BACKGROUND ON POLYNOMIALS

JOSEPH W. IVERSON

This document contains background on polynomials, as necessary for graduate linear algebra. We might skip some proofs in lecture, and you can find them here instead. Our treatment is completely unoriginal to us, and it is based almost entirely on Chapter 4 of [3]. We do not claim to improve upon [3], and any errors that occur here are our own. Other standard references include [1, 2].

1. Polynomials and their arithmetic

We begin with the definition of a *field*. Our main examples will be $\mathbb{R}$ and $\mathbb{C}$.

Definition 1. Let R be a set equipped with *addition* and *multiplication* operations that map $R \times R \to R$. We say R is a **commutative ring with identity** if it has distinct elements $0 \neq 1$ for which the following axioms hold:

(A1) $(a+b)+c = a+(b+c)$ for every $a,b,c \in R$,
(A2) $a+0 = a = 0+a$ for every $a \in R$,
(A3) $a+b = b+a$ for every $a,b \in R$,
(A4) for every $a \in R$, there exists $-a \in R$ such that $a+(-a)=0$,
(M1) $(ab)c = a(bc)$ for every $a,b,c \in R$,
(M2) $a1 = a = 1a$ for every $a \in R$,
(M3) $ab = ba$ for every $a,b \in R$, and
(D) $a(b+c) = (ab)+(ac)$ for every $a,b,c \in R$.

A **field** is a commutative ring with identity that also satisfies:

(M4) for every $a \in R$ with $a \neq 0$, there exists $a^{-1} \in R$ such that $aa^{-1} = 1$.

Example 2. Each of $\mathbb{Q}$, $\mathbb{R}$, and $\mathbb{C}$ is a field with its usual addition and multiplication, but there are other, more exotic fields. For instance, if p is any prime, one can show that $\mathbb{Z}_p = \{0,1,\dots,p-1\}$ is a field under modular addition and multiplication.

We usually denote $\mathbb{F}$ for a field. In this document, we are less interested in the field itself and more interested in its corresponding *polynomials*.

Definition 3. Given a field $\mathbb{F}$, $\mathbb{F}[x]$ denotes the vector space of formal polynomials with unknown x and coefficients in $\mathbb{F}$. It has a basis consisting of formal monomials $1 = x^0, x^1, x^2, \dots$, and any nonzero $f \in \mathbb{F}[x]$ can be written uniquely as a finite linear combination of monomials

$$f(x) = c_0 + c_1 x + c_2 x^2 + \cdots + c_n x^n \qquad (c_1,\dots,c_n \in \mathbb{F})$$

with $c_n \neq 0$. Here, c_n is the **leading coefficient** of f, and f is said to have **degree** $\deg f = n$. (We do not consider the zero polynomial to have a degree.) A **linear** polynomial has degree 1, and a **constant** (or **scalar**) polynomial has degree 0. A **monic** polynomial has leading coefficient $c_n = 1$.

Version: February 11, 2023.

We emphasize that a polynomial is a formal linear combination of symbols, and is not itself a function. However, every polynomial $f \in \mathbb{F}[x]$ *determines* a function $\mathbb{F} \to \mathbb{F}$ in the obvious way (see Definition 9). It will turn out that the mapping of a polynomial to its associated function is injective if and only if $\mathbb{F}$ is infinite. Then for infinite fields, no trouble arises when identifying a polynomial with a function, but for finite fields no such identification is possible.

Polynomial arithmetic is defined as expected. Specifically, given polynomials $f(x) = \sum_{i=0}^{n} a_i x^i$ and $g(x) = \sum_{j=0}^{m} b_j x^j$, their **product** is

$$(fg)(x) = \sum_{i=0}^{n} \sum_{j=0}^{m} a_i b_j x^{i+j} = \sum_{k=0}^{m+n} \sum_{i=0}^{k} a_i b_{k-i} x^k,$$

where we interpret $a_i = 0$ for $i > n$ and $b_j = 0$ for $j > m$. The sum of polynomials is determined by the vector space structure of $\mathbb{F}[x]$.

Remark 4. We consider $\mathbb{F}$ as a subset of $\mathbb{F}[x]$ by identifying $c \in \mathbb{F}$ with the constant polynomial $cx^0 \in \mathbb{F}[x]$. This does not lead to any confusion: given $c \in \mathbb{F}$ and $f \in \mathbb{F}[x]$, we get the same thing whether we interpret cf as a product of polynomials or as the scalar c times the vector f.

