2.1 Polynomial Basics

Definition: A **polynomial** in the variable x has the following form:

$$f(x) = a_d x^d + a_{d-1} x^{d-1} + \cdots + a_1 x + a_0$$

where the **coefficients** $a_0, a_1, ..., a_d$ are elements of a field.

Note: We have seen three fields so far: $\mathbb{Q}, \mathbb{R}, \mathbb{C}$. We will see many other fields! The set of all polynomials with coefficients in a given field F will be denoted:

$$F[x]$$

so for example $\mathbb{Q}[x]$ is the set of all polynomials with rational coefficients.

Examples: Some polynomials have special names:

(a) The zero polynomial is $f(x) = 0$.

(b) The constants are $f(x) = a$ with $a \neq 0$.

(c) The linear polynomials are $f(x) = ax + b$, with $a \neq 0$.

(d) The quadratic polynomials are $f(x) = ax^2 + bx + c$, with $a \neq 0$.

(e) The cubic polynomials are $f(x) = ax^3 + bx^2 + cx + d$, with $a \neq 0$.

Padding polynomials: The two polynomials:

$$f(x) \quad \text{and} \quad 0x^d + f(x)$$

will be considered to be equivalent. That is why, for example, we do not consider $0x + b$ to be a linear polynomial. It is a constant that has been padded with the fake linear term $0x$, and similarly $0x^2 + ax + b$ is a padded linear polynomial, not a quadratic polynomial. In every equivalence class of polynomials except the zero polynomial, there is exactly one "unpadded" polynomial:

$$f(x) = a_d x^d + \cdots + a_0 \text{ with } a_d \neq 0$$

and the **degree** of this $f(x)$ (or any padding of it) is well-defined to be d. Thus only the zero polynomial does not have a well-defined degree (after you unpad all the zero coefficients, it completely disappears!). Some texts set the degree of the zero polynomial to be $-\infty$. We won't do that here. We will simply leave the degree of the zero polynomial undefined.

Examples: The special names correspond to low degree polynomials:

(b) The constants are the polynomials of degree 0.

(c) The linear polynomials are the polynomials of degree 1.

(d) The quadratic polynomials are the polynomials of degree 2.

(e) The cubic polynomials are the polynomials of degree 3.

Definition of Addition. This is done coefficient by coefficient:

$$\begin{array}{cccccccc} & a_d x^d & + & \cdots & + & a_0 \\ + & b_d x^d & + & \cdots & + & b_0 \\ \hline = & (a_d + b_d)x^d & + & \cdots & + & (a_0 + b_0) \end{array}$$

More on Addition: $F[x]$ is a vector space with (infinite) basis: $1, x, x^2, x^3, \ldots$. The addition is vector addition, which is associative and commutative with additive identity element 0, and additive inverses always exist. Also:

$$b(a_d x^n + \cdots + a_0) = (ba_d)x^n + \cdots + (ba_0)$$

is scalar multiplication (and F is often called the **scalar field**.)

Note: You may have only seen linear algebra in the case of the scalar field $\mathbb{R}$. In this course, it will be important to consider other scalar fields. See §3.1.

Definition of Multiplication. This is **not** just scalar multiplication above. It is determined by "foil" (the distributive law) and the rule for exponents $x^d x^e = x^{d+e}$. The bookkeeping may be done in the following way:

$$\begin{array}{cccccccccc} & & a_d x^d & + & \cdots & + & \cdots & + & a_1 x & + & a_0 \\ & \times & & & b_e x^e & + & \cdots & + & b_1 x & + & b_0 \\ \hline & & (a_d b_0)x^d & + & \cdots & + & \cdots & + & (a_1 b_0)x & + & a_0 b_0 \\ (a_d b_1)x^{d+1} & + & (a_{d-1} b_1)x^d & + & \cdots & + & \cdots & + & (a_0 b_1)x & & \\ & & \vdots & & & & & & & & \end{array}$$

just as you do the bookkeeping when you multiply many-digit numbers, adding up each of the columns under the bar. Notice that in the far left column, there will only be one term to add, namely $a_d b_e x^{d+e}$ just as on the far right there is only the constant term $a_0 b_0$, so the final answer looks like:

$$a_d b_e x^{d+e} + (a_d b_{e-1} + a_{d-1} b_e)x^{d+e-1} + \cdots + (a_1 b_0 + a_0 b_1)x + a_0 b_0$$

with a jumble of terms in the middle. In summation notation (from calculus) the final answer is written really simply like this:

$$\sum_{i=0}^{d} \sum_{j=0}^{e} a_i b_j x^{i+j}$$

Example: Calculate $(x^2+x+1)(x-1) = x^3-1$:

		x^2	$+$	x	$+$	1
	$\times$			x	$+$	(-1)
		$(-1)x^2$	$+$	$(-1)x$	$+$	(-1)
x^3	$+$	x^2	$+$	x		
x^3	$+$	$0x^2$	$+$	$0x$	$+$	(-1)

