8. Math 235 Fall 2023, Worksheet 8: Some algebraic properties of polynomials

Polynomials and their many variants have been occupying mathematicians since the Middle Ages. Somewhere between 1/4 and 3/4 of a typical textbook on abstract algebra is devoted to the study of polynomials, and yet even some of the most basic questions that one might ask usually remain unanswered.[1]

This worksheet is not meant to be any comprehensive survey of the subject (this would be near-impossible), nor even an introduction to all its bases (we will treat only univariate polynomials with numbers as coefficients). Its goal is rather to expose some elementary properties and features that are often useful. Obviously, we can only scratch the surface of the subject. For further results, techniques and problems, see [Zeitz17, §5.4], [GelAnd17, §2.2], [Engel98, Chapter 10], [AndDos10, Chapters 10–11], [AndDos12, Chapters 10–11], [Barbea89] and [Prasol04] (listed roughly in the order of increasing sophistication).

As before, $\mathbb{N}$ means the set $\{0, 1, 2, \ldots\}$.

8.1. Definitions

I assume that you have already encountered polynomials in your life, although maybe not their precise definition. I will outline it without going into too many details (see [Grinbe19a, Chapter 7] for the latter). The following is more of a reminder than an introduction.

8.1.1. The informal definition

Informally speaking, a *polynomial with real coefficients* is an "expression" of the form $c_0 X^0 + c_1 X^1 + \cdots + c_n X^n$, where $c_0, c_1, \ldots, c_n$ are real numbers, and where X is "an indeterminate" (a symbol with no fixed value other than itself; but you will soon be able to substitute numbers and other things for it). Examples of such polynomials are

$$2X^0 + 7X^1 + \frac{3}{8}X^2 + \sqrt{2}X^3 + (-\pi)\, X^4 \qquad \text{and}$$
$$1X^0 + 2X^1 \qquad \text{and} \qquad 1X^0 + 0X^1 + (-1)\, X^2.$$

We treat such expressions as if they were actual sums of powers of X multiplied with real factors, and we allow them to be manipulated accordingly: For example, an addend of the form $0X^i$ can be omitted (or, conversely, inserted), an addend of the form $1X^i$ can be rewritten as X^i, and we can replace a "$+c_i X^i$" by a "$-\,(-c_i)\, X^i$" whenever we find it convenient. Thus, for example, the polynomial $1X^0 + 0X^1 +$

[1] Indeed, we saw such a question on worksheet #4: How can we factor a polynomial with integer coefficients into irreducible polynomials?

$(-1)\,X^2$ can be rewritten in any of the forms

$$X^0 + (-1)\,X^2, \qquad X^0 - 1X^2, \qquad X^0 - X^2$$

(as well as more complicated forms such as $1X^0 + 0X^1 + (-1)\,X^2 + 0X^3 + 0X^4$, because we can insert $0X^i$ addends). We can also put the monomials in any order (so, e.g., we can rewrite $X^3 + 2X^7$ as $2X^7 + X^3$). Moreover, we abbreviate the monomials X^0 and X^1 as 1 and X, respectively. Thus, the polynomial $X^0 - 2X^1 + X^2$ can be simplified to $1 - 2X + X^2$. We shall say a few words about how to make all this rigorous further below.

For theoretical purposes, we often write a polynomial $c_0X^0 + c_1X^1 + \cdots + c_nX^n$ as an infinite sum $c_0X^0 + c_1X^1 + c_2X^2 + \cdots$ by setting $c_i := 0$ for all $i > n$. For example, $X^0 - X^2 = 1X^0 + 0X^1 + (-1)\,X^2 + 0X^3 + 0X^4 + 0X^5 + \cdots$. When a polynomial is written in this infinite-sum form $c_0X^0 + c_1X^1 + c_2X^2 + \cdots$, the numbers $c_0, c_1, c_2, \ldots$ are called its *coefficients*.

Having introduced polynomials, we need to explain how they can be added and multiplied. This is best done using their infinite-sum form: Addition is defined by

$$\begin{aligned}&\left(c_0X^0 + c_1X^1 + c_2X^2 + \cdots\right) + \left(d_0X^0 + d_1X^1 + d_2X^2 + \cdots\right)\\ &= (c_0 + d_0)\,X^0 + (c_1 + d_1)\,X^1 + (c_2 + d_2)\,X^2 + \cdots\end{aligned}$$

(this is called "coefficientwise addition"), whereas multiplication is defined by the more complicated formula

$$\begin{aligned}&\left(c_0X^0 + c_1X^1 + c_2X^2 + \cdots\right) \cdot \left(d_0X^0 + d_1X^1 + d_2X^2 + \cdots\right)\\ &= e_0X^0 + e_1X^1 + e_2X^2 + \cdots, \qquad \text{where } e_n := \sum_{i=0}^{n} c_i d_{n-i} \text{ for each } n \in \mathbb{N}\end{aligned}$$

(this is the formula you would obtain if you expanded the product on the left hand side and used distributivity and $X^iX^j = X^{i+j}$ to collect equal powers of X together).

The *degree* of a polynomial $c_0X^0 + c_1X^1 + c_2X^2 + \cdots$ is defined to be the largest number $n \in \mathbb{N}$ such that $c_n \neq 0$. If no such n exists (i.e., if the polynomial is $0X^0 + 0X^1 + 0X^2 + \cdots$), then the degree is defined to be $-\infty$ (a symbol that is understood to be smaller than every integer). The degree of a polynomial P is denoted by $\deg P$.

Each real number $r \in \mathbb{R}$ is commonly identified with the polynomial $rX^0 + 0X^1 + 0X^2 + 0X^3 + \cdots$, which is known as a *constant polynomial* (and has degree 0 if $r \neq 0$ and degree $-\infty$ if $r = 0$).

Nothing forces us to use the letter X for the indeterminate; and indeed, many other symbols are commonly used in its stead, such as the lowercase x or t. On this worksheet, I will generally use uppercase letters (such as P, Q and R) for polynomials (and in particular, X for the indeterminate, which itself is a polynomial, namely the polynomial $0X^0 + 1X^1 + 0X^2 + 0X^3 + \cdots$), and lowercase letters for numbers. But you should be prepared to encounter other notations on the Putnam contest and elsewhere.

We let $\mathbb{R}[X]$ denote the set of all polynomials with real coefficients. As we know, these polynomials can be added and multiplied. They can also be subtracted, namely by the rule $P - Q = P + (-Q)$, where $-Q$ is $(-1)\,Q$. Explicitly, this boils down to subtracting them coefficientwise. We will soon see how they can be divided (with remainder).

All the above was stated for polynomials with real coefficients. Similarly, we can define polynomials with integer coefficients, polynomials with rational coefficients, and polynomials with complex coefficients. The sets of such polynomials are denoted by $\mathbb{Z}[X]$, $\mathbb{Q}[X]$ and $\mathbb{C}[X]$, respectively. More generally, if $\mathbb{K}$ is any set of numbers, then the notation $\mathbb{K}[X]$ means the set of all polynomials with coefficients in $\mathbb{K}$. Clearly, $\mathbb{Z}[X] \subseteq \mathbb{Q}[X] \subseteq \mathbb{R}[X] \subseteq \mathbb{C}[X]$.

One important feature of polynomials is the possibility of *substituting* things (e.g., numbers, matrices, other polynomials) into them (aka *evaluating* them at these things). Let me remind how this works: If $P = c_0X^0 + c_1X^1 + \cdots + c_nX^n$ is a polynomial, and if a is a "number-like thing" (i.e., a number, a square matrix, a polynomial, or more generally an element of a ring[2]), then $P(a)$ denotes the sum $c_0a^0 + c_1a^1 + \cdots + c_na^n$. Roughly speaking, $P(a)$ is obtained from P by "plugging a for each X". We say that $P(a)$ is obtained by *substituting a* into P or by *evaluating* P at a.

Since we can substitute numbers a into polynomials P, you may be tempted to view polynomials as functions. And indeed, polynomials *give rise to* functions: For example, any polynomial $P \in \mathbb{R}[X]$ gives rise to the function

$$\begin{aligned} f_P : \mathbb{R} &\to \mathbb{R}, \\ a &\mapsto P(a), \end{aligned}$$

called the *polynomial function* corresponding to P. Moreover, the polynomial P can be uniquely reconstructed from the polynomial function f_P (see Corollary 8.5.8 below). Analogous things hold for polynomials in $\mathbb{Z}[X]$, $\mathbb{Q}[X]$ and $\mathbb{C}[X]$. Thus, one might be tempted to equate a polynomial P with its corresponding polynomial function f_P, and perhaps even to **define** polynomials as such functions (as Axler does in his otherwise superb textbook [Axler23]). However, this is bad for two reasons: Firstly, it lets you forget that polynomials can take not just numbers but also other things (e.g., square matrices) as inputs. Secondly, this turns out to be the wrong definition of polynomials once you generalize them to other kinds of coefficients (e.g., coefficients in a finite field). I personally think of polynomials as things that lie somewhat upstream from functions, but can be made into functions at the slightest need and often can be recovered back from the latter functions.

[2] This will make sense if you know some abstract algebra. (In that case, you will realize that I'm being a bit sloppy here.)

8.1.2. The formal definition

The definition of a polynomial we gave above was not quite up to the modern standards of rigorous mathematics. It referred to a mystical "indeterminate" and to a vague concept of an "expression". There is a well-known way to make this definition bulletproof, which I will outline just for the sake of completeness; the details can be found in [Grinbe19a, Chapter 7]. This rigorous definition proceeds in multiple steps:

1. We let $\mathbb{K}$ be one of the four sets $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$ and $\mathbb{C}$. We define a *polynomial with coefficients in* $\mathbb{K}$ (that is, a polynomial with integer, rational, real or complex coefficients, depending on what $\mathbb{K}$ is) as an infinite sequence $(c_0, c_1, c_2, \ldots)$ of numbers in $\mathbb{K}$ such that only finitely many entries of this sequence are nonzero (i.e., only finitely many $i \in \mathbb{N}$ satisfy $c_i \neq 0$). For example, $\left(1, 0, 4, -1, \underbrace{0, 0, 0, \ldots}_{\text{zeroes}}\right)$ is a polynomial, but $(1, 1, 1, 1, \ldots)$ is not. This definition may look strange, but it is completely precise (you know what a sequence is, right?). Later on, such a polynomial-as-sequence $(c_0, c_1, c_2, \ldots)$ will be rewritten as $c_0 X^0 + c_1 X^1 + c_2 X^2 + \cdots$, so we will recover the usual notation for polynomials.

2. We define addition and multiplication of polynomials by the formulas

$$(c_0, c_1, c_2, \ldots) + (d_0, d_1, d_2, \ldots) = (c_0 + d_0,\ c_1 + d_1,\ c_2 + d_2,\ \ldots)$$

   and

$$(c_0, c_1, c_2, \ldots) \cdot (d_0, d_1, d_2, \ldots) = (e_0, e_1, e_2, \ldots),$$
$$\text{where } e_n := \sum_{i=0}^{n} c_i d_{n-i} \text{ for each } n \in \mathbb{N}.$$

   (These are exactly the usual definitions of addition and multiplication, rewritten in terms of infinite sequences.)

   We identify each number $r \in \mathbb{K}$ with the polynomial $\left(r, \underbrace{0, 0, 0, \ldots}_{\text{zeroes}}\right)$, and we define subtraction of polynomials by $P - Q := P + (-1)\, Q$. We also set $\dfrac{P}{r} := \dfrac{1}{r} P$ for any nonzero $r \in \mathbb{K}$ and any polynomial P.

3. If $P = (c_0, c_1, c_2, \ldots)$ is a polynomial, then the entries $c_0, c_1, c_2, \ldots$ are called its *coefficients*. Specifically, c_i is called its *i-th coefficient* (or its *coefficient before X^i*).

4. The *degree* of a polynomial $P = (c_0, c_1, c_2, \ldots)$ is defined to be the largest $n \in \mathbb{N}$ such that $c_n \neq 0$ (or, if no such n exists, then it is defined to be $-\infty$).

5. We define the *indeterminate* X to be the polynomial $\left(0, 1, \underbrace{0, 0, 0, \ldots}_{\text{zeroes}}\right)$.

6. We define powers of polynomials in the usual way: $P^i = \underbrace{PP \cdots P}_{i \text{ times}}$ and $P^0 = 1$.

7. If a is an appropriate "number-like thing" (we will explain what this means in a moment), and if $P = (c_0, c_1, c_2, \ldots)$ is a polynomial, then $P(a)$ denotes the sum $c_0 a^0 + c_1 a^1 + c_2 a^2 + \cdots$. This sum looks infinite, but actually it has only finitely many nonzero addends (since only finitely many c_i are nonzero), and thus it can be computed in finite time simply by throwing away its zero addends and summing the rest.

   What is an "appropriate" "number-like thing"? The short answer is "anything for which the sum $c_0 a^0 + c_1 a^1 + c_2 a^2 + \cdots$ would make sense". For example, a can be a number in $\mathbb{K}$, or a square matrix with entries in $\mathbb{K}$, or another polynomial with coefficients in $\mathbb{K}$. Unlike a function, a polynomial has no a-priori restriction on what can be substituted into it.

   The value $P(a)$ is called the *value* of P at a. It is said to be obtained by *substituting* a into P or by *evaluating* P at a.

Strictly speaking, some things need to be checked for this definition to work. For instance, it is not obvious that the multiplication of polynomials is associative – i.e., that $(PQ)R = P(QR)$ for any three polynomials P, Q, R. Likewise, some things that look obvious actually need some proof: For instance, if P and Q are two polynomials, and if a is a "number-like thing", then

$$(P+Q)(a) = P(a) + Q(a) \qquad \text{and} \qquad (1)$$

$$(PQ)(a) = P(a) \cdot Q(a) \qquad \text{and} \qquad (2)$$

$$(P(Q))(a) = P(Q(a)). \qquad (3)$$

You have probably used these facts tacitly many times before realizing that they are theorems rather than tautologies. Don't worry too much: Their proofs are not very hard (see [Grinbe19a, Theorem 7.2.5 and Theorem 7.6.3 and Proposition 7.6.14] for the facts just mentioned)[3].

It is also easy to see that if $P \in \mathbb{K}[X]$ is a polynomial, then $P(X) = P$. In other words, if $P = (c_0, c_1, c_2, \ldots)$ is a polynomial, then

$$P = P(X) = c_0 X^0 + c_1 X^1 + c_2 X^2 + \cdots.$$

This fact completes the connection between our rigorous definition of polynomials as sequences and our informal understanding of polynomials as "expressions".

[3] The proofs of (1) and (2) are very easy.

It is common to write polynomials P as "$P\left(X\right)$", even though the equality $P\left(X\right) = P$ shows that the "$\left(X\right)$" part is redundant. The advantage of this redundant notation is that it stresses the fact that P is a polynomial and that the indeterminate is being denoted by X.

Note that, as a consequence of our definition, two polynomials $P = \left(c_0, c_1, c_2, \ldots\right)$ and $Q = \left(d_0, d_1, d_2, \ldots\right)$ are equal if and only if their respective coefficients are equal (i.e., if and only if $c_i = d_i$ for all $i \in \mathbb{N}$). We will later see that (for $\mathbb{K}$ being one of the sets $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$ and $\mathbb{C}$) this is equivalent to equality of values (more precisely, to requiring that $P\left(x\right) = Q\left(x\right)$ for all $x \in \mathbb{K}$), but this is not obvious at this point (and does not generalize to more exotic cases such as $\mathbb{K}$ being a finite field).

8.1.3. A caveat

Keep in mind that a polynomial $P \in \mathbb{Q}\left[X\right]$ might be *integer-valued* (i.e., satisfy $P\left(a\right) \in \mathbb{Z}$ for all $a \in \mathbb{Z}$) without having integer coefficients (i.e., without belonging to $\mathbb{Z}\left[X\right]$). For instance, the polynomial $P = \dfrac{X\left(X-1\right)}{2} = \dfrac{-1}{2}X + \dfrac{1}{2}X^2$ is integer-valued (since $\dfrac{a\left(a-1\right)}{2} \in \mathbb{Z}$ for all $a \in \mathbb{Z}$), but has two non-integer coefficients. This again illustrates the downsides of thinking of polynomials as functions.

