Notes on the combinatorial fundamentals of algebra[*]

Darij Grinberg

September 15, 2022
(with minor corrections September 15, 2022)[†]

Abstract. This is a detailed survey – with rigorous and self-contained proofs – of some of the basics of elementary combinatorics and algebra, including the properties of finite sums, binomial coefficients, permutations and determinants. It is entirely expository (and written to a large extent as a repository for folklore proofs); no new results (and few, if any, new proofs) appear.

Contents

[*] old title: PRIMES 2015 reading project: problems and solutions

[†] The numbering in this version is compatible with that in the version of 10 January 2019 and in all intermediate versions.

Note

This is the version without solutions. See `http://www.cip.ifi.lmu.de/~grinberg/primes2015/sols.pdf` for the complete version.

1. Introduction

These notes are a detailed introduction to some of the basic objects of combinatorics and algebra: finite sums, binomial coefficients, permutations and determinants (from a combinatorial viewpoint – no linear algebra is presumed). To a lesser extent, modular arithmetic and recurrent integer sequences are treated as well. The reader is assumed to be proficient in high-school mathematics, and mature enough to understand nontrivial mathematical proofs. Familiarity with "contest mathematics" is also useful.

One feature of these notes is their focus on rigorous and detailed proofs. Indeed, so extensive are the details that a reader with experience in mathematics will probably be able to skip whole paragraphs of proof without losing the thread. (As a consequence of this amount of detail, the notes contain far less material than might be expected from their length.) Rigorous proofs mean that (with some minor exceptions) no "handwaving" is used; all relevant objects are defined in mathematical (usually set-theoretical) language, and are manipulated in logically well-defined ways. (In particular, some things that are commonly taken for granted in the literature – e.g., the fact that the sum of n numbers is well-defined without specifying in what order they are being added – are unpacked and proven in a rigorous way.)

These notes are split into several chapters:

- Chapter 1 collects some basic facts and notations that are used in later chapters. This chapter is **not** meant to be read first; it is best consulted when needed.

- Chapter 2 is an in-depth look at mathematical induction (in various forms, including strong and two-sided induction) and several of its applications (including basic modular arithmetic, division with remainder, Bezout's theorem, some properties of recurrent sequences, the well-definedness of compositions of n maps and sums of n numbers, and various properties thereof).

- Chapter 3 surveys binomial coefficients and their basic properties. Unlike most texts on combinatorics, our treatment of binomial coefficients leans to the algebraic side, relying mostly on computation and manipulations of sums; but some basics of counting are included.

- Chapter 4 treats some more properties of Fibonacci-like sequences, including explicit formulas (à la Binet) for two-term recursions of the form $x_n = ax_{n-1} + bx_{n-2}$.

- Chapter 5 is concerned with permutations of finite sets. The coverage is heavily influenced by the needs of the next chapter (on determinants); thus, a great role is played by transpositions and the inversions of a permutation.

- Chapter 6 is a comprehensive introduction to determinants of square matrices over a commutative ring[1], from an elementary point of view. This is probably the most unique feature of these notes: I define determinants using Leibniz's formula (i.e., as sums over permutations) and prove all their properties (Laplace expansion in one or several rows; the Cauchy-Binet, Desnanot-Jacobi and Plücker identities; the Vandermonde and Cauchy determinants; and several more) from this vantage point, thus treating them as an elementary object unmoored from its linear-algebraic origins and applications. No use is made of modules (or vector spaces), exterior powers, eigenvalues, or of the "universal coefficients" trick[2]. (This means that all proofs are done through combinatorics and manipulation of sums – a rather restrictive requirement!) This is a conscious and (to a large extent) aesthetic choice on my part, and I do **not** consider it the best way to learn about determinants; but I do regard it as a road worth charting, and these notes are my attempt at doing so.

The notes include numerous exercises of varying difficulty, many of them solved. The reader should treat exercises and theorems (and propositions, lemmas and

[1] The notion of a commutative ring is defined (and illustrated with several examples) in Section 6.1, but I don't delve deeper into abstract algebra.

[2] This refers to the standard trick used for proving determinant identities (and other polynomial identities), in which one first replaces the entries of a matrix (or, more generally, the variables appearing in the identity) by indeterminates, then uses the "genericity" of these indeterminates (e.g., to invert the matrix, or to divide by an expression that could otherwise be 0), and finally substitutes the old variables back for the indeterminates.

corollaries) as interchangeable to some extent; it is perfectly reasonable to read the solution of an exercise, or conversely, to prove a theorem on one's own instead of reading its proof. The reader's experience will be the strongest determinant of their success in solving the exercises independently.

I have not meant these notes to be a textbook on any particular subject. For one thing, their content does not map to any of the standard university courses, but rather straddles various subjects:

- Much of Chapter 3 (on binomial coefficients) and Chapter 5 (on permutations) is seen in a typical combinatorics class; but my focus is more on the algebraic side and not so much on the combinatorics.

- Chapter 6 studies determinants far beyond what a usual class on linear algebra would do; but it does not include any of the other topics that a linear algebra class usually covers (such as row reduction, vector spaces, linear maps, eigenvectors, tensors or bilinear forms).

- Being devoted to mathematical induction, Chapter 2 appears to cover the same ground as a typical "introduction to proofs" textbook or class (or at least one of its main topics). In reality, however, it complements rather than competes with most "introduction to proofs" texts I have seen; the examples I give are (with a few exceptions) nonstandard, and the focus different.

- While the notions of rings and groups are defined in Chapter 6, I cannot claim to really be doing any abstract algebra: I am merely working *in* rings (i.e., doing computations with elements of rings or with matrices over rings), rather than working *with* rings. Nevertheless, Chapter 6 might help familiarize the reader with these concepts, facilitating proper learning of abstract algebra later on.

All in all, these notes are probably more useful as a repository of detailed proofs than as a textbook to be read cover-to-cover. Indeed, one of my motives in writing them was to have a reference for certain folklore results – one in which these results are proved elementary and without appeal to the reader's problem-solving acumen.

These notes began as worksheets for the PRIMES reading project I have mentored in 2015; they have since been greatly expanded with new material (some of it originally written for my combinatorics classes, some in response to math.stackexchange questions).

The notes are in flux, and probably have their share of misprints. I thank Anya Zhang and Karthik Karnik (the two students taking part in the 2015 PRIMES project) for finding some errors, and Christian Krattenthaler for comments. Thanks also to the PRIMES project at MIT, which gave the impetus for the writing of this notes; and to George Lusztig for the sponsorship of my mentoring position in this project.

1.1. Prerequisites

Let me first discuss the prerequisites for a reader of these notes. At the current moment, I assume that the reader

- has a good grasp on basic school-level mathematics (integers, rational numbers, etc.);
- has some experience with proofs (mathematical induction, proof by contradiction, the concept of "WLOG", etc.) and mathematical notation (functions, subscripts, cases, what it means for an object to be "well-defined", etc.)[3];
- knows what a polynomial is (at least over $\mathbb{Z}$ and $\mathbb{Q}$) and how polynomials differ from polynomial functions[4];
- is somewhat familiar with the summation sign ($\sum$) and the product sign ($\prod$) and knows how to transform them (e.g., interchanging summations, and substituting the index)[5];
- has some familiarity with matrices (i.e., knows how to add and to multiply them)[6].

Probably a few more requirements creep in at certain points of the notes, which I have overlooked. Some examples and remarks rely on additional knowledge (such as analysis, graph theory, abstract algebra); however, these can be skipped.

1.2. Notations

- In the following, we use $\mathbb{N}$ to denote the set $\{0, 1, 2, \ldots\}$. (Be warned that some other authors use the letter $\mathbb{N}$ for $\{1, 2, 3, \ldots\}$ instead.)

[3] A great introduction into these matters (and many others!) is the free book [LeLeMe16] by Lehman, Leighton and Meyer. (**Practical note:** As of 2018, this book is still undergoing frequent revisions; thus, the version I am citing below might be outdated by the time you are reading this. I therefore suggest searching for possibly newer versions on the internet. Unfortunately, you will also find many older versions, often as the first google hits. Try searching for the title of the book along with the current year to find something up-to-date.)

Another introduction to proofs and mathematical workmanship is Day's [Day16] (but beware that the definition of polynomials in [Day16, Chapter 5] is the wrong one for our purposes). Two others are Hammack's [Hammac15] and Doud's and Nielsen's [DouNie19]. Yet another is Newstead's [Newste19] (currently a work in progress, but promising to become one of the most interesting and sophisticated texts of this kind). There are also several books on this subject; an especially popular one is Velleman's [Vellem06].

[4] This is used only in a few sections and exercises, so it is not an unalienable requirement. See Section 1.5 below for a quick survey of polynomials, and for references to sources in which precise definitions can be found.

[5] See Section 1.4 below for a quick overview of the notations that we will need.

[6] See, e.g., [Grinbe16b, Chapter 2] or any textbook on linear algebra for an introduction.

- We let $\mathbb{Q}$ denote the set of all rational numbers; we let $\mathbb{R}$ be the set of all real numbers; we let $\mathbb{C}$ be the set of all complex numbers[7].
- If X and Y are two sets, then we shall use the notation "$X \to Y,\ x \mapsto E$" (where x is some symbol which has no specific meaning in the current context, and where E is some expression which usually involves x) for "the map from X to Y which sends every $x \in X$ to E".

  For example, "$\mathbb{N} \to \mathbb{N},\ x \mapsto x^2 + x + 6$" means the map from $\mathbb{N}$ to $\mathbb{N}$ which sends every $x \in \mathbb{N}$ to $x^2 + x + 6$.

  For another example, "$\mathbb{N} \to \mathbb{Q},\ x \mapsto \dfrac{x}{1+x}$" denotes the map from $\mathbb{N}$ to $\mathbb{Q}$ which sends every $x \in \mathbb{N}$ to $\dfrac{x}{1+x}$. [8]
- If S is a set, then the *powerset* of S means the set of all subsets of S. This powerset will be denoted by $\mathcal{P}(S)$. For example, the powerset of $\{1,2\}$ is $\mathcal{P}(\{1,2\}) = \{\varnothing, \{1\}, \{2\}, \{1,2\}\}$.
- The letter i will **not** denote the imaginary unit $\sqrt{-1}$ (except when we explicitly say so).

Further notations will be defined whenever they arise for the first time.

1.3. Injectivity, surjectivity, bijectivity

In this section[9], we recall some basic properties of maps – specifically, what it means for a map to be injective, surjective and bijective. We begin by recalling basic definitions:

[7] See [Swanso20, Section 3.9] or [AmaEsc05, Section I.11] for a quick introduction to complex numbers. We will rarely use complex numbers. Most of the time we use them, you can instead use real numbers.

[8] A word of warning: Of course, the notation "$X \to Y,\ x \mapsto E$" does not always make sense; indeed, the map that it stands for might sometimes not exist. For instance, the notation "$\mathbb{N} \to \mathbb{Q},\ x \mapsto \dfrac{x}{1-x}$" does not actually define a map, because the map that it is supposed to define (i.e., the map from $\mathbb{N}$ to $\mathbb{Q}$ which sends every $x \in \mathbb{N}$ to $\dfrac{x}{1-x}$) does not exist (since $\dfrac{x}{1-x}$ is not defined for $x = 1$). For another example, the notation "$\mathbb{N} \to \mathbb{Z},\ x \mapsto \dfrac{x}{1+x}$" does not define a map, because the map that it is supposed to define (i.e., the map from $\mathbb{N}$ to $\mathbb{Z}$ which sends every $x \in \mathbb{N}$ to $\dfrac{x}{1+x}$) does not exist (for $x = 2$, we have $\dfrac{x}{1+x} = \dfrac{2}{1+2} \notin \mathbb{Z}$, which shows that a map from $\mathbb{N}$ to $\mathbb{Z}$ cannot send this x to this $\dfrac{x}{1+x}$). Thus, when defining a map from X to Y (using whatever notation), do not forget to check that it is well-defined (i.e., that your definition specifies precisely one image for each $x \in X$, and that these images all lie in Y). In many cases, this is obvious or very easy to check (I will usually not even mention this check), but in some cases, this is a difficult task.

[9] a significant part of which is copied from [Grinbe16b, §3.21]

- The words "map", "mapping", "function", "transformation" and "operator" are synonyms in mathematics.[10]

- A map $f : X \to Y$ between two sets X and Y is said to be *injective* if it has the following property:

  - If x_1 and x_2 are two elements of X satisfying $f(x_1) = f(x_2)$, then $x_1 = x_2$. (In words: If two elements of X are sent to one and the same element of Y by f, then these two elements of X must have been equal in the first place. In other words: An element of X is uniquely determined by its image under f.)

  Injective maps are often called "one-to-one maps" or "injections".

  For example:

  - The map $\mathbb{Z} \to \mathbb{Z},\ x \mapsto 2x$ (this is the map that sends each integer x to $2x$) is injective, because if x_1 and x_2 are two integers satisfying $2x_1 = 2x_2$, then $x_1 = x_2$.
  - The map $\mathbb{Z} \to \mathbb{Z},\ x \mapsto x^2$ (this is the map that sends each integer x to x^2) is **not** injective, because if x_1 and x_2 are two integers satisfying $x_1^2 = x_2^2$, then we do not necessarily have $x_1 = x_2$. (For example, if $x_1 = -1$ and $x_2 = 1$, then $x_1^2 = x_2^2$ but not $x_1 = x_2$.)

- A map $f : X \to Y$ between two sets X and Y is said to be *surjective* if it has the following property:

  - For each $y \in Y$, there exists some $x \in X$ satisfying $f(x) = y$. (In words: Each element of Y is an image of some element of X under f.)

  Surjective maps are often called "onto maps" or "surjections".

  For example:

  - The map $\mathbb{Z} \to \mathbb{Z},\ x \mapsto x + 1$ (this is the map that sends each integer x to $x + 1$) is surjective, because each integer y has some integer satisfying $x + 1 = y$ (namely, $x = y - 1$).
  - The map $\mathbb{Z} \to \mathbb{Z},\ x \mapsto 2x$ (this is the map that sends each integer x to $2x$) is **not** surjective, because not each integer y has some integer x satisfying $2x = y$. (For instance, $y = 1$ has no such x, since y is odd.)
  - The map $\{1, 2, 3, 4\} \to \{1, 2, 3, 4, 5\},\ x \mapsto x$ (this is the map sending each x to x) is **not** surjective, because not each $y \in \{1, 2, 3, 4, 5\}$ has some $x \in \{1, 2, 3, 4\}$ satisfying $x = y$. (Namely, $y = 5$ has no such x.)

[10] That said, mathematicians often show some nuance by using one of them and not the other. However, we do not need to concern ourselves with this here.

- A map $f : X \to Y$ between two sets X and Y is said to be *bijective* if it is both injective and surjective. Bijective maps are often called "one-to-one correspondences" or "bijections".

  For example:

  - The map $\mathbb{Z} \to \mathbb{Z},\ x \mapsto x + 1$ is bijective, since it is both injective and surjective.
  - The map $\{1, 2, 3, 4\} \to \{1, 2, 3, 4, 5\},\ x \mapsto x$ is **not** bijective, since it is not surjective. (However, it is injective.)
  - The map $\mathbb{Z} \to \mathbb{N},\ x \mapsto |x|$ is **not** bijective, since it is not injective. (However, it is surjective.)
  - The map $\mathbb{Z} \to \mathbb{Z},\ x \mapsto x^2$ is **not** bijective, since it is not injective. (It also is not surjective.)

- If X is a set, then id_X denotes the map from X to X that sends each $x \in X$ to x itself. (In words: id_X denotes the map which sends each element of X to itself.) The map id_X is often called the *identity map on* X, and often denoted by id (when X is clear from the context or irrelevant). The identity map id_X is always bijective.

- If $f : X \to Y$ and $g : Y \to Z$ are two maps, then the *composition* $g \circ f$ of the maps g and f is defined to be the map from X to Z that sends each $x \in X$ to $g(f(x))$. (In words: The composition $g \circ f$ is the map from X to Z that applies the map f **first** and **then** applies the map g.) You might find it confusing that this map is denoted by $g \circ f$ (rather than $f \circ g$), given that it proceeds by applying f first and g last; however, this has its reasons: It satisfies $(g \circ f)(x) = g(f(x))$. Had we denoted it by $f \circ g$ instead, this equality would instead become $(f \circ g)(x) = g(f(x))$, which would be even more confusing.

- If $f : X \to Y$ is a map between two sets X and Y, then an *inverse* of f means a map $g : Y \to X$ satisfying $f \circ g = \operatorname{id}_Y$ and $g \circ f = \operatorname{id}_X$. (In words, the condition "$f \circ g = \operatorname{id}_Y$" means "if you start with some element $y \in Y$, then apply g, then apply f, then you get y back", or equivalently "the map f undoes the map g". Similarly, the condition "$g \circ f = \operatorname{id}_X$" means "if you start with some element $x \in X$, then apply f, then apply g, then you get x back", or equivalently "the map g undoes the map f". Thus, an inverse of f means a map $g : Y \to X$ that both undoes and is undone by f.)

  The map $f : X \to Y$ is said to be *invertible* if and only if an inverse of f exists. If an inverse of f exists, then it is unique[11], and thus is called *the inverse of* f, and is denoted by f^{-1}.

[11] *Proof.* Let g_1 and g_2 be two inverses of f. We shall show that $g_1 = g_2$.

We know that g_1 is an inverse of f. In other words, g_1 is a map $Y \to X$ satisfying $f \circ g_1 = \operatorname{id}_Y$ and $g_1 \circ f = \operatorname{id}_X$.

For example:

- The map $\mathbb{Z} \to \mathbb{Z}$, $x \mapsto x+1$ is invertible, and its inverse is $\mathbb{Z} \to \mathbb{Z}$, $x \mapsto x-1$.
- The map $\mathbb{Q} \setminus \{1\} \to \mathbb{Q} \setminus \{0\}$, $x \mapsto \dfrac{1}{1-x}$ is invertible, and its inverse is the map $\mathbb{Q} \setminus \{0\} \to \mathbb{Q} \setminus \{1\}$, $x \mapsto 1 - \dfrac{1}{x}$.

- If $f : X \to Y$ is a map between two sets X and Y, then the following notations will be used:
  - For any subset U of X, we let $f(U)$ be the subset $\{f(u) \mid u \in U\}$ of Y. This set $f(U)$ is called the *image* of U under f. This should not be confused with the image $f(x)$ of a single element $x \in X$ under f.

    Note that the map $f : X \to Y$ is surjective if and only if $Y = f(X)$. (This is easily seen to be a restatement of the definition of "surjective".)
  - For any subset V of Y, we let $f^{-1}(V)$ be the subset $\{u \in X \mid f(u) \in V\}$ of X. This set $f^{-1}(V)$ is called the *preimage* of V under f. This should not be confused with the image $f^{-1}(y)$ of a single element $y \in Y$ under the inverse f^{-1} of f (when this inverse exists).

    (Note that in general, $f\left(f^{-1}(V)\right) \neq V$ and $f^{-1}(f(U)) \neq U$. However, $f\left(f^{-1}(V)\right) \subseteq V$ and $U \subseteq f^{-1}(f(U))$.)
  - For any subset U of X, we let $f\mid_U$ be the map from U to Y which sends each $u \in U$ to $f(u) \in Y$. This map $f\mid_U$ is called the *restriction* of f to the subset U.

The following facts are fundamental:

We know that g_2 is an inverse of f. In other words, g_2 is a map $Y \to X$ satisfying $f \circ g_2 = \operatorname{id}_Y$ and $g_2 \circ f = \operatorname{id}_X$.

A well-known fact (known as *associativity of map composition*, and stated explicitly as Proposition 2.82 below) says that if X, Y, Z and W are four sets, and if $c : X \to Y$, $b : Y \to Z$ and $a : Z \to W$ are three maps, then

$$(a \circ b) \circ c = a \circ (b \circ c).$$

Applying this fact to Y, X, Y, X, g_1, f and g_2 instead of X, Y, Z, W, c, b and a, we obtain $(g_2 \circ f) \circ g_1 = g_2 \circ (f \circ g_1)$.

Hence, $g_2 \circ (f \circ g_1) = \underbrace{(g_2 \circ f)}_{=\operatorname{id}_X} \circ g_1 = \operatorname{id}_X \circ g_1 = g_1$. Comparing this with $g_2 \circ \underbrace{(f \circ g_1)}_{=\operatorname{id}_Y} = g_2 \circ \operatorname{id}_Y = g_2$, we obtain $g_1 = g_2$.

Now, forget that we fixed g_1 and g_2. We thus have shown that if g_1 and g_2 are two inverses of f, then $g_1 = g_2$. In other words, any two inverses of f must be equal. In other words, if an inverse of f exists, then it is unique.

Theorem 1.1. A map $f : X \to Y$ is invertible if and only if it is bijective.

Theorem 1.2. Let U and V be two finite sets. Then, $|U| = |V|$ if and only if there exists a bijective map $f : U \to V$.

Theorem 1.2 holds even if the sets U and V are infinite, but to make sense of this we would need to define the size of an infinite set, which is a much subtler issue than the size of a finite set. We will only need Theorem 1.2 for finite sets.

Let us state some more well-known and basic properties of maps between finite sets:

Lemma 1.3. Let U and V be two finite sets. Let $f : U \to V$ be a map.

(a) We have $|f(S)| \leq |S|$ for each subset S of U.

(b) If $|f(U)| \geq |U|$, then the map f is injective.

(c) If f is injective, then $|f(S)| = |S|$ for each subset S of U.

Lemma 1.4. Let U and V be two finite sets such that $|U| \leq |V|$. Let $f : U \to V$ be a map. Then, we have the following logical equivalence:

$$(f \text{ is surjective}) \iff (f \text{ is bijective}).$$

Lemma 1.5. Let U and V be two finite sets such that $|U| \geq |V|$. Let $f : U \to V$ be a map. Then, we have the following logical equivalence:

$$(f \text{ is injective}) \iff (f \text{ is bijective}).$$

Exercise 1.1. Prove Lemma 1.3, Lemma 1.4 and Lemma 1.5.

Let us make one additional observation about maps:

Remark 1.6. Composition of maps is associative: If X, Y, Z and W are three sets, and if $c : X \to Y$, $b : Y \to Z$ and $a : Z \to W$ are three maps, then $(a \circ b) \circ c = a \circ (b \circ c)$. (This shall be proven in Proposition 2.82 below.)

In Section 2.13, we shall prove a more general fact: If $X_1, X_2, \ldots, X_{k+1}$ are $k+1$ sets for some $k \in \mathbb{N}$, and if $f_i : X_i \to X_{i+1}$ is a map for each $i \in \{1, 2, \ldots, k\}$, then the composition $f_k \circ f_{k-1} \circ \cdots \circ f_1$ of all k maps $f_1, f_2, \ldots, f_k$ is a well-defined map from X_1 to X_{k+1}, which sends each element $x \in X_1$ to $f_k(f_{k-1}(f_{k-2}(\cdots(f_2(f_1(x)))\cdots)))$ (in other words, which transforms each element $x \in X_1$ by first applying f_1, then applying f_2, then applying f_3, and so on); this composition $f_k \circ f_{k-1} \circ \cdots \circ f_1$ can also be written as $f_k \circ (f_{k-1} \circ (f_{k-2} \circ (\cdots \circ (f_2 \circ f_1) \cdots)))$ or as $(((\cdots (f_k \circ f_{k-1}) \circ \cdots) \circ f_3) \circ f_2) \circ f_1$. An important particular case is when $k = 0$; in this case, $f_k \circ f_{k-1} \circ \cdots \circ f_1$ is a composition of 0 maps. It is defined to be id_{X_1} (the identity map of the set X_1), and it is called the "empty composition of maps $X_1 \to X_1$". (The logic behind this definition is that the composition $f_k \circ f_{k-1} \circ \cdots \circ f_1$ should transform each element $x \in X_1$ by first applying f_1, then applying f_2, then applying f_3, and so on; however, for $k = 0$, there are no maps to apply, and so x just remains unchanged.)

1.4. Sums and products: a synopsis

In this section, I will recall the definitions of the $\sum$ and $\prod$ signs and collect some of their basic properties (without proofs). When I say "recall", I am implying that the reader has at least some prior acquaintance (and, ideally, experience) with these signs; for a first introduction, this section is probably too brief and too abstract. Ideally, you should use this section to familiarize yourself with my (sometimes idiosyncratic) notations.

Throughout Section 1.4, we let $\mathbb{A}$ be one of the sets $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$ and $\mathbb{C}$.

1.4.1. Definition of $\sum$

Let us first define the $\sum$ sign. There are actually several (slightly different, but still closely related) notations involving the $\sum$ sign; let us define the most important of them:

- If S is a finite set, and if a_s is an element of $\mathbb{A}$ for each $s \in S$, then $\sum\limits_{s \in S} a_s$ denotes the sum of all of these elements a_s. Formally, this sum is defined by recursion on $|S|$, as follows:
  - If $|S| = 0$, then $\sum\limits_{s \in S} a_s$ is defined to be 0.
  - Let $n \in \mathbb{N}$. Assume that we have defined $\sum\limits_{s \in S} a_s$ for every finite set S with $|S| = n$ (and every choice of elements a_s of $\mathbb{A}$). Now, if S is a finite set with $|S| = n + 1$ (and if $a_s \in \mathbb{A}$ are chosen for all $s \in S$), then $\sum\limits_{s \in S} a_s$ is defined by picking any $t \in S$ [12] and setting

    $$\sum_{s \in S} a_s = a_t + \sum_{s \in S \setminus \{t\}} a_s. \tag{1}$$

    It is not immediately clear why this definition is legitimate: The right hand side of (1) is defined using a choice of t, but we want our value of $\sum\limits_{s \in S} a_s$ to depend only on S and on the a_s (not on some arbitrarily chosen $t \in S$). However, it is possible to prove that the right hand side of (1) is actually independent of t (that is, any two choices of t will lead to the

[12] This is possible, because S is nonempty (in fact, $|S| = n + 1 > n \geq 0$).

same result). See Section 2.14 below (and Theorem 2.118 **(a)** in particular) for the proof of this fact.

Examples:

- If $S = \{4, 7, 9\}$ and $a_s = \dfrac{1}{s^2}$ for every $s \in S$, then $\sum\limits_{s \in S} a_s = a_4 + a_7 + a_9 = \dfrac{1}{4^2} + \dfrac{1}{7^2} + \dfrac{1}{9^2} = \dfrac{6049}{63504}$.
- If $S = \{1, 2, \ldots, n\}$ (for some $n \in \mathbb{N}$) and $a_s = s^2$ for every $s \in S$, then $\sum\limits_{s \in S} a_s = \sum\limits_{s \in S} s^2 = 1^2 + 2^2 + \cdots + n^2$. (There is a formula saying that the right hand side of this equality is $\dfrac{1}{6} n \left(2n+1\right) \left(n+1\right)$.)
- If $S = \varnothing$, then $\sum\limits_{s \in S} a_s = 0$ (since $|S| = 0$).

Remarks:

- The sum $\sum\limits_{s \in S} a_s$ is usually pronounced "sum of the a_s over all $s \in S$" or "sum of the a_s with s ranging over S" or "sum of the a_s with s running through all elements of S". The letter "s" in the sum is called the "summation index"[13], and its exact choice is immaterial (for example, you can rewrite $\sum\limits_{s \in S} a_s$ as $\sum\limits_{t \in S} a_t$ or as $\sum\limits_{\Phi \in S} a_\Phi$ or as $\sum\limits_{\spadesuit \in S} a_\spadesuit$), as long as it does not already have a different meaning outside of the sum[14]. (Ultimately, a summation index is the same kind of placeholder variable as the "s" in the statement "for all $s \in S$, we have $a_s + 2a_s = 3a_s$", or as a loop variable in a for-loop in programming.) The sign $\sum$ itself is called "the summation sign" or "the $\sum$ sign". The numbers a_s are called the *addends* (or *summands*) of the sum $\sum\limits_{s \in S} a_s$. More precisely, for any given $t \in S$, we can refer to the number a_t as the "addend corresponding to the index t" (or as the "addend for $s = t$", or as the "addend for t") of the sum $\sum\limits_{s \in S} a_s$.
- When the set S is empty, the sum $\sum\limits_{s \in S} a_s$ is called an *empty sum*. Our definition implies that any empty sum is 0. This convention is used throughout mathematics, except in rare occasions where a slightly subtler version of it is used[15]. Ignore anyone who tells you that empty sums are undefined!

[13] The plural of the word "index" here is "indices", not "indexes".

[14] If it already has a different meaning, then it must not be used as a summation index! For example, you must not write "every $n \in \mathbb{N}$ satisfies $\sum\limits_{n \in \{0,1,\ldots,n\}} n = \dfrac{n\left(n+1\right)}{2}$", because here the summation index n clashes with a different meaning of the letter n.

[15] Do not worry about this subtler version for the time being. If you really want to know what it is: Our above definition is tailored to the cases when the a_s are numbers (i.e., elements of one

– The summation index does not always have to be a single letter. For instance, if S is a set of pairs, then we can write $\sum\limits_{(x,y)\in S} a_{(x,y)}$ (meaning the same as $\sum\limits_{s\in S} a_s$). Here is an example of this notation:

$$\sum_{(x,y)\in\{1,2,3\}^2} \frac{x}{y} = \frac{1}{1}+\frac{1}{2}+\frac{1}{3}+\frac{2}{1}+\frac{2}{2}+\frac{2}{3}+\frac{3}{1}+\frac{3}{2}+\frac{3}{3}$$

(here, we are using the notation $\sum\limits_{(x,y)\in S} a_{(x,y)}$ with $S = \{1,2,3\}^2$ and $a_{(x,y)} = \frac{x}{y}$). Note that we could not have rewritten this sum in the form $\sum\limits_{s\in S} a_s$ with a single-letter variable s without introducing an extra notation such as $a_{(x,y)}$ for the quotients $\frac{x}{y}$.

– Mathematicians don't seem to have reached an agreement on the operator precedence of the $\sum$ sign. By this I mean the following question: Does $\sum\limits_{s\in S} a_s + b$ (where b is some other element of $\mathbb{A}$) mean $\sum\limits_{s\in S} (a_s + b)$ or $\left(\sum\limits_{s\in S} a_s\right) + b$? In my experience, the second interpretation (i.e., reading it as $\left(\sum\limits_{s\in S} a_s\right) + b$) is more widespread, and this is the interpretation that I will follow. Nevertheless, be on the watch for possible misunderstandings, as someone might be using the first interpretation when you expect it the least![16]

However, the situation is different for products and nested sums. For instance, the expression $\sum\limits_{s\in S} ba_sc$ is understood to mean $\sum\limits_{s\in S} (ba_sc)$, and a nested sum like $\sum\limits_{s\in S}\sum\limits_{t\in T} a_{s,t}$ (where S and T are two sets, and where $a_{s,t}$ is an element of $\mathbb{A}$ for each pair $(s,t) \in S \times T$) is to be read as $\sum\limits_{s\in S}\left(\sum\limits_{t\in T} a_{s,t}\right)$.

of the sets $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$ and $\mathbb{C}$). In more advanced settings, one tends to take sums of the form $\sum\limits_{s\in S} a_s$ where the a_s are not numbers but (for example) elements of a commutative ring $\mathbb{K}$. (See Definition 6.2 for the definition of a commutative ring.) In such cases, one wants the sum $\sum\limits_{s\in S} a_s$ for an empty set S to be not the integer 0, but the zero of the commutative ring $\mathbb{K}$ (which is sometimes distinct from the integer 0). This has the slightly confusing consequence that the meaning of the sum $\sum\limits_{s\in S} a_s$ for an empty set S depends on what ring $\mathbb{K}$ the a_s belong to, even if (for an empty set S) there are no a_s to begin with! But in practice, the choice of $\mathbb{K}$ is always clear from context, so this is not ambiguous.

A similar caveat applies to the other versions of the $\sum$ sign, as well as to the $\prod$ sign defined further below; I shall not elaborate on it further.

[16] This is similar to the notorious disagreement about whether a/bc means $(a/b)\cdot c$ or $a/(bc)$.

- Speaking of nested sums: they mean exactly what they seem to mean. For instance, $\sum_{s\in S}\sum_{t\in T} a_{s,t}$ is what you get if you compute the sum $\sum_{t\in T} a_{s,t}$ for each $s \in S$, and then sum up all of these sums together. In a nested sum $\sum_{s\in S}\sum_{t\in T} a_{s,t}$, the first summation sign ($\sum_{s\in S}$) is called the "outer summation", and the second summation sign ($\sum_{t\in T}$) is called the "inner summation".
- An expression of the form " $\sum_{s\in S} a_s$" (where S is a finite set) is called a *finite sum*.
- We have required the set S to be finite when defining $\sum_{s\in S} a_s$. Of course, this requirement was necessary for our definition, and there is no way to make sense of infinite sums such as $\sum_{s\in\mathbb{Z}} s^2$. However, **some** infinite sums can be made sense of. The simplest case is when the set S might be infinite, but only finitely many among the a_s are nonzero. In this case, we can define $\sum_{s\in S} a_s$ simply by discarding the zero addends and summing the finitely many remaining addends. Other situations in which infinite sums make sense appear in analysis and in topological algebra (e.g., power series).
- The sum $\sum_{s\in S} a_s$ always belongs to $\mathbb{A}$. [17] For instance, a sum of elements of $\mathbb{N}$ belongs to $\mathbb{N}$; a sum of elements of $\mathbb{R}$ belongs to $\mathbb{R}$, and so on.

- A slightly more complicated version of the summation sign is the following: Let S be a finite set, and let $\mathcal{A}(s)$ be a logical statement defined for every $s \in S$ [18]. For example, S can be $\{1,2,3,4\}$, and $\mathcal{A}(s)$ can be the statement "s is even". For each $s \in S$ satisfying $\mathcal{A}(s)$, let a_s be an element of $\mathbb{A}$. Then, the sum $\sum_{\substack{s\in S;\\ \mathcal{A}(s)}} a_s$ is defined by

$$\sum_{\substack{s\in S;\\ \mathcal{A}(s)}} a_s = \sum_{s\in\{t\in S \mid \mathcal{A}(t)\}} a_s.$$

In other words, $\sum_{\substack{s\in S;\\ \mathcal{A}(s)}} a_s$ is the sum of the a_s for all $s \in S$ which satisfy $\mathcal{A}(s)$.

Examples:

[17] Recall that we have assumed $\mathbb{A}$ to be one of the sets $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$ and $\mathbb{C}$, and that we have assumed the a_s to belong to $\mathbb{A}$.

[18] Formally speaking, this means that $\mathcal{A}$ is a map from S to the set of all logical statements. Such a map is called a *predicate*.

  - If $S = \{1,2,3,4,5\}$, then $\sum\limits_{\substack{s\in S;\\ s \text{ is even}}} a_s = a_2 + a_4$. (Of course, $\sum\limits_{\substack{s\in S;\\ s \text{ is even}}} a_s$ is $\sum\limits_{\substack{s\in S;\\ \mathcal{A}(s)}} a_s$ when $\mathcal{A}(s)$ is defined to be the statement "s is even".)
  - If $S = \{1,2,\ldots,n\}$ (for some $n \in \mathbb{N}$) and $a_s = s^2$ for every $s \in S$, then $\sum\limits_{\substack{s\in S;\\ s \text{ is even}}} a_s = a_2 + a_4 + \cdots + a_k$, where k is the largest even number among $1,2,\ldots,n$ (that is, $k = n$ if n is even, and $k = n-1$ otherwise).

Remarks:

  - The sum $\sum\limits_{\substack{s\in S;\\ \mathcal{A}(s)}} a_s$ is usually pronounced "sum of the a_s over all $s \in S$ satisfying $\mathcal{A}(s)$". The semicolon after "$s \in S$" is often omitted or replaced by a colon or a comma. Many authors often omit the "$s \in S$" part (so they simply write $\sum\limits_{\mathcal{A}(s)} a_s$) when it is clear enough what the S is. (For instance, they would write $\sum\limits_{1\leq s\leq 5} s^2$ instead of $\sum\limits_{\substack{s\in\mathbb{N};\\ 1\leq s\leq 5}} s^2$.)
  - The set S needs not be finite in order for $\sum\limits_{\substack{s\in S;\\ \mathcal{A}(s)}} a_s$ to be defined; it suffices that the set $\{t \in S \mid \mathcal{A}(t)\}$ be finite (i.e., that only finitely many $s \in S$ satisfy $\mathcal{A}(s)$).
  - The sum $\sum\limits_{\substack{s\in S;\\ \mathcal{A}(s)}} a_s$ is said to be *empty* whenever the set $\{t \in S \mid \mathcal{A}(t)\}$ is empty (i.e., whenever no $s \in S$ satisfies $\mathcal{A}(s)$).

- Finally, here is the simplest version of the summation sign: Let u and v be two integers. We agree to understand the set $\{u, u+1, \ldots, v\}$ to be empty when $u > v$. Let a_s be an element of $\mathbb{A}$ for each $s \in \{u, u+1, \ldots, v\}$. Then, $\sum\limits_{s=u}^{v} a_s$ is defined by

$$\sum_{s=u}^{v} a_s = \sum_{s\in\{u,u+1,\ldots,v\}} a_s.$$

Examples:

  - We have $\sum\limits_{s=3}^{8} \dfrac{1}{s} = \sum\limits_{s\in\{3,4,\ldots,8\}} \dfrac{1}{s} = \dfrac{1}{3} + \dfrac{1}{4} + \dfrac{1}{5} + \dfrac{1}{6} + \dfrac{1}{7} + \dfrac{1}{8} = \dfrac{341}{280}$.
  - We have $\sum\limits_{s=3}^{3} \dfrac{1}{s} = \sum\limits_{s\in\{3\}} \dfrac{1}{s} = \dfrac{1}{3}$.

- We have $\sum_{s=3}^{2} \frac{1}{s} = \sum_{s \in \varnothing} \frac{1}{s} = 0$.

Remarks:

- The sum $\sum_{s=u}^{v} a_s$ is usually pronounced "sum of the a_s for all s from u to v (inclusive)". It is often written $a_u + a_{u+1} + \cdots + a_v$, but this latter notation has its drawbacks: In order to understand an expression like $a_u + a_{u+1} + \cdots + a_v$, one needs to correctly guess the pattern (which can be unintuitive when the a_s themselves are complicated: for example, it takes a while to find the "moving parts" in the expression $\frac{2 \cdot 7}{3+2} + \frac{3 \cdot 7}{3+3} + \cdots + \frac{7 \cdot 7}{3+7}$, whereas the notation $\sum_{s=2}^{7} \frac{s \cdot 7}{3+s}$ for the same sum is perfectly clear).

- In the sum $\sum_{s=u}^{v} a_s$, the integer u is called the *lower limit* (of the sum), whereas the integer v is called the *upper limit* (of the sum). The sum is said to *start* (or *begin*) at u and *end* at v.

- The sum $\sum_{s=u}^{v} a_s$ is said to be *empty* whenever $u > v$. In other words, a sum of the form $\sum_{s=u}^{v} a_s$ is empty whenever it "ends before it has begun". However, a sum which "ends right after it begins" (i.e., a sum $\sum_{s=u}^{v} a_s$ with $u = v$) is not empty; it just has one addend only. (This is unlike integrals, which are 0 whenever their lower and upper limit are equal.)

- Let me stress once again that a sum $\sum_{s=u}^{v} a_s$ with $u > v$ is empty and equals 0. It does not matter how much greater u is than v. So, for example, $\sum_{s=1}^{-5} s = 0$. The fact that the upper bound (-5) is much smaller than the lower bound (1) does not mean that you have to subtract rather than add.

Thus we have introduced the main three forms of the summation sign. Some mild variations on them appear in the literature (e.g., there is a slightly awkward notation $\sum_{\substack{s=u;\\ \mathcal{A}(s)}}^{v} a_s$ for $\sum_{\substack{s \in \{u,u+1,\ldots,v\};\\ \mathcal{A}(s)}} a_s$).

1.4.2. Properties of $\sum$

Let me now show some basic properties of summation signs that are important in making them useful:

- **Splitting-off:** Let S be a finite set. Let $t \in S$. Let a_s be an element of $\mathbb{A}$ for each $s \in S$. Then,
$$\sum_{s\in S} a_s = a_t + \sum_{s\in S\setminus\{t\}} a_s. \qquad (2)$$
(This is precisely the equality (1) (applied to $n = |S \setminus \{t\}|$), because $|S| = |S \setminus \{t\}| + 1$.) This formula (2) allows us to "split off" an addend from a sum.

  Example: If $n \in \mathbb{N}$, then
$$\sum_{s\in\{1,2,\ldots,n+1\}} a_s = a_{n+1} + \sum_{s\in\{1,2,\ldots,n\}} a_s$$
(by (2), applied to $S = \{1,2,\ldots,n+1\}$ and $t = n+1$), but also
$$\sum_{s\in\{1,2,\ldots,n+1\}} a_s = a_1 + \sum_{s\in\{2,3,\ldots,n+1\}} a_s$$
(by (2), applied to $S = \{1,2,\ldots,n+1\}$ and $t = 1$).

- **Splitting:** Let S be a finite set. Let X and Y be two subsets of S such that $X \cap Y = \varnothing$ and $X \cup Y = S$. (Equivalently, X and Y are two subsets of S such that each element of S lies in **exactly** one of X and Y.) Let a_s be an element of $\mathbb{A}$ for each $s \in S$. Then,
$$\sum_{s\in S} a_s = \sum_{s\in X} a_s + \sum_{s\in Y} a_s. \qquad (3)$$
(Here, as we explained, $\sum_{s\in X} a_s + \sum_{s\in Y} a_s$ stands for $\left(\sum_{s\in X} a_s\right) + \left(\sum_{s\in Y} a_s\right)$.) The idea behind (3) is that if we want to add a bunch of numbers (the a_s for $s \in S$), we can proceed by splitting it into two "sub-bunches" (one "sub-bunch" consisting of the a_s for $s \in X$, and the other consisting of the a_s for $s \in Y$), then take the sum of each of these two sub-bunches, and finally add together the two sums. For a rigorous proof of (3), see Theorem 2.130 below.

  Examples:

  - If $n \in \mathbb{N}$, then
$$\sum_{s\in\{1,2,\ldots,2n\}} a_s = \sum_{s\in\{1,3,\ldots,2n-1\}} a_s + \sum_{s\in\{2,4,\ldots,2n\}} a_s$$
(by (3), applied to $S = \{1,2,\ldots,2n\}$, $X = \{1,3,\ldots,2n-1\}$ and $Y = \{2,4,\ldots,2n\}$).
  - If $n \in \mathbb{N}$ and $m \in \mathbb{N}$, then
$$\sum_{s\in\{-m,-m+1,\ldots,n\}} a_s = \sum_{s\in\{-m,-m+1,\ldots,0\}} a_s + \sum_{s\in\{1,2,\ldots,n\}} a_s$$
(by (3), applied to $S = \{-m,-m+1,\ldots,n\}$, $X = \{-m,-m+1,\ldots,0\}$ and $Y = \{1,2,\ldots,n\}$).

  - If u, v and w are three integers such that $u-1 \leq v \leq w$, and if a_s is an element of $\mathbb{A}$ for each $s \in \{u, u+1, \ldots, w\}$, then

$$\sum_{s=u}^{w} a_s = \sum_{s=u}^{v} a_s + \sum_{s=v+1}^{w} a_s. \tag{4}$$

    This follows from (3), applied to $S = \{u, u+1, \ldots, w\}$, $X = \{u, u+1, \ldots, v\}$ and $Y = \{v+1, v+2, \ldots, w\}$. Notice that the requirement $u-1 \leq v \leq w$ is important; otherwise, the $X \cap Y = \varnothing$ and $X \cup Y = S$ condition would not hold!

- **Splitting using a predicate:** Let S be a finite set. Let $\mathcal{A}(s)$ be a logical statement for each $s \in S$. Let a_s be an element of $\mathbb{A}$ for each $s \in S$. Then,

$$\sum_{s\in S} a_s = \sum_{\substack{s\in S;\\ \mathcal{A}(s)}} a_s + \sum_{\substack{s\in S;\\ \text{not } \mathcal{A}(s)}} a_s \tag{5}$$

  (where "not $\mathcal{A}(s)$" means the negation of $\mathcal{A}(s)$). This simply follows from (3), applied to $X = \{s \in S \mid \mathcal{A}(s)\}$ and $Y = \{s \in S \mid \text{not } \mathcal{A}(s)\}$.

  Example: If $S \subseteq \mathbb{Z}$, then

$$\sum_{s\in S} a_s = \sum_{\substack{s\in S;\\ s \text{ is even}}} a_s + \sum_{\substack{s\in S;\\ s \text{ is odd}}} a_s$$

  (because "s is odd" is the negation of "s is even").

- **Summing equal values:** Let S be a finite set. Let a be an element of $\mathbb{A}$. Then,

$$\sum_{s\in S} a = |S| \cdot a. \tag{6}$$

  [19] In other words, if all addends of a sum are equal to one and the same element a, then the sum is just the number of its addends times a. In particular,

$$\sum_{s\in S} 1 = |S| \cdot 1 = |S|.$$

- **Splitting an addend:** Let S be a finite set. For every $s \in S$, let a_s and b_s be elements of $\mathbb{A}$. Then,

$$\sum_{s\in S} (a_s + b_s) = \sum_{s\in S} a_s + \sum_{s\in S} b_s. \tag{7}$$

  For a rigorous proof of this equality, see Theorem 2.122 below.

[19] This is easy to prove by induction on $|S|$.

Remark: Of course, similar rules hold for other forms of summations: If $\mathcal{A}(s)$ is a logical statement for each $s \in S$, then

$$\sum_{\substack{s\in S;\\ \mathcal{A}(s)}} (a_s + b_s) = \sum_{\substack{s\in S;\\ \mathcal{A}(s)}} a_s + \sum_{\substack{s\in S;\\ \mathcal{A}(s)}} b_s.$$

If u and v are two integers, then

$$\sum_{s=u}^{v} (a_s + b_s) = \sum_{s=u}^{v} a_s + \sum_{s=u}^{v} b_s. \tag{8}$$

- **Factoring out:** Let S be a finite set. For every $s \in S$, let a_s be an element of $\mathbb{A}$. Also, let λ be an element of $\mathbb{A}$. Then,

  $$\sum_{s\in S} \lambda a_s = \lambda \sum_{s\in S} a_s. \tag{9}$$

  For a rigorous proof of this equality, see Theorem 2.124 below.

  Again, similar rules hold for the other types of summation sign.

  Remark: Applying (9) to $\lambda = -1$, we obtain

  $$\sum_{s\in S} (-a_s) = -\sum_{s\in S} a_s.$$

- **Zeroes sum to zero:** Let S be a finite set. Then,

  $$\sum_{s\in S} 0 = 0. \tag{10}$$

  That is, any sum of zeroes is zero.

  For a rigorous proof of this equality, see Theorem 2.126 below.

  Remark: This applies even to infinite sums! Do not be fooled by the infiniteness of a sum: There are no reasonable situations where an infinite sum of zeroes is defined to be anything other than zero. The infinity does not "compensate" for the zero.

- **Dropping zeroes:** Let S be a finite set. Let a_s be an element of $\mathbb{A}$ for each $s \in S$. Let T be a subset of S such that every $s \in T$ satisfies $a_s = 0$. Then,

  $$\sum_{s\in S} a_s = \sum_{s\in S\setminus T} a_s. \tag{11}$$

  (That is, any addends which are zero can be removed from a sum without changing the sum's value.) See Corollary 2.131 below for a proof of (11).

- **Renaming the index:** Let S be a finite set. Let a_s be an element of $\mathbb{A}$ for each $s \in S$. Then,
$$\sum_{s\in S} a_s = \sum_{t\in S} a_t.$$
This is just saying that the summation index in a sum can be renamed at will, as long as its name does not clash with other notation.

- **Substituting the index I:** Let S and T be two finite sets. Let $f : S \to T$ be a **bijective** map. Let a_t be an element of $\mathbb{A}$ for each $t \in T$. Then,
$$\sum_{t\in T} a_t = \sum_{s\in S} a_{f(s)}. \qquad (12)$$
(The idea here is that the sum $\sum_{s\in S} a_{f(s)}$ contains the same addends as the sum $\sum_{t\in T} a_t$.) A rigorous proof of (12) can be found in Theorem 2.132 below.

  Examples:

  - For any $n \in \mathbb{N}$, we have
$$\sum_{t\in\{1,2,\ldots,n\}} t^3 = \sum_{s\in\{-n,-n+1,\ldots,-1\}} (-s)^3 .$$
(This follows from (12), applied to $S = \{-n, -n+1, \ldots, -1\}$, $T = \{1, 2, \ldots, n\}$, $f(s) = -s$, and $a_t = t^3$.)
  - The sets S and T in (12) may well be the same. For example, for any $n \in \mathbb{N}$, we have
$$\sum_{t\in\{1,2,\ldots,n\}} t^3 = \sum_{s\in\{1,2,\ldots,n\}} (n+1-s)^3 .$$
(This follows from (12), applied to $S = \{1, 2, \ldots, n\}$, $T = \{1, 2, \ldots, n\}$, $f(s) = n+1-s$ and $a_t = t^3$.)
  - More generally: Let u and v be two integers. Then, the map $\{u, u+1, \ldots, v\} \to \{u, u+1, \ldots, v\}$ sending each $s \in \{u, u+1, \ldots, v\}$ to $u+v-s$ is a bijection[20]. Hence, we can substitute $u+v-s$ for s in the sum $\sum_{s=u}^{v} a_s$ whenever an element a_s of $\mathbb{A}$ is given for each $s \in \{u, u+1, \ldots, v\}$. We thus obtain the formula
$$\sum_{s=u}^{v} a_s = \sum_{s=u}^{v} a_{u+v-s}.$$

  Remark:

[20] Check this!

- When I use (12) to rewrite the sum $\sum\limits_{t\in T} a_t$ as $\sum\limits_{s\in S} a_{f(s)}$, I say that I have "substituted $f(s)$ for t in the sum". Conversely, when I use (12) to rewrite the sum $\sum\limits_{s\in S} a_{f(s)}$ as $\sum\limits_{t\in T} a_t$, I say that I have "substituted t for $f(s)$ in the sum".
- For convenience, I have chosen s and t as summation indices in (12). But as before, they can be chosen to be any letters not otherwise used. It is perfectly okay to use one and the same letter for both of them, e.g., to write $\sum\limits_{s\in T} a_s = \sum\limits_{s\in S} a_{f(s)}$.
- Here is the probably most famous example of substitution in a sum: Fix a nonnegative integer n. Then, we can substitute $n-i$ for i in the sum $\sum\limits_{i=0}^{n} i$ (since the map $\{0,1,\ldots,n\} \to \{0,1,\ldots,n\}$, $i \mapsto n-i$ is a bijection). Thus, we obtain

$$\sum_{i=0}^{n} i = \sum_{i=0}^{n} (n-i).$$

Now,

$$\begin{aligned} 2\sum_{i=0}^{n} i &= \sum_{i=0}^{n} i + \underbrace{\sum_{i=0}^{n} i}_{=\sum\limits_{i=0}^{n}(n-i)} && (\text{since } 2q = q+q \text{ for every } q \in \mathbb{Q}) \\ &= \sum_{i=0}^{n} i + \sum_{i=0}^{n} (n-i) \\ &= \sum_{i=0}^{n} \underbrace{(i+(n-i))}_{=n} && (\text{here, we have used (8) backwards}) \\ &= \sum_{i=0}^{n} n = (n+1)\,n && (\text{by (6)}) \\ &= n(n+1), \end{aligned}$$

and therefore

$$\sum_{i=0}^{n} i = \frac{n(n+1)}{2}. \tag{13}$$

Since $\sum\limits_{i=0}^{n} i = 0 + \sum\limits_{i=1}^{n} i = \sum\limits_{i=1}^{n} i$, this rewrites as

$$\sum_{i=1}^{n} i = \frac{n(n+1)}{2}. \tag{14}$$

This is the famous "Little Gauss formula" (supposedly discovered by Carl Friedrich Gauss in primary school, but already known to the Pythagoreans).

- **Substituting the index II:** Let S and T be two finite sets. Let $f : S \to T$ be a **bijective** map. Let a_s be an element of $\mathbb{A}$ for each $s \in S$. Then,

$$\sum_{s\in S} a_s = \sum_{t\in T} a_{f^{-1}(t)}. \tag{15}$$

  This is, of course, just (12) but applied to T, S and f^{-1} instead of S, T and f. (Nevertheless, I prefer to mention (15) separately because it often is used in this very form.)

- **Telescoping sums:** Let u and v be two integers such that $u-1 \leq v$. Let a_s be an element of $\mathbb{A}$ for each $s \in \{u-1, u, \ldots, v\}$. Then,

$$\sum_{s=u}^{v} (a_s - a_{s-1}) = a_v - a_{u-1}. \tag{16}$$

  Examples:

  - Let us give a new proof of (14). Indeed, fix a nonnegative integer n. An easy computation reveals that

$$s = \frac{s(s+1)}{2} - \frac{(s-1)((s-1)+1)}{2} \tag{17}$$

    for each $s \in \mathbb{Z}$. Thus,

$$\begin{aligned}\sum_{i=1}^{n} i &= \sum_{s=1}^{n} s = \sum_{s=1}^{n} \left(\frac{s(s+1)}{2} - \frac{(s-1)((s-1)+1)}{2} \right) \qquad \text{(by (17))} \\ &= \frac{n(n+1)}{2} - \underbrace{\frac{(1-1)((1-1)+1)}{2}}_{=0} \\ &\qquad \left(\text{by (16), applied to } u=1,\ v=n \text{ and } a_s = \frac{s(s+1)}{2} \right) \\ &= \frac{n(n+1)}{2}.\end{aligned}$$

    Thus, (14) is proven again. This kind of proof works often when we need to prove a formula like (14); the only tricky part was to "guess" the right value of a_s, which is straightforward if you know what you are looking for (you want $a_n - a_0$ to be $\frac{n(n+1)}{2}$), but rather tricky if you don't.

  - Another application of (16) is a proof of the well-known formula

$$\sum_{i=1}^{n} i^2 = \frac{n(n+1)(2n+1)}{6} \qquad \text{for all } n \in \mathbb{N}.$$

Indeed, an easy computation reveals that

$$s^2 = \frac{s(s+1)(2s+1)}{6} - \frac{(s-1)((s-1)+1)(2(s-1)+1)}{6}$$

for each $s \in \mathbb{Z}$; now, as in the previous example, we can sum this equality over all $s \in \{1, 2, \ldots, n\}$ and apply (16) to obtain our claim.

– Here is another important identity that follows from (16): If a and b are any elements of $\mathbb{A}$, and if $m \in \mathbb{N}$, then

$$(a-b) \sum_{i=0}^{m-1} a^i b^{m-1-i} = a^m - b^m. \tag{18}$$

(This is one of the versions of the "geometric series formula".) To prove (18), we observe that

$$\begin{aligned}
&(a-b) \sum_{i=0}^{m-1} a^i b^{m-1-i} \\
&= \sum_{i=0}^{m-1} \underbrace{(a-b)\, a^i b^{m-1-i}}_{=aa^i b^{m-1-i} - ba^i b^{m-1-i}} \qquad \text{(this follows from (9))} \\
&= \sum_{i=0}^{m-1} \left(aa^i b^{m-1-i} - \underbrace{ba^i}_{=a^i b} b^{m-1-i} \right) \\
&= \sum_{i=0}^{m-1} \left(\underbrace{aa^i}_{=a^{i+1}} b^{m-1-i} - \underbrace{a^i}_{\substack{=a^{(i-1)+1} \\ \text{(since } i=(i-1)+1)}} \underbrace{bb^{m-1-i}}_{\substack{=b^{(m-1-i)+1}=b^{m-1-(i-1)} \\ \text{(since } (m-1-i)+1=m-1-(i-1))}} \right) \\
&= \sum_{i=0}^{m-1} \left(a^{i+1} b^{m-1-i} - a^{(i-1)+1} b^{m-1-(i-1)} \right) \\
&= \sum_{s=0}^{m-1} \left(a^{s+1} b^{m-1-s} - a^{(s-1)+1} b^{m-1-(s-1)} \right) \\
&\qquad \text{(here, we have renamed the summation index } i \text{ as } s) \\
&= \underbrace{a^{(m-1)+1}}_{\substack{=a^m \\ \text{(since } (m-1)+1=m)}} \underbrace{b^{m-1-(m-1)}}_{\substack{=b^0 \\ \text{(since } m-1-(m-1)=0)}} - \underbrace{a^{(0-1)+1}}_{\substack{=a^0 \\ \text{(since } (0-1)+1=0)}} \underbrace{b^{m-1-(0-1)}}_{\substack{=b^m \\ \text{(since } m-1-(0-1)=m)}} \\
&\qquad \left(\text{by (16) (applied to } u=0,\ v=m-1 \text{ and } a_s = a^{s+1} b^{m-1-s}) \right) \\
&= a^m \underbrace{b^0}_{=1} - \underbrace{a^0}_{=1} b^m = a^m - b^m.
\end{aligned}$$

– Other examples for the use of (16) can be found on the Wikipedia page for "telescoping series". Let me add just one more example: Given $n \in$

$\mathbb{N}$, we want to compute $\sum\limits_{i=1}^{n} \dfrac{1}{\sqrt{i}+\sqrt{i+1}}$. (Here, of course, we need to take $\mathbb{A} = \mathbb{R}$ or $\mathbb{A} = \mathbb{C}$.) We proceed as follows: For every positive integer i, we have

$$\frac{1}{\sqrt{i}+\sqrt{i+1}} = \frac{\left(\sqrt{i+1}-\sqrt{i}\right)}{\left(\sqrt{i}+\sqrt{i+1}\right)\left(\sqrt{i+1}-\sqrt{i}\right)} = \sqrt{i+1}-\sqrt{i}$$

(since $\left(\sqrt{i}+\sqrt{i+1}\right)\left(\sqrt{i+1}-\sqrt{i}\right) = \left(\sqrt{i+1}\right)^2 - \left(\sqrt{i}\right)^2 = (i+1) - i = 1$). Thus,

$$\begin{aligned}
&\sum_{i=1}^{n} \frac{1}{\sqrt{i}+\sqrt{i+1}} \\
&= \sum_{i=1}^{n} \left(\sqrt{i+1}-\sqrt{i}\right) = \sum_{s=2}^{n+1} \left(\sqrt{s}-\sqrt{s-1}\right) \\
&\qquad \left(\begin{array}{c}\text{here, we have substituted } s-1 \text{ for } i \text{ in the sum,} \\ \text{since the map } \{2,3,\ldots,n+1\} \to \{1,2,\ldots,n\},\ s \mapsto s-1 \\ \text{is a bijection}\end{array}\right) \\
&= \sqrt{n+1} - \underbrace{\sqrt{2-1}}_{=\sqrt{1}=1} \\
&\qquad \left(\text{by (16), applied to } u=2,\ v=n+1 \text{ and } a_s = \sqrt{s}-\sqrt{s-1}\right) \\
&= \sqrt{n+1} - 1.
\end{aligned}$$

Remarks:

- When we use the equality (16) to rewrite the sum $\sum\limits_{s=u}^{v} (a_s - a_{s-1})$ as $a_v - a_{u-1}$, we can say that the sum $\sum\limits_{s=u}^{v} (a_s - a_{s-1})$ "telescopes" to $a_v - a_{u-1}$. A sum like $\sum\limits_{s=u}^{v} (a_s - a_{s-1})$ is said to be a "telescoping sum". This terminology references the idea that the sum $\sum\limits_{s=u}^{v} (a_s - a_{s-1})$ "shrink" to the simple difference $a_v - a_{u-1}$ like a telescope does when it is collapsed.
- Here is a *proof of (16):* Let u and v be two integers such that $u - 1 \leq v$. Let a_s be an element of $\mathbb{A}$ for each $s \in \{u-1, u, \ldots, v\}$. Then, (8) (applied to $a_s - a_{s-1}$ and a_{s-1} instead of a_s and b_s) yields

$$\sum_{s=u}^{v} \left((a_s - a_{s-1}) + a_{s-1}\right) = \sum_{s=u}^{v} (a_s - a_{s-1}) + \sum_{s=u}^{v} a_{s-1}.$$

Solving this equation for $\sum\limits_{s=u}^{v}(a_s - a_{s-1})$, we obtain

$$\sum_{s=u}^{v}(a_s - a_{s-1}) = \sum_{s=u}^{v}\underbrace{((a_s - a_{s-1}) + a_{s-1})}_{=a_s} - \underbrace{\sum_{s=u}^{v} a_{s-1}}_{\substack{=\sum\limits_{s=u-1}^{v-1} a_s \\ \text{(here, we have substituted } s \text{ for } s-1 \\ \text{in the sum)}}}$$
$$= \sum_{s=u}^{v} a_s - \sum_{s=u-1}^{v-1} a_s. \tag{19}$$

But $u-1 \leq v$. Hence, we can split off the addend for $s = u-1$ from the sum $\sum\limits_{s=u-1}^{v} a_s$. We thus obtain

$$\sum_{s=u-1}^{v} a_s = a_{u-1} + \sum_{s=u}^{v} a_s.$$

Solving this equation for $\sum\limits_{s=u}^{v} a_s$, we obtain

$$\sum_{s=u}^{v} a_s = \sum_{s=u-1}^{v} a_s - a_{u-1}. \tag{20}$$

Also, $u-1 \leq v$. Hence, we can split off the addend for $s = v$ from the sum $\sum\limits_{s=u-1}^{v} a_s$. We thus obtain

$$\sum_{s=u-1}^{v} a_s = a_v + \sum_{s=u-1}^{v-1} a_s.$$

Solving this equation for $\sum\limits_{s=u-1}^{v-1} a_s$, we obtain

$$\sum_{s=u-1}^{v-1} a_s = \sum_{s=u-1}^{v} a_s - a_v. \tag{21}$$

Now, (19) becomes

$$\sum_{s=u}^{v}(a_s - a_{s-1}) = \underbrace{\sum_{s=u}^{v} a_s}_{\substack{=\sum\limits_{s=u-1}^{v} a_s - a_{u-1} \\ \text{(by (20))}}} - \underbrace{\sum_{s=u-1}^{v-1} a_s}_{\substack{=\sum\limits_{s=u-1}^{v} a_s - a_v \\ \text{(by (21))}}}$$
$$= \left(\sum_{s=u-1}^{v} a_s - a_{u-1}\right) - \left(\sum_{s=u-1}^{v} a_s - a_v\right) = a_v - a_{u-1}.$$

This proves (16).

- **Restricting to a subset:** Let S be a finite set. Let T be a subset of S. Let a_s be an element of $\mathbb{A}$ for each $s \in T$. Then,

  $$\sum_{\substack{s \in S; \\ s \in T}} a_s = \sum_{s \in T} a_s .$$

  This is because the $s \in S$ satisfying $s \in T$ are exactly the elements of T.

  Remark: Here is a slightly more general form of this rule: Let S be a finite set. Let T be a subset of S. Let $\mathcal{A}(s)$ be a logical statement for each $s \in S$. Let a_s be an element of $\mathbb{A}$ for each $s \in T$ satisfying $\mathcal{A}(s)$. Then,

  $$\sum_{\substack{s \in S; \\ s \in T; \\ \mathcal{A}(s)}} a_s = \sum_{\substack{s \in T; \\ \mathcal{A}(s)}} a_s .$$

- **Splitting a sum by a value of a function:** Let S be a finite set. Let W be a set. Let $f : S \to W$ be a map. Let a_s be an element of $\mathbb{A}$ for each $s \in S$. Then,

  $$\sum_{s \in S} a_s = \sum_{w \in W} \sum_{\substack{s \in S; \\ f(s) = w}} a_s . \qquad (22)$$

  The idea behind this formula is the following: The left hand side is the sum of all a_s for $s \in S$. The right hand side is the same sum, but split in a particular way: First, for each $w \in W$, we sum the a_s for all $s \in S$ satisfying $f(s) = w$, and then we take the sum of all these "partial sums". For a rigorous proof of (22), see Theorem 2.127 (for the case when W is finite) and Theorem 2.147 (for the general case).

  Examples:

  - Let $n \in \mathbb{N}$. Then,

    $$\sum_{s \in \{-n, -(n-1), \ldots, n\}} s^3 = \sum_{w \in \{0,1,\ldots,n\}} \sum_{\substack{s \in \{-n, -(n-1), \ldots, n\}; \\ |s| = w}} s^3 . \qquad (23)$$

    (This follows from (22), applied to $S = \{-n, -(n-1), \ldots, n\}$, $W = \{0, 1, \ldots, n\}$ and $f(s) = |s|$.) You might wonder what you gain by this observation. But actually, it allows you to compute the sum: For any $w \in \{0, 1, \ldots, n\}$, the sum $\sum_{\substack{s \in \{-n, -(n-1), \ldots, n\}; \\ |s| = w}} s^3$ is 0 [21], and therefore (23)

[21] *Proof.* If $w = 0$, then this sum $\sum_{\substack{s \in \{-n, -(n-1), \ldots, n\}; \\ |s| = w}} s^3$ consists of one addend only, and this addend is 0^3. If $w > 0$, then this sum has two addends, namely $(-w)^3$ and w^3. In either case, the sum is 0 (because $0^3 = 0$ and $(-w)^3 + w^3 = -w^3 + w^3 = 0$).

becomes

$$\sum_{s\in\{-n,-(n-1),\ldots,n\}} s^3 = \sum_{w\in\{0,1,\ldots,n\}} \underbrace{\sum_{\substack{s\in\{-n,-(n-1),\ldots,n\};\\ |s|=w}} s^3}_{=0} = \sum_{w\in\{0,1,\ldots,n\}} 0 = 0.$$

Thus, a strategic application of (22) can help in evaluating a sum.

– Let S be a finite set. Let W be a set. Let $f : S \to W$ be a map. If we apply (22) to $a_s = 1$, then we obtain

$$\sum_{s\in S} 1 = \sum_{w\in W} \underbrace{\sum_{\substack{s\in S;\\ f(s)=w}} 1}_{\substack{=|\{s\in S \ \mid\ f(s)=w\}|\cdot 1\\ =|\{s\in S \ \mid\ f(s)=w\}|}} = \sum_{w\in W} |\{s \in S \ \mid\ f(s) = w\}|.$$

Since $\sum\limits_{s\in S} 1 = |S| \cdot 1 = |S|$, this rewrites as follows:

$$|S| = \sum_{w\in W} |\{s \in S \ \mid\ f(s) = w\}|. \qquad (24)$$

This equality is often called the *shepherd's principle*, because it is connected to the joke that "in order to count a flock of sheep, just count the legs and divide by 4". The connection is somewhat weak, actually; the equality (24) is better regarded as a formalization of the (less funny) idea that in order to count all legs of a flock of sheep, you can count the legs of every single sheep, and then sum the resulting numbers over all sheep in the flock. Think of the S in (24) as the set of all legs of all sheep in the flock; think of W as the set of all sheep in the flock; and think of f as the function which sends every leg to the (hopefully uniquely determined) sheep it belongs to.

Remarks:

– If $f : S \to W$ is a map between two sets S and W, and if w is an element of W, then it is common to denote the set $\{s \in S \ \mid\ f(s) = w\}$ by $f^{-1}(w)$. (Formally speaking, this notation might clash with the notation $f^{-1}(w)$ for the actual preimage of w when f happens to be bijective; but in practice, this causes far less confusion than it might seem to.) Using this notation, we can rewrite (22) as follows:

$$\sum_{s\in S} a_s = \sum_{w\in W} \underbrace{\sum_{\substack{s\in S;\\ f(s)=w}}}_{=\sum\limits_{s\in f^{-1}(w)}} a_s = \sum_{w\in W} \sum_{s\in f^{-1}(w)} a_s. \qquad (25)$$

- When I rewrite a sum $\sum\limits_{s\in S} a_s$ as $\sum\limits_{w\in W} \sum\limits_{\substack{s\in S;\\ f(s)=w}} a_s$ (or as $\sum\limits_{w\in W} \sum\limits_{s\in f^{-1}(w)} a_s$), I say that I am "splitting the sum according to the value of $f(s)$". (Though, most of the time, I shall be doing such manipulations without explicit mention.)

- **Splitting a sum into subsums:** Let S be a finite set. Let $S_1, S_2, \ldots, S_n$ be finitely many subsets of S. Assume that these subsets $S_1, S_2, \ldots, S_n$ are pairwise disjoint (i.e., we have $S_i \cap S_j = \varnothing$ for any two distinct elements i and j of $\{1, 2, \ldots, n\}$) and their union is S. (Thus, every element of S lies in precisely one of the subsets $S_1, S_2, \ldots, S_n$.) Let a_s be an element of $\mathbb{A}$ for each $s \in S$. Then,

$$\sum_{s\in S} a_s = \sum_{w=1}^{n} \sum_{s\in S_w} a_s. \tag{26}$$

This is a generalization of (3) (indeed, (3) is obtained from (26) by setting $n = 2$, $S_1 = X$ and $S_2 = Y$). It is also a consequence of (22): Indeed, set $W = \{1, 2, \ldots, n\}$, and define a map $f : S \to W$ to send each $s \in S$ to the unique $w \in \{1, 2, \ldots, n\}$ for which $s \in S_w$. Then, every $w \in W$ satisfies $\sum\limits_{\substack{s\in S;\\ f(s)=w}} a_s = \sum\limits_{s\in S_w} a_s$; therefore, (22) becomes (26).

Example: If we set $a_s = 1$ for each $s \in S$, then (26) becomes

$$\sum_{s\in S} 1 = \sum_{w=1}^{n} \underbrace{\sum_{s\in S_w} 1}_{=|S_w|} = \sum_{w=1}^{n} |S_w| .$$

Hence,

$$\sum_{w=1}^{n} |S_w| = \sum_{s\in S} 1 = |S| \cdot 1 = |S| .$$

- **Fubini's theorem (interchanging the order of summation):** Let X and Y be two finite sets. Let $a_{(x,y)}$ be an element of $\mathbb{A}$ for each $(x, y) \in X \times Y$. Then,

$$\sum_{x\in X} \sum_{y\in Y} a_{(x,y)} = \sum_{(x,y)\in X\times Y} a_{(x,y)} = \sum_{y\in Y} \sum_{x\in X} a_{(x,y)}. \tag{27}$$

This is called *Fubini's theorem for finite sums*, and is a lot easier to prove than what analysts tend to call Fubini's theorem. I shall sketch a proof shortly (in the Remark below); but first, let me give some intuition for the statement. Imagine that you have a rectangular table filled with numbers. If you want to sum the numbers in the table, you can proceed in several ways. One way is to sum the numbers in each row, and then sum all the sums you have obtained. Another way is to sum the numbers in each column, and then sum all the obtained sums. Either way, you get the same result – namely, the sum of all numbers in the table. This is essentially what (27) says, at least when $X = \{1, 2, \ldots, n\}$ and $Y = \{1, 2, \ldots, m\}$ for some integers n and m. In this case, the numbers $a_{(x,y)}$ can be viewed as forming a table, where $a_{(x,y)}$ is placed in the cell at the intersection of row x with column y. When X and Y are arbitrary finite sets (not necessarily $\{1, 2, \ldots, n\}$ and $\{1, 2, \ldots, m\}$), then you need to slightly stretch your imagination in order to see the $a_{(x,y)}$ as "forming a table"; in fact, there is no obvious order in which the numbers appear in a row or column, but there is still a notion of rows and columns.

Examples:

– Let $n \in \mathbb{N}$ and $m \in \mathbb{N}$. Let $a_{(x,y)}$ be an element of $\mathbb{A}$ for each $(x, y) \in \{1, 2, \ldots, n\} \times \{1, 2, \ldots, m\}$. Then,

$$\sum_{x=1}^{n} \sum_{y=1}^{m} a_{(x,y)} = \sum_{(x,y)\in\{1,2,\ldots,n\}\times\{1,2,\ldots,m\}} a_{(x,y)} = \sum_{y=1}^{m} \sum_{x=1}^{n} a_{(x,y)}. \tag{28}$$

(This follows from (27), applied to $X = \{1, 2, \ldots, n\}$ and $Y = \{1, 2, \ldots, m\}$.) We can rewrite the equality (28) without using $\sum$ signs; it then takes the following form:

$$\begin{aligned}
&\left(a_{(1,1)} + a_{(1,2)} + \cdots + a_{(1,m)}\right) \\
&\qquad + \left(a_{(2,1)} + a_{(2,2)} + \cdots + a_{(2,m)}\right) \\
&\qquad + \cdots \\
&\qquad + \left(a_{(n,1)} + a_{(n,2)} + \cdots + a_{(n,m)}\right) \\
&= a_{(1,1)} + a_{(1,2)} + \cdots + a_{(n,m)} \qquad \left(\text{this is the sum of all } nm \text{ numbers } a_{(x,y)}\right) \\
&= \left(a_{(1,1)} + a_{(2,1)} + \cdots + a_{(n,1)}\right) \\
&\qquad + \left(a_{(1,2)} + a_{(2,2)} + \cdots + a_{(n,2)}\right) \\
&\qquad + \cdots \\
&\qquad + \left(a_{(1,m)} + a_{(2,m)} + \cdots + a_{(n,m)}\right).
\end{aligned}$$

In other words, we can sum the entries of the rectangular table

$$\boxed{\begin{array}{cccc}
a_{(1,1)} & a_{(1,2)} & \cdots & a_{(1,m)} \\
a_{(2,1)} & a_{(2,2)} & \cdots & a_{(2,m)} \\
\vdots & \vdots & \ddots & \vdots \\
a_{(n,1)} & a_{(n,2)} & \cdots & a_{(n,m)}
\end{array}}$$

in three different ways:

(a) row by row (i.e., first summing the entries in each row, then summing up the n resulting tallies);

(b) arbitrarily (i.e., just summing all entries of the table in some arbitrary order);

(c) column by column (i.e., first summing the entries in each column, then summing up the m resulting tallies);

and each time, we get the same result.

- Here is a concrete application of (28): Let $n \in \mathbb{N}$ and $m \in \mathbb{N}$. We want to compute $\sum\limits_{(x,y)\in\{1,2,\ldots,n\}\times\{1,2,\ldots,m\}} xy$. (This is the sum of all entries of the $n \times m$ multiplication table.) Applying (28) to $a_{(x,y)} = xy$, we obtain

$$\sum_{x=1}^{n}\sum_{y=1}^{m} xy = \sum_{(x,y)\in\{1,2,\ldots,n\}\times\{1,2,\ldots,m\}} xy = \sum_{y=1}^{m}\sum_{x=1}^{n} xy.$$

Hence,

$$\begin{aligned}
&\sum_{(x,y)\in\{1,2,\ldots,n\}\times\{1,2,\ldots,m\}} xy \\
&= \sum_{x=1}^{n} \underbrace{\sum_{y=1}^{m} xy}_{\substack{=\sum_{s=1}^{m} xs = x\sum_{s=1}^{m} s \\ \text{(by (9), applied to } S=\{1,2,\ldots,m\}, \\ a_s=s \text{ and } \lambda=x)}} = \sum_{x=1}^{n} x \underbrace{\sum_{s=1}^{m} s}_{\substack{=\sum_{i=1}^{m} i=\frac{m(m+1)}{2} \\ \text{(by (14), applied to } m \\ \text{instead of } n)}} \\
&= \sum_{x=1}^{n} x\frac{m(m+1)}{2} = \sum_{x=1}^{n} \frac{m(m+1)}{2}x = \sum_{s=1}^{n} \frac{m(m+1)}{2}s \\
&= \frac{m(m+1)}{2} \underbrace{\sum_{s=1}^{n} s}_{\substack{=\sum_{i=1}^{n} i=\frac{n(n+1)}{2} \\ \text{(by (14))}}} \\
&\qquad \left(\text{by (9), applied to } S = \{1,2,\ldots,n\},\ a_s = s \text{ and } \lambda = \frac{m(m+1)}{2}\right) \\
&= \frac{m(m+1)}{2} \cdot \frac{n(n+1)}{2}.
\end{aligned}$$

Remarks:

- I have promised to outline a proof of (27). Here it comes: Let $S = X \times Y$ and $W = Y$, and let $f : S \to W$ be the map which sends every pair (x,y)

to its second entry y. Then, (25) shows that

$$\sum_{s\in X\times Y} a_s = \sum_{w\in Y} \sum_{s\in f^{-1}(w)} a_s. \tag{29}$$

However, for every given $w \in Y$, the set $f^{-1}(w)$ is simply the set of all pairs (x, w) with $x \in X$. Thus, for every given $w \in Y$, there is a bijection $g_w : X \to f^{-1}(w)$ given by

$$g_w(x) = (x, w) \qquad \text{for all } x \in X.$$

Hence, for every given $w \in Y$, we can substitute $g_w(x)$ for s in the sum $\sum\limits_{s\in f^{-1}(w)} a_s$, and thus obtain

$$\sum_{s\in f^{-1}(w)} a_s = \sum_{x\in X} \underbrace{a_{g_w(x)}}_{\substack{=a_{(x,w)} \\ (\text{since } g_w(x)=(x,w))}} = \sum_{x\in X} a_{(x,w)}.$$

Hence, (29) becomes

$$\sum_{s\in X\times Y} a_s = \sum_{w\in Y} \underbrace{\sum_{s\in f^{-1}(w)} a_s}_{=\sum\limits_{x\in X} a_{(x,w)}} = \sum_{w\in Y} \sum_{x\in X} a_{(x,w)} = \sum_{y\in Y} \sum_{x\in X} a_{(x,y)}$$

(here, we have renamed the summation index w as y in the outer sum). Therefore,

$$\sum_{y\in Y} \sum_{x\in X} a_{(x,y)} = \sum_{s\in X\times Y} a_s = \sum_{(x,y)\in X\times Y} a_{(x,y)}$$

(here, we have renamed the summation index s as (x, y)). Thus, we have proven the second part of the equality (27). The first part can be proven similarly.

– I like to abbreviate the equality (28) as follows:

$$\sum_{x=1}^{n} \sum_{y=1}^{m} = \sum_{(x,y)\in\{1,2,...,n\}\times\{1,2,...,m\}} = \sum_{y=1}^{m} \sum_{x=1}^{n}. \tag{30}$$

This is an "equality between summation signs"; it should be understood as follows: Every time you see an "$\sum\limits_{x=1}^{n} \sum\limits_{y=1}^{m}$" in an expression, you can replace it by a "$\sum\limits_{(x,y)\in\{1,2,...,n\}\times\{1,2,...,m\}}$" or by a "$\sum\limits_{y=1}^{m} \sum\limits_{x=1}^{n}$", and similarly the other ways round.

- **Triangular Fubini's theorem I:** The equality (28) formalizes the idea that we can sum the entries of a rectangular table by first tallying each row and then adding together, or first tallying each column and adding together. The same holds for triangular tables. More precisely: Let $n \in \mathbb{N}$. Let T_n be the set $\left\{(x,y) \in \{1,2,3,\ldots\}^2 \mid x+y \leq n\right\}$. (For instance, if $n=3$, then $T_n = T_3 = \{(1,1), (1,2), (2,1)\}$.) Let $a_{(x,y)}$ be an element of $\mathbb{A}$ for each $(x,y) \in T_n$. Then,

$$\sum_{x=1}^{n} \sum_{y=1}^{n-x} a_{(x,y)} = \sum_{(x,y)\in T_n} a_{(x,y)} = \sum_{y=1}^{n} \sum_{x=1}^{n-y} a_{(x,y)}. \qquad (31)$$

Examples:

- In the case when $n=4$, the formula (31) (rewritten without the use of $\sum$ signs) looks as follows:

$$\begin{aligned} &\left(a_{(1,1)} + a_{(1,2)} + a_{(1,3)}\right) + \left(a_{(2,1)} + a_{(2,2)}\right) + a_{(3,1)} + (\text{empty sum}) \\ &= \left(\text{the sum of the } a_{(x,y)} \text{ for all } (x,y) \in T_4\right) \\ &= \left(a_{(1,1)} + a_{(2,1)} + a_{(3,1)}\right) + \left(a_{(1,2)} + a_{(2,2)}\right) + a_{(1,3)} + (\text{empty sum}). \end{aligned}$$

In other words, we can sum the entries of the triangular table

$a_{(1,1)}$	$a_{(1,2)}$	$a_{(1,3)}$
$a_{(2,1)}$	$a_{(2,2)}$	
$a_{(3,1)}$		

in three different ways:

(a) row by row (i.e., first summing the entries in each row, then summing up the resulting tallies);

(b) arbitrarily (i.e., just summing all entries of the table in some arbitrary order);

(c) column by column (i.e., first summing the entries in each column, then summing up the resulting tallies);

and each time, we get the same result.

- Let us use (31) to compute $|T_n|$. Indeed, we can apply (31) to $a_{(x,y)} = 1$. Thus, we obtain

$$\sum_{x=1}^{n} \sum_{y=1}^{n-x} 1 = \sum_{(x,y)\in T_n} 1 = \sum_{y=1}^{n} \sum_{x=1}^{n-y} 1.$$

Hence,

$$\sum_{x=1}^{n} \sum_{y=1}^{n-x} 1 = \sum_{(x,y)\in T_n} 1 = |T_n|,$$

so that

$$\begin{aligned}
|T_n| &= \sum_{x=1}^{n} \underbrace{\sum_{y=1}^{n-x} 1}_{=n-x} = \sum_{x=1}^{n} (n-x) = \sum_{i=0}^{n-1} i \\
&\qquad \left(\begin{array}{c}\text{here, we have substituted } i \text{ for } n-x \text{ in the sum,} \\ \text{since the map } \{1,2,\ldots,n\} \to \{0,1,\ldots,n-1\},\ x \mapsto n-x \\ \text{is a bijection}\end{array}\right) \\
&= \frac{(n-1)((n-1)+1)}{2} \qquad \text{(by (13), applied to } n-1 \text{ instead of } n) \\
&= \frac{(n-1)n}{2}.
\end{aligned}$$

Remarks:

- The sum $\sum_{(x,y)\in T_n} a_{(x,y)}$ in (31) can also be rewritten as $\sum_{\substack{(x,y)\in\{1,2,3,\ldots\}^2;\\ x+y\leq n}} a_{(x,y)}$.

- Let us prove (31). Indeed, the proof will be very similar to our proof of (27) above. Let $S = T_n$ and $W = \{1,2,\ldots,n\}$, and let $f : S \to W$ be the map which sends every pair (x,y) to its second entry y. Then, (25) shows that

$$\sum_{s\in T_n} a_s = \sum_{w\in W} \sum_{s\in f^{-1}(w)} a_s. \tag{32}$$

  However, for every given $w \in W$, the set $f^{-1}(w)$ is simply the set of all pairs (x,w) with $x \in \{1,2,\ldots,n-w\}$. Thus, for every given $w \in W$, there is a bijection $g_w : \{1,2,\ldots,n-w\} \to f^{-1}(w)$ given by

$$g_w(x) = (x,w) \qquad \text{for all } x \in \{1,2,\ldots,n-w\}.$$

  Hence, for every given $w \in W$, we can substitute $g_w(x)$ for s in the sum $\sum_{s\in f^{-1}(w)} a_s$, and thus obtain

$$\sum_{s\in f^{-1}(w)} a_s = \underbrace{\sum_{x\in\{1,2,\ldots,n-w\}}}_{=\sum_{x=1}^{n-w}} \underbrace{a_{g_w(x)}}_{\substack{=a_{(x,w)}\\ \text{(since } g_w(x)=(x,w))}} = \sum_{x=1}^{n-w} a_{(x,w)}.$$

Hence, (32) becomes

$$\sum_{s\in T_n} a_s = \underbrace{\sum_{w\in W}}_{\substack{=\sum_{w=1}^{n} \\ (\text{since } W=\{1,2,\ldots,n\})}} \underbrace{\sum_{s\in f^{-1}(w)} a_s}_{=\sum_{x=1}^{n-w} a_{(x,w)}} = \sum_{w=1}^{n} \sum_{x=1}^{n-w} a_{(x,w)} = \sum_{y=1}^{n} \sum_{x=1}^{n-y} a_{(x,y)}$$

(here, we have renamed the summation index w as y in the outer sum). Therefore,

$$\sum_{y=1}^{n} \sum_{x=1}^{n-y} a_{(x,y)} = \sum_{s\in T_n} a_s = \sum_{(x,y)\in T_n} a_{(x,y)}.$$

Thus, we have proven the second part of the equality (31). The first part can be proven similarly.

- **Triangular Fubini's theorem II:** Here is another equality similar to (31). Let $n \in \mathbb{N}$. Let Q_n be the set $\left\{(x,y) \in \{1,2,\ldots,n\}^2 \mid x \leq y\right\}$. (For instance, if $n = 3$, then $Q_n = Q_3 = \{(1,1), (1,2), (1,3), (2,2), (2,3), (3,3)\}$.) Let $a_{(x,y)}$ be an element of $\mathbb{A}$ for each $(x,y) \in Q_n$. Then,

$$\sum_{x=1}^{n} \sum_{y=x}^{n} a_{(x,y)} = \sum_{(x,y)\in Q_n} a_{(x,y)} = \sum_{y=1}^{n} \sum_{x=1}^{y} a_{(x,y)}. \tag{33}$$

Examples:

  - In the case when $n = 4$, the formula (33) (rewritten without the use of $\sum$ signs) looks as follows:

$$\begin{aligned} &\left(a_{(1,1)} + a_{(1,2)} + a_{(1,3)} + a_{(1,4)}\right) \\ &\qquad + \left(a_{(2,2)} + a_{(2,3)} + a_{(2,4)}\right) \\ &\qquad + \left(a_{(3,3)} + a_{(3,4)}\right) \\ &\qquad + a_{(4,4)} \\ &= \left(\text{the sum of the } a_{(x,y)} \text{ for all } (x,y) \in Q_4\right) \\ &= a_{(1,1)} \\ &\qquad + \left(a_{(1,2)} + a_{(2,2)}\right) \\ &\qquad + \left(a_{(1,3)} + a_{(2,3)} + a_{(3,3)}\right) \\ &\qquad + \left(a_{(1,4)} + a_{(2,4)} + a_{(3,4)} + a_{(4,4)}\right). \end{aligned}$$

In other words, we can sum the entries of the triangular table

$a_{(1,1)}$	$a_{(1,2)}$	$a_{(1,3)}$	$a_{(1,4)}$
	$a_{(2,2)}$	$a_{(2,3)}$	$a_{(2,4)}$
		$a_{(3,3)}$	$a_{(3,4)}$
			$a_{(4,4)}$

in three different ways:

(a) row by row (i.e., first summing the entries in each row, then summing up the resulting tallies);

(b) arbitrarily (i.e., just summing all entries of the table in some arbitrary order);

(c) column by column (i.e., first summing the entries in each column, then summing up the resulting tallies);

and each time, we get the same result.

– Let us use (33) to compute $|Q_n|$. Indeed, we can apply (33) to $a_{(x,y)} = 1$. Thus, we obtain

$$\sum_{x=1}^{n} \sum_{y=x}^{n} 1 = \sum_{(x,y)\in Q_n} 1 = \sum_{y=1}^{n} \sum_{x=1}^{y} 1.$$

Hence,

$$\sum_{y=1}^{n} \sum_{x=1}^{y} 1 = \sum_{(x,y)\in Q_n} 1 = |Q_n| ,$$

so that

$$|Q_n| = \sum_{y=1}^{n} \underbrace{\sum_{x=1}^{y} 1}_{=y} = \sum_{y=1}^{n} y = \sum_{i=1}^{n} i = \frac{n(n+1)}{2} \qquad \text{(by (14))} .$$

Remarks:

– The sum $\sum_{(x,y)\in Q_n} a_{(x,y)}$ in (33) can also be rewritten as $\sum_{\substack{(x,y)\in\{1,2,\ldots,n\}^2; \\ x\leq y}} a_{(x,y)}$.

It is also often written as $\sum_{1\leq x\leq y\leq n} a_{(x,y)}$.

– The proof of (33) is similar to that of (31).

- **Fubini's theorem with a predicate:** Let X and Y be two finite sets. For every pair $(x,y) \in X \times Y$, let $\mathcal{A}(x,y)$ be a logical statement. For each $(x,y) \in X \times Y$ satisfying $\mathcal{A}(x,y)$, let $a_{(x,y)}$ be an element of $\mathbb{A}$. Then,

  $$\sum_{x\in X} \sum_{\substack{y\in Y;\\ \mathcal{A}(x,y)}} a_{(x,y)} = \sum_{\substack{(x,y)\in X\times Y;\\ \mathcal{A}(x,y)}} a_{(x,y)} = \sum_{y\in Y} \sum_{\substack{x\in X;\\ \mathcal{A}(x,y)}} a_{(x,y)}. \qquad (34)$$

  Examples:

  - For any $n \in \mathbb{N}$ and $m \in \mathbb{N}$, we have

    $$\sum_{x\in\{1,2,\ldots,n\}} \sum_{\substack{y\in\{1,2,\ldots,m\};\\ x+y \text{ is even}}} xy = \sum_{\substack{(x,y)\in\{1,2,\ldots,n\}\times\{1,2,\ldots,m\};\\ x+y \text{ is even}}} xy = \sum_{y\in\{1,2,\ldots,m\}} \sum_{\substack{x\in\{1,2,\ldots,n\};\\ x+y \text{ is even}}} xy.$$

    (This follows from (34), applied to $X = \{1,2,\ldots,n\}$, $Y = \{1,2,\ldots,m\}$ and $\mathcal{A}(x,y) = (\text{“}x+y \text{ is even”})$.)

  Remarks:

  - We have assumed that the sets X and Y are finite. But (34) is still valid if we replace this assumption by the weaker assumption that only finitely many $(x,y) \in X \times Y$ satisfy $\mathcal{A}(x,y)$.
  - It is not hard to prove (34) by suitably adapting our proof of (27).
  - The equality (31) can be derived from (34) by setting $X = \{1,2,\ldots,n\}$, $Y = \{1,2,\ldots,n\}$ and $\mathcal{A}(x,y) = (\text{“}x+y \leq n\text{”})$. Similarly, the equality (33) can be derived from (34) by setting $X = \{1,2,\ldots,n\}$, $Y = \{1,2,\ldots,n\}$ and $\mathcal{A}(x,y) = (\text{“}x \leq y\text{”})$.

- **Interchange of predicates:** Let S be a finite set. For every $s \in S$, let $\mathcal{A}(s)$ and $\mathcal{B}(s)$ be two equivalent logical statements. ("Equivalent" means that $\mathcal{A}(s)$ holds if and only if $\mathcal{B}(s)$ holds.) Let a_s be an element of $\mathbb{A}$ for each $s \in S$. Then,

  $$\sum_{\substack{s\in S;\\ \mathcal{A}(s)}} a_s = \sum_{\substack{s\in S;\\ \mathcal{B}(s)}} a_s.$$

  (If you regard equivalent logical statements as identical, then you will see this as a tautology. If not, it is still completely obvious, since the equivalence of $\mathcal{A}(s)$ with $\mathcal{B}(s)$ shows that $\{t \in S \mid \mathcal{A}(t)\} = \{t \in S \mid \mathcal{B}(t)\}$.)

- **Substituting the index I with a predicate:** Let S and T be two finite sets. Let $f : S \to T$ be a **bijective** map. Let a_t be an element of $\mathbb{A}$ for each $t \in T$. For every $t \in T$, let $\mathcal{A}(t)$ be a logical statement. Then,

$$\sum_{\substack{t \in T; \\ \mathcal{A}(t)}} a_t = \sum_{\substack{s \in S; \\ \mathcal{A}(f(s))}} a_{f(s)}. \qquad (35)$$

Remarks:

- The equality (35) is a generalization of (12). There is a similar generalization of (15).
- The equality (35) can be easily derived from (12). Indeed, let S' be the subset $\{s \in S \mid \mathcal{A}(f(s))\}$ of S, and let T' be the subset $\{t \in T \mid \mathcal{A}(t)\}$ of T. Then, the map $S' \to T'$, $s \mapsto f(s)$ is well-defined and a bijection[22], and thus (12) (applied to S', T' and this map instead of S, T and f) yields $\sum\limits_{t \in T'} a_t = \sum\limits_{s \in S'} a_{f(s)}$. But this is precisely the equality (35), because clearly we have $\sum\limits_{t \in T'} = \sum\limits_{\substack{t \in T; \\ \mathcal{A}(t)}}$ and $\sum\limits_{s \in S'} = \sum\limits_{\substack{s \in S; \\ \mathcal{A}(f(s))}}$.

1.4.3. Definition of $\prod$

We shall now define the $\prod$ sign. Since the $\prod$ sign is (in many aspects) analogous to the $\sum$ sign, we shall be brief and confine ourselves to the bare necessities; we trust the reader to transfer most of what we said about $\sum$ to the case of $\prod$. In particular, we shall give very few examples and no proofs.

- If S is a finite set, and if a_s is an element of $\mathbb{A}$ for each $s \in S$, then $\prod_{s \in S} a_s$ denotes the product of all of these elements a_s. Formally, this product is defined by recursion on $|S|$, as follows:
  - If $|S| = 0$, then $\prod\limits_{s \in S} a_s$ is defined to be 1.
  - Let $n \in \mathbb{N}$. Assume that we have defined $\prod\limits_{s \in S} a_s$ for every finite set S with $|S| = n$ (and every choice of elements a_s of $\mathbb{A}$). Now, if S is a finite set with $|S| = n + 1$ (and if $a_s \in \mathbb{A}$ are chosen for all $s \in S$), then $\prod\limits_{s \in S} a_s$ is defined by picking any $t \in S$ and setting

    $$\prod_{s \in S} a_s = a_t \cdot \prod_{s \in S \setminus \{t\}} a_s. \qquad (36)$$

    As for $\sum\limits_{s \in S} a_s$, this definition is not obviously legitimate, but it can be proven to be legitimate nevertheless. (The proof is analogous to the proof for $\sum\limits_{s \in S} a_s$; see Subsection 2.14.14 for details.)

[22] This is easy to see.

Examples:

- If $S = \{1, 2, \ldots, n\}$ (for some $n \in \mathbb{N}$) and $a_s = s$ for every $s \in S$, then $\prod\limits_{s \in S} a_s = \prod\limits_{s \in S} s = 1 \cdot 2 \cdot \cdots \cdot n$. This number $1 \cdot 2 \cdot \cdots \cdot n$ is denoted by $n!$ and called the *factorial of n*.

  In particular,

  $$\begin{aligned} 0! &= \prod_{s \in \{1,2,\ldots,0\}} s = \prod_{s \in \varnothing} s \qquad (\text{since } \{1, 2, \ldots, 0\} = \varnothing) \\ &= 1 \qquad (\text{since } |\varnothing| = 0); \\ 1! &= \prod_{s \in \{1,2,\ldots,1\}} s = \prod_{s \in \{1\}} s = 1; \\ 2! &= \prod_{s \in \{1,2,\ldots,2\}} s = \prod_{s \in \{1,2\}} s = 1 \cdot 2 = 2; \\ 3! &= \prod_{s \in \{1,2,\ldots,3\}} s = \prod_{s \in \{1,2,3\}} s = 1 \cdot 2 \cdot 3 = 6; \end{aligned}$$

  similarly,

  $$4! = 1 \cdot 2 \cdot 3 \cdot 4 = 24; \qquad 5! = 120; \qquad 6! = 720; \qquad 7! = 5040.$$

  Notice that

  $$n! = n \cdot (n-1)! \qquad \text{for any positive integer } n. \tag{37}$$

  (This can be obtained from (36), applied to $S = \{1, 2, \ldots, n\}$, $a_s = s$ and $t = n$.)

Remarks:

- The product $\prod\limits_{s \in S} a_s$ is usually pronounced "product of the a_s over all $s \in S$" or "product of the a_s with s ranging over S" or "product of the a_s with s running through all elements of S". The letter "s" in the product is called the "product index", and its exact choice is immaterial, as long as it does not already have a different meaning outside of the product. The sign $\prod$ itself is called "the product sign" or "the $\prod$ sign". The numbers a_s are called the *factors* of the product $\prod\limits_{s \in S} a_s$. More precisely, for any given $t \in S$, we can refer to the number a_t as the "factor corresponding to the index t" (or as the "factor for $s = t$", or as the "factor for t") of the product $\prod\limits_{s \in S} a_s$.
- When the set S is empty, the product $\prod\limits_{s \in S} a_s$ is called an *empty product*. Our definition implies that any empty product is 1. This convention is

used throughout mathematics, except in rare occasions where a slightly subtler version of it is used[23].

- If $a \in \mathbb{A}$ and $n \in \mathbb{N}$, then the n-th power of a (written a^n) is defined by

$$a^n = \underbrace{a \cdot a \cdot \cdots \cdot a}_{n \text{ times}} = \prod_{i \in \{1,2,\ldots,n\}} a.$$

  Thus, a^0 is an empty product, and therefore equal to 1. This holds for any $a \in \mathbb{A}$, including 0; thus, $0^0 = 1$. **There is nothing controversial about the equality** $0^0 = 1$; it is a consequence of the only reasonable definition of the n-th power of a number. Ignore anyone who tells you that 0^0 is "undefined" or "indeterminate" or "can be 0 or 1 or anything, depending on the context".[24]

- The product index (just like a summation index) needs not be a single letter; it can be a pair or a triple, for example.

- Mathematicians don't seem to have reached an agreement on the operator precedence of the $\prod$ sign. My convention is that the product sign has higher precedence than the plus sign (so an expression like $\prod_{s \in S} a_s + b$ must be read as $\left(\prod_{s \in S} a_s \right) + b$, and not as $\prod_{s \in S} (a_s + b)$); this is, of course, in line with the standard convention that multiplication-like operations have higher precedence than addition-like operations ("PEMDAS"). Be warned that some authors disagree even with this convention. I strongly advise against writing things like $\prod_{s \in S} a_s b$, since it might mean both $\left(\prod_{s \in S} a_s \right) b$ and $\prod_{s \in S} (a_s b)$ depending on the weather. In particular, I advise against writing things like $\prod_{s \in S} a_s \cdot \prod_{s \in S} b_s$ without parentheses (although I do use a similar convention for sums, namely $\sum_{s \in S} a_s + \sum_{s \in S} b_s$, and I find it to be fairly harmless). These rules are not carved in stone, and you should use whatever conventions make **you** safe from ambiguity; either way, you should keep in mind that other authors make different choices.

- An expression of the form "$\prod_{s \in S} a_s$" (where S is a finite set) is called a *finite product*.

- We have required the set S to be finite when defining $\prod_{s \in S} a_s$. Such products are not generally defined when S is infinite. However, **some** infinite

[23] Just as with sums, the subtlety lies in the fact that mathematicians sometimes want an empty product to be not the integer 1 but the unity of some ring. As before, this does not matter for us right now.

[24] I am talking about the **number** 0^0 here. There is also something called "the indeterminate form 0^0", which is a much different story.

products can be made sense of. The simplest case is when the set S might be infinite, but only finitely many among the a_s are distinct from 1. In this case, we can define $\prod\limits_{s\in S} a_s$ simply by discarding the factors which equal 1 and multiplying the finitely many remaining factors. Other situations in which infinite products make sense appear in analysis and in topological algebra.

- The product $\prod\limits_{s\in S} a_s$ always belongs to $\mathbb{A}$.

