ALGEBRAIC NUMBER THEORY

J.S. MILNE

ABSTRACT. These are the notes for a course taught at the University of Michigan in F92 as Math 676. They are available at `www.math.lsa.umich.edu/~jmilne/`.

Please send comments and corrections to me at jmilne@umich.edu.

v2.01 (August 14, 1996.) First version on the web.

v2.10 (August 31, 1998.) Fixed many minor errors; added exercises and index.

CONTENTS

It is standard to use Gothic (fraktur) letters for ideals:

$\mathfrak{a}$	$\mathfrak{b}$	$\mathfrak{c}$	$\mathfrak{m}$	$\mathfrak{n}$	$\mathfrak{p}$	$\mathfrak{q}$	$\mathfrak{A}$	$\mathfrak{B}$	$\mathfrak{C}$	$\mathfrak{M}$	$\mathfrak{N}$	$\mathfrak{P}$	$\mathfrak{Q}$
a	b	c	m	n	p	q	A	B	C	M	N	P	Q

I use the following notations:

$X \approx Y$	X and Y are isomorphic;
$X \cong Y$	X and Y are canonically isomorphic or there is a given or unique isomorphism;
$X \stackrel{\text{df}}{=} Y$	X is defined to be Y, or equals Y by definition;
$X \subset Y$	X is a subset of Y (not necessarily proper).

Introduction

An *algebraic number field* is a finite extension of $\mathbb{Q}$; an *algebraic number* is an element of an algebraic number field. Algebraic number theory studies the arithmetic of algebraic number fields — the ring of integers in the number field, the ideals in the ring of integers, the units, the extent to which the ring of integers fails to be have unique factorization, and so on. One important tool for this is "localization", in which we complete the number field relative to a metric attached to a prime ideal of the number field. The completed field is called a *local field* — its arithmetic is much simpler than that of the number field, and sometimes we can answer questions by first solving them locally, that is, in the local fields.

An *abelian extension* of a field is a Galois extension of the field with abelian Galois group. Global class field theory classifies the abelian extensions of a number field K in terms of the arithmetic of K; local class field theory does the same for local fields.

This course is concerned with algebraic number theory. Its sequel is on class field theory (see my notes CFT).

I now give a quick sketch of what the course will cover. The *fundamental theorem of arithmetic* says that integers can be uniquely factored into products of prime powers: an $m \neq 0$ in $\mathbb{Z}$ can be written in the form,

$$m = up_1^{r_1} \cdots p_n^{r_n}, \quad u = \pm 1, \quad p_i \text{ prime number}, r_i > 0,$$

and this factorization is essentially unique.

Consider more generally an integral domain A. An element $a \in A$ is said to be a *unit* if it has an inverse in A; I write $A^\times$ for the multiplicative group of units in A. An element p of A is said to *prime* if it is neither zero nor a unit, and if

$$p|ab \Rightarrow p|a \text{ or } p|b.$$

If A is a principal ideal domain, then every nonzero nonunit element a of A can be written in the form,

$$a = p_1^{r_1} \cdots p_n^{r_n}, \quad p_i \text{ prime element}, r_i > 0,$$

and the factorization is unique up to order and replacing each p_i with an associate, i.e., with its product with a unit.

Our first task will be to discover to what extent unique factorization holds, or fails to hold, in number fields. Three problems present themselves. First, factorization in a field only makes sense with respect to a subring, and so we must define the "ring of integers" $\mathcal{O}_K$ in our number field K. Secondly, since unique factorization will in general fail, we shall need to find a way of measuring by how much it fails. Finally, since factorization is only considered up to units, in order to fully understand the arithmetic of K, we need to understand the structure of the group of units U_K in $\mathcal{O}_K$. Resolving these three problems will occupy the first five sections of the course.

The ring of integers. Let K be an algebraic number field. Because K is of finite degree over $\mathbb{Q}$, every element α of K is a root of a monic polynomial

$$f(X) = X^n + a_1X^{n-1} + \cdots + a_0, \quad a_i \in \mathbb{Q}.$$

If α is a root of a monic polynomial with *integer* coefficients, then α is called an *algebraic integer* of K. We shall see that the algebraic integers form a subring $\mathcal{O}_K$ of K.

The criterion as stated is difficult to apply. We shall see that to prove that α is an algebraic integer, it suffices to check that its minimum polynomial (relative to $\mathbb{Q}$) has integer coefficients.

Consider for example the field $K = \mathbb{Q}[\sqrt{d}]$, where d is a square-free integer. The minimum polynomial of $\alpha = a + b\sqrt{d}$, $b \neq 0$, $a, b \in \mathbb{Q}$, is

$$(X - (a + b\sqrt{d}))(X - (a - b\sqrt{d})) = X^2 - 2aX + (a^2 - b^2 d).$$

Thus α is an algebraic integer if and only if

$$2a \in \mathbb{Z}, \quad a^2 - b^2 d \in \mathbb{Z}.$$

From this it follows easily that

$$\mathcal{O}_K = \mathbb{Z}[\sqrt{d}] = \{m + n\sqrt{d} \mid m, n \in \mathbb{Z}\} \text{ if } d \equiv 2, 3 \mod 4,$$

and

$$\mathcal{O}_K = \{m + n\frac{1 + \sqrt{d}}{2} \mid m, n \in \mathbb{Z}\} \text{ if } d \equiv 1 \mod 4,$$

i.e., $\mathcal{O}_K$ is the set of sums $m' + n'\sqrt{d}$ with m' and n' either both integers or both half-integers.

Let ζ_d be a primitive d^{th} root of 1, for example, $\zeta_d = \exp(2\pi i/d)$, and let $K = \mathbb{Q}[\zeta_d]$. Then we shall see that

$$\mathcal{O}_K = \mathbb{Z}[\zeta_d] = \{\textstyle\sum m_i \zeta_d^i \mid m_i \in \mathbb{Z}\}.$$

as one would hope.

Factorization. An element p of an integral domain A is said to be *irreducible* if it is neither zero nor a unit, and can't be written as a product of two nonunits. For example, a prime element is (obviously) irreducible. A ring A is a *unique factorization domain* if every nonzero nonunit element of A can be expressed as a product of irreducible elements in essentially one way. Is $\mathcal{O}_K$ a unique factorization domain? No, not in general!

In fact, we shall see that each element of $\mathcal{O}_K$ can be written as a product of irreducible elements (this is true for all Noetherian rings) — it is the uniqueness that fails.

For example, in $\mathbb{Z}[\sqrt{-5}]$ we have

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

To see that 2, 3, $1 + \sqrt{-5}$, $1 - \sqrt{-5}$ are irreducible, and no two are associates, we use the norm map

$$\mathrm{Nm} : \mathbb{Q}[\sqrt{-5}] \to \mathbb{Q}, \quad a + b\sqrt{-5} \mapsto a^2 + 5b^2.$$

For $\alpha \in \mathcal{O}_K$, we have

$$\mathrm{Nm}(\alpha) = 1 \iff \alpha\bar{\alpha} = 1 \iff \alpha \text{ is a unit.} \qquad (*)$$

If $1+\sqrt{-5}=\alpha\beta$, then $\mathrm{Nm}(\alpha\beta)=\mathrm{Nm}(1+\sqrt{-5})=6$. Thus $\mathrm{Nm}(\alpha)=1,2,3$, or 6. In the first case, α is a unit, the second and third cases don't occur, and in the fourth case β is a unit. A similar argument shows that $2,3$, and $1-\sqrt{-5}$ are irreducible. Next note that (*) implies that associates have the same norm, and so it remains to show that $1+\sqrt{-5}$ and $1-\sqrt{-5}$ are not associates, but

$$1+\sqrt{-5}=(a+b\sqrt{-5})(1-\sqrt{-5})$$

has no solution with $a,b\in\mathbb{Z}$.

Why does unique factorization fail in $\mathcal{O}_K$? The problem is that irreducible elements in $\mathcal{O}_K$ need not be prime. In the above example, $1+\sqrt{-5}$ divides $2\cdot 3$ but it divides neither 2 nor 3. In fact, in an integral domain in which factorizations exist (e.g. a Noetherian ring), factorization is unique if all irreducible elements are prime.

What can we recover? Consider

$$210=6\cdot 35=10\cdot 21.$$

If we were naive, we might say this shows factorization is not unique in $\mathbb{Z}$; instead, we recognize that there is a unique factorization underlying these two decompositions, namely,

$$210=(2\cdot 3)(5\cdot 7)=(2\cdot 5)(3\cdot 7).$$

The idea of Kummer and Dedekind was to enlarge the set of "prime numbers" so that, for example, in $\mathbb{Z}[\sqrt{-5}]$ there is a unique factorization,

$$6=(\mathfrak{p}_1\cdot\mathfrak{p}_2)(\mathfrak{p}_3\cdot\mathfrak{p}_4)=(\mathfrak{p}_1\cdot\mathfrak{p}_3)(\mathfrak{p}_2\cdot\mathfrak{p}_4),$$

underlying the above factorization; here the $\mathfrak{p}_i$ are "ideal prime factors".

How do we define "ideal factors"? Clearly, an ideal factor should be characterized by the algebraic integers it divides. Moreover divisibility by $\mathfrak{a}$ should have the following properties:

$$\mathfrak{a}|0;\quad \mathfrak{a}|a,\mathfrak{a}|b\Rightarrow\mathfrak{a}|a\pm b;\quad \mathfrak{a}|a\Rightarrow\mathfrak{a}|ab \text{ for all } b\in\mathcal{O}_K.$$

If in addition division by $\mathfrak{a}$ has the property that

$$\mathfrak{a}|ab\Rightarrow\mathfrak{a}|a \text{ or } \mathfrak{a}|b,$$

then we call $\mathfrak{a}$ a "prime ideal factor". Since all we know about an ideal factor is the set of elements it divides, we may as well identify it with this set. Thus an ideal factor is a set of elements $\mathfrak{a}\subset\mathcal{O}_K$ such that

$$0\in\mathfrak{a};\quad a,b\in\mathfrak{a}\Rightarrow a\pm b\in\mathfrak{a};\quad a\in\mathfrak{a}\Rightarrow ab\in\mathfrak{a} \text{ for all } b\in\mathcal{O}_K;$$

it is prime if an addition,

$$ab\in\mathfrak{a}\Rightarrow a\in\mathfrak{a} \text{ or } b\in\mathfrak{a}.$$

Many of you will recognize that an ideal factor is what we now call an *ideal*, and a prime ideal factor is a *prime ideal*.

There is an obvious notion of the product of two ideals:

$$\mathfrak{ab}|c\iff c=\sum a_ib_i,\quad \mathfrak{a}|a_i,\quad \mathfrak{b}|b_i.$$

In other words,

$$\mathfrak{ab}=\{\sum a_ib_i\mid a_i\in\mathfrak{a},\quad b_i\in\mathfrak{b}\}.$$

One see easily that this is again an ideal, and that if

$$\mathfrak{a} = (a_1, ..., a_m) \text{ and } \mathfrak{b} = (b_1, ..., b_n)$$

then

$$\mathfrak{a} \cdot \mathfrak{b} = (a_1b_1, a_1b_2, ..., a_ib_j, ..., a_mb_n).$$

With these definitions, one recovers unique factorization: if $a \neq 0$, then there is an essentially unique factorization:

$$(a) = \mathfrak{p}_1^{r_1} \cdots \mathfrak{p}_n^{r_n} \text{ with each } \mathfrak{p}_i \text{ a prime ideal.}$$

In the above example,

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

In fact, I claim

$$\begin{aligned}
(2, 1 + \sqrt{-5})(2, 1 - \sqrt{-5}) &= (2) \\
(3, 1 + \sqrt{-5})(3, 1 - \sqrt{-5}) &= (3) \\
(2, 1 + \sqrt{-5})(3, 1 + \sqrt{-5}) &= (1 + \sqrt{-5}) \\
(2, 1 - \sqrt{-5})(3, 1 - \sqrt{-5}) &= (1 - \sqrt{-5}).
\end{aligned}$$

For example, $(2, 1 + \sqrt{-5})(2, 1 - \sqrt{-5}) = (4, 2 + 2\sqrt{-5}, 2 - 2\sqrt{-5}, 6)$. Since every generator is divisible by 2, $(2, 1 + \sqrt{-5})(2, 1 - \sqrt{-5}) \subset (2)$. Conversely,

$$2 = 6 - 4 \in (4, 2 + 2\sqrt{-5}, 2 - 2\sqrt{-5}, 6)$$

and so $(2, 1 + \sqrt{-5})(2, 1 - \sqrt{-5}) = (2)$. Moreover, the four ideals $(2, 1 + \sqrt{-5})$, $(2, 1 - \sqrt{-5})$, $(3, 1 + \sqrt{-5})$, and $(3, 1 - \sqrt{-5})$ are all prime. For example

$$\mathbb{Z}[\sqrt{-5}]/(3, 1 - \sqrt{-5}) = \mathbb{Z}/(3),$$

which is an integral domain.

How far is this from what we want, namely, unique factorization of elements? In other words, how many "ideal" elements have we had to add to our "real" elements to get unique factorization. In a certain sense, only a finite number: we shall see that there is a finite set of ideals $\mathfrak{a}_1, ..., \mathfrak{a}_h$ such that every ideal is of the form $\mathfrak{a}_i \cdot (a)$ for some i and some $a \in \mathcal{O}_K$. Better, we shall construct a group I of "fractional" ideals in which the principal fractional ideals (a), $a \in K^\times$, form a subgroup P of finite index. The index is called the *class number* h_K of K. We shall see that

$$h_K = 1 \iff \mathcal{O}_K \text{ is a principal ideal domain} \iff \mathcal{O}_K \text{ is a unique factorization domain.}$$

Units. Unlike $\mathbb{Z}$, $\mathcal{O}_K$ can have an infinite number of units. For example, $(1 + \sqrt{2})$ is a unit of infinite order in $\mathbb{Z}[\sqrt{2}]$:

$$(1 + \sqrt{2})(-1 + \sqrt{2}) = 1; \quad (1 + \sqrt{2})^m \neq 1 \text{ for } m \geq 1.$$

In fact $\mathbb{Z}[\sqrt{2}]^\times = \{\pm(1 + \sqrt{2})^m \mid m \in \mathbb{Z}\}$, and so

$$\mathbb{Z}[\sqrt{2}]^\times \approx \{\pm 1\} \times \{\text{free abelian group of rank 1}\}.$$

In general, we shall show (unit theorem) that the roots of 1 in K form a finite group $\mu(K)$, and that

$$\mathcal{O}_K^\times \approx \mu(K) \times \mathbb{Z}^r \quad \text{(as an abelian group)};$$

moreover, we shall find r.

Applications. I hope to give some applications. One motivation for the development of algebraic number theory was the attempt to prove Fermat's last "theorem", i.e., that there are no integer solutions to the equation

$$X^m + Y^m = Z^m$$

when $m \geq 3$, except for the obvious solutions.

When $m = 3$, this can proved by the method of "infinite descent", i.e., from one solution, you show that you can construct a smaller solution, which leads to a contradiction[1]. The proof makes use of the factorization

$$Y^3 = Z^3 - X^3 = (Z - X)(Z^2 + XZ + X^2),$$

and it was recognized that a stumbling block to proving the theorem for larger m is that no such factorization exists into polynomials with integer coefficients. This led people to look at more general factorizations.

In a very famous incident, the French mathematician Lamé gave a talk at the Paris Academy in 1847 in which he claimed to prove Fermat's last theorem using the following ideas. Let $p > 2$ be a prime, and suppose x, y, z are nonzero integers such that

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

Write

$$x^p = z^p - y^p = \prod(z - \zeta^i y), \quad 0 \leq i \leq p-1, \quad \zeta = e^{2\pi i/p}.$$

He then showed how to obtain a smaller solution to the equation, and hence a contradiction. Liouville immediately questioned a step in Lamé's proof in which he assumed that, in order to show that each factor $(z - \zeta^i y)$ is a p^{th} power, it suffices to show that the factors are relatively prime in pairs and their product is a p^{th} power. In fact, Lamé couldn't justify his step ($\mathbb{Z}[\zeta]$ is not always a principal ideal domain), and Fermat's last theorem remains unproven to the present day[2]. However, shortly after Lamé's embarrassing lecture, Kummer used his results on the arithmetic of the fields $\mathbb{Q}[\zeta]$ to prove Fermat's last theorem for all "regular primes".

Another application is to finding Galois groups. The splitting field of a polynomial $f(X) \in \mathbb{Q}[X]$ is a Galois extension of $\mathbb{Q}$. In the basic graduate algebra course (see FT), we learn how to compute the Galois group only when the degree is very small (e.g., ≤ 3). By using algebraic number theory one can write down an algorithm to do it for any degree.

[1]The simplest proof by infinite descent is that showing that $\sqrt{2}$ is irrational.

[2]Written in 1992.

A brief history of numbers. PREHISTORY (??-1600). Basic arithmetic was developed in many parts of the world thousands of years ago. For example, 3,500 years ago the Babylonians apparently knew how to construct the solutions to

$$X^2 + Y^2 = Z^2.$$

At least they knew that

$$(4961)^2 + (6480)^2 = (8161)^2$$

which could scarcely be found by trial and error. The Chinese remainder theorem was known in China, thousands of years ago. The Greeks knew the fundamental theorem of arithmetic, and, of course, Euclid's algorithm.

FERMAT (1601–1665). Apart from his famous last "theorem", he invented the method of infinite descent. He also posed the problem of finding integer solutions to the equation,

$$X^2 - AY^2 = 1, \quad A \in \mathbb{Z}, \qquad (*)$$

which is essentially the problem[3] of finding the units in $\mathbb{Z}[\sqrt{A}]$. The English mathematicians found an algorithm for solving the problem, but neglected to show that the algorithm always works.

EULER (1707–1783). Among many other works, he discovered the quadratic reciprocity law.

LAGRANGE (1736–1813). He proved that the algorithm for solving (*) always leads to a solution.

LEGENDRE (1752–1833). He proved the "Hasse principle" for quadratic forms in three variables over $\mathbb{Q}$: the quadratic form $Q(X, Y, Z)$ has a nontrivial zero in $\mathbb{Q}$ if and only if it has one in $\mathbb{R}$ and the congruence $Q \equiv 0 \bmod p^n$ has a nontrivial solution for all p and n.

GAUSS (1777–1855). He found many proofs of the quadratic reciprocity law:

$$\left(\frac{p}{q}\right)\left(\frac{q}{p}\right) = (-1)^{(p-1)(q-1)/4}, \quad p, q \text{ odd primes.}$$

He studied the Gaussian integers $\mathbb{Z}[i]$ in order to find a quartic reciprocity law. He studied the classification of binary quadratic forms over $\mathbb{Z}$ which, as we shall see, is closely related to the problem of finding the class numbers of quadratic fields.

DIRICHLET (1805–1859). He proved the following "unit theorem": let α be a root of a monic irreducible polynomial $f(X)$ with integer coefficients; suppose that $f(X)$ has r real roots and $2s$ complex roots; then $\mathbb{Z}[\alpha]^\times$ is a finitely generated group of rank $r + s - 1$. He proved a famous analytic formula for the class number.

KUMMER (1810–1893). He made a deep study of the arithmetic of cyclotomic fields, motivated by a search for higher reciprocity laws. His general result on Fermat's last theorem is the most important to date.

HERMITE (1822–1901).

EISENSTEIN (1823–1852).

[3]The Indian mathematician Bhaskara (12th century) knew general rules for finding solutions to the equation.

Kronecker (1823–1891). He developed an alternative to Dedekind's ideals. He also had one of the most beautiful ideas in mathematics, the *Kronecker liebster Jugendtraum*, for generating abelian extensions of number fields.

Riemann (1826–1866). Made the Riemann hypothesis.

Dedekind (1831–1916). He was the first mathematician to formally define fields — many of the basic theorems on fields in basic graduate algebra courses were proved by him. He also found the correct general definition of the ring of integers in a number field, and he proved that ideals factor uniquely into products of prime ideals. Moreover, he improved the Dirichlet unit theorem.

Weber (1842–1913). Made important progress in class field theory and the Kronecker Jugendtraum.

Hensel (1861–1941). He introduced the notion of the p-adic completion of a field.

Hilbert (1862–1943). He wrote a very influential book on algebraic number theory in 1897, which gave the first systematic account of the theory. Some of his famous problems were on number theory, and have also been influential.

Takagi (1875–1960). He made very important advances in class field theory.

Hecke (1887–1947). Introduced Hecke L-series.

Artin (1898–1962). He found the "Artin reciprocity law", which is the main theorem of class field theory.

Hasse (1898–1979). Proved the Hasse principle for all quadratic forms over number fields.

Weil (1906–1998). Defined the Weil group, which enabled him to give a common generalization of Artin L-series and Hecke L-series.

Chevalley (1909–??). The main statements of class field theory are purely algebraic, but all the earlier proofs used analysis. Chevalley gave a purely algebraic proof.

Iwasawa (1917–). He introduced an important new approach into the study of algebraic number theory which was suggested by the theory of curves over finite fields.

Tate (1925–). With Artin, he gave a complete cohomological treatment of class field theory. With Lubin he introduced a concrete way of generating abelian extensions of local fields.

Langlands (1936–). "Langlands's philosophy" is a vast series of conjectures that, among other things, contains a *nonabelian* class field theory.

References. *Books on algebraic number theory.*

Artin, E., *Theory of Algebraic Numbers*, Göttingen notes, 1959. Elegant; good examples; but he adopts a valuation approach rather than the ideal-theoretic approach we use in this course.

Artin, E., *Algebraic Numbers and Algebraic Functions*, Nelson, 1968. Covers both the number field and function field case.

Borevich, Z. I., and Shafarevich, I. R., *Number Theory*, Academic Press, 1966. In addition to basic algebraic number theory, it contains material on diophantine equations.

Cassels, J.W.S., and Fröhlich, A., Eds., *Algebraic Number Theory*, Academic Press, 1967. The proceedings of an instructional conference. Many of the articles are excellent, for example, those of Serre and Tate on class field theory.

Cassels, J.W.S., *Local fields*, London Math. Soc., 1986. Concentrates on local fields, but does also deal with number fields, and it gives some interesting applications.

Cohn, P.M., *Algebraic Numbers and Algebraic Functions*, Chapman and Hall, 1991. The valuation approach.

Dedekind, R., *Theory of Algebraic Integers*, Cambridge Univ. Press, 1996 (translation of the 1877 French original). Develops the basic theory through the finiteness of the class number in a way that is surprising close to modern approach in, for example, these notes.

Edwards, H., *Fermat's Last Theorem: A Genetic Introduction to Algebraic Number Theory*, Springer, 1977. A history of algebraic number theory, concentrating on the efforts to prove Fermat's last theorem. Edwards is one of the most reliable writers on the history of number theory.

Fröhlich, A., and Taylor, M.J., *Algebraic Number Theory*, Cambridge Univ. Press, 1991. Lots of good problems.

Goldstein, L.J., *Analytic Number Theory*, Prentice-Hall, 1971. Similar approach to Lang 1970, but the writing is a bit careless. Sometimes includes more details than Lang, and so it is probably easier to read.

Janusz, G. *Algebraic Number Fields*, Second Edn, Amer. Math. Soc., 1996. It covers both algebraic number theory and class field theory, which it treats from a lowbrow analytic/algebraic approach. In the past, I sometimes used the first edition as a text for this course and its sequel.

Lang, S. *Algebraic Numbers Theory*, Addison-Wesley, 1970. Difficult to read unless you already know it, but it does contain an enormous amount of material. Covers algebraic number theory, and it does class field theory from a highbrow analytic/algebraic approach.

Marcus, D. *Number Fields*, Springer, 1977. This is a rather pleasant down-to-earth introduction to algebraic number theory.

Narkiewicz, W. *Algebraic Numbers*, Springer, 1990. Encyclopedic coverage of algebraic number theory.

Samuel, P., *Algebraic Theory of Numbers*, Houghton Mifflin, 1970. A very easy treatment, with lots of good examples, but doesn't go very far.

Serre, J.-P. *Corps Locaux*, Hermann, 1962 (Translated as Local Fields). A classic. An excellent account of local fields, cohomology of groups, and local class field theory. The local class field theory is bit dated (Lubin-Tate groups weren't known when the book was written) but this is the best book for the other two topics.

Weil, A., *Basic Number Theory*, Springer, 1967. Very heavy going, but you will learn a lot if you manage to read it (covers algebraic number theory and class field theory).

Weiss, R., *Algebraic Number Theory*, McGraw-Hill, 1963. Very detailed; in fact a bit too fussy and pedantic.

Weyl, H., *Algebraic Theory of Numbers*, Princeton Univ. Press, 1940. One of the first books in English; by one of the great mathematicians of the twentieth century. Idiosyncratic — Weyl prefers Kronecker to Dedekind, e.g., see the section "Our disbelief in ideals".

Computational Number Theory.

Cohen, H., *A Course in Computational Number Theory*, Springer, 1993.

Lenstra, H., *Algorithms in Algebraic Number Theory*, Bull. Amer. Math. Soc., 26, 1992, 211–244.

Pohst and Zassenhaus, *Algorithmic Algebraic Number Theory*, Cambridge Univ. Press, 1989.

The two books provide algorithms for most of the constructions we make in this course. The first assumes the reader knows number theory, whereas the second develops the whole subject algorithmically. Cohen's book is the more useful as a supplement to this course, but wasn't available when these notes were first written, and so the references are to Pohst and Zassenhaus. While the books are concerned with more-or-less practical algorithms for fields of small degree and small discriminant, Lenstra's article concentrates on finding "good" general algorithms.

Additional references

Atiyah, M.F., and MacDonald, I.G., *Introduction to Commutative Algebra*, Addison-Wesley, 1969. I use this as a reference on commutative algebra.

Washington, L., *Introduction to Cyclotomic Fields*, 1982. This is the best book on cyclotomic fields.

I will sometimes refer to my other course notes:

GT: Group Theory (594)
FT: Fields and Galois Theory (594)
EC: Elliptic Curves (679).
CFT: Class Field Theory (776).

1. Preliminaries from Commutative Algebra

Many results that were first proved for rings of integers in number fields are true for more general commutative rings, and it is more natural to prove them in that context.

Basic definitions. All rings will be commutative, and have an identity element (i.e., an element 1 such that $1a = a$ for all $a \in A$), and a homomorphism of rings will map the identity element to the identity element.

A ring B together with a homomorphism of rings $A \to B$ will be referred to as an *A-algebra*. We use this terminology mainly when A is a subring of B. In this case, for elements $\beta_1, ..., \beta_m$ of B, $A[\beta_1, ..., \beta_m]$ denotes the smallest subring of B containing A and the β_i. It consists of all polynomials in the β_i with coefficients in A, i.e., elements of the form

$$\sum a_{i_1 \dots i_m} \beta_1^{i_1} \dots \beta_m^{i_m}, \quad a_{i_1 \dots i_m} \in A.$$

We also refer to $A[\beta_1, ..., \beta_m]$ as the A-subalgebra of B *generated* by the β_i, and when $B = A[\beta_1, ..., \beta_m]$ we say that the β_i *generate* B as an A-algebra.

For elements $a_1, a_2, \dots$ of A, $(a_1, a_2, \dots)$ denotes the smallest ideal containing the a_i. It consists of finite sums $\sum c_i a_i$, $c_i \in A$, and it is called the *ideal generated by* $a_1, a_2, \dots$. When $\mathfrak{a}$ and $\mathfrak{b}$ are ideals in A, we define

$$\mathfrak{a} + \mathfrak{b} = \{a + b \mid a \in \mathfrak{a},\, b \in \mathfrak{b}\}.$$

It is again an ideal in A — in fact, it is the smallest ideal containing both $\mathfrak{a}$ and $\mathfrak{b}$. If $\mathfrak{a} = (a_1, ..., a_m)$ and $\mathfrak{b} = (b_1, ..., b_n)$, then $\mathfrak{a} + \mathfrak{b} = (a_1, ..., a_m, b_1, ..., b_n)$.

Given an ideal $\mathfrak{a}$ in A, we can form the quotient ring $A/\mathfrak{a}$. Let $f\colon A \to A/\mathfrak{a}$ be the homomorphism $a \mapsto a + \mathfrak{a}$; then $\mathfrak{b} \mapsto f^{-1}(\mathfrak{b})$ defines a one-to-one correspondence between the ideals of $A/\mathfrak{a}$ and the ideals of A containing $\mathfrak{a}$, and

$$A/f^{-1}(\mathfrak{b}) \xrightarrow{\approx} (A/\mathfrak{a})/\mathfrak{b}.$$

A proper ideal $\mathfrak{a}$ of A is *prime* if $ab \in \mathfrak{a} \Rightarrow a$ or $b \in \mathfrak{a}$. An ideal $\mathfrak{a}$ is prime if and only if the quotient ring $A/\mathfrak{a}$ is an integral domain. An element p of A is said to be *prime* if (p) is a prime ideal; equivalently, if $p|ab \Rightarrow p|a$ or $p|b$.

A proper ideal $\mathfrak{a}$ in A is *maximal* if there does not exist an ideal $\mathfrak{b}$, $\mathfrak{a} \subsetneqq \mathfrak{b} \subsetneqq A$. An ideal $\mathfrak{a}$ is maximal if and only if $A/\mathfrak{a}$ is a field. Every proper ideal $\mathfrak{a}$ of A is contained in a maximal ideal — if A is Noetherian (see below) this is obvious; otherwise the proof requires Zorn's lemma. In particular, every nonunit in A is contained in a maximal ideal.

There are the implications: A is a Euclidean domain $\Rightarrow$ A is a principal ideal domain $\Rightarrow$ A is a unique factorization domain (see any good graduate algebra course).

Noetherian rings.

Lemma 1.1. *The following conditions on a ring A are equivalent:*

(a) *Every ideal in A is finitely generated.*

(b) *Every ascending chain of ideals*

$$\mathfrak{a}_1 \subset \mathfrak{a}_2 \subset \cdots \subset \mathfrak{a}_n \subset \cdots$$

becomes stationary, i.e., after a certain point $\mathfrak{a}_n = \mathfrak{a}_{n+1} = \cdots$.

(c) *every nonempty set* S *of ideals in* A *has a maximal element* $\mathfrak{a}$*, i.e., there is an ideal* $\mathfrak{a}$ *in* S *that is not contained in any other ideal in* S.

PROOF. (a)⇒(b): Let $\mathfrak{a} = \cup\mathfrak{a}_i$; it is an ideal, and hence is finitely generated, say $\mathfrak{a} = (a_1, \dots, a_r)$. For some n, $\mathfrak{a}_n$ will contain all the a_i, and so $\mathfrak{a}_n = \mathfrak{a}_{n+1} = \cdots = \mathfrak{a}$.

(b)⇒(a): Consider an ideal $\mathfrak{a}$. If $\mathfrak{a} = (0)$, then $\mathfrak{a}$ is generated by the empty set, which is finite. Otherwise there is an element $a_1 \in \mathfrak{a}$, $a_1 \neq 0$. If $\mathfrak{a} = (a_1)$, then $\mathfrak{a}$ is certainly finitely generated. If not, there is an element $a_2 \in \mathfrak{a}$ such that $(a_1) \subsetneqq (a_1, a_2)$. Continuing in this way, we obtain a chain of ideals

$$(a_1) \subsetneqq (a_1, a_2) \subsetneqq \cdots .$$

This process must eventually stop with $(a_1, \dots, a_n) = \mathfrak{a}$.

(b)⇒(c): Let $\mathfrak{a}_1 \in S$. If $\mathfrak{a}_1$ is not a maximal element of S, then there is an $\mathfrak{a}_2 \in S$ such that $\mathfrak{a}_1 \subsetneqq \mathfrak{a}_2$. If $\mathfrak{a}_2$ is not maximal, then there is an $\mathfrak{a}_3$ etc.. From (b) we know that this process will lead to a maximal element after only finitely many steps.

(c)⇒(b): Apply (c) to the set $S = \{\mathfrak{a}_1, \mathfrak{a}_2, \dots\}$. □

A ring A satisfying the equivalent conditions of the lemma is said to be *Noetherian*[4]

A famous theorem of Hilbert states that $k[X_1, ..., X_n]$ is Noetherian. In practice, almost all the rings that arise naturally in algebraic number theory or algebraic geometry are Noetherian, but not all rings are Noetherian. For example, $k[X_1, \dots, X_n, \dots]$ is not Noetherian: $X_1, \dots, X_n$ is a minimal set of generators for the ideal $(X_1, \dots, X_n)$ in $k[X_1, \dots, X_n]$, and $X_1, \dots, X_n, \dots$ is a minimal set of generators for the ideal $(X_1, \dots, X_n, \dots)$ in $k[X_1, \dots, X_n, \dots]$

PROPOSITION 1.2. *Every nonzero nonunit element of a Noetherian integral domain can be written as a product of irreducible elements.*

PROOF. We shall need to use that

$$(a) \subset (b) \iff b|a, \text{ with equality } \iff b = a \times \text{unit}.$$

The first assertion is obvious. For the second, note that if $a = bc$ and $b = ad$ then $a = bc = adc$, and so $dc = 1$. Hence both c and d are units.

Suppose the statement is false, and choose an element $a \in A$ which contradicts the statement and is such that (a) is maximal among the ideals generated by such elements (here we use that A is Noetherian). Since a can not be written as a product of irreducible elements, it is not itself irreducible, and so $a = bc$ with b and c nonunits. Clearly $(b) \supset (a)$, and the ideals can't be equal for otherwise c would be a unit. From the maximality of (a), we deduce that b can be written as a product of irreducible elements, and similarly for c. Thus a is a product of irreducible elements, and we have a contradiction. □

[4]After Emmy Noether (1882–1935).

Local rings. A ring A is said to *local* if it has exactly one maximal ideal $\mathfrak{m}$. In this case, $A^\times = A \smallsetminus \mathfrak{m}$ (complement of $\mathfrak{m}$ in A).

LEMMA 1.3 (Nakayama's lemma). *Let A be a local Noetherian ring, and let $\mathfrak{a}$ be a proper ideal in A. Let M be a finitely generated A-module, and define*

$$\mathfrak{a}M = \{\textstyle\sum a_i m_i \mid a_i \in \mathfrak{a}, \quad m_i \in M\}.$$

(a) *If $\mathfrak{a}M = M$, then $M = 0$.*
(b) *If N is a submodule of M such that $N + \mathfrak{a}M = M$, then $N = M$.*

PROOF. (a) Suppose $M \neq 0$. Among the finite sets of generators for M, choose one $\{m_1, ..., m_k\}$ having the fewest elements. From the hypothesis, we know that we can write

$$m_k = a_1 m_1 + a_2 m_2 + ...a_k m_k \text{ some } a_i \in \mathfrak{a}.$$

Then

$$(1 - a_k)m_k = a_1 m_1 + a_2 m_2 + ... + a_{k-1} m_{k-1}.$$

As $1 - a_k$ is not in $\mathfrak{m}$, it is a unit, and so $\{m_1, ..., m_{k-1}\}$ generates M. This contradicts our choice of $\{m_1, ..., m_k\}$, and so $M = 0$.

(b) We shall show that $\mathfrak{a}(M/N) = M/N$, and then apply the first part of the lemma to deduce that $M/N = 0$. Consider $m + N$, $m \in M$. From the assumption, we can write

$$m = n + \textstyle\sum a_i m_i, \text{ with } a_i \in \mathfrak{a},\ m_i \in M.$$

Whence

$$m + N = \textstyle\sum a_i m_i + N = \sum a_i(m_i + N) \text{ (definition of the action of } A \text{ on } M/N),$$

and so $m + N \in \mathfrak{a}(M/N)$. □

The hypothesis that M be finitely generated in the lemma is crucial. For example, if A is a local integral domain with maximal ideal $\mathfrak{m} \neq 0$, then $\mathfrak{m}M = M$ for any field M containing A but $M \neq 0$.

Rings of fractions. Let A be an integral domain; there is a field $K \supset A$, called the *field of fractions* of A, with the property that every $c \in K$ can be written in the form $c = ab^{-1}$, $a, b \in A$, $b \neq 0$. For example, $\mathbb{Q}$ is the field of fractions of $\mathbb{Z}$, and $k(X)$ is the field of fractions of $k[X]$.

Let A be an integral domain with field of fractions K. A subset S of A is said to be *multiplicative* if $0 \notin S$, $1 \in S$, and S is closed under multiplication. If S is a multiplicative subset, then we define

$$S^{-1}A = \{a/b \in K \mid b \in S\}.$$

It is obviously a subring of K.

EXAMPLE 1.4. (a) Let t be a nonzero element of A; then

$$S_t \stackrel{\text{df}}{=} \{1, t, t^2, ...\}$$

is a multiplicative subset of A, and we (sometimes) write A_t for $S_t^{-1}A$. For example, if d is a nonzero integer,

$$\mathbb{Z}_d = \{a/d^n \in \mathbb{Q} \mid a \in \mathbb{Z},\ n \geq 0\}.$$

It consists of those elements of $\mathbb{Q}$ whose denominator divides some power of d.

(b) If $\mathfrak{p}$ is a prime ideal, then $S_{\mathfrak{p}} = A \smallsetminus \mathfrak{p}$ is a multiplicative set (if neither a nor b belongs to $\mathfrak{p}$, then ab does not belong to $\mathfrak{p}$). We write $A_{\mathfrak{p}}$ for $S_{\mathfrak{p}}^{-1}A$. For example,

$$\mathbb{Z}_{(p)} = \{m/n \in \mathbb{Q} \mid n \text{ is not divisible by } p\}.$$

PROPOSITION 1.5. *Let A be an integral domain, and let S be a multiplicative subset of A. The map*

$$\mathfrak{p} \mapsto S^{-1}\mathfrak{p} \stackrel{df}{=} \{a/s \mid a \in \mathfrak{p},\ s \in S\}$$

is a bijection from the set of prime ideals in A such that $\mathfrak{p} \cap S = \varnothing$ to the set of prime ideals in $S^{-1}A$; the inverse map is $\mathfrak{q} \mapsto \mathfrak{q} \cap A$.

PROOF. It is easy to see that

$$\mathfrak{p} \text{ a prime ideal disjoint from } S \Rightarrow S^{-1}\mathfrak{p} \text{ is a prime ideal,}$$

$$\mathfrak{q} \text{ a prime ideal in } S^{-1}A \Rightarrow \mathfrak{q} \cap A \text{ is a prime ideal disjoint from } S,$$

and so we only have to show that the two maps are inverse, i.e.,

$$(S^{-1}\mathfrak{p}) \cap A = \mathfrak{p} \text{ and } S^{-1}(\mathfrak{q} \cap A) = \mathfrak{q}.$$

$(S^{-1}\mathfrak{p})\cap A = \mathfrak{p}$: Clearly $(S^{-1}\mathfrak{p})\cap A \supset \mathfrak{p}$. For the reverse inclusion, let $a/s \in (S^{-1}\mathfrak{p})\cap A$, $a \in \mathfrak{p}$, $s \in S$. Consider the equation $\frac{a}{s} \cdot s = a \in \mathfrak{p}$. Both a/s and s are in A, and so at least one of a/s or s is in $\mathfrak{p}$ (because it is prime); but $s \notin \mathfrak{p}$ (by assumption), and so $a/s \in \mathfrak{p}$.

$S^{-1}(\mathfrak{q} \cap A) = \mathfrak{q}$: Clearly $S^{-1}(\mathfrak{q} \cap A) \subset \mathfrak{q}$ because $\mathfrak{q} \cap A \subset \mathfrak{q}$ and $\mathfrak{q}$ is an ideal in $S^{-1}A$. For the reverse inclusion, let $b \in \mathfrak{q}$. We can write it $b = a/s$ with $a \in A$, $s \in S$. Then $a = s \cdot (a/s) \in \mathfrak{q} \cap A$, and so $a/s = (s \cdot (a/s))/s \in S^{-1}(\mathfrak{q} \cap A)$. □

EXAMPLE 1.6. (a) If $\mathfrak{p}$ is a prime ideal in A, then $A_{\mathfrak{p}}$ is a local ring (because $\mathfrak{p}$ contains every prime ideal disjoint from $S_{\mathfrak{p}}$).

(b) We list the prime ideals in some rings:

$\mathbb{Z}$: $(2), (3), (5), (7), (11), \dots, (0)$;

$\mathbb{Z}_2$: $(3), (5), (7), (11), \dots, (0)$;

$\mathbb{Z}_{(2)}$: $(2), (0)$;

$\mathbb{Z}_{42}$: $(5), (11), (13), \dots, (0)$;

$\mathbb{Z}/(42)$: $(2), (3), (7)$.

Note that in general, for t a nonzero element of an integral domain,

$$\{\text{prime ideals of } A_t\} \quad \leftrightarrow \quad \{\text{prime ideals of } A \text{ not containing } t\}$$

$$\{\text{prime ideals of } A/(t)\} \quad \leftrightarrow \quad \{\text{prime ideals of } A \text{ containing } t\}.$$

The Chinese remainder theorem. Recall the classical form of the theorem: let $d_1, ..., d_n$ be integers, relatively prime in pairs; then for any integers $x_1, ..., x_n$, the equations

$$x \equiv x_i \pmod{d_i}$$

have a simultaneous solution $x \in \mathbb{Z}$; if x is one solution, then the other solutions are the integers of the form $x + md$, $m \in \mathbb{Z}$, where $d = \prod d_i$.

We want to translate this in terms of ideals. Integers m and n are relatively prime if and only if $(m, n) = \mathbb{Z}$, i.e., if and only if $(m) + (n) = \mathbb{Z}$. This suggests defining ideals $\mathfrak{a}$ and $\mathfrak{b}$ in a ring A to be *relatively prime* if $\mathfrak{a} + \mathfrak{b} = A$.

If $m_1, ..., m_k$ are integers, then $\cap(m_i) = (m)$ where m is the least common multiple of the m_i. Thus $\cap(m_i) \supset (\prod m_i) = \prod(m_i)$. If the m_i are relatively prime in pairs, then $m = \prod m_i$, and so we have $\cap(m_i) = \prod(m_i)$. Note that in general,

$$\mathfrak{a}_1 \cdot \mathfrak{a}_2 \cdot \cdots \mathfrak{a}_n \subset \mathfrak{a}_1 \cap \mathfrak{a}_2 \cap ... \cap \mathfrak{a}_n.$$

These remarks suggest the following statement.

THEOREM 1.7. *Let $\mathfrak{a}_1, ..., \mathfrak{a}_n$ be ideals in a ring A, relatively prime in pairs. Then for any elements $x_1, ..., x_n$ of A, the equations*

$$x \equiv x_i \quad (mod\ \mathfrak{a}_i)$$

have a simultaneous solution $x \in A$; if x is one solution, then the other solutions are the elements of the form $x + a$ with $a \in \cap\mathfrak{a}_i$; moreover, $\cap\mathfrak{a}_i = \prod \mathfrak{a}_i$. In other words, the natural maps give an exact sequence

$$0 \to \mathfrak{a} \to A \to \prod_{i=1}^{n} A/\mathfrak{a}_i \to 0$$

with $\mathfrak{a} = \cap\mathfrak{a}_i = \prod \mathfrak{a}_i$.

PROOF. Suppose first that $n = 2$. As $\mathfrak{a}_1 + \mathfrak{a}_2 = A$, there are elements $a_i \in \mathfrak{a}_i$ such that $a_1 + a_2 = 1$. The element $x =_{df} a_1x_2 + a_2x_1$ has the required property.

For each i we can find elements $a_i \in \mathfrak{a}_1$ and $b_i \in \mathfrak{a}_i$ such that

$$a_i + b_i = 1, \text{ all } i \geq 2.$$

The product $\prod_{i\geq 2}(a_i + b_i) = 1$, and lies in $\mathfrak{a}_1 + \prod_{i\geq 2} \mathfrak{a}_i$, and so

$$\mathfrak{a}_1 + \prod_{i\geq 2} \mathfrak{a}_i = A.$$

We can now apply the theorem in the case $n = 2$ to obtain an element y_1 of A such that

$$y_1 \equiv 1 \bmod \mathfrak{a}_1, \quad y_1 \equiv 0 \bmod \prod_{i\geq 2} \mathfrak{a}_i.$$

These conditions imply

$$y_1 \equiv 1 \bmod \mathfrak{a}_1, \quad y_1 \equiv 0 \bmod \mathfrak{a}_j, \text{ all } j > 1.$$

Similarly, there exist elements $y_2, ..., y_n$ such that

$$y_i \equiv 1 \bmod \mathfrak{a}_i, \quad y_i \equiv 0 \bmod \mathfrak{a}_j \text{ for } j \neq i.$$

The element $x = \sum x_iy_i$ now satisfies the requirements.

It remains to prove that $\cap\mathfrak{a}_i = \prod \mathfrak{a}_i$. We have already noted that $\cap\mathfrak{a}_i \supset \prod \mathfrak{a}_i$. First suppose that $n = 2$, and let $a_1 + a_2 = 1$, as before. For $c \in \mathfrak{a}_1 \cap \mathfrak{a}_2$, we have

$$c = a_1c + a_2c \in \mathfrak{a}_1 \cdot \mathfrak{a}_2$$

which proves that $\mathfrak{a}_1 \cap \mathfrak{a}_2 = \mathfrak{a}_1\mathfrak{a}_2$. We complete the proof by induction. This allows us to assume that $\prod_{i\geq 2} \mathfrak{a}_i = \cap_{i\geq 2}\mathfrak{a}_i$. We showed above that $\mathfrak{a}_1$ and $\prod_{i\geq 2} \mathfrak{a}_i$ are relatively prime, and so

$$\mathfrak{a}_1 \cdot (\prod_{i\geq 2} \mathfrak{a}_i) = \mathfrak{a}_1 \cap (\prod_{i\geq 2} \mathfrak{a}_i) = \cap\mathfrak{a}_i.$$

□

The theorem extends to A-modules.

THEOREM 1.8. *Let $\mathfrak{a}_1, ..., \mathfrak{a}_n$ be ideals in A, relatively prime in pairs, and let M be an A-module. There is an exact sequence:*

$$0 \to \mathfrak{a}M \to M \to \prod_i M/\mathfrak{a}_iM \to 0$$

with $\mathfrak{a} = \prod \mathfrak{a}_i = \cap\mathfrak{a}_i$.

This has an elementary proof (see Janusz 1996, p. 9), but I prefer to use tensor products, which I now review.

Review of tensor products. Let M, N, and P be A-modules. A mapping $f\colon M \times N \to P$ is said to be *A-bilinear* if

$$\begin{aligned} f(m + m', n) &= f(m, n) + f(m', n); \quad f(m, n + n') = f(m, n) + f(m, n') \\ f(am, n) &= af(m, n) = f(m, an), \quad a \in A, \quad m, m' \in M, \quad n, n' \in N, \end{aligned}$$

i.e., if it is linear in each variable. A pair (Q, f) consisting of an A-module Q and an A-bilinear map $f : M \times N \to Q$ is called the *tensor product* of M and N if any other A-bilinear map $f' : M \times N \to P$ factors uniquely into $f' = \alpha \circ f$ with $\alpha : Q \to P$ A-linear. The tensor product exists, and is unique (up to a unique isomorphism). We denote it by $M \otimes_A N$, and we write $(m, n) \mapsto m \otimes n$ for f. The pair $(M \otimes_A N, (m, n) \mapsto m \otimes n)$ is characterized by each of the following two conditions:

(a) The map $M \times N \to M \otimes_A N$ is A-bilinear, and any other A-bilinear map $M\times N \to P$ is of the form $(m, n) \mapsto \alpha(m\otimes n)$ for a unique A-linear map $\alpha\colon M\otimes_A N \to P$; thus

$$\mathrm{Bilin}_A(M \times N, P) = \mathrm{Hom}_A(M \otimes_A N, P).$$

(b) As an A-module, $M \otimes_A N$ generated by the symbols $m \otimes n$, $m \in M$, $n \in N$, which satisfy the relations

$$\begin{aligned} (m + m') \otimes n &= m \otimes n + m' \otimes n; \quad m \otimes (n + n') = m \otimes n + m \otimes n' \\ am \otimes n &= a(m \otimes n) = m \otimes an. \end{aligned}$$

Tensor products commute with direct sums: there is a canonical isomorphism

$$(\oplus_i M_i) \otimes_A (\oplus_j N_j) \xrightarrow{\approx} \oplus_{i,j} M_i \otimes_A N_j, \quad (\sum m_i) \otimes (\sum n_j) \mapsto \sum m_i \otimes n_j.$$

It follows that if M and N are free A-modules[5] with bases (e_i) and (f_j) respectively, then $M \otimes_A N$ is a free A-module with basis $(e_i \otimes f_j)$. In particular, if V and W are vector spaces over a field k of dimensions m and n respectively, then $V \otimes_k W$ is a vector space over k of dimension mn.

Let $\alpha\colon M \to N$ and $\beta\colon M' \to N'$ be A-linear maps. Then

$$(m,n) \mapsto \alpha(m) \otimes \beta(n)\colon M \times N \to M' \otimes_A N'$$

is A-bilinear, and therefore factors through $M \times N \to M \otimes_A N$. Thus there is an A-linear map $\alpha \otimes \beta : M \otimes_A N \to M' \otimes_A N'$ such that

$$(\alpha \otimes \beta)(m \otimes n) = \alpha(m) \otimes \beta(n).$$

REMARK 1.9. Let $\alpha\colon k^m \to k^m$ and $\beta\colon k^n \to k^n$ be two matrices, regarded as a linear maps. Then $\alpha \otimes \beta$ is a linear map $k^{mn} \to k^{mn}$. Its matrix with respect to the canonical basis is called the *Kronecker product* of the two matrices. (Kronecker products of matrices pre-date tensor products by about 70 years.)

LEMMA 1.10. *If $\alpha\colon M \to N$ and $\beta\colon M' \to N'$ are surjective, then so also is*

$$\alpha \otimes \beta\colon M \otimes_A N \to M' \otimes_A N'.$$

PROOF. Recall that $M' \otimes N'$ is generated as an A-module by the elements $m' \otimes n'$, $m' \in M'$, $n' \in N'$. By assumption $m' = \alpha(m)$ for some $m \in M$ and $n' = \beta(n)$ for some $n \in N$, and so $m' \otimes n' = \alpha(m) \otimes \beta(n) = (\alpha \otimes \beta)(m \otimes n)$. Therefore $\mathrm{Im}(\alpha \otimes \beta)$ contains a set of generators for $M' \otimes_A N'$ and so it is equal to it. □

One can also show that if

$$M\prime \to M \to M'' \to 0$$

is exact, then so also is

$$M' \otimes_A P \to M \otimes_A P \to M'' \otimes_A P \to 0.$$

For example, if we tensor the exact sequence

$$0 \to \mathfrak{a} \to A \to A/\mathfrak{a} \to 0$$

with M, we obtain an exact sequence

$$\mathfrak{a} \otimes_A M \to M \to (A/\mathfrak{a}) \otimes_A M \to 0$$

The image of $\mathfrak{a} \otimes M$ in M is

$$\mathfrak{a}M \stackrel{\text{df}}{=} \{\textstyle\sum a_i m_i \mid a_i \in \mathfrak{a},\, m_i \in M\},$$

and so we obtain from the exact sequence that

$$M/\mathfrak{a}M \cong (A/\mathfrak{a}A) \otimes_A M \qquad (1.11).$$

By way of contrast, if $M \to N$ is injective, then $M \otimes_A P \to N \otimes_A P$ need not be injective. For example, take $A = \mathbb{Z}$, and note that $(\mathbb{Z} \stackrel{m}{\to} \mathbb{Z}) \otimes_{\mathbb{Z}} (\mathbb{Z}/m\mathbb{Z})$ equals $\mathbb{Z}/m\mathbb{Z} \stackrel{m}{\to} \mathbb{Z}/m\mathbb{Z}$, which is the zero map.

[5]Let M be an A-module. Elements $e_1, \dots, e_m$ form a *basis* for M if every element of M can be expressed uniquely as a linear combination of the e_i's with coefficients in A. Then $A^m \to M$, $(a_1, \dots, a_m) \mapsto \sum a_i e_i$, is an isomorphism of A-modules, and M is said to be a *free A-module of rank m*.

PROOF OF THEOREM 1.8. Return to the situation of the theorem. When we tensor the isomorphism

$$A/\mathfrak{a} \xrightarrow{\cong} \prod A/\mathfrak{a}_i$$

with M, we get an isomorphism

$$M/\mathfrak{a}M \cong (A/\mathfrak{a}) \otimes_A M \xrightarrow{\cong} \prod (A/\mathfrak{a}_i) \otimes_A M \cong \prod M/\mathfrak{a}_i M,$$

as required. □

Extension of scalars. If $A \to B$ is an A-algebra and M is an A-module, then $B \otimes_A M$ has a natural structure of a B-module for which

$$b(b' \otimes m) = bb' \otimes m, \quad b, b' \in B, \quad m \in M.$$

We say that $B \otimes_A M$ is the B-module obtained from M by *extension of scalars.* The map $m \mapsto 1 \otimes m \colon M \to B \otimes_A M$ is uniquely determined by the following universal property: it is A-linear, and for any A-linear map $\alpha \colon M \to N$ from M into a B-module N, there is a unique B-linear map $\alpha' \colon B \otimes_A M \to N$ such that $\alpha'(1 \otimes m) = \alpha(m)$. Thus $\alpha \mapsto \alpha'$ defines an isomorphism

$$\mathrm{Hom}_A(M, N) \to \mathrm{Hom}_B(B \otimes_A M, N), N \text{ a } B\text{-module}).$$

For example, $A \otimes_A M = M$. If M is a free A-module with basis $e_1, \dots, e_m$, then $B \otimes_A M$ is a free B-module with basis $1 \otimes e_1, \dots, 1 \otimes e_m$.

Tensor products of algebras. If $f \colon A \to B$ and $g \colon A \to C$ are A-algebras, then $B \otimes_A C$ has a natural structure of an A-algebra: the product structure is determined by the rule

$$(b \otimes c)(b' \otimes c') = bb' \otimes cc'$$

and the map $A \to B \otimes_A C$ is $a \mapsto f(a) \otimes 1 = 1 \otimes g(a)$.

For example, there is a canonical isomorphism

$$a \otimes f \mapsto af : K \otimes_k k[X_1, \dots, X_m] \to K[X_1, \dots, X_m] \qquad (1.12).$$

Tensor products of fields. We are now able to compute $K \otimes_k \Omega$ if K is a finite separable field extension of k and Ω is an arbitrary field extension of k. According to the primitive element theorem (FT, 5.1), $K = k[\alpha]$ for some $\alpha \in K$. Let $f(X)$ be the minimum polynomial of α. By definition this means that the map $g(X) \mapsto g(\alpha)$ determines an isomorphism

$$k[X]/(f(X)) \to K.$$

Hence

$$K \otimes_k \Omega \cong (k[X]/(f(X))) \otimes \Omega \cong \Omega[X]/(f(X))$$

by (1.11) and (1.12). Because K is separable over k, $f(X)$ has distinct roots. Therefore $f(X)$ factors in $\Omega[X]$ into monic irreducible polynomials

$$f(X) = f_1(X) \cdots f_r(X)$$

that are relatively prime in pairs. We can apply the Chinese Remainder Theorem to deduce that

$$\Omega[X]/(f(X)) = \prod_{i=1}^{r} \Omega[X]/(f_i(X)).$$

Finally, $\Omega[X]/(f_i(X))$ is a finite separable field extension of Ω of degree $\deg f_i$. Thus we have proved the following result:

THEOREM 1.13. *Let K be a finite separable field extension of k, and let Ω be an arbitrary field extension. Then $K \otimes_k \Omega$ is a product of finite separable field extensions of Ω,*

$$K \otimes_k \Omega = \prod_{i=1}^{r} \Omega_i.$$

If α is a primitive element for K/k, then the image α_i of α in Ω_i is a primitive element for Ω_i/Ω, and if $f(X)$ and $f_i(X)$ are the minimum polynomials for α and α_i respectively, then

$$f(X) = \prod_{i=1}^{r} f_i(X).$$

EXAMPLE 1.14. Let $K = \mathbb{Q}[\alpha]$ with α algebraic over $\mathbb{Q}$. Then

$$\mathbb{C} \otimes_{\mathbb{Q}} K \cong \mathbb{C} \otimes_{\mathbb{Q}} (\mathbb{Q}[X]/(f(X))) \cong \mathbb{C}[X]/((f(X)) \cong \prod \mathbb{C}[X]/(X - \alpha_i) \approx \mathbb{C}^r.$$

Here $\alpha_1, \dots, \alpha_r$ are the conjugates of α in $\mathbb{C}$. The composite of $\beta \mapsto 1 \otimes \beta \colon K \to \mathbb{C} \otimes_{\mathbb{Q}} K$ with projection onto the ith factor is $\sum a_j \alpha^j \mapsto \sum a_j \alpha_i^j$.

Finally we note that it is essential to assume in (1.13) that K is separable over k. If not, there will be an $\alpha \in K$ such that $\alpha^p = a \in k$ but $\alpha \notin k$. The ring $K \otimes_k K$ contains an element $\beta = (\alpha \otimes 1 - 1 \otimes \alpha) \neq 0$ such that

$$\beta^p = a \otimes 1 - 1 \otimes a = a(1 \otimes 1) - a(1 \otimes 1) = 0.$$

Hence $K \otimes_k K$ contains a nonzero nilpotent element, and so can't be a product of fields.

2. Rings of Integers

Let A be an integral domain, and let L be a field containing A. An element α of L is said to be *integral* over A if it is a root of a *monic* polynomial with coefficients in A, i.e., if it satisfies an equation

$$\alpha^n + a_1\alpha^{n-1} + ... + a_n = 0, \quad a_i \in A.$$

Before proving that the elements of L integral over A form a ring, we need to review symmetric polynomials.

Symmetric polynomials. A polynomial $P(X_1, ..., X_r) \in A[X_1, \dots, X_r]$ is said to be *symmetric* if it is unchanged when its variables are permuted, i.e., if

$$P(X_{\sigma(1)}, \dots, X_{\sigma(r)}) = P(X_1, \dots, X_r), \quad \text{all } \sigma \in \mathrm{Sym}_r.$$

For example

$$S_1 = \sum X_i, \quad S_2 = \sum_{i<j} X_iX_j, \quad \dots, \quad S_r = X_1 \cdots X_r,$$

are all symmetric. These particular polynomials are called the *elementary symmetric polynomials*.