8.1.4. What we are missing

We are deliberately restricting ourselves to *univariate polynomials* (i.e., polynomials in one indeterminate) here, and specifically to those that have coefficients in $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$ or $\mathbb{C}$. The reason is that there is enough to say about these "simple" cases already! In particular, these cases cover the majority of polynomials that appear on contests like the Putnam (even though polynomials with coefficients in finite fields sometimes appear as well).

8.2. Basics about degrees and coefficients

We introduce some more notations regarding polynomials:

- If $P \in \mathbb{K}\left[X\right]$ is a polynomial (where $\mathbb{K}$ is one of $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$ and $\mathbb{C}$), then $\deg P$ denotes the degree of P.

- A polynomial P is called *nonzero* if it is distinct from the polynomial $0 = \left(0, 0, 0, \ldots\right)$. (This does not mean that all its values $P\left(a\right)$ are nonzero!)

- The *leading coefficient* of a nonzero polynomial $P = \left(c_0, c_1, c_2, \ldots\right)$ is its coefficient $c_{\deg P}$. In other words, it is the last nonzero coefficient of P.

- If $P = \left(c_0, c_1, c_2, \ldots\right)$ is a polynomial, and $i \in \mathbb{N}$, then $\left[X^i\right] P$ denotes the coefficient c_i of P.

For example, if $P = (X+1)^n$ for some $n \in \mathbb{N}$, then $\left[X^i\right] P = \dbinom{n}{i}$ by the binomial formula.

- The *constant coefficient* of a polynomial P is defined to be $\left[X^0\right] P$. This equals the value $P(0)$.
- A polynomial P is said to be *monic* if it is nonzero and its leading coefficient is 1.

  For instance, the polynomial $X^2 - 5X + 4$ is monic, but the polynomial $4X^2 - 5X + 1$ is not.
- As we know, the polynomials of degree ≤ 0 are called constant. The polynomials of degree 1 (or sometimes of degree ≤ 1) are called *linear*; the polynomials of degree 2 (or sometimes ≤ 2) are called *quadratic*; the polynomials of degree 3 (or sometimes ≤ 3) are called *cubic*.

Degrees and leading coefficients behave nicely when polynomials are added, subtracted and multiplied:[4]

Proposition 8.2.1. Let $\mathbb{K}$ be one of the sets $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$ and $\mathbb{C}$. Let P and Q be two polynomials in $\mathbb{K}[X]$. Then:

(a) We have $\deg(P+Q) \leq \max\{\deg P, \deg Q\}$ and $\deg(P-Q) \leq \max\{\deg P, \deg Q\}$.

(b) We have $\deg(PQ) = \deg P + \deg Q$.

(c) If P and Q are nonzero, then PQ is nonzero as well, and the leading coefficient of PQ is the product of the leading coefficients of P and Q.

(d) If P and Q are monic, then PQ is also monic.

(e) If $PQ = 0$, then $P = 0$ or $Q = 0$.

(f) The constant coefficient of PQ is the product of the constant coefficients of P and Q. In other words, $\left[X^0\right](PQ) = \left[X^0\right] P \cdot \left[X^0\right] Q$.

Proof. This is essentially [Grinbe19a, Theorem 7.4.7], but the proof is easy enough that you don't need a reference. Do note that parts **(b)**, **(c)** and **(e)** rely on the fact that a product of two nonzero numbers is nonzero! □

[4] The symbol $-\infty$ is understood to be smaller than every integer.

Sums of the form $(-\infty) + k$ or $k + (-\infty)$ (where k is an integer or $-\infty$) are understood to be $-\infty$.

8.3. Division with remainder

We now know that polynomials can be added, subtracted and multiplied. But the type of polynomials we are considering – that is, univariate polynomials – have an additional feature: They can also be divided with remainder, just like integers. To be more precise, polynomials in $\mathbb{Q}[X]$, $\mathbb{R}[X]$ and $\mathbb{C}[X]$ can be divided with remainder. For polynomials in $\mathbb{Z}[X]$, the situation is more complicated, and we discuss this briefly below.

We first state the main result for $\mathbb{Q}[X]$, $\mathbb{R}[X]$ and $\mathbb{C}[X]$ (see [Grinbe19a, Theorem 7.5.4] for a rigorous proof, but it is actually quite easy):

Theorem 8.3.1. Let $\mathbb{K}$ be one of the sets $\mathbb{Q}$, $\mathbb{R}$ and $\mathbb{C}$. Let P and N be two polynomials in $\mathbb{K}[X]$, where N is nonzero. Then, there exists a **unique** pair (Q, R) of two polynomials in $\mathbb{K}[X]$ satisfying

$$P = QN + R \qquad \text{and} \qquad \deg R < \deg N.$$ [5]

The two entries of this pair (Q, R) have names: The first entry Q is called the *quotient* of P upon division by N, whereas the second entry R is called the *remainder* of P upon division by N. We also use the notations

$$P//N := Q \qquad \text{and} \qquad P\%N := R$$

for them. Of course, this terminology is analogous to the standard terminology used for quotients and remainders of integers (Definition 2.5.1 on Worksheet 2), and the condition $\deg R < \deg N$ on the remainder is the analogue of the $r \in \{0, 1, \ldots, n-1\}$ condition on the remainder for integers.

Example 8.3.2. Let $\mathbb{K} = \mathbb{Q}$ and $P = 2X^6 + 3X^2 - 7$ and $N = 2X^3 - 5$. Then,

$$P//N = X^3 + \frac{5}{2} \qquad \text{and} \qquad P\%N = 3X^2 + \frac{11}{2},$$

because the pair $(Q, R) = \left(X^3 + \frac{5}{2}, 3X^2 + \frac{11}{2}\right)$ satisfies $P = QN + R$ and $\deg R < \deg N$.

The way to find this pair (both in this example and in general) is to keep reducing the degree of P by subtracting appropriate multiples of N (every time choosing the multiple whose subtraction will cancel the leading coefficient of P), until the degree falls below $\deg N$. This method is called *polynomial long division* and is similar to the analogous method for integers.

[5] The inequality $\deg R < \deg N$ allows for the case $R = 0$, in which case $\deg R = -\infty$ is automatically smaller than $\deg N$.

This method works if $\mathbb{K}$ is $\mathbb{Q}$, $\mathbb{R}$ or $\mathbb{C}$ (it does not really care what $\mathbb{K}$ is, as long as we can add, subtract, multiply and divide numbers in $\mathbb{K}$). However, as we saw in Example 8.3.2, if we apply this method to two polynomials with integer coefficients, it can produce non-integer coefficients in the output, since the "appropriate multiples of N" that must be subtracted can be non-integer multiples. Thus, Theorem 8.3.1 does not hold for $\mathbb{K} = \mathbb{Z}$. However, it can be salvaged for $\mathbb{K} = \mathbb{Z}$ if we require N to be monic:

Theorem 8.3.3. Let P and N be two polynomials in $\mathbb{Z}[X]$, where N is **monic**. Then, there exists a **unique** pair (Q, R) of two polynomials in $\mathbb{Z}[X]$ satisfying

$$P = QN + R \qquad \text{and} \qquad \deg R < \deg N.$$

In other words, the polynomials $P//N$ and $P\%N$ belong to $\mathbb{Z}[X]$.

(Sometimes, of course, this claim will hold even if N is not monic, but this cannot be guaranteed.)

Just like for integers, we can define divisibility and congruence for polynomials:

Definition 8.3.4. Let $\mathbb{K}$ be one of the sets $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$ and $\mathbb{C}$.

(a) Given two polynomials P and Q in $\mathbb{K}[X]$, we write $P \mid Q$ (and say that P *divides* Q, or that Q is *divisible by* P) if and only if there exists a polynomial $F \in \mathbb{K}[X]$ satisfying $Q = PF$. To be more precise, we write "$P \mid Q$ in $\mathbb{K}[X]$" in this case, because this relation depends on $\mathbb{K}$.

(b) Given three polynomials P, Q and N in $\mathbb{K}[X]$, we write $P \equiv Q \bmod N$ (and say that P is *congruent to* Q *modulo* N) if and only if $N \mid P - Q$ in $\mathbb{K}[X]$. To be more precise, we write "$P \equiv Q \bmod N$ in $\mathbb{K}[X]$" in this case, because this relation depends on $\mathbb{K}$.

Just as an example of how these relations depend on $\mathbb{K}$, note that we have $2X \mid X^2 - X$ in $\mathbb{Q}[X]$ (since $X^2 - X = 2X \cdot \dfrac{X-1}{2}$), but we don't have $2X \mid X^2 - X$ in $\mathbb{Z}[X]$ (since $\dfrac{X-1}{2}$ does not belong to $\mathbb{Z}[X]$). Similarly, $X^2 \equiv X \bmod 2X$ in $\mathbb{Q}[X]$ but not in $\mathbb{Z}[X]$.

Divisibility and congruence satisfy the same rules for polynomials as they do for integers (in particular, all claims in §2.2 and §2.4 of Worksheet 2 hold equally well for polynomials). For example, congruences can be multiplied (i.e., if $A \equiv B \bmod N$ and $C \equiv D \bmod N$, then $AC \equiv BD \bmod N$), and so can divisibilities (i.e., if $A \mid B$ and $C \mid D$, then $AC \mid BD$).

Exactly as for integers, congruence for polynomials is closely related to remainders:

Proposition 8.3.5. Let $\mathbb{K}$ be one of the sets $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$ and $\mathbb{C}$. Let P, Q and N be three polynomials in $\mathbb{K}[X]$, where N is nonzero. If $\mathbb{K} = \mathbb{Z}$, then we furthermore assume that N is monic.

Then, $P \equiv Q \bmod N$ if and only if $P\%N = Q\%N$.

Proof idea. Just like for integers. □

Exercise 8.3.1. Let $n \in \mathbb{N}$. Compute the remainder $X^n\% \left(X^2 - X - 1\right)$.

Solution idea. For the sake of brevity, let us set $N := X^2 - X - 1 \in \mathbb{Z}[X]$. This is a monic polynomial, so that Proposition 8.3.5 can be used (with $\mathbb{K} = \mathbb{Z}$).

We are looking for the remainder $X^n\% \left(X^2 - X - 1\right) = X^n\%N$. Computing this remainder for small values of n might give us an idea:

$$\begin{array}{l|l|l}
X^0\%N = 1; & X^3\%N = 2X + 1; & X^6\%N = 8X + 5; \\
X^1\%N = X; & X^4\%N = 3X + 2; & X^7\%N = 13X + 8; \\
X^2\%N = X + 1; & X^5\%N = 5X + 3; & X^8\%N = 21X + 13.
\end{array}$$

Note that the easiest way to compute all these is by working modulo N: Indeed, a polynomial of degree ≤ 1 that is congruent to X^n modulo N must automatically be the remainder $X^n\%N$ (why?). Hence, in order to find the latter remainder, we just need to simplify X^n modulo N. We can do this by repeatedly replacing X^2 by $X + 1$ (since $X^2 \equiv X + 1 \bmod N$); here, for instance, is this procedure for $n = 4$:

$$\begin{aligned}
X^4 &= \underbrace{X^2}_{\equiv X+1 \bmod N} X^2 \equiv (X+1)\,X^2 = X^3 + X^2 = \underbrace{X^2}_{\equiv X+1 \bmod N} X + \underbrace{X^2}_{\equiv X+1 \bmod N} \\
&\equiv (X+1)\,X + (X+1) = \underbrace{X^2}_{\equiv X+1 \bmod N} + 2X + 1 \\
&\equiv X + 1 + 2X + 1 = 3X + 2 \bmod N,
\end{aligned}$$

so that $X^4\%N = 3X + 2$. Moreover, once we have simplified X^n modulo N, we can easily simplify X^{n+1} as well (by multiplying the result for X^n by X and then doing one more X^2-replacement); thus, we need not start the computation from scratch for each new n.

At this point, you will have realized what the answer must be: For each $n \geq 1$, we have

$$X^n\%N = f_n X + f_{n-1}, \qquad (4)$$

where $(f_0, f_1, f_2, \ldots)$ is the Fibonacci sequence (defined recursively by $f_0 = 0$ and $f_1 = 1$ and $f_n = f_{n-1} + f_{n-2}$ for all $n \geq 2$). (The case $n = 0$ also fits under this pattern if we set $f_{-1} := 1$.)

It remains to prove (4). There are several ways to do so; the shortest is probably the following:

We shall first show that

$$X^n \equiv f_n X + f_{n-1} \bmod N \qquad \text{for each } n \geq 1. \tag{5}$$

[*Proof of (5):* We proceed by induction on n.

Base case: The claim (5) is true for $n = 1$, since $\underbrace{f_1}_{=1} X + \underbrace{f_{1-1}}_{=f_0=0} = X$ is literally equal to X^1.

Induction step: Let $n \geq 1$ be an integer. Assume (as the induction hypothesis) that (5) holds for this n. We must then prove that (5) holds for $n+1$ as well, i.e., that we have $X^{n+1} \equiv f_{n+1} X + f_{(n+1)-1} \bmod N$.

By the induction hypothesis, we know that (5) holds for n, so that we have $X^n \equiv f_n X + f_{n-1} \bmod N$. Furthermore, $X^2 - (X+1) = X^2 - X - 1 = N$ is clearly divisible by N, so that we have $X^2 \equiv X + 1 \bmod N$. Thus,

$$\begin{aligned} X^{n+1} &= \underbrace{X^n}_{\equiv f_n X + f_{n-1} \bmod N} X \equiv (f_n X + f_{n-1}) X = f_n \underbrace{X^2}_{\equiv X+1 \bmod N} + f_{n-1} X \\ &\equiv f_n (X+1) + f_{n-1} X = \underbrace{(f_n + f_{n-1})}_{\substack{= f_{n+1} \\ \text{(by the definition of the} \\ \text{Fibonacci sequence)}}} X + \underbrace{f_n}_{= f_{(n+1)-1}} \\ &= f_{n+1} X + f_{(n+1)-1} \bmod N. \end{aligned}$$

This proves that (5) holds for $n+1$. Thus, the induction step is complete, and (5) is proved by induction.]

Now, let $n \geq 1$ be an integer. Let R be the polynomial $f_n X + f_{n-1} \in \mathbb{Z}[X]$; its degree is $\deg R \leq 1 < 2$. Now, (5) shows that $X^n \equiv f_n X + f_{n-1} = R \bmod N$. In other words, $X^n - R$ is divisible by N. In other words, there exists a polynomial $Q \in \mathbb{Z}[X]$ such that $X^n - R = NQ$. Consider this Q. From $X^n - R = NQ$, we obtain

$$X^n = NQ + R = QN + R.$$

Since $\deg R < 2 = \deg N$, this equality shows that Q is the quotient and R the remainder upon division X^n by N. In other words, $Q = X^n // N$ and $R = X^n \% N$. In particular, $X^n \% N = R = f_n X + f_{n-1}$. This proves (4), so the problem is solved.

Remark: We can also compute the quotient $X^n // N$: It equals $\sum\limits_{k=0}^{n-2} f_{n-k-1} X^k$. The simplest proof of this is by showing that

$$X^n = \left(\sum_{k=0}^{n-2} f_{n-k-1} X^k \right) N + (f_n X + f_{n-1}) \qquad \text{for each } n \geq 1.$$

(This can be shown, e.g., by induction on n, or directly using a form of the telescope principle.) □

8.4. Gcds and lcms

We have now seen that polynomials have quotients and remainders just like integers do, and can satisfy divisibilities and congruences just like integers can. Some of the basic arithmetical properties have analogues for polynomials. Here are three examples (we always assume $\mathbb{K}$ to be one of the sets $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$ and $\mathbb{C}$):

- If two polynomials P and Q in $\mathbb{K}[X]$ satisfy $P \mid Q$, then $\deg P \leq \deg Q$ unless $Q = 0$.
- If two polynomials P and Q in $\mathbb{K}[X]$ satisfy $P \mid Q$ and $Q \mid P$, then $Q = \lambda P$ for some $\lambda \in \mathbb{K}$. (This is the analogue of "if two integers a and b satisfy $a \mid b$ and $b \mid a$, then $|a| = |b|$".)
- If two polynomials P and Q in $\mathbb{K}[X]$ satisfy $P \mid Q$ and $\deg Q \leq \deg P$, then $Q = \lambda P$ for some $\lambda \in \mathbb{K}$.