We omit the straightforward (but lengthy) proofs of the next two propositions.

Proposition 5. $\mathbb{F}[x]$ *is a commutative ring with identity (see Definition 1).*

Proposition 6. *The following hold for any nonzero polynomials $f, g \in \mathbb{F}[x]$.*

(a) $fg \neq 0$.

(b) $\deg(fg) = \deg f + \deg g$.

(c) *The lead coefficient of fg is the product of the lead coefficients of f and g. In particular, if f and g are monic, then so is fg.*

(d) *Either $f + g = 0$ or else* $\deg(f+g) \leq \max\{\deg f, \deg g\}$.

As a consequence of Proposition 6(a), the multiplication in $\mathbb{F}[x]$ satisfies a *cancellation law.*

Proposition 7 (Cancellation)**.** *Let $f \in \mathbb{F}[x]$ be nonzero. If $g, h \in \mathbb{F}[x]$ satisfy $fg = fh$, then $g = h$.*

Proof. We have $f(g - h) = 0$, and so $g - h = 0$ by Proposition 6(a). □

It does not usually make sense to talk about a quotient of polynomials. However, we can perform long division on polynomials to divide and leave a remainder, just as you may have learned in middle school. That algorithm is embedded in the proof of the following theorem, which records one of the most important features of $\mathbb{F}[x]$.

Theorem 8 (Division algorithm)**.** *Let $f, d \in \mathbb{F}[x]$ be polynomials with $d \neq 0$. Then there exist unique polynomials $q, r \in \mathbb{F}[x]$ such that $f = dq + r$ and either $r = 0$ or else* $\deg r < \deg d$.

Proof. First we establish existence. We may assume that $f \neq 0$ and $\deg d \leq \deg f$, since otherwise we could take $q = 0$ and $r = f$. In what follows, we will repeatedly apply the following claim: given nonzero polynomials $g, h \in \mathbb{F}[x]$ with $\deg g \geq \deg h$, there exists $k \in \mathbb{F}[x]$ for which $g - hk$ is either zero or has $\deg(g - kh) < \deg g$. Indeed, we can write

$$g(x) = a_0 + a_1 x + \cdots + a_m x^m \quad \text{and} \quad h(x) = b_0 + b_1 x + \cdots + b_n x^n$$

with $m = \deg g \geq \deg h = n$, and then we can take $k(x) = \frac{a_m}{b_n} x^{m-n}$.

To begin the process, apply the claim to d and f to produce $q_1 \in \mathbb{F}[x]$ for which $r_1 := f - dq_1$ satisfies $r_1 = 0$ or $\deg r_1 < \deg f$. If $r_1 = 0$ or $\deg r_1 < \deg d$, then we are done. Otherwise, we have $\deg r_1 \geq \deg d$, and we can apply the claim again to produce $q_2 \in \mathbb{F}[x]$ for which $r_2 := r_1 - dq_2 = f - d(q_1 + q_2)$ satisfies $r_2 = 0$ or $\deg r_2 < \deg r_1$. Once again, if $r_2 = 0$ or $\deg r_2 < \deg d$, then we are done, and otherwise we repeat to find q_3 and $r_3 := r_2 - dq_3 = f - d(q_1 + q_2 + q_3)$. Continuing in this way, we eventually find r_k such that $r_k = 0$ or $\deg r_k < \deg d$, since the degree goes down every time. As soon as this happens, we define $r := r_k = f - d(q_1 + \cdots + q_k)$ and $q = q_1 + \cdots + q_k$. This proves existence.

For uniqueness, suppose we have another expression $f = dq' + r'$ with $r' = 0$ or $\deg r' < \deg d$. Then $dq' + r' = f = dq + r$, and so

$$d(q' - q) = r - r'. \tag{1}$$

It now suffices to prove $q' - q = 0$. Suppose not. Then the left-hand side of (1) is nonzero and has degree

$$\deg d + \deg(q' - q) \geq \deg d,$$

while the right-hand side is either zero or has degree

$$\deg(r - r') \leq \max\{\deg r, \deg r'\} < \deg d.$$

This is a contradiction. Therefore $q' - q = 0$, and (1) implies $r - r' = 0$ as well. □

2. Factorization and ideals

2.1. **Roots.**

Definition 9. Given a polynomial $f(x) = a_0 + a_1 x + \cdots + a_n x^n \in \mathbb{F}[x]$ and a constant $c \in \mathbb{F}$, we denote

$$f(c) := a_0 + a_1 c + \cdots + a_n c^n \in \mathbb{F}.$$

If $f(c) = 0$, then we say that c is a **root** (or a **zero**) of f.