- A slightly more complicated version of the product sign is the following: Let S be a finite set, and let $\mathcal{A}(s)$ be a logical statement defined for every $s \in S$. For each $s \in S$ satisfying $\mathcal{A}(s)$, let a_s be an element of $\mathbb{A}$. Then, the product $\prod\limits_{\substack{s\in S;\\ \mathcal{A}(s)}} a_s$ is defined by

$$\prod_{\substack{s\in S;\\ \mathcal{A}(s)}} a_s = \prod_{s\in\{t\in S \mid \mathcal{A}(t)\}} a_s .$$

- Finally, here is the simplest version of the product sign: Let u and v be two integers. As before, we understand the set $\{u, u+1, \ldots, v\}$ to be empty when $u > v$. Let a_s be an element of $\mathbb{A}$ for each $s \in \{u, u+1, \ldots, v\}$. Then, $\prod\limits_{s=u}^{v} a_s$ is defined by

$$\prod_{s=u}^{v} a_s = \prod_{s\in\{u,u+1,\ldots,v\}} a_s .$$

Examples:

  - We have $\prod\limits_{s=1}^{n} s = 1 \cdot 2 \cdot \cdots \cdot n = n!$ for each $n \in \mathbb{N}$.

Remarks:

  - The product $\prod\limits_{s=u}^{v} a_s$ is usually pronounced "product of the a_s for all s from u to v (inclusive)". It is often written $a_u \cdot a_{u+1} \cdot \cdots \cdot a_v$ (or just $a_u a_{u+1} \cdots a_v$), but this latter notation has the same drawbacks as the similar notation $a_u + a_{u+1} + \cdots + a_v$ for $\sum\limits_{s=u}^{v} a_s$.

  - The product $\prod\limits_{s=u}^{v} a_s$ is said to be *empty* whenever $u > v$. As with sums, it does not matter how much smaller v is than u; as long as v is smaller than u, the product is empty and equals 1.

Thus we have introduced the main three forms of the product sign.

1.4.4. Properties of $\prod$

Now, let me summarize the most important properties of the $\prod$ sign. These properties mirror the properties of $\sum$ discussed before; thus, I will again be brief.

- **Splitting-off:** Let S be a finite set. Let $t \in S$. Let a_s be an element of $\mathbb{A}$ for each $s \in S$. Then,
$$\prod_{s\in S} a_s = a_t \cdot \prod_{s\in S\setminus\{t\}} a_s.$$

- **Splitting:** Let S be a finite set. Let X and Y be two subsets of S such that $X \cap Y = \varnothing$ and $X \cup Y = S$. (Equivalently, X and Y are two subsets of S such that each element of S lies in **exactly** one of X and Y.) Let a_s be an element of $\mathbb{A}$ for each $s \in S$. Then,
$$\prod_{s\in S} a_s = \left(\prod_{s\in X} a_s\right) \cdot \left(\prod_{s\in Y} a_s\right).$$

- **Splitting using a predicate:** Let S be a finite set. Let $\mathcal{A}(s)$ be a logical statement for each $s \in S$. Let a_s be an element of $\mathbb{A}$ for each $s \in S$. Then,
$$\prod_{s\in S} a_s = \left(\prod_{\substack{s\in S;\\ \mathcal{A}(s)}} a_s\right) \cdot \left(\prod_{\substack{s\in S;\\ \text{not } \mathcal{A}(s)}} a_s\right).$$

- **Multiplying equal values:** Let S be a finite set. Let a be an element of $\mathbb{A}$. Then,
$$\prod_{s\in S} a = a^{|S|}.$$

- **Splitting a factor:** Let S be a finite set. For every $s \in S$, let a_s and b_s be elements of $\mathbb{A}$. Then,
$$\prod_{s\in S} (a_s b_s) = \left(\prod_{s\in S} a_s\right) \cdot \left(\prod_{s\in S} b_s\right). \tag{38}$$

  Examples:

  - Here is a frequently used particular case of (38): Let S be a finite set. For every $s \in S$, let b_s be an element of $\mathbb{A}$. Let a be an element of $\mathbb{A}$. Then, (38) (applied to $a_s = a$) yields
$$\prod_{s\in S} (a b_s) = \underbrace{\left(\prod_{s\in S} a\right)}_{=a^{|S|}} \cdot \left(\prod_{s\in S} b_s\right) = a^{|S|} \cdot \left(\prod_{s\in S} b_s\right). \tag{39}$$

- Here is an even further particular case: Let S be a finite set. For every $s \in S$, let b_s be an element of $\mathbb{A}$. Then,

  $$\prod_{s\in S} \underbrace{(-b_s)}_{=(-1)b_s} = \prod_{s\in S} ((-1) b_s) = (-1)^{|S|} \cdot \left(\prod_{s\in S} b_s \right)$$

  (by (39), applied to $a = -1$).

- **Factoring out an exponent:** Let S be a finite set. For every $s \in S$, let a_s be an element of $\mathbb{A}$. Also, let $\lambda \in \mathbb{N}$. Then,

  $$\prod_{s\in S} a_s^{\lambda} = \left(\prod_{s\in S} a_s \right)^{\lambda}.$$

- **Factoring out an integer exponent:** Let S be a finite set. For every $s \in S$, let a_s be a nonzero element of $\mathbb{A}$. Also, let $\lambda \in \mathbb{Z}$. Then,

  $$\prod_{s\in S} a_s^{\lambda} = \left(\prod_{s\in S} a_s \right)^{\lambda}.$$

  Remark: Applying this to $\lambda = -1$, we obtain

  $$\prod_{s\in S} a_s^{-1} = \left(\prod_{s\in S} a_s \right)^{-1}.$$

  In other words,

  $$\prod_{s\in S} \frac{1}{a_s} = \frac{1}{\prod_{s\in S} a_s}.$$

- **Ones multiply to one:** Let S be a finite set. Then,

  $$\prod_{s\in S} 1 = 1.$$

- **Dropping ones:** Let S be a finite set. Let a_s be an element of $\mathbb{A}$ for each $s \in S$. Let T be a subset of S such that every $s \in T$ satisfies $a_s = 1$. Then,

  $$\prod_{s\in S} a_s = \prod_{s\in S\setminus T} a_s.$$

- **Renaming the index:** Let S be a finite set. Let a_s be an element of $\mathbb{A}$ for each $s \in S$. Then,

  $$\prod_{s\in S} a_s = \prod_{t\in S} a_t.$$

- **Substituting the index I:** Let S and T be two finite sets. Let $f : S \to T$ be a **bijective** map. Let a_t be an element of $\mathbb{A}$ for each $t \in T$. Then,

$$\prod_{t \in T} a_t = \prod_{s \in S} a_{f(s)}.$$

- **Substituting the index II:** Let S and T be two finite sets. Let $f : S \to T$ be a **bijective** map. Let a_s be an element of $\mathbb{A}$ for each $s \in S$. Then,

$$\prod_{s \in S} a_s = \prod_{t \in T} a_{f^{-1}(t)}.$$

- **Telescoping products:** Let u and v be two integers such that $u - 1 \leq v$. Let a_s be an element of $\mathbb{A}$ for each $s \in \{u-1, u, \ldots, v\}$. Then,

$$\prod_{s=u}^{v} \frac{a_s}{a_{s-1}} = \frac{a_v}{a_{u-1}} \tag{40}$$

(provided that $a_{s-1} \neq 0$ for all $s \in \{u, u+1, \ldots, v\}$).

Examples:

  - Let n be a positive integer. Then,

$$\prod_{s=2}^{n} \underbrace{\left(1 - \frac{1}{s}\right)}_{= \frac{s-1}{s} = \frac{1/s}{1/(s-1)}} = \prod_{s=2}^{n} \frac{1/s}{1/(s-1)} = \frac{1/n}{1/(2-1)}$$

$$\left(\text{by (40), applied to } u = 2,\ v = n \text{ and } a_s = 1/s\right)$$

$$= \frac{1}{n}.$$

- **Restricting to a subset:** Let S be a finite set. Let T be a subset of S. Let a_s be an element of $\mathbb{A}$ for each $s \in T$. Then,

$$\prod_{\substack{s \in S; \\ s \in T}} a_s = \prod_{s \in T} a_s.$$

Remark: Here is a slightly more general form of this rule: Let S be a finite set. Let T be a subset of S. Let $\mathcal{A}(s)$ be a logical statement for each $s \in S$. Let a_s be an element of $\mathbb{A}$ for each $s \in T$ satisfying $\mathcal{A}(s)$. Then,

$$\prod_{\substack{s \in S; \\ s \in T; \\ \mathcal{A}(s)}} a_s = \prod_{\substack{s \in T; \\ \mathcal{A}(s)}} a_s.$$

- **Splitting a product by a value of a function:** Let S be a finite set. Let W be a set. Let $f : S \to W$ be a map. Let a_s be an element of $\mathbb{A}$ for each $s \in S$. Then,
$$\prod_{s\in S} a_s = \prod_{w\in W} \prod_{\substack{s\in S;\\ f(s)=w}} a_s.$$
(The right hand side is to be read as $\prod\limits_{w\in W} \left(\prod\limits_{\substack{s\in S;\\ f(s)=w}} a_s \right)$.)

- **Splitting a product into subproducts:** Let S be a finite set. Let $S_1, S_2, \ldots, S_n$ be finitely many subsets of S. Assume that these subsets $S_1, S_2, \ldots, S_n$ are pairwise disjoint (i.e., we have $S_i \cap S_j = \varnothing$ for any two distinct elements i and j of $\{1, 2, \ldots, n\}$) and their union is S. (Thus, every element of S lies in precisely one of the subsets $S_1, S_2, \ldots, S_n$.) Let a_s be an element of $\mathbb{A}$ for each $s \in S$. Then,
$$\prod_{s\in S} a_s = \prod_{w=1}^{n} \prod_{s\in S_w} a_s.$$

- **Fubini's theorem (interchanging the order of multiplication):** Let X and Y be two finite sets. Let $a_{(x,y)}$ be an element of $\mathbb{A}$ for each $(x, y) \in X \times Y$. Then,
$$\prod_{x\in X} \prod_{y\in Y} a_{(x,y)} = \prod_{(x,y)\in X\times Y} a_{(x,y)} = \prod_{y\in Y} \prod_{x\in X} a_{(x,y)}.$$
In particular, if n and m are two elements of $\mathbb{N}$, and if $a_{(x,y)}$ is an element of $\mathbb{A}$ for each $(x, y) \in \{1, 2, \ldots, n\} \times \{1, 2, \ldots, m\}$, then
$$\prod_{x=1}^{n} \prod_{y=1}^{m} a_{(x,y)} = \prod_{(x,y)\in\{1,2,\ldots,n\}\times\{1,2,\ldots,m\}} a_{(x,y)} = \prod_{y=1}^{m} \prod_{x=1}^{n} a_{(x,y)}.$$

- **Triangular Fubini's theorem I:** Let $n \in \mathbb{N}$. Let T_n be the set $\left\{ (x, y) \in \{1, 2, 3, \ldots\}^2 \mid x + y \leq n \right\}$. Let $a_{(x,y)}$ be an element of $\mathbb{A}$ for each $(x, y) \in T_n$. Then,
$$\prod_{x=1}^{n} \prod_{y=1}^{n-x} a_{(x,y)} = \prod_{(x,y)\in T_n} a_{(x,y)} = \prod_{y=1}^{n} \prod_{x=1}^{n-y} a_{(x,y)}.$$

- **Triangular Fubini's theorem II:** Let $n \in \mathbb{N}$. Let Q_n be the set $\left\{ (x, y) \in \{1, 2, \ldots, n\}^2 \mid x \leq y \right\}$. Let $a_{(x,y)}$ be an element of $\mathbb{A}$ for each $(x, y) \in Q_n$. Then,
$$\prod_{x=1}^{n} \prod_{y=x}^{n} a_{(x,y)} = \prod_{(x,y)\in Q_n} a_{(x,y)} = \prod_{y=1}^{n} \prod_{x=1}^{y} a_{(x,y)}.$$

- **Fubini's theorem with a predicate:** Let X and Y be two finite sets. For every pair $(x,y) \in X \times Y$, let $\mathcal{A}(x,y)$ be a logical statement. For each $(x,y) \in X \times Y$ satisfying $\mathcal{A}(x,y)$, let $a_{(x,y)}$ be an element of $\mathbb{A}$. Then,

$$\prod_{x \in X} \prod_{\substack{y \in Y; \\ \mathcal{A}(x,y)}} a_{(x,y)} = \prod_{\substack{(x,y) \in X \times Y; \\ \mathcal{A}(x,y)}} a_{(x,y)} = \prod_{y \in Y} \prod_{\substack{x \in X; \\ \mathcal{A}(x,y)}} a_{(x,y)}.$$

- **Interchange of predicates:** Let S be a finite set. For every $s \in S$, let $\mathcal{A}(s)$ and $\mathcal{B}(s)$ be two equivalent logical statements. ("Equivalent" means that $\mathcal{A}(s)$ holds if and only if $\mathcal{B}(s)$ holds.) Let a_s be an element of $\mathbb{A}$ for each $s \in S$. Then,

$$\prod_{\substack{s \in S; \\ \mathcal{A}(s)}} a_s = \prod_{\substack{s \in S; \\ \mathcal{B}(s)}} a_s.$$

- **Substituting the index I with a predicate:** Let S and T be two finite sets. Let $f : S \to T$ be a **bijective** map. Let a_t be an element of $\mathbb{A}$ for each $t \in T$. For every $t \in T$, let $\mathcal{A}(t)$ be a logical statement. Then,

$$\prod_{\substack{t \in T; \\ \mathcal{A}(t)}} a_t = \prod_{\substack{s \in S; \\ \mathcal{A}(f(s))}} a_{f(s)}.$$

1.5. Polynomials: a precise definition

As I have already mentioned in the above list of prerequisites, the notion of polynomials (in one and in several indeterminates) will be occasionally used in these notes. Most likely, the reader already has at least a vague understanding of this notion (e.g., from high school); this vague understanding is probably sufficient for reading most of these notes. But polynomials are one of the most important notions in algebra (if not to say in mathematics), and the reader will likely encounter them over and over; sooner or later, it will happen that the vague understanding is not sufficient and some subtleties do matter. For that reason, anyone serious about doing abstract algebra should know a complete and correct definition of polynomials and have some experience working with it. I shall not give a complete definition of the most general notion of polynomials in these notes, but I will comment on some of the subtleties and define an important special case (that of polynomials in one variable with rational coefficients) in the present section. A reader is probably best advised to skip this section on their first read.

It is not easy to find a good (formal and sufficiently general) treatment of polynomials in textbooks. Various authors tend to skimp on subtleties and technical points such as the notion of an "indeterminate", or the precise meaning of "formal expression" in the slogan "a polynomial is a formal expression" (the best texts do not use this vague slogan at all), or the definition of the degree of the zero polynomial, or the difference between regarding polynomials as sequences (which is the classical viewpoint and particularly useful for polynomials in one variable) and regarding polynomials as elements of a monoid ring (which is important in the case of several variables, since it allows us to regard the polynomial rings $\mathbb{Q}[X]$ and $\mathbb{Q}[Y]$ as two distinct subrings of $\mathbb{Q}[X, Y]$). They also tend to take some questionable shortcuts, such as defining polynomials in n variables (by induction over n) as one-variable polynomials over the ring of $(n-1)$-variable polynomials (this shortcut has several shortcomings, such as making the symmetric role of the n variables opaque, and functioning only for finitely many variables).

More often than not, the polynomials we will be using will be polynomials in one variable. These are usually handled well in good books on abstract algebra – e.g., in [Walker87, §4.5], in [Hunger14, Appendix G], in [Hunger03, Chapter III, §5], in [Rotman15, Chapter A-3], in [HofKun71, §4.1, §4.2] (although in [HofKun71, §4.1, §4.2], only polynomials over fields are studied, but the definition applies to commutative rings mutatis mutandis), in [AmaEsc05, §8], and in [BirMac99, Chapter III, §6]. Most of these treatments rely on the notion of a *commutative ring*, which is not difficult but somewhat abstract (I shall introduce it below in Section 6.1).

Let me give a brief survey of the notion of univariate polynomials (i.e., polynomials in one variable). I shall define them as sequences. For the sake of simplicity, I shall only talk of polynomials with rational coefficients. Similarly, one can define polynomials with integer coefficients, with real coefficients, or with complex coefficients; of course, one then has to replace each "$\mathbb{Q}$" by a "$\mathbb{Z}$", an "$\mathbb{R}$" or a "$\mathbb{C}$".

The rough idea behind the definition of a polynomial is that a polynomial with rational coefficients should be a "formal expression" which is built out of rational numbers, an "indeterminate" X as well as addition, subtraction and multiplication signs, such as $X^4 - 27X + \dfrac{3}{2}$ or $-X^3 + 2X + 1$ or $\dfrac{1}{3}(X-3) \cdot X^2$ or $X^4 + 7X^3(X-2)$ or -15. We have not explicitly allowed powers, but we understand X^n to mean the product $\underbrace{XX \cdots X}_{n \text{ times}}$ (which is 1 when $n = 0$). Notice that division is not allowed, so we cannot get $\dfrac{X}{X+1}$ (but we can get $\dfrac{3}{2}X$, because $\dfrac{3}{2}$ is a rational number). Notice also that a polynomial can be a single rational number, since we never said that X must necessarily be used; for instance, -15 and 0 are polynomials.

This is, of course, not a valid definition. One problem with it that it does not explain what a "formal expression" is. For starters, we want an expression that is well-defined – i.e., into that we can substitute a rational number for X and obtain a valid term. For example, $X - + \cdot 5$ is not well-defined, so it does not fit our bill; neither is the "empty expression". Furthermore, when do we want two "formal expressions" to be viewed as one and the same polynomial? Do we want to equate $X(X+2)$ with $X^2 + 2X$? Do we want to equate $0X^3 + 2X + 1$ with $2X + 1$? The answer is "yes" both times, but a general rule is not easy to give if we keep talking of "formal expressions".

We *could* define two polynomials $p(X)$ and $q(X)$ to be equal if and only if, for every number $\alpha \in \mathbb{Q}$, the values $p(\alpha)$ and $q(\alpha)$ (obtained by substituting α for

X in p and in q, respectively) are equal. This would be tantamount to treating polynomials as *functions*: it would mean that we identify a polynomial $p\left(X\right)$ with the function $\mathbb{Q} \to \mathbb{Q}$, $\alpha \mapsto p\left(\alpha\right)$. Such a definition would work well as long as we would do only rather basic things with it[25], but as soon as we would try to go deeper, we would encounter technical issues which would make it inadequate and painful[26]. Also, if we equated polynomials with the functions they describe, then we would waste the word "polynomial" on a concept (a function described by a polynomial) that already has a word for it (namely, *polynomial function*).

[25] And some authors, such as Axler in [Axler15, Chapter 4], do use this definition.

[26] Here are three of these issues:

- One of the strengths of polynomials is that we can evaluate them not only at numbers, but also at many other things, e.g., at square matrices: Evaluating the polynomial $X^2 - 3X$ at the square matrix $\begin{pmatrix} 1 & 3 \\ -1 & 2 \end{pmatrix}$ gives $\begin{pmatrix} 1 & 3 \\ -1 & 2 \end{pmatrix}^2 - 3\begin{pmatrix} 1 & 3 \\ -1 & 2 \end{pmatrix} = \begin{pmatrix} -5 & 0 \\ 0 & -5 \end{pmatrix}$. However, a function must have a well-defined domain, and does not make sense outside of this domain. So, if the polynomial $X^2 - 3X$ is regarded as the function $\mathbb{Q} \to \mathbb{Q}$, $\alpha \mapsto \alpha^2 - 3\alpha$, then it makes no sense to evaluate this polynomial at the matrix $\begin{pmatrix} 1 & 3 \\ -1 & 2 \end{pmatrix}$, just because this matrix does not lie in the domain $\mathbb{Q}$ of the function. We could, of course, extend the domain of the function to (say) the set of square matrices over $\mathbb{Q}$, but then we would still have the same problem with other things that we want to evaluate polynomials at. At some point we want to be able to evaluate polynomials at functions and at other polynomials, and if we would try to achieve this by extending the domain, we would have to do this over and over, because each time we extend the domain, we get even more polynomials to evaluate our polynomials at; thus, the definition would be eternally "hunting its own tail"! (We could resolve this difficulty by defining polynomials as *natural transformations* in the sense of category theory. I do not want to even go into this definition here, as it would take several pages to properly introduce. At this point, it is not worth the hassle.)
- Let $p\left(X\right)$ be a polynomial with real coefficients. Then, it should be obvious that $p\left(X\right)$ can also be viewed as a polynomial with complex coefficients: For instance, if $p\left(X\right)$ was defined as $3X + \frac{7}{2}X\left(X-1\right)$, then we can view the numbers 3, $\frac{7}{2}$ and -1 appearing in its definition as complex numbers, and thus get a polynomial with complex coefficients. But wait! What if two polynomials $p\left(X\right)$ and $q\left(X\right)$ are equal when viewed as polynomials with real coefficients, but become distinct when viewed as polynomials with complex coefficients (because when we view them as polynomials with complex coefficients, their domains grow larger to include complex numbers, and a new complex α might perhaps no longer satisfy $p\left(\alpha\right) = q\left(\alpha\right)$)? This does not actually happen, but ruling this out is not obvious if you regard polynomials as functions.
- (This requires some familiarity with finite fields:) Treating polynomials as functions works reasonably well for polynomials with integer, rational, real and complex coefficients (as long as one is not too demanding). But we will eventually want to consider polynomials with coefficients in any arbitrary commutative ring $\mathbb{K}$. An example for a commutative ring $\mathbb{K}$ is the finite field $\mathbb{F}_p$ with p elements, where p is a prime. (This finite field $\mathbb{F}_p$ is better known as the ring of integers modulo p.) If we define polynomials with coefficients in $\mathbb{F}_p$ as functions $\mathbb{F}_p \to \mathbb{F}_p$, then we really run into problems; for example, the polynomials X and X^p over this field become identical as functions!

The preceding paragraphs indicate that it is worth defining "polynomials" in a way that, on the one hand, conveys the idea that they are more "formal expressions" than "functions", but on the other hand, is less nebulous than "formal expression". Here is one such definition:

Definition 1.7. **(a)** A *univariate polynomial with rational coefficients* means a sequence $(p_0, p_1, p_2, \ldots) \in \mathbb{Q}^\infty$ of elements of $\mathbb{Q}$ such that

$$\text{all but finitely many } k \in \mathbb{N} \text{ satisfy } p_k = 0. \tag{41}$$

Here, the phrase "all but finitely many $k \in \mathbb{N}$ satisfy $p_k = 0$" means "there exists some finite subset J of $\mathbb{N}$ such that every $k \in \mathbb{N} \setminus J$ satisfies $p_k = 0$". (See Definition 5.17 for the general definition of "all but finitely many", and Section 5.4 for some practice with this concept.) More concretely, the condition (41) can be rewritten as follows: The sequence $(p_0, p_1, p_2, \ldots)$ contains only zeroes from some point on (i.e., there exists some $N \in \mathbb{N}$ such that $p_N = p_{N+1} = p_{N+2} = \cdots = 0$).

For the remainder of this definition, "univariate polynomial with rational coefficients" will be abbreviated as "polynomial".

For example, the sequences $(0, 0, 0, \ldots)$, $(1, 3, 5, 0, 0, 0, \ldots)$, $\left(4, 0, -\frac{2}{3}, 5, 0, 0, 0, \ldots\right)$, $\left(0, -1, \frac{1}{2}, 0, 0, 0, \ldots\right)$ (where the "$\ldots$" stand for infinitely many zeroes) are polynomials, but the sequence $(1, 1, 1, \ldots)$ (where the "$\ldots$" stands for infinitely many 1's) is not (since it does not satisfy (41)).

So we have defined a polynomial as an infinite sequence of rational numbers with a certain property. So far, this does not seem to reflect any intuition of polynomials as "formal expressions". However, we shall soon (namely, in Definition 1.7 **(j)**) identify the polynomial $(p_0, p_1, p_2, \ldots) \in \mathbb{Q}^\infty$ with the "formal expression" $p_0 + p_1 X + p_2 X^2 + \cdots$ (this is an infinite sum, but due to (41) all but its first few terms are 0 and thus can be neglected). For instance, the polynomial $(1, 3, 5, 0, 0, 0, \ldots)$ will be identified with the "formal expression" $1 + 3X + 5X^2 + 0X^3 + 0X^4 + 0X^5 + \cdots = 1 + 3X + 5X^2$. Of course, we cannot do this identification right now, since we do not have a reasonable definition of X.

(b) We let $\mathbb{Q}[X]$ denote the set of all univariate polynomials with rational coefficients. Given a polynomial $p = (p_0, p_1, p_2, \ldots) \in \mathbb{Q}[X]$, we denote the numbers $p_0, p_1, p_2, \ldots$ as the *coefficients* of p. More precisely, for every $i \in \mathbb{N}$, we shall refer to p_i as the *i-th coefficient* of p. (Do not forget that we are counting from 0 here: any polynomial "begins" with its 0-th coefficient.) The 0-th coefficient of p is also known as the *constant term* of p.

Instead of "the i-th coefficient of p", we often also say "the *coefficient before X^i of p*" or "the *coefficient of X^i in p*".

Thus, any polynomial $p \in \mathbb{Q}[X]$ is the sequence of its coefficients.

(c) We denote the polynomial $(0, 0, 0, \ldots) \in \mathbb{Q}[X]$ by $\mathbf{0}$. We will also write 0 for it when no confusion with the number 0 is possible. The polynomial $\mathbf{0}$ is called the *zero polynomial*. A polynomial $p \in \mathbb{Q}[X]$ is said to be *nonzero* if $p \neq \mathbf{0}$.

(d) We denote the polynomial $(1,0,0,0,\ldots) \in \mathbb{Q}[X]$ by $\mathbf{1}$. We will also write 1 for it when no confusion with the number 1 is possible.

(e) For any $\lambda \in \mathbb{Q}$, we denote the polynomial $(\lambda,0,0,0,\ldots) \in \mathbb{Q}[X]$ by $\operatorname{const}\lambda$. We call it the *constant polynomial with value* λ. It is often useful to identify $\lambda \in \mathbb{Q}$ with $\operatorname{const}\lambda \in \mathbb{Q}[X]$. Notice that $\mathbf{0} = \operatorname{const} 0$ and $\mathbf{1} = \operatorname{const} 1$.

(f) Now, let us define the sum, the difference and the product of two polynomials. Indeed, let $a = (a_0, a_1, a_2, \ldots) \in \mathbb{Q}[X]$ and $b = (b_0, b_1, b_2, \ldots) \in \mathbb{Q}[X]$ be two polynomials. Then, we define three polynomials $a+b$, $a-b$ and $a \cdot b$ in $\mathbb{Q}[X]$ by

$$\begin{aligned} a+b &= (a_0+b_0, a_1+b_1, a_2+b_2, \ldots); \\ a-b &= (a_0-b_0, a_1-b_1, a_2-b_2, \ldots); \\ a \cdot b &= (c_0, c_1, c_2, \ldots), \end{aligned}$$

where

$$c_k = \sum_{i=0}^{k} a_i b_{k-i} \qquad \text{for every } k \in \mathbb{N}.$$

We call $a+b$ the *sum* of a and b; we call $a-b$ the *difference* of a and b; we call $a \cdot b$ the *product* of a and b. We abbreviate $a \cdot b$ by ab, and we abbreviate $\mathbf{0} - a$ by $-a$.

For example,

$$\begin{aligned} (1,2,2,0,0,\ldots) + (3,0,-1,0,0,0,\ldots) &= (4,2,1,0,0,0,\ldots); \\ (1,2,2,0,0,\ldots) - (3,0,-1,0,0,0,\ldots) &= (-2,2,3,0,0,0,\ldots); \\ (1,2,2,0,0,\ldots) \cdot (3,0,-1,0,0,0,\ldots) &= (3,6,5,-2,-2,0,0,0,\ldots). \end{aligned}$$

The definition of $a+b$ essentially says that "polynomials are added coefficientwise" (i.e., in order to obtain the sum of two polynomials a and b, it suffices to add each coefficient of a to the corresponding coefficient of b). Similarly, the definition of $a-b$ says the same thing about subtraction. The definition of $a \cdot b$ is more surprising. However, it loses its mystique when we identify the polynomials a and b with the "formal expressions" $a_0 + a_1X + a_2X^2 + \cdots$ and $b_0 + b_1X + b_2X^2 + \cdots$ (although, at this point, we do not know what these expressions really mean); indeed, it simply says that

$$\left(a_0 + a_1X + a_2X^2 + \cdots\right)\left(b_0 + b_1X + b_2X^2 + \cdots\right) = c_0 + c_1X + c_2X^2 + \cdots,$$

where $c_k = \sum\limits_{i=0}^{k} a_i b_{k-i}$ for every $k \in \mathbb{N}$. This is precisely what one would expect, because if you expand $\left(a_0 + a_1X + a_2X^2 + \cdots\right)\left(b_0 + b_1X + b_2X^2 + \cdots\right)$ using the distributive law and collect equal powers of X, then you get precisely $c_0 + c_1X + c_2X^2 + \cdots$. Thus, the definition of $a \cdot b$ has been tailored to make the distributive law hold.

(By the way, why is $a \cdot b$ a polynomial? That is, why does it satisfy (41) ? The proof is easy, but we omit it.)

Addition, subtraction and multiplication of polynomials satisfy some of the same rules as addition, subtraction and multiplication of numbers. For example, the commutative laws $a+b=b+a$ and $ab=ba$ are valid for polynomials just as they are for numbers; the same holds for the associative laws $(a+b)+c=a+(b+c)$ and $(ab)c=a(bc)$ and the distributive laws $(a+b)c=ac+bc$ and $a(b+c)=ab+ac$. Moreover, each polynomial a satisfies $a+\mathbf{0}=\mathbf{0}+a=a$ and $a\cdot\mathbf{0}=\mathbf{0}\cdot a=\mathbf{0}$ and $a\cdot\mathbf{1}=\mathbf{1}\cdot a=a$ and $a+(-a)=(-a)+a=\mathbf{0}$.

Using the notations of Definition 6.2, we can summarize this as follows: The set $\mathbb{Q}[X]$, endowed with the operations $+$ and $\cdot$ just defined, and with the elements $\mathbf{0}$ and $\mathbf{1}$, is a commutative ring. It is called the *(univariate) polynomial ring over* $\mathbb{Q}$.

(g) Let $a=(a_0,a_1,a_2,\ldots)\in\mathbb{Q}[X]$ and $\lambda\in\mathbb{Q}$. Then, λa denotes the polynomial $(\lambda a_0,\lambda a_1,\lambda a_2,\ldots)\in\mathbb{Q}[X]$. (This equals the polynomial $(\operatorname{const}\lambda)\cdot a$; thus, identifying λ with $\operatorname{const}\lambda$ does not cause any inconsistencies here.)

(h) If $p=(p_0,p_1,p_2,\ldots)\in\mathbb{Q}[X]$ is a nonzero polynomial, then the *degree* of p is defined to be the maximum $i\in\mathbb{N}$ satisfying $p_i\neq 0$. If $p\in\mathbb{Q}[X]$ is the zero polynomial, then the degree of p is defined to be $-\infty$. (Here, $-\infty$ is just a fancy symbol, not a number.) The degree of a polynomial $p\in\mathbb{Q}[X]$ is denoted $\deg p$. For example, $\deg(0,4,0,-1,0,0,0,\ldots)=3$.

(i) If $a=(a_0,a_1,a_2,\ldots)\in\mathbb{Q}[X]$ and $n\in\mathbb{N}$, then a polynomial $a^n\in\mathbb{Q}[X]$ is defined to be the product $\underbrace{aa\cdots a}_{n\text{ times}}$. (This is understood to be $\mathbf{1}$ when $n=0$. In general, an empty product of polynomials is always understood to be $\mathbf{1}$.)

(j) We let X denote the polynomial $(0,1,0,0,0,\ldots)\in\mathbb{Q}[X]$. (This is the polynomial whose 1-st coefficient is 1 and whose other coefficients are 0.) This polynomial is called the *indeterminate* of $\mathbb{Q}[X]$. It is easy to see that, for any $n\in\mathbb{N}$, we have

$$X^n=\left(\underbrace{0,0,\ldots,0}_{n\text{ zeroes}},1,0,0,0,\ldots\right).$$

This polynomial X finally provides an answer to the questions "what is an indeterminate" and "what is a formal expression". Namely, let $(p_0,p_1,p_2,\ldots)\in\mathbb{Q}[X]$ be any polynomial. Then, the sum $p_0+p_1X+p_2X^2+\cdots$ is well-defined (it is an infinite sum, but due to (41) it has only finitely many nonzero addends), and it is easy to see that this sum equals $(p_0,p_1,p_2,\ldots)$. Thus,

$$(p_0,p_1,p_2,\ldots)=p_0+p_1X+p_2X^2+\cdots\qquad\text{for every }(p_0,p_1,p_2,\ldots)\in\mathbb{Q}[X].$$

This finally allows us to write a polynomial $(p_0,p_1,p_2,\ldots)$ as a sum $p_0+p_1X+p_2X^2+\cdots$ while remaining honest; the sum $p_0+p_1X+p_2X^2+\cdots$ is no longer a "formal expression" of unclear meaning, nor a function, but it is just an alternative way to write the sequence $(p_0,p_1,p_2,\ldots)$. So, at last, our notion of a polynomial resembles the intuitive notion of a polynomial!

Of course, we can write polynomials as finite sums as well. Indeed, if $(p_0,p_1,p_2,\ldots)\in\mathbb{Q}[X]$ is a polynomial and N is a nonnegative integer such that every $n > N$ satisfies $p_n = 0$, then

$$(p_0, p_1, p_2, \ldots) = p_0 + p_1 X + p_2 X^2 + \cdots = p_0 + p_1 X + \cdots + p_N X^N$$

(because addends can be discarded when they are 0). For example,

$$\begin{aligned} (4, 1, 0, 0, 0, \ldots) &= 4 + 1X = 4 + X \qquad \text{and} \\ \left(\frac{1}{2}, 0, \frac{1}{3}, 0, 0, 0, \ldots\right) &= \frac{1}{2} + 0X + \frac{1}{3}X^2 = \frac{1}{2} + \frac{1}{3}X^2. \end{aligned}$$

(k) For our definition of polynomials to be fully compatible with our intuition, we are missing only one more thing: a way to evaluate a polynomial at a number, or some other object (e.g., another polynomial or a function). This is easy: Let $p = (p_0, p_1, p_2, \ldots) \in \mathbb{Q}[X]$ be a polynomial, and let $\alpha \in \mathbb{Q}$. Then, $p(\alpha)$ means the number $p_0 + p_1\alpha + p_2\alpha^2 + \cdots \in \mathbb{Q}$. (Again, the infinite sum $p_0 + p_1\alpha + p_2\alpha^2 + \cdots$ makes sense because of (41).) Similarly, we can define $p(\alpha)$ when $\alpha \in \mathbb{R}$ (but in this case, $p(\alpha)$ will be an element of $\mathbb{R}$) or when $\alpha \in \mathbb{C}$ (in this case, $p(\alpha) \in \mathbb{C}$) or when α is a square matrix with rational entries (in this case, $p(\alpha)$ will also be such a matrix) or when α is another polynomial (in this case, $p(\alpha)$ is such a polynomial as well).

For example, if $p = (1, -2, 0, 3, 0, 0, 0, \ldots) = 1 - 2X + 3X^3$, then $p(\alpha) = 1 - 2\alpha + 3\alpha^3$ for every α.

The map $\mathbb{Q} \to \mathbb{Q}$, $\alpha \mapsto p(\alpha)$ is called the *polynomial function described by p*. As we said above, this function is not p, and it is not a good idea to equate it with p.

If α is a number (or a square matrix, or another polynomial), then $p(\alpha)$ is called the result of *evaluating p at $X = \alpha$* (or, simply, evaluating p at α), or the result of *substituting α for X in p*. This notation, of course, reminds of functions; nevertheless, (as we already said a few times) p is **not a function**.

Probably the simplest three cases of evaluation are the following ones:

- We have $p(0) = p_0 + p_1 0^1 + p_2 0^2 + \cdots = p_0$. In other words, evaluating p at $X = 0$ yields the constant term of p.

- We have $p(1) = p_0 + p_1 1^1 + p_2 1^2 + \cdots = p_0 + p_1 + p_2 + \cdots$. In other words, evaluating p at $X = 1$ yields the sum of all coefficients of p.

- We have $p(X) = p_0 + p_1 X^1 + p_2 X^2 + \cdots = p_0 + p_1 X + p_2 X^2 + \cdots = p$. In other words, evaluating p at $X = X$ yields p itself. This allows us to write $p(X)$ for p. Many authors do so, just in order to stress that p is a polynomial and that the indeterminate is called X. It should be kept in mind that X is **not a variable** (just as p is **not a function**); it is the (fixed!) sequence $(0, 1, 0, 0, 0, \ldots) \in \mathbb{Q}[X]$ which serves as the indeterminate for polynomials in $\mathbb{Q}[X]$.

(l) Often, one wants (or is required) to give an indeterminate a name other than X. (For instance, instead of polynomials with rational coefficients, we could be considering polynomials whose coefficients themselves are polynomials in $\mathbb{Q}[X]$; and then, we would not be allowed to use the letter X for the "new" indeterminate anymore, as it already means the indeterminate of $\mathbb{Q}[X]$!) This can be done, and the rules are the following: Any letter (that does not already have a meaning) can be used to denote the indeterminate; but then, the set of all polynomials has to be renamed as $\mathbb{Q}[\eta]$, where η is this letter. For instance, if we want to denote the indeterminate as x, then we have to denote the set by $\mathbb{Q}[x]$.

It is furthermore convenient to regard the sets $\mathbb{Q}[\eta]$ for different letters η as distinct. Thus, for example, the polynomial $3X^2 + 1$ is not the same as the polynomial $3Y^2 + 1$. (The reason for doing so is that one sometimes wishes to view both of these polynomials as polynomials in the two variables X and Y.) Formally speaking, this means that we should define a polynomial in $\mathbb{Q}[\eta]$ to be not just a sequence $(p_0, p_1, p_2, \ldots)$ of rational numbers, but actually a pair $((p_0, p_1, p_2, \ldots), \text{"}\eta\text{"})$ of a sequence of rational numbers and the letter η. (Here, "η" really means the letter η, not the sequence $(0, 1, 0, 0, 0, \ldots)$.) This is, of course, a very technical point which is of little relevance to most of mathematics; it becomes important when one tries to implement polynomials in a programming language.

(m) As already explained, we can replace $\mathbb{Q}$ by $\mathbb{Z}$, $\mathbb{R}$, $\mathbb{C}$ or any other commutative ring $\mathbb{K}$ in the above definition. (See Definition 6.2 for the definition of a commutative ring.) When $\mathbb{Q}$ is replaced by a commutative ring $\mathbb{K}$, the notion of "univariate polynomials with rational coefficients" becomes "univariate polynomials with coefficients in $\mathbb{K}$" (also known as "univariate polynomials over $\mathbb{K}$"), and the set of such polynomials is denoted by $\mathbb{K}[X]$ rather than $\mathbb{Q}[X]$.

So much for univariate polynomials.

Polynomials in multiple variables are (in my opinion) treated the best in [Lang02, Chapter II, §3], where they are introduced as elements of a monoid ring. However, this treatment is rather abstract and uses a good deal of algebraic language[27]. The treatments in [Walker87, §4.5], in [Rotman15, Chapter A-3] and in [BirMac99, Chapter IV, §4] use the above-mentioned recursive shortcut that makes them inferior (in my opinion). A neat (and rather elementary) treatment of polynomials in n variables (for finite n) can be found in [Hunger03, Chapter III, §5], in [Loehr11, §7.16], in [GalQua22, §30.2], in [ZarSam67, §18] and in [AmaEsc05, §I.8]; it generalizes the viewpoint we used in Definition 1.7 for univariate polynomials above[28].

[27] Also, the book [Lang02] is notorious for its unpolished writing; it is best read with Bergman's companion [Bergma15] at hand.

[28] You are reading right: The analysis textbook [AmaEsc05] is one of the few sources I am aware of to define the (algebraic!) notion of polynomials precisely and well.

2. A closer look at induction

In this chapter, we shall recall several versions of the *induction principle* (the principle of mathematical induction) and provide examples for their use. We assume that the reader is at least somewhat familiar with mathematical induction[29]; we shall present some nonstandard examples of its use (including a proof of the legitimacy of the definition of a sum $\sum\limits_{s\in S} a_s$ given in Section 1.4).

2.1. Standard induction

2.1.1. The Principle of Mathematical Induction

We first recall the classical principle of mathematical induction[30]:

> **Theorem 2.1.** For each $n \in \mathbb{N}$, let $\mathcal{A}(n)$ be a logical statement.
>
> Assume the following:
>
> *Assumption 1:* The statement $\mathcal{A}(0)$ holds.
>
> *Assumption 2:* If $m \in \mathbb{N}$ is such that $\mathcal{A}(m)$ holds, then $\mathcal{A}(m+1)$ also holds.
>
> Then, $\mathcal{A}(n)$ holds for each $n \in \mathbb{N}$.

Theorem 2.1 is commonly taken to be one of the axioms of mathematics (the "axiom of induction"), or (in type theory) as part of the definition of $\mathbb{N}$. Intuitively, Theorem 2.1 should be obvious: For example, if you want to prove (under the assumptions of Theorem 2.1) that $\mathcal{A}(4)$ holds, you can argue as follows:

- By Assumption 1, the statement $\mathcal{A}(0)$ holds.
- Thus, by Assumption 2 (applied to $m = 0$), the statement $\mathcal{A}(1)$ holds.
- Thus, by Assumption 2 (applied to $m = 1$), the statement $\mathcal{A}(2)$ holds.
- Thus, by Assumption 2 (applied to $m = 2$), the statement $\mathcal{A}(3)$ holds.
- Thus, by Assumption 2 (applied to $m = 3$), the statement $\mathcal{A}(4)$ holds.

A similar (but longer) argument shows that the statement $\mathcal{A}(5)$ holds. Likewise, you can show that the statement $\mathcal{A}(15)$ holds, if you have the patience to apply Assumption 2 a total of 15 times. It is thus not surprising that $\mathcal{A}(n)$ holds for each $n \in \mathbb{N}$; but if you don't assume Theorem 2.1 as an axiom, you would need to write

[29] If not, introductions can be found in [LeLeMe16, Chapter 5], [Day16], [Vellem06, Chapter 6], [Hammac15, Chapter 10], [Vorobi02] and various other sources.

[30] Keep in mind that $\mathbb{N}$ means the set $\{0, 1, 2, \ldots\}$ for us.

down a different proof for each value of n (which becomes the longer the larger n is), and thus would never reach the general result (i.e., that $\mathcal{A}(n)$ holds for **each** $n \in \mathbb{N}$), because you cannot write down infinitely many proofs. What Theorem 2.1 does is, roughly speaking, to apply Assumption 2 for you as many times as it is needed for each $n \in \mathbb{N}$.

(Authors of textbooks like to visualize Theorem 2.1 by envisioning an infinite sequence of dominos (numbered $0, 1, 2, \ldots$) placed in row, sufficiently close to each other that if domino m falls, then domino $m+1$ will also fall. Now, assume that you kick domino 0 over. What Theorem 2.1 then says is that each domino will fall. See, e.g., [Hammac15, Chapter 10] for a detailed explanation of this metaphor. Here is another metaphor for Theorem 2.1: Assume that there is a virus that infects nonnegative integers. Once it has infected some $m \in \mathbb{N}$, it will soon spread to $m+1$ as well. Now, assume that 0 gets infected. Then, Theorem 2.1 says that each $n \in \mathbb{N}$ will eventually be infected.)

Theorem 2.1 is called the *principle of induction* or *principle of complete induction* or *principle of mathematical induction*, and we shall also call it *principle of standard induction* in order to distinguish it from several variant "principles of induction" that we will see later. Proofs that use this principle are called *proofs by induction* or *induction proofs*. Usually, in such proofs, we don't explicitly cite Theorem 2.1, but instead say certain words that signal that Theorem 2.1 is being applied and that (ideally) also indicate what statements $\mathcal{A}(n)$ it is being applied to[31]. However, for our very first example of a proof by induction, we are going to use Theorem 2.1 explicitly. We shall show the following fact:

Proposition 2.2. Let q and d be two real numbers such that $q \neq 1$. Let $(a_0, a_1, a_2, \ldots)$ be a sequence of real numbers. Assume that

$$a_{n+1} = qa_n + d \qquad \text{for each } n \in \mathbb{N}. \tag{42}$$

Then,

$$a_n = q^n a_0 + \frac{q^n - 1}{q - 1} d \qquad \text{for each } n \in \mathbb{N}. \tag{43}$$

Proof of Proposition 2.2. For each $n \in \mathbb{N}$, we let $\mathcal{A}(n)$ be the statement $\left(a_n = q^n a_0 + \dfrac{q^n - 1}{q - 1} d\right)$. Thus, our goal is to prove the statement $\mathcal{A}(n)$ for each $n \in \mathbb{N}$.

We first notice that the statement $\mathcal{A}(0)$ holds[32].

Now, we claim that

$$\text{if } m \in \mathbb{N} \text{ is such that } \mathcal{A}(m) \text{ holds, then } \mathcal{A}(m+1) \text{ also holds.} \tag{44}$$

[31] We will explain this in Convention 2.3 below.

[32] *Proof.* This is easy to verify: We have $q^0 = 1$, thus $q^0 - 1 = 0$, and therefore $\dfrac{q^0 - 1}{q - 1} = \dfrac{0}{q - 1} = 0$.

[*Proof of (44):* Let $m \in \mathbb{N}$ be such that $\mathcal{A}(m)$ holds. We must show that $\mathcal{A}(m+1)$ also holds.

We have assumed that $\mathcal{A}(m)$ holds. In other words, $a_m = q^m a_0 + \frac{q^m - 1}{q-1} d$ holds[33]. Now, (42) (applied to $n = m$) yields

$$\begin{aligned}
a_{m+1} &= q \underbrace{a_m}_{= q^m a_0 + \frac{q^m-1}{q-1}d} + d = q\left(q^m a_0 + \frac{q^m-1}{q-1}d\right) + d \\
&= \underbrace{qq^m}_{=q^{m+1}} a_0 + \underbrace{q \cdot \frac{q^m-1}{q-1}d + d}_{=\left(q\cdot\frac{q^m-1}{q-1}+1\right)d} \\
&= q^{m+1}a_0 + \underbrace{\left(q \cdot \frac{q^m-1}{q-1} + 1\right)}_{\substack{=\frac{q(q^m-1)+(q-1)}{q-1}=\frac{q^{m+1}-1}{q-1} \\ (\text{since } q(q^m-1)+(q-1)=qq^m-q+q-1=qq^m-1=q^{m+1}-1)}} d \\
&= q^{m+1}a_0 + \frac{q^{m+1}-1}{q-1}d.
\end{aligned}$$

So we have shown that $a_{m+1} = q^{m+1}a_0 + \frac{q^{m+1}-1}{q-1}d$. But this is precisely the statement $\mathcal{A}(m+1)$ [34]. Thus, the statement $\mathcal{A}(m+1)$ holds.

Now, forget that we fixed m. We thus have shown that if $m \in \mathbb{N}$ is such that $\mathcal{A}(m)$ holds, then $\mathcal{A}(m+1)$ also holds. This proves (44).]

Now, both assumptions of Theorem 2.1 are satisfied (indeed, Assumption 1 holds because the statement $\mathcal{A}(0)$ holds, whereas Assumption 2 holds because of (44)). Thus, Theorem 2.1 shows that $\mathcal{A}(n)$ holds for each $n \in \mathbb{N}$. In other

Now,

$$\underbrace{q^0}_{=1} a_0 + \underbrace{\frac{q^0-1}{q-1}}_{=0} d = 1a_0 + 0d = a_0,$$

so that $a_0 = q^0 a_0 + \frac{q^0-1}{q-1}d$. But this is precisely the statement $\mathcal{A}(0)$ (since $\mathcal{A}(0)$ is defined to be the statement $\left(a_0 = q^0 a_0 + \frac{q^0-1}{q-1}d\right)$). Hence, the statement $\mathcal{A}(0)$ holds.

[33] because $\mathcal{A}(m)$ is defined to be the statement $\left(a_m = q^m a_0 + \frac{q^m-1}{q-1}d\right)$

[34] because $\mathcal{A}(m+1)$ is defined to be the statement $\left(a_{m+1} = q^{m+1}a_0 + \frac{q^{m+1}-1}{q-1}d\right)$

words, $a_n = q^n a_0 + \frac{q^n - 1}{q - 1} d$ holds for each $n \in \mathbb{N}$ (since $\mathcal{A}(n)$ is the statement $\left(a_n = q^n a_0 + \frac{q^n - 1}{q - 1} d\right)$). This proves Proposition 2.2. □

2.1.2. Conventions for writing induction proofs

Now, let us introduce some standard language that is commonly used in proofs by induction:

> **Convention 2.3.** For each $n \in \mathbb{N}$, let $\mathcal{A}(n)$ be a logical statement. Assume that you want to prove that $\mathcal{A}(n)$ holds for each $n \in \mathbb{N}$.
>
> Theorem 2.1 offers the following strategy for proving this: First show that Assumption 1 of Theorem 2.1 is satisfied; then, show that Assumption 2 of Theorem 2.1 is satisfied; then, Theorem 2.1 automatically completes your proof.
>
> A proof that follows this strategy is called a *proof by induction on n* (or *proof by induction over n*) or (less precisely) an *inductive proof*. When you follow this strategy, you say that you are *inducting on n* (or *over n*). The proof that Assumption 1 is satisfied is called the *induction base* (or *base case*) of the proof. The proof that Assumption 2 is satisfied is called the *induction step* of the proof.
>
> In order to prove that Assumption 2 is satisfied, you will usually want to fix an $m \in \mathbb{N}$ such that $\mathcal{A}(m)$ holds, and then prove that $\mathcal{A}(m+1)$ holds. In other words, you will usually want to fix $m \in \mathbb{N}$, assume that $\mathcal{A}(m)$ holds, and then prove that $\mathcal{A}(m+1)$ holds. When doing so, it is common to refer to the assumption that $\mathcal{A}(m)$ holds as the *induction hypothesis* (or *induction assumption*).

Using this language, we can rewrite our above proof of Proposition 2.2 as follows:

Proof of Proposition 2.2 (second version). For each $n \in \mathbb{N}$, we let $\mathcal{A}(n)$ be the statement $\left(a_n = q^n a_0 + \frac{q^n - 1}{q - 1} d\right)$. Thus, our goal is to prove the statement $\mathcal{A}(n)$ for each $n \in \mathbb{N}$.

We shall prove this by induction on n:

Induction base: We have $q^0 = 1$, thus $q^0 - 1 = 0$, and therefore $\frac{q^0 - 1}{q - 1} = \frac{0}{q - 1} = 0$. Now,

$$\underbrace{q^0}_{=1} a_0 + \underbrace{\frac{q^0 - 1}{q - 1}}_{=0} d = 1 a_0 + 0 d = a_0,$$

so that $a_0 = q^0 a_0 + \frac{q^0 - 1}{q - 1} d$. But this is precisely the statement $\mathcal{A}(0)$ (since $\mathcal{A}(0)$ is defined to be the statement $\left(a_0 = q^0 a_0 + \frac{q^0 - 1}{q - 1} d\right)$). Hence, the statement $\mathcal{A}(0)$ holds. This completes the induction base.

Induction step: Let $m \in \mathbb{N}$. Assume that $\mathcal{A}(m)$ holds. We must show that $\mathcal{A}(m+1)$ also holds.

We have assumed that $\mathcal{A}(m)$ holds (this is our induction hypothesis). In other words, $a_m = q^m a_0 + \dfrac{q^m - 1}{q-1} d$ holds[35]. Now, (42) (applied to $n = m$) yields

$$\begin{aligned}
a_{m+1} &= q \underbrace{a_m}_{= q^m a_0 + \frac{q^m - 1}{q-1} d} + d = q \left(q^m a_0 + \frac{q^m - 1}{q-1} d \right) + d \\
&= \underbrace{qq^m}_{= q^{m+1}} a_0 + \underbrace{q \cdot \frac{q^m - 1}{q-1} d + d}_{= \left(q \cdot \frac{q^m - 1}{q-1} + 1 \right) d} \\
&= q^{m+1} a_0 + \underbrace{\left(q \cdot \frac{q^m - 1}{q-1} + 1 \right)}_{\substack{= \frac{q(q^m-1)+(q-1)}{q-1} = \frac{q^{m+1}-1}{q-1} \\ (\text{since } q(q^m-1)+(q-1) = qq^m - q + q - 1 = qq^m - 1 = q^{m+1} - 1)}} d \\
&= q^{m+1} a_0 + \frac{q^{m+1} - 1}{q-1} d.
\end{aligned}$$

So we have shown that $a_{m+1} = q^{m+1} a_0 + \dfrac{q^{m+1} - 1}{q-1} d$. But this is precisely the statement $\mathcal{A}(m+1)$ [36]. Thus, the statement $\mathcal{A}(m+1)$ holds.

Now, forget that we fixed m. We thus have shown that if $m \in \mathbb{N}$ is such that $\mathcal{A}(m)$ holds, then $\mathcal{A}(m+1)$ also holds. This completes the induction step.

Thus, we have completed both the induction base and the induction step. Hence, by induction, we conclude that $\mathcal{A}(n)$ holds for each $n \in \mathbb{N}$. This proves Proposition 2.2. □

The proof we just gave still has a lot of "boilerplate" text. For example, we have explicitly defined the statement $\mathcal{A}(n)$, but it is not really necessary, since it is clear what this statement should be (viz., it should be the claim we are proving, without the "for each $n \in \mathbb{N}$" part). Allowing ourselves some imprecision, we could say this statement is simply (43). (This is a bit imprecise, because (43) contains the words "for each $n \in \mathbb{N}$", but it should be clear that we don't mean to include these words, since there can be no "for each $n \in \mathbb{N}$" in the statement $\mathcal{A}(n)$.) Furthermore, we don't need to write the sentence

[35] because $\mathcal{A}(m)$ is defined to be the statement $\left(a_m = q^m a_0 + \dfrac{q^m - 1}{q-1} d \right)$

[36] because $\mathcal{A}(m+1)$ is defined to be the statement $\left(a_{m+1} = q^{m+1} a_0 + \dfrac{q^{m+1} - 1}{q-1} d \right)$

> "Thus, we have completed both the induction base and the induction step"

before we declare our inductive proof to be finished; it is clear enough that we have completed them. We also can remove the following two sentences:

> "Now, forget that we fixed m. We thus have shown that if $m \in \mathbb{N}$ is such that $\mathcal{A}(m)$ holds, then $\mathcal{A}(m+1)$ also holds.".

In fact, these sentences merely say that we have completed the induction step; they carry no other information (since the induction step always consists in fixing $m \in \mathbb{N}$ such that $\mathcal{A}(m)$ holds, and proving that $\mathcal{A}(m+1)$ also holds). So once we say that the induction step is completed, we don't need these sentences anymore.

So we can shorten our proof above a bit further:

Proof of Proposition 2.2 (third version). We shall prove (43) by induction on n:

Induction base: We have $q^0 = 1$, thus $q^0 - 1 = 0$, and therefore $\dfrac{q^0 - 1}{q-1} = \dfrac{0}{q-1} = 0$. Now,

$$\underbrace{q^0}_{=1} a_0 + \underbrace{\frac{q^0 - 1}{q-1}}_{=0} d = 1a_0 + 0d = a_0,$$

so that $a_0 = q^0 a_0 + \dfrac{q^0 - 1}{q-1} d$. In other words, (43) holds for $n = 0$. [37] This completes the induction base.

Induction step: Let $m \in \mathbb{N}$. Assume that (43) holds for $n = m$. [38] We must show that (43) holds for $n = m+1$. [39]

We have assumed that (43) holds for $n = m$. In other words, $a_m = q^m a_0 + \dfrac{q^m - 1}{q-1} d$

[37] Note that the statement "(43) holds for $n = 0$" (which we just proved) is precisely the statement $\mathcal{A}(0)$ in the previous two versions of our proof.

[38] Note that the statement "(43) holds for $n = m$" (which we just assumed) is precisely the statement $\mathcal{A}(m)$ in the previous two versions of our proof.

[39] Note that this statement "(43) holds for $n = m+1$" is precisely the statement $\mathcal{A}(m+1)$ in the previous two versions of our proof.

holds. Now, (42) (applied to $n = m$) yields

$$\begin{aligned} a_{m+1} &= q \underbrace{a_m}_{= q^m a_0 + \frac{q^m - 1}{q-1} d} + d = q \left(q^m a_0 + \frac{q^m - 1}{q-1} d \right) + d \\ &= \underbrace{qq^m}_{= q^{m+1}} a_0 + \underbrace{q \cdot \frac{q^m - 1}{q-1} d + d}_{= \left(q \cdot \frac{q^m - 1}{q-1} + 1 \right) d} \\ &= q^{m+1} a_0 + \underbrace{\left(q \cdot \frac{q^m - 1}{q-1} + 1 \right)}_{\substack{= \frac{q(q^m - 1) + (q-1)}{q-1} = \frac{q^{m+1} - 1}{q-1} \\ (\text{since } q(q^m-1)+(q-1) = qq^m - q + q - 1 = qq^m - 1 = q^{m+1} - 1)}} d \\ &= q^{m+1} a_0 + \frac{q^{m+1} - 1}{q-1} d. \end{aligned}$$

So we have shown that $a_{m+1} = q^{m+1} a_0 + \frac{q^{m+1} - 1}{q-1} d$. In other words, (43) holds for $n = m + 1$. This completes the induction step. Hence, (43) is proven by induction. This proves Proposition 2.2. □

2.2. Examples from modular arithmetic

2.2.1. Divisibility of integers

We shall soon give some more examples of inductive proofs, including some that will include slightly new tactics. These examples come from the realm of *modular arithmetic*, which is the study of congruences modulo integers. Before we come to these examples, we will introduce the definition of such congruences. But first, let us recall the definition of divisibility:

Definition 2.4. Let u and v be two integers. Then, we say that u *divides* v if and only if there exists an integer w such that $v = uw$. Instead of saying "u divides v", we can also say "v is *divisible by* u" or "v is a *multiple* of u" or "u is a *divisor* of v" or "$u \mid v$".

Thus, two integers u and v satisfy $u \mid v$ if and only if there is some $w \in \mathbb{Z}$ such that $v = uw$. For example, $1 \mid v$ holds for every integer v (since $v = 1v$), whereas $0 \mid v$ holds only for $v = 0$ (since $v = 0w$ is equivalent to $v = 0$). An integer v satisfies $2 \mid v$ if and only if v is even.

Definition 2.4 is fairly common in the modern literature (e.g., it is used in [Day16], [LeLeMe16], [Mulhol16] and [Rotman15]), but there are also some books that define these notations differently. For example, in [GrKnPa94], the notation "u divides v" is defined differently (it requires not only the existence of an integer w such that $v = uw$, but also that u is positive), whereas the notation "v is a multiple of u" is defined as it is here (i.e., it just means that there exists an integer w such that $v = uw$); thus, these two notations are not mutually interchangeable in [GrKnPa94].

Let us first prove some basic properties of divisibility:

Proposition 2.5. Let a, b and c be three integers such that $a \mid b$ and $b \mid c$. Then, $a \mid c$.

Proof of Proposition 2.5. We have $a \mid b$. In other words, there exists an integer w such that $b = aw$ (by the definition of "divides"). Consider this w, and denote it by k. Thus, k is an integer such that $b = ak$.

We have $b \mid c$. In other words, there exists an integer w such that $c = bw$ (by the definition of "divides"). Consider this w, and denote it by j. Thus, j is an integer such that $c = bj$.

Now, $c = \underbrace{b}_{=ak} j = akj$. Hence, there exists an integer w such that $c = aw$ (namely, $w = kj$). In other words, a divides c (by the definition of "divides"). In other words, $a \mid c$. This proves Proposition 2.5. □

Proposition 2.6. Let a, b and c be three integers such that $a \mid b$. Then, $ac \mid bc$.

Proof of Proposition 2.6. We have $a \mid b$. In other words, there exists an integer w such that $b = aw$ (by the definition of "divides"). Consider this w, and denote it by k. Thus, k is an integer such that $b = ak$. Hence, $\underbrace{b}_{=ak} c = akc = ack$. Thus, there exists an integer w such that $bc = acw$ (namely, $w = k$). In other words, ac divides bc (by the definition of "divides"). In other words, $ac \mid bc$. This proves Proposition 2.6. □

Proposition 2.7. Let a, b, g, x and y be integers such that $g = ax + by$. Let d be an integer such that $d \mid a$ and $d \mid b$. Then, $d \mid g$.

Proof of Proposition 2.7. We have $d \mid a$. In other words, there exists an integer w such that $a = dw$ (by the definition of "divides"). Consider this w, and denote it by p. Thus, p is an integer and satisfies $a = dp$.

Similarly, there is an integer q such that $b = dq$. Consider this q.

Now, $g = \underbrace{a}_{=dp} x + \underbrace{b}_{=dq} y = dpx + dqy = d(px + qy)$. Hence, there exists an integer w such that $g = dw$ (namely, $w = px + qy$). In other words, $d \mid g$ (by the definition of "divides"). This proves Proposition 2.7. □

It is easy to characterize divisibility in terms of fractions:

Proposition 2.8. Let a and b be two integers such that $a \neq 0$. Then, $a \mid b$ if and only if b/a is an integer.

Proof of Proposition 2.8. We first claim the following logical implication[40]:

$$(a \mid b) \implies (b/a \text{ is an integer}) . \tag{45}$$

[*Proof of (45):* Assume that $a \mid b$. In other words, there exists an integer w such that $b = aw$ (by the definition of "divides"). Consider this w. Now, dividing the equality $b = aw$ by a, we obtain $b/a = w$ (since $a \neq 0$). Hence, b/a is an integer (since w is an integer). This proves the implication (45).]

Next, we claim the following logical implication:

$$(b/a \text{ is an integer}) \implies (a \mid b) . \tag{46}$$

[*Proof of (46):* Assume that b/a is an integer. Let k denote this integer. Thus, $b/a = k$, so that $b = ak$. Hence, there exists an integer w such that $b = aw$ (namely, $w = k$). In other words, a divides b (by the definition of "divides"). In other words, $a \mid b$. This proves the implication (46).]

Combining the implications (45) and (46), we obtain the equivalence $(a \mid b) \iff (b/a \text{ is an integer})$. In other words, $a \mid b$ if and only if b/a is an integer. This proves Proposition 2.8. □

2.2.2. Definition of congruences

We can now define congruences:

Definition 2.9. Let a, b and n be three integers. Then, we say that *a is congruent to b modulo n* if and only if $n \mid a - b$. We shall use the notation "$a \equiv b \mod n$" for "a is congruent to b modulo n". Relations of the form "$a \equiv b \mod n$" (for integers a, b and n) are called *congruences modulo n*.

Thus, three integers a, b and n satisfy $a \equiv b \mod n$ if and only if $n \mid a - b$.

Hence, in particular:

- Any two integers a and b satisfy $a \equiv b \mod 1$. (Indeed, any two integers a and b satisfy $a - b = 1(a - b)$, thus $1 \mid a - b$, thus $a \equiv b \mod 1$.)
- Two integers a and b satisfy $a \equiv b \mod 0$ if and only if $a = b$. (Indeed, $a \equiv b \mod 0$ is equivalent to $0 \mid a - b$, which in turn is equivalent to $a - b = 0$, which in turn is equivalent to $a = b$.)
- Two integers a and b satisfy $a \equiv b \mod 2$ if and only if they have the same parity (i.e., they are either both odd or both even). This is not obvious at this point yet, but follows from Proposition 2.159 further below.

[40] A *logical implication* (or, short, *implication*) is a logical statement of the form "if $\mathcal{A}$, then $\mathcal{B}$" (where $\mathcal{A}$ and $\mathcal{B}$ are two statements).

We have

$$4 \equiv 10 \bmod 3 \qquad \text{and} \qquad 5 \equiv -35 \bmod 4.$$

Note that Day, in [Day16], writes "$a \equiv_n b$" instead of "$a \equiv b \bmod n$". Also, other authors (particularly of older texts) write "$a \equiv b \pmod{n}$" instead of "$a \equiv b \bmod n$".

Let us next introduce notations for the negations of the statements "$u \mid v$" and "$a \equiv b \bmod n$":

Definition 2.10. **(a)** If u and v are two integers, then the notation "$u \nmid v$" shall mean "not $u \mid v$" (that is, "u does not divide v").

(b) If a, b and n are three integers, then the notation "$a \not\equiv b \bmod n$" shall mean "not $a \equiv b \bmod n$" (that is, "a is not congruent to b modulo n").

Thus, three integers a, b and n satisfy $a \not\equiv b \bmod n$ if and only if $n \nmid a - b$. For example, $1 \not\equiv -1 \bmod 3$, since $3 \nmid 1 - (-1)$.

2.2.3. Congruence basics

Let us now state some of the basic laws of congruences (so far, not needing induction to prove):

Proposition 2.11. Let a and n be integers. Then:

(a) We have $a \equiv 0 \bmod n$ if and only if $n \mid a$.

(b) Let b be an integer. Then, $a \equiv b \bmod n$ if and only if $a \equiv b \bmod (-n)$.

(c) Let m and b be integers such that $m \mid n$. If $a \equiv b \bmod n$, then $a \equiv b \bmod m$.

Proof of Proposition 2.11. **(a)** We have the following chain of logical equivalences:

$$\begin{aligned}
&(a \equiv 0 \bmod n) \\
\Longleftrightarrow\ &(a \text{ is congruent to } 0 \text{ modulo } n) \\
&\qquad (\text{since "}a \equiv 0 \bmod n\text{" is just a notation for "}a \text{ is congruent to } 0 \text{ modulo } n\text{"}) \\
\Longleftrightarrow\ &\left(n \mid \underbrace{a - 0}_{=a}\right) \qquad (\text{by the definition of "congruent"}) \\
\Longleftrightarrow\ &(n \mid a).
\end{aligned}$$

Thus, we have $a \equiv 0 \bmod n$ if and only if $n \mid a$. This proves Proposition 2.11 **(a)**.

(b) Let us first assume that $a \equiv b \bmod n$. Thus, a is congruent to b modulo n. In other words, $n \mid a - b$ (by the definition of "congruent"). In other words, n divides $a - b$. In other words, there exists an integer w such that $a - b = nw$ (by the definition of "divides"). Consider this w, and denote it by k. Thus, k is an integer such that $a - b = nk$.

Thus, $a - b = nk = (-n)(-k)$. Hence, there exists an integer w such that $a - b = (-n)w$ (namely, $w = -k$). In other words, $-n$ divides $a - b$ (by the definition of "divides"). In other words, $-n \mid a-b$. In other words, a is congruent to b modulo $-n$ (by the definition of "congruent"). In other words, $a \equiv b \operatorname{mod} (-n)$.

Now, forget that we assumed that $a \equiv b \operatorname{mod} n$. We thus have shown that

$$\text{if } a \equiv b \operatorname{mod} n, \text{ then } a \equiv b \operatorname{mod} (-n) . \tag{47}$$

The same argument (applied to $-n$ instead of n) shows that

$$\text{if } a \equiv b \operatorname{mod} (-n), \text{ then } a \equiv b \operatorname{mod} (-(-n)) .$$

Since $-(-n) = n$, this rewrites as follows:

$$\text{if } a \equiv b \operatorname{mod} (-n), \text{ then } a \equiv b \operatorname{mod} n.$$

Combining this implication with (47), we conclude that $a \equiv b \operatorname{mod} n$ if and only if $a \equiv b \operatorname{mod} (-n)$. This proves Proposition 2.11 **(b)**.

(c) Assume that $a \equiv b \operatorname{mod} n$. Thus, a is congruent to b modulo n. In other words, $n \mid a-b$ (by the definition of "congruent"). Hence, Proposition 2.5 (applied to m, n and $a-b$ instead of a, b and c) yields $m \mid a-b$ (since $m \mid n$). In other words, a is congruent to b modulo m (by the definition of "congruent"). Thus, $a \equiv b \operatorname{mod} m$. This proves Proposition 2.11 **(c)**. □

Proposition 2.12. Let n be an integer.

(a) For any integer a, we have $a \equiv a \operatorname{mod} n$.

(b) For any integers a and b satisfying $a \equiv b \operatorname{mod} n$, we have $b \equiv a \operatorname{mod} n$.

(c) For any integers a, b and c satisfying $a \equiv b \operatorname{mod} n$ and $b \equiv c \operatorname{mod} n$, we have $a \equiv c \operatorname{mod} n$.

Proof of Proposition 2.12. **(a)** Let a be an integer. Then, $a - a = 0 = n \cdot 0$. Hence, there exists an integer w such that $a - a = nw$ (namely, $w = 0$). In other words, n divides $a - a$ (by the definition of "divides"). In other words, $n \mid a - a$. In other words, a is congruent to a modulo n (by the definition of "congruent"). In other words, $a \equiv a \operatorname{mod} n$. This proves Proposition 2.12 **(a)**.

(b) Let a and b be two integers satisfying $a \equiv b \operatorname{mod} n$. Thus, a is congruent to b modulo n (since $a \equiv b \operatorname{mod} n$). In other words, $n \mid a-b$ (by the definition of "congruent"). In other words, n divides $a-b$. In other words, there exists an integer w such that $a - b = nw$ (by the definition of "divides"). Consider this w, and denote it by q. Thus, q is an integer such that $a - b = nq$. Now, $b - a = -\underbrace{(a-b)}_{=nq} = -nq = n(-q)$. Hence, there exists an integer w such that $b - a = nw$ (namely, $w = -q$). In other words, n divides $b - a$ (by the definition of "divides"). In other words, $n \mid b - a$. In other words, b is congruent to a modulo n (by the definition of "congruent"). In other words, $b \equiv a \operatorname{mod} n$. This proves Proposition 2.12 **(b)**.

(c) Let a, b and c be three integers satisfying $a \equiv b \operatorname{mod} n$ and $b \equiv c \operatorname{mod} n$.

Just as in the above proof of Proposition 2.12 **(b)**, we can use the assumption $a \equiv b \mod n$ to construct an integer q such that $a - b = nq$. Similarly, we can use the assumption $b \equiv c \mod n$ to construct an integer r such that $b - c = nr$. Consider these q and r.

Now,

$$a - c = \underbrace{(a - b)}_{=nq} + \underbrace{(b - c)}_{=nr} = nq + nr = n(q + r).$$

Hence, there exists an integer w such that $a - c = nw$ (namely, $w = q + r$). In other words, n divides $a - c$ (by the definition of "divides"). In other words, $n \mid a - c$. In other words, a is congruent to c modulo n (by the definition of "congruent"). In other words, $a \equiv c \mod n$. This proves Proposition 2.12 **(c)**. □

Simple as they are, the three parts of Proposition 2.12 have names: Proposition 2.12 **(a)** is called the *reflexivity of congruence (modulo n)*; Proposition 2.12 **(b)** is called the *symmetry of congruence (modulo n)*; Proposition 2.12 **(c)** is called the *transitivity of congruence (modulo n)*.

Proposition 2.12 **(b)** allows the following definition:

Definition 2.13. Let n, a and b be three integers. Then, we say that *a and b are congruent modulo n* if and only if $a \equiv b \mod n$. Proposition 2.12 **(b)** shows that a and b actually play equal roles in this relation (i.e., the statement "a and b are congruent modulo n" is equivalent to "b and a are congruent modulo n").

Proposition 2.14. Let n be an integer. Then, $n \equiv 0 \mod n$.

Proof of Proposition 2.14. We have $n = n \cdot 1$. Thus, there exists an integer w such that $n = nw$ (namely, $w = 1$). Therefore, $n \mid n$ (by the definition of "divides"). Proposition 2.11 **(a)** (applied to $a = n$) shows that we have $n \equiv 0 \mod n$ if and only if $n \mid n$. Hence, we have $n \equiv 0 \mod n$ (since $n \mid n$). This proves Proposition 2.14. □

2.2.4. Chains of congruences

Proposition 2.12 shows that congruences (modulo n) behave like equalities – in that we can turn them around (since Proposition 2.12 **(b)** says that $a \equiv b \mod n$ implies $b \equiv a \mod n$) and we can chain them together (by Proposition 2.12 **(c)**) and in that every integer is congruent to itself (by Proposition 2.12 **(a)**). This leads to the following notation:

Definition 2.15. If $a_1, a_2, \ldots, a_k$ and n are integers, then the statement "$a_1 \equiv a_2 \equiv \cdots \equiv a_k \mod n$" shall mean that

$$(a_i \equiv a_{i+1} \mod n \text{ holds for each } i \in \{1, 2, \ldots, k - 1\}).$$

Such a statement is called a *chain of congruences modulo n* (or, less precisely, a *chain of congruences*). We shall refer to the integers $a_1, a_2, \ldots, a_k$ (but not n) as the *members* of this chain.

For example, the chain $a \equiv b \equiv c \equiv d \operatorname{mod} n$ (for five integers a, b, c, d, n) means that $a \equiv b \operatorname{mod} n$ and $b \equiv c \operatorname{mod} n$ and $c \equiv d \operatorname{mod} n$.

The usefulness of such chains lies in the following fact:

Proposition 2.16. Let $a_1, a_2, \ldots, a_k$ and n be integers such that $a_1 \equiv a_2 \equiv \cdots \equiv a_k \operatorname{mod} n$. Let u and v be two elements of $\{1, 2, \ldots, k\}$. Then,

$$a_u \equiv a_v \operatorname{mod} n.$$

In other words, any two members of a chain of congruences modulo n are congruent to each other modulo n. Thus, chains of congruences are like chains of equalities: From any chain of congruences modulo n with k members, you can extract k^2 congruences modulo n by picking any two members of the chain.

Example 2.17. Proposition 2.16 shows (among other things) that if a, b, c, d, e, n are integers such that $a \equiv b \equiv c \equiv d \equiv e \operatorname{mod} n$, then $a \equiv d \operatorname{mod} n$ and $b \equiv d \operatorname{mod} n$ and $e \equiv b \operatorname{mod} n$ (and various other congruences).

Unsurprisingly, Proposition 2.16 can be proven by induction, although not in an immediately obvious manner: We cannot directly prove it by induction on n, on k, on u or on v. Instead, we will first introduce an auxiliary statement (the statement (49) in the following proof) which will be tailored to an inductive proof. This is a commonly used tactic, and particularly helpful to us now as we only have the most basic form of the principle of induction available. (Soon, we will see more versions of that principle, which will obviate the need for some of the tailoring.)

Proof of Proposition 2.16. By assumption, we have $a_1 \equiv a_2 \equiv \cdots \equiv a_k \operatorname{mod} n$. In other words,

$$\left(a_i \equiv a_{i+1} \operatorname{mod} n \text{ holds for each } i \in \{1, 2, \ldots, k-1\}\right) \tag{48}$$

(since this is what "$a_1 \equiv a_2 \equiv \cdots \equiv a_k \operatorname{mod} n$" means).

Fix $p \in \{1, 2, \ldots, k\}$. For each $i \in \mathbb{N}$, we let $\mathcal{A}(i)$ be the statement

$$\left(\text{if } p + i \in \{1, 2, \ldots, k\}, \text{ then } a_p \equiv a_{p+i} \operatorname{mod} n\right). \tag{49}$$

We shall prove that this statement $\mathcal{A}(i)$ holds for each $i \in \mathbb{N}$.

In fact, let us prove this by induction on i: [41]

Induction base: The statement $\mathcal{A}(0)$ holds[42]. This completes the induction base.

Induction step: Let $m \in \mathbb{N}$. Assume that $\mathcal{A}(m)$ holds. We must show that $\mathcal{A}(m+1)$ holds.

[41] Thus, the letter "i" plays the role of the "n" in Theorem 2.1 (since we are already using "n" for a different thing).

[42] *Proof.* Proposition 2.12 **(a)** (applied to $a = a_p$) yields $a_p \equiv a_p \operatorname{mod} n$. In view of $p = p + 0$, this rewrites as $a_p \equiv a_{p+0} \operatorname{mod} n$. Hence, $\left(\text{if } p + 0 \in \{1, 2, \ldots, k\}, \text{ then } a_p \equiv a_{p+0} \operatorname{mod} n\right)$. But this is precisely the statement $\mathcal{A}(0)$. Hence, the statement $\mathcal{A}(0)$ holds.

We have assumed that $\mathcal{A}(m)$ holds. In other words,

$$\left(\text{if } p+m \in \{1,2,\ldots,k\}\text{, then } a_p \equiv a_{p+m} \operatorname{mod} n\right). \tag{50}$$

Next, let us assume that $p+(m+1) \in \{1,2,\ldots,k\}$. Thus, $p+(m+1) \leq k$, so that $p+m+1 = p+(m+1) \leq k$ and therefore $p+m \leq k-1$. Also, $p \in \{1,2,\ldots,k\}$, so that $p \geq 1$ and thus $\underbrace{p}_{\geq 1} + \underbrace{m}_{\geq 0} \geq 1+0 = 1$. Combining this with $p+m \leq k-1$, we obtain $p+m \in \{1,2,\ldots,k-1\} \subseteq \{1,2,\ldots,k\}$. Hence, (50) shows that $a_p \equiv a_{p+m} \operatorname{mod} n$. But (48) (applied to $p+m$ instead of i) yields $a_{p+m} \equiv a_{(p+m)+1} \operatorname{mod} n$ (since $p+m \in \{1,2,\ldots,k-1\}$).

So we know that $a_p \equiv a_{p+m} \operatorname{mod} n$ and $a_{p+m} \equiv a_{(p+m)+1} \operatorname{mod} n$. Hence, Proposition 2.12 **(c)** (applied to $a = a_p$, $b = a_{p+m}$ and $c = a_{(p+m)+1}$) yields $a_p \equiv a_{(p+m)+1} \operatorname{mod} n$. Since $(p+m)+1 = p+(m+1)$, this rewrites as $a_p \equiv a_{p+(m+1)} \operatorname{mod} n$.

Now, forget that we assumed that $p+(m+1) \in \{1,2,\ldots,k\}$. We thus have shown that

$$\left(\text{if } p+(m+1) \in \{1,2,\ldots,k\}\text{, then } a_p \equiv a_{p+(m+1)} \operatorname{mod} n\right).$$

But this is precisely the statement $\mathcal{A}(m+1)$. Thus, $\mathcal{A}(m+1)$ holds.

Now, forget that we fixed m. We thus have shown that if $m \in \mathbb{N}$ is such that $\mathcal{A}(m)$ holds, then $\mathcal{A}(m+1)$ also holds. This completes the induction step.

Thus, we have completed both the induction base and the induction step. Hence, by induction, we conclude that $\mathcal{A}(i)$ holds for each $i \in \mathbb{N}$. In other words, (49) holds for each $i \in \mathbb{N}$.

We are not done yet, since our goal is to prove Proposition 2.16, not merely to prove $\mathcal{A}(i)$. But this is now easy.

First, let us forget that we fixed p. Thus, we have shown that (49) holds for each $p \in \{1,2,\ldots,k\}$ and $i \in \mathbb{N}$.

But we have either $u \leq v$ or $u > v$. In other words, we are in one of the following two cases:

Case 1: We have $u \leq v$.

Case 2: We have $u > v$.

Let us first consider Case 1. In this case, we have $u \leq v$. Thus, $v-u \geq 0$, so that $v-u \in \mathbb{N}$. But recall that (49) holds for each $p \in \{1,2,\ldots,k\}$ and $i \in \mathbb{N}$. Applying this to $p = u$ and $i = v-u$, we conclude that (49) holds for $p = u$ and $i = v-u$ (since $u \in \{1,2,\ldots,k\}$ and $v-u \in \mathbb{N}$). In other words,

$$\left(\text{if } u+(v-u) \in \{1,2,\ldots,k\}\text{, then } a_u \equiv a_{u+(v-u)} \operatorname{mod} n\right).$$

Since $u+(v-u) = v$, this rewrites as

$$\left(\text{if } v \in \{1,2,\ldots,k\}\text{, then } a_u \equiv a_v \operatorname{mod} n\right).$$

Since $v \in \{1,2,\ldots,k\}$ holds (by assumption), we conclude that $a_u \equiv a_v \operatorname{mod} n$. Thus, Proposition 2.16 is proven in Case 1.

Let us now consider Case 2. In this case, we have $u > v$. Thus, $u - v > 0$, so that $u - v \in \mathbb{N}$. But recall that (49) holds for each $p \in \{1, 2, \ldots, k\}$ and $i \in \mathbb{N}$. Applying this to $p = v$ and $i = u - v$, we conclude that (49) holds for $p = v$ and $i = u - v$ (since $v \in \{1, 2, \ldots, k\}$ and $u - v \in \mathbb{N}$). In other words,

$$\left(\text{if } v + (u - v) \in \{1, 2, \ldots, k\}, \text{ then } a_v \equiv a_{v+(u-v)} \bmod n\right).$$

Since $v + (u - v) = u$, this rewrites as

$$(\text{if } u \in \{1, 2, \ldots, k\}, \text{ then } a_v \equiv a_u \bmod n).$$

Since $u \in \{1, 2, \ldots, k\}$ holds (by assumption), we conclude that $a_v \equiv a_u \bmod n$. Therefore, Proposition 2.12 **(b)** (applied to $a = a_v$ and $b = a_u$) yields that $a_u \equiv a_v \bmod n$. Thus, Proposition 2.16 is proven in Case 2.

Hence, Proposition 2.16 is proven in both Cases 1 and 2. Since these two Cases cover all possibilities, we thus conclude that Proposition 2.16 always holds. □

2.2.5. Chains of inequalities (a digression)

Much of the above proof of Proposition 2.16 was unremarkable and straightforward reasoning – but this proof is nevertheless fundamental and important. More or less the same argument can be used to show the following fact about chains of inequalities:

Proposition 2.18. Let $a_1, a_2, \ldots, a_k$ be integers such that $a_1 \leq a_2 \leq \cdots \leq a_k$. (Recall that the statement "$a_1 \leq a_2 \leq \cdots \leq a_k$" means that ($a_i \leq a_{i+1}$ holds for each $i \in \{1, 2, \ldots, k-1\}$).) Let u and v be two elements of $\{1, 2, \ldots, k\}$ such that $u \leq v$. Then,

$$a_u \leq a_v.$$

Proposition 2.18 is similar to Proposition 2.16, with the congruences replaced by inequalities; but note that the condition "$u \leq v$" is now required. Make sure you understand where you need this condition when adapting the proof of Proposition 2.16 to Proposition 2.18!

For future use, let us prove a corollary of Proposition 2.18 which essentially observes that the inequality sign in $a_u \leq a_v$ can be made strict if there is any strict inequality sign between a_u and a_v in the chain $a_1 \leq a_2 \leq \cdots \leq a_k$:

Corollary 2.19. Let $a_1, a_2, \ldots, a_k$ be integers such that $a_1 \leq a_2 \leq \cdots \leq a_k$. Let u and v be two elements of $\{1, 2, \ldots, k\}$ such that $u \leq v$. Let $p \in \{u, u+1, \ldots, v-1\}$ be such that $a_p < a_{p+1}$. Then,

$$a_u < a_v.$$

Proof of Corollary 2.19. From $u \in \{1, 2, \ldots, k\}$, we obtain $u \geq 1$. From $v \in \{1, 2, \ldots, k\}$, we obtain $v \leq k$. From $p \in \{u, u+1, \ldots, v-1\}$, we obtain $p \geq u$ and $p \leq v-1$. From $p \leq v-1$, we obtain $p+1 \leq v \leq k$. Combining this with $p+1 \geq p \geq u \geq 1$, we obtain $p+1 \in \{1, 2, \ldots, k\}$ (since $p+1$ is an integer). Combining $p \leq v-1 \leq v \leq k$ with $p \geq u \geq 1$, we obtain $p \in \{1, 2, \ldots, k\}$ (since p is an integer). We thus know that both p and $p+1$ are elements of $\{1, 2, \ldots, k\}$.

We have $p \geq u$, thus $u \leq p$. Hence, Proposition 2.18 (applied to p instead of v) yields $a_u \leq a_p$. Combining this with $a_p < a_{p+1}$, we find $a_u < a_{p+1}$.

We have $p+1 \leq v$. Hence, Proposition 2.18 (applied to $p+1$ instead of u) yields $a_{p+1} \leq a_v$. Combining $a_u < a_{p+1}$ with $a_{p+1} \leq a_v$, we obtain $a_u < a_v$. This proves Corollary 2.19. □

In particular, we see that the inequality sign in $a_u \leq a_v$ is strict when $u < v$ holds and **all** inequality signs in the chain $a_1 \leq a_2 \leq \cdots \leq a_k$ are strict:

Corollary 2.20. Let $a_1, a_2, \ldots, a_k$ be integers such that $a_1 < a_2 < \cdots < a_k$. (Recall that the statement "$a_1 < a_2 < \cdots < a_k$" means that ($a_i < a_{i+1}$ holds for each $i \in \{1, 2, \ldots, k-1\}$).) Let u and v be two elements of $\{1, 2, \ldots, k\}$ such that $u < v$. Then,

$$a_u < a_v.$$

Proof of Corollary 2.20. From $u < v$, we obtain $u \leq v-1$ (since u and v are integers). Combining this with $u \geq u$, we conclude that $u \in \{u, u+1, \ldots, v-1\}$. Also, from $a_1 < a_2 < \cdots < a_k$, we obtain $a_1 \leq a_2 \leq \cdots \leq a_k$.

We have $u \leq v-1$, thus $u+1 \leq v \leq k$ (since $v \in \{1, 2, \ldots, k\}$), so that $u \leq k-1$. Combining this with $u \geq 1$ (which is a consequence of $u \in \{1, 2, \ldots, k\}$), we find $u \in \{1, 2, \ldots, k-1\}$. Hence, from $a_1 < a_2 < \cdots < a_k$, we obtain $a_u < a_{u+1}$. Hence, Corollary 2.19 (applied to $p = u$) yields $a_u < a_v$ (since $u \leq v$ (because $u < v$)). This proves Corollary 2.20. □

2.2.6. Addition, subtraction and multiplication of congruences

Let us now return to the topic of congruences.

Chains of congruences can include equality signs. For example, if a, b, c, d, n are integers, then "$a \equiv b = c \equiv d \bmod n$" means that $a \equiv b \bmod n$ and $b = c$ and $c \equiv d \bmod n$. Such a chain is still a chain of congruences, because $b = c$ implies $b \equiv c \bmod n$ (by Proposition 2.12 **(a)**).

Let us continue with basic properties of congruences:

Proposition 2.21. Let a, b, c, d and n be integers such that $a \equiv b \bmod n$ and $c \equiv d \bmod n$. Then:

(a) We have $a + c \equiv b + d \bmod n$.

(b) We have $a - c \equiv b - d \bmod n$.

(c) We have $ac \equiv bd \bmod n$.

Note that Proposition 2.21 does **not** claim that $a/c \equiv b/d \mod n$. Indeed, this would not be true in general. One reason for this is that a/c and b/d aren't always integers. But even when they are, they may not satisfy $a/c \equiv b/d \mod n$. For example, $6 \equiv 4 \mod 2$ and $2 \equiv 2 \mod 2$, but $6/2 \not\equiv 4/2 \mod 2$. Likewise, Proposition 2.21 does **not** claim that $a^c \equiv b^d \mod n$ even when a, b, c, d are nonnegative; that too would not be true. But we will soon see that a weaker statement (Proposition 2.22) holds. First, let us prove Proposition 2.21:

Proof of Proposition 2.21. From $a \equiv b \mod n$, we conclude that a is congruent to b modulo n. In other words, $n \mid a - b$ (by the definition of "congruent"). In other words, n divides $a - b$. In other words, there exists an integer w such that $a - b = nw$ (by the definition of "divides"). Consider this w, and denote it by q. Thus, q is an integer such that $a - b = nq$.

Similarly, from $c \equiv d \mod n$, we can construct an integer r such that $c - d = nr$. Consider this r.

(a) We have

$$(a+c) - (b+d) = \underbrace{(a-b)}_{=nq} + \underbrace{(c-d)}_{=nr} = nq + nr = n(q+r).$$

Hence, there exists an integer w such that $(a+c) - (b+d) = nw$ (namely, $w = q + r$). In other words, n divides $(a+c) - (b+d)$ (by the definition of "divides"). In other words, $n \mid (a+c) - (b+d)$. In other words, $a + c \equiv b + d \mod n$ (by the definition of "congruent"). This proves Proposition 2.21 **(a)**.

(b) We have

$$(a-c) - (b-d) = \underbrace{(a-b)}_{=nq} - \underbrace{(c-d)}_{=nr} = nq - nr = n(q-r).$$

Hence, there exists an integer w such that $(a-c) - (b-d) = nw$ (namely, $w = q - r$). In other words, n divides $(a-c) - (b-d)$ (by the definition of "divides"). In other words, $n \mid (a-c) - (b-d)$. In other words, $a - c \equiv b - d \mod n$ (by the definition of "congruent"). This proves Proposition 2.21 **(b)**.

(c) We have $ac - ad = a \underbrace{(c-d)}_{=nr} = anr = n(ar)$. Hence, there exists an integer w such that $ac - ad = nw$ (namely, $w = ar$). In other words, n divides $ac - ad$ (by the definition of "divides"). In other words, $n \mid ac - ad$. In other words, $ac \equiv ad \mod n$ (by the definition of "congruent").

We have $ad - bd = \underbrace{(a-b)}_{=nq} d = nqd = n(qd)$. Hence, there exists an integer w such that $ad - bd = nw$ (namely, $w = qd$). In other words, n divides $ad - bd$ (by the definition of "divides"). In other words, $n \mid ad - bd$. In other words, $ad \equiv bd \mod n$ (by the definition of "congruent").

Now, we know that $ac \equiv ad \mod n$ and $ad \equiv bd \mod n$. Hence, Proposition 2.12 **(c)** (applied to ac, ad and bd instead of a, b and c) shows that $ac \equiv bd \mod n$. This proves Proposition 2.21 **(c)**. □

Proposition 2.21 shows yet another aspect in which congruences (modulo n) behave like equalities: They can be added, subtracted and multiplied, in the following sense:

- We can add two congruences modulo n (in the sense of adding each side of one congruence to the corresponding side of the other); this yields a new congruence modulo n (because of Proposition 2.21 **(a)**).
- We can subtract two congruences modulo n; this yields a new congruence modulo n (because of Proposition 2.21 **(b)**).
- We can multiply two congruences modulo n; this yields a new congruence modulo n (because of Proposition 2.21 **(c)**).

2.2.7. Substitutivity for congruences

Combined with Proposition 2.12, these observations lead to a further feature of congruences, which is even more important: the principle of *substitutivity for congruences*. We are not going to state it fully formally (as it is a meta-mathematical principle), but merely explain what it means.

Recall that the *principle of substitutivity for equalities* says the following:

> *Principle of substitutivity for equalities:* If two objects[43] x and x' are equal, and if we have any expression A that involves the object x, then we can replace this x (or, more precisely, any arbitrary appearance of x in A) in A by x'; the value of the resulting expression A' will be equal to the value of A.

Here are two examples of how this principle can be used:

- If a, b, c, d, e, c' are numbers such that $c = c'$, then the principle of substitutivity for equalities says that we can replace c by c' in the expression $a\left(b-(c+d)e\right)$, and the value of the resulting expression $a\left(b-(c'+d)e\right)$ will be equal to the value of $a\left(b-(c+d)e\right)$; that is, we have
$$a\left(b-(c+d)e\right) = a\left(b-\left(c'+d\right)e\right). \qquad (51)$$
- If a, b, c, a' are numbers such that $a = a'$, then
$$(a-b)(a+b) = \left(a'-b\right)(a+b), \qquad (52)$$
because the principle of substitutivity allows us to replace the first a appearing in the expression $(a-b)(a+b)$ by an a'. (We can also replace the second a by a', of course.)

[43] "Objects" can be numbers, sets, tuples or any other mathematical objects.

More generally, we can make several such replacements at the same time.

The principle of substitutivity for equalities is one of the headstones of mathematical logic; it is the essence of what it means for two objects to be equal.

The *principle of substitutivity for congruences* is similar, but far less fundamental; it says the following:

> *Principle of substitutivity for congruences:* Fix an integer n. If two numbers x and x' are congruent to each other modulo n (that is, $x \equiv x' \mod n$), and if we have any expression A that involves only integers, addition, subtraction and multiplication, and involves the object x, then we can replace this x (or, more precisely, any arbitrary appearance of x in A) in A by x'; the value of the resulting expression A' will be congruent to the value of A modulo n.

Note that this principle is less general than the principle of substitutivity for equalities, because it only applies to expressions that are built from integers and certain operations (note that division is not one of these operations). But it still lets us prove analogues of our above examples (51) and (52):

- If n is any integer, and if a, b, c, d, e, c' are integers such that $c \equiv c' \mod n$, then the principle of substitutivity for congruences says that we can replace c by c' in the expression $a\left(b-(c+d)\,e\right)$, and the value of the resulting expression $a\left(b-(c'+d)\,e\right)$ will be congruent to the value of $a\left(b-(c+d)\,e\right)$ modulo n; that is, we have

$$a\left(b-(c+d)\,e\right) \equiv a\left(b-\left(c'+d\right)e\right) \mod n. \tag{53}$$

- If n is any integer, and if a, b, c, a' are integers such that $a \equiv a' \mod n$, then

$$(a-b)\,(a+b) \equiv \left(a'-b\right)(a+b) \mod n, \tag{54}$$

  because the principle of substitutivity allows us to replace the first a appearing in the expression $(a-b)\,(a+b)$ by an a'. (We can also replace the second a by a', of course.)

We shall not prove the principle of substitutivity for congruences, since we have not formalized it (after all, we have not defined what an "expression" is). But we shall prove the specific congruences (53) and (54) using Proposition 2.21 and Proposition 2.12; the way in which we prove these congruences is symptomatic: Every congruence obtained from the principle of substitutivity for congruences can be proven in a manner like these. Thus, we hope that the proofs of (53) and (54) given below serve as templates which can easily be adapted to any other situation in which an application of the principle of substitutivity for congruences needs to be justified.

Proof of (53). Let n be any integer, and let a, b, c, d, e, c' be integers such that $c \equiv c' \mod n$.

Adding the congruence $c \equiv c' \mod n$ with the congruence $d \equiv d \mod n$ (which follows from Proposition 2.12 **(a)**), we obtain $c + d \equiv c' + d \mod n$. Multiplying this congruence with the congruence $e \equiv e \mod n$ (which follows from Proposition 2.12 **(a)**), we obtain $(c + d) e \equiv (c' + d) e \mod n$. Subtracting this congruence from the congruence $b \equiv b \mod n$ (which, again, follows from Proposition 2.12 **(a)**), we obtain $b - (c + d) e \equiv b - (c' + d) e \mod n$. Multiplying the congruence $a \equiv a \mod n$ (which follows from Proposition 2.12 **(a)**) with this congruence, we obtain $a (b - (c + d) e) \equiv a (b - (c' + d) e) \mod n$. This proves (53). □

Proof of (54). Let n be any integer, and let a, b, c, a' be integers such that $a \equiv a' \mod n$.

Subtracting the congruence $b \equiv b \mod n$ (which follows from Proposition 2.12 **(a)**) from the congruence $a \equiv a' \mod n$, we obtain $a - b \equiv a' - b \mod n$. Multiplying this congruence with the congruence $a + b \equiv a + b \mod n$ (which follows from Proposition 2.12 **(a)**), we obtain $(a - b)(a + b) \equiv (a' - b)(a + b) \mod n$. This proves (54). □

As we said, these two proofs are exemplary: Any congruence obtained from the principle of substitutivity for congruences can be proven in such a way (starting with the congruence $x \equiv x' \mod n$, and then "wrapping" it up in the expression A by repeatedly adding, multiplying and subtracting congruences that follow from Proposition 2.12 **(a)**).

When we apply the principle of substitutivity for congruences, we shall use underbraces to point out which integers we are replacing. For example, when deriving (53) from this principle, we shall write

$$a \left(b - \left(\underbrace{c}_{\equiv c' \mod n} + d \right) e \right) \equiv a \left(b - \left(c' + d \right) e \right) \mod n,$$

in order to stress that we are replacing c by c'. Likewise, when deriving (54) from this principle, we shall write

$$\left(\underbrace{a}_{\equiv a' \mod n} - b \right) (a + b) \equiv \left(a' - b \right) (a + b) \mod n,$$

in order to stress that we are replacing the first a (but not the second a) by a'.

The principle of substitutivity for congruences allows us to replace a **single** integer x appearing in an expression by another integer x' that is congruent to x modulo n. Applying this principle many times, we thus conclude that we can also replace **several** integers at the same time (because we can get to the same result by performing these replacements one at a time, and Proposition 2.16 shows that the final result will be congruent to the original result).

For example, if seven integers a, a', b, b', c, c', n satisfy $a \equiv a' \mod n$ and $b \equiv b' \mod n$ and $c \equiv c' \mod n$, then

$$bc + ca + ab \equiv b'c' + c'a' + a'b' \mod n, \tag{55}$$

because we can replace all the six integers b, c, c, a, a, b in the expression $bc + ca + ab$ (listed in the order of their appearance in this expression) by b', c', c', a', a', b', respectively. If we want to derive this from the principle of substitutivity for congruences, we must perform the replacements one at a time, e.g., as follows:

$$\begin{aligned} \underbrace{b}_{\equiv b' \mod n} c + ca + ab &\equiv b' \underbrace{c}_{\equiv c' \mod n} + ca + ab \equiv b'c' + \underbrace{c}_{\equiv c' \mod n} a + ab \\ &\equiv b'c' + c' \underbrace{a}_{\equiv a' \mod n} + ab \equiv b'c' + c'a' + \underbrace{a}_{\equiv a' \mod n} b \\ &\equiv b'c' + c'a' + a' \underbrace{b}_{\equiv b' \mod n} \equiv b'c' + c'a' + a'b' \mod n. \end{aligned}$$

Of course, we shall always just show the replacements as a single step:

$$\underbrace{b}_{\equiv b' \mod n} \underbrace{c}_{\equiv c' \mod n} + \underbrace{c}_{\equiv c' \mod n} \underbrace{a}_{\equiv a' \mod n} + \underbrace{a}_{\equiv a' \mod n} \underbrace{b}_{\equiv b' \mod n} \equiv b'c' + c'a' + a'b' \mod n.$$

2.2.8. Taking congruences to the k-th power

We have seen that congruences (like equalities) can be added, subtracted and multiplied (but, unlike equalities, they cannot be divided). One other thing we can do with congruences is taking powers of them, as long as the exponent is a nonnegative integer. This relies on the following fact:

Proposition 2.22. Let a, b and n be three integers such that $a \equiv b \mod n$. Then, $a^k \equiv b^k \mod n$ for each $k \in \mathbb{N}$.

The following proof of Proposition 2.22 is an example of a straightforward inductive proof; the only thing to keep in mind is that it uses induction on k, not induction on n as some of our previous proofs did.

Proof of Proposition 2.22. We claim that

$$a^k \equiv b^k \mod n \qquad \text{for each } k \in \mathbb{N}. \tag{56}$$

We shall prove (56) by induction on k:

Induction base: We have $1 \equiv 1 \mod n$ (by Proposition 2.12 **(a)**). In view of $a^0 = 1$ and $b^0 = 1$, this rewrites as $a^0 \equiv b^0 \mod n$. In other words, (56) holds for $k = 0$. This completes the induction base.

Induction step: Let $m \in \mathbb{N}$. Assume that (56) holds for $k = m$. We must show that (56) holds for $k = m + 1$.

