$

$$= 9d^2(\zeta_3^2 - \zeta_3)^2 = 9d^2(\frac{-1+\sqrt{-3}}{2} + \frac{1+\sqrt{-3}}{2})^2 = -27d^2$$

The integer primes whose square divide $-27d^2$ are 3 and the prime divisors of d. We call these primes *suitable* in this subsection. We need to check when $\frac{a+b\sqrt[3]{d}+c\sqrt[3]{d^2}}{p}$ is an algebraic integer for a suitable prime p and integers $0 \leq a, b, c < p$. We can do so by computing the minimal polynomial using remark 1.5.9. Fix $\{1, \sqrt[3]{d}, \sqrt[3]{d^2}\}$ as a $\mathbb{Q}$-basis for the K. We find that the matrix corresponding to multiplication by $\frac{a+b\sqrt[3]{d}+c\sqrt[3]{d^2}}{p}$ is

$$M = \begin{bmatrix} \frac{a}{p} & \frac{cd}{p} & \frac{bd}{p} \\ \frac{b}{p} & \frac{a}{p} & \frac{cd}{p} \\ \frac{c}{p} & \frac{b}{p} & \frac{a}{p} \end{bmatrix}$$

The characteristic polynomial of M is

$$f = x^3 - \frac{3a}{p}x^2 + \frac{3bcd - 3a^3}{p^2}x - \frac{a^3 + b^3d + c^3d^2 - 3abcd}{p^3}$$

Suppose $p \neq 3$. Then we must have $a = 0$ in order for $\frac{3a}{p}$ to be an integer. This gives

$$f = x^3 + \frac{3bcd}{p^2}x - \frac{b^3d + c^3d^2}{p^3}$$

$\frac{3bcd}{p^2}$ must be an integer and since $p \neq 3$ it must divide b, c or d. p divides b or c if and only if they equal 0 by our restriction. However if either of them equals 0 then either $\frac{c^3d^2}{p^3}$ or $\frac{b^3d}{p^3}$ must equal 0. Since $\frac{b^3d+c^3d^2}{p^3}$ must be an integer then either $\frac{c^3d^2}{p^3}$ or $\frac{b^3d}{p^3}$ is an integer. Either way, p must divide d otherwise both b and c would be 0, in which case $f = x^3$.

As a result $\frac{b^3+c^3d}{p}$ must be an integer since d is cube-free. We must set $b = 0$ because p divides c^3d and so it divides b^3. In this case $\frac{c^3d^2}{p^3}$ must be an integer so we must have $p^2|d$ in order to avoid

p dividing c. But in this case f becomes $x^3 - \frac{c^3d^2}{p^3}$, whose root is $\frac{\sqrt[3]{d^2}}{p}$. This is already in our order because $p^3|d^2$ so the fraction cancels. Therefore we get no new algebraic integers in the case $p \neq 3$.

In the case $p = 3$ we get

$$f = x^3 - ax^2 + \frac{bcd - a^3}{3}x - \frac{a^3 + b^3d + c^3d^2 - 3abcd}{27}$$

In the case $3|d$ we must have $a = 0$ anyway since $\frac{bcd-a^3}{3}$ is to be an integer, implying that 3 divides a. This case is exactly the same as before and so it gives no new algebraic integers. This leaves us with the case $p = 3$ and $3 \nmid d$, where the order is at most one algebraic integer away from becoming the ring of integers, whose discriminant would have to be $-3d^2$. This can be seen since the only suitable prime left is 3 and $3 \nmid d$.

We focus on the remaining case $p = 3$ and $3 \nmid d$. In the case $a = 0$, $\frac{bcd-a^3}{3}$ must be an integer so either $b = 0$ or $c = 0$ since 3 divides bcd. In these cases f becomes either $x^3 - \frac{b^3d}{27}$ or $x^3 - \frac{c^3d^2}{27}$. Since $3 \nmid d$ we get that both b and c are 0 which makes $f = x^3$.

In the case that either $b = 0$ or $c = 0$ we must have $a = 0$ since $\frac{bcd-a^3}{3}$ is an integer. This is the previous case which gives us $f = x^3$. Therefore we can assume that none of the coefficients are 0. We can also assume that $a = 1$. Now we have some casework to do.

- In the case $b = c = 1$, we get $f = x^3 - x^2 + \frac{d-1}{3}x - \frac{1-2d+d^2}{27}$. $\frac{d-1}{3}$ must be an integer so $d \equiv 1 \pmod 3$. Also $\frac{1-2d+d^2-3d}{27} = \frac{(d-1)^2}{27}$ must be an integer which occurs if and only if $d \equiv 1 \pmod 9$ which automatically implies $d \equiv 1 \pmod 3$. So in the case $d \equiv 1 \pmod 9$ we get a nontrivial algebraic integer $\frac{1+\sqrt[3]{d}+\sqrt[3]{d^2}}{3}$.

- In the case $b = c = 2$, we get $f = x^3 - x^2 + \frac{4d-1}{3}x - \frac{1-4d+8d^2}{27}$. $\frac{4d-1}{3}$ must be an integer so $d \equiv 1 \pmod 3$. Also $\frac{8d^2-4d+1}{27}$ must be an integer but $d \equiv 1 \pmod 3$ which means $8d^2 - 4d + 1 \equiv 2 \pmod 3$ so the numerator is not divisible by 3. This case fails to give an algebraic integer.

- In the case $b = 2, c = 1$, we get $f = x^3 - x^2 + \frac{2d-1}{3}x - \frac{1+2d+d^2}{27}$. $\frac{2d-1}{3}$ must be an integer so $d \equiv 2 \pmod 3$. Also $\frac{1+2d+d^2}{27} = \frac{(d+1)^2}{27}$ must be an integer which occurs if and only if $d \equiv 8 \pmod 9$ which automatically implies $d \equiv 2 \pmod 3$. So in the case $d \equiv 8 \pmod 9$ we have the nontrivial algebraic integer $\frac{1+2\sqrt[3]{d}+\sqrt[3]{d^2}}{3}$.

- In the case $b = 1, c = 2$, we get $f = x^3 - x^2 + \frac{2d-1}{3}x - \frac{1-5d+8d^2}{27}$. $\frac{2d-1}{3}$ must be an integer so $d \equiv 2 \pmod 3$. Also $\frac{8d^2-5d+1}{27}$ must be an integer but $d \equiv 2 \pmod 3$ which means $8d^2 - 5d + 1 \equiv 1 \pmod 3$ so the numerator is not divisible by 3. This case fails to give an algebraic integer.

Now we show that in the special cases $d \equiv \pm 1 \pmod 9$ the new algebraic integers are primitive elements for the ring of integers. In the case $d \equiv 1 \pmod 9$ the integral basis for the ring of integers of $\mathbb{Q}(\sqrt[3]{d})$ is $\{1, \sqrt[3]{d}, \frac{1+\sqrt[3]{d}+\sqrt[3]{d^2}}{3}\}$. The minimal polynomial of $\frac{1+\sqrt[3]{d}+\sqrt[3]{d^2}}{3}$ is $x^3 - x^2 + \frac{d-1}{3}x - \frac{(d-1)^2}{27}$. We have $(\frac{1+\sqrt[3]{d}+\sqrt[3]{d^2}}{3})^2 = \frac{2d+1+(d+2)\sqrt[3]{d}+3\sqrt[3]{d^2}}{9}$ and $(\frac{1+\sqrt[3]{d}+\sqrt[3]{d^2}}{3})^2 - \frac{1+\sqrt[3]{d}+\sqrt[3]{d^2}}{3} = \frac{2d-2+(d-1)\sqrt[3]{d}}{9}$ where

$2d-2$ and $d-1$ are integers divisible by 9. Therefore $\sqrt[3]{d} \in \mathbb{Z}[\frac{1+\sqrt[3]{d}+\sqrt[3]{d^2}}{3}]$ and so this is the ring of integers.

In the case $d \equiv -1 \pmod 9$ the integral basis for the ring of integers of $\mathbb{Q}(\sqrt[3]{d})$ is $\{1, \sqrt[3]{d}, \frac{1+2\sqrt[3]{d}+\sqrt[3]{d^2}}{3}\}$. The minimal polynomial of $\frac{1+2\sqrt[3]{d}+\sqrt[3]{d^2}}{3}$ is $x^3 - x^2 + \frac{2d-1}{3}x - \frac{(d+1)^2}{27}$. We have $(\frac{1+2\sqrt[3]{d}+\sqrt[3]{d^2}}{3})^2 = \frac{4d+1+(d+4)\sqrt[3]{d}+6\sqrt[3]{d^2}}{9}$ and $(\frac{1+2\sqrt[3]{d}+\sqrt[3]{d^2}}{3})^2 - 2\frac{1+2\sqrt[3]{d}+\sqrt[3]{d^2}}{3} = \frac{4d-5+(d-8)\sqrt[3]{d}}{9}$ where $4d-5$ and $d-8$ are integers divisible by 9. Therefore $\sqrt[3]{d} \in \mathbb{Z}[\frac{1+2\sqrt[3]{d}+\sqrt[3]{d^2}}{3}]$ and so this is the ring of integers.

Theorem 2.2.1. *We have the following classification for the discriminant and ring of integers of a pure cubic number field $\mathbb{Q}(\sqrt[3]{d})$ for a positive, cube-free integer d:*

- $d \equiv 1 \pmod 9 \implies \mathcal{O}_K = \mathbb{Z}[\frac{1+\sqrt[3]{d}+\sqrt[3]{d^2}}{3}], \Delta(K) = -3d^2$
- $d \equiv -1 \pmod 9 \implies \mathcal{O}_K = \mathbb{Z}[\frac{1+2\sqrt[3]{d}+\sqrt[3]{d^2}}{3}], \Delta(K) = -3d^2$
- $d \not\equiv \pm 1 \pmod 9 \implies \mathcal{O}_K = \mathbb{Z}[\sqrt[3]{d}], \Delta(K) = -27d^2$

We will now look at how integer primes decompose in each of these extensions. Again we are lucky because in all cases there is a primitive element, so we can use the original Dedekind–Kummer theorem. However the special cases are quite difficult so we will be using the generalized version.

- Let $d \not\equiv \pm 1 \pmod 9$, in which case the ring of integers is $\mathbb{Z}[\sqrt[3]{d}]$. The primitive element has minimal polynomial $x^3 - d$. Let p be an integer prime not dividing the discriminant $-27d^2$. Then the roots of $x^3 \equiv d \pmod p$ will determine the decomposition of p.

  If $3| \ (p-1) = |(\mathbb{Z}/(p))^\times|$ then $\mathbb{Z}/(p)^\times$ has a subgroup of order 3 by Cauchy's theorem for groups. Therefore the equation $x^3 \equiv 1 \pmod p$ has 3 roots, namely the elements of the subgroup of order 3. As a result the equation $x^3 \equiv d \pmod p$ either has no roots or 3 roots, because $\mathbb{Z}/(p)$ has the third roots of unity. These correspond to the prime p being inert or split completely respectively. If $x^3 \equiv d \pmod p$ has a root we say d is a cubic residue modulo p.

  If $3 \not| (p-1)$ then it turns out that every element of $(\mathbb{Z}/(p))^\times$ is a cubic residue with each equation of the form $x^3 \equiv a \pmod p$ for $a \in (\mathbb{Z}/(p))^\times$ having exactly one root. Suppose not, then one of these equations will have at least 2 roots by the pigeonhole principle, and by factoring this polynomial we find that in fact it has 3 roots. Let $x^3 - a$ be such a polynomial, with roots a_1, a_2, a_3 modulo p. Then the quotients $\frac{a_1}{a_2}, \frac{a_2}{a_3}, \frac{a_3}{a_1}$ are roots of unity. These form a subgroup of order 3 and so actually $3|(p-1)$, a contradiction.

  Therefore in the case $3 \not| (p-1)$ when p does not divide the discriminant, the polynomial $x^3 - d$ splits into two factors modulo p. Hence the prime p will split incompletely into 2 prime ideals.

  Now let p be an integer prime dividing the discriminant. If it divides d then $x^3 - d \equiv x^3 \pmod 3$ and so the prime p totally ramifies as

$$(p) = (p, \sqrt[3]{d})^3$$

If $p = 3$ and it does not divide d, then $x^3 - d \equiv (x - d)^3 \pmod 3$ and so we get total ramification of 3 as

$$(3) = (3, \sqrt[3]{d} - d)^3$$

- Let $d \equiv 1 \pmod 9$, in which case the ring of integers is $\mathbb{Z}[\frac{1+\sqrt[3]{d}+\sqrt[3]{d^2}}{3}]$. The primitive element has minimal polynomial $x^3 - x^2 + \frac{d-1}{3}x - \frac{(d-1)^2}{27}$. We have $x^3 - x^2 + \frac{d-1}{3}x - \frac{(d-1)^2}{27} \equiv x^3 - x^2 \equiv x^2(x-1) \pmod 3$ and so

$$(3) = (3, \sqrt[3]{d})^2(3, \sqrt[3]{d} - 1)$$

  Note that the ideal (3) of $\mathbb{Z}[\frac{1+\sqrt[3]{d}+\sqrt[3]{d^2}}{3}]$ is clearly contained in $\mathbb{Z}[\sqrt[3]{d}]$. Since the conductor ideal corresponding to the order $\mathbb{Z}[\sqrt[3]{d}]$ contains (3), all other integer primes, being coprime to 3, will factor in the same way as in the first case.

- Let $d \equiv -1 \pmod 9$, in which case the ring of integers is $\mathbb{Z}[\frac{1+2\sqrt[3]{d}+\sqrt[3]{d^2}}{3}]$. The primitive element has minimal polynomial $x^3 - x^2 + \frac{2d-1}{3}x - \frac{(d+1)^2}{27}$. We have $x^3 - x^2 + \frac{2d-1}{3}x - \frac{(d+1)^2}{27} \equiv x^3 - x^2 \equiv x^2(x-1) \pmod 3$ and so

$$(3) = (3, \sqrt[3]{d})^2(3, \sqrt[3]{d} - 1)$$

  Note that the ideal (3) of $\mathbb{Z}[\frac{1+2\sqrt[3]{d}+\sqrt[3]{d^2}}{3}]$ is clearly contained in $\mathbb{Z}[\sqrt[3]{d}]$. Since the conductor ideal corresponding to the order $\mathbb{Z}[\sqrt[3]{d}]$ contains (3), all other integer primes, being coprime to 3, will factor in the same way as in the first case.

Theorem 2.2.2. *We have the following classification for prime decomposition of an integer prime p in a pure cubic number field $\mathbb{Q}(\sqrt[3]{d})$, for d a positive integer which is cube-free.*

- *In the case $p = 3$, if $d \equiv \pm 1 \pmod 9$ then*

$$(3) = (3, \sqrt[3]{d})^2(3, \sqrt[3]{d} - 1)$$

  and otherwise if $d \not\equiv \pm 1 \pmod 9$ then

$$(3) = (3, \sqrt[3]{d} - d)^3$$

- *If $p|d$ then p totally ramifies as*

$$(p) = (p, \sqrt[3]{d})^3$$

- *If $p \equiv 1 \pmod 3$ and d is a cubic residue modulo p, then p splits completely.*
- *If $p \equiv 1 \pmod 3$ and d is a cubic non-residue modulo p, then p is inert.*
- *If $p \equiv 2 \pmod 3$ and $p \not| d$, then p splits incompletely as the product of 2 prime ideals, one with inertia degree 1 and the other with inertia degree 2.*

2.3 Prime decomposition in cyclotomic fields

A cyclotomic field in general is a field of the form $\mathbb{Q}(\zeta_n)$ for some primitive n^{th} root of unity ζ_n. Firstly we will compute the discriminant and ring of integers of cyclotomic fields in general. Then we will focus on prime decomposition in prime power cyclotomic fields, which are of the form $\mathbb{Q}(\zeta_{p^m})$ for some prime $p \geq 3$ and a positive integer m.

We know that ζ_n is a root of $x^n - 1$, but this polynomial is not irreducible of course. For every number d that divides n, the polynomial $x^d - 1$ divides $x^n - 1$.

We denote by Φ_n the minimal polynomial of ζ_n. If n is prime we know that $\Phi_n = \frac{x^n-1}{x-1}$ by using Eisenstein's irreducibility criterion. We also know that

$$\frac{x^{p^m} - 1}{x - 1} = \prod_{d=1}^{m} \frac{x^{p^d} - 1}{x^{p^{d-1}} - 1}$$

where each term on the RHS is irreducible, again by Eisenstein. Therefore we must have

$$\Phi_{p^m} = \frac{x^{p^m} - 1}{x^{p^{m-1}} - 1}$$

since the other factors are minimal polynomials for ζ_{p^k} with $k < m$. Now look at the general case. Factorize n into primes as $\prod_{i=1}^{s} p_i^{e_i}$. Then $\Phi_{p_i^{e_i}} | \Phi_n$ for each i as we've discussed. Moreover, the $\Phi_{p_i^{e_i}}$ are pairwise coprime (since they do not share any roots) and so we must have

$$\Phi_n = \prod_{i=1}^{s} \Phi_{p_i^{e_i}}$$

Looking at degrees, we have $\deg(\Phi_{p_i^{e_i}}) = p_i^{e_i} - p_i^{e_i - 1} = \phi(p_i^{e_i})$, which is Euler's Totient function. Hence $\deg(\Phi_n) = \phi(n)$ by the weak multiplicativity of ϕ. As a result $[\mathbb{Q}(\zeta_n) : \mathbb{Q}] = \phi(n)$.

It is clear that cyclotomic fields are Galois extensions of $\mathbb{Q}$. For the cyclotomic field $\mathbb{Q}(\zeta_n)$, we will start with the order $\mathbb{Z}[\zeta_n]$, which is monogenic. Therefore its discriminant will equal the discriminant of the minimal polynomial of ζ_n, which is Φ_n. We will first compute the discriminant in the case $n = p^m$. We start with

$$disc(\Phi_{p^m}) = \prod_{j<k:(j,p)=(k,p)=1} (\zeta_{p^m}^j - \zeta_{p^m}^k)^2 =$$
$$= (-1)^{\frac{\phi(p^m)(\phi(p^m)-1)}{2}} \prod_{j \neq k:(j,p)=(k,p)=1} (\zeta_{p^m}^j - \zeta_{p^m}^k)$$

We have $(-1)^{\frac{\phi(p^m)(\phi(p^m)-1)}{2}} = (-1)^{\frac{\phi(p^m)}{2}}$ since $\phi(p^m)-1$ is odd. Note that $(x^{p^{m-1}}-1)\Phi_{p^m} = x^{p^m} - 1$ and differentiating gives us

$$p^m x^{p^m - 1} = (x^{p^{m-1}} - 1)\Phi'_{p^m} + p^{m-1} x^{p^{m-1} - 1} \Phi_{p^m}$$

However, we have $\Phi_{p^m}(\zeta_{p^m}^j) = 0$ for all j coprime to p and also

$$\Phi'_{p^m} = \sum_{j:(j,p)=1} \prod_{k \neq j:(k,p)=1} (x - \zeta_{p^m}^k)$$

which means that for all j coprime to p we have

$$\Phi'_{p^m}(\zeta_{p^m}^j) = \prod_{k \neq j:(k,p)=1} (\zeta_{p^m}^j - \zeta_{p^m}^k)$$

Putting it all together gives us

$$\prod_{k \neq j:(k,p)=1} (\zeta_{p^m}^j - \zeta_{p^m}^k) = \frac{p^m \zeta_{p^m}^{j(p^m-1)}}{\zeta_{p^m}^{jp^{m-1}} - 1}$$

so the formula for the discriminant becomes

$$disc(\Phi_{p^m}) = (-1)^{\frac{\phi(p^m)}{2}} \prod_{j:(j,p)=1} \frac{p^m \zeta_{p^m}^{j(p^m-1)}}{\zeta_{p^m}^{jp^{m-1}} - 1} = (-1)^{\frac{\phi(p^m)}{2}} \frac{p^{m \cdot \phi(p^m)}}{\prod_{j:(j,p)=1} (\zeta_{p^m}^{jp^{m-1}} - 1)}$$

where $\zeta_{p^m}^{jp^{m-1}}$ is a primitive p^{th} root of unity and so

$$\prod_{j:(j,p)=1} (\zeta_{p^m}^{jp^{m-1}} - 1) = N_{\mathbb{Q}(\zeta_p)}(\zeta_p - 1)^{\frac{\phi(p^m)}{p-1}} = \Phi_p(1)^{\frac{\phi(p^m)}{p-1}}$$

Note that $\Phi_p(1) = \sum_{i=0}^{p-1} 1^i = p$ and so finally

$$disc(\Phi_{p^m}) = (-1)^{\frac{\phi(p^m)}{2}} \frac{p^{m\phi(p^m)}}{p^{\frac{\phi(p^m)}{p-1}}} = (-1)^{\frac{\phi(p^m)}{2}} p^{\phi(p^m)(m - \frac{1}{p-1})}$$

We keep the discriminant in this format so that we can generalize to all cyclotomic fields. There is a formula for the discriminant of a product of polynomials in terms of their discriminants and pairwise resultants. For $n = \prod_{i=1}^s p_i^{e_i}$ it tells us that

$$disc(\Phi_n) = disc(\prod_{i=1}^s \Phi_{p_i^{e_i}}) = (\prod_{i=1}^s disc(\Phi_{p_i^{e_i}}))(\prod_{i<j} res(\Phi_{p_i^{e_i}}, \Phi_{p_j^{e_j}}))$$

Note that the resultant $res(\Phi_{p_i^{e_i}}, \Phi_{p_j^{e_j}})$ is invariant under addition of a multiple of one term to the other. Therefore we technically need to carry out the Euclidean algorithm on $(\Phi_{p_i^{e_i}}, \Phi_{p_j^{e_j}})$. Firstly we do it on $(x^{p_i^{e_i}} - 1, x^{p_j^{e_j}} - 1)$ to get $x^{(p_i^{e_i}, p_j^{e_j})} - 1 = x - 1$, because this is actually the same as doing the Euclidean algorithm on $(p_i^{e_i}, p_j^{e_j})$ which are coprime. Hence there are integer polynomials A, B so that

$$A \frac{x^{p_i^{e_i}} - 1}{x - 1} + B \frac{x^{p_j^{e_j}} - 1}{x - 1} = 1$$

Since $\Phi_{p_i^{e_i}} | \frac{x^{p_i^{e_i}}-1}{x-1}$ and $\Phi_{p_j^{e_j}} | \frac{x^{p_j^{e_j}}-1}{x-1}$ we get some corresponding integer polynomials A', B' satisfying

$$A'\Phi_{p_i^{e_i}} + B'\Phi_{p_j^{e_j}} = 1$$

so the resultant $res(\Phi_{p_i^{e_i}}, \Phi_{p_j^{e_j}})$ is either 1 or -1. Brill's theorem will tell us the sign of the discriminant so we do not need to worry. $\frac{\phi(n)}{2}$ is the number of pairs of complex conjugate embeddings and so the sign is $(-1)^{\frac{\phi(n)}{2}}$. Therefore we have

$$disc(\Phi_n) = (-1)^{\frac{\phi(n)}{2}} |\prod_{i=1}^{s} disc(\Phi_{p_i^{e_i}})| = (-1)^{\frac{\phi(n)}{2}} \prod_{i=1}^{s} p_i^{\phi(p_i^{e_i})(e_i - \frac{1}{p_i - 1})}$$

We will find the ring of integers of prime power cyclotomic fields first. The only prime whose square might divide the discriminant of Φ_{p^e} is p. Therefore the conductor of $\mathbb{Z}[\zeta_{p^e}]$ contains the ideal (p) by Corollary 1.5.22. We factorize (p) in $\mathbb{Z}[\zeta_{p^e}]$ as

$$\Phi_{p^e}(1) = p = \prod_{k:(k,p=1)} (\zeta_{p^e}^k - 1) = (\zeta_{p^e} - 1)^{\phi(p^e)} \prod_{k:(k,p=1)} \frac{(\zeta_{p^e}^k - 1)}{(\zeta_{p^e} - 1)}$$

where $\frac{(\zeta_{p^e}^k - 1)}{(\zeta_{p^e} - 1)} = \sum_{j=0}^{k-1} \zeta_{p^e}^j$ is a unit in the order, since it has norm 1. Therefore we have

$$(p) = ((\zeta_{p^e} - 1)^{\phi(p^e)} \prod_{k:(k,p=1)} \frac{(\zeta_{p^e}^k - 1)}{(\zeta_{p^e} - 1)}) = (\zeta_{p^e} - 1)^{\phi(p^e)}$$

This factorization will also occur in the ring of integers. In particular, the ideal $(\zeta_{p^e} - 1)$ must be prime in the ring of integers because $\phi(p^e) = [\mathbb{Q}(\zeta_{p^e}) : \mathbb{Q}]$. It follows that the conductor ideal is some power of this ideal. Every prime ideal except $(\zeta_{p^e} - 1)$ is therefore invertible because it is coprime to the conductor ideal. However, $(\zeta_{p^e} - 1)$ is also invertible since it is principal, and so every prime ideal of $\mathbb{Z}[\zeta_{p^e}]$ is invertible. It follows that $\mathbb{Z}[\zeta_{p^e}]$ is the ring of integers.

Now in general $\mathbb{Q}(\zeta_n)$ is the compositum of the fields $\mathbb{Q}(\zeta_{p_i^{e_i}})$ and so the discriminant of $\mathbb{Q}(\zeta_{p_i^{e_i}})$ will divide the discriminant of $\mathbb{Q}(\zeta_n)$ for each prime power divisor $p_i^{e_i}$ of n. Moreover we have shown that the discriminant of $\mathbb{Q}(\zeta_n)$ divides the product of the discriminants of $\mathbb{Q}(\zeta_{p_i^{e_i}})$, which are pairwise coprime. Therefore we must have that the discriminant of $\mathbb{Q}(\zeta_n)$ is the discriminant of Φ_n up to sign. But they have the same sign by checking with Brill's theorem. Therefore the ring of integers of $\mathbb{Q}(\zeta_n)$ is indeed $\mathbb{Z}[\zeta_n]$.

Theorem 2.3.1. *Let $\mathbb{Q}(\zeta_n)$ be a cyclotomic field. Factorize n into integer primes as $\prod_{i=1}^{s} p_i^{e_i}$. Then its discriminant is*

$$\Delta(\mathbb{Q}(\zeta_n)) = (-1)^{\frac{\phi(n)}{2}} \prod_{i=1}^{s} p_i^{\phi(p_i^{e_i})(e_i - \frac{1}{p_i - 1})}$$

and its ring of integers is $\mathcal{O}_{\mathbb{Q}(\zeta_n)} = \mathbb{Z}[\zeta_n]$.

Now we will classify the prime decomposition in prime power cyclotomic fields. The ring of integers $\mathbb{Z}[\zeta_{p^m}]$ is monogenic with minimal polynomial Φ_{p^m}. We have already classified the ramified primes; the only prime that ramifies is p and it ramifies completely as

$$(p) = (\zeta_{p^m} - 1)^{\phi(p^m)}$$

Let q be an integer prime not equal to p. Then Φ_{p^m} will factor into d irreducible polynomials modulo q for some d that divides $\phi(p^m)$, each with degree $c = \frac{\phi(p^m)}{d}$. This is due to fact that we are in a Galois extension. Let's say that the factorization is

$$\Phi_{p^m} \equiv \prod_{i=1}^{d} f_i \pmod{q}$$

Then the f_i are pairwise coprime, since q does not ramify. There must be a Galois extension of $\mathbb{Z}/(q) = \mathbb{F}_q$ of degree c where f_1 splits. This extension is the finite field $\mathbb{F}_{q^c}$ with q^c elements. Note that two finite fields are isomorphic if and only if they have the same number of elements. Therefore each of the f_i splits in $\mathbb{F}_{q^c}$, along with Φ_{p^m}.

If Φ_{p^m} has a root in $\mathbb{F}_{q^c}$, it will split completely there and so we can infer that in $\mathbb{F}_q$ it will factor into at least d irreducible factors. Since $\Phi_{p^m} = \frac{x^{p^m}-1}{x^{p^{m-1}}-1}$ we need a root of $x^{p^m} - 1$ which is not a root of $x^{p^{m-1}} - 1$. The roots of $x^{p^m} - 1$ form a subgroup of $\mathbb{F}_{q^c}^\times$ for all c.

A root of $x^{p^m} - 1$ which is not a root of $x^{p^{m-1}} - 1$ will generate all the other roots. Therefore in order for $x^{p^m} - 1$ to split in $\mathbb{F}_{q^c}^\times$, we need p^m distinct roots of $x^{p^m} - 1$. This is equivalent to finding a subgroup of $\mathbb{F}_{q^c}^\times$ of order p^m, for which we require $p^m | q^c - 1$. Conversely, if $p^m | q^c - 1$, then there is a subgroup of order p^m by Sylow's theorem, and all of its elements will be roots of $x^{p^m} - 1$.

Therefore we get the following result. q splits into a product of d ideals if and only if $p^m | q^{\frac{\phi(p^m)}{d}} - 1$ and $p^m \nmid q^{\frac{\phi(p^m)}{k}} - 1$ for any k divisible by d. This will ensure that q splits into d ideals and no more than d ideals. This condition can be rewritten as $\mathrm{ord}_{\mathbb{Z}/(p^m)}(q) = \frac{\phi(p^m)}{d}$.

Theorem 2.3.2. *We get the following classification of prime decomposition in a prime power cyclotomic field $\mathbb{Q}(\zeta_{p^m})$:*

- *The prime p ramifies completely as*

$$(p) = (\zeta_{p^m} - 1)^{\phi(p^m)}$$

- *Let q be an integer prime not equal to p. Write $\mathrm{ord}_{\mathbb{Z}/(p^m)}(q) = \frac{\phi(p^m)}{d}$ for some positive integer d. Then q factors into a product of d prime ideals, each with inertia degree $\frac{\phi(p^m)}{d}$.*

2.4 Cubic fields in general

There are three types of cubic fields, sorted by the nature of their embeddings.

- Cyclic cubic fields, which are Galois cubic fields, and must necessarily be totally real.
- Totally real cubic fields which are not Galois.
- Cubic fields with one real embedding and a pair of complex conjugate embeddings.

For non-Galois cubic fields, their normal closure is a degree 2 extension by Galois theory. In these cases we will find the quadratic field that we must append to the cubic field in order to get the normal closure.

Proposition 2.4.1. *Let K be a number field with absolute discriminant $\Delta(K)$. Then the normal closure of K contains the quadratic subfield $\mathbb{Q}(\sqrt{\Delta(K)})$ when $\Delta(K)$ is not a perfect square.*

Proof. The determinant $\sqrt{\Delta(K)}$ of the discriminant matrix is written in terms of elements of the different embeddings of K. The normal closure contains these elements and hence $\sqrt{\Delta(K)}$. □

Corollary 2.4.2. *A number field K where $[K : \mathbb{Q}]$ is odd and whose absolute discriminant is not a perfect square cannot be Galois over $\mathbb{Q}$.*

In particular, any cubic field is an extension of $\mathbb{Q}$ by a root of some irreducible polynomial $x^3 + ax + b$ whose discriminant is $-4a^3 - 27b^2$. If this discriminant is not a square, then the cubic field is not Galois. The following theorem proves the converse.

Theorem 2.4.3. *The number field $\mathbb{Q}(x)/(x^3 + ax + b)$ is Galois over $\mathbb{Q}$ for some irreducible $x^3 + ax + b$ if and only if the discriminant $-4a^3 - 27b^2$ is a perfect square.*

Proof. We've already proven the forward direction. Now assume that the above field is not Galois. Let's say we have a root θ of x^3+ax+b so that $x^3+ax+b = (x-\theta)(x^2+vx+w)$ for some algebraic numbers v, w. Then we must have $v = \theta$, $w - \theta^2 = a$ and $b = -\theta w$ by comparing coefficients. Let θ_1, θ_2 be the other roots so that $\theta_1 + \theta_2 = -\theta$ and $\theta_1\theta_2 = a + \theta^2$ by Vieta formulae. Then

$$-4a^3 - 27b^2 = (\theta - \theta_1)^2(\theta - \theta_2)^2(\theta_1 - \theta_2)^2 = (3\theta^2 + a)^2\Delta$$

by various substitutions, where Δ is the discriminant of the quadratic $x^2 + vx + w$. $(3\theta^2 + a)^2$ is a square in our cubic field and so Δ is a square if and only if the discriminant of $x^3 + ax + b$ is a square. But a root of $(x^2 + vx + w)$ exists in our cubic field if and only if Δ is a square. We assumed the cubic field is not Galois so the discriminant cannot be a square. We are done. □

To tell whether the cubic field is totally real or has a pair of complex conjugate embeddings, just look at the sign of the discriminant. Now we can tell what kind of cubic field we have just by staring at the corresponding cubic polynomial's discriminant.

Example 2.4.4. *Let $K = \mathbb{Q}(x)/(x^3 + 5x + 10)$. This is a cubic extension, as $x^3 + 5x + 10$ is irreducible by Eisenstein's criterion. Its discriminant is $-600 - 2700 = -3300$, so we can say that K is not Galois and has a pair of complex conjugate embeddings. Its Galois closure must be its compositum with the quadratic field $\mathbb{Q}(\sqrt{-33})$. We can work out its ring of integers abstractly.*

We start with the order $\mathbb{Z}[\theta]$ for a root θ of $x^3 + 5x + 10$, whose discriminant is also -3300 by considering the Vandermonde matrix of $x^3 + 5x + 10$. Note that $\frac{-3300}{4} = -825$ is not a valid discriminant by Stickelberger's theorem since it is congruent to 3 modulo 4. Hence we only need to check the prime 5 as a denominator. We need to check if $\frac{a+b\theta+c\theta^2}{5}$ is an algebraic integer for $0 \leq a, b, c < 5$.

The linear map corresponding to multiplication by this element is the matrix

$$M = \begin{bmatrix} \frac{a}{5} & -2c & -2b \\ \frac{b}{5} & \frac{a-5c}{5} & -(b+2c) \\ \frac{c}{5} & \frac{b}{5} & \frac{a-5c}{5} \end{bmatrix}$$

using the integral basis $1, \theta, \theta^2$ and the relation $\theta^3 = -5\theta - 10$. We only need to look at the trace and norm in order to show that it cannot be an algebraic integer. The trace is $\frac{3a-10c}{5}$, and so a must be 0 in order for the trace to be an integer. The norm is the determinant which is

$$-2c(\frac{-bc}{5} + \frac{bc + 2c^2}{5}) - 2b(\frac{b^2}{25} + \frac{c^2}{5}) = \frac{1}{25}(20c^3 - 2b^3 + 10bc^2)$$

In order for this to be an integer, 5 must divide b^3 and so b must be 0. Then c must also be zero as the determinant becomes $\frac{4c^3}{5}$. As a result $\mathbb{Z}[\theta]$ really is the ring of integers.

Note that we did not need to find the ring of integers to factor any integer primes other than $2, 3, 5, 11$. For example we could have factored 7 by seeing that

$$x^3 + 5x + 10 \equiv (x - 2)(x^2 + 2x + 2) \pmod{7}$$

Since $x^2 + 2x + 2$ has no roots modulo 7, we have

$$(7) = (7, \theta - 2)(7, \theta^2 + 2\theta + 2)$$

But now that we know the ring of integers, we can factor any integer prime, in particular those which ramify. For example, 5 totally ramifies as $(5) = (5, \theta)^3$.

We will give a criterion in certain cases that describes the type of ramification that occurs in cubic fields. We either get total ramification, or what we call *partial* ramification.

Proposition 2.4.5. *Let $\mathbb{Q}(x)/(x^3 + ax + b)$ be a cubic field (with $x^3 + ax + b$ irreducible) so that its ring of integers is $\mathbb{Z}[\theta]$ for a root θ of $x^3 + ax + b$. Let $p \geq 5$ be an integer prime dividing $-4a^3 - 27b^2$. Then p totally ramifies if and only if $p|(a, b)$, and ramifies partially otherwise.*

Proof. p totally ramifies if and only if $x^3 + ax + b \equiv (x + c)^3 \pmod{p}$ for some $c \in \mathbb{Z}/(p)$ by the original Dedekind–Kummer method. Differentiation respects reduction modulo p and so $(x^3 + ax + b)' \equiv 3x^2 + a \equiv 3(x + c)^2 \pmod{p}$ in this case. Since 3 is invertible modulo p this occurs if and only if $(x^3 + ax + b)' | x^3 + ax + b \pmod{p}$. We now prove the converse, that $(x^3 + ax + b)' | x^3 + ax + b \pmod{p}$ implies total ramification. Assuming this, we get that the polynomial shares two roots with the derivative. Let $\theta_0, \theta_1, \theta_2$ be the roots of $x^3 + ax + b$. Then by assumption we have

$$(x - \theta_0)(x - \theta_1) + (x - \theta_0)(x - \theta_2) + (x - \theta_1)(x - \theta_2) | (x - \theta_0)(x - \theta_1)(x - \theta_2) \pmod{p}$$

Let's assume W.L.O.G that θ_0, θ_1 are the two roots of $3x^2 + a \bmod p$. Then we get

$$(\theta_0 - \theta_1)(\theta_0 - \theta_2) \equiv (\theta_1 - \theta_0)(\theta_1 - \theta_2) \equiv 0 \pmod{p}$$

Let's assume $\theta_0 \equiv \theta_1 \pmod{p}$ because all other cases imply total ramification. Then $x^3 + ax + b$ becomes $(x - \theta_1)^2(x - \theta_2)$ and its derivative becomes $(x - \theta_1)(x - \frac{\theta_1 + 2\theta_2}{3}) \equiv (x - \theta_1)^2 \pmod{p}$. In particular this means that $\frac{\theta_1 + 2\theta_2}{3} = \theta_1$ and so $\theta_1 = \theta_2$. This also implies total ramification.

