Chapter 7

Introducing Algebraic Number Theory

(Commutative Algebra 1)

The general theory of commutative rings is known as *commutative algebra*. The main applications of this discipline are to algebraic number theory, to be discussed in this chapter, and algebraic geometry, to be introduced in Chapter 8.

Techniques of abstract algebra have been applied to problems in number theory for a long time, notably in the effort to prove Fermat's Last Theorem. As an introductory example, we will sketch a problem for which an algebraic approach works very well. If p is an odd prime and $p \equiv 1 \bmod 4$, we will prove that p is the sum of two squares, that is, p can be expressed as $x^2 + y^2$ where x and y are integers. Since $\frac{p-1}{2}$ is even, it follows that -1 is a quadratic residue (that is, a square) mod p. [Pair each of the numbers 2,3, $\ldots$,$p-2$ with its inverse mod p and pair 1 with $p-1 \equiv -1 \bmod p$. The product of the numbers 1 through $p-1$ is, mod p,

$$1 \times 2 \times \cdots \times \frac{p-1}{2} \times -1 \times -2 \times \cdots \times -\frac{p-1}{2}$$

and therefore $\left[\left(\frac{p-1}{2}\right)!\right]^2 \equiv -1 \bmod p$.]

If $-1 \equiv x^2 \bmod p$, then p divides $x^2 + 1$. Now we enter the ring of Gaussian integers and factor $x^2 + 1$ as $(x + i)(x - i)$. Since p can divide neither factor, it follows that p is not prime in $\mathbb{Z}[i]$, so we can write $p = \alpha\beta$ where neither α nor β is a unit.

Define the *norm* of $\gamma = a + bi$ as $N(\gamma) = a^2 + b^2$. Then $N(\gamma) = 1$ iff $\gamma = \pm 1$ or $\pm i$ iff γ is a unit. (See Section 2.1, Problem 5.) Thus

$$p^2 = N(p) = N(\alpha)N(\beta) \text{ with } N(\alpha) > 1 \text{ and } N(\beta) > 1,$$

so $N(\alpha) = N(\beta) = p$. If $\alpha = x + iy$, then $p = x^2 + y^2$.

Conversely, if p is an odd prime and $p = x^2 + y^2$, then p is congruent to 1 mod 4. (If x is even, then $x^2 \equiv 0$ mod 4, and if x is odd, then $x^2 \equiv 1$ mod 4. We cannot have x and y both even or both odd, since p is odd.)

It is natural to conjecture that we can identify those primes that can be represented as $x^2 + |d|y^2$, where d is a negative integer, by working in the ring $\mathbb{Z}[\sqrt{d}]$. But the Gaussian integers ($d = -1$) form a Euclidean domain, in particular a unique factorization domain. On the other hand, unique factorization fails for $d \leq -3$ (Section 2.7, Problem 7), so the above argument collapses. [Recall from (2.6.4) that in a UFD, an element p that is not prime must be reducible.] Difficulties of this sort led Kummer to invent "ideal numbers", which later became ideals at the hands of Dedekind. We will see that although a ring of algebraic integers need not be a UFD, unique factorization of ideals will always hold.

7.1 Integral Extensions

If E/F is a field extension and $\alpha \in E$, then α is algebraic over F iff α is a root of a polynomial with coefficients in F. We can assume if we like that the polynomial is monic, and this turns out to be crucial in generalizing the idea to ring extensions.

7.1.1 Definitions and Comments

In this chapter, unless otherwise specified, *all rings are assumed commutative*. Let A be a subring of the ring R, and let $x \in R$. We say that x is *integral over* A if x is a root of a monic polynomial f with coefficients in A. The equation $f(X) = 0$ is called an *equation of integral dependence* for x over A. If x is a real or complex number that is integral over $\mathbb{Z}$, then x is called an *algebraic integer*. Thus for every integer d, $\sqrt{d}$ is an algebraic integer, as is any n^{th} root of unity. (The monic polynomials are, respectively, $X^2 - d$ and $X^n - 1$.) In preparation for the next result on conditions equivalent to integrality, note that $A[x]$, the set of polynomials in x with coefficients in A, is an A-module. (The sum of two polynomials is a polynomial, and multiplying a polynomial by a member of A produces another polynomial over A.)

7.1.2 Proposition

Let A be a subring of R, with $x \in R$. The following conditions are equivalent:

(i) x is integral over A;

(ii) The A-module $A[x]$ is finitely generated;

(iii) x belongs to a subring B of R such that $A \subseteq B$ and B is a finitely generated A-module.

Proof. (i) implies (ii). If x is a root of a monic polynomial over A of degree n, then x^n and all higher powers of x can be expressed as linear combinations of lower powers of x. Thus $1, x, x^2, , \ldots, x^{n-1}$ generate $A[x]$ over A.

(ii) implies (iii). Take $B = A[x]$.

(iii) implies (i). If $\beta_1, \ldots, \beta_n$ generate B over A, then $x\beta_i$ is a linear combination of the β_j, say $x\beta_i = \sum_{j=1}^{n} c_{ij}\beta_j$. Thus if β is a column vector whose components are the β_i, I is an n by n identity matrix, and $C = [c_{ij}]$, then

$$(xI - C)\beta = 0,$$

and if we premultiply by the adjoint matrix of $xI - C$ (as in Cramer's rule), we get

$$[\det(xI - C)]I\beta = 0,$$

hence $[\det(xI - C)]b = 0$ for every $b \in B$. Since B is a ring we may set $b = 1$ and conclude that x is a root of the monic polynomial $\det(XI - C)$ in $A[X]$. ♣

For other equivalent conditions, see Problems 1 and 2.

We are going to prove a transitivity property for integral extensions (analogous to (3.3.5)), and the following result will be helpful.

7.1.3 Lemma

Let A be a subring of R, with $x_1, \ldots, x_n \in R$. If x_1 is integral over A, x_2 is integral over $A[x_1]$, ... , and x_n is integral over $A[x_1, \ldots, x_{n-1}]$, then $A[x_1, \ldots, x_n]$ is a finitely generated A-module.

Proof. The $n = 1$ case follows from (7.1.2), part (ii). Going from $n - 1$ to n amounts to proving that if A, B and C are rings, with C a finitely generated B-module and B a finitely generated A-module, then C is a finitely generated A-module. This follows by a brief computation:

$$C = \sum_{j=1}^{r} By_j, \quad B = \sum_{k=1}^{s} Az_k \quad \text{so} \quad C = \sum_{j=1}^{r}\sum_{k=1}^{s} Ay_j z_k. \quad ♣$$

7.1.4 Transitivity of Integral Extensions

Let A, B and C be subrings of R. If C is integral over B, that is, each element of C is integral over B, and B is integral over A, then C is integral over A.

Proof. Let $x \in C$, with $x^n + b_{n-1}x^{n-1} + \cdots + b_1 x + b_0 = 0$, $b_i \in B$. Then x is integral over $A[b_0, \ldots, b_{n-1}]$. Each b_i is integral over A, hence over $A[b_0, \ldots, b_{i-1}]$. By (7.1.3), $A[b_0, \ldots, b_{n-1}, x]$ is a finitely generated A-module. By (7.1.2), part (iii), x is integral over A. ♣

7.1.5 Definitions and Comments

If A is a subring of R, the *integral closure of A in R* is the set A_c of elements of R that are integral over A. Note that $A \subseteq A_c$ because each $a \in A$ is a root of $X - a$. We say that A is *integrally closed* in R if $A_c = A$. If we simply say that A is *integrally closed* without reference to R, we assume that A is an integral domain with quotient field K, and A is integrally closed in K.

If x and y are integral over A, then just as in the proof of (7.1.4), it follows from (7.1.3) that $A[x,y]$ is a finitely generated A-module. Since $x+y, x-y$ and xy belong to this module, they are integral over A by (7.1.2) part (iii). The important conclusion is that

$$A_c \text{ is a subring of } R \text{ containing } A.$$

If we take the integral closure of the integral closure, we get nothing new.

7.1.6 Proposition

The integral closure A_c of A in R is integrally closed in R.

Proof. By definition, A_c is integral over A. If x is integral over A_c, then as in the proof of (7.1.4), x is integral over A, so that $x \in A_c$. ♣

We can identify a large class of integrally closed rings.

7.1.7 Proposition

If A is a UFD, then A is integrally closed.

Proof. If x belongs to the quotient field K, then we can write $x = a/b$ where $a, b \in A$, with a and b relatively prime. If x is integral over A, then there is an equation of the form

$$(a/b)^n + a_{n-1}(a/b)^{n-1} + \cdots + a_1(a/b) + a_0 = 0$$

with $a_i \in A$. Multiplying by b^n, we have $a^n + bc = 0$, with $c \in A$. Thus b divides a^n, which cannot happen for relatively prime a and b unless b has no prime factors at all, in other words, b is a unit. But then $x = ab^{-1} \in A$. ♣

We can now discuss one of the standard setups for doing algebraic number theory.

7.1.8 Definitions and Comments

A *number field* is a subfield L of the complex numbers $\mathbb{C}$ such that L is a finite extension of the rationals $\mathbb{Q}$. Thus the elements of L are algebraic numbers. The integral closure of $\mathbb{Z}$ in L is called the ring of *algebraic integers* (or simply *integers*) of L. In the next section, we will find the algebraic integers explicitly when L is a quadratic extension.

Problems For Section 7.1

1. Show that in (7.1.2) another equivalent condition is the following:

   (iv) There is a subring B of R such that B is a finitely generated A-module and $xB \subseteq B$.

   If R is a field, show that the assumption that B is a subring can be dropped (as long as $B \neq 0$).

2. A module is said to be *faithful* if its annihilator is 0. Show that in (7.1.2) the following is another equivalent condition:
   (v) There is a faithful $A[x]$-module B that is finitely generated as an A-module.

Let A be a subring of the integral domain B, with B integral over A. In Problems 3–5 we are going to show that A is a field if and only if B is a field.

3. Assume that B is a field, and let a be a nonzero element of A. Then since $a^{-1} \in B$, there is an equation of the form
$$(a^{-1})^n + c_{n-1}(a^{-1})^{n-1} + \cdots + c_1 a^{-1} + c_0 = 0$$
with $c_i \in A$. Show that $a^{-1} \in A$, proving that A is a field.
4. Now assume that A is a field, and let b be a nonzero element of B. By (7.1.2) part (ii), $A[b]$ is a finite-dimensional vector space over A. Let f be the A-linear transformation on this vector space given by multiplication by b, in other words, $f(z) = bz, z \in A[b]$. Show that f is injective.
5. Show that f is surjective as well, and conclude that B is a field.

In Problems 6–8, let A be a subring of B, with B integral over A. Let Q be a prime ideal of B and let $P = Q \cap A$.

6. Show that P is a prime ideal of A, and that A/P can be regarded as a subring of B/Q.
7. Show that B/Q is integral over A/P.
8. Show that P is a maximal ideal of A if and only if Q is a maximal ideal of B.

7.2 Quadratic Extensions of the Rationals

We will determine the algebraic integers of $L = \mathbb{Q}(\sqrt{d})$, where d is a square-free integer (a product of distinct primes). The restriction on d involves no loss of generality; for example, $\mathbb{Q}(\sqrt{12}) = \mathbb{Q}(\sqrt{3})$. The minimal polynomial of $\sqrt{d}$ over $\mathbb{Q}$ is $X^2 - d$, which has roots $\pm\sqrt{d}$. The extension $L/\mathbb{Q}$ is Galois, and the Galois group consists of the identity and the automorphism $\sigma(a + b\sqrt{d}) = a - b\sqrt{d},\ a, b \in \mathbb{Q}$.