Multiplication is associative, commutative and distributes with addition. The constant 1 is the multiplicative identity element, but $F[x]$ is not a field because not every non-zero polynomial has a multiplicative inverse. In fact, only the constants have multiplicative inverses. All but the last of these statements follow from the rules of arithmetic for F, and can be seen quite elegantly with the summation notation (see Exercise 5-7). The last statement is a corollary of the following:

Proposition 2.1.1. *If $f(x)$ has degree d and $g(x)$ has degree e, then*

$$f(x)g(x) \text{ has degree } d+e$$

Proof: Unpad all the zero terms from $f(x)$ and $g(x)$. Then

$$f(x)g(x) = (a_d x^d + \cdots + a_0)(b_e x^e + \cdots + b_0) = a_d b_e x^{d+e} + \cdots + a_0 b_0$$

and $a_d \neq 0$ and $b_e \neq 0$ by the definition of degree. Since $a_d \neq 0$ and $b_e \neq 0$ and F is a field, there are multiplicative inverses $\frac{1}{a_d}$ and $\frac{1}{b_e}$ in F, and then $\left(\frac{1}{a_d}\right)\left(\frac{1}{b_e}\right)$ is the multiplicative inverse of $a_d b_e$, so $a_d b_e \neq 0$. Thus $d+e$ is the degree of $f(x)g(x)$.

Corollary 2.1.2. *Only the constants have multiplicative inverses in $F[x]$.*

Proof: If $f(x)$ and $g(x)$ have degrees d and e and are multiplicative inverses of each other, then $f(x)g(x) = 1$, which has degree 0, so $d+e=0$. But $d \geq 0$ and $e \geq 0$, so $d+e=0$ can only happen if both $d=0$ and $e=0$. That is, $f(x)$ and $g(x)$ can only be multiplicative inverses if they are both constants.

Corollary 2.1.3. *If $f(x)g(x) = 0$, then $f(x) = 0$ or $g(x) = 0$.*

Proof: If $f(x)g(x) = 0$, then its degree is undefined. By Proposition 2.1.1, this means that either the degree of $f(x)$ or the degree of $g(x)$ is undefined (otherwise, the degree of $f(x)g(x)$ would be $d+e$!). Thus either $f(x) = 0$ or $g(x) = 0$ (or both).

Corollary 2.1.4. *If $f(x) \neq 0$ and $f(x)g(x) = f(x)h(x)$, then $g(x) = h(x)$. (in other words, $f(x)$ can be cancelled from both sides).*

Proof: After subtracting and distributing: $f(x)(g(x)-h(x)) = 0$, so either $f(x) = 0$ or $g(x)-h(x) = 0$ by Corollary 2.1.3. But $f(x) \neq 0$ by assumption, so $g(x)-h(x) = 0$, and adding $h(x)$ to both sides gives $g(x) = h(x)$.

No multiplicative inverses (of non-constant polynomials) means we don't have an honest division. But as in §1.1 there is a consolation prize:

Division with Remainders: If $f(x)$ and $g(x)$ have degrees d and e, with $d \leq e$, then:

$$g(x) = f(x)q(x) + r(x)$$

where $r(x)$ is either 0 or else a polynomial of smaller degree than d.

I won't burden you with the proof. Suffice it to say that as with the natural numbers, you can, using induction, turn the familiar long division into a mathematical proof of division with remainders.

Example: In $\mathbb{Q}[x]$, long divide $x^3 + 1$ by $2x + 1$:

$$\begin{array}{rl} & \frac{1}{2}x^2 - \frac{1}{4}x + \frac{1}{8} \\ 2x+1 \quad & \overline{|x^3 + 0x^2 + 0x + 1} \\ & \underline{x^3 + \frac{1}{2}x^2} \\ & \quad -\frac{1}{2}x^2 + 0x \\ & \quad \underline{-\frac{1}{2}x^2 - \frac{1}{4}x} \\ & \qquad\qquad \frac{1}{4}x + 1 \\ & \qquad\qquad \underline{\frac{1}{4}x + \frac{1}{8}} \\ & \qquad\qquad\quad \frac{7}{8} \end{array}$$

to get the quotient $q(x) = \frac{1}{2}x^2 - \frac{1}{4}x + \frac{1}{8}$ and remainder $r(x) = \frac{7}{8}$.