We have shown that total ramification occurs if and only if $3x^2 + a$ divides $x^3 + ax + b$ modulo p, which occurs if and only if we can write $x^3 + ax + b \equiv (x^2 + \frac{a}{3})(x + c) \pmod{p}$ for some integer c. But comparing coefficients tells us that $c \equiv 0 \pmod{p}$ and $a \equiv b \equiv 0 \pmod{p}$ as long as $3 \not\equiv 1 \pmod{p}$. This condition is met since $p \neq 2$. Conversely, $a \equiv b \equiv 0 \pmod{p}$ implies that $x^3 + ax + b \equiv x^3 \pmod{p}$ and so we get total ramification. This proves that total ramification of $p \geq 5$ occurs if and only if it divides a and b. □

Using Proposition 2.4.5, we can get unramified extensions of degree 3 for certain quadratic fields. This is important because by class field theory, it tells us that the class group of the quadratic field has 3-torsion. We will state a criterion and give some examples.

Theorem 2.4.6. *Let $K = \mathbb{Q}(x)/(x^3 + ax + b)$ be a cubic field and let $\Delta = -4a^3 - 27b^2$ so that*

- $(a, b) = 1$
- Δ *is square-free*
- Δ *is not divisible by* 2 *or* 3

Let $L = \mathbb{Q}(\sqrt{\Delta})$ be the associated quadratic field. Then $LK : L$ is an unramified extension of degree 3*, and so $Cl(L)$ has* 3*-torsion.*

Proof. Δ is square-free and so K is monogenic, and has ring of integers generated by a root θ of the polynomial $x^3 + ax + b$. We apply Proposition 2.4.5. The prime divisors of Δ are not 2 or 3 nor do they divide $(a, b) = 1$. Hence all integer primes that ramify in $K : \mathbb{Q}$ will ramify partially, and so they cannot totally ramify in $LK : \mathbb{Q}$ either.

Now suppose some prime ideal $\mathfrak{p}$ of $\mathcal{O}_L$ is ramified in $LK : L$. Then it must be totally ramified since it is a Galois extension of prime degree. As a result the integer prime p below $\mathfrak{p}$ has ramification

degree divisible by 3 in $LK : \mathbb{Q}$. However p can only ramify partially in $K : \mathbb{Q}$ with ramification degree 2, so its ramification degree over $LK : \mathbb{Q}$ is either $0, 2$ or 4. This gives a contradiction. Therefore $LK : L$ is unramified. □

Example 2.4.7. *Here are some examples. The two simplest examples are taking $a = 1, b = 1$ and $a = -1, b = 1$. These will give discriminants -31 and -23 respectively, and all conditions of Theorem 2.4.6 are satisfied. As a result, 3 divides the class numbers of $\mathbb{Q}(\sqrt{-31})$ and $\mathbb{Q}(\sqrt{-23})$. In fact, the class numbers of these quadratic fields is 3, so the unramified extensions given by Theorem 2.4.6 are the Hilbert class fields of these quadratic fields.*

Here is a table of further examples:

Cubic Polynomial	Discriminant	Associated quadratic field	Class number of quadratic field
$x^3 + 2x + 1$	-59	$\mathbb{Q}(\sqrt{-59})$	3
$x^3 + 4x + 1$	-283	$\mathbb{Q}(\sqrt{-283})$	3
$x^3 - 4x + 1$	229	$\mathbb{Q}(\sqrt{229})$	3
$x^3 + 1x + 3$	-247	$\mathbb{Q}(\sqrt{-247})$	6
$x^3 - 1x + 3$	-239	$\mathbb{Q}(\sqrt{-239})$	15
$x^3 - 2x + 3$	-211	$\mathbb{Q}(\sqrt{-211})$	3
$x^3 - 5x + 1$	473	$\mathbb{Q}(\sqrt{473})$	3
$x^3 - 5x + 3$	257	$\mathbb{Q}(\sqrt{257})$	3

2.5 Quadratic reciprocity via prime decomposition

In this short subsection we prove quadratic reciprocity. Let $L = \mathbb{Q}(\zeta_p)$ be a prime cyclotomic field. It will have discriminant

$$\Delta(L) = (-1)^{\frac{p-1}{2}} p^{p-2}$$

p divides $\Delta(L)$ with odd multiplicity. Because the cyclotomic field is Galois over $\mathbb{Q}$, it will have a quadratic subfield $K = \mathbb{Q}(\sqrt{\Delta(L)}) = \mathbb{Q}(\sqrt{(-1)^{\frac{p-1}{2}} p})$. The discriminant of K is always $\Delta(K) = (-1)^{\frac{p-1}{2}} p$ since $(-1)^{\frac{p-1}{2}} p \equiv 1 \pmod 4$ by Stickelberger's theorem.

- Let q be an integral prime. Firstly, q ramifies in K if and only if $q = p$. Furthermore, q ramifies in L if and only if $q = p$. As a result $(\frac{p}{q}) = 0 \iff (\frac{q}{p}) = 0$ from our classification of prime decomposition in cyclotomic and quadratic fields. From now on assume $q \neq p$.

- If $p \equiv 1 \pmod 4$ and $(\frac{p}{q}) = 1$, then $\Delta(K) = p$ and so q will split in K. As a result q will split into an even number of primes in $L : \mathbb{Q}$ and so $\text{ord}_p(q) = \frac{p-1}{d}$ for some d even. Then $q^{\frac{p-1}{2}} \equiv 1 \pmod p$ and so $(\frac{q}{p}) = 1$ by Euler's criterion.

- If $p \equiv 3 \pmod 4$ and $(\frac{p}{q}) = 1$, then $\Delta(K) = -p$ and so q will split in K if and only if $q \equiv 1 \pmod 4$ since $(\frac{-p}{q}) = (-1)^{\frac{q-1}{2}}(\frac{p}{q}) = (-1)^{\frac{q-1}{2}}$. When q splits we have $(\frac{q}{p}) = 1$ in a similar fashion to the above case. If q does not split in K then it can only split into an odd number of primes in L since $[L : K] = \frac{p-1}{2}$ is odd. Thus $\text{ord}_p(q) = \frac{p-1}{d}$ for some d odd and so $(\frac{q}{p}) = -1$ by Euler's criterion.

- If $p \equiv 3 \pmod 4$ and $(\frac{p}{q}) = -1$, then $\Delta(K) = -p$ and so q will be inert in K if and only if $q \equiv 1 \pmod 4$ since $(\frac{-p}{q}) = (-1)^{\frac{q-1}{2}}(\frac{p}{q}) = (-1)^{\frac{q+1}{2}}$. If q remains inert in K then it can only split into an odd number of primes in L since $[L : K] = \frac{p-1}{2}$ is odd. As a result $\text{ord}_p(q) = \frac{p-1}{d}$ for some d odd and so $(\frac{p}{q}) = -1$ by Euler's criterion. If q splits in K then $(\frac{p}{q}) = 1$ in a similar fashion to the second case.

- If $p \equiv 1 \pmod 4$ and $(\frac{p}{q}) = -1$, then $\Delta(K) = p$ and so q will be inert in K. This is the hardest case and we will have to consider the Artin symbol of the prime q in the extension $L : \mathbb{Q}$. We claim that it is the automorphism σ sending ζ_p to ζ_p^q. To see this, note that the order of this automorphism is $\text{ord}_p(q) = \frac{p-1}{d}$ where d is the number of primes that q splits into in the extension $L : \mathbb{Q}$. Since we are in a Galois extension, $\frac{p-1}{d}$ is the inertia degree of q in $L : \mathbb{Q}$. Let $\mathfrak{q}$ be a prime above q in this extension. The fact that σ is the Artin symbol is a simple consequence of the Freshman's dream, since $(\sum_{i=0}^{p-1} a_i \zeta_p^i)^q \equiv \sum_{i=0}^{p-1} a_i^q \zeta_p^{qi} \equiv \sigma(\sum_{i=0}^{p-1} a_i \zeta_p^i) \pmod{\mathfrak{q}}$ for any integers a_i.

  The Artin symbol of q in $L : \mathbb{Q}$ will restrict to the Artin symbol of q in $K : \mathbb{Q}$ under the quotient map $\text{Gal}(L : \mathbb{Q}) \to \text{Gal}(K : \mathbb{Q})$. As we have seen, q splitting into an even number of factors in L is equivalent to d being even, which occurs if and only if $(\frac{q}{p}) = 1$ by Euler's criterion. From our classification of prime decomposition in quadratic fields, the fiber of the trivial automorphism in the quotient $\text{Gal}(L : \mathbb{Q}) \to \text{Gal}(K : \mathbb{Q})$ must correspond to the $\frac{p-1}{2}$ quadratic residues modulo p, when realizing $\text{Gal}(L : \mathbb{Q})$ as $(\mathbb{Z}/(p))^\times$. It follows that $(\frac{q}{p}) = 1$ would imply that q splits in K in this case. Hence q being inert would imply that $(\frac{q}{p}) = -1$.

Summarizing all of the above cases, we get quadratic reciprocity for odd positive primes p, q.

$$(\frac{q}{p}) = (-1)^{\frac{(p-1)(q-1)}{4}}(\frac{p}{q})$$

3 Ring of adeles of a number field

The ring of adeles $\mathbb{A}_K$ of a number field K is what we get by looking at the completion of K at each of its places (absolute values) simultaneously. Each prime ideal of $\mathcal{O}_K$ and every embedding of K into $\mathbb{C}$ gives rise to an absolute value, and we will sometimes write the prime ideal or embedding in place of their induced absolute value. A theorem of Ostrowski states that all absolute values are equivalent to these. For a proof of Ostrowski's theorem see [Gim], although we will not be using it in this section.

Loosely speaking, $\mathbb{A}_K$ is a restricted direct product of all completions of K, where we make a restriction so that $\mathbb{A}_K$ can be a locally compact topological ring. We will use the theory developed to then prove Dirichlet's unit theorem and the finiteness of the ideal class group. Throughout this section we follow [Wes].

3.1 Definitions of adeles and ideles

Definition 3.1.1. *Let K be a number field. Denote by $\mathrm{Spec}(\mathcal{O}_K)$ the set of prime ideals of $\mathcal{O}_K$.*

Definition 3.1.2. *Let K be a number field. For each real embedding σ we get the archimedean absolute value*

$$||\cdot||_\sigma := |\sigma(\cdot)|$$

For each complex embedding σ we get the archimedean absolute value

$$||\cdot||_\sigma := |\sigma(\cdot)| \cdot |\bar{\sigma}(\cdot)|$$

Definition 3.1.3. *Let $\mathfrak{p} \in \mathrm{Spec}(\mathcal{O}_K)$ be a nonzero prime ideal. This give rise to the valuation*

$$v_\mathfrak{p}(a) := \max\{n \in \mathbb{Z} : \mathfrak{p}^n | (a), \mathfrak{p}^{n+1} \nmid (a)\}$$

which in turn gives rise to the non-archimedean absolute value

$$||\cdot||_\mathfrak{p} := N(\mathfrak{p})^{v_\mathfrak{p}(\cdot)}$$

Definition 3.1.4. *In this section, we define $\mathcal{O}_{K,\mathfrak{p}}$ to be the $\mathfrak{p}$-adic integers in $\mathcal{O}_K$. This is the projective limit of the diagram*

$$\cdots \to \mathcal{O}_K/\mathfrak{p}^{n+1} \to \mathcal{O}_K/\mathfrak{p}^n \to \cdots$$

Definition 3.1.5. *The completion of K with respect to a place v is denoted by K_v. For archimedean places this will either be $\mathbb{R}$ or $\mathbb{C}$. For a non-archimedean place $v = v_\mathfrak{p}$, we have*

$$K_v = \mathrm{Frac}(\mathcal{O}_{K,\mathfrak{p}}) = \mathcal{O}_{K,\mathfrak{p}}[\mathfrak{p}^{-1}]$$

The absolute values turn these completions into metric topological spaces, whose basic open sets are open balls of the form

$$B(a,r) := \{x \in K_v : ||a-x||_v < r\}$$

When v is archimedean, r can be any positive real number. When $v = v_{\mathfrak{p}}$ is non-archimedean, $r = N(\mathfrak{p})^m$ for some integer m.

Proposition 3.1.6. *Let K be a number field and v be a place of K. Then K_v is a topological field.*

Proof. 1. We show that addition is a continuous map. Define $f : K_v \times K_v \to K_v$ by $f(x,y) = x+y$. It suffices to show that the preimage of an open ball $B(a,r)$ under this map is open in the product topology $K_v \times K_v$.

Select some $(x,y) \in K_v \times K_v$ so that $f(x,y) \in B(a,r)$. Thus $||a-f(x,y)||_v = ||a-x-y||_v < r$. We choose the open set $B(x,\epsilon) \times B(y,\epsilon)$ for small enough ϵ. We have to show that the image of this open set under f is contained in $B(a,r)$. ϵ can be made as small as needed in both the archimedean and non-archimedean cases. Hence we can fix some ϵ so that for any $(x',y') \in B(x,\epsilon) \times B(y,\epsilon)$, we have

$$||a-x'-y'||_v \leq ||a-x-y||_v + ||x-x'||_v + ||y-y'||_v < r$$

In particular, choose $\epsilon < \frac{r-||a-x-y||_v}{2}$. The above shows that the image is contained in $B(a,r)$ and so addition is a continuous map.

2. We show that negation is a continuous map. Define $f : K_v \to K_v$ by $f(x) = -x$. We will show that the preimage of the open ball $B(a,r)$ under this map is open. Choose some element $x \in f^{-1}(B(a,r))$, so that $||a-f(x)||_v = ||a+x||_v < r$. We find an intermediate open ball $B(x,\epsilon)$ and show that its image is contained in $B(a,r)$ for small enough ϵ. Let $x' \in B(x,\epsilon)$. Hence by choosing some $\epsilon < r - ||a+x||_v$ we get

$$||a+x'||_v \leq ||a+x||_v + ||x'-x||_v < r$$

This shows that the image is indeed contained in $B(a,r)$ and so negation is a continuous map.

3. We show that multiplication is a continuous map. Define $f : K_v \times K_v \to K_v$ by $f(x,y) = xy$. We will show that the preimage of the open ball $B(a,r)$ under this map is open. Choose an element $(x,y) \in f^{-1}(B(a,r))$ so that $||a-f(x,y)||_v = ||a-xy||_v < r$. Again we choose the intermediate open set to be $B(x,\epsilon) \times B(y,\epsilon)$ for some small enough ϵ. Let $(x',y') \in B(x,\epsilon) \times B(y,\epsilon)$, in which case $||(x-x')||_v < \epsilon$ and $||(y-y')||_v < \epsilon$. This means we have $\epsilon||x'||_v - \epsilon||x||_v \leq \epsilon||x-x'||_v < \epsilon^2$ and so

$$\epsilon^2 + \epsilon||x||_v > \epsilon||x'||_v > ||yx'-y'x'||_v \wedge \epsilon||y||_v > ||yx-yx'||_v$$

Combining the two inequalities gives

$$\epsilon^2 + \epsilon(||x||_v + ||y||_v) > ||yx' - y'x'||_v + ||yx - yx'||_v \geq ||yx - y'x'||_v$$

Now we need to choose ϵ so that $\epsilon^2 + \epsilon(||x||_v + ||y||_v) < r - ||a - xy||_v$. This is possible by making ϵ small enough. Assuming this, it follows then that

$$||a - x'y'||_v \leq ||a - xy||_v + ||yx - y'x'||_v < ||a - xy||_v + \epsilon^2 + \epsilon(||x||_v + ||y||_v) < r$$

so the image is contained in $B(a, r)$ and hence multiplication is a continuous map.

4. We show that inversion is a continuous map. Define $f : K_v^\times \to K_v^\times$ by $f(x) = x^{-1}$. We will show that the preimage of the open ball $B(a, r)$ under this map is open. Choose an element $x \in f^{-1}(B(a, r))$ so that $||a - f(x)||_v = ||a - x^{-1}||_v < r$. The intermediate set is chosen to be $B(x, \epsilon)$ for ϵ small enough. Let $x' \in B(x, \epsilon)$, in which case $||x - x'||_v < \epsilon$. Then

$$||x'^{-1} - x^{-1}||_v = ||\frac{x}{xx'} - \frac{x'}{xx'}||_v < \frac{\epsilon}{||xx'||_v}$$

$$||x^2||_v - ||x'x||_v \leq ||x^2 - x'x||_v < ||x||_v\epsilon$$

The second inequality implies that $||x'x||_v > ||x^2||_v - ||x||_v\epsilon$. We have $||x||_v > \epsilon$ and so both $||x'x||_v$ and $||x^2||_v - ||x||_v\epsilon$ are positive real numbers. It follows that $\frac{1}{||x'x||_v} < \frac{1}{||x^2||_v - ||x||_v\epsilon}$. Applying this to the first inequality yields

$$||x'^{-1} - x^{-1}||_v < \frac{\epsilon}{||x^2||_v - ||x||_v\epsilon} = ||x||_v^{-1}\frac{\epsilon}{||x||_v - \epsilon}$$

The right hand side can be made as small as needed by choosing a small enough ϵ. For some such ϵ we get

$$||a - x'^{-1}||_v \leq ||a - x^{-1}||_v + ||x'^{-1} - x^{-1}||_v < r$$

so the image is contained in $B(a, r)$ and thus the preimage is open. As a result inversion is a continuous map on the subspace topology $K_v^\times$ and we have completed the proof that K_v is a topological field.

□

Lemma 3.1.7. *For non-archimedean local fields, the open balls $B(a, N(\mathfrak{p})^m)$ are compact sets.*

Proof. We claim that $B(a, N(\mathfrak{p})^m)$ is the coset $a + \mathfrak{p}^{1-m}\mathcal{O}_{K,\mathfrak{p}}$. This can be seen since $x - a \in \mathfrak{p}^{1-m}\mathcal{O}_{K,\mathfrak{p}}$ if and only if $||x - a||_\mathfrak{p} < N(\mathfrak{p})^m$. Since $K_\mathfrak{p}$ is a topological field, it suffices to prove that $\mathfrak{p}^{1-m}\mathcal{O}_{K,\mathfrak{p}}$ is compact. We construct a homeomorphism $\prod_{i=2-m}^{\infty} \mathcal{O}_K/\mathfrak{p}\mathcal{O}_K \simeq \mathfrak{p}^{1-m}\mathcal{O}_{K,\mathfrak{p}}$ as

$$(a_1, a_2, a_3 \dots) \mapsto a_1\pi^{2-m} + a_2\pi^{3-m} + a_3\pi^{4-m} + \dots$$

where π is a fixed uniformizer for $\mathfrak{p}$. Here $\mathcal{O}_K/\mathfrak{p}\mathcal{O}_K$ is given the discrete topology and thus it is compact as it is finite. By Tychonoff's theorem, a product of compact sets such as $\prod_{i=2-m}^{\infty} \mathcal{O}_K/\mathfrak{p}\mathcal{O}_K$ is compact, so it remains to demonstrate that the map above is a homeomorphism.

The open balls of the form $B(a, N(\mathfrak{p})^j)$ are basic in $K_\mathfrak{p}$. Therefore they are also basic in the induced subspace topology of $\mathfrak{p}^{1-m}\mathcal{O}_{K,\mathfrak{p}}$. Hence we only need to show that the preimages of these open balls under the above map is open.

Consider the preimage of an open ball $B(a, N(\mathfrak{p})^j)$ under the above map. If $a = \sum_{i=1}^{\infty} a_i \pi^{1-m+i}$, then the preimage of the open ball is

$$\prod_{i=2-m}^{1-j} \{a_i\} \times \prod_{i=2-j}^{\infty} \mathcal{O}_K/\mathfrak{p}\mathcal{O}_K$$

The basic open sets in $\prod_{i=2-m}^{\infty} \mathcal{O}_K/\mathfrak{p}\mathcal{O}_K$ are of the form $U = \prod_{i=2-m}^{\infty} U_i$ for U_i open in $\mathcal{O}_K/\mathfrak{p}\mathcal{O}_K$ such that $U_i \neq \mathcal{O}_K/\mathfrak{p}\mathcal{O}_K$ for finitely many i. Thus the preimage of the open ball above is open. □

Proposition 3.1.8. *Let K be a number field and v a place of K. Then K_v is a locally compact, Hausdorff topological field.*

Proof. We have shown in Proposition 3.1.6 that K_v is a topological field for any place v of K.

For non-archimedean v, every element x of K_v is contained in some open ball $B(x, r)$ which is compact by Lemma 3.1.7. Therefore K_v is locally compact. To show that K_v is Hausdorff, fix a uniformizer π and take two distinct elements $a, b \in K_v$. Write them out as $a = \sum_{i=n}^{\infty} a_i \pi^i$ and $b = \sum_{i=m}^{\infty} b_i \pi^i$ for some integers m, n and $a_i, b_i \in \mathcal{O}_K/\mathfrak{p}\mathcal{O}_K$. Since they are distinct we must have $a_j \neq b_j$ for some integer j. Then two open sets that separate a and b are $B(a, N(\mathfrak{p})^{-j})$ and $B(b, N(\mathfrak{p})^{-j})$. Therefore K_v is Hausdorff.

For archimedean v, K_v is homeomorphic to $\mathbb{R}$ or $\mathbb{C}$ with the standard metric. We know that $\mathbb{R}$ and $\mathbb{C}$ with the standard metric are locally compact and Hausdorff. It follows that all archimedean local fields are locally compact and Hausdorff. □

Definition 3.1.9. *Let K be a number field. The ring of adeles $\mathbb{A}_K$ is defined as the restricted product*

$$\mathbb{A}_K := \prod_v{}' K_v$$

over all places v of K. The restriction is that any element of $\mathbb{A}_K$ must be a v-adic integer in all but finitely many non-archimedean places v. This makes $\mathbb{A}_K$ into a ring with pointwise addition and multiplication. We define a topology on $\mathbb{A}_K$ by letting the basic open sets take the form

$$U = \prod_{v \in S} U_v \times \prod_{v \notin S} \mathcal{O}_{K,v}$$

for U_v open sets in K_v with the standard topology and S a finite set of places containing all the archimedean places. For any adele $a \in \mathbb{A}_K$ and any place v of K, we denote by a_v the restriction of a to K_v.

Proposition 3.1.10. *Let K be a number field. Then $\mathbb{A}_K$ is a locally compact, Hausdorff topological ring with the above topology.*

Proof. Let $a \in \mathbb{A}_K$ be an adele, and let S be the finite set of places containing all the archimedean places of K, so that a_v is a v-adic integer for any place $v \notin S$. For each $v \notin S$, the set of v-adic integers $\mathcal{O}_{K,v}$ is just the open compact ball $B(0, N(\mathfrak{p}))$. For each $v \in S$, K_v is locally compact by Proposition 3.1.8 and so we can still embed a_v into an open compact ball $B(a_v, r_v)$ for some r_v. As a result a is contained in the open set

$$\prod_{v \in S} B(a_v, r_v) \prod_{v \notin S} B(0, N(\mathfrak{p}))$$

which is a product of compact sets. This is compact by Tychonoff's theorem and so the ring of adeles is locally compact, since we can embed any adele a into a compact open set like above.

It is an easy result that the product of Hausdorff spaces is Hausdorff. Loosely speaking, we can separate any two adeles by separating them in a place where they differ. Therefore the the ring of adeles is also Hausdorff.

Addition, multiplication and negation is continuous in each component of the above restricted product by Proposition 3.1.6. Therefore these operations are continuous on the ring of adeles. It follows that $\mathbb{A}_K$ is a topological ring. □

Note that inversion is not a continuous map on the adeles, because the inverse of an adele may not necessarily be an adele. This is because the inverse of a v-adic integer may not be a v-adic integer for non-archimedean v, and so the inverse of an adele may not be a v-adic integer in all but finitely many places. We will now look at the units adeles; those adeles whose inverse is an adele as well. These from a group under multiplication called the group of ideles.

Definition 3.1.11. *Let K be a number field. The group of ideles $\mathbb{J}_K$ is defined as the units of $\mathbb{A}_K$. An idele is necessarily an adele which is a v-adic unit for all but finitely many non-archimedean places v. We don't want the ideles to inherit the topology of the adeles, as inversion would not be continuous. Instead we give them a new topology, this time the basic open sets take the form*

$$U = \prod_{v \in S} U_v \times \prod_{v \notin S} \mathcal{O}_{K,v}^{\times}$$

for U_v open sets in $K_v^{\times}$ with the standard subspace topology and S is a finite set of places containing all the archimedean places.

Proposition 3.1.12. *Let K be a number field. Then $\mathbb{J}_K$ a locally compact, Hausdorff topological group with the above topology.*

Proof. Let $i \in \mathbb{J}_K$ be an idele, and let S be the finite set of places containing all the archimedean places of K, so that i_v is a v-adic unit for any place $v \notin S$.

Let $\mathfrak{p} = v \notin S$ be a non-archimedean place. The set of v-adic units $\mathcal{O}_{K,v}^\times$ is simply the closed subset $B(0, N(\mathfrak{p})) - B(0,1)$ of the compact set $B(0, N(\mathfrak{p}))$. As a result $\mathcal{O}_{K,v}^\times$ is also compact. Moreover, $B(0, N(\mathfrak{p})) - B(0,1)$ is open because it has an open cover $\{p + \mathfrak{p}\mathcal{O}_{K,\mathfrak{p}} : p \in (\mathcal{O}_K/\mathfrak{p})^\times\}$.

Let $v \in S$. Then K_v is locally compact, so i can be embedded into an open compact ball $B(i_v, r_v)$ for some r_v. Then i is contained in the open set

$$\prod_{v \in S} B(i_v, r_v) \prod_{v \notin S} \mathcal{O}_{K,v}^\times$$

This is a product of compact open sets, which is compact by Tychonoff's theorem. Therefore the group of ideles is locally compact, as every idele is contained in an open compact set like above.

The product of Hausdorff spaces is Hausdorff, and so the group of ideles is Hausdorff.

It remains to show that multiplication and inversion is continuous. We have shown that multiplication and inversion are continuous on $K_v^\times$ for every place v. It is therefore sufficient to show that $\mathbb{J}_K$ is closed under multiplication and inversion, but this is immediate from the definition. Therefore $\mathbb{J}_K$ is a topological group. □

We introduce the notion of S-ideles and S-units for any finite set of places S containing all the archimedean places. This will be some subgroup of the group of ideles.

Definition 3.1.13. *Let K be a number field. Let S be a finite set of places of K containing all the archimedean places. We denote by $\mathbb{J}_S$ the S-ideles. This is the subgroup*

$$\mathbb{J}_S := \prod_{v \in S} K_v^\times \times \prod_{v \notin S} \mathcal{O}_{K,v}^\times$$

of $\mathbb{J}_K$. The topology of $\mathbb{J}_S$ is the induced subspace topology, making $\mathbb{J}_S$ into a locally compact, Hausdorff topological group.

Definition 3.1.14. *Let K be a number field. Let S be a finite set of places of K containing all the archimedean places. We define the S-units of K as*

$$K_S := K^\times \cap \mathbb{J}_S$$

where $K^\times$ is embedded diagonally into $\mathbb{A}_K$.

From now on denote by $V(K)$ the set of places of K and by $V_\infty(K)$ the set of archimedean places of a K. A special case of the above definition is the global units

$$K_{V_\infty(K)} = K^\times \cap \mathbb{J}_{V_\infty(K)} = \mathcal{O}_K^\times$$

We define two important maps on the group of ideles. This will allow us to define a ton of other groups, which will be of use later.

Definition 3.1.15. *Let K be a number field. The idele norm map is the map*

$$||.|| : \mathbb{J}_K \to \mathbb{R}^+$$

$$||a|| = \prod_{v \in V(K)} ||a_v||_v$$

In particular, the idele norm map is bounded. To see this, let a be an idele. Then $||a_v||_v = 1$ for all but finitely many non-archimedean places v. This is because a must be a v-adic unit in all but finitely many non-archimedean places v.

Proposition 3.1.16. *Let K be a number field. Then the idele norm map is continuous.*

Proof. Take an open interval $(a, b) \subset \mathbb{R}^+$ and select an element $x \in \mathbb{J}_K$ in the preimage, so that $||x|| \in (a, b)$. We must find an intermediate open set U. We select the open set

$$U = \prod_{v \in S} B(x, \epsilon) \times \prod_{v \notin S} \mathcal{O}_{K,v}^{\times}$$

where S is the set of all non-archimedean places v where x is not a v-adic unit, plus all the archimedean places. Select an element $y \in U$. Then there exist some polynomial functions f, g whose constant coefficient is 0 so that $||y||$ is bounded as

$$||x|| - f(\epsilon) = \prod_{v \in S} (||x_v||_v - \epsilon) < ||y|| = \prod_{v \in S} ||y_v||_v < \prod_{v \in S} (\epsilon + ||x_v||_v) = ||x|| + g(\epsilon)$$

where $f(\epsilon), g(\epsilon) \to 0$ as $\epsilon \to 0$. Therefore for small enough ϵ we have $||y|| \in (a, b)$. This implies that the image of U lies in (a, b) for small enough ϵ and so $|| \cdot ||$ must be continuous. □

Definition 3.1.17. *We denote the second map by $(\cdot)$. This realizes every idele as a fractional ideal and is defined by*

$$(\cdot) : \mathbb{J}_K \to \mathcal{I}_K$$

$$(a) = \prod_{\mathfrak{p} \in V(K) \setminus V_\infty(K)} \mathfrak{p}^{v_{\mathfrak{p}}(a)}$$

Definition 3.1.18. *Let K be a number field, and let S be a finite set of places containing all archimedean places. We define the groups*

$$\mathbb{J}_K^0 := \ker(|| \cdot ||) = \{a \in \mathbb{J}_K : ||a|| = 1\}$$

$$\mathbb{J}_S^0 := \mathbb{J}_K^0 \cap \mathbb{J}_S$$

Keep in mind that $K^\times \in \mathbb{J}_K^0$ and $K_S \in \mathbb{J}_S^0$ due to the product formula. We finally arrive at the definitions of the idele class groups.

Definition 3.1.19. *Let K be a number field, and let S be a finite set of places containing all archimedean places. We define the idele class group as*

$$\mathcal{C}_K := \mathbb{J}_K / K^\times$$

We define the reduced idele class group as

$$\mathcal{C}_K^0 := \mathbb{J}_K^0 / K^\times$$

We define the S-idele class group as

$$\mathcal{C}_S := \mathbb{J}_S / K_S$$

We define the reduced S-idele class group as

$$\mathcal{C}_S^0 := \mathbb{J}_S^0 / K_S$$

Proposition 3.1.20. *Let K be a number field. We have the isomorphisms*

$$C_K / C_{V_\infty(K)} \cong Cl(\mathcal{O}_k)$$

$$C_K^0 / C_{V_\infty(K)}^0 \cong Cl(\mathcal{O}_k)$$

Proof. There is a map $C_K \to Cl(\mathcal{O}_k)$ induced by $(\cdot)$. This is well-defined as $(\cdot)$ sends $K^\times$ to principal fractional ideals. To show that it is surjective, take a fractional ideal $\mathfrak{i}$ of K. This can be factored into prime ideals as

$$\mathfrak{i} = \prod_{i=1}^{n} \mathfrak{p}_i^{e_i}$$

for integers e_i. Let π_i be a uniformizer for $\mathcal{O}_{K,\mathfrak{p}_i}$ for each $i = 1 \dots n$. Then the idele

$$\prod_{v \notin \{\mathfrak{p}_i : i=1\dots n\}} \{1\}_v \prod_{i=1}^{n} \{\pi_i^{e_i}\}_{\mathfrak{p}_i}$$

will map to the class of $\mathfrak{i}$ in $Cl(\mathcal{O}_k)$ under $(\cdot)$. Now we have to find the kernel of this map. It is induced by those ideles whose image under $(\cdot)$ is a principal fractional ideal. Let i be such an idele. Then there is some element $k \in K^\times$ so that $(i) = (k)$. This idele is equivalent to $k^{-1}i$ in C_K because $C_K = \mathbb{J}_K / K^\times$. Now $k^{-1}i$ is a v-adic unit for all non-archimedean v and so $k^{-1}i \in \mathbb{J}_{V_\infty(K)}$. It follows that the class of i in C_K lies in $C_{V_\infty(K)}$ and so the kernel lies in $C_{V_\infty(K)}$. Conversely, $C_{V_\infty(K)}$ lies in the kernel as its image consists of principal fractional ideals. Therefore $C_{V_\infty(K)}$ is the kernel and we can conclude that

$$C_K / C_{V_\infty(K)} \cong Cl(\mathcal{O}_k)$$

There is also an induced map $C_K^0 \to Cl(\mathcal{O}_k)$, also induced by $(\cdot)$. To show that is is surjective, note that the archimedean places do not affect the output of this map. Therefore we can start with

some class $c \in C_K$ where $||c|| = l$. Now there exists some $j \in \mathbb{J}_{V_\infty(K)}$ so that $||j|| = l^{-1}$. j does not affect the image as it is a unit in all non-archimedean places. Therefore $jc \equiv c$ in C_k, but $||jc|| = 1$. Hence $jc \in C_K^0$, which is sent to the same class as c. We've already shown that the induced map on C_K is surjective, and so the induced map on C_K^0 is surjective too.

Now we find the kernel, which is induced by those ideles whose image under $(\cdot)$ is a principal fractional ideal. The same strategy as above works here, and so the kernel is equivalent to $\mathbb{J}^0_{V_\infty(K)}$ modulo $K^\times$. This means that the kernel is contained in $C^0_{V_\infty(K)}$. However $C^0_{V_\infty(K)}$ is also contained in the kernel, and so

$$C_K^0 / C^0_{V_\infty(K)} \cong Cl(\mathcal{O}_k)$$

□

3.2 Compactness of the reduced idele class group

The singleton $\{1\}$ is closed in $\mathbb{R}$ and therefore its preimage $\mathbb{J}_K^0$ under the idele norm map is closed in $\mathbb{J}_K$. This is a topological group if we give it the subspace topology. Now $K^\times$ is a subgroup of $\mathbb{J}_K^0$ and so we can define a quotient topology on the topological group $\mathbb{J}_K^0 / K^\times = C_K^0$. In this subsection we prove that C_K^0 is a compact topological group with this topology.

The more classical proof that the ideal class group of a number field is finite involves Minkowski's bound. See [Jan96] for this approach. Our progression through this subsection mimics the classical picture involving Minkowski's bound. Proposition 3.2.4 is essentially the idelic counterpart to Minkowski's bound. Given an idele a, we wish to find all elements of K which are less than or equal to a in all places.

Definition 3.2.1. *Let K be a number field and let $a \in \mathbb{J}_K$ be an idele. We define*

$$L(a) := \{x \in K : ||x||_v \leq ||a||_v \, \forall v \in V(K)\}$$

Then we define $\lambda(a) := |L(a)|$.

Proposition 3.2.4 gives a lower bound on $\lambda(a)$ based on the idele norm $||a||$. This is much worse than what one would get from Minkowski's bound, but it will suffice in our case. We will use the following two results from [Wes].

Theorem 3.2.2 (Product formula)**.** *Let K be a number field and let $a \in K^\times$. Let $V(K)$ be the set of places of K. Then*

$$\prod_{v \in V(K)} ||a||_v = 1$$

Proof. See [Wes, Theorem 4.3]. □

Theorem 3.2.3 (Weak approximation theorem)**.** *Let K be a number field and $\{v_1 \dots v_n\}$ a finite set of places of K. Let $\{a_1 \dots a_n\}$ be a set of elements of K. Then for any real $\epsilon > 0$, there exists some $a \in K$ so that*

$$||a - a_i||_{v_i} < \epsilon$$

for all $i = 1 \dots n$.

Proof. See [Wes, Theorem 4.8]. □

Proposition 3.2.4. *Let $K : \mathbb{Q}$ be a field extension of degree n. There is a real positive constant c such that $\lambda(a) \geq c||a||$ for any idele $a \in \mathbb{J}_K$.*

Proof. Choose a basis $\{\omega_1 \dots \omega_n\}$ for K over $\mathbb{Q}$ lying in $\mathcal{O}_K$. Let

$$c_0 := n \cdot \sup\{||\omega_i||_v : v \in V_\infty(K), i = 1 \dots n\}$$

We will see later on that we wish to minimize c_0 to get the best lower bound, and so we desire an integral basis that is as small as possible with respect to all the archimedean places.