A remark on notation: To make sure that there is no confusion between algebraic integers and ordinary integers, we will use the term *rational integer* for a member of $\mathbb{Z}$.

7.2.1 Lemma

If a and b are rational numbers, then $a + b\sqrt{d}$ is an algebraic integer if and only if $2a$ and $a^2 - db^2$ belong to $\mathbb{Z}$. In this case, $2b$ is also in $\mathbb{Z}$.

Proof. Let $x = a + b\sqrt{d}$, so that $\sigma(x) = a - b\sqrt{d}$. Then $x + \sigma(x) = 2a \in \mathbb{Q}$ and $x\sigma(x) = a^2 - db^2 \in \mathbb{Q}$. Now if x is an algebraic integer, then x is a root of a monic polynomial $f \in \mathbb{Z}[X]$. But $f(\sigma(x)) = \sigma(f(x))$ since σ is an automorphism, so $\sigma(x)$ is also a root of f and hence an algebraic integer. By (7.1.5), $2a$ and $a^2 - db^2$ are also algebraic integers, as well as rational numbers. By (7.1.7), $\mathbb{Z}$ is integrally closed, so $2a$

and $a^2 - db^2$ belong to $\mathbb{Z}$. The converse holds because $a + b\sqrt{d}$ is a root of $(X - a)^2 = db^2$, i.e., $X^2 - 2aX + a^2 - db^2 = 0$.

Now if $2a$ and $a^2 - db^2$ are rational integers, then $(2a)^2 - d(2b)^2 = 4(a^2 - db^2) \in \mathbb{Z}$, so $d(2b)^2 \in \mathbb{Z}$. If $2b \notin \mathbb{Z}$, then its denominator would include a prime factor p , which would appear as p^2 in the denominator of $(2b)^2$. Multiplication of $(2b)^2$ by d cannot cancel the p^2 because d is square-free, and the result follows. ♣

7.2.2 Corollary

The set B of algebraic integers of $\mathbb{Q}(\sqrt{d})$, d square-free, can be described as follows.

(i) If $d \not\equiv 1 \mod 4$, then B consists of all $a + b\sqrt{d}$, $a, b \in \mathbb{Z}$;

(ii) If $d \equiv 1 \mod 4$, then B consists of all $\frac{u}{2} + \frac{v}{2}\sqrt{d}$, $u, v \in Z$, where u and v have the same parity (both even or both odd).

[Note that since d is square-free, it is not divisible by 4, so the condition in (i) is $d \equiv 2$ or 3 mod 4.]

Proof. By (7.2.1), the algebraic integers are of the form $\frac{u}{2} + \frac{v}{2}\sqrt{d}$ where $u, v \in \mathbb{Z}$ and $\frac{u^2}{4} - \frac{dv^2}{4} \in \mathbb{Z}$, i.e., $u^2 - dv^2 \equiv 0 \mod 4$. It follows that u and v have the same parity. (The square of an even number is congruent to 0 and the square of an odd number to 1 mod 4.) Moreover, the "both odd" case can only occur when $d \equiv 1 \mod 4$. The "both even" case is equivalent to $\frac{u}{2}, \frac{v}{2} \in \mathbb{Z}$, and the result follows. ♣

We can express these results in a more convenient form. We will show in (7.4.10) that the set B of algebraic integers in any number field L is a free $\mathbb{Z}$-module of rank $n = [L : \mathbb{Q}]$. A basis for this module is called an *integral basis* or $\mathbb{Z}$-*basis* for B.

7.2.3 Theorem

Let B be the algebraic integers of $\mathbb{Q}(\sqrt{d})$, d square-free.

(i) If $d \not\equiv 1 \mod 4$, then 1 and $\sqrt{d}$ form an integral basis of B;

(ii) If $d \equiv 1 \mod 4$, then 1 and $\frac{1}{2}(1 + \sqrt{d})$ form an integral basis.

Proof. (i) By (7.2.2), 1 and $\sqrt{d}$ span B over $\mathbb{Z}$, and they are linearly independent because $\sqrt{d}$ is irrational.

(ii) By (7.2.2), 1 and $\frac{1}{2}(1 + \sqrt{d})$ are algebraic integers. To show that they span B, consider $\frac{1}{2}(u + v\sqrt{d})$, where u and v have the same parity. Then

$$\frac{1}{2}(u + v\sqrt{d}) = \left(\frac{u - v}{2}\right)(1) + v\left[\frac{1}{2}(1 + \sqrt{d})\right]$$

with $\frac{u-v}{2}$ and v in $\mathbb{Z}$. Finally, to show linear independence, assume that $a, b \in \mathbb{Z}$ and

$$a + b\left[\frac{1}{2}(1 + \sqrt{d})\right] = 0.$$

Then $2a + b + b\sqrt{d} = 0$, which forces $a = b = 0$. ♣

Problems For Section 7.2

1. Let $L = \mathbb{Q}(\alpha)$, where α is a root of the irreducible quadratic $X^2 + bX + c$, with $b, c \in \mathbb{Q}$. Show that $L = \mathbb{Q}(\sqrt{d})$ for some square-free integer d. Thus the analysis of this section covers all possible quadratic extensions of $\mathbb{Q}$.
2. Show that the quadratic extensions $\mathbb{Q}(\sqrt{d})$, d square-free, are all distinct.
3. Continuing Problem 2, show that in fact no two distinct quadratic extensions of $\mathbb{Q}$ are $\mathbb{Q}$-isomorphic.

Cyclotomic fields do not exhibit the same behavior. Let $\omega_n = e^{i2\pi/n}$, a primitive n^{th} root of unity. By a direct computation, we have $\omega_{2n}^2 = \omega_n$, and

$$-\omega_{2n}^{n+1} = -e^{i\pi(n+1)/n} = e^{i\pi}e^{i\pi}e^{i\pi/n} = \omega_{2n}.$$

4. Show that if n is odd, then $\mathbb{Q}(\omega_n) = \mathbb{Q}(\omega_{2n})$.
5. If x is an algebraic integer, show that the minimal polynomial of x over $\mathbb{Q}$ has coefficients in $\mathbb{Z}$. (This will be a consequence of the general theory to be developed in this chapter, but it is accessible now without heavy machinery.) Consequently, an algebraic integer that belongs to $\mathbb{Q}$ in fact belongs to $\mathbb{Z}$. (The minimal polynomial of $r \in \mathbb{Q}$ over $\mathbb{Q}$ is $X - r$.)
6. Give an example of a quadratic extension of $\mathbb{Q}$ that is also a cyclotomic extension.
7. Show that an integral basis for the ring of algebraic integers of a number field L is, in particular, a basis for L over $\mathbb{Q}$.

7.3 Norms and Traces

7.3.1 Definitions and Comments

If E/F is a field extension of finite degree n, then in particular, E is an n-dimensional vector space over F, and the machinery of basic linear algebra becomes available. If x is any element of E, we can study the F-linear transformation $m(x)$ given by multiplication by x, that is, $m(x)y = xy$. We define the *norm* and the *trace* of x, relative to the extension E/F, as

$$N[E/F](x) = \det m(x) \text{ and } T[E/F](x) = \text{ trace } m(x).$$

We will write $N(x)$ and $T(x)$ if E/F is understood. If the matrix $A(x) = [a_{ij}(x)]$ represents $m(x)$ with respect to some basis for E over F, then the norm of x is the determinant of $A(x)$ and the trace of x is the trace of $A(x)$, that is, the sum of the main diagonal entries. The *characteristic polynomial* of x is defined as the characteristic polynomial of the matrix $A(x)$, that is,

$$\text{char}[E/F](x) = \det[XI - A(x)]$$

where I is an n by n identity matrix. If E/F is understood, we will refer to the *characteristic polynomial of* x, written $\text{char}(x)$.

7.3.2 Example

Let $E = \mathbb{C}$ and $F = \mathbb{R}$. A basis for $\mathbb{C}$ over $\mathbb{R}$ is $\{1, i\}$ and, with $x = a + bi$, we have

$$(a + bi)(1) = a(1) + b(i) \text{ and } (a + bi)(i) = -b(1) + a(i).$$

Thus

$$A(a + bi) = \begin{bmatrix} a & -b \\ b & a \end{bmatrix}.$$

The norm, trace and characteristic polynomial of $a + bi$ are

$$N(a + bi) = a^2 + b^2,\ T(a + bi) = 2a,\ \text{char}(a + bi) = X^2 - 2aX + a^2 + b^2.$$

The computation is exactly the same if $E = \mathbb{Q}(i)$ and $F = \mathbb{Q}$. Notice that the coefficient of X is minus the trace and the constant term is the norm. In general, it follows from the definition of characteristic polynomial that

$$\text{char}(x) = X^n - T(x)X^{n-1} + \cdots + (-1)^n N(x).$$

[The only terms multiplying X^{n-1} in the expansion of the determinant are $-a_{ii}(x)$, $i = 1, \ldots, n$. Set $X = 0$ to show that the constant term of $\text{char}(x)$ is $(-1)^n$ det $A(x)$.]

7.3.3 Lemma

If E is an extension of F and $x \in E$, then $N(x)$, $T(x)$ and the coefficients of $\text{char}(x)$ belong to F. If $a \in F$, then

$$N(a) = a^n,\ T(a) = na,\ \text{and}\ \text{char}(a) = (X - a)^n.$$

Proof. The first assertion follows because the entries of the matrix $A(x)$ are in F. The second statement holds because if $a \in F$, the matrix representing multiplication by a is aI. ♣

It is natural to look for a connection between the characteristic polynomial of x and the minimal polynomial of x over F.

7.3.4 Proposition

$$\text{char}[E/F](x) = \ [\min(x, F)]^r$$

where $r = [E : F(x)]$.

Proof. First assume that $r = 1$, so that $E = F(x)$. By the Cayley-Hamilton theorem, the linear transformation $m(x)$ satisfies $\text{char}(x)$, and since $m(x)$ is multiplication by x, x itself is a root of $\text{char}(x)$. Thus $\min(x, F)$ divides $\text{char}(x)$. But both polynomials have degree n, and the result follows. In the general case, let $y_1, \ldots, y_s$ be a basis for $F(x)$ over F, and let $z_1, \ldots, z_r$ be a basis for E over $F(x)$. Then the $y_i z_j$ form a basis

for E over F. Let $A = A(x)$ be the matrix representing multiplication by x in the extension $F(x)/F$, so that $xy_i = \sum_k a_{ki}y_k$, and $x(y_iz_j) = \sum_k a_{ki}(y_kz_j)$. Order the basis for E/F as $y_1z_1, y_2z_1, \ldots, y_sz_1;\ y_1z_2, y_2z_2, \ldots, y_sz_2; \ldots;\ y_1z_r, y_2z_r, \ldots, y_sz_r$. Then $m(x)$ is represented in E/F as

$$\begin{bmatrix} A & 0 & \ldots & 0 \\ 0 & A & \ldots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \ldots & A \end{bmatrix}$$

Thus $\mathrm{char}[E/F](x) = [\det(XI - A)]^r$, which by the $r = 1$ case coincides with $[\min(x, F)]^r$. ♣

7.3.5 Corollary

Let $[E : F] = n$, and $[F(x) : F] = d$. Let $x_1, \ldots, x_d$ be the roots of $\min(x, F)$ in a splitting field (counting multiplicity). Then

$$N(x) = \left(\prod_{i=1}^{d} x_i\right)^{n/d}, \quad T(x) = \frac{n}{d}\sum_{i=1}^{d} x_i$$

and

$$\mathrm{char}(x) = \left[\prod_{i=1}^{d}(X - x_i)\right]^{n/d}.$$

Proof. The formula for the characteristic polynomial follows from (7.3.4). The norm is $(-1)^n$ times the constant term of $\mathrm{char}(x)$ (see (7.3.2)), hence is

$$(-1)^n(-1)^n\left(\prod_{i=1}^{d} x_i\right)^{n/d}.$$

Finally, if $\min(x, F) = X^d + a_{d-1}X^{d-1} + \cdots + a_1X + a_0$, then the coefficient of X^{n-1} in $[\min(x, F)]^{n/d}$ is $\frac{n}{d}a_{d-1} = -\frac{n}{d}\sum_{i=1}^{d} x_i$. Since the trace is the negative of this coefficient [see (7.3.2)], the result follows. ♣

If E is a separable extension of F, there are very useful alternative expressions for the trace and norm.

