CHAPTER 6

Polynomials

Before continuing with our treatment of linear operators and transformations, we must make a digression and consider the theory of polynomials in some detail. The subject matter of this chapter will be quite important throughout much of the remainder of this text. Our basic goal is to discuss the factorization of polynomials in detail, including many of the elementary properties that we all learned in high school.

6.1 DEFINITIONS

Let $\mathcal{F}$ be a field. In high school (or earlier), we all learned that a polynomial $p(x)$ in the indeterminate (or variable) x is basically an expression of the form

$$p(x) = a_0 + a_1x + a_2x^2 + \cdots + a_nx^n$$

where n is any nonnegative integer and $a_0, \ldots, a_n$ are all elements of $\mathcal{F}$. Note that our elementary experience with polynomials tells us that if

$$q(x) = b_0 + b_1x + \cdots + b_mx^m$$

is another polynomial in x then, assuming without loss of generality that $n \geq m$, we have (where we define $a_j = 0$ for $j > n$ and $b_j = 0$ for $j > m$)

$$p(x) + q(x) = (a_0 + b_0) + (a_1 + b_1)x + \cdots + (a_n + b_n)x^n$$

and

$$\begin{aligned} p(x)q(x) = {} & (a_0b_0) + (a_0b_1 + a_1b_0)x + (a_0b_2 + a_1b_1 + a_2b_0)x^2 \\ & + \cdots + (a_0b_k + a_1b_{k-1} + \cdots + a_{k-1}b_1 + a_kb_0)x^k \\ & + \cdots + a_nb_mx^{n+m} \ . \end{aligned}$$

While this has a definite intuitive appeal (to previous experience), the term "expression" in the above definition of polynomial is rather nebulous, and it is worth making this definition somewhat more precise. To accomplish this, we focus our attention on the coefficients a_i.

We define a **polynomial** over $\mathcal{F}$ to be an (infinite) sequence of scalars

$$p = \{a_0, a_1, a_2, \ldots\}$$

such that $a_n = 0$ for all but finitely many n. The scalars a_i are called the **coefficients** of the polynomial. If

$$q = \{b_0, b_1, b_2, \ldots\}$$

is another polynomial in $\mathcal{F}$, then p = q if and only if $a_i = b_i$ for every i. As we did for vector n-tuples, we define the addition of two polynomials p and q by

$$p + q = \{a_0 + b_0, a_1 + b_1, \ldots\} \ .$$

Furthermore, we now also define the multiplication of p and q by

$$pq = \{c_0, c_1, c_2, \ldots\}$$

where

$$c_k = \sum_{i+j=k} a_ib_j = \sum_{t=0}^{k} a_tb_{k-t} = a_0b_k + a_1b_{k-1} + \cdots + a_kb_0 \ .$$

Since p and q have a finite number of nonzero terms, so do both p + q and pq, and hence both p + q and pq are also polynomials.

We claim that the set of all polynomials over $\mathcal{F}$ forms a ring. Indeed, if we define the zero polynomial to be the sequence {0, 0, . . . }, and the negative of any polynomial $\{a_0, a_1, \ldots\}$ to be the polynomial $\{-a_0, -a_1, \ldots\}$, then axioms (R1) – (R6) for a ring given in Section 1.4 are clearly satisfied. As to axiom (R7), let $p = \{a_0, a_1, \ldots\}$, $q = \{b_0, b_1, \ldots\}$ and $r = \{c_0, c_1, \ldots\}$. Then the k*th* coefficient of (pq)r is the sum (using the associative property of $\mathcal{F}$)

$$\sum_{i+j=k}\left(\sum_{m+n=i} a_m b_n\right)c_j = \sum_{m+n+j=k}(a_m b_n)c_j = \sum_{m+n+j=k} a_m(b_n c_j)$$
$$= \sum_{m+i=k} a_m\left(\sum_{n+j=i} b_n c_j\right) .$$

But this last expression is just the k*th* coefficient of p(qr). Finally, to prove axiom (R8), we use the distributive property of $\mathcal{F}$ to see that the k*th* coefficient of p(q + r) is

$$\sum_{i+j=k} a_i(b_j + c_j) = \sum_{i+j=k} a_i b_j + \sum_{i+j=k} a_i c_j \quad .$$

Again we see that this last expression is just the k*th* coefficient of pq + pr. Similarly, it is easy to see that (p + q)r = pr + qr.

It should be clear that the ring of polynomials is commutative, and that if 1 is the unit element of $\mathcal{F}$, then {1, 0, 0, . . . } is a unit element for the ring of polynomials. However, since an arbitrary polynomial does not have a multiplicative inverse, the ring of polynomials does not form a field (see Theorem 6.2, Corollary 3 below).

Example 6.1 Consider the polynomials

$$\begin{aligned} p &= \{0, 1, 0, 0, \ldots\} \\ q &= \{1, 2, -1, 0, \ldots\} . \end{aligned}$$

Then

$$p + q = \{1, 3, -1, 0, \ldots\}$$

and

$$\begin{aligned} pq = \{&0(1), 0(2) + 1(1), 0(-1) + 1(2) + 0(1), \\ &0(0) + 1(-1) + 0(2) + 0(1), \ldots\} \end{aligned}$$

$$= \{0, 1, 2, -1, 0, \ldots\} \ . \ //$$

Since the reader probably thought he (or she) already knew what a polynomial was, and since our definition may not be what it was he (or she) had in mind, what we will do now is relate our formal definition to our earlier elementary experience with polynomials. We will explain shortly why we are going through all of this apparently complicated formalism.

Given any element $a \in \mathcal{F}$, we associate a polynomial a' defined by

$$a' = \{a, 0, 0, \ldots\} .$$

This is clearly a one-to-one mapping of $\mathcal{F}$ into the set of all polynomials with coefficients in $\mathcal{F}$. We also note that if $a, b \in \mathcal{F}$, then $a' = \{a, 0, \ldots\}$ and $b' = \{b, 0, \ldots\}$ so that

$$(a + b)' = \{a + b, 0, \ldots\} = a' + b'$$

and

$$(ab)' = \{ab, 0, \ldots\} = a'\ b' .$$

If $\mathcal{F}'$ denotes the set of all polynomials a' obtained in this manner, then $\mathcal{F}'$ is a field isomorphic to $\mathcal{F}$. Because of this isomorphism, we shall identify the elements of $\mathcal{F}$ with their corresponding polynomials, and write $a = \{a, 0, \ldots\}$.

Now let the symbol x denote the polynomial $\{0, 1, 0, 0, \ldots\}$. We call the symbol x an **indeterminate**. Applying our definition of polynomial multiplication, we see that $x^2 = \{0, 0, 1, 0, \ldots\}$ and, in general,

$$x^n = \{0, \ldots, 0, 1, 0, \ldots\}$$

where the 1 is in the n*th* position (remember that we start our numbering with 0). We also see that for any $a \in \mathcal{F}$ we have (applying our multiplication rule)

$$ax^n = \{a, 0, \ldots\}\{0, \ldots, 1, 0, \ldots\} = \{0, \ldots, a, 0, \ldots\} .$$

This means that an arbitrary polynomial $p = \{a_0, a_1, \ldots, a_n, 0, \ldots\}$ can be uniquely expressed in the familiar form

$$p = a_0 + a_1x + a_2x^2 + \cdots + a_nx^n .$$

This discussion has now established a precise meaning to the term "expression" used at the beginning of this chapter. We will denote the commutative ring of all polynomials over $\mathcal{F}$ by $\mathcal{F}[x]$. In summary, we see that that while a polynomial was actually defined as a sequence, we showed that any polynomial $p = \{a_0, a_1, \ldots\} \in \mathcal{F}[x]$ could be uniquely expressed in terms of the indeterminate x as

$$p = a_0 + a_1x + \cdots + a_nx^n = \sum_{i=1}^{n} a_ix^i \quad .$$

Now suppose we are given both a polynomial $p = \sum a_ix^i \in \mathcal{F}[x]$ and any element $c \in \mathcal{F}$. We define the element $p(c) \in \mathcal{F}$ by $p(c) = \sum a_ic^i$. In other words, given a polynomial $p \in \mathcal{F}[x]$, the **polynomial function** p(x) is the mapping from $\mathcal{F}$ to $\mathcal{F}$ that takes $c \in \mathcal{F}$ into the element $p(c) \in \mathcal{F}$. We call p(c)

the **value** of the polynomial p when c is **substituted** for x. Because of this, a polynomial $p \in \mathcal{F}[x]$ is frequently denoted by p(x).

The reason for this apparently complicated technical definition is that it is possible for two different polynomials in $\mathcal{F}[x]$ to result in the same polynomial function (see Exercise 6.1.1).

Theorem 6.1 Suppose $p, q \in \mathcal{F}[x]$ and $c \in \mathcal{F}$. Then

(a) $(p \pm q)(c) = p(c) \pm q(c)$.

(b) $(pq)(c) = p(c)q(c)$.

Proof (a) Writing $p = a_0 + a_1x + \cdots + a_mx^m$ and $q = b_0 + b_1x + \cdots + b_nx^n$ we have

$$\begin{aligned}(p \pm q)(c) &= (a_0 \pm b_0) + (a_1 \pm b_1)c + (a_2 \pm b_2)c^2 + \cdots \\ &= (a_0 + a_1c + a_2c^2 + \cdots) \pm (b_0 + b_1c + b_2c^2 + \cdots) \\ &= p(c) \pm q(c) \ .\end{aligned}$$

(b) Using p and q from part (a) and the definition of pq, we have

$$\begin{aligned}(pq)(c) &= a_0b_0 + (a_0b_1 + a_1b_0)c + (a_0b_2 + a_1b_1 + a_2b_0)c^2 + \cdots \\ &= (a_0 + a_1c + a_2c^2 + \cdots)(b_0 + b_1c + b_2c^2 + \cdots) \\ &= p(c)q(c) \ . \ \blacksquare\end{aligned}$$

It should now be clear that the definitions given above for the algebraic properties of polynomials in terms of sequences are just those that we all learned in high school for adding and multiplying polynomials together. It should be easy for the reader to show that Example 6.1 may be repeated in terms of our elementary notion of polynomial addition and multiplication.

If $p = a_0 + a_1x + \cdots + a_nx^n \neq 0$ and $a_n \neq 0$, then we say that the **degree** of the polynomial p is n, and write we write deg p = n. The term a_n is called the **leading** coefficient of the polynomial, and if $a_n = 1$, then the polynomial is said to be **monic**. If deg p = 0, then p is said to be **constant**. By convention, the degree of the zero polynomial is not defined.

Theorem 6.2 Suppose $p, q \in \mathcal{F}[x]$ are nonzero. Then

(a) $\deg(p + q) \leq \max\{\deg p, \deg q\}$ (where $p + q \neq 0$).

(b) $\deg(pq) = \deg p + \deg q$.

Proof (a) Let $p = a_0 + a_1x + \cdots + a_mx^m$ and $q = b_0 + b_1x + \cdots + b_nx^n$ where $a_m, b_n \neq 0$. Then

$$p + q = (a_0 + b_0) + (a_1 + b_1)x + \cdots + (a_k + b_k)x^k$$

where $k = \max\{m, n\}$. Therefore $\deg(p + q) \leq \max\{\deg p, \deg q\}$ where the inequality follows since $a_k + b_k$ could equal 0.

(b) From the definition of pq, we see that (using the same p and q from part (a)) the k*th* term is $c_k x^k$ where $c_k = \sum_{i+j=k} a_i b_j$. But if $k > m + n$, then necessarily either a_i or b_j is zero, and therefore $c_k = 0$ for $k > m + n$. Since $\mathcal{F}$ is a field and therefore also a division ring, it follows that $a_m, b_n \neq 0$ implies $a_m b_n \neq 0$ (if $a_m b_n = 0$, then multiplying from the left by a_m^{-1} says that $b_n = 0$, a contradiction). Thus $\deg pq = m + n = \deg p + \deg q$. ∎

Corollary 1 If $p, q \in \mathcal{F}[x]$ are both nonzero, then $\deg p \leq \deg pq$.

Proof Since p and q are nonzero, $\deg p \geq 0$ and $\deg q \geq 0$, and hence $\deg p \leq \deg p + \deg q = \deg pq$. ∎

Corollary 2 If $p, q, r \in \mathcal{F}[x]$, then

(a) $pq = 0$ implies that either $p = 0$ or $q = 0$.

(b) If $pq = rq$ where $q \neq 0$, then $p = r$.

Proof (a) If $p \neq 0$ and $q \neq 0$, then $\deg pq \geq 0$ implies that $pq \neq 0$.

(b) Since $\mathcal{F}[x]$ is a ring, we see that if $pq = rq$, then $(p - r)q = 0$. But $q \neq 0$, so that by (a) we must have $p - r = 0$, or $p = r$. ∎

We note that part (a) of this corollary shows that $\mathcal{F}[x]$ has no zero divisors, and hence $\mathcal{F}[x]$ forms an integral domain (see Section 1.5).

Corollary 3 Let $p \neq 0$ be an element of $\mathcal{F}[x]$. Then there exists $q \in \mathcal{F}[x]$ such that $pq = 1$ if and only if $\deg p = 0$, and hence $\mathcal{F}[x]$ is not a field.