This analogy can be taken further: We can define greatest common divisors and lowest common multiples of polynomials, at least when $\mathbb{K}$ is $\mathbb{Q}$ or $\mathbb{R}$ or $\mathbb{C}$ (the case $\mathbb{K} = \mathbb{Z}$ is more complicated). Here are the definitions:

Definition 8.4.1. Let $\mathbb{K}$ be one of the sets $\mathbb{Q}$, $\mathbb{R}$ and $\mathbb{C}$ (not $\mathbb{Z}$). Let P and Q be two polynomials in $\mathbb{K}[X]$ that are not both zero. Then:

(a) The *greatest common divisor* of P and Q is defined to be the monic polynomial of largest possible degree that divides each of P and Q. It is denoted by $\gcd(P, Q)$.

(b) The *lowest common multiple* of P and Q is defined to be the monic polynomial of smallest possible degree that is divisible by each of P and Q. It is denoted by $\operatorname{lcm}(P, Q)$.

We also define $\gcd(0, 0)$ and $\operatorname{lcm}(0, 0)$ to be 0.

Note that the word "monic" serves to make these polynomials unique, just like gcds and lcms of integers are required to be nonnegative. The uniqueness of $\operatorname{lcm}(P, Q)$ is easy to see[6]; the uniqueness of $\gcd(P, Q)$ is trickier[7]. Moreover, analogues of the classical properties of gcds and lcms for integers hold: If $\mathbb{K}$ is one of

[6] *Hint.* Argue that if U and V are two monic polynomials of the same degree that are divisible by each of P and Q, then their difference $U - V$ is a polynomial of smaller degree that is also divisible by each of P and Q. By properly scaling it, we can make it monic as well (unless it is 0).

[7] *Hint.* This can be proved by strong induction on $\deg P + \deg Q$, where the induction step proceeds by replacing P by $P\%Q$ (if $\deg P \geq \deg Q$) or replacing Q by $Q\%P$ (if $\deg Q > \deg P$). The reason why this works is that the common divisors of P and Q are exactly the common divisors of $P\%Q$ and Q and also are the common divisors of P and $Q\%P$. (Once again, this should all sound familiar from the number theory of integers, specifically from the Euclidean algorithm.)

the sets $\mathbb{Q}$, $\mathbb{R}$ and $\mathbb{C}$ (but not $\mathbb{Z}$), and if P and Q are two polynomials in $\mathbb{K}[X]$, then we have the following facts:

- *Universal property of the gcd:* Every common divisor of P and Q divides $\gcd(P, Q)$.
- *Universal property of the lcm:* Every common multiple of P and Q is a multiple of $\operatorname{lcm}(P, Q)$.
- *Bezout's identity:* There exist two polynomials A and B in $\mathbb{K}[X]$ that satisfy
$$AP + BQ = \gcd(P, Q). \tag{6}$$
- *Gcd-lcm relation:* We have
$$PQ = \lambda \cdot \gcd(P, Q) \cdot \operatorname{lcm}(P, Q), \tag{7}$$
where λ is the leading coefficient of PQ.
- *Invariance of the gcd under adding a multiple of one argument to the other:* If $N \in \mathbb{K}[X]$ is a polynomial such that $P \equiv Q \bmod N$, then
$$\gcd(N, P) = \gcd(N, Q). \tag{8}$$

Actually, a more civilized definition of $\gcd(P, Q)$ than the one we gave above would stipulate the universal property (as opposed to "smallest degree") as a defining property. (This would make the uniqueness of $\gcd(P, Q)$ easy, but the existence would become nontrivial.)

Just like for integers, $\gcd(P, Q)$ can be computed efficiently using a form of the Euclidean algorithm, in which we repeatedly replace the pair (P, Q) by $(Q, P\%Q)$ until the second entry of the pair becomes 0. (The "cave-man's Euclidean algorithm" from §6.1.3 on Worksheet 6 does not work here, since mere subtraction is not enough to decrease the degree; we actually need to take remainders.) Then, $\operatorname{lcm}(P, Q)$ can be computed using the gcd-lcm relation.

We reiterate that gcds and lcms break down for $\mathbb{K} = \mathbb{Z}$, at least according to our above definition. The definitions can be adapted, but Bezout's identity cannot be saved, as the following example shows.

Example 8.4.2. Let $\mathbb{K} = \mathbb{Q}$ and $P = X^3 - 2X^2$ and $Q = X^2 - 4$. Then, it is not hard to see that $\gcd(P, Q) = X - 2$. Bezout's identity thus says that there exist two polynomials A and B in $\mathbb{K}[X]$ that satisfy $AP + BQ = X - 2$. Such A and B can indeed be easily found: for instance, $A = \frac{1}{4}$ and $B = -\frac{1}{4}X + \frac{1}{2}$ will do.

Note that, even though P and Q are monic polynomials in $\mathbb{Z}[X]$, the A and the B cannot be found in $\mathbb{Z}[X]$. This illustrates why Bezout's identity was not stated for $\mathbb{K} = \mathbb{Z}$.

Exercise 8.4.1. Let $n, m \in \mathbb{N}$. Let $\mathbb{K} = \mathbb{Q}$. Prove that:

(a) If m is positive, then $(X^n - 1) \,\%\, (X^m - 1) = X^{n\%m} - 1$.

(b) We have $\gcd(X^n - 1,\ X^m - 1) = X^{\gcd(n,m)} - 1$.

Solution idea. **(a)** Assume that m is positive. Then, we can write n as $n = qm + r$ with $q = n//m \in \mathbb{Z}$ and $r = n\%m \in \{0, 1, \ldots, m-1\}$ (this is just classical division with remainder of integers). Consider these q and r. Note that $q \in \mathbb{N}$ (why?). Define the two polynomials

$$Q := X^r \left(X^{0m} + X^{1m} + X^{2m} + \cdots + X^{(q-1)m}\right) \qquad \text{and} \qquad R := X^r - 1$$

in $\mathbb{Q}[X]$. The definition of Q yields

$$\begin{aligned} Q \cdot (X^m - 1) &= X^r \underbrace{\left(X^{0m} + X^{1m} + X^{2m} + \cdots + X^{(q-1)m}\right) \cdot (X^m - 1)}_{\substack{= X^{qm} - 1 \\ \text{(by the geometric series formula – i.e., Theorem 4.1.2 on Worksheet 4)}}} \\ &= X^r \left(X^{qm} - 1\right) = X^{qm+r} - X^r = X^n - X^r \qquad (\text{since } qm + r = n) \\ &= (X^n - 1) - \underbrace{(X^r - 1)}_{=R} = (X^n - 1) - R, \end{aligned}$$

so that

$$X^n - 1 = Q \cdot (X^m - 1) + R.$$

Since $\deg R \leq r < m = \deg(X^m - 1)$, this equality entails that Q and R are the quotient and the remainder upon division of $X^n - 1$ by $X^m - 1$. In other words, $Q = (X^n - 1) // (X^m - 1)$ and $R = (X^n - 1) \,\%\, (X^m - 1)$. Hence,

$$(X^n - 1) \,\%\, (X^m - 1) = R = X^r - 1 = X^{n\%m} - 1 \qquad (\text{since } r = n\%m).$$

This solves part **(a)** of this exercise.

(b) Do not fix n and m. Instead, we proceed by strong induction on m:

Base case: The case of $m = 0$ is obvious, because in this case we have $X^m - 1 = X^0 - 1 = 0$ and thus $\gcd(X^n - 1,\ X^m - 1) = \gcd(X^n - 1,\ 0) = X^n - 1$.

Induction step: Let m be a positive integer. We assume (as the induction hypothesis) that part **(b)** of the exercise is already known to be true for all integers $m' < m$ in the place of m. Hence, in particular, part **(b)** of the exercise is proved for m and $n\%m$ instead of n and m (since $n\%m \in \{0, 1, \ldots, m-1\}$ and thus $n\%m < m$). In other words, we have

$$\gcd\left(X^m - 1,\ X^{n\%m} - 1\right) = X^{\gcd(m,\ n\%m)} - 1. \qquad (9)$$

However, recalling the proof of correctness of the usual Euclidean algorithm for integers, we know that $\gcd(m,\ n\%m) = \gcd(n, m)$. (Indeed, this is easy to check: We have $n\%m \equiv n \operatorname{mod} m$. Thus, the analogue of (8) for integers yields $\gcd\left(m,\ n\%m\right) = \gcd\left(m,n\right) = \gcd\left(n,m\right)$.)

Part **(a)** of this exercise shows that $\left(X^n - 1\right) \% \left(X^m - 1\right) = X^{n\%m} - 1$. Hence,

$$X^{n\%m} - 1 = \left(X^n - 1\right) \% \left(X^m - 1\right) \equiv X^n - 1 \operatorname{mod} X^m - 1$$

(because if P and N are two polynomials with $N \neq 0$, then $P\%N \equiv P \operatorname{mod} N$). Therefore, (8) (applied to $P = X^{n\%m} - 1$ and $Q = X^n - 1$ and $N = X^m - 1$) yields

$$\gcd\left(X^m - 1,\ X^{n\%m} - 1\right) = \gcd\left(X^m - 1,\ X^n - 1\right) = \gcd\left(X^n - 1,\ X^m - 1\right),$$

so that

$$\begin{aligned} \gcd\left(X^n - 1,\ X^m - 1\right) &= \gcd\left(X^m - 1,\ X^{n\%m} - 1\right) \\ &= X^{\gcd(m,\ n\%m)} - 1 \qquad \text{(by (9))} \\ &= X^{\gcd(n,m)} - 1 \qquad \left(\text{since } \gcd\left(m,\ n\%m\right) = \gcd\left(n,m\right)\right). \end{aligned}$$

In other words, part **(b)** of the exercise holds for our two numbers n and m. This completes the induction step, and thus part **(b)** of the exercise is solved by induction. □

The "number theory" of polynomials (i.e., their properties regarding divisibility) is a fruitful field of research, but we will leave it now. Let us only briefly mention that there is a polynomial analogue of coprime integers (unsurprisingly called *coprime polynomials*) and a polynomial analogue of prime numbers (known as *irreducible polynomials*[8]). Much has been written about irreducible polynomials (see, e.g., [Prasol04, Chapter 2]), but most of it would go beyond the reach of this worksheet.

8.5. Roots and the polynomial identity trick

8.5.1. Roots of polynomials

Polynomials first appeared in mathematics as an abstraction for algebraic equations. For instance, solving the equation $x^3 - 2x^2 + 7 = 0$ means finding a number x such that substituting x into the polynomial $X^3 - 2X^2 + 7$ gives 0. In general, a number x that becomes 0 when substituted into a polynomial P is called a *root* of P. In other words:

[8] Specifically, a polynomial $P \in \mathbb{K}\left[X\right]$ (where $\mathbb{K}$ is $\mathbb{Q}$ or $\mathbb{R}$ or $\mathbb{C}$) is said to be *irreducible* if it is non-constant and is not divisible by any non-constant polynomials other than those of the form λP with $\lambda \in \mathbb{K}$. Note that this depends very strongly on $\mathbb{K}$: For example, the polynomial $X^2 - 2$ is irreducible when regarded as a polynomial in $\mathbb{Q}\left[X\right]$, but not when regarded as a polynomial in $\mathbb{R}\left[X\right]$ (since it is divisible by $X - \sqrt{2}$).

The analogy between prime numbers and irreducible polynomials is almost perfect; one minor difference is that irreducible polynomials are not commonly required to be monic, so that an irreducible polynomial P automatically entails irreducible polynomials λP for all nonzero $\lambda \in \mathbb{K}$.

Definition 8.5.1. Let $\mathbb{K}$ be one of the sets $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$ and $\mathbb{C}$. Let $P \in \mathbb{K}[X]$ be a polynomial. A *root* of P means a complex number $r \in \mathbb{C}$ such that $P(r) = 0$.

Note that even if a polynomial P belongs to $\mathbb{Z}[X]$ (that is, all its coefficients are integers), one is commonly interested in complex roots of P as well, since they often shed light on P (and they include all the integer roots as a subset, so nothing is lost by extending the scope). Here are some examples:

- The polynomial $X^3 - 4X^2 + 3X$ has three roots: 0, 1 and 3. All of these roots are integers.
- The polynomial $X^3 - 2X^2 + 1$ has three roots: 1, $\dfrac{1+\sqrt{5}}{2}$ and $\dfrac{1-\sqrt{5}}{2}$. Only the first of these roots is an integer.
- The polynomial $X^3 - 2X^2 + X$ has two roots: 0 and 1 (both integers).
- The polynomial $X^3 + X^2 + X$ has three roots: 0, $\dfrac{-1+\sqrt{3}i}{2}$ and $\dfrac{-1-\sqrt{3}i}{2}$, where $i = \sqrt{-1}$. The first of them is a real number, while the other two are not.
- The polynomial $X^3 - 3X + 1$ has three real roots, which are approximately equal to -1.879, 0.347 and 1.532. The fact that these roots exist can be easily derived from the analytic properties of the corresponding polynomial **function**[9], but the only exact expressions for them are intricate and inconvenient (see Cardano's formula). The question of finding a formula for the roots of an arbitrary cubic polynomial (i.e., solving cubic equations) has been one of the guiding problems in medieval mathematics, and its solution by Tartaglia and Cardano in the 1530s has reestablished Europe as the center of mathematical research for the first time after the Dark Ages.
- The polynomial $2X + 1$ has one root: $-1/2$.
- The polynomial 3 has no roots, since substituting anything into 3 gives 3.
- The polynomial 0 has infinitely many roots: Any complex number is a root of 0.

[9] To wit: The function $f : \mathbb{R} \to \mathbb{R}$ given by $f(x) = x^3 - 3x + 1$ is continuous (just like any polynomial function), and it takes a negative value at -2 (since $f(-2) = -1$), a positive value at -1 (since $f(-1) = 3$), a negative value at 1 (since $f(1) = -1$) and a positive value at 2 (since $f(2) = 3$). Hence, by the intermediate value theorem, it must take the value 0 in each of the open intervals $(-2, -1)$, $(-1, 1)$ and $(1, 2)$. Thus, it takes the value 0 at least three times, i.e., the polynomial $X^3 - 3X + 1$ has at least three real roots. Approximate values for these roots can be found by the bisection method.

8.5.2. Roots vs. divisors

As these examples show, roots can behave rather unpredictably. Nevertheless, a few simple but important things can be said. For one, the roots of a polynomial are directly related to its linear divisors:

> **Theorem 8.5.2** ("root = factor" theorem)**.** Let $\mathbb{K}$ be one of the sets $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$ and $\mathbb{C}$. Let $P \in \mathbb{K}[X]$ be a polynomial. Let $r \in \mathbb{K}$ be a number. Then, r is a root of P if and only if P is divisible by $X - r$ (in $\mathbb{K}[X]$).

Proof. We must prove the logical equivalence

$$(r \text{ is a root of } P) \iff (P \text{ is divisible by } X - r).$$

We shall prove the "$\Longleftarrow$" and "$\Longrightarrow$" directions of this equivalence separately:

$\Longleftarrow$: Assume that P is divisible by $X - r$. We must prove that r is a root of P.

Since P is divisible by $X - r$, we can write P as $P = (X - r) \cdot Q$ for some polynomial Q. Consider this Q. Substituting r for X on both sides of the equality $P = (X - r) \cdot Q$, we obtain

$$\begin{aligned} P(r) &= ((X - r) \cdot Q)(r) && \text{(this is to be read as "the value of } (X - r) \cdot Q \text{ at } r\text{")} \\ &= \underbrace{(r - r)}_{=0} \cdot Q(r) && \text{(here, we have used (1) and (2))} \\ &= 0, \end{aligned}$$

which shows that r is a root of P. Thus, the "$\Longleftarrow$" direction of our equivalence is proved.

$\Longrightarrow$: Assume that r is a root of P. We must show that P is divisible by $X - r$.