Now let a be an idele. We wish to find some $b \in K^\times$ so that

$$c_0 \leq ||ba_v||_v \leq 2c_0$$

for all archimedean places v. We will do so using the weak approximation theorem. Applying it to $\frac{3}{4}a$ gives that for any real $\epsilon > 0$ there is some element $z \in K^\times$ so that

$$||\frac{3}{4}a_v - z||_v < \epsilon$$

for any place v. We denote $b = \frac{\lceil c_0 \rceil}{z}$. In this case we get that

$$c_0 \cdot ||\frac{1}{b}||_v - \frac{3}{4}||a_v||_v \leq \lceil c_0 \rceil \cdot ||\frac{1}{b}||_v - \frac{3}{4}||a_v||_v = ||\frac{\lceil c_0 \rceil}{b}||_v - ||\frac{3}{4}a_v||_v \leq ||\frac{3}{4}a_v - z||_v < \epsilon$$

for all archimedean v since $\lceil c_0 \rceil$ is a positive integer unaffected by Galois actions. Now we can choose ϵ to be any positive real and we will get some corresponding value of b satisfying the above. By choosing $\epsilon = \frac{1}{4}||a_v||_v$ we get

$$c_0 \cdot ||\frac{1}{b}||_v \leq ||a_v||_v$$

Now by comparing the initial inequality in a different way, we get

$$\frac{3}{4}||a_v||_v - \lceil c_0 \rceil \cdot ||\frac{1}{b}||_v = ||\frac{3}{4}a_v||_v - ||\frac{\lceil c_0 \rceil}{b}||_v \leq ||\frac{3}{4}a_v - z||_v < \epsilon$$

$$\frac{3}{4}||a_v||_v - \epsilon = \frac{1}{2}||a_v||_v \leq \lceil c_0 \rceil \cdot ||\frac{1}{b}||_v$$

Together this gives us

$$\frac{1}{2}||a_v||_v \leq c_0 \cdot ||\frac{1}{b}||_v \leq ||a_v||_v$$

Since all these terms are positive and nonzero, we can invert to get:

$$2\frac{1}{||a_v||_v} \geq \frac{1}{c_0}||b||_v \geq \frac{1}{||a_v||_v}$$

Multiplying throughout by $c_0||a_v||_v$ then gives our desired result. Now there exists a positive integer m so that $||mba_v||_v \leq 1$ for any non-archimedean v, since ba is an idele and we can "cancel denominators". For the archimedean places v we get

$$mc_0 \leq ||mba_v||_v \leq 2mc_0$$

Since $mb \in K^\times$, the product formula yields $||mba|| = ||a||$, and $mba \equiv a$ in the idele class group. Moreover, (mba) is an ideal of $\mathcal{O}_K$. Let's denote it as $\mathfrak{i} = (mba)$. Additionally, $\lambda(mba) = \lambda(a)$ since there is a bijection between $L(mba)$ and $L(a)$ given by multiplication by $mb \in K^\times$. Hence it suffices to solve the problem for the idele mba, which has been well calibrated.

We start with the set

$$\Lambda := \{\sum_{i=1}^{n} f_i\omega_i : 0 \leq f_i \leq m, f_i \in \mathbb{Z}\}$$

We clearly have $|\Lambda| = (m+1)^n$ due to the linear independence of the integral basis over $\mathbb{Z}$. Define $N := N(\mathfrak{i}) = |\mathcal{O}_K/\mathfrak{i}|$, then by the pigeonhole principle, there is a subset $S \subset \Lambda$ of at least $\frac{m^n}{N}$ elements in Λ that are in the same class in the quotient ring $\mathcal{O}_K/\mathfrak{i}$. Any two different elements of S gives rise to an element of $L(mba)$. To see this, let $x, y \in \Lambda$ so that $x \neq y$ but $x \equiv y \pmod{\mathfrak{i}}$. Then

$$||x-y||_v \leq ||mba_v||_v$$

for all non-archimedean v since $x - y \in \mathfrak{i}$. Also

$$||x-y||_v = ||\sum_{i=1}^{n}(f_{x,i} - f_{y,i})\omega_i||_v \leq \sum_{i=1}^{n} m \cdot ||\omega_i||_v \leq mc_0 \leq ||mba_v||_v$$

for all archimedean v. This tells us that for each element in S, we can get a corresponding element of $L(mba)$, by taking away some fixed $s_0 \in S$. Thus we get the lower bound

$$\lambda(mba) \geq \frac{m^n}{N}$$

Note that the product of the non-archimedean norms of mba gives the inverse of the norm of the ideal $\mathfrak{i}$. Hence $||mba|| = N^{-1} \cdot \prod_{v \in V_\infty(K)} ||mba_v||_v$. But for all archimedean places we have $2mc_0 \geq ||mba_v||_v$ and so

$$\prod_{v \in V_\infty(K)} ||mba_v||_v \leq (2mc_0)^{|V_\infty(K)|} \leq (2mc_0)^n$$

$$||mba|| \leq N^{-1}(2mc_0)^n \implies m^n \geq \frac{N||mba||}{(2c_0)^n}$$

By using this inequality on our lower bound, we get that

$$\lambda(mba) \geq \frac{m^n}{N} \geq (2c_o)^{-n}||mba||$$

so our constant is $c = (2c_0)^{-n}$. This completes the lemma. □

Lemma 3.2.5. *Let K be a number field and let c be a positive real number so that $\lambda(a) \geq c||a||$ for every $a \in \mathbb{J}_K$. Let a be an idele of K with $||a|| \geq 2/c$. Then there is some $b \in K^\times$ so that*

$$1 \leq ||ba_v||_v \leq ||a||$$

for all places v of K.

Proof. By Proposition 3.2.4, together with our assumption, we get

$$\lambda(a) \geq c||a|| \geq 2$$

There must be a non-zero element $k \in L(a)$ such that $||k||_v \leq ||a_v||_v$ for all places v. We let $b = k^{-1}$ in which case $1 \leq ||ba_v||_v$ for all places v. Now $||b|| = 1$ by the product formula and so $||ba|| = ||a||$. Therefore for any particular place v' we get

$$\prod_{v \neq v'} ||ba_v||_v \geq 1 \implies ||ba_{v'}||_{v'} = \frac{||ba||}{\prod_{v \neq v'} ||ba_v||_v} \leq \frac{||ba||}{1} = ||a||$$

so altogether $1 \leq ||ba_v||_v \leq ||a||$ for any place v. Then b satisfies the conditions of this lemma and so we are done. □

Theorem 3.2.6. *Let K be a number field. The group $\mathcal{C}_K^0$ is compact.*

Proof. Because of the product formula, there is a well-defined map $\mathbb{J}_K/K^\times \to \mathbb{R}^+$ induced by $||.||$. The kernel of this map is by definition $\mathbb{J}_K^0 \cap (\mathbb{J}_K/K^\times) = \mathbb{J}_K^0/K^\times = C_K^0$. We know that this map is surjective and so we can find for each real positive number $p \in \mathbb{R}^+$ an idele a so that $||a|| = p$. We know that the fiber of p is the coset aC_K^0. Since $\mathbb{J}_K/K^\times$ is a topological group, $aC_K^0 \simeq C_K^0$ topologically. Therefore to prove that C_K^0 is compact it is sufficient to show that the fiber of some positive real number p under the idele norm is compact.

By Proposition 3.2.4 there is some positive real number c such that $\lambda(a) \geq c||a||$ for every idele a. We pick some real number $p > 2/c$ and select an idele a in the fiber of p under the idele norm. By the Lemma 3.2.5 there is some $b \in K^\times$ so that

$$1 \leq ||ba_v||_v \leq p = ||a||$$

for all places v of K. The number of prime ideals in $\mathcal{O}_K$ with absolute norm less than or equal to p is finite. Define the finite set of places S as

$$S := \{v_\mathfrak{p} \in V(K) \backslash V_\infty(K) : N(\mathfrak{p}) \leq p\} \cup V_\infty(K)$$

Define now the subset $T \subset \mathbb{J}_K$ as

$$T := \prod_{v \in S} (\overline{B(0,p)} - B(0,1)) \times \prod_{v \notin S} \mathcal{O}_{K,v}^{\times}$$

We know that the sets $\mathcal{O}_{K,v}^{\times}$ are compact. $(\overline{B(0,p)} - B(0,1))$ is a closed subset of a compact set and therefore also compact. Therefore T is compact by Tychonoff's theorem.

It is easy to see that the idele ba lies in T. Note that the set T is independent of the idele a, and relies only on the value of p. Therefore any idele in aC_K^0 can be multiplied by some element of $K^{\times}$ in order to get it in T. As a result T maps onto aC_K^0 under the quotient map $\mathbb{J}_K \to C_K$.

aC_K^0 is the preimage of a closed set (the singleton p) and is therefore closed. The image of T in C_K is also compact and contains aC_K^0. As a result aC_K^0 is a closed subset of a compact set, hence compact. This shows that C_K^0 is compact. □

3.3 Applications to finiteness of ideal class group and Dirichlet's unit theorem

Theorem 3.3.1. *Let K be a number field. Then $Cl(\mathcal{O}_k)$ is finite.*

Proof. Recall that

$$C_K^0 / C_{V_\infty(K)}^0 \cong \mathcal{I}_K / \mathcal{P}_K \cong Cl(\mathcal{O}_k)$$

A quotient of a compact set is compact, and since C_K^0 is compact, we have that $C_K^0 / C_{V_\infty(K)}^0$ is compact. Now $C_{V_\infty(K)}^0$ is an open subgroup of C_K^0 and so the quotient $C_K^0 / C_{V_\infty(K)}^0$ must also be discrete. A compact discrete set must be finite, and so the class group $Cl(\mathcal{O}_k)$ is finite. □

As for Dirichlet's unit theorem, we will prove a more general theorem called the S-unit theorem, of which Dirichlet's unit theorem is a special case. It states that for a finite set of places S containing all the archimedean places, the $\mathbb{Z}$-rank of the S-units is $|S| - 1$. We will do so by constructing a homomorphism from the group of norm 1 S-ideles to a hyperplane of codimension 1 in $\mathbb{R}^{|S|}$. The fact that the subgroup of S-units spans this hyperplane will follow from the compactness of the reduced idele class group. First we need a lemma on discrete subgroups of real vector spaces.

Lemma 3.3.2. *Let Λ be a discrete subgroup of $\mathbb{R}^n$. Then Λ is free abelian with $\mathbb{Z}$-rank $\dim_{\mathbb{R}}(\mathbb{R}\Lambda)$.*

Proof. We prove this by induction. For $n = 1$, let Λ be a discrete subgroup of $\mathbb{R}$. The case $\Lambda = 0$ is easy so assume that Λ has some nonzero element. Let $\lambda_0 \in \Lambda \cap \mathbb{R}^+$ be the positive element of least absolute value, which exists since Λ is discrete. We claim $\Lambda = \lambda_0 \mathbb{Z}$. Let $\omega \in \Lambda$. Then by the division algorithm there are integers n, r so that $\omega = n\lambda_0 + r$ with $0 \leq r < \lambda_0$ and $r \in \Lambda$ since Λ is an additive group. In the case that r is non-zero we get $r < \lambda_0$ which contradicts the fact that λ_0 is the least positive element of Λ. We must therefore have $r = 0$ and so actually $\omega \in \lambda_0 \mathbb{Z}$. We conclude that Λ has $\mathbb{Z}$-rank 1.

Let $\Lambda \subset \mathbb{R}^n$ be a discrete subgroup. Let $\dim_{\mathbb{R}}(\mathbb{R}\Lambda) = m$ and choose a basis $\{\lambda_1 \dots \lambda_m\}$ for $\mathbb{R}\Lambda$ contained in Λ. This is possible since Λ spans $\mathbb{R}\Lambda$ so some basis is contained in Λ. By the inductive hypothesis we have that $\Lambda_0 = \bigoplus_{i=1}^{m-1} \lambda_i$ is free abelian of rank $m-1$. Define

$$B := \Lambda \cap \{\sum_{i=1}^{m} a_i\lambda_i : 0 \leq a_m \leq 1 \wedge 0 \leq a_i < 1 \; \forall i = 1 \dots m-1\}$$

where the a_i are real numbers. This is a bounded subset of a discrete set, and thus it is finite. We select an element $b = \sum_{i=1}^{m} b_i\lambda_i \in B$ with the minimal nonzero coefficient b_m of λ_m. This can be done as B is non-empty, containing λ_m itself. We carry out a procedure similar to the base case.

Let $\lambda = \sum_{i=1}^{m} a_i\lambda_i \in \Lambda$. By the division algorithm there exists some integer t and element $r = \sum_{i=1}^{m} r_i\lambda_i \in \Lambda$ so that $\lambda = tb + r$ with $0 \leq r_m < b_m$. Furthermore by the division algorithm there is an element $\lambda_0 = \sum_{i=1}^{m-1} c_i\lambda_i \in \Lambda_0$ so that $0 \leq r_i - c_i < 1$ for each $i = 1 \dots m-1$. This implies that $r - \lambda_0 \in B$. Since b has the minimal λ_m coefficient in B, we must have $r_m = 0$ in order to avoid a contradiction. As a result $r \in \Lambda_0$. It follows that $b\mathbb{Z} + \Lambda_0 = \Lambda$. Linear independence of $\{\lambda_1 \dots \lambda_{m-1}, b\}$ over $\mathbb{R}$ then implies that $b\mathbb{Z} \oplus \Lambda_0 = \Lambda$ and so Λ has $\mathbb{Z}$-rank m. The lemma then follows by induction. □

Theorem 3.3.3. *Let $S = \{v_1 \dots v_s\}$ be a finite set of places of K containing all the archimedean places, ordered so that v_s is archimedean. Let K_S be the S-units of K, then K_S has $\mathbb{Z}$-rank $s-1$.*

Proof. We define the injective map

$$Log : \mathbb{J}_S \to \mathbb{R}^s$$

$$a \mapsto (\log ||a_{v_1}||_{v_1} \dots \log ||a_{v_s}||_{v_s})$$

This is a continuous map, as each component is the composition of two continuous functions. Recall that $\mathbb{J}_S^0$ is the set of unit norm S-ideles, and therefore

$$\sum_{v \in S} \log ||a_v||_v = \log(\prod_{v \in S} ||a_v||_v) = \log(\frac{\prod_{v \in V(K)} ||a_v||_v)}{\prod_{v \notin S} ||a_v||_v)} = \log(1) = 0$$

for every $a \in \mathbb{J}_S^0$. Hence the Log-image of $\mathbb{J}_S^0$ lies in the hyperplane

$$H := \{(x_1 \dots x_s) \in \mathbb{R}^s : x_1 + \dots + x_s = 0\}$$

In particular, so do the S-units $K_S \subset \mathbb{J}_S^0$. We claim that the image of K_S is discrete. By Lemma 3.3.2 it will follow that $Log(K_S)$ is a free abelian subgroup of H with $\mathbb{Z}$-rank $\dim_{\mathbb{R}}(\mathbb{R}Log(K_S))$. Since H has dimension $s-1$, it will only remain to show that $Log(K_S)$ spans H.

Let $k \in Log(K_S)$, and choose a bounded open U of $\mathbb{R}^s$ so that $k \in U \subset \mathbb{R}^s$. We need to show that $U \cap Log(K_S)$ is finite, from which discreteness of $Log(K_S)$ will follow. We first prove it in the case $S = V_\infty(K)$. Here $K_S = \mathcal{O}_K^\times$ are the global units, each with a minimal monic polynomial with integer coefficients. These coefficients are determined by the value of the global unit in all

the embeddings of K, by Vieta formulae. The global units in the preimage of $U \cap Log(K_S)$ are bounded in all archimedean places. Therefore the coefficients of their minimal polynomials are also bounded. Since the space of monic polynomials with integer coefficients with some bounded degree is discrete, there must be a finite number of polynomials whose roots could be in the preimage of $U \cap Log(K_S)$. Therefore $U \cap Log(K_S)$ is indeed finite and we are done in this case.

In the general case, we can assume that S has a non-archimedean place. Take a non-archimedean place in S with corresponding prime $\mathfrak{p}$ and look at the image of Log restricted to its component. For any integer m, we have $\log(N(\mathfrak{p})^m) = m \cdot \log(N(\mathfrak{p}))$ so all possible values that the image can take in this component are integer multiples of $\log(N(\mathfrak{p}))$. The restriction of the image to this component and all other non-archimedean components is therefore discrete. The image of the open U is bounded in the restriction to any non-archimedean component and therefore finite. There are a finite number of non-archimedean places in S and so overall there is a finite number of possible values for the restriction of $U \cap Log(K_S)$ to the non-archimedean components of $\mathbb{R}^s$. For each such possible value, find a fixed representative a in K_S which takes that value under the Log map. Let $\overline{a}$ denote the preimage of the elements of $U \cap Log(K_S)$ which have the same value as a in all non-archimedean places. Inverse multiplication by a yields the set $a^{-1}\overline{a}$, which consists of global units because they have absolute value 1 in every non-archimedean place. $Log(a^{-1}\overline{a}) = Log(\overline{a}) - Log(a)$ is bounded because $Log(\overline{a})$ is a subset of the bounded set U. Therefore $a^{-1}\overline{a}$ is finite from the case $S = V_\infty(K)$ above. Hence $\overline{a}$ is finite since multiplication by a^{-1} is bijective. There are a finite number of a to consider and so $U \cap Log(K_S)$ is finite. This finally proves that $Log(K_S)$ is discrete.

Let $W = \mathbb{R}\log(K_S)$, then H/W is a real vector space. There is a continuous surjective map $C_S^0 \to H/W$ induced by Log. Now C_S^0 is compact, because it is a closed subset of the compact set C_K^0. As a result H/W is also compact. But as a real vector space, this only occurs when $H/W = 0$ and so $H = W$. This means that K_S spans H and so indeed the $\mathbb{Z}$-rank of K_S is $s - 1$. □

Dirichlet's unit theorem is a corollary of the above by setting $S = V_\infty(K)$, in which case the S-units become the global units.

Corollary 3.3.4 (Dirichlet)**.** *Let K be a number field. Let r_1 be the number of real embeddings and r_2 the pairs of complex conjugate embeddings of K. Then $\mathcal{O}_K^\times$ has $\mathbb{Z}$-rank equal to $r_1 + r_2 - 1$.*

4 L-series and zeta functions

In this section we develop some of the theory of Dirichlet L-series and Dedekind zeta functions. We will define Dirichlet characters, which are homomorphisms from abelian groups to the multiplicative group of roots of unity. They can be thought of as realizations of the abelian group over $\mathbb{C}$. To each Dirichlet character we can attach a Dirichlet L-series, which is a meromorphic function on $\mathbb{C}$. This function captures some essential data about the original abelian group.

We will also introduce Dedekind zeta functions, which play the same role as the Riemann zeta function, but for arbitrary number fields. We will see that Galois extensions over $\mathbb{Q}$ which have abelian Galois group are special. We will study the Dirichlet characters over their Galois groups and the associated Dirichlet L-series. One interesting result is that the Dedekind zeta function for an abelian number field can be factorized as the product of the Dirichlet L-series associated to their Galois group. Along the way we will also prove Dirichlet's theorem on primes in arithmetic progression, as an application of the theory of Dirichlet characters and L-series.

The most important fact we need, however, is the analytic class number formula. This is an explicit formula for the residue of the Dedekind zeta function at 1. It consists of important arithmetical invariants of the associated number field, including the class number. This will be used throughout future sections.

4.1 Definitions and first properties

In Sections 4.1 and 4.2 we follow [IR90, Chapter 16].

Definition 4.1.1. *Let K be a number field. Then its associated Dedekind zeta function is*

$$\zeta_K(s) = \prod_{\mathfrak{p} \in \mathrm{Spec}(\mathcal{O}_K) \setminus (0)} \frac{1}{1 - N(\mathfrak{p})^{-s}}$$

A Maclaurin expansion of the Dedekind zeta function of a number field K gives

$$\zeta_K(s) = \prod_{\mathfrak{p} \in \mathrm{Spec}(\mathcal{O}_K) \setminus (0)} \sum_{e=0}^{\infty} (\frac{1}{N(\mathfrak{p})^s})^e$$

By unique prime factorization of ideals, and multiplicativity of absolute ideal norm, we get

$$\zeta_K(s) = \sum_{\mathfrak{i} \subset \mathcal{O}_K} \frac{1}{N(\mathfrak{i})^s}$$

Note that the Riemann zeta function is a special case of the Dedekind zeta function, where the number field K is the field of rational numbers $\mathbb{Q}$.

Definition 4.1.2. *Let k be a positive integer. A Dirichlet character χ modulo k is a multiplicative group homomorphism*

$$\chi : (\mathbb{Z}/(k))^\times \to \mathbb{C}^\times$$

which also takes the value 0 for any element in $\mathbb{Z}/(k)$ not in $(\mathbb{Z}/(k))^\times$.

Let χ be a Dirichlet character modulo k and let n be a positive integer so that $k|n$. Then χ *induces* a Dirichlet character ψ modulo n, by defining

$$\psi(a) = \chi(\bar{a})$$

for all a coprime to n, where $\bar{\cdot}$ denotes reduction modulo k.

Definition 4.1.3. *A Dirichlet character χ is called primitive if it is not induced by any Dirichlet character other than itself. Note that every Dirichlet character is induced by a unique primitive Dirichlet character. The conductor f_χ of a Dirichlet character χ is the modulus of the unique primitive character that induces it.*

Every Dirichlet character χ modulo k induces a multiplicative map $\mathbb{Z} \to \mathbb{C}$ in the same way as above. You could view this map as a Dirichlet character modulo 0. Then every Dirichlet character gives rise to an associated Dirichlet L-series as follows.

Definition 4.1.4. *Let χ be a Dirichlet character. Then the associated Dirichlet L-series is*

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

Example 4.1.5. *The trivial Dirichlet character χ_0 modulo k takes the value 1 for all integers coprime to k and 0 otherwise. Its conductor is defined to be 1. The primitive trivial Dirichlet character induces a map $\chi_0 : \mathbb{Z} \to \mathbb{C}^\times$ that takes the value 1 everywhere. The associated Dirichlet L-series is the Riemann zeta function.*

Proposition 4.1.6. *Let χ_0 be the trivial Dirichlet character modulo k. Then*

$$L(s, \chi_0) = \prod_{p|k} (1 - p^{-s}) \zeta_{\mathbb{Q}}(s)$$

Proof. We start with the definition of trivial character modulo k. This gives

$$L(s, \chi_0) = \sum_{n=1}^{\infty} \frac{\chi_0(n)}{n^{-s}} = \sum_{n \in \mathbb{Z}^+ : (n,k)=1} \frac{1}{n^{-s}}$$

Then we notice that due to the multiplicativity of the Euler factors, we simply need to remove those factors corresponding to primes dividing k. As a result

$$L(s, \chi_0) = \prod_{p \nmid k} (1 - p^{-s})^{-1} = \prod_{p|k} (1 - p^{-s}) \zeta_{\mathbb{Q}}(s)$$

□

Since Dirichlet characters are multiplicative, there is also an Euler product form for Dirichlet L-series. For a Dirichlet character χ we get:

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

Proposition 4.1.7. *Denote by* $\widehat{(\mathbb{Z}/(k))}^\times$ *the set of Dirichlet characters modulo* k*. Then they form a group, where multiplication is defined by* $(\chi \cdot \psi)(a) = \chi(a)\psi(a)$

Proof. It is easy to see that if χ, ψ are characters modulo k, then $(\chi \cdot \psi)$ is also a character modulo k, since $(\chi \cdot \psi)(ab) = \chi(ab)\psi(ab) = \chi(a)\chi(b)\psi(a)\psi(b) = (\chi \cdot \psi)(a)(\chi \cdot \psi)(b)$. The identity element is then the trivial character χ_0 modulo k. Let χ be a character modulo k. We define χ^{-1} by setting $\chi^{-1}(a) = \chi(a)^{-1}$. Then it is clear that $(\chi \cdot \chi^{-1}) = \chi_0$. This completes the proof. □

The image of a Dirichlet character is always a subgroup of the group of roots of unity. The multiplicative inverse of a root of unity is just its complex conjugate. Therefore, we will interchangeably write the complex conjugate of a character for its inverse.

Proposition 4.1.8. *For any positive integer* k *we have* $\widehat{(\mathbb{Z}/(k))}^\times \cong (\mathbb{Z}/(k))^\times$*.*

Proof. By the fundamental theorem of finitely generated abelian groups, we may write $(\mathbb{Z}/(k))^\times$ as a direct sum of cyclic multiplicative groups as

$$(\mathbb{Z}/(k))^\times \cong \bigoplus_m (\mathbb{Z}/(m))$$

Taking character groups respects direct sum. In other words

$$\widehat{(\mathbb{Z}/(k))}^\times \cong \bigoplus_m \widehat{\mathbb{Z}/(m)}$$

Now for cyclic groups, every character is determined uniquely by its value on the generator. It follows that $\widehat{\mathbb{Z}/(m)} \cong \mathbb{Z}/(m)$ for every m. Putting it all together gives

$$\widehat{(\mathbb{Z}/(k))}^\times \cong \bigoplus_m \widehat{\mathbb{Z}/(m)} \cong \bigoplus_m (\mathbb{Z}/(m)) \cong (\mathbb{Z}/(k))^\times$$

□

Proposition 4.1.9. *We have the following orthogonality relations for Dirichlet characters, where* δ *is the Kronecker delta.*

1. $\sum_{a \in (\mathbb{Z}/(k))^\times} \chi(a)\overline{\psi(a)} = \phi(k)\delta(\chi, \psi)$ *for any* $\chi, \psi \in \widehat{\mathbb{Z}/(k)}$
2. $\sum_{\chi \in \widehat{\mathbb{Z}/(k)}} \chi(a)\overline{\chi(b)} = \phi(k)\delta(a, b)$ *for any* $a, b \in (\mathbb{Z}/(k))^\times$

Proof. In the first case, $\chi = \psi$ implies that

$$\sum_{a \in (\mathbb{Z}/(k))^\times} \chi(a)\overline{\psi(a)} = \sum_{a \in (\mathbb{Z}/(k))^\times} \chi(a)\overline{\chi(a)} = \sum_{a \in (\mathbb{Z}/(k))^\times} 1 = \phi(k)$$

Now assume $\chi \neq \psi$. Multiplying the entirety of $(\mathbb{Z}/(k))^\times$ by any of its elements b will simply permute the group. Hence

$$\sum_{a \in (\mathbb{Z}/(k))^\times} \chi(a)\overline{\psi(a)} = \sum_{a \in (\mathbb{Z}/(k))^\times} \chi(ab)\overline{\psi(ab)} = (\chi \cdot \overline{\psi})(b) \sum_{a \in (\mathbb{Z}/(k))^\times} \chi(a)\overline{\psi(a)}$$

In particular we must have

$$(1 - (\chi \cdot \overline{\psi})(b)) \sum_{a \in (\mathbb{Z}/(k))^\times} \chi(a)\overline{\psi(a)} = 0$$

so either $(1 - (\chi \cdot \overline{\psi})(b)) = 0$ or $\sum_{a \in (\mathbb{Z}/(k))^\times} \chi(a)\overline{\psi(a)} = 0$. $(1 - (\chi \cdot \overline{\psi})(b)) = 0$ cannot occur for all b, unless we have $\chi = \psi$, which is contradictory to our assumption. Therefore

$$\sum_{a \in (\mathbb{Z}/(k))^\times} \chi(a)\overline{\psi(a)} = 0$$

The second case follows from the first by Pontryagin duality. □

4.2 Dirichlet's theorem on arithmetic progressions

A standard analytic result is that $\zeta_{\mathbb{Q}}(s)$ converges for $\{s \in \mathbb{C} : Re(s) > 1\}$. The next proposition tells us something about the value close to $s = 1$. It says that $\zeta_{\mathbb{Q}}(s)$ has a simple pole with residue 1 at $s = 1$.

Proposition 4.2.1. $\lim_{s \to 1^+}((s-1)\zeta_{\mathbb{Q}}(s)) = 1$

Proof. We apply a standard trick often used in integral tests. We have for positive integers n that

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

because t^{-s} as a function of t is monotonically decreasing for fixed $s > 0$. We sum this inequality from $n = 1$ to infinity, giving us

$$\zeta_{\mathbb{Q}}(s) - 1 < \int_1^{\infty} t^{-s} dt < \zeta_{\mathbb{Q}}(s)$$

This integral is evaluated as $\int_1^\infty t^{-s}dt = \frac{t^{1-s}}{1-s}\Big|_1^\infty = \frac{1}{s-1}$ for any real $s > 1$. By inequality manipulation we get that

$$1 < (s-1)\zeta_{\mathbb{Q}}(s) < s$$

for any $s > 1$, so the one sided limit $\lim_{s \to 1^+}((s-1)\zeta_{\mathbb{Q}}(s))$ is 1. □

We are going to give a proof that there are infinitely many primes using the Riemann zeta function. The method is important because we will generalize it in order to prove Dirichlet's theorem on primes in arithmetic progressions.

Lemma 4.2.2. $\log(\zeta_{\mathbb{Q}}(s)) < \sum_p p^{-s} + 2\zeta_{\mathbb{Q}}(2)$ *for real* $s > 1$.

Proof. We apply the natural logarithm to the Euler product form of the zeta function to get

$$\log(\zeta_{\mathbb{Q}}(s)) = \log(\prod_p \frac{1}{1-p^{-s}}) = \sum_p -\log(1-p^{-s})$$

We will now apply the Taylor expansion

$$-\log(1-x) = \sum_{n=1}^{\infty} \frac{x^n}{n}$$

to each $x = p^{-s}$. This will yield

$$\log(\zeta_{\mathbb{Q}}(s)) = \sum_p \sum_{n=1}^{\infty} \frac{p^{-sn}}{n} = \sum_p p^{-s} + \sum_p \sum_{n=2}^{\infty} \frac{p^{-sn}}{n}$$

Finally we bound the second sum on the RHS to get

$$\sum_p \sum_{n=2}^{\infty} \frac{p^{-sn}}{n} < \sum_p \sum_{n=2}^{\infty} p^{-sn} = \sum_p p^{-2s}(1-p^{-s})^{-1} < (1-2^{-s})^{-1} \sum_p p^{-2s} < 2\zeta_{\mathbb{Q}}(2)$$

Altogether this gives us the desired result. □

Corollary 4.2.3. *From Proposition 4.2.1 we have* $\lim_{s\to 1^+}(\log(s-1) + \log(\zeta_{\mathbb{Q}}(s))) = 0$ *since the logarithm is continuous on* $\{x \in \mathbb{R} : x > 0\}$. *Then we get* $\lim_{s\to 1^+}(\log(\zeta_{\mathbb{Q}}(s))) = \lim_{s\to 1^+}(\log(\frac{1}{s-1}))$. *Finally we get*

$$\lim_{s\to 1^+}(\frac{\log(\zeta_{\mathbb{Q}}(s))}{\log(\frac{1}{s-1})}) = 1$$

From Lemma 4.2.2 we get $\log(\zeta_{\mathbb{Q}}(s)) = \sum_p p^{-s} + 2\zeta_{\mathbb{Q}}(2) - a$ *for some finite* a *and so*

$$\lim_{s\to 1^+}(\frac{\log(\zeta_{\mathbb{Q}}(s))}{\log(\frac{1}{s-1})}) = \lim_{s\to 1^+}(\frac{\sum_p p^{-s}}{\log((s-1)^{-1})} + \frac{2\zeta_{\mathbb{Q}}(2)-a}{\log((s-1)^{-1})}) = \lim_{s\to 1^+}(\frac{\sum_p p^{-s}}{\log((s-1)^{-1})}) = 1$$

The $\frac{2\zeta_{\mathbb{Q}}(2)-a}{\log((s-1)^{-1})}$ *term disappears since the numerator is finite, but the denominator diverges as* $s \to 1$. *It follows that the sum* $\sum_p p^{-s}$ *must diverge as* $s \to 1$ *because the limit is nonzero, and so there are infinitely many primes.*

Of course we could prove there are infinitely many primes more directly. However, using this train of thought motivates the following definition, which will be used to prove Dirichlet's theorem.

Definition 4.2.4. *Let* S *be a set of positive integer primes. If the limit*

$$d(S) := \lim_{s\to 1^+}(\frac{\sum_{p\in S} p^{-s}}{\log((s-1)^{-1})})$$

exists, then S *is said to have Dirichlet density* $d(S)$.

We've shown that the set of all primes has Dirichlet density 1. When S is finite, the numerator $\sum_{p\in S} p^{-s}$ converges and so $d(S) = 0$ as we've discussed before. As a result, if the Dirichlet density of S is greater than 0, then S must contain infinitely many prime numbers. This is the gadget we will be using to prove the main result, but we will be applying it to Dirichlet L-series in general.

Proposition 4.2.5. *We have the expansion*

$$\log(L(s,\chi)) = \sum_p \sum_{k=1}^{\infty} \frac{\chi(p^k)}{kp^{ks}}$$

for the natural logarithm of $L(s,\chi)$.

Proof. Recall the Taylor expansion for $\log((1-z)^{-1})$. We apply the exponential to both sides to get

$$\frac{1}{1-z} = \exp(\sum_{k=1}^{\infty} \frac{z^k}{k})$$

Now substitute $z = \chi(p)p^{-s}$ and take a product over all integral primes p to get

$$\prod_p \frac{1}{1-\chi(p)p^{-s}} = \prod_p \exp(\sum_{k=1}^{\infty} \frac{\chi(p^k)p^{-ks}}{k}) = \exp(\sum_p \sum_{k=1}^{\infty} \frac{\chi(p^k)p^{-ks}}{k})$$

Taking the natural logarithm of both sides gives the desired result. □

Keep in mind that this result is analogous to Lemma 4.2.2, for Dirichlet L-series in general. Now we proceed as in the base case, by showing that $\log(L(s,\chi))$ is $\sum_p \chi(p)p^{-s}$ up to a finite difference.

Lemma 4.2.6. *Let χ be a Dirichlet character. The natural logarithm of the corresponding Dirichlet L-function for real $s > 1$ can be approximated as*

$$|\log(L(s,\chi)) - \sum_p \chi(p)p^{-s}| < 2\zeta_{\mathbb{Q}}(2)$$

Proof. We begin with the result of Proposition 4.2.5 that says

$$\log(L(s,\chi)) = \sum_p \sum_{k=1}^{\infty} \frac{\chi(p^k)}{kp^{ks}} = \sum_p \chi(p)p^{-s} + \sum_p \sum_{k=2}^{\infty} \frac{\chi(p^k)}{kp^{ks}}$$

Then we apply Lemma 4.2.2 to bound the second sum on the RHS as

$$|\sum_p \sum_{k=2}^{\infty} \frac{\chi(p^k)}{kp^{ks}}| \leq \sum_p \sum_{k=2}^{\infty} \frac{|\chi(p^k)|}{kp^{ks}} \leq \sum_p \sum_{k=2}^{\infty} \frac{1}{kp^{ks}} < 2\zeta_{\mathbb{Q}}(2)$$

where the lemma was applied to get the last inequality. The result then follows. □

Here is where the magic happens. We need to filter through the primes congruent to a modulo k for some coprime integers a, k. We will make use of the orthogonality relations in Proposition 4.1.9. We start with the natural logarithm expansion of $L(s, \chi)$ which is

$$\log(L(s,\chi)) = \sum_p \sum_{k=1}^{\infty} \frac{\chi(p^k)}{kp^{ks}} = \sum_p p^{-s}\chi(p) + R(s)$$

where $R(s)$ is some finite value for $s > 1$ by Lemma 4.2.6. We multiply both sides by $\overline{\chi(a)}$ and sum over all characters χ modulo k, which gives

$$\sum_\chi \overline{\chi(a)} ln(L(s,\chi)) = \sum_p p^{-s} \sum_\chi \overline{\chi(a)}\chi(p) + \sum_\chi \overline{\chi(a)} R(s)$$

Since there are a finite number of characters modulo k, the sum $R'(s) := \sum_\chi \overline{\chi(a)} R(s)$ is still finite. The orthogonality relations will do the filtering for us, giving

$$\sum_\chi \overline{\chi(a)} ln(L(s,\chi)) = \sum_p p^{-s}\phi(k)\delta(\chi(a),\chi(p)) + R'(s) = \phi(k) \sum_{p \equiv a \pmod k} p^{-s} + R'(s)$$

Now we divide both sides by $\log((1-s)^{-1})$ and take the limit as $s \to 1^+$. The finite term $R'(s)$ will disappear. Let $S_{a,k}$ be the set of primes congruent to a modulo k. Then

$$\lim_{s\to 1^+} \frac{\sum_\chi \overline{\chi(a)} \log(L(s,\chi))}{\log((1-s)^{-1})} = \lim_{s\to 1^+} \frac{\phi(k) \sum_{p\equiv a \pmod k} p^{-s}}{\log((1-s)^{-1})} = \phi(k) d(S_{a,k})$$

The Dirichlet density of $S_{a,k}$ exists because $S_{a,k}$ is a subset of the set of all primes, which has Dirichlet density 1. We are now ready to prove Dirichlet's theorem on primes in arithmetic progressions.