Proof If $\deg p = 0$ then $p = a \in \mathcal{F}$ with $a \neq 0$, and thus there exists $q = a^{-1} \in \mathcal{F}$ (hence $q = a^{-1} \in \mathcal{F}[x]$) such that $pq = aa^{-1} = 1$. On the other hand, if $pq = 1$ we have

$$0 = \deg 1 = \deg pq = \deg p + \deg q \ .$$

But $p \neq 0$ implies that $\deg p \geq 0$, and if $q \neq 0$, we must also have $\deg q \geq 0$. It then follows that $\deg p = 0$. ∎

We now prove a very useful and fundamental result known as the division algorithm. Essentially, we will show the existence of the polynomial quotient f/g (although technically this symbol as of yet has no meaning). The polyno-

mials q and r defined in the following theorem are called the **quotient** and **remainder** respectively. After the proof, we will give an example which shows that this is just the "long division" we all learned in elementary school.

Theorem 6.3 (**Division Algorithm**) Given f, g $\in \mathcal{F}$[x] with g $\neq$ 0, there exist unique polynomials q, r $\in \mathcal{F}$[x] such that

$$f = qg + r$$

where either r = 0 or deg r < deg g.

Proof The basic idea of the proof is to consider all possible degrees for the polynomials f and g, and show that the theorem can be satisfied in each case. After proving the existence of the polynomials q and r, we shall prove their uniqueness.

If f = 0 we simply choose q = r = 0. Now suppose that

$$f = a_0 + a_1x + \cdots + a_mx^m$$
$$g = b_0 + b_1x + \cdots + b_nx^n$$

where a_m, $b_n \neq 0$. If m = n = 0, then by Corollary 3 of Theorem 6.2, there exists $g^{-1} \in \mathcal{F}$[x] such that $g^{-1}g = 1$, and therefore $f = f(g^{-1}g) = (fg^{-1})g + 0$ satisfies our requirements. Next, if m = 0 and n > 0, then we may write f = 0g + f with deg r = deg f < deg g.

We now assume that m > 0 and proceed by induction on m. In other words, we assume that q and r can be found for all polynomials f with deg f $\leq$ m − 1 and proceed to construct new polynomials q and r for deg f = m. First note that if n > m we may again take f = 0g + f to satisfy the theorem. Thus we need only consider the case of n $\leq$ m.

Define the polynomial

$$f_1 = f - (a_m/b_n)x^{m-n}g \ .$$

Then the coefficient of the x^m term in f_1 is 0 (it cancels out on the right hand side), and hence deg $f_1 \leq$ m − 1. Therefore, by our induction hypothesis, there exist polynomials q_1 and r_1 in $\mathcal{F}$[x] with either $r_1 = 0$ or deg r_1 < deg g such that $f_1 = q_1g + r_1$. Substituting the definition of f_1 in this equation yields

$$f = [(a_m/b_n)x^{m-n} + q_1]g + r_1 \ .$$

If we define $r = r_1$ and $q = (a_m/b_n)x^{m-n} + q_1$, we see that $f = qg + r$ where either $r = 0$ or $\deg r < \deg g$. This proves the existence of the polynomials q and r, and all that remains is to prove their uniqueness.

Suppose that

$$f = qg + r = q'g + r'$$

where both r and r' satisfy the theorem, and assume that $r \neq r'$. Then

$$r - r' = (q' - q)g \neq 0$$

where

$$\deg(r - r') < \deg g$$

by Theorem 6.2(a). On the other hand, from Theorem 6.2(b) we see that

$$\deg(r - r') = \deg[(q' - q)g] = \deg(q' - q) + \deg g \geq \deg g \ .$$

This contradiction shows that in fact $r' = r$. We now have $(q' - q)g = 0$ with $g \neq 0$, and hence by Corollary 2(a) of Theorem 6.2, we have $q' - q = 0$, and therefore $q' = q$. ∎

Let us give an example of the division algorithm that should clarify what was done in the theorem.

Example 6.2 Consider the polynomials

$$f = 2x^4 + x^2 - x + 1$$
$$g = 2x - 1 \ .$$

Following the proof of Theorem 6.3 we have

$$f_1 = f - x^3 g = x^3 + x^2 - x + 1 \ .$$

Now let

$$f_2 = f_1 - (1/2)x^2 g = (3/2)x^2 - x + 1 \ .$$

Again, we let

$$f_3 = f_2 - (3/4)xg = (-1/4)x + 1$$

so that

$$f_4 = f_3 + (1/8)g = 7/8 \ .$$

Since deg(7/8) < deg g, we are finished with the division. Combining the above polynomials we see that

$$f = [x^3 + (1/2)x^2 + (3/4)x - (1/8)]g + f_4$$

and therefore

$$q = x^3 + (1/2)x^2 + (3/4)x - (1/8)$$
$$r = 7/8 \ .$$

This may also be written out in a more familiar form as

$$\begin{array}{r} x^3 + (1/2)x^2 + (3/4)x - (1/8) \\ 2x-1\overline{)\,2x^4 \quad + \quad x^2 - \quad x+1} \\ \underline{2x^4 - \quad x^3} \\ \quad x^3 + \quad x^2 - \quad x+1 \\ \quad \underline{x^3 - (1/2)x^2} \\ \quad\quad (3/2)x^2 - \quad x+1 \\ \quad\quad \underline{(3/2)x^2 - (3/4)x} \\ \quad\quad\quad -(1/4)x+1 \\ \quad\quad\quad \underline{-(1/4)x+(1/8)} \\ \quad\quad\quad\quad 7/8 \end{array}$$

It should be noted that at each step in the division, we eliminated the highest remaining power of f by subtracting the appropriate multiple of g. //

Exercises

1. Let $\mathcal{F} = \{0, 1\}$ be the field consisting of only two elements, and define addition and multiplication on these elements in the obvious way (see Exercise 1.5.17). Show that the distinct polynomials $x^2 - x$ and 0 define the same polynomial function.

2. Use the division algorithm to find the quotient and remainder when $f = 2x^4 - x^3 + x - 1 \in \mathbb{R}[x]$ is divided by $g = 3x^3 - x^2 + 3 \in \mathbb{R}[x]$.

3. Consider the polynomials p = {2, 0, 1, 1} and q = {1, 1, −1} over $\mathbb{R}$. Evaluate the product pq by applying the definition. Show that this yields the same result as directly multiplying together the polynomial functions p and q.

4. Given a polynomial $p = a_n x^n + \cdots + a_1 x + a_0$, we define its **formal derivative** to be the polynomial $Dp = na_n x^{n-1} + \cdots + 2a_2 x + a_1$. In other words, $D: \mathcal{F}[x] \to \mathcal{F}[x]$ is a **differentiation operator**. Prove $D(p + q) = Dp + Dq$ and $D(pq) = p(Dq) + (Dp)q$.

5. Find the remainder when $ix^9 + 3x^7 + x^6 - 2ix + 1 \in \mathbb{C}[x]$ is divided by $x + i \in \mathbb{C}[x]$.

6.2 FACTORIZATION OF POLYNOMIALS

If f(x) is a polynomial in $\mathcal{F}[x]$, then $c \in \mathcal{F}$ is said to be a **zero** (or **root**) of f if f(c) = 0. We shall also sometimes say that c is a **solution** of the polynomial equation f(x) = 0. We will see that information about the roots of a polynomial plays an extremely important role throughout much of the remainder of this text.

If $f, g \in \mathcal{F}[x]$ and $g \neq 0$, then we say that f is **divisible** by g (or g **divides** f) over $\mathcal{F}$ if f = qg for some $q \in \mathcal{F}[x]$. In other words, f is divisible by g if the remainder in the division of f by g is zero. In this case we also say that g is a **factor** of f (over $\mathcal{F}$). It is standard notation to write g|f when we wish to say that g divides f, or to write $g \nmid f$ when g does not divide f.

The next theorem is known as the remainder theorem, and its corollary is known as the factor theorem.

Theorem 6.4 (**Remainder Theorem**) Suppose $f \in \mathcal{F}[x]$ and $c \in \mathcal{F}$. Then the remainder in the division of f by x − c is f(c). In other words,

$$f(x) = (x - c)q + f(c) \ .$$

Proof We see from the division algorithm that f = (x − c)q + r where either r = 0 or deg r < deg (x − c) = 1, and hence either r = 0 or deg r = 0 (in which case $r \in \mathcal{F}$). In either case, we may substitute c for x to obtain

$$f(c) = (c - c)q(c) + r = r \ . \ \blacksquare$$

Corollary (**Factor Theorem**) If $f \in \mathcal{F}[x]$ and $c \in \mathcal{F}$, then $x - c$ is a factor of f if and only if $f(c) = 0$.

Proof Rephrasing the statement of the corollary as $f = q(x - c)$ if and only if $f(c) = 0$, it is clear that this follows directly from the theorem. ∎

Example 6.3 If we divide $f = x^3 - 5x^2 + 7x$ by $g = x - 2$, we obtain $q = x^2 - 3x + 1$ and $r = 2$. It is also easy see that $f(2) = 8 - 5(4) + 7(2) = 2$ as it should according to Theorem 6.4. //

Let R be a commutative ring with unit element. An element $u \in R$ is called a **unit** (*not* a unit element) if there exists $r \in R$ such that $ur = 1$. Other ways to say this are that u divides 1, or that a unit is an element whose inverse is also in the ring. We leave it to the reader to show that $u \in R$ is a unit if and only if it is a factor of every element of R (see Exercise 6.2.1). An element $p \in R$ that is neither zero nor a unit is said to be **prime** if $p = ab$ implies that either a or b is a unit. Thus a prime element is one that can not be factored in a nontrivial way.

Example 6.4 Since for any integer $n \neq \pm 1$ the number $1/n$ is not an integer, it should be clear that the ring of integers $\mathbb{Z}$ has only the units 1 and -1. On the other hand, if $\mathcal{F}$ is any field and $a \in \mathcal{F}$, then a^{-1} is also a member of $\mathcal{F}$, and hence any nonzero element of a field is a unit. In particular, the units of the ring $\mathcal{F}[x]$ are just the polynomials of degree zero (i.e., the nonzero constant polynomials).

If we consider the ring of integers $\mathbb{Z}$, then a number $p \in \mathbb{Z}$ with $p \neq \pm 1$ or 0 will be prime if the only divisors of p are ± 1 and $\pm p$. However, if we consider the field $\mathbb{R}$, then any nonzero element of $\mathbb{R}$ (i.e., any nonzero real number) is a unit, and hence the notion of a prime real number is not very useful. //

In the particular case of $R = \mathcal{F}[x]$, a prime polynomial is frequently called an **irreducible** polynomial. A polynomial that is not irreducible is said to be **reducible**. Two polynomials $f, g \in \mathcal{F}[x]$ are said to be **associates** if $f = cg$ for some nonzero $c \in \mathcal{F}$, and in general, two elements $a, b \in R$ are said to be **associates** if $a = ub$ where u is a unit of R. We leave it to the reader to show that this defines an equivalence relation on a commutative ring with unit element (Exercise 6.2.3).

It should be clear that any nonzero polynomial has exactly one monic polynomial as an associate since we can always write

$$a_0 + a_1x + \cdots + a_nx^n = a_n(a_0a_n^{-1} + a_1a_n^{-1}x + \cdots + x^n) .$$

It should also be clear that any polynomial f with deg f ≥ 1 has its associates and the set of nonzero constant polynomials as divisors. Thus we see that a nonzero polynomial f is prime if and only if f = gh implies that either g or h is of degree zero, and hence the other is an associate of f.

It is important to realize that whether or not a polynomial is prime depends on the particular field $\mathcal{F}$. For example, since $x^2 - 2 = (x - \sqrt{2})(x + \sqrt{2})$, we see that $x^2 - 2$ is prime over $\mathbb{Q}$, but is not prime over $\mathbb{R}$.

Returning to our commutative ring R with unit element, we say that an element $d \in R$ is a **greatest common divisor** (frequently denoted simply by gcd) of the elements $a_1, \ldots, a_n \in R$ if

(1) $d|a_i$ for every $i = 1, \ldots, n$ (i.e., d is a **common divisor** of the a_i);
(2) If $c \in R$ is such that $c|a_i$ for every $i = 1, \ldots, n$, then $c|d$.

Two distinct elements $a, b \in R$ are said to be **relatively prime** if their greatest common divisor is a unit of R. Note that we have referred to *a* greatest common divisor, implying that there may be more than one. This is only true in a certain sense as the next theorem shows.

Theorem 6.5 Let $f_1, \ldots, f_n$ be nonzero polynomials in $\mathcal{F}[x]$. Then there exists at least one greatest common divisor d of the set $\{f_1, \ldots, f_n\}$. Moreover, this greatest common divisor is unique up to a unit factor, and can be expressed in the form $d = \sum_{i=1}^{n} h_i f_i$ for some set of polynomials $h_i \in \mathcal{F}[x]$.

Proof Consider the set S of all polynomials in $\mathcal{F}[x]$ of the form $g_1 f_1 + \cdots + g_n f_n$ where each g_i is an *arbitrary* element of $\mathcal{F}[x]$. Then in particular, each f_i is an element of S, and in addition, if $p \in S$ and $q \in \mathcal{F}[x]$, then $pq \in S$. Let D be the set of degrees of all nonzero polynomials in S. Then D is just a collection of nonnegative integers, and hence by the well-ordering principle (Section 0.5), D has a least element α. This means that there exists a nonzero polynomial $d = h_1 f_1 + \cdots + h_n f_n \in S$ such that $\alpha = \deg d \leq \deg c$ for all nonzero polynomials $c \in S$. We first show that $d|f_i$ for every $i = 1, \ldots, n$.