Theorem 2.1. *(Symmetric function theorem) Let A be a ring. Every symmetric polynomial $P(X_1, ..., X_r)$ in $A[X_1, ..., X_r]$ is equal to a polynomial in the symmetric elementary polynomials with coefficients in A, i.e., $P \in A[S_1, ..., S_r]$.*

Proof. We define an ordering on the monomials in the X_i by requiring that

$$X_1^{i_1}X_2^{i_2}\cdots X_r^{i_r} > X_1^{j_1}X_2^{j_2}\cdots X_r^{j_r}$$

if either

$$i_1 + i_2 + \cdots + i_r > j_1 + j_2 + \cdots + j_r$$

or equality holds and, for some s,

$$i_1 = j_1, \quad \dots, \quad i_s = j_s, \text{ but } i_{s+1} > j_{s+1}.$$

Let $X_1^{k_1}\cdots X_r^{k_r}$ be the highest monomial occurring in P with a coefficient $c \neq 0$. Because P is symmetric, it contains all monomials obtained from $X_1^{k_1}\cdots X_r^{k_r}$ by permuting the X's. Hence $k_1 \geq k_2 \geq \cdots \geq k_r$.

Clearly, the highest monomial in S_i is $X_1 \cdots X_i$, and it follows easily that the highest monomial in $S_1^{d_1}\cdots S_r^{d_r}$ is

$$X_1^{d_1+d_2+\cdots+d_r}X_2^{d_2+\cdots+d_r}\cdots X_r^{d_r}.$$

Therefore

$$P(X_1, \dots, X_r) - cS_1^{k_1-k_2}S_2^{k_2-k_3}\cdots S_r^{k_r} < P(X_1, \dots, X_r).$$

We can repeat this argument with the polynomial on the left, and after a finite number of steps, we will arrive at a representation of P as a polynomial in $S_1, \dots, S_r$. □

Let $f(X) = X^n + a_1X^{n-1} + \cdots + a_n \in A[X]$, and let $\alpha_1, \ldots, \alpha_n$ be the roots of $f(X)$ in some ring containing A, so that $f(X) = \prod(X - \alpha_i)$ in some larger ring. Then

$$a_1 = -S_1(\alpha_1, \ldots, \alpha_n), \quad a_2 = S_2(\alpha_1, \ldots, \alpha_n), \quad \ldots, \quad a_n = \pm S_n(\alpha_1, \ldots, \alpha_n).$$

Thus the *elementary* symmetric polynomials in the roots of $f(X)$ lie in A, and so the theorem implies that *every* symmetric polynomial in the roots of $f(X)$ lies in A.

Integral elements.

THEOREM 2.2. *The set of elements of L integral over A forms a ring.*

PROOF. I shall give two proofs, first an old-fashioned proof, and later the slick modern proof. Suppose α and β are integral over A; I'll prove only that $\alpha + \beta$ is integral over A since the same proof works for $\alpha - \beta$ and $\alpha\beta$. Let Ω be an algebraically closed field containing L.

We are given that α is a root of a polynomial $f(X) = X^m + a_1X^{m-1} + \cdots + a_m$, $a_i \in A$. Write

$$f(X) = \prod(X - \alpha_i),\ \alpha_i \in \Omega.$$

Similarly, β is a root of polynomial $g(X) = X^n + b_1X^{n-1} + \cdots + b_n$, $b_i \in A$, and we write

$$f(X) = \prod(X - \beta_i),\ \beta_i \in \Omega.$$

Let $\gamma_1, \gamma_2, ..., \gamma_{mn}$ be the family of numbers of the form $\alpha_i + \beta_j$ (or $\alpha_i - \beta_j$, or $\alpha_i\beta_j$). I claim that $h(X) =_{df} \prod(X - \gamma_{ij})$ has coefficients in A. This will prove that $\alpha + \beta$ is integral over A because h is monic and $h(\alpha + \beta) = 0$.

The coefficients of h are symmetric in the α_i and β_j. Let $P(\alpha_1, ..., \alpha_m, \beta_1, ..., \beta_n)$ be one of these coefficients, and regard it as a polynomial $Q(\beta_1, ..., \beta_n)$ in the β's with coefficients in $A[\alpha_1, ..., \alpha_m]$; then its coefficients are symmetric in the α_i, and so lie in A. Thus $P(\alpha_1, ..., \alpha_m, \beta_1, ..., \beta_n)$ is a symmetric polynomial in the β's with coefficients in A — it therefore lies in A, as claimed. □

DEFINITION 2.3. The ring of elements of L integral over A is called the *integral closure* of A in L. The integral closure of $\mathbb{Z}$ in an algebraic number field L is called the *ring of integers* $\mathcal{O}_L$ in L.

Next we want to see that L is the field of fractions of $\mathcal{O}_L$; in fact we can prove more.

PROPOSITION 2.4. *Let K be the field of fractions of A, and let L be a field containing K. If $\alpha \in L$ is algebraic over K, then there exists a $d \in A$ such that $d\alpha$ is integral over A.*

PROOF. By assumption, α satisfies an equation

$$\alpha^m + a_1\alpha^{m-1} + \cdots + a_m = 0,\ a_i \in K.$$

Let d be a common denominator for the a_i, so that $da_i \in A$ for all i, and multiply through the equation by d^m :

$$d^m\alpha^m + a_1d^m\alpha^{m-1} + ... + a_md^m = 0.$$

We can rewrite this as

$$(d\alpha)^m + a_1d(d\alpha)^{m-1} + \cdots + a_md^m = 0.$$

As a_1d, ... , $a_md^m \in A$, this shows that $d\alpha$ is integral over A. □

COROLLARY 2.5. *Let A be an integral domain with field of fractions K, and let L be an algebraic extension of K. If B is the integral closure of A in L, then L is the field of fractions of B.*

PROOF. The proposition shows that every $\alpha \in L$ can be written $\alpha = \beta/d$ with $\beta \in B$, $d \in A$. □

DEFINITION 2.6. A ring A is *integrally closed* if it is its own integral closure in its field of fractions K, i.e., if

$$\alpha \in K, \quad \alpha \text{ integral over } A \Rightarrow \alpha \in A.$$

PROPOSITION 2.7. *A unique factorization domain (e.g. a principal ideal domain) is integrally closed.*

PROOF. Suppose a/b, $a, b \in A$, is an element of the field of fractions of A that is integral over A. If b is a unit, then $a/b \in A$. Otherwise we may suppose that there is an irreducible element p of A dividing b but not a. As a/b is integral over A, it satisfies an equation

$$(a/b)^n + a_1(a/b)^{n-1} + \cdots + a_n = 0,\ a_i \in A.$$

On multiplying through by b^n, we obtain the equation

$$a^n + a_1a^{n-1}b + ... + a_nb^n = 0.$$

The element p then divides every term on the left except a^n, and hence must divide a^n. Since it doesn't divide a, this is a contradiction. □

Hence it is easy to get examples where unique factorization fails — take any ring which is not integrally closed, for example, $\mathbb{Z}[\sqrt{5}]$.

EXAMPLE 2.8. (a) The rings $\mathbb{Z}$ and $\mathbb{Z}[i]$ are integrally closed — both are principal ideal domains.

(b) Let k be a field. I claim that the integral closure of $k[S_1, ..., S_m]$ in $k(X_1, \ldots, X_m)$ is $k[X_1, \ldots, X_m]$ (here the S_i are the elementary symmetric polynomials).

Let $f \in k(X_1, \ldots, X_m)$ be integral over $k[S_1, \ldots, S_m]$. Then f is integral over $k[X_1, \ldots, X_m]$, which is a unique factorization domain, and hence is integrally closed in its field of fractions. Thus $f \in k[X_1, \ldots, X_m]$.

Conversely, let $f \in k[X_1, \ldots, X_m]$. Then f is a root of the monic polynomial

$$\prod_{\sigma \in \mathrm{Sym}_m} (T - f(X_{\sigma(1)}, \ldots, X_{\sigma(m)})).$$

The coefficients of this polynomial are symmetric polynomials in the X_i, and therefore (see 2.1) lie in $k[S_1, \dots, S_r]$.

PROPOSITION 2.9. *Let K be the field of fractions of A, and let L be an extension of K of finite degree. Assume A is integrally closed. An element α of L is integral over A if and only if its minimum polynomial over K has coefficients in A.*

PROOF. Assume α is integral over A, so that

$$\alpha^m + a_1\alpha^{m-1} + \dots + a_m = 0, \quad \text{some } a_i \in A.$$

Let α' be a conjugate of α, i.e., a root of the minimum polynomial of α over K. Then there is an K-isomorphism

$$\sigma : K[\alpha] \to K[\alpha'], \quad \sigma(\alpha) = \alpha';$$

see[6] FT. On applying σ to the above equation we obtain the equation

$$\alpha'^m + a_1\alpha'^{m-1} + \dots + a_m = 0,$$

which shows that α' is integral over A. Hence all the conjugates of α are integral over A, and it follows from (2.2) that the coefficients of $f(X)$ are integral over A. They lie in K, and A is integrally closed, and so they lie in A. This proves the "only if" part of the statement, and the "if" part is obvious. □

REMARK 2.10. As we noted in the introduction, this makes it easy to compute some rings of integers. For example, an element $\alpha \in \mathbb{Q}[\sqrt{d}]$ is integral over $\mathbb{Z}$ if and only if its trace and norm both lie in $\mathbb{Z}$.

PROPOSITION 2.11. *Let L be a field containing A. An element α of L is integral over A if and only if there is a nonzero finitely generated A-submodule of L such that $\alpha M \subset M$ (in fact, we can take $M = A[\alpha]$, the A-subalgebra generated by α).*

PROOF. $\Rightarrow$: Suppose

$$\alpha^n + a_1\alpha^{n-1} + \dots + a_n = 0, \quad a_i \in A.$$

Then the A-submodule M of L generated by 1, α, ..., α^{n-1} has the property that $\alpha M \subset M$.

$\Leftarrow$: We shall need to apply Cramer's rule. As usually stated (in linear algebra courses) this says that, if

$$\sum_{j=1}^{m} c_{ij}x_j = d_i, \quad i = 1, \dots, m,$$

then

$$x_j = \det(C_j)/\det(C)$$

where $C = (c_{ij})$ and C_j is obtained from C by replacing the elements of the j^{th} column with the d_is. When one restates the equation as

$$\det(C) \cdot x_j = \det(C_j)$$

[6]If $f(X)$ is the minimum polynomial of α, hence also of α', over K, then the map $h(X) \mapsto h(\alpha)\colon K[X] \to K[\alpha]$ induces an isomorphism $\tau\colon K[X]/(f(X)) \to K[\alpha]$. Similarly, $h(X) \mapsto h(\alpha')\colon K[X] \to K[\alpha']$ induces an isomorphism $\tau'\colon K[X]/(f(X)) \to K[\alpha']$, and we set $\sigma = \tau' \circ \tau^{-1}$.

it becomes true over any ring (whether or not $\det(C)$ is invertible). The proof is elementary—essentially it is what you wind up with when you eliminate the other variables (try it for $m = 2$). Alternatively, expand out

$$\det C_j = \begin{vmatrix} c_{11} & \dots & -\sum c_{1j}x_j & \dots & c_{1m} \\ & \cdots & & \cdots & \\ c_{m1} & \dots & -\sum c_{mj}x_j & \dots & c_{mm} \end{vmatrix}$$

using standard properties of determinants.

Now let M be a nonzero A-module in L such that $\alpha M \subset M$, and let $v_1, \dots, v_n$ be a finite set of generators for M. Then, for each i,

$$\alpha v_i = \sum a_{ij} v_j, \text{ some } a_{ij} \in A.$$

We can rewrite this system of equations as

$$\begin{aligned} (\alpha - a_{11})v_1 - a_{12}v_2 - a_{13}v_3 - \cdots &= 0 \\ -a_{21}v_1 + (\alpha - a_{22})v_2 - a_{23}v_3 - \cdots &= 0 \\ \cdots &= 0. \end{aligned}$$

Let C be the matrix of coefficients on the left-hand side. Then Cramer's rule tells us that $\det(C) \cdot v_i = 0$ for all i. Since at least one v_i is nonzero and we are working inside the field L, this implies that $\det(C) = 0$. On expanding out the determinant, we obtain an equation

$$\alpha^n + c_1\alpha^{n-1} + c_2\alpha^{n-2} + \cdots + c_n = 0, \quad c_i \in A.$$

□

We now give a second proof that if A is a subring of a field L and B is the set of elements of L that are integral over A, then B is a ring. Let α and β be two elements of L integral over A, and let M and N be finitely generated A-modules in L such that $\alpha M \subset M$ and $\beta N \subset N$. Define

$$MN = \{\sum m_i n_i \mid m_i \in M, \quad n_i \in N\}.$$

Then:

(a) MN is an A-submodule of L (easy);
(b) it is finitely generated — if $\{e_1, \dots, e_m\}$ generates M and $\{f_1, \dots, f_n\}$ generates N, then $\{e_1f_1, \dots, e_if_j, \dots, e_mf_n\}$ generates MN;
(c) it is stable under multiplication by $\alpha\beta$ and by $\alpha \pm \beta$.

We can now apply (2.11) to deduce that $\alpha\beta$ and $\alpha \pm \beta$ are integral over A.

PROPOSITION 2.12. *If B is integral over A and finitely generated as an A-algebra, then it is finitely generated as an A-module.*

PROOF. First consider the case that B is generated as an A-algebra by a single element, say $B = A[\beta]$. By assumption

$$\beta^n + a_1\beta^{n-1} + \cdots + a_n = 0, \text{ some } a_i \in A.$$

Every element of B is a finite sum

$$c_0 + c_1\beta + c_2\beta^2 + \cdots + c_N\beta^N,$$

and we can exploit the preceding equality to replace β^n (successively) with a linear combination of lower powers of β. Thus every element of B is of the form

$$c_0 + c_1\beta + c_2\beta^2 + \cdots + c_{n-1}\beta^{n-1},$$

i.e., $1, \beta, \beta^2, \ldots, \beta^{n-1}$ generate B as an A-module. In order to pass to the general case, we need a lemma. □

LEMMA 2.13. *Let $A \subset B \subset C$ be rings. If B is finitely generated as an A-module, and C is finitely generated as a B-module, then C is finitely generated as an A-module.*

PROOF. If $\{\beta_1, ..., \beta_m\}$ is a set of generators for B as an A-module, and $\{\gamma_1, ..., \gamma_n\}$ is a set of generators for C as a B-module, then $\{\beta_i\gamma_j\}$ is a set of generators for C as an A-module. □

We now complete the proof of (2.12). Let $\beta_1, \ldots, \beta_m$ generate B as an A-algebra, and consider

$$A \subset A[\beta_1] \subset A[\beta_1, \beta_2] \subset \cdots \subset A[\beta_1, ..., \beta_m] = B.$$

We saw above that $A[\beta_1]$ is finitely generated as an A-module. Since $A[\beta_1, \beta_2] = A[\beta_1][\beta_2]$, and β_2 is integral over $A[\beta_1]$ (because it is over A), the same observation shows that $A[\beta_1, \beta_2]$ is finitely generated as a $A[\beta_1]$-module. Now the lemma shows that $A[\beta_1, \beta_2]$ is finitely generated as an A-module. Continuing in this fashion, we find that B is finitely generated as an A-module.

PROPOSITION 2.14. *Consider integral domains $A \subset B \subset C$; if B is integral over A, and C is integral over B, then C is integral over A.*

PROOF. Let $\gamma \in C$; it satisfies an equation

$$\gamma^n + b_1\gamma^{n-1} + \cdots + b_n = 0, \quad b_i \in B.$$

Let $B' = A[b_1, ..., b_n]$. Then B' is finitely generated as an A-module (by the last proposition), and γ is integral over B' (by our choice of the b_i), and so $B'[\gamma]$ is finitely generated as an A-module. Since $\gamma B'[\gamma] \subset B'[\gamma]$, Proposition 2.11 shows that γ is integral over A. □

COROLLARY 2.15. *The integral closure of A in an algebraic extension L of its field of fractions is integrally closed.*

PROOF. Let B be the integral closure of A in L. We know from (2.5) that L is the field of fractions of B. If $\gamma \in L$ is integral over B, then the proposition shows that it is integral over A, and so lies in B. □

REMARK 2.16. In particular, the ring of integers in a number field is integrally closed. Clearly we want this, since we want our ring of integers to have the best chance of being a unique factorization domain (see 2.7).

EXAMPLE 2.17. Let k be a finite field, and let K be a finite extension of $k(X)$. Let $\mathcal{O}_K$ be the integral closure of $k[X]$ in K. The arithmetic of $\mathcal{O}_K$ is very similar to that of the ring of integers in a number field.

Review of bases of A-modules. Let M be an A-module. Recall that a set of elements $e_1, ..., e_n$ is a *basis* for M if

(a) $\sum a_i e_i = 0$, $a_i \in A \Rightarrow$ all $a_i = 0$, and
(b) every element x of M can be expressed in the form $x = \sum a_i e_i$, $a_i \in A$.

Let $\{e_1, ..., e_n\}$ be a basis for M, and let $\{f_1, ..., f_n\}$ be a second set of n elements in M. Then we can write $f_i = \sum a_{ij} e_j$, $a_{ij} \in A$, and f_i is also a basis if and only if the matrix (a_{ij}) is invertible in the ring $M_n(A)$ of $n \times n$ matrices with coefficients in A (this is obvious). Moreover (a_{ij}) is invertible in $M_n(A)$ if and only if its determinant is a unit in A, and in this case, the inverse is given by the usual formula:

$$(a_{ij})^{-1} = \mathrm{adj}(a_{ij}) \cdot \det(a_{ij})^{-1}.$$

In the case that $A = \mathbb{Z}$, the index of $N =_{df} \mathbb{Z}f_1 + \mathbb{Z}f_2 + \cdots + \mathbb{Z}f_n$ in M is $|\det(a_{ij})|$ (assuming this is nonzero). To prove this, recall from basic graduate algebra that we can choose bases $\{e_i'\}$ for M and $\{f_i'\}$ for N such that $f_i' = m_i e_i'$, $m_i \in \mathbb{Z}$. If $(e_i') = U \cdot (e_i)$ and $(f_i') = V \cdot (f_i)$, then $(f_i) = V^{-1} D U (e_i)$ where $D = \mathrm{diag}(m_1, \ldots, m_n)$, and

$$\det(V^{-1}DU) = \det(V^{-1}) \cdot \det(D) \cdot \det(U) = \pm \prod m_i = \pm(M : N).$$

Review of norms and traces. Let $A \subset B$ be rings, and assume that B is a free A-module of rank n. Then any $\beta \in B$ defines an A-linear map

$$x \mapsto \beta x : B \to B,$$

and the trace and determinant of this map are well-defined. We call them the *trace* $\mathrm{Tr}_{B/A}\,\beta$ and *norm* $\mathrm{Nm}_{B/A}\,\beta$ of β in the extension B/A. Thus if $\{e_1, ..., e_n\}$ is a basis for B over A, and $\beta e_i = \sum a_{ij} e_j$, then $\mathrm{Tr}_{B/A}(\beta) = \sum a_{ii}$ and $\mathrm{Nm}_{B/A}(\beta) = \det(a_{ij})$. When $B \supset A$ is a finite field extension, this agrees with the usual definition. The following hold:

$$\mathrm{Tr}(\beta + \beta') = \mathrm{Tr}(\beta) + \mathrm{Tr}(\beta'); \quad \mathrm{Tr}(a\beta) = a\,\mathrm{Tr}(\beta); \quad \mathrm{Tr}(a) = na \quad (a \in A);$$

$$\mathrm{Nm}(\beta\beta') = \mathrm{Nm}(\beta) \cdot \mathrm{Nm}(\beta'); \quad \mathrm{Nm}(a) = a^n \quad (a \in A).$$

PROPOSITION 2.18. *Let L/K be an extension of fields of degree n, and let $\beta \in L$. Let $f(X)$ be the minimum polynomial of β over K and let $\beta_1 = \beta$, β_2, ... , β_m be the roots of $f(X)$. Then*

$$\mathrm{Tr}_{L/K}\,\beta = r(\beta_1 + \cdots + \beta_m), \quad \mathrm{Nm}_{L/K}\,\beta = (\beta_1 \cdots \beta_m)^r$$

where $r = [L : K[\beta]] = n/m$.

PROOF. Suppose first that $L = K[\beta]$, and compute the matrix of $x \mapsto \beta x$ relative to the basis $\{1, \beta, \ldots, \beta^{n-1}\}$—one sees easily that it has trace $\sum \beta_i$ and determinant $\prod \beta_i$. For the general case, use the transitivity of norms and traces (see FT, Proposition 5.37). □

COROLLARY 2.19. *Assume L is separable of degree n over K, and let $\{\sigma_1, ..., \sigma_n\}$ be the set of distinct K-homomorphisms $L \hookrightarrow \Omega$ where Ω is some big Galois extension of K (e.g., the Galois closure of L over K). Then*

$$\mathrm{Tr}_{L/K}\,\beta = \sigma_1\beta + \cdots + \sigma_n\beta, \quad \mathrm{Nm}_{L/K}\,\beta = \sigma_1\beta \cdots \sigma_n\beta.$$

PROOF. Each β_i occurs exactly r times in the family $\{\sigma_i\beta\}$—see FT §5.9. □

COROLLARY 2.20. *Let A be an integrally closed integral domain, and let L be a finite extension of the field of fractions K of A; if $\beta \in L$ is integral over A, then $\mathrm{Tr}_{L/K}\,\beta$ and $\mathrm{Nm}_{L/K}\,\beta$ are in A.*

PROOF. We know that if β is integral, then so also is each of its conjugates. Alternatively, apply 2.9. □

Review of bilinear forms. Let V be a finite-dimensional vector space over a field K. A *bilinear form* on V is a map

$$\psi\colon V \times V \to K$$

such that $x \mapsto \psi(x, v)$ and $x \mapsto \psi(v, x)$ are both linear maps $V \to K$ for all $v \in V$. The *discriminant* of a symmetric bilinear form relative to a basis $\{e_1, ..., e_m\}$ of V is $\det(\psi(e_i, e_j))$. If $\{f_1, ..., f_m\}$ is a set of elements of V, and $f_j = \sum a_{ji}e_i$, then

$$\psi(f_k, f_l) = \sum_{i,j} \psi(a_{ki}e_i, a_{lj}e_j) = \sum_{i,j} a_{ki} \cdot \psi(e_i, e_j) \cdot a_{lj},$$

and so

$$(\psi(f_k, f_l)) = (a_{ki}) \cdot (\psi(e_i, e_j)) \cdot (a_{jl})^{\mathrm{tr}}$$

(equality of $m \times m$ matrices). Hence

$$\det(\psi(f_i, f_j)) = \det(a_{ij})^2 \cdot \det(\psi(e_i, e_j)) \qquad (2.21)$$

The form ψ is said to be *nondegenerate* if it satisfies each of the following equivalent conditions:

(a) ψ has a nonzero discriminant relative to one (hence every) basis of V;
(b) the left kernel $\{v \in V \mid \psi(v, x) = 0 \text{ for all } x \in V\}$ is zero;
(c) the right kernel of ψ is zero.

Thus if ψ is nondegenerate, the map $v \mapsto (x \mapsto \psi(v, x))$ from V onto the dual vector space $V^\vee \overset{\mathrm{df}}{=} \mathrm{Hom}(V, K)$ is an isomorphism. Let $\{e_1, ..., e_m\}$ be a basis for V, and let $f_1, ..., f_m$ be the dual basis in $V^\vee$, i.e., $f_i(e_j) = \delta_{ij}$ (Kronecker delta). We can use the isomorphism $V \to V^\vee$ given by a nondegenerate form ψ to transfer $\{f_1, ..., f_m\}$ to a basis $\{e'_1, ..., e'_m\}$ of V; it has the property that

$$\psi(e'_i, e_j) = \delta_{ij}.$$

For example, suppose $\{e_1, ..., e_m\}$ is a basis such that $(\psi(e_i, e_j))$ is a diagonal matrix — the Gram-Schmidt process always allows us to find such a basis — then $e'_i = e_i/\psi(e_i, e_i)$.

Discriminants. If L is a finite extension of K (L and K fields), then

$$(\alpha, \beta) \mapsto \mathrm{Tr}_{L/K}(\alpha\beta)\colon L \times L \to K$$

is a symmetric bilinear form on L (regarded as a vector space over K), and the discriminant of this form (relative to any basis for L as a K-vector space) is called the *discriminant* of L/K.

More generally, let $B \supset A$ be rings, and assume B is free of rank m as an A-module. Let $\beta_1, ..., \beta_m$ be elements of B. We define their *discriminant* to be

$$D(\beta_1, ..., \beta_m) = \det(\mathrm{Tr}_{B/A}(\beta_i \beta_j)).$$

LEMMA 2.22. *If* $\gamma_j = \sum a_{ji}\beta_i$, $a_{ij} \in A$, *then*

$$D(\gamma_1, ..., \gamma_m) = \det(a_{ij})^2 \cdot D(\beta_1, ..., \beta_m).$$

PROOF. See the proof of (2.21). □

If the β's and γ's both form a basis for B over A, then $\det(a_{ij})$ is a unit (see p. 26). Thus the discriminant $D(\beta_1, ..., \beta_m)$ of a basis $\{\beta_1, ..., \beta_m\}$ of B is well-defined up to multiplication by the square of a unit in A. In particular, the ideal in A that it generates is independent of the choice of the basis. This ideal, or $D(\beta_1, ..., \beta_m)$ itself regarded as an element of $A/A^{\times 2}$, is called the *discriminant* $\mathrm{disc}(B/A)$ of B over A.

For example, when we have a finite extension of fields L/K, $\mathrm{disc}(L/K)$ is an element of K, well-defined up to multiplication by a nonzero square in K.

When $A = \mathbb{Z}$, $\mathrm{disc}(B/A)$ is a well-defined integer, because 1 is the only square of a unit in $\mathbb{Z}$.

Warning: We shall see shortly that, when K is a number field of degree m over $\mathbb{Q}$, the ring of integers $\mathcal{O}_K$ in K is free of rank m over $\mathbb{Z}$, and so $\mathrm{disc}(\mathcal{O}_K/\mathbb{Z})$ is a well-defined integer. Sometimes this is loosely referred to as the discriminant of $K/\mathbb{Q}$ — strictly speaking, $\mathrm{disc}(K/\mathbb{Q})$ is the element of $\mathbb{Q}^\times/\mathbb{Q}^{\times 2}$ represented by the integer $\mathrm{disc}(\mathcal{O}_K/\mathbb{Z})$.

PROPOSITION 2.23. *Let* $A \subset B$ *be integral domains and assume that* B *is a free* A*-module of rank* m *and that* $\mathrm{disc}(B/A) \neq 0$. *Elements* $\gamma_1, ..., \gamma_m$ *form a basis for* B *as an* A*-module if and only if*

$$(D(\gamma_1, ..., \gamma_m)) = (\mathrm{disc}(B/A)) \quad \textit{(as ideals in } A\textit{)}.$$

PROOF. Let $\{\beta_1, ..., \beta_m\}$ be a basis for B as an A-module, and let $\gamma_1, ..., \gamma_m$ be any elements of B. Write $\gamma_j = \sum a_{ji}\beta_i$, $a_{ji} \in A$. Then $D(\gamma_1, ..., \gamma_m) = \det(a_{ij})^2 \cdot D(\beta_1, ..., \beta_m)$ and as we noted in the subsection "Review of bases of A-modules", $\{\gamma_1, \dots, \gamma_m\}$ is a basis if and only if $\det(a_{ij})$ is a unit. □

REMARK 2.24. Take $A = \mathbb{Z}$ in (2.23). Elements $\gamma_1, \gamma_2, \dots, \gamma_m$ generate a submodule N of finite index in B if and only if $D(\gamma_1, \dots, \gamma_m) \neq 0$, in which case

$$D(\gamma_1, \dots, \gamma_m) = (B\colon N)^2 \cdot \mathrm{disc}(B/\mathbb{Z}).$$

To prove this, choose a basis $\beta_1, \dots, \beta_m$ for B as a $\mathbb{Z}$-module, and write $\gamma_j = \sum a_{ji}\beta_i$. Then both sides equal $\det(a_{ij})^2 \cdot D(\beta_1, \dots, \beta_m)$.

PROPOSITION 2.25. *Let* L *be a finite separable extension of the field* K *of degree* m, *and let* $\sigma_1, ..., \sigma_m$ *be the distinct* K*-homomorphisms of* L *into some large Galois extension* Ω *of* L. *Then, for any basis* $\beta_1, ..., \beta_m$ *of* L *over* K,

$$D(\beta_1, ..., \beta_m) = \det(\sigma_i \beta_j)^2 \neq 0.$$

PROOF. By direct calculation, we have

$$\begin{aligned} D(\beta_1,\dots,\beta_m) &\stackrel{\text{df}}{=} \det(\mathrm{Tr}(\beta_i\beta_j)) \\ &= \det(\textstyle\sum_k \sigma_k(\beta_i\beta_j)) \qquad \text{(by 2.19)} \\ &= \det(\textstyle\sum_k \sigma_k(\beta_i)\cdot\sigma_k(\beta_j)) \\ &= \det(\sigma_k(\beta_i))\cdot\det(\sigma_k(\beta_j)) \\ &= \det(\sigma_k(\beta_i))^2. \end{aligned}$$

Suppose that $\det(\sigma_i\beta_j)=0$. Then there exist $c_1,\dots,c_m\in\Omega$ such that

$$\sum_i c_i\sigma_i(\beta_j)=0 \text{ all } j.$$

By linearity, it follows that $\sum_i c_i\sigma_i(\beta)=0$ for all $\beta\in L$, but this contradicts the following result. (Apply it with $G=L^\times$.) □

LEMMA 2.26 (Dedekind's Lemma). *Let G be a group and Ω a field, and let $\sigma_1,\dots,\sigma_m$ be distinct homomorphisms $G\to\Omega^\times$; then $\sigma_1,\dots,\sigma_m$ are linearly independent over Ω, i.e., there do not exist $c_i\in\Omega$ such that $x\mapsto\sum_i c_i\sigma_i(x)\colon G\to\Omega$ is the zero map.*

PROOF. See FT, Theorem 5.13 (the proof is easy and elementary). □

COROLLARY 2.27. *Let K be the field of fractions of A, and let L be a finite separable extension of K of degree m. If the integral closure B of A in L is free of rank m over A, then* $\mathrm{disc}(B/A)\neq 0$.

PROOF. If $\{\beta_1,\dots,\beta_m\}$ is a basis for B as an A-module, then it follows easily from (2.4) that it is also a basis for L as a K-vector space. Hence $\mathrm{disc}(B/A)$ represents $\mathrm{disc}(L/K)$. □

REMARK 2.28. (a) The proposition shows that the K-bilinear pairing

$$(\beta,\beta')\mapsto\mathrm{Tr}(\beta\cdot\beta'):L\times L\to K$$

is nondegenerate (its discriminant is $\mathrm{disc}(L/K)$).

(b) The assumption that L/K is separable is essential; in fact, if L/K is not separable, then $\mathrm{disc}(L/K)=0$ (see exercises).

Rings of integers are finitely generated. We now show that $\mathcal{O}_K$ is finitely generated as a $\mathbb{Z}$-module.

PROPOSITION 2.29. *Let A be an integrally closed integral domain with field of fractions K, and let B the integral closure of A in a separable extension L of K of degree m. Then B is contained[7] in a free A-module of rank m. If A is a principal ideal domain, then B is itself a free A-module of rank m.*

PROOF. Let $\{\beta_1,\dots,\beta_m\}$ be a basis for L over K. According to (2.4), there is a $d\in A$ such that $d\cdot\beta_i\in B$ for all i. Clearly $\{d\cdot\beta_1,\dots,d\cdot\beta_m\}$ is still a basis for L as a vector space over K, and so we can assume that each $\beta_i\in B$. Because the trace

[7]This implies that B is finitely generated as an A-module — see 3.31 below.

pairing is nondegenerate, there is a "dual" basis $\{\beta_1', ..., \beta_m'\}$ of L over K such that $\mathrm{Tr}(\beta_i \cdot \beta_j') = \delta_{ij}$ (see the discussion following (2.21)). We shall show that

$$A\beta_1 + A\beta_2 + \cdots + A\beta_m \subset B \subset A\beta_1' + A\beta_2' + \cdots + A\beta_m'. \qquad (2.29.1)$$

Only the second inclusion requires proof. Let $\beta \in B$. Then β can be written uniquely as a linear combination $\beta = \sum b_j\beta_j'$ of the β_j' with coefficients $b_j \in K$, and we have to show that each $b_j \in A$. As β_i and β are in B, so also is $\beta \cdot \beta_i$, and so $\mathrm{Tr}(\beta \cdot \beta_i) \in A$ (see 2.20). But

$$\mathrm{Tr}(\beta \cdot \beta_i) = \mathrm{Tr}(\sum_j b_j\beta_j' \cdot \beta_i) = \sum_j b_j \,\mathrm{Tr}(\beta_j' \cdot \beta_i) = \sum_j b_j \cdot \delta_{ij} = b_i.$$

Hence $b_i \in A$.

If A is a principal ideal domain, then B is free of rank $\leq m$ as an A-module because it is contained in a free A-module of rank m (see any basic graduate algebra course), and it has rank $\geq m$ because it contains a free A-module of rank m. □

Corollary 2.30. *The ring of integers in a number field L is the largest subring that is finitely generated as a $\mathbb{Z}$-module.*

Proof. We have just seen that $\mathcal{O}_L$ is a finitely generated $\mathbb{Z}$-module. Let B be another subring of L that is finitely generated as a $\mathbb{Z}$-module; then every element of B is integral over $\mathbb{Z}$ (by 2.11), and so $B \subset \mathcal{O}_L$. □

Remark 2.31. (a) The hypothesis that L/K be separable is necessary to conclude that B is a finitely generated A-module (we used that the trace pairing was nondegenerate). However it is still true that the integral closure of $k[X]$ in any finite extension of $k(X)$ (not necessarily separable) is a finitely generated $k[X]$-module.

(b) The hypothesis that A be a principal ideal domain is necessary to conclude from (2.29.1) that B is a free A-module — there do exist examples of number fields L/K such that $\mathcal{O}_L$ is not a free $\mathcal{O}_K$-module.

(c) Here is an example of a finitely generated module that is not free. Let $A = \mathbb{Z}[\sqrt{-5}]$, and consider the A-modules

$$(2) \subset (2, 1 + \sqrt{-5}) \subset \mathbb{Z}[\sqrt{-5}].$$

Both (2) and $\mathbb{Z}[\sqrt{-5}]$ are free $\mathbb{Z}[\sqrt{-5}]$-modules of rank 1, but $(2, 1+\sqrt{-5})$ is *not* a free $\mathbb{Z}[\sqrt{-5}]$-module of rank 1, because it is not a principal ideal (see the Introduction). In fact, it is not a free module of any rank.

When K is a number field, a basis $\alpha_1, ..., \alpha_m$ for $\mathcal{O}_K$ as a $\mathbb{Z}$-module is called an *integral basis* for K.

Remark 2.32. We retain the notations of the proposition and its proof.

(a) Let $C = \sum A\beta_i \subset B$, with β_i a basis for L over K. Define

$$C^* = \{\beta \in L \mid \mathrm{Tr}(\beta\gamma) \in A \text{ for all } \gamma \in C\}.$$

By linearity,

$$\beta \in C^* \iff \mathrm{Tr}(\beta\beta_i) \in A \text{ for } i = 1, ..., m,$$

and it follows that

$$C^* = \sum A\beta_i'.$$

Thus we have:

$$C = \sum A\beta_i \subset B \subset \sum A\beta_i' = C^*.$$

(b) Write $L = \mathbb{Q}[\beta]$ with $\beta \in B$, and let $f(X)$ be the minimum polynomial of β. Let $C = \mathbb{Z}[\beta] = \mathbb{Z}1 + \mathbb{Z}\beta + \cdots + \mathbb{Z}\beta^{m-1}$. We want to find C^*.

One can show (see Fröhlich and Taylor 1991, p. 128) that

$$\mathrm{Tr}(\beta^i/f'(\beta)) = 0 \text{ if } 0 \le i \le m-2, \text{ and } \mathrm{Tr}(\beta^{m-1}/f'(\beta)) = 1$$

(these formulas go back to Euler). It follows from this that

$$\det(\mathrm{Tr}(\beta^i \cdot \beta^j/f'(\beta)) = (-1)^m$$

(the only term contributing to the determinant is the product of the elements on the *other* diagonal). If $\beta_1', ..., \beta_m'$ is the dual basis to $1, \beta, \dots, \beta^{m-1}$, so that $\mathrm{Tr}(\beta^i \cdot \beta_j') = \delta_{ij}$, then

$$\det(\mathrm{Tr}(\beta^i \cdot \beta_j')) = 1.$$

On comparing these formulas, one sees that the matrix relating the family $\{1/f'(\beta), ..., \beta^{m-1}/f'(\beta)\}$ to the basis $\beta_1', ..., \beta_m'$ has determinant ± 1, and so it is invertible in $M_n(A)$. Thus we see that C^* is a free A-module with basis $\{1/f'(\beta), \dots, \beta^{m-1}/f'(\beta)\}$:

$$C = A[\beta] \subset B \subset f'(\beta)^{-1}A[\beta] = C^*.$$

Finding the ring of integers. We now assume K to be a field of characteristic zero.

PROPOSITION 2.33. *Let $L = K[\beta]$ some β, and let $f(X)$ be the minimum polynomial of β over K. Suppose that $f(X)$ factors into $\prod(X - \beta_i)$ over the Galois closure of L. Then*

$$D(1, \beta, \beta^2, \dots, \beta^{m-1}) = \prod_{1 \le i < j \le m} (\beta_i - \beta_j)^2 = (-1)^{m(m-1)/2} \cdot \mathrm{Nm}_{L/K}(f'(\beta)).$$

PROOF. We have

$$\begin{aligned} D(1, \beta, \beta^2, \dots, \beta^{m-1}) &= \det(\sigma_i(\beta^j))^2 && (2.25)\\ &= \det(\beta_i^j)^2 \\ &= (\textstyle\prod_{i<j}(\beta_i - \beta_j))^2 && (\text{Vandermonde})\\ &= (-1)^{m(m-1)/2} \cdot \textstyle\prod_i(\prod_{j\ne i}(\beta_i - \beta_j)) \\ &= (-1)^{m(m-1)/2} \cdot \textstyle\prod_j f'(\beta_j) \\ &= (-1)^{m(m-1)/2}\,\mathrm{Nm}(f'(\beta)). \end{aligned}$$

□

The number in (2.33) is called the *discriminant* of $f(X)$. It can also be defined as the resultant of $f(X)$ and $f'(X)$. The discriminant of f lies in K, and it is zero if and only if f has a repeated root. It is a symmetric polynomial in the β_i with coefficients

in A, and so (by 2.1) it can be expressed in terms of the coefficients of $f(X)$, but the formulas are quite complicated.

EXAMPLE 2.34. We compute the discriminant of

$$f(X) = X^n + aX + b, \quad a, b \in K,$$

assumed to be irreducible and separable. Let β be a root of $f(X)$, and let $\gamma = f'(\beta) = n\beta^{n-1} + a$. We compute $\mathrm{Nm}(\gamma)$. On multiplying the equation

$$\beta^n + a\beta + b = 0$$

by $n\beta^{-1}$ and rearranging, we obtain the equation

$$n\beta^{n-1} = -na - nb\beta^{-1}.$$

Hence

$$\gamma = n\beta^{n-1} + a = -(n-1)a - nb\beta^{-1}.$$

Solving for β gives

$$\beta = \frac{-nb}{\gamma + (n-1)a},$$

from which it is clear that $K[\beta] = K[\gamma]$, and so the minimum polynomial of γ over K has degree n also. If we write

$$f(\frac{-nb}{X + (n-1)a}) = P(X)/Q(X),$$

then $P(\gamma)/Q(\gamma) = f(\beta) = 0$ and so $P(\gamma) = 0$. Since

$$P(X) = (X + (n-1)a)^n - na(X + (n-1)a)^{n-1} + (-1)^n n^n b^{n-1}$$

is monic of degree n, it must be the minimum polynomial of γ. Therefore $\mathrm{Nm}(\gamma)$ is $(-1)^n$ times the constant term of this polynomial, and so we find that

$$\mathrm{Nm}(\gamma) = n^n b^{n-1} + (-1)^{n-1}(n-1)^{n-1}a^n.$$

Finally we obtain the formula:

$$\mathrm{disc}(X^n + aX + b) = (-1)^{n(n-1)/2}(n^n b^{n-1} + (-1)^{n-1}(n-1)^{n-1}a^n)$$

For example:

$\mathrm{disc}(X^2 + aX + b) = -4b + a^2$,
$\mathrm{disc}(X^3 + aX + b) = -27b^2 - 4a^3$,
$\mathrm{disc}(X^4 + aX + b) = 256b^3 - 27a^4$,
$\mathrm{disc}(X^5 + aX + b) = 5^5 b^4 + 4^4 a^5$.

For any polynomials more complicated than the above, use Maple (or Mathematica). For example, after starting Maple, type:

```
discrim(X^3 + a*X^2 + b*X + c, X);
```

Don't forget the semicolon at the end! The program displays:

$-27c^2 + 18cab + a^2b^2 - 4a^3c - 4b^3$.

Since it is awkward to write a polynomial in the notation Maple understands, in future I'll use normal notation and leave you to insert asterisks and hats. To compute discriminants with Mathematica, you compute the resultant of $f(X)$ and $f'(X)$.

The general strategy for finding the ring of integers of K is to write $K = \mathbb{Q}[\alpha]$ with α an integer in K, and compute $D(1, \alpha, ..., \alpha^{m-1})$. It is an integer, and if it is square-free, then $\{1, \alpha, ..., \alpha^{m-1}\}$ is automatically an integral basis, because (see 2.24)

$$D(1, \alpha, \dots, \alpha^{m-1}) = \text{disc}(\mathcal{O}_K/\mathbb{Z}) \cdot (\mathcal{O}_K : \mathbb{Z}[\alpha])^2.$$

If it is not square-free, $\{1, \alpha, ..., \alpha^{m-1}\}$ may still be a basis, and sometimes one can tell this by using Stickelberger's theorem (see 2.39 below) or by looking at how primes ramify (see later). If $\{1, \alpha, ..., \alpha^{m-1}\}$ is not an integral basis, one has to look for algebraic integers not in $\sum \mathbb{Z} \cdot \alpha^i$ (we describe an algorithm below).

EXAMPLE 2.35. Let α be a root of the polynomial $X^3 - X - 1$. Check that $X^3 - X - 1$ is irreducible[8] in $\mathbb{Q}[X]$ (if it factored, it would have a root in $\mathbb{Q}$, which would be an integer dividing 1). We have

$$D(1, \alpha, \alpha^2) = \text{disc}(f(X)) = -23,$$

which contains no square factor, and so $\mathbb{Z}[\alpha]$ is the ring of integers in $\mathbb{Q}[\alpha]$ — $\{1, \alpha, \alpha^2\}$ is an integral basis for $\mathbb{Q}[\alpha]$.

EXAMPLE 2.36. Let α be a root of the polynomial X^3+X+1. Then $D(1, \alpha, \alpha^2) = \text{disc}(f(X)) = -31$, which contains no square factor, and so again $\{1, \alpha, \alpha^2\}$ is an integral basis for $\mathbb{Q}[\alpha]$.

EXAMPLE 2.37. This example goes back to Dedekind. Let $K = \mathbb{Q}[\alpha]$, where α is a root of

$$f(X) = X^3 + X^2 - 2X + 8.$$

Maple computes $\text{disc}(f(X)) = -4 \cdot 503$, but Dedekind showed that $\mathcal{O}_K \neq \mathbb{Z}[\beta]$, and so $\text{disc}(\mathcal{O}/\mathbb{Z}) = -503$. In fact Dedekind showed that there is no integral basis of the form 1, β, β^2 (Weiss 1963, p. 170; for another example of this type, see Problems 2, no. 2.)

EXAMPLE 2.38. Consider the field $\mathbb{Q}[\alpha]$ where α is a root of $f(X) = X^5 - X - 1$. This polynomial is irreducible, because it is irreducible in $\mathbb{F}_3[X]$. The discriminant of $f(X)$ is $2869 = 19 \cdot 151$, and so the ring of integers is $\mathbb{Z}[\alpha]$.

PROPOSITION 2.39. *Let K be an algebraic number field.*

(a) *The sign of* $\text{disc}(K/\mathbb{Q})$ *is* $(-1)^s$, *where* $2s$ *is the number of homomorphisms* $K \hookrightarrow \mathbb{C}$ *whose image is not contained in* $\mathbb{R}$.
(b) *(Stickelberger's theorem)* $\text{disc}(\mathcal{O}_K/\mathbb{Z}) \equiv 0$ *or* $1 \mod 4$.

PROOF. (a) Let $K = \mathbb{Q}[\alpha]$, and let $\alpha_1 = \alpha$, α_2, ..., α_r be the real conjugates of α and α_{r+1}, $\bar{\alpha}_{r+1}$..., $\alpha_{r+s}, \bar{\alpha}_{r+s}$ the complex conjugates. One sees easily that

$$\text{sign}(\text{disc}(1, ..., \alpha^{m-1})) = \text{sign}(\prod_{1 \le i \le s} (\alpha_{r+i-s} - \bar{\alpha}_{r+i-s}))^2$$

(the other terms are either squares of real numbers or occur in conjugate pairs), and this equals $(-1)^s$.

(b) Recall that $\text{disc}(\mathcal{O}_K/\mathbb{Z}) = \det(\sigma_i \alpha_j)^2$, where $\alpha_1, ..., \alpha_m$ is an integral basis. Let P be the sum of the terms in the expansion of $\det(\sigma_i \alpha_j)$ corresponding to even

[8] In fact, this is the monic irreducible cubic polynomial in $\mathbb{Z}[X]$ with the smallest discriminant.

permutations, and $-N$ the sum of the terms corresponding to odd permutations. Then

$$\operatorname{disc}(\mathcal{O}_K/\mathbb{Z}) = (P-N)^2 = (P+N)^2 - 4PN.$$

If τ is an element of the Galois group of the Galois closure of K over $\mathbb{Q}$, then either $\tau P = P$ and $\tau N = N$, or $\tau P = N$ and $\tau N = P$. In either case, τ fixes $P+N$ and PN, and so they are rational numbers. As they are integral over $\mathbb{Z}$, they must in fact be integers, from which it follows that

$$\operatorname{disc}(\mathcal{O}_K/\mathbb{Z}) \equiv (P+N)^2 \equiv 0 \text{ or } 1 \mod 4.$$

□

EXAMPLE 2.40. Consider the field $\mathbb{Q}[\sqrt{m}]$, where m is a square-free integer.

Case $m \equiv 2, 3 \bmod 4$. Here $D(1, \sqrt{m}) = \operatorname{disc}(X^2 - m) = 4m$, and so Stickelberger's theorem shows that $\operatorname{disc}(\mathcal{O}_K/\mathbb{Z}) = 4m$, and hence $\{1, \sqrt{m}\}$ is an integral basis.

Case $m \equiv 1 \bmod 4$. First verify that $(1+\sqrt{m})/2$ is integral. Then $D(1, (1+\sqrt{m})/2) = m$, and so $\{1, (1+\sqrt{m})/2\}$ is an integral basis.

REMARK 2.41. Let K and K' be number fields. If K and K' are isomorphic, then $[K:\mathbb{Q}] = [K':\mathbb{Q}]$ and $\operatorname{disc}(\mathcal{O}_K/\mathbb{Z}) = \operatorname{disc}(\mathcal{O}_{K'}/\mathbb{Z})$, but the converse is not true. For example, there are four nonisomorphic cubic number fields with discriminant -4027 (4027 is prime) (see later for two of them).

The curious may wonder why we didn't give an example of a field generated by an integral element whose minimum polynomial has discriminant ± 1. The reason is that there is no such polynomial of degree > 1 — see the discussion following Theorem 4.8 below.

Algorithms for finding the ring of integers. By an *algorithm* I mean a procedure that could (in principle) be put on a computer and is guaranteed to lead to the answer in a finite number of steps. Suppose the input requires N digits to express it. A *good algorithm* is one whose running time is $< N^c$ for some c. For example, there is no known good algorithm for factoring an integer. By a *practical algorithm* I mean one that has been (or should have been) put on a computer, and is actually useful.

The following variant of (2.29) is useful. Let A be a principal ideal domain with field of fractions K, and let B be the integral closure of A in a finite separable extension L of K of degree m.

PROPOSITION 2.42. *Let $\beta_1, ..., \beta_m$ be a basis for L over K consisting of elements of B, and let $d = \operatorname{disc}(\beta_1, ..., \beta_m)$. Then*

$$A\cdot\beta_1 + ... + A\cdot\beta_m \subset B \subset A\cdot(\beta_1/d) + ... + A\cdot(\beta_m/d).$$

PROOF. Let $\beta \in B$, and write

$$\beta = x_1\beta_1 + \cdots + x_m\beta_m, \quad x_i \in K.$$

Let $\sigma_1, \dots, \sigma_m$ be the distinct K-embeddings of L into some large Galois extension Ω of K. On applying the σ's to this equation, we obtain a system of linear equations:

$$\sigma_i\beta = x_1\sigma_i\beta_1 + x_2\sigma_i\beta_2 + \cdots + x_m\sigma_i\beta_m, \quad i = 1, \dots, m.$$

Hence by Cramer's rule

$$x_i = \gamma_i/\delta$$

where $\delta = \det(\sigma_i\beta_j)$ and γ_i is the determinant of the same matrix, but with the ith column replaced with $(\sigma_i\beta)$. From (2.33), we know that $\delta^2 = d$. Thus $x_i = \gamma_i\delta/d$, and $\gamma_i\delta$ is an element of K (because it equals dx_i) and is integral over A. Therefore $\gamma_i\delta \in A$, which completes the proof. □

Thus there is the following algorithm for finding the ring of integers in a number field K. Write $K = \mathbb{Q}[\alpha]$ where α is integral over $\mathbb{Q}$. Compute $d = \operatorname{disc}(1, \alpha, ..., \alpha^{m-1})$. Then

$$\mathbb{Z}[\alpha] \subset \mathcal{O}_K \subset d^{-1}\mathbb{Z}[\alpha].$$

Note that $(d^{-1}\mathbb{Z}[\alpha] \colon \mathbb{Z}[\alpha]) = d^m$, which is huge but finite. Each coset $\beta + \mathbb{Z}[\alpha]$, $\beta \in d^{-1}\mathbb{Z}[\alpha]$, consists entirely of algebraic integers or contains no algebraic integer. Find a set of representatives $\beta_1, ..., \beta_n$ for $\mathbb{Z}[\alpha]$ in $d^{-1}\mathbb{Z}[\alpha]$, and test each to see whether it is integral over $\mathbb{Z}$ (the coefficients of its minimum polynomial will have denominators bounded by a power of d, and so it is possible to tell whether or not they are integers by computing them with sufficient accuracy).

Unfortunately this method is not practical. For example, the polynomial

$$f(X) = X^5 + 17X^4 + 3X^3 + 2X^2 + X + 1$$

is irreducible [9] and has discriminant 285401001. Hence, if α is a root of $f(X)$ and $K = \mathbb{Q}[\alpha]$, then the index of $\mathbb{Z}[\alpha]$ in $\mathbb{Z}\frac{1}{d} + \mathbb{Z}\frac{\alpha}{d} + \cdots + \mathbb{Z}\frac{\alpha^4}{d}$ is $(285401001)^5$. [Actually, as luck would have it,

$$285401001 = 3 \cdot 179 \cdot 233 \cdot 2281$$

is square-free, and so $\mathcal{O}_K = \mathbb{Z}[\alpha]$.]

Note Maple can compute minimal polynomials over $\mathbb{Q}$. For example,

readlib(lattice);

minpoly(1.41421356,3);

gives the output

$$X^2 - 2.$$

The first line loads the appropriate library, and the second finds the polynomial (not necessarily monic) with small integer coefficients of degree ≤ 3 that comes closest to having $1.414\ldots$ as a root.

[9] In Maple, type "factor(f(X));" to factor a polynomial over $\mathbb{Q}$. It is obvious that there is an algorithm for factoring a polynomial $f(X)$ in $\mathbb{Q}[X]$. First divide through by the leading coefficient to make $f(X)$ monic. Then proceed as in the proof of (2.4) to obtain a monic polynomial with coefficients in $\mathbb{Z}$. There is a bound on the absolute value of any root of α of the polynomial in terms of the degree and the coefficients (if $|\alpha|$ is too big, then the remaining terms can't cancel the leading term α^m). Therefore there is a bound on the absolute values of the coefficients of the factors of the polynomial, and since these coefficients are integers, it is possible to simply search for them. Alternatively, note that two polynomials in $\mathbb{Z}[X]$ can be distinguished by looking modulo a sufficiently large prime. Hence factoring polynomials in $\mathbb{Q}[X]$ is something that can be safely left to the computer.

I now discuss a practical algorithm for finding $\mathcal{O}_K$ for small degrees and small discriminants from Pohst and Zassenhaus 1989 (see the additional references at the end of this section). The next result will help us get an idea of what should be possible.

LEMMA 2.43. *Let (A, δ) be Euclidean domain, and let M be an $m \times m$ matrix with coefficients in A. Then it is possible to put M into upper triangular form by elementary row operations of the following type:*

(i) add a multiple of one row to a second;

(ii) swap two rows.

PROOF. By definition $\delta : A \to \mathbb{Z}$ is a function with the following property: for any two elements $a, b \in A$, $a \neq 0$, there exist elements q and r such that

$$b = qa + r, \text{ with } r = 0 \text{ or } \delta(r) < \delta(a).$$

Apply an operation of type (ii) so that the element of the first column with the minimum δ is in the $(1,1)$-position. If a_{11} divides all elements in the first column, we can use operations of type (i) to make all the remaining elements of the first column zero. If not, we can use (i) to get an element in the first column that has smaller δ-value than a_{11}, and put that in the $(1,1)$ position. Repeat — eventually, we will have the gcd of the original elements in the first column in the $(1,1)$ position and zeros elsewhere. Then move onto the next column... □

REMARK 2.44. (a) The operations (i) and (ii) are invertible in matrices with coefficients in A, and they correspond to multiplying on the left with an invertible matrix in $M_n(A)$. Hence we have shown that there exists an invertible matrix U in $M_n(A)$ such that UM is upper triangular.

On taking transposes, we find that for any matrix $M \in M_n(A)$, there is an invertible matrix U in $M_n(A)$ such that MU is lower triangular.

(b) Take $A = \mathbb{Z}$ (for simplicity), and add the (invertible) operation:

(iii) multiply a row by -1.

Then it is possible to make the triangular matrix $T = UM$ satisfy the following conditions (assuming $\det(M) \neq 0$):

$a_{ii} > 0$ for all i;
the elements a_{ij} of the j^{th} column satisfy $0 \leq a_{ij} < a_{jj}$.

Then T is unique. It is said to be in *Hermite normal form.*

Consider the field $K = \mathbb{Q}[\alpha]$ generated over $\mathbb{Q}$ by the algebraic integer α with minimum polynomial $f(X)$. Let $\{\omega_1, ..., \omega_n\}$ be a basis for $\mathcal{O}_K$ as a $\mathbb{Z}$-module, and write

$$A = M \cdot \Omega$$

where $A = (1, \alpha, ..., \alpha^{n-1})^{\text{tr}}$ and $\Omega = (\omega_1, ..., \omega_n)^{\text{tr}}$. Choose U such that MU is lower triangular (and in Hermite normal form), and write

$$A = MU \cdot U^{-1}\Omega = T \cdot \Omega'.$$

Here $\Omega' =_{df} U^{-1}\Omega$ is again a $\mathbb{Z}$-basis for $\mathcal{O}_K$, and $\Omega' = T^{-1} \cdot A$ with T^{-1} also lower triangular (but not necessarily with integer coefficients). Thus

$\omega_1' = a_{11}1$;
$\omega_2' = a_{21}1 + a_{22}\alpha$;
etc.,
where $d \cdot a_{ij} \in \mathbb{Z}$, $d = |\det(M)| = |\det(T)|$.

EXAMPLE 2.45. Let $K = \mathbb{Q}[\sqrt{m}]$, m square-free, $m \equiv 1 \pmod 4$. The integral basis

$$1, \frac{1+\sqrt{m}}{2}$$

is of the above form.

In (Pohst and Zassenhaus 1989, 4.6), there is an algorithm that, starting from a monic irreducible polynomial

$$f(X) = X^n + a_1X^{n-1} + \cdots + a_n, \quad a_n \in \mathbb{Z},$$

constructs an integral basis $\omega_1, ..., \omega_n$, such that

$$\omega_i = (\sum_{k=1}^{i} a_{ik}\alpha^i)/N_i$$

where

$$\alpha \text{ is a root of } f(X), \quad a_{ik} \in \mathbb{Z}, \quad N_i \in \mathbb{Z}, \quad \gcd(a_{i1}, ..., a_{ii}) = 1.$$

In an Appendix, they use it to show that $\mathbb{Q}[\alpha]$, where α is a root of

$$f(X) = X^{11} + 101X^{10} + 4151X^9 + \cdots - 332150625,$$

has an integral basis

$\omega_1 = 1$,
$\omega_2 = (1/2)\alpha + 1/2$
$\omega_3 = (1/4)\alpha^2 - 1/4$
$\omega_4 = (1/8)\alpha^3 + (1/8)\alpha^2 - (1/8)\alpha - 1/8$
.....
$\omega_{11} = (1/9103145472000)\alpha^{10} + \cdots - 4064571/49948672$.

The discriminant of f is $2^{130} \times 3^{12} \times 5^{12} \times 29^{18} \times 82231^6$, and the index of $\mathbb{Z}[\alpha]$ in $\mathcal{O}_K$ is $2^{56} \times 3^6 \times 5^3 \times 29^9$.

