Algebraic Number Theory, a Computational Approach

William Stein

September 30, 2015

Contents

Preface

This book is based on notes the author created for a one-semester undergraduate course on Algebraic Number Theory, which the author taught at Harvard during Spring 2004 and Spring 2005. This book was mainly inspired by the [SD01, Ch. 1] and Cassels's article *Global Fields* in [Cas67]

- Copyright: William Stein, 2005, 2007.

License: Creative Commons Attribution-Share Alike 3.0 License

Please send any typos or corrections to `wstein@gmail.com`.

Acknowledgement: This book closely builds on Swinnerton-Dyer's book [SD01] and Cassels's article [Cas67]. Many of the students of Math 129 at Harvard during Spring 2004 and 2005 made helpful comments: Jennifer Balakrishnan, Peter Behrooz, Jonathan Bloom, David Escott Jayce Getz, Michael Hamburg, Deniz Kural, Danielle Li, Andrew Ostergaard, Gregory Price, Grant Schoenebeck, Jennifer Sinnott, Stephen Walker, Daniel Weissman, and Inna Zakharevich in 2004; Mauro Braunstein, Steven Byrnes, William Fithian, Frank Kelly, Alison Miller, Nizameddin Ordulu, Corina Patrascu, Anatoly Preygel, Emily Riehl, Gary Sivek, Steven Sivek, Kaloyan Slavov, Gregory Valiant, and Yan Zhang in 2005. Also the course assistants Matt Bainbridge and Andrei Jorza made many helpful comments. The mathemtical software [S$^+$11], [PAR], and [BCP97] were used in writing this book.

This material is based upon work supported by the National Science Foundation under Grant No. 0400386.

Chapter 1

Introduction

1.1 Mathematical background

In addition to general mathematical maturity, this book assumes you have the following background:

- Basics of finite group theory
- Commutative rings, ideals, quotient rings
- Some elementary number theory
- Basic Galois theory of fields
- Point set topology
- Basics of topological rings, groups, and measure theory

For example, if you have never worked with finite groups before, you should read another book first. If you haven't seen much elementary ring theory, there is still hope, but you will have to do some additional reading and exercises. We will briefly review the basics of the Galois theory of number fields.

Some of the homework problems involve using a computer, but there are examples which you can build on. We will not assume that you have a programming background or know much about algorithms. Most of the book uses Sage (`http://sagemath.org`), which is free open source mathematical software. The following is an example Sage session:

```
2 + 2
```

```
4
```

```
k.<a> = NumberField(x^2 + 1); k
```

```
Number Field in a with defining polynomial x^2 + 1
```

1.2 What is algebraic number theory?

A number field K is a finite degree algebraic extension of the rational numbers $\mathbb{Q}$. The primitive element theorem from Galois theory asserts that every such extension can be represented as the set of all polynomials of degree at most $d = [K : \mathbb{Q}] = \dim_{\mathbb{Q}} K$ in a single root α of some polynomial with coefficients in $\mathbb{Q}$:

$$K = \mathbb{Q}(\alpha) = \left\{ \sum_{n=0}^{m} a_n \alpha^n : a_n \in \mathbb{Q} \right\}.$$

Note that $\mathbb{Q}(\alpha)$ is non-canonically isomorphic to $\mathbb{Q}[x]/(f)$, where f is the minimal polynomial of α. The homomorphism $\mathbb{Q}[x] \to \mathbb{Q}(\alpha)$ that sends x to α has kernel (f), hence it induces an isomorphism between $\mathbb{Q}[x]/(f)$ and $\mathbb{Q}(\alpha)$. It is not canonical, since $\mathbb{Q}(\alpha)$ could have nontrivial automorphisms. For example, if $\alpha = \sqrt{2}$, then $\mathbb{Q}(\sqrt{2})$ is isomorphic as a field to $\mathbb{Q}(-\sqrt{2})$ via $\sqrt{2} \mapsto -\sqrt{2}$. There are two isomorphisms $\mathbb{Q}[x]/(x^2 - 2) \to \mathbb{Q}(\sqrt{2})$.

Algebraic number theory involves using techniques from (mostly commutative) algebra and finite group theory to gain a deeper understanding of the arithmetic of number fields and related objects (e.g., functions fields, elliptic curves, etc.). The main objects that we study in this book are number fields, rings of integers of number fields, unit groups, ideal class groups, norms, traces, discriminants, prime ideals, Hilbert and other class fields and associated reciprocity laws, zeta and L-functions, and algorithms for computing with each of the above.

1.2.1 Topics in this book

These are some of the main topics that are discussed in this book:

- Rings of integers of number fields
- Unique factorization of nonzero ideals in Dedekind domains
- Structure of the group of units of the ring of integers
- Finiteness of the abelian group of equivalence classes of nonzero ideals of the ring of integers (the "class group")
- Decomposition and inertia groups, Frobenius elements
- Ramification
- Discriminant and different
- Quadratic and biquadratic fields
- Cyclotomic fields (and applications)
- How to use a computer to compute with many of the above objects
- Valuations on fields
- Completions (p-adic fields)
- Adeles and Ideles

Note that we will not do anything nontrivial with zeta functions or L-functions.

1.3 Some applications of algebraic number theory

The following examples illustrate some of the power, depth and importance of algebraic number theory.

1. **Integer factorization** using the number field sieve. The number field sieve is the asymptotically fastest known algorithm for factoring general large integers (that don't have too special of a form). On December 12, 2009, the number field sieve was used to factor the RSA-768 challenge, which is a 232 digit number that is a product of two primes:

```
rsa768 = 123018668453011775513049495838496272077285356959533479\
219732245215172640050726365751874520219978646938995647494277406384592\
519255732630345373154826850791702612214291346167042921431160222124047\
927473779408066535141959745985690214341 3
n = 33478071698956898786044169848212690817704794983713768568912\
431388982883793878002287614711652531743087737814467999489
m = 36746043666799590428244633799627952632279158164343087642676\
032283815739666511279233373417143396810270092798736308917
n*m == rsa768
```

```
True
```

This record integer factorization cracked a certain 768-bit public key cryptosystem (see `http://eprint.iacr.org/2010/006`), thus establishing a lower bound on one's choice of key size:

```
$ man ssh-keygen   # in ubuntu-12.04
...
     -b bits
             Specifies the number of bits in the key to create.
             For RSA keys, the minimum size is 768 bits ...
```

2. **Primality testing:** Agrawal and his students Saxena and Kayal from India found in 2002 the first ever deterministic polynomial-time (in the number of digits) primality test. Their methods involve arithmetic in quotients of $(\mathbb{Z}/n\mathbb{Z})[x]$, which are best understood in the context of algebraic number theory.

3. **Deeper point of view** on questions in number theory:

   (a) Pell's Equation $x^2 - dy^2 = 1$ can be reinterpreted in terms of units in real quadratic fields, which leads to a study of unit groups of number fields.

   (b) Integer factorization leads to factorization of nonzero ideals in rings of integers of number fields.

   (c) The Riemann hypothesis about the zeros of $\zeta(s)$ generalizes to zeta functions of number fields.

   (d) Reinterpreting Gauss's quadratic reciprocity law in terms of the arithmetic of cyclotomic fields $\mathbb{Q}(e^{2\pi i/n})$ leads to class field theory, which in turn leads to the Langlands program.

4. Wiles's proof of **Fermat's Last Theorem**, i.e., that the equation $x^n + y^n = z^n$ has no solutions with x, y, z, n all positive integers and $n \geq 3$, uses methods from algebraic number theory extensively, in addition to many other deep techniques. Attempts to prove Fermat's Last Theorem long ago were hugely influential in the development of algebraic number theory by Dedekind, Hilbert, Kummer, Kronecker, and others.

5. **Arithmetic geometry:** This is a huge field that studies solutions to polynomial equations that lie in arithmetically interesting rings, such as the integers or number fields. A famous major triumph of arithmetic geometry is Faltings's proof of Mordell's Conjecture.

   Theorem 1.3.1 (Faltings)**.** *Let X be a nonsingular plane algebraic curve over a number field K. Assume that the manifold $X(\mathbb{C})$ of complex solutions to X has genus at least 2 (i.e., $X(\mathbb{C})$ is topologically a donut with at least two holes). Then the set $X(K)$ of points on X with coordinates in K is finite.*

   For example, Theorem 1.3.1 implies that for any $n \geq 4$ and any number field K, there are only finitely many solutions in K to $x^n + y^n = 1$.

   A major open problem in arithmetic geometry is the *Birch and Swinnerton-Dyer conjecture*. An *elliptic curves* E is an algebraic curve with at least one point with coordinates in K such that the set of complex points $E(\mathbb{C})$ is a topological torus. The Birch and Swinnerton-Dyer conjecture gives a criterion for whether or not $E(K)$ is infinite in terms of analytic properties of the L-function $L(E, s)$. See `http://www.claymath.org/millennium/Birch_and_Swinnerton-Dyer_Conjecture/`.

Part I

Algebraic Number Fields

Chapter 2

Basic Commutative Algebra

The commutative algebra in this chapter provides a foundation for understanding the more refined number-theoretic structures associated to number fields.

First we prove the structure theorem for finitely generated abelian groups. Then we establish the standard properties of Noetherian rings and modules, including a proof of the Hilbert basis theorem. We also observe that finitely generated abelian groups are Noetherian $\mathbb{Z}$-modules. After establishing properties of Noetherian rings, we consider rings of algebraic integers and discuss some of their properties.

2.1 Finitely Generated Abelian Groups

Finitely generated abelian groups arise all over algebraic number theory. For example, they will appear in this book as class groups, unit groups, and the underlying additive groups of rings of integers, and as Mordell-Weil groups of elliptic curves.

In this section, we prove the structure theorem for finitely generated abelian groups, since it will be crucial for much of what we will do later.

Let $\mathbb{Z} = \{0, \pm 1, \pm 2, \ldots\}$ denote the ring of (rational) integers, and for each positive integer n, let $\mathbb{Z}/n\mathbb{Z}$ denote the ring of integers modulo n, which is a cyclic abelian group of order n under addition.

Definition 2.1.1 (Finitely Generated)**.** A group G is *finitely generated* if there exists $g_1, \ldots, g_n \in G$ such that every element of G can be expressed as a finite product (or sum, if we write G additively) of positive or negative powers of the g_i.

For example, the group $\mathbb{Z}$ is finitely generated, since it is generated by 1.

Theorem 2.1.2 (Structure Theorem for Finitely Generated Abelian Groups)**.** *Let G be a finitely generated abelian group. Then there is an isomorphism*

$$G \approx (\mathbb{Z}/n_1\mathbb{Z}) \oplus (\mathbb{Z}/n_2\mathbb{Z}) \oplus \cdots \oplus (\mathbb{Z}/n_s\mathbb{Z}) \oplus \mathbb{Z}^r,$$

where $r, s \geq 0$, $n_1 > 1$ and $n_1 \mid n_2 \mid \cdots \mid n_s$. Furthermore, the n_i and r are uniquely determined by G.

Exercise 2.1.3. Quick! Guess how many abelian groups there are of order less than 12. Use Theorem 2.1.2 to classify all abelian groups of order less than 12. How many do you think there are? How many are there?

We will prove the theorem as follows. We first remark that any subgroup of a finitely generated free abelian group is finitely generated. Then we see how to represent finitely generated abelian groups as quotients of finite rank free abelian groups, and how to reinterpret such a presentation in terms of matrices over the integers. Next we describe how to use row and column operations over the integers to show that every matrix over the integers is equivalent to one in a canonical diagonal form, called the Smith normal form. We obtain a proof of the theorem by reinterpreting the Smith normal form in terms of groups. Finally, we observe that the representation in the theorem is necessarily unique.

Proposition 2.1.4. *If H is a subgroup of a finitely generated abelian group, then H is finitely generated.*

The key reason that this is true is that G is a finitely generated module over the principal ideal domain $\mathbb{Z}$. We defer the proof of Proposition 2.1.4 to Section 2.2, where we will give a complete proof of a beautiful generalization in the context of Noetherian rings (the Hilbert basis theorem).

Corollary 2.1.5. *Suppose G is a finitely generated abelian group. Then there are finitely generated free abelian groups F_1 and F_2 and there is a homomorphism $\psi : F_2 \to F_1$ such that $G \approx F_1/\psi(F_2)$.*

Proof. Let $x_1, \ldots, x_m$ be generators for G. Let $F_1 = \mathbb{Z}^m$ and let $\varphi : F_1 \to G$ be the homomorphism that sends the ith generator $(0, 0, \ldots, 1, \ldots, 0)$ of $\mathbb{Z}^m$ to x_i. Then φ is surjective, and by Proposition 2.1.4 the kernel $\ker(\varphi)$ of φ is a finitely generated abelian group. Suppose there are n generators for $\ker(\varphi)$, let $F_2 = \mathbb{Z}^n$ and fix a surjective homomorphism $\psi : F_2 \to \ker(\varphi)$. Then $F_1/\psi(F_2)$ is isomorphic to G. $\square$

An *sequence* of homomorphisms of abelian groups

$$H \xrightarrow{f} G \xrightarrow{g} K$$

is exact if $\mathrm{im}(f) = \ker(g)$. Given a finitely generated abelian group G, Corollary 2.1.5 provides an exact sequence

$$F_2 \xrightarrow{\psi} F_1 \to G \to 0.$$

Suppose G is a nonzero finitely generated abelian group. By the corollary, there are free abelian groups F_1 and F_2 and there is a homomorphism $\psi : F_2 \to F_1$ such that $G \approx F_1/\psi(F_2)$. Upon choosing a basis for F_1 and F_2, we obtain isomorphisms $F_1 \approx \mathbb{Z}^n$ and $F_2 \approx \mathbb{Z}^m$ for integers n and m. Just as in linear algebra, we view $\psi : F_2 \to F_1$ as being given by left multiplication by the $n \times m$ matrix A whose columns are the images of the generators of F_2 in $\mathbb{Z}^n$. We visualize this as follows:

$$\mathbb{Z}^m \xrightarrow{A} \mathbb{Z}^n \to G \to 0$$

The *cokernel* of the homomorphism defined by A is the quotient of $\mathbb{Z}^n$ by the image of A (i.e., the $\mathbb{Z}$-span of the columns of A), and this cokernel is isomorphic to G.

The following proposition implies that we may choose a bases for F_1 and F_2 such that the matrix of A only has nonzero entries along the diagonal, so that the structure of the cokernel of A is trivial to understand.

Proposition 2.1.6 (Smith normal form)**.** *Suppose A is an $n \times m$ integer matrix. Then there exist invertible integer matrices P and Q such that $A' = PAQ$ only has nonzero entries along the diagonal, and these entries are $n_1, n_2, \ldots, n_s, 0, \ldots, 0$, where $s \geq 0$, $n_1 \geq 1$ and $n_1 \mid n_2 \mid \cdots \mid n_s$. Here P and Q are invertible as integer matrices, so* $\det(P)$ *and* $\det(Q)$ *are ± 1. The matrix A' is called the* Smith normal form *of A.*

We will see in the proof of Theorem 2.1.2 that A' is uniquely determined by A. An example of a matrix in Smith normal form is

$$A = \begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & 6 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Proof of Proposition 2.1.6. The matrix P will be a product of matrices that define elementary row operations and Q will be a product corresponding to elementary column operations. The elementary row and column operations over $\mathbb{Z}$ are as follows:

1. [**Add multiple**] Add an integer multiple of one row to another (or a multiple of one column to another).

2. [**Swap**] Interchange two rows or two columns.

3. [**Rescale**] Multiply a row by -1.

Each of these operations is given by left or right multiplying by an invertible matrix E with integer entries, where E is the result of applying the given operation to the identity matrix, and E is invertible because each operation can be reversed using another row or column operation over the integers.

To see that the proposition must be true, assume $A \neq 0$ and perform the following steps (compare [Art91, Pg. 459]):

1. By permuting rows and columns, move a nonzero entry of A with smallest absolute value to the upper left corner of A. Now "attempt" (as explained in detail below) to make all other entries in the first row and column 0 by adding multiples of the top row or first column to other rows or columns, as follows:

Suppose a_{i1} is a nonzero entry in the first column, with $i > 1$. Using the division algorithm, write $a_{i1} = a_{11}q + r$, with $0 \leq r < a_{11}$. Now add $-q$ times the first row to the ith row. If $r > 0$, then go to step 1 (so that an entry with absolute value at most r is the upper left corner).

If at any point this operation produces a nonzero entry in the matrix with absolute value smaller than $|a_{11}|$, start the process over by permuting rows and columns to move that entry to the upper left corner of A. Since the integers $|a_{11}|$ are a decreasing sequence of positive integers, we will not have to move an entry to the upper left corner infinitely often, so when this step is done the upper left entry of the matrix is nonzero, and all entries in the first row and column are 0.

2. We may now assume that a_{11} is the only nonzero entry in the first row and column. If some entry a_{ij} of A is not divisible by a_{11}, add the column of A containing a_{ij} to the first column, thus producing an entry in the first column that is nonzero. When we perform step 2, the remainder r will be greater than 0. Permuting rows and columns results in a smaller $|a_{11}|$. Since $|a_{11}|$ can only shrink finitely many times, eventually we will get to a point where every a_{ij} is divisible by a_{11}. If a_{11} is negative, multiple the first row by -1.

After performing the above operations, the first row and column of A are zero except for a_{11} which is positive and divides all other entries of A. We repeat the above steps for the matrix B obtained from A by deleting the first row and column. The upper left entry of the resulting matrix will be divisible by a_{11}, since every entry of B is. Repeating the argument inductively proves the proposition. □

Example 2.1.7. The matrix $\begin{pmatrix} -1 & 2 \\ -3 & 4 \end{pmatrix}$ has Smith normal form $\begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}$, and the matrix $\begin{pmatrix} 1 & 4 & 9 \\ 16 & 25 & 36 \\ 49 & 64 & 81 \end{pmatrix}$ has Smith normal form $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 72 \end{pmatrix}$. As a double check, note that the determinants of a matrix and its Smith normal form match, up to sign. This is because

$$\det(PAQ) = \det(P)\det(A)\det(Q) = \pm\det(A).$$

We compute each of the above Smith forms using Sage, along with the corresponding transformation matrices. First the 2×2 matrix.

```
A = matrix(ZZ, 2, [-1,2, -3,4])
S, U, V = A.smith_form(); S
```

```
[1 0]
[0 2]
```

```
U*A*V
```

```
[1 0]
[0 2]
```

```
U
```

```
[ 0  1]
[ 1 -1]
```

```
V
```

```
[1 4]
[1 3]
```

The Sage `matrix` command takes as input the base ring, the number of rows, and the entries. Next we compute with a 3×3 matrix.

```
A = matrix(ZZ, 3, [1,4,9,  16,25,36, 49,64,81])
S, U, V = A.smith_form(); S
```

```
[ 1  0  0]
[ 0  3  0]
[ 0  0 72]
```

```
U*A*V
```

```
[ 1  0  0]
[ 0  3  0]
[ 0  0 72]
```

```
U
```

```
[  0   0   1]
[  0   1  -1]
[  1 -20 -17]
```

```
V
```

```
[  47   74   93]
[ -79 -125 -156]
[  34   54   67]
```

Finally we compute the Smith form of a matrix of rank 2:

```
m = matrix(ZZ, 3, [2..10]); m
```

```
[ 2  3  4]
[ 5  6  7]
[ 8  9 10]
```

```
m.smith_form()[0]
```

```
[1 0 0]
[0 3 0]
[0 0 0]
```

Proof of Theorem 2.1.2. Suppose G is a finitely generated abelian group, which we may assume is nonzero. As in the paragraph before Proposition 2.1.6, we use Corollary 2.1.5 to write G as the cokernel of an $n \times m$ integer matrix A. By Proposition 2.1.6 there are isomorphisms $Q : \mathbb{Z}^m \to \mathbb{Z}^m$ and $P : \mathbb{Z}^n \to \mathbb{Z}^n$ such that $A' = PAQ$ has diagonal entries $n_1, n_2, \ldots, n_s, 0, \ldots, 0$, where $n_1 > 1$ and $n_1 \mid n_2 \mid \ldots \mid n_s$. Then G is isomorphic to the cokernel of the diagonal matrix A', so

$$G \cong (\mathbb{Z}/n_1\mathbb{Z}) \oplus (\mathbb{Z}/n_2\mathbb{Z}) \oplus \cdots \oplus (\mathbb{Z}/n_s\mathbb{Z}) \oplus \mathbb{Z}^r, \tag{2.1.1}$$

as claimed. The n_i are determined by G, because n_i is the smallest positive integer n such that nG requires at most $s + r - i$ generators. We see from the representation (2.1.1) of G as a product that n_i has this property and that no smaller positive integer does. □

Exercise 2.1.8. Recall Smith normal form defined in Proposition 2.1.6. With only minor modifications, then the proposition and proof will work over any principle ideal domain. Find and apply these modifications then find the Smith normal form of the matrix $\begin{pmatrix} 1 & 2 & 3 \\ 0 & 1+i & 2 \\ 0 & 1 & 5 \end{pmatrix}$.

[*Hint*: You can use Sage to verify your answer. However, you will need to make explicitly construct the Gaussian integers in order to input the matrix. You can do this by the following code.]

```
K.<i> = QuadraticField(-1)
R = K.maximal_order()
M = matrix(R, 3, [1,2,3,0,1+i,2,0,1,5]); show(M)
#show(M.smith_form()[0]) #uncomment for the answer
```

Exercise 2.1.9. Let $A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}$.

1. Find the Smith normal form of A.
2. Prove that the cokernel of the map $\mathbb{Z}^3 \to \mathbb{Z}^3$ given by multiplication by A is isomorphic to $\mathbb{Z}/3\mathbb{Z} \oplus \mathbb{Z}$.

2.2 Noetherian Rings and Modules

A module M over a commutative ring R with unit element is much like a vector space, but with more subtle structure. In this book, most of the modules we encounter will be noetherian, which is a generalization of the "finite dimensional" property of vector spaces. This section is about properties of noetherian modules (and rings), which are crucial to much of this book. We thus give complete proofs of these properties, so you will have a solid foundation on which to learn algebraic number theory.

We first define noetherian rings and modules, then introduce several equivalent characterizations of them. We prove that when the base ring is noetherian, a module is finitely generated if and only if it is noetherian. Next we define short exact sequences, and prove that the middle module in a sequence is noetherian if and only if the first and last modules are noetherian. Finally, we prove the Hilbert basis theorem, which asserts that adjoining finitely many elements to a noetherian ring results in a noetherian ring.

Let R be a commutative ring with unity. An *R-module* is an additive abelian group M equipped with a map $R \times M \to M$ such that for all $r, r' \in R$ and all $m, m' \in M$ we have $(rr')m = r(r'm)$, $(r+r')m = rm + r'm$, $r(m+m') = rm + rm'$, and $1m = m$. A *submodule* of M is a subgroup of M that is preserved by the action of R. For example, R is a module over itself, and any ideal I in R is an R-submodule of R.

Example 2.2.1. Abelian groups are the same as $\mathbb{Z}$-modules, and vector spaces over a field K are the same as K-modules.

An R-module M is finitely generated if there are elements $m_1, \ldots, m_n \in M$ such that every element of M is an R-linear combination of the m_i. The noetherian property is stronger than just being finitely generated:

Definition 2.2.2 (Noetherian)**.** An R-module M is *noetherian* if every submodule of M is finitely generated. A ring R is *noetherian* if R is noetherian as a module over itself, i.e., if every ideal of R is finitely generated.

Any submodule M' of a noetherian module M is also noetherian. Indeed, if every submodule of M is finitely generated then so is every submodule of M', since submodules of M' are also submodules of M.

Example 2.2.3. Let $R = M = \mathbb{Q}[x_1, x_2, \ldots]$ be a polynomial ring over $\mathbb{Q}$ in infinitely many indeterminants x_i. Then M is finitely generated *as an R-module* (!), since it is generated by 1. Consider the submodule $I = (x_1, x_2, \ldots)$ of polynomials with 0 constant term, and suppose it is generated by polynomials $f_1, \ldots, f_n$. Let x_i be an indeterminant that does not appear in any f_j, and suppose there are $h_k \in R$ such that $\sum_{k=1}^{n} h_k f_k = x_i$. Setting $x_i = 1$ and all other $x_j = 0$ on both sides of this equation and using that the f_k all vanish (they have 0 constant term), yields $0 = 1$, a contradiction. We conclude that the ideal I is not finitely generated, hence M is not a noetherian R-module, despite being finitely generated.

Definition 2.2.4 (Ascending chain condition)**.** An R-module M satisfies the *ascending chain condition* if every sequence $M_1 \subset M_2 \subset M_3 \subset \cdots$ of submodules of M eventually stabilizes, i.e., there is some n such that $M_n = M_{n+1} = M_{n+2} = \cdots$.

We will use the notion of maximal element below. If $\mathcal{X}$ is a set of subsets of a set S, ordered by inclusion, then a *maximal element* $A \in \mathcal{X}$ is a set such that no superset of A is contained in $\mathcal{X}$. Note that $\mathcal{X}$ may contain many different maximal elements.

Proposition 2.2.5. *If M is an R-module, then the following are equivalent:*

1. *M is noetherian,*
2. *M satisfies the ascending chain condition, and*
3. *Every nonempty set of submodules of M contains at least one maximal element.*

Proof. 1 $\implies$ **2:** Suppose $M_1 \subset M_2 \subset \cdots$ is a sequence of submodules of M. Then $M_\infty = \cup_{n=1}^{\infty} M_n$ is a submodule of M. Since M is noetherian and M_∞ is a submodule of M, there is a finite set $a_1, \ldots, a_m$ of generators for M_∞. Each a_i must be contained in some M_j, so there is an n such that $a_1, \ldots, a_m \in M_n$. But then $M_k = M_n$ for all $k \geq n$, which proves that the chain of M_i stabilizes, so the ascending chain condition holds for M.
2 $\implies$ **3:** Suppose 3 were false, so there exists a nonempty set S of submodules of M that does not contain a maximal element. We will use S to construct an infinite ascending chain of submodules of M that does not stabilize. Note that S is infinite, otherwise it would contain a maximal element. Let M_1 be any element of S. Then there is an M_2 in S that contains M_1, otherwise S would contain the maximal element M_1. Continuing inductively in this way we find an M_3 in S that properly contains M_2, etc., and we produce an infinite ascending chain of submodules of M, which contradicts the ascending chain condition.
3 $\implies$ **1:** Suppose 1 is false, so there is a submodule M' of M that is not finitely generated. We will show that the set S of all finitely generated submodules of M' does not have a maximal element, which will be a contradiction. Suppose S does have a maximal element L. Since L is finitely generated and $L \subset M'$, and M' is not finitely generated, there is an $a \in M'$ such that $a \notin L$. Then $L' = L + Ra$ is an element of S that strictly contains the presumed maximal element L, a contradiction. $\square$

A *homomorphism* of R-modules $\varphi : M \to N$ is an abelian group homomorphism such that for any $r \in R$ and $m \in M$ we have $\varphi(rm) = r\varphi(m)$. A sequence

$$L \xrightarrow{f} M \xrightarrow{g} N,$$

where f and g are homomorphisms of R-modules, is *exact* if $\mathrm{im}(f) = \ker(g)$. A *short exact sequence* of R-modules is a sequence

$$0 \to L \xrightarrow{f} M \xrightarrow{g} N \to 0$$

that is exact at each point; thus f is injective, g is surjective, and $\mathrm{im}(f) = \ker(g)$.

Example 2.2.6. The sequence

$$0 \to \mathbb{Z} \xrightarrow{2} \mathbb{Z} \to \mathbb{Z}/2\mathbb{Z} \to 0$$

is an exact sequence, where the first map sends 1 to 2, and the second is the natural quotient map.

Lemma 2.2.7. *If*

$$0 \to L \xrightarrow{f} M \xrightarrow{g} N \to 0$$

is a short exact sequence of R-modules, then M is noetherian if and only if both L and N are noetherian.

Proof. First suppose that M is noetherian. Then L is a submodule of M, so L is noetherian. Let N' be a submodule of N; then the inverse image of N' in M is a submodule of M, so it is finitely generated, hence its image N' is also finitely generated. Thus N is noetherian as well.

Next assume nothing about M, but suppose that both L and N are noetherian. Suppose M' is a submodule of M; then $M_0 = f(L) \cap M'$ is isomorphic to a submodule of the noetherian module L, so M_0 is generated by finitely many elements $a_1, \ldots, a_n$. The quotient M'/M_0 is isomorphic (via g) to a submodule of the noetherian module N, so M'/M_0 is generated by finitely many elements $b_1, \ldots, b_m$. For each $i \leq m$, let c_i be a lift of b_i to M', modulo M_0. Then the elements $a_1, \ldots, a_n, c_1, \ldots, c_m$ generate M', for if $x \in M'$, then there is some element $y \in M_0$ such that $x-y$ is an R-linear combination of the c_i, and y is an R-linear combination of the a_i. $\square$

Proposition 2.2.8. *Suppose R is a noetherian ring. Then an R-module M is noetherian if and only if it is finitely generated.*

Proof. If M is noetherian then every submodule of M is finitely generated so M itself is finitely generated. Conversely, suppose M is finitely generated, say by elements $a_1, \ldots, a_n$. Then there is a surjective homomorphism from $R^n = R \oplus \cdots \oplus R$ to M that sends $(0, \ldots, 0, 1, 0, \ldots, 0)$ (1 in the ith factor) to a_i. Using Lemma 2.2.7 and exact sequences of R-modules such as $0 \to R \to R \oplus R \to R \to 0$, we see inductively that R^n is noetherian. Again by Lemma 2.2.7, homomorphic images of noetherian modules are noetherian, so M is noetherian. $\square$

Lemma 2.2.9. *Suppose $\varphi : R \to S$ is a surjective homomorphism of rings and R is noetherian. Then S is noetherian.*

Proof. The kernel of φ is an ideal I in R, and we have an exact sequence

$$0 \to I \to R \to S \to 0$$

with R noetherian. This is an exact sequence of R-modules, where S has the R-module structure induced from φ (if $r \in R$ and $s \in S$, then we define $rs = \varphi(r)s$).

By Lemma 2.2.7, it follows that S is a noetherian R-modules. Suppose J is an ideal of S. Since J is an R-submodule of S, if we view J as an R-module, then J is finitely generated. Since R acts on J through S, the R-generators of J are also S-generators of J, so J is finitely generated as an ideal. Thus S is noetherian. $\square$

Theorem 2.2.10 (Hilbert Basis Theorem). *If R is a noetherian ring and S is finitely generated as a ring over R, then S is noetherian. In particular, for any n the polynomial ring $R[x_1, \ldots, x_n]$ and any of its quotients are noetherian.*

Proof. Assume first that we have already shown that for any n the polynomial ring $R[x_1, \ldots, x_n]$ is noetherian. Suppose S is finitely generated as a ring over R, so there are generators $s_1, \ldots, s_n$ for S. Then the map $x_i \mapsto s_i$ extends uniquely to a surjective homomorphism $\pi : R[x_1, \ldots, x_n] \twoheadrightarrow S$, and Lemma 2.2.9 implies that S is noetherian.

The rings $R[x_1, \ldots, x_n]$ and $(R[x_1, \ldots, x_{n-1}])[x_n]$ are isomorphic, so it suffices to prove that if R is noetherian then $R[x]$ is also noetherian. (Our proof follows [Art91, §12.5].) Thus suppose I is an ideal of $R[x]$ and that R is noetherian. We will show that I is finitely generated.

Let A be the set of leading coefficients of polynomials in I. (The leading coefficient of a polynomial is the coefficient of the highest degree monomial, or 0 if the polynomial is 0; thus $3x^7 + 5x^2 - 4$ has leading coefficient 3.) We will first show that A is an ideal of R. Suppose $a, b \in A$ are nonzero with $a + b \neq 0$. Then there are polynomials f and g in I with leading coefficients a and b. If $\deg(f) \leq \deg(g)$, then $a+b$ is the leading coefficient of $x^{\deg(g)-\deg(f)} f + g$, so $a+b \in A$; the argument when $\deg(f) > \deg(g)$ is analogous. Suppose $r \in R$ and $a \in A$ with $ra \neq 0$. Then ra is the leading coefficient of rf, so $ra \in A$. Thus A is an ideal in R.

Since R is noetherian and A is an ideal of R, there exist nonzero $a_1, \ldots, a_n \in A$ that generate A as an ideal. Since A is the set of leading coefficients of elements of I, and the a_j are in A, we can choose for each $j \leq n$ an element $f_j \in I$ with leading coefficient a_j. By multipying the f_j by some power of x, we may assume that the f_j all have the same degree $d \geq 1$.

Let $S_{<d}$ be the set of elements of I that have degree strictly less than d. This set is closed under addition and under multiplication by elements of R, so $S_{<d}$ is a module over R. The module $S_{<d}$ is the submodule of the R-module of polynomials of degree less than n, which is noetherian by Proposition 2.2.8 because it is generated by $1, x, \ldots, x^{n-1}$. Thus $S_{<d}$ is finitely generated, and we may choose generators $h_1, \ldots, h_m$ for $S_{<d}$.

We finish by proving using induction on the degree that every $g \in I$ is an $R[x]$-linear combination of $f_1, \ldots, f_n, h_1, \ldots, h_m$. If $g \in I$ has degree 0, then $g \in S_{<d}$, since $d \geq 1$, so g is a linear combination of $h_1, \ldots, h_m$. Next suppose $g \in I$ has degree e, and that we have proven the statement for all elements of I of degree $< e$. If $e \leq d$, then $g \in S_{<d}$, so g is in the $R[x]$-ideal generated by $h_1, \ldots, h_m$. Next suppose that $e \geq d$. Then the leading coefficient b of g lies in the ideal A of leading coefficients of elements of I, so there exist $r_i \in R$ such that $b = r_1 a_1 + \cdots + r_n a_n$. Since f_i has

leading coefficient a_i, the difference $g - x^{e-d} r_i f_i$ has degree less than the degree e of g. By induction $g - x^{e-d} r_i f_i$ is an $R[x]$ linear combination of $f_1, \ldots, f_n, h_1, \ldots, h_m$, so g is also an $R[x]$ linear combination of $f_1, \ldots, f_n, h_1, \ldots, h_m$. Since each f_i and h_j lies in I, it follows that I is generated by $f_1, \ldots, f_n, h_1, \ldots, h_m$, so I is finitely generated, as required. □

2.2.1 The Ring $\mathbb{Z}$ is noetherian

The ring $\mathbb{Z}$ is noetherian since every ideal of $\mathbb{Z}$ is generated by one element.

Proposition 2.2.11. *Every ideal of the ring $\mathbb{Z}$ is principal.*

Proof. Suppose I is a nonzero ideal in $\mathbb{Z}$. Let d be the least positive element of I. Suppose that $a \in I$ is any nonzero element of I. Using the division algorithm, we write $a = dq + r$, where q is an integer and $0 \leq r < d$. We have $r = a - dq \in I$ and $r < d$, so our assumption that d is minimal implies that $r = 0$, hence $a = dq$ is in the ideal generated by d. Thus I is the principal ideal generated by d. □

Example 2.2.12. Let $I = (12, 18)$ be the ideal of $\mathbb{Z}$ generated by 12 and 18. If $n = 12a + 18b \in I$, with $a, b \in \mathbb{Z}$, then $6 \mid n$, since $6 \mid 12$ and $6 \mid 18$. Also, $6 = 18 - 12 \in I$, so $I = (6)$.

The ring $\mathbb{Z}$ in Sage is ZZ, which is Noetherian.

```
ZZ.is_noetherian()
```

```
True
```

We create the ideal I in Sage as follows, and note that it is principal:

```
I = ideal(12,18); I
```

```
Principal ideal (6) of Integer Ring
```

```
I.is_principal()
```

```
True
```

We could also create I as follows:

```
ZZ.ideal(12,18)
```

```
Principal ideal (6) of Integer Ring
```

Propositions 2.2.8 and 2.2.11 together imply that any finitely generated abelian group is noetherian. This means that subgroups of finitely generated abelian groups are finitely generated, which provides the missing step in our proof of the structure theorem for finitely generated abelian groups.

Exercise 2.2.13. There is another way to show every principle ideal domain (for example $\mathbb{Z}$) is noetherian (contrast to the proof in Section 2.2.1). Let R be a PID and (a) an arbitrary ideal. Use the facts that $(b) \supseteq (a)$ if and only if $b \mid a$ and R is a UFD to show that ascending chain of ideals starting with (a) must stabilize.

2.3 Rings of Algebraic Integers

In this section we introduce the central objects of this book, which are the rings of algebraic integers. These are noetherian rings with an enormous amount of structure. We also introduce a function field analogue of these rings.

An *algebraic number* is a root of some nonzero polynomial $f(x) \in \mathbb{Q}[x]$. For example, $\sqrt{2}$ and $\sqrt{5}$ are both algebraic numbers, being roots of $x^2 - 2$ and $x^2 - 5$, respectively. But is $\sqrt{2} + \sqrt{5}$ necessarily the root of some polynomial in $\mathbb{Q}[x]$? This isn't quite so obvious.

Proposition 2.3.1. *An element α of a field extension of $\mathbb{Q}$ is an algebraic number if and only if the ring $\mathbb{Q}[\alpha]$ generated by α is finite dimensional as a $\mathbb{Q}$ vector space.*

Proof. Suppose α is an algebraic number, so there is a nonzero polynomial $f(x) \in \mathbb{Q}[x]$, so that $f(\alpha) = 0$. The equation $f(\alpha) = 0$ implies that $\alpha^{\deg(f)}$ can be written in terms of smaller powers of α, so $\mathbb{Q}[\alpha]$ is spanned by the finitely many numbers $1, \alpha, \ldots, \alpha^{\deg(f)-1}$, hence finite dimensional. Conversely, suppose $\mathbb{Q}[\alpha]$ is finite dimensional. Then for some $n \geq 1$, we have that α^n is in the $\mathbb{Q}$-vector space spanned by $1, \alpha, \ldots, \alpha^{n-1}$. Thus α satisfies a polynomial $f(x) \in \mathbb{Q}[x]$ of degree n. □

Proposition 2.3.2. *Suppose K is a field and $\alpha, \beta \in K$ are two algebraic numbers. Then $\alpha\beta$ and $\alpha + \beta$ are also algebraic numbers.*

Proof. Let $f, g \in \mathbb{Q}[x]$ be polynomials that are satisfied by α, β, respectively. The subring $\mathbb{Q}[\alpha, \beta] \subset K$ is a $\mathbb{Q}$-vector space that is spanned by the numbers $\alpha^i\beta^j$, where $0 \leq i < \deg(f)$ and $0 \leq j < \deg(g)$. Thus $\mathbb{Q}[\alpha, \beta]$ is finite dimensional, and since $\alpha + \beta$ and $\alpha\beta$ are both in $\mathbb{Q}[\alpha, \beta]$, we conclude by Proposition 2.3.1 that both are algebraic numbers. □

Suppose C is a field extension of $\mathbb{Q}$ such that every polynomial $f(x) \in \mathbb{Q}[x]$ factors completely in C. The algebraic closure $\overline{\mathbb{Q}}$ of $\mathbb{Q}$ inside C is the field generated by all roots in C of polynomials in $\mathbb{Q}[x]$. The fundamental theorem of algebra tells us that $C = \mathbb{C}$ is one choice of field C as above. There are other fields C, e.g., constructed using p-adic numbers. One can show that any two choices of $\overline{\mathbb{Q}}$ are isomorphic; however, there will be *many* isomorphisms between them.

Definition 2.3.3 (Algebraic Integer)**.** An element $\alpha \in \overline{\mathbb{Q}}$ is an *algebraic integer* if it is a root of some monic polynomial with coefficients in $\mathbb{Z}$.

For example, $\sqrt{2}$ is an algebraic integer, since it is a root of the monic integral polynomial $x^2 - 2$. As we will see below, $1/2$ is not an algebraic integer.

The following two propositions are analogous to Propositions 2.3.1–2.3.2 above, with the proofs replacing basic facts about vector spaces with facts we proved above about noetherian rings and modules.

Proposition 2.3.4. *An element $\alpha \in \overline{\mathbb{Q}}$ is an algebraic integer if and only if $\mathbb{Z}[\alpha]$ is finitely generated as a $\mathbb{Z}$-module.*

Proof. Suppose α is integral and let $f \in \mathbb{Z}[x]$ be a monic integral polynomial such that $f(\alpha) = 0$. Then, as a $\mathbb{Z}$-module, $\mathbb{Z}[\alpha]$ is generated by $1, \alpha, \alpha^2, \ldots, \alpha^{d-1}$, where d is the degree of f. Conversely, suppose $\alpha \in \overline{\mathbb{Q}}$ is such that $\mathbb{Z}[\alpha]$ is finitely generated as a module over $\mathbb{Z}$, say by elements $f_1(\alpha), \ldots, f_n(\alpha)$. Let d be any integer bigger than the degrees of all f_i. Then there exist integers a_i such that $\alpha^d = \sum_{i=1}^n a_i f_i(\alpha)$, hence α satisfies the monic polynomial $x^d - \sum_{i=1}^n a_i f_i(x) \in \mathbb{Z}[x]$, so α is an algebraic integer. $\square$

The proof of the following proposition uses repeatedly that any submodule of a finitely generated $\mathbb{Z}$-module is finitely generated, which uses that $\mathbb{Z}$ is noetherian and that finitely generated modules over a noetherian ring are noetherian.

Proposition 2.3.5. *Suppose K is a field and $\alpha, \beta \in K$ are two algebraic integers. Then $\alpha\beta$ and $\alpha + \beta$ are also algebraic integers.*

Proof. Let m, n be the degrees of monic integral polynomials that have α, β as roots, respectively. Then we can write α^m in terms of smaller powers of α and likewise for β^n, so the elements $\alpha^i \beta^j$ for $0 \leq i < m$ and $0 \leq j < n$ span the $\mathbb{Z}$-module $\mathbb{Z}[\alpha, \beta]$. Since $\mathbb{Z}[\alpha + \beta]$ is a submodule of the finitely-generated $\mathbb{Z}$-module $\mathbb{Z}[\alpha, \beta]$, it is finitely generated, so $\alpha + \beta$ is integral. Likewise, $\mathbb{Z}[\alpha\beta]$ is a submodule of $\mathbb{Z}[\alpha, \beta]$, so it is also finitely generated, and $\alpha\beta$ is integral. $\square$

2.3.1 Minimal Polynomials

Definition 2.3.6 (Minimal Polynomial)**.** The *minimal polynomial* of $\alpha \in \overline{\mathbb{Q}}$ is the monic polynomial $f \in \mathbb{Q}[x]$ of least positive degree such that $f(\alpha) = 0$.

It is a consequence of Lemma 2.3.9 below that "the" minimal polynomial of α is unique. The minimal polynomial of $1/2$ is $x - 1/2$, and the minimal polynomial of $\sqrt[3]{2}$ is $x^3 - 2$.

Example 2.3.7. We compute the minimal polynomial of a number expressed in terms of $\sqrt[4]{2}$:

```
k.<a> = NumberField(x^4 - 2)
a^4
```

```
2
```

```
(a^2 + 3).minpoly()
```

```
x^2 - 6*x + 7
```

Exercise 2.3.8. Find the minimal polynomial of $\sqrt{2}+\sqrt{3}$ by hand. Check your result with Sage.

Lemma 2.3.9. *Suppose $\alpha \in \overline{\mathbb{Q}}$. Then the minimal polynomial of α divides any polynomial h such that $h(\alpha) = 0$.*

Proof. Let f be a choice of minimal polynomial of α, as in Definition 2.3.6, and let h be a polynomial with $h(\alpha) = 0$. Use the division algorithm to write $h = qf + r$, where $0 \leq \deg(r) < \deg(f)$. We have

$$r(\alpha) = h(\alpha) - q(\alpha)f(\alpha) = 0,$$

so α is a root of r. However, f is a polynomial of least positive degree with root α, so $r = 0$. $\square$

Exercise 2.3.10. Show that the minimal polynomial of an algebraic number $\alpha \in \overline{\mathbb{Q}}$ is unique.

Lemma 2.3.11. *Suppose $\alpha \in \overline{\mathbb{Q}}$. Then α is an algebraic integer if and only if the minimal polynomial f of α has coefficients in $\mathbb{Z}$.*

Proof. ($\Longleftarrow$) Since $f \in \mathbb{Z}[x]$ is monic and $f(\alpha) = 0$, we see immediately that α is an algebraic integer.

($\Longrightarrow$) Since α is an algebraic integer, there is some nonzero monic $g \in \mathbb{Z}[x]$ such that $g(\alpha) = 0$. By Lemma 2.3.9, we have $g = fh$, for some $h \in \mathbb{Q}[x]$, and h is monic because f and g are. If $f \notin \mathbb{Z}[x]$, then some prime p divides the denominator of some coefficient of f. Let p^i be the largest power of p that divides some denominator of some coefficient f, and likewise let p^j be the largest power of p that divides some denominator of a coefficient of h. Then $p^{i+j}g = (p^i f)(p^j h)$, and if we reduce both sides modulo p, then the left hand side is 0 but the right hand side is a product of two nonzero polynomials in $\mathbb{F}_p[x]$, hence nonzero, a contradiction. $\square$

Exercise 2.3.12. Which of the following numbers are algebraic integers?

1. The number $(1+\sqrt{5})/2$.
2. The number $(2+\sqrt{5})/2$.
3. The value of the infinite sum $\sum_{n=1}^{\infty} 1/n^2$.
4. The number $\alpha/3$, where α is a root of $x^4 + 54x + 243$.

Example 2.3.13. We compute some minimal polynomials in Sage. The minimal polynomial of $1/2$:

```
(1/2).minpoly()
```

```
x - 1/2
```

We construct a root a of $x^2 - 2$ and compute its minimal polynomial:

```
k.<a> = NumberField(x^2 - 2)
a^2 - 2
```

```
0
```

```
a.minpoly()
```

```
x^2 - 2
```

Finally we compute the minimal polynomial of $\alpha = \sqrt{2}/2 + 3$, which is not integral, hence Proposition 2.3.4 implies that α is not an algebraic integer:

```
(a/2 + 3).minpoly()
```

```
x^2 - 6*x + 17/2
```

The only elements of $\mathbb{Q}$ that are algebraic integers are the usual integers $\mathbb{Z}$, since $\mathbb{Z}[1/d]$ is not finitely generated as a $\mathbb{Z}$-module. Watch out since there are elements of $\overline{\mathbb{Q}}$ that seem to *appear* to have denominators when written down, but are still algebraic integers. This is an artifact of how we write them down, e.g., if we wrote our integers as a multiple of $\alpha = 2$, then we would write 1 as $\alpha/2$. For example,

$$\alpha = \frac{1 + \sqrt{5}}{2}$$

is an algebraic integer, since it is a root of the monic integral polynomial $x^2 - x - 1$. We verify this using Sage below, though of course this is easy to do by hand (you should try much more complicated examples in Sage).

```
k.<a> = QuadraticField(5)
a^2
```

```
5
```

```
alpha = (1 + a)/2
alpha.minpoly()
```

```
x^2 - x - 1
```

```
alpha.is_integral()
```

```
True
```

Since $\sqrt{5}$ can be expressed in terms of radicals, we can also compute this minimal polynomial using the symbolic functionality in Sage.

```
alpha = (1+sqrt(5))/2
alpha.minpoly()
```

```
x^2 - x - 1
```

Here is a more complicated example using a similar approach:

```
alpha = sqrt(2) + 3^(1/4)
alpha.minpoly()
```

```
x^8 - 8*x^6 + 18*x^4 - 104*x^2 + 1
```

Example 2.3.14. We illustrate an example of a sum and product of two algebraic integers being an algebraic integer. We first make the relative number field obtained by adjoining a root of $x^3 - 5$ to the field $\mathbb{Q}(\sqrt{2})$:

```
k.<a, b> = NumberField([x^2 - 2, x^3 - 5])
k
```

```
Number Field in a with defining polynomial x^2 + -2 over its base field
```

Here a and b are roots of $x^2 - 2$ and $x^3 - 5$, respectively.

```
a^2
```

```
2
```

```
b^3
```

```
5
```

We compute the minimal polynomial of the sum and product of $\sqrt[3]{5}$ and $\sqrt{2}$. The command `absolute_minpoly` gives the minimal polynomial of the element over the rational numbers.

```
(a+b).absolute_minpoly()
```

```
x^6 - 6*x^4 - 10*x^3 + 12*x^2 - 60*x + 17
```

```
(a*b).absolute_minpoly()
```

```
x^6 - 200
```

The minimal polynomial of the product is $\sqrt[3]{5}\sqrt{2}$ is trivial to compute by hand. In light of the Cayley-Hamilton theorem, we can compute the minimal polynomial of

$\alpha = \sqrt[3]{5} + \sqrt{2}$ by hand by computing the determinant of the matrix given by left multiplication by α on the basis

$$1, \sqrt{2}, \sqrt[3]{5}, \sqrt[3]{5}\sqrt{2}, \sqrt[3]{5}^2, \sqrt[3]{5}^2\sqrt{2}.$$

The following is an alternative, more symbolic way to compute the minimal polynomials above, though it is not provably correct. We compute α to 100 bits precision (via the `n` command), then use the LLL algorithm (via the `algdep` command) to heuristically find a linear relation between the first 6 powers of α (see Section 2.5 below for more about LLL).

```
a = 5^(1/3); b = sqrt(2)
c = a+b; c
```

```
5^(1/3) + sqrt(2)
```

```
(a+b).n(100).algdep(6)
```

```
x^6 - 6*x^4 - 10*x^3 + 12*x^2 - 60*x + 17
```

```
(a*b).n(100).algdep(6)
```

```
x^6 - 200
```

Exercise 2.3.15. Let $\alpha = \sqrt{2} + \frac{1+\sqrt{5}}{2}$.

1. Is α an algebraic integer?

2. Explicitly write down the minimal polynomial of α as an element of $\mathbb{Q}[x]$.

2.3.2 Number fields, rings of integers, and orders

Definition 2.3.16 (Number field)**.** A *number field* is a field K that contains the rational numbers $\mathbb{Q}$ such that the degree $[K : \mathbb{Q}] = \dim_{\mathbb{Q}}(K)$ is finite.

If K is a number field, then by the primitive element theorem there is an $\alpha \in K$ so that $K = \mathbb{Q}(\alpha)$. Let $f(x) \in \mathbb{Q}[x]$ be the minimal polynomial of α. Fix a choice of algebraic closure $\overline{\mathbb{Q}}$ of $\mathbb{Q}$. Associated to each of the $\deg(f)$ roots $\alpha' \in \overline{\mathbb{Q}}$ of f, we obtain a field embedding $K \hookrightarrow \overline{\mathbb{Q}}$ that sends α to α'. Thus any number field can be embedded in $[K : \mathbb{Q}] = \deg(f)$ distinct ways in $\overline{\mathbb{Q}}$.

Definition 2.3.17 (Ring of Integers)**.** The *ring of integers* of a number field K is the ring

$$\mathcal{O}_K = \{x \in K : \; x \text{ satisfies a monic polynomial with integer coefficients}\}.$$

Proposition 2.3.5 implies that $\mathcal{O}_K$ is a ring.

Example 2.3.18. The field $\mathbb{Q}$ of rational numbers is a number field of degree 1, and the ring of integers of $\mathbb{Q}$ is $\mathbb{Z}$. The field $K = \mathbb{Q}(i)$ of Gaussian integers has degree 2 and $\mathcal{O}_K = \mathbb{Z}[i]$.

Example 2.3.19. The golden ratio $\varphi = (1+\sqrt{5})/2$ is in the quadratic number field $K = \mathbb{Q}(\sqrt{5}) = \mathbb{Q}(\varphi)$; notice that φ satisfies $x^2 - x - 1$, so $\varphi \in \mathcal{O}_K$. To see that $\mathcal{O}_K = \mathbb{Z}[\varphi]$ directly, we proceed as follows. By Proposition 2.3.4, the algebraic integers K are exactly the elements $a + b\sqrt{5} \in K$, with $a, b \in \mathbb{Q}$ that have integral minimal polynomial. The matrix of $a + b\sqrt{5}$ with respect to the basis $1, \sqrt{5}$ for K is $m = \left(\begin{smallmatrix} a & 5b \\ b & a \end{smallmatrix}\right)$. The characteristic polynomial of m is $f = (x-a)^2 - 5b^2 = x^2 - 2ax + a^2 - 5b^2$, which is in $\mathbb{Z}[x]$ if and only if $2a \in \mathbb{Z}$ and $a^2 - 5b^2 \in \mathbb{Z}$. Thus $a = a'/2$ with $a' \in \mathbb{Z}$, and $(a'/2)^2 - 5b^2 \in \mathbb{Z}$, so $5b^2 \in \frac{1}{4}\mathbb{Z}$, so $b \in \frac{1}{2}\mathbb{Z}$ as well. If a has a denominator of 2, then b must also have a denominator of 2 to ensure that the difference $a^2 - 5b^2$ is an integer. This proves that $\mathcal{O}_K = \mathbb{Z}[\varphi]$.

Example 2.3.20. The ring of integers of $K = \mathbb{Q}(\sqrt[3]{9})$ is $\mathbb{Z}[\sqrt[3]{3}]$, where $\sqrt[3]{3} = \frac{1}{3}(\sqrt[3]{9})^2 \notin \sqrt[3]{9}$. As we will see, in general the problem of computing $\mathcal{O}_K$ given K may be very hard, since it requires factoring a certain potentially large integer.

Exercise 2.3.21. From basic definitions, find the rings of integers of the fields $\mathbb{Q}(\sqrt{11})$ and $\mathbb{Q}(\sqrt{-6})$.

Definition 2.3.22 (Order)**.** An *order* in $\mathcal{O}_K$ is any subring R of $\mathcal{O}_K$ such that the quotient $\mathcal{O}_K/R$ of abelian groups is finite. (By definition R must contain 1 because it is a ring.)

As noted above, $\mathbb{Z}[i]$ is the ring of integers of $\mathbb{Q}(i)$. For every nonzero integer n, the subring $\mathbb{Z}+ni\mathbb{Z}$ of $\mathbb{Z}[i]$ is an order. The subring $\mathbb{Z}$ of $\mathbb{Z}[i]$ is not an order, because $\mathbb{Z}$ does not have finite index in $\mathbb{Z}[i]$. Also the subgroup $2\mathbb{Z} + i\mathbb{Z}$ of $\mathbb{Z}[i]$ is not an order because it is not a ring.

Exercise 2.3.23. Let K be a quadratic extension of $\mathbb{Q}$ and R be any order in $\mathcal{O}_K$. Show that $\mathcal{O}_K/R$ is cyclic as an abelian group and that there is a bijection between orders of $\mathcal{O}_K$ containing R and divisors of $[\mathcal{O}_K : R]$.

Remark 2.3.24. Exercise 2.3.23 is used in elliptic curve cryptography to measure the number of isogenies, see [**?**, §11.2] for an example.

We define the number field $\mathbb{Q}(i)$ and compute its ring of integers.

```
K.<i> = NumberField(x^2 + 1)
OK = K.ring_of_integers(); OK
```

```
Order with module basis 1, i in Number Field in i with
defining polynomial x^2 + 1
```

Next we compute the order $\mathbb{Z} + 3i\mathbb{Z}$.

```
O3 = K.order(3*i); O3
```

```
Order with module basis 1, 3*i in Number Field in i with
defining polynomial x^2 + 1
```

```
O3.gens()
```

```
[1, 3*i]
```

We test whether certain elements are in the order.

```
5 + 9*i in O3
```

```
True
```

```
1 + 2*i in O3
```

```
False
```

We will frequently consider orders because they are often much easier to write down explicitly than $\mathcal{O}_K$. For example, if $K = \mathbb{Q}(\alpha)$ and α is an algebraic integer, then $\mathbb{Z}[\alpha]$ is an order in $\mathcal{O}_K$, but frequently $\mathbb{Z}[\alpha] \neq \mathcal{O}_K$.

Example 2.3.25. In this example $[\mathcal{O}_K : \mathbb{Z}[a]] = 2197$. First we define the number field $K = \mathbb{Q}(a)$ where a is a root of $x^3 - 15x^2 - 94x - 3674$, then we compute the order $\mathbb{Z}[a]$ generated by a.

```
K.<a> = NumberField(x^3 - 15*x^2 - 94*x - 3674)
Oa = K.order(a); Oa
```

```
Order with module basis 1, a, a^2 in Number Field in a with defining
polynomial x^3 - 15*x^2 - 94*x - 3674
```

```
Oa.basis()
```

```
[1, a, a^2]
```

Next we compute a $\mathbb{Z}$-basis for the maximal order $\mathcal{O}_K$ of K, and compute that the index of $\mathbb{Z}[a]$ in $\mathcal{O}_K$ is $2197 = 13^3$.

```
OK = K.maximal_order()
OK.basis()
```

```
[25/169*a^2 + 10/169*a + 1/169, 5/13*a^2 + 1/13*a, a^2]
```

```
Oa.index_in(OK)
```

```
2197
```

Lemma 2.3.26. *Let $\mathcal{O}_K$ be the ring of integers of a number field. Then $\mathcal{O}_K \cap \mathbb{Q} = \mathbb{Z}$ and $\mathbb{Q}\mathcal{O}_K = K$.*

Proof. Suppose $\alpha \in \mathcal{O}_K \cap \mathbb{Q}$ with $\alpha = a/b \in \mathbb{Q}$ in lowest terms and $b > 0$. Since α is integral, $\mathbb{Z}[a/b]$ is finitely generated as a module, so $b = 1$.

To prove that $\mathbb{Q}\mathcal{O}_K = K$, suppose $\alpha \in K$, and let $f(x) \in \mathbb{Q}[x]$ be the minimal monic polynomial of α. For any positive integer d, the minimal monic polynomial of $d\alpha$ is $d^{\deg(f)} f(x/d)$, i.e., the polynomial obtained from $f(x)$ by multiplying the coefficient of $x^{\deg(f)}$ by 1, multiplying the coefficient of $x^{\deg(f)-1}$ by d, multiplying the coefficient of $x^{\deg(f)-2}$ by d^2, etc. If d is the least common multiple of the denominators of the coefficients of f, then the minimal monic polynomial of $d\alpha$ has integer coefficients, so $d\alpha$ is integral and $d\alpha \in \mathcal{O}_K$. This proves that $\mathbb{Q}\mathcal{O}_K = K$. $\square$

Exercise 2.3.27. Which are the following rings are orders in the given number field, i.e. orders in the ring of integers of the given number field.

1. The ring $R = \mathbb{Z}[i]$ in the number field $\mathbb{Q}(i)$.
2. The ring $R = \mathbb{Z}[i/2]$ in the number field $\mathbb{Q}(i)$.
3. The ring $R = \mathbb{Z}[17i]$ in the number field $\mathbb{Q}(i)$.
4. The ring $R = \mathbb{Z}[i]$ in the number field $\mathbb{Q}(\sqrt[4]{-1})$.

2.3.3 Function fields

Let k be any field. We can also make the same definitions, but with $\mathbb{Q}$ replaced by the field $k(t)$ of rational functions in an indeterminate t, and $\mathbb{Z}$ replaced by $k[t]$. The analogue of a number field is called a *function field*; it is a finite algebraic extension field K of $k(t)$. Elements of K have a unique minimal polynomial as above, and the ring of integers of K consists of those elements whose monic minimal polynomial has coefficients in the polynomial ring $k[t]$.

Geometrically, if $F(x,t) = 0$ is an affine equation that defines (via projective closure) a nonsingular projective curve C, then $K = k(t)[x]/(F(x,t))$ is a function field. We view the field K as the field of all rational functions on the projective closure of the curve C. The ring of integers $\mathcal{O}_K$ is the subring of rational functions that have no poles on the affine curve $F(x,t) = 0$, though they may have poles at infinity, i.e., at the extra points we introduce when passing to the projective closure C. The algebraic arguments we gave above prove that $\mathcal{O}_K$ is a ring. This is also geometrically intuitive, since the sum and product of two functions with no poles also have no poles.

Exercise 2.3.28. Let $k = \mathbb{F}_p$ be the finite field with p elements where p is some prime. Find all automorphisms of $k(t)$. Note that an automorphism is completely characterized by its value on t. How many such automorphisms are there?

[*Hint*: For some people, it is easier to think about the equivalent question: What rational functions $f \in k(t)$ is the map $k(t) \to k(t)$ given by $t \mapsto f(t)$ an automorphism?]

2.4 Norms and Traces

In this section we develop some basic properties of norms, traces, and discriminants, and give more properties of rings of integers in the general context of Dedekind domains.

Before discussing norms and traces we introduce some notation for field extensions. If $K \subset L$ are number fields, we let $[L : K]$ denote the dimension of L viewed as a K-vector space. If K is a number field and $a \in \overline{\mathbb{Q}}$, let $K(a)$ be the extension of K generated by a, which is the smallest number field that contains both K and a. If $a \in \overline{\mathbb{Q}}$ then a has a minimal polynomial $f(x) \in \mathbb{Q}[x]$, and the *Galois conjugates* of a are the roots of f. These are called the Galois conjugates because they are the orbit of a under the action of $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$.

Example 2.4.1. The element $\sqrt{2}$ has minimal polynomial $x^2 - 2$ and the Galois conjugates of $\sqrt{2}$ are $\sqrt{2}$ and $-\sqrt{2}$. The cube root $\sqrt[3]{2}$ has minimial polynomial $x^3 - 2$ and three Galois conjugates $\sqrt[3]{2}, \zeta_3\sqrt[3]{2}, \zeta_3^2\sqrt[3]{2}$, where ζ_3 is a cube root of unity.

We create the extension $\mathbb{Q}(\zeta_3)(\sqrt[3]{2})$ in Sage.

```
L.<cuberoot2> = CyclotomicField(3).extension(x^3 - 2)
cuberoot2^3
```

```
2
```

Then we list the Galois conjugates of $\sqrt[3]{2}$.

```
cuberoot2.galois_conjugates(L)
```

```
[cuberoot2, (-zeta3 - 1)*cuberoot2, zeta3*cuberoot2]
```

Note that $\zeta_3^2 = -\zeta_3 - 1$:

```
zeta3 = L.base_field().0
zeta3^2
```

```
-zeta3 - 1
```

Suppose $K \subset L$ is an inclusion of number fields and let $a \in L$. Then left multiplication by a defines a K-linear transformation $\ell_a : L \to L$. (The transformation ℓ_a is K-linear because L is commutative.)

Definition 2.4.2 (Norm and Trace)**.** The *norm* and *trace* of a from L to K are

$$\mathrm{Norm}_{L/K}(a) = \det(\ell_a) \quad \text{and} \quad \mathrm{tr}_{L/K}(a) = \mathrm{tr}(\ell_a).$$

We know from linear algebra that determinants are multiplicative and traces are additive, so for $a, b \in L$ we have

$$\mathrm{Norm}_{L/K}(ab) = \mathrm{Norm}_{L/K}(a) \cdot \mathrm{Norm}_{L/K}(b)$$

and

$$\mathrm{tr}_{L/K}(a+b) = \mathrm{tr}_{L/K}(a) + \mathrm{tr}_{L/K}(b).$$

Note that if $f \in \mathbb{Q}[x]$ is the characteristic polynomial of ℓ_a, then the constant term of f is $(-1)^{\deg(f)} \det(\ell_a)$, and the coefficient of $x^{\deg(f)-1}$ is $-\mathrm{tr}(\ell_a)$.

Proposition 2.4.3. *Let $a \in L$ and let $\sigma_1, \ldots, \sigma_d$, where $d = [L : K]$, be the distinct field embeddings $L \hookrightarrow \overline{\mathbb{Q}}$ that fix every element of K. Then*

$$\mathrm{Norm}_{L/K}(a) = \prod_{i=1}^{d} \sigma_i(a) \quad \textit{and} \quad \mathrm{tr}_{L/K}(a) = \sum_{i=1}^{d} \sigma_i(a).$$

Proof. We prove the proposition by computing the characteristic polynomial of a. Let $f \in K[x]$ be the minimal polynomial of a over K, and note that f has distinct roots and is irreducible, since it is the polynomial in $K[x]$ of least degree that is satisfied by a and K has characteristic 0. Since f is irreducible, we have $K(a) \cong K[x]/(f)$, so $[K(a) : K] = \deg(f)$. Also a satisfies a polynomial if and only if ℓ_a does, so the characteristic polynomial of ℓ_a acting on $K(a)$ is f. Let $b_1, \ldots, b_n$ be a basis for L over $K(a)$ and note that $1, \ldots, a^m$ is a basis for $K(a)/K$, where $m = \deg(f) - 1$. Then $a^i b_j$ is a basis for L over K, and left multiplication by a acts the same way on the span of $b_j, ab_j, \ldots, a^m b_j$ as on the span of $b_k, ab_k, \ldots, a^m b_k$, for any pair $j, k \le n$. Thus the matrix of ℓ_a on L is a block direct sum of copies of the matrix of ℓ_a acting on $K(a)$, so the characteristic polynomial of ℓ_a on L is $f^{[L:K(a)]}$. The proposition follows because the roots of $f^{[L:K(a)]}$ are exactly the images $\sigma_i(a)$, with multiplicity $[L : K(a)]$, since each embedding of $K(a)$ into $\overline{\mathbb{Q}}$ extends in exactly $[L : K(a)]$ ways to L. □

It is important in Proposition 2.4.3 that the product and sum be over *all* the images $\sigma_i(a)$, not over just the distinct images. For example, if $a = 1 \in L$, then $\mathrm{Tr}_{L/K}(a) = [L : K]$, whereas the sum of the distinct conjugates of a is 1.

The following corollary asserts that the norm and trace behave well in towers.

Corollary 2.4.4. *Suppose $K \subset L \subset M$ is a tower of number fields, and let $a \in M$. Then*

$$\mathrm{Norm}_{M/K}(a) = \mathrm{Norm}_{L/K}(\mathrm{Norm}_{M/L}(a)) \quad \textit{and} \quad \mathrm{tr}_{M/K}(a) = \mathrm{tr}_{L/K}(\mathrm{tr}_{M/L}(a)).$$

Proof. The proof uses that every embedding $L \hookrightarrow \overline{\mathbb{Q}}$ extends in exactly $[M : L]$ way to an embedding $M \hookrightarrow \overline{\mathbb{Q}}$. This is clear if we view M as $L[x]/(h(x))$ for some irreducicble polynomial $h(x) \in L[x]$ of degree $[M : L]$, and note that the extensions of $L \hookrightarrow \overline{\mathbb{Q}}$ to M correspond to the roots of h, of which there are $\deg(h)$, since $\overline{\mathbb{Q}}$ is algebraically closed.

For the first equation, both sides are the product of $\sigma_i(a)$, where σ_i runs through the embeddings of M into $\overline{\mathbb{Q}}$ that fix K. To see this, suppose $\sigma : L \to \overline{\mathbb{Q}}$ fixes K. If σ' is an extension of σ to M, and $\tau_1, \ldots, \tau_d$ are the embeddings of M into $\overline{\mathbb{Q}}$ that fix L, then $\sigma'\tau_1, \ldots, \sigma'\tau_d$ are exactly the extensions of σ to M. For the second statement, both sides are the sum of the $\sigma_i(a)$. □

Let $K \subset L \subset M$ be as in Corollary 2.4.4. If $\alpha \in M$, then the formula of Proposition 2.4.3 implies that the norm and trace down to L of α is an element of $\mathcal{O}_L$, because the sum and product of algebraic integers is an algebraic integer.

Proposition 2.4.5. *Let K be a number field. The ring of integers $\mathcal{O}_K$ is a lattice in K, i.e., $\mathbb{Q}\mathcal{O}_K = K$ and $\mathcal{O}_K$ is an abelian group of rank $[K : \mathbb{Q}]$.*

Proof. We saw in Lemma 2.3.26 that $\mathbb{Q}\mathcal{O}_K = K$. Thus there exists a basis $a_1, \ldots, a_n$ for K, where each a_i is in $\mathcal{O}_K$. Suppose that as $x = \sum_{i=1}^n c_i a_i \in \mathcal{O}_K$ varies over all elements of $\mathcal{O}_K$ the denominators of the coefficients c_i are not all uniformly bounded. Then subtracting off integer multiples of the a_i, we see that as $x = \sum_{i=1}^n c_i a_i \in \mathcal{O}_K$ varies over elements of $\mathcal{O}_K$ with c_i between 0 and 1, the denominators of the c_i are also arbitrarily large. This implies that there are infinitely many elements of $\mathcal{O}_K$ in the bounded subset

$$S = \{c_1 a_1 + \cdots + c_n a_n : c_i \in \mathbb{Q},\ 0 \leq c_i \leq 1\} \subset K.$$

Thus for any $\varepsilon > 0$, there are elements $a, b \in \mathcal{O}_K$ such that the coefficients of $a - b$ are all less than ε (otherwise the elements of $\mathcal{O}_K$ would all be a “distance” of least ε from each other, so only finitely many of them would fit in S).

As mentioned above, the norms of elements of $\mathcal{O}_K$ are integers. Since the norm of an element is the determinant of left multiplication by that element, the norm is a homogenous polynomial of degree n in the indeterminate coefficients c_i, which is 0 only on the element 0, so the constant term of this polynomial is 0. If the c_i get arbitrarily small for elements of $\mathcal{O}_K$, then the values of the norm polynomial get arbitrarily small, which would imply that there are elements of $\mathcal{O}_K$ with positive norm too small to be in $\mathbb{Z}$, a contradiction. So the set S contains only finitely many elements of $\mathcal{O}_K$. Thus the denominators of the c_i are bounded, so for some d, we have that $\mathcal{O}_K$ has finite index in $A = \frac{1}{d}\mathbb{Z}a_1 + \cdots + \frac{1}{d}\mathbb{Z}a_n$. Since A is isomorphic to $\mathbb{Z}^n$, it follows from the structure theorem for finitely generated abelian groups that $\mathcal{O}_K$ is isomorphic as a $\mathbb{Z}$-module to $\mathbb{Z}^n$, as claimed. □

Corollary 2.4.6. *The ring of integers $\mathcal{O}_K$ of a number field is noetherian.*

Proof. By Proposition 2.4.5, the ring $\mathcal{O}_K$ is finitely generated as a module over $\mathbb{Z}$, so it is certainly finitely generated as a ring over $\mathbb{Z}$. By Theorem 2.2.10, $\mathcal{O}_K$ is noetherian. □

2.5 Recognizing Algebraic Numbers using LLL

Suppose we somehow compute a decimal approximation α to some rational number $\beta \in \mathbb{Q}$ and from this wish to recover β. For concreteness, say

$$\beta = 22/389 = 0.056555269922879177377892030848329048843187660668380 46\ldots$$

and we compute

$$\alpha = 0.056555.$$

Now suppose given only α that you would like to recover β. A standard technique is to use continued fractions, which yields a sequence of good rational approximations for α; by truncating right before a surprisingly big partial quotient, we obtain β:

```
v = continued_fraction(0.056555)
continued_fraction(0.056555)
```

```
[0, 17, 1, 2, 6, 1, 23, 1, 1, 1, 1, 1, 2]
```

```
convergents([0, 17, 1, 2, 6, 1])
```

```
[0, 1/17, 1/18, 3/53, 19/336, 22/389]
```

Generalizing this, suppose next that somehow you numerically approximate an algebraic number, e.g., by evaluating a special function and get a decimal approximation $\alpha \in \mathbb{C}$ to an algebraic number $\beta \in \overline{\mathbb{Q}}$. For concreteness, suppose $\beta = \frac{1}{3} + \sqrt[4]{3}$:

```
N(1/3 + 3^(1/4), digits=50)
```

```
1.6494073462858257941525522351303323884934019235391 6
```

Now suppose you very much want to find the (rescaled) minimal polynomial $f(x) \in \mathbb{Z}[x]$ of β just given this numerical approximation α. This is of great value even without proof, since often in practice once you know a potential minimal polynomial you can verify that it is in fact right. Exactly this situation arises in the explicit construction of class fields (a more advanced topic in number theory) and in the construction of Heegner points on elliptic curves. As we will see, the LLL algorithm provides a polynomial time way to solve this problem, assuming α has been computed to sufficient precision.