Theorem 4.2.7 (Dirichlet)**.** *For any coprime integers a, k, there are infinitely many primes congruent to a modulo k. In fact, the Dirichlet density of the set $S_{a,k}$ is $\frac{1}{\phi(k)}$.*

Proof. This will follow from the fact that

$$\lim_{s\to 1^+} \frac{\sum_\chi \overline{\chi(a)} \log(L(s,\chi))}{\log((1-s)^{-1})} = 1$$

Let χ_0 be the trivial character modulo k. Then from Proposition 4.1.6 we have

$$\overline{\chi_0(a)} \log(L(s,\chi_0)) = \log(\prod_{p|k}(1-p^{-s})\zeta_{\mathbb{Q}}(s)) = \log(\prod_{p|k}(1-p^{-s})) + \log(\zeta_{\mathbb{Q}}(s))$$

$\log(\prod_{p|k}(1-p^{-s}))$ is clearly finite and will disappear in this limit. Therefore we have

$$\lim_{s\to 1^+} \frac{\overline{\chi_0(a)} \log(L(s,\chi_0))}{\log((1-s)^{-1})} = \lim_{s\to 1^+} \frac{\log(\zeta_{\mathbb{Q}}(s))}{\log((1-s)^{-1})} = 1$$

from our base case result. Now let χ be a nontrivial character modulo k and assume $L(1,\chi) \neq 0$, which follows from Corollary 4.4.2. Therefore $\log(L(1,\chi))$ will converge to a finite value. This occurs for all nontrivial characters for some fixed branch of the logarithm. As a result we get that

$$\lim_{s\to 1^+} \frac{\overline{\chi(a)}\log(L(s,\chi))}{\log((1-s)^{-1})} = 0$$

so when summing over all characters modulo k, only the trivial character contributes, and so

$$\lim_{s\to 1^+} \frac{\sum_\chi \overline{\chi(a)}\log(L(s,\chi))}{\log((1-s)^{-1})} = 1$$

Then we can finally compare with previous results to get

$$\lim_{s\to 1^+} \frac{\sum_\chi \overline{\chi(a)}\log(L(s,\chi))}{\log((1-s)^{-1})} = \phi(k)d(S_{a,k}) = 1 \implies d(S_{a,k}) = \frac{1}{\phi(k)}$$

□

4.3 The analytic class number formula

We have shown previously that the Riemann zeta function converges for input with real value greater than 1. The same is true for Dedekind zeta functions, since the number of prime ideals of a given norm n is bounded above by the dimension of the number field over $\mathbb{Q}$. We have shown that the Riemann zeta function has a simple pole at $s = 1$ whose residue is 1, and that this can be used to deduce unique prime factorization in $\mathbb{Z}$. What can we deduce from the residue of the Dedekind zeta function at $s = 1$ in general? This question is answered by the analytic class number formula. In this subsection we follow [UiO].

Theorem 4.3.1. *$\zeta_K(s)$ has a simple pole at $s = 1$ whose residue is*

$$\operatorname{res}(\zeta_K, 1) = \frac{2^{r_1}(2\pi)^{r_2} h_K \mathcal{R}_K}{\omega_K \sqrt{|\Delta(K)|}}$$

where

- *r_1 is the number of real embeddings of K*
- *r_2 is the number of pairs of complex conjugate embeddings of K*
- *h_K is the class number of K*
- *$\mathcal{R}_K$ is the regulator of K*
- *ω_K is the number of roots of unity in K*
- *$\Delta(K)$ is the discriminant of K*

We will speak about the meromorphic continuation of ζ_K later. For now, we shall prove the result above for $\lim_{s\to 1^+}((s-1)\zeta_K) = \text{res}(\zeta_K, 1)$. There are several steps to the proof. Let's start with the Dedekind zeta function

$$\zeta_K(s) = \sum_{\mathfrak{i}\subset\mathcal{O}_K} \frac{1}{N(\mathfrak{i})^s}$$

We want to relate the ideals in the above expression to actual elements of $\mathcal{O}_K$. There are two obstructions for us, given by the exact sequence

$$1 \to \mathcal{O}_K^\times \to \mathcal{O}_K \to \mathcal{I}_K \to Cl(K) \to 1$$

Firstly, the units $\mathcal{O}_K^\times$ give multiple elements that could generate a principal ideal. Secondly, there are multiple ideal classes, each of which could be related to a principal ideal via multiplication by a fixed ideal in the inverse class.

- We will start by resolving the second issue. We can split the Dedekind zeta function by the ideal classes of $\mathcal{O}_K$ as

  $$\zeta_K(s) = \sum_{\mathfrak{i}\subset\mathcal{O}_K} \frac{1}{N(\mathfrak{i})^s} = \sum_{c\in Cl(K)} \sum_{\bar{\mathfrak{i}}=c} \frac{1}{N(\mathfrak{i})^s}$$

  Now for each class c we fix an ideal $\mathfrak{a}_{c^{-1}}$ lying in the inverse class c^{-1}. For every ideal $\mathfrak{i}$ in the class c we have that $\mathfrak{i}\mathfrak{a}_{c^{-1}}$ is principal. In fact, the set of ideals of the form $\mathfrak{i}\mathfrak{a}_{c^{-1}}$ is the same as the set of principal ideals contained in $\mathfrak{a}_{c^{-1}}$. Hence we can write

  $$\zeta_K(s) = \sum_{c\in Cl(K)} N(\mathfrak{a}_{c^{-1}})^s \sum_{\bar{\mathfrak{i}}=c} \frac{1}{N(\mathfrak{i}\mathfrak{a}_{c^{-1}})^s} = \sum_{c\in Cl(K)} N(\mathfrak{a}_{c^{-1}})^s \sum_{(m)\subset\mathfrak{a}_{c^{-1}}} \frac{1}{N((m))^s}$$

  We cannot pass from the principal ideals (m) to the elements m yet, because $(m) = (um)$ for any unit u. In the next part we will sort out the issue of units.

- Recall that K can be embedded diagonally in the product of the archimedean completions $V = \mathbb{R}^{r_1} \oplus \mathbb{C}^{r_2}$, by sending

  $$x \to (\sigma_1(x) \dots \sigma_{r_1}(x), \sigma_{r_1+1}(x) \dots \sigma_{r_1+r_2}(x))$$

  From now on when we mention elements of K, we will actually refer to their embedding into V. In this embedding, addition and multiplication are respected, and so the group of units $\mathcal{O}_K^\times$ acts multiplicatively on the entirety of V. Suppose we have a fundamental domain D for this action in V. This means that D contains precisely one representative for each orbit.

  $\mathcal{O}_K$ embeds as a lattice into V, and so every ideal $\mathfrak{i} \subset \mathcal{O}_K$ also embeds as a lattice into V. Suppose $\mathfrak{i} = (m)$ is principal, and consider the points in $\mathcal{O}_K \cap D$. since D is fundamental, there is only one element in $\mathcal{O}_K \cap D$ which generates (m). To see this, the orbit of the element

m embedded into V is precisely the set of generators of (m), of which only one can lie in D. Therefore we can finally write $\zeta_K(s)$ using elements as

$$\zeta_K(s) = \sum_{c \in Cl(K)} N(\mathfrak{a}_{c^{-1}})^s \sum_{m \in \mathfrak{a}_{c^{-1}} \cap D} \frac{1}{|N(m)|^s}$$

We will now actually find a fundamental domain D for $\mathcal{O}_K^\times \curvearrowright V$. There is a map

$$Log : V_0 := (\mathbb{R}^\times)^{r_1} \oplus (\mathbb{C}^\times)^{r_2} \to \mathbb{R}^{r_1+r_2}$$

$$(x_1 \dots x_{r_1+r_2}) \mapsto (\log|x_1| \dots \log|x_{r_1}|, 2\log|x_{r_1+1}| \dots 2\log|x_{r_1+r_2}|)$$

where the absolute value here measures standard distance, and so the coefficient 2 for the image of the complex places is justified. We have a norm map on V by taking the product of the absolute values of each component. It is defined as

$$N : V \to \mathbb{R}$$

$$(x_1 \dots x_{r_1+r_2}) \mapsto \prod_{i=1}^{r_1} |x_i| \prod_{i=r_1+1}^{r_1+r_2} |x_i|^2$$

Note that for K this will just be the standard norm, so this is notationally sound. We also have a trace map

$$Tr : \mathbb{R}^{r_1+r_2} \to \mathbb{R}$$

$$(x_1 \dots x_{r_1+r_2}) \mapsto \sum_{i=1}^{r_1+r_2} x_i$$

There is a relation $\log(N(x)) = Tr(Log(x))$ which is easily checked.

Recall also Dirichlet's unit theorem, which tells us that there are fundamental units $\{\eta_1 \dots \eta_{r_1+r_2-1}\}$ which generate $\mathcal{O}_K^\times$. The torsion subgroup of $\mathcal{O}_K^\times$ has size ω_K. Let $\epsilon_j = Log(\eta_j)$ for each $j = 1 \dots r_1 + r_2 - 1$. We know that the ϵ_j generate a hyperplane of codimension 1 in $\mathbb{R}^{r_1+r_2}$, defined as

$$H := \{x \in \mathbb{R}^{r_1+r_2} : Tr(x) = 0\}$$

Let $\epsilon_0 = (1, 1 \dots 1, 2 \dots 2)$, whose components are 1 in the image of all real places and 2 in the image of all complex places. Then $Tr(\epsilon_0) = r_1 + 2r_2$ so ϵ_0 does not lie on H. Therefore $\{\epsilon_0 \dots \epsilon_{r_1+r_2-1}\}$ is a basis for $\mathbb{R}^{r_1+r_2}$. We denote $d = [K : \mathbb{Q}] = r_1 + 2r_2$. With this setup we are ready to define the fundamental domain.

Proposition 4.3.2. *A fundamental domain D for the action of $\mathcal{O}_K^\times$ on V defined above can be defined as the subset $D \subset V$ of elements x satisfying*

- $Log(x) = \sum_{i=0}^{r_1+r_2-1} b_i \epsilon_i$ *with* $0 \le b_i < 1$ *for all* $1 \le i \le r_1 + r_2 - 1$
- $x_1 > 0$ *if K has a real place, and* $0 \le \arg(x_1) < \frac{2\pi}{\omega_K}$ *if K is totally complex.*

Proof. Let $y \in V$ and write $Log(y) = \sum_{i=0}^{r_1+r_2-1} a_i\epsilon_i$. Let $t = \lfloor \frac{\omega_K \arg(y_1)}{2\pi} \rfloor$ and define the unit

$$u = e^{i\frac{2\pi t}{\omega_K}} \prod_{i=1}^{r_1+r_2-1} \eta_i^{\lfloor a_1 \rfloor}$$

Then we claim $u^{-1}y$ lies in D. We check the second condition first.

$$\arg(u^{-1}y_1) = \arg(y_1) - \arg(e^{i\frac{2\pi t}{\omega_K}}) = \frac{2\pi}{\omega_K}\frac{\omega_K \arg(y_1)}{2\pi} - \frac{2\pi}{\omega_K}\lfloor \frac{\omega_K \arg(y_1)}{2\pi} \rfloor = \frac{2\pi}{\omega_K}\{\frac{\omega_K \arg(y_1)}{2\pi}\}$$

where $\{\cdot\}$ denotes the fractional part. It follows that $0 \leq \arg(u^{-1}y_1) < \frac{2\pi}{\omega_K}$. If K has a real place, this translates to $u^{-1}y_1$ being positive. Now we check the second condition.

$$Log(u^{-1}y) = Log(y) - Log(u) = a_0\epsilon_0 + \sum_{i=1}^{r_1+r_2-1} (a_i - \lfloor a_i \rfloor)\epsilon_i = a_0\epsilon_0 + \sum_{i=1}^{r_1+r_2-1} \{a_i\}\epsilon_i$$

It remains to show that there is only one representative of each orbit in D. Suppose $x, ux \in D$ where u is a unit of $\mathcal{O}_K$. Then $Log(u) = Log(ux) - Log(x)$. The coefficients of ϵ_i for $i \neq 0$ for both $Log(ux)$ and $Log(x)$ are bounded by $[0, 1)$. As a result the coefficients of ϵ_i for $i \neq 0$ for $Log(u)$ are bounded by $(-1, 1)$, and so they must be 0 because $Log(u)$ is contained in the $\mathbb{Z}$-span of the ϵ_i for $i \neq 0$. As a result u must be a root of unity. However the arguments of both ux and x in their first component are bounded by $[0, \frac{2\pi}{\omega_K})$, and so the first component of u must be 1. As a result $u = 1$. □

The next proposition relates the residue of ζ_K at $s = 1$ to a ratio of certain volumes. Upon computing these volumes, we will get the analytic class number formula.

Definition 4.3.3. *Let $S \subset \mathbb{R}^n$. Then S is called a cone if for any $r \in \mathbb{R}^+$ and $s \in S$, we have $rs \in S$.*

An important thing to note is that the set D is a cone. This means that for any positive real number x, and element $y \in D$, we have $xy \in D$. To see this, note that multiplication by x does not alter the argument of the first component, and so xy satisfies the second condition. Also, $Log(xy) = Log(y) + x\epsilon_0$ so x does not alter the hyperplane components in the image under Log. As a result the first condition is also satisfied, and D is indeed a cone.

Proposition 4.3.4. *Let T be the subset of D of elements with norm having absolute value at most 1. Let λ be the volume of T and let Γ be the covolume of the lattice $\mathfrak{a}_{c^{-1}}$ embedded into V. Then we have*

$$\lim_{s \to 1^+} ((s-1) \sum_{m \in \mathfrak{a}_{c^{-1}} \cap D} \frac{1}{|N(m)|^s}) = \frac{\lambda}{\Gamma}$$

Proof. We wish to approximate the volume of T using the lattice $L = \mathfrak{a}_{c^{-1}}$. We can do this by letting μ be the number of L-points in T, and then writing $\lambda \sim \mu\Gamma$. This is a terrible bound in

general, but the finer the lattice is, the better the bound. Note that $Vol(\frac{1}{r}L) = \frac{1}{r^d}\Gamma$ for any positive real r by scaling, and we get a finer lattice this way. Let $\mu(r)$ be the number of $\frac{1}{r}L$-points in T. Then we can conclude, by letting the lattice become infinitely fine, that

$$\lim_{r\to\infty}(\mu(r)\frac{1}{r^d}\Gamma) = \lambda \implies \frac{\lambda}{\Gamma} = \lim_{r\to\infty}(\mu(r)\frac{1}{r^d})$$

However, scaling down L by a factor of r is the same as scaling up T by a factor of r, when considering the two objects relative to each other. In other words

$$\mu(r) = \#(\frac{1}{r}L \cap T) = \#(L \cap rT)$$

We claim that rT is the set $\{x \in D : |N(x)| < r^d\}$. Since D is a cone, and r is a positive real, the set rT is contained in D. We have $T = \{x \in D : |N(x)| < 1\}$ and so

$$rT = \{rx \in D : |N(x)| < 1\} = \{x \in D : |N(\frac{x}{r})| < 1\} =$$
$$= \{x \in D : \frac{|N(x)|}{r^d} < 1\} = \{x \in D : |N(x)| < r^d\}$$

because the norm map $N(\cdot)$ on V is homogeneous of degree d. Every finite dimensional lattice is countable. In particular, the set $L \cap D$ is countable. Let us order this set according to the function $|N(\cdot)| : V \to \mathbb{R}^+$. Write $L \cap D = \{x_1, x_2 \dots\}$ so that $|N(x_i)| \leq |N(x_j)|$ whenever $i \leq j$. The number k now approximates the number of elements of L with absolute norm less than $|N(x_k)|$. In other words, the elements of $L \cap \sqrt[d]{|N(x_k)|}T$. In particular, we can write $\mu(\sqrt[d]{|N(x_k)|} - \epsilon) < k \leq \mu(\sqrt[d]{|N(x_k)|})$ for any $\epsilon > 0$. As a result we have

$$\frac{\mu(\sqrt[d]{|N(x_k)|} - \epsilon)}{|N(x_k)|}\Gamma < \frac{k}{|N(x_k)|}\Gamma \leq \frac{\mu(\sqrt[d]{|N(x_k)|})}{|N(x_k)|}\Gamma$$

Then taking the limit as k goes to infinity, ϵ becomes negligible and so

$$\lim_{k\to\infty}(\frac{k}{|N(x_k)|}\Gamma) = \lim_{k\to\infty}(\frac{\mu(\sqrt[d]{|N(x_k)|})}{|N(x_k)|}\Gamma) = \lim_{r\to\infty}(\mu(r)\frac{1}{r^d}\Gamma) = \lambda$$

Therefore for any $\epsilon > 0$, and $s > 1$, there is a large enough k_0 so that for all $k \geq k_0$ we have

$$(\frac{\lambda}{\Gamma} - \epsilon)^s < \frac{k^s}{|N(x_k)|^s} < (\frac{\lambda}{\Gamma} + \epsilon)^s$$

Consequently, by dividing everything by k^s and summing over all $k \geq k_0$, we get

$$(\frac{\lambda}{\Gamma} - \epsilon)^s(\sum_{k=k_0}^{\infty}\frac{1}{k^s}) < \sum_{k=k_0}^{\infty}\frac{1}{|N(x_k)|^s} < (\frac{\lambda}{\Gamma} + \epsilon)^s(\sum_{k=k_0}^{\infty}\frac{1}{k^s})$$

Let $A(s) = \sum_{k=1}^{k_0-1} \frac{1}{k^s}$ and $B(s) = \sum_{k=1}^{k_0-1} \frac{1}{|N(x_k)|^s}$. By multiplying the above by $s-1$ and taking the limit as $s \to 1^+$, we get

$$\lim_{s\to 1^+}((s-1)(\frac{\lambda}{\Gamma}-\epsilon)^s(\zeta_{\mathbb{Q}}(s)-A(s)) < \lim_{s\to 1^+}((s-1)(\sum_{k=1}^{\infty}\frac{1}{|N(x_k)|^s}-B(s))) <$$
$$< \lim_{s\to 1^+}((s-1)(\frac{\lambda}{\Gamma}+\epsilon)^s(\zeta_{\mathbb{Q}}(s)-A(s)))$$

$A(s)$ and $B(s)$ are finite quantities for s close to 1, and so $\lim_{s\to 1^+}((s-1)A(s)) = \lim_{s\to 1^+}((s-1)B(s)) = 0$. Additionally, we know that $\lim_{s\to 1^+}((s-1)\zeta_{\mathbb{Q}}(s)) = 1$, and the sum $\sum_{k=1}^{\infty}\frac{1}{|N(x_k)|^s}$ is simply $\sum_{m\in \mathfrak{a}_{c^{-1}}\cap D}\frac{1}{|N(m)|^s}$. The inequality boils down to

$$\frac{\lambda}{\Gamma}-\epsilon < \lim_{s\to 1^+}((s-1)\sum_{m\in \mathfrak{a}_{c^{-1}}\cap D}\frac{1}{|N(m)|^s}) < \frac{\lambda}{\Gamma}+\epsilon$$

Since ϵ can be chosen to be arbitrarily small, the proposition follows. □

What follows now is the computation of the two volumes Γ and λ. It will then be straightforward to put everything together. The formula will simplify heavily once we know Γ.

Lemma 4.3.5. *The covolume of the ideal $\mathfrak{a}$ when embedded as a lattice into V is $\frac{N(\mathfrak{a})\sqrt{|\Delta(K)|}}{2^{r_2}}$.*

Proof. We prove first that $\frac{\sqrt{|\Delta(K)|}}{2^{r_2}}$ is the covolume of $\mathcal{O}_K$. The covolume of $\mathcal{O}_K$ is the determinant of a $\mathbb{Z}$-basis for $\mathcal{O}_K$ when considering V as a real vector space $\mathbb{R}^d$. This is closely related to the discriminant, but realizing each complex place as a real plane will cost us a factor of 2.

Let $\{a_1 \dots a_d\}$ be a $\mathbb{Z}$-basis basis for $\mathcal{O}_K$. For a real place given by an embedding σ, the row in the discriminant matrix corresponding to that embedding is $\sigma(a_1)\dots\sigma(a_d)$. These are coincidently also the σ-components of the vectors $a_1 \dots a_d$ when considering $\mathbb{R}^{r_1}\oplus\mathbb{C}^{r_2}$ as a real vector space $\mathbb{R}^d$.

For complex embeddings σ it is more complicated. In the discriminant matrix, we get two rows $\sigma(a_1)\dots\sigma(a_d)$ and $\overline{\sigma(a_1)}\dots\overline{\sigma(a_d)}$. However, the complex plane $\mathbb{C}$ corresponding to σ is realized as a real plane with real and imaginary coordinates, forming the two rows $Re(\sigma(a_1))\dots Re(\sigma(a_n))$ and $Im(\sigma(a_1))\dots Im(\sigma(a_n))$. The determinant calculation for this specific complex place is

$$\det\begin{bmatrix}\dots & \dots & \dots\\ Re(\sigma(a_1)) & \dots & Re(\sigma(a_n))\\ Im(\sigma(a_1)) & \dots & Im(\sigma(a_n))\end{bmatrix} = \det\begin{bmatrix}\dots & \dots & \dots\\ \sigma(a_1) & \dots & \sigma(a_n)\\ Im(\sigma(a_1)) & \dots & Im(\sigma(a_n))\end{bmatrix} =$$
$$= \frac{1}{2}\det\begin{bmatrix}\dots & \dots & \dots\\ \sigma(a_1) & \dots & \sigma(a_n)\\ 2Im(\sigma(a_1)) & \dots & 2Im(\sigma(a_n))\end{bmatrix} = -\frac{1}{2}\det\begin{bmatrix}\dots & \dots & \dots\\ \sigma(a_1) & \dots & \sigma(a_n)\\ \overline{\sigma(a_1)} & \dots & \overline{\sigma(a_n)}\end{bmatrix}$$

but the sign will not contribute anything and so indeed every complex place loses us a factor of 2. There are r_2 complex places and so $\frac{\sqrt{|\Delta(K)|}}{2^{r_2}}$ is the covolume of $\mathcal{O}_K$.

Now $\mathfrak{a}$ is a sublattice of $\mathcal{O}_K$ and we need $[\mathcal{O}_K : \mathfrak{a}] = |\mathcal{O}_K/\mathfrak{a}| = N(\mathfrak{a})$ translates of $\mathfrak{a}$ to cover $\mathcal{O}_K$. It follows that the covolume of $\mathfrak{a}$ is $N(\mathfrak{a})\frac{\sqrt{|\Delta(K)|}}{2^{r_2}}$ and we are done. □

Before computing λ, which will require more work, we simplify our current formula for the residue as:

$$\lim_{s\to 1^+}((s-1)\zeta_K(s)) = \sum_{c\in Cl(K)} N(\mathfrak{a}_{c^{-1}}) \lim_{s\to 1^+}((s-1)\sum_{m\in\mathfrak{a}_{c^{-1}}\cap D}\frac{1}{|N(m)|^s}) =$$

$$= \sum_{c\in Cl(K)} N(\mathfrak{a}_{c^{-1}})\frac{\lambda}{\Gamma_{\mathfrak{a}_{c^{-1}}}} = \sum_{c\in Cl(K)} N(\mathfrak{a}_{c^{-1}})\frac{2^{r_2}\lambda}{N(\mathfrak{a}_{c^{-1}})\sqrt{|\Delta(K)|}} = \frac{2^{r_2}\lambda h_K}{\sqrt{|\Delta(K)|}}$$

We now compute the volume $\lambda = Vol(T)$. Recall that T is the subset of V of elements x satisfying

- $|N(x)| \leq 1$
- $Log(x) = \sum_{i=0}^{r_1+r_2-1} b_i\epsilon_i$ with $0 \leq b_i < 1$ for all $1 \leq i \leq r_1 + r_2 - 1$
- $x_1 > 0$ if K has a real place, and $0 \leq \arg(x_1) < \frac{2\pi}{\omega_K}$ if K is totally complex.

Lemma 4.3.6.

$$\lambda = \mathrm{Vol}(T) = \frac{\pi^{r_2}2^{r_1}\mathcal{R}_K}{\omega_K}$$

Proof. The volume of T can be computed using the coordinates of V realized as the real vector space $\mathbb{R}^d$. We will first compute the volume of T_0, the subset of T whose real places are positive and the third constraint above is removed.

The first variable change is to use polar coordinates for all the complex coordinates. Let $g_1 \dots g_{r_1}$ be the real variables and $g_{r_1+1} \dots g_{r_1+r_2}$ the radii of the complex variables. Let $\theta_1 \dots \theta_{r_2}$ be the corresponding arguments of the complex variables. The Jacobian of this transformation is known to be the product of the radii $J_1 = \prod_{i=1}^{r_2} g_{r_1+i}$.

The second variable change is to relate the radii to our basis $\{\epsilon_0 \dots \epsilon_{r_1+r_2-1}\}$ of the *Log*-image of V. There is a reason we chose ϵ_0 as we did and that reason will be shown now. Let $x \in V_0$ so that it has positive real components. Write $Log(x) = \sum_{i=0}^{r_1+r_2-1} b_i\epsilon_i$ and $x = \sqrt[d]{|N(x)|}y$ where y has norm 1. Then $Log(x) = Log(y) + \frac{1}{d}\log(N(x))\epsilon_0$ but y is a unit so its ϵ_0 component is 0. As a result $\frac{1}{d}\log(N(x)) = b_0$. From now on we will denote $c_0 = N(x)$ and $c_i = b_i$ for all other i, so that all the variables c_j have the constraint of lying in the interval $[0, 1)$.

Now we can give some relations between the radii g_j and the coefficients c_j as

$$\log(g_i) = \frac{1}{d}\log(c_0) + \sum_{j=1}^{r_1+r_2-1} c_j \log(|\sigma_i(\eta_j)|)$$

This will constitute a variable change, whose Jacobian is

$$J_2 = \det\begin{bmatrix} \frac{\partial g_1}{\partial c_0} & \dots & \frac{\partial g_1}{\partial c_{r_1+r_2-1}} \\ \dots & \dots & \dots \\ \frac{\partial g_{r_1+r_2}}{\partial c_0} & \dots & \frac{\partial g_{r_1+r_2}}{\partial c_{r_1+r_2-1}} \end{bmatrix} == \frac{\prod_{i=1}^{r_1+r_2} g_i}{dc_0}\det\begin{bmatrix} 1 & \dots & \log(|\sigma_1(\eta_{r_1+r_2-1})|) \\ \dots & \dots & \dots \\ 1 & \dots & \log(|\sigma_{r_1+r_2}(\eta_{r_1+r_2-1})|) \end{bmatrix}$$

Remember that $c_0 = N(x) = \prod_{i=1}^{r_1} g_i \prod_{i=1}^{r_2} g_{r_1+i}^2$. Also, the rightmost determinant is the regulator minus the factors of 2 for the complex places, but having an extra factor of d for the column of 1's. Therefore it is equal to $\frac{d\mathcal{R}_K}{2^{r_2}}$. We get that $J_2 = \frac{\mathcal{R}_K}{2^{r_2}\prod_{i=1}^{r_2} g_{r_1+i}}$. The product of the two Jacobians is $J_1 J_2 = \frac{\mathcal{R}_K}{2^{r_2}}$ and the integral is now simple due to our constraints on the c_i. It becomes

$$\mathrm{Vol}(T_0) = J_0 J_1 \int_{c=0}^{1} \int_{\theta=0}^{2/pi} dc_0 \cdots dc_{r_1+r_2-1} d\theta_1 \cdots d\theta_{r_2} = J_0 J_1 (2\pi)^{r_2}$$

which is evaluated as $(\pi)^{r_2}\mathcal{R}_K$. To get $\mathrm{Vol}(T)$ we need to multiply by 2 for each real place and divide by ω_K to satisfy the third constraint. This gives $\lambda = \mathrm{Vol}(T) = \frac{2^{r_1}(\pi)^{r_2}\mathcal{R}_K}{\omega_K}$ as required. □

We are done. By plugging λ into our most recent formula we get

$$lim_{s\to 1^+}((s-1)\zeta_K(s)) = \frac{2^{r_2}\lambda h_K}{\sqrt{|\Delta(K)|}} = \frac{2^{r_1}(2\pi)^{r_2}\mathcal{R}_K h_K}{\omega_K\sqrt{|\Delta(K)|}}$$

This completes the proof of the analytic class number formula.

4.4 Applications and examples of the analytic class number formula

We will show in some cases that by understanding the decomposition of primes in field extensions $K : \mathbb{Q}$, we can factorize the Dedekind zeta function of K into a product of Dirichlet L-series. In the next subsection we will investigate this connection further. In general, it will work for any abelian extension K of $\mathbb{Q}$, which is a subfield of some cyclotomic field $\mathbb{Q}(\zeta)$ by Kronecker-Weber. As a result K will be induced by a group of Dirichlet characters with modulus equal to the order of ζ, and the Dedekind zeta function of K will factor into a product of the corresponding Dirichlet L-series. We begin by proving this result for cyclotomic fields themselves.

Proposition 4.4.1. *Let $K = \mathbb{Q}(\zeta_m)$ be a cyclotomic field, for an integer $m > 2$. Then*

$$\zeta_K(s) = \prod_{\chi\in\widehat{\mathbb{Z}/(m)}} L(s,\chi) = \zeta_{\mathbb{Q}}(s) \prod_{\chi\in\widehat{\mathbb{Z}/(m)}\setminus\chi_0} L(s,\chi)$$

where χ_0 is the trivial character modulo m.

Proof. We start with the Euler product form for Dirichlet L-series, which is

$$\prod_{\chi\in\widehat{\mathbb{Z}/(m)}} L(s,\chi) = \prod_p \prod_{\chi\in\widehat{\mathbb{Z}/(m)}} \frac{1}{1-\chi(p)p^{-s}}$$

Note that the images of the characters in $\widehat{\mathbb{Z}/(m)}$ will be m^{th} roots of unity. Let $\sigma \in \mathrm{Gal}(K:\mathbb{Q}) = (\mathbb{Z}/(m))^\times$ be the Artin symbol for some prime p not dividing m. The order of σ is the inertia degree of p and it also equals $\mathrm{ord}_m(p) = f_p$. Hence the characters evaluated at p are f_p^{th} roots of

unity. For each f_p^{th} root of unity there are $\frac{\phi(m)}{f_p} = g_p$ characters that send p to it, where g_p is the number of primes above p. As a result we can write

$$\prod_p \prod_{\chi \in \widehat{\mathbb{Z}/(m)}} \frac{1}{1-\chi(p)p^{-s}} = \prod_p \prod_{k=0}^{f_p-1} (\frac{1}{1-\zeta_{f_p}^k p^{-s}})^{g_p}$$

We know that $\prod_{k=0}^{f_p-1}(1-\zeta_{f_p}^k p^{-s}) = 1-p^{-fs}$ and p^{f_p} is the norm of the primes above p, of which there are g_p. From this we finally get

$$\prod_p \prod_{k=0}^{f_p-1} (\frac{1}{1-\zeta_{f_p}^k p^{-s}})^{g_p} = \prod_{\mathfrak{p}} \frac{1}{1-N(\mathfrak{p})^{-s}} = \zeta_K(s)$$

□

Corollary 4.4.2. *Let χ be a nontrivial character modulo m. Then $L(s,\chi)$ does not vanish at $s=1$.*

Proof. Let $K = \mathbb{Q}(\zeta_m)$ be a cyclotomic field. $\zeta_K(s)$ has a simple pole at $s=1$ which is contributed to by $\zeta_{\mathbb{Q}}(s)$ in its factorization in Proposition 4.4.1. As a result $\prod_{\chi \in \widehat{\mathbb{Z}/(m)} \setminus \chi_0} L(s,\chi)$ has no pole or zero at $s=1$ and the L-series attached to the nontrivial character χ in question divides this product. If it had a zero at $s=1$, then some other L-series attached to some nontrivial character modulo m would have to diverge (have a pole) at $s=1$, but we know that this is not possible. Hence every L-series attached to a nontrivial character must not vanish at $s=1$. □

We produce a result for quadratic fields, which follows from quadratic reciprocity. This will allow us to analytically determine the class numbers of quadratic fields.

Proposition 4.4.3. *Let $K = \mathbb{Q}(\sqrt{a})$ be a quadratic field with discriminant d. Then*

$$\zeta_K(s) = \zeta_{\mathbb{Q}}(s)L(s,\chi)$$

where χ is the Kronecker symbol $(\frac{\cdot}{d})$.

Proof. Let $K = \mathbb{Q}(\sqrt{a})$ be a quadratic field with discriminant d. Then

$$\zeta_K(s) = \prod_{p:(\frac{d}{p})=0} \frac{1}{1-p^{-s}} \prod_{p:(\frac{d}{p})=-1} \frac{1}{1-p^{-2s}} \prod_{p:(\frac{d}{p})=1} (\frac{1}{1-p^{-s}})^2$$

by our understanding of prime decomposition in K. We can already take out a factor of the Riemann zeta function to give

$$\zeta_K(s) = \zeta_{\mathbb{Q}}(s) \prod_{p:(\frac{d}{p})=-1} \frac{1}{1+p^{-s}} \prod_{p:(\frac{d}{p})=1} \frac{1}{1-p^{-s}}$$

We can see that for special cases we can use quadratic reciprocity to write the sign in front of p^{-s} as a Dirichlet character. By general quadratic reciprocity we get

$$\zeta_K(s) = \zeta_{\mathbb{Q}}(s) \prod_p \frac{1}{1 - (\frac{p}{d})p^{-s}} = \zeta_{\mathbb{Q}}(s) L(s, (\frac{\cdot}{d}))$$

□

We give easy examples of the above result for illustrative purposes.

Example 4.4.4. *Let $K = \mathbb{Q}(\sqrt{-3})$. We need to compute $L(1, (\frac{\cdot}{3}))$ which equals*

$$\sum_{n \equiv 1 \pmod 3} \frac{1}{n} - \sum_{n \equiv 2 \pmod 3} \frac{1}{n}$$

We will use generating functions to compute this sum. We have

$$L(1, (\frac{\cdot}{3})) = \left[\sum_{n \equiv 1 \pmod 3} \frac{x^n}{n} - \sum_{n \equiv 2 \pmod 3} \frac{x^n}{n} \right]_0^1$$

which comes from the integral

$$L(1, (\frac{\cdot}{3})) = \int_0^1 (\sum_{n \equiv 0 \pmod 3} x^n - \sum_{n \equiv 1 \pmod 3} x^n) dx = \int_0^1 \frac{1-x}{1-x^3} dx$$

This particular integral is easily evaluated by hand as

$$\int_0^1 \frac{1-x}{1-x^3} dx = \int_0^1 \frac{1}{1+x+x^2} dx = \int_0^{\frac{1}{2}} \frac{1}{\frac{3}{4}+x^2} dx = \frac{1}{\sqrt{3}} \arctan(\sqrt{3}) = \frac{\pi}{3\sqrt{3}}$$

Now using the analytic class number formula gives

$$\lim_{s \to 1^+} ((s-1)\zeta_K(s)) = \lim_{s \to 1^+} ((s-1)\zeta_{\mathbb{Q}}(s) L(s, (\frac{\cdot}{3}))) = L(1, (\frac{\cdot}{3})) = \frac{\pi}{3\sqrt{3}} = \frac{2^{r_1}(2\pi)^{r_2} \mathcal{R}_K h_K}{\omega_K \sqrt{|\Delta(K)|}}$$

We have $r_1 = 0$, $r_2 = 1$ and $\omega_K = 6$. The regulator is trivial because imaginary quadratic fields have no non-torsion units. Finally, the discriminant is -3 and so we get

$$\frac{\pi}{3\sqrt{3}} = \frac{2\pi h_K}{6\sqrt{3}}$$

It follows that $h_K = 1$.

Here is the simplest example for real quadratic fields, where the regulator is nontrivial.

Example 4.4.5. *Let $K = \mathbb{Q}(\sqrt{5})$. We need to compute $L(1, (\frac{\cdot}{5}))$ which equals*

$$\sum_{n \equiv 1,4 \pmod 5} \frac{1}{n} - \sum_{n \equiv 2,3 \pmod 5} \frac{1}{n}$$

We will use generating functions to compute this sum. We have

$$L(1, (\frac{\cdot}{5})) = \left[\sum_{n \equiv 1,4 \pmod 5} \frac{x^n}{n} - \sum_{n \equiv 2,3 \pmod 5} \frac{x^n}{n} \right]_0^1$$

which comes from the integral

$$L(1, (\frac{\cdot}{5})) = \int_0^1 (\sum_{n \equiv 0,3 \pmod 5} x^n - \sum_{n \equiv 1,2 \pmod 5} x^n) dx = \int_0^1 \frac{1 - x - x^2 + x^3}{1 - x^5} dx$$

This integral can be evaluated by computer to give

$$L(1, \chi) = \int_0^1 \frac{1 - x - x^2 + x^3}{1 - x^5} dx = \frac{2^{r_1} (2\pi)^{r_2} \mathcal{R}_K h_K}{\omega_K \sqrt{|\Delta(K)|}} \approx 0.43041$$

For $\mathbb{Q}(\sqrt{5})$ we have $\omega_K = 2, r_1 = 2$ and $r_2 = 0$. The discriminant is 5, and it remains to compute the regulator. We have to solve Pell's equation which is

$$x^2 - 5y^2 = \pm 1$$

in this case. The smallest solution is $x = \frac{1}{2}, y = \frac{1}{2}$ which gives the fundamental unit $\frac{1+\sqrt{5}}{2}$. The regulator is therefore $\log(\frac{1+\sqrt{5}}{2}) \approx 0.4812$. We finally get

$$0.43041 \approx \frac{4 h_K}{2\sqrt{5}} \cdot 0.4812 \approx 0.43041 h_K$$

so $h_K = 1$ since it must be a positive integer.