Definitions: (a) $f(x)$ **divides** $g(x)$ if $g(x) = f(x)q(x)$ (with zero remainder). In this case, $f(x)$ is said to be a **factor** of $g(x)$.

(b) $f(x)$ is a **prime** polynomial of degree $d > 0$ if the only factors of $f(x)$ have either degree d, or else degree 0 (in this case, the constants are all considered to be uninteresting, and in particular are not primes!).

Examples: (a) All linear polynomials are prime.

(b) Whether a polynomial is a prime or not may depend upon the coefficients. For example, x^2+1 is prime in $\mathbb{R}[x]$, but not in $\mathbb{C}[x]$, where $x^2+1 = (x-i)(x+i)$.

The Fundamental Theorem for Polynomials: Each non-constant polynomial in $F[x]$ factors as a product of finitely many prime polynomials.

Proof: Let $S \subset \mathbb{N}$ be the set of **degrees** of all the polynomials that do not factor as a product of finitely many prime polynomials. Then $S = \emptyset$ or else S has a smallest element d, by the well-ordered axiom. If $f(x)$ is any polynomial of degree d, then either $f(x)$ is prime or else $f(x) = g(x)h(x)$ so that $g(x)$ and $h(x)$ have smaller degree. But then $g(x)$ and $h(x)$ must both be products of finitely many primes because their degrees are not elements of S, and so $f(x)$ itself factors as a product of finitely many primes. But this tells us that every polynomial of degree d must factor as a product of primes, so there can be no such d, and therefore $S = \emptyset$, meaning that all polynomials factor.

Euclid's Theorem: There are infinitely many primes in each $F[x]$.

Proof: Same as the proof for $\mathbb{N}$. Given a finite number of prime polynomials:

$$p_1(x), p_2(x), ..., p_n(x)$$

the fundamental theorem tells us we can factor the polynomial:

$$g(x) = p_1(x)p_2(x)....p_n(x) + 1$$

and each of the prime factors of $g(x)$ is "new" (not one of the $p_i(x)$) because none of the $p_i(x)$ divide $g(x)$. So however many primes we start with, we know there are more, and this can only be true if there are infinitely many.

Remark: Every field F we have seen so far is already infinite, so in this case there are already infinitely many **linear** polynomials:

$$f(x) = x + a$$

and Euclid's theorem isn't really telling us much. We will, however, soon see fields with only a finite number of elements, where Euclid's theorem is definitely telling us something interesting!

Euclid's Algorithm: Start with $f(x)$ and $g(x)$ of degrees $d \leq e$, and apply division with remainders according to the following prescription:

$$\begin{array}{rcrcl} g(x) & = & f(x)q_1(x) & + & r_1(x) \\ f(x) & = & r_1(x)q_2(x) & + & r_2(x) \\ r_1(x) & = & r_2(x)q_3(x) & + & r_3(x) \\ & \vdots & & & \end{array}$$

until we reach a remainder of zero. Then the last non-zero remainder $r_{k+1}(x)$ is a common divisor of $f(x)$ and $g(x)$ of greatest degree.

Remark: If $d(x)$ is a common divisor of $f(x)$ and $g(x)$, then so is any constant multiple of $d(x)$. So there is no single gcd of two polynomials.

Example: If $f(x) = x^6 + 1$ and $g(x) = x^{10} + 1$ then:

$$\begin{array}{rcrcl} x^{10} + 1 & = & (x^6 + 1)(x^4) & + & (-x^4 + 1) \\ x^6 + 1 & = & (-x^4 + 1)(-x^2) & + & (x^2 + 1) \\ -x^4 + 1 & = & (x^2 + 1)(-x^2 + 1) & & \end{array}$$

so $x^2 + 1$ is a common divisor of largest degree. But so are $2x^2 + 2$ and $\frac{1}{2}x^2 + \frac{1}{2}$.

Rational Functions: The set of rational functions with coefficients in F is

$$F(x) = \left\{ \text{equivalence classes of fractions } \tfrac{f(x)}{g(x)} \right\}$$

where $f(x)$ and $g(x)$ are elements of $F[x]$, and $g(x) \neq 0$.