The symbol $f(c)$ is a slight abuse of notation since f is a polynomial and not a function, but in practice, no confusion should arise. It is easy to check the following, and we omit the proof.

Proposition 10. *For any $f, g \in \mathbb{F}[x]$ and any $c \in \mathbb{F}$,*

$$(f + g)(c) = f(c) + g(c) \quad \text{and} \quad (fg)(c) = f(c)g(c).$$

Definition 11. If $d, q, p \in \mathbb{F}[x]$ and $p = dq$, then we say that d **divides** p, that d is a **factor** of p, that p is **divisible** by d, and that p is a **multiple** of d. In case $d \neq 0$, we also say that $q = p/d$ is the **quotient** of p by d. (Note that such a quotient is unique when it exists, by the uniqueness part of Theorem 8.)

Proposition 12. *For any $f \in \mathbb{F}[x]$ and any $c \in \mathbb{F}$, $x - c$ is a factor of f if and only if $f(c) = 0$.*

Proof. Put $d := x - c$ and apply the division algorithm to obtain $f = dq + r$, where r is the unique such remainder satisfying $r = 0$ or $\deg r = 0$. In particular, $r \in \mathbb{F}$ is a constant. Since $d(c) = 0$, when we evaluate at c we obtain $f(c) = r$. As such, d divides f if and only if $r = 0$, if and only if $f(c) = 0$. □

Definition 13. Let $f \in \mathbb{F}[x]$ be nonzero. The **multiplicity** of a scalar $c \in \mathbb{F}$ in f is the largest integer $m \geq 0$ for which $(x-c)^m$ divides f, where we interpret $(x-c)^0 = 1$. Thus, c is a root of f if and only if it has multiplicity at least 1 in f. When $f(c) = 0$, one also says c has multiplicity m "as a root" of f.

Proposition 14. *Let $c \in \mathbb{F}$ have multiplicities m and n in nonzero $f, g \in \mathbb{F}[x]$, respectively. Then c has multiplicity $m+n$ in fg.*

Proof. Let p be the multiplicity of c in fg. Find (necessarily nonzero) quotients $h, k, q \in \mathbb{F}[x]$ to write $f(x) = (x-c)^m h(x)$, $g(x) = (x-c)^n k(x)$, and

$$(x-c)^p q(x) = f(x)g(x) = (x-c)^{m+n} h(x)k(x).$$

Then $p \geq m+n$ since $(x-c)^{m+n}$ is a factor of $f(x)g(x)$, and we can divide the equation above to obtain

$$(x-c)^{p-m-n} q(x) = h(x)k(x). \tag{2}$$

Observe that $h(c) \neq 0$, since otherwise we could find $h' \in \mathbb{F}[x]$ to write $h(x) = (x-c)h'(x)$, and then the multiplicity of c in $f(x) = (x-c)^{m+1}h'(x)$ would be greater than m. Similarly, $k(c) \neq 0$. Then $h(c)k(c) \neq 0$, and $x-c$ does not divide $h(x)k(x)$. In (2), the exponent on $x-c$ must be 0, so $p = m+n$. □

Proposition 15. *A nonzero polynomial of degree n has at most n distinct roots, counting multiplicities. That is, if $c_1, \ldots, c_k \in \mathbb{F}$ are distinct roots of nonzero $f \in \mathbb{F}[x]$ with respective multiplicities $m_1, \ldots, m_k \geq 1$, and if $n = \deg(f)$, then $k \leq \sum_{i=1}^k m_i \leq n$.*