2.5.1 LLL Reduced Basis

Given a basis $b_1, \ldots, b_n$ for $\mathbb{R}^n$, the *Gramm-Schmidt orthogonalization* process produces an orthogonal basis $b_1^*, \ldots, b_n^*$ for $\mathbb{R}^n$ as follows. Define inductively

$$b_i^* = b_i - \sum_{j<i} \mu_{i,j} b_j^*$$

where

$$\mu_{i,j} = \frac{b_i \cdot b_j^*}{b_j^* \cdot b_j^*}.$$

Example 2.5.1. We compute the Gramm-Schmidt orthogonal basis of the rows of a matrix. Note that no square roots are introduced in the process; there would be square roots if we constructed an orthonormal basis.

```
A = matrix(ZZ, 2, [1,2, 3,4]); A
```

```
[1 2]
[3 4]
```

```
Bstar, mu = A.gramm_schmidt()
```

The rows of the matrix B^* are obtained from the rows of A by the Gramm-Schmidt procedure.

```
Bstar
```

```
[   1     2]
[ 4/5  -2/5]
```

```
mu
```

```
[   0     0]
[11/5     0]
```

A *lattice* $L \subset \mathbb{R}^n$ is a subgroup that is free of rank n such that $\mathbb{R}L = \mathbb{R}^n$.

Definition 2.5.2 (LLL-reduced basis)**.** The basis $b_1, \ldots, b_n$ for a lattice $L \subset \mathbb{R}^n$ is *LLL reduced* if for all i, j,

$$|\mu_{i,j}| \leq \frac{1}{2}$$

and for each $i \geq 2$,

$$|b_i^*|^2 \geq \left(\frac{3}{4} - \mu_{i,i-1}^2\right) |b_{i-1}^*|^2$$

For example, the basis $b_1 = (1, 2)$, $b_2 = (3, 4)$ for a lattice L is *not* LLL reduced because $b_1^* = b_1$ and

$$\mu_{2,1} = \frac{b_2 \cdot b_1^*}{b_1^* \cdot b_1^*} = \frac{11}{5} > \frac{1}{2}.$$

However, the basis $b_1 = (1, 0)$, $b_2 = (0, 2)$ for L is LLL reduced, since

$$\mu_{2,1} = \frac{b_2 \cdot b_1^*}{b_1^* \cdot b_1^*} = 0,$$

and

$$2^2 \geq (3/4) \cdot 1^2.$$

```
A = matrix(ZZ, 2, [1,2, 3,4])
A.LLL()
```

```
[1 0]
[0 2]
```

2.5.2 What LLL really means

The following theorem is not too difficult to prove.

Let $b_1, \ldots, b_n$ be an LLL reduced basis for a lattice $L \subset \mathbb{R}^n$. Let $d(L)$ denote the absolute value of the determinant of any matrix whose rows are basis for L. Then the vectors b_i are "nearly orthogonal" and "short" in the sense of the following theorem:

Theorem 2.5.3. *We have*

1. $d(L) \leq \prod_{i=1}^{n} |b_i| \leq 2^{n(n-1)/4} d(L)$.
2. *For* $1 \leq j \leq i \leq n$, *we have*
$$|b_j| \leq 2^{(i-1)/2} |b_i^*|.$$
3. *The vector* b_1 *is very short in the sense that*
$$|b_1| \leq 2^{(n-1)/4} d(L)^{1/n}$$
and for every nonzero $x \in L$ *we have*
$$|b_1| \leq 2^{(n-1)/2} |x|.$$
4. *More generally, for any linearly independent* $x_1, \ldots, x_t \in L$, *we have*
$$|b_j| \leq 2^{(n-1)/2} \max(|x_1|, \ldots, |x_t|)$$
for $1 \leq j \leq t$.

Perhaps the most amazing thing about the idea of an LLL reduced basis is that there is an algorithm (in fact many) that given a basis for a lattice L produce an LLL reduced basis for L, and do so *quickly*, i.e., in polynomial time in the number of digits of the input. The current optimal implementation (and practically optimal algorithms) for computing LLL reduced basis are due to Damien Stehle, and are included standard in Magma in Sage. Stehle's code is amazing – it can LLL reduce a random lattice in $\mathbb{R}^n$ for $n < 1000$ in a matter of minutes!

```
A = random_matrix(ZZ, 200)
t = cputime()
B = A.LLL()
cputime(t)      # random output
```

```
3.0494159999999999
```

There is even a very fast variant of Stehle's implementation that computes a basis for L that is very likely LLL reduced but may in rare cases fail to be LLL reduced.

```
t = cputime()
B = A.LLL(algorithm="fpLLL:fast")   # not tested
cputime(t)      # random output
```

```
0.96842699999999837
```

2.5.3 Applying LLL

The LLL definition and algorithm has many application in number theory, e.g., to cracking lattice-based cryptosystems, to enumerating all short vectors in a lattice, to finding relations between decimal approximations to complex numbers, to very fast univariate polynomial factorization in $\mathbb{Z}[x]$ and more generally in $K[x]$ where K is a number fields, and to computation of kernels and images of integer matrices. LLL can also be used to solve the problem of recognizing algebraic numbers mentioned at the beginning of Section 2.5.

Suppose as above that α is a decimal approximation to some algebraic number β, and to for simplicity assume that $\alpha \in \mathbb{R}$ (the general case of $\alpha \in \mathbb{C}$ is described in [Coh93]). We finish by explaining how to use LLL to find a polynomial $f(x) \in \mathbb{Z}[x]$ such that $f(\alpha)$ is small, hence has a shot at being the minimal polynomial of β.

Given a real number decimal approximation α, an integer d (the degree), and an integer K (a function of the precision to which α is known), the following steps produce a polynomial $f(x) \in \mathbb{Z}[x]$ of degree at most d such that $f(\alpha)$ is small.

1. Form the lattice in $\mathbb{R}^{d+2}$ with basis the rows of the matrix A whose first $(d+1)\times(d+1)$ part is the identity matrix, and whose last column has entries

   $$K, \lfloor K\alpha \rfloor, \lfloor K\alpha^2 \rfloor, \ldots, \lfloor K\alpha^d \rfloor. \tag{2.5.1}$$

   (Note this matrix is $(d+1) \times (d+2)$ so the lattice is not of full rank in $\mathbb{R}^{d+2}$, which isn't a problem, since the LLL definition also makes sense for less vectors.)

2. Compute an LLL reduced basis for the $\mathbb{Z}$-span of the rows of A, and let B be the corresponding matrix. Let $b_1 = (a_0, a_1, \ldots, a_{d+1})$ be the first row of B and notice that B is obtained from A by left multiplication by an invertible integer matrix. Thus $a_0, \ldots, a_d$ are the linear combination of the (2.5.1) that equals a_{d+1}. Moreover, since B is LLL reduced we expect that a_{d+1} is relatively small.

3. Output $f(x) = a_0 + a_1x + \cdots + a_dx^d$. We have that $f(\alpha) \sim a_{d+1}/K$, which is small. Thus $f(x)$ may be a very good candidate for the minimal polynomial of β (the algebraic number we are approximating), assuming d was chosen minimally and α was computed to sufficient precision.

The following is a complete implementation of the above algorithm in Sage:

```
def myalgdep(a, d, K=10^6):
    aa = [floor(K*a^i) for i in range(d+1)]
    A = identity_matrix(ZZ, d+1)
    B = matrix(ZZ, d+1, 1, aa)
    A = A.augment(B)
    L = A.LLL()
    v = L[0][:-1].list()
    return ZZ['x'](v)
```

Here is an example of using it:

```
R.<x> = RDF[]
f = 2*x^3 - 3*x^2 + 10*x - 4
a = f.roots()[0][0]; a
myalgdep(a, 3, 10^6)        # not tested
```

```
2*x^3 - 3*x^2 + 10*x - 4
```

Chapter 3

Unique Factorization of Ideals

Unique factorization into irreducible elements frequently fails for rings of integers of number fields. In this chapter we will deduce a central property of the ring of integers $\mathcal{O}_K$ of an algebraic number field, namely that every nonzero *ideal* factors uniquely as a products of prime ideals. Along the way, we will introduce fractional ideals and prove that they form a free abelian group under multiplication. Factorization of *elements* of $\mathcal{O}_K$ (and much more!) is governed by the class group of $\mathcal{O}_K$, which is the quotient of the group of fractional ideals by the principal fractional ideals (see Chapter 7).

3.1 Dedekind Domains

Recall (Corollary 2.4.6) that we proved that the ring of integers $\mathcal{O}_K$ of a number field is noetherian as follows. As we saw before using norms, the ring $\mathcal{O}_K$ is finitely generated as a module over $\mathbb{Z}$, so it is certainly finitely generated as a ring over $\mathbb{Z}$. By the Hilbert Basis Theorem (Theorem 2.2.10), $\mathcal{O}_K$ is noetherian.

If R is an integral domain, the *field of fractions* $\mathrm{Frac}(R)$ of R is the field of all equivalence classes of formal quotients a/b, where $a, b \in R$ with $b \neq 0$, and $a/b \sim c/d$ if $ad = bc$. For example, the field of fractions of $\mathbb{Z}$ is (canonically isomorphic to) $\mathbb{Q}$ and the field of fractions of $\mathbb{Z}[(1+\sqrt{5})/2]$ is $\mathbb{Q}(\sqrt{5})$. The field of fractions of the ring $\mathcal{O}_K$ of integers of a number field K is just the number field K (see Lemma 2.3.26).

Example 3.1.1. We compute the fraction fields mentioned above.

```
Frac(ZZ)
```

```
Rational Field
```

In Sage the `Frac` command usually returns a field canonically isomorphic to the fraction field (not a formal construction).

```
K.<a> = QuadraticField(5)
OK = K.ring_of_integers(); OK
```

```
Maximal Order in Number Field in a with defining polynomial x^2 - 5
```

```
OK.basis()
```

```
[1/2*a + 1/2, a]
```

```
Frac(OK)
```

```
Number Field in a with defining polynomial x^2 - 5
```

Remark 3.1.2. Note that in computers `1/2 * x` means the same as `(1/2)*x`. For more information about the order of operations in programming see `http://en.wikipedia.org/wiki/Order_of_operations`. In Sage the ^ symbol is replaced with python's exponentiation `**` at execution.[1]

The fraction field of an *order* – i.e., a subring of $\mathcal{O}_K$ of finite index – is also the number field again.

```
O2 = K.order(2*a); O2
```

```
Order in Number Field in a with defining polynomial x^2 - 5
```

```
Frac(O2)
```

```
Number Field in a with defining polynomial x^2 - 5
```

Definition 3.1.3 (Integrally Closed)**.** An integral domain R is *integrally closed in its field of fractions* if whenever α is in the field of fractions of R and α satisfies a monic polynomial $f \in R[x]$, then $\alpha \in R$.

For example, every field is integrally closed in its field of fractions, as is the ring $\mathbb{Z}$ of integers. However, $\mathbb{Z}[\sqrt{5}]$ is not integrally closed in its field of fractions, since $(1+\sqrt{5})/2$ is integrally over $\mathbb{Z}$ and lies in $\mathbb{Q}(\sqrt{5})$, but not in $\mathbb{Z}[\sqrt{5}]$

Proposition 3.1.4. *If K is any number field, then $\mathcal{O}_K$ is integrally closed. Also, the ring $\overline{\mathbb{Z}}$ of all algebraic integers (in a fixed choice of $\overline{\mathbb{Q}}$) is integrally closed.*

Proof. We first prove that $\overline{\mathbb{Z}}$ is integrally closed. Suppose $\alpha \in \overline{\mathbb{Q}}$ is integral over $\overline{\mathbb{Z}}$, so there is a monic polynomial $f(x) = x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$ with $a_i \in \overline{\mathbb{Z}}$ and $f(\alpha) = 0$. The a_i all lie in the ring of integers $\mathcal{O}_K$ of the number field $K = \mathbb{Q}(a_0, a_1, \ldots a_{n-1})$, and $\mathcal{O}_K$ is finitely generated as a $\mathbb{Z}$-module, so $\mathbb{Z}[a_0, \ldots, a_{n-1}]$ is finitely generated as a $\mathbb{Z}$-module. Since $f(\alpha) = 0$, we can write α^n

[1] Another source for order of operations specific to python is `https://docs.python.org/2/reference/expressions.html#operator-precedence`.

as a $\mathbb{Z}[a_0, \ldots, a_{n-1}]$-linear combination of α^i for $i < n$, so the ring $\mathbb{Z}[a_0, \ldots, a_{n-1}, \alpha]$ is also finitely generated as a $\mathbb{Z}$-module. Thus $\mathbb{Z}[\alpha]$ is finitely generated as a $\mathbb{Z}$-module because it is a submodule of a finitely generated $\mathbb{Z}$-module, which implies that α is integral over $\mathbb{Z}$.

Without loss we may assume that $K \subset \overline{\mathbb{Q}}$, so that $\mathcal{O}_K = \overline{\mathbb{Z}} \cap K$. Suppose $\alpha \in K$ is integral over $\mathcal{O}_K$. Then since $\overline{\mathbb{Z}}$ is integrally closed, α is an element of $\overline{\mathbb{Z}}$, so $\alpha \in K \cap \overline{\mathbb{Z}} = \mathcal{O}_K$, as required. □

Exercise 3.1.5. Prove that $\overline{\mathbb{Z}}$ is not noetherian.

[*Hint*: Consider an ideal generated by fractional powers of a prime.]

Definition 3.1.6 (Dedekind Domain)**.** An integral domain R is a *Dedekind domain* if it is noetherian, integrally closed in its field of fractions, and every nonzero prime ideal of R is maximal.

Exercise 3.1.7. Let K be a field.

(a) Prove that the polynomial ring $K[x]$ is a Dedekind domain.

(b) Is $\mathbb{Z}[x]$ a Dedekind domain?

The ring $\mathbb{Z} \oplus \mathbb{Z}$ is not a Dedekind domain because it is not an integral domain. The ring $\mathbb{Z}[\sqrt{5}]$ is not a Dedekind domain because it is not integrally closed in its field of fractions. The ring $\mathbb{Z}$ is a Dedekind domain, as is any ring of integers $\mathcal{O}_K$ of a number field, as we will see below. Also, any field K is a Dedekind domain, since it is an integral domain, it is trivially integrally closed in itself, and there are no nonzero prime ideals so the condition that they be maximal is empty.

Exercise 3.1.8. In Proposition 3.1.4 we showed that $\overline{\mathbb{Z}}$ is integrally closed in its field of fractions. Prove that and every nonzero prime ideal of $\overline{\mathbb{Z}}$ is maximal. Together with Exercise 3.1.5, this shows $\overline{\mathbb{Z}}$ is not a Dedekind domain only because it is not noetherian.

Exercise* 3.1.9. Show that Dedekind domains are closed under localization. This means the following: given any non-zero prime $\mathfrak{p}$ in R, the *localization* $R_\mathfrak{p}$ of R at $\mathfrak{p}$ is the ring formed by inverting all elements of R not contained in $\mathfrak{p}$. Thus $R_\mathfrak{p}$ is a subring of the field of fractions K of R which contains R. For example, $\mathbb{Z}_{(2)}$ is the localization of $\mathbb{Z}$ at the prime ideal (2). Note $\mathbb{Z}_{(2)}$ contains $\frac{1}{3}$ but not $\frac{1}{2}$. This exercise will show $R_\mathfrak{p}$ is again a Dedekind domain. In general, any element of $R_\mathfrak{p}$ can be written as a quotient $\frac{a}{b}$ for some $a \in R$ and $b \in R \setminus \mathfrak{p}$.

[*Hint*: It is a standard fact of localizations that set of prime ideals in $R_\mathfrak{p}$ is in bijection with the set of prime ideals of R contained in P. Use this to show $R_\mathfrak{p}$ is noetherian and all prime ideals of $R_\mathfrak{p}$ are maximal. It remains to show $R_\mathfrak{p}$ is integrally closed. Let $\alpha \in K$ satisfy a monic polynomial with coefficients in $R_\mathfrak{p}$. By clearing denominators show that $s\alpha \in R$ for some $s \in R \setminus \mathfrak{p}$.]

Proposition 3.1.10. *The ring of integers $\mathcal{O}_K$ of a number field is a Dedekind domain.*

Proof. By Proposition 3.1.4, the ring $\mathcal{O}_K$ is integrally closed, and by Proposition 2.4.6 it is noetherian. Suppose that $\mathfrak{p}$ is a nonzero prime ideal of $\mathcal{O}_K$. Let $\alpha \in \mathfrak{p}$ be a nonzero element, and let $f(x) \in \mathbb{Z}[x]$ be the minimal polynomial of α. Then

$$f(\alpha) = \alpha^n + a_{n-1}\alpha^{n-1} + \cdots + a_1\alpha + a_0 = 0,$$

so $a_0 = -(\alpha^n + a_{n-1}\alpha^{n-1} + \cdots + a_1\alpha) \in \mathfrak{p}$. Since f is irreducible, a_0 is a nonzero element of $\mathbb{Z}$ that lies in $\mathfrak{p}$. Every element of the finitely generated abelian group $\mathcal{O}_K/\mathfrak{p}$ is killed by a_0, so $\mathcal{O}_K/\mathfrak{p}$ is a finite set. Since $\mathfrak{p}$ is prime, $\mathcal{O}_K/\mathfrak{p}$ is an integral domain. Every finite integral domain is a field (see Exercise 2), so $\mathfrak{p}$ is maximal, which completes the proof. □

3.2 Factorization of Ideals

If I and J are ideals in a ring R, the product IJ is the ideal *generated by* all products of elements in I with elements in J:

$$IJ = (ab : a \in I, b \in J) \subset R.$$

Note that the set of all products ab, with $a \in I$ and $b \in J$, need not be an ideal, so it is important to take the ideal generated by that set (see Exercise 3).

Definition 3.2.1 (Fractional Ideal)**.** A *fractional ideal* is a nonzero $\mathcal{O}_K$-submodule I of K that is finitely generated as an $\mathcal{O}_K$-module.

We will sometimes call a genuine ideal $I \subset \mathcal{O}_K$ an *integral ideal*. The notion of fractional ideal makes sense for an arbitrary Dedekind domain R – it is an R-module $I \subset K = \mathrm{Frac}(R)$ that is finitely generated as an R-module.

Example 3.2.2. We multiply two fractional ideals in Sage:

```
K.<a> = NumberField(x^2 + 23)
I = K.fractional_ideal(2, 1/2*a - 1/2)
J = I^2
I
```

```
Fractional ideal (2, 1/2*a - 1/2)
```

```
J
```

```
Fractional ideal (4, 1/2*a + 3/2)
```

```
I*J
```

```
Fractional ideal (1/2*a + 3/2)
```

Since fractional ideals I are finitely generated, we can clear denominators of a generating set to see that there exists some nonzero $\alpha \in K$ such that

$$\alpha I = J \subset \mathcal{O}_K,$$

with J an integral ideal. Thus dividing by α, we see that every fractional ideal is of the form

$$aJ = \{ab : b \in J\}$$

for some $a \in K$ and integral ideal $J \subset \mathcal{O}_K$.

For example, the set $\frac{1}{2}\mathbb{Z}$ of rational numbers with denominator 1 or 2 is a fractional ideal of $\mathbb{Z}$.

Theorem 3.2.3. *The set of fractional ideals of a Dedekind domain R is an abelian group under ideal multiplication with identity element R.*

Note that fractional ideals are nonzero by definition, so it is not necessary to write "nonzero fractional ideals" in the statement of the theorem. We will *only* prove Theorem 3.2.3 in the case when $R = \mathcal{O}_K$ is the ring of integers of a number field K. The general case can be found in many algebraic number theory books such as [Mar77, Ch. 3]. Before proving Theorem 3.2.3 we prove a lemma. For the rest of this section $\mathcal{O}_K$ is the ring of integers of a number field K.

Definition 3.2.4 (Divides for Ideals)**.** Suppose that I, J are ideals of $\mathcal{O}_K$. Then we say that I *divides* J if $I \supset J$.

To see that this notion of divides is sensible, suppose $K = \mathbb{Q}$, so $\mathcal{O}_K = \mathbb{Z}$. Then $I = (n)$ and $J = (m)$ for some integer n and m, and I divides J means that $(n) \supset (m)$, i.e., that there exists an integer c such that $m = cn$, which exactly means that n divides m, as expected.

Lemma 3.2.5. *Suppose I is a nonzero ideal of $\mathcal{O}_K$. Then there exist prime ideals $\mathfrak{p}_1, \ldots, \mathfrak{p}_n$ such that $\mathfrak{p}_1 \cdot \mathfrak{p}_2 \cdots \mathfrak{p}_n \subset I$, i.e., I divides a product of prime ideals.*

Proof. Let S be the set of nonzero ideals of $\mathcal{O}_K$ that do not satisfy the conclusion of the lemma. The key idea is to use that $\mathcal{O}_K$ is noetherian to show that S is the empty set. If S is nonempty, then since $\mathcal{O}_K$ is noetherian, there is an ideal $I \in S$ that is maximal as an element of S. If I were prime, then I would trivially contain a product of primes, so we may assume that I is not prime. Thus there exists $a, b \in \mathcal{O}_K$ such that $ab \in I$ but $a \notin I$ and $b \notin I$. Let $J_1 = I + (a)$ and $J_2 = I + (b)$. Then neither J_1 nor J_2 is in S, since I is maximal, so both J_1 and J_2 contain a product of prime ideals, say $\mathfrak{p}_1 \cdots \mathfrak{p}_r \subset J_1$ and $\mathfrak{q}_1 \cdots \mathfrak{q}_s \subset J_2$. Then

$$\mathfrak{p}_1 \cdots \mathfrak{p}_r \cdot \mathfrak{q}_1 \cdots \mathfrak{q}_s \subset J_1 J_2 = I^2 + I(b) + (a)I + (ab) \subset I,$$

so I contains a product of primes. This is a contradiction, since we assumed $I \in S$. Thus S is empty, which completes the proof. $\square$

We are now ready to prove the theorem.

Proof of Theorem 3.2.3. Note that we will *only* prove Theorem 3.2.3 in the case when $R = \mathcal{O}_K$ is the ring of integers of a number field K.

The product of two fractional ideals is again finitely generated, so it is a fractional ideal, and $I\mathcal{O}_K = I$ for any ideal I, so to prove that the set of fractional ideals under multiplication is a group it suffices to show the existence of inverses. We will first prove that if $\mathfrak{p}$ is a prime ideal, then $\mathfrak{p}$ has an inverse, then we will prove that all nonzero integral ideals have inverses, and finally observe that every fractional ideal has an inverse. (Note: Once we know that the set of fractional ideals is a group, it will follows that inverses are unique; until then we will be careful to write "an" instead of "the".)

Suppose $\mathfrak{p}$ is a nonzero prime ideal of $\mathcal{O}_K$. We will show that the $\mathcal{O}_K$-module

$$I = \{a \in K : a\mathfrak{p} \subset \mathcal{O}_K\}$$

is a fractional ideal of $\mathcal{O}_K$ such that $I\mathfrak{p} = \mathcal{O}_K$, so that I is an inverse of $\mathfrak{p}$.

For the rest of the proof, fix a nonzero element $b \in \mathfrak{p}$. Since I is an $\mathcal{O}_K$-module, $bI \subset \mathcal{O}_K$ is an $\mathcal{O}_K$ ideal, hence I is a fractional ideal. Since $\mathcal{O}_K \subset I$ we have $\mathfrak{p} \subset I\mathfrak{p} \subset \mathcal{O}_K$, hence since $\mathfrak{p}$ is maximal, either $\mathfrak{p} = I\mathfrak{p}$ or $I\mathfrak{p} = \mathcal{O}_K$. If $I\mathfrak{p} = \mathcal{O}_K$, we are done since then I is an inverse of $\mathfrak{p}$. Thus suppose that $I\mathfrak{p} = \mathfrak{p}$. Our strategy is to show that there is some $d \in I$, with $d \notin \mathcal{O}_K$. Since $I\mathfrak{p} = \mathfrak{p}$, such a d would leave $\mathfrak{p}$ invariant, i.e., $d\mathfrak{p} \subset \mathfrak{p}$. Since $\mathfrak{p}$ is a finitely generated $\mathcal{O}_K$-module we will see that it will follow that $d \in \mathcal{O}_K$, a contradiction.

By Lemma 3.2.5, we can choose a product $\mathfrak{p}_1, \ldots, \mathfrak{p}_m$, with m minimal, with

$$\mathfrak{p}_1\mathfrak{p}_2\cdots\mathfrak{p}_m \subset (b) \subset \mathfrak{p}.$$

If no $\mathfrak{p}_i$ is contained in $\mathfrak{p}$, then we can choose for each i an $a_i \in \mathfrak{p}_i$ with $a_i \notin \mathfrak{p}$; but then $\prod a_i \in \mathfrak{p}$, which contradicts that $\mathfrak{p}$ is a prime ideal. Thus some $\mathfrak{p}_i$, say $\mathfrak{p}_1$, is contained in $\mathfrak{p}$, which implies that $\mathfrak{p}_1 = \mathfrak{p}$ since every nonzero prime ideal is maximal. Because m is minimal, $\mathfrak{p}_2\cdots\mathfrak{p}_m$ is not a subset of (b), so there exists $c \in \mathfrak{p}_2\cdots\mathfrak{p}_m$ that does not lie in (b). Then $\mathfrak{p}(c) \subset (b)$, so by definition of I we have $d = c/b \in I$. However, $d \notin \mathcal{O}_K$, since if it were then c would be in (b). We have thus found our element $d \in I$ that does not lie in $\mathcal{O}_K$.

To finish the proof that $\mathfrak{p}$ has an inverse, we observe that d preserves the finitely generated $\mathcal{O}_K$-module $\mathfrak{p}$, and is hence in $\mathcal{O}_K$, a contradiction. More precisely, if $b_1, \ldots, b_n$ is a basis for $\mathfrak{p}$ as a $\mathbb{Z}$-module, then the action of d on $\mathfrak{p}$ is given by a matrix with entries in $\mathbb{Z}$, so the minimal polynomial of d has coefficients in $\mathbb{Z}$ (because d satisfies the minimal polynomial of ℓ_d, by the Cayley-Hamilton theorem – here we also use that $\mathbb{Q} \otimes \mathfrak{p} = K$, since $\mathcal{O}_K/\mathfrak{p}$ is a finite set). This implies that d is integral over $\mathbb{Z}$, so $d \in \mathcal{O}_K$, since $\mathcal{O}_K$ is integrally closed by Proposition 3.1.4. (Note how this argument depends strongly on the fact that $\mathcal{O}_K$ is integrally closed!)

So far we have proved that if $\mathfrak{p}$ is a prime ideal of $\mathcal{O}_K$, then

$$\mathfrak{p}^{-1} = \{a \in K : a\mathfrak{p} \subset \mathcal{O}_K\}$$

is the inverse of $\mathfrak{p}$ in the monoid of nonzero fractional ideals of $\mathcal{O}_K$. As mentioned after Definition 3.2.1, every nonzero fractional ideal is of the form aI for $a \in K$

and I an integral ideal, so since (a) has inverse $(1/a)$, it suffices to show that every integral ideal I has an inverse. If not, then there is a nonzero integral ideal I that is maximal among all nonzero integral ideals that do not have an inverse. Every ideal is contained in a maximal ideal, so there is a nonzero prime ideal $\mathfrak{p}$ such that $I \subset \mathfrak{p}$. Multiplying both sides of this inclusion by $\mathfrak{p}^{-1}$ and using that $\mathcal{O}_K \subset \mathfrak{p}^{-1}$, we see that

$$I \subset \mathfrak{p}^{-1}I \subset \mathfrak{p}^{-1}\mathfrak{p} = \mathcal{O}_K.$$

If $I = \mathfrak{p}^{-1}I$, then arguing as in the proof that $\mathfrak{p}^{-1}$ is an inverse of $\mathfrak{p}$, we see that each element of $\mathfrak{p}^{-1}$ preserves the finitely generated $\mathbb{Z}$-module I and is hence integral. But then $\mathfrak{p}^{-1} \subset \mathcal{O}_K$, which, upon multiplying both sides by $\mathfrak{p}$, implies that $\mathcal{O}_K = \mathfrak{p}\mathfrak{p}^{-1} \subset \mathfrak{p}$, a contradiction. Thus $I \neq \mathfrak{p}^{-1}I$. Because I is maximal among ideals that do not have an inverse, the ideal $\mathfrak{p}^{-1}I$ does have an inverse J. Then $\mathfrak{p}^{-1}J$ is an inverse of I, since $(J\mathfrak{p}^{-1})I = J(\mathfrak{p}^{-1}I) = \mathcal{O}_K$. □

We can finally deduce the crucial Theorem 3.2.6, which will allow us to show that any nonzero ideal of a Dedekind domain can be expressed uniquely as a product of primes (up to order). Thus unique factorization holds for ideals in a Dedekind domain, and it is this unique factorization that initially motivated the introduction of ideals to mathematics over a century ago.

Theorem 3.2.6. *Suppose I is a nonzero integral ideal of $\mathcal{O}_K$. Then I can be written as a product*

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

of prime ideals of $\mathcal{O}_K$, and this representation is unique up to order.

Proof. Suppose I is an ideal that is maximal among the set of all ideals in $\mathcal{O}_K$ that cannot be written as a product of primes. Every ideal is contained in a maximal ideal, so I is contained in a nonzero prime ideal $\mathfrak{p}$. If $I\mathfrak{p}^{-1} = I$, then by Theorem 3.2.3 we can cancel I from both sides of this equation to see that $\mathfrak{p}^{-1} = \mathcal{O}_K$, a contradiction. Since $\mathcal{O}_K \subset \mathfrak{p}^{-1}$, we have $I \subset I\mathfrak{p}^{-1}$, and by the above observation I is strictly contained in $I\mathfrak{p}^{-1}$. By our maximality assumption on I, there are maximal ideals $\mathfrak{p}_1, \ldots, \mathfrak{p}_n$ such that $I\mathfrak{p}^{-1} = \mathfrak{p}_1 \cdots \mathfrak{p}_n$. Then $I = \mathfrak{p} \cdot \mathfrak{p}_1 \cdots \mathfrak{p}_n$, a contradiction. Thus every ideal can be written as a product of primes.

Suppose $\mathfrak{p}_1 \cdots \mathfrak{p}_n = \mathfrak{q}_1 \cdots \mathfrak{q}_m$. If no $\mathfrak{q}_i$ is contained in $\mathfrak{p}_1$, then for each i there is an $a_i \in \mathfrak{q}_i$ such that $a_i \notin \mathfrak{p}_1$. But the product of the a_i is in $\mathfrak{p}_1 \cdots \mathfrak{p}_n$, which is a subset of $\mathfrak{p}_1$, which contradicts that $\mathfrak{p}_1$ is a prime ideal. Thus $\mathfrak{q}_i = \mathfrak{p}_1$ for some i. We can thus cancel $\mathfrak{q}_i$ and $\mathfrak{p}_1$ from both sides of the equation by multiplying both sides by the inverse. Repeating this argument finishes the proof of uniqueness. □

Theorem 3.2.7. *If I is a fractional ideal of $\mathcal{O}_K$ then there exists prime ideals $\mathfrak{p}_1, \ldots, \mathfrak{p}_n$ and $\mathfrak{q}_1, \ldots, \mathfrak{q}_m$, unique up to order, such that*

$$I = (\mathfrak{p}_1 \cdots \mathfrak{p}_n)(\mathfrak{q}_1 \cdots \mathfrak{q}_m)^{-1}.$$

Proof. We have $I = (a/b)J$ for some $a, b \in \mathcal{O}_K$ and integral ideal J. Applying Theorem 3.2.6 to (a), (b), and J gives an expression as claimed. For uniqueness, if one has two such product expressions, multiply through by the denominators and use the uniqueness part of Theorem 3.2.6 □

Example 3.2.8. The ring of integers of $K = \mathbb{Q}(\sqrt{-6})$ is $\mathcal{O}_K = \mathbb{Z}[\sqrt{-6}]$. We have

$$6 = -\sqrt{-6}\sqrt{-6} = 2 \cdot 3.$$

If $ab = \sqrt{-6}$, with $a, b \in \mathcal{O}_K$ and neither a unit, then $\mathrm{Norm}(a)\,\mathrm{Norm}(b) = 6$, so without loss $\mathrm{Norm}(a) = 2$ and $\mathrm{Norm}(b) = 3$. If $a = c + d\sqrt{-6}$, then $\mathrm{Norm}(a) = c^2 + 6d^2$; since the equation $c^2 + 6d^2 = 2$ has no solution with $c, d \in \mathbb{Z}$, there is no element in $\mathcal{O}_K$ with norm 2, so $\sqrt{-6}$ is irreducible. Also, $\sqrt{-6}$ is not a unit times 2 or times 3, since again the norms would not match up. Thus 6 cannot be written uniquely as a product of irreducibles in $\mathcal{O}_K$. Theorem 3.2.7, however, implies that the principal ideal (6) can, however, be written uniquely as a product of prime ideals. An explicit decomposition is

$$(6) = (2, 2 + \sqrt{-6})^2 \cdot (3, 3 + \sqrt{-6})^2, \tag{3.2.1}$$

where each of the ideals $(2, 2 + \sqrt{-6})$ and $(3, 3 + \sqrt{-6})$ is prime. We will discuss algorithms for computing such a decomposition in detail in Chapter 4. The first idea is to write $(6) = (2)(3)$, and hence reduce to the case of writing the (p), for $p \in \mathbb{Z}$ prime, as a product of primes. Next one decomposes the finite (as a set) ring $\mathcal{O}_K/p\mathcal{O}_K$.

The factorization (3.2.1) can be compute using Sage as follows:

```
K.<a> = NumberField(x^2 + 6); K
```

```
Number Field in a with defining polynomial x^2 + 6
```

```
K.factor(6)
```

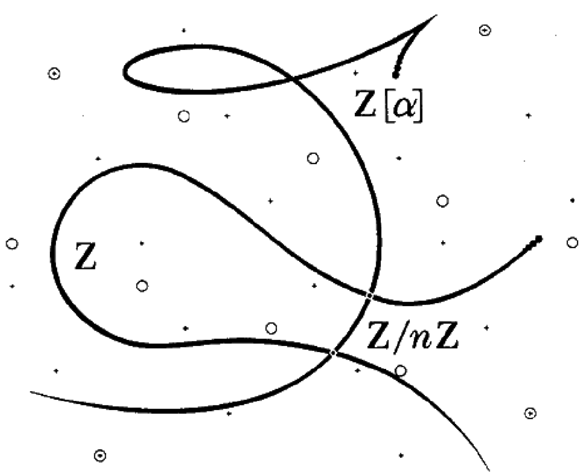
```
(Fractional ideal (2, a))^2 * \
(Fractional ideal (3, a))^2
```

Chapter 4

Factoring Primes

Let p be a prime and $\mathcal{O}_K$ the ring of integers of a number field. This chapter is about how to write $p\mathcal{O}_K$ as a product of prime ideals of $\mathcal{O}_K$. Paradoxically, computing the explicit prime ideal factorization of $p\mathcal{O}_K$ is easier than computing $\mathcal{O}_K$.

4.1 The Problem



A diagram from [LL93].

"The obvious mathematical breakthrough would be development of an easy way to factor large prime numbers."
– Bill Gates, *The Road Ahead*, 1st ed., pg 265

Bill Gates meant[1] factoring products of two primes, which would break the RSA cryptosystem (see e.g. [Ste09, §3.2]). However, perhaps Gates is an algebraic number theorist, and he really meant what he said: then we might imagine that he meant factorization of primes of $\mathbb{Z}$ in rings of integers of number fields. For example, $2^{16}+1 = 65537$ is a "large" prime, and in $\mathbb{Z}[i]$ we have

$$(65537) = (65537, 2^8 + i) \cdot (65537, 2^8 - i).$$

4.1.1 Geometric Intuition

Let $K = \mathbb{Q}(\alpha)$ be a number field, and let $\mathcal{O}_K$ be the ring of integers of K. To employ our geometric intuition, as the Lenstras did on the cover of [LL93], it is helpful to view $\mathcal{O}_K$ as a 1-dimensional scheme

$$X = \mathrm{Spec}(\mathcal{O}_K) = \{\text{all prime ideals of } \mathcal{O}_K\}$$

over

$$Y = \mathrm{Spec}(\mathbb{Z}) = \{(0)\} \cup \{p\mathbb{Z} : p \in \mathbb{Z}_{>0} \text{ is prime}\}.$$

There is a natural map $\pi : X \to Y$ that sends a prime ideal $\mathfrak{p} \in X$ to $\mathfrak{p} \cap \mathbb{Z} \in Y$. For example, if

$$\mathfrak{p} = (65537, 2^8 + i) \subset \mathbb{Z}[i],$$

then $\mathfrak{p} \cap \mathbb{Z} = (65537)$. For more on this viewpoint, see [Har77] and [EH00, Ch. 2].

If $p \in \mathbb{Z}$ is a prime number, then the ideal $p\mathcal{O}_K$ of $\mathcal{O}_K$ factors uniquely as a product $\prod \mathfrak{p}_i^{e_i}$, where the $\mathfrak{p}_i$ are maximal ideals of $\mathcal{O}_K$. We may imagine the decomposition of $p\mathcal{O}_K$ into prime ideals geometrically as the fiber $\pi^{-1}(p\mathbb{Z})$, where the exponents e_i are the multiplicities of the fibers. Notice that the elements of $\pi^{-1}(p\mathbb{Z})$ are the prime ideals of $\mathcal{O}_K$ that contain p, i.e., the primes that divide $p\mathcal{O}_K$. This chapter is about how to compute the $\mathfrak{p}_i$ and e_i.

Remark 4.1.1. More technically, in algebraic geometry one defines the inverse image of the point $p\mathbb{Z}$ to be the spectrum of the tensor product $\mathcal{O}_K \otimes_{\mathbb{Z}} \mathbb{Z}/\mathfrak{p}\mathbb{Z}$; by a generalization of the Chinese Remainder Theorem, we have

$$\mathcal{O}_K \otimes_{\mathbb{Z}} (\mathbb{Z}/\mathfrak{p}\mathbb{Z}) \cong \oplus \mathcal{O}_K/\mathfrak{p}_i^{e_i}.$$

4.1.2 Examples

The following Sage session shows the commands needed to compute the factorization of $p\mathcal{O}_K$ for K the number field defined by a root of $x^5 + 7x^4 + 3x^2 - x + 1$ and $p = 2$ and 5. We first create an element $f \in \mathbb{Q}[x]$ in Sage:

[1]This quote is on page 265 of the first edition. In the second edition, on page 303, this sentence is changed to "The obvious mathematical breakthrough that would defeat our public key encryption would be the development of an easy way to factor large numbers." This is less nonsensical; however, fast factoring is *not* known to break all commonly used public-key cryptosystem. For example, there are cryptosystems based on the difficulty of computing discrete logarithms in $\mathbb{F}_p^*$ and on elliptic curves over $\mathbb{F}_p$, which (presumably) would not be broken even if one could factor large numbers quickly.

```
R.<x> = QQ[]
f = x^5 + 7*x^4 + 3*x^2 - x + 1
```

Then we create the corresponding number field obtained by adjoining a root of f, and find its ring of integers.

```
K.<a> = NumberField(f)
OK = K.ring_of_integers()
OK.basis()
```

```
[1, a, a^2, a^3, a^4]
```

We define the ideal $2\mathcal{O}_K$ and factor – it turns out to be prime.

```
I = K.fractional_ideal(2); I
```

```
Fractional ideal (2)
```

```
I.factor()
```

```
Fractional ideal (2)
```

```
I.is_prime()
```

```
True
```

Finally we factor $5\mathcal{O}_K$, which factors as a product of three primes.

```
I = K.fractional_ideal(5); I
```

```
Fractional ideal (5)
```

```
I.factor()
```

```
(Fractional ideal (5, -2*a^4 - 13*a^3 + 7*a^2 - 6*a + 2)) * \
(Fractional ideal (5, a^4 + 7*a^3 + 3*a + 1)) * \
(Fractional ideal (5, a^4 + 7*a^3 + 3*a - 3))^2
```

Notice that the polynomial f factors in a similar way:

```
f.factor_mod(5)
```

```
(x + 2) * (x + 3)^2 * (x^2 + 4*x + 2)
```

Thus $2\mathcal{O}_K$ is already a prime ideal, and

$$5\mathcal{O}_K = (5, 2+a) \cdot (5, 3+a)^2 \cdot (5, 2+4a+a^2).$$

Notice that in this example $\mathcal{O}_K = \mathbb{Z}[a]$. (Warning: There are examples of $\mathcal{O}_K$ such that $\mathcal{O}_K \neq \mathbb{Z}[a]$ for any $a \in \mathcal{O}_K$, as Example 4.3.2 below illustrates.) When

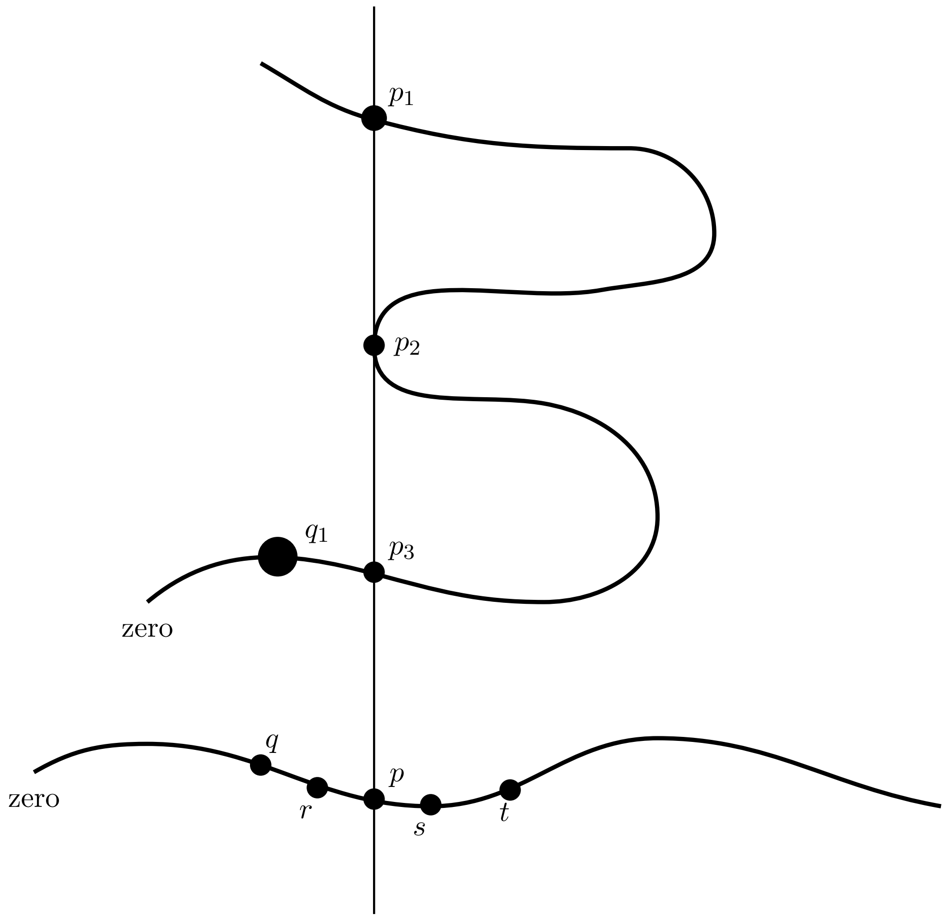


Figure 4.1.1: Diagram of $\mathrm{Spec}(\mathcal{O}_K) \to \mathrm{Spec}(\mathbb{Z})$

$\mathcal{O}_K = \mathbb{Z}[a]$ it is relatively easy to factor $p\mathcal{O}_K$, at least assuming one can factor polynomials in $\mathbb{F}_p[x]$. The following factorization gives a hint as to why:

$$x^5 + 7x^4 + 3x^2 - x + 1 \equiv (x+2) \cdot (x+3)^2 \cdot (x^2 + 4x + 2) \pmod 5.$$

The exponent 2 of $(5, 3+a)^2$ in the factorization of $5\mathcal{O}_K$ above suggests "ramification", in the sense that the cover $X \to Y$ has less points (counting their "size", i.e., their residue class degree) in its fiber over 5 than it has generically. See Figure 4.1.1.

4.2 A Method for Factoring Primes that Often Works

Suppose $a \in \mathcal{O}_K$ is such that $K = \mathbb{Q}(a)$, and let $f(x) \in \mathbb{Z}[x]$ be the minimal polynomial of a. Then $\mathbb{Z}[a] \subset \mathcal{O}_K$, and we have a diagram of schemes

$$\begin{tikzcd}
\bigcup \mathrm{Spec}(\mathcal{O}_K/\mathfrak{p}_i^{e_i}) \arrow[r, hook] \arrow[d] & \mathrm{Spec}(\mathcal{O}_K) \arrow[d] \\
\bigcup \mathrm{Spec}(\mathbb{F}_p[x]/(\overline{f}_i^{e_i})) \arrow[r, hook] \arrow[d] & \mathrm{Spec}(\mathbb{Z}[a]) \arrow[d] \\
\mathrm{Spec}(\mathbb{F}_p) \arrow[r, hook] & \mathrm{Spec}(\mathbb{Z})
\end{tikzcd}$$

where $\overline{f} = \prod_i \overline{f}_i^{e_i}$ is the factorization of the image of f in $\mathbb{F}_p[x]$, and $p\mathcal{O}_K = \prod \mathfrak{p}_i^{e_i}$ is the factorization of $p\mathcal{O}_K$ in terms of prime ideals of $\mathcal{O}_K$. On the level of rings, the bottom horizontal map is the quotient map $\mathbb{Z} \to \mathbb{Z}/p\mathbb{Z} \cong \mathbb{F}_p$. The middle horizontal map is induced by

$$\mathbb{Z}[x] \to \bigoplus_i \mathbb{F}_p[x]/(\overline{f}_i^{e_i}),$$

and the top horizontal map is induced by

$$\mathcal{O}_K \to \mathcal{O}_K/p\mathcal{O}_K \cong \bigoplus \mathcal{O}_K/\mathfrak{p}_i^{e_i},$$

where the isomorphism is by the Chinese Remainder Theorem, which is Theorem 5.1.4 below. The left vertical maps come from the inclusions

$$\mathbb{F}_p \hookrightarrow \mathbb{F}_p[x]/(\overline{f}_i^{e_i}) \hookrightarrow \mathcal{O}_K/\mathfrak{p}_i^{e_i},$$

and the right from the inclusions $\mathbb{Z} \hookrightarrow \mathbb{Z}[a] \hookrightarrow \mathcal{O}_K$.

The cover $\pi : \mathrm{Spec}(\mathbb{Z}[a]) \to \mathrm{Spec}(\mathbb{Z})$ is easy to understand because it is defined by the single equation $f(x)$, in the sense that $\mathbb{Z}[a] \cong \mathbb{Z}[x]/(f(x))$. To give a maximal ideal $\mathfrak{p}$ of $\mathbb{Z}[a]$ such that $\pi(\mathfrak{p}) = p\mathbb{Z}$ is the same as giving a homomorphism $\varphi : \mathbb{Z}[x]/(f) \to \overline{\mathbb{F}}_p$ up to automorphisms of the image, which is in turn the same as giving a root of f in $\overline{\mathbb{F}}_p$ up to automorphism, which is the same as giving an irreducible factor of the reduction of f modulo p.

Lemma 4.2.1. *Suppose the index of $\mathbb{Z}[a]$ in $\mathcal{O}_K$ is coprime to p. Then the primes $\mathfrak{p}_i$ in the factorization of $p\mathbb{Z}[a]$ do not decompose further going from $\mathbb{Z}[a]$ to $\mathcal{O}_K$, so finding the prime ideals of $\mathbb{Z}[a]$ that contain p yields the primes that appear in the factorization of $p\mathcal{O}_K$.*

Proof. Fix a basis for $\mathcal{O}_K$ and for $\mathbb{Z}[a]$ as $\mathbb{Z}$-modules. Form the matrix A whose columns express each basis element of $\mathbb{Z}[a]$ as a $\mathbb{Z}$-linear combination of the basis for $\mathcal{O}_K$. Then

$$\det(A) = \pm[\mathcal{O}_K : \mathbb{Z}[a]]$$

is coprime to p, by hypothesis. Thus the reduction of A modulo p is invertible, so it defines an isomorphism $\mathbb{Z}[a]/p\mathbb{Z}[a] \cong \mathcal{O}_K/p\mathcal{O}_K$.

Let $\overline{\mathbb{F}}_p$ denote a fixed algebraic closure of $\mathbb{F}_p$; thus $\overline{\mathbb{F}}_p$ is an algebraically closed field of characteristic p, over which all polynomials in $\mathbb{F}_p[x]$ factor into linear factors. Any homomorphism $\mathcal{O}_K \to \overline{\mathbb{F}}_p$ sends p to 0, so is the composition of a homomorphism $\mathcal{O}_K \to \mathcal{O}_K/p\mathcal{O}_K$ with a homomorphism $\mathcal{O}_K/p\mathcal{O}_K \to \overline{\mathbb{F}}_p$. Since $\mathcal{O}_K/p\mathcal{O}_K \cong \mathbb{Z}[a]/p\mathbb{Z}[a]$, the homomorphisms $\mathcal{O}_K \to \overline{\mathbb{F}}_p$ are in bijection with the homomorphisms $\mathbb{Z}[a] \to \overline{\mathbb{F}}_p$. The homomorphisms $\mathbb{Z}[a] \to \overline{\mathbb{F}}_p$ are in bijection with the roots of the reduction modulo p of the minimal polynomial of a in $\overline{\mathbb{F}}_p$. $\square$

Remark 4.2.2. Here is a "high-brow" proof of Lemma 4.2.1. By hypothesis we have an exact sequence of abelian groups

$$0 \to \mathbb{Z}[a] \to \mathcal{O}_K \to H \to 0,$$

where H is a finite abelian group of order coprime to p. Tensor product is right exact, and there is an exact sequence

$$\mathrm{Tor}_1(H, \mathbb{F}_p) \to \mathbb{Z}[a] \otimes \mathbb{F}_p \to \mathcal{O}_K \otimes \mathbb{F}_p \to H \otimes \mathbb{F}_p \to 0,$$

and $\mathrm{Tor}_1(H, \mathbb{F}_p) = 0$ (since H has no p-torsion), so $\mathbb{Z}[a] \otimes \mathbb{F}_p \cong \mathcal{O}_K \otimes \mathbb{F}_p$.

As suggested in the proof of the lemma, we find all homomorphisms $\mathcal{O}_K \to \overline{\mathbb{F}}_p$ by finding all homomorphism $\mathbb{Z}[a] \to \overline{\mathbb{F}}_p$. In terms of ideals, if $\mathfrak{p} = (f(a), p)\mathbb{Z}[a]$ is a maximal ideal of $\mathbb{Z}[a]$, then the ideal $\mathfrak{p}' = (f(a), p)\mathcal{O}_K$ of $\mathcal{O}_K$ is also maximal, since

$$\mathcal{O}_K/\mathfrak{p}' \cong (\mathcal{O}_K/p\mathcal{O}_K)/(f(\tilde{a})) \cong (\mathbb{Z}[a]/p\mathbb{Z}[a])/(f(\tilde{a})) \subset \overline{\mathbb{F}}_p,$$

where $\tilde{a}$ denotes the image of a in $\mathcal{O}_K/p\mathcal{O}_K$.

We formalize the above discussion in the following theorem (note: we will not prove that the powers are e_i here):

Theorem 4.2.3. *Let $f \in \mathbb{Z}[x]$ be the minimal polynomial of a over $\mathbb{Z}$. Suppose that $p \nmid [\mathcal{O}_K : \mathbb{Z}[a]]$ is a prime. Let*

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

where the $\overline{f}_i$ are distinct monic irreducible polynomials. Let $\mathfrak{p}_i = (p, f_i(a))$ where $f_i \in \mathbb{Z}[x]$ is a lift of $\overline{f}_i$ in $\mathbb{F}_p[x]$. Then

$$p\mathcal{O}_K = \prod_{i=1}^{t} \mathfrak{p}_i^{e_i}.$$

We return to the example from above, in which $K = \mathbb{Q}(a)$, where a is a root of $f = x^5 + 7x^4 + 3x^2 - x + 1$. The ring of integers $\mathcal{O}_K$ has discriminant $2945785 = 5 \cdot 353 \cdot 1669$, as the following Sage code shows.

```
K.<a> = NumberField(x^5 + 7*x^4 + 3*x^2 - x + 1)
D = K.discriminant(); D
```

```
2945785
```

```
factor(D)
```

```
5 * 353 * 1669
```

The order $\mathbb{Z}[a]$ has the same discriminant as $f(x)$, which is the same as the discriminant of $\mathcal{O}_K$, so $\mathbb{Z}[a] = \mathcal{O}_K$ and we can apply the above theorem. (Here we use that the index of $\mathbb{Z}[a]$ in $\mathcal{O}_K$ is the square of the quotient of their discriminants, a fact we will prove later in Section 6.2.)

```
R.<x> = QQ[]
discriminant(x^5 + 7*x^4 + 3*x^2 - x + 1)
```

```
2945785
```

We have

$$x^5 + 7x^4 + 3x^2 - x + 1 \equiv (x + 2) \cdot (x + 3)^2 \cdot (x^2 + 4x + 2) \pmod{5},$$

which yields the factorization of $5\mathcal{O}_K$ given before the theorem.

If we replace a by $b = 7a$, then the index of $\mathbb{Z}[b]$ in $\mathcal{O}_K$ will be a power of 7, which is coprime to 5, so the above method will still work.

```
K.<a> = NumberField(x^5 + 7*x^4 + 3*x^2 - x + 1)
f = (7*a).minpoly('x')
f
```

```
x^5 + 49*x^4 + 1029*x^2 - 2401*x + 16807
```

```
f.disc()
```

```
235050861175510968365785
```

```
factor(f.disc() / K.disc())
```

```
7^20
```

```
f.factor_mod(5)
```

```
(x + 4) * (x + 1)^2 * (x^2 + 3*x + 3)
```

Thus 5 factors in $\mathcal{O}_K$ as

$$5\mathcal{O}_K = (5, 7a + 1)^2 \cdot (5, 7a + 4) \cdot (5, (7a)^2 + 3(7a) + 3).$$

If we replace a by $b = 5a$ and try the above algorithm with $\mathbb{Z}[b]$, then the method fails because the index of $\mathbb{Z}[b]$ in $\mathcal{O}_K$ is divisible by 5.

```
K.<a> = NumberField(x^5 + 7*x^4 + 3*x^2 - x + 1)
f = (5*a).minpoly('x')
f
```

```
x^5 + 35*x^4 + 375*x^2 - 625*x + 3125
```

```
f.factor_mod(5)
```

```
x^5
```

4.3 A General Method

There are numbers fields K such that $\mathcal{O}_K$ is not of the form $\mathbb{Z}[a]$ for any $a \in K$. Even worse, Dedekind found a field K such that $2 \mid [\mathcal{O}_K : \mathbb{Z}[a]]$ for *all* $a \in \mathcal{O}_K$, so there is no choice of a such that Theorem 4.2.3 can be used to factor 2 for K (see Example 4.3.2 below).

4.3.1 Inessential Discriminant Divisors

Definition 4.3.1. A prime p is an *inessential discriminant divisor* if $p \mid [\mathcal{O}_K : \mathbb{Z}[a]]$ for *every* $a \in \mathcal{O}_K$.

See Example 6.2.7 below for why it is called an inessential "discriminant divisor" instead of an inessential "index divisor".

Since $[\mathcal{O}_K : \mathbb{Z}[a]]^2$ is the absolute value of $\mathrm{Disc}(f(x))/\mathrm{Disc}(\mathcal{O}_K)$, where $f(x)$ is the characteristic polynomial of $f(x)$, an inessential discriminant divisor divides the discriminant of the characteristic polynomial of any element of $\mathcal{O}_K$.

Example 4.3.2 (Dedekind). Let $K = \mathbb{Q}(a)$ be the cubic field defined by a root a of the polynomial $f = x^3 + x^2 - 2x + 8$. We will use Sage to show that 2 is an inessential discriminant divisor for K.

```
K.<a> = NumberField(x^3 + x^2 - 2*x + 8); K
```

```
Number Field in a with defining polynomial x^3 + x^2 - 2*x + 8
```

```
K.factor(2)
```

```
(Fractional ideal (1/2*a^2 - 1/2*a + 1)) * \
(Fractional ideal (-a^2 + 2*a - 3)) * \
(Fractional ideal (-3/2*a^2 + 5/2*a - 4))
```

Thus $2\mathcal{O}_K = \mathfrak{p}_1\mathfrak{p}_2\mathfrak{p}_3$, with the $\mathfrak{p}_i$ distinct, and one sees directly from the above expressions that $\mathcal{O}_K/\mathfrak{p}_i \cong \mathbb{F}_2$ for each i. If $\mathcal{O}_K = \mathbb{Z}[a]$ for some $a \in \mathcal{O}_K$ with minimal polynomial f, then $\overline{f}(x) \in \mathbb{F}_2[x]$ must be a product of three *distinct* linear factors, which is impossible, since the only linear polynomials in $\mathbb{F}_2[x]$ are x and $x + 1$.

4.3.2 Remarks on Ideal Factorization in General

Recall (from Definition 2.3.22) that an *order* in $\mathcal{O}_K$ is a subring $\mathcal{O}$ of $\mathcal{O}_K$ that has finite index in $\mathcal{O}_K$. For example, if $\mathcal{O}_K = \mathbb{Z}[i]$, then $\mathcal{O} = \mathbb{Z} + 5\mathbb{Z}[i]$ is an order in $\mathcal{O}_K$, and as an abelian group $\mathcal{O}_K/\mathcal{O}$ is cyclic of order 5.

Most algebraic number theory books do not describe an algorithm for decomposing primes in the general case. Fortunately, Cohen's book [Coh93, Ch. 6] does describe how to solve the general problem, in more than one way. The algorithms are nontrivial, and occupy a substantial part of Chapter 6 of Cohen's book. Our goal for the rest of this section is to give a hint as to what goes into them.

The general solutions to prime ideal factorization are somewhat surprising, since the algorithms are much more sophisticated than the one suggested by Theorem 4.2.3. However, these complicated algorithms all run very quickly in practice, even without assuming the maximal order is already known. In fact, they avoid computing $\mathcal{O}_K$ altogether, and instead compute only an order $\mathcal{O}$ that is *p-maximal*, i.e., is such that $p \nmid [\mathcal{O}_K : \mathcal{O}]$.

For simplicity we consider the following slightly easier problem whose solution illustrates the key ideas needed in the general case.

Problem 4.3.3. Let $\mathcal{O}$ be any order in $\mathcal{O}_K$ and let p be a prime of $\mathbb{Z}$. Find the prime ideals of $\mathcal{O}$ that contain p.

Given a prime p that we wish to factor in $\mathcal{O}_K$, we first find a p-maximal order $\mathcal{O}$. We then use a solution to Problem 4.3.3 to find the prime ideals $\mathfrak{p}$ of $\mathcal{O}$ that contain p. Second, we find the exponents e such that $\mathfrak{p}^e$ exactly divides $p\mathcal{O}$. The resulting factorization in $\mathcal{O}$ completely determines the factorization of $p\mathcal{O}_K$.

A p-maximal order can be found reasonably quickly in practice using algorithms called "round 2" and "round 4". To compute $\mathcal{O}_K$, given an order $\mathbb{Z}[\alpha] \subset \mathcal{O}_K$, one takes a sum of p-maximal orders, one for every p such that p^2 divides $\mathrm{Disc}(\mathbb{Z}[\alpha])$. The time-consuming part of this computation is finding the primes p such that $p^2 \mid \mathrm{Disc}(\mathbb{Z}[\alpha])$, not finding the p-maximal orders. This example illustrates that a fast algorithm for factoring integers would not only break the RSA cryptosystems, but would massively speed up computation of the ring of integers of a number field.

Remark 4.3.4. The MathSciNet review of [BL94] by J. Buhler contains the following:

> A result of Chistov says that finding the ring of integers $\mathcal{O}_K$ in an algebraic number field K is equivalent, under certain polynomial time reductions, to the problem of finding the largest squarefree divisor of a positive integer. No feasible (i.e., polynomial time) algorithm is known for the latter problem, and it is possible that it is no easier than the more general problem of factoring integers.

Thus it appears that computing the ring $\mathcal{O}_K$ is quite hard.

4.3.3 Finding a p-Maximal Order

Before describing the general factorization algorithm, we sketch some of the theory behind the general algorithms for computing a p-maximal order $\mathcal{O}$ in $\mathcal{O}_K$. The main input is the following theorem:

Theorem 4.3.5 (Pohst-Zassenhaus)**.** *Let $\mathcal{O}$ be an order in the the ring of integers $\mathcal{O}_K$ of a number field, let $p \in \mathbb{Z}$ be a prime, and let*

$$I_p = \{x \in \mathcal{O} : x^m \in p\mathcal{O} \text{ for some } m \geq 1\} \subset \mathcal{O}$$

be the radical of $p\mathcal{O}$, which is an ideal of $\mathcal{O}$. Let

$$\mathcal{O}' = \{x \in K : xI_p \subset I_p\}.$$

Then $\mathcal{O}'$ is an order and either $\mathcal{O}' = \mathcal{O}$, in which case $\mathcal{O}$ is p-maximal, or $\mathcal{O} \subset \mathcal{O}'$ and p divides $[\mathcal{O}' : \mathcal{O}]$.

Proof. We prove here only that $[\mathcal{O}' : \mathcal{O}] \mid p^n$, where n is the degree of K. We have $p \in I_p$, so if $x \in \mathcal{O}'$, then $xp \in I_p \subset \mathcal{O}$, which implies that $x \in \frac{1}{p}\mathcal{O}$. Since $(\frac{1}{p}\mathcal{O})/\mathcal{O}$ is of order p^n, the claim follows.

To complete the proof, we would show that if $\mathcal{O}' = \mathcal{O}$, then $\mathcal{O}$ is already p-maximal. See [Coh93, §6.1.1] for the rest if this proof. □

After deciding on how to represent elements of K and orders and ideals in K, one can give an efficient algorithm to compute the $\mathcal{O}'$ of the theorem. The algorithm mainly involves linear algebra over finite fields. It is complicated to describe, but efficient in practice, and is conceptually simple—just compute $\mathcal{O}'$. The trick for reducing the computation of $\mathcal{O}'$ to linear algebra is the following lemma:

Lemma 4.3.6. *Define a homomorphism $\psi : \mathcal{O} \hookrightarrow \mathrm{End}(I_p/pI_p)$ given by sending $\alpha \in \mathcal{O}$ to left multiplication by the reduction of α modulo p. Then*

$$\mathcal{O}' = \frac{1}{p}\,\mathrm{Ker}(\psi).$$

Proof. If $x \in \mathcal{O}'$, then $xI_p \subset I_P$, so $\psi(x)$ is the 0 endomorphism. Conversely, if $\psi(x)$ acts as 0 on I_p/pI_p, then clearly $xI_p \subset I_p$. □

Note that to give an algorithm one must also figure out how to explicitly compute I_p/pI_p and the kernel of this map (see the next section for more details).

4.3.4 General Factorization Algorithm of Buchman-Lenstra

We finally give an algorithm to factor $p\mathcal{O}_K$ in general. This is a summary of the algorithm described in more detail in [Coh93, §6.2].

Algorithm 4.3.7 (Factoring a Finite Separable Algebra)**.** Let A be a finite separable algebra over $\mathbb{F}_p$. This algorithm either shows that A is a field or finds a nontrivial idempotent in A, i.e., an $\varepsilon \in A$ such that $\varepsilon^2 = \varepsilon$ with $\varepsilon \neq 0$ and $\varepsilon \neq 1$.

1. The dimension of the kernel V of the map $x \mapsto x^p - x$ is equal to k. This is because abstractly we have that $A \approx A_1 \times \cdots \times A_k$, with each A_i a finite field extension of $\mathbb{F}_p$.

2. If $k = 1$ we are done. Terminate.

3. Otherwise, choose $\alpha \in V$ with $\alpha \notin \mathbb{F}_p$. (Think of $\mathbb{F}_p$ as the diagonal embedding of $\mathbb{F}_p$ in $A_1 \times \cdots \times A_k$). Compute powers of α and find the minimal polynomial $m(X)$ of α.

4. Since $V \approx \mathbb{F}_p \times \cdots \times F_p$ (k factors), the polynomial $m(X)$ is a square-free product of linear factors, that has degree > 1 since $\alpha \notin \mathbb{F}_p$. Thus we can compute a splitting $m(X) = m_1(X) \cdot m_2(X)$, where both $m_i(X)$ have positive degree.

5. Use the Euclidean algorithm in $\mathbb{F}_p[X]$ to find $U_1(X)$ and $U_2(X)$ such that
$$U_1 m_1 + U_2 m_2 = 1.$$

6. Let $\varepsilon = (U_1 m_1)(\alpha)$. Then we have
$$U_1 m_1 U_1 m_1 + U_2 m_2 U_1 m_1 = U_1 m_1,$$
so since $(m_1 m_2)(\alpha) = m(\alpha) = 01$, we have $\varepsilon^2 = \varepsilon$. Also, since $\gcd(U_1, m_2) = \gcd(U_2, m_1) = 1$, we have $\varepsilon \neq 0$ and $\varepsilon \neq 1$.

Given Algorithm 4.3.7, we compute an idempotent $\varepsilon \in A$, and observe that
$$A \cong \mathrm{Ker}(1 - \varepsilon) \oplus \mathrm{Ker}(\varepsilon).$$
Since $(1 - \varepsilon) + \varepsilon = 1$, we see that $(1 - \varepsilon)v + \varepsilon v = v$, so that the sum of the two kernels equals A. Also, if v is in the intersection of the two kernels, then $\varepsilon(v) = 0$ and $(1 - \varepsilon)(v) = 0$, so $0 = (1 - \varepsilon)(v) = v - \varepsilon(v) = v$, so the sum is direct.

Remark 4.3.8. The beginning of [Coh93, §6.2.4] suggests that one can just randomly find an $\alpha \in A$ such that $A \cong \mathbb{F}_p[x]/(m(x))$ where m is the minimal polynomial of α. This is usually the case, but is *wrong in general*, since there need *not* be an $\alpha \in A$ such that $A \cong \mathbb{F}_p[\alpha]$. For example, let $p = 2$ and K be as in Example 4.3.2. Then $A \cong \mathbb{F}_2 \times \mathbb{F}_2 \times \mathbb{F}_2$, which as a ring is not generated by a single element, since there are only 2 distinct linear polynomials over $\mathbb{F}_2[x]$.

Algorithm 4.3.9 (Factoring a General Prime Ideal)**.** Let $K = \mathbb{Q}(a)$ be a number field given by an algebraic integer a as a root of its minimal monic polynomial f of degree n. We assume that an order $\mathcal{O}$ has been given by a basis $w_1, \ldots, w_n$ and that $\mathcal{O}$ that contains $\mathbb{Z}[a]$. For any prime $p \in \mathbb{Z}$, the following algorithm computes the set of maximal ideals of $\mathcal{O}$ that contain p.

1. [Check if easy] If $p \nmid \mathrm{disc}(\mathbb{Z}[a])/\mathrm{disc}(\mathcal{O})$ (so $p \nmid [\mathcal{O} : \mathbb{Z}[a]]$), then using Theorem 4.2.3 we factor $p\mathcal{O}$.

2. [Compute radical] Let I be the *radical* of $p\mathcal{O}$, which is the ideal of elements $x \in \mathcal{O}$ such that $x^m \in p\mathcal{O}$ for some positive integer m. Note that $p\mathcal{O} \subset I$, i.e., $I \mid p\mathcal{O}$; also I is the product of the primes that divide p, without multiplicity. Using linear algebra over the finite field $\mathbb{F}_p$, we compute a basis for $I/p\mathcal{O}$ by computing the abelian subgroup of $\mathcal{O}/p\mathcal{O}$ of all nilpotent elements. This computes I, since $p\mathcal{O} \subset I$.

3. [Compute quotient by radical] Compute an $\mathbb{F}_p$ basis for

$$A = \mathcal{O}/I = (\mathcal{O}/p\mathcal{O})/(I/p\mathcal{O}).$$

   The second equality comes from the fact that $p\mathcal{O} \subset I$. Note that $\mathcal{O}/p\mathcal{O}$ is obtained by simply reducing the basis $w_1, \ldots, w_n$ modulo p. Thus this step entirely involves linear algebra modulo p.

4. [Decompose quotient] The ring A is isomorphic to the quotient of $\mathcal{O}$ by a radical ideal, so it decomposes as a product $A \cong A_1 \times \cdots \times A_k$ of finite fields. We find such a decomposition explicitly using Algorithm 4.3.7.