7.3.6 Proposition

Let E/F be a separable extension of degree n, and let $\sigma_1, \ldots, \sigma_n$ be the distinct F-monomorphisms of E into an algebraic closure of E, or equally well into a normal extension L of F containing E. Then

$$T[E/F](x) = \sum_{i=1}^{n}\sigma_i(x) \text{ and } N[E/F](x) = \prod_{i=1}^{n}\sigma_i(x).$$

Consequently, $T(ax+by) = aT(x)+bT(y)$ and $N(xy) = N(x)N(y)$ for $x, y \in E, a, b \in F$.

Proof. Each of the d distinct F-embeddings τ_i of $F(x)$ into L takes x into a unique conjugate x_i, and extends to exactly $\frac{n}{d} = [E : F(x)]$ F-embeddings of E into L, all of which also take x to x_i[see (3.5.1), (3.2.3) and (3.5.2)]. Thus

$$\sum_{i=1}^{n} \sigma_i(x) = \frac{n}{d} \sum_{i=1}^{d} \tau_i(x) = T(x)$$

and

$$\prod_{i=1}^{n} \sigma_i(x) = \left[\prod_{i=1}^{d} \tau_i(x) \right]^{n/d} = N(x)$$

by (7.3.5). ♣

The linearity of T and the multiplicativity of N hold without any assumption of separability, since in (7.3.1) we have $m(ax + by) = am(x) + bm(y)$ and $m(xy) = m(x) \circ m(y)$.

7.3.7 Corollary (Transitivity of Trace and Norm)

If $F \leq K \leq E$, where E/F is finite and separable, then

$$T[E/F] = T[K/F] \circ T[E/K] \text{ and } N[E/F] = N[K/F] \circ N[E/K].$$

Proof. Let $\sigma_1, \ldots, \sigma_n$ be the distinct F-embeddings of K into L, and let $\tau_1, \ldots, \tau_m$ be the distinct K-embeddings of E into L, where L is the normal closure of E over F. By (6.3.1) and (3.5.11), L/F is Galois, and by (3.5.2), (3.5.5) and (3.5.6), each mapping σ_i and τ_j extends to an automorphism of L. Therefore it makes sense to allow the mappings to be composed. By (7.3.6),

$$T[K/F](T[E/K])(x) = \sum_{i=1}^{n} \sigma_i \Big(\sum_{j=1}^{m} \tau_j(x) \Big) = \sum_{i=1}^{n} \sum_{j=1}^{m} \sigma_i \tau_j(x).$$

Now each $\sigma_i\tau_j$ is an F-embedding of E into L, and the number of mappings is $mn = [E : K][K : F] = [E : F]$. Furthermore, the $\sigma_i\tau_j$ are distinct when restricted to E. For if $\sigma_i\tau_j = \sigma_k\tau_l$ on E, hence on K, then $\sigma_i = \sigma_k$ on K (because $\tau_j = \tau_l =$ the identity on K). Thus $i = k$, so that $\tau_j = \tau_l$ on E. But then $j = l$. By (7.3.6), $T[K/F](T[E/K])(x) = T[E/F](x)$. The norm is handled the same way, with sums replaced by products. ♣

7.3.8 Corollary

If E/F is a finite separable extension, then $T[E/F](x)$ cannot be 0 for all $x \in E$.

Proof. If $T(x) = 0$ for all x, then by (7.3.6), $\sum_{i=1}^{n} \sigma_i(x) = 0$ for all x. This contradicts Dedekind's lemma (6.1.6). ♣

A statement equivalent to (7.3.8) is that if E/F is finite and separable, then the "trace form" [the bilinear form $(x,y) \to T[E/F](xy)$] is nondegenerate, i.e., if $T(xy) = 0$ for all y, then $x = 0$. For if $x \neq 0, T(x_0) \neq 0$, and $T(xy) = 0$ for all y, choose y so that $xy = x_0$ to reach a contradiction.

7.3.9 The Basic Setup For Algebraic Number Theory

Let A be an integral domain with quotient field K, and let L be a finite separable extension of K. Let B be the set of elements of L that are integral over A, that is, B is the integral closure of A in L. The diagram below summarizes all the information.

$$\begin{array}{ccc} L & \text{—} & B \\ | & & | \\ K & \text{—} & A \end{array}$$

In the most important special case, $A = \mathbb{Z}$, $K = \mathbb{Q}$, L is a number field, and B is the ring of algebraic integers of L. From now on, we will refer to (7.3.9) as the *AKLB setup*.

7.3.10 Proposition

If $x \in B$, then the coefficients of $\mathrm{char}[L/K](x)$ and $\min(x,K)$ are integral over A. In particular, $T[L/K](x)$ and $N[L/K](x)$ are integral over A, by (7.3.2). If A is integrally closed, then by (7.3.3), the coefficients belong to A.

Proof. The coefficients of $\min(x,K)$ are sums of products of the roots x_i, so by (7.1.5) and (7.3.4), it suffices to show that the x_i are integral over A. Each x_i is a conjugate of x over K, so by (3.2.3) there is a K-isomorphism $\tau_i : K(x) \to K(x_i)$ such that $\tau_i(x) = x_i$. If we apply τ_i to an equation of integral dependence for x over A, we get an equation of integral dependence for x_i over A. ♣

Problems For Section 7.3

1. If $E = \mathbb{Q}(\sqrt{d})$ and $x = a + b\sqrt{d} \in E$, find the norm and trace of x.
2. If $E = \mathbb{Q}(\theta)$ where θ is a root of the irreducible cubic $X^3 - 3X + 1$, find the norm and trace of θ^2.
3. Find the trace of the primitive 6^{th} root of unity ω in the cyclotomic extension $\mathbb{Q}_6$.

We will now prove *Hilbert's Theorem 90*: If E/F is a cyclic extension with $[E:F] = n$ and Galois group $G = \{1, \sigma, \ldots, \sigma^{n-1}\}$ generated by σ, and $x \in E$, then

(i) $N(x) = 1$ if and only if there exists $y \in E$ such that $x = y/\sigma(y)$;

(ii) $T(x) = 0$ if and only if there exists $z \in E$ such that $x = z - \sigma(z)$.

4. Prove the "if" parts of (i) and (ii) by direct computation.

By Dedekind's lemma, $1, \sigma, \sigma^2, \ldots, \sigma^{n-1}$ are linearly independent over E, so

$$1 + x\sigma + x\sigma(x)\sigma^2 + \cdots + x\sigma(x)\cdots\sigma^{n-2}(x)\sigma^{n-1}$$

is not identically 0 on E.

5. Use this to prove the "only if" part of (i).

   By (7.3.8), there is an element $u \in E$ whose trace is not 0. Let

$$w = x\sigma(u) + (x + \sigma(x))\sigma^2(u) + \cdots + (x + \sigma(x) + \cdots + \sigma^{n-2}(x))\sigma^{n-1}(u)$$

hence

$$\sigma(w) = \sigma(x)\sigma^2(u) + (\sigma(x) + \sigma^2(x))\sigma^3(u) + \cdots + (\sigma(x) + \sigma^2(x) + \cdots + \sigma^{n-1}(x))\sigma^n(u)$$

6. If $T(x) = 0$, show that $w - \sigma(w) = xT(u)$.
7. If $z = w/T(u)$, show that $z - \sigma(z) = x$, proving the "only if" part of (ii).
8. In Hilbert's Theorem 90, are the elements y and z unique?
9. Let θ be a root of $X^4 - 2$ over $\mathbb{Q}$. Find the trace over $\mathbb{Q}$ of θ, θ^2, θ^3 and $\sqrt{3}\theta$.
10. Continuing Problem 9, show that $\sqrt{3}$ cannot belong to $\mathbb{Q}[\theta]$.

7.4 The Discriminant

We have met the discriminant of a polynomial in connection with Galois theory (Section 6.6). There is also a discriminant in algebraic number theory. The two concepts are unrelated at first glance, but there is a connection between them. We assume the basic $AKLB$ setup of (7.3.9), with $n = [L : K]$.

7.4.1 Definition

The *discriminant* of the n-tuple $x = (x_1, \ldots, x_n)$ of elements of L is

$$D(x) = \det(T[L/K](x_i x_j)).$$

Thus we form a matrix whose ij element is the trace of $x_i x_j$, and take the determinant of the matrix. By (7.3.3) and (7.3.10), $D(x)$ belongs to K and is integral over A, hence belongs to A if A is integrally closed.

The discriminant behaves quite reasonably under linear transformation:

7.4.2 Lemma

If $y = Cx$, where C is an n by n matrix over K and x and y are n-tuples written as column vectors, then $D(y) = (\det C)^2 D(x)$.

Proof. The trace of $y_r y_s$ is

$$T\Big(\sum_{i,j} c_{ri} c_{sj} x_i x_j\Big) = \sum_{i,j} c_{ri} T(x_i x_j) c_{sj}$$

hence

$$(T(y_r y_s)) = C(T(x_i x_j))C'$$

where C' is the transpose of C. The result follows upon taking determinants. ♣

Here is an alternative expression for the discriminant.

7.4.3 Lemma

Let $\sigma_1, \ldots, \sigma_n$ be the K-embeddings of L into an algebraic closure of L, as in (7.3.6). Then $D(x) = [\det(\sigma_i(x_j))]^2$.

Thus we form the matrix whose ij element is $\sigma_i(x_j)$, take the determinant and square the result.

Proof. By (7.3.6),

$$T(x_i x_j) = \sum_k \sigma_k(x_i x_j) = \sum_k \sigma_k(x_i)\sigma_k(x_j)$$

so if C is the matrix whose ij entry is $\sigma_i(x_j)$, then

$$(T(x_i x_j)) = C'C$$

and again the result follows upon taking determinants. ♣

The discriminant "discriminates" between bases and non-bases, as follows.

7.4.4 Proposition

If $x = (x_1, \ldots, x_n)$, then the x_i form a basis for L over K if and only if $D(x) \neq 0$.