Proof. To begin, find (necessarily nonzero) $q_1 \in \mathbb{F}[x]$ to write $f(x) = (x-c_1)^{m_1} q_1(x)$. For each $i \in \{2, \ldots, k\}$, c_i has multiplicity m_i in $f(x)$ and 0 in $(x-c_1)^{m_1}$, so Proposition 14 implies c_i has multiplicity m_i in $q_1(x)$. Then we can find nonzero $q_2 \in \mathbb{F}[x]$ to write $q_1(x) = (x-c_2)^{m_2} q_2(x)$ and $f(x) = (x-c_1)^{m_1}(x-c_2)^{m_2} q_2(x)$. As above, $c_3, \ldots, c_k$ have respective multiplicities $m_3, \ldots, m_k$ in q_2. Then we can find nonzero $q_3(x)$ to write $q_2(x) = (x-c_3)^{m_3} q_3(x)$, and so on. Continuing in this way, we eventually find nonzero $q_k \in \mathbb{F}[x]$ for which

$$f(x) = (x-c_1)^{m_1} \cdots (x-c_k)^{m_k} q_k(x).$$

Now take degrees of both sides to find

$$n = \deg(q_k) + \sum_{i=1}^k m_i \geq \sum_{i=1}^k m_i \geq k,$$

as desired. □

2.2. **Ideals.**

The following notion plays a prominent role in the theory of polynomial rings.

Definition 16. An **ideal** of $\mathbb{F}[x]$ is a subset $M \subseteq \mathbb{F}[x]$ satisfying the following conditions:

(I1) $0 \in M$,

(I2) if $f \in M$ and $g \in M$, then $f + g \in M$, and

(I3) if $f \in M$ and $g \in \mathbb{F}[x]$ is arbitrary, then $fg \in M$.

Taking g to be a scalar in (I3), we find that every ideal is a subspace; conversely, every subspace M that satisfies (I3) is an ideal. Notice that we consider the entire space $\mathbb{F}[x]$ to be an ideal.

Example 17. Let V be a vector space over $\mathbb{F}$, and let $T \in L(V)$ be an operator. Given $f(x) = c_0 + c_1x + \cdots + c_nx^n \in \mathbb{F}[x]$, we define $f(T) := c_0I + c_1T + \cdots + c_nT^n$. It is easy to check that $(f+g)(T) = f(T)+g(T)$ and $(fg)(T) = f(T)g(T)$ whenever $f, g \in \mathbb{F}[x]$. The **annihilator**

$$\operatorname{Ann} T := \{f \in \mathbb{F}[x] : f(T) = 0\} \subseteq \mathbb{F}[x]$$

is an ideal. Indeed, (I1) is trivially satisfied, and (I2) is easy to see. For (I3), choose any $f \in \operatorname{Ann} T$ and $g \in \mathbb{F}[x]$, and observe that

$$(fg)(T) = f(T)g(T) = 0g(T) = 0.$$

Example 18. Given any $f \in \mathbb{F}[x]$, the set

$$\langle f\rangle := \{fg : g \in \mathbb{F}[x]\}$$

is easily seen to be an ideal. We call $\langle f\rangle$ the **principal ideal generated by** f.

Example 19. More generally, for any choice of finitely many polynomials $f_1, \ldots, f_n \in \mathbb{F}[x]$, the set

$$\langle f_1, \ldots, f_n\rangle := \{f_1g_1 + \cdots + f_ng_n : g_1, \ldots, g_n \in \mathbb{F}[x]\}$$

is an ideal, and it is said to be **generated by** $f_1, \ldots, f_n$.

Remark 20. Principal ideals encode the notion of division, since d divides p if and only if $p \in \langle d\rangle$. Moreover, since $\langle d\rangle$ is an ideal, it is easy to show that $p \in \langle d\rangle$ if and only if $\langle p\rangle \subseteq \langle d\rangle$.

In abstract algebra, an analog of Definition 16 defines ideals in any commutative ring. For general commutative rings, the ideals can be extremely complicated. The situation for $\mathbb{F}[x]$ is much nicer, and the division algorithm ensures that every nonzero ideal of $\mathbb{F}[x]$ enjoys a simple description, as in the theorem below. This will be the crucial property of polynomials for our study of operators on finite-dimensional vector spaces.