For quadratic fields, this method gives a more general formula for $L(1, (\frac{\cdot}{D_K}))$, which is

$$L(1, (\frac{\cdot}{D_K})) = \int_0^1 \frac{\sum_{m=1}^{D_K - 1} (\frac{m}{D_K}) x^{m-1}}{1 - x^{D_K}} dx$$

We will find a different way to compute $L(1, \chi)$ for general Dirichlet characters χ in Section 5, using generalized Bernoulli numbers.

4.5 Dirichlet characters and associated number fields

We begin by describing how one could associate abelian number fields to groups of Dirichlet characters for the Galois group $\mathrm{Gal}(\mathbb{Q}(\zeta_m) : \mathbb{Q}) \cong (\mathbb{Z}/(m))^\times$ and vice versa. We do not give proofs, but instead refer the reader to [Was97, Chapter 3].

Definition 4.5.1. *Let X be a group of Dirichlet characters modulo m, for m minimal (set m to be the LCM of the conductors of the characters in X). Let $G = \mathrm{Gal}(\mathbb{Q}(\zeta_m) : \mathbb{Q})$. This is canonically isomorphic to $(\mathbb{Z}/(m))^\times$, where the element $k \pmod m$ corresponds to the automorphism that sends $\zeta_m \to \zeta_m^k$. We can then associate X to a number field K_X in the following way. We define the normal subgroup $\ker(X) \triangleleft G$ as*

$$\ker(X) := \{h \in G : \chi(h) = 1 \, \forall \chi \in X\}$$

then we set $K_X = \mathbb{Q}(\zeta_m)^{\ker(X)}$. This will be an abelian number field since G is abelian.

If a group of Dirichlet characters has a modulus ab which is not minimal as above, then it will be induced by a group of Dirichlet characters for a smaller modulus, say b. By Galois theory, the associated number field will be the same.

Theorem 4.5.2. *There is a one-to-one inclusion preserving bijection between groups of Dirichlet characters and abelian Galois extensions of $\mathbb{Q}$.*

The above definition gives the forward association. We illustrate the other direction. Let M be an abelian number field. By the Kronecker-Weber theorem, we have $M = \mathbb{Q}(\zeta_m)^H$ for some minimal integer m and subgroup H of $\mathrm{Gal}(\mathbb{Q}(\zeta_m)) = (\mathbb{Z}/(m))^\times$. Let $\mathrm{cok}(H)$ be the defined as

$$\mathrm{cok}(H) := \{\chi \in \widehat{(\mathbb{Z}/(m))^\times} : \chi(h) = 1 \, \forall h \in H\}$$

Then we say $\mathrm{cok}(H)$ is the group of Dirichlet characters associated to M.

Theorem 4.5.3. *Let X be a group of Dirichlet characters and let K_X be its associated field. Then the integral prime p ramifies in K_X if and only if $\chi(p) = 0$ for every $\chi \in X$.*

Proof. See [Was97, Corollary 3.6]. □

Proposition 4.5.4. *Let X be a group of Dirichlet characters and let K_X be their associated field. Then we have*

$$\zeta_{K_X}(s) = \prod_{\chi \in X} L(s, \chi)$$

Proof. We begin as in the proof for the factorization of the cyclotomic zeta function starting with

$$\zeta_{K_X}(s) = \prod_p \prod_{\mathfrak{p}|p} \frac{1}{1 - N(\mathfrak{p})^{-s}} = (\prod_p \frac{1}{1 - p^{-f_p s}})^{n_p}$$

where f_p is the residue field degree of the primes above p and n_s is the number of primes that p factors into, including ramified primes. Remember that K_X is an abelian Galois extension of $\mathbb{Q}$. As a result, for unramified primes p the Artin symbol σ_p can be defined on p and it will have order

f_p. X becomes the group of characters for the group $G = \mathrm{Gal}(K_X : \mathbb{Q}) \cong \mathrm{Gal}(\mathbb{Q}(\zeta_m) : \mathbb{Q})/\ker(X)$. As a result X will send σ_p to the f_p^{th} roots of unity. We can therefore factor, for unramified p,

$$(\frac{1}{1-p^{-f_p s}})^{n_p} = \prod_{j:(j,f_p)=1} (\frac{1}{1-\zeta_{f_p}^j p^s})^{n_p} = \prod_{\chi \in X} \frac{1}{1-\chi(p)p^s}$$

By Theorem 4.5.3, $\chi(p) = 0$ for any ramified primes p so we can safely write

$$\zeta_{K_X}(s) = \prod_p \prod_{\chi \in X} \frac{1}{1-\chi(p)p^s} = \prod_{\chi \in X} L(s,\chi)$$

□

5 Arithmetic of cyclotomic fields and Fermat's last theorem

We will use the tools of analytic and algebraic number theory to prove Fermat's last theorem (FLT) for regular primes p. A prime p is called regular if it does not divide the class number of $\mathbb{Z}(\zeta_p)$. Conjecturally the regular primes make up about 61% of all primes so this is a significant result. In the process we will develop a wider picture of the arithmetic of cyclotomic fields. Fermat's last theorem for integers x, y, z and an integer $p > 2$ states that

$$x^p + y^p = z^p \implies xyz = 0$$

It is easy to see that if FLT holds for some positive integer a, then it also holds for any positive integer b divisible by a. Therefore to prove FLT it is sufficient to prove it for all odd primes p and for 4.

Additionally, the correct number field to work with this problem is the cyclotomic field $\mathbb{Q}(\zeta_p)$. In this field we can factorize the left hand side as

$$\prod_{k=0}^{p-1}(x + \zeta_p^k y) = z^p$$

This is the setup we will continuously refer back to. There are two additional elementary assumptions that could be made on the x, y, z. Firstly, we can assume that they are pairwise coprime. Otherwise, all of them would be divisible by some integer n but we could get a simpler solution by using $\frac{x}{n}, \frac{y}{n}, \frac{z}{n}$ instead.

Secondly, we can assume that $x \not\equiv y \pmod p$. Otherwise we can simply use the equation $x^p + (-z)^p = (-y)^p$ instead, since $-z \equiv (-z)^p \equiv (-x)^p + (-y)^p \equiv -(x + y) \equiv -2x$ and if $z \equiv x$ as well then $-2x \equiv x \pmod p$. This would imply that either $p = 3$ or p divides (x, y, z). The latter case is not possible because we assumed that x, y, z are pairwise coprime. In the case $p = 3$ we know that cubes must be congruent to $0, 1, 8$ modulo 9 and the only viable counter-examples to our assumption are $\{x, y, z\} \equiv \{1, 1, -1\} \pmod 3$ and $\{x, y, z\} \equiv \{2, 2, -2\} \pmod 3$. However their cubes would correspond modulo 9 to $\{x^3, y^3, z^3\} \equiv \{1, 1, 8\} \pmod 9$ and $\{x^3, y^3, z^3\} \equiv \{8, 8, 1\} \pmod p$ respectively which are not valid.

These assumptions can be assumed simultaneously and will be assumed from now on. We split up the proof into two cases. The first case is when $p \geq 5$ does not divide xyz, and the second case is when $p \geq 5$ divides xyz. We will prove the cases $p = 4$ and $p = 3$ separately.

5.1 Arithmetic of cyclotomic fields

To start off with we will prove some more general facts about CM fields. These are defined below.

Definition 5.1.1. *Let K be a number field. The maximally real subfield K^+ of K is the subfield generated by all real elements of K. K is called a CM field if it is totally complex and $[K : K^+] = 2$.*

Lemma 5.1.2. *Let K be a Galois CM field with maximally real subfield K^+. Let $H < G = \mathrm{Gal}(K : \mathbb{Q})$ be the subgroup generated by complex conjugation. Then K^+ is the fixed field of H and H is central in G.*

Proof. K^+ is clearly fixed by complex conjugation, and it has codimension 2 in K because K is a CM field. We have the tower $K : K^H : K^+$ and since $K \neq K^H$ we must have $K^H = K^+$.

Now we will show that K^+ is normal, so that $H \triangleleft G$. Let $a \in K^+$ and let f be its minimal polynomial. Suppose that f does not split in K^+, then it must have as a factor a degree 2 polynomial g which is irreducible in K^+ but splits in K. Any other irreducible factor of f will split in K so it must have degree 2 or be linear. It will suffice to prove that g actually splits in K^+.

The roots of g must be real, since K^+ is totally real. Let L be the splitting field of g so that we have a tower $K : L : K^+$. Then L is totally real still so we cannot have $L = K$ which implies that $K = K^+$. As a result g splits in K^+ so every irreducible polynomial with a root in K^+ will split in K^+, and so K^+ is normal.

A normal subgroup of degree 2 must be central. To see this, let τ denote complex conjugation and let $a \in G$. Then $a\tau a^{-1} = 1$ or τ. The first case would imply $a\tau = a$ which could not occur as $\tau \neq 1$. As a result we must have $a\tau a^{-1} = \tau$ for all a and so $\langle \tau \rangle$ is central. □

Lemma 5.1.3. *Let K be a number field and let $a \in \mathcal{O}_K$. If $||a||_\sigma = 1$ for all archimedean places σ, then a must be a root of unity.*

Proof. The key fact is that if $||a||_\sigma = 1$ for all archimedean places σ, then $||a^k||_\sigma = 1$ for all σ and integers k as well by multiplicativity of absolute values. We will show that the subset of $\mathcal{O}_K$ satisfying this property is finite. Since every finite cancellative monoid is a group, this subset must be the torsion subgroup of the group of units, which is the group of roots of unity.

Recall the method used to prove that the *Log*-embedding of K is discrete in Theorem 3.3.3. A similar method is used here. The subset of $\mathcal{O}_K$ satisfying the property above have minimal monic polynomials with bounded coefficients due to Vieta formulae. Hence only a finite number of polynomials could be a minimal polynomial to such an element and so this subset must be finite. □

Proposition 5.1.4. *Let K be a Galois CM field with maximally real subfield K^+. Let $r \in \mathcal{O}_K$ and suppose that $u = \frac{r}{\overline{r}} \in \mathcal{O}_K$. Then u is a root of unity.*

Proof. Since every place is complex and the subgroup generated by complex conjugation is central in $\mathrm{Gal}(K : \mathbb{Q})$ we get

$$||u||_\sigma = \sigma(u)\overline{\sigma}(u) = \frac{\sigma(r)}{\sigma(\overline{r})}\frac{\overline{\sigma}(r)}{\overline{\sigma}(\overline{r})} = \frac{\sigma(r)}{\overline{\sigma}(r)}\frac{\overline{\sigma}(r)}{\sigma(r)} = 1$$

and so u is an integral element which is 1 in all archimedean places. As a result it must be a root of unity by Lemma 5.1.3. □

Proposition 5.1.5. *Let $K = \mathbb{Q}(\zeta_m)$ be a cyclotomic field. Then K is a CM field with totally real subfield $\mathbb{Q}(\zeta_m + \zeta_m^{-1})$.*

Proof. Firstly note that $\zeta_m + \zeta_m^{-1} = \zeta_m + \tau(\zeta_m)$ is real, where τ is complex conjugation. We have as a result the inclusion $\mathbb{Q}(\zeta_m + \zeta_m^{-1}) \subset K^H$ where $H = \langle\tau\rangle$. To turn this into an equality, which would complete the proof, we must show that $K : \mathbb{Q}(\zeta_m + \zeta_m^{-1})$ is a degree 2 field extension.

ζ_m is the root of the quadratic $x^2 - (\zeta_m + \zeta_m^{-1})x + 1$ with coefficients in $\mathbb{Q}(\zeta_m + \zeta_m^{-1})$, so it follows that $K : \mathbb{Q}(\zeta_m + \zeta_m^{-1})$ is a degree 2 field extension. □

The next proof is from [WE, Proposition 13].

Proposition 5.1.6. *Let $K = \mathbb{Q}(\zeta_p)$ for an odd prime p. Then $\mathcal{O}_K^\times = \langle\zeta_p\rangle\mathcal{O}_{K^+}^\times$.*

Proof. Let $\epsilon \in \mathcal{O}_K^\times$ and let $u = \frac{\epsilon}{\tau(\epsilon)} \in \mathcal{O}_K^\times$. Then $u = \frac{\epsilon}{\tau(\epsilon)} = (-\zeta_p)^k$ for some integer k by Proposition 5.1.4. We will show that u is in fact a p^{th} root of unity.

Write $u = \sum_{i=0}^{p-2} a_i\zeta_p^i$ for integers a_i and note that

$$u^p \equiv (\sum_{i=0}^{p-2} a_i\zeta_p^i)^p \equiv \sum_{i=0}^{p-2}(a_i\zeta_p^i)^p \equiv \sum_{i=0}^{p-2} a_i^p \equiv \pm 1 \pmod{p}$$

due to the freshman's dream. In fact, any element of $\mathcal{O}_K$ raised to the p^{th} power will be congruent to a rational integer modulo p. Now write

$$\epsilon^p \equiv \pm\tau(\epsilon^p)$$

but note that ϵ^p is also a rational integer modulo p, not congruent to 0 and fixed by τ, so that $\pm$ becomes $+$. It follows that u is a p^{th} root of unity.

Write $\frac{\epsilon}{\tau(\epsilon)} = \zeta_p^k$ assuming that ϵ is not real in which case ζ_p^k is not 1. Then we can set $r \equiv -\frac{k}{2} \pmod{p}$ since p is an odd prime. Now set $\delta = \zeta_p^r\epsilon$. We find that $\tau(\delta) = \tau(\zeta_p^r\epsilon) = \zeta_p^{-r}\tau(\epsilon) = \zeta_p^{-r}\epsilon\zeta_p^{-k} = \zeta_p^{-r}\epsilon\zeta_p^{2r} = \zeta_p^r\epsilon = \delta$. This tells us that δ is a real unit and so the unit $\epsilon = \delta\zeta_p^{-r}$ can be written as the product of a real unit and a root of unity. It follows that $\mathcal{O}_K^\times = \langle\zeta_p\rangle\mathcal{O}_{K^+}^\times$. □

Lemma 5.1.7. *Let $K = \mathbb{Q}(\zeta_p)$ for an odd prime p. Then $\frac{1-\zeta_p^j}{1-\zeta_p^k}$ is a unit for any $j, k \not\equiv 0 \pmod{p}$.*

Proof. As ideals we have $(1 - \zeta_p^j) = (1 - \zeta_p^k)$ for any $j, k \not\equiv 0 \pmod{p}$, because they both lie above p which totally ramifies. As a result their quotient will be a unit. □

Proposition 5.1.8. *Let $K = \mathbb{Q}(\zeta_p)$ for an odd prime p. Then $\mathcal{O}_{K^+} = \mathbb{Z}[\zeta_p + \zeta_p^{-1}]$*

Proof. The discriminant of K^+ divides that of K and so we only need to consider p as a potential denominator in our algorithm for finding the ring of integers. Assume we can add an algebraic integer $x = \frac{1}{p}\sum_{j=0}^{\frac{p-1}{2}} a_j(\zeta_p + \zeta_p^{-1})^j$ for $0 \le a_j < p$. Then p would have to divide each of the coefficients of the ζ_p^i else we would also get a new algebraic integer for K. Now the coefficients to $\zeta_p^{\frac{p-1}{2}}$ and $\zeta_p^{-\frac{p-1}{2}}$ is contributed to in the above sum only by the summand

$$\frac{a_{\frac{p-1}{2}}}{p}(\zeta_p + \zeta_p^{-1})^{\frac{p-1}{2}}$$

These coefficients are both $\frac{a_{\frac{p-1}{2}}}{p}$ which leads to $a_{\frac{p-1}{2}} = 0$. This was the base case. The sum now becomes $x = \frac{1}{p}\sum_{j=0}^{\frac{p-3}{2}} a_j(\zeta_p + \zeta_p^{-1})^j$. Proceed by induction, eliminating each a_j so that no new algebraic integers can be found. □

For the next two results we follow [Was97, Theorem 5.36].

Lemma 5.1.9. *Let $K = \mathbb{Q}(\zeta_p)$ for an odd prime p. Let u be a unit in K congruent to a rational integer a modulo p. Then u is a real unit.*

Proof. Write $u = \zeta_p^b\epsilon$ for some real unit ϵ. Then we have

$$\zeta_p^b \equiv (1-(1-\zeta_p))^b \equiv 1 - b(1-\zeta_p) \pmod{(1-\zeta_p)^2}$$

As ideals we have $(1-\zeta_p)^2 = (1-\zeta_p)(1-\zeta_p^{-1}) = (2-(\zeta_p+\zeta_p^{-1}))$. The ring of integers of $\mathbb{Q}(\zeta_p+\zeta_p^{-1})$ is $\mathbb{Z}[\zeta_p+\zeta_p^{-1}]$ by Proposition 5.1.8. As a result every element of $\mathbb{Z}[\zeta_p+\zeta_p^{-1}]$, including ϵ, is congruent to some rational integer modulo $(1-\zeta_p)^2$ because $\zeta_p + \zeta_p^{-1} \equiv 2 \pmod{(1-\zeta_p)^2}$.

Putting it all together gives us a rational integer c satisfying

$$u \equiv \zeta_p^b\epsilon \equiv (1-b(1-\zeta_p))c \equiv a \pmod{(1-\zeta_p)^2}$$

Now a must be a unit modulo $(1-\zeta_p)^2$ because u is a global unit. As a result c is also a unit modulo $(1-\zeta_p)^2$ and we have $c - a \equiv b(1-\zeta_p) \pmod{(1-\zeta_p)^2}$. Thus $(1-\zeta_p)$ divides $c-a$, but this is a rational integer and so in fact p divides $c-a$. Now $(1-\zeta_p)$ must divide b which is also a rational integer and so p divides b. As a result we finally get $u = \zeta_p^p\epsilon = \epsilon$ is real. □

Lemma 5.1.10 (Kummer's lemma)**.** *Let u be a unit of $\mathbb{Q}(\zeta_p)$ for a regular prime p such that $u \equiv a \pmod{p}$ for some rational integer a. Then u is the p^{th} power of some unit $v = \sqrt[p]{u}$.*

Proof. Since p is regular, the class group of $\mathbb{Z}[\zeta_p]$ has no p-torsion. By class field theory, this means that there is no unramified abelian extension of $\mathbb{Q}(\zeta_p)$ of degree p. Consider now the abelian extension $\mathbb{Q}(\zeta_p, \sqrt[p]{u}) : \mathbb{Q}(\zeta_p)$ which is the splitting field of the polynomial $x^p - u$. Then the degree of this extension divides p so it is either p or 1. We will show that this extension is unramified, and as a result the degree cannot be p, so it must be 1. It will follow that $\sqrt[p]{u} \in \mathbb{Q}(\zeta_p)$.

We will first show that every prime except $(1-\zeta_p)$ is unramified in this extension by considering the discriminant. A computation gives us

$$\Delta(x^p - u) = \prod_{i<j}(\zeta_p^i\sqrt[p]{u} - \zeta_p^j\sqrt[p]{u})^2 = u^{p-1}\prod_{i<j}(\zeta_p^i - \zeta_p^j)^2 = u^{p-1}(-1)^{\frac{p-1}{2}}p^{p-2}$$

Moreover the relative discriminant of the extension will divide this and so the only prime that may be ramified is the prime $(1-\zeta_p)$ above p in $\mathbb{Q}(\zeta_p)$. We will use local methods to show that this is also unramified.

Note that u is a p^{th} power if and only if $e = u^{p-1}$ is a p^{th} power because p and $p-1$ are coprime. Also $e \equiv a^{p-1} \equiv 1 \pmod{p}$ because a is a rational integer, and so we may write $e = pb + 1$ for some $b \in \mathbb{Z}[\zeta_p]$. Recall that all elements of $\mathbb{Z}[\zeta_p]$ are congruent to some rational integer modulo $\lambda = 1 - \zeta_p$. Hence we may write $b = c + y\lambda$ for some $c \in \mathbb{Z}$ and some $y \in \mathbb{Z}[\zeta_p]$. Then we get $e = 1+pc+py\lambda$. The norm of e is 1 because we have $e \equiv 1 \pmod{p}$. Modulo $p\lambda$ we get the relation

$$1 \equiv N(e) \equiv (1+pc)^{p-1} \equiv 1 + (p-1)pc \equiv 1 - pc \pmod{p\lambda}$$

since $\lambda | p$. As a result $pc \equiv 0 \pmod{p\lambda}$ so λ divides c. It follows that $e \equiv 1+pc+py\lambda \equiv 1 \pmod{p\lambda}$.

We may assume that e is a real unit by Lemma 5.1.9. From this assumption $e-1$ is also real and we have $e - 1 \equiv 0 \pmod{\lambda^2}$ by our calculations thus far. Now $(\lambda)^2$ is the prime above p in $\mathbb{Q}(\zeta_p + \zeta_p^{-1})$ and so $v_p(e-1)$ is a multiple of $\frac{2}{p-1}$. As a result we can strengthen our result $v_p(e-1) \geq \frac{p}{p-1}$ to $v_p(e-1) \geq \frac{p+1}{p-1}$ because p is odd.

Consider the monic polynomial $F(x) = \frac{(\lambda x - 1)^p + e}{\lambda^p}$. The constant term is $\frac{e-1}{\lambda^p}$ which lies in $\mathbb{Z}[\zeta_p]$. All other middle terms of $(\lambda x - 1)^p$ are multiples of $p\lambda$ and so all coefficients of F lie in $\mathbb{Z}[\zeta_p]$.

We will now invoke Hensel's lemma. We have $F(0) = \frac{e-1}{\lambda^p} \equiv 0 \pmod{\lambda}$, and $F'(0) = \frac{p\lambda}{\lambda^p} = \frac{p}{\lambda^{p-1}} \not\equiv 0 \pmod{\lambda}$ because it must be a unit. It follows that there is a root of F in $\mathbb{Z}[\zeta_p]_\lambda = \mathbb{Z}_p[\zeta_p]$. However this is a Galois extension of $\mathbb{Z}_p$ and so F splits in $\mathbb{Z}_p[\zeta_p]$ because it is the minimal polynomial of $\frac{1 - \sqrt[p]{e}}{\lambda}$. This means $\sqrt[p]{e} \in \mathbb{Z}[\zeta_p]_\lambda$ and so $\mathbb{Z}_p[\zeta_p, \sqrt[p]{e}] = \mathbb{Z}_p[\zeta_p]$. Local Galois groups correspond to global decomposition groups and so it follows that (λ) must split completely in the global extension $\mathbb{Q}(\zeta_p, \sqrt[p]{e}) : \mathbb{Q}(\zeta_p)$ and so it is an unramified extension. □

5.2 Case 1 of Fermat's last theorem

In the next two subsections we follow [Conc].

Theorem 5.2.1. *If $x^p + y^p = z^p$ for integers x, y, z and a regular prime $p \geq 5$ so that $p \nmid xyz$, then $xyz = 0$.*

Proof. This is the base case. Recall the factorization of ideals

$$\prod_{k=0}^{p-1} (x + \zeta_p^k y) = (z)^p$$

Given the assumptions, we will prove that the ideals on the LHS are pairwise coprime. Let $(x + \zeta_p^j y, x + \zeta_p^k y) = I_{jk}$ as ideals for some $0 \leq k < j < p$. Then we have as elements

$$(x + \zeta_p^k y) - (x + \zeta_p^j y) = y(\zeta_p^k - \zeta_p^j) = y\zeta_p^k(1 - \zeta_p^{j-k}) \in I_{jk}$$

$$(x + \zeta_p^j y) - \zeta_p^{j-k}(x + \zeta_p^k y) = x(1 - \zeta_p^{j-k}) \in I_{jk}$$

In particular, since ζ_p^k is a unit, we get $y(1 - \zeta_p^{j-k}) \in I_{jk}$. From the assumption that $(x, y) = (1)$, it follows that $(1 - \zeta_p^{j-k}) \in I_{jk}$. since $j \neq k$, we know that $(1 - \zeta_p^{j-k}) = (1 - \zeta_p)$ is a prime ideal

and so either $I_{jk} = (1-\zeta_p)$ or $I_{jk} = (1)$. The first case implies that $(1-\zeta_p)$ divides $(x+\zeta_p^j y)$ for every j and in particular $(1-\zeta_p)^{p-1} = (p)$ divides $(z)^p$. This contradicts the assumptions of this case. We must have $I_{jk} = (1)$ for every $j \neq k$ and so the ideals on the LHS are pairwise coprime.

The ideals on the LHS must be p^{th} powers of ideals, since the RHS is a p^{th} power of an ideal. We focus on a particular ideal $(x+\zeta_p y)$. Since p is regular the class group has no p-torsion and so the p^{th} root of the ideal $(x+\zeta_p y)$ would have to be principal. Therefore we get

$$(x+\zeta_p y) = (a)^p \implies x+\zeta_p y = ua^p$$

for some element $a \in \mathcal{O}_K$ and unit $u \in \mathcal{O}_K^\times$. Write $u = \zeta_p^k \epsilon$ for some real unit ϵ, which can be done by Proposition 5.1.6. Additionally let $a^p \equiv \alpha \pmod{p}$ for some rational integer α. Then

$$x+\zeta_p y \equiv \zeta_p^k \epsilon\alpha \pmod{p}$$

$$x+\zeta_p^{-1} y \equiv \zeta_p^{-k} \epsilon\alpha \pmod{p}$$

where we applied complex conjugation in the second case. Combining these two gives us

$$x+\zeta_p y - \zeta_p^{2k}(x+\zeta_p^{-1}y) \equiv x+\zeta_p y - \zeta_p^{2k}x - \zeta_p^{2k-1}y \equiv 0 \pmod{p}$$

This gives us an algebraic integer $\frac{x+\zeta_p y-\zeta_p^{2k}x-\zeta_p^{2k-1}y}{p}$. We must show that this cannot be an algebraic integer to get a contradiction. If all the roots of unity involved are distinct, then p would have to divide each of the coefficients, namely x and y. This would contradict our assumptions. Now we focus on the various cases where the roots of unity are not distinct. We start with the case $\zeta_p^{2k-1} = 1$ which gives

$$(x-y) - \zeta_p(x-y) \equiv 0 \pmod{p}$$

which requires $x \equiv y \pmod{p}$, but this contradicts our assumptions. The case $\zeta_p^{2k} = 1$ gives

$$\zeta_p y - \zeta_p^{-1} y \equiv 0 \pmod{p}$$

which requires p to divide y, contradicting the assumption $p \nmid xyz$. The last case $\zeta_p^{2k-1} = \zeta_p$ gives

$$x - \zeta_p^2 x \equiv 0 \pmod{p}$$

which requires p to divide x, contradicting the assumption $p \nmid xyz$. We have exhausted all the cases, each one leading to a contradiction and so we are done. □

5.3 Case 2 of Fermat's last theorem

Suppose now that $p|xyz$, in which case p can only divide one of x, y, z without breaching our assumptions. Then we can suppose $p|z$ without loss of generality, otherwise we can just rearrange the equation since p is odd. In particular we can still assume $x \not\equiv y \pmod{p}$ but this assumption

will not be required in this case. For convenience we denote $\lambda = 1 - \zeta_p$. We will prove this case by descent. Fix a solution in integers to

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

Let m be the highest power of λ dividing z. Then we can write

$$x^p + y^p + \lambda^{pm} z_0^p = 0$$

where $\lambda^{pm} z_0^p = (-z)^p$ and as a result z_0 contains no factor of λ. In fact, none of x, y, z_0 contain a factor of λ. The proof of this case will therefore follow from the following theorem.

Theorem 5.3.1. *Let p be an odd regular prime. Suppose we have a solution to $x^p+y^p+u\lambda^{pm}z_0^p = 0$ for elements $x, y, z_0 \in \mathbb{Z}[\zeta_p]$, a unit $u \in \mathbb{Z}[\zeta_p]^\times$ and an integer $m \geq 1$, so that x, y, z_0 are not divisible by $1 - \zeta_p$. Then $xyz_0 = 0$.*

Proof. Fix such a solution where m is minimal. Note that modulo λ we get

$$\zeta_p \equiv 1 \pmod{\lambda} \implies x + \zeta_p^k y \equiv x + y \pmod{\lambda}$$

for all k. Now λ must divide $x + \zeta_p^k y$ for some k so it will divide $x + \zeta_p^k y$ for all k. By similar computations as in case 1, we get the containment of ideals

$$(x, y)(\lambda) \subset (x + \zeta_p^k y, x + \zeta_p^j y) \subset (\lambda)$$

for $j \neq k$. But by assumption we have $(x, y) = (1) \not\subset (\lambda)$ and so λ divides all but one ideal of the form $(x + \zeta_p^k y)$ exactly once. We can assume that the exceptional ideal is $(x + y)$, since if it was $(x + \zeta_p^k y)$ instead, then we could just set $y = y\zeta_p^k$ in the above theorem. Now we pass to elements and write

$$\frac{x + y}{\lambda^{pm-(p-1)}} \prod_{k=1}^{p-1} \frac{x + \zeta_p^k y}{\lambda} + z_0^p = 0$$

As ideals, $\frac{(x+\zeta_p^k y)}{(\lambda)}$ are pairwise coprime and also coprime to $\frac{(x+y)}{(\lambda)^{pm-(p-1)}}$. As a result they are all p^{th} powers of principal ideals when p is regular. Write as ideals then pass to elements to get

$$\frac{(x + y)}{(\lambda)^{pm-(p-1)}} = (a_0)^p \implies \frac{x + y}{\lambda^{pm-(p-1)}} = u_0 a_0^p$$

$$\frac{(x + \zeta_p^k y)}{(\lambda)} = (a_k)^p \implies \frac{x + \zeta_p^k y}{\lambda} = u_k a_k^p$$

for elements $a_0 \dots a_{p-1} \in \mathbb{Z}[\zeta_p]$ and units $u_0 \dots u_{p-1} \in \mathbb{Z}[\zeta_p]^\times$. We have the simple relation $(x + \zeta_p y)(1 + \zeta_p) - (x + \zeta_p^2 y) = \zeta_p(x + y)$, into which we substitute the above forms to get

$$u_1 a_1^p \lambda(1 + \zeta_p) - u_2 a_2^p \lambda = \zeta_p u_0 \lambda^{pm-(p-1)} a_0^p$$

Then we rearrange a little to get

$$a_1^p - \frac{u_2}{u_1(1+\zeta_p)}a_2^p = \frac{\zeta_p u_0}{u_1(1+\zeta_p)}a_0^p\lambda^{p(m-1)}$$

we need to show that $m \geq 2$. Suppose that $m = 1$. Then λ would divide each of the $(x + \zeta_p^k y)$ exactly once. However, $[\mathbb{Z}[\zeta_p]/(\lambda)^2 : \mathbb{Z}[\zeta_p]/(\lambda)] = p - 1$ and so by pigeonhole principle

$$x + \zeta_p^k y \equiv x + \zeta_p^j y \pmod{\lambda^2}$$

for at least one set of distinct pairs j, k. From this we get

$$y(\zeta_p^k - \zeta_p^j) \equiv 0 \pmod{\lambda^2}$$

Now λ will divide $(\zeta_p^k - \zeta_p^j)$ once and so it must also divide y, which contradicts our assumptions.

Therefore $m \geq 2$ and so p will divide the RHS of $a_1^p - \frac{u_2}{u_1(1+\zeta_p)}a_2^p = \frac{\zeta_p u_0}{u_1(1+\zeta_p)}a_0^p\lambda^{p(m-1)}$ because $p(m-1) \geq p-1$. Both a_1 and a_2 are units modulo p which satisfy $a_1^p - \frac{u_2}{u_1(1+\zeta_p)}a_2^p \equiv 0 \pmod{p}$. We rearrange to get

$$\frac{u_2}{u_1(1+\zeta_p)} \equiv (\frac{a_1}{a_2})^p \pmod{p}$$

which is in turn congruent to some rational integer by previous results. Now we invoke Kummer's lemma to write $\frac{u_2}{u_1(1+\zeta_p)} = \eta^p$ for some unit η. We get

$$a_1^p + (-\eta a_2)^p + (-\frac{\zeta_p u_0 u^{1-m}}{u_1(1+\zeta_p)})a_0^p\lambda^{p(m-1)} = 0$$

contradicting the minimality of our previous solution, since $(-\frac{\zeta_p u_0 u^{1-m}}{u_1(1+\zeta_p)})$ is a unit and $a_1, a_0, (-\eta a_2)$ have no factor of λ. □

This completes Fermat's last theorem for regular primes $p \geq 5$.

5.4 Cases $p = 3$ and $p = 4$

In this short subsection we present proofs of Fermat's last theorem for the cases $p = 3$ and $p = 4$.

Theorem 5.4.1. *Suppose there are integers x, y, z so that $x^3 + y^3 = z^3$. Then $xyz = 0$.*

Proof. Suppose there is a non-trivial solution in integers x, y, z to $x^3 + y^3 = z^3$. Then 3 must divide one of x, y, z by looking modulo 9, since the only cubes are $\{0, 1, 8\}$. By rearranging the terms, this constitutes a solution to $x^3 + y^3 + (-z)^3 = 0$ where 3 divides z. In particular, by setting $z = 3^m z_0$ where $3 \nmid z_0$, we get a solution to $x^3 + y^3 + (1 - \zeta_3)^{6m}(-z_0)^3 = 0$ with $m \geq 1$. However this contradicts Theorem 5.3.1 because 3 is a regular prime. In fact $h(\mathbb{Z}[\zeta_3]) = 1$. □

The case $p = 4$ follows immediately from the following more general theorem.

Theorem 5.4.2. *Suppose there are integers x, y, z so that $x^4 + y^4 = z^2$. Then $xyz = 0$.*

Proof. This proof is drastically different from the other cases that we dealt with. We can prove this case by making use of the ring $\mathbb{Z}[i]$, but this would be far too complicated given that 4 is even. We can assume as usual that x, y, z are pairwise coprime and form a solution to the above equation. Note that (x^2, y^2, z) is a Pythagorean triple. We can assume W.L.O.G. that x is odd so that we can use the parametrization

$$x^2 = a^2 - b^2$$

$$y^2 = 2ab$$

$$z = a^2 + b^2$$

for integers a, b. Then (b, x, a) is a Pythagorean triple, and can be parametrized W.L.O.G. as

$$b = 2mn$$

$$x = m^2 - n^2$$

$$a = m^2 + n^2$$

for integers m, n, since x is odd. This gives

$$y^2 = 4mn(m^2 + n^2)$$

a and b must be coprime in order to ensure $(z, x) = 1$. m and n must also be coprime to ensure $(a, b) = 1$. As a result m, n and $m^2 + n^2$ have no common factors, and must all be perfect squares. This converts the equation $m^2 + n^2 = a$ into one of the form $f^4 + g^4 = h^2$ where $m = f^2, n = g^2$ and $a = h^2$. This makes (f, g, h) a nontrivial solution to the original equation where $h < a < z$. By infinite descent the solution must have been trivial to begin with. □

5.5 The relative class number formula for prime cyclotomic fields

In this subsection we derive the relative class number formula for prime cyclotomic fields. This is an explicit formula for the quotient $\frac{h(\mathbb{Q}(\zeta_p))}{h(\mathbb{Q}(\zeta_p)^+)}$ which turns out to be an integer. This quantity is known as the relative class number and is denoted by $h^-(\mathbb{Q}(\zeta_p))$.

The first thing we need to do is to prove the analytic continuation of Dirichlet L-series to the entire complex plane. For this we follow [IR90, Chapter 16.6]. In doing so, we will automatically get the value of Dirichlet L-series at non-positive integers. The proof is similar to Riemann's first proof of the analytic continuation of the Riemann zeta function, and makes use of the gamma function, which is defined as

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

Integration by parts yields the functional equation:

$$\Gamma(s+1) = s\Gamma(s)$$

which can be used to analytically continue Γ to a meromorphic function on $\mathbb{C}$, with simple poles at the non-positive integers.