The polynomial $X - r \in \mathbb{K}[X]$ is monic. Hence, using Theorem 8.3.3 or Theorem 8.3.1 (depending on whether $\mathbb{K}$ is $\mathbb{Z}$ or not), we see that the quotient $P // (X - r)$ and the remainder $P \% (X - r)$ are well-defined. Let us set $Q := P // (X - r)$ and $R := P \% (X - r)$. By the definition of quotient and remainder, we thus have

$$P = Q \cdot (X - r) + R \qquad \text{and} \qquad \deg R < \deg(X - r).$$

From $\deg R < \deg(X - r) = 1$, we see that the polynomial R is constant, i.e., satisfies $R = b$ for some number $b \in \mathbb{K}$. Consider this b. Thus, $P = Q \cdot (X - r) + R$ can be rewritten as $P = Q \cdot (X - r) + b$.

Since r is a root of P, we have $P(r) = 0$. But substituting r for X on both sides of the equality $P = Q \cdot (X - r) + b$, we obtain

$$\begin{aligned} P(r) &= (Q \cdot (X - r) + b)(r) && \text{(this is to be read as "the value of } Q \cdot (X - r) + b \text{ at } r\text{")} \\ &= Q(r) \cdot \underbrace{(r - r)}_{=0} + b && \text{(here, we have used (1) and (2))} \\ &= b, \end{aligned}$$

so that $b = P(r) = 0$. Therefore, $P = Q \cdot (X - r) + \underbrace{b}_{=0} = Q \cdot (X - r) = (X - r) \cdot Q$. This shows that P is divisible by $X - r$. Thus, the "$\Longrightarrow$" direction of our equivalence is proved.

We have now proved both directions of our equivalence. Thus, Theorem 8.5.2 is proved. □

Theorem 8.5.3 ("roots = factors" theorem)**.** Let $\mathbb{K}$ be one of the sets $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$ and $\mathbb{C}$. Let $P \in \mathbb{K}[X]$ be a polynomial. Let $r_1, r_2, \ldots, r_m \in \mathbb{K}$ be distinct numbers. Then, $r_1, r_2, \ldots, r_m$ are roots of P if and only if P is divisible by $(X - r_1)(X - r_2)\cdots(X - r_m)$ (in $\mathbb{K}[X]$).

Proof idea. $\Longleftarrow$: Assume that P is divisible by $(X - r_1)(X - r_2)\cdots(X - r_m)$. We must show that $r_1, r_2, \ldots, r_m$ are roots of P.

Our assumption yields that

$$P = (X - r_1)(X - r_2)\cdots(X - r_m) \cdot Q \qquad (10)$$

for some polynomial Q. Consider this Q.

For each $i \in \{1, 2, \ldots, m\}$, we have

$$\begin{aligned} P(r_i) &= ((X - r_1)(X - r_2)\cdots(X - r_m) \cdot Q)(r_i) \qquad \text{(by (10))} \\ &= \underbrace{(r_i - r_1)(r_i - r_2)\cdots(r_i - r_m)}_{\substack{=0 \\ \text{(since one factor of this product is } r_i - r_i = 0)}} \cdot Q(r_i) = 0, \end{aligned}$$

which shows that r_i is a root of P. In other words, $r_1, r_2, \ldots, r_m$ are roots of P. This proves the "$\Longleftarrow$" direction of Theorem 8.5.3.

$\Longrightarrow$: We must show that if $r_1, r_2, \ldots, r_m$ are roots of P, then P is divisible by $(X - r_1)(X - r_2)\cdots(X - r_m)$.

We will prove this claim by induction on m:

Base case: For $m = 0$, the claim is obvious (since the product $(X - r_1)(X - r_2)\cdots(X - r_m)$ is an empty product in this case, hence equal to 1, and of course P is divisible by 1).

Induction step: Let m be a positive integer. Assume (as the induction hypothesis) that our claim is proved for $m - 1$ instead of m. We must now prove it for m.

So let us assume that $r_1, r_2, \ldots, r_m$ are roots of P. Thus, in particular, r_m is a root of P. Hence, Theorem 8.5.2 (applied to $r = r_m$) shows that P is divisible by $X - r_m$ (in $\mathbb{K}[X]$). In other words, there exists a polynomial $Q \in \mathbb{K}[X]$ such that $P = (X - r_m) \cdot Q$. Consider this Q.

Now, for each $i \in \{1, 2, \ldots, m - 1\}$, we have $r_i \neq r_m$ (since the numbers $r_1, r_2, \ldots, r_m$ are distinct), so that $r_i - r_m \neq 0$. Moreover, for each $i \in \{1, 2, \ldots, m - 1\}$, we have

$$P(r_i) = (r_i - r_m) \cdot Q(r_i)$$

(this follows by substituting r_i for X into the equality $P = (X - r_m) \cdot Q$) and thus

$$(r_i - r_m) \cdot Q(r_i) = P(r_i) = 0 \qquad \text{(since } r_i \text{ is a root of } P\text{)},$$

which entails that $Q(r_i) = 0$ (because $r_i - r_m \neq 0$); in other words, r_i is a root of Q. Thus we have shown that the $m-1$ distinct numbers $r_1, r_2, \ldots, r_{m-1}$ are roots of Q. Hence, by our induction hypothesis, we can apply the "$\Longrightarrow$" direction of Theorem 8.5.3 to Q and $m-1$ instead of P and m. We thus conclude that Q is divisible by $(X - r_1)(X - r_2) \cdots (X - r_{m-1})$. In other words, $Q = (X - r_1)(X - r_2) \cdots (X - r_{m-1}) \cdot R$ for some polynomial $R \in \mathbb{K}[X]$. Consider this R. Now,

$$\begin{aligned} P &= (X - r_m) \cdot \underbrace{Q}_{=(X-r_1)(X-r_2)\cdots(X-r_{m-1})\cdot R} \\ &= \underbrace{(X - r_m) \cdot (X - r_1)(X - r_2) \cdots (X - r_{m-1})}_{=(X-r_1)(X-r_2)\cdots(X-r_m)} \cdot R \\ &= (X - r_1)(X - r_2) \cdots (X - r_m) \cdot R, \end{aligned}$$

so that P is divisible by $(X - r_1)(X - r_2) \cdots (X - r_m)$. This is precisely what we wanted to show. Thus, our claim holds for m, and the induction step is complete. □

Corollary 8.5.4. Let $\mathbb{K}$ be one of the sets $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$ and $\mathbb{C}$. Let $n \in \mathbb{N}$. Let $P \in \mathbb{K}[X]$ be a nonzero polynomial of degree $\leq n$. Then, P has at most n roots.

Proof idea. Assume the contrary. Thus, P has at least $n+1$ roots. In other words, there exist $n+1$ distinct roots $r_1, r_2, \ldots, r_{n+1} \in \mathbb{C}$ of P. Consider these $n+1$ roots.

The set $\mathbb{K}$ is a subset of $\mathbb{C}$. Thus, $\mathbb{K}[X]$ is a subset of $\mathbb{C}[X]$. Hence, $P \in \mathbb{K}[X] \subseteq \mathbb{C}[X]$. Therefore, applying Theorem 8.5.3 to $n+1$ and $\mathbb{C}$ instead of n and $\mathbb{K}$, we see that $r_1, r_2, \ldots, r_{n+1}$ are roots of P if and only if P is divisible by $(X - r_1)(X - r_2) \cdots (X - r_{n+1})$ (in $\mathbb{C}[X]$). We thus conclude that P is divisible by $(X - r_1)(X - r_2) \cdots (X - r_{n+1})$ (since $r_1, r_2, \ldots, r_{n+1}$ are roots of P). In other words, P can be written as

$$P = (X - r_1)(X - r_2) \cdots (X - r_{n+1}) \cdot Q \qquad (11)$$

for some polynomial $Q \in \mathbb{C}[X]$. Consider this Q. The equality (11) shows that Q is nonzero (since P is nonzero). Hence, $\deg Q \geq 0$.

Now, from (11), we obtain

$$\begin{aligned} \deg P &= \deg((X - r_1)(X - r_2) \cdots (X - r_{n+1}) \cdot Q) \\ &= \underbrace{\deg((X - r_1)(X - r_2) \cdots (X - r_{n+1}))}_{=n+1} + \underbrace{\deg Q}_{\geq 0} \\ &\qquad \text{(by Proposition 8.2.1 \textbf{(b)})} \\ &\geq n + 1 > n, \end{aligned}$$

which contradicts the fact that P has degree $\leq n$. This contradiction shows that our assumption was false. Hence, Corollary 8.5.4 is proved. □

Theorem 8.5.2, Theorem 8.5.3 and Corollary 8.5.4 are surprisingly useful for their simplicity. Here is a classical competition problem that has spawned many variants (in particular, some on the Putnam contest):

Exercise 8.5.1. Let $n \in \mathbb{N}$. Let $P \in \mathbb{Q}[X]$ be a polynomial of degree $\leq n$ that satisfies $P(k) = \dfrac{1}{k}$ for each $k \in \{1, 2, \ldots, n+1\}$. Find $P(n+2)$.

Solution idea. How can we exploit the "$P(k) = \dfrac{1}{k}$ for each $k \in \{1, 2, \ldots, n+1\}$" condition? One way to do so is by multiplying it by k, obtaining $k \cdot P(k) = 1$, or, equivalently, $k \cdot P(k) - 1 = 0$. This shows that each $k \in \{1, 2, \ldots, n+1\}$ is a root of the polynomial $XP - 1$ (because $(XP-1)(k) = k \cdot P(k) - 1 = 0$). In other words, the $n+1$ distinct numbers $1, 2, \ldots, n+1$ are roots of the polynomial $XP - 1$.

The polynomial P has degree $\leq n$. Thus, the polynomial XP has degree $\leq n+1$ (by Proposition 8.2.1 **(b)**, or just by observing that multiplication by X turns every monomial X^i into X^{i+1} and thus raises the degree by 1). Therefore, the polynomial $XP - 1$ has degree $\leq n+1$ as well. In other words, $\deg(XP-1) \leq n+1$.

However, Theorem 8.5.3 (applied to $\mathbb{Q}$, $n+1$, $XP-1$ and $\dfrac{1}{k}$ instead of $\mathbb{K}$, m, P and r_k) yields that $1, 2, \ldots, n+1$ are roots of the polynomial $XP-1$ if and only if $XP-1$ is divisible by $(X-1)(X-2)\cdots(X-(n+1))$. Thus, $XP-1$ is divisible by $(X-1)(X-2)\cdots(X-(n+1))$ (since $1, 2, \ldots, n+1$ are roots of the polynomial $XP-1$). In other words, we can write $XP-1$ as

$$XP - 1 = (X-1)(X-2)\cdots(X-(n+1)) \cdot Q \tag{12}$$

for some $Q \in \mathbb{Q}[X]$. Consider this Q.

From (12), we obtain

$$\begin{aligned}\deg(XP-1) &= \deg((X-1)(X-2)\cdots(X-(n+1)) \cdot Q) \\ &= \underbrace{\deg((X-1)(X-2)\cdots(X-(n+1)))}_{\substack{=n+1 \\ \text{(by Proposition 8.2.1 \textbf{(b)})}}} + \deg Q \\ &= (n+1) + \deg Q,\end{aligned}$$

so that

$$\deg Q = \underbrace{\deg(XP-1)}_{\leq n+1} - (n+1) \leq (n+1) - (n+1) = 0.$$

This shows that the polynomial Q is constant. In other words, $Q = q$ for some $q \in \mathbb{Q}$. Consider this q. Since $Q = q$, we can rewrite (12) as

$$XP - 1 = (X-1)(X-2)\cdots(X-(n+1)) \cdot q. \tag{13}$$

How can we compute q ? We can compare constant coefficients on both sides of (13). In fact:

- The constant coefficient of $XP-1$ is -1 (since the product XP contains only positive powers of X and thus contributes nothing to the constant coefficient). In other words, the constant coefficient on the left hand side of (13) is -1.
- The constant coefficient of $(X-1)(X-2)\cdots(X-(n+1))$ is $(-1)(-2)\cdots(-(n+1))$ (since we can expand the product $(X-1)(X-2)\cdots(X-(n+1))$ as a sum, and the only addend without an X factor will be $(-1)(-2)\cdots(-(n+1))$). Thus, the constant coefficient of $(X-1)(X-2)\cdots(X-(n+1))\cdot q$ is
$$\begin{aligned}(-1)(-2)\cdots(-(n+1))\cdot q &= (-1)^{n+1}\underbrace{(1\cdot 2\cdot\cdots\cdot(n+1))}_{=(n+1)!}\cdot q\\ &= (-1)^{n+1}(n+1)!\cdot q.\end{aligned}$$
  In other words, the constant coefficient on the right hand side of (13) is $(-1)^{n+1}(n+1)!\cdot q$.

Since the two sides of (13) are equal, their constant coefficients must be equal as well. In other words,
$$-1=(-1)^{n+1}(n+1)!\cdot q$$
(since we just found that these constant coefficients are -1 and $(-1)^{n+1}(n+1)!\cdot q$, respectively). Solving this for q, we obtain
$$q=\frac{-1}{(-1)^{n+1}(n+1)!}=\frac{(-1)^n}{(n+1)!}. \tag{14}$$

Now, recall that we are looking for $P(n+2)$. We can obtain this by substituting $n+2$ for X in (13). This results in
$$\begin{aligned}(n+2)\cdot P(n+2)-1 &= \underbrace{((n+2)-1)((n+2)-2)\cdots((n+2)-(n+1))}_{\substack{=(n+1)n\cdots 1\\ =1\cdot 2\cdot\cdots\cdot(n+1)\\ =(n+1)!}}\cdot q\\ &= (n+1)!\cdot q=(n+1)!\cdot\frac{(-1)^n}{(n+1)!} \qquad \text{(by (14))}\\ &= (-1)^n.\end{aligned}$$

Solving this for $P(n+2)$, we find
$$P(n+2)=\frac{1+(-1)^n}{n+2}.$$

□

Remark 8.5.5. In Exercise 8.5.1, the existence of a polynomial P satisfying the imposed conditions is assumed. Does such a P actually exist? Yes. Indeed, we can solve the equality (13) for P, obtaining

$$P = \frac{1 + (X-1)(X-2)\cdots(X-(n+1)) \cdot q}{X},$$

where q is given by (14). The division by X is allowed here, since the numerator $1 + (X-1)(X-2)\cdots(X-(n+1)) \cdot q$ is a polynomial with constant term 0 (why?) and thus can be divided by X without remainder (i.e., the remainder will be 0). Having defined P in this way, we can easily check that P has degree $\leq n$ (actually exactly n) and satisfies $P(k) = \dfrac{1}{k}$ for each $k \in \{1, 2, \ldots, n+1\}$.

Alternatively, the existence of P can be easily obtained from Lagrange interpolation (see Theorem 8.5.10 below).

Remark 8.5.6. Exercise 8.5.1 had us compute $P(n+2)$. We could have similarly computed $P(a)$ for any $a \neq 0$ (although the answers would be less elegant). Computing $P(0)$ is trickier, however. In fact, substituting 0 for X in (13) yields $0 \cdot P(0) - 1$ on the left hand side, which reveals nothing about $P(0)$. There is a trick that helps, and the answer is surprising: see Exercise 8.8.5 below.

8.5.3. The polynomial identity trick

Corollary 8.5.4 has the following trivial but rather useful consequence:

Corollary 8.5.7 (polynomial identity trick)**.** Let $\mathbb{K}$ be one of the sets $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$ and $\mathbb{C}$. Let $n \in \mathbb{N}$. Let $P, Q \in \mathbb{K}[X]$ be two polynomials of degree $\leq n$. Assume that there are more than n numbers $x \in \mathbb{K}$ that satisfy $P(x) = Q(x)$. Then, $P = Q$.

Proof idea. Assume the contrary. Thus, the difference $P - Q$ is nonzero.

The polynomials P and Q have degree $\leq n$. Hence, their difference $P - Q$ has degree $\leq n$ as well. Since it is furthermore nonzero, we can thus conclude that $P - Q$ has at most n roots (by Corollary 8.5.4, applied to $P - Q$ instead of P).

