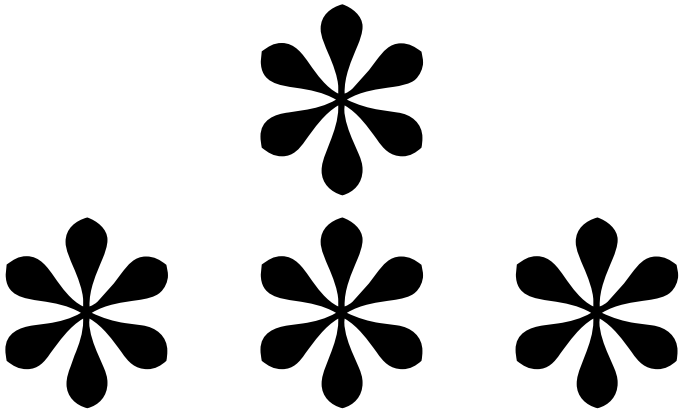
Chapter 4

Polynomials

This short chapter contains no linear algebra. It does contain the background material on polynomials that we will need in our study of linear maps from a vector space to itself. Many of the results in this chapter will already be familiar to you from other courses; they are included here for completeness. Because this chapter is not about linear algebra, your instructor may go through it rapidly. You may not be asked to scrutinize all the proofs. Make sure, however, that you at least read and understand the statements of all the results in this chapter—they will be used in the rest of the book.

Recall that **F** denotes **R** or **C**.



Degree

Recall that a function $p: \mathbf{F} \to \mathbf{F}$ is called a polynomial with coefficients in $\mathbf{F}$ if there exist $a_0, \ldots, a_m \in \mathbf{F}$ such that

$$p(z) = a_0 + a_1 z + a_2 z^2 + \cdots + a_m z^m$$

for all $z \in \mathbf{F}$. If p can be written in the form above with $a_m \neq 0$, then we say that p has degree m. If all the coefficients $a_0, \ldots, a_m$ equal 0, then we say that p has degree $-\infty$. For all we know at this stage, a polynomial may have more than one degree because we have not yet proved that the coefficients in the equation above are uniquely determined by the function p.

When necessary, use the obvious arithmetic with $-\infty$. For example, $-\infty < m$ and $-\infty + m = -\infty$ for every integer m. The 0 polynomial is declared to have degree $-\infty$ so that exceptions are not needed for various reasonable results. For example, the degree of pq equals the degree of p plus the degree of q even if $p = 0$.

Recall that $\mathcal{P}(\mathbf{F})$ denotes the vector space of all polynomials with coefficients in $\mathbf{F}$ and that $\mathcal{P}_m(\mathbf{F})$ is the subspace of $\mathcal{P}(\mathbf{F})$ consisting of the polynomials with coefficients in $\mathbf{F}$ and degree at most m. A number $\lambda \in \mathbf{F}$ is called a ***root*** of a polynomial $p \in \mathcal{P}(\mathbf{F})$ if

$$p(\lambda) = 0.$$

Roots play a crucial role in the study of polynomials. We begin by showing that λ is a root of p if and only if p is a polynomial multiple of $z - \lambda$.

4.1 Proposition: *Suppose $p \in \mathcal{P}(\mathbf{F})$ is a polynomial with degree $m \geq 1$. Let $\lambda \in \mathbf{F}$. Then λ is a root of p if and only if there is a polynomial $q \in \mathcal{P}(\mathbf{F})$ with degree $m - 1$ such that*

4.2 $$p(z) = (z - \lambda)q(z)$$

for all $z \in \mathbf{F}$.

PROOF: One direction is obvious. Namely, suppose there is a polynomial $q \in \mathcal{P}(\mathbf{F})$ such that 4.2 holds. Then

$$p(\lambda) = (\lambda - \lambda)q(\lambda) = 0,$$

and hence λ is a root of p, as desired.