We substitute nt into t in the above integral to get

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

The factor n^s is moved to the other side and used in a sum to create the L-series

$$\sum_{n=1}^{\infty} \chi(n)n^{-s}\Gamma(s) = L(s,\chi)\Gamma(s) = \int_0^\infty t^{s-1}\sum_{n=1}^{\infty}\chi(n)e^{-nt}dt$$

The characters in the sum $\sum_{n=1}^{\infty}\chi(n)e^{-nt}$ are periodic with period f_χ and so we can rewrite this sum as a sum of f_χ generating functions as:

$$\sum_{n=1}^{\infty}\chi(n)e^{-nt} = \sum_{n=1}^{f_\chi}\frac{\chi(n)e^{-nt}}{1-e^{-f_\chi t}}$$

Our aim is to ultimately integrate the RHS by parts, so that we may extend $L(s,\chi)$. However this is currently not possible because $\sum_{n=1}^{f_\chi}\frac{\chi(n)e^{-nt}}{1-e^{-f_\chi t}}$ cannot be evaluated at $t=0$. A simple trick is required. Substitute $2t$ into t to get

$$L(s,\chi)\Gamma(s) = \int_0^\infty (2t)^{s-1}\sum_{n=1}^{f_\chi}\frac{\chi(n)e^{-2nt}}{1-e^{-2f_\chi t}}2dt = 2^{s-1}\int_0^\infty t^{s-1}\sum_{n=1}^{f_\chi}\frac{2\chi(n)e^{-2nt}}{1-e^{-2f_\chi t}}dt$$

Now take away this equation from the original integral, after moving the 2^{s-1} to the other side, to get

$$(1-2^{1-s})L(s,\chi)\Gamma(s) = \int_0^\infty t^{s-1}\sum_{n=1}^{f_\chi}\left(\frac{\chi(n)e^{-nt}}{1-e^{-f_\chi t}} - \frac{2\chi(n)e^{-2nt}}{1-e^{-2f_\chi t}}\right)dt =$$

$$= \int_0^\infty t^{s-1}\sum_{n=1}^{f_\chi}\frac{(1+e^{-f_\chi t})\chi(n)e^{-nt} - 2\chi(n)e^{-2nt}}{1-e^{-2f_\chi t}}dt =$$

$$= \int_0^\infty t^{s-1}\sum_{n=1}^{f_\chi}\frac{\chi(n)e^{-nt}(1+e^{-f_\chi t} - 2\chi(n)e^{-nt})}{1-e^{-2f_\chi t}}dt = \int_0^\infty t^{s-1}\sum_{n=1}^{f_\chi}\frac{\chi(n)P_0(e^{-t})}{\sum_{k=0}^{2f_\chi-1}e^{-kt}}dt$$

for some integer polynomial P_0. Now the expression $\sum_{n=1}^{f_\chi}\frac{\chi(n)e^{-nt}P(e^{-t})}{\sum_{k=0}^{2f_\chi-1}e^{-kt}}$ clearly vanishes at $t=0$ and $t\to\infty$. Integrating by parts is now possible and it gives us

$$(1-2^{1-s})L(s,\chi)\Gamma(s) = -\frac{1}{s}\int_0^\infty t^s\sum_{n=1}^{f_\chi}\frac{\chi(n)P_1(e^{-t})}{(\sum_{k=0}^{2f_\chi-1}e^{-kt})^2}dt$$

for another integral polynomial P_1. For notational purposes define

$$R_k = \sum_{n=1}^{f_\chi} \frac{\chi(n) P_k(e^{-t})}{(\sum_{k=0}^{2f_\chi - 1} e^{-kt})^{k+1}}$$

for some integer polynomials P_k so that we may write

$$(1 - 2^{1-s}) L(s,\chi)\Gamma(s) = (-1)^k \frac{1}{\prod_{j=0}^{k-1}(s+j)} \int_0^\infty t^{s+k-1} R_k dt$$

after integrating by parts k times. Applying the gamma functional equation gives

$$(1 - 2^{1-s}) L(s,\chi)\Gamma(s+k) = (-1)^k \int_0^\infty t^{s+k-1} R_k dt$$

This will extend $L(s,\chi)$ to $\{s \in \mathbb{C} : Re(s) > -k\}$ since the integral converges for $s > -k$. By choosing k to be large enough we can analytically continue any Dirichlet L-series to the entire complex plane. As a result any Dedekind zeta function associated to abelian Galois extensions of $\mathbb{Q}$ can also be extended to the entire complex plane.

We are more interested in the values of L-series at non-positive integers. By setting $s = 1 - k$ in the integral involving R_k, we get

$$(1 - 2^k) L(1-k,\chi)\Gamma(1) = (-1)^k \int_0^\infty R_k dt$$

We will now relate R_0 to the generalized Bernoulli numbers of χ.

Definition 5.5.1. *Generalized Bernoulli numbers are defined as coefficients of the Taylor expansion*

$$F(t) = \sum_{n=1}^{f_\chi} \chi(n) \frac{te^{nt}}{e^{f_\chi t} - 1} = \sum_{k=0}^{\infty} \frac{B_{k,\chi}}{k!} t^k$$

Note that

$$F(-t) = \sum_{k=0}^{\infty} \frac{B_{k,\chi}}{k!} (-t)^k = \sum_{n=1}^{f_\chi} \chi(n) \frac{te^{-nt}}{1 - e^{-f_\chi t}}$$

so we can write:

$$R_0 = \sum_{n=1}^{f_\chi} \chi(n) \frac{e^{-2nt}}{1 - e^{-2f_\chi t}} - \sum_{n=1}^{f_\chi} \chi(n) \frac{2e^{-nt}}{1 - e^{-f_\chi t}} = \frac{1}{t}(F(-t) - F(-2t))$$

Finally we can relate R_0 to generalized Bernoulli numbers as

$$R_0 = \frac{1}{t}(\sum_{k=0}^{\infty} \frac{B_{k,\chi}}{k!}((-t)^k - (-2t)^k)) = \sum_{k=0}^{\infty} (-1)^k \frac{B_{k,\chi}}{k!} t^{k-1}(1 - 2^k)$$

It is clear that $R_k(t) = \frac{d^k R_0(t)}{dt^k}$ so by basic analysis we have

$$R_{k+1}(0) = (-1)^k \frac{B_{k,\chi}}{k}(1-2^k)$$

Then, because $R_k(t)$ vanishes as $t \to \infty$, we get

$$(1-2^k)L(1-k,\chi) = (-1)^k \int_0^\infty R_k dt = (-1)^{k+1} R_{k-1}(0) = (1-2^k)(-1)^{2k+1}\frac{B_{k,\chi}}{k}$$

The following theorem is proved as a result.

Theorem 5.5.2. *Let χ be a Dirichlet character and let k be a nonnegative integer. Then*

$$L(1-k,\chi) = -\frac{B_{k,\chi}}{k}$$

We need the following extra results which will not be proven here.

Theorem 5.5.3 (Legendre duplication formula)**.** *Let s be a complex number not equal to a negative integer. We have*

$$\Gamma(\frac{s}{2})\Gamma(\frac{s+1}{2}) = \Gamma(s)2^{1-s}\sqrt{\pi}$$

Proof. See [Chi, Theorem 3.24]. □

Theorem 5.5.4. *(Functional equation for Dedekind zeta functions) Let K be a number field and s a complex number. Then*

$$(2^{-r_2}\pi^{-\frac{[K:\mathbb{Q}]}{2}}\sqrt{|D_k|})^s\Gamma(\frac{s}{2})^{r_1}\Gamma(s)^{r_2}\zeta_K(s) = (2^{-r_2}\pi^{-\frac{[K:\mathbb{Q}]}{2}}\sqrt{|D_k|})^{1-s}\Gamma(\frac{1-s}{2})^{r_1}\Gamma(1-s)^{r_2}\zeta_K(1-s)$$

Theorem 5.5.5. *(Functional equation for Dirichlet L-series) Let χ be a Dirichlet character and set $a = 1$ when χ is odd and $a = 0$ when χ is even. Let s be a complex number. Then*

$$(\frac{\pi}{f_\chi})^{-\frac{(1-s)+a}{2}}\Gamma(\frac{(1-s)+a}{2})L(1-s,\bar{\chi}) = \frac{i^a\sqrt{f_\chi}}{\tau(\chi)}(\frac{\pi}{f_\chi})^{-\frac{s+a}{2}}\Gamma(\frac{s+a}{2})L(s,\chi)$$

where $\tau(\chi)$ is the standard Gauss sum associated to χ. See Section 6.2 for a definition of this.

For proofs of these functional equations, see Tate's thesis [Tat67].

We will make use of Proposition 4.5.4 to link these two functional equations together for an abelian number field K, following [Was97, Chapter 4]. Let X be the associated group of Dirichlet characters. Then K is either totally real or totally complex so we will separate the calculation over these two cases.

- Let K be totally real. Then $r_2 = 0$ and $r_1 = [K:\mathbb{Q}]$. Furthermore all associated characters are even and so $a = 0$. We take the product over all functional equations for the corresponding Dirichlet L-series to get

$$\zeta_K(1-s)\prod_{\chi\in X}(\frac{\pi}{f_\chi})^{-\frac{(1-s)}{2}}\Gamma(\frac{(1-s)}{2}) = \zeta_K(s)\prod_{\chi\in X}\frac{\sqrt{f_\chi}}{\tau(\chi)}(\frac{\pi}{f_\chi})^{-\frac{s}{2}}\Gamma(\frac{s}{2})$$

We compare with the functional equation for the Dedekind zeta which reads

$$(\pi^{-\frac{r_1}{2}}\sqrt{|D_k|})^s\Gamma(\frac{s}{2})^{r_1}\zeta_K(s) = (\pi^{-\frac{r_1}{2}}\sqrt{|D_k|})^{1-s}\Gamma(\frac{1-s}{2})^{r_1}\zeta_K(1-s)$$

We substitute the formula for $\zeta_K(s)$ from this functional equation into the previous one to get

$$\zeta_K(1-s)\prod_{\chi\in X}(\frac{\pi}{f_\chi})^{-\frac{(1-s)}{2}}\Gamma(\frac{(1-s)}{2}) = \frac{(\pi^{-\frac{r_1}{2}}\sqrt{|D_k|})^{1-s}\Gamma(\frac{1-s}{2})^{r_1}\zeta_K(1-s)}{(\pi^{-\frac{r_1}{2}}\sqrt{|D_k|})^s\Gamma(\frac{s}{2})^{r_1}}\prod_{\chi\in X}\frac{\sqrt{f_\chi}}{\tau(\chi)}(\frac{\pi}{f_\chi})^{-\frac{s}{2}}\Gamma(\frac{s}{2})$$

We first set $s=\frac{1}{2}$ to get

$$\zeta_K(\frac{1}{2})\prod_{\chi\in X}(\frac{\pi}{f_\chi})^{-\frac{1}{4}}\Gamma(\frac{1}{4}) = \frac{(\pi^{-\frac{r_1}{2}}\sqrt{|D_k|})^{\frac{1}{2}}\Gamma(\frac{1}{4})^{r_1}\zeta_K(\frac{1}{2})}{(\pi^{-\frac{r_1}{2}}\sqrt{|D_k|})^{\frac{1}{2}}\Gamma(\frac{1}{4})^{r_1}}\prod_{\chi\in X}\frac{\sqrt{f_\chi}}{\tau(\chi)}(\frac{\pi}{f_\chi})^{-\frac{1}{4}}\Gamma(\frac{1}{4})$$

Major cancellation occurs giving us

$$1 = \prod_{\chi\in X}\frac{\sqrt{f_\chi}}{\tau(\chi)}$$

Now we set $s=-\frac{1}{2}$ to get

$$\zeta_K(\frac{3}{2})\prod_{\chi\in X}(\frac{\pi}{f_\chi})^{-\frac{3}{4}}\Gamma(\frac{3}{4}) = \frac{(\pi^{-\frac{r_1}{2}}\sqrt{|D_k|})^{\frac{3}{2}}\Gamma(\frac{3}{4})^{r_1}\zeta_K(\frac{3}{2})}{(\pi^{-\frac{r_1}{2}}\sqrt{|D_k|})^{-\frac{1}{2}}\Gamma(\frac{-1}{4})^{r_1}}\prod_{\chi\in X}\frac{\sqrt{f_\chi}}{\tau(\chi)}(\frac{\pi}{f_\chi})^{\frac{1}{4}}\Gamma(\frac{-1}{4})$$

Recall that $|X|=r_1$. Major cancellation occurs giving us

$$\prod_{\chi\in X}(\frac{\pi}{f_\chi})^{-1} = \frac{(\pi^{-\frac{r_1}{2}}\sqrt{|D_k|})^{\frac{3}{2}}}{(\pi^{-\frac{r_1}{2}}\sqrt{|D_k|})^{-\frac{1}{2}}}\prod_{\chi\in X}\frac{\sqrt{f_\chi}}{\tau(\chi)}$$

Using the result from the $s=\frac{1}{2}$ substitution and rearranging a bit gives us

$$\prod_{\chi\in X} f_\chi = \pi^{r_1}(\pi^{-\frac{r_1}{2}}\sqrt{|D_k|})^2 = |D_k|$$

- Now let K be totally complex. Then $r_1=0$ and $2r_2=[K:\mathbb{Q}]$. Half the characters are even and half the characters are odd. Let X_0 be the set of even characters and X_1 the set of odd characters. Taking the product over all functional equations for the corresponding Dirichlet L-series gives

$$\zeta_K(1-s)\prod_{\chi\in X_0}(\frac{\pi}{f_\chi})^{-\frac{(1-s)}{2}}\Gamma(\frac{(1-s)}{2})\prod_{\chi\in X_1}(\frac{\pi}{f_\chi})^{-\frac{(2-s)}{2}}\Gamma(\frac{(2-s)}{2}) =$$
$$= \zeta_K(s)\prod_{\chi\in X_0}\frac{\sqrt{f_\chi}}{\tau(\chi)}(\frac{\pi}{f_\chi})^{-\frac{s}{2}}\Gamma(\frac{s}{2})\prod_{\chi\in X_1}\frac{i\sqrt{f_\chi}}{\tau(\chi)}(\frac{\pi}{f_\chi})^{-\frac{1+s}{2}}\Gamma(\frac{1+s}{2})$$

The functional equation for the Dedekind zeta function in this case reads

$$(2^{-r_2}\pi^{-r_2}\sqrt{|D_k|})^s\Gamma(s)^{r_2}\zeta_K(s) = (2^{-r_2}\pi^{-r_2}\sqrt{|D_k|})^{1-s}\Gamma(1-s)^{r_2}\zeta_K(1-s)$$

We substitute the formula for $\zeta_K(s)$ from this equation into the previous one to get

$$\zeta_K(1-s)\prod_{\chi\in X_0}(\frac{\pi}{f_\chi})^{-\frac{(1-s)}{2}}\Gamma(\frac{(1-s)}{2})\prod_{\chi\in X_1}(\frac{\pi}{f_\chi})^{-\frac{(2-s)}{2}}\Gamma(\frac{(2-s)}{2}) =$$

$$= \frac{(2^{-r_2}\pi^{-r_2}\sqrt{|D_k|})^{1-s}\Gamma(1-s)^{r_2}\zeta_K(1-s)}{(2^{-r_2}\pi^{-r_2}\sqrt{|D_k|})^s\Gamma(s)^{r_2}}\prod_{\chi\in X_0}\frac{\sqrt{f_\chi}}{\tau(\chi)}(\frac{\pi}{f_\chi})^{-\frac{s}{2}}\Gamma(\frac{s}{2})\prod_{\chi\in X_1}\frac{i\sqrt{f_\chi}}{\tau(\chi)}(\frac{\pi}{f_\chi})^{-\frac{1+s}{2}}\Gamma(\frac{1+s}{2})$$

Setting $s = \frac{1}{2}$ first will give us

$$\zeta_K(\frac{1}{2})\prod_{\chi\in X_0}(\frac{\pi}{f_\chi})^{-\frac{1}{4}}\Gamma(\frac{1}{4})\prod_{\chi\in X_1}(\frac{\pi}{f_\chi})^{-\frac{3}{4}}\Gamma(\frac{3}{4}) =$$

$$= \frac{(2^{-r_2}\pi^{-r_2}\sqrt{|D_k|})^{\frac{1}{2}}\Gamma(\frac{1}{2})^{r_2}\zeta_K(\frac{1}{2})}{(2^{-r_2}\pi^{-r_2}\sqrt{|D_k|})^{\frac{1}{2}}\Gamma(\frac{1}{2})^{r_2}}\prod_{\chi\in X_0}\frac{\sqrt{f_\chi}}{\tau(\chi)}(\frac{\pi}{f_\chi})^{-\frac{1}{4}}\Gamma(\frac{1}{4})\prod_{\chi\in X_1}\frac{i\sqrt{f_\chi}}{\tau(\chi)}(\frac{\pi}{f_\chi})^{-\frac{3}{4}}\Gamma(\frac{3}{4})$$

Major cancellation occurs giving us

$$1 = \prod_{\chi\in X_0}\frac{\sqrt{f_\chi}}{\tau(\chi)}\prod_{\chi\in X_1}\frac{i\sqrt{f_\chi}}{\tau(\chi)}$$

Now we can set $s = -\frac{1}{2}$ to give us

$$\zeta_K(\frac{3}{2})\prod_{\chi\in X_0}(\frac{\pi}{f_\chi})^{-\frac{3}{4}}\Gamma(\frac{3}{4})\prod_{\chi\in X_1}(\frac{\pi}{f_\chi})^{-\frac{5}{4}}\Gamma(\frac{5}{4}) =$$

$$= \frac{(2^{-r_2}\pi^{-r_2}\sqrt{|D_k|})^{\frac{3}{2}}\Gamma(\frac{3}{2})^{r_2}\zeta_K(\frac{3}{2})}{(2^{-r_2}\pi^{-r_2}\sqrt{|D_k|})^{-\frac{1}{2}}\Gamma(-\frac{1}{2})^{r_2}}\prod_{\chi\in X_0}\frac{\sqrt{f_\chi}}{\tau(\chi)}(\frac{\pi}{f_\chi})^{\frac{1}{4}}\Gamma(-\frac{1}{4})\prod_{\chi\in X_1}\frac{i\sqrt{f_\chi}}{\tau(\chi)}(\frac{\pi}{f_\chi})^{-\frac{1}{4}}\Gamma(\frac{1}{4})$$

We apply the result from the $s = \frac{1}{2}$ substitution. Some cancellation occurs giving us

$$\prod_{\chi\in X_0}(\frac{\pi}{f_\chi})^{-\frac{3}{4}}\Gamma(\frac{3}{4})\prod_{\chi\in X_1}(\frac{\pi}{f_\chi})^{-\frac{5}{4}}\Gamma(\frac{5}{4}) =$$

$$= \frac{(2^{-r_2}\pi^{-r_2}\sqrt{|D_k|})^{\frac{3}{2}}\Gamma(\frac{3}{2})^{r_2}}{(2^{-r_2}\pi^{-r_2}\sqrt{|D_k|})^{-\frac{1}{2}}\Gamma(-\frac{1}{2})^{r_2}}\prod_{\chi\in X_0}(\frac{\pi}{f_\chi})^{\frac{1}{4}}\Gamma(-\frac{1}{4})\prod_{\chi\in X_1}(\frac{\pi}{f_\chi})^{-\frac{1}{4}}\Gamma(\frac{1}{4})$$

We rearrange, keeping in mind that $|X_0| = |X_1| = r_2$, to get

$$\prod_{\chi\in X_0}\frac{f_\chi}{\pi}\prod_{\chi\in X_1}\frac{f_\chi}{\pi} = (2^{-r_2}\pi^{-r_2}\sqrt{|D_k|})^2\frac{\Gamma(\frac{3}{2})^{r_2}\Gamma(\frac{1}{4})^{r_2}\Gamma(-\frac{1}{4})^{r_2}}{\Gamma(-\frac{1}{2})^{r_2}\Gamma(\frac{3}{4})^{r_2}\Gamma(\frac{5}{4})^{r_2}}$$

Here we will make use of the Legendre duplication formula to cancel out the gamma factors. We make the substitutions $\Gamma(\frac{3}{2}) = \frac{\Gamma(\frac{3}{4})\Gamma(\frac{5}{4})}{2^{-\frac{1}{2}}\sqrt{\pi}}$ and $\Gamma(-\frac{1}{2}) = \frac{\Gamma(-\frac{1}{4})\Gamma(\frac{1}{4})}{2^{\frac{3}{2}}\sqrt{\pi}}$ into the above formula to get

$$\prod_{\chi \in X} f_\chi = \pi^{2r_2}(2^{-r_2}\pi^{-r_2}\sqrt{|D_k|})^2 2^{2r_2} = |D_K|$$

The above calculation proves the following crucial theorem.

Theorem 5.5.6 (Conductor-Discriminant formula)**.** *Let X be a group of Dirichlet characters and K its associated field. Then*

$$\prod_{\chi \in X} f_\chi = |D_K|$$

We also got the following along the way.

Proposition 5.5.7. *Let X be a group of Dirichlet characters. Let X_0 be the subset of even characters and let X_1 be the subset of odd characters. Then*

$$\prod_{\chi \in X_0} \frac{\sqrt{f_\chi}}{\tau(\chi)} \prod_{\chi \in X_1} \frac{i\sqrt{f_\chi}}{\tau(\chi)} = 1$$

Now we can begin working towards the relative class number formula.

Proposition 5.5.8. *Let $K = \mathbb{Q}(\zeta_p)$ be a cyclotomic field with p an odd prime. Then the ratio of its regulator with the regulator of the maximally real subfield is*

$$\frac{\mathcal{R}_K}{\mathcal{R}_{K^+}} = 2^{\frac{p-3}{2}}$$

Proof. We have previously shown in Proposition 5.1.6 that $\mathcal{O}_K^\times = \langle \zeta_p \rangle \mathcal{O}_{K^+}^\times$, but the torsion subgroup $\langle \zeta_p \rangle$ is killed in the *Log*-embedding. As a result we can use the same set of fundamental units for both fields. However, in the regulator matrix for K, a coefficient of 2 is added to the logarithm attached to each complex embedding, of which there are $\frac{p-3}{2}$. The result follows. □

We look at the analytic class number formula for $K = \mathbb{Q}(\zeta_p)$, whose associated group of Dirichlet characters minus the trivial character is X_0. It says that

$$\lim_{s \to 1}((s-1)\zeta_K(s)) = \prod_{\chi \in X_0} L(1, \chi) = \frac{(2\pi)^{\frac{p-1}{2}} \mathcal{R}_K h_K}{2p\sqrt{|D_K|}}$$

We compare it with the analytic class number formula for the maximally real subfield K^+, whose associated group of Dirichlet characters minus the trivial character is X_0^+. It says that

$$\lim_{s \to 1}((s-1)\zeta_{K^+}(s)) = \prod_{\chi \in X_0^+} L(1, \chi) = \frac{2^{\frac{p-1}{2}} \mathcal{R}_{K^+} h_{K^+}}{2\sqrt{|D_{K^+}|}}$$

We note that X_0^+ just consists of the even characters of X_0 minus the trivial character. We can therefore take the quotient of the analytic class number formulae and write

$$\prod_{\chi odd} L(1,\chi) = \frac{(2\pi)^{\frac{p-1}{2}} \mathcal{R}_K h_K 2\sqrt{|D_{K^+}|}}{2^{\frac{p-1}{2}} \mathcal{R}_{K^+} h_{K^+} 2p\sqrt{|D_K|}} = \frac{\pi^{\frac{p-1}{2}} 2^{\frac{p-3}{2}} h_K \sqrt{|D_{K^+}|}}{h_{K^+} p\sqrt{|D_K|}}$$

where we applied Proposition 5.5.8 in the last step. From Theorem 5.5.6 we get that

$$|D_{K^+}| = \prod_{\chi even} f_\chi = p^{\frac{p-3}{2}}$$

due to the fact that the group of even characters has size $\frac{p-1}{2}$ and each is of conductor p, except the trivial character which has conductor 1. We also know that $|D_K| = p^{p-2}$. After rearranging, the formula for the relative class number becomes

$$h_K^- := \frac{h_K}{h_{K^+}} = \frac{\sqrt{p^{\frac{p+3}{2}}} \prod_{\chi odd} L(1,\chi)}{\pi^{\frac{p-1}{2}} 2^{\frac{p-3}{2}}}$$

Applying the functional equation for odd Dirichlet characters at $s = 1$ gives us

$$(\frac{\pi}{f_\chi})^{-\frac{1}{2}} \Gamma(\frac{1}{2}) L(0,\bar{\chi}) = \frac{i\sqrt{f_\chi}}{\tau(\chi)} (\frac{\pi}{f_\chi})^{-1} \Gamma(1) L(1,\chi)$$

We know that $\Gamma(\frac{1}{2}) = \sqrt{\pi}$, $\Gamma(1) = 1$ and $f_\chi = p$. Then taking the product across all odd characters gives

$$\prod_{\chi odd} L(1,\chi) = \prod_{\chi odd} L(0,\bar{\chi}) \frac{\pi}{p} \frac{\tau(\chi)}{i} = \prod_{\chi odd} L(0,\bar{\chi}) \frac{\pi}{\sqrt{p}} = (\frac{\pi}{\sqrt{p}})^{\frac{p-1}{2}} \prod_{\chi odd} L(0,\chi)$$

The conjugate of an odd character is an odd character, which explains the last equality. The penultimate equality follows from Lemma 5.5.7.

The relative class number formula in terms of L-series valued at 0 becomes

$$h_K^- = \frac{p \prod_{\chi odd} L(0,\chi)}{2^{\frac{p-3}{2}}}$$

The value of L-series at 0 can be computed using the generalized Bernoulli numbers from Theorem 5.5.2. This updates our formula to the relative class number formula

$$h_K^- = 2p \prod_{\chi odd} -\frac{1}{2} B_{1,\chi}$$

Why do we actually care about the relative class number formula? We will show later that p is regular if and only if it does not divide $h^-_{\mathbb{Q}(\zeta_p)}$. Therefore the existence of p-torsion is solely determined by the relative class number. For now, let's show that h_{K^+} divides h_K, so that h_K^- is a positive integer. This follows immediately from the following result in [Was97, Proposition 4.11].

Proposition 5.5.9. *Let $L : K$ be an extension of number fields containing no intermediate abelian unramified extensions of K. Then $h_K | h_L$.*

Proof. Let H_K be the Hilbert class field of K. Let $H_K \cap L = M$. Then $M : K$ is unramified because M is contained in the Hilbert class field. By the assumptions of the problem, we must have $M = K$. $H_K : K$ is unramified and so its relative discriminant is 1. As a result we can apply Theorem 1.8.14 on the compositum $H_K L$, because $H_K \cap L = K$ and the relative discriminants of $H_K : K$ and $L : K$ are coprime. This tells us that the relative discriminant of $H_K L : L$ is also 1 so it is an unramified extension. It is also abelian because its Galois group is isomorphic to $\mathrm{Gal}(H_K : K)$.

$H_K L : L$ is unramified abelian so it is contained in the Hilbert class field of L. This means $[H_K L : L]$ divides $[H_L : L]$. As a result we have $h_K | h_L$. □

6 More arithmetic of cyclotomic fields

In this section we give two presentations for a refined criterion on whether p divides the relative class number of $\mathbb{Q}(\zeta_p)$. One is analytic, making use of p-adic L-functions, and one is algebraic, making use of Stickelberger's relation and Herbrand's theorem. We start with the analytic presentation. The algebraic presentation will take us further and we will prove Kummer's criterion as well. Altogether this gives a relatively easy method to determine when a prime is regular.

The main sources for this section are [Was97] and [IR90]. We will clarify which sources are used at the beginning of each subsection.

6.1 Construction of p-adic L-functions

We begin by giving an alternate derivation for the values of ordinary Dirichlet L-functions at non-positive integers, which will help us motivate the definition of p-adic L-functions.

Definition 6.1.1. *The Hurwitz zeta function is defined as*

$$\zeta(s,b) = \sum_{n=0}^{\infty} \frac{1}{(n+b)^s}$$

for some rational b in the range $0 \leq b < 1$.

We can patch up Hurwitz zeta functions to form Dirichlet L-series via

$$\sum_{n=0}^{\infty} \frac{1}{(xn+y)^s} = x^{-s} \sum_{n=0}^{\infty} \frac{1}{(n+\frac{y}{x})^s} = x^{-s}\zeta(s, \frac{y}{x})$$

$$L(s,\chi) = \sum_{a=1}^{f_\chi} \sum_{n=0}^{\infty} \frac{\chi(a)}{(f_\chi n + a)^s} = \sum_{a=1}^{f_\chi} \chi(a) f_\chi^{-s} \zeta(s, \frac{a}{f_\chi})$$

Definition 6.1.2. *The Bernoulli polynomials $B_k(x)$ are defined from the Taylor expansion*

$$\frac{te^{xt}}{e^t - 1} = \sum_{k=0}^{\infty} \frac{B_k(x)}{k!} t^k$$

From the above definition we can rewrite generalized Bernoulli numbers using

$$\sum_{a=1}^{f_\chi} \chi(a) \frac{te^{at}}{e^{f_\chi t} - 1} = \sum_{a=1}^{f_\chi} \chi(a) f_\chi^{-1} \sum_{k=0}^{\infty} \frac{B_k(\frac{a}{f_\chi})}{k!} (f_\chi t)^k$$

$$B_{k,\chi} = \sum_{a=1}^{f_\chi} \chi(a) f_\chi^{k-1} B_k(\frac{a}{f_\chi})$$

So another way to derive the values of L-functions at non-positive integers is to derive the values of the Hurwitz zeta function at non-positive integers. We should get

$$\zeta(1-k,\frac{y}{x}) = -\frac{B_k(\frac{y}{x})}{k}$$

for x, y positive integers so that $0 \leq y < x$. We will prove this now using the same idea we used in the proof for L-series. We first need the simple result $B_k(1-m) = (-1)^k B_k(m)$. This is due to

$$\sum_{k=0}^{\infty} \frac{B_k(x)}{k!}(-t)^k = \frac{-te^{-xt}}{e^{-t}-1} = \frac{te^{(1-x)t}}{e^t-1} = \sum_{k=0}^{\infty} \frac{B_k(x)}{k!}t^k = \sum_{k=0}^{\infty} \frac{B_k(1-x)}{k!}t^k$$

Proposition 6.1.3. *The Hurwitz zeta function $\zeta(s,\frac{y}{x})$ for x, y positive integers with $0 \leq y < x$ can be analytically extended to $\mathbb{C}$ so that its value at $s = 1-k$ for positive integers k reads*

$$\zeta(1-k,\frac{y}{x}) = -\frac{B_k(\frac{y}{x})}{k}$$

Proof. We start with the gamma function defined as the integral

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

We substitute $(n+\frac{y}{x})t$ into t to get

$$\Gamma(s) = (n+\frac{y}{x})^s \int_0^{\infty} e^{-(n+\frac{y}{x})t}t^{s-1}dt$$

We move the $(n+\frac{y}{x})^s$ to the other side and sum over nonnegative integers n to get

$$\zeta(s,\frac{y}{x})\Gamma(s) = \int_0^{\infty} \sum_{n=0}^{\infty} e^{-(n+\frac{y}{x})t}t^{s-1}dt = \int_0^{\infty} \frac{e^{-\frac{y}{x}t}}{1-e^{-t}}t^{s-1}dt$$

We substitute xt into t to make all exponents integers, giving us

$$\zeta(s,\frac{y}{x})\Gamma(s) = x^s \int_0^{\infty} \frac{e^{-yt}}{1-e^{-xt}}t^{s-1}dt$$

Now we can do the same trick as before to get

$$\begin{aligned}(1-2^{1-s})x^{-s}\zeta(s,\frac{y}{x})\Gamma(s) &= \int_0^{\infty}(\frac{e^{-yt}}{1-e^{-xt}} - \frac{2e^{-2yt}}{1-e^{-2xt}})t^{s-1}dt = \\ &= \int_0^{\infty} \frac{e^{-yt}P_0(e^{-t})}{1+e^{-xt}}t^{s-1}dt = \int_0^{\infty} R_0 t^{s-1}dt\end{aligned}$$

for some integer polynomial P_0. Now integrating by parts will extend the Hurwitz zeta function to give

$$(1-2^{1-s})x^{-s}\zeta(s,\frac{y}{x})\Gamma(s+k) = (-1)^k \int_0^{\infty} R_k t^{s+k-1}dt$$

Setting $s = 1 - k$ above just as before gives us

$$(1-2^k)x^{k-1}\zeta(1-k,\frac{y}{x}) = (-1)^k \int_0^\infty R_k dt = (-1)^{k+1} R_{k-1}(0)$$

We need to know what R_0 is in order to compute the above. We have

$$R_0 = \frac{e^{-yt}}{1-e^{-xt}} - \frac{2e^{-2yt}}{1-e^{-2xt}} = \frac{1}{xt}(\frac{xte^{(x-y)t}}{e^{xt}-1} - \frac{2xte^{2(x-y)t}}{e^{2xt}-1}) =$$
$$= \frac{1}{xt}(\sum_{k=0}^{\infty}(1-2^k)\frac{B_k(1-\frac{y}{x})}{k!}(xt)^k)$$

giving us the desired result

$$R_{k-1}(0) = (1-2^k)\frac{B_k(1-\frac{y}{x})}{k}x^k = (1-2^k)(-1)^k\frac{B_k(\frac{y}{x})}{k}x^{k-1}$$

Altogether this gives us

$$\zeta(1-k,\frac{y}{x}) = -\frac{B_k(\frac{y}{x})}{k}$$

□

The main idea for creating p-adic L-series is similar. We will define some function analogous to the Hurwitz zeta function and patch them up to define p-adic L-series. We do so in a way so that the special values (at non-positive integers) are related to the special values of the ordinary L-series. This can be seen as an interpolation of the special values. This definition of p-adic L-series is called the Kubota–Leopoldt p-adic L-series. We need some preliminaries before we get started.

Lemma 6.1.4. *For some prime number p, we have*

$$\sum_{a=1}^{p} a^n \equiv 0 \pmod{p}$$

in the case that $p-1$ does not divide n. Otherwise we have

$$\sum_{a=1}^{p} a^n \equiv -1 \pmod{p}$$

Proof. The second case is easy. If $p-1$ divides n then

$$\sum_{a=1}^{p} a^n \equiv \sum_{a=1}^{p} 1 \equiv p-1 \pmod{p}$$

Otherwise, $a^n \not\equiv 1$ for at least one value of a. Let $a^n = b$ for that value. The n^{th} powers modulo p form a multiplicative subgroup which is not trivial as a result. Multiplication by b will permute its elements so we get

$$b\sum_{a=1}^{p} a^n \equiv \sum_{a=1}^{p} a^n \implies (b-1)\sum_{a=1}^{p} a^n \equiv 0 \pmod{p}$$

Since $b \not\equiv 1 \pmod{p}$, we must get $\sum_{a=1}^{p} a^n \equiv 0 \pmod{p}$. □

For the remainder of this subsection we follow [Was97, Chapter 5].

Theorem 6.1.5 (von Staudt-Clausen)**.** *Let n be an even positive integer. Then*

$$B_n + \sum_{p-1|n} \frac{1}{p} \in \mathbb{Z}$$

Proof. First notice that the standard even Bernoulli numbers agree with the generalized even Bernoulli numbers for the trivial character $\chi = 1$, since

$$\frac{t}{e^t - 1} + t = \frac{te^t}{e^t - 1}$$

and so $B_n = B_{n,1}$ for positive even integers n. Recall the formula for generalized Bernoulli numbers in terms of Bernoulli polynomials and apply it to $\chi = 1$ to get

$$B_{n,1} = B_n(1)$$

We may write

$$\sum_{n=0}^{\infty} \frac{B_{n,1}}{n!} t^n = \frac{te^t}{e^{pt} - 1} \frac{e^{pt} - 1}{e^t - 1} = \sum_{a=1}^{p} \frac{te^{at}}{e^{pt} - 1} = \frac{1}{p} \sum_{n=0}^{\infty} \sum_{a=1}^{p} \frac{B_n(\frac{a}{p})}{n!} (pt)^n$$

by using definitions of generalized Bernoulli numbers and Bernoulli polynomials. Then reading off the coefficients for even positive integers n gives

$$B_n = B_{n,1} = p^{n-1} \sum_{a=1}^{p} B_n(\frac{a}{p})$$

Another way to relate Bernoulli polynomials and Bernoulli numbers is to write

$$\sum_{k=0}^{\infty} \frac{B_k(x)}{k!} t^k = \frac{te^{xt}}{e^t - 1} = \sum_{k=0}^{\infty} \frac{B_k}{k!} t^k \sum_{m=0}^{\infty} \frac{(xt)^m}{m!}$$

When we read off the coefficients to t^k we get:

$$\frac{B_k(x)}{k!} = \sum_{i=0}^{k} \frac{B_i x^{k-i}}{i!(k-i)!}$$

giving us the expression for the Bernoulli polynomials in terms of Bernoulli numbers

$$B_k(x) = \sum_{i=0}^{k} \binom{k}{i} B_i x^{k-i}$$

Substituting this into our formula for Bernoulli numbers in terms of Bernoulli polynomials gives:

$$B_n = p^{n-1} \sum_{a=1}^{p} \sum_{i=0}^{n} \binom{n}{i} (B_i)(\frac{a}{p})^{n-i} = \sum_{a=1}^{p} \sum_{i=0}^{n} \binom{n}{i} B_i a^{n-i} p^{i-1}$$

We will now commence proof by induction. Suppose the result is true for all even Bernoulli numbers up to but excluding n. Take the above expression modulo p to get

$$B_n \equiv \sum_{a=1}^{p} B_0 a^n p^{-1} + nB_1 a^{n-1} + B_n p^{n-1}$$

because for all m positive even with $m < n$ we have, looking modulo p, that either $B_m \in \mathbb{Z}_p$ or $B_m + \frac{1}{p} \in \mathbb{Z}_p$. This means that pB_m is certainly p-integral. Now $B_0 = 1$ and $B_1 = -\frac{1}{2}$ so

$$B_n \equiv \sum_{a=1}^{p} a^n p^{-1} - \frac{na^{n-1}}{2} + B_n p^{n-1}$$

It follows that

$$(1-p^n)B_n = \frac{1}{p}\sum_{a=1}^{p} a^n - \frac{n}{2}\sum_{a=1}^{p} a^{n-1}$$

Invoking Lemma 6.1.4 gives, if $p-1$ does not divide n, that $(1-p^n)B_n$ is a p-adic integer. Even if $p = 2$, n is even so $\frac{n}{2}$ is a 2-adic integer. This means that B_n is a p-adic integer as $1 - p^n \equiv 1 \pmod{p}$.