By the division algorithm, for each $i = 1, \ldots, n$ we have $f_i = q_i d + r_i$ where either $r_i = 0$ or $\deg r_i < \deg d$. But $r_i = f_i - q_i d \in S$ so that if $r_i \neq 0$, we would have $\deg r_i < \deg d$ which contradicts the definition of d. Therefore we must have $r_i = 0$ and hence $f_i = q_i d$ so that $d|f_i$ for every $i = 1, \ldots, n$. While this shows that d is a common divisor of $\{f_i\}$, we must show that it is in fact a greatest common divisor.

If c is any other common divisor of $\{f_1, \ldots, f_n\}$, then by definition there exist polynomials g_i such that $f_i = g_i c$ for each $i = 1, \ldots, n$. But then

$$d = \sum h_i f_i = \sum h_i(g_i c) = (\sum h_i g_i)c$$

so that $c|d$. This proves that d is a greatest common divisor.

Now suppose d' is another greatest common divisor of the set $\{f_1, \ldots, f_n\}$. Then by definition of greatest common divisor we must have both $d|d'$ and $d'|d$, so that $d' = ud$ and $d = vd'$ for some polynomials $u, v \in \mathcal{F}[x]$. Multiplying the second of these equations by u, we see that $uvd' = ud = d'$, and therefore $d'(1 - uv) = 0$. By Corollary 2(a) of Theorem 6.2, we then have $1 - uv = 0$ so that $uv = 1$ and hence u and v are units. (Alternatively, the fact that $\deg d = \deg d'$ implies that u and v must be of degree zero, and hence are units.) ∎

What we have shown is that a gcd exists, and is unique up to its associates. Therefore, if we restrict ourselves to *monic* greatest common divisors, then we have proved the existence of a unique gcd.

Corollary 1 Let $q_1, \ldots, q_n \in \mathcal{F}[x]$ be relatively prime (i.e., they have no common divisors other than units). Then there exist elements $h_1, \ldots, h_n \in \mathcal{F}[x]$ such that $h_1q_1 + \cdots + h_nq_n = 1$.

Proof This is an obvious special case of Theorem 6.5. ∎

Corollary 2 Suppose $f, g, p \in \mathcal{F}[x]$ where p is prime and $p|fg$. Then either $p|f$ or $p|g$.

Proof Since p is prime, its only divisors are units and its associates. Therefore, if we assume that $p \nmid f$, then the only greatest common divisor of p and f is a unit, and thus p and f are relatively prime. Applying Theorem 6.5, we may write $up + vf = 1$ for some $u, v \in \mathcal{F}[x]$, and hence multiplying by g yields $pug + fgv = g$. But $p|fg$ so that $fg = qp$ for some $q \in \mathcal{F}[x]$, and thus we have $p(ug + qv) = g$ so that $p|g$. It is obvious that had we started with the assumption that $p \nmid g$, we would have found that $p|f$. ∎

By choosing $f = f_1$ and $g = f_2 \cdots f_n$ in Corollary 2 we have the following obvious generalization.

Corollary 2′ If $p, f_1, f_2, \ldots, f_n \in \mathcal{F}[x]$ where p is prime and $p|f_1f_2 \cdots f_n$, then $p|f_i$ for some $i = 1, \ldots, n$.

While Theorem 6.5 proves the existence of a greatest common divisor, it is not of any help in actually computing one. Given two polynomials, we can find their gcd by a procedure known as the Euclidean algorithm (compare

Section 0.7). This approach, illustrated in the next example, is also an alternative proof of Theorem 6.5, the general case as stated in the theorem following by induction.

Example 6.5 (**Euclidean algorithm**) Suppose f, g $\in \mathcal{F}[x]$ and f $\neq$ 0. We show the existence of a unique monic polynomial d $\in \mathcal{F}[x]$ such that

(1) d|f and d|g.

(2) If c $\in \mathcal{F}[x]$ is such that c|f and c|g, then c|d.

First note that if g = 0 and a_m is the leading coefficient of f, then the monic polynomial $d = a_m^{-1}f$ satisfies both requirements (1) and (2). Now assume that g $\neq$ 0 also. By the division algorithm, there exist *unique* polynomials q_1 and r_1 such that

$$f = gq_1 + r_1$$

with either $r_1 = 0$ or deg r_1 < deg g. If $r_1 = 0$, then g|f and d = g satisfies (1) and (2). If $r_1 \neq 0$, then we apply the division algorithm again to obtain polynomials q_2 and r_2 such that

$$g = r_1q_2 + r_2 \ .$$

If $r_2 = 0$, then $r_1|g$ which implies that $r_1|f$, and thus $d = r_1$ is a common divisor. (It still remains to be shown that this d is a *greatest* common divisor.) If $r_2 \neq 0$, then we continue the process, thus obtaining the following progression:

$$\begin{aligned}
f &= gq_1 + r_1 && \deg r_1 < \deg g \\
g &= r_1q_2 + r_2 && \deg r_2 < \deg r_1 \\
r_1 &= r_2q_3 + r_3 && \deg r_3 < \deg r_2 \\
&\vdots && \vdots \\
r_{k-2} &= r_{k-1}q_k + r_k && \deg r_k < \deg r_{k-1} \\
r_{k-1} &= r_kq_{k+1}
\end{aligned}$$

This progression must terminate as shown since the degree of any polynomial is a positive integer and deg r_1 > deg r_2 > · · · > deg $r_k \geq 0$. Letting r_k be the last nonzero remainder, we claim that $r_k = d$.

To see this, first note that $r_k|r_{k-1}$ since $r_{k-1} = r_kq_{k+1}$. Next, we see that

$$r_{k-2} = r_{k-1}\, q_k + r_k = r_k\ q_{k+1}\, q_k + r_k$$

and therefore $r_k|r_{k-2}$. Continuing this procedure, we find that $r_k|r_{k-1}$, $r_k|r_{k-2}$, . . . , $r_k|r_1$, $r_k|g$ and finally $r_k|f$. This shows that $d = r_k$ satisfies (1). Now suppose that c|f and c|g. Then, since $r_1 = f - gq_1$ we must have $c|r_1$. Similarly $r_2 = g - r_1q_2$ so that $c|r_2$. Continuing this process, it is clear that we eventually

arrive at the conclusion that $c|r_k$, thus proving (2) for the choice $d = r_k$. Finally, if r is the leading coefficient of r_k, then $r^{-1}r_k$ is a monic polynomial satisfying (1) and (2), and its uniqueness follows exactly as in the proof of Theorem 6.5. //

Example 6.6 As a specific illustration of the preceding example, consider the polynomials $f = x^4 - x^3 - x^2 + 1$ and $g = x^3 - 1$ over the field $\mathbb{Q}$. Dividing f by g we obtain

$$x^4 - x^3 - x^2 + 1 = (x^3 - 1)(x - 1) + (-x^2 + x) .$$

Now divide g by $r_1 = -x^2 + x$ to obtain

$$x^3 - 1 = (-x^2 + x)(-x - 1) + (x - 1) .$$

Lastly, we divide r_1 by $r_2 = x - 1$ to find

$$-x^2 + x = (x - 1)(-x)$$

and therefore the gcd of f and g is $x - 1$. //

Our next very important result is known as the unique factorization theorem. Recall that by definition, a prime polynomial is not a unit, and thus has positive degree.

Theorem 6.6 (**Unique Factorization Theorem**) Every nonzero element $f \in \mathcal{F}[x]$ is either a unit, or is expressible as a unique (up to associates) finite product of prime elements.

Proof We first show that $f \in \mathcal{F}[x]$ is expressible as a product of prime polynomials. Afterwards we will prove uniqueness. Our approach is by induction on deg f; in other words, we assume that deg $f > 1$ (if deg $f = 0$ then f is a unit, and if deg $f = 1$ the theorem is obvious), and suppose that the theorem is true for all $g \in \mathcal{F}[x]$ with deg $g <$ deg f. We will show that the theorem is true for f.

Assume f is reducible (or else there is nothing to prove) so that $f = pq$ where neither p nor q is a unit. By Theorem 6.2(b) we have deg $p <$ deg $p +$ deg $q =$ deg f, and similarly deg $q <$ deg f. Therefore, by our induction hypothesis, both p and q can be written as a finite product of prime elements in $\mathcal{F}[x]$, and hence the same is true of $f = pq$.

To prove the uniqueness of the product, assume that

$$f = p_1 p_2 \cdots p_n = q_1 q_2 \cdots q_m$$

where each of the p_i and q_j are prime. Since $p_1 | p_1 p_2 \cdots p_n$, it follows that $p_1 | q_1 q_2 \cdots q_m$. By Corollary 2′ of Theorem 6.5, it then follows that $p_1 | q_j$ for some $j = 1, \ldots, m$. But since both p_1 and q_j are prime and $p_1 | q_j$, they must be associates, and hence $q_j = u_1 p_1$ where u_1 is a unit in $\mathcal{F}[x]$. This means that

$$p_1 p_2 \cdots p_n = q_1 q_2 \cdots q_m = q_1 q_2 \cdots q_{j-1} u_1 p_1 q_{j+1} \cdots q_m .$$

Cancelling p_1 from both sides of this equation (using Theorem 6.2, Corollary 2) results in

$$p_2 \cdots p_n = u_1 q_1 \cdots q_{j-1} q_{j+1} \cdots q_m .$$

Repeating this argument, we next eliminate p_2 and one of the remaining factors on the right. Continuing in this manner, we pairwise eliminate one of the p_i and one of the q_k with each operation, always replacing a q_k with a corresponding u_k. But the primes on one side of this equation can not be eliminated before those on the other side because this would imply that a product of prime polynomials was equal to 1 which is impossible. Therefore $n = m$, and the expansion of f as a product of prime elements must be unique up to an associate. ∎

Note that the expansion proved in this theorem for $f \in \mathcal{F}[x]$ is completely unique (except for order) if we require that the prime polynomials be monic.

Example 6.7 Consider the polynomial $p = 3x^4 - 3x^2 - 6 \in \mathcal{F}[x]$. Using the fields $\mathbb{Q}$, $\mathbb{R}$ and $\mathbb{C}$ we can factor p three different ways depending on $\mathcal{F}$:

$$\begin{aligned} 3(x^2 - 2)(x^2 + 1) &\quad \text{in } \mathbb{Q}[x] \\ 3(x + \sqrt{2})(x - \sqrt{2})(x^2 + 1) &\quad \text{in } \mathbb{R}[x] \\ 3(x + \sqrt{2})(x - \sqrt{2})(x + i)(x - i) &\quad \text{in } \mathbb{C}[x] \end{aligned}$$

In each case, p is a product of prime polynomials relative to the appropriate field. //

Exercises

1. Let R be a commutative ring with unit element. Show that $u \in R$ is a unit of R if and only if u is a factor of every element of R.

2. Let R be an integral domain with unit element and suppose it is true that $a|b$ and $b|a$ for some $a, b \in R$. Show that $a = ub$ where u is a unit in R.

3. Show that the property of being associates defines an equivalence relation on a commutative ring with unit element.

4. Let $\mathcal{F}$ be an arbitrary field, and suppose $p \in \mathcal{F}[x]$ is of degree ≤ 3. Prove that p is prime in $\mathcal{F}[x]$ if and only if p is either of degree 1, or has no zeros in $\mathcal{F}$. Give an example to show that this result is not true if $\deg p > 3$.

5. Factor the following polynomials into their prime factors in both $\mathbb{R}[x]$ and $\mathbb{Q}[x]$:
   (a) $2x^3 - x^2 + x + 1$.
   (b) $3x^3 + 2x^2 - 4x + 1$.
   (c) $x^6 + 1$.
   (d) $x^4 + 16$.

6. Let $\mathcal{F} = \{0, 1\}$ be the field consisting of only two elements, and define addition and multiplication on $\mathcal{F}$ in the obvious way. Factor the following polynomials into primes in $\mathcal{F}[x]$:
   (a) $x^2 + x + 1$.
   (b) $x^3 + 1$.
   (c) $x^4 + x^2 + 1$.
   (d) $x^4 + 1$.

7. Let $\mathcal{F}$ be as in the previous problem. Find the greatest common divisor of $x^3 + x^2 + x + 1$ and $x^5 + x^4 + x^3 + x^2 + x + 1$ over $\mathcal{F}[x]$.

8. Find the greatest common divisor of the following pairs of polynomials over $\mathbb{R}[x]$. Express your result in the form defined in Theorem 6.5.
   (a) $4x^3 + 2x^2 - 2x - 1$ and $2x^3 - x^2 + x + 1$.
   (b) $x^3 - x + 1$ and $2x^4 + x^2 + x - 5$.
   (c) $x^4 + 3x^2 + 2$ and $x^5 - x$.
   (d) $x^3 + x^2 - 2x - 2$ and $x^4 - 2x^3 + 3x^2 - 6x$.

9. Use the remainder theorem to find the remainder when $2x^5 - 3x^3 + 2x + 1 \in \mathbb{R}[x]$ is divided by:
   (a) $x - 2 \in \mathbb{R}[x]$.
   (b) $x + 3 \in \mathbb{R}[x]$.

10. (a) Is $x - 3$ a factor of $3x^3 - 9x^2 - 7x + 21$ over $\mathbb{Q}[x]$?
    (b) Is $x + 2$ a factor of $x^3 + 8x^2 + 6x - 8$ over $\mathbb{R}[x]$?
    (c) For which $k \in \mathbb{Q}$ is $x - 1$ a factor of $x^3 + 2x^2 + x + k$ over $\mathbb{Q}[x]$?
    (d) For which $k \in \mathbb{C}$ is $x + i$ a factor of $ix^9 + 3x^7 + x^6 - 2ix + k$ over $\mathbb{C}[x]$?