To prove the other direction, suppose that $\lambda \in \mathbf{F}$ is a root of p. Let $a_0, \ldots, a_m \in \mathbf{F}$ be such that $a_m \neq 0$ and

$$p(z) = a_0 + a_1 z + a_2 z^2 + \cdots + a_m z^m$$

for all $z \in \mathbf{F}$. Because $p(\lambda) = 0$, we have

$$0 = a_0 + a_1\lambda + a_2\lambda^2 + \cdots + a_m\lambda^m.$$

Subtracting the last two equations, we get

$$p(z) = a_1(z - \lambda) + a_2(z^2 - \lambda^2) + \cdots + a_m(z^m - \lambda^m)$$

for all $z \in \mathbf{F}$. For each $j = 2, \ldots, m$, we can write

$$z^j - \lambda^j = (z - \lambda)q_{j-1}(z)$$

for all $z \in \mathbf{F}$, where q_{j-1} is a polynomial with degree $j - 1$ (specifically, take $q_{j-1}(z) = z^{j-1} + z^{j-2}\lambda + \cdots + z\lambda^{j-2} + \lambda^{j-1}$). Thus

$$p(z) = (z - \lambda)\underbrace{(a_1 + a_2q_2(z) + \cdots + a_mq_{m-1}(z))}_{q(z)}$$

for all $z \in \mathbf{F}$. Clearly q is a polynomial with degree $m - 1$, as desired. ■

Now we can prove that polynomials do not have too many roots.

4.3 Corollary: *Suppose $p \in \mathcal{P}(\mathbf{F})$ is a polynomial with degree $m \geq 0$. Then p has at most m distinct roots in $\mathbf{F}$.*

PROOF: If $m = 0$, then $p(z) = a_0 \neq 0$ and so p has no roots. If $m = 1$, then $p(z) = a_0 + a_1z$, with $a_1 \neq 0$, and p has exactly one root, namely, $-a_0/a_1$. Now suppose $m > 1$. We use induction on m, assuming that every polynomial with degree $m - 1$ has at most $m - 1$ distinct roots. If p has no roots in $\mathbf{F}$, then we are done. If p has a root $\lambda \in \mathbf{F}$, then by 4.1 there is a polynomial q with degree $m - 1$ such that

$$p(z) = (z - \lambda)q(z)$$

for all $z \in \mathbf{F}$. The equation above shows that if $p(z) = 0$, then either $z = \lambda$ or $q(z) = 0$. In other words, the roots of p consist of λ and the roots of q. By our induction hypothesis, q has at most $m - 1$ distinct roots in $\mathbf{F}$. Thus p has at most m distinct roots in $\mathbf{F}$. ■

The next result states that if a polynomial is identically 0, then all its coefficients must be 0.

4.4 Corollary: *Suppose $a_0, \dots, a_m \in \mathbf{F}$. If*

$$a_0 + a_1 z + a_2 z^2 + \cdots + a_m z^m = 0$$

for all $z \in \mathbf{F}$, then $a_0 = \cdots = a_m = 0$.

PROOF: Suppose $a_0 + a_1 z + a_2 z^2 + \cdots + a_m z^m$ equals 0 for all $z \in \mathbf{F}$. By 4.3, no nonnegative integer can be the degree of this polynomial. Thus all the coefficients equal 0. ■

The corollary above implies that $(1, z, \dots, z^m)$ is linearly independent in $\mathcal{P}(\mathbf{F})$ for every nonnegative integer m. We had noted this earlier (in Chapter 2), but now we have a complete proof. This linear independence implies that each polynomial can be represented in only one way as a linear combination of functions of the form z^j. In particular, the degree of a polynomial is unique.

If p and q are nonnegative integers, with $p \neq 0$, then there exist nonnegative integers s and r such that

$$q = sp + r.$$

and $r < p$. Think of dividing q by p, getting s with remainder r. Our next task is to prove an analogous result for polynomials.

Let $\deg p$ denote the degree of a polynomial p. The next result is often called the division algorithm, though as stated here it is not really an algorithm, just a useful lemma.

Think of 4.6 as giving the remainder r when q is divided by p.