Theorem 21 (Every nonzero ideal is principal)**.** *If $M \subseteq \mathbb{F}[x]$ is a nonzero ideal, then there is a unique monic polynomial $p \in \mathbb{F}[x]$ for which $M = \langle p\rangle$.*

Proof. First we establish existence. Put $m := \min\{\deg f : f \in M,\ f \neq 0\}$, and let $g(x) = c_0 + c_1x + \cdots + c_mx^m \in M$ have degree m. Then $p = 1/c_mg$ is monic with degree m, and we claim that $M = \langle p\rangle$. The inclusion $\langle p\rangle \subseteq M$ is satisfied since M is an ideal and $p \in M$. In the other direction, choose any $f \in M$ and apply the division algorithm to write $f = pq + r$, where either $r = 0$ or $\deg r < \deg p = m$. Since M is an ideal and both f and p belong to M, so does $r = f - pq$. From the minimality of m, we conclude that $r = 0$, and therefore $f = pq \in \langle p\rangle$. This proves existence.

For uniqueness, suppose $p' \in M$ is another choice of monic generator for which $M = \langle p'\rangle$. Since $p \in M = \langle p'\rangle$, there exists $q' \in \mathbb{F}[x]$ such that $p = p'q'$. By Proposition 6(b) and the minimality of m, we have

$$m = \deg p = \deg p' + \deg q' \geq \deg p' \geq m,$$

and equality holds throughout. In particular, q' is scalar. Since p and p' are both monic with $p = p'q'$, we conclude that $q' = 1$. Therefore, $p' = p$. □

Example 22. Let V be a vector space over $\mathbb{F}$ with finite dimension n, and let $T \in L(V)$ be an operator. Then $I = T^0, T^1, T^2, \ldots, T^{n^2}$ are n^2+1 vectors in a vector space of dimension n^2, so they enjoy a nontrivial linear dependence

$$0 = c_0 I + c_1 T + c_2 T^2 + \cdots + c_{n^2} T^{n^2} = f(T),$$

where $f(x) = \sum_{j=0}^{n^2} c_j x^j \in \mathbb{F}[x]$ is not zero. By definition, $f \in \operatorname{Ann} T \subseteq \mathbb{F}[x]$. In particular, the annihilator is a nonzero ideal. The unique generator $p \in \operatorname{Ann} T$ from Theorem 21 is called the **minimal polynomial** of T. From the proof of Theorem 21, we see that $p \in \mathbb{F}[x]$ is uniquely determined by the following attributes:

(i) $p(T) = 0$,
(ii) p is monic, and
(iii) $\deg(p) = \min\{\deg f : f(T) = 0\}$.

An important consequence of Theorem 21 is that the ideal $\langle f_1, \ldots, f_n \rangle$ from Example 19 can also be generated by one polynomial instead of n.

Definition 23. Given finitely many polynomials $f_1, \ldots, f_n \in \mathbb{F}[x]$, not all of which are zero, their **greatest common divisor** is the unique monic polynomial $d =: \mathrm{gcd(f_1, \ldots, f_n)}$ for which $\langle f_1, \ldots, f_n \rangle = \langle d \rangle$.

This terminology is justified by the combination of (ii)–(iii) below.

Proposition 24. *If $f_1, \ldots, f_n \in \mathbb{F}[x]$ are not all zero, then $d := \gcd(f_1, \ldots, f_n)$ satisfies the following:*

(i) *there exist $g_1, \ldots, g_n \in \mathbb{F}[x]$ such that $d = f_1 g_1 + \cdots + f_n g_n$,*
(ii) *d divides each of $f_1, \ldots, f_n$, and*
(iii) *any other polynomial p that divides each of $f_1, \ldots, f_n$ also divides d.*

Furthermore, if d' is any monic polynomial that satisfies either (i)–(ii) or (ii)–(iii), then $d' = \gcd(f_1, \ldots, f_n)$.

Proof. First we prove (i)–(iii). Recall that d is the unique monic polynomial for which $\langle f_1, \ldots, f_n \rangle = \langle d \rangle$. With this in mind, (i) holds since $d \in \langle f_1, \ldots, f_n \rangle$, and (ii) holds since each $f_j \in \langle d \rangle$. For (iii), choose $h_1, \ldots, h_n$ such that $f_j = p h_j$ for each $j \in [n]$, and use (i) to write

$$d = p h_1 g_1 + \cdots + p h_n g_n = p(h_1 g_1 + \cdots + h_n g_n).$$

This gives (iii).