We have assumed that (56) holds for $k = m$. In other words, we have $a^m \equiv b^m \mod n$. Now,

$$a^{m+1} = \underbrace{a^m}_{\equiv b^m \mod n} \underbrace{a}_{\equiv b \mod n} \equiv b^m b = b^{m+1} \mod n.$$

[44] In other words, (56) holds for $k = m + 1$. This completes the induction step. Hence, (56) is proven by induction. This proves Proposition 2.22. □

2.3. A few recursively defined sequences

2.3.1. $a_n = a_{n-1}^q + r$

We next proceed to give some more examples of proofs by induction.

Example 2.23. Let $(a_0, a_1, a_2, \ldots)$ be a sequence of integers defined recursively by

$$a_0 = 0, \qquad \text{and}$$
$$a_n = a_{n-1}^2 + 1 \qquad \text{for each } n \geq 1.$$

("Defined recursively" means that we aren't defining each entry a_n of our sequence by an explicit formula, but rather defining a_n in terms of the previous entries $a_0, a_1, \ldots, a_{n-1}$. Thus, in order to compute some entry a_n of our sequence, we need to compute all the previous entries $a_0, a_1, \ldots, a_{n-1}$. This means that if we want to compute a_n, we should first compute a_0, then compute a_1 (using our value of a_0), then compute a_2 (using our values of a_0 and a_1), and so on, until we reach a_n. For example, in order to compute a_6, we proceed as follows:

$$\begin{aligned}
a_0 &= 0; \\
a_1 &= a_0^2 + 1 = 0^2 + 1 = 1; \\
a_2 &= a_1^2 + 1 = 1^2 + 1 = 2; \\
a_3 &= a_2^2 + 1 = 2^2 + 1 = 5; \\
a_4 &= a_3^2 + 1 = 5^2 + 1 = 26; \\
a_5 &= a_4^2 + 1 = 26^2 + 1 = 677; \\
a_6 &= a_5^2 + 1 = 677^2 + 1 = 458\,330.
\end{aligned}$$

And similarly we can compute a_n for any $n \in \mathbb{N}$.)

This sequence $(a_0, a_1, a_2, \ldots)$ is not unknown: It is the sequence A003095 in the Online Encyclopedia of Integer Sequences.

[44] This computation relied on the principle of substitutivity for congruences. Here is how to rewrite this argument in a more explicit way (without using this principle): We have $a^m \equiv b^m \mod n$ and $a \equiv b \mod n$. Hence, Proposition 2.21 **(c)** (applied to a^m, b^m, a and b instead of a, b, c and d) yields $a^m a \equiv b^m b \mod n$. This rewrites as $a^{m+1} \equiv b^{m+1} \mod n$ (since $a^{m+1} = a^m a$ and $b^{m+1} = b^m b$).

A look at the first few entries of the sequence makes us realize that both a_2 and a_3 divide a_6, just as the integers 2 and 3 themselves divide 6. This suggests that we might have $a_u \mid a_v$ whenever u and v are two nonnegative integers satisfying $u \mid v$. We shall soon prove this observation (which was found by Michael Somos in 2013) in greater generality.

Theorem 2.24. Fix some $q \in \mathbb{N}$ and $r \in \mathbb{Z}$. Let $(a_0, a_1, a_2, \ldots)$ be a sequence of integers defined recursively by

$$\begin{aligned} a_0 &= 0, \qquad \text{and} \\ a_n &= a_{n-1}^q + r \qquad \text{for each } n \geq 1. \end{aligned}$$

(Note that if $q = 2$ and $r = 1$, then this sequence $(a_0, a_1, a_2, \ldots)$ is precisely the sequence $(a_0, a_1, a_2, \ldots)$ from Example 2.23. If $q = 3$ and $r = 1$, then our sequence $(a_0, a_1, a_2, \ldots)$ is the sequence A135361 in the Online Encyclopedia of Integer Sequences. If $q = 0$, then our sequence $(a_0, a_1, a_2, \ldots)$ is $(0, r+1, r+1, r+1, \ldots)$. If $q = 1$, then our sequence $(a_0, a_1, a_2, \ldots)$ is $(0, r, 2r, 3r, 4r, \ldots)$, as can be easily proven by induction.)

(a) For any $k \in \mathbb{N}$ and $n \in \mathbb{N}$, we have $a_{k+n} \equiv a_k \mod a_n$.

(b) For any $n \in \mathbb{N}$ and $w \in \mathbb{N}$, we have $a_n \mid a_{nw}$.

(c) If u and v are two nonnegative integers satisfying $u \mid v$, then $a_u \mid a_v$.

Proof of Theorem 2.24. **(a)** Let $n \in \mathbb{N}$. We claim that

$$a_{k+n} \equiv a_k \mod a_n \qquad \text{for every } k \in \mathbb{N}. \tag{57}$$

We shall prove (57) by induction on k:

Induction base: Proposition 2.14 (applied to a_n instead of n) yields $a_n \equiv 0 \mod a_n$. This rewrites as $a_n \equiv a_0 \mod a_n$ (since $a_0 = 0$). In other words, $a_{0+n} \equiv a_0 \mod a_n$ (since $0 + n = n$). In other words, (57) holds for $k = 0$. This completes the induction base.

Induction step: Let $m \in \mathbb{N}$. Assume that (57) holds for $k = m$. We must prove that (57) holds for $k = m + 1$.

We have assumed that (57) holds for $k = m$. In other words, we have

$$a_{m+n} \equiv a_m \mod a_n.$$

Hence, Proposition 2.22 (applied to a_{m+n}, a_m, a_n and q instead of a, b, n and k) shows that $a_{m+n}^q \equiv a_m^q \mod a_n$. Hence, $a_{m+n}^q + r \equiv a_m^q + r \mod a_n$. (Indeed, this follows by adding the congruence $a_{m+n}^q \equiv a_m^q \mod a_n$ to the congruence $r \equiv r \mod a_n$; the latter congruence is a consequence of Proposition 2.12 **(a)**.)

Now, $(m+1) + n = (m+n) + 1 \geq 1$. Hence, the recursive definition of the sequence $(a_0, a_1, a_2, \ldots)$ yields

$$a_{(m+1)+n} = a_{((m+1)+n)-1}^q + r = a_{m+n}^q + r$$

(since $((m+1)+n)-1=m+n$). Also, $m+1\geq 1$. Hence, the recursive definition of the sequence $(a_0,a_1,a_2,\ldots)$ yields

$$a_{m+1}=a_{(m+1)-1}^q+r=a_m^q+r.$$

The congruence $a_{m+n}^q+r\equiv a_m^q+r \operatorname{mod} a_n$ rewrites as $a_{(m+1)+n}\equiv a_{m+1} \operatorname{mod} a_n$ (since $a_{(m+1)+n}=a_{m+n}^q+r$ and $a_{m+1}=a_m^q+r$). In other words, (57) holds for $k=m+1$. This completes the induction step. Thus, (57) is proven by induction.

Therefore, Theorem 2.24 **(a)** is proven.

(b) Let $n\in\mathbb{N}$. We claim that

$$a_n\mid a_{nw} \qquad \text{for every } w\in\mathbb{N}. \tag{58}$$

We shall prove (58) by induction on w:

Induction base: We have $a_{n\cdot 0}=a_0=0$. But $a_n\mid 0$ (since $0=0a_n$). This rewrites as $a_n\mid a_{n\cdot 0}$ (since $a_{n\cdot 0}=0$). In other words, (58) holds for $w=0$. This completes the induction base.

Induction step: Let $m\in\mathbb{N}$. Assume that (58) holds for $w=m$. We must prove that (58) holds for $w=m+1$.

We have assumed that (58) holds for $w=m$. In other words, we have $a_n\mid a_{nm}$.

Proposition 2.11 **(a)** (applied to a_{nm} and a_n instead of a and n) shows that we have $a_{nm}\equiv 0 \operatorname{mod} a_n$ if and only if $a_n\mid a_{nm}$. Hence, we have $a_{nm}\equiv 0 \operatorname{mod} a_n$ (since $a_n\mid a_{nm}$).

Theorem 2.24 **(a)** (applied to $k=nm$) yields $a_{nm+n}\equiv a_{nm} \operatorname{mod} a_n$. Thus, $a_{nm+n}\equiv a_{nm}\equiv 0 \operatorname{mod} a_n$. This is a chain of congruences; hence, an application of Proposition 2.16 shows that $a_{nm+n}\equiv 0 \operatorname{mod} a_n$. (In the future, we shall no longer explicitly say things like this; we shall leave it to the reader to apply Proposition 2.16 to any chain of congruences that we write down.)

Proposition 2.11 **(a)** (applied to a_{nm+n} and a_n instead of a and n) shows that we have $a_{nm+n}\equiv 0 \operatorname{mod} a_n$ if and only if $a_n\mid a_{nm+n}$. Hence, we have $a_n\mid a_{nm+n}$ (since $a_{nm+n}\equiv 0 \operatorname{mod} a_n$). In view of $nm+n=n(m+1)$, this rewrites as $a_n\mid a_{n(m+1)}$. In other words, (58) holds for $w=m+1$. This completes the induction step. Thus, (58) is proven by induction.

Therefore, Theorem 2.24 **(b)** is proven.

(c) Let u and v be two nonnegative integers satisfying $u\mid v$. We must prove that $a_u\mid a_v$. If $v=0$, then this is obvious (because if $v=0$, then $a_v=a_0=0=0a_u$ and therefore $a_u\mid a_v$). Hence, for the rest of this proof, we can WLOG assume that we don't have $v=0$. Assume this.

Thus, we don't have $v=0$. Hence, $v\neq 0$, so that $v>0$ (since v is nonnegative).

But u divides v (since $u\mid v$). In other words, there exists an integer w such that $v=uw$. Consider this w. If we had $w<0$, then we would have $uw\leq 0$ (since u is nonnegative), which would contradict $uw=v>0$. Hence, we cannot have $w<0$. Thus, we must have $w\geq 0$. Therefore, $w\in\mathbb{N}$. Hence, Theorem 2.24 **(b)** (applied to $n=u$) yields $a_u\mid a_{uw}$. In view of $v=uw$, this rewrites as $a_u\mid a_v$. This proves Theorem 2.24 **(c)**. □

Applying Theorem 2.24 **(c)** to $q = 2$ and $r = 1$, we obtain the observation about divisibility made in Example 2.23.

2.3.2. The Fibonacci sequence and a generalization

Another example of a recursively defined sequence is the famous Fibonacci sequence:

Example 2.25. The *Fibonacci sequence* is the sequence $(f_0, f_1, f_2, \ldots)$ of integers which is defined recursively by

$$f_0 = 0, \qquad f_1 = 1, \qquad \text{and}$$
$$f_n = f_{n-1} + f_{n-2} \qquad \text{for all } n \geq 2.$$

Let us compute its first few entries:

$$\begin{aligned}
f_0 &= 0; \\
f_1 &= 1; \\
f_2 &= \underbrace{f_1}_{=1} + \underbrace{f_0}_{=0} = 1 + 0 = 1; \\
f_3 &= \underbrace{f_2}_{=1} + \underbrace{f_1}_{=1} = 1 + 1 = 2; \\
f_4 &= \underbrace{f_3}_{=2} + \underbrace{f_2}_{=1} = 2 + 1 = 3; \\
f_5 &= \underbrace{f_4}_{=3} + \underbrace{f_3}_{=2} = 3 + 2 = 5; \\
f_6 &= \underbrace{f_5}_{=5} + \underbrace{f_4}_{=3} = 5 + 3 = 8.
\end{aligned}$$

Again, we observe (as in Example 2.23) that $f_2 \mid f_6$ and $f_3 \mid f_6$, which suggests that we might have $f_u \mid f_v$ whenever u and v are two nonnegative integers satisfying $u \mid v$.

Some further experimentation may suggest that the equality $f_{n+m+1} = f_n f_m + f_{n+1} f_{m+1}$ holds for all $n \in \mathbb{N}$ and $m \in \mathbb{N}$.

Both of these conjectures will be shown in the following theorem, in greater generality.

Theorem 2.26. Fix some $a \in \mathbb{Z}$ and $b \in \mathbb{Z}$. Let $(x_0, x_1, x_2, \ldots)$ be a sequence of integers defined recursively by

$$x_0 = 0, \qquad x_1 = 1, \qquad \text{and}$$
$$x_n = a x_{n-1} + b x_{n-2} \qquad \text{for each } n \geq 2.$$

(Note that if $a = 1$ and $b = 1$, then this sequence $(x_0, x_1, x_2, \ldots)$ is precisely the Fibonacci sequence $(f_0, f_1, f_2, \ldots)$ from Example 2.25. If $a = 0$ and $b = 1$, then our sequence $(x_0, x_1, x_2, \ldots)$ is the sequence $\left(0, 1, 0, b, 0, b^2, 0, b^3, \ldots\right)$ that alternates between 0's and powers of b. The reader can easily work out further examples.)

(a) We have $x_{n+m+1} = bx_n x_m + x_{n+1} x_{m+1}$ for all $n \in \mathbb{N}$ and $m \in \mathbb{N}$.

(b) For any $n \in \mathbb{N}$ and $w \in \mathbb{N}$, we have $x_n \mid x_{nw}$.

(c) If u and v are two nonnegative integers satisfying $u \mid v$, then $x_u \mid x_v$.

Before we prove this theorem, let us discuss how **not** to prove it:

Remark 2.27. The proof of Theorem 2.26 **(a)** below illustrates an important aspect of induction proofs: Namely, when devising an induction proof, we often have not only a choice of what variable to induct on (e.g., we could try proving Theorem 2.26 **(a)** by induction on n or by induction on m), but also a choice of whether to leave the other variables fixed. For example, let us try to prove Theorem 2.26 **(a)** by induction on n while leaving the variable m fixed. That is, we fix some $m \in \mathbb{N}$, and we define $\mathcal{A}(n)$ (for each $n \in \mathbb{N}$) to be the following statement:

$$\left(x_{n+m+1} = bx_n x_m + x_{n+1} x_{m+1}\right).$$

Then, it is easy to check that $\mathcal{A}(0)$ holds, so the induction base is complete. For the induction step, we fix some $k \in \mathbb{N}$. (This k serves the role of the "m" in Theorem 2.1, but we cannot call it m here since m already stands for a fixed number.) We assume that $\mathcal{A}(k)$ holds, and we intend to prove $\mathcal{A}(k+1)$.

Our induction hypothesis says that $\mathcal{A}(k)$ holds; in other words, we have $x_{k+m+1} = bx_k x_m + x_{k+1} x_{m+1}$. We want to prove $\mathcal{A}(k+1)$; in other words, we want to prove that $x_{(k+1)+m+1} = bx_{k+1} x_m + x_{(k+1)+1} x_{m+1}$.

A short moment of deliberation shows that we cannot do this (at least not with our current knowledge). There is no direct way of deriving $\mathcal{A}(k+1)$ from $\mathcal{A}(k)$. **However**, if we knew that the statement $\mathcal{A}(k)$ holds "for $m+1$ instead of m" (that is, if we knew that $x_{k+(m+1)+1} = bx_k x_{m+1} + x_{k+1} x_{(m+1)+1}$), then we could derive $\mathcal{A}(k+1)$. But we cannot just "apply $\mathcal{A}(k)$ to $m+1$ instead of m"; after all, m is a fixed number, so we cannot have it take different values in $\mathcal{A}(k)$ and in $\mathcal{A}(k+1)$.

So we are at an impasse. We got into this impasse by fixing m. So let us try **not** fixing $m \in \mathbb{N}$ right away, but instead defining $\mathcal{A}(n)$ (for each $n \in \mathbb{N}$) to be the following statement:

$$\left(x_{n+m+1} = bx_n x_m + x_{n+1} x_{m+1} \text{ for all } m \in \mathbb{N}\right).$$

Thus, $\mathcal{A}(n)$ is not a statement about a specific integer m any more, but rather a statement about all nonnegative integers m. This allows us to apply $\mathcal{A}(k)$ to $m+1$ instead of m in the induction step. (We can still fix $m \in \mathbb{N}$ **during the induction step**; this doesn't prevent us from applying $\mathcal{A}(k)$ to $m+1$ instead of m, since $\mathcal{A}(k)$ has been formulated before m was fixed.) This way, we arrive at the following proof:

Proof of Theorem 2.26. **(a)** We claim that for each $n \in \mathbb{N}$, we have

$$\left(x_{n+m+1} = bx_n x_m + x_{n+1} x_{m+1} \text{ for all } m \in \mathbb{N}\right). \tag{59}$$

Indeed, let us prove (59) by induction on n:

Induction base: We have $x_{0+m+1} = bx_0 x_m + x_{0+1} x_{m+1}$ for all $m \in \mathbb{N}$ [45]. In other words, (59) holds for $n = 0$. This completes the induction base.

Induction step: Let $k \in \mathbb{N}$. Assume that (59) holds for $n = k$. We must prove that (59) holds for $n = k + 1$.

We have assumed that (59) holds for $n = k$. In other words, we have

$$\left(x_{k+m+1} = bx_k x_m + x_{k+1} x_{m+1} \text{ for all } m \in \mathbb{N}\right). \tag{60}$$

Now, let $m \in \mathbb{N}$. We have $m + 2 \geq 2$; thus, the recursive definition of the sequence $(x_0, x_1, x_2, \ldots)$ yields

$$x_{m+2} = a \underbrace{x_{(m+2)-1}}_{=x_{m+1}} + b \underbrace{x_{(m+2)-2}}_{=x_m} = ax_{m+1} + bx_m. \tag{61}$$

The same argument (with m replaced by k) yields

$$x_{k+2} = ax_{k+1} + bx_k. \tag{62}$$

But we can apply (60) to $m + 1$ instead of m. Thus, we obtain

$$\begin{aligned}
x_{k+(m+1)+1} &= bx_k x_{m+1} + x_{k+1} \underbrace{x_{(m+1)+1}}_{\substack{=x_{m+2}=ax_{m+1}+bx_m \\ \text{(by (61))}}} \\
&= bx_k x_{m+1} + \underbrace{x_{k+1}\left(ax_{m+1} + bx_m\right)}_{=ax_{k+1}x_{m+1}+bx_{k+1}x_m} = \underbrace{bx_k x_{m+1} + ax_{k+1} x_{m+1}}_{=(ax_{k+1}+bx_k)x_{m+1}} + bx_{k+1} x_m \\
&= \underbrace{\left(ax_{k+1} + bx_k\right)}_{\substack{=x_{k+2} \\ \text{(by (62))}}} x_{m+1} + bx_{k+1} x_m = x_{k+2} x_{m+1} + bx_{k+1} x_m \\
&= bx_{k+1} x_m + \underbrace{x_{k+2}}_{=x_{(k+1)+1}} x_{m+1} = bx_{k+1} x_m + x_{(k+1)+1} x_{m+1}.
\end{aligned}$$

In view of $k + (m + 1) + 1 = (k + 1) + m + 1$, this rewrites as

$$x_{(k+1)+m+1} = bx_{k+1} x_m + x_{(k+1)+1} x_{m+1}.$$

Now, forget that we fixed m. We thus have shown that $x_{(k+1)+m+1} = bx_{k+1} x_m + x_{(k+1)+1} x_{m+1}$ for all $m \in \mathbb{N}$. In other words, (59) holds for $n = k + 1$. This completes the induction step. Thus, (59) is proven.

[45] *Proof.* Let $m \in \mathbb{N}$. Then, $x_{0+m+1} = x_{m+1}$. Comparing this with $b \underbrace{x_0}_{=0} x_m + \underbrace{x_{0+1}}_{=x_1=1} x_{m+1} = b0x_m + 1x_{m+1} = x_{m+1}$, we obtain $x_{0+m+1} = bx_0 x_m + x_{0+1} x_{m+1}$, qed.

Hence, Theorem 2.26 **(a)** holds.

(b) Fix $n \in \mathbb{N}$. We claim that

$$x_n \mid x_{nw} \qquad \text{for each } w \in \mathbb{N}. \tag{63}$$

Indeed, let us prove (63) by induction on w:

Induction base: We have $x_{n \cdot 0} = x_0 = 0 = 0x_n$ and thus $x_n \mid x_{n \cdot 0}$. In other words, (63) holds for $w = 0$. This completes the induction base.

Induction step: Let $k \in \mathbb{N}$. Assume that (63) holds for $w = k$. We must now prove that (63) holds for $w = k+1$. In other words, we must prove that $x_n \mid x_{n(k+1)}$.

If $n = 0$, then this is true[46]. Hence, for the rest of this proof, we can WLOG assume that we don't have $n = 0$. Assume this.

We have assumed that (63) holds for $w = k$. In other words, we have $x_n \mid x_{nk}$. In other words, $x_{nk} \equiv 0 \bmod x_n$. [47] Likewise, from $x_n \mid x_n$, we obtain $x_n \equiv 0 \bmod x_n$.

We have $n \in \mathbb{N}$ but $n \neq 0$ (since we don't have $n = 0$). Hence, n is a positive integer. Thus, $n - 1 \in \mathbb{N}$. Therefore, Theorem 2.26 **(a)** (applied to nk and $n-1$ instead of n and m) yields

$$x_{nk+(n-1)+1} = bx_{nk}x_{n-1} + x_{nk+1}x_{(n-1)+1}.$$

In view of $nk + (n-1) + 1 = n(k+1)$, this rewrites as

$$x_{n(k+1)} = b \underbrace{x_{nk}}_{\equiv 0 \bmod x_n} x_{n-1} + x_{nk+1} \underbrace{x_{(n-1)+1}}_{= x_n \equiv 0 \bmod x_n} \equiv b0x_{n-1} + x_{nk+1}0 = 0 \bmod x_n.$$

[48] Thus, we have shown that $x_{n(k+1)} \equiv 0 \bmod x_n$. In other words, $x_n \mid x_{n(k+1)}$ (again, this follows from Proposition 2.11 **(a)**). In other words, (63) holds for $w = k+1$. This completes the induction step. Hence, (63) is proven by induction.

[46] *Proof.* Let us assume that $n = 0$. Then, $x_{n(k+1)} = x_{0(k+1)} = x_0 = 0 = 0x_n$, and thus $x_n \mid x_{n(k+1)}$, qed.

[47] Here, again, we have used Proposition 2.11 **(a)** (applied to x_{nk} and x_n instead of a and n). This argument is simple enough that we will leave it unsaid in the future.

[48] We have used substitutivity for congruences in this computation. Here is, again, a way to rewrite it without this use:

We have $x_{n(k+1)} = bx_{nk}x_{n-1} + x_{nk+1}x_{(n-1)+1}$. But $b \equiv b \bmod x_n$ (by Proposition 2.12 **(a)**) and $x_{n-1} \equiv x_{n-1} \bmod x_n$ (for the same reason) and $x_{nk+1} \equiv x_{nk+1} \bmod x_n$ (for the same reason). Now, Proposition 2.21 **(c)** (applied to b, b, x_{nk}, 0 and x_n instead of a, b, c, d and n) yields $bx_{nk} \equiv b0 \bmod x_n$ (since $b \equiv b \bmod x_n$ and $x_{nk} \equiv 0 \bmod x_n$). Hence, Proposition 2.21 **(c)** (applied to bx_{nk}, $b0$, x_{n-1}, x_{n-1} and x_n instead of a, b, c, d and n) yields $bx_{nk}x_{n-1} \equiv b0x_{n-1} \bmod x_n$ (since $bx_{nk} \equiv b0 \bmod x_n$ and $x_{n-1} \equiv x_{n-1} \bmod x_n$). Also, $x_{nk+1}x_{(n-1)+1} \equiv x_{nk+1}x_{(n-1)+1} \bmod x_n$ (by Proposition 2.12 **(a)**). Hence, Proposition 2.21 **(a)** (applied to $bx_{nk}x_{n-1}$, $b0x_{n-1}$, $x_{nk+1}x_{(n-1)+1}$, $x_{nk+1}x_{(n-1)+1}$ and x_n instead of a, b, c, d and n) yields

$$bx_{nk}x_{n-1} + x_{nk+1}x_{(n-1)+1} \equiv b0x_{n-1} + x_{nk+1}x_{(n-1)+1} \bmod x_n$$

(since $bx_{nk}x_{n-1} \equiv b0x_{n-1} \bmod x_n$ and $x_{nk+1}x_{(n-1)+1} \equiv x_{nk+1}x_{(n-1)+1} \bmod x_n$).

Also, Proposition 2.21 **(c)** (applied to x_{nk+1}, x_{nk+1}, $x_{(n-1)+1}$, 0 and x_n instead of a, b, c, d and n) yields $x_{nk+1}x_{(n-1)+1} \equiv x_{nk+1}0 \bmod x_n$ (since $x_{nk+1} \equiv x_{nk+1} \bmod x_n$ and $x_{(n-1)+1} = x_n \equiv 0 \bmod x_n$). Furthermore, $b0x_{n-1} \equiv b0x_{n-1} \bmod x_n$ (by Proposition 2.12 **(a)**). Finally, Proposition

This proves Theorem 2.26 **(b)**.

(c) Theorem 2.26 **(c)** can be derived from Theorem 2.26 **(b)** in the same way as Theorem 2.24 **(c)** was derived from Theorem 2.24 **(b)**. □

Applying Theorem 2.26 **(a)** to $a = 1$ and $b = 1$, we obtain the equality $f_{n+m+1} = f_n f_m + f_{n+1} f_{m+1}$ noticed in Example 2.25. Applying Theorem 2.26 **(c)** to $a = 1$ and $b = 1$, we obtain the observation about divisibility made in Example 2.25.

Note that part **(a)** of Theorem 2.26 still works if a and b are real numbers (instead of being integers). But of course, in this case, $(x_0, x_1, x_2, \ldots)$ will be merely a sequence of real numbers (rather than a sequence of integers), and thus parts **(b)** and **(c)** of Theorem 2.26 will no longer make sense (since divisibility is only defined for integers).

2.4. The sum of the first n positive integers

We now come to one of the most classical examples of a proof by induction: Namely, we shall prove the fact that for each $n \in \mathbb{N}$, the sum of the first n positive integers (that is, the sum $1 + 2 + \cdots + n$) equals $\dfrac{n(n+1)}{2}$. However, there is a catch here, which is easy to overlook if one isn't trying to be completely rigorous: We don't really know yet whether there is such a thing as "the sum of the first n positive integers"! To be more precise, we have introduced the $\sum$ sign in Section 1.4, which would allow us to define the sum of the first n positive integers (as $\sum\limits_{i=1}^{n} i$); but our definition of the $\sum$ sign relied on a fact which we have not proved yet (namely, the fact that the right hand side of (1) does not depend on the choice of t). We shall prove this fact later (Theorem 2.118 **(a)**), but for now we prefer not to use it. Instead, let us replace the notion of "the sum of the first n positive integers" by a recursively defined sequence:

Proposition 2.28. Let $(t_0, t_1, t_2, \ldots)$ be a sequence of integers defined recursively by

$$t_0 = 0, \qquad \text{and}$$
$$t_n = t_{n-1} + n \qquad \text{for each } n \geq 1.$$

2.21 **(a)** (applied to $b0x_{n-1}$, $b0x_{n-1}$, $x_{nk+1}x_{(n-1)+1}$, $x_{nk+1}0$ and x_n instead of a, b, c, d and n) yields

$$b0x_{n-1} + x_{nk+1}x_{(n-1)+1} \equiv b0x_{n-1} + x_{nk+1}0 \mod x_n$$

(since $b0x_{n-1} \equiv b0x_{n-1} \mod x_n$ and $x_{nk+1}x_{(n-1)+1} \equiv x_{nk+1}0 \mod x_n$). Thus,

$$\begin{aligned} x_{n(k+1)} &= bx_{nk}x_{n-1} + x_{nk+1}x_{(n-1)+1} \equiv b0x_{n-1} + x_{nk+1}x_{(n-1)+1} \\ &\equiv b0x_{n-1} + x_{nk+1}0 = 0 \mod x_n. \end{aligned}$$

So we have proven that $x_{n(k+1)} \equiv 0 \mod x_n$.

Then,

$$t_n = \frac{n(n+1)}{2} \qquad \text{for each } n \in \mathbb{N}. \tag{64}$$

The sequence $(t_0, t_1, t_2, \ldots)$ defined in Proposition 2.28 is known as the *sequence of triangular numbers*. Its definition shows that

$$\begin{aligned}
t_0 &= 0; \\
t_1 &= \underbrace{t_0}_{=0} + 1 = 0 + 1 = 1; \\
t_2 &= \underbrace{t_1}_{=1} + 2 = 1 + 2; \\
t_3 &= \underbrace{t_2}_{=1+2} + 3 = (1+2) + 3; \\
t_4 &= \underbrace{t_3}_{=(1+2)+3} + 4 = ((1+2)+3) + 4; \\
t_5 &= \underbrace{t_4}_{=((1+2)+3)+4} + 5 = (((1+2)+3)+4) + 5
\end{aligned}$$

[49] and so on; this explains why it makes sense to think of t_n as the sum of the first n positive integers. (This is legitimate even when $n = 0$, because the sum of the first 0 positive integers is an empty sum, and an empty sum is always defined to be equal to 0.) Once we have convinced ourselves that "the sum of the first n positive integers" is a well-defined concept, it will be easy to see (by induction) that t_n **is** the sum of the first n positive integers whenever $n \in \mathbb{N}$. Therefore, Proposition 2.28 will tell us that the sum of the first n positive integers equals $\frac{n(n+1)}{2}$ whenever $n \in \mathbb{N}$.

For now, let us prove Proposition 2.28:

Proof of Proposition 2.28. We shall prove (64) by induction on n:

Induction base: Comparing $t_0 = 0$ with $\frac{0(0+1)}{2} = 0$, we obtain $t_0 = \frac{0(0+1)}{2}$. In other words, (64) holds for $n = 0$. This completes the induction base.

Induction step: Let $m \in \mathbb{N}$. Assume that (64) holds for $n = m$. We must prove that (64) holds for $n = m + 1$.

We have assumed that (64) holds for $n = m$. In other words, we have $t_m = \frac{m(m+1)}{2}$.

[49] Note that we write "$(((1+2)+3)+4)+5$" and not "$1+2+3+4+5$". The reason for this is that we haven't proven yet that the expression "$1+2+3+4+5$" is well-defined. (This expression **is** well-defined, but this will only be clear once we have proven Theorem 2.118 **(a)** below.)

Recall that $t_n = t_{n-1} + n$ for each $n \geq 1$. Applying this to $n = m+1$, we obtain

$$\begin{aligned} t_{m+1} &= \underbrace{t_{(m+1)-1}}_{=t_m = \dfrac{m(m+1)}{2}} + (m+1) = \frac{m(m+1)}{2} + (m+1) = \frac{m(m+1) + 2(m+1)}{2} \\ &= \frac{(m+2)(m+1)}{2} = \frac{(m+1)(m+2)}{2} = \frac{(m+1)((m+1)+1)}{2} \end{aligned}$$

(since $m + 2 = (m+1) + 1$). In other words, (64) holds for $n = m+1$. This completes the induction step. Hence, (64) is proven by induction. This proves Proposition 2.28. □

2.5. Induction on a derived quantity: maxima of sets

2.5.1. Defining maxima

We have so far been applying the Induction Principle in fairly obvious ways: With the exception of our proof of Proposition 2.16, we have mostly been doing induction on a variable (n or k or i) that already appeared in the claim that we were proving. But sometimes, it is worth doing induction on a variable that does **not** explicitly appear in this claim (which, formally speaking, means that we introduce a new variable to do induction on). For example, the claim might be saying "Each nonempty finite set S of integers has a largest element", and we prove it by induction on $|S| - 1$. This means that instead of directly proving the claim itself, we rather prove the equivalent claim "For each $n \in \mathbb{N}$, each nonempty finite set S of integers satisfying $|S| - 1 = n$ has a largest element" by induction on n. We shall show this proof in more detail below (see Theorem 2.35). First, we prepare by discussing largest elements of sets in general.

Definition 2.29. Let S be a set of integers (or rational numbers, or real numbers). A *maximum* of S is defined to be an element $s \in S$ that satisfies

$$(s \geq t \text{ for each } t \in S).$$

In other words, a maximum of S is defined to be an element of S which is greater or equal to each element of S.

(The plural of the word "maximum" is "maxima".)

Example 2.30. The set $\{2, 4, 5\}$ has exactly one maximum: namely, 5.

The set $\mathbb{N} = \{0, 1, 2, \ldots\}$ has no maximum: If k was a maximum of $\mathbb{N}$, then we would have $k \geq k + 1$, which is absurd.

The set $\{0, -1, -2, \ldots\}$ has a maximum: namely, 0.

The set $\varnothing$ has no maximum, since a maximum would have to be an element of $\varnothing$.

In Theorem 2.35, we shall soon show that every nonempty finite set of integers has a maximum. First, we prove that a maximum is unique if it exists:

Proposition 2.31. Let S be a set of integers (or rational numbers, or real numbers). Then, S has **at most one** maximum.

Proof of Proposition 2.31. Let s_1 and s_2 be two maxima of S. We shall show that $s_1 = s_2$.

Indeed, s_1 is a maximum of S. In other words, s_1 is an element $s \in S$ that satisfies $(s \geq t \text{ for each } t \in S)$ (by the definition of a maximum). In other words, s_1 is an element of S and satisfies

$$(s_1 \geq t \text{ for each } t \in S) . \tag{65}$$

The same argument (applied to s_2 instead of s_1) shows that s_2 is an element of S and satisfies

$$(s_2 \geq t \text{ for each } t \in S) . \tag{66}$$

Now, s_1 is an element of S. Hence, (66) (applied to $t = s_1$) yields $s_2 \geq s_1$. But the same argument (with the roles of s_1 and s_2 interchanged) shows that $s_1 \geq s_2$. Combining this with $s_2 \geq s_1$, we obtain $s_1 = s_2$.

Now, forget that we fixed s_1 and s_2. We thus have shown that if s_1 and s_2 are two maxima of S, then $s_1 = s_2$. In other words, any two maxima of S are equal. In other words, S has **at most one** maximum. This proves Proposition 2.31. □

Definition 2.32. Let S be a set of integers (or rational numbers, or real numbers). Proposition 2.31 shows that S has **at most one** maximum. Thus, if S has a maximum, then this maximum is the unique maximum of S; we shall thus call it *the maximum* of S or *the largest element* of S. We shall denote this maximum by $\max S$.

Thus, if S is a set of integers (or rational numbers, or real numbers) that has a maximum, then this maximum $\max S$ satisfies

$$\max S \in S \tag{67}$$

and

$$(\max S \geq t \text{ for each } t \in S) \tag{68}$$

(because of the definition of a maximum).

Let us next show two simple facts:

Lemma 2.33. Let x be an integer (or rational number, or real number). Then, the set $\{x\}$ has a maximum, namely x.

Proof of Lemma 2.33. Clearly, $x \geq x$. Thus, $x \geq t$ for each $t \in \{x\}$ (because the only $t \in \{x\}$ is x). In other words, x is an element $s \in \{x\}$ that satisfies $(s \geq t \text{ for each } t \in \{x\})$ (since $x \in \{x\}$).

But recall that a maximum of $\{x\}$ means an element $s \in \{x\}$ that satisfies $(s \geq t \text{ for each } t \in \{x\})$ (by the definition of a maximum). Hence, x is a maximum of $\{x\}$ (since x is such an element). Thus, the set $\{x\}$ has a maximum, namely x. This proves Lemma 2.33. □

Proposition 2.34. Let P and Q be two sets of integers (or rational numbers, or real numbers). Assume that P has a maximum, and assume that Q has a maximum. Then, the set $P \cup Q$ has a maximum.

Proof of Proposition 2.34. We know that P has a maximum; it is denoted by $\max P$. We also know that Q has a maximum; it is denoted by $\max Q$. The sets P and Q play symmetric roles in Proposition 2.34 (since $P \cup Q = Q \cup P$). Thus, we can WLOG assume that $\max P \geq \max Q$ (since otherwise, we can simply swap P with Q, without altering the meaning of Proposition 2.34). Assume this.

Now, (67) (applied to $S = P$) shows that $\max P \in P \subseteq P \cup Q$. Furthermore, we claim that

$$(\max P \geq t \text{ for each } t \in P \cup Q). \tag{69}$$

[*Proof of (69):* Let $t \in P \cup Q$. We must show that $\max P \geq t$.

We have $t \in P \cup Q$. In other words, $t \in P$ or $t \in Q$. Hence, we are in one of the following two cases:

Case 1: We have $t \in P$.

Case 2: We have $t \in Q$.

(These two cases might have overlap, but there is nothing wrong about this.)

Let us first consider Case 1. In this case, we have $t \in P$. Hence, (68) (applied to $S = P$) yields $\max P \geq t$. Hence, $\max P \geq t$ is proven in Case 1.

Let us next consider Case 2. In this case, we have $t \in Q$. Hence, (68) (applied to $S = Q$) yields $\max Q \geq t$. Hence, $\max P \geq \max Q \geq t$. Thus, $\max P \geq t$ is proven in Case 2.

We have now proven $\max P \geq t$ in each of the two Cases 1 and 2. Since these two Cases cover all possibilities, we thus conclude that $\max P \geq t$ always holds. This proves (69).]

Now, $\max P$ is an element $s \in P \cup Q$ that satisfies $(s \geq t \text{ for each } t \in P \cup Q)$ (since $\max P \in P \cup Q$ and $(\max P \geq t \text{ for each } t \in P \cup Q)$).

But recall that a maximum of $P \cup Q$ means an element $s \in P \cup Q$ that satisfies $(s \geq t \text{ for each } t \in P \cup Q)$ (by the definition of a maximum). Hence, $\max P$ is a maximum of $P \cup Q$ (since $\max P$ is such an element). Thus, the set $P \cup Q$ has a maximum. This proves Proposition 2.34. □

2.5.2. Nonempty finite sets of integers have maxima

Theorem 2.35. Let S be a nonempty finite set of integers. Then, S has a maximum.

First proof of Theorem 2.35. First of all, let us forget that we fixed S. So we want to prove that if S is a nonempty finite set of integers, then S has a maximum.

For each $n \in \mathbb{N}$, we let $\mathcal{A}(n)$ be the statement

$$\left(\begin{array}{c}\text{if } S \text{ is a nonempty finite set of integers satisfying } |S|-1=n,\\ \text{then } S \text{ has a maximum}\end{array}\right).$$

We claim that $\mathcal{A}(n)$ holds for all $n \in \mathbb{N}$.

Indeed, let us prove this by induction on n:

Induction base: If S is a nonempty finite set of integers satisfying $|S|-1=0$, then S has a maximum[50]. But this is exactly the statement $\mathcal{A}(0)$. Hence, $\mathcal{A}(0)$ holds. This completes the induction base.

Induction step: Let $m \in \mathbb{N}$. Assume that $\mathcal{A}(m)$ holds. We shall now show that $\mathcal{A}(m+1)$ holds.

We have assumed that $\mathcal{A}(m)$ holds. In other words,

$$\left(\begin{array}{c}\text{if } S \text{ is a nonempty finite set of integers satisfying } |S|-1=m,\\ \text{then } S \text{ has a maximum}\end{array}\right) \qquad (70)$$

(because this is what the statement $\mathcal{A}(m)$ says).

Now, let S be a nonempty finite set of integers satisfying $|S|-1=m+1$. There exists some $t \in S$ (since S is nonempty). Consider this t. We have $(S \setminus \{t\}) \cup \{t\} = S \cup \{t\} = S$ (since $t \in S$).

From $t \in S$, we obtain $|S \setminus \{t\}| = |S|-1 = m+1 > m \geq 0$ (since $m \in \mathbb{N}$). Hence, the set $S \setminus \{t\}$ is nonempty. Furthermore, this set $S \setminus \{t\}$ is finite (since S is finite) and satisfies $|S \setminus \{t\}|-1 = m$ (since $|S \setminus \{t\}| = m+1$). Hence, (70) (applied to $S \setminus \{t\}$ instead of S) shows that $S \setminus \{t\}$ has a maximum. Also, Lemma 2.33 (applied to $x=t$) shows that the set $\{t\}$ has a maximum, namely t. Hence, Proposition 2.34 (applied to $P = S \setminus \{t\}$ and $Q = \{t\}$) shows that the set $(S \setminus \{t\}) \cup \{t\}$ has a maximum. Since $(S \setminus \{t\}) \cup \{t\} = S$, this rewrites as follows: The set S has a maximum.

Now, forget that we fixed S. We thus have shown that if S is a nonempty finite set of integers satisfying $|S|-1=m+1$, then S has a maximum. But this is precisely the statement $\mathcal{A}(m+1)$. Hence, we have shown that $\mathcal{A}(m+1)$ holds. This completes the induction step.

Thus, we have proven (by induction) that $\mathcal{A}(n)$ holds for all $n \in \mathbb{N}$. In other words, for all $n \in \mathbb{N}$, the following holds:

$$\left(\begin{array}{c}\text{if } S \text{ is a nonempty finite set of integers satisfying } |S|-1=n,\\ \text{then } S \text{ has a maximum}\end{array}\right) \qquad (71)$$

[50] *Proof.* Let S be a nonempty finite set of integers satisfying $|S|-1=0$. We must show that S has a maximum.

Indeed, $|S|=1$ (since $|S|-1=0$). In other words, S is a 1-element set. In other words, $S=\{x\}$ for some integer x. Consider this x. Lemma 2.33 shows that the set $\{x\}$ has a maximum. In other words, the set S has a maximum (since $S=\{x\}$). This completes our proof.

(because this is what $\mathcal{A}(n)$ says).

Now, let S be a nonempty finite set of integers. We shall prove that S has a maximum.

Indeed, $|S| \in \mathbb{N}$ (since S is finite) and $|S| > 0$ (since S is nonempty); hence, $|S| \geq 1$. Thus, $|S| - 1 \geq 0$, so that $|S| - 1 \in \mathbb{N}$. Hence, we can define $n \in \mathbb{N}$ by $n = |S| - 1$. Consider this n. Thus, $|S| - 1 = n$. Hence, (71) shows that S has a maximum. This proves Theorem 2.35. □

2.5.3. Conventions for writing induction proofs on derived quantities

Let us take a closer look at the proof we just gave. The definition of the statement $\mathcal{A}(n)$ was not exactly unmotivated: This statement simply says that Theorem 2.35 holds under the condition that $|S| - 1 = n$. Thus, by introducing $\mathcal{A}(n)$, we have "sliced" Theorem 2.35 into a sequence of statements $\mathcal{A}(0), \mathcal{A}(1), \mathcal{A}(2), \ldots$, which then allowed us to prove these statements by induction on n even though no "n" appeared in Theorem 2.35 itself. This kind of strategy applies to various other problems. Again, we don't need to explicitly define the statement $\mathcal{A}(n)$ if it is simply saying that the claim we are trying to prove (in our case, Theorem 2.35) holds under the condition that $|S| - 1 = n$; we can just say that we are doing "induction on $|S| - 1$". More generally:

Convention 2.36. Let $\mathcal{B}$ be a logical statement that involves some variables $v_1, v_2, v_3, \ldots$. (For example, $\mathcal{B}$ can be the statement of Theorem 2.35; then, there is only one variable, namely S.)

Let q be some expression (involving the variables $v_1, v_2, v_3, \ldots$ or some of them) that has the property that whenever the variables $v_1, v_2, v_3, \ldots$ satisfy the assumptions of $\mathcal{B}$, the expression q evaluates to some nonnegative integer. (For example, if $\mathcal{B}$ is the statement of Theorem 2.35, then q can be the expression $|S| - 1$, because it is easily seen that if S is a nonempty finite set of integers, then $|S| - 1$ is a nonnegative integer.)

Assume that you want to prove the statement $\mathcal{B}$. Then, you can proceed as follows: For each $n \in \mathbb{N}$, define $\mathcal{A}(n)$ to be the statement saying that[51]

$$(\text{the statement } \mathcal{B} \text{ holds under the condition that } q = n).$$

Then, prove $\mathcal{A}(n)$ by induction on n. Thus:

- The *induction base* consists in proving that the statement $\mathcal{B}$ holds under the condition that $q = 0$.
- The *induction step* consists in fixing $m \in \mathbb{N}$, and showing that if the statement $\mathcal{B}$ holds under the condition that $q = m$, then the statement $\mathcal{B}$ holds under the condition that $q = m + 1$.

Once this induction proof is finished, it immediately follows that the statement $\mathcal{B}$ always holds.

This strategy of proof is called "induction on q" (or "induction over q"). Once you have specified what q is, you don't need to explicitly define $\mathcal{A}(n)$, nor do you ever need to mention n.

Using this convention, we can rewrite our above proof of Theorem 2.35 as follows:

First proof of Theorem 2.35 (second version). It is easy to see that $|S| - 1 \in \mathbb{N}$ [52]. Hence, we can apply induction on $|S| - 1$ to prove Theorem 2.35:

Induction base: Theorem 2.35 holds under the condition that $|S| - 1 = 0$ [53]. This completes the induction base.

Induction step: Let $m \in \mathbb{N}$. Assume that Theorem 2.35 holds under the condition that $|S| - 1 = m$. We shall now show that Theorem 2.35 holds under the condition that $|S| - 1 = m + 1$.

We have assumed that Theorem 2.35 holds under the condition that $|S| - 1 = m$. In other words,

$$\left(\begin{array}{c}\text{if } S \text{ is a nonempty finite set of integers satisfying } |S| - 1 = m, \\ \text{then } S \text{ has a maximum}\end{array}\right). \qquad (72)$$

Now, let S be a nonempty finite set of integers satisfying $|S| - 1 = m + 1$. There exists some $t \in S$ (since S is nonempty). Consider this t. We have $(S \setminus \{t\}) \cup \{t\} = S \cup \{t\} = S$ (since $t \in S$).

From $t \in S$, we obtain $|S \setminus \{t\}| = |S| - 1 = m + 1 > m \geq 0$ (since $m \in \mathbb{N}$). Hence, the set $S \setminus \{t\}$ is nonempty. Furthermore, this set $S \setminus \{t\}$ is finite (since S is finite) and satisfies $|S \setminus \{t\}| - 1 = m$ (since $|S \setminus \{t\}| = m + 1$). Hence, (72) (applied to $S \setminus \{t\}$ instead of S) shows that $S \setminus \{t\}$ has a maximum. Also, Lemma 2.33 (applied to $x = t$) shows that the set $\{t\}$ has a maximum, namely t. Hence, Proposition 2.34 (applied to $P = S \setminus \{t\}$ and $Q = \{t\}$) shows that the set $(S \setminus \{t\}) \cup \{t\}$ has a maximum. Since $(S \setminus \{t\}) \cup \{t\} = S$, this rewrites as follows: The set S has a maximum.

Now, forget that we fixed S. We thus have shown that if S is a nonempty finite set of integers satisfying $|S| - 1 = m + 1$, then S has a maximum. In other words, Theorem 2.35 holds under the condition that $|S| - 1 = m + 1$. This completes the induction step. Thus, the induction proof of Theorem 2.35 is complete. □

[51] We assume that no variable named "n" appears in the statement $\mathcal{B}$; otherwise, we need a different letter for our new variable in order to avoid confusion.

[52] *Proof.* We have $|S| \in \mathbb{N}$ (since S is finite) and $|S| > 0$ (since S is nonempty); hence, $|S| \geq 1$. Thus, $|S| - 1 \in \mathbb{N}$, qed.

[53] *Proof.* Let S be as in Theorem 2.35, and assume that $|S| - 1 = 0$. We must show that the claim of Theorem 2.35 holds.

Indeed, $|S| = 1$ (since $|S| - 1 = 0$). In other words, S is a 1-element set. In other words, $S = \{x\}$ for some integer x. Consider this x. Lemma 2.33 shows that the set $\{x\}$ has a maximum. In other words, the set S has a maximum (since $S = \{x\}$). In other words, the claim of Theorem 2.35 holds. This completes our proof.

We could have shortened this proof even further if we didn't explicitly state (72), but rather (instead of applying (72)) said that "we can apply Theorem 2.35 to $S \setminus \{t\}$ instead of S".

Let us stress again that, in order to prove Theorem 2.35 by induction on $|S| - 1$, we had to check that $|S| - 1 \in \mathbb{N}$ whenever S satisfies the assumptions of Theorem 2.35.[54] This check was necessary. For example, if we had instead tried to proceed by induction on $|S| - 2$, then we would only have proven Theorem 2.35 under the condition that $|S| - 2 \in \mathbb{N}$; but this condition isn't always satisfied (indeed, it misses the case when S is a 1-element set).

2.5.4. Vacuous truth and induction bases

Can we also prove Theorem 2.35 by induction on $|S|$ (instead of $|S| - 1$)? This seems a bit strange, since $|S|$ can never be 0 in Theorem 2.35 (because S is required to be nonempty), so that the induction base would be talking about a situation that never occurs. However, there is nothing wrong about it, and we already do talk about such situations oftentimes (for example, every time we make a proof by contradiction). The following concept from basic logic explains this:

Convention 2.37. **(a)** A logical statement of the form "if $\mathcal{A}$, then $\mathcal{B}$" (where $\mathcal{A}$ and $\mathcal{B}$ are two statements) is said to be *vacuously true* if $\mathcal{A}$ does not hold. For example, the statement "if $0 = 1$, then every set is empty" is vacuously true, because $0 = 1$ is false. The statement "if $0 = 1$, then $1 = 1$" is also vacuously true, although its truth can also be seen as a consequence of the fact that $1 = 1$ is true.

By the laws of logic, a vacuously true statement is always true! This may sound counterintuitive, but actually makes perfect sense: A statement "if $\mathcal{A}$, then $\mathcal{B}$" only says anything about situations where $\mathcal{A}$ holds. If $\mathcal{A}$ never holds, then it therefore says nothing. And when you are saying nothing, you are certainly not lying.

The principle that a vacuously true statement always holds is known as "*ex falso quodlibet*" (literal translation: "from the false, anything") or "*principle of explosion*". It can be restated as follows: From a false statement, any statement follows.

(b) Now, let X be a set, and let $\mathcal{A}(x)$ and $\mathcal{B}(x)$ be two statements defined for each $x \in X$. A statement of the form "for each $x \in X$ satisfying $\mathcal{A}(x)$, we have $\mathcal{B}(x)$" will automatically hold if there exists no $x \in X$ satisfying $\mathcal{A}(x)$. (Indeed, this statement can be rewritten as "for each $x \in X$, we have (if $\mathcal{A}(x)$, then $\mathcal{B}(x)$)"; but this holds because the statement "if $\mathcal{A}(x)$, then $\mathcal{B}(x)$" is vacuously true for each $x \in X$.) Such a statement will also be called *vacuously true*.

[54] In our first version of the above proof, we checked this at the end; in the second version, we checked it at the beginning of the proof.

For example, the statement "if $n \in \mathbb{N}$ is both odd and even, then $n = n + 1$" is vacuously true, since no $n \in \mathbb{N}$ can be both odd and even at the same time.

(c) Now, let X be the empty set (that is, $X = \varnothing$), and let $\mathcal{B}(x)$ be a statement defined for each $x \in X$. Then, a statement of the form "for each $x \in X$, we have $\mathcal{B}(x)$" will automatically hold. (Indeed, this statement can be rewritten as "for each $x \in X$, we have (if $x \in X$, then $\mathcal{B}(x)$)"; but this holds because the statement "if $x \in X$, then $\mathcal{B}(x)$" is vacuously true for each $x \in X$, since its premise ($x \in X$) is false.) Again, such a statement is said to be *vacuously true*.

For example, the statement "for each $x \in \varnothing$, we have $x \neq x$" is vacuously true (because there exists no $x \in \varnothing$).

Thus, if we try to prove Theorem 2.35 by induction on $|S|$, then the induction base becomes vacuously true. However, the induction step becomes more complicated, since we can no longer argue that $S \setminus \{t\}$ is nonempty, but instead have to account for the case when $S \setminus \{t\}$ is empty as well. So we gain and we lose at the same time. Here is how this proof looks like:

Second proof of Theorem 2.35. Clearly, $|S| \in \mathbb{N}$ (since S is a finite set). Hence, we can apply induction on $|S|$ to prove Theorem 2.35:

Induction base: Theorem 2.35 holds under the condition that $|S| = 0$ [55]. This completes the induction base.

Induction step: Let $m \in \mathbb{N}$. Assume that Theorem 2.35 holds under the condition that $|S| = m$. We shall now show that Theorem 2.35 holds under the condition that $|S| = m + 1$.

We have assumed that Theorem 2.35 holds under the condition that $|S| = m$. In other words,

$$\left(\begin{array}{c} \text{if } S \text{ is a nonempty finite set of integers satisfying } |S| = m, \\ \text{then } S \text{ has a maximum} \end{array} \right). \tag{73}$$

Now, let S be a nonempty finite set of integers satisfying $|S| = m + 1$. We want to prove that S has a maximum.

There exists some $t \in S$ (since S is nonempty). Consider this t. We have $(S \setminus \{t\}) \cup \{t\} = S \cup \{t\} = S$ (since $t \in S$). Lemma 2.33 (applied to $x = t$) shows that the set $\{t\}$ has a maximum, namely t.

We are in one of the following two cases:

Case 1: We have $S \setminus \{t\} = \varnothing$.

Case 2: We have $S \setminus \{t\} \neq \varnothing$.

Let us first consider Case 1. In this case, we have $S \setminus \{t\} = \varnothing$. Hence, $S \subseteq \{t\}$. Thus, either $S = \varnothing$ or $S = \{t\}$ (since the only subsets of $\{t\}$ are $\varnothing$ and $\{t\}$). Since

[55] *Proof.* Let S be as in Theorem 2.35, and assume that $|S| = 0$. We must show that the claim of Theorem 2.35 holds.

Indeed, $|S| = 0$, so that S is the empty set. This contradicts the assumption that S be nonempty. From this contradiction, we conclude that everything holds (by the "ex falso quodlibet" principle). Thus, in particular, the claim of Theorem 2.35 holds. This completes our proof.

$S = \varnothing$ is impossible (because S is nonempty), we thus have $S = \{t\}$. But the set $\{t\}$ has a maximum. In view of $S = \{t\}$, this rewrites as follows: The set S has a maximum. Thus, our goal (to prove that S has a maximum) is achieved in Case 1.

Let us now consider Case 2. In this case, we have $S \setminus \{t\} \neq \varnothing$. Hence, the set $S \setminus \{t\}$ is nonempty. From $t \in S$, we obtain $|S \setminus \{t\}| = |S| - 1 = m$ (since $|S| = m + 1$). Furthermore, the set $S \setminus \{t\}$ is finite (since S is finite). Hence, (73) (applied to $S \setminus \{t\}$ instead of S) shows that $S \setminus \{t\}$ has a maximum. Also, recall that the set $\{t\}$ has a maximum. Hence, Proposition 2.34 (applied to $P = S \setminus \{t\}$ and $Q = \{t\}$) shows that the set $(S \setminus \{t\}) \cup \{t\}$ has a maximum. Since $(S \setminus \{t\}) \cup \{t\} = S$, this rewrites as follows: The set S has a maximum. Hence, our goal (to prove that S has a maximum) is achieved in Case 2.

We have now proven that S has a maximum in each of the two Cases 1 and 2. Therefore, S always has a maximum (since Cases 1 and 2 cover all possibilities).

Now, forget that we fixed S. We thus have shown that if S is a nonempty finite set of integers satisfying $|S| = m + 1$, then S has a maximum. In other words, Theorem 2.35 holds under the condition that $|S| = m + 1$. This completes the induction step. Thus, the induction proof of Theorem 2.35 is complete. □

2.5.5. Further results on maxima and minima

We can replace "integers" by "rational numbers" or "real numbers" in Theorem 2.35; all the proofs given above still apply then. Thus, we obtain the following:

Theorem 2.38. Let S be a nonempty finite set of integers (or rational numbers, or real numbers). Then, S has a maximum.

Hence, if S is a nonempty finite set of integers (or rational numbers, or real numbers), then $\max S$ is well-defined (because Theorem 2.38 shows that S has a maximum, and Proposition 2.31 shows that this maximum is unique).

Moreover, just as we have defined maxima (i.e., largest elements) of sets, we can define minima (i.e., smallest elements) of sets, and prove similar results about them:

Definition 2.39. Let S be a set of integers (or rational numbers, or real numbers). A *minimum* of S is defined to be an element $s \in S$ that satisfies

$$(s \leq t \text{ for each } t \in S).$$

In other words, a minimum of S is defined to be an element of S which is less or equal to each element of S.

(The plural of the word "minimum" is "minima".)

Example 2.40. The set $\{2, 4, 5\}$ has exactly one minimum: namely, 2.

The set $\mathbb{N} = \{0, 1, 2, \ldots\}$ has exactly one minimum: namely, 0.

The set $\{0, -1, -2, \ldots\}$ has no minimum: If k was a minimum of this set, then we would have $k \leq k - 1$, which is absurd.

The set $\varnothing$ has no minimum, since a minimum would have to be an element of $\varnothing$.

The analogue of Proposition 2.31 for minima instead of maxima looks exactly as one would expect it:

Proposition 2.41. Let S be a set of integers (or rational numbers, or real numbers). Then, S has **at most one** minimum.

Proof of Proposition 2.41. To obtain a proof of Proposition 2.41, it suffices to replace every "$\geq$" sign by a "$\leq$" sign (and every word "maximum" by "minimum") in the proof of Proposition 2.31 given above. □

Definition 2.42. Let S be a set of integers (or rational numbers, or real numbers). Proposition 2.41 shows that S has **at most one** minimum. Thus, if S has a minimum, then this minimum is the unique minimum of S; we shall thus call it *the minimum* of S or *the smallest element* of S. We shall denote this minimum by $\min S$.

The analogue of Theorem 2.38 is the following:

Theorem 2.43. Let S be a nonempty finite set of integers (or rational numbers, or real numbers). Then, S has a minimum.

Proof of Theorem 2.43. To obtain a proof of Theorem 2.43, it suffices to replace every "$\geq$" sign by a "$\leq$" sign (and every word "maximum" by "minimum") in the proof of Theorem 2.38 given above (and also in the proofs of all the auxiliary results that were used in said proof).[56] □

Alternatively, Theorem 2.43 can be obtained from Theorem 2.38 by applying the latter theorem to the set $\{-s \mid s \in S\}$. In fact, it is easy to see that a number x is the minimum of S if and only if $-x$ is the maximum of the set $\{-s \mid s \in S\}$. We leave the details of this simple argument to the reader.

We also should mention that Theorem 2.43 holds **without** requiring that S be finite, if we instead require that S consist of nonnegative integers:

Theorem 2.44. Let S be a nonempty set of nonnegative integers. Then, S has a minimum.

But S does not necessarily have a maximum in this situation; the nonnegativity requirement has "broken the symmetry" between maxima and minima.

[56] To be technically precise: not every "$\geq$" sign, of course. The "$\geq$" sign in "$m \geq 0$" should stay unchanged.

We note that the word "integers" is crucial in Theorem 2.44. If we replaced "integers" by "rational numbers", then the theorem would no longer hold (for example, the set of all positive rational numbers has no minimum, since positive rational numbers can get arbitrarily close to 0 yet cannot equal 0).

Proof of Theorem 2.44. The set S is nonempty. Thus, there exists some $p \in S$. Consider this p.

We have $p \in S \subseteq \mathbb{N}$ (since S is a set of nonnegative integers). Thus, $p \in \{0, 1, \ldots, p\}$. Combining this with $p \in S$, we obtain $p \in \{0, 1, \ldots, p\} \cap S$. Hence, the set $\{0, 1, \ldots, p\} \cap S$ contains the element p, and thus is nonempty. Moreover, this set $\{0, 1, \ldots, p\} \cap S$ is a subset of the finite set $\{0, 1, \ldots, p\}$, and thus is finite.

Now we know that $\{0, 1, \ldots, p\} \cap S$ is a nonempty finite set of integers. Hence, Theorem 2.43 (applied to $\{0, 1, \ldots, p\} \cap S$ instead of S) shows that the set $\{0, 1, \ldots, p\} \cap S$ has a minimum. Denote this minimum by m.

Hence, m is a minimum of the set $\{0, 1, \ldots, p\} \cap S$. In other words, m is an element $s \in \{0, 1, \ldots, p\} \cap S$ that satisfies

$$(s \leq t \text{ for each } t \in \{0, 1, \ldots, p\} \cap S)$$

(by the definition of a minimum). In other words, m is an element of $\{0, 1, \ldots, p\} \cap S$ and satisfies

$$(m \leq t \text{ for each } t \in \{0, 1, \ldots, p\} \cap S) . \tag{74}$$

Hence, $m \in \{0, 1, \ldots, p\} \cap S \subseteq \{0, 1, \ldots, p\}$, so that $m \leq p$.

Furthermore, $m \in \{0, 1, \ldots, p\} \cap S \subseteq S$. Moreover, we have

$$(m \leq t \text{ for each } t \in S) . \tag{75}$$

[*Proof of (75):* Let $t \in S$. We must prove that $m \leq t$.

If $t \in \{0, 1, \ldots, p\} \cap S$, then this follows from (74). Hence, for the rest of this proof, we can WLOG assume that we don't have $t \in \{0, 1, \ldots, p\} \cap S$. Assume this. Thus, $t \notin \{0, 1, \ldots, p\} \cap S$. Combining $t \in S$ with $t \notin \{0, 1, \ldots, p\} \cap S$, we obtain

$$t \in S \setminus (\{0, 1, \ldots, p\} \cap S) = S \setminus \{0, 1, \ldots, p\} .$$

Hence, $t \notin \{0, 1, \ldots, p\}$, so that $t > p$ (since $t \in \mathbb{N}$). Therefore, $t \geq p \geq m$ (since $m \leq p$), so that $m \leq t$. This completes the proof of (75).]

Now, we know that m is an element of S (since $m \in S$) and satisfies $(m \leq t \text{ for each } t \in S)$ (by (75)). In other words, m is an $s \in S$ that satisfies $(s \leq t \text{ for each } t \in S)$. In other words, m is a minimum of S (by the definition of a minimum). Thus, S has a minimum (namely, m). This proves Theorem 2.44. □

2.6. Increasing lists of finite sets

We shall next study (again using induction) another basic feature of finite sets.

We recall that "list" is just a synonym for "tuple"; i.e., a list is a k-tuple for some $k \in \mathbb{N}$. Note that tuples and lists are always understood to be finite and ordered.

Definition 2.45. Let S be a set of integers. An *increasing list* of S shall mean a list $(s_1, s_2, \ldots, s_k)$ of elements of S such that $S = \{s_1, s_2, \ldots, s_k\}$ and $s_1 < s_2 < \cdots < s_k$.

In other words, if S is a set of integers, then an increasing list of S means a list such that

- the set S consists of all elements of this list, and
- the elements of this list are strictly increasing.

For example, $(2,4,6)$ is an increasing list of the set $\{2,4,6\}$, but neither $(2,6)$ nor $(2,4,4,6)$ nor $(4,2,6)$ nor $(2,4,5,6)$ is an increasing list of this set. For another example, $(1,4,9,16)$ is an increasing list of the set $\left\{i^2 \mid i \in \{1,2,3,4\}\right\} = \{1,4,9,16\}$. For yet another example, the empty list $()$ is an increasing list of the empty set $\varnothing$.

Now, it is intuitively obvious that any finite set S of integers has a unique increasing list – we just need to list all the elements of S in increasing order, with no repetitions. But from the viewpoint of rigorous mathematics, this needs to be proven. Let us state this as a theorem:

Theorem 2.46. Let S be a finite set of integers. Then, S has exactly one increasing list.

Before we prove this theorem, let us show some auxiliary facts:

Proposition 2.47. Let S be a set of integers. Let $(s_1, s_2, \ldots, s_k)$ be an increasing list of S. Then:
(a) The set S is finite.
(b) We have $|S| = k$.
(c) The elements $s_1, s_2, \ldots, s_k$ are distinct.

Proof of Proposition 2.47. We know that $(s_1, s_2, \ldots, s_k)$ is an increasing list of S. In other words, $(s_1, s_2, \ldots, s_k)$ is a list of elements of S such that $S = \{s_1, s_2, \ldots, s_k\}$ and $s_1 < s_2 < \cdots < s_k$ (by the definition of an "increasing list").

From $S = \{s_1, s_2, \ldots, s_k\}$, we conclude that the set S has at most k elements. Thus, the set S is finite. This proves Proposition 2.47 **(a)**.

We have $s_1 < s_2 < \cdots < s_k$. Hence, if u and v are two elements of $\{1, 2, \ldots, k\}$ such that $u < v$, then $s_u < s_v$ (by Corollary 2.20, applied to $a_i = s_i$) and therefore $s_u \neq s_v$. In other words, the elements $s_1, s_2, \ldots, s_k$ are distinct. This proves Proposition 2.47 **(c)**.

The k elements $s_1, s_2, \ldots, s_k$ are distinct; thus, the set $\{s_1, s_2, \ldots, s_k\}$ has size k. In other words, the set S has size k (since $S = \{s_1, s_2, \ldots, s_k\}$). In other words, $|S| = k$. This proves Proposition 2.47 **(b)**. □

Proposition 2.48. The set $\varnothing$ has exactly one increasing list: namely, the empty list $()$.

Proof of Proposition 2.48. The empty list $()$ satisfies $\varnothing = \{\}$. Thus, the empty list $()$ is a list $(s_1, s_2, \ldots, s_k)$ of elements of $\varnothing$ such that $\varnothing = \{s_1, s_2, \ldots, s_k\}$ and $s_1 < s_2 < \cdots < s_k$ (indeed, the chain of inequalities $s_1 < s_2 < \cdots < s_k$ is vacuously true for the empty list $()$, because it contains no inequality signs). In other words, the empty list $()$ is an increasing list of $\varnothing$ (by the definition of an increasing list). It remains to show that it is the only increasing list of $\varnothing$.

Let $(s_1, s_2, \ldots, s_k)$ be any increasing list of $\varnothing$. Then, Proposition 2.47 **(b)** (applied to $S = \varnothing$) yields $|\varnothing| = k$. Hence, $k = |\varnothing| = 0$, so that $(s_1, s_2, \ldots, s_k) = (s_1, s_2, \ldots, s_0) = ()$.

Now, forget that we fixed $(s_1, s_2, \ldots, s_k)$. We thus have shown that if $(s_1, s_2, \ldots, s_k)$ is any increasing list of $\varnothing$, then $(s_1, s_2, \ldots, s_k) = ()$. In other words, any increasing list of $\varnothing$ is $()$. Therefore, the set $\varnothing$ has exactly one increasing list: namely, the empty list $()$ (since we already know that $()$ is an increasing list of $\varnothing$). This proves Proposition 2.48. □

Proposition 2.49. Let S be a nonempty finite set of integers. Let $m = \max S$. Let $(s_1, s_2, \ldots, s_k)$ be any increasing list of S. Then:

(a) We have $k \geq 1$ and $s_k = m$.

(b) The list $(s_1, s_2, \ldots, s_{k-1})$ is an increasing list of $S \setminus \{m\}$.

Proof of Proposition 2.49. We know that $(s_1, s_2, \ldots, s_k)$ is an increasing list of S. In other words, $(s_1, s_2, \ldots, s_k)$ is a list of elements of S such that $S = \{s_1, s_2, \ldots, s_k\}$ and $s_1 < s_2 < \cdots < s_k$ (by the definition of an "increasing list").

Proposition 2.47 **(b)** yields $|S| = k$. Hence, $k = |S| > 0$ (since S is nonempty). Thus, $k \geq 1$ (since k is an integer). Therefore, s_k is well-defined. Clearly, $k \in \{1, 2, \ldots, k\}$ (since $k \geq 1$), so that $s_k \in \{s_1, s_2, \ldots, s_k\} = S$.

We have $s_1 < s_2 < \cdots < s_k$ and thus $s_1 \leq s_2 \leq \cdots \leq s_k$. Hence, Proposition 2.18 (applied to $a_i = s_i$) shows that if u and v are two elements of $\{1, 2, \ldots, k\}$ such that $u \leq v$, then

$$s_u \leq s_v. \tag{76}$$

Thus, we have $(s_k \geq t \text{ for each } t \in S)$ [57]. Hence, s_k is an element $s \in S$ that satisfies $(s \geq t \text{ for each } t \in S)$ (since $s_k \in S$). In other words, s_k is a maximum of S (by the definition of a maximum). Since we know that S has at most one maximum (by Proposition 2.31), we thus conclude that s_k is **the** maximum of S. In other words, $s_k = \max S$. Hence, $s_k = \max S = m$. This completes the proof of Proposition 2.49 **(a)**.

(b) From $s_1 < s_2 < \cdots < s_k$, we obtain $s_1 < s_2 < \cdots < s_{k-1}$. Furthermore, the elements $s_1, s_2, \ldots, s_k$ are distinct (according to Proposition 2.47 **(c)**). In other

[57] *Proof.* Let $t \in S$. Thus, $t \in S = \{s_1, s_2, \ldots, s_k\}$. Hence, $t = s_u$ for some $u \in \{1, 2, \ldots, k\}$. Consider this u. Now, u and k are elements of $\{1, 2, \ldots, k\}$ such that $u \leq k$ (since $u \in \{1, 2, \ldots, k\}$). Hence, (76) (applied to $v = k$) yields $s_u \leq s_k$. Hence, $s_k \geq s_u = t$ (since $t = s_u$), qed.

words, for any two distinct elements u and v of $\{1, 2, \ldots, k\}$, we have

$$s_u \neq s_v. \tag{77}$$

Hence, $s_k \notin \{s_1, s_2, \ldots, s_{k-1}\}$ [58]. Now,

$$\underbrace{S}_{\substack{=\{s_1,s_2,\ldots,s_k\} \\ =\{s_1,s_2,\ldots,s_{k-1}\}\cup\{s_k\}}} \setminus \left\{ \underbrace{m}_{=s_k} \right\} = (\{s_1, s_2, \ldots, s_{k-1}\} \cup \{s_k\}) \setminus \{s_k\}$$
$$= \{s_1, s_2, \ldots, s_{k-1}\} \setminus \{s_k\} = \{s_1, s_2, \ldots, s_{k-1}\}$$

(since $s_k \notin \{s_1, s_2, \ldots, s_{k-1}\}$). Hence, the elements $s_1, s_2, \ldots, s_{k-1}$ belong to the set $S \setminus \{m\}$ (since they clearly belong to the set $\{s_1, s_2, \ldots, s_{k-1}\} = S \setminus \{m\}$). In other words, $(s_1, s_2, \ldots, s_{k-1})$ is a list of elements of $S \setminus \{m\}$.

Now, we know that $(s_1, s_2, \ldots, s_{k-1})$ is a list of elements of $S \setminus \{m\}$ such that $S \setminus \{m\} = \{s_1, s_2, \ldots, s_{k-1}\}$ and $s_1 < s_2 < \cdots < s_{k-1}$. In other words, $(s_1, s_2, \ldots, s_{k-1})$ is an increasing list of $S \setminus \{m\}$. This proves Proposition 2.49 **(b)**. □

We are now ready to prove Theorem 2.46:

Proof of Theorem 2.46. We shall prove Theorem 2.46 by induction on $|S|$:

Induction base: Theorem 2.46 holds under the condition that $|S| = 0$ [59]. This completes the induction base.

Induction step: Let $g \in \mathbb{N}$. Assume that Theorem 2.46 holds under the condition that $|S| = g$. We shall now show that Theorem 2.46 holds under the condition that $|S| = g + 1$.

We have assumed that Theorem 2.46 holds under the condition that $|S| = g$. In other words,

$$\left(\begin{array}{c} \text{if } S \text{ is a finite set of integers satisfying } |S| = g, \\ \text{then } S \text{ has exactly one increasing list} \end{array} \right). \tag{78}$$

Now, let S be a finite set of integers satisfying $|S| = g + 1$. We want to prove that S has exactly one increasing list.

The set S is nonempty (since $|S| = g + 1 > g \geq 0$). Thus, S has a maximum (by Theorem 2.35). Hence, $\max S$ is well-defined. Set $m = \max S$. Thus, $m = \max S \in S$

[58] *Proof.* Assume the contrary. Thus, $s_k \in \{s_1, s_2, \ldots, s_{k-1}\}$. In other words, $s_k = s_u$ for some $u \in \{1, 2, \ldots, k-1\}$. Consider this u. We have $u \in \{1, 2, \ldots, k-1\} \subseteq \{1, 2, \ldots, k\}$.

Now, $u \in \{1, 2, \ldots, k-1\}$, so that $u \leq k - 1 < k$ and thus $u \neq k$. Hence, the elements u and k of $\{1, 2, \ldots, k\}$ are distinct. Thus, (77) (applied to $v = k$) yields $s_u \neq s_k = s_u$. This is absurd. This contradiction shows that our assumption was wrong, qed.

[59] *Proof.* Let S be as in Theorem 2.46, and assume that $|S| = 0$. We must show that the claim of Theorem 2.46 holds.

Indeed, $|S| = 0$, so that S is the empty set. Thus, $S = \varnothing$. But Proposition 2.48 shows that the set $\varnothing$ has exactly one increasing list. Since $S = \varnothing$, this rewrites as follows: The set S has exactly one increasing list. Thus, the claim of Theorem 2.46 holds. This completes our proof.

(by (67)). Therefore, $|S \setminus \{m\}| = |S| - 1 = g$ (since $|S| = g + 1$). Hence, (78) (applied to $S \setminus \{m\}$ instead of S) shows that $S \setminus \{m\}$ has exactly one increasing list. Let $(t_1, t_2, \ldots, t_j)$ be this list. We extend this list to a $(j+1)$-tuple $(t_1, t_2, \ldots, t_{j+1})$ by setting $t_{j+1} = m$.

We have defined $(t_1, t_2, \ldots, t_j)$ as an increasing list of the set $S \setminus \{m\}$. In other words, $(t_1, t_2, \ldots, t_j)$ is a list of elements of $S \setminus \{m\}$ such that $S \setminus \{m\} = \{t_1, t_2, \ldots, t_j\}$ and $t_1 < t_2 < \cdots < t_j$ (by the definition of an "increasing list").

We claim that

$$t_1 < t_2 < \cdots < t_{j+1}. \tag{79}$$

[*Proof of (79):* If $j + 1 \leq 1$, then the chain of inequalities (79) is vacuously true (since it contains no inequality signs). Thus, for the rest of this proof of (79), we WLOG assume that we don't have $j + 1 \leq 1$. Hence, $j + 1 > 1$, so that $j > 0$ and thus $j \geq 1$ (since j is an integer). Hence, t_j is well-defined. We have $j \in \{1, 2, \ldots, j\}$ (since $j \geq 1$) and thus $t_j \in \{t_1, t_2, \ldots, t_j\} = S \setminus \{m\} \subseteq S$. Hence, (68) (applied to $t = t_j$) yields $\max S \geq t_j$. Hence, $t_j \leq \max S = m$. Moreover, $t_j \notin \{m\}$ (since $t_j \in S \setminus \{m\}$); in other words, $t_j \neq m$. Combining this with $t_j \leq m$, we obtain $t_j < m = t_{j+1}$. Combining the chain of inequalities $t_1 < t_2 < \cdots < t_j$ with the single inequality $t_j < t_{j+1}$, we obtain the longer chain of inequalities $t_1 < t_2 < \cdots < t_j < t_{j+1}$. In other words, $t_1 < t_2 < \cdots < t_{j+1}$. This proves (79).]

Next, we shall prove that

$$S = \{t_1, t_2, \ldots, t_{j+1}\}. \tag{80}$$

[*Proof of (80):* We have $(S \setminus \{m\}) \cup \{m\} = S \cup \{m\} = S$ (since $m \in S$). Thus,

$$\begin{aligned} S &= \left(\underbrace{S \setminus \{m\}}_{= \{t_1, t_2, \ldots, t_j\}} \right) \cup \left\{ \underbrace{m}_{= t_{j+1}} \right\} = \{t_1, t_2, \ldots, t_j\} \cup \{t_{j+1}\} = \{t_1, t_2, \ldots, t_j, t_{j+1}\} \\ &= \{t_1, t_2, \ldots, t_{j+1}\}. \end{aligned}$$

This proves (80).]

Clearly, $t_1, t_2, \ldots, t_{j+1}$ are elements of the set $\{t_1, t_2, \ldots, t_{j+1}\}$. In other words, $t_1, t_2, \ldots, t_{j+1}$ are elements of the set S (since $S = \{t_1, t_2, \ldots, t_{j+1}\}$).

Hence, $(t_1, t_2, \ldots, t_{j+1})$ is a list of elements of S. Thus, $(t_1, t_2, \ldots, t_{j+1})$ is a list of elements of S such that $S = \{t_1, t_2, \ldots, t_{j+1}\}$ (by (80)) and $t_1 < t_2 < \cdots < t_{j+1}$ (by (79)). In other words, $(t_1, t_2, \ldots, t_{j+1})$ is an increasing list of S (by the definition of an "increasing list"). Hence, the set S has **at least** one increasing list (namely, $(t_1, t_2, \ldots, t_{j+1})$).

We shall next show that $(t_1, t_2, \ldots, t_{j+1})$ is the only increasing list of S. Indeed, let $(s_1, s_2, \ldots, s_k)$ be any increasing list of S. Then, Proposition 2.49 **(a)** shows that $k \geq 1$ and $s_k = m$. Also, Proposition 2.49 **(b)** shows that the list $(s_1, s_2, \ldots, s_{k-1})$ is an increasing list of $S \setminus \{m\}$.

But recall that $S \setminus \{m\}$ has exactly one increasing list. Thus, in particular, $S \setminus \{m\}$ has **at most** one increasing list. In other words, any two increasing lists of $S \setminus \{m\}$ are equal. Hence, the lists $(s_1, s_2, \ldots, s_{k-1})$ and $(t_1, t_2, \ldots, t_j)$ must be equal (since both of these lists are increasing lists of $S \setminus \{m\}$). In other words, $(s_1, s_2, \ldots, s_{k-1}) = (t_1, t_2, \ldots, t_j)$. In other words, $k-1 = j$ and

$$(s_i = t_i \text{ for each } i \in \{1, 2, \ldots, k-1\}) . \tag{81}$$

From $k - 1 = j$, we obtain $k = j + 1$. Hence, $t_k = t_{j+1} = m$. Next, we claim that

$$s_i = t_i \text{ for each } i \in \{1, 2, \ldots, k\} . \tag{82}$$

[*Proof of (82):* Let $i \in \{1, 2, \ldots, k\}$. We must prove that $s_i = t_i$. If $i \in \{1, 2, \ldots, k-1\}$, then this follows from (81). Hence, for the rest of this proof, we WLOG assume that we don't have $i \in \{1, 2, \ldots, k-1\}$. Hence, $i \notin \{1, 2, \ldots, k-1\}$. Combining $i \in \{1, 2, \ldots, k\}$ with $i \notin \{1, 2, \ldots, k-1\}$, we obtain

$$i \in \{1, 2, \ldots, k\} \setminus \{1, 2, \ldots, k-1\} = \{k\} .$$

In other words, $i = k$. Hence, $s_i = s_k = m = t_k$ (since $t_k = m$). In view of $k = i$, this rewrites as $s_i = t_i$. This proves (82).]

From (82), we obtain $(s_1, s_2, \ldots, s_k) = (t_1, t_2, \ldots, t_k) = (t_1, t_2, \ldots, t_{j+1})$ (since $k = j + 1$).

Now, forget that we fixed $(s_1, s_2, \ldots, s_k)$. We thus have proven that if $(s_1, s_2, \ldots, s_k)$ is any increasing list of S, then $(s_1, s_2, \ldots, s_k) = (t_1, t_2, \ldots, t_{j+1})$. In other words, any increasing list of S equals $(t_1, t_2, \ldots, t_{j+1})$. Thus, the set S has **at most** one increasing list. Since we also know that the set S has **at least** one increasing list, we thus conclude that S has exactly one increasing list.

Now, forget that we fixed S. We thus have shown that

$$\left(\begin{array}{c} \text{if } S \text{ is a finite set of integers satisfying } |S| = g+1, \\ \text{then } S \text{ has exactly one increasing list} \end{array} \right) .$$

In other words, Theorem 2.46 holds under the condition that $|S| = g + 1$. This completes the induction step. Hence, Theorem 2.46 is proven by induction. □

Definition 2.50. Let S be a finite set of integers. Theorem 2.46 shows that S has exactly one increasing list. This increasing list is called *the increasing list* of S. It is also called *the list of all elements of S in increasing order (with no repetitions)*. (The latter name, of course, is descriptive.)

The increasing list of S has length $|S|$. (Indeed, if we denote this increasing list by $(s_1, s_2, \ldots, s_k)$, then its length is $k = |S|$, because Proposition 2.47 **(b)** shows that $|S| = k$.)

For each $j \in \{1, 2, \ldots, |S|\}$, we define the *$j$-th smallest element of S* to be the j-th entry of the increasing list of S. In other words, if $(s_1, s_2, \ldots, s_k)$ is the increasing list of S, then the j-th smallest element of S is s_j. Some say "j-th lowest element of S" instead of "j-th smallest element of S".

Remark 2.51. (a) Clearly, we can replace the word "integer" by "rational number" or by "real number" in Proposition 2.18, Corollary 2.19, Corollary 2.20, Definition 2.45, Theorem 2.46, Proposition 2.47, Proposition 2.48, Proposition 2.49 and Definition 2.50, because we have not used any properties specific to integers.

(b) If we replace all the "$<$" signs in Definition 2.45 by "$>$" signs, then we obtain the notion of a *decreasing list* of S. There are straightforward analogues of Theorem 2.46, Proposition 2.47, Proposition 2.48 and Proposition 2.49 for decreasing lists (where, of course, the analogue of Proposition 2.49 uses $\min S$ instead of $\max S$). Thus, we can state an analogue of Definition 2.50 as well. In this analogue, the word "increasing" is replaced by "decreasing" everywhere, the word "smallest" is replaced by "largest", and the word "lowest" is replaced by "highest".

(c) That said, the decreasing list and the increasing list are closely related: If S is a finite set of integers (or rational numbers, or real numbers), and if $(s_1, s_2, \ldots, s_k)$ is the increasing list of S, then $(s_k, s_{k-1}, \ldots, s_1)$ is the decreasing list of S. (The proof is very simple.)

(d) Let S be a nonempty finite set of integers (or rational numbers, or real numbers), and let $(s_1, s_2, \ldots, s_k)$ be the increasing list of S. Proposition 2.49 **(a)** (applied to $m = \max S$) shows that $k \geq 1$ and $s_k = \max S$. A similar argument can be used to show that $s_1 = \min S$. Thus, the increasing list of S begins with the smallest element of S and ends with the largest element of S (as one would expect).

2.7. Induction with shifted base

2.7.1. Induction starting at g

All the induction proofs we have done so far were applications of Theorem 2.1 (even though we have often written them up in ways that hide the exact statements $\mathcal{A}(n)$ to which Theorem 2.1 is being applied). We are soon going to see several other "induction principles" which can also be used to make proofs. Unlike Theorem 2.1, these other principles need not be taken on trust; instead, they can themselves be proven using Theorem 2.1. Thus, they merely offer convenience, not new logical opportunities.

Our first such "alternative induction principle" is Theorem 2.53 below. First, we introduce a simple notation:

Definition 2.52. Let $g \in \mathbb{Z}$. Then, $\mathbb{Z}_{\geq g}$ denotes the set $\{g, g+1, g+2, \ldots\}$; this is the set of all integers that are $\geq g$.

For example, $\mathbb{Z}_{\geq 0} = \{0, 1, 2, \ldots\} = \mathbb{N}$ is the set of all nonnegative integers, whereas $\mathbb{Z}_{\geq 1} = \{1, 2, 3, \ldots\}$ is the set of all positive integers.

Now, we state our first "alternative induction principle":

Theorem 2.53. Let $g \in \mathbb{Z}$. For each $n \in \mathbb{Z}_{\geq g}$, let $\mathcal{A}(n)$ be a logical statement.

Assume the following:

Assumption 1: The statement $\mathcal{A}(g)$ holds.

Assumption 2: If $m \in \mathbb{Z}_{\geq g}$ is such that $\mathcal{A}(m)$ holds, then $\mathcal{A}(m+1)$ also holds.

Then, $\mathcal{A}(n)$ holds for each $n \in \mathbb{Z}_{\geq g}$.

Again, Theorem 2.53 is intuitively clear: For example, if you have $g = 4$, and you want to prove (under the assumptions of Theorem 2.53) that $\mathcal{A}(8)$ holds, you can argue as follows:

- By Assumption 1, the statement $\mathcal{A}(4)$ holds.
- Thus, by Assumption 2 (applied to $m = 4$), the statement $\mathcal{A}(5)$ holds.
- Thus, by Assumption 2 (applied to $m = 5$), the statement $\mathcal{A}(6)$ holds.
- Thus, by Assumption 2 (applied to $m = 6$), the statement $\mathcal{A}(7)$ holds.
- Thus, by Assumption 2 (applied to $m = 7$), the statement $\mathcal{A}(8)$ holds.

A similar (but longer) argument shows that the statement $\mathcal{A}(9)$ holds; likewise, $\mathcal{A}(n)$ can be shown to hold for each $n \in \mathbb{Z}_{\geq g}$ by means of an argument that takes $n - g + 1$ steps.

Theorem 2.53 generalizes Theorem 2.1. Indeed, Theorem 2.1 is the particular case of Theorem 2.53 for $g = 0$ (since $\mathbb{Z}_{\geq 0} = \mathbb{N}$). However, Theorem 2.53 can also be derived from Theorem 2.1. In order to do this, we essentially need to "shift" the index n in Theorem 2.53 down by g – that is, we need to rename our sequence $(\mathcal{A}(g), \mathcal{A}(g+1), \mathcal{A}(g+2), \ldots)$ of statements as $(\mathcal{B}(0), \mathcal{B}(1), \mathcal{B}(2), \ldots)$, and apply Theorem 2.1 to $\mathcal{B}(n)$ instead of $\mathcal{A}(n)$. In order to make this renaming procedure rigorous, let us first restate Theorem 2.1 as follows:

Corollary 2.54. For each $n \in \mathbb{N}$, let $\mathcal{B}(n)$ be a logical statement.

Assume the following:

Assumption A: The statement $\mathcal{B}(0)$ holds.

Assumption B: If $p \in \mathbb{N}$ is such that $\mathcal{B}(p)$ holds, then $\mathcal{B}(p+1)$ also holds.

Then, $\mathcal{B}(n)$ holds for each $n \in \mathbb{N}$.

Proof of Corollary 2.54. Corollary 2.54 is exactly Theorem 2.1, except that some names have been changed:

- The statements $\mathcal{A}(n)$ have been renamed as $\mathcal{B}(n)$.
- Assumption 1 and Assumption 2 have been renamed as Assumption A and Assumption B.
- The variable m in Assumption B has been renamed as p.

Thus, Corollary 2.54 holds (since Theorem 2.1 holds). □

Let us now derive Theorem 2.53 from Theorem 2.1:

Proof of Theorem 2.53. For any $n \in \mathbb{N}$, we have $n + g \in \mathbb{Z}_{\geq g}$ [60]. Hence, for each $n \in \mathbb{N}$, we can define a logical statement $\mathcal{B}(n)$ by

$$\mathcal{B}(n) = \mathcal{A}(n+g).$$

Consider this $\mathcal{B}(n)$.

Now, let us consider the Assumptions A and B from Corollary 2.54. We claim that both of these assumptions are satisfied.

Indeed, the statement $\mathcal{A}(g)$ holds (by Assumption 1). But the definition of the statement $\mathcal{B}(0)$ shows that $\mathcal{B}(0) = \mathcal{A}(0+g) = \mathcal{A}(g)$. Hence, the statement $\mathcal{B}(0)$ holds (since the statement $\mathcal{A}(g)$ holds). In other words, Assumption A is satisfied.

Now, we shall show that Assumption B is satisfied. Indeed, let $p \in \mathbb{N}$ be such that $\mathcal{B}(p)$ holds. The definition of the statement $\mathcal{B}(p)$ shows that $\mathcal{B}(p) = \mathcal{A}(p+g)$. Hence, the statement $\mathcal{A}(p+g)$ holds (since $\mathcal{B}(p)$ holds).

Also, $p \in \mathbb{N}$, so that $p \geq 0$ and thus $p + g \geq g$. In other words, $p + g \in \mathbb{Z}_{\geq g}$ (since $\mathbb{Z}_{\geq g}$ is the set of all integers that are $\geq g$).

Recall that Assumption 2 holds. In other words, if $m \in \mathbb{Z}_{\geq g}$ is such that $\mathcal{A}(m)$ holds, then $\mathcal{A}(m+1)$ also holds. Applying this to $m = p + g$, we conclude that $\mathcal{A}((p+g)+1)$ holds (since $\mathcal{A}(p+g)$ holds).

But the definition of $\mathcal{B}(p+1)$ yields $\mathcal{B}(p+1) = \mathcal{A}\left(\underbrace{p+1+g}_{=(p+g)+1}\right) = \mathcal{A}((p+g)+1)$. Hence, the statement $\mathcal{B}(p+1)$ holds (since the statement $\mathcal{A}((p+g)+1)$ holds).

Now, forget that we fixed p. We thus have shown that if $p \in \mathbb{N}$ is such that $\mathcal{B}(p)$ holds, then $\mathcal{B}(p+1)$ also holds. In other words, Assumption B is satisfied.

We now know that both Assumption A and Assumption B are satisfied. Hence, Corollary 2.54 shows that

$$\mathcal{B}(n) \text{ holds for each } n \in \mathbb{N}. \tag{83}$$

Now, let $n \in \mathbb{Z}_{\geq g}$. Thus, n is an integer such that $n \geq g$ (by the definition of $\mathbb{Z}_{\geq g}$). Hence, $n - g \geq 0$, so that $n - g \in \mathbb{N}$. Thus, (83) (applied to $n - g$

[60] *Proof.* Let $n \in \mathbb{N}$. Thus, $n \geq 0$, so that $\underbrace{n}_{\geq 0} + g \geq 0 + g = g$. Hence, $n + g$ is an integer $\geq g$. In other words, $n + g \in \mathbb{Z}_{\geq g}$ (since $\mathbb{Z}_{\geq g}$ is the set of all integers that are $\geq g$). Qed.

instead of n) yields that $\mathcal{B}(n-g)$ holds. But the definition of $\mathcal{B}(n-g)$ yields $\mathcal{B}(n-g) = \mathcal{A}\left(\underbrace{(n-g)+g}_{=n}\right) = \mathcal{A}(n)$. Hence, the statement $\mathcal{A}(n)$ holds (since $\mathcal{B}(n-g)$ holds).

Now, forget that we fixed n. We thus have shown that $\mathcal{A}(n)$ holds for each $n \in \mathbb{Z}_{\geq g}$. This proves Theorem 2.53. □

Theorem 2.53 is called the *principle of induction starting at* g, and proofs that use it are usually called *proofs by induction* or *induction proofs*. As with the standard induction principle (Theorem 2.1), we don't usually explicitly cite Theorem 2.53, but instead say certain words that signal that it is being applied and that (ideally) also indicate what integer g and what statements $\mathcal{A}(n)$ it is being applied to[61]. However, for our very first example of the use of Theorem 2.53, we are going to reference it explicitly:

Proposition 2.55. Let a and b be integers. Then, every positive integer n satisfies

$$(a+b)^n \equiv a^n + na^{n-1}b \mod b^2. \tag{84}$$

Note that we have chosen not to allow $n=0$ in Proposition 2.55, because it is not clear what "a^{n-1}" would mean when $n=0$ and $a=0$. (Recall that $0^{0-1} = 0^{-1}$ is not defined!) In truth, it is easy to convince oneself that this is not a serious hindrance, since the expression "na^{n-1}" has a meaningful interpretation even when its subexpression "a^{n-1}" does not (one just has to interpret it as 0 when $n=0$, without regard to whether "a^{n-1}" is well-defined). Nevertheless, we prefer to rule out the case of $n=0$ by requiring n to be positive, in order to avoid having to discuss such questions of interpretation. (Of course, this also gives us an excuse to apply Theorem 2.53 instead of the old Theorem 2.1.)

Proof of Proposition 2.55. For each $n \in \mathbb{Z}_{\geq 1}$, we let $\mathcal{A}(n)$ be the statement

$$\left((a+b)^n \equiv a^n + na^{n-1}b \mod b^2\right).$$

Our next goal is to prove the statement $\mathcal{A}(n)$ for each $n \in \mathbb{Z}_{\geq 1}$.

We first notice that the statement $\mathcal{A}(1)$ holds[62].

Now, we claim that

$$\text{if } m \in \mathbb{Z}_{\geq 1} \text{ is such that } \mathcal{A}(m) \text{ holds, then } \mathcal{A}(m+1) \text{ also holds.} \tag{85}$$

[61] We will explain this in Convention 2.56 below.

[62] *Proof.* We have $(a+b)^1 = a+b$. Comparing this with $\underbrace{a^1}_{=a} + 1 \underbrace{a^{1-1}}_{=a^0=1} b = a+b$, we obtain $(a+b)^1 = a^1 + 1a^{1-1}b$. Hence, $(a+b)^1 \equiv a^1 + 1a^{1-1}b \mod b^2$. But this is precisely the statement $\mathcal{A}(1)$ (since $\mathcal{A}(1)$ is defined to be the statement $\left((a+b)^1 \equiv a^1 + 1a^{1-1}b \mod b^2\right)$). Hence, the statement $\mathcal{A}(1)$ holds.

[*Proof of (85):* Let $m \in \mathbb{Z}_{\geq 1}$ be such that $\mathcal{A}(m)$ holds. We must show that $\mathcal{A}(m+1)$ also holds.

We have assumed that $\mathcal{A}(m)$ holds. In other words,

$$(a+b)^m \equiv a^m + ma^{m-1}b \bmod b^2$$

holds[63]. Now,

$$\begin{aligned}
(a+b)^{m+1} &= \underbrace{(a+b)^m}_{\equiv a^m + ma^{m-1}b \bmod b^2} (a+b) \\
&\equiv \left(a^m + ma^{m-1}b\right)(a+b) \\
&= \underbrace{a^m a}_{=a^{m+1}} + a^m b + m \underbrace{a^{m-1}ba}_{\substack{=a^{m-1}ab=a^m b \\ (\text{since } a^{m-1}a=a^m)}} + \underbrace{ma^{m-1}bb}_{\substack{=ma^{m-1}b^2\equiv 0 \bmod b^2 \\ (\text{since } b^2 \mid ma^{m-1}b^2)}} \\
&\equiv a^{m+1} + \underbrace{a^m b + ma^m b}_{=(m+1)a^m b} + 0 \\
&= a^{m+1} + (m+1) \underbrace{a^m}_{\substack{=a^{(m+1)-1} \\ (\text{since } m=(m+1)-1)}} b = a^{m+1} + (m+1)\,a^{(m+1)-1}b \bmod b^2.
\end{aligned}$$

So we have shown that $(a+b)^{m+1} \equiv a^{m+1} + (m+1)\,a^{(m+1)-1}b \bmod b^2$. But this is precisely the statement $\mathcal{A}(m+1)$ [64]. Thus, the statement $\mathcal{A}(m+1)$ holds.

Now, forget that we fixed m. We thus have shown that if $m \in \mathbb{Z}_{\geq 1}$ is such that $\mathcal{A}(m)$ holds, then $\mathcal{A}(m+1)$ also holds. This proves (85).]

Now, both assumptions of Theorem 2.53 (applied to $g = 1$) are satisfied (indeed, Assumption 1 is satisfied because the statement $\mathcal{A}(1)$ holds, whereas Assumption 2 is satisfied because of (85)). Thus, Theorem 2.53 (applied to $g = 1$) shows that $\mathcal{A}(n)$ holds for each $n \in \mathbb{Z}_{\geq 1}$. In other words, $(a+b)^n \equiv a^n + na^{n-1}b \bmod b^2$ holds for each $n \in \mathbb{Z}_{\geq 1}$ (since $\mathcal{A}(n)$ is the statement $\left((a+b)^n \equiv a^n + na^{n-1}b \bmod b^2\right)$). In other words, $(a+b)^n \equiv a^n + na^{n-1}b \bmod b^2$ holds for each positive integer n (because the positive integers are exactly the $n \in \mathbb{Z}_{\geq 1}$). This proves Proposition 2.55. □

2.7.2. Conventions for writing proofs by induction starting at g

Now, let us introduce some standard language that is commonly used in proofs by induction starting at g:

[63] because $\mathcal{A}(m)$ is defined to be the statement $\left((a+b)^m \equiv a^m + ma^{m-1}b \bmod b^2\right)$

[64] because $\mathcal{A}(m+1)$ is defined to be the statement $\left((a+b)^{m+1} \equiv a^{m+1} + (m+1)\,a^{(m+1)-1}b \bmod b^2\right)$

Convention 2.56. Let $g \in \mathbb{Z}$. For each $n \in \mathbb{Z}_{\geq g}$, let $\mathcal{A}(n)$ be a logical statement. Assume that you want to prove that $\mathcal{A}(n)$ holds for each $n \in \mathbb{Z}_{\geq g}$.

Theorem 2.53 offers the following strategy for proving this: First show that Assumption 1 of Theorem 2.53 is satisfied; then, show that Assumption 2 of Theorem 2.53 is satisfied; then, Theorem 2.53 automatically completes your proof.

A proof that follows this strategy is called a *proof by induction on n* (or *proof by induction over n*) *starting at g* or (less precisely) an *inductive proof*. Most of the time, the words "starting at g" are omitted, since they merely repeat what is clear from the context anyway: For example, if you make a claim about all integers $n \geq 3$, and you say that you are proving it by induction on n, then it is clear that you are using induction on n starting at 3. (And if this isn't clear from the claim, then the induction base will make it clear.)

The proof that Assumption 1 is satisfied is called the *induction base* (or *base case*) of the proof. The proof that Assumption 2 is satisfied is called the *induction step* of the proof.

In order to prove that Assumption 2 is satisfied, you will usually want to fix an $m \in \mathbb{Z}_{\geq g}$ such that $\mathcal{A}(m)$ holds, and then prove that $\mathcal{A}(m+1)$ holds. In other words, you will usually want to fix $m \in \mathbb{Z}_{\geq g}$, assume that $\mathcal{A}(m)$ holds, and then prove that $\mathcal{A}(m+1)$ holds. When doing so, it is common to refer to the assumption that $\mathcal{A}(m)$ holds as the *induction hypothesis* (or *induction assumption*).

Unsurprisingly, this language parallels the language introduced in Convention 2.3 for proofs by "standard" induction.

Again, we can shorten our inductive proofs by omitting some sentences that convey no information. In particular, we can leave out the explicit definition of the statement $\mathcal{A}(n)$ when this statement is precisely the claim that we are proving (without the "for each $n \in \mathbb{Z}_{\geq g}$" part). Thus, we can rewrite our above proof of Proposition 2.55 as follows:

Proof of Proposition 2.55 (second version). We must prove (84) for every positive integer n. In other words, we must prove (84) for every $n \in \mathbb{Z}_{\geq 1}$ (since the positive integers are precisely the $n \in \mathbb{Z}_{\geq 1}$). We shall prove this by induction on n starting at 1:

Induction base: We have $(a+b)^1 = a+b$. Comparing this with $\underbrace{a^1}_{=a} + 1 \underbrace{a^{1-1}}_{=a^0=1} b = a+b$, we obtain $(a+b)^1 = a^1 + 1a^{1-1}b$. Hence, $(a+b)^1 \equiv a^1 + 1a^{1-1}b \mod b^2$. In other words, (84) holds for $n=1$. This completes the induction base.

Induction step: Let $m \in \mathbb{Z}_{\geq 1}$. Assume that (84) holds for $n = m$. We must show that (84) also holds for $n = m+1$.