Proof. If $\sum_j c_j x_j = 0$, with the $c_j \in K$ and not all 0, then $\sum_j c_j \sigma_i(x_j) = 0$ for all i, so the columns of the matrix $B = (\sigma_i(x_j))$ are linearly dependent. Thus linear dependence of the x_i implies that $D = 0$. Conversely, assume that the x_i are linearly independent (and therefore a basis since $n = [L : K]$). If $D = 0$, then the rows of B are linearly dependent, so for some $c_i \in K$, not all 0, we have $\sum_i c_i \sigma_i(x_j) = 0$ for all j. Since the x_j form a basis, we have $\sum_i c_i \sigma_i(u) = 0$ for all $u \in L$, so the monomorphisms σ_i are linearly dependent. This contradicts Dedekind's lemma. ♣

We now make the connection between the discriminant defined above and the discriminant of a polynomial defined previously.

7.4.5 Proposition

Assume that $L = K(x)$, and let f be the minimal polynomial of x over K. Let D be the discriminant of the basis $1, x, x^2, \ldots, x^{n-1}$ for L over K. Then D is the discriminant of the polynomial f.

Proof. Let $x_1, \ldots, x_n$ be the roots of f in a splitting field, with $x_1 = x$. Let σ_i be the K-embedding that takes x to x_i, $i = 1, \ldots, n$. Then $\sigma_i(x^j) = x_i^j$, so by (7.4.3), D is the

square of the determinant of the matrix

$$\begin{bmatrix} 1 & x_1 & x_1^2 & \cdots & x_1^{n-1} \\ 1 & x_2 & x_2^2 & \cdots & x_2^{n-1} \\ \vdots & \vdots & & \vdots & \\ 1 & x_n & x_n^2 & \dots & x_n^{n-1} \end{bmatrix}$$

and the result follows from the formula for a Vandermonde determinant; see (A6.2). ♣

7.4.6 Corollary

Under the hypothesis of (7.4.5),

$$D = (-1)^{\binom{n}{2}} N[L/K](f'(x))$$

where f' is the derivative of f.

Proof. Let $a = (-1)^{\binom{n}{2}}$. By (7.4.5),

$$D = \prod_{i<j}(x_i - x_j)^2 = a\prod_{i\neq j}(x_i - x_j) = a\prod_i\prod_{j\neq i}(x_i - x_j).$$

But $f(X) = (X - x_1)\cdots(X - x_n)$, so

$$f'(x_i) = \sum_k \prod_{j\neq k}(X - x_j)$$

with X replaced by x_i. When X is replaced by x_i, only the $k = i$ term is nonzero, hence

$$f'(x_i) = \prod_{j\neq i}(x_i - x_j).$$

Consequently,

$$D = a\prod_{i=1}^{n} f'(x_i).$$

But

$$f'(x_i) = f'(\sigma_i(x)) = \sigma_i(f'(x))$$

so by (7.3.6),

$$D = aN[L/K](f'(x)). \quad ♣$$

The discriminant of an integral basis for a number field has special properties. We will get at these results by considering the general $AKLB$ setup, adding some additional conditions as we go along.

7.4.7 Lemma

There is a basis for L/K consisting entirely of elements of B.

Proof. Let $x_1, \ldots, x_n$ be a basis for L over K. Each x_i is algebraic over K, and therefore satisfies a polynomial equation of the form

$$a_m x_i^m + \cdots + a_1 x_i + a_0 = 0$$

with $a_m \neq 0$ and the $a_i \in A$. (Initially, we only have $a_i \in K$, but then a_i is the ratio of two elements in A, and we can form a common denominator.) Multiply the equation by a_m^{m-1} to obtain an equation of integral dependence for $y_i = a_m x_i$ over A. The y_i form the desired basis. ♣

7.4.8 Theorem

Suppose we have a nondegenerate symmetric bilinear form on an n-dimensional vector space V, written for convenience using inner product notation (x, y). If $x_1, \ldots, x_n$ is any basis for V, then there is a basis $y_1, \ldots, y_n$ for V, called the *dual basis referred to* V, such that

$$(x_i, y_j) = \delta_{ij} = \begin{cases} 1, & i = j; \\ 0, & i \neq j. \end{cases}$$

This is a standard (and quite instructive) result in linear algebra, and it will be developed in the exercises.

7.4.9 Theorem

If A is a principal ideal domain, then B is a free A-module of rank n.

Proof. By (7.3.8), the trace is a nondegenerate symmetric bilinear form on the n-dimensional vector space L over K. By (7.1.7), A is integrally closed, so by (7.3.10), the trace of any element of L belongs to A. Now let $x_1, \ldots, x_n$ be any basis for L over K, and let $y_1, \ldots, y_n$ be the dual basis referred to L (see (7.4.8)). If $z \in B$, then we can write $z = \sum_{j=1}^n a_j y_j$ with $a_j \in K$. We know that the trace of $x_i z$ belongs to A, and we also have

$$T(x_i z) = T\Big(\sum_{j=1}^n a_j x_i y_j\Big) = \sum_{j=1}^n a_j T(x_i y_j) = \sum_{j=1}^n a_j \delta_{ij} = a_i.$$

Thus each a_i belongs to A, so that B is an A-submodule of the free A-module $\oplus_{j=1}^n A y_j$. By (4.6.2), B is a free A-module of rank at most n. But by (7.4.7), B contains a basis for L over K, and if we wish, we can assume that this basis is $x_1, \ldots, x_n$. Then B contains the free A-module $\oplus_{j=1}^n A x_j$, so the rank of B as an A-module is at least n, and hence exactly n. ♣

7.4.10 Corollary

The set B of algebraic integers in any number field L is a free $\mathbb{Z}$-module of rank $n = [L : \mathbb{Q}]$. Therefore B has an integral basis. The discriminant is the same for every integral basis; it is known as the *field discriminant*.

Proof. Take $A = \mathbb{Z}$ in (7.4.9) to show that B has an integral basis. The transformation matrix C between two integral bases (see (7.4.2)) is invertible, and both C and C^{-1} have rational integer coefficients. Take determinants in the equation $CC^{-1} = I$ to conclude that $\det C$ is a unit in $\mathbb{Z}$. Therefore $\det C = \pm 1$, so by (7.4.2), all integral bases have the same discriminant. ♣

Problems For Section 7.4

Let $x_1, \ldots, x_n$ be a basis for the vector space V, and let (x, y) be a nondegenerate symmetric bilinear form on V. We now supply the details of the proof of (7.4.8).

1. For any $y \in V$, the mapping $x \to (x, y)$ is a linear form $l(y)$, i.e., a linear map from V to the field of scalars. Show that the linear transformation $y \to l(y)$ from V to V^*, the dual space of V (i.e., the space of all linear forms on V), is injective.
2. Show that any linear form on V is $l(y)$ for some $y \in V$.
3. Let $f_1, \ldots, f_n$ be the dual basis corresponding to $x_1, \ldots, x_n$. Thus each f_j belongs to V^* (*not* V) and $f_j(x_i) = \delta_{ij}$. If $f_j = l(y_j)$, show that $y_1, \ldots, y_n$ is the required dual basis referred to V.
4. Show that $x_i = \sum_{j=1}^{n} (x_i, x_j) y_j$. Thus in order to compute the dual basis referred to V in terms of the original basis, we must invert the matrix $((x_i, x_j))$.
5. A matrix C with coefficients in $\mathbb{Z}$ is said to be *unimodular* if its determinant is ± 1. Show that C is unimodular if and only if C is invertible and its inverse has coefficients in $\mathbb{Z}$.
6. Show that the field discriminant of the quadratic extension $\mathbb{Q}(\sqrt{d})$, d square-free, is
$$D = \begin{cases} 4d & \text{if } d \not\equiv 1 \bmod 4; \\ d & \text{if } d \equiv 1 \bmod 4. \end{cases}$$
7. Let $x_1, \ldots, x_n$ be arbitrary algebraic integers in a number field, and consider the determinant of the matrix $(\sigma_i(x_j))$, as in (7.4.3). The direct expansion of the determinant has $n!$ terms. Let P be the sum of those terms in the expansion that have plus signs in front of them, and N the sum of those terms prefixed by minus signs. Thus the discriminant D of $(x_1, \ldots, x_n)$ is $(P - N)^2$. Show that $P + N$ and PN are fixed by each σ_i, and deduce that $P + N$ and PN are rational numbers.
8. Continuing Problem 7, show that $P + N$ and PN are rational integers.
9. Continuing Problem 8, prove *Stickelberger's theorem*: $D \equiv 0$ or $1 \bmod 4$.
10. Let L be a number field of degree n over $\mathbb{Q}$, and let $y_1, \ldots, y_n$ be a basis for L over $\mathbb{Q}$ consisting of algebraic integers. Let $x_1, \ldots, x_n$ be an integral basis. Show that if the discriminant $D(y_1, \ldots, y_n)$ is square-free, then each x_i can be expressed as a linear combination of the y_j with integer coefficients.

11. Continuing Problem 10, show that if $D(y_1, \ldots, y_n)$ is square-free, then $y_1, \ldots, y_n$ is an integral basis.

12. Is the converse of the result of Problem 11 true?

13. In the standard $AKLB$ setup (see (7.3.9)), show that L is the quotient field of B.

7.5 Noetherian and Artinian Modules and Rings

7.5.1 Definitions and Comments

In this section, rings are *not* assumed commutative. Let M be an R-module, and suppose that we have an increasing sequence of submodules $M_1 \leq M_2 \leq M_3 \leq \ldots$, or a decreasing sequence $M_1 \geq M_2 \geq M_3 \geq \ldots$. We say that the sequence *stabilizes* if for some t, $M_t = M_{t+1} = M_{t+2} = \ldots$. The question of stabilization of sequences of submodules appears in a fundamental way in many areas of abstract algebra and its applications. We have already made contact with the idea; see (2.6.6) and the introductory remarks in Section 4.6.

The module M is said to satisfy the *ascending chain condition (acc)* if every increasing sequence of submodules stabilizes; M satisfies the *descending chain condition (dcc)* if every decreasing sequence of submodules stabilizes.

7.5.2 Proposition

The following conditions on an R-module M are equivalent, and define a *Noetherian module*:

(1) M satisfies the acc;

(2) Every nonempty collection of submodules of M has a maximal element (with respect to inclusion).

The following conditions on M are equivalent, and define an *Artinian module*:

(1') M satisfies the dcc;

(2') Every nonempty collection of submodules of M has a minimal element.

Proof. Assume (1), and let $\mathcal{S}$ be a nonempty collection of submodules. Choose $M_1 \in \mathcal{S}$. If M_1 is maximal, we are finished; otherwise we have $M_1 < M_2$ for some $M_2 \in \mathcal{S}$. If we continue inductively, the process must terminate at a maximal element; otherwise the acc would be violated.

Conversely, assume (2), and let $M_1 \leq M_2 \leq \ldots$. The sequence must stabilize; otherwise $\{M_1, M_2, \ldots\}$ would be a nonempty collection of submodules with no maximal element. The proof is exactly the same in the Artinian case, with all inequalities reversed. ♣

There is another equivalent condition in the Noetherian case.

7.5.3 Proposition

M is Noetherian iff every submodule of M is finitely generated.