4.5 Division Algorithm: *Suppose $p, q \in \mathcal{P}(\mathbf{F})$, with $p \neq 0$. Then there exist polynomials $s, r \in \mathcal{P}(\mathbf{F})$ such that*

4.6
$$q = sp + r$$

and $\deg r < \deg p$.

PROOF: Choose $s \in \mathcal{P}(\mathbf{F})$ such that $q - sp$ has degree as small as possible. Let $r = q - sp$. Thus 4.6 holds, and all that remains is to show that $\deg r < \deg p$. Suppose that $\deg r \geq \deg p$. If $c \in \mathbf{F}$ and j is a nonnegative integer, then

$$q - (s + cz^j)p = r - cz^j p.$$

Choose j and c so that the polynomial on the right side of this equation has degree less than $\deg r$ (specifically, take $j = \deg r - \deg p$ and then

choose c so that the coefficients of $z^{\deg r}$ in r and in $cz^j p$ are equal). This contradicts our choice of s as the polynomial that produces the smallest degree for expressions of the form $q - sp$, completing the proof. ■

Complex Coefficients

So far we have been handling polynomials with complex coefficients and polynomials with real coefficients simultaneously through our convention that $\mathbf{F}$ denotes $\mathbf{R}$ or $\mathbf{C}$. Now we will see some differences between these two cases. In this section we treat polynomials with complex coefficients. In the next section we will use our results about polynomials with complex coefficients to prove corresponding results for polynomials with real coefficients.

Though this chapter contains no linear algebra, the results so far have nonetheless been proved using algebra. The next result, though called the fundamental theorem of algebra, requires analysis for its proof. The short proof presented here uses tools from complex analysis. If you have not had a course in complex analysis, this proof will almost certainly be meaningless to you. In that case, just accept the fundamental theorem of algebra as something that we need to use but whose proof requires more advanced tools that you may learn in later courses.

4.7 Fundamental Theorem of Algebra: *Every nonconstant polynomial with complex coefficients has a root.*

This is an existence theorem. The quadratic formula gives the roots explicitly for polynomials of degree 2. Similar but more complicated formulas exist for polynomials of degree 3 and 4. No such formulas exist for polynomials of degree 5 and above.

PROOF: Let p be a nonconstant polynomial with complex coefficients. Suppose that p has no roots. Then $1/p$ is an analytic function on $\mathbf{C}$. Furthermore, $p(z) \to \infty$ as $z \to \infty$, which implies that $1/p \to 0$ as $z \to \infty$. Thus $1/p$ is a bounded analytic function on $\mathbf{C}$. By Liouville's theorem, any such function must be constant. But if $1/p$ is constant, then p is constant, contradicting our assumption that p is nonconstant. ■

The fundamental theorem of algebra leads to the following factorization result for polynomials with complex coefficients. Note that in this factorization, the numbers $\lambda_1, \ldots, \lambda_m$ are precisely the roots of p, for these are the only values of z for which the right side of 4.9 equals 0.

4.8 Corollary: *If $p \in \mathcal{P}(\mathbf{C})$ is a nonconstant polynomial, then p has a unique factorization (except for the order of the factors) of the form*

4.9 $$p(z) = c(z - \lambda_1) \ldots (z - \lambda_m),$$

where $c, \lambda_1, \ldots, \lambda_m \in \mathbf{C}$.

PROOF: Let $p \in \mathcal{P}(\mathbf{C})$ and let m denote the degree of p. We will use induction on m. If $m = 1$, then clearly the desired factorization exists and is unique. So assume that $m > 1$ and that the desired factorization exists and is unique for all polynomials of degree $m - 1$.

First we will show that the desired factorization of p exists. By the fundamental theorem of algebra (4.7), p has a root λ. By 4.1, there is a polynomial q with degree $m - 1$ such that

$$p(z) = (z - \lambda)q(z)$$

for all $z \in \mathbf{C}$. Our induction hypothesis implies that q has the desired factorization, which when plugged into the equation above gives the desired factorization of p.