The first step is to compute $D(1, \alpha, \alpha^2, \dots) = \text{disc}(f(X))$ and to find its square factors. Finding the square factors of $\text{disc}(f(X))$ is the most time-consuming part of the algorithm. The time taken to factor an N-digit number is exponential in the number of digits of N. Every computer can factor a 25 digit number easily, but after that it becomes rapidly more difficult. Hundred digit numbers are extremely difficult. Thus this is not a good algorithm in the above sense. Once one has found the square factors of $\text{disc}(f(X))$ the algorithm for computing an integral basis of the above form is good.

3. Dedekind Domains; Factorization

Presently, we shall define the notion of a Dedekind domain; then we'll prove:

(i) ideals in Dedekind domains factor uniquely into products of prime ideals;

(ii) rings of integers in number fields are Dedekind domains.

First we consider a local version of a Dedekind domain.

Discrete valuation rings. The following conditions on a principal ideal domain are equivalent:

(a) A has exactly one nonzero prime ideal;
(b) up to associates, A has exactly one prime element;
(c) A is local and is not a field.

A ring satisfying these conditions is called a *discrete valuation ring*.

Example 3.1. The ring $\mathbb{Z}_{(p)} =_{df} \{\frac{m}{n} \in \mathbb{Q} \mid n \text{ not divisible by } p\}$ is a discrete valuation ring with prime elements $\pm p$ and prime ideal (p).

Later we shall define discrete valuations, and so justify the name.

If A is a discrete valuation ring and π is a prime element in A, then each nonzero ideal in A is of the form (π^m) for a unique $m \in \mathbb{N}$. Thus, if $\mathfrak{a}$ is an ideal in A and $\mathfrak{p}$ denotes the (unique) maximal ideal of A, then $\mathfrak{a} = \mathfrak{p}^m$ for a well-defined integer $m \geq 0$.

Recall that, for an A-module M and an $m \in M$, the *annihilator* of m

$$\mathrm{Ann}(m) = \{a \in A \mid am = 0\}.$$

It is an ideal in A, and it is a proper ideal if $m \neq 0$. Suppose A is a discrete valuation ring, and let c be a nonzero element of A. Let $M = A/(c)$. What is the annihilator of a nonzero $b + (c)$ of M. Fix a prime element π of A, and let $c = u\pi^m$, $b = v\pi^n$ with u and v units. Then $n < m$ (else $b + (c) = 0$ in M), and

$$\mathrm{Ann}(b + (c)) = (\pi^{m-n}).$$

Thus, a b for which $\mathrm{Ann}(b+(c))$ is maximal, is of the form $v\pi^{m-1}$, and for this choice $\mathrm{Ann}(b + (c))$ is a prime ideal generated by $\frac{c}{b}$. We shall exploit these observations in the proof of the next proposition, which gives a criterion for a ring to be a discrete valuation ring.

Proposition 3.2. *An integral domain A is a discrete valuation ring if and only if*

(i) A is Noetherian,

(ii) A is integrally closed, and

(iii) A has exactly one nonzero prime ideal.

Proof. The necessity of the three conditions is obvious, so let A be an integral domain satisfying (i), (ii), and (iii). We have to show that every ideal in A is principal. As a first step, we prove that the nonzero prime ideal in principal. Note that the conditions imply that A is a local ring.

Choose an element $c \in A$, $c \neq 0$, $c \neq$ unit, and consider the A-module $M =_{df} A/(c)$. For any nonzero $m \in M$, the annihilator of m,

$$\mathrm{Ann}(m) = \{a \in A \mid am = 0\}$$

is a proper ideal in A. Because A is Noetherian (here we use (i)), we can choose an m such that $\mathrm{Ann}(m)$ is maximal among these ideals. Write $m = b + (c)$ and $\mathfrak{p} = \mathrm{Ann}(b + (c))$. Note that $c \in \mathfrak{p}$, and so $\mathfrak{p} \neq 0$, and that

$$\mathfrak{p} = \{a \in A \mid c|ab\}.$$

I claim that $\mathfrak{p}$ is prime. If not there exist elements x, $y \in A$ such that $xy \in \mathfrak{p}$ but neither x nor $y \in \mathfrak{p}$. Then $yb+(c)$ is a nonzero element of M because $y \notin \mathfrak{p}$. Consider $\mathrm{Ann}(yb + (c))$. Obviously it contains $\mathfrak{p}$ and it contains x, but this contradicts the maximality of $\mathfrak{p}$ among ideals of the form $\mathrm{Ann}(m)$. Hence $\mathfrak{p}$ is prime.

I claim $\frac{b}{c} \notin A$. Otherwise $b = c \cdot \frac{b}{c} \in (c)$, and $m = 0$ (in M).

I claim that $\frac{c}{b} \in A$, and $\mathfrak{p} = (\frac{c}{b})$. By definition, $\mathfrak{p}b \subset (c)$, and so $\mathfrak{p} \cdot \frac{b}{c} \subset A$, and it is an ideal in A. If $\mathfrak{p} \cdot \frac{b}{c} \subset \mathfrak{p}$, then $\frac{b}{c}$ is integral over A (by 2.11, since $\mathfrak{p}$ is finitely generated), and so $\frac{b}{c} \in A$ (because of condition (ii)), but we know $\frac{b}{c} \notin A$. . Thus $\mathfrak{p} \cdot \frac{b}{c} = A$ (by (iii)), and this implies that $\mathfrak{p} = (\frac{c}{b})$.

Let $\pi = \frac{c}{b}$, so that $\mathfrak{p} = (\pi)$. Let $\mathfrak{a}$ be a proper ideal of A, and consider the sequence

$$\mathfrak{a} \subset \mathfrak{a}\pi^{-1} \subset \mathfrak{a}\pi^{-2} \subset \dots .$$

If $\mathfrak{a}\pi^{-r} = \mathfrak{a}\pi^{-r-1}$ for some r, then $\pi^{-1}(\mathfrak{a}\pi^{-r}) = \mathfrak{a}\pi^{-r}$, and π^{-1} is integral over A, and so lies in A — this is impossible (π is not a unit in A). Therefore the sequence is strictly increasing, and (again because A is Noetherian) it can't be contained in A. Let m be the smallest integer such that $\mathfrak{a}\pi^{-m} \subset A$ but $\mathfrak{a}\pi^{-m-1} \not\subseteq A$. Then $\mathfrak{a}\pi^{-m} \not\subseteq \mathfrak{p}$, and so $\mathfrak{a}\pi^{-m} = A$. Hence $\mathfrak{a} = (\pi^m)$. □

Dedekind domains. A *Dedekind domain* is an integral domain $A \neq$ field such that

(i) A is Noetherian;

(ii) A is integrally closed;

(iii) every nonzero prime ideal is maximal.

Thus Proposition 3.2 says that a local integral domain is a Dedekind domain if and only if it is a discrete valuation ring.

PROPOSITION 3.3. *Let A be a Dedekind domain, and let S be a multiplicative subset of A. Then $S^{-1}A$ is either a Dedekind domain or a field.*

PROOF. Condition (iii) says that there is no containment relation between nonzero prime ideals of A. If this condition holds for A, then (1.5) shows that it holds for $S^{-1}A$. Conditions (i) and (ii) follow from the next lemma. □

PROPOSITION 3.4. *Let A be an integral domain, and let S be a multiplicative subset of A.*

(a) *If A is Noetherian, then so also is $S^{-1}A$.*
(b) *If A is integrally closed, then so also is $S^{-1}A$.*

PROOF. (a) Let $\mathfrak{a}$ be an ideal in $S^{-1}A$. Then $\mathfrak{a} = S^{-1}(\mathfrak{a} \cap A)$ (the proof of this in 1.5 didn't use that $\mathfrak{a}$ is prime), and so $\mathfrak{a}$ is generated by any (finite) set of generators for $\mathfrak{a} \cap A$.

(b) Let α be an element of the field of fractions of A (= field of fractions of $S^{-1}A$) that is integral over $S^{-1}A$. Then

$$\alpha^m + a_1\alpha^{m-1} + \cdots + a_m = 0, \text{ some } a_i \in S^{-1}A.$$

For each i, there exists an $s_i \in S$ such that $s_i a_i \in A$. Set $s = s_1 \cdots s_m \in S$, and multiply through the equation by s^m :

$$(s\alpha)^m + sa_1(s\alpha)^{m-1} + \cdots + s^m a_m = 0.$$

This equation shows that $s\alpha$ is integral over A, and so lies in A. Hence $\alpha = (s\alpha)/s \in S^{-1}A$. □

COROLLARY 3.5. *For any nonzero prime ideal $\mathfrak{p}$ in a Dedekind domain, $A_\mathfrak{p}$ is a discrete valuation ring.*

PROOF. We saw in (1.6a) that $A_\mathfrak{p}$ is local, and the proposition implies that it is Dedekind. □

Unique factorization. The main result concerning Dedekind domains is the following.

THEOREM 3.6. *Let A be a Dedekind domain. Every proper nonzero ideal $\mathfrak{a}$ of A can be written in the form*

$$\mathfrak{a} = \mathfrak{p}_1^{r_1} \cdots \mathfrak{p}_n^{r_n}$$

with the $\mathfrak{p}_i$ distinct prime ideals and the $r_i > 0$; the $\mathfrak{p}_i$ and the r_i are uniquely determined.

The proof will require several lemmas.

LEMMA 3.7. *Let A be a Noetherian ring; then every ideal $\mathfrak{a}$ in A contains a product of nonzero prime ideals.*

PROOF. (Note the similarity to the proof of 1.2.) Suppose not, and choose a maximal counterexample $\mathfrak{a}$. Then $\mathfrak{a}$ itself can not be prime, and so there exist elements x and y of A such that $xy \in \mathfrak{a}$ but neither x nor $y \in \mathfrak{a}$. The ideals $\mathfrak{a} + (x)$ and $\mathfrak{a} + (y)$ strictly contain $\mathfrak{a}$, but their product is contained in $\mathfrak{a}$. Because $\mathfrak{a}$ is a maximal counterexample to the statement of the lemma, each of $\mathfrak{a} + (x)$ and $\mathfrak{a} + (y)$ contains a product of prime ideals, and it follows that $\mathfrak{a}$ contains a product of prime ideals. □

LEMMA 3.8. *Let A be a ring, and let $\mathfrak{a}$ and $\mathfrak{b}$ be relatively prime ideals in A; for any m, $n \in \mathbb{N}$, $\mathfrak{a}^m$ and $\mathfrak{b}^n$ are relatively prime.*

PROOF. If $\mathfrak{a}^m$ and $\mathfrak{b}^n$ are not relatively prime, then they are both contained in some prime (even maximal) ideal $\mathfrak{p}$. But if a prime ideal contains a power of an element, then it contains the element, and so $\mathfrak{p} \supset \mathfrak{a}^m \Rightarrow \mathfrak{p} \supset \mathfrak{a}$ and $\mathfrak{p} \supset \mathfrak{b}^n \Rightarrow \mathfrak{p} \supset \mathfrak{b}$. Thus $\mathfrak{a}$ and $\mathfrak{b}$ are both contained in $\mathfrak{p}$, and so they are not relatively prime.

Alternative proof: We are given that there exist elements $a \in A$ and $b \in B$ such that $a + b = 1$. Consider

$$1 = (a+b)^r = a^r + \begin{pmatrix} r \\ 1 \end{pmatrix} a^{r-1}b + \cdots + b^r.$$

If $r \geq m + n$, then the term on the right is the sum of an element of $\mathfrak{a}^m$ with an element of $\mathfrak{b}^n$. □

If $\mathfrak{p}$ and $\mathfrak{p}'$ are distinct prime ideals of a Dedekind domain, then condition (iii) of the definition implies that $\mathfrak{p}$ and $\mathfrak{p}'$ are relatively prime, and the lemma shows that $\mathfrak{p}^m$ and $\mathfrak{p}'^n$ are also relatively prime for all $m, n \geq 1$.

LEMMA 3.9. *Consider a product of rings $A \times B$. If $\mathfrak{a}$ and $\mathfrak{b}$ are ideals in A and B respectively, then $\mathfrak{a} \times \mathfrak{b}$ is an ideal in $A \times B$, and every ideal in $A \times B$ is of this form. The prime ideals of $A \times B$ are the ideals of the form*

$$\mathfrak{p} \times B \quad (\mathfrak{p} \text{ a prime ideal of } A), \quad A \times \mathfrak{p} \quad (\mathfrak{p} \text{ a prime ideal of } B).$$

PROOF. Let $\mathfrak{c}$ be an ideal in $A \times B$, and let

$$\mathfrak{a} = \{a \in A \mid (a,0) \in \mathfrak{c}\}, \quad \mathfrak{b} = \{b \in B \mid (0,b) \in \mathfrak{c}\}.$$

Clearly $\mathfrak{a} \times \mathfrak{b} \subset \mathfrak{c}$. Conversely, let $(a,b) \in \mathfrak{c}$. Then $(a,0) = (a,b) \cdot (1,0) \in \mathfrak{a}$ and $(0,b) = (a,b) \cdot (0,1) \in \mathfrak{b}$, and so $(a,b) \in \mathfrak{a} \times \mathfrak{b}$.

Recall that an ideal $\mathfrak{c} \subset C$ is prime if and only if $C/\mathfrak{c}$ is an integral domain. The map

$$A \times B \to A/\mathfrak{a} \times B/\mathfrak{b}, \quad (a,b) \mapsto (a + \mathfrak{a}, b + \mathfrak{b})$$

has kernel $\mathfrak{a} \times \mathfrak{b}$, and hence induces an isomorphism

$$A \times B/(\mathfrak{a} \times \mathfrak{b}) \approx A/\mathfrak{a} \times B/\mathfrak{b}.$$

The product of two nonzero rings always has nonzero zero-divisors, and so in order for $A \times B/(\mathfrak{a} \times \mathfrak{b})$ to be prime, we must have $\mathfrak{a} = A$ or $\mathfrak{b} = B$. Suppose the latter holds. Then $A \times B/(\mathfrak{a} \times \mathfrak{b}) \approx A/\mathfrak{a}$, and this is an integral domain if and only if $\mathfrak{a}$ is prime. □

REMARK 3.10. The lemma extends in an obvious way to a finite product of rings: the ideals in $A_1 \times \cdots \times A_m$ are of the form $\mathfrak{a}_1 \times \cdots \times \mathfrak{a}_m$ with $\mathfrak{a}_i$ an ideal in A_i; moreover, $\mathfrak{a}_1 \times \cdots \times \mathfrak{a}_m$ is prime if and only if there is a j such that $\mathfrak{a}_j$ is a prime ideal in A_j and $\mathfrak{a}_i = A_i$ for $i \neq j$.

LEMMA 3.11. *Let $\mathfrak{p}$ be a maximal ideal of a ring A, and let $\mathfrak{q}$ be the ideal it generates in $A_{\mathfrak{p}}$, $\mathfrak{q} = \mathfrak{p}A_{\mathfrak{p}}$. The map*

$$a + \mathfrak{p}^m \mapsto a + \mathfrak{q}^m \colon A/\mathfrak{p}^m \to A_{\mathfrak{p}}/\mathfrak{q}^m$$

is an isomorphism.

PROOF. We first show that the map is one-to-one. For this we have to show that $\mathfrak{q}^m \cap A = \mathfrak{p}^m$. But $\mathfrak{q}^m = S^{-1}\mathfrak{p}^m$, $S = A - \mathfrak{p}$, and so we have to show that $\mathfrak{p}^m = (S^{-1}\mathfrak{p}^m) \cap A$. An element of $(S^{-1}\mathfrak{p}^m) \cap A$ can be written $a = b/s$ with $b \in \mathfrak{p}^m$, $s \in S$, and $a \in A$. Then $sa \in \mathfrak{p}^m$, and so $sa = 0$ in $A/\mathfrak{p}^m$. The only maximal ideal containing $\mathfrak{p}^m$ is $\mathfrak{p}$ (because $\mathfrak{m} \supset \mathfrak{p}^m \Rightarrow \mathfrak{m} \supset \mathfrak{p}$), and so the only maximal ideal in

$A/\mathfrak{p}^m$ is $\mathfrak{p}/\mathfrak{p}^m$; in particular, $A/\mathfrak{p}^m$ is a local ring. As $s+\mathfrak{p}^m$ is not in $\mathfrak{p}/\mathfrak{p}^m$, it is a unit in $A/\mathfrak{p}^m$, and so $sa=0$ in $A/\mathfrak{p}^m \Rightarrow a=0$ in $A/\mathfrak{p}^m$, i.e., $a\in\mathfrak{p}^m$.

We now prove that the map is surjective. Let $\frac{a}{s}\in A_{\mathfrak{p}}$. Because $s\notin\mathfrak{p}$ and $\mathfrak{p}$ is maximal, we have that $(s)+\mathfrak{p}=A$, i.e., (s) and $\mathfrak{p}$ are relatively prime. Therefore (s) and $\mathfrak{p}^m$ are relatively prime, and so there exist $b\in A$ and $q\in\mathfrak{p}^m$ such that $bs+q=1$. Then b maps to s^{-1} in $A_{\mathfrak{p}}/\mathfrak{q}^m$ and so ba maps to $\frac{a}{s}$. More precisely: because s is invertible in $A_{\mathfrak{p}}/\mathfrak{q}^m$, $\frac{a}{s}$ is the *unique* element of this ring such that $s\frac{a}{s}=a$; since $s(ba)=a(1-q)$, the image of ba in $A_{\mathfrak{p}}$ also has this property and therefore equals $\frac{a}{s}$. □

REMARK 3.12. Consider an integral domain A and a multiplicative subset S of A. For an ideal $\mathfrak{a}$ of A, write $\mathfrak{a}^e$ for the ideal it generates in $S^{-1}A$; for an ideal $\mathfrak{a}$ of $S^{-1}A$, write $\mathfrak{a}^c$ for $\mathfrak{a}\cap A$. Then we have shown (1.5; proof of 3.4; proof of 3.11):

$\mathfrak{a}^{ce}=\mathfrak{a}$ (all ideals $\mathfrak{a}$ of $S^{-1}A$);

$\mathfrak{a}^{ec}=\mathfrak{a}$ if $\mathfrak{a}$ is a prime ideal disjoint from S, or if $\mathfrak{a}$ is a power of a maximal ideal $\mathfrak{p}$ and $S=A-\mathfrak{p}$.

We now prove that a nonzero ideal $\mathfrak{a}$ of A can be factored into a product of prime ideals. According to 3.7 (applied to A), $\mathfrak{a}$ contains a product of nonzero prime ideals,

$$\mathfrak{b}=\mathfrak{p}_1^{r_1}\cdots\mathfrak{p}_m^{r_m}.$$

We may suppose that the $\mathfrak{p}_i$ are distinct. Then

$$A/\mathfrak{b}\approx A/\mathfrak{p}_1^{r_1}\times\cdots\times A/\mathfrak{p}_m^{r_m}\approx A_{\mathfrak{p}_1}/\mathfrak{q}_1^{r_1}\times\cdots\times A_{\mathfrak{p}_m}/\mathfrak{q}_m^{r_m}$$

where $\mathfrak{q}_i=\mathfrak{p}_iA_{\mathfrak{p}_i}$ is the maximal ideal of $A_{\mathfrak{p}_i}$. The first isomorphism is given by the Chinese Remainder Theorem (and 3.8), and the second is given by (3.11). Under this isomorphism, $\mathfrak{a}/\mathfrak{b}$ corresponds to $\mathfrak{q}_1^{s_1}/\mathfrak{q}_1^{r_1}\times\cdots\times\mathfrak{q}_m^{s_m}/\mathfrak{q}_m^{r_m}$ for some $s_i\le r_i$ (recall that the rings $A_{\mathfrak{p}_i}$ are all discrete valuation rings). Since this ideal is also the image of $\mathfrak{p}_1^{s_1}\cdots\mathfrak{p}_m^{s_m}$ under the isomorphism, we see that

$$\mathfrak{a}=\mathfrak{p}_1^{s_1}\cdots\mathfrak{p}_m^{s_m}\text{ in }A/\mathfrak{b}.$$

Both of these ideals contain $\mathfrak{b}$, and so this implies that

$$\mathfrak{a}=\mathfrak{p}_1^{s_1}\cdots\mathfrak{p}_m^{s_m}$$

in A (because there is a one-to-one correspondence between the ideals of $A/\mathfrak{b}$ and the ideals of A containing $\mathfrak{b}$).

To complete the proof of Theorem 3.6, we have to prove that the above factorization is unique, but in the course of the proof, we showed that s_i is determined by the condition,

$$\mathfrak{a}A_{\mathfrak{p}_i}=\mathfrak{q}_i^{s_i},\quad \mathfrak{q}_i\text{ the maximal ideal in }A_{\mathfrak{p}_i}.$$

REMARK 3.13. Note that

$$s_i>0\iff\mathfrak{a}A_{\mathfrak{p}_i}\ne A_{\mathfrak{p}_i}\iff\mathfrak{a}\subset\mathfrak{p}_i.$$

COROLLARY 3.14. *Let $\mathfrak{a}$ and $\mathfrak{b}$ be ideals in A; then*

$$\mathfrak{a}\subset\mathfrak{b}\iff\mathfrak{a}A_{\mathfrak{p}}\subset\mathfrak{b}A_{\mathfrak{p}}$$

for all ideals nonzero prime ideals $\mathfrak{p}$ of A. In particular, $\mathfrak{a}=\mathfrak{b}$ if and only if $\mathfrak{a}A_{\mathfrak{p}}=\mathfrak{b}A_{\mathfrak{p}}$ for all $\mathfrak{p}$.

PROOF. The necessity is obvious. For the sufficiency, factor $\mathfrak{a}$ and $\mathfrak{b}$

$$\mathfrak{a} = \mathfrak{p}_1^{r_1} \cdots \mathfrak{p}_m^{r_m}, \quad \mathfrak{b} = \mathfrak{p}_1^{s_1} \cdots \mathfrak{p}_m^{s_m}, \quad r_i, s_i \geq 0.$$

Then

$$\mathfrak{a}A_{\mathfrak{p}_i} \subset \mathfrak{b}A_{\mathfrak{p}_i} \iff r_i \geq s_i,$$

(recall that $A_{\mathfrak{p}_i}$ is a discrete valuation ring) and $r_i \geq s_i$ all i implies $\mathfrak{a} \subset \mathfrak{b}$. □

COROLLARY 3.15. *Let A be an integral domain with only finitely many prime ideals; then A is a Dedekind domain if and only if it is a principal ideal domain.*

PROOF. Assume A is a Dedekind domain. After (3.6), to show that A is principal, it suffices to show that the prime ideals are principal. Let $\mathfrak{p}_1, \ldots, \mathfrak{p}_m$ be these ideals. Choose an element $x_1 \in \mathfrak{p}_1 - \mathfrak{p}_1^2$. According to the Chinese Remainder Theorem (1.7), there is an element $x \in A$ such that

$$x \equiv x_1 \mod \mathfrak{p}_1^2, \quad x \equiv 1 \mod \mathfrak{p}_i, \quad i \neq 1.$$

Now the ideals $\mathfrak{p}_1$ and (x) generate the same ideals in $A_{\mathfrak{p}_i}$ for all i, and so they are equal in A (by 3.14). □

COROLLARY 3.16. *Let $\mathfrak{a} \supset \mathfrak{b} \neq 0$ be two ideals in a Dedekind domain; then $\mathfrak{a} = \mathfrak{b} + (a)$ for some $a \in A$.*

PROOF. Let $\mathfrak{b} = \mathfrak{p}_1^{r_1} \ldots \mathfrak{p}_m^{r_m}$ and $\mathfrak{a} = \mathfrak{p}_1^{s_1} \ldots \mathfrak{p}_m^{s_m}$ with $r_i, s_j \geq 0$. Because $\mathfrak{b} \subset \mathfrak{a}$, $s_i \leq r_i$ for all i. For $1 \leq i \leq m$, choose an $x_i \in A$ such that $x_i \in \mathfrak{p}_i^{s_i}$, $x_i \notin \mathfrak{p}_i^{s_i+1}$. By the Chinese Remainder Theorem, there is an $a \in A$ such that

$$a \equiv x_i \mod \mathfrak{p}_i^{r_i}, \text{ for all } i.$$

Now one sees that $\mathfrak{b} + (a) = \mathfrak{a}$ by looking at the ideals they generate in $A_{\mathfrak{p}}$ for all $\mathfrak{p}$. □

COROLLARY 3.17. *Let $\mathfrak{a}$ be an ideal in a Dedekind domain, and let a be any nonzero element of $\mathfrak{a}$; then there exists a $b \in \mathfrak{a}$ such that $\mathfrak{a} = (a, b)$.*

PROOF. Apply (3.16) to $\mathfrak{a} \supset (a)$. □

COROLLARY 3.18. *Let $\mathfrak{a}$ be a nonzero ideal in a Dedekind domain; then there exists a nonzero ideal $\mathfrak{a}^*$ in A such that $\mathfrak{a}\mathfrak{a}^*$ is principal. Moreover, $\mathfrak{a}^*$ can be chosen to be relatively prime to any particular ideal $\mathfrak{c}$, and it can be chosen so that $\mathfrak{a}\mathfrak{a}^* = (a)$ with a any particular element of $\mathfrak{a}$ (but not both).*

PROOF. Let $a \in \mathfrak{a}$, $a \neq 0$; then $\mathfrak{a} \supset (a)$, and so we have

$$(a) = \mathfrak{p}_1^{r_1} \cdots \mathfrak{p}_m^{r_m} \text{ and } \mathfrak{a} = \mathfrak{p}_1^{s_1} \cdots \mathfrak{p}_m^{s_m}, \quad s_i \leq r_i.$$

If $\mathfrak{a}^* = \mathfrak{p}_1^{r_1 - s_1} \cdots \mathfrak{p}_m^{r_m - s_m}$, then $\mathfrak{a}\mathfrak{a}^* = (a)$.

We now show that $\mathfrak{a}^*$ can be chosen to be prime to $\mathfrak{c}$. We have $\mathfrak{a} \supset \mathfrak{a}\mathfrak{c}$, and so (by 3.16) there exists an $a \in \mathfrak{a}$ such that $\mathfrak{a} = \mathfrak{a}\mathfrak{c} + (a)$. As $\mathfrak{a} \supset (a)$, we have $(a) = \mathfrak{a} \cdot \mathfrak{a}^*$ for some ideal $\mathfrak{a}^*$ (by the above argument); now, $\mathfrak{a}\mathfrak{c} + \mathfrak{a}\mathfrak{a}^* = \mathfrak{a}$, and so $\mathfrak{c} + \mathfrak{a}^* = A$. (Otherwise $\mathfrak{c} + \mathfrak{a}^* \subset \mathfrak{p}$ some prime ideal, and $\mathfrak{a}\mathfrak{c} + \mathfrak{a}\mathfrak{a}^* = \mathfrak{a}(\mathfrak{c} + \mathfrak{a}^*) \subset \mathfrak{a}\mathfrak{p} \neq \mathfrak{a}$.) □

In basic graduate algebra courses, it is shown that

$$A \text{ a principal ideal domain} \Rightarrow A \text{ is a unique factorization domain.}$$

The converse is false (e.g., $k[X,Y]$, k a field, is a unique factorization domain but the ideal (X,Y) is not principal), but it is true for Dedekind domains.

PROPOSITION 3.19. *A Dedekind domain is a unique factorization domain if and only if it is a principal ideal domain.*

PROOF. Certainly, a principal ideal domain is Dedekind. Conversely, let A be a Dedekind domain with unique factorization. It suffices to show that the nonzero prime ideals are principal—let $\mathfrak{p}$ be such an ideal. It will contain a nonzero element, which (because of 1.2) is a product of irreducible elements. Because $\mathfrak{p}$ is prime, it will contain one of the irreducible factors π, and we know from (3.18) that there exists an ideal $\mathfrak{p}^*$ such that $\mathfrak{p}\mathfrak{p}^* = (\pi)$. I will show that $\mathfrak{p}^* = A$, and so $\mathfrak{p} = (\pi)$. From (3.18) we know that there are ideals $\mathfrak{q}$ and $\mathfrak{q}^*$ such that

$$\mathfrak{p}\mathfrak{q} = (a), \quad \mathfrak{q} + \mathfrak{p}^* = A; \quad \mathfrak{q}\mathfrak{q}^* = (b), \quad \mathfrak{q}^* + \mathfrak{p} = A$$

for some $a, b \in A$. Since $(\pi b) = \mathfrak{p}\mathfrak{p}^*\mathfrak{q}\mathfrak{q}^* = (a)\mathfrak{p}^*\mathfrak{q}^*$, we see that $a|\pi b$, and so $c = \frac{\pi b}{a} \in A$. Then $\pi b = ac$, and because A is a unique factorization domain, this implies that $\pi|a$ or $\pi|c$.

If $\pi|a$, then $\frac{a}{\pi} \in A$, and $(\frac{a}{\pi})\mathfrak{p}^* = \mathfrak{q}$. Thus any prime ideal dividing $\mathfrak{p}^*$ will also divide $\mathfrak{q}$, and this is impossible because $\mathfrak{q}$ and $\mathfrak{p}^*$ are relatively prime. Therefore, there is no such ideal, and $\mathfrak{p}^* = A$ in this case.

Similarly, if $\pi|c$, then $(\frac{c}{\pi})\mathfrak{p} = \mathfrak{q}^*$, which is impossible because $\mathfrak{p}$ does not divide $\mathfrak{q}^*$ ($\mathfrak{q}^*$ is relatively prime to $\mathfrak{p}$). Thus this case does not occur. □

The ideal class group. Let A be a Dedekind domain. A *fractional ideal* of A is a nonzero A-submodule $\mathfrak{a}$ of K such that

$$d\mathfrak{a} =_{df} \{da \mid a \in \mathfrak{a}\} \subset A$$

for some $d \in A$ (or K), i.e., it is a nonzero A-submodule of K whose elements have a common denominator. Note that a fractional ideal is *not* an ideal — when necessary to avoid confusion, we refer to ideals in A as *integral* ideals.

Equivalently, a fractional ideal of A can be defined to be a nonzero finitely generated A-submodule of K: a common denominator for the generators will be a common denominator for all the elements of the module, and, conversely, if $d\mathfrak{a}$ is an integral ideal, it is finitely generated, and this implies that $\mathfrak{a}$ is finitely generated.

Every nonzero element b of K defines a fractional ideal

$$(b) \stackrel{\text{df}}{=} bA \stackrel{\text{df}}{=} \{ba \mid a \in A\}.$$

An fractional ideal of this type is said to be *principal*.

The product of two fractional ideals is defined in the same way as for (integral) ideals

$$\mathfrak{a} \cdot \mathfrak{b} = \{\sum a_i b_i \mid a_i \in \mathfrak{a}, \quad b_i \in \mathfrak{b}\}.$$

This is again a fractional ideal: it is obviously an A-module, and if $d\mathfrak{a} \subset A$ and $e\mathfrak{b} \subset A$, then $de\mathfrak{a}\mathfrak{b} \subset A$. For principal fractional ideals, $(a)(b) = (ab)$.

EXAMPLE 3.20. Let A be a discrete valuation ring with maximal ideal $\mathfrak{p}$ and field of fractions K. Write π for a generator of $\mathfrak{p}$. Every nonzero element of K can be written uniquely in the form $a = u\pi^m$ with u a unit in A and $m \in \mathbb{Z}$. Let $\mathfrak{a}$ be a

fractional ideal of A. Then $d\mathfrak{a} \subset A$ for some $d \in A$, and we can suppose $d = \pi^n$. Thus $\pi^n\mathfrak{a}$ is an ideal in A, and so it is of the form (π^m) for some $m \geq 0$. Clearly, $\mathfrak{a} = (\pi^{m-n})$. Thus the fractional ideals of A are of the form (π^m), $m \in \mathbb{Z}$. They form a free abelian group of rank 1, and the map

$$m \mapsto (\pi^m)\colon \mathbb{Z} \to \mathrm{Id}(A)$$

is an isomorphism.

THEOREM 3.21. *Let A be a Dedekind domain. The set Id(A) of fractional ideals is a group; in fact, it is the free abelian group on the set of prime ideals.*

PROOF. We have noted that the law of composition is well-defined. It is obviously commutative. For associativity, one checks that

$$(\mathfrak{a}\mathfrak{b})\mathfrak{c} = \{\sum a_ib_ic_i \mid a_i \in \mathfrak{a}, \quad b_i \in \mathfrak{b}, \quad c_i \in \mathfrak{c}\} = \mathfrak{a}(\mathfrak{b}\mathfrak{c}).$$

The ring A plays the role of an identity element: $\mathfrak{a}A = \mathfrak{a}$. In order to show that $\mathrm{Id}(A)$ is a group, it remains to show that inverses exist.

Let $\mathfrak{a}$ be a nonzero integral ideal. According to (3.18), there is an ideal $\mathfrak{a}^*$ and an $a \in A$ such that $\mathfrak{a}\mathfrak{a}^* = (a)$. Clearly $\mathfrak{a} \cdot (a^{-1}\mathfrak{a}^*) = A$, and so $a^{-1}\mathfrak{a}^*$ is an inverse of $\mathfrak{a}$. If $\mathfrak{a}$ is a fractional ideal, then $d\mathfrak{a}$ is an integral ideal for some d, and $d \cdot (d\mathfrak{a})^{-1}$ will be an inverse for $\mathfrak{a}$.

It remains to show that the group $\mathrm{Id}(A)$ is freely generated by the prime ideals, i.e., that each fractional ideal can be expressed in a unique way as a product of powers of prime ideals. Let $\mathfrak{a}$ be a fractional ideal. Then $d\mathfrak{a}$ is an integral ideal for some $d \in A$, and we can write

$$d\mathfrak{a} = \mathfrak{p}_1^{r_1} \cdots \mathfrak{p}_m^{r_m}, \quad (d) = \mathfrak{p}_1^{s_1} \cdots \mathfrak{p}_m^{s_m}.$$

Thus $\mathfrak{a} = \mathfrak{p}_1^{r_1-s_1} \cdots \mathfrak{p}_m^{r_m-s_m}$. The uniqueness follows from the uniqueness of the factorization for integral ideals. □

REMARK 3.22. (a) Conversely, E. Noether showed that an integral domain whose fractional ideals form a group under ideal multiplication is a Dedekind domain (see Cohn 1991, p. 4.6).

(b) Let S be a multiplicative subset in a Dedekind domain A, and let $A_S = S^{-1}A$. It is an integral domain with the same field of fractions as A:

$$A \subset A_S \subset K.$$

For any fractional ideal $\mathfrak{a}$ of A, $S^{-1}\mathfrak{a} =_{df} \{\frac{a}{s} \mid a \in \mathfrak{a},\ s \in S\}$ is a fractional ideal of A_S. It is the A_S-module generated by $\mathfrak{a}$. The following hold for any fractional ideals $\mathfrak{a}$ and $\mathfrak{b}$,

$$S^{-1}(\mathfrak{a}\mathfrak{b}) = (S^{-1}\mathfrak{a})(S^{-1}\mathfrak{b}), \quad S^{-1}\mathfrak{a}^{-1} = (\mathfrak{a}A_S)^{-1}.$$

For any fractional ideal $\mathfrak{a}$, define

$$\mathfrak{a}' = \{a \in K \mid a\mathfrak{a} \subset A\}.$$

This is an A-module, and if $d \in \mathfrak{a}$, $d \neq 0$, then $d\mathfrak{a}' \subset A$, and so $\mathfrak{a}$ is a fractional ideal. From the definition of $\mathfrak{a}'$, we see that $\mathfrak{a}\mathfrak{a}'$ is an ideal in A. If it is not equal to A, then it is contained in some prime ideal $\mathfrak{p}$. When we pass to $A_\mathfrak{p}$, the inclusion $\mathfrak{a}\mathfrak{a}' \subset \mathfrak{p}$

becomes $\mathfrak{b}\mathfrak{b}' \subset \mathfrak{q}$, where $\mathfrak{b}$, $\mathfrak{b}'$, and $\mathfrak{q}$ are the ideals in $A_\mathfrak{p}$ generated by $\mathfrak{a}$, $\mathfrak{a}'$, and $\mathfrak{p}$. Moreover,

$$\mathfrak{b}' = \{a \in K \mid a\mathfrak{b} \subset A_\mathfrak{p}\}.$$

But $\mathfrak{q} = (\pi)$, and $\mathfrak{b} = (\pi^m) = \pi^m \cdot A_\mathfrak{p}$ for some $m \in \mathbb{Z}$. Clearly $\mathfrak{b}' = \pi^{-m}A$, and so $\mathfrak{b}\mathfrak{b}' = A_\mathfrak{p}$ — we have a contradiction.

We define the *ideal class group* $\mathrm{Cl}(A)$ of A to be the quotient $\mathrm{Cl}(A) = \mathrm{Id}(A)/\mathrm{P}(A)$ of $\mathrm{Id}(A)$ by the subgroup of principal ideals. The *class number* of A is the order of $\mathrm{Cl}(A)$ (when finite). In the case that A is the ring of integers $\mathcal{O}_K$ in a number field K, we often refer to $\mathrm{Cl}(\mathcal{O}_K)$ as the *ideal class group* of K, and its order as the *class number* of K.

One of the main theorems of this course will be that the class number h_K of a number field K is finite. Understanding how the class numbers of number fields vary remains an interesting problem. For example, the class number of $\mathbb{Q}[\sqrt{-m}]$ for $m > 0$ and square-free is 1 if and only if $m = 1, 2, 3, 7, 11, 19, 43, 67, 163$. It not difficult to show that these fields have class number 1, but it was not until 1954 that it was shown (by Heegner) that there were no more (and for more than 15 years, no one believed Heegner's proof to be correct). We have seen that $\mathbb{Z}[\sqrt{-5}]$ is not a principal ideal domain, and so can't have class number 1— in fact it has class number 2. The method we use to prove that the class number is finite is effective: it provides an algorithm for computing it. There are expected to be an infinite number of real quadratic fields with class number one, but this has not been proved. Gauss showed that the class group of a quadratic field $\mathbb{Q}[\sqrt{d}]$ can have arbitrarily many cyclic factors of even order, and the same is expected to be true (but is not proved) for cyclic factors of order divisible by 3 — see the thesis of M. DeLong (Michigan 1998).

It is known that every abelian group can be realized as the class group of a Dedekind domain (not necessarily the ring of integers in a number field). See Claborn, L., Every abelian group is a class group, Pacific J. Math. 18, pp. 219–222.

Example 3.23. Consider the affine elliptic curve

$$Y^2 = X^3 + aX + b, \quad \Delta = -4a^3 - 27b^2 \neq 0.$$

The associated ring $A = \mathbb{C}[X, Y]/(Y^2 - X^3 - aX - b)$ of regular functions on A is a Dedekind domain, and its class group is uncountable. In fact, it is isomorphic in a natural way to $\mathbb{C}/\Lambda$ for some lattice Λ in $\mathbb{C}$. (Exercise for those familiar with the theory of elliptic curves.)

Proposition 3.24. *Let A be a Dedekind domain, and let S be a multiplicative set in A. Then $\mathfrak{a} \mapsto S^{-1}\mathfrak{a}$ defines an isomorphism from the subgroup of Id(A) generated by prime ideals not meeting S to the group Id$(S^{-1}A)$.*

Proof. Immediate consequence of 1.5 and 3.21. □

Remark 3.25. Let A be a Dedekind domain with finite ideal class group. There is then a finite set of ideals $\mathfrak{a}_1, ..., \mathfrak{a}_m$ which is a set of representatives for the ideal classes. Clearly we may take the $\mathfrak{a}_i$ to be integral. Let b be any element in $\cap\mathfrak{a}_i$, and let S be the multiplicative set generated by b, $S = \{1, b, b^2, \dots\}$. I claim that $S^{-1}A$ is a principal ideal domain.

By assumption, any ideal $\mathfrak{a} \subset A$ can be written $\mathfrak{a} = (a) \cdot \mathfrak{a}_i$ for some $a \in K^\times$ and i, $1 \le i \le m$. Because the map $\mathfrak{b} \mapsto S^{-1}\mathfrak{b}$ is a homomorphism we have $S^{-1}\mathfrak{a} = (a) \cdot S^{-1}\mathfrak{a}_i$ where (a) now denotes the ideal generated by a in $S^{-1}A$. Since $S^{-1}\mathfrak{a}_i$ contains a unit, it is the whole ring. Thus $S^{-1}\mathfrak{a} = (a)$, and we see that every ideal in $S^{-1}A$ of the form $S^{-1}\mathfrak{a}$ is principal. According to (3.12), all ideals of $S^{-1}A$ are of this form.

REMARK 3.26. The following conditions on an integral domain A are equivalent:

(a) A is a Dedekind domain;
(b) for every prime ideal $\mathfrak{p}$ of A, $A_\mathfrak{p}$ is a discrete valuation ring;
(c) the fractional ideals of A form a group;
(d) for every fractional ideal $\mathfrak{a}$ of A, there is an ideal $\mathfrak{b}$ such that $\mathfrak{ab} = A$.

We have seen that (a) implies (b) and (c), and (d) is certainly implied by (c). The converses can be found in several books (e.g., Atiyah and MacDonald 1969).

Discrete valuations. Let K be a field. A *discrete valuation* on K is a nonzero homomorphism $v\colon K^\times \to \mathbb{Z}$ such that $v(a+b) \ge \min(v(a), v(b))$. As v is not the zero homomorphism, its image is a nonzero subgroup of $\mathbb{Z}$, and is therefore of the form $m\mathbb{Z}$ for some $m \in \mathbb{Z}$. If $m = 1$, then $v\colon K^\times \to \mathbb{Z}$ is surjective, and v is said to be *normalized*; otherwise, $x \mapsto m^{-1} \cdot v(x)$ will be a normalized discrete valuation.

EXAMPLE 3.27. (a) Let $\mathcal{M}$ be the field of meromorphic functions on a connected open subset U of the complex plane (or, better, a compact Riemann surface). For each $P \in U$ and nonzero $f \in \mathcal{M}$, define $\operatorname{ord}_P(f)$ to be $-m$, m, or 0 according as f has a pole of order m at P, a zero of order m at P, or neither a pole nor a zero at P. Then ord_P is a normalized discrete valuation on $\mathcal{M}$.

(b) Let A be a principal ideal domain with field of fractions K, and let π be a prime element of A. Then each element c of $K^\times$ can be written $c = \pi^m \frac{a}{b}$ with a and b elements of A relatively prime to π. Define $v(c) = m$. Then v is a normalized discrete valuation on K.

(c) Let A be a Dedekind domain and let $\mathfrak{p}$ be a prime ideal in A. For any $c \in K^\times$, let $\mathfrak{p}^{v(c)}$ be the power of $\mathfrak{p}$ in the factorization of (c). Then v is a normalized discrete valuation on K.

In all these examples, we have that $v(a+b) = v(b)$ if $v(a) > v(b)$. This is in fact a general property of discrete valuations. First note that $v(\zeta) = 0$ for any element of $K^\times$ of finite order (v is a homomorphism and $\mathbb{Z}$ has no elements of finite order); hence $v(-a) = v(-1) + v(a) = v(a)$. Therefore, if $v(a) > v(b)$, we have

$$v(b) = v(a + b - a)) \ge \min(v(a+b), v(a)) \ge \min(v(a), v(b)) = v(b),$$

and so equality must hold throughout, and this implies $v(a+b) = v(b)$.

We often use "ord" rather than "v" to denote a discrete valuation; for example, we often use $\operatorname{ord}_\mathfrak{p}$ to denote the discrete valuation defined by $\mathfrak{p}$ in (c).

Example (b) shows that every discrete valuation ring gives rise to a discrete valuation on its field of fractions. There is a converse to this statement.

PROPOSITION 3.28. *Let v be a discrete valuation on K, then*

$$A \overset{df}{=} \{a \in K \mid v(a) \ge 0\}$$

is a principal ideal domain with maximal ideal

$$\mathfrak{m} \stackrel{df}{=} \{a \in K \mid v(a) > 0\}.$$

If $v(K^\times) = m\mathbb{Z}$*, then the ideal* $\mathfrak{m}$ *is generated by any element* π *such that* $v(\pi) = m$*.*

PROOF. Routine. □

Later we shall see that a discrete valuation ord defines a topology on K for which two elements x and y are close if $\mathrm{ord}(x-y)$ is large. The Chinese Remainder Theorem can be restated as an approximation theorem.

PROPOSITION 3.29. *Let* $x_1, ..., x_m$ *be elements of a Dedekind domain* A*, and let* $\mathfrak{p}_1, ..., \mathfrak{p}_m$ *be distinct prime ideals of* A*. For any integer* n*, there is an* $x \in A$ *such that*

$$ord_{\mathfrak{p}_i}(x - x_i) > n, \quad i = 1, 2, ..., m.$$

PROOF. From (3.8) we know that the ideals $\mathfrak{p}_i^{n+1}$ are relatively prime in pairs, and so (1.7) provides us with an element $x \in A$ such that

$$x \equiv x_i \mod \mathfrak{p}_i^{n+1}, \quad i = 1, 2, \dots, m,$$

i.e., such that

$$\mathrm{ord}_{\mathfrak{p}_i}(x - x_i) > n, \quad i = 1, 2, ..., m.$$

□

Integral closures of Dedekind domains. We now prove a result that implies that rings of integers in number fields are Dedekind domains, and hence that their ideals factor uniquely into products of prime ideals.

THEOREM 3.30. *Let* A *be a Dedekind domain with field of fractions* K*, and let* B *be the integral closure of* A *in a finite separable extension* L *of* K*. Then* B *is a Dedekind domain.*

We have to check the three conditions in the definition of a Dedekind domain (second page of this section).

Let R be a ring (not necessarily an integral domain). An R-module M is said to be *Noetherian* if every submodule is finitely generated. (Equivalent conditions: every ascending chain of submodules becomes stationary; every nonempty set of submodules contains a maximal element.)

LEMMA 3.31. *Let* R *be a Noetherian ring. Then any finitely generated* R*-module is Noetherian.*

PROOF. (Sketch) First show that if

$$0 \to M' \to M \to M'' \to 0$$

is exact and M' and M'' are Noetherian, then so also is M; then use induction on the number of generators of M. □

We now show that B is Noetherian. In (2.29) we showed that B is contained in a finitely generated A-module. It follows that every ideal in B is finitely generated when regarded as an A-module (being a submodule of a Noetherian A-module) and *a fortiori* as an ideal (= B-module).

Next B is integrally closed because of (2.15).

It remains to prove that every nonzero prime ideal $\mathfrak{q}$ of B is maximal. Let $\beta \in \mathfrak{q}$, $\beta \neq 0$. Then β is integral over A, and so there is an equation

$$\beta^n + a_1\beta^{n-1} + \cdots + a_n = 0, \quad a_i \in A,$$

which we may suppose to have the minimum possible degree. Then $a_n \neq 0$. As $a_n \in \beta B \cap A$, we have that $\mathfrak{q} \cap A \neq (0)$. But $\mathfrak{q} \cap A$ is a prime ideal (obviously), and so it is a maximal ideal $\mathfrak{p}$ of A, and $A/\mathfrak{p}$ is a field. We know $B/\mathfrak{q}$ is an integral domain, and the map

$$a + \mathfrak{p} \mapsto a + \mathfrak{q}$$

identifies $A/\mathfrak{p}$ with a subfield of $B/\mathfrak{q}$. As B is integral over A, $B/\mathfrak{q}$ is algebraic over $A/\mathfrak{p}$. The next lemma shows that $B/\mathfrak{q}$ is a field, and hence that $\mathfrak{q}$ is maximal.

LEMMA 3.32. *Any integral domain B containing a field k and algebraic over k is itself a field.*

PROOF. Let β be a nonzero element of B — we have to prove that it has an inverse in B. Because β is algebraic over k, the ring $k[\beta]$ is finite-dimensional as a k-vector space, and the map $x \mapsto \beta x\colon k[\beta] \to k[\beta]$ is injective (because B is an integral domain). From linear algebra we deduce that the map is surjective, and so there is an element $\beta' \in k[\beta]$ such that $\beta\beta' = 1$. □

This completes the proof of Theorem 3.30.

In fact, Theorem 3.30 is true without the assumption that L be separable over K — see Janusz 1996, I.6 for a proof of the more general result. The difficulty is that, without the separability condition, B may fail to be finitely generated as an A-module, and so the proof that it is Noetherian is more difficult.

Modules over Dedekind domains (sketch). The structure theorem for finitely generated modules over principal ideal domains has an interesting extension to modules over Dedekind domains. Throughout this subsection, A is a Dedekind domain.

First, note that a finitely generated torsion-free A-module M need not be free. For example, every nonzero fractional ideal is finitely generated and torsion-free, but it is free if and only if it is principal. Thus the best we can hope for is the following.

THEOREM 3.33. *Let A be a Dedekind domain.*

(a) *Every finitely generated torsion-free A-module M is isomorphic to a direct sum of fractional ideals,*

$$M \approx \mathfrak{a}_1 \oplus \cdots \oplus \mathfrak{a}_m.$$

(b) *Two finitely generated torsion-free A-modules $M \approx \mathfrak{a}_1 \oplus \cdots \oplus \mathfrak{a}_m$ and $N \approx \mathfrak{b}_1 \oplus \cdots \oplus \mathfrak{b}_n$ are isomorphic if and only if $m = n$ and $\prod \mathfrak{a}_i \equiv \prod \mathfrak{b}_i$ modulo principal ideals.*

Hence,

$$M \approx \mathfrak{a}_1 \oplus \cdots \oplus \mathfrak{a}_m \approx A \oplus \cdots \oplus A \oplus \mathfrak{a}_1 \cdots \mathfrak{a}_m.$$

Moreover, two fractional ideals $\mathfrak{a}$ and $\mathfrak{b}$ of A are isomorphic as A-modules if and only they define the same element of the class group of A.

The *rank* of a module M over an integral domain R is the dimension of $K \otimes_R M$ as a K-vector space, where K is the field of fractions of R. Clearly the rank of $M \approx \mathfrak{a}_1 \oplus \cdots \oplus \mathfrak{a}_m$ is m.

These remarks show that the set of isomorphism classes of finitely generated torsion-free R-modules of rank 1 can be identified with the class group of A. Multiplication of elements in $\mathrm{Cl}(A)$ corresponds to the formation of tensor product of modules. The Grothendieck group of the category of finitely generated A-modules is $\mathrm{Cl}(A) \oplus \mathbb{Z}$.

THEOREM 3.34 (Invariant factor theorem). *Let $M \supset N$ be finitely generated torsion-free A-modules of the same rank m. Then there exist elements $e_1, ..., e_m$ of M, fractional ideals $\mathfrak{a}_1, ..., \mathfrak{a}_m$, and integral ideals $\mathfrak{b}_1 \supset \mathfrak{b}_2 \supset ... \supset \mathfrak{b}_m$ such that*

$$M = \mathfrak{a}_1 e_1 \oplus \cdots \oplus \mathfrak{a}_m e_m, \quad N = \mathfrak{a}_1\mathfrak{b}_1 e_1 \oplus \cdots \oplus \mathfrak{a}_m\mathfrak{b}_m e_m.$$

The ideals $\mathfrak{b}_1$, $\mathfrak{b}_2$, ..., $\mathfrak{b}_m$ are uniquely determined by the pair $M \supset N$, and are called the *invariant factors* of N in M.

The last theorem also yields a description of finitely generated torsion A-modules.

For proofs of the above results, see Curtis, C., and Reiner, I., Representation Theory of Finite Groups and Associative Algebras, 1962, III, 22, or Narkiewicz 1990, I.3.

Factorization in extensions. Let A be a Dedekind domain with field of fractions K, and let B be the integral closure of A in a finite separable extension L of K.

A prime ideal $\mathfrak{p}$ of A will factor in B,

$$\mathfrak{p}B = \mathfrak{P}_1^{e_1} \cdots \mathfrak{P}_g^{e_g}, \quad e_i \geq 1.$$

If any of the numbers is > 1, then we say that $\mathfrak{p}$ is *ramified* in B (or L). The number e_i is called the *ramification index*. We say $\mathfrak{P}$ divides $\mathfrak{p}$ (written $\mathfrak{P}|\mathfrak{p}$) if $\mathfrak{P}$ occurs in the factorization of $\mathfrak{p}$ in B. We then write $e(\mathfrak{P}/\mathfrak{p})$ for the ramification index and $f(\mathfrak{P}/\mathfrak{p})$ for the degree of the field extension $[B/\mathfrak{P} : A/\mathfrak{p}]$ (called the *residue class degree*).

LEMMA 3.35. *A prime ideal $\mathfrak{P}$ of B divides $\mathfrak{p}$ if and only if $\mathfrak{p} = \mathfrak{P} \cap K$.*

PROOF. $\Rightarrow$: Clearly $\mathfrak{p} \subset \mathfrak{P} \cap K$, and $\mathfrak{P} \cap K \neq A$.

$\Leftarrow$: If $\mathfrak{p} \subset \mathfrak{P}$ then $\mathfrak{p}B \subset \mathfrak{P}$, and we have seen (3.13) that this implies that $\mathfrak{P}$ occurs in the factorization of $\mathfrak{p}B$. □

THEOREM 3.36. *Let m be the degree of L over K, and let $\mathfrak{P}_1, ..., \mathfrak{P}_g$ be the primes dividing $\mathfrak{p}$; then*

$$\sum_{i=1}^{g} e_i f_i = m.$$

If L is Galois over K, then all the ramification numbers are equal, and all the residue class degrees are equal, and so

$$efg = m.$$

PROOF. To prove the first part of the theorem we shall show that

$$\sum e_i f_i = [B/\mathfrak{p}B : A/\mathfrak{p}] = m.$$

For the first equality, note that $B/\mathfrak{p}B = B/\prod \mathfrak{P}_i^{e_i} \cong \prod B/\mathfrak{P}_i^{e_i}$ (Chinese Remainder Theorem), and so we have to show that $[B/\mathfrak{P}_i^{e_i} : A/\mathfrak{p}] = e_i f_i$. From the definition of f_i, we know that $B/\mathfrak{p}_i$ is a field of degree f_i over $A/\mathfrak{p}$. For each r, $\mathfrak{P}_i^{r_i}/\mathfrak{P}_i^{r_{i+1}}$ is a $B/\mathfrak{P}_i$-module, and because there is no ideal between $\mathfrak{P}_i^{r_i}$ and $\mathfrak{P}_i^{r_{i+1}}$, it must have dimension one as a $B/\mathfrak{P}_i$-vector space, and hence dimension f_i as an $A/\mathfrak{p}_i$-vector space. Therefore each quotient in the chain

$$B \supset \mathfrak{P}_i \supset \mathfrak{P}_i^2 \supset \cdots \supset \mathfrak{P}_i^{e_i}$$

has dimension f_i over $A/\mathfrak{p}$, and so the dimension of $B/\mathfrak{P}_i^{e_i}$ is $e_i f_i$.

The proof of the second equality is easy if A is a principal ideal domain: a basis $x_1, ..., x_m$ for the A-module B is also a basis for the K-vector space L and gives, by reduction mod $\mathfrak{p}$, a basis for $B/\mathfrak{p}B$ over $A/\mathfrak{p}$. [To prove the second statement, note that to say $\{x_1, ..., x_m\}$ is a basis for an A-module M means that

$$A^m \to M, \quad (a_i) \mapsto \sum a_i x_i$$

is an isomorphism. When we tensor this isomorphism with $A/\mathfrak{a}$, we obtain an isomorphism

$$(A/\mathfrak{a})^m \to M/\mathfrak{a}M, \quad (a_i) \mapsto \sum a_i \bar{x}_i$$

(see 1.11), and so $\{\bar{x}_1, ..., \bar{x}_m\}$ is a basis for $M/\mathfrak{a}M$ as an $A/\mathfrak{a}$-module.]

Now let S be a multiplicative subset of A disjoint from $\mathfrak{p}$ and such that $S^{-1}A$ is principal (e.g., $S = A - \mathfrak{p}$). Write $B' = S^{-1}B$ and $A' = S^{-1}A$. Then $\mathfrak{p}B' = \prod(\mathfrak{p}_i B')^{e_i}$ (see 3.24), and so $\sum e_i f_i = [B'/\mathfrak{p}B' \colon A'/\mathfrak{p}A']$; but A' is principal, and so $[B'/\mathfrak{p}B' \colon A'/\mathfrak{p}A'] = m$. This completes the proof of the first part of the theorem.

Now assume L is Galois over K. Clearly B is stable under the action of $\mathrm{Gal}(L/K)$, and if $\sigma \in \mathrm{Gal}(L/K)$ and $\mathfrak{P}$ is a prime ideal of B, then $\sigma\mathfrak{P}$ is also a prime ideal. Moreover, if $\mathfrak{P}$ divides $\mathfrak{p}$, then it follows from (3.35) that $\sigma\mathfrak{P}$ divides $\mathfrak{p}$. Clearly $e(\sigma\mathfrak{P}/\mathfrak{p}) = e(\mathfrak{P}/\mathfrak{p})$ and $f(\sigma\mathfrak{P}/\mathfrak{p}) = f(\mathfrak{P}/\mathfrak{p})$, and so it remains to show that $\mathrm{Gal}(L/K)$ acts transitively on the prime ideals of B dividing $\mathfrak{p}$.

Suppose $\mathfrak{P}$ and $\mathfrak{Q}$ both divide $\mathfrak{p}$, and suppose $\mathfrak{Q}$ is not conjugate to $\mathfrak{P}$, i.e., that for all $\sigma \in \mathrm{Gal}(L/K)$, $\sigma\mathfrak{P} \neq \mathfrak{Q}$. According to the Chinese Remainder Theorem, we can find an element $\beta \in \mathfrak{Q}$ such that $\beta \notin \sigma\mathfrak{P}$ for any $\sigma \in \mathrm{Gal}(L/K)$. Consider $b = \mathrm{Nm}(\beta) \stackrel{\mathrm{df}}{=} \prod \sigma\beta$. Then $b \in A$, and as $\beta \in \mathfrak{Q}$, it also lies in $\mathfrak{Q}$; hence $b \in \mathfrak{Q} \cap A = \mathfrak{p}$. On the other hand, for all $\sigma \in \mathrm{Gal}(L/K)$, $\beta \notin \sigma^{-1}\mathfrak{P}$, and so $\sigma\beta \notin \mathfrak{P}$; the fact that $\prod \sigma\beta \in \mathfrak{p} \subset \mathfrak{P}$ contradicts the primality of $\mathfrak{P}$. □

The primes that ramify. In this subsection, we obtain a description of the primes that ramify in an extension.