5. [Compute the maximal ideals over p] Each maximal ideal $\mathfrak{p}_i$ lying over p is the kernel of one of the compositions

$$\mathcal{O} \to A \approx A_1 \times \cdots \times A_k \to A_i.$$

Algorithm 4.3.9 finds all primes of $\mathcal{O}$ that contain the radical I of $p\mathcal{O}$. Every such prime clearly contains p, so to see that the algorithm is correct, we prove that the primes $\mathfrak{p}$ of $\mathcal{O}$ that contain p also contain I. If $\mathfrak{p}$ is a prime of $\mathcal{O}$ that contains p, then $p\mathcal{O} \subset \mathfrak{p}$. If $x \in I$ then $x^m \in p\mathcal{O}$ for some m, so $x^m \in \mathfrak{p}$ which implies that $x \in \mathfrak{p}$ by the primality of $\mathfrak{p}$. Thus $\mathfrak{p}$ contains I, as required. Note that we do not find the powers of primes that divide p in Algorithm 4.3.9; that's left to another algorithm that we will not discuss in this book.

Algorithm 4.3.9 was invented by J. Buchmann and H. W. Lenstra, though their paper seems to have never been published; however, the algorithm is described in detail in [Coh93, §6.2.5]. Incidentally, this chapter is based on Chapters 4 and 6 of [Coh93], which is highly recommended, and goes into much more detail about these algorithms.

Chapter 5

The Chinese Remainder Theorem

In this chapter, we prove the Chinese Remainder Theorem (CRT) for arbitrary commutative rings, then apply CRT to prove that every ideal in a Dedekind domain R is generated by at most two elements. We also prove that $\mathfrak{p}^n/\mathfrak{p}^{n+1}$ is (noncanonically) isomorphic to $R/\mathfrak{p}$ as an R-module, for any nonzero prime ideal $\mathfrak{p}$ of R. The tools we develop in this chapter will be used frequently to prove other results later.

5.1 The Chinese Remainder Theorem

5.1.1 CRT in the Integers

The classical CRT asserts that if $n_1, \ldots, n_r$ are integers that are coprime in pairs, and $a_1, \ldots, a_r$ are integers, then there exists an integer a such that $a \equiv a_i \pmod{n_i}$ for each $i = 1, \ldots, r$. Here "coprime in pairs" means that $\gcd(n_i, n_j) = 1$ whenever $i \neq j$; it does *not* mean that $\gcd(n_1, \ldots, n_r) = 1$, though it implies this. In terms of rings, CRT asserts that the natural map

$$\mathbb{Z}/(n_1 \cdots n_r)\mathbb{Z} \to (\mathbb{Z}/n_1\mathbb{Z}) \oplus \cdots \oplus (\mathbb{Z}/n_r\mathbb{Z}) \tag{5.1.1}$$

that sends $a \in \mathbb{Z}$ to its reduction modulo each n_i, is an isomorphism.

This map is *never* an isomorphism if the n_i are not coprime. Indeed, the cardinality of the image of the left hand side of (5.1.1) is $\mathrm{lcm}(n_1, \ldots, n_r)$, since it is the image of a cyclic group and $\mathrm{lcm}(n_1, \ldots, n_r)$ is the largest order of an element of the right hand side, whereas the cardinality of the right hand side is $n_1 \cdots n_r$.

The isomorphism (5.1.1) can alternatively be viewed as asserting that any system of linear congruences

$$x \equiv a_1 \pmod{n_1}, \quad x \equiv a_2 \pmod{n_2}, \quad \ldots, \quad x \equiv a_r \pmod{n_r}$$

with pairwise coprime moduli has a unique solution modulo $n_1 \cdots n_r$.

Before proving the CRT in more generality, we prove (5.1.1). There is a natural map

$$\phi : \mathbb{Z} \to (\mathbb{Z}/n_1\mathbb{Z}) \oplus \cdots \oplus (\mathbb{Z}/n_r\mathbb{Z})$$

given by projection onto each factor. Its kernel is

$$n_1\mathbb{Z} \cap \cdots \cap n_r\mathbb{Z}.$$

If n and m are integers, then $n\mathbb{Z} \cap m\mathbb{Z}$ is the set of multiples of both n and m, so $n\mathbb{Z} \cap m\mathbb{Z} = \mathrm{lcm}(n,m)\mathbb{Z}$. Since the n_i are coprime,

$$n_1\mathbb{Z} \cap \cdots \cap n_r\mathbb{Z} = n_1 \cdots n_r\mathbb{Z}.$$

Thus we have proved there is an inclusion

$$i : \mathbb{Z}/(n_1 \cdots n_r)\mathbb{Z} \hookrightarrow (\mathbb{Z}/n_1\mathbb{Z}) \oplus \cdots \oplus (\mathbb{Z}/n_r\mathbb{Z}). \tag{5.1.2}$$

This is half of the CRT; the other half is to prove that this map is surjective. In this case, it is clear that i is also surjective, because i is an injective map between finite sets of the same cardinality. We will, however, give a proof of surjectivity that doesn't use finiteness of the above two sets.

To prove surjectivity of i, note that since the n_i are coprime in pairs,

$$\gcd(n_1, n_2 \cdots n_r) = 1,$$

so there exists integers x, y such that

$$xn_1 + yn_2 \cdots n_r = 1.$$

To complete the proof, observe that $yn_2 \cdots n_r = 1 - xn_1$ is congruent to 1 modulo n_1 and 0 modulo $n_2 \cdots n_r$. Thus $(1, 0, \ldots, 0) = i(yn_2 \cdots n_r)$ is in the image of i. By a similar argument, we see that $(0, 1, \ldots, 0)$ and the other similar elements are all in the image of i, so i is surjective, which proves CRT.

5.1.2 CRT in General

Recall that *all rings in this book are commutative with unity.* Let R be such a ring.

Definition 5.1.1 (Coprime)**.** Ideals I and J of R are *coprime* if $I + J = (1)$.

For example, if I and J are nonzero ideals in a Dedekind domain, then they are coprime precisely when the prime ideals that appear in their two (unique) factorizations are disjoint.

Lemma 5.1.2. *If I and J are coprime ideals in a ring R, then $I \cap J = IJ$.*

Proof. Choose $x \in I$ and $y \in J$ such that $x + y = 1$. If $c \in I \cap J$ then

$$c = c \cdot 1 = c \cdot (x + y) = cx + cy \in IJ + IJ = IJ,$$

so $I \cap J \subset IJ$. The other inclusion is obvious by the definition of an ideal. $\square$

Lemma 5.1.3. *Suppose $I_1, \ldots, I_s$ are pairwise coprime ideals. Then I_1 is coprime to the product $I_2 \cdots I_s$.*

Proof. In the special case of a Dedekind domain, we could easily prove this lemma using unique factorization of ideals as products of primes (Theorem 3.2.6); instead, we give a direct general argument.

It suffices to prove the lemma in the case $s = 3$, since the general case then follows from induction. By assumption, there are $x_1 \in I_1, y_2 \in I_2$ and $a_1 \in I_1, b_3 \in I_3$ such

$$x_1 + y_2 = 1 \qquad \text{and} \qquad a_1 + b_3 = 1.$$

Multiplying these two relations yields

$$x_1a_1 + x_1b_3 + y_2a_1 + y_2b_3 = 1 \cdot 1 = 1.$$

The first three terms are in I_1 and the last term is in $I_2I_3 = I_2 \cap I_3$ (by Lemma 5.1.2), so I_1 is coprime to I_2I_3. □

Next we prove the general Chinese Remainder Theorem. We will apply this result with $R = \mathcal{O}_K$ in the rest of this chapter.

Theorem 5.1.4 (Chinese Remainder Theorem)**.** *Suppose $I_1, \ldots, I_r$ are nonzero ideals of a ring R such I_m and I_n are coprime for any $m \neq n$. Then the natural homomorphism $R \to \bigoplus_{n=1}^{r} R/I_n$ induces an isomorphism*

$$\psi : R/\prod_{n=1}^{r} I_n \to \bigoplus_{n=1}^{r} R/I_n.$$

Thus given any $a_n \in R$, for $n = 1, \ldots, r$, there exists some $a \in R$ such that $a \equiv a_n$ $(\mathrm{mod}\ I_n)$ for $n = 1, \ldots, r$; moreover, a is unique modulo $\prod_{n=1}^{r} I_n$.

Proof. Let $\varphi : R \to \bigoplus_{n=1}^{r} R/I_n$ be the natural map induced by reduction modulo the I_n. An inductive application of Lemma 5.1.2 implies that the kernel $\cap_{n=1}^{r} I_n$ of φ is equal to $\prod_{n=1}^{r} I_n$, so the map ψ of the theorem is injective.

Each projection $R \to R/I_n$ is surjective, so to prove that ψ is surjective, it suffices to show that $(1, 0, \ldots, 0)$ is in the image of φ, and similarly for the other factors. By Lemma 5.1.3, $J = \prod_{n=2}^{r} I_n$ is coprime to I_1, so there exists $x \in I_1$ and $y \in J$ such that $x + y = 1$. Then $y = 1 - x$ maps to 1 in R/I_1 and to 0 in R/J, hence to 0 in R/I_n for each $n \geq 2$, since $J \subset I_n$. □

5.2 Structural Applications of the CRT

Let $\mathcal{O}_K$ be the ring of integers of some number field K, and suppose I is a nonzero ideal of $\mathcal{O}_K$. As an abelian group $\mathcal{O}_K$ is free of rank $[K : \mathbb{Q}]$, and I is of finite index in $\mathcal{O}_K$, so I is generated by $[K : \mathbb{Q}]$ generators as an abelian group, so as an R-ideal I requires at most $[K : \mathbb{Q}]$ generators. The main result of this section asserts something better, namely that I can be generated *as an ideal* by at most two elements. Moreover, our result is more general, since it applies to an arbitrary Dedekind domain R. Thus, for the rest of this section, R is any Dedekind domain, e.g., the ring of integers of either a number field or function field. We use CRT to prove that every ideal of R can be generated by two elements.

Remark 5.2.1. Caution – If we replace R by an order in a Dedekind domain, i.e., by a subring of finite index,then there may be ideals that require far more than 2 generators.

Suppose that I is a nonzero integral ideal of R. If $a \in I$, then $(a) \subset I$, so I divides (a) and the quotient $(a)I^{-1}$ is an integral ideal. The following lemma asserts that (a) can be chosen so the quotient $(a)I^{-1}$ is coprime to any given ideal.

Lemma 5.2.2. *If I and J are nonzero integral ideals in R, then there exists an $a \in I$ such that the integral ideal $(a)I^{-1}$ is coprime to J.*

Before we give the proof in general, note that the lemma is trivial when I is principal, since if $I = (b)$, just take $a = b$, and then $(a)I^{-1} = (a)(a^{-1}) = (1)$ is coprime to every ideal.

Proof. Let $\mathfrak{p}_1, \ldots, \mathfrak{p}_r$ be the prime divisors of J. For each n, let v_n be the largest power of $\mathfrak{p}_n$ that divides I. Since $\mathfrak{p}_n^{v_n} \neq \mathfrak{p}_n^{v_n+1}$, we can choose an element $a_n \in \mathfrak{p}_n^{v_n}$ that is not in $\mathfrak{p}_n^{v_n+1}$. By Theorem 5.1.4 applied to the $r+1$ coprime integral ideals

$$\mathfrak{p}_1^{v_1+1}, \ldots, \mathfrak{p}_r^{v_r+1}, \, I \cdot \left(\prod \mathfrak{p}_n^{v_n}\right)^{-1},$$

there exists $a \in R$ such that

$$a \equiv a_n \pmod{\mathfrak{p}_n^{v_n+1}}$$

for all $n = 1, \ldots, r$ and also

$$a \equiv 0 \quad \left(\text{mod } I \cdot \left(\prod \mathfrak{p}_n^{v_n}\right)^{-1}\right).$$

To complete the proof we show that $(a)I^{-1}$ is not divisible by any $\mathfrak{p}_n$, or equivalently, that each $\mathfrak{p}_n^{v_n}$ exactly divides (a). First we show that $\mathfrak{p}_n^{v_n}$ divides (a). Because $a \equiv a_n \pmod{\mathfrak{p}_n^{v_n+1}}$, there exists $b \in \mathfrak{p}_n^{v_n+1}$ such that $a = a_n + b$. Since $a_n \in \mathfrak{p}_n^{v_n}$ and $b \in \mathfrak{p}_n^{v_n+1} \subset \mathfrak{p}_n^{v_n}$, it follows that $a \in \mathfrak{p}_n^{v_n}$, so $\mathfrak{p}_n^{v_n}$ divides (a). Now assume for the sake of contradiction that $\mathfrak{p}_n^{v_n+1}$ divides (a); then $a_n = a - b \in \mathfrak{p}_n^{v_n+1}$, which contradicts that we chose $a_n \notin \mathfrak{p}_n^{v_n+1}$. Thus $\mathfrak{p}_n^{v_n+1}$ does not divide (a), as claimed. □

Proposition 5.2.3. *Suppose I is a fractional ideal in a Dedekind domain R. Then there exist $a, b \in K$ such that $I = (a, b) = \{\alpha a + \beta b : \alpha, \beta \in R\}$.*

Proof. If $I = (0)$, then I is generated by 1 element and we are done. If I is not an integral ideal, then there is an $x \in K$ such that xI is an integral ideal, and the number of generators of xI is the same as the number of generators of I, so we may assume that I is an integral ideal.

Let a be *any* nonzero element of the integral ideal I. We will show that there is some $b \in I$ such that $I = (a, b)$. Let $J = (a)$. By Lemma 5.2.2, there exists $b \in I$ such that $(b)I^{-1}$ is coprime to (a). Since $a, b \in I$, we have $I \mid (a)$ and $I \mid (b)$, so $I \mid (a, b)$. Suppose $\mathfrak{p}^n \mid (a, b)$ with $\mathfrak{p}$ prime and $n \geq 1$. Then $\mathfrak{p}^n \mid (a)$ and $\mathfrak{p}^n \mid (b)$, so $\mathfrak{p} \nmid (b)I^{-1}$, since $(b)I^{-1}$ is coprime to (a). We have $\mathfrak{p}^n \mid (b) = I \cdot (b)I^{-1}$ and $\mathfrak{p} \nmid (b)I^{-1}$, so $\mathfrak{p}^n \mid I$. Thus by unique factorization of ideals in R we have that $(a, b) \mid I$. Since $I \mid (a, b)$ we conclude that $I = (a, b)$, as claimed. □

We can also use Theorem 5.1.4 to determine the R-module structure of $\mathfrak{p}^n/\mathfrak{p}^{n+1}$.

Proposition 5.2.4. *Let $\mathfrak{p}$ be a nonzero prime ideal of R, and let $n \geq 0$ be an integer. Then $\mathfrak{p}^n/\mathfrak{p}^{n+1} \cong R/\mathfrak{p}$ as R-modules.*

Proof [1]*.* Since $\mathfrak{p}^n \neq \mathfrak{p}^{n+1}$, by unique factorization, there is an element $b \in \mathfrak{p}^n$ such that $b \notin \mathfrak{p}^{n+1}$. Let $\varphi : R \to \mathfrak{p}^n/\mathfrak{p}^{n+1}$ be the R-module morphism defined by $\varphi(a) = ab$. The kernel of φ is $\mathfrak{p}$ since clearly $\varphi(\mathfrak{p}) = 0$ and if $\varphi(a) = 0$ then $ab \in \mathfrak{p}^{n+1}$, so $\mathfrak{p}^{n+1} \mid (a)(b)$, so $\mathfrak{p} \mid (a)$, since $\mathfrak{p}^{n+1}$ does not divide (b). Thus φ induces an injective R-module homomorphism $R/\mathfrak{p} \hookrightarrow \mathfrak{p}^n/\mathfrak{p}^{n+1}$.

It remains to show that φ is surjective, and this is where we will use Theorem 5.1.4. Suppose $c \in \mathfrak{p}^n$. By Theorem 5.1.4 there exists $d \in R$ such that

$$d \equiv c \pmod{\mathfrak{p}^{n+1}} \qquad \text{and} \qquad d \equiv 0 \pmod{(b)/\mathfrak{p}^n}.$$

We have $\mathfrak{p}^n \mid (d)$ since $d \in \mathfrak{p}^n$ and $(b)/\mathfrak{p}^n \mid (d)$ by the second displayed condition, so since $\mathfrak{p} \nmid (b)/\mathfrak{p}^n$, we have $(b) = \mathfrak{p}^n \cdot (b)/\mathfrak{p}^n \mid (d)$, hence $d/b \in R$. Finally

$$\varphi\left(\frac{d}{b}\right) \equiv \frac{d}{b} \cdot b \pmod{\mathfrak{p}^{n+1}} \equiv d \pmod{\mathfrak{p}^{n+1}} \equiv c \pmod{\mathfrak{p}^{n+1}},$$

so φ is surjective. □

Exercise 5.2.5. (See [Mar77, Thm. 22(a)]) Let R be a Dedekind domain and $\mathfrak{p}$ a nonzero prime ideal in R. Show that $\#(R/\mathfrak{p}^m) = \#(R/\mathfrak{p})^m$.

Note: $\#(R/\mathfrak{p})$ is not finite in general! For example, The ring of formal power series $k[[t]]$ for some field k is a Dedekind domain and the residue field at the prime (t) is k.

[*Hint*: Consider the exact sequence

$$0 \to \mathfrak{p}/\mathfrak{p}^m \to R/\mathfrak{p}^m \to R/\mathfrak{p}^{m-1} \to 0$$

and the chain

$$\mathfrak{p}^m \subseteq \mathfrak{p}^{m-1} \subseteq \cdots \subseteq \mathfrak{p}^2 \subseteq \mathfrak{p}.$$

]

Remark 5.2.6. There is one special case of the previous exercise that you probably have seen before: the size of $\mathbb{Z}/4\mathbb{Z}$ is the same as $(\mathbb{Z}/2\mathbb{Z})^2$. In fact you might have seen a proof of the fact that $\mathbb{Z}/n^m\mathbb{Z}$ has the same cardinality as $(\mathbb{Z}/n\mathbb{Z})^m$ in a standard group theory or abstract algebra course.

5.3 Computing Using the CRT

In order to explicitly compute an a as given by Theorem 5.1.4, usually one first precomputes elements $v_1, \ldots, v_r \in R$ such that $v_1 \mapsto (1, 0, \ldots, 0)$, $v_2 \mapsto (0, 1, \ldots, 0)$, etc. Then given any $a_n \in R$, for $n = 1, \ldots, r$, we obtain an $a \in R$ with $a_n \equiv a \pmod{I_n}$ by taking

$$a = a_1 v_1 + \cdots + a_r v_r.$$

How to compute the v_i depends on the ring R. It reduces to the following problem: Given coprimes ideals $I, J \subset R$, find $x \in I$ and $y \in J$ such that $x + y = 1$. If R is torsion free and of finite rank as a $\mathbb{Z}$-module, so $R \approx \mathbb{Z}^n$, then I, J can be represented by giving a basis in terms of a basis for R, and finding x, y such that $x + y = 1$ can then be reduced to a problem in linear algebra over $\mathbb{Z}$. More precisely, let A be the matrix whose columns are the concatenation of a basis for I with a basis for J. Suppose $v \in \mathbb{Z}^n$ corresponds to $1 \in \mathbb{Z}^n$. Then finding x, y such that $x + y = 1$ is equivalent to finding a solution $z \in \mathbb{Z}^n$ to the matrix equation $Az = v$. This latter linear algebra problem can be solved using Hermite normal form (see [Coh93, §4.7.1]), which is a generalization over $\mathbb{Z}$ of reduced row echelon form.

5.3.1 Sage

[[TODO]]

5.3.2 Magma

The Magma command `ChineseRemainderTheorem` implements the algorithm suggested by Theorem 5.1.4. In the following example, we compute a prime over (3) and a prime over (5) of the ring of integers of $\mathbb{Q}(\sqrt[3]{2})$, and find an element of $\mathcal{O}_K$ that is congruent to $\sqrt[3]{2}$ modulo one prime and 1 modulo the other.

```
> R<x> := PolynomialRing(RationalField());
> K<a> := NumberField(x^3-2);
> OK := RingOfIntegers(K);
> I := Factorization(3*OK)[1][1];
> J := Factorization(5*OK)[1][1];
> I;
Prime Ideal of OK
Two element generators:
    [3, 0, 0]
    [4, 1, 0]
```

```
> J;
Prime Ideal of OK
Two element generators:
    [5, 0, 0]
    [7, 1, 0]
> b := ChineseRemainderTheorem(I, J, OK!a, OK!1);
> K!b;
-4
> b - a in I;
true
> b - 1 in J;
true
```

5.3.3 PARI

There is also a CRT algorithm for number fields in PARI, but it is more cumbersome to use. First we defined $\mathbb{Q}(\sqrt[3]{2})$ and factor the ideals (3) and (5).

```
? f = x^3 - 2;
? k = nfinit(f);
? i = idealfactor(k,3);
? j = idealfactor(k,5);
```

Next we form matrix whose rows correspond to a product of two primes, one dividing 3 and one dividing 5:

```
? m = matrix(2,2);
? m[1,] = i[1,];
? m[1,2] = 1;
? m[2,] = j[1,];
```

Note that we set `m[1,2] = 1`, so the exponent is 1 instead of 3. We apply the CRT to obtain a lift in terms of the basis for $\mathcal{O}_K$.

```
? ?idealchinese
idealchinese(nf,x,y): x being a prime ideal factorization and y
a vector of elements, gives an element b such that
v_p(b-y_p)>=v_p(x) for all prime ideals p dividing x,
and v_p(b)>=0 for all other p.
? idealchinese(k, m, [x,1])
[0, 0, -1]~
? nfbasis(f)
[1, x, x^2]
```

Thus PARI finds the lift $-(\sqrt[3]{2})^2$, and we finish by verifying that this lift is correct. I couldn't figure out how to test for ideal membership in PARI, so here we just check that the prime ideal plus the element is not the unit ideal, which since the ideal is prime, implies membership.

```
? idealadd(k, i[1,1], -x^2 - x)
[3 1 2]
[0 1 0]
[0 0 1]
? idealadd(k, j[1,1], -x^2-1)
[5 2 1]
[0 1 0]
[0 0 1]
```

Chapter 6

Discriminants and Norms

In this chapter we give a geometric interpretation of the discriminant of an order in a number field. We also define norms of ideals and prove that the norm function is multiplicative. Discriminants of orders and norms of ideals will play a crucial role in our proof of finiteness of the class group in the next chapter.

6.1 Viewing $\mathcal{O}_K$ as a Lattice in a Real Vector Space

Let K be a number field of degree n. By the primitive element theorem, $K = \mathbb{Q}(\alpha)$ for some α, so we can write $K \cong \mathbb{Q}[x]/(f)$, where $f \in \mathbb{Q}[x]$ is the minimal polynomial of α. Because $\mathbb{C}$ is algebraically closed and f is irreducible, it has exactly $n = [K : \mathbb{Q}]$ complex roots. Each of these roots $z \in \mathbb{C}$ induces a homomorphism $\mathbb{Q}[x] \to \mathbb{C}$ given by $x \mapsto z$, whose kernel is the ideal (f). Thus we obtain n embeddings of $K \cong \mathbb{Q}[x]/(f)$ into $\mathbb{C}$:

$$\sigma_1, \ldots, \sigma_n : K \hookrightarrow \mathbb{C}.$$

Example 6.1.1. We compute the embeddings listed above for $K = \mathbb{Q}(\sqrt[3]{2})$.

```
K = QQ[2^(1/3)]; K
```

```
Number Field in a with defining polynomial x^3 - 2
```

```
K.complex_embeddings()
```

```
[Ring morphism: ...
  Defn: a |--> -0.629960524947 - 1.09112363597*I,
 Ring morphism: ...
  Defn: a |--> -0.629960524947 + 1.09112363597*I,
 Ring morphism: ...
  Defn: a |--> 1.25992104989]
```

Let $\sigma : K \hookrightarrow \mathbb{C}^n$ be the map $a \mapsto (\sigma_1(a), \ldots, \sigma_n(a))$, and let $V = \mathbb{R}\sigma(K)$ be the $\mathbb{R}$-span of the image $\sigma(K)$ of K inside $\mathbb{C}^n$.

Lemma 6.1.2. *Suppose $L \subset \mathbb{R}^n$ is a subgroup of the vector space $\mathbb{R}^n$. Then the induced topology on L is discrete if and only if for every $H > 0$ the set*

$$X_H = \{v \in L : \max\{|v_1|, \dots, |v_n|\} \leq H\}$$

is finite.

Proof. If L is not discrete, then there is a point $x \in L$ such that for every $\varepsilon > 0$ there is $y \in L$ such that $0 < |x - y| < \varepsilon$. By choosing smaller and smaller ε, we find infinitely many elements $x - y \in L$ all of whose coordinates are smaller than 1. The set X_1 is thus not finite. Thus if the sets X_H are all finite, L must be discrete.

Next assume that L is discrete and let $H > 0$ be any positive number. Then for every $x \in X_H$ there is an open ball B_x that contains x but no other element of L. Since X_H is closed and bounded, the Heine-Borel theorem implies that X_H is compact, so the open covering $\cup B_x$ of X_H has a finite subcover, which implies that X_H is finite, as claimed. □

Lemma 6.1.3. *If L if a free abelian group that is discrete in a finite-dimensional real vector space V and $\mathbb{R}L = V$, then the rank of L equals the dimension of V.*

Proof. Let $x_1, \dots, x_m \in L$ be an $\mathbb{R}$-vector space basis for $\mathbb{R}L$, and consider the $\mathbb{Z}$-submodule $M = \mathbb{Z}x_1 + \cdots + \mathbb{Z}x_m$ of L. If the quotient L/M is infinite, then there are infinitely many distinct elements of L that all lie in a fundamental domain for M, so Lemma 6.1.2 implies that L is not discrete. This is a contradiction, so L/M is finite, and the rank of L is $m = \dim(\mathbb{R}L)$, as claimed. □

Proposition 6.1.4. *The $\mathbb{R}$-vector space $V = \mathbb{R}\sigma(K)$ spanned by the image $\sigma(K)$ of K has dimension n.*

Proof. We prove this by showing that the image $\sigma(\mathcal{O}_K)$ is discrete. If $\sigma(\mathcal{O}_K)$ were not discrete it would contain elements all of whose coordinates are simultaneously arbitrarily small. The norm of an element $a \in \mathcal{O}_K$ is the product of the entries of $\sigma(a)$, so the norms of nonzero elements of $\mathcal{O}_K$ would go to 0. This is a contradiction, since the norms of nonzero elements of $\mathcal{O}_K$ are nonzero integers.

Since $\sigma(\mathcal{O}_K)$ is discrete in $\mathbb{C}^n$, Lemma 6.1.3 implies that $\dim(V)$ equals the rank of $\sigma(\mathcal{O}_K)$. Since σ is injective, $\dim(V)$ is the rank of $\mathcal{O}_K$, which equals n by Proposition 2.4.5. □

6.1.1 A Determinant

Suppose $w_1, \dots, w_n$ is a basis for $\mathcal{O}_K$, and let A be the matrix whose ith row is $\sigma(w_i)$. Consider the determinant $\det(A)$.

Example 6.1.5. The ring $\mathcal{O}_K = \mathbb{Z}[i]$ of integers of $K = \mathbb{Q}(i)$ has $\mathbb{Z}$-basis $w_1 = 1$, $w_2 = i$. The map $\sigma : K \to \mathbb{C}^2$ is given by

$$\sigma(a + bi) = (a + bi, a - bi) \in \mathbb{C}^2.$$

The image $\sigma(\mathcal{O}_K)$ is spanned by $(1,1)$ and $(i,-i)$. The determinant is

$$\left|\begin{pmatrix} 1 & 1 \\ i & -i \end{pmatrix}\right| = -2i.$$

Let $\mathcal{O}_K = \mathbb{Z}[\sqrt{2}]$ be the ring of integers of $K = \mathbb{Q}(\sqrt{2})$. The map σ is

$$\sigma(a + b\sqrt{2}) = (a + b\sqrt{2}, a - b\sqrt{2}) \in \mathbb{R}^2,$$

and

$$A = \begin{pmatrix} 1 & 1 \\ \sqrt{2} & -\sqrt{2} \end{pmatrix},$$

which has determinant $-2\sqrt{2}$.

As the above example illustrates, the determinant $\det(A)$ most certainly need not be an integer. However, as we will see, it's square is an integer that does not depend on our choice of basis for $\mathcal{O}_K$.

6.2 Discriminants

Suppose $w_1, \ldots, w_n$ is a basis for $\mathcal{O}_K$ as a $\mathbb{Z}$-module, which we view as a $\mathbb{Q}$-vector space. Let $\sigma : K \hookrightarrow \mathbb{C}^n$ be the embedding $\sigma(a) = (\sigma_1(a), \ldots, \sigma_n(a))$, where $\sigma_1, \ldots, \sigma_n$ are the distinct embeddings of K into $\mathbb{C}$. Let A be the matrix whose rows are $\sigma(w_1), \ldots, \sigma(w_n)$.

Changing our choice of basis for $\mathcal{O}_K$ is the same as left multiplying A by an integer matrix U of determinant ± 1, which changes $\det(A)$ by ± 1. This leads us to consider $\det(A)^2$ instead, which does not depend on the choice of basis; moreover, as we will see, $\det(A)^2$ is an integer. Note that

$$\begin{aligned}
\det(A)^2 &= \det(AA) = \det(A)\det(A) = \det(A)\det(A^t) = \det(AA^t) \\
&= \det\left(\sum_{k=1,\ldots,n} \sigma_k(w_i)\sigma_k(w_j)\right) = \det\left(\sum_{k=1,\ldots,n} \sigma_k(w_i w_j)\right) \\
&= \det(\mathrm{Tr}(w_i w_j)_{1 \leq i,j \leq n}),
\end{aligned}$$

so $\det(A)^2$ can be defined purely in terms of the trace without mentioning the embeddings σ_i. Moreover, if we change basis hence multiplying A by some U with determinant ± 1, then $\det(UA)^2 = \det(U)^2 \det(A)^2 = \det(A)^2$. Because $\det(A)$ is an algebraic integer and $\mathrm{Tr}(w_i w_j) \in \mathbb{Q}$, it follows that $\det(A)^2$ is an algebraic integer in $\mathbb{Q}$. Thus $\det(A)^2 \in \mathbb{Z}$ is well defined as a quantity associated to $\mathcal{O}_K$.

If we view K as a $\mathbb{Q}$-vector space, then $(x, y) \mapsto \mathrm{Tr}(xy)$ defines a bilinear pairing $K \times K \to \mathbb{Q}$ on K, which we call the *trace pairing*. The following lemma asserts that this pairing is nondegenerate, so $\det(\mathrm{Tr}(w_i w_j)) \neq 0$ hence $\det(A) \neq 0$.

Lemma 6.2.1. *The trace pairing is nondegenerate.*

Proof. If the trace pairing is degenerate, then there exists $0 \neq a \in K$ such that for every $b \in K$ we have $\mathrm{Tr}(ab) = 0$. In particularly, taking $b = a^{-1}$ we see that $0 = \mathrm{Tr}(aa^{-1}) = \mathrm{Tr}(1) = [K : \mathbb{Q}] > 0$, which is absurd. □

Definition 6.2.2 (Discriminant)**.** Suppose $a_1, \ldots, a_n$ is any $\mathbb{Q}$-basis of K. The *discriminant* of $a_1, \ldots, a_n$ is

$$\mathrm{Disc}(a_1, \ldots, a_n) = \det(\mathrm{Tr}(a_i a_j)_{1 \leq i,j \leq n}) \in \mathbb{Q}.$$

The *discriminant* $\mathrm{Disc}(\mathcal{O})$ of an order $\mathcal{O}$ in $\mathcal{O}_K$ is the discriminant of any $\mathbb{Z}$-basis for $\mathcal{O}$. The *discriminant* $d_K = \mathrm{Disc}(K)$ of the number field K is the discriminant of $\mathcal{O}_K$. Note that these discriminants are all nonzero by Lemma 6.2.1.

Remark 6.2.3. It is also standard to define the discriminant of a monic polynomial to be the product of the differences of the roots. If $\alpha \in \mathcal{O}_K$ with $\mathbb{Z}[\alpha]$ of finite index in $\mathcal{O}_K$, and f is the minimal polynomial of α, then $\mathrm{Disc}(f) = \mathrm{Disc}(\mathbb{Z}[\alpha])$. To see this, note that if we choose the basis $1, \alpha, \ldots, \alpha^{n-1}$ for $\mathbb{Z}[\alpha]$, then both discriminants are the square of the same Vandermonde determinant.

Remark 6.2.4. If S/R is an extension of Dedekind domains, with S a free R module of finite rank, then the above definition of a *relative* discriminant of S/R does not make sense in general. The problem is that R may have more units than $\{\pm 1\}$, in which case $\det(A^2)$ is not well defined. To generalize the notion of discriminant to arbitrary finite extensions of Dedekind domains, one must instead introduce a discriminant *ideal*.

Example 6.2.5. In Sage, we compute the discriminant of a number field or order using the discriminant command:

```
K.<a> = NumberField(x^2 - 5)
K.discriminant()
```

```
5
```

This also works for orders (notice the square factor below, which will be explained by Proposition 6.2.6):

```
R = K.order([7*a]); R
```

```
Order in Number Field in a with defining polynomial x^2 - 5
```

```
factor(R.discriminant())
```

```
2^2 * 5 * 7^2
```

Warning: In Magma $\mathrm{Disc}(K)$ is defined to be the discriminant of the polynomial you happened to use to define K.

```
> K := NumberField(x^2-5);
> Discriminant(K);
20
```

This is an intentional choice done for efficiency reasons, since computing the maximal order can take a long time. Nonetheless, it conflicts with standard mathematical usage, so beware.

The following proposition asserts that the discriminant of an order $\mathcal{O}$ in $\mathcal{O}_K$ is bigger than $\mathrm{disc}(\mathcal{O}_K)$ by a factor of the square of the index.

Proposition 6.2.6. *Suppose $\mathcal{O}$ is an order in $\mathcal{O}_K$. Then*

$$\mathrm{Disc}(\mathcal{O}) = \mathrm{Disc}(\mathcal{O}_K) \cdot [\mathcal{O}_K : \mathcal{O}]^2.$$

Proof. Let A be a matrix whose rows are the images via σ of a basis for $\mathcal{O}_K$, and let B be a matrix whose rows are the images via σ of a basis for $\mathcal{O}$. Since $\mathcal{O} \subset \mathcal{O}_K$ has finite index, there is an integer matrix C such that $CA = B$, and $|\det(C)| = [\mathcal{O}_K : \mathcal{O}]$. Then

$$\mathrm{Disc}(\mathcal{O}) = \det(B)^2 = \det(CA)^2 = \det(C)^2 \det(A)^2 = [\mathcal{O}_K : \mathcal{O}]^2 \cdot \mathrm{Disc}(\mathcal{O}_K).$$

□

Example 6.2.7. Let K be a number field and consider the quantity

$$D(K) = \gcd\{\mathrm{Disc}(\alpha) : \alpha \in \mathcal{O}_K \text{ and } [\mathcal{O}_K : \mathbb{Z}[\alpha]] < \infty\}.$$

One might hope that $D(K)$ is equal to the discriminant $\mathrm{Disc}(\mathcal{O}_K)$ of K, but this is not the case in general. Recall Example 4.3.2, in which we considered the field K generated by a root of $f = x^3 + x^2 - 2x + 8$. In that example, the discriminant of $\mathcal{O}_K$ is -503 with 503 prime:

```
K.<a> = NumberField(x^3 + x^2 - 2*x + 8)
factor(K.discriminant())
```

```
-1 * 503
```

For every $\alpha \in \mathcal{O}_K$, we have $2 \mid [\mathcal{O}_K : \mathbb{Z}[\alpha]]$, since $\mathcal{O}_K$ fails to be monogenic at 2. By Proposition 6.2.6, the discriminant of $\mathbb{Z}[\alpha]$ is divisible by 4 for all α, so $\mathrm{Disc}(\alpha)$ is also divisible by 4. This is why 2 is called an "inessential *discriminant* divisor".

Proposition 6.2.6 gives an algorithm for computing $\mathcal{O}_K$, albeit a slow one. Given K, find some order $\mathcal{O} \subset K$, and compute $d = \mathrm{Disc}(\mathcal{O})$. Factor d, and use the factorization to write $d = s \cdot f^2$, where f^2 is the largest square that divides d. Then the index of $\mathcal{O}$ in $\mathcal{O}_K$ is a divisor of f, and we (tediously) can enumerate all rings R with $\mathcal{O} \subset R \subset K$ and $[R : \mathcal{O}] \mid f$, until we find the largest one all of whose elements are integral. A much better algorithm is to proceed exactly as just described, except use the ideas of Section 4.3.3 to find a p-maximal order for each prime divisor of f, then add these p-maximal orders together.

Example 6.2.8. Consider the ring $\mathcal{O}_K = \mathbb{Z}[(1+\sqrt{5})/2]$ of integers of $K = \mathbb{Q}(\sqrt{5})$. The discriminant of the basis $1, a = (1+\sqrt{5})/2$ is

$$\mathrm{Disc}(\mathcal{O}_K) = \left|\begin{pmatrix} 2 & 1 \\ 1 & 3 \end{pmatrix}\right| = 5.$$

Let $\mathcal{O} = \mathbb{Z}[\sqrt{5}]$ be the order generated by $\sqrt{5}$. Then $\mathcal{O}$ has basis $1, \sqrt{5}$, so

$$\mathrm{Disc}(\mathcal{O}) = \left|\begin{pmatrix} 2 & 0 \\ 0 & 10 \end{pmatrix}\right| = 20 = [\mathcal{O}_K : \mathcal{O}]^2 \cdot 5,$$

hence $[\mathcal{O}_K : \mathcal{O}] = 2$.

Example 6.2.9. Consider the cubic field $K = \mathbb{Q}(\sqrt[3]{2})$, and let $\mathcal{O}$ be the order $\mathbb{Z}[\sqrt[3]{2}]$. Relative to the base $1, \sqrt[3]{2}, (\sqrt[3]{2})^2$ for $\mathcal{O}$, the matrix of the trace pairing is

$$A = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 0 & 6 \\ 0 & 6 & 0 \end{pmatrix}.$$

Thus

$$\mathrm{disc}(\mathcal{O}) = \det(A) = 108 = 2^2 \cdot 3^3.$$

Suppose we do not know that the ring of integers $\mathcal{O}_K$ is equal to $\mathcal{O}$. By Proposition 6.2.6, we have

$$\mathrm{Disc}(\mathcal{O}_K) \cdot [\mathcal{O}_K : \mathcal{O}]^2 = 2^2 \cdot 3^3,$$

so $3 \mid \mathrm{disc}(\mathcal{O}_K)$, and $[\mathcal{O}_K : \mathcal{O}] \mid 6$. Thus to prove $\mathcal{O} = \mathcal{O}_K$ it suffices to prove that $\mathcal{O}$ is 2-maximal and 3-maximal, which could be accomplished as described in Section 4.3.3.

6.3 Norms of Ideals

In this section we extend the notion of norm to ideals. This will be helpful in the next chapter, where we will prove that the group of fractional ideals modulo principal fractional ideals of a number field is finite by showing that every ideal is equivalent to an ideal with norm at most some bound. This is enough, because as we will see below there are only finitely many ideals of bounded norm.

Definition 6.3.1 (Lattice Index)**.** If L and M are two lattices in a vector space V, then the *lattice index* $[L : M]$ is by definition the absolute value of the determinant of any linear automorphism A of V such that $A(L) = M$.

For example, if $L = 2\mathbb{Z}$ and $M = 10\mathbb{Z}$, then

$$[L : M] = [2\mathbb{Z} : 10\mathbb{Z}] = \det([5]) = 5,$$

since 5 multiplies $2\mathbb{Z}$ onto $10\mathbb{Z}$.

The lattice index has the following properties:

- If $M \subset L$, then $[L : M] = \#(L/M)$.
- If M, L, N are any lattices in V, then
$$[L : N] = [L : M] \cdot [M : N].$$

Definition 6.3.2 (Norm of Fractional Ideal)**.** Suppose I is a fractional ideal of $\mathcal{O}_K$. The *norm* of I is the lattice index

$$\mathrm{Norm}(I) = [\mathcal{O}_K : I] \in \mathbb{Q}_{\geq 0},$$

or 0 if $I = 0$.

Note that if I is an integral ideal, then $\mathrm{Norm}(I) = \#(\mathcal{O}_K/I)$.

Lemma 6.3.3. *Suppose $a \in K$ and I is an integral ideal. Then*

$$\mathrm{Norm}(aI) = \left|\mathrm{Norm}_{K/\mathbb{Q}}(a)\right| \mathrm{Norm}(I).$$

Proof. By properties of the lattice index mentioned above we have

$$[\mathcal{O}_K : aI] = [\mathcal{O}_K : I] \cdot [I : aI] = \mathrm{Norm}(I) \cdot \left|\mathrm{Norm}_{K/\mathbb{Q}}(a)\right|.$$

Here we have used that $[I : aI] = \left|\mathrm{Norm}_{K/\mathbb{Q}}(a)\right|$, which is because left multiplication ℓ_a by a is an automorphism of K that sends I onto aI, so

$$[I : aI] = |\det(\ell_a)| = \left|\mathrm{Norm}_{K/\mathbb{Q}}(a)\right|.$$

$\square$

Proposition 6.3.4. *If I and J are fractional ideals, then*

$$\mathrm{Norm}(IJ) = \mathrm{Norm}(I) \cdot \mathrm{Norm}(J).$$

Proof. By Lemma 6.3.3, it suffices to prove this when I and J are integral ideals. If I and J are coprime, then Theorem 5.1.4 (the Chinese Remainder Theorem) implies that $\mathrm{Norm}(IJ) = \mathrm{Norm}(I) \cdot \mathrm{Norm}(J)$. Thus we reduce to the case when $I = \mathfrak{p}^m$ and $J = \mathfrak{p}^k$ for some prime ideal $\mathfrak{p}$ and integers m, k. By Proposition 5.2.4, which is a consequence of CRT, the filtration of $\mathcal{O}_K/\mathfrak{p}^n$ given by powers of $\mathfrak{p}$ has successive quotients isomorphic to $\mathcal{O}_K/\mathfrak{p}$. Thus we see that $\#(\mathcal{O}_K/\mathfrak{p}^n) = \#(\mathcal{O}_K/\mathfrak{p})^n$, which proves that $\mathrm{Norm}(\mathfrak{p}^n) = \mathrm{Norm}(\mathfrak{p})^n$. $\square$

Example 6.3.5. We compute some ideal norms using Sage.

```
K.<a> = NumberField(x^2 - 5)
I = K.fractional_ideal(a)
I.norm()
```

```
5
```

```
J = K.fractional_ideal(17)
J.norm()
```

```
289
```

We can also use functional notation:

```
norm(I*J)
```

```
1445
```

We will use the following proposition in the next chapter when we prove finiteness of class groups.

Proposition 6.3.6. *Fix a number field K. Let B be a positive integer. There are only finitely many integral ideals I of $\mathcal{O}_K$ with norm at most B.*

Proof. An integral ideal I is a subgroup of $\mathcal{O}_K$ of index equal to the norm of I. If G is any finitely generated abelian group, then there are only finitely many subgroups of G of index at most B. This is because the subgroups of index dividing an integer n are all subgroups of G that contain nG, and the group G/nG is finite. $\square$

Chapter 7

Finiteness of the Class Group

Frequently $\mathcal{O}_K$ is not a principal ideal domain. This chapter is about a way to understand how badly $\mathcal{O}_K$ fails to be a principal ideal domain. The class group of $\mathcal{O}_K$ measures this failure. As one sees in a course on Class Field Theory, the class group and its generalizations also yield deep insight into the extensions of K that are Galois with abelian Galois group.

In Section 7.1, we define the class group and state the main theorem of this chapter. We then illustrate the implications of this theorem in detail for the field $\mathbb{Q}(\sqrt{10})$, proving that it has class group of order 2. Next, we prove several geometric lemmas, building very heavily on ours results from Chapter 6. Finally, we close the section by giving a complete proof of finiteness of the class group, but leave an explicit upper bound as an exercise in calculus. In Section 7.2 we very briefly discuss how often number fields have class number 1. Finally, in Section 7.3 we further discuss how to compute class groups, though nothing we do in this book begins to approach the state of the art regarding such computations – for that, see Cohen's books.

7.1 The Class Group

Definition 7.1.1 (Class Group)**.** Let $\mathcal{O}_K$ be the ring of integers of a number field K. The *class group* C_K of K is the group of fractional ideals modulo the sugroup of principal fractional ideals (a), for $a \in K$.

Note that if we let $\mathrm{Div}(\mathcal{O}_K)$ denote the group of fractional ideals, then we have an exact sequence

$$0 \to \mathcal{O}_K^* \to K^* \to \mathrm{Div}(\mathcal{O}_K) \to C_K \to 0.$$

That the class group C_K is finite follows from the first part of the following theorem and that there are only finitely many ideals of norm less than a given integer (Proposition 6.3.6).

Theorem 7.1.2 (Finiteness of the Class Group)**.** *Let K be a number field. There is a constant $C_{r,s}$ that depends only on the number r, s of real and pairs of complex conjugate embeddings of K such that every ideal class of $\mathcal{O}_K$ contains an integral ideal of norm at most $C_{r,s}\sqrt{|d_K|}$, where $d_K = \mathrm{Disc}(\mathcal{O}_K)$. Thus by Proposition 6.3.6 the class group C_K of K is finite. In fact, one can take*

$$C_{r,s} = \left(\frac{4}{\pi}\right)^s \frac{n!}{n^n}.$$

The explicit bound in the theorem

$$M_K = \left(\frac{4}{\pi}\right)^s \frac{n!}{n^n} \cdot \sqrt{|d_K|}$$

is called the *Minkowski bound*. There are other better bounds, but they depend on unproven conjectures.

The following two examples illustrate how to apply Theorem 7.1.2 to compute C_K in simple cases.

Example 7.1.3. Let $K = \mathbb{Q}[i]$. Then $n = 2$, $s = 1$, and $|d_K| = 4$, so the Minkowski bound is

$$\sqrt{4} \cdot \left(\frac{4}{\pi}\right)^1 \frac{2!}{2^2} = \frac{4}{\pi} < 2.$$

Thus every fractional ideal is equivalent to an ideal of norm 1. Since (1) is the only ideal of norm 1, every ideal is principal, so C_K is trivial.

Example 7.1.4. Let $K = \mathbb{Q}(\sqrt{10})$. We have $\mathcal{O}_K = \mathbb{Z}[\sqrt{10}]$, so $n = 2$, $s = 0$, $|d_K| = 40$, and the Minkowski bound is

$$\sqrt{40} \cdot \left(\frac{4}{\pi}\right)^0 \cdot \frac{2!}{2^2} = 2 \cdot \sqrt{10} \cdot \frac{1}{2} = \sqrt{10} = 3.162277\ldots.$$

We compute the Minkowski bound in Sage as follows:

```
K = QQ[sqrt(10)]; K
```

```
Number Field in sqrt10 with defining polynomial x^2 - 10
```

```
B = K.minkowski_bound(); B
```

```
sqrt(10)
```

```
B.n()
```

```
3.16227766016838
```

Theorem 7.1.2 implies that every ideal class has a representative that is an integral ideal of norm 1, 2, or 3. The ideal $2\mathcal{O}_K$ is ramified in $\mathcal{O}_K$, so

$$2\mathcal{O}_K = (2, \sqrt{10})^2.$$

If $(2, \sqrt{10})$ were principal, say (α), then $\alpha = a + b\sqrt{10}$ would have norm ± 2. Then the equation

$$x^2 - 10y^2 = \pm 2, \tag{7.1.1}$$

would have an integer solution. But the squares mod 5 are $0, \pm 1$, so (7.1.1) has no solutions. Thus $(2, \sqrt{10})$ defines a nontrivial element of the class group, and it has order 2 since its square is the principal ideal $2\mathcal{O}_K$. Thus $2 \mid \#C_K$.

To find the integral ideals of norm 3, we factor $x^2 - 10$ modulo 3, and see that

$$3\mathcal{O}_K = (3, 2 + \sqrt{10}) \cdot (3, 4 + \sqrt{10}).$$

If either of the prime divisors of $3\mathcal{O}_K$ were principal, then the equation $x^2 - 10y^2 = \pm 3$ would have an integer solution. Since it does not have one mod 5, the prime divisors of $3\mathcal{O}_K$ are both nontrivial elements of the class group. Let

$$\alpha = \frac{4 + \sqrt{10}}{2 + \sqrt{10}} = \frac{1}{3} \cdot (1 + \sqrt{10}).$$

Then

$$(3, 2 + \sqrt{10}) \cdot (\alpha) = (3\alpha, 4 + \sqrt{10}) = (1 + \sqrt{10}, 4 + \sqrt{10}) = (3, 4 + \sqrt{10}),$$

so the classes over 3 are equal.

In summary, we now know that every element of C_K is equivalent to one of

$$(1), \quad (2, \sqrt{10}), \quad \text{or} \quad (3, 2 + \sqrt{10}).$$

Thus the class group is a group of order at most 3 that contains an element of order 2. Thus it must have order 2. We verify this in Sage below, where we also check that $(3, 2 + \sqrt{10})$ generates the class group.

```
K.<sqrt10> = QQ[sqrt(10)]; K
```

```
Number Field in sqrt10 with defining polynomial x^2 - 10
```

```
G = K.class_group(); G
```

```
Class group of order 2 with structure C2 of Number Field ...
```

```
G.0
```

```
Fractional ideal class (3, sqrt10 + 1)
```

```
G.0^2
```

```
Trivial principal fractional ideal class
```

```
G.0 == G( (3, 2 + sqrt10) )
```

```
True
```

Before proving Theorem 7.1.2, we prove a few lemmas. The strategy of the proof is to start with any nonzero ideal I, and prove that there is some nonzero $a \in K$ having very small norm, such that aI is an integral ideal. Then $\mathrm{Norm}(aI) = \mathrm{Norm}_{K/\mathbb{Q}}(a)\,\mathrm{Norm}(I)$ will be small, since $\mathrm{Norm}_{K/\mathbb{Q}}(a)$ is small. The trick is to determine precisely how small an a we can choose subject to the condition that aI is an integral ideal, i.e., that $a \in I^{-1}$.

Let S be a subset of $V = \mathbb{R}^n$. Then S is *convex* if whenever $x, y \in S$ then the line connecting x and y lies entirely in S. We say that S is *symmetric about the origin* if whenever $x \in S$ then $-x \in S$ also. If L is a lattice in the real vector space $V = \mathbb{R}^n$, then the *volume* of V/L is the volume of the compact real manifold V/L, which is the same thing as the absolute value of the determinant of any matrix whose rows form a basis for L.

Lemma 7.1.5 (Blichfeld)**.** *Let L be a lattice in $V = \mathbb{R}^n$, and let S be a bounded closed convex subset of V that is symmetric about the origin. If $\mathrm{Vol}(S) \geq 2^n \mathrm{Vol}(V/L)$, then S contains a nonzero element of L.*

Proof. First assume that $\mathrm{Vol}(S) > 2^n \mathrm{Vol}(V/L)$. If the map $\pi : \frac{1}{2}S \to V/L$ is injective, then

$$\frac{1}{2^n}\mathrm{Vol}(S) = \mathrm{Vol}\left(\frac{1}{2}S\right) \leq \mathrm{Vol}(V/L),$$

a contradiction. Thus π is not injective, so there exist $P_1 \neq P_2 \in \frac{1}{2}S$ such that $P_1 - P_2 \in L$. Because S is symmetric about the origin, $-P_2 \in \frac{1}{2}S$. By convexity, the average $\frac{1}{2}(P_1 - P_2)$ of P_1 and $-P_2$ is also in $\frac{1}{2}S$. Thus $0 \neq P_1 - P_2 \in S \cap L$, as claimed.

Next assume that $\mathrm{Vol}(S) = 2^n \cdot \mathrm{Vol}(V/L)$. Then for all $\varepsilon > 0$ there is $0 \neq Q_\varepsilon \in L \cap (1+\varepsilon)S$, since $\mathrm{Vol}((1+\varepsilon)S) > \mathrm{Vol}(S) = 2^n \cdot \mathrm{Vol}(V/L)$. If $\varepsilon < 1$ then the Q_ε are all in $L \cap 2S$, which is finite since $2S$ is bounded and L is discrete. Hence there exists nonzero $Q = Q_\varepsilon \in L \cap (1+\varepsilon)S$ for arbitrarily small ε. Since S is closed, $Q \in L \cap S$. □

Lemma 7.1.6. *If L_1 and L_2 are lattices in V, then*

$$\mathrm{Vol}(V/L_2) = \mathrm{Vol}(V/L_1) \cdot [L_1 : L_2].$$

Proof. Let A be an automorphism of V such that $A(L_1) = L_2$. Then A defines an isomorphism of real manifolds $V/L_1 \to V/L_2$ that changes volume by a factor of $|\det(A)| = [L_1 : L_2]$. The claimed formula then follows, since $[L_1 : L_2] = |\det(A)|$, by definition. □

Fix a number field K with ring of integers $\mathcal{O}_K$. Let $\sigma_1, \ldots, \sigma_r$ be the real embeddings of K and $\sigma_{r+1}, \ldots, \sigma_{r+s}$ be half the complex embeddings of K, with one representative of each pair of complex conjugate embeddings. Let $\sigma : K \to V = \mathbb{R}^n$ be the embedding

$$\sigma(x) = \big(\sigma_1(x), \sigma_2(x), \ldots, \sigma_r(x),$$
$$\mathrm{Re}(\sigma_{r+1}(x)), \ldots, \mathrm{Re}(\sigma_{r+s}(x)), \mathrm{Im}(\sigma_{r+1}(x)), \ldots, \mathrm{Im}(\sigma_{r+s}(x))\big),$$

Warning 7.1.7. Note that this σ is *not* exactly the same as the one at the beginning of Section 6.2 if $s > 0$.

Lemma 7.1.8. *Let σ be the map described above. Then*

$$\mathrm{Vol}(V/\sigma(\mathcal{O}_K)) = 2^{-s}\sqrt{|d_K|}.$$

Proof. Let $L = \sigma(\mathcal{O}_K)$. From a basis $w_1, \ldots, w_n$ for $\mathcal{O}_K$ we obtain a matrix A whose ith row is

$$(\sigma_1(w_i), \cdots, \sigma_r(w_i), \mathrm{Re}(\sigma_{r+1}(w_i)), \ldots, \mathrm{Re}(\sigma_{r+s}(w_i)), \mathrm{Im}(\sigma_{r+1}(w_i)), \ldots, \mathrm{Im}(\sigma_{r+s}(w_i)))$$

and whose determinant has absolute value equal to the volume of V/L. By doing the following three column operations, we obtain a matrix whose rows are exactly the images of the w_i under *all* embeddings of K into $\mathbb{C}$, which is the matrix that came up when we defined $d_K = \mathrm{Disc}(\mathcal{O}_K)$ in Section 6.2.

1. Add $i = \sqrt{-1}$ times each column with entries $\mathrm{Im}(\sigma_{r+j}(w_i))$ to the column with entries $\mathrm{Re}(\sigma_{r+j}(w_i))$.

2. Multiply all columns with entries $\mathrm{Im}(\sigma_{r+j}(w_i))$ by $-2i$, thus changing the determinant by $(-2i)^s$.

3. Add each column that now has entries $\mathrm{Re}(\sigma_{r+j}(w_i)) + i\mathrm{Im}(\sigma_{r+j}(w_i))$ to the the column with entries $-2i\mathrm{Im}(\sigma_{r+j}(w_i))$ to obtain columns $\mathrm{Re}(\sigma_{r+j}(w_i)) - i\mathrm{Im}(\sigma_{r+j}(w_i))$.

Recalling the definition of discriminant, we see that if B is the matrix constructed by doing the above three operations to A, then $\left|\det(B)^2\right| = |d_K|$. Thus

$$\mathrm{Vol}(V/L) = |\det(A)| = \left|(-2i)^{-s} \cdot \det(B)\right| = 2^{-s}\sqrt{|d_K|}.$$

□

Lemma 7.1.9. *If I is a fractional $\mathcal{O}_K$-ideal, then $\sigma(I)$ is a lattice in V and*

$$\mathrm{Vol}(V/\sigma(I)) = 2^{-s}\sqrt{|d_K|} \cdot \mathrm{Norm}(I).$$

Proof. Since $\sigma(\mathcal{O}_K)$ has rank n as an abelian group, and Lemma 7.1.8 implies that $\sigma(\mathcal{O}_K)$ also spans V, it follows that $\sigma(\mathcal{O}_K)$ is a lattice in V. For some nonzero integer m we have $m\mathcal{O}_K \subset I \subset \frac{1}{m}\mathcal{O}_K$, so $\sigma(I)$ is also a lattice in V. To prove the displayed volume formula, combine Lemmas 7.1.6 and 7.1.8 to get

$$\mathrm{Vol}(V/\sigma(I)) = \mathrm{Vol}(V/\sigma(\mathcal{O}_K)) \cdot [\mathcal{O}_K : I] = 2^{-s}\sqrt{|d_K|}\,\mathrm{Norm}(I).$$

□

Proof of Theorem 7.1.2. Let K be a number field with ring of integers $\mathcal{O}_K$, let $\sigma : K \hookrightarrow V \cong \mathbb{R}^n$ be as above, and let $f : V \to \mathbb{R}$ be the function defined by

$$f(x_1, \ldots, x_n) = |x_1 \cdots x_r \cdot (x_{r+1}^2 + x_{(r+1)+s}^2) \cdots (x_{r+s}^2 + x_n^2)|.$$

Notice that if $x \in K$ then $f(\sigma(x)) = |\operatorname{Norm}_{K/\mathbb{Q}}(x)|$, and for any $a \in \mathbb{R}$,

$$f(ax_1, \ldots, ax_n) = |a|^n f(x_1, \ldots, x_n).$$

Let $S \subset V$ be any fixed choice of closed, bounded, convex, subset with positive volume that is symmetric with respect to the origin. Since S is closed and bounded,

$$M = \max\{f(x) : x \in S\}$$

exists.

Suppose I is any fractional ideal of $\mathcal{O}_K$. Our goal is to prove that there is an integral ideal aI with small norm. We will do this by finding an appropriate $a \in I^{-1}$. By Lemma 7.1.9,

$$c = \operatorname{Vol}(V/\sigma(I^{-1})) = 2^{-s}\sqrt{|d_K|} \cdot \operatorname{Norm}(I)^{-1} = \frac{2^{-s}\sqrt{|d_K|}}{\operatorname{Norm}(I)}.$$

Let $\lambda = 2 \cdot \left(\frac{c}{v}\right)^{1/n}$, where $v = \operatorname{Vol}(S)$. Then

$$\operatorname{Vol}(\lambda S) = \lambda^n \operatorname{Vol}(S) = 2^n \cdot \frac{c}{v} \cdot v = 2^n \cdot c = 2^n \operatorname{Vol}(V/\sigma(I^{-1})),$$

so by Lemma 7.1.5 there exists $0 \neq b \in \sigma(I^{-1}) \cap \lambda S$. Let $a \in I^{-1}$ be such that $\sigma(a) = b$. Since M is the largest norm of an element of S, the largest norm of an element of $\sigma(I^{-1}) \cap \lambda S$ is at most $\lambda^n M$, so

$$\left|\operatorname{Norm}_{K/\mathbb{Q}}(a)\right| \leq \lambda^n M.$$

Since $a \in I^{-1}$, we have $aI \subset \mathcal{O}_K$, so aI is an integral ideal of $\mathcal{O}_K$ that is equivalent to I, and

$$\begin{aligned}
\operatorname{Norm}(aI) &= \left|\operatorname{Norm}_{K/\mathbb{Q}}(a)\right| \cdot \operatorname{Norm}(I) \\
&\leq \lambda^n M \cdot \operatorname{Norm}(I) \\
&\leq 2^n \frac{c}{v} M \cdot \operatorname{Norm}(I) \\
&= 2^n \cdot 2^{-s}\sqrt{|d_K|} \cdot M \cdot v^{-1} \\
&= 2^{r+s}\sqrt{|d_K|} \cdot M \cdot v^{-1}.
\end{aligned}$$

Notice that the right hand side is independent of I. It depends only on r, s, $|d_K|$, and our choice of S. This completes the proof of the theorem, except for the assertion that S can be chosen to give the claim at the end of the theorem which is shown in Exercise 7.1.10. $\square$

Exercise 7.1.10. Show that in the proof of Theorem 7.1.2, S can be chosen so that the final bound matches the statement of the theorem. This means S can be chosen so that

$$\mathrm{Norm}(aI) \le \left(\frac{4}{\pi}\right)^s \frac{n!}{n^n}\sqrt{|d_K|}.$$

[*Hint*: Consider the subset S of $\mathbb{R}^n$ defined by

$$|x_1| + \cdots + |x_r| + 2\left(\sqrt{x_{r+1}^2 + x_{(r+1)+s}^2} + \cdots + \sqrt{x_{r+s}^2 + x_{(r+s)+s}^2}\right) \le 1.$$

Suppose $a \in \mathcal{O}_K$ such that $\sigma(a) \in S$. What can you say about $\mathrm{Norm}_{K/\mathbb{Q}}(a)$? What is $\mathrm{Vol}(S)$?]

Corollary 7.1.11. *Suppose that $K \neq \mathbb{Q}$ is a number field. Then $|d_K| > 1$.*

Proof. Applying Theorem 7.1.2 to the unit ideal, we get the bound

$$1 \le \sqrt{|d_K|} \cdot \left(\frac{4}{\pi}\right)^s \frac{n!}{n^n}.$$

Thus

$$\sqrt{|d_K|} \ge \left(\frac{\pi}{4}\right)^s \frac{n^n}{n!},$$

and the right hand quantity is strictly bigger than 1 for any $s \le n/2$ and any $n > 1$, see Exercise 7.1.12. □

Exercise 7.1.12. Prove the statement at the end of the proof for Corollary 7.1.11, i.e. suppose $n > 1$ and $s \le \frac{n}{2}$ as above. Show that $\left(\frac{\pi}{4}\right)^s \frac{n^n}{n!} > 1$.

A prime p ramifies in $\mathcal{O}_K$ if and only if $d \mid d_K$, so the corollary implies that every nontrivial extension of $\mathbb{Q}$ is ramified at some prime.

7.2 Class Number 1

The fields of class number 1 are exactly the fields for which $\mathcal{O}_K$ is a principal ideal domain. How many such number fields are there? We still don't know.

Conjecture 7.2.1. *There are infinitely many number fields K such that the class group of K has order* 1.

For example, if we consider real quadratic fields $K = \mathbb{Q}(\sqrt{d})$, with d positive and square free, many class numbers are probably 1, as suggested by the Sage output below. It looks like 1's will keep appearing infinitely often, and indeed Cohen and Lenstra conjecture that they do ([CL84]).

```
for d in [2..1000]:
    if is_fundamental_discriminant(d):
        h = QuadraticField(d, 'a').class_number()
        if h == 1:
            print d,
```

```
5 8 12 13 17 21 24 28 29 33 37 41 44 53 56 57 61 69
73 76 77 88 89 92 93 97 101 109 113 124 129 133 137
141 149 152 157 161 172 173 177 181 184 188 193 197
201 209 213 217 233 236 237 241 248 249 253 268 269
277 281 284 293 301 309 313 317 329 332 337 341 344
349 353 373 376 381 389 393 397 409 412 413 417 421
428 433 437 449 453 457 461 472 489 497 501 508 509
517 521 524 536 537 541 553 556 557 569 573 581 589
593 597 601 604 613 617 632 633 641 649 652 653 661
664 668 669 673 677 681 701 709 713 716 717 721 737
749 753 757 764 769 773 781 789 796 797 809 813 821
824 829 844 849 853 856 857 869 877 881 889 893 908
913 917 921 929 933 937 941 953 956 973 977 989 997
```

In contrast, if we look at class numbers of quadratic imaginary fields, only a few at the beginning have class number 1.

```
for d in [-1,-2..-1000]:
    if is_fundamental_discriminant(d):
        h = QuadraticField(d, 'a').class_number()
        if h == 1:
            print d
```

```
-3 -4 -7 -8 -11 -19 -43 -67 -163
```

It is a theorem that was proved independently and in different ways by Heegner, Stark, and Baker that the above list of 9 fields is the complete list with class number 1. More generally, it is possible, using deep work of Gross, Zagier, and Goldfeld involving zeta functions and elliptic curves, to enumerate all quadratic number fields with a given class number (Mark Watkins has done very substantial work in this direction).

7.3 More About Computing Class Groups

If $\mathfrak{p}$ is a prime of $\mathcal{O}_K$, then the intersection $\mathfrak{p} \cap \mathbb{Z} = p\mathbb{Z}$ is a prime ideal of $\mathbb{Z}$. We say that $\mathfrak{p}$ *lies over* $p \in \mathbb{Z}$. Note $\mathfrak{p}$ lies over $p \in \mathbb{Z}$ if and only if $\mathfrak{p}$ is one of the prime factors in the factorization of the ideal $p\mathcal{O}_K$. Geometrically, $\mathfrak{p}$ is a point of $\mathrm{Spec}(\mathcal{O}_K)$ that lies over the point $p\mathbb{Z}$ of $\mathrm{Spec}(\mathbb{Z})$ under the map induced by the inclusion $\mathbb{Z} \hookrightarrow \mathcal{O}_K$ as described in Section 4.1.1.

Lemma 7.3.1. *Let K be a number field with ring of integers $\mathcal{O}_K$. Then the class group $\mathrm{Cl}(K)$ is generated by the prime ideals $\mathfrak{p}$ of $\mathcal{O}_K$ lying over primes $p \in \mathbb{Z}$ with $p \leq B_K = \sqrt{|d_K|} \cdot \left(\frac{4}{\pi}\right)^s \cdot \frac{n!}{n^n}$, where s is the number of complex conjugate pairs of embeddings $K \hookrightarrow \mathbb{C}$.*

Proof. Theorem 7.1.2 asserts that every ideal class in $\mathrm{Cl}(K)$ is represented by an ideal I with $\mathrm{Norm}(I) \le B_K$. Write $I = \prod_{i=1}^{m} \mathfrak{p}_i^{e_i}$, with each $e_i \ge 1$. Then by multiplicativity of the norm, each $\mathfrak{p}_i$ also satisfies $\mathrm{Norm}(\mathfrak{p}_i) \le B_K$. If $\mathfrak{p}_i \cap \mathbb{Z} = p\mathbb{Z}$, then $p \mid \mathrm{Norm}(\mathfrak{p}_i)$, since p is the residue characteristic of $\mathcal{O}_K/\mathfrak{p}$, so $p \le B_K$. Thus I is a product of primes $\mathfrak{p}$ that satisfies the norm bound of the lemma. $\square$

This is a sketch of how to compute $\mathrm{Cl}(K)$:

1. Use the algorithms of Chapter 4 to list all prime ideals $\mathfrak{p}$ of $\mathcal{O}_K$ that appear in the factorization of a prime $p \in \mathbb{Z}$ with $p \le B_K$.

2. Find the group generated by the ideal classes $[\mathfrak{p}]$, where the $\mathfrak{p}$ are the prime ideals found in step 1. (In general, this step can become fairly complicated.)

The following three examples illustrate computation of $\mathrm{Cl}(K)$ for $K = \mathbb{Q}(i), \mathbb{Q}(\sqrt{5})$ and $\mathbb{Q}(\sqrt{-6})$.

Example 7.3.2. We compute the class group of $K = \mathbb{Q}(i)$. We have

$$n = 2, \quad r = 0, \quad s = 1, \quad d_K = -4,$$

so

$$B_K = \sqrt{4} \cdot \left(\frac{4}{\pi}\right)^1 \cdot \left(\frac{2!}{2^2}\right) = \frac{8}{\pi} < 3.$$

Thus $\mathrm{Cl}(K)$ is generated by the prime divisors of 2. We have

$$2\mathcal{O}_K = (1+i)^2,$$

so $\mathrm{Cl}(K)$ is generated by the principal prime ideal $\mathfrak{p} = (1+i)$. Thus $\mathrm{Cl}(K) = 0$ is trivial.

Example 7.3.3. We compute the class group of $K = \mathbb{Q}(\sqrt{5})$. We have

$$n = 2, \quad r = 2, \quad s = 0, \quad d_K = 5,$$

so

$$B = \sqrt{5} \cdot \left(\frac{4}{\pi}\right)^0 \cdot \left(\frac{2!}{2^2}\right) < 3.$$

Thus $\mathrm{Cl}(K)$ is generated by the primes that divide 2. We have $\mathcal{O}_K = \mathbb{Z}[\gamma]$, where $\gamma = \frac{1+\sqrt{5}}{2}$ satisfies $x^2 - x - 1$. The polynomial $x^2 - x - 1$ is irreducible mod 2, so $2\mathcal{O}_K$ is prime. Since it is principal, we see that $\mathrm{Cl}(K) = 1$ is trivial.

Example 7.3.4. In this example, we compute the class group of $K = \mathbb{Q}(\sqrt{-6})$. We have

$$n = 2, \quad r = 0, \quad s = 1, \quad d_K = -24,$$

so

$$B = \sqrt{24} \cdot \frac{4}{\pi} \cdot \left(\frac{2!}{2^2}\right) \sim 3.1.$$

Thus $\mathrm{Cl}(K)$ is generated by the prime ideals lying over 2 and 3. We have $\mathcal{O}_K = \mathbb{Z}[\sqrt{-6}]$, and $\sqrt{-6}$ satisfies $x^2+6=0$. Factoring x^2+6 modulo 2 and 3 we see that the class group is generated by the prime ideals

$$\mathfrak{p}_2 = (2, \sqrt{-6}) \qquad \text{and} \qquad \mathfrak{p}_3 = (3, \sqrt{-6}).$$

Also, $\mathfrak{p}_2^2 = 2\mathcal{O}_K$ and $\mathfrak{p}_3^2 = 3\mathcal{O}_K$, so $\mathfrak{p}_2$ and $\mathfrak{p}_3$ define elements of order dividing 2 in $\mathrm{Cl}(K)$.

Is either $\mathfrak{p}_2$ or $\mathfrak{p}_3$ principal? Fortunately, there is an easier norm trick that allows us to decide. Suppose $\mathfrak{p}_2 = (\alpha)$, where $\alpha = a + b\sqrt{-6}$. Then

$$2 = \mathrm{Norm}(\mathfrak{p}_2) = |\mathrm{Norm}(\alpha)| = (a+b\sqrt{-6})(a-b\sqrt{-6}) = a^2+6b^2.$$

Trying the first few values of $a, b \in \mathbb{Z}$, we see that this equation has no solutions, so $\mathfrak{p}_2$ can not be principal. By a similar argument, we see that $\mathfrak{p}_3$ is not principal either. Thus $\mathfrak{p}_2$ and $\mathfrak{p}_3$ define elements of order 2 in $\mathrm{Cl}(K)$.