Now we turn to the question of uniqueness. Clearly c is uniquely determined by 4.9—it must equal the coefficient of z^m in p. So we need only show that except for the order, there is only one way to choose $\lambda_1, \ldots, \lambda_m$. If

$$(z - \lambda_1) \ldots (z - \lambda_m) = (z - \tau_1) \ldots (z - \tau_m)$$

for all $z \in \mathbf{C}$, then because the left side of the equation above equals 0 when $z = \lambda_1$, one of the τ's on the right side must equal λ_1. Relabeling, we can assume that $\tau_1 = \lambda_1$. Now for $z \neq \lambda_1$, we can divide both sides of the equation above by $z - \lambda_1$, getting

$$(z - \lambda_2) \ldots (z - \lambda_m) = (z - \tau_2) \ldots (z - \tau_m)$$

for all $z \in \mathbf{C}$ except possibly $z = \lambda_1$. Actually the equation above must hold for all $z \in \mathbf{C}$ because otherwise by subtracting the right side from the left side we would get a nonzero polynomial that has infinitely many roots. The equation above and our induction hypothesis imply that except for the order, the λ's are the same as the τ's, completing the proof of the uniqueness. ■

Real Coefficients

Before discussing polynomials with real coefficients, we need to learn a bit more about the complex numbers.

Suppose $z = a + bi$, where a and b are real numbers. Then a is called the ***real part*** of z, denoted $\operatorname{Re} z$, and b is called the ***imaginary part*** of z, denoted $\operatorname{Im} z$. Thus for every complex number z, we have

$$z = \operatorname{Re} z + (\operatorname{Im} z)i.$$

The ***complex conjugate*** of $z \in \mathbf{C}$, denoted $\bar{z}$, is defined by

$$\bar{z} = \operatorname{Re} z - (\operatorname{Im} z)i.$$

Note that $z = \bar{z}$ if and only if z is a real number.

For example, $\overline{2 + 3i} = 2 - 3i$.

The ***absolute value*** of a complex number z, denoted $|z|$, is defined by

$$|z| = \sqrt{(\operatorname{Re} z)^2 + (\operatorname{Im} z)^2}.$$

For example, $|1 + 2i| = \sqrt{5}$. Note that $|z|$ is always a nonnegative number.

You should verify that the real and imaginary parts, absolute value, and complex conjugate have the following properties:

additivity of real part
$\operatorname{Re}(w + z) = \operatorname{Re} w + \operatorname{Re} z$ for all $w, z \in \mathbf{C}$;

additivity of imaginary part
$\operatorname{Im}(w + z) = \operatorname{Im} w + \operatorname{Im} z$ for all $w, z \in \mathbf{C}$;

sum of z and $\bar{z}$
$z + \bar{z} = 2 \operatorname{Re} z$ for all $z \in \mathbf{C}$;

difference of z and $\bar{z}$
$z - \bar{z} = 2(\operatorname{Im} z)i$ for all $z \in \mathbf{C}$;

product of z and $\bar{z}$
$z\bar{z} = |z|^2$ for all $z \in \mathbf{C}$;

additivity of complex conjugate
$\overline{w + z} = \bar{w} + \bar{z}$ for all $w, z \in \mathbf{C}$;

multiplicativity of complex conjugate
$\overline{wz} = \bar{w}\bar{z}$ for all $w, z \in \mathbf{C}$;

conjugate of conjugate
$\overline{\overline{z}} = z$ for all $z \in \mathbf{C}$;

multiplicativity of absolute value
$|wz| = |w|\,|z|$ for all $w, z \in C$.

In the next result, we need to think of a polynomial with real coefficients as an element of $\mathcal{P}(\mathbf{C})$. This makes sense because every real number is also a complex number.

A polynomial with real coefficients may have no real roots. For example, the polynomial $1 + x^2$ has no real roots. The failure of the fundamental theorem of algebra for $\mathbf{R}$ accounts for the differences between operators on real and complex vector spaces, as we will see in later chapters.