11. (a) Construct an example to show that the division algorithm is not true if $\mathcal{F}$ is replaced by the integral domain $\mathbb{Z}$.
    (b) Prove that if the division algorithm is true for polynomials over an integral domain D, then D must be a field.

12. Determine the monic associate of:
    (a) $2x^3 - x + 1 \in \mathbb{Q}[x]$.
    (b) $-ix^2 + x + 1 \in \mathbb{C}[x]$.

13. Let $f = a_0 + a_1x + \cdots + a_nx^n \in \mathbb{Z}[x]$ be a polynomial with integer coefficients, and suppose $r/s \in \mathbb{Q}$ is a rational root of f. Assume that r and s are relatively prime. Prove that $r|a_0$ and $s|a_n$.

6.3 POLYNOMIAL IDEALS

We now apply the formalism of Section 1.5 to polynomials. If the reader has not yet studied that section (or does not remember it), now is the time to go back and do so (again if necessary).

Example 6.8 Let A and B be ideals of $\mathcal{F}[x]$. We show that

$$A \cap B = \{f \in \mathcal{F}[x]: f \in A \text{ and } f \in B\}$$

and

$$A + B = \{f + g \in \mathcal{F}[x]: f \in A \text{ and } g \in B\}$$

are both ideals of $\mathcal{F}[x]$. Indeed, if $f, g \in A \cap B$, then $f \pm g \in A$ and $f \pm g \in B$ since A and B are ideals. Therefore $f \pm g \in A \cap B$ so that $A \cap B$ is a subgroup of $\mathcal{F}[x]$ under addition. Similarly, if $f \in A \cap B$ and $g \in \mathcal{F}[x]$, then $fg \in A$ and $fg \in B$ so that $fg \in A \cap B$, and hence $A \cap B$ is an ideal. Now suppose $f_1 + g_1, f_2 + g_2 \in A + B$. Then

$$(f_1 + g_1) \pm (f_2 + g_2) = (f_1 \pm f_2) + (g_1 \pm g_2) \in A + B$$

so that A + B is also a subgroup of $\mathcal{F}[x]$ under addition. Finally, suppose that $f_1 + g_1 \in A + B$ and $h \in \mathcal{F}[x]$. Then

$$(f_1 + g_1)h = f_1h + g_1h \in A + B$$

so that A + B is an ideal also. //

Recall that $\mathcal{F}[x]$ is not only a ring, it is in fact an integral domain (Corollary 2 of Theorem 6.2). Since $\mathcal{F}[x]$ is a ring, we see that if $p \in \mathcal{F}[x]$, then p can be used to generate the principal ideal (p) of $\mathcal{F}[x]$. Note that by construction, the ideal (p) can not contain any prime polynomials (other than associates of p) even if p itself is prime. This is because any $q \in (p)$ can be written in the form q = pr for some $r \in \mathcal{F}[x]$, and hence p|q. We will also write the principal ideal (p) in the form

$$(p) = p\mathcal{F}[x] = \{pf: f \in \mathcal{F}[x]\} .$$

Our next theorem is frequently useful when working with quotient rings (see Theorem 1.13) of the general form $\mathcal{F}[x]/(p)$.

Theorem 6.7 Suppose $p = a_0 + a_1x + \cdots + a_nx^n \in \mathcal{F}[x]$, $a_n \neq 0$, and let I = (p). Then every element of $\mathcal{F}[x]/I$ can be uniquely expressed in the form

$$I + (b_0 + b_1x + \cdots + b_{n-1}x^{n-1})$$

where $b_0, \ldots, b_{n-1} \in \mathcal{F}$.

Proof Choose any $I + f \in \mathcal{F}[x]/I$. By the division algorithm (Theorem 6.3) we can write f = pq + r for some $q, r \in \mathcal{F}[x]$ with either r = 0 or deg r < deg p. But by definition of I, we have $pq \in I$ so that

$$I + f = I + (pq + r) = I + r .$$

This shows that I + f has the form desired.

To prove the uniqueness of this representation, suppose that

$$I + (b_0 + b_1 x + \cdots + b_{n-1}x^{n-1}) = I + (c_0 + c_1 x + \cdots + c_{n-1}x^{n-1}) .$$

Then (by adding $I + (-c_0 - c_1x - \cdots - c_{n-1}x^{n-1})$ to both sides) we see that

$$(b_0 - c_0) + (b_1 - c_1)x + \cdots + (b_{n-1} - c_{n-1})x^{n-1} \in I .$$

But the degree of any nonzero polynomial in I must be greater than or equal to n = deg p (by definition of I and Theorem 6.2(b)), and therefore it follows that

$$(b_0 - c_0) + (b_1 - c_1)x + \cdots + (b_{n-1} - c_{n-1})x^{n-1} = 0 \ .$$

Since two polynomials are equal if and only if their coefficients are equal, this means that $b_i = c_i$ for every $i = 0, \ldots, n - 1$. ∎

It is an interesting fact that every ideal of $\mathcal{F}[x]$ is actually a principal ideal. We prove this in our next theorem.

Theorem 6.8 Every ideal of the ring $\mathcal{F}[x]$ is a principal ideal.

Proof Let I be any ideal of $\mathcal{F}[x]$. If $I = \{0\}$, then I is just the principal ideal (0). Now assume that $I \neq \{0\}$ and let g be any nonzero polynomial of least degree in I. (That g exists follows from the well-ordering principle. In other words, if S is the set of degrees of all polynomials in $\mathcal{F}[x]$, then S has a least element.) From the definitions it is clear that $(g) \subset I$. We now show that $I \subset (g)$ which will then prove that $I = (g)$.

By the division algorithm, for any $f \in I$ there exist polynomials $q, r \in \mathcal{F}[x]$ such that $f = gq + r$ where either $r = 0$ or $\deg r < \deg g$. Since $f \in I$ and $gq \in I$, it follows that $r = f - gq \in I$. But if $r \neq 0$ we have $\deg r < \deg g$ which contradicts the definition of g as the polynomial of least degree in I. Therefore $r = 0$ so that $f = gq \in (g)$, and hence $I \subset (g)$. ∎

Corollary Every ideal of $\mathcal{F}[x]$ is generated by a unique monic polynomial.

Proof Since every ideal of $\mathcal{F}[x]$ is principal, suppose that $(p) = (q)$ or, equivalently, $p\,\mathcal{F}[x] = q\,\mathcal{F}[x]$. Then clearly $p \in p\,\mathcal{F}[x]$ so that $p = qf_1$ for some $f_1 \in \mathcal{F}[x]$. Similarly, we see that $q = pf_2$ for some $f_2 \in \mathcal{F}[x]$ and hence

$$p = qf_1 = pf_2f_1 \ .$$

Since $\mathcal{F}[x]$ is an integral domain it follows that $f_1f_2 = 1$. But this means that f_1 and f_2 are units (i.e., constant polynomials), and hence $q = cp$ for some $c \in \mathcal{F}$. Noting that $cp\,\mathcal{F}[x]$ is the same as $p\,\mathcal{F}[x]$, we see that any ideal $p\,\mathcal{F}[x]$ may be written in the form $cp\,\mathcal{F}[x]$ for arbitrary $c \in \mathcal{F}$. By choosing c to be the inverse of the leading coefficient of p, we have shown that any ideal of $\mathcal{F}[x]$ has a unique monic generator. ∎

In Section 6.2 we discussed the greatest common divisor of a collection of polynomials. We now treat a related concept that the reader may be wondering about. If f, g $\in \mathcal{F}[x]$ then, by the **least common multiple** (or simply lcm) of f and g, we mean the polynomial m $\in \mathcal{F}[x]$ such that f|m and g|m, and if m′ $\in \mathcal{F}[x]$ is another polynomial that satisfies f|m′ and g|m′, then m|m′.

As a useful observation, note that if f, g $\in \mathcal{F}[x]$ and f|g, then g = fq for some q $\in \mathcal{F}[x]$. But (f) = f $\mathcal{F}[x]$ and hence

$$(g) = g\mathcal{F}[x] = fq\mathcal{F}[x] \subset (f) .$$

In other words, f|g implies that (g) $\subset$ (f).

Example 6.9 Let A and B be ideals of $\mathcal{F}[x]$. By Theorem 6.8 and Example 6.8 we may write A = h $\mathcal{F}[x]$ and B = k $\mathcal{F}[x]$, and also A $\cap$ B = m $\mathcal{F}[x]$ and A + B = d $\mathcal{F}[x]$. We claim that d is a greatest common divisor of h and k, and that m is a least common multiple of h and k.

To see this, first note that since h $\in$ h $\mathcal{F}[x]$ = A, it follows that h = h + 0 $\in$ A + B, and hence h = dh_1 for some $h_1 \in \mathcal{F}[x]$. Similarly, we must have k = dk_1 for some $k_1 \in \mathcal{F}[x]$. Therefore d|h and d|k so that d is a common divisor of h and k. We must show that if d′|h and d′|k, then d′|d. Now, if d′|h and d′|k, then A = (h) $\subset$ (d′) and B = (k) $\subset$ (d′). But then A + B $\subset$ (d′) because for any f + g $\in$ A + B we have f $\in$ A $\subset$ (d′) and g $\in$ B $\subset$ (d′), and therefore f + g $\in$ (d′) since (d′) is an ideal. This means that d $\in$ A + B $\subset$ (d′) so that d = d′p for some p $\in \mathcal{F}[x]$, and hence d′|d.

Now note that m $\in$ A $\cap$ B so that m $\in$ A implies m = hm_1 and m $\in$ B implies m = km_2 for some polynomials m_1, $m_2 \in \mathcal{F}[x]$. This means that h|m and k|m so that m is a common multiple of h and k. Next, note that if h|m′ then (m′) $\subset$ (h) = A, and if k|m′ then (m′) $\subset$ (k) = B. Therefore we see that m′ $\in$ (m′) $\subset$ A $\cap$ B = (m) so that m′ = mq for q $\in \mathcal{F}[x]$. Hence m|m′ and m is a least common multiple of h and k. //

Greatest common divisors and least common multiples will be of considerable use to us in the next chapter. The following important theorem relates least common multiples and greatest common divisors. Rather than prove it directly, we simply note that it follows easily from Theorem 6.10 below.

Theorem 6.9 Suppose h, k $\in \mathcal{F}[x]$ and let d and m be the greatest common divisor and least common multiple respectively of h and k. Then hk = dm.

Recall from Theorem 6.6 that any polynomial $h \in \mathcal{F}[x]$ is expressible as a unique product of prime polynomials. If h contains the prime factor $g_i \in \mathcal{F}[x]$ repeated r_i times, then we write $g_i^{r_i}$ as one of the factors of h. Therefore we may write the decomposition of h in the form $h = \prod_i g_i^{r_i}$. If $k \in \mathcal{F}[x]$ is another polynomial, then it may also be factored in the same manner as $k = \prod_j q_j^{s_j}$. In fact, we may write both h and k as a product of the *same* factors if we allow the exponent to be zero for any factor that does not appear in that expansion. In other words, we may write $h = \prod_{i=1}^{n} p_i^{r_i}$ and $k = \prod_{i=1}^{n} p_i^{s_i}$ where $r_i \geq 0$ and $s_i \geq 0$ for each $i = 1, \ldots, n$.

The next theorem contains Theorem 6.9 as an immediate and obvious corollary.

Theorem 6.10 Suppose that $h, k \in \mathcal{F}[x]$ and write

$$h = \prod_{i=1}^{n} p_i^{r_i} \qquad k = \prod_{i=1}^{n} p_i^{s_i}$$

where each $r_i \geq 0$ and each $s_i \geq 0$. For the given r_i and s_i, define the polynomials

$$\alpha = \prod_{r_i > s_i} p_i^{r_i} \qquad \beta = \prod_{r_i \leq s_i} p_i^{r_i} \qquad \gamma = \prod_{s_i < r_i} p_i^{s_i} \qquad \delta = \prod_{s_i \geq r_i} p_i^{s_i}$$

so that $h = \alpha\beta$ and $k = \gamma\delta$. Then the least common multiple m of h and k is given by

$$m = \alpha\delta = \prod_{i=1}^{n} p_i^{\max(r_i, s_i)}$$

and the greatest common divisor d is given by

$$d = \beta\gamma = \prod_{i=1}^{n} p_i^{\min(r_i, s_i)} \quad .$$

Proof Since the expressions for h and k are given in terms of the same set of prime polynomials p_i, a moment's thought should make it clear that the least common multiple is given by

$$m = \prod_{i=1}^{n} p_i^{\max(r_i, s_i)}$$

Formally, we see that $h|m$ and $k|m$, and if m' is another polynomial such that $h|m'$ and $k|m'$, then by Theorem 6.6 again we can write $m' = \prod_{i=1}^{n} p_i^{t_i}$ where we must have $t_i \geq r_i$ and $t_i \geq s_i$ for each $i = 1, \ldots, n$ in order that m' be a common multiple of h and k. But this means that $m|m'$ so that m is the least common multiple of h and k. In any case, m is exactly the same as the product $\alpha\delta$.

That the greatest common divisor is given by $\prod_{i=1}^{n} p_i^{\min(r_i, s_i)} = \beta\gamma$ follows from a similar argument. ∎

By Theorem 1.13, the quotient structure $\mathcal{F}[x]/(p)$ is a ring for any p, and we now show that it is actually a field for appropriate p.

Theorem 6.11 Suppose $p \in \mathcal{F}[x]$, and let $I = (p)$. Then $\mathcal{F}[x]/I$ is a field if and only if p is prime over $\mathcal{F}$.