Proof. If the sequence $M_1 \leq M_2 \leq \dots$ does not stabilize, let $N = \cup_{r=1}^{\infty} M_r$. Then N is a submodule of M, and it cannot be finitely generated. For if $x_1, \dots, x_s$ generate N, then for sufficiently large t, all the x_i belong to M_t. But then $N \subseteq M_t \subseteq M_{t+1} \subseteq \cdots \subseteq N$, so $M_t = M_{t+1} = \dots$. Conversely, assume that the acc holds, and let $N \leq M$. If $N \neq 0$, choose $x_1 \in N$. If $Rx_1 = N$, then N is finitely generated. Otherwise, there exists $x_2 \notin Rx_1$. If x_1 and x_2 generate N, we are finished. Otherwise, there exists $x_3 \notin Rx_1 + Rx_2$. The acc forces the process to terminate at some stage t, in which case $x_1, \dots, x_t$ generate N. ♣

The analogous equivalent condition in the Artinian case (see Problem 8) is that every quotient module M/N is *finitely cogenerated*, that is, if the intersection of a collection of submodules of M/N is 0, then there is a finite subcollection whose intersection is 0.

7.5.4 Definitions and Comments

A ring R is *Noetherian* [resp. *Artinian*] if it is Noetherian [resp. Artinian] as a module over itself. If we need to distinguish between R as a left, as opposed to right, R-module, we will refer to a *left Noetherian* and a *right Noetherian* ring, and similarly for Artinian rings. This problem will not arise until Chapter 9.

7.5.5 Examples

1. Every PID is Noetherian.

   This follows from (7.5.3), since every ideal is generated by a single element.

2. $\mathbb{Z}$ is Noetherian (a special case of Example 1) but not Artinian. There are many descending chains of ideals that do not stabilize, e.g.,

   $$\mathbb{Z} \supset (2) \supset (4) \supset (8) \supset \dots.$$

   We will prove in Chapter 9 that an Artinian ring must also be Noetherian.

3. If F is a field, then the polynomial ring $F[X]$ is Noetherian (another special case of Example 1) but not Artinian. A descending chain of ideals that does not stabilize is

   $$(X) \supset (X^2) \supset (X^3) \supset \dots.$$

4. The ring $F[X_1, X_2, \dots]$ of polynomials over F in infinitely many variables is neither Artinian nor Noetherian. A descending chain of ideals that does not stabilize is constructed as in Example 3, and an ascending chain of ideals that does not stabilize is

   $$(X_1) \subset (X_1, X_2) \subset (X_1, X_2, X_3) \subset \dots.$$

7.5.6 Remark

The following observations will be useful in deriving properties of Noetherian and Artinian modules. If $N \leq M$, then a submodule L of M that contains N can always be written in the form $K+N$ for some submodule K. ($K = L$ is one possibility.) By the correspondence theorem,

$$(K_1 + N)/N = (K_2 + N)/N \text{ implies } K_1 + N = K_2 + N \text{ and}$$
$$(K_1 + N)/N \leq (K_2 + N)/N \text{ implies } K_1 + N \leq K_2 + N.$$

7.5.7 Proposition

If N is a submodule of M, then M is Noetherian [resp. Artinian] if and only if N and M/N are Noetherian [resp. Artinian].

Proof. Assume M is Noetherian. Then N is Noetherian by (2) of (7.5.2), since a submodule of N must also be a submodule of M. By (7.5.6), an ascending chain of submodules of M/N looks like $(M_1 + N)/N \leq (M_2 + N)/N \leq \cdots$. But then the $M_i + N$ form an ascending sequence of submodules of M, which must stabilize. Consequently, the sequence $(M_i + N)/N, i = 1, 2, \ldots$, must stabilize.

Conversely, assume that N and M/N are Noetherian, and let $M_1 \leq M_2 \leq \cdots$ be an increasing sequence of submodules of M. Take i large enough so that both sequences $\{M_i \cap N\}$ and $\{M_i + N\}$ have stabilized. If $x \in M_{i+1}$, then $x + N \in M_{i+1} + N = M_i + N$, so $x = y + z$ where $y \in M_i$ and $z \in N$. Thus $x - y \in M_{i+1} \cap N = M_i \cap N$, and since $y \in M_i$ we have $x \in M_i$ as well. Consequently, $M_i = M_{i+1}$ and the sequence of M_i's has stabilized. The Artinian case is handled by reversing inequalities (and interchanging indices i and $i + 1$ in the second half of the proof). ♣

7.5.8 Corollary

If $M_1, \ldots, M_n$ are Noetherian [resp. Artinian] R-modules, then so is $M_1 \oplus M_2 \oplus \cdots \oplus M_n$.

Proof. It suffices to consider $n = 2$ (induction will take care of higher values of n). The submodule $N = M_1$ of $M = M_1 \oplus M_2$ is Noetherian by hypothesis, and $M/N \cong M_2$ is also Noetherian (apply the first isomorphism theorem to the natural projection of M onto M_2). By (7.5.7), M is Noetherian. The Artinian case is done the same way. ♣

7.5.9 Corollary

If M is a finitely generated module over the Noetherian [resp. Artinian] ring R, then M is Noetherian [resp. Artinian]

Proof. By (4.3.6), M is a quotient of a free module L of finite rank. Since L is the direct sum of a finite number of copies of R, the result follows from (7.5.8) and (7.5.7). ♣

Ascending and descending chains of submodules are reminiscent of normal and subnormal series in group theory, and in fact we can make a precise connection.

7.5.10 Definitions

A *series* of *length* n for a module M is a sequence of the form

$$M = M_0 \geq M_1 \geq \cdots \geq M_n = 0.$$

The series is called a *composition series* if each factor module M_i/M_{i+1} is *simple*. [A module is simple if it is nonzero and has no submodules except itself and 0. We will study simple modules in detail in Chapter 9.] Thus we are requiring the series to have no proper refinement. Two series are *equivalent* if they have the same length and the same factor modules, up to isomorphism and rearrangement.

By convention, the zero module has a composition series, namely $\{0\}$ itself.

7.5.11 Jordan-Hölder Theorem For Modules

If M has a composition series, then any two composition series for M are equivalent. Furthermore, any strictly decreasing sequence of submodules can be refined to a composition series.

Proof. The development of the Jordan-Hölder theorem for groups can be taken over verbatim if we change multiplicative to additive notation. In particular, we can reproduce the preliminary lemma (5.6.2), the Zassenhaus lemma (5.6.3), the Schreier refinement theorem (5.6.5), and the Jordan-Hölder Theorem (5.6.6). We need not worry about normality of subgroups because in an abelian group, all subgroups are normal. As an example of the change in notation, the Zassenhaus lemma becomes

$$\frac{A + (B \cap D)}{A + (B \cap C)} \cong \frac{C + (D \cap B)}{C + (D \cap A)}.$$

This type of proof can be irritating, because it forces readers to look at the earlier development and make sure that everything does carry over. A possible question is "Why can't a composition series $\mathcal{S}$ of length n coexist with an *infinite* ascending or descending chain?" But if such a situation occurs, we can form a series $\mathcal{T}$ for M of length $n+1$. By Schreier, $\mathcal{S}$ and $\mathcal{T}$ have equivalent refinements. Since $\mathcal{S}$ has no proper refinements, and equivalent refinement have the same length, we have $n \geq n+1$, a contradiction. ♣

We can now relate the ascending and descending chain conditions to composition series.

7.5.12 Theorem

The R-module M has a composition series if and only if M is both Noetherian and Artinian.

Proof. The "only if" part was just done at the end of the proof of (7.5.11). Thus assume that M is Noetherian and Artinian. Assuming (without loss of generality) that $M \neq 0$, it follows from (2) of (7.5.2) that $M_0 = M$ has a maximal proper submodule M_1. Now M_1 is Noetherian by (7.5.7), so if $M_1 \neq 0$, then M_1 has a maximal proper submodule M_2. Continuing inductively, we must reach 0 at some point because M is Artinian. By construction, each M_i/M_{i+1} is simple, and we have a composition series for M. ♣

Here is a connection with algebraic number theory.

7.5.13 Proposition

In the basic $AKLB$ setup of (7.3.9), assume that A is integrally closed. If A is a Noetherian ring, then so is B. In particular, the ring of algebraic integers in a number field is Noetherian.

Proof. By the proof of (7.4.9), B is a submodule of a free A-module M of finite rank. (The assumption that A is a PID in (7.4.9) is used to show that A is integrally closed, and we have this by hypothesis. The PID assumption is also used to show that B is a free A-module, but we do not need this in the present argument.) By (7.5.8), M is Noetherian, so by (7.5.7), B is a Noetherian A-module. An ideal of B is, in particular, an A-submodule of B, hence is finitely generated over A and therefore over B. Thus B is a Noetherian ring. ♣

Problems For Section 7.5

1. Let p be a fixed prime, and let A be the abelian group of all rational numbers a/p^n, $n = 0, 1, \ldots, a \in \mathbb{Z}$, where all calculations are modulo 1, in other words, A is a subgroup of $\mathbb{Q}/\mathbb{Z}$. Let A_n be the subgroup $\{0, 1/p^n, 2/p^n, \ldots, (p^n - 1)/p^n\}$. Show that A is not a Noetherian $\mathbb{Z}$-module.
2. Continuing Problem 1, if B is a proper subgroup of A, show that B must be one of the A_n. Thus A is an Artinian $\mathbb{Z}$-module. [This situation cannot arise for *rings*, where Artinian implies Noetherian.]
3. If V is a vector space, show that V is finite-dimensional iff V is Noetherian iff V is Artinian iff V has a composition series.
4. Define the *length* of a module M [notation $l(M)$] as the length of a composition series for M. (If M has no composition series, take $l(M) = \infty$.) Suppose that we have a short exact sequence

$$0 \longrightarrow N \xrightarrow{f} M \xrightarrow{g} M/N \longrightarrow 0$$

   Show that $l(M)$ is finite if and only if $l(N)$ and $l(M/N)$ are both finite.
5. Show that l is *additive*, that is,

$$l(M) = l(N) + l(M/N).$$

6. Let S be a subring of the ring R, and assume that S is a Noetherian ring. If R is finitely generated as a module over S, show that R is also a Noetherian ring.
7. Let R be a ring, and assume that the polynomial ring $R[X]$ is Noetherian. Does it follow that R is Noetherian?
8. Show that a module M is Artinian if and only if every quotient module M/N is finitely cogenerated.

7.6 Fractional Ideals

Our goal is to establish unique factorization of ideals in a Dedekind domain, and to do this we will need to generalize the notion of ideal. First, some preliminaries.

7.6.1 Definition

If $I_1, \ldots, I_n$ are ideals, the *product* $I_1 \cdots I_n$ is the set of all finite sums $\sum_i a_{1i}a_{2i}\cdots a_{ni}$, where $a_{ki} \in I_k$, $k = 1, \ldots, n$. It follows from the definition that the product is an ideal contained in each I_j.

7.6.2 Lemma

If P is a prime ideal that contains a product $I_1 \cdots I_n$ of ideals, then P contains I_j for some j.

Proof. If not, let $a_j \in I_j \setminus P$, $j = 1, \ldots, n$. Then $a_1 \cdots a_n$ belongs to $I_1 \cdots I_n \subseteq P$, and since P is prime, some a_j belongs to P, a contradiction. ♣

7.6.3 Proposition

If I is a nonzero ideal of the Noetherian integral domain R, then I contains a product of nonzero prime ideals.

