Chapter 4
Polynomials

This chapter contains material on polynomials that we will use to investigate linear maps from a vector space to itself. Many results in this chapter will already be familiar to you from other courses; they are included here for completeness.

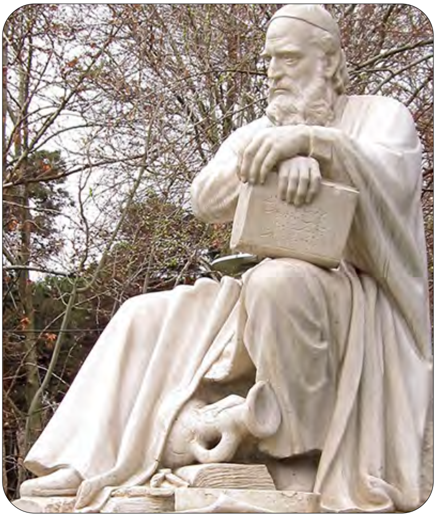
Because this chapter is not about linear algebra, your instructor may go through it rapidly. You may not be asked to scrutinize all the proofs. Make sure, however, that you at least read and understand the statements of all results in this chapter—they will be used in later chapters.

This chapter begins with a brief discussion of algebraic properties of the complex numbers. Then we prove that a nonconstant polynomial cannot have more zeros than its degree. We also give a linear-algebra-based proof of the division algorithm for polynomials, which is worth reading even if you are already familiar with a proof that does not use linear algebra.

As we will see, the fundamental theorem of algebra leads to a factorization of every polynomial into degree-one factors if the scalar field is **C** or to factors of degree at most two if the scalar field is **R**.

standing assumption for this chapter

- **F** denotes **R** or **C**.



Alireza Javaheri CC BY

Statue of mathematician and poet Omar Khayyam (1048–1131), whose algebra book written in 1070 contained the first serious study of cubic polynomials.

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

4.1 definition: *real part,* $\operatorname{Re} z$*, imaginary part,* $\operatorname{Im} z$

Suppose $z = a + bi$, where a and b are real numbers.

- The *real part* of z, denoted by $\operatorname{Re} z$, is defined by $\operatorname{Re} z = a$.
- The *imaginary part* of z, denoted by $\operatorname{Im} z$, is defined by $\operatorname{Im} z = b$.

Thus for every complex number z, we have

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

4.2 definition: *complex conjugate,* $\bar{z}$*, absolute value,* $|z|$

Suppose $z \in \mathbf{C}$.

- The *complex conjugate* of $z \in \mathbf{C}$, denoted by $\bar{z}$, is defined by
$$\bar{z} = \operatorname{Re} z - (\operatorname{Im} z)\, i.$$
- The *absolute value* of a complex number z, denoted by $|z|$, is defined by
$$|z| = \sqrt{(\operatorname{Re} z)^2 + (\operatorname{Im} z)^2}.$$

4.3 example: *real and imaginary part, complex conjugate, absolute value*

Suppose $z = 3 + 2i$. Then

- $\operatorname{Re} z = 3$ and $\operatorname{Im} z = 2$;
- $\bar{z} = 3 - 2i$;
- $|z| = \sqrt{3^2 + 2^2} = \sqrt{13}$.

Identifying a complex number $z \in \mathbf{C}$ with the ordered pair $(\operatorname{Re} z, \operatorname{Im} z) \in \mathbf{R}^2$ identifies $\mathbf{C}$ with $\mathbf{R}^2$. Note that $\mathbf{C}$ is a one-dimensional complex vector space, but we can also think of $\mathbf{C}$ (identified with $\mathbf{R}^2$) as a two-dimensional real vector space.

The absolute value of each complex number is a nonnegative number. Specifically, if $z \in \mathbf{C}$, then $|z|$ equals the distance from the origin in $\mathbf{R}^2$ to the point $(\operatorname{Re} z, \operatorname{Im} z) \in \mathbf{R}^2$.

The real and imaginary parts, complex conjugate, and absolute value have the properties listed in the following multipart result.

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

4.4 *properties of complex numbers*

Suppose $w, z \in \mathbf{C}$. Then the following equalities and inequalities hold.

sum of z and $\bar{z}$
$z + \bar{z} = 2\operatorname{Re} z$.

difference of z and $\bar{z}$
$z - \bar{z} = 2(\operatorname{Im} z)i$.

product of z and $\bar{z}$
$z\bar{z} = |z|^2$.

additivity and multiplicativity of complex conjugate
$\overline{w + z} = \bar{w} + \bar{z}$ and $\overline{wz} = \bar{w}\,\bar{z}$.

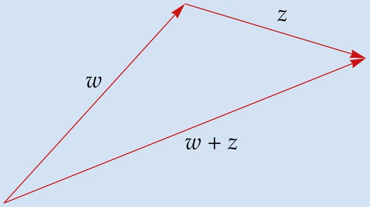
double complex conjugate
$\bar{\bar{z}} = z$.

real and imaginary parts are bounded by $|z|$
$|\operatorname{Re} z| \le |z|$ and $|\operatorname{Im} z| \le |z|$.

absolute value of the complex conjugate
$|\bar{z}| = |z|$.

multiplicativity of absolute value
$|wz| = |w|\,|z|$.

triangle inequality
$|w + z| \le |w| + |z|$.



Geometric interpretation of triangle inequality: The length of each side of a triangle is less than or equal to the sum of the lengths of the two other sides.