For the uniqueness statements, let $d' \in \mathbb{F}[x]$ be monic and divide each of $f_1, \ldots, f_n$, as in (ii). We will show that $d' = \gcd(f_1, \ldots, f_n)$ if it satisfies either (i) or (iii) with d suitably replaced by d'. First, we show that $\langle f_1, \ldots, f_n \rangle \subseteq \langle d' \rangle$. Indeed, choose $h_1, \ldots, h_n \in \mathbb{F}[x]$ such that $f_j = d' h_j$ for each $j \in [n]$. Given any $p \in \langle f_1, \ldots, f_n \rangle$, there exist $k_1, \ldots, k_n \in \mathbb{F}[x]$ to write

$$p = f_1 h_1 + \cdots + f_n h_n = d' k_1 h_1 + \cdots + d' k_n h_n = d'(k_1 h_1 + \cdots + k_n h_n).$$

As such, $p \in \langle d' \rangle$, as desired. If (i) holds for d', then $d' \in \langle f_1, \ldots, f_n \rangle$, and since the latter is an ideal it follows easily that $\langle d' \rangle \subseteq \langle f_1, \ldots, f_n \rangle$. Similarly, if (iii) holds for d', then in particular $d = \gcd(f_1, \ldots, f_n)$ divides d', so that $d' \in \langle d \rangle = \langle f_1, \ldots, f_n \rangle$ and $\langle d' \rangle \subseteq \langle f_1, \ldots, f_n \rangle$. Overall, if either (i) or (iii) holds, then $\langle d \rangle = \langle d' \rangle$, and the uniqueness in Theorem 21 implies that $d' = d$. □

Definition 25. Let $f_1, \ldots, f_n \in \mathbb{F}[x]$ be not all zero. We say they are **relatively prime**, or **coprime**, if $\gcd(f_1, \ldots, f_n) = 1$. Equivalently, $\langle f_1, \ldots, f_n \rangle = \mathbb{F}[x]$.

Since 1 divides everything, Proposition 24 implies that $f_1, \dots, f_n$ are relatively prime if and only if there exist $g_1, \dots, g_n \in \mathbb{F}[x]$ for which $f_1 g_1 + \cdots + f_n g_n = 1$.

2.3. **Prime factorization.** By this point, you may have noticed how closely our language surrounding divisibility mirrors the situation in $\mathbb{Z}$. In that setting, there is a straightforward way to find the greatest common divisor of any collection of integers: find their prime factorizations, and collect everything they have in common. Next, we develop the analogous technology for polynomials.

Definition 26. A polynomial $f \in \mathbb{F}[x]$ is **reducible over** $\mathbb{F}$ if there exist nonzero $g, h \in \mathbb{F}[x]$ both having degree at least 1 such that $f = gh$. If there are no such g and h, then f is **irreducible over** $\mathbb{F}$. A **prime in** $\mathbb{F}[x]$ is an irreducible polynomial in $\mathbb{F}$ that is not a scalar.

Example 27. Every linear polynomial $f(x) = a(x - c)$ is irreducible, since if $g, h \in \mathbb{F}[x]$ are polynomials with degrees at least 1, then $\deg(gh) \geq 2$ implies $gh \neq f$.

Example 28. A polynomial $f \in \mathbb{F}[x]$ of degree two is reducible in $\mathbb{F}[x]$ if and only if it has a root in $\mathbb{F}$. This is a consequence of Proposition 12. Explicitly, if f is reducible in $\mathbb{F}[x]$ and $f = gh$ with $\deg g, \deg h \geq 1$, then we can factor out a constant from gh to write $f(x) = c(x - a)(x - b)$ for some $a, b, c \in \mathbb{F}$. Then $a, b \in \mathbb{F}$ are roots of f. Conversely, if $a \in \mathbb{F}$ is a root of f, then Proposition 12 gives a factorization $f(x) = (x - a)h(x)$ for some $h \in \mathbb{F}[x]$, and since $\deg f = 2$ we conclude that $\deg h = 1$. Therefore, f is reducible in $\mathbb{F}[x]$.

Example 29. Irreducibility of a polynomial depends on the ambient field $\mathbb{F}$. For example, the polynomial $f(x) = x^2 + 1$ is irreducible in $\mathbb{R}[x]$ because it is has no roots in $\mathbb{R}$. But in the larger field $\mathbb{C}$, it has roots $\pm i$, so it is reducible in $\mathbb{C}[x]$. Explicitly, $f(x) = (x - i)(x + i)$.