Proof. Assume the contrary. If $\mathcal{S}$ is the collection of all nonzero ideals that do not contain a product of nonzero prime ideals, then since R is Noetherian, $\mathcal{S}$ has a maximal element J, and J cannot be prime because it belongs to $\mathcal{S}$. Thus there are elements $a, b \in R$ with $a \notin J$, $b \notin J$, and $ab \in J$. By maximality of J, the ideals $J + Ra$ and $J + Rb$ each contain a product of nonzero prime ideals, hence so does $(J + Ra)(J + Rb) \subseteq J + Rab = J$. This is a contradiction. [Notice that we must use the fact that a product of nonzero ideals is nonzero, and this is where the hypothesis that R is an integral domain comes in.] ♣

7.6.4 Corollary

If I is an ideal of the Noetherian ring R (not necessarily an integral domain), then I contains a product of prime ideals.

Proof. Repeat the proof of (7.6.3) with the word "nonzero" deleted. ♣

Ideals in the ring of integers are of the form $n\mathbb{Z}$, the set of multiples of n. A set of the form $\frac{3}{2}\mathbb{Z}$ is not an ideal because it is not a subset of $\mathbb{Z}$, yet it behaves in a similar manner. The set is closed under addition and multiplication by an integer, and it becomes an ideal of $\mathbb{Z}$ if we simply multiply all the elements by 2. It will be profitable to study sets of this type.

7.6.5 Definitions

Let R be an integral domain, with K its quotient field, and let I be an R-submodule of K. We say that I is a *fractional ideal* of R if $rI \subseteq R$ for some nonzero $r \in R$. We will call r a *denominator* of I. An ordinary ideal of R is a fractional ideal (take $r = 1$), and will often be referred to an as *integral ideal*.

7.6.6 Lemma

(i) If I is a finitely generated R-submodule of K, then I is a fractional ideal.

(ii) If R is Noetherian and I is a fractional ideal of R, then I is a finitely generated R-submodule of K.

(iii) If I and J are fractional ideals with denominators r and s respectively, then $I \cap J$, $I + J$ and IJ are fractional ideals with respective denominators r (or s), rs and rs. [The product of fractional ideals is defined exactly as in (7.6.1).]

Proof. (i) If $x_1 = a_1/b_1, \ldots, x_n = a_n/b_n$ generate I and $b = b_1 \cdots b_n$, then $bI \subseteq R$.

(ii) If $rI \subseteq R$, then $I \subseteq r^{-1}R$. As an R-module, $r^{-1}R$ is isomorphic to R and is therefore Noetherian. Consequently, I is finitely generated.

(iii) It follows from the definition (7.6.5) that the intersection, sum and product of fractional ideals are fractional ideals. The assertions about denominators are proved by noting that $r(I \cap J) \subseteq rI \subseteq R$, $rs(I + J) \subseteq rI + sJ \subseteq R$, and $rsIJ = (rI)(sJ) \subseteq R$. ♣

The product of two nonzero fractional ideals is a nonzero fractional ideal, and the multiplication is associative (since multiplication in R is associative). There is an identity element , namely R, since $RI \subseteq I = 1I \subseteq RI$. We will show that if R is a Dedekind domain, then every nonzero fractional ideal has a multiplicative inverse, so the nonzero fractional ideals form a group.

7.6.7 Definitions and Comments

A *Dedekind domain* is an integral domain R such that

(1) R is Noetherian,

(2) R is integrally closed, and

(3) Every nonzero prime ideal of R is maximal.

Every PID is a Dedekind domain, by (7.5.5), (7.1.7), (2.6.8) and (2.6.9). We will prove that the algebraic integers of a number field form a Dedekind domain. But as we know, the ring of algebraic integers need not be a PID, or even a UFD (see the discussion at the beginning of this chapter, and the exercises in Section 7.7).

7.6.8 Lemma

Let I be a nonzero prime ideal of the Dedekind domain R, and let $J = \{x \in K : xI \subseteq R\}$. Then $R \subset J$.

Proof. Since $RI \subseteq R$, it follows that R is a subset of J. Pick a nonzero element $a \in I$, so that I contains the principal ideal Ra. Let n be the smallest positive integer such that Ra contains a product $P_1 \cdots P_n$ of n nonzero prime ideals. Since R is Noetherian, there is such an n by (7.6.3), and by (7.6.2), I contains one of the P_i, say P_1. But in a Dedekind domain, every nonzero prime ideal is maximal, so $I = P_1$. Assuming $n \geq 2$, set $I_1 = P_2 \cdots P_n$, so that $Ra \not\supseteq I_1$ by minimality of n. Choose $b \in I_1$ with $b \notin Ra$. Now $II_1 \subseteq Ra$, in particular, $Ib \subseteq Ra$, hence $Iba^{-1} \subseteq R$. (Note that a has an inverse in K, but not necessarily in R.) Thus $ba^{-1} \in J$, but $ba^{-1} \notin R$, for if so, $b \in Ra$, contradicting the choice of b.

The case $n = 1$ must be handled separately. In this case, $P_1 = I \supseteq Ra \supseteq P_1$, so $I = Ra$. Thus Ra is a proper ideal, and we can choose $b \in R$ with $b \notin Ra$. Then $ba^{-1} \notin R$, but $ba^{-1}I = ba^{-1}Ra = bR \subseteq R$, so $ba^{-1} \in J$. ♣

We now prove that in (7.6.8), J is the inverse of I.

7.6.9 Proposition

Let I be a nonzero prime ideal of the Dedekind domain R, and let $J = \{x \in K : xI \subseteq R\}$. Then J is a fractional ideal and $IJ = R$.

Proof. By definition, J is an R-submodule of K. If r is a nonzero element of I and $x \in J$, then $rx \in R$, so $rJ \subseteq R$ and J is a fractional ideal. Now $IJ \subseteq R$ by definition of J, so IJ is an integral ideal. Since (using (7.6.8)) $I = IR \subseteq IJ \subseteq R$, maximality of I implies that either $IJ = I$ or $IJ = R$. In the latter case, we are finished, so assume $IJ = I$.

If $x \in J$, then $xI \subseteq IJ = I$, and by induction, $x^nI \subseteq I$ for all $n = 1, 2, \ldots$. Let r be any nonzero element of I. Then $rx^n \in x^nI \subseteq I \subseteq R$, so $R[x]$ is a fractional ideal. Since R is Noetherian, part (ii) of (7.6.6) implies that $R[x]$ is a finitely generated R-submodule of K. By (7.1.2), x is integral over R. But R, a Dedekind domain, is integrally closed, so $x \in R$. Therefore $J \subseteq R$, contradicting (7.6.8). ♣

Problems For Section 7.6

1. Show that a proper ideal P is prime if and only if for all ideals A and B, $P \supseteq AB$ implies that $P \supseteq A$ or $P \supseteq B$.

We are going to show that if an ideal I is contained in the union of the prime ideals $P_1, \ldots, P_n$, then I is contained in some P_i. Equivalently, if for all $i = 1, \ldots, n$, we have $I \not\subseteq P_i$, then $I \not\subseteq \cup_{i=1}^n P_i$. There is no problem when $n = 1$, so assume the result holds for $n - 1$ prime ideals. By the induction hypothesis, for each i there exists $x_i \in I$ with $x_i \notin P_j$, $j \neq i$.

2. Show that we can assume without loss of generality that $x_i \in P_i$ for all i.
3. Continuing Problem 2, let $x = \sum_{i=1}^n x_1 \cdots x_{i-1}x_{i+1} \cdots x_n$. Show that $x \in I$ but $x \notin \cup_{i=1}^n P_i$, completing the proof.
4. If I and J are relatively prime ideals ($I+J = R$), show that $IJ = I \cap J$. More generally, if $I_1, \ldots, I_n$ are relatively prime in pairs (see (2.3.7)), show that $I_1 \cdots I_n = \cap_{i=1}^n I_i$.

5. Show that if a Dedekind domain R is a UFD, then R is a PID.
6. Suppose that in (7.6.9), we would like to invert every maximal ideal of R, rather than the nonzero prime ideals. What is a reasonable hypothesis to add about R?
7. Let R be an integral domain with quotient field K. If K is a fractional ideal of R, show that $R = K$.
8. Let P_1 and P_2 be relatively prime ideals in the ring R. Show that P_1^r and P_2^s are relatively prime for arbitrary positive integers r and s.

7.7 Unique Factorization of Ideals in a Dedekind Domain

In the previous section, we inverted nonzero prime ideals in a Dedekind domain. We must now extend this result to nonzero fractional ideals.

7.7.1 Theorem

If I is a nonzero fractional ideal of the Dedekind domain R, then I can be factored uniquely as $P_1^{n_1}P_2^{n_2}\cdots P_r^{n_r}$ where the n_i are integers. Consequently, the nonzero fractional ideals form a group under multiplication.

Proof. First consider the existence of such a factorization. Without loss of generality, we can restrict to integral ideals. [Note that if $r \neq 0$ and $rI \subseteq R$, then $I = (Rr)^{-1}(rI)$.] By convention, we regard R as the product of the empty collection of prime ideals, so let $\mathcal{S}$ be the set of all nonzero proper ideals of R that cannot be factored in the given form, with all n_i *positive* integers. [This trick will yield the useful result that the factorization of integral ideals only involves positive exponents.] Since R is Noetherian, $\mathcal{S}$, if nonempty, has a maximal element I_0, which is contained in a maximal ideal I. By (7.6.9), I has an inverse fractional ideal J. Thus by (7.6.8) and (7.6.9),

$$I_0 = I_0R \subseteq I_0J \subseteq IJ = R.$$

Therefore I_0J is an integral ideal, and we claim that $I_0 \subset I_0J$. For if $I_0 = I_0J$, the last paragraph of the proof of (7.6.9) can be reproduced with I replaced by I_0 to reach a contradiction. By maximality of I_0, I_0J is a product of prime ideals, say $I_0J = P_1\cdots P_r$ (with repetition allowed). Multiply both sides by the prime ideal I to conclude that I_0 is a product of prime ideals, contradicting $I_0 \in \mathcal{S}$. Thus $\mathcal{S}$ must be empty, and the existence of the desired factorization is established.

To prove uniqueness, suppose that we have two prime factorizations

$$P_1^{n_1}\cdots P_r^{n_r} = Q_1^{t_1}\cdots Q_s^{t_s}$$

where again we may assume without loss of generality that all exponents are positive. [If P^{-n} appears, multiply both sides by P^n.] Now P_1 contains the product of the $P_i^{n_i}$, so by (7.6.2), P_1 contains Q_j for some j. By maximality of Q_j, $P_1 = Q_j$, and we may renumber so that $P_1 = Q_1$. Multiply by the inverse of P_1 (a fractional ideal, but there is no problem) to cancel P_1 and Q_1, and continue inductively to complete the proof. ♣

7.7.2 Corollary

A nonzero fractional ideal I is an integral ideal if and only if all exponents in the prime factorization of I are nonnegative.

Proof. The "only if" part was noted in the proof of (7.7.1). The "if" part follows because a power of an integral ideal is still an integral ideal. ♣

7.7.3 Corollary

Denote by $n_P(I)$ the exponent of the prime ideal P in the factorization of I. (If P does not appear, take $n_P(I) = 0$.) If I_1 and I_2 are nonzero fractional ideals, then $I_1 \supseteq I_2$ if and only if for every prime ideal P of R, $n_P(I_1) \leq n_P(I_2)$.

Proof. We have $I_2 \subseteq I_1$ iff $I_2 I_1^{-1} \subseteq R$, and by (7.7.2), this happens iff for every P, $n_P(I_2) - n_P(I_1) \geq 0$. ♣

7.7.4 Definition

Let I_1 and I_2 be nonzero integral ideals. We say that I_1 *divides* I_2 if $I_2 = JI_1$ for some integral ideal J. Just as with integers, an equivalent statement is that each prime factor of I_1 is a factor of I_2.