However, we assumed that there are more than n numbers $x \in \mathbb{K}$ that satisfy $P(x) = Q(x)$. All such numbers $x \in \mathbb{K}$ are roots of $P - Q$ (since they satisfy $(P-Q)(x) = \underbrace{P(x)}_{=Q(x)} - Q(x) = Q(x) - Q(x) = 0$). Hence, $P - Q$ has more than n roots. This contradicts the fact that $P - Q$ has at most n roots. This contradiction shows that our assumption was wrong. Thus, Corollary 8.5.7 is proved. □

Corollary 8.5.7 entails that polynomials have rather little "freedom" as far as their values are concerned: A polynomial of degree $\leq n$ is uniquely determined by any $n+1$ of its values (on $n+1$ distinct numbers x). For example, a linear polynomial is uniquely determined by two values (no surprise here: a line is uniquely determined by two distinct points on it), whereas a quadratic polynomial is uniquely determined by three values.

This yields, in particular, the following:

Corollary 8.5.8. Let $\mathbb{K}$ be one of the sets $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$ and $\mathbb{C}$. Any polynomial $P \in \mathbb{K}[X]$ is uniquely determined by the corresponding polynomial function

$$f_P : \mathbb{K} \to \mathbb{K},$$
$$a \mapsto P(a).$$

In other words: If P and Q are two polynomials in $\mathbb{K}[X]$ such that $f_P = f_Q$, then $P = Q$.

Proof idea. Let P and Q be two polynomials in $\mathbb{K}[X]$ such that $f_P = f_Q$. We must prove that $P = Q$.

Let $n = \max\{\deg P, \deg Q\}$. Thus, P and Q are polynomials of degree $\leq n$. But the set $\mathbb{K}$ is infinite, and thus contains more than n numbers. Each number $x \in \mathbb{K}$ satisfies $P(x) = \underbrace{f_P}_{=f_Q}(x) = f_Q(x) = Q(x)$. Hence, there are more than n numbers $x \in \mathbb{K}$ that satisfy $P(x) = Q(x)$ (since there are more than n numbers $x \in \mathbb{K}$). Therefore, Corollary 8.5.7 yields $P = Q$. □

These facts can be made more concrete: We can give an explicit formula for reconstructing a polynomial P of degree $\leq n$ from $n+1$ values $P(a_0), P(a_1), \ldots, P(a_n)$. This formula is Theorem 8.5.9 below. First, however, we shall show how Corollary 8.5.7 can be used already in its present (non-concrete) form.

Recall that the *binomial coefficient* $\dbinom{n}{k}$ is defined for any number $n \in \mathbb{C}$ and any nonnegative integer $k \in \mathbb{N}$ by the formula

$$\binom{n}{k} := \frac{n(n-1)(n-2)\cdots(n-k+1)}{k!}.$$

We could have just as well defined the polynomial $\dbinom{X}{k} \in \mathbb{Q}[X]$ by

$$\binom{X}{k} := \frac{X(X-1)(X-2)\cdots(X-k+1)}{k!},$$

and then defined $\dbinom{n}{k}$ to be the value of this polynomial $\dbinom{X}{k}$ at n. Much can be said about binomial coefficients (see, e.g., [Grinbe19b, Chapter 2], [BenQui03], [GrKnPa94, Chapter 5], [Granvi05]), but here is not the time and space. We will just solve a toy problem:

Exercise 8.5.2. Prove that $\dbinom{\dbinom{n}{2}}{2} = 3\dbinom{n}{4} + 3\dbinom{n}{3}$ for each $n \in \mathbb{C}$.

Solution idea. Obviously, this is an identity that can be proved mechanically by expanding both sides and comparing. But here is a slicker proof using Corollary 8.5.7:

Consider the two polynomials

$$P := \dbinom{\dbinom{X}{2}}{2} \qquad \text{and} \qquad Q := 3\dbinom{X}{4} + 3\dbinom{X}{3}$$

in $\mathbb{Q}[X]$. Thus, the exercise wants us to prove that $P(n) = Q(n)$ for each $n \in \mathbb{C}$. Obviously, it suffices to show that $P = Q$ (and because of Corollary 8.5.8, this is equivalent to the claim that $P(n) = Q(n)$ for each $n \in \mathbb{C}$, so that we don't need to worry about trying to prove something false).

It is easy to see that both polynomials P and Q have degree ≤ 4. (Indeed, Q has degree ≤ 4 because the polynomial $\dbinom{X}{k}$ has degree k for each $k \in \mathbb{N}$. As far as P is concerned, it is not hard to see that substituting a degree-n polynomial into a degree-m polynomial yields a degree-nm polynomial (as long as both polynomials are nonzero). Hence, the polynomial P (which is obtained by substituting the degree-2 polynomial $\dbinom{X}{2}$ into itself) is a degree-4 polynomial.)

Corollary 8.5.7 (applied to $n = 4$ and $\mathbb{K} = \mathbb{Q}$) thus says that if there are more than 4 numbers $x \in \mathbb{Q}$ that satisfy $P(x) = Q(x)$, then $P = Q$. Therefore, it suffices to find more than 4 such numbers $x \in \mathbb{Q}$.

The easiest way to find them is to take the five numbers $0, 1, 2, 3, 4$. Thus, we need to check that $P(x) = Q(x)$ for each $x \in \{0, 1, 2, 3, 4\}$. This is not only straightforward but actually goes very quick if you remember Pascal's triangle (which you should). For example, for $x = 4$, we have

$$P(x) = P(4) = \dbinom{\dbinom{4}{2}}{2} = \dbinom{6}{2} = 15 \qquad \text{and}$$
$$Q(x) = Q(4) = 3\underbrace{\dbinom{4}{4}}_{=1} + 3\underbrace{\dbinom{4}{3}}_{=4} = 3 + 3 \cdot 4 = 15,$$

so that $P(x) = Q(x)$. The other four options for x are even easier (e.g., the cases $x \in \{0, 1, 2\}$ boil down to $0 = 0$). Thus, by "probing" the equality $P = Q$ at our five numbers x, we have ensured that it holds. This solves the exercise.

Remark: Our choice of $0, 1, 2, 3, 4$ is just one possible option; any five distinct numbers would have worked. For instance, you can make your job a bit easier by using $x = -1$ instead of $x = 4$, as long as you remember (which again you should) that $\dbinom{-1}{k} = (-1)^k$ for each $k \in \mathbb{N}$. Furthermore, you can save yourself some work if you check that the polynomials P and Q have the same coefficient before X^4 (namely, $\dfrac{1}{8}$), so that $P - Q$ has degree ≤ 3 (which means that we need to "probe" the equality $P = Q$ at four points only). □

This was arguably a toy example (we could have solved Exercise 8.5.2 by dully expanding both sides), but the "probing" method we have just used is worth remembering, as it has various other uses, many of them much less gratuitous. Examples of such uses can be found in [Grinbe20, §7.5.3], [Grinbe19b, §2.6] and [GrKnPa94, Chapter 5].

8.5.4. Lagrange interpolation

Corollaries 8.5.7 and 8.5.8 show that a polynomial $P \in \mathbb{C}[X]$ of degree $\leq n$ is uniquely determined by its values on $n+1$ given distinct numbers $a_0, a_1, \ldots, a_n$. But is this just an abstract uniqueness statement, or can we actually reconstruct P given these $n+1$ values? It turns out that we can, and in fact there is a rather explicit formula:

> **Theorem 8.5.9** (Lagrange interpolation formula, I)**.** Let $\mathbb{K}$ be one of the sets $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$ and $\mathbb{C}$. Let $n \in \mathbb{N}$. Let $P \in \mathbb{K}[X]$ be a polynomial of degree $\leq n$. Let $a_0, a_1, \ldots, a_n$ be $n+1$ distinct numbers in $\mathbb{K}$. Let $b_i := P(a_i)$ for each $i \in \{0, 1, \ldots, n\}$. Then,
>
> $$P(X) = \sum_{j=0}^{n} b_j \frac{\prod\limits_{k \neq j} (X - a_k)}{\prod\limits_{k \neq j} (a_j - a_k)}$$
>
> (where the "$\prod\limits_{k \neq j}$" sign means a product over all $k \in \{0, 1, \ldots, n\}$ satisfying $k \neq j$).

Proof idea. We define a polynomial $Q \in \mathbb{K}[X]$ by

$$Q = \sum_{j=0}^{n} b_j \frac{\prod\limits_{k \neq j} (X - a_k)}{\prod\limits_{k \neq j} (a_j - a_k)}. \qquad (15)$$

Our goal is thus to prove that $P(X) = Q$. In other words, we must prove that $P = Q$ (since $P(X)$ is just another way to say P).

The polynomial Q has degree $\leq n$ (since each product $\prod\limits_{k \neq j} (X - a_k)$ has n factors, thus is a polynomial of degree n, and our polynomial Q is a sum of all these products multiplied with certain numbers). We shall now show that it satisfies

$$Q(a_i) = b_i \qquad \text{for each } i \in \{0, 1, \ldots, n\}. \qquad (16)$$

Proof of (16). Let $i \in \{0, 1, \ldots, n\}$. Substituting a_i for X on both sides of (15), we find

$$\begin{aligned}
Q(a_i) &= \left(\sum_{j=0}^{n} b_j \frac{\prod\limits_{k \neq j} (X - a_k)}{\prod\limits_{k \neq j} (a_j - a_k)} \right)(a_i) \\
&= \sum_{j=0}^{n} b_j \frac{\prod\limits_{k \neq j} (a_i - a_k)}{\prod\limits_{k \neq j} (a_j - a_k)} \qquad \text{(by (1) and (2), used many times)} \\
&= b_i \underbrace{\frac{\prod\limits_{k \neq i} (a_i - a_k)}{\prod\limits_{k \neq i} (a_i - a_k)}}_{=1} + \sum_{\substack{j \in \{0,1,\ldots,n\}; \\ j \neq i}} b_j \underbrace{\frac{\prod\limits_{k \neq j} (a_i - a_k)}{\prod\limits_{k \neq j} (a_j - a_k)}}_{\substack{=0 \\ \text{(because the product in the numerator} \\ \text{has a factor } a_i - a_i \text{ (since } i \neq j\text{),} \\ \text{thus a factor equal to 0)}}} \\
&\qquad \left(\begin{array}{c} \text{here, we have split off the addend for } j = i \\ \text{from the sum} \end{array} \right) \\
&= b_i + \underbrace{\sum_{\substack{j \in \{0,1,\ldots,n\}; \\ j \neq i}} b_j 0}_{=0} = b_i.
\end{aligned}$$

This proves (16).] □

Now we are almost done: For each $i \in \{0, 1, \ldots, n\}$, we have $P(a_i) = Q(a_i)$ (since (16) yields $Q(a_i) = b_i = P(a_i)$ by the definition of b_i). In other words, the equality $P(x) = Q(x)$ holds for each $x \in \{a_0, a_1, \ldots, a_n\}$. Since the $n+1$ numbers $a_0, a_1, \ldots, a_n$ are all distinct, this shows that $P(x) = Q(x)$ holds for at least $n+1$ many numbers $x \in \mathbb{K}$. In other words, there are more than n numbers $x \in \mathbb{K}$ that satisfy $P(x) = Q(x)$. Corollary 8.5.7 thus yields $P = Q$, and this completes our proof. □

Theorem 8.5.9 tells us how to **reconstruct** a polynomial P from any $n+1$ of its values; but we can use the same method to **construct** a polynomial P with given values at $n+1$ given points:

> **Theorem 8.5.10** (Lagrange interpolation formula, II)**.** Let $\mathbb{K}$ be one of the sets $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$ and $\mathbb{C}$. Let $n \in \mathbb{N}$. Let $a_0, a_1, \ldots, a_n$ be $n+1$ distinct numbers in $\mathbb{K}$. Let $b_0, b_1, \ldots, b_n$ be $n+1$ numbers in $\mathbb{K}$ (not necessarily distinct). Define a polynomial $Q \in \mathbb{K}[X]$ by
>
> $$Q = \sum_{j=0}^{n} b_j \frac{\prod\limits_{k \neq j} (X - a_k)}{\prod\limits_{k \neq j} (a_j - a_k)}$$

(where the "$\prod\limits_{k\neq j}$" sign means a product over all $k \in \{0, 1, \ldots, n\}$ satisfying $k \neq j$).

Then,

$$Q(a_i) = b_i \qquad \text{for each } i \in \{0, 1, \ldots, n\}.$$

Proof. This is precisely the equality (16), which we showed during our proof of Theorem 8.5.9. Its proof still applies here, since it used nothing about P. □

Theorem 8.5.9 is known as the *Lagrange interpolation formula*, and has many applications. The following two exercises are just a little selection:

Exercise 8.5.3. Let $\mathbb{K}$ be one of the sets $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$ and $\mathbb{C}$. Let $n \in \mathbb{N}$. Let $P \in \mathbb{K}[X]$ be a polynomial of degree $\leq n$. Let $c_n = [X^n] P$ be the coefficient of X^n in P. (Note that $c_n = 0$ if P has degree $< n$.) Let $a_0, a_1, \ldots, a_n$ be $n+1$ distinct numbers in $\mathbb{K}$. Prove that

$$\sum_{j=0}^{n} \frac{P(a_j)}{\prod\limits_{k\neq j} (a_j - a_k)} = c_n$$

(where the "$\prod\limits_{k\neq j}$" sign means a product over all $k \in \{0, 1, \ldots, n\}$ satisfying $k \neq j$).

Solution idea. Let $b_i := P(a_i)$ for each $i \in \{0, 1, \ldots, n\}$. Then, Theorem 8.5.9 yields

$$P(X) = \sum_{j=0}^{n} b_j \frac{\prod\limits_{k\neq j} (X - a_k)}{\prod\limits_{k\neq j} (a_j - a_k)}. \tag{17}$$

Let us compare the coefficients of X^n on both sides of this equality.

- The coefficient of X^n on the left hand side is c_n (since c_n is defined as the coefficient of X^n in $P = P(X)$).
- The right hand side is a bit more complicated. For each $j \in \{0, 1, \ldots, n\}$, the product $\prod\limits_{k\neq j} (X - a_k)$ is a product of n monic polynomials of degree 1 (since each $X - a_k$ is monic of degree 1), and thus itself is a monic polynomial of degree n (by Proposition 8.2.1 **(b)** and **(d)**, applied many times). The coefficient of X^n in this product is therefore 1. Thus, the coefficient of X^n in $b_j \frac{\prod\limits_{k\neq j} (X - a_k)}{\prod\limits_{k\neq j} (a_j - a_k)}$ is

$$\underbrace{b_j}_{\substack{=P(a_j)\\ \text{(by the definition}\\ \text{of } b_j)}} \frac{1}{\prod\limits_{k\neq j} (a_j - a_k)} = P(a_j) \frac{1}{\prod\limits_{k\neq j} (a_j - a_k)} = \frac{P(a_j)}{\prod\limits_{k\neq j} (a_j - a_k)}.$$

This holds for each $j \in \{0,1,\ldots,n\}$. Hence, the coefficient of X^n in the sum $\sum\limits_{j=0}^{n} b_j \dfrac{\prod\limits_{k \neq j} (X - a_k)}{\prod\limits_{k \neq j} (a_j - a_k)}$ is

$$\sum_{j=0}^{n} \frac{P(a_j)}{\prod\limits_{k \neq j} (a_j - a_k)}.$$

This is therefore the coefficient of X^n on the right hand side of (17).