We have assumed that (84) holds for $n = m$. In other words,

$$(a+b)^m \equiv a^m + ma^{m-1}b \mod b^2$$

holds. Now,

$$\begin{aligned}
(a+b)^{m+1} &= \underbrace{(a+b)^m}_{\equiv a^m + ma^{m-1}b \operatorname{mod} b^2} (a+b) \\
&\equiv \left(a^m + ma^{m-1}b\right)(a+b) \\
&= \underbrace{a^m a}_{=a^{m+1}} + a^m b + m \underbrace{a^{m-1}ba}_{\substack{=a^{m-1}ab = a^m b \\ (\text{since } a^{m-1}a = a^m)}} + \underbrace{ma^{m-1}bb}_{\substack{=ma^{m-1}b^2 \equiv 0 \operatorname{mod} b^2 \\ (\text{since } b^2 \mid ma^{m-1}b^2)}} \\
&\equiv a^{m+1} + \underbrace{a^m b + ma^m b}_{=(m+1)a^m b} + 0 \\
&= a^{m+1} + (m+1) \underbrace{a^m}_{\substack{=a^{(m+1)-1} \\ (\text{since } m=(m+1)-1)}} b = a^{m+1} + (m+1)\, a^{(m+1)-1} b \operatorname{mod} b^2.
\end{aligned}$$

So we have shown that $(a+b)^{m+1} \equiv a^{m+1} + (m+1)\, a^{(m+1)-1} b \operatorname{mod} b^2$. In other words, (84) holds for $n = m+1$.

Now, forget that we fixed m. We thus have shown that if $m \in \mathbb{Z}_{\geq 1}$ is such that (84) holds for $n = m$, then (84) also holds for $n = m+1$. This completes the induction step. Hence, (84) is proven by induction. This proves Proposition 2.55. □

Proposition 2.55 can also be seen as a consequence of the binomial formula (Proposition 3.21 further below).

2.7.3. More properties of congruences

Let us use this occasion to show two corollaries of Proposition 2.55:

Corollary 2.57. Let a, b and n be three integers such that $a \equiv b \operatorname{mod} n$. Let $d \in \mathbb{N}$ be such that $d \mid n$. Then, $a^d \equiv b^d \operatorname{mod} nd$.

Proof of Corollary 2.57. We have $a \equiv b \operatorname{mod} n$. In other words, a is congruent to b modulo n. In other words, $n \mid a-b$ (by the definition of "congruent"). In other words, there exists an integer w such that $a - b = nw$. Consider this w. From $a - b = nw$, we obtain $a = b + nw$. Also, $d \mid n$, thus $dn \mid nn$ (by Proposition 2.6, applied to d, n and n instead of a, b and c). On the other hand, $nn \mid (nw)^2$ (since $(nw)^2 = nwnw = nnww$). Hence, Proposition 2.5 (applied to dn, nn and $(nw)^2$ instead of a, b and c) yields $dn \mid (nw)^2$ (since $dn \mid nn$ and $nn \mid (nw)^2$). In other words, $nd \mid (nw)^2$ (since $dn = nd$).

Next, we claim that

$$nd \mid a^d - b^d. \tag{86}$$

[*Proof of (86):* If $d = 0$, then (86) holds (because if $d = 0$, then $a^d - b^d = \underbrace{a^0}_{=1} - \underbrace{b^0}_{=1} = 1 - 1 = 0 = 0nd$, and thus $nd \mid a^d - b^d$). Hence, for the rest of this proof of (86), we WLOG assume that we don't have $d = 0$. Thus, $d \neq 0$. Hence, d is a positive integer (since $d \in \mathbb{N}$). Thus, Proposition 2.55 (applied to d, b and nw instead of n, a and b) yields

$$(b + nw)^d \equiv b^d + db^{d-1} nw \mod (nw)^2 .$$

In view of $a = b + nw$, this rewrites as

$$a^d \equiv b^d + db^{d-1} nw \mod (nw)^2 .$$

Hence, Proposition 2.11 **(c)** (applied to a^d, $b^d + db^{d-1}nw$, $(nw)^2$ and nd instead of a, b, n and m) yields

$$a^d \equiv b^d + db^{d-1} nw \mod nd$$

(since $nd \mid (nw)^2$). Hence,

$$a^d \equiv b^d + \underbrace{db^{d-1}nw}_{\substack{=ndb^{d-1}w \equiv 0 \mod nd \\ \text{(since } nd \mid ndb^{d-1}w\text{)}}} \equiv b^d + 0 = b^d \mod nd.$$

In other words, $nd \mid a^d - b^d$. This proves (86).]

From (86), we immediately obtain $a^d \equiv b^d \mod nd$ (by the definition of "congruent"). This proves Corollary 2.57. □

For the next corollary, we need a convention:

Convention 2.58. Let a, b and c be three integers. Then, the expression "a^{b^c}" shall always be interpreted as "$a^{(b^c)}$", never as "$\left(a^b\right)^c$".

Thus, for example, "3^{3^3}" means $3^{(3^3)} = 3^{27} = 7625\,597\,484\,987$, not $\left(3^3\right)^3 = 27^3 = 19\,683$. The reason for this convention is that $\left(a^b\right)^c$ can be simplified to a^{bc} and thus there is little use in having yet another notation for it. Of course, this convention applies not only to integers, but to any other numbers a, b, c.

We can now state the following fact, which is sometimes known as "lifting-the-exponent lemma":

Corollary 2.59. Let $n \in \mathbb{N}$. Let a and b be two integers such that $a \equiv b \mod n$. Let $k \in \mathbb{N}$. Then,

$$a^{n^k} \equiv b^{n^k} \mod n^{k+1}. \tag{87}$$

We shall give two **different** proofs of Corollary 2.59 by induction on k, to illustrate once again the point (previously made in Remark 2.27) that we have a choice of what precise statement we are proving by induction. In the first proof, the statement will be the congruence (87) for three **fixed** integers a, b and n, whereas in the second proof, it will be the statement

$$\left(a^{n^k} \equiv b^{n^k} \mod n^{k+1} \text{ for } \textbf{all} \text{ integers } a \text{ and } b \text{ and } \textbf{all } n \in \mathbb{N} \text{ satisfying } a \equiv b \mod n\right).$$

First proof of Corollary 2.59. Forget that we fixed k. We thus must prove (87) for each $k \in \mathbb{N}$.

We shall prove this by induction on k:

Induction base: We have $n^0 = 1$ and thus $a^{n^0} = a^1 = a$. Similarly, $b^{n^0} = b$. Thus, $a^{n^0} = a \equiv b = b^{n^0} \operatorname{mod} n$. In other words, $a^{n^0} \equiv b^{n^0} \operatorname{mod} n^{0+1}$ (since $n^{0+1} = n^1 = n$). In other words, (87) holds for $k = 0$. This completes the induction base.

Induction step: Let $m \in \mathbb{N}$. Assume that (87) holds for $k = m$. We must prove that (87) holds for $k = m + 1$.

We have $n^{m+1} = nn^m$. Hence, $n \mid n^{m+1}$.

We have assumed that (87) holds for $k = m$. In other words, we have

$$a^{n^m} \equiv b^{n^m} \operatorname{mod} n^{m+1}.$$

Hence, Corollary 2.57 (applied to a^{n^m}, b^{n^m}, n^{m+1} and n instead of a, b, n and d) yields

$$\left(a^{n^m}\right)^n \equiv \left(b^{n^m}\right)^n \operatorname{mod} n^{m+1}n.$$

Now, $n^{m+1} = n^m n$, so that

$$a^{n^{m+1}} = a^{n^m n} = \left(a^{n^m}\right)^n \equiv \left(b^{n^m}\right)^n = b^{n^m n} = b^{n^{m+1}} \operatorname{mod} n^{m+1}n$$

(since $n^m n = n^{m+1}$). In view of $n^{m+1}n = n^{(m+1)+1}$, this rewrites as

$$a^{n^{m+1}} \equiv b^{n^{m+1}} \operatorname{mod} n^{(m+1)+1}.$$

In other words, (87) holds for $k = m + 1$. This completes the induction step. Thus, (87) is proven by induction. Hence, Corollary 2.59 holds. □

Second proof of Corollary 2.59. Forget that we fixed a, b, n and k. We thus must prove

$$\left(a^{n^k} \equiv b^{n^k} \operatorname{mod} n^{k+1} \text{ for all integers } a \text{ and } b \text{ and all } n \in \mathbb{N} \text{ satisfying } a \equiv b \operatorname{mod} n\right) \tag{88}$$

for all $k \in \mathbb{N}$.

We shall prove this by induction on k:

Induction base: Let $n \in \mathbb{N}$. Let a and b be two integers such that $a \equiv b \operatorname{mod} n$. We have $n^0 = 1$ and thus $a^{n^0} = a^1 = a$. Similarly, $b^{n^0} = b$. Thus, $a^{n^0} = a \equiv b = b^{n^0} \operatorname{mod} n$. In other words, $a^{n^0} \equiv b^{n^0} \operatorname{mod} n^{0+1}$ (since $n^{0+1} = n^1 = n$).

Now, forget that we fixed n, a and b. We thus have proven that $a^{n^0} \equiv b^{n^0} \operatorname{mod} n^{0+1}$ for all integers a and b and all $n \in \mathbb{N}$ satisfying $a \equiv b \operatorname{mod} n$. In other words, (88) holds for $k = 0$. This completes the induction base.

Induction step: Let $m \in \mathbb{N}$. Assume that (88) holds for $k = m$. We must prove that (88) holds for $k = m + 1$.

Let $n \in \mathbb{N}$. Let a and b be two integers such that $a \equiv b \operatorname{mod} n$. Now,

$$\begin{aligned} \left(n^2\right)^{m+1} &= n^{2(m+1)} = n^{(m+2)+m} \qquad \left(\text{since } 2\left(m+1\right) = \left(m+2\right)+m\right) \\ &= n^{m+2}n^m, \end{aligned}$$

so that $n^{m+2} \mid \left(n^2\right)^{m+1}$.

We have $n \mid n$. Hence, Corollary 2.57 (applied to $d = n$) yields $a^n \equiv b^n \operatorname{mod} nn$. In other words, $a^n \equiv b^n \operatorname{mod} n^2$ (since $nn = n^2$).

We have assumed that (88) holds for $k = m$. Hence, we can apply (88) to a^n, b^n, n^2 and m instead of a, b, n and k (since $a^n \equiv b^n \operatorname{mod} n^2$). We thus conclude that

$$\left(a^n\right)^{n^m} \equiv \left(b^n\right)^{n^m} \operatorname{mod} \left(n^2\right)^{m+1}.$$

Now, $n^{m+1} = nn^m$, so that

$$a^{n^{m+1}} = a^{nn^m} = \left(a^n\right)^{n^m} \equiv \left(b^n\right)^{n^m} = b^{nn^m} = b^{n^{m+1}} \operatorname{mod} \left(n^2\right)^{m+1}$$

(since $nn^m = n^{m+1}$). Hence, Proposition 2.11 **(c)** (applied to $a^{n^{m+1}}$, $b^{n^{m+1}}$, $\left(n^2\right)^{m+1}$ and n^{m+2} instead of a, b, n and m) yields $a^{n^{m+1}} \equiv b^{n^{m+1}} \operatorname{mod} n^{m+2}$ (since $n^{m+2} \mid \left(n^2\right)^{m+1}$). In view of $m+2 = (m+1)+1$, this rewrites as

$$a^{n^{m+1}} \equiv b^{n^{m+1}} \operatorname{mod} n^{(m+1)+1}.$$

Now, forget that we fixed n, a and b. We thus have proven that $a^{n^{m+1}} \equiv b^{n^{m+1}} \operatorname{mod} n^{(m+1)+1}$ for all integers a and b and all $n \in \mathbb{N}$ satisfying $a \equiv b \operatorname{mod} n$. In other words, (88) holds for $k = m+1$. This completes the induction step. Thus, (88) is proven by induction. Hence, Corollary 2.59 is proven again. □

2.8. Strong induction

2.8.1. The strong induction principle

We shall now show another "alternative induction principle", which is known as the *strong induction principle* because it feels stronger than Theorem 2.1 (in the sense that it appears to get the same conclusion from weaker assumptions). Just as Theorem 2.53, this principle is not a new axiom, but rather a consequence of the standard induction principle; we shall soon deduce it from Theorem 2.53.

Theorem 2.60. Let $g \in \mathbb{Z}$. For each $n \in \mathbb{Z}_{\geq g}$, let $\mathcal{A}(n)$ be a logical statement.

Assume the following:

Assumption 1: If $m \in \mathbb{Z}_{\geq g}$ is such that

$$\left(\mathcal{A}(n) \text{ holds for every } n \in \mathbb{Z}_{\geq g} \text{ satisfying } n < m\right),$$

then $\mathcal{A}(m)$ holds.

Then, $\mathcal{A}(n)$ holds for each $n \in \mathbb{Z}_{\geq g}$.

Notice that Theorem 2.60 has only one assumption (unlike Theorem 2.1 and Theorem 2.53). We shall soon see that this one assumption "incorporates" both an induction base and an induction step.

Let us first explain why Theorem 2.60 is intuitively clear. For example, if you have $g = 4$, and you want to prove (under the assumptions of Theorem 2.60) that $\mathcal{A}(7)$ holds, you can argue as follows:

- We know that $\mathcal{A}(n)$ holds for every $n \in \mathbb{Z}_{\geq 4}$ satisfying $n < 4$. (Indeed, this is vacuously true, since there is no $n \in \mathbb{Z}_{\geq 4}$ satisfying $n < 4$.)

  Hence, Assumption 1 (applied to $m = 4$) shows that the statement $\mathcal{A}(4)$ holds.

- Thus, we know that $\mathcal{A}(n)$ holds for every $n \in \mathbb{Z}_{\geq 4}$ satisfying $n < 5$ (because $\mathcal{A}(4)$ holds).

  Hence, Assumption 1 (applied to $m = 5$) shows that the statement $\mathcal{A}(5)$ holds.

- Thus, we know that $\mathcal{A}(n)$ holds for every $n \in \mathbb{Z}_{\geq 4}$ satisfying $n < 6$ (because $\mathcal{A}(4)$ and $\mathcal{A}(5)$ hold).

  Hence, Assumption 1 (applied to $m = 6$) shows that the statement $\mathcal{A}(6)$ holds.

- Thus, we know that $\mathcal{A}(n)$ holds for every $n \in \mathbb{Z}_{\geq 4}$ satisfying $n < 7$ (because $\mathcal{A}(4)$, $\mathcal{A}(5)$ and $\mathcal{A}(6)$ hold).

  Hence, Assumption 1 (applied to $m = 7$) shows that the statement $\mathcal{A}(7)$ holds.

A similar (but longer) argument shows that the statement $\mathcal{A}(8)$ holds; likewise, $\mathcal{A}(n)$ can be shown to hold for each $n \in \mathbb{Z}_{\geq g}$ by means of an argument that takes $n - g + 1$ steps.

It is easy to see that Theorem 2.60 generalizes Theorem 2.53 (because if the two Assumptions 1 and 2 of Theorem 2.53 hold, then so does Assumption 1 of Theorem 2.60). More interesting for us is the converse implication: We shall show that Theorem 2.60 can be derived from Theorem 2.53. This will allow us to use Theorem 2.60 without having to taking it on trust.

Before we derive Theorem 2.60, let us restate Theorem 2.53 as follows:

Corollary 2.61. Let $g \in \mathbb{Z}$. For each $n \in \mathbb{Z}_{\geq g}$, let $\mathcal{B}(n)$ be a logical statement.

Assume the following:

Assumption A: The statement $\mathcal{B}(g)$ holds.

Assumption B: If $p \in \mathbb{Z}_{\geq g}$ is such that $\mathcal{B}(p)$ holds, then $\mathcal{B}(p+1)$ also holds.

Then, $\mathcal{B}(n)$ holds for each $n \in \mathbb{Z}_{\geq g}$.

Proof of Corollary 2.61. Corollary 2.61 is exactly Theorem 2.53, except that some names have been changed:

- The statements $\mathcal{A}(n)$ have been renamed as $\mathcal{B}(n)$.
- Assumption 1 and Assumption 2 have been renamed as Assumption A and Assumption B.
- The variable m in Assumption B has been renamed as p.

Thus, Corollary 2.61 holds (since Theorem 2.53 holds). □

Let us now derive Theorem 2.60 from Theorem 2.53:

Proof of Theorem 2.60. For each $n \in \mathbb{Z}_{\geq g}$, we let $\mathcal{B}(n)$ be the statement

$$\left(\mathcal{A}(q) \text{ holds for every } q \in \mathbb{Z}_{\geq g} \text{ satisfying } q < n\right).$$

Now, let us consider the Assumptions A and B from Corollary 2.61. We claim that both of these assumptions are satisfied.

The statement $\mathcal{B}(g)$ holds[65]. Thus, Assumption A is satisfied.

Next, let us prove that Assumption B is satisfied. Indeed, let $p \in \mathbb{Z}_{\geq g}$ be such that $\mathcal{B}(p)$ holds. We shall show that $\mathcal{B}(p+1)$ also holds.

Indeed, we have assumed that $\mathcal{B}(p)$ holds. In other words,

$$\mathcal{A}(q) \text{ holds for every } q \in \mathbb{Z}_{\geq g} \text{ satisfying } q < p \tag{89}$$

(because the statement $\mathcal{B}(p)$ is defined as $\left(\mathcal{A}(q) \text{ holds for every } q \in \mathbb{Z}_{\geq g} \text{ satisfying } q < p\right)$). Renaming the variable q as n in this statement, we conclude that

$$\mathcal{A}(n) \text{ holds for every } n \in \mathbb{Z}_{\geq g} \text{ satisfying } n < p. \tag{90}$$

Hence, Assumption 1 (applied to $m = p$) yields that $\mathcal{A}(p)$ holds.

Now, we claim that

$$\mathcal{A}(q) \text{ holds for every } q \in \mathbb{Z}_{\geq g} \text{ satisfying } q < p+1. \tag{91}$$

[*Proof of (91):* Let $q \in \mathbb{Z}_{\geq g}$ be such that $q < p+1$. We must prove that $\mathcal{A}(q)$ holds.

[65] *Proof.* Let $q \in \mathbb{Z}_{\geq g}$ be such that $q < g$. Then, $q \geq g$ (since $q \in \mathbb{Z}_{\geq g}$); but this contradicts $q < g$.

Now, forget that we fixed q. We thus have found a contradiction for each $q \in \mathbb{Z}_{\geq g}$ satisfying $q < g$. Hence, there exists no $q \in \mathbb{Z}_{\geq g}$ satisfying $q < g$. Thus, the statement

$$\left(\mathcal{A}(q) \text{ holds for every } q \in \mathbb{Z}_{\geq g} \text{ satisfying } q < g\right)$$

is vacuously true, and therefore true. In other words, the statement $\mathcal{B}(g)$ is true (since $\mathcal{B}(g)$ is defined as the statement $\left(\mathcal{A}(q) \text{ holds for every } q \in \mathbb{Z}_{\geq g} \text{ satisfying } q < g\right)$). Qed.

If $q = p$, then this follows from the fact that $\mathcal{A}(p)$ holds. Hence, for the rest of this proof, we WLOG assume that we don't have $q = p$. Thus, $q \neq p$. But $q < p+1$ and therefore $q \leq (p+1) - 1$ (since q and $p+1$ are integers). Hence, $q \leq (p+1) - 1 = p$. Combining this with $q \neq p$, we obtain $q < p$. Hence, (89) shows that $\mathcal{A}(q)$ holds. This completes the proof of (91).]

But the statement $\mathcal{B}(p+1)$ is defined as $\left(\mathcal{A}(q) \text{ holds for every } q \in \mathbb{Z}_{\geq g} \text{ satisfying } q < p+1\right)$. In other words, the statement $\mathcal{B}(p+1)$ is precisely the statement (91). Hence, the statement $\mathcal{B}(p+1)$ holds (since (91) holds).

Now, forget that we fixed p. We thus have shown that if $p \in \mathbb{Z}_{\geq g}$ is such that $\mathcal{B}(p)$ holds, then $\mathcal{B}(p+1)$ also holds. In other words, Assumption B is satisfied.

We now know that both Assumption A and Assumption B are satisfied. Hence, Corollary 2.61 shows that

$$\mathcal{B}(n) \text{ holds for each } n \in \mathbb{Z}_{\geq g}. \tag{92}$$

Now, let $n \in \mathbb{Z}_{\geq g}$. Thus, n is an integer such that $n \geq g$ (by the definition of $\mathbb{Z}_{\geq g}$). Hence, $n+1$ is also an integer and satisfies $n+1 \geq n \geq g$, so that $n+1 \in \mathbb{Z}_{\geq g}$. Hence, (92) (applied to $n+1$ instead of n) shows that $\mathcal{B}(n+1)$ holds. In other words,

$$\mathcal{A}(q) \text{ holds for every } q \in \mathbb{Z}_{\geq g} \text{ satisfying } q < n+1$$

(because the statement $\mathcal{B}(n+1)$ is defined as $\left(\mathcal{A}(q) \text{ holds for every } q \in \mathbb{Z}_{\geq g} \text{ satisfying } q < n+1\right)$). We can apply this to $q = n$ (because $n \in \mathbb{Z}_{\geq g}$ satisfies $n < n+1$), and conclude that $\mathcal{A}(n)$ holds.

Now, forget that we fixed n. We thus have shown that $\mathcal{A}(n)$ holds for each $n \in \mathbb{Z}_{\geq g}$. This proves Theorem 2.60. □

Thus, proving a sequence of statements $\mathcal{A}(0), \mathcal{A}(1), \mathcal{A}(2), \ldots$ using Theorem 2.60 is tantamount to proving a slightly different sequence of statements $\mathcal{B}(0), \mathcal{B}(1), \mathcal{B}(2), \ldots$ using Corollary 2.61 and then deriving the former from the latter.

Theorem 2.53 is called the *principle of strong induction starting at g*, and proofs that use it are usually called *proofs by strong induction*. We illustrate its use on the following easy property of the Fibonacci sequence:

Proposition 2.62. Let $(f_0, f_1, f_2, \ldots)$ be the Fibonacci sequence (defined as in Example 2.25). Then,

$$f_n \leq 2^{n-1} \tag{93}$$

for each $n \in \mathbb{N}$.

Proof of Proposition 2.62. For each $n \in \mathbb{Z}_{\geq 0}$, we let $\mathcal{A}(n)$ be the statement $\left(f_n \leq 2^{n-1}\right)$.

Thus, $\mathcal{A}(0)$ is the statement $\left(f_0 \leq 2^{0-1}\right)$; hence, this statement holds (since $f_0 = 0 \leq 2^{0-1}$).

Also, $\mathcal{A}(1)$ is the statement $\left(f_1 \leq 2^{1-1}\right)$ (by the definition of $\mathcal{A}(1)$); hence, this statement also holds (since $f_1 = 1 = 2^{1-1}$).

Now, we claim the following:

Claim 1: If $m \in \mathbb{Z}_{\geq 0}$ is such that

$$\left(\mathcal{A}\left(n\right) \text{ holds for every } n \in \mathbb{Z}_{\geq 0} \text{ satisfying } n < m\right),$$

then $\mathcal{A}\left(m\right)$ holds.

[*Proof of Claim 1:* Let $m \in \mathbb{Z}_{\geq 0}$ be such that

$$\left(\mathcal{A}\left(n\right) \text{ holds for every } n \in \mathbb{Z}_{\geq 0} \text{ satisfying } n < m\right). \tag{94}$$

We must prove that $\mathcal{A}\left(m\right)$ holds.

This is true if $m \in \{0,1\}$ (because we have shown that both statements $\mathcal{A}\left(0\right)$ and $\mathcal{A}\left(1\right)$ hold). Thus, for the rest of the proof of Claim 1, we WLOG assume that we don't have $m \in \{0,1\}$. Hence, $m \in \mathbb{N} \setminus \{0,1\} = \{2,3,4,\ldots\}$, so that $m \geq 2$.

From $m \geq 2$, we conclude that $m-1 \geq 2-1 = 1 \geq 0$ and $m-2 \geq 2-2 = 0$. Thus, both $m-1$ and $m-2$ belong to $\mathbb{N}$; therefore, f_{m-1} and f_{m-2} are well-defined.

We have $m-1 \in \mathbb{N} = \mathbb{Z}_{\geq 0}$ and $m-1 < m$. Hence, (94) (applied to $n = m-1$) yields that $\mathcal{A}\left(m-1\right)$ holds. In other words, $f_{m-1} \leq 2^{(m-1)-1}$ (because this is what the statement $\mathcal{A}\left(m-1\right)$ says).

We have $m-2 \in \mathbb{N} = \mathbb{Z}_{\geq 0}$ and $m-2 < m$. Hence, (94) (applied to $n = m-2$) yields that $\mathcal{A}\left(m-2\right)$ holds. In other words, $f_{m-2} \leq 2^{(m-2)-1}$ (because this is what the statement $\mathcal{A}\left(m-2\right)$ says).

We have $(m-1)-1 = m-2$ and thus $2^{(m-1)-1} = 2^{m-2} = 2 \cdot 2^{(m-2)-1} \geq 2^{(m-2)-1}$ (since $2 \cdot 2^{(m-2)-1} - 2^{(m-2)-1} = 2^{(m-2)-1} \geq 0$). Hence, $2^{(m-2)-1} \leq 2^{(m-1)-1}$.

But the recursive definition of the Fibonacci sequence yields $f_m = f_{m-1} + f_{m-2}$ (since $m \geq 2$). Hence,

$$f_m = \underbrace{f_{m-1}}_{\leq 2^{(m-1)-1}} + \underbrace{f_{m-2}}_{\leq 2^{(m-2)-1} \leq 2^{(m-1)-1}} \leq 2^{(m-1)-1} + 2^{(m-1)-1} = 2 \cdot 2^{(m-1)-1} = 2^{m-1}.$$

In other words, the statement $\mathcal{A}\left(m\right)$ holds (since the statement $\mathcal{A}\left(m\right)$ is defined to be $\left(f_m \leq 2^{m-1}\right)$). This completes the proof of Claim 1.]

Claim 1 shows that Assumption 1 of Theorem 2.60 (applied to $g = 0$) is satisfied. Hence, Theorem 2.60 (applied to $g = 0$) shows that $\mathcal{A}\left(n\right)$ holds for each $n \in \mathbb{Z}_{\geq 0}$. In other words, $f_n \leq 2^{n-1}$ holds for each $n \in \mathbb{Z}_{\geq 0}$ (since the statement $\mathcal{A}\left(n\right)$ is defined to be $\left(f_n \leq 2^{n-1}\right)$). In other words, $f_n \leq 2^{n-1}$ holds for each $n \in \mathbb{N}$ (since $\mathbb{Z}_{\geq 0} = \mathbb{N}$). This proves Proposition 2.62. □

2.8.2. Conventions for writing strong induction proofs

Again, when using the principle of strong induction, one commonly does not directly cite Theorem 2.60; instead one uses the following language:

Convention 2.63. Let $g \in \mathbb{Z}$. For each $n \in \mathbb{Z}_{\geq g}$, let $\mathcal{A}(n)$ be a logical statement. Assume that you want to prove that $\mathcal{A}(n)$ holds for each $n \in \mathbb{Z}_{\geq g}$.

Theorem 2.60 offers the following strategy for proving this: Show that Assumption 1 of Theorem 2.60 is satisfied; then, Theorem 2.60 automatically completes your proof.

A proof that follows this strategy is called a *proof by strong induction on n starting at g*. The proof that Assumption 1 is satisfied is called the *induction step* of the proof. This kind of proof does not have an "induction base" (unlike proofs that use Theorem 2.1 or Theorem 2.53).[66]

In order to prove that Assumption 1 is satisfied, you will usually want to fix an $m \in \mathbb{Z}_{\geq g}$ such that

$$\left(\mathcal{A}(n) \text{ holds for every } n \in \mathbb{Z}_{\geq g} \text{ satisfying } n < m\right), \tag{95}$$

and then prove that $\mathcal{A}(m)$ holds. In other words, you will usually want to fix $m \in \mathbb{Z}_{\geq g}$, assume that (95) holds, and then prove that $\mathcal{A}(m)$ holds. When doing so, it is common to refer to the assumption that (95) holds as the *induction hypothesis* (or *induction assumption*).

Using this language, we can rewrite our above proof of Proposition 2.62 as follows:

Proof of Proposition 2.62 (second version). For each $n \in \mathbb{Z}_{\geq 0}$, we let $\mathcal{A}(n)$ be the statement $\left(f_n \leq 2^{n-1}\right)$. Thus, our goal is to prove the statement $\mathcal{A}(n)$ for each $n \in \mathbb{N}$. In other words, our goal is to prove the statement $\mathcal{A}(n)$ for each $n \in \mathbb{Z}_{\geq 0}$ (since $\mathbb{N} = \mathbb{Z}_{\geq 0}$).

We shall prove this by strong induction on n starting at 0:

Induction step: Let $m \in \mathbb{Z}_{\geq 0}$. Assume that

$$\left(\mathcal{A}(n) \text{ holds for every } n \in \mathbb{Z}_{\geq 0} \text{ satisfying } n < m\right). \tag{96}$$

We must then show that $\mathcal{A}(m)$ holds. In other words, we must show that $f_m \leq 2^{m-1}$ holds (since the statement $\mathcal{A}(m)$ is defined as $\left(f_m \leq 2^{m-1}\right)$).

This is true if $m = 0$ (since $f_0 = 0 \leq 2^{0-1}$) and also true if $m = 1$ (since $f_1 = 1 = 2^{1-1}$ and thus $f_1 \leq 2^{1-1}$). In other words, this is true if $m \in \{0, 1\}$. Thus, for the rest of the induction step, we WLOG assume that we don't have $m \in \{0, 1\}$. Hence, $m \notin \{0, 1\}$, so that $m \in \mathbb{N} \setminus \{0, 1\} = \{2, 3, 4, \ldots\}$. Hence, $m \geq 2$.

From $m \geq 2$, we conclude that $m - 1 \geq 2 - 1 = 1 \geq 0$ and $m - 2 \geq 2 - 2 = 0$. Thus, both $m - 1$ and $m - 2$ belong to $\mathbb{N}$; therefore, f_{m-1} and f_{m-2} are well-defined.

We have $m - 1 \in \mathbb{N} = \mathbb{Z}_{\geq 0}$ and $m - 1 < m$. Hence, (96) (applied to $n = m - 1$) yields that $\mathcal{A}(m-1)$ holds. In other words, $f_{m-1} \leq 2^{(m-1)-1}$ (because this is what the statement $\mathcal{A}(m-1)$ says).

[66] There is a version of strong induction which does include an induction base (or even several). But the version we are using does not.

We have $m-2 \in \mathbb{N} = \mathbb{Z}_{\geq 0}$ and $m-2 < m$. Hence, (96) (applied to $n = m-2$) yields that $\mathcal{A}(m-2)$ holds. In other words, $f_{m-2} \leq 2^{(m-2)-1}$ (because this is what the statement $\mathcal{A}(m-2)$ says).

We have $(m-1)-1 = m-2$ and thus $2^{(m-1)-1} = 2^{m-2} = 2 \cdot 2^{(m-2)-1} \geq 2^{(m-2)-1}$ (since $2 \cdot 2^{(m-2)-1} - 2^{(m-2)-1} = 2^{(m-2)-1} \geq 0$). Hence, $2^{(m-2)-1} \leq 2^{(m-1)-1}$.

But the recursive definition of the Fibonacci sequence yields $f_m = f_{m-1} + f_{m-2}$ (since $m \geq 2$). Hence,

$$f_m = \underbrace{f_{m-1}}_{\leq 2^{(m-1)-1}} + \underbrace{f_{m-2}}_{\leq 2^{(m-2)-1} \leq 2^{(m-1)-1}} \leq 2^{(m-1)-1} + 2^{(m-1)-1} = 2 \cdot 2^{(m-1)-1} = 2^{m-1}.$$

In other words, the statement $\mathcal{A}(m)$ holds (since the statement $\mathcal{A}(m)$ is defined to be $\left(f_m \leq 2^{m-1}\right)$).

Now, forget that we fixed m. We thus have shown that if $m \in \mathbb{Z}_{\geq 0}$ is such that (96) holds, then $\mathcal{A}(m)$ holds. This completes the induction step. Hence, by strong induction, we conclude that $\mathcal{A}(n)$ holds for each $n \in \mathbb{Z}_{\geq 0}$. This completes our proof of Proposition 2.62. □

The proof that we just showed still has a lot of "boilerplate" text that conveys no information. For example, we have again explicitly defined the statement $\mathcal{A}(n)$, which is unnecessary: This statement is exactly what one would expect (namely, the claim that we are proving, without the "for each $n \in \mathbb{N}$" part). Thus, in our case, this statement is simply (93). Furthermore, we can remove the two sentences

> "Now, forget that we fixed m. We thus have shown that if $m \in \mathbb{Z}_{\geq 0}$ is such that (96) holds, then $\mathcal{A}(m)$ holds.".

In fact, these sentences merely say that we have completed the induction step; but this is clear anyway when we say that the induction step is completed.

We said that we are proving our statement "by strong induction on n starting at 0". Again, we can omit the words "starting at 0" here, since this is the only option (because our statement is about all $n \in \mathbb{Z}_{\geq 0}$).

Finally, we can remove the words "*Induction step:*", because a proof by strong induction (unlike a proof by standard induction) does not have an induction base (so the induction step is all that it consists of).

Thus, our above proof can be shortened to the following:

Proof of Proposition 2.62 (third version). We shall prove (93) by strong induction on n:

Let $m \in \mathbb{Z}_{\geq 0}$. Assume that (93) holds for every $n \in \mathbb{Z}_{\geq 0}$ satisfying $n < m$. We must then show that (93) holds for $n = m$. In other words, we must show that $f_m \leq 2^{m-1}$ holds.

This is true if $m = 0$ (since $f_0 = 0 \leq 2^{0-1}$) and also true if $m = 1$ (since $f_1 = 1 = 2^{1-1}$ and thus $f_1 \leq 2^{1-1}$). In other words, this is true if $m \in \{0, 1\}$. Thus, for the rest of the induction step, we WLOG assume that we don't have $m \in \{0,1\}$. Hence, $m \notin \{0,1\}$, so that $m \in \mathbb{N} \setminus \{0,1\} = \{2,3,4,\ldots\}$. Hence, $m \geq 2$.

From $m \geq 2$, we conclude that $m-1 \geq 2-1 = 1 \geq 0$ and $m-2 \geq 2-2 = 0$. Thus, both $m-1$ and $m-2$ belong to $\mathbb{N}$; therefore, f_{m-1} and f_{m-2} are well-defined.

We have $m-1 \in \mathbb{N} = \mathbb{Z}_{\geq 0}$ and $m-1 < m$. Hence, (93) (applied to $n = m-1$) yields that $f_{m-1} \leq 2^{(m-1)-1}$ (since we have assumed that (93) holds for every $n \in \mathbb{Z}_{\geq 0}$ satisfying $n < m$).

We have $m-2 \in \mathbb{N} = \mathbb{Z}_{\geq 0}$ and $m-2 < m$. Hence, (93) (applied to $n = m-2$) yields that $f_{m-2} \leq 2^{(m-2)-1}$ (since we have assumed that (93) holds for every $n \in \mathbb{Z}_{\geq 0}$ satisfying $n < m$).

We have $(m-1)-1 = m-2$ and thus $2^{(m-1)-1} = 2^{m-2} = 2 \cdot 2^{(m-2)-1} \geq 2^{(m-2)-1}$ (since $2 \cdot 2^{(m-2)-1} - 2^{(m-2)-1} = 2^{(m-2)-1} \geq 0$). Hence, $2^{(m-2)-1} \leq 2^{(m-1)-1}$.

But the recursive definition of the Fibonacci sequence yields $f_m = f_{m-1} + f_{m-2}$ (since $m \geq 2$). Hence,

$$f_m = \underbrace{f_{m-1}}_{\leq 2^{(m-1)-1}} + \underbrace{f_{m-2}}_{\leq 2^{(m-2)-1} \leq 2^{(m-1)-1}} \leq 2^{(m-1)-1} + 2^{(m-1)-1} = 2 \cdot 2^{(m-1)-1} = 2^{m-1}.$$

In other words, (93) holds for $n = m$. This completes the induction step. Hence, by strong induction, we conclude that (93) holds for each $n \in \mathbb{Z}_{\geq 0}$. In other words, (93) holds for each $n \in \mathbb{N}$ (since $\mathbb{Z}_{\geq 0} = \mathbb{N}$). This completes our proof of Proposition 2.62. □

2.9. Two unexpected integralities

2.9.1. The first integrality

We shall illustrate strong induction on two further examples, which both have the form of an "unexpected integrality": A sequence of rational numbers is defined recursively by an equation that involves fractions, but it turns out that all the entries of the sequence are nevertheless integers. There is by now a whole genre of such results (see [Gale98, Chapter 1] for an introduction[67]), and many of them are connected with recent research in the theory of cluster algebras (see [Lampe13] for an introduction).

The first of these examples is the following result:

Proposition 2.64. Define a sequence $(t_0, t_1, t_2, \ldots)$ of positive rational numbers recursively by setting

$$t_0 = 1, \qquad t_1 = 1, \qquad t_2 = 1, \qquad \text{and}$$
$$t_n = \frac{1 + t_{n-1}t_{n-2}}{t_{n-3}} \qquad \text{for each } n \geq 3.$$

[67] See also [FomZel02] for a seminal research paper at a more advanced level.

(Thus,

$$\begin{aligned} t_3 &= \frac{1+t_2t_1}{t_0} = \frac{1+1\cdot 1}{1} = 2; \\ t_4 &= \frac{1+t_3t_2}{t_1} = \frac{1+2\cdot 1}{1} = 3; \\ t_5 &= \frac{1+t_4t_3}{t_2} = \frac{1+3\cdot 2}{1} = 7; \\ t_6 &= \frac{1+t_5t_4}{t_3} = \frac{1+7\cdot 3}{2} = 11, \end{aligned}$$

and so on.) Then:

(a) We have $t_{n+2} = 4t_n - t_{n-2}$ for each $n \in \mathbb{Z}_{\geq 2}$.

(b) We have $t_n \in \mathbb{N}$ for each $n \in \mathbb{N}$.

Note that the sequence $(t_0, t_1, t_2, \ldots)$ in Proposition 2.64 is clearly well-defined, because the expression $\dfrac{1+t_{n-1}t_{n-2}}{t_{n-3}}$ always yields a well-defined positive rational number when $t_{n-1}, t_{n-2}, t_{n-3}$ are positive rational numbers. (In particular, the denominator t_{n-3} of this fraction is $\neq 0$ because it is positive.) In contrast, if we had set $t_n = \dfrac{1-t_{n-1}t_{n-2}}{t_{n-3}}$ instead of $t_n = \dfrac{1+t_{n-1}t_{n-2}}{t_{n-3}}$, then the sequence would **not** be well-defined (because then, we would get $t_3 = \dfrac{1-1\cdot 1}{1} = 0$ and $t_6 = \dfrac{1-t_5t_4}{t_3} = \dfrac{1-t_5t_4}{0}$, which is undefined).

Remark 2.65. The sequence $(t_0, t_1, t_2, \ldots)$ defined in Proposition 2.64 is the sequence A005246 in the OEIS (Online Encyclopedia of Integer Sequences). Its first entries are

$$\begin{array}{lllll} t_0 = 1, & t_1 = 1, & t_2 = 1, & t_3 = 2, & t_4 = 3, \\ t_5 = 7, & t_6 = 11, & t_7 = 26, & t_8 = 41, & t_9 = 97. \end{array}$$

Proposition 2.64 **(b)** is an instance of the *Laurent phenomenon* (see, e.g., [FomZel02, Example 3.2]).

Part **(a)** of Proposition 2.64 is proven by a (regular) induction; it is part **(b)** where strong induction comes handy:

Proof of Proposition 2.64. First, we notice that the recursive definition of the sequence $(t_0, t_1, t_2, \ldots)$ yields

$$t_3 = \frac{1+t_{3-1}t_{3-2}}{t_{3-3}} = \frac{1+t_2t_1}{t_0} = \frac{1+1\cdot 1}{1} \qquad (\text{since } t_0 = 1 \text{ and } t_1 = 1 \text{ and } t_2 = 1)$$
$$= 2.$$

Furthermore, the recursive definition of the sequence $(t_0, t_1, t_2, \ldots)$ yields

$$t_4 = \frac{1 + t_{4-1}t_{4-2}}{t_{4-3}} = \frac{1 + t_3 t_2}{t_1} = \frac{1 + 2 \cdot 1}{1} \qquad \text{(since } t_1 = 1 \text{ and } t_2 = 1 \text{ and } t_3 = 2\text{)}$$
$$= 3.$$

Thus, $t_{2+2} = t_4 = 3$. Comparing this with $4 \underbrace{t_2}_{=1} - \underbrace{t_{2-2}}_{=t_0=1} = 4 \cdot 1 - 1 = 3$, we obtain $t_{2+2} = 4t_2 - t_{2-2}$.

(a) We shall prove Proposition 2.64 **(a)** by induction on n starting at 2:

Induction base: We have already shown that $t_{2+2} = 4t_2 - t_{2-2}$. In other words, Proposition 2.64 **(a)** holds for $n = 2$. This completes the induction base.

Induction step: Let $m \in \mathbb{Z}_{\geq 2}$. Assume that Proposition 2.64 **(a)** holds for $n = m$. We must prove that Proposition 2.64 **(a)** holds for $n = m + 1$.

We have assumed that Proposition 2.64 **(a)** holds for $n = m$. In other words, we have $t_{m+2} = 4t_m - t_{m-2}$.

We have $m \in \mathbb{Z}_{\geq 2}$. Thus, m is an integer that is ≥ 2. Hence, $m \geq 2$ and thus $m + 1 \geq 2 + 1 = 3$. Thus, the recursive definition of the sequence $(t_0, t_1, t_2, \ldots)$ yields

$$t_{m+1} = \frac{1 + t_{(m+1)-1}t_{(m+1)-2}}{t_{(m+1)-3}} = \frac{1 + t_m t_{m-1}}{t_{m-2}}.$$

Multiplying this equality by t_{m-2}, we obtain $t_{m-2}t_{m+1} = 1 + t_m t_{m-1}$. In other words,

$$t_{m-2}t_{m+1} - 1 = t_m t_{m-1}. \tag{97}$$

Hence,

$$1 + \underbrace{t_{m+2}}_{=4t_m - t_{m-2}} t_{m+1} = 1 + (4t_m - t_{m-2})\, t_{m+1} = 4t_m t_{m+1} - \underbrace{(t_{m-2}t_{m+1} - 1)}_{\substack{=t_m t_{m-1} \\ \text{(by (97))}}}$$
$$= 4t_m t_{m+1} - t_m t_{m-1} = t_m (4t_{m+1} - t_{m-1}). \tag{98}$$

Also, $m + 3 \geq 3$. Thus, the recursive definition of the sequence $(t_0, t_1, t_2, \ldots)$ yields

$$t_{m+3} = \frac{1 + t_{(m+3)-1}t_{(m+3)-2}}{t_{(m+3)-3}} = \frac{1 + t_{m+2}t_{m+1}}{t_m} = \frac{1}{t_m} \underbrace{(1 + t_{m+2}t_{m+1})}_{\substack{=t_m(4t_{m+1} - t_{m-1}) \\ \text{(by (98))}}}$$
$$= \frac{1}{t_m} t_m (4t_{m+1} - t_{m-1}) = 4t_{m+1} - t_{m-1}.$$

In view of $m + 3 = (m + 1) + 2$ and $m - 1 = (m + 1) - 2$, this rewrites as $t_{(m+1)+2} = 4t_{m+1} - t_{(m+1)-2}$. In other words, Proposition 2.64 **(a)** holds for $n = m + 1$. This completes the induction step. Hence, Proposition 2.64 **(a)** is proven by induction.

(b) We shall prove Proposition 2.64 **(b)** by strong induction on n starting at 0:

Induction step: Let $m \in \mathbb{N}$. [68] Assume that Proposition 2.64 **(b)** holds for every $n \in \mathbb{N}$ satisfying $n < m$. We must now show that Proposition 2.64 **(b)** holds for $n = m$.

We have assumed that Proposition 2.64 **(b)** holds for every $n \in \mathbb{N}$ satisfying $n < m$. In other words, we have

$$t_n \in \mathbb{N} \text{ for every } n \in \mathbb{N} \text{ satisfying } n < m. \tag{99}$$

We must now show that Proposition 2.64 **(b)** holds for $n = m$. In other words, we must show that $t_m \in \mathbb{N}$.

Recall that $(t_0, t_1, t_2, \ldots)$ is a sequence of positive rational numbers. Thus, t_m is a positive rational number.

We are in one of the following five cases:

Case 1: We have $m = 0$.

Case 2: We have $m = 1$.

Case 3: We have $m = 2$.

Case 4: We have $m = 3$.

Case 5: We have $m > 3$.

Let us first consider Case 1. In this case, we have $m = 0$. Thus, $t_m = t_0 = 1 \in \mathbb{N}$. Hence, $t_m \in \mathbb{N}$ is proven in Case 1.

Similarly, we can prove $t_m \in \mathbb{N}$ in Case 2 (using $t_1 = 1$) and in Case 3 (using $t_2 = 1$) and in Case 4 (using $t_3 = 2$). It thus remains to prove $t_m \in \mathbb{N}$ in Case 5.

So let us consider Case 5. In this case, we have $m > 3$. Thus, $m \geq 4$ (since m is an integer), so that $m - 2 \geq 4 - 2 = 2$. Thus, $m - 2$ is an integer that is ≥ 2. In other words, $m - 2 \in \mathbb{Z}_{\geq 2}$. Hence, Proposition 2.64 **(a)** (applied to $n = m - 2$) yields $t_{(m-2)+2} = 4t_{m-2} - t_{(m-2)-2}$. In view of $(m-2)+2 = m$ and $(m-2)-2 = m-4$, this rewrites as $t_m = 4t_{m-2} - t_{m-4}$.

But $m \geq 4$, so that $m - 4 \in \mathbb{N}$, and $m - 4 < m$. Hence, (99) (applied to $n = m - 4$) yields $t_{m-4} \in \mathbb{N} \subseteq \mathbb{Z}$. Similarly, $t_{m-2} \in \mathbb{Z}$.

So we know that t_{m-2} and t_{m-4} are both integers (since $t_{m-2} \in \mathbb{Z}$ and $t_{m-4} \in \mathbb{Z}$). Hence, $4t_{m-2} - t_{m-4}$ is an integer as well. In other words, t_m is an integer (because $t_m = 4t_{m-2} - t_{m-4}$). Since t_m is positive, we thus conclude that t_m is a positive integer. Hence, $t_m \in \mathbb{N}$. This shows that $t_m \in \mathbb{N}$ in Case 5.

We now have proven $t_m \in \mathbb{N}$ in each of the five Cases 1, 2, 3, 4 and 5. Since these five Cases cover all possibilities, we thus conclude that $t_m \in \mathbb{N}$ always holds. In other words, Proposition 2.64 **(b)** holds for $n = m$. This completes the induction step. Thus, Proposition 2.64 **(b)** is proven by strong induction. □

2.9.2. The second integrality

Our next example of an "unexpected integrality" is the following fact:

[68] In order to match the notations used in Theorem 2.60, we should be saying "Let $m \in \mathbb{Z}_{\geq 0}$" here, rather than "Let $m \in \mathbb{N}$". But of course, this amounts to the same thing, since $\mathbb{N} = \mathbb{Z}_{\geq 0}$.

Proposition 2.66. Fix a positive integer r. Define a sequence $(b_0, b_1, b_2, \ldots)$ of positive rational numbers recursively by setting

$$b_0 = 1, \qquad b_1 = 1, \qquad \text{and}$$
$$b_n = \frac{b_{n-1}^r + 1}{b_{n-2}} \qquad \text{for each } n \geq 2.$$

(Thus,

$$\begin{aligned} b_2 &= \frac{b_1^r + 1}{b_0} = \frac{1^r + 1}{1} = 2; \\ b_3 &= \frac{b_2^r + 1}{b_1} = \frac{2^r + 1}{1} = 2^r + 1; \\ b_4 &= \frac{b_3^r + 1}{b_2} = \frac{(2^r + 1)^r + 1}{2}, \end{aligned}$$

and so on.) Then:

(a) We have $b_n \in \mathbb{N}$ for each $n \in \mathbb{N}$.

(b) If $r \geq 2$, then $b_n \mid b_{n-2} + b_{n+2}$ for each $n \in \mathbb{Z}_{\geq 2}$.

Remark 2.67. If $r = 1$, then the sequence $(b_0, b_1, b_2, \ldots)$ defined in Proposition 2.66 is

$$(1, 1, 2, 3, 2, 1, 1, 2, 3, 2, 1, 1, 2, 3, 2, \ldots)$$

(this is a periodic sequence, which consists of the five terms $1, 1, 2, 3, 2$ repeated over and over); this can easily be proven by strong induction. Despite its simplicity, this sequence is the sequence A076839 in the OEIS.

If $r = 2$, then the sequence $(b_0, b_1, b_2, \ldots)$ defined in Proposition 2.66 is

$$(1, f_1, f_3, f_5, f_7, \ldots) = (1, 1, 2, 5, 13, 34, 89, 233, 610, 1597, \ldots)$$

consisting of all Fibonacci numbers at odd positions (i.e., Fibonacci numbers of the form f_{2n-1} for $n \in \mathbb{N}$) with an extra 1 at the front. This, again, can be proven by induction. Also, this sequence satisfies the recurrence relation $b_n = 3b_{n-1} - b_{n-2}$ for all $n \geq 2$. This is the sequence A001519 in the OEIS.

If $r = 3$, then the sequence $(b_0, b_1, b_2, \ldots)$ defined in Proposition 2.66 is

$$(1, 1, 2, 9, 365, 5403014, 432130991537958813, \ldots);$$

its entries grow so fast that the next entry would need a separate line. This is the sequence A003818 in the OEIS. Unlike the cases of $r = 1$ and $r = 2$, not much can be said about this sequence, other than what has been said in Proposition 2.66.

Proposition 2.66 **(a)** is an instance of the *Laurent phenomenon for cluster algebras* (see, e.g., [FomZel01, Example 2.5]; also, see [Marsh13] and [FoWiZe16] for expositions). See also [MusPro07] for a study of the specific recurrence equation from Proposition 2.66 (actually, a slightly more general equation).

Before we prove Proposition 2.66, let us state an auxiliary fact:

Lemma 2.68. Let $r \in \mathbb{N}$. For every nonzero $x \in \mathbb{Q}$, we set $H(x) = \dfrac{(x+1)^r - 1}{x}$. Then, $H(x) \in \mathbb{Z}$ whenever x is a nonzero integer.

Proof of Lemma 2.68. Let x be a nonzero integer. Then, $x \mid (x+1) - 1$ (because $(x+1) - 1 = x$). In other words, $x + 1 \equiv 1 \bmod x$ (by the definition of "congruent"). Hence, Proposition 2.22 (applied to $a = x+1$, $b = 1$, $n = x$ and $k = r$) shows that $(x+1)^r \equiv 1^r = 1 \bmod x$. In other words, $(x+1)^r - 1$ is divisible by x. In other words, $\dfrac{(x+1)^r - 1}{x}$ is an integer. In other words, $\dfrac{(x+1)^r - 1}{x} \in \mathbb{Z}$. Thus, $H(x) = \dfrac{(x+1)^r - 1}{x} \in \mathbb{Z}$. This proves Lemma 2.68. □

Proof of Proposition 2.66. First, we notice that the recursive definition of the sequence $(b_0, b_1, b_2, \ldots)$ yields

$$\begin{aligned} b_2 &= \frac{b_{2-1}^r + 1}{b_{2-2}} = \frac{b_1^r + 1}{b_0} = \frac{1^r + 1}{1} \qquad (\text{since } b_0 = 1 \text{ and } b_1 = 1) \\ &= \frac{1+1}{1} \qquad (\text{since } 1^r = 1) \\ &= 2. \end{aligned}$$

Furthermore, the recursive definition of the sequence $(b_0, b_1, b_2, \ldots)$ yields

$$\begin{aligned} b_3 &= \frac{b_{3-1}^r + 1}{b_{3-2}} = \frac{b_2^r + 1}{b_1} = \frac{2^r + 1}{1} \qquad (\text{since } b_1 = 1 \text{ and } b_2 = 2) \\ &= 2^r + 1. \end{aligned}$$

For every nonzero $x \in \mathbb{Q}$, we set $H(x) = \dfrac{(x+1)^r - 1}{x}$.

For every integer $m \geq 1$, we have

$$b_m^r + 1 = b_{m+1} b_{m-1}. \tag{100}$$

[*Proof of (100):* Let $m \geq 1$ be an integer. From $m \geq 1$, we obtain $m + 1 \geq 1 + 1 = 2$. Hence, the recursive definition of the sequence $(b_0, b_1, b_2, \ldots)$ yields

$$b_{m+1} = \frac{b_{(m+1)-1}^r + 1}{b_{(m+1)-2}} = \frac{b_m^r + 1}{b_{m-1}}.$$

Multiplying both sides of this equality by b_{m-1}, we obtain $b_{m+1} b_{m-1} = b_m^r + 1$. This proves (100).]

Let us first prove the following observation:

Observation 1: Each integer $n \geq 2$ satisfies $b_{n+2} = b_{n-2} b_{n+1}^r - b_n^{r-1} H(b_n^r)$.

[*Proof of Observation 1:* Let $n \geq 2$ be an integer. Thus, $n \geq 2 \geq 1$. Thus, (100) (applied to $m = n$) yields

$$b_n^r + 1 = b_{n+1} b_{n-1}. \tag{101}$$

On the other hand, $n + 1 \geq n \geq 2 \geq 1$. Hence, (100) (applied to $m = n + 1$) yields

$$b_{n+1}^r + 1 = b_{(n+1)+1} b_{(n+1)-1} = b_{n+2} b_n.$$

Hence,

$$b_{n+1}^r = b_{n+2} b_n - 1. \tag{102}$$

Also, $n - 1 \geq 1$ (since $n \geq 2 = 1 + 1$). Hence, (100) (applied to $m = n - 1$) yields

$$b_{n-1}^r + 1 = b_{(n-1)+1} b_{(n-1)-1} = b_n b_{n-2}.$$

Hence,

$$b_{n-1}^r = b_n b_{n-2} - 1. \tag{103}$$

But b_n is a positive rational number (since $(b_0, b_1, b_2, \ldots)$ is a sequence of positive rational numbers). Thus, b_n^r is also a positive rational number. Hence, $b_n^r \in \mathbb{Q}$ is nonzero. The definition of $H(b_n^r)$ yields $H(b_n^r) = \dfrac{(b_n^r + 1)^r - 1}{b_n^r}$; therefore,

$$\begin{aligned}
b_n^{r-1} \underbrace{H(b_n^r)}_{= \frac{(b_n^r + 1)^r - 1}{b_n^r}} &= b_n^{r-1} \cdot \frac{(b_n^r + 1)^r - 1}{b_n^r} = \underbrace{\frac{b_n^{r-1}}{b_n^r}}_{= \frac{1}{b_n}} \cdot \left(\left(\underbrace{b_n^r + 1}_{\substack{= b_{n+1} b_{n-1} \\ \text{(by (101))}}} \right)^r - 1 \right) \\
&= \frac{1}{b_n} \cdot \left(\underbrace{(b_{n+1} b_{n-1})^r}_{= b_{n+1}^r b_{n-1}^r} - 1 \right) = \frac{1}{b_n} \cdot \left(\underbrace{b_{n+1}^r}_{\substack{= b_{n+2} b_n - 1 \\ \text{(by (102))}}} \underbrace{b_{n-1}^r}_{\substack{= b_n b_{n-2} - 1 \\ \text{(by (103))}}} - 1 \right) \\
&= \frac{1}{b_n} \cdot \underbrace{\left((b_{n+2} b_n - 1)(b_n b_{n-2} - 1) - 1 \right)}_{= b_n (b_n b_{n+2} b_{n-2} - b_{n+2} - b_{n-2})} \\
&= \frac{1}{b_n} \cdot b_n (b_n b_{n+2} b_{n-2} - b_{n+2} - b_{n-2}) \\
&= b_n b_{n+2} b_{n-2} - b_{n+2} - b_{n-2} = b_{n-2} \underbrace{(b_{n+2} b_n - 1)}_{\substack{= b_{n+1}^r \\ \text{(by (102))}}} - b_{n+2} \\
&= b_{n-2} b_{n+1}^r - b_{n+2}.
\end{aligned}$$

Solving this equation for b_{n+2}, we obtain $b_{n+2} = b_{n-2} b_{n+1}^r - b_n^{r-1} H(b_n^r)$. This proves Observation 1.]

(a) We shall prove Proposition 2.66 **(a)** by strong induction on n:

Induction step: Let $m \in \mathbb{N}$. Assume that Proposition 2.66 **(a)** holds for every $n \in \mathbb{N}$ satisfying $n < m$. We must now prove that Proposition 2.66 **(a)** holds for $n = m$.

We have assumed that Proposition 2.66 **(a)** holds for every $n \in \mathbb{N}$ satisfying $n < m$. In other words, we have

$$b_n \in \mathbb{N} \text{ for every } n \in \mathbb{N} \text{ satisfying } n < m. \tag{104}$$

We must now show that Proposition 2.66 **(a)** holds for $n = m$. In other words, we must show that $b_m \in \mathbb{N}$.

Recall that $(b_0, b_1, b_2, \ldots)$ is a sequence of positive rational numbers. Thus, b_m is a positive rational number.

We are in one of the following five cases:

Case 1: We have $m = 0$.

Case 2: We have $m = 1$.

Case 3: We have $m = 2$.

Case 4: We have $m = 3$.

Case 5: We have $m > 3$.

Let us first consider Case 1. In this case, we have $m = 0$. Thus, $b_m = b_0 = 1 \in \mathbb{N}$. Hence, $b_m \in \mathbb{N}$ is proven in Case 1.

Similarly, we can prove $b_m \in \mathbb{N}$ in Case 2 (using $b_1 = 1$) and in Case 3 (using $b_2 = 2$) and in Case 4 (using $b_3 = 2^r + 1$). It thus remains to prove $b_m \in \mathbb{N}$ in Case 5.

So let us consider Case 5. In this case, we have $m > 3$. Thus, $m \geq 4$ (since m is an integer), so that $m - 2 \geq 4 - 2 = 2$. Hence, Observation 1 (applied to $n = m - 2$) yields $b_{(m-2)+2} = b_{(m-2)-2} b^r_{(m-2)+1} - b^{r-1}_{m-2} H\left(b^r_{m-2}\right)$. In view of $(m-2)+2 = m$ and $(m-2)-2 = m-4$ and $(m-2)+1 = m-1$, this rewrites as

$$b_m = b_{m-4} b^r_{m-1} - b^{r-1}_{m-2} H\left(b^r_{m-2}\right). \tag{105}$$

But $m - 2 \in \mathbb{N}$ (since $m \geq 4 \geq 2$) and $m - 2 < m$. Hence, (104) (applied to $n = m - 2$) yields $b_{m-2} \in \mathbb{N} \subseteq \mathbb{Z}$. Also, b_{m-2} is a positive rational number (since $(b_0, b_1, b_2, \ldots)$ is a sequence of positive rational numbers) and thus a positive integer (since $b_{m-2} \in \mathbb{N}$), hence a nonzero integer. Thus, b^r_{m-2} is a nonzero integer as well. Therefore, Lemma 2.68 (applied to $x = b^r_{m-2}$) shows that $H\left(b^r_{m-2}\right) \in \mathbb{Z}$. In other words, $H\left(b^r_{m-2}\right)$ is an integer. Also, $r - 1 \geq 0$ (since $r \geq 1$), and thus $r - 1 \in \mathbb{N}$. Hence, b^{r-1}_{m-2} is an integer (since b_{m-2} is an integer).

Also, $m - 4 \in \mathbb{N}$ (since $m \geq 4$) and $m - 4 < m$. Hence, (104) (applied to $n = m - 4$) yields $b_{m-4} \in \mathbb{N} \subseteq \mathbb{Z}$. In other words, b_{m-4} is an integer.

Similarly, b_{m-1} is an integer. Thus, b^r_{m-1} is an integer.

We now know that the four numbers b_{m-4}, b^r_{m-1}, b^{r-1}_{m-2} and $H\left(b^r_{m-2}\right)$ are integers. Thus, the number $b_{m-4} b^r_{m-1} - b^{r-1}_{m-2} H\left(b^r_{m-2}\right)$ also is an integer (since it is obtained from these four numbers by multiplication and subtraction). In view of (105), this rewrites as follows: The number b_m is an integer. Since b_m is positive, we thus conclude that b_m is a positive integer. Hence, $b_m \in \mathbb{N}$. This shows that $b_m \in \mathbb{N}$ in Case 5.

We now have proven $b_m \in \mathbb{N}$ in each of the five Cases 1, 2, 3, 4 and 5. Thus, $b_m \in \mathbb{N}$ always holds. In other words, Proposition 2.66 **(a)** holds for $n = m$. This completes the induction step. Thus, Proposition 2.66 **(a)** is proven by strong induction.

(b) Assume that $r \geq 2$. We must prove that $b_n \mid b_{n-2} + b_{n+2}$ for each $n \in \mathbb{Z}_{\geq 2}$.

So let $n \in \mathbb{Z}_{\geq 2}$. We must show that $b_n \mid b_{n-2} + b_{n+2}$.

From $n \in \mathbb{Z}_{\geq 2}$, we obtain $n \geq 2$, so that $n - 2 \in \mathbb{N}$.

Proposition 2.66 **(a)** (applied to $n - 2$ instead of n) yields $b_{n-2} \in \mathbb{N}$. Similarly, $b_n \in \mathbb{N}$ and $b_{n+1} \in \mathbb{N}$ and $b_{n+2} \in \mathbb{N}$. Thus, all of b_{n-2}, b_{n+1}, b_n are b_{n+2} are integers.

But b_n is a positive rational number (since $(b_0, b_1, b_2, \ldots)$ is a sequence of positive rational numbers), and therefore a positive integer (since b_n is an integer). Hence, b_n^r is a positive integer, and thus a nonzero integer. Therefore, Lemma 2.68 (applied to $x = b_n^r$) shows that $H(b_n^r) \in \mathbb{Z}$. In other words, $H(b_n^r)$ is an integer.

We have $n + 2 \geq 2$. Hence, the recursive definition of the sequence $(b_0, b_1, b_2, \ldots)$ yields $b_{n+2} = \dfrac{b_{(n+2)-1}^r + 1}{b_{(n+2)-2}} = \dfrac{b_{n+1}^r + 1}{b_n}$. Multiplying this equality by b_n, we obtain

$$b_n b_{n+2} = b_{n+1}^r + 1. \tag{106}$$

We have $r - 2 \in \mathbb{N}$ (since $r \geq 2$) and $b_n \in \mathbb{N}$. Hence, b_n^{r-2} is an integer.

Observation 1 yields $b_{n+2} = b_{n-2} b_{n+1}^r - b_n^{r-1} H(b_n^r)$. Thus,

$$\begin{aligned}
&\underbrace{b_{n+2}}_{= b_{n-2} b_{n+1}^r - b_n^{r-1} H(b_n^r)} + b_{n-2} \\
&= b_{n-2} b_{n+1}^r - b_n^{r-1} H(b_n^r) + b_{n-2} = \underbrace{b_{n-2} b_{n+1}^r + b_{n-2}}_{= b_{n-2}(b_{n+1}^r + 1)} - b_n^{r-1} H(b_n^r) \\
&= b_{n-2} \underbrace{\left(b_{n+1}^r + 1\right)}_{\substack{= b_n b_{n+2} \\ \text{(by (106))}}} - \underbrace{b_n^{r-1}}_{= b_n b_n^{r-2}} H(b_n^r) = b_{n-2} b_n b_{n+2} - b_n b_n^{r-2} H(b_n^r) \\
&= b_n \left(b_{n-2} b_{n+2} - b_n^{r-2} H(b_n^r) \right).
\end{aligned} \tag{107}$$

But $b_{n-2} b_{n+2} - b_n^{r-2} H(b_n^r)$ is an integer (because b_{n-2}, b_{n+2}, b_n^{r-2} and $H(b_n^r)$ are integers). Denote this integer by z. Thus, $z = b_{n-2} b_{n+2} - b_n^{r-2} H(b_n^r)$. Since b_n and

z are integers, we have

$$\begin{aligned} b_n \mid b_n & \underbrace{z}_{=b_{n-2}b_{n+2}-b_n^{r-2}H(b_n^r)} \\ &= b_n \left(b_{n-2}b_{n+2} - b_n^{r-2} H\left(b_n^r\right)\right) = b_{n+2} + b_{n-2} \qquad \text{(by (107))} \\ &= b_{n-2} + b_{n+2}. \end{aligned}$$

This proves Proposition 2.66 **(b)**. □

For a (slightly) different proof of Proposition 2.66, see `http://artofproblemsolving.com/community/c6h428645p3705719` .

Remark 2.69. You might wonder what happens if we replace "$b_{n-1}^r + 1$" by "$b_{n-1}^r + q$" in Proposition 2.66, where q is some fixed nonnegative integer. The answer turns out to be somewhat disappointing in general: For example, if we set $r = 3$ and $q = 2$, then our sequence $(b_0, b_1, b_2, \ldots)$ begins with

$$b_0 = 1, \qquad b_1 = 1, \qquad b_2 = \frac{1^3+2}{1} = 3,$$
$$b_3 = \frac{3^3+2}{1} = 29, \qquad b_4 = \frac{29^3+2}{3} = \frac{24\,391}{3},$$

at which point it becomes clear that $b_n \in \mathbb{N}$ no longer holds for all $n \in \mathbb{N}$. The same happens for all $r > 2$ as long as $q = 3$. (More precisely, if we take $r > 2$ and $q = 2$, then the first n that violates $b_n \in \mathbb{N}$ is 4 or 5 depending on whether r is odd or even. Proving this is a nice exercise!)

However, $b_n \in \mathbb{N}$ still holds for all $n \in \mathbb{N}$ when $r = 2$. This follows from Exercise 2.1 below.

Exercise 2.1. Fix a nonnegative integer q. Define a sequence $(b_0, b_1, b_2, \ldots)$ of positive rational numbers recursively by setting

$$b_0 = 1, \qquad b_1 = 1, \qquad \text{and}$$
$$b_n = \frac{b_{n-1}^2 + q}{b_{n-2}} \qquad \text{for each } n \geq 2.$$

(Thus,

$$\begin{aligned} b_2 &= \frac{b_1^2+q}{b_0} = \frac{1^2+q}{1} = q+1; \\ b_3 &= \frac{b_2^2+q}{b_1} = \frac{(q+1)^2+q}{1} = q^2+3q+1; \\ b_4 &= \frac{b_3^2+q}{b_2} = \frac{\left(q^2+3q+1\right)^2+q}{q+1} = q^3+5q^2+6q+1; \end{aligned}$$

and so on.) Prove that:

(a) We have $b_n = (q+2)\, b_{n-1} - b_{n-2}$ for each $n \in \mathbb{Z}_{\geq 2}$.

(b) We have $b_n \in \mathbb{N}$ for each $n \in \mathbb{N}$.

2.10. Strong induction on a derived quantity: Bezout's theorem

2.10.1. Strong induction on a derived quantity

In Section 2.5, we have seen how to use induction on a variable that does not explicitly appear in the claim. In the current section, we shall show the same for strong induction. This time, the fact that we shall be proving is the following:

Theorem 2.70. Let $a \in \mathbb{N}$ and $b \in \mathbb{N}$. Then, there exist $g \in \mathbb{N}$, $x \in \mathbb{Z}$ and $y \in \mathbb{Z}$ such that $g = ax + by$ and $g \mid a$ and $g \mid b$.

Example 2.71. (a) If $a = 3$ and $b = 5$, then Theorem 2.70 says that there exist $g \in \mathbb{N}$, $x \in \mathbb{Z}$ and $y \in \mathbb{Z}$ such that $g = 3x + 5y$ and $g \mid 3$ and $g \mid 5$. And indeed, it is easy to find such g, x and y: for example, $g = 1$, $x = -3$ and $y = 2$ will do (since $1 = 3(-3) + 5 \cdot 2$ and $1 \mid 3$ and $1 \mid 5$).

(b) If $a = 4$ and $b = 6$, then Theorem 2.70 says that there exist $g \in \mathbb{N}$, $x \in \mathbb{Z}$ and $y \in \mathbb{Z}$ such that $g = 4x + 6y$ and $g \mid 4$ and $g \mid 6$. And indeed, it is easy to find such g, x and y: for example, $g = 2$, $x = -1$ and $y = 1$ will do (since $2 = 4(-1) + 6 \cdot 1$ and $2 \mid 4$ and $2 \mid 6$).

Theorem 2.70 is one form of *Bezout's theorem for integers*, and its real significance might not be clear at this point; it becomes important when the greatest common divisor of two integers is studied. For now, we observe that the g in Theorem 2.70 is obviously a common divisor of a and b (that is, an integer that divides both a and b); but it is also divisible by every common divisor of a and b (because of Proposition 2.7).

Let us now focus on the proof of Theorem 2.70. It is natural to try proving it by induction (or perhaps strong induction) on a or on b, but neither option leads to success. It may feel like "induction on a and on b at the same time" could help, and this is indeed a viable approach[69]. But there is a simpler and shorter method

[69] Of course, it needs to be done right: An induction proof always requires choosing **one** variable to do induction on; but it is possible to nest an induction proof inside the induction step (or inside the induction base) of a different induction proof. For example, imagine that we are trying to prove that

$$ab = ba \qquad \text{for any } a \in \mathbb{N} \text{ and } b \in \mathbb{N}.$$

We can prove this by induction on a. More precisely, for each $a \in \mathbb{N}$, we let $\mathcal{A}(a)$ be the statement ($ab = ba$ for all $b \in \mathbb{N}$). We then prove $\mathcal{A}(a)$ by induction on a. In the induction step, we fix $m \in \mathbb{N}$, and we assume that $\mathcal{A}(m)$ holds; we now need to prove that $\mathcal{A}(m+1)$ holds. In other words, we need to prove that $(m+1)\, b = b\,(m+1)$ for all $b \in \mathbb{N}$. We can now prove this statement by induction on b (although there are easier options, of course). Thus, the induction proof of this statement happens inside the induction step of another induction available: strong induction on $a+b$. As in Section 2.5, the way to formalize such a strong induction is by introducing auxiliary statements $\mathcal{A}(n)$, which say as much as "Theorem 2.70 holds under the requirement that $a+b=n$":

Proof of Theorem 2.70. First of all, let us forget that we fixed a and b. So we want to prove that if $a \in \mathbb{N}$ and $b \in \mathbb{N}$, then there exist $g \in \mathbb{N}$, $x \in \mathbb{Z}$ and $y \in \mathbb{Z}$ such that $g = ax+by$ and $g \mid a$ and $g \mid b$.

For each $n \in \mathbb{N}$, we let $\mathcal{A}(n)$ be the statement

$$\left(\begin{array}{c}\text{if } a \in \mathbb{N} \text{ and } b \in \mathbb{N} \text{ satisfy } a+b=n\text{, then there exist } g \in \mathbb{N},\ x \in \mathbb{Z} \\ \text{and } y \in \mathbb{Z} \text{ such that } g = ax+by \text{ and } g \mid a \text{ and } g \mid b\end{array}\right). \tag{108}$$

We claim that $\mathcal{A}(n)$ holds for all $n \in \mathbb{N}$.

Indeed, let us prove this by strong induction on n starting at 0:

Induction step: Let $m \in \mathbb{N}$. Assume that

$$\left(\mathcal{A}(n) \text{ holds for every } n \in \mathbb{N} \text{ satisfying } n < m\right). \tag{109}$$

We must then show that $\mathcal{A}(m)$ holds.

To do this, we shall prove the following claim:

> *Claim 1:* Let $a \in \mathbb{N}$ and $b \in \mathbb{N}$ satisfy $a+b=m$. Then, there exist $g \in \mathbb{N}$, $x \in \mathbb{Z}$ and $y \in \mathbb{Z}$ such that $g = ax+by$ and $g \mid a$ and $g \mid b$.

Before we prove Claim 1, let us show a slightly weaker version of it, in which we rename a and b as u and v and add the assumption that $u \geq v$:

> *Claim 2:* Let $u \in \mathbb{N}$ and $v \in \mathbb{N}$ satisfy $u+v=m$ and $u \geq v$. Then, there exist $g \in \mathbb{N}$, $x \in \mathbb{Z}$ and $y \in \mathbb{Z}$ such that $g = ux+vy$ and $g \mid u$ and $g \mid v$.

[*Proof of Claim 2:* We are in one of the following two cases:

Case 1: We have $v = 0$.

Case 2: We have $v \neq 0$.

Let us first consider Case 1. In this case, we have $v = 0$. Hence, $v = 0 = 0u$, so that $u \mid v$. Also, $u \cdot 1 + v \cdot 0 = u$. Thus, $u = u \cdot 1 + v \cdot 0$ and $u \mid u$ and $u \mid v$. Hence, there exist $g \in \mathbb{N}$, $x \in \mathbb{Z}$ and $y \in \mathbb{Z}$ such that $g = ux+vy$ and $g \mid u$ and $g \mid v$ (namely, $g = u$, $x = 1$ and $y = 0$). Thus, Claim 2 is proven in Case 1.

Let us now consider Case 2. In this case, we have $v \neq 0$. Hence, $v > 0$ (since $v \in \mathbb{N}$). Thus, $u + v > u + 0 = u$, so that $u < u + v = m$. Hence, (109) (applied to $n = u$) yields that $\mathcal{A}(u)$ holds. In other words,

$$\left(\begin{array}{c}\text{if } a \in \mathbb{N} \text{ and } b \in \mathbb{N} \text{ satisfy } a+b=u\text{, then there exist } g \in \mathbb{N},\ x \in \mathbb{Z} \\ \text{and } y \in \mathbb{Z} \text{ such that } g = ax+by \text{ and } g \mid a \text{ and } g \mid b\end{array}\right) \tag{110}$$

proof. This nesting of induction proofs is legitimate (and even has a name: it is called *double induction*), but tends to be rather confusing (just think about what the sentence "The induction base is complete" means: is it about the induction base of the first induction proof, or that of the second?), and is best avoided when possible.

(because this is what the statement $\mathcal{A}(u)$ says).

Also, $u - v \in \mathbb{N}$ (since $u \geq v$) and $(u - v) + v = u$. Hence, (110) (applied to $a = u - v$ and $b = v$) shows that there exist $g \in \mathbb{N}$, $x \in \mathbb{Z}$ and $y \in \mathbb{Z}$ such that $g = (u - v)x + vy$ and $g \mid u - v$ and $g \mid v$. Consider these g, x and y, and denote them by g', x' and y'. Thus, g' is an element of $\mathbb{N}$, and x' and y' are elements of $\mathbb{Z}$ satisfying $g' = (u - v)x' + vy'$ and $g' \mid u - v$ and $g' \mid v$.

Now, we have $g' \mid u - v$; in other words, $u \equiv v \mod g'$. Also, $g' \mid v$; in other words, $v \equiv 0 \mod g'$. Hence, $u \equiv v \equiv 0 \mod g'$, so that $u \equiv 0 \mod g'$. In other words, $g' \mid u$. Furthermore,

$$g' = (u - v)x' + vy' = ux' - vx' + vy' = ux' + v(y' - x').$$

Hence, there exist $g \in \mathbb{N}$, $x \in \mathbb{Z}$ and $y \in \mathbb{Z}$ such that $g = ux + vy$ and $g \mid u$ and $g \mid v$ (namely, $g = g'$, $x = x'$ and $y = y' - x'$). Thus, Claim 2 is proven in Case 2.

We have now proven Claim 2 in each of the two Cases 1 and 2. Thus, Claim 2 always holds (since Cases 1 and 2 cover all possibilities).]

Now, we can prove Claim 1 as well:

[*Proof of Claim 1:* We are in one of the following two cases:

Case 1: We have $a \geq b$.

Case 2: We have $a < b$.

Let us first consider Case 1. In this case, we have $a \geq b$. Hence, Claim 2 (applied to $u = a$ and $v = b$) shows that there exist $g \in \mathbb{N}$, $x \in \mathbb{Z}$ and $y \in \mathbb{Z}$ such that $g = ax + by$ and $g \mid a$ and $g \mid b$. Thus, Claim 1 is proven in Case 1.

Let us next consider Case 2. In this case, we have $a < b$. Hence, $a \leq b$, so that $b \geq a$. Also, $b + a = a + b = m$. Hence, Claim 2 (applied to $u = b$ and $v = a$) shows that there exist $g \in \mathbb{N}$, $x \in \mathbb{Z}$ and $y \in \mathbb{Z}$ such that $g = bx + ay$ and $g \mid b$ and $g \mid a$. Consider these g, x and y, and denote them by g', x' and y'. Thus, g' is an element of $\mathbb{N}$, and x' and y' are elements of $\mathbb{Z}$ satisfying $g' = bx' + ay'$ and $g' \mid b$ and $g' \mid a$. Now, $g' = bx' + ay' = ay' + bx'$. Hence, there exist $g \in \mathbb{N}$, $x \in \mathbb{Z}$ and $y \in \mathbb{Z}$ such that $g = ax + by$ and $g \mid a$ and $g \mid b$ (namely, $g = g'$, $x = y'$ and $y = x'$). Thus, Claim 1 is proven in Case 2.

We have now proven Claim 1 in each of the two Cases 1 and 2. Thus, Claim 1 always holds (since Cases 1 and 2 cover all possibilities).]

But $\mathcal{A}(m)$ is defined as the statement

$$\left(\begin{array}{c}\text{if } a \in \mathbb{N} \text{ and } b \in \mathbb{N} \text{ satisfy } a + b = m, \text{ then there exist } g \in \mathbb{N}, x \in \mathbb{Z} \\ \text{and } y \in \mathbb{Z} \text{ such that } g = ax + by \text{ and } g \mid a \text{ and } g \mid b\end{array}\right).$$

Thus, $\mathcal{A}(m)$ is precisely Claim 1. Hence, $\mathcal{A}(m)$ holds (since Claim 1 holds). This completes the induction step. Thus, we have proven by strong induction that $\mathcal{A}(n)$ holds for all $n \in \mathbb{N}$. In other words, the statement (108) holds for all $n \in \mathbb{N}$ (since this statement is precisely $\mathcal{A}(n)$).

Now, let $a \in \mathbb{N}$ and $b \in \mathbb{N}$. Then, $a + b \in \mathbb{N}$. Hence, we can apply (108) to $n = a + b$ (since $a + b = a + b$). We thus conclude that there exist $g \in \mathbb{N}$, $x \in \mathbb{Z}$ and $y \in \mathbb{Z}$ such that $g = ax + by$ and $g \mid a$ and $g \mid b$. This proves Theorem 2.70. □

2.10.2. Conventions for writing proofs by strong induction on derived quantities

Let us take a closer look at the proof we just gave. The statement $\mathcal{A}(n)$ that we defined was unsurprising: It simply says that Theorem 2.70 holds under the condition that $a+b=n$. Thus, by introducing $\mathcal{A}(n)$, we have "sliced" Theorem 2.70 into a sequence of statements $\mathcal{A}(0), \mathcal{A}(1), \mathcal{A}(2), \ldots$, which then allowed us to prove these statements by strong induction on n even though no "n" appeared in Theorem 2.70 itself. This strong induction can be simply called a "strong induction on $a+b$". More generally:

Convention 2.72. Let $\mathcal{B}$ be a logical statement that involves some variables $v_1, v_2, v_3, \ldots$. (For example, $\mathcal{B}$ can be the statement of Theorem 2.70; then, these variables are a and b.)

Let $g \in \mathbb{Z}$. (This g has nothing to do with the g from Theorem 2.70.)

Let q be some expression (involving the variables $v_1, v_2, v_3, \ldots$ or some of them) that has the property that whenever the variables $v_1, v_2, v_3, \ldots$ satisfy the assumptions of $\mathcal{B}$, the expression q evaluates to some element of $\mathbb{Z}_{\geq g}$. (For example, if $\mathcal{B}$ is the statement of Theorem 2.70 and $g=0$, then q can be the expression $a+b$, because $a+b \in \mathbb{N} = \mathbb{Z}_{\geq 0}$ whenever a and b are as in Theorem 2.70.)

Assume that you want to prove the statement $\mathcal{B}$. Then, you can proceed as follows: For each $n \in \mathbb{Z}_{\geq g}$, define $\mathcal{A}(n)$ to be the statement saying that[70]

$$(\text{the statement } \mathcal{B} \text{ holds under the condition that } q=n).$$

Then, prove $\mathcal{A}(n)$ by strong induction on n starting at g. Thus:

- The *induction step* consists in fixing $m \in \mathbb{Z}_{\geq g}$, and showing that if

  $$\left(\mathcal{A}(n) \text{ holds for every } n \in \mathbb{Z}_{\geq g} \text{ satisfying } n<m\right), \tag{111}$$

  then

  $$(\mathcal{A}(m) \text{ holds}). \tag{112}$$

  In other words, it consists in fixing $m \in \mathbb{Z}_{\geq g}$, and showing that if

  $$(\text{the statement } \mathcal{B} \text{ holds under the condition that } q<m), \tag{113}$$

  then

  $$(\text{the statement } \mathcal{B} \text{ holds under the condition that } q=m). \tag{114}$$

  (Indeed, the previous two sentences are equivalent, because of the logical equivalences

$$
\begin{aligned}
& \left(\mathcal{A}\left(n\right) \text{ holds for every } n \in \mathbb{Z}_{\geq g} \text{ satisfying } n < m\right) \\
\Longleftrightarrow\ & \left(\begin{array}{c} \left(\text{the statement } \mathcal{B} \text{ holds under the condition that } q = n\right) \\ \text{holds for every } n \in \mathbb{Z}_{\geq g} \text{ satisfying } n < m \end{array}\right) \\
& \qquad \left(\begin{array}{c} \text{since the statement } \mathcal{A}\left(n\right) \text{ is defined as} \\ \left(\text{the statement } \mathcal{B} \text{ holds under the condition that } q = n\right) \end{array}\right) \\
\Longleftrightarrow\ & \left(\text{the statement } \mathcal{B} \text{ holds under the condition that } q < m\right)
\end{aligned}
$$

and

$$
\begin{aligned}
& \left(\mathcal{A}\left(m\right) \text{ holds}\right) \\
\Longleftrightarrow\ & \left(\text{the statement } \mathcal{B} \text{ holds under the condition that } q = m\right) \\
& \qquad \left(\begin{array}{c} \text{since the statement } \mathcal{A}\left(m\right) \text{ is defined as} \\ \left(\text{the statement } \mathcal{B} \text{ holds under the condition that } q = m\right) \end{array}\right).
\end{aligned}
$$

)

In practice, this induction step will usually be organized as follows: We fix $m \in \mathbb{Z}_{\geq g}$, then we assume that the statement $\mathcal{B}$ holds under the condition that $q < m$ (this is the induction hypothesis), and then we prove that the statement $\mathcal{B}$ holds under the condition that $q = m$.

Once this induction proof is finished, it immediately follows that the statement $\mathcal{B}$ always holds (because the induction proof has shown that, whatever $n \in \mathbb{Z}_{\geq g}$ is, the statement $\mathcal{B}$ holds under the condition that $q = n$).

This strategy of proof is called "strong induction on q" (or "strong induction over q"). Once you have specified what q is, you don't need to explicitly define $\mathcal{A}\left(n\right)$, nor do you ever need to mention n.

Using this convention, we can rewrite our above proof of Theorem 2.70 as follows (remembering once again that $\mathbb{Z}_{\geq 0} = \mathbb{N}$):

Proof of Theorem 2.70 (second version). Let us prove Theorem 2.70 by strong induction on $a + b$ starting at 0:

Induction step: Let $m \in \mathbb{N}$. Assume that Theorem 2.70 holds under the condition that $a + b < m$. We must then show that Theorem 2.70 holds under the condition that $a + b = m$. This is tantamount to proving the following claim:

Claim 1: Let $a \in \mathbb{N}$ and $b \in \mathbb{N}$ satisfy $a + b = m$. Then, there exist $g \in \mathbb{N}$, $x \in \mathbb{Z}$ and $y \in \mathbb{Z}$ such that $g = ax + by$ and $g \mid a$ and $g \mid b$.

[70] We assume that no variable named "n" appears in the statement $\mathcal{B}$; otherwise, we need a different letter for our new variable in order to avoid confusion.

Before we prove Claim 1, let us show a slightly weaker version of it, in which we rename a and b as u and v and add the assumption that $u \geq v$:

> *Claim 2:* Let $u \in \mathbb{N}$ and $v \in \mathbb{N}$ satisfy $u + v = m$ and $u \geq v$. Then, there exist $g \in \mathbb{N}$, $x \in \mathbb{Z}$ and $y \in \mathbb{Z}$ such that $g = ux + vy$ and $g \mid u$ and $g \mid v$.

[*Proof of Claim 2:* We are in one of the following two cases:

Case 1: We have $v = 0$.

Case 2: We have $v \neq 0$.

Let us first consider Case 1. In this case, we have $v = 0$. Hence, $v = 0 = 0u$, so that $u \mid v$. Also, $u \cdot 1 + v \cdot 0 = u$. Thus, $u = u \cdot 1 + v \cdot 0$ and $u \mid u$ and $u \mid v$. Hence, there exist $g \in \mathbb{N}$, $x \in \mathbb{Z}$ and $y \in \mathbb{Z}$ such that $g = ux + vy$ and $g \mid u$ and $g \mid v$ (namely, $g = u$, $x = 1$ and $y = 0$). Thus, Claim 2 is proven in Case 1.

Let us now consider Case 2. In this case, we have $v \neq 0$. Hence, $v > 0$ (since $v \in \mathbb{N}$). Thus, $u + v > u + 0 = u$, so that $u < u + v = m$. Also, $u - v \in \mathbb{N}$ (since $u \geq v$) and $(u - v) + v = u$.

But we assumed that Theorem 2.70 holds under the condition that $a + b < m$. Thus, we can apply Theorem 2.70 to $a = u - v$ and $b = v$ (since $u - v \in \mathbb{N}$ and $(u - v) + v = u < m$). We thus conclude that there exist $g \in \mathbb{N}$, $x \in \mathbb{Z}$ and $y \in \mathbb{Z}$ such that $g = (u - v) x + vy$ and $g \mid u - v$ and $g \mid v$. Consider these g, x and y, and denote them by g', x' and y'. Thus, g' is an element of $\mathbb{N}$, and x' and y' are elements of $\mathbb{Z}$ satisfying $g' = (u - v) x' + vy'$ and $g' \mid u - v$ and $g' \mid v$.

Now, we have $g' \mid u - v$; in other words, $u \equiv v \mod g'$. Also, $g' \mid v$; in other words, $v \equiv 0 \mod g'$. Hence, $u \equiv v \equiv 0 \mod g'$, so that $u \equiv 0 \mod g'$. In other words, $g' \mid u$. Furthermore,

$$g' = (u - v) x' + vy' = ux' - vx' + vy' = ux' + v\left(y' - x'\right).$$

Hence, there exist $g \in \mathbb{N}$, $x \in \mathbb{Z}$ and $y \in \mathbb{Z}$ such that $g = ux + vy$ and $g \mid u$ and $g \mid v$ (namely, $g = g'$, $x = x'$ and $y = y' - x'$). Thus, Claim 2 is proven in Case 2.

We have now proven Claim 2 in each of the two Cases 1 and 2. Thus, Claim 2 always holds (since Cases 1 and 2 cover all possibilities).]

Now, we can prove Claim 1 as well:

[*Proof of Claim 1:* Claim 1 can be derived from Claim 2 in the same way as we derived it in the first version of the proof above. We shall not repeat this argument, since it just applies verbatim.]

But Claim 1 is simply saying that Theorem 2.70 holds under the condition that $a + b = m$. Thus, by proving Claim 1, we have shown that Theorem 2.70 holds under the condition that $a + b = m$. This completes the induction step. Thus, Theorem 2.70 is proven by strong induction. □

2.11. Induction in an interval

2.11.1. The induction principle for intervals

The induction principles we have seen so far were tailored towards proving statements whose variables range over infinite sets such as $\mathbb{N}$ and $\mathbb{Z}_{\geq g}$. Sometimes, one instead wants to do an induction on a variable that ranges over a finite interval, such as $\{g, g+1, \ldots, h\}$ for some integers g and h. We shall next state an induction principle tailored to such situations. First, we make an important convention:

Convention 2.73. If g and h are two integers such that $g > h$, then the set $\{g, g+1, \ldots, h\}$ is understood to be the empty set.

Thus, for example, $\{2, 3, \ldots, 1\} = \varnothing$ and $\{2, 3, \ldots, 0\} = \varnothing$ and $\{5, 6, \ldots, -100\} = \varnothing$. (But $\{5, 6, \ldots, 5\} = \{5\}$ and $\{5, 6, \ldots, 6\} = \{5, 6\}$.)

We now state our induction principle for intervals:

Theorem 2.74. Let $g \in \mathbb{Z}$ and $h \in \mathbb{Z}$. For each $n \in \{g, g+1, \ldots, h\}$, let $\mathcal{A}(n)$ be a logical statement.

Assume the following:

Assumption 1: If $g \leq h$, then the statement $\mathcal{A}(g)$ holds.

Assumption 2: If $m \in \{g, g+1, \ldots, h-1\}$ is such that $\mathcal{A}(m)$ holds, then $\mathcal{A}(m+1)$ also holds.

Then, $\mathcal{A}(n)$ holds for each $n \in \{g, g+1, \ldots, h\}$.

Theorem 2.74 is, in a sense, the closest one can get to Theorem 2.53 when having only finitely many statements $\mathcal{A}(g), \mathcal{A}(g+1), \ldots, \mathcal{A}(h)$ instead of an infinite sequence of statements $\mathcal{A}(g), \mathcal{A}(g+1), \mathcal{A}(g+2), \ldots$. It is easy to derive Theorem 2.74 from Corollary 2.61:

Proof of Theorem 2.74. For each $n \in \mathbb{Z}_{\geq g}$, we define $\mathcal{B}(n)$ to be the logical statement

$$(\text{if } n \in \{g, g+1, \ldots, h\}, \text{ then } \mathcal{A}(n) \text{ holds}).$$

Now, let us consider the Assumptions A and B from Corollary 2.61. We claim that both of these assumptions are satisfied.

Assumption 1 says that if $g \leq h$, then the statement $\mathcal{A}(g)$ holds. Thus, $\mathcal{B}(g)$ holds[71]. In other words, Assumption A is satisfied.

Next, we shall prove that Assumption B is satisfied. Indeed, let $p \in \mathbb{Z}_{\geq g}$ be such that $\mathcal{B}(p)$ holds. We shall now show that $\mathcal{B}(p+1)$ also holds.

Indeed, assume that $p+1 \in \{g, g+1, \ldots, h\}$. Thus, $p+1 \leq h$, so that $p \leq p+1 \leq h$. Combining this with $p \geq g$ (since $p \in \mathbb{Z}_{\geq g}$), we conclude that $p \in \{g, g+1, \ldots, h\}$ (since p is an integer). But we have assumed that $\mathcal{B}(p)$ holds. In other words,

$$\text{if } p \in \{g, g+1, \ldots, h\}, \text{ then } \mathcal{A}(p) \text{ holds}$$

[71] *Proof.* Assume that $g \in \{g, g+1, \ldots, h\}$. Thus, $g \leq h$. But Assumption 1 says that if $g \leq h$, then the statement $\mathcal{A}(g)$ holds. Hence, the statement $\mathcal{A}(g)$ holds (since $g \leq h$).

Now, forget that we assumed that $g \in \{g, g+1, \ldots, h\}$. We thus have proven that if $g \in \{g, g+1, \ldots, h\}$, then $\mathcal{A}(g)$ holds. In other words, $\mathcal{B}(g)$ holds (because the statement $\mathcal{B}(g)$ is defined as (if $g \in \{g, g+1, \ldots, h\}$, then $\mathcal{A}(g)$ holds)). Qed.

(because the statement $\mathcal{B}(p)$ is defined as (if $p \in \{g, g+1, \ldots, h\}$, then $\mathcal{A}(p)$ holds)). Thus, $\mathcal{A}(p)$ holds (since we have $p \in \{g, g+1, \ldots, h\}$). Also, from $p+1 \leq h$, we obtain $p \leq h-1$. Combining this with $p \geq g$, we find $p \in \{g, g+1, \ldots, h-1\}$. Thus, we know that $p \in \{g, g+1, \ldots, h-1\}$ is such that $\mathcal{A}(p)$ holds. Hence, Assumption 2 (applied to $m = p$) shows that $\mathcal{A}(p+1)$ also holds.

Now, forget that we assumed that $p+1 \in \{g, g+1, \ldots, h\}$. We thus have proven that if $p+1 \in \{g, g+1, \ldots, h\}$, then $\mathcal{A}(p+1)$ holds. In other words, $\mathcal{B}(p+1)$ holds (since the statement $\mathcal{B}(p+1)$ is defined as (if $p+1 \in \{g, g+1, \ldots, h\}$, then $\mathcal{A}(p+1)$ holds)).

Now, forget that we fixed p. We thus have proven that if $p \in \mathbb{Z}_{\geq g}$ is such that $\mathcal{B}(p)$ holds, then $\mathcal{B}(p+1)$ also holds. In other words, Assumption B is satisfied.

We now know that both Assumption A and Assumption B are satisfied. Hence, Corollary 2.61 shows that

$$\mathcal{B}(n) \text{ holds for each } n \in \mathbb{Z}_{\geq g}. \tag{115}$$

Now, let $n \in \{g, g+1, \ldots, h\}$. Thus, $n \geq g$, so that $n \in \mathbb{Z}_{\geq g}$. Hence, (115) shows that $\mathcal{B}(n)$ holds. In other words,

$$\text{if } n \in \{g, g+1, \ldots, h\}, \text{ then } \mathcal{A}(n) \text{ holds}$$

(since the statement $\mathcal{B}(n)$ was defined as (if $n \in \{g, g+1, \ldots, h\}$, then $\mathcal{A}(n)$ holds)). Thus, $\mathcal{A}(n)$ holds (since we have $n \in \{g, g+1, \ldots, h\}$).

Now, forget that we fixed n. We thus have shown that $\mathcal{A}(n)$ holds for each $n \in \{g, g+1, \ldots, h\}$. This proves Theorem 2.74. □

Theorem 2.74 is called the *principle of induction starting at g and ending at h*, and proofs that use it are usually called *proofs by induction* or *induction proofs*. As with all the other induction principles seen so far, we don't usually explicitly cite Theorem 2.74, but instead say certain words that signal that it is being applied and that (ideally) also indicate what integers g and h and what statements $\mathcal{A}(n)$ it is being applied to[72]. However, we shall reference it explicitly in our very first example of the use of Theorem 2.74:

Proposition 2.75. Let g and h be integers such that $g \leq h$. Let $b_g, b_{g+1}, \ldots, b_h$ be any $h-g+1$ nonzero integers. Assume that $b_g \geq 0$. Assume further that

$$|b_{i+1} - b_i| \leq 1 \qquad \text{for every } i \in \{g, g+1, \ldots, h-1\}. \tag{116}$$

Then, $b_n > 0$ for each $n \in \{g, g+1, \ldots, h\}$.

Proposition 2.75 is often called the *"discrete intermediate value theorem"* or the *"discrete continuity principle"*. Its intuitive meaning is that if a finite list of nonzero integers starts with a nonnegative integer, and every further entry of this list differs

[72] We will explain this in Convention 2.76 below.

from its preceding entry by at most 1, then all entries of this list must be positive. An example of such a list is $(2,3,3,2,3,4,4,3,2,3,2,3,2,1)$. Notice that Proposition 2.75 is, again, rather obvious from an intuitive perspective: It just says that it isn't possible to go from a nonnegative integer to a negative integer by steps of 1 without ever stepping at 0. The rigorous proof of Proposition 2.75 is not much harder – but because it is a statement about elements of $\{g,g+1,\ldots,h\}$, it naturally relies on Theorem 2.74:

Proof of Proposition 2.75. For each $n \in \{g,g+1,\ldots,h\}$, we let $\mathcal{A}(n)$ be the statement $(b_n > 0)$.

Our next goal is to prove the statement $\mathcal{A}(n)$ for each $n \in \{g,g+1,\ldots,h\}$.

All the $h-g+1$ integers $b_g, b_{g+1}, \ldots, b_h$ are nonzero (by assumption). Thus, in particular, b_g is nonzero. In other words, $b_g \neq 0$. Combining this with $b_g \geq 0$, we obtain $b_g > 0$. In other words, the statement $\mathcal{A}(g)$ holds (since this statement $\mathcal{A}(g)$ is defined to be $(b_g > 0)$). Hence,

$$\text{if } g \leq h \text{, then the statement } \mathcal{A}(g) \text{ holds.} \tag{117}$$

Now, we claim that

$$\text{if } m \in \{g,g+1,\ldots,h-1\} \text{ is such that } \mathcal{A}(m) \text{ holds, then } \mathcal{A}(m+1) \text{ also holds.} \tag{118}$$

[*Proof of (118):* Let $m \in \{g,g+1,\ldots,h-1\}$ be such that $\mathcal{A}(m)$ holds. We must show that $\mathcal{A}(m+1)$ also holds.

We have assumed that $\mathcal{A}(m)$ holds. In other words, $b_m > 0$ holds (since $\mathcal{A}(m)$ is defined to be the statement $(b_m > 0)$). Now, (116) (applied to $i = m$) yields $|b_{m+1} - b_m| \leq 1$. But it is well-known (and easy to see) that every integer x satisfies $-x \leq |x|$. Applying this to $x = b_{m+1} - b_m$, we obtain $-(b_{m+1} - b_m) \leq |b_{m+1} - b_m| \leq 1$. In other words, $1 \geq -(b_{m+1} - b_m) = b_m - b_{m+1}$. In other words, $1 + b_{m+1} \geq b_m$. Hence, $1 + b_{m+1} \geq b_m > 0$, so that $1 + b_{m+1} \geq 1$ (since $1 + b_{m+1}$ is an integer). In other words, $b_{m+1} \geq 0$.

But all the $h-g+1$ integers $b_g, b_{g+1}, \ldots, b_h$ are nonzero (by assumption). Thus, in particular, b_{m+1} is nonzero. In other words, $b_{m+1} \neq 0$. Combining this with $b_{m+1} \geq 0$, we obtain $b_{m+1} > 0$. But this is precisely the statement $\mathcal{A}(m+1)$ (because $\mathcal{A}(m+1)$ is defined to be the statement $(b_{m+1} > 0)$). Thus, the statement $\mathcal{A}(m+1)$ holds.

Now, forget that we fixed m. We thus have shown that if $m \in \{g,g+1,\ldots,h-1\}$ is such that $\mathcal{A}(m)$ holds, then $\mathcal{A}(m+1)$ also holds. This proves (118).]

Now, both assumptions of Theorem 2.74 are satisfied (indeed, Assumption 1 holds because of (117), whereas Assumption 2 holds because of (118)). Thus, Theorem 2.74 shows that $\mathcal{A}(n)$ holds for each $n \in \{g,g+1,\ldots,h\}$. In other words, $b_n > 0$ holds for each $n \in \{g,g+1,\ldots,h\}$ (since $\mathcal{A}(n)$ is the statement $(b_n > 0)$). This proves Proposition 2.75. □

2.11.2. Conventions for writing induction proofs in intervals

Next, we shall introduce some standard language that is commonly used in proofs by induction starting at g and ending at h. This language closely imitates the one we use for proofs by standard induction:

Convention 2.76. Let $g \in \mathbb{Z}$ and $h \in \mathbb{Z}$. For each $n \in \{g, g+1, \ldots, h\}$, let $\mathcal{A}(n)$ be a logical statement. Assume that you want to prove that $\mathcal{A}(n)$ holds for each $n \in \{g, g+1, \ldots, h\}$.