Proof We first show that if p is reducible over $\mathcal{F}$, then $\mathcal{F}[x]/I$ is not a field. (This is the contrapositive to the statement that if $\mathcal{F}[x]/I$ is a field, then p must be prime.) To see this, assume that $p = ab$ where neither a nor b is a unit, and each is of degree less than that of p. From the definition of a principal ideal, we see that the degree of any polynomial in I must be greater than or equal to deg p, and hence neither a nor b can be an element of I. Since I is the zero element of $\mathcal{F}[x]/I$, we see that $I + a \neq I$ and $I + b \neq I$ must both be nonzero elements of $\mathcal{F}[x]/I$. But then

$$(I + a)(I + b) = I + ab = I + p = I$$

where we used the fact that $p \in I$. This shows that the product of two nonzero elements of $\mathcal{F}[x]/I$ yields the zero element of $\mathcal{F}[x]/I$, and thus the set of nonzero elements of $\mathcal{F}[x]/I$ is not closed under multiplication. Hence $\mathcal{F}[x]/I$ is neither a division ring nor an integral domain, so it certainly is not a field.

Conversely, suppose that p is prime. Since $\mathcal{F}[x]$ is a commutative ring, it follows that for any $a, b \in \mathcal{F}[x]$ we have that $\mathcal{F}[x]/I$ is also a commutative ring. The identity element in $\mathcal{F}[x]/I$ is easily seen to be $I + 1$ where 1 is the unit element for the field $\mathcal{F}$. Therefore, all that remains is to show the existence of a multiplicative inverse for each nonzero element in $\mathcal{F}[x]/I$.

If $I + f$ is any nonzero element in $\mathcal{F}[x]/I$, then $f \notin I$ so that $p \nmid f$. Since p is prime, its only divisors are units and its associates, and therefore the greatest common divisor of p and f is a unit (i.e., p and f are relatively prime). Applying Corollary 1 of Theorem 6.5 we see there exist $u, v \in \mathcal{F}[x]$ such that $up + vf = 1$. Then $1 - vf = up \in I$ and hence

$$I + 1 = I + up + vf = I + vf = (I + v)(I + f) .$$

This shows that $I + v$ is a multiplicative inverse of $I + f$. ∎

Exercises

1. Referring to Theorem 6.7, show that $\{I + c: c \in \mathcal{F}\}$ is a subfield of $\mathcal{F}[x]/I$ isomorphic to $\mathcal{F}$.

2. Let $p = 1 + x^2 \in \mathbb{R}[x]$ and let $I = (p)$.
   (a) Show $\mathbb{R}[x]/I$ is a field.
   (b) Show $\mathbb{R}[x]/I$ is isomorphic to $\mathbb{C}$. [*Hint*: Justify defining the mapping θ: $\mathbb{R}[x]/I \to \mathbb{C}$ by $\theta(I + (a + bx)) = a + ib$. Show that θ is bijective and preserves addition. To show that θ preserves multiplication, note that

$$(I + (a + bx))(I + (c + dx)) = I + (ac + (ad + bc)x + bdx^2) .$$

   Write this in the form $I + (u + vx)$ by following the first part of the proof of Theorem 6.7. Now show that

$$\theta[(I + (a + bx))(I + (c + dx))] = \theta(I + (a + bx))\theta(I + (c + dx)) .]$$

3. Suppose $f, g \in \mathcal{F}[x]$. Prove that $(f) = (g)$ if and only if f and g are associates.

4. Suppose $f, g \in \mathcal{F}[x]$. Prove or disprove the following:
   (a) If $(f) = (g)$, then $\deg f = \deg g$.
   (b) If $\deg f = \deg g$, then $(f) = (g)$.
   (c) If $f \in (g)$ and $\deg f = \deg g$, then $(f) = (g)$.

5. Find the greatest common divisor and least common multiple of the following pairs of polynomials:
   (a) $(x - 1)(x + 2)^2$ and $(x + 2)(x - 4)$.
   (b) $(x - 2)^2(x - 3)^4(x - i)$ and $(x - 1)(x - 2)(x - 3)^3$.
   (c) $(x^2 + 1)(x^2 - 1)$ and $(x + i)^3(x^3 - 1)$.

6. (a) Suppose $f_1, \ldots, f_n \in \mathcal{F}[x]$, and let $I = f_1\mathcal{F} + \cdots + f_n\mathcal{F}$ be the set of all polynomials of the form $g = f_1g_1 + \cdots + f_ng_n$ where $g_i \in \mathcal{F}[x]$. Show that I is an ideal. This is called the ideal **generated** by $\{f_1, \ldots, f_n\}$.
   (b) Show, in particular, that $\mathcal{F}[x]$ is an ideal generated by $\{1\}$. This is called the **unit** ideal.
   (c) Let d be the unique monic generator of I. Show that d divides each of the f_i.
   (d) If $c \in \mathcal{F}[x]$ divides each of the f_i, show that $c|d$.

(e) Suppose $\{f_1, f_2, f_3\}$ generates the unit ideal. Show that we can always find polynomials $f_{ij} \in \mathcal{F}[x]$ such that

$$\begin{vmatrix} f_1 & f_2 & f_3 \\ f_{21} & f_{22} & f_{23} \\ f_{31} & f_{32} & f_{33} \end{vmatrix} = 1 \ .$$

[*Hint*: Show there exists $g_1, g_2, g_3 \in \mathcal{F}[x]$ such that $\sum g_i f_i = 1$, and let $\alpha = \gcd\{g_1, g_2\}$. Next, show there exists $h_1, h_2 \in \mathcal{F}[x]$ such that $(g_1/\alpha)h_1 + (g_2/\alpha)h_2 = 1$. Now use the polynomials g_i, h_i and α to form the f_{ij}.]

6.4 POLYNOMIALS OVER ALGEBRAICALLY CLOSED FIELDS

We now turn to a discussion of polynomials over the fields $\mathbb{R}$ and $\mathbb{C}$. These are well worth considering in more detail since most practical applications in mathematics and physics deal with these two special cases. By way of terminology, a field $\mathcal{F}$ is said to be **algebraically closed** if every polynomial $f \in \mathcal{F}[x]$ with $\deg f > 0$ has at least one zero (or root) in $\mathcal{F}$.

Our next theorem is called the Fundamental Theorem of Algebra. While most proofs of this theorem involve the theory of complex variables, this result is so fundamental to our work that we present a proof in Appendix A that depends only on some relatively elementary properties of metric spaces. Basically, if the reader knows that a continuous function defined on a compact space takes its maximum and minimum values on the space, then there should be no problem understanding the proof. However, if the reader does not even know what a compact space is, then Appendix A presents all of the necessary formalism for a reasonably complete understanding of the concepts involved.

Theorem 6.12 (**Fundamental Theorem of Algebra**) The complex number field $\mathbb{C}$ is algebraically closed.

Proof See Appendix A. ∎

Let $p = a_0 + a_1 x + \cdots + a_n x^n$ be a polynomial of degree $n \geq 1$ over an algebraically closed field $\mathcal{F}$. Then there exists an element $\alpha_1 \in \mathcal{F}$ such that $p(\alpha_1) = 0$. Hence applying the factor theorem (Corollary to Theorem 6.4) we have

$$p = (x - \alpha_1)q$$

where q is a polynomial of degree n − 1. Again, if n − 1 > 0, we see that q has a zero α_2 in $\mathcal{F}$, and continuing this process we obtain

$$p = c(x - \alpha_1)(x - \alpha_2) \cdots (x - \alpha_n)$$

where c is a unit of $\mathcal{F}$. We thus see that any polynomial of degree n ≥ 1 over an algebraically closed field has exactly n roots (although they need not be distinct). We repeat this statement as part of the next theorem.

Theorem 6.13 Let $\mathcal{F}$ be an algebraically closed field. Then every prime polynomial $p \in \mathcal{F}[x]$ has (up to a unit factor) the form x − a where $a \in \mathcal{F}$. Moreover, every monic polynomial $f \in \mathcal{F}[x]$ can be factored into the form

$$f = \prod_{i=1}^{n} (x - a_i)$$

where each $a_i \in \mathcal{F}$.

Proof Let $p \in \mathcal{F}[x]$ be prime. Since $\mathcal{F}$ is algebraically closed, there exists $a \in \mathcal{F}$ such that p(a) = 0. By the factor theorem, x − a must be a factor of p. But p is prime so its only factors are its associates and units. This proves the first part of the theorem.

Now let $f \in \mathcal{F}[x]$ be of degree n ≥ 1. The second part of the theorem is essentially obvious from the first part and Theorem 6.6. However, we may proceed as follows. Since $\mathcal{F}$ is algebraically closed there exists $a_1 \in \mathcal{F}$ such that $f(a_1) = 0$, and hence by the factor theorem,

$$f = (x - a_1)q_1$$

where $q_1 \in \mathcal{F}[x]$ and deg $q_1 = n - 1$ (Theorem 6.2(b)). Now, by the algebraic closure of $\mathcal{F}$ there exists $a_2 \in \mathcal{F}$ such that $q_1(a_2) = 0$, and therefore

$$q_1 = (x - a_2)q_2$$

where deg $q_2 = n - 2$. It is clear that we can continue this process a total of n times, finally arriving at

$$f = c(x - a_1)(x - a_2) \cdots (x - a_n)$$

where $c \in \mathcal{F}$ is a unit. In particular, c = 1 if q_{n-1} is monic. ∎

While Theorem 6.13 shows that any polynomial of degree n over an algebraically closed field has exactly n (not necessarily distinct) roots, a more general result is the following.

Theorem 6.14 Any polynomial $p \in \mathcal{F}[x]$ of degree $n \geq 1$ over $\mathcal{F}$ has at most n roots in $\mathcal{F}$.

Proof We proceed by induction on the degree n of p. If $n = 1$, then $p = a_0 + a_1x$ so that $-a_1/a_0$ is the unique root of p, and the theorem thus holds in this case. Now assume that $n > 1$ and that the theorem holds for all polynomials of degree less than n. If p has no roots, then the conclusion of the theorem is valid, so we assume that p has at least one root c. Then $(x - c)|p$ so that $p = (x - c)\,q$ for some $q \in \mathcal{F}[x]$ with $\deg q = n - 1$ (Theorem 6.2(b)). By our induction hypothesis, q has at most $n - 1$ roots in $\mathcal{F}$, so the proof will be finished if we can show that p has no roots in $\mathcal{F}$ other than c and the roots of q. Suppose that $b \in \mathcal{F}$ is such that $p(b) = (b - c)q(b) = 0$. Since the field $\mathcal{F}$ can have no zero divisors (Exercise 1.5.12), it must be true that either $b - c = 0$ or $q(b) = 0$. In other words, if $p(b) = 0$, then either $b = c$ or else b is a root of q. ∎

Corollary Every polynomial p of degree $n \geq 1$ over an algebraically closed field $\mathcal{F}$ has n roots in $\mathcal{F}$.

Proof While this was proved in Theorem 6.13, we repeat it here in a slightly different manner. As was done in the proof of Theorem 6.14, we proceed by induction on the degree of p. The case $n = 1$ is true as above, so we assume that $n > 1$. Since $\mathcal{F}$ is algebraically closed, there exists at least one root $c \in \mathcal{F}$ such that $p = (x - c)q$ where $\deg q = n - 1$. By our induction hypothesis, q has $n - 1$ roots in $\mathcal{F}$ which are also clearly roots of p. It therefore follows that p has at least $n - 1 + 1 = n$ roots in $\mathcal{F}$, while Theorem 6.14 shows that p has at most n roots in $\mathcal{F}$. Therefore p must have exactly n roots in $\mathcal{F}$. ∎

While we proved in Theorem 6.12 that the field $\mathbb{C}$ is algebraically closed, it is not true that $\mathbb{R}$ is algebraically closed. This should be obvious because any quadratic equation of the form $ax^2 + bx + c = 0$ has solutions given by the quadratic formula

$$\frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

and if $b^2 - 4ac < 0$, then there is no solution for x in the real number system. (Recall that the quadratic formula follows by writing

$$0 = x^2 + bx/a + c/a = (x + b/2a)^2 - b^2/4a^2 + c/a$$

and solving for x.) However, in the case of $\mathbb{R}[x]$, we do have the following result.

Theorem 6.15 Suppose $f = a_0 + a_1x + \cdots + a_nx^n \in \mathbb{R}[x]$. If $\alpha \in \mathbb{C}$ is a root of f, then so is α^*. Furthermore, if $\alpha \neq \alpha^*$, then $(x - \alpha)(x - \alpha^*)$ is a factor of f.

Proof If $\alpha \in \mathbb{C}$ is a root of f, then $a_0 + a_1\alpha + \cdots + a_n\alpha^n = 0$. Taking the complex conjugate of this equation and remembering that each $a_i \in \mathbb{R}$, we obtain $a_0 + a_1\alpha^* + \cdots + a_n\alpha^{*n} = 0$ so that α^* is also a root of f. The second part of the theorem now follows directly from the factor theorem. ∎

Corollary Every prime polynomial in $\mathbb{R}[x]$ is (up to a unit factor) either of the form $x - a$ or $x^2 + ax + b$ where $a, b \in \mathbb{R}$ and $a^2 - 4b < 0$.