Does the class of $\mathfrak{p}_2$ equal the class of $\mathfrak{p}_3$? Since $\mathfrak{p}_2$ and $\mathfrak{p}_3$ define classes of order 2, we can decide this by finding the class of $\mathfrak{p}_2 \cdot \mathfrak{p}_3$. We have

$$\mathfrak{p}_2 \cdot \mathfrak{p}_3 = (2, \sqrt{-6}) \cdot (3, \sqrt{-6}) = (6, 2\sqrt{-6}, 3\sqrt{-6}) \subset (\sqrt{-6}).$$

The ideals on both sides of the inclusion have norm 6, so by multiplicativity of the norm, they must be the same ideal. Thus $\mathfrak{p}_2 \cdot \mathfrak{p}_3 = (\sqrt{-6})$ is principal, which shows $\mathfrak{p}_3$ is the inverse of $\mathfrak{p}_2$ in $\mathrm{Cl}(K)$. But $\mathfrak{p}_2$ had order 2, so $\mathfrak{p}_2$ and $\mathfrak{p}_3$ represent the same element of $\mathrm{Cl}(K)$. We conclude that

$$\mathrm{Cl}(K) = \langle \mathfrak{p}_2 \rangle = \mathbb{Z}/2\mathbb{Z}.$$

Chapter 8

Dirichlet's Unit Theorem

In this chapter we will prove Dirichlet's unit theorem, which is a structure theorem for the group of units of the ring of integers of a number field. The answer is remarkably simple: if K has r real and s pairs of complex conjugate embeddings, then

$$\mathcal{O}_K^* \approx \mathbb{Z}^{r+s-1} \times T,$$

where T is a finite cyclic group.

Many questions can be encoded as questions about the structure of the group of units. For example, Dirichlet's unit theorem explains the structure of the integer solutions (x, y) to Pell's equation $x^2 - dy^2 = 1$ (see Section 8.2.1).

8.1 The Group of Units

Definition 8.1.1 (Unit Group)**.** The *group of units* U_K associated to a number field K is the group of elements of $\mathcal{O}_K$ that have an inverse in $\mathcal{O}_K$.

Theorem 8.1.2 (Dirichlet)**.** *The group U_K is the product of a finite cyclic group of roots of unity with a free abelian group of rank $r+s-1$, where r is the number of real embeddings of K and s is the number of complex conjugate pairs of embeddings.*

(Note that we will prove a generalization of Theorem 8.1.2 in Section 12.1 below.)

We prove the theorem by defining a map $\varphi : U_K \to \mathbb{R}^{r+s}$, and showing that the kernel of φ is finite and the image of φ is a lattice in a hyperplane in $\mathbb{R}^{r+s}$. The trickiest part of the proof is showing that the image of φ spans a hyperplane, and we do this by a clever application of Blichfeld's Lemma 7.1.5.

Remark 8.1.3. Theorem 8.1.2 is due to Dirichlet who lived 1805–1859. Thomas Hirst described Dirichlet thus:

> He is a rather tall, lanky-looking man, with moustache and beard about to turn grey with a somewhat harsh voice and rather deaf. He was unwashed, with his cup of coffee and cigar. One of his failings is forgetting time, he pulls his watch out, finds it past three, and runs out without even finishing the sentence.

Koch wrote that:

> ... important parts of mathematics were influenced by Dirichlet. His proofs characteristically started with surprisingly simple observations, followed by extremely sharp analysis of the remaining problem.

I think Koch's observation nicely describes the proof we will give of Theorem 8.1.2.

Units have a simple characterization in terms of their norm.

Proposition 8.1.4. *An element $a \in \mathcal{O}_K$ is a unit if and only if* $\mathrm{Norm}_{K/\mathbb{Q}}(a) = \pm 1$.

Proof. Write $\mathrm{Norm} = \mathrm{Norm}_{K/\mathbb{Q}}$. If a is a unit, then a^{-1} is also a unit, and $1 = \mathrm{Norm}(a)\,\mathrm{Norm}(a^{-1})$. Since both $\mathrm{Norm}(a)$ and $\mathrm{Norm}(a^{-1})$ are integers, it follows that $\mathrm{Norm}(a) = \pm 1$. Conversely, if $a \in \mathcal{O}_K$ and $\mathrm{Norm}(a) = \pm 1$, then the equation $aa^{-1} = 1 = \pm\,\mathrm{Norm}(a)$ implies that $a^{-1} = \pm\,\mathrm{Norm}(a)/a$. But $\mathrm{Norm}(a)$ is the product of the images of a in $\mathbb{C}$ by all embeddings of K into $\mathbb{C}$, so $\mathrm{Norm}(a)/a$ is also a product of images of a in $\mathbb{C}$, hence a product of algebraic integers, hence an algebraic integer. Thus $a^{-1} \in K \cap \overline{\mathbb{Z}} = \mathcal{O}_K$, which proves that a is a unit. $\square$

Remark 8.1.5. Proposition 8.1.4 is false if we replace $\mathcal{O}_K$ by K. For example, if α is a root of $x^2 - \frac{1}{2}x + 1$, then α has norm ± 1, but α is not a unit of $\mathcal{O}_K$, since $\alpha \notin \mathcal{O}_K$. To general Proposition 8.1.4 to an arbitrary finite extension R/S of Dedekind domains, we replace ± 1 by "an element of S^*".

Let r be the number of real and s the number of complex conjugate embeddings of K into $\mathbb{C}$, so $n = [K : \mathbb{Q}] = r + 2s$. Define the *log map*

$$\varphi : U_K \to \mathbb{R}^{r+s}$$

by

$$\varphi(a) = (\log|\sigma_1(a)|, \ldots, \log|\sigma_{r+s}(a)|).$$

Here $|z|$ is the usual absolute value of $z = x + iy \in \mathbb{C}$ (so $|z| = \sqrt{x^2 + y^2}$), and the maps σ_i are the same as those described in Lemma 7.1.8. In particular, $\sigma_1, \ldots, \sigma_r$ represent all real embeddings $K \to \mathbb{R}$ and $\sigma_{r+1}, \ldots, \sigma_{r+s}$ represent half of the complex embeddings $K \to \mathbb{C}$, with one representative for each pair of complex conjugate embeddings.

Lemma 8.1.6. *The image of φ lies in the hyperplane*

$$H = \{(x_1, \ldots, x_{r+s}) \in \mathbb{R}^{r+s} : x_1 + \cdots + x_r + 2x_{r+1} + \cdots + 2x_{r+s} = 0\}. \qquad (8.1.1)$$

Proof. If $a \in U_K$, then by Proposition 8.1.4,

$$\left(\prod_{i=1}^{r} |\sigma_i(a)|\right) \cdot \left(\prod_{i=r+1}^{r+s} |\sigma_i(a)|^2\right) = \left|\mathrm{Norm}_{K/\mathbb{Q}}(a)\right| = 1.$$

Taking logs of both sides proves the lemma. □

Lemma 8.1.7. *The kernel of φ is finite.*

Proof. We have

$$\begin{aligned} \mathrm{Ker}(\varphi) &\subset \{a \in \mathcal{O}_K \colon |\sigma_i(a)| = 1 \text{ for } i = 1, \ldots, r+s\} \\ \sigma(\mathrm{Ker}(\varphi)) &\subset \sigma(\mathcal{O}_K) \cap X \end{aligned}$$

where $\sigma : \mathcal{O}_K \to \mathbb{C}^{r+s}$ is given by $\sigma(a) = (\sigma_1(a), \ldots, \sigma_{r+s}(a))$ and X is the set $\{(z_1, \ldots, z_{r+s}) \in \mathbb{C}^{r+s} \colon |z_i| \leq 1\}$. Since $\sigma(\mathcal{O}_K)$ is a lattice (see Proposition 2.4.5) and X is compact, the intersection $\sigma(\mathcal{O}_K) \cap X$ is finite. This implies $\mathrm{Ker}(\varphi)$ is finite. □

Lemma 8.1.8. *The kernel of φ is a finite cyclic group.*

Proof. Lemma 8.1.7 implies that $\ker(\varphi)$ is a finite group. It is a general fact that any finite subgroup G of the multiplicative group K^* of a field is cyclic (see Exercise 8.1.9). □

Exercise 8.1.9. Finish the proof of Lemma 8.1.8 by showing that for a field K, every finite subgroup G of the multiplicative group K^* is cyclic.

[*Hint*: Every element in G satisfies a polynomial of the form $x^n - 1$. Recall that a polynomial of degree n over a field has at most n distinct roots. Now consider the orders of the elements of G.]

To prove Theorem 8.1.2, it suffices to prove that $\mathrm{Im}(\varphi)$ is a lattice in the hyperplane H of (8.1.1), which we view as a vector space of dimension $r+s-1$.

Define an embedding

$$\sigma : K \hookrightarrow \mathbb{R}^n \tag{8.1.2}$$

given by $\sigma(x) = (\sigma_1(x), \ldots, \sigma_{r+s}(x))$, where we view $\mathbb{C} \cong \mathbb{R} \times \mathbb{R}$ via $a + bi \mapsto (a, b)$. Thus this is the embedding

$$\begin{aligned} x \mapsto \big(&\sigma_1(x), \sigma_2(x), \ldots, \sigma_r(x), \\ &\mathrm{Re}(\sigma_{r+1}(x)), \mathrm{Im}(\sigma_{r+1}(x)), \ldots, \mathrm{Re}(\sigma_{r+s}(x)), \mathrm{Im}(\sigma_{r+s}(x))\big). \end{aligned}$$

Lemma 8.1.10. *The image $\varphi : U_K \to \mathbb{R}^{r+s}$ is discrete.*

Proof. Let X be a bounded subset of $\mathbb{R}^{r+s}$. We will show that the intersection $\varphi(U_K) \cap X$ is finite. Since X is bounded, for any $u \in Y = \varphi^{-1}(X) \subset U_K$ the coordinates of $\sigma(u)$ are bounded, since $|\log(x)|$ is bounded on bounded subsets of $[1, \infty)$. Thus $\sigma(Y)$ is a bounded subset of $\mathbb{R}^n$. Since $\sigma(Y) \subset \sigma(\mathcal{O}_K)$, and $\sigma(\mathcal{O}_K)$ is a lattice in $\mathbb{R}^n$, it follows that $\sigma(Y)$ is finite; moreover, σ is injective, so Y is finite. Thus $\varphi(U_K) \cap X \subset \varphi(Y) \cap X$ is finite. □

We will use the following lemma in our proof of Theorem 8.1.2.

Lemma 8.1.11. *Let $n \geq 2$ be an integer, suppose $w_1, \ldots, w_n \in \mathbb{R}$ are not all equal, and suppose $A, B \in \mathbb{R}$ are positive. Then there exist $d_1, \ldots, d_n \in \mathbb{R}_{>0}$ such that*

$$|w_1 \log(d_1) + \cdots + w_n \log(d_n)| > B$$

and $d_1 \cdots d_n = A$.

Proof. Order the w_i so that $w_1 \neq 0$. By hypothesis there exists a w_j such that $w_j \neq w_1$, and again re-ordering we may assume that $j = 2$. Set $d_3 = \cdots = d_{r+s} = 1$. Suppose d_1, d_2 are any positive real numbers with $d_1 d_2 = A$. Since $\log(1) = 0$,

$$\begin{aligned} \left| \sum_{i=1}^n w_i \log(d_i) \right| &= |w_1 \log(d_1) + w_2 \log(d_2)| \\ &= |w_1 \log(d_1) + w_2 \log(A/d_1)| \\ &= |(w_1 - w_2) \log(d_1) + w_2 \log(A)| \end{aligned}$$

Since $w_1 \neq w_2$, we have $|(w_1 - w_2) \log(d_1) + w_2 \log(A)| \to \infty$ as $d_1 \to \infty$. It is thus possible to choose the d_i as in the lemma. □

Proof of Theorem 8.1.2. By Lemma 8.1.10, the image $\varphi(U_K)$ is discrete, so it remains to show that $\varphi(U_K)$ spans H. Let W be the $\mathbb{R}$-span of the image $\varphi(U_K)$, and note that W is a subspace of H, by Lemma 8.1.6. We will show that $W = H$ indirectly by showing that if $v \notin H^\perp$, where $\perp$ is the orthogonal complement with respect to the dot product on $\mathbb{R}^{r+s}$, then $v \notin W^\perp$. This will show that $W^\perp \subset H^\perp$, hence that $H \subset W$, as required.

Thus suppose $z = (z_1, \ldots, z_{r+s}) \not\in H^\perp$. Define a function $f : K^* \to \mathbb{R}$ by

$$f(x) = z_1 \log|\sigma_1(x)| + \cdots + z_{r+s} \log|\sigma_{r+s}(x)|. \tag{8.1.3}$$

Note that $f(U_K) = \{0\}$ if and only if $z \in W^\perp$, so to show that $z \not\in W^\perp$ we show that there exists some $u \in U_K$ with $f(u) \neq 0$.

Let

$$A = \sqrt{|d_K|} \cdot \left(\frac{2}{\pi}\right)^s \in \mathbb{R}_{>0}.$$

Choose any positive real numbers $c_1, \ldots, c_{r+s} \in \mathbb{R}_{>0}$ such that

$$c_1 \cdots c_r \cdot (c_{r+1} \cdots c_{r+s})^2 = A.$$

Let

$$\begin{aligned} S = \{(x_1, \ldots, x_n) \in \mathbb{R}^n : \\ &|x_i| \leq c_i \text{ for } 1 \leq i \leq r, \\ &|x_i^2 + x_{i+s}^2| \leq c_i^2 \text{ for } r < i \leq r+s\} \subset \mathbb{R}^n. \end{aligned}$$

Then S is closed, bounded, convex, symmetric with respect to the origin, and of dimension $r + 2s$, since S is a product of r intervals and s discs, each of which has these properties. Viewing S as a product of intervals and discs, we see that the volume of S is

$$\mathrm{Vol}(S) = \prod_{i=1}^{r}(2c_i) \cdot \prod_{i=1}^{s}(\pi c_i^2) = 2^r \cdot \pi^s \cdot A = 2^{r+s}\sqrt{|d_K|} = 2^n \cdot 2^{-s}\sqrt{|d_K|}.$$

Recall Blichfeldt's Lemma 7.1.5, which asserts that if L is a lattice and S is closed, bounded, etc., and has volume at least $2^n \cdot \mathrm{Vol}(V/L)$, then $S \cap L$ contains a nonzero element. To apply this lemma, we take $L = \sigma(\mathcal{O}_K) \subset \mathbb{R}^n$, where σ is as in (8.1.2). By Lemma 7.1.8, we have $\mathrm{Vol}(\mathbb{R}^n/L) = 2^{-s}\sqrt{|d_K|}$. To check the hypothesis of Blichfeld's lemma, note that

$$\mathrm{Vol}(S) = 2^n \cdot 2^{-s}\sqrt{|d_K|} = 2^n\,\mathrm{Vol}(\mathbb{R}^n/L).$$

Thus there exists a nonzero element x in $S \cap \sigma(\mathcal{O}_K)$. Let $a \in \mathcal{O}_K$ with $\sigma(a) = x$, then $\sigma(a) \in S$, so $|\sigma_i(a)| \leq c_i$ for $1 \leq i \leq r+s$. We then have

$$\begin{aligned} \left|\mathrm{Norm}_{K/\mathbb{Q}}(a)\right| &= \left|\prod_{i=1}^{r+2s} \sigma_i(a)\right| \\ &= \prod_{i=1}^{r} |\sigma_i(a)| \cdot \prod_{i=r+1}^{s} |\sigma_i(a)|^2 \\ &\leq c_1 \cdots c_r \cdot (c_{r+1} \cdots c_{r+s})^2 = A. \end{aligned}$$

Since $a \in \mathcal{O}_K$ is nonzero, we also have

$$\left|\mathrm{Norm}_{K/\mathbb{Q}}(a)\right| \geq 1.$$

Moreover, if for any $i \leq r$, we have $|\sigma_i(a)| < \frac{c_i}{A}$, then

$$1 \leq \left|\mathrm{Norm}_{K/\mathbb{Q}}(a)\right| < c_1 \cdots \frac{c_i}{A} \cdots c_r \cdot (c_{r+1} \cdots c_{r+s})^2 = \frac{A}{A} = 1,$$

a contradiction, so $|\sigma_i(a)| \geq \frac{c_i}{A}$ for $i = 1, \ldots, r$. Likewise, $|\sigma_i(a)|^2 \geq \frac{c_i^2}{A}$, for $i = r+1, \ldots, r+s$. Rewriting this we have

$$\frac{c_i}{|\sigma_i(a)|} \leq A \quad \text{for } i \leq r \quad \text{and} \quad \left(\frac{c_i}{|\sigma_i(a)|}\right)^2 \leq A \quad \text{for } i = r+1, \ldots, r+s. \quad (8.1.4)$$

Recall that our overall strategy is to use an appropriately chosen a to construct a unit $u \in U_K$ such $f(u) \neq 0$. First, let $b_1, \ldots, b_m$ be representative generators for the finitely many nonzero principal ideals of $\mathcal{O}_K$ of norm at most A. Since $\left|\mathrm{Norm}_{K/\mathbb{Q}}(a)\right| \leq A$, we have $(a) = (b_j)$, for some j, so there is a unit $u \in \mathcal{O}_K$ such that $a = ub_j$.

Let

$$t = t_{c_1,\ldots,c_{r+s}} = z_1 \log(c_1) + \cdots + z_{r+s} \log(c_{r+s}),$$

and recall $f : K^* \to \mathbb{R}$ defined in (8.1.3) above. We have

$$\begin{aligned}
|f(u) - t| &= |f(a) - f(b_j) - t| \\
&\leq |f(b_j)| + |t - f(a)| \\
&= |f(b_j)| + |z_1(\log(c_1) - \log(|\sigma_1(a)|)) + \cdots + z_{r+s}(\log(c_{r+s}) - \log(|\sigma_{r+s}(a)|))| \\
&= |f(b_j)| + |z_1 \cdot \log(c_1/|\sigma_1(a)|) + \cdots + \frac{z_{r+s}}{2} \cdot \log((c_{r+s}/|\sigma_{r+s}(a)|)^2)| \\
&\leq |f(b_j)| + \log(A) \cdot \left(\sum_{i=1}^{r} |z_i| + \frac{1}{2} \cdot \sum_{i=r+1}^{s} |z_i|\right) \stackrel{\text{def}}{=} B_j.
\end{aligned}$$

In the last step we use (8.1.4).

Let $B = \max_j B_j$, and note that B does not depend on the choice of the c_i; in fact, it only depends our *fixed* choice of z and on the field K. Moreover, for any choice of the c_i as above, we have

$$|f(u) - t| \leq B.$$

If we can choose positive real numbers c_i such that

$$\begin{aligned}
c_1 \cdots c_r \cdot (c_{r+1} \cdots c_{r+s})^2 &= A \\
|t_{c_1,\ldots,c_{r+s}}| &> B,
\end{aligned}$$

then the fact that $|f(u) - t| \leq B$ would then imply that $|f(u)| > 0$, which is exactly what we aimed to prove.

If $r+s=1$, then we are trying to prove that $\varphi(U_K)$ is a lattice in $\mathbb{R}^0 = \mathbb{R}^{r+s-1}$, which is automatically true, so assume $r+s>1$. To finish the proof, we explain how to use Lemma 8.1.11 to choose c_i such that $|t|>B$. We have

$$\begin{aligned} t &= z_1 \log(c_1) + \cdots + z_{r+s}\log(c_{r+s}) \\ &= z_1 \log(c_1) + \cdots + z_r \log(c_r) + \frac{1}{2}\cdot z_{r+1}\log(c_{r+1}^2) + \cdots + \frac{1}{2}\cdot z_{r+s}\log(c_{r+s}^2) \\ &= w_1 \log(d_1) + \cdots + w_r\log(d_r) + w_{r+1}\log(d_{r+1}) + \cdots + \cdot w_{r+s}\log(d_{r+s}), \end{aligned}$$

where $w_i = z_i$ and $d_i = c_i$ for $i \leq r$, and $w_i = \frac{1}{2}z_i$ and $d_i = c_i^2$ for $r < i \leq r+s$. The condition that $z \notin H^\perp$ is that the w_i are not all the same, and in our new coordinates the lemma is equivalent to showing that $|\sum_{i=1}^{r+s} w_i \log(d_i)| > B$, subject to the condition that $\prod_{i=1}^{r+s} d_i = A$. But this is exactly what Lemma 8.1.11 shows. It is thus possible to find a unit u such that $|f(u)|>0$. Thus $z \notin W^\perp$, so $W^\perp \subset H^\perp$, whence $H \subset W$, which finishes the proof of Theorem 8.1.2. □

8.2 Examples with Sage

8.2.1 Pell's Equation

The so-called "Pell's equation" is $x^2 - dy^2 = 1$ with $d>0$ square free, and we seek integer solutions x, y to this equation. If $x + y\sqrt{d} \in K = \mathbb{Q}(\sqrt{d})$, then

$$\mathrm{Norm}(x+y\sqrt{d}) = (x+y\sqrt{d})(x-y\sqrt{d}) = x^2 - dy^2.$$

Thus if (x,y) are integers such that $x^2-dy^2=1$, then $\alpha = x + \sqrt{d}y \in \mathcal{O}_K$ has norm 1, so by Proposition 8.1.4 we have $\alpha \in U_K$. The integer solutions to Pell's equation thus form a finite-index subgroup of the group of units in the ring of integers of $\mathbb{Q}(\sqrt{d})$. Dirichlet's unit theorem implies that for any d the solutions to Pell's equation with x, y not both negative forms an infinite cyclic group, which is a fact that takes substantial work to prove using only elementary number theory (for example, using continued fractions).

We first solve Pell's equation $x^2 - 5y^2 = 1$ with $d = 5$ by finding the units of the ring of integers of $\mathbb{Q}(\sqrt{5})$ using Sage. Recall from Example 2.3.19 that the ring of integers of $\mathbb{Q}(\sqrt{5})$ is $\mathbb{Z}[\frac{1+\sqrt{5}}{2}]$

```
K.<sqrt5> = QuadraticField(5)
G = K.unit_group(); G
```

```
Unit group with structure C2 x Z of Number Field in sqrt5 with
defining polynomial x^2 - 5
```

```
u = G.1.value(); v = G.0.value(); (u, v)
```

```
(1/2*sqrt5 + 1/2, -1)
```

The subgroup of cubes of u gives us the units with integer x, y (not both negative).

```
[u^(3*i) for i in [0..9]]
```

```
[1, sqrt5 + 2, 4*sqrt5 + 9, 17*sqrt5 + 38, 72*sqrt5 + 161, \
305*sqrt5 + 682, 1292*sqrt5 + 2889, 5473*sqrt5 + 12238, \
23184*sqrt5 + 51841, 98209*sqrt5 + 219602]
```

However, the norm of $u = \frac{1+\sqrt{5}}{2}$ is -1. So the 6th powers of u will generate solutions to Pell's Equation. We can also list the coefficients for these powers as follows.

```
[list(u^(6*i)) for i in [0..7]]
```

```
[[1, 0], [9, 4], [161, 72], [2889, 1292], [51841, 23184], \
[930249, 416020], [16692641, 7465176], [299537289, 133957148]]
```

Remark 8.2.1. A great article about Pell's equation is [Len02]. The MathSciNet review begins: "This wonderful article begins with history and some elementary facts and proceeds to greater and greater depth about the existence of solutions to Pell equations and then later the algorithmic issues of finding those solutions. The cattle problem is discussed, as are modern smooth number methods for solving Pell equations and the algorithmic issues of representing very large solutions in a reasonable way."

The simplest solutions to Pell's equation can be huge, even when d is quite small. Read Lenstra's paper for some examples from over two thousand years ago. Here is one example for $d = 10000019$.

```
K.<a> = QuadraticField(next_prime(10^7))
G = K.unit_group(); G.1.value()
```

```
16358025988034632822559223812109462549914267769314291550674725300\
340064100365767872890438816249271266423998175030309436575610631639\
272377601680603795883791477817611974184075445702823789975945910042\
8895693238165048098039*a - \
517286692885814967470170672368346798303629034373575202975075605058\
714958080893991274427903448098643836512878351227856269086856679078\
304979321047765031073345259902622712059164969008633603603640331175\
66345622041829362222240930
```

Exercise 8.2.2. Let U be the group of units of the ring of integers of $K = \mathbb{Q}(\sqrt{5})$.

(a) Prove that the set S of units $x + y\sqrt{5} \in U$ with $x, y \in \mathbb{Z}$ is a subgroup of U. (The main point is to show that the inverse of a unit with $x, y \in \mathbb{Z}$ again has coefficients in $\mathbb{Z}$.)

(b) Let U^3 denote the subgroup of cubes of elements of U. Prove that $S = U^3$ by showing that $U^3 \subset S \subsetneq U$ and that there are no groups H with $U^3 \subsetneq H \subsetneq U$.

8.2.2 Examples with Various Signatures

In this section we give examples for various (r, s) pairs. First we consider $K = \mathbb{Q}(i)$.

```
K.<a> = QuadraticField(-1)
K.signature()
```

```
(0, 1)
```

```
U = K.unit_group(); U
```

```
Unit group with structure C4 of Number Field in a with
defining polynomial x^2 + 1
```

```
U.0.value()
```

```
a
```

The `signature` method returns the number of real and complex conjugate embeddings of K into $\mathbb{C}$. The `unit_group` method, which we used above, returns the unit group U_K as an abstract abelian group and a homomorphism $U_K \to \mathcal{O}_K$.

Next we consider $K = \mathbb{Q}(\sqrt[3]{2})$.

```
K.<a> = NumberField(x^3 - 2)
K.signature()
```

```
(1, 1)
```

```
U = K.unit_group(); U
```

```
Unit group with structure C2 x Z of Number Field in a with
defining polynomial x^3 - 2
```

```
[u.value() for u in U.gens()]
```

```
[-1, a - 1]
```

```
u = U.1.value(); u
```

```
a - 1
```

Below we use the `places` command, which returns the real embeddings and representatives for the complex conjugate embeddings. We use the places to define the log map φ, which plays such a big role in this chapter.

```
S = K.places(prec=53); S
```

```
[Ring morphism:
  From: Number Field in a with defining polynomial x^3 - 2
  To:   Real Double Field
  Defn: a |--> 1.25992104989, \
Ring morphism:
  From: Number Field in a with defining polynomial x^3 - 2
  To:   Complex Double Field
  Defn: a |--> -0.629960524947 + 1.09112363597*I]
```

```
def phi(z):
    return [log(abs(sigma(z))) for sigma in S]
phi(u)
```

```
[-1.3473773483293832, 0.673688674164692]
```

```
phi(K(-1))
```

```
[0.0, 0.0]
```

Note that $\varphi : U_K \to \mathbb{R}^2$, and the image lands in the 1-dimensional subspace of (x_1, x_2) such that $x_1 + 2x_2 = 0$. Also, note that $\varphi(-1) = (0, 0)$.

Let's try a field such that $r + s - 1 = 2$. First, one with $r = 0$ and $s = 3$:

```
K.<a> = NumberField(x^6 + x + 1)
K.signature()
```

```
(0, 3)
```

```
U = K.unit_group(); U
```

```
Unit group with structure C2 x Z x Z of Number Field in a with
defining polynomial x^6 + x + 1
```

```
u1 = U.1.value(); u1
```

```
a
```

```
u2 = U.2.value(); u2
```

```
a^3 + a
```

```
S = K.places(prec=53)
def phi(z):
    return [log(abs(sigma(z))) for sigma in S]
phi(u1)
```

```
[-0.16741548328589614, 0.04864390975267338, 0.11877157353322298]
```

```
phi(u2)
```

```
[0.30678570892329504, -1.0725146505489758, 0.7657289416256803]
```

```
phi(K(-1))
```

```
[0.0, 0.0, 0.0]
```

```
sum(phi(u1))
```

```
2.220446049250313e-16
```

```
sum(phi(u2))
```

```
-4.440892098500626e-16
```

Notice that the log image of u_1 is clearly not a real multiple of the log image of u_2 (e.g., the scalar would have to be positive because of the first coefficient, but negative because of the second). This illustrates the fact that the log images of u_1 and u_2 span a two-dimensional space.

Next we compute a field with $r = 3$ and $s = 0$. (A field with $s = 0$ is called totally real.)

```
K.<a> = NumberField(x^3 + x^2 - 5*x - 1)
K.signature()
```

```
(3, 0)
```

```
U = K.unit_group(); U
```

```
Unit group with structure C2 x Z x Z of Number Field in a with
defining polynomial x^3 + x^2 - 5*x - 1
```

```
u1 = U.1.value(); u1
```

```
1/2*a^2 + a - 1/2
```

```
u2 = U.2.value(); u2
```

```
a
```

```
S = K.places(prec=53)
def phi(z):
    return [log(abs(sigma(z))) for sigma in S]
phi(u1)
```

```
[-0.7747670223461895, -0.3928487245813982, 1.1676157469275887]
```

```
phi(u2)
```

```
[0.9966812040934553, -1.6402241503223172, 0.6435429462288627]
```

A field with $r = 0$ is called totally complex. For example, the *cyclotomic fields* $\mathbb{Q}(\zeta_n)$ are totally complex, where ζ_n is a primitive nth root of unity. The degree of $\mathbb{Q}(\zeta_n)$ over $\mathbb{Q}$ is $\varphi(n)$ and $r = 0$, so $s = \varphi(n)/2$ (assuming $n > 2$). Here φ is the Euler Totient function which on n is defined as the number of integers k such that $0 < k \leq n$ and $\gcd(k, n) = 1$.

```
K.<a> = CyclotomicField(11); K
```

```
Cyclotomic Field of order 11 and degree 10
```

```
K.signature()
```

```
(0, 5)
```

```
U = K.unit_group(); U
```

```
Unit group with structure C22 x Z x Z x Z x Z of \
Cyclotomic Field of order 11 and degree 10
```

```
u = U.1.value(); u
```

```
a^7 + a^6
```

```
S = K.places(prec=20)
def phi(z):
    return [log(abs(sigma(z))) for sigma in S]
phi(u)
```

```
[-1.2566, -0.18533, 0.26982, 0.52028, 0.65180]
```

```
for u in U.gens():
    print phi(u.value())
```

```
[0.00000, 0.00000, 0.00000, -9.5367e-7, 0.00000]
[-1.2566, -0.18533, 0.26982, 0.52028, 0.65180]
[-0.26981, -0.52028, 0.18533, -0.65180, 1.2566]
[0.65180, 0.26981, -1.2566, -0.18533, 0.52029]
[-0.084486, -1.1721, -0.33496, 0.60477, 0.98675]
```

How far can we go computing unit groups of cyclotomic fields directly with Sage?

```
%time U = CyclotomicField(11).unit_group()
```

```
CPU time: 0.01 s, Wall time: 0.01 s
```

```
%time U = CyclotomicField(13).unit_group()
```

```
CPU time: 0.30 s, Wall time: 0.30 s
```

```
%time U = CyclotomicField(17).unit_group()
```

```
CPU time: 1.13 s, Wall time: 1.31 s
```

```
%time U = CyclotomicField(23).unit_group()
```

```
.... I waited a few minutes and gave up....
```

However, if you are willing to assume some conjectures (something related to the Generalized Riemann Hypothesis), you can go further:

```
proof.number_field(False)
%time U = CyclotomicField(11).unit_group()
```

```
CPU time: 0.07 s, Wall time: 0.07 s
```

```
%time U = CyclotomicField(13).unit_group()
```

```
CPU time: 0.03 s, Wall time: 0.03 s
```

```
%time U = CyclotomicField(17).unit_group()
```

```
CPU time: 0.06 s, Wall time: 0.06 s
```

```
%time U = CyclotomicField(23).unit_group()
```

```
CPU time: 0.26 s, Wall time: 0.31 s
```

```
%time U = CyclotomicField(29).unit_group()
```

```
CPU time: 0.60 s, Wall time: 0.62 s
```

The generators of the units for $\mathbb{Q}(\zeta_{29})$ are

$$\begin{aligned}
u_0 &= -\zeta_{29}^3\\
u_1 &= \zeta_{29}^{26} + \zeta_{29}^{25} + \zeta_{29}^{22} + \zeta_{29}^{21} + \zeta_{29}^{19} + \zeta_{29}^{18} + \zeta_{29}^{15} + \zeta_{29}^{14} + \zeta_{29}^{11} + \zeta_{29}^{8} + \zeta_{29}^{7} + \zeta_{29}^{4} + \zeta_{29}^{3} + \zeta_{29} + 1\\
u_2 &= \zeta_{29}^{14} + \zeta_{29}^{3}\\
u_3 &= \zeta_{29}^{3} + 1\\
u_4 &= \zeta_{29}^{26} + \zeta_{29}^{20} + \zeta_{29}^{3}\\
u_5 &= \zeta_{29}^{22} + \zeta_{29}^{11} + \zeta_{29}^{2}\\
u_6 &= \zeta_{29}^{10} + \zeta_{29}^{9} + \zeta_{29}^{8}\\
u_7 &= \zeta_{29}^{23} + \zeta_{29}\\
u_8 &= \zeta_{29}^{17} + \zeta_{29}^{11}\\
u_9 &= \zeta_{29}^{22} + \zeta_{29}^{3}\\
u_{10} &= \zeta_{29}^{24} + \zeta_{29}^{19} + \zeta_{29}^{5} + 1\\
u_{11} &= \zeta_{29}^{19} + \zeta_{29}^{6}\\
u_{12} &= \zeta_{29}^{27} + \zeta_{29}^{19} + \zeta_{29}^{11} + \zeta_{29}^{6} + \zeta_{29}^{3}\\
u_{13} &= \zeta_{29}^{26} + \zeta_{29}^{15} + \zeta_{29}^{4}.
\end{aligned}$$

There are better ways to compute units in cyclotomic fields than to just use general purpose software. For example, there are explicit *cyclotomic units* that can be written down and generate a finite subgroup of U_K. See [Was97, Ch. 8], which would be a great book to read now that you've gone this far in the present book. Also, using ideas explained in that book, it is probably possible to make the `unit_group` command in Sage for cyclotomic fields extremely fast, which would be an interesting project for a reader who also likes to code.

Chapter 9

Decomposition and Inertia Groups

In this chapter we will study extra structure in the case when K is Galois over $\mathbb{Q}$. We will learn about Frobenius elements, the Artin symbol, decomposition groups, and how the Galois group of K is related to Galois groups of residue class fields. These are the basic structures needed to attach L-function to representations of $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$, which will play a central role in the next few chapters.

9.1 Galois Extensions

In this section we give a survey (no proofs) of the basic facts about Galois extensions of $\mathbb{Q}$ that will be needed in the rest of this chapter.

Definition 9.1.1 (Galois)**.** An extension K/L of number fields is *Galois* if

$$\#\,\mathrm{Aut}(K/L) = [K : L],$$

where $\mathrm{Aut}(K/L)$ is the group of automorphisms of K that fix L. We write

$$\mathrm{Gal}(K/L) = \mathrm{Aut}(K/L).$$

For example, if $K \subset \mathbb{C}$ is a number field embedded in the complex numbers, then K is *Galois* over $\mathbb{Q}$ if every field homomorphism $K \to \mathbb{C}$ has image K. As another example, any quadratic extension K/L is Galois over L, since it is of the form $L(\sqrt{a})$, for some $a \in L$, and the nontrivial automorphism is induced by $\sqrt{a} \mapsto -\sqrt{a}$, so there is always one nontrivial automorphism. If $f \in L[x]$ is an irreducible cubic polynomial, and a is a root of f, then one proves in a course on Galois theory that $L(a)$ is Galois over L if and only if the discriminant of f is a perfect square in L. "Random" number fields of degree bigger than 2 are rarely Galois.

If $K \subset \mathbb{C}$ is a number field, then the *Galois closure* K^{gc} of K in $\mathbb{C}$ is the field generated by all images of K under all embeddings in $\mathbb{C}$ (more generally, if K/L

is an extension, the Galois closure of K over L is the field generated by images of embeddings $K \to \mathbb{C}$ that are the identity map on L). If $K = \mathbb{Q}(a)$, then K^{gc} is the field generated by all of the conjugates of a, and is hence Galois over $\mathbb{Q}$, since the image under an embedding of any polynomial in the conjugates of a is again a polynomial in conjugates of a.

How much bigger can the degree of K^{gc} be as compared to the degree of $K = \mathbb{Q}(a)$? There is an embedding of $\mathrm{Gal}(K^{\mathrm{gc}}/\mathbb{Q})$ into the group of permutations of the conjugates of a. If a has n conjugates, then this is an embedding $\mathrm{Gal}(K^{\mathrm{gc}}/\mathbb{Q}) \hookrightarrow S_n$, where S_n is the symmetric group on n symbols, which has order $n!$. Thus the degree of the K^{gc} over $\mathbb{Q}$ is a divisor of $n!$. Also $\mathrm{Gal}(K^{\mathrm{gc}}/\mathbb{Q})$ is a transitive subgroup of S_n, which constrains the possibilities further. When $n = 2$, we recover the fact that quadratic extensions are Galois. When $n = 3$, we see that the Galois closure of a cubic extension is either the cubic extension or a quadratic extension of the cubic extension. One can show that the Galois closure of a cubic extension is obtained by adjoining the square root of the discriminant, which is why an irreducible cubic defines a Galois extension if and only if the discriminant is a perfect square.

For an extension K of $\mathbb{Q}$ of degree 5, it is "frequently" the case that the Galois closure has degree 120, and in fact it is an interesting problem to enumerate examples of degree 5 extension in which the Galois closure has degree smaller than 120. For example, the only possibilities for the order of a transitive proper subgroup of S_5 are 5, 10, 20, and 60; there are also proper subgroups of S_5 order $2, 3, 4, 6, 8, 12$, and 24, but none are transitive.

Let n be a positive integer. Consider the field $K = \mathbb{Q}(\zeta_n)$, where $\zeta_n = e^{2\pi i/n}$ is a primitive nth root of unity. If $\sigma : K \to \mathbb{C}$ is an embedding, then $\sigma(\zeta_n)$ is also an nth root of unity, and the group of nth roots of unity is cyclic, so $\sigma(\zeta_n) = \zeta_n^m$ for some m which is invertible modulo n. Thus K is Galois and $\mathrm{Gal}(K/\mathbb{Q}) \hookrightarrow (\mathbb{Z}/n\mathbb{Z})^*$. However, $[K : \mathbb{Q}] = \varphi(n)$, so this map is an isomorphism. (Remark: Taking a limit using the maps $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}) \to \mathrm{Gal}(\mathbb{Q}(\zeta_{p^r})/\mathbb{Q})$, we obtain a homomorphism $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}) \to \mathbb{Z}_p^*$, which is called the *p-adic cyclotomic character*.)

Compositums of Galois extensions are Galois. For example, the biquadratic field $K = \mathbb{Q}(\sqrt{5}, \sqrt{-1})$ is a Galois extension of $\mathbb{Q}$ of degree 4, which is the compositum of the Galois extensions $\mathbb{Q}(\sqrt{5})$ and $\mathbb{Q}(\sqrt{-1})$ of $\mathbb{Q}$.

Fix a number field K that is Galois over a subfield L. Then the Galois group $G = \mathrm{Gal}(K/L)$ acts on many of the object that we have associated to K.

Exercise 9.1.2. Describe the natural action of G on the following objects:

- The ring of integers $\mathcal{O}_K$
- The group units U_K
- The set of ideals of $\mathcal{O}_K$
- The group of fractional ideals of $\mathcal{O}_K$
- The class group $\mathrm{Cl}(K)$

- The set $S_{\mathfrak{p}}$ of prime ideals lying over a given nonzero prime ideal $\mathfrak{p}$ of $\mathcal{O}_L$, i.e., the prime divisors of $\mathfrak{p}\mathcal{O}_K$

In the next section we will be concerned with the action of $\mathrm{Gal}(K/L)$ on $S_{\mathfrak{p}}$, though actions on each of the other objects, especially $\mathrm{Cl}(K)$, are also of great interest. Understanding the action of $\mathrm{Gal}(K/L)$ on $S_{\mathfrak{p}}$ will enable us to associate, in a natural way, a holomorphic L-function to any complex representation $\mathrm{Gal}(K/L) \to \mathrm{GL}_n(\mathbb{C})$.

9.2 Decomposition of Primes: $efg = n$

If $I \subset \mathcal{O}_K$ is any ideal in the ring of integers of a Galois extension K of $\mathbb{Q}$ and $\sigma \in \mathrm{Gal}(K/\mathbb{Q})$, then

$$\sigma(I) = \{\sigma(x) : x \in I\}$$

is also an ideal of $\mathcal{O}_K$.

Fix a prime $\mathfrak{p} \subset \mathcal{O}_K$ and write $\mathfrak{p}\mathcal{O}_K = \mathfrak{P}_1^{e_1} \cdots \mathfrak{P}_g^{e_g}$, so $S_{\mathfrak{p}} = \{\mathfrak{P}_1, \ldots, \mathfrak{P}_g\}$.

Definition 9.2.1 (Residue class degree)**.** Suppose $\mathfrak{P}$ is a prime of $\mathcal{O}_K$ lying over $\mathfrak{p}$. Then the *residue class degree* of $\mathfrak{P}$ is

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

i.e., the degree of the extension of residue class fields.

If $M/K/L$ is a tower of field extensions and $\mathfrak{q}$ is a prime of M over $\mathfrak{P}$, then

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

so the residue class degree is multiplicative in towers.

Note that if $\sigma \in \mathrm{Gal}(K/L)$ and $\mathfrak{P} \in S_p$, then σ induces an isomorphism of finite fields $\mathcal{O}_K/\mathfrak{P} \to \mathcal{O}_K/\sigma(\mathfrak{P})$ that fixes the common subfield $\mathcal{O}_L/\mathfrak{p}$. Thus the residue class degrees of $\mathfrak{P}$ and $\sigma(\mathfrak{P})$ are the same. In fact, much more is true.

Theorem 9.2.2. *Suppose K/L is a Galois extension of number fields, and let $\mathfrak{p}$ be a prime of $\mathcal{O}_L$. Write $\mathfrak{p}\mathcal{O}_K = \prod_{i=1}^{g} \mathfrak{P}_i^{e_i}$, and let $f_i = f_{\mathfrak{P}_i/\mathfrak{p}}$. Then $G = \mathrm{Gal}(K/L)$ acts transitively on the set $S_{\mathfrak{p}}$ of primes $\mathfrak{P}_i$, and*

$$e_1 = \cdots = e_g, \qquad f_1 = \cdots = f_g.$$

Morever, if we let e be the common value of the e_i, f the common value of the f_i, and $n = [K : L]$, then

$$efg = n.$$

Proof. For simplicity, we will give the proof only in the case $L = \mathbb{Q}$, but the proof works in general. Suppose $p \in \mathbb{Z}$ and $p\mathcal{O}_K = \mathfrak{p}_1^{e_1} \cdots \mathfrak{p}_g^{e_g}$, and $S = \{\mathfrak{p}_1, \ldots, \mathfrak{p}_g\}$. We will first prove that G acts transitively on S. Let $\mathfrak{p} = \mathfrak{p}_i$ for some i. Recall

Lemma 5.2.2 which we proved long ago using the Chinese Remainder Theorem (Theorem 5.1.4). It showed there exists $a \in \mathfrak{p}$ such that $(a)/\mathfrak{p}$ is an integral ideal that is coprime to $p\mathcal{O}_K$. The product

$$I = \prod_{\sigma \in G} \sigma((a)/\mathfrak{p}) = \prod_{\sigma \in G} \frac{(\sigma(a))\mathcal{O}_K}{\sigma(\mathfrak{p})} = \frac{(\mathrm{Norm}_{K/\mathbb{Q}}(a))\mathcal{O}_K}{\prod_{\sigma \in G} \sigma(\mathfrak{p})} \tag{9.2.1}$$

is a nonzero integral $\mathcal{O}_K$ ideal since it is a product of nonzero integral $\mathcal{O}_K$ ideals. Since $a \in \mathfrak{p}$ we have that $\mathrm{Norm}_{K/\mathbb{Q}}(a) \in \mathfrak{p} \cap \mathbb{Z} = p\mathbb{Z}$. Thus the numerator of the rightmost expression in (9.2.1) is divisible by $p\mathcal{O}_K$. Also, because $(a)/\mathfrak{p}$ is coprime to $p\mathcal{O}_K$, each $\sigma((a)/\mathfrak{p})$ is coprime to $p\mathcal{O}_K$ as well. Thus I is coprime to $p\mathcal{O}_K$. This means the denominator of the rightmost expression in (9.2.1) must also be divisible by $p\mathcal{O}_K$ in order to cancel the $p\mathcal{O}_K$ in the numerator. Thus we have shown that for any i,

$$\prod_{j=1}^{g} \mathfrak{p}_j^{e_j} = p\mathcal{O}_K \;\Big|\; \prod_{\sigma \in G} \sigma(\mathfrak{p}_i).$$

By unique factorization, since every $\mathfrak{p}_j$ appears in the left hand side, we must have that for each j there is a σ with $\sigma(\mathfrak{p}_i) = \mathfrak{p}_j$, i.e., G acts transitively on S.

Choose some j and suppose that $k \neq j$ is another index. Because G acts transitively, there exists $\sigma \in G$ such that $\sigma(\mathfrak{p}_k) = \mathfrak{p}_j$. Applying σ to the factorization $p\mathcal{O}_K = \prod_{i=1}^{g} \mathfrak{p}_i^{e_i}$, we see that

$$\prod_{i=1}^{g} \mathfrak{p}_i^{e_i} = \prod_{i=1}^{g} \sigma(\mathfrak{p}_i)^{e_i}.$$

Using unique factorization, we get $e_j = e_k$. Thus $e_1 = e_2 = \cdots = e_g$.

As was mentioned right before the statement of the theorem, for any $\sigma \in G$ we have $\mathcal{O}_K/\mathfrak{p}_i \cong \mathcal{O}_K/\sigma(\mathfrak{p}_i)$. Since G acts transitively it follows that $f_1 = f_2 = \cdots = f_g$. We have, upon applying the Chinese Remainder Theorem and noting $\#(\mathcal{O}_K/(\mathfrak{p}^m)) = \#(\mathcal{O}_K/\mathfrak{p})^m$ (see Exercise 5.2.5), that

$$\begin{aligned}[K:\mathbb{Q}] &= \dim_{\mathbb{Z}} \mathcal{O}_K = \dim_{\mathbb{F}_p} \mathcal{O}_K/p\mathcal{O}_K \\ &= \dim_{\mathbb{F}_p} \left(\bigoplus_{i=1}^{g} \mathcal{O}_K/\mathfrak{p}_i^{e_i} \right) = \sum_{i=1}^{g} e_i f_i = efg,\end{aligned}$$

which completes the proof. □

The rest of this section illustrates the theorem for quadratic fields and a cubic field and its Galois closure.

9.2.1 Special Cases

Quadratic Extensions

Suppose $K/\mathbb{Q}$ is a quadratic field. Then K is Galois, so for each prime $p \in \mathbb{Z}$ we have $2 = efg$. There are exactly three possibilities:

Ramified: $e = 2$, $f = g = 1$: The prime p *ramifies* in $\mathcal{O}_K$, which means $p\mathcal{O}_K = \mathfrak{p}^2$. Let α be a generator for $\mathcal{O}_K$ and $h \in \mathbb{Z}[x]$ a minimal polynomial for α. By Theorem 4.2.3 a prime p is ramified in $\mathcal{O}_K$ if and only if h has a double root modulo p, which is equivalent to p dividing the discriminant of h. This shows there are only finitely many ramified primes. More generally, the ramified primes are exactly the ones that divide the discriminant (see [Mar77, Thm. 24] or [**?**, Cor. III.2.12]).

Inert: $e = 1$, $f = 2$, $g = 1$: The prime p is *inert* in $\mathcal{O}_K$, which means $p\mathcal{O}_K = \mathfrak{p}$ is prime. It is a nontrivial theorem that this happens half of the time, as we will see illustrated below for a particular example.

Split: $e = f = 1$, $g = 2$: The prime p *splits* in $\mathcal{O}_K$, which means $p\mathcal{O}_K = \mathfrak{p}_1\mathfrak{p}_2$ with $\mathfrak{p}_1 \neq \mathfrak{p}_2$. This happens the other half of the time.

Example 9.2.3. Let $K = \mathbb{Q}(\sqrt{5})$, so $\mathcal{O}_K = \mathbb{Z}[\gamma]$, where $\gamma = (1+\sqrt{5})/2$. Then $p = 5$ is ramified, since $5\mathcal{O}_K = (\sqrt{5})^2$. More generally, the order $\mathbb{Z}[\sqrt{5}]$ has index 2 in $\mathcal{O}_K$, so for any prime $p \neq 2$ we can determine the factorization of p in $\mathcal{O}_K$ by finding the factorization of the polynomial $x^2 - 5 \in \mathbb{F}_p[x]$. The polynomial $x^2 - 5$ splits as a product of two distinct factors in $\mathbb{F}_p[x]$ if and only if $e = f = 1$ and $g = 2$. For $p \neq 2, 5$ this is the case if and only if 5 is a square in $\mathbb{F}_p$, i.e., if $\left(\frac{5}{p}\right) = 1$, where $\left(\frac{5}{p}\right)$ is $+1$ if 5 is a square mod p and -1 if 5 is not. By quadratic reciprocity,

$$\left(\frac{5}{p}\right) = (-1)^{\frac{5-1}{2}\cdot\frac{p-1}{2}} \cdot \left(\frac{p}{5}\right) = \left(\frac{p}{5}\right) = \begin{cases} +1 & \text{if } p \equiv \pm 1 \pmod 5 \\ -1 & \text{if } p \equiv \pm 2 \pmod 5. \end{cases}$$

Thus whether p splits or is inert in $\mathcal{O}_K$ is determined by the residue class of p modulo 5. It is a theorem of Dirichlet, which was massively generalized by Chebotarev, that $p \equiv \pm 1$ half the time and $p \equiv \pm 2$ the other half the time.[1]

The Cube Root of Two

Suppose $K/\mathbb{Q}$ is not Galois. Then e_i, f_i, and g are defined for each prime $p \in \mathbb{Z}$, but we need not have $e_1 = \cdots = e_g$ or $f_1 = \cdots = f_g$. We do still have that $\sum_{i=1}^{g} e_i f_i = n$, by the Chinese Remainder Theorem. For a proof of this identity, see [Mar77, Thm. 21], or, for a slightly more general version, [**?**, Prop. I.8.2]

Consider the case where $K = \mathbb{Q}(\sqrt[3]{2})$. We know that $\mathcal{O}_K = \mathbb{Z}[\sqrt[3]{2}]$. Thus $2\mathcal{O}_K = (\sqrt[3]{2})^3$, so for 2 we have $e = 3$ and $f = g = 1$.

[1] For a technical statement and proof of this theorem, see [**?**] Theorem VII.13.4.

Working modulo 5 we have

$$x^3 - 2 = (x+2)(x^2+3x+4) \in \mathbb{F}_5[x],$$

and the quadratic factor is irreducible. Thus

$$5\mathcal{O}_K = (5, \sqrt[3]{2}+2) \cdot (5, \sqrt[3]{2}^2 + 3\sqrt[3]{2} + 4).$$

Thus here $g = 2$, $e_1 = e_2 = 1$, $f_1 = 1$, and $f_2 = 2$. Thus when K is not Galois we need not have that the f_i are all equal.

9.2.2 Definitions and Terminology

In the previous sections we used words like "ramify", "inert", and "split" to describe the decomposition of a prime in an extension. This section will define the generalizations of these concepts which will be used in later sections.

Let K/L be an extension of number fields. Let $\mathcal{O}_K, \mathcal{O}_L$ denote the respective ring of integers and $\mathfrak{q}$ a prime in $\mathcal{O}_L$. By Theorem 3.2.6 we know that the ideal $\mathfrak{q}\mathcal{O}_K$ factors uniquely into a product of primes $\mathfrak{p}_i$ in $\mathcal{O}_K$ given by

$$\mathfrak{q}\mathcal{O}_K = \mathfrak{p}_1^{e_1} \cdots \mathfrak{p}_g^{e_g}.$$

Let f_i be the degree of the extension of residue fields, i.e.,

$$f_i = [\mathcal{O}_K/\mathfrak{p}_i : \mathcal{O}_L/\mathfrak{q}].$$

Definition 9.2.4. The prime $\mathfrak{q}$ *ramifies* in L if $e_i > 1$ for some $1 \leq i \leq g$. Otherwise $\mathfrak{q}$ is *unramified*. If $\mathfrak{q}$ is ramified and moreover $f_i = 1$ for all i, then $\mathfrak{q}$ is *totally ramified*.

Definition 9.2.5. The prime $\mathfrak{p}$ is *inert* in L if $\mathfrak{p}\mathcal{O}_L$ is prime. In this case we have $g = 1$, $\mathfrak{q}_1 = \mathfrak{p}\mathcal{O}_L$, and $e_1 = 1$.

Definition 9.2.6. The prime $\mathfrak{p}$ is *split* in L if $g > 1$. If moreover $g = [L : K]$, then $\mathfrak{p}$ *splits completely* or is *totally split*.

It will sometimes be helpful to emphasize which prime we are referring to. To do this we will use the notation $e(\mathfrak{p}/\mathfrak{q})$ to represent the power of $\mathfrak{p}$ appearing in the factorization of $\mathfrak{q}\mathcal{O}_K$. The number $e(\mathfrak{p}/\mathfrak{q})$ is called the *ramification index* of $\mathfrak{p}$ over $\mathfrak{q}$. In this notation we could write $\mathfrak{q}\mathcal{O}_K = \prod \mathfrak{p}^{e(\mathfrak{p}/\mathfrak{q})}$ where the product ranges over all primes $\mathfrak{p}$ in $\mathcal{O}_K$. We will similarly denote $f(\mathfrak{p}/\mathfrak{q})$ to be the degree of the extension of residue fields $[\mathcal{O}_K/\mathfrak{p} : \mathcal{O}_L/\mathfrak{q}]$. The number $f(\mathfrak{p}/\mathfrak{q})$ is called the *inertia degree* of $\mathfrak{p}/\mathfrak{q}$. Because the number of primes over $\mathfrak{q}$ depends on the field K, we sometimes denote g by $g_K(\mathfrak{q})$.

Exercise 9.2.7. The following are some basic properties of decompositions. For each one, compare the result with previous examples we have seen such as Example 9.2.3.

Let $K/L/\mathbb{Q}$ be a tower of number fields. Let p be a prime in $\mathbb{Z}$, $\mathfrak{q}$ a prime in $\mathcal{O}_L$ lying over p, and $\mathfrak{p}$ a prime in $\mathcal{O}_K$ lying over $\mathfrak{q}$.

(a) Show that e is multiplicative, that is $e(\mathfrak{p}/p) = e(\mathfrak{p}/\mathfrak{q}) \cdot e(\mathfrak{q}/p)$.

(b) Show that f is multiplicative, that is $f(\mathfrak{p}/p) = f(\mathfrak{p}/\mathfrak{q}) \cdot f(\mathfrak{q}/p)$.

(c) Let $g_L(p)$ be the number of primes of $\mathcal{O}_L$ lying over p. Show that $g_K(p) = \sum_{\mathfrak{q} \text{ lies over } p} g_L(\mathfrak{q})$.

Exercise 9.2.8 (See [Mar77, Ch. 4, Exercise 24])**.** Continue the notation from the previous exercise.

(a) If p it totally ramified in K then it is totally ramified in L.

(b) Let K' be another extension of L. If $\mathfrak{p}$ is totally ramified in K and unramified in K' then $K \cap K' = L$.

9.3 The Decomposition Group

Suppose K is a number field that is Galois over $\mathbb{Q}$ with group $G = \mathrm{Gal}(K/\mathbb{Q})$. Fix a prime $\mathfrak{p} \subset \mathcal{O}_K$ lying over $p \in \mathbb{Z}$.

Definition 9.3.1 (Decomposition group)**.** The *decomposition group* of $\mathfrak{p}$ is the subgroup

$$D_{\mathfrak{p}} = \{\sigma \in G : \sigma(\mathfrak{p}) = \mathfrak{p}\} \subset G.$$

Note that $D_{\mathfrak{p}}$ is the stabilizer of $\mathfrak{p}$ for the action of G on the set of primes lying over p.

It also makes sense to define decomposition groups for relative extensions K/L, but for simplicity and to fix ideas in this section we only define decomposition groups for a Galois extension $K/\mathbb{Q}$.

Let $k_{\mathfrak{p}} = \mathcal{O}_K/\mathfrak{p}$ denote the residue class field of $\mathfrak{p}$. In this section we will prove that there is an exact sequence

$$1 \to I_{\mathfrak{p}} \to D_{\mathfrak{p}} \to \mathrm{Gal}(k_{\mathfrak{p}}/\mathbb{F}_p) \to 1,$$

where $I_{\mathfrak{p}}$ is the *inertia subgroup* of $D_{\mathfrak{p}}$, and $\#I_{\mathfrak{p}} = e = e(\mathfrak{p}/p)$. The most interesting part of the proof is showing that the natural map $D_{\mathfrak{p}} \to \mathrm{Gal}(k_{\mathfrak{p}}/\mathbb{F}_p)$ is surjective. We will also discuss the structure of $D_{\mathfrak{p}}$ and introduce Frobenius elements, which play a crucial role in understanding Galois representations.

Recall from Theorem 9.2.2 that G acts transitively on the set of primes $\mathfrak{p}$ lying over p. The orbit-stabilizer theorem implies that $[G : D_{\mathfrak{p}}]$ equals the cardinality of the orbit of $\mathfrak{p}$, which by Theorem 9.2.2 equals the number g of primes lying over p, so $[G : D_{\mathfrak{p}}] = g$.

Lemma 9.3.2. *The decomposition subgroups $D_{\mathfrak{p}}$ corresponding to primes $\mathfrak{p}$ lying over a given p are all conjugate as subgroups of G.*

Proof. See Exercise 9.3.3. □

Exercise 9.3.3. Prove Lemma 9.3.2.

[*Hint*: For $\sigma, \tau \in G$ you need to show $\tau D_{\mathfrak{p}} \tau^{-1} = D_{\tau\mathfrak{p}}$. Start by writing down what it means for $\sigma \in D_{\mathfrak{p}}$ and $\tau\sigma\tau^{-1} \in D_{\tau\mathfrak{p}}$.]

The decomposition group is useful because it allows us to refine the extension $K/\mathbb{Q}$ into a tower of extensions, such that at each step in the tower we understand the splitting behavior of the primes lying over p.

Recall the correspondence between subgroups of the Galois group G and subfields of K. The fixed fields corresponding to the decomposition and inertia subgroups have an important description in terms of the splitting behavior of the prime $\mathfrak{p}$. We characterize the fixed field of $D = D_{\mathfrak{p}}$ as follows.

Proposition 9.3.4. *The fixed field*

$$K^D = \{a \in K : \sigma(a) = a \text{ for all } \sigma \in D\}$$

of D is the smallest subfield $L \subset K$ such that the prime ideal $\mathfrak{q} = \mathfrak{p} \cap \mathcal{O}_L$ has $g_K(\mathfrak{q}) = 1$, i.e., there is a unique prime of $\mathcal{O}_K$ lying over $\mathfrak{q}$.

Proof. First suppose $L = K^D$, and note that by Galois theory $\mathrm{Gal}(K/L) \cong D$, and by Theorem 9.2.2, the group D acts transitively on the primes of K lying over $\mathfrak{q}$. One of these primes is $\mathfrak{p}$, and D fixes $\mathfrak{p}$ by definition, so there is only one prime of K lying over $\mathfrak{q}$, that is $g = 1$. Conversely, if $L \subset K$ is such that $\mathfrak{q}$ has $g = 1$, then $\mathrm{Gal}(K/L)$ fixes $\mathfrak{p}$ (since it is the only prime over $\mathfrak{q}$), so $\mathrm{Gal}(K/L) \subset D$, hence $K^D \subset L$. $\square$

Thus p does not split in going from K^D to K—it does some combination of ramifying and staying inert. To fill in more of the picture, the following proposition asserts that p splits completely and does not ramify in $K^D/\mathbb{Q}$.

Proposition 9.3.5. *Fix a finite Galois extension K of $\mathbb{Q}$, let $\mathfrak{p}$ be a prime lying over p with decomposition group D, and set $L = K^D$ and $\mathfrak{q} = \mathfrak{p} \cap \mathcal{O}_L$. Then $e(\mathfrak{q}/p) = f(\mathfrak{q}/p) = 1$, $g_L(p) = [L : \mathbb{Q}]$, $e(\mathfrak{p}/p) = e(\mathfrak{p}/\mathfrak{q})$ and $f(\mathfrak{p}/p) = f(\mathfrak{p}/\mathfrak{q})$.*

Proof. As mentioned right after Definition 9.3.1, the orbit-stabilizer theorem implies that $g_K(p) = [G : D]$, and by Galois theory $[G : D] = [L : \mathbb{Q}]$, so $g_K(p) = [L : \mathbb{Q}]$. By Proposition 9.3.4, we have $g_K(\mathfrak{q}) = 1$ so by Theorem 9.2.2,

$$\begin{aligned} e(\mathfrak{p}/\mathfrak{q}) \cdot f(\mathfrak{p}/\mathfrak{q}) = [K : L] &= \frac{[K : \mathbb{Q}]}{[L : \mathbb{Q}]} \\ &= \frac{e(\mathfrak{p}/p) \cdot f(\mathfrak{p}/p) \cdot g_K(p)}{[L : \mathbb{Q}]} \\ &= e(\mathfrak{p}/p) \cdot f(\mathfrak{p}/p). \end{aligned}$$

Now $e(\mathfrak{p}/\mathfrak{q}) \leq e(\mathfrak{p}/p)$ and $f(\mathfrak{p}/\mathfrak{q}) \leq f(\mathfrak{p}/p)$, so we must have $e(\mathfrak{p}/\mathfrak{q}) = e(\mathfrak{p}/p)$ and $f(\mathfrak{p}/\mathfrak{q}) = f(\mathfrak{p}/p)$. Since from Exercise 9.2.7 we have $e(\mathfrak{p}/p) = e(\mathfrak{p}/\mathfrak{q}) \cdot e(\mathfrak{q}/p)$ and $f(\mathfrak{p}/q) = f(\mathfrak{p}/\mathfrak{q}) \cdot f(\mathfrak{q}/p)$, it follows that $e(\mathfrak{q}/p) = f(\mathfrak{q}/p) = 1$. $\square$

We summarize the results of the decomposition of a prime in the tower $K \supseteq L = K^D \supseteq \mathbb{Q}$ in Table 9.3.1. This table shows the ramification indices, inertia degrees, and the number of primes at each step of the tower.

Ramification (e)	Inertia (f)	Splitting (g)	Primes	Fields
			$\mathfrak{p}$	K
$e(\mathfrak{p}/p)$	$f(\mathfrak{p}/p)$	1	$\mid$	$\mid$
			$\mathfrak{q}$	L
1	1	$[L : \mathbb{Q}]$	$\mid$	$\mid$
			p	$\mathbb{Q}$

Table 9.3.1: Decomposition in the fixed field $L = K^D$.

9.3.1 Galois groups of finite fields

Each $\sigma \in D = D_{\mathfrak{p}}$ acts in a well-defined way on the finite field $k_{\mathfrak{p}} = \mathcal{O}_K/\mathfrak{p}$, so we obtain a homomorphism

$$\varphi : D_{\mathfrak{p}} \to \mathrm{Aut}(k_{\mathfrak{p}}/\mathbb{F}_p).$$

We pause for a moment and review a few basic properties of extensions of finite fields. In particular, they turn out to be Galois so the map φ above is actually a map $D_{\mathfrak{p}} \to \mathrm{Gal}(k_{\mathfrak{p}}/\mathbb{F}_p)$. The properties in this section are general properties of Galois groups for finite fields.

Definition 9.3.6. Let k be any field of characteristic p. Define $\mathrm{Frob}_p : k \to k$ to be the homomorphism given by $a \mapsto a^p$. The map Frob_p is called the *Frobenius* homomorphism.

Exercise 9.3.7.

(a) Show the map Frob_p is in fact a field homomorphism, that is $\mathrm{Frob}_p(a+b) = \mathrm{Frob}_p(a) + \mathrm{Frob}_p(b)$ and $\mathrm{Frob}_p(ab) = \mathrm{Frob}_p(a)\,\mathrm{Frob}_p(b)$.

(b) Suppose $k = \mathbb{F}_p$. Then show $\mathrm{Frob}_p = id$, i.e., $a^p = a$ for any $a \in \mathbb{F}_p$.

(c) Suppose $k = \mathbb{F}_q$ where $q = p^f$ for some $f \geq 1$. Show that $\mathrm{Frob}_p : k \to k$ is an automorphism.

(d) Continuing part (c), note that by Exercise 8.1.9 k^* is cyclic. Let $a \in k$ be a generator for k^*, so a has multiplicative order $p^f - 1$ and $k = \mathbb{F}_p(a)$. Show that

$$\mathrm{Frob}_p^n(a) = a^{p^n} = a \quad \Leftrightarrow \quad (p^f - 1) \mid p^n - 1 \quad \Leftrightarrow \quad f \mid n$$

Remark 9.3.8. Exercise 9.3.7 shows that all finite fields are *perfect*. For more on perfect fields see a standard abstract algebra text such as [**?**].

By Exercise 9.3.7(b,c) the map Frob_p is an automorphism of $k_{\mathfrak{p}}$ fixing $\mathbb{F}_p$ and hence defines an element in $\mathrm{Gal}(k_{\mathfrak{p}}/\mathbb{F}_p)$. Let $f = f_{\mathfrak{p}/p}$ be the residue degree of $\mathfrak{p}$, i.e., $f = [k_{\mathfrak{p}} : \mathbb{F}_p]$. Exercise 9.3.7(d) shows the order of Frob_p is f. Since the order of the automorphism group of a field extension is at most the degree of the extension, we conclude that $\mathrm{Aut}(k_{\mathfrak{p}}/\mathbb{F}_p)$ is generated by Frob_p. This shows $\mathrm{Aut}(k_{\mathfrak{p}}/\mathbb{F}_p)$ has order equal to the degree $[k_{\mathfrak{p}}/\mathbb{F}_p]$ so we conclude that $k_{\mathfrak{p}}/\mathbb{F}_p$ is Galois. We summarize the discussion into the following theorem.

Theorem 9.3.9. *The extension $k_{\mathfrak{p}}/\mathbb{F}_p$ is Galois and moreover,* $\mathrm{Gal}(k_{\mathfrak{p}}/\mathbb{F}_p)$ *is generated by the Frobenius map* Frob_p *defined by $a \mapsto a^p$.*

Exercise 9.3.10. Prove that up to isomorphism there is exactly one finite field of each degree.

[*Hint*: By Theorem 9.3.9 all elements in a finite field satisfy an equation of the form $x^{p^f} - 1$ where p is the characteristic and f is the degree over the field $\mathbb{F}_p$.]

9.3.2 The Exact Sequence

Because $D_{\mathfrak{p}}$ preserves $\mathfrak{p}$, there is a natural reduction homomorphism

$$\varphi : D_{\mathfrak{p}} \to \mathrm{Gal}(k_{\mathfrak{p}}/\mathbb{F}_p).$$

Theorem 9.3.11. *The homomorphism φ is surjective.*

Proof. Let $D = D_{\mathfrak{p}}$ and $\tilde{a} \in k_{\mathfrak{p}}$ be an element such that $k_{\mathfrak{p}} = \mathbb{F}_p(\tilde{a})$. Lift $\tilde{a}$ to an algebraic integer $a \in \mathcal{O}_K$, and let $h = \prod_{\sigma \in D}(x - \sigma(a)) \in K^D[x]$. Let $\tilde{h}$ be the reduction of h modulo $\mathfrak{p}$. Note that $h(a) = 0$ so $\tilde{h}(\tilde{a}) = 0$.

Note that the coefficients of h lie in $\mathcal{O}_{K^D}$. By Proposition 9.3.5, the residue field of $\mathcal{O}_{K^D}$ is $\mathbb{F}_p$ so $\tilde{h} \in \mathbb{F}_p[x]$. Therefore $\tilde{h}$ is a multiple of the minimal polynomial of $\tilde{a}$ over $\mathbb{F}_p$. In particular, $\mathrm{Frob}_p(\tilde{a})$ must also be a root of $\tilde{h}$. Since the roots of $\tilde{h}$ are of the form $\widetilde{\sigma(a)}$ this shows that $\widetilde{\sigma(a)} = \mathrm{Frob}(\tilde{a})$ for some $\sigma \in D$. Hence $\varphi(\sigma)(\tilde{a}) = \mathrm{Frob}(\tilde{a})$. Since elements of $\mathrm{Gal}(K_{\mathfrak{p}}/\mathbb{F}_p)$ are determined by their action on $\tilde{a}$ by choice of $\tilde{a}$, it follows that $\varphi(\sigma) = \mathrm{Frob}$ and hence φ is surjective because Frob_p generates $\mathrm{Gal}(k_{\mathfrak{p}}/\mathbb{F}_p)$. $\square$

Definition 9.3.12 (Inertia Group)**.** The *inertia group associated to* $\mathfrak{p}$ is the kernel $I_{\mathfrak{p}}$ of $D_{\mathfrak{p}} \to \mathrm{Gal}(k_{\mathfrak{p}}/\mathbb{F}_p)$.

We have an exact sequence of groups

$$1 \to I_{\mathfrak{p}} \to D_{\mathfrak{p}} \to \mathrm{Gal}(k_{\mathfrak{p}}/\mathbb{F}_p) \to 1. \tag{9.3.1}$$

The inertia group is a measure of how p ramifies in K.

Corollary 9.3.13. *We have $\#I_{\mathfrak{p}} = e = e(\mathfrak{p}/p)$.*

Proof. The exact sequence (9.3.1) implies that $\#I_{\mathfrak{p}} = \#D_{\mathfrak{p}}/f$ where $f = f(\mathfrak{p}/p) = [k_{\mathfrak{p}} : \mathbb{F}_p]$. Applying Propositions 9.3.4 and 9.3.5, we have

$$\#D_{\mathfrak{p}} = [K : L] = \frac{[K : \mathbb{Q}]}{g} = \frac{efg}{g} = ef.$$

Dividing both sides by f proves the corollary. $\square$

We have the following characterization of $I_{\mathfrak{p}}$.