However, if $p-1$ divides n, then we get

$$(1-p^n)B_n + \frac{1}{p} \equiv \frac{1}{p}(1 + \sum_{a=1}^{p} a^n) - \frac{n}{2}\sum_{a=1}^{p} a^{n-1}$$

The right hand side is a p-adic integer. In particular, $B_n + \frac{1}{p}$ is a p-adic integer. It follows that the sum $B_n + \sum_{p-1|n} \frac{1}{p} \in \mathbb{Z}$ is a p-adic integer for all primes p and hence it is an integer. □

The above theorem proves that $|B_n|_p \leq p$ and will be essential in our construction of p-adic L-functions. We will give the construction now, and then develop the p-adic analysis necessary in order to prove that p-adic L-functions converge where we want them to.

Until now we have only discussed Dirichlet characters, which take complex values, specifically some root of unity. By Hensel's lemma, for every non-zero equivalence class modulo a prime p, we get a corresponding $(p-1)^{\text{th}}$ root of unity in $\mathbb{Z}_p$. Therefore we can also define characters which take values of roots of unity in p-adic rings.

Definition 6.1.6. *The Teichmuller character is the group homomorphism*

$$\omega : (\mathbb{Z}/p\mathbb{Z})^\times \to \mathbb{Z}_p^\times$$

so that $\omega(a) \equiv a \pmod{p}$ *for all* $a \in (\mathbb{Z}/p\mathbb{Z})^\times$ *and* $\omega(0) = 0$. *This property uniquely characterizes* ω.

Due to the above definition we can normalize any p-adic unit so that it has value 1 modulo p. Let a be a p-adic unit. We define

$$\langle a \rangle = \omega^{-1}(a)a$$

This will be a p-adic unit which has value 1 modulo p.

Definition 6.1.7. *The p-adic Hurwitz zeta function is defined as*

$$H_p(s,a,F) = \frac{1}{s-1}\frac{1}{F}\langle a\rangle^{1-s}\sum_{k=0}^{\infty}\binom{1-s}{k}(B_k)\frac{F^k}{a^k}$$

for some integers a and F so that $p \not| a$ and $p|F$. If $p=2$, we require 4 to divide F.

We will show that the non-positive integer values s of the p-adic Hurwitz zeta function are related to the ordinary Hurwitz zeta function. Substituting $s = 1-n$ for some integer $n \geq 1$ gives

$$H_p(1-n,a,F) = -\frac{1}{nF}\langle a\rangle^n\sum_{k=0}^{n}\binom{n}{k}(B_k)(\frac{F}{a})^k = -\frac{F^{n-1}}{n}\omega^{-n}(a)\sum_{k=0}^{n}\binom{n}{k}(B_k)(\frac{a}{F})^{n-k}$$

Now the sum gives the n^{th} Bernoulli polynomial evaluated at $\frac{a}{F}$ and so

$$H_p(1-n,a,F) = -\frac{F^{n-1}\omega^{-n}(a)}{n}B_n(\frac{a}{F}) = F^{n-1}\omega^{-n}(a)\zeta(1-n,\frac{a}{F})$$

ω has order $p-1$, so in particular if $p-1|n$ then

$$H_p(1-n,a,F) = F^{n-1}\zeta(1-n,\frac{a}{F})$$

We now need some analytic results.

Lemma 6.1.8. *We have the following bounds, for a prime p and an integer n.*

$$\frac{n-1}{p-1} \leq v_p(n!) \leq \frac{n}{p-1}$$

Proof. We start with the result

$$v_p(n!) = \sum_{k=1}^{\infty}\lfloor\frac{n}{p^k}\rfloor$$

Then write $n = \sum_{k=0}^{m} a_k p^k$ so that

$$v_p(n!) = \sum_{k=1}^{m}\sum_{k=1}^{\infty}\lfloor\frac{\sum_{k=0}^{m}a_k p^k}{p^k}\rfloor = \sum_{k=1}^{m}\sum_{i=0}^{k-1}a_k p^i = \sum_{k=1}^{m}a_k\frac{1-p^k}{1-p}$$

The coefficients a_i can be chosen so that $0 \leq a_i < p$ and thus

$$v_p(n!) = \frac{\sum_{k=1}^{m}a_k(1-p^k)}{1-p} = \frac{(\sum_{k=1}^{m}a_k) - p\lfloor\frac{n}{p}\rfloor)}{1-p}$$

It is easy to check the bounds now. We get

$$\frac{n-1}{p-1} \leq v_p(n!) = \frac{(\sum_{k=1}^{m}a_k) - p\lfloor\frac{n}{p}\rfloor)}{1-p} \leq \frac{m(p-1)-(n-(p-1))}{1-p} = \frac{n}{p-1}$$

□

Proposition 6.1.9. *Consider a p-adic function written as*

$$f(x) = \sum_{n=0}^{\infty} a_n \binom{x}{n}$$

where the a_n are constants that decrease exponentially w.r.t. the p-adic metric as

$$|a_n|_p \leq Mr^n$$

for some positive real constant M and $r < p^{-\frac{1}{p-1}}$. Then f may be expressed as a power series with radius of convergence at least $R = \frac{p^{-\frac{1}{p-1}}}{r} > 1$.

Proof. We define the partial sums

$$P_i(x) = \sum_{n \leq i} a_n \binom{x}{n} = \sum_{n \leq i} a_{n,i} x^n$$

for some constants $a_{n,i}$, since each $\binom{x}{n}$ is a polynomial of degree n. We may write

$$a_{n,i} = \sum_{k=0}^{i-n} a_{n+k} \frac{c(i,n,k)}{(n+k)!}$$

where $c(i,n,k)$ are p-integral and so $|c(i,n,k)|_p \leq 1$. As a result

$$|a_{n,i}|_p \leq \max_{k=0\ldots i-n} \{|\frac{a_{n+k}}{(n+k)!}|_p\}$$

Recall that $|(n+k)!|_p \geq p^{\frac{n+k}{1-p}}$ and $|a_{n+k}|_p \leq Mr^{n+k}$ and so

$$|\frac{a_{n+k}}{(n+k)!}|_p \leq Mr^{n+k} p^{\frac{n+k}{1-p}} = M(\frac{p^{-\frac{1}{p-1}}}{r})^{-(n+k)} \leq MR^{-n}$$

It follows that $|a_{n,i}|_p \leq MR^{-n}$. Now we compare coefficients across the partial sums to get

$$P_{i+k}(x) - P_i(x) = \sum_{n=i+1}^{i+k} a_n \binom{x}{n}$$

$$|a_{n,i} - a_{n,i+k}|_p = |\sum_{j=1}^{k} a_{i+j} \frac{d(i,n,j,k)}{(i+j)!}|_p$$

where the $d(i,n,j,k)$ are p-integral. Similarly to above, we get

$$|a_{n,i} - a_{n,i+k}|_p \leq MR^{-(i+1)}$$

which converges to 0 as i increases. Therefore the sequences $\{a_{n,i}\}_i$ are Cauchy. The limit

$$a_{n,0} = \lim_{i \to \infty} (a_{n,i})$$

exists for each n because $\mathbb{Q}_p$ is complete. Since $|a_{n,i}|_p \leq MR^{-n}$ for each element of the sequence, we get that $|a_{n,0}|_p \leq MR^{-n}$ also. Therefore the function

$$P_0(x) = \sum_{n=0}^{\infty} a_{n,0}x^n$$

Will clearly converge for $\{x \in \mathbb{Q}_p : |x|_p < R\}$. It remains to show that $P_0(x) = f(x)$.

We know that $f(x) = \lim_{i\to\infty}(P_i(x))$. We've shown that

$$|P_0(x) - P_i(x)|_p \leq \max(|(a_{n,i} - a_{n,0})x^n|_p) \leq \max\{MR^{-(n)}|x^n|_p\}$$

The right hand side converges to 0 as n grows, for x in our specified range, and so eventually for large enough i we can instead use the bound

$$|P_0(x) - P_i(x)|_p \leq \max\{|(a_{n,i} - a_{n,0})x^n|_p\} \leq \max\{MR^{-(i+1)}|x^n|_p\}$$

so that P_i converges uniformly to P_0 as the right hand side goes to zero for large enough i. It follows that $f = P_0$. □

Proposition 6.1.10. *The p-adic Hurwitz zeta function is analytic on* $\{s \in \mathbb{Q}_p : |s|_p < qp^{-\frac{1}{p-1}}\}$ *except for a simple pole at* $s = 1$. *Where we define* $q = p$ *for odd* p *and* $q = 4$ *for* $p = 2$.

Proof. We require $q|F$ and $p \nmid a$ and so as a result we can say $|(F/a)^j|_p \leq q^{-j}$. Together with our result from von-Staudt Clausen we can say that $|B_j(F/a)^j|_p \leq pq^{-j}$.

We can therefore invoke Proposition 6.1.9 on

$$\sum_{j=0}^{\infty} \binom{s}{j}(B_j)(F/a)^j$$

as a function in s, setting $r = q^{-1}$. This gives us that the function is analytic on $D = \{s \in \mathbb{Q}_p : |s|_p < qp^{-\frac{1}{p-1}}\}$ since $q^{-1} < p^{-\frac{1}{p-1}}$. The function

$$\sum_{j=0}^{\infty} \binom{1-s}{j}(B_j)(F/a)^j$$

is also analytic on D, because 1 is a unit and the absolute value is non-archimedean. The function $\langle a\rangle^s = a^s\omega^{-s}(a)$ is also analytic on D because a^s is, since

$$a^s = \exp(s\log_p(a))$$

and by standard convergence facts about the p-adic logarithm and exponential, this is analytic on D. By the same reasoning we get that $\langle a\rangle^{1-s}$ is analytic on D. It follows finally that

$$\frac{1}{F}\langle a\rangle^{1-s}\sum_{j=0}^{\infty}\binom{1-s}{j}(B_j)(\frac{F}{a})^j = (s-1)H_p(s,a,F)$$

is analytic on D. Therefore $H_p(s,a,F)$ has a simple pole at $s=1$ with residue

$$\mathrm{res}_{s=1}(H_p(s,a,F)) = \frac{1}{F}\langle a\rangle^0 \sum_{j=0}^{\infty}\binom{0}{j}(B_j)(\frac{F}{a})^j = \frac{1}{F}$$

□

We are now ready to define p-adic L-functions.

Definition 6.1.11. *The p-adic L-function attached to a Dirichlet character χ of conductor f is defined as*

$$L_p(s,\chi) = \sum_{1\le a\le F:(a,p)=1}\chi(a)H_p(s,a,F)$$

for some F divisible by both q and f.

Theorem 6.1.12. *$L_p(s,\chi)$ as defined above is analytic on $D=\{s\in\mathbb{Q}_p : |s|_p < qp^{-\frac{1}{p-1}}\}$ except for a simple pole at $s=1$ with residue $1-p^{-1}$ when χ is trivial. It takes the value*

$$L_p(1-n,\chi) = -(1-\chi\omega^{-n}(p)p^{n-1})\frac{B_{n,\chi\omega^{-n}}}{n}$$

for positive integers n.

Proof. The analytic property follows from Proposition 6.1.10. The residue of the pole at $s=1$ is

$$\mathrm{res}_{s=1}(L_p(s,\chi)) = \sum_{1\le a\le F:(a,p)=1}\frac{\chi(a)}{F} = \frac{1}{F}\sum_{a=1}^{F}\chi(a) - \frac{1}{F}\sum_{b=1}^{F/p}\chi(pb)$$

When χ is trivial the above sum equals $1-\frac{1}{F}\frac{F}{p} = 1-p^{-1}$. If not then

$$\frac{1}{F}\sum_{a=1}^{F}\chi(a) - \frac{1}{F}\sum_{b=1}^{F/p}\chi(pb) = -\frac{\chi(p)}{F}\sum_{b=1}^{F/p}\chi(b)$$

Since p divides F, it must either divide f otherwise f will divide F/p. In the first case we have $\chi(p)=0$, and in the second case we have $\sum_{b=1}^{F/p}\chi(b)=0$, so if χ is non-trivial we get that $L_p(s,\chi)$ has no pole at $s=1$.

For a positive integer n we have

$$L_p(1-n,\chi) = \sum_{1\le a\le F:(a,p)=1}\chi(a)H_p(1-n,a,F) = -\sum_{1\le a\le F:(a,p)=1}\chi(a)\frac{F^{n-1}\omega^{-n}(a)}{n}B_n(\frac{a}{F})$$

We can write

$$L_p(1-n,\chi) = -\frac{F^{n-1}}{n}\sum_{1\le a\le F:(a,p)=1}\chi\omega^{-n}(a)B_n(\frac{a}{F})$$

$$L_p(1-n,\chi) = -\frac{F^{n-1}}{n}\sum_{a=1}^{F}\chi\omega^{-n}(a)B_n(\frac{a}{F}) + -\frac{F^{n-1}}{n}\sum_{b=1}^{F/p}\chi\omega^{-n}(bp)B_n(\frac{b}{F/p})$$

since $\omega^{-n}(a) \neq 0$ for all a in the sum. By our results on generalized Bernoulli numbers this gives

$$L_p(1-n,\chi) = -\frac{1}{n}(1-\chi\omega^{-n}(p)p^{n-1})B_{n,\chi\omega^{-n}}$$

□

Recall that when χ is non-trivial, the p-adic L-function is analytic on D. We can therefore obtain its expansion about $s = 1$.

Proposition 6.1.13. *Suppose we have the expansion*

$$L_p(s,\chi) = \sum_{i=0}^{\infty} a_i(s-1)^i$$

for χ non-trivial so that pq does not divide its conductor. Then $|a_i|_p < 1$ for $i > 0$ and $|a_0| \leq 1$.

Proof. Recall that $q = p$ when p odd and $q = 4$ when $p = 2$. In the definition of $L_p(s,\chi)$ choose F so that q divides F but pq does not, which can be done by assumption. Then we have

$$|\frac{B_jF^{j-1}}{j!a^j}|_p \leq \frac{p^{\frac{j}{p-1}}\cdot p}{q^{j-1}} = \frac{p^{\frac{p+j-1}{p-1}}}{q^{j-1}} = p^{-\frac{p(j-2)-2(j-1)}{p-1}}$$

for $j \geq 1$ and p odd. For $j \geq 6$, we automatically get that the right hand side is less that p^{-1}. For $p = 2$ we can check that

$$|\frac{B_jF^{j-1}}{j!a^j}|_p \leq \frac{p^{\frac{j}{p-1}}\cdot p}{q^{j-1}} = \frac{p^{\frac{p+j-1}{p-1}}}{q^{j-1}} = p^{-\frac{p(2j-3)-3(j-1)}{p-1}} \leq q^{-1}$$

for $j \geq 6$. Overall, for $j \geq 6$ we get that the right hand side is less that or equal to q^{-1}. We check by hand the cases $j = 3, 4, 5$.

$$|\frac{B_3F^2}{3!a^3}|_p = |\frac{B_5F^4}{5!a^5}|_p = |0|_p = 0 \leq q^{-1}$$

$$|\frac{B_4F^3}{4!a^4}|_p = |-\frac{F^4}{3600a^5}|_p \leq \frac{p^{v_p(3600)}}{q^3} \leq q^{-1}$$

This tells us that the coefficients in

$$\frac{1}{F}\sum_{j\geq 3}\binom{1-s}{j}(B_j)(\frac{F}{a})^j$$

are divisible by p. We also have the expansion

$$\langle a\rangle^{1-s} = \exp((1-s)\log_p(\langle a\rangle)) = \sum_{j=0}^{\infty}\frac{1}{j!}(1-s)^j(\log_p(\langle a\rangle))^j$$

By standard results in p-adic analysis we have that q divides $\log_p(\langle a\rangle)$. Therefore every coefficient in the above expansion is bounded above p-adically by $p^{\frac{j}{p-1}}q^{-j} < 1$ so they are p-integral. In particular for $j \geq 3$ the bound becomes $p^{\frac{3}{p-1}}q^{-3} \leq (pq)^{-1}$. For $j = 2$ we manually compute the upper bound to be $p^{v_p(2)}q^{-2} < (pq)^{-1}$ still. It remains to study the cases $j = 0, 1, 2$ from the original series. We have

$$|\frac{B_2F^1}{2!a^2}|_p = |\frac{F}{12a^2}|_p \leq \frac{p^{v_p(12)}}{q} \leq 1$$
$$|\frac{B_1}{1!a}|_p = |-\frac{1}{2a}|_p \leq p^{v_p(2)} \leq p$$
$$|\frac{B_0F^{-1}}{1}|_p = |0|_p = 0$$

As a result, by removing all contributors whose coefficients are already divisible by p, we only need to consider the finite sum

$$L_p(s,\chi) \equiv \frac{1}{s-1}\sum_{1\leq a\leq F:(a,p)=1}\chi(a)(1+(1-s)\log_p(\langle a\rangle))(\frac{1}{F}-\frac{1-s}{2a}+\frac{(1-s)(1-s-1)F}{12a^2}) \pmod p$$

This gives us the coefficients a_0, a_1 and a_2 modulo p as

$$a_0 \equiv -\sum_{1\leq a\leq F:(a,p)=1}\chi(a)(\frac{1}{F}\log_p(\langle a\rangle)-\frac{1}{2a}-\frac{F}{12a^2}) \pmod p$$

$$a_1 \equiv -\sum_{1\leq a\leq F:(a,p)=1}\chi(a)(\frac{F}{12a^2}-\frac{\log_p(\langle a\rangle)}{2a}--\frac{F\log_p(\langle a\rangle)}{12a^2}) \pmod p$$

$$a_2 \equiv -\sum_{1\leq a\leq F:(a,p)=1}\chi(a)(\frac{F\log_p(\langle a\rangle)}{12a^2}) \pmod p$$

q divides $\frac{F\log_p(\langle a\rangle)}{12a^2}$ so obviously $a_2 \equiv 0 \pmod p$. p also divides $\frac{\log_p(\langle a\rangle)}{2a}$ so we can write

$$a_1 \equiv -\frac{F}{12}\sum_{1\leq a\leq F:(a,p)=1}\chi(a)a^{-2} \pmod p$$

If p is not 2 or 3 then q divides $\frac{F}{12}$ and so $a_1 \equiv 0 \pmod p$. If p is 2 or 3 then $a^2 \equiv 1 \pmod p$ for all units a. Hence the sum becomes $-\frac{F}{12}\sum_{1\leq a\leq F:(a,p)=1}\chi(a) \equiv 0 \pmod p$ again.

$\frac{1}{F}\log_p(\langle a\rangle)$ and $\frac{F}{12a^2}$ are both p-integral so showing that a_0 is p-integral reduces to showing that

$$\frac{1}{2}\sum_{1\leq a\leq F:(a,p)=1}\frac{\chi(a)}{a}$$

is p-integral. This is only an issue when $p = 2$. We can write this as

$$\frac{1}{2}\sum_{a=1}^{F}\frac{\chi(a)}{a}-\frac{1}{2}\sum_{b=1}^{F/p}\frac{\chi(bp)}{bp} = \frac{1}{2}\sum_{a=1}^{F}\frac{\chi(a)}{a} \equiv \frac{1}{2}\sum_{a=1}^{F}\chi\omega^{-1}(a) \pmod p$$

The right hand side is then p-integral by standard results on characters. This completes the proof. □

Corollary 6.1.14. *If χ is nontrivial and pq does not divide its conductor then*

$$L_p(m,\chi) \equiv L_p(n,\chi) \pmod{p}$$

for any p-integral m and and n.

Proposition 6.1.15. *If m, n are positive integers with $m \equiv n \pmod{p-1}$ and neither of which is divisible by $p-1$ then*

$$\frac{B_m}{m} \equiv \frac{B_n}{n} \pmod{p}$$

Proof.

$$L_p(1-m,\omega^m) = -(1-\omega^m(p)\omega^{-m}(p)p^{m-1})\frac{B_m}{m} = -(1-p^{m-1})\frac{B_m}{m}$$

If $m \equiv n \pmod{p-1}$ are integers not divisible by $p-1$ then $\omega^m = \omega^n$ are not trivial. By the previous corollary, since pq does not divide the conductor p, we have

$$L_p(1-m,\omega^m) \equiv L_p(1-n,\omega^n) \pmod{p}$$

so $-(1-p^{m-1})\frac{B_m}{m} \equiv -(1-p^{n-1})\frac{B_n}{n} \pmod{p}$ and the result follows. □

Proposition 6.1.16. *If m is a positive integer with $m+1$ not divisible by $p-1$ then*

$$B_{1,\omega^m} \equiv \frac{B_{m+1}}{m+1} \pmod{p}$$

with both sides p-integral.

Proof. From our corollary we have

$$L_p(0,\omega^{m+1}) \equiv L_p(-m,\omega^{m+1}) \pmod{p}$$

The Bernoulli number computation gives

$$L_p(0,\omega^{m+1}) = -(1-\omega^m(p))B_{1,\omega^m} = -B_{1,\omega^m}$$

$$L_p(-m,\omega^{m+1}) = -(1-\omega^m(p)p^m)\frac{B_{m+1,\omega^m\omega^{-m}}}{m+1} = -\frac{B_{m+1}}{m+1}$$

The result then follows. □

We will now apply this result to refine our condition for the relative class number being divisible by p. Recall the explicit formula for the relative class number

$$h_K^- = 2p\prod_{\chi odd} -\frac{1}{2}B_{1,\chi}$$

The Bernoulli numbers $B_{1,\chi}$ for odd χ correspond to B_{1,ω^m} for odd m, when taken modulo p. The only Bernoulli number we need to worry about is $B_{1,\omega^{-1}}$ which is calculated as

$$B_{1,\omega^{-1}} = \frac{1}{p}\sum_{a=1}^{p}\omega^{-1}(a)a \equiv \frac{p-1}{p} \pmod{p}$$

As for the other Bernoulli numbers, we can use our results from p-adic L-functions to get

$$h_K^- \equiv 2(p-1)\prod_{m=1,3\ldots p-4} -\frac{1}{2}B_{1,\omega^m} \equiv 2(p-1)(-\frac{1}{2})^{\frac{p-3}{2}}\prod_{m=1,3\ldots p-4}\frac{B_{m+1}}{m+1} \pmod{p}$$

Therefore, if p does not divide any of the Bernoulli numbers $B_2 \ldots B_{p-3}$, then the relative class number is not divisible by p.

6.2 Gauss sums and the Stickelberger relation

Throughout this subsection we follow [IR90, Chapter 8] and [IR90, Chapter 14].

Definition 6.2.1. *Let F be a finite field. Let χ be a multiplicative character on F and let ψ be an additive character on F. Then the associated Gauss sum is defined as*

$$g(\chi,\psi) := \sum_{t\in F}\chi(t)\psi(t)$$

Recall that a multiplicative character on F is a group homomorphism

$$\chi : F^\times \to \mathbb{C}^\times$$

whose image must be $(|F|-1)^{\text{th}}$ roots of unity. An additive character is a group homomorphism

$$\psi : (F,+) \to \mathbb{C}^\times$$

whose image must be $|F|^{\text{th}}$ roots of unity. In particular, if the characteristic of the field F is a prime p, then the image must be the p^{th} roots of unity. Let $|F| = p^f$. If no additive character is given, the default additive character is

$$\psi(t) = \zeta_p^{Tr(t)}$$

Where $Tr(\cdot)$ is the trace map that sends F to $\mathbb{Z}/(p)$ via

$$Tr(t) := \sum_{i=0}^{f-1} t^{p^i}$$

To see why this is the trace map, note that $F \cong \mathbb{Z}[\zeta_{p^f-1}]/\mathfrak{p}$ where $\mathfrak{p}$ is the prime above p, by referring back to our classification of prime decomposition in cyclotomic fields. Then the Galois group of the extension $F : \mathbb{Z}/p$ is generated by the Frobenius map $t \mapsto t^p$, so the above definition of trace is correct.

Proposition 6.2.2. *Let $g(\chi,\psi)$ be a Gauss sum on F, with ψ non-trivial. Then $|g(\chi,\psi)|^2 = |F|$.*

Proof. For $1 \leq a < p$ define the modified additive characters ψ_a on F as

$$\psi_a(t) = \psi(at)$$

We will prove the proposition by evaluating $\sum_{1\leq a<p} g(\chi,\psi_a)\overline{g(\chi,\psi_a)}$ in two different ways. Firstly we note that

$$g(\chi,\psi_a)\overline{g(\chi,\psi_a)} = \sum_{x\in F}\sum_{y\in F}\chi(x)\overline{\chi(y)}\psi(a(x-y)) =$$
$$\overline{\chi(a)}\chi(a)\sum_{x\in F}\sum_{y\in F}\chi(ax)\overline{\chi(ay)}\psi(a(x-y)) = \sum_{x\in F}\sum_{y\in F}\chi(x)\overline{\chi(y)}\psi(x-y)$$

because multiplication by a, which is invertible, permutes the elements of F. It follows that $|g(\chi,\psi_a)| = |g(\chi,\psi)|$ for every a and so the sum above is $(p-1)|g(\chi,\psi)|^2$.

Alternatively, we can write

$$\sum_{1\leq a<p} g(\chi,\psi_a)\overline{g(\chi,\psi_a)} = \sum_{x\in F}\sum_{y\in F}\chi(x)\overline{\chi(y)}\sum_{1\leq a<p}\psi(a(x-y))$$

$\psi(0) = 1$ and so if $x = y$, then $\sum_{1\leq a<p}\psi(a(x-y)) = p-1$. Otherwise, $\sum_{1\leq a<p}\psi(a(x-y))$ will run through the p^{th} roots of unity and so this sum will clearly be 0. Therefore

$$\sum_{1\leq a<p}\psi(a(x-y)) = (p-1)\delta(x,y)$$

where δ is the Kronecker delta. It follows that

$$\sum_{x\in F}\sum_{y\in F}\chi(x)\overline{\chi(y)}\sum_{1\leq a<p}\psi(a(x-y)) = \sum_{x\in F}\chi(x)\overline{\chi(x)}(p-1) = (p-1)|F|$$

Equating the two different formulae gives $|g(\chi,\psi)|^2 = |F|$. □

Definition 6.2.3. *Let F be a finite field. Let χ,ψ be multiplicative characters on F. We define the associated Jacobi sum as*

$$J(\chi,\psi) := \sum_{a+b=1}\chi(a)\psi(b)$$

Proposition 6.2.4. *Let χ,ψ be multiplicative characters on F so that $\chi \neq \overline{\psi}$. Then*

$$J(\chi,\psi) = \frac{g(\chi)g(\psi)}{g(\chi\psi)}$$

Proof. We begin by writing

$$g(\chi)g(\psi) = \sum_{x\in F}\sum_{y\in F}\chi(x)\psi(y)\zeta_p^{Tr(x+y)} = \sum_{0\leq t<p}\zeta_p^t\sum_{Tr(x+y)=t}\chi(x)\psi(y)$$

Then we can relate this sum to the Jacobi sum using

$$\sum_{Tr(x+y)=t} \chi(x)\psi(y) = \sum_{i=0}^{f-1} \sum_{x+y=t+r_i} \chi(x)\psi(y) = \sum_{i=0}^{f-1} \chi\psi(t+r_i) \sum_{x+y=1} \chi(x)\psi(y) = J(\chi,\psi) \sum_{i=0}^{f-1} \chi\psi(t+r_i)$$

where the r_i are the zero trace elements of F. Next we have

$$g(\chi\psi) = \sum_{x\in F} \chi\psi(x)\zeta_p^{Tr(x)} = \sum_{0\leq t<p} \zeta_p^t \sum_{i=0}^{f-1} \chi\psi(t+r_i)$$

because $\sum_{0\leq t<p} \sum_{i=0}^{f-1} \chi\psi(t+r_i)$ will run through all the elements of F, classifying them by trace. Putting it all together gives the desired result. $g(\chi\psi)$ does not vanish since $\chi \neq \overline{\psi}$. □

Corollary 6.2.5. *For any multiplicative characters χ, ψ on F so that $\chi \neq \overline{\psi}$, we have*

$$|J(\chi,\psi)|^2 = \frac{|g(\chi)|^2|g(\psi)|^2}{|g(\chi\psi)|^2} = |F|$$

Proposition 6.2.6. *Let χ be a multiplicative character on F. Let $\mathrm{ord}(\chi) = m$. Then*

$$g(\chi)^m = \chi(-1)|F| \prod_{i=1}^{m-2} J(\chi,\chi^i)$$

Proof. At the beginning we have

$$J(\chi,\chi) = \frac{g(\chi)^2}{g(\chi^2)} \implies g(\chi)^2 = J(\chi,\chi)g(\chi^2)$$

Now assume that

$$g(\chi)^k = g(\chi^k) \prod_{i=1}^{k-1} J(\chi,\chi^i)$$

for $k < m-1$. Then we can multiply both sides by $g(\chi)$ and use

$$g(\chi^k)g(\chi) = J(\chi,\chi^k)g(\chi^{k+1})$$

since $\chi^k \neq \chi^{-1}$. This gives

$$g(\chi)^{k+1} = g(\chi^{k+1}) \prod_{i=1}^{k} J(\chi,\chi^i)$$

Inductively, we get to the expression

$$g(\chi)^{m-1} = g(\chi^{m-1}) \prod_{i=1}^{m-2} J(\chi,\chi^i) \implies g(\chi)^m = g(\chi^{-1})g(\chi) \prod_{i=1}^{m-2} J(\chi,\chi^i)$$

Then we need to make use of

$$g(\chi^{-1}) = \sum_{x\in F} \chi^{-1}(x)\zeta_p^{Tr(x)} = \overline{\chi(-1)} \sum_{x\in F} \overline{\chi(-x)}\overline{\zeta_p^{Tr(-x)}} = \overline{\chi(-1)g(\chi)}$$

But $\chi(-1)$ is either 1 or -1, so that $\overline{\chi(-1)} = \chi(-1)$. This means that $g(\chi^{-1})g(\chi) = \chi(-1)|F|$ and the result follows directly. $\square$

We will now work on factoring certain Gauss sums into prime ideals. The factorization will give us the Stickelberger relation, which finds an element in the group ring of the Galois group of cyclotomic fields that annihilates the ideal class group. This, together with result from the next subsection, will help us complete Kummer's criterion.

Definition 6.2.7. *Let $\mathfrak{p}$ be a prime ideal in $\mathbb{Q}(\zeta_m)$ not containing m. We define the power residue symbol as a multiplicative character on $\mathbb{Z}[\zeta_m]/\mathfrak{p}$ which sends a unit a to*

$$\left(\frac{a}{\mathfrak{p}}\right)_m = \zeta_m^{k(a)}$$

for the unique m^{th} root of unity $\zeta_m^{k(a)}$ so that $a^{\frac{N(\mathfrak{p})-1}{m}} \equiv \zeta_m^{k(a)} \pmod{\mathfrak{p}}$

A few explanations are in order. The roots of unity are distinct modulo $\mathfrak{p}$, otherwise we get some equivalence of the form $1-\zeta_m^k \equiv 0 \pmod{\mathfrak{p}}$ for some k not divisible by m. However, m is contained in $(1-\zeta_m^k)$ which is in turn contained in $\mathfrak{p}$ and we get a contradiction. m divides $N(\mathfrak{p})-1$ because $N(\mathfrak{p}) = p^f$ for f the residue field degree of $\mathfrak{p}$ which satisfies $\text{ord}_m(p) = f$ by our classification of prime decomposition in cyclotomic fields.

Let $p^f = N(\mathfrak{p})$ where p is the prime below $\mathfrak{p}$. Then the finite field we are concerned about is $F = \mathbb{Z}[\zeta_m]/\mathfrak{p}$. We associate the multiplicative character

$$\chi_{\mathfrak{p}}(\cdot) = \left(\frac{\cdot}{\mathfrak{p}}\right)_m^{-1}$$

to this field. Then we will work on factoring the Gauss sum $g(\chi_{\mathfrak{p}}(\cdot))$, which is an element of $\mathbb{Q}(\zeta_{p^f-1}, \zeta_p)$. The order of $\chi_{\mathfrak{p}}$ is m so we can apply Proposition 6.2.6 to write

$$\Theta(\mathfrak{p}) := g(\chi_{\mathfrak{p}})^m = \chi_{\mathfrak{p}}(-1)|F| \prod_{i=1}^{m-2} J(\chi_{\mathfrak{p}}, \chi_{\mathfrak{p}}^i)$$

which is an expression in $\chi_{\mathfrak{p}}$ and therefore $g(\chi_{\mathfrak{p}})^m \in \mathbb{Q}(\zeta_m)$. We will factor this $\Theta(\mathfrak{p})$ first. We have $|\Theta(\mathfrak{p})|^2 = p^{mf}$ and so the only primes that could possibly divide $\Theta(\mathfrak{p})$ are the primes above p.

We will work in the tower of fields $\mathbb{Q}(\zeta_{p^f-1}, \zeta_p) : \mathbb{Q}(\zeta_{p^f-1}) : \mathbb{Q}(\zeta_m) : \mathbb{Q}$. Let's give a summary of the decomposition of p in these fields. By assumption p will split in $\mathbb{Q}(\zeta_m)$ and a fixed prime $\mathfrak{p}$ above p will have residue field degree f.

In the extension $\mathbb{Q}(\zeta_{p^f-1}) : \mathbb{Q}(\zeta_m)$, the prime $\mathfrak{p}$ splits completely because $\text{ord}_{p^f-1}(p) = f$ and so the residue field degree remains unchanged. Fix a prime $\mathfrak{P}$ above $\mathfrak{p}$.

In the extension $\mathbb{Q}(\zeta_{p^f-1}, \zeta_p) : \mathbb{Q}(\zeta_{p^f-1})$, the prime $\mathfrak{P}$ must ramify completely. To see this, note that p has ramification degree at least $p-1$ in the extension $\mathbb{Q}(\zeta_{p^f-1}, \zeta_p) : \mathbb{Q}$ but it is unramified in the extension $\mathbb{Q}(\zeta_{p^f-1}) : \mathbb{Q}$. Let $\mathcal{P}$ be the unique prime above $\mathfrak{P}$.

It is sufficient to know the orders of multiplicity $\mathrm{ord}_{\mathcal{P}}(\Theta(\mathfrak{p}))$ for each prime $\mathcal{P}$ above p in order to factor it. From the decomposition of p above we know that $\mathbb{Z}[\zeta_m]/\mathfrak{p} \cong \mathbb{Z}[\zeta_{p^f-1}]/\mathfrak{P}$ and all the $(p^f-1)^{\text{th}}$ roots of unity are distinct modulo $\mathfrak{P}$. Therefore we can define the power residue symbol modulo $\mathfrak{P}$ as

$$\gamma(\cdot) = \left(\frac{\cdot}{\mathfrak{P}}\right)_{p^f-1}$$

It follows that $\gamma^{-\frac{p^f-1}{m}} = \chi_{\mathfrak{p}}$ due to the isomorphism of fields. To compute the orders it is sufficient to compute the orders $\mathrm{ord}_{\mathcal{P}}(g(\gamma^{-a}))$ for all a. We make use of the following facts.