Theorem 2.74 offers the following strategy for proving this: First show that Assumption 1 of Theorem 2.74 is satisfied; then, show that Assumption 2 of Theorem 2.74 is satisfied; then, Theorem 2.74 automatically completes your proof.

A proof that follows this strategy is called a *proof by induction on n* (or *proof by induction over n*) *starting at g and ending at h* or (less precisely) an *inductive proof*. Most of the time, the words "starting at g and ending at h" are omitted, since they merely repeat what is clear from the context anyway: For example, if you make a claim about all integers $n \in \{3, 4, 5, 6\}$, and you say that you are proving it by induction on n, it is clear that you are using induction on n starting at 3 and ending at 6.

The proof that Assumption 1 is satisfied is called the *induction base* (or *base case*) of the proof. The proof that Assumption 2 is satisfied is called the *induction step* of the proof.

In order to prove that Assumption 2 is satisfied, you will usually want to fix an $m \in \{g, g+1, \ldots, h-1\}$ such that $\mathcal{A}(m)$ holds, and then prove that $\mathcal{A}(m+1)$ holds. In other words, you will usually want to fix $m \in \{g, g+1, \ldots, h-1\}$, assume that $\mathcal{A}(m)$ holds, and then prove that $\mathcal{A}(m+1)$ holds. When doing so, it is common to refer to the assumption that $\mathcal{A}(m)$ holds as the *induction hypothesis* (or *induction assumption*).

Unsurprisingly, this language parallels the language introduced in Convention 2.3 and in Convention 2.56.

Again, we can shorten our inductive proofs by omitting some sentences that convey no information. In particular, we can leave out the explicit definition of the statement $\mathcal{A}(n)$ when this statement is precisely the claim that we are proving (without the "for each $n \in \{g, g+1, \ldots, h\}$" part). Furthermore, it is common to leave the "If $g \leq h$" part of Assumption 1 unsaid (i.e., to pretend that Assumption 1 simply says that $\mathcal{A}(g)$ holds). Strictly speaking, this is somewhat imprecise, since $\mathcal{A}(g)$ is not defined when $g > h$; but of course, the whole claim that is being proven is moot anyway when $g > h$ (because there exist no $n \in \{g, g+1, \ldots, h\}$ in this case), so this imprecision doesn't matter.

Thus, we can rewrite our above proof of Proposition 2.75 as follows:

Proof of Proposition 2.75 (second version). We claim that

$$b_n > 0 \tag{119}$$

for each $n \in \{g, g+1, \ldots, h\}$.

Indeed, we shall prove (119) by induction on n:

Induction base: All the $h - g + 1$ integers $b_g, b_{g+1}, \ldots, b_h$ are nonzero (by assumption). Thus, in particular, b_g is nonzero. In other words, $b_g \neq 0$. Combining this with $b_g \geq 0$, we obtain $b_g > 0$. In other words, (119) holds for $n = g$. This completes the induction base.

Induction step: Let $m \in \{g, g+1, \ldots, h-1\}$. Assume that (119) holds for $n = m$. We must show that (119) also holds for $n = m + 1$.

We have assumed that (119) holds for $n = m$. In other words, $b_m > 0$. Now, (116) (applied to $i = m$) yields $|b_{m+1} - b_m| \leq 1$. But it is well-known (and easy to see) that every integer x satisfies $-x \leq |x|$. Applying this to $x = b_{m+1} - b_m$, we obtain $-(b_{m+1} - b_m) \leq |b_{m+1} - b_m| \leq 1$. In other words, $1 \geq -(b_{m+1} - b_m) = b_m - b_{m+1}$. In other words, $1 + b_{m+1} \geq b_m$. Hence, $1 + b_{m+1} \geq b_m > 0$, so that $1 + b_{m+1} \geq 1$ (since $1 + b_{m+1}$ is an integer). In other words, $b_{m+1} \geq 0$.

But all the $h - g + 1$ integers $b_g, b_{g+1}, \ldots, b_h$ are nonzero (by assumption). Thus, in particular, b_{m+1} is nonzero. In other words, $b_{m+1} \neq 0$. Combining this with $b_{m+1} \geq 0$, we obtain $b_{m+1} > 0$. In other words, (119) holds for $n = m + 1$. This completes the induction step. Thus, (119) is proven by induction. This proves Proposition 2.75. □

2.12. Strong induction in an interval

2.12.1. The strong induction principle for intervals

We shall next state yet another induction principle – one that combines the idea of strong induction (as in Theorem 2.60) with the idea of working inside an interval $\{g, g+1, \ldots, h\}$ (as in Theorem 2.74):

> **Theorem 2.77.** Let $g \in \mathbb{Z}$ and $h \in \mathbb{Z}$. For each $n \in \{g, g+1, \ldots, h\}$, let $\mathcal{A}(n)$ be a logical statement.
>
> Assume the following:
>
> *Assumption 1:* If $m \in \{g, g+1, \ldots, h\}$ is such that
>
> $$(\mathcal{A}(n) \text{ holds for every } n \in \{g, g+1, \ldots, h\} \text{ satisfying } n < m),$$
>
> then $\mathcal{A}(m)$ holds.
>
> Then, $\mathcal{A}(n)$ holds for each $n \in \{g, g+1, \ldots, h\}$.

Our proof of Theorem 2.77 will be similar to the proof of Theorem 2.74, except that we shall be using Theorem 2.60 instead of Corollary 2.61. Or, to be more precise, we shall be using the following restatement of Theorem 2.60:

Corollary 2.78. Let $g \in \mathbb{Z}$. For each $n \in \mathbb{Z}_{\geq g}$, let $\mathcal{B}(n)$ be a logical statement.

Assume the following:

Assumption A: If $p \in \mathbb{Z}_{\geq g}$ is such that

$$\left(\mathcal{B}(n) \text{ holds for every } n \in \mathbb{Z}_{\geq g} \text{ satisfying } n < p\right),$$

then $\mathcal{B}(p)$ holds.

Then, $\mathcal{B}(n)$ holds for each $n \in \mathbb{Z}_{\geq g}$.

Proof of Corollary 2.78. Corollary 2.78 is exactly Theorem 2.60, except that some names have been changed:

- The statements $\mathcal{A}(n)$ have been renamed as $\mathcal{B}(n)$.
- Assumption 1 has been renamed as Assumption A.
- The variable m in Assumption A has been renamed as p.

Thus, Corollary 2.78 holds (since Theorem 2.60 holds). □

We can now prove Theorem 2.77:

Proof of Theorem 2.77. For each $n \in \mathbb{Z}_{\geq g}$, we define $\mathcal{B}(n)$ to be the logical statement

$$(\text{if } n \in \{g, g+1, \ldots, h\}, \text{ then } \mathcal{A}(n) \text{ holds}).$$

Now, let us consider the Assumption A from Corollary 2.78. We claim that this assumption is satisfied.

Indeed, let $p \in \mathbb{Z}_{\geq g}$ be such that

$$\left(\mathcal{B}(n) \text{ holds for every } n \in \mathbb{Z}_{\geq g} \text{ satisfying } n < p\right). \tag{120}$$

We shall now show that $\mathcal{B}(p)$ holds.

Indeed, assume that $p \in \{g, g+1, \ldots, h\}$. Thus, $p \leq h$.

Now, let $n \in \{g, g+1, \ldots, h\}$ be such that $n < p$. Then, $n \in \{g, g+1, \ldots, h\} \subseteq \{g, g+1, g+2, \ldots\} = \mathbb{Z}_{\geq g}$ and $n < p$. Hence, (120) shows that $\mathcal{B}(n)$ holds. In other words, $(\text{if } n \in \{g, g+1, \ldots, h\}, \text{ then } \mathcal{A}(n) \text{ holds})$ (because the statement $\mathcal{B}(n)$ is defined as $(\text{if } n \in \{g, g+1, \ldots, h\}, \text{ then } \mathcal{A}(n) \text{ holds})$). Therefore, $\mathcal{A}(n)$ holds (since we know that $n \in \{g, g+1, \ldots, h\}$).

Now, forget that we fixed n. We thus have proven that

$$(\mathcal{A}(n) \text{ holds for every } n \in \{g, g+1, \ldots, h\} \text{ satisfying } n < p).$$

Hence, Assumption 1 (applied to $m = p$) yields that $\mathcal{A}(p)$ holds.

Now, forget that we assumed that $p \in \{g, g+1, \ldots, h\}$. We thus have proven that

$$(\text{if } p \in \{g, g+1, \ldots, h\}, \text{ then } \mathcal{A}(p) \text{ holds}).$$

In other words, $\mathcal{B}(p)$ holds (since the statement $\mathcal{B}(p)$ was defined as $(\text{if } p \in \{g, g+1, \ldots, h\}, \text{ then } \mathcal{A}(p) \text{ holds})$).

Now, forget that we fixed p. We thus have shown that if $p \in \mathbb{Z}_{\geq g}$ is such that

$$\left(\mathcal{B}(n) \text{ holds for every } n \in \mathbb{Z}_{\geq g} \text{ satisfying } n < p\right),$$

then $\mathcal{B}(p)$ holds. In other words, Assumption A is satisfied.

Hence, Corollary 2.78 shows that

$$\mathcal{B}(n) \text{ holds for each } n \in \mathbb{Z}_{\geq g}. \tag{121}$$

Now, let $n \in \{g, g+1, \ldots, h\}$. Thus, $n \geq g$, so that $n \in \mathbb{Z}_{\geq g}$. Hence, (121) shows that $\mathcal{B}(n)$ holds. In other words,

$$\text{if } n \in \{g, g+1, \ldots, h\}, \text{ then } \mathcal{A}(n) \text{ holds}$$

(since the statement $\mathcal{B}(n)$ was defined as $(\text{if } n \in \{g, g+1, \ldots, h\}, \text{ then } \mathcal{A}(n) \text{ holds})$). Thus, $\mathcal{A}(n)$ holds (since we have $n \in \{g, g+1, \ldots, h\}$).

Now, forget that we fixed n. We thus have shown that $\mathcal{A}(n)$ holds for each $n \in \{g, g+1, \ldots, h\}$. This proves Theorem 2.77. □

Theorem 2.77 is called the *principle of strong induction starting at g and ending at h*, and proofs that use it are usually called *proofs by strong induction*. Once again, we usually don't explicitly cite Theorem 2.77 in such proofs, and we usually don't say explicitly what g and h are and what the statements $\mathcal{A}(n)$ are when it is clear from the context. But (as with all the other induction principles considered so far) we shall be explicit about all these details in our first example:

Proposition 2.79. Let g and h be integers such that $g \leq h$. Let $b_g, b_{g+1}, \ldots, b_h$ be any $h - g + 1$ nonzero integers. Assume that $b_g \geq 0$. Assume that for each $p \in \{g+1, g+2, \ldots, h\}$,

$$\text{there exists some } j \in \{g, g+1, \ldots, p-1\} \text{ such that } b_p \geq b_j - 1. \tag{122}$$

(Of course, the j can depend on p.) Then, $b_n > 0$ for each $n \in \{g, g+1, \ldots, h\}$.

Proposition 2.79 is a more general (although less intuitive) version of Proposition 2.75; indeed, it is easy to see that the condition (116) is stronger than the condition (122) (when required for all $p \in \{g+1, g+2, \ldots, h\}$).

Example 2.80. For this example, set $g = 3$ and $h = 7$. Then, if we set $(b_3, b_4, b_5, b_6, b_7) = (4, 5, 3, 4, 2)$, then the condition (122) holds for all $p \in \{g+1, g+2, \ldots, h\}$. (For example, it holds for $p = 5$, since $b_5 = 3 \geq 4 - 1 = b_1 - 1$ and $1 \in \{g, g+1, \ldots, 5-1\}$.) On the other hand, if we set $(b_3, b_4, b_5, b_6, b_7) = (4, 5, 2, 4, 3)$, then this condition does not hold (indeed, it fails for $p = 5$, since $b_5 = 2$ is neither $\geq 4 - 1$ nor $\geq 5 - 1$).

Let us now prove Proposition 2.79 using Theorem 2.77:

Proof of Proposition 2.79. For each $n \in \{g, g+1, \ldots, h\}$, we let $\mathcal{A}(n)$ be the statement $(b_n > 0)$.

Our next goal is to prove the statement $\mathcal{A}(n)$ for each $n \in \{g, g+1, \ldots, h\}$.

All the $h - g + 1$ integers $b_g, b_{g+1}, \ldots, b_h$ are nonzero (by assumption). Thus, in particular, b_g is nonzero. In other words, $b_g \neq 0$. Combining this with $b_g \geq 0$, we obtain $b_g > 0$. In other words, the statement $\mathcal{A}(g)$ holds (since this statement $\mathcal{A}(g)$ is defined to be $(b_g > 0)$).

Now, we make the following claim:

> *Claim 1:* If $m \in \{g, g+1, \ldots, h\}$ is such that
>
> $$(\mathcal{A}(n) \text{ holds for every } n \in \{g, g+1, \ldots, h\} \text{ satisfying } n < m),$$
>
> then $\mathcal{A}(m)$ holds.

[*Proof of Claim 1:* Let $m \in \{g, g+1, \ldots, h\}$ be such that

$$(\mathcal{A}(n) \text{ holds for every } n \in \{g, g+1, \ldots, h\} \text{ satisfying } n < m). \tag{123}$$

We must show that $\mathcal{A}(m)$ holds.

If $m = g$, then this follows from the fact that $\mathcal{A}(g)$ holds. Thus, for the rest of the proof of Claim 1, we WLOG assume that we don't have $m = g$. Hence, $m \neq g$. Combining this with $m \in \{g, g+1, \ldots, h\}$, we obtain $m \in \{g, g+1, \ldots, h\} \setminus \{g\} \subseteq \{g+1, g+2, \ldots, h\}$. Hence, (122) (applied to $p = m$) shows that there exists some $j \in \{g, g+1, \ldots, m-1\}$ such that $b_m \geq b_j - 1$. Consider this j. From $m \in \{g+1, g+2, \ldots, h\}$, we obtain $m \leq h$.

From $j \in \{g, g+1, \ldots, m-1\}$, we obtain $j \leq m - 1 < m$. Also, $j \in \{g, g+1, \ldots, m-1\} \subseteq \{g, g+1, \ldots, h\}$ (since $m - 1 \leq m \leq h$). Thus, (123) (applied to $n = j$) yields that $\mathcal{A}(j)$ holds. In other words, $b_j > 0$ holds (since $\mathcal{A}(j)$ is defined to be the statement $(b_j > 0)$). Thus, $b_j \geq 1$ (since b_j is an integer), so that $b_j - 1 \geq 0$. But recall that $b_m \geq b_j - 1 \geq 0$.

But all the $h - g + 1$ integers $b_g, b_{g+1}, \ldots, b_h$ are nonzero (by assumption). Thus, in particular, b_m is nonzero. In other words, $b_m \neq 0$. Combining this with $b_m \geq 0$, we obtain $b_m > 0$. But this is precisely the statement $\mathcal{A}(m)$ (because $\mathcal{A}(m)$ is defined to be the statement $(b_m > 0)$). Thus, the statement $\mathcal{A}(m)$ holds. This completes the proof of Claim 1.]

Claim 1 says that Assumption 1 of Theorem 2.77 is satisfied. Thus, Theorem 2.77 shows that $\mathcal{A}(n)$ holds for each $n \in \{g, g+1, \ldots, h\}$. In other words, $b_n > 0$ holds for each $n \in \{g, g+1, \ldots, h\}$ (since $\mathcal{A}(n)$ is the statement $(b_n > 0)$). This proves Proposition 2.79. □

2.12.2. Conventions for writing strong induction proofs in intervals

Next, we shall introduce some standard language that is commonly used in proofs by strong induction starting at g and ending at h. This language closely imitates the one we use for proofs by "usual" strong induction:

Convention 2.81. Let $g \in \mathbb{Z}$ and $h \in \mathbb{Z}$. For each $n \in \{g, g+1, \ldots, h\}$, let $\mathcal{A}(n)$ be a logical statement. Assume that you want to prove that $\mathcal{A}(n)$ holds for each $n \in \{g, g+1, \ldots, h\}$.

Theorem 2.77 offers the following strategy for proving this: Show that Assumption 1 of Theorem 2.77 is satisfied; then, Theorem 2.77 automatically completes your proof.

A proof that follows this strategy is called a *proof by strong induction on n starting at g and ending at h*. Most of the time, the words "starting at g and ending at h" are omitted. The proof that Assumption 1 is satisfied is called the *induction step* of the proof. This kind of proof has no "induction base".

In order to prove that Assumption 1 is satisfied, you will usually want to fix an $m \in \{g, g+1, \ldots, h\}$ such that

$$(\mathcal{A}(n) \text{ holds for every } n \in \{g, g+1, \ldots, h\} \text{ satisfying } n < m), \tag{124}$$

and then prove that $\mathcal{A}(m)$ holds. In other words, you will usually want to fix $m \in \{g, g+1, \ldots, h\}$, assume that (124) holds, and then prove that $\mathcal{A}(m)$ holds. When doing so, it is common to refer to the assumption that (124) holds as the *induction hypothesis* (or *induction assumption*).

Unsurprisingly, this language parallels the language introduced in Convention 2.63.

As before, proofs using strong induction can be shortened by leaving out some uninformative prose. In particular, the explicit definition of the statement $\mathcal{A}(n)$ can often be omitted when this statement is precisely the claim that we are proving (without the "for each $n \in \{g, g+1, \ldots, h\}$" part). The values of g and h can also be inferred from the statement of the claim, so they don't need to be specified explicitly. And once again, we don't need to write "*Induction step:*", since our strong induction has no induction base.

This leads to the following abridged version of our above proof of Proposition 2.79:

Proof of Proposition 2.79 (second version). We claim that

$$b_n > 0 \tag{125}$$

for each $n \in \{g, g+1, \ldots, h\}$.

Indeed, we shall prove (125) by strong induction on n:

Let $m \in \{g, g+1, \ldots, h\}$. Assume that (125) holds for every $n \in \{g, g+1, \ldots, h\}$ satisfying $n < m$. We must show that (125) also holds for $n = m$. In other words, we must show that $b_m > 0$.

All the $h-g+1$ integers $b_g, b_{g+1}, \ldots, b_h$ are nonzero (by assumption). Thus, in particular, b_g is nonzero. In other words, $b_g \neq 0$. Combining this with $b_g \geq 0$, we obtain $b_g > 0$.

We have assumed that (125) holds for every $n \in \{g, g+1, \ldots, h\}$ satisfying $n < m$. In other words, we have

$$b_n > 0 \text{ for every } n \in \{g, g+1, \ldots, h\} \text{ satisfying } n < m. \tag{126}$$

Recall that we must prove that $b_m > 0$. If $m = g$, then this follows from $b_g > 0$. Thus, for the rest of this induction step, we WLOG assume that we don't have $m = g$. Hence, $m \neq g$. Combining this with $m \in \{g, g+1, \ldots, h\}$, we obtain $m \in \{g, g+1, \ldots, h\} \setminus \{g\} \subseteq \{g+1, g+2, \ldots, h\}$. Hence, (122) (applied to $p = m$) shows that there exists some $j \in \{g, g+1, \ldots, m-1\}$ such that $b_m \geq b_j - 1$. Consider this j. From $m \in \{g+1, g+2, \ldots, h\}$, we obtain $m \leq h$.

From $j \in \{g, g+1, \ldots, m-1\}$, we obtain $j \leq m-1 < m$. Also, $j \in \{g, g+1, \ldots, m-1\} \subseteq \{g, g+1, \ldots, h\}$ (since $m-1 \leq m \leq h$). Thus, (126) (applied to $n = j$) yields that $b_j > 0$. Thus, $b_j \geq 1$ (since b_j is an integer), so that $b_j - 1 \geq 0$. But recall that $b_m \geq b_j - 1 \geq 0$.

But all the $h-g+1$ integers $b_g, b_{g+1}, \ldots, b_h$ are nonzero (by assumption). Thus, in particular, b_m is nonzero. In other words, $b_m \neq 0$. Combining this with $b_m \geq 0$, we obtain $b_m > 0$.

Thus, we have proven that $b_m > 0$. In other words, (125) holds for $n = m$. This completes the induction step. Thus, (125) is proven by strong induction. This proves Proposition 2.79. □

2.13. General associativity for composition of maps

2.13.1. Associativity of map composition

Recall that if $f : X \to Y$ and $g : Y \to Z$ are two maps, then the *composition* $g \circ f$ of the maps g and f is defined to be the map

$$X \to Z, \ x \mapsto g(f(x)).$$

Now, if we have four sets X, Y, Z and W and three maps $c : X \to Y$, $b : Y \to Z$ and $a : Z \to W$, then we can build two possible compositions that use all three of these maps: namely, the two compositions $(a \circ b) \circ c$ and $a \circ (b \circ c)$. It turns out that these two compositions are the same map:[73]

[73] Of course, when some of the four sets X, Y, Z and W are equal, then more compositions can be built: For example, if $Y = Z = W$, then we can also build the composition $(b \circ a) \circ c$ or the composition $((b \circ b) \circ a) \circ c$. But these compositions are not the same map as the two that we previously constructed.

Proposition 2.82. Let X, Y, Z and W be four sets. Let $c : X \to Y$, $b : Y \to Z$ and $a : Z \to W$ be three maps. Then,

$$(a \circ b) \circ c = a \circ (b \circ c) .$$

Proposition 2.82 is called the *associativity of map composition*, and is proven straightforwardly:

Proof of Proposition 2.82. Let $x \in X$. Then, the definition of $b \circ c$ yields $(b \circ c)(x) = b(c(x))$. But the definition of $(a \circ b) \circ c$ yields

$$((a \circ b) \circ c)(x) = (a \circ b)(c(x)) = a(b(c(x))) \qquad \text{(by the definition of } a \circ b\text{)} .$$

On the other hand, the definition of $a \circ (b \circ c)$ yields

$$(a \circ (b \circ c))(x) = a\left(\underbrace{(b \circ c)(x)}_{=b(c(x))}\right) = a(b(c(x))) .$$

Comparing these two equalities, we obtain $((a \circ b) \circ c)(x) = (a \circ (b \circ c))(x)$.

Now, forget that we fixed x. We thus have shown that

$$((a \circ b) \circ c)(x) = (a \circ (b \circ c))(x) \qquad \text{for each } x \in X.$$

In other words, $(a \circ b) \circ c = a \circ (b \circ c)$. This proves Proposition 2.82. □

2.13.2. Composing more than 3 maps: exploration

Proposition 2.82 can be restated as follows: If a, b and c are three maps such that the compositions $a \circ b$ and $b \circ c$ are well-defined, then $(a \circ b) \circ c = a \circ (b \circ c)$. This allows us to write "$a \circ b \circ c$" for each of the compositions $(a \circ b) \circ c$ and $a \circ (b \circ c)$ without having to disambiguate this expression by means of parentheses. It is natural to ask whether we can do the same thing for more than three maps. For example, let us consider four maps a, b, c and d for which the compositions $a \circ b$, $b \circ c$ and $c \circ d$ are well-defined:

Example 2.83. Let X, Y, Z, W and U be five sets. Let $d : X \to Y$, $c : Y \to Z$, $b : Z \to W$ and $a : W \to U$ be four maps. Then, there we can construct five compositions that use all four of these maps; these five compositions are

$$((a \circ b) \circ c) \circ d, \qquad (a \circ (b \circ c)) \circ d, \qquad (a \circ b) \circ (c \circ d) , \tag{127}$$

$$a \circ ((b \circ c) \circ d) , \qquad a \circ (b \circ (c \circ d)) . \tag{128}$$

It turns out that these five compositions are all the same map. Indeed, this follows by combining the following observations:

- We have $((a \circ b) \circ c) \circ d = (a \circ (b \circ c)) \circ d$ (since Proposition 2.82 yields $(a \circ b) \circ c = a \circ (b \circ c)$).
- We have $a \circ ((b \circ c) \circ d) = a \circ (b \circ (c \circ d))$ (since Proposition 2.82 yields $(b \circ c) \circ d = b \circ (c \circ d)$).
- We have $(a \circ (b \circ c)) \circ d = a \circ ((b \circ c) \circ d)$ (by Proposition 2.82, applied to W, U, $b \circ c$ and d instead of Z, W, b and c).
- We have $((a \circ b) \circ c) \circ d = (a \circ b) \circ (c \circ d)$ (by Proposition 2.82, applied to U, $a \circ b$, c and d instead of W, a, b and c).

Hence, all five compositions are equal. Thus, we can write "$a \circ b \circ c \circ d$" for each of these five compositions, again dropping the parentheses.

We shall refer to the five compositions listed in (127) and (128) as the "complete parenthesizations of $a \circ b \circ c \circ d$". Here, the word "parenthesization" means a way to put parentheses into the expression "$a \circ b \circ c \circ d$", whereas the word "complete" means that these parentheses unambiguously determine which two maps any given $\circ$ sign is composing. (For example, the parenthesization "$(a \circ b \circ c) \circ d$" is not complete, because the first $\circ$ sign in it could be either composing a with b or composing a with $b \circ c$. But the parenthesization "$((a \circ b) \circ c) \circ d$" is complete, because its first $\circ$ sign composes a and b, whereas its second $\circ$ sign composes $a \circ b$ with c, and finally its third $\circ$ sign composes $(a \circ b) \circ c$ with d.)

Thus, we have seen that all five complete parenthesizations of $a \circ b \circ c \circ d$ are the same map.

What happens if we compose more than four maps? Clearly, the more maps we have, the more complete parenthesizations can be constructed. We have good reasons to suspect that these parenthesizations will all be the same map (so we can again drop the parentheses); but if we try to prove it in the ad-hoc way we did in Example 2.83, then we have more and more work to do the more maps we are composing. Clearly, if we want to prove our suspicion for arbitrarily many maps, we need a more general approach.

2.13.3. Formalizing general associativity

So let us make a general statement; but first, let us formally define the notion of a "complete parenthesization":

Definition 2.84. Let n be a positive integer. Let $X_1, X_2, \ldots, X_{n+1}$ be $n+1$ sets. For each $i \in \{1, 2, \ldots, n\}$, let $f_i : X_i \to X_{i+1}$ be a map. Then, we want to define the notion of a *complete parenthesization* of $f_n \circ f_{n-1} \circ \cdots \circ f_1$. We define this notion by recursion on n as follows:

- For $n = 1$, there is only one complete parenthesization of $f_n \circ f_{n-1} \circ \cdots \circ f_1$, and this is simply the map $f_1 : X_1 \to X_2$.
- If $n > 1$, then the complete parenthesizations of $f_n \circ f_{n-1} \circ \cdots \circ f_1$ are all the maps of the form $\alpha \circ \beta$, where
  - k is some element of $\{1, 2, \ldots, n-1\}$;
  - α is a complete parenthesization of $f_n \circ f_{n-1} \circ \cdots \circ f_{k+1}$;
  - β is a complete parenthesization of $f_k \circ f_{k-1} \circ \cdots \circ f_1$.

Example 2.85. Let us see what this definition yields for small values of n:

- For $n = 1$, the only complete parenthesization of f_1 is f_1.
- For $n = 2$, the only complete parenthesization of $f_2 \circ f_1$ is the composition $f_2 \circ f_1$ (because here, the only possible values of k, α and β are 1, f_2 and f_1, respectively).
- For $n = 3$, the complete parenthesizations of $f_3 \circ f_2 \circ f_1$ are the two compositions $(f_3 \circ f_2) \circ f_1$ and $f_3 \circ (f_2 \circ f_1)$ (because here, the only possible values of k are 1 and 2, and each value of k uniquely determines α and β). Proposition 2.82 shows that they are equal (as maps).
- For $n = 4$, the complete parenthesizations of $f_4 \circ f_3 \circ f_2 \circ f_1$ are the five compositions

$$\begin{array}{lll} ((f_4 \circ f_3) \circ f_2) \circ f_1, & (f_4 \circ (f_3 \circ f_2)) \circ f_1, & (f_4 \circ f_3) \circ (f_2 \circ f_1), \\ f_4 \circ ((f_3 \circ f_2) \circ f_1), & f_4 \circ (f_3 \circ (f_2 \circ f_1)). & \end{array}$$

  (These are exactly the five compositions listed in (127) and (128), except that the maps d, c, b, a are now called f_1, f_2, f_3, f_4.) We have seen in Example 2.83 that these five compositions are equal as maps.
- For $n = 5$, the complete parenthesizations of $f_5 \circ f_4 \circ f_3 \circ f_2 \circ f_1$ are 14 compositions, one of which is $(f_5 \circ f_4) \circ (f_3 \circ (f_2 \circ f_1))$. Again, it is laborious but not difficult to check that all the 14 compositions are equal as maps.

Now, we want to prove the following general statement:

Theorem 2.86. Let n be a positive integer. Let $X_1, X_2, \ldots, X_{n+1}$ be $n + 1$ sets. For each $i \in \{1, 2, \ldots, n\}$, let $f_i : X_i \to X_{i+1}$ be a map. Then, all complete parenthesizations of $f_n \circ f_{n-1} \circ \cdots \circ f_1$ are the same map (from X_1 to X_{n+1}).

Theorem 2.86 is sometimes called the *general associativity* theorem, and is often proved in the context of monoids (see, e.g., [Artin10, Proposition 2.1.4]); while the context is somewhat different from ours, the proofs usually given still apply in ours.

2.13.4. Defining the "canonical" composition $C(f_n, f_{n-1}, \ldots, f_1)$

We shall prove Theorem 2.86 in a slightly indirect way: We first define a *specific* complete parenthesization of $f_n \circ f_{n-1} \circ \cdots \circ f_1$, which we shall call $C(f_n, f_{n-1}, \ldots, f_1)$; then we will show that it satisfies certain equalities (Proposition 2.89), and then prove that every complete parenthesization of $f_n \circ f_{n-1} \circ \cdots \circ f_1$ equals this map $C(f_n, f_{n-1}, \ldots, f_1)$ (Proposition 2.90). Each step of this strategy will rely on induction.

We begin with the definition of $C(f_n, f_{n-1}, \ldots, f_1)$:

Definition 2.87. Let n be a positive integer. Let $X_1, X_2, \ldots, X_{n+1}$ be $n+1$ sets. For each $i \in \{1, 2, \ldots, n\}$, let $f_i : X_i \to X_{i+1}$ be a map. Then, we want to define a map $C(f_n, f_{n-1}, \ldots, f_1) : X_1 \to X_{n+1}$. We define this map by recursion on n as follows:

- If $n = 1$, then we define $C(f_n, f_{n-1}, \ldots, f_1)$ to be the map $f_1 : X_1 \to X_2$. (Note that in this case, $C(f_n, f_{n-1}, \ldots, f_1) = C(f_1)$, because $(f_n, f_{n-1}, \ldots, f_1) = (f_1, f_{1-1}, \ldots, f_1) = (f_1)$.)
- If $n > 1$, then we define $C(f_n, f_{n-1}, \ldots, f_1) : X_1 \to X_{n+1}$ by

$$C(f_n, f_{n-1}, \ldots, f_1) = f_n \circ C(f_{n-1}, f_{n-2}, \ldots, f_1). \tag{129}$$

Example 2.88. Consider the situation of Definition 2.87.

(a) If $n = 1$, then

$$C(f_1) = f_1 \tag{130}$$

(by the $n = 1$ case of the definition).

(b) If $n = 2$, then

$$\begin{aligned} C(f_2, f_1) &= f_2 \circ \underbrace{C(f_1)}_{\substack{= f_1 \\ \text{(by (130))}}} \qquad \text{(by (129), applied to } n = 2\text{)} \\ &= f_2 \circ f_1. \end{aligned} \tag{131}$$

(c) If $n = 3$, then

$$\begin{aligned} C(f_3, f_2, f_1) &= f_3 \circ \underbrace{C(f_2, f_1)}_{\substack{= f_2 \circ f_1 \\ \text{(by (131))}}} \qquad \text{(by (129), applied to } n = 3\text{)} \\ &= f_3 \circ (f_2 \circ f_1). \end{aligned} \tag{132}$$

(d) If $n = 4$, then

$$\begin{aligned} C\left(f_4, f_3, f_2, f_1\right) &= f_4 \circ \underbrace{C\left(f_3, f_2, f_1\right)}_{\substack{=f_3\circ(f_2\circ f_1) \\ \text{(by (132))}}} \qquad \text{(by (129), applied to } n = 4\text{)} \\ &= f_4 \circ \left(f_3 \circ \left(f_2 \circ f_1\right)\right). \end{aligned} \tag{133}$$

(e) For an arbitrary $n \geq 1$, we can informally write $C\left(f_n, f_{n-1}, \ldots, f_1\right)$ as

$$C\left(f_n, f_{n-1}, \ldots, f_1\right) = f_n \circ \left(f_{n-1} \circ \left(f_{n-2} \circ \left(\cdots \circ \left(f_2 \circ f_1\right) \cdots\right)\right)\right).$$

The right hand side of this equality is a complete parenthesization of $f_n \circ f_{n-1} \circ \cdots \circ f_1$, where all the parentheses are "concentrated as far right as possible" (i.e., there is an opening parenthesis after each "$\circ$" sign except for the last one; and there are $n-2$ closing parentheses at the end of the expression). This is merely a visual restatement of the recursive definition of $C\left(f_n, f_{n-1}, \ldots, f_1\right)$ we gave above.

2.13.5. The crucial property of $C\left(f_n, f_{n-1}, \ldots, f_1\right)$

The following proposition will be key to our proof of Theorem 2.86:

Proposition 2.89. Let n be a positive integer. Let $X_1, X_2, \ldots, X_{n+1}$ be $n+1$ sets. For each $i \in \{1, 2, \ldots, n\}$, let $f_i : X_i \to X_{i+1}$ be a map. Then,

$$C\left(f_n, f_{n-1}, \ldots, f_1\right) = C\left(f_n, f_{n-1}, \ldots, f_{k+1}\right) \circ C\left(f_k, f_{k-1}, \ldots, f_1\right)$$

for each $k \in \{1, 2, \ldots, n-1\}$.

Proof of Proposition 2.89. Forget that we fixed n, $X_1, X_2, \ldots, X_{n+1}$ and the maps f_i. We shall prove Proposition 2.89 by induction on n: [74]

Induction base: If $n = 1$, then $\{1, 2, \ldots, n-1\} = \{1, 2, \ldots, 1-1\} = \varnothing$. Hence, if $n = 1$, then there exists no $k \in \{1, 2, \ldots, n-1\}$. Thus, if $n = 1$, then Proposition 2.89 is vacuously true (since Proposition 2.89 has a "for each $k \in \{1, 2, \ldots, n-1\}$" clause). This completes the induction base.

Induction step: Let $m \in \mathbb{Z}_{\geq 1}$. Assume that Proposition 2.89 holds under the condition that $n = m$. We must now prove that Proposition 2.89 holds under the condition that $n = m+1$. In other words, we must prove the following claim:

> *Claim 1:* Let $X_1, X_2, \ldots, X_{(m+1)+1}$ be $(m+1)+1$ sets. For each $i \in \{1, 2, \ldots, m+1\}$, let $f_i : X_i \to X_{i+1}$ be a map. Then,
>
> $$C\left(f_{m+1}, f_{(m+1)-1}, \ldots, f_1\right) = C\left(f_{m+1}, f_{(m+1)-1}, \ldots, f_{k+1}\right) \circ C\left(f_k, f_{k-1}, \ldots, f_1\right)$$

[74] The induction principle that we are applying here is Theorem 2.53 with $g = 1$ (since $\mathbb{Z}_{\geq 1}$ is the set of all positive integers).

for each $k \in \{1,2,\ldots,(m+1)-1\}$.

[*Proof of Claim 1:* Let $k \in \{1,2,\ldots,(m+1)-1\}$. Thus, $k \in \{1,2,\ldots,(m+1)-1\} = \{1,2,\ldots,m\}$ (since $(m+1)-1=m$).

We know that $X_1, X_2, \ldots, X_{(m+1)+1}$ are $(m+1)+1$ sets. In other words, $X_1, X_2, \ldots, X_{m+2}$ are $m+2$ sets (since $(m+1)+1 = m+2$). We have $m \in \mathbb{Z}_{\geq 1}$, thus $m \geq 1 > 0$; hence, $m+1 > 1$. Thus, (129) (applied to $n = m+1$) yields

$$\begin{aligned} C\left(f_{m+1}, f_{(m+1)-1}, \ldots, f_1\right) &= f_{m+1} \circ C\left(f_{(m+1)-1}, f_{(m+1)-2}, \ldots, f_1\right) \\ &= f_{m+1} \circ C\left(f_m, f_{m-1}, \ldots, f_1\right) \end{aligned} \tag{134}$$

(since $(m+1)-1 = m$ and $(m+1)-2 = m-1$).

But we are in one of the following two cases:

Case 1: We have $k = m$.

Case 2: We have $k \neq m$.

Let us first consider Case 1. In this case, we have $k = m$. Hence,

$$C\left(f_{m+1}, f_{(m+1)-1}, \ldots, f_{k+1}\right) = C\left(f_{m+1}, f_{(m+1)-1}, \ldots, f_{m+1}\right) = C\left(f_{m+1}\right) = f_{m+1}$$

(by (130), applied to X_{m+1}, X_{m+2} and f_{m+1} instead of X_1, X_2 and f_1), so that

$$\underbrace{C\left(f_{m+1}, f_{(m+1)-1}, \ldots, f_{k+1}\right)}_{=f_{m+1}} \circ \underbrace{C\left(f_k, f_{k-1}, \ldots, f_1\right)}_{\substack{=C(f_m, f_{m-1}, \ldots, f_1) \\ \text{(since } k=m\text{)}}} = f_{m+1} \circ C\left(f_m, f_{m-1}, \ldots, f_1\right).$$

Comparing this with (134), we obtain

$$C\left(f_{m+1}, f_{(m+1)-1}, \ldots, f_1\right) = C\left(f_{m+1}, f_{(m+1)-1}, \ldots, f_{k+1}\right) \circ C\left(f_k, f_{k-1}, \ldots, f_1\right).$$

Hence, Claim 1 is proven in Case 1.

Let us now consider Case 2. In this case, we have $k \neq m$. Combining $k \in \{1,2,\ldots,m\}$ with $k \neq m$, we obtain

$$k \in \{1,2,\ldots,m\} \setminus \{m\} = \{1,2,\ldots,m-1\}.$$

Hence, $k \leq m-1 < m$, so that $m+1-\underbrace{k}_{<m} > m+1-m = 1$.

But we assumed that Proposition 2.89 holds under the condition that $n = m$. Hence, we can apply Proposition 2.89 to m instead of n. We thus obtain

$$C\left(f_m, f_{m-1}, \ldots, f_1\right) = C\left(f_m, f_{m-1}, \ldots, f_{k+1}\right) \circ C\left(f_k, f_{k-1}, \ldots, f_1\right)$$

(since $k \in \{1,2,\ldots,m-1\}$). Now, (134) yields

$$\begin{aligned} & C\left(f_{m+1}, f_{(m+1)-1}, \ldots, f_1\right) \\ & = f_{m+1} \circ \underbrace{C\left(f_m, f_{m-1}, \ldots, f_1\right)}_{=C(f_m, f_{m-1}, \ldots, f_{k+1}) \circ C(f_k, f_{k-1}, \ldots, f_1)} \\ & = f_{m+1} \circ \left(C\left(f_m, f_{m-1}, \ldots, f_{k+1}\right) \circ C\left(f_k, f_{k-1}, \ldots, f_1\right)\right). \end{aligned} \tag{135}$$

On the other hand, $m+1-k>1$. Hence, (129) (applied to $m+1-k$, X_{k+i} and f_{k+i} instead of n, X_i and f_i) yields

$$\begin{aligned} & C\left(f_{k+(m+1-k)}, f_{k+((m+1-k)-1)}, \ldots, f_{k+1}\right) \\ & = f_{k+(m+1-k)} \circ C\left(f_{k+((m+1-k)-1)}, f_{k+((m+1-k)-2)}, \ldots, f_{k+1}\right) \\ & = f_{m+1} \circ C\left(f_m, f_{m-1}, \ldots, f_{k+1}\right) \\ & \qquad \left(\begin{array}{c} \text{since } k+(m+1-k)=m+1 \text{ and } k+((m+1-k)-1)=m \\ \text{and } k+((m+1-k)-2)=m-1 \end{array}\right). \end{aligned}$$

Since $k+(m+1-k)=m+1$ and $k+((m+1-k)-1)=(m+1)-1$, this rewrites as

$$C\left(f_{m+1}, f_{(m+1)-1}, \ldots, f_{k+1}\right) = f_{m+1} \circ C\left(f_m, f_{m-1}, \ldots, f_{k+1}\right).$$

Hence,

$$\begin{aligned} & \underbrace{C\left(f_{m+1}, f_{(m+1)-1}, \ldots, f_{k+1}\right)}_{=f_{m+1} \circ C(f_m, f_{m-1}, \ldots, f_{k+1})} \circ C\left(f_k, f_{k-1}, \ldots, f_1\right) \\ & = \left(f_{m+1} \circ C\left(f_m, f_{m-1}, \ldots, f_{k+1}\right)\right) \circ C\left(f_k, f_{k-1}, \ldots, f_1\right) \\ & = f_{m+1} \circ \left(C\left(f_m, f_{m-1}, \ldots, f_{k+1}\right) \circ C\left(f_k, f_{k-1}, \ldots, f_1\right)\right) \end{aligned}$$

(by Proposition 2.82, applied to $X = X_1$, $Y = X_{k+1}$, $Z = X_{m+1}$, $W = X_{m+2}$, $c = C\left(f_k, f_{k-1}, \ldots, f_1\right)$, $b = C\left(f_m, f_{m-1}, \ldots, f_{k+1}\right)$ and $a = f_{m+1}$). Comparing this with (135), we obtain

$$C\left(f_{m+1}, f_{(m+1)-1}, \ldots, f_1\right) = C\left(f_{m+1}, f_{(m+1)-1}, \ldots, f_{k+1}\right) \circ C\left(f_k, f_{k-1}, \ldots, f_1\right).$$

Hence, Claim 1 is proven in Case 2.

We have now proven Claim 1 in each of the two Cases 1 and 2. Since these two Cases cover all possibilities, we thus conclude that Claim 1 always holds.]

Now, we have proven Claim 1. In other words, we have proven that Proposition 2.89 holds under the condition that $n = m+1$. This completes the induction step. Hence, Proposition 2.89 is proven by induction. □

2.13.6. Proof of general associativity

Proposition 2.90. Let n be a positive integer. Let $X_1, X_2, \ldots, X_{n+1}$ be $n+1$ sets. For each $i \in \{1, 2, \ldots, n\}$, let $f_i : X_i \to X_{i+1}$ be a map. Then, every complete parenthesization of $f_n \circ f_{n-1} \circ \cdots \circ f_1$ equals $C\left(f_n, f_{n-1}, \ldots, f_1\right)$.

Proof of Proposition 2.90. Forget that we fixed n, $X_1, X_2, \ldots, X_{n+1}$ and the maps f_i. We shall prove Proposition 2.90 by strong induction on n: [75]

[75] The induction principle that we are applying here is Theorem 2.60 with $g = 1$ (since $\mathbb{Z}_{\geq 1}$ is the set of all positive integers).

Induction step: Let $m \in \mathbb{Z}_{\geq 1}$. Assume that Proposition 2.90 holds under the condition that $n < m$. We must prove that Proposition 2.90 holds under the condition that $n = m$. In other words, we must prove the following claim:

> *Claim 1:* Let $X_1, X_2, \ldots, X_{m+1}$ be $m + 1$ sets. For each $i \in \{1, 2, \ldots, m\}$, let $f_i : X_i \to X_{i+1}$ be a map. Then, every complete parenthesization of $f_m \circ f_{m-1} \circ \cdots \circ f_1$ equals $C(f_m, f_{m-1}, \ldots, f_1)$.

[*Proof of Claim 1:* Let γ be a complete parenthesization of $f_m \circ f_{m-1} \circ \cdots \circ f_1$. Thus, we must prove that $\gamma = C(f_m, f_{m-1}, \ldots, f_1)$.

We have $m \in \mathbb{Z}_{\geq 1}$, thus $m \geq 1$. Hence, either $m = 1$ or $m > 1$. Thus, we are in one of the following two cases:

Case 1: We have $m = 1$.

Case 2: We have $m > 1$.

Let us first consider Case 1. In this case, we have $m = 1$. Thus, we have $C(f_m, f_{m-1}, \ldots, f_1) = f_1$ (by the definition of $C(f_m, f_{m-1}, \ldots, f_1)$).

Recall that $m = 1$. Thus, the definition of a "complete parenthesization of $f_m \circ f_{m-1} \circ \cdots \circ f_1$" shows that there is only one complete parenthesization of $f_m \circ f_{m-1} \circ \cdots \circ f_1$, and this is simply the map $f_1 : X_1 \to X_2$. Hence, γ is simply the map $f_1 : X_1 \to X_2$ (since γ is a complete parenthesization of $f_m \circ f_{m-1} \circ \cdots \circ f_1$). Thus, $\gamma = f_1 = C(f_m, f_{m-1}, \ldots, f_1)$ (since $C(f_m, f_{m-1}, \ldots, f_1) = f_1$). Thus, $\gamma = C(f_m, f_{m-1}, \ldots, f_1)$ is proven in Case 1.

Now, let us consider Case 2. In this case, we have $m > 1$. Hence, the definition of a "complete parenthesization of $f_m \circ f_{m-1} \circ \cdots \circ f_1$" shows that any complete parenthesization of $f_m \circ f_{m-1} \circ \cdots \circ f_1$ is a map of the form $\alpha \circ \beta$, where

- k is some element of $\{1, 2, \ldots, m-1\}$;
- α is a complete parenthesization of $f_m \circ f_{m-1} \circ \cdots \circ f_{k+1}$;
- β is a complete parenthesization of $f_k \circ f_{k-1} \circ \cdots \circ f_1$.

Thus, γ is a map of this form (since γ is a complete parenthesization of $f_m \circ f_{m-1} \circ \cdots \circ f_1$). In other words, we can write γ in the form $\gamma = \alpha \circ \beta$, where k is some element of $\{1, 2, \ldots, m-1\}$, where α is a complete parenthesization of $f_m \circ f_{m-1} \circ \cdots \circ f_{k+1}$, and where β is a complete parenthesization of $f_k \circ f_{k-1} \circ \cdots \circ f_1$. Consider these k, α and β.

We have $k \in \{1, 2, \ldots, m-1\}$, thus $k \leq m - 1 < m$. Hence, we can apply Proposition 2.90 to $n = k$ (since we assumed that Proposition 2.90 holds under the condition that $n < m$). We thus conclude that every complete parenthesization of $f_k \circ f_{k-1} \circ \cdots \circ f_1$ equals $C(f_k, f_{k-1}, \ldots, f_1)$. Hence, β equals $C(f_k, f_{k-1}, \ldots, f_1)$ (since β is a complete parenthesization of $f_k \circ f_{k-1} \circ \cdots \circ f_1$). In other words,

$$\beta = C(f_k, f_{k-1}, \ldots, f_1). \tag{136}$$

We have $k \in \{1, 2, \ldots, m-1\}$, thus $k \geq 1$ and therefore $m - \underbrace{k}_{\geq 1} \leq m - 1 < m$. Hence, we can apply Proposition 2.90 to $m-k$, X_{k+i} and f_{k+i} instead of n, X_i and f_i (since we assumed that Proposition 2.90 holds under the condition that $n < m$). We thus conclude that every complete parenthesization of $f_{k+(m-k)} \circ f_{k+(m-k-1)} \circ \cdots \circ f_{k+1}$ equals $C\left(f_{k+(m-k)}, f_{k+(m-k-1)}, \ldots, f_{k+1}\right)$.

Since $k + (m-k) = m$ and $k + (m-k-1) = m-1$, this rewrites as follows: Every complete parenthesization of $f_m \circ f_{m-1} \circ \cdots \circ f_{k+1}$ equals $C(f_m, f_{m-1}, \ldots, f_{k+1})$. Thus, α equals $C(f_m, f_{m-1}, \ldots, f_{k+1})$ (since α is a complete parenthesization of $f_m \circ f_{m-1} \circ \cdots \circ f_{k+1}$). In other words,

$$\alpha = C(f_m, f_{m-1}, \ldots, f_{k+1}). \qquad (137)$$

But Proposition 2.89 (applied to $n = m$) yields

$$C(f_m, f_{m-1}, \ldots, f_1) = \underbrace{C(f_m, f_{m-1}, \ldots, f_{k+1})}_{\substack{=\alpha \\ \text{(by (137))}}} \circ \underbrace{C(f_k, f_{k-1}, \ldots, f_1)}_{\substack{=\beta \\ \text{(by (136))}}} = \alpha \circ \beta = \gamma$$

(since $\gamma = \alpha \circ \beta$), so that $\gamma = C(f_m, f_{m-1}, \ldots, f_1)$. Hence, $\gamma = C(f_m, f_{m-1}, \ldots, f_1)$ is proven in Case 2.

We now have shown that $\gamma = C(f_m, f_{m-1}, \ldots, f_1)$ in each of the two Cases 1 and 2. Since these two Cases cover all possibilities, this yields that $\gamma = C(f_m, f_{m-1}, \ldots, f_1)$ always holds.

Now, forget that we fixed γ. We thus have shown that $\gamma = C(f_m, f_{m-1}, \ldots, f_1)$ whenever γ is a complete parenthesization of $f_m \circ f_{m-1} \circ \cdots \circ f_1$. In other words, every complete parenthesization of $f_m \circ f_{m-1} \circ \cdots \circ f_1$ equals $C(f_m, f_{m-1}, \ldots, f_1)$. This proves Claim 1.]

Now, we have proven Claim 1. In other words, we have proven that Proposition 2.90 holds under the condition that $n = m$. This completes the induction step. Hence, Proposition 2.90 is proven by strong induction. □

Proof of Theorem 2.86. Proposition 2.90 shows that every complete parenthesization of $f_n \circ f_{n-1} \circ \cdots \circ f_1$ equals $C(f_n, f_{n-1}, \ldots, f_1)$. Thus, all complete parenthesizations of $f_n \circ f_{n-1} \circ \cdots \circ f_1$ are the same map. This proves Theorem 2.86. □

2.13.7. Compositions of multiple maps without parentheses

Definition 2.91. Let n be a positive integer. Let $X_1, X_2, \ldots, X_{n+1}$ be $n+1$ sets. For each $i \in \{1, 2, \ldots, n\}$, let $f_i : X_i \to X_{i+1}$ be a map. Then, the map $C(f_n, f_{n-1}, \ldots, f_1) : X_1 \to X_{n+1}$ is denoted by $f_n \circ f_{n-1} \circ \cdots \circ f_1$ and called the *composition* of $f_n, f_{n-1}, \ldots, f_1$. This notation $f_n \circ f_{n-1} \circ \cdots \circ f_1$ may conflict with existing notations in two cases:

- In the case when $n = 1$, this notation $f_n \circ f_{n-1} \circ \cdots \circ f_1$ simply becomes f_1, which looks exactly like the map f_1 itself. Fortunately, this conflict of notation is harmless, because the new meaning that we are giving to f_1 in this case (namely, $C(f_n, f_{n-1}, \ldots, f_1) = C(f_1)$) agrees with the map f_1 (because of (130)).

- In the case when $n = 2$, this notation $f_n \circ f_{n-1} \circ \cdots \circ f_1$ simply becomes $f_2 \circ f_1$, which looks exactly like the composition $f_2 \circ f_1$ of the two maps f_2 and f_1. Fortunately, this conflict of notation is harmless, because the new meaning that we are giving to $f_2 \circ f_1$ in this case (namely, $C(f_n, f_{n-1}, \ldots, f_1) = C(f_2, f_1)$) agrees with the latter composition (because of (131)).

Thus, in both cases, the conflict with existing notations is harmless (the conflicting notations actually stand for the same thing).

Remark 2.92. Let n, $X_1, X_2, \ldots, X_{n+1}$ and f_i be as in Definition 2.91. Then, Proposition 2.90 shows that every complete parenthesization of $f_n \circ f_{n-1} \circ \cdots \circ f_1$ equals $C(f_n, f_{n-1}, \ldots, f_1)$. In other words, every complete parenthesization of $f_n \circ f_{n-1} \circ \cdots \circ f_1$ equals $f_n \circ f_{n-1} \circ \cdots \circ f_1$ (because $f_n \circ f_{n-1} \circ \cdots \circ f_1$ was defined to be $C(f_n, f_{n-1}, \ldots, f_1)$ in Definition 2.91). In other words, **we can drop the parentheses** in every complete parenthesization of $f_n \circ f_{n-1} \circ \cdots \circ f_1$. For example, for $n = 7$, we get

$$(f_7 \circ (f_6 \circ f_5)) \circ (f_4 \circ ((f_3 \circ f_2) \circ f_1)) = f_7 \circ f_6 \circ f_5 \circ f_4 \circ f_3 \circ f_2 \circ f_1,$$

and a similar equality for any other complete parenthesization.

Definition 2.91 and Remark 2.92 finally give us the justification to write compositions of multiple maps (like $f_n \circ f_{n-1} \circ \cdots \circ f_1$ for $n \geq 1$) without the need for parentheses. We shall now go one little step further and extend this notation to the case of $n = 0$ – that is, we shall define the composition of **no** maps:

Definition 2.93. Let $n \in \mathbb{N}$. Let $X_1, X_2, \ldots, X_{n+1}$ be $n+1$ sets. For each $i \in \{1, 2, \ldots, n\}$, let $f_i : X_i \to X_{i+1}$ be a map. In Definition 2.91, we have defined the composition $f_n \circ f_{n-1} \circ \cdots \circ f_1$ of $f_n, f_{n-1}, \ldots, f_1$ when n is a positive integer. We shall now extend this definition to the case when $n = 0$ (so that it will be defined for all $n \in \mathbb{N}$, not just for all positive integers n). Namely, we extend it by setting

$$f_n \circ f_{n-1} \circ \cdots \circ f_1 = \mathrm{id}_{X_1} \qquad \text{when } n = 0. \tag{138}$$

That is, we say that the composition of 0 maps is the identity map $\mathrm{id}_{X_1} : X_1 \to X_1$. This composition of 0 maps is also known as the *empty composition of maps*. Thus, the empty composition of maps is defined to be id_{X_1}. (This is similar to the well-known conventions that a sum of 0 numbers is 0, and that a product of 0 numbers is 1.)

This definition is slightly dangerous, because it entails that the composition of 0 maps depends on the set X_1, but of course the 0 maps being composed know nothing about this set X_1. Thus, when we speak of an empty composition, we should always specify the set X_1 or ensure that it is clear from the context. (See Definition 2.94 below for an example where it is clear from the context.)

2.13.8. Composition powers

Having defined the composition of n maps, we get the notion of composition powers of maps for free:

Definition 2.94. Let $n \in \mathbb{N}$. Let X be a set. Let $f : X \to X$ be a map. Then, $f^{\circ n}$ shall denote the map

$$\underbrace{f \circ f \circ \cdots \circ f}_{n \text{ times } f} : X \to X.$$

Thus, in particular,

$$f^{\circ 0} = \underbrace{f \circ f \circ \cdots \circ f}_{0 \text{ times } f} = \operatorname{id}_X \tag{139}$$

(by Definition 2.93). Also, $f^{\circ 1} = f$, $f^{\circ 2} = f \circ f$, $f^{\circ 3} = f \circ f \circ f$, etc.

The map $f^{\circ n}$ is called the *n-th composition power* of f (or simply the *n-th power* of f).

Before we study composition powers in detail, let us show a general rule that allows us to "split" compositions of maps:

Theorem 2.95. Let $n \in \mathbb{N}$. Let $X_1, X_2, \ldots, X_{n+1}$ be $n+1$ sets. For each $i \in \{1, 2, \ldots, n\}$, let $f_i : X_i \to X_{i+1}$ be a map.

(a) We have

$$f_n \circ f_{n-1} \circ \cdots \circ f_1 = (f_n \circ f_{n-1} \circ \cdots \circ f_{k+1}) \circ (f_k \circ f_{k-1} \circ \cdots \circ f_1)$$

for each $k \in \{0, 1, \ldots, n\}$.

(b) If $n \geq 1$, then

$$f_n \circ f_{n-1} \circ \cdots \circ f_1 = f_n \circ (f_{n-1} \circ f_{n-2} \circ \cdots \circ f_1) .$$

(c) If $n \geq 1$, then

$$f_n \circ f_{n-1} \circ \cdots \circ f_1 = (f_n \circ f_{n-1} \circ \cdots \circ f_2) \circ f_1.$$

Proof of Theorem 2.95. **(a)** Let $k \in \{0, 1, \ldots, n\}$. We are in one of the following three cases:

Case 1: We have $k = 0$.

Case 2: We have $k = n$.

Case 3: We have neither $k = 0$ nor $k = n$.

(Of course, Cases 1 and 2 overlap when $n = 0$.)

Let us first consider Case 1. In this case, we have $k = 0$. Thus, $f_k \circ f_{k-1} \circ \cdots \circ f_1$ is an empty composition of maps, and therefore equals id_{X_1}. In other words, $f_k \circ f_{k-1} \circ \cdots \circ f_1 = \operatorname{id}_{X_1}$.

On the other hand, $k = 0$, so that $k + 1 = 1$. Hence, $f_n \circ f_{n-1} \circ \cdots \circ f_{k+1} = f_n \circ f_{n-1} \circ \cdots \circ f_1$. Thus,

$$\underbrace{(f_n \circ f_{n-1} \circ \cdots \circ f_{k+1})}_{=f_n \circ f_{n-1} \circ \cdots \circ f_1} \circ \underbrace{(f_k \circ f_{k-1} \circ \cdots \circ f_1)}_{=\operatorname{id}_{X_1}}$$
$$= (f_n \circ f_{n-1} \circ \cdots \circ f_1) \circ \operatorname{id}_{X_1} = f_n \circ f_{n-1} \circ \cdots \circ f_1.$$

In other words,

$$f_n \circ f_{n-1} \circ \cdots \circ f_1 = (f_n \circ f_{n-1} \circ \cdots \circ f_{k+1}) \circ (f_k \circ f_{k-1} \circ \cdots \circ f_1).$$

Hence, Theorem 2.95 **(a)** is proven in Case 1.

Let us next consider Case 2. In this case, we have $k = n$. Thus, $f_n \circ f_{n-1} \circ \cdots \circ f_{k+1}$ is an empty composition of maps, and therefore equals $\operatorname{id}_{X_{n+1}}$. In other words, $f_n \circ f_{n-1} \circ \cdots \circ f_{k+1} = \operatorname{id}_{X_{n+1}}$. Thus,

$$\underbrace{(f_n \circ f_{n-1} \circ \cdots \circ f_{k+1})}_{=\operatorname{id}_{X_{n+1}}} \circ \underbrace{(f_k \circ f_{k-1} \circ \cdots \circ f_1)}_{\substack{=f_n \circ f_{n-1} \circ \cdots \circ f_1 \\ \text{(since } k=n\text{)}}}$$
$$= \operatorname{id}_{X_{n+1}} \circ (f_n \circ f_{n-1} \circ \cdots \circ f_1) = f_n \circ f_{n-1} \circ \cdots \circ f_1.$$

In other words,

$$f_n \circ f_{n-1} \circ \cdots \circ f_1 = (f_n \circ f_{n-1} \circ \cdots \circ f_{k+1}) \circ (f_k \circ f_{k-1} \circ \cdots \circ f_1).$$

Hence, Theorem 2.95 **(a)** is proven in Case 2.

Let us finally consider Case 3. In this case, we have neither $k = 0$ nor $k = n$. In other words, we have $k \neq 0$ and $k \neq n$. Combining $k \in \{0, 1, \ldots, n\}$ with $k \neq 0$, we find $k \in \{0, 1, \ldots, n\} \setminus \{0\} \subseteq \{1, 2, \ldots, n\}$. Combining this with $k \neq n$, we find $k \in \{1, 2, \ldots, n\} \setminus \{n\} \subseteq \{1, 2, \ldots, n-1\}$. Hence, $1 \leq k \leq n - 1$, so that $n - 1 \geq 1$ and thus $n \geq 2 \geq 1$. Hence, n is a positive integer. Thus, Proposition 2.89 yields

$$C(f_n, f_{n-1}, \ldots, f_1) = C(f_n, f_{n-1}, \ldots, f_{k+1}) \circ C(f_k, f_{k-1}, \ldots, f_1). \qquad (140)$$

Now, k is a positive integer (since $k \in \{1, 2, \ldots, n-1\}$). Hence, Definition 2.91 (applied to k instead of n) yields

$$f_k \circ f_{k-1} \circ \cdots \circ f_1 = C(f_k, f_{k-1}, \ldots, f_1). \qquad (141)$$

Also, $n-k$ is an integer satisfying $n - \underbrace{k}_{\leq n-1} \geq n-(n-1) = 1$. Hence, $n-k$ is a positive integer. Thus, Definition 2.91 (applied to $n-k$, X_{k+i} and f_{k+i} instead of n, X_i and f_i) yields

$$f_{k+(n-k)} \circ f_{k+(n-k-1)} \circ \cdots \circ f_{k+1} = C\left(f_{k+(n-k)}, f_{k+(n-k-1)}, \ldots, f_{k+1}\right).$$

In view of $k+(n-k) = n$ and $k+(n-k-1) = n-1$, this rewrites as follows:

$$f_n \circ f_{n-1} \circ \cdots \circ f_{k+1} = C\left(f_n, f_{n-1}, \ldots, f_{k+1}\right). \qquad (142)$$

But n is a positive integer. Thus, Definition 2.91 yields

$$\begin{aligned} f_n \circ f_{n-1} \circ \cdots \circ f_1 &= C\left(f_n, f_{n-1}, \ldots, f_1\right) \\ &= \underbrace{C\left(f_n, f_{n-1}, \ldots, f_{k+1}\right)}_{\substack{=f_n \circ f_{n-1} \circ \cdots \circ f_{k+1} \\ \text{(by (142))}}} \circ \underbrace{C\left(f_k, f_{k-1}, \ldots, f_1\right)}_{\substack{=f_k \circ f_{k-1} \circ \cdots \circ f_1 \\ \text{(by (141))}}} \qquad \text{(by (140))} \\ &= \left(f_n \circ f_{n-1} \circ \cdots \circ f_{k+1}\right) \circ \left(f_k \circ f_{k-1} \circ \cdots \circ f_1\right). \end{aligned}$$

Hence, Theorem 2.95 **(a)** is proven in Case 3.

We have now proven Theorem 2.95 **(a)** in each of the three Cases 1, 2 and 3. Since these three Cases cover all possibilities, we thus conclude that Theorem 2.95 **(a)** always holds.

(b) Assume that $n \geq 1$. Hence, $n-1 \geq 0$, so that $n-1 \in \{0, 1, \ldots, n\}$ (since $n-1 \leq n$). Hence, Theorem 2.95 **(a)** (applied to $k = n-1$) yields

$$\begin{aligned} f_n \circ f_{n-1} \circ \cdots \circ f_1 &= \underbrace{\left(f_n \circ f_{n-1} \circ \cdots \circ f_{(n-1)+1}\right)}_{=f_n \circ f_{n-1} \circ \cdots \circ f_n = f_n} \circ \underbrace{\left(f_{n-1} \circ f_{(n-1)-1} \circ \cdots \circ f_1\right)}_{=f_{n-1} \circ f_{n-2} \circ \cdots \circ f_1} \\ &= f_n \circ \left(f_{n-1} \circ f_{n-2} \circ \cdots \circ f_1\right). \end{aligned}$$

This proves Theorem 2.95 **(b)**.

(c) Assume that $n \geq 1$. Hence, $1 \in \{0, 1, \ldots, n\}$. Hence, Theorem 2.95 **(a)** (applied to $k = 1$) yields

$$\begin{aligned} f_n \circ f_{n-1} \circ \cdots \circ f_1 &= \underbrace{\left(f_n \circ f_{n-1} \circ \cdots \circ f_{1+1}\right)}_{=f_n \circ f_{n-1} \circ \cdots \circ f_2} \circ \underbrace{\left(f_1 \circ f_{1-1} \circ \cdots \circ f_1\right)}_{=f_1} \\ &= \left(f_n \circ f_{n-1} \circ \cdots \circ f_2\right) \circ f_1. \end{aligned}$$

This proves Theorem 2.95 **(c)**. □

We can draw some consequences about composition powers of maps from Theorem 2.95:

Proposition 2.96. Let X be a set. Let $f : X \to X$ be a map. Let n be a positive integer.

(a) We have $f^{\circ n} = f \circ f^{\circ(n-1)}$.

(b) We have $f^{\circ n} = f^{\circ(n-1)} \circ f$.

Proof of Proposition 2.96. The definition of $f^{\circ n}$ yields

$$f^{\circ n} = \underbrace{f \circ f \circ \cdots \circ f}_{n \text{ times } f}. \tag{143}$$

The definition of $f^{\circ(n-1)}$ yields

$$f^{\circ(n-1)} = \underbrace{f \circ f \circ \cdots \circ f}_{n-1 \text{ times } f}. \tag{144}$$

(a) Theorem 2.95 **(b)** (applied to $X_i = X$ and $f_i = f$) yields

$$\underbrace{f \circ f \circ \cdots \circ f}_{n \text{ times } f} = f \circ \left(\underbrace{f \circ f \circ \cdots \circ f}_{n-1 \text{ times } f} \right).$$

In view of (143) and (144), this rewrites as $f^{\circ n} = f \circ f^{\circ(n-1)}$. This proves Proposition 2.96 **(a)**.

(b) Theorem 2.95 **(c)** (applied to $X_i = X$ and $f_i = f$) yields

$$\underbrace{f \circ f \circ \cdots \circ f}_{n \text{ times } f} = \left(\underbrace{f \circ f \circ \cdots \circ f}_{n-1 \text{ times } f} \right) \circ f.$$

In view of (143) and (144), this rewrites as $f^{\circ n} = f^{\circ(n-1)} \circ f$. This proves Proposition 2.96 **(b)**. □

Proposition 2.97. Let X be a set. Let $f : X \to X$ be a map.

(a) We have $f^{\circ(a+b)} = f^{\circ a} \circ f^{\circ b}$ for every $a \in \mathbb{N}$ and $b \in \mathbb{N}$.

(b) We have $f^{\circ(ab)} = \left(f^{\circ a}\right)^{\circ b}$ for every $a \in \mathbb{N}$ and $b \in \mathbb{N}$.

Proof of Proposition 2.97. **(a)** Let $a \in \mathbb{N}$ and $b \in \mathbb{N}$. Thus, $a \geq 0$ and $b \geq 0$, so that $0 \leq b \leq a + b$ (since $\underbrace{a}_{\geq 0} + b \geq b$). Hence, $b \in \{0, 1, \ldots, a + b\}$. Thus, Theorem 2.95 **(a)** (applied to $n = a + b$, $X_i = X$, $f_i = f$ and $k = b$) yields

$$\underbrace{f \circ f \circ \cdots \circ f}_{a+b \text{ times } f} = \left(\underbrace{f \circ f \circ \cdots \circ f}_{(a+b)-b \text{ times } f} \right) \circ \left(\underbrace{f \circ f \circ \cdots \circ f}_{b \text{ times } f} \right).$$

In view of $(a+b)-b=a$, this rewrites as

$$\underbrace{f\circ f\circ\cdots\circ f}_{a+b\text{ times }f}=\left(\underbrace{f\circ f\circ\cdots\circ f}_{a\text{ times }f}\right)\circ\left(\underbrace{f\circ f\circ\cdots\circ f}_{b\text{ times }f}\right). \tag{145}$$

But the definition of $f^{\circ a}$ yields

$$f^{\circ a}=\underbrace{f\circ f\circ\cdots\circ f}_{a\text{ times }f}. \tag{146}$$

Also, the definition of $f^{\circ b}$ yields

$$f^{\circ b}=\underbrace{f\circ f\circ\cdots\circ f}_{b\text{ times }f}. \tag{147}$$

Finally, the definition of $f^{\circ(a+b)}$ yields

$$f^{\circ(a+b)}=\underbrace{f\circ f\circ\cdots\circ f}_{a+b\text{ times }f}. \tag{148}$$

In view of these three equalities (146), (147) and (148), we can rewrite the equality (145) as $f^{\circ(a+b)}=f^{\circ a}\circ f^{\circ b}$. This proves Proposition 2.97 **(a)**.

(Alternatively, it is easy to prove Proposition 2.97 **(a)** by induction on a.)

(b) Let $a\in\mathbb{N}$. We claim that

$$f^{\circ(ab)}=\left(f^{\circ a}\right)^{\circ b}\qquad\text{for every }b\in\mathbb{N}. \tag{149}$$

We shall prove (149) by induction on b:

Induction base: We have $a\cdot 0=0$ and thus $f^{\circ(a\cdot 0)}=f^{\circ 0}=\operatorname{id}_X$ (by (139)). Comparing this with $\left(f^{\circ a}\right)^{\circ 0}=\operatorname{id}_X$ (which follows from (139), applied to $f^{\circ a}$ instead of f), we obtain $f^{\circ(a\cdot 0)}=\left(f^{\circ a}\right)^{\circ 0}$. In other words, (149) holds for $b=0$. This completes the induction base.

Induction step: Let $m\in\mathbb{N}$. Assume that (149) holds for $b=m$. We must prove that (149) holds for $b=m+1$.

We have assumed that (149) holds for $b=m$. In other words, we have $f^{\circ(am)}=\left(f^{\circ a}\right)^{\circ m}$.

But $m+1$ is a positive integer (since $m+1>m\geq 0$). Hence, Proposition 2.96 **(b)** (applied to $m+1$ and $f^{\circ a}$ instead of n and f) yields

$$\left(f^{\circ a}\right)^{\circ(m+1)}=\left(f^{\circ a}\right)^{\circ((m+1)-1)}\circ f^{\circ a}=\left(f^{\circ a}\right)^{\circ m}\circ f^{\circ a} \tag{150}$$

(since $(m+1)-1=m$).

But $a(m+1) = am + a$. Thus,

$$f^{\circ(a(m+1))} = f^{\circ(am+a)} = \underbrace{f^{\circ(am)}}_{=(f^{\circ a})^{\circ m}} \circ f^{\circ a} \qquad \left(\begin{array}{c}\text{by Proposition 2.97 \textbf{(a)}}\\ \text{(applied to } am \text{ and } a \text{ instead of } a \text{ and } b\text{)}\end{array}\right)$$
$$= (f^{\circ a})^{\circ m} \circ f^{\circ a} = (f^{\circ a})^{\circ(m+1)} \qquad \text{(by (150))}.$$

In other words, (149) holds for $b = m+1$. This completes the induction step. Thus, (149) is proven by induction. Hence, Proposition 2.97 **(b)** is proven. □

Note that Proposition 2.97 is similar to the rules of exponents

$$n^{a+b} = n^a n^b \qquad \text{and} \qquad n^{ab} = (n^a)^b$$

that hold for $n \in \mathbb{Q}$ and $a, b \in \mathbb{N}$ (and for various other situations). Can we find similar analogues for other rules of exponents, such as $(mn)^a = m^a n^a$? The simplest analogue one could think of for this rule would be $(f \circ g)^{\circ a} = f^{\circ a} \circ g^{\circ a}$; but this does not hold in general (unless $a \leq 1$). However, it turns out that this does hold if we assume that $f \circ g = g \circ f$ (which is not automatically true, unlike the analogous equality $mn = nm$ for integers). Let us prove this:

Proposition 2.98. Let X be a set. Let $f : X \to X$ and $g : X \to X$ be two maps such that $f \circ g = g \circ f$. Then:
(a) We have $f \circ g^{\circ b} = g^{\circ b} \circ f$ for each $b \in \mathbb{N}$.
(b) We have $f^{\circ a} \circ g^{\circ b} = g^{\circ b} \circ f^{\circ a}$ for each $a \in \mathbb{N}$ and $b \in \mathbb{N}$.
(c) We have $(f \circ g)^{\circ a} = f^{\circ a} \circ g^{\circ a}$ for each $a \in \mathbb{N}$.

Example 2.99. Let us see why the requirement $f \circ g = g \circ f$ is needed in Proposition 2.98:

Let X be the set $\mathbb{Z}$. Let $f : X \to X$ be the map that sends every integer x to $-x$. Let $g : X \to X$ be the map that sends every integer x to $1 - x$. Then, $f^{\circ 2} = \operatorname{id}_X$ (since $f^{\circ 2}(x) = f(f(x)) = -(-x) = x$ for each $x \in X$) and $g^{\circ 2} = \operatorname{id}_X$ (since $g^{\circ 2}(x) = g(g(x)) = 1 - (1-x) = x$ for each $x \in X$). But the map $f \circ g$ satisfies $(f \circ g)(x) = f(g(x)) = -(1-x) = x - 1$ for each $x \in X$. Hence, $(f \circ g)^{\circ 2}(x) = (f \circ g)((f \circ g)(x)) = (x-1) - 1 = x - 2$ for each $x \in X$. Thus, $(f \circ g)^{\circ 2} \neq \operatorname{id}_X$. Comparing this with $\underbrace{f^{\circ 2}}_{=\operatorname{id}_X} \circ \underbrace{g^{\circ 2}}_{=\operatorname{id}_X} = \operatorname{id}_X \circ \operatorname{id}_X = \operatorname{id}_X$, we obtain $(f \circ g)^{\circ 2} \neq f^{\circ 2} \circ g^{\circ 2}$. This shows that Proposition 2.98 **(c)** would not hold without the requirement $f \circ g = g \circ f$.

Proof of Proposition 2.98. **(a)** We claim that

$$f \circ g^{\circ b} = g^{\circ b} \circ f \qquad \text{for each } b \in \mathbb{N}. \tag{151}$$

Indeed, let us prove (151) by induction on b:

Induction base: We have $g^{\circ 0} = \operatorname{id}_X$ (by (139), applied to g instead of f). Hence, $f \circ \underbrace{g^{\circ 0}}_{=\operatorname{id}_X} = f \circ \operatorname{id}_X = f$ and $\underbrace{g^{\circ 0}}_{=\operatorname{id}_X} \circ f = \operatorname{id}_X \circ f = f$. Comparing these two equalities, we obtain $f \circ g^{\circ 0} = g^{\circ 0} \circ f$. In other words, (151) holds for $b = 0$. This completes the induction base.

Induction step: Let $m \in \mathbb{N}$. Assume that (151) holds for $b = m$. We must prove that (151) holds for $b = m + 1$.

We have assumed that (151) holds for $b = m$. In other words,

$$f \circ g^{\circ m} = g^{\circ m} \circ f. \tag{152}$$

Proposition 2.82 (applied to $Y = X$, $Z = X$, $W = X$, $c = g$, $b = g^{\circ m}$ and $a = f$) yields

$$(f \circ g^{\circ m}) \circ g = f \circ (g^{\circ m} \circ g). \tag{153}$$

Proposition 2.97 **(a)** (applied to g, m and 1 instead of f, a and b) yields

$$g^{\circ(m+1)} = g^{\circ m} \circ \underbrace{g^{\circ 1}}_{=g} = g^{\circ m} \circ g. \tag{154}$$

Hence,

$$\begin{aligned} f \circ \underbrace{g^{\circ(m+1)}}_{=g^{\circ m} \circ g} &= f \circ (g^{\circ m} \circ g) = \underbrace{(f \circ g^{\circ m})}_{\substack{=g^{\circ m} \circ f \\ \text{(by (152))}}} \circ g \qquad \text{(by (153))} \\ &= (g^{\circ m} \circ f) \circ g = g^{\circ m} \circ \underbrace{(f \circ g)}_{=g \circ f} \\ &\qquad \left(\begin{array}{c} \text{by Proposition 2.82 (applied} \\ \text{to } Y = X,\ Z = X,\ W = X,\ c = g,\ b = f \text{ and } a = g^{\circ m}) \end{array} \right) \\ &= g^{\circ m} \circ (g \circ f). \end{aligned} \tag{155}$$

On the other hand, Proposition 2.82 (applied to $Y = X$, $Z = X$, $W = X$, $c = f$, $b = g$ and $a = g^{\circ m}$) yields

$$(g^{\circ m} \circ g) \circ f = g^{\circ m} \circ (g \circ f).$$

Comparing this with (155), we obtain

$$f \circ g^{\circ(m+1)} = \underbrace{(g^{\circ m} \circ g)}_{\substack{=g^{\circ(m+1)} \\ \text{(by (154))}}} \circ f = g^{\circ(m+1)} \circ f.$$

In other words, (151) holds for $b = m + 1$. This completes the induction step. Thus, (151) is proven by induction.

Therefore, Proposition 2.98 **(a)** follows.

(b) Let $a \in \mathbb{N}$ and $b \in \mathbb{N}$. From $f \circ g = g \circ f$, we obtain $g \circ f = f \circ g$. Hence, Proposition 2.98 **(a)** (applied to g, f and a instead of f, g and b) yields $g \circ f^{\circ a} = f^{\circ a} \circ g$. In other words, $f^{\circ a} \circ g = g \circ f^{\circ a}$. Hence, Proposition 2.98 **(a)** (applied to $f^{\circ a}$ instead of f) yields $f^{\circ a} \circ g^{\circ b} = g^{\circ b} \circ f^{\circ a}$. This proves Proposition 2.98 **(b)**.

(c) We claim that

$$(f \circ g)^{\circ a} = f^{\circ a} \circ g^{\circ a} \qquad \text{for each } a \in \mathbb{N}. \tag{156}$$

Indeed, let us prove (156) by induction on a:

Induction base: From (139), we obtain $f^{\circ 0} = \operatorname{id}_X$ and $g^{\circ 0} = \operatorname{id}_X$ and $(f \circ g)^{\circ 0} = \operatorname{id}_X$. Thus,

$$(f \circ g)^{\circ 0} = \operatorname{id}_X = \underbrace{\operatorname{id}_X}_{=f^{\circ 0}} \circ \underbrace{\operatorname{id}_X}_{=g^{\circ 0}} = f^{\circ 0} \circ g^{\circ 0}.$$

In other words, (156) holds for $a = 0$. This completes the induction base.

Induction step: Let $m \in \mathbb{N}$. Assume that (156) holds for $a = m$. We must prove that (156) holds for $a = m + 1$.

We have assumed that (156) holds for $a = m$. In other words,

$$(f \circ g)^{\circ m} = f^{\circ m} \circ g^{\circ m}. \tag{157}$$

But Proposition 2.97 **(a)** (applied to g, m and 1 instead of f, a and b) yields

$$g^{\circ(m+1)} = g^{\circ m} \circ \underbrace{g^{\circ 1}}_{=g} = g^{\circ m} \circ g.$$

The same argument (applied to f instead of g) yields $f^{\circ(m+1)} = f^{\circ m} \circ f$. Hence,

$$\underbrace{f^{\circ(m+1)}}_{=f^{\circ m} \circ f} \circ g^{\circ(m+1)} = \left(f^{\circ m} \circ f\right) \circ g^{\circ(m+1)} = f^{\circ m} \circ \left(f \circ g^{\circ(m+1)}\right) \tag{158}$$

(by Proposition 2.82 (applied to $Y = X$, $Z = X$, $W = X$, $c = g^{\circ(m+1)}$, $b = f$ and $a = f^{\circ m}$)).

But Proposition 2.98 **(a)** (applied to $b = m + 1$) yields

$$f \circ g^{\circ(m+1)} = \underbrace{g^{\circ(m+1)}}_{=g^{\circ m} \circ g} \circ f = \left(g^{\circ m} \circ g\right) \circ f = g^{\circ m} \circ (g \circ f)$$

(by Proposition 2.82 (applied to $Y = X$, $Z = X$, $W = X$, $c = f$, $b = g$ and $a = g^{\circ m}$)). Hence,

$$f \circ g^{\circ(m+1)} = g^{\circ m} \circ \underbrace{(g \circ f)}_{\substack{=f \circ g \\ (\text{since } f \circ g = g \circ f)}} = g^{\circ m} \circ (f \circ g). \tag{159}$$

On the other hand, Proposition 2.97 **(a)** (applied to $f \circ g$, m and 1 instead of f, a and b) yields

$$(f \circ g)^{\circ(m+1)} = \underbrace{(f \circ g)^{\circ m}}_{\substack{= f^{\circ m} \circ g^{\circ m} \\ \text{(by (157))}}} \circ \underbrace{(f \circ g)^{\circ 1}}_{= f \circ g} = (f^{\circ m} \circ g^{\circ m}) \circ (f \circ g) = f^{\circ m} \circ (g^{\circ m} \circ (f \circ g))$$

(by Proposition 2.82 (applied to $Y = X$, $Z = X$, $W = X$, $c = f \circ g$, $b = g^{\circ m}$ and $a = f^{\circ m}$)). Hence,

$$(f \circ g)^{\circ(m+1)} = f^{\circ m} \circ \underbrace{(g^{\circ m} \circ (f \circ g))}_{\substack{= f \circ g^{\circ(m+1)} \\ \text{(by (159))}}} = f^{\circ m} \circ \left(f \circ g^{\circ(m+1)} \right) = f^{\circ(m+1)} \circ g^{\circ(m+1)}$$

(by (158)). In other words, (156) holds for $a = m + 1$. This completes the induction step. Thus, (156) is proven by induction. Therefore, Proposition 2.98 **(c)** follows. □

Remark 2.100. In our above proof of Proposition 2.98, we have not used the notation $f_n \circ f_{n-1} \circ \cdots \circ f_1$ introduced in Definition 2.91, but instead relied on parentheses and compositions of two maps (i.e., we have never composed more than two maps at the same time). Thus, for example, in the proof of Proposition 2.98 **(a)**, we wrote "$(g^{\circ m} \circ g) \circ f$" and "$g^{\circ m} \circ (g \circ f)$" rather than "$g^{\circ m} \circ g \circ f$". But Remark 2.92 says that we could have just as well dropped all the parentheses. This would have saved us the trouble of explicitly applying Proposition 2.82 (since if we drop all parentheses, then there is no difference between "$(g^{\circ m} \circ g) \circ f$" and "$g^{\circ m} \circ (g \circ f)$" any more). This way, the induction step in the proof of Proposition 2.98 **(a)** could have been made much shorter:

Induction step (second version): Let $m \in \mathbb{N}$. Assume that (151) holds for $b = m$. We must prove that (151) holds for $b = m + 1$.

We have assumed that (151) holds for $b = m$. In other words,

$$f \circ g^{\circ m} = g^{\circ m} \circ f. \tag{160}$$

Proposition 2.97 **(a)** (applied to g, m and 1 instead of f, a and b) yields $g^{\circ(m+1)} = g^{\circ m} \circ \underbrace{g^{\circ 1}}_{= g} = g^{\circ m} \circ g$. Hence,

$$f \circ \underbrace{g^{\circ(m+1)}}_{= g^{\circ m} \circ g} = \underbrace{f \circ g^{\circ m}}_{\substack{= g^{\circ m} \circ f \\ \text{(by (160))}}} \circ g = g^{\circ m} \circ \underbrace{f \circ g}_{= g \circ f} = \underbrace{g^{\circ m} \circ g}_{= g^{\circ(m+1)}} \circ f = g^{\circ(m+1)} \circ f.$$

In other words, (151) holds for $b = m + 1$. This completes the induction step.

Similarly, we can simplify the proof of Proposition 2.98 **(c)** by dropping the parentheses. (The details are left to the reader.)

2.13.9. Composition of invertible maps

The composition of two invertible maps is always invertible, and its inverse can be computed by the following formula:

Proposition 2.101. Let X, Y and Z be three sets. Let $b : X \to Y$ and $a : Y \to Z$ be two invertible maps. Then, the map $a \circ b : X \to Z$ is invertible as well, and its inverse is

$$(a \circ b)^{-1} = b^{-1} \circ a^{-1}.$$

Proof of Proposition 2.101. As we have explained in Definition 2.91, we can drop the parentheses when composing several maps. This will allow us to write expressions like $b^{-1} \circ a^{-1} \circ a \circ b$ without specifying where parentheses should be placed, and then pretending that they are placed wherever we would find them most convenient.

The equalities

$$\left(b^{-1} \circ a^{-1}\right) \circ (a \circ b) = b^{-1} \circ \underbrace{a^{-1} \circ a}_{=\mathrm{id}_Y} \circ b = b^{-1} \circ b = \mathrm{id}_X$$

and

$$(a \circ b) \circ \left(b^{-1} \circ a^{-1}\right) = a \circ \underbrace{b \circ b^{-1}}_{=\mathrm{id}_Y} \circ a^{-1} = a \circ a^{-1} = \mathrm{id}_Z$$

show that the map $b^{-1} \circ a^{-1}$ is an inverse of $a \circ b$. Thus, the map $a \circ b$ has an inverse (namely, $b^{-1} \circ a^{-1}$). In other words, the map $a \circ b$ is invertible. Its inverse is $(a \circ b)^{-1} = b^{-1} \circ a^{-1}$ (since $b^{-1} \circ a^{-1}$ is an inverse of $a \circ b$). This completes the proof of Proposition 2.101. □

By a straightforward induction, we can extend Proposition 2.101 to compositions of n invertible maps:

Proposition 2.102. Let $n \in \mathbb{N}$. Let $X_1, X_2, \ldots, X_{n+1}$ be $n+1$ sets. For each $i \in \{1, 2, \ldots, n\}$, let $f_i : X_i \to X_{i+1}$ be an invertible map. Then, the map $f_n \circ f_{n-1} \circ \cdots \circ f_1 : X_1 \to X_{n+1}$ is invertible as well, and its inverse is

$$(f_n \circ f_{n-1} \circ \cdots \circ f_1)^{-1} = f_1^{-1} \circ f_2^{-1} \circ \cdots \circ f_n^{-1}.$$

Exercise 2.2. Prove Proposition 2.102.

In particular, Proposition 2.102 shows that any composition of invertible maps is invertible. Since invertible maps are the same as bijective maps, we can rewrite this as follows: Any composition of bijective maps is bijective.

2.14. General commutativity for addition of numbers

2.14.1. The setup and the problem

Throughout Section 2.14, we let $\mathbb{A}$ be one of the sets $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$ and $\mathbb{C}$. The elements of $\mathbb{A}$ will be simply called *numbers*.

There is an analogue of Proposition 2.82 for numbers:

Proposition 2.103. Let a, b and c be three numbers (i.e., elements of $\mathbb{A}$). Then, $(a+b)+c = a+(b+c)$.

Proposition 2.103 is known as the *associativity of addition* (in $\mathbb{A}$), and is fundamental; its proof can be found in any textbook on the construction of the number system[76].

In Section 2.13, we have used Proposition 2.82 to show that we can "drop the parentheses" in a composition $f_n \circ f_{n-1} \circ \cdots \circ f_1$ of maps (i.e., all possible complete parenthesizations of this composition are actually the same map). Likewise, we can use Proposition 2.103 to show that we can "drop the parentheses" in a sum $a_1 + a_2 + \cdots + a_n$ of numbers (i.e., all possible complete parenthesizations of this sum are actually the same number). For example, if a, b, c, d are four numbers, then the complete parenthesizations of $a+b+c+d$ are

$$\begin{array}{lll} ((a+b)+c)+d, & (a+(b+c))+d, & (a+b)+(c+d), \\ a+((b+c)+d), & a+(b+(c+d)), & \end{array}$$

and all of these five complete parenthesizations are the same number.

However, numbers behave better than maps. In particular, along with Proposition 2.103, they satisfy another law that maps (generally) don't satisfy:

Proposition 2.104. Let a and b be two numbers (i.e., elements of $\mathbb{A}$). Then, $a+b = b+a$.

Proposition 2.104 is known as the *commutativity of addition* (in $\mathbb{A}$), and again is a fundamental result whose proofs are found in standard textbooks[77].

Furthermore, numbers can **always** be added, whereas maps can only be composed if the domain of one is the codomain of the other. Thus, when we want to take the sum of n numbers $a_1, a_2, \ldots, a_n$, we can not only choose where to put the parentheses, but also in what order the numbers should appear in the sum. It turns

[76] For example, Proposition 2.103 is proven in [Swanso20, Theorem 3.2.3 (3)] for the case when $\mathbb{A} = \mathbb{N}$; in [Swanso20, Theorem 3.5.4 (3)] for the case when $\mathbb{A} = \mathbb{Z}$; in [Swanso20, Theorem 3.6.4 (3)] for the case when $\mathbb{A} = \mathbb{Q}$; in [Swanso20, Theorem 3.7.11] for the case when $\mathbb{A} = \mathbb{R}$; in [Swanso20, Theorem 3.9.2] for the case when $\mathbb{A} = \mathbb{C}$.

[77] For example, Proposition 2.104 is proven in [Swanso20, Theorem 3.2.3 (4)] for the case when $\mathbb{A} = \mathbb{N}$; in [Swanso20, Theorem 3.5.4 (4)] for the case when $\mathbb{A} = \mathbb{Z}$; in [Swanso20, Theorem 3.6.4 (4)] for the case when $\mathbb{A} = \mathbb{Q}$; in [Swanso20, Theorem 3.7.11] for the case when $\mathbb{A} = \mathbb{R}$; in [Swanso20, Theorem 3.9.2] for the case when $\mathbb{A} = \mathbb{C}$.

out that neither of these choices affects the result. For example, if a, b, c are three numbers, then all 12 possible sums

$$\begin{array}{llll} (a+b)+c, & a+(b+c), & (a+c)+b, & a+(c+b), \\ (b+a)+c, & b+(a+c), & (b+c)+a, & b+(c+a), \\ (c+a)+b, & c+(a+b), & (c+b)+a, & c+(b+a) \end{array}$$

are actually the same number. The reader can easily verify this for three numbers a, b, c (using Proposition 2.103 and Proposition 2.104), but of course the general case (with n numbers) is more difficult. The independence of the result on the parenthesization can be proven using the same arguments that we gave in Section 2.13 (except that the $\circ$ symbol is now replaced by $+$), but the independence on the order cannot easily be shown (or even stated) in this way.

Thus, we shall proceed differently: We shall rigorously define the sum of n numbers without specifying an order in which they are added or using parentheses. Unlike the composition of n maps, which was defined for an *ordered list* of n maps, we shall define the sum of n numbers for a *family* of n numbers (see the next subsection for the definition of a "family"). Families don't come with an ordering chosen in advance, so we cannot single out any specific ordering for use in the definition. Thus, the independence on the order will be baked right into the definition.

Different solutions to the problem of formalizing the concept of the sum of n numbers can be found in [Bourba74, Chapter 1, §1.5][78], in [GalQua22, §3.3] and in [Clemen22, §2.4].

2.14.2. Families

Let us first define what we mean by a "family" of n numbers. More generally, we can define a family of elements of any set, or even a family of elements of **different** sets. To motivate the definition, we first recall a concept of an n-tuple:

Remark 2.105. Let $n \in \mathbb{N}$.

(a) Let A be a set. Then, to specify an *n-tuple of elements of A* means specifying an element a_i of A for each $i \in \{1, 2, \ldots, n\}$. This n-tuple is then denoted by $(a_1, a_2, \ldots, a_n)$ or by $(a_i)_{i \in \{1,2,\ldots,n\}}$. For each $i \in \{1, 2, \ldots, n\}$, we refer to a_i as the *i-th entry* of this n-tuple.

The set of all n-tuples of elements of A is denoted by A^n or by $A^{\times n}$; it is called the *n-th Cartesian power* of the set A.

(b) More generally, we can define n-tuples of elements from **different** sets: For each $i \in \{1, 2, \ldots, n\}$, let A_i be a set. Then, to specify an *n-tuple of elements of $A_1, A_2, \ldots, A_n$* means specifying an element a_i of A_i for each $i \in \{1, 2, \ldots, n\}$.

[78] Bourbaki, in [Bourba74, Chapter 1, §1.5], define something more general than a sum of n numbers: They define the "composition" of a finite family of elements of a commutative magma. The sum of n numbers is a particular case of this concept when the magma is the set $\mathbb{A}$ (endowed with its addition).

This n-tuple is (again) denoted by $(a_1, a_2, \ldots, a_n)$ or by $(a_i)_{i\in\{1,2,\ldots,n\}}$. For each $i \in \{1, 2, \ldots, n\}$, we refer to a_i as the *i-th entry* of this n-tuple.

The set of all n-tuples of elements of $A_1, A_2, \ldots, A_n$ is denoted by $A_1 \times A_2 \times \cdots \times A_n$ or by $\prod_{i=1}^{n} A_i$; it is called the *Cartesian product* of the n sets $A_1, A_2, \ldots, A_n$. These n sets $A_1, A_2, \ldots, A_n$ are called the *factors* of this Cartesian product.

Example 2.106. **(a)** The 3-tuple $(7, 8, 9)$ is a 3-tuple of elements of $\mathbb{N}$, and also a 3-tuple of elements of $\mathbb{Z}$. It can also be written in the form $(6+i)_{i\in\{1,2,3\}}$. Thus, $(6+i)_{i\in\{1,2,3\}} = (6+1, 6+2, 6+3) = (7, 8, 9) \in \mathbb{N}^3$ and also $(6+i)_{i\in\{1,2,3\}} \in \mathbb{Z}^3$.

(b) The 5-tuple $(\{1\}, \{2\}, \{3\}, \varnothing, \mathbb{N})$ is a 5-tuple of elements of the powerset of $\mathbb{N}$ (since $\{1\}, \{2\}, \{3\}, \varnothing, \mathbb{N}$ are subsets of $\mathbb{N}$, thus elements of the powerset of $\mathbb{N}$).

(c) The 0-tuple $()$ can be viewed as a 0-tuple of elements of **any** set A.

(d) If we let $[n]$ be the set $\{1, 2, \ldots, n\}$ for each $n \in \mathbb{N}$, then $(1, 2, 2, 3, 3)$ is a 5-tuple of elements of $[1], [2], [3], [4], [5]$ (because $1 \in [1]$, $2 \in [2]$, $2 \in [3]$, $3 \in [4]$ and $3 \in [5]$). In other words, $(1, 2, 2, 3, 3) \in [1] \times [2] \times [3] \times [4] \times [5]$.

(e) A 2-tuple is the same as an ordered pair. A 3-tuple is the same as an ordered triple. A 1-tuple of elements of a set A is "almost" the same as a single element of A; more precisely, there is a bijection

$$A \to A^1, \qquad a \mapsto (a)$$

from A to the set of 1-tuples of elements of A.

The notation "$(a_i)_{i\in\{1,2,\ldots,n\}}$" in Remark 2.105 should be pronounced as "the n-tuple whose i-th entry is a_i for each $i \in \{1, 2, \ldots, n\}$". The letter "$i$" is used as a variable in this notation (similar to the "i" in the expression "$\sum_{i=1}^{n} i$" or in the expression "the map $\mathbb{N} \to \mathbb{N}$, $i \mapsto i+1$" or in the expression "for all $i \in \mathbb{N}$, we have $i+1 > i$"); it does not refer to any specific element of $\{1, 2, \ldots, n\}$. As usual, it does not matter which letter we are using for this variable (as long as it does not already have a different meaning); thus, for example, the 3-tuples $(6+i)_{i\in\{1,2,3\}}$ and $(6+j)_{j\in\{1,2,3\}}$ and $(6+x)_{x\in\{1,2,3\}}$ are all identical (and equal $(7, 8, 9)$).

We also note that the "$\prod$" sign in Remark 2.105 **(b)** has a different meaning than the "$\prod$" sign in Section 1.4. The former stands for a Cartesian product of sets, whereas the latter stands for a product of numbers. In particular, a product $\prod_{i=1}^{n} a_i$ of numbers does not change when its factors are swapped, whereas a Cartesian product $\prod_{i=1}^{n} A_i$ of sets does. (In particular, if A and B are two sets, then $A \times B$ and $B \times A$ are different sets in general. The 2-tuple $(1, -1)$ belongs to $\mathbb{N} \times \mathbb{Z}$, but not to $\mathbb{Z} \times \mathbb{N}$.)

Thus, the purpose of an n-tuple is storing several elements (possibly of different sets) in one "container". This is a highly useful notion, but sometimes one wants a more general concept, which can store several elements but not necessarily organized in a "linear order". For example, assume you want to store four integers a, b, c, d in the form of a rectangular table $\left(\begin{array}{cc} a & b \\ c & d \end{array}\right)$ (also known as a "2×2-table of integers"). Such a table doesn't have a well-defined "1-st entry" or "2-nd entry" (unless you agree on a specific order in which you read it); instead, it makes sense to speak of a "$(1,2)$-th entry" (i.e., the entry in row 1 and column 2, which is b) or of a "$(2,2)$-th entry" (i.e., the entry in row 2 and column 2, which is d). Thus, such tables work similarly to n-tuples, but they are "indexed" by pairs (i, j) of appropriate integers rather than by the numbers $1, 2, \ldots, n$.

The concept of a "family" generalizes both n-tuples and rectangular tables: It allows the entries to be indexed by the elements of an arbitrary (possibly infinite) set I instead of the numbers $1, 2, \ldots, n$. Here is its definition (watch the similarities to Remark 2.105):

Definition 2.107. Let I be a set.

(a) Let A be a set. Then, to specify an *I-family of elements of A* means specifying an element a_i of A for each $i \in I$. This I-family is then denoted by $(a_i)_{i \in I}$. For each $i \in I$, we refer to a_i as the *i-th entry* of this I-family. (Unlike the case of n-tuples, there is no notation like $(a_1, a_2, \ldots, a_n)$ for I-families, because there is no natural way in which their entries should be listed.)

An I-family of elements of A is also called an *A-valued I-family*.

The set of all I-families of elements of A is denoted by A^I or by $A^{\times I}$. (Note that the notation A^I is also used for the set of all maps from I to A. But this set is more or less the same as the set of all I-families of elements of A; see Remark 2.109 below for the details.)

(b) More generally, we can define I-families of elements from **different** sets: For each $i \in I$, let A_i be a set. Then, to specify an *I-family of elements of $(A_i)_{i \in I}$* means specifying an element a_i of A_i for each $i \in I$. This I-family is (again) denoted by $(a_i)_{i \in I}$. For each $i \in I$, we refer to a_i as the *i-th entry* of this I-family.

The set of all I-families of elements of $(A_i)_{i \in I}$ is denoted by $\prod_{i \in I} A_i$.

The word "I-family" (without further qualifications) means an I-family of elements of $(A_i)_{i \in I}$ for some sets A_i.

The word "family" (without further qualifications) means an I-family for some set I.

Example 2.108. **(a)** The family $(6 + i)_{i \in \{0,3,5\}}$ is a $\{0, 3, 5\}$-family of elements of $\mathbb{N}$ (that is, an $\mathbb{N}$-valued $\{0, 3, 5\}$-family). It has three entries: Its 0-th entry is $6 + 0 = 6$; its 3-rd entry is $6 + 3 = 9$; its 5-th entry is $6 + 5 = 11$. Of course, this family is also a $\{0, 3, 5\}$-family of elements of $\mathbb{Z}$. If we squint hard enough, we can pretend that this family is simply the 3-tuple $(6, 9, 11)$; but this is not advisable, and also does not extend to situations in which there is no natural order on the set I.

(b) Let X be the set $\{\text{"cat"}, \text{"chicken"}, \text{"dog"}\}$ consisting of three words. Then, we can define an X-family $(a_i)_{i \in X}$ of elements of $\mathbb{N}$ by setting

$$a_{\text{"cat"}} = 4, \qquad a_{\text{"chicken"}} = 2, \qquad a_{\text{"dog"}} = 4.$$

This family has 3 entries, which are 4, 2 and 4; but there is no natural order on the set X, so we cannot identify it with a 3-tuple.