Proposition 9.3.14. *Let $K/\mathbb{Q}$ be a Galois extension with group G, and let $\mathfrak{p}$ be a prime of $\mathcal{O}_K$ lying over a prime p. Then*

$$I_{\mathfrak{p}} = \{\sigma \in G : \sigma(a) \equiv a \pmod{\mathfrak{p}} \text{ for all } a \in \mathcal{O}_K\}.$$

Proof. By definition $I_{\mathfrak{p}} = \{\sigma \in D_{\mathfrak{p}} : \sigma(a) \equiv a \pmod{\mathfrak{p}} \text{ for all } a \in \mathcal{O}_K\}$, so it suffices to show that if $\sigma \notin D_{\mathfrak{p}}$, then there exists $a \in \mathcal{O}_K$ such that $\sigma(a) \not\equiv a \pmod{\mathfrak{p}}$. If $\sigma \notin D_{\mathfrak{p}}$, then $\sigma^{-1} \notin D_{\mathfrak{p}}$, so $\sigma^{-1}(\mathfrak{p}) \neq \mathfrak{p}$. Since both are maximal ideals, there exists $a \in \mathfrak{p}$ with $a \notin \sigma^{-1}(\mathfrak{p})$, i.e., $\sigma(a) \notin \mathfrak{p}$. Thus $\sigma(a) \not\equiv a \pmod{\mathfrak{p}}$. $\square$

9.4 Frobenius Elements

Suppose that $K/\mathbb{Q}$ is a finite Galois extension with group G and p is a prime such that $e = 1$ (i.e., an unramified prime). Then $I = I_{\mathfrak{p}} = 1$ for any $\mathfrak{p} \mid p$, so the map φ of Theorem 9.3.11 is a canonical isomorphism $D_{\mathfrak{p}} \cong \mathrm{Gal}(k_{\mathfrak{p}}/\mathbb{F}_p)$. By Section 9.3.1, the group $\mathrm{Gal}(k_{\mathfrak{p}}/\mathbb{F}_p)$ is cyclic with canonical generator Frob_p. The *Frobenius element* corresponding to $\mathfrak{p}$ is $\mathrm{Frob}_{\mathfrak{p}} \in D_{\mathfrak{p}}$. It is the unique (see Exercise 9.4.1) element of G such that for all $a \in \mathcal{O}_K$ we have

$$\mathrm{Frob}_{\mathfrak{p}}(a) \equiv a^p \pmod{\mathfrak{p}}.$$

Exercise 9.4.1. With the notation above, prove that $\mathrm{Frob}_{\mathfrak{p}}$ is unique. That is, if σ satisfies $\sigma(a) \equiv a^p \pmod{\mathfrak{p}}$ for all $a \in \mathcal{O}_K$ then $\sigma = \mathrm{Frob}_{\mathfrak{p}}$.

[*Hint*: First show $\sigma \in D_{\mathfrak{p}}$, then argue as in the proof of Proposition 9.3.14.]

Just as the primes $\mathfrak{p}$ and decomposition groups $D_{\mathfrak{p}}$ are all conjugate, the Frobenius elements corresponding to primes $\mathfrak{p} \mid p$ are all conjugate as elements of G.

Proposition 9.4.2. *For each $\sigma \in G$, we have*

$$\mathrm{Frob}_{\sigma\mathfrak{p}} = \sigma \, \mathrm{Frob}_{\mathfrak{p}} \, \sigma^{-1}.$$

In particular, the Frobenius elements lying over a given prime are all conjugate.

Proof. Fix $\sigma \in G$. For any $a \in \mathcal{O}_K$ we have $\mathrm{Frob}_{\mathfrak{p}}(\sigma^{-1}(a)) - \sigma^{-1}(a)^p \in \mathfrak{p}$. Applying σ to both sides, we see that $\sigma \, \mathrm{Frob}_{\mathfrak{p}}(\sigma^{-1}(a)) - a^p \in \sigma\mathfrak{p}$, so $\sigma \, \mathrm{Frob}_{\mathfrak{p}} \, \sigma^{-1} = \mathrm{Frob}_{\sigma\mathfrak{p}}$. $\square$

Thus the conjugacy class of $\mathrm{Frob}_{\mathfrak{p}}$ in G is a well-defined function of p. For example, if G is abelian, then $\mathrm{Frob}_{\mathfrak{p}}$ does not depend on the choice of $\mathfrak{p}$ lying over p and we obtain a well defined symbol $\left(\frac{K/\mathbb{Q}}{p}\right) = \mathrm{Frob}_{\mathfrak{p}} \in G$ called the *Artin symbol*. It extends to a homomorphism from the free abelian group on unramified primes p to G. Class field theory (for $\mathbb{Q}$) sets up a natural bijection between abelian Galois extensions of $\mathbb{Q}$ and certain maps from certain subgroups of the group of fractional ideals for $\mathbb{Z}$ (i.e., $\mathbb{Q}^*$). We have just described one direction of this bijection, which associates to an abelian extension the Artin symbol (which is a homomorphism). The Kronecker-Weber theorem asserts that the abelian extensions of $\mathbb{Q}$ are exactly the subfields of the fields $\mathbb{Q}(\zeta_n)$, as n varies over all positive integers. By Galois theory there is a correspondence between the subfields of the field $\mathbb{Q}(\zeta_n)$, which has Galois group $(\mathbb{Z}/n\mathbb{Z})^*$, and the subgroups of $(\mathbb{Z}/n\mathbb{Z})^*$. If $H \subseteq (\mathbb{Z}/n\mathbb{Z})^*$ is the subgroup corresponding to $K \subset \mathbb{Q}(\zeta_n)$ then the Artin reciprocity map $p \mapsto \left(\frac{K/\mathbb{Q}}{p}\right)$ is given by $p \mapsto [p] \in (\mathbb{Z}/n\mathbb{Z})^*/H$.

Remark 9.4.3. Notice above that the n used is not unique. That is, if K is an abelian extension of $\mathbb{Q}$ then it lies in some $\mathbb{Q}(\zeta_n)$. But then it also lies inside of $\mathbb{Q}(\zeta_{dn})$ for any positive integer d. However, a different choice of n would mean a different choice of H. Note that the quotient $(\mathbb{Z}/n\mathbb{Z})^*/H$ used is not dependent on n since it is isomorphic to the Galois group of $K/\mathbb{Q}$.

9.5 The Artin Conjecture

The Galois group $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ is an object of central importance in number theory, and we can interpret much of number theory as the study of this group. A good way to study a group is to study how it acts on various objects, that is, to study its representations.

Endow $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ with the topology which has as a basis of open neighborhoods of the origin the subgroups $\mathrm{Gal}(\overline{\mathbb{Q}}/K)$, where K varies over finite Galois extensions of $\mathbb{Q}$. Fix a positive integer n and let $\mathrm{GL}_n(\mathbb{C})$ be the group of $n \times n$ invertible matrices over $\mathbb{C}$ with the discrete topology.

Warning 9.5.1. The topology on $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ is **not** the topology induced by taking as a basis of open neighborhoods around the origin the collection of finite-index normal subgroups of $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$, see [**?**, Ch. 7] or Exercise 9.5.5. In particular, there exist nonopen normal subgroups of finite index which do not correspond to subgroups $\mathrm{Gal}(\overline{\mathbb{Q}}/K)$ for some finite Galois extension $K/\mathbb{Q}$.

Definition 9.5.2. A *complex n-dimensional representation* of $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ is a continuous homomorphism

$$\rho : \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}) \to \mathrm{GL}_n(\mathbb{C}).$$

For ρ to be continuous means that if K is the fixed field of $\mathrm{Ker}(\rho)$, then $K/\mathbb{Q}$ is a finite Galois extension. We have a diagram

$$\begin{tikzcd}
\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}) \arrow[rr, "\rho"] \arrow[dr] & & \mathrm{GL}_n(\mathbb{C}) \\
& \mathrm{Gal}(K/\mathbb{Q}) \arrow[ur, hook, "\rho'"'] &
\end{tikzcd}$$

Exercise 9.5.3. Suppose $\rho : \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}) \to \mathrm{GL}_n(\mathbb{C})$ is continuous. Show that the image is finite.

Remark 9.5.4. The converse to Exercise 9.5.3 is **false** in general (see Exercise 9.5.5). This is essentially the same warning as Warning 9.5.1, however it is worth pointing out to avoid mistakes.[2]

Exercise 9.5.5. Find a nonopen subgroup of index 2 in $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$. Note this is also an example of a non-continuous homomorphism $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}) \to \mathrm{GL}_n(\mathbb{C})$ with finite image.

[*Hint*: Use Zorn's lemma to show that there are homomorphisms $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}) \to \{\pm 1\}$ with finite image that are not continuous, since they do not factor through the Galois group of any finite Galois extension.]

[*Hint*: The extension $\mathbb{Q}(\sqrt{d}, d \in \mathbb{Q}^*/(\mathbb{Q}^*)^2)$ is an extension of $\mathbb{Q}$ with Galois group $X \approx \prod \mathbb{F}_2$. The index-two open subgroups of X correspond to the quadratic extensions of $\mathbb{Q}$. However, Zorn's lemma implies that X contains many index-two subgroups that do not correspond to quadratic extensions of $\mathbb{Q}$.]

Fix a Galois representation ρ and let K be the fixed field of $\ker(\rho)$, so ρ factors through $\mathrm{Gal}(K/\mathbb{Q})$. For each prime $p \in \mathbb{Z}$ that is not ramified in K, there is an element $\mathrm{Frob}_\mathfrak{p} \in \mathrm{Gal}(K/\mathbb{Q})$ that is well-defined up to conjugation by elements of $\mathrm{Gal}(K/\mathbb{Q})$. This means that $\rho'(\mathrm{Frob}_p) \in \mathrm{GL}_n(\mathbb{C})$ is well-defined up to conjugation. Thus the characteristic polynomial $F_p(x) \in \mathbb{C}[x]$ of $\rho'(\mathrm{Frob}_p)$ is a well-defined invariant of p and ρ. Let

$$R_p(x) = x^{\deg(F_p)} \cdot F_p(1/x) = 1 + \cdots + \det(\mathrm{Frob}_p) \cdot x^{\deg(F_p)}$$

be the polynomial obtain by reversing the order of the coefficients of F_p. Following E. Artin [Art23, Art30], set

$$L(\rho, s) = \prod_{p \text{ unramified}} \frac{1}{R_p(p^{-s})}. \tag{9.5.1}$$

We view $L(\rho, s)$ as a function of a single complex variable s. One can prove that $L(\rho, s)$ is holomorphic on some right half plane, and extends to a meromorphic function on all $\mathbb{C}$.

[2] See [**?**, Pg. 1].

Conjecture 9.5.6 (Artin)**.** *The L-function of any continuous representation*

$$\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}) \to \mathrm{GL}_n(\mathbb{C})$$

is an entire function on all $\mathbb{C}$, except possibly at 1.

This conjecture asserts that there is some way to analytically continue $L(\rho, s)$ to the whole complex plane, except possibly at 1. (A standard fact from complex analysis is that this analytic continuation must be unique.) The simple pole at $s = 1$ corresponds to the trivial representation (the Riemann zeta function), and if $n \geq 2$ and ρ is irreducible, then the conjecture is that ρ extends to a holomorphic function on all $\mathbb{C}$.

The conjecture is known when $n = 1$. Assume for the rest of this paragraph that ρ is odd, i.e., if $c \in \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ is complex conjugation, then $\det(\rho(c)) = -1$. When $n = 2$ and the image of ρ in $\mathrm{PGL}_2(\mathbb{C})$ is a solvable group, the conjecture is known, and is a deep theorem of Langlands and others (see [Lan80]), which played a crucial roll in Wiles's proof of Fermat's Last Theorem. When $n = 2$ and the image of ρ in $\mathrm{PGL}_2(\mathbb{C})$ is not solvable, the only possibility is that the projective image is isomorphic to the alternating group A_5. Because A_5 is the symmetry group of the icosahedron, these representations are called *icosahedral*. In this case, Joe Buhler's Harvard Ph.D. thesis [Buh78] gave the first example in which ρ was shown to satisfy Conjecture 9.5.6. There is a book [Fre94], which proves Artin's conjecture for 7 icosahedral representation (none of which are twists of each other). Kevin Buzzard and the author proved the conjecture for 8 more examples [BS02]. Subsequently, Richard Taylor, Kevin Buzzard, Nick Shepherd-Barron, and Mark Dickinson proved the conjecture for an infinite class of icosahedral Galois representations (disjoint from the examples) [BDSBT01]. The general problem for $n = 2$ is in fact now completely solved, due to recent work of Khare and Wintenberger [KW08] that proves Serre's conjecture.

Chapter 10

Elliptic Curves, Galois Representations, and L-functions

This chapter is about elliptic curves and the central role they play in algebraic number theory. Our approach will be less systematic and more a survey than most of the rest of this book. The goal is to give you a glimpse of the forefront of research by assuming many basic facts that can be found in other books (see, e.g., [Sil92]).

10.1 Groups Attached to Elliptic Curves

Definition 10.1.1 (Elliptic Curve)**.** An *elliptic curve* over a field K is a genus one curve E defined over K equipped with a distinguished point $\mathcal{O} \in E(K)$. Here $E(K)$ is the set of all points on E defined over K.

We will not define *genus* in this book, except to note that a nonsingular curve over K has genus one if and only if over $\overline{K}$ it can be realized as a nonsingular plane cubic curve.[1] Moreover, one can show (using the Riemann-Roch formula) that over any field a genus one curve with a rational point can always be defined by a projective cubic equation of the form

$$Y^2Z + a_1XYZ + a_3YZ^2 = X^3 + a_2X^2Z + a_4XZ^2 + a_6Z^3.$$

In this form the distinguished point $\mathcal{O}$ is $(X : Y : Z) = (0 : 1 : 0)$. Note that $\mathcal{O}$ is the only point on the curve with $Z = 0$. So we can consider the rest of the curve in the affine coordinates by projecting onto the affine plane defined by $Z \neq 0$. This gives the equation

$$y^2 + a_1xy + a_3y = x^3 + a_2x^2 + a_4x + a_6. \qquad (10.1.1)$$

[1] For a detailed and technical explanation of genus see [Har77, Ch. II.8] or [**?**, Ch. 7.3]

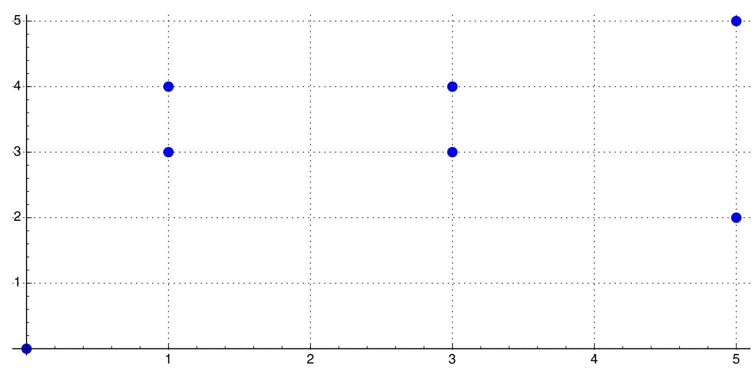
Thus one often presents an elliptic curve by giving a *Weierstrass equation* (10.1.1), though there are significant computational advantages to other equations for curves (e.g., Edwards coordinates – see work of Bernstein and Lange in [?]).

Using Sage we plot an elliptic curve over the finite field $\mathbb{F}_7$ and an elliptic curve curve defined over $\mathbb{Q}$.

```
E = EllipticCurve(GF(7), [1,0])
E
```

```
Elliptic Curve defined by y^2 = x^3 + x over
    Finite Field of size 7
```

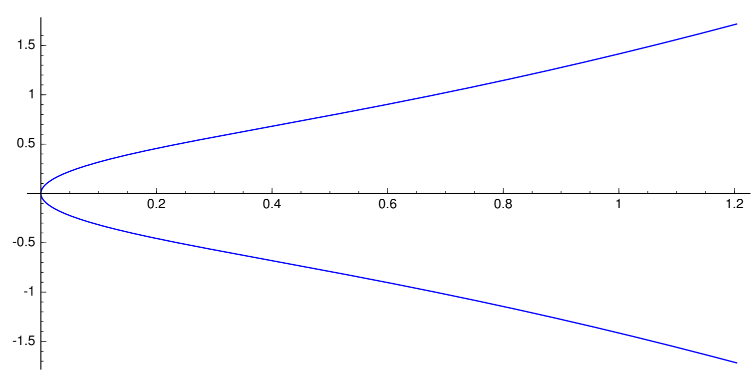
```
E.plot(pointsize=60, gridlines=True)
```




```
E = EllipticCurve([1,0])
E
```

```
Elliptic Curve defined by y^2 = x^3 + x over
    Rational Field
```

```
E.plot()
```



Note that both plots above are of the affine equation $y^2 = x^3 + x$, and do not include the distinguished point $\mathcal{O}$, which lies at infinity.

Remark 10.1.2. The command `EllipticCurve` in Sage can take as input a list `[a4,a6]` of coefficients and returns an elliptic curve given by a Weirstrass equation with $a_1 = a_2 = a_3 = 0$ and a_4, a_6 as specified.

10.1.1 Abelian Groups Attached to Elliptic Curves

If E is an elliptic curve over K, then we give the set $E(K)$ of all K-rational points on E the structure of abelian group with identity element $\mathcal{O}$.[2] If we embed E in the projective plane, then this group is determined by the condition that three points sum to the zero element $\mathcal{O}$ if and only if they lie on a common line (some care needs to be taken when the points are not distinct). In our affine picture, a line will intersect the point at infinity if it is vertical, or equivalently if it of the form $x = a$ for some fixed $a \in K$.

Example 10.1.3. On the curve $y^2 = x^3 - 5x + 4$, we have $(0,2) + (1,0) = (3,4)$. This is because $(0,2)$, $(1,0)$, and $(3,-4)$ are on a common line (given by the equation $y = 2 - 2x$) hence they sum to zero:

$$(0,2) + (1,0) + (3,-4) = \mathcal{O}.$$

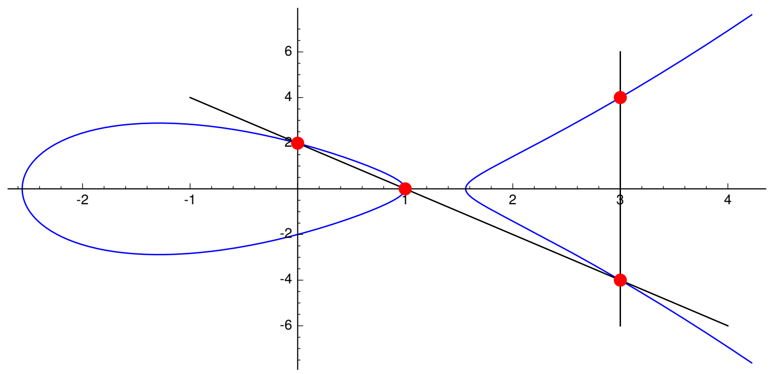
Notice $(3,4)$, $(3,-4)$, and $\mathcal{O}$ (the point at infinity on the curve) are also on a common line (given by $x = 3$), so $(3,4) = -(3,-4)$. We can illustration this in Sage:

```
E = EllipticCurve([-5,4])
E(0,2) + E(1,0)
```

```
(3 : 4 : 1)
```

[2] As a reminder, we will not give rigorous proofs of any facts in this section. For a more detailed and technical explanation of the group structure for elliptic curves see [Sil92, Ch. III.2].

```
G = E.plot()
G += points ([(0,2) , (1,0) , (3,4) , (3,-4)],
    pointsize=90 , color='red', zorder=10)
G += line ([(-1,4) , (4,-6)] , color='black')
G += line ([(3,-6) , (3,6)] , color='black')
G.show()
```



Iterating the group operation often leads quickly to very complicated points:

```
7*E(0,2)
```

```
(14100601873051200/48437552041038241 :
-170870044187066778452359222/10660394576906522772066289 :
 1)
```

Remark 10.1.4. In the previous example we saw that iterating the group operation led to points which used a lot of digits to write down. This notion can be made formal and is called the *height* of the point. The height function is used to prove the general Mordell-Weil theorem, see [Sil92, Ch. VIII.4]

Exercise 10.1.5. Let E be an elliptic curve given by a Weirstrass equation such as (10.1.1). Show that the points of order two are exactly the points on E with y-coordinate equal to 0.

[*Hint*: Recall that a point P has order 2 if $P + P + \mathcal{O} = \mathcal{O}$, which means the tangent line at P goes through the point at infinity.]

That the above condition—three points on a line sum to zero—defines an abelian group structure on $E(K)$ is not obvious. Depending on your perspective, the trickiest part is seeing that the operation satisfies the associative axiom. The best way to understand the group operation on $E(K)$ is to view $E(K)$ as being related to a class group. As a first observation, note that the ring

$$R = K[x, y]/(y^2 + a_1xy + a_3y - (x^3 + a_2x^2 + a_4x + a_6))$$

is a Dedekind domain, so $\mathrm{Cl}(R)$ is defined, and every nonzero fractional ideal can be written uniquely in terms of prime ideals. When K is a perfect field, the prime ideals correspond to the Galois orbits of affine points of $E(\overline{K})$. Note that these do not include the point at infinity.

Let $\mathrm{Div}(E/K)$ be the free abelian group on the Galois orbits of points of $E(\overline{K})$, which as explained above is analogous to the group of fractional ideals of a number field (here we *do* include the point at infinity). We call the elements of $\mathrm{Div}(E/K)$ *divisors*. Let $\mathrm{Pic}(E/K)$ be the quotient of $\mathrm{Div}(E/K)$ by the *principal divisors*, i.e., the divisors associated to rational functions $f \in K(E)^*$ via

$$f \mapsto (f) = \sum_P \mathrm{ord}_P(f)[P].$$

Here $K(E)$ is the fraction field of the ring R defined above. Note that the principal divisor associated to f is analogous to the principal fractional ideal associated to a nonzero element of a number field. The definition of $\mathrm{ord}_P(f)$ is analogous to the "power of P that divides the principal ideal generated by f". Define the *class group* $\mathrm{Pic}(E/K)$ to be the quotient of the divisors by the principal divisors, so we have an exact sequence:

$$1 \to K(E)^*/K^* \to \mathrm{Div}(E/K) \to \mathrm{Pic}(E/K) \to 0.$$

A key difference between elliptic curves and algebraic number fields is that the principal divisors in the context of elliptic curves all have degree 0, i.e., the sum of the coefficients of the divisor (f) is always 0. This might be a familiar fact to you: the number of zeros of a nonzero rational function on a projective curve equals the number of poles, counted with multiplicity. If we let $\mathrm{Div}^0(E/K)$ denote the subgroup of divisors of degree 0, then we have an exact sequence

$$1 \to K(E)^*/K^* \to \mathrm{Div}^0(E/K) \to \mathrm{Pic}^0(E/K) \to 0.$$

To connect this with the group law on $E(K)$, note that there is a natural map

$$E(K) \to \mathrm{Pic}^0(E/K), \qquad P \mapsto [P - \mathcal{O}].$$

Using the Riemann-Roch theorem, one can prove that this map is a bijection, which is moreover an isomorphism of abelian groups. Thus really when we discuss the group of K-rational points on an E, we are talking about the class group $\mathrm{Pic}^0(E/K)$.

Recall that we proved (Theorem 7.1.2) that the class group $\mathrm{Cl}(\mathcal{O}_K)$ of a number field is finite. The group $\mathrm{Pic}^0(E/K) = E(K)$ of an elliptic curve can be either finite (e.g., for $y^2 + y = x^3 - x + 1$) or infinite (e.g., for $y^2 + y = x^3 - x$), and determining which is the case for any particular curve is one of the central unsolved problems in number theory.

The Mordell-Weil theorem (see Chapter 12) asserts that if E is an elliptic curve over a number field K, then there is a nonnegative integer r, referred to as the *algebraic rank of* E, such that

$$E(\mathbb{Q}) \approx \mathbb{Z}^r \oplus T, \qquad (10.1.2)$$

where T is a finite group. This is similar to Dirichlet's unit theorem, which gives the structure of the unit group of the ring of integers of a number field. The main difference is that T need not be cyclic, and computing r appears to be much more difficult than just finding the number of real and complex roots of a polynomial!

Example 10.1.6. Sage has algorithms which can compute this rank for us. For example we can compute the ranks of the curves $y^2+y = x^3-x+1$ and $y^2+y = x^3-x$ respectively.

```
EllipticCurve([0,0,1,-1,1]).rank()
```

```
0
```

```
EllipticCurve([0,0,1,-1,0]).rank()
```

```
1
```

Also, if L/K is an arbitrary extension of fields, and E is an elliptic curve over K, then there is a natural inclusion homomorphism $E(K) \hookrightarrow E(L)$. Thus instead of just obtaining one group attached to an elliptic curve, we obtain a whole collection, one for each extension of L. Even more generally, if S/K is an arbitrary scheme, then $E(S)$ is a group, and the association $S \mapsto E(S)$ defines a functor from the category of schemes to the category of groups. Thus each elliptic curve gives rise to map:

$$\{\text{Schemes over } K\} \longrightarrow \{\text{Abelian Groups}\}$$

Remark 10.1.7. Elliptic curves are not the only objects that induce a functor from schemes to groups. *Abelian varieties* are a larger class of schemes, which includes elliptic curves, that also induce such a functor. For more on Abelian varieties see [**?**].

10.1.2 A Formula for Adding Points

We close this section with an explicit formula for adding two points in $E(K)$. If E is an elliptic curve over a field K, given by an equation $y^2 = x^3 + ax + b$, then we can compute the group addition using the following algorithm.

Algorithm 10.1.8 (Elliptic Curve Group Law)**.** Given $P_1, P_2 \in E(K)$, this algorithm computes the sum $R = P_1 + P_2 \in E(K)$.

1. [One Point $\mathcal{O}$] If $P_1 = \mathcal{O}$ set $R = P_2$ or if $P_2 = \mathcal{O}$ set $R = P_1$ and terminate. Otherwise write $P_i = (x_i, y_i)$.
2. [Negatives] If $x_1 = x_2$ and $y_1 = -y_2$, set $R = \mathcal{O}$ and terminate.
3. [Compute λ] Set $\lambda = \begin{cases} (3x_1^2 + a)/(2y_1) & \text{if } P_1 = P_2, \\ (y_1 - y_2)/(x_1 - x_2) & \text{otherwise.} \end{cases}$
   Note: If $y_1 = 0$ and $P_1 = P_2$, output $\mathcal{O}$ and terminate.

4. [Compute Sum] Then $R = \left(\lambda^2 - x_1 - x_2, -\lambda x_3 - \nu\right)$, where $\nu = y_1 - \lambda x_1$ and x_3 is the x coordinate of R.

10.1.3 Other Groups

There are other abelian groups attached to elliptic curves, such as the torsion subgroup $E(K)_{\mathrm{tor}}$ of elements of $E(K)$ of finite order. The torsion subgroup is (isomorphic to) the group T that appeared in Equation (10.1.2) above). When K is a number field, there is a group called the Shafarevich-Tate group $\text{Ш}(E/K)$ attached to E, which plays a role similar to that of the class group of a number field (though it is an open problem to prove that $\text{Ш}(E/K)$ is finite in general). The definition of $\text{Ш}(E/K)$ involves Galois cohomology, so we wait until Chapter 11 to define it. There are also component groups attached to E, one for each prime of $\mathcal{O}_K$. These groups all come together in the Birch and Swinnerton-Dyer conjecture (see `http://wstein.org/books/bsd/`).

10.2 Galois Representations Attached to Elliptic Curves

Let E be an elliptic curve over a number field K. In this section we attach representations of $G_K = \mathrm{Gal}(\overline{K}/K)$ to E, and use them to define an L-function $L(E,s)$. This L-function is yet another generalization of the Riemann Zeta function, that is different from the L-functions attached to complex representations $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}) \to \mathrm{GL}_n(\mathbb{C})$, which we encountered before in Section 9.5.

There is a natural action of G_K on the points of $E(\overline{K})$. Given a point $P = (a,b) \in E(\overline{K})$ we define $\sigma(P)$ to be the point $(\sigma(a), \sigma(b))$. Since E is defined over K the point $\sigma(P)$ will again lie on E so the action is well defined. Note that the group structure on E is defined by algebraic formulas with coefficients in K. It follows that the action commutes with point addition meaning that $\sigma(P+Q) = \sigma(P) + \sigma(Q)$. Now fix an integer n. From what we have seen, the subgroup

$$E[n] = \{P \in E(\overline{K}) : nP = \mathcal{O}\}$$

is invariant under the action of G_K. We thus obtain a homomorphism

$$\overline{\rho}_{E,n} : G_K \to \mathrm{Aut}(E[n]).$$

Warning 10.2.1. Though the action of G_K leaves the group $E[n]$ fixed, it may act non-trivially on individual elements! Otherwise $\overline{\rho}_{E,n}$ would not be very interesting.

For any positive integer n, the group $E[n]$ is isomorphic as an abstract abelian group to $(\mathbb{Z}/n\mathbb{Z})^2$. There are various related ways to see why this is true. One is to use the Weierstrass $\wp$-theory to parametrize $E(\mathbb{C})$ by the the complex numbers, i.e., to find an isomorphism $\mathbb{C}/\Lambda \cong E(\mathbb{C})$, where Λ is a lattice in $\mathbb{C}$ and the isomorphism is given by $z \mapsto (\wp(z), \wp'(z))$ with respect to an appropriate choice of coordinates on $E(\mathbb{C})$. It is then an easy exercise to verify that $(\mathbb{C}/\Lambda)[n] \cong (\mathbb{Z}/n\mathbb{Z})^2$. For a detailed and rigorous walk through of this method see [**?**, Ch. 1.4].

Another way to understand $E[n]$ is to use the fact that $E(\mathbb{C})_{\text{tor}}$ is isomorphic to the quotient

$$\mathrm{H}_1(E(\mathbb{C}), \mathbb{Q}) / \mathrm{H}_1(E(\mathbb{C}), \mathbb{Z})$$

of homology groups and that the homology of a curve of genus g is isomorphic to $\mathbb{Z}^{2g}$. Then we have a non-canonical isomorphism

$$E[n] \approx (\mathbb{Q}/\mathbb{Z})^2[n] = (\mathbb{Z}/n\mathbb{Z})^2.$$

Technically the previous arguments have shown $E(\mathbb{C})[n] \approx (\mathbb{Z}/n\mathbb{Z})^2$. However, our definition of $E[n]$ used points in $E(\overline{K})$. So we need to show the points $E(\mathbb{C})[n]$ are actually defined over $\overline{K}$. Note that $E(\mathbb{C})[n]$ is finite and invariant under $\mathrm{Aut}(\mathbb{C}/\overline{K})$ for the same reason as $E[n]$ was invariant under $\mathrm{Gal}(\overline{K}/K)$ (point addition is defined by algebraic formulas with coefficients in K). It follows that $E(\mathbb{C})[n]$ is indeed defined over $E(\overline{K})$ so the arguments above show that $E[n] \approx (\mathbb{Z}/n\mathbb{Z})^2$.

Remark 10.2.2. Notice that the arguments above used many analytic facts about geometry over $\mathbb{C}$ (e.g. homology, analytic structure) in order to prove algebraic facts (e.g. the number of torsion points) about $E(\overline{K})$. This is part of a more general concept called the *Lefschetz principle* which generally relates geometry over an algebraically closed field of characteristic 0 to geometry over $\mathbb{C}$. For more on this see [Sil92, Ch. VI.6].

Remark 10.2.3. In fact, if p is a prime that does not divide n then $E[n] \approx (\mathbb{Z}/n\mathbb{Z})^2$ over fields of characteristic p. However, the methods we used above do not apply to the case of positive characteristic. Another method is to show the multiplication by n map is separable and has degree n^2. For a detailed proof see [Sil92, Cor. III.6.4].

Exercise 10.2.4. Let E be an elliptic curve defined over a number field K. Fix an integer n and consider the extension of K given by

$$K(E[n]) = K(\{a, b : (a, b) \in E[n]\}).$$

Show that $K(E[n])/K$ is a finite Galois extension.

Hint: By the arguments above $\#E[n] = n^2$ which shows the extension is finite. Next recall that $E[n]$ is left invariant by the action of $\mathrm{Gal}(\overline{K}/K)$. What can you say about the embeddings from $K(E[n])$ into $\overline{K}$ which leave K fixed?

Example 10.2.5. Consider the case when $n = 2$. From Exercise 10.1.5 we know that the points in $E[2]$ are exactly the points with y-coordinate 0. Let E be the elliptic curve given by $E : y^2 = x^3+x+1$. If $y = 0$ then x has to be a root of the polynomial $x^3 + x + 1$, so the points in $E[2]$ are defined over the splitting field of $x^3 + x + 1$. We can compute these points in Sage.

```
E = EllipticCurve([1,1]); E
```

```
Elliptic Curve defined by y^2 = x^3 + x + 1 over
    Rational Field
```

```
R.<x> = QQ[]; R
```

```
Univariate Polynomial Ring in x over Rational Field
```

```
f = x^3 + x + 1
K.<a> = NumberField(f)
M.<b> = K.galois_closure(); M
```

```
Number Field in b with defining polynomial
    x^6 + 6*x^4 + 9*x^2 + 31
```

```
F = E.change_ring(M)
T = F.torsion_subgroup(); T
```

```
Torsion Subgroup isomorphic to Z/2 + Z/2 associated
    to the Elliptic Curve defined by y^2 = x^3 + x + 1
    over Number Field in b with defining polynomial
    x^6 + 6*x^4 + 9*x^2 + 31
```

```
T.gens()
```

```
((1/18*b^4 + 5/18*b^2 + 1/2*b + 2/9 : 0 : 1),
    (1/18*b^4 + 5/18*b^2 - 1/2*b + 2/9 : 0 : 1))
```

Note that this matches with what we expected: we computed two generators for $E[2]$ (the output of the last cell) corresponding to two generators of $(\mathbb{Z}/2\mathbb{Z})^2$.

If $n = p$ is a prime, then upon chosing a basis for the two-dimensional $\mathbb{F}_p$-vector space $E[p]$, we obtain an isomorphism $\mathrm{Aut}(E[p]) \cong \mathrm{GL}_2(\mathbb{F}_p)$. We thus obtain a mod p Galois representation

$$\overline{\rho}_{E,p} : G_K \to \mathrm{GL}_2(\mathbb{F}_p).$$

This representation $\overline{\rho}_{E,p}$ is continuous if $\mathrm{GL}_2(\mathbb{F}_p)$ is endowed with the discrete topology, because the field $K(E[p])$ is a Galois extension of K of finite degree by Exercise 10.2.4.

In order to attach an L-function to E, one could try to embed $\mathrm{GL}_2(\mathbb{F}_p)$ into $\mathrm{GL}_2(\mathbb{C})$ and use the construction of Artin L-functions from Section 9.5. Unfortunately, this approach is doomed in general, since $\mathrm{GL}_2(\mathbb{F}_p)$ frequently does not embed in $\mathrm{GL}_2(\mathbb{C})$. The following Sage session shows that for $p = 5, 7$, there are no 2-dimensional irreducible representations of $\mathrm{GL}_2(\mathbb{F}_p)$, so $\mathrm{GL}_2(\mathbb{F}_p)$ does not embed in $\mathrm{GL}_2(\mathbb{C})$. The notation in the output below is `[degree of rep, number of times it occurs]`.

```
GL(2,GF(2)).gap().CharacterTable().CharacterDegrees()
```

```
[ [ 1, 2 ], [ 2, 1 ] ]
```

```
GL(2,GF(3)).gap().CharacterTable().CharacterDegrees()
```

```
[ [ 1, 2 ], [ 2, 3 ], [ 3, 2 ], [ 4, 1 ] ]
```

```
GL(2,GF(5)).gap().CharacterTable().CharacterDegrees()
```

```
[ [ 1, 4 ], [ 4, 10 ], [ 5, 4 ], [ 6, 6 ] ]
```

```
GL(2,GF(7)).gap().CharacterTable().CharacterDegrees()
```

```
[ [ 1, 6 ], [ 6, 21 ], [ 7, 6 ], [ 8, 15 ] ]
```

Instead of using the complex numbers, we use the *p-adic numbers*[3], as follows. For each power p^m of p, we have defined a homomorphism

$$\overline{\rho}_{E,p^m} : G_K \to \mathrm{Aut}(E[p^m]) \approx \mathrm{GL}_2(\mathbb{Z}/p^m\mathbb{Z}).$$

We combine together all of these representations (for all $m \geq 1$) using the inverse limit. Recall that the p-adic numbers are

$$\mathbb{Z}_p = \varprojlim \mathbb{Z}/p^m\mathbb{Z},$$

which is the set of all compatible choices of integers modulo p^m for all m. We obtain a (continuous) homomorphism

$$\rho_{E,p} : G_K \to \mathrm{Aut}(\varprojlim E[p^m]) \cong \mathrm{GL}_2(\mathbb{Z}_p),$$

where $\mathbb{Z}_p$ is the ring of p-adic integers. The composition of this homomorphism with the reduction map $\mathrm{GL}_2(\mathbb{Z}_p) \to \mathrm{GL}_2(\mathbb{F}_p)$ is the representation $\overline{\rho}_{E,p}$, which we defined above, which is why we denoted it by $\overline{\rho}_{E,p}$. We next try to mimic the construction of $L(\rho, s)$ from Section 9.5 in the context of a p-adic Galois representation $\rho_{E,p}$.

Definition 10.2.6 (Tate module)**.** The *p-adic Tate module of E* is

$$T_p(E) = \varprojlim E[p^n].$$

Let M be the fixed field of $\ker(\rho_{E,p})$. The image of $\rho_{E,p}$ is infinite, so M is an infinite extension of K. Fortunately, one can prove that M is ramified at only finitely many primes (the primes of *bad reduction* for E and p—see [ST68]). If ℓ is a prime of K, let D_ℓ be a choice of decomposition group for some prime $\mathfrak{p}$ of M lying over ℓ, and let I_ℓ be the inertia group. We haven't defined inertia and decomposition groups for infinite Galois extensions, but the definitions are almost

[3] For a review of p-adic numbers and p-adic analysis see [**?**].

the same: choose a prime of $\mathcal{O}_M$ over ℓ, and let D_ℓ be the subgroup of $\mathrm{Gal}(M/K)$ that leaves $\mathfrak{p}$ invariant. Then the submodule $T_p(E)^{I_\ell}$ of inertia invariants is a module for D_ℓ and the characteristic polynomial $F_\ell(x)$ of Frob_ℓ on $T_p(E)^{I_\ell}$ is well defined (since inertia acts trivially). Let $R_\ell(x)$ be the polynomial obtained by reversing the coefficients of $F_\ell(x)$. One can prove that $R_\ell(x) \in \mathbb{Z}[x]$ and that $R_\ell(x)$, for $\ell \neq p$ does not depend on the choice of p. Define $R_\ell(x)$ for $\ell = p$ using a different prime $q \neq p$, so the definition of $R_\ell(x)$ does not depend on the choice of p.

Definition 10.2.7. The L-series of E is

$$L(E,s) = \prod_\ell \frac{1}{R_\ell(\ell^{-s})}.$$

A prime $\mathfrak{p}$ of $\mathcal{O}_K$ is a prime of *good reduction* for E if there is an equation for E such that $E \mod \mathfrak{p}$ is an elliptic curve over the field $\mathcal{O}_K/\mathfrak{p}$. If $K = \mathbb{Q}$ and ℓ is a prime of good reduction for E, then one can show that that $R_\ell(\ell^{-s}) = 1 - a_\ell \ell^{-s} + \ell^{1-2s}$, where $a_\ell = \ell + 1 - \#\tilde{E}(\mathbb{F}_\ell)$ and $\tilde{E}$ is the reduction of a local minimal model for E modulo ℓ. (There is a similar statement for $K \neq \mathbb{Q}$.)

One can prove using fairly general techniques that the product expression for $L(E,s)$ defines a holomorphic function in some right half plane of $\mathbb{C}$, i.e., the product converges for all s with $\mathrm{Re}(s) > \alpha$, for some real number α.

Recall that the Artin L-function from Section 9.5 (see Equation 9.5.1) extended to meromorphic function on the entire complex plane and Artin conjectured that the L-function of any continuous representation of $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}) \to \mathrm{GL}_n(\mathbb{C})$ also extends to a meromorphic function on $\mathbb{C}$. We could ask the same question for the L-functions attached to elliptic curves. However, we will instead ask for something stronger:

Does the L-function $L(E,s)$ attached to an elliptic curve E extends to a holomorphic function on $\mathbb{C}$?

This question was one of the central topics in number theory in the late 1990s and early 2000s. An amazing fact is that the question has been answered in the affirmative.

Theorem 10.2.8. *The function $L(E,s)$ extends to a holomorphic function on all $\mathbb{C}$.*

This is a corollary to the modularity theorem described in the next section, see Corollary 10.2.10.

10.2.1 Modularity of Elliptic Curves over $\mathbb{Q}$

Fix an elliptic curve E over $\mathbb{Q}$. In this section we will explain what it means for E to be modular, and note the connection with Conjecture 10.2.8 from the previous section.

First, we give the general definition of modular form (of weight 2). The complex *upper half plane* is $\mathfrak{h} = \{z \in \mathbb{C} : \mathrm{Im}(z) > 0\}$. A *cuspidal modular form f* of level N

(of weight 2) is a holomorphic function $f : \mathfrak{h} \to \mathbb{C}$ such that $\lim_{z\to i\infty} f(z) = 0$ and for every integer matrix $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ with determinant 1 and $c \equiv 0 \pmod{N}$, we have

$$f\left(\frac{az+b}{cz+d}\right) = (cz+d)^{-2} f(z).$$

For each prime number ℓ of good reduction, let $a_\ell = \ell + 1 - \#\tilde{E}(\mathbb{F}_\ell)$. If ℓ is a prime of bad reduction let $a_\ell = 0, 1, -1$, depending on how singular the reduction $\tilde{E}$ of E is over $\mathbb{F}_\ell$. If $\tilde{E}$ has a cusp, then $a_\ell = 0$, and $a_\ell = 1$ or -1 if $\tilde{E}$ has a node; in particular, let $a_\ell = 1$ if and only if the tangents at the cusp are defined over $\mathbb{F}_\ell$.

Extend the definition of the a_ℓ to a_n for all positive integers n as follows. If $\gcd(n, m) = 1$ let $a_{nm} = a_n \cdot a_m$. If p^r is a power of a prime p of good reduction, let

$$a_{p^r} = a_{p^{r-1}} \cdot a_p \;-\; p \cdot a_{p^{r-2}}.$$

If p is a prime of bad reduction let $a_{p^r} = (a_p)^r$.

Attach to E the function

$$f_E(z) = \sum_{n=1}^{\infty} a_n e^{2\pi i z}.$$

It is an extremely deep theorem that $f_E(z)$ is actually a cuspidal modular form, and not just some random function.

The following theorem is called the modularity theorem for elliptic curves over $\mathbb{Q}$. Before it was proved it was known as the Taniyama-Shimura-Weil conjecture.

Theorem 10.2.9 (Wiles, Brueil, Conrad, Diamond, Taylor)**.** *Every elliptic curve over $\mathbb{Q}$ is modular, i.e, the function $f_E(z)$ is a cuspidal modular form.*

Corollary 10.2.10 (Hecke)**.** *If E is an elliptic curve over $\mathbb{Q}$, then the L-function $L(E, s)$ has an analytic continuous to the whole complex plane.*

Chapter 11

Galois Cohomology

Let G be a group and suppose G acts on an abelian group A (defined below). In this chapter we will study abelian groups attached to the action of G on A. These are called *cohomology groups* and denoted by $\mathrm{H}^n(G, A)$. The theory of these groups is referred to as *group cohomology.* In the later sections G will represent the Galois group of a field extension. This is called *Galois cohomology.* Studying Galois cohomology helps us understand the structure of Galois groups such as $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$.

11.1 Group Rings and Modules

In this section we define group modules, which are analogous to modules over a ring. For a review of the theory of modules over a ring see [**?**, Ch. 10].

Definition 11.1.1. Let G be any group. The *group ring* $\mathbb{Z}[G]$ of G is the free abelian group (equivalently the free $\mathbb{Z}$-module) on the elements of G equipped with multiplication given by the group structure on G. Note that $\mathbb{Z}[G]$ is a commutative ring if and only if G is abelian.

Example 11.1.2. For example, the group ring of the cyclic group $C_n = \langle a \rangle$ of order n is the free $\mathbb{Z}$-module on $1, a, \ldots, a^{n-1}$, and the multiplication is induced by $a^i a^j = a^{i+j} = a^{i+j \pmod{n}}$ extended linearly. For example, in $\mathbb{Z}[C_3]$ we have

$$(1+2a)(1-a^2) = 1 - a^2 + 2a - 2a^3 = 1 + 2a - a^2 - 2 = -1 + 2a - a^2.$$

Since $a^3 = 1$ you might think that $\mathbb{Z}[C_3]$ is isomorphic to the ring $\mathbb{Z}[\zeta_3]$ of integers of $\mathbb{Q}(\zeta_3)$, but you would be wrong, since the ring of integers is isomorphic to $\mathbb{Z}^2$ as an abelian group, but $\mathbb{Z}[C_3]$ is isomorphic to $\mathbb{Z}^3$ as abelian group. Note that $\mathbb{Q}(\zeta_3)$ is a quadratic extension of $\mathbb{Q}$.

Exercise 11.1.3. Is $\mathbb{Z}[\zeta_3]$ isomorphic to the group ring of some group?

Hint: Note that the rank of the group ring as a $\mathbb{Z}$-module is equal to the size of the group. If $\mathbb{Z}[\zeta_3]$ was a group ring then it would have to be isomorphic to $\mathbb{Z}[C_2]$.

Exercise 11.1.4.

(a) Write down an two elements of $\mathbb{Z}[\mathbb{Z}]$ and multiply them. This is not hard, but is good practice with the concept of a group ring.

(b) Show $\mathbb{Z}[\mathbb{Z}]$ is isomorphic to $\mathbb{Z}\left[x, \frac{1}{x}\right]$.

Definition 11.1.5. Let G be a finite group. A *G-module* is an abelian group A equipped with a left action of G, i.e., a group homomorphism $G \to \mathrm{Aut}(A)$, where $\mathrm{Aut}(A)$ denotes the group of group isomorphisms $A \to A$ with the operation of function composition.

Exercise 11.1.6. Fix an abelian group A. Show the following are equivalent sets of data. Specifically, given any one of the following objects, there is a natural way to construct another.

(a) A group homomorphism $G \to \mathrm{Aut}(A)$.

(b) A map $\rho : G \times A \to A$ such that for all $g, h \in G$ and $a, b \in A$,

  (i) $\rho(g, a + b) = \rho(g, a) + \rho(g, b)$

  (ii) $\rho(e, a) = a$ where e is the identity in G.

  (iii) $\rho(gh, a) = \rho(g, \rho(h, a))$

(c) A ring homomorphism $\mathbb{Z}[G] \to \mathrm{End}(A)$.

(d) A map $\rho : \mathbb{Z}[G] \times A \to A$ with the same properties listed in (b).

Remark 11.1.7. In Exercise 11.1.6, part (a) is our definition of a G-module and parts (c) and (d) are the data of a $\mathbb{Z}[G]$-module. This shows that a G-module in the above sense is the same as a $\mathbb{Z}[G]$-module in the usual module sense.

Example 11.1.8. If G is any finite group and A any abelian group then we can always make A into a G-module by giving it the trivial action. In particular, $\mathbb{Z}$ with the trivial action is a module over any group G, as is $\mathbb{Z}/m\mathbb{Z}$ for any positive integer m. Another example is $G = (\mathbb{Z}/n\mathbb{Z})^*$, which acts via multiplication on $A = \mathbb{Z}/n\mathbb{Z}$.

Remark 11.1.9. The construction $\mathbb{Z}[G]$ from G is natural, in the sense that it defines a functor between categories. Moreover, $\mathbb{Z}[G]$ is the most natural way to construct a ring from a group in the sense that the group ring functor is a left adjoint to the forgetful functor from rings to groups. These types of functors are sometimes called "free" functors. If you are interested in free objects, see if you can come up with a natural way to add structure to other objects. Could you make a set into a group? How about a vector space?

11.2 Group Cohomology

Let G be a finite group and A a G-module. For each integer $n \geq 0$ there is an abelian group $\mathrm{H}^n(G, A)$ called the *nth cohomology group of G acting on A*. The general definition is somewhat complicated, but the definition for $n \leq 1$ is fairly concrete. For example, the 0*th cohomology group*

$$\mathrm{H}^0(G, A) = \{x \in A : \sigma x = x \text{ for all } \sigma \in G\} = G^A$$

is the subgroup of elements of A that are fixed by every element of G.

The *first cohomology group*

$$\mathrm{H}^1(G, A) = C^1(G, A)/B^1(G, A)$$

is the group C^1 of 1-*cocycles* modulo the group B^1 of 1-*coboundaries*, where

$$C^1(G, A) = \{f : G \to A \text{ such that } f(\sigma\tau) = f(\sigma) + \sigma f(\tau)\}$$

where the maps $f : G \to A$ range over all set-theoretic maps. If we let $f_a : G \to A$ denote the set-theoretic map $f_a(\sigma) = \sigma(a) - a$, then

$$B^1(G, A) = \{f_a : a \in A\}.$$

There are also explicit, and increasingly complicated, definitions of $\mathrm{H}^n(G, A)$ for each $n \geq 2$ in terms of *crossed homomorphisms*, which are certain maps $G \times \cdots \times G \to A$ modulo a subgroup. We will not need these maps, but for more information about them see [Cp86, Ch. IV.2].

Exercise 11.2.1. Suppose G acts trivially on A. Show that $B^1(G, A) = 0$ and $C^1(G, A) \cong \mathrm{Hom}(G, A)$. In particular, this shows $\mathrm{H}^1(G, \mathbb{Z}) \cong \mathrm{Hom}(G, A)$. Deduce that if $A = \mathbb{Z}$ then $\mathrm{H}^1(G, \mathbb{Z}) = 0$.

Hint: For any $\sigma \in G$ we have $f_a(\sigma) = \sigma(a) - a = a - a = 0$. Also for any finite group G, show that $\mathrm{Hom}(G, \mathbb{Z}) = 0$.

Example 11.2.2. The groups $H^n(G, \mathbb{Z})$ and $H^n(G, \mathbb{Z}/p\mathbb{Z})$ (where p is a prime) are computable in Sage. For example we can compute $H^{10}(A_5, \mathbb{Z})$ and $H^7(A_5, \mathbb{Z}/5\mathbb{Z})$ where A_5 is the alternating group of order 120 and $\mathbb{Z}/5\mathbb{Z}$ is given the trivial A_5-module structure.

```
G = AlternatingGroup(5); G
```

```
Alternating group of order 5!/2 as a permutation group
```

```
G.cohomology(10)
```

```
Multiplicative Abelian group isomorphic to C2 x C2
```

```
G.cohomology(7,5)
```

```
Multiplicative Abelian group isomorphic to C5
```

11.2.1 The Main Theorem

Definition 11.2.3. If X is any abelian group, then $A = \mathrm{Hom}(\mathbb{Z}[G], X)$ is a G-module. We call a module constructed in this way *coinduced*.

The following theorem gives three properties of group cohomology, which uniquely determine group cohomology.

Theorem 11.2.4. *Suppose G is a finite group. Then*

1. *We have $\mathrm{H}^0(G, A) = A^G$.*

2. *If A is a coinduced G-module, then $\mathrm{H}^n(G, A) = 0$ for all $n \geq 1$.*

3. *If $0 \to A \to B \to C \to 0$ is any exact sequence of G-modules, then there is a long exact sequence*

$$\begin{tikzcd}
0 \arrow[r] & \mathrm{H}^0(G,A) \arrow[r] & \mathrm{H}^0(G,B) \arrow[r] & \mathrm{H}^0(G,C) \arrow[dll, out=-10, in=170] \\
& \mathrm{H}^1(G,A) \arrow[r] & \mathrm{H}^1(G,B) \arrow[r] & \mathrm{H}^1(G,C) \arrow[dl, out=-10, in=170] \\
& & \cdots \arrow[dl, out=-10, in=170] & \\
& \mathrm{H}^n(G,A) \arrow[r] & \mathrm{H}^n(G,B) \arrow[r] & \mathrm{H}^n(G,C) \arrow[dll, out=-10, in=170] \\
& \mathrm{H}^{n+1}(G,A) \arrow[r] & \mathrm{H}^{n+1}(G,B) \arrow[r] & \mathrm{H}^{n+1}(G,C) \arrow[r] & \cdots
\end{tikzcd}$$

Moreover, the functor $\mathrm{H}^n(G, -)$ is uniquely determined by these three properties.

We will not prove this theorem. For proofs see [Cp86, Atiyah-Wall] and [Ser79, Ch. 7]. The properties of the theorem uniquely determine group cohomology, so one should in theory be able to use them to deduce anything that can be deduced about cohomology groups. Indeed, in practice one frequently proves results about higher cohomology groups $\mathrm{H}^n(G, A)$ by writing down appropriate exact sequences, using explicit knowledge of H^0, and chasing diagrams.

Remark 11.2.5. Alternatively, we could view the defining properties of the theorem as the definition of group cohomology, and could state a theorem that asserts that group cohomology exists.

Remark 11.2.6. For those familiar with commutative and homological algebra, we have

$$\mathrm{H}^n(G, A) = \mathrm{Ext}^n_{\mathbb{Z}[G]}(\mathbb{Z}, A),$$

where $\mathbb{Z}$ is the trivial G-module.

Remark 11.2.7. One can interpret $\mathrm{H}^2(G, A)$ as the group of equivalence classes of extensions of G by A, where an extension is an exact sequence

$$0 \to A \to M \to G \to 1$$

such that the induced conjugation action of G on A is the given action of G on A. (Note that G acts by conjugation, as A is a normal subgroup since it is the kernel of a homomorphism.)

11.2.2 Example Application of the Theorem

For example, let's see what we get from the exact sequence

$$0 \to \mathbb{Z} \xrightarrow{m} \mathbb{Z} \to \mathbb{Z}/m\mathbb{Z} \to 0,$$

where m is a positive integer, and $\mathbb{Z}$ has the structure of trivial G module. By definition we have $\mathrm{H}^0(G, \mathbb{Z}) = \mathbb{Z}$ and $\mathrm{H}^0(G, \mathbb{Z}/m\mathbb{Z}) = \mathbb{Z}/m\mathbb{Z}$. The long exact sequence begins

$$\begin{array}{l}
0 \longrightarrow \mathbb{Z} \xrightarrow{\quad m \quad} \mathbb{Z} \longrightarrow \mathbb{Z}/m\mathbb{Z} \\
\quad \longrightarrow \mathrm{H}^1(G, \mathbb{Z}) \xrightarrow{[m]} \mathrm{H}^1(G, \mathbb{Z}) \longrightarrow \mathrm{H}^1(G, \mathbb{Z}/m\mathbb{Z}) \\
\quad \longrightarrow \mathrm{H}^2(G, \mathbb{Z}) \xrightarrow{[m]} \mathrm{H}^2(G, \mathbb{Z}) \longrightarrow \mathrm{H}^2(G, \mathbb{Z}/m\mathbb{Z}) \longrightarrow \cdots
\end{array}$$

From the first few terms of the sequence and the fact that $\mathbb{Z}$ surjects onto $\mathbb{Z}/m\mathbb{Z}$, we see that $[m] : \mathrm{H}^1(G, \mathbb{Z}) \to \mathrm{H}^1(G, \mathbb{Z})$ is injective. This is consistent with Exercise 11.2.1 above that showed $\mathrm{H}^1(G, \mathbb{Z}) = 0$. Using this vanishing and the right side of the exact sequence we obtain an isomorphism

$$\mathrm{H}^1(G, \mathbb{Z}/m\mathbb{Z}) \cong \mathrm{H}^2(G, \mathbb{Z})[m]$$

where $\mathrm{H}^2(G, \mathbb{Z})[m]$ is the kernel of the map $[m] : \mathrm{H}^2(G, \mathbb{Z}) \to \mathrm{H}^2(G, \mathbb{Z})$. By Exercise 11.2.1, when a group acts trivially the H^1 is Hom, so

$$\mathrm{H}^2(G, \mathbb{Z})[m] \cong \mathrm{Hom}(G, \mathbb{Z}/m\mathbb{Z}). \tag{11.2.1}$$

One can prove that for any $n > 0$ and any module A that the group $\mathrm{H}^n(G, A)$ has order dividing $\#G$ (see Remark 11.3.5). Thus (11.2.1) allows us to understand $\mathrm{H}^2(G, \mathbb{Z})$, and this comprehension arose naturally from the properties in Theorem 11.2.4 that determine the cohomology groups H^n.

11.3 Inflation and Restriction

Suppose H is a subgroup of a finite group G and A is a G-module.

For each $n \geq 0$, there is a natural map

$$\mathrm{res}_H : \mathrm{H}^n(G, A) \to \mathrm{H}^n(H, A)$$

called *restriction*. Elements of $\mathrm{H}^n(G, A)$ can be viewed as classes of n-cocycles, which are certain maps $G \times \cdots \times G \to A$. From this perspective res_H takes a map to its restriction $H \times \cdots \times H \to A$. This is equivalent to precomposing with the natural inclusion $H \times \cdots \times H \to G \times \cdots \times G$.

If H is a normal subgroup of G, there is also an *inflation* map

$$\mathrm{inf}_H : \mathrm{H}^n(G/H, A^H) \to \mathrm{H}^n(G, A),$$

given by taking a cocycle $f : G/H \times \cdots \times G/H \to A^H$ and precomposing with the quotient map $G \to G/H$ to obtain a cocycle for G.

Exercise 11.3.1. Show that if A is a G-module then A^H is naturally a G/H-module for any normal subgroup H. Convince yourself that G/H does not in general naturally act on all of A.

The following proposition will be useful when proving the weak Mordell-Weil theorem (see Theorem 12.2.3).

Proposition 11.3.2. *Suppose H is a normal subgroup of G. Then there is an exact sequence*

$$0 \to \mathrm{H}^1(G/H, A^H) \xrightarrow{\mathrm{inf}_H} \mathrm{H}^1(G, A) \xrightarrow{\mathrm{res}_H} \mathrm{H}^1(H, A).$$

Proof. Our proof follows [Ser79, Pg. 117] closely.

We see that $\mathrm{res} \circ \mathrm{inf} = 0$ since on cocycles the composition is defined by precomposing with $H \to G \to G/H$, which gives the trivial map. It remains to prove that inf_H is injective and that the image of inf_H contains the kernel of res_H.

1. *(That inf_H is injective):* Suppose $f : G/H \to A^H$ is a cocycle whose image in $\mathrm{H}^1(G, A)$ is equivalent to 0 modulo coboundaries. Then there is an $a \in A$ such that $f(\sigma) = \sigma a - a$, where we identify f with the map $G \to A$ that is constant on the costs of H. But f depends only on the coset of σ modulo H, so $\sigma a - a = \sigma\tau a - a$ for all $\tau \in H$, i.e., $\tau a = a$ (as we see by adding a to both sides and multiplying by σ^{-1}). Thus $a \in A^H$, so f is equivalent to 0 in $\mathrm{H}^1(G/H, A^H)$.

2. *(The image of inf_H contains the kernel of res_H):* Suppose $f : G \to A$ is a cocycle whose restriction to H is a coboundary, i.e., there is $a \in A$ such that $f(\tau) = \tau a - a$ for all $\tau \in H$. Subtracting the coboundary $g(\sigma) = \sigma a - a$ for $\sigma \in G$ from f, we may assume $f(\tau) = 0$ for all $\tau \in H$. Examing the equation

$f(\sigma\tau) = f(\sigma) + \sigma f(\tau)$ with $\tau \in H$ shows that f is constant on the cosets of H. Again using this formula, but with $\sigma \in H$ and $\tau \in G$, we see that

$$f(\tau) = f(\sigma\tau) = f(\sigma) + \sigma f(\tau) = \sigma f(\tau),$$

so the image of f is contained in A^H. Thus f defines a cocycle $G/H \to A^H$, i.e., is in the image of $\inf_H$.

□

Example 11.3.3. The sequence of Proposition 11.3.2 need not be surjective on the right. For example, suppose $H = A_3 \subset S_3$, and let S_3 act trivially on the group $\mathbb{Z}/3\mathbb{Z}$. Using the Hom interpretation of H^1, we see that $\mathrm{H}^1(S_3/A_3, \mathbb{Z}/3\mathbb{Z}) = \mathrm{H}^1(S_3, \mathbb{Z}/3\mathbb{Z}) = 0$, but $\mathrm{H}^1(A_3, \mathbb{Z}/3\mathbb{Z})$ has order 3. We can compute this example in Sage as follows.

```
S3 = SymmetricGroup(3); S3
```

```
Symmetric group of order 3! as a permutation group
```

```
S3.cohomology(1,3)
```

```
Trivial Abelian group
```

```
A3 = AlternatingGroup(3); A3
```

```
Alternating group of order 3!/2 as a permutation group
```

```
A3.cohomology(1,3)
```

```
Multiplicative Abelian group isomorphic to C3
```

Remark 11.3.4. One generalization of Proposition 11.3.2 is to a more complicated exact sequence involving the "transgression map" tr:

$$0 \to \mathrm{H}^1(G/H, A^H) \xrightarrow{\inf_H} \mathrm{H}^1(G, A) \xrightarrow{\mathrm{res}_H} \mathrm{H}^1(H, A)^{G/H} \xrightarrow{\mathrm{tr}} \mathrm{H}^2(G/H, A^H) \to \mathrm{H}^2(G, A).$$

Another generalization of Proposition 11.3.2 is that if $\mathrm{H}^m(H, A) = 0$ for $1 \leq m < n$, then there is an exact sequence

$$0 \to \mathrm{H}^n(G/H, A^H) \xrightarrow{\inf_H} \mathrm{H}^n(G, A) \xrightarrow{\mathrm{res}_H} \mathrm{H}^n(H, A).$$

For more information see [Ser79, Ch. VII.6].

Remark 11.3.5. If H is a not-necessarily-normal subgroup of G, there are also maps

$$\mathrm{cores}_H : \mathrm{H}^n(H, A) \to \mathrm{H}^n(G, A)$$

for each n. For $n = 0$ this is the trace map $a \mapsto \sum_{\sigma \in G/H} \sigma a$, but the definition for $n \geq 1$ is more involved. One has $\mathrm{cores}_H \circ \mathrm{res}_H = [\#(G/H)]$. Taking $H = 1$ this implies that for each $n \geq 1$ the group $\mathrm{H}^n(G, A)$ is annihilated by $[\#G]$.

11.4 Galois Cohomology

Suppose L/K is a finite Galois extension of fields (recall that Galois here means is normal and separable), and A is a $\mathrm{Gal}(L/K)$-module. Put

$$\mathrm{H}^n(L/K, A) = \mathrm{H}^n(\mathrm{Gal}(L/K), A).$$

Following Section 9.5, we can put a topology on $\mathrm{Gal}(K^{\mathrm{sep}}/K)$ by taking as a basis of the origin, subgroups of the form $\mathrm{Gal}(K^{\mathrm{sep}}/L)$ where L/K is a finite Galois extension.

Exercise 11.4.1. Let H be a subgroup of $G = \mathrm{Gal}(K^{\mathrm{sep}}/K)$. Show that H is open if and only if H is closed and has finite index in G. [*Hint*: If H is open then it contains a basis element N. By definition of the basis described above, N is finite index in G. What does this say about the index of H in G? What about the compliment of H?]

Definition 11.4.2. Let A be a $\mathrm{Gal}(K^{\mathrm{sep}}/K)$-module. We say that A is a *continuous* $\mathrm{Gal}(K^{\mathrm{sep}}/K)$-module if the map $\mathrm{Gal}(K^{\mathrm{sep}}/K) \times A \to A$ (see Exercise 11.1.6) is continuous when A has the discrete topology.

Exercise 11.4.3. Let $G = \mathrm{Gal}(K^{\mathrm{sep}}/K)$ and A be a G-module. Show that A is a continuous G-module if and only if the subgroup $G_a = \{\sigma \in G : \sigma(a) = a\}$ is open for every $a \in A$.

Now let A be a continuous $\mathrm{Gal}(K^{\mathrm{sep}}/K)$-module. Let

$$A(L) = A^{\mathrm{Gal}(K^{\mathrm{sep}}/L)} = \{x \in A : \sigma(x) = x \text{ for all } \sigma \in \mathrm{Gal}(K^{\mathrm{sep}}/L)\}.$$

and define

$$\mathrm{H}^n(K, A) = \varinjlim_{L/K} \mathrm{H}^n(L/K, A(L)),$$

where the limit is taken over all finite Galois extensions L/K.

It is not obvious that the groups $\mathrm{H}^n(K, A)$ are actually cohomology groups, i.e. they satisfy the conclusion of Theorem 11.2.4. However one can show they have analogous properties, see [Ser79, Ch. X.3] for references.

Remark 11.4.4. Those familiar with algebraic geometry should compare the groups $\mathrm{H}^n(K, A)$ with the Čech cohomology groups on the étale site over $\mathrm{Spec}\, K$. One can show that Čech cohomology agrees with the derived functor groups of $A \mapsto A^G$, see [**?**, Ch. 10]. Therefore $\mathrm{H}^n(K, A)$ do indeed define a cohomology theory.

Example 11.4.5. The following are examples of continuous $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$-modules:

$$\overline{\mathbb{Q}}, \quad \overline{\mathbb{Q}}^*, \quad \overline{\mathbb{Z}}, \quad \overline{\mathbb{Z}}^*, \quad E(\overline{\mathbb{Q}}), \quad E(\overline{\mathbb{Q}})[n], \quad \mathrm{Tate}_\ell(E),$$

where E is an elliptic curve over $\mathbb{Q}$. Can you identify the action for each module A? What about $A(L)$ for any finite Galois extension $L/\mathbb{Q}$. It is important to notice $\overline{\mathbb{Q}}^*(L) = L^*$.

Theorem 11.4.6 (Hilbert 90)**.** *We have* $\mathrm{H}^1(K, \overline{K}^*) = 0$.

Proof. Our proof follows [Ser79, Pg. 150] closely.

Because $\mathrm{H}^1(K, \overline{K}^*) = \varinjlim_{L/K} \mathrm{H}^1(L/K, L^*)$ It suffices to prove $\mathrm{H}^1(L/K, L^*) = 0$ for every finite Galois extension L/K. Let $G = \mathrm{Gal}(L/K)$ and f be a 1-cocycle so that $f : G \to L^*$ such that $f(\sigma\tau) = f(\sigma) \cdot \sigma(f(\tau))$. Here " $\cdot$ " represents multiplication in L^*. A standard fact from Galois theory is that the elements of G are L linearly independent. Hence we can find some $c \in L$ such that

$$b = \sum_{\tau \in G} f(\tau) \cdot \tau(c) \neq 0.$$

Now apply σ to both sides to get

$$\begin{aligned} \sigma(b) &= \sum_{\tau \in G} \sigma(f(\tau)) \cdot \sigma\tau(c) \\ &= \sum_{\tau \in G} f(\sigma)^{-1} \cdot f(\sigma\tau) \cdot \sigma\tau(c) \\ &= f(\sigma)^{-1} \cdot \sum_{\tau \in G} f(\sigma\tau) \cdot (\sigma\tau)(c) \\ &= f(\sigma)^{-1} \cdot b. \end{aligned}$$

This shows f is a coboundary. Specifically, it shows $f = f_{b^{-1}}$ in the notation we used to define coboundaries above. $\square$

Chapter 12

The Weak Mordell-Weil Theorem

12.1 Kummer Theory of Number Fields

Suppose K is a number field and fix a positive integer n. Let μ_n denote the nth roots of unity in $\overline{K}$ as a group under multiplication. Consider the exact sequence

$$1 \to \mu_n \to \overline{K}^* \xrightarrow{n} \overline{K}^* \to 1$$

where n denotes the map $a \mapsto a^n$.

The corresponding long exact sequence from Theorem 11.2.4 is

$$1 \to \mu_n(K) \to K^* \xrightarrow{n} K^* \to \mathrm{H}^1(K, \mu_n) \to \mathrm{H}^1(K, \overline{K}^*) = 0,$$

where $\mu_n(K)$ is the nth roots of unity contained in K. The last equality follows from Theorem 11.4.6.

Assume now that the group μ_n is contained in K. Using Galois cohomology we obtain a relatively simple classification of all abelian extensions of K with cyclic Galois group of order dividing n. Moreover, since the action of $\mathrm{Gal}(\overline{K}/K)$ on μ_n is trivial, by our hypothesis that $\mu_n \subset K$, Exercise 11.2.1 implies

$$\mathrm{H}^1(K, \mu_n) = \mathrm{Hom}(\mathrm{Gal}(\overline{K}/K), \mu_n).$$

Thus we obtain an exact sequence

$$1 \to \mu_n \to K^* \xrightarrow{n} K^* \to \mathrm{Hom}(\mathrm{Gal}(\overline{K}/K), \mu_n) \to 1,$$

or equivalently, an isomorphism

$$K^*/(K^*)^n \cong \mathrm{Hom}(\mathrm{Gal}(\overline{K}/K), \mu_n).$$

By Galois theory, homomorphisms $\mathrm{Gal}(\overline{K}/K) \to \mu_n$ (up to automorphisms of μ_n) correspond to cyclic abelian extensions of K with Galois group a subgroup of the cyclic group μ_n. Unwinding the definitions, this says that every cyclic abelian extension of K of degree dividing n is of the form $K(a^{1/n})$ for some element $a \in K$.

One can prove via calculations that $K(a^{1/n})$ is unramified outside n and the primes that divide $\mathrm{Norm}(a)$. Moreover, and this is a much bigger result, one can combine this with facts about class groups and unit groups to prove the following theorem:

Theorem 12.1.1. *Suppose K is a number field with $\mu_n \subset K$, where n is a positive integer. Let L be the maximal extension of K such that*

(i) $\mathrm{Gal}(L/K)$ *is abelian,*

(ii) $n \cdot \mathrm{Gal}(L/K) = 0$, *and*

(iii) *L is unramified outside a finite set S of primes.*

Then L/K is of finite degree.

Sketch of Proof. Note that we may enlarge S as needed. To see this, choose a finite set $S' \supseteq S$ and let L' the maximal extension with respect to S' as in the statement of the theorem. Because L is unramified outside of S, it is certainly unramified outside of S'. By maximality of L' this implies $L' \subseteq L$. Therefore it's sufficient to show the larger extension L'/K is finite.

We first argue that we can enlarge S so that the ring

$$\mathcal{O}_{K,S} = \{a \in K^* \colon \mathrm{ord}_{\mathfrak{p}}(a\mathcal{O}_K) \geq 0 \text{ for all } \mathfrak{p} \notin S\} \cup \{0\}$$

is a principal ideal domain. One can show that for any S, the ring $\mathcal{O}_{K,S}$ is a Dedekind domain. The condition $\mathrm{ord}_{\mathfrak{p}}(a\mathcal{O}_K) \geq 0$ means that in the prime ideal factorization of the fractional ideal $a\mathcal{O}_K$, we have that $\mathfrak{p}$ occurs to a nonnegative power. Thus we are allowing denominators at the primes in S. Since the class group of $\mathcal{O}_K$ is finite, there are primes $\mathfrak{p}_1, \ldots, \mathfrak{p}_r$ that generate the class group as a group (for example, take all primes with norm up to the Minkowski bound). Enlarge S to contain the primes $\mathfrak{p}_i$.