Proof Except for the last item above, the routine verifications of the assertions above are left to the reader. To verify the triangle inequality, we have

$$\begin{aligned}
|w + z|^2 &= (w + z)(\bar{w} + \bar{z}) \\
&= w\bar{w} + z\bar{z} + w\bar{z} + z\bar{w} \\
&= |w|^2 + |z|^2 + w\bar{z} + \overline{w\bar{z}} \\
&= |w|^2 + |z|^2 + 2\operatorname{Re}(w\bar{z}) \\
&\le |w|^2 + |z|^2 + 2|w\bar{z}| \\
&= |w|^2 + |z|^2 + 2|w|\,|z| \\
&= (|w| + |z|)^2.
\end{aligned}$$

Taking square roots now gives the desired inequality $|w + z| \le |w| + |z|$. ■

See Exercise 2 for the reverse triangle inequality.

Zeros of Polynomials

Recall that a function $p\colon \mathbf{F} \to \mathbf{F}$ is called a polynomial of degree m if there exist $a_0, \dots, a_m \in \mathbf{F}$ with $a_m \neq 0$ such that

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

for all $z \in \mathbf{F}$. A polynomial could have more than one degree if the representation of p in the form above were not unique. Our first task is to show that this cannot happen.

The solutions to the equation $p(z) = 0$ play a crucial role in the study of a polynomial $p \in \mathcal{P}(\mathbf{F})$. Thus these solutions have a special name.

4.5 definition: *zero of a polynomial*

A number $\lambda \in \mathbf{F}$ is called a *zero* (or *root*) of a polynomial $p \in \mathcal{P}(\mathbf{F})$ if

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

The next result is the key tool that we will use to show that the degree of a polynomial is unique.

4.6 *each zero of a polynomial corresponds to a degree-one factor*

Suppose m is a positive integer and $p \in \mathcal{P}(\mathbf{F})$ is a polynomial of degree m. Suppose $\lambda \in \mathbf{F}$. Then $p(\lambda) = 0$ if and only if there exists a polynomial $q \in \mathcal{P}(\mathbf{F})$ of degree $m - 1$ such that

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

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

Proof First suppose $p(\lambda) = 0$. Let $a_0, a_1, \dots, a_m \in \mathbf{F}$ be such that

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

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

$$p(z) = p(z) - p(\lambda) = a_1(z - \lambda) + \cdots + a_m\bigl(z^m - \lambda^m\bigr) \tag{4.7}$$

for all $z \in \mathbf{F}$. For each $k \in \{1, \dots, m\}$, the equation

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

shows that $z^k - \lambda^k$ equals $z - \lambda$ times some polynomial of degree $k - 1$. Thus 4.7 shows that p equals $z - \lambda$ times some polynomial of degree $m - 1$, as desired.

To prove the implication in the other direction, now suppose that there is a polynomial $q \in \mathcal{P}(\mathbf{F})$ such that $p(z) = (z - \lambda) q(z)$ for every $z \in \mathbf{F}$. Then $p(\lambda) = (\lambda - \lambda) q(\lambda) = 0$, as desired. ■

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

4.8 *degree m implies at most m zeros*

Suppose m is a positive integer and $p \in \mathcal{P}(\mathbf{F})$ is a polynomial of degree m. Then p has at most m zeros in $\mathbf{F}$.

Proof We will use induction on m. The desired result holds if $m = 1$ because if $a_1 \neq 0$ then the polynomial $a_0 + a_1 z$ has only one zero (which equals $-a_0/a_1$). Thus assume that $m > 1$ and the desired result holds for $m - 1$.

If p has no zeros in $\mathbf{F}$, then the desired result holds and we are done. Thus suppose p has a zero $\lambda \in \mathbf{F}$. By 4.6, there is polynomial $q \in \mathcal{P}(\mathbf{F})$ of degree $m - 1$ such that

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

for every $z \in \mathbf{F}$. Our induction hypothesis implies that q has at most $m - 1$ zeros in $\mathbf{F}$. The equation above shows that the zeros of p in $\mathbf{F}$ are exactly the zeros of q in $\mathbf{F}$ along with λ. Thus p has at most m zeros in $\mathbf{F}$. ■

The result above implies that the coefficients of a polynomial are uniquely determined (because if a polynomial had two different sets of coefficients, then subtracting the two representations of the polynomial would give a polynomial with some nonzero coefficients but infinitely many zeros). In particular, the degree of a polynomial is uniquely defined.

Recall that the degree of the 0 polynomial is defined to be $-\infty$. When necessary, use the expected 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 such as $\deg(pq) = \deg p + \deg q$.

Division Algorithm for Polynomials

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

$$p = sq + r$$

and $r < s$. Think of dividing p by s, getting quotient q with remainder r. Our next result gives an analogous result for polynomials. Thus the next result is often called the division algorithm for polynomials, although as stated here it is not really an algorithm, just a useful result.

The division algorithm for polynomials could be proved without using any linear algebra. However, as is appropriate for a linear algebra textbook, the proof given here uses linear algebra techniques and makes nice use of a basis of $\mathcal{P}_n(\mathbf{F})$, which is the $(n+1)$-dimensional vector space of polynomials with coefficients in $\mathbf{F}$ and of degree at most n.

Think of the division algorithm for polynomials as giving a remainder polynomial r when the polynomial p is divided by the polynomial s.