4.10 Proposition: *Suppose p is a polynomial with real coefficients. If $\lambda \in \mathbf{C}$ is a root of p, then so is $\bar{\lambda}$.*

PROOF: Let

$$p(z) = a_0 + a_1 z + \cdots + a_m z^m,$$

where $a_0, \ldots, a_m$ are real numbers. Suppose $\lambda \in \mathbf{C}$ is a root of p. Then

$$a_0 + a_1\lambda + \cdots + a_m\lambda^m = 0.$$

Take the complex conjugate of both sides of this equation, obtaining

$$a_0 + a_1\bar{\lambda} + \cdots + a_m\bar{\lambda}^m = 0,$$

where we have used some of the basic properties of complex conjugation listed earlier. The equation above shows that $\bar{\lambda}$ is a root of p. ■

We want to prove a factorization theorem for polynomials with real coefficients. To do this, we begin by characterizing the polynomials with real coefficients and degree 2 that can be written as the product of two polynomials with real coefficients and degree 1.

Think about the connection between the quadratic formula and this proposition.

4.11 Proposition: *Let $\alpha, \beta \in \mathbf{R}$. Then there is a polynomial factorization of the form*

4.12
$$x^2 + \alpha x + \beta = (x - \lambda_1)(x - \lambda_2),$$

with $\lambda_1, \lambda_2 \in \mathbf{R}$, if and only if $\alpha^2 \geq 4\beta$.

PROOF: Notice that

4.13
$$x^2 + \alpha x + \beta = (x + \frac{\alpha}{2})^2 + (\beta - \frac{\alpha^2}{4}).$$

First suppose that $\alpha^2 < 4\beta$. Then clearly the right side of the equation above is positive for every $x \in \mathbf{R}$, and hence the polynomial $x^2 + \alpha x + \beta$ has no real roots. Thus no factorization of the form 4.12, with $\lambda_1, \lambda_2 \in \mathbf{R}$, can exist.

Conversely, now suppose that $\alpha^2 \geq 4\beta$. Thus there is a real number c such that $c^2 = \frac{\alpha^2}{4} - \beta$. From 4.13, we have

$$\begin{aligned} x^2 + \alpha x + \beta &= (x + \frac{\alpha}{2})^2 - c^2 \\ &= (x + \frac{\alpha}{2} + c)(x + \frac{\alpha}{2} - c), \end{aligned}$$

which gives the desired factorization. ■

In the following theorem, each term of the form $x^2 + \alpha_j x + \beta_j$, with ${\alpha_j}^2 < 4\beta_j$, cannot be factored into the product of two polynomials with real coefficients and degree 1 (by 4.11). Note that in the factorization below, the numbers $\lambda_1, \ldots, \lambda_m$ are precisely the real roots of p, for these are the only real values of x for which the right side of the equation below equals 0.

4.14 Theorem: *If $p \in \mathcal{P}(\mathbf{R})$ is a nonconstant polynomial, then p has a unique factorization (except for the order of the factors) of the form*

$$p(x) = c(x - \lambda_1) \ldots (x - \lambda_m)(x^2 + \alpha_1 x + \beta_1) \ldots (x^2 + \alpha_M x + \beta_M),$$

where $c, \lambda_1, \ldots, \lambda_m \in \mathbf{R}$ and $(\alpha_1, \beta_1), \ldots, (\alpha_M, \beta_M) \in \mathbf{R}^2$ with ${\alpha_j}^2 < 4\beta_j$ for each j.

Here either m or M may equal 0.

PROOF: Let $p \in \mathcal{P}(\mathbf{R})$ be a nonconstant polynomial. We can think of p as an element of $\mathcal{P}(\mathbf{C})$ (because every real number is a complex number). The idea of the proof is to use the factorization 4.8 of p as a polynomial with complex coefficients. Complex but nonreal roots of p come in pairs; see 4.10. Thus if the factorization of p as an element of $\mathcal{P}(\mathbf{C})$ includes terms of the form $(x - \lambda)$ with λ a nonreal complex number, then $(x - \bar{\lambda})$ is also a term in the factorization. Combining these two terms, we get a quadratic term of the required form.