THEOREM 3.37. *Let L be a finite extension of a number field K, let A be a Dedekind domain in K with field of fractions K (e.g., $A = \mathcal{O}_K$), and let B be the integral closure of A in L. Assume that K is a number field and that B is a free A-module (this is true for example if A is principal ideal domain). Then a prime*

$\mathfrak{p}$ ramifies in L if and only if $\mathfrak{p} | \operatorname{disc}(B/A)$. In particular, only finitely many prime ideals ramify.

We obtain this as the consequence of a series of lemmas.

LEMMA 3.38. *Let A be a ring and let B be a ring containing A and admitting a finite basis $\{e_1, ..., e_m\}$ as an A-module. For any ideal $\mathfrak{a}$ of A, $\{\bar{e}_1, ..., \bar{e}_m\}$ is a basis for the $A/\mathfrak{a}$-module $B/\mathfrak{a}B$, and*

$$D(\bar{e}_1, ..., \bar{e}_m) \equiv D(e_1, ..., e_m) \mod \mathfrak{a}.$$

PROOF. We noted in the proof of (3.36) that $\bar{e}_1, ..., \bar{e}_m$ is a basis for $B/\mathfrak{a}B$. The second assertion is obvious from the definitions. □

LEMMA 3.39. *Let A be a ring and let $B_1, ..., B_g$ be rings containing A and free of finite rank as A-modules. Then*

$$\operatorname{disc}((\Pi B_i)/A) = \prod \operatorname{disc}(B_i/A).$$

PROOF. Choose bases ε_i for each of the B_i (as A-modules), and compute the discriminant of B/A using the basis $\cup_i \varepsilon_i$. □

An element α of a ring is said to be *nilpotent* if $\alpha^m = 0$ for some $m > 1$. A ring is said to be *reduced* if it has no nonzero nilpotent elements.

LEMMA 3.40. *Let k be a perfect field, and let B be a k-algebra of finite dimension. Then B is reduced if and only if* $\operatorname{disc}(B/k) \neq 0$.

PROOF. Let $\beta \neq 0$ be a nilpotent element of B, and choose a basis $e_1, \dots, e_m$ for B with $e_1 = \beta$. Then βe_i is nilpotent for all i, and so the k-linear map

$$x \mapsto \beta e_i x \colon B \to B$$

is nilpotent. Its matrix is also nilpotent, but a nilpotent matrix has trace zero—its minimum polynomial (and hence its characteristic polynomial) is of the form X^r—and so the first row of the matrix $(\operatorname{Tr}(e_i e_j))$ is zero. Therefore its determinant is zero.

Conversely, suppose B is reduced. We first show that the intersection $\mathfrak{N}$ of the prime ideals of B is zero (this, in fact, is true for any reduced Noetherian ring). Let $b \in B$, $b \neq 0$. Let Σ be the set of ideals of B containing no power of b. Because b is not nilpotent, Σ contains the zero ideal, and hence is nonempty. Because B is Noetherian, Σ has a maximal element $\mathfrak{p}$. We shall show that $\mathfrak{p}$ is prime. Since $b \notin \mathfrak{p}$, this will show that $b \notin \mathfrak{N}$.

Let x, y be elements of B not in $\mathfrak{p}$. Then $\mathfrak{p} + (x)$ and $\mathfrak{p} + (y)$ strictly contain $\mathfrak{p}$, and so

$$b^m \in \mathfrak{p} + (x), \quad b^n \in \mathfrak{p} + (y)$$

for some m, n, say,

$$b^m = p + cx, \quad b^n = p' + c'y, \quad p, p' \in \mathfrak{p}, \quad c, c' \in B.$$

Then $b^{m+n} = pp' + pc'y + p'cx + cc'xy \in \mathfrak{p} + (xy)$, and so $\mathfrak{p} + (xy)$ is not in Σ; in particular, $\mathfrak{p} + (xy) \neq \mathfrak{p}$, and $xy \notin \mathfrak{p}$. Therefore $\mathfrak{p}$ is prime ideal, which completes the proof that $\mathfrak{N} = 0$.

Let $\mathfrak{p}$ be a prime ideal of B. Then $B/\mathfrak{p}$ is an integral domain, algebraic over k, and hence is a field (by 3.32). Therefore $\mathfrak{p}$ is maximal. Let $\mathfrak{p}_1, \mathfrak{p}_2, \dots, \mathfrak{p}_r$ be prime ideals of B. Since they are all maximal, they are relatively prime in pairs. Therefore the Chinese remainder theorem shows that

$$B/\cap \mathfrak{p}_i = \prod B/\mathfrak{p}_i \qquad (*).$$

Note that

$$[B:k] \geq [B/\cap \mathfrak{p}_i : k] = \sum [B/\mathfrak{p}_i : k] \geq r.$$

Therefore B has only finitely many prime ideals, say $\mathfrak{p}_1, \dots, \mathfrak{p}_g$ where $g \leq [B\colon k]$, and $\cap \mathfrak{p}_i = 0$. When we take $r = g$ in (*) we find that

$$B = \prod_{i=1}^{g} B/\mathfrak{p}_i.$$

For each i, $B/\mathfrak{p}_i$ is a field, and it is a finite extension of k. Because k is perfect, it is even a separable extension of k. Now we can apply (2.25) to deduce that $\operatorname{disc}((B/\mathfrak{p}_i)/k) \neq 0$, and we can apply the preceding lemma to deduce that $\operatorname{disc}(B/k) \neq 0$. □

We now prove the theorem. From the first lemma, we see that

$$\operatorname{disc}(B/A) \mod \mathfrak{p} = \operatorname{disc}((B/\mathfrak{p}B)/(A/\mathfrak{p})),$$

and from the last lemma that $\operatorname{disc}((B/\mathfrak{p}B)/(A/\mathfrak{p})) = 0$ if and only $B/\mathfrak{p}B$ is not reduced. Let $\mathfrak{p}B = \prod \mathfrak{P}_i^{e_i}$. Then $B/\mathfrak{p}B \cong \prod B/\mathfrak{P}^{e_i}$, and

$$\prod B/\mathfrak{P}^{e_i} \text{ is reduced} \iff \text{each } B/\mathfrak{P}^{e_i} \text{ is reduced} \iff \text{each } e_i = 1.$$

Remark 3.41. (a) In fact there is a precise, but complicated, relation between the power of $\mathfrak{p}$ dividing $\operatorname{Disc}(B/A)$ and the extent to which $\mathfrak{p}$ ramifies in B. It implies for example that $\operatorname{ord}_{\mathfrak{p}}(\operatorname{disc}(B/A)) \geq \sum f_i(e_i - 1)$, and that equality holds if no e_i is divisible by the characteristic of $A/\mathfrak{p}$. (See Serre 1962, III 6.)

(b) Let A be the ring of integers in a number field K, and let B be the integral closure of A in a finite extension L of K. It is possible to define $\operatorname{disc}(B/A)$ (as an ideal) without assuming B to be a free A-module. Let $\mathfrak{p}$ be an ideal in A, and let $S = A - \mathfrak{p}$. Then $S^{-1}A = A_{\mathfrak{p}}$ is principal, and so we can define $\operatorname{Disc}(S^{-1}B/S^{-1}A)$. It is a power $(\mathfrak{p}A_{\mathfrak{p}})^{m(\mathfrak{p})}$ of $\mathfrak{p}A_{\mathfrak{p}}$. Define

$$\operatorname{disc}(B/A) = \prod \mathfrak{p}^{m(\mathfrak{p})}.$$

The index $m(\mathfrak{p})$ is nonzero for only finitely many $\mathfrak{p}$, and so this formula does define an ideal in A. Clearly this definition agrees with the usual one when B is a free A-module, and the above proof shows that a prime ideal $\mathfrak{p}$ ramifies in B if and only if it divides $\operatorname{Disc}(B/A)$.

Example 3.42. (For experts on Riemann surfaces.) Let X and Y be compact connected Riemann surfaces, and let $\alpha\colon Y \to X$ be a nonconstant holomorphic mapping. Write $\mathcal{M}(X)$ and $\mathcal{M}(Y)$ for the fields of meromorphic functions on X and Y. The map $f \mapsto f \circ \alpha$ is an inclusion $\mathcal{M}(X) \hookrightarrow \mathcal{M}(Y)$ which makes $\mathcal{M}(Y)$ into a field of finite degree over $\mathcal{M}(X)$; let m be this degree. Geometrically, the map is $m\colon 1$ except at a finite number of branch points.

Let $P \in X$ and let $\mathcal{O}_P$ be the set of meromorphic functions on X that are holomorphic at P — it is the discrete valuation ring attached to the discrete valuation ord_P, and its maximal ideal is the set of meromorphic functions on X that are zero at P. Let B be the integral closure of $\mathcal{O}_P$ in $\mathcal{M}(Y)$. Let $\alpha^{-1}(P) = \{Q_1, ..., Q_g\}$ and let e_i be the number of sheets of Y over X that coincide at Q_i. Then $\mathfrak{p}B = \prod \mathfrak{q}_i^{e_i}$ where $\mathfrak{q}_i$ is the prime ideal $\{f \in B \mid f(Q_i) = 0\}$.

Finding factorizations. The following result often makes it very easy to factor an ideal in an extension field. Again A is a Dedekind domain with field of fractions K, and B is the integral closure of A in a finite separable extension L of K.

THEOREM 3.43. *Suppose that $B = A[\alpha]$, and let $f(X)$ be the minimum polynomial of α over K. Let $\mathfrak{p}$ be a prime ideal in A. Choose monic polynomials $g_1(X), \dots, g_r(X)$ in $A[X]$ that are distinct and irreducible modulo $\mathfrak{p}$, and such that $f(X) \equiv \prod g_i(X)^{e_i}$ modulo $\mathfrak{p}$. Then*

$$\mathfrak{p}B = \prod (\mathfrak{p}, g_i(\alpha))^{e_i}$$

is the factorization of $\mathfrak{p}B$ into a product of powers of distinct prime ideals. Moreover, the residue field $B/(\mathfrak{p}, g_i(\alpha)) \approx (A/\mathfrak{p})[X]/(\bar{g}_i)$, and so the residue class degree f_i is equal to the degree of g_i.

PROOF. Our assumption is that the map $X \mapsto \alpha$ defines an isomorphism

$$A[X]/(f(X)) \to B.$$

When we divide out by $\mathfrak{p}$ (better, tensor with $A/\mathfrak{p}$), this becomes an isomorphism

$$k[X]/(\bar{f}(X)) \to B/\mathfrak{p}B, \quad X \mapsto \alpha.$$

where $k = A/\mathfrak{p}$. The ring $k[X]/(\bar{f})$ has maximal ideals $(\bar{g}_1), ..., (\bar{g}_r)$, and $\prod(\bar{g}_i)^{e_i} = 0$ (but no product with smaller exponents is zero). The ideal $(\bar{g}_i)$ in $k[X]/(\bar{f})$ corresponds to the ideal $(g_i(\alpha)) + \mathfrak{p}B$ in $B/\mathfrak{p}B$, and this corresponds to the ideal $\mathfrak{p}_i =_{df} (\mathfrak{p}, g_i(\alpha))$ in B. Thus $\mathfrak{p}_1, ..., \mathfrak{p}_r$ is the complete set of prime ideals containing $\mathfrak{p}B$, and hence is the complete set of prime divisors of $\mathfrak{p}$ (see 3.13). When we write $\mathfrak{p}B = \prod \mathfrak{P}_i^{e_i}$, then the e_i are characterized by the fact that $\mathfrak{p}B \supset \prod \mathfrak{P}_i^{e_i}$, but $\mathfrak{p}B$ does not contain the product if any e_i is replaced with a smaller value. Thus it follows from the above (parenthetical) statement that e_i is the exponent of $\bar{g}_i$ occurring in the factorization of $\bar{f}$. □

REMARK 3.44. When it applies the last theorem can be used to prove (3.36) and (3.37). For example, $m = \deg(f)$, and so the equation $m = \sum e_i f_i$ is simply the equation $\deg(f) = \sum e_i \cdot \deg(g_i)$. Also, $\text{disc}(B/A) = \text{disc}(f(X))$, and this is divisible by $\mathfrak{p}$ if and only if $\bar{f}(X)$ has multiple factors (when regarded as an element of $(A/\mathfrak{p})[X]$), i.e., if and only if some $e_i > 0$.

REMARK 3.45. The conclusion of the theorem holds for a particular prime $\mathfrak{p}$ of A under the following weaker hypothesis: $\text{disc}(1, \alpha, ..., \alpha^{m-1}) = \mathfrak{a}\cdot\text{Disc}(B/A)$ with $\mathfrak{a}$ an ideal of A not divisible by $\mathfrak{p}$. To prove this, invert any element of $\mathfrak{a}$ not in $\mathfrak{p}$, and apply the theorem to the new ring and its integral closure.

Examples of factorizations. We use Theorem 3.43 to obtain some factorizations.

EXAMPLE 3.46. Let $m \neq 1$ be a square-free integer. We consider the factorization of prime integers in $K = \mathbb{Q}[\sqrt{m}]$. Recall that $\operatorname{disc}(1, \sqrt{m}) = 4m$, and that $\operatorname{disc}(\mathcal{O}_K/\mathbb{Z}) = \operatorname{disc}(1, \sqrt{m})$ if $m \equiv 2, 3 \bmod 4$, and that $\operatorname{disc}(\mathcal{O}_K/\mathbb{Z}) = \operatorname{disc}(1, \sqrt{m})/4$ if $m \equiv 1 \bmod 4$. In both cases, we can use the set $\{1, \sqrt{m}\}$ to compute the factorization of an odd prime p (see 3.45). Note that (3.36) allows only three possible factorizations of (p) in $\mathcal{O}_K$, namely,

$(p) = \mathfrak{p}^2$: (p) ramifies, $e = 2$, $f = 1$, $g = 1$;

$(p) = \mathfrak{p}$: (p) stays prime, $e = 1$, $f = 2$, $g = 1$;

$(p) = \mathfrak{p}_1\mathfrak{p}_2$: (p) splits, $e = 1$, $f = 1$, $g = 2$.

One obtains the following result.

(i) If $p|\operatorname{disc}(\mathcal{O}_K/\mathbb{Z})$, then (p) ramifies in $\mathcal{O}_K$.

(ii) For an odd prime p not dividing the m, we have

(p) is the product of two distinct ideals $\iff$ m is a square mod p, i.e., $(\frac{m}{p}) = 1$;

(p) is a prime ideal in $\mathbb{Q}[\sqrt{m}]$ $\iff$ m is not a square mod p, i.e., $(\frac{m}{p}) = -1$.

(iii) For the prime 2 when $m \equiv 1 \bmod 4$, we have

(p) is the product of two distinct ideals $\iff$ $m \equiv 1 \bmod 8$;

(p) is a prime ideal in $\mathbb{Q}[\sqrt{m}]$ $\iff$ $m \equiv 5 \bmod 8$.

To prove (iii), we must use the integral basis $\{1, \alpha\}$, $\alpha = (1 + \sqrt{m})/2$. The minimum polynomial of α is $X^2 - X + (1 - m)/4$. If $m \equiv 1 \bmod 8$, this factors as $X^2 + X = X(X + 1) \bmod 2$, and so $(2) = (2, \alpha)(2, 1 + \alpha)$. If $m \equiv 5 \bmod 8$, then $X^2 - X + (1 - m)/4 \equiv X^2 + X + 1 \bmod 2$, which is irreducible, and so $(2) = (2, 1 + \alpha + \alpha^2) = (2)$.

EXAMPLE 3.47. It is proved in basic graduate algebra courses that $\mathbb{Z}[i]$, the Gaussian integers, is a principal ideal domain. I claim that the following conditions on an odd prime p are equivalent:

(a) $p \equiv 1 \bmod 4$;
(b) (p) splits in $\mathbb{Z}[i]$;
(c) there exist integers a and b such that $p = a^2 + b^2$.

We know that (p) splits in $\mathbb{Z}[i]$ if and only if -1 is a square mod p, but this is so if and only if $\mathbb{F}_p$ contains a 4^{th} root of 1, i.e., if and only if the group $\mathbb{F}_p^\times$ contains an element of order 4. As $\mathbb{F}_p^\times$ is a cyclic group (any finite subgroup of the multiplicative group of a field is cyclic — exercise) of order $p - 1$, this is so if and only if $4|p - 1$. Thus we have shown that (a) and (b) are equivalent.

Suppose (p) splits in $\mathbb{Z}[i]$, say $(p) = \mathfrak{p}_1\mathfrak{p}_2$. Then $\mathfrak{p}_1$ and $\mathfrak{p}_2$ are principal, and if $\mathfrak{p}_1 = (a + ib)$ then $\mathfrak{p}_2 = (a - ib)$. Therefore $a^2 + b^2 = p$ up to multiplication by a unit in $\mathbb{Z}[i]$. But the only units in $\mathbb{Z}[i]$ are ± 1, $\pm i$, and so obviously[10] $a^2 + b^2 = p$. Conversely, if $p = a^2 + b^2$ with $a, b \in \mathbb{Z}$, then $(p) = (a + ib)(a - ib)$ in $\mathbb{Z}[i]$.

[The fact that every prime of the form $4n+1$ is a sum of two squares was stated as a theorem by Fermat in a letter in 1654. Euler, who was almost certainly unaware of Fermat's letter, found a proof. For some history, and a discussion of algorithms for finding a and b, see Edwards 1977, p. 55.]

[10]Following the usual convention, we generally take a prime p in $\mathbb{Z}$ to be *positive*.

REMARK 3.48. (a) From (3.43) and (3.45) we see that, for almost all p, factoring (p) in $\mathcal{O}_K$ amounts to factoring a polynomial $f(X)$ modulo p into a product of powers of irreducible polynomials. Clearly, this can always be done, but may require a lot of hard work (but not much intelligence). Hence it can safely be left to the computer. In Maple, type:

Factors(f(X)) mod p;

In Mathematica, type:

Factor[f(X), Modulus->p]

(b) In the next section, we shall show, not only that the class group of a number field is finite, but that it is generated by the prime ideals dividing a certain small set of prime numbers. Finding the class number therefore involves finding the prime ideal factors of these prime numbers, and the relations among them.

EXAMPLE 3.49. Let α be a root of $X^3+10X+1$. Recall that the discriminant of the polynomial is -4027, and so the ring of integers in $\mathbb{Q}[\alpha]$ is $\mathbb{Z}+\mathbb{Z}\alpha+\mathbb{Z}\alpha^2$. There are the following factorizations:

2	$(1+X)(1+X+X^2)$	(2)	$=$	$(2,1+\alpha)(2,1+\alpha+\alpha^2)$
3	$(2+X)(2+X+X^2)$	(3)	$=$	$(3,2+\alpha)(3,2+\alpha+\alpha^2)$
5	$(1+X)(1+4X+X^2)$	(5)	$=$	$(5,1+\alpha)(5,1+4\alpha+\alpha^2)$
7	$(3+X)(5+4X+X^2)$	(7)	$=$	$(7,3+\alpha)(7,5+4\alpha+\alpha^2)$
11	$(6+X)(2+5X+X^2)$	(11)	$=$	$(11,6+\alpha)(11,2+5\alpha+\alpha^2)$
13	$1+10X+X^3$	(13)	$=$	$(13,1+10\alpha+\alpha^2)=(13)$
17	$1+10X+X^3$	(17)	$=$	prime ideal.
4027	$(2215+X)^2(3624+X)$	(4027)	$=$	$(4027,2215+\alpha)^2(4027,3624+\alpha)$.

EXAMPLE 3.50. Let α be a root of $X^3-8X+15$. Here again, the discriminant of the polynomial is -4027, and so the ring of integers in $\mathbb{Q}[\alpha]$ is $\mathbb{Z}+\mathbb{Z}\alpha+\mathbb{Z}\alpha^2$. There are the following factorizations:

2	$(1+X)(1+X+X^2)$	(2)	$=$	$(2,1+\alpha)(2,1+\alpha+\alpha^2)$
3	$X(1+X^2)$	(3)	$=$	$(3,\alpha)(3,1+\alpha^2)$
5	$X(2+X^2)$	(5)	$=$	$(5,\alpha)(5,2+\alpha^2)$
7	$(5+X)(3+2X+X^2)$	(7)	$=$	$(7,\alpha)(7,3+2\alpha+\alpha^2)$
11	$(1+X)(4+10X+X^2)$	(11)	$=$	$(11,\alpha)(11,4+10\alpha+\alpha^2)$
13	$2+5X+X^3$	(13)	$=$	(13)
17	$(4+X)(6+X)(7+X)$	(17)	$=$	$(17,4+\alpha)(17,6+\alpha)(17,7+\alpha)$
4027	$(509+X)(1759+X)^2$.	(4027)	$=$	$(4027,509+\alpha)(4027,1759+\alpha)^2$

On comparing the factorizations of (17) in the fields in the last two examples, we see that the fields can't be isomorphic.

REMARK 3.51. When K is a number field, it is interesting to have a description of the set $\mathrm{Spl}(K)$ of prime numbers that split in K. For $K=\mathbb{Q}[\sqrt{m}]$, this is the set of p for which $(\frac{m}{p})=1$, and we shall see later that the quadratic reciprocity law gives a good description of the set. For any abelian Galois extension K of $\mathbb{Q}$, class field theory gives a similarly good description, but for an arbitrary extension very little is known about what sets can occur. There is a theorem that says that two *Galois* extensions K and K' of $\mathbb{Q}$ are isomorphic if and only if $\mathrm{Spl}(K)=\mathrm{Spl}(K')$. Moreover, this can be made into an effective procedure for determining when fields are isomorphic. See Theorem 8.38 below.

EXAMPLE 3.52. In (2.38), we saw that $f(X) = X^5 - X - 1$ is irreducible in $\mathbb{Q}[X]$, and that its discriminant is $19 \cdot 151$, which is square-free, and so, if α is a root of $f(X)$, then $\mathbb{Z}[\alpha]$ is the ring of integers in $\mathbb{Q}[\alpha]$. We have

$19 \quad (6+X)^2(10+13X+17X^2+X^3)$

$\quad (19) = (19, 6+\alpha)^2(19, 10+13\alpha+17\alpha^2+\alpha^3)$

$151 \quad (9+X)(39+X)^2\dots$

$\quad (151) = (151, 9+\alpha)(151, 39+\alpha)^2\dots$

$4027 \quad (1261+X)(2592+X)(790+3499X+174X^2+X^3).$

Thus (19) and (151) are ramified in $\mathbb{Q}[\alpha]$, and 4027 isn't, which is what Theorem 3.37 predicts.

EXAMPLE 3.53. According to Maple,

$$1 + X + X^2 + X^3 + X^4 \equiv (4+X)^4 \mod 5$$

Why is this obvious?

Eisenstein extensions. Recall that Eisenstein's Criterion says that a polynomial

$$X^m + a_1X^{m-1} + \cdots + a_m,$$

such that $a_i \in \mathbb{Z}$, $p|a_i$ all i, and p^2 does not divide a_m, is irreducible in $\mathbb{Q}[X]$. We will improve this result, but first we need to make two observations about discrete valuations.

Let A be a Dedekind domain, and let B be its integral closure in a finite extension L of its field of fractions K. Let $\mathfrak{p}$ be a prime ideal of A and let $\mathfrak{p}$ be an ideal of B dividing $\mathfrak{p}$, say $\mathfrak{p}B = \mathfrak{P}^e \cdots$. Write $\mathrm{ord}_\mathfrak{p}$ and $\mathrm{ord}_\mathfrak{P}$ for the normalized valuations on K and L defined by $\mathfrak{p}$ and $\mathfrak{P}$. Then

$$\mathrm{ord}_\mathfrak{P}|K = e \cdot \mathrm{ord}_\mathfrak{p}$$

because, if $(a) = \mathfrak{p}^m \cdots$ in A, then $(a) = \mathfrak{P}^{me} \cdots$ in B.

Next I claim that if

$$a_1 + \cdots + a_n = 0,$$

then the minimum value of $\mathrm{ord}(a_i)$ must be attained for at least two i's. Suppose not, say $\mathrm{ord}(a_1) < \mathrm{ord}(a_i)$ for all $i > 1$. Then

$$\mathrm{ord}(a_2 + a_3 + \ldots + a_m) \geq \min(\mathrm{ord}(a_2), \mathrm{ord}(a_3 + \ldots + a_m)) \geq \cdots \geq \min_{2 \leq i \leq m}(\mathrm{ord}(a_i)),$$

but $-a_1 = \sum a_i$ implies $\mathrm{ord}(a_1) = \mathrm{ord}(\sum a_i)$, which is a contradiction.

Let A be a Dedekind domain and let $\mathfrak{p}$ be a prime ideal in A. A polynomial

$$X^m + a_1X^{m-1} + \cdots + a_m, \quad a_i \in A,$$

is said to be *Eisenstein relative to* $\mathfrak{p}$ if

$$\mathrm{ord}_\mathfrak{p}(a_1) > 0, \dots, \mathrm{ord}_\mathfrak{p}(a_{m-1}) > 0, \mathrm{ord}_\mathfrak{p}(a_m) = 1.$$

PROPOSITION 3.54. *Let $f(X) \in A[X]$ be an Eisenstein polynomial with respect to $\mathfrak{p}$. Then $f(X)$ is irreducible, and if α is a root of $f(X)$, then $\mathfrak{p}$ is totally ramified in $K[\alpha]$; in fact $\mathfrak{p}B = \mathfrak{P}^m$ with $\mathfrak{P} = (\mathfrak{p}, \alpha)$ and $m = \deg(f)$.*

PROOF. Let $L = K[\alpha]$ — we have $[L\colon K] \leq m$. Let $\mathfrak{P}$ be a prime ideal dividing $\mathfrak{p}$, with ramification index e say. Consider the equation

$$\alpha^m + a_1\alpha^{m-1} + \cdots + a_m = 0.$$

Because $f(X)$ is Eisenstein,

$$\begin{aligned} \mathrm{ord}_{\mathfrak{P}}(\alpha^m) &= m \cdot \mathrm{ord}_{\mathfrak{P}}(\alpha); \\ \mathrm{ord}_{\mathfrak{P}}(a_i\alpha^{m-i}) &\geq (m-i) \cdot \mathrm{ord}_{\mathfrak{P}}(\alpha) + e; \\ \mathrm{ord}_{\mathfrak{P}}(a_m) &= e. \end{aligned}$$

If $\mathrm{ord}_{\mathfrak{P}}(\alpha) = 0$, then the minimum value of $\mathrm{ord}_{\mathfrak{P}}$ is taken for a single term, namely α^m. This is impossible, and so $\mathrm{ord}_{\mathfrak{P}}(\alpha) \geq 1$, and $\mathrm{ord}_{\mathfrak{P}}(a_i\alpha^{m-i}) > \mathrm{ord}_{\mathfrak{P}}(a_m) = e$ for $i = 1, ..., m$. From the remark preceding the proposition, we see that $m \cdot \mathrm{ord}_{\mathfrak{P}}(\alpha) = e$. Then

$$m \cdot \mathrm{ord}_{\mathfrak{P}}(\alpha) = e \leq [K[\alpha] : K] \leq m,$$

and we must have equalities throughout: $\mathrm{ord}_{\mathfrak{P}}(\alpha) = 1$, $[K(\alpha)\colon K] = m = e$. □

4. The Finiteness of the Class Number

In this section we prove the first main theorem of the course: the class number of a number field is finite. The method of proof is effective: it gives an algorithm for computing the class group.

Norms of ideals. Let A be a Dedekind domain with field of fractions K, and let B be the integral closure of A in a finite separable extension L. We want to define a homomorphism $\mathrm{Nm}\colon \mathrm{Id}(B) \to \mathrm{Id}(A)$ which is compatible with taking norms of elements, i.e., such that the following diagram commutes:

$$\begin{array}{ccc} L^\times & \to & \mathrm{Id}(B) \\ \downarrow \mathrm{Nm} & & \downarrow \mathrm{Nm} \\ K^\times & \to & \mathrm{Id}(A) \end{array} \qquad (*)$$

Because $\mathrm{Id}(B)$ is the free abelian group on the set of prime ideals, we only have to define $\mathrm{Nm}(\mathfrak{p})$ for $\mathfrak{p}$ prime.

Let $\mathfrak{p}$ be a prime ideal A, and factor $\mathfrak{p}B = \prod \mathfrak{P}_i^{e_i}$. If $\mathfrak{p}$ is principal, say $\mathfrak{p} = (\pi)$, then we should have

$$\mathrm{Nm}(\mathfrak{p}B) = \mathrm{Nm}(\pi \cdot B) = \mathrm{Nm}(\pi) \cdot A = (\pi^m) = \mathfrak{p}^m, \quad m = [L\colon K].$$

Also, because Nm is to be a homomorphism, we should have

$$\mathrm{Nm}(\mathfrak{p}B) = \mathrm{Nm}(\textstyle\prod \mathfrak{P}_i^{e_i}) = \prod \mathrm{Nm}(\mathfrak{P}_i)^{e_i}.$$

On comparing these two formulas, and recalling (3.36) that $m = \sum e_i f_i$, we see that we should define $\mathrm{Nm}(\mathfrak{P}_i) = \mathfrak{p}^{f_i}$. We take this as our definition:

$$\mathrm{Nm}(\mathfrak{P}) = \mathfrak{P}^{f(\mathfrak{P}/\mathfrak{p})} \text{ where } \mathfrak{p} = \mathfrak{P} \cap A \text{ and } f(\mathfrak{P}/\mathfrak{p}) = [B/\mathfrak{P} : A/\mathfrak{p}].$$

To avoid confusion, I sometimes use $\mathcal{N}$ to denote norms of ideals.

If we have a tower of fields $M \supset L \supset K$, then

$$\mathcal{N}_{L/K}(\mathcal{N}_{M/L}\mathfrak{a}) = \mathcal{N}_{M/K}\mathfrak{a}$$

because $f(\mathfrak{Q}/\mathfrak{P}) \cdot f(\mathfrak{P}/\mathfrak{p}) = f(\mathfrak{Q}/\mathfrak{p})$, i.e., $[C/\mathfrak{Q} : B/\mathfrak{P}] \cdot [B/\mathfrak{P} : A/\mathfrak{p}] = [C/\mathfrak{Q} : A/\mathfrak{p}]$ where $C \supset B \supset A$ are the integral closures of A in M, L, and K respectively.

Proposition 4.1. *Let $A \subset B$ and $K \subset L$ be as above.*

(a) *For any nonzero ideal $\mathfrak{a} \subset A$, $\mathcal{N}_{L/K}(\mathfrak{a}B) = \mathfrak{a}^m$, where $m = [L : K]$.*
(b) *Suppose L is Galois over K. Let $\mathfrak{P}$ be a nonzero prime ideal of B and let $\mathfrak{p} = \mathfrak{P} \cap A$. Write $\mathfrak{p} \cdot B = (\mathfrak{P}_1 \cdots \mathfrak{P}_g)^e$ (cf. 3.36). Then*

$$\mathcal{N}\mathfrak{P} \cdot B = (\mathfrak{P}_1 \cdots \mathfrak{P}_g)^{ef} = \prod_{\sigma \in \mathrm{Gal}(L/K)} \sigma\mathfrak{P}.$$

(c) *For any nonzero element $\beta \in B$, $\mathrm{Nm}(\beta) \cdot A = \mathrm{Nm}(\beta \cdot B)$ (i.e., (*) commutes).*

Proof. (a) It suffices to prove this for a prime ideal $\mathfrak{p}$, and for such an ideal we have that

$$\mathcal{N}(\mathfrak{p}B) = \mathcal{N}(\textstyle\prod \mathfrak{P}_i^{e_i}) =_{df} \mathfrak{p}^{\sum e_i f_i} = \mathfrak{p}^m \quad \text{(by 3.36)}.$$

(b) Since $\mathcal{N}\mathfrak{P}_i = \mathfrak{p}^f$ for each i, the first equality is obvious. In the course of the proof of (3.36), we showed that $\mathrm{Gal}(L/K)$ acts transitively on the set $\{\mathfrak{P}_1, ..., \mathfrak{P}_g\}$, and it follows that each $\mathfrak{P}_i$ occurs $\frac{m}{g} = ef$ times in the family $\{\sigma\mathfrak{P} \mid \sigma \in \mathrm{Gal}(L/K)\}$.

(c) Suppose first that L is Galois over K, and let $\beta \cdot B = \mathfrak{b}$. The map $\mathfrak{a} \mapsto \mathfrak{a} \cdot B \colon \mathrm{Id}(A) \to \mathrm{Id}(B)$ is injective (remember they are free abelian groups on the nonzero prime ideals), and so it suffices to show that $\mathrm{Nm}(\beta) \cdot B = \mathrm{Nm}(\mathfrak{b}) \cdot B$. But

$$\mathrm{Nm}(\mathfrak{b}) \cdot B \overset{(b)}{=} \prod \sigma\mathfrak{b} = \prod(\sigma\beta \cdot B) = (\prod \sigma\beta) \cdot B = \mathrm{Nm}(\beta) \cdot B$$

as required.

In the general case, let E be a finite Galois extension of K containing L, and let $d = [E \colon L]$. Let C be the integral closure of B in E. From (a), the Galois case, and the transitivity of $\mathcal{N}$ we have that

$$\mathcal{N}_{L/K}(\beta \cdot B)^d = \mathcal{N}_{E/K}(\beta \cdot C) = \mathrm{Nm}_{E/K}(\beta) \cdot A = \mathrm{Nm}_{L/K}(\beta)^d \cdot A.$$

As the group of ideals $\mathrm{Id}(A)$ is torsion-free, this implies that $\mathcal{N}_{L/K}(\beta \cdot B) = \mathrm{Nm}_{L/K}(\beta) \cdot A$. □

Let $\mathfrak{a}$ be a nonzero ideal in the ring of integers $\mathcal{O}_K$ of a number field K. Then $\mathfrak{a}$ is of finite index in $\mathcal{O}_K$, and we let $\mathbb{N}\mathfrak{a}$, the *numerical norm of* $\mathfrak{a}$, be this index:

$$\mathbb{N}\mathfrak{a} = (\mathcal{O}_K : \mathfrak{a}).$$

PROPOSITION 4.2. *Let $\mathcal{O}_K$ be the ring of integers in a number field K.*

(a) *For any ideal $\mathfrak{a}$ in $\mathcal{O}_K$, $\mathcal{N}_{K/\mathbb{Q}}(\mathfrak{a}) = (\mathbb{N}(\mathfrak{a}))$; therefore $\mathbb{N}(\mathfrak{a}\mathfrak{b}) = \mathbb{N}(\mathfrak{a})\mathbb{N}(\mathfrak{b})$.*
(b) *Let $\mathfrak{b} \subset \mathfrak{a}$ be fractional ideals in K; then*

$$(\mathfrak{a} : \mathfrak{b}) = \mathbb{N}(\mathfrak{a}^{-1}\mathfrak{b}).$$

PROOF. (a) Write $\mathfrak{a} = \prod \mathfrak{p}_i^{r_i}$, and let $f_i = f(\mathfrak{p}_i/p_i)$ where $(p_i) = \mathbb{Z} \cap \mathfrak{p}_i$; then $\mathrm{Nm}(\mathfrak{p}_i) = (p_i)^{f_i}$. From the Chinese remainder theorem, $\mathcal{O}_K/\mathfrak{a} \approx \prod \mathcal{O}_K/\mathfrak{p}_i^{r_i}$, and so $(\mathcal{O}_K : \mathfrak{a}) = \prod(\mathcal{O}_K : \mathfrak{p}_i^{r_i})$. In the course of the proof of (3.36), we showed that $\mathcal{O}_K/\mathfrak{p}_i^{r_i}$ is a vector space of dimension $f_i r_i$ over $\mathbb{F}_{p_i}$, and so $(\mathcal{O}_K : \mathfrak{p}_i^{r_i}) = p_i^{f_i r_i}$. On taking the product over i, we find that $(\mathcal{O}_K : \mathfrak{a}) = \prod(p_i^{f_i r_i}) = \mathcal{N}_{K/\mathbb{Q}}\mathfrak{a}$. When we identify the set of nonzero ideals in $\mathbb{Z}$ with the set of positive integers, then $\mathcal{N}$ becomes identified with $\mathbb{N}$, and so the multiplicativity of $\mathbb{N}$ follows from that of $\mathcal{N}$.

(b) For any nonzero $d \in K$, the map $x \mapsto dx : K \to K$ is an additive isomorphism, and so $(d\mathfrak{a} : d\mathfrak{b}) = (\mathfrak{a} : \mathfrak{b})$. Since $(d\mathfrak{a})(d\mathfrak{b})^{-1} = \mathfrak{a}\mathfrak{b}^{-1}$, we may suppose that $\mathfrak{a}$ and $\mathfrak{b}$ are integral ideals. The required formula then follows from (a) and the formulas

$$(\mathcal{O}_K : \mathfrak{a})(\mathfrak{a} : \mathfrak{b}) = (\mathcal{O}_K : \mathfrak{b})$$

and

$$\mathbb{N}(\mathfrak{a}) \cdot \mathbb{N}(\mathfrak{a}^{-1}\mathfrak{b}) = \mathbb{N}(\mathfrak{b}).$$

□

Statement of the main theorem and its consequences. We now state the main theorem of this section and discuss some of its consequences.

THEOREM 4.3. *Let K be an extension of degree n of $\mathbb{Q}$, and let Δ_K be the discriminant of $K/\mathbb{Q}$. Let $2s$ be the number of nonreal complex embeddings of K. Then*

there exists a set of representatives for the ideal class group of K consisting of integral ideals $\mathfrak{a}$ with

$$\mathbb{N}(\mathfrak{a}) \le \frac{n!}{n^n}\left(\frac{4}{\pi}\right)^s |\Delta_K|^{\frac{1}{2}}.$$

The number on the right is called the *Minkowski bound* — we sometimes denote it by B_K. The term $C_K = \frac{n!}{n^n}\left(\frac{4}{\pi}\right)^s$ is called the *Minkowski constant.* It takes the following values:

N	r	s	C
2	0	1	0.636
2	2	0	0.500
3	1	1	0.283
3	3	0	0.222
4	0	2	0.152
4	2	1	0.119
4	4	0	0.094
5	1	2	0.062
5	3	1	0.049
5	5	0	0.038
...	...	...	...
100	100	0	0.93×10^{-42}

Here r is the number of real embeddings of K. We have

$$K \otimes_{\mathbb{Q}} \mathbb{R} \approx \mathbb{R}^r \times \mathbb{C}^s,$$

and, if $K = \mathbb{Q}[\alpha]$ and $f(X)$ is the minimum polynomial of α, then r is the number of real roots of $f(X)$ and $2s$ is the number of its nonreal roots. To see that these descriptions of r and s are agree, apply (1.13).

Before proving (4.3), we give some applications and examples.

THEOREM 4.4. *The class number of K is finite.*

PROOF. It suffices to show that there are only finitely many integral ideals $\mathfrak{a}$ in $\mathcal{O}_K$ such that $\mathbb{N}(\mathfrak{a})$ is less than the Minkowski bound — in fact, we shall show that, for any integer M, there are only finitely many integral ideals $\mathfrak{a}$ with $\mathbb{N}(\mathfrak{a}) < M$. If $\mathfrak{a} = \prod \mathfrak{p}_i^{r_i}$, then $\mathbb{N}(\mathfrak{a}) = \prod p_i^{r_i f_i}$ where $(p_i) = \mathfrak{p}_i \cap \mathbb{Q}$. As $\mathbb{N}(\mathfrak{a}) < M$, this allows only finitely many possibilities for the p_i (and hence for the $\mathfrak{p}_i$), and only finitely many possibilities for the exponents r_i. □

Let S be the set of integral ideals in K with norm $< B_K$. Then S is a finite set, and $\mathrm{Cl}(\mathcal{O}_K) = S/\sim$, where $\mathfrak{a} \sim \mathfrak{b}$ if one ideal is the product of the other with a principal (fractional) ideal. There is an algorithm for finding S, and an algorithm for deciding whether $\mathfrak{a} \sim \mathfrak{b}$, and so there is an algorithm for finding $\mathrm{Cl}(\mathcal{O}_K)$ (the group, not just it's order). To find S, find the prime ideal factors of enough prime numbers, and form some of their products. To decide whether $\mathfrak{a} \sim \mathfrak{b}$, one has to decide whether $\mathfrak{c} = \mathfrak{a}\mathfrak{b}^{-1}$ is principal. From (4.2b) we know that, for $\gamma \in \mathfrak{c}$,

$$\mathfrak{c} = (\gamma) \iff \mathbb{N}\mathfrak{c} = |\operatorname{Nm}\gamma|$$

and so we have to solve the equation:

$$\operatorname{Nm}\gamma = \text{constant}.$$

When we express γ in terms of an integral basis, this becomes a (very special) type of diophantine equation. For a description of an algorithm for finding $\mathrm{Cl}(\mathcal{O}_K)$, see Pohst and Zassenhaus 1989, p424.

EXAMPLE 4.5. Let $K = \mathbb{Q}[i]$. The condition in the theorem is that $\mathbb{N}(\mathfrak{a}) \leq \frac{2}{4}\frac{4}{\pi}2 < 1.27$. There are no such ideals other than $\mathbb{Z}[i]$, and so $\mathbb{Z}[i]$ is a principal ideal domain. (Of course, the elementary proof of this shows more, namely, that $\mathbb{Z}[i]$ is a Euclidean domain. Even for rings of integers in number fields, it is *not* true that all principal ideal domains are Euclidean domains. For example, $\mathbb{Q}[\sqrt{-19}]$ has class number 1, but its ring of integers is not a Euclidean domain. For more on such things, see the survey article: Lemmermeyer, F., The Euclidean algorithm in algebraic number fields, Exposition. Math., 13 (1995).)

EXAMPLE 4.6. Let $K = \mathbb{Q}[\sqrt{-5}]$. Here $\mathbb{N}(\mathfrak{a}) \leq 0.63 \times \sqrt{20} < 3$. Any ideal satisfying this must divide (2). In fact, $(2) = \mathfrak{p}^2$ where $\mathfrak{p} = (2, 1+\sqrt{-5})$, and $\mathbb{N}\mathfrak{p}^2 = \mathbb{N}(2) = 4$, and so $\mathbb{N}\mathfrak{p} = 2$. The ideals $\mathcal{O}_K$ and $\mathfrak{p}$ form a set of representatives for $\mathrm{Cl}(\mathbb{Z}[\sqrt{-5}])$. The ideal $\mathfrak{p}$ can't be principal because there does not exist an element $\alpha = m + n\sqrt{-5}$ such that $\operatorname{Nm}(\alpha) = m^2 + 5n^2 = 2$, and so $\mathrm{Cl}(\mathbb{Z}[\sqrt{-5}])$ has order 2.

EXAMPLE 4.7. Let K be a cubic field with discriminant < 0. Since the sign of Δ_K is $(-1)^s$, and $[K:\mathbb{Q}] = r + 2s$, we have $s = 1$, $r = 1$. The Minkowski bound is

$$B < 0.283|\Delta_K|^{\frac{1}{2}}.$$

For $|\Delta_K| \leq 49$, $B < 2$, and so for cubic fields with $-49 \leq \Delta_K < 0$, the class number $h = 1$. For example, this is true for the number fields with discriminants -23 and -31 discussed earlier (see 2.35, 2.36).

For the field generated by a root of $X^3 + 10X + 1$, the discriminant is -4027, and the Minkowski bound is < 18. Recall from (3.49) that

$$(2) = (2, 1+\alpha)(2, 1+\alpha+\alpha^2).$$

Let $\mathfrak{p} = (2, 1+\alpha)$—its norm is 2. One can show that it generates the class group, and that it has order 6 in the class group, i.e., $\mathfrak{p}^6$ but no smaller power is principal. Hence the class group is cyclic of order 6. (The proof takes quite a bit of hard work if you do it by hand — see Artin 1959, pp. 160-162, 170-172.)

An extension L of a number field K is said to be *unramified over K* if no prime ideal of $\mathcal{O}_K$ ramifies in $\mathcal{O}_L$.

THEOREM 4.8. *There does not exist an unramified extension of $\mathbb{Q}$.*

PROOF. Let K be a finite extension of $\mathbb{Q}$. Since a set of representatives for the class group must have at least one element, and that element will have numerical norm ≥ 1, Theorem 4.3 shows that

$$|\Delta|^{\frac{1}{2}} \geq \frac{n^n}{n!}\left(\frac{\pi}{4}\right)^s \geq \frac{n^n}{n!}\left(\frac{\pi}{4}\right)^{n/2}.$$

Let $a_n = \text{rhs}$. Then $a_2 > 1$, and $\frac{a_{n+1}}{a_n} = \left(\frac{\pi}{4}\right)^{\frac{1}{2}}(1+\frac{1}{n})^n > 1$, and so the sequence a_n is monotonically increasing. Hence the discriminant of K has absolute value > 1, and we know from (3.37) that any prime dividing the discriminant ramifies. □

We can now prove that there is no irreducible monic polynomial $f(X) \in \mathbb{Z}[X]$ of degree > 1 with discriminant ± 1. Let $f(X)$ be such a polynomial, and let α be a root of $f(X)$. Then $\text{disc}(\mathbb{Z}[\alpha]/\mathbb{Z}) = \pm 1$, and so $\mathbb{Z}[\alpha]$ is the ring of integers in $K \stackrel{\text{df}}{=} \mathbb{Q}[\alpha]$ and $\text{disc}(\mathcal{O}_K/\mathbb{Z}) = \pm 1$, which we have just seen can't happen.

REMARK 4.9. There may exist unramified extensions of number fields other than $\mathbb{Q}$. In fact, class field theory says that the maximal abelian unramified[11] extension of K (called the *Hilbert class field* of K) has Galois group canonically isomorphic to $\text{Cl}(\mathcal{O}_K)$. For example, the theory says that $\mathbb{Q}[\sqrt{-5}]$ has an unramified extension of degree 2, and one verifies that $\mathbb{Q}[\sqrt{-1}, \sqrt{-5}]$ is unramified over $\mathbb{Q}[\sqrt{-5}]$.

ASIDE 4.10. Let K_1 be a number field with class number $h_{K_1} > 1$. Its Hilbert class field is an abelian unramified extension K_2 of K_1 with $\text{Gal}(K_2/K_1) \cong \text{Cl}(K_1)$. Let K_3 be the Hilbert class field of K_2, and so on. In this way, we obtain a tower of fields,

$$K_1 \subset K_2 \subset K_3 \subset \cdots$$

It was a famous question (class field tower problem) to decide whether this tower can be infinite, or must always terminate with a field of class number 1 after a finite number of steps. It was shown by Golod and Shafarevich in the early 60s that the tower is frequently infinite. (See the article by Roquette in Cassels and Fröhlich 1967.)

EXAMPLE 4.11. Let α be a root of $f(X) = X^5 - X + 1$. As we saw in (2.38) that $f(X)$ is irreducible and its discriminant is 19×151, and so the ring of integers of $K \stackrel{\text{df}}{=} \mathbb{Q}[\alpha]$ is $\mathbb{Z}[\alpha]$.

According to Theorem 4.3, every class of ideals for $\mathbb{Q}[\alpha]$ contains an integral ideal $\mathfrak{a}$ with $\mathbb{N}(\mathfrak{a}) < 0.062 \times \sqrt{19 \times 151} = 3.3 < 4$. If $\mathfrak{p}$ is a prime ideal with $\mathbb{N}(\mathfrak{p}) = 2$, then $f(\mathfrak{p}/(2)) = 1$, and $f(X)$ must have a root mod 2, but it doesn't, and so $\mathfrak{p}$ can't exist. Similarly, there is no prime ideal $\mathfrak{p}$ with $\mathbb{N}(\mathfrak{p}) = 3$, and so $\mathcal{O}_K$ is a principal ideal domain!

The Galois group of the splitting field M of $f(X)$ is S_5 (later we shall see how to find Galois groups; for the moment type "galois($X^5 - X + 1$);" in Maple), and hence $[M\colon \mathbb{Q}] = 120$. It is possible to show that M is unramified over $\mathbb{Q}[\sqrt{19 \times 151}]$.

Lattices. Let V be a vector space of dimension n over $\mathbb{R}$. A *lattice* Λ in V is a subgroup of the form

$$\Lambda = \mathbb{Z}e_1 + \cdots + \mathbb{Z}e_r$$

with $e_1, \ldots, e_r$ linearly independent elements of V. Thus a lattice is the free abelian subgroup of V generated by elements of V that are linearly independent over $\mathbb{R}$. When $r = n$, the lattice is said to be *full*. At the opposite extreme, $\Lambda = \{0\}$ is a

[11]The Hilbert class field L of K is required to be unramified even at the infinite primes — this means that every real embedding of K extends to a real embedding of L.

lattice (generated by the empty set of elements). In terms of tensor products, one can say that a full lattice in V is a subgroup Λ of V such that

$$\sum r_i \otimes x_i \mapsto \sum r_i x_i, \quad \mathbb{R} \otimes_{\mathbb{Z}} \Lambda \to V,$$

is an isomorphism.

NONEXAMPLE 4.12. The subgroup $\mathbb{Z} + \mathbb{Z}\sqrt{2}$ of $\mathbb{R}$ is a free abelian group of rank 2 (because $\sqrt{2}$ is not rational), but it is *not* a lattice in $\mathbb{R}$.

We shall need another criterion for a subgroup Λ of V to be a lattice. The choice of a basis for V determines an isomorphism of V with $\mathbb{R}^n$, and hence a topology on V; the topology is independent of the basis, because any linear automorphism of $\mathbb{R}^n$ is a homeomorphism. A subgroup Λ of V is said to be *discrete* if it is discrete in the induced topology. Recall that a topological space is discrete if its points (hence all subsets) are open. Thus to say that Λ is discrete means that every point α of Λ has a neighbourhood U in V such that $U \cap \Lambda = \{\alpha\}$.

LEMMA 4.13. *The following conditions on a subgroup Λ of a finite-dimensional real vector space V are equivalent:*

(a) *Λ is a discrete subgroup;*
(b) *there is an open subset U of V such that $U \cap \Lambda = \{0\}$;*
(c) *each compact subset of V intersects Λ in a finite set;*
(d) *each bounded subset of V intersects Λ in a finite set.*

PROOF. (a) $\iff$ (b). Obviously (a) implies (b). For the converse, note that the translation map $x \mapsto \alpha + x \colon V \to V$ is a homeomorphism, and that therefore if U is a neighbourhood of 0 such that $U \cap \Lambda = \{0\}$, then $\alpha + U$ is a neighbourhood of α such that $(\alpha + U) \cap \Lambda = \{\alpha\}$.

(a)$\Rightarrow$(c). Condition (a) says that Λ is a discrete space for the induced topology. Hence, if C is compact, then $C \cap \Lambda$ is both discrete and compact, and therefore must be finite.

(c)$\Rightarrow$(d). The closure of a bounded set in $\mathbb{R}^n$ (hence in V) is compact, and so this is obvious.

(d)$\Rightarrow$(b). Let U be a bounded open neighbourhood of 0. Then $S = U \cap \Lambda \smallsetminus \{0\}$ is finite and hence closed, and so $U \smallsetminus S$ is an open neighbourhood of $\{0\}$ such that $(U \smallsetminus S) \cap \Lambda = \{0\}$. □

PROPOSITION 4.14. *A subgroup Λ of V is a lattice if and only if it is discrete.*

PROOF. Clearly, a lattice is discrete. For the converse, let Λ be a discrete subgroup of V, and let $e_1, ..., e_r$ be a maximal $\mathbb{R}$-linearly independent subset of Λ. We shall use induction on r.

If $r = 0$, $\Lambda = 0$, and there is nothing to prove.

If $r = 1$, then every $\alpha \in \Lambda$ can be written $\alpha = ae_1$, some $a \in \mathbb{R}$. Because Λ is discrete, $\{ae_1 \mid |a| < M\} \cap \Lambda$ is finite, and so there is an $f_1 \in \Lambda$ such that, when we write $f_1 = ae_1$, a attains its minimum value > 0. I claim $\Lambda = \mathbb{Z}f_1$. If $\alpha \in \Lambda$, $\alpha \notin \mathbb{Z}f_1$, then we can find an integer m such that $\alpha - mf_1 = bf_1$ with $0 < b < 1$; but then $(\alpha - mf_1) = bf_1 = abe_1$, and $0 < ab < a$, which contradicts our choice of f_1.

Let $\Lambda' = \Lambda \cap (\mathbb{R}e_1 + \cdots + \mathbb{R}e_{r-1})$. Clearly this is a discrete subgroup of the vector space $V' =_{df} \mathbb{R}e_1 + \cdots + \mathbb{R}e_{r-1}$ and so, by induction, $\Lambda' = \mathbb{Z}f_1 + \cdots + \mathbb{Z}f_{r-1}$ for some f_i that are linearly independent over $\mathbb{R}$ (and hence also form a basis for V'). Every $\alpha \in \Lambda$ can be written uniquely

$$\alpha = a_1 f_1 + \cdots + a_{r-1} f_{r-1} + a e_r, \quad a_i, a \in \mathbb{R}.$$

Let $\varphi\colon \Lambda \to \mathbb{R}$ be the map $\alpha \mapsto a$, and let $\Lambda'' = \mathrm{Im}(\varphi)$. Note that a is also the image of

$$(a_1 - [a_1])f_1 + \cdots + (a_{r-1} - [a_{r-1}])f_{r-1} + a e_r, \quad [*] = \text{integer part},$$

and so each element $a \in \Lambda''$ in a bounded set, say with $0 \le |a| < M$, is the image of an element of Λ in a bounded set,

$$0 \le a_i < 1, \quad i = 1, \dots, r-1, \quad |a| < M.$$

Thus there are only finitely many such a's, and so Λ'' is a lattice in $\mathbb{R}$, say $\Lambda'' = \mathbb{Z} \cdot \varphi(f_r)$, $f_r \in \Lambda$.

Let $\alpha \in \Lambda$. Then $\varphi(\alpha) = a\varphi(f_r)$ for some $a \in \mathbb{Z}$, and $\varphi(\alpha - a f_r) = 0$. Therefore $\alpha - a f_r \in \Lambda'$, and so it can be written

$$\alpha - a f_r = a_1 f_1 + \cdots + a_{r-1} f_{r-1}, \quad a_i \in \mathbb{Z}.$$

Hence

$$\alpha = a_1 f_1 + \cdots + a_{r-1} f_{r-1} + a f_r, \quad a_i, a \in \mathbb{Z},$$

which proves that $\Lambda = \sum \mathbb{Z} f_i$. □

Let V be a real vector space of dimension n, and let Λ be a full lattice in V, say $\Lambda = \sum \mathbb{Z} e_i$. For any $\lambda_0 \in \Lambda$, let

$$D = \{\lambda_0 + \textstyle\sum a_i e_i \mid 0 \le a_i < 1\}.$$

Such a set is called a *fundamental parallelopiped* for Λ. The shape of the parallelopiped depends on the choice of the basis (e_i), but if we fix the basis and vary $\lambda_0 \in \Lambda$, then the parallelopipeds cover $\mathbb{R}^n$ without overlaps.

Remark 4.15. (a) For a fundamental parallelopiped D of a full lattice

$$\Lambda = \mathbb{Z}f_1 + \cdots + \mathbb{Z}f_n$$

in $\mathbb{R}^n$, the volume of D

$$\mu(D) = |\det(f_1, \cdots, f_n)|.$$

(See any good book on calculus.) If also

$$\Lambda = \mathbb{Z}f_1' + \mathbb{Z}f_2' + \cdots + \mathbb{Z}f_n',$$

then the determinant of the matrix relating $\{f_i\}$ and $\{f_i'\}$ has determinant ± 1, and so the volume of the fundamental parallelopiped doesn't depend on the choice of the basis for Λ.

(b) When $\Lambda \supset \Lambda'$ are two full lattices $\mathbb{R}^n$, we can choose bases $\{e_i\}$ and $\{f_i\}$ for Λ and Λ' such that $f_i = m_i e_i$ with m_i a positive integer. With this choice of bases,

the fundamental parallelopiped D of Λ is a disjoint union of $(\Lambda : \Lambda')$ fundamental parallelopipeds D' of Λ'. Hence

$$\frac{\mu(D')}{\mu(D)} = (\Lambda : \Lambda') \qquad (*).$$

As we noted above, the choice of a basis for V determines an isomorphism $V \approx \mathbb{R}^n$, and hence a measure μ on V. This measure is translation invariant (because the Lebesgue measure on $\mathbb{R}^n$ is translation invariant), and well-defined up to multiplication by a nonzero constant (depending on the choice of the basis) [12]. Thus the ratio of the measures of two sets is well-defined, and the equation (*) holds for two full lattices $\Lambda \supset \Lambda'$ in V.

THEOREM 4.16. *Let D_0 be a fundamental parallelopiped for a full lattice in V, and let S be a measurable subset in V. If $\mu(S) > \mu(D_0)$, then S contains distinct points α and β such that $\beta - \alpha \in \Lambda$.*

PROOF. The set $S \cap D$ is measurable for all fundamental parallelopipeds D, and

$$\mu(S) = \sum \mu(S \cap D)$$

(sum over translates of D by elements of Λ). For each D, a (unique) translate of $S \cap D$ by an element of Λ will be a subset of D_0. Since $\mu(S) > \mu(D_0)$, at least two of these sets will overlap, i.e., there exist elements $\alpha, \beta \in S$ such that

$$\alpha - \lambda = \beta - \lambda', \quad \text{some } \lambda, \lambda' \in \Lambda.$$

Then $\beta - \alpha \in \Lambda$. □

REMARK 4.17. In the language of differential geometry, the theorem can be given a more geometric statement. Let $M = V/\Lambda$; it is an n-dimensional torus. The measure μ on V defines a measure on M for which M has measure $\mu(M) = \mu(D)$. The theorem says that if $\mu(S) > \mu(M)$, then the restriction of the quotient map $V \to M$ to S can't be injective.

Let T be a set such that

$$\alpha,\beta \in T \Rightarrow \frac{1}{2}(\alpha - \beta) \in T, \qquad (*)$$

and let $S = \frac{1}{2}T$. Then T contains the difference of any two points of S, and so T will contain a point of Λ other than the origin whenever

$$\mu(D) < \mu(\frac{1}{2}T) = 2^{-n}\mu(T),$$

i.e., whenever

$$\mu(T) > 2^n\mu(D).$$

We say that a set T is *convex* if, with any two points, it contains the line joining the two points, and that T is *symmetric in the origin* if $\alpha \in T$ implies $-\alpha \in T$. A convex set, symmetric in the origin, obviously satisfies (*), and so it will contain a point of $\Lambda \setminus \{0\}$ if its volume is greater than $2^n\mu(D)$.