This time, we define the equivalence relation on fractions by the:

Cross Multiplication Rule:

$$\frac{f(x)}{g(x)} \sim \frac{a(x)}{b(x)} \text{ if } f(x)b(x) = g(x)a(x)$$

Unlike the rational numbers, there is no nice geometric way of picturing the equivalence classes as lines through the origin. But this doesn't matter! The "algebraic" cross multiplication rule is all that we need to create the field of rational functions. Recall that we have to verify three properties before we are technically allowed to talk about equivalence classes. Namely:

(i) Reflexivity:

$$\frac{f(x)}{g(x)} \sim \frac{f(x)}{g(x)} \text{ because } f(x)g(x) = g(x)f(x) \text{ Check.}$$

(ii) Symmetry:

$$\text{If } \frac{f(x)}{g(x)} \sim \frac{a(x)}{b(x)} \text{ then } \frac{a(x)}{b(x)} \sim \frac{f(x)}{g(x)}$$

because $f(x)b(x) = g(x)a(x)$ rearranges as $a(x)g(x) = b(x)f(x)$. Check.

(iii) Transitivity:

$$\text{If } \frac{f(x)}{g(x)} \sim \frac{a(x)}{b(x)} \text{ and } \frac{a(x)}{b(x)} \sim \frac{c(x)}{d(x)} \text{ then } \frac{f(x)}{g(x)} \sim \frac{c(x)}{d(x)}$$

because, from $f(x)b(x) = g(x)a(x)$ and $a(x)d(x) = b(x)c(x)$ we conclude:

$$g(x)b(x)c(x) = g(x)a(x)d(x) = f(x)b(x)d(x)$$

and then we can cancel $b(x)$ from both sides using Corollary 5.4 to finally get: $f(x)d(x) = g(x)c(x)$. Check.

Now we can finish as with rational numbers.

The formulas for addition and multiplication are the same:

$$\left[\frac{f(x)}{g(x)}\right] + \left[\frac{a(x)}{b(x)}\right] = \left[\frac{f(x)b(x) + a(x)g(x)}{g(x)b(x)}\right]$$

$$\left[\frac{f(x)}{g(x)}\right]\left[\frac{a(x)}{b(x)}\right] = \left[\frac{f(x)a(x)}{g(x)b(x)}\right]$$

Using the cross multiplication rule, these formulas are seen to be well-defined in precisely the same way as the formulas for addition and multiplication were seen to be well-defined in §1.2. The rules of arithmetic are also verified in the same way, showing that $F(x)$ (like $\mathbb{Q}$) is a field.

Next we turn our attention in a completely different direction, to:

The Simplest Field. This has just the two elements 0 and 1:

$$F_2 = \{0, 1\}$$

and the only unusual thing about the field F_2 is the definition:

$$1 + 1 = 0$$

(the other additions and multiplications are the familiar ones) For polynomials with coefficients in F_2, you are allowed to make the "Freshman's mistake:"

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

Let's think about how some of the results we've been discussing play out for this field. First, we can list all the polynomials in low degrees!

Constants in $F_2[x]$: There is only 1 (0 isn't a constant)

Linear Polynomials in $F_2[x]$: There are two: x and $x + 1$

Quadratic Polynomials: There are four: $x^2, x^2 + 1, x^2 + x$ and $x^2 + x + 1$

Cubic Polynomials: There are 8 of them: $x^3,\ x^3 + 1,$

$$x^3 + x,\ x^3 + x + 1,\ x^3 + x^2,\ x^3 + x^2 + 1,\ x^3 + x^2 + x,\ x^3 + x^2 + x + 1$$

and there are exactly 2^d polynomials of each degree d. (Can you see why?)

Now for the **prime** polynomials:

Linear Primes: Linear polynomials are always prime.

Quadratic Primes: If a quadratic polynomial is **not** prime, then it must factor as a product of two linear polynomials. That means that if we form all the products of linear polynomials, then whatever is left over is prime!! There are three products we can take:

$$x \cdot x = x^2,\ x \cdot (x+1) = x^2 + x \text{ and } (x+1)^2 = x^2 + 1$$

and this leaves $x^2 + x + 1$ as the only quadratic prime.