Note that we have used *the class group of $\mathcal{O}_K$ is finite.*

Next we want to show $\mathfrak{p}_i\mathcal{O}_{K,S}$ is the unit ideal. To see this, let m be the order of $\mathfrak{p}_i$ in the class group of $\mathcal{O}_K$ so that $\mathfrak{p}_i^m = (\alpha)$ for some $\alpha \in \mathcal{O}_K$. Note the factorization of $\frac{1}{\alpha}\mathcal{O}_K$ is $\mathfrak{p}_i^{-m}$ so by construction $\frac{1}{\alpha} \in \mathcal{O}_{K,S}$. Since $\alpha \in (\mathfrak{p}_i\mathcal{O}_{K,S})^m$ this shows $(\mathfrak{p}_i\mathcal{O}_{K,S})^m$ is the unit ideal. It follows from the unique factorization of ideals in the Dedekind domain $\mathcal{O}_{K,S}$ that $\mathfrak{p}_i\mathcal{O}_{K,S}$ is the unit ideal.

Now we can show $\mathcal{O}_{K,S}$ is a principal ideal domain. Let $\mathfrak{P}$ be a prime ideal of $\mathcal{O}_{K,S}$. Since the $\mathfrak{p}_i$ generate the class group of $\mathcal{O}_K$, the restriction of $\mathfrak{P}$ to $\mathcal{O}_K$ is equivalent modulo a principal ideal to a product of the primes $\mathfrak{p}_i$. Therefore $\mathfrak{P}$ is equivalent modulo a principal ideal to a product of ideals of the form $\mathfrak{p}_i\mathcal{O}_{K,S}$. Because we showed $\mathfrak{p}_i\mathcal{O}_{K,S}$ was the unit ideal, this means $\mathfrak{P}$ is principal.

Next enlarge S so that all primes over $n\mathcal{O}_K$ are in S. Note that $\mathcal{O}_{K,S}$ is still a PID. Let

$$K(S,n) = \{a \in K^*/(K^*)^n \colon n \mid \mathrm{ord}_{\mathfrak{p}}(a) \text{ for all } \mathfrak{p} \notin S\}.$$

Then a refinement of the arguments at the beginning of this section show that L is generated by all nth roots of the elements of $K(S,n)$ (specifically, their representatives in K). Thus it suffices to prove that $K(S,n)$ is finite.

If $a \in \mathcal{O}_{K,S}^*$ then $\mathrm{ord}_{\mathfrak{p}}(a) = 0$ for all $\mathfrak{p} \notin S$. So there is a natural map

$$\phi : \mathcal{O}_{K,S}^* \to K(S,n)$$

sending a to it's residue class in $K^*/(K^*)^n$. Suppose $a \in K^*$ is a representative of an element in $K(S,n)$. The ideal $a\mathcal{O}_{K,S}$ has a factorization which is a product of nth powers, so it is an nth power of an ideal. Since $\mathcal{O}_{K,S}$ is a PID, there is $b \in \mathcal{O}_{K,S}$ and $u \in \mathcal{O}_{K,S}^*$ such that

$$a = b^n \cdot u.$$

Thus $u \in \mathcal{O}_{K,S}^*$ maps to $[a] \in K(S,n)$. This shows ϕ is surjective.

Recall *Dirichlet's unit theorem* (Theorem 8.1.2), which asserts that the group $\mathcal{O}_K^*$ is a finitely generated abelian group of rank $r+s-1$. More generally, we now show that $\mathcal{O}_{K,S}^*$ is a finitely generated abelian group of rank $r+s+\#S-1$. This result with above would imply $K(S,n)$ is a torsion group which is a quotient of a finitely generated group and hence finite, thus proving the theorem.

The fact that $\mathcal{O}_{K,S}^*$ has rank $r+s-1+\#S$ is sometimes referred to as the *S-unit theorem* or the *Dirichlet S-unit theorem*. To prove this theorem, let $\mathfrak{p}_1,\ldots,\mathfrak{p}_m$ be the primes in S and define a map $\phi : \mathcal{O}_{K,S}^* \to \mathbb{Z}^m$ by

$$\phi(u) = (\mathrm{ord}_{\mathfrak{p}_1}(u),\ldots,\mathrm{ord}_{\mathfrak{p}_m}(u)).$$

First we show that $\mathrm{Ker}(\phi) = \mathcal{O}_K^*$. We have that $u \in \mathrm{Ker}(\phi)$ if and only if $u \in \mathcal{O}_{K,S}^*$ and $\mathrm{ord}_{\mathfrak{p}_i}(u) = 0$ for all i; but the latter condition implies that u is a unit at each prime in S. But $u \in \mathcal{O}_{K,S}^*$ implies $\mathrm{ord}_{\mathfrak{p}}(u) = 0$ for all $\mathfrak{p} \notin S$, so it follows that $\mathrm{ord}_{\mathfrak{p}}(u) = 0$ for all primes $\mathfrak{p}$ in $\mathcal{O}_K$ and therefore $u \in \mathcal{O}_K^*$. Thus we have an exact sequence

$$1 \to \mathcal{O}_K^* \to \mathcal{O}_{K,S}^* \xrightarrow{\phi} \mathbb{Z}^m.$$

Next we show that the image of ϕ has finite index in $\mathbb{Z}^m$. Let h be the class number of $\mathcal{O}_K$. For each i there exists $\alpha_i \in \mathcal{O}_K$ such that $\mathfrak{p}_i^h = (\alpha_i)$. But $\alpha_i \in \mathcal{O}_{K,S}^*$ since $\mathrm{ord}_{\mathfrak{p}}(\alpha_i) = 0$ for all $\mathfrak{p} \notin S$ (by unique factorization). Then

$$\phi(\alpha_i) = (0,\ldots,0,h,0,\ldots,0).$$

It follows that $(h\mathbb{Z})^m \subset \mathrm{Im}(\phi)$, so the image of ϕ has finite index in $\mathbb{Z}^m$. It follows that $\mathcal{O}_{K,S}^*$ has rank equal to $r+s-1+\#S$. $\square$

12.2 Proof of the Weak Mordell-Weil Theorem

Suppose E is an elliptic curve over a number field K, and fix a positive integer n. Just as with number fields, we have an exact sequence

$$0 \to E[n] \to E \xrightarrow{n} E \to 0.$$

Then we have an exact sequence

$$0 \to E[n](K) \to E(K) \xrightarrow{n} E(K) \to \mathrm{H}^1(K, E[n]) \to \mathrm{H}^1(K, E)[n] \to 0.$$

Note the last term comes from replacing the codomain of $\mathrm{H}^1(K, E[n]) \to \mathrm{H}^1(K, E)$ by the kernel of $\mathrm{H}^1(K, E) \xrightarrow{n} \mathrm{H}^1(K, E)$. From this we obtain a short exact sequence

$$0 \to E(K)/nE(K) \to \mathrm{H}^1(K, E[n]) \to \mathrm{H}^1(K, E)[n] \to 0. \tag{12.2.1}$$

Now assume, in analogy with Section 12.1, that $E[n] \subset E(K)$, i.e., all n-torsion points are defined over K. Then the Galois action on $E[n]$ is trivial so by exercise 11.2.1 we have

$$\mathrm{H}^1(K, E[n]) = \mathrm{Hom}(\mathrm{Gal}(\overline{K}/K), E[n]) \cong \mathrm{Hom}(\mathrm{Gal}(\overline{K}/K), (\mathbb{Z}/n\mathbb{Z})^2),$$

and the sequence (12.2.1) induces an inclusion

$$E(K)/nE(K) \hookrightarrow \mathrm{Hom}(\mathrm{Gal}(\overline{K}/K), (\mathbb{Z}/n\mathbb{Z})^2). \tag{12.2.2}$$

Explicitly, this homomorphism sends a point P to the homomorphism defined as follows: Choose $Q \in E(\overline{K})$ such that $nQ = P$; then send each $\sigma \in \mathrm{Gal}(\overline{K}/K)$ to $\sigma(Q) - Q \in E[n]$.

Exercise 12.2.1. Consider the map $E(K) \to \mathrm{Hom}(\mathrm{Gal}(\overline{K}/K), E[n])$ defined above. First show this map is well defined, i.e., $\sigma(Q) - Q \in E[n]$. Then show it does not depend on the choice of P modulo $nE(K)$ so it in fact descends to a homomorphism on $E(K)/nE(K)$.

Because $E[n] \cong (\mathbb{Z}/n\mathbb{Z})^2$, given a point $P \in E(K)$, we obtain a homomorphism $\varphi : \mathrm{Gal}(\overline{K}/K) \to (\mathbb{Z}/n\mathbb{Z})^2$, whose kernel defines an abelian extension L of K that has exponent n. The amazing fact is that L can be ramified only at the primes of bad reduction for E and the primes that divide n. Thus we can apply theorem 12.1.1 to see that there are only finitely many such L.

Theorem 12.2.2. *If $P \in E(K)$ is a point, then the field L obtained by adjoining to K all coordinates of all choices of $Q = \frac{1}{n}P$ is unramified outside n and the primes of bad reduction for E.*

Sketch of Proof. First one proves that if $\mathfrak{p} \nmid n$ is a prime of good reduction for E, then the natural reduction map $\pi : E(K)[n] \to \tilde{E}(\mathcal{O}_K/\mathfrak{p})$ is injective. The argument that π is injective uses *formal groups*, whose development is outside the scope of this course.[1] Next fix any Q as in the statement of the theorem. Above we saw $\sigma(Q) - Q \in E[n]$ for all $\sigma \in \mathrm{Gal}(\overline{K}/K)$. Let $I_{\mathfrak{p}} \subset \mathrm{Gal}(L/K)$ be the inertia group at $\mathfrak{p}$. By definition of interia group, $I_{\mathfrak{p}}$ acts trivially on $\tilde{E}(\mathcal{O}_K/\mathfrak{p})$. Thus for each $\sigma \in I_{\mathfrak{p}}$ we have

$$\pi(\sigma(Q) - Q) = \sigma(\pi(Q)) - \pi(Q) = \pi(Q) - \pi(Q) = 0.$$

Since π is injective, it follows that $\sigma(Q) = Q$ for $\sigma \in I_{\mathfrak{p}}$, i.e., that Q is fixed under all $I_{\mathfrak{p}}$. Repeating for all Q this shows $I_{\mathfrak{p}} = 1$ so L is unramified at $\mathfrak{p}$.

Note that we technically only defined π on $E(K)[n]$ and $\sigma(Q) - Q$ may not lie in $E(K)$. However, given a finite extension K'/K and prime $\mathfrak{p}'$ lying over $\mathfrak{p}$, E the reduction map $E(K') \to \tilde{E}(\mathcal{O}_{K'}/\mathfrak{p}')$ is still injective. So we could apply the argument to the field K' given by adjoining the coordinates of Q to K. $\square$

Theorem 12.2.3 (Weak Mordell-Weil)**.** *Let E be an elliptic curve over a number field K, and let n be any positive integer. Then $E(K)/nE(K)$ is finitely generated.*

Proof. First suppose all elements of $E[n]$ have coordinates in K. Then the homomorphism (12.2.2) provides an injection of $E(K)/nE(K)$ into

$$\mathrm{Hom}(\mathrm{Gal}(\overline{K}/K), (\mathbb{Z}/n\mathbb{Z})^2).$$

By Theorem 12.2.2, the image consists of homomorphisms whose kernels cut out an abelian extension of K unramified outside n and primes of bad reduction for E. Since this is a finite set of primes, Theorem 12.1.1 implies that the homomorphisms all factor through a finite quotient $\mathrm{Gal}(L/K)$ of $\mathrm{Gal}(\overline{\mathbb{Q}}/K)$. Thus there can be only finitely many such homomorphisms, so the image of $E(K)/nE(K)$ is finite. Thus $E(K)/nE(K)$ itself is finite, which proves the theorem in this case.

Next suppose E is an elliptic curve over a number field, but do *not* make the hypothesis that the elements of $E[n]$ have coordinates in K. Since the group $E[n](\mathbb{C})$ is finite and its elements are defined over $\overline{\mathbb{Q}}$, the extension L of K got by adjoining to K all coordinates of elements of $E[n](\mathbb{C})$ is a finite extension. It is also Galois, as we saw when constructing Galois representations attached to elliptic curves. By Proposition 11.3.2, we have an exact sequence

$$0 \to \mathrm{H}^1(L/K, E[n](L)) \to \mathrm{H}^1(K, E[n]) \to \mathrm{H}^1(L, E[n]).$$

The kernel of the restriction map $\mathrm{H}^1(K, E[n]) \to \mathrm{H}^1(L, E[n])$ is finite, since it is isomorphic to the finite cohomology group $\mathrm{H}^1(L/K, E[n](L))$. By the argument of the previous paragraph, the image of $E(K)/nE(K)$ in $\mathrm{H}^1(L, E[n])$ under

$$E(K)/nE(K) \hookrightarrow \mathrm{H}^1(K, E[n]) \xrightarrow{\mathrm{res}} \mathrm{H}^1(L, E[n])$$

is finite, since it is contained in the image of $E(L)/nE(L)$. Thus $E(K)/nE(K)$ is finite, since we just proved the kernel of res is finite. $\square$

[1] For an introduction to formal groups of elliptic curves see [Sil92, Ch. IV].

Part II

Adelic Viewpoint

Chapter 13

Valuations

The rest of this book is a partial rewrite of [Cas67] meant to make it more accessible. I have attempted to add examples and details of the implicit exercises and remarks that are left to the reader.

13.1 Valuations

Definition 13.1.1 (Valuation)**.** A *valuation* $|\cdot|$ on a field K is a function defined on K with values in $\mathbb{R}_{\geq 0}$ satisfying the following axioms:

(1) $|a| = 0$ if and only if $a = 0$,

(2) $|ab| = |a|\,|b|$, and

(3) there is a constant $C \geq 1$ such that $|1 + a| \leq C$ whenever $|a| \leq 1$.

The *trivial valuation* is the valuation for which $|a| = 1$ for all $a \neq 0$. We will often tacitly exclude the trivial valuation from consideration.

From (2) we have

$$|1| = |1| \cdot |1|,$$

so $|1| = 1$ by (1). If $w \in K$ and $w^n = 1$, then $|w| = 1$ by (2). In particular, the only valuation of a finite field is the trivial one. The same argument shows that $|-1| = |1|$, so

$$|-a| = |a| \qquad \text{all } a \in K.$$

Definition 13.1.2 (Equivalent)**.** Two valuations $|\cdot|_1$ and $|\cdot|_2$ on the same field are *equivalent* if there exists $c > 0$ such that

$$|a|_2 = |a|_1^c$$

for all $a \in K$.

Note that if $|\cdot|_1$ is a valuation, then $|\cdot|_2 = |\cdot|_1^c$ is also a valuation. Also, equivalence of valuations is an equivalence relation.

If $|\cdot|$ is a valuation and $C > 1$ is the constant from Axiom (3), then there is a $c > 0$ such that $C^c = 2$ (i.e., $c = \log(2)/\log(C)$). Then we can take 2 as constant for the equivalent valuation $|\cdot|^c$. Thus every valuation is equivalent to a valuation with $C = 2$. Note that if $C = 1$, e.g., if $|\cdot|$ is the trivial valuation, then we could simply take $C = 2$ in Axiom (3).

Proposition 13.1.3. *Suppose $|\cdot|$ is a valuation with $C = 2$. Then for all $a, b \in K$ we have*

$$|a + b| \leq |a| + |b| \qquad \textit{(triangle inequality).} \tag{13.1.1}$$

Proof. Suppose $a_1, a_2 \in K$ with $|a_1| \geq |a_2|$. Then $a = a_2/a_1$ satisfies $|a| \leq 1$. By Axiom (3) we have $|1 + a| \leq 2$, so multiplying by a_1 we see that

$$|a_1 + a_2| \leq 2|a_1| = 2 \cdot \max\{|a_1|, |a_2|\}.$$

Also we have

$$|a_1 + a_2 + a_3 + a_4| \leq 2 \cdot \max\{|a_1 + a_2|, |a_3 + a_4|\} \leq 4 \cdot \max\{|a_1|, |a_2|, |a_3|, |a_4|\},$$

and inductively we have for any $r > 0$ that

$$|a_1 + a_2 + \cdots + a_{2^r}| \leq 2^r \cdot \max |a_j|.$$

If n is any positive integer, let r be such that $2^{r-1} \leq n \leq 2^r$. Thenn

$$|a_1 + a_2 + \cdots + a_n| \leq 2^r \cdot \max\{|a_j|\} \leq 2n \cdot \max\{|a_j|\},$$

since $2^r \leq 2n$. In particular,

$$|n| \leq 2n \cdot |1| = 2n \qquad (\text{for } n > 0). \tag{13.1.2}$$

Applying (13.1.2) to $\left|\binom{n}{j}\right|$ and using the binomial expansion, we have for any $a, b \in K$ that

$$\begin{aligned}
|a + b|^n &= \left|\sum_{j=0}^{n} \binom{n}{j} a^j b^{n-j}\right| \\
&\leq 2(n+1) \max_j \left\{ \left|\binom{n}{j}\right| |a|^j \, |b|^{n-j} \right\} \\
&\leq 2(n+1) \max_j \left\{ 2\binom{n}{j} |a|^j \, |b|^{n-j} \right\} \\
&\leq 4(n+1) \max_j \left\{ \binom{n}{j} |a|^j \, |b|^{n-j} \right\} \\
&\leq 4(n+1)(|a| + |b|)^n.
\end{aligned}$$

Now take nth roots of both sides to obtain

$$|a+b| \leq \sqrt[n]{4(n+1)} \cdot (|a|+|b|).$$

We have by elementary calculus that

$$\lim_{n\to\infty} \sqrt[n]{4(n+1)} = 1,$$

so $|a+b| \leq |a|+|b|$. (The "elementary calculus": We instead prove that $\sqrt[n]{n} \to 1$, since the argument is the same and the notation is simpler. First, for any $n \geq 1$ we have $\sqrt[n]{n} \geq 1$, since upon taking nth powers this is equivalent to $n \geq 1^n$, which is true by hypothesis. Second, suppose there is an $\varepsilon > 0$ such that $\sqrt[n]{n} \geq 1+\varepsilon$ for all $n \geq 1$. Then taking logs of boths sides we see that $\frac{1}{n}\log(n) \geq \log(1+\varepsilon) > 0$. But $\log(n)/n \to 0$, so there is no such ε. Thus $\sqrt[n]{n} \to 1$ as $n \to \infty$.) □

Note that Axioms (1), (2) and Equation (13.1.1) imply Axiom (3) with $C = 2$. We take Axiom (3) instead of Equation (13.1.1) for the technical reason that we will want to call the square of the absolute value of the complex numbers a valuation.

Lemma 13.1.4. *Suppose $a, b \in K$, and $|\cdot|$ is a valuation on K with $C \leq 2$. Then*

$$\Big||a|-|b|\Big| \leq |a-b|\,.$$

(Here the big absolute value on the outside of the left-hand side of the inequality is the usual absolute value on real numbers, but the other absolute values are a valuation on an arbitrary field K.)

Proof. We have

$$|a| = |b+(a-b)| \leq |b|+|a-b|,$$

so $|a|-|b| \leq |a-b|$. The same argument with a and b swapped implies that $|b|-|a| \leq |a-b|$, which proves the lemma. □

13.2 Types of Valuations

We define two important properties of valuations, both of which apply to equivalence classes of valuations (i.e., the property holds for $|\cdot|$ if and only if it holds for a valuation equivalent to $|\cdot|$).

Definition 13.2.1 (Discrete)**.** A valuation $|\cdot|$ is *discrete* if there is a $\delta > 0$ such that for any $a \in K$

$$1-\delta < |a| < 1+\delta \implies |a| = 1.$$

Thus the absolute values are bounded away from 1.

To say that $|\cdot|$ is discrete is the same as saying that the set

$$G = \left\{\log|a| : a \in K, a \neq 0\right\} \subset \mathbb{R}$$

forms a discrete subgroup of the reals under addition (because the elements of the group G are bounded away from 0).

Proposition 13.2.2. *A nonzero discrete subgroup G of $\mathbb{R}$ is free on one generator.*

Proof. Since G is discrete there is a positive $m \in G$ such that for any positive $x \in G$ we have $m \leq x$. Suppose $x \in G$ is an arbitrary positive element. By subtracting off integer multiples of m, we find that there is a unique n such that

$$0 \leq x - nm < m.$$

Since $x - nm \in G$ and $0 < x - nm < m$, it follows that $x - nm = 0$, so x is a multiple of m. $\square$

By Proposition 13.2.2, the set of $\log|a|$ for nonzero $a \in K$ is free on one generator, so there is a $c < 1$ such that $|a|$, for $a \neq 0$, runs precisely through the set

$$c^{\mathbb{Z}} = \{c^m : m \in \mathbb{Z}\}$$

(Note: we can replace c by c^{-1} to see that we can assume that $c < 1$).

Definition 13.2.3 (Order)**.** If $|a| = c^m$, we call $m = \mathrm{ord}(a)$ the *order* of a.

Axiom (2) of valuations translates into

$$\mathrm{ord}(ab) = \mathrm{ord}(a) + \mathrm{ord}(b).$$

Definition 13.2.4 (Non-archimedean)**.** A valuation $|\cdot|$ is *non-archimedean* if we can take $C = 1$ in Axiom (3), i.e., if

$$|a+b| \leq \max\left\{|a|, |b|\right\}. \tag{13.2.1}$$

If $|\cdot|$ is not non-archimedean then it is *archimedean.*

Note that if we can take $C = 1$ for $|\cdot|$ then we can take $C = 1$ for any valuation equivalent to $|\cdot|$. To see that (13.2.1) is equivalent to Axiom (3) with $C = 1$, suppose $|b| \leq |a|$. Then $|b/a| \leq 1$, so Axiom (3) asserts that $|1 + b/a| \leq 1$, which implies that $|a + b| \leq |a| = \max\{|a|, |b|\}$, and conversely.

We note at once the following consequence:

Lemma 13.2.5. *Suppose $|\cdot|$ is a non-archimedean valuation. If $a, b \in K$ with $|b| < |a|$, then $|a + b| = |a|$.*

Proof. Note that $|a+b| \le \max\{|a|,|b|\} = |a|$, which is true even if $|b| = |a|$. Also,

$$|a| = |(a+b) - b| \le \max\{|a+b|, |b|\} = |a+b|,$$

where for the last equality we have used that $|b| < |a|$ (if $\max\{|a+b|,|b|\} = |b|$, then $|a| \le |b|$, a contradiction).

□

Definition 13.2.6 (Ring of Integers)**.** Suppose $|\cdot|$ is a non-archimedean absolute value on a field K. Then

$$\mathcal{O} = \{a \in K : |a| \le 1\}$$

is a ring called the *ring of integers* of K with respect to $|\cdot|$.

Lemma 13.2.7. *Two non-archimedean valuations $|\cdot|_1$ and $|\cdot|_2$ are equivalent if and only if they give the same $\mathcal{O}$.*

We will prove this modulo the claim (to be proved later in Section 14.1) that valuations are equivalent if (and only if) they induce the same topology.

Proof. Suppose suppose $|\cdot|_1$ is equivalent to $|\cdot|_2$, so $|\cdot|_1 = |\cdot|_2^c$, for some $c > 0$. Then $|c|_1 \le 1$ if and only if $|c|_2^c \le 1$, i.e., if $|c|_2 \le 1^{1/c} = 1$. Thus $\mathcal{O}_1 = \mathcal{O}_2$.

Conversely, suppose $\mathcal{O}_1 = \mathcal{O}_2$. Then $|a|_1 < |b|_1$ if and only if $a/b \in \mathcal{O}_1$ and $b/a \notin \mathcal{O}_1$, so

$$|a|_1 < |b|_1 \iff |a|_2 < |b|_2. \tag{13.2.2}$$

The topology induced by $|\ |_1$ has as basis of open neighborhoods the set of open balls

$$B_1(z,r) = \{x \in K : |x - z|_1 < r\},$$

for $r > 0$, and likewise for $|\ |_2$. Since the absolute values $|b|_1$ get arbitrarily close to 0, the set $\mathcal{U}$ of open balls $B_1(z, |b|_1)$ also forms a basis of the topology induced by $|\ |_1$ (and similarly for $|\ |_2$). By (13.2.2) we have

$$B_1(z, |b|_1) = B_2(z, |b|_2),$$

so the two topologies both have $\mathcal{U}$ as a basis, hence are equal. That equal topologies imply equivalence of the corresponding valuations will be proved in Section 14.1. □

The set of $a \in \mathcal{O}$ with $|a| < 1$ forms an ideal $\mathfrak{p}$ in $\mathcal{O}$. The ideal $\mathfrak{p}$ is maximal, since if $a \in \mathcal{O}$ and $a \notin \mathfrak{p}$ then $|a| = 1$, so $|1/a| = 1/|a| = 1$, hence $1/a \in \mathcal{O}$, so a is a unit.

Lemma 13.2.8. *A non-archimedean valuation $|\cdot|$ is discrete if and only if $\mathfrak{p}$ is a principal ideal.*

Proof. First suppose that $|\cdot|$ is discrete. Choose $\pi \in \mathfrak{p}$ with $|\pi|$ maximal, which we can do since

$$S = \{\log|a| : a \in \mathfrak{p}\} \subset (-\infty, 1],$$

so the discrete set S is bounded above. Suppose $a \in \mathfrak{p}$. Then

$$\left|\frac{a}{\pi}\right| = \frac{|a|}{|\pi|} \leq 1,$$

so $a/\pi \in \mathcal{O}$. Thus

$$a = \pi \cdot \frac{a}{\pi} \in \pi\mathcal{O}.$$

Conversely, suppose $\mathfrak{p} = (\pi)$ is principal. For any $a \in \mathfrak{p}$ we have $a = \pi b$ with $b \in \mathcal{O}$. Thus

$$|a| = |\pi| \cdot |b| \leq |\pi| < 1.$$

Thus $\{|a| : |a| < 1\}$ is bounded away from 1, which is exactly the definition of discrete. $\square$

Example 13.2.9. For any prime p, define the p-adic valuation $|\cdot|_p : \mathbb{Q} \to \mathbb{R}$ as follows. Write a nonzero $\alpha \in K$ as $p^n \cdot \frac{a}{b}$, where $\gcd(a,p) = \gcd(b,p) = 1$. Then

$$\left|p^n \cdot \frac{a}{b}\right|_p := p^{-n} = \left(\frac{1}{p}\right)^n.$$

This valuation is both discrete and non-archimedean. The ring $\mathcal{O}$ is the local ring

$$\mathbb{Z}_{(p)} = \left\{\frac{a}{b} \in \mathbb{Q} : p \nmid b\right\},$$

which has maximal ideal generated by p. Note that $\mathrm{ord}(p^n \cdot \frac{a}{b}) = p^n$.

We will using the following lemma later (e.g., in the proof of Corollary 14.2.4 and Theorem 13.3.2).

Lemma 13.2.10. *A valuation $|\cdot|$ is non-archimedean if and only if $|n| \leq 1$ for all n in the ring generated by 1 in K.*

Note that we cannot identify the ring generated by 1 with $\mathbb{Z}$ in general, because K might have characteristic $p > 0$.

Proof. If $|\cdot|$ is non-archimedean, then $|1| \leq 1$, so by Axiom (3) with $a = 1$, we have $|1+1| \leq 1$. By induction it follows that $|n| \leq 1$.

Conversely, suppose $|n| \leq 1$ for all integer multiples n of 1. This condition is also true if we replace $|\cdot|$ by any equivalent valuation, so replace $|\cdot|$ by one with $C \leq 2$, so that the triangle inequality holds. Suppose $a \in K$ with $|a| \leq 1$. Then by the triangle inequality,

$$\begin{aligned}|1+a|^n &= |(1+a)^n| \\ &\leq \sum_{j=0}^{n} \left|\binom{n}{j}\right| |a| \\ &\leq 1 + 1 + \cdots + 1 = n.\end{aligned}$$

Now take nth roots of both sides to get

$$|1+a| \leq \sqrt[n]{n},$$

and take the limit as $n \to \infty$ to see that $|1+a| \leq 1$. This proves that one can take $C = 1$ in Axiom (3), hence that $|\cdot|$ is non-archimedean. $\square$

13.3 Examples of Valuations

The archetypal example of an archimedean valuation is the absolute value on the complex numbers. It is essentially the only one:

Theorem 13.3.1 (Gelfand-Tornheim). *Any field K with an archimedean valuation is isomorphic to a subfield of $\mathbb{C}$, the valuation being equivalent to that induced by the usual absolute value on $\mathbb{C}$.*

We do not prove this here as we do not need it. For a proof, see [Art59, pg. 45, 67].

There are many non-archimedean valuations. On the rationals $\mathbb{Q}$ there is one for every prime $p > 0$, the p-adic valuation, as in Example 13.2.9.

Theorem 13.3.2 (Ostrowski). *The nontrivial valuations on $\mathbb{Q}$ are those equivalent to $|\cdot|_p$, for some prime p, and the usual absolute value $|\cdot|_\infty$.*

Remark 13.3.3. Before giving the proof, we pause with a brief remark about Ostrowski. According to

`http://www-gap.dcs.st-and.ac.uk/~history/Mathematicians/Ostrowski.html`

Ostrowski was a Ukrainian mathematician who lived 1893–1986. Gautschi writes about Ostrowski as follows: "... you are able, on the one hand, to emphasise the abstract and axiomatic side of mathematics, as for example in your theory of general norms, or, on the other hand, to concentrate on the concrete and constructive aspects of mathematics, as in your study of numerical methods, and to do both with equal ease. *You delight in finding short and succinct proofs, of which you have given many examples* ..." [italics mine]

We will now give an example of one of these short and succinct proofs.

Proof. Suppose $|\cdot|$ is a nontrivial valuation on $\mathbb{Q}$.

Nonarchimedean case: Suppose $|c| \leq 1$ for all $c \in \mathbb{Z}$, so by Lemma 13.2.10, $|\cdot|$ is nonarchimedean. Since $|\cdot|$ is nontrivial, the set

$$\mathfrak{p} = \{a \in \mathbb{Z} : |a| < 1\}$$

is nonzero. Also $\mathfrak{p}$ is an ideal and if $|ab| < 1$, then $|a|\,|b| = |ab| < 1$, so $|a| < 1$ or $|b| < 1$, so $\mathfrak{p}$ is a prime ideal of $\mathbb{Z}$. Thus $\mathfrak{p} = p\mathbb{Z}$, for some prime number p. Since every element of $\mathbb{Z}$ has valuation at most 1, if $u \in \mathbb{Z}$ with $\gcd(u,p) = 1$, then $u \notin \mathfrak{p}$, so $|u| = 1$. Let $\alpha = \log_{|p|} \frac{1}{p}$, so $|p|^\alpha = \frac{1}{p}$. Then for any r and any $u \in \mathbb{Z}$ with $\gcd(u,p) = 1$, we have

$$|up^r|^\alpha = |u|^\alpha\, |p|^{\alpha r} = |p|^{\alpha r} = p^{-r} = |up^r|_p\,.$$

Thus $|\cdot|^\alpha = |\cdot|_p$ on $\mathbb{Z}$, hence on $\mathbb{Q}$ by multiplicativity, so $|\cdot|$ is equivalent to $|\cdot|_p$, as claimed.

Archimedean case: By replacing $|\cdot|$ by a power of $|\cdot|$, we may assume without loss that $|\cdot|$ satisfies the triangle inequality. We first make some general remarks about any valuation that satisfies the triangle inequality. Suppose $a \in \mathbb{Z}$ is greater than 1. Consider, for any $b \in \mathbb{Z}$ the base-a expansion of b:

$$b = b_m a^m + b_{m-1}a^{m-1} + \cdots + b_0,$$

where

$$0 \le b_j < a \qquad (0 \le j \le m),$$

and $b_m \neq 0$. Since $a^m \le b$, taking logs we see that $m\log(a) \le \log(b)$, so

$$m \le \frac{\log(b)}{\log(a)}.$$

Let $M = \max\limits_{1 \le d < a} |d|$. Then by the triangle inequality for $|\cdot|$, we have

$$\begin{aligned}
|b| &\le |b_m|\, a^m + \cdots + |b_1|\,|a| + |b_0| \\
&\le M \cdot (|a|^m + \cdots + |a| + 1) \\
&\le M \cdot (m+1) \cdot \max(1, |a|^m) \\
&\le M \cdot \left(\frac{\log(b)}{\log(a)} + 1\right) \cdot \max\left(1, |a|^{\log(b)/\log(a)}\right),
\end{aligned}$$

where in the last step we use that $m \le \frac{\log(b)}{\log(a)}$. Setting $b = c^n$, for $c \in \mathbb{Z}$, in the above inequality and taking nth roots, we have

$$\begin{aligned}
|c| &\le \left(M \cdot \left(\frac{\log(c^n)}{\log(a)} + 1\right) \cdot \max(1, |a|^{log(c^n)/\log(a)})\right)^{1/n} \\
&= M^{1/n} \cdot \left(\frac{\log(c^n)}{\log(a)} + 1\right)^{1/n} \cdot \max\left(1, |a|^{\log(c^n)/\log(a)}\right)^{1/n}.
\end{aligned}$$

The first factor $M^{1/n}$ converges to 1 as $n \to \infty$, since $M \ge 1$ (because $|1| = 1$). The second factor is

$$\left(\frac{\log(c^n)}{\log(a)} + 1\right)^{1/n} = \left(n \cdot \frac{\log(c)}{\log(a)} + 1\right)^{1/n}$$

which also converges to 1, for the same reason that $n^{1/n} \to 1$ (because $\log(n^{1/n}) = \frac{1}{n}\log(n) \to 0$ as $n \to \infty$). The third factor is

$$\max\left(1, |a|^{\log(c^n)/\log(a)}\right)^{1/n} = \begin{cases} 1 & \text{if } |a| < 1, \\ |a|^{\log(c)/\log(a)} & \text{if } |a| \ge 1. \end{cases}$$

Putting this all together, we see that

$$|c| \leq \max\left(1, |a|^{\frac{\log(c)}{\log(a)}}\right).$$

Our assumption that $|\cdot|$ is nonarchimedean implies that there is $c \in \mathbb{Z}$ with $c > 1$ and $|c| > 1$. Then for all $a \in \mathbb{Z}$ with $a > 1$ we have

$$1 < |c| \leq \max\left(1, |a|^{\frac{\log(c)}{\log(a)}}\right), \tag{13.3.1}$$

so $1 < |a|^{\log(c)/\log(a)}$, so $1 < |a|$ as well (i.e., any $a \in \mathbb{Z}$ with $a > 1$ automatically satisfies $|a| > 1$). Also, taking the $1/\log(c)$ power on both sides of (13.3.1) we see that

$$|c|^{\frac{1}{\log(c)}} \leq |a|^{\frac{1}{\log(a)}}. \tag{13.3.2}$$

Because, as mentioned above, $|a| > 1$, we can interchange the roll of a and c to obtain the reverse inequality of (13.3.2). We thus have

$$|c| = |a|^{\frac{\log(c)}{\log(a)}}.$$

Letting $\alpha = \log(2) \cdot \log_{|2|}(e)$ and setting $a = 2$, we have

$$|c|^{\alpha} = |2|^{\frac{\alpha}{\log(2)} \cdot \log(c)} = \left(|2|^{\log_{|2|}(e)}\right)^{\log(c)} = e^{\log(c)} = c = |c|_{\infty}.$$

Thus for all integers $c \in \mathbb{Z}$ with $c > 1$ we have $|c|^{\alpha} = |c|_{\infty}$, which implies that $|\cdot|$ is equivalent to $|\cdot|_{\infty}$. □

Let k be any field and let $K = k(t)$, where t is transcendental. Fix a real number $c > 1$. If $p = p(t)$ is an irreducible polynomial in the ring $k[t]$, we define a valuation by

$$\left|p^a \cdot \frac{u}{v}\right|_p = c^{-\deg(p)\cdot a}, \tag{13.3.3}$$

where $a \in \mathbb{Z}$ and $u, v \in k[t]$ with $p \nmid u$ and $p \nmid v$.

Remark 13.3.4. This definition differs from the one page 46 of [Cassels-Frohlich, Ch. 2] in two ways. First, we assume that $c > 1$ instead of $c < 1$, since otherwise $|\cdot|_p$ does not satisfy Axiom 3 of a valuation. Also, we write $c^{-\deg(p)\cdot a}$ instead of c^{-a}, so that the product formula will hold. (For more about the product formula, see Section 18.1.)

In addition there is a a non-archimedean valuation $|\cdot|_{\infty}$ defined by

$$\left|\frac{u}{v}\right|_{\infty} = c^{\deg(u)-\deg(v)}. \tag{13.3.4}$$

This definition differs from the one in [Cas67, pg. 46] in two ways. First, we assume that $c > 1$ instead of $c < 1$, since otherwise $|\cdot|_p$ does not satisfy Axiom 3 of a valuation. Here's why: Recall that Axiom 3 for a non-archimedean valuation on K asserts that whenever $a \in K$ and $|a| \leq 1$, then $|a+1| \leq 1$. Set $a = p - 1$, where $p = p(t) \in K[t]$ is an irreducible polynomial. Then $|a| = c^0 = 1$, since $\mathrm{ord}_p(p-1) = 0$. However, $|a+1| = |p-1+1| = |p| = c^1 < 1$, since $\mathrm{ord}_p(p) = 1$. If we take $c > 1$ instead of $c < 1$, as I propose, then $|p| = c^1 > 1$, as required.

Note the (albeit imperfect) analogy between $K = k(t)$ and $\mathbb{Q}$. If $s = t^{-1}$, so $k(t) = k(s)$, the valuation $|\cdot|_\infty$ is of the type (13.3.3) belonging to the irreducible polynomial $p(s) = s$.

The reader is urged to prove the following lemma as a homework problem.

Lemma 13.3.5. *The only nontrivial valuations on $k(t)$ which are trivial on k are equivalent to the valuation (13.3.3) or (13.3.4).*

For example, if k is a finite field, there are no nontrivial valuations on k, so the only nontrivial valuations on $k(t)$ are equivalent to (13.3.3) or (13.3.4).

Chapter 14

Topology and Completeness

14.1 Topology

A valuation $|\cdot|$ on a field K induces a topology in which a basis for the neighborhoods of a are the *open balls*

$$B(a, d) = \{x \in K : |x - a| < d\}$$

for $d > 0$.

Lemma 14.1.1. *Equivalent valuations induce the same topology.*

Proof. If $|\cdot|_1 = |\cdot|_2^r$, then $|x - a|_1 < d$ if and only if $|x - a|_2^r < d$ if and only if $|x - a|_2 < d^{1/r}$ so $B_1(a, d) = B_2(a, d^{1/r})$. Thus the basis of open neighborhoods of a for $|\cdot|_1$ and $|\cdot|_2$ are identical. $\square$

A valuation satisfying the triangle inequality gives a metric for the topology on defining the distance from a to b to be $|a - b|$. Assume for the rest of this section that we only consider valuations that satisfy the triangle inequality.

Lemma 14.1.2. *A field with the topology induced by a valuation is a* topological field*, i.e., the operations sum, product, and reciprocal are continuous.*

Proof. For example (product) the triangle inequality implies that

$$|(a + \varepsilon)(b + \delta) - ab| \leq |\varepsilon|\,|\delta| + |a|\,|\delta| + |b|\,|\varepsilon|$$

is small when $|\varepsilon|$ and $|\delta|$ are small (for fixed a, b). $\square$

Lemma 14.1.3. *Suppose two valuations $|\cdot|_1$ and $|\cdot|_2$ on the same field K induce the same topology. Then for any sequence $\{x_n\}$ in K we have*

$$|x_n|_1 \to 0 \iff |x_n|_2 \to 0.$$

Proof. It suffices to prove that if $|x_n|_1 \to 0$ then $|x_n|_2 \to 0$, since the proof of the other implication is the same. Let $\varepsilon > 0$. The topologies induced by the two absolute values are the same, so $B_2(0,\varepsilon)$ can be covered by open balls $B_1(a_i,r_i)$. One of these open balls $B_1(a,r)$ contains 0. There is $\varepsilon' > 0$ such that

$$B_1(0,\varepsilon') \subset B_1(a,r) \subset B_2(0,\varepsilon).$$

Since $|x_n|_1 \to 0$, there exists N such that for $n \geq N$ we have $|x_n|_1 < \varepsilon'$. For such n, we have $x_n \in B_1(0,\varepsilon')$, so $x_n \in B_2(0,\varepsilon)$, so $|x_n|_2 < \varepsilon$. Thus $|x_n|_2 \to 0$. $\square$

Proposition 14.1.4. *If two valuations $|\cdot|_1$ and $|\cdot|_2$ on the same field induce the same topology, then they are equivalent in the sense that there is a positive real α such that $|\cdot|_1 = |\cdot|_2^\alpha$.*

Proof. If $x \in K$ and $i = 1,2$, then $|x^n|_i \to 0$ if and only if $|x|_i^n \to 0$, which is the case if and only if $|x|_i < 1$. Thus Lemma 14.1.3 implies that $|x|_1 < 1$ if and only if $|x|_2 < 1$. On taking reciprocals we see that $|x|_1 > 1$ if and only if $|x|_2 > 1$, so finally $|x|_1 = 1$ if and only if $|x|_2 = 1$.

Let now $w, z \in K$ be nonzero elements with $|w|_i \neq 1$ and $|z|_i \neq 1$. On applying the foregoing to

$$x = w^m z^n \qquad (m,n \in \mathbb{Z})$$

we see that

$$m \log |w|_1 + n \log |z|_1 \geq 0$$

if and only if

$$m \log |w|_2 + n \log |z|_2 \geq 0.$$

Dividing through by $\log |z|_i$, and rearranging, we see that for every rational number $\alpha = -n/m$,

$$\frac{\log |w|_1}{\log |z|_1} \geq \alpha \iff \frac{\log |w|_2}{\log |z|_2} \geq \alpha.$$

Thus

$$\frac{\log |w|_1}{\log |z|_1} = \frac{\log |w|_2}{\log |z|_2},$$

so

$$\frac{\log |w|_1}{\log |w|_2} = \frac{\log |z|_1}{\log |z|_2}.$$

Since this equality does not depend on the choice of z, we see that there is a constant c $(= \log |z|_1 / \log |z|_2)$ such that $\log |w|_1 / \log |w|_2 = c$ for all w. Thus $\log |w|_1 = c \cdot \log |w|_2$, so $|w|_1 = |w|_2^c$, which implies that $|\cdot|_1$ is equivalent to $|\cdot|_2$. $\square$

14.2 Completeness

We recall the definition of metric on a set X.

Definition 14.2.1 (Metric)**.** A *metric* on a set X is a map

$$d : X \times X \to \mathbb{R}$$

such that for all $x, y, z \in X$,

1. $d(x, y) \geq 0$ and $d(x, y) = 0$ if and only if $x = y$,
2. $d(x, y) = d(y, x)$, and
3. $d(x, z) \leq d(x, y) + d(y, z)$.

A *Cauchy sequence* is a sequence (x_n) in X such that for all $\varepsilon > 0$ there exists M such that for all $n, m > M$ we have $d(x_n, x_m) < \varepsilon$. The *completion* of X is the set of Cauchy sequences (x_n) in X modulo the equivalence relation in which two Cauchy sequences (x_n) and (y_n) are equivalent if $\lim_{n\to\infty} d(x_n, y_n) = 0$. A metric space is *complete* if every Cauchy sequence converges, and one can show that the completion of X with respect to a metric is complete.

For example, $d(x, y) = |x - y|$ (usual archimedean absolute value) defines a metric on $\mathbb{Q}$. The completion of $\mathbb{Q}$ with respect to this metric is the field $\mathbb{R}$ of real numbers. More generally, whenever $|\cdot|$ is a valuation on a field K that satisfies the triangle inequality, then $d(x, y) = |x - y|$ defines a metric on K. Consider for the rest of this section only valuations that satisfy the triangle inequality.

Definition 14.2.2 (Complete)**.** A field K is *complete* with respect to a valuation $|\cdot|$ if given any Cauchy sequence a_n, $(n = 1, 2, \ldots)$, i.e., one for which

$$|a_m - a_n| \to 0 \qquad (m, n \to \infty, \infty),$$

there is an $a^* \in K$ such that

$$a_n \to a^* \qquad \text{w.r.t. } |\cdot|$$

(i.e., $|a_n - a^*| \to 0$).

Theorem 14.2.3. *Every field K with valuation $v = |\cdot|$ can be embedded in a complete field K_v with a valuation $|\cdot|$ extending the original one in such a way that K_v is the closure of K with respect to $|\cdot|$. Further K_v is unique up to a unique isomorphism fixing K.*

Proof. Define K_v to be the completion of K with respect to the metric defined by $|\cdot|$. Thus K_v is the set of equivalence classes of Cauchy sequences, and there is a natural injective map from K to K_v sending an element $a \in K$ to the constant Cauchy sequence (a). Because the field operations on K are continuous, they induce well-defined field operations on equivalence classes of Cauchy sequences componentwise. Also, define a valuation on K_v by

$$|(a_n)_{n=1}^{\infty}| = \lim_{n\to\infty} |a_n|,$$

and note that this is well defined and extends the valuation on K.

To see that K_v is unique up to a unique isomorphism fixing K, we observe that there are no nontrivial continuous automorphisms $K_v \to K_v$ that fix K. This is because, by denseness, a continuous automorphism $\sigma : K_v \to K_v$ is determined by what it does to K, and by assumption σ is the identity map on K. More precisely, suppose $a \in K_v$ and n is a positive integer. Then by continuity there is $\delta > 0$ (with $\delta < 1/n$) such that if $a_n \in K_v$ and $|a - a_n| < \delta$ then $|\sigma(a) - \sigma(a_n)| < 1/n$. Since K is dense in K_v, we can choose the a_n above to be an element of K. Then by hypothesis $\sigma(a_n) = a_n$, so $|\sigma(a) - a_n| < 1/n$. Thus $\sigma(a) = \lim_{n\to\infty} a_n = a$. $\square$

Corollary 14.2.4. *The valuation $|\cdot|$ is non-archimedean on K_v if and only if it is so on K. If $|\cdot|$ is non-archimedean, then the set of values taken by $|\cdot|$ on K and K_v are the same.*

Proof. The first part follows from Lemma 13.2.10 which asserts that a valuation is non-archimedean if and only if $|n| < 1$ for all integers n. Since the valuation on K_v extends the valuation on K, and all n are in K, the first statement follows.

For the second, suppose that $|\cdot|$ is non-archimedean (but not necessarily discrete). Suppose $b \in K_v$ with $b \neq 0$. First I claim that there is $c \in K$ such that $|b - c| < |b|$. To see this, let $c' = b - \frac{b}{a}$, where a is some element of K_v with $|a| > 1$, note that $|b - c'| = \left|\frac{b}{a}\right| < |b|$, and choose $c \in K$ such that $|c - c'| < |b - c'|$, so

$$|b - c| = \left|b - c' - (c - c')\right| \le \max\left(\left|b - c'\right|, \left|c - c'\right|\right) = \left|b - c'\right| < |b|.$$

Since $|\cdot|$ is non-archimedean, we have

$$|b| = |(b - c) + c| \le \max(|b - c|, |c|) = |c|,$$

where in the last equality we use that $|b - c| < |b|$. Also,

$$|c| = |b + (c - b)| \le \max(|b|, |c - b|) = |b|,$$

so $|b| = |c|$, which is in the set of values of $|\cdot|$ on K. $\square$

14.2.1 p-adic Numbers

This section is about the p-adic numbers $\mathbb{Q}_p$, which are the completion of $\mathbb{Q}$ with respect to the p-adic valuation. Alternatively, to give a p-adic *integer* in $\mathbb{Z}_p$ is the same as giving for every prime power p^r an element $a_r \in \mathbb{Z}/p^r\mathbb{Z}$ such that if $s \le r$ then a_s is the reduction of a_r modulo p^s. The field $\mathbb{Q}_p$ is then the field of fractions of $\mathbb{Z}_p$.

We begin with the definition of the N-adic numbers for any positive integer N. Section 14.2.1 is about the N-adics in the special case $N = 10$; these are fun because they can be represented as decimal expansions that go off infinitely far to the left. Section 14.2.3 is about how the topology of $\mathbb{Q}_N$ is nothing like the topology of $\mathbb{R}$. Finally, in Section 14.2.4 we state the Hasse-Minkowski theorem, which shows how to use p-adic numbers to decide whether or not a quadratic equation in n variables has a rational zero.

The N-adic Numbers

Lemma 14.2.5. *Let N be a positive integer. Then for any nonzero rational number α there exists a unique $e \in \mathbb{Z}$ and integers a, b, with b positive, such that $\alpha = N^e \cdot \frac{a}{b}$ with $N \nmid a$, $\gcd(a,b) = 1$, and $\gcd(N,b) = 1$.*

Proof. Write $\alpha = c/d$ with $c, d \in \mathbb{Z}$ and $d > 0$. First suppose d is exactly divisible by a power of N, so for some r we have $N^r \mid d$ but $\gcd(N, d/N^r) = 1$. Then

$$\frac{c}{d} = N^{-r}\frac{c}{d/N^r}.$$

If N^s is the largest power of N that divides c, then $e = s - r$, $a = c/N^s$, $b = d/N^r$ satisfy the conclusion of the lemma.

By unique factorization of integers, there is a smallest multiple f of d such that fd is exactly divisible by N. Now apply the above argument with c and d replaced by cf and df. $\square$

Definition 14.2.6 (N-adic valuation)**.** Let N be a positive integer. For any positive $\alpha \in \mathbb{Q}$, the *N-adic valuation* of α is e, where e is as in Lemma 14.2.5. The N-adic valuation of 0 is ∞.

We denote the N-adic valuation of α by $\mathrm{ord}_N(\alpha)$. (Note: Here we are using "valuation" in a different way than in the rest of the text. This valuation is not an absolute value, but the logarithm of one.)

Definition 14.2.7 (N-adic metric)**.** For $x, y \in \mathbb{Q}$ the *N-adic distance* between x and y is

$$d_N(x,y) = N^{-\mathrm{ord}_N(x-y)}.$$

We let $d_N(x,x) = 0$, since $\mathrm{ord}_N(x-x) = \mathrm{ord}_N(0) = \infty$.

For example, $x, y \in \mathbb{Z}$ are close in the N-adic metric if their difference is divisible by a large power of N. E.g., if $N = 10$ then 93427 and 13427 are close because their difference is 80000, which is divisible by a large power of 10.

Proposition 14.2.8. *The distance d_N on $\mathbb{Q}$ defined above is a metric. Moreover, for all $x, y, z \in \mathbb{Q}$ we have*

$$d(x,z) \le \max(d(x,y), d(y,z)).$$

(This is the "nonarchimedean" triangle inequality.)

Proof. The first two properties of Definition 14.2.1 are immediate. For the third, we first prove that if $\alpha, \beta \in \mathbb{Q}$ then

$$\mathrm{ord}_N(\alpha + \beta) \geq \min(\mathrm{ord}_N(\alpha), \mathrm{ord}_N(\beta)).$$

Assume, without loss, that $\mathrm{ord}_N(\alpha) \leq \mathrm{ord}_N(\beta)$ and that both α and β are nonzero. Using Lemma 14.2.5 write $\alpha = N^e(a/b)$ and $\beta = N^f(c/d)$ with a or c possibly negative. Then

$$\alpha + \beta = N^e \left(\frac{a}{b} + N^{f-e}\frac{c}{d}\right) = N^e \left(\frac{ad + bcN^{f-e}}{bd}\right).$$

Since $\gcd(N, bd) = 1$ it follows that $\mathrm{ord}_N(\alpha + \beta) \geq e$. Now suppose $x, y, z \in \mathbb{Q}$. Then

$$x - z = (x - y) + (y - z),$$

so

$$\mathrm{ord}_N(x - z) \geq \min(\mathrm{ord}_N(x - y), \mathrm{ord}_N(y - z)),$$

hence $d_N(x, z) \leq \max(d_N(x, y), d_N(y, z))$. □

We can finally define the N-adic numbers.

Definition 14.2.9 (The N-adic Numbers)**.** The set of N*-adic numbers*, denoted $\mathbb{Q}_N$, is the completion of $\mathbb{Q}$ with respect to the metric d_N.

The set $\mathbb{Q}_N$ is a ring, but it need not be a field as you will show in Exercises 11 and 12. It is a field if and only if N is prime. Also, $\mathbb{Q}_N$ has a “bizarre” topology, as we will see in Section 14.2.3.

The 10-adic Numbers

It’s a familiar fact that every real number can be written in the form

$$d_n \ldots d_1 d_0 . d_{-1} d_{-2} \ldots = d_n 10^n + \cdots + d_1 10 + d_0 + d_{-1} 10^{-1} + d_{-2} 10^{-2} + \cdots$$

where each digit d_i is between 0 and 9, and the sequence can continue indefinitely to the right.

The 10-adic numbers also have decimal expansions, but everything is backward! To get a feeling for why this might be the case, we consider Euler’s nonsensical series

$$\sum_{n=1}^{\infty} (-1)^{n+1} n! = 1! - 2! + 3! - 4! + 5! - 6! + \cdots .$$

One can prove (see Exercise 9) that this series converges in $\mathbb{Q}_{10}$ to some element $\alpha \in \mathbb{Q}_{10}$.

What is α? How can we write it down? First note that for all $M \geq 5$, the terms of the sum are divisible by 10, so the difference between α and $1! - 2! + 3! - 4!$ is divisible by 10. Thus we can compute α modulo 10 by computing $1! - 2! + 3! - 4!$ modulo 10. Likewise, we can compute α modulo 100 by compute $1! - 2! + \cdots + 9! - 10!$, etc. We obtain the following table:

α	mod 10^r
1	mod 10
81	mod 10^2
981	mod 10^3
2981	mod 10^4
22981	mod 10^5
422981	mod 10^6

Continuing we see that

$$1! - 2! + 3! - 4! + \cdots = \ldots 637838364422981 \qquad \text{in } \mathbf{Q}_{10}\ !$$

Here's another example. Reducing $1/7$ modulo larger and larger powers of 10 we see that

$$\frac{1}{7} = \ldots 857142857143 \qquad \text{in } \mathbf{Q}_{10}.$$

Here's another example, but with a decimal point.

$$\frac{1}{70} = \frac{1}{10} \cdot \frac{1}{7} = \ldots 85714285714.3$$

We have

$$\frac{1}{3} + \frac{1}{7} = \ldots 66667 + \ldots 57143 = \frac{10}{21} = \ldots 23810,$$

which illustrates that addition with carrying works as usual.

Fermat's Last Theorem in $\mathbf{Z}_{10}$

An amusing observation, which people often argued about on USENET news back in the 1990s, is that Fermat's last theorem is false in $\mathbf{Z}_{10}$. For example, $x^3 + y^3 = z^3$ has a nontrivial solution, namely $x = 1$, $y = 2$, and $z = \ldots 60569$. Here z is a cube root of 9 in $\mathbf{Z}_{10}$. Note that it takes some work to prove that there is a cube root of 9 in $\mathbf{Z}_{10}$ (see Exercise 10).

14.2.2 The Field of p-adic Numbers

The ring $\mathbf{Q}_{10}$ of 10-adic numbers is isomorphic to $\mathbf{Q}_2 \times \mathbf{Q}_5$ (see Exercise 12), so it is not a field. For example, the element $\ldots 8212890625$ corresponding to $(1, 0)$ under this isomorphism has no inverse. (To compute n digits of $(1, 0)$ use the Chinese remainder theorem to find a number that is 1 modulo 2^n and 0 modulo 5^n.)

If p is prime then $\mathbf{Q}_p$ is a field (see Exercise 11). Since $p \neq 10$ it is a little more complicated to write p-adic numbers down. People typically write p-adic numbers in the form

$$\frac{a_{-d}}{p^d} + \cdots + \frac{a_{-1}}{p} + a_0 + a_1 p + a_2 p^2 + a_3 p^3 + \cdots$$

where $0 \le a_i < p$ for each i.

14.2.3 The Topology of $\mathbb{Q}_N$ (is Weird)

Definition 14.2.10 (Connected)**.** Let X be a topological space. A subset S of X is *disconnected* if there exist open subsets $U_1, U_2 \subset X$ with $U_1 \cap U_2 \cap S = \emptyset$ and $S = (S \cap U_1) \cup (S \cap U_2)$ with $S \cap U_1$ and $S \cap U_2$ nonempty. If S is not disconnected it is *connected*.

The topology on $\mathbb{Q}_N$ is induced by d_N, so every open set is a union of open balls

$$B(x, r) = \{y \in \mathbb{Q}_N : d_N(x, y) < r\}.$$

Recall Proposition 14.2.8, which asserts that for all x, y, z,

$$d(x, z) \leq \max(d(x, y), d(y, z)).$$

This translates into the following shocking and bizarre lemma:

Lemma 14.2.11. *Suppose $x \in \mathbb{Q}_N$ and $r > 0$. If $y \in \mathbb{Q}_N$ and $d_N(x, y) \geq r$, then $B(x, r) \cap B(y, r) = \emptyset$.*

Proof. Suppose $z \in B(x, r)$ and $z \in B(y, r)$. Then

$$r \leq d_N(x, y) \leq \max(d_N(x, z), d_N(z, y)) < r,$$

a contradiction. $\square$

You should draw a picture to illustrates Lemma 14.2.11.

Lemma 14.2.12. *The open ball $B(x, r)$ is also closed.*

Proof. Suppose $y \notin B(x, r)$. Then $r \leq d(x, y)$ so

$$B(y, d(x, y)) \cap B(x, r) \subset B(y, d(x, y)) \cap B(x, d(x, y)) = \emptyset.$$

Thus the complement of $B(x, r)$ is a union of open balls. $\square$

The lemmas imply that $\mathbb{Q}_N$ is *totally disconnected*, in the following sense.

Proposition 14.2.13. *The only connected subsets of $\mathbb{Q}_N$ are the singleton sets $\{x\}$ for $x \in \mathbb{Q}_N$ and the empty set.*

Proof. Suppose $S \subset \mathbb{Q}_N$ is a nonempty connected set and x, y are distinct elements of S. Let $r = d_N(x, y) > 0$. Let $U_1 = B(x, r)$ and U_2 be the complement of U_1, which is open by Lemma 14.2.12. Then U_1 and U_2 satisfies the conditions of Definition 14.2.10, so S is not connected, a contradiction. $\square$

14.2.4 The Local-to-Global Principle of Hasse and Minkowski

Section 14.2.3 might have convinced you that $\mathbb{Q}_N$ is a bizarre pathology. In fact, $\mathbb{Q}_N$ is omnipresent in number theory, as the following two fundamental examples illustrate.

In the statement of the following theorem, a *nontrivial solution* to a homogeneous polynomial equation is a solution where not all indeterminates are 0.

Theorem 14.2.14 (Hasse-Minkowski)**.** *The quadratic equation*

$$a_1x_1^2 + a_2x_2^2 + \cdots + a_nx_n^2 = 0, \tag{14.2.1}$$

with $a_i \in \mathbb{Q}^\times$, has a nontrivial solution with $x_1, \ldots, x_n$ in $\mathbb{Q}$ if and only if (14.2.1) has a solution in $\mathbb{R}$ and in $\mathbb{Q}_p$ for all primes p.

This theorem is very useful in practice because the p-adic condition turns out to be easy to check. For more details, including a complete proof, see [Ser73, IV.3.2].

The analogue of Theorem 14.2.14 for cubic equations is false. For example, Selmer proved that the cubic

$$3x^3 + 4y^3 + 5z^3 = 0$$

has a solution other than $(0, 0, 0)$ in $\mathbb{R}$ and in $\mathbb{Q}_p$ for all primes p but has no solution other than $(0, 0, 0)$ in $\mathbb{Q}$ (for a proof see [Cas91, §18]).

Open Problem. Give an algorithm that decides whether or not a cubic

$$ax^3 + by^3 + cz^3 = 0$$

has a nontrivial solution in $\mathbb{Q}$.

This open problem is closely related to the Birch and Swinnerton-Dyer Conjecture for elliptic curves. The truth of the conjecture would follow if we knew that "Shafarevich-Tate Groups" of certain elliptic curves are finite.

14.3 Weak Approximation

The following theorem asserts that inequivalent valuations are in fact almost totally independent. For our purposes it will be superseded by the strong approximation theorem (Theorem 18.4.4).

Theorem 14.3.1 (Weak Approximation)**.** *Let $|\cdot|_n$, for $1 \leq n \leq N$, be inequivalent nontrivial valuations of a field K. For each n, let K_n be the topological space consisting of the set of elements of K with the topology induced by $|\cdot|_n$. Let Δ be the image of K in the topological product*

$$A = \prod_{1 \leq n \leq N} K_n$$

equipped with the product topology. Then Δ is dense in A.

The conclusion of the theorem may be expressed in a less topological manner as follows: given any $a_n \in K$, for $1 \leq n \leq N$, and real $\varepsilon > 0$, there is an $b \in K$ such that simultaneously

$$|a_n - b|_n < \varepsilon \qquad (1 \leq n \leq N).$$

If $K = \mathbb{Q}$ and the $|\cdot|$ are p-adic valuations, Theorem 14.3.1 is related to the Chinese Remainder Theorem (Theorem 5.1.4), but the strong approximation theorem (Theorem 18.4.4) is the real generalization.

Proof. We note first that it will be enough to find, for each n, an element $c_n \in K$ such that

$$|c_n|_n > 1 \quad \text{and} \quad |c_n|_m < 1 \quad \text{for } n \neq m,$$

where $1 \leq n, m \leq N$. For then as $r \to +\infty$, we have

$$\frac{c_n^r}{1 + c_n^r} = \frac{1}{1 + \left(\frac{1}{c_n}\right)^r} \to \begin{cases} 1 & \text{with respect to } |\cdot|_n \text{ and} \\ 0 & \text{with respect to } |\cdot|_m\text{, for } m \neq n. \end{cases}$$

It is then enough to take

$$b = \sum_{n=1}^{N} \frac{c_n^r}{1 + c_n^r} \cdot a_n$$

By symmetry it is enough to show the existence of $c = c_1$ with

$$|c|_1 > 1 \qquad \text{and} \qquad |c|_n < 1 \quad \text{for} \quad 2 \leq n \leq N.$$

We will do this by induction on N.

First suppose $N = 2$. Since $|\cdot|_1$ and $|\cdot|_2$ are inequivalent (and all absolute values are assumed nontrivial) there is an $a \in K$ such that

$$|a|_1 < 1 \qquad \text{and} \qquad |a|_2 \geq 1 \tag{14.3.1}$$

and similarly a b such that

$$|b|_1 \geq 1 \qquad \text{and} \qquad |b|_2 < 1.$$

Then $c = \dfrac{b}{a}$ will do.

Remark 14.3.2. It is not completely clear that one can choose an a such that (14.3.1) is satisfied. Suppose it were impossible. Then because the valuations are nontrivial, we would have that for any $a \in K$ if $|a|_1 < 1$ then $|a|_2 < 1$. This implies the converse statement: if $a \in K$ and $|a|_2 < 1$ then $|a|_1 < 1$. To see this, suppose there is an $a \in K$ such that $|a|_2 < 1$ and $|a|_1 \geq 1$. Choose $y \in K$ such that $|y|_1 < 1$. Then for any integer $n > 0$ we have $|y/a^n|_1 < 1$, so by hypothesis $|y/a^n|_2 < 1$. Thus $|y|_2 < |a|_2^n < 1$ for all n. Since $|a|_2 < 1$ we have $|a|_2^n \to 0$ as $n \to \infty$, so $|y|_2 = 0$, a contradiction since $y \neq 0$. Thus $|a|_1 < 1$ if and only if $|a|_2 < 1$, and we have proved before that this implies that $|\cdot|_1$ is equivalent to $|\cdot|_2$.

Next suppose $N \geq 3$. By the case $N-1$, there is an $a \in K$ such that

$$|a|_1 > 1 \qquad \text{and} \qquad |a|_n < 1 \quad \text{for} \quad 2 \leq n \leq N-1.$$

By the case for $N = 2$ there is a $b \in K$ such that

$$|b|_1 > 1 \qquad \text{and} \qquad |b|_N < 1.$$

Then put

$$c = \begin{cases} a & \text{if } |a|_N < 1 \\ a^r \cdot b & \text{if } |a|_N = 1 \\ \dfrac{a^r}{1+a^r} \cdot b & \text{if } |a|_N > 1 \end{cases}$$

where $r \in \mathbb{Z}$ is sufficiently large so that $|c|_1 > 1$ and $|c|_n < 1$ for $2 \leq n \leq N$. $\square$

Example 14.3.3. Suppose $K = \mathbb{Q}$, let $|\cdot|_1$ be the archimedean absolute value and let $|\cdot|_2$ be the 2-adic absolute value. Let $a_1 = -1$, $a_2 = 8$, and $\varepsilon = 1/10$, as in the remark right after Theorem 14.3.1. Then the theorem implies that there is an element $b \in \mathbb{Q}$ such that

$$|-1-b|_1 < \frac{1}{10} \qquad \text{and} \qquad |8-b|_2 < \frac{1}{10}.$$

As in the proof of the theorem, we can find such a b by finding a $c_1, c_2 \in \mathbb{Q}$ such that $|c_1|_1 > 1$ and $|c_1|_2 < 1$, and a $|c_2|_1 < 1$ and $|c_2|_2 > 1$. For example, $c_1 = 2$ and $c_2 = 1/2$ works, since $|2|_1 = 2$ and $|2|_2 = 1/2$ and $|1/2|_1 = 1/2$ and $|1/2|_2 = 2$. Again following the proof, we see that for sufficiently large r we can take

$$\begin{aligned} b_r &= \frac{c_1^r}{1+c_1^r} \cdot a_1 + \frac{c_2^r}{1+c_2^r} \cdot a_2 \\ &= \frac{2^r}{1+2^r} \cdot (-1) + \frac{(1/2)^r}{1+(1/2)^r} \cdot 8. \end{aligned}$$

We have $b_1 = 2$, $b_2 = 4/5$, $b_3 = 0$, $b_4 = -8/17$, $b_5 = -8/11$, $b_6 = -56/55$. None of the b_i work for $i < 6$, but b_6 works.

Chapter 15

Adic Numbers: The Finite Residue Field Case

15.1 Finite Residue Field Case

Let K be a field with a non-archimedean valuation $v = |\cdot|$. Recall that the set of $a \in K$ with $|a| \leq 1$ forms a ring $\mathcal{O}$, the ring of integers for v. The set of $u \in K$ with $|u| = 1$ are a group U under multiplication, the group of units for v. Finally, the set of $a \in K$ with $|a| < 1$ is a maximal ideal $\mathfrak{p}$, so the quotient ring $\mathcal{O}/\mathfrak{p}$ is a field. In this section we consider the case when $\mathcal{O}/\mathfrak{p}$ is a finite field of order a prime power q. For example, K could be $\mathbb{Q}$ and $|\cdot|$ could be a p-adic valuation, or K could be a number field and $|\cdot|$ could be the valuation corresponding to a maximal ideal of the ring of integers. Among other things, we will discuss in more depth the topological and measure-theoretic nature of the completion of K at v.

Suppose further for the rest of this section that $|\cdot|$ is discrete. Then by Lemma 13.2.8, the ideal $\mathfrak{p}$ is a principal ideal (π), say, and every $a \in K$ is of the form $a = \pi^n \varepsilon$, where $n \in \mathbb{Z}$ and $\varepsilon \in U$ is a unit. We call

$$n = \mathrm{ord}(a) = \mathrm{ord}_\pi(a) = \mathrm{ord}_\mathfrak{p}(a) = \mathrm{ord}_v(a)$$

the ord of a at v. (Some authors, including me (!) also call this integer the *valuation* of a with respect to v.) If $\mathfrak{p} = (\pi')$, then π/π' is a unit, and conversely, so $\mathrm{ord}(a)$ is independent of the choice of π.

Let $\mathcal{O}_v$ and $\mathfrak{p}_v$ be defined with respect to the completion K_v of K at v.

Lemma 15.1.1. *There is a natural isomorphism*

$$\varphi : \mathcal{O}_v/\mathfrak{p}_v \to \mathcal{O}/\mathfrak{p},$$

and $\mathfrak{p}_v = (\pi)$ *as an* $\mathcal{O}_v$*-ideal.*

Proof. We may view $\mathcal{O}_v$ as the set of equivalence classes of Cauchy sequences (a_n) in K such that $a_n \in \mathcal{O}$ for n sufficiently large. For any ε, given such a sequence

(a_n), there is N such that for $n, m \geq N$, we have $|a_n - a_m| < \varepsilon$. In particular, we can choose N such that $n, m \geq N$ implies that $a_n \equiv a_m \pmod{\mathfrak{p}}$. Let $\varphi((a_n)) = a_N \pmod{\mathfrak{p}}$, which is well-defined. The map φ is surjective because the constant sequences are in $\mathcal{O}_v$. Its kernel is the set of Cauchy sequences whose elements are eventually all in $\mathfrak{p}$, which is exactly $\mathfrak{p}_v$. This proves the first part of the lemma. The second part is true because any element of $\mathfrak{p}_v$ is a sequence all of whose terms are eventually in $\mathfrak{p}$, hence all a multiple of π (we can set to 0 a finite number of terms of the sequence without changing the equivalence class of the sequence). $\square$

Assume for the rest of this section that K is complete with respect to $|\cdot|$.

Lemma 15.1.2. *Then ring $\mathcal{O}$ is precisely the set of infinite sums*

$$a = \sum_{j=0}^{\infty} a_j \cdot \pi^j \tag{15.1.1}$$

where the a_j run independently through some set $\mathcal{R}$ of representatives of $\mathcal{O}$ in $\mathcal{O}/\mathfrak{p}$.

By (15.1.1) is meant the limit of the Cauchy sequence $\sum_{j=0}^{n} a_j \cdot \pi^j$ as $j \to \infty$.