The idea sketched in the paragraph above almost provides a proof of the existence of our desired factorization. However, we need to be careful about one point. Suppose λ is a nonreal complex number

and $(x - \lambda)$ is a term in the factorization of p as an element of $\mathcal{P}(\mathbf{C})$. We are guaranteed by 4.10 that $(x - \bar{\lambda})$ also appears as a term in the factorization, but 4.10 does not state that these two factors appear the same number of times, as needed to make the idea above work. However, all is well. We can write

$$\begin{aligned} p(x) &= (x - \lambda)(x - \bar{\lambda})q(x) \\ &= (x^2 - 2(\operatorname{Re}\lambda)x + |\lambda|^2)q(x) \end{aligned}$$

for some polynomial $q \in \mathcal{P}(\mathbf{C})$ with degree two less than the degree of p. If we can prove that q has real coefficients, then, by using induction on the degree of p, we can conclude that $(x - \lambda)$ appears in the factorization of p exactly as many times as $(x - \bar{\lambda})$.

To prove that q has real coefficients, we solve the equation above for q, getting

Here we are not dividing by 0 because the roots of $x^2 - 2(\operatorname{Re}\lambda)x + |\lambda|^2$ are λ and $\bar{\lambda}$, neither of which is real.

$$q(x) = \frac{p(x)}{x^2 - 2(\operatorname{Re}\lambda)x + |\lambda|^2}$$

for all $x \in \mathbf{R}$. The equation above implies that $q(x) \in \mathbf{R}$ for all $x \in \mathbf{R}$. Writing

$$q(x) = a_0 + a_1 x + \cdots + a_{n-2}x^{n-2},$$

where $a_0, \ldots, a_{n-2} \in \mathbf{C}$, we thus have

$$0 = \operatorname{Im} q(x) = (\operatorname{Im} a_0) + (\operatorname{Im} a_1)x + \cdots + (\operatorname{Im} a_{n-2})x^{n-2}$$

for all $x \in \mathbf{R}$. This implies that $\operatorname{Im} a_0, \ldots, \operatorname{Im} a_{n-2}$ all equal 0 (by 4.4). Thus all the coefficients of q are real, as desired, and hence the desired factorization exists.

Now we turn to the question of uniqueness of our factorization. A factor of p of the form $x^2 + \alpha x + \beta$ with $\alpha^2 < 4\beta$ can be uniquely written as $(x - \lambda)(x - \bar{\lambda})$ with $\lambda \in \mathbf{C}$. A moment's thought shows that two different factorizations of p as an element of $\mathcal{P}(\mathbf{R})$ would lead to two different factorizations of p as an element of $\mathcal{P}(\mathbf{C})$, contradicting 4.8. ■

Exercises

1. Suppose m and n are positive integers with $m \leq n$. Prove that there exists a polynomial $p \in \mathcal{P}_n(\mathbf{F})$ with exactly m distinct roots.

2. Suppose that $z_1, \ldots, z_{m+1}$ are distinct elements of $\mathbf{F}$ and that $w_1, \ldots, w_{m+1} \in \mathbf{F}$. Prove that there exists a unique polynomial $p \in \mathcal{P}_m(\mathbf{F})$ such that
$$p(z_j) = w_j$$
for $j = 1, \ldots, m+1$.

3. Prove that if $p, q \in \mathcal{P}(\mathbf{F})$, with $p \neq 0$, then there exist unique polynomials $s, r \in \mathcal{P}(\mathbf{F})$ such that
$$q = sp + r$$
and $\deg r < \deg p$. In other words, add a uniqueness statement to the division algorithm (4.5).

4. Suppose $p \in \mathcal{P}(\mathbf{C})$ has degree m. Prove that p has m distinct roots if and only if p and its derivative p' have no roots in common.

5. Prove that every polynomial with odd degree and real coefficients has a real root.