1. $\mathrm{ord}_{\mathcal{P}}(g(\gamma^{-1})) = 1$

   We start with

   $$g(\gamma^{-1}) = \sum_{t \in \mathbb{Z}[\zeta_{p^f-1}]/\mathcal{P}} \left(\frac{t}{\mathfrak{P}}\right)_{p^f-1}^{-1} \zeta_p^{Tr(t)} \equiv \sum_{t=1}^{p^f-1} \zeta_{p^f-1}^{-t} \zeta_p^{Tr(\zeta_{p^f-1}^t)} \pmod{\mathfrak{P}}$$

   The ideal $(1-\zeta_p)$ is contained in $\mathcal{P}$ and so we can expand binomially

   $$\zeta_p^{Tr(\zeta_{p^f-1}^t)} = (1-(1-\zeta_p))^{Tr(\zeta_{p^f-1}^t)} \equiv 1 - Tr(\zeta_{p^f-1}^t)(1-\zeta_p) \pmod{\mathcal{P}^2}$$

   Altogether this gives us

   $$g(\gamma^{-1}) \equiv \sum_{t=1}^{p^f-1} \zeta_{p^f-1}^{-t}(1-(1-\zeta_p)\sum_{i=0}^{f-1} \zeta_{p^f-1}^{tp^i}) \pmod{\mathcal{P}^2}$$

   However we know that $\sum_{t=1}^{p^f-1} \zeta_{p^f-1}^{-t} = 0$ and so in fact

   $$g(\gamma^{-1}) \equiv -(1-\zeta_p) \sum_{t=1}^{p^f-1} \zeta_{p^f-1}^{-t} \sum_{i=0}^{f-1} \zeta_{p^f-1}^{tp^i} \equiv -(1-\zeta_p) \sum_{t=1}^{p^f-1} \sum_{i=0}^{f-1} \zeta_{p^f-1}^{t(p^i-1)} \pmod{\mathcal{P}^2}$$

   Now $\sum_{t=1}^{p^f-1} \zeta_{p^f-1}^{t(p^i-1)} = 0$ unless $i = 0$ and so

   $$g(\gamma^{-1}) \equiv -(1-\zeta_p) \sum_{t=1}^{p^f-1} 1 \equiv -(p^f-1)(1-\zeta_p) \equiv (1-\zeta_p) \pmod{\mathcal{P}^2}$$

   By ramification degree considerations, we know that $(1-\zeta_p) \not\subset \mathcal{P}^2$ and so $\mathrm{ord}_{\mathcal{P}}(g(\gamma^{-1})) = 1$.

2. $\mathrm{ord}_{\mathcal{P}}(g(\gamma^{-(a+b)})) \leq \mathrm{ord}_{\mathcal{P}}(g(\gamma^{-a})) + \mathrm{ord}_{\mathcal{P}}(g(\gamma^{-b}))\ \forall a, b$

   We have $J(\gamma^{-a}, \gamma^{-b}) g(\gamma^{-(a+b)}) = g(\gamma^{-a}) g(\gamma^{-b})$ by applying Proposition 6.2.4. Taking the order of $\mathcal{P}$ dividing both sides gives the desired result.

3. $\mathrm{ord}_{\mathcal{P}}(g(\gamma^{-(a+b)})) \equiv \mathrm{ord}_{\mathcal{P}}(g(\gamma^{-a})) + \mathrm{ord}_{\mathcal{P}}(g(\gamma^{-b})) \pmod{p-1}\ \forall a, b$

   Recall that $J(\gamma^{-a}, \gamma^{-b})$ belongs to $\mathbb{Q}(\zeta_{p^f-1})$ as it is an expression in γ^{-a} and γ^{-b}. $\mathbb{Q}(\zeta_{p^f-1}, \zeta_p) : \mathbb{Q}(\zeta_{p^f-1})$ is totally ramified of degree $p-1$ so if $\mathcal{P}$ divides $J(\gamma^{-a}, \gamma^{-b})$, then $\mathcal{P}^{p-1} = \mathfrak{P}$ divides $J(\gamma^{-a}, \gamma^{-b})$. It follows that $\mathrm{ord}_{\mathcal{P}}(J(\gamma^{-a}, \gamma^{-b}))$ is a multiple of $p-1$.

   Taking the order of $\mathcal{P}$ dividing both sides of $J(\gamma^{-a}, \gamma^{-b})g(\gamma^{-(a+b)}) = g(\gamma^{-a})g(\gamma^{-b})$ then reducing modulo $p-1$ gives the desired result.

4. $\mathrm{ord}_{\mathcal{P}}(g(\gamma^{-pa})) = \mathrm{ord}_{\mathcal{P}}(g(\gamma^{-a}))\ \forall a$

   We start with

$$g(\gamma^{-pa}) = \sum_{t \in \mathbb{Z}[\zeta_{p^f-1}]/\mathcal{P}} \left(\frac{t}{\mathfrak{P}}\right)_{p^f-1}^{-pa} \zeta_p^{Tr(t)} = \sum_t \left(\frac{t^p}{\mathfrak{P}}\right)_{p^f-1}^{-a} \zeta_p^{Tr(t)}$$

   since $(p, p^f - 1) = 1$ and so p can be moved inside the power residue symbol. Then

$$\sum_t \left(\frac{t^p}{\mathfrak{P}}\right)_{p^f-1}^{-a} \zeta_p^{Tr(t)} = \sum_t \left(\frac{t^p}{\mathfrak{P}}\right)_{p^f-1}^{-a} \zeta_p^{Tr(t^p)} = g(\gamma^{-a})$$

   because t and t^p are conjugates modulo $\mathfrak{P}$. The roots of unity being distinct modulo $\mathfrak{P}$ and spanning its quotient field means that t^p runs through $\mathbb{Z}[\zeta_{p^f-1}]/\mathcal{P}$ also. The result follows.

These four facts are enough to determine the orders of multiplicity for all a. It is given by the following proposition.

Proposition 6.2.8. *Let $a \in \mathbb{Z}$ and write $a \equiv \sum_{i=0}^{f-1} a_i p^i \pmod{p^f - 1}$ with $0 \le a_i < p$ Then*

$$\mathrm{ord}_{\mathcal{P}}(g(\gamma^{-a})) = \sum_{i=0}^{f-1} a_i = (p-1) \sum_{i=0}^{f-1} \{\frac{p^i a}{p^f - 1}\}$$

Proof. We begin by proving the first equality. This is true for $a = 1$ by fact 1. Fact 2 gives us

$$\mathrm{ord}_{\mathcal{P}}(g(\gamma^{-a})) \le \sum_{i=1}^{a} \mathrm{ord}_{\mathcal{P}}(g(\gamma^{-1})) = a$$

As a consequence, fact 3 tells us that for $1 \le a < p$ we have $\mathrm{ord}_{\mathcal{P}}(g(\gamma^{-a})) = a$. Now take a in general as above. Applying fact 2 again also gives

$$\mathrm{ord}_{\mathcal{P}}(g(\gamma^{-a})) \le \sum_{i=0}^{f-1} \mathrm{ord}_{\mathcal{P}}(g(\gamma^{-p^i a_i})) = \sum_{i=0}^{f-1} \mathrm{ord}_{\mathcal{P}}(g(\gamma^{-a_i})) = \sum_{i=0}^{f-1} a_i$$

where we repeatedly applied fact 4 to each summand in the middle equality. If $a_i = 0$ then it is easy to show that the corresponding order is 0. To show that the above inequality is an equality, it would be sufficient to prove

$$\sum_{a=1}^{p^f-2} \mathrm{ord}_{\mathcal{P}}(g(\gamma^{-a})) = \sum_{a=1}^{p^f-2} \sum_{i=0}^{f-1} a_i = \frac{f(p^f-2)(p-1)}{2}$$

where the last equality is easy to demonstrate using the same trick by Gauss on summing all the integers from 1 to n. The left sum is simply

$$\frac{1}{2}\sum_{a=1}^{p^f-2} \mathrm{ord}_{\mathcal{P}}(g(\gamma^{-a})) + \mathrm{ord}_{\mathcal{P}}(g(\gamma^{-(p^f-1-a)}))$$

However $\gamma^{-(p^f-1-a)} = \overline{\gamma^{-a}}$ and we have shown before that $g(\overline{\chi}) = \overline{\chi(-1)g(\chi)}$ and so $g(\chi)g(\overline{\chi}) = \overline{\chi(-1)}|F|$. As a result $\mathrm{ord}_{\mathcal{P}}(g(\gamma^{-(p^f-1-a)})g(\gamma^{-a})) = (p-1)f$ and so the above sum becomes $\frac{f(p^f-2)(p-1)}{2}$. This is because $|F| = p^f$ and $\mathcal{P}$ goes into p a total of $p-1$ times.

To prove the second equality, note that

$$\sum_{i=0}^{f-1} p^i a \equiv (\sum_{i=0}^{f-1} a_i)\sum_{i=0}^{f-1} p^i \pmod{p^f-1}$$

We convert this statement into

$$\sum_{i=0}^{f-1}\{\frac{p^i a}{p^f-1}\} = \frac{(\sum_{i=0}^{f-1} a_i)\sum_{i=0}^{f-1} p^i}{p^f-1} = \frac{\sum_{i=0}^{f-1} a_i}{p-1}$$

which is about fractional parts. The result follows. □

We are now ready to factor $\Theta(\mathfrak{p})$. Its factors in $\mathbb{Q}(\zeta_m)$ are $\mathfrak{p}$ and its distinct conjugates. We know that $\mathrm{ord}_{\mathfrak{p}}(\Theta(\mathfrak{p})) = m\sum_{i=0}^{f-1}\{\frac{p^i\frac{p^f-1}{m}}{p^f-1}\}$. As for its conjugates, when $(t,m)=1$ we get

$$\mathrm{ord}_{\sigma_t^{-1}(\mathfrak{p})}(\Theta(\mathfrak{p})) = \mathrm{ord}_{\mathfrak{p}}(\sigma_t(\Theta(\mathfrak{p}))) = m\sum_{i=0}^{f-1}\{\frac{tp^i\frac{p^f-1}{m}}{p^f-1}\} = m\sum_{i=0}^{f-1}\{\frac{tp^i}{m}\}$$

The Artin symbol for $\mathfrak{p}$ is σ_p and so the distinct conjugates are given by choosing representatives from each coset of $\langle\sigma_p\rangle$ in $\mathrm{Gal}(\mathbb{Q}(\zeta_m):\mathbb{Q})$. If t is a representative then its coset will look like $\sigma_t\langle\sigma_p\rangle = \{\sigma_{tp^i} : i = 0\dots f-1\}$. This means that the sum $m\sum_{i=0}^{f-1}\{\frac{tp^i}{m}\}$ accounts for the entire coset corresponding to t and so in fact we can write

$$\Theta(\mathfrak{p}) = \prod_{t\in(\mathbb{Z}/(m))^\times/\langle\sigma_p\rangle} \sigma_t^{-1}(\mathfrak{p})^{m\sum_{i=0}^{f-1}\{\frac{tp^i}{m}\}} = \prod_{i=0}^{f-1}\prod_{t\in(\mathbb{Z}/(m))^\times/\langle\sigma_p\rangle} \sigma_t^{-1}(\mathfrak{p})^{m\{\frac{tp^i}{m}\}} = \prod_{(t,m)=1}\sigma_t^{-1}(\mathfrak{p})^t$$

Definition 6.2.9. *Let $\mathbb{Q}(\zeta_m)$ be a cyclotomic field with Galois group $G = (\mathbb{Z}/(m))^\times$. Let $\mathbb{Z}[G]$ be the group ring of G. Then the Stickelberger element is defined as the element*

$$\theta = \frac{1}{m}\sum_{(t,m)=1} t\sigma_t^{-1}$$

of $\mathbb{Q}[G]$. For any subfield K of $\mathbb{Q}(\zeta_m)$, its Stickelberger element is defined as the restriction of θ under the quotient of group rings of Galois groups.

We have proven that we can factorize

$$(\Theta(\mathfrak{p})) = (g(\chi_{\mathfrak{p}}))^m = (m\theta)(\mathfrak{p})$$

for every prime ideal $\mathfrak{p}$ of $\mathbb{Q}(\zeta_m)$. As a result the element $m\theta$ applied to any ideal of $\mathbb{Z}[\zeta_m]$ will always give a principal ideal, and therefore it annihilates the ideal class group of $\mathbb{Z}[\zeta_m]$.

We want more annihilators of the ideal class group. Consider the ideal $\mathcal{I} = (\theta)\mathbb{Z}[G] \cap \mathbb{Z}[G]$. Applying these elements to fractional ideals in $\mathbb{Q}(\zeta_m)$ will give us an ideal in $\mathbb{Q}(\zeta_m)$ which is principal. The ideal $\mathcal{I}$ is called the Stickelberger ideal and it annihilates the ideal class group. It is defined for abelian number fields analogously using their Stickelberger elements.

Proposition 6.2.10. *Let K be the m^{th} cyclotomic field with Galois group G over the rationals and Stickelberger ideal $\mathcal{I}$. Then*

$$\theta\mathfrak{i} \subset \mathcal{I}$$

where $\mathfrak{i}$ is the ideal generated by elements of the form $c - \sigma_c$ for integers c coprime to m.

Proof.

$$(c - \sigma_c)\theta = \sum_{(k,m)=1} \frac{kc\sigma_k^{-1} - k\sigma_{kc}^{-1}}{m} = \sum_{(k,m)=1} (c\{\frac{k}{m}\} - \{\frac{kc}{m}\})\sigma_k^{-1} \in \mathbb{Z}[G] \cap (\theta)\mathbb{Z}[G]$$

This is true for each integer c coprime to m and so $\theta\mathfrak{i} \subset \mathcal{I}$. □

6.3 Herbrand's theorem

In this subsection we follow [Was97, Chapter 6.3].

Take the ideal class group C of the p^{th} cyclotomic field $\mathbb{Q}(\zeta_p)$, for an odd prime p. Let $G = (\mathbb{Z}/(p))^\times$ be its Galois group. We know that $\mathbb{Z}[G]$ acts on C. This action can sometimes tell us about the nonexistence of some particular type of torsion. We can make $\mathbb{Z}_p[G]$ act on $C[p]$, the p-torsion subgroup of the ideal class group, as follows. Let $c \in C[p]$ and let $e = \sum_{k=1}^{p-1} a_k\sigma_k$ for some p-adic integers a_k. Then

$$e(c) = \prod_{k=1}^{p-1} \sigma_k(c)^{a_k}$$

This is well-defined as $c^p = 1$ for any $c \in C[p]$ and so $\sigma_k(c)^{a_k} = \sigma_k(c)^{\overline{a_k}}$ where $\overline{a_k}$ is the reduction of a_k modulo p. The Stickelberger ideal can be seen as an ideal I of $\mathbb{Z}_p[G]$ which annihilates $C[p]$. We will introduce some representation theory before we continue. Let ω be the Teichmuller character on G. We define the element

$$\epsilon_i := \frac{1}{|G|} \sum_{g \in G} \omega^i(g)g^{-1} = \frac{1}{p-1} \sum_{k=1}^{p-1} \omega^i(k)\sigma_k^{-1}$$

of the group ring $\mathbb{Z}_p[G]$ for each character ω^i. We now have an important theorem from representation theory.

Theorem 6.3.1. *The elements ϵ_i are orthogonal idempotents of the group ring $\mathbb{Z}_p[G]$.*

Proof. First of all, these elements exist because $|G| = p-1$ is invertible in $\mathbb{Z}_p$. To show that they are (clearly non-trivial) idempotents, we write

$$\epsilon_i^2 := \frac{1}{|G|^2}\sum_{g\in G}\sum_{h\in G}\omega^i(gh)(gh)^{-1} = \frac{1}{|G|^2}\sum_{g\in G}|G|\omega^i(g)g^{-1} = \frac{1}{|G|}\sum_{g\in G}\omega^i(g)g^{-1} = \epsilon_i$$

Orthogonality is also simple. Suppose $i \neq j$, then

$$\epsilon_i\epsilon_j = \frac{1}{|G|^2}\sum_{g\in G}\sum_{h\in G}\omega^i(g)\omega^j(h)(gh)^{-1} = \frac{1}{|G|^2}\sum_{g\in G}g^{-1}\sum_{h\in G}\omega^i(h)\omega^j(gh^{-1}) =$$
$$= \frac{1}{|G|^2}\sum_{g\in G}g^{-1}\omega^j(g)\sum_{h\in G}\omega^i(h)\omega^j(h)^{-1}$$

From orthogonality relations we know that the sum $\sum_{h\in G}\omega^i(h)\omega^j(h)^{-1}$ is zero when $i \neq j$. It follows that $\epsilon_i\epsilon_j = 0$. □

$\mathbb{Z}_p[G]$ is a free $\mathbb{Z}_p$-module of rank $|G| = p-1$. It is then a theorem in algebra that for any $\mathbb{Z}_p[G]$-module M we have

$$M = \bigoplus_{i=0}^{p-2}\epsilon_i M = \bigoplus_{i=0}^{p-2} M_i$$

so the p-part of the ideal class group, being a $\mathbb{Z}_p[G]$-module, has a decomposition

$$C[p] = \bigoplus_{i=0}^{p-2} C[p]_i$$

Lemma 6.3.2. *We have $\epsilon_i\sigma = \omega^i(\sigma)\epsilon_i$ for all $\sigma \in G$.*

Proof.

$$\epsilon_i\sigma = \frac{1}{|G|}\sum_{\tau\in G}\omega^i(\tau)\tau^{-1}\sigma = \frac{1}{|G|}\omega^i(\sigma)\sum_{\tau\in G}\omega^i(\sigma^{-1}\tau)(\tau\sigma^{-1})^{-1} = \omega^i(\sigma)\epsilon_i \quad \square$$

It follows from this lemma that the Stickelberger element will act by scalar multiplication on each piece $C[p]_i$. If these scalars are not 0 (mod p), then we must have $C[p]_i = 0$ because the Stickelberger element annihilates $C[p]_i$.

Theorem 6.3.3. *Let θ be the Stickelberger element of $\mathbb{Q}(\zeta_p)$. By taking the p-adic Bernoulli number $B_{1,\omega^{-i}}$ modulo p for some i, we have*

$$\epsilon_i\theta \equiv B_{1,\omega^{-i}}\epsilon_i \pmod{p}$$

Proof.

$$\epsilon_i\theta = \frac{1}{p}\sum_{k=1}^{p-1}k\epsilon_i\sigma_k^{-1} = \frac{1}{p}\sum_{k=1}^{p-1}k\omega^i(\sigma_k^{-1})\epsilon_i = B_{1,\omega^{-i}}\epsilon_i \quad \square$$

Corollary 6.3.4. *By Proposition 6.2.10, $(c-\sigma_c)\theta \in \mathcal{I}$ annihilates $C[p]$ for any integer c coprime to p. As a result $(c-\omega^i(\sigma_c))B_{1,\omega^{-i}}$ annihilates $C[p]_i$ for any c coprime to p.*

Theorem 6.3.5 (Herbrand's theorem)**.** *Let C be the ideal class group of $\mathbb{Q}(\zeta_p)$. Let*

$$C[p] = \bigoplus_{i=0}^{p-2} C[p]_i$$

be the decomposition of $C[p]$ with the above notation. Then $C[p]_0 = C[p]_1 = 0$ and if $p \nmid B_{p-i}$ for i odd, then $C[p]_i = 0$.

Proof. The above corollary states that $(c-\omega^i(\sigma_c))B_{1,\omega^{-i}}$ annihilates $C[p]_i$. By results in Section 6.1, we know that $B_{1,\omega^{-i}} = 0$ for even $i \neq 0$, so we get no information about $C[p]_i$ for nonzero even i. Furthermore, we've shown that $B_{1,\omega^{-i}} \equiv \frac{B_{p-i}}{p-i} \pmod{p}$ and $p-i$ is invertible modulo p for nonzero i. As a result $(c-\omega^i(\sigma_c))B_{p-i}$ annihilates $C[p]_i$ for $i \neq 0$.

- Let $i = 0$. Then $(2-\omega^0(2))B_{1,\omega^0} = B_{1,\omega^0} = \frac{p-1}{2} \not\equiv 0 \pmod{p}$ and so $C[p]_0 = 0$.
- Let $i = 1$. Then $(p+1-\omega(p+1))B_{1,\omega^{-1}} = pB_{1,\omega^{-1}} \equiv p-1 \not\equiv \pmod{p}$. As a result $C[p]_1 = 0$.
- Let i be odd and not equal to 1. Then we can choose c so that $(c-\omega^i(\sigma_c)) \equiv \omega(c)(1-\omega(c)^{i-1}) \not\equiv 0 \pmod{p}$. As a result B_{p-i} annihilates $C[p]_i$ and we are done.

□

We have just obtained the result from the end of Section 6.1 algebraically, and in fact we obtained a strengthened result, which tells us information about individual pieces of the p-torsion of the class group. However, we only know information about $C[p]_i$ for odd i, and in order to check that a prime is regular, we need to show that $C[p]_i$ is trivial for all i. We will show in the next subsection that the odd and even parts of the p-torsion of the class group are linked, and it is sufficient to show that $C[p]_i$ is trivial for odd i.

6.4 Kummer's criterion for the regularity of primes

Let E be the group of units of $\mathbb{Z}[\zeta_p]$ and let G be the Galois group of the p^{th} cyclotomic field over the rationals. We will study, just as we did with the class group in the previous subsection, the action of $\mathbb{Z}_p[G]$ on E/pE. We can write

$$E/pE = \bigoplus_{i=0}^{p-2} \epsilon_i(E/pE) = \bigoplus_{i=0}^{p-2} (E/pE)_i$$

with the same notation as in the previous subsection. For $i = 0$ we have

$$\epsilon_0(u) = (\frac{1}{p-1}\sum_{k=1}^{p-1} \sigma_k^{-1})(u) = N(u)^{\frac{1}{p-1}} = 1$$

for any unit u, because ϵ_0 acts as some power of the norm. As a result $(E/pE)_0 = 0$. The following result is from [Was97, Proposition 8.10].

Proposition 6.4.1. *Let $E/pE = \bigoplus_{i=0}^{p-2}(E/pE)_i$ as above. Then we have*

- $(E/pE)_1 = \langle \zeta_p \rangle$
- $(E/pE)_i = 0$ *for odd* $i \neq 1$.

Proof. Let u be a unit. Write $u = \zeta_p^j r$ for some real unit r and integer k. Then

$$\epsilon_1(r) = (\frac{1}{p-1}\sum_{k=1}^{p-1}\omega(k)\sigma_k^{-1})(r) = \prod_{k=1}^{p-1}\sigma_k^{-1}(r)^{\frac{\omega(k)}{p-1}} =$$

$$= \prod_{k=1}^{\frac{p-1}{2}}\sigma_k^{-1}(r)^{\frac{\omega(k)}{p-1}}\sigma_k^{-1}(r)^{\frac{p-\omega(k)}{p-1}} = \prod_{k=1}^{\frac{p-1}{2}}\sigma_k^{-1}(r)^{\frac{p}{p-1}} \equiv 0 \pmod{pE}$$

Meanwhile, the root of unity is mapped to

$$\epsilon_1(\zeta_p) = \prod_{k=1}^{p-1}\sigma_k^{-1}(\zeta_p)^{\frac{\omega(k)}{p-1}} = \zeta_p$$

It follows that $(E/pE)_1 = \langle \zeta_p \rangle$. Now let $i \neq 1$ be odd. Then

$$\epsilon_i(\zeta_p^j r) = \epsilon_i(r) = \prod_{k=1}^{p-1}\sigma_k^{-1}(r)^{\frac{\omega^i(k)}{p-1}} = \prod_{k=1}^{\frac{p-1}{2}}\sigma_k^{-1}(r)^{\frac{\omega^i(k)}{p-1}}\sigma_k^{-1}(r)^{\frac{p-\omega^i(k)}{p-1}}$$

The roots of unity are killed because they belong to the ϵ_1 part, and the ϵ_i are orthogonal idempotents. The same cancellation occurs as in the $i = 1$ case and so we get $\epsilon_i(r) \equiv 0 \pmod{pE}$. □

We will now proceed to prove a special case of the reflection theorem, which will be sufficient to imply Kummer's criterion. We will assume some knowledge of Kummer theory and make use of class field theory. We require the following results in particular.

Proposition 6.4.2. *Let K be a number field with ideal class group C. Then the following are true.*

1. *There exists a unique number field extension $H : K$ called the Hilbert class field of K which is the maximally unramified abelian extension of K, in the sense that any other unramified abelian extension of K is intermediate.*
2. *For any intermediate number field $H : L : K$, the Artin symbol $(\frac{L:K}{\cdot})$ induces an isomorphism between $Gal(L : K)$ and a subgroup of C. This constitutes a one-to-one inclusion preserving bijection between subgroups of $Gal(H : K)$ and subgroups of C.*

Proof. This is class field theory for unramified abelian extensions. It is a special case of [Jan96, Theorem 9.16]. □

Proposition 6.4.3. *Let K be a number field containing the p^{th} roots of unity.*

1. *There is a one-to-one inclusion preserving bijection between p-extensions of K and subgroups of $K^\times/(K^\times)^p$ as follows. Given a subgroup $B \triangleleft K^\times/(K^\times)^p$, we get a p-extension $K(\sqrt[p]{B}) : K$.*
2. *Given a p-extension $K(\sqrt[p]{B}) : K$ as above with Galois group $H = Gal(K(\sqrt[p]{B}) : K)$, there is a perfect bilinear pairing*

$$B \times H \to \langle \zeta_p \rangle$$
$$(b, h) \mapsto \frac{h(\sqrt[p]{b})}{\sqrt[p]{b}}$$

Proof. For an elaboration on the statements see [Stec]. For proofs see [Har]. □

For the remainder of this subsection we follow [Was97, Chapter 10.2].

Let $G = \text{Gal}(\mathbb{Q}(\zeta_p) : \mathbb{Q})$. Let L be the maximally unramified elementary abelian p-extension of $\mathbb{Q}(\zeta_p)$. In other words, the maximally unramified abelian extension with Galois group of the form $\sum_{i=1}^{n} \mathbb{Z}/(p)$ for some n. Let $H = Gal(L : \mathbb{Q}(\zeta_p))$. Then by Proposition 6.4.2 we have $H \cong C/pC$, where C is the ideal class group of $\mathbb{Q}(\zeta_p)$. H is a normal subgroup of $\text{Gal}(L : \mathbb{Q})$ and so G acts on H by conjugation. In fact, H becomes a $\mathbb{Z}[G]$-module, and the isomorphism $H \cong C/pC$ is $\mathbb{Z}[G]$-linear. To see this, for any $g \in G$ we have

$$(\frac{L : \mathbb{Q}(\zeta_P)}{g(\mathfrak{i})}) = g(\frac{L : \mathbb{Q}(\zeta_P)}{\mathfrak{i}})g^{-1}$$

as Artin symbols, which we've shown in Proposition 1.6.9.

Now $L : \mathbb{Q}(\zeta_p)$ is a Kummer extension, and Proposition 6.4.3 states that we can get this extension by adjoining some p^{th} roots of elements in $\mathbb{Q}(\zeta_p)^\times$. In particular there is a subgroup $B \triangleleft \mathbb{Q}(\zeta_p)^\times/(\mathbb{Q}(\zeta_p)^\times)^p$ so that $L = \mathbb{Q}(\zeta_p, \sqrt[p]{B})$. It also states that there is a pairing

$$B \times H \to \langle \zeta_p \rangle$$

sending (b, h) to $\frac{h(\sqrt[p]{b})}{\sqrt[p]{b}}$, and that this pairing is perfect and bilinear. This implies that $B \cong \widehat{H}$ canonically by sending b to the map that sends $h \to \frac{h(\sqrt[p]{b})}{\sqrt[p]{b}}$. This is $\mathbb{Z}[G]$-linear because

$$g((b, h)) = g(\frac{h(\sqrt[p]{b})}{\sqrt[p]{b}}) = \frac{ghg^{-1}(\sqrt[p]{g(b)})}{\sqrt[p]{g(b)}} = (g(b), ghg^{-1})$$

for all $g \in G$. Since H is finite we also have some non-canonical isomorphism $B \simeq H$.

$L : \mathbb{Q}(\zeta_p)$ is unramified so for each $b \in B$ we claim that $(b) = \mathfrak{i}^p$ for some ideal $\mathfrak{i}$ in $\mathbb{Q}(\zeta_p)$. Suppose not, then $(b) = (\sqrt[p]{b})^p$ in L whilst the prime ideals dividing (b) in $\mathbb{Q}(\zeta_p)$ will do so with multiplicity coprime to p. Therefore the prime ideals dividing (b) must be totally ramified in $L : \mathbb{Q}(\zeta_p)$, a contradiction.

Now consider the map $\phi : B \to C/pC$ induced by sending b to the ideal $\mathfrak{i}$ discussed above. This is well-defined because $\phi((\mathbb{Q}(\zeta_p)^\times)^p) \equiv 0 \pmod{pC}$. It is also $\mathbb{Z}[G]$-linear because

$$\phi(g(b)) = \phi(g(\mathfrak{i})^p) = \overline{g(\mathfrak{i})} = g(\bar{\mathfrak{i}})$$

for any $g \in G$. We will also need the kernel of this map. Suppose $\phi(b) = 1$. Then we have $(b) = (a)^p$ for some $a \in \mathbb{Q}(\zeta_p)$, and so $b = ua^p$ for some unit u. However $b \equiv ba^{-p} \equiv u \pmod{(\mathbb{Q}(\zeta_p)^\times)^p}$. Therefore the kernel of ϕ is induced by a subgroup of the group of units. In particular

$$\ker(\phi) \subset E/pE$$

because $pE = E \cap (\mathbb{Q}(\zeta_p)^\times)^p$. This containment is $\mathbb{Z}[G]$-linear, since $g(ua^p) = g(u)g(a)^p \equiv g(u) \pmod{(\mathbb{Q}(\zeta_p)^\times)^p}$. We now have all the tools we need to finish Kummer's criterion.

Theorem 6.4.4. *Suppose p does not divide $h^-(\mathbb{Q}(\zeta_p))$. Then p does not divide $h(\mathbb{Q}(\zeta_p))$.*

Proof. Adopt the notation developed above. Recall that we have $H \cong C/pC \cong C[p]$ as $\mathbb{Z}[G]$-modules. Then it follows that $\epsilon_i H \cong C[p]_i$ G-linearly for each ϵ_i as defined in Section 6.3.

Let $h \in \epsilon_i H$. Then $\sigma_a h \sigma_a^{-1} = h^{\omega^i(a)}$ for each $\sigma \in G$ because $\epsilon_i \sigma_a = \omega^i(a)\epsilon_i$. Let $b \in \epsilon_k B$. We apply the Kummer pairing to these elements and study the action of G on the result. We get

$$(b, h)^{\omega(a)} = \sigma_a((b, h))$$

because (b, h) is a root of unity realized in $\mathbb{Z}_p$. Next we have

$$\sigma_a((b, h)) = (\sigma_a(b), \sigma_a h \sigma_a^{-1})$$

since we've shown that $B \cong \widehat{H}$ is G-linear. Then

$$(\sigma_a(b), \sigma_a h \sigma_a^{-1}) = (b^{\omega^k(a)}, h^{\omega^i(a)})$$

as discussed earlier. Finally

$$(b^{\omega^k(a)}, h^{\omega^i(a)}) = (b, h)^{\omega^{i+k}(a)} = (b, h)^{\omega(a)}$$

because the Kummer pairing is bilinear. Now if $(b, h) \neq 1$, we must have $i + k \equiv 1 \pmod{p-1}$. It follows that when there is an induced perfect bilinear pairing

$$\epsilon_k B \times \epsilon_i H \to \langle \zeta_p \rangle$$

whenever $i + k \equiv 1 \pmod{p-1}$. This implies, as before, that

$$\epsilon_k B \cong \epsilon_i H \cong C[p]_i$$

Here is where we will use the G-linear map $\psi : B \to C[p]$. This induces a map

$$\psi_k : \epsilon_k B \to C[p]_k$$

$$\ker(\phi_k) \subset \epsilon_k(E/pE)$$

since the containment $\ker(\phi) \subset E/pE$ is also G-linear. Now we can combine everything to write

$$\dim_{\mathbb{F}_p}(C[p]_i) = \dim_{\mathbb{F}_p}(\epsilon_k B) \leq \dim_{\mathbb{F}_p}(C[p]_k) + \dim_{\mathbb{F}_p}(\ker(\phi_k))$$

In particular, we get the bound

$$\dim_{\mathbb{F}_p}(C[p]_i) \leq \dim_{\mathbb{F}_p}(C[p]_k) + \dim_{\mathbb{F}_p}(\epsilon_k(E/pE))$$

From Proposition 6.4.1, when $k = 1$ we get

$$\dim_{\mathbb{F}_p}(C[p]_0) \leq \dim_{\mathbb{F}_p}(C[p]_1) + 1$$

But we already know $C[p]_0 = 0$ and $C[p]_1 = 0$. When $k \neq 1$ is odd we get

$$\dim_{\mathbb{F}_p}(C[p]_i) \leq \dim_{\mathbb{F}_p}(C[p]_k)$$

Suppose p does not divide B_{p-k} for $k \neq 1$ odd. Then $C[p]_k = 0$ by Herbrand's theorem and so the above bound gives $C[p]_{p-k} = 0$. Therefore if p does not divide the relative class number $h^-(\mathbb{Q}(\zeta_p))$, it will not divide the class number $h(\mathbb{Q}(\zeta_p))$ so p would be regular. □

Combining this theorem with our condition for the divisibility of the relative class number by p will give us Kummer's criterion.

Corollary 6.4.5 (Kummer's criterion)**.** *If an odd prime p does not divide B_{p-i} for i odd in the range $3 \leq i \leq p-2$, then p is regular.*

This gives a very nice way of proving Fermat's last theorem in many cases.

Corollary 6.4.6. *Let $p \geq 5$ be a prime number. If p does not divide B_{p-i} for i odd in the range $3 \leq i \leq p-2$ then FLT holds for exponent p.*

7 Acknowledgements

I would like to thank my supervisor Dr Carl Wang-Erickson for the many insightful discussions, which enriched my knowledge of number theory and were essential in the development of this document.

References

[BA] Robert B. Ash. Norms, traces and discriminants. https://faculty.math.illinois.edu/~r-ash/Ant/AntChapter2.pdf. Accessed : 2018-08-31.

[Chi] Edmund Y. M. Chiang. Classical analysis. http://www.math.ust.hk/~machiang/391N/Classical_Analysis.pdf. Accessed : 2018-08-31.

[Cona] Keith Conrad. The conductor ideal. http://www.math.uconn.edu/~kconrad/blurbs/gradnumthy/conductor.pdf. Accessed : 2018-08-31.

[Conb] Keith Conrad. The different ideal. http://www.math.uconn.edu/~kconrad/blurbs/gradnumthy/different.pdf. Accessed : 2018-08-31.

[Conc] Keith Conrad. Fermat's last theorem for regular primes. http://www.math.uconn.edu/~kconrad/blurbs/gradnumthy/fltreg.pdf. Accessed : 2018-08-31.

[Cond] Keith Conrad. Ideal factorization. http://www.math.uconn.edu/~kconrad/blurbs/gradnumthy/idealfactor.pdf. Accessed : 2018-08-31.

[Cox13] David A. Cox. *Primes of the form $x^2 + ny^2$*. Pure and Applied Mathematics (Hoboken). John Wiley & Sons, Inc., Hoboken, NJ, second edition, 2013. Fermat, class field theory, and complex multiplication.

[Gim] Geunho Gim. Ostrowski's theorem. http://www.math.ucla.edu/~ggim/F12-205A.pdf. Accessed : 2018-08-31.

[Har] Kris Harper. Group cohomology and kummer theory. http://www.math.uchicago.edu/~may/VIGRE/VIGRE2010/REUPapers/Harper.pdf. Accessed : 2018-08-31.

[IR90] Kenneth Ireland and Michael Rosen. *A classical introduction to modern number theory*, volume 84 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, second edition, 1990.

[Jan96] Gerald J. Janusz. *Algebraic number fields*, volume 7 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, second edition, 1996.

[Lan94] Serge Lang. *Algebraic number theory*, volume 110 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, second edition, 1994.

[Oss] Brian Osserman. The discriminant and ramification. https://www.math.ucdavis.edu/~osserman/classes/254a/lectures/13.pdf. Accessed : 2018-08-31.

[Stea] William Stein. The decomposition group. https://www.williamstein.org/papers/ant/html/node51.html. Accessed : 2018-08-31.

[Steb] William Stein. The exact sequence. https://www.williamstein.org/papers/ant/html/node53.html. Accessed : 2018-08-31.

[Stec] William Stein. Lecture 12: Kummer theory. https://wstein.org/edu/2010/582e/lectures/582e-2010-02-08/582e-2010-02-08.pdf. Accessed : 2018-08-31.

[Tat67] J. T. Tate. Fourier analysis in number fields, and Hecke's zeta-functions. In *Algebraic Number Theory (Proc. Instructional Conf., Brighton, 1965)*, pages 305–347. Thompson, Washington, D.C., 1967.

[UiO] 2013 Universitetet i Oslo, MAT4250. The class number formula. https://www.uio.no/studier/emner/matnat/math/MAT4250/h13/zetafu.pdf. Accessed : 2018-08-31.

[Was97] Lawrence C. Washington. *Introduction to cyclotomic fields*, volume 83 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, second edition, 1997.

[WE] Carl Wang-Erickson. Cyclotomic fields. http://wwwf.imperial.ac.uk/~cwangeri/pdfs/cyclotomic_fields_part_iii.pdf. Accessed : 2018-08-31.

[Wes] Tom Weston. The idelic aproach to number theory. http://people.math.umass.edu/~weston/oldpapers/idele.pdf. Accessed : 2018-08-31.