We can also rewrite this family as

$$(\text{the number of legs of a typical specimen of animal } i)_{i \in X}.$$

Of course, not every family will have a description like this; sometimes a family is just a choice of elements without any underlying pattern.

(c) If I is the empty set $\varnothing$, and if A is any set, then there is exactly one I-family of elements of A; namely, the *empty family*. Indeed, specifying such a family means specifying no elements at all, and there is just one way to do that. We can denote the empty family by $()$, just like the empty 0-tuple.

(d) The family $(|i|)_{i \in \mathbb{Z}}$ is a $\mathbb{Z}$-family of elements of $\mathbb{N}$ (because $|i|$ is an element of $\mathbb{N}$ for each $i \in \mathbb{Z}$). It can also be regarded as a $\mathbb{Z}$-family of elements of $\mathbb{Z}$.

(e) If I is the set $\{1, 2, \ldots, n\}$ for some $n \in \mathbb{N}$, and if A is any set, then an I-family $(a_i)_{i \in \{1,2,\ldots,n\}}$ of elements of A is the same as an n-tuple of elements of A. The same holds for families and n-tuples of elements from different sets. Thus, any n sets $A_1, A_2, \ldots, A_n$ satisfy $\prod\limits_{i \in \{1,2,\ldots,n\}} A_i = \prod\limits_{i=1}^{n} A_i$.

The notation "$(a_i)_{i \in I}$" in Definition 2.107 should be pronounced as "the I-family whose i-th entry is a_i for each $i \in I$". The letter "i" is used as a variable in this notation (similar to the "i" in the expression "$\sum\limits_{i=1}^{n} i$"); it does not refer to any specific element of I. As usual, it does not matter which letter we are using for this variable (as long as it does not already have a different meaning); thus, for example, the $\mathbb{Z}$-families $(|i|)_{i \in \mathbb{Z}}$ and $(|p|)_{p \in \mathbb{Z}}$ and $(|w|)_{w \in \mathbb{Z}}$ are all identical.

Remark 2.109. Let I and A be two sets. What is the difference between an A-valued I-family and a map from I to A ? Both of these objects consist of a choice of an element of A for each $i \in I$.

The main difference is terminological: e.g., when we speak of a family, the elements of A that constitute it are called its "entries", whereas for a map they are called its "images" or "values". Also, the notations for them are different: The A-valued I-family $(a_i)_{i \in I}$ corresponds to the map $I \to A,\ i \mapsto a_i$.

There is also another, subtler difference: A map from I to A "knows" what the set A is (so that, for example, the maps $\mathbb{N} \to \mathbb{N},\ i \mapsto i$ and $\mathbb{N} \to \mathbb{Z},\ i \mapsto i$ are considered different, even though they map every element of $\mathbb{N}$ to the same value); but an A-valued I-family does not "know" what the set A is (so that, for example, the $\mathbb{N}$-valued $\mathbb{N}$-family $(i)_{i\in\mathbb{N}}$ is considered identical with the $\mathbb{Z}$-valued $\mathbb{N}$-family $(i)_{i\in\mathbb{N}}$). This matters occasionally when one wants to consider maps or families for different sets simultaneously; it is not relevant if we just work with A-valued I-families (or maps from I to A) for two fixed sets I and A. And either way, these conventions are not universal across the mathematical literature; for some authors, maps from I to A do not "know" what A is, whereas other authors want families to "know" this too.

What is certainly true, independently of any conventions, is the following fact: If I and A are two sets, then the map

$$\{\text{maps from } I \text{ to } A\} \to \{A\text{-valued } I\text{-families}\},$$
$$f \mapsto (f(i))_{i\in I}$$

is bijective. (Its inverse map sends every A-valued I-family $(a_i)_{i\in I}$ to the map $I \to A,\ i \mapsto a_i$.) Thus, there is little harm in equating $\{\text{maps from } I \text{ to } A\}$ with $\{A\text{-valued } I\text{-families}\}$.

We already know from Example 2.108 **(e)** that n-tuples are a particular case of families; the same holds for rectangular tables:

Definition 2.110. Let A be a set. Let $n \in \mathbb{N}$ and $m \in \mathbb{N}$. Then, an *$n \times m$-table* of elements of A means an A-valued $\{1,2,\ldots,n\} \times \{1,2,\ldots,m\}$-family. According to Remark 2.109, this is tantamount to saying that an $n \times m$-table of elements of A means a map from $\{1,2,\ldots,n\} \times \{1,2,\ldots,m\}$ to A, except for notational differences (such as referring to the elements that constitute the $n \times m$-table as "entries" rather than "values") and for the fact that an $n \times m$-table does not "know" A (whereas a map would do).

In future chapters, we shall consider "$n \times m$-matrices", which are defined as maps from $\{1,2,\ldots,n\} \times \{1,2,\ldots,m\}$ to A rather than as A-valued $\{1,2,\ldots,n\} \times \{1,2,\ldots,m\}$-families. We shall keep using the same notations for them as for $n \times m$-tables, but unlike $n \times m$-tables, they will "know" A (that is, two $n \times m$-matrices with the same entries but different sets A will be considered different). Anyway, this difference is minor.

2.14.3. A desirable definition

We now know what an $\mathbb{A}$-valued S-family is (for some set S): It is just a way of choosing some element of $\mathbb{A}$ for each $s \in S$. When this element is called a_s, the S-family is called $(a_s)_{s\in S}$.

We now want to define the sum of an $\mathbb{A}$-valued S-family $(a_s)_{s\in S}$ when the set S is finite. Actually, we have already seen a definition of this sum (which is called $\sum\limits_{s\in S} a_s$) in Section 1.4. The only problem with that definition is that we don't know yet that it is legitimate. Let us nevertheless recall it (rewriting it using the notion of an $\mathbb{A}$-valued S-family):

Definition 2.111. If S is a finite set, and if $(a_s)_{s\in S}$ is an $\mathbb{A}$-valued S-family, then we want to define the number $\sum\limits_{s\in S} a_s$. We define this number by recursion on $|S|$ as follows:

- If $|S| = 0$, then $\sum\limits_{s\in S} a_s$ is defined to be 0.
- Let $n \in \mathbb{N}$. Assume that we have defined $\sum\limits_{s\in S} a_s$ for every finite set S with $|S| = n$ and any $\mathbb{A}$-valued S-family $(a_s)_{s\in S}$. Now, if S is a finite set with $|S| = n + 1$, and if $(a_s)_{s\in S}$ is any $\mathbb{A}$-valued S-family, then $\sum\limits_{s\in S} a_s$ is defined by picking any $t \in S$ and setting
$$\sum_{s\in S} a_s = a_t + \sum_{s\in S\setminus\{t\}} a_s. \qquad (161)$$

As we already observed in Section 1.4, it is not obvious that this definition is legitimate: The right hand side of (161) is defined using a choice of t, but we want our value of $\sum\limits_{s\in S} a_s$ to depend only on S and $(a_s)_{s\in S}$ (not on some arbitrarily chosen $t \in S$). Thus, we cannot use this definition yet. Our main goal in this section is to prove that it is indeed legitimate.

2.14.4. The set of all possible sums

There are two ways to approach this goal. One is to prove the legitimacy of Definition 2.111 by strong induction on $|S|$; the statement $\mathcal{A}(n)$ that we would be proving for each $n \in \mathbb{N}$ here would be saying that Definition 2.111 is legitimate for all finite sets S satisfying $|S| = n$. This is not hard, but conceptually confusing, as it would require us to use Definition 2.111 for **some** sets S while its legitimacy for other sets S is yet unproven.

We prefer to proceed in a different way: We shall first define a set $\operatorname{Sums}\left((a_s)_{s\in S}\right)$ for any $\mathbb{A}$-valued S-family $(a_s)_{s\in S}$; this set shall consist (roughly speaking) of "all possible values that $\sum\limits_{s\in S} a_s$ could have according to Definition 2.111". This set will be defined recursively, more or less following Definition 2.111, but instead of relying on a choice of **some** $t \in S$, it will use **all** possible elements $t \in S$. (See Definition 2.112 for the precise definition.) Unlike $\sum\limits_{s\in S} a_s$ itself, it will be a set of numbers, not a single number; however, it has the advantage that the legitimacy of its definition will be immediately obvious. Then, we will prove (in Theorem 2.114) that this set $\operatorname{Sums}\left((a_s)_{s\in S}\right)$ is actually a 1-element set; this will allow us to define $\sum\limits_{s\in S} a_s$ to be the unique element of $\operatorname{Sums}\left((a_s)_{s\in S}\right)$ for any $\mathbb{A}$-valued S-family $(a_s)_{s\in S}$ (see Definition 2.116). Then, we will retroactively legitimize Definition 2.111 by showing that Definition 2.111 leads to the same value of $\sum\limits_{s\in S} a_s$ as Definition 2.116 (no matter which $t \in S$ is chosen). Having thus justified Definition 2.111, we will forget about the set $\operatorname{Sums}\left((a_s)_{s\in S}\right)$ and about Definition 2.116.

In later subsections, we shall prove some basic properties of sums.

Let us define the set $\operatorname{Sums}\left((a_s)_{s\in S}\right)$, as promised:

Definition 2.112. If S is a finite set, and if $(a_s)_{s\in S}$ is an $\mathbb{A}$-valued S-family, then we want to define the set $\operatorname{Sums}\left((a_s)_{s\in S}\right)$ of numbers. We define this set by recursion on $|S|$ as follows:

- If $|S| = 0$, then $\operatorname{Sums}\left((a_s)_{s\in S}\right)$ is defined to be $\{0\}$.
- Let $n \in \mathbb{N}$. Assume that we have defined $\operatorname{Sums}\left((a_s)_{s\in S}\right)$ for every finite set S with $|S| = n$ and any $\mathbb{A}$-valued S-family $(a_s)_{s\in S}$. Now, if S is a finite set with $|S| = n+1$, and if $(a_s)_{s\in S}$ is any $\mathbb{A}$-valued S-family, then $\operatorname{Sums}\left((a_s)_{s\in S}\right)$ is defined by

$$\begin{aligned}&\operatorname{Sums}\left((a_s)_{s\in S}\right)\\&= \left\{a_t + b \mid t \in S \text{ and } b \in \operatorname{Sums}\left((a_s)_{s\in S\setminus\{t\}}\right)\right\}.\end{aligned} \tag{162}$$

  (The sets $\operatorname{Sums}\left((a_s)_{s\in S\setminus\{t\}}\right)$ on the right hand side of this equation are well-defined, because for each $t \in S$, we have $|S \setminus \{t\}| = |S| - 1 = n$ (since $|S| = n+1$), and therefore $\operatorname{Sums}\left((a_s)_{s\in S\setminus\{t\}}\right)$ is well-defined by our assumption.)

Example 2.113. Let S be a finite set. Let $(a_s)_{s\in S}$ be an $\mathbb{A}$-valued S-family. Let us see what Definition 2.112 says when S has only few elements:

(a) If $S = \varnothing$, then

$$\operatorname{Sums}\left((a_s)_{s\in\varnothing}\right) = \{0\} \tag{163}$$

(directly by Definition 2.112, since $|S| = |\varnothing| = 0$ in this case).

(b) If $S = \{x\}$ for some element x, then Definition 2.112 yields

$$\begin{aligned}&\operatorname{Sums}\left((a_s)_{s\in\{x\}}\right)\\&= \left\{a_t + b \mid t \in \{x\} \text{ and } b \in \operatorname{Sums}\left((a_s)_{s\in\{x\}\setminus\{t\}}\right)\right\}\\&= \left\{a_x + b \mid b \in \operatorname{Sums}\left((a_s)_{s\in\{x\}\setminus\{x\}}\right)\right\} \qquad (\text{since the only } t \in \{x\} \text{ is } x)\\&= \left\{a_x + b \mid b \in \underbrace{\operatorname{Sums}\left((a_s)_{s\in\varnothing}\right)}_{=\{0\}}\right\} \qquad (\text{since } \{x\} \setminus \{x\} = \varnothing)\\&= \{a_x + b \mid b \in \{0\}\} = \{a_x + 0\} = \{a_x\}.\end{aligned} \tag{164}$$

(c) If $S = \{x, y\}$ for two distinct elements x and y, then Definition 2.112 yields

$$\begin{aligned}
&\operatorname{Sums}\left(\left(a_s\right)_{s\in\{x,y\}}\right) \\
&= \left\{a_t + b \mid t \in \{x,y\} \text{ and } b \in \operatorname{Sums}\left(\left(a_s\right)_{s\in\{x,y\}\setminus\{t\}}\right)\right\} \\
&= \left\{a_x + b \mid b \in \operatorname{Sums}\left(\left(a_s\right)_{s\in\{x,y\}\setminus\{x\}}\right)\right\} \\
&\qquad \cup \left\{a_y + b \mid b \in \operatorname{Sums}\left(\left(a_s\right)_{s\in\{x,y\}\setminus\{y\}}\right)\right\} \\
&= \left\{a_x + b \mid b \in \underbrace{\operatorname{Sums}\left(\left(a_s\right)_{s\in\{y\}}\right)}_{\substack{=\{a_y\} \\ \text{(by (164), applied to } y \text{ instead of } x)}}\right\} \\
&\qquad \cup \left\{a_y + b \mid b \in \underbrace{\operatorname{Sums}\left(\left(a_s\right)_{s\in\{x\}}\right)}_{\substack{=\{a_x\} \\ \text{(by (164))}}}\right\} \\
&\qquad \left(\text{since } \{x,y\} \setminus \{x\} = \{y\} \text{ and } \{x,y\} \setminus \{y\} = \{x\}\right) \\
&= \underbrace{\left\{a_x + b \mid b \in \{a_y\}\right\}}_{=\{a_x+a_y\}} \cup \underbrace{\left\{a_y + b \mid b \in \{a_x\}\right\}}_{=\{a_y+a_x\}} \\
&= \left\{a_x + a_y\right\} \cup \left\{a_y + a_x\right\} = \left\{a_x + a_y, a_y + a_x\right\} = \left\{a_x + a_y\right\}
\end{aligned}$$

(since $a_y + a_x = a_x + a_y$).

(d) Similar reasoning shows that if $S = \{x, y, z\}$ for three distinct elements x, y and z, then

$$\operatorname{Sums}\left(\left(a_s\right)_{s\in\{x,y,z\}}\right) = \left\{a_x + \left(a_y + a_z\right), a_y + \left(a_x + a_z\right), a_z + \left(a_x + a_y\right)\right\}.$$

It is not hard to check (using Proposition 2.103 and Proposition 2.104) that the three elements $a_x + \left(a_y + a_z\right)$, $a_y + \left(a_x + a_z\right)$ and $a_z + \left(a_x + a_y\right)$ of this set are equal, so we may call them $a_x + a_y + a_z$; thus, we can rewrite this equality as

$$\operatorname{Sums}\left(\left(a_s\right)_{s\in\{x,y,z\}}\right) = \left\{a_x + a_y + a_z\right\}.$$

(e) Going further, we can see that if $S = \{x, y, z, w\}$ for four distinct elements x, y, z and w, then

$$\begin{aligned}
\operatorname{Sums}\left(\left(a_s\right)_{s\in\{x,y,z,w\}}\right) = \{&a_x + \left(a_y + a_z + a_w\right), a_y + \left(a_x + a_z + a_w\right), \\
&a_z + \left(a_x + a_y + a_w\right), a_w + \left(a_x + a_y + a_z\right)\}.
\end{aligned}$$

Again, it is not hard to prove that

$$\begin{aligned} a_x + \left(a_y + a_z + a_w\right) &= a_y + \left(a_x + a_z + a_w\right) \\ &= a_z + \left(a_x + a_y + a_w\right) = a_w + \left(a_x + a_y + a_z\right), \end{aligned}$$

and thus the set $\operatorname{Sums}\left(\left(a_s\right)_{s\in\{x,y,z,w\}}\right)$ is again a 1-element set, whose unique element can be called $a_x + a_y + a_z + a_w$.

These examples suggest that the set $\operatorname{Sums}\left(\left(a_s\right)_{s\in S}\right)$ should always be a 1-element set. This is precisely what we are going to claim now:

Theorem 2.114. If S is a finite set, and if $\left(a_s\right)_{s\in S}$ is an $\mathbb{A}$-valued S-family, then the set $\operatorname{Sums}\left(\left(a_s\right)_{s\in S}\right)$ is a 1-element set.

2.14.5. The set of all possible sums is a 1-element set: proof

Before we step to the proof of Theorem 2.114, we observe an almost trivial lemma:

Lemma 2.115. Let a, b and c be three numbers (i.e., elements of $\mathbb{A}$). Then, $a + (b+c) = b + (a+c)$.

Proof of Lemma 2.115. Proposition 2.103 (applied to b and a instead of a and b) yields $(b+a)+c = b+(a+c)$. Also, Proposition 2.103 yields $(a+b)+c = a+(b+c)$. Hence,

$$a + (b+c) = \underbrace{(a+b)}_{\substack{=b+a \\ \text{(by Proposition 2.104)}}} + c = (b+a)+c = b+(a+c).$$

This proves Lemma 2.115. □

Proof of Theorem 2.114. We shall prove Theorem 2.114 by strong induction on $|S|$:

Let $m \in \mathbb{N}$. Assume that Theorem 2.114 holds under the condition that $|S| < m$. We must now prove that Theorem 2.114 holds under the condition that $|S| = m$.

We have assumed that Theorem 2.114 holds under the condition that $|S| < m$. In other words, the following claim holds:

> *Claim 1:* Let S be a finite set satisfying $|S| < m$. Let $\left(a_s\right)_{s\in S}$ be an $\mathbb{A}$-valued S-family. Then, the set $\operatorname{Sums}\left(\left(a_s\right)_{s\in S}\right)$ is a 1-element set.

Now, we must now prove that Theorem 2.114 holds under the condition that $|S| = m$. In other words, we must prove the following claim:

> *Claim 2:* Let S be a finite set satisfying $|S| = m$. Let $\left(a_s\right)_{s\in S}$ be an $\mathbb{A}$-valued S-family. Then, the set $\operatorname{Sums}\left(\left(a_s\right)_{s\in S}\right)$ is a 1-element set.

Before we start proving Claim 2, let us prove two auxiliary claims:

> *Claim 3:* Let S be a finite set satisfying $|S| < m$. Let $(a_s)_{s\in S}$ be an $\mathbb{A}$-valued S-family. Let $r \in S$. Let $g \in \operatorname{Sums}\left((a_s)_{s\in S}\right)$ and $c \in \operatorname{Sums}\left((a_s)_{s\in S\setminus\{r\}}\right)$. Then, $g = a_r + c$.

[*Proof of Claim 3:* The set $S \setminus \{r\}$ is a subset of the finite set S, and thus itself is finite. Moreover, $r \in S$, so that $|S \setminus \{r\}| = |S| - 1$. Thus, $|S| = |S \setminus \{r\}| + 1$. Hence, the definition of $\operatorname{Sums}\left((a_s)_{s\in S}\right)$ yields

$$\operatorname{Sums}\left((a_s)_{s\in S}\right) = \left\{a_t + b \mid t \in S \text{ and } b \in \operatorname{Sums}\left((a_s)_{s\in S\setminus\{t\}}\right)\right\}. \tag{165}$$

But recall that $r \in S$ and $c \in \operatorname{Sums}\left((a_s)_{s\in S\setminus\{r\}}\right)$. Hence, the number $a_r + c$ has the form $a_t + b$ for some $t \in S$ and $b \in \operatorname{Sums}\left((a_s)_{s\in S\setminus\{t\}}\right)$ (namely, for $t = r$ and $b = c$). In other words,

$$a_r + c \in \left\{a_t + b \mid t \in S \text{ and } b \in \operatorname{Sums}\left((a_s)_{s\in S\setminus\{t\}}\right)\right\}.$$

In view of (165), this rewrites as $a_r + c \in \operatorname{Sums}\left((a_s)_{s\in S}\right)$.

But Claim 1 shows that the set $\operatorname{Sums}\left((a_s)_{s\in S}\right)$ is a 1-element set. Hence, any two elements of $\operatorname{Sums}\left((a_s)_{s\in S}\right)$ are equal. In other words, any $x \in \operatorname{Sums}\left((a_s)_{s\in S}\right)$ and $y \in \operatorname{Sums}\left((a_s)_{s\in S}\right)$ satisfy $x = y$. Applying this to $x = g$ and $y = a_r + c$, we obtain $g = a_r + c$ (since $g \in \operatorname{Sums}\left((a_s)_{s\in S}\right)$ and $a_r + c \in \operatorname{Sums}\left((a_s)_{s\in S}\right)$). This proves Claim 3.]

> *Claim 4:* Let S be a finite set satisfying $|S| = m$. Let $(a_s)_{s\in S}$ be an $\mathbb{A}$-valued S-family. Let $p \in S$ and $q \in S$. Let $f \in \operatorname{Sums}\left((a_s)_{s\in S\setminus\{p\}}\right)$ and $g \in \operatorname{Sums}\left((a_s)_{s\in S\setminus\{q\}}\right)$. Then, $a_p + f = a_q + g$.

[*Proof of Claim 4:* We have $p \in S$, and thus $|S \setminus \{p\}| = |S| - 1 < |S| = m$. Hence, Claim 1 (applied to $S \setminus \{p\}$ instead of S) yields that the set $\operatorname{Sums}\left((a_s)_{s\in S\setminus\{p\}}\right)$ is a 1-element set. In other words, $\operatorname{Sums}\left((a_s)_{s\in S\setminus\{p\}}\right)$ can be written in the form $\operatorname{Sums}\left((a_s)_{s\in S\setminus\{p\}}\right) = \{h\}$ for some number h. Consider this h.

We are in one of the following two cases:

Case 1: We have $p = q$.

Case 2: We have $p \neq q$.

Let us first consider Case 1. In this case, we have $p = q$. Hence, $q = p$, so that $a_q = a_p$.

We have $f \in \operatorname{Sums}\left(\left(a_s\right)_{s\in S\setminus\{p\}}\right) = \{h\}$, so that $f = h$. Also,

$$\begin{aligned} g &\in \operatorname{Sums}\left(\left(a_s\right)_{s\in S\setminus\{q\}}\right) = \operatorname{Sums}\left(\left(a_s\right)_{s\in S\setminus\{p\}}\right) \qquad (\text{since } q = p) \\ &= \{h\}, \end{aligned}$$

so that $g = h$. Comparing $a_p + \underbrace{f}_{=h} = a_p + h$ with $\underbrace{a_q}_{=a_p} + \underbrace{g}_{=h} = a_p + h$, we obtain $a_p + f = a_q + g$. Hence, Claim 4 is proven in Case 1.

Let us now consider Case 2. In this case, we have $p \neq q$. Thus, $q \neq p$, so that $q \notin \{p\}$.

We have $S \setminus \{p, q\} \subseteq S \setminus \{p\}$ (since $\{p\} \subseteq \{p, q\}$) and thus $|S \setminus \{p, q\}| \leq |S \setminus \{p\}| < m$. Hence, Claim 1 (applied to $S \setminus \{p, q\}$ instead of S) shows that the set $\operatorname{Sums}\left(\left(a_s\right)_{s\in S\setminus\{p,q\}}\right)$ is a 1-element set. In other words, $\operatorname{Sums}\left(\left(a_s\right)_{s\in S\setminus\{p,q\}}\right)$ has the form

$$\operatorname{Sums}\left(\left(a_s\right)_{s\in S\setminus\{p,q\}}\right) = \{c\}$$

for some number c. Consider this c. Hence,

$$c \in \{c\} = \operatorname{Sums}\left(\left(a_s\right)_{s\in S\setminus\{p,q\}}\right) = \operatorname{Sums}\left(\left(a_s\right)_{s\in (S\setminus\{p\})\setminus\{q\}}\right)$$

(since $S \setminus \{p, q\} = (S \setminus \{p\}) \setminus \{q\}$). Also, $q \in S \setminus \{p\}$ (since $q \in S$ and $q \notin \{p\}$). Thus, Claim 3 (applied to $S \setminus \{p\}$, q and f instead of S, r and g) yields $f = a_q + c$ (since $|S \setminus \{p\}| < m$ and $f \in \operatorname{Sums}\left(\left(a_s\right)_{s\in S\setminus\{p\}}\right)$ and $c \in \operatorname{Sums}\left(\left(a_s\right)_{s\in (S\setminus\{p\})\setminus\{q\}}\right)$).

The same argument (but with p, q, f and g replaced by q, p, g and f) yields $g = a_p + c$.

Now,

$$a_p + \underbrace{f}_{=a_q+c} = a_p + \left(a_q + c\right) = a_q + \left(a_p + c\right)$$

(by Lemma 2.115, applied to $a = a_p$ and $b = a_q$). Comparing this with

$$a_q + \underbrace{g}_{=a_p+c} = a_q + \left(a_p + c\right),$$

we obtain $a_p + f = a_q + g$. Thus, Claim 4 is proven in Case 2.

We have now proven Claim 4 in both Cases 1 and 2. Since these two Cases cover all possibilities, we thus conclude that Claim 4 always holds.]

We can now prove Claim 2:

[*Proof of Claim 2:* If $|S| = 0$, then Claim 2 holds[79]. Hence, for the rest of this proof of Claim 2, we can WLOG assume that we don't have $|S| = 0$. Assume this.

[79] *Proof.* Assume that $|S| = 0$. Hence, Definition 2.112 yields $\operatorname{Sums}\left(\left(a_s\right)_{s\in S}\right) = \{0\}$. Hence, the set $\operatorname{Sums}\left(\left(a_s\right)_{s\in S}\right)$ is a 1-element set (since the set $\{0\}$ is a 1-element set). In other words, Claim 2 holds. Qed.

We have $|S| \neq 0$ (since we don't have $|S| = 0$). Hence, $|S|$ is a positive integer. Thus, $|S| - 1 \in \mathbb{N}$. Also, the set S is nonempty (since $|S| \neq 0$). Hence, there exists some $p \in S$. Consider this p.

We have $p \in S$ and thus $|S \setminus \{p\}| = |S| - 1 < |S| = m$. Hence, Claim 1 (applied to $S \setminus \{p\}$ instead of S) shows that the set $\operatorname{Sums}\left((a_s)_{s \in S \setminus \{p\}}\right)$ is a 1-element set. In other words, $\operatorname{Sums}\left((a_s)_{s \in S \setminus \{p\}}\right)$ has the form

$$\operatorname{Sums}\left((a_s)_{s \in S \setminus \{p\}}\right) = \{f\}$$

for some number f. Consider this f. Thus,

$$f \in \{f\} = \operatorname{Sums}\left((a_s)_{s \in S \setminus \{p\}}\right). \tag{166}$$

Define $n \in \mathbb{N}$ by $n = |S| - 1$. (This is allowed, since $|S| - 1 \in \mathbb{N}$.) Then, from $n = |S| - 1$, we obtain $|S| = n + 1$. Hence, the definition of $\operatorname{Sums}\left((a_s)_{s \in S}\right)$ yields

$$\operatorname{Sums}\left((a_s)_{s \in S}\right) = \left\{a_t + b \mid t \in S \text{ and } b \in \operatorname{Sums}\left((a_s)_{s \in S \setminus \{t\}}\right)\right\}. \tag{167}$$

But recall that $p \in S$ and $f \in \operatorname{Sums}\left((a_s)_{s \in S \setminus \{p\}}\right)$. Hence, the number $a_p + f$ has the form $a_t + b$ for some $t \in S$ and $b \in \operatorname{Sums}\left((a_s)_{s \in S \setminus \{t\}}\right)$ (namely, for $t = p$ and $b = f$). In other words,

$$a_p + f \in \left\{a_t + b \mid t \in S \text{ and } b \in \operatorname{Sums}\left((a_s)_{s \in S \setminus \{t\}}\right)\right\}.$$

In view of (167), this rewrites as

$$a_p + f \in \operatorname{Sums}\left((a_s)_{s \in S}\right).$$

Thus,

$$\left\{a_p + f\right\} \subseteq \operatorname{Sums}\left((a_s)_{s \in S}\right). \tag{168}$$

Next, we shall show the reverse inclusion (i.e., we shall show that $\operatorname{Sums}\left((a_s)_{s \in S}\right) \subseteq \left\{a_p + f\right\}$).

Indeed, let $w \in \operatorname{Sums}\left((a_s)_{s \in S}\right)$. Thus,

$$w \in \operatorname{Sums}\left((a_s)_{s \in S}\right) = \left\{a_t + b \mid t \in S \text{ and } b \in \operatorname{Sums}\left((a_s)_{s \in S \setminus \{t\}}\right)\right\}$$

(by (167)). In other words, w can be written as $w = a_t + b$ for some $t \in S$ and $b \in \operatorname{Sums}\left((a_s)_{s \in S \setminus \{t\}}\right)$. Consider these t and b, and denote them by q and g. Thus, $q \in S$ and $g \in \operatorname{Sums}\left((a_s)_{s \in S \setminus \{q\}}\right)$ satisfy $w = a_q + g$.

But Claim 4 yields $a_p + f = a_q + g$. Comparing this with $w = a_q + g$, we obtain $w = a_p + f$. Thus, $w \in \left\{a_p + f\right\}$.

Now, forget that we fixed w. We thus have proven that $w \in \left\{a_p + f\right\}$ for each $w \in \operatorname{Sums}\left(\left(a_s\right)_{s\in S}\right)$. In other words, $\operatorname{Sums}\left(\left(a_s\right)_{s\in S}\right) \subseteq \left\{a_p + f\right\}$. Combining this with (168), we conclude that $\operatorname{Sums}\left(\left(a_s\right)_{s\in S}\right) = \left\{a_p + f\right\}$. Hence, the set $\operatorname{Sums}\left(\left(a_s\right)_{s\in S}\right)$ is a 1-element set. This proves Claim 2.]

Now, we have proven Claim 2. But Claim 2 says precisely that Theorem 2.114 holds under the condition that $|S| = m$. Hence, we have proven that Theorem 2.114 holds under the condition that $|S| = m$. This completes the induction step. Thus, Theorem 2.114 is proven by strong induction. □

2.14.6. Sums of numbers are well-defined

We can now give a new definition of the sum $\sum\limits_{s\in S} a_s$ (for any finite set S and any $\mathbb{A}$-valued S-family $\left(a_s\right)_{s\in S}$), which is different from Definition 2.111 and (unlike the latter) is clearly legitimate:

Definition 2.116. Let S be a finite set, and let $\left(a_s\right)_{s\in S}$ be an $\mathbb{A}$-valued S-family. Then, the set $\operatorname{Sums}\left(\left(a_s\right)_{s\in S}\right)$ is a 1-element set (by Theorem 2.114). We define $\sum\limits_{s\in S} a_s$ to be the unique element of this set $\operatorname{Sums}\left(\left(a_s\right)_{s\in S}\right)$.

However, we have not reached our goal yet: After all, we wanted to prove that Definition 2.111 is legitimate, rather than replace it by a new definition!

Fortunately, we are very close to achieving this goal (after having done all the hard work in the proof of Theorem 2.114 above); we are soon going to show that Definition 2.111 is justified **and** that it is equivalent to Definition 2.116 (that is, both definitions yield the same value of $\sum\limits_{s\in S} a_s$). First, we need a simple lemma, which says that the notation $\sum\limits_{s\in S} a_s$ defined in Definition 2.116 "behaves" like the one defined in Definition 2.111:

Lemma 2.117. In this lemma, we shall use Definition 2.116 (not Definition 2.111).

Let S be a finite set, and let $\left(a_s\right)_{s\in S}$ be an $\mathbb{A}$-valued S-family.

(a) If $|S| = 0$, then

$$\sum_{s\in S} a_s = 0.$$

(b) For any $t \in S$, we have

$$\sum_{s\in S} a_s = a_t + \sum_{s\in S\setminus\{t\}} a_s.$$

Proof of Lemma 2.117. **(a)** Assume that $|S| = 0$. Thus, $\operatorname{Sums}\left(\left(a_s\right)_{s\in S}\right) = \{0\}$ (by the definition of $\operatorname{Sums}\left(\left(a_s\right)_{s\in S}\right)$). Hence, the unique element of the set $\operatorname{Sums}\left(\left(a_s\right)_{s\in S}\right)$ is 0.

But Definition 2.116 yields that $\sum_{s\in S} a_s$ is the unique element of the set Sums $\left((a_s)_{s\in S}\right)$. Thus, $\sum_{s\in S} a_s$ is 0 (since the unique element of the set Sums $\left((a_s)_{s\in S}\right)$ is 0). In other words, $\sum_{s\in S} a_s = 0$. This proves Lemma 2.117 **(a)**.

(b) Let $p \in S$. Thus, $|S \setminus \{p\}| = |S| - 1$, so that $|S| = |S \setminus \{p\}| + 1$. Hence, the definition of Sums $\left((a_s)_{s\in S}\right)$ yields

$$\operatorname{Sums}\left((a_s)_{s\in S}\right) = \left\{a_t + b \mid t \in S \text{ and } b \in \operatorname{Sums}\left((a_s)_{s\in S\setminus\{t\}}\right)\right\}. \tag{169}$$

Definition 2.116 yields that $\sum_{s\in S\setminus\{p\}} a_s$ is the unique element of the set Sums $\left((a_s)_{s\in S\setminus\{p\}}\right)$. Thus, $\sum_{s\in S\setminus\{p\}} a_s \in \operatorname{Sums}\left((a_s)_{s\in S\setminus\{p\}}\right)$. Thus, $a_p + \sum_{s\in S\setminus\{p\}} a_s$ is a number of the form $a_t + b$ for some $t \in S$ and some $b \in \operatorname{Sums}\left((a_s)_{s\in S\setminus\{t\}}\right)$ (namely, for $t = p$ and $b = \sum_{s\in S\setminus\{p\}} a_s$). In other words,

$$a_p + \sum_{s\in S\setminus\{p\}} a_s \in \left\{a_t + b \mid t \in S \text{ and } b \in \operatorname{Sums}\left((a_s)_{s\in S\setminus\{t\}}\right)\right\}.$$

In view of (169), this rewrites as

$$a_p + \sum_{s\in S\setminus\{p\}} a_s \in \operatorname{Sums}\left((a_s)_{s\in S}\right). \tag{170}$$

But Definition 2.116 yields that $\sum_{s\in S} a_s$ is the unique element of the set Sums $\left((a_s)_{s\in S}\right)$. Hence, the set Sums $\left((a_s)_{s\in S}\right)$ consists only of the element $\sum_{s\in S} a_s$. In other words,

$$\operatorname{Sums}\left((a_s)_{s\in S}\right) = \left\{\sum_{s\in S} a_s\right\}.$$

Thus, (170) rewrites as

$$a_p + \sum_{s\in S\setminus\{p\}} a_s \in \left\{\sum_{s\in S} a_s\right\}.$$

In other words, $a_p + \sum_{s\in S\setminus\{p\}} a_s = \sum_{s\in S} a_s$. Thus, $\sum_{s\in S} a_s = a_p + \sum_{s\in S\setminus\{p\}} a_s$.

Now, forget that we fixed p. We thus have proven that for any $p \in S$, we have $\sum_{s\in S} a_s = a_p + \sum_{s\in S\setminus\{p\}} a_s$. Renaming the variable p as t in this statement, we obtain the following: For any $t \in S$, we have $\sum_{s\in S} a_s = a_t + \sum_{s\in S\setminus\{t\}} a_s$. This proves Lemma 2.117 **(b)**. □

We can now finally state what we wanted to state:

Theorem 2.118. **(a)** Definition 2.111 is legitimate: i.e., the value of $\sum\limits_{s\in S} a_s$ in Definition 2.111 does not depend on the choice of t.

(b) Definition 2.111 is equivalent to Definition 2.116: i.e., both of these definitions yield the same value of $\sum\limits_{s\in S} a_s$.

It makes sense to call Theorem 2.118 **(a)** the *general commutativity theorem*, as it says that a sum of n numbers can be computed in an arbitrary order.

Proof of Theorem 2.118. Let us first use Definition 2.116 (not Definition 2.111). Then, for any finite set S and any $\mathbb{A}$-valued S-family $(a_s)_{s\in S}$, we can compute the number $\sum\limits_{s\in S} a_s$ by the following algorithm (which uses recursion on $|S|$):

- If $|S| = 0$, then $\sum\limits_{s\in S} a_s = 0$. (This follows from Lemma 2.117 **(a)**.)
- Otherwise, we have $|S| = n+1$ for some $n \in \mathbb{N}$. Consider this n. Thus, $|S| = n+1 \geq 1 > 0$, so that the set S is nonempty. Fix any $t \in S$. (Such a t exists, since the set S is nonempty.) We have $|S \setminus \{t\}| = |S| - 1 = n$ (since $|S| = n+1$), so that we can assume (because we are using recursion) that $\sum\limits_{s\in S\setminus\{t\}} a_s$ has already been computed. Then, $\sum\limits_{s\in S} a_s = a_t + \sum\limits_{s\in S\setminus\{t\}} a_s$. (This follows from Lemma 2.117 **(b)**.)

We can restate this algorithm as an alternative definition of $\sum\limits_{s\in S} a_s$; it then takes the following form:

> *Alternative definition of $\sum\limits_{s\in S} a_s$ for any finite set S and any $\mathbb{A}$-valued S-family $(a_s)_{s\in S}$:* If S is a finite set, and if $(a_s)_{s\in S}$ is an $\mathbb{A}$-valued S-family, then we define $\sum\limits_{s\in S} a_s$ by recursion on $|S|$ as follows:
>
> - If $|S| = 0$, then $\sum\limits_{s\in S} a_s$ is defined to be 0.
> - Let $n \in \mathbb{N}$. Assume that we have defined $\sum\limits_{s\in S} a_s$ for every finite set S with $|S| = n$ and any $\mathbb{A}$-valued S-family $(a_s)_{s\in S}$. Now, if S is a finite set with $|S| = n+1$, and if $(a_s)_{s\in S}$ is any $\mathbb{A}$-valued S-family, then $\sum\limits_{s\in S} a_s$ is defined by picking any $t \in S$ and setting
>
> $$\sum_{s\in S} a_s = a_t + \sum_{s\in S\setminus\{t\}} a_s. \tag{171}$$

This alternative definition of $\sum\limits_{s\in S} a_s$ merely follows the above algorithm for computing $\sum\limits_{s\in S} a_s$. Thus, it is guaranteed to always yield the same value of $\sum\limits_{s\in S} a_s$ as Definition 2.116, independently of the choice of t. Hence, we obtain the following:

> *Claim 1:* This alternative definition is legitimate (i.e., the value of $\sum\limits_{s\in S} a_s$ in (171) does not depend on the choice of t), and is equivalent to Definition 2.116.

But on the other hand, this alternative definition is precisely Definition 2.111. Hence, Claim 1 rewrites as follows: Definition 2.111 is legitimate (i.e., the value of $\sum\limits_{s\in S} a_s$ in Definition 2.111 does not depend on the choice of t), and is equivalent to Definition 2.116. This proves both parts **(a)** and **(b)** of Theorem 2.118. □

Theorem 2.118 **(a)** shows that Definition 2.111 is legitimate.

Thus, at last, we have vindicated the notation $\sum\limits_{s\in S} a_s$ that was introduced in Section 1.4 (because the definition of this notation we gave in Section 1.4 was precisely Definition 2.111). We can now forget about Definition 2.116, since it has served its purpose (which was to justify Definition 2.111). (Of course, we could also forget about Definition 2.111 instead, and use Definition 2.116 as our definition of $\sum\limits_{s\in S} a_s$ (after all, these two definitions are equivalent, as we now know). Then, we would have to replace every reference to the definition of $\sum\limits_{s\in S} a_s$ by a reference to Lemma 2.117; in particular, we would have to replace every use of (1) by a use of Lemma 2.117 **(b)**. Other than this, everything would work the same way.)

The notation $\sum\limits_{s\in S} a_s$ has several properties, many of which were collected in Section 1.4. We shall prove some of these properties later in this section.

From now on, we shall be using all the conventions and notations regarding sums that we introduced in Section 1.4. In particular, expressions of the form "$\sum\limits_{s\in S} a_s + b$" shall always be interpreted as $\left(\sum\limits_{s\in S} a_s\right) + b$, not as $\sum\limits_{s\in S} (a_s + b)$; but expressions of the form "$\sum\limits_{s\in S} ba_sc$" shall always be understood to mean $\sum\limits_{s\in S} (ba_sc)$.

2.14.7. Triangular numbers revisited

Recall one specific notation we introduced in Section 1.4: If u and v are two integers, and if a_s is a number for each $s \in \{u, u+1, \ldots, v\}$, then $\sum\limits_{s=u}^{v} a_s$ is defined by

$$\sum_{s=u}^{v} a_s = \sum_{s\in\{u,u+1,\ldots,v\}} a_s.$$

This sum $\sum\limits_{s=u}^{v} a_s$ is also denoted by $a_u + a_{u+1} + \cdots + a_v$.

We are now ready to do something that we evaded in Section 2.4: namely, to speak of the sum of the first n positive integers without having to define it recursively. Indeed, we can now interpret this sum as $\sum\limits_{i\in\{1,2,\ldots,n\}} i$, an expression which has a well-defined meaning because we have shown that the notation $\sum_{s\in S} a_s$ is well-defined. We can also rewrite this expression as $\sum_{i=1}^{n} i$ or as $1+2+\cdots+n$.

Thus, the classical fact that the sum of the first n positive integers is $\frac{n(n+1)}{2}$ can now be stated as follows:

Proposition 2.119. We have

$$\sum_{i\in\{1,2,\ldots,n\}} i = \frac{n(n+1)}{2} \qquad \text{for each } n \in \mathbb{N}. \tag{172}$$

Proof of Proposition 2.119. We shall prove (172) by induction on n:

Induction base: We have $\{1,2,\ldots,0\} = \varnothing$ and thus $|\{1,2,\ldots,0\}| = |\varnothing| = 0$. Hence, the definition of $\sum_{i\in\{1,2,\ldots,0\}} i$ yields

$$\sum_{i\in\{1,2,\ldots,0\}} i = 0. \tag{173}$$

(To be more precise, we have used the first bullet point of Definition 2.111 here, which says that $\sum_{s\in S} a_s = 0$ whenever the set S and the $\mathbb{A}$-valued S-family $(a_s)_{s\in S}$ satisfy $|S| = 0$. If you are using Definition 2.116 instead of Definition 2.111, you should instead be using Lemma 2.117 **(a)** to argue this.)

Comparing (173) with $\frac{0(0+1)}{2} = 0$, we obtain $\sum_{i\in\{1,2,\ldots,0\}} i = \frac{0(0+1)}{2}$. In other words, (172) holds for $n = 0$. This completes the induction base.

Induction step: Let $m \in \mathbb{N}$. Assume that (172) holds for $n = m$. We must prove that (172) holds for $n = m+1$.

We have assumed that (172) holds for $n = m$. In other words, we have

$$\sum_{i\in\{1,2,\ldots,m\}} i = \frac{m(m+1)}{2}. \tag{174}$$

Now, $|\{1,2,\ldots,m+1\}| = m+1$ and $m+1 \in \{1,2,\ldots,m+1\}$ (since $m+1$ is a positive integer (since $m \in \mathbb{N}$)). Hence, (1) (applied to $n = m$, $S = \{1,2,\ldots,m+1\}$, $t = m+1$ and $(a_s)_{s\in S} = (i)_{i\in\{1,2,\ldots,m+1\}}$) yields

$$\sum_{i\in\{1,2,\ldots,m+1\}} i = (m+1) + \sum_{i\in\{1,2,\ldots,m+1\}\setminus\{m+1\}} i. \tag{175}$$

(Here, we have relied on the equality (1), which appears verbatim in Definition 2.111. If you are using Definition 2.116 instead of Definition 2.111, you should instead be using Lemma 2.117 **(b)** to argue this.)

Now, (175) becomes

$$\begin{aligned}\sum_{i\in\{1,2,\ldots,m+1\}} i &= (m+1) + \sum_{i\in\{1,2,\ldots,m+1\}\setminus\{m+1\}} i = (m+1) + \underbrace{\sum_{i\in\{1,2,\ldots,m\}} i}_{\substack{=\dfrac{m(m+1)}{2} \\ \text{(by (174))}}} \\ &\qquad \left(\text{since } \{1,2,\ldots,m+1\}\setminus\{m+1\} = \{1,2,\ldots,m\}\right) \\ &= (m+1) + \frac{m(m+1)}{2} = \frac{2(m+1)+m(m+1)}{2} \\ &= \frac{(m+1)((m+1)+1)}{2}\end{aligned}$$

(since $2(m+1)+m(m+1) = (m+1)((m+1)+1)$). In other words, (172) holds for $n = m+1$. This completes the induction step. Thus, the induction proof of (172) is finished. Hence, Proposition 2.119 holds. □

2.14.8. Sums of a few numbers

Merely for the sake of future convenience, let us restate (1) in a slightly more direct way (without mentioning $|S|$):

Proposition 2.120. Let S be a finite set, and let $(a_s)_{s\in S}$ be an $\mathbb{A}$-valued S-family. Let $t \in S$. Then,

$$\sum_{s\in S} a_s = a_t + \sum_{s\in S\setminus\{t\}} a_s.$$

Proof of Proposition 2.120. Let $n = |S\setminus\{t\}|$; thus, $n \in \mathbb{N}$ (since $S\setminus\{t\}$ is a finite set). Also, $n = |S\setminus\{t\}| = |S|-1$ (since $t \in S$), and thus $|S| = n+1$. Hence, (1) yields $\sum\limits_{s\in S} a_s = a_t + \sum\limits_{s\in S\setminus\{t\}} a_s$. This proves Proposition 2.120. □

(Alternatively, we can argue that Proposition 2.120 is the same as Lemma 2.117 **(b)**, except that we are now using Definition 2.111 instead of Definition 2.116 to define the sums involved – but this difference is insubstantial, since we have shown that these two definitions are equivalent.)

In Section 1.4, we have introduced $a_u + a_{u+1} + \cdots + a_v$ as an abbreviation for the sum $\sum\limits_{s=u}^{v} a_s = \sum\limits_{s\in\{u,u+1,\ldots,v\}} a_s$ (whenever u and v are two integers, and a_s is a number for each $s \in \{u, u+1, \ldots, v\}$). In order to ensure that this abbreviation does not create any nasty surprises, we need to check that it behaves as we would expect – i.e., that it satisfies the following four properties:

- If the sum $a_u + a_{u+1} + \cdots + a_v$ has no addends (i.e., if $u > v$), then it equals 0.

- If the sum $a_u + a_{u+1} + \cdots + a_v$ has exactly one addend (i.e., if $u = v$), then it equals a_u.

- If the sum $a_u + a_{u+1} + \cdots + a_v$ has exactly two addends (i.e., if $u = v - 1$), then it equals $a_u + a_v$.

- If $v \geq u$, then

$$\begin{aligned} a_u + a_{u+1} + \cdots + a_v &= \left(a_u + a_{u+1} + \cdots + a_{v-1}\right) + a_v \\ &= a_u + \left(a_{u+1} + a_{u+2} + \cdots + a_v\right). \end{aligned}$$

The first of these four properties follows from the definition (indeed, if $u > v$, then the set $\{u, u+1, \ldots, v\}$ is empty and thus satisfies $|\{u, u+1, \ldots, v\}| = 0$; but this yields $\sum\limits_{s \in \{u,u+1,\ldots,v\}} a_s = 0$). The fourth of these four properties can easily be obtained from Proposition 2.120[80]. The second and third properties follow from the following fact:

[80] In more detail: Assume that $v \geq u$. Thus, both u and v belong to the set $\{u, u+1, \ldots, v\}$. Hence, Proposition 2.120 (applied to $S = \{u, u+1, \ldots, v\}$ and $t = v$) yields $\sum\limits_{s \in \{u,u+1,\ldots,v\}} a_s = a_v + \sum\limits_{s \in \{u,u+1,\ldots,v\} \setminus \{v\}} a_s$. Thus,

$$\begin{aligned} & a_u + a_{u+1} + \cdots + a_v \\ &= \sum_{s \in \{u,u+1,\ldots,v\}} a_s = a_v + \sum_{s \in \{u,u+1,\ldots,v\} \setminus \{v\}} a_s \\ &= a_v + \sum_{s \in \{u,u+1,\ldots,v-1\}} a_s \qquad \left(\text{since } \{u, u+1, \ldots, v\} \setminus \{v\} = \{u, u+1, \ldots, v-1\}\right) \\ &= a_v + \left(a_u + a_{u+1} + \cdots + a_{v-1}\right) = \left(a_u + a_{u+1} + \cdots + a_{v-1}\right) + a_v. \end{aligned}$$

Also, Proposition 2.120 (applied to $S = \{u, u+1, \ldots, v\}$ and $t = u$) yields $\sum\limits_{s \in \{u,u+1,\ldots,v\}} a_s = a_u + \sum\limits_{s \in \{u,u+1,\ldots,v\} \setminus \{u\}} a_s$. Thus,

$$\begin{aligned} & a_u + a_{u+1} + \cdots + a_v \\ &= \sum_{s \in \{u,u+1,\ldots,v\}} a_s = a_u + \sum_{s \in \{u,u+1,\ldots,v\} \setminus \{u\}} a_s \\ &= a_u + \sum_{s \in \{u+1,u+2,\ldots,v\}} a_s \qquad \left(\text{since } \{u, u+1, \ldots, v\} \setminus \{u\} = \{u+1, u+2, \ldots, v\}\right) \\ &= a_u + \left(a_{u+1} + a_{u+2} + \cdots + a_v\right). \end{aligned}$$

Hence,

$$a_u + a_{u+1} + \cdots + a_v = \left(a_u + a_{u+1} + \cdots + a_{v-1}\right) + a_v = a_u + \left(a_{u+1} + a_{u+2} + \cdots + a_v\right).$$

Proposition 2.121. Let S be a finite set. For every $s \in S$, let a_s be an element of $\mathbb{A}$.

(a) If $S = \{p\}$ for some element p, then this p satisfies

$$\sum_{s\in S} a_s = a_p.$$

(b) If $S = \{p, q\}$ for two distinct elements p and q, then these p and q satisfy

$$\sum_{s\in S} a_s = a_p + a_q.$$

Proof of Proposition 2.121. **(a)** Assume that $S = \{p\}$ for some element p. Consider this p.

The first bullet point of Definition 2.111 shows that $\sum_{s\in\varnothing} a_s = 0$ (since $|\varnothing| = 0$). But $p \in \{p\} = S$. Hence, Proposition 2.120 (applied to $t = p$) yields

$$\sum_{s\in S} a_s = a_p + \sum_{s\in S\setminus\{p\}} a_s = a_p + \underbrace{\sum_{s\in\varnothing} a_s}_{=0} \qquad \left(\text{since } \underbrace{S}_{=\{p\}} \setminus \{p\} = \{p\} \setminus \{p\} = \varnothing\right)$$
$$= a_p + 0 = a_p.$$

This proves Proposition 2.121 **(a)**.

(b) Assume that $S = \{p, q\}$ for two distinct elements p and q. Consider these p and q. Thus, $q \neq p$ (since p and q are distinct), so that $q \notin \{p\}$.

Proposition 2.121 **(a)** (applied to $\{p\}$ instead of S) yields $\sum_{s\in\{p\}} a_s = a_p$ (since $\{p\} = \{p\}$).

We have $\underbrace{S}_{=\{p,q\}=\{p\}\cup\{q\}} \setminus \{q\} = (\{p\} \cup \{q\}) \setminus \{q\} = \{p\} \setminus \{q\} = \{p\}$ (since $q \notin \{p\}$). Also, $q \in \{p, q\} = S$. Hence, Proposition 2.120 (applied to $t = q$) yields

$$\sum_{s\in S} a_s = a_q + \sum_{s\in S\setminus\{q\}} a_s = a_q + \underbrace{\sum_{s\in\{p\}} a_s}_{=a_p} \qquad (\text{since } S \setminus \{q\} = \{p\})$$
$$= a_q + a_p = a_p + a_q.$$

This proves Proposition 2.121 **(b)**. □

2.14.9. Linearity of sums

We shall now prove some general properties of finite sums. We begin with the equality (7) from Section 1.4:

Theorem 2.122. Let S be a finite set. For every $s \in S$, let a_s and b_s be elements of $\mathbb{A}$. Then,

$$\sum_{s \in S} (a_s + b_s) = \sum_{s \in S} a_s + \sum_{s \in S} b_s.$$

Before we prove this theorem, let us show a simple lemma:

Lemma 2.123. Let x, y, u and v be four numbers (i.e., elements of $\mathbb{A}$). Then,

$$(x+y) + (u+v) = (x+u) + (y+v).$$

Proof of Lemma 2.123. Proposition 2.103 (applied to $a = y$, $b = u$ and $c = v$) yields

$$(y+u) + v = y + (u+v). \tag{176}$$

Also, Proposition 2.103 (applied to $a = x$, $b = y$ and $c = u + v$) yields

$$\begin{aligned}(x+y)+(u+v) &= x + \underbrace{(y+(u+v))}_{\substack{=(y+u)+v \\ \text{(by (176))}}} \\ &= x + ((y+u)+v). \end{aligned} \tag{177}$$

The same argument (with y and u replaced by u and y) yields

$$(x+u) + (y+v) = x + ((u+y)+v). \tag{178}$$

But Proposition 2.104 (applied to $a = y$ and $b = u$) yields $y + u = u + y$. Thus, (177) becomes

$$(x+y)+(u+v) = x + \left(\underbrace{(y+u)}_{=u+y} + v \right) = x + ((u+y)+v) = (x+u)+(y+v)$$

(by (178)). This proves Lemma 2.123. □

Proof of Theorem 2.122. Forget that we fixed S, a_s and b_s. We shall prove Theorem 2.122 by induction on $|S|$:

Induction base: Theorem 2.122 holds under the condition that $|S| = 0$ [81]. This completes the induction base.

[81] *Proof.* Let S, a_s and b_s be as in Theorem 2.122. Assume that $|S| = 0$. Thus, the first bullet point of Definition 2.111 yields $\sum_{s \in S} a_s = 0$ and $\sum_{s \in S} b_s = 0$ and $\sum_{s \in S} (a_s + b_s) = 0$. Hence,

$$\sum_{s \in S} (a_s + b_s) = 0 = \underbrace{0}_{=\sum_{s \in S} a_s} + \underbrace{0}_{=\sum_{s \in S} b_s} = \sum_{s \in S} a_s + \sum_{s \in S} b_s.$$

Now, forget that we fixed S, a_s and b_s. We thus have proved that if S, a_s and b_s are as in Theorem 2.122, and if $|S| = 0$, then $\sum_{s \in S} (a_s + b_s) = \sum_{s \in S} a_s + \sum_{s \in S} b_s$. In other words, Theorem 2.122 holds under the condition that $|S| = 0$. Qed.

Induction step: Let $m \in \mathbb{N}$. Assume that Theorem 2.122 holds under the condition that $|S| = m$. We must now prove that Theorem 2.122 holds under the condition that $|S| = m + 1$.

We have assumed that Theorem 2.122 holds under the condition that $|S| = m$. In other words, the following claim holds:

> *Claim 1:* Let S be a finite set such that $|S| = m$. For every $s \in S$, let a_s and b_s be elements of $\mathbb{A}$. Then,
>
> $$\sum_{s\in S} (a_s + b_s) = \sum_{s\in S} a_s + \sum_{s\in S} b_s.$$

Next, we shall show the following claim:

> *Claim 2:* Let S be a finite set such that $|S| = m + 1$. For every $s \in S$, let a_s and b_s be elements of $\mathbb{A}$. Then,
>
> $$\sum_{s\in S} (a_s + b_s) = \sum_{s\in S} a_s + \sum_{s\in S} b_s.$$

[*Proof of Claim 2:* We have $|S| = m + 1 > m \geq 0$. Hence, the set S is nonempty. Thus, there exists some $t \in S$. Consider this t.

From $t \in S$, we obtain $|S \setminus \{t\}| = |S| - 1 = m$ (since $|S| = m + 1$). Hence, Claim 1 (applied to $S \setminus \{t\}$ instead of S) yields

$$\sum_{s\in S\setminus\{t\}} (a_s + b_s) = \sum_{s\in S\setminus\{t\}} a_s + \sum_{s\in S\setminus\{t\}} b_s. \qquad (179)$$

Now, Proposition 2.120 yields

$$\sum_{s\in S} a_s = a_t + \sum_{s\in S\setminus\{t\}} a_s. \qquad (180)$$

Also, Proposition 2.120 (applied to b_s instead of a_s) yields

$$\sum_{s\in S} b_s = b_t + \sum_{s\in S\setminus\{t\}} b_s. \qquad (181)$$

Finally, Proposition 2.120 (applied to $a_s + b_s$ instead of a_s) yields

$$\begin{aligned}
\sum_{s\in S}(a_s+b_s) &= (a_t+b_t) + \underbrace{\sum_{s\in S\setminus\{t\}}(a_s+b_s)}_{\substack{=\sum\limits_{s\in S\setminus\{t\}} a_s + \sum\limits_{s\in S\setminus\{t\}} b_s \\ \text{(by (179))}}} \\
&= (a_t+b_t) + \left(\sum_{s\in S\setminus\{t\}} a_s + \sum_{s\in S\setminus\{t\}} b_s\right) \\
&= \underbrace{\left(a_t + \sum_{s\in S\setminus\{t\}} a_s\right)}_{\substack{=\sum\limits_{s\in S} a_s \\ \text{(by (180))}}} + \underbrace{\left(b_t + \sum_{s\in S\setminus\{t\}} b_s\right)}_{\substack{=\sum\limits_{s\in S} b_s \\ \text{(by (181))}}} \\
&\qquad \left(\begin{array}{c}\text{by Lemma 2.123 (applied} \\ \text{to } x = a_t,\ y = b_t,\ u = \sum\limits_{s\in S\setminus\{t\}} a_s \text{ and } v = \sum\limits_{s\in S\setminus\{t\}} b_s)\end{array}\right) \\
&= \sum_{s\in S} a_s + \sum_{s\in S} b_s.
\end{aligned}$$

This proves Claim 2.]

But Claim 2 says precisely that Theorem 2.122 holds under the condition that $|S| = m+1$. Hence, we conclude that Theorem 2.122 holds under the condition that $|S| = m+1$ (since Claim 2 is proven). This completes the induction step. Thus, Theorem 2.122 is proven by induction. □

We shall next prove (9):

Theorem 2.124. Let S be a finite set. For every $s \in S$, let a_s be an element of $\mathbb{A}$. Also, let λ be an element of $\mathbb{A}$. Then,

$$\sum_{s\in S} \lambda a_s = \lambda \sum_{s\in S} a_s.$$

To prove this theorem, we need the following fundamental fact of arithmetic:

Proposition 2.125. Let x, y and z be three numbers (i.e., elements of $\mathbb{A}$). Then, $x(y+z) = xy + xz$.

Proposition 2.125 is known as the *distributivity* (or *left distributivity*) in $\mathbb{A}$. It is a fundamental result, and its proof can be found in standard textbooks[82].

[82] For example, Proposition 2.125 is proven in [Swanso20, Theorem 3.2.3 (6)] for the case when $\mathbb{A} = \mathbb{N}$; in [Swanso20, Theorem 3.5.4 (6)] for the case when $\mathbb{A} = \mathbb{Z}$; in [Swanso20, Theorem 3.6.4 (6)] for the case when $\mathbb{A} = \mathbb{Q}$; in [Swanso20, Theorem 3.7.14] for the case when $\mathbb{A} = \mathbb{R}$; in [Swanso20, Theorem 3.9.2] for the case when $\mathbb{A} = \mathbb{C}$.

Proof of Theorem 2.124. Forget that we fixed S, a_s and λ. We shall prove Theorem 2.124 by induction on $|S|$:

Induction base: The induction base (i.e., proving that Theorem 2.124 holds under the condition that $|S| = 0$) is similar to the induction base in the proof of Theorem 2.122 above; we thus leave it to the reader.

Induction step: Let $m \in \mathbb{N}$. Assume that Theorem 2.124 holds under the condition that $|S| = m$. We must now prove that Theorem 2.124 holds under the condition that $|S| = m + 1$.

We have assumed that Theorem 2.124 holds under the condition that $|S| = m$. In other words, the following claim holds:

> *Claim 1:* Let S be a finite set such that $|S| = m$. For every $s \in S$, let a_s be an element of $\mathbb{A}$. Also, let λ be an element of $\mathbb{A}$. Then,
>
> $$\sum_{s\in S} \lambda a_s = \lambda \sum_{s\in S} a_s.$$

Next, we shall show the following claim:

> *Claim 2:* Let S be a finite set such that $|S| = m + 1$. For every $s \in S$, let a_s be an element of $\mathbb{A}$. Also, let λ be an element of $\mathbb{A}$. Then,
>
> $$\sum_{s\in S} \lambda a_s = \lambda \sum_{s\in S} a_s.$$

[*Proof of Claim 2:* We have $|S| = m + 1 > m \geq 0$. Hence, the set S is nonempty. Thus, there exists some $t \in S$. Consider this t.

From $t \in S$, we obtain $|S \setminus \{t\}| = |S| - 1 = m$ (since $|S| = m + 1$). Hence, Claim 1 (applied to $S \setminus \{t\}$ instead of S) yields

$$\sum_{s\in S\setminus\{t\}} \lambda a_s = \lambda \sum_{s\in S\setminus\{t\}} a_s. \tag{182}$$

Now, Proposition 2.120 yields

$$\sum_{s\in S} a_s = a_t + \sum_{s\in S\setminus\{t\}} a_s.$$

Multiplying both sides of this equality by λ, we obtain

$$\lambda \sum_{s\in S} a_s = \lambda \left(a_t + \sum_{s\in S\setminus\{t\}} a_s \right) = \lambda a_t + \lambda \sum_{s\in S\setminus\{t\}} a_s$$

(by Proposition 2.125 (applied to $x = \lambda$, $y = a_t$ and $z = \sum\limits_{s\in S\setminus\{t\}} a_s$)). Also, Proposition 2.120 (applied to λa_s instead of a_s) yields

$$\sum_{s\in S} \lambda a_s = \lambda a_t + \underbrace{\sum_{s\in S\setminus\{t\}} \lambda a_s}_{\substack{=\lambda \sum\limits_{s\in S\setminus\{t\}} a_s \\ \text{(by (182))}}} = \lambda a_t + \lambda \sum_{s\in S\setminus\{t\}} a_s.$$

Comparing the preceding two equalities, we find

$$\sum_{s\in S} \lambda a_s = \lambda \sum_{s\in S} a_s.$$

This proves Claim 2.]

But Claim 2 says precisely that Theorem 2.124 holds under the condition that $|S| = m+1$. Hence, we conclude that Theorem 2.124 holds under the condition that $|S| = m+1$ (since Claim 2 is proven). This completes the induction step. Thus, Theorem 2.124 is proven by induction. □

Finally, let us prove (10):

Theorem 2.126. Let S be a finite set. Then,

$$\sum_{s\in S} 0 = 0.$$

Proof of Theorem 2.126. It is completely straightforward to prove Theorem 2.126 by induction on $|S|$ (as we proved Theorem 2.124, for example). But let us give an even shorter argument: Theorem 2.124 (applied to $a_s = 0$ and $\lambda = 0$) yields

$$\sum_{s\in S} 0 \cdot 0 = 0 \sum_{s\in S} 0 = 0.$$

In view of $0 \cdot 0 = 0$, this rewrites as $\sum\limits_{s\in S} 0 = 0$. This proves Theorem 2.126. □

2.14.10. Splitting a sum by a value of a function

We shall now prove a more complicated (but crucial) property of finite sums – namely, the equality (22) in the case when W is finite[83]:

Theorem 2.127. Let S be a finite set. Let W be a finite set. Let $f : S \to W$ be a map. Let a_s be an element of $\mathbb{A}$ for each $s \in S$. Then,

$$\sum_{s\in S} a_s = \sum_{w\in W} \sum_{\substack{s\in S;\\ f(s)=w}} a_s.$$

Here, we are using the following convention (made in Section 1.4):

[83] We prefer to only treat the case when W is finite for now. The case when W is infinite would require us to properly introduce the notion of an infinite sum with only finitely many nonzero terms. While this is not hard to do, we aren't quite ready for it yet (see Theorem 2.147 further below for this).

Convention 2.128. Let S be a finite set. Let $\mathcal{A}(s)$ be a logical statement defined for every $s \in S$. For each $s \in S$ satisfying $\mathcal{A}(s)$, let a_s be a number (i.e., an element of $\mathbb{A}$). Then, we set

$$\sum_{\substack{s\in S;\\ \mathcal{A}(s)}} a_s = \sum_{s\in\{t\in S \mid \mathcal{A}(t)\}} a_s.$$

Thus, the sum $\sum_{\substack{s\in S;\\ f(s)=w}} a_s$ in Theorem 2.127 can be rewritten as $\sum_{s\in\{t\in S \mid f(t)=w\}} a_s$.

Our proof of Theorem 2.127 relies on the following simple set-theoretic fact:

Lemma 2.129. Let S and W be two sets. Let $f : S \to W$ be a map. Let $q \in S$. Let g be the restriction $f\mid_{S\setminus\{q\}}$ of the map f to $S \setminus \{q\}$. Let $w \in W$. Then,

$$\{t \in S\setminus\{q\} \mid g(t) = w\} = \{t \in S \mid f(t) = w\} \setminus \{q\}.$$

Proof of Lemma 2.129. We know that g is the restriction $f\mid_{S\setminus\{q\}}$ of the map f to $S \setminus \{q\}$. Thus, g is a map from $S \setminus \{q\}$ to W and satisfies

$$g(t) = f(t) \qquad \text{for each } t \in S\setminus\{q\}. \tag{183}$$

Now,

$$\begin{aligned}
&\left\{t \in S\setminus\{q\} \mid \underbrace{g(t)}_{\substack{=f(t)\\ \text{(by (183))}}} = w\right\}\\
&= \{t \in S\setminus\{q\} \mid f(t) = w\} = \{t \in S \mid f(t) = w \text{ and } t \in S\setminus\{q\}\}\\
&= \{t \in S \mid f(t) = w\} \cap \underbrace{\{t \in S \mid t \in S\setminus\{q\}\}}_{=S\setminus\{q\}}\\
&= \{t \in S \mid f(t) = w\} \cap (S\setminus\{q\}) = \{t \in S \mid f(t) = w\} \setminus \{q\}.
\end{aligned}$$

This proves Lemma 2.129. □

Proof of Theorem 2.127. We shall prove Theorem 2.127 by induction on $|S|$:

Induction base: Theorem 2.127 holds under the condition that $|S| = 0$ [84]. This completes the induction base.

[84] *Proof.* Let S, W, f and a_s be as in Theorem 2.127. Assume that $|S| = 0$. Thus, the first bullet point of Definition 2.111 yields $\sum_{s\in S} a_s = 0$. Moreover, $S = \varnothing$ (since $|S| = 0$). Hence, each $w \in W$

Induction step: Let $m \in \mathbb{N}$. Assume that Theorem 2.127 holds under the condition that $|S| = m$. We must now prove that Theorem 2.127 holds under the condition that $|S| = m+1$.

We have assumed that Theorem 2.127 holds under the condition that $|S| = m$. In other words, the following claim holds:

> *Claim 1:* Let S be a finite set such that $|S| = m$. Let W be a finite set. Let $f : S \to W$ be a map. Let a_s be an element of $\mathbb{A}$ for each $s \in S$. Then,
>
> $$\sum_{s\in S} a_s = \sum_{w\in W} \sum_{\substack{s\in S;\\ f(s)=w}} a_s.$$

Next, we shall show the following claim:

> *Claim 2:* Let S be a finite set such that $|S| = m+1$. Let W be a finite set. Let $f : S \to W$ be a map. Let a_s be an element of $\mathbb{A}$ for each $s \in S$. Then,
>
> $$\sum_{s\in S} a_s = \sum_{w\in W} \sum_{\substack{s\in S;\\ f(s)=w}} a_s.$$

[*Proof of Claim 2:* We have $|S| = m+1 > m \geq 0$. Hence, the set S is nonempty. Thus, there exists some $q \in S$. Consider this q.

From $q \in S$, we obtain $|S \setminus \{q\}| = |S| - 1 = m$ (since $|S| = m+1$).

Let g be the restriction $f\mid_{S\setminus\{q\}}$ of the map f to $S \setminus \{q\}$. Thus, g is a map from $S \setminus \{q\}$ to W.

satisfies

$$\sum_{\substack{s\in S;\\ f(s)=w}} a_s = \sum_{s\in\{t\in S \mid f(t)=w\}} a_s = \sum_{s\in\varnothing} a_s \qquad \left(\begin{array}{c}\text{since } \{t\in S \mid f(s)=w\} = \varnothing \\ \text{(because } \{t\in S \mid f(s)=w\} \subseteq S = \varnothing)\end{array}\right)$$
$$= (\text{empty sum}) = 0.$$

Summing these equalities over all $w \in W$, we obtain

$$\sum_{w\in W} \sum_{\substack{s\in S;\\ f(s)=w}} a_s = \sum_{w\in W} 0 = 0$$

(by an application of Theorem 2.126). Comparing this with $\sum_{s\in S} a_s = 0$, we obtain $\sum_{s\in S} a_s = \sum_{w\in W} \sum_{\substack{s\in S;\\ f(s)=w}} a_s$.

Now, forget that we fixed S, W, f and a_s. We thus have proved that if S, W, f and a_s are as in Theorem 2.127, and if $|S| = 0$, then $\sum_{s\in S} a_s = \sum_{w\in W} \sum_{\substack{s\in S;\\ f(s)=w}} a_s$. In other words, Theorem 2.127 holds under the condition that $|S| = 0$. Qed.

For each $w \in W$, we define a number b_w by

$$b_w = \sum_{\substack{s\in S;\\ f(s)=w}} a_s. \tag{184}$$

Furthermore, for each $w \in W$, we define a number c_w by

$$c_w = \sum_{\substack{s\in S\setminus\{q\};\\ g(s)=w}} a_s. \tag{185}$$

Recall that $|S \setminus \{q\}| = m$. Hence, Claim 1 (applied to $S \setminus \{q\}$ and g instead of S and f) yields

$$\sum_{s\in S\setminus\{q\}} a_s = \sum_{w\in W} \underbrace{\sum_{\substack{s\in S\setminus\{q\};\\ g(s)=w}} a_s}_{\substack{=c_w\\ \text{(by (185))}}} = \sum_{w\in W} c_w. \tag{186}$$

Every $w \in W \setminus \{f(q)\}$ satisfies

$$b_w = c_w. \tag{187}$$

[*Proof of (187):* Let $w \in W \setminus \{f(q)\}$. Thus, $w \in W$ and $w \notin \{f(q)\}$.

If we had $q \in \{t \in S \mid f(t) = w\}$, then we would have $f(q) = w$, which would lead to $w = f(q) \in \{f(q)\}$; but this would contradict $w \notin \{f(q)\}$. Hence, we cannot have $q \in \{t \in S \mid f(t) = w\}$. Hence, we have $q \notin \{t \in S \mid f(t) = w\}$.

But $w \in W$; thus, Lemma 2.129 yields

$$\begin{aligned}\{t \in S\setminus\{q\} \mid g(t) = w\} &= \{t \in S \mid f(t) = w\} \setminus \{q\}\\ &= \{t \in S \mid f(t) = w\}\end{aligned} \tag{188}$$

(since $q \notin \{t \in S \mid f(t) = w\}$).

On the other hand, the definition of b_w yields

$$b_w = \sum_{\substack{s\in S;\\ f(s)=w}} a_s = \sum_{s\in\{t\in S \mid f(t)=w\}} a_s \tag{189}$$

(by the definition of the "$\sum_{\substack{s\in S;\\ f(s)=w}}$" symbol). Also, the definition of c_w yields

$$\begin{aligned}c_w &= \sum_{\substack{s\in S\setminus\{q\};\\ g(s)=w}} a_s = \sum_{s\in\{t\in S\setminus\{q\} \mid g(t)=w\}} a_s = \sum_{s\in\{t\in S \mid f(t)=w\}} a_s\\ &\qquad \left(\begin{array}{c}\text{since } \{t \in S\setminus\{q\} \mid g(t) = w\} = \{t \in S \mid f(t) = w\}\\ \text{(by (188))}\end{array}\right)\\ &= b_w \qquad \text{(by (189))}.\end{aligned}$$

Thus, $b_w = c_w$. This proves (187).]

Also,

$$b_{f(q)} = a_q + c_{f(q)}. \tag{190}$$

[*Proof of (190):* Define a subset U of S by

$$U = \{t \in S \mid f(t) = f(q)\}. \tag{191}$$

We can apply Lemma 2.129 to $w = f(q)$. We thus obtain

$$\begin{aligned} \{t \in S \setminus \{q\} \mid g(t) = f(q)\} &= \underbrace{\{t \in S \mid f(t) = f(q)\}}_{\substack{=U \\ \text{(by (191))}}} \setminus \{q\} \\ &= U \setminus \{q\}. \end{aligned} \tag{192}$$

We know that q is a $t \in S$ satisfying $f(t) = f(q)$ (since $q \in S$ and $f(q) = f(q)$). In other words, $q \in \{t \in S \mid f(t) = f(q)\}$. In other words, $q \in U$ (since $U = \{t \in S \mid f(t) = f(q)\}$). Thus, Proposition 2.120 (applied to U and q instead of S and t) yields

$$\sum_{s \in U} a_s = a_q + \sum_{s \in U \setminus \{q\}} a_s. \tag{193}$$

But (192) shows that $U \setminus \{q\} = \{t \in S \setminus \{q\} \mid g(t) = f(q)\}$. Thus,

$$\sum_{s \in U \setminus \{q\}} a_s = \sum_{s \in \{t \in S \setminus \{q\} \mid g(t) = f(q)\}} a_s = c_{f(q)} \tag{194}$$

(since the definition of $c_{f(q)}$ yields $c_{f(q)} = \sum_{\substack{s \in S \setminus \{q\}; \\ g(s) = f(q)}} a_s = \sum_{s \in \{t \in S \setminus \{q\} \mid g(t) = f(q)\}} a_s$).

On the other hand, the definition of $b_{f(q)}$ yields

$$\begin{aligned} b_{f(q)} &= \sum_{\substack{s \in S; \\ f(s) = f(q)}} a_s = \sum_{s \in \{t \in S \mid f(t) = f(q)\}} a_s \\ &= \sum_{s \in U} a_s \qquad \left(\text{since } \{t \in S \mid f(t) = f(q)\} = U\right) \\ &= a_q + \underbrace{\sum_{s \in U \setminus \{q\}} a_s}_{\substack{= c_{f(q)} \\ \text{(by (194))}}} \qquad \text{(by (193))} \\ &= a_q + c_{f(q)}. \end{aligned}$$

This proves (190).]

Now, recall that $q \in S$. Hence, Proposition 2.120 (applied to $t = q$) yields

$$\sum_{s \in S} a_s = a_q + \sum_{s \in S \setminus \{q\}} a_s. \tag{195}$$

Also, $f(q) \in W$. Hence, Proposition 2.120 (applied to W, $(c_w)_{w\in W}$ and $f(q)$ instead of S, $(a_s)_{s\in S}$ and t) yields

$$\sum_{w\in W} c_w = c_{f(q)} + \sum_{w\in W\setminus\{f(q)\}} c_w.$$

Hence, (186) becomes

$$\sum_{s\in S\setminus\{q\}} a_s = \sum_{w\in W} c_w = c_{f(q)} + \sum_{w\in W\setminus\{f(q)\}} c_w. \qquad (196)$$

Also, Proposition 2.120 (applied to W, $(b_w)_{w\in W}$ and $f(q)$ instead of S, $(a_s)_{s\in S}$ and t) yields

$$\sum_{w\in W} b_w = \underbrace{b_{f(q)}}_{\substack{=a_q+c_{f(q)} \\ \text{(by (190))}}} + \sum_{w\in W\setminus\{f(q)\}} \underbrace{b_w}_{\substack{=c_w \\ \text{(by (187))}}}$$
$$= \left(a_q + c_{f(q)}\right) + \sum_{w\in W\setminus\{f(q)\}} c_w = a_q + \left(c_{f(q)} + \sum_{w\in W\setminus\{f(q)\}} c_w\right)$$

(by Proposition 2.103, applied to a_q, $c_{f(q)}$ and $\sum\limits_{w\in W\setminus\{f(q)\}} c_w$ instead of a, b and c). Thus,

$$\sum_{w\in W} b_w = a_q + \underbrace{\left(c_{f(q)} + \sum_{w\in W\setminus\{f(q)\}} c_w\right)}_{\substack{=\sum\limits_{s\in S\setminus\{q\}} a_s \\ \text{(by (196))}}} = a_q + \sum_{s\in S\setminus\{q\}} a_s = \sum_{s\in S} a_s$$

(by (195)). Hence,

$$\sum_{s\in S} a_s = \sum_{w\in W} \underbrace{b_w}_{\substack{=\sum\limits_{\substack{s\in S;\\ f(s)=w}} a_s \\ \text{(by (184))}}} = \sum_{w\in W} \sum_{\substack{s\in S;\\ f(s)=w}} a_s.$$

This proves Claim 2.]

But Claim 2 says precisely that Theorem 2.127 holds under the condition that $|S| = m+1$. Hence, we conclude that Theorem 2.127 holds under the condition that $|S| = m+1$ (since Claim 2 is proven). This completes the induction step. Thus, Theorem 2.127 is proven by induction. □

2.14.11. Splitting a sum into two

Next, we shall prove the equality (3):

Theorem 2.130. Let S be a finite set. Let X and Y be two subsets of S such that $X \cap Y = \varnothing$ and $X \cup Y = S$. (Equivalently, X and Y are two subsets of S such that each element of S lies in **exactly** one of X and Y.) Let a_s be a number (i.e., an element of $\mathbb{A}$) for each $s \in S$. Then,

$$\sum_{s\in S} a_s = \sum_{s\in X} a_s + \sum_{s\in Y} a_s .$$

Proof of Theorem 2.130. From the assumptions $X \cap Y = \varnothing$ and $X \cup Y = S$, we can easily obtain $S \setminus X = Y$.

We define a map $f : S \to \{0,1\}$ by setting

$$\left(f(s) = \begin{cases} 0, & \text{if } s \in X; \\ 1, & \text{if } s \notin X \end{cases} \qquad \text{for every } s \in S \right).$$

For each $w \in \{0,1\}$, we define a number b_w by

$$b_w = \sum_{\substack{s\in S;\\ f(s)=w}} a_s . \tag{197}$$

Proposition 2.121 **(b)** (applied to $\{0,1\}$, 0, 1 and $(b_w)_{w\in\{0,1\}}$ instead of S, p, q and $(a_s)_{s\in S}$) yields $\sum_{w\in\{0,1\}} b_w = b_0 + b_1$.

Now, Theorem 2.127 (applied to $W = \{0,1\}$) yields

$$\sum_{s\in S} a_s = \sum_{w\in\{0,1\}} \underbrace{\sum_{\substack{s\in S;\\ f(s)=w}} a_s}_{\substack{=b_w \\ \text{(by (197))}}} = \sum_{w\in\{0,1\}} b_w = b_0 + b_1 . \tag{198}$$

On the other hand,

$$b_0 = \sum_{s\in X} a_s . \tag{199}$$

[*Proof of (199):* The definition of the map f shows that an element $t \in S$ satisfies $f(t) = 0$ **if and only if** it belongs to X. Hence, the set of all elements $t \in S$ that satisfy $f(t) = 0$ is precisely X. In other words,

$$\{t \in S \mid f(t) = 0\} = X.$$

But the definition of b_0 yields

$$b_0 = \sum_{\substack{s\in S;\\ f(s)=0}} a_s = \sum_{s\in\{t\in S \mid f(t)=0\}} a_s = \sum_{s\in X} a_s$$

(since $\{t \in S \mid f(t) = 0\} = X$). This proves (199).]

Furthermore,

$$b_1 = \sum_{s \in Y} a_s. \tag{200}$$

[*Proof of (200):* The definition of the map f shows that an element $t \in S$ satisfies $f(t) = 1$ **if and only if** $t \notin X$. Thus, for each $t \in S$, we have the following chain of equivalences:

$$(f(t) = 1) \iff (t \notin X) \iff (t \in S \setminus X) \iff (t \in Y)$$

(since $S \setminus X = Y$). In other words, an element $t \in S$ satisfies $f(t) = 1$ **if and only if** t belongs to Y. Hence, the set of all elements $t \in S$ that satisfy $f(t) = 1$ is precisely Y. In other words,

$$\{t \in S \mid f(t) = 1\} = Y.$$

But the definition of b_1 yields

$$b_1 = \sum_{\substack{s \in S; \\ f(s)=1}} a_s = \sum_{s \in \{t \in S \mid f(t)=1\}} a_s = \sum_{s \in Y} a_s$$

(since $\{t \in S \mid f(t) = 1\} = Y$). This proves (200).]

Now, (198) becomes

$$\sum_{s \in S} a_s = \underbrace{b_0}_{\substack{= \sum_{s \in X} a_s \\ \text{(by (199))}}} + \underbrace{b_1}_{\substack{= \sum_{s \in Y} a_s \\ \text{(by (200))}}} = \sum_{s \in X} a_s + \sum_{s \in Y} a_s.$$

This proves Theorem 2.130. □

Similarly, we can prove the equality (26). (This proof was already outlined in Section 1.4.)

A consequence of Theorem 2.130 is the following fact, which has appeared as the equality (11) in Section 1.4:

Corollary 2.131. Let S be a finite set. Let a_s be an element of $\mathbb{A}$ for each $s \in S$. Let T be a subset of S such that every $s \in T$ satisfies $a_s = 0$. Then,

$$\sum_{s \in S} a_s = \sum_{s \in S \setminus T} a_s.$$

Proof of Corollary 2.131. We have assumed that every $s \in T$ satisfies $a_s = 0$. Thus, $\sum_{s \in T} \underbrace{a_s}_{=0} = \sum_{s \in T} 0 = 0$ (by Theorem 2.126 (applied to T instead of S)).

But T and $S \setminus T$ are subsets of S. These two subsets satisfy $T \cap (S \setminus T) = \varnothing$ and $T \cup (S \setminus T) = S$ (since $T \subseteq S$). Hence, Theorem 2.130 (applied to $X = T$ and $Y = S \setminus T$) yields

$$\sum_{s \in S} a_s = \underbrace{\sum_{s \in T} a_s}_{=0} + \sum_{s \in S \setminus T} a_s = \sum_{s \in S \setminus T} a_s.$$

This proves Corollary 2.131. □

2.14.12. Substituting the summation index

Next, we shall show the equality (12):

> **Theorem 2.132.** Let S and T be two finite sets. Let $f : S \to T$ be a **bijective** map. Let a_t be an element of $\mathbb{A}$ for each $t \in T$. Then,
>
> $$\sum_{t \in T} a_t = \sum_{s \in S} a_{f(s)}.$$

Proof of Theorem 2.132. Each $w \in T$ satisfies

$$\sum_{\substack{s \in S; \\ f(s) = w}} a_{f(s)} = a_w. \qquad (201)$$

[*Proof of (201):* Let $w \in T$.

The map f is bijective; thus, it is invertible. In other words, its inverse map $f^{-1} : T \to S$ exists. Hence, $f^{-1}(w)$ is a well-defined element of S, and is the only element $t \in S$ satisfying $f(t) = w$. Therefore,

$$\{t \in S \mid f(t) = w\} = \left\{ f^{-1}(w) \right\}. \qquad (202)$$

Now,

$$\begin{aligned}
& \sum_{\substack{s \in S; \\ f(s) = w}} a_{f(s)} \\
&= \sum_{s \in \{t \in S \mid f(t) = w\}} a_{f(s)} = \sum_{s \in \left\{ f^{-1}(w) \right\}} a_{f(s)} \qquad (\text{by (202)}) \\
&= a_{f\left(f^{-1}(w)\right)} \qquad \left(\begin{array}{c} \text{by Proposition 2.121 \textbf{(a)} (applied to } \left\{ f^{-1}(w) \right\}, a_{f(s)} \\ \text{and } f^{-1}(w) \text{ instead of } S, a_s \text{ and } p) \end{array} \right) \\
&= a_w \qquad \left(\text{since } f\left(f^{-1}(w) \right) = w \right).
\end{aligned}$$

This proves (201).]

Renaming the summation index w as t in the sum $\sum_{w\in T} a_w$ does not change the sum (since $(a_w)_{w\in T}$ and $(a_t)_{t\in T}$ are the same $\mathbb{A}$-valued T-family). In other words, $\sum_{w\in T} a_w = \sum_{t\in T} a_t$.

Theorem 2.127 (applied to T and $a_{f(s)}$ instead of W and a_s) yields

$$\sum_{s\in S} a_{f(s)} = \sum_{w\in T} \underbrace{\sum_{\substack{s\in S;\\ f(s)=w}} a_{f(s)}}_{\substack{=a_w \\ \text{(by (201))}}} = \sum_{w\in T} a_w = \sum_{t\in T} a_t .$$

This proves Theorem 2.132. □

2.14.13. Sums of congruences

Proposition 2.21 **(a)** says that we can add two congruences modulo an integer n. We shall now see that we can add **any** number of congruences modulo an integer n:

> **Theorem 2.133.** Let n be an integer. Let S be a finite set. For each $s \in S$, let a_s and b_s be two integers. Assume that
>
> $$a_s \equiv b_s \operatorname{mod} n \qquad \text{for each } s \in S.$$
>
> Then,
>
> $$\sum_{s\in S} a_s \equiv \sum_{s\in S} b_s \operatorname{mod} n.$$

Proof of Theorem 2.133. We forget that we fixed n, S, a_s and b_s. We shall prove Theorem 2.133 by induction on $|S|$:

Induction base: The induction base (i.e., proving that Theorem 2.133 holds under the condition that $|S| = 0$) is left to the reader (as it boils down to the trivial fact that $0 \equiv 0 \operatorname{mod} n$).

Induction step: Let $m \in \mathbb{N}$. Assume that Theorem 2.133 holds under the condition that $|S| = m$. We must now prove that Theorem 2.133 holds under the condition that $|S| = m + 1$.

We have assumed that Theorem 2.133 holds under the condition that $|S| = m$. In other words, the following claim holds:

> *Claim 1:* Let n be an integer. Let S be a finite set such that $|S| = m$. For each $s \in S$, let a_s and b_s be two integers. Assume that
>
> $$a_s \equiv b_s \operatorname{mod} n \qquad \text{for each } s \in S.$$
>
> Then,
>
> $$\sum_{s\in S} a_s \equiv \sum_{s\in S} b_s \operatorname{mod} n.$$

Next, we shall show the following claim:

> *Claim 2:* Let n be an integer. Let S be a finite set such that $|S| = m+1$. For each $s \in S$, let a_s and b_s be two integers. Assume that
>
> $$a_s \equiv b_s \operatorname{mod} n \qquad \text{for each } s \in S. \tag{203}$$
>
> Then,
>
> $$\sum_{s\in S} a_s \equiv \sum_{s\in S} b_s \operatorname{mod} n.$$

[*Proof of Claim 2:* We have $|S| = m+1 > m \geq 0$. Hence, the set S is nonempty. Thus, there exists some $t \in S$. Consider this t.

From $t \in S$, we obtain $|S \setminus \{t\}| = |S| - 1 = m$ (since $|S| = m+1$). Also, every $s \in S \setminus \{t\}$ satisfies $s \in S \setminus \{t\} \subseteq S$ and thus $a_s \equiv b_s \operatorname{mod} n$ (by (203)). In other words, we have

$$a_s \equiv b_s \operatorname{mod} n \qquad \text{for each } s \in S \setminus \{t\}.$$

Hence, Claim 1 (applied to $S \setminus \{t\}$ instead of S) yields

$$\sum_{s\in S\setminus\{t\}} a_s \equiv \sum_{s\in S\setminus\{t\}} b_s \operatorname{mod} n. \tag{204}$$

But $t \in S$. Hence, (203) (applied to $s = t$) yields $a_t \equiv b_t \operatorname{mod} n$.

Now, Proposition 2.120 (applied to b_s instead of a_s) yields

$$\sum_{s\in S} b_s = b_t + \sum_{s\in S\setminus\{t\}} b_s. \tag{205}$$

But Proposition 2.120 yields

$$\sum_{s\in S} a_s = \underbrace{a_t}_{\equiv b_t \operatorname{mod} n} + \underbrace{\sum_{s\in S\setminus\{t\}} a_s}_{\substack{\equiv \sum_{s\in S\setminus\{t\}} b_s \operatorname{mod} n \\ \text{(by (204))}}} \equiv b_t + \sum_{s\in S\setminus\{t\}} b_s = \sum_{s\in S} b_s \operatorname{mod} n$$

(by (205)). This proves Claim 2.]

But Claim 2 says precisely that Theorem 2.133 holds under the condition that $|S| = m+1$. Hence, we conclude that Theorem 2.133 holds under the condition that $|S| = m+1$ (since Claim 2 is proven). This completes the induction step. Thus, Theorem 2.133 is proven by induction. □

As we said, Theorem 2.133 shows that we can sum up several congruences. Thus, we can extend our principle of substitutivity for congruences as follows:

> *Principle of substitutivity for congruences (stronger version):* Fix an integer n. If two numbers x and x' are congruent to each other modulo n (that is, $x \equiv x' \mod n$), and if we have any expression A that involves only integers, addition, subtraction, multiplication **and summation signs**, and involves the object x, then we can replace this x (or, more precisely, any arbitrary appearance of x in A) in A by x'; the value of the resulting expression A' will be congruent to the value of A modulo n.

For example, if $p \in \mathbb{N}$, then

$$\sum_{s\in\{1,2,\ldots,p\}} s^2 (5-3s) \equiv \sum_{s\in\{1,2,\ldots,p\}} s(5-3s) \mod 2$$

(here, we have replaced the "s^2" inside the sum by "s"), because every $s \in \{1,2,\ldots,p\}$ satisfies $s^2 \equiv s \mod 2$ (this is easy to check[85]).

2.14.14. Finite products

Proposition 2.103 is a property of the addition of numbers; it has an analogue for multiplication of numbers:

Proposition 2.134. Let a, b and c be three numbers (i.e., elements of $\mathbb{A}$). Then, $(ab)c = a(bc)$.

Proposition 2.134 is known as the *associativity of multiplication* (in $\mathbb{A}$), and is fundamental; its proof can be found in any textbook on the construction of the number system[86].

Proposition 2.104 also has an analogue for multiplication:

Proposition 2.135. Let a and b be two numbers (i.e., elements of $\mathbb{A}$). Then, $ab = ba$.

Proposition 2.135 is known as the *commutativity of multiplication* (in $\mathbb{A}$), and again is a fundamental result whose proofs are found in standard textbooks[87].

[85] *Proof.* Let $p \in \mathbb{N}$ and $s \in \{1,2,\ldots,p\}$. We must prove that $s^2 \equiv s \mod 2$.

We have $s \in \{1,2,\ldots,p\}$ and thus $s-1 \in \{0,1,\ldots,p-1\} \subseteq \mathbb{N}$. Hence, (172) (applied to $n = s-1$) yields $\sum_{i\in\{1,2,\ldots,s-1\}} i = \frac{(s-1)((s-1)+1)}{2} = \frac{(s-1)s}{2}$. Hence, $\frac{(s-1)s}{2}$ is an integer (since $\sum_{i\in\{1,2,\ldots,s-1\}} i$ is an integer). In other words, $2 \mid (s-1)s$. In other words, $2 \mid s^2 - s$ (since $(s-1)s = s^2 - s$). In other words, $s^2 \equiv s \mod 2$ (by the definition of "congruent"), qed.

[86] For example, Proposition 2.134 is proven in [Swanso20, Theorem 3.2.3 (7)] for the case when $\mathbb{A} = \mathbb{N}$; in [Swanso20, Theorem 3.5.4 (7)] for the case when $\mathbb{A} = \mathbb{Z}$; in [Swanso20, Theorem 3.6.4 (7)] for the case when $\mathbb{A} = \mathbb{Q}$; in [Swanso20, Theorem 3.7.14] for the case when $\mathbb{A} = \mathbb{R}$; in [Swanso20, Theorem 3.9.2] for the case when $\mathbb{A} = \mathbb{C}$.

[87] For example, Proposition 2.135 is proven in [Swanso20, Theorem 3.2.3 (8)] for the case when $\mathbb{A} = \mathbb{N}$; in [Swanso20, Theorem 3.5.4 (8)] for the case when $\mathbb{A} = \mathbb{Z}$; in [Swanso20, Theorem 3.6.4 (8)] for the case when $\mathbb{A} = \mathbb{Q}$; in [Swanso20, Theorem 3.7.14] for the case when $\mathbb{A} = \mathbb{R}$; in [Swanso20, Theorem 3.9.2] for the case when $\mathbb{A} = \mathbb{C}$.

Proposition 2.125 has an analogue for multiplication as well (but note that x now needs to be in $\mathbb{N}$, in order to guarantee that the powers are well-defined):

Proposition 2.136. Let $x \in \mathbb{N}$. Let y and z be two numbers (i.e., elements of $\mathbb{A}$). Then, $(yz)^x = y^x z^x$.

Proposition 2.136 is one of the laws of exponents, and can easily be shown by induction on x (using Proposition 2.135 and Proposition 2.134).

So far in Section 2.14, we have been studying **sums** of $\mathbb{A}$-valued S-families (when S is a finite set): We have proven that the definition of $\sum\limits_{s \in S} a_s$ given in Section 1.4 is legitimate, and we have proven several properties of such sums. By the exact same reasoning (but with addition replaced by multiplication), we can study **products** of $\mathbb{A}$-valued S-families. In particular, we can similarly prove that the definition of $\prod\limits_{s \in S} a_s$ given in Section 1.4 is legitimate, and we can prove properties of such products that are analogous to the properties of sums proven above (except for Proposition 2.119, which does not have an analogue for products)[88]. For example, the following theorems are analogues of Theorem 2.122, Theorem 2.124, Theorem 2.126, Theorem 2.127, Theorem 2.132 and Theorem 2.133, respectively:

Theorem 2.137. Let S be a finite set. For every $s \in S$, let a_s and b_s be elements of $\mathbb{A}$. Then,

$$\prod_{s \in S} (a_s b_s) = \left(\prod_{s \in S} a_s \right) \cdot \left(\prod_{s \in S} b_s \right).$$

[88] We need to be slightly careful when we adapt our above proofs to products instead of sums: Apart from replacing addition by multiplication everywhere, we need to:

- replace the number 0 by 1 whenever it appears in a computation inside $\mathbb{A}$ (but, of course, not when it appears as the size of a set);
- replace every $\sum$ sign by a $\prod$ sign;
- replace "let λ be an element of $\mathbb{A}$" by "let λ be an element of $\mathbb{N}$" in Theorem 2.124;
- replace any expression of the form "λb" by "b^λ" in Theorem 2.124 (so that the claim of Theorem 2.124 becomes $\prod\limits_{s \in S} (a_s)^\lambda = \left(\prod\limits_{s \in S} a_s \right)^\lambda$) and in its proof;
- replace every reference to Proposition 2.103 by a reference to Proposition 2.134;
- replace every reference to Proposition 2.104 by a reference to Proposition 2.135;
- replace every reference to Proposition 2.125 by a reference to Proposition 2.136.

And, to be fully precise: We should not replace addition by multiplication **everywhere** (e.g., we should not replace "$|S| = m + 1$" by "$|S| = m \cdot 1$" in the proof of Theorem 2.127), but of course only where it stands for the addition **inside** $\mathbb{A}$.

Theorem 2.138. Let S be a finite set. For every $s \in S$, let a_s be an element of $\mathbb{A}$. Also, let λ be an element of $\mathbb{N}$. Then,

$$\prod_{s \in S} (a_s)^{\lambda} = \left(\prod_{s \in S} a_s \right)^{\lambda}.$$

Theorem 2.139. Let S be a finite set. Then,

$$\prod_{s \in S} 1 = 1.$$

Theorem 2.140. Let S be a finite set. Let W be a finite set. Let $f : S \to W$ be a map. Let a_s be an element of $\mathbb{A}$ for each $s \in S$. Then,

$$\prod_{s \in S} a_s = \prod_{w \in W} \prod_{\substack{s \in S; \\ f(s) = w}} a_s.$$

Theorem 2.141. Let S and T be two finite sets. Let $f : S \to T$ be a **bijective** map. Let a_t be an element of $\mathbb{A}$ for each $t \in T$. Then,

$$\prod_{t \in T} a_t = \prod_{s \in S} a_{f(s)}.$$

Theorem 2.142. Let n be an integer. Let S be a finite set. For each $s \in S$, let a_s and b_s be two integers. Assume that

$$a_s \equiv b_s \operatorname{mod} n \qquad \text{for each } s \in S.$$

Then,

$$\prod_{s \in S} a_s \equiv \prod_{s \in S} b_s \operatorname{mod} n.$$

2.14.15. Finitely supported (but possibly infinite) sums

In Section 1.4, we mentioned that a sum of the form $\sum_{s \in S} a_s$ can be well-defined even when the set S is not finite. Indeed, for it to be well-defined, it suffices that **only finitely many among the a_s are nonzero** (or, more rigorously: only finitely many $s \in S$ satisfy $a_s \neq 0$). As we already mentioned, the sum $\sum_{s \in S} a_s$ in this case is defined by discarding the zero addends and summing the finitely many addends that remain. Let us briefly discuss such sums (without focussing on advanced properties):

Definition 2.143. Let S be any set. An $\mathbb{A}$-valued S-family $(a_s)_{s\in S}$ is said to be *finitely supported* if only finitely many $s \in S$ satisfy $a_s \neq 0$.

So the sums we want to discuss are sums $\sum\limits_{s\in S} a_s$ for which the set S may be infinite but the S-family $(a_s)_{s\in S}$ is finitely supported. Let us repeat the definition of such sums in more rigorous language:

Definition 2.144. Let S be any set. Let $(a_s)_{s\in S}$ be a finitely supported $\mathbb{A}$-valued S-family. Thus, there exists a **finite** subset T of S such that

$$\text{every } s \in S \setminus T \text{ satisfies } a_s = 0. \tag{206}$$

(This is because only finitely many $s \in S$ satisfy $a_s \neq 0$.) We then define the sum $\sum\limits_{s\in S} a_s$ to be $\sum\limits_{s\in T} a_s$. (This definition is legitimate, because Proposition 2.145 **(a)** below shows that $\sum\limits_{s\in T} a_s$ does not depend on the choice of T.)

This definition formalizes what we said above about making sense of $\sum\limits_{s\in S} a_s$: Namely, we discard zero addends (namely, the addends corresponding to $s \in S \setminus T$) and only sum the finitely many addends that remain (these are the addends corresponding to $s \in T$); thus, we get $\sum\limits_{s\in T} a_s$. Note that we are not requiring that every $s \in T$ satisfies $a_s \neq 0$; that is, we are not necessarily discarding **all** the zero addends from our sum (but merely discarding enough of them to ensure that only finitely many remain). This may appear like a strange choice (why introduce extra freedom into the definition?), but is reasonable from the viewpoint of constructive mathematics (where it is not always decidable if a number is 0 or not).

Proposition 2.145. Let S be any set. Let $(a_s)_{s\in S}$ be a finitely supported $\mathbb{A}$-valued S-family.

(a) If T is a finite subset of S such that (206) holds, then the sum $\sum\limits_{s\in T} a_s$ does not depend on the choice of T. (That is, if T_1 and T_2 are two finite subsets T of S satisfying (206), then $\sum\limits_{s\in T_1} a_s = \sum\limits_{s\in T_2} a_s$.)