Hence, comparing the coefficients of X^n on both sides of (17), we find

$$c_n = \sum_{j=0}^{n} \frac{P(a_j)}{\prod\limits_{k \neq j} (a_j - a_k)}.$$

This solves the exercise. □

> **Exercise 8.5.4.** Let $\mathbb{K}$ be one of the sets $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$ and $\mathbb{C}$. Let $n \in \mathbb{N}$. Let $P \in \mathbb{K}[X]$ be a polynomial of degree $\leq n$. Let $c_n = [X^n] P$ be the coefficient of X^n in P. (Note that $c_n = 0$ if P has degree $< n$.) Prove that
>
> $$\sum_{i=0}^{n} (-1)^{n-i} \binom{n}{i} P(i) = c_n \cdot n!.$$

Solution idea. Exercise 8.5.3 (applied to $a_i = i$) yields

$$\sum_{j=0}^{n} \frac{P(j)}{\prod\limits_{k \neq j} (j-k)} = c_n$$

(where the "$\prod\limits_{k \neq j}$" sign means a product over all $k \in \{0,1,\ldots,n\}$ satisfying $k \neq j$). Multiplying this equality by $n!$, we obtain

$$\sum_{j=0}^{n} \frac{P(j)}{\prod\limits_{k \neq j} (j-k)} \cdot n! = c_n \cdot n!. \tag{18}$$

However, each $j \in \{0,1,\ldots,n\}$ satisfies

$$\prod_{k \neq j} (j-k) = \underbrace{\left(\prod_{k=0}^{j-1} (j-k) \right)}_{\substack{=(j-0)(j-1)\cdots(j-(j-1)) \\ =j(j-1)\cdots 1 \\ =1\cdot 2\cdot\cdots\cdot j \\ =j!}} \cdot \underbrace{\left(\prod_{k=j+1}^{n} (j-k) \right)}_{\substack{=(j-(j+1))(j-(j+2))\cdots(j-n) \\ =(-1)(-2)\cdots(-(n-j)) \\ =(-1)^{n-j}(1\cdot 2\cdot\cdots\cdot(n-j)) \\ =(-1)^{n-j}(n-j)!}} = j! \cdot (-1)^{n-j} (n-j)!$$

and therefore

$$\frac{P(j)}{\prod\limits_{k\neq j}(j-k)}\cdot n! = \frac{P(j)}{j!\cdot(-1)^{n-j}(n-j)!}\cdot n! = (-1)^{n-j}\cdot \underbrace{\frac{n!}{j!\cdot(n-j)!}}_{\substack{=\binom{n}{j} \\ \text{(by the factorial formula} \\ \text{for binomial coefficients)}}} P(j)$$

$$= (-1)^{n-j}\binom{n}{j}P(j). \tag{19}$$

Hence, (18) can be rewritten as

$$\sum_{j=0}^{n}(-1)^{n-j}\binom{n}{j}P(j) = c_n\cdot n!.$$

Renaming the index i as j on the left hand side, we obtain precisely the claim of Exercise 8.5.4. □

The above two exercises just scratch the surface. Further applications of Lagrange interpolation can be found in [Nica23], [AndDos10, Chapter 11], [AndDos12, Chapter 11] and [Grinbe23, §4.3.7].

8.5.5. The integer and rational root tests

As we already said, finding roots of a polynomial (both exactly and numerically) is a difficult task in general. However, some simple cases of this task can be done in easy ways. One such case is when you are looking for rational roots of a polynomial with rational coefficients. By multiplying such a polynomial with the lowest common denominator of its coefficients, we can always make its coefficients become integers, so that our task becomes to find the rational roots of a polynomial with integer coefficients. And for this, there is a theorem ([Grinbe20, Theorem 3.5.14]):

> **Theorem 8.5.11** (rational root test)**.** Let $P = a_0X^0 + a_1X^1 + \cdots + a_nX^n \in \mathbb{Z}[X]$ be a polynomial with integer coefficients (i.e., all of $a_0, a_1, \ldots, a_n$ are integers). Let r be a rational root of P (that is, a rational number satisfying $P(r) = 0$). Write r as a reduced fraction, i.e., in the form $r = p/q$ for some coprime integers p and q. Then, $p \mid a_0$ and $q \mid a_n$.

When the two coefficients a_0 and a_n in Theorem 8.5.11 are nonzero, the theorem leads to an algorithm for finding all roots of P, since the divisibilities $p \mid a_0$ and $q \mid a_n$ leave only finitely many options for p and q (because a nonzero integer has only finitely many divisors), and we can just try all possible combinations of divisors of p and q and check which of them are roots of P. (Don't forget that divisors can be negative!) The case when a_0 or a_n is zero can be easily dealt with[10].

[10] If $a_0 = 0$, then we can divide P by X. If $a_n = 0$, then we can drop the zero coefficient and replace n by $n-1$.

The proof of Theorem 8.5.11 is easy and nice, and can be found in [Grinbe20, Theorem 3.5.14], but I recommend treating it as a good exercise.

As a particular case of Theorem 8.5.11, we obtain the following:

Corollary 8.5.12. Let $P \in \mathbb{Z}[X]$ be a **monic** polynomial with integer coefficients. Then, any rational root of P is an integer.

Proof idea. Write P as $P = a_0X^0 + a_1X^1 + \cdots + a_nX^n$, where $n = \deg P$. Then, $a_n = 1$ (since P is monic).

Now, let r be a rational root of P. We must prove that r is an integer.

Write r as a reduced fraction, i.e., in the form $r = p/q$ for some coprime integers p and q. Then, Theorem 8.5.11 yields $p \mid a_0$ and $q \mid a_n$. Hence, $q \mid a_n = 1$, so that $q = \pm 1$, and thus $p/q = p/(\pm 1) = \pm p \in \mathbb{Z}$. Hence, $r = p/q \in \mathbb{Z}$. In other words, r is an integer, qed. □

This corollary easily yields the following classical result (generalizing the irrationality of $\sqrt{2}$):

Corollary 8.5.13. Let a be an integer. Let r be a positive integer. If $\sqrt[r]{a}$ is not an integer, then $\sqrt[r]{a}$ is irrational.

Proof idea. The polynomial $X^r - a$ is monic. Thus, Corollary 8.5.12 shows that any rational root of $X^r - a$ is an integer. Hence, if $\sqrt[r]{a}$ is rational, then $\sqrt[r]{a}$ is an integer (since $\sqrt[r]{a}$ is a root of $X^r - a$). Taking the contrapositive, we conclude that if $\sqrt[r]{a}$ is not an integer, then $\sqrt[r]{a}$ is irrational. Qed. □

For another proof of Corollary 8.5.13, see [Grinbe20, Exercise 9.3.2].

8.5.6. Real roots

Having discussed integer and rational roots, let us turn to real roots. There are algorithms for determining the number of real roots of a polynomial in $\mathbb{Q}[X]$, but these are more advanced than would be appropriate for this worksheet[11]. I will thus restrict myself to a simple example and a simple result:

Example 8.5.14. Let $n \in \mathbb{N}$. Then, the polynomial $X^{2n} + 1$ has no real roots, since every real x satisfies $x^{2n} + 1 > x^{2n} = (x^n)^2 \geq 0$ (because squares of reals are always ≥ 0).

Thus, you should not expect a polynomial of even degree to have any real roots (although many such polynomials have them). But polynomials of odd degree are forced to have at least one:

[11] The reason why I am stating this for $\mathbb{Q}[X]$ instead of $\mathbb{R}[X]$ is technical: Real numbers cannot be precisely represented on a computer (due to their infinitary nature), so it is not clear what an algorithm would even be that takes a polynomial in $\mathbb{R}[X]$ as input. Of course, numerical algorithms exist, but the problem is not always well-posed.

> **Proposition 8.5.15.** Let $P \in \mathbb{R}[X]$ be a polynomial of odd degree. Then, P has at least one real root.

Proof sketch. We WLOG assume that P is monic (otherwise, we divide P by its leading coefficient; this makes P monic without changing its roots). Thus, we can write P as

$$P = c_0 X^0 + c_1 X^1 + \cdots + c_{n-1} X^{n-1} + X^n$$

for some $c_0, c_1, \ldots, c_{n-1} \in \mathbb{R}$. Hence, for each $x \in \mathbb{R}$, we have

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

As the real number x increases towards $+\infty$, the x^n addend here becomes large and outgrows all other $c_i x^i$ addends, and thus the sum $P(x)$ eventually becomes positive[12]. On the other hand, as x decreases towards $-\infty$, the x^n addend becomes small (since n is odd) and outgrows (in the negative sense) all other $c_i x^i$ addends, and therefore the sum $P(x)$ eventually becomes negative[13]. This shows that the polynomial function

$$\begin{aligned} f_P : \mathbb{R} &\to \mathbb{R}, \\ x &\mapsto P(x) \end{aligned}$$

takes both positive and negative values. Since this function is continuous, it must therefore take the value 0 as well (by the intermediate value theorem). In other words, P has a real root. □

8.6. Viete's theorem and the fundamental theorem of algebra

8.6.1. The multiplicity of a root

Theorems 8.5.2 and 8.5.3 show that each root of a polynomial P corresponds to a linear factor $X - r$ that can be "split off" from P (that is, P can be divided by this factor without remainder). For example, the polynomial $(X-4)(X-5)(X-6)$ has the three roots $4, 5, 6$, each corresponding to one of its three factors. But this is not always a one-to-one correspondence: For instance, the polynomial $(X-4)^2(X-5)$ has the two roots 4 and 5, because its two $X - 4$ factors correspond to the same root. This seems somewhat unfair. It appears more appropriate to count the root 4 twice, as it corresponds to two linear factors. This is done by the following definition:

> **Definition 8.6.1.** Let $\mathbb{K}$ be one of the sets $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$ and $\mathbb{C}$. Let $P \in \mathbb{K}[X]$ be a polynomial. Let $r \in \mathbb{K}$ be a number. Then, the *multiplicity of r as a root of P* is defined to be the largest $m \in \mathbb{N}$ such that $(X-r)^m \mid P$ (in $\mathbb{K}[X]$). In other words, it is the number of times we can divide P by $X - r$.

[12] This will definitely happen for $x > |c_0| + |c_1| + \cdots + |c_{n-1}| + 1$.

[13] This will definitely happen for $x < -(|c_0| + |c_1| + \cdots + |c_{n-1}| + 1)$.

Note that we did not require r to be a root of P in the first place. If it isn't, then its multiplicity as a root of P is 0. Conversely, if r is a root of P, then its multiplicity as a root of P is ≥ 1.

For example, the polynomial $(X-4)^2 (X-5) (X-6)^3$ has three roots: 4, 5 and 6, with respective multiplicities 2, 1 and 3. Equivalently, we can say that it has the roots $4, 4, 5, 6, 6, 6$ (listed with multiplicities).

The following theorem generalizes Theorem 8.5.3 to roots with multiplicities:

> **Theorem 8.6.2** ("roots = factors" theorem with multiplicities)**.** Let $\mathbb{K}$ be one of the sets $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$ and $\mathbb{C}$. Let $P \in \mathbb{K}[X]$ be a polynomial. Let $r_1, r_2, \ldots, r_m \in \mathbb{K}$ be distinct numbers. Let $a_1, a_2, \ldots, a_m$ be nonnegative integers. Then, each r_i (for each $i \in \{1, 2, \ldots, m\}$) has multiplicity $\geq a_i$ as a root of P if and only if P is divisible by $(X-r_1)^{a_1} (X-r_2)^{a_2} \cdots (X-r_m)^{a_m}$ (in $\mathbb{K}[X]$).

Proof idea. $\Longleftarrow$: The "$\Longleftarrow$" direction is easy (just as in Theorem 8.5.3).

$\Longrightarrow$: We induct on m, similarly to Theorem 8.5.3. The *base case* is again obvious.

In the *induction step*, we now start out by writing P as $P = (X-r_m)^{a_m} \cdot Q$ for some $Q \in \mathbb{K}[X]$ (this can be done, since r_m has multiplicity $\geq a_m$ as a root of P, so that $(X-r_m)^{a_m} \mid P$). Consider this Q. Now, the tricky part is to argue that the numbers $r_1, r_2, \ldots, r_{m-1}$ still have multiplicities $\geq a_1, \ \geq a_2, \ \ldots, \ \geq a_{m-1}$ as roots of Q. To this purpose, we show the following fact:

> *Claim 1:* Let r and s be two distinct elements of $\mathbb{K}$. Let a and b be nonnegative integers. Let $S \in \mathbb{K}[X]$ be a polynomial such that $(X-s)^b \mid (X-r)^a \cdot S$. Then, $(X-s)^b \mid S$.

Proof of Claim 1. We induct on b.

Base case: For $b = 0$, we have $(X-s)^b = (X-s)^0 = 1 \mid S$, so that Claim 1 holds.

Induction step: Let b be a positive integer. Assume (as the induction hypothesis) that Claim 1 has already been proved for $b-1$ instead of b.

Now assume that $(X-s)^b \mid (X-r)^a \cdot S$. We must show that $(X-s)^b \mid S$.

Since b is positive, we have $X - s \mid (X-s)^b \mid (X-r)^a \cdot S$, and thus s is a root of the polynomial $(X-r)^a \cdot S$ (by Theorem 8.5.2). In other words, $(s-r)^a \cdot S(s) = 0$. Since the $(s-r)^a$ factor is nonzero (because r and s are distinct), we can cancel it and obtain $S(s) = 0$. Hence, s is a root of S, and therefore we have $X - s \mid S$ (by Theorem 8.5.2). Hence, $S = (X-s) \cdot T$ for some $T \in \mathbb{K}[X]$. Consider this T. We have

$$\begin{aligned} (X-s) \cdot (X-s)^{b-1} &= (X-s)^b \\ &\mid (X-r)^a \cdot \underbrace{S}_{=(X-s)\cdot T} = (X-r)^a \cdot (X-s) \cdot T \\ &= (X-s) \cdot (X-r)^a \cdot T. \end{aligned} \tag{20}$$

However, it is easy to see that a nonzero polynomial can always be cancelled from a divisibility: If three polynomials A, B, C satisfy $AB \mid AC$ and $A \neq 0$, then $B \mid C$.

Thus, cancelling $X-s$ from the divisibility (20), we obtain $(X-s)^{b-1} \mid (X-r)^a \cdot T$. By our induction hypothesis, we can thus apply Claim 1 to $b-1$ and T instead of b and S. We conclude that $(X-s)^{b-1} \mid T$. Hence,

$$(X-s)^b = (X-s) \cdot \underbrace{(X-s)^{b-1}}_{\mid T} \mid (X-s) \cdot T = S.$$

This concludes the induction step. Thus, Claim 1 is proved. □

Now, for each $i \in \{1,2,\ldots,m-1\}$, the number r_i has a multiplicity $\geq a_i$ as a root of P. Hence, for each $i \in \{1,2,\ldots,m-1\}$, we have $(X-r_i)^{a_i} \mid P = (X-r_m)^{a_m} \cdot Q$, and thus $(X-r_m)^{a_m} \mid Q$ (by Claim 1, applied to $r = r_m$ and $s = r_i$ and $a = a_m$ and $b = a_i$, since r_m and r_i are distinct). In other words, the numbers $r_1, r_2, \ldots, r_{m-1}$ still have multiplicities $\geq a_1$, $\geq a_2$, $\ldots$, $\geq a_{m-1}$ as roots of Q. Hence, we can apply our induction hypothesis, and this easily finishes the induction step. □

The multiplicity of a root of a polynomial can also be described using the derivative. Let us first define this derivative:

Definition 8.6.3. Let $\mathbb{K}$ be one of the sets $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$ and $\mathbb{C}$. The *derivative* P' of a polynomial $P = c_0X^0 + c_1X^1 + c_2X^2 + \cdots \in \mathbb{K}[X]$ is defined to be the polynomial $1c_1X^0 + 2c_2X^1 + 3c_3X^2 + \cdots \in \mathbb{R}[X]$.

This definition, of course, has been chosen deliberately to match the formula for the derivative of a polynomial **function** known from calculus. Standard laws of differentiation (such as the Leibniz rule $(PQ)' = P'Q + PQ'$ and the chain rule $(P(Q))' = P'(Q) \cdot Q'$) are easily shown to hold for polynomials.

Using the notion of a derivative, we can define the m-th derivative for each $m \in \mathbb{N}$:

Definition 8.6.4. Let $\mathbb{K}$ be one of the sets $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$ and $\mathbb{C}$. Let $m \in \mathbb{N}$. The *m-th derivative* $P^{(m)}$ of a polynomial $P \in \mathbb{K}[X]$ is defined to be the polynomial $\left(\left(\left((P')'\right)'\right)\cdots\right)'$, with m many primes. (In other words, it is defined recursively by $P^{(0)} = P$ and $P^{(m)} = \left(P^{(m-1)}\right)'$ for each $m > 0$.)

For example, if $P = 2X^3 - 7X + 1$, then $P^{(0)} = P = 2X^3 - 7X + 1$ and $P^{(1)} = P' = 6X^2 - 7$ and $P^{(2)} = P'' = 12X$ and $P^{(3)} = P''' = 12$ and $P^{(m)} = 0$ for all $m > 3$.

The connection between higher derivatives and multiplicities of roots is as follows:

Theorem 8.6.5. Let $\mathbb{K}$ be one of the sets $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$ and $\mathbb{C}$. Let $P \in \mathbb{K}[X]$ be a polynomial. Let $r \in \mathbb{K}$ be a number. Then, the multiplicity of r as a root of P is the smallest $m \in \mathbb{N}$ such that $P^{(m)}(r) \neq 0$.