Proof Let $f \in \mathbb{R}[x]$ be prime, and let $\alpha \in \mathbb{C}$ be a root of f (that α exists follows from Theorem 6.13). Then $x - \alpha$ is a factor of f so that if $\alpha \in \mathbb{R}$, then $f = c(x - \alpha)$ where $c \in \mathbb{R}$ (since f is prime). But if $\alpha \notin \mathbb{R}$, then $\alpha \in \mathbb{C}$ and $\alpha^* \neq \alpha$ so that by Theorem 6.15, f has the factor

$$(x - \alpha)(x - \alpha^*) = x^2 - (\alpha + \alpha^*)x + \alpha\alpha^* \ .$$

Writing $\alpha = u + iv$ we see that

$$-a = \alpha + \alpha^* = 2u \in \mathbb{R}$$

and

$$b = \alpha\alpha^* = u^2 + v^2 \in \mathbb{R}$$

so that f has the form (up to a unit factor) $x^2 + ax + b$. Finally, note that

$$a^2 - 4b = 4u^2 - 4(u^2 + v^2) = -4v^2 < 0 \ . \ ∎$$

Exercises

1. Suppose $x^n + a_{n-1}x^{n-1} + \cdots + a_1x + a_0 \in \mathbb{C}[x]$ has zeros $\alpha_1, \ldots, \alpha_n$. Prove that $a_0 = \pm\alpha_1 \cdots \alpha_n$ and $a_{n-1} = -(\alpha_1 + \cdots + \alpha_n)$.

2. Let $V_n \subset \mathcal{F}[x]$ denote the set of all polynomials of degree $\leq n$, and let $a_0, a_1, \ldots, a_n \in \mathcal{F}$ be distinct.

(a) Show that V_n is a vector space over $\mathcal{F}$ with basis $\{1, x, x^2, \ldots, x^n\}$, and hence that dim $V_n = n + 1$.
(b) For each $i = 0, \ldots, n$, define the mapping $T_i: V_n \to \mathcal{F}$ by $T_i(f) = f(a_i)$. Show that the T_i are linear functionals on V_n, i.e., that $T_i \in V_n^*$.
(c) For each $k = 0, \ldots, n$ define the polynomial

$$p_k(x) = \frac{(x-a_0)\cdots(x-a_{k-1})(x-a_{k+1})\cdots(x-a_n)}{(a_k-a_0)\cdots(a_k-a_{k-1})(a_k-a_{k+1})\cdots(a_k-a_n)}$$
$$= \prod_{i\neq k}\left(\frac{x-a_i}{a_k-a_i}\right) \in V_n \quad .$$

Show that $T_i(p_j) = \delta_{ij}$.
(d) Show that $p_0, \ldots, p_n$ forms a basis for V_n, and hence that any $f \in V_n$ may be written as

$$f = \sum_{i=0}^{n} f(a_i)p_i \quad .$$

(e) Now let $b_0, b_1, \ldots, b_n \in \mathcal{F}$ be arbitrary, and define $f = \sum b_i p_i$. Show that $f(a_j) = b_j$ for $0 \leq j \leq n$. Thus there exists a polynomial of degree $\leq n$ that takes on given values at $n + 1$ distinct points.
(f) Now assume that $f, g \in \mathcal{F}[x]$ are of degree $\leq n$ and satisfy $f(a_j) = b_j = g(a_j)$ for $0 \leq j \leq n$. Prove that $f = g$, and hence that the polynomial defined in part (e) is unique. This is called the **Lagrange interpolation formula**.

3. Suppose $Q \subset M_2(\mathbb{C})$ is the set of all complex matrices of the form

$$\begin{pmatrix} z & w \\ -w^* & z^* \end{pmatrix} .$$

(a) Prove that Q is a division ring (i.e., that the nonzero elements of Q form a multiplicative group). Q is called the ring of **quaternions**.
(b) Prove that Q is not a field.
(c) Prove that $x^2 + 1 \in Q[x]$ has infinitely many roots in Q (where 1 denotes the unit element of Q, i.e., the 2 x 2 identity matrix).

4. Prove that $f, g \in \mathbb{C}[x]$ are relatively prime if and only if they have no root in common.

5. Let D be the differentiation operator defined in Problem 6.1.4, and suppose $f \in \mathbb{C}[x]$ is a monic polynomial. Prove that $f = (x - \alpha_1) \cdots (x - \alpha_n)$

where $\alpha_1, \ldots, \alpha_n \in \mathbb{C}$ are distinct if and only if f and Df are relatively prime.

6. If $f \in \mathcal{F}[x]$ has a root α, and $f = (x - \alpha)^m g$ where $g(\alpha) \neq 0$, then α is said to be a root of **multiplicity** m. In other words, m is the largest integer such that $(x - \alpha)^m | f$. Let α be a root of $f \in \mathcal{F}[x]$ and assume that $\deg f \geq 1$. Show that the multiplicity of f is > 1 if and only if $Df(\alpha) = 0$, and hence that the multiplicity of α is 1 if $Df(\alpha) \neq 0$. (See Problem 6.1.4 for the definition of Df.)

7. Show that the following polynomials have no multiple roots in $\mathbb{C}$ (see the previous problem for the definition of multiplicity):
   (a) $x^4 + x$.
   (b) $x^5 - 5x + 1$.
   (c) $x^2 + bx + c$ where $b, c \in \mathbb{C}$ and $b^2 - 4c \neq 0$.

6.5 THE FIELD OF QUOTIENTS

What we will do in this section is show how to construct a field out of the ring $\mathcal{F}[x]$. Rather than talk about polynomials specifically, we use the fact that $\mathcal{F}[x]$ is an integral domain and treat the problem on a more general footing.

Notice that the set $\mathbb{Z}$ of all integers has the property that if $ab = 0$ for some $a, b \in \mathbb{Z}$, then either a or b must equal 0. Since $\mathbb{Z}$ is a ring, this shows that $\mathbb{Z}$ is in fact an integral domain. Also note though, for any $a \in \mathbb{Z}$, $a \neq 1$, we have $a^{-1} = 1/a \notin \mathbb{Z}$, so that $\mathbb{Z}$ is not a field. However, if we enlarge the set $\mathbb{Z}$ to include all of the rational numbers, then we do indeed obtain a field. What this really entails is taking all pairs $a, b \in \mathbb{Z}$ and forming the object a/b with the appropriate algebraic operations defined on it. In this particular case, we say that $a/b = c/d$ if and only if $ad = bc$, and we define the operations of addition and multiplication by

$$a/b + c/d = (ad + bc)/bd$$

and

$$(a/b)(c/d) = (ac)/(bd) \ .$$

In order to generalize this result, we make the following definition. Let D be an integral domain (i.e., a commutative ring with no zero divisors), let D' denote the set of all *nonzero* elements of D, and let Q be the set of all ordered pairs

$$Q = \{(a, b) \in D \times D'\} \ .$$

(You may think of (a, b) as the quotient a/b.) We define a relation ~ on Q by (a, b) ~ (c, d) if ad = bc, and we claim that this is an equivalence relation (for example, 2/3 is "equivalent" to 8/12). To prove this, we must verify the three requirements given in Section 0.3. First, for any (a, b) $\in Q$ we have (a, b) ~ (a, b) since ab = ba. Next, for any (a, b), (c, d) $\in Q$ we see that (a, b) ~ (c, d) implies ad = bc, and hence cb = da which thus implies (c, d) ~ (a, b). Finally, suppose (a, b), (c, d), (e, f) $\in Q$ where (a, b) ~ (c, d) and (c, d) ~ (e, f). Then ad = bc and cf = de, and therefore bde = bcf = adf. But D is commutative, and hence this is just afd = bed. By assumption $d \neq 0$ so that, since D is an integral domain, we must have af = be and thus (a, b) ~ (e, f).

We are now in a position to show that any integral domain can be enlarged in a similar manner to form a field. By way of terminology, if there is a one-to-one homomorphism (i.e., an isomorphism) of a ring R into a ring R′, then we say that R can be **embedded** in R′. Furthermore, if R and R′ are both rings with unit elements 1 and 1′ respectively, then we require that the embedding take 1 into 1′. The ring R′ is also called an **extension** of R.

The proof of the next theorem appears to be quite involved, but it is actually nothing more than a long series of simple steps.

Theorem 6.16 Every integral domain D can be embedded in a field.

Proof Let D′ be the nonzero elements of D, and let Q be the set of all ordered pairs (a, b) $\in$ D × D′ as defined above. We let [a, b] denote the equivalence class in Q of (a, b) as constructed above. In other words,

$$[a, b] = \{(x, y) \in D \times D' : (a, b) \sim (x, y)\} .$$

We claim that the set $\mathcal{F}_D$ of all such equivalence classes forms a field. To prove this, we must first define addition and multiplication in $\mathcal{F}_D$.

Guided by the properties of $\mathbb{Z}$, we define addition in $\mathcal{F}_D$ by the rule

$$[a, b] + [c, d] = [ad + bc, bd] .$$

Since D is an integral domain, we know that $bd \neq 0$ for any nonzero b, d $\in$ D, and hence [ad + bc, bd] $\in \mathcal{F}_D$. We still must show that this addition is well-defined, i.e., if [a, b] = [a′, b′] and [c, d] = [c′, d′], then

$$[a, b] + [c, d] = [a', b'] + [c', d'] .$$

This is equivalent to showing that

$$[ad + bc, bd] = [a'd' + b'c', b'd']$$

or, alternatively, that

$$(ad + bc)b'd' = bd(a'd' + b'c') \ .$$

From $[a, b] = [a', b']$ we have $ab' = ba'$, and similarly $cd' = dc'$. Therefore we indeed have

$$\begin{aligned}(ad + bc)b'd' &= adb'd' + bcb'd' = ab'dd' + bb'cd' = ba'dd' + bb'dc' \\ &= bda'd' + bdb'c' = bd(a'd' + b'c') \ .\end{aligned}$$

Since D is commutative, it should be clear that if $c \neq 0$ then

$$[a, b] = [ac, bc] = [ca, cb] \ .$$

Therefore

$$[a, b] + [0, c] = [ac + b0, bc] = [ac, bc] = [a, b]$$

so that $[0, c]$ is a zero element for addition. We now see that

$$[a, b] + [-a, b] = [ab - ba, bb] = [0, b]$$

and hence $[-a, b]$ is the negative of $[a, b]$. The reader should now have no trouble showing that $\mathcal{F}_D$ is an abelian group under addition.

To complete our ring structure, we define multiplication in $\mathcal{F}_D$ by the rule

$$[a, b][c, d] = [ac, bd] \ .$$

As was the case with addition, the fact that $b, d \neq 0$ means that $bd \neq 0$, and hence the product is an element of $\mathcal{F}_D$. We leave it to the reader to show that the product is well-defined (see Exercise 6.5.1). It is also easy to see that $[x, x]$ is a unit element in $\mathcal{F}_D$ for any nonzero $x \in D$, and that the nonzero elements of $\mathcal{F}_D$ (i.e., those of the form $[a, b]$ with $a \neq 0$) form an abelian group under multiplication with the inverse of an element $[a, b]$ given by $[a, b]^{-1} = [b, a]$ (where $[b, a] \in \mathcal{F}_D$ since $a \neq 0$).

As to the ring axioms, we show only one of the distributive laws, and leave the others to the reader. This will complete the proof that $\mathcal{F}_D$ forms a field. If $[a, b], [c, d], [e, f] \in \mathcal{F}_D$, then

$$\begin{aligned}[a, b]([c, d]+[e, f]) &= [a, b][cf + de, df] = [acf + ade, bdf] \\ &= [b(acf + ade), bbdf] = [(ac)(bf) + (bd)(ae), (bd)(bf)] \\ &= [ac, bd] + [ae, bf] = [a, b][c, d] + [a, b][e, f]\end{aligned}$$

What we have accomplished up to this point is the construction of a field $\mathcal{F}_D$ from an arbitrary integral domain D. It still must be shown that D can be embedded in $\mathcal{F}_D$. As was noted above, for any nonzero x, y $\in$ D we have [ax, x] = [ay, y] because (ax)y = x(ay). This means that we can denote the element [ax, x] $\in \mathcal{F}_D$ by [a, 1]. (It is important to realize that the ring D does not necessarily contain a unit element, so that there need not exist a unit element $1 \in D$. What we have just done is *define* the element [a, 1] $\in \mathcal{F}_D$. Everything that follows in the remainder of this proof holds if we replace the symbol 1 by an arbitrary nonzero element $x \in D$.)

We now define the mapping ϕ: D $\to \mathcal{F}_D$ by $\phi(a) = [a, 1]$ for all $a \in D$. If $\phi(a) = \phi(a')$, then [a, 1] = [a′, 1] so that a1 = 1a′ and hence a = a′, thus proving that ϕ is one-to-one. (As we just mentioned, the symbol 1 could actually be replaced by any $x \neq 0$ since the fact that D is an integral domain then says that if ax = xa′, then x(a − a′) = 0 which also implies that a = a′.) To finish the proof, we need only show that ϕ is a homomorphism. But for any a, b $\in$ D we have

$$\phi(a + b) = [a + b, 1] = [a1 + b1, 1\cdot 1] = [a, 1] + [b, 1] = \phi(a) + \phi(b)$$

and

$$\phi(ab) = [ab, 1] = [ab, 1\cdot 1] = [a, 1][b, 1] = \phi(a)\phi(b) \ . \ \blacksquare$$

The field $\mathcal{F}_D$ constructed in this theorem is usually called the **field of quotients** of D. If we start with the ring of integers $\mathbb{Z}$, then this construction yields the rational field $\mathbb{Q}$. While we have shown that any integral domain D can be embedded in its field of quotients, there can be other fields in which D can also be embedded. However, it can be shown that $\mathcal{F}_D$ is the "smallest" field in which D can be embedded (see Exercise 6.5.2).

Exercises

1. Referring to the proof of Theorem 6.16, show that the product in $\mathcal{F}_D$ is well-defined.

2. Show that $\mathcal{F}_D$ is the smallest field in which an integral domain D can be embedded. In other words, show that if $\mathcal{K}$ is any field containing an integral domain isomorphic to D, then $\mathcal{K}$ contains a field isomorphic to $\mathcal{F}_D$.