7.7.5 Corollary

If I_1 and I_2 are nonzero integral ideals, then I_1 divides I_2 if and only if $I_1 \supseteq I_2$. In other words, for these ideals,

DIVIDES MEANS CONTAINS.

Proof. By (7.7.4), I_1 divides I_2 iff $n_P(I_1) \leq n_P(I_2)$ for every prime ideal P. By (7.7.3), this is equivalent to $I_1 \supseteq I_2$. ♣

The next result explains why Dedekind domains are important in algebraic number theory.

7.7.6 Theorem

In the basic $AKLB$ setup of (7.3.9), if A is a Dedekind domain, then so is B. In particular, the ring of algebraic integers in a number field is a Dedekind domain. In addition, B is a finitely generated A-module and the quotient field of B is L.

Proof. By (7.1.6), B is integrally closed in L. The proof of (7.4.7), with x_i replaced by an arbitrary element of L, shows that L is the quotient field of B. Therefore B is integrally closed. By (7.5.13), B is a Noetherian ring, and the proof of (7.5.13) shows that B is a Noetherian, hence finitely generated, A-module.

It remains to prove that every nonzero prime ideal Q of B is maximal. Choose any nonzero element x of Q. Since $x \in B$, x satisfies a polynomial equation

$$x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0 = 0$$

with the $a_i \in A$. If we take the positive integer n as small as possible, then $a_0 \neq 0$ by minimality of n. Solving for a_0, we see that $a_0 \in Bx \cap A \subseteq Q \cap A$, so $P = Q \cap A \neq 0$. But P is the preimage of the prime ideal Q under the inclusion map of A into B. Therefore P is a nonzero prime, hence maximal, ideal of the Dedekind domain A. Consequently, A/P is a field.

Now A/P can be identified with a subring of the integral domain B/Q via $y + P \rightarrow y + Q$. Moreover, B/Q is integral over A/P. [B is integral over A, and we can simply use the same equation of integral dependence.] It follows from Section 7.1, Problem 5, that B/Q is a field, so Q is a maximal ideal. ♣

Problems For Section 7.7

By (7.2.3), the ring B of algebraic integers in $\mathbb{Q}(\sqrt{-5})$ is $\mathbb{Z}[\sqrt{-5}]$. We will show that $\mathbb{Z}[\sqrt{-5}]$ is not a unique factorization domain. (For a different approach, see Section 2.7, Problems 5–7.) Consider the factorization

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

1. By computing norms, verify that all four of the above factors are irreducible.
2. Show that the only units of B are ± 1.
3. Show that no factor on one side of the above equation is an associate of a factor on the other side, so unique factorization fails.
4. We can use the prime factorization of ideals in a Dedekind domain to compute the greatest common divisor and the least common multiple of two nonzero ideals I and J, exactly as with integers. Show that the greatest common divisor of I and J is $I + J$ and the least common multiple is $I \cap J$.
5. A Dedekind domain R comes close to being a PID in the following sense. (All ideals are assumed nonzero.) If I is an integral ideal, in fact if I is a fractional ideal, show that there is an integral ideal J such that IJ is a principal ideal of R.
6. Show that the ring of algebraic integers in $\mathbb{Q}(\sqrt{-17})$ is not a unique factorization domain.
7. In Problem 6, the only algebraic integers of norm 1 are ± 1. Show that this property does not hold for the algebraic integers in $\mathbb{Q}(\sqrt{-3})$.

7.8 Some Arithmetic in Dedekind Domains

Unique factorization of ideals in a Dedekind domain permits calculations that are analogous to familiar manipulations involving ordinary integers. In this section, we illustrate some of the ideas.

Let $P_1, \ldots, P_n$ be distinct nonzero prime ideals of the Dedekind domain R, and let $J = P_1 \cdots P_n$. Let Q_i be the product of the P_j with P_i omitted, that is,

$$Q_i = P_1 \cdots P_{i-1} P_{i+1} \cdots P_n.$$

(If $n = 1$, we take $Q_i = R$.) If I is any nonzero ideal of R, then by unique factorization, $IQ_i \supset IJ$. For each $i = 1, \ldots, n$, choose an element a_i belonging to IQ_i but not to IJ, and let $a = \sum_{i=1}^{n} a_i$.

7.8.1 Lemma

$a \in I$ but for each i, $a \notin IP_i$. (In particular, $a \neq 0$.)

Proof. Since each a_i belongs to $IQ_i \subseteq I$, we have $a \in I$. Now a_i cannot belong to IP_i, for if so, $a_i \in IP_i \cap IQ_i$, which is the least common multiple of IP_i and IQ_i (see Section 7.7, Problem 4). But by definition of Q_i, the least common multiple is simply IJ, and this contradicts the choice of a_i. We break up the sum defining a as follows:

$$a = (a_1 + \cdots + a_{i-1}) + a_i + (a_{i+1} + \cdots + a_n). \tag{1}$$

If $j \neq i$, then $a_j \in IQ_j \subseteq IP_i$, so the first and third terms of (1) belong to IP_i. Since $a_i \notin IP_i$, as found above, we have $a \notin IP_i$. ♣

In Section 7.7, Problem 5, we found that any nonzero ideal is a factor of a principal ideal. We can sharpen this result as follows.

7.8.2 Proposition

Let I be a nonzero ideal of the Dedekind domain R. Then there is a nonzero ideal I' such that II' is a principal ideal (a). Moreover, if J is an arbitrary nonzero ideal of R, I' can be chosen to be relatively prime to J.

Proof. Let $P_1, \ldots, P_n$ be the distinct prime divisors of J, and choose a as in (7.8.1). Then $a \in I$, so $(a) \subseteq I$. Since divides means contains (see (7.7.5)), I divides (a), so $(a) = II'$ for some nonzero ideal I'. If I' is divisible by P_i, then $I' = P_i I_0$ for some nonzero ideal I_0, and $(a) = IP_i I_0$. Consequently, $a \in IP_i$, contradicting (7.8.1). ♣

7.8.3 Corollary

A Dedekind domain with only finitely many prime ideals is a PID.

Proof. Let J be the product of all the nonzero prime ideals. If I is any nonzero ideal, then by (7.8.2) there is a nonzero ideal I' such that II' is a principal ideal (a), with I' relatively prime to J. But then the set of prime factors of I' is empty, so that $I' = R$. Thus $(a) = II' = IR = I$. ♣

The next result shows that a Dedekind domain is not too far away from a principal ideal domain.

7.8.4 Corollary

Let I be a nonzero ideal of the Dedekind domain R, and let a be any nonzero element of I. Then I can be generated by two elements, one of which is a.

Proof. Since $a \in I$, we have $(a) \subseteq I$, so I divides (a), say $(a) = IJ$. By (7.8.2) there is a nonzero ideal I' such that II' is a principal ideal (b) and I' is relatively prime to J. If gcd stands for greatest common divisor, then the ideal generated by a and b is

$$\gcd((a),(b)) = \gcd(IJ, II') = I$$

since $\gcd(J, I') = (1)$. ♣

Problems For Section 7.8

1. Let $I(R)$ be the group of nonzero fractional ideals of a Dedekind domain R. If $P(R)$ is the subset of $I(R)$ consisting of all nonzero *principal fractional ideals* Rx, $x \in K$, show that $P(R)$ is a subgroup of $I(R)$. The quotient group $C(R) = I(R)/P(R)$ is called the *ideal class group* of R. Since R is commutative, $C(R)$ is abelian, and it can be shown that in the number field case, $C(R)$ is finite.
2. Continuing Problem 1, show that $C(R)$ is trivial iff R is a PID.

We will now go through the factorization of an ideal in a number field. The necessary background is developed in a course in algebraic number theory, but some of the manipulations are accessible to us now. By (7.2.3), the ring B of algebraic integers of the number field $\mathbb{Q}(\sqrt{-5})$ is $\mathbb{Z}[\sqrt{-5}]$. (Note that $-5 \equiv 3 \bmod 4$.) If we wish to factor the ideal $(2) = 2B$ in B, the idea is to factor $x^2 + 5 \bmod 2$, and the result is $x^2 + 5 \equiv (x+1)^2 \bmod 2$. Identifying x with $\sqrt{-5}$, we form the ideal $P_2 = (2, 1 + \sqrt{-5})$, which turns out to be prime. The desired factorization is $(2) = P_2^2$. This technique works if $B = \mathbb{Z}[\alpha]$, where the number field L is $\mathbb{Q}(\alpha)$.

3. Show that $1 - \sqrt{-5} \in P_2$, and conclude that $6 \in P_2^2$.
4. Show that $2 \in P_2^2$, hence $(2) \subseteq P_2^2$.
5. Expand $P_2^2 = (2, 1 + \sqrt{-5})(2, 1 + \sqrt{-5})$, and conclude that $P_2^2 \subseteq (2)$.
6. Following the technique suggested in the above problems, factor $x^2 + 5 \bmod 3$, and conjecture that the prime factorization of (3) in the ring of integers of $\mathbb{Q}(\sqrt{-5})$ is $(3) = P_3 P_3'$ for appropriate P_3 and P_3'.
7. With P_3 and P_3' as found in Problem 6, verify that $(3) = P_3 P_3'$.

7.9 p-adic Numbers

We will give a very informal introduction to this basic area of number theory. (For further details, see Gouvea, "p-adic Numbers".) Throughout the discussion, p is a fixed prime.

7.9.1 Definitions and Comments

A *p-adic integer* can be described in several ways. One representation is via a series

$$x = a_0 + a_1 p + a_2 p^2 + \cdots, \ a_i \in \mathbb{Z}. \tag{1}$$

(Let's ignore the problem of convergence for now.) The partial sums are $x_n = a_0 + a_1 p + \cdots + a_n p^n$, so that $x_n - x_{n-1} = a_n p^n$. A p-adic integer can also be defined as a sequence of integers $x = \{x_0, x_1, \ldots, \}$ satisfying

$$x_n \equiv x_{n-1} \mod p^n, \ n = 1, 2, \ldots. \tag{2}$$

Given a sequence satisfying (2), we can recover the coefficients of the series (1) by

$$a_0 = x_0, \ a_1 = \frac{x_1 - x_0}{p}, \ a_2 = \frac{x_2 - x_1}{p^2}, \ldots.$$

The sequences x and y are regarded as defining the same p-adic integer iff $x_n \equiv y_n \mod p^{n+1}$, $n = 0, 1, \ldots$. Replacing each x_n by the smallest nonnegative integer congruent to it mod p^{n+1} is equivalent to restricting the a_i in (1) to $\{0, 1, \ldots, p-1\}$. [We call this the *standard representation*.] Thus (1) is the limiting case in some sense of an expansion in base p.

Sums and products of p-adic integers can be defined by polynomial multiplication if (1) is used. With the representation (2), we take

$$x + y = \{x_n + y_n\}, \quad xy = \{x_n y_n\}.$$

With addition and multiplication defined in this way, we get the *ring of p-adic integers*, denoted by θ_p. (A more common notation is $\mathbb{Z}_p$, with the ring of integers modulo p written as $\mathbb{Z}/p\mathbb{Z}$. Since the integers mod p occur much more frequently in this text than the p-adic integers, and $\mathbb{Z}_p$ is a bit simpler than $\mathbb{Z}/p\mathbb{Z}$, I elected to use $\mathbb{Z}_p$ for the integers mod p.) The rational integers $\mathbb{Z}$ form a subring of θ_p via $x = \{x, x, x, \ldots\}$.