Proof. There is a uniquely defined $a_0 \in \mathcal{R}$ such that $|a - a_0| < 1$. Then $a' = \pi^{-1} \cdot (a - a_0) \in \mathcal{O}$. Now define $a_1 \in \mathcal{R}$ by $|a' - a_1| < 1$. And so on. $\square$

Example 15.1.3. Suppose $K = \mathbb{Q}$ and $|\cdot| = |\cdot|_p$ is the p-adic valuation, for some prime p. We can take $\mathcal{R} = \{0, 1, \ldots, p-1\}$. The lemma asserts that

$$\mathcal{O} = \mathbb{Z}_p = \left\{ \sum_{j=0}^{\infty} a_n p^n : 0 \leq a_n \leq p-1 \right\}.$$

Notice that $\mathcal{O}$ is uncountable since there are p choices for each p-adic "digit". We can do arithmetic with elements of $\mathbb{Z}_p$, which can be thought of "backwards" as numbers in base p. For example, with $p = 3$ we have

$$\begin{aligned}
&(1 + 2\cdot 3 + 3^2 + \cdots) + (2 + 2\cdot 3 + 3^2 + \cdots) \\
&= 3 + 4\cdot 3 + 2\cdot 3^2 + \cdots \qquad \text{not in canonical form} \\
&= 0 + 2\cdot 3 + 3\cdot 3 + 2\cdot 3^2 + \cdots \qquad \text{still not canonical} \\
&= 0 + 2\cdot 3 + 0\cdot 3^2 + \cdots
\end{aligned}$$

Here is an example of doing basic arithmetic with p-adic numbers in Sage:

```
sage: a = 1 + 2*3 + 3^2 + O(3^3)
sage: b = 2 + 2*3 + 3^2 + O(3^3)
sage: a + b
2*3 + O(3^3)
sage: sqrt(a)
```

```
1 + 3 + O(3^3)
sage: sqrt(a)^2
1 + 2*3 + 3^2 + O(3^3)
sage: a * b
2 + O(3^3)
```

Type `Zp?` and `Qp?` in Sage for much more information about the various computer models of p-adic arithmetic that are available.

Theorem 15.1.4. *Under the conditions of the preceding lemma, $\mathcal{O}$ is compact with respect to the $|\cdot|$-topology.*

Proof. Let V_λ, for λ running through some index set Λ, be some family of open sets that cover $\mathcal{O}$. We must show that there is a finite subcover. We suppose not.

Let $\mathcal{R}$ be a set of representatives for $\mathcal{O}/\mathfrak{p}$. Then $\mathcal{O}$ is the union of the finite number of cosets $a + \pi\mathcal{O}$, for $a \in \mathcal{R}$. Hence for at lest one $a_0 \in \mathcal{R}$ the set $a_0 + \pi\mathcal{O}$ is not covered by finitely many of the V_λ. Then similarly there is an $a_1 \in \mathcal{R}$ such that $a_0 + a_1\pi + \pi^2\mathcal{O}$ is not finitely covered. And so on. Let

$$a = a_0 + a_1\pi + a_2\pi^2 + \cdots \in \mathcal{O}.$$

Then $a \in V_{\lambda_0}$ for some $\lambda_0 \in \Lambda$. Since V_{λ_0} is an open set, $a + \pi^J \cdot \mathcal{O} \subset V_{\lambda_0}$ for some J (since those are exactly the open balls that form a basis for the topology). This is a contradiction because we constructed a so that none of the sets $a + \pi^n \cdot \mathcal{O}$, for each n, are not covered by any finite subset of the V_λ. $\square$

Definition 15.1.5 (Locally compact)**.** A topological space X is *locally compact* at a point x if there is some compact subset C of X that contains a neighborhood of x. The space X is locally compact if it is locally compact at each point in X.

Corollary 15.1.6. *The complete local field K is locally compact.*

Proof. If $x \in K$, then $x \in C = x + \mathcal{O}$, and C is a compact subset of K by Theorem 15.1.4. Also C contains the neighborhood $x + \pi\mathcal{O} = B(x, 1)$ of x. Thus K is locally compact at x. $\square$

Remark 15.1.7. The converse is also true. If K is locally compact with respect to a non-archimedean valuation $|\cdot|$, then

1. K is complete,
2. the residue field is finite, and
3. the valuation is discrete.

For there is a compact neighbourhood C of 0. Let π be any nonzero with $|\pi| < 1$. Then $\pi^n \cdot \mathcal{O} \subset C$ for sufficiently large n, so $\pi^n \cdot \mathcal{O}$ is compact, being closed. Hence $\mathcal{O}$ is compact. Since $|\cdot|$ is a metric, $\mathcal{O}$ is sequentially compact, i.e., every fundamental sequence in $\mathcal{O}$ has a limit, which implies (1). Let a_λ (for $\lambda \in \Lambda$) be a set of representatives in $\mathcal{O}$ of $\mathcal{O}/\mathfrak{p}$. Then $\mathcal{O}_\lambda = \{z : |z - a_\lambda| < 1\}$ is an open covering of $\mathcal{O}$. Thus (2) holds since $\mathcal{O}$ is compact. Finally, $\mathfrak{p}$ is compact, being a closed subset of $\mathcal{O}$. Let S_n be the set of $a \in K$ with $|a| < 1 - 1/n$. Then S_n (for $1 \leq n < \infty$) is an open covering of $\mathfrak{p}$, so $\mathfrak{p} = S_n$ for some n, i.e., (3) is true.

If we allow $|\cdot|$ to be archimedean the only further possibilities are $k = \mathbb{R}$ and $k = \mathbb{C}$ with $|\cdot|$ equivalent to the usual absolute value.

We denote by K^+ the commutative topological group whose points are the elements of K, whose group law is addition and whose topology is that induced by $|\cdot|$. General theory tells us that there is an invariant Haar measure defined on K^+ and that this measure is unique up to a multiplicative constant.

Definition 15.1.8 (Haar Measure)**.** A *Haar measure* on a locally compact topological group G is a translation invariant measure such that every open set can be covered by open sets with finite measure.

Lemma 15.1.9. *Haar measure of any compact subset C of G is finite.*

Proof. The whole group G is open, so there is a covering U_α of G by open sets each of which has finite measure. Since C is compact, there is a finite subset of the U_α that covers C. The measure of C is at most the sum of the measures of these finitely many U_α, hence finite. □

Remark 15.1.10. Usually one defined Haar measure to be a translation invariant measure such that the measure of compact sets is finite. Because of local compactness, this definition is equivalent to Definition 15.1.8. We take this alternative viewpoint because Haar measure is constructed naturally on the topological groups we will consider by defining the measure on each member of a basis of open sets for the topology.

We now deduce what any such measure μ on $G = K^+$ must be. Since $\mathcal{O}$ is compact (Theorem 15.1.4), the measure of $\mathcal{O}$ is finite. Since μ is translation invariant,

$$\mu_n = \mu(a + \pi^n \mathcal{O})$$

is independent of a. Further,

$$a + \pi^n \mathcal{O} = \bigcup_{1 \leq j \leq q} a + \pi^n a_j + \pi^{n+1}\mathcal{O}, \qquad \text{(disjoint union)}$$

where a_j (for $1 \leq j \leq q$) is a set of representatives of $\mathcal{O}/\mathfrak{p}$. Hence

$$\mu_n = q \cdot \mu_{n+1}.$$

If we normalize μ by putting

$$\mu(\mathcal{O}) = 1$$

we have $\mu_0 = 1$, hence $\mu_1 = q$, and in general

$$\mu_n = q^{-n}.$$

Conversely, without the theory of Haar measure, we could *define* μ to be the necessarily unique measure on K^+ such that $\mu(\mathcal{O}) = 1$ that is translation invariant. This would have to be the μ we just found above.

Everything so far in this section has depended not on the valuation $|\cdot|$ but only on its equivalence class. The above considerations now single out one valuation in the equivalence class as particularly important.

Definition 15.1.11 (Normalized valuation)**.** Let K be a field equipped with a discrete valuation $|\cdot|$ and residue class field with $q < \infty$ elements. We say that $|\cdot|$ is *normalized* if

$$|\pi| = \frac{1}{q},$$

where $\mathfrak{p} = (\pi)$ is the maximal ideal of $\mathcal{O}$.

Example 15.1.12. The normalized valuation on the p-adic numbers $\mathbb{Q}_p$ is $|u \cdot p^n| = p^{-n}$, where u is a rational number whose numerator and denominator are coprime to p.

Next suppose $K = \mathbb{Q}_p(\sqrt{p})$. Then the p-adic valuation on $\mathbb{Q}_p$ extends uniquely to one on K such that $\left|\sqrt{p}\right|^2 = |p| = 1/p$. Since $\pi = \sqrt{p}$ for K, this valuation is not normalized. (Note that the ord of $\pi = \sqrt{p}$ is $1/2$.) The normalized valuation is $v = |\cdot|' = |\cdot|^2$. Note that $|\cdot|'\,p = 1/p^2$, or $\mathrm{ord}_v(p) = 2$ instead of 1.

Finally suppose that $K = \mathbb{Q}_p(\sqrt{q})$ where $x^2 - q$ has not root mod p. Then the residue class field degree is 2, and the normalized valuation must satisfy $\left|\sqrt{q}\right| = 1/p^2$.

The following proposition makes clear why this is the best choice of normalization.

Theorem 15.1.13. *Suppose further that K is complete with respect to the normalized valuation $|\cdot|$. Then*

$$\mu(a + b\mathcal{O}) = |b|\,,$$

where μ is the Haar measure on K^+ normalized so that $\mu(\mathcal{O}) = 1$.

Proof. Since μ is translation invariant, $\mu(a + b\mathcal{O}) = \mu(b\mathcal{O})$. Write $b = u \cdot \pi^n$, where u is a unit. Then since $u \cdot \mathcal{O} = \mathcal{O}$, we have

$$\mu(b\mathcal{O}) = \mu(u \cdot \pi^n \cdot \mathcal{O}) = \mu(\pi^n \cdot u \cdot \mathcal{O}) = \mu(\pi^n \cdot \mathcal{O}) = q^{-n} = |\pi^n| = |b|\,.$$

Here we have $\mu(\pi^n \cdot \mathcal{O}) = q^{-n}$ by the discussion before Definition 15.1.11. □

We can express the result of the theorem in a more suggestive way. Let $b \in K$ with $b \neq 0$, and let μ be a Haar measure on K^+ (not necessarily normalized as in the theorem). Then we can define a new Haar measure μ_b on K^+ by putting $\mu_b(E) = \mu(bE)$ for $E \subset K^+$. But Haar measure is unique up to a multiplicative constant and so $\mu_b(E) = \mu(bE) = c \cdot \mu(E)$ for all measurable sets E, where the factor c depends only on b. Putting $E = \mathcal{O}$, shows that the theorem implies that c is just $|b|$, when $|\cdot|$ is the normalized valuation.

Remark 15.1.14. The theory of locally compact topological groups leads to the consideration of the dual (character) group of K^+. It turns out that it is isomorphic to K^+. We do not need this fact for class field theory, so do not prove it here. For a proof and applications see Tate's thesis or Lang's *Algebraic Numbers*, and for generalizations see Weil's *Adeles and Algebraic Groups* and Godement's Bourbaki seminars 171 and 176. The determination of the character group of K^* is local class field theory.

The set of nonzero elements of K is a group K^* under multiplication. Multiplication and inverses are continuous with respect to the topology induced on K^* as a subset of K, so K^* is a topological group with this topology. We have

$$U_1 \subset U \subset K^*$$

where U is the group of units of $\mathcal{O} \subset K$ and U_1 is the group of 1-units, i.e., those units $\varepsilon \in U$ with $|\varepsilon - 1| < 1$, so

$$U_1 = 1 + \pi\mathcal{O}.$$

The set U is the open ball about 0 of radius 1, so is open, and because the metric is nonarchimedean U is also closed. Likewise, U_1 is both open and closed.

The quotient $K^*/U = \{\pi^n \cdot U : n \in \mathbb{Z}\}$ is isomorphic to the additive group $\mathbb{Z}^+$ of integers with the discrete topology, where the map is

$$\pi^n \cdot U \mapsto n \qquad \text{for } n \in \mathbb{Z}.$$

The quotient U/U_1 is isomorphic to the multiplicative group $\mathbb{F}^*$ of the nonzero elements of the residue class field, where the finite gorup $\mathbb{F}^*$ has the discrete topology. Note that $\mathbb{F}^*$ is cyclic of order $q-1$, and Hensel's lemma implies that K^* contains a primitive $(q-1)$th root of unity ζ. Thus K^* has the following structure:

$$K^* = \{\pi^n \zeta^m \varepsilon : n \in \mathbb{Z}, m \in \mathbb{Z}/(q-1)\mathbb{Z}, \varepsilon \in U_1\} \cong \mathbb{Z} \ \times \ \mathbb{Z}/(q-1)\mathbb{Z} \ \times \ U_1.$$

(How to apply Hensel's lemma: Let $f(x) = x^{q-1} - 1$ and let $a \in \mathcal{O}$ be such that a mod $\mathfrak{p}$ generates K^*. Then $|f(a)| < 1$ and $|f'(a)| = 1$. By Hensel's lemma there is a $\zeta \in K$ such that $f(\zeta) = 0$ and $\zeta \equiv a \pmod{\mathfrak{p}}$.)

Since U is compact and the cosets of U cover K, we see that K^* is locally compact.

Lemma 15.1.15. *The additive Haar measure μ on K^+, when restricted to U_1 gives a measure on U_1 that is also invariant under multiplication, so gives a Haar measure on U_1.*

Proof. It suffices to show that

$$\mu(1+\pi^n\mathcal{O}) = \mu(u\cdot(1+\pi^n\mathcal{O})),$$

for any $u \in U_1$ and $n > 0$. Write $u = 1 + a_1\pi + a_2\pi^2 + \cdots$. We have

$$\begin{aligned} u\cdot(1+\pi^n\mathcal{O}) &= (1 + a_1\pi + a_2\pi^2 + \cdots)\cdot(1+\pi^n\mathcal{O}) \\ &= 1 + a_1\pi + a_2\pi^2 + \cdots + \pi^n\mathcal{O} \\ &= a_1\pi + a_2\pi^2 + \cdots + (1+\pi^n\mathcal{O}), \end{aligned}$$

which is an additive translate of $1+\pi^n\mathcal{O}$, hence has the same measure. □

Thus μ gives a Haar measure on K^* by translating U_1 around to cover K^*.

Lemma 15.1.16. *The topological spaces K^+ and K^* are totally disconnected (the only connected sets are points).*

Proof. The proof is the same as that of Proposition 14.2.13. The point is that the non-archimedean triangle inequality forces the complement an open disc to be open, hence any set with at least two distinct elements "falls apart" into a disjoint union of two disjoint open subsets. □

Remark 15.1.17. Note that K^* and K^+ are locally isomorphic if K has characteristic 0. We have the exponential map

$$a \mapsto \exp(a) = \sum_{n=0}^{\infty} \frac{a^n}{n!}$$

defined for all sufficiently small a with its inverse

$$\log(a) = \sum_{n=1}^{\infty} \frac{(-1)^{n-1}(a-1)^n}{n},$$

which is defined for all a sufficiently close to 1.

Chapter 16

Normed Spaces and Tensor Products

Much of this chapter is preparation for what we will do later when we will prove that if K is complete with respect to a valuation (and locally compact) and L is a finite extension of K, then there is a *unique* valuation on L that extends the valuation on K. Also, if K is a number field, $v = |\cdot|$ is a valuation on K, K_v is the completion of K with respect to v, and L is a finite extension of K, we'll prove that

$$K_v \otimes_K L = \bigoplus_{j=1}^{J} L_j,$$

where the L_j are the completions of L with respect to the equivalence classes of extensions of v to L. In particular, if L is a number field defined by a root of $f(x) \in \mathbb{Q}[x]$, then

$$\mathbb{Q}_p \otimes_{\mathbb{Q}} L = \bigoplus_{j=1}^{J} L_j,$$

where the L_j correspond to the irreducible factors of the polynomial $f(x) \in \mathbb{Q}_p[x]$ (hence the extensions of $|\cdot|_p$ correspond to irreducible factors of $f(x)$ over $\mathbb{Q}_p[x]$).

In preparation for this clean view of the local nature of number fields, we will prove that the norms on a finite-dimensional vector space over a complete field are all equivalent. We will also explicitly construct tensor products of fields and deduce some of their properties.

16.1 Normed Spaces

Definition 16.1.1 (Norm)**.** Let K be a field with valuation $|\cdot|$ and let V be a vector space over K. A real-valued function $\|\cdot\|$ on V is called a *norm* if

1. $\|v\| > 0$ for all nonzero $v \in V$ (positivity).

2. $\|v+w\| \leq \|v\| + \|w\|$ for all $v, w \in V$ (triangle inequality).

3. $\|av\| = |a|\,\|v\|$ for all $a \in K$ and $v \in V$ (homogeneity).

Note that setting $\|v\| = 1$ for all $v \neq 0$ does *not* define a norm unless the absolute value on K is trivial, as $1 = \|av\| = |a|\,\|v\| = |a|$. We assume for the rest of this section that $|\cdot|$ is not trivial.

Definition 16.1.2 (Equivalent)**.** Two norms $\|\cdot\|_1$ and $\|\cdot\|_2$ on the same vector space V are *equivalent* if there exists positive real numbers c_1 and c_2 such that for all $v \in V$

$$\|v\|_1 \leq c_1 \|v\|_2 \qquad \text{and} \qquad \|v\|_2 \leq c_2 \|v\|_1 .$$

Lemma 16.1.3. *Suppose that K is a field that is complete with respect to a valuation $|\cdot|$ and that V is a finite dimensional K vector space. Continue to assume, as mentioned above, that K is complete with respect to $|\cdot|$. Then any two norms on V are equivalent.*

Remark 16.1.4. As we shall see soon (see Theorem 17.1.8), the lemma is usually false if we do not assume that K is complete. For example, when $K = \mathbb{Q}$ and $|\cdot|_p$ is the p-adic valuation, and V is a number field, then there may be several extensions of $|\cdot|_p$ to inequivalent norms on V.

If two norms are equivalent then the corresponding topologies on V are equal, since very open ball for $\|\cdot\|_1$ is contained in an open ball for $\|\cdot\|_2$, and conversely. (The converse is also true, since, as we will show, all norms on V are equivalent.)

Proof. Let $v_1, \ldots, v_N$ be a basis for V. Define the max norm $\|\cdot\|_0$ by

$$\left\|\sum_{n=1}^{N} a_n v_n\right\|_0 = \max\{|a_n| : n = 1, \ldots, N\}.$$

It is enough to show that any norm $\|\cdot\|$ is equivalent to $\|\cdot\|_0$. We have

$$\begin{aligned}\left\|\sum_{n=1}^{N} a_n v_n\right\| &\leq \sum_{n=1}^{N} |a_n|\,\|v_n\| \\ &\leq \sum_{n=1}^{N} \max|a_n|\,\|v_n\| \\ &= c_1 \cdot \left\|\sum_{n=1}^{N} a_n v_n\right\|_0 ,\end{aligned}$$

where $c_1 = \sum_{n=1}^{N} \|v_n\|$.

To finish the proof, we show that there is a $c_2 \in \mathbb{R}$ such that for all $v \in V$,

$$\|v\|_0 \leq c_2 \cdot \|v\| .$$

We will only prove this in the case when K is not just merely complete with respect to $|\cdot|$ but also locally compact. This will be the case of primary interest to us. For a proof in the general case, see the original article by Cassels (page 53).

By what we have already shown, the function $\|v\|$ is continuous in the $\|\cdot\|_0$-topology, so by local compactness it attains its lower bound δ on the unit circle $\{v \in V : \|v\|_0 = 1\}$. (Why is the unit circle compact? With respect to $\|\cdot\|_0$, the topology on V is the same as that of a product of copies of K. If the valuation is archimedean then $K \cong \mathbb{R}$ or $\mathbb{C}$ with the standard topology and the unit circle is compact. If the valuation is non-archimedean, then we saw (see Remark 15.1.7) that if K is locally compact, then the valuation is discrete, in which case we showed that the unit disc is compact, hence the unit circle is also compact since it is closed.) Note that $\delta > 0$ by part 1 of Definition 16.1.1. Also, by definition of $\|\cdot\|_0$, for any $v \in V$ there exists $a \in K$ such that $\|v\|_0 = |a|$ (just take the max coefficient in our basis). Thus we can write any $v \in V$ as $a \cdot w$ where $a \in K$ and $w \in V$ with $\|w\|_0 = 1$. We then have

$$\frac{\|v\|_0}{\|v\|} = \frac{\|aw\|_0}{\|aw\|} = \frac{|a|\,\|w\|_0}{|a|\,\|w\|} = \frac{1}{\|w\|} \le \frac{1}{\delta}.$$

Thus for all v we have

$$\|v\|_0 \le c_2 \cdot \|v\|,$$

where $c_2 = 1/\delta$, which proves the theorem. $\square$

16.2 Tensor Products

We need only a special case of the tensor product construction. Let A and B be commutative rings containing a field K and suppose that B is of finite dimension N over K, say, with basis

$$1 = w_1, w_2, \ldots, w_N.$$

Then B is determined up to isomorphism as a ring over K by the multiplication table $(c_{i,j,n})$ defined by

$$w_i \cdot w_j = \sum_{n=1}^{N} c_{i,j,n} \cdot w_n.$$

We define a new ring C containing K whose elements are the set of all expressions

$$\sum_{n=1}^{N} a_n \underline{w}_n$$

where the $\underline{w}_n$ have the same multiplication rule

$$\underline{w}_i \cdot \underline{w}_j = \sum_{n=1}^{N} c_{i,j,n} \cdot \underline{w}_n$$

as the w_n.

There are injective ring homomorphisms

$$i : A \hookrightarrow C, \qquad i(a) = a\underline{w}_1 \qquad \text{(note that } \underline{w}_1 = 1\text{)}$$

and

$$j : B \hookrightarrow C, \qquad j\left(\sum_{n=1}^{N} c_n w_n\right) = \sum_{n=1}^{N} c_n \underline{w}_n.$$

Moreover C is defined, up to isomorphism, by A and B and is independent of the particular choice of basis w_n of B (i.e., a change of basis of B induces a canonical isomorphism of the C defined by the first basis to the C defined by the second basis). We write

$$C = A \otimes_K B$$

since C is, in fact, a special case of the ring tensor product.

Let us now suppose, further, that A is a topological ring, i.e., has a topology with respect to which addition and multiplication are continuous. Then the map

$$C \to A \oplus \cdots \oplus A, \qquad \sum_{m=1}^{N} a_m \underline{w}_m \mapsto (a_1, \ldots, a_N)$$

defines a bijection between C and the product of N copies of A (considered as sets). We give C the product topology. It is readily verified that this topology is independent of the choice of basis $w_1, \ldots, w_N$ and that multiplication and addition on C are continuous, so C is a topological ring. We call this topology on C the *tensor product topology.*

Now drop our assumption that A and B have a topology, but suppose that A and B are not merely rings but fields. Recall that a finite extension L/K of fields is *separable* if the number of embeddings $L \hookrightarrow \overline{K}$ that fix K equals the degree of L over K, where $\overline{K}$ is an algebraic closure of K. The primitive element theorem from Galois theory asserts that any such extension is generated by a single element, i.e., $L = K(a)$ for some $a \in L$.

Lemma 16.2.1. *Let A and B be fields containing the field K and suppose that B is a separable extension of finite degree $N = [B : K]$. Then $C = A \otimes_K B$ is the direct sum of a finite number of fields K_j, each containing an isomorphic image of A and an isomorphic image of B.*

Proof. By the primitive element theorem, we have $B = K(b)$, where b is a root of some separable irreducible polynomial $f(x) \in K[x]$ of degree N. Then $1, b, \ldots, b^{N-1}$ is a basis for B over K, so

$$A \otimes_K B = A[\underline{b}] \cong A[x]/(f(x))$$

where $1, \underline{b}, \underline{b}^2, \ldots, \underline{b}^{N-1}$ are linearly independent over A and $\underline{b}$ satisfies $f(\underline{b}) = 0$.

Although the polynomial $f(x)$ is irreducible as an element of $K[x]$, it need not be irreducible in $A[x]$. Since A is a field, we have a factorization

$$f(x) = \prod_{j=1}^{J} g_j(x)$$

where $g_j(x) \in A[x]$ is irreducible. The $g_j(x)$ are distinct because $f(x)$ is separable (i.e., has distinct roots in any algebraic closure).

For each j, let $\underline{b}_j \in \overline{A}$ be a root of $g_j(x)$, where $\overline{A}$ is a fixed algebraic closure of the field A. Let $K_j = A(\underline{b}_j)$. Then the map

$$\varphi_j : A \otimes_K B \to K_j \tag{16.2.1}$$

given by sending any polynomial $h(\underline{b})$ in $\underline{b}$ (where $h \in A[x]$) to $h(\underline{b}_j)$ is a ring homomorphism, because the image of $\underline{b}$ satisfies the polynomial $f(x)$, and $A \otimes_K B \cong A[x]/(f(x))$.

By the Chinese Remainder Theorem, the maps from (16.2.1) combine to define a ring isomorphism

$$A \otimes_K B \cong A[x]/(f(x)) \cong \bigoplus_{j=1}^{J} A[x]/(g_j(x)) \cong \bigoplus_{j=1}^{J} K_j.$$

Each K_j is of the form $A[x]/(g_j(x))$, so contains an isomorphic image of A. It thus remains to show that the ring homomorphisms

$$\lambda_j : B \xrightarrow{b \mapsto 1 \otimes b} A \otimes_K B \xrightarrow{\varphi_j} K_j$$

are injections. Since B and K_j are both fields, λ_j is either the 0 map or injective. However, λ_j is not the 0 map since $\lambda_j(1) = 1 \in K_j$. $\square$

Example 16.2.2. If A and B are finite extensions of $\mathbb{Q}$, then $A \otimes_{\mathbb{Q}} B$ is an algebra of degree $[A : \mathbb{Q}] \cdot [B : \mathbb{Q}]$. For example, suppose A is generated by a root of $x^2 + 1$ and B is generated by a root of $x^3 - 2$. We can view $A \otimes_{\mathbb{Q}} B$ as either $A[x]/(x^3 - 2)$ or $B[x]/(x^2 + 1)$. The polynomial $x^2 + 1$ is irreducible over $\mathbb{Q}$, and if it factored over the cubic field B, then there would be a root of $x^2 + 1$ in B, i.e., the quadratic field $A = \mathbb{Q}(i)$ would be a subfield of the cubic field $B = \mathbb{Q}(\sqrt[3]{2})$, which is impossible. Thus $x^2 + 1$ is irreducible over B, so $A \otimes_{\mathbb{Q}} B = A.B = \mathbb{Q}(i, \sqrt[3]{2})$ is a degree 6 extension of $\mathbb{Q}$. Notice that $A.B$ contains a copy A and a copy of B. By the primitive element theorem the composite field $A.B$ can be generated by the root of a single polynomial. For example, the minimal polynomial of $i + \sqrt[3]{2}$ is $x^6 + 3x^4 - 4x^3 + 3x^2 + 12x + 5$, hence $\mathbb{Q}(i + \sqrt[3]{2}) = A.B$.

Example 16.2.3. The case $A \cong B$ is even more exciting. For example, suppose $A = B = \mathbb{Q}(i)$. Using the Chinese Remainder Theorem we have that

$$\mathbb{Q}(i) \otimes_{\mathbb{Q}} \mathbb{Q}(i) \cong \mathbb{Q}(i)[x]/(x^2 + 1) \cong \mathbb{Q}(i)[x]/((x - i)(x + i)) \cong \mathbb{Q}(i) \oplus \mathbb{Q}(i),$$

since $(x-i)$ and $(x+i)$ are coprime. The last isomorphism sends $a+bx$, with $a,b \in \mathbb{Q}(i)$, to $(a+bi, a-bi)$. Since $\mathbb{Q}(i) \oplus \mathbb{Q}(i)$ has zero divisors, the tensor product $\mathbb{Q}(i) \otimes_{\mathbb{Q}} \mathbb{Q}(i)$ must also have zero divisors. For example, $(1,0)$ and $(0,1)$ is a zero divisor pair on the right hand side, and we can trace back to the elements of the tensor product that they define. First, by solving the system

$$a + bi = 1 \qquad \text{and} \qquad a - bi = 0$$

we see that $(1,0)$ corresponds to $a = 1/2$ and $b = -i/2$, i.e., to the element

$$\frac{1}{2} - \frac{i}{2}x \in \mathbb{Q}(i)[x]/(x^2+1).$$

This element in turn corresponds to

$$\frac{1}{2} \otimes 1 - \frac{i}{2} \otimes i \in \mathbb{Q}(i) \otimes_{\mathbb{Q}} \mathbb{Q}(i).$$

Similarly the other element $(0,1)$ corresponds to

$$\frac{1}{2} \otimes 1 + \frac{i}{2} \otimes i \in \mathbb{Q}(i) \otimes_{\mathbb{Q}} \mathbb{Q}(i).$$

As a double check, observe that

$$\begin{aligned}\left(\frac{1}{2} \otimes 1 - \frac{i}{2} \otimes i\right) \cdot \left(\frac{1}{2} \otimes 1 + \frac{i}{2} \otimes i\right) &= \frac{1}{4} \otimes 1 + \frac{i}{4} \otimes i - \frac{i}{4} \otimes i - \frac{i^2}{4} \otimes i^2 \\ &= \frac{1}{4} \otimes 1 - \frac{1}{4} \otimes 1 = 0 \in \mathbb{Q}(i) \otimes_{\mathbb{Q}} \mathbb{Q}(i).\end{aligned}$$

Clearing the denominator of 2 and writing $1 \otimes 1 = 1$, we have $(1 - i \otimes i)(1 + i \otimes i) = 0$, so $i \otimes i$ is a root of the polynomimal $x^2 - 1$, and $i \otimes i$ is not ± 1, so $x^2 - 1$ has more than 2 roots.

In general, to understand $A \otimes_K B$ explicitly is the same as factoring either the defining polynomial of B over the field A, or factoring the defining polynomial of A over B.

Corollary 16.2.4. *Let $a \in B$ be any element and let $f(x) \in K[x]$ be the characteristic polynomials of a over K and let $g_j(x) \in A[x]$ (for $1 \leq j \leq J$) be the characteristic polynomials of the images of a under $B \to A \otimes_K B \to K_j$ over A, respectively. Then*

$$f(x) = \prod_{j=1}^{J} g_j(X). \tag{16.2.2}$$

Proof. We show that both sides of (16.2.2) are the characteristic polynomial $T(x)$ of the image of a in $A \otimes_K B$ over A. That $f(x) = T(x)$ follows at once by computing the characteristic polynomial in terms of a basis $\underline{w}_1, \ldots, \underline{w}_N$ of $A \otimes_K B$, where $w_1, \ldots, w_N$ is a basis for B over K (this is because the matrix of left multiplication by b on $A \otimes_K B$ is exactly the same as the matrix of left multiplication on B, so the characteristic polynomial doesn't change). To see that $T(X) = \prod g_j(X)$, compute the action of the image of a in $A \otimes_K B$ with respect to a basis of

$$A \otimes_K B \cong \bigoplus_{j=1}^{J} K_j \tag{16.2.3}$$

composed of basis of the individual extensions K_j of A. The resulting matrix will be a block direct sum of submatrices, each of whose characteristic polynomials is one of the $g_j(X)$. Taking the product gives the claimed identity (16.2.2). □

Corollary 16.2.5. *For $a \in B$ we have*

$$\mathrm{Norm}_{B/K}(a) = \prod_{j=1}^{J} \mathrm{Norm}_{K_j/A}(a),$$

and

$$\mathrm{Tr}_{B/K}(a) = \sum_{j=1}^{J} \mathrm{Tr}_{K_j/A}(a),$$

Proof. This follows from Corollary 16.2.4. First, the norm is $\pm$ the constant term of the characteristic polynomial, and the constant term of the product of polynomials is the product of the constant terms (and one sees that the sign matches up correctly). Second, the trace is minus the second coefficient of the characteristic polynomial, and second coefficients add when one multiplies polynomials:

$$(x^n + a_{n-1}x^{n-1} + \cdots) \cdot (x^m + a_{m-1}x^{m-1} + \cdots) = x^{n+m} + x^{n+m-1}(a_{m-1} + a_{n-1}) + \cdots .$$

One could also see both the statements by considering a matrix of left multiplication by a first with respect to the basis of $\underline{w}_n$ and second with respect to the basis coming from the left side of (16.2.3).

□

Chapter 17

Extensions and Normalizations of Valuations

17.1 Extensions of Valuations

In this section we continue to tacitly assume that all valuations are nontrivial. We do not assume all our valuations satisfy the triangle

Suppose $K \subset L$ is a finite extension of fields, and that $|\cdot|$ and $\|\cdot\|$ are valuations on K and L, respectively.

Definition 17.1.1 (Extends)**.** We say that $\|\cdot\|$ *extends* $|\cdot|$ if $|a| = \|a\|$ for all $a \in K$.

Theorem 17.1.2. *Suppose that K is a field that is complete with respect to $|\cdot|$ and that L is a finite extension of K of degree $N = [L : K]$. Then there is precisely one extension of $|\cdot|$ to K, namely*

$$\|a\| = \left|\mathrm{Norm}_{L/K}(a)\right|^{1/N}, \tag{17.1.1}$$

where the Nth root is the non-negative real Nth root of the nonnegative real number $\left|\mathrm{Norm}_{L/K}(a)\right|$.

Proof. We may assume that $|\cdot|$ is normalized so as to satisfy the triangle inequality. Otherwise, normalize $|\cdot|$ so that it does, prove the theorem for the normalized valuation $|\cdot|^c$, then raise both sides of (17.1.1) to the power $1/c$. In the uniqueness proof, by the same argument we may assume that $\|\cdot\|$ also satisfies the triangle inequality.

Uniqueness. View L as a finite-dimensional vector space over K. Then $\|\cdot\|$ is a norm in the sense defined earlier (Definition 16.1.1). Hence any two extensions $\|\cdot\|_1$ and $\|\cdot\|_2$ of $|\cdot|$ are equivalent as norms, so induce the same topology on K. But as we have seen (Proposition 14.1.4), two valuations which induce the same topology are equivalent valuations, i.e., $\|\cdot\|_1 = \|\cdot\|_2^c$, for some positive real c. Finally $c = 1$ since $\|a\|_1 = |a| = \|a\|_2$ for all $a \in K$.

Existence. We do not give a proof of existence in the general case. Instead we give a proof, which was suggested by Dr. Geyer at the conference out of which [Cas67] arose. It is valid when K is locally compact, which is the only case we will use later.

We see at once that the function defined in (17.1.1) satisfies the condition (i) that $\|a\| \geq 0$ with equality only for $a = 0$, and (ii) $\|ab\| = \|a\| \cdot \|b\|$ for all $a, b \in L$. The difficult part of the proof is to show that there is a constant $C > 0$ such that

$$\|a\| \leq 1 \implies \|1 + a\| \leq C.$$

Note that we do not know (and will not show) that $\|\cdot\|$ as defined by (17.1.1) is a norm as in Definition 16.1.1, since showing that $\|\cdot\|$ is a norm would entail showing that it satisfies the triangle inequality, which is not obvious.

Choose a basis $b_1, \ldots, b_N$ for L over K. Let $\|\cdot\|_0$ be the max norm on L, so for $a = \sum_{i=1}^{N} c_i b_i$ with $c_i \in K$ we have

$$\|a\|_0 = \left\|\sum_{i=1}^{N} c_i b_i\right\|_0 = \max\{|c_i| : i = 1, \ldots, N\}.$$

(Note: in Cassels's original article he let $\|\cdot\|_0$ be *any* norm, but we don't because the rest of the proof does not work, since we can't use homogeneity as he claims to do. This is because it need not be possible to find, for any nonzero $a \in L$ some element $c \in K$ such that $\|ac\|_0 = 1$. This would fail, e.g., if $\|a\|_0 \neq |c|$ for any $c \in K$.) The rest of the argument is very similar to our proof from Lemma 16.1.3 of uniqueness of norms on vector spaces over complete fields.

With respect to the $\|\cdot\|_0$-topology, L has the product topology as a product of copies of K. The function $a \mapsto \|a\|$ is a composition of continuous functions on L with respect to this topology (e.g., $\mathrm{Norm}_{L/K}$ is the determinant, hence polynomial), hence $\|\cdot\|$ defines nonzero continuous function on the compact set

$$S = \{a \in L : \|a\|_0 = 1\}.$$

By compactness, there are real numbers $\delta, \Delta \in \mathbb{R}_{>0}$ such that

$$0 < \delta \leq \|a\| \leq \Delta \qquad \text{for all } a \in S.$$

For any nonzero $a \in L$ there exists $c \in K$ such that $\|a\|_0 = |c|$; to see this take c to be a c_i in the expression $a = \sum_{i=1}^{N} c_i b_i$ with $|c_i| \geq |c_j|$ for any j. Hence $\|a/c\|_0 = 1$, so $a/c \in S$ and

$$0 \leq \delta \leq \frac{\|a/c\|}{\|a/c\|_0} \leq \Delta.$$

Then by homogeneity

$$0 \leq \delta \leq \frac{\|a\|}{\|a\|_0} \leq \Delta.$$

Suppose now that $\|a\| \leq 1$. Then $\|a\|_0 \leq \delta^{-1}$, so

$$\begin{aligned}\|1+a\| &\leq \Delta \cdot \|1+a\|_0 \\ &\leq \Delta \cdot (\|1\|_0 + \|a\|_0) \\ &\leq \Delta \cdot \left(\|1\|_0 + \delta^{-1}\right) \\ &= C \quad \text{(say)},\end{aligned}$$

as required. $\square$

Example 17.1.3. Consider the extension $\mathbb{C}$ of $\mathbb{R}$ equipped with the archimedean valuation. The unique extension is the ordinary absolute value on $\mathbb{C}$:

$$\|x+iy\| = \left(x^2+y^2\right)^{1/2}.$$

Example 17.1.4. Consider the extension $\mathbb{Q}_2(\sqrt{2})$ of $\mathbb{Q}_2$ equipped with the 2-adic absolute value. Since x^2-2 is irreducible over $\mathbb{Q}_2$ we can do some computations by working in the subfield $\mathbb{Q}(\sqrt{2})$ of $\mathbb{Q}_2(\sqrt{2})$.

```
sage: K.<a> = NumberField(x^2 - 2); K
Number Field in a with defining polynomial x^2 - 2
sage: norm = lambda z: math.sqrt(2^(-z.norm().valuation(2)))
sage: norm(1 + a)
1.0
sage: norm(1 + a + 1)
0.70710678118654757
sage: z = 3 + 2*a
sage: norm(z)
1.0
sage: norm(z + 1)
0.35355339059327379
```

Remark 17.1.5. Geyer's existence proof gives (17.1.1). But it is perhaps worth noting that in any case (17.1.1) is a consequence of unique existence, as follows. Suppose L/K is as above. Suppose M is a finite Galois extension of K that contains L. Then by assumption there is a unique extension of $|\cdot|$ to M, which we shall also denote by $\|\cdot\|$. If $\sigma \in \mathrm{Gal}(M/K)$, then

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

is also an extension of $|\cdot|$ to M, so $\|\cdot\|_\sigma = \|\cdot\|$, i.e.,

$$\|\sigma(a)\| = \|a\| \qquad \text{for all } a \in M.$$

But now

$$\mathrm{Norm}_{L/K}(a) = \sigma_1(a) \cdot \sigma_2(a) \cdots \sigma_N(a)$$

for $a \in K$, where $\sigma_1, \ldots, \sigma_N \in \mathrm{Gal}(M/K)$ extend the embeddings of L into M. Hence

$$\begin{aligned} \left|\mathrm{Norm}_{L/K}(a)\right| &= \left\|\mathrm{Norm}_{L/K}(a)\right\| \\ &= \prod_{1 \le n \le N} \|\sigma_n(a)\| \\ &= \|a\|^N, \end{aligned}$$

as required.

Corollary 17.1.6. *Let $w_1, \ldots, w_N$ be a basis for L over K. Then there are positive constants c_1 and c_2 such that*

$$c_1 \le \frac{\left\| \sum_{n=1}^{N} b_n w_n \right\|}{\max\{|b_n| : n = 1, \ldots, N\}} \le c_2$$

for any $b_1, \ldots, b_N \in K$ not all 0.

Proof. For $\left|\sum_{n=1}^{N} b_n w_n\right|$ and $\max |b_n|$ are two norms on L considered as a vector space over K.

I don't believe this proof, which I copied from Cassels's article. My problem with it is that the proof of Theorem 17.1.2 does not give that $C \le 2$, i.e., that the triangle inequality holds for $\|\cdot\|$. By changing the basis for L/K one can make any nonzero vector $a \in L$ have $\|a\|_0 = 1$, so if we choose a such that $|a|$ is very large, then the Δ in the proof will also be very large. One way to fix the corollary is to only claim that there are positive constants c_1, c_2, c_3, c_4 such that

$$c_1 \le \frac{\left\| \sum_{n=1}^{N} b_n w_n \right\|^{c_3}}{\max\{|b_n|^{c_4} : n = 1, \ldots, N\}} \le c_2.$$

Then choose c_3, c_4 such that $\|\cdot\|^{c_3}$ and $|\cdot|^{c_4}$ satisfies the triangle inequality, and prove the modified corollary using the proof suggested by Cassels. $\square$

Corollary 17.1.7. *A finite extension of a completely valued field K is complete with respect to the extended valuation.*

Proof. By the proceeding corollary it has the topology of a finite-dimensional vector space over K. (The problem with the proof of the previous corollary is not an issue, because we can replace the extended valuation by an inequivalent one that satisfies the triangle inequality and induces the same topology.) $\square$

When K is no longer complete under $|\cdot|$ the position is more complicated:

Theorem 17.1.8. *Let L be a separable extension of K of finite degree $N = [L : K]$. Then there are at most N extensions of a valuation $|\cdot|$ on K to L, say $\|\cdot\|_j$, for $1 \leq j \leq J$. Let K_v be the completion of K with respect to $|\cdot|$, and for each j let L_j be the completion of L with respect to $\|\cdot\|_j$. Then*

$$K_v \otimes_K L \cong \bigoplus_{1\leq j\leq J} L_j \tag{17.1.2}$$

algebraically and topologically, where the right hand side is given the product topology.

Proof. We already know (Lemma 16.2.1) that $K_v \otimes_K L$ is of the shape (17.1.2), where the L_j are finite extensions of K_v. Hence there is a unique extension $|\cdot|_j^*$ of $|\cdot|$ to the L_j, and by Corollary 17.1.7 the L_j are complete with respect to the extended valuation. Further, the ring homomorphisms

$$\lambda_j : L \to K_v \otimes_K L \to L_j$$

are injections. Hence we get an extension $\|\cdot\|_j$ of $|\cdot|$ to L by putting

$$\|b\|_j = |\lambda_j(b)|_j^* .$$

Further, $L \cong \lambda_j(L)$ is dense in L_j with respect to $\|\cdot\|_j$ because $L = K \otimes_K L$ is dense in $K_v \otimes_K L$ (since K is dense in K_v). Hence L_j is exactly the completion of L.

It remains to show that the $\|\cdot\|_j$ are distinct and that they are the only extensions of $|\cdot|$ to L.

Suppose $\|\cdot\|$ is any valuation of L that extends $|\cdot|$. Then $\|\cdot\|$ extends by continuity to a real-valued function on $K_v \otimes_K L$, which we also denote by $\|\cdot\|$. (We are again using that L is dense in $K_v \otimes_K L$.) By continuity we have for all $a, b \in K_v \otimes_K L$,

$$\|ab\| = \|a\| \cdot \|b\|$$

and if C is the constant in axiom (iii) for L and $\|\cdot\|$, then

$$\|a\| \leq 1 \implies \|1 + a\| \leq C.$$

(In Cassels, he inexplicable assume that $C = 1$ at this point in the proof.)

We consider the restriction of $\|\cdot\|$ to one of the L_j. If $\|a\| \neq 0$ for some $a \in L_j$, then $\|a\| = \|b\| \cdot \left\|ab^{-1}\right\|$ for every $b \neq 0$ in L_j so $\|b\| \neq 0$. Hence either $\|\cdot\|$ is identically 0 on L_j or it induces a valuation on L_j.

Further, $\|\cdot\|$ cannot induce a valuation on two of the L_j. For

$$(a_1, 0, \ldots, 0) \cdot (0, a_2, 0, \ldots, 0) = (0, 0, 0, \ldots, 0),$$

so for any $a_1 \in L_1$, $a_2 \in L_2$,

$$\|a_1\| \cdot \|a_2\| = 0.$$

Hence $\|\cdot\|$ induces a valuation in precisely one of the L_j, and it extends the given valuation $|\cdot|$ of K_v. Hence $\|\cdot\| = \|\cdot\|_j$ for precisely one j.

It remains only to show that (17.1.2) is a topological homomorphism. For

$$(b_1, \ldots, b_J) \in L_1 \oplus \cdots \oplus L_J$$

put

$$\|(b_1, \ldots, b_J)\|_0 = \max_{1 \le j \le J} \|b_j\|_j \,.$$

Then $\|\cdot\|_0$ is a norm on the right hand side of (17.1.2), considered as a vector space over K_v and it induces the product topology. On the other hand, any two norms are equivalent, since K_v is complete, so $\|\cdot\|_0$ induces the tensor product topology on the left hand side of (17.1.2). □

Corollary 17.1.9. *Suppose $L = K(a)$, and let $f(x) \in K[x]$ be the minimal polynomial of a. Suppose that*

$$f(x) = \prod_{1 \le j \le J} g_j(x)$$

in $K_v[x]$, where the g_j are irreducible. Then $L_j = K_v(b_j)$, where b_j is a root of g_j.

17.2 Extensions of Normalized Valuations

Let K be a complete field with valuation $|\cdot|$. We consider the following three cases:

(1) $|\cdot|$ is discrete non-archimedean and the residue class field is finite.

(2i) The completion of K with respect to $|\cdot|$ is $\mathbb{R}$.

(2ii) The completion of K with respect to $|\cdot|$ is $\mathbb{C}$.

(Alternatively, these cases can be subsumed by the hypothesis that the completion of K is locally compact.)

In case (1) we defined the normalized valuation to be the one such that if Haar measure of the ring of integers $\mathcal{O}$ is 1, then $\mu(a\mathcal{O}) = |a|$ (see Definition 15.1.11). In case (2i) we say that $|\cdot|$ is normalized if it is the ordinary absolute value, and in (2ii) if it is the *square* of the ordinary absolute value:

$$|x + iy| = x^2 + y^2 \qquad \text{(normalized)}.$$

In every case, for every $a \in K$, the map

$$a : x \mapsto ax$$

on K^+ multiplies any choice of Haar measure by $|a|$, and this characterizes the normalized valuations among equivalent ones.

We have already verified the above characterization for non-archimedean valuations, and it is clear for the ordinary absolute value on $\mathbb{R}$, so it remains to verify it for $\mathbb{C}$. The additive group $\mathbb{C}^+$ is topologically isomorphic to $\mathbb{R}^+ \oplus \mathbb{R}^+$, so a choice of Haar measure of $\mathbb{C}^+$ is the usual area measure on the Euclidean plane. Multiplication by $x + iy \in \mathbb{C}$ is the same as rotation followed by scaling by a factor of $\sqrt{x^2+y^2}$, so if we rescale a region by a factor of $x + iy$, the area of the region changes by a factor of the square of $\sqrt{x^2+y^2}$. This explains why the normalized valuation on $\mathbb{C}$ is the square of the usual absolute value. Note that the normalized valuation on $\mathbb{C}$ does not satisfy the triangle inequality:

$$|1+(1+i)| = |2+i| = 2^2+1^2 = 5 \not\leq 3 = 1^2 + (1^2+1^2) = |1| + |1+i|\,.$$

The constant C in axiom (3) of a valuation for the ordinary absolute value on $\mathbb{C}$ is 2, so the constant for the normalized valuation $|\cdot|$ is $C \leq 4$:

$$|x+iy| \leq 1 \implies |x+iy+1| \leq 4.$$

Note that $x^2+y^2 \leq 1$ implies

$$(x+1)^2 + y^2 = x^2 + 2x + 1 + y^2 \leq 1 + 2x + 1 \leq 4$$

since $x \leq 1$.

Lemma 17.2.1. *Suppose K is a field that is complete with respect to a normalized valuation $|\cdot|$ and let L be a finite extension of K of degree $N = [L : K]$. Then the normalized valuation $\|\cdot\|$ on L which is equivalent to the unique extension of $|\cdot|$ to L is given by the formula*

$$\|a\| = \left|\mathrm{Norm}_{L/K}(a)\right| \qquad \textit{all } a \in L. \tag{17.2.1}$$

Proof. Let $\|\cdot\|$ be the normalized valuation on L that extends $|\cdot|$. Our goal is to identify $\|\cdot\|$, and in particular to show that it is given by (17.2.1).

By the preceding section there is a positive real number c such that for all $a \in L$ we have

$$\|a\| = \left|\mathrm{Norm}_{L/K}(a)\right|^c.$$

Thus all we have to do is prove that $c = 1$. In case 2 the only nontrivial situation is $L = \mathbb{C}$ and $K = \mathbb{R}$, in which case $\left|\mathrm{Norm}_{\mathbb{C}/\mathbb{R}}(x+iy)\right| = \left|x^2+y^2\right|$, which is the normalized valuation on $\mathbb{C}$ defined above.

One can argue in a unified way in all cases as follows. Let $w_1, \ldots, w_N$ be a basis for L/K. Then the map

$$\varphi : L^+ \to \bigoplus_{n=1}^{N} K^+, \qquad \sum a_n w_n \mapsto (a_1, \ldots, a_N)$$

is an isomorphism between the additive group L^+ and the direct sum $\oplus_{n=1}^{N} K^+$, and this is a homeomorphism if the right hand side is given the product topology. In particular, the Haar measures on L^+ and on $\oplus_{n=1}^{N} K^+$ are the same up to a multiplicative constant in $\mathbb{Q}^*$.

Let $b \in K$. Then the left-multiplication-by-b map

$$b : \sum a_n w_n \mapsto \sum b a_n w_n$$

on L^+ is the same as the map

$$(a_1, \ldots, a_N) \mapsto (ba_1, \ldots, ba_N)$$

on $\oplus_{n=1}^N K^+$, so it multiplies the Haar measure by $|b|^N$, since $|\cdot|$ on K is assumed normalized (the measure of each factor is multiplied by $|b|$, so the measure on the product is multiplied by $|b|^N$). Since $\|\cdot\|$ is assumed normalized, so multiplication by b rescales by $\|b\|$, we have

$$\|b\| = |b|^N .$$

But $b \in K$, so $\mathrm{Norm}_{L/K}(b) = b^N$. Since $|\cdot|$ is nontrivial and for $a \in K$ we have

$$\|a\| = |a|^N = \left|a^N\right| = \left|\mathrm{Norm}_{L/K}(a)\right| ,$$

so we must have $c = 1$ in (17.2.1), as claimed. □

In the case when K need not be complete with respect to the valuation $|\cdot|$ on K, we have the following theorem.

Theorem 17.2.2. *Suppose $|\cdot|$ is a (nontrivial as always) normalized valuation of a field K and let L be a finite extension of K. Then for any $a \in L$,*

$$\prod_{1 \le j \le J} \|a\|_j = \left|\mathrm{Norm}_{L/K}(a)\right|$$

where the $\|\cdot\|_j$ are the normalized valuations equivalent to the extensions of $|\cdot|$ to K.

Proof. Let K_v denote the completion of K with respect to $|\cdot|$. Write

$$K_v \otimes_K L = \bigoplus_{1 \le j \le J} L_j .$$

Then Theorem 17.2.2 asserts that

$$\mathrm{Norm}_{L/K}(a) = \prod_{1 \le j \le J} \mathrm{Norm}_{L_j/K_v}(a). \tag{17.2.2}$$

By Theorem 17.1.8, the $\|\cdot\|_j$ are exactly the normalizations of the extensions of $|\cdot|$ to the L_j (i.e., the L_j are in bijection with the extensions of valuations, so there are no other valuations missed). By Lemma 17.1.1, the normalized valuation $\|\cdot\|_j$ on L_j is $|a| = \left|\mathrm{Norm}_{L_J/K_v}(a)\right|$. The theorem now follows by taking absolute values of both sides of (17.2.2). □

What next?! We'll building up to giving a new proof of finiteness of the class group that uses that the class group naturally has the discrete topology and is the continuous image of a compact group.

Chapter 18

Global Fields and Adeles

18.1 Global Fields

Definition 18.1.1 (Global Field)**.** A *global field* is a number field or a finite separable extension of $\mathbb{F}(t)$, where $\mathbb{F}$ is a finite field, and t is transcendental over $\mathbb{F}$.

In this chapter, we will focus attention on number fields, and leave the function field case to the reader.

The following lemma essentially says that the denominator of an element of a global field is only "nontrivial" at a finite number of valuations.

Lemma 18.1.2. *Let $a \in K$ be a nonzero element of a global field K. Then there are only finitely many inequivalent valuations $|\cdot|$ of K for which*

$$|a| > 1.$$

Proof. If $K = \mathbb{Q}$ or $\mathbb{F}(t)$ then the lemma follows by Ostrowski's classification of all the valuations on K (see Theorem 13.3.2). For example, when $a = \frac{n}{d} \in \mathbb{Q}$, with $n, d \in \mathbb{Z}$, then the valuations where we could have $|a| > 1$ are the archimedean one, or the p-adic valuations $|\cdot|_p$ for which $p \mid d$.

Suppose now that K is a finite extension of $\mathbb{Q}$, so a satisfies a monic polynomial

$$a^n + c_{n-1}a^{n-1} + \cdots + c_0 = 0,$$

for some n and $c_0, \ldots, c_{n-1} \in \mathbb{Q}$. If $|\cdot|$ is a non-archimedean valuation on K, we have

$$\begin{aligned} |a|^n &= \left|-(c_{n-1}a^{n-1} + \cdots + c_0)\right| \\ &\leq \max(1, |a|^{n-1}) \cdot \max(|c_0|, \ldots, |c_{n-1}|). \end{aligned}$$

Dividing each side by $|a|^{n-1}$, we have that

$$|a| \leq \max(|c_0|, \ldots, |c_{n-1}|),$$

so in all cases we have

$$|a| \leq \max(1, |c_0|, \ldots, |c_{n-1}|)^{1/(n-1)}. \tag{18.1.1}$$

We know the lemma for $\mathbb{Q}$, so there are only finitely many valuations $|\cdot|$ on $\mathbb{Q}$ such that the right hand side of (18.1.1) is bigger than 1. Since each valuation of $\mathbb{Q}$ has finitely many extensions to K, and there are only finitely many archimedean valuations, it follows that there are only finitely many valuations on K such that $|a| > 1$. □

Any valuation on a global field is either archimedean, or discrete non-archimedean with finite residue class field, since this is true of $\mathbb{Q}$ and $\mathbb{F}(t)$ and is a property preserved by extending a valuation to a finite extension of the base field. Hence it makes sense to talk of normalized valuations. Recall that the normalized p-adic valuation on $\mathbb{Q}$ is $|x|_p = p^{-\operatorname{ord}_p(x)}$, and if v is a valuation on a number field K equivalent to an extension of $|\cdot|_p$, then the normalization of v is the composite of the sequence of maps

$$K \hookrightarrow K_v \xrightarrow{\text{Norm}} \mathbb{Q}_p \xrightarrow{|\cdot|_p} \mathbb{R},$$

where K_v is the completion of K at v.

Example 18.1.3. Let $K = \mathbb{Q}(\sqrt{2})$, and let $p = 2$. Because $\sqrt{2} \notin \mathbb{Q}_2$, there is exactly one extension of $|\cdot|_2$ to K, and it sends $a = 1/\sqrt{2}$ to

$$\left|\text{Norm}_{\mathbb{Q}_2(\sqrt{2})/\mathbb{Q}_2}(1/\sqrt{2})\right|_2^{1/2} = \sqrt{2}.$$

Thus the normalized valuation of a is 2.

There are two extensions of $|\cdot|_7$ to $\mathbb{Q}(\sqrt{2})$, since $\mathbb{Q}(\sqrt{2}) \otimes_{\mathbb{Q}} \mathbb{Q}_7 \cong \mathbb{Q}_7 \oplus \mathbb{Q}_7$, as $x^2 - 2 = (x-3)(x-4) \pmod 7$. The image of $\sqrt{2}$ under each embedding into $\mathbb{Q}_7$ is a unit in $\mathbb{Z}_7$, so the normalized valuation of $a = 1/\sqrt{2}$ is, in both cases, equal to 1. More generally, for any valuation of K of characteristic an odd prime p, the normalized valuation of a is 1.

Since $K = \mathbb{Q}(\sqrt{2}) \hookrightarrow \mathbb{R}$ in two ways, there are exactly two normalized archimedean valuations on K, and both of their values on a equal $1/\sqrt{2}$. Notice that the product of the absolute values of a with respect to all normalized valuations is

$$2 \cdot \frac{1}{\sqrt{2}} \cdot \frac{1}{\sqrt{2}} \cdot 1 \cdot 1 \cdot 1 \cdots = 1.$$

This "product formula" holds in much more generality, as we will now see.

Theorem 18.1.4 (Product Formula). *Let $a \in K$ be a nonzero element of a global field K. Let $|\cdot|_v$ run through the normalized valuations of K. Then $|a|_v = 1$ for almost all v, and*

$$\prod_{\text{all } v} |a|_v = 1 \qquad \text{(the product formula).}$$

We will later give a more conceptual proof of this using Haar measure (see Remark 18.3.9).

Proof. By Lemma 18.1.2, we have $|a|_v \leq 1$ for almost all v. Likewise, $1/|a|_v = |1/a|_v \leq 1$ for almost all v, so $|a|_v = 1$ for almost all v.

Let w run through all normalized valuations of $\mathbb{Q}$ (or of $\mathbb{F}(t)$), and write $v \mid w$ if the restriction of v to $\mathbb{Q}$ is equivalent to w. Then by Theorem 17.2.2,

$$\prod_v |a|_v = \prod_w \left(\prod_{v|w} |a|_v\right) = \prod_w \left|\mathrm{Norm}_{K/\mathbb{Q}}(a)\right|_w,$$

so it suffices to prove the theorem for $K = \mathbb{Q}$.

By multiplicativity of valuations, if the theorem is true for b and c then it is true for the product bc and quotient b/c (when $c \neq 0$). The theorem is clearly true for -1, which has valuation 1 at all valuations. Thus to prove the theorem for $\mathbb{Q}$ it suffices to prove it when $a = p$ is a prime number. Then we have $|p|_\infty = p$, $|p|_p = 1/p$, and for primes $q \neq p$ that $|p|_q = 1$. Thus

$$\prod_v |p|_v = p \cdot \frac{1}{p} \cdot 1 \cdot 1 \cdot 1 \cdots = 1,$$

as claimed. $\square$

If v is a valuation on a field K, recall that we let K_v denote the completion of K with respect to v. Also when v is non-archimedean, let

$$\mathcal{O}_v = \mathcal{O}_{K,v} = \{x \in K_v : |x| \leq 1\}$$

be the ring of integers of the completion.

Definition 18.1.5 (Almost All)**.** We say a condition holds for *almost all* elements of a set if it holds for all but finitely many elements.

We will use the following lemma later (see Lemma 18.3.3) to prove that formation of the adeles of a global field is compatible with base change.

Lemma 18.1.6. *Let $\omega_1, \ldots, \omega_n$ be a basis for L/K, where L is a finite separable extension of the global field K of degree n. Then for almost all normalized non-archimedean valuations v on K we have*

$$\omega_1\mathcal{O}_v \oplus \cdots \oplus \omega_n\mathcal{O}_v = \mathcal{O}_{w_1} \oplus \cdots \oplus \mathcal{O}_{w_g} \subset K_v \otimes_K L, \tag{18.1.2}$$

where $w_1, \ldots, w_g$ are the extensions of v to L. Here we have identified $a \in L$ with its canonical image in $K_v \otimes_K L$, and the direct sum on the left is the sum taken inside the tensor product (so directness means that the intersections are trivial).

Proof. The proof proceeds in two steps. First we deduce easily from Lemma 18.1.2 that for almost all v the left hand side of (18.1.2) is contained in the right hand side. Then we use a trick involving discriminants to show the opposite inclusion for all but finitely many primes.

Since $\mathcal{O}_v \subset \mathcal{O}_{w_i}$ for all i, the left hand side of (18.1.2) is contained in the right hand side if $|\omega_i|_{w_j} \leq 1$ for $1 \leq i \leq n$ and $1 \leq j \leq g$. Thus by Lemma 18.1.2, for all but finitely many v the left hand side of (18.1.2) is contained in the right hand side. We have just eliminated the finitely many primes corresponding to "denominators" of some ω_i, and now only consider v such that $\omega_1, \ldots, \omega_n \in \mathcal{O}_w$ for all $w \mid v$.

For any elements $a_1, \ldots, a_n \in K_v \otimes_K L$, consider the discriminant

$$D(a_1, \ldots, a_n) = \det(\mathrm{Tr}(a_i a_j)) \in K_v,$$

where the trace is induced from the L/K trace. Since each ω_i is in each $\mathcal{O}_w$, for $w \mid v$, the traces lie in $\mathcal{O}_v$, so

$$d = D(\omega_1, \ldots, \omega_n) \in \mathcal{O}_v.$$

Also note that $d \in K$ since each ω_i is in L. Now suppose that

$$\alpha = \sum_{i=1}^{n} a_i \omega_i \in \mathcal{O}_{w_1} \oplus \cdots \oplus \mathcal{O}_{w_g},$$

with $a_i \in K_v$. Then by properties of determinants for any m with $1 \leq m \leq n$, we have

$$D(\omega_1, \ldots, \omega_{m-1}, \alpha, \omega_{m+1}, \ldots, \omega_n) = a_m^2 D(\omega_1, \ldots, \omega_n). \tag{18.1.3}$$

The left hand side of (18.1.3) is in $\mathcal{O}_v$, so the right hand side is well, i.e.,

$$a_m^2 \cdot d \in \mathcal{O}_v, \qquad (\text{for } m = 1, \ldots, n),$$

where $d \in K$. Since $\omega_1, \ldots, \omega_n$ are a basis for L over K and the trace pairing is nondegenerate, we have $d \neq 0$, so by Theorem 18.1.4 we have $|d|_v = 1$ for all but finitely many v. Then for all but finitely many v we have that $a_m^2 \in \mathcal{O}_v$. For these v, that $a_m^2 \in \mathcal{O}_v$ implies $a_m \in \mathcal{O}_v$ since $a_m \in K_v$, i.e., α is in the left hand side of (18.1.2). □

Example 18.1.7. Let $K = \mathbb{Q}$ and $L = \mathbb{Q}(\sqrt{2})$. Let $\omega_1 = 1/3$ and $\omega_2 = 2\sqrt{2}$. In the first stage of the above proof we would eliminate $|\cdot|_3$ because ω_2 is not integral at 3. The discriminant is

$$d = D\left(\frac{1}{3}, 2\sqrt{2}\right) = \det\begin{pmatrix} \frac{2}{9} & 0 \\ 0 & 16 \end{pmatrix} = \frac{32}{9}.$$

As explained in the second part of the proof, as long as $v \neq 2, 3$, we have equality of the left and right hand sides in (18.1.2).

18.2 Restricted Topological Products

In this section we describe a topological tool, which we need in order to define adeles (see Definition 18.3.1).

Definition 18.2.1 (Restricted Topological Products)**.** Let X_λ, for $\lambda \in \Lambda$, be a family of topological spaces, and for almost all λ let $Y_\lambda \subset X_\lambda$ be an open subset of X_λ. Consider the space X whose elements are sequences $\mathbf{x} = \{x_\lambda\}_{\lambda \in \Lambda}$, where $x_\lambda \in X_\lambda$ for every λ, and $x_\lambda \in Y_\lambda$ for almost all λ. We give X a topology by taking as a basis of open sets the sets $\prod U_\lambda$, where $U_\lambda \subset X_\lambda$ is open for all λ, and $U_\lambda = Y_\lambda$ for almost all λ. We call X with this topology the *restricted topological product* of the X_λ with respect to the Y_λ.

Corollary 18.2.2. *Let S be a finite subset of Λ, and let X_S be the set of $\mathbf{x} \in X$ with $x_\lambda \in Y_\lambda$ for all $\lambda \notin S$, i.e.,*

$$X_S = \prod_{\lambda \in S} X_\lambda \times \prod_{\lambda \notin S} Y_\lambda \subset X.$$

Then X_S is an open subset of X, and the topology induced on X_S as a subset of X is the same as the product topology.

The restricted topological product depends on the totality of the Y_λ, but not on the individual Y_λ:

Lemma 18.2.3. *Let $Y'_\lambda \subset X_\lambda$ be open subsets, and suppose that $Y_\lambda = Y'_\lambda$ for almost all λ. Then the restricted topological product of the X_λ with respect to the Y'_λ is canonically isomorphic to the restricted topological product with respect to the Y_λ.*

Lemma 18.2.4. *Suppose that the X_λ are locally compact and that the Y_λ are compact. Then the restricted topological product X of the X_λ is locally compact.*

Proof. For any finite subset S of Λ, the open subset $X_S \subset X$ is locally compact, because by Lemma 18.2.2 it is a product of finitely many locally compact sets with an infinite product of compact sets. (Here we are using Tychonoff's theorem from topology, which asserts that an arbitrary product of compact topological spaces is compact (see Munkres's *Topology, a first course*, chapter 5).) Since $X = \cup_S X_S$, and the X_S are open in X, the result follows. $\square$

The following measure will be extremely important in deducing topological properties of the ideles, which will be used in proving finiteness of class groups. See, e.g., the proof of Lemma 18.4.1, which is a key input to the proof of strong approximation (Theorem 18.4.4).

Definition 18.2.5 (Product Measure)**.** For all $\lambda \in \Lambda$, suppose μ_λ is a measure on X_λ with $\mu_\lambda(Y_\lambda) = 1$ when Y_λ is defined. We define the *product measure* μ on X to be that for which a basis of measurable sets is

$$\prod_\lambda M_\lambda$$

where each $M_\lambda \subset X_\lambda$ has finite μ_λ-measure and $M_\lambda = Y_\lambda$ for almost all λ, and where

$$\mu\left(\prod_\lambda M_\lambda\right) = \prod_\lambda \mu_\lambda(M_\lambda).$$

18.3 The Adele Ring

Let K be a global field. For each normalized valuation $|\cdot|_v$ of K, let K_v denote the completion of K. If $|\cdot|_v$ is non-archimedean, let $\mathcal{O}_v$ denote the ring of integers of K_v.

Definition 18.3.1 (Adele Ring)**.** The *adele ring* $\mathbb{A}_K$ of K is the topological ring whose underlying topological space is the restricted topological product of the K_v with respect to the $\mathcal{O}_v$, and where addition and multiplication are defined componentwise:

$$(\mathbf{xy})_v = \mathbf{x}_v\mathbf{y}_v \qquad (\mathbf{x}+\mathbf{y})_v = \mathbf{x}_v + \mathbf{y}_v \qquad \text{for } \mathbf{x},\mathbf{y} \in \mathbb{A}_K. \tag{18.3.1}$$

It is readily verified that (i) this definition makes sense, i.e., if $\mathbf{x},\mathbf{y} \in \mathbb{A}_K$, then $\mathbf{xy}$ and $\mathbf{x}+\mathbf{y}$, whose components are given by (18.3.1), are also in $\mathbb{A}_K$, and (ii) that addition and multiplication are continuous in the $\mathbb{A}_K$-topology, so $\mathbb{A}_K$ is a topological ring, as asserted. Also, Lemma 18.2.4 implies that $\mathbb{A}_K$ is locally compact because the K_v are locally compact (Corollary 15.1.6), and the $\mathcal{O}_v$ are compact (Theorem 15.1.4).

There is a natural continuous ring inclusion

$$K \hookrightarrow \mathbb{A}_K \tag{18.3.2}$$

that sends $x \in K$ to the adele every one of whose components is x. This is an adele because $x \in \mathcal{O}_v$ for almost all v, by Lemma 18.1.2. The map is injective because each map $K \to K_v$ is an inclusion.

Definition 18.3.2 (Principal Adeles)**.** The image of (18.3.2) is the ring of *principal adeles.*

It will cause no trouble to identify K with the principal adeles, so we shall speak of K as a subring of $\mathbb{A}_K$.

Formation of the adeles is compatibility with base change, in the following sense.

Lemma 18.3.3. *Suppose L is a finite (separable) extension of the global field K. Then*

$$\mathbb{A}_K \otimes_K L \cong \mathbb{A}_L \tag{18.3.3}$$

both algebraically and topologically. Under this isomorphism,

$$L \cong K \otimes_K L \subset \mathbb{A}_K \otimes_K L$$

maps isomorphically onto $L \subset \mathbb{A}_L$.

Proof. Let $\omega_1, \ldots, \omega_n$ be a basis for L/K and let v run through the normalized valuations on K. The left hand side of (18.3.3), with the tensor product topology, is the restricted product of the tensor products

$$K_v \otimes_K L \cong K_v \cdot \omega_1 \oplus \cdots \oplus K_v \cdot \omega_n$$

with respect to the integers

$$\mathcal{O}_v \cdot \omega_1 \oplus \cdots \oplus \mathcal{O}_v \cdot \omega_n. \tag{18.3.4}$$

(An element of the left hand side is a finite linear combination $\sum \mathbf{x}_i \otimes a_i$ of adeles $\mathbf{x}_i \in \mathbb{A}_K$ and coefficients $a_i \in L$, and there is a natural isomorphism from the ring of such formal sums to the restricted product of the $K_v \otimes_K L$.)

We proved before (Theorem 17.1.8) that

$$K_v \otimes_K L \cong L_{w_1} \oplus \cdots \oplus L_{w_g},$$

where $w_1, \ldots, w_g$ are the normalizations of the extensions of v to L. Furthermore, as we proved using discriminants (see Lemma 18.1.6), the above identification identifies (18.3.4) with

$$\mathcal{O}_{L_{w_1}} \oplus \cdots \oplus \mathcal{O}_{L_{w_g}},$$

for almost all v. Thus the left hand side of (18.3.3) is the restricted product of the $L_{w_1} \oplus \cdots \oplus L_{w_g}$ with respect to the $\mathcal{O}_{L_{w_1}} \oplus \cdots \oplus \mathcal{O}_{L_{w_g}}$. But this is canonically isomorphic to the restricted product of all completions L_w with respect to $\mathcal{O}_w$, which is the right hand side of (18.3.3). This establishes an isomorphism between the two sides of (18.3.3) as topological spaces. The map is also a ring homomorphism, so the two sides are algebraically isomorphic, as claimed. □

Corollary 18.3.4. *Let $\mathbb{A}_K^+$ denote the topological group obtained from the additive structure on $\mathbb{A}_K$. Suppose L is a finite seperable extension of K. Then*

$$\mathbb{A}_L^+ = \mathbb{A}_K^+ \oplus \cdots \oplus \mathbb{A}_K^+, \qquad ([L:K] \text{ summands}).$$

In this isomorphism the additive group $L^+ \subset \mathbb{A}_L^+$ of the principal adeles is mapped isomorphically onto $K^+ \oplus \cdots \oplus K^+$.