For example, if $P = (X-4)^2(X-5)$, then $P(4) = P'(4) = 0$ but $P''(4) \neq 0$, so that the multiplicity of 4 as a root of P is 2.

Theorem 8.6.5 is not hard to prove using the Leibniz rule, but we omit this proof.

8.6.2. The Fundamental Theorem of Algebra

As we have seen in Example 8.5.14, a polynomial over $\mathbb{R}$ can be of arbitrarily high degree without having a single real root. The situation for complex roots is completely different, as witnessed by the celebrated *fundamental theorem of algebra*:

> **Theorem 8.6.6** (fundamental theorem of algebra)**.** Let $n \in \mathbb{N}$. Let $P \in \mathbb{C}[X]$ be a polynomial of degree n. Let λ be the leading coefficient of P. Then, P can be written as
> $$P = \lambda (X - r_1)(X - r_2) \cdots (X - r_n)$$
> for n complex numbers $r_1, r_2, \ldots, r_n \in \mathbb{C}$. Thus, P has n complex roots, counted with multiplicity.

This theorem (first proved by Argand in 1806, after many attempts and half-proofs by Euler, Gauss, Lagrange and other luminaries) is one of the main reasons for the popularity of the complex numbers! In a sense, it reveals that $\mathbb{C}$ is the best place to look for roots, even if one is only interested in polynomials in $\mathbb{Z}[X]$ or $\mathbb{Q}[X]$ or $\mathbb{R}[X]$.

Proofs of Theorem 8.6.6 can be found in [LaNaSc16, Chapter 3], [Aluffi16, Theorem 7.1], [Knapp16, Chapter IX, §10], [Warner90, Theorem 44.8], [Steinb06, Theorem 11.6.7] and many other places[14]. More exotic proofs are listed in `https://mathoverflow.net/questions/10535` .

Ironically, Theorem 8.6.6 is nowhere near fundamental to modern algebra. Algebraists don't need to hunt for their roots in $\mathbb{C}$; they can always adjoin them to their favorite field[15]! But Theorem 8.6.6 is fundamentally important to anything analytic done with roots of polynomials, such as inequalities, and such things can be useful even in seemingly unrelated situations. For example, Problem A6 of the Putnam contest 2021 is a purely number-theoretical statement about polynomials with integer coefficients, but the only solutions known involve taking its complex roots, whose existence is guaranteed by Theorem 8.6.6. This is not an isolated occurrence; "pick a root" is a useful technique whenever it comes to proving properties of polynomials. The following exercise illustrates this technique on a baby example which is easy enough to solve by hand but even easier using roots:

> **Exercise 8.6.1.** Let $P \in \mathbb{Z}[X]$ be a monic quadratic polynomial. Show that there exist two quadratic polynomials $Q, R \in \mathbb{Z}[X]$ such that $P(X) \cdot P(X+1) = Q(R(X))$.

[14] Some of these sources only prove that any non-constant polynomial in $\mathbb{C}[X]$ has at least one complex root. This fact sounds weaker than Theorem 8.6.6, but actually it suffices to prove the whole theorem, since you can apply this fact to find a root r_1 of P, then divide P by $X - r_1$, then apply this fact again to the quotient to find a further root r_2, then divide the quotient by $X - r_2$, then apply this fact again to the new quotient, and so on, until you have found n roots and divided P by n linear factors.

[15] See [Grinbe23, §4.5] for how this works.

Solution idea. The polynomial P has degree 2 and leading coefficient 1. Thus, Theorem 8.6.6 (applied to $n = 2$ and $\lambda = 1$) shows that P can be written as

$$P = (X - r_1)(X - r_2)$$

for 2 complex numbers $r_1, r_2 \in \mathbb{C}$. Consider these r_1, r_2. (Arguably, this can also be obtained from the quadratic formula; our use of Theorem 8.6.6 was overkill. But imagine a problem about a polynomial of higher degree...)

Substituting $X + 1$ for X in $P = (X - r_1)(X - r_2)$, we obtain

$$P(X+1) = (X + 1 - r_1)(X + 1 - r_2).$$

Thus,

$$\begin{aligned}
&\underbrace{P(X)}_{=P=(X-r_1)(X-r_2)} \cdot \underbrace{P(X+1)}_{=(X+1-r_1)(X+1-r_2)} \\
&= (X - r_1)(X - r_2) \cdot (X + 1 - r_1)(X + 1 - r_2) \\
&= \underbrace{(X - r_1)(X + 1 - r_2)}_{=X^2-(r_1+r_2-1)X+r_1r_2-r_1} \cdot \underbrace{(X - r_2)(X + 1 - r_1)}_{=X^2-(r_1+r_2-1)X+r_1r_2-r_2} \\
&= \left(X^2 - (r_1 + r_2 - 1)X + r_1r_2 - r_1\right) \cdot \left(X^2 - (r_1 + r_2 - 1)X + r_1r_2 - r_2\right).
\end{aligned}$$

Setting $R := X^2 - (r_1 + r_2 - 1)X + r_1r_2$, we can rewrite this as

$$P(X) \cdot P(X+1) = (R - r_1) \cdot (R - r_2).$$

Setting $Q := (X - r_1)(X - r_2)$, we can furthermore rewrite this as

$$P(X) \cdot P(X+1) = Q(R) = Q(R(X)) \qquad (\text{since } R = R(X)).$$

Are we done? Not quite, since we want our polynomials Q and R to belong to $\mathbb{Z}[X]$, but their construction does not make this obvious (after all, the roots r_1 and r_2 are not usually integers).

But this is not overly hard to check either: The polynomial Q belongs to $\mathbb{Z}[X]$ since $Q = (X - r_1)(X - r_2) = P \in \mathbb{Z}[X]$. The polynomial R belongs to $\mathbb{Z}[X]$ since

$$R = X^2 - (r_1 + r_2 - 1)X + r_1r_2 = \underbrace{(X - r_1)(X - r_2)}_{=P} + X = P + X$$

(and again since $P \in \mathbb{Z}[X]$).

At this point, we have solved the problem. Of course, you might wonder whether we actually need the roots r_1 and r_2 if we found out that $Q = P$ and $R = P + X$ at the end. And indeed, we don't. We could just as well have defined the polynomials Q and P by $Q := P$ and $R := P + X$, and checked that $P(X) \cdot P(X+1) = Q(R(X))$

by direct computation: Writing the monic quadratic polynomial P as $P = X^2 + aX + b$ (for $a, b \in \mathbb{Z}$), we have

$$\begin{aligned} P(X+1) &= (X+1)^2 + a(X+1) + b = \left(X^2 + 2X + 1\right) + (aX + a) + b \\ &= \underbrace{X^2 + aX + b}_{=P} + 2X + 1 + a \\ &= P + 2X + 1 + a \qquad (21) \end{aligned}$$

and

$$\begin{aligned} Q(R(X)) &= \underbrace{P(P+X)}_{\substack{\text{the value of } P \text{ at } P+X, \\ \text{not the product of } P \text{ and } P+X!}} \qquad (\text{since } Q = P \text{ and } R(X) = R = P + X) \\ &= (P+X)^2 + a(P+X) + b \qquad \left(\text{since } P = X^2 + aX + b\right) \\ &= \left(P^2 + 2PX + X^2\right) + (aP + aX) + b = P(P + 2X + a) + \underbrace{\left(X^2 + aX + b\right)}_{=P} \\ &= P(P + 2X + a) + P = \underbrace{P}_{=P(X)} \cdot \underbrace{(P + 2X + a + 1)}_{\substack{=P+2X+1+a \\ =P(X+1) \\ (\text{by (21)})}} = P(X) \cdot P(X+1), \end{aligned}$$

so that $P(X) \cdot P(X+1) = Q(R(X))$. In hindsight, this is a more direct solution than the one using r_1 and r_2. But it would probably not have been easy to find without motivation. □

8.6.3. Viete's theorem

Viete's identities (also known as *Viete's formulas* or *Viete's theorem*[16]) are relations between the coefficients of a polynomial and its n complex roots. For the sake of simplicity, we state them for monic polynomials only[17]:

Theorem 8.6.7. Let $r_1, r_2, \ldots, r_n \in \mathbb{C}$ be n complex numbers. Consider the polynomial

$$P := (X - r_1)(X - r_2) \cdots (X - r_n).$$

Write this polynomial (which is clearly monic of degree n) in the form

$$P = a_0 X^0 + a_1 X^1 + \cdots + a_n X^n$$

[16] Viete (François Viète, 1540–1603) is also often spelled "Viète" or "Vieta".

[17] This is sufficient for all practical purposes, since we can make any nonzero polynomial monic by dividing it by its leading coefficient. (Clearly, this operation does not change the roots of the polynomial.)

with $a_0, a_1, \ldots, a_n \in \mathbb{C}$. Then, we have the $n+1$ identities

$$\begin{aligned}
1 &= a_n; \\
r_1 + r_2 + \cdots + r_n &= -a_{n-1}; \\
\underbrace{r_1r_2 + r_1r_3 + \cdots + r_{n-1}r_n}_{\text{This is the sum of all products } r_ir_j \text{ with } i<j} &= a_{n-2}; \\
\underbrace{r_1r_2r_3 + r_1r_2r_4 + \cdots + r_{n-2}r_{n-1}r_n}_{\text{This is the sum of all products } r_ir_jr_k \text{ with } i<j<k} &= -a_{n-3}; \\
&\ldots; \\
r_1r_2\cdots r_n &= (-1)^n a_0.
\end{aligned}$$

In other words, for each $k \in \{0, 1, \ldots, n\}$, we have

$$\sum_{i_1<i_2<\cdots<i_k} r_{i_1}r_{i_2}\cdots r_{i_k} = (-1)^k a_{n-k},$$

where the sum on the left hand side ranges over all k-tuples $(i_1, i_2, \ldots, i_k) \in \{1, 2, \ldots, n\}^k$ satisfying $i_1 < i_2 < \cdots < i_k$.

Example 8.6.8. To get a bit of intuition, let us see what Theorem 8.6.7 says for $n = 3$. For simplicity, let us rename the numbers r_1, r_2, r_3 as u, v, w. Thus, Theorem 8.6.7 is saying that if the polynomial

$$P := (X-u)(X-v)(X-w).$$

is written in the form

$$P = a_0X^0 + a_1X^1 + a_2X^2 + a_3X^3$$

with $a_0, a_1, a_2, a_3 \in \mathbb{C}$, then

$$1 = a_3; \tag{22}$$
$$u + v + w = -a_2; \tag{23}$$
$$uv + uw + vw = a_1; \tag{24}$$
$$uvw = -a_0. \tag{25}$$

Proof idea for Theorem 8.6.7. Let $k \in \{0, 1, \ldots, n\}$. When we expand the product

$$(X-r_1)(X-r_2)\cdots(X-r_n),$$

we obtain a sum of 2^n addends, each of which has the form "a power of X times a product of some $-r_i$'s" (because from each factor $X - r_i$, we get to pick either the X or the $-r_i$). Moreover, the number of $-r_i$'s in the product plus the exponent over the X will always equal n (since there are n factors in total). Thus, the power of X in the addend will be $n-k$ precisely when the product contains k many $-r_i$'s. Hence, the coefficient of X^{n-k} in the product $(X-r_1)(X-r_2)\cdots(X-r_n)$ is

$$\sum_{i_1<i_2<\cdots<i_k} \underbrace{(-r_{i_1})(-r_{i_2})\cdots(-r_{i_k})}_{=(-1)^k r_{i_1}r_{i_2}\cdots r_{i_k}} = (-1)^k \sum_{i_1<i_2<\cdots<i_k} r_{i_1}r_{i_2}\cdots r_{i_k}.$$

Thus, comparing the coefficients of X^{n-k} on both sides of the equality

$$(X-r_1)(X-r_2)\cdots(X-r_n) = P = a_0X^0 + a_1X^1 + \cdots + a_nX^n,$$

we obtain

$$(-1)^k \sum_{i_1<i_2<\cdots<i_k} r_{i_1}r_{i_2}\cdots r_{i_k} = a_{n-k}$$

(since the coefficient on the right hand side is clearly a_{n-k}). Multiplying this equality by $(-1)^k$, we readily obtain

$$\sum_{i_1<i_2<\cdots<i_k} r_{i_1}r_{i_2}\cdots r_{i_k} = (-1)^k a_{n-k}.$$

This proves Theorem 8.6.7. □

Viete's theorem is helpful in a slew of contest-style problems; examples can be found in [AndEne12, §1.8] and various other places. Let me just give two examples:

> **Exercise 8.6.2.** Assume that the cubic polynomial $X^3+aX^2+bX+c \in \mathbb{C}[X]$ has three roots u,v,w (listed with multiplicities). Express $\dfrac{u^2}{u+1}+\dfrac{v^2}{v+1}+\dfrac{w^2}{w+1}$ in terms of a,b,c.

Solution idea. Let P be the cubic polynomial X^3+aX^2+bX+c. By assumption, this polynomial P has roots u,v,w (listed with multiplicities). Thus, by Theorem 8.6.2, we have

$$(X-u)(X-v)(X-w) \mid P.$$

That is, there exists a polynomial $Q \in \mathbb{C}[X]$ such that

$$P = (X-u)(X-v)(X-w)\cdot Q. \qquad (26)$$

Consider this Q. Then, comparing the degrees on both sides of (26), we see that $\deg Q = 0$ [18]. Hence, Q is constant, i.e., we have $Q=\lambda$ for some $\lambda \in \mathbb{C}$. Consider

[18] Here is this argument in more detail: From (26), we find

$$\begin{aligned}\deg P &= \deg\left((X-u)(X-v)(X-w)\cdot Q\right)\\ &= \underbrace{\deg\left((X-u)(X-v)(X-w)\right)}_{=3} + \deg Q \qquad \text{(by Proposition 8.2.1 \textbf{(b)})}\\ &= 3+\deg Q,\end{aligned}$$

this λ. Therefore, the leading coefficient on the right hand side of (26) is λ (since the polynomial $(X-u)(X-v)(X-w)$ has leading coefficient 1, while Q is just λ). Comparing this with the leading coefficient on the left hand side of (26) (which is visibly 1), we obtain $\lambda = 1$. Thus, $Q = \lambda = 1$. Hence, (26) simplifies to

$$P = (X-u)(X-v)(X-w). \tag{27}$$

Compare this to $P = X^3 + aX^2 + bX + c = cX^0 + bX^1 + aX^2 + 1X^3$. Therefore, we can apply Viete's theorem (specifically, the formulas (23), (24) and (25)) to $a_0 = c$ and $a_1 = b$ and $a_2 = a$ and $a_3 = 1$. We obtain

$$u+v+w = -a; \tag{28}$$

$$uv+uw+vw = b; \tag{29}$$

$$uvw = -c. \tag{30}$$

But we are looking for the sum $\dfrac{u^2}{u+1} + \dfrac{v^2}{v+1} + \dfrac{w^2}{w+1}$. Thus, we could use a way to express this sum in terms of $u+v+w$ and $uv+uw+vw$ and uvw.