[*Hint*: For simplicity, assume that D is actually a subring of $\mathcal{K}$. Now, for any a, b ∈ D show that the map $\phi: \mathcal{F}_D \to \mathcal{K}$ defined by $\phi([a, b]) = ab^{-1}$ is one-to-one and preserves addition and multiplication. Thus ϕ is an isomorphism of $\mathcal{F}_D$ onto a subfield of $\mathcal{K}$.]

3. Show that $\mathcal{F}_D$ obeys all of the axioms for a ring.

6.6 POLYNOMIALS OVER FINITE FIELDS *

With very few exceptions (e.g., Exercise 1.5.15), the fields we have been using (such as $\mathbb{R}$ and $\mathbb{C}$) contain an infinite number of elements. However, it is also possible to construct many fields that contain only a finite number of elements. This section is meant to be only an introduction to the theory of finite fields.

Recall from Example 1.11 that two integers a, b ∈ $\mathbb{Z}$ are said to be **congruent modulo n** (where $n \in \mathbb{Z}^+$) if n|(a – b), and we write this as

$$a \equiv b \pmod{n} \ .$$

We also saw in Exercise 1.5.2 that for each $n \in \mathbb{Z}^+$, this defines an equivalence relation on $\mathbb{Z}$ that decomposes $\mathbb{Z}$ into n distinct congruence classes. We shall denote the congruence class (for a fixed n) of an integer $k \in \mathbb{Z}$ by [k]. If there is any possible ambiguity as to the value of n under discussion, we will write $[k]_n$. For example, if n = 5 we have

$$[2] = [7] = [-33] = \{\ldots, -8, -3, 2, 7, 12, \ldots\} \ .$$

We also refer to any of the integers in [k] as a representative of [k]. For example, 2 is the smallest positive representative of the class [7] (or the class [–33] etc.). We emphasize that [k] is a *subset* of $\mathbb{Z}$ for each k.

From Example 1.11, we say that the collection

$$\{[0], [1], [2], \ldots, [n-1]\}$$

forms a **complete set** of congruence classes modulo n, and we denote this collection by $\mathbb{Z}_n$. We first show that $\mathbb{Z}_n$ can be made into an abelian group. For any [a], [b] ∈ $\mathbb{Z}_n$ we define a group "addition" operation $[a] \oplus [b] \in \mathbb{Z}_n$ by

$$[a] \oplus [b] = [a + b] \ .$$

(Note that the symbol $\oplus$ is used in an entirely different context when we talk about direct sums.) It should be obvious that [0] will serve as the additive identity element since [a] $\oplus$ [0] = [a + 0] = [a] and [0] $\oplus$ [a] = [0 + a] = [a]. Noting that, for example with n = 5 again, we have [3] = [18] and [4] = [−1], we must be sure that [3] $\oplus$ [4] = [18] $\oplus$ [−1]. Clearly this is true because [3] $\oplus$ [4] = [7] = [2] and [18] $\oplus$ [−1] = [17] = [2]. In other words, we must be sure that this addition operation is well-defined. That this is in fact the case is included in the next theorem.

Theorem 6.17 The set $\mathbb{Z}_n$ is an abelian group with respect to the operation $\oplus$ defined above.

Proof We leave it to the reader to show that $\oplus$ is indeed well-defined (see Exercise 6.6.1). As to the group properties, we prove associativity, leaving the rest of the proof to the reader (see Exercise 6.6.2). We have

$$[a]\oplus([b]\oplus[c]) = [a]\oplus[b+c] = [a+(b+c)] = [(a+b)+c]$$
$$= [a+b]\oplus[c] = ([a]\oplus[b])\oplus[c] \ . \ \blacksquare$$

By virtue of this theorem, $\mathbb{Z}_n$ is called the **group of integers modulo n** (or simply mod n). We can also define another operation on $\mathbb{Z}_n$ that is analogous to multiplication. Thus, we define the "multiplication" operation $\otimes$ on $\mathbb{Z}_n$ by

$$[a] \otimes [b] = [ab] \ .$$

(Again, this symbol should not be confused with the tensor product to be introduced in Chapter 11.) For example, if n = 6 we have [2] $\otimes$ [5] = [10] = [4] and [3] $\otimes$ [−4] = [−12] = [0]. The closest analogue to Theorem 6.17 that we have for $\otimes$ is the following.

Theorem 6.18 The operation $\otimes$ defined above on $\mathbb{Z}_n$ is well-defined, obeys the associative and commutative laws, and has [1] as the identity element.

Proof See Exercise 6.6.3. $\blacksquare$

Since [1] is the identity element for $\otimes$, it is easy to see that [0] has no multiplicative inverse in $\mathbb{Z}_n$, and hence $\mathbb{Z}_n$ can not possibly form a group under $\otimes$. Let us denote the set $\mathbb{Z}_n - [0] = \{[1], [2], \ldots, [n-1]\}$ by $\mathbb{Z}_n{}^+$. It turns out that for some (but not all) values of n, $\mathbb{Z}_n{}^+$ will in fact form a group with respect to $\otimes$. We will leave specific examples of this to the exercises at the end of this section.

With the operations $\oplus$ and $\otimes$ defined, it is now easy to see that $\mathbb{Z}_n$ actually forms a commutative ring. All we must do is verify the axioms given in Section 1.4. We will show that the first half of axiom (R8) is obeyed, and leave it to the reader to verify the rest of the ring axioms (see Exercise 6.6.4). We therefore have

$$\begin{aligned}[a]\otimes([b]\oplus[c]) &= [a]\otimes[b+c] = [a(b+c)] = [ab+ac] \\ &= [ab]\oplus[ac] = ([a]\otimes[b])\oplus([a]\otimes[c]) \ .\end{aligned}$$

Now consider the ring $\mathbb{Z}_n$ and assume that n is not prime. Then we may write n = rs where r, s > 1. But then [r] $\otimes$ [s] = [rs] = [n] = [0] where [r], [s] $\neq$ [0]. Since [0] is the zero element of $\mathbb{Z}_n$, this shows that $\mathbb{Z}_n$ is not an integral domain if n is not prime. On the other hand, suppose that n = p is prime. We claim that $\mathbb{Z}_p$ is an integral domain. For, suppose [a] $\in \mathbb{Z}_p$ and [a] $\neq$ [0]. We may assume that a is the smallest positive representative of the equivalence class [a], and hence a < p. Now assume that [b] $\in \mathbb{Z}_p$ is such that [a] $\otimes$ [b] = [ab] = [0] (where we again choose b < p to be the smallest positive representative of the class [b]). Then by definition we have p|ab. But p is prime so (by Theorem 0.9) this implies that either p|a or p|b. Since a < p, it is impossible for p to divide a, and therefore p|b. Since $0 \leq b < p$, we must have b = 0, and thus $\mathbb{Z}_p$ is an integral domain. This proves the next result.

Theorem 6.19 The ring $\mathbb{Z}_n$ is an integral domain if and only if n is prime.

Noting that $\mathbb{Z}_n$ consists of n equivalence classes, we now claim that $\mathbb{Z}_n$ is in fact a field if n is prime. This is an immediate consequence of the following general result. Recall that a field is a commutative ring with identity element in which the nonzero elements form a multiplicative group (i.e., a commutative division ring). Furthermore, any field is necessarily an integral domain (see Exercise 1.5.6 or 1.5.12).

Theorem 6.20 Every finite integral domain is a field.

Proof Let D be a finite integral domain (which is commutative by definition). We must show that $1 \in D$, and that every nonzero $a \in D$ has a multiplicative inverse that is also in D. In other words, we must show that for every nonzero $a \in D$ there exists $b \in D$ such that $ab = 1 \in D$. Let $\{x_1, \ldots, x_n\}$ denote all the elements of D, and consider the set $\{ax_1, \ldots, ax_n\}$ where $a \in D$ and $a \neq 0$. If $ax_i = ax_j$ for $i \neq j$, then $a(x_i - x_j) = 0$ which (since D has no zero divisors) implies that $x_i = x_j$, contradicting the assumption that $i \neq j$. Thus

$ax_1, \ldots, ax_n$ are all distinct. Since D contains n elements, it follows that in fact we have $D = \{ax_1, \ldots, ax_n\}$. In other words, every $y \in D$ can be written in the form $ax_i = x_i a$ for some $i = 1, \ldots, n$. In particular, we must have $a = ax_{i_0}$ for some $i_0 = 1, \ldots, n$. Then for any $y = x_i a \in D$ we have

$$yx_{i_0} = (x_i a)x_{i_0} = x_i(ax_{i_0}) = x_i a = y$$

so that x_{i_0} may be taken as the identity element 1 in D. Finally, since we have now shown that $1 \in D$, it follows that $1 = ax_j$ for some particular $j = 1, \ldots, n$. Defining $b = x_j$ yields $1 = ab$ and completes the proof. ∎

Corollary $\mathbb{Z}_n$ is a field if and only if n is prime.

We now turn our attention to the question of whether there exist any finite fields that do not contain a prime number of elements. As with groups, we refer to the number of elements in a finite field as its **order**. This is not surprising since any field is a ring, and any ring is an additive group. We will frequently denote a finite field by F rather than by $\mathcal{F}$.

Example 6.10 Let $S_2(\mathcal{F}) \subset M_2(\mathcal{F})$ denote the set of all matrices of the form

$$\begin{pmatrix} x & y \\ -y & x \end{pmatrix}.$$

We will show that $S_2(\mathcal{F})$ is a field when $\mathcal{F} = \mathbb{Z}_3$ but not when $\mathcal{F} = \mathbb{Z}_5$.

We leave it as a simple exercise for the reader to show that if $A, B \in S_2(\mathcal{F})$, then $A + B$ and AB are also in $S_2(\mathcal{F})$. Furthermore, $AB = BA$ so that $S_2(\mathcal{F})$ is commutative. Note that $S_2(\mathcal{F})$ also contains the zero and identity matrices, and if $A \in S_2(\mathcal{F})$, then so is $-A$. Thus $S_2(\mathcal{F})$ is easily seen to be a subring of $M_2(\mathcal{F})$. We now consider the problem of inverses. If

$$A = \begin{pmatrix} x & y \\ -y & x \end{pmatrix}$$

has inverse

$$A^{-1} = \begin{pmatrix} a & b \\ -b & a \end{pmatrix}$$

then $AA^{-1} = I$ yields the two simultaneous equations

$$ax - by = 1$$
$$ay + bx = 0 \quad .$$

Since x, y, a, b $\in \mathcal{F}$ we can (formally) solve these for a and b to obtain

$$a = \; x(x^2 + y^2)^{-1}$$
$$b = -y(x^2 + y^2)^{-1} \quad .$$

The element $(x^2 + y^2)^{-1}$ will exist as long as $x^2 + y^2 \neq 0$.

In $\mathbb{Z}_3$ we have x, y = {0, 1, 2} and it is easy to see by direct calculation that $x^2 + y^2 \neq 0$ as long as (x, y) ≠ (0, 0). For example, if x = 1 and y = 2 we have $x^2 + y^2 = 1 + 1 = 2$. Thus every nonzero matrix in $S_2(\mathbb{Z}_3)$ is invertible, and hence $S_2(\mathbb{Z}_3)$ is a field with 9 elements.

On the other hand, in $\mathbb{Z}_5$ we see that $1^2 + 2^2 = 0$ so that the matrix with x = 1 and y = 2 is not invertible, and hence $S_2(\mathbb{Z}_5)$ is not a field. //

Example 6.11 Since $\mathbb{Z}_3$ is a field, we can consider polynomials in $\mathbb{Z}_3[x]$. These may be used to generate a field of order 9 as follows. We define the set $F_9 \subset \mathbb{Z}_3[x]$ consisting of the nine polynomials of degree ≤ 1 with coefficients in $\mathbb{Z}_3$:

$$F_9 \; = \; \{0, 1, 2, x, x + 1, x + 2, 2x, 2x + 1, 2x + 2\} \quad .$$

It is easy to see that this set is closed under ordinary polynomial addition. For example, remembering that our scalars lie in $\mathbb{Z}_3$ we have (x + 1) + (2x + 1) = 2. However, we must be careful in defining multiplication. This is because, for example, $(x + 1)(2x + 1) = 2x^2 + 1 \notin F_9$ even though it is in $\mathbb{Z}_3[x]$. To ensure that multiplication is closed, we multiply as usual in $\mathbb{Z}_3[x]$ and then **reduce modulo** $x^2 + 1$. In other words, we subtract off multiples of $x^2 + 1$. For example, we have

$$\begin{aligned} (x+1)(2x+1) &= 2x^2 + 1 && (\text{in } \mathbb{Z}_3[x]) \\ &= 2(x^2+1) + 2 \\ &= 2 && (\text{in } F_9) \quad . \end{aligned}$$

As another example,

$$\begin{aligned} (2x+1)(x) &= 2x^2 + x && (\text{in } \mathbb{Z}_3[x]) \\ &= 2(x^2+1) + (x+1) \\ &= x + 1 && (\text{in } F_9) \quad . \end{aligned}$$

Using the constant polynomials 0 and 1 as the 0 and 1 elements of a ring, it is easy to show that F_9 forms a commutative ring. That F_9 in fact is a field follows from the observation that each nonzero element of F_9 has the inverse shown below:

Element:	1	2	x	x+1	x+2	2x	2x+1	2x+2
Inverse:	1	2	2x	x+2	x+1	x	2x+2	2x+1

We leave verification of these facts to the reader. //

Let G be any group, and let e be the identity element of G. If there exists an element $x \in G$ such that every element in G is a power of x, then G is said to be a **cyclic** group **generated** by x. The cyclic group generated by x is usually denoted by $\langle x \rangle$. If we consider the set of all powers of the generator x, then this set will consist of either all distinct elements, or else some elements will be repeated. In the case of repeated elements, there will exist a positive integer $m > 0$ such that $x^m = e$ while no smaller nonzero power of x is also equal to e. For example, if $i \leq j$ and $x^i = x^j$, we have $e = x^j(x^i)^{-1} = x^j\, x^{-i} = x^{j-i}$. Then let m be the smallest nonzero value of all such differences $j - i$.