(b) If the set S is finite, then the sum $\sum\limits_{s\in S} a_s$ defined in Definition 2.144 is identical with the sum $\sum\limits_{s\in S} a_s$ defined in Definition 2.111. (Thus, Definition 2.144 does not conflict with the previous definition of $\sum\limits_{s\in S} a_s$ for finite sets S.)

Proposition 2.145 is fairly easy to prove using Corollary 2.131; this proof is part of Exercise 2.3 below.

Most properties of finite sums have analogues for sums of finitely supported $\mathbb{A}$-valued S-families. For example, here is an analogue of Theorem 2.122:

Theorem 2.146. Let S be a set. Let $(a_s)_{s\in S}$ and $(b_s)_{s\in S}$ be two finitely supported $\mathbb{A}$-valued S-families. Then, the $\mathbb{A}$-valued S-family $(a_s+b_s)_{s\in S}$ is finitely supported as well, and we have

$$\sum_{s\in S}(a_s+b_s)=\sum_{s\in S}a_s+\sum_{s\in S}b_s.$$

The proof of Theorem 2.146 is fairly simple (it relies prominently on the fact that the union of two finite sets is finite), and again is part of Exercise 2.3 below.

It is also easy to state and prove analogues of Theorem 2.124 and Theorem 2.126.

We can next prove (22) in full generality (not only when W is finite):

Theorem 2.147. Let S be a finite set. Let W be a set. Let $f : S \to W$ be a map. Let a_s be an element of $\mathbb{A}$ for each $s \in S$. Then, the $\mathbb{A}$-valued W-family $\left(\sum\limits_{\substack{s\in S;\\ f(s)=w}} a_s\right)_{w\in W}$ is finitely supported and satisfies

$$\sum_{s\in S}a_s=\sum_{w\in W}\sum_{\substack{s\in S;\\ f(s)=w}}a_s.$$

Note that the sum on the right hand side of Theorem 2.147 makes sense even when W is infinite, because the W-family $\left(\sum\limits_{\substack{s\in S;\\ f(s)=w}} a_s\right)_{w\in W}$ is finitely supported (i.e., only finitely many $w \in W$ satisfy $\sum\limits_{\substack{s\in S;\\ f(s)=w}} a_s \neq 0$). The easiest way to prove Theorem 2.147 is probably by reducing it to Theorem 2.127 (since $f(S)$ is a finite subset of W, and every $w \in W \setminus f(S)$ satisfies $\sum\limits_{\substack{s\in S;\\ f(s)=w}} a_s = (\text{empty sum}) = 0$). Again, we leave the details to the interested reader.

Again, we refer to Exercise 2.3 for the proof of Theorem 2.147.

Actually, Theorem 2.147 can be generalized even further:

Theorem 2.148. Let S be a set. Let W be a set. Let $f : S \to W$ be a map. Let $(a_s)_{s\in S}$ be a finitely supported $\mathbb{A}$-valued S-family. Then, for each $w \in W$, the $\mathbb{A}$-valued $\{t \in S \mid f(t) = w\}$-family $(a_s)_{s\in\{t\in S \mid f(t)=w\}}$ is finitely supported as well (so that the sum $\sum\limits_{\substack{s\in S;\\ f(s)=w}} a_s$ is well-defined). Furthermore, the $\mathbb{A}$-valued W-family $\left(\sum\limits_{\substack{s\in S;\\ f(s)=w}} a_s\right)_{w\in W}$ is also finitely supported. Finally,

$$\sum_{s\in S} a_s = \sum_{w\in W} \sum_{\substack{s\in S;\\ f(s)=w}} a_s.$$

Again, see Exercise 2.3 for the proof. This theorem can be used to obtain an analogue of Theorem 2.130 for finitely supported $\mathbb{A}$-valued S-families.

Exercise 2.3. Prove Proposition 2.145, Theorem 2.146, Theorem 2.147 and Theorem 2.148.

Thus, we have defined the values of certain infinite sums (although not nearly as many infinite sums as analysis can make sense of). We can similarly define the values of certain infinite products: In order for $\prod\limits_{s\in S} a_s$ to be well-defined, it suffices that **only finitely many among the** a_s **are distinct from** 1 (or, more rigorously: only finitely many $s \in S$ satisfy $a_s \neq 1$). We leave the details and properties of this definition to the reader.

2.15. Two-sided induction

2.15.1. The principle of two-sided induction

Let us now return to studying induction principles. We have seen several induction principles that allow us to prove statements about nonnegative integers, integers in $\mathbb{Z}_{\geq g}$ or integers in an interval. What about proving statements about **arbitrary** integers? The induction principles we have seen so far do not suffice to prove such statements directly, since our induction steps always "go up" (in the sense that they begin by assuming that our statement $\mathcal{A}(k)$ holds for some integers k, and involve proving that it also holds for a **larger** value of k), but it is impossible to traverse all the integers by starting at some integer g and going up (you will never get to $g-1$ this way). In contrast, the following induction principle includes both an "upwards" and a "downwards" induction step, which makes it suited for proving statements about all integers:

Theorem 2.149. Let $g \in \mathbb{Z}$. Let $\mathbb{Z}_{\leq g}$ be the set $\{g, g-1, g-2, \ldots\}$ (that is, the set of all integers that are $\leq g$).

For each $n \in \mathbb{Z}$, let $\mathcal{A}(n)$ be a logical statement.

Assume the following:

> *Assumption 1:* The statement $\mathcal{A}(g)$ holds.
>
> *Assumption 2:* If $m \in \mathbb{Z}_{\geq g}$ is such that $\mathcal{A}(m)$ holds, then $\mathcal{A}(m+1)$ also holds.
>
> *Assumption 3:* If $m \in \mathbb{Z}_{\leq g}$ is such that $\mathcal{A}(m)$ holds, then $\mathcal{A}(m-1)$ also holds.
>
> Then, $\mathcal{A}(n)$ holds for each $n \in \mathbb{Z}$.

Theorem 2.149 is known as the *principle of two-sided induction*. Roughly speaking, a proof using Theorem 2.149 will involve two induction steps: one that "goes up" (proving that Assumption 2 holds), and one that "goes down" (proving that Assumption 3 holds). However, in practice, Theorem 2.149 is seldom used, which is why we shall not make any conventions about how to write proofs using Theorem 2.149. We will only give one example for such a proof.

Let us first prove Theorem 2.149 itself:

Proof of Theorem 2.149. Assumptions 1 and 2 of Theorem 2.149 are exactly Assumptions 1 and 2 of Theorem 2.53. Hence, Assumptions 1 and 2 of Theorem 2.53 hold (since Assumptions 1 and 2 of Theorem 2.149 hold). Thus, Theorem 2.53 shows that

$$\mathcal{A}(n) \text{ holds for each } n \in \mathbb{Z}_{\geq g}. \tag{207}$$

On the other hand, for each $n \in \mathbb{Z}$, we define a logical statement $\mathcal{B}(n)$ by $\mathcal{B}(n) = \mathcal{A}(2g-n)$. We shall now consider the Assumptions A and B of Corollary 2.61.

The definition of $\mathcal{B}(g)$ yields $\mathcal{B}(g) = \mathcal{A}(2g-g) = \mathcal{A}(g)$ (since $2g-g=g$). Hence, the statement $\mathcal{B}(g)$ holds (since the statement $\mathcal{A}(g)$ holds (by Assumption 1)). In other words, Assumption A is satisfied.

Next, let $p \in \mathbb{Z}_{\geq g}$ be such that $\mathcal{B}(p)$ holds. We shall show that $\mathcal{B}(p+1)$ holds.

Indeed, we have $\mathcal{B}(p) = \mathcal{A}(2g-p)$ (by the definition of $\mathcal{B}(p)$). Thus, $\mathcal{A}(2g-p)$ holds (since $\mathcal{B}(p)$ holds). But $p \in \mathbb{Z}_{\geq g}$; hence, p is an integer that is $\geq g$. Thus, $p \geq g$, so that $2g - \underbrace{p}_{\geq g} \leq 2g - g = g$. Hence, $2g-p$ is an integer that is $\leq g$. In other words, $2g-p \in \mathbb{Z}_{\leq g}$. Therefore, Assumption 3 (applied to $m = 2g-p$) shows that $\mathcal{A}(2g-p-1)$ also holds (since $\mathcal{A}(2g-p)$ holds). But the definition of $\mathcal{B}(p+1)$ yields $\mathcal{B}(p+1) = \mathcal{A}\left(\underbrace{2g-(p+1)}_{=2g-p-1}\right) = \mathcal{A}(2g-p-1)$. Hence, $\mathcal{B}(p+1)$ holds (since $\mathcal{A}(2g-p-1)$ holds).

Now, forget that we fixed p. We thus have shown that if $p \in \mathbb{Z}_{\geq g}$ is such that $\mathcal{B}(p)$ holds, then $\mathcal{B}(p+1)$ also holds. In other words, Assumption B is satisfied.

We now have shown that both Assumptions A and B are satisfied. Hence, Corollary 2.61 shows that

$$\mathcal{B}(n) \text{ holds for each } n \in \mathbb{Z}_{\geq g}. \tag{208}$$

Now, let $n \in \mathbb{Z}$. We shall prove that $\mathcal{A}(n)$ holds.

Indeed, we have either $n \geq g$ or $n < g$. Hence, we are in one of the following two cases:

Case 1: We have $n \geq g$.

Case 2: We have $n < g$.

Let us first consider Case 1. In this case, we have $n \geq g$. Hence, $n \in \mathbb{Z}_{\geq g}$ (since n is an integer). Thus, (207) shows that $\mathcal{A}(n)$ holds. We thus have proven that $\mathcal{A}(n)$ holds in Case 1.

Let us now consider Case 2. In this case, we have $n < g$. Thus, $n \leq g$. Hence, $2g - \underbrace{n}_{\leq g} \geq 2g - g = g$. Thus, $2g - n \in \mathbb{Z}_{\geq g}$ (since $2g - n$ is an integer). Hence, (208) (applied to $2g - n$ instead of n) shows that $\mathcal{B}(2g-n)$ holds. But the definition of $\mathcal{B}(2g-n)$ yields $\mathcal{B}(2g-n) = \mathcal{A}\left(\underbrace{2g-(2g-n)}_{=n}\right) = \mathcal{A}(n)$. Hence, $\mathcal{A}(n)$ holds (since $\mathcal{B}(2g-n)$ holds). Thus, we have proven that $\mathcal{A}(n)$ holds in Case 2.

We now have shown that $\mathcal{A}(n)$ holds in each of the two Cases 1 and 2. Since these two Cases cover all possibilities, we thus conclude that $\mathcal{A}(n)$ always holds.

Now, forget that we fixed n. We thus have proven that $\mathcal{A}(n)$ holds for each $n \in \mathbb{Z}$. This proves Theorem 2.149. □

As an example for the use of Theorem 2.149, we shall prove the following fact:

Proposition 2.150. Let N be a positive integer. For each $n \in \mathbb{Z}$, there exist $q \in \mathbb{Z}$ and $r \in \{0, 1, \ldots, N-1\}$ such that $n = qN + r$.

We shall soon see (in Theorem 2.153) that these q and r are actually uniquely determined by N and n; they are called the *quotient* and the *remainder of the division of n by N*. This is fundamental to all of number theory.

Example 2.151. If we apply Proposition 2.150 to $N = 4$ and $n = 10$, then we conclude that there exist $q \in \mathbb{Z}$ and $r \in \{0, 1, 2, 3\}$ such that $10 = q \cdot 4 + r$. And indeed, such q and r can easily be found ($q = 2$ and $r = 2$).

Proof of Proposition 2.150. First, we notice that $N - 1 \in \mathbb{N}$ (since N is a positive integer). Hence, $0 \in \{0, 1, \ldots, N-1\}$ and $N - 1 \in \{0, 1, \ldots, N-1\}$.

For each $n \in \mathbb{Z}$, we let $\mathcal{A}(n)$ be the statement

$$\left(\text{there exist } q \in \mathbb{Z} \text{ and } r \in \{0, 1, \ldots, N-1\} \text{ such that } n = qN + r\right).$$

Let $g = 0$; thus, $g \in \mathbb{Z}$. Define the set $\mathbb{Z}_{\leq g}$ as in Theorem 2.149. (Thus, $\mathbb{Z}_{\leq g} = \{g, g-1, g-2, \ldots\} = \{0, -1, -2, \ldots\}$ is the set of all nonpositive integers.) We shall now show that Assumptions 1, 2 and 3 of Theorem 2.149 are satisfied.

[*Proof that Assumption 1 is satisfied:* We have $0 \in \{0, 1, \ldots, N-1\}$ and $0 = 0N + 0$. Hence, there exist $q \in \mathbb{Z}$ and $r \in \{0, 1, \ldots, N-1\}$ such that $0 = qN + r$ (namely,

$q = 0$ and $r = 0$). In other words, the statement $\mathcal{A}(0)$ holds[89]. In other words, the statement $\mathcal{A}(g)$ holds (since $g = 0$). In other words, Assumption 1 is satisfied.]

[*Proof that Assumption 2 is satisfied:* Let $m \in \mathbb{Z}_{\geq g}$ be such that $\mathcal{A}(m)$ holds. We shall show that $\mathcal{A}(m+1)$ also holds.

We know that $\mathcal{A}(m)$ holds. In other words, there exist $q \in \mathbb{Z}$ and $r \in \{0, 1, \ldots, N-1\}$ such that $m = qN + r$ [90]. Consider these q and r, and denote them by q_0 and r_0. Thus, $q_0 \in \mathbb{Z}$ and $r_0 \in \{0, 1, \ldots, N-1\}$ and $m = q_0 N + r_0$.

Now, we are in one of the following two cases:

Case 1: We have $r_0 = N - 1$.

Case 2: We have $r_0 \neq N - 1$.

Let us first consider Case 1. In this case, we have $r_0 = N - 1$. Hence, $r_0 + 1 = N$. Now,

$$\underbrace{m}_{=q_0 N + r_0} + 1 = q_0 N + \underbrace{r_0 + 1}_{=N} = q_0 N + N = (q_0 + 1) N = (q_0 + 1) N + 0.$$

Since $0 \in \{0, 1, \ldots, N-1\}$, we can therefore conclude that there exist $q \in \mathbb{Z}$ and $r \in \{0, 1, \ldots, N-1\}$ such that $m + 1 = qN + r$ (namely, $q = q_0 + 1$ and $r = 0$). In other words, the statement $\mathcal{A}(m+1)$ holds[91]. Thus, we have proven that $\mathcal{A}(m+1)$ holds in Case 1.

Let us next consider Case 2. In this case, we have $r_0 \neq N - 1$. Combining $r_0 \in \{0, 1, \ldots, N-1\}$ with $r_0 \neq N - 1$, we obtain

$$r_0 \in \{0, 1, \ldots, N-1\} \setminus \{N-1\} = \{0, 1, \ldots, N-2\},$$

so that $r_0 + 1 \in \{1, 2, \ldots, N-1\} \subseteq \{0, 1, \ldots, N-1\}$. Also,

$$\underbrace{m}_{=q_0 N + r_0} + 1 = q_0 N + r_0 + 1 = q_0 N + (r_0 + 1).$$

Since $r_0 + 1 \in \{0, 1, \ldots, N-1\}$, we can therefore conclude that there exist $q \in \mathbb{Z}$ and $r \in \{0, 1, \ldots, N-1\}$ such that $m + 1 = qN + r$ (namely, $q = q_0$ and $r = r_0 + 1$). In other words, the statement $\mathcal{A}(m+1)$ holds[92]. Thus, we have proven that $\mathcal{A}(m+1)$ holds in Case 2.

We have now proven that $\mathcal{A}(m+1)$ holds in each of the two Cases 1 and 2. Since these two Cases cover all possibilities, we thus conclude that $\mathcal{A}(m+1)$ always holds.

[89] since the statement $\mathcal{A}(0)$ is defined as
$$(\text{there exist } q \in \mathbb{Z} \text{ and } r \in \{0, 1, \ldots, N-1\} \text{ such that } 0 = qN + r)$$

[90] since the statement $\mathcal{A}(m)$ is defined as
$$(\text{there exist } q \in \mathbb{Z} \text{ and } r \in \{0, 1, \ldots, N-1\} \text{ such that } m = qN + r)$$

[91] since the statement $\mathcal{A}(m+1)$ is defined as
$$(\text{there exist } q \in \mathbb{Z} \text{ and } r \in \{0, 1, \ldots, N-1\} \text{ such that } m + 1 = qN + r)$$

[92] since the statement $\mathcal{A}(m+1)$ is defined as
$$(\text{there exist } q \in \mathbb{Z} \text{ and } r \in \{0, 1, \ldots, N-1\} \text{ such that } m + 1 = qN + r)$$

Now, forget that we fixed m. We thus have shown that if $m \in \mathbb{Z}_{\geq g}$ is such that $\mathcal{A}(m)$ holds, then $\mathcal{A}(m+1)$ also holds. In other words, Assumption 2 is satisfied.]

[*Proof that Assumption 3 is satisfied:* Let $m \in \mathbb{Z}_{\leq g}$ be such that $\mathcal{A}(m)$ holds. We shall show that $\mathcal{A}(m-1)$ also holds.

We know that $\mathcal{A}(m)$ holds. In other words, there exist $q \in \mathbb{Z}$ and $r \in \{0,1,\ldots,N-1\}$ such that $m = qN + r$ [93]. Consider these q and r, and denote them by q_0 and r_0. Thus, $q_0 \in \mathbb{Z}$ and $r_0 \in \{0,1,\ldots,N-1\}$ and $m = q_0N + r_0$.

Now, we are in one of the following two cases:

Case 1: We have $r_0 = 0$.

Case 2: We have $r_0 \neq 0$.

Let us first consider Case 1. In this case, we have $r_0 = 0$. Now,

$$\begin{aligned} \underbrace{m}_{=q_0N+r_0} -1 &= q_0N + \underbrace{r_0}_{=0} -1 = q_0N - 1 = \underbrace{q_0N - N}_{=(q_0-1)N} + (N-1) \\ &= (q_0-1)N + (N-1). \end{aligned}$$

Since $N-1 \in \{0,1,\ldots,N-1\}$, we can therefore conclude that there exist $q \in \mathbb{Z}$ and $r \in \{0,1,\ldots,N-1\}$ such that $m-1 = qN+r$ (namely, $q = q_0 - 1$ and $r = N-1$). In other words, the statement $\mathcal{A}(m-1)$ holds[94]. Thus, we have proven that $\mathcal{A}(m-1)$ holds in Case 1.

Let us next consider Case 2. In this case, we have $r_0 \neq 0$. Combining $r_0 \in \{0,1,\ldots,N-1\}$ with $r_0 \neq 0$, we obtain

$$r_0 \in \{0,1,\ldots,N-1\} \setminus \{0\} = \{1,2,\ldots,N-1\},$$

so that $r_0 - 1 \in \{0,1,\ldots,N-2\} \subseteq \{0,1,\ldots,N-1\}$. Also,

$$\underbrace{m}_{=q_0N+r_0} -1 = q_0N + r_0 - 1 = q_0N + (r_0 - 1).$$

Since $r_0 - 1 \in \{0,1,\ldots,N-1\}$, we can therefore conclude that there exist $q \in \mathbb{Z}$ and $r \in \{0,1,\ldots,N-1\}$ such that $m-1 = qN+r$ (namely, $q = q_0$ and $r = r_0 - 1$). In other words, the statement $\mathcal{A}(m-1)$ holds[95]. Thus, we have proven that $\mathcal{A}(m-1)$ holds in Case 2.

We have now proven that $\mathcal{A}(m-1)$ holds in each of the two Cases 1 and 2. Since these two Cases cover all possibilities, we thus conclude that $\mathcal{A}(m-1)$ always holds.

Now, forget that we fixed m. We thus have shown that if $m \in \mathbb{Z}_{\leq g}$ is such that $\mathcal{A}(m)$ holds, then $\mathcal{A}(m-1)$ also holds. In other words, Assumption 3 is satisfied.]

[93] since the statement $\mathcal{A}(m)$ is defined as
(there exist $q \in \mathbb{Z}$ and $r \in \{0,1,\ldots,N-1\}$ such that $m = qN + r$)

[94] since the statement $\mathcal{A}(m-1)$ is defined as
(there exist $q \in \mathbb{Z}$ and $r \in \{0,1,\ldots,N-1\}$ such that $m - 1 = qN + r$)

[95] since the statement $\mathcal{A}(m-1)$ is defined as
(there exist $q \in \mathbb{Z}$ and $r \in \{0,1,\ldots,N-1\}$ such that $m - 1 = qN + r$)

We have now shown that all three Assumptions 1, 2 and 3 of Theorem 2.149 are satisfied. Thus, Theorem 2.149 yields that

$$\mathcal{A}(n) \text{ holds for each } n \in \mathbb{Z}. \tag{209}$$

Now, let $n \in \mathbb{Z}$. Then, $\mathcal{A}(n)$ holds (by (209)). In other words, there exist $q \in \mathbb{Z}$ and $r \in \{0, 1, \ldots, N-1\}$ such that $n = qN + r$ (because the statement $\mathcal{A}(n)$ is defined as (there exist $q \in \mathbb{Z}$ and $r \in \{0, 1, \ldots, N-1\}$ such that $n = qN + r$)). This proves Proposition 2.150. □

2.15.2. Division with remainder

We need one more lemma:

> **Lemma 2.152.** Let N be a positive integer. Let $(q_1, r_1) \in \mathbb{Z} \times \{0, 1, \ldots, N-1\}$ and $(q_2, r_2) \in \mathbb{Z} \times \{0, 1, \ldots, N-1\}$. Assume that $q_1 N + r_1 = q_2 N + r_2$. Then, $(q_1, r_1) = (q_2, r_2)$.

Proof of Lemma 2.152. We have $(q_1, r_1) \in \mathbb{Z} \times \{0, 1, \ldots, N-1\}$. Thus, $q_1 \in \mathbb{Z}$ and $r_1 \in \{0, 1, \ldots, N-1\}$. Similarly, $q_2 \in \mathbb{Z}$ and $r_2 \in \{0, 1, \ldots, N-1\}$.

From $r_1 \in \{0, 1, \ldots, N-1\}$, we obtain $r_1 \geq 0$. From $r_2 \in \{0, 1, \ldots, N-1\}$, we obtain $r_2 \leq N-1$. Hence, $\underbrace{r_2}_{\leq N-1} - \underbrace{r_1}_{\geq 0} \leq (N-1) - 0 = N-1 < N$.

Next, we shall prove that

$$q_1 \leq q_2. \tag{210}$$

[*Proof of (210):* Assume the contrary. Thus, $q_1 > q_2$, so that $q_1 - q_2 > 0$. Hence, $q_1 - q_2 \geq 1$ (since $q_1 - q_2$ is an integer). Therefore, $q_1 - q_2 - 1 \geq 0$, so that $N(q_1 - q_2 - 1) \geq 0$ (because $N > 0$ and $q_1 - q_2 - 1 \geq 0$).

But $q_1 N + r_1 = q_2 N + r_2$, so that $q_2 N + r_2 = q_1 N + r_1$. Hence,

$$r_2 - r_1 = q_1 N - q_2 N = N(q_1 - q_2) = \underbrace{N(q_1 - q_2 - 1)}_{\geq 0} + N \geq N.$$

This contradicts $r_2 - r_1 < N$. This contradiction shows that our assumption was wrong. Hence, (210) is proven.]

Thus, we have proven that $q_1 \leq q_2$. The same argument (with the roles of (q_1, r_1) and (q_2, r_2) interchanged) shows that $q_2 \leq q_1$. Combining the inequalities $q_1 \leq q_2$ and $q_2 \leq q_1$, we obtain $q_1 = q_2$.

Also, $q_2 N + r_2 = \underbrace{q_1}_{=q_2} N + r_1 = q_2 N + r_1$. Subtracting $q_2 N$ from both sides of this equality, we find $r_2 = r_1$. Hence, $r_1 = r_2$.

Thus, $\left(\underbrace{q_1}_{=q_2}, \underbrace{r_1}_{=r_2}\right) = (q_2, r_2)$. This proves Lemma 2.152. □

As we have already mentioned, Proposition 2.150 is just a part of a crucial result from number theory:

Theorem 2.153. Let N be a positive integer. Let $n \in \mathbb{Z}$. Then, there is a unique pair $(q, r) \in \mathbb{Z} \times \{0, 1, \ldots, N-1\}$ such that $n = qN + r$.

Proving this theorem will turn out rather easy, since we have already done the hard work with our proofs of Proposition 2.150 and Lemma 2.152:

Proof of Theorem 2.153. Proposition 2.150 shows that there exist $q \in \mathbb{Z}$ and $r \in \{0, 1, \ldots, N-1\}$ such that $n = qN + r$. Consider these q and r, and denote them by q_0 and r_0. Thus, $q_0 \in \mathbb{Z}$ and $r_0 \in \{0, 1, \ldots, N-1\}$ and $n = q_0 N + r_0$. From $q_0 \in \mathbb{Z}$ and $r_0 \in \{0, 1, \ldots, N-1\}$, we obtain $(q_0, r_0) \in \mathbb{Z} \times \{0, 1, \ldots, N-1\}$. Hence, there exists **at least one** pair $(q, r) \in \mathbb{Z} \times \{0, 1, \ldots, N-1\}$ such that $n = qN + r$ (namely, $(q, r) = (q_0, r_0)$).

Now, let (q_1, r_1) and (q_2, r_2) be two pairs $(q, r) \in \mathbb{Z} \times \{0, 1, \ldots, N-1\}$ such that $n = qN + r$. We shall prove that $(q_1, r_1) = (q_2, r_2)$.

We have assumed that (q_1, r_1) is a pair $(q, r) \in \mathbb{Z} \times \{0, 1, \ldots, N-1\}$ such that $n = qN + r$. In other words, (q_1, r_1) is a pair in $\mathbb{Z} \times \{0, 1, \ldots, N-1\}$ and satisfies $n = q_1 N + r_1$. Similarly, (q_2, r_2) is a pair in $\mathbb{Z} \times \{0, 1, \ldots, N-1\}$ and satisfies $n = q_2 N + r_2$.

Hence, $q_1 N + r_1 = n = q_2 N + r_2$. Thus, Lemma 2.152 yields $(q_1, r_1) = (q_2, r_2)$.

Let us now forget that we fixed (q_1, r_1) and (q_2, r_2). We thus have shown that if (q_1, r_1) and (q_2, r_2) are two pairs $(q, r) \in \mathbb{Z} \times \{0, 1, \ldots, N-1\}$ such that $n = qN + r$, then $(q_1, r_1) = (q_2, r_2)$. In other words, any two pairs $(q, r) \in \mathbb{Z} \times \{0, 1, \ldots, N-1\}$ such that $n = qN + r$ must be equal. In other words, there exists **at most one** pair $(q, r) \in \mathbb{Z} \times \{0, 1, \ldots, N-1\}$ such that $n = qN + r$. Since we also know that there exists **at least one** such pair, we can therefore conclude that there exists **exactly one** such pair. In other words, there is a unique pair $(q, r) \in \mathbb{Z} \times \{0, 1, \ldots, N-1\}$ such that $n = qN + r$. This proves Theorem 2.153. □

Definition 2.154. Let N be a positive integer. Let $n \in \mathbb{Z}$. Theorem 2.153 says that there is a unique pair $(q, r) \in \mathbb{Z} \times \{0, 1, \ldots, N-1\}$ such that $n = qN + r$. Consider this pair (q, r). Then, q is called the *quotient of the division of n by N* (or the *quotient obtained when n is divided by N*), whereas r is called the *remainder of the division of n by N* (or the *remainder obtained when n is divided by N*).

For example, the quotient of the division of 7 by 3 is 2, whereas the remainder of the division of 7 by 3 is 1 (because $(2, 1)$ is a pair in $\mathbb{Z} \times \{0, 1, 2\}$ such that $7 = 2 \cdot 3 + 1$).

We collect some basic properties of remainders:

Corollary 2.155. Let N be a positive integer. Let $n \in \mathbb{Z}$. Let $n\%N$ denote the remainder of the division of n by N.

(a) Then, $n\%N \in \{0, 1, \ldots, N-1\}$ and $n\%N \equiv n \bmod N$.

(b) We have $N \mid n$ if and only if $n\%N = 0$.
(c) Let $c \in \{0, 1, \ldots, N-1\}$ be such that $c \equiv n \mod N$. Then, $c = n\%N$.

Proof of Corollary 2.155. Theorem 2.153 says that there is a unique pair $(q, r) \in \mathbb{Z} \times \{0, 1, \ldots, N-1\}$ such that $n = qN + r$. Consider this pair (q, r). Then, the remainder of the division of n by N is r (because this is how this remainder was defined). In other words, $n\%N$ is r (since $n\%N$ is the remainder of the division of n by N). Thus, $n\%N = r$. But $N \mid qN$ (since q is an integer), so that $qN \equiv 0 \mod N$. Hence, $\underbrace{qN}_{\equiv 0 \mod N} + r \equiv 0 + r = r \mod N$. Hence, $r \equiv qN + r = n \mod N$, so that $n\%N = r \equiv n \mod N$. Furthermore, $n\%N = r \in \{0, 1, \ldots, N-1\}$ (since $(q, r) \in \mathbb{Z} \times \{0, 1, \ldots, N-1\}$). This completes the proof of Corollary 2.155 **(a)**.

(b) We have the following implication:

$$(N \mid n) \Longrightarrow (n\%N = 0) . \tag{211}$$

[*Proof of (211):* Assume that $N \mid n$. We must prove that $n\%N = 0$.

We have $N \mid n$. In other words, there exists some integer w such that $n = Nw$. Consider this w.

We have $N - 1 \in \mathbb{N}$ (since N is a positive integer), thus $0 \in \{0, 1, \ldots, N-1\}$. From $w \in \mathbb{Z}$ and $0 \in \{0, 1, \ldots, N-1\}$, we obtain $(w, 0) \in \mathbb{Z} \times \{0, 1, \ldots, N-1\}$. Also, $wN + 0 = wN = Nw = n = qN + r$. Hence, Lemma 2.152 (applied to $(q_1, r_1) = (w, 0)$ and $(q_2, r_2) = (q, r)$) yields $(w, 0) = (q, r)$. In other words, $w = q$ and $0 = r$. Hence, $r = 0$, so that $n\%N = r = 0$. This proves the implication (211).]

Next, we have the following implication:

$$(n\%N = 0) \Longrightarrow (N \mid n) . \tag{212}$$

[*Proof of (212):* Assume that $n\%N = 0$. We must prove that $N \mid n$.

We have $n = qN + \underbrace{r}_{= n\%N = 0} = qN$. Thus, $N \mid n$. This proves the implication (212).]

Combining the two implications (211) and (212), we obtain the logical equivalence $(N \mid n) \iff (n\%N = 0)$. In other words, we have $N \mid n$ if and only if $n\%N = 0$. This proves Corollary 2.155 **(b)**.

(c) We have $c \equiv n \mod N$. In other words, $N \mid c - n$. In other words, there exists some integer w such that $c - n = Nw$. Consider this w.

From $-w \in \mathbb{Z}$ and $c \in \{0, 1, \ldots, N-1\}$, we obtain $(-w, c) \in \mathbb{Z} \times \{0, 1, \ldots, N-1\}$. Also, from $c - n = Nw$, we obtain $n = c - Nw = (-w) N + c$, so that $(-w) N + c = n = qN + r$. Hence, Lemma 2.152 (applied to $(q_1, r_1) = (-w, c)$ and $(q_2, r_2) = (q, r)$) yields $(-w, c) = (q, r)$. In other words, $-w = q$ and $c = r$. Hence, $c = r = n\%N$. This proves Corollary 2.155 **(c)**. □

Note that parts **(a)** and **(c)** of Corollary 2.155 (taken together) characterize the remainder $n\%N$ as the unique element of $\{0, 1, \ldots, N-1\}$ that is congruent to n modulo N. Corollary 2.155 **(b)** provides a simple algorithm to check whether a given integer n is divisible by a given positive integer N; namely, it suffices to compute the remainder $n\%N$ and check whether $n\%N = 0$.

Let us further illustrate the usefulness of Theorem 2.153 by proving a fundamental property of odd numbers. Recall the following standard definitions:

Definition 2.156. Let $n \in \mathbb{Z}$.

(a) We say that the integer n is *even* if and only if n is divisible by 2.

(b) We say that the integer n is *odd* if and only if n is not divisible by 2.

This definition shows that any integer n is either even or odd (but not both at the same time).

It is clear that an integer n is even if and only if it can be written in the form $n = 2m$ for some $m \in \mathbb{Z}$. Moreover, this m is unique (because $n = 2m$ implies $m = n/2$). Let us prove a similar property for odd numbers:

Proposition 2.157. Let $n \in \mathbb{Z}$.

(a) The integer n is odd if and only if n can be written in the form $n = 2m + 1$ for some $m \in \mathbb{Z}$.

(b) This m is unique if it exists. (That is, any two integers $m \in \mathbb{Z}$ satisfying $n = 2m + 1$ must be equal.)

We shall use Theorem 2.153 several times in the below proof (far more than necessary), mostly to illustrate how it can be applied.

Proof of Proposition 2.157. **(a)** Let us first prove the logical implication

$$(n \text{ is odd}) \implies (\text{there exists an } m \in \mathbb{Z} \text{ such that } n = 2m+1). \tag{213}$$

[*Proof of (213):* Assume that n is odd. We must prove that there exists an $m \in \mathbb{Z}$ such that $n = 2m + 1$.

Theorem 2.153 (applied to $N = 2$) yields that there is a unique pair $(q, r) \in \mathbb{Z} \times \{0, 1, \ldots, 2-1\}$ such that $n = q \cdot 2 + r$. Consider this (q, r). From $(q, r) \in \mathbb{Z} \times \{0, 1, \ldots, 2-1\}$, we obtain $q \in \mathbb{Z}$ and $r \in \{0, 1, \ldots, 2-1\} = \{0, 1\}$.

We know that n is odd; in other words, n is not divisible by 2 (by the definition of "odd"). If we had $n = 2q$, then n would be divisible by 2, which would contradict the fact that n is not divisible by 2. Hence, we cannot have $n = 2q$. If we had $r = 0$, then we would have $n = \underbrace{q \cdot 2}_{=2q} + \underbrace{r}_{=0} = 2q$, which would contradict the fact that we cannot have $n = 2q$. Hence, we cannot have $r = 0$. Thus, $r \neq 0$.

Combining $r \in \{0, 1\}$ with $r \neq 0$, we obtain $r \in \{0, 1\} \setminus \{0\} = \{1\}$. Thus, $r = 1$. Hence, $n = \underbrace{q \cdot 2}_{=2q} + \underbrace{r}_{=1} = 2q + 1$. Thus, there exists an $m \in \mathbb{Z}$ such that $n = 2m + 1$ (namely, $m = q$). This proves the implication (213).]

Next, we shall prove the logical implication

$$(\text{there exists an } m \in \mathbb{Z} \text{ such that } n = 2m+1) \implies (n \text{ is odd}). \tag{214}$$

[*Proof of (214):* Assume that there exists an $m \in \mathbb{Z}$ such that $n = 2m + 1$. We must prove that n is odd.

We have assumed that there exists an $m \in \mathbb{Z}$ such that $n = 2m + 1$. Consider this m. Thus, the pair $(m, 1)$ belongs to $\mathbb{Z} \times \{0, 1, \ldots, 2 - 1\}$ (since $m \in \mathbb{Z}$ and $1 \in \{0, 1, \ldots, 2 - 1\}$) and satisfies $n = m \cdot 2 + 1$ (since $n = \underbrace{2m}_{= m \cdot 2} + 1 = m \cdot 2 + 1$). In other words, the pair $(m, 1)$ is a pair $(q, r) \in \mathbb{Z} \times \{0, 1, \ldots, 2 - 1\}$ such that $n = q \cdot 2 + r$.

Now, assume (for the sake of contradiction) that n is divisible by 2. Thus, there exists some integer w such that $n = 2w$. Consider this w. Thus, the pair $(w, 0)$ belongs to $\mathbb{Z} \times \{0, 1, \ldots, 2 - 1\}$ (since $w \in \mathbb{Z}$ and $0 \in \{0, 1, \ldots, 2 - 1\}$) and satisfies $n = w \cdot 2 + 0$ (since $n = 2w = w \cdot 2 = w \cdot 2 + 0$). In other words, the pair $(w, 0)$ is a pair $(q, r) \in \mathbb{Z} \times \{0, 1, \ldots, 2 - 1\}$ such that $n = q \cdot 2 + r$.

Theorem 2.153 (applied to $N = 2$) yields that there is a unique pair $(q, r) \in \mathbb{Z} \times \{0, 1, \ldots, 2 - 1\}$ such that $n = q \cdot 2 + r$. Thus, there exists **at most** one such pair. In other words, any two such pairs must be equal. Hence, the two pairs $(m, 1)$ and $(w, 0)$ must be equal (since $(m, 1)$ and $(w, 0)$ are two pairs $(q, r) \in \mathbb{Z} \times \{0, 1, \ldots, 2 - 1\}$ such that $n = q \cdot 2 + r$). In other words, $(m, 1) = (w, 0)$. In other words, $m = w$ and $1 = 0$. But $1 = 0$ is clearly absurd. Thus, we have obtained a contradiction. This shows that our assumption (that n is divisible by 2) was wrong. Hence, n is not divisible by 2. In other words, n is odd (by the definition of "odd"). This proves the implication (214).]

Combining the two implications (213) and (214), we obtain the logical equivalence

$$\begin{aligned} (n \text{ is odd}) &\iff (\text{there exists an } m \in \mathbb{Z} \text{ such that } n = 2m + 1) \\ &\iff (n \text{ can be written in the form } n = 2m + 1 \text{ for some } m \in \mathbb{Z}) . \end{aligned}$$

In other words, the integer n is odd if and only if n can be written in the form $n = 2m + 1$ for some $m \in \mathbb{Z}$. This proves Proposition 2.157 **(a)**.

(b) This is easy to prove in any way, but let us prove this using Theorem 2.153 just in order to illustrate the use of the latter theorem.

We must prove that any two integers $m \in \mathbb{Z}$ satisfying $n = 2m + 1$ must be equal.

Let m_1 and m_2 be two integers $m \in \mathbb{Z}$ satisfying $n = 2m + 1$. We shall show that $m_1 = m_2$.

We know that m_1 is an integer $m \in \mathbb{Z}$ satisfying $n = 2m + 1$. In other words, m_1 is an integer in $\mathbb{Z}$ and satisfies $n = 2m_1 + 1$. Thus, the pair $(m_1, 1)$ belongs to $\mathbb{Z} \times \{0, 1, \ldots, 2 - 1\}$ (since $m_1 \in \mathbb{Z}$ and $1 \in \{0, 1, \ldots, 2 - 1\}$) and satisfies $n = m_1 \cdot 2 + 1$ (since $n = \underbrace{2m_1}_{= m_1 \cdot 2} + 1 = m_1 \cdot 2 + 1$). In other words, the pair $(m_1, 1)$ is a pair $(q, r) \in \mathbb{Z} \times \{0, 1, \ldots, 2 - 1\}$ such that $n = q \cdot 2 + r$. The same argument (applied to m_2 instead of m_1) shows that $(m_2, 1)$ is a pair $(q, r) \in \mathbb{Z} \times \{0, 1, \ldots, 2 - 1\}$ such that $n = q \cdot 2 + r$.

Theorem 2.153 (applied to $N = 2$) yields that there is a unique pair $(q, r) \in \mathbb{Z} \times \{0, 1, \ldots, 2 - 1\}$ such that $n = q \cdot 2 + r$. Thus, there exists **at most** one such pair. In other words, any two such pairs must be equal. Hence, the two pairs $(m_1, 1)$ and $(m_2, 1)$ must be equal (since $(m_1, 1)$ and $(m_2, 1)$ are two pairs $(q, r) \in \mathbb{Z} \times \{0, 1, \ldots, 2-1\}$ such that $n = q \cdot 2 + r$). In other words, $(m_1, 1) = (m_2, 1)$. In other words, $m_1 = m_2$ and $1 = 1$. Hence, we have shown that $m_1 = m_2$.

Now, forget that we fixed m_1 and m_2. We thus have proven that if m_1 and m_2 are two integers $m \in \mathbb{Z}$ satisfying $n = 2m + 1$, then $m_1 = m_2$. In other words, any two integers $m \in \mathbb{Z}$ satisfying $n = 2m + 1$ must be equal. In other words, the m in Proposition 2.157 **(a)** is unique. This proves Proposition 2.157 **(b)**. □

We can use this to obtain the following fundamental fact:

Corollary 2.158. Let $n \in \mathbb{Z}$.

(a) If n is even, then $(-1)^n = 1$.

(b) If n is odd, then $(-1)^n = -1$.

Proof of Corollary 2.158. **(a)** Assume that n is even. In other words, n is divisible by 2 (by the definition of "even"). In other words, $2 \mid n$. In other words, there exists an integer w such that $n = 2w$. Consider this w. From $n = 2w$, we obtain $(-1)^n = (-1)^{2w} = \left(\underbrace{(-1)^2}_{=1}\right)^w = 1^w = 1$. This proves Corollary 2.158 **(a)**.

(b) Assume that n is odd. Proposition 2.157 **(a)** shows that the integer n is odd if and only if n can be written in the form $n = 2m + 1$ for some $m \in \mathbb{Z}$. Hence, n can be written in the form $n = 2m + 1$ for some $m \in \mathbb{Z}$ (since the integer n is odd). Consider this m. From $n = 2m + 1$, we obtain

$$(-1)^n = (-1)^{2m+1} = (-1)^{2m}(-1) = -\underbrace{(-1)^{2m}}_{=\left((-1)^2\right)^m} = -\left(\underbrace{(-1)^2}_{=1}\right)^m = -\underbrace{1^m}_{=1} = -1.$$

This proves Corollary 2.158 **(b)**. □

Let us state two more fundamental facts, which are proven in Exercise 2.4:

Proposition 2.159. Let u and v be two integers. Then, we have the following chain of logical equivalences:

$$\begin{aligned}(u \equiv v \bmod 2) &\iff (u \text{ and } v \text{ are either both even or both odd}) \\ &\iff \left((-1)^u = (-1)^v\right).\end{aligned}$$

Proposition 2.160. Let $n \in \mathbb{Z}$.

(a) The integer n is even if and only if $n \equiv 0 \bmod 2$.

(b) The integer n is odd if and only if $n \equiv 1 \bmod 2$.

Exercise 2.4. Prove Proposition 2.159 and Proposition 2.160.

Exercise 2.5. Let N be a positive integer. Let $p \in \mathbb{Z}$ and $h \in \mathbb{Z}$. Prove that there exists a **unique** element $g \in \{p+1, p+2, \ldots, p+N\}$ satisfying $g \equiv h \mod N$.

Exercise 2.6. Let $k \in \mathbb{N}$, $a \in \mathbb{Z}$ and $b \in \mathbb{Z}$.

(a) Prove that $a - b \mid a^k - b^k$.

(b) Assume that k is odd. Prove that $a + b \mid a^k + b^k$.

Exercise 2.7. Fix an **odd** positive integer r. Consider the sequence $(b_0, b_1, b_2, \ldots)$ defined in Proposition 2.66.

Prove that $b_n + 1 \mid b_{n-1}(b_{n+2} + 1)$ for each positive integer n. (Note that the statement "$b_n + 1 \mid b_{n-1}(b_{n+2} + 1)$" makes sense, since Proposition 2.66 **(a)** yields that all three numbers b_n, b_{n-1}, b_{n+2} belong to $\mathbb{N}$.)

2.15.3. Backwards induction principles

When we use Theorem 2.149 to prove a statement, we can regard the proof of Assumption 2 as a (regular) induction step ("forwards induction step"), and regard the proof of Assumption 3 as a sort of "backwards induction step". There are also "backwards induction principles" which include a "backwards induction step" but no "forwards induction step". Here are two such principles:

Theorem 2.161. Let $g \in \mathbb{Z}$. Let $\mathbb{Z}_{\leq g}$ be the set $\{g, g-1, g-2, \ldots\}$ (that is, the set of all integers that are $\leq g$). For each $n \in \mathbb{Z}_{\leq g}$, let $\mathcal{A}(n)$ be a logical statement.

Assume the following:

Assumption 1: The statement $\mathcal{A}(g)$ holds.

Assumption 2: If $m \in \mathbb{Z}_{\leq g}$ is such that $\mathcal{A}(m)$ holds, then $\mathcal{A}(m-1)$ also holds.

Then, $\mathcal{A}(n)$ holds for each $n \in \mathbb{Z}_{\leq g}$.

Theorem 2.162. Let $g \in \mathbb{Z}$ and $h \in \mathbb{Z}$. For each $n \in \{g, g+1, \ldots, h\}$, let $\mathcal{A}(n)$ be a logical statement.

Assume the following:

Assumption 1: If $g \leq h$, then the statement $\mathcal{A}(h)$ holds.

Assumption 2: If $m \in \{g+1, g+2, \ldots, h\}$ is such that $\mathcal{A}(m)$ holds, then $\mathcal{A}(m-1)$ also holds.

Then, $\mathcal{A}(n)$ holds for each $n \in \{g, g+1, \ldots, h\}$.

Theorem 2.161 is an analogue of Theorem 2.53, while Theorem 2.162 is an analogue of Theorem 2.74. However, it is not hard to derive these theorems from the induction principles we already know:

Exercise 2.8. Prove Theorem 2.161 and Theorem 2.162.

It is also easy to state and prove a "backwards" analogue of Theorem 2.60. (We leave this to the reader.)

A proof using Theorem 2.161 or using Theorem 2.162 is usually called a *proof by descending induction* or a *proof by backwards induction*.

2.16. Induction from $k-1$ to k

2.16.1. The principle

Let us next show yet another "alternative induction principle", which differs from Theorem 2.53 in a mere notational detail:

Theorem 2.163. Let $g \in \mathbb{Z}$. For each $n \in \mathbb{Z}_{\geq g}$, let $\mathcal{A}(n)$ be a logical statement.

Assume the following:

Assumption 1: The statement $\mathcal{A}(g)$ holds.

Assumption 2: If $k \in \mathbb{Z}_{\geq g+1}$ is such that $\mathcal{A}(k-1)$ holds, then $\mathcal{A}(k)$ also holds.

Then, $\mathcal{A}(n)$ holds for each $n \in \mathbb{Z}_{\geq g}$.

Roughly speaking, this Theorem 2.163 is just Theorem 2.53, except that the variable m in Assumption 2 has been renamed as $k-1$. Consequently, it stands to reason that Theorem 2.163 can easily be derived from Theorem 2.53. Here is the derivation in full detail:

Proof of Theorem 2.163. For each $n \in \mathbb{Z}_{\geq g}$, we define the logical statement $\mathcal{B}(n)$ to be the statement $\mathcal{A}(n)$. Thus, $\mathcal{B}(n) = \mathcal{A}(n)$ for each $n \in \mathbb{Z}_{\geq g}$. Applying this to $n = g$, we obtain $\mathcal{B}(g) = \mathcal{A}(g)$ (since $g \in \mathbb{Z}_{\geq g}$).

We shall now show that the two Assumptions A and B of Corollary 2.61 are satisfied.

Indeed, recall that Assumption 1 is satisfied. In other words, the statement $\mathcal{A}(g)$ holds. In other words, the statement $\mathcal{B}(g)$ holds (since $\mathcal{B}(g) = \mathcal{A}(g)$). In other words, Assumption A is satisfied.

We shall next show that Assumption B is satisfied. Indeed, let $p \in \mathbb{Z}_{\geq g}$ be such that $\mathcal{B}(p)$ holds. Recall that the statement $\mathcal{B}(p)$ was defined to be the statement $\mathcal{A}(p)$. Thus, $\mathcal{B}(p) = \mathcal{A}(p)$. Hence, $\mathcal{A}(p)$ holds (since $\mathcal{B}(p)$ holds). Now, let $k = p+1$. We know that $p \in \mathbb{Z}_{\geq g}$; in other words, p is an integer and satisfies

$p \geq g$. Hence, $k = p + 1$ is an integer as well and satisfies $k = \underbrace{p}_{\geq g} + 1 \geq g + 1$. In other words, $k \in \mathbb{Z}_{\geq g+1}$. Moreover, from $k = p + 1$, we obtain $k - 1 = p$. Hence, $\mathcal{A}(k-1) = \mathcal{A}(p)$. Thus, $\mathcal{A}(k-1)$ holds (since $\mathcal{A}(p)$ holds). Thus, Assumption 2 shows that $\mathcal{A}(k)$ also holds. But the statement $\mathcal{B}(k)$ was defined to be the statement $\mathcal{A}(k)$. Hence, $\mathcal{B}(k) = \mathcal{A}(k)$, so that $\mathcal{A}(k) = \mathcal{B}(k) = \mathcal{B}(p+1)$ (since $k = p + 1$). Thus, the statement $\mathcal{B}(p+1)$ holds (since $\mathcal{A}(k)$ holds). Now, forget that we fixed p. We thus have shown that if $p \in \mathbb{Z}_{\geq g}$ is such that $\mathcal{B}(p)$ holds, then $\mathcal{B}(p+1)$ also holds. In other words, Assumption B is satisfied.

We have now proven that both Assumptions A and B of Corollary 2.61 are satisfied. Hence, Corollary 2.61 shows that $\mathcal{B}(n)$ holds for each $n \in \mathbb{Z}_{\geq g}$. In other words, $\mathcal{A}(n)$ holds for each $n \in \mathbb{Z}_{\geq g}$ (because each $n \in \mathbb{Z}_{\geq g}$ satisfies $\mathcal{B}(n) = \mathcal{A}(n)$ (by the definition of $\mathcal{B}(n)$)). This proves Theorem 2.163. □

Proofs that use Theorem 2.163 are usually called *proofs by induction* or *induction proofs*. As an example of such a proof, let us show the following identity:

Proposition 2.164. For every $n \in \mathbb{N}$, we have

$$\sum_{i=1}^{2n} \frac{(-1)^{i-1}}{i} = \sum_{i=n+1}^{2n} \frac{1}{i}. \tag{215}$$

The equality (215) can be rewritten as

$$\frac{1}{1} - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} \pm \cdots + \frac{1}{2n-1} - \frac{1}{2n} = \frac{1}{n+1} + \frac{1}{n+2} + \cdots + \frac{1}{2n}$$

(where all the signs on the right hand side are + signs, whereas the signs on the left hand side alternate between + signs and − signs).

Proof of Proposition 2.164. For each $n \in \mathbb{Z}_{\geq 0}$, we let $\mathcal{A}(n)$ be the statement

$$\left(\sum_{i=1}^{2n} \frac{(-1)^{i-1}}{i} = \sum_{i=n+1}^{2n} \frac{1}{i} \right).$$

Our next goal is to prove the statement $\mathcal{A}(n)$ for each $n \in \mathbb{Z}_{\geq 0}$.

We first notice that the statement $\mathcal{A}(0)$ holds[96].

[96] *Proof.* We have $\sum_{i=1}^{2\cdot 0} \frac{(-1)^{i-1}}{i} = (\text{empty sum}) = 0$. Comparing this with $\sum_{i=0+1}^{2\cdot 0} \frac{1}{i} = (\text{empty sum}) = 0$, we obtain $\sum_{i=1}^{2\cdot 0} \frac{(-1)^{i-1}}{i} = \sum_{i=0+1}^{2\cdot 0} \frac{1}{i}$. But this is precisely the statement $\mathcal{A}(0)$ (since $\mathcal{A}(0)$ is defined to be the statement $\left(\sum_{i=1}^{2\cdot 0} \frac{(-1)^{i-1}}{i} = \sum_{i=0+1}^{2\cdot 0} \frac{1}{i} \right)$). Hence, the statement $\mathcal{A}(0)$ holds.

Now, we claim that

$$\text{if } k \in \mathbb{Z}_{\geq 0+1} \text{ is such that } \mathcal{A}\left(k-1\right) \text{ holds, then } \mathcal{A}\left(k\right) \text{ also holds.} \tag{216}$$

[*Proof of (216):* Let $k \in \mathbb{Z}_{\geq 0+1}$ be such that $\mathcal{A}(k-1)$ holds. We must show that $\mathcal{A}(k)$ also holds.

We have $k \in \mathbb{Z}_{\geq 0+1}$. Thus, k is an integer and satisfies $k \geq 0+1 = 1$.

We have assumed that $\mathcal{A}(k-1)$ holds. In other words,

$$\sum_{i=1}^{2(k-1)} \frac{(-1)^{i-1}}{i} = \sum_{i=(k-1)+1}^{2(k-1)} \frac{1}{i} \tag{217}$$

holds[97].

We have $(-1)^{2(k-1)} = \left(\underbrace{(-1)^2}_{=1}\right)^{k-1} = 1^{k-1} = 1$. But $2k-1 = 2(k-1)+1$. Thus, $(-1)^{2k-1} = (-1)^{2(k-1)+1} = \underbrace{(-1)^{2(k-1)}}_{=1} \underbrace{(-1)^1}_{=-1} = -1$.

Now, $k \geq 1$, so that $2k \geq 2$ and therefore $2k-1 \geq 1$. Hence, we can split off the addend for $i = 2k-1$ from the sum $\sum_{i=1}^{2k-1} \frac{(-1)^{i-1}}{i}$. We thus obtain

$$\begin{aligned} \sum_{i=1}^{2k-1} \frac{(-1)^{i-1}}{i} &= \sum_{i=1}^{(2k-1)-1} \frac{(-1)^{i-1}}{i} + \frac{(-1)^{(2k-1)-1}}{2k-1} \\ &= \underbrace{\sum_{i=1}^{2(k-1)} \frac{(-1)^{i-1}}{i}}_{\substack{= \sum_{i=(k-1)+1}^{2(k-1)} \frac{1}{i} \\ \text{(by (217))}}} + \underbrace{\frac{(-1)^{2(k-1)}}{2k-1}}_{\substack{= \frac{1}{2k-1} \\ \text{(since } (-1)^{2(k-1)}=1)}} \\ &\qquad \left(\text{since } (2k-1)-1 = 2(k-1)\right) \\ &= \sum_{i=(k-1)+1}^{2(k-1)} \frac{1}{i} + \frac{1}{2k-1} = \sum_{i=k}^{2k-2} \frac{1}{i} + \frac{1}{2k-1} \end{aligned} \tag{218}$$

(since $(k-1)+1 = k$ and $2(k-1) = 2k-2$).

On the other hand, $2k \geq 2 \geq 1$. Hence, we can split off the addend for $i = 2k$

[97] because $\mathcal{A}(k-1)$ is defined to be the statement $\left(\sum_{i=1}^{2(k-1)} \frac{(-1)^{i-1}}{i} = \sum_{i=(k-1)+1}^{2(k-1)} \frac{1}{i}\right)$

from the sum $\sum\limits_{i=1}^{2k} \frac{(-1)^{i-1}}{i}$. We thus obtain

$$\begin{aligned}\sum_{i=1}^{2k} \frac{(-1)^{i-1}}{i} &= \underbrace{\sum_{i=1}^{2k-1} \frac{(-1)^{i-1}}{i}}_{\substack{=\sum\limits_{i=k}^{2k-2} \frac{1}{i}+\frac{1}{2k-1} \\ \text{(by (218))}}} + \underbrace{\frac{(-1)^{2k-1}}{2k}}_{\substack{=\frac{-1}{2k} \\ \text{(since } (-1)^{2k-1}=-1\text{)}}} \\ &= \sum_{i=k}^{2k-2} \frac{1}{i} + \frac{1}{2k-1} + \frac{-1}{2k}. \end{aligned} \tag{219}$$

But we have $(2k-1)-k = k-1 \geq 0$ (since $k \geq 1$). Thus, $2k-1 \geq k$. Thus, we can split off the addend for $i = 2k-1$ from the sum $\sum\limits_{i=k}^{2k-1} \frac{1}{i}$. We thus obtain

$$\sum_{i=k}^{2k-1} \frac{1}{i} = \sum_{i=k}^{(2k-1)-1} \frac{1}{i} + \frac{1}{2k-1} = \sum_{i=k}^{2k-2} \frac{1}{i} + \frac{1}{2k-1} \tag{220}$$

(since $(2k-1)-1 = 2k-2$). Hence, (219) becomes

$$\sum_{i=1}^{2k} \frac{(-1)^{i-1}}{i} = \underbrace{\sum_{i=k}^{2k-2} \frac{1}{i} + \frac{1}{2k-1}}_{\substack{=\sum\limits_{i=k}^{2k-1} \frac{1}{i} \\ \text{(by (220))}}} + \frac{-1}{2k} = \sum_{i=k}^{2k-1} \frac{1}{i} + \frac{-1}{2k}. \tag{221}$$

But we have $k+1 \leq 2k$ (since $2k-(k+1) = k-1 \geq 0$). Thus, we can split off the addend for $i = 2k$ from the sum $\sum\limits_{i=k+1}^{2k} \frac{1}{i}$. We thus obtain

$$\sum_{i=k+1}^{2k} \frac{1}{i} = \sum_{i=k+1}^{2k-1} \frac{1}{i} + \frac{1}{2k}.$$

Hence,

$$\sum_{i=k+1}^{2k-1} \frac{1}{i} = \sum_{i=k+1}^{2k} \frac{1}{i} - \frac{1}{2k}. \tag{222}$$

Also, $k \leq 2k-1$ (since $(2k-1)-k = k-1 \geq 0$). Thus, we can split off the addend for $i = k$ from the sum $\sum\limits_{i=k}^{2k-1} \frac{1}{i}$. We thus obtain

$$\sum_{i=k}^{2k-1} \frac{1}{i} = \frac{1}{k} + \underbrace{\sum_{i=k+1}^{2k-1} \frac{1}{i}}_{\substack{=\sum\limits_{i=k+1}^{2k} \frac{1}{i}-\frac{1}{2k} \\ \text{(by (222))}}} = \frac{1}{k} + \sum_{i=k+1}^{2k} \frac{1}{i} - \frac{1}{2k} = \sum_{i=k+1}^{2k} \frac{1}{i} + \underbrace{\frac{1}{k} - \frac{1}{2k}}_{=\frac{1}{2k}} = \sum_{i=k+1}^{2k} \frac{1}{i} + \frac{1}{2k}.$$

Subtracting $\dfrac{1}{2k}$ from this equality, we obtain

$$\sum_{i=k}^{2k-1} \frac{1}{i} - \frac{1}{2k} = \sum_{i=k+1}^{2k} \frac{1}{i}.$$

Hence,

$$\sum_{i=k+1}^{2k} \frac{1}{i} = \sum_{i=k}^{2k-1} \frac{1}{i} - \frac{1}{2k} = \sum_{i=k}^{2k-1} \frac{1}{i} + \frac{-1}{2k}.$$

Comparing this with (221), we obtain

$$\sum_{i=1}^{2k} \frac{(-1)^{i-1}}{i} = \sum_{i=k+1}^{2k} \frac{1}{i}. \tag{223}$$

But this is precisely the statement $\mathcal{A}(k)$ [98]. Thus, the statement $\mathcal{A}(k)$ holds.

Now, forget that we fixed k. We thus have shown that if $k \in \mathbb{Z}_{\geq 0+1}$ is such that $\mathcal{A}(k-1)$ holds, then $\mathcal{A}(k)$ also holds. This proves (216).]

Now, both assumptions of Theorem 2.163 (applied to $g = 0$) are satisfied (indeed, Assumption 1 holds because the statement $\mathcal{A}(0)$ holds, whereas Assumption 2 holds because of (216)). Thus, Theorem 2.163 (applied to $g = 0$) shows that $\mathcal{A}(n)$ holds for each $n \in \mathbb{Z}_{\geq 0}$. In other words, $\sum\limits_{i=1}^{2n} \dfrac{(-1)^{i-1}}{i} = \sum\limits_{i=n+1}^{2n} \dfrac{1}{i}$ holds for each $n \in \mathbb{Z}_{\geq 0}$ (since $\mathcal{A}(n)$ is the statement $\left(\sum\limits_{i=1}^{2n} \dfrac{(-1)^{i-1}}{i} = \sum\limits_{i=n+1}^{2n} \dfrac{1}{i}\right)$). In other words, $\sum\limits_{i=1}^{2n} \dfrac{(-1)^{i-1}}{i} = \sum\limits_{i=n+1}^{2n} \dfrac{1}{i}$ holds for each $n \in \mathbb{N}$ (because $\mathbb{Z}_{\geq 0} = \mathbb{N}$). This proves Proposition 2.164. □

2.16.2. Conventions for writing proofs using "$k-1$ to k" induction

Just like most of the other induction principles that we have so far introduced, Theorem 2.163 is not usually invoked explicitly when it is used; instead, its use is signalled by certain words:

Convention 2.165. Let $g \in \mathbb{Z}$. For each $n \in \mathbb{Z}_{\geq g}$, let $\mathcal{A}(n)$ be a logical statement. Assume that you want to prove that $\mathcal{A}(n)$ holds for each $n \in \mathbb{Z}_{\geq g}$.

Theorem 2.163 offers the following strategy for proving this: First show that Assumption 1 of Theorem 2.163 is satisfied; then, show that Assumption 2 of Theorem 2.163 is satisfied; then, Theorem 2.163 automatically completes your proof.

[98] because $\mathcal{A}(k)$ is defined to be the statement $\left(\sum\limits_{i=1}^{2k} \dfrac{(-1)^{i-1}}{i} = \sum\limits_{i=k+1}^{2k} \dfrac{1}{i}\right)$

A proof that follows this strategy is called a *proof by induction on n* (or *proof by induction over n*) *starting at g* or (less precisely) an *inductive proof*. Most of the time, the words "starting at g" are omitted, since the value of g is usually clear from the statement that is being proven. Usually, the statements $\mathcal{A}(n)$ are not explicitly stated in the proof either, since they can also be inferred from the context.

The proof that Assumption 1 is satisfied is called the *induction base* (or *base case*) of the proof. The proof that Assumption 2 is satisfied is called the *induction step* of the proof.

In order to prove that Assumption 2 is satisfied, you will usually want to fix a $k \in \mathbb{Z}_{\geq g+1}$ such that $\mathcal{A}(k-1)$ holds, and then prove that $\mathcal{A}(k)$ holds. In other words, you will usually want to fix $k \in \mathbb{Z}_{\geq g+1}$, assume that $\mathcal{A}(k-1)$ holds, and then prove that $\mathcal{A}(k)$ holds. When doing so, it is common to refer to the assumption that $\mathcal{A}(k-1)$ holds as the *induction hypothesis* (or *induction assumption*).

This language is exactly the same that was introduced in Convention 2.56 for proofs by "standard" induction starting at g. The only difference between proofs that use Theorem 2.53 and proofs that use Theorem 2.163 is that the induction step in the former proofs assumes $\mathcal{A}(m)$ and proves $\mathcal{A}(m+1)$, whereas the induction step in the latter proofs assumes $\mathcal{A}(k-1)$ and proves $\mathcal{A}(k)$. (Of course, the letters "m" and "k" are not set in stone; any otherwise unused letters can be used in their stead. Thus, what distinguishes proofs that use Theorem 2.53 from proofs that use Theorem 2.163 is not the letter they use, but the "$+1$" versus the "-1".)

Let us repeat the above proof of Proposition 2.164 (or, more precisely, its non-computational part) using this language:

Proof of Proposition 2.164 (second version). We must prove (215) for every $n \in \mathbb{N}$. In other words, we must prove (215) for every $n \in \mathbb{Z}_{\geq 0}$ (since $\mathbb{N} = \mathbb{Z}_{\geq 0}$). We shall prove this by induction on n starting at 0:

Induction base: We have $\sum\limits_{i=1}^{2\cdot 0} \dfrac{(-1)^{i-1}}{i} = (\text{empty sum}) = 0$. Comparing this with $\sum\limits_{i=0+1}^{2\cdot 0} \dfrac{1}{i} = (\text{empty sum}) = 0$, we obtain $\sum\limits_{i=1}^{2\cdot 0} \dfrac{(-1)^{i-1}}{i} = \sum\limits_{i=0+1}^{2\cdot 0} \dfrac{1}{i}$. In other words, (215) holds for $n = 0$. This completes the induction base.

Induction step: Let $k \in \mathbb{Z}_{\geq 1}$. Assume that (215) holds for $n = k-1$. We must show that (215) holds for $n = k$.

We have $k \in \mathbb{Z}_{\geq 1}$. In other words, k is an integer and satisfies $k \geq 1$.

We have assumed that (215) holds for $n = k-1$. In other words,

$$\sum_{i=1}^{2(k-1)} \frac{(-1)^{i-1}}{i} = \sum_{i=(k-1)+1}^{2(k-1)} \frac{1}{i}. \tag{224}$$

From here, we can obtain

$$\sum_{i=1}^{2k} \frac{(-1)^{i-1}}{i} = \sum_{i=k+1}^{2k} \frac{1}{i}. \tag{225}$$

(Indeed, we can derive (225) from (224) in exactly the same way as we derived (223) from (217) in the above first version of the proof of Proposition 2.164; nothing about this argument needs to be changed, so we have no reason to repeat it.)

But the equality (225) shows that (215) holds for $n = k$. This completes the induction step. Hence, (215) is proven by induction. This proves Proposition 2.164. □

Exercise 2.9. Let $n \in \mathbb{N}$. Prove that

$$\sum_{k=0}^{n} (-1)^k (k+1) = \begin{cases} n/2+1, & \text{if } n \text{ is even;} \\ -(n+1)/2, & \text{if } n \text{ is odd} \end{cases}.$$

The claim of Exercise 2.9 can be rewritten as

$$1 - 2 + 3 - 4 \pm \cdots + (-1)^n (n+1) = \begin{cases} n/2+1, & \text{if } n \text{ is even;} \\ -(n+1)/2, & \text{if } n \text{ is odd} \end{cases}.$$

3. On binomial coefficients

The present chapter is about *binomial coefficients*. They are used in almost every part of mathematics, and studying them provides good opportunities to practice the arts of mathematical induction and of finding combinatorial bijections.

Identities involving binomial coefficients are legion, and books have been written about them (let me mention [GrKnPa94, Chapter 5] as a highly readable introduction; but, e.g., Henry W. Gould's website goes far deeper down the rabbit hole). We shall only study a few of these identities.

3.1. Definitions and basic properties

3.1.1. The definition

Let us first define binomial coefficients:

Definition 3.1. Let $n \in \mathbb{N}$ and $m \in \mathbb{Q}$. Recall that $n!$ is a positive integer; thus, $n! \neq 0$. (Keep in mind that $0! = 1$.)

We define a rational number $\binom{m}{n}$ by

$$\binom{m}{n} = \frac{m(m-1)\cdots(m-n+1)}{n!}. \tag{226}$$

(This fraction is well-defined, since $n! \neq 0$. When $n = 0$, the numerator of this fraction (i.e., the product $m(m-1)\cdots(m-n+1)$) is an empty product. Recall that, by convention, an empty product is always defined to be 1.)

This number $\dbinom{m}{n}$ is called a *binomial coefficient*, and is often pronounced "m choose n".

We can extend this definition to the case when $m \in \mathbb{R}$ or $m \in \mathbb{C}$ (rather than $m \in \mathbb{Q}$) by using the same equality (226). Of course, in that case, $\dbinom{m}{n}$ will not be a rational number anymore.

Example 3.2. The formula (226) yields

$$\binom{4}{2} = \frac{4(4-1)}{2!} = \frac{4(4-1)}{2} = 6;$$
$$\binom{5}{1} = \frac{5}{1!} = \frac{5}{1} = 5;$$
$$\binom{8}{3} = \frac{8(8-1)(8-2)}{3!} = \frac{8(8-1)(8-2)}{6} = 56.$$

Here is a table of the binomial coefficients $\dbinom{n}{k}$ for all values $n \in \{-3,-2,-1,\ldots,6\}$ and some of the values $k \in \{0,1,2,3,4,5\}$. In the following table, each row corresponds to a value of n, while each southwest-northeast diagonal corresponds to a value of k:

											k=0 ↙		k=1 ↙		k=2 ↙		k=3 ↙
$n = -3$ →										1		−3		6		−10	
$n = -2$ →									1		−2		3		−4		
$n = -1$ →								1		−1		1		−1		1	
$n = 0$ →							1		0		0		0		0		
$n = 1$ →						1		1		0		0		0		0	
$n = 2$ →					1		2		1		0		0		0		
$n = 3$ →				1		3		3		1		0		0		0	
$n = 4$ →			1		4		6		4		1		0		0		
$n = 5$ →		1		5		10		10		5		1		0		0	
$n = 6$ →	1		6		15		20		15		6		1		0		

The binomial coefficients $\dbinom{m}{n}$ form the so-called *Pascal's triangle*[99]. Let us state

[99] More precisely, the numbers $\dbinom{m}{n}$ for $m \in \mathbb{N}$ and $n \in \{0,1,\ldots,m\}$ form Pascal's triangle. Nev-

a few basic properties of these numbers:

3.1.2. Simple formulas

Proposition 3.3. Let $m \in \mathbb{Q}$.

(a) We have

$$\binom{m}{0} = 1. \tag{227}$$

(b) We have

$$\binom{m}{1} = m. \tag{228}$$

Proof of Proposition 3.3. **(a)** The definition of $\binom{m}{0}$ yields

$$\binom{m}{0} = \frac{m(m-1)\cdots(m-0+1)}{0!}.$$

Since $m(m-1)\cdots(m-0+1) = (\text{a product of 0 integers}) = 1$, this rewrites as $\binom{m}{0} = \frac{1}{0!} = 1$ (since $0! = 1$). This proves Proposition 3.3 **(a)**.

(b) The definition of $\binom{m}{1}$ yields

$$\binom{m}{1} = \frac{m(m-1)\cdots(m-1+1)}{1!}.$$

Since $m(m-1)\cdots(m-1+1) = m$ (because the product $m(m-1)\cdots(m-1+1)$ consists of 1 factor only, and this factor is m), this rewrites as $\binom{m}{1} = \frac{m}{1!} = m$ (since $1! = 1$). This proves Proposition 3.3 **(b)**. □

Proposition 3.4. Let $m \in \mathbb{N}$ and $n \in \mathbb{N}$ be such that $m \geq n$. Then,

$$\binom{m}{n} = \frac{m!}{n!(m-n)!}. \tag{229}$$

Remark 3.5. Caution: The formula (229) holds only for $m \in \mathbb{N}$ and $n \in \mathbb{N}$ satisfying $m \geq n$. Thus, neither $\binom{-3}{2}$ nor $\binom{1/3}{3}$ nor $\binom{2}{5}$ can be computed using this formula! Definition 3.1 thus can be used to compute $\binom{m}{n}$ in many more cases than (229) does.

ertheless, the "other" binomial coefficients (particularly the ones where m is a negative integer) are highly useful, too.

Proof of Proposition 3.4. Multiplying both sides of the equality (226) with $n!$, we obtain

$$n! \cdot \binom{m}{n} = m(m-1)\cdots(m-n+1). \tag{230}$$

But

$$m! = m(m-1)\cdots 1 = (m(m-1)\cdots(m-n+1)) \cdot \underbrace{((m-n)(m-n-1)\cdots 1)}_{=(m-n)!}$$
$$= (m(m-1)\cdots(m-n+1)) \cdot (m-n)!,$$

so that $\dfrac{m!}{(m-n)!} = m(m-1)\cdots(m-n+1)$. Comparing this with (230), we obtain $n! \cdot \dbinom{m}{n} = \dfrac{m!}{(m-n)!}$. Dividing this equality by $n!$, we obtain $\dbinom{m}{n} = \dfrac{m!}{n!(m-n)!}$. Thus, Proposition 3.4 is proven. □

Proposition 3.6. Let $m \in \mathbb{N}$ and $n \in \mathbb{N}$ be such that $m < n$. Then,

$$\binom{m}{n} = 0. \tag{231}$$

Remark 3.7. Caution: The formula (231) is not true if we drop the condition $m \in \mathbb{N}$. For example, $\dbinom{-3}{2} = 6 \neq 0$ despite $-3 < 2$.

Proof of Proposition 3.6. We have $m \geq 0$ (since $m \in \mathbb{N}$). Also, $m < n$, so that $m \leq n-1$ (since m and n are integers). Thus, $m \in \{0, 1, \ldots, n-1\}$. Hence, $m - m$ is one of the n integers $m, m-1, \ldots, m-n+1$. Thus, one of the n factors of the product $m(m-1)\cdots(m-n+1)$ is $m - m = 0$. Therefore, the whole product $m(m-1)\cdots(m-n+1)$ is 0 (because if one of the factors of a product is 0, then the whole product must be 0). Thus, $m(m-1)\cdots(m-n+1) = 0$. Hence, (226) becomes

$$\binom{m}{n} = \frac{m(m-1)\cdots(m-n+1)}{n!} = \frac{0}{n!} \qquad (\text{since } m(m-1)\cdots(m-n+1) = 0)$$
$$= 0.$$

This proves Proposition 3.6. □

Proposition 3.8. Let $m \in \mathbb{N}$ and $n \in \mathbb{N}$ be such that $m \geq n$. Then,

$$\binom{m}{n} = \binom{m}{m-n}. \tag{232}$$

Proposition 3.8 is commonly known as the *symmetry identity for the binomial coefficients*. Notice that Proposition 3.8 becomes false (and, with our definitions, actually meaningless) if the requirement that $m \in \mathbb{N}$ is dropped.

Proof of Proposition 3.8. We have $m - n \in \mathbb{N}$ (since $m \geq n$) and $m \geq m - n$ (since $n \geq 0$ (since $n \in \mathbb{N}$)). Hence, (229) (applied to $m - n$ instead of n) yields

$$\binom{m}{m-n} = \frac{m!}{(m-n)!\,(m-(m-n))!} = \frac{m!}{(m-(m-n))!\,(m-n)!} = \frac{m!}{n!\,(m-n)!}$$

(since $m - (m - n) = n$). Compared with (229), this yields $\binom{m}{n} = \binom{m}{m-n}$. Proposition 3.8 is thus proven. □

Proposition 3.9. Let $m \in \mathbb{N}$. Then,

$$\binom{m}{m} = 1. \tag{233}$$

Proof of Proposition 3.9. The equality (232) (applied to $n = m$) yields $\binom{m}{m} = \binom{m}{m-m} = \binom{m}{0} = 1$ (according to (227)). This proves Proposition 3.9. □

Exercise 3.1. Let $m \in \mathbb{N}$ and $(k_1, k_2, \ldots, k_m) \in \mathbb{N}^m$. Prove that $\dfrac{(k_1 + k_2 + \cdots + k_m)!}{k_1! k_2! \cdots k_m!}$ is a positive integer.

Remark 3.10. Let $m \in \mathbb{N}$ and $(k_1, k_2, \ldots, k_m) \in \mathbb{N}^m$. Exercise 3.1 shows that $\dfrac{(k_1 + k_2 + \cdots + k_m)!}{k_1! k_2! \cdots k_m!}$ is a positive integer. This positive integer is called a *multinomial coefficient*, and is often denoted by $\binom{n}{k_1, k_2, \ldots, k_m}$, where $n = k_1 + k_2 + \cdots + k_m$. (We shall avoid this particular notation, since it makes the meaning of $\binom{n}{k}$ slightly ambiguous: It could mean both the binomial coefficient $\binom{n}{k}$ and the multinomial coefficient $\binom{n}{k_1, k_2, \ldots, k_m}$ for $(k_1, k_2, \ldots, k_m) = (k)$. Fortunately, the ambiguity is not really an issue, because the only situation in which both meanings make sense is when $k = n \in \mathbb{N}$, but in this case both interpretations give the same value 1.)

Exercise 3.2. Let $n \in \mathbb{N}$.

(a) Prove that

$$(2n-1) \cdot (2n-3) \cdot \cdots \cdot 1 = \frac{(2n)!}{2^n n!}.$$

(The left hand side is understood to be the product of all odd integers from 1 to $2n-1$.)

(b) Prove that

$$\binom{-1/2}{n} = \left(\frac{-1}{4}\right)^n \binom{2n}{n}.$$

(c) Prove that

$$\binom{-1/3}{n}\binom{-2/3}{n} = \frac{(3n)!}{(3^n n!)^3}.$$

3.1.3. The recurrence relation of the binomial coefficients

Proposition 3.11. Let $m \in \mathbb{Q}$ and $n \in \{1,2,3,\ldots\}$. Then,

$$\binom{m}{n} = \binom{m-1}{n-1} + \binom{m-1}{n}. \tag{234}$$

Proof of Proposition 3.11. From $n \in \{1,2,3,\ldots\}$, we obtain $n! = n \cdot (n-1)!$, so that $(n-1)! = n!/n$ and thus $\dfrac{1}{(n-1)!} = \dfrac{1}{n!/n} = \dfrac{1}{n!} \cdot n$.

The definition of $\dbinom{m}{n-1}$ yields

$$\begin{aligned}\binom{m}{n-1} &= \frac{m(m-1)\cdots(m-(n-1)+1)}{(n-1)!} \\ &= \frac{1}{(n-1)!} \cdot (m(m-1)\cdots(m-(n-1)+1)).\end{aligned}$$

The same argument (applied to $m-1$ instead of m) yields

$$\begin{aligned}\binom{m-1}{n-1} &= \underbrace{\frac{1}{(n-1)!}}_{=\frac{1}{n!}\cdot n} \cdot \left((m-1)\underbrace{((m-1)-1)}_{=m-2}\cdots\underbrace{((m-1)-(n-1)+1)}_{=m-n+1}\right) \\ &= \frac{1}{n!} \cdot n \cdot ((m-1)(m-2)\cdots(m-n+1)). \end{aligned} \tag{235}$$

On the other hand,

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

The same argument (applied to $m-1$ instead of m) yields

$$\begin{aligned}\dbinom{m-1}{n} &= \dfrac{1}{n!}\left((m-1)\underbrace{((m-1)-1)}_{=m-2}\cdots\underbrace{((m-1)-n+1)}_{=m-n}\right)\\ &= \dfrac{1}{n!}\underbrace{((m-1)(m-2)\cdots(m-n))}_{=((m-1)(m-2)\cdots(m-n+1))\cdot(m-n)}\\ &= \dfrac{1}{n!}((m-1)(m-2)\cdots(m-n+1))\cdot(m-n)\\ &= \dfrac{1}{n!}(m-n)\cdot((m-1)(m-2)\cdots(m-n+1)).\end{aligned}$$

Adding (235) to this equality, we obtain

$$\begin{aligned}&\dbinom{m-1}{n}+\dbinom{m-1}{n-1}\\ &= \dfrac{1}{n!}(m-n)\cdot((m-1)(m-2)\cdots(m-n+1))\\ &\qquad + \dfrac{1}{n!}\cdot n\cdot((m-1)(m-2)\cdots(m-n+1))\\ &= \dfrac{1}{n!}\underbrace{((m-n)+n)}_{=m}\cdot((m-1)(m-2)\cdots(m-n+1))\\ &= \dfrac{1}{n!}\underbrace{m\cdot((m-1)(m-2)\cdots(m-n+1))}_{=m(m-1)\cdots(m-n+1)} = \dfrac{1}{n!}(m(m-1)\cdots(m-n+1))\\ &= \dbinom{m}{n} \qquad \left(\text{since } \dbinom{m}{n} = \dfrac{1}{n!}(m(m-1)\cdots(m-n+1))\right).\end{aligned}$$

This proves Proposition 3.11. □

The formula (234) is known as the *recurrence relation of the binomial coefficients*[100].

Exercise 3.3. (a) Prove that every $n \in \mathbb{N}$ and $q \in \mathbb{Q}$ satisfy

$$\sum_{r=0}^{n}\dbinom{r+q}{r} = \dbinom{n+q+1}{n}.$$

(b) Prove that every $n \in \{-1, 0, 1, \ldots\}$ and $k \in \mathbb{N}$ satisfy

$$\sum_{i=0}^{n}\dbinom{i}{k} = \sum_{i=k}^{n}\dbinom{i}{k} = \dbinom{n+1}{k+1}.$$

[100] Often it is extended to the case $n = 0$ by setting $\dbinom{m}{-1} = 0$. It then follows from (227) in this case.

The formula (234) is responsible for the fact that "every number in Pascal's triangle is the sum of the two numbers above it". (Of course, if you use this fact as a *definition* of Pascal's triangle, then (234) is conversely responsible for the fact that the numbers in this triangle are the binomial coefficients.)

(Keep in mind that $\sum_{i=k}^{n} \dbinom{i}{k}$ is an empty sum whenever $n < k$.)

The claim of Exercise 3.3 **(b)** is one of several formulas known as the *hockey-stick identity* (due to the fact that marking the binomial coefficients appearing in it in Pascal's triangle results in a shape resembling a hockey stick[101]); it appears, e.g., in [Galvin17, Identity 11.10] (or, rather, the second equality sign of Exercise 3.3 **(b)** appears there, but the rest is easy).

3.1.4. The combinatorial interpretation of binomial coefficients

Proposition 3.12. If $m \in \mathbb{N}$ and $n \in \mathbb{N}$, and if S is an m-element set, then

$$\dbinom{m}{n} \text{ is the number of all } n\text{-element subsets of } S. \tag{236}$$

In less formal terms, Proposition 3.12 says the following: If $m \in \mathbb{N}$ and $n \in \mathbb{N}$, then $\dbinom{m}{n}$ is the number of ways to pick out n among m given objects, without replacement[102] and without regard for the order in which they are picked out. (Probabilists call this "unordered samples without replacement".)

Example 3.13. Proposition 3.12 (applied to $m = 4$, $n = 2$ and $S = \{0, 1, 2, 3\}$) shows that $\dbinom{4}{2}$ is the number of all 2-element subsets of $\{0, 1, 2, 3\}$ (since $\{0, 1, 2, 3\}$ is a 4-element set). And indeed, this is easy to verify by brute force: The 2-element subsets of $\{0, 1, 2, 3\}$ are

$$\{0, 1\}, \{0, 2\}, \{0, 3\}, \{1, 2\}, \{1, 3\} \text{ and } \{2, 3\},$$

so there are $6 = \dbinom{4}{2}$ of them.

Remark 3.14. Caution: Proposition 3.12 says nothing about binomial coefficients $\dbinom{m}{n}$ with negative m. Indeed, there are no m-element sets S when m is negative; thus, Proposition 3.12 would be vacuously true[103] when m is negative, but this would not help us computing binomial coefficients $\dbinom{m}{n}$ with negative m.

Actually, when $m \in \mathbb{Z}$ is negative, the number $\dbinom{m}{n}$ is positive for n even and negative for n odd (easy exercise), and so an interpretation of $\dbinom{m}{n}$ as a number

[101] See `https://math.stackexchange.com/q/1490794` for an illustration.

[102] That is, one must not pick out the same object twice.

of ways to do something is rather unlikely. (On the other hand, $(-1)^n \binom{m}{n}$ does have such an interpretation.)

Remark 3.15. Some authors (for example, those of [LeLeMe16] and of [Galvin17]) use (236) as the *definition* of $\binom{m}{n}$. This is a legitimate definition of $\binom{m}{n}$ in the case when m and n are nonnegative integers (and, of course, equivalent to our definition); but it is not as general as ours, since it does not extend to negative (or non-integer) m.

Exercise 3.4. Prove Proposition 3.12.

Proposition 3.12 is one of the most basic facts of *enumerative combinatorics* – the part of mathematics that is mostly concerned with counting problems (i.e., the study of the sizes of finite sets). We will encounter some further results from enumerative combinatorics below (e.g., Exercise 3.15 and Exercise 4.3); but we shall not go deep into this subject. More serious expositions of enumerative combinatorics include Loehr's textbook [Loehr11], Galvin's lecture notes [Galvin17], Aigner's book [Aigner07], and Stanley's two-volume treatise ([Stanle11] and [Stanle01]).

3.1.5. Upper negation

Proposition 3.16. Let $m \in \mathbb{Q}$ and $n \in \mathbb{N}$. Then,

$$\binom{m}{n} = (-1)^n \binom{n-m-1}{n}. \tag{237}$$

[103] Recall that a mathematical statement of the form "if $\mathcal{A}$, then $\mathcal{B}$" is said to be *vacuously true* if $\mathcal{A}$ never holds. For example, the statement "if $0 = 1$, then every integer is odd" is vacuously true, because $0 = 1$ is false. Proposition 3.12 is vacuously true when m is negative, because the condition "S is an m-element set" never holds when m is negative.

By the laws of logic, a vacuously true statement is always true! See Convention 2.37 for a discussion of this principle.

Proof of Proposition 3.16. The equality (226) (applied to $n-m-1$ instead of m) yields

$$\begin{aligned}
\binom{n-m-1}{n} &= \frac{(n-m-1)\,((n-m-1)-1)\cdots((n-m-1)-n+1)}{n!} \\
&= \frac{1}{n!}(n-m-1)\,((n-m-1)-1)\cdots\underbrace{((n-m-1)-n+1)}_{=-m} \\
&= \frac{1}{n!}\underbrace{(n-m-1)\,((n-m-1)-1)\cdots(-m)}_{\substack{=(-m)(-m+1)\cdots(n-m-1)\\ \text{(here, we have just reversed the order of the factors in the product)}}} \\
&= \frac{1}{n!}\underbrace{(-m)}_{=(-1)m}\underbrace{(-m+1)}_{=(-1)(m-1)}\cdots\underbrace{(n-m-1)}_{=(-1)(m-n+1)} \\
&= \frac{1}{n!}((-1)\,m)\,((-1)\,(m-1))\cdots((-1)\,(m-n+1)) \\
&= \frac{1}{n!}(-1)^n\,(m\,(m-1)\cdots(m-n+1)),
\end{aligned}$$

so that

$$\begin{aligned}
(-1)^n\binom{n-m-1}{n} &= (-1)^n\cdot\frac{1}{n!}(-1)^n\,(m\,(m-1)\cdots(m-n+1)) \\
&= \underbrace{(-1)^n\,(-1)^n}_{\substack{=(-1)^{n+n}=(-1)^{2n}=1\\ \text{(since } 2n \text{ is even)}}}\cdot\frac{1}{n!}(m\,(m-1)\cdots(m-n+1)) \\
&= \frac{1}{n!}(m\,(m-1)\cdots(m-n+1)) = \frac{m\,(m-1)\cdots(m-n+1)}{n!}.
\end{aligned}$$

Compared with (226), this yields $\binom{m}{n} = (-1)^n\binom{n-m-1}{n}$. Proposition 3.16 is therefore proven. □

The formula (237) is known as the *upper negation formula*.

Corollary 3.17. Let $n \in \mathbb{N}$. Then,

$$\binom{-1}{n} = (-1)^n.$$

Proof of Corollary 3.17. Proposition 3.16 (applied to $m = -1$) yields

$$\binom{-1}{n} = (-1)^n \underbrace{\binom{n-(-1)-1}{n}}_{\substack{=\binom{n}{n} \\ \text{(since } n-(-1)-1=n)}} = (-1)^n \underbrace{\binom{n}{n}}_{\substack{=1 \\ \text{(by Proposition 3.9} \\ \text{(applied to } m=n))}} = (-1)^n .$$

This proves Corollary 3.17. □

Exercise 3.5. **(a)** Show that $\binom{-1}{k} = (-1)^k$ for each $k \in \mathbb{N}$.

(b) Show that $\binom{-2}{k} = (-1)^k (k+1)$ for each $k \in \mathbb{N}$.

(c) Show that $\dfrac{1! \cdot 2! \cdot \cdots \cdot (2n)!}{n!} = 2^n \cdot \left(\prod_{i=1}^{n} ((2i-1)!) \right)^2$ for each $n \in \mathbb{N}$.

Remark 3.18. Parts **(a)** and **(b)** of Exercise 3.5 are known facts (actually, part **(a)** is just a repetition of Corollary 3.17, for the purpose of making the analogy to part **(b)** more visible). Part **(c)** is a generalization of a puzzle posted on `https://www.reddit.com/r/math/comments/7rybhp/factorial_problem/` . (The puzzle boils down to the fact that $\dfrac{1! \cdot 2! \cdot \cdots \cdot (2n)!}{n!}$ is a perfect square when $n \in \mathbb{N}$ is even. But this follows from Exercise 3.5 **(c)**, because when $n \in \mathbb{N}$ is even, both factors 2^n and $\left(\prod_{i=1}^{n} ((2i-1)!) \right)^2$ on the right hand side of Exercise 3.5 **(c)** are perfect squares.)

3.1.6. Binomial coefficients of integers are integers

Lemma 3.19. Let $m \in \mathbb{N}$ and $n \in \mathbb{N}$. Then,

$$\binom{m}{n} \in \mathbb{N}.$$

Proof of Lemma 3.19. We have $m \in \mathbb{N}$. Thus, there exists an m-element set S (for example, $S = \{1, 2, \ldots, m\}$). Consider such an S. Then, $\binom{m}{n}$ is the number of all n-element subsets of S (because of (236)). Hence, $\binom{m}{n}$ is a nonnegative integer, so that $\binom{m}{n} \in \mathbb{N}$. This proves Lemma 3.19. □

It is also easy to prove Lemma 3.19 by induction on m, using (227) and (231) in the induction base and using (234) in the induction step.

Proposition 3.20. Let $m \in \mathbb{Z}$ and $n \in \mathbb{N}$. Then,

$$\binom{m}{n} \in \mathbb{Z}. \tag{238}$$

Proof of Proposition 3.20. We need to show (238). We are in one of the following two cases:

Case 1: We have $m \geq 0$.

Case 2: We have $m < 0$.

Let us first consider Case 1. In this case, we have $m \geq 0$. Hence, $m \in \mathbb{N}$. Thus, Lemma 3.19 yields $\binom{m}{n} \in \mathbb{N} \subseteq \mathbb{Z}$. This proves (238) in Case 1.

Let us now consider Case 2. In this case, we have $m < 0$. Thus, $m \leq -1$ (since m is an integer), so that $m+1 \leq 0$, so that $n-m-1 = n - \underbrace{(m+1)}_{\leq 0} \geq n \geq 0$. Hence, $n-m-1 \in \mathbb{N}$. Therefore, Lemma 3.19 (applied to $n-m-1$ instead of m) yields $\binom{n-m-1}{n} \in \mathbb{N} \subseteq \mathbb{Z}$. Now, (237) shows that $\binom{m}{n} = \underbrace{(-1)^n}_{\in\mathbb{Z}} \underbrace{\binom{n-m-1}{n}}_{\in\mathbb{Z}} \in \mathbb{Z}$ (here, we have used the fact that the product of two integers is an integer). This proves (238) in Case 2.

We thus have proven (238) in each of the two Cases 1 and 2. We can therefore conclude that (238) always holds. Thus, Proposition 3.20 is proven. □

The above proof of Proposition 3.20 may well be the simplest one. There is another proof, which uses Theorem 2.149, but it is more complicated[104]. There is yet another proof using basic number theory (specifically, checking how often a prime p appears in the numerator and the denominator of $\binom{m}{n} = \frac{m(m-1)\cdots(m-n+1)}{n!}$), but this is not quite easy.

3.1.7. The binomial formula

Proposition 3.21. Let x and y be two rational numbers (or real numbers, or complex numbers). Then,

$$(x+y)^n = \sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k} \tag{239}$$

for every $n \in \mathbb{N}$.

[104] It requires an induction on n nested inside the induction step of the induction on m.

Proposition 3.21 is the famous *binomial formula* (also known as the *binomial theorem*) and has a well-known standard proof by induction over n (using (234) and (227))[105]. Some versions of it hold for negative n as well (but not in the exact form (239), and not without restrictions).

Exercise 3.6. Prove Proposition 3.21.

There is an analogue of Proposition 3.21 for a sum of m rational numbers (rather than 2 rational numbers); it is called the "multinomial formula" (and involves the multinomial coefficients from Remark 3.10). We shall state it in a more general setting in Exercise 6.2.

3.1.8. The absorption identity

Proposition 3.22. Let $n \in \{1, 2, 3, \ldots\}$ and $m \in \mathbb{Q}$. Then,

$$\binom{m}{n} = \frac{m}{n}\binom{m-1}{n-1}. \tag{240}$$

Proof of Proposition 3.22. The definition of $\binom{m-1}{n-1}$ yields

$$\begin{aligned}\binom{m-1}{n-1} &= \frac{(m-1)\left((m-1)-1\right)\cdots\left((m-1)-(n-1)+1\right)}{(n-1)!} \\ &= \frac{(m-1)(m-2)\cdots(m-n+1)}{(n-1)!}\end{aligned}$$

(since $(m-1)-1 = m-2$ and $(m-1)-(n-1)+1 = m-n+1$). Multiplying both sides of this equality by $\frac{m}{n}$, we obtain

$$\begin{aligned}\frac{m}{n}\binom{m-1}{n-1} &= \frac{m}{n} \cdot \frac{(m-1)(m-2)\cdots(m-n+1)}{(n-1)!} \\ &= \frac{m(m-1)(m-2)\cdots(m-n+1)}{n(n-1)!} = \frac{m(m-1)\cdots(m-n+1)}{n!}\end{aligned}$$

(since $m(m-1)(m-2)\cdots(m-n+1) = m(m-1)\cdots(m-n+1)$ and $n(n-1)! = n!$). Compared with (226), this yields $\binom{m}{n} = \frac{m}{n}\binom{m-1}{n-1}$. This proves Proposition 3.22. □

The relation (240) is called the *absorption identity* in [GrKnPa94, §5.1].

[105] See Exercise 3.6 for this proof.

Exercise 3.7. Let k, a and b be three positive integers such that $k \leq a \leq b$. Prove that

$$\frac{k-1}{k} \sum_{n=a}^{b} \frac{1}{\dbinom{n}{k}} = \frac{1}{\dbinom{a-1}{k-1}} - \frac{1}{\dbinom{b}{k-1}}.$$

(In particular, all fractions appearing in this equality are well-defined.)

3.1.9. Trinomial revision

Proposition 3.23. Let $m \in \mathbb{Q}$, $a \in \mathbb{N}$ and $i \in \mathbb{N}$ be such that $i \geq a$. Then,

$$\binom{m}{i}\binom{i}{a} = \binom{m}{a}\binom{m-a}{i-a}. \qquad (241)$$

Proof of Proposition 3.23. Let $g = m - a$. Then, $m - a = g$. Therefore,

$$\begin{aligned}
&\binom{m}{a}\binom{m-a}{i-a} \\
&= \underbrace{\binom{m}{a}}_{\substack{=\dfrac{m(m-1)\cdots(m-a+1)}{a!} \\ \left(\text{by the definition of } \binom{m}{a}\right)}} \underbrace{\binom{g}{i-a}}_{\substack{=\dfrac{g(g-1)\cdots(g-(i-a)+1)}{(i-a)!} \\ \left(\text{by the definition of } \binom{g}{i-a}\right)}} \\
&= \frac{m(m-1)\cdots(m-a+1)}{a!} \cdot \frac{g(g-1)\cdots(g-(i-a)+1)}{(i-a)!} \\
&= \frac{m(m-1)\cdots(g+1)}{a!} \cdot \frac{g(g-1)\cdots(m-i+1)}{(i-a)!} \\
&\qquad \left(\text{since } m-a=g \text{ and } \underbrace{g}_{=m-a} - (i-a) = (m-a)-(i-a) = m-i\right) \\
&= \frac{1}{a! \cdot (i-a)!} \cdot \underbrace{(m(m-1)\cdots(g+1)) \cdot (g(g-1)\cdots(m-i+1))}_{=m(m-1)\cdots(m-i+1)} \\
&= \frac{1}{a! \cdot (i-a)!} \cdot m(m-1)\cdots(m-i+1).
\end{aligned}$$

Compared with

$$\underbrace{\dbinom{m}{i}}_{\substack{=\dfrac{m(m-1)\cdots(m-i+1)}{i!} \\ \text{(by the definition of } \binom{m}{i}\text{)}}} \underbrace{\dbinom{i}{a}}_{\substack{=\dfrac{i!}{a!(i-a)!} \\ \text{(by (229), applied to } i \text{ and } a \\ \text{instead of } m \text{ and } n\text{)}}}$$

$$= \frac{m(m-1)\cdots(m-i+1)}{i!} \cdot \frac{i!}{a!(i-a)!} = \frac{1}{a! \cdot (i-a)!} \cdot m(m-1)\cdots(m-i+1),$$

this yields $\dbinom{m}{i}\dbinom{i}{a} = \dbinom{m}{a}\dbinom{m-a}{i-a}$. This proves Proposition 3.23.

[Notice that we used (229) to simplify $\dbinom{i}{a}$ in this proof. Do not be tempted to use (229) to simplify $\dbinom{m}{i}$, $\dbinom{m}{a}$ and $\dbinom{m-a}{i-a}$: The m in these expressions may fail to be an integer!] □

Proposition 3.23 is a simple and yet highly useful formula, which Graham, Knuth and Patashnik call *trinomial revision* in [GrKnPa94, Table 174].

3.2. Binomial coefficients and polynomials

We have so far defined the binomial coefficient $\dbinom{m}{n}$ in the case when $n \in \mathbb{N}$ while m is some number (rational, real or complex). However, we can take this definition even further: For example, we can define $\dbinom{m}{n}$ when m is a polynomial with rational or real coefficients. Let us do this now:

Definition 3.24. Let $n \in \mathbb{N}$. Let m be a polynomial whose coefficients are rational numbers (or real numbers, or complex numbers).

We define a polynomial $\dbinom{m}{n}$ by the equality (226). This is a polynomial whose coefficients will be rational numbers or real numbers or complex numbers, depending on the coefficients of m.

Thus, in particular, for the polynomial $X \in \mathbb{Q}[X]$, we have

$$\binom{X}{n} = \frac{X(X-1)\cdots(X-n+1)}{n!} \qquad \text{for every } n \in \mathbb{N}.$$

In particular,

$$\binom{X}{0} = \frac{X(X-1)\cdots(X-0+1)}{0!} = \frac{\text{(empty product)}}{1} = 1;$$
$$\binom{X}{1} = \frac{X(X-1)\cdots(X-1+1)}{1!} = \frac{X}{1} = X;$$
$$\binom{X}{2} = \frac{X(X-1)}{2!} = \frac{X(X-1)}{2} = \frac{1}{2}X^2 - \frac{1}{2}X;$$
$$\binom{X}{3} = \frac{X(X-1)(X-2)}{3!} = \frac{X(X-1)(X-2)}{6} = \frac{1}{6}X^3 - \frac{1}{2}X^2 + \frac{1}{3}X.$$

The polynomial $\binom{X}{n}$ lets us compute the binomial coefficients $\binom{m}{n}$ for all $m \in \mathbb{N}$, because of the following:

Proposition 3.25. Let $m \in \mathbb{Q}$ and $n \in \mathbb{N}$. Then, the rational number $\binom{m}{n}$ is the result of evaluating the polynomial $\binom{X}{n}$ at $X = m$.