We now identify the units of θ_p.

7.9.2 Proposition

The p-adic integer $x = \{x_n\}$ is a unit of θ_p (also called a *p-adic unit*) if and only if $x_0 \not\equiv 0 \mod p$. In particular, a rational integer a is a p-adic unit if and only if $a \not\equiv 0 \mod p$.

Proof. If $(a_0 + a_1 p + \cdots)(b_0 + b_1 p + \cdots) = 1$, then $a_0 b_0 = 1$, so $a_0 \not\equiv 0 \mod p$, proving the "only if" part. Thus assume that $x_0 \not\equiv 0 \mod p$. By (2), $x_n \equiv x_{n-1} \equiv \cdots \equiv x_0 \mod p$, so $x_n \not\equiv 0 \mod p$. Therefore x_n and p^{n+1} are relatively prime, so there exists y_n such that $x_n y_n \equiv 1 \mod p^{n+1}$, hence mod p^n. Now by (2), $x_n \equiv x_{n-1} \mod p^n$, so $x_n y_{n-1} \equiv x_{n-1} y_{n-1} \equiv 1 \mod p^n$. Thus $x_n y_n \equiv x_n y_{n-1} \mod p^n$, so $y_n \equiv y_{n-1} \mod p^n$. The sequence $y = \{y_n\}$ is therefore a p-adic integer, and by construction, $xy = 1$. ♣

7.9.3 Corollary

Every nonzero p-adic integer has the form $x = p^n u$ where $n \geq 0$ and u is a p-adic unit. Consequently, θ_p is an integral domain. Furthermore, θ_p has only one prime element p, and every $x \in \theta_p$ is a power of p, up to multiplication by a unit.

Proof. The series representation for x has a nonzero term $a_n p^n$ of lowest degree n, where a_n can be taken between 1 and $p-1$. Factor out p^n to obtain $x = p^n u$, where u is a unit by (7.9.2). ♣

7.9.4 Definitions and Comments

The quotient field $\mathbb{Q}_p$ of θ_p is called the field of *p-adic numbers*. By (7.9.3), each $\alpha \in \mathbb{Q}_p$ has the form $p^m u$, where m is an integer (possibly negative) and u is a unit in θ_p. Thus α has a "Laurent expansion"

$$\frac{a_{-r}}{p^r} + \cdots + \frac{a_{-1}}{p} + a_0 + a_1 p + \cdots .$$

Another representation is $\alpha = x/p^r$, where x is a p-adic integer and $r \geq 0$. This version is convenient for doing addition and multiplication in $\mathbb{Q}_p$.

The rationals $\mathbb{Q}$ are a subfield of $\mathbb{Q}_p$. To see this, let a/b be a rational number in lowest terms (a and b relatively prime). If p does not divide b, then by (7.9.2), b is a unit of θ_p. Since $a \in \mathbb{Z} \subseteq \theta_p$, we have $a/b \in \theta_p$. If $b = p^t b'$ where p does not divide b', we can factor out p^t and reduce to the previous case. Thus a/b always belongs to $\mathbb{Q}_p$, and $a/b \in \theta_p$ iff p does not divide b. Rational numbers belonging to θ_p are sometimes called *p-integers*.

We now outline a procedure for constructing the p-adic numbers formally.

7.9.5 Definitions and Comments

The *p-adic valuation* on $\mathbb{Q}_p$ is defined by

$$v_p(p^m u) = m.$$

In general, a *valuation* v on a field F is a real-valued function on $F \setminus \{0\}$ satisfying:

(a) $v(xy) = v(x) + v(y)$;

(b) $v(x+y) \geq \min(v(x), v(y))$.

By convention, we take $v(0) = +\infty$.

The representation $x = p^m u$ shows that v_p is indeed a valuation on $\mathbb{Q}_p$. If c is any real number greater than 1, then the valuation v induces an *absolute value* on F, namely,

$$|x| = c^{-v(x)}.$$

When $v = v_p$, the constant c is usually taken to be p, and we obtain the *p-adic absolute value*

$$|x|_p = p^{-v_p(x)}.$$

Thus the p-adic absolute value of p^n is p^{-n}, which approaches 0 exponentially as n approaches infinity.

In general, an *absolute value* on a field F is a real-valued function on F such that:

(i) $|x| \geq 0$, with equality if and only if $x = 0$;

(ii) $|xy| = |x||y|$;

(iii) $|x + y| \leq |x| + |y|$.

By (b), an absolute value induced by a valuation satisfies a property that is stronger than (iii):

(iv) $|x + y| \leq \max(|x|, |y|)$.

An absolute value satisfying (iv) is said to be *nonarchimedian*.

7.9.6 Proposition

Let F be the quotient field of an integral domain R. The absolute value $|\ |$ on F is nonarchimedian if and only if $|n| \leq 1$ for every integer $n = 1 \pm \cdots \pm 1 \in R$.

Proof. Assume a nonarchimedian absolute value. By property (ii) of (7.9.5), $|\pm 1| = 1$. If $|n| \leq 1$, then by property (iv), $|n \pm 1| \leq 1$, and the desired conclusion follows by induction. Conversely, assume that the absolute value of every integer is at most 1. To prove (iv), it suffices to show that $|x + 1| \leq \max(|x|, 1)$ for every $x \in F$. [If $y \neq 0$ in (iv), divide by $|y|$.] By the binomial theorem,

$$|x + 1|^n = |\sum_{r=0}^{n} \binom{n}{r} x^r| \leq \sum_{r=0}^{n} |\binom{n}{r}||x|^r.$$

By hypothesis, the integer $\binom{n}{r}$ has absolute value at most 1. If $|x| > 1$, then $|x|^r \leq |x|^n$ for all $r = 0, 1, \ldots, n$. If $|x| \leq 1$, then $|x|^r \leq 1$. Consequently,

$$|x + 1|^n \leq (n + 1) \max(|x|^n, 1).$$

Take n^{th} roots and let $n \to \infty$ to get $|x + 1| \leq \max(|x|, 1)$. ♣

The next result may seem innocuous, but it leads to a remarkable property of nonarchimedian absolute values.

7.9.7 Proposition

If $|\ |$ is a nonarchimedian absolute value, then

$$|x| \neq |y| \text{ implies } |x + y| = \max(|x|, |y|).$$

Proof. First note that $|-y| = |(-1)y| = |-1||y| = |y|$. We can assume without loss of generality that $|x| > |y|$. Using property (iv) of (7.9.5), we have

$$|x| = |x + y - y| \leq \max(|x + y|, |y|) = |x + y|.$$

[Otherwise, $\max(|x + y|, |y|) = |y|$, hence $|x| \leq |y| < |x|$, a contradiction.] Since $|x + y| \leq \max(|x|, |y|) = |x|$, the result follows. ♣

Any absolute value determines a metric via $d(x, y) = |x - y|$. This distance function can be used to measure the length of the sides of a triangle.

7.9.8 Corollary

With respect to the metric induced by a nonarchimedian absolute value, all triangles are isosceles.

Proof. Let the vertices of the triangle be x, y and z. Then

$$|x - y| = |(x - z) + (z - y)|.$$

If $|x - z| = |z - y|$, then two side lengths are equal. If $|x - z| \neq ||z - y|$, then by (7.9.7), $|x - y| = \max(|x - z|, |z - y|)$, and again two side lengths are equal. ♣

We now look at the p-adic numbers from the viewpoint of valuation theory.

7.9.9 Definitions and Comments

Let $|\ |$ be a nonarchimedian absolute value on the field F. The *valuation ring* of $|\ |$ is

$$\theta = \{x \in F \colon |x| \leq 1\}.$$

In the p-adic case, $\theta = \{x \in \mathbb{Q}_p \colon v_p(x) \geq 0\} = \theta_p$. By properties (ii) and (iv) of (7.9.5), θ is a subring of F.

The *valuation ideal* of $|\ |$ is

$$\beta = \{x \in F \colon |x| < 1\}.$$

In the p-adic case, $\beta = \{x \in \mathbb{Q}_p \colon v_p(x) \geq 1\} = p\theta_p$, those p-adic integers whose series representation has no constant term. To verify that β is an ideal of θ, note that if $x, y \in \beta$ and $r \in \theta$, then $|rx| = |r||x| \leq |x| < 1$ and $|x + y| \leq \max(|x|, |y|) < 1$.

Now if $x \in \theta \setminus \beta$, then $|x| = 1$, hence $|x^{-1}| = 1/|x| = 1$, so $x^{-1} \in \theta$ and x is a unit of θ. On the other hand, if $x \in \beta$, then x cannot be a unit of θ. [If $xy = 1$, then $1 = |x||y| \leq |x| < 1$, a contradiction.] Thus the ideal β is the set of all nonunits of θ. No proper ideal I of θ can contain a unit, so $I \subseteq \beta$. It follows that β is the unique maximal ideal of θ. A ring with a unique maximal ideal is called a *local ring*. We will meet such rings again when we examine the localization process in Section 8.5.

To construct the p-adic numbers, we start with the p-adic valuation on the integers, and extend it to the rationals in the natural way: $v_p(a/b) = v_p(a) - v_p(b)$. The p-adic valuation then determines the p-adic absolute value, which induces a metric d on $\mathbb{Q}$.

[Because d comes from a nonarchimedian absolute value, it will satisfy the *ultrametric inequality* $d(x,y) \leq \max(d(x,z), d(z,y))$, which is stronger than the triangle inequality.] The process of constructing the real numbers by completing the rationals using equivalence classes of Cauchy sequences is familiar. The same process can be carried out using the p-adic absolute value rather than the usual absolute value on $\mathbb{Q}$. The result is a complete metric space, the field of p-adic numbers, in which $\mathbb{Q}$ is dense.

Ostrowski's theorem says that the usual absolute value $|\ |_\infty$ on $\mathbb{Q}$, along with the p-adic absolute values $|\ |_p$ for all primes p, and the *trivial* absolute value ($|0| = 0; |x| = 1$ for $x \neq 0$), essentially exhaust all possibilities. To be more precise, two absolute values on a field F are *equivalent* if the corresponding metrics on F induce the same topology. Any nontrivial absolute value on $\mathbb{Q}$ is equivalent to $|\ |_\infty$ or to one of the $|\ |_p$.

Problems For Section 7.9

1. Take $p = 3$, and compute the standard representation of $(2 + p + p^2)(2 + p^2)$ in two ways, using (1) and (2) of (7.9.1). Check the result by computing the product using ordinary multiplication of two integers, and then expanding in base $p = 3$.
2. Express the p-adic integer -1 as an infinite series of the form (1), using the standard representation.
3. Show that every absolute value on a finite field is trivial.
4. Show that an absolute value is archimedian iff the set $S = \{|n| : n \in \mathbb{Z}\}$ is unbounded.
5. Show that a field that has an archimedian absolute value must have characteristic 0.
6. Show that an infinite series $\sum z_n$ of p-adic numbers converges if and only if $z_n \to 0$ as $n \to \infty$.
7. Show that the sequence $a_n = n!$ of p-adic integers converges to 0.
8. Does the sequence $a_n = n$ converge in $\mathbb{Q}_p$?