There are several ways to find such an expression; here is the easiest one: Let us reduce the numerators of our three fractions. Polynomial long division shows that $X^2 = (X-1)(X+1)+1$. Thus, $u^2 = (u-1)(u+1)+1$, so that $\dfrac{u^2}{u+1} = u - 1 + \dfrac{1}{u+1}$. Similarly, $\dfrac{v^2}{v+1} = v - 1 + \dfrac{1}{v+1}$ and $\dfrac{w^2}{w+1} = w - 1 + \dfrac{1}{w+1}$. Summing these three equalities, we obtain

$$\begin{aligned}
&\frac{u^2}{u+1} + \frac{v^2}{v+1} + \frac{w^2}{w+1} \\
&= \left(u - 1 + \frac{1}{u+1}\right) + \left(v - 1 + \frac{1}{v+1}\right) + \left(w - 1 + \frac{1}{w+1}\right) \\
&= \underbrace{(u+v+w)}_{=-a} - 3 + \left(\frac{1}{u+1} + \frac{1}{v+1} + \frac{1}{w+1}\right).
\end{aligned}$$

It thus remains to compute $\dfrac{1}{u+1} + \dfrac{1}{v+1} + \dfrac{1}{w+1}$. This we can just do by brute

so that

$$\deg Q = \underbrace{\deg P}_{\substack{=3 \\ \text{(since } P \text{ is cubic)}}} - 3 = 3 - 3 = 0.$$

force:

$$\begin{aligned}\frac{1}{u+1}+\frac{1}{v+1}+\frac{1}{w+1} &= \frac{(v+1)(w+1)+(u+1)(w+1)+(u+1)(v+1)}{(u+1)(v+1)(w+1)}\\ &= \frac{(vw+v+w+1)+(uw+u+w+1)+(uv+u+v+1)}{uvw+uv+vw+wu+u+v+w+1}\\ &= \frac{(uv+uw+vw)+2(u+v+w)+3}{uvw+(uv+uw+vw)+(u+v+w)+1}\\ &= \frac{b+2(-a)+3}{(-c)+b+(-a)+1} \qquad \text{(by (28), (29) and (30))}\\ &= \frac{b-2a+3}{b+1-c-a}.\end{aligned}$$

Hence,

$$\begin{aligned}\frac{u^2}{u+1}+\frac{v^2}{v+1}+\frac{w^2}{w+1} &= \underbrace{(u+v+w)}_{=-a}-3+\underbrace{\left(\frac{1}{u+1}+\frac{1}{v+1}+\frac{1}{w+1}\right)}_{=\frac{b-2a+3}{b+1-c-a}}\\ &= -a-3+\frac{b-2a+3}{b+1-c-a}.\end{aligned}$$

Not the nicest expression, but it does answer the problem! □

> **Exercise 8.6.3.** Assume that the cubic polynomial $X^3+aX^2+bX+c\in\mathbb{C}[X]$ has three roots u,v,w (listed with multiplicities). Find a cubic polynomial that has the three roots u^2,v^2,w^2, and express it in terms of a,b,c.

Solution idea. As in the solution to Exercise 8.6.2, we can find that

$$u+v+w=-a; \tag{31}$$

$$uv+uw+vw=b; \tag{32}$$

$$uvw=-c. \tag{33}$$

Now, we are looking for a cubic polynomial that has the three roots u^2,v^2,w^2. Clearly, $(X-u^2)(X-v^2)(X-w^2)$ is such a polynomial, but we want to express it in terms of a,b,c. Expanding it, we find

$$\begin{aligned}&\left(X-u^2\right)\left(X-v^2\right)\left(X-w^2\right)\\ &= X^3-\left(u^2+v^2+w^2\right)X^2+\left(u^2v^2+u^2w^2+v^2w^2\right)X-u^2v^2w^2.\end{aligned}$$

Thus, we must express the three coefficients $u^2+v^2+w^2$, $u^2v^2+u^2w^2+v^2w^2$ and $u^2v^2w^2$ in terms of a,b,c.

For $u^2v^2w^2$, this is very easy: We have $u^2v^2w^2 = (uvw)^2 = (-c)^2$ (by (33)).

For $u^2 + v^2 + w^2$, we recall the well-known trinomial formula

$$(u+v+w)^2 = u^2 + v^2 + w^2 + 2(uv + uw + vw)$$

(which follows by expanding the left hand side). Solving this for $u^2 + v^2 + w^2$, we find

$$\begin{aligned} u^2 + v^2 + w^2 &= (u+v+w)^2 - 2(uv + uw + vw) \qquad (34) \\ &= (-a)^2 - 2b \end{aligned}$$

(by (31) and (32)).

Finally, for $u^2v^2 + u^2w^2 + v^2w^2$, we note that

$$\begin{aligned} u^2v^2 + u^2w^2 + v^2w^2 &= (uv)^2 + (uw)^2 + (vw)^2 \\ &= (uv + uw + vw)^2 - 2\underbrace{(uv \cdot uw + uv \cdot vw + uw \cdot vw)}_{=uvw(u+v+w)} \\ &\qquad \left(\begin{array}{c}\text{by (34), applied to } uv,\ uw,\ vw \\ \text{instead of } u,\ v,\ w\end{array}\right) \\ &= (uv + uw + vw)^2 - 2uvw(u+v+w) \\ &= b^2 - 2(-c)(-a) \qquad \text{(by (31), (32) and (33))}. \end{aligned}$$

Altogether, the polynomial we are looking for is

$$\begin{aligned} &\left(X - u^2\right)\left(X - v^2\right)\left(X - w^2\right) \\ &= X^3 - \underbrace{\left(u^2 + v^2 + w^2\right)}_{\substack{=(-a)^2-2b \\ =a^2-2b}} X^2 + \underbrace{\left(u^2v^2 + u^2w^2 + v^2w^2\right)}_{\substack{=b^2-2(-c)(-a) \\ =b^2-2ca}} X - \underbrace{u^2v^2w^2}_{\substack{=(-c)^2 \\ =c^2}} \\ &= X^3 - \left(a^2 - 2b\right)X^2 + \left(b^2 - 2ca\right)X - c^2. \end{aligned}$$

□

8.7. Class problems

The following problems are to be discussed during class.

Exercise 8.7.1. Let $P \in \mathbb{K}[X]$ be a polynomial, where $\mathbb{K}$ is either $\mathbb{Z}$ or $\mathbb{Q}$ or $\mathbb{R}$ or $\mathbb{C}$. Prove the following:

(a) If $P(X) = P(-X)$, then P can be written as $P = Q\left(X^2\right)$ for some $Q \in \mathbb{K}[X]$.

(b) If $P(X) = -P(-X)$, then P can be written as $P = X \cdot Q\left(X^2\right)$ for some $Q \in \mathbb{K}[X]$.

Exercise 8.7.2. Let $n \in \mathbb{N}$. Let $P \in \mathbb{Q}[X]$ be a polynomial of degree $\leq n$ that satisfies $P(k) = 2^k$ for each $k \in \{0, 1, \ldots, n\}$. Find $P(n+1)$.

Exercise 8.7.3. Let $P \in \mathbb{R}[X]$ be a polynomial such that $P(X^2+1) = (P(X))^2 + 1$. Prove that P is an iterate of X^2+1 – that is, a polynomial of the form $\underbrace{S(S(\cdots(S(X))))}_{k \text{ times}}$ for some $k \in \mathbb{N}$, where $S := X^2 + 1$.

Exercise 8.7.4.

(a) Let $P \in \mathbb{C}[X]$ be an even polynomial (i.e., a polynomial satisfying $P(X) = P(-X)$). Prove that P can be written as $P = Q(X) \cdot Q(-X)$ for some $Q \in \mathbb{C}[X]$.

(b) Does this hold if $\mathbb{C}$ is replaced by $\mathbb{R}$?

Exercise 8.7.5. The polynomials $P_n \in \mathbb{R}[X]$ for all $n \in \mathbb{N}$ are defined recursively as follows: We set $P_0 = 1$ and define the polynomial P_n implicitly by the requirements

$$P_n' = nP_{n-1}(X+1) \qquad \text{and} \qquad P_n(0) = 0 \qquad \text{for all } n \geq 1$$

(where P_n' denotes the derivative of P_n). Factor the value $P_{100}(1)$ into prime numbers.

(46th Putnam contest 1985, Problem B2)

8.8. Homework exercises

This homework set is optional; it will not be graded.

Exercise 8.8.1. Recall the Fibonacci sequence $(f_0, f_1, f_2, \ldots)$, defined by $f_0 = 0$ and $f_1 = 1$ and $f_n = f_{n-1} + f_{n-2}$ for all $n \geq 2$.

Let $n \in \mathbb{N}$. Let $P \in \mathbb{Q}[X]$ be a polynomial of degree $\leq n$ that satisfies $P(k) = f_k$ for each $k \in \{0, 1, \ldots, n\}$. Find $P(n+1)$.

Exercise 8.8.2. Let $P \in \mathbb{R}[X]$ be a polynomial such that $P(\cos\alpha) = P(\sin\alpha)$ for all $\alpha \in \mathbb{R}$. Prove that $P = Q(X^4 - X^2)$ for some $Q \in \mathbb{R}[X]$.

Exercise 8.8.3. A polynomial $P \in \mathbb{Z}[X]$ is said to be *primitive* if the gcd of its coefficients is 1.

Let Q and R be two primitive polynomials in $\mathbb{Z}[X]$. Prove that their product QR is also primitive.

[Hint: A bunch of integers are coprime if and only if they have no common prime divisor.]

Exercise 8.8.4. Let $n \in \mathbb{N}$. Prove that we have $X^2 + X + 1 \mid X^{2n} + X^n + 1$ in $\mathbb{Z}[X]$ if and only if $3 \nmid n$ in $\mathbb{Z}$.

[Hint: First, define $N := X^2 + X + 1$, and show that $X^3 \equiv 1 \bmod N$ in $\mathbb{Z}[X]$.]

Exercise 8.8.5. Let $n \in \mathbb{N}$. Let $P \in \mathbb{Q}[X]$ be a polynomial of degree $\leq n$ that satisfies $P(k) = \dfrac{1}{k}$ for each $k \in \{1, 2, \ldots, n+1\}$. Show that $P(0) = \dfrac{1}{1} + \dfrac{1}{2} + \cdots + \dfrac{1}{n+1}$.

[Hint: Take the derivative on both sides of (12).]

Exercise 8.8.6. Let n be an even positive integer. Let $u_1, u_2, \ldots, u_n$ be the n roots of the polynomial $X^n - nX + 1$. Prove that

$$\frac{1}{u_1+1} + \frac{1}{u_2+1} + \cdots + \frac{1}{u_n+1} = \frac{2n}{n+2}.$$

[Hint: Find a degree-n polynomial with roots $\dfrac{1}{u_1+1}, \dfrac{1}{u_2+1}, \ldots, \dfrac{1}{u_n+1}$.]

Exercise 8.8.7. Let $P \in \mathbb{Z}[X]$ be a polynomial of even degree whose all coefficients (not counting the zero coefficients in front of powers that are larger than the degree) are odd. Prove that P has no rational root.

Exercise 8.8.8. Let $P = c_0 X^0 + c_1 X^1 + \cdots + c_n X^n$ be a polynomial in $\mathbb{Z}[X]$ (with $c_0, c_1, \ldots, c_n \in \mathbb{Z}$). Let $r \in \mathbb{Q}$ be a root of P. Prove that $c_n r^i + c_{n-1} r^{i-1} + \cdots + c_{n-i} r^0 \in \mathbb{Z}$ for each $i \in \{0, 1, \ldots, n\}$.

Exercise 8.8.9. Let $a_0, a_1, \ldots, a_n$ be n pairwise distinct integers (where $n \in \mathbb{N}$). Prove that for any $s \in \mathbb{N}$, the number

$$\sum_{j=0}^{n} \frac{a_j^s}{\prod\limits_{k \neq j} (a_j - a_k)}$$

(where the "$\prod\limits_{k \neq j}$" sign means a product over all $k \in \{0, 1, \ldots, n\}$ satisfying $k \neq j$) is an integer.

Exercise 8.8.10. Let $n \in \mathbb{N}$. Prove that

$$\sum_{i=0}^{n} (-1)^{n-i} \dbinom{n}{i} i^{n+1} = \frac{n\,(n+1)!}{2}.$$

[**Hint:** Alas, the polynomial $P := X^{n+1}$ has degree $n+1$, which is too high for an obvious application of Lagrange interpolation. Can you find a polynomial Q that has degree $\leq n$ but a sum $\sum\limits_{i=0}^{n} (-1)^{n-i} \dbinom{n}{i} Q\,(i)$ closely related to the sum on the left hand side?]

Exercise 8.8.11. Let $P \in \mathbb{R}[X]$ be a polynomial such that $P\,(x) \geq 0$ for all $x \in \mathbb{R}$. Prove that P can be written in the form $P = Q_1^2 + Q_2^2 + \cdots + Q_k^2$ for some polynomials $Q_1, Q_2, \ldots, Q_k \in \mathbb{R}[X]$.

[**Hint:** First, show that if two polynomials P_1 and P_2 can be written in this form, then so can their product $P_1 P_2$. Next, show that every polynomial of the type $(X - z)(X - \overline{z})$ for two conjugate complex numbers $z = a + bi$ and $\overline{z} = a - bi$ can be rewritten in this form. Finally, use the Fundamental Theorem of Algebra on P.]

References

[Aluffi16] Paolo Aluffi, *Algebra: Chapter 0*, Graduate Studies in Mathematics **104**, 2nd printing, AMS 2016.

[Aluffi21] Paolo Aluffi, *Algebra: Notes from the Underground*, Cambridge University Press 2021.

[AndDos10] Titu Andreescu, Gabriel Dospinescu, *Problems from the Book*, 2nd edition, XYZ Press 2010.

[AndDos12] Titu Andreescu, Gabriel Dospinescu, *Straight from the Book*, XYZ Press 2012.

[AndEne12] Titu Andreescu, Bogdan Enescu, *Mathematical Olympiad Treasures*, 2nd edition, Springer 2012.

[Axler23] Sheldon Axler, *Linear Algebra Done Right*, 4th edition, Springer 2023.

[Barbea89] Edward J. Barbeau, *Polynomials*, Springer 1989.

[BenQui03] Arthur T. Benjamin and Jennifer J. Quinn, *Proofs that Really Count: The Art of Combinatorial Proof*, Dolciani Mathematical Expositions **27**, The Mathematical Association of America, 2003.

[Engel98] Arthur Engel, *Problem-Solving Strategies*, Springer 1998.

[GelAnd17] Răzvan Gelca, Titu Andreescu, *Putnam and Beyond*, 2nd edition, Springer 2017.

[Granvi05] Andrew Granville, *Binomial coefficients modulo prime powers*, preprint. `https://web.archive.org/web/20181024055320/http://ebooks.bharathuniv.ac.in/gdlc1/gdlc1/EngineeringMergedLibraryv3.0/AndrewGranville/BinomialCoefficientsModuloPrimePowers(5579)/BinomialCoefficientsModuloPrimePowers-AndrewGranville.pdf`

[Grinbe19a] Darij Grinberg, *Introduction to Modern Algebra (UMN Spring 2019 Math 4281 notes)*, 1 October 2020.
`http://www.cip.ifi.lmu.de/~grinberg/t/19s/notes.pdf`

[Grinbe19b] Darij Grinberg, *Enumerative Combinatorics: class notes (Drexel Fall 2019 Math 222 notes)*, 11 September 2022.
`http://www.cip.ifi.lmu.de/~grinberg/t/19fco/n/n.pdf`

[Grinbe20] Darij Grinberg, *Math 235: Mathematical Problem Solving*, 10 August 2021.
`https://www.cip.ifi.lmu.de/~grinberg/t/20f/mps.pdf`

[Grinbe23] Darij Grinberg, *An introduction to the algebra of rings and fields (Text for Math 332 Winter 2023 at Drexel University)*, 27 August 2023.
`https://www.cip.ifi.lmu.de/~grinberg/t/23wa/23wa.pdf`

[GrKnPa94] Ronald L. Graham, Donald E. Knuth, Oren Patashnik, *Concrete Mathematics*, 2nd edition, Addison–Wesley 1994.

[Knapp16] Anthony W. Knapp, *Basic Algebra*, Digital 2nd edition 2016.
`http://www.math.stonybrook.edu/~aknapp/download.html`

[LaNaSc16] Isaiah Lankham, Bruno Nachtergaele, Anne Schilling, *Linear Algebra As an Introduction to Abstract Mathematics*, 2016.
`https://www.math.ucdavis.edu/~anne/linear_algebra/mat67_course_notes.pdf`

[Nica23] Bogdan Nica, *On an identity of Sylvester*, Expositiones Mathematicae **41**, Issue 4, December 2023, 125511.
Also available at arXiv:2212.13624v2.

[Prasol04] Victor V. Prasolov, *Polynomials*, Springer 2004.

[Steinb06] Mark Steinberger, *Algebra*, 31 August 2006.
`https://math.hawaii.edu/~tom/algebra.pdf`

[Warner90] Seth Warner, *Modern Algebra: two volumes bound as one*, Dover 1990.

[Zeitz17] Paul Zeitz, *The Art and Craft of Problem Solving*, 3rd edition, Wiley 2017.