[12] The experts will recognize μ as being a *Haar measure* on V.

THEOREM 4.18 (Minkowski's). *Let T be a subset of V that is compact, convex, and symmetric in the origin. If*

$$\mu(T) \geq 2^n \mu(D)$$

then T contains a point of the lattice other than the origin.

PROOF. Replace T with $(1+\varepsilon)T$, $\varepsilon > 0$. Then $\mu((1+\varepsilon)T) = (1+\varepsilon)^n \mu(T) > 2^n \mu(D)$, and so $(1+\varepsilon)T$ contains a point of Λ other than the origin (see the preceding remark). It will contain only finitely many such points because Λ is discrete and $(1+\varepsilon)T$ is compact. Because T is closed

$$T = \cap_{\varepsilon>0}(1+\varepsilon)T.$$

If none of the points of $\Lambda \cap (1+\varepsilon)T$ is in T, we will be able to shrink $(1+\varepsilon)T$ (keeping $\varepsilon > 0$) so that it contains no point of Λ other than the origin—which is a contradiction. □

REMARK 4.19. Theorem 4.18 was discovered by Minkowski in 1896. Although it is almost trivial to prove, it has lots of nontrivial consequences, and was the starting point for the branch of number theory (now almost defunct) called the "geometry of numbers". We give one immediate application of it to prove that every positive integer is a sum of four squares of integers.

From the identity

$$\begin{aligned}(a^2+b^2+\quad & c^2+d^2)(A^2+B^2+C^2+D^2) = \\ & (aA-bB-cC-dD)^2 + (aB+bA+cD-dC)^2 + \\ & (aC-bD+cA+dB)^2 + (aD+bC-cB+dA)^2,\end{aligned}$$

we see that it suffices to prove that a prime p is a sum of four squares.

Since

$$2 = 1^2+1^2+0^2+0^2,$$

we can suppose that p is odd. I claim that the congruence

$$m^2+n^2+1 \equiv 0 \mod p$$

has a solution in $\mathbb{Z}$. As m runs through $0, 1, \ldots, p-1$, m^2 takes exactly $(p+1)/2$ distinct values modulo p, and similarly for $-1-n^2$. For the congruence to have no solution, all these values, $p+1$ in total, must be distinct, but this is impossible.

Fix a solution m, n to the congruence, and consider the lattice $\Lambda \subset \mathbb{Z}^4$ consisting of (a,b,c,d) such that

$$c \equiv ma+nb, \quad d \equiv mb-na \mod p.$$

Then $\mathbb{Z}^4 \supset \Lambda \supset p\mathbb{Z}^4$ and $\Lambda/p\mathbb{Z}^4$ is a 2-dimensional subspace of $\mathbb{F}_p^4$ (the a and b can be arbitrary mod p, but then c and d are determined). Hence Λ has index p^2 in $\mathbb{Z}^4$, and so the volume of a fundamental parallelopiped is p^2. Let T be a closed ball of radius r centered at the origin. Then T has volume $\pi^2 r^4/2$, and so if we choose $r^2 > 1.9p$ say, then

$$\mu(T) > 16\mu(D).$$

According to Minkowski's theorem, there is a point $(a,b,c,d) \in (\Lambda \setminus \{0\}) \cap T$. Because $(a,b,c,d) \in \Lambda$,

$$a^2+b^2+c^2+d^2 \equiv a^2(1+m^2+n^2)+b^2(1+m^2+n^2) \equiv 0 \mod p,$$

and because $(a,b,c,d) \in T$,

$$a^2+b^2+c^2+d^2 < 2p.$$

As $a^2+b^2+c^2+d^2$ is a positive integer, these conditions imply that it equals p.

This result was stated by Fermat. Euler tried to prove it over a period of 40 years, and Lagrange succeeded in 1770.

Some calculus. Let V be a finite-dimensional real vector space. A *norm* on V is a function $\|\cdot\|: V \to \mathbb{R}$ such that

(4.20.1) for all $\mathbf{x} \in V$, $\|\mathbf{x}\| \geq 0$, and $\|\mathbf{x}\| = 0 \iff \mathbf{x} = 0$;
(4.20.2) for $r \in \mathbb{R}$ and $\mathbf{x} \in V$, $\|r\mathbf{x}\| = |r|\|\mathbf{x}\|$;
(4.20.3) (triangle law) for $\mathbf{x}, \mathbf{y} \in V$, $\|\mathbf{x}+\mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|$.

Let $V = \mathbb{R}^r \times \mathbb{C}^s$ — it is a real vector space of dimension $n = r+2s$. Define a norm on V by

$$\|\mathbf{x}\| = \sum_{i=1}^{r} |x_i| + 2\sum_{i=r+1}^{s} |z_i|$$

if $\mathbf{x} = (x_1, ..., x_r, z_{r+1}, ..., z_{r+s})$.

LEMMA 4.21. *For any real number $t > 0$, let*

$$X(t) = \{\mathbf{x} \in V \mid \|\mathbf{x}\| \leq t\}.$$

Then

$$\mu(X(t)) = 2^r(\pi/2)^s t^n/n!.$$

PROOF. Since $X(t)$ is symmetric with respect to the r real axes, we have

$$\mu(X(t)) = 2^r \cdot \mu(Y(t))$$

where $Y(t) = \{\mathbf{x} \mid \|\mathbf{x}\| \leq t,\ x_1, ..., x_r \geq 0\}$. For the complex variables, we make the change of variable

$$z_j = x_j + iy_j = \frac{1}{2}\rho_j(\cos\theta_j + i\sin\theta_j).$$

The Jacobian of this change of variables is $\rho_j/4$. After integrating over the θ_j, for $0 \leq \theta_j \leq 2\pi$, we find that

$$\mu(X(t)) = 2^r \cdot 4^{-s} \cdot (2\pi)^s \int_Z \rho_{r+1} \cdots \rho_{r+s} dx_1 \cdots dx_r d\rho_{r+1} \cdots d\rho_{r+s}$$

where

$$Z = \{(\mathbf{x}, \rho) \in \mathbb{R}^{r+s} \mid x_i, \rho_i \geq 0, \quad \sum x_i + \sum \rho_i \leq t\}.$$

The result now follows from the next lemma by taking: $m = r+s$; $a_i = 0$, $1 \leq i \leq r$; $a_i = 1$, $r+1 \leq i \leq m$; for then

$$\mu(X(t)) = 2^r \cdot 4^{-s} \cdot (2\pi)^s \cdot t^n/n!$$

as required. □

LEMMA 4.22. *For $a_i > 0 \in \mathbb{R}$, let*

$$I(a_1, ..., a_m, t) = \int_{Z(t)} x_1^{a_1} \cdots x_m^{a_m} dx_1 \cdots dx_m,$$

where $Z(t) = \{x \in \mathbb{R}^m \mid x_i \geq 0, \quad \sum x_i \leq t\}$. Then

$$I(a_1, \ldots, a_m; t) = t^{\sum a_i + m} \cdot \frac{\Gamma(a_1 + 1) \cdots \Gamma(a_m + 1)}{\Gamma(a_1 + \cdots + a_m + m + 1)}.$$

PROOF. Recall that, by definition, (e.g., Widder, D., Advanced Calculus, 1961, Chapter 11),

$$\Gamma(x) = \int_{0+}^{\infty} e^{-t} t^{x-1} dt.$$

It takes the value $\Gamma(n) = (n-1)!$ for n a nonnegative integer.

By making the change of variables $x_i' = tx_i$ in I, we see that

$$I(a_1, \ldots, a_m; t) = t^{\sum a_i + m} I(a_1, \ldots, a_m; 1).$$

Therefore it suffices to prove the formula for $t = 1$. We prove this case by induction on m. First, we have

$$I(a_1; 1) = \int_0^1 x_1^{a_1} dx_1 = \frac{1}{a_1 + 1} = \frac{\Gamma(a_1 + 1)}{\Gamma(a_1 + 2)}.$$

Let

$$Z(x_m)' = \{\mathbf{x} \in \mathbb{R}^{m-1} \mid x_i \geq 0, \quad \sum x_i \leq 1 - x_m\}.$$

Then

$$\begin{aligned}
I(a_1, ..., a_m; 1) = & \quad \int_0^1 x_m^{a_m} \left(\int_{Z(x_m)'} x_1^{a_1} \cdots x_{m-1}^{a_{m-1}} dx_1 \cdots dx_{m-1} \right) dx_m, \\
= & \quad \int_0^1 x_m^{a_m} I(a_1, ..., a_{m-1}; 1 - x_m) dx_m \\
= & \quad I(a_1, ..., a_{m-1}; 1) \int_0^1 x_m^{a_m} (1 - x_m)^{\sum a_i + m - 1} dx_m \\
= & \quad I(a_1, ..., a_{m-1}; 1) \frac{\Gamma(a_m + 1)\Gamma(a_1 + \cdots + a_{m-1} + m)}{\Gamma(a_1 + \cdots + a_m + m + 1)}.
\end{aligned}$$

In the last step, we used the standard formula

$$\int_0^1 x^{m-1} (1 - x)^{n-1} dx = B(m, n) = \frac{\Gamma(m)\Gamma(n)}{\Gamma(m + n)}.$$

□

EXAMPLE 4.23. (a) Case $r = 2$, $s = 0$. Then $X(t)$ is defined by $|x| + |y| \leq t$. It is a square of side $\sqrt{2}t$, and so $\mu(X(t)) = 2t^2$.

(b) Case $r = 0$, $s = 1$. Then $X(t)$ is the circle of radius $t/2$, which has area $\pi t^2/4$.

LEMMA 4.24. *Let $a_1, \dots, a_n$ be positive real numbers. Then*

$$(\prod a_i)^{1/n} \le (\sum a_i)/n;$$

equivalently,

$$\prod a_i \le (\sum a_i)^n/n^n.$$

(The geometric mean is less than or equal to the arithmetic mean.)

PROOF. See any good course on advanced calculus. □

Finiteness of the class number. Let K be number field of degree n over $\mathbb{Q}$. Suppose that K has r real embeddings $\{\sigma_1, \dots, \sigma_r\}$ and 2s complex embedding $\{\sigma_{r+1}, \bar{\sigma}_{r+1}, \dots, \sigma_{r+s}, \bar{\sigma}_{r+s}\}$. Thus $n = r + 2s$. We have an embedding

$$\sigma : K \hookrightarrow \mathbb{R}^r \times \mathbb{C}^s, \quad \alpha \mapsto (\sigma_1\alpha, \dots, \sigma_{r+s}\alpha).$$

We identify $V \stackrel{\text{df}}{=} \mathbb{R}^r \times \mathbb{C}^s$ with $\mathbb{R}^n$ using the basis $\{1, i\}$ for $\mathbb{C}$.

PROPOSITION 4.25. *Let $\mathfrak{a}$ be an ideal in $\mathcal{O}_K$; then $\sigma(\mathfrak{a})$ is a full lattice in V, and the volume of a fundamental parallelopiped of $\sigma(\mathfrak{a})$ is $2^{-s} \cdot \mathbb{N}\mathfrak{a} \cdot |\Delta_K|^{\frac{1}{2}}$.*

PROOF. Let $\alpha_1, \dots, \alpha_n$ be a basis for $\mathfrak{a}$ as a $\mathbb{Z}$-module. To prove that $\sigma(\mathfrak{a})$ is a lattice we show that the vectors $\sigma(\alpha_1), \dots, \sigma(\alpha_n)$ are linearly independent, and we prove this by showing that the matrix A, whose i^{th} row is

$$(\sigma_1(\alpha_i), \dots, \sigma_r(\alpha_i), \Re(\sigma_{r+1}\alpha_i), \Im(\sigma_{r+1}\alpha_i), \dots)$$

has nonzero determinant.

First consider the matrix B whose i^{th} row is

$$(\sigma_1(\alpha_i), \dots, \sigma_r(\alpha_i), \sigma_{r+1}(\alpha_i), \overline{\sigma_{r+1}(\alpha_i)}, \dots, \overline{\sigma_{r+s}(\alpha_i)}).$$

We saw in (2.25) that $\det(B)^2 = \operatorname{disc}(\alpha_1, \dots, \alpha_n) \neq 0$.

What is the relation between the determinants of A and B? Add column $r+2$ in B to column $r+1$, and then subtract $1/2$ column $r+1$ from column $r+2$. This gives us $2\Re(\sigma_{r+1}(\alpha_i))$ in column $r+1$ and $-i\Im(\sigma_{r+1}(\alpha_i))$ in column $r+2$. Repeat for the other pairs of columns. These column operations don't change the determinant of B, and so

$$\det(B) = (-2i)^s \det(A),$$

or

$$\det(A) = (-2i)^{-s} \det(B) = (-2i)^{-s} \operatorname{disc}(\alpha_1, \dots, \alpha_n)^{1/2} \neq 0.$$

Thus $\sigma(\mathfrak{a})$ is a lattice in V.

Since $\sigma(\mathfrak{a}) = \sum_{i=1}^n \mathbb{Z}\sigma(\alpha_i)$, the volume of a fundamental parallelopiped D for $\sigma(\mathfrak{a})$ is $|\det(A)|$, and from (2.24) we know that

$$|\operatorname{disc}(\alpha_1, \dots, \alpha_n)| = (\mathcal{O}_K : \mathfrak{a})^2 \cdot |\operatorname{disc}(\mathcal{O}_K/\mathbb{Z})|.$$

Hence

$$\mu(D) = 2^{-s} \cdot |\operatorname{disc}(\alpha_1, \dots, \alpha_n)|^{\frac{1}{2}} = 2^{-s} \cdot \mathbb{N}\mathfrak{a} \cdot |\Delta_K|^{\frac{1}{2}}.$$

□

PROPOSITION 4.26. *Let $\mathfrak{a}$ be an ideal in $\mathcal{O}_K$. Then $\mathfrak{a}$ contains a nonzero element α of K with*

$$|\operatorname{Nm}(\alpha)| \leq B_K \cdot \mathbb{N}\mathfrak{a} = \left(\frac{4}{\pi}\right)^s \frac{n!}{n^n} \mathbb{N}\mathfrak{a}|\Delta|^{\frac{1}{2}}.$$

PROOF. Let $X(t)$ be as in (4.21), and let D be a fundamental domain for the lattice $\sigma(\mathfrak{a})$. The set $X(t)$ is compact convex and symmetric in the origin, and so, when we choose t so large that $\mu(X(t)) \geq 2^n \cdot \mu(D)$, Minkowski's Theorem shows that $X(t)$ contains a point $\sigma(\alpha) \neq 0$ of $\sigma(\mathfrak{a})$. For this $\alpha \in \mathfrak{a}$,

$$\begin{aligned} |\operatorname{Nm}(\alpha)| &= |\sigma_1(\alpha)| \cdots |\sigma_r(\alpha)||\sigma_{r+1}(\alpha)|^2 \cdots |\sigma_{r+s}(\alpha)|^2 \\ &\leq (\sum |\sigma_i \alpha| + \sum 2|\sigma_i \alpha|)^n / n^n \text{ (by 4.24)} \\ &\leq t^n / n^n. \end{aligned}$$

In order to have $\mu(X(t)) \geq 2^n \cdot \mu(D)$, we need

$$2^r (\pi/2)^s t^n / n! \geq 2^n \cdot 2^{-s} \cdot \mathbb{N}\mathfrak{a} \cdot |\Delta_K|^{\frac{1}{2}},$$

i.e.,

$$t^n \geq n! \cdot \frac{2^{n-r}}{\pi^s} \cdot \mathbb{N}\mathfrak{a} \cdot |\Delta_K|^{\frac{1}{2}}.$$

When we take t^n to equal the expression on the right, we find that

$$|\operatorname{Nm}(\alpha)| \leq \frac{n!}{n^n} \cdot \frac{2^{n-r}}{\pi^s} \cdot \mathbb{N}\mathfrak{a} \cdot |\Delta_K|^{\frac{1}{2}}.$$

As $n - r = 2s$, this is the required formula. □

PROOF OF THEOREM 4.3. Let $\mathfrak{c}$ be a fractional ideal in K — we have to show that the class of $\mathfrak{c}$ in the ideal class group is represented by an integral ideal $\mathfrak{a}$ with

$$\mathbb{N}\mathfrak{a} \leq B_K \stackrel{\text{df}}{=} \frac{n!}{n^n} \left(\frac{4}{\pi}\right)^s |\Delta_K|^{\frac{1}{2}}.$$

For some $d \in K^\times$, $d\mathfrak{c}^{-1}$ is an integral ideal, say $(d) \cdot \mathfrak{c}^{-1} = \mathfrak{b}$. According to the result just proved, there is a $\beta \in \mathfrak{b}$, $\beta \neq 0$, with

$$|\operatorname{Nm}(\beta)| \leq B_K \cdot \mathbb{N}\mathfrak{b}.$$

Now $\beta\mathcal{O}_K \subset \mathfrak{b} \Rightarrow \beta\mathcal{O}_K = \mathfrak{a}\mathfrak{b}$ with $\mathfrak{a}$ integral, and $\mathfrak{a} \sim \mathfrak{b}^{-1} \sim \mathfrak{c}$. Moreover,

$$\mathbb{N}\mathfrak{a} \cdot \mathbb{N}\mathfrak{b} = |\operatorname{Nm}_{L/K} \beta| \leq B_K \cdot \mathbb{N}\mathfrak{b}.$$

On cancelling $\mathbb{N}\mathfrak{b}$, we find that $\mathbb{N}\mathfrak{a} \leq B$. □

REMARK 4.27. Proposition 4.26 can be useful in deciding whether an integral ideal is principal.

Binary quadratic forms. The first person to consider class numbers (implicitly) was Gauss. Rather than working with ideals, which hadn't been defined then, he worked with binary quadratic forms.

By a *binary quadratic form* we mean an expression of the form

$$Q(X,Y) = aX^2 + bXY + cY^2.$$

We call the form *integral* if $Q(m,n)$ is an integer whenever m and n are integers, or, equivalently, if $a, b, c \in \mathbb{Z}$. The *discriminant* of Q is

$$d_Q = b^2 - 4ac.$$

A form is said to be *nondegenerate* if its discriminant is nonzero. Two integral binary quadratic forms Q and Q' are said to be *equivalent* if there exists a matrix $A = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z})$ such that

$$Q'(X,Y) = Q(\alpha X + \beta Y, \gamma X + \delta Y).$$

Clearly, equivalent forms have the same discriminant, but there exist inequivalent forms with the same discriminant. The question considered by Gauss was to try to describe the set of equivalence classes of forms with a fixed discriminant. As we shall explain, this question can be interpreted in terms of ideals.

Let $d \neq 1$ be a square-free integer, let $K = \mathbb{Q}[\sqrt{d}]$, and let $d_K = \mathrm{disc}(\mathcal{O}_K/\mathbb{Z})$. Define the *norm form* q_K by

$$q_K(X,Y) = \mathrm{Nm}_{K/\mathbb{Q}}(X + Y\sqrt{d}) = X^2 - dY^2, \quad \text{if } d \equiv 2,3 \mod 4$$

or

$$q_K(X,Y) = \mathrm{Nm}_{K/\mathbb{Q}}(X + Y\frac{1+\sqrt{d}}{2}) = X^2 + XY + \frac{1-d}{4}Y^2, \quad \text{if } d \equiv 1 \mod 4.$$

In both cases q_K has discriminant d_K ($= 4d$ or d).

In general, if Q is an integral binary quadratic form, then $d_Q = d_K f^2$, some integer f, where $K = \mathbb{Q}[\sqrt{d_Q}]$. Moreover, if $d_Q = d_K$, then Q is primitive, i.e., $\gcd(a,b,c) = 1$.

Fix a field $K = \mathbb{Q}[\sqrt{d}]$ and an embedding $K \hookrightarrow \mathbb{C}$. We choose $\sqrt{d}$ to be positive if $d > 0$ and to have positive imaginary part if d is negative. Set $\sqrt{d_K} = 2\sqrt{d}$ or $\sqrt{d}$. Write $\mathrm{Gal}(K/\mathbb{Q}) = \{1, \sigma\}$. If $d < 0$, define $\mathrm{Cl}^+(K) = \mathrm{Cl}(K)$ (usual class group of K) and if $d > 0$, define

$$\mathrm{Cl}^+(K) = \mathrm{Id}(K)/\mathrm{P}^+(K)$$

where $P^+(K)$ is the group of principal ideals of the form (α) with $\alpha > 0$ under every embedding of K into $\mathbb{R}$.

Let $\mathfrak{a}$ be a fractional ideal in K, and let a_1, a_2 be a basis for $\mathfrak{a}$ as a $\mathbb{Z}$-module. From (2.24) we know that

$$\begin{vmatrix} a_1 & a_2 \\ \sigma a_1 & \sigma a_2 \end{vmatrix}^2 = d_K \mathbb{N}\mathfrak{a}^2.$$

After possibly reordering the pair a_1, a_2 we will have

$$\begin{vmatrix} a_1 & a_2 \\ \sigma a_1 & \sigma a_2 \end{vmatrix} = \sqrt{d_K}\mathbb{N}\mathfrak{a}.$$

For such a pair, define

$$Q_{a_1,a_2}(X,Y) = \mathbb{N}\mathfrak{a}^{-1} \cdot \mathrm{Nm}_{K/\mathbb{Q}}(a_1X + a_2Y).$$

This is an integral binary quadratic form with discriminant d_K.

THEOREM 4.28. *The equivalence class of $Q_{a_1,a_2}(X,Y)$ depends only on the image of $\mathfrak{a}$ in $Cl^+(K)$; moreover, the map sending $\mathfrak{a}$ to the equivalence class of Q_{a_1,a_2} defines a bijection from $Cl^+(K)$ to the set of equivalence classes of integral binary quadratic forms with discriminant d_K.*

PROOF. See Fröhlich and Taylor 1991, VII.2 (and elsewhere). □

In particular, the set of equivalence classes is finite, and has the structure of an abelian group. This was known to Gauss, even though groups had not yet been defined. (Gauss even knew it was a direct sum of cyclic groups.)

REMARK 4.29. Write h_d for the class number of $\mathbb{Q}[\sqrt{d}]$, d a square-free integer $\neq 1$. In modern terminology, Gauss conjectured that, for a fixed h, there are only finitely many negative d such that $h_d = h$. (Actually, because of a difference of terminology, this is not quite what Gauss conjectured.)

In 1935, Siegel showed that, for every $\varepsilon > 0$, there exists a constant $c > 0$ such that

$$h_d > c|d|^{\frac{1}{2}-\varepsilon}, \quad d < 0.$$

This proves Gauss's conjecture. Unfortunately, the c in Siegel's theorem is not effectively computable, and so Siegel's theorem gives no way of computing the d's for a given h.

In 1951, Tatuzawa showed that Siegel's theorem is true with an effectively computable c except for at most one exceptional d.

It is easy to show that $h_d = 1$ for $-d = 1, 2, 3, 7, 11, 19, 43, 67, 163$ (exercise!). Thus in 1951 it was known that there exist these 9 quadratic imaginary number fields with class number 1, and possibly 1 more.

In 1952 Heegner proved that there was no 10^{th} such field, but his proof was not recognized to be correct for many years.

More recently (1983), Goldfeld, Gross, and Zagier showed, using completely different methods from Siegel, that there is an effective procedure for finding all $d < 0$ with h_d equal to a given h. For an expository article on this, see Goldfeld, Bull. Amer. Math. Soc. 13 (1985), 23–37.

By way of contrast, it is conjectured that there are infinitely many real quadratic fields with class number 1, but this has not been proved.

There are tables of class numbers at the back of Borevich and Shafarevich 1966 (and elsewhere).

5. The Unit Theorem

In this section we prove the second main theorem of the course.

Statement of the theorem. Recall that a finitely generated abelian group A is isomorphic to $A_{\text{tors}} \oplus \mathbb{Z}^r$ for some r where A_{tors} is the (finite) subgroup of torsion elements of A (i.e., of elements of finite order). The number r is uniquely determined by A, and is called the *rank* of A.

Theorem 5.1. *The group of units in a number field K is finitely generated with rank equal to $r + s - 1$.*

The theorem is usually referred to as the "Dirichlet Unit Theorem" although Dirichlet in fact proved it for rings of the form $\mathbb{Z}[\alpha]$ rather than $\mathcal{O}_K$.

Write U_K $(= \mathcal{O}_K^\times)$ for the group of units in K. The torsion subgroup of U_K is (obviously) the group $\mu(K)$ of roots of 1 in K.

A set of units $u_1, \dots, u_{r+s-1}$ is called a *fundamental system of units* if it forms a basis for U_K modulo torsion, i.e., if every unit u can be written uniquely in the form

$$u = \zeta u_1^{m_1} \cdots u_{r+s-1}^{m_{r+s-1}}, \quad \zeta \in \mu(K), \quad m_i \in \mathbb{Z}.$$

The theorem implies that $\mu(K)$ is finite (and hence cyclic). This can be proved directly. In §7, we shall see that, if ζ_m is a primitive m^{th} root of 1, then $\mathbb{Q}[\zeta]$ is a Galois extension of $\mathbb{Q}$ with Galois group isomorphic to $(\mathbb{Z}/m\mathbb{Z})^\times$. Note that

$$\begin{aligned} \#(\mathbb{Z}/m\mathbb{Z})^\times &= \#\{n \mid 0 \le n \le m-1, \gcd(n,m) = 1\} \\ &\overset{\text{df}}{=} \varphi(m) \quad \text{(Euler } \varphi\text{-function)} \end{aligned}$$

and

$$\varphi(m) = \prod \varphi(p_i^{r_i}) = \prod p_i^{r_i - 1}(p_i - 1), \quad \text{for } m = \prod p_i^{r_i},$$

which increases with m. Since

$$\zeta_m \in K \Rightarrow \mathbb{Q}[\zeta_m] \subset K \Rightarrow \varphi(m) | [K : \mathbb{Q}],$$

this implies that $\mu(K)$ is finite.

Lemma 5.2. *An element $\alpha \in K$ is a unit if and only if $\alpha \in \mathcal{O}_K$ and $\operatorname{Nm}_{K/\mathbb{Q}} \alpha = \pm 1$.*

Proof. If α is a unit, then there is a $\beta \in \mathcal{O}_K$ such that $\alpha\beta = 1$, and then $\operatorname{Nm}(\alpha)$ and $\operatorname{Nm}(\beta)$ lie in $\mathbb{Z}$ and $1 = \operatorname{Nm}(\alpha\beta) = \operatorname{Nm}(\alpha) \cdot \operatorname{Nm}(\beta)$. Hence $\operatorname{Nm} \alpha \in \mathbb{Z}^\times = \{\pm 1\}$.

For the converse, fix an embedding σ_0 of K into $\mathbb{C}$, and use it to identify K with a subfield of $\mathbb{C}$. Recall (2.19) that

$$\operatorname{Nm} \alpha = \prod_\sigma \sigma\alpha = \alpha \cdot \prod_{\sigma \ne \sigma_0} \sigma\alpha, \quad \sigma \colon K \hookrightarrow \mathbb{C}.$$

Let $\beta = \prod_{\sigma \ne \sigma_0} \sigma\alpha$. If $\alpha \in \mathcal{O}_K$ then β is an algebraic integer. If $\operatorname{Nm} \alpha = \pm 1$, then $\beta = \pm\alpha^{-1}$ and so belongs to K. Therefore, if α satisfies both conditions, it has an inverse $\pm\beta$ in $\mathcal{O}_K$, and so is a unit. □

For all real fields, i.e., fields with an embedding into $\mathbb{R}$, $\mu(K) = \{\pm 1\}$; for "most" nonreal fields, this is also true.

EXAMPLE 5.3. Let K be a quadratic field $\mathbb{Q}[\sqrt{d}]$. Then $\mathcal{O}_K = \{m+n\sqrt{d} \mid m,n \in \mathbb{Z}\}$ or $\{m+n(1+\sqrt{d})/2 \mid m,n \in \mathbb{Z}\}$. The units in K are given by the solutions of the equation:

$$m^2 - n^2 d = \pm 1, \text{ or } (2m+n)^2 - dn^2 = 4.$$

When $d<0$, these equations (obviously) have only finitely many solutions, and so $U_K = \mu(K)$. Note that ζ_m lies in a quadratic field if and only if $\varphi(m) \le 2$. This only happens for m dividing 4 or 6. Thus $\mu(K) = \{\pm 1\}$ except for the following fields:

$\mathbb{Q}[i]$, $\mu(K) = \{\pm 1, \pm i\}$;

$\mathbb{Q}[\sqrt{-3}]$, $\mu(K) = \{\pm 1, \pm\rho, \pm\rho^2\}$, with $\rho = (1+\sqrt{-3})/2)$.

When $d>0$, the theorem shows that there are infinitely many solutions, and that $U_K = \pm u^{\mathbb{Z}}$ for some element u (called the *fundamental unit*). As H. Cohn (A Classical Invitation...) puts it, "the actual computation of quadratic units lies in the realm of popularized elementary number theory, including devices such as continued fractions." The method is surprisingly effective, and yields some remarkably large numbers — see later.

EXAMPLE 5.4. Let $K = \mathbb{Q}[\alpha]$, where α is a root of $X^3 + 10X + 1$. We know that the discriminant $\Delta_K = -4027$. Since $\operatorname{sign}(\Delta_K) = (-1)^s$ and $r+2s = 3$, we must have $r = 1 = s$. From its minimum equation, we see that $\operatorname{Nm}\alpha = -1$, and so α is a unit. Later we shall show that α is a fundamental unit, and so $U_K = \{\pm\alpha^m \mid m \in \mathbb{Z}\}$.

Proof that U_K is finitely generated. We first need a simple lemma.

LEMMA 5.5. *For any integers m and M, the set of all algebraic integers α such that*

(i) the degree of α is $\le m$, and

(ii) $|\alpha'| < M$ for all conjugates α' of α

is finite.

PROOF. The first condition says that α is a root of a monic irreducible polynomial of degree $\le m$, and the second condition implies that the coefficients of the polynomial are bounded. Since the coefficients are integers, there are only finitely many such polynomials, and hence only finitely many α's. □

Recall that we previously considered the map

$$\sigma : K \to \mathbb{R}^r \times \mathbb{C}^s, \quad \alpha \mapsto (\sigma_1\alpha, \dots, \sigma_r\alpha, \sigma_{r+1}\alpha, \dots, \sigma_{r+s}\alpha)$$

where $\{\sigma_1, \dots, \sigma_r, \sigma_{r+1}, \bar{\sigma}_{r+1}, \dots, \sigma_{r+s}, \bar{\sigma}_{r+s}\}$ is the complete set of embeddings of K into $\mathbb{C}$. It takes sums to sums. Now we want a map that takes products to sums, and so we take logarithms. Thus we consider the map:

$$L : K^\times \to \mathbb{R}^{r+s}, \quad \alpha \mapsto (\log|\sigma_1\alpha|, \dots, \log|\sigma_r\alpha|, 2\log|\sigma_{r+1}\alpha|, \dots, 2\log|\sigma_{r+s}\alpha|).$$

It is a homomorphism. If u is a unit in $\mathcal{O}_K$, then $\operatorname{Nm}_{K/\mathbb{Q}} u = \pm 1$, i.e.,

$$|\sigma_1\alpha| \cdots |\sigma_r\alpha||\sigma_{r+1}\alpha|^2 \cdots |\sigma_{r+s}\alpha|^2 = 1.$$

On taking logs, we see that $L(u)$ is contained in the hyperplane

$$H : x_1 + \cdots + x_r + 2x_{r+1} + \cdots + 2x_{r+s} = 0.$$

Dropping the last coordinate defines an isomorphism $H \approx \mathbb{R}^{r+s-1}$.

PROPOSITION 5.6. *The image of $L\colon U \to H$ is a lattice in H, and the kernel of L is a finite group (hence is $\mu(K)$).*

PROOF. Let C be a bounded subset of H containing 0, say

$$C \subset \{\mathbf{x} \in H \mid |x_i| \leq M\}.$$

If $L(u) \in C$, then $|\sigma_j u| \leq e^M$ for all j, and Lemma 5.5 implies that there are only finitely many such u's. Thus $L(U) \cap C$ is finite, and this implies that $L(U)$ is a lattice in H (by 4.14). Since everything in the kernel maps into C, the kernel is finite. □

Since the kernel of L is finite, we have

$$\operatorname{rank}(U) = \operatorname{rank}(L(U)) \leq \dim H = r + s - 1.$$

Computation of the rank. We now prove the unit theorem.

THEOREM 5.7. *The image $L(U)$ of U in H is a full lattice; thus U has rank $r+s-1$.*

PROOF. To prove the theorem, we have to find a way to construct units. We work again with the embedding

$$\sigma : K \hookrightarrow \mathbb{R}^r \times \mathbb{C}^s \approx \mathbb{R}^{r+2s}.$$

For $\mathbf{x} = (x_1, ..., x_r, x_{r+1}, ...) \in \mathbb{R}^r \times \mathbb{C}^s$, define

$$\operatorname{Nm}(\mathbf{x}) = x_1 \cdots x_r \cdot x_{r+1} \cdot \bar{x}_{r+1} \cdots x_{r+s} \cdot \bar{x}_{r+s}.$$

Then $\operatorname{Nm}(\sigma(\alpha)) = \operatorname{Nm}(\alpha)$. Note that $|\operatorname{Nm}(\mathbf{x})| = |x_1| \cdots |x_r||x_{r+1}|^2 \cdots |x_{r+s}|^2$.

Recall from (4.25), that $\sigma(\mathcal{O}_K)$ is a full lattice in $\mathbb{R}^r \times \mathbb{C}^s$, and the volume of its fundamental parallelopiped is $2^{-s} \cdot |\Delta|^{\frac{1}{2}}$; in more detail, if $\alpha_1, \ldots, \alpha_n$ is a $\mathbb{Z}$-basis for $\mathcal{O}_K$, then we showed that the absolute value of the determinant of the matrix whose i^{th} row is

$$\sigma(\alpha_i) = (\sigma_1(\alpha_i), \ldots, \Re(\sigma_{r+1}(\alpha_i)), \Im(\sigma_{r+1}(\alpha_i)), \ldots)$$

is $2^{-s} \cdot |\Delta|^{\frac{1}{2}}$. In fact, we showed that we could get this matrix from the matrix whose i^{th} row is

$$(\sigma_1(\alpha_i), \ldots, \sigma_{r+1}(\alpha_i), \bar{\sigma}_{r+1}(\alpha_i), \ldots)$$

by some elementary column operations that multiplied the absolute value of the determinant by 2^{-s}, and we know that the determinant of the second matrix is $\pm|\Delta|^{\frac{1}{2}}$.

In the rest of the proof, $\mathbf{x}$ will be a point of $\mathbb{R}^r \times \mathbb{C}^s$ with

$$1/2 \leq |\operatorname{Nm}(\mathbf{x})| \leq 1.$$

Define

$$\mathbf{x} \cdot \sigma(\mathcal{O}_K) = \{\mathbf{x} \cdot \sigma(\alpha) \mid \alpha \in \mathcal{O}_K\}.$$

Since $\mathbb{R}^r \times \mathbb{C}^s$ is a ring, this product makes sense. This is again a lattice in $\mathbb{R}^r \times \mathbb{C}^s$, and the volume of its fundamental parallelopiped is the determinant of the matrix whose i^{th} row is

$$(x_1\sigma_1(\alpha_i), \ldots, \Re(x_{r+1}\sigma_{r+1}(\alpha_i)), \Im(x_{r+1}\sigma_{r+1}(\alpha_i)), \ldots).$$

As before, the absolute value of the determinant of this matrix is 2^{-s} times the absolute value of the determinant of the matrix whose i^{th} row is

$$(x_1\sigma_1(\alpha_i), \dots, x_{r+1}\cdot\sigma_{r+1}(\alpha_i), \bar{x}_{r+1}\cdot\bar{\sigma}_{r+1}(\alpha_i), \dots),$$

which is

$$|\Delta|^{\frac{1}{2}}\cdot|\operatorname{Nm}(\mathbf{x})|.$$

Therefore $\mathbf{x}\cdot\sigma(\mathcal{O}_K)$ is a lattice with $2^{-s}|\Delta|^{\frac{1}{2}}|\operatorname{Nm}(\mathbf{x})|$ as the volume of its fundamental domain. Note that as $\mathbf{x}$ ranges over our set these volumes remain bounded.

Let T be a compact convex subset of $\mathbb{R}^r\times\mathbb{C}^s$, which is symmetric in the origin, and whose volume is so large that, for every $\mathbf{x}$ in the above set, Minkowski's theorem (4.18) implies there is a point γ of $\mathcal{O}_K$, $\gamma\neq 0$, such that $\mathbf{x}\cdot\sigma(\gamma)\in T$. The points of T have bounded coordinates, and hence bounded norms, and so

$$\mathbf{x}\cdot\sigma(\gamma)\in T\Rightarrow|\operatorname{Nm}(\mathbf{x}\cdot\sigma(\gamma))|\leq M,$$

for some M (depending on T); thus

$$|\operatorname{Nm}(\gamma)|\leq M/\operatorname{Nm}(\mathbf{x})\leq 2M.$$

Consider the set of ideals $\gamma\cdot\mathcal{O}_K$, where γ runs through the γ's in $\mathcal{O}_K$ for which $\mathbf{x}\cdot\sigma(\gamma)\in T$ for some $\mathbf{x}$ in our set. The norm $\mathbb{N}$ of such an ideal is $\leq 2M$, and so there can only be finitely many such ideals, say $\gamma_1\cdot\mathcal{O}_K,\dots,\gamma_t\cdot\mathcal{O}_K$. Now if γ is any element of $\mathcal{O}_K$ with $\mathbf{x}\cdot\sigma(\gamma)\in T$, some $\mathbf{x}$, then $\gamma\cdot\mathcal{O}_K=\gamma_i\cdot\mathcal{O}_K$ for some i, and so there exists a unit ε such that $\gamma=\gamma_i\cdot\varepsilon$. Then $\mathbf{x}\cdot\sigma(\varepsilon)\in\sigma(\gamma_i^{-1})\cdot T$. The set $T'=\sigma(\gamma_1^{-1})\cdot T\cup\ldots\cup\sigma(\gamma_t^{-1})\cdot T$ is bounded, and so we have shown that, for each $\mathbf{x}$ in our set there exists a unit ε such that the coordinates of $\mathbf{x}\cdot\sigma(\varepsilon)$ are bounded uniformly in $\mathbf{x}$ (the set T' doesn't depend on $\mathbf{x}$).

We are now ready to prove that $L(U)$ is a full lattice in H. If $r+s-1=0$, there is nothing to prove, and so we assume $r+s-1\geq 1$.

For each i, $1\leq i\leq r+s$, we choose an $\mathbf{x}$ in our set such that all the coordinates of $\mathbf{x}$ except x_i are very large (compared with T'), and x_i is sufficiently small that $|\operatorname{Nm}\mathbf{x}|=1$. We know that there exists a unit ε_i such that $\mathbf{x}\cdot\sigma(\varepsilon_i)$ has bounded coordinates, and we deduce that $|\sigma_j\varepsilon_i|<1$ for $j\neq i$, and hence that $\log|\sigma_j\varepsilon_i|<0$.

I claim that $L(\varepsilon_1),\ldots,L(\varepsilon_{r+s-1})$ are linearly independent vectors in the lattice $L(U)$. For this we have to prove that the matrix whose i^{th} row is

$$(l_1(\varepsilon_i),\ldots,l_{r+s-1}(\varepsilon_i)),\quad l_i(\varepsilon)=\log|\sigma_i\varepsilon|,$$

is invertible. The elements of the matrix except those on the diagonal are negative, but the sum

$$l_1(\varepsilon_i)+\cdots+l_{r+s-1}(\varepsilon_i)+l_{r+s}(\varepsilon_i)=0,$$

and so the sum of the terms in the i^{th} row

$$l_1(\varepsilon_i)+\cdots+l_{r+s-1}(\varepsilon_i)=-l_{r+s}(\varepsilon_i)>0.$$

The next lemma implies that the matrix is invertible, and so completes the proof of Theorem 5.7. □

LEMMA 5.8. *Let (a_{ij}) be a real $m\times m$ matrix such that*

(a) $a_{ij}<0$ *for* $i\neq j$;

(b) $\sum_j a_{ij} > 0$ *for* $i = 1, 2, ..., m$.

Then (a_{ij}) *is invertible.*

PROOF. If it isn't, then the system of equations

$$\sum a_{ij}x_j = 0$$

has a nontrivial solution. Write $x_1, ..., x_m$ for such a solution, and suppose i is such that $|x_i| = \max\{|x_j|\}$. We can scale the solution so that $x_i = 1$. Then $|x_j| \leq 1$ for $j \neq i$, and the i^{th} equation is

$$0 = \sum a_{ij}x_j = a_{ii} + \sum_{j\neq i} a_{ij}x_j \geq a_{ii} + \sum_{j\neq i} a_{ij} > 0.$$

□

S-units. Let S be a finite set of prime ideals of K, and define the ring of *S-integers* to be

$$\mathcal{O}_K(S) = \cap_{\mathfrak{p}\notin S}\mathcal{O}_\mathfrak{p} = \{\alpha \in K \mid \mathrm{ord}_\mathfrak{p}(\alpha) \geq 0, \text{ all } \mathfrak{p} \notin S\}.$$

For example, if $S = \emptyset$, then $\mathcal{O}_K(S) = \mathcal{O}_K$.

Define the group of *S-units*, to be

$$U(S) = \mathcal{O}_K(S)^\times = \{\alpha \in K \mid \mathrm{ord}_\mathfrak{p}(\alpha) = 0, \text{ all } \mathfrak{p} \notin S\}.$$

Clearly, the torsion subgroup of $U(S)$ is again $\mu(K)$.

THEOREM 5.9. *The group of S-units is finitely generated with rank $r+s+\#S-1$.*

PROOF. Let $\mathfrak{p}_1, \mathfrak{p}_2, \ldots, \mathfrak{p}_t$ be the elements of S. The homomorphism

$$u \mapsto (\ldots, \mathrm{ord}_{\mathfrak{p}_i}(u), \ldots) \colon U(S) \to \mathbb{Z}^t$$

obviously has kernel U. To complete the proof, it suffices to show that the image of $U(S)$ in $\mathbb{Z}^t$ has rank t. Let h be the class number of K. Then $\mathfrak{p}_i^h$ is principal, say $\mathfrak{p}_i^h = (\pi_i)$, and π_i is an S-unit with image

$$(0, \ldots, h, \ldots, 0) \quad (h \text{ in the } i^{th} \text{ position}).$$

Clearly these elements generate a subgroup of rank t. □

For example, if $K = \mathbb{Q}$ and $S = \{(2), (3), (5)\}$ then

$$U(S) = \{\pm 2^k 3^m 5^n \mid k, m, n \in \mathbb{Z}\},$$

and the statement is obvious in this case.

Finding fundamental units in real quadratic fields. An expression

$$a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{a_3 + \ldots}}}$$

is called a *continued fraction*. We abbreviate the expression on the right as

$$[a_0, a_1, a_2, \ldots].$$

We shall always assume that the a_i are integers with $a_1 > 0$, $a_2 > 0, \ldots$. The integers a_i are called the *quotients,* and $[a_0, a_1, ..., a_n]$ is called the n^{th} *convergent.* Every irrational number α can be expressed in just one way as an infinite continued fraction, and the continued fraction is periodic if and only if α has degree 2 over $\mathbb{Q}$. (See any book on elementary number theory, for example, Hardy, G. H., and Wright, E. M., An Introduction to the Theory of Numbers, Oxford Univ. Press, 1960 (4th edition), Chapter X.)

Now let d be a square-free positive integer, and let ε be the (unique) fundamental unit for $\mathbb{Q}[\sqrt{d}]$ with $\varepsilon > 1$. Let s be the period of the continued fraction for $\sqrt{d}$ and let p/q be the $(s-1)^{th}$ convergent of it; then

$$\varepsilon = p + q\sqrt{d} \text{ if } d \equiv 2, 3 \mod 4, \text{ or } d \equiv 1 \mod 8,$$

and

$$\varepsilon = p + q\sqrt{d} \text{ or } \varepsilon^3 = p + q\sqrt{d} \text{ otherwise.}$$

Using Maple or Mathematica, it is very easy to carry this out, and one obtains some spectacularly large numbers.

For example, to find the fundamental unit in $\mathbb{Q}[\sqrt{94}]$, first compute $\sqrt{94} = 9.6954\ldots$. Then compute the continued fraction of $\sqrt{94}$. One gets

$$\{9, 1, 2, 3, 1, 1, 5, 1, 8, 1, 5, 1, 1, 3, 2, 1, 18, 1, 2, 3, \ldots\}.$$

This suggests the period is 16. Now evaluate the 15^{th} convergent. One gets

$$\frac{2143295}{221064}$$

Hence the fundamental unit > 1 is

$$\varepsilon = 2143295 + 221064 \cdot \sqrt{94}.$$

Compute that

$$(2143295)^2 - (221064)^2 \cdot 94 = 1,$$

which verifies that ε is a unit.

When one carries out this procedure for $\mathbb{Q}[\sqrt{9199}]$, the first coefficient of the fundamental unit has 88 digits! The computer has no problem finding the fundamental unit — the only problem is counting the length of the period, which is about 180.

Units in cubic fields with negative discriminant. Since the sign of the discriminant is $(-1)^s$ (see 2.39), a cubic field K will have negative discriminant if and only if $r = 1 = s$. We identify K with a subfield of $\mathbb{R}$ using its unique real embedding. We have $\Delta < 0$, and the group of units is $\{\pm\varepsilon^m\}$ for some ε (fundamental unit). We want to find ε. Since -ε, $-\varepsilon^{-1}$, and ε^{-1} are also fundamental units, we may suppose that $\varepsilon > 1$.

LEMMA 5.10. *Let K be a cubic extension of $\mathbb{Q}$ with negative discriminant, and let ε be the fundamental unit with $\varepsilon > 1$. Then*

$$|\Delta_K| < 4\varepsilon^3 + 24.$$

PROOF. Since $\varepsilon \notin \mathbb{Q}$, it must generate K. The two conjugates of ε (other than ε itself) must be complex conjugates, and so the product of ε with its conjugates must be $+1$ (rather than -1). Write $\varepsilon = u^2$, $u \in \mathbb{R}$, $u > 1$. Then the remaining conjugates of ε can be written

$$u^{-1}e^{i\theta}, \quad u^{-1}e^{-i\theta} \quad (0 \le \theta \le \pi).$$

Let $\Delta' = D(1, \varepsilon, \varepsilon^2)$ be the discriminant of the minimum equation of ε. Then

$$\Delta'^{\frac{1}{2}} = (u^2 - u^{-1}e^{i\theta})(u^2 - u^{-1}e^{-i\theta})(u^{-1}e^{i\theta} - u^{-1}e^{-i\theta}) = 2i(u^3 + u^{-3} - 2\cos\theta)\sin\theta.$$

If we set $2\xi = u^3 + u^{-3}$, then

$$|\Delta'|^{\frac{1}{2}} = 4(\xi - \cos\theta)\sin\theta,$$

which, for a given u, has a maximum where

$$\xi\cos\theta - \cos^2\theta + \sin^2\theta = 0,$$

or

$$-g(x) \stackrel{\text{df}}{=} \xi x - 2x^2 + 1 = 0, \quad |x| \le 1, \quad x = \cos\theta.$$

We seek a root of $g(x)$ with $|x| < 1$. But $g(1) = 1 - \xi < 0$ (because $u > 1$ implies $\xi = \frac{u^3 - u^{-3}}{2} > 1$), and $g(-\frac{1}{2u^3}) = \frac{3}{4}(u^{-6} - 1) < 0$. Since $g(x) = 2x^2 + \cdots$, it follows $g(x)$ has one root > 1, and that the desired root x_0, with $|x_0| \le 1$, is $< -\frac{1}{2u^3}$. But then

$$x_0^2 > \frac{1}{4u^6} \Rightarrow u^{-6} - 4x_0^2 < 0 \Rightarrow u^{-6} - 4x_0^{-2} - 4x_0^4 < 0. \qquad (*)$$

This maximum yields

$$|\Delta'| \le 16(\xi^2 - 2\xi x_0 + x_0^2)(1 - x_0^2),$$

and, on applying the conditions $\xi x_0 = 2x_0^2 - 1$, $\xi^2 x_0^2 = 4x_0^4 - 4x_0^2 + 1$, and the inequality (*) we find that

$$|\Delta'| \le 16(\xi^2 + 1 - x_0^2 - x_0^4) = 4u^6 + 24 + 4(u^{-6} - 4x_0^2 - 4x_0^4) < 4u^6 + 24.$$

Hence

$$|\Delta'| < 4\varepsilon^3 + 24.$$

Since $\Delta' = \Delta_K \cdot$ (square of an integer), this completes the proof. □

EXAMPLE 5.11. Let $K = \mathbb{Q}[\alpha]$ where α is a real root of $X^3 + 10X + 1$. Here the discriminant is -4027, and so $\varepsilon > \sqrt[3]{\frac{4027-24}{4}} > 10$ for ε the fundamental unit with $\varepsilon > 1$. Note that $\mathrm{Nm}(\alpha) = -1$, and so α is a unit. Moreover, $\alpha = -0.0999003...$ and so $\beta = -\alpha^{-1} = 10.00998....$ Since β is a power of ε, we must have $\beta = \varepsilon$; i.e., $-\alpha^{-1}$ is the fundamental unit > 1. Thus

$$U_K = \{\pm\alpha^m \mid m \in \mathbb{Z}\}.$$

Once one knows ε, it becomes easier to compute the class group. We know (see 3.49) that there is a prime ideal $\mathfrak{p} = (2, 1 + \alpha)$ such that $\mathbb{N}(\mathfrak{p}) = 2$. One shows that $\mathfrak{p}$ generates the class group, and it then remains to find the order of $\mathfrak{p}$. One verifies that $\mathfrak{p}^6$ is the ideal generated by $\frac{(\alpha-1)^3}{\alpha+2}$, and so it remains to show that $\mathfrak{p}^2$ and $\mathfrak{p}^3$ are nonprincipal.

Suppose $\mathfrak{p}^3 = (\gamma)$. Then $\gamma^2 = \pm\alpha^m \cdot \frac{(\alpha-1)^3}{\alpha+2}$ for some m and choice of signs. But this says that at least one of the numbers $\frac{\alpha-1}{\alpha+2}$, $-\frac{\alpha-1}{\alpha+2}$, $\alpha\frac{\alpha-1}{\alpha+2}$, $-\alpha\frac{\alpha-1}{\alpha+2}$ is a square. Let β be that number. If $\mathfrak{q}$ is a prime ideal such that $\beta \in \mathcal{O}_\mathfrak{q}$ (i.e., such that $\mathrm{ord}_\mathfrak{q}(\beta) \geq 0$), then we can look at β mod $\mathfrak{q}$ and ask if it is a square.

We first work modulo 29. We have

$$X^3 + 10X + 1 \equiv (X+5)(X-3)(X-2) \mod 29.$$

Take $\mathfrak{q}$ to be the ideal $(29, \alpha - 2)$. The residue field $\mathcal{O}_K/\mathfrak{q}$ is $\mathbb{F}_{29} = \mathbb{Z}/(29)$, and the map $\mathbb{Z}[\alpha] \to \mathbb{F}_{29}$ is $\alpha \mapsto 2 \pmod{29}$. Thus

$$\alpha - 1 \mapsto 1, \quad \alpha + 2 \mapsto 4, \quad (\alpha+2)^{-1} \mapsto 22, \quad -1 \mapsto -1.$$

The numbers 1, 4, and $-1 \equiv 12^2$ are squares modulo 29, but 2 is not; hence m must be 0. Since $\frac{\alpha-1}{\alpha+2} < 0$ it can't be a square in K (since it isn't even in $\mathbb{R}$), and so the only possibility for β is $-\frac{\alpha-1}{\alpha+2}$. We eliminate this by looking mod 7.

Take $\mathfrak{q} = (7, \alpha + 3)$ (see 3.49). Then in the map $\mathbb{Z}[\alpha] \to \mathbb{Z}[\alpha]/\mathfrak{q} = \mathbb{F}_7$,

$$\alpha \mapsto -3 = 4, \quad -\frac{\alpha-1}{\alpha+2} \mapsto \frac{-3}{6} \equiv -\frac{1}{2} \equiv -4 \equiv 3 \mod 7,$$

and 3 is not a square modulo 7. Thus $-\frac{\alpha-1}{\alpha+2}$ is not a square in $\mathbb{Q}[\alpha]$.

Similarly, $\mathfrak{p}^2 = (\gamma)$ can be shown to be impossible. Thus $\mathrm{Cl}(\mathcal{O}_K)$ is a cyclic group of order 6.

Finding $\mu(K)$. If $\mathbb{Q}[\zeta_m] \subset K$, where ζ_m is a primitive m^{th} root of 1, then $\varphi(m) | [K\colon \mathbb{Q}]$. Thus there are only finitely many possibilities for m. For each of them, use the test in the later section on algorithms to determine whether the minimum polynomial Φ_m for ζ_m has a root in K.

Finding a system of fundamental units. The strategy for finding units in the general case seems to be to find lots of solutions to equations $\mathrm{Nm}(\alpha) = m$ for m a fixed small number, and then take quotients of solutions. Note that there can be only finitely many ideals $\mathfrak{a}$ with $\mathbb{N}(\mathfrak{a}) = m$; thus if we have lots of elements α_i with $\mathrm{Nm}(\alpha_i) = m$, then frequently $\alpha_i \cdot \mathcal{O}_K = \alpha_j \cdot \mathcal{O}_K$, and this implies that α_i and α_j differ by a unit — note that this was the strategy used to prove the unit theorem. See Pohst and Zassenhaus 1989, Chapter 5.

Regulators. There is one other important invariant that we should define. Let $t = r + s - 1$, and let $u_1, ..., u_t$ be a system of fundamental units. Then the vectors

$$L(u_i) \stackrel{\mathrm{df}}{=} (\log|\sigma_1 u_i|, ..., \log|\sigma_r u_i|, 2\cdot\log|\sigma_{r+1}u_i|, \dots, 2\log|\sigma_t u_i|) \in \mathbb{R}^t$$

generate the lattice $L(U)$ in $\mathbb{R}^t$. The *regulator* is defined to be determinant of the matrix whose i^{th} row is $L(u_i)$. Thus, up to sign, the regulator is the volume of a fundamental domain for $L(U)$ (regarded as a full lattice in $\mathbb{R}^t$).

The regulator plays the same role for the group of units (mod torsion) that the discriminant plays for $\mathcal{O}_K$. One can similarly define the regulator of any set $\{\varepsilon_1, ..., \varepsilon_t\}$ of independent units, and the square of the index of the group generated by the ε_i and $\mu(K)$ in the full group of units is measured by ratio

$$|\mathrm{Reg}(\varepsilon_1, \dots, \varepsilon_t)| / |\mathrm{Reg}(U)|.$$

There are lower bounds for the regulator (see Pohst and Zassenhaus 1989, p 365) similar to the one we proved for a cubic field with one real embedding.

6. Cyclotomic Extensions; Fermat's Last Theorem.

Cyclotomic extensions of $\mathbb{Q}$, i.e., extensions generated by a root of 1, provide interesting examples of the theory we have developed, but, more importantly, their arithmetic has important applications.

The basic results. An element ζ of a field K is said to be a *primitive* n^{th} root of 1 if $\zeta^n = 1$ but $\zeta^d \neq 1$ for any $d < n$, i.e., if ζ is an element of order n in $K^\times$. For example, the n^{th} roots of 1 in $\mathbb{C}$ are the numbers $e^{2\pi im/n}$, $0 \leq m \leq n-1$, and the next lemma shows that $e^{2\pi im/n}$ is a primitive n^{th} root of 1 if and only if m is relatively prime to n.

Lemma 6.1. *Let ζ be a primitive n^{th} root of 1. Then ζ^m is again a primitive n^{th} root of 1 if and only if m is relatively prime to n.*

Proof. This is a consequence of a more general fact: if α is an element of order n in a group, then α^m is also of order n if and only if m is relatively prime to n. Here is the proof. If $d|m, n$, then $(\alpha^m)^{\frac{n}{d}} = \alpha^{n\frac{m}{d}} = 1$. Conversely, if m and n are relatively prime, then there are integers a and b such that

$$am + bn = 1.$$

Now $\alpha^{am} = \alpha$ and so $(\alpha^m)^d = 1 \Rightarrow \alpha^d = (\alpha^{am})^d = 1 \Rightarrow n|d$. □

Let $K = \mathbb{Q}[\zeta]$, where ζ is a primitive n^{th} root of 1. Then K is the splitting field of $X^n - 1$, and so it is Galois over $\mathbb{Q}$. Let $G = \mathrm{Gal}(\mathbb{Q}[\zeta]/\mathbb{Q})$. It permutes the set of primitive n^{th} roots of 1 in K, and so, for any $\sigma \in G$, $\sigma\zeta = \zeta^m$ for some integer m relatively prime to n; moreover, m is well-defined modulo n. The map $\sigma \mapsto [m]$ is an injective homomorphism $G \to (\mathbb{Z}/n\mathbb{Z})^\times$. In FT, Proposition 5.7, it is proved that this map is an isomorphism, and so $[K : \mathbb{Q}] = \varphi(n) \overset{\text{df}}{=} \#(\mathbb{Z}/n\mathbb{Z})^\times$. We shall give another proof, and at the same time obtain many results concerning the arithmetic of $\mathbb{Q}[\zeta]$.