Proof of Proposition 3.25. We have $\binom{X}{n} = \frac{X(X-1)\cdots(X-n+1)}{n!}$. Hence, the result of evaluating the polynomial $\binom{X}{n}$ at $X = m$ is

$$\frac{m(m-1)\cdots(m-n+1)}{n!} = \binom{m}{n} \qquad \text{(by (226))}.$$

This proves Proposition 3.25. □

We note the following properties of the polynomials $\binom{X}{n}$:

Proposition 3.26. **(a)** We have

$$\binom{X}{0} = 1.$$

(b) We have

$$\binom{X}{1} = X.$$

(c) For every $n \in \{1, 2, 3, \ldots\}$, we have

$$\binom{X}{n} = \binom{X-1}{n} + \binom{X-1}{n-1}.$$

(d) For every $n \in \mathbb{N}$, we have

$$\binom{X}{n} = (-1)^n \binom{n-X-1}{n}.$$

(e) For every $n \in \{1, 2, 3, \ldots\}$, we have

$$\binom{X}{n} = \frac{X}{n} \binom{X-1}{n-1}.$$

(f) Let $a \in \mathbb{N}$ and $i \in \mathbb{N}$ be such that $i \geq a$. Then,

$$\binom{X}{i} \binom{i}{a} = \binom{X}{a} \binom{X-a}{i-a}.$$

Proof of Proposition 3.26. **(a)** To obtain a proof of Proposition 3.26 **(a)**, replace every appearance of "m" by "X" in the proof of Proposition 3.3 **(a)**.

(b) To obtain a proof of Proposition 3.26 **(b)**, replace every appearance of "m" by "X" in the proof of Proposition 3.3 **(b)**.

(c) To obtain a proof of Proposition 3.26 **(c)**, replace every appearance of "m" by "X" in the proof of Proposition 3.11.

(d) To obtain a proof of Proposition 3.26 **(d)**, replace every appearance of "m" by "X" in the proof of Proposition 3.16.

(e) To obtain a proof of Proposition 3.26 **(e)**, replace every appearance of "m" by "X" in the proof of Proposition 3.22.

(f) To obtain a proof of Proposition 3.26 **(f)**, replace every appearance of "m" by "X" in the proof of Proposition 3.23. □

Recall that any polynomial $P \in \mathbb{Q}[X]$ (that is, any polynomial in the indeterminate X with rational coefficients) can be quasi-uniquely written in the form $P(X) = \sum_{i=0}^{d} c_i X^i$ with rational $c_0, c_1, \ldots, c_d$. The word "quasi-uniquely" here means that the coefficients $c_0, c_1, \ldots, c_d$ are uniquely determined when $d \in \mathbb{N}$ is specified; they are not literally unique because we can always increase d by adding new 0 coefficients (for example, the polynomial $(1+X)^2$ can be written both as $1 + 2X + X^2$ and as $1 + 2X + X^2 + 0X^3 + 0X^4$).

It is not hard to check that an analogue of this statement holds with the X^i replaced by the $\binom{X}{i}$:

Proposition 3.27. **(a)** Any polynomial $P \in \mathbb{Q}[X]$ can be quasi-uniquely written in the form $P(X) = \sum_{i=0}^{d} c_i \binom{X}{i}$ with rational $c_0, c_1, \ldots, c_d$. (Again, "quasi-uniquely" means that we can always increase d by adding new 0 coefficients, but apart from this the $c_0, c_1, \ldots, c_d$ are uniquely determined.)

(b) The polynomial P is *integer-valued* (i.e., its values at integers are integers) if and only if these rationals $c_0, c_1, \ldots, c_d$ are integers.

We will not use this fact below, but it gives context to Theorem 3.30 and Exercise 3.8 further below. The "if" part of Proposition 3.27 **(b)** follows from (238). For a full proof of Proposition 3.27 **(b)**, see [AndDos12, Theorem 10.3]. See also [daSilv12] for a proof of the "only if" part.

We shall now prove some facts and give some exercises about binomial coefficients; but let us first prove a fundamental property of polynomials:

Lemma 3.28. **(a)** Let P be a polynomial in the indeterminate X with rational coefficients. Assume that $P(x) = 0$ for all $x \in \mathbb{N}$. Then, $P = 0$ as polynomials[106].

(b) Let P and Q be two polynomials in the indeterminate X with rational coefficients. Assume that $P(x) = Q(x)$ for all $x \in \mathbb{N}$. Then, $P = Q$ as polynomials.

(c) Let P be a polynomial in the indeterminates X and Y with rational coefficients. Assume that $P(x,y) = 0$ for all $x \in \mathbb{N}$ and $y \in \mathbb{N}$. Then, $P = 0$ as polynomials.

(d) Let P and Q be two polynomials in the indeterminates X and Y with rational coefficients. Assume that $P(x,y) = Q(x,y)$ for all $x \in \mathbb{N}$ and $y \in \mathbb{N}$. Then, $P = Q$ as polynomials.

Lemma 3.28 is a well-known property of polynomials with rational coefficients; let us still prove it for the sake of completeness.

Proof of Lemma 3.28. **(a)** The polynomial P satisfies $P(x) = 0$ for every $x \in \mathbb{N}$. Hence, every $x \in \mathbb{N}$ is a root of P. Thus, the polynomial P has infinitely many roots. But a nonzero polynomial in one variable (with rational coefficients) can only have finitely many roots[107]. If P was nonzero, this would force a contradiction with the sentence before. So P must be zero. In other words, $P = 0$. Lemma 3.28 **(a)** is proven.

(b) Every $x \in \mathbb{N}$ satisfies $(P - Q)(x) = P(x) - Q(x) = 0$ (since $P(x) = Q(x)$). Hence, Lemma 3.28 **(a)** (applied to $P - Q$ instead of P) yields $P - Q = 0$. Thus, $P = Q$. Lemma 3.28 **(b)** is thus proven.

(c) Every $x \in \mathbb{N}$ and $y \in \mathbb{N}$ satisfy

$$P(x,y) = 0. \tag{242}$$

We can write the polynomial P in the form $P = \sum_{k=0}^{d} P_k(X) Y^k$, where d is an integer and where each $P_k(X)$ (for $0 \leq k \leq d$) is a polynomial in the single variable X. Consider this d and these $P_k(X)$.

[106] Recall that two polynomials are said to be equal if and only if their respective coefficients are equal.

[107] In fact, a stronger statement holds: A nonzero polynomial in one variable (with rational coefficients) having degree $n \geq 0$ has at most n roots. See, for example, [Goodma15, Corollary 1.8.24] or [Joyce17, Theorem 1.58] or [Walker87, Corollary 4.5.10] or [Elman19, Corollary 33.7] or [Swanso20, Theorem 2.4.15] or [Knapp16, Corollary 1.14] for a proof. Note that Swanson, in [Swanso20], works with polynomial functions instead of polynomials; but as far as roots are concerned, the difference does not matter (since the roots of a polynomial are precisely the roots of the corresponding polynomial function).

Fix $\alpha \in \mathbb{N}$. Every $x \in \mathbb{N}$ satisfies

$$P(\alpha, x) = \sum_{k=0}^{d} P_k(\alpha) x^k \qquad \left(\text{here, we substituted } \alpha \text{ and } x \text{ for } X \text{ and } Y \text{ in } P = \sum_{k=0}^{d} P_k(X) Y^k \right),$$

so that $\sum_{k=0}^{d} P_k(\alpha) x^k = P(\alpha, x) = 0$ (by (242), applied to α and x instead of x and y).

Therefore, Lemma 3.28 **(a)** (applied to $\sum_{k=0}^{d} P_k(\alpha) X^k$ instead of P) yields that

$$\sum_{k=0}^{d} P_k(\alpha) X^k = 0$$

as polynomials (in the indeterminate X). In other words, all coefficients of the polynomial $\sum_{k=0}^{d} P_k(\alpha) X^k$ are 0. In other words, $P_k(\alpha) = 0$ for all $k \in \{0, 1, \ldots, d\}$.

Now, let us forget that we fixed α. We thus have shown that $P_k(\alpha) = 0$ for all $k \in \{0, 1, \ldots, d\}$ and $\alpha \in \mathbb{N}$.

Let us now fix $k \in \{0, 1, \ldots, d\}$. Then, $P_k(\alpha) = 0$ for all $\alpha \in \mathbb{N}$. In other words, $P_k(x) = 0$ for all $x \in \mathbb{N}$. Hence, Lemma 3.28 **(a)** (applied to $P = P_k$) yields that $P_k = 0$ as polynomials.

Let us forget that we fixed k. We thus have proven that $P_k = 0$ as polynomials for each $k \in \{0, 1, \ldots, d\}$. Hence, $P = \sum_{k=0}^{d} \underbrace{P_k(X)}_{=0} Y^k = 0$. This proves Lemma 3.28 **(c)**.

(d) Every $x \in \mathbb{N}$ and $y \in \mathbb{N}$ satisfy

$$(P - Q)(x, y) = P(x, y) - Q(x, y) = 0 \qquad (\text{since } P(x, y) = Q(x, y)).$$

Hence, Lemma 3.28 **(c)** (applied to $P - Q$ instead of P) yields $P - Q = 0$. Thus, $P = Q$. Lemma 3.28 **(d)** is proven. □

Of course, Lemma 3.28 can be generalized to polynomials in more than two variables (the proof of Lemma 3.28 **(c)** essentially suggests how to prove this generalization by induction over the number of variables).[108]

[108] If you know what a commutative ring is, you might wonder whether Lemma 3.28 can also be generalized to polynomials with coefficients from other commutative rings (e.g., from $\mathbb{R}$ or $\mathbb{C}$) instead of rational coefficients. In other words, what happens if we replace "rational coefficients" by "coefficients in R" throughout Lemma 3.28, where R is some commutative ring? (Of course, we will then have to also replace $P(x)$ by $P(x \cdot 1_R)$ and so on.)

The answer is that Lemma 3.28 becomes generally false if we don't require anything more specific on R. However, there are certain conditions on R that make Lemma 3.28 remain valid.

3.3. The Chu-Vandermonde identity

3.3.1. The statements

The following fact is known as the *Chu-Vandermonde identity*[109]:

Theorem 3.29. Let $n \in \mathbb{N}$, $x \in \mathbb{Q}$ and $y \in \mathbb{Q}$. Then,

$$\binom{x+y}{n} = \sum_{k=0}^{n} \binom{x}{k} \binom{y}{n-k}.$$

Let us also give an analogous statement for polynomials:

Theorem 3.30. Let $n \in \mathbb{N}$. Then,

$$\binom{X+Y}{n} = \sum_{k=0}^{n} \binom{X}{k} \binom{Y}{n-k}$$

(an equality between polynomials in two variables X and Y).

We will give two proofs of this theorem: one algebraic, and one combinatorial.[110]

3.3.2. An algebraic proof

First proof of Theorem 3.29. Forget that we fixed n, x and y. We thus must prove that for each $n \in \mathbb{N}$, we have

$$\left(\binom{x+y}{n} = \sum_{k=0}^{n} \binom{x}{k} \binom{y}{n-k} \text{ for all } x \in \mathbb{Q} \text{ and } y \in \mathbb{Q} \right). \tag{243}$$

We shall prove (243) by induction over n:

For instance, Lemma 3.28 remains valid for $R = \mathbb{Z}$, for $R = \mathbb{R}$ and for $R = \mathbb{C}$, as well as for R being any polynomial ring over $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$ or $\mathbb{C}$. More generally, Lemma 3.28 is valid if R is any field of characteristic 0 (i.e., any field such that the elements $n \cdot 1_R$ for n ranging over $\mathbb{N}$ are pairwise distinct), or any subring of such a field.

[109] See the Wikipedia page for part of its history. Usually, the equality $\binom{x+y}{n} = \sum_{k=0}^{n} \binom{x}{k} \binom{y}{n-k}$ for two **nonnegative integers** x and y (this is a particular case of Theorem 3.29) is called the *Vandermonde identity* (or the *Vandermonde convolution identity*), whereas the name "*Chu-Vandermonde identity*" is used for the identity $\binom{X+Y}{n} = \sum_{k=0}^{n} \binom{X}{k} \binom{Y}{n-k}$ in which X and Y are **indeterminates** (this is Theorem 3.30). However, this seems to be mostly a matter of convention (which isn't even universally followed); and anyway the two identities are easily derived from one another as we will see in the second proof of Theorem 3.30.

[110] Note that Theorem 3.29 appears in [GrKnPa94, (5.27)], where it is called *Vandermonde's convolution*. The second proof of Theorem 3.29 we shall show below is just a more detailed writeup of the proof given there.

Induction base: Let $x \in \mathbb{Q}$ and $y \in \mathbb{Q}$. Proposition 3.3 **(a)** (applied to $m = x$) yields $\dbinom{x}{0} = 1$. Proposition 3.3 **(a)** (applied to $m = y$) yields $\dbinom{y}{0} = 1$. Now,

$$\sum_{k=0}^{0} \binom{x}{k} \binom{y}{0-k} = \underbrace{\binom{x}{0}}_{=1} \underbrace{\binom{y}{0-0}}_{=\binom{y}{0}=1} = 1. \tag{244}$$

But Proposition 3.3 **(a)** (applied to $m = x+y$) yields $\dbinom{x+y}{0} = 1$. Compared with (244), this yields $\dbinom{x+y}{0} = \sum\limits_{k=0}^{0} \dbinom{x}{k} \dbinom{y}{0-k}$.

Now, forget that we fixed x and y. We thus have shown that

$$\binom{x+y}{0} = \sum_{k=0}^{0} \binom{x}{k} \binom{y}{0-k} \qquad \text{for all } x \in \mathbb{Q} \text{ and } y \in \mathbb{Q}.$$

In other words, (243) holds for $n = 0$. This completes the induction base.

Induction step: Let N be a positive integer. Assume that (243) holds for $n = N-1$. We need to prove that (243) holds for $n = N$. In other words, we need to prove that

$$\binom{x+y}{N} = \sum_{k=0}^{N} \binom{x}{k} \binom{y}{N-k} \qquad \text{for all } x \in \mathbb{Q} \text{ and } y \in \mathbb{Q}. \tag{245}$$

We have assumed that (243) holds for $n = N-1$. In other words, we have

$$\binom{x+y}{N-1} = \sum_{k=0}^{N-1} \binom{x}{k} \binom{y}{(N-1)-k} \qquad \text{for all } x \in \mathbb{Q} \text{ and } y \in \mathbb{Q}. \tag{246}$$

Now, let us prove (245):

[*Proof of* (245): Fix $x \in \mathbb{Q}$ and $y \in \mathbb{Q}$. Then, (246) (applied to $x-1$ instead of x) yields

$$\begin{aligned} \binom{x-1+y}{N-1} &= \sum_{k=0}^{N-1} \binom{x-1}{k} \binom{y}{(N-1)-k} \\ &= \sum_{k=1}^{N} \binom{x-1}{k-1} \underbrace{\binom{y}{(N-1)-(k-1)}}_{=\binom{y}{N-k}} \\ &\qquad \text{(here, we have substituted } k-1 \text{ for } k \text{ in the sum)} \\ &= \sum_{k=1}^{N} \binom{x-1}{k-1} \binom{y}{N-k}. \end{aligned}$$

Since $x-1+y=x+y-1$, this rewrites as

$$\binom{x+y-1}{N-1}=\sum_{k=1}^{N}\binom{x-1}{k-1}\binom{y}{N-k}. \tag{247}$$

On the other hand, (246) (applied to $y-1$ instead of y) shows that

$$\binom{x+y-1}{N-1}=\sum_{k=0}^{N-1}\binom{x}{k}\underbrace{\binom{y-1}{(N-1)-k}}_{=\binom{y-1}{N-k-1}}=\sum_{k=0}^{N-1}\binom{x}{k}\binom{y-1}{N-k-1}. \tag{248}$$

Next, we notice a simple consequence of (240): We have

$$\frac{x}{N}\binom{x-1}{a-1}=\frac{a}{N}\binom{x}{a} \qquad \text{for every } a\in\{1,2,3,\ldots\} \tag{249}$$

[111]. The same argument (applied to y instead of x) shows that

$$\frac{y}{N}\binom{y-1}{a-1}=\frac{a}{N}\binom{y}{a} \qquad \text{for every } a\in\{1,2,3,\ldots\}. \tag{250}$$

We have

$$\frac{x}{N}\underbrace{\binom{x+y-1}{N-1}}_{\substack{=\sum_{k=1}^{N}\binom{x-1}{k-1}\binom{y}{N-k}\\ \text{(by (247))}}}=\frac{x}{N}\sum_{k=1}^{N}\binom{x-1}{k-1}\binom{y}{N-k}$$

$$=\sum_{k=1}^{N}\underbrace{\frac{x}{N}\binom{x-1}{k-1}}_{\substack{=\frac{k}{N}\binom{x}{k}\\ \text{(by (249),}\\ \text{applied to } a=k)}}\binom{y}{N-k}=\sum_{k=1}^{N}\frac{k}{N}\binom{x}{k}\binom{y}{N-k}.$$

[111] *Proof of (249):* Let $a\in\{1,2,3,\ldots\}$. Then, (240) (applied to $m=x$ and $n=a$) yields $\binom{x}{a}=\frac{x}{a}\binom{x-1}{a-1}$. Hence,

$$\frac{a}{N}\underbrace{\binom{x}{a}}_{=\frac{x}{a}\binom{x-1}{a-1}}=\underbrace{\frac{a}{N}\cdot\frac{x}{a}}_{=\frac{x}{N}}\binom{x-1}{a-1}=\frac{x}{N}\binom{x-1}{a-1}.$$

This proves (249).

Compared with

$$\begin{aligned}\sum_{k=0}^{N}\frac{k}{N}\binom{x}{k}\binom{y}{N-k} &= \underbrace{\frac{0}{N}\binom{x}{0}\binom{y}{N-0}}_{=0}+\sum_{k=1}^{N}\frac{k}{N}\binom{x}{k}\binom{y}{N-k}\\ &\qquad \left(\text{here, we have split off the addend for } k=0\right)\\ &= \sum_{k=1}^{N}\frac{k}{N}\binom{x}{k}\binom{y}{N-k},\end{aligned}$$

this yields

$$\frac{x}{N}\binom{x+y-1}{N-1}=\sum_{k=0}^{N}\frac{k}{N}\binom{x}{k}\binom{y}{N-k}. \tag{251}$$

We also have

$$\begin{aligned}\frac{y}{N}\underbrace{\binom{x+y-1}{N-1}}_{\substack{=\sum\limits_{k=0}^{N-1}\binom{x}{k}\binom{y-1}{N-k-1}\\ \text{(by (248))}}} &= \frac{y}{N}\sum_{k=0}^{N-1}\binom{x}{k}\binom{y-1}{N-k-1}\\ &= \sum_{k=0}^{N-1}\binom{x}{k}\underbrace{\frac{y}{N}\binom{y-1}{N-k-1}}_{\substack{=\frac{N-k}{N}\binom{y}{N-k}\\ \text{(by (250), applied to } a=N-k\\ \text{(since } N-k\in\{1,2,3,\ldots\} \text{ (because } k\in\{0,1,\ldots,N-1\})))}}\\ &= \sum_{k=0}^{N-1}\binom{x}{k}\frac{N-k}{N}\binom{y}{N-k}=\sum_{k=0}^{N-1}\frac{N-k}{N}\binom{x}{k}\binom{y}{N-k}.\end{aligned}$$

Compared with

$$\begin{aligned}\sum_{k=0}^{N}\frac{N-k}{N}\binom{x}{k}\binom{y}{N-k} &= \sum_{k=0}^{N-1}\frac{N-k}{N}\binom{x}{k}\binom{y}{N-k}+\underbrace{\frac{N-N}{N}}_{=0}\binom{x}{N}\binom{y}{N-N}\\ &\qquad \left(\text{here, we have split off the addend for } k=N\right)\\ &= \sum_{k=0}^{N-1}\frac{N-k}{N}\binom{x}{k}\binom{y}{N-k}+\underbrace{0\binom{x}{N}\binom{y}{N-N}}_{=0}\\ &= \sum_{k=0}^{N-1}\frac{N-k}{N}\binom{x}{k}\binom{y}{N-k},\end{aligned}$$

this yields

$$\frac{y}{N}\binom{x+y-1}{N-1} = \sum_{k=0}^{N} \frac{N-k}{N}\binom{x}{k}\binom{y}{N-k}. \tag{252}$$

Now, (240) (applied to $m = x+y$ and $n = N$) yields

$$\begin{aligned}
\binom{x+y}{N} &= \underbrace{\frac{x+y}{N}}_{=\frac{x}{N}+\frac{y}{N}} \binom{x+y-1}{N-1} = \left(\frac{x}{N}+\frac{y}{N}\right)\binom{x+y-1}{N-1} \\
&= \underbrace{\frac{x}{N}\binom{x+y-1}{N-1}}_{\substack{=\sum_{k=0}^{N}\frac{k}{N}\binom{x}{k}\binom{y}{N-k} \\ \text{(by (251))}}} + \underbrace{\frac{y}{N}\binom{x+y-1}{N-1}}_{\substack{=\sum_{k=0}^{N}\frac{N-k}{N}\binom{x}{k}\binom{y}{N-k} \\ \text{(by (252))}}} \\
&= \sum_{k=0}^{N} \frac{k}{N}\binom{x}{k}\binom{y}{N-k} + \sum_{k=0}^{N} \frac{N-k}{N}\binom{x}{k}\binom{y}{N-k} \\
&= \sum_{k=0}^{N} \left(\underbrace{\frac{k}{N}+\frac{N-k}{N}}_{=1}\right)\binom{x}{k}\binom{y}{N-k} = \sum_{k=0}^{N}\binom{x}{k}\binom{y}{N-k}.
\end{aligned}$$

This proves (245).]

We have thus proven (245). In other words, (243) holds for $n = N$. This completes the induction step. Thus, the induction proof of (243) is complete. Hence, Theorem 3.29 is proven. □

The above proof has the advantage of being completely algebraic; it thus does not rely on what x and y actually are. It works equally well if x and y are assumed to be real numbers or complex numbers or polynomials. Thus, it can also be used to prove Theorem 3.30:

First proof of Theorem 3.30. To obtain a proof of Theorem 3.30, replace every appearance of "x" by "X" and every appearance of "y" by "Y" in the above First proof of Theorem 3.29. □

3.3.3. A combinatorial proof

We shall next give a different, combinatorial proof of Theorems 3.29 and 3.30. This proof is somewhat indirect, as it begins by showing the following particular case of Theorem 3.29:

Lemma 3.31. Let $n \in \mathbb{N}$, $x \in \mathbb{N}$ and $y \in \mathbb{N}$. Then,

$$\binom{x+y}{n} = \sum_{k=0}^{n} \binom{x}{k} \binom{y}{n-k}. \tag{253}$$

This lemma is less general than Theorem 3.29, since it requires x and y to belong to $\mathbb{N}$.

Proof of Lemma 3.31. For every $N \in \mathbb{N}$, we let $[N]$ denote the N-element set $\{1, 2, \ldots, N\}$.

Recall that $\binom{x+y}{n}$ is the number of n-element subsets of a given $(x+y)$-element set[112]. Since $[x+y]$ is an $(x+y)$-element set, we thus conclude that $\binom{x+y}{n}$ is the number of n-element subsets of $[x+y]$.

But let us count the n-element subsets of $[x+y]$ in a different way (i.e., find a different expression for the number of n-element subsets of $[x+y]$). Namely, we can choose an n-element subset S of $[x+y]$ by means of the following process:

1. We decide how many elements of this subset S will be among the numbers $1, 2, \ldots, x$. Let k be the number of these elements. Clearly, k must be an integer between 0 and n (inclusive)[113].

2. Then, we choose these k elements of S among the numbers $1, 2, \ldots, x$. This can be done in $\binom{x}{k}$ different ways (because we are choosing k out of x numbers, with no repetitions, and with no regard for their order; in other words, we are choosing a k-element subset of $\{1, 2, \ldots, x\}$).

3. Then, we choose the remaining $n-k$ elements of S (because S should have n elements in total) among the remaining numbers $x+1, x+2, \ldots, x+y$. This can be done in $\binom{y}{n-k}$ ways (because we are choosing $n-k$ out of y numbers, with no repetitions, and with no regard for their order).

This process makes it clear that the total number of ways to choose an n-element subset S of $[x+y]$ is $\sum_{k=0}^{n} \binom{x}{k} \binom{y}{n-k}$. In other words, the number of n-element subsets of $[x+y]$ is $\sum_{k=0}^{n} \binom{x}{k} \binom{y}{n-k}$. But earlier, we have shown that the same

[112] This follows from (236).

[113] Because the subset S will have n elements in total, and thus at most n of them can be among the numbers $1, 2, \ldots, x$.

number is $\dbinom{x+y}{n}$. Comparing these two results, we conclude that $\dbinom{x+y}{n} = \sum\limits_{k=0}^{n} \dbinom{x}{k}\dbinom{y}{n-k}$. Thus, Lemma 3.31 is proven. □

We now shall leverage Lemma 3.28 to derive Theorem 3.30 from this lemma:

Second proof of Theorem 3.30. We define two polynomials P and Q in the indeterminates X and Y with rational coefficients by setting

$$P = \binom{X+Y}{n};$$
$$Q = \sum_{k=0}^{n} \binom{X}{k}\binom{Y}{n-k}$$

[114]. The equality (253) (which we have proven) states that $P(x,y) = Q(x,y)$ for all $x \in \mathbb{N}$ and $y \in \mathbb{N}$. Thus, Lemma 3.28 **(d)** yields that $P = Q$. Recalling how P and Q are defined, we can rewrite this as $\dbinom{X+Y}{n} = \sum\limits_{k=0}^{n} \dbinom{X}{k}\dbinom{Y}{n-k}$. This proves Theorem 3.30. □

The argument that we used at the end of the above proof to derive Theorem 3.30 from (253) is a very common argument that appears in proofs of equalities for binomial coefficients. The binomial coefficients $\dbinom{m}{n}$ are defined for arbitrary rational, real or complex m [115], but their combinatorial interpretation (via counting subsets) only makes sense when m and n are nonnegative integers. Thus, if we want to prove an identity of the form $P = Q$ (where P and Q are two polynomials, say, in two indeterminates X and Y) using the combinatorial interpretation of binomial coefficients, then a reasonable tactic is to first show that $P(x,y) = Q(x,y)$ for all $x \in \mathbb{N}$ and $y \in \mathbb{N}$ (using combinatorics), and then to use something like Lemma 3.28 in order to conclude that P and Q are equal as polynomials. We shall see this tactic used a few more times.[116]

[114] These are both polynomials since $\dbinom{X+Y}{n}$, $\dbinom{X}{k}$ and $\dbinom{Y}{n-k}$ are polynomials in X and Y.

[115] For example, terms like $\dbinom{-1/2}{3}$, $\dbinom{2+\sqrt{3}}{5}$ and $\dbinom{-7}{0}$ make perfect sense. (But we cannot substitute arbitrary complex numbers for n in $\dbinom{m}{n}$. So far we have only defined $\dbinom{m}{n}$ for $n \in \mathbb{N}$. It is usual to define $\dbinom{m}{n}$ to mean 0 for negative integers n, and using analysis (specifically, the Γ function) it is possible to give a reasonable meaning to $\dbinom{m}{n}$ for m and n being reals, but this will no longer be a polynomial in m.)

[116] This tactic is called "the polynomial argument" in [GrKnPa94, §5.1].

Second proof of Theorem 3.29. Theorem 3.30 yields

$$\binom{X+Y}{n} = \sum_{k=0}^{n} \binom{X}{k} \binom{Y}{n-k}$$

(an equality between polynomials in two variables X and Y). Now, let us evaluate both sides of this equality at $X = x$ and $Y = y$. As a result, we obtain

$$\binom{x+y}{n} = \sum_{k=0}^{n} \binom{x}{k} \binom{y}{n-k}$$

(because of Proposition 3.25). □

3.3.4. Some applications

Let us give some sample applications of Theorem 3.29:

Proposition 3.32. (a) For every $x \in \mathbb{Z}$ and $y \in \mathbb{Z}$ and $n \in \mathbb{N}$, we have

$$\binom{x+y}{n} = \sum_{k=0}^{n} \binom{x}{k} \binom{y}{n-k}.$$

(b) For every $x \in \mathbb{N}$ and $y \in \mathbb{Z}$, we have

$$\binom{x+y}{x} = \sum_{k=0}^{x} \binom{x}{k} \binom{y}{k}.$$

(c) For every $n \in \mathbb{N}$, we have

$$\binom{2n}{n} = \sum_{k=0}^{n} \binom{n}{k}^2.$$

(d) For every $x \in \mathbb{Z}$ and $y \in \mathbb{Z}$ and $n \in \mathbb{N}$, we have

$$\binom{x-y}{n} = \sum_{k=0}^{n} (-1)^k \binom{x}{n-k} \binom{k+y-1}{k}.$$

(e) For every $x \in \mathbb{N}$ and $y \in \mathbb{Z}$ and $n \in \mathbb{N}$ with $x \leq n$, we have

$$\binom{y-x-1}{n-x} = \sum_{k=0}^{n} (-1)^{k-x} \binom{k}{x} \binom{y}{n-k}.$$

(f) For every $x \in \mathbb{N}$ and $y \in \mathbb{N}$ and $n \in \mathbb{N}$, we have

$$\binom{n+1}{x+y+1} = \sum_{k=0}^{n} \binom{k}{x} \binom{n-k}{y}.$$

(g) For every $x \in \mathbb{Z}$ and $y \in \mathbb{N}$ and $n \in \mathbb{N}$ satisfying $x + y \geq 0$ and $n \geq x$, we have

$$\binom{x+y}{n} = \sum_{k=0}^{x+y} \binom{x}{k} \binom{y}{n+k-x}.$$

Remark 3.33. I have learnt Proposition 3.32 **(f)** from the AoPS forum. Proposition 3.32 **(g)** is a generalization of Proposition 3.32 **(b)**.

Note that if we apply Proposition 3.32 **(f)** to $y = 0$, then we obtain the identity $\dbinom{n+1}{x+1} = \sum\limits_{k=0}^{n} \dbinom{k}{x}$ for all $n \in \mathbb{N}$ and $x \in \mathbb{N}$. This identity is also a particular case of Exercise 3.3 **(b)**.

Proof of Proposition 3.32. **(a)** Proposition 3.32 **(a)** is a particular case of Theorem 3.29.

(b) Let $x \in \mathbb{N}$ and $y \in \mathbb{Z}$. Proposition 3.32 **(a)** (applied to y, x and x instead of x, y and n) yields

$$\binom{y+x}{x} = \sum_{k=0}^{x} \binom{y}{k} \binom{x}{x-k}.$$

Compared with

$$\sum_{k=0}^{x} \underbrace{\binom{x}{k}}_{\substack{=\binom{x}{x-k} \\ \text{(by (232), applied to } m=x \text{ and } n=k \\ \text{(since } x \geq k \text{ (because } k \leq x)))}} \binom{y}{k} = \sum_{k=0}^{x} \binom{x}{x-k} \binom{y}{k} = \sum_{k=0}^{x} \binom{y}{k} \binom{x}{x-k},$$

this yields $\dbinom{y+x}{x} = \sum\limits_{k=0}^{x} \dbinom{x}{k} \dbinom{y}{k}$. Since $y + x = x + y$, this rewrites as $\dbinom{x+y}{x} = \sum\limits_{k=0}^{x} \dbinom{x}{k} \dbinom{y}{k}$. This proves Proposition 3.32 **(b)**.

(c) Let $n \in \mathbb{N}$. Applying Proposition 3.32 **(b)** to $x = n$ and $y = n$, we obtain

$$\binom{n+n}{n} = \sum_{k=0}^{n} \underbrace{\binom{n}{k} \binom{n}{k}}_{=\binom{n}{k}^2} = \sum_{k=0}^{n} \binom{n}{k}^2.$$

Since $n + n = 2n$, this rewrites as $\dbinom{2n}{n} = \sum\limits_{k=0}^{n} \dbinom{n}{k}^2$. This proves Proposition 3.32 **(c)**.

(d) Let $x \in \mathbb{Z}$ and $y \in \mathbb{Z}$ and $n \in \mathbb{N}$. Proposition 3.32 **(a)** (applied to $-y$ instead of y) yields

$$\begin{aligned}\binom{x+(-y)}{n} &= \sum_{k=0}^{n} \binom{x}{k}\binom{-y}{n-k} = \sum_{k=0}^{n} \binom{x}{n-k} \underbrace{\binom{-y}{n-(n-k)}}_{=\binom{-y}{k}} \\ &\qquad \text{(here, we substituted } n-k \text{ for } k \text{ in the sum)} \\ &= \sum_{k=0}^{n} \binom{x}{n-k} \underbrace{\binom{-y}{k}}_{\substack{=(-1)^k \binom{k-(-y)-1}{k} \\ \text{(by (237), applied to } k \text{ and } -y \text{ instead of } n \text{ and } m)}} \\ &= \sum_{k=0}^{n} \underbrace{\binom{x}{n-k}(-1)^k}_{=(-1)^k\binom{x}{n-k}} \underbrace{\binom{k-(-y)-1}{k}}_{=\binom{k+y-1}{k}} \\ &= \sum_{k=0}^{n} (-1)^k \binom{x}{n-k}\binom{k+y-1}{k}.\end{aligned}$$

Since $x+(-y) = x-y$, this rewrites as $\binom{x-y}{n} = \sum_{k=0}^{n} (-1)^k \binom{x}{n-k}\binom{k+y-1}{k}$. This proves Proposition 3.32 **(d)**.

(e) Let $x \in \mathbb{N}$ and $y \in \mathbb{Z}$ and $n \in \mathbb{N}$ be such that $x \leq n$. From $x \in \mathbb{N}$, we obtain $0 \leq x$ and thus $0 \leq x \leq n$. We notice that every integer $k \geq x$ satisfies

$$\binom{k}{k-x} = \binom{k}{x} \tag{254}$$

[117]. Furthermore, $n - x \in \mathbb{N}$ (since $x \leq n$). Hence, we can apply Proposition 3.32

[117] *Proof of (254):* Let k be an integer such that $k \geq x$. Thus, $k - x \in \mathbb{N}$. Also, $k \geq x \geq 0$ (since $x \in \mathbb{N}$), and thus $k \in \mathbb{N}$. Now, recall that $k \geq x$. Hence, (232) (applied to k and x instead of m and n) yields $\binom{k}{x} = \binom{k}{k-x}$. This proves (254).

(a) to y, $-x-1$ and $n-x$ instead of x, y and n. As a result, we obtain

$$\begin{aligned}
\binom{y+(-x-1)}{n-x} &= \sum_{k=0}^{n-x} \binom{y}{k} \underbrace{\binom{-x-1}{(n-x)-k}}_{\substack{=(-1)^{(n-x)-k}\binom{((n-x)-k)-(-x-1)-1}{(n-x)-k} \\ \text{(by (237), applied to } -x-1 \text{ and } (n-x)-k \\ \text{instead of } m \text{ and } n\text{)}}} \\
&= \sum_{k=0}^{n-x} \binom{y}{k} (-1)^{(n-x)-k} \underbrace{\binom{((n-x)-k)-(-x-1)-1}{(n-x)-k}}_{\substack{=\binom{n-k}{(n-x)-k} \\ \text{(since } ((n-x)-k)-(-x-1)-1=n-k\text{)}}} \\
&= \sum_{k=0}^{n-x} \binom{y}{k} (-1)^{(n-x)-k} \binom{n-k}{(n-x)-k} \\
&= \underbrace{\sum_{k=n-(n-x)}^{n}}_{\substack{=\sum_{k=x}^{n} \\ \text{(since } n-(n-x)=x\text{)}}} \binom{y}{n-k} \underbrace{(-1)^{(n-x)-(n-k)}}_{\substack{=(-1)^{k-x} \\ \text{(since } (n-x)-(n-k)=k-x\text{)}}} \underbrace{\binom{n-(n-k)}{(n-x)-(n-k)}}_{\substack{=\binom{k}{k-x} \\ \text{(since } n-(n-k)=k \\ \text{and } (n-x)-(n-k)=k-x\text{)}}} \\
&\qquad \text{(here, we have substituted } n-k \text{ for } k \text{ in the sum)} \\
&= \sum_{k=x}^{n} \binom{y}{n-k} (-1)^{k-x} \underbrace{\binom{k}{k-x}}_{\substack{=\binom{k}{x} \\ \text{(by (254))}}} = \sum_{k=x}^{n} \binom{y}{n-k} (-1)^{k-x} \binom{k}{x} \\
&= \sum_{k=x}^{n} (-1)^{k-x} \binom{k}{x} \binom{y}{n-k}.
\end{aligned}$$

Compared with

$$\begin{aligned}
&\sum_{k=0}^{n}(-1)^{k-x}\binom{k}{x}\binom{y}{n-k} \\
&= \sum_{k=0}^{x-1}(-1)^{k-x}\underbrace{\binom{k}{x}}_{\substack{=0 \\ \text{(by (231), applied to } k \text{ and } x \\ \text{instead of } m \text{ and } n \text{ (since } k\leq x-1<x\text{))}}}\binom{y}{n-k}+\sum_{k=x}^{n}(-1)^{k-x}\binom{k}{x}\binom{y}{n-k} \\
&\qquad (\text{since } 0\leq x\leq n) \\
&= \underbrace{\sum_{k=0}^{x-1}(-1)^{k-x}0\binom{y}{n-k}}_{=0}+\sum_{k=x}^{n}(-1)^{k-x}\binom{k}{x}\binom{y}{n-k}=\sum_{k=x}^{n}(-1)^{k-x}\binom{k}{x}\binom{y}{n-k},
\end{aligned}$$

this yields

$$\binom{y+(-x-1)}{n-x}=\sum_{k=0}^{n}(-1)^{k-x}\binom{k}{x}\binom{y}{n-k}.$$

In other words,

$$\binom{y-x-1}{n-x}=\sum_{k=0}^{n}(-1)^{k-x}\binom{k}{x}\binom{y}{n-k}$$

(since $y+(-x-1)=y-x-1$). This proves Proposition 3.32 **(e)**.

(f) Let $x\in\mathbb{N}$ and $y\in\mathbb{N}$ and $n\in\mathbb{N}$. We must be in one of the following two cases:

Case 1: We have $n<x+y$.

Case 2: We have $n\geq x+y$.

Let us first consider Case 1. In this case, we have $n<x+y$. Thus, $n+1<x+y+1$. Therefore, $\binom{n+1}{x+y+1}=0$ (by (231), applied to $n+1$ and $x+y+1$ instead of m and n). But every $k\in\{0,1,\ldots,n\}$ satisfies $\binom{k}{x}\binom{n-k}{y}=0$ [118].

[118] *Proof.* Let $k\in\{0,1,\ldots,n\}$. We need to show that $\binom{k}{x}\binom{n-k}{y}=0$.

If we have $k<x$, then we have $\binom{k}{x}=0$ (by (231), applied to k and x instead of m and n). Therefore, if we have $k<x$, then $\underbrace{\binom{k}{x}}_{=0}\binom{n-k}{y}=0$. Hence, for the rest of this proof of $\binom{k}{x}\binom{n-k}{y}=0$, we can WLOG assume that we don't have $k<x$. Assume this.

We have $k\leq n$ (since $k\in\{0,1,\ldots,n\}$) and thus $n-k\in\mathbb{N}$.

We have $k\geq x$ (since we don't have $k<x$), and thus $n-\underbrace{k}_{\geq x}\leq n-x<y$ (since $n<$

Hence, $\sum_{k=0}^{n} \underbrace{\binom{k}{x}\binom{n-k}{y}}_{=0} = \sum_{k=0}^{n} 0 = 0$. Compared with $\binom{n+1}{x+y+1} = 0$, this yields $\binom{n+1}{x+y+1} = \sum_{k=0}^{n} \binom{k}{x}\binom{n-k}{y}$. Thus, Proposition 3.32 **(f)** is proven in Case 1.

Let us now consider Case 2. In this case, we have $n \geq x+y$. Hence, $n-y \geq x$ (since $x \in \mathbb{N}$), so that $(n-y)-x \in \mathbb{N}$. Also, $n-y \geq x \geq 0$ and thus $n-y \in \mathbb{N}$. Moreover, $x \leq n-y$. Therefore, we can apply Proposition 3.32 **(e)** to $-y-1$ and $n-y$ instead of y and n. As a result, we obtain

$$\begin{aligned}
&\binom{(-y-1)-x-1}{(n-y)-x} \\
&= \sum_{k=0}^{n-y} (-1)^{k-x} \binom{k}{x} \underbrace{\binom{-y-1}{(n-y)-k}}_{\substack{=(-1)^{(n-y)-k}\binom{((n-y)-k)-(-y-1)-1}{(n-y)-k} \\ \text{(by (237), applied to } -y-1 \text{ and } (n-y)-k \\ \text{instead of } m \text{ and } n)}} \\
&= \sum_{k=0}^{n-y} \underbrace{(-1)^{k-x} \binom{k}{x} (-1)^{(n-y)-k}}_{=(-1)^{(n-y)-k}\binom{k}{x}} \underbrace{\binom{((n-y)-k)-(-y-1)-1}{(n-y)-k}}_{\substack{=\binom{n-k}{(n-y)-k} \\ \text{(since } ((n-y)-k)-(-y-1)-1=n-k)}} \\
&= \sum_{k=0}^{n-y} \underbrace{(-1)^{k-x}(-1)^{(n-y)-k}}_{\substack{=(-1)^{(k-x)+((n-y)-k)}=(-1)^{n-x-y} \\ \text{(since } (k-x)+((n-y)-k)=n-x-y)}} \binom{k}{x}\binom{n-k}{(n-y)-k} \\
&= \sum_{k=0}^{n-y} (-1)^{n-x-y} \binom{k}{x}\binom{n-k}{(n-y)-k} \\
&= (-1)^{n-x-y} \sum_{k=0}^{n-y} \binom{k}{x}\binom{n-k}{(n-y)-k}. \qquad (255)
\end{aligned}$$

But every $k \in \{0, 1, \ldots, n-y\}$ satisfies

$$\binom{n-k}{(n-y)-k} = \binom{n-k}{y} \qquad (256)$$

$x+y$). Hence, $\binom{n-k}{y} = 0$ (by (231), applied to $n-k$ and y instead of m and n). Therefore, $\underbrace{\binom{k}{x}\binom{n-k}{y}}_{=0} = 0$, qed.

[119]. Thus, (255) yields

$$\binom{(-y-1)-x-1}{(n-y)-x} = (-1)^{n-x-y} \sum_{k=0}^{n-y} \binom{k}{x} \underbrace{\binom{n-k}{(n-y)-k}}_{\substack{=\binom{n-k}{y} \\ \text{(by (256))}}} = (-1)^{n-x-y} \sum_{k=0}^{n-y} \binom{k}{x} \binom{n-k}{y}.$$

Compared with

$$\binom{(-y-1)-x-1}{(n-y)-x} = \underbrace{(-1)^{(n-y)-x}}_{\substack{=(-1)^{n-x-y} \\ \text{(since } (n-y)-x=n-x-y)}} \underbrace{\binom{((n-y)-x)-((-y-1)-x-1)-1}{(n-y)-x}}_{\substack{=\binom{n+1}{n-x-y} \\ \text{(since } ((n-y)-x)-((-y-1)-x-1)-1=n+1 \\ \text{and } (n-y)-x=n-x-y)}}$$

$$\left(\begin{array}{c}\text{by (237), applied to } (-y-1)-x-1 \text{ and } (n-y)-x \\ \text{instead of } m \text{ and } n\end{array}\right)$$

$$= (-1)^{n-x-y} \binom{n+1}{n-x-y},$$

this yields

$$(-1)^{n-x-y} \binom{n+1}{n-x-y} = (-1)^{n-x-y} \sum_{k=0}^{n-y} \binom{k}{x} \binom{n-k}{y}.$$

We can cancel $(-1)^{n-x-y}$ from this equality (because $(-1)^{n-x-y} \neq 0$). As a result, we obtain

$$\binom{n+1}{n-x-y} = \sum_{k=0}^{n-y} \binom{k}{x} \binom{n-k}{y}. \tag{257}$$

But $0 \leq n-y$ (since $n-y \in \mathbb{N}$) and $n-y \leq n$ (since $y \in \mathbb{N}$). Also, every $k \in \{n-y+1, n-y+2, \ldots, n\}$ satisfies

$$\binom{n-k}{y} = 0 \tag{258}$$

[119] *Proof of (256):* Let $k \in \{0, 1, \ldots, n-y\}$. Then, $k \in \mathbb{N}$ and $n-y \geq k$. From $n-y \geq k$, we obtain $n \geq y+k$, so that $n-k \geq y$. Thus, $n-k \geq y \geq 0$, so that $n-k \in \mathbb{N}$. Hence, (232) (applied to $n-k$ and y instead of m and n) yields $\binom{n-k}{y} = \binom{n-k}{(n-k)-y} = \binom{n-k}{(n-y)-k}$ (since $(n-k)-y = (n-y)-k$). This proves (256).

[120]. Hence,

$$\sum_{k=0}^{n}\binom{k}{x}\binom{n-k}{y}=\sum_{k=0}^{n-y}\binom{k}{x}\binom{n-k}{y}+\sum_{k=n-y+1}^{n}\binom{k}{x}\underbrace{\binom{n-k}{y}}_{\substack{=0\\ \text{(by (258))}}}$$

$$\left(\text{since } 0\leq n-y\leq n\right)$$

$$=\sum_{k=0}^{n-y}\binom{k}{x}\binom{n-k}{y}+\underbrace{\sum_{k=n-y+1}^{n}\binom{k}{x}0}_{=0}=\sum_{k=0}^{n-y}\binom{k}{x}\binom{n-k}{y}$$

$$=\binom{n+1}{n-x-y}\qquad\text{(by (257))}. \tag{259}$$

Finally, $n+1\in\mathbb{N}$ and $x+y+1\in\mathbb{N}$ (since $x\in\mathbb{N}$ and $y\in\mathbb{N}$) and $\underbrace{n}_{\geq x+y}+1\geq x+y+1$. Hence, (232) (applied to $n+1$ and $x+y+1$ instead of m and n) yields

$$\binom{n+1}{x+y+1}=\binom{n+1}{(n+1)-(x+y+1)}=\binom{n+1}{n-x-y}$$

(since $(n+1)-(x+y+1)=n-x-y$). Comparing this with (259), we obtain

$$\binom{n+1}{x+y+1}=\sum_{k=0}^{n}\binom{k}{x}\binom{n-k}{y}.$$

Thus, Proposition 3.32 **(f)** is proven in Case 2.

We have now proven Proposition 3.32 **(f)** in both Cases 1 and 2; thus, Proposition 3.32 **(f)** always holds.

(g) Let $x\in\mathbb{Z}$ and $y\in\mathbb{N}$ and $n\in\mathbb{N}$ be such that $x+y\geq 0$ and $n\geq x$. We have $x+y\in\mathbb{N}$ (since $x+y\geq 0$). We must be in one of the following two cases:

Case 1: We have $x+y<n$.

Case 2: We have $x+y\geq n$.

Let us first consider Case 1. In this case, we have $x+y<n$. Thus, $\binom{x+y}{n}=0$ (by (231), applied to $m=x+y$). But every $k\in\{0,1,\ldots,x+y\}$ satisfies $\binom{y}{n+k-x}=$

[120] *Proof of (258):* Let $k\in\{n-y+1,n-y+2,\ldots,n\}$. Then, $k\leq n$ and $k>n-y$. Hence, $n-k\in\mathbb{N}$ (since $k\leq n$) and $n-\underbrace{k}_{>n-y}<n-(n-y)=y$. Therefore, (231) (applied to $n-k$ and y instead of m and n) yields $\binom{n-k}{y}=0$. This proves (258).

0 [121]. Thus, $\sum_{k=0}^{x+y} \binom{x}{k} \underbrace{\binom{y}{n+k-x}}_{=0} = \sum_{k=0}^{x+y} \binom{x}{k} 0 = 0$. Compared with $\binom{x+y}{n} = 0$, this yields $\binom{x+y}{n} = \sum_{k=0}^{x+y} \binom{x}{k} \binom{y}{n+k-x}$. Thus, Proposition 3.32 **(g)** is proven in Case 1.

Let us now consider Case 2. In this case, we have $x+y \geq n$. Hence, $\binom{x+y}{n} = \binom{x+y}{x+y-n}$ (by (232), applied to $m = x+y$). Also, $x+y-n \in \mathbb{N}$ (since $x+y \geq n$). Therefore, Proposition 3.32 **(a)** (applied to $x+y-n$ instead of n) yields

$$\binom{x+y}{x+y-n} = \sum_{k=0}^{x+y-n} \binom{x}{k} \binom{y}{x+y-n-k}.$$

Since $\binom{x+y}{n} = \binom{x+y}{x+y-n}$, this rewrites as

$$\binom{x+y}{n} = \sum_{k=0}^{x+y-n} \binom{x}{k} \binom{y}{x+y-n-k}. \qquad (260)$$

But every $k \in \{0,1,\ldots,x+y-n\}$ satisfies $\binom{y}{x+y-n-k} = \binom{y}{n+k-x}$ [122]. Hence, (260) becomes

$$\binom{x+y}{n} = \sum_{k=0}^{x+y-n} \binom{x}{k} \underbrace{\binom{y}{x+y-n-k}}_{=\binom{y}{n+k-x}} = \sum_{k=0}^{x+y-n} \binom{x}{k} \binom{y}{n+k-x}. \qquad (261)$$

On the other hand, we have $0 \leq n \leq x+y$ and thus $0 \leq x+y-n \leq x+y$. But every $k \in \mathbb{N}$ satisfying $k > x+y-n$ satisfies

$$\binom{y}{n+k-x} = 0 \qquad (262)$$

[121] *Proof.* Let $k \in \{0,1,\ldots,x+y\}$. Then, $k \geq 0$, so that $n + \underbrace{k}_{\geq 0} - x \geq n - x > y$ (since $n > x+y$ (since $x+y < n$)). In other words, $y < n+k-x$. Also, $n+k-x > y \geq 0$, so that $n+k-x \in \mathbb{N}$. Hence, $\binom{y}{n+k-x} = 0$ (by (231), applied to y and $n+k-x$ instead of m and n). Qed.

[122] *Proof.* Let $k \in \{0,1,\ldots,x+y-n\}$. Then, $0 \leq k \leq x+y-n$. Hence, $x+y-n \geq k$, so that $x+y-n-k \in \mathbb{N}$. Also, $y \geq x+y-n-k$ (since $y-(x+y-n-k) = \underbrace{n}_{\geq x} + \underbrace{k}_{\geq 0} - x \geq x+0-x = 0$). Therefore, (232) (applied to y and $x+y-n-k$ instead of m and n) yields $\binom{y}{x+y-n-k} = \binom{y}{y-(x+y-n-k)} = \binom{y}{n+k-x}$ (since $y-(x+y-n-k) = n+k-x$), qed.

[123]. Hence,

$$\begin{aligned}
&\sum_{k=0}^{x+y} \binom{x}{k} \binom{y}{n+k-x} \\
&= \sum_{k=0}^{x+y-n} \binom{x}{k} \binom{y}{n+k-x} + \sum_{k=(x+y-n)+1}^{x+y} \binom{x}{k} \underbrace{\binom{y}{n+k-x}}_{\substack{=0 \\ \text{(by (262) (since } k \geq (x+y-n)+1 > x+y-n))}} \\
&\qquad \text{(since } 0 \leq x+y-n \leq x+y) \\
&= \sum_{k=0}^{x+y-n} \binom{x}{k} \binom{y}{n+k-x} + \underbrace{\sum_{k=(x+y-n)+1}^{x+y} \binom{x}{k} 0}_{=0} = \sum_{k=0}^{x+y-n} \binom{x}{k} \binom{y}{n+k-x}.
\end{aligned}$$

Compared with (261), this yields

$$\binom{x+y}{n} = \sum_{k=0}^{x+y} \binom{x}{k} \binom{y}{n+k-x}.$$

This proves Proposition 3.32 **(g)** in Case 2.

Proposition 3.32 **(g)** is thus proven in each of the two Cases 1 and 2. Therefore, Proposition 3.32 **(g)** holds in full generality. □

Remark 3.34. The proof of Proposition 3.32 given above illustrates a useful technique: the use of upper negation (i.e., the equality (237)) to transform one equality into another. In a nutshell,

- we have proven Proposition 3.32 **(d)** by applying Proposition 3.32 **(a)** to $-y$ instead of y, and then rewriting the result using upper negation;
- we have proven Proposition 3.32 **(e)** by applying Proposition 3.32 **(a)** to y, $-x-1$ and $n-x$ instead of x, y and n, and then rewriting the resulting identity using upper negation;
- we have proven Proposition 3.32 **(f)** by applying Proposition 3.32 **(e)** to $-y-1$ and $n-y$ instead of y and n, and rewriting the resulting identity using upper negation.

Thus, by substitution and rewriting using upper negation, one single equality (namely, Proposition 3.32 **(a)**) has morphed into three other equalities. Note, in particular, that no negative numbers appear in Proposition 3.32 **(f)**, but yet we proved it by substituting negative values for x and y.

[123] *Proof.* Let $k \in \mathbb{N}$ be such that $k > x+y-n$. Then, $n + \underbrace{k}_{>x+y-n} - x > n + (x+y-n) - x = y$. In other words, $y < n+k-x$. Also, $n+k-x > y \geq 0$, so that $n+k-x \in \mathbb{N}$. Hence, (231) (applied to y and $n+k-x$ instead of m and n) yields $\binom{y}{n+k-x} = 0$, qed.

3.4. Further results

Exercise 3.8. Let n be a nonnegative integer. Prove that there exist **nonnegative** integers $c_{i,j}$ for all $0 \leq i \leq n$ and $0 \leq j \leq n$ such that

$$\binom{XY}{n} = \sum_{i=0}^{n} \sum_{j=0}^{n} c_{i,j} \binom{X}{i} \binom{Y}{j} \tag{263}$$

(an equality between polynomials in two variables X and Y).

Notice that the integers $c_{i,j}$ in Exercise 3.8 can depend on the n (besides depending on i and j). We just have not included the n in the notation because it is fixed.

We shall now state two results that are used by Lee and Schiffler in their celebrated proof of positivity for cluster algebras [LeeSch13] (one of the recent breakthroughs in cluster algebra theory). Specifically, our Exercise 3.9 is (essentially) [LeeSch13, Lemma 5.11], and our Proposition 3.35 is (essentially) [LeeSch13, Lemma 5.12][124].

Exercise 3.9. Let a, b and c be three nonnegative integers. Prove that the polynomial $\binom{aX+b}{c}$ in the variable X (this is a polynomial in X of degree $\leq c$) can be written as a sum $\sum_{i=0}^{c} d_i \binom{X}{i}$ with **nonnegative** d_i.

Proposition 3.35. Let a and b be two nonnegative integers. There exist **nonnegative** integers $e_0, e_1, \ldots, e_{a+b}$ such that

$$\binom{X}{a} \binom{X}{b} = \sum_{i=0}^{a+b} e_i \binom{X}{i}$$

(an equality between polynomials in X).

First proof of Proposition 3.35. For every $N \in \mathbb{N}$, we let $[N]$ denote the N-element set $\{1, 2, \ldots, N\}$.

For every set S, we let an *S-junction* mean a pair (A, B), where A is an a-element subset of S and where B is a b-element subset of S such that $A \cup B = S$. (We do not mention a and b in our notation, because a and b are fixed.)

For example, if $a = 2$ and $b = 3$, then $(\{1,4\}, \{2,3,4\})$ is a [4]-junction, and $(\{2,4\}, \{1,4,6\})$ is a $\{1,2,4,6\}$-junction, but $(\{1,3\}, \{2,3,5\})$ is not a [5]-junction (since $\{1,3\} \cup \{2,3,5\} \neq [5]$).

[124] We say "essentially" because the X in [LeeSch13, Lemma 5.11] and in [LeeSch13, Lemma 5.12] is a variable ranging over the nonnegative integers rather than an indeterminate. But this does not make much of a difference (indeed, Lemma 3.28 **(b)** allows us to easily derive our Exercise 3.9 and Proposition 3.35 from [LeeSch13, Lemma 5.11] and [LeeSch13, Lemma 5.12], and of course the converse implication is obvious).

For every $i \in \mathbb{N}$, we let e_i be the number of all $[i]$-junctions. Then, if S is any i-element set, then

$$e_i \text{ is the number of all } S\text{-junctions} \tag{264}$$

[125].

Now, let us show that

$$\binom{x}{a}\binom{x}{b} = \sum_{i=0}^{a+b} e_i \binom{x}{i} \tag{265}$$

for every $x \in \mathbb{N}$.

[*Proof of (265):* Let $x \in \mathbb{N}$. How many ways are there to choose a pair (A, B) consisting of an a-element subset A of $[x]$ and a b-element subset B of $[x]$?

Let us give two different answers to this question. The first answer is the straightforward one: To choose a pair (A, B) consisting of an a-element subset A of $[x]$ and a b-element subset B of $[x]$, we need to choose an a-element subset A of $[x]$ and a b-element subset B of $[x]$. There are $\binom{x}{a}\binom{x}{b}$ total ways to do this (since there are $\binom{x}{a}$ choices for A [126], and $\binom{x}{b}$ choices for B [127], and these choices are independent). In other words, the number of all pairs (A, B) consisting of an a-element subset A of $[x]$ and a b-element subset B of $[x]$ equals $\binom{x}{a}\binom{x}{b}$.

On the other hand, here is a more imaginative procedure to choose a pair (A, B) consisting of an a-element subset A of $[x]$ and a b-element subset B of $[x]$:

1. We choose how many elements the union $A \cup B$ will have. In other words, we choose an $i \in \mathbb{N}$ that will satisfy $|A \cup B| = i$. This i must be an integer between 0 and $a + b$ (inclusive)[128].

[125] *Proof of (264):* Let S be any i-element set. We know that e_i is the number of all $[i]$-junctions. We want to prove that e_i is the number of all S-junctions. Roughly speaking, this is obvious, because we can "relabel the elements of S as $1, 2, \ldots, i$" (since S is an i-element set), and then the S-junctions become precisely the $[i]$-junctions.

Here is a formal way to make this argument: The sets $[i]$ and S have the same number of elements (indeed, both are i-element sets). Hence, there exists a bijection $\phi : S \to [i]$. Fix such a ϕ. Now, the S-junctions are in a 1-to-1 correspondence with the $[i]$-junctions (namely, to every S-junction (A, B) corresponds the $[i]$-junction $(\phi(A), \phi(B))$, and conversely, to every $[i]$-junction (A', B') corresponds the S-junction $(\phi^{-1}(A'), \phi^{-1}(B'))$). Hence, the number of all S-junctions equals the number of $[i]$-junctions. Since the latter number is e_i, this shows that the former number is also e_i. This proves (264).

[126] This follows from (236).

[127] Again, this follows from (236).

[128] *Proof.* Clearly, i cannot be smaller than 0. But i also cannot be larger than $a + b$ (since i will have to satisfy $i = |A \cup B| \leq \underbrace{|A|}_{=a} + \underbrace{|B|}_{=b} = a + b$). Thus, i must be an integer between 0 and $a + b$ (inclusive).

2. We choose a subset S of $[x]$, which will serve as the union $A \cup B$. This subset S must be an i-element subset of $[x]$ (because we will have $\left|\underbrace{S}_{=A\cup B}\right| = |A \cup B| = i$). Thus, there are $\dbinom{x}{i}$ ways to choose it (since we need to choose an i-element subset of $[x]$).

3. Now, it remains to choose the pair (A, B) itself. This pair must be a pair of subsets of $[x]$ satisfying $|A| = a$, $|B| = b$, $A \cup B = S$ and $|A \cup B| = i$. We can forget about the $|A \cup B| = i$ condition, since it automatically follows from $A \cup B = S$ (because $|S| = i$). So we need to choose a pair (A, B) of subsets of $[x]$ satisfying $|A| = a$, $|B| = b$ and $A \cup B = S$. In other words, we need to choose a pair (A, B) of subsets of S satisfying $|A| = a$, $|B| = b$ and $A \cup B = S$ [129]. In other words, we need to choose an S-junction (since this is how an S-junction was defined). This can be done in exactly e_i ways (according to (264)).

Thus, in total, there are $\sum\limits_{i=0}^{a+b} \dbinom{x}{i} e_i$ ways to perform this procedure. Hence, the total number of all pairs (A, B) consisting of an a-element subset A of $[x]$ and a b-element subset B of $[x]$ equals $\sum\limits_{i=0}^{a+b} \dbinom{x}{i} e_i$. But earlier, we have shown that this number is $\dbinom{x}{a}\dbinom{x}{b}$. Comparing these two results, we conclude that $\dbinom{x}{a}\dbinom{x}{b} = \sum\limits_{i=0}^{a+b} \dbinom{x}{i} e_i = \sum\limits_{i=0}^{a+b} e_i \dbinom{x}{i}$. Thus, (265) is proven.]

Now, we define two polynomials P and Q in the indeterminate X with rational coefficients by setting

$$P = \binom{X}{a}\binom{X}{b}; \qquad Q = \sum_{i=0}^{a+b} e_i \binom{X}{i}.$$

The equality (265) (which we have proven) states that $P(x) = Q(x)$ for all $x \in \mathbb{N}$. Thus, Lemma 3.28 **(b)** yields that $P = Q$. Recalling how P and Q are defined, we see that this rewrites as $\dbinom{X}{a}\dbinom{X}{b} = \sum\limits_{i=0}^{a+b} e_i \dbinom{X}{i}$. This proves Proposition 3.35. □

Our second proof of Proposition 3.35 is algebraic, and is based on a suggestion of math.stackexchange user tcamps in a comment on question #1342384. It proceeds by way of the following, more explicit result:

[129] Here, we have replaced "subsets of $[x]$" by "subsets of S", because the condition $A \cup B = S$ forces A and B to be subsets of S.

Proposition 3.36. Let a and b be two nonnegative integers. Then,

$$\binom{X}{a}\binom{X}{b} = \sum_{i=a}^{a+b} \binom{i}{a}\binom{a}{a+b-i}\binom{X}{i}.$$

Let us also state the analogue of this proposition in which the indeterminate X is replaced by a rational number m:

Proposition 3.37. Let a and b be two nonnegative integers. Let $m \in \mathbb{Q}$. Then,

$$\binom{m}{a}\binom{m}{b} = \sum_{i=a}^{a+b} \binom{i}{a}\binom{a}{a+b-i}\binom{m}{i}.$$

Proof of Proposition 3.37. Theorem 3.29 (applied to b, $m-a$ and a instead of n, x and y) yields

$$\binom{(m-a)+a}{b} = \sum_{k=0}^{b} \binom{m-a}{k}\binom{a}{b-k}.$$

Since $(m-a)+a = m$, this rewrites as

$$\binom{m}{b} = \sum_{k=0}^{b} \binom{m-a}{k}\binom{a}{b-k} = \sum_{i=a}^{a+b} \binom{m-a}{i-a} \underbrace{\binom{a}{b-(i-a)}}_{\substack{=\binom{a}{a+b-i} \\ \text{(since } b-(i-a)=a+b-i\text{)}}}$$

$$\text{(here, we substituted } i-a \text{ for } k \text{ in the sum)}$$

$$= \sum_{i=a}^{a+b} \binom{m-a}{i-a}\binom{a}{a+b-i}.$$

Multiplying both sides of this identity with $\binom{m}{a}$, we obtain

$$\binom{m}{a}\binom{m}{b} = \binom{m}{a}\sum_{i=a}^{a+b} \binom{m-a}{i-a}\binom{a}{a+b-i} = \sum_{i=a}^{a+b} \underbrace{\binom{m}{a}\binom{m-a}{i-a}}_{\substack{=\binom{m}{i}\binom{i}{a} \\ \text{(by Proposition 3.23)}}} \binom{a}{a+b-i}$$

$$= \sum_{i=a}^{a+b} \binom{m}{i}\binom{i}{a}\binom{a}{a+b-i} = \sum_{i=a}^{a+b} \binom{i}{a}\binom{a}{a+b-i}\binom{m}{i}.$$

This proves Proposition 3.37. □

Proof of Proposition 3.36. To obtain a proof of Proposition 3.36, replace every appearance of "m" by "X" in the above proof of Proposition 3.37. (Of course, this requires knowing that Theorem 3.29 holds when x and y are polynomials rather than numbers. But this is true, because both proofs that we gave for Theorem 3.29 still apply in this case.) □

Second proof of Proposition 3.35. Let us define $a+b+1$ nonnegative integers $e_0, e_1, \ldots, e_{a+b}$ by

$$e_i = \begin{cases} \dbinom{i}{a}\dbinom{a}{a+b-i}, & \text{if } i \geq a; \\ 0, & \text{otherwise} \end{cases} \qquad \text{for all } i \in \{0,1,\ldots,a+b\}. \tag{266}$$

Then,

$$\begin{aligned} \sum_{i=0}^{a+b} e_i \binom{X}{i} &= \sum_{i=a}^{a+b} \binom{i}{a}\binom{a}{a+b-i}\binom{X}{i} && \left(\text{by our definition of } e_0, e_1, \ldots, e_{a+b}\right) \\ &= \binom{X}{a}\binom{X}{b} && \left(\text{by Proposition 3.36}\right). \end{aligned}$$

Thus, Proposition 3.35 is proven again. □

Remark 3.38. Comparing our two proofs of Proposition 3.35, it is natural to suspect that the $e_0, e_1, \ldots, e_{a+b}$ defined in the First proof are identical with the $e_0, e_1, \ldots, e_{a+b}$ defined in the Second proof. This actually follows from general principles (namely, from the word "unique" in Proposition 3.27 **(a)**), but there is also a simple combinatorial reason. Namely, let $i \in \{0,1,\ldots,a+b\}$. We shall show that the e_i defined in the First proof equals the e_i defined in the Second proof.

The e_i defined in the First proof is the number of all $[i]$-junctions. An $[i]$-junction is a pair (A,B), where A is an a-element subset of $[i]$ and where B is a b-element subset of $[i]$ such that $A \cup B = [i]$. Here is a way to construct an $[i]$-junction:

- First, we pick the set A. There are $\dbinom{i}{a}$ ways to do this, since A has to be an a-element subset of the i-element set $[i]$.
- Then, we pick the set B. This has to be a b-element subset of the i-element set $[i]$ satisfying $A \cup B = [i]$. The equality $A \cup B = [i]$ means that B has to contain the $i-a$ elements of $[i] \setminus A$; but the remaining $b-(i-a) = a+b-i$ elements of B can be chosen arbitrarily among the a elements of A. Thus, there are $\dbinom{a}{a+b-i}$ ways to choose B (since we have to choose $a+b-i$ elements of B among the a elements of A).

Thus, the number of all $[i]$-junctions is $\dbinom{i}{a}\dbinom{a}{a+b-i}$. This can be rewritten in the form $\begin{cases} \dbinom{i}{a}\dbinom{a}{a+b-i}, & \text{if } i \geq a; \\ 0, & \text{otherwise} \end{cases}$ (because if $i < a$, then $\dbinom{i}{a} = 0$ and thus $\dbinom{i}{a}\dbinom{a}{a+b-i} = 0$). Thus, we have shown that the number of all $[i]$-junctions is $\begin{cases} \dbinom{i}{a}\dbinom{a}{a+b-i}, & \text{if } i \geq a; \\ 0, & \text{otherwise} \end{cases}$. In other words, the e_i defined in the First proof equals the e_i defined in the Second proof.

Here is an assortment of other identities that involve binomial coefficients:

Proposition 3.39. **(a)** Every $x \in \mathbb{Z}$, $y \in \mathbb{Z}$ and $n \in \mathbb{N}$ satisfy $(x+y)^n = \sum\limits_{k=0}^{n} \dbinom{n}{k} x^k y^{n-k}$.

(b) Every $n \in \mathbb{N}$ satisfies $\sum\limits_{k=0}^{n} \dbinom{n}{k} = 2^n$.

(c) Every $n \in \mathbb{N}$ satisfies $\sum\limits_{k=0}^{n} (-1)^k \dbinom{n}{k} = \begin{cases} 1, & \text{if } n = 0; \\ 0, & \text{if } n \neq 0 \end{cases}$.

(d) Every $n \in \mathbb{Z}$, $i \in \mathbb{N}$ and $a \in \mathbb{N}$ satisfying $i \geq a$ satisfy $\dbinom{n}{i}\dbinom{i}{a} = \dbinom{n}{a}\dbinom{n-a}{i-a}$.

(e) Every $n \in \mathbb{N}$ and $m \in \mathbb{Z}$ satisfy $\sum\limits_{i=0}^{n} \dbinom{n}{i}\dbinom{m+i}{n} = \sum\limits_{i=0}^{n} \dbinom{n}{i}\dbinom{m}{i} 2^i$.

(f) Every $a \in \mathbb{N}$, $b \in \mathbb{N}$ and $x \in \mathbb{Z}$ satisfy $\sum\limits_{i=0}^{b} \dbinom{a}{i}\dbinom{b}{i}\dbinom{x+i}{a+b} = \dbinom{x}{a}\dbinom{x}{b}$.

(g) Every $a \in \mathbb{N}$, $b \in \mathbb{N}$ and $x \in \mathbb{Z}$ satisfy $\sum\limits_{i=0}^{b} \dbinom{a}{i}\dbinom{b}{i}\dbinom{a+b+x-i}{a+b} = \dbinom{a+x}{a}\dbinom{b+x}{b}$.

(I have learnt parts **(e)** and **(f)** of Proposition 3.39 from AoPS, but they are fairly classical results. Part **(e)** is equivalent to a claim in [Comtet74, Chapter I, Exercise 21]. Part **(f)** is [Riorda68, §1.4, (10)]. Part **(g)** is a restatement of [Gould10, (6.93)].)

Proof of Proposition 3.39. **(a)** Proposition 3.39 **(a)** is clearly a particular case of Proposition 3.21.

(b) Let $n \in \mathbb{N}$. Applying Proposition 3.39 **(a)** to $x = 1$ and $y = 1$, we obtain

$$(1+1)^n = \sum_{k=0}^{n} \binom{n}{k} \underbrace{1^k}_{=1} \underbrace{1^{n-k}}_{=1} = \sum_{k=0}^{n} \binom{n}{k},$$

thus

$$\sum_{k=0}^{n} \binom{n}{k} = \Bigg(\underbrace{1+1}_{=2}\Bigg)^n = 2^n.$$

This proves Proposition 3.39 **(b)**.

(c) Let $n \in \mathbb{N}$. Applying Proposition 3.39 **(a)** to $x = -1$ and $y = 1$, we obtain

$$(-1+1)^n = \sum_{k=0}^{n} \binom{n}{k} (-1)^k \underbrace{1^{n-k}}_{=1} = \sum_{k=0}^{n} \binom{n}{k} (-1)^k = \sum_{k=0}^{n} (-1)^k \binom{n}{k},$$

thus

$$\sum_{k=0}^{n} (-1)^k \binom{n}{k} = \Bigg(\underbrace{-1+1}_{=0}\Bigg)^n = 0^n = \begin{cases} 1, & \text{if } n = 0; \\ 0, & \text{if } n \neq 0 \end{cases}.$$

This proves Proposition 3.39 **(c)**.

(d) Let $n \in \mathbb{Z}$, $i \in \mathbb{N}$ and $a \in \mathbb{N}$ be such that $i \geq a$. Proposition 3.23 (applied to n instead of m) yields $\binom{n}{i}\binom{i}{a} = \binom{n}{a}\binom{n-a}{i-a}$. This proves Proposition 3.39 **(d)**.

(e) Let $n \in \mathbb{N}$ and $m \in \mathbb{Z}$. Clearly, every $p \in \mathbb{N}$ satisfies

$$\begin{aligned} \sum_{i=0}^{p} \binom{p}{i} &= \sum_{k=0}^{p} \binom{p}{k} \qquad \text{(here, we renamed the summation index } i \text{ as } k\text{)} \\ &= 2^p \end{aligned} \tag{267}$$

(by Proposition 3.39 **(b)**, applied to p instead of n).

Now, let $i \in \{0, 1, \ldots, n\}$. Applying Proposition 3.32 **(a)** to $x = i$ and $y = m$, we obtain

$$\begin{aligned} &\binom{i+m}{n} \\ &= \sum_{k=0}^{n} \binom{i}{k}\binom{m}{n-k} \\ &= \sum_{k=0}^{i} \binom{i}{k}\binom{m}{n-k} + \sum_{k=i+1}^{n} \underbrace{\binom{i}{k}}_{\substack{=0 \\ \text{(by (231), applied to } i \text{ and } k \\ \text{instead of } m \text{ and } n \text{ (since } i<k \\ \text{(because } k \geq i+1 > i\text{)))}}} \binom{m}{n-k} \qquad (\text{since } 0 \leq i \leq n) \\ &= \sum_{k=0}^{i} \binom{i}{k}\binom{m}{n-k} + \underbrace{\sum_{k=i+1}^{n} 0 \binom{m}{n-k}}_{=0} = \sum_{k=0}^{i} \binom{i}{k}\binom{m}{n-k}. \end{aligned} \tag{268}$$

Now, let us forget that we fixed i. We thus have proven (268) for every $i \in \{0,1,\ldots,n\}$. Now,

$$\begin{aligned}
&\sum_{i=0}^{n} \binom{n}{i} \underbrace{\binom{m+i}{n}}_{\substack{=\binom{i+m}{n}=\sum_{k=0}^{i}\binom{i}{k}\binom{m}{n-k} \\ \text{(by (268))}}} \\
&= \sum_{i=0}^{n} \binom{n}{i} \left(\sum_{k=0}^{i} \binom{i}{k} \binom{m}{n-k} \right) = \underbrace{\sum_{i=0}^{n} \sum_{k=0}^{i}}_{=\sum_{k=0}^{n}\sum_{i=k}^{n}} \binom{n}{i} \binom{i}{k} \binom{m}{n-k} \\
&= \sum_{k=0}^{n} \sum_{i=k}^{n} \underbrace{\binom{n}{i}\binom{i}{k}}_{\substack{=\binom{n}{k}\binom{n-k}{i-k} \\ \text{(by Proposition 3.39 \textbf{(d)},} \\ \text{applied to } a=k \text{ (since } i\geq k))}} \binom{m}{n-k} \\
&= \sum_{k=0}^{n} \sum_{i=k}^{n} \binom{n}{k} \binom{n-k}{i-k} \binom{m}{n-k} = \sum_{k=0}^{n} \binom{n}{k} \binom{m}{n-k} \sum_{i=k}^{n} \binom{n-k}{i-k} \\
&= \sum_{k=0}^{n} \underbrace{\binom{n}{k}}_{\substack{=\binom{n}{n-k} \\ \text{(by (232), applied to } n \text{ and } k \\ \text{instead of } m \text{ and } n \text{ (since } n\geq k))}} \binom{m}{n-k} \sum_{i=0}^{n-k} \binom{n-k}{i} \\
&\qquad \text{(here, we have substituted } i \text{ for } i-k \text{ in the second sum)} \\
&= \sum_{k=0}^{n} \binom{n}{n-k} \binom{m}{n-k} \sum_{i=0}^{n-k} \binom{n-k}{i} = \sum_{k=0}^{n} \binom{n}{k} \binom{m}{k} \underbrace{\sum_{i=0}^{k} \binom{k}{i}}_{\substack{=2^k \\ \text{(by (267), applied to } p=k)}} \\
&\qquad \text{(here, we have substituted } k \text{ for } n-k \text{ in the first sum)} \\
&= \sum_{k=0}^{n} \binom{n}{k} \binom{m}{k} 2^k = \sum_{i=0}^{n} \binom{n}{i} \binom{m}{i} 2^i
\end{aligned}$$

(here, we have renamed the summation index k as i). This proves Proposition 3.39 **(e)**.

(f) Let $a \in \mathbb{N}$, $b \in \mathbb{N}$ and $x \in \mathbb{Z}$. Proposition 3.37 (applied to $m = x$) yields

$$\binom{x}{a}\binom{x}{b} = \sum_{i=a}^{a+b} \binom{i}{a} \binom{a}{a+b-i} \binom{x}{i}. \tag{269}$$

Clearly,

$$\begin{aligned}
&\sum_{i=0}^{a+b} \binom{a}{i} \binom{b}{i} \binom{x+i}{a+b} \\
&= \sum_{i=0}^{b} \binom{a}{i} \binom{b}{i} \binom{x+i}{a+b} + \sum_{i=b+1}^{a+b} \binom{a}{i} \underbrace{\binom{b}{i}}_{\substack{=0 \\ \text{(by (231), applied to } b \text{ and } i \\ \text{instead of } m \text{ and } n \text{ (since } b<i \\ \text{(because } i\geq b+1>b\text{)))}}} \binom{x+i}{a+b} \\
&\qquad (\text{since } 0 \leq b \leq a+b) \\
&= \sum_{i=0}^{b} \binom{a}{i} \binom{b}{i} \binom{x+i}{a+b} + \underbrace{\sum_{i=b+1}^{a+b} \binom{a}{i} 0 \binom{x+i}{a+b}}_{=0} \\
&= \sum_{i=0}^{b} \binom{a}{i} \binom{b}{i} \binom{x+i}{a+b}. \qquad (270)
\end{aligned}$$

For every $i \in \{0, 1, \ldots, b\}$, we have

$$\binom{x+i}{a+b} = \sum_{k=0}^{a+b} \binom{x}{k} \binom{i}{a+b-k}. \qquad (271)$$

(This follows from Theorem 3.29 (applied to $a+b$ and i instead of n and y).) Hence,

$$\begin{aligned}
&\sum_{i=0}^{a+b} \binom{a}{i} \binom{b}{i} \underbrace{\binom{x+i}{a+b}}_{\substack{=\sum_{k=0}^{a+b} \binom{x}{k} \binom{i}{a+b-k} \\ \text{(by (271))}}} \\
&= \sum_{i=0}^{a+b} \binom{a}{i} \binom{b}{i} \sum_{k=0}^{a+b} \binom{x}{k} \binom{i}{a+b-k} \\
&= \underbrace{\sum_{i=0}^{a+b} \sum_{k=0}^{a+b}}_{=\sum_{k=0}^{a+b} \sum_{i=0}^{a+b}} \binom{a}{i} \underbrace{\binom{b}{i} \binom{x}{k} \binom{i}{a+b-k}}_{=\binom{i}{a+b-k} \binom{b}{i} \binom{x}{k}} \\
&= \sum_{k=0}^{a+b} \sum_{i=0}^{a+b} \binom{a}{i} \binom{i}{a+b-k} \binom{b}{i} \binom{x}{k} = \sum_{k=0}^{a+b} \binom{x}{k} \sum_{i=0}^{a+b} \binom{a}{i} \binom{i}{a+b-k} \binom{b}{i}.
\end{aligned}$$

Compared with (270), this yields

$$\begin{aligned}&\sum_{i=0}^{b}\binom{a}{i}\binom{b}{i}\binom{x+i}{a+b}\\&=\sum_{k=0}^{a+b}\binom{x}{k}\sum_{i=0}^{a+b}\binom{a}{i}\binom{i}{a+b-k}\binom{b}{i}.\end{aligned} \tag{272}$$

However, for every $k\in\{0,1,\ldots,a+b\}$, we have

$$\sum_{i=0}^{a+b}\binom{a}{i}\binom{i}{a+b-k}\binom{b}{i}=\binom{a}{a+b-k}\sum_{j=0}^{k}\binom{k-b}{k-j}\binom{b}{a+b-j}. \tag{273}$$

[*Proof of (273):* Let $k\in\{0,1,\ldots,a+b\}$. Then, $a+b-k\in\{0,1,\ldots,a+b\}$, so that

$0 \leq a+b-k \leq a+b$. Now,

$$\begin{aligned}
&\sum_{i=0}^{a+b} \binom{a}{i} \binom{i}{a+b-k} \binom{b}{i} \\
&= \sum_{i=0}^{(a+b-k)-1} \binom{a}{i} \underbrace{\binom{i}{a+b-k}}_{\substack{=0 \\ \text{(by (231), applied to } i \text{ and} \\ a+b-k \text{ instead of } m \text{ and } n \\ \text{(since } i \leq (a+b-k)-1 < a+b-k\text{))}}} \binom{b}{i} + \sum_{i=a+b-k}^{a+b} \binom{a}{i} \binom{i}{a+b-k} \binom{b}{i} \\
&\qquad \left(\text{since } 0 \leq a+b-k \leq a+b\right) \\
&= \underbrace{\sum_{i=0}^{(a+b-k)-1} \binom{a}{i} 0 \binom{b}{i}}_{=0} + \sum_{i=a+b-k}^{a+b} \binom{a}{i} \binom{i}{a+b-k} \binom{b}{i} \\
&= \sum_{i=a+b-k}^{a+b} \underbrace{\binom{a}{i} \binom{i}{a+b-k}}_{\substack{= \binom{a}{a+b-k} \binom{a-(a+b-k)}{i-(a+b-k)} \\ \text{(by Proposition 3.39 } \textbf{(d)}\text{, applied to} \\ a \text{ and } a+b-k \text{ instead of } n \text{ and } a \text{ (since } i \geq a+b-k\text{))}}} \binom{b}{i} \\
&= \sum_{i=a+b-k}^{a+b} \binom{a}{a+b-k} \binom{a-(a+b-k)}{i-(a+b-k)} \binom{b}{i} \\
&= \sum_{j=0}^{k} \binom{a}{a+b-k} \underbrace{\binom{a-(a+b-k)}{(a+b-j)-(a+b-k)}}_{\substack{= \binom{k-b}{k-j} \\ \text{(since } a-(a+b-k)=k-b \text{ and} \\ (a+b-j)-(a+b-k)=k-j\text{)}}} \binom{b}{a+b-j} \\
&\qquad \left(\text{here, we have substituted } a+b-j \text{ for } i \text{ in the sum}\right) \\
&= \sum_{j=0}^{k} \binom{a}{a+b-k} \binom{k-b}{k-j} \binom{b}{a+b-j} = \binom{a}{a+b-k} \sum_{j=0}^{k} \binom{k-b}{k-j} \binom{b}{a+b-j},
\end{aligned}$$

and this proves (273).]

Now, (272) becomes

$$\begin{aligned}
&\sum_{i=0}^{b} \binom{a}{i} \binom{b}{i} \binom{x+i}{a+b} \\
&= \sum_{k=0}^{a+b} \binom{x}{k} \underbrace{\sum_{i=0}^{a+b} \binom{a}{i} \binom{i}{a+b-k} \binom{b}{i}}_{\substack{=\binom{a}{a+b-k} \sum_{j=0}^{k} \binom{k-b}{k-j} \binom{b}{a+b-j} \\ \text{(by (273))}}} \\
&= \sum_{k=0}^{a+b} \binom{x}{k} \binom{a}{a+b-k} \sum_{j=0}^{k} \binom{k-b}{k-j} \binom{b}{a+b-j} \\
&= \sum_{i=0}^{a+b} \binom{x}{i} \binom{a}{a+b-i} \sum_{j=0}^{i} \binom{i-b}{i-j} \binom{b}{a+b-j} \qquad (274)
\end{aligned}$$

(here, we renamed the summation index k as i in the first sum).

Furthermore, every $i \in \{0, 1, \ldots, a+b\}$ satisfies

$$\sum_{j=0}^{i} \binom{i-b}{i-j} \binom{b}{a+b-j} = \binom{i}{a}. \qquad (275)$$

[*Proof of (275):* Let $i \in \{0, 1, \ldots, a+b\}$. Thus, $0 \leq i \leq a+b$. We have

$$\begin{aligned}
\sum_{j=0}^{i} \binom{i-b}{i-j} \binom{b}{a+b-j} &= \sum_{k=0}^{i} \underbrace{\binom{i-b}{i-(i-k)}}_{\substack{=\binom{i-b}{k} \\ \text{(since } i-(i-k)=k\text{)}}} \underbrace{\binom{b}{a+b-(i-k)}}_{\substack{=\binom{b}{(a+b)+k-i} \\ \text{(since } a+b-(i-k)=(a+b)+k-i\text{)}}} \\
&\qquad \text{(here, we have substituted } i-k \text{ for } j \text{ in the sum)} \\
&= \sum_{k=0}^{i} \binom{i-b}{k} \binom{b}{(a+b)+k-i}. \qquad (276)
\end{aligned}$$

On the other hand, we have $b \in \mathbb{N}$, $(i-b)+b = i \geq 0$ and $a \geq i-b$ (since $a+b \geq i$). Therefore, we can apply Proposition 3.32 **(g)** to $i-b$, b and a instead of

x, y and n. As a result, we obtain

$$\begin{aligned}\binom{(i-b)+b}{a} &= \sum_{k=0}^{(i-b)+b} \binom{i-b}{k} \underbrace{\binom{b}{a+k-(i-b)}}_{\substack{=\binom{b}{(a+b)+k-i} \\ \text{(since } a+k-(i-b)=(a+b)+k-i\text{)}}} \\ &= \sum_{k=0}^{(i-b)+b} \binom{i-b}{k} \binom{b}{(a+b)+k-i}.\end{aligned}$$

Since $(i-b)+b=i$, this rewrites as

$$\binom{i}{a} = \sum_{k=0}^{i} \binom{i-b}{k} \binom{b}{(a+b)+k-i}.$$

Compared with (276), this yields

$$\sum_{j=0}^{i} \binom{i-b}{i-j} \binom{b}{a+b-j} = \binom{i}{a}.$$

This proves (275).]

Hence, (274) becomes

$$\begin{aligned}
&\sum_{i=0}^{b} \binom{a}{i} \binom{b}{i} \binom{x+i}{a+b} \\
&= \sum_{i=0}^{a+b} \binom{x}{i} \binom{a}{a+b-i} \underbrace{\sum_{j=0}^{i} \binom{i-b}{i-j} \binom{b}{a+b-j}}_{\substack{=\binom{i}{a} \\ \text{(by (275))}}} \\
&= \sum_{i=0}^{a+b} \underbrace{\binom{x}{i} \binom{a}{a+b-i} \binom{i}{a}}_{=\binom{i}{a} \binom{a}{a+b-i} \binom{x}{i}} = \sum_{i=0}^{a+b} \binom{i}{a} \binom{a}{a+b-i} \binom{x}{i} \\
&= \sum_{i=0}^{a-1} \underbrace{\binom{i}{a}}_{\substack{=0 \\ \text{(by (231), applied to } i \text{ and } a \\ \text{instead of } m \text{ and } n \text{ (since } i<a\text{))}}} \binom{a}{a+b-i} \binom{x}{i} + \sum_{i=a}^{a+b} \binom{i}{a} \binom{a}{a+b-i} \binom{x}{i} \\
&\qquad \left(\text{since } 0 \leq a \leq a+b\right) \\
&= \underbrace{\sum_{i=0}^{a-1} 0 \binom{a}{a+b-i} \binom{x}{i}}_{=0} + \sum_{i=a}^{a+b} \binom{i}{a} \binom{a}{a+b-i} \binom{x}{i} \\
&= \sum_{i=a}^{a+b} \binom{i}{a} \binom{a}{a+b-i} \binom{x}{i} = \binom{x}{a} \binom{x}{b} \qquad \left(\text{by (269)}\right).
\end{aligned}$$

This proves Proposition 3.39 **(f)**.

(g) Let $a \in \mathbb{N}$, $b \in \mathbb{N}$ and $x \in \mathbb{Z}$. From (237) (applied to $m = -x-1$ and $n = a$), we obtain $\binom{-x-1}{a} = (-1)^a \binom{a-(-x-1)-1}{a} = (-1)^a \binom{a+x}{a}$ (since $a-(-x-1)-1 = a+x$). The same argument (applied to b instead of a) shows that $\binom{-x-1}{b} = (-1)^b \binom{b+x}{b}$.

Now, Proposition 3.39 **(f)** (applied to $-x-1$ instead of x) shows that

$$\begin{aligned}\sum_{i=0}^{b}\binom{a}{i}\binom{b}{i}\binom{(-x-1)+i}{a+b} &= \underbrace{\binom{-x-1}{a}}_{=(-1)^a\binom{a+x}{a}}\underbrace{\binom{-x-1}{b}}_{=(-1)^b\binom{b+x}{b}}\\ &= (-1)^a\binom{a+x}{a}(-1)^b\binom{b+x}{b}\\ &= \underbrace{(-1)^a(-1)^b}_{=(-1)^{a+b}}\binom{a+x}{a}\binom{b+x}{b}\\ &= (-1)^{a+b}\binom{a+x}{a}\binom{b+x}{b}. \qquad (277)\end{aligned}$$

But every $i \in \{0, 1, \ldots, b\}$ satisfies

$$\begin{aligned}\binom{(-x-1)+i}{a+b} &= (-1)^{a+b}\binom{a+b-((-x-1)+i)-1}{a+b}\\ &\qquad \left(\text{by (237), applied to } m=(-x-1)+i \text{ and } n=a+b\right)\\ &= (-1)^{a+b}\binom{a+b+x-i}{a+b}\\ &\qquad \left(\text{since } a+b-((-x-1)+i)-1 = a+b+x-i\right).\end{aligned}$$

Hence,

$$\begin{aligned}&\sum_{i=0}^{b}\binom{a}{i}\binom{b}{i}\underbrace{\binom{(-x-1)+i}{a+b}}_{=(-1)^{a+b}\binom{a+b+x-i}{a+b}}\\ &= \sum_{i=0}^{b}\binom{a}{i}\binom{b}{i}(-1)^{a+b}\binom{a+b+x-i}{a+b} = (-1)^{a+b}\sum_{i=0}^{b}\binom{a}{i}\binom{b}{i}\binom{a+b+x-i}{a+b}.\end{aligned}$$

Comparing this with (277), we obtain

$$(-1)^{a+b}\sum_{i=0}^{b}\binom{a}{i}\binom{b}{i}\binom{a+b+x-i}{a+b} = (-1)^{a+b}\binom{a+x}{a}\binom{b+x}{b}.$$

We can cancel $(-1)^{a+b}$ from this equality (since $(-1)^{a+b} \neq 0$), and thus obtain $\sum\limits_{i=0}^{b}\binom{a}{i}\binom{b}{i}\binom{a+b+x-i}{a+b} = \binom{a+x}{a}\binom{b+x}{b}$. This proves Proposition 3.39 **(g)**.

□

Many more examples of equalities with binomial coefficients, as well as advanced tactics for proving such equalities, can be found in [GrKnPa94, Chapter 5].

Exercise 3.10. Let $n \in \mathbb{Q}$, $a \in \mathbb{N}$ and $b \in \mathbb{N}$.

(a) Prove that every integer $j \geq a$ satisfies

$$\binom{n}{j}\binom{j}{a}\binom{n-j}{b} = \binom{n}{a}\binom{n-a}{b}\binom{n-a-b}{j-a}.$$

(b) Compute the sum $\sum\limits_{j=a}^{n} \binom{n}{j}\binom{j}{a}\binom{n-j}{b}$ for every integer $n \geq a$. (The result should contain no summation signs.)

3.5. The principle of inclusion and exclusion

We shall next discuss the *principle of inclusion and exclusion*, and some of its generalizations. This is a crucial result in combinatorics, which can help both in answering enumerative questions (i.e., questions of the form "how many objects of a given kind satisfy a certain set of properties") and in proving combinatorial identities (such as, to give a simple example, Proposition 3.39 **(c)**, but also various deeper results). We shall not dwell on the applications of this principle; the reader can easily find them in textbooks on enumerative combinatorics (such as [Aigner07, §5.1] or [Galvin17, §16] or [Loehr11, Chapter 4] or [Comtet74, Chapter IV] or [LeLeMe16, §15.9] or [AndFen04, Chapter 6]). We will, however, prove the principle and a few of its generalizations.

The principle itself (in one of its most basic forms) answers the following simple question: Given n finite sets $A_1, A_2, \ldots, A_n$, how do we compute the size of their union $A_1 \cup A_2 \cup \cdots \cup A_n$ if we know the sizes of all of their intersections (not just the intersection $A_1 \cap A_2 \cap \cdots \cap A_n$, but also the intersections of some of the sets only)? Let us first answer this question for specific small values of n:

- For $n = 1$, we have the tautological equality
$$|A_1| = |A_1|. \tag{278}$$
(Only a true mathematician would begin a study with such a statement.)

- For $n = 2$, we have the known formula
$$|A_1 \cup A_2| = |A_1| + |A_2| - |A_1 \cap A_2|. \tag{279}$$
Notice that $|A_1|$ and $|A_2|$ are sizes of intersections of some of the sets A_1, A_2: Namely, A_1 is the intersection of the single set A_1, while A_2 is the intersection of the single set A_2.

- For $n = 3$, we have
$$\begin{aligned} |A_1 \cup A_2 \cup A_3| &= |A_1| + |A_2| + |A_3| - |A_1 \cap A_2| - |A_1 \cap A_3| - |A_2 \cap A_3| \\ &\qquad + |A_1 \cap A_2 \cap A_3| . \end{aligned} \tag{280}$$
This is not as well-known as (279), but can be easily derived by applying (279) twice. (In fact, first apply (279) to $A_1 \cup A_2$ and A_3 instead of A_1 and A_2; then, apply (279) directly to rewrite $|A_1 \cup A_2|$.)

- For $n = 4$, we have
$$\begin{aligned} &|A_1 \cup A_2 \cup A_3 \cup A_4| \\ &= |A_1| + |A_2| + |A_3| + |A_4| \\ &\qquad - |A_1 \cap A_2| - |A_1 \cap A_3| - |A_1 \cap A_4| - |A_2 \cap A_3| - |A_2 \cap A_4| - |A_3 \cap A_4| \\ &\qquad + |A_1 \cap A_2 \cap A_3| + |A_1 \cap A_2 \cap A_4| + |A_1 \cap A_3 \cap A_4| + |A_2 \cap A_3 \cap A_4| \\ &\qquad - |A_1 \cap A_2 \cap A_3 \cap A_4| . \end{aligned} \tag{281}$$
Again, this can be derived by applying (279) many times.

The four equalities (278), (279), (280) and (281) all follow the same pattern: On the left hand side is the size $|A_1 \cup A_2 \cup \cdots \cup A_n|$ of the union $A_1 \cup A_2 \cup \cdots \cup A_n$ of all the n sets $A_1, A_2, \ldots, A_n$. On the right hand side is an "alternating sum" (i.e., a sum, but with minus signs in front of some of its addends), whose addends are the sizes of the intersections of all possible choices of **some** of the n sets $A_1, A_2, \ldots, A_n$ (except for the choice where none of the n sets are chosen; this does not have a well-defined intersection). Notice that there are $2^n - 1$ such choices, so the right hand side is an "alternating sum" of $2^n - 1$ addends. In other words, each addend on the right hand side has the form $\left|A_{i_1} \cap A_{i_2} \cap \cdots \cap A_{i_m}\right|$ for some m-tuple $(i_1, i_2, \ldots, i_m)$ of integers between 1 and n (inclusive) such that $m \geq 1$ and $i_1 < i_2 < \cdots < i_m$. The sign in front of this addend is a plus sign if m is odd, and is a minus sign if m is even. Thus, we can replace this sign by a factor of $(-1)^{m-1}$.

We can try and generalize the pattern as follows:
$$\begin{aligned} &|A_1 \cup A_2 \cup \cdots \cup A_n| \\ &= \sum_{m=1}^{n} \sum_{1 \leq i_1 < i_2 < \cdots < i_m \leq n} (-1)^{m-1} \left|A_{i_1} \cap A_{i_2} \cap \cdots \cap A_{i_m}\right| . \end{aligned} \tag{282}$$
Here, the summation sign "$\sum\limits_{1 \leq i_1 < i_2 < \cdots < i_m \leq n}$" on the right hand side is an abbreviation for $\sum\limits_{\substack{(i_1, i_2, \ldots, i_m) \in \{1,2,\ldots,n\}^m; \\ i_1 < i_2 < \cdots < i_m}}$.

The equality (282) is indeed correct; it is one of several (equivalent) versions of the principle of inclusion and exclusion. For example, it appears in [Loehr11, §4.7]. We shall, however, state it differently, for the sake of better generalizability. First, we will index the intersections of **some** of the n sets $A_1, A_2, \ldots, A_n$ not by m-tuples $(i_1, i_2, \ldots, i_m) \in \{1, 2, \ldots, n\}^m$ satisfying $i_1 < i_2 < \cdots < i_m$, but rather by nonempty subsets of $\{1, 2, \ldots, n\}$. Second, our sets $A_1, A_2, \ldots, A_n$ will be labelled not by the numbers $1, 2, \ldots, n$, but rather by elements of a finite set G. This will result in a more abstract, but also more flexible version of (282).

First, we introduce some notations:

Definition 3.40. Let I be a set. For each $i \in I$, we let A_i be a set.

(a) Then, $\bigcup\limits_{i \in I} A_i$ denotes the union of all the sets A_i for $i \in I$. This union is defined by

$$\bigcup_{i \in I} A_i = \{x \mid \text{there exists an } i \in I \text{ such that } x \in A_i\}.$$

For example, if $I = \{i_1, i_2, \ldots, i_k\}$ is a finite set, then

$$\bigcup_{i \in I} A_i = A_{i_1} \cup A_{i_2} \cup \cdots \cup A_{i_k}.$$

Notice that $A_j \subseteq \bigcup\limits_{i \in I} A_i$ for each $j \in I$. If I is finite, and if each of the sets A_i is finite, then their union $\bigcup\limits_{i \in I} A_i$ is also finite.

Note that $\bigcup\limits_{i \in \varnothing} A_i = \varnothing$.

(b) Assume that I is nonempty. Then, $\bigcap\limits_{i \in I} A_i$ denotes the intersection of all the sets A_i for $i \in I$. This intersection is defined by

$$\bigcap_{i \in I} A_i = \{x \mid x \in A_i \text{ for all } i \in I\}.$$

For example, if $I = \{i_1, i_2, \ldots, i_k\}$ is a finite set, then

$$\bigcap_{i \in I} A_i = A_{i_1} \cap A_{i_2} \cap \cdots \cap A_{i_k}.$$

Notice that $\bigcap\limits_{i \in I} A_i \subseteq A_j$ for each $j \in I$. If each of the sets A_i is finite, then their intersection $\bigcap\limits_{i \in I} A_i$ is also finite.

Caution: The intersection $\bigcap\limits_{i \in I} A_i$ is not defined when I is empty, because this intersection would have to contain every object in the universe (which is impossible for a set).

Definition 3.41. If G is any finite set, then the sign $\sum\limits_{I \subseteq G}$ shall mean $\sum\limits_{I \in \mathcal{P}(G)}$, where

$\mathcal{P}(G)$ denotes the powerset of G. For example,

$$\sum_{I\subseteq\{7,8\}}\prod_{i\in I} i = \sum_{I\in\mathcal{P}(\{7,8\})}\prod_{i\in I} i = \underbrace{\prod_{i\in\varnothing} i}_{=(\text{empty product})=1} + \underbrace{\prod_{i\in\{7\}} i}_{=7} + \underbrace{\prod_{i\in\{8\}} i}_{=8} + \underbrace{\prod_{i\in\{7,8\}} i}_{=7\cdot 8}$$
$$\left(\text{since the subsets of } \{7,8\} \text{ are } \varnothing, \{7\}, \{8\}, \{7,8\}\right)$$
$$= 1+7+8+7\cdot 8.$$

We are now ready to state one of the forms of the principle:

Theorem 3.42. Let G be a finite set. For each $i \in G$, let A_i be a finite set. Then,

$$\left|\bigcup_{i\in G} A_i\right| = \sum_{\substack{I\subseteq G;\\ I\neq\varnothing}} (-1)^{|I|-1}\left|\bigcap_{i\in I} A_i\right|.$$

If $G = \{1,2,\ldots,n\}$ for some $n \in \mathbb{N}$, then the formula in Theorem 3.42 is a restatement of (282) (because the nonempty subsets of $\{1,2,\ldots,n\}$ are in 1-to-1 correspondence with the m-tuples $(i_1, i_2, \ldots, i_m) \in \{1,2,\ldots,n\}^m$ satisfying $i_1 < i_2 < \cdots < i_m$ and $m \in \{1,2,\ldots,n\}$).

A statement equivalent to Theorem 3.42 is the following:

Theorem 3.43. Let S be a finite set. Let G be a finite set. For each $i \in G$, let