Cubic Primes: Again, we notice that if a cubic polynomial is **not** prime, then it must factor as a product of a linear polynomial and a quadratic polynomial. If we look at our list, there are 2 linears and 4 quadratics, so it looks like we aren't going to have any cubic polynomials left over! But we do, because some of these products turn out to be the same. Let's see:

$$x \cdot x^2 = x^3\ (1), x \cdot (x^2 + 1) = x^3 + x\ (2), x \cdot (x^2 + x) = x^3 + x^2\ (3)$$

$$x \cdot (x^2 + x + 1) = x^3 + x^2 + x\ (4), (x+1) \cdot x^2 = x^3 + x^2\ (5)$$

$$(x+1)(x^2+1) = x^3+x^2+x+1 \ (6)$$

$$(x+1)(x^2+x) = x^3+x \ (7), (x+1)(x^2+x+1) = x^3+1 \ (8)$$

Indeed, (3)=(5) and (2)=(7) and there are two polynomials left out:

$$x^3+x+1 \text{ and } x^3+x^2+1 \text{ are the cubic primes!}$$

One could go on. A quartic (degree four) polynomial that is not prime must **either** factor as a linear times a cubic or as a quadratic times a quadratic. Again, it looks like there are plenty of products to use up all 16 quartic polynomials, but in fact many of the products are the same, and several are once again left over. Thus in this setting, Euclid's theorem is very powerful indeed, since it tells us that there are infinitely many primes, and in particular there are primes of larger and larger degrees in $F_2[x]$.

A Last Remark about this field F_2. The negation transformation here is a bit of a surprise, because $-0 = 0$ (as always) but also $-1 = 1$ since $1+1 = 0$. That is, the negation transformation on F_2 **does nothing!**. The smallest field where the negation transformation does something interesting will be:

$$F_3 = \{-1, 0, 1\}$$

but that is a story for the exercises.

2.1.1 Polynomial Exercises

5-1 Prove by induction that:

$$(x+c)^n =$$

$$x^n + \frac{n!}{(n-1)!1!}x^{n-1}c + \frac{n!}{(n-2)!2!}x^{n-2}c^2 + \cdots + \frac{n!}{1!(n-1)!}xc^{n-1} + c^n$$

(Hint: Where have we seen something like this before?)

5-2 Factor each of the following polynomials as a product of primes in four ways, regarding them first as elements of $\mathbb{Q}[x]$, then $\mathbb{R}[x], \mathbb{C}[x]$ and finally $F_2[x]$.

(a) x^2-1

(b) x^2+1

(c) x^3-1

(d) x^3+1

(e) x^4-1

(f) x^4+1

(g) x^5-1

(h) x^6-1

5-3 Find a common divisor of largest degree of each of the following pairs of polynomials. Does it matter whether we regard them as polynomials in $\mathbb{Q}[x]$ or as polynomials in $\mathbb{C}[x]$ or $F_2[x]$? If so, what's the difference?

(a) $x^9 - 1$ and $x^6 - 1$

(b) $x^9 + 1$ and $x^6 + 1$

(c) $x^{10} - 1$ and $x^8 - 1$

(d) $x^{10} + 1$ and $x^8 + 1$

5-4 Find all the prime quartic polynomials in $F_2[x]$.

5-5 (a) Find the only possible definitions of addition and multiplication that will make:

$$F_3 = \{-1, 0, 1\}$$

into a field (this is the second simplest field).

(b) Find all the prime quadratic polynomials in $F_3[x]$.

5-6 Explain why:

(a) There are no prime quadratic polynomials in $\mathbb{C}[x]$.

(b) The prime quadratic polynomials $ax^2 + bx + c$ in $\mathbb{R}[x]$ all satisfy:

$$b^2 - 4ac < 0$$

5-7 Let $p(x) = \sum_{i=0}^{d} a_i x^i$, $q(x) = \sum_{j=0}^{e} b_j x^j$ and $r(x) = \sum_{k=0}^{f} c_k x^k$ be three polynomials (written in summation notation). Then:

$$p(x)q(x) = \sum_{i=0}^{d} \sum_{j=0}^{e} a_i b_j x^{i+j} \text{ and}$$

$$q(x)p(x) = \sum_{j=0}^{e} \sum_{i=0}^{d} b_j a_i x^{j+i}$$

It is a fact of the summation notation that (finite) sums can be reversed:

$$\sum_{j=0}^{e} \sum_{i=0}^{d} b_j a_i x^{j+i} = \sum_{i=0}^{d} \sum_{i=0}^{d} b_j a_i x^{j+i}$$

and then $a_i b_j = b_j a_i$ and $i + j = j + i$ (for all i and j) explain the commutative law of multiplication for polynomials.

Using the above discussion as a model, use facts about summation notation to:

(a) State and explain the distributive law for polynomials.

(b) State and explain the associative law of multiplication for polynomials.