The *cyclotomic polynomial* Φ_n is defined to be,

$$\Phi_n(X) = \prod(X - \zeta^m)$$

where the product runs over a set of representatives m for the elements of $(\mathbb{Z}/n\mathbb{Z})^\times$, for example, over the integers m, $0 \leq m \leq n-1$, relatively prime to n. Alternatively,

$$\Phi_n(X) = \prod(X - \zeta')$$

where ζ' runs over the primitive n^{th} roots of 1. Because G permutes the ζ', $\Phi_n(X) \in \mathbb{Q}[X]$, and clearly $\Phi_n(\zeta) = 0$. Therefore, $\Phi_n(X)$ is the minimum polynomial of ζ if and only if it is irreducible, in which case $[K : \mathbb{Q}] = \varphi(n)$ and the map $G \to (\mathbb{Z}/n\mathbb{Z})^\times$ is an isomorphism. Hence the following statements are equivalent:

(a) the map $\mathrm{Gal}(\mathbb{Q}[\zeta]/\mathbb{Q}) \to (\mathbb{Z}/n\mathbb{Z})^\times$ is an isomorphism;
(b) $[\mathbb{Q}[\zeta] : \mathbb{Q}] = \varphi(n)$;
(c) $\mathrm{Gal}(\mathbb{Q}[\zeta]/\mathbb{Q})$ acts transitively on the set of primitive n^{th} roots of 1 (i.e., they are conjugates);
(d) $\Phi_n(X)$ is irreducible (and so $\Phi_n(X)$ is the minimum polynomial of ζ).

We shall see that all these statements are true.

Note that each n^{th} root of 1 is a primitive d^{th} root of 1 for exactly one $d|n$, and so

$$X^n - 1 = \prod_{d|n} \Phi_d(X) = (X-1)\cdots\Phi_n(X).$$

We first examine a cyclotomic extension in the case that n is a power p^r of a prime.

PROPOSITION 6.2. *Let ζ be a primitive p^rth root of 1, and let $K = \mathbb{Q}[\zeta]$.*

(a) *The field $\mathbb{Q}[\zeta]$ is of degree $\varphi(p^r) = p^{r-1}(p-1)$ over $\mathbb{Q}$.*
(b) *The ring of integers in $\mathbb{Q}[\zeta]$ is $\mathbb{Z}[\zeta]$.*
(c) *The element $\pi \overset{df}{=} 1-\zeta$ is a prime element of $\mathcal{O}_K$, and $(p) = (\pi)^e$ with $e = \varphi(p^r)$.*
(d) *The discriminant of $\mathcal{O}_K$ over $\mathbb{Z}$ is $\pm p^c$, some c (in fact, $c = p^{r-1}(pr-r-1)$); therefore, p is the only prime to ramify in $\mathbb{Q}[\zeta]$.*

PROOF. Observe first that $\mathbb{Z}[\zeta] \subset \mathcal{O}_K$ (obviously).

If ζ' is another primitive p^rth root of 1, then $\zeta' = \zeta^s$ and $\zeta = \zeta'^t$ for some integers s and t not divisible by p, and so $\mathbb{Q}[\zeta'] = \mathbb{Q}[\zeta]$, $\mathbb{Z}[\zeta'] = \mathbb{Z}[\zeta]$. Moreover,

$$\frac{1-\zeta'}{1-\zeta} = 1 + \zeta + \cdots + \zeta^{s-1} \in \mathbb{Z}[\zeta].$$

Similarly, $(1-\zeta)/(1-\zeta') \in \mathbb{Z}[\zeta]$, and so $(1-\zeta')/(1-\zeta)$ is a unit in $\mathbb{Z}[\zeta]$ (hence also in $\mathcal{O}_K$). Note that

$$\Phi_{p^r}(X) = \frac{X^{p^r}-1}{X^{p^{r-1}}-1} = \frac{t^p-1}{t-1} = 1 + t + \cdots + t^{p-1}, \quad t = X^{p^{r-1}},$$

and so

$$\Phi_{p^r}(1) = p.$$

For its definition, we see that

$$\Phi_{p^r}(1) = \prod(1-\zeta') = \prod \frac{1-\zeta'}{1-\zeta}(1-\zeta) = u\cdot(1-\zeta)^{\varphi(p^r)},$$

with u a unit in $\mathbb{Z}[\zeta]$. Therefore we have an equality of ideals in $\mathcal{O}_K$,

$$(p) = (\pi)^e, \quad \pi \overset{\text{df}}{=} 1-\zeta, \quad e = \varphi(p^r),$$

and so (p) has at least $\varphi(p^r)$ prime factors in $\mathcal{O}_K$. Now (3.36) implies that $[\mathbb{Q}[\zeta] : \mathbb{Q}] \geq \varphi(p^r)$. This proves (a) of the Proposition since we know $[\mathbb{Q}[\zeta] : \mathbb{Q}] \leq \varphi(p^r)$.

Moreover we see that π must generate a prime ideal in $\mathcal{O}_K$, otherwise, again, (p) would have too many prime-ideal factors. This completes the proof of (c).

For future reference, we note that, in $\mathcal{O}_K$,

$$(p) = \mathfrak{p}^{\varphi(p^r)}, \quad \mathfrak{p} = (\pi), \quad f(\mathfrak{p}/p) = 1.$$

The last equality means that the map $\mathbb{Z}/(p) \to \mathcal{O}_K/(\pi)$ is an isomorphism.

We next show that (up to sign) $\operatorname{disc}(\mathbb{Z}[\zeta]/\mathbb{Z})$ is a power of p. Since

$$\operatorname{disc}(\mathcal{O}_K/\mathbb{Z})\cdot(\mathcal{O}_K : \mathbb{Z}[\zeta])^2 = \operatorname{disc}(\mathbb{Z}[\zeta]/\mathbb{Z}),$$

this will imply:

(i) $\operatorname{disc}(\mathcal{O}_K/\mathbb{Z})$ is a power of p;

(ii) $(\mathcal{O}_K : \mathbb{Z}[\zeta])$ is a power of p, and therefore $p^M(\mathcal{O}_K/\mathbb{Z}[\zeta]) = 0$ for some M, i.e., $p^M\mathcal{O}_K \subset \mathbb{Z}[\zeta]$.

To compute $\operatorname{disc}(\mathbb{Z}[\zeta]/\mathbb{Z})$, we shall use the formula in (2.33), which in our case reads:

$$\operatorname{disc}(\mathbb{Z}[\zeta]/\mathbb{Z}) = \pm \operatorname{Nm}_{K/\mathbb{Q}}(\Phi'_{p^r}(\zeta)).$$

On differentiating the equation

$$(X^{p^{r-1}} - 1) \cdot \Phi_{p^r}(X) = X^{p^r} - 1$$

and substituting ζ for X, we find that $\Phi'_{p^r}(\zeta) = p^r\zeta^{p^r-1}/(\zeta^{p^{r-1}} - 1)$. Clearly

$$\operatorname{Nm}_{K/\mathbb{Q}}\zeta = \pm 1, \quad \operatorname{Nm}_{K/\mathbb{Q}}p^r = (p^r)^{\varphi(p^r)} = p^{r\varphi(p^r)}.$$

We shall show that

$$\operatorname{Nm}_{K/\mathbb{Q}}(1 - \zeta^{p^s}) = p^{p^s}, \quad 0 \le s < r,$$

and so

$$\operatorname{Nm}_{K/\mathbb{Q}}\Phi'_{p^r}(\zeta) = \pm p^c, \quad c = r(p-1)p^{r-1} - p^{r-1} = p^{r-1}(pr - r - 1).$$

First we compute $\operatorname{Nm}_{K/\mathbb{Q}}(1-\zeta)$. The minimum polynomial of $1-\zeta$ is $\Phi_{p^r}(1-X)$, which has constant term $\Phi_{p^r}(1) = p$, and so $\operatorname{Nm}_{K/\mathbb{Q}}(1-\zeta) = \pm p$.

We next compute $\operatorname{Nm}_{K/\mathbb{Q}}(1-\zeta^{p^s})$ some $s < r$. Because ζ^{p^s} is a primitive p^{r-s}th root of 1, the computation just made (with r replaced by $r-s$) shows that

$$\operatorname{Nm}_{\mathbb{Q}[\zeta^{p^s}]/\mathbb{Q}}(1 - \zeta^{p^s}) = \pm p.$$

Using that

$$\operatorname{Nm}_{M/K} = \operatorname{Nm}_{L/K} \circ \operatorname{Nm}_{M/L} \text{ and } \operatorname{Nm}_{M/L}\alpha = \alpha^{[M:L]} \text{ if } \alpha \in L,$$

we see that

$$\operatorname{Nm}_{K/\mathbb{Q}}(1 - \zeta^{p^s}) = p^a \text{ where } a = [\mathbb{Q}[\zeta] : \mathbb{Q}[\zeta^{p^s}]] = \varphi(p^r)/\varphi(p^{r-s}) = p^s.$$

This completes the proof of (d).

We are now ready to prove (b). As we observed above the inclusion $\mathbb{Z} \hookrightarrow \mathcal{O}_K$ induces an isomorphism $\mathbb{Z}/(p) \to \mathcal{O}_K/(\pi)$. Thus

$$\mathbb{Z} + \pi\mathcal{O}_K = \mathcal{O}_K,$$

and *a fortiori*

$$\mathbb{Z}[\zeta] + \pi\mathcal{O}_K = \mathcal{O}_K. \qquad (*)$$

On multiplying through by π, we obtain the equality

$$\pi\mathbb{Z}[\zeta] + \pi^2\mathcal{O}_K = \pi\mathcal{O}_K. \qquad (**)$$

Let $\alpha \in \mathcal{O}_K$; equation (*) shows that we can write $\alpha = \alpha' + \gamma$ with $\alpha' \in \pi\mathcal{O}_K$ and $\gamma \in \mathbb{Z}[\zeta]$, and (**) shows that we can write $\alpha' = \alpha'' + \gamma'$ with $\alpha'' \in \pi^2\mathcal{O}_K$ and $\gamma' \in \mathbb{Z}[\zeta]$. Hence $\alpha = (\gamma + \gamma') + \alpha''$, and so

$$\mathbb{Z}[\zeta] + \pi^2\mathcal{O}_K = \mathcal{O}_K.$$

On repeating these arguments, we find that

$$\mathbb{Z}[\zeta] + \pi^m\mathcal{O}_K = \mathcal{O}_K$$

for all $m \in \mathbb{N}$. Since $\pi^{\varphi(p^r)} = p \times (\text{unit})$, this implies that

$$\mathbb{Z}[\zeta] + p^m \cdot \mathcal{O}_K = \mathcal{O}_K$$

for all $m \in \mathbb{N}$. But for m large enough, we know that $p^m \mathcal{O}_K \subset \mathbb{Z}[\zeta]$, and so it can be dropped from the equation. Hence $\mathbb{Z}[\zeta] = \mathcal{O}_K$, and this completes the proof of (b). □

REMARK 6.3. (a) The sign of the $\operatorname{disc}(\mathbb{Q}[\zeta]/\mathbb{Q})$, ζ any root of 1, can be computed most easily by using (2.39a). Clearly $\mathbb{Q}[\zeta]$ has no real embeddings unless $\zeta = \pm 1$ (and $\mathbb{Q}[\zeta] = \mathbb{Q}$), and so, except for this case,

$$\operatorname{sign}(\operatorname{disc}(\mathbb{Q}[\zeta]/\mathbb{Q})) = (-1)^s, \quad s = [\mathbb{Q}[\zeta] : \mathbb{Q}]/2.$$

If ζ is a primitive p^rth root of 1, $p^r > 2$, then

$$[\mathbb{Q}[\zeta] : \mathbb{Q}]/2 = (p-1)p^{r-1}/2$$

which is odd if and only if $p^r = 4$ or $p \equiv 3 \mod 4$.

(b) Let ζ and ζ' be primitive p^rth and q^sth roots of 1. If p and q are distinct primes, then $\mathbb{Q}[\zeta] \cap \mathbb{Q}[\zeta'] = \mathbb{Q}$, because $K \subset \mathbb{Q}[\zeta] \Rightarrow p$ ramifies totally in K and q does not, and $K \subset \mathbb{Q}[\zeta'] \Rightarrow q$ ramifies totally in K and p does not, and these are contradictory unless $K = \mathbb{Q}$.

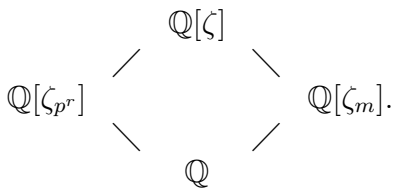
THEOREM 6.4. *Let ζ be a primitive n^{th} root of 1.*

(a) *The field $\mathbb{Q}[\zeta]$ is of degree $\varphi(n)$ over $\mathbb{Q}$.*
(b) *The ring of integers in $\mathbb{Q}[\zeta]$ is $\mathbb{Z}[\zeta]$.*
(c) *If p ramifies in $\mathbb{Q}[\zeta]$ then $p|n$; more precisely, if $n = p^r \cdot m$ with m relatively prime to p, then*

$$(p) = (\mathfrak{p}_1 \cdots \mathfrak{p}_s)^{\varphi(p^r)}$$

in $\mathbb{Q}[\zeta]$ with the $\mathfrak{p}_i$ distinct primes in $\mathbb{Q}[\zeta]$.

PROOF. We use induction on the number of primes dividing n. Write $n = p^r \cdot m$ with m not divisible by p. We may assume the theorem for m. Note that $\zeta_{p^r} \overset{\text{df}}{=} \zeta^m$ is a primitive p^rth root of 1, $\zeta_m = \zeta^{p^r}$ is a primitive mth root of 1, and that $\mathbb{Q}[\zeta] = \mathbb{Q}[\zeta_{p^r}] \cdot \mathbb{Q}[\zeta_m]$. Consider the fields:



The prime ideal (p) ramifies totally in $\mathbb{Q}[\zeta_{p^r}]$, say $(p) = \mathfrak{p}^{\varphi(p^r)}$, but doesn't ramify in $\mathbb{Q}[\zeta_m]$, say $(p) = \prod \mathfrak{p}_i$ with the $\mathfrak{p}_i$ distinct primes. On comparing the factorization of (p) in $\mathbb{Q}[\zeta]$ obtained by going two different ways up the tower, one finds that $[\mathbb{Q}[\zeta] : \mathbb{Q}[\zeta_m]] = \varphi(p^r)$, and that $(p) = (\prod \mathfrak{q}_i)^{\varphi(p^r)}$, where

$$\mathfrak{p}_i \mathcal{O}_{\mathbb{Q}[\zeta]} = \mathfrak{q}_i^{\varphi(p^r)}, \quad \mathfrak{p}\mathcal{O}_{\mathbb{Q}[\zeta]} = \prod \mathfrak{q}_i.$$

Again we are using (3.36). Therefore $[\mathbb{Q}[\zeta] : \mathbb{Q}[\zeta_m]] = \varphi(p^r)$, and $[\mathbb{Q}[\zeta] : \mathbb{Q}] = \varphi(p^r) \cdot \varphi(m) = \varphi(n)$. The following lemma completes the proof of the theorem (because it shows that $\mathcal{O}_{\mathbb{Q}[\zeta]} = \mathbb{Z}[\zeta_{p^r}, \zeta_m] = \mathbb{Z}[\zeta]$). □

LEMMA 6.5. *Let K and L be finite extensions of $\mathbb{Q}$ such that*

$$[KL\colon \mathbb{Q}] = [K\colon \mathbb{Q}] \cdot [L\colon \mathbb{Q}],$$

and let $d = \gcd(\operatorname{disc}(\mathcal{O}_K/\mathbb{Z}), \operatorname{disc}(\mathcal{O}_L/\mathbb{Z}))$. Then

$$\mathcal{O}_{K\cdot L} \subset d^{-1}\mathcal{O}_K \cdot \mathcal{O}_L.$$

PROOF. Let $\{\alpha_1, ..., \alpha_m\}$ and $\{\beta_1, ..., \beta_m\}$ be integral bases for K and L respectively. Then $\alpha_i\beta_j$ is a basis for $K \cdot L$ over $\mathbb{Q}$. Thus every $\gamma \in \mathcal{O}_{K\cdot L}$ can be written uniquely in the form

$$\gamma = \sum_{ij} \frac{a_{ij}}{r}\alpha_i\beta_j, \quad a_{ij}, r \in \mathbb{Z}.$$

After dividing out any common factors from top and bottom, no prime factor of r will divide all the a_{ij}, and we then have to show that $r|d$.

Identify L with a subfield of $\mathbb{C}$, and let σ be an embedding of K into $\mathbb{C}$. Then σ extends uniquely to an embedding of $K \cdot L$ into $\mathbb{C}$ fixing the elements of L (to see this, write $K = \mathbb{Q}[\alpha]$; then $K \cdot L = L[\alpha]$, and the hypothesis on the degrees implies that the minimum polynomial of α doesn't change when we pass from $\mathbb{Q}$ to L; there is therefore a unique L-homomorphism $L[\alpha] \to \mathbb{C}$ sending α to $\sigma\alpha$). On applying σ to the above equation, we obtain an equation

$$\sigma(\gamma) = \sum_{ij} \frac{a_{ij}}{r}\sigma(\alpha_i)\beta_j.$$

Write $x_i = \sum_j (a_{ij}/r)\beta_j$, and let $\sigma_1, \sigma_2, ..., \sigma_m$ be the distinct embeddings of K into $\mathbb{C}$. We obtain a system of m linear equations

$$\sum_i \sigma_k(\alpha_i)x_i = \sigma_k(\gamma), \quad k = 1, 2, ..., m,$$

and Cramer's rule tells us that

$$Dx_i = D_i$$

where $D = \det(\sigma_j(\alpha_i))$ and $D_i \in \mathcal{O}_{K\cdot L}$. According to (2.25), $D^2 = \Delta \overset{\text{df}}{=} \operatorname{disc}(\mathcal{O}_K/\mathbb{Z})$, and so

$$\Delta \cdot x_i = DD_i \in \mathcal{O}_{K\cdot L}.$$

But $\Delta x_i = \sum \frac{\Delta a_{ij}}{r}\beta_j$, and the β_js form an integral basis for $\mathcal{O}_L$, and so $\frac{\Delta a_{ij}}{r} \in \mathbb{Z}$. Hence $r|\Delta a_{ij}$ all i, j, and, because of our assumption on r and the a_{ij}s, this implies that $r|\Delta$.

Similarly, $r|\operatorname{disc}(\mathcal{O}_L/\mathbb{Z})$, and so r divides the greatest common divisor of $\operatorname{disc}(\mathcal{O}_K/\mathbb{Z})$ and $\operatorname{disc}(\mathcal{O}_L/\mathbb{Z})$. □

REMARK 6.6. (a) Statement (c) of the theorem shows that if $p|n$ then p ramifies unless $\varphi(p^r) = 1$. Since $\varphi(p^r) = p^{r-1}(p-1)$, this can only happen if $p^r = 2$. Thus $p|n \Rightarrow p$ ramifies in $\mathbb{Q}[\zeta_n]$ except when $p = 2$ and $n = 2 \cdot$ (odd number). Note that $\mathbb{Q}[\zeta_n] = \mathbb{Q}[\zeta_{2n}]$ if n is odd.

(b) In the situation of the lemma,

$$\operatorname{disc}(KL/\mathbb{Q}) = \operatorname{disc}(K/\mathbb{Q})^{[L\colon \mathbb{Q}]} \cdot \operatorname{disc}(L/\mathbb{Q})^{[K\colon \mathbb{Q}]},$$

provided the discriminants on the right are relatively prime. The example $\mathbb{Q}[i,\sqrt{5}] = \mathbb{Q}[i]\cdot\mathbb{Q}[\sqrt{-5}]$ shows that the condition is necessary, because the extensions have discriminants $4^2 5^2$, 4, and 20 respectively. Using this, one can show that, for ζ_n a primitive n^{th} root of 1,

$$\operatorname{disc}(\mathbb{Q}[\zeta_n]/\mathbb{Q}) = (-1)^{\varphi(n)/2} n^{\varphi(n)} / \prod_{p|n} p^{\varphi(n)/(p-1)}.$$

Class numbers of cyclotomic fields. Let ζ be a primitive p^{th} root of 1, p an odd prime. It is known that the class number of $\mathbb{Q}[\zeta]$ grows quite rapidly with p, and that in fact the class number is 1 if and only if $p \le 19$.

Here is how to prove that $\mathbb{Q}[\zeta]$ has class number > 1 when $p = 23$. The Galois group of $\mathbb{Q}[\zeta]$ over $\mathbb{Q}$ is cyclic of order 22, and therefore has a unique subgroup of index 2. Hence $\mathbb{Q}[\zeta]$ contains a unique quadratic extension K of $\mathbb{Q}$. Since 23 is the only prime ramifying in $\mathbb{Q}[\zeta]$, it must also be the only prime ramifying in K, and this implies that $K = \mathbb{Q}[\sqrt{-23}]$. One checks that (2) splits in $\mathbb{Q}[\sqrt{-23}]$, say $(2) = \mathfrak{pq}$, that $\mathfrak{p}$ is not principal, and that $\mathfrak{p}^3$ is principal. Let $\mathfrak{P}$ be a prime ideal of $\mathbb{Z}[\zeta]$ lying over $\mathfrak{p}$. Then $\mathcal{N}\mathfrak{P} = \mathfrak{p}^f$, where f is the residue class degree. Since f divides $[\mathbb{Q}[\zeta] : \mathbb{Q}[\sqrt{-23}]] = 11$, we see that $f = 1$ or 11 (in fact, $f = 11$). In either case, $\mathfrak{p}^f$ is not principal, and this implies that $\mathfrak{P}$ is not principal, because the norm of a principal ideal is principal.

Because of the connection to Fermat's last theorem, primes p such that p does not divide the class number of $\mathbb{Q}[\zeta]$ are of particular interest. They are called *regular*. Kummer found a simple test for when a prime is regular: define the Bernoulli numbers B_n by the formula

$$\frac{t}{e^t - 1} = \sum_{n=0}^{\infty} B_n \frac{t^n}{n!}, \quad B_n \in \mathbb{Q};$$

then p is regular if and only if p divides the numerator of some B_k with $k = 2, 4, \ldots, p-3$. It has long between known that (unfortunately) there are infinitely many irregular primes, and it still not proved that there are infinitely many regular primes (although the first case of Fermat's theorem is known for infinitely many primes). It is expected that 61% of primes are regular and 39% are irregular.

Units in cyclotomic fields. Let ζ be a primitive n^{th} root of 1, $n > 2$. Define

$$\mathbb{Q}[\zeta]^+ = \mathbb{Q}[\zeta + \zeta^{-1}].$$

For example, if $\zeta = e^{2\pi i/n}$, then $\mathbb{Q}[\zeta]^+ = \mathbb{Q}[\cos\frac{2\pi}{n}]$. Under any embedding of $\mathbb{Q}[\zeta]$ into $\mathbb{C}$, ζ^{-1} maps to the complex conjugate of ζ, and therefore the image of $\mathbb{Q}[\zeta]^+$ is fixed under complex conjugation and hence lies in $\mathbb{R}$. Thus $\mathbb{Q}[\zeta]^+$ has $\varphi(n)/2$ real embeddings (and no nonreal embeddings), whereas $\mathbb{Q}[\zeta]$ has $\varphi(n)$ nonreal complex embeddings. Therefore the unit theorem (5.1) shows that the groups of units in $\mathbb{Q}[\zeta]$ and $\mathbb{Q}[\zeta]^+$ have the same rank, and so, if u is a unit in $\mathbb{Q}[\zeta]$, then $u^m \in \mathbb{Q}[\zeta]^+$ for some m. In fact a more precise result is known.

PROPOSITION 6.7. *Assume n is a prime power; then every unit $u \in \mathbb{Q}[\zeta]$ can be written*

$$u = \zeta \cdot v$$

with ζ a root of unity and v a unit in $\mathbb{Q}[\zeta]^+$.

PROOF. This is not difficult — see Fröhlich and Taylor 1991, VI.1.19, or Washington 1982. □

Fermat's last theorem. Fermat's last theorem is known to be true for p a regular prime. Here we prove a weaker result, known as the *first case* of Fermat's last theorem.

THEOREM 6.8. *Let p be an odd prime. If the class number of $\mathbb{Q}[\zeta]$ is not divisible by p, then there is no integer solution (x, y, z) to*

$$X^p + Y^p = Z^p$$

with $p \nmid xyz$.

Let (x, y, z) be a solution of Fermat's equation with $p \nmid xyz$. After removing any common factor, we may suppose that $\gcd(x, y, z) = 1$.

We first treat the case $p = 3$. The only cubes modulo 9 are -1, 0, 1, and so

$$x^3 + y^3 \equiv -2, 0, 2 \mod 9, \quad z^3 \equiv -1, 1 \mod 9,$$

which are contradictory. Similarly we may eliminate the case $p = 5$ by looking modulo 25. Henceforth we assume $p > 5$.

If $x \equiv y \equiv -z \mod p$, then $-2z^p \equiv z^p$ and $p|3z$, contradicting our hypotheses. Hence one of the congruences can't hold, and after rewriting the equation $x^p+(-z)^p = (-y)^p$ if necessary, we may assume that $p \nmid x - y$.

The roots of $X^p + 1$ are $-1, -\zeta, \ldots, -\zeta^{p-1}$, and so

$$X^p + 1 = \prod_{i=0}^{p-1}(X + \zeta^i).$$

Hence

$$\prod_{i=0}^{p-1}(x + \zeta^i y) = z^p.$$

The idea of the proof is to exploit this factorization and what we know of the arithmetic of $\mathbb{Q}[\zeta]$ to obtain a contradiction.

Let $\mathfrak{p}$ be the unique prime ideal of $\mathbb{Z}[\zeta]$ dividing (p); thus (see 6.2) $\mathfrak{p} = (1 - \zeta^i)$ for any i, $1 \le i \le p - 1$.

LEMMA 6.9. *The elements $x + \zeta^i y$ of $\mathbb{Z}[\zeta]$ are relatively prime in pairs.*

PROOF. We have to show that there does not exist a prime ideal $\mathfrak{q}$ dividing $x+\zeta^i y$ and $x + \zeta^j y$ for $i \ne j$. Suppose there does. Then $\mathfrak{q}|((\zeta^i - \zeta^j)y) = \mathfrak{p}y$, and $\mathfrak{q}|((\zeta^j - \zeta^i)x) = \mathfrak{p}x$. By assumption, x and y are relatively prime, and therefore $\mathfrak{q} = \mathfrak{p}$. Thus $x+y \equiv x+\zeta^i y \equiv 0 \mod \mathfrak{p}$. Hence $x+y \in \mathfrak{p} \cap \mathbb{Z} = (p)$. But $z^p = x^p + y^p \equiv x + y \equiv 0 \mod p$, and so $p|z$, which contradicts our hypotheses. □

LEMMA 6.10. *For any $\alpha \in \mathbb{Z}[\zeta]$, $\alpha^p \in \mathbb{Z} + p\mathbb{Z}[\zeta]$.*

PROOF. Write

$$\alpha = a_0 + a_1\zeta + \cdots + a_{p-2}\zeta^{p-2}, \quad a_i \in \mathbb{Z}.$$

Then

$$\alpha^p \equiv a_0^p + a_1^p + \cdots + a_{p-1}^p \mod p,$$

which lies in $\mathbb{Z}$. □

LEMMA 6.11. *Suppose $\alpha = a_0 + a_1\zeta + \cdots + a_{p-1}\zeta^{p-1}$ with $a_i \in \mathbb{Z}$ and at least one $a_i = 0$. If α is divisible by an integer n, i.e., if $\alpha \in n\mathbb{Z}[\zeta]$, then each a_i is divisible by n.*

PROOF. Since $1 + \zeta + \cdots + \zeta^{p-1} = 0$, any subset of $\{1, \zeta, \ldots, \zeta^{p-1}\}$ with $p-1$ elements will be a $\mathbb{Z}$-basis for $\mathbb{Z}[\zeta]$. The result is now obvious. □

We can now complete the proof of Theorem 6.8. Consider the equation

$$\prod_{i=0}^{p-1}(x + \zeta^i y) = (z)^p$$

as an equality of ideals in $\mathbb{Z}[\zeta]$. Since the factors on the left are relatively prime in pairs, each one must be the p^{th} power of an ideal, say

$$(x + \zeta^i y) = \mathfrak{a}_i^p$$

for some ideal $\mathfrak{a}_i$ in $\mathbb{Z}[\zeta]$. This equation says $\mathfrak{a}_i$ has order dividing p in the class group, but we are assuming that the class group of $\mathbb{Z}[\zeta]$ is of order prime to p, and so $\mathfrak{a}_i$ itself is principal, say $\mathfrak{a}_i = (\alpha_i)$.

Take $i = 1$, and omit subscripts. Then we have that $x + \zeta y = u\alpha^p$ for some unit u in $\mathbb{Z}[\zeta]$. We apply (6.7) to write $u = \zeta^r v$ where $\bar{v} = v$. According to (6.10), there is an $a \in \mathbb{Z}$ such that $\alpha^p \equiv a \mod p$. Therefore

$$x + \zeta y = \zeta^r v\alpha^p \equiv \zeta^r va \mod p.$$

Also

$$x + \bar{\zeta} y = \zeta^{-r} v\bar{\alpha}^p \equiv \zeta^{-r} va \mod p.$$

On combining these statements, we find that

$$\zeta^{-r}(x + \zeta y) \equiv \zeta^r(x + \zeta^{-1}y) \mod p,$$

or

$$x + \zeta y - \zeta^{2r}x - \zeta^{2r-1}y \equiv 0 \mod p. \qquad (*)$$

If $1, \zeta, \zeta^{2r-1}, \zeta^{2r}$ are distinct, then, because $p \geq 5$, Lemma 6.11 implies that p divides x and y, which is contrary to our original assumption. The only remaining possibilities are:

(a) $1 = \zeta^{2r}$; but then (*) says

$$\zeta y - \zeta^{-1}y \equiv 0 \mod p,$$

and Lemma 6.11 implies $p|y$, which contradicts our original assumption.

(b) $1 = \zeta^{2r-1}$; then $\zeta = \zeta^{2r}$, and (*) says

$$(x - y) - (x - y)\zeta \equiv 0 \mod p,$$

and Lemma 6.11 implies that $p|x - y$, which contradicts the choice of x and y made at the start of the proof.

(c) $\zeta = \zeta^{2r-1}$; but then (*) says

$$x - \zeta^2 x \equiv 0 \mod p,$$

and Lemma 6.11 implies that $p|x$, which contradicts our original assumption.

This completes the proof.

7. Valuations; Local Fields

In this section, we define the notion of a valuation and study the completions of number fields with respect to valuations.

Valuations. A *(multiplicative) valuation* on a field K is a function $x \mapsto |x| \colon K \to \mathbb{R}$ such that

(a) $|x| > 0$ except that $|0| = 0$;
(b) $|xy| = |x||y|$
(c) $|x + y| \leq |x| + |y|$ (triangle inequality).

If the stronger condition

(c′) $|x + y| \leq \max\{|x|, |y|\}$

holds, then $|\ |$ is called a *nonarchimedean valuation.*

Note that (a) and (b) imply that $|\ |$ is a homomorphism $K^\times \to \mathbb{R}_{>0}$ (multiplicative group of positive real numbers). Since $\mathbb{R}_{>0}$ is torsion-free, $|\ |$ maps all roots of unity in $K^\times$ to 1. In particular, $|-1| = 1$, and $|-x| = |x|$ for all x.

Example 7.1. (a) For any number field K, and embedding $\sigma \colon K \hookrightarrow \mathbb{C}$, we get a valuation on K by putting $|a| = |\sigma a|$.

(b) Let $\mathrm{ord} \colon K^\times \to \mathbb{Z}$ be an (additive) discrete valuation, and let e be a real number with $e > 1$; then

$$|a| = (1/e)^{\mathrm{ord}(a)}, \quad a \neq 0, \quad |0| = 0$$

is a nonarchimedean valuation on K. For example, for any prime number p, we have the *p-adic valuation* $|\ |_p$ on $\mathbb{Q}$:

$$|a|_p = (1/e)^{\mathrm{ord}_p(a)}.$$

Usually we normalize this by taking $e = p$; thus

$$|a|_p = (1/p)^{\mathrm{ord}_p(a)} = 1/p^r \text{ if } a = a_0 \cdot p^r \text{ with } \mathrm{ord}_p(a_0) = 0.$$

Similarly, for any prime ideal $\mathfrak{p}$ in a number field K, we have a *normalized $\mathfrak{p}$-adic valuation*

$$|a|_\mathfrak{p} = (1/\mathbb{N}\mathfrak{p})^{\mathrm{ord}_\mathfrak{p}(a)}.$$

(c) On any field we can define the *trivial valuation*: $|a| = 1$ for all $a \neq 0$. When K is finite, there is no other (because *all* nonzero elements of a finite field are roots of 1).

Nonarchimedean valuations. Recall that this means that, instead of the triangle inequality, we have

$$|x + y| \leq \max\{|x|, |y|\}.$$

By induction, this condition implies that

$$|\sum x_i| \leq \max\{|x_i|\}. \qquad (*)$$

Proposition 7.2. *A valuation $|\ |$ is nonarchimedean if and only if it takes bounded values on $\{m1 \mid m \in \mathbb{Z}\}$.*

PROOF. If $|\ |$ is nonarchimedean, then, for $m > 0$,

$$|m1| = |1 + 1 + \cdots + 1| \leq |1| = 1.$$

As we noted above, $|-1| = |1|$, and so $|-m1| = |m1| \leq 1$.

Conversely, suppose $|m1| \leq N$ for all m. Then

$$|x+y|^n = |\ \sum \binom{n}{r} x^r y^{n-r}\ | \leq \sum_r\ |\binom{n}{r}|\ |x|^r |y|^{n-r}.$$

Clearly $|x|^r|y|^{n-r} \leq \max\{|x|^n, |y|^n\} = \max\{|x|, |y|\}^n$ and $\binom{n}{r}$ is an integer, and so

$$|x+y|^n \leq N(n+1)\max\{|x|, |y|\}^n.$$

On taking n^{th} roots we find that

$$|x+y| \leq N^{1/n}(n+1)^{1/n}\max\{|x|, |y|\}.$$

When we let $n \to \infty$, the terms involving n tend to 1 (to see this, take logs). □

COROLLARY 7.3. *If char $K \neq 0$, then K has only nonarchimedean valuations.*

PROOF. In this case, the set $\{m \cdot 1 \mid m \in \mathbb{Z}\}$ is finite. □

ASIDE 7.4. The classical *archimedean axiom* states that if a and b are nonzero elements of K, then there is an $n \in \mathbb{N}$ such that $|b| < |na|$. The proposition shows that the nonarchimedean valuations are precisely those for which the archimedean axiom fails, whence the name.

As we noted above, a discrete (additive) valuation ord on K determines a valuation by

$$|x| = e^{-\mathrm{ord}(x)},$$

any $e > 1$. Taking logs gives $\log_e |x| = -\mathrm{ord}(x)$, or $\mathrm{ord}(x) = -\log_e |x|$. This suggests how we might pass from multiplicative valuations to additive valuations.

PROPOSITION 7.5. *Let $|\ |$ be a nontrivial nonarchimedean valuation, and put $v(x) = -\log|x|$, $x \neq 0$ (log to base e for any real $e > 1$). Then $v\colon K^\times \to \mathbb{R}$ satisfies the following conditions:*

(a) $v(xy) = v(x) + v(y)$;
(b) $v(x+y) \geq \min\{v(x), v(y)\}$.

If $v(K^\times)$ is a discrete in $\mathbb{R}$, then it is a multiple of a discrete valuation ord: $K^\times \twoheadrightarrow \mathbb{Z} \subset \mathbb{R}$.

PROOF. That v satisfies (a) and (b) is obvious. For the last statement, note that $v(K^\times)$ is a subgroup of $\mathbb{R}$ (under addition). If it is a discrete subgroup, then it is a lattice (by 4.14), which means that $v(K^\times) = \mathbb{Z}c$ for some c. Now $\mathrm{ord} \stackrel{\mathrm{df}}{=} c^{-1} \cdot v$ is an additive discrete valuation $K^\times \twoheadrightarrow \mathbb{Z}$. □

We shall say $|\ |$ is *discrete* when $|K^\times|$ is a discrete subgroup of $\mathbb{R}_{>0}$. Note that, even when $|K^\times|$ is discrete in $\mathbb{R}$, $|K|$ usually won't be, because 0 will be a limit point for the set $|K^\times|$. For example, $|p^n|_p = p^{-n}$, which converges to 0 as $n \to \infty$.

PROPOSITION 7.6. *Let $|\ |$ be a nonarchimedean valuation. Then*

$A \stackrel{df}{=} \{a \in K \mid |a| \leq 1\}$ *is a subring of K, with*

$U \stackrel{df}{=} \{a \in K \mid |a| = 1\}$ *as its group of units, and*

$\mathfrak{m} \stackrel{df}{=} \{a \in K \mid |a| < 1\}$ *as its unique maximal ideal.*

The valuation $|\ |$ is discrete if and only if $\mathfrak{m}$ is principal, in which case A is a discrete valuation ring.

PROOF. The first assertion is obvious. If $|\ |$ is discrete, then A and $\mathfrak{m}$ are the pair associated (as in 3.28) with the additive valuation $-\log|\ |$, and so A is a discrete valuation ring and $\mathfrak{m}$ is generated by any element $\pi \in K^\times$ such that $|\pi|$ is the largest element of $|K^\times|$ less than one. Conversely, if $\mathfrak{m} = (\pi)$, then $|K^\times|$ is the subgroup of $\mathbb{R}_{>0}$ generated by $|\pi|$. □

REMARK 7.7. There do exist nondiscrete nonarchimedean valuations. For example, let $\mathbb{Q}^{\mathrm{al}}$ be an algebraic closure of $\mathbb{Q}$. We shall see later that the p-adic valuation $|\ |_p : \mathbb{Q} \to \mathbb{R}$ extends to $\mathbb{Q}^{\mathrm{al}}$ (in many different ways). Since $\mathbb{Q}^{\mathrm{al}}$ contains an element $p^{1/n}$ for all n, we see that $|\mathbb{Q}^{\mathrm{al}\times}| \ni (p^{-1})^{1/n} = 1/\sqrt[n]{p}$ for all n, and $1/\sqrt[n]{p} \to 1$ as $n \to \infty$. In fact, one can show that $|\mathbb{Q}^{\mathrm{al}\times}| = \{p^r \mid r \in \mathbb{Q}\}$, which is not discrete in $\mathbb{R}_{>0}$.

Equivalent valuations. Note that a valuation $|\ |$ defines a metric on K, with distance function

$$d(a,b) = |a-b|,$$

and hence a topology on K. In more detail, for $a \in K$, the sets

$$U(a,\varepsilon) = \{x \in K \mid |x-a| < \varepsilon\}, \quad \varepsilon > 0,$$

form a fundamental system of open neighbourhoods of a. A set is open if and only if it is a union of sets of the form $U(a,\varepsilon)$.

For example, for the topology on $\mathbb{Q}$ defined by $|\ |_p$, a and b are close if their difference is divisible by a high power of p. In particular, the sequence

$$1, p, p^2, \dots, p^n, \dots$$

converges to 0.

The topology defined by the $\mathfrak{p}$-adic valuation $|\ |_{\mathfrak{p}}$ is called the $\mathfrak{p}$-*adic topology* on K.

PROPOSITION 7.8. *Let $|\ |_1$, $|\ |_2$ be valuations on K, with $|\ |_1$ nontrivial. The following conditions are equivalent:*

(a) $|\ |_1$, $|\ |_2$ *define the same topology on K;*
(b) $|\alpha|_1 < 1 \Rightarrow |\alpha|_2 < 1$;
(c) $|\ |_2 = |\ |_1^a$ *for some $a > 0$.*

PROOF. (a) ⇒ (b): Since $|\alpha^n| = |\alpha|^n$, clearly $\alpha^n \to 0$ if and only if $|\alpha| < 1$. Therefore (a) implies that

$$|\alpha|_1 < 1 \iff |\alpha|_2 < 1.$$

(b) $\Rightarrow$ (c): Because $|\ |_1$ is nontrivial, there exists a $y \in K$ such that $|y| > 1$. Let

$$a = \log |y|_2 / \log |y|_1,$$

so that

$$\log |y|_2 = a \cdot \log |y|_1,$$

or

$$|y|_2 = |y|_1^a.$$

Now let x be any nonzero element of K. There is a real number b such that

$$|x|_1 = |y|_1^b.$$

To prove (c), it suffices to prove that

$$|x|_2 = |y|_2^b,$$

because then

$$|x|_2 = |y|_2^b = |y|_1^{ab} = |x|_1^a.$$

Let m/n, $n > 0$, be a rational number $> b$. Then

$$|x|_1 = |y|_1^b < |y|_1^{\frac{m}{n}}$$

and so

$$|x^n/y^m|_1 < 1.$$

From our assumption (b), this implies that

$$|x^n/y^m|_2 < 1$$

and so

$$|x|_2 < |y|_2^{\frac{m}{n}}.$$

This is true for all rational numbers $\frac{m}{n} > b$, and so

$$|x|_2 \leq |y|_2^b.$$

A similar argument with rational numbers $\frac{m}{n} < b$ shows that

$$|x|_2 \geq |y|_2^b,$$

and so we have equality, which completes the proof of (a). □

Two valuations are said to be *equivalent* if they satisfy the conditions of the proposition.

Properties of discrete valuations. We make some easy, but important, observations about discrete valuations.

(7.9.1) For an additive valuation, we are given that

$$\mathrm{ord}(a+b) \geq \min\{\mathrm{ord}(a), \mathrm{ord}(b)\}$$

and we checked (3.27 et seq.) that this implies that equality holds if $\mathrm{ord}(a) \neq \mathrm{ord}(b)$. For multiplicative valuations, we are given that

$$|a+b| \leq \max\{|a|, |b|\},$$

and a similar argument shows that equality holds if $|a| \neq |b|$. This has the following consequences.

(7.9.2) Recall that we define a metric on K by setting $d(a,b) = |a-b|$. I claim that if x is closer to b than it is to a, then $d(a,x) = d(a,b)$. For we are given that

$$|x-b| < |x-a|,$$

and this implies that

$$|b-a| = |b-x+x-a| = |x-a|.$$

(7.9.3) Suppose

$$a_1 + a_2 + \cdots + a_n = 0.$$

Then an argument as in the subsection on Eisenstein extensions (end §3) shows that the maximum value of the summands must be attained for at least two values of the subscript.

Complete list of valuations for $\mathbb{Q}$. We now give a complete list of the valuations on $\mathbb{Q}$ (up to equivalence)**.** We write $|\ |_\infty$ for the valuation on $\mathbb{Q}$ defined by the usual absolute value on $\mathbb{R}$, and we say that $|\ |_\infty$ is *normalized*.

THEOREM 7.10 (Ostrowski). *Let $|\ |$ be a nontrivial valuation on $\mathbb{Q}$.*

(a) *If $|\ |$ is archimedean, then $|\ |$ is equivalent to $|\ |_\infty$.*
(b) *If $|\ |$ is nonarchimedean, then it is equivalent to $|\ |_p$ for exactly one prime p.*

PROOF. Let m, n be integers > 1. Then we can write

$$m = a_0 + a_1 n + \cdots + a_r n^r$$

with the a_i integers, $0 \leq a_i < n$, $n^r \leq m$. Let $N = \max\{1, |n|\}$. By the triangle inequality,

$$|m| \leq \sum |a_i||n|^i \leq \sum |a_i| N^r.$$

We know

$$r \leq \log(m)/\log(n),$$

(log relative to some $e > 1$) and the triangle inequality shows that

$$|a_i| \leq |1 + \cdots + 1| = a_i|1| = a_i \leq n.$$

On putting these into the first inequality, we find that

$$|m| \leq (1+r) n N^r \leq (1 + \frac{\log m}{\log n}) n N^{\frac{\log m}{\log n}}.$$

In this inequality, replace m with m^t (t an integer), and take t^{th} roots:

$$|m| \leq \left(1 + \frac{t \log m}{\log n}\right)^{\frac{1}{t}} n^{\frac{1}{t}} N^{\frac{\log m}{\log n}}.$$

Now let $t \to \infty$. The terms involving t tend to 1, and so

$$|m| \leq N^{\frac{\log m}{\log n}}. \qquad (*)$$

Case (i): For all integers $n > 1$, $|n| > 1$.

In this case $N = |n|$, and (*) yields:

$$|m|^{1/\log m} \leq |n|^{1/\log n}.$$

By symmetry, we must have equality, and so there is an $c > 1$ such that

$$c = |m|^{1/\log m} = |n|^{1/\log n}$$

for all integers $m, n > 1$. Hence

$$|n| = c^{\log n} = e^{\log c \log n} = n^{\log c}, \text{ all integers } n > 1.$$

Let $a = \log c$, and rewrite this

$$|n| = |n|_\infty^a, \text{ all integers } n > 1,$$

where $|\ |_\infty$ is the usual absolute value on $\mathbb{Q}$. Since both $|\ |$ and $|\ |_\infty^a$ are homomorphisms $\mathbb{Q}^\times \to \mathbb{R}_{>0}$, the fact that they agree on a set of generators for the group $\mathbb{Q}^\times$ (the primes and -1) implies that they agree on all of $\mathbb{Q}^\times$.

Case (ii): For some $n > 1$, $|n| \leq 1$.

In this case, $N = 1$, and (*) implies $|m| \leq 1$ for all integers m. Therefore the valuation is nonarchimedean. Let A be the associated local ring and $\mathfrak{m}$ its maximal ideal. From the definition of A, we know that $\mathbb{Z} \subset A$. Then $\mathfrak{m} \cap \mathbb{Z}$ is a prime ideal in $\mathbb{Z}$ (because $\mathfrak{m}$ is a prime ideal), and it is nonzero for otherwise the valuation would be trivial. Hence $\mathfrak{m} \cap \mathbb{Z} = (p)$ for some p. This implies that $|m| = 1$ if m is an integer not divisible by p, and so $|np^r| = |p|^r$ if n is a rational number whose numerator and denominator are not divisible by p. If a is such that $|p| = (1/p)^a$; then $|x| = |x|_p^a$ for all $x \in \mathbb{Q}$. □

THEOREM 7.11 (Product Formula). *For $p = 2, 3, 5, 7, ..., \infty$, let $|\ |_p$ be the corresponding normalized valuation on $\mathbb{Q}$. For any nonzero rational number a*

$$\prod |a|_p = 1 \text{ (product over all } p \text{ including } \infty).$$

PROOF. Let $\alpha = a/b$, $a, b \in \mathbb{Z}$. Then $|\alpha|_p = 1$ unless $p|a$ or $p|b$. Therefore $|\alpha|_v = 1$ for all but finite many v's, and so the product is really finite.

Let $\pi(a) = \prod |a|_v$. Then π is a homomorphism $\mathbb{Q}^\times \to \mathbb{R}^\times$, and so it suffices to show that $\pi(-1) = 1$ and $\pi(p) = 1$ for each prime number p. The first is obvious, because $|-1| = 1$ for all valuations $|\ |$. For the second, note that

$$|p|_p = 1/p, \quad |p|_q = 1, \quad q \text{ a prime } \neq p, \quad |p|_\infty = p.$$

The product of these numbers is 1. □

The primes of a number field. Let K be an algebraic number field. An equivalence class of valuations on K is called a *prime* of K.

THEOREM 7.12. *Let K be an algebraic number field. There exists exactly one prime of K*

(a) *for each prime ideal $\mathfrak{p}$;*
(b) *for each real embedding;*
(c) *for each conjugate pair of complex embeddings.*

PROOF. See §8. □

In each equivalence class of valuations of K we select a normalized valuation[13] as follows:

for a prime ideal $\mathfrak{p}$ of $\mathcal{O}_K$, $|a|_{\mathfrak{p}} = (1/\mathbb{N}\mathfrak{p})^{\mathrm{ord}_{\mathfrak{p}}(a)} = (\mathcal{O}_{\mathfrak{p}} : (a))^{-1}$;

for a real embedding $\sigma\colon K \hookrightarrow \mathbb{R}$, $|a| = |\sigma a|$;

for a nonreal complex embedding $\sigma\colon K \hookrightarrow \mathbb{C}$, $|a| = |\sigma a|^2$.

Note that this last is not actually a valuation, because it doesn't satisfy the triangle law. There are various ways of getting around this problem the best of which is simply to ignore it.

Notations. We generally write v for a prime. If it corresponds to a prime ideal $\mathfrak{p}$ of $\mathcal{O}_K$, then we call it a *finite prime*, and we write $\mathfrak{p}_v$ for the ideal. If it corresponds to a (real or nonreal) embedding of K, then we call it an infinite (real or complex) prime. We write $|\ |_v$ for a valuation in the equivalence class. If $L \supset K$ and w and v are primes of L and K such that $|\ |_w$ restricted to K is equivalent to $|\ |_v$, then we say that w *divides* v, or w *lies over* v, and we write $w|v$. For a finite prime, this means $\mathfrak{P}_w \cap \mathcal{O}_K = \mathfrak{p}_v$ or, equivalently, that $\mathfrak{P}_w$ divides $\mathfrak{p}_v \cdot \mathcal{O}_L$. For an infinite prime, it means that w corresponds to an embedding $\sigma\colon L \hookrightarrow \mathbb{C}$ that extends the embedding corresponding to v (or its complex conjugate).

THEOREM 7.13 (Product Formula). *For each prime v, let $|\ |_v$ be the normalized valuation. For any nonzero $\alpha \in K$,*

$$\prod |\alpha|_v = 1 \text{ (product over all primes of } K).$$

PROOF. The product formula for a general number field follows from the product formula for $\mathbb{Q}$ and the next result. □

LEMMA 7.14. *Let L be a finite extension of a number field K.*

(a) *Each prime on K extends to a finite number of primes of L.*
(b) *For any prime v of K and $\alpha \in L^\times$,*

$$\prod_{w|v} |\alpha|_w = |\operatorname{Nm}_{L/K} \alpha|_v.$$

[13]These are the most natural definitions for which the product formula hold. Alternatively, let K_v be the completion of K with respect to the valuation v, and let μ be a Haar measure on $(K_v, +)$ — it is uniquely determined up to a nonzero constant. For any nonzero $a \in K_v$, $\mu_a(U) \overset{\text{df}}{=} \mu(aU)$ is also a Haar measure on $(K_v, +)$, and so $\mu_a = c(a)\mu$ for some constant $c(a)$. In fact, $c(a) = |a|$, the normalized valuation of a.

PROOF. See §8. □

REMARK 7.15. The product formula is true in two other important situations.

(a) Let K be a finite extension of $k(T)$ where k is a finite field. According to (7.3), the valuations of K are all discrete, and hence correspond to discrete valuation rings in K. As in the number field case, we can normalize a valuation by setting $|a|_v = (1/\mathbb{N}v)^{\mathrm{ord}_v(a)}$ where $\mathbb{N}v$ is the number of elements in the residue field of the discrete valuation ring and $\mathrm{ord}_v\colon K^\times \twoheadrightarrow \mathbb{Z}$. Then $\prod_v |a|_v = 1$. The proof of this is easy when $K = k(T)$, and the general case is obtained by means of a result like (7.14).

(b) Let K be a finite extension of $k(T)$ where k is an algebraically closed field. In this case we only look at primes that are trivial when restricted to k. All such primes are nonarchimedean, and hence correspond to discrete valuations $\mathrm{ord}_v\colon K^\times \twoheadrightarrow \mathbb{Z}$. Fix an $e > 1$, and define $|a|_v = (1/e)^{\mathrm{ord}_v(a)}$ for every v. Then $\prod |a|_v = 1$ for all $a \in K^\times$. This of course is equivalent to the statement

$$\sum \mathrm{ord}_v(a) = 0.$$

For example, let X be a compact Riemann surface, and let K be the field of meromorphic functions on X. For each point P of X we have a discrete valuation, defined by $\mathrm{ord}_P(f) = m$ or $-m$ according as f has a zero or pole of order m at P. The valuations ord_P are precisely the valuations on K trivial on $\mathbb{C} \subset K$, and so the product formula for K is simply the statement that f has as many zeros as poles.

The proof of this runs as follows: the Cauchy integral formula implies that if f is a nonconstant meromorphic function on an open set U in $\mathbb{C}$, and Γ is the oriented boundary of a compact set C contained in U, then

$$\int_\Gamma \frac{f'(z)}{f(z)} dz = 2\pi i(Z - P)$$

where Z is the number of zeros of f in C and P is the number of poles of f, both counted with multiplicities. This formula also holds for compact subsets of manifolds. If the manifold M is itself compact, then we can take $C = M$, which has no boundary, and so the formula becomes

$$Z - P = 0,$$

i.e.,

$$\sum \mathrm{ord}_P(f) = 0, \quad P \in M.$$

Completions. Let K be a field with a nontrivial valuation. A sequence (a_n) of elements in K is called a *Cauchy sequence* if, for every $\varepsilon > 0$, there is an N such that

$$|a_n - a_m| < \varepsilon, \text{ all } m, n > N.$$

The field K is said to be *complete* if every Cauchy sequence has a limit in K. (The limit is necessarily unique.)

EXAMPLE 7.16. Consider the sequence in $\mathbb{Z}$

$$4, 34, 334, 3334, \dots .$$

As

$$|a_m - a_n|_5 = 5^{-n} \quad (m > n),$$

this is a Cauchy sequence for the 5-adic topology on $\mathbb{Q}$. Note that

$$3 \cdot 4 = 12, \quad 3 \cdot 34 = 102, \quad 3 \cdot 334 = 1002, \quad 3 \cdot 3334 = 10002, \ldots$$

and so $3 \cdot a_n - 2 \to 0$ as $n \to \infty$. Thus $\lim_{n\to\infty} a_n = 2/3 \in \mathbb{Q}$.

There is a similar notion of Cauchy series. For example, any series of the form

$$a_{-n}p^{-n} + \cdots + a_0 + a_1 p + \cdots + a_m p^m + \cdots, \quad 0 \le a_i < p,$$

is a Cauchy series in $\mathbb{Q}$ for the p-adic topology.

THEOREM 7.17. *Let K be a field with a valuation $|\ |$. Then there exists a complete valued field $(\hat{K}, |\ |)$ and a homomorphism $K \to \hat{K}$ preserving the valuation that is universal in the following sense: any homomorphism $K \to L$ from K into a complete valued field $(L, |\ |)$ preserving the valuation, extends uniquely to a homomorphism $\hat{K} \to L$. The image of K in $\hat{K}$ is dense.*

PROOF. (Sketch) The uniqueness of $(\hat{K}, |\ |)$ is obvious from the universal property. Let $\bar{K}$ be the closure of K in $\hat{K}$. Then $\bar{K}$ is complete, and so the homomorphism $K \to \bar{K}$ extends to $\hat{K}$ — this implies that $\bar{K} = \hat{K}$, and so K is dense in $\hat{K}$.

We now construct $\hat{K}$. Every point of $\hat{K}$ will be the limit of a sequence of points in K, and the sequence will be Cauchy. Two Cauchy sequences will converge to the same point in $\hat{K}$ if and only if they are *equivalent* in the sense that

$$\lim_{n\to\infty} |a_n - b_n| = 0.$$

This suggests defining $\hat{K}$ to be the set of equivalence classes of Cauchy sequences in K. Define addition and multiplication of Cauchy sequences in the obvious way, and verify that $\hat{K}$ is a field. There is a canonical map $K \to \hat{K}$ sending a to the constant Cauchy sequence $a, a, a, \ldots$, which we use to identify K with a subfield of $\hat{K}$. We can extend a homomorphism from K into a second complete valued field L to $\hat{K}$ by mapping the limit of a Cauchy sequence in $\hat{K}$ to its limit in L. □

For a prime v of K, we write K_v for the completion of K with respect to v. When v corresponds to a prime ideal $\mathfrak{p}$, we write $K_\mathfrak{p}$ for the completion, and $\hat{\mathcal{O}}_\mathfrak{p}$ for the ring of integers in $K_\mathfrak{p}$. For example, $\mathbb{Q}_p$ is the completion of $\mathbb{Q}$ with respect to the p-adic valuation $|\ |_p$. We write $\mathbb{Z}_p$ (not $\hat{\mathbb{Z}}_p$) for the ring of integers in $\mathbb{Q}_p$ (the ring of p-adic integers).

Completions in the nonarchimedean case. Let $|\ |$ be a discrete nonarchimedean valuation on K, and let π be an element of K with largest value < 1 (therefore π generates the maximal ideal $\mathfrak{m}$ in the valuation ring A). Such a π is called a *local uniformizing parameter.*

The set of values is

$$|K| = \{c^m \mid m \in \mathbb{Z}\} \cup \{0\}, \quad c = |\pi|.$$

Let $a \in \hat{K}^\times$, and let a_n be a sequence in K converging to a. Then $|a_n| \to |a|$ (because $|\ |$ is a continuous map), and so $|a|$ is a limit point for the set $|K^\times|$. But $|K^\times|$ is closed (being discrete), and so $|a| \in |K^\times|$. Thus $|\hat{K}| = |K|$, and so $|\ |$ is a discrete valuation on $\hat{K}$ also. Let $\mathrm{ord}\colon K^\times \twoheadrightarrow \mathbb{Z}$ be a normalized discrete additive valuation corresponding to $|\ |$; then ord extends to a normalized discrete valuation on $\hat{K}$.

Note that if $a_n \to a \neq 0$, then $|a_n| \to |a| \neq 0$, and (because $|K^\times|$ is discrete), $|a_n| = |a|$ for all n large enough.

The ring associated with $|\ |$ in $\hat{K}$ is

$$\hat{A} = \{a \in \hat{K} \mid |a| \leq 1\}.$$

Clearly $\hat{A}$ is the set of limits of Cauchy sequences in A, and it is therefore the closure of A in $\hat{K}$. The maximal ideal in $\hat{A}$ is

$$\hat{\mathfrak{m}} = \{a \in \hat{K} \mid |a| < 1\}.$$

Again it is the set of limits of Cauchy sequences in $\mathfrak{m}$, and so it is the closure of $\mathfrak{m}$. Similarly, $\hat{\mathfrak{m}}^n$ is the closure of $\mathfrak{m}^n$. Let π be an element with $\mathrm{ord}(\pi) = 1$; then π generates $\mathfrak{m}$ in A and $\hat{\mathfrak{m}}$ in $\hat{A}$.