Proof. For any nonzero $\omega \in L$, the subgroup $\omega \cdot \mathbb{A}_K^+$ of $\mathbb{A}_L^+$ is isomorphic as a topological group to $\mathbb{A}_K^+$ (the isomorphism is multiplication by $1/\omega$). By Lemma 18.3.3, we have isomorphisms

$$\mathbb{A}_L^+ = \mathbb{A}_K^+ \otimes_K L \cong \omega_1 \cdot \mathbb{A}_K^+ \oplus \cdots \oplus \omega_n \cdot \mathbb{A}_K^+ \cong \mathbb{A}_K^+ \oplus \cdots \oplus \mathbb{A}_K^+.$$

If $a \in L$, write $a = \sum b_i \omega_i$, with $b_i \in K$. Then a maps via the above map to

$$x = (\omega_1 \cdot \{b_1\}, \ldots, \omega_n \cdot \{b_n\}),$$

where $\{b_i\}$ denotes the principal adele defined by b_i. Under the final map, x maps to the tuple

$$(b_1, \ldots, b_n) \in K \oplus \cdots \oplus K \subset \mathbb{A}_K^+ \oplus \cdots \oplus \mathbb{A}_K^+.$$

The dimensions of L and of $K \oplus \cdots \oplus K$ over K are the same, so this proves the final claim of the corollary. $\square$

Theorem 18.3.5. *The global field K is discrete in $\mathbb{A}_K$ and the quotient $\mathbb{A}_K^+/K^+$ of additive groups is compact in the quotient topology.*

At this point Cassels remarks

> "It is impossible to conceive of any other uniquely defined topology on K. This metamathematical reason is more persuasive than the argument that follows!"

Proof. Corollary 18.3.4, with K for L and $\mathbb{Q}$ or $\mathbb{F}(t)$ for K, shows that it is enough to verify the theorem for $\mathbb{Q}$ or $\mathbb{F}(t)$, and we shall do it here for $\mathbb{Q}$.

To show that $\mathbb{Q}^+$ is discrete in $\mathbb{A}_\mathbb{Q}^+$ it is enough, because of the group structure, to find an open set U that contains $0 \in \mathbb{A}_\mathbb{Q}^+$, but which contains no other elements of $\mathbb{Q}^+$. (If $\alpha \in \mathbb{Q}^+$, then $U + \alpha$ is an open subset of $\mathbb{A}_\mathbb{Q}^+$ whose intersection with $\mathbb{Q}^+$ is $\{\alpha\}$.) We take for U the set of $\mathbf{x} = \{x_v\}_v \in \mathbb{A}_\mathbb{Q}^+$ with

$$|x_\infty|_\infty < 1 \qquad \text{and} \qquad |x_p|_p \leq 1 \quad (\text{all } p),$$

where $|\cdot|_p$ and $|\cdot|_\infty$ are respectively the p-adic and the usual archimedean absolute values on $\mathbb{Q}$. If $b \in \mathbb{Q} \cap U$, then in the first place $b \in \mathbb{Z}$ because $|b|_p \leq 1$ for all p, and then $b = 0$ because $|b|_\infty < 1$. This proves that K^+ is discrete in $\mathbb{A}_\mathbb{Q}^+$. (If we leave out one valuation, as we will see later (Theorem 18.4.4), this theorem is false—what goes wrong with the proof just given?)

Next we prove that the quotient $\mathbb{A}_\mathbb{Q}^+/\mathbb{Q}^+$ is compact. Let $W \subset \mathbb{A}_\mathbb{Q}^+$ consist of the $\mathbf{x} = \{x_v\}_v \in \mathbb{A}_\mathbb{Q}^+$ with

$$|x_\infty|_\infty \leq \frac{1}{2} \qquad \text{and} \qquad |x_p|_p \leq 1 \qquad \text{for all primes } p.$$

We show that every adele $\mathbf{y} = \{y_v\}_v$ is of the form

$$\mathbf{y} = a + \mathbf{x}, \qquad a \in \mathbb{Q}, \quad \mathbf{x} \in W,$$

which will imply that the compact set W maps surjectively onto $\mathbb{A}_{\mathbb{Q}}^{+}/\mathbb{Q}^{+}$. Fix an adele $\mathbf{y} = \{y_v\} \in \mathbb{A}_{\mathbb{Q}}^{+}$. Since $\mathbf{y}$ is an adele, for each prime p we can find a rational number

$$r_p = \frac{z_p}{p^{n_p}} \qquad \text{with} \quad z_p \in \mathbb{Z} \quad \text{and} \quad n_p \in \mathbb{Z}_{\geq 0}$$

such that

$$|y_p - r_p|_p \leq 1,$$

and

$$r_p = 0 \qquad \text{almost all } p.$$

More precisely, for the finitely many p such that

$$y_p = \sum_{n \geq -|s|} a_n p^n \notin \mathbb{Z}_p,$$

choose r_p to be a rational number that is the value of an appropriate truncation of the p-adic expansion of y_p, and when $y_p \in \mathbb{Z}_p$ just choose $r_p = 0$. Hence $r = \sum_p r_p \in \mathbb{Q}$ is well defined. The r_q for $q \neq p$ do not mess up the inequality $|y_p - r|_p \leq 1$ since the valuation $|\cdot|_p$ is non-archimedean and the r_q do not have any p in their denominator:

$$|y_p - r|_p = \left| y_p - r_p - \sum_{q \neq p} r_q \right|_p \leq \max \left(|y_p - r_p|_p \,, \left| \sum_{q \neq p} r_q \right|_p \right) \leq \max(1, 1) = 1.$$

Now choose $s \in \mathbb{Z}$ such that

$$|b_\infty - r - s| \leq \frac{1}{2}.$$

Then $a = r + s$ and $\mathbf{x} = \mathbf{y} - a$ do what is required, since $\mathbf{y} - a = \mathbf{y} - r - s$ has the desired property (since $s \in \mathbb{Z}$ and the p-adic valuations are non-archimedean).

Hence the continuous map $W \to \mathbb{A}_{\mathbb{Q}}^{+}/\mathbb{Q}^{+}$ induced by the quotient map $\mathbb{A}_{\mathbb{Q}}^{+} \to \mathbb{A}_{\mathbb{Q}}^{+}/\mathbb{Q}^{+}$ is surjective. But W is compact (being the topological product of the compact spaces $|x_\infty|_\infty \leq 1/2$ and the $\mathbb{Z}_p$ for all p), hence $\mathbb{A}_{\mathbb{Q}}^{+}/\mathbb{Q}^{+}$ is also compact. $\square$

Corollary 18.3.6. *There is a subset W of $\mathbb{A}_K$ defined by inequalities of the type $|x_v|_v \leq \delta_v$, where $\delta_v = 1$ for almost all v, such that every $\mathbf{y} \in \mathbb{A}_K$ can be put in the form*

$$\mathbf{y} = a + \mathbf{x}, \qquad a \in K, \quad \mathbf{x} \in W,$$

i.e., $\mathbb{A}_K = K + W$.

Proof. We constructed such a set for $K = \mathbb{Q}$ when proving Theorem 18.3.5. For general K the W coming from the proof determines compenent-wise a subset of $\mathbb{A}_K^{+} \cong \mathbb{A}_{\mathbb{Q}}^{+} \oplus \cdots \oplus \mathbb{A}_{\mathbb{Q}}^{+}$ that is a subset of a set with the properties claimed by the corollary. $\square$

As already remarked, $\mathbb{A}_K^+$ is a locally compact group, so it has an invariant Haar measure. In fact one choice of this Haar measure is the product of the Haar measures on the K_v, in the sense of Definition 18.2.5.

Corollary 18.3.7. *The quotient $\mathbb{A}_K^+/K^+$ has finite measure in the quotient measure induced by the Haar measure on $\mathbb{A}_K^+$.*

Remark 18.3.8. This statement is independent of the particular choice of the multiplicative constant in the Haar measure on $\mathbb{A}_K^+$. We do not here go into the question of finding the measure $\mathbb{A}_K^+/K^+$ in terms of our explicitly given Haar measure. (See Tate's thesis, [Cp86, Chapter XV].)

Proof. This can be reduced similarly to the case of $\mathbb{Q}$ or $\mathbb{F}(t)$ which is immediate, e.g., the W defined above has measure 1 for our Haar measure.

Alternatively, finite measure follows from compactness. To see this, cover $\mathbb{A}_K/K^+$ with the translates of U, where U is a nonempty open set with finite measure. The existence of a finite subcover implies finite measure. □

Remark 18.3.9. We give an alternative proof of the product formula $\prod |a|_v = 1$ for nonzero $a \in K$. We have seen that if $x_v \in K_v$, then multiplication by x_v magnifies the Haar measure in K_v^+ by a factor of $|x_v|_v$. Hence if $\mathbf{x} = \{x_v\} \in \mathbb{A}_K$, then multiplication by $\mathbf{x}$ magnifies the Haar measure in $\mathbb{A}_K^+$ by $\prod |x_v|_v$. But now multiplication by $a \in K$ takes $K^+ \subset \mathbb{A}_K^+$ into K^+, so gives a well-defined bijection of $\mathbb{A}_K^+/K^+$ onto $\mathbb{A}_K^+/K^+$ which magnifies the measure by the factor $\prod |a|_v$. Hence $\prod |a|_v = 1$ Corollary 18.3.7. (The point is that if μ is the measure of $\mathbb{A}_K^+/K^+$, then $\mu = \prod |a|_v \cdot \mu$, so because μ is finite we must have $\prod |a|_v = 1$.)

18.4 Strong Approximation

We first prove a technical lemma and corollary, then use them to deduce the strong approximation theorem, which is an extreme generalization of the Chinese Remainder Theorem; it asserts that K^+ is dense in the analogue of the adeles with one valuation removed.

The proof of Lemma 18.4.1 below will use in a crucial way the normalized Haar measure on $\mathbb{A}_K$ and the induced measure on the compact quotient $\mathbb{A}_K^+/K^+$. Since I am not formally developing Haar measure on locally compact groups, and since I didn't explain induced measures on quotients well in the last chapter, hopefully the following discussion will help clarify what is going on.

The real numbers $\mathbb{R}^+$ under addition is a locally compact topological group. Normalized Haar measure μ has the property that $\mu([a,b]) = b - a$, where $a \leq b$ are real numbers and $[a,b]$ is the closed interval from a to b. The subset $\mathbb{Z}^+$ of $\mathbb{R}^+$ is discrete, and the quotient $S^1 = \mathbb{R}^+/\mathbb{Z}^+$ is a compact topological group, which thus has a Haar measure. Let $\overline{\mu}$ be the Haar measure on S^1 normalized so that the natural quotient $\pi : \mathbb{R}^+ \to S^1$ preserves the measure, in the sense that if $X \subset \mathbb{R}^+$ is a measurable set that maps injectively into S^1, then $\mu(X) = \overline{\mu}(\pi(X))$. This determine $\overline{\mu}$ and we have $\overline{\mu}(S^1) = 1$ since $X = [0, 1)$ is a measurable set that maps bijectively onto S^1 and has measure 1. The situation for the map $\mathbb{A}_K \to \mathbb{A}_K/K^+$ is pretty much the same.

Lemma 18.4.1. *There is a constant $C > 0$ that depends only on the global field K with the following property:*

Whenever $\mathbf{x} = \{x_v\}_v \in \mathbb{A}_K$ is such that

$$\prod_v |x_v|_v > C, \tag{18.4.1}$$

then there is a nonzero principal adele $a \in K \subset \mathbb{A}_K$ such that

$$|a|_v \leq |x_v|_v \qquad \text{for all } v.$$

Proof. This proof is modelled on Blichfeldt's proof of Minkowski's Theorem in the Geometry of Numbers, and works in quite general circumstances.

First we show that (18.4.1) implies that $|x_v|_v = 1$ for almost all v. Because $\mathbf{x}$ is an adele, we have $|x_v|_v \leq 1$ for almost all v. If $|x_v|_v < 1$ for infinitely many v, then the product in (18.4.1) would have to be 0. (We prove this only when K is a finite extension of $\mathbb{Q}$.) Excluding archimedean valuations, this is because the normalized valuation $|x_v|_v = |\mathrm{Norm}(x_v)|_p$, which if less than 1 is necessarily $\leq 1/p$. Any infinite product of numbers $1/p_i$ must be 0, whenever p_i is a sequence of primes.

Let c_0 be the Haar measure of $\mathbb{A}_K^+/K^+$ induced from normalized Haar measure on $\mathbb{A}_K^+$, and let c_1 be the Haar measure of the set of $\mathbf{y} = \{y_v\}_v \in \mathbb{A}_K^+$ that satisfy

$$\begin{aligned} |y_v|_v &\leq \frac{1}{2} && \text{if } v \text{ is real archimedean,} \\ |y_v|_v &\leq \frac{1}{2} && \text{if } v \text{ is complex archimedean,} \\ |y_v|_v &\leq 1 && \text{if } v \text{ is non-archimedean.} \end{aligned}$$

(As we will see, any positive real number $\leq 1/2$ would suffice in the definition of c_1 above. For example, in Cassels's article he uses the mysterious $1/10$. He also doesn't discuss the subtleties of the complex archimedean case separately.)

Then $0 < c_0 < \infty$ since $\mathbb{A}_K/K^+$ is compact, and $0 < c_1 < \infty$ because the number of archimedean valuations v is finite. We show that

$$C = \frac{c_0}{c_1}$$

will do. Thus suppose $\mathbf{x}$ is as in (18.4.1).

The set T of $\mathbf{t} = \{t_v\}_v \in \mathbb{A}_K^+$ such that

$$\begin{aligned} |t_v|_v &\leq \frac{1}{2}\,|x_v|_v && \text{if } v \text{ is real archimedean,} \\ |t_v|_v &\leq \frac{1}{2}\sqrt{|x_v|_v} && \text{if } v \text{ is complex archimedean,} \\ |t_v|_v &\leq |x_v|_v && \text{if } v \text{ is non-archimedean} \end{aligned}$$

has measure

$$c_1 \cdot \prod_v |x_v|_v > c_1 \cdot C = c_0. \tag{18.4.2}$$

(Note: If there are complex valuations, then the some of the $|x_v|_v$'s in the product must be squared.)

Because of (18.4.2), in the quotient map $\mathbb{A}_K^+ \to \mathbb{A}_K^+/K^+$ there must be a pair of distinct points of T that have the same image in $\mathbb{A}_K^+/K^+$, say

$$\mathbf{t}' = \{t'_v\}_v \in T \quad \text{and} \quad \mathbf{t}'' = \{t''_v\}_v \in T$$

and

$$a = \mathbf{t}' - \mathbf{t}'' \in K^+$$

is nonzero. Then

$$|a|_v = \left|t'_v - t''_v\right|_v \leq \begin{cases} |t'_v| + |t''_v| \leq 2 \cdot \frac{1}{2}\,|x_v|_v \leq |x_v|_v & \text{if } v \text{ is real archimedean, or} \\ \max(|t'_v|\,, |t''_v|) \leq |x_v|_v & \text{if } v \text{ is non-archimedean,} \end{cases}$$

for all v. In the case of complex archimedean v, we must be careful because the normalized valuation $|\cdot|_v$ is the *square* of the usual archimedean complex valuation $|\cdot|_\infty$ on $\mathbb{C}$, so e.g., it does not satisfy the triangle inequality. In particular, the quantity $|t'_v - t''_v|_v$ is at most the square of the maximum distance between two points in the disc in $\mathbb{C}$ of radius $\frac{1}{2}\sqrt{|x_v|_v}$, where by distance we mean the usual distance. This maximum distance in such a disc is at most $\sqrt{|x_v|_v}$, so $|t'_v - t''_v|_v$ is at most $|x_v|_v$, as required. Thus a satisfies the requirements of the lemma. □

Corollary 18.4.2. *Let v_0 be a normalized valuation and let $\delta_v > 0$ be given for all $v \neq v_0$ with $\delta_v = 1$ for almost all v. Then there is a nonzero $a \in K$ with*

$$|a|_v \leq \delta_v \qquad (\textit{all } v \neq v_0).$$

Proof. This is just a degenerate case of Lemma 18.4.1. Choose $x_v \in K_v$ with $0 < |x_v|_v \leq \delta_v$ and $|x_v|_v = 1$ if $\delta_v = 1$. We can then choose $x_{v_0} \in K_{v_0}$ so that

$$\prod_{\text{all } v \text{ including } v_0} |x_v|_v > C.$$

Then Lemma 18.4.1 does what is required. □

Remark 18.4.3. The character group of the locally compact group $\mathbb{A}_K^+$ is isomorphic to $\mathbb{A}_K^+$ and K^+ plays a special role. See Chapter XV of [Cp86], Lang's [Lan64], Weil's [Wei82], and Godement's Bourbaki seminars 171 and 176. This duality lies behind the functional equation of ζ and L-functions. Iwasawa has shown [Iwa53] that the rings of adeles are characterized by certain general topologico-algebraic properties.

We proved before that K is discrete in $\mathbb{A}_K$. If one valuation is removed, the situation is much different.

Theorem 18.4.4 (Strong Approximation). *Let v_0 be any normalized nontrivial valuation of the global field K. Let $\mathbb{A}_{K,v_0}$ be the restricted topological product of the K_v with respect to the $\mathcal{O}_v$, where v runs through all normalized valuations $v \neq v_0$. Then K is dense in $\mathbb{A}_{K,v_0}$.*

Proof. This proof was suggested by Prof. Kneser at the Cassels-Frohlich conference.

Recall that if $\mathbf{x} = \{x_v\}_v \in \mathbb{A}_{K,v_0}$ then a basis of open sets about $\mathbf{x}$ is the collection of products

$$\prod_{v \in S} B(x_v, \varepsilon_v) \times \prod_{v \notin S,\ v \neq v_0} \mathcal{O}_v,$$

where $B(x_v, \varepsilon_v)$ is an open ball in K_v about x_v, and S runs through finite sets of normalized valuations (not including v_0). Thus denseness of K in $\mathbb{A}_{K,v_0}$ is equivalent to the following statement about elements. Suppose we are given (i) a finite set S of valuations $v \neq v_0$, (ii) elements $x_v \in K_v$ for all $v \in S$, and (iii) an $\varepsilon > 0$. Then there is an element $b \in K$ such that $|b - x_v|_v < \varepsilon$ for all $v \in S$ and $|b|_v \leq 1$ for all $v \notin S$ with $v \neq v_0$.

By the corollary to our proof that $\mathbb{A}_K^+/K^+$ is compact (Corollary 18.3.6), there is a $W \subset \mathbb{A}_K$ that is defined by inequalities of the form $|y_v|_v \leq \delta_v$ (where $\delta_v = 1$ for almost all v) such that ever $\mathbf{z} \in \mathbb{A}_K$ is of the form

$$\mathbf{z} = \mathbf{y} + c, \qquad \mathbf{y} \in W, \quad c \in K. \tag{18.4.3}$$

By Corollary 18.4.2, there is a nonzero $a \in K$ such that

$$\begin{aligned} |a|_v &< \frac{1}{\delta_v} \cdot \varepsilon && \text{for } v \in S, \\ |a|_v &\leq \frac{1}{\delta_v} && \text{for } v \notin S,\ v \neq v_0. \end{aligned}$$

Hence on putting $\mathbf{z} = \frac{1}{a} \cdot \mathbf{x}$ in (18.4.3) and multiplying by a, we see that every $\mathbf{x} \in \mathbb{A}_K$ is of the shape

$$\mathbf{x} = \mathbf{w} + b, \qquad \mathbf{w} \in a \cdot W, \quad b \in K,$$

where $a \cdot W$ is the set of $a\mathbf{y}$ for $\mathbf{y} \in W$. If now we let $\mathbf{x}$ have components the given x_v at $v \in S$, and (say) 0 elsewhere, then $b = \mathbf{x} - \mathbf{w}$ has the properties required. $\square$

Remark 18.4.5. The proof gives a quantitative form of the theorem (i.e., with a bound for $|b|_{v_0}$). For an alternative approach, see [Mah64].

In the next chapter we'll introduce the ideles $\mathbb{A}_K^*$. Finally, we'll relate ideles to ideals, and use everything so far to give a new interpretation of class groups and their finiteness.

Chapter 19

Ideles and Ideals

In this chapter, we introduce the ideles $\mathbb{I}_K$, and relate ideles to ideals, and use what we've done so far to give an alternative interpretation of class groups and their finiteness, thus linking the adelic point of view with the classical point of view of the first part of this course.

19.1 The Idele Group

The invertible elements of any commutative topological ring R are a group R^* under multiplication. In general R^* is not a topological group if it is endowed with the subset topology because inversion need not be continuous (only multiplication and addition on R are required to be continuous). It is usual therefore to give R^* the following topology. There is an injection

$$x \mapsto \left(x, \ \frac{1}{x}\right) \tag{19.1.1}$$

of R^* into the topological product $R \times R$. We give R^* the corresponding subset topology. Then R^* with this topology is a topological group and the inclusion map $R^* \hookrightarrow R$ is continous. To see continuity of inclusion, note that this topology is finer (has at least as many open sets) than the subset topology induced by $R^* \subset R$, since the projection maps $R \times R \to R$ are continuous.

Example 19.1.1. This is a "non-example". The inverse map on $\mathbb{Z}_p^*$ is continuous with respect to the p-adic topology. If $a, b \in \mathbb{Z}_p^*$, then $|a| = |b| = 1$, so if $|a - b| < \varepsilon$, then

$$\left|\frac{1}{a} - \frac{1}{b}\right| = \left|\frac{b-a}{ab}\right| = \frac{|b-a|}{|ab|} < \frac{\varepsilon}{1} = \varepsilon.$$

Definition 19.1.2 (Idele Group)**.** The *idele group* $\mathbb{I}_K$ of K is the group $\mathbb{A}_K^*$ of invertible elements of the adele ring $\mathbb{A}_K$.

We shall usually speak of $\mathbb{I}_K$ as a subset of $\mathbb{A}_K$, and will have to distinguish between the $\mathbb{I}_K$ and $\mathbb{A}_K$-topologies.

Example 19.1.3. For a rational prime p, let $\mathbf{x}_p \in \mathbb{A}_\mathbb{Q}$ be the adele whose pth component is p and whose vth component, for $v \neq p$, is 1. Then $\mathbf{x}_p \to 1$ as $p \to \infty$ in $\mathbb{A}_\mathbb{Q}$, for the following reason. We must show that if U is a basic open set that contains the adele $1 = \{1\}_v$, the $\mathbf{x}_p$ for all sufficiently large p are contained in U. Since U contains 1 and is a basic open set, it is of the form

$$\prod_{v \in S} U_v \times \prod_{v \notin S} \mathbb{Z}_v,$$

where S if a finite set, and the U_v, for $v \in S$, are arbitrary open subsets of $\mathbb{Q}_v$ that contain 1. If q is a prime larger than any prime in S, then $\mathbf{x}_p$ for $p \geq q$, is in U. This proves convergence. If the inverse map were continuous on $\mathbb{I}_K$, then the sequence of $\mathbf{x}_p^{-1}$ would converge to $1^{-1} = 1$. However, if U is an open set as above about 1, then for sufficiently large p, *none* of the adeles $\mathbf{x}_p$ are contained in U.

Lemma 19.1.4. *The group of ideles $\mathbb{I}_K$ is the restricted topological project of the K_v^* with respect to the units $U_v = \mathcal{O}_v^* \subset K_v$, with the restricted product topology.*

We omit the proof of Lemma 19.1.4, which is a matter of thinking carefully about the definitions. The main point is that inversion is continuous on $\mathcal{O}_v^*$ for each v. (See Example 19.1.1.)

We have seen that K is naturally embedded in $\mathbb{A}_K$, so K^* is naturally embedded in $\mathbb{I}_K$.

Definition 19.1.5 (Principal Ideles)**.** We call K^*, considered as a subgroup of $\mathbb{I}_K$, the *principal ideles.*

Lemma 19.1.6. *The principal ideles K^* are discrete as a subgroup of $\mathbb{I}_K$.*

Proof. For K is discrete in $\mathbb{A}_K$, so K^* is embedded in $\mathbb{A}_K \times \mathbb{A}_K$ by (19.1.1) as a discrete subset. (Alternatively, the subgroup topology on $\mathbb{I}_K$ is finer than the topology coming from $\mathbb{I}_K$ being a subset of $\mathbb{A}_K$, and K is already discrete in $\mathbb{A}_K$.) $\square$

Definition 19.1.7 (Content of an Idele)**.** The *content* of $\mathbf{x} = \{x_v\}_v \in \mathbb{I}_K$ is

$$c(\mathbf{x}) = \prod_{\text{all } v} |x_v|_v \in \mathbb{R}_{>0}.$$

Lemma 19.1.8. *The map $\mathbf{x} \to c(\mathbf{x})$ is a continuous homomorphism of the topological group $\mathbb{I}_K$ into $\mathbb{R}_{>0}$, where we view $\mathbb{R}_{>0}$ as a topological group under multiplication. If K is a number field, then c is surjective.*

Proof. That the content map c satisfies the axioms of a homomorphisms follows from the multiplicative nature of the defining formula for c. For continuity, suppose (a, b) is an open interval in $\mathbb{R}_{>0}$. Suppose $\mathbf{x} \in \mathbb{I}_K$ is such that $c(\mathbf{x}) \in (a, b)$. By considering small intervals about each non-unit component of $\mathbf{x}$, we find an open neighborhood $U \subset \mathbb{I}_K$ of $\mathbf{x}$ such that $c(U) \subset (a, b)$. It follows the $c^{-1}((a, b))$ is open.

For surjectivity, use that each archimedean valuation is surjective, and choose an idele that is 1 at all but one archimedean valuation. $\square$

Remark 19.1.9. Note also that the $\mathbb{I}_K$-topology is that appropriate to a group of operators on $\mathbb{A}_K^+$: a basis of open sets is the $S(C,U)$, where $C, U \subset \mathbb{A}_K^+$ are, respectively, $\mathbb{A}_K$-compact and $\mathbb{A}_K$-open, and S consists of the $\mathbf{x} \in \mathbb{I}_J$ such that $(1-\mathbf{x})C \subset U$ and $(1-\mathbf{x}^{-1})C \subset U$.

Definition 19.1.10 (1-Ideles)**.** The subgroup $\mathbb{I}_K^1$ of 1-*ideles* is the subgroup of ideles $\mathbf{x} = \{x_v\}$ such that $c(\mathbf{x}) = 1$. Thus $\mathbb{I}_K^1$ is the kernel of c, so we have an exact sequence

$$1 \to \mathbb{I}_K^1 \to \mathbb{I}_K \xrightarrow{c} \mathbb{R}_{>0} \to 1,$$

where the surjectivity on the right is only if K is a number field.

Lemma 19.1.11. *The subset $\mathbb{I}_K^1$ of $\mathbb{A}_K$ is closed as a subset, and the $\mathbb{A}_K$-subset topology on $\mathbb{I}_K^1$ coincides with the $\mathbb{I}_K$-subset topology on $\mathbb{I}_K^1$.*

Proof. Let $\mathbf{x} \in \mathbb{A}_K$ with $\mathbf{x} \notin \mathbb{I}_K^1$. To prove that $\mathbb{I}_K^1$ is closed in $\mathbb{A}_K$, we find an $\mathbb{A}_K$-neighborhood W of $\mathbf{x}$ that does not meet $\mathbb{I}_K^1$.

1st Case. Suppose that $\prod_v |x_v|_v < 1$ (possibly $= 0$). Then there is a finite set S of v such that

1. S contains all the v with $|x_v|_v > 1$, and
2. $\prod_{v\in S} |x_v|_v < 1$.

Then the set W can be defined by

$$\begin{aligned} |w_v - x_v|_v < \varepsilon \qquad & v \in S \\ |w_v|_v \leq 1 \qquad & v \notin S \end{aligned}$$

for sufficiently small ε.

2nd Case. Suppose that $C := \prod_v |x_v|_v > 1$. Then there is a finite set S of v such that

1. S contains all the v with $|x_v|_v > 1$, and
2. if $v \notin S$ an inequality $|w_v|_v < 1$ implies $|w_v|_v < \frac{1}{2C}$. (This is because for a non-archimedean valuation, the largest absolute value less than 1 is $1/p$, where p is the residue characteristic. Also, the upper bound in Cassels's article is $\frac{1}{2}C$ instead of $\frac{1}{2C}$, but I think he got it wrong.)

We can choose ε so small that $|w_v - x_v|_v < \varepsilon$ (for $v \in S$) implies $1 < \prod_{v\in S} |w_v|_v < 2C$. Then W may be defined by

$$\begin{aligned} |w_v - x_v|_v < \varepsilon \qquad & v \in S \\ |w_v|_v \leq 1 \qquad & v \notin S. \end{aligned}$$

This works because if $\mathbf{w} \in W$, then either $|w_v|_v = 1$ for all $v \notin S$, in which case $1 < c(\mathbf{w}) < 2c$, so $\mathbf{w} \notin \mathbb{I}_K^1$, or $|w_{v_0}|_{v_0} < 1$ for some $v_0 \notin S$, in which case

$$c(\mathbf{w}) = \left(\prod_{v \in S} |w_v|_v\right) \cdot |w_{v_0}| \cdots < 2C \cdot \frac{1}{2C} \cdots < 1,$$

so again $\mathbf{w} \notin \mathbb{I}_K^1$.

We next show that the $\mathbb{I}_K$- and $\mathbb{A}_K$-topologies on $\mathbb{I}_K^1$ are the same. If $\mathbf{x} \in \mathbb{I}_K^1$, we must show that every $\mathbb{A}_K$-neighborhood of $\mathbf{x}$ contains an $\mathbb{A}_K$-neighborhood and vice-versa.

Let $W \subset \mathbb{I}_K^1$ be an $\mathbb{A}_K$-neighborhood of $\mathbf{x}$. Then it contains an $\mathbb{A}_K$-neighborhood of the type

$$|w_v - x_v|_v < \varepsilon \qquad v \in S \tag{19.1.2}$$

$$|w_v|_v \leq 1 \qquad v \notin S \tag{19.1.3}$$

where S is a finite set of valuations v. This contains the $\mathbb{I}_K$-neighborhood in which $\leq$ in (19.1.2) is replaced by $=$.

Next let $H \subset \mathbb{I}_K^1$ be an $\mathbb{I}_K$-neighborhood. Then it contains an $\mathbb{I}_K$-neighborhood of the form

$$|w_v - x_v|_v < \varepsilon \qquad v \in S \tag{19.1.4}$$

$$|w_v|_v = 1 \qquad v \notin S, \tag{19.1.5}$$

where the finite set S contains at least all archimedean valuations v and all valuations v with $|x_v|_v \neq 1$. Since $\prod |x_v|_v = 1$, we may also suppose that ε is so small that (19.1.4) implies

$$\prod_v |w_v|_v < 2.$$

Then the intersection of (19.1.4) with $\mathbb{I}_K^1$ is the same as that of (19.1.2) with $\mathbb{I}_K^1$, i.e., (19.1.4) defines an $\mathbb{A}_K$-neighborhood. $\square$

By the product formula we have that $K^* \subset \mathbb{I}_K^1$. The following result is of vital importance in class field theory.

Theorem 19.1.12. *The quotient $\mathbb{I}_K^1/K^*$ with the quotient topology is compact.*

Proof. After the preceeding lemma, it is enough to find an $\mathbb{A}_K$-compact set $W \subset \mathbb{A}_K$ such that the map

$$W \cap \mathbb{I}_K^1 \to \mathbb{I}_K^1/K^*$$

is surjective. We take for W the set of $\mathbf{w} = \{w_v\}_v$ with

$$|w_v|_v \leq |x_v|_v\,,$$

where $\mathbf{x} = \{x_v\}_v$ is any idele of content greater than the C of Lemma 18.4.1.

Let $\mathbf{y} = \{y_v\}_v \in \mathbb{I}_K^1$. Then the content of $\mathbf{x}/\mathbf{y}$ equals the content of $\mathbf{x}$, so by Lemma 18.4.1 there is an $a \in K^*$ such that

$$|a|_v \leq \left|\frac{x_v}{y_v}\right|_v \qquad \text{all } v.$$

Then $a\mathbf{y} \in W$, as required. $\square$

Remark 19.1.13. The quotient $\mathbb{I}_K^1/K^*$ is totally disconnected in the function field case. For the structure of its connected component in the number field case, see papers of Artin and Weil in the "Proceedings of the Tokyo Symposium on Algebraic Number Theory, 1955" (Science Council of Japan) or [AT90]. The determination of the character group of $\mathbb{I}_K/K^*$ is global class field theory.

19.2 Ideals and Divisors

Suppose that K is a finite extension of $\mathbb{Q}$. Let F_K be the the free abelian group on a set of symbols in bijection with the non-archimedean valuation v of K. Thus an element of F_K is a formal linear combination

$$\sum_{v \text{ non arch.}} n_v \cdot v$$

where $n_v \in \mathbb{Z}$ and all but finitely many n_v are 0.

Lemma 19.2.1. *There is a natural bijection between F_K and the group of nonzero fractional ideals of $\mathcal{O}_K$. The correspondence is induced by*

$$v \mapsto \wp_v = \{x \in \mathcal{O}_K : v(x) < 1\},$$

where v is a non-archimedean valuation.

Endow F_K with the discrete topology. Then there is a natural continuous map $\pi : \mathbb{I}_K \to F_K$ given by

$$\mathbf{x} = \{x_v\}_v \mapsto \sum_v \mathrm{ord}_v(x_v) \cdot v.$$

This map is continuous since the inverse image of a valuation v (a point) is the product

$$\pi^{-1}(v) = \pi\mathcal{O}_v^* \quad \times \prod_{w \text{ archimedean}} K_w^* \quad \times \prod_{w \neq v \text{ non-arch.}} \mathcal{O}_w^*,$$

which is an open set in the restricted product topology on $\mathbb{I}_K$. Moreover, the image of K^* in F_K is the group of nonzero principal fractional ideals.

Recall that the *class group* C_K of the number field K is by definition the quotient of F_K by the image of K^*.

Theorem 19.2.2. *The class group C_K of a number field K is finite.*

Proof. We first prove that the map $\mathbb{I}_K^1 \to F_K$ is surjective. Let ∞ be an archimedean valuation on K. If v is a non-archimedean valuation, let $\mathbf{x} \in \mathbb{I}_K^1$ be a 1-idele such that $x_w = 1$ at ever valuation w except v and ∞. At v, choose $x_v = \pi$ to be a generator for the maximal ideal of $\mathcal{O}_v$, and choose x_∞ to be such that $|x_\infty|_\infty = 1/|x_v|_v$. Then $\mathbf{x} \in \mathbb{I}_K$ and $\prod_w |x_w|_w = 1$, so $\mathbf{x} \in \mathbb{I}_K^1$. Also $\mathbf{x}$ maps to $v \in F_K$.

Thus the group of ideal classes is the continuous image of the compact group $\mathbb{I}_K^1/K^*$ (see Theorem 19.1.12), hence compact. But a compact discrete group is finite. □

19.2.1 The Function Field Case

When K is a finite separable extension of $\mathbb{F}(t)$, we define the divisor group D_K of K to be the free abelian group on all the valuations v. For each v the number of elements of the residue class field $\mathbb{F}_v = \mathcal{O}_v/\wp_v$ of v is a power, say q^{n_v}, of the number q of elements in $\mathbb{F}_v$. We call n_v the degree of v, and similarly define $\sum n_v d_v$ to be the degree of the divisor $\sum n_v \cdot v$. The divisors of degree 0 form a group D_K^0. As before, the principal divisor attached to $a \in K^*$ is $\sum \mathrm{ord}_v(a) \cdot v \in D_K$. The following theorem is proved in the same way as Theorem 19.2.2.

Theorem 19.2.3. *The quotient of D_K^0 modulo the principal divisors is a finite group.*

19.2.2 Jacobians of Curves

For those familiar with algebraic geometry and algebraic curves, one can prove Theorem 19.2.3 from an alternative point of view. There is a bijection between nonsingular geometrically irreducible projective curves over $\mathbb{F}$ and function fields K over $\mathbb{F}$ (which we assume are finite separable extensions of $\mathbb{F}(t)$ such that $\overline{\mathbb{F}} \cap K = \mathbb{F}$). Let X be the curve corresponding to K. The group D_K^0 is in bijection with the divisors of degree 0 on X, a group typically denoted $\mathrm{Div}^0(X)$. The quotient of $\mathrm{Div}^0(X)$ by principal divisors is denoted $\mathrm{Pic}^0(X)$. The *Jacobian* of X is an abelian variety $J = \mathrm{Jac}(X)$ over the finite field $\mathbb{F}$ whose dimension is equal to the genus of X. Moreover, assuming X has an $\mathbb{F}$-rational point, the elements of $\mathrm{Pic}^0(X)$ are in natural bijection with the $\mathbb{F}$-rational points on J. In particular, with these hypothesis, the class group of K, which is isomorphic to $\mathrm{Pic}^0(X)$, is in bijection with the group of $\mathbb{F}$-rational points on an algebraic variety over a finite field. This gives an alternative more complicated proof of finiteness of the degree 0 class group of a function field.

Without the degree 0 condition, the divisor class group won't be finite. It is an extension of $\mathbb{Z}$ by a finite group.

$$0 \to \mathrm{Pic}^0(X) \to \mathrm{Pic}(X) \xrightarrow{\deg} n\mathbb{Z} \to 0,$$

where n is the greatest common divisor of the degrees of elements of $\mathrm{Pic}(X)$, which is 1 when X has a rational point.

Chapter 20

Exercises

1. Which of the following rings have infinitely many prime ideals?

   (a) The integers $\mathbb{Z}$.

   (b) The ring $\mathbb{Z}[x]$ of polynomials over $\mathbb{Z}$.

   (c) The quotient ring $\mathbb{C}[x]/(x^{2005} - 1)$.

   (d) The ring $(\mathbb{Z}/6\mathbb{Z})[x]$ of polynomials over the ring $\mathbb{Z}/6\mathbb{Z}$.

   (e) The quotient ring $\mathbb{Z}/n\mathbb{Z}$, for a fixed positive integer n.

   (f) The rational numbers $\mathbb{Q}$.

   (g) The polynomial ring $\mathbb{Q}[x, y, z]$ in three variables.

2. Prove that every finite integral domain is a field.

3. (a) Give an example of two ideals I, J in a commutative ring R whose product is *not* equal to the set $\{ab : a \in I, b \in J\}$.

   (b) Suppose R is a principal ideal domain. Is it always the case that

   $$IJ = \{ab : a \in I, b \in J\}$$

   for all ideals I, J in R?

4. Is the set $\mathbb{Z}[\frac{1}{2}]$ of rational numbers with denominator a power of 2 a fractional ideal?

5. Suppose you had the choice of the following two jobs[1]:

   Job 1 Starting with an annual salary of \$1000, and a \$200 increase every year.

   Job 2 Starting with a semiannual salary of \$500, and an increase of \$50 every 6 months.

[1]From *The Education of T.C. MITS* (1942).

In all other respects, the two jobs are exactly alike. Which is the better offer (after the first year)? Write a Sage program that creates a table showing how much money you will receive at the end of each year for each job. (Of course you could easily do this by hand – the point is to get familiar with Sage.)

6. Let $\mathcal{O}_K$ be the ring of integers of a number field. Let F_K denote the abelian group of fractional ideals of $\mathcal{O}_K$.
   (a) Prove that F_K is torsion free.
   (b) Prove that F_K is not finitely generated.
   (c) Prove that F_K is countable.
   (d) Conclude that if K and L are number fields, then there exists some (non-canonical) isomorphism of groups $F_K \approx F_L$.

7. In this problem, you will give an example to illustrate the failure of unique factorization in the ring $\mathcal{O}_K$ of integers of $\mathbb{Q}(\sqrt{-6})$.
   (a) Give an element $\alpha \in \mathcal{O}_K$ that factors in two distinct ways into irreducible elements.
   (b) Observe explicitly that the (α) factors uniquely, i.e., the two distinct factorization in the previous part of this problem do not lead to two distinct factorization of the ideal (α) into prime ideals.

8. Factor the ideal (10) as a product of primes in the ring of integers of $\mathbb{Q}(\sqrt{11})$. You're allowed to use a computer, as long as you show the commands you use.

9. Let $\mathcal{O}_K$ be the ring of integers of a number field K, and let $p \in \mathbb{Z}$ be a prime number. What is the cardinality of $\mathcal{O}_K/(p)$ in terms of p and $[K : \mathbb{Q}]$, where (p) is the ideal of $\mathcal{O}_K$ generated by p?

10. Give an example of each of the following, with proof:
    (a) A non-principal ideal in a ring.
    (b) A module that is not finitely generated.
    (c) The ring of integers of a number field of degree 3.
    (d) An order in the ring of integers of a number field of degree 5.
    (e) The matrix on K of left multiplication by an element of K, where K is a degree 3 number field.
    (f) An integral domain that is not integrally closed in its field of fractions.
    (g) A Dedekind domain with finite cardinality.
    (h) A fractional ideal of the ring of integers of a number field that is not an integral ideal.

11. Let $\varphi : R \to S$ be a homomorphism of (commutative) rings.

(a) Prove that if $I \subset S$ is an ideal, then $\varphi^{-1}(I)$ is an ideal of R.

(b) Prove moreover that if I is prime, then $\varphi^{-1}(I)$ is also prime.

12. Let $\mathcal{O}_K$ be the ring of integers of a number field. The Zariski topology on the set $X = \mathrm{Spec}(\mathcal{O}_K)$ of all prime ideals of $\mathcal{O}_K$ has closed sets the sets of the form
$$V(I) = \{\mathfrak{p} \in X : \mathfrak{p} \mid I\},$$
where I varies through *all* ideals of $\mathcal{O}_K$, and $\mathfrak{p} \mid I$ means that $I \subset \mathfrak{p}$.

    (a) Prove that the collection of closed sets of the form $V(I)$ is a topology on X.

    (b) Let Y be the subset of nonzero prime ideals of $\mathcal{O}_K$, with the induced topology. Use unique factorization of ideals to prove that the closed subsets of Y are exactly the finite subsets of Y along with the set Y.

    (c) Prove that the conclusion of (a) is still true if $\mathcal{O}_K$ is replaced by an order in $\mathcal{O}_K$, i.e., a subring that has finite index in $\mathcal{O}_K$ as a $\mathbb{Z}$-module.

13. Explicitly factor the ideals generated by each of 2, 3, and 5 in the ring of integers of $\mathbb{Q}(\sqrt[3]{2})$. (Thus you'll factor 3 separate ideals as products of prime ideals.) You may assume that the ring of integers of $\mathbb{Q}(\sqrt[3]{2})$ is $\mathbb{Z}[\sqrt[3]{2}]$, but do *not* simply use a computer command to do the factorizations.

14. Let $K = \mathbb{Q}(\zeta_{13})$,where ζ_{13} is a primitive 13th root of unity. Note that K has ring of integers $\mathcal{O}_K = \mathbb{Z}[\zeta_{13}]$.

    (a) Factor 2, 3, 5, 7, 11, and 13 in the ring of integers $\mathcal{O}_K$. You may use a computer.

    (b) For $p \neq 13$, find a conjectural relationship between the number of prime ideal factors of $p\mathcal{O}_K$ and the order of the reduction of p in $(\mathbb{Z}/13\mathbb{Z})^*$.

    (c) Compute the minimal polynomial $f(x) \in \mathbb{Z}[x]$ of ζ_{13}. Reinterpret your conjecture as a conjecture that relates the degrees of the irreducible factors of $f(x) \pmod{p}$ to the order of p modulo 13. Does your conjecture remind you of quadratic reciprocity?

15. (a) Find by hand and with proof the ring of integers of each of the following two fields: $\mathbb{Q}(\sqrt{5})$, $\mathbb{Q}(i)$.

    (b) Find the ring of integers of $\mathbb{Q}(a)$, where $a^5+7a+1 = 0$ using a computer.

16. Let p be a prime. Let $\mathcal{O}_K$ be the ring of integers of a number field K, and suppose $a \in \mathcal{O}_K$ is such that $[\mathcal{O}_K : \mathbb{Z}[a]]$ is finite and coprime to p. Let $f(x)$ be the minimal polynomial of a. We proved in class that if the reduction $\overline{f} \in \mathbb{F}_p[x]$ of f factors as
$$\overline{f} = \prod g_i^{e_i},$$

where the g_i are distinct irreducible polynomials in $\mathbb{F}_p[x]$, then the primes appearing in the factorization of $p\mathcal{O}_K$ are the ideals $(p, g_i(a))$. In class, we did not prove that the exponents of these primes in the factorization of $p\mathcal{O}_K$ are the e_i. Prove this.

17. Let $a_1 = 1+i$, $a_2 = 3+2i$, and $a_3 = 3+4i$ as elements of $\mathbb{Z}[i]$.

    (a) Prove that the ideals $I_1 = (a_1)$, $I_2 = (a_2)$, and $I_3 = (a_3)$ are coprime in pairs.

    (b) Compute $\#\mathbb{Z}[i]/(I_1 I_2 I_3)$.

    (c) Find a single element in $\mathbb{Z}[i]$ that is congruent to n modulo I_n, for each $n \leq 3$.

18. Find an example of a field K of degree at least 4 such that the ring $\mathcal{O}_K$ of integers of K is not of the form $\mathbb{Z}[a]$ for any $a \in \mathcal{O}_K$.

19. Let $\mathfrak{p}$ be a prime ideal of $\mathcal{O}_K$, and suppose that $\mathcal{O}_K/\mathfrak{p}$ is a finite field of characteristic $p \in \mathbb{Z}$. Prove that there is an element $\alpha \in \mathcal{O}_K$ such that $\mathfrak{p} = (p, \alpha)$. This justifies why we can represent prime ideals of $\mathcal{O}_K$ as pairs (p, α), as is done in Sage. (More generally, if I is an ideal of $\mathcal{O}_K$, we can choose one of the elements of I to be *any* nonzero element of I.)

20. (*) Give an example of an order $\mathcal{O}$ in the ring of integers of a number field and an ideal I such that I cannot be generated by 2 elements as an ideal. Does the Chinese Remainder Theorem hold in $\mathcal{O}$? [The (*) means that this problem is more difficult than usual.]

21. For each of the following three fields, determining if there is an order of discriminant 20 contained in its ring of integers:

$$K = \mathbb{Q}(\sqrt{5}), \quad K = \mathbb{Q}(\sqrt[3]{2}), \quad \text{and} \ldots$$

    K any extension of $\mathbb{Q}$ of degree 2005. [Hint: for the last one, apply the exact form of our theorem about finiteness of class groups to the unit ideal to show that the discriminant of a degree 2005 field must be large.]

22. Prove that the quantity $C_{r,s}$ in our theorem about finiteness of the class group can be taken to be $\left(\frac{4}{\pi}\right)^s \frac{n!}{n^n}$, as follows (adapted from [SD01, pg. 19]): Let S be the set of elements $(x_1, \ldots, x_n) \in \mathbb{R}^n$ such that

$$|x_1| + \cdots |x_r| + 2 \sum_{v=r+1}^{r+s} \sqrt{x_v^2 + x_{v+s}^2} \leq 1.$$

    (a) Prove that S is convex and that $M = n^{-n}$, where

$$M = \max\{|x_1 \cdots x_r \cdot (x_{r+1}^2 + x_{(r+1)+s}^2) \cdots (x_{r+s}^2 + x_n^2)| : (x_1, \ldots, x_n) \in S\}.$$

[Hint: For convexity, use the triangle inequality and that for $0 \leq \lambda \leq 1$, we have

$$\lambda\sqrt{x_1^2+y_1^2}+(1-\lambda)\sqrt{x_2^2+y_2^2} \geq \sqrt{(\lambda x_1+(1-\lambda)x_2)^2+(\lambda y_1+(1-\lambda)y_2)^2}$$

for $0 \leq \lambda \leq 1$. In polar coordinates this last inequality is

$$\lambda r_1+(1-\lambda)r_2 \geq \sqrt{\lambda^2 r_1^2+2\lambda(1-\lambda)r_1 r_2\cos(\theta_1-\theta_2)+(1-\lambda)^2 r_2^2},$$

which is trivial. That $M \leq n^{-n}$ follows from the inequality between the arithmetic and geometric means.

(b) Transforming pairs x_v, x_{v+s} from Cartesian to polar coordinates, show also that $v = 2^r(2\pi)^s D_{r,s}(1)$, where

$$D_{\ell,m}(t) = \int\cdots\int_{\mathcal{R}_{\ell,m}(t)} y_1\cdots y_m dx_1\cdots dx_\ell dy_1\cdots dy_m$$

and $\mathcal{R}_{\ell,\Updownarrow}(t)$ is given by $x_\rho \geq 0$ $(1 \leq \rho \leq \ell)$, $y_\rho \geq 0$ $(1 \leq \rho \leq m)$ and

$$x_1+\cdots+x_\ell+2(y_1+\cdots+y_m) \leq t.$$

(c) Prove that

$$D_{\ell,m}(t) = \int_0^t D_{\ell-1,m}(t-x)dx = \int_0^{t/2} D_{\ell,m-1}(t-2y)ydy$$

and deduce by induction that

$$D_{\ell,m}(t) = \frac{4^{-m}t^{\ell+2m}}{(\ell+2m)!}$$

23. Let K vary through all number fields. What torsion subgroups $(U_K)_{\mathrm{tor}}$ actually occur?

24. If $U_K \approx \mathbb{Z}^n \times (U_K)_{\mathrm{tor}}$, we say that U_K has rank n. Let K vary through all number fields. What ranks actually occur?

25. Let K vary through all number fields such that the group U_K of units of K is a finite group. What finite groups U_K actually occur?

26. Let $K = \mathbb{Q}(\zeta_5)$.

(a) Show that $r = 0$ and $s = 2$.

(b) Find explicit generators for the group of units U_K.

(c) Draw an illustration of the log map $\varphi : U_K \to \mathbb{R}^2$, including the hyperplane $x_1+x_2 = 0$ and the lattice in the hyperplane spanned by the image of U_K.

27. Let K be a number field. Prove that $p \mid d_K$ if and only if p ramifies in K. (Note: This fact is proved in many books.)

28. (a) Give an example of a finite nontrivial Galois extension K of $\mathbb{Q}$ and a prime ideal $\mathfrak{p}$ such that $D_{\mathfrak{p}} = \mathrm{Gal}(K/\mathbb{Q})$.

    (b) Give an example of a finite nontrivial Galois extension K of $\mathbb{Q}$ and a prime ideal $\mathfrak{p}$ such that $D_{\mathfrak{p}}$ has order 1.

    (c) Give an example of a finite Galois extension K of $\mathbb{Q}$ and a prime ideal $\mathfrak{p}$ such that $D_{\mathfrak{p}}$ is not a normal subgroup of $\mathrm{Gal}(K/\mathbb{Q})$.

    (d) Give an example of a finite Galois extension K of $\mathbb{Q}$ and a prime ideal $\mathfrak{p}$ such that $I_{\mathfrak{p}}$ is not a normal subgroup of $\mathrm{Gal}(K/\mathbb{Q})$.

29. Let S_3 by the symmetric group on three symbols, which has order 6.

    (a) Observe that $S_3 \cong D_3$, where D_3 is the dihedral group of order 6, which is the group of symmetries of an equilateral triangle.

    (b) Use (29a) to write down an explicit embedding $S_3 \hookrightarrow \mathrm{GL}_2(\mathbb{C})$.

    (c) Let K be the number field $\mathbb{Q}(\sqrt[3]{2}, \omega)$, where $\omega^3 = 1$ is a nontrivial cube root of unity. Show that K is a Galois extension with Galois group isomorphic to S_3.

    (d) We thus obtain a 2-dimensional irreducible complex Galois representation

    $$\rho : \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}) \to \mathrm{Gal}(K/\mathbb{Q}) \cong S_3 \subset \mathrm{GL}_2(\mathbb{C}).$$

    Compute a representative matrix of Frob_p and the characteristic polynomial of Frob_p for $p = 5, 7, 11, 13$.

30. Look up the Riemann-Roch theorem in a book on algebraic curves.

    (a) Write it down in your own words.

    (b) Let E be an elliptic curve over a field K. Use the Riemann-Roch theorem to deduce that the natural map

    $$E(K) \to \mathrm{Pic}^0(E/K)$$

    is an isomorphism.

31. Suppose G is a finite group and A is a finite G-module. Prove that for any q, the group $\mathrm{H}^q(G, A)$ is a torsion abelian group of exponent dividing the order $\#A$ of A.

32. Let $K = \mathbb{Q}(\sqrt{5})$ and let $A = U_K$ be the group of units of K, which is a module over the group $G = \mathrm{Gal}(K/\mathbb{Q})$. Compute the cohomology groups $\mathrm{H}^0(G, A)$ and $\mathrm{H}^1(G, A)$. (You shouldn't use a computer, except maybe to determine U_K.)

33. Let $K = \mathbb{Q}(\sqrt{-23})$ and let C be the class group of $\mathbb{Q}(\sqrt{-23})$, which is a module over the Galois group $G = \mathrm{Gal}(K/\mathbb{Q})$. Determine $\mathrm{H}^0(G, C)$ and $\mathrm{H}^1(G, C)$.

34. Let E be the elliptic curve $y^2 = x^3 + x + 1$. Let $E[2]$ be the group of points of order dividing 2 on E. Let

$$\overline{\rho}_{E,2} : \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}) \to \mathrm{Aut}(E[2])$$

be the mod 2 Galois representation associated to E.

(a) Find the fixed field K of $\ker(\overline{\rho}_{E,2})$.

(b) Is $\overline{\rho}_{E,2}$ surjective?

(c) Find the group $\mathrm{Gal}(K/\mathbb{Q})$.

(d) Which primes are ramified in K?

(e) Let I be an inertia group above 2, which is one of the ramified primes. Determine $E[2]^I$ explicitly for your choice of I. What is the characteristic polynomial of Frob_2 acting on $E[2]^I$.

(f) What is the characteristic polynomial of Frob_3 acting on $E[2]$?

(g) Let K be a number field. Prove that there is a finite set S of primes of K such that

$$\mathcal{O}_{K,S} = \{a \in K^* : \mathrm{ord}_{\mathfrak{p}}(a\mathcal{O}_K) \geq 0 \text{ all } \mathfrak{p} \notin S\} \cup \{0\}$$

is a prinicipal ideal domain. The condition $\mathrm{ord}_{\mathfrak{p}}(a\mathcal{O}_K) \geq 0$ means that in the prime ideal factorization of the fractional ideal $a\mathcal{O}_K$, we have that $\mathfrak{p}$ occurs to a nonnegative power.

(h) Let $a \in K$ and n a positive integer. Prove that $L = K(a^{1/n})$ is unramified outside the primes that divide n and the norm of a. This means that if $\mathfrak{p}$ is a prime of $\mathcal{O}_K$, and $\mathfrak{p}$ is coprime to $n\,\mathrm{Norm}_{L/K}(a)\mathcal{O}_K$, then the prime factorization of $\mathfrak{p}\mathcal{O}_L$ involves no primes with exponent bigger than 1.

(i) Write down a proof of Hilbert's Theorem 90, formulated as the statement that for any number field K, we have

$$\mathrm{H}^1(K, \overline{K}^*) = 0.$$

1. Let k be any field. Prove that the only nontrivial valuations on $k(t)$ which are trivial on k are equivalent to the valuation (13.3.3) or (13.3.4) of page 157.

2. A field with the topology induced by a valuation is a topological field, i.e., the operations sum, product, and reciprocal are continuous.

3. Give an example of a non-archimedean valuation on a field that is not discrete.

4. Prove that the field $\mathbb{Q}_p$ of p-adic numbers is uncountable.

5. Prove that the polynomial $f(x) = x^3 - 3x^2 + 2x + 5$ has all its roots in $\mathbb{Q}_5$, and find the 5-adic valuations of each of these roots. (You might need to use Hensel's lemma, which we don't discuss in detail in this book. See [Cas67, App. C].)

6. In this problem you will compute an example of weak approximation, like I did in the Example 14.3.3. Let $K = \mathbb{Q}$, let $|\cdot|_7$ be the 7-adic absolute value, let $|\cdot|_{11}$ be the 11-adic absolute value, and let $|\cdot|_\infty$ be the usual archimedean absolute value. Find an element $b \in \mathbb{Q}$ such that $|b - a_i|_i < \frac{1}{10}$, where $a_7 = 1$, $a_{11} = 2$, and $a_\infty = -2004$.

7. Prove that -9 has a cube root in $\mathbb{Q}_{10}$ using the following strategy (this is a special case of Hensel's Lemma, which you can read about in an appendix to Cassel's article).

   (a) Show that there is an element $\alpha \in \mathbb{Z}$ such that $\alpha^3 \equiv 9 \pmod{10^3}$.

   (b) Suppose $n \geq 3$. Use induction to show that if $\alpha_1 \in \mathbb{Z}$ and $\alpha^3 \equiv 9 \pmod{10^n}$, then there exists $\alpha_2 \in \mathbb{Z}$ such that $\alpha_2^3 \equiv 9 \pmod{10^{n+1}}$. (Hint: Show that there is an integer b such that $(\alpha_1 + b \cdot 10^n)^3 \equiv 9 \pmod{10^{n+1}}$.)

   (c) Conclude that 9 has a cube root in $\mathbb{Q}_{10}$.

8. Compute the first 5 digits of the 10-adic expansions of the following rational numbers:
$$\frac{13}{2}, \quad \frac{1}{389}, \quad \frac{17}{19}, \quad \text{the 4 square roots of 41.}$$

9. Let $N > 1$ be an integer. Prove that the series
$$\sum_{n=1}^{\infty} (-1)^{n+1} n! = 1! - 2! + 3! - 4! + 5! - 6! + \cdots .$$
converges in $\mathbb{Q}_N$.

10. Prove that -9 has a cube root in $\mathbb{Q}_{10}$ using the following strategy (this is a special case of "Hensel's Lemma").

(a) Show that there is $\alpha \in \mathbb{Z}$ such that $\alpha^3 \equiv 9 \pmod{10^3}$.

(b) Suppose $n \geq 3$. Use induction to show that if $\alpha_1 \in \mathbb{Z}$ and $\alpha^3 \equiv 9 \pmod{10^n}$, then there exists $\alpha_2 \in \mathbb{Z}$ such that $\alpha_2^3 \equiv 9 \pmod{10^{n+1}}$. (Hint: Show that there is an integer b such that $(\alpha_1 + b10^n)^3 \equiv 9 \pmod{10^{n+1}}$.)

(c) Conclude that 9 has a cube root in $\mathbb{Q}_{10}$.

11. Let $N > 1$ be an integer.

(a) Prove that $\mathbb{Q}_N$ is equipped with a natural ring structure.

(b) If N is prime, prove that $\mathbb{Q}_N$ is a field.

12. (a) Let p and q be distinct primes. Prove that $\mathbb{Q}_{pq} \cong \mathbb{Q}_p \times \mathbb{Q}_q$.

(b) Is $\mathbb{Q}_{p^2}$ isomorphic to either of $\mathbb{Q}_p \times \mathbb{Q}_p$ or $\mathbb{Q}_p$?

13. Prove that every finite extension of $\mathbb{Q}_p$ "comes from" an extension of $\mathbb{Q}$, in the following sense. Given an irreducible polynomial $f \in \mathbb{Q}_p[x]$ there exists an irreducible polynomial $g \in \mathbb{Q}[x]$ such that the fields $\mathbb{Q}_p[x]/(f)$ and $\mathbb{Q}_p[x]/(g)$ are isomorphic. [Hint: Choose each coefficient of g to be sufficiently close to the corresponding coefficient of f, then use Hensel's lemma to show that g has a root in $\mathbb{Q}_p[x]/(f)$.]

14. Find the 3-adic expansion to precision 4 of each root of the following polynomial over $\mathbb{Q}_3$:

$$f = x^3 - 3x^2 + 2x + 3 \in \mathbb{Q}_3[x].$$

Your solution should conclude with three expressions of the form

$$a_0 + a_1 \cdot 3 + a_2 \cdot 3^2 + a_3 \cdot 3^3 + O(3^4).$$

15. (a) Find the normalized Haar measure of the following subset of $\mathbb{Q}_7^+$:

$$U = B\left(28, \frac{1}{50}\right) = \left\{x \in \mathbb{Q}_7 : |x - 28| < \frac{1}{50}\right\}.$$

(b) Find the normalized Haar measure of the subset $\mathbb{Z}_7^*$ of $\mathbb{Q}_7^*$.

16. Suppose that K is a finite extension of $\mathbb{Q}_p$ and L is a finite extension of $\mathbb{Q}_q$, with $p \neq q$ and assume that K and L have the same degree. Prove that there is a polynomial $g \in \mathbb{Q}[x]$ such that $\mathbb{Q}_p[x]/(g) \cong K$ and $\mathbb{Q}_q[x]/(g) \cong L$. [Hint: Combine your solution to 13 with the weak approximation theorem.]

17. Prove that the ring C defined in Section 9 really is the tensor product of A and B, i.e., that it satisfies the defining universal mapping property for tensor products. Part of this problem is for you to look up a functorial definition of tensor product.

18. Find a zero divisor pair in $\mathbb{Q}(\sqrt{5}) \otimes_{\mathbb{Q}} \mathbb{Q}(\sqrt{5})$.

19. (a) Is $\mathbb{Q}(\sqrt{5}) \otimes_{\mathbb{Q}} \mathbb{Q}(\sqrt{-5})$ a field?
    (b) Is $\mathbb{Q}(\sqrt[4]{5}) \otimes_{\mathbb{Q}} \mathbb{Q}(\sqrt[4]{-5}) \otimes_{\mathbb{Q}} \mathbb{Q}(\sqrt{-1})$ a field?

20. Suppose ζ_5 denotes a primitive 5th root of unity. For any prime p, consider the tensor product $\mathbb{Q}_p \otimes_{\mathbb{Q}} \mathbb{Q}(\zeta_5) = K_1 \oplus \cdots \oplus K_{n(p)}$. Find a simple formula for the number $n(p)$ of fields appearing in the decomposition of the tensor product $\mathbb{Q}_p \otimes_{\mathbb{Q}} \mathbb{Q}(\zeta_5)$. To get full credit on this problem your formula must be correct, but you do *not* have to prove that it is correct.

21. Suppose $\|\cdot\|_1$ and $\|\cdot\|_2$ are equivalent norms on a finite-dimensional vector space V over a field K (with valuation $|\cdot|$). Carefully prove that the topology induced by $\|\cdot\|_1$ is the same as that induced by $\|\cdot\|_2$.

22. Suppose K and L are number fields (i.e., finite extensions of $\mathbb{Q}$). Is it possible for the tensor product $K \otimes_{\mathbb{Q}} L$ to contain a nilpotent element? (A nonzero element a in a ring R is *nilpotent* if there exists $n > 1$ such that $a^n = 0$.)

23. Let K be the number field $\mathbb{Q}(\sqrt[5]{2})$.

    (a) In how many ways does the 2-adic valuation $|\cdot|_2$ on $\mathbb{Q}$ extend to a valuation on K?
    (b) Let $v = |\cdot|$ be a valuation on K that extends $|\cdot|_2$. Let K_v be the completion of K with respect to v. What is the residue class field $\mathbb{F}$ of K_v?

24. Prove that the product formula holds for $\mathbb{F}(t)$ similar to the proof we gave in class using Ostrowski's theorem for $\mathbb{Q}$. You may use the analogue of Ostrowski's theorem for $\mathbb{F}(t)$, which you had on a previous homework assignment. (Don't give a measure-theoretic proof.)

25. Prove Theorem 18.3.5, that "The global field K is discrete in $\mathbb{A}_K$ and the quotient $\mathbb{A}_K^+/K^+$ of additive groups is compact in the quotient topology." in the case when K is a finite extension of $\mathbb{F}(t)$, where $\mathbb{F}$ is a finite field.

Bibliography

[Art23] E. Artin, *Über eine neue Art von L-reihen*, Abh. Math. Sem. Univ. Hamburg **3** (1923), 89–108.

[Art30] E Artin, *Zur Theorie der L-Reihen mit allgemeinen Gruppencharakteren*, Abh. math. Semin. Univ. Hamburg **8** (1930), 292–306.

[Art59] E. Artin, *Theory of algebraic numbers*, Notes by Gerhard Würges from lectures held at the Mathematisches Institut, Göttingen, Germany, in the Winter Semester, vol. 1956/7, George Striker, Schildweg 12, Göttingen, 1959. MR 24 #A1884

[Art91] M. Artin, *Algebra*, Prentice Hall Inc., Englewood Cliffs, NJ, 1991. MR 92g:00001

[AT90] E. Artin and J. Tate, *Class field theory*, second ed., Advanced Book Classics, Addison-Wesley Publishing Company Advanced Book Program, Redwood City, CA, 1990. MR 91b:11129

[BCP97] W. Bosma, J. Cannon, and C. Playoust, *The Magma algebra system. I. The user language*, J. Symbolic Comput. **24** (1997), no. 3–4, 235–265, Computational algebra and number theory (London, 1993). MR 1 484 478

[BDSBT01] Kevin Buzzard, Mark Dickinson, Nick Shepherd-Barron, and Richard Taylor, *On icosahedral Artin representations*, Duke Math. J. **109** (2001), no. 2, 283–318. MR 1845181 (2002k:11078)

[BL94] J. A. Buchmann and H. W. Lenstra, Jr., *Approximating rings of integers in number fields*, J. Théor. Nombres Bordeaux **6** (1994), no. 2, 221–260. MR 1360644 (96m:11092)

[BS02] K. Buzzard and W. A. Stein, *A mod five approach to modularity of icosahedral Galois representations*, Pacific J. Math. **203** (2002), no. 2, 265–282. MR 2003c:11052

[Buh78] J. P. Buhler, *Icosahedral Galois representations*, Springer-Verlag, Berlin, 1978, Lecture Notes in Mathematics, Vol. 654.

[Cas67] J. W. S. Cassels, *Global fields*, Algebraic Number Theory (Proc. Instructional Conf., Brighton, 1965), Thompson, Washington, D.C., 1967, pp. 42–84.

[Cas91] ______, *Lectures on elliptic curves*, London Mathematical Society Student Texts, vol. 24, Cambridge University Press, Cambridge, 1991. MR 92k:11058

[CL84] H. Cohen and H. W. Lenstra, Jr., *Heuristics on class groups of number fields*, Number theory, Noordwijkerhout 1983 (Noordwijkerhout, 1983), Lecture Notes in Math., vol. 1068, Springer, Berlin, 1984, pp. 33–62. MR 756082 (85j:11144)

[Coh93] H. Cohen, *A course in computational algebraic number theory*, Springer-Verlag, Berlin, 1993. MR 94i:11105

[Cp86] J. W. S. Cassels and A. Fröhlich (eds.), *Algebraic number theory*, London, Academic Press Inc. [Harcourt Brace Jovanovich Publishers], 1986, Reprint of the 1967 original.

[EH00] D. Eisenbud and J. Harris, *The geometry of schemes*, Springer-Verlag, New York, 2000. MR 2001d:14002

[Fre94] G. Frey (ed.), *On Artin's conjecture for odd 2-dimensional representations*, Springer-Verlag, Berlin, 1994, 1585. MR 95i:11001

[Har77] R. Hartshorne, *Algebraic Geometry*, Springer-Verlag, New York, 1977, Graduate Texts in Mathematics, No. 52.

[Iwa53] K. Iwasawa, *On the rings of valuation vectors*, Ann. of Math. (2) **57** (1953), 331–356. MR 14,849a

[KW08] C. Khare and J.-P. Wintenberger, *Serre's modularity conjecture (i)*, Preprint (2008).

[Lan64] S. Lang, *Algebraic numbers*, Addison-Wesley Publishing Co., Inc., Reading, Mass.-Palo Alto-London, 1964. MR 28 #3974

[Lan80] R. P. Langlands, *Base change for* GL(2), Princeton University Press, Princeton, N.J., 1980.

[Len02] H. W. Lenstra, Jr., *Solving the Pell equation*, Notices Amer. Math. Soc. **49** (2002), no. 2, 182–192. MR 2002i:11028

[LL93] A. K. Lenstra and H. W. Lenstra, Jr. (eds.), *The development of the number field sieve*, Springer-Verlag, Berlin, 1993. MR 96m:11116

[Mah64] K. Mahler, *Inequalities for ideal bases in algebraic number fields*, J. Austral. Math. Soc. **4** (1964), 425–448. MR 31 #1243

[Mar77] Daniel A. Marcus, *Number fields*, Universitext (1979), Springer, 1977.

[PAR] PARI, *A computer algebra system designed for fast computations in number theory*, `http://pari.math.u-bordeaux.fr/`.

[S+11] W. A. Stein et al., *Sage Mathematics Software (Version 4.6.2)*, The Sage Development Team, 2011, `http://www.sagemath.org`.

[SD01] H. P. F. Swinnerton-Dyer, *A brief guide to algebraic number theory*, London Mathematical Society Student Texts, vol. 50, Cambridge University Press, Cambridge, 2001. MR 2002a:11117

[Ser73] J-P. Serre, *A Course in Arithmetic*, Springer-Verlag, New York, 1973, Translated from the French, Graduate Texts in Mathematics, No. 7.

[Ser79] ———, *Local fields*, Springer-Verlag, New York, 1979, Translated from the French by Marvin Jay Greenberg.

[Sil92] J. H. Silverman, *The arithmetic of elliptic curves*, Springer-Verlag, New York, 1992, Corrected reprint of the 1986 original.

[ST68] J-P. Serre and J. T. Tate, *Good reduction of abelian varieties*, Ann. of Math. (2) **88** (1968), 492–517, `http://wstein.org/papers/bib/Serre-Tate-Good_Reduction_of_Abelian_Varieties.pdf`.

[Ste09] William Stein, *Elementary number theory: primes, congruences, and secrets*, Undergraduate Texts in Mathematics, Springer, New York, 2009, A computational approach. MR 2464052 (2009i:11002)

[Was97] Lawrence C. Washington, *Introduction to cyclotomic fields*, second ed., Graduate Texts in Mathematics, vol. 83, Springer-Verlag, New York, 1997. MR 1421575 (97h:11130)

[Wei82] A. Weil, *Adeles and algebraic groups*, Progress in Mathematics, vol. 23, Birkhäuser Boston, Mass., 1982, With appendices by M. Demazure and Takashi Ono. MR 83m:10032