We say that $G = \{e, x, x^2, \ldots, x^{m-1}\}$ is a cyclic group of **order** m, and we denote this group by C_m. We are using the letter e to denote the identity element of a general group so that we may distinguish between the multiplicative identity (usually written as 1) of some groups and the additive identity (usually written as 0) of other groups. Note that given any $k \in \mathbb{Z}$ we may write $k = qm + r$ (where $0 \leq r < m$), and hence

$$x^k = (x^m)^q\, x^r = e^q\, x^r = x^r \ .$$

Thus all powers of x can indeed be written in terms of the first m powers of x.

In the case that *all* powers of x are distinct, then the group

$$G = \{ \ldots, x^{-2}, x^{-1}, e, x, x^2, \ldots \}$$

contains an infinite number of elements. We denote such an infinite cyclic group by C_∞. Of importance to us is the fact that many apparently diverse groups are isomorphic to cyclic groups (either finite or infinite). We will see a simple example of this below.

If $x \in G$ and $m > 0$ is the smallest integer such that $x^m = 1$, then m is called the **order** of x, and will also be denoted by o(x). (This should not be confused with the order of G which is the number of elements in G. It is for this reason that o(G) is frequently denoted by $|G|$.) If $o(x) \leq o(G)$, then it is

easy to see that $\langle x \rangle$ is simply a subgroup of G, and furthermore that $o(x) = o(\langle x \rangle)$ (see Exercise 6.6.9).

Example 6.12 Let us show that the set of integers $\mathbb{Z}$ under addition is isomorphic to $C_\infty = \langle x \rangle$. We define the mapping $\phi: \mathbb{Z} \to C_\infty$ by $\phi(n) = x^n$. Clearly ϕ is a homomorphism of groups since

$$\phi(n + m) = x^{n+m} = x^n x^m = \phi(n)\phi(m) \ .$$

(Note that $n + m$ in $\phi(n + m)$ denotes the "product" of two group elements in $\mathbb{Z}$ while $\phi(n)\phi(m)$ denotes the product of two elements in C_∞.) It should also be obvious that ϕ is surjective because every $x^k \in C_\infty$ is just the image of $k \in \mathbb{Z}$. Finally, ϕ is injective since $\text{Ker}\ \phi = \{0\}$. We have thus constructed an isomorphism of $\mathbb{Z}$ onto C_∞.

We leave it to the reader to show that $\mathbb{Z}_m$ is isomorphic to C_m (see Exercise 6.6.7). //

Let $\mathcal{F}$ be any field. Since $\mathcal{F}$ is closed under addition and contains the multiplicative identity element 1, we see that for any $n \in \mathbb{Z}^+$, the sum $1 + \cdots + 1$ of n 1's is also in $\mathcal{F}$. For example, $2 = 1 + 1 \in \mathcal{F}$ as is $3 = 1 + 1 + 1$ and so on. (However, it is important to stress that in an arbitrary field these elements need not be distinct.) Therefore the positive integers form a (not necessarily infinite) cyclic subgroup $\langle 1 \rangle$ of the *additive* group of $\mathcal{F}$. In other words

$$\langle 1 \rangle = \{0, 1, 2, \ldots\}$$

where each $n \in \langle 1 \rangle$ denotes $1 + \cdots + 1$ (n times) in $\mathcal{F}$.

Now consider the special case where F is a finite field. Note that since $\langle 1 \rangle$ is a subgroup of F, Theorem 1.9 tells us that $o(\langle 1 \rangle) | o(F)$. The number $o(\langle 1 \rangle)$ is called the **characteristic** of F (see Section 1.5). For example, if $F = \mathbb{Z}_p$, then $\langle 1 \rangle = \{0, 1, 2, \ldots, p - 1\} = \mathbb{Z}_p$ and hence the characteristic of $\mathbb{Z}_p$ is $o(\langle 1 \rangle) = p = o(F)$. The characteristic of an infinite field may or may not be finite.

Example 6.13 Let us show that the characteristic of any finite field must be a prime number (if it is nonzero). By definition, the characteristic of F is the smallest $m \in \mathbb{Z}^+$ such that $m \cdot 1 = 1 + \cdots + 1 = 0$. If $m \neq 0$ is not prime, then we may write $m = rs$ for some integers $0 < r < m$ and $0 < s < m$. But then we have $rs = 0$ which implies (since we are in a field) that either $r = 0$ or $s = 0$. In either case this contradicts the definition of m as the least positive integer such that $m = 0$. Thus m must be prime. Note that this proof actually applies to any finite integral domain with an identity element. //

We now wish to prove that if a finite field exists, then its order must be a prime power. For example, we have seen that $\mathbb{Z}_n$ is a field if and only if n is prime, and the field F_9 discussed above is of order $9 = 3^2$. Our claim will follow as an immediate corollary of the next theorem. The reader should recall that the direct product of two groups was defined in Exercise 1.1.5. The direct product of a finite number of groups follows by an obvious induction argument.

Example 6.14 Consider the cyclic groups $C_2 = \langle x \rangle = \{1, x\}$ and $C_3 = \langle y \rangle = \{1, y, y^2\}$. Then the product $C_2 \times C_3$ consists of the six elements

$$C_2 \times C_3 = \{(1, 1), (1, y), (1, y^2), (x, 1), (x, y), (x, y^2)\} .$$

To show that this product group is isomorphic to C_6, let $z = (x, y) \in C_2 \times C_3$. Then by definition of the group product in $C_2 \times C_3$ we have $z^2 = (1, y^2)$, $z^3 = (x, 1)$, $z^4 = (1, y)$, $z^5 = (x, y^2)$, $z^6 = (1, 1)$. Therefore $C_2 \times C_3 = \{z, z^2, \ldots, z^6\}$ which is just a cyclic group of order 6. ∥

Theorem 6.21 If F is a finite field of characteristic p, then for some $r \geq 1$, the additive group of F is isomorphic to the r-fold direct product $(C_p)^r$. Thus $o(F) = p^r$.

Proof We leave it to the reader to show that $\mathbb{Z}_p$ is isomorphic to a subfield of F, and hence that F may be considered to be a vector space V over the field $\mathbb{Z}_p$. Since F is finite, $r = \dim V$ must also be finite. By the corollary to Theorem 2.8, V is isomorphic to $(\mathbb{Z}_p)^r = \mathbb{Z}_p \times \cdots \times \mathbb{Z}_p$. Noting that $\mathbb{Z}_p = \{0, 1, \ldots, p - 1\} = \langle 1 \rangle$ is isomorphic to the (additive) cyclic group C_p, it follows that V is isomorphic to $(C_p)^r$ (i.e., the set of all r-tuples of field elements). Thus V has p^r elements. ∎

Corollary The order of a finite field must be a prime power.

Proof This follows from Example 6.13 and Theorem 6.21. ∎

We now comment briefly on the construction of finite fields. What we shall do is generalize the procedure demonstrated in Example 6.11 where we constructed the field F_9. Recall that the problem came in defining a closed multiplication in $\mathbb{Z}_3[x]$. A more general way to view the solution to this problem is to define an equivalence relation $\approx$ on $\mathbb{Z}_3[x]$ by the requirement that $a \approx b$ if $(x^2 + 1) | (a - b)$. Note that $x^2 + 1$ is prime in $\mathbb{Z}_3[x]$. We may now

take the elements of F_9 to be the equivalence classes of the nine polynomials that were previously used in Example 6.11 to define F_9. Another way to say this is that these nine polynomials of degree ≤ 1 form a complete set of representatives of the classes. In this case, addition and multiplication are defined as expected by

$$[a] + [b] = [a + b]$$

and

$$[a][b] = [ab] .$$

Note that if $(x^2 + 1)|(a - b)$, then the remainder of a divided by $x^2 + 1$ must be the same as the remainder of b divided by $x^2 + 1$. Since there are only a finite number of polynomials in $\mathbb{Z}_3[x]$, there can be only a finite number of distinct remainders, and the degree of each remainder must be less than that of $x^2 + 1$. Referring to Theorem 6.7 and its proof, a moments thought should convince you that all we are doing is considering the cosets $\mathbb{Z}_3[x]/(x^2 + 1)$ where $(x^2 + 1)$ denotes the principal ideal generated by $k = x^2 + 1 \in \mathbb{Z}_3[x]$. This is because any $p \in \mathbb{Z}_3[x]/(k) = \mathbb{Z}_3[x]/I$ is of the form $I + h$ where $h \in \mathbb{Z}_3[x]$. But $h = qk + r$ for some $q \in \mathbb{Z}_3[x]$ and where deg r < deg k, and hence $qk \in I$. Therefore p must actually be of the form $I + r$, and thus there can be only as many distinct such p as there are distinct r.

The next theorem shows that this approach works in general.

Theorem 6.22 Let $k \in \mathbb{Z}_p[x]$ be a prime polynomial of degree r, and define an equivalence relation on $\mathbb{Z}_p[x]$ by $a \approx b$ if and only if $k|(a - b)$. Then the corresponding set of equivalence classes in $\mathbb{Z}_p[x]$ is a field of order p^r.

Proof First note that if $a_i \in \mathbb{Z}_p[x]$, then there are p^r distinct polynomials of the form $a_0 + a_1x + \cdots + a_{r-1}x^{r-1}$. This set of p^r polynomials (which consists of all distinct polynomials of degrees 0, 1, 2, . . . , r − 1) forms a complete set of representatives of the classes, and hence there are p^r classes in all. Since these equivalence classes are just the cosets $\mathbb{Z}_p[x]/(k)$ where k is prime, it follows from Theorem 6.11 that $\mathbb{Z}_p[x]/(k)$ is a field. ∎

One consequence of this theorem is that to construct a field of order p^r we need only find a prime polynomial of degree r in $\mathbb{Z}_p[x]$. While this is easy enough to do in most common cases, it is fairly hard to prove that there exists at least one prime polynomial for every choice of p and r. We refer the interested reader to e.g., the very readable book by Biggs (1985).

Exercises

1. Show that the operation $\oplus$ defined on $\mathbb{Z}_n$ is well-defined. In other words, show that if $[a_1] = [a_2]$ and $[b_1] = [b_2]$ then $[a_1 + b_1] = [a_2 + b_2]$.

2. Finish the proof that $\mathbb{Z}_n$ forms an additive group.

3. Prove Theorem 6.18.

4. Finish the proof that $\mathbb{Z}_n$ forms a ring.

5. Finish the details in Example 6.10.

6. Finish the details in Example 6.11.

7. Prove that $\mathbb{Z}_m$ is isomorphic to C_m.

8. If m and n are relatively prime positive integers, prove that $C_m \times C_n$ is isomorphic to C_{mn}. [*Hint*: Suppose $C_m = \langle x \rangle$ and $C_n = \langle y \rangle$. Let $z = (x, y) \in C_m \times C_n$ have order r. Show that $r = mn$ and then conclude that $C_m \times C_n$ must be a cyclic group.]

9. Let G be a group, and suppose $\langle x \rangle \subset G$. If $o(x) \leq o(G)$, show that $\langle x \rangle$ is a subgroup of G and that $o(x) = o(\langle x \rangle)$.

10. Fill in the details in the proof of Theorem 6.22.

11. For each of the following expressions in $\mathbb{Z}_5$, write the answer as [0], [1], [2], [3] or [4]:
    (a) $[3] \oplus [4]$
    (b) $[2] \oplus [-7]$
    (c) $[17] \oplus [76]$
    (d) $[3] \otimes [4]$
    (e) $[2] \otimes [-7]$
    (f) $[17] \otimes [76]$
    (g) $[3] \otimes ([2] \oplus [4])$
    (h) $([3] \otimes [2]) \oplus ([3] \otimes [4])$

12. Repeat the previous problem in $\mathbb{Z}_6$.

13. (a) Which elements of $\mathbb{Z}_4$ are zero divisors?
    (b) Which elements of $\mathbb{Z}_{10}$ are zero divisors?

14. Show that ([2], [0]) is a zero divisor in $\mathbb{Z}_3 \times \mathbb{Z}_3$.

15. (a) Show that $1 + x + x^2 \in \mathbb{Z}_2[x]$ is prime over $\mathbb{Z}_2$.
    (b) Show that $1 + x^2 \in \mathbb{Z}_3[x]$ is prime over $\mathbb{Z}_3$.
    [*Hint*: Use the factor theorem.]

16. (a) Show that $1 + x^2 + x^3 \in \mathbb{Z}_2[x]$ is prime over $\mathbb{Z}_2$, and use this to construct a field of order 8.
    (b) What is the order of its multiplicative group?

17. Prove that for every prime number p there exist fields of order p^2 and p^3.

18. For which of the following primes p can we construct a field of order p^2 by using the polynomial $1 + x^2$?

$$p = 3, 5, 7, 11, 13, 19, 23 \ .$$

    Describe the multiplicative group for the first two cases in which the field can be constructed.