LEMMA 7.18. *For any n, the map $A/\mathfrak{m}^n \to \hat{A}/\hat{\mathfrak{m}}^n$ is an isomorphism.*

PROOF. Note that

$$\mathfrak{m}^n = \{a \in A \mid |a| \leq |\pi|^n\} = \{a \in A \mid |a| < |\pi|^{n-1}\}$$

is both open and closed in A. Because it is closed, the map is injective; because $\hat{\mathfrak{m}}^n$ is open, the map is surjective. □

PROPOSITION 7.19. *Choose a set S of representatives for $A/\mathfrak{m}$, and let π generate $\mathfrak{m}$. The series*

$$a_{-n}\pi^{-n} + \cdots + a_0 + a_1\pi + \cdots + a_m\pi^m + \cdots, \quad a_i \in S$$

is a Cauchy series, and every Cauchy series is equivalent to exactly one of this form. Thus each element of $\hat{K}$ has a unique representative of this form.

PROOF. Let $s_M = \sum_{i=-n}^{M} a_i\pi^i$. Then

$$|s_M - s_N| \leq |\pi|^{M+1}, \text{ if } M < N,$$

which shows that the sequence s_M is Cauchy. Let $\alpha \in \hat{K}$. Because $|\hat{K}| = |K|$, we can write $\alpha = \pi^n\alpha_0$ with α_0 a unit in $\hat{A}$. From the definition of S, we see that there exists an $a_0 \in S$ such that $\alpha_0 - a_0 \in \hat{\mathfrak{m}}$. Now $\frac{\alpha_0 - a_0}{\pi} \in \hat{A}$, and so there exists an $a_1 \in S$ such that $\frac{\alpha_0 - a_0}{\pi} - a_1 \in \hat{\mathfrak{m}}$. Now there exists an a_2 such that $\frac{\alpha_0 - a_0 - a_1\pi}{\pi^2} - a_2 \in \hat{\mathfrak{m}}$, etc. In the limit,

$$\alpha_0 = a_0 + a_1\pi + \cdots, \quad \alpha = \pi^n\alpha_0.$$

Note that

$$|\textstyle\sum a_i\pi^i| = |\pi^m|$$

if a_m is the first nonzero coefficient. Therefore $\sum a_i\pi^i = 0$ (if and) only if $a_i = 0$ for all i. This proves the uniqueness. □

Thus, for example, every equivalence class of Cauchy sequences in $\mathbb{Q}$ for $|\ |_p$ has a unique representative of the form

$$a_{-n}p^{-n} + \cdots + a_0 + a_1p + a_2p^2 + \cdots, \quad 0 \le a_i < p.$$

Note that the partial sums of such a series are rational numbers. It is as easy to work with such series as with decimal expansions of real numbers — just remember high powers of p are small, and hence the first to be ignored.

We explain this in more detail. The maps

$$\mathbb{Z}/(p^n) \to \mathbb{Z}_{(p)}/(p^n) \to \mathbb{Z}_p/(p^n)$$

are both bijective (see 3.11 for the first map). Let $\alpha \in \mathbb{Z}_p$. Because the map is bijective, for all n, there is an $a_n \in \mathbb{Z}$ such that $\alpha \equiv a_n \mod p^n$. Note that, if $n < m$, $a_n \equiv a_m \mod p^n$, which implies that (a_n) is a Cauchy sequence. Let

$$a_n \equiv c_0 + c_1p + \cdots + c_{n-1}p^{n-1} \mod p^n, \quad 0 \le c_i \le p-1;$$

then

$$\alpha = \sum_{i \ge 0} c_i p^i.$$

Conversely, if $\alpha = \sum c_i p^i$, $0 \le c_i \le p-1$, then $c_0, c_1, \ldots$ is the unique sequence of integers, $0 \le c_i \le p-1$, such that

$$\alpha \equiv \sum_{i=0}^{n-1} c_i p^i \mod p^n.$$

If $\alpha \in \mathbb{Q}_p$ but not $\mathbb{Z}_p$, then $p^m\alpha \in \mathbb{Z}_p$ for a sufficiently large m, and the above arguments can be applied to it.

Example 7.20. To illustrate how to work with p-adic numbers, I prove that -1 is a square in $\mathbb{Q}_5$. We have to find a series

$$a_0 + a_1 5 + a_2 5^2 + \cdots, \quad a_i = 0, 1, 2, 3, \text{ or } 4$$

such that

$$(a_0 + a_1 5 + a_2 5^2 + \ldots)^2 + 1 = 0.$$

We first need that

$$a_0^2 + 1 \equiv 0 \bmod 5.$$

Thus we must take $a_0 = 2$ or 3; we choose 2 (choosing 3 would lead to the other root). Next we need

$$(2 + a_1 5)^2 + 1 \equiv 0 \bmod 5^2,$$

and so we want

$$5 + 20a_1 \equiv 0 \pmod{5^2}.$$

We must take $a_1 = 1$. Suppose we have found

$$c_n = a_0 + a_1 5 + a_2 5^2 + \cdots + a_n 5^n$$

such that

$$c_n^2 + 1 \equiv 0 \pmod{5^{n+1}},$$

and consider $c_n + a_{n+1}5^{n+1}$. We want

$$(c_n + a_{n+1}5^{n+1})^2 + 1 \equiv 0 \pmod{5^{n+2}},$$

for which we need that

$$c_n^2 + 1 + 2c_n a_{n+1}5^{n+1} \equiv 0 \pmod{5^{n+2}},$$

or that

$$2c_n a_{n+1}5^{n+1} \equiv (-1 - c_n^2) \pmod{5^{n+2}},$$

or that

$$2c_n a_{n+1} \equiv (-1 - c_n^2)/5^{n+1} \pmod{5},$$

or that

$$4a_{n+1} = (-1 - c_n^2)/5^{n+1} \pmod{5}.$$

Since 4 is invertible modulo 5, we can always achieve this. Hence we obtain a series converging to -1.

There is a leisurely, and very detailed, discussion of $\mathbb{Q}_p$ in the first chapter of N. Koblitz, p-adic Numbers, p-adic Analysis, and Zeta-Functions, Springer, 1977.

Aside 7.21. Those who have taken a course in commutative algebra will know another method of completing a local ring R, namely

$$R' = \lim_{\leftarrow} R/\mathfrak{m}^n = \{(a_n) \mid a_n \in R/\mathfrak{m}^n, \quad a_{n+1} \equiv a_n \bmod \mathfrak{m}^n\}.$$

In the case that R is a discrete valuation ring, this definition agrees with the above. There is an injective homomorphism

$$R \to R', \quad a \mapsto (a_n), \quad a_n = a \bmod \pi^n.$$

We can define a homomorphism $R' \to \hat{R}$ as follows: let $(a_n) \in R'$, and choose a representative a'_n for a_n in R; then (a'_n) is an Cauchy sequence whose equivalence class is independent of the choices of the a'_n, and we can map (a_n) to (a'_n). It is easy to see that the map $R' \to \hat{R}$ is surjective, and it follows that it is an isomorphism.

Newton's lemma. The argument in the above example works much more generally. Let $f(X) = X^2 + 1$. Then all we in fact used was that $f(X)$ has a simple root modulo 5.

In the rest of this subsection, A is a complete discrete valuation ring and π generates its maximal ideal (unless we say otherwise).

Proposition 7.22. *Let $f(X) \in A[X]$, and let a_0 be a simple root of $f(X)$ mod π. Then there is a unique root a of $f(X)$ with $a \equiv a_0$ mod π.*

Proof. Suppose we have found $a_n \equiv a_0 \mod \pi$ such that

$$f(a_n) \equiv 0 \mod \pi^{n+1}.$$

Let $a_{n+1} = a_n + h\pi^{n+1}$, $h \in A$. We want

$$f(a_n + h\pi^{n+1}) \equiv 0 \bmod \pi^{n+2}.$$

Recall (trivial Taylor's expansion) that, for any polynomial f,

$$f(c + t) = f(c) + t \cdot f'(c) + \cdots$$

where $f'(X)$ is the formal derivative of $f(X)$. Then

$$f(a_n + h\pi^{n+1}) = f(a_n) + h\pi^{n+1} \cdot f'(a_n) + \cdots ,$$

which we want $\equiv 0 \mod \pi^{n+2}$. Hence we must take h so that

$$h = -\frac{f(a_n)}{\pi^{n+1}} \cdot f'(a_n)^{-1} \mod \pi.$$

This is possible because $\pi^n | f(a_n)$ and

$$f'(a_n) \equiv f'(a_0) \mod \pi,$$

which is nonzero, and hence invertible, $\mod \pi$. □

There is a stronger form of the proposition. Recall Newton's approximation method for finding a solution to $f(x) = 0$, where f is a function of a real variable. Starting from an a_0 such that $f(a_0)$ is small, define a sequence $a_1, a_2, ...$ by putting

$$a_{n+1} = a_n - f(a_n)/f'(a_n).$$

Often a_n converges to a root of $f(x)$. In the above proof, this is what we did, but the same argument can be made to work more generally.

THEOREM 7.23 (Newton's lemma). *Let $f(X) \in A[X]$. Let $a_0 \in A$ satisfy*

$$|f(a_0)| < |f'(a_0)|^2.$$

Then there is a unique root a of $f(X)$ such that

$$|a - a_0| \leq \left| \frac{f(a_0)}{f'(a_0)^2} \right| .$$

PROOF. Define a sequence $a_0, a_1, \ldots$ by setting

$$a_{n+1} = a_n - \frac{f(a_n)}{f'(a_n)}$$

and prove that it is a Cauchy sequence converging to a root of $f(X)$. See, for example, EC 2.8. [In fact, it is not necessary to assume that $|\ |$ is discrete — see Lang 1970, p. 42.] □

Proposition 7.22 shows that a simple factor of degree 1 of $f(X) \mod \pi$ lifts to a factor of $f(X)$. This generalizes.

THEOREM 7.24 (Hensel's lemma). *Let k be the residue field of A; for $f(X) \in A[X]$, write $\bar{f}(X)$ for the image of f in $k[X]$. Consider a monic polynomial $f(X) \in A[X]$. If $\bar{f}(X)$ factors as $\bar{f} = g_0 h_0$ with g_0 and h_0 monic and relatively prime (in $k[X]$), then f itself factors as $f = gh$ with g and h monic and such that $\bar{g} = g_0$ and $\bar{h} = h_0$. Moreover, g and h are uniquely determined, and $(g, h) = A[X]$.*

We first prove that $(g, h) = A[X]$ (such a pair is said to be *strictly coprime*; in $k[X]$ strictly coprime just means coprime, i.e., relatively prime).

LEMMA 7.25. *Let A be a local ring with residue field k. If $f, g \in A[X]$ are such that $\bar{f}$ and $\bar{g}$ are relatively prime and f is monic, then $(f, g) = A[X]$.*

PROOF. Let $M = A[X]/(f, g)$. As f is monic, this is a finitely generated A-module. As $(\bar{f}, \bar{g}) = k[X]$, we have that $(f, g) + \mathfrak{m}A[X] = A[X]$ and so $\mathfrak{m}M = M$. Now Nakayama's Lemma (1.3) implies that $M = 0$. □

We next prove uniqueness of g and h.

LEMMA 7.26. *Let A be a local ring with residue field k. Suppose $f = gh = g'h'$ with g, h, g', h' all monic, and $\bar{g} = \bar{g}'$, $\bar{h} = \bar{h}'$ with $\bar{g}$ and $\bar{h}$ relatively prime. Then $g = g'$ and $h = h'$.*

PROOF. From the preceding lemma we know that $(g, h') = A[X]$, and so there exist $r, s \in A[X]$ such that $gr + h's = 1$. Now

$$g' = g'gr + g'h's = g'gr + ghs,$$

and so g divides g'. As both are monic and have the same degree, they must be equal. □

Finally, we prove the existence of g and h. We are given that there exist monic polynomials g_0, $h_0 \in A[X]$ such that

$$f - g_0h_0 \in \pi \cdot A[X].$$

Suppose we have constructed monic polynomials g_n, h_n such that

$$f - g_nh_n \equiv 0 \bmod \pi^{n+1}A[X]$$

and $g_n \equiv g_0$, $h_n \equiv h_0 \bmod \pi A[X]$. We want to find u, $v \in A[X]$ such that

$$f - (g_n + \pi^{n+1}u)(h_n + \pi^{n+1}v) \equiv 0 \bmod \pi^{n+2}A[X],$$

i.e., we want

$$(f - g_nh_n) - \pi^{n+1}(uh_n + g_nv) \equiv 0 \bmod \pi^{n+2}A[X].$$

Thus we are looking for polynomials u, v in $A[X]$ such that

$$uh_n + g_nv \equiv (f - g_nh_n)/\pi^{n+1} \bmod \pi A[X].$$

From (7.25), we know that h_n and g_n are strictly coprime, and so we can always find such polynomials u, v.

REMARK 7.27. By induction, the theorem shows that a factorization of f into a product of relatively prime polynomials in $k[X]$ lifts to a factorization in $A[X]$. For example, in $\mathbb{F}_p[X]$, $X^p - X$ splits into p distinct factors, and so it also splits in $\mathbb{Z}_p[X]$. Hence $\mathbb{Z}_p$ contains the $p - 1^{st}$ roots of 1. More generally, if K has a residue field k with q elements, then K contains q roots of the polynomial $X^q - X$. Let S be the set of these roots. Then

$$a \mapsto \bar{a} \colon S \to k,$$

is a bijection preserving multiplication (but not, of course, addition) – the elements of S are called the *Teichmüller representatives* for the elements of the residue field.

REMARK 7.28. Theorems 7.23 and 7.24 are both stronger versions of 7.22. There is in fact a stronger version of 7.23. For a polynomial $h = \sum c_iX^i$, define

$$\|h\| = \max |c_i|.$$

Let

$$f(X) = a_nX^n + a_{n-1}X^{n-1} + \cdots + a_0 \in A[X]$$

have $|a_n| = 1$ (i.e., a_n is a unit). Let $g_0(X)$ and $h_0(X)$ be polynomials in $A[X]$ with degrees r and s respectively, and suppose that

$$\|f(X) - g_0(X)h_0(X)\| < |\operatorname{Res}(g_0(X), h_0(X))|^2$$

where Res denotes the resultant. Then $f(X)$ factors in $A[X]$ as the product of a polynomial of degree r and a polynomial of degree s. The proof follows the same general lines as the above proofs. In fact, the hypothesis can be replaced by

$$\|f(X) - g_0(X)h_0(X)\| < |\operatorname{disc}(f)|.$$

(For this, see Cassels 1986, p107.)

Note that, this gives an algorithm for factoring polynomials in $\mathbb{Q}_p[X]$ (for example). Given $f(X)$, compute $\operatorname{disc}(f)$. If this is zero, then f and f' have a common factor (which we can find by the Euclidean algorithm). Otherwise $\operatorname{ord}(\operatorname{disc}(f)) = m$ for some m, and it is enough to consider factorizations of f into polynomials with coefficients in the finite ring $\mathbb{Z}/p^m\mathbb{Z}$. Apparently the fastest algorithms for factoring polynomials in $\mathbb{Z}[X]$ begin by factoring in $\mathbb{Z}_p[X]$ for an appropriate prime p — computers seem to have no problem handling polynomials of degree 200. (But Problems 10, no. 3, shows that there are irreducible polynomials in $\mathbb{Z}[X]$ of arbitrarily large degree that factor in all the rings $\mathbb{Z}_p[X]$ into polynomials of low degree.)

Extensions of nonarchimedean valuations. We explain how to extend a valuation to a larger field.

THEOREM 7.29. *Let K be complete with respect to a discrete valuation $|\ |_K$, and let L be a finite separable extension of K of degree n. Then $|\ |$ extends uniquely to a discrete valuation $|\ |_L$ on L, and L is complete for the extended valuation. For all $\beta \in L$,*

$$|\beta|_L = |\operatorname{Nm}_{L/K} \beta|_K^{1/n}.$$

PROOF. Let A be the discrete valuation ring in K, and let B be its integral closure in L. Let $\mathfrak{p}$ be the maximal ideal of A. We know from (3.30) that B is a Dedekind domain, and the valuations of L extending $|\ |_\mathfrak{p}$ correspond to the ideals of B lying over $\mathfrak{p}$.

Suppose that there are distinct prime ideals $\mathfrak{P}_1$ and $\mathfrak{P}_2$ in B dividing $\mathfrak{p}$. There will be a $\beta \in B$ such that $\mathfrak{P}_1 \cap A[\beta] \neq \mathfrak{P}_2 \cap A[\beta]$; for example, choose $\beta \in B$ such that $\beta \in \mathfrak{P}_1$, $\beta \notin \mathfrak{P}_2$. Let $f(X)$ be the minimum polynomial of β over K, so that $A[\beta] \cong A[X]/(f(X))$. Because $f(X)$ is irreducible in $A[X]$ and A is complete, Hensel's lemma shows that $\bar{f}(X)$ (image of $f(X)$ in $k[X]$, $k = A/\mathfrak{p}$) must be a power of an irreducible polynomial. Then

$$A[\beta]/\mathfrak{p}A[\beta] \approx k[X]/(\bar{f}(X))$$

is a local ring, which contradicts the fact that $A[\beta]$ has two prime ideals containing $\mathfrak{p}$.

Hence $|\ |_\mathfrak{p}$ extends uniquely to a valuation $|\ |$ on L.

Clearly, $|\ |_\mathfrak{p}$ also extends uniquely to the Galois closure L' of L. For each $\sigma \in \operatorname{Gal}(L/K)$, consider the map $L \hookrightarrow \mathbb{C}$, $\beta \mapsto |\sigma\beta|$. This is again a valuation of L, and

so the uniqueness implies that $|\beta| = |\sigma\beta|$. Now

$$|\operatorname{Nm}(\beta)| = |\prod \sigma\beta| = |\beta|^n$$

which implies the formula.

Finally, we have to show that L is complete. Let $e_1, \dots, e_n$ be a basis for B as an A-module, and let $(\alpha(m))$ be a Cauchy sequence in L. Write $\alpha(m) = a_1(m)e_1 + \cdots + a_n(m)e_n$, with $a_i(m) \in K$. For each i, $a_i(m)$ is a Cauchy sequence, and if a_i denotes its limit, then $\alpha \overset{\text{df}}{=} a_1e_1 + \cdots + a_ne_n$ is the limit of the sequence $\alpha(m)$. □

Remark 7.30. It is obvious from the criterion (7.2) that a nonarchimedean valuation can only extend to a nonarchimedean valuation. It is possible to prove (7.29) without assuming that the valuation $|\ |$ on K is discrete or even nonarchimedean, but the proof is then completely different, and much longer — we shall in fact need this in the §8, and so I should have included it. The formula $|\beta|_L = |\operatorname{Nm}_{L/K}\beta|_K^{1/n}$ shows that $|\ |_L$ is discrete if and only if $|\ |_K$ is discrete.

Corollary 7.31. *Let K be as in the theorem, and let Ω be a (possibly infinite) algebraic extension of $\mathbb{Q}$. Then $|\ |$ extends in a unique way to a valuation $|\ |$ on Ω.*

Proof. The theorem shows that $|\ |$ extends in a unique way to any finite subextension of Ω, and hence it extends uniquely to Ω. □

Remark 7.32. In the last corollary, the extended valuation is still nonarchimedean, but it need not be discrete, and Ω need not be complete. However, the completion of Ω is again algebraically closed.

For example as we noted in (7.7), the valuation on the algebraic closure $\mathbb{Q}_p^{\text{al}}$ of $\mathbb{Q}_p$ is not discrete, and Problems 10, no. 4, shows that $\mathbb{Q}_p^{\text{al}}$ is not complete. The completion of $\mathbb{Q}_p^{\text{al}}$ is often denoted $\mathbb{C}_p$ because it plays the same role for the p-adic valuation on $\mathbb{Q}$ that $\mathbb{C}$ plays for the real valuation. (In fact $\mathbb{C}_p \approx \mathbb{C}$ as abstract fields because they are both algebraically closed, and they both have a transcendence basis with cardinality equal to that of $\mathbb{R}$. The isomorphism is as far from being canonical as it is possible to get — its construction requires the axiom of choice.)

Corollary 7.33. *Let K and L be as in the theorem; then $n = ef$ where $n = [L : K]$, e is the ramification index, and f is the degree of the residue field extension.*

Proof. We know from (3.36) that $n = \sum e_if_i$. In this case, there is only one prime dividing $\mathfrak{p}$ and so the formula becomes $n = ef$. □

When $e = n$, so that $\mathfrak{p}B = \mathfrak{P}^n$, we say that L is *totally ramified* over K; when $f = n$, we say that L is *unramified* over K.

Note that the valuation ring B of L is the integral closure of the valuation ring A of K.

Many of the results proved above for complete discrete valuation rings hold also for Henselian local rings (see §4 of my notes on Etale Cohomology).

Remark 7.34. Let K be complete with respect to a discrete valuation, and let L be a finite extension of K. Let $\mathfrak{P}$ and $\mathfrak{p}$ be the maximal ideals in the rings of integers A and B of K and L. Then $\mathfrak{p}B = \mathfrak{P}^e$ where e is the ramification index. Let π and

Π be generators of $\mathfrak{p}$ and $\mathfrak{P}$. The normalized valuations ord_K and ord_L on K and L are characterized by equations:

$$\mathrm{ord}_K(\pi) = 1, \qquad \mathrm{ord}_L(\Pi) = 1.$$

Note that $\pi = \Pi^e \times \text{unit}$, and so

$$\mathrm{ord}_K = e^{-1}\mathrm{ord}_L.$$

If we denote the extension of ord_K to L by ord, then

$$\mathrm{ord}(L^\times) = e^{-1}\mathbb{Z}.$$

This characterizes the ramification index.

Newton's polygon. Let K be complete with respect to a discrete valuation. Let ord be the corresponding additive valuation $\mathrm{ord}\colon K^\times \twoheadrightarrow \mathbb{Z}$, and extend ord to a valuation $\mathrm{ord} : K^{\mathrm{al}\times} \to \mathbb{Q}$. For a polynomial

$$f(X) = X^n + a_1X^{n-1} + \cdots + a_n, \quad a_i \in K,$$

define the *Newton polygon*[14] of $f(X)$ to be the lower convex hull of the set of points

$$P_i \stackrel{\mathrm{df}}{=} (i, \mathrm{ord}(a_i)),\ i = 0, ..., n.$$

In more detail, rotate the negative y-axis counter-clockwise about $P_0 = (0, 0)$ until it hits a P_i — the first segment of the Newton polygon is the line $P_0P_{i_1}$ where P_{i_1} is the point furthest from P_0 on the rotated y-axis. Repeat the process rotating about P_{i_1}, etc.. The resulting polygon starts at P_0 and ends at P_n; each of its segments begins and ends at a P_i; each P_i either lies on the polygon or is above it; any line joining two points of the polygon has no point that is below the polygon (this is what we mean by the Newton polygon being lower convex).

PROPOSITION 7.35. *Suppose that the Newton polygon of $f(X) \in K[X]$ has segments of x-length n_i and slope s_i. Then $f(X)$ has exactly n_i roots α (in K^{al}) with*

$$ord(\alpha) = s_i.$$

Moreover, the polynomial $f_i(X) \stackrel{df}{=} \Pi_{ord(\alpha_i)=s_i}(X - \alpha_i)$ has coefficients in K.

PROOF. In proving the first part, we don't have to assume that $f(X)$ has coefficients in K—any finite extension of K will do. Thus it suffices to prove the following statement: let $f(X) = \prod(X - \alpha_j)$; if exactly n_i of the α_j's have $\mathrm{ord} s_i$, then the Newton polygon of $f(X)$ has a segment of slope s_i and x-length n_i.

We prove this by induction on $n = \deg(f)$. If $n = 1$, then it is obvious. Assume it for n, and put

$$g(X) = (X - \alpha)f(X) = X^{n+1} + b_1X^n + b_2X^{n-1} + \cdots + b_{n+1}.$$

Note that $b_i = a_i - \alpha a_{i-1}$.

Case (i). $\mathrm{ord}(\alpha) < s_1$. Recall $\mathrm{ord}(a + b) \geq \min\{\mathrm{ord}(a), \mathrm{ord}(b)\}$, with equality if $\mathrm{ord}(a) \neq \mathrm{ord}(b)$. Using this, one finds that

[14]Most people write the polynomial $a_0 + a_1X + \cdots + X^n$ when they define Newton polygons. This is slightly less convenient than the way I do it, but allows you to define the Newton polygon of a power series.

the Newton polygon of g is obtained from that of f by adding a segment of slope $\mathrm{ord}(\alpha)$ and x-length 1, and moving the Newton polygon of f to start at $(1, \mathrm{ord}(\alpha))$. This is what the proposition predicts.

Case (ii). $\mathrm{ord}(\alpha) = s_1$. In this case, the initial segment of slope s_1 is lengthened by 1, and the rest of the polygon is as before. This is what the proposition predicts.

The remaining cases are similar.

We now prove the second statement. Let α be a root of $f(X)$, and let $m_\alpha(X)$ be the minimum polynomial of α. As we saw in the proof of (7.29), $\mathrm{ord}(\alpha') = \mathrm{ord}(\alpha)$ for all conjugates α' of α, i.e., for all roots of $m_\alpha(X)$. Because $f(\alpha) = 0$, $m_\alpha(X)|f(X)$, and the remark just made implies that in fact $m_\alpha(X)|f_i(X)$ where $s_i = \mathrm{ord}(\alpha)$. If β is a root of $f_i(X)/m_\alpha(X)$, then a similar argument shows that $m_\beta(X)|(f_i/m_\alpha)$. Continuing in this way, we find that $f_i(X)$ is a product of polynomials with coefficients in K. □

Example 7.36. Consider the polynomial

$$f(X) \stackrel{\mathrm{df}}{=} X^3 + X^2 + 2X - 8.$$

By testing ±1, ±2, ±4, ±8 (actually, by asking Maple) one sees that this polynomial is irreducible over $\mathbb{Q}$. The Newton polygon of f relative to ord_2 has slopes $0, 1, 2$, each with x-length 1. Therefore f splits in $\mathbb{Q}_2[X]$, and it has roots α_1, α_2, α_3 with ords 0, 1, 2.

Locally compact fields. We now look at the compactness properties of our fields.

Proposition 7.37. *Let K be complete with respect to a nonarchimedean discrete valuation. Let A be the ring of integers in K and let $\mathfrak{m}$ be the maximal ideal in A. Then A is compact if and only if $A/\mathfrak{m}$ is finite.*

Proof. Let S be a set of representatives for $A/\mathfrak{m}$. We have to show that A is compact if and only if S is finite.

$\Rightarrow$: Clearly $\mathfrak{m} = \{x \in K \mid |x| < 1\}$ is open in K. As A is the disjoint union of the open sets $s + \mathfrak{m}$, $s \in S$, S must be finite if A is compact.

$\Leftarrow$: Recall that a metric space X is compact if and only if it is complete and totally bounded (this means that for any $r > 0$, there is a finite covering of X by open balls of radius r). But every element of A can be written

$$s_0 + s_1\pi + s_2\pi^2 + \cdots + s_n\pi^n + \cdots, \quad s_i \in S.$$

For a fixed n, there are only finitely many sums

$$s_0 + s_1\pi + s_2\pi^2 + \cdots + s_n\pi^n, \quad s_i \in S,$$

and every element of A is within $|\pi^{n+1}|$ of such an element. □

Corollary 7.38. *Assume that the residue field is finite. Then $\mathfrak{p}^n$, $1 + \mathfrak{p}^n$, and $A^\times$ are all compact.*

Proof. They are all closed subsets of A. □

Definition 7.39. A *local field* is a field K with a nontrivial valuation $|\ |$ (as defined at the start of this section) such that K is locally compact (and hence complete).

REMARK 7.40. It is possible to give a complete classification of local fields.

(a) Let K be a field that is complete with respect to an archimedean valuation $|\ |$; then K is isomorphic to $\mathbb{R}$ of $\mathbb{C}$, and the valuation is equivalent to the usual absolute value (Theorem of Ostrowski, see Janusz 1996, II.4). Thus for archimedean valuations, completeness implies local compactness.

(b) A nonarchimedean local field K of characteristic zero is isomorphic to a finite extension of $\mathbb{Q}_p$, and the valuation is equivalent to the (unique) extension of the p-adic valuation. (To prove this, note that, by assumption, K contains $\mathbb{Q}$. The restriction of $|\ |$ to $\mathbb{Q}$ can't be the trivial valuation, because otherwise $A^\times$ wouldn't be compact. Therefore (see 7.10) $|\ |$ induces a valuation on $\mathbb{Q}$ equivalent to the p-adic valuation for some prime number p. The closure of $\mathbb{Q}$ in K is therefore $\mathbb{Q}_p$. If K has infinite degree over $\mathbb{Q}_p$, it will not be locally compact.)

(c) A nonarchimedean local field K of characteristic $p \neq 0$ is isomorphic to the field of formal Laurent series $k((T))$ over a finite field k. The field $k((T))$ is the completion of $k(T)$ for the valuation defined by the ideal $(T) \subset k[T]$; it consists of finite-tailed formal power series:

$$\sum_{i \geq -n}^{\infty} a_i T^i.$$

Unramified extensions of a local field. Again K is a field complete with respect to a discrete valuation $|\ |$. To avoid problems with separability, we assume that K and the residue field k are both perfect[15]—of course in the case we are particularly interested in, K has characteristic zero and k is finite. Let A be the discrete valuation ring in K corresponding to $|\ |$.

If L is an algebraic (possibly infinite) extension of K, we can still define

$$B = \{\alpha \in L \mid |\alpha| \leq 1\}$$

$$\mathfrak{p} = \{\alpha \in B \mid |\alpha| < 1\}$$

and call $B/\mathfrak{p}$ the residue field of L.

PROPOSITION 7.41. *Let L be an algebraic extension of K, and let l be the residue field of L. The map $K' \mapsto k'$ sending an unramified extension K' of K contained in L to its residue field k' is a one-to-one correspondence between the sets*

$$\{K' \subset L, \textit{ finite and unramified over } K\} \leftrightarrow \{k' \subset l, \textit{ finite over } k\}.$$

Moreover:

(a) *if $K' \leftrightarrow k'$ and $K'' \leftrightarrow k''$, then $K' \subset K'' \iff k' \subset k''$;*
(b) *if $K' \leftrightarrow k'$, then K' is Galois over K if and only if k' is Galois over k, in which case there is a canonical isomorphism*

$$\mathrm{Gal}(K'/K) \to \mathrm{Gal}(k'/k).$$

[15]When k is not perfect, we should define L/K to be unramified if (a) the ramification index is 1, and (b) the residue field extension is separable. These conditions imply that L/K is separable. With this definition, (7.41) continues to hold without K and k being assumed to be perfect

PROOF. Let k' be a finite extension of k. We can write it $k' = k[a]$. Let $f_0(X)$ be the minimum polynomial of a over k, and let $f(X)$ be any lifting of $f_0(X)$ to $A[X]$. As a is a simple root of $f_0(X)$, Newton's lemma (7.22) shows that there is a (unique) $\alpha \in L$ such that $f(\alpha) = 0$ and $\alpha \equiv a \mod \mathfrak{p}$. Now $K' =_{df} K[\alpha]$ has residue field k'. Thus $K' \mapsto k'$ is surjective. Suppose that K' and K'' are unramified extensions of K in L with the same residue field k'. Then $K' \cdot K''$ is an unramified extension of K (see 6.5 and 6.6b) with residue field k'. Hence

$$[K' \cdot K'' \colon K] = [k' \colon k] = [K' \colon K],$$

and so $K'' = K'$.

Statement (a) is obvious.

Assume K' is Galois over K; then $\mathrm{Gal}(K'/K)$ preserves A' (the valuation ring in K') and its maximal ideal, and so we get a map $\mathrm{Gal}(K'/K) \to \mathrm{Aut}(k'/k)$. Write $k' = k[a]$, and let $g(X) \in A[X]$ be such that $\bar{g}(X) \in k[X]$ is the minimum polynomial of a. Let $\alpha \in A'$ be the unique root of $g(X)$ such that $\bar{\alpha} = a$. Because K' is Galois over K, $g(X)$ splits in $A'[X]$, and this implies that $\bar{g}(X)$ splits in $k'[X]$, and so k' is Galois over k. Let $f = [k' \colon k] = [K' \colon K]$, and let $\alpha_1, \dots, \alpha_f$ be the roots of $g(X)$. Then

$$\{\alpha_1, ..., \alpha_f\} = \{\sigma\alpha \mid \sigma \in \mathrm{Gal}(L/K)\}.$$

Because $\bar{g}(X)$ is separable, the α_i are distinct modulo $\mathfrak{p}$, and this shows that the image of the map $\mathrm{Gal}(K'/K) \to \mathrm{Gal}(k'/k)$ has order f, and hence is an isomorphism. Conversely, suppose k'/k is Galois. Again write $k' = k[a]$, and $\alpha \in A'$ lift a. It follows from Hensel's lemma that A' contains the conjugates of α, and hence that K' is Galois over K. □

COROLLARY 7.42. *There is a field $K_0 \subset L$ containing all unramified extensions of K in L (called the* largest unramified extension *of K in L). In fact, it is obtained from K by adjoining all roots of 1 of order prime to the characteristic of k.*

PROOF. This is an obvious consequence of the theorem. □

COROLLARY 7.43. *The residue field of K^{al} is k^{al}; there is a subfield K^{un} of K^{al} such that a subfield L of K^{al}, finite over K, is unramified if and only if $L \subset K^{un}$. (Recall that we are assuming k and K to be perfect.)*

PROOF. Let $f_0(X)$ be any polynomial in $k[X]$, and let $f(X)$ be any lift of $f_0(X)$ to $A[X]$. Then K^{al} contains all the roots of $f(X)$, and so the residue field k' of K^{al} contains all the roots of $f_0(X)$. Hence k' is algebraic over k, and every polynomial in $k[X]$ splits in k', and so it must be the algebraic closure of k. □

REMARK 7.44. For those familiar with the language of category theory, we can be a little more precise: there is an equivalence between the category of finite unramified extensions of K and the category of finite (separable) extensions of k.

EXAMPLE 7.45. Let K be a local field of characteristic zero (hence a finite extension of $\mathbb{Q}_p$ for some p), and let q be the order of the residue field k of K.

Recall from (FT §4.6) that, for each n, there is an extension k_n of k of degree n, and that k_n is unique up to k-isomorphism; it is the splitting field of $X^{q^n} - X$. The

Galois group $\mathrm{Gal}(k_n/k)$ is a cyclic group of order n, having as canonical generator the *Frobenius element* $x \mapsto x^q$.

Therefore, for each n, there is an unramified extension K_n of K of degree n, and it is unique up to K-isomorphism; it is the splitting field of $X^{q^n} - X$; the Galois group $\mathrm{Gal}(K_n/K)$ is a cyclic group of order n, having as canonical generator the *Frobenius element* σ which is determined by the property

$$\sigma\beta \equiv \beta^q \pmod{\mathfrak{p}},$$

all $\beta \in B$. (Here B is the discrete valuation ring in K_n, and $\mathfrak{p}$ is the nonzero prime ideal in B.)

Totally ramified extensions of K. Let K be a complete discretely-valued nonarchimedean field, and let π be a local uniformizing parameter for K. A polynomial $f(X) \in K[X]$ is said to be *Eisenstein* if it is Eisenstein for the maximal ideal of the ring of integers in K, i.e., if

$$f(X) = a_0X^n + a_1X^{n-1} + \cdots + a_n, \text{ with } |a_0| = 1, \quad |a_i| < 1, \quad |a_n| = |\pi|.$$

Equivalently,

$$\mathrm{ord}(a_0) = 0, \quad \mathrm{ord}(a_i) > 0, \quad \mathrm{ord}(a_n) = 1,$$

for the normalized additive valuation. Equivalently, the Newton polygon of $f(X)$ has only one segment, which has slope $\frac{1}{n}$, $n = \deg f$. Eisenstein polynomials allow us to give an explicit description of all totally ramified extensions of K.

Proposition 7.46. *Let L be a finite extension of K. Then L/K is totally ramified if and only if $L = K[\alpha]$ with α a root of an Eisenstein polynomial.*

Proof. $\Leftarrow$: Suppose $L = K[\alpha]$ with α a root of an Eisenstein polynomial $f(X)$ of degree n. If ord is the extension of the normalized discrete (additive) valuation on K to L, then $\mathrm{ord}(\alpha) = 1/n$. This implies that the ramification index of L/K is $\geq n$. But it can't be greater than n, and so it is exactly n — L is totally ramified over K. (Compare the proof of 6.2.)

$\Rightarrow$: Suppose L is a totally ramified extension of K of degree n. Let α be a generator of the maximal ideal in the ring of integers in L; thus $\mathrm{ord}(\alpha) = 1/n$ if ord extends the normalized discrete valuation on K. The elements $1, \alpha, ..., \alpha^{n-1}$ represent different cosets of $\mathrm{ord}(K^\times)$ in $\mathrm{ord}(L^\times)$, and so it is impossible to have a nontrivial relation

$$a_0 + a_1\alpha + \cdots + a_{n-1}\alpha^{n-1} = 0, \quad a_i \in K$$

(because of 7.9.3). Hence $L = K[\alpha]$. The elements $1, \alpha, \ldots, \alpha^{n-1}, \alpha^n$ are linearly dependent over K, and so we have a relation:

$$\alpha^n + a_1\alpha^{n-1} + \cdots + a_n = 0, \quad a_i \in K.$$

Applying (7.9.3) again, we see that the minimum ord of a summand must be attained for two terms. The only way this can happen is if $\mathrm{ord}(a_i) > 0$ for all i and $\mathrm{ord}(a_n) = \mathrm{ord}(\alpha^n) = 1$, i.e., if $\sum a_iX^i$ is an Eisenstein polynomial. □

Remark 7.47. Let L be a finite totally ramified extension of K. Let A and B be the discrete valuation rings in K and L, and let π and Π be a prime elements in A

and B. I claim that $B = A[\Pi]$. The argument is the same as in the proof of 6.2 (see also Problems 8, no. 1). Because B and A have the same residue field,

$$A[\Pi] + \Pi B = B.$$

The discriminant of $1, \Pi, \Pi^2, \dots$ is a unit$\times\pi^m$ for some m, and so

$$\mathfrak{p}^c B \subset A[\Pi] \subset B$$

for some c. As before, these two conditions suffice to imply that $B = A[\Pi]$.

Ramification groups. Let L be a finite Galois extension of K, and assume that the residue field k of K is perfect. As we have noted, $G \stackrel{\text{df}}{=} \text{Gal}(L/K)$ preserves the valuation on L. In particular, it preserves

$$B = \{\alpha \in L \mid |\alpha| \leq 1\}, \quad \mathfrak{p} = \{\alpha \in L \mid |\alpha| < 1\}.$$

Let Π be a prime element of L (so that $\mathfrak{p} = (\Pi)$). We define a sequence of subgroups $G \supset G_0 \supset G_1 \supset \cdots$ by the condition:

$$\sigma \in G_i \iff |\sigma\alpha - \alpha| < |\Pi|^i, \text{ all } \alpha \in B.$$

The group G_0 is called the *inertia group*, the group G_1 is called the *ramification group,* and the groups G_i, $i > 1$, are called the *higher ramification groups* of L over K.

LEMMA 7.48. *The G_i are normal subgroups of G, and $G_i = \{1\}$ for i large enough.*

PROOF. (a) For $\sigma, \tau \in G$, we have

$$|\tau^{-1}\sigma\tau\alpha - \alpha| = |\sigma(\tau\alpha) - (\tau\alpha)|$$

(because $|x| = |\tau x|$). As α runs through B, so also does $\tau\alpha$, and so $\tau^{-1}\sigma\tau \in G_i$ exactly when σ does.

(b) If $\sigma \neq 1$, then $\sigma\alpha \neq \alpha$ for some $\alpha \in B$. Hence $\sigma \notin G_i$ as soon as $|\sigma\alpha - \alpha| \geq |\Pi|^i$. □

THEOREM 7.49. *Let L/K be a Galois extension, and assume that the residue field extension l/k is separable.*

(a) *The fixed field of G_0 is the largest unramified extension K_0 of K in L, and*

$$G/G_0 = \text{Gal}(K_0/K) = \text{Gal}(l/k).$$

(b) *For $i \geq 1$, the group*

$$G_i = \{\sigma \in G_0 \mid |\sigma\Pi - \Pi| < |\Pi|^i\}.$$

PROOF. (a) Let K_0 be the largest unramified extension in L (see 7.42). Then σK_0 is also unramified, and so it is contained in K_0. Thus K_0 is Galois over K, and the canonical map $\text{Gal}(K_0/K) \to \text{Gal}(l/k)$ is an isomorphism (see 7.41). By definition G_0 is the kernel of $G \to \text{Gal}(l/k)$, and so K_0 is its fixed field.

(b) Let A_0 be the discrete valuation ring in K_0. Then $B = A_0[\Pi]$ (by 7.45). Since G_0 leaves A_0 fixed, in order to check that $\sigma \in G_i$ it suffices to check that $|\sigma\alpha - \alpha| < |\Pi|^i$ for the element $\alpha = \Pi$. □

COROLLARY 7.50. *We have an exhaustive filtration $G \supset G_0 \supset \cdots$ such that*
$G/G_0 = \mathrm{Gal}(l/k)$;
$G_0/G_1 \hookrightarrow l^\times$;
$G_i/G_{i+1} \hookrightarrow l$.
Therefore, if k is finite, then $\mathrm{Gal}(L/K)$ is solvable.

PROOF. Let $\sigma \in G_0$; then $\sigma\Pi$ is also a prime element and so $\sigma\Pi = u\Pi$ with u a unit in B. The map $\sigma \mapsto u \bmod \mathfrak{p}$ is a homomorphism $G_0 \to l^\times$ with kernel G_1.

Let $\sigma \in G_i$. Then $|\sigma\Pi - \Pi| \leq |\Pi|^{i+1}$, and so $\sigma\Pi = \Pi + a\Pi^{i+1}$ some $a \in B$. The map $\sigma \mapsto a \pmod{\mathfrak{p}}$ is a homomorphism $G_i \to l$ with kernel G_{i+1}. □

An extension L/K is said to be *wildly ramified* if $p|e$ where $p = \mathrm{char}(k)$. Otherwise it is said to be *tamely ramified*. Hence for a Galois extension

$$L/K \text{ is unramified} \iff G_0 = \{1\},$$

and

$$L/K \text{ is tamely ramified} \iff G_1 = \{1\}.$$

Krasner's lemma and applications. Again let K be complete with respect to a discrete nonarchimedean valuation $|\ |$, and extend the valuation (uniquely) to a valuation on K^{al}. It is clear from our discussion of unramified extensions of K that roots of distinct polynomials $f(X)$ and $g(X)$ will often generate the same extension of K; in fact, this will be true if $\bar{f} = \bar{g}$ and both are irreducible in $k[X]$. Krasner's lemma and its consequences show that the roots of two polynomials will generate the same extension if they are sufficiently close.

PROPOSITION 7.51 (Krasner's lemma). . *Let $\alpha, \beta \in K^{al}$, and assume that α is separable over $K[\beta]$. If α is closer to β than to any conjugate of α (over K), then $K[\alpha] \subset K[\beta]$.*

PROOF. Let σ be an embedding of $K[\alpha, \beta]$ into K^{al} fixing $K[\beta]$. By Galois theory, it suffices to show that $\sigma\alpha = \alpha$. But

$$|\sigma\alpha - \beta| = |\sigma\alpha - \sigma\beta| = |\alpha - \beta|$$

because $\sigma\beta = \beta$ and $|\sigma * | = | * |$. Hence

$$|\sigma\alpha - \alpha| = |\sigma\alpha - \beta + \beta - \alpha| \leq |\alpha - \beta|.$$

Since $\sigma\alpha$ is a conjugate of α over K, the hypothesis now implies that $\sigma\alpha = \alpha$. □

Now assume K has characteristic zero (to avoid complications). As before, for $h(X) = \sum c_i X^i$, we define $\|h\| = \max\{|c_i|\}$. Note that if $h(X)$ varies in a family of monic polynomials for which $\|h\|$ remains bounded, then the maximum value of a root of h is bounded; in fact, if

$$\sum c_i \beta^i = 0,$$

we must have $|\beta^n| \leq |c_j \beta^j|$ for some $j < n$, and so $|\beta|^{n-j} \leq |c_j|$.

Fix a monic irreducible polynomial $f(X)$ in $K[X]$, and let

$$f(X) = \prod (X - \alpha_i), \quad \alpha_i \in K^{\mathrm{al}}.$$

The α_i must be distinct. Let $g(X)$ be a second monic polynomial in $K[X]$, and suppose that $\|f-g\|$ is small. For any root β of $g(X)$, $|f(\beta)| = |(f-g)(\beta)|$ is small (because $\|f-g\|$ small implies that $\|g\|$ is bounded, and hence $|\beta|$ is bounded). But

$$|f(\beta)| = \prod |\beta - \alpha_i|.$$

In order for this to be small, at least one term $|\beta - \alpha_i|$ must be small. By taking $\|f-g\|$ small enough, we can force β to be closer to one root α_i than α_i is to any other α_j. That is, we can achieve:

$$|\beta - \alpha_i| < |\alpha_i - \alpha_j|, \text{ all } j \neq i.$$

In this case, we say that β *belongs to* α_i. Krasner's lemma then says that $K[\alpha_i] \subset K[\beta]$, and because f and g have the same degree, they must be equal. We have proved:

PROPOSITION 7.52. *Let $f(X)$ be a monic irreducible polynomial of $K[X]$, and let α be a root of f. Then any monic polynomial $g(X) \in K[X]$ sufficiently close to $f(X)$ is also irreducible, and it has a root β that belongs to α. For such a root $K[\alpha] = K[\beta]$.*

COROLLARY 7.53. *Let K be a finite extension of $\mathbb{Q}_p$. Then there is a finite extension L of $\mathbb{Q}$ contained in K such that $[L\colon \mathbb{Q}] = [K\colon \mathbb{Q}_p]$ and $L \cdot \mathbb{Q}_p = K$.*

PROOF. Write $K = \mathbb{Q}_p[\alpha]$, and let $f(X)$ be the minimum polynomial of α over $\mathbb{Q}_p$. Choose $g(X) \in \mathbb{Q}[X]$ sufficiently close to $f(X)$, and let $L = \mathbb{Q}[\beta]$ for β a root of $g(X)$ belonging to α. □

Now consider two monic polynomials f and g, and write α_i for the roots of f and β_i for the roots of g. For $\|f-g\|$ sufficiently small, every root of g will belong to a root of f, and I claim that they will belong to distinct roots, i.e., the roots can be numbered so that β_i belongs to α_i. For each choice s of n elements from $\{\alpha_1, ..., \alpha_n\}$ (possibly with repetitions), we form the polynomial $f_s(X) = \prod_{\alpha_i \in s}(X - \alpha_i)$. If two roots of $g(X)$ belong to the same root of $f(X)$, then $g(X)$ will be close to $f_s(X)$ for some $f_s \neq f$. But if we choose g to be closer to f than f is to any f_s, this will be impossible. We have proved:

PROPOSITION 7.54. *Assume K is of characteristic zero. If two monic irreducible polynomials f and g are sufficiently close, then each root of g will belong to exactly one root of f, and so*

$$\{K[\alpha] \mid \alpha \text{ a root of } f\} = \{K[\beta] \mid \beta \text{ a root of } g\}.$$

PROPOSITION 7.55. *Assume K has characteristic zero and has finite residue field. Then, up to isomorphism, there are only finitely many totally ramified extensions of $\mathbb{Q}_p$ of a given degree.*

PROOF. We fix an n and show that there are only finite many extensions of degree $\leq n$. Each point of

$$(a_1, ..., a_n) \in \mathfrak{p} \times \mathfrak{p} \times \mathfrak{p} \times \cdots \times A^\times \pi$$

defines an Eisenstein polynomial of degree n, namely,

$$f(X) = X^n + a_1 X^{n-1} + \cdots + a_n,$$

and hence a finite set of totally ramified extensions of degree n, namely, those generated by the roots of $f(X)$. According to the last proposition, each point of $\mathfrak{p} \times \mathfrak{p} \times \mathfrak{p} \times \cdots \times A^\times\pi$ has a neighbourhood such that the points in the neighbourhood all give the same extensions of K. In (7.38) we showed that the factors of $\mathfrak{p} \times \mathfrak{p} \times \mathfrak{p} \times \cdots \times A^\times\pi$ are compact, hence the product is compact, and so a finite number of these neighbourhoods will cover it. □

REMARK 7.56. We proved above that

(i) every finite extension L of K contains a largest unramified extension of K;

(ii) for each $m \geq 1$, there is an unramified extension of degree m of K, and any two such extensions are K-isomorphic.

Fix an n; then each extension L of K of degree n can be realized as a totally ramified extension of degree n/m of the (unique) unramified extension of degree m, some m dividing n. Clearly there are only finitely many such L's (up to K-isomorphism).

A Brief Introduction to PARI. Pari is a program designed for computations in number theory. It was written by H. Cohen and others, and is available from `ftp://megrez.math.u-bordeaux.fr/pub/pari/`. See also the Pari home page `http://pari.home.ml.org/`. It runs under Windows 95 and other operating systems. The following are a few commands for version 2 (they have been changed from version 1.x).

To start PARI on a network, type `gp` To quit PARI, type $\backslash q$

`nfbasis(f)` finds an integral basis for the field generated by a root of f.

`nfdisc(f)` finds the discriminant of f .

`polcyclo(n)` finds the n^{th} cyclotomic polynomial.

`polgalois(f)` finds the Galois group of f (f irreducible of degree ≤ 11).

`newtonpoly(f,p)` finds the Newton polygon of f.

`factor(f)` finds the factors of f .

`factormod(f,p)` factor f modulo p.

`quadunit(x)` finds the fundamental unit in the real quadratic field with discriminant x.

The syntax for polynomials is similar to that in Maple, e.g., `x^2+3*x+5.`

8. Global Fields

A *global field* is defined to be an algebraic number field (finite extension of $\mathbb{Q}$) or a function field in one variable over a finite field (finite extension of $\mathbb{F}_q(T)$ for some q). We are mainly interested in the number field case.

Extending valuations. Let K be a field with a valuation $|\ |$ (archimedean or discrete nonarchimedean), and let L be a finite separable extension of K. When K is complete, we know that there is a unique extension of $|\ |$ to L (see 7.29, 7.30), and we want to understand the extensions when K is not complete.

Write $L = K[\alpha]$, and let $f(X)$ be the minimum polynomial of α over K. Let $|\ |'$ be an extension of $|\ |$ to L. Then we can form the completion $\hat{L}$ of L with respect to $|\ |'$, and obtain a diagram:

$$\begin{array}{ccc} L & \hookrightarrow & \hat{L} \\ | & & | \\ K & \hookrightarrow & \hat{K}. \end{array}$$

Then $\hat{L} = \hat{K}[\alpha]$ (because every element ξ of $\hat{L}$ is the limit of a Cauchy sequence $\xi(n)$ in L; write $\xi(n) = \sum a_i(n)\alpha^n$, $a_i(n) \in K$; then each $a_i(n)$ is a Cauchy sequence in K, with limit a_i say in $\hat{K}$, and $\xi = \sum a_i\alpha^i$). Let $g(X)$ be the minimum polynomial of α over $\hat{K}$. Since $f(\alpha) = 0$, $g(X)|f(X)$, and so with each extension of $|\ |$, we have associated an irreducible factor of $f(X)$ in $\hat{K}[X]$.

Conversely, let $g(X)$ be a monic irreducible factor of $f(X)$ in $\hat{K}[X]$, and let $\hat{K}[x] = \hat{K}[X]/(g(X))$. Then we obtain a diagram:

$$\begin{array}{ccc} L & \overset{\alpha\mapsto x}{\to} & \hat{K}[x] \\ | & & | \\ K & \to & \hat{K}. \end{array}$$

According to (7.29, 7.30), the valuation on $\hat{K}$ extends uniquely to $\hat{K}[x]$, and this induces a valuation on L extending $|\ |$.

These two operations are inverse, and so we have proved the following result:

Proposition 8.1. *Let $L = K[\alpha]$ be a finite separable extension of K, and let $f(X)$ be the minimum polynomial of α over K. Then there is a natural one-to-one correspondence between the extensions of $|\ |$ to L and the irreducible factors of $f(X)$ in $\hat{K}[X]$.*

There is a more canonical way of obtaining the completions of L for the various extensions of $|\ |$.

Proposition 8.2. *Let $|\ |$ be a valuation on K (archimedean or discrete nonarchimedean) and let L be a finite separable extension of K. Let $\hat{K}$ be the completion of K with respect to $|\ |$. Then $|\ |$ has finitely many extensions $|\ |_1, \dots, |\ |_g$ to L; if L_i denotes the completion of L with respect to the valuation $|\ |_i$, then*

$$L \otimes_K \hat{K} \approx \prod L_i.$$

PROOF. Since L is separable over K, $L = K[\alpha] \approx K[X]/(f(X))$ for some element $\alpha \in L$ and its minimum polynomial $f(X)$. Suppose $f(X)$ factors in $\hat{K}[X]$ as

$$f(X) = f_1(X) \cdot f_2(X) \cdots f_g(X)$$

with $f_i(X)$ monic and irreducible. Then (see 1.13)

$$L \otimes_K \hat{K} = K[\alpha] \otimes_K \hat{K} \approx \hat{K}[X]/((f(X)) \approx \prod \hat{K}[X]/(f_i(X))$$

and so the proposition follows from (8.1). □

REMARK 8.3. Suppose now that K is a number field, that $\mathcal{O}_L = \mathcal{O}_K[\alpha]$, and that $|\ | = |\ |_{\mathfrak{p}}$ for some prime ideal $\mathfrak{p}$ in $\mathcal{O}_K$. Because $f_i(X)$ is irreducible in $\hat{K}[X]$, Hensel's lemma shows that, modulo $\hat{\mathfrak{p}}$, $f_i(X)$ is a power of an irreducible polynomial, say,

$$\bar{f}_i(X) = g_i(X)^{e_i}.$$

Then

$$\bar{f}(X) = \prod_{i=1}^{g} g_i(X)^{e_i},$$

and (3.43) tells us that

$$\mathfrak{p}\mathcal{O}_L = \prod \mathfrak{P}_i^{e_i}, \quad \mathfrak{P}_i = (\mathfrak{p}, g_i(\alpha)).$$

The valuations extending $|\ |_{\mathfrak{p}}$ correspond to the primes $\mathfrak{P}_i$, and so the two descriptions of the extensions agree.

COROLLARY 8.4. *In the situation of the Proposition, for any element $\alpha \in L$,*

$$\mathrm{Nm}_{L/K}(\alpha) = \prod \mathrm{Nm}_{L_i/\hat{K}}(\alpha), \quad \mathrm{Tr}_{L/K}(\alpha) = \sum \mathrm{Tr}_{L_i/\hat{K}}(\alpha).$$

(in the i^{th} factor or summand on the right, α is regarded as an element of L_i).

PROOF. By definition the norm and trace of α are the determinant and trace of the K-linear map $x \mapsto \alpha x \colon L \to L$. These don't change when L is tensored with $\hat{K}$, and it easy to see that norms and traces in products break up into products and sums respectively. □

EXAMPLE 8.5. According to Maple

$$f(X) = X^6 + 5X^5 + 5X^3 + 25X + 125$$

is irreducible in $\mathbb{Q}[X]$. Its Newton polygon for ord_5 has three segments of x-lengths 3, 2, 1 respectively, and so it has at least three factors in $\mathbb{Q}_5$. The discriminant of $f(X)$ is

$$2^4 5^{11} (59)(365587),$$

and so according to (7.28), to find the number of factors of $f(X)$ in $\mathbb{Q}_5[X]$, it suffices to factor in modulo 5^{11}. Better, according to Pari, version 2, $f(X)$ has exactly 3 irreducible factors in $\mathbb{Q}_5[X]$. (Type `factorpadic(f,p,r)` where r is the precision required.)

Suppose we find a factorization

$$f(X) = f_1(X) f_2(X) f_3(X)$$

(to whatever degree of accuracy we wish). To compute $|\beta|_i$, map $\beta = \sum c_j\alpha^j$ to $\beta_i = \sum c_j\alpha_i^j \in L_i \stackrel{\text{df}}{=} \mathbb{Q}_5[\alpha_i]$, α_i a root of $f_i(X)$, and use that

$$|\beta|_i = |\beta_i|_i = |\operatorname{Nm}_{L_i/\mathbb{Q}_5}\beta|_i^{1/\deg f_i}.$$

The product formula. Before proving the product formula for a number field, we need one extra fact for local fields.

Let K be a local field with normalized valuation $|\ |$. Recall that this means that $|\ |$ is the usual absolute value if K is $\mathbb{R}$, the square of the usual valuation if K is $\mathbb{C}$, and $|a| = (1/\mathbb{N}\mathfrak{p})^{\operatorname{ord}(a)}$ if the valuation is defined by a prime ideal $\mathfrak{p}$.

Let L be a finite separable extension of K, and let $|\ |$ be the unique extension of $|\ |$ to L. Let $\|\ \|$ be the normalized valuation on L corresponding to $|\ |$. What is the relation of $\|\ \|$ to $|\ |$?

LEMMA 8.6. *In the above situation, $\|a\| = |a|^n$, where $n = [L\colon K]$.*

PROOF. When K is archimedean, there are only two cases to consider, and both are obvious. Thus, assume K is nonarchimedean. Since, by assumption, $\|\ \| = |\ |^c$ for some c, we only have to check that the formula holds for a prime element π of K. Let Π be a prime element of L, and let $\mathfrak{P} = (\Pi)$ and $\mathfrak{p} = (\pi)$; then $\pi = (\text{unit}) \times \Pi^e$, and so

$$\|\pi\| = \|\Pi^e\| = (1/\mathbb{N}\mathfrak{P})^e = (1/\mathbb{N}\mathfrak{p})^{ef} = |\pi|^n,$$

as required.

Alternatively, use (7.34). For $a \in K$, we have

$$\|a\| \stackrel{\text{df}}{=} \mathbb{N}\mathfrak{P}^{-\operatorname{ord}_L a} \stackrel{7.34}{=} (\mathbb{N}\mathfrak{p}^f)^{-e\cdot\operatorname{ord}_K a} = |a|^{ef} = |a|^n.$$

□

PROPOSITION 8.7. *Let L/K be a finite extension of number fields. For any prime v of K and $\alpha \in L$,*

$$\prod_{w|v} \|\alpha\|_w = \|\operatorname{Nm}_{L/K}\alpha\|_v.$$

Here $\|\ \|_w$ and $\|\ \|_v$ denote the normalized valuations for the primes w and v.