Theorem 30. *If $f_1, \dots, f_n \in \mathbb{F}[x]$ are polynomials and $p \in \mathbb{F}[x]$ is a prime that divides $f_1 \cdots f_n$, then p divides at least one of $f_1, \dots, f_n$.*

Proof. We prove the case $n = 2$, and the general case follows by induction on n. Suppose p is a prime that divides $f_1 f_2$ but does not divide f_1. We must prove that p divides f_2. After replacing p by a suitable scalar multiple, we may assume it is monic. In that case, the only monic polynomials that divide p are 1 and p itself. In particular, $d := \gcd(p, f_1)$ is either p or 1, and since p does not divide f_1, we conclude that $d = 1$. Apply Proposition 24(i) to write $1 = gp + hf_1$ with $g, h \in \mathbb{F}[x]$. Multiplying both sides by f_2, we find $f_2 = gpf_2 + hf_1 f_2$. In this expression, p divides the right-hand side since it divides $f_1 f_2$. Therefore, p divides f_2. □

Theorem 31 (Prime factorization)**.** *Each nonscalar monic polynomial $f \in \mathbb{F}[x]$ can be factored as a product of monic primes $f = p_1 \cdots p_n$, and this factorization is unique up to reordering the factors $p_1, \dots, p_n$.*

Proof. We proceed by induction on $\deg f$. In the base case $\deg f = 1$, $f(x) = x - c$ is already prime, and by considering degrees it is easy to see that $f = f$ is its only factorization into monic primes. Now suppose $\deg f = n \geq 2$, and that we proved the theorem for all polynomials of degree at most $n - 1$.

First we prove existence. If f is irreducible, then it is already prime, and $f = f$ is a factorization. If f is reducible, choose a factorization $f = gh$ with $\deg g, \deg h \geq 1$.

Then g and h each have degree less than n, and after scaling both if necessary, we can assume they are monic. By the inductive hypothesis, each of g and h factors as a product of monic primes, and combining gives a factorization for f.

Now we prove uniqueness. Suppose we have two factorizations

$$f = p_1 \cdots p_m = q_1 \cdots q_k$$

with every $p_i, q_j \in \mathbb{F}[x]$ a monic prime. Then

$$n = \deg f = \sum_{i=1}^{m} \deg p_i = \sum_{j=1}^{k} \deg g_j, \tag{3}$$

and in particular, $m, k \geq 1$. Since p_m divides $q_1 \cdots q_k$, it must divide some q_j. Furthermore, since p_m and q_j are both monic primes, they must be equal. After rearranging the q's if necessary, we may assume that $p_m = q_k$. If either m or k equals 1, then (3) implies they both equal 1, in which case we are done. Now assume $m, k \geq 2$. Then $(p_1 \cdots p_{m-1})p_m = (q_1 \cdots q_{k-1})p_m$, and by the cancellation property Proposition 7, we conclude that

$$p_1 \cdots p_{m-1} = q_1 \cdots q_{k-1} =: g,$$

where $\deg g = \deg f - \deg p_m \leq n - 1$. By the inductive hypothesis, g has a unique factorization into monic primes up to reordering. As such, $m - 1 = k - 1$ and the list $q_1, \dots, q_{k-1}$ can be reordered to obtain $p_1, \dots, p_{m-1}$. Adding the factor $q_k = p_m$ back into these lists, we conclude that our two factorizations for f agree up to reordering. □

Remark 32. Let f be a monic prime with $\deg f \geq 1$. In the factorization of f into monic primes from Theorem 31, some of the factors may be repeated. Grouping them together, we obtain the **primary decomposition** $f = p_1^{n_1} \cdots p_r^{n_r}$ with $p_1, \dots, p_r$ distinct monic primes. This factorization is also unique up to reordering, and each $p_j^{n_j}$ is called a **primary factor** of f.

Primary decompositions determine divisibility. If g is another nonscalar monic polynomial, then g divides f if and only if $g = p_1^{m_1} \cdots p_r^{m_r}$ with $0 \leq m_j \leq n_j$ for each $j \in [r]$, if and only if every primary factor of g divides a primary factor of f (i.e., it has the same prime base p_j as some primary factor of f, and its exponent in the factorization of g is no larger than its exponent in the factorization of f).