PROOF. Let $|\ |_i$, $i = 1, 2, \dots, g$, be the extensions of $\|\ \|_v$ to L, and let $\|\ \|_i$ be the normalized valuation corresponding to $|\ |_i$. Then

$$\begin{aligned}\|\operatorname{Nm}_{L/K}\alpha\|_v \stackrel{8.4}{=} \|\textstyle\prod_{i=1}^g \operatorname{Nm}_{L_i/\hat{K}}\alpha\|_v &= \textstyle\prod_{i=1}^g \|\operatorname{Nm}_{L_i/\hat{K}}\alpha\|_v \\ &\stackrel{7.29}{=} \textstyle\prod_{i=1}^g |\alpha|_i^{n(i)} \stackrel{8.6}{=} \prod_{i=1}^g \|\alpha\|_w,\end{aligned}$$

where $n_i = [L_i\colon \hat{K}]$. □

THEOREM 8.8 (Product formula). *Let K be an algebraic number field; for all nonzero $\alpha \in K$,*

$$\prod_w \|\alpha\|_w = 1,$$

where the product is over the primes of K and $\|\ \|_w$ is the normalized valuation for the prime w.

PROOF. We have

$$\prod_w \|\alpha\|_w = \prod_v (\prod_{w|v} \|\alpha\|_w) = \prod_v \|\operatorname{Nm}\alpha\|_v$$

where v runs through the primes $2, 3, 5, 7, ..., \infty$ of $\mathbb{Q}$. The last product is 1 by (7.11). □

REMARK 8.9. E. Artin and Whaples (Bull. Amer. Math. Soc. **51** (1946), 469–492) proved the following characterization of global fields. Let K be a field with a set $\mathfrak{V}$ of primes (equivalence classes of valuations) satisfying the following axioms.

Axiom I. There is a set of representatives $|\ |_v$ for the primes such that, for any nonzero $a \in K$, $|a|_v \neq 1$ for only finitely many v and

$$\Pi_v |a|_v = 1 \text{ (product over all } v \in \mathfrak{V}).$$

Axiom II. There exists at least one prime v for which K_v is a local field.

Then K is a global field, and $\mathfrak{V}$ consists of all the primes for K.

Throughout his career, E. Artin promoted the idea that if only one could understand the similarities between function fields and number fields sufficiently well, then one could transfer proofs from function fields to number fields (e.g. the proof of the Riemann hypothesis!). This hasn't worked as well as he hoped, but the analogy has still been very fruitful. In the above paper, he suggested one should develop number theory and class field theory as much as possible working only from the axioms.

Decomposition groups. Let L be a finite Galois extension of a number field K, and let $G = \operatorname{Gal}(L/K)$. For a valuation w of L, we write σw for the valuation such that $|\sigma\alpha|_{\sigma w} = |\alpha|_w$, i.e., $|\alpha|_{\sigma w} = |\sigma^{-1}\alpha|_w$. For example, if w is the prime defined by a prime ideal $\mathfrak{P}$, then σw is the prime defined by the prime ideal $\sigma\mathfrak{P}$, because

$$|\alpha|_{\sigma w} < 1 \iff \sigma^{-1}\alpha \in \mathfrak{P} \iff \alpha \in \sigma\mathfrak{P}.$$

The group G acts on the set of primes of L lying over a fixed prime v of K, and we define the *decomposition (or splitting) group* of w to be the stabilizer of w in G; thus

$$G_w = \{\sigma \in G \mid \sigma w = w\}.$$

Equivalently, G_w is the set of elements of G that act continuously for the topology defined by $|\ |_w$. Each $\sigma \in G_w$ extends uniquely to a continuous automorphism of L_w. Note that $G_{\tau w} = \tau G_w \tau^{-1}$.

PROPOSITION 8.10. *The homomorphism $G_w \to \operatorname{Gal}(L_w/K_v)$ just defined is an isomorphism.*

PROOF. Clearly the map is injective, and so $(G_w : 1) \leq [L_w : K_v]$. The valuation σw has decomposition group $\sigma G_w \sigma^{-1}$, which has the same order as G_w, and so we also have $(G_w : 1) \leq [L_{\sigma w} : K_v]$. The number of distinct w's dividing v is $(G : G_w)$, and so

$$(G : 1) = (G : G_w)(G_w : 1) \leq \sum_{\sigma \in G/G_w} [L_{\sigma w} : K_v] \overset{(8.2)}{\leq} [L : K].$$

Hence equality holds: $(G_w : 1) = [L_w : K_v]$ (and G acts transitively on the primes dividing v, which we knew already from the proof of 3.36). □

Let $D(\mathfrak{P})$ (or $G(\mathfrak{P})$) be the decomposition group of $\mathfrak{P}$, so that $D(\mathfrak{P}) = \mathrm{Gal}(L_{\mathfrak{P}}/K_{\mathfrak{p}})$, and let $I(\mathfrak{P}) \subset D(\mathfrak{P})$ be the inertia group. We have the following picture:

$$\begin{array}{cccccccc}
 & \mathfrak{P} & L & - & L_{\mathfrak{P}} & & & \\
e & & | & & | & \diagdown & & \\
 & \mathfrak{P}_I & L^{I(\mathfrak{P})} & - & L_{\mathfrak{P}}^{I(\mathfrak{P})} & - & l & \\
f & & | & & | & & | & D(\mathfrak{P})/I(\mathfrak{P}) \\
 & \mathfrak{P}_D & L^{D(\mathfrak{P})} & - & K_{\mathfrak{p}} & - & k & \\
g & & | & \diagup & & & & \\
 & \mathfrak{p} & K & & & & &
\end{array}$$

Here:

$\mathfrak{P}_I = \mathfrak{P} \cap L^{I(\mathfrak{P})}$, $\mathfrak{P}_D = \mathfrak{P} \cap L^{D(\mathfrak{P})}$, $\mathfrak{p} = \mathfrak{P} \cap K$;

the fields in the second column are the completions of those in the first;

the fields in the third column are the residue fields of those in the second.

Proposition 8.11. *(a) The only prime ideal of L lying over $\mathfrak{P}_D$ is $\mathfrak{P}$.*

(b) The prime ideal $\mathfrak{P}_D$ is unramified in L^I, and $f(\mathfrak{P}_I/\mathfrak{P}_D) = f(\mathfrak{P}/\mathfrak{p})$.

(c) The prime ideal $\mathfrak{P}_I$ is totally ramified in L, and $e(\mathfrak{P}/\mathfrak{P}_I) = e(\mathfrak{P}/\mathfrak{p})$.

(d) If $D(\mathfrak{P})$ is normal in G, then

$$\mathfrak{p}\mathcal{O}_{L^D} = \prod \sigma\mathfrak{P}_D$$

where the product is over a set of representatives for $G/D(\mathfrak{P})$.

Proof. (a) Because L is Galois over $L^{D(\mathfrak{P})}$, its Galois group $D(\mathfrak{P})$ acts transitively on the set of prime ideals of L lying over $\mathfrak{P}_D$. Thus (a) is obvious from the definition of $D(\mathfrak{P})$.

(b), (c), (d) are similarly straightforward. □

The diagram, and the proposition, show that we can construct a chain of fields

$$L \supset L^I \supset L^D \supset K$$

such that all the ramification of $\mathfrak{P}$ over $\mathfrak{p}$ takes place in the top extension, all the residue field extension takes place in the middle extension, and, when L^D is normal over K, all the splitting takes place in the bottom extension. One should be a little careful about the last assertion when $D(\mathfrak{P})$ is not normal in G; all we know in general is that

$$\mathfrak{p} \cdot \mathcal{O}_{L^D} = \prod \mathfrak{P}_i^{e_i},\ \mathfrak{P}_1 = \mathfrak{P}_D$$

with $e_1 = 1 = f_1$ (i.e., in general $\mathfrak{p}$ will *not* split completely in L^D).

Remark 8.12. Let L be a Galois extension of $\mathbb{Q}$, with Galois group G. Suppose that $\mathcal{O}_L = \mathbb{Z}[\alpha]$ for some $\alpha \in L$. Let $f(X)$ be the minimum polynomial of α over $\mathbb{Q}$, and write $\bar{f}(X)$ for $f(X)$ modulo p. Choose an irreducible factor $g_1(X)$ of $\bar{f}(X)$, and let $g_1(X)^{e_1}$ be the largest power of $g_1(X)$ dividing $\bar{f}(X)$. According to Hensel's lemma, $g_1(X)^{e_1}$ lifts to an irreducible factor $f_1(X)$ of $f(X)$ in $\mathbb{Q}_p[X]$, which can be found to any desired degree of accuracy by factoring $f(X)$ modulo a high power of

p (essentially using the method of proof of Hensel's lemma). Let $\mathfrak{P}_1 = (p, h_1(\alpha))$ for any lifting h_1 of g_1 to $\mathbb{Z}[X]$. Then

$$D(\mathfrak{P}_1) = \{\sigma \in G \mid \sigma\mathfrak{P}_1 = \mathfrak{P}_1\},$$

which can be computed easily (provided G has been found explicitly as a subgroup of the symmetric group on the set of roots of $f(X)$). Let $\bar{\alpha}$ be the image of α in $\mathcal{O}_L/\mathfrak{P}_1 = \mathbb{F}_p[\bar{\alpha}]$. Then $g_1(X)$ is the minimum polynomial of $\bar{\alpha}$ over $\mathbb{F}_p$, and $I(\mathfrak{P}_1)$ is the subgroup of $D(\mathfrak{P}_1)$ fixing $\bar{\alpha}$. Finally $D(\mathfrak{P}_1)/I(\mathfrak{P}_1) = \mathrm{Gal}(\mathbb{F}_p[\bar{\alpha}]/\mathbb{F}_p)$.

Consider a tower of fields

$$\begin{array}{ccc} 1 & M & \mathfrak{P} \\ & | & \\ H & L & \mathfrak{P}_L \\ & | & \\ G & K & \mathfrak{p} \end{array}$$

Assume M is Galois over K with Galois group G, and that H is the subgroup of G fixing L. (Recall $D(\mathfrak{P})$ and $G(\mathfrak{P})$ are two notations for the same object.)

PROPOSITION 8.13. *Let $\mathfrak{P}$ be a prime ideal in $\mathcal{O}_M$, and let $\mathfrak{P}_L = \mathfrak{P} \cap L$.*

(a) The decomposition group $H(\mathfrak{P})$ of $\mathfrak{P}$ over L is $G(\mathfrak{P}) \cap H$.

(b) Suppose further that H is a normal subgroup of G, so that G/H is the Galois group of L/K. The decomposition group of $\mathfrak{P}_L$ over K is the image of $G(\mathfrak{P})$ in G/H.

PROOF. (a) Clearly

$$H(\mathfrak{P}) = \{\sigma \in G \mid \sigma \in H, \quad \sigma\mathfrak{P} = \mathfrak{P}\} = H \cap G(\mathfrak{P}).$$

(b) This is equally obvious. □

The Frobenius element. Let L/K be a Galois extension of number fields with Galois group G. Given an ideal $\mathfrak{P}$ of L that is unramified in L/K we define the Frobenius element $\sigma = (\mathfrak{P}, L/K)$ to be the element of $G(\mathfrak{P})$ that acts as the Frobenius automorphism on the residue field. Thus σ is uniquely determined by the following two conditions:

(a) $\sigma \in G(\mathfrak{P})$, i.e., $\sigma\mathfrak{P} = \mathfrak{P}$;

(b) for all $\alpha \in \mathcal{O}_L$, $\sigma\alpha \equiv \alpha^q \bmod \mathfrak{P}$, where q is the number of elements the residue field $\mathcal{O}_K/\mathfrak{p}$, $\mathfrak{p} = \mathfrak{P} \cap K$.

We now list the basic properties of $(\mathfrak{P}, L/K)$.

8.14. *Let $\tau\mathfrak{P}$ be a second prime dividing $\mathfrak{p}$, $\tau \in G$. Then $G(\tau\mathfrak{P}) = \tau G(\mathfrak{P})\tau^{-1}$, and*

$$(\tau\mathfrak{P}, L/K) = \tau(\mathfrak{P}, L/K)\tau^{-1}.$$

PROOF. Let $\alpha \in \mathcal{O}_L$; then

$$\tau\sigma\tau^{-1}(\alpha) = \tau((\tau^{-1}\alpha)^q + a), \text{ some } a \in \mathfrak{P}, \text{ and}$$

$$\tau((\tau^{-1}\alpha)^q + a) = \alpha^q + \tau a \equiv \alpha^q \bmod \tau\mathfrak{P}.$$

□

Thus if $\mathrm{Gal}(L/K)$ is abelian, then $(\mathfrak{P}, L/K) = (\mathfrak{P}', L/K)$ for all primes $\mathfrak{P}$, $\mathfrak{P}'$ dividing $\mathfrak{p}$, and we write $(\mathfrak{p}, L/K)$ for this element. If $\mathrm{Gal}(L/K)$ is not abelian, then

$$\{(\mathfrak{P}, L/K) \mid \mathfrak{P}|\mathfrak{p}\}$$

is a conjugacy class in G, which (by an abuse of notation) we again denote $(\mathfrak{p}, L/K)$. Thus, for a prime $\mathfrak{p}$ of K, $(\mathfrak{p}, L/K)$ is either an element of $\mathrm{Gal}(L/K)$ or a conjugacy class depending on whether $\mathrm{Gal}(L/K)$ is abelian or nonabelian.

8.15. *Consider a tower of fields*

$$\begin{array}{cc} M & \mathfrak{Q} \\ | & \\ L & \mathfrak{P} \\ | & \\ K & \mathfrak{p} \end{array}$$

and assume that $\mathfrak{Q}$ is unramified over $\mathfrak{p}$; then

$$(\mathfrak{Q}, M/K)^{f(\mathfrak{P}/\mathfrak{p})} = (\mathfrak{Q}, M/L).$$

Proof. Let $k(\mathfrak{Q}) \supset k(\mathfrak{P}) \supset k(\mathfrak{p})$ be the corresponding sequence of residue fields. Then $f(\mathfrak{P}/\mathfrak{p}) = [k(\mathfrak{P}) : k(\mathfrak{p})]$, and the Frobenius element in $\mathrm{Gal}(k(\mathfrak{Q})/k(\mathfrak{P}))$ is the $f(\mathfrak{P}/\mathfrak{p})^{th}$ power of the Frobenius element in $\mathrm{Gal}(k(\mathfrak{Q})/k(\mathfrak{p}))$. □

8.16. *In (8.15), assume that L is Galois over K; then*

$$(\mathfrak{Q}, M/K)|L = (\mathfrak{P}, L/K).$$

Proof. Obvious. □

Let L_1 and L_2 be Galois extensions of K contained in some field Ω, and let $M = L_1 \cdot L_2$. Then M is Galois over K, and there is a canonical homomorphism

$$\sigma \mapsto (\sigma|L_1, \sigma|L_2)\colon\ \mathrm{Gal}(M/K) \to \mathrm{Gal}(L_1/K) \times \mathrm{Gal}(L_2/K)$$

which is injective.

8.17. *Under the above map,*

$$(\mathfrak{Q}, M/K) \mapsto (\mathfrak{P}_1, L_1/K) \times (\mathfrak{P}_2, L_2/K).$$

Proof. This follows from (8.16). □

Note that $\mathfrak{p}$ splits completely in L if and only if $(\mathfrak{P}, L/K) = 1$ for one (hence all) primes $\mathfrak{P}$ lying over it. Hence, in the situation of (8.17), $\mathfrak{p}$ splits completely in M if and only if it splits completely in L_1 and L_2.

Examples. We find the Frobenius maps for quadratic and cyclotomic fields, and obtain a surprisingly simple proof of the quadratic reciprocity law.

Example 8.18. Let $K = \mathbb{Q}[\zeta_n]$, where ζ_n is a primitive n^{th} root of 1. If $p|n$ then p ramifies in K, and $(p, K/\mathbb{Q})$ is not defined. Otherwise $\sigma = (p, K/\mathbb{Q})$ is the unique element of $\mathrm{Gal}(K/\mathbb{Q})$ such that

$$\sigma\alpha \equiv \alpha^p \bmod \mathfrak{p}, \quad \text{for all } \alpha \in \mathbb{Z}[\zeta_n],$$

for any prime ideal $\mathfrak{p}$ lying over p.

I claim that σ is the element of the Galois group such that $\sigma(\zeta_n) = \zeta_n^p$: let $\mathfrak{p}$ be a prime lying over p in $\mathbb{Z}[\zeta_n]$; then modulo $\mathfrak{p}$, we have,

$$\sigma(\sum a_i\zeta_n^i) = \sum a_i\zeta_n^{ip} \equiv \sum a_i^p\zeta_n^{ip} \equiv (\sum a_i\zeta_n^i)^p$$

as required.

Note that $(p, K/\mathbb{Q})$ has order f where f is the smallest integer such that $n|p^f - 1$ (because this is the order of p in $(\mathbb{Z}/(n))^\times$).

EXAMPLE 8.19. Let $K = \mathbb{Q}[\sqrt{d}]$, and let p be a prime that is unramified in K. Identify $\mathrm{Gal}(K/\mathbb{Q})$ with $\{\pm 1\}$. Then $(p, K/\mathbb{Q}) = +1$ or -1 according as p does, or does not, split in K, i.e., according as d is, or is not, a square modulo p. Thus $(p, K/\mathbb{Q}) = (\frac{d}{p})$.

Application: the quadratic reciprocity law. Let $K = \mathbb{Q}[\zeta]$, where ζ is a primitive p^{th} root of 1, $p \neq 2$. Because $\mathrm{Gal}(K/\mathbb{Q}) \cong (\mathbb{Z}/p\mathbb{Z})^\times$ is cyclic of order $p-1$, it contains a unique subgroup of order $(p-1)/2$ (consisting of the elements of $(\mathbb{Z}/p\mathbb{Z})^\times$ that are squares), and hence K contains a unique quadratic extension F of $\mathbb{Q}$. If $p \equiv 1 \mod 4$, then p is the only prime ramifying in $\mathbb{Q}[\sqrt{p}]$, and $\mathbb{Q}[\sqrt{p}]$ is the only quadratic field for which this is true. Similarly if $p \equiv 3 \mod 4$, then $-p \equiv 1 \mod 4$, and $-p$ is the only prime ramifying in $\mathbb{Q}[\sqrt{-p}]$. Thus $F = \mathbb{Q}[\sqrt{d}]$ where $d = (-1)^{(p-1)/2} \cdot p$.

If q is an odd prime $\neq p$; then

$$(q, K/\mathbb{Q})(\zeta) = \zeta^q.$$

Thus $(q, K/\mathbb{Q})$ restricts to the identity element of $\mathrm{Gal}(\mathbb{Q}[\sqrt{d}]/\mathbb{Q})$ or not according as q is a square in $(\mathbb{Z}/p\mathbb{Z})^\times$ or not. Thus $(q, K/\mathbb{Q})|\mathbb{Q}[\sqrt{d}] = (\frac{q}{p})$. But we know that it is also equal to $(\frac{d}{q})$. Hence

$$(\frac{q}{p}) = (\frac{-1}{q})^{(p-1)/2} \cdot (\frac{p}{q}) = (-1)^{(p-1)(q-1)/4} \cdot (\frac{p}{q}).$$

Here we have used that -1 is square in $\mathbb{F}_q$ if and only if $4|q-1$, so that $(\frac{-1}{q}) = (-1)^{(q-1)/2}$. The displayed formula, together with the statements

$$(\frac{-1}{p}) = (-1)^{(p-1)/2},\ (\frac{2}{p}) = (-1)^{(p^2-1)/8}$$

constitute the *quadratic reciprocity law.* (For the last formula, see Serre, Cours d'Arithmétique I.3.2; the proof of the rest of the reciprocity law there, ibid. 3.3., is essentially the above proof, made elementary.)

Computing Galois groups (the hard way). Let $f(X)$ be a polynomial over a field K, and let $\alpha_1, \dots, \alpha_n$ be the roots of $f(X)$ in K^{al}. We want to determine the Galois group of f as a subgroup of the group of permutations S_n of $\{\alpha_1, \dots, \alpha_n\}$.

Introduce variables $t_1, \dots, t_n$. For any $\sigma \in S_n$ and polynomial $f(t_1, \dots, t_n)$, define $\sigma_t f = f(t_{\sigma(1)}, \dots, t_{\sigma(n)})$. Let $\theta = \sum \alpha_i t_i$, and define a polynomial

$$F(X, t) = \prod(X - \sigma_t\theta) \quad (\text{product over } \sigma \in S_n).$$

The coefficients of this polynomial are symmetric polynomials in the α_i, and so lie in K. Now factor

$$F(X, t) = F_1(X, t) \cdots F_r(X, t)$$

in $K[X, t_1, \dots, t_n]$.

THEOREM 8.20. *Let G be the set of $\sigma \in S_n$ such that σ_t fixes $F_1(X,t)$; then G is the Galois group of f.*

PROOF. See van der Waerden, Algebra, Vol 1, §61 (Calculation of the Galois group). □

This theorem gives an algorithm (unfortunately impractical) for computing the Galois group of a polynomial $f(X) \in \mathbb{Q}[X]$. We may suppose $f(X)$ to be monic with integer coefficients. First find the roots of $f(X)$ to a high degree of accuracy. Then compute $F(X,t)$ exactly, noting that this has coefficients in $\mathbb{Z}$. Factor $F(X,t)$, and take one of the factors $F_1(X,t)$. Finally list the elements σ of S_n such that σ_t fixes $F_1(X,t)$. The problem with this approach is that $F(X,t)$ has degree $n!$. It will probably work (on a computer) if $n \leq 5$, but otherwise it is like trying to compute a determinant directly from the definition as a sum of products.

Computing Galois groups (the easy way). We now give a more practical procedure (also largely in van der Waerden with a more direct proof).

PROPOSITION 8.21. *Let $f(X)$ be a monic separable polynomial of degree n over a field K, and suppose that the Galois group G of $f(X)$ has s orbits (as a group of permutations of the roots of f) with $n_1, \dots, n_s$ elements respectively (so that $n_1 + n_2 + \cdots + n_s = n$); then there is a factorization*

$$f(X) = f_1(X) \cdots f_r(X)$$

with $f_i(X)$ an irreducible polynomial in $K[X]$ of degree n_i.

PROOF. Write $f(X) = \prod(X - \alpha_i)$. For $S \subset \{1, 2, \dots, n\}$, consider $f_S = \prod_{i \in S}(X - \alpha_i)$. This polynomial divides $f(X)$, and it is fixed under the action of G (and hence has coefficients in K) if and only if S is stable under G. Therefore the irreducible factors of $f(X)$ are the polynomials f_S with S a minimal subset of $\{1, \dots, n\}$ stable under G, but such sets S are precisely the orbits of G in $\{1, 2, \dots, n\}$. □

Let $\sigma \in S_n$. In GT, §4, it is proved that σ is a product of disjoint cycles. More precisely, if

$$o_1 = \{m_{11}, \dots, m_{1n_1}\}, \quad o_2 = \{m_{21}, \cdots, m_{2n_2}\}, \quad \dots$$

are the orbits of $\langle\sigma\rangle$ acting on $\{1, 2, ..., n\}$, numbered in such a way that $\sigma m_{ij} = m_{i\,j+1}$, then

$$\sigma = (m_{11} \dots m_{1n_1}) \cdot (m_{21} \dots m_{2n_2}) \cdot \dots .$$

This remark, together with (8.21), gives us the following result.

COROLLARY 8.22. *Let $f(X)$ be a monic separable polynomial of degree n over a finite field k, and let ℓ be the splitting field of $f(X)$. Suppose that the Frobenius element $\sigma \in \mathrm{Gal}(\ell/k)$ (when regarded as a permutation of the roots of $f(X)$) is a product of disjoint cycles $\sigma = c_1 \cdots c_s$ with c_i of length n_i (so that $\sum n_i = n$). Then $f(X)$ factors as a product of irreducible polynomials in $k[X]$*

$$f(X) = f_1(X) \cdots f_r(X)$$

with f_i of degree n_i.

In other words, the type of the cycle decomposition of σ can be read off from the factorization of $f(X)$.

THEOREM 8.23 (Dedekind). *Let $f(X)$ be a polynomial of degree n over a number field K, and let G be the Galois group of f. Assume $f(X) \in \mathcal{O}_K[X]$ and is monic. Let $\mathfrak{p}$ be a prime ideal of K, and suppose that*

$$f(X) \equiv f_1(X) \cdots f_r(X) \bmod \mathfrak{p}$$

with the f_i distinct irreducible polynomials in $k[X]$ and f_i of degree n_i, $k = \mathcal{O}_K/\mathfrak{p}$. Then G contains a permutation σ that is a product of disjoint cycles of length n_i.

PROOF. Take σ to be the Frobenius element of any prime lying over $\mathfrak{p}$ — the hypothesis on the factorization of $f(X)$ mod $\mathfrak{p}$ implies that $\mathfrak{p}$ is unramified in the splitting field (because it implies that $\mathfrak{p}$ doesn't divide the discriminant of f). □

REMARK 8.24. There is a similar statement for real primes, namely, if $f(X) = f_1(X) \cdots f_r(X)$ in $\mathbb{R}[X]$ with $f_1, \dots, f_j$ of degree 2 and the remainder of the degree 1, then G contains a permutation σ that is a product of disjoint j cycles of length 2.

This suggests the following strategy for factoring a polynomial $\mathbb{Q}[X]$: factor $f(X)$ modulo many primes p; discard the result if $f(X)$ mod p has multiple factors; continue until a sequence of, say n, primes has yielded no new cycle types for the elements. Then attempt to read off the type of the group from tables. We discuss how effective this is later.

EXAMPLE 8.25. Let $f(X) = X^5 - X - 1$. Modulo 2 this factors as $(X^2 + X + 1)(X^3 + X^2 + 1)$; modulo 3 it is irreducible. Hence G contains (12345) and $(ik)(\ell mn)$ for some numbering of the roots. It also contains $((ik)(\ell mn))^3 = (ik)$, and this implies that $G = S_5$ (see 8.28 below).

LEMMA 8.26. *Let H be a subgroup of S_n; if H is transitive (for example, contains an n-cycle) and contains an $(n-1)$-cycle and a transposition, then $H = S_n$.*

PROOF. After possibly renumbering, we may suppose that the $(n-1)$-cycle is (1 2 3 ... $n-1$). By virtue of the transitivity, the transposition can be transformed into (in), some $i \le n-1$. Now the $(n-1)$-cycle and its powers will transform this into $(1\ n)$, $(2\ n)$, ..., $(n-1\ n)$, and these elements obviously generate S_n (because S_n is generated by transpositions). □

EXAMPLE 8.27. Select monic polynomials of degree n, f_1, f_2, f_3 with coefficients in $\mathbb{Z}$ such that

(a) f_1 is irreducible modulo 2;

(b) $f_2 =$(degree 1)(irreducible of degree $n-1$) mod 3;

(c) $f_3 =$(irreducible of degree 2)(product of one or two irreducible polynomials of odd degree) mod 5. We need to choose f_3 to have distinct roots modulo 5.

Take

$$f = -15f_1 + 10f_2 + 6f_3,$$

and let G be the Galois group of f. Then

(a′) G is transitive (it contains an n-cycle because of (a));

(b′) G contains a cycle of length $n-1$;

(c′) G contains a transposition (because it contains the product of a transposition with a commuting element of odd order).

The above lemma shows that $G = S_n$.

REMARK 8.28. There are other criteria for a subgroup H of S_n to be all of S_n. For example, a subgroup H of S_p, p prime, that contains an element of order p and a transposition is equal to S_p. (See FT, Lemma 4.12, or Jacobson, Basic Algebra I, 4.10.)

REMARK 8.29. In Pohst and Zassenhaus 1989, p. 73, there are suggestions for constructing irreducible polynomials $f(X)$ of degree n in $\mathbb{F}_p[X]$. A root of such a polynomial will generate $\mathbb{F}_q$, $q = p^n$, and so every such $f(X)$ will divide $X^q - X$. One can therefore find all $f(X)$s by factoring $X^q - X$.

For example, consider $X^{125} - X \in \mathbb{F}_5[X]$. Its splitting field is $\mathbb{F}_{125}$, which has degree 3 over $\mathbb{F}_5$. The factors of $X^{125} - X$ are the minimum polynomials of the elements of $\mathbb{F}_{125}$. They therefore have degree 1 or 3. There are 5 linear factors, X, $X-1$, $X-2$, $X-3$, $X-4$, and 40 cubic factors, which constitute a complete list of all the monic irreducible cubic polynomials in $\mathbb{F}_5[X]$. (Maple factored $X^{125} - X$ in 2 seconds and $X^{625} - X$ in 13 seconds on 1992 notebook computer. PARI version 2 is much slower, and tends to run out of memory.)

However, if you only want one irreducible polynomial of degree n, it is easier to write down a polynomial at random, and check whether it is irreducible.

Cubic polynomials. The group S_3 has the following subgroups:

order	group	group elements
1	1	1
2	C_2	$1 \times 1 + 1 \times 2$
3	A_3	$1 \times 1 + 2 \times 3$
6	S_3	$1 \times 1 + 3 \times 2 + 2 \times 3$.

By the last row, I mean S_3 has one 1-cycle, three 2-cycles, and two 3-cycles.

Note that any subgroup of S_3 containing cycles of length 2 and 3 is the whole of S_3; thus if f is irreducible modulo some prime and has an irreducible factor of degree 2 modulo a second prime, then its Galois group is S_3. On the other hand, if factorizing f modulo many primes doesn't turn up a factor of degree 2, but f is irreducible, then expect the Galois group of f to be A_3. This can be checked by seeing whether $\operatorname{disc}(f)$ is a square. For example, the calculations on p. 61 show that the polynomials $X^3 + 10X + 1$ and $X^3 - 8X + 15$ both have Galois group S_3.

To make this more effective (in the technical sense), we need the Chebotarev density theorem.

Chebotarev density theorem.

DEFINITION 8.30. Let S be a set of finite primes in a number field K, and let P be the set of all finite primes. We say that S has *natural density* δ if

$$\lim_{N\to\infty} \frac{\#\{\mathfrak{p} \in S \mid \mathbb{N}\mathfrak{p} \le N\}}{\#\{\mathfrak{p} \mid \mathbb{N}\mathfrak{p} \le N\}} = \delta.$$

THEOREM 8.31 (Chebotarev density theorem). *Let L be a finite Galois extension of the number field K, with Galois group G, and let C be a conjugacy class in G. The set of prime ideals $\mathfrak{p}$ of K such that $(\mathfrak{p}, L/K) = C$ has density $\delta = \#C/\#G$.*

PROOF. See my notes CFT (in fact, normally one proves this result with a slightly weaker notion of density). □

For example, if G is abelian, then for each $\sigma \in G$, the set of $\mathfrak{p}$ such that $(\mathfrak{p}, L/K) = \sigma$ has density $1/\#G$.

COROLLARY 8.32. *The primes that split in L have density $1/[L:K]$. In particular, there exist infinitely many primes of K not splitting in L.*

REMARK 8.33. There is a bound for the error in (8.26) (in terms of the discriminant of the polynomial), but it is too large to be of practical importance. However the existence of the bound has the following consequence: given a polynomial $f(X) \in \mathbb{Q}[X]$ (say), there exists a bound B such that, if a given cycle type doesn't occur as the Frobenius element of some $p \leq B$, then it doesn't occur at all. [For a discussion of the effective Chebotarev density theorem, see Lagarias and Odlysko, in *Algebraic Number Fields,* ed. A Fröhlich.]

EXAMPLE 8.34. Let $K = \mathbb{Q}[\zeta_n]$. Then $\mathrm{Gal}(K/\mathbb{Q}) = (\mathbb{Z}/n\mathbb{Z})^\times$ and $(p, K/\mathbb{Q}) = [p]$. The Chebotarev density theorem says that the primes are equidistributed among the congruence classes. In other words, each of the arithmetic progression

$$k,\ k+n,\ k+2n,\ k+3n, \ldots \quad \gcd(k,n) = 1,$$

contains $1/\varphi(n)$ of the primes. In particular, each of the arithmetic progressions contains infinitely many primes. This statement was conjectured by Legendre and proved by Dirichlet (using Dirichlet series). The proof of the Chebotarev density theorem is a generalization of that of Dirichlet.

EXAMPLE 8.35. In a quadratic extension, half the primes split and half the primes remain prime.

EXAMPLE 8.36. Let f be a cubic polynomial with coefficients in $\mathbb{Q}$. The Chebotarev density theorem implies the following statements (see the above table):

$G = 1$: f splits modulo all primes.

$G = C_2$: f splits for $1/2$ of the primes and has an irreducible factor of degree 2 for $1/2$ of the primes.

$G = A_3$: f splits for $1/3$ of the primes and f remains irreducible for $2/3$ of the primes.

$G = S_3$: f splits for $1/6$ of the primes, has a factor of degree 2 for $1/2$ of the primes, and remains prime for $1/3$ of the primes.

EXAMPLE 8.37. Let f be a quartic polynomial with no linear factor.

(a) When $\mathrm{Disc}(f)$ is a square, the possible Galois groups are:

order	group	elements
2	C_2	$1 \times 1 + 1 \times 2^2$
4	V_4	$1 \times 1 + 3 \times 2^2$
12	A_4	$1 \times 1 + 3 \times 2^2 + 8 \times 3$

(b) When $\mathrm{Disc}(f)$ is not a square, the possible Galois groups are:

order	group	elements
4	C_4	$1\times 1+1\times 2^2+2\times 4$
8	D_8	$1\times 1+2\times 2+3\times 2^2+2\times 4$
24	S_4	$1\times 1+3\times 2^2+6\times 2+8\times 3+6\times 4$

Thus if f is a polynomial of degree 4 with Galois group D_8, then it will split modulo p for 1/8 of the primes, factor as the product of a quadratic and two linear polynomials for 1/4 of the primes, factor as the product of two quadratics for 3/8 of the primes, and remain irreducible for 1/4 of the primes.

For a similar table for polynomials of degree 5, see Pohst and Zassenhaus 1989, p132.

The strategy for determining the Galois group of a polynomial is

(a) test whether f is irreducible over $\mathbb{Q}$;

(b) compute the discriminant of f;

(c) factor f modulo good primes (i.e., those not dividing the discriminant) until you seem to be getting no new cycle types;

(d) compute the orbit lengths on the r-sets of roots (these are the degrees of the irreducible factors in $\mathbb{Q}[X]$ of the polynomial whose roots are the sums of r roots of f);

(e) ad hoc methods.

As late as 1984, it had not been proved that the Mathieu group M_{11} occurs as a Galois group over $\mathbb{Q}$ (M_{11} is subgroup of S_{11} of order $11!/5040 = 7920$).

References.

Butler, G., and McKay, J., The transitive groups of degree up to eleven, Comm. Algebra 11 (1983), 863-911. (This lists all transitive subgroups of S_n, $n \le 11$, and gives the cycle types of their elements and the orbit lengths of the subgroup acting on the r-sets of roots; with a few exceptions, these invariants are sufficient to determine the subgroup up to isomorphism.)

Cohen 1993, Section 6.3.

Ford, D., and McKay, J., in Computer Algebra, QA155.7 E4 C6491.

Pohst and Zassenhaus 1989. Chapter 2 is entirely concerned with computing Galois groups; for example, II.10.8 discusses the following problem: given $G \subset H \subset S_n$, determine whether G is contained in a given smaller subgroup J of G.)

Soicher, L., An algorithm for computing Galois groups, in Computational Group Theory, ed M. Atkinson, 1984, 291-296.

Soicher and McKay, Computing Galois groups over the rationals, J. Number Theory, 20 (1985) 273-281.

Programs for finding the Galois group of a polynomial of degree ≤ 7 are implemented in Maple and in PARI (PARI now claims ≤ 11).

Applications of the Chebotarev density theorem. We now discuss some other applications of the Chebotarev density theorem.

For any extension L/K of number fields, write $\mathrm{Spl}(L/K)$ for the set of primes that split completely in L, and write $\mathrm{Spl}'(L/K)$ for the set of primes that have at least one split factor. Then $\mathrm{Spl}(L/K) \subset \mathrm{Spl}'(L/K)$ always, and equality holds if L/K is Galois, in which case the Chebotarev density theorem shows that $\mathrm{Spl}(L/K)$ has density $1/[L:K]$.

THEOREM 8.38. *If L and M are Galois over K, then*

$$L \subset M \iff \mathit{Spl}(L/K) \supset \mathit{Spl}(M/K).$$

PROOF. $\Rightarrow$: This is obvious.

$\Leftarrow$: We have

$$\mathrm{Spl}(LM/K) = \mathrm{Spl}(L/K) \cap \mathrm{Spl}(M/K).$$

To see this, note that

$$\begin{aligned} \mathfrak{p} \in \mathrm{Spl}(LM/K) &\iff (\mathfrak{p}, LM/K) = 1 \\ &\iff (\mathfrak{p}, LM/K)|L = 1 \text{ and } (\mathfrak{p}, LM/K)|M = 1; \end{aligned}$$

but $(\mathfrak{p}, LM/K)|L = (\mathfrak{p}, L/K)$ and $(\mathfrak{p}, LM/K)|M = (\mathfrak{p}, M/K)$. Now

$$\begin{aligned} \mathrm{Spl}(M/K) \subset \mathrm{Spl}(L/K) &\Rightarrow \mathrm{Spl}(LM/K) = \mathrm{Spl}(M/K) \\ &\overset{8.31}{\Rightarrow} [LM:K] = [M:K] \Rightarrow L \subset M. \end{aligned}$$

□

COROLLARY 8.39. *If L and M are Galois over K, then*

$$L = M \iff \mathit{Spl}(M/K) = \mathit{Spl}(L/K).$$

PROOF. Obvious from the Proposition. □

REMARK 8.40. (a) In fact, $L = M$ if $\mathrm{Spl}(M/K)$ and $\mathrm{Spl}(L/K)$ differ by at worst a finite set of primes (or if they differ by at worst a set of primes of density zero).

(b) The effective form of the Chebotarev density theorem shows that (8.38) is effective: in order to show that $L \subset M$ it suffices to check that

$$\mathfrak{p} \text{ splits in } M \Rightarrow \mathfrak{p} \text{ splits in } L$$

for all primes $\mathfrak{p}$ less than some bound.

(c) Proposition 8.39 is not true without the Galois assumptions: there exist non-isomorphic extensions L and M of $\mathbb{Q}$ such that $\mathrm{Spl}(L/K) = \mathrm{Spl}(M/K)$. In fact there exist nonisomorphic extensions L and M of $\mathbb{Q}$ of the same degree such that

(i) L and M have the same discriminant;

(ii) a prime p not dividing the common discriminant decomposes in exactly the same way in the two fields.

(d) It is clear from (8.39) that if a separable polynomial $f(X) \in K[X]$ splits into linear factors mod $\mathfrak{p}$ for all but finitely many primes $\mathfrak{p}$ of K, then $f(X)$ splits into linear factors in $K[X]$. With a little more work, one can show that an *irreducible* polynomial $f(X) \in K[X]$ can not have a root mod $\mathfrak{p}$ for all but a finite number of primes. This last statement is false for reducible polynomials—consider for example,

$$(X^2-2)(X^2-3)(X^2-6).$$

For more on these questions, see Cassels and Fröhlich 1967, Exercise 6, p361.

EXAMPLE 8.41. Fix a number field K. According to (8.39), a Galois extension L of K is determined by the set $\mathrm{Spl}(L/K)$. Thus, in order to classify the Galois extensions of K, it suffices to classify the sets of primes in K that can occur as $\mathrm{Spl}(L/K)$. For abelian extensions of K, class field theory does this — see CFT (they are determined by congruence conditions). For nonabelian extensions the sets are still a mystery — they are not determined by congruence conditions — but Langlands's conjectures shed some light.

Topics not covered.

More algorithms. Let K be a number field. There is a rather simple algorithm for factoring a polynomial $f(X) \in K[X]$ which involves only:

(a) forming resultants of polynomials over $\mathbb{Q}$, and
(b) factoring polynomials over $\mathbb{Q}$,

both of which Maple and Mathematica can do. However, Maple knows how to factor polynomials over number fields other than $\mathbb{Q}$.

The Hasse principle for quadratic forms. Consider a quadratic form

$$Q(X_1, \dots, X_n) = \sum a_{ij} X_i X_j$$

over a field k. By a *nontrivial zero* of Q we mean an n-tuple $(a_1, \dots, a_n) \neq (0, \dots, 0)$ such that

$$Q(a_1, \dots, a_n) = 0.$$

THEOREM 8.42. *A quadratic form Q over $\mathbb{Q}$ has a nontrivial zero in $\mathbb{Q}$ if and only if it has a nontrivial zero in $\mathbb{R}$ and in $\mathbb{Q}_p$ for each p.*

PROOF. The necessity is obvious. The key point in the sufficiency is the quadratic reciprocity law. See Serre, Cours d'Arithmétique, Ch. IV. □

The same theorem holds for any number field K, with much the same proof, except that there is no longer an elementary proof of the quadratic reciprocity law. Instead, one obtains it as a special case of the Artin reciprocity law, which is proved in CFT — see CFT VIII.1 for a proof of the theorem over an arbitrary number field.

Algebraic function fields. Appropriately interpreted, everything we have proved for algebraic number fields also holds for finite extensions of $\mathbb{F}_p(X)$.

Suppose K is such a field, and let k be the algebraic closure of $\mathbb{F}_p$ in K. It is a finite field, and it is possible to find an element $X \in K$ such that K is a finite *separable* extension of $k(X)$. Let $\mathcal{O}_K$ be the integral closure of $k[X]$ in K. According to (3.30), it is a Dedekind domain, and one can show that it has finite class number. Moreover, $K^\times$ contains only finitely many roots of unity, and the rank of $\mathcal{O}_K^\times$ is $s-1$ where s is the "number of primes at infinity". Let $\mathcal{O}'_K$ be the integral closure of $k[X^{-1}]$ in K; then the primes at infinity are those corresponding to the prime ideals in $\mathcal{O}'_K$ lying over the prime ideal $(X^{-1}) \subset k[X^{-1}]$. See for example, Cohn 1991. Generally, the same proofs work in the two cases.

I haven't discussed function fields in this course from lack of time and because it is possible to give a more geometric approach in the function field case. Every K

as above can be realized as the field of algebraic functions on a unique nonsingular projective algebraic curve C over k, and it is more illuminating to discuss the arithmetic of K in terms of the geometry of C than to simply translate the proofs from the number field case.

> And after the first year [as an undergraduate at Göttingen] I went home with Hilbert's Zahlbericht under my arm, and during the summer vacation I worked my way through it — without any previous knowledge of elementary number theory or Galois theory. These were the happiest months of my life, whose shine, across years burdened with our common share of doubt and failure, still comforts my soul.
>
> Hermann Weyl, Bull. Amer. Math. Soc. 50 (1944), 612–654.

Exercises

During the course, weekly exercise sets were given to the students. Those marked with an asterisk were not to be handed in. If you would like a dvi-file with the solutions to the exercises or a copy of the final two-hour exam please e-mail me.

Problems 1. [Introduction, §1, §2]

1. Complete the verification that, in $\mathbb{Z}[\sqrt{-5}]$

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

is a factorization of (6) into a product of prime ideals (see the Introduction).

2. Let d be a square-free integer. Show that the ring of integers in $\mathbb{Q}[\sqrt{d}]$ has the description in the Introduction.

3. Let A be an integral domain. A multiplicative subset S of A is said to be *saturated* if

$$ab \in S \Rightarrow a \text{ and } b \in S.$$

(a) Show that S is saturated $\iff$ its complement is a union of prime ideals.

(b) Show that given a multiplicative system S, there is a unique smallest saturated multiplicative system S' containing S, and that $S' = A \smallsetminus \cup\mathfrak{p}$, where $\mathfrak{p}$ runs over the prime ideals disjoint from S. Show that $S'^{-1}A = S^{-1}A$. Deduce that $S^{-1}A$ is characterized by the set of prime ideals of A that remain prime in $S^{-1}A$.

4. Since $\mathbb{Z}[\sqrt{5}]$ is not integrally closed, it can not be a UFD. Give an example of an element of $\mathbb{Z}[\sqrt{5}]$ that has two distinct factorizations into irreducible elements.

5*. Let A be an integrally closed ring, and let K be its field of fractions. Let $f(X) \in A[X]$ be a monic polynomial. If $f(X)$ is reducible in $K[X]$, show that it is reducible in $A[X]$.

Problems 2. [§2]

1. Show that if L/K is not separable, then $\operatorname{disc}(L/K) = 0$.

2. Let $\mathfrak{a} = (2, 1+\sqrt{-3})$ in $\mathbb{Z}[\sqrt{-3}]$. Show that $\mathfrak{a} \neq (2)$, but $\mathfrak{a}^2 = (2)\mathfrak{a}$. Conclude that ideals in $\mathbb{Z}[\sqrt{-3}]$ do not factor uniquely into prime ideals. (Hence $\mathbb{Z}[\sqrt{-3}]$ is the wrong choice for the ring of integers in $\mathbb{Q}[\sqrt{-3}]$.)

3. Let A be a subring of a ring B, and let β be a unit in B. Show that every $\alpha \in A[\beta] \cap A[\beta^{-1}]$ is integral over A. [This has a short solution, but it's hard to find it.]

4*. Let $K = \mathbb{Q}[\sqrt{7}, \sqrt{10}]$, and let α be an algebraic integer in K. The following argument will show that $\mathcal{O}_K \neq \mathbb{Z}[\alpha]$.

(a) Consider the four algebraic integers:

$$\begin{aligned}\alpha_1 &= (1+\sqrt{7})(1+\sqrt{10});\\ \alpha_2 &= (1+\sqrt{7})(1-\sqrt{10});\\ \alpha_3 &= (1-\sqrt{7})(1+\sqrt{10});\\ \alpha_4 &= (1-\sqrt{7})(1-\sqrt{10}).\end{aligned}$$

Show that all the products $\alpha_i\alpha_j$, $i \neq j$, are divisible by 3 in $\mathcal{O}_K$, but that 3 does not divide any power of any α_i. [Hint: Show that $\alpha_i^n/3$ is not an algebraic integer by considering its trace: show that $\mathrm{Tr}(\alpha_i^n) \equiv (\sum \alpha_j^n) \equiv 4^n \pmod 3$ in $\mathbb{Z}[\alpha]$; deduce $\mathrm{Tr}(\alpha_i^n) \equiv 1 \pmod 3$ in $\mathbb{Z}$.]

(b) Assume now that $\mathcal{O}_K = \mathbb{Z}[\alpha]$ — we shall derive a contradiction. Let $f(X)$ be the minimum polynomial of α over $\mathbb{Q}$. For $g(X) \in \mathbb{Z}[X]$, let $\bar{g}(X)$ denote the image of g in $\mathbb{F}_3[X]$, $\mathbb{F}_3 = \mathbb{Z}/(3)$. Show that $g(\alpha)$ is divisible by 3 in $\mathbb{Z}[\alpha]$ if and only if $\bar{g}$ is divisible by $\bar{f}$ in $\mathbb{F}_3[X]$.

(c) For each i, $1 \leq i \leq 4$, let f_i be a polynomial in $\mathbb{Z}[X]$ such that $\alpha_i = f_i(\alpha)$. Show that $\bar{f} | \bar{f}_i\bar{f}_j$ $(i \neq j)$ in $\mathbb{F}_3[X]$, but that $\bar{f}$ does not divide $\bar{f}_i^n$ for any n. Conclude that for each i, $\bar{f}$ has an irreducible factor which does not divide $\bar{f}_i$ but does divide all $\bar{f}_j$, $j \neq i$.

(d) This shows that $\bar{f}$ has at least four distinct irreducible factors over $\mathbb{F}_3$. On the other hand, f has degree at most 4. Why is this a contradiction?

Problems 3 [§3]

1. Let k be a field. Is $k[X,Y]$ a Dedekind domain? (Explain).

2. Show that $\mathbb{Z}[\sqrt{3}]$ is the ring of integers in $\mathbb{Q}[\sqrt{3}]$ and $\mathbb{Z}[\sqrt{7}]$ is the ring of integers in $\mathbb{Q}[\sqrt{7}]$, but $\mathbb{Z}[\sqrt{3},\sqrt{7}]$ is not the ring of integers in $\mathbb{Q}[\sqrt{3},\sqrt{7}]$. (Hint: look at $(\sqrt{3}+\sqrt{7})/2$.)

3*. Let k be a field, and let A be the subring $k[X^2, X^3]$ of $k[X]$.

(a) Show that $k[X]$ is a finitely generated $k[X^2]$-module, and hence is a Noetherian $k[X^2]$-module. Deduce that A is Noetherian. [This requires facts about modules over Noetherian rings.]

(b) Show that every nonzero prime ideal of A is maximal, but that A is not a Dedekind domain.

Hence A satisfies conditions (i) and (iii) to be a Dedekind domain, but fails (ii). There are also rings that satisfy (ii) and (iii) but fail (i), and rings that satisfy (i) and (ii) but not (iii) (in fact $k[X,Y]$).

Problems 4 [§4]

1. Give an example of an integral domain B, a nonzero prime ideal $\mathfrak{p}$ in B, and a subring A of B such that $\mathfrak{p} \cap A = 0$. (Note that this can't happen if B is integral over A — see the paragraph preceding 3.32.)

2. Let $F \subset K \subset L$ be a sequence of number fields, and let $A \subset B \subset C$ be their rings of integers. If $\mathfrak{Q}|\mathfrak{P}$ and $\mathfrak{P}|\mathfrak{p}$ (prime ideals in C, B, and A respectively), show that

$$e(\mathfrak{Q}/\mathfrak{P}) \cdot e(\mathfrak{P}/\mathfrak{p}) = e(\mathfrak{Q}/\mathfrak{p}), \quad f(\mathfrak{Q}/\mathfrak{P}) \cdot f(\mathfrak{P}/\mathfrak{p}) = f(\mathfrak{Q}/\mathfrak{p}).$$

Problems 5 [§§3,4]

1. Let $K = \mathbb{Q}[\alpha]$ where α is a root of $X^3 + X + 1$ (see 2.36). According to (3.36), what are the possible ways that (p) can factor in $\mathcal{O}_K$ as a product of prime ideals. Which of these possibilities actually occur? (Illustrate by examples.)

2. Show that $\mathbb{Q}[\sqrt{-23}]$ has class number 3, and that $\mathbb{Q}[\sqrt{-47}]$ has class number 5.

3.* Let K be an algebraic number field. Prove that there is a finite extension L of K such that every ideal in $\mathcal{O}_K$ becomes principal in $\mathcal{O}_L$. [Hint: Use the finiteness of the class number.]

Problems 6 [§4]

1. Let $K = \mathbb{Q}[\alpha]$ where α is a root of $X^3 - X + 2$. Show that $\mathcal{O}_K = \mathbb{Z}[\alpha]$ and that K has class number 1. [One approach is to consider the square factors of the discriminant of $X^3 - X + 2$, and show that $\frac{1}{2}(a + b\alpha + c\alpha^2)$ is an algebraic integer if and only if a, b, and c are all even, but you may be able to find a better one.]

2. Let $K = \mathbb{Q}[\sqrt{-1}, \sqrt{5}]$. Show that $\mathcal{O}_K = \mathbb{Z}[\sqrt{-1}, \frac{1+\sqrt{5}}{2}]$. Show that the only primes (in $\mathbb{Z}$) that ramify in K are 2 and 5, and that their ramification indexes are both 2. Deduce that K is unramified over $\mathbb{Q}[\sqrt{-5}]$. Prove that $\mathbb{Q}[\sqrt{-5}]$ has class number 2, and deduce that K is the Hilbert class field of $\mathbb{Q}[\sqrt{-5}]$. (Cf. 4.9.)

Problems 7 [§5]

1. Fix an m and and M. Is it necessarily true that the set of algebraic integers α in $\mathbb{C}$ of degree $< m$ and with $|\alpha| < M$ is finite? [Either prove, or give a counterexample.]

2. Find a fundamental unit for the field $\mathbb{Q}[\sqrt{67}]$.

3. Let α be an element of a number field K. Does $\mathrm{Nm}_{K/\mathbb{Q}} = \pm 1$ imply that α is unit in $\mathcal{O}_K$. [Either prove, or give a counterexample.]

Problems 8 [§6]

1. Show that $X^3 - 3X + 1$ is an irreducible polynomial in $\mathbb{Q}[X]$ with three real roots. Let α be one of them, and let $K = \mathbb{Q}[\alpha]$. Compute $\mathrm{disc}(\mathbb{Z}[\alpha]/\mathbb{Z})$, and deduce that

$$\mathcal{O}_K \supset \mathbb{Z}[\alpha] \supset 3^m \mathcal{O}_K$$

for some m. Show that α and $\alpha + 2$ are units in $\mathbb{Z}[\alpha]$ and $\mathcal{O}_K$, and that $(\alpha + 1)^3 = 3\alpha(\alpha + 2)$. Deduce that $(\alpha + 1)$ is a prime ideal in $\mathcal{O}_K$, and show that $\mathcal{O}_K = \mathbb{Z}[\alpha] + (\alpha + 1)\mathcal{O}_K$. Use this to show that $\mathcal{O}_K = \mathbb{Z}[\alpha]$. Show that (2) is a prime ideal in $\mathcal{O}_K$, and deduce that $\mathcal{O}_K$ is a principal ideal domain.

2. Show that the ring of integers in $\mathbb{Q}[\cos\frac{2\pi}{m}]$ is $\mathbb{Z}[2\cos\frac{2\pi}{m}]$.

Problems 9 [§7]

1. Let $|\ |$ be nonarchimedean valuation on a field K.

(a) Define an open disk with radius r and centre a to be

$$D(a,r) = \{x \in K \mid |x-a| < r\}.$$

Prove that $D(a,r) = D(b,r)$ for any $b \in D(a,r)$. Deduce that if two disks meet, then the large disk contains the smaller.

(b) Assume K to be complete. Show that the series $\sum a_n$ converges if and only if $a_n \to 0$.

(This problem illustrates the weirdness of the topology defined by a nonarchimedean valuation.)

2. For which $a \in \mathbb{Z}$ is $7X^2 = a$ solvable in $\mathbb{Z}_7$? For which $a \in \mathbb{Q}$ is it solvable in $\mathbb{Q}_7$?

Problems 10 [§7]

1. (a) Show that $(X^2-2)(X^2-17)(X^2-34)$ has a root in $\mathbb{Z}_p$ for every p.

(b) Show that $5X^3 - 7X^2 + 3X + 6$ has a root α in $\mathbb{Z}_7$ with $|\alpha - 1|_7 < 1$. Find an $a \in \mathbb{Z}$ such that $|\alpha - a|_7 \leq 7^{-4}$.

2. Find all the quadratic extensions of $\mathbb{Q}_2$. Hint: there are exactly 7 (up to isomorphism).

3. Let $p_1, \dots, p_m$ be distinct prime numbers, and let $\alpha_i = \sqrt{p}$. Let $K = \mathbb{Q}[\alpha_1, \dots, \alpha_m]$. Show that $[K\colon \mathbb{Q}] = 2^m$. Let $\gamma = \sum \alpha_i$. Show that $K = \mathbb{Q}[\gamma]$, and deduce that the minimum polynomial $f(X)$ of γ over $\mathbb{Q}$ has degree 2^m. Show that $f(X)$ factors in $\mathbb{Z}_p[X]$ into a product of polynomials of degree ≤ 4 ($p \neq 2$) or of degree ≤ 8 ($p = 2$).

4*. Fix an algebraic closure $\mathbb{Q}_p^{\mathrm{al}}$ of $\mathbb{Q}_p$, and for each n prime to p, let ζ_n be a primitive n^{th} root of 1. Show that a finite extension K of $\mathbb{Q}_p$ can contain only finitely many ζ_n's. Deduce that the Cauchy sequence $\sum \zeta_n p^n$ does not converge to an element of $\mathbb{Q}_p^{\mathrm{al}}$.

Problems 11 [§§7,8]

1. (a) Find two monic polynomials of degree 3 in $\mathbb{Q}_5[X]$ with the same Newton polygon, but with one irreducible and the other not.

(b) Find a monic irreducible polynomial in $\mathbb{Z}[X]$ of degree 6 which factors in $\mathbb{Q}_5[X]$ into a product of 3 irreducible polynomials of degree 2.

2. Let $K = \mathbb{Q}[\alpha]$ where α is a root of $X^3 - X^2 - 2X - 8$. Show that there are three extensions of the 2-adic valuation to K. Deduce that $2 | \operatorname{disc}(\mathbb{Z}[\alpha]/\mathbb{Z})$ but not $\operatorname{disc}(\mathcal{O}_K/\mathbb{Z})$.

3*. Let L be a finite Galois extension of the local field K, and let G_i, $i \geq 0$, be the i^{th} ramification group. Let Π generate the maximal ideal in $\mathcal{O}_L$. For $\sigma \in G_i$, write $\sigma\Pi = \Pi + a(\sigma)\Pi^{i+1}$, and consider the map $G_i \to l$, $\sigma \mapsto a(\sigma) \mod (\Pi)$, where

$l = \mathcal{O}_L/(\Pi)$. Show that this is a homomorphism (additive structure on l) if and only if $i > 0$.

4*. "It is a thought-provoking question that few graduate students would know how to approach the question of determining the Galois group of, say,

$$X^6 + 2X^5 + 3X^4 + 4X^3 + 5X^2 + 6X + 7."$$

(a) Can you find it?

(b) Can you find it without using a computer?

5*. Let $K = k(X)$ where k is a finite field. Assume that every valuation of K comes from a prime ideal of $k[X]$ or $k[X^{-1}]$, and prove the product formula.

I recommend also Problems 5–38 of Chapter IV of Marcus 1977, which guide the reader through a proof of the Kronecker-Weber Theorem (every abelian extension of the rationals is contained in a cyclotomic extension) which is probably close to the Hilbert's original proof.

Also, Arakelov theory suggests a different way of viewing the classical results in algebraic number theory. Sometime I intend to write a sequence of problems illustrating this. For the moment, I can only refer the reader to Chapter III of Neukirch, J., Algebraische Zahlentheorie, Springer 1992.