This makes it easy to find the greatest common divisor of monic polynomials when their primary factorizations are known. If $f_1, \dots, f_n$ are nonscalar monic polynomials, then the primary decomposition of $\gcd(f_1, \dots, f_n)$ is obtained by collecting all the distinct primes that occur in the factorizations of $f_1, \dots, f_n$, and raising each to the highest power that divides all of $f_1, \dots, f_n$. In particular, $f_1, \dots, f_n$ are relatively prime if and only if there is no prime that appears in all of their primary decompositions. This implies the following result, which we will find useful in lecture.

Corollary 33. *Let $f \in \mathbb{F}[x]$ be a nonscalar monic polynomial with primary decomposition $f = p_1^{n_1} \cdots p_r^{n_r}$. For each $j \in [r]$, define $f_j = f/p_j^{n_j} = \prod_{i \neq j} p_i^{n_i}$. Then $f_1, \dots, f_r$ are relatively prime.*

Example 34. The *Fundamental Theorem of Algebra* states that every prime polynomial in $\mathbb{C}[x]$ is linear. As such, the primary decomposition of any nonscalar monic polynomial $f \in \mathbb{C}[x]$ takes the form

$$f(x) = (x - c_1)^{n_1} \cdots (x - c_r)^{n_r},$$

where $c_1, \ldots, c_r \in \mathbb{C}$ are distinct roots of f. As a consequence of Theorem 31, nonscalar polynomials are relatively prime in $\mathbb{C}[x]$ if and only if they have no common root. Corollary 33 states that this holds in particular for the polynomials $f_1, \ldots, f_r$ given by $f_j(x) = f(x)/(x - c_i)^{n_i} = \prod_{i \neq j}(x - c_i)^{n_i}$.

Example 35. The Fundamental Theorem of Algebra also implies a characterization of primes in $\mathbb{R}[x]$. Let $f(x) \in \mathbb{R}[x]$ be a nonscalar monic polynomial. When viewed as a polynomial in $\mathbb{C}[x]$, it factors as

$$f(x) = (x - c_1)^{n_1} \cdots (x - c_r)^{n_r},$$

with $c_1, \ldots, c_r \in \mathbb{C}$. Applying complex conjugation to the left-hand side does nothing since $f(x) \in \mathbb{R}[x]$, and when we apply it on the right-hand side we find

$$f(x) = (x - \overline{c_1})^{n_1} \cdots (x - \overline{c_r})^{n_r}.$$

By uniqueness of prime factorizations in $\mathbb{C}[x]$, we conclude that the roots $c_1, \ldots, c_r \in \mathbb{C}$ come in conjugate pairs. For every non-real root $c_j \in \mathbb{C}$, $(x - c_j)(x - \overline{c_j}) = x^2 - (c_j + \overline{c_j})x + |c_j|^2$ has real coefficients. As such, combining non-real linear factors of $f(x)$ with their complex conjugates produces a factorization

$$f(x) = (x^2 + a_1 x + b_1) \cdots (x^2 + a_m x + b_m)(x - d_1) \cdots (x - d_n)$$

with every factor in $\mathbb{R}[x]$. In case f is prime, we deduce that every prime monic in $\mathbb{R}[x]$ takes the form $x^2 + ax + b$ or $x - d$ with $a, b, d \in \mathbb{R}$, where the quadratic is prime if and only if it has no roots in $\mathbb{R}$. Consequently, the primary decomposition of a nonscalar monic polynomial $f \in \mathbb{R}[x]$ takes the form

$$f(x) = (x^2 + a_1 x + b_1)^{n_1} \cdots (x^2 + a_k x + b_k)^{n_k} (x - d_1)^{m_1} \cdots (x - d_r)^{m_r}$$

where the base factors $x^2 + a_i x + b_i$ and $x - d_j$ are distinct real polynomials, and $a_i^2 - 4b_i < 0$ for every $i \in [k]$.

References

[1] D. S. Dummit, R. M. Foote, Abstract algebra, 3rd edition, Wiley, 2003.
[2] J. Fraleigh, A first course in abstract algebra, 7th edition, Pearson, 2002.
[3] K. Hoffman, R. Kunze, Linear algebra, 2nd edition, Prentice–Hall, 1971.

Email address: jwi@iastate.edu

Department of Mathematics, Iowa State University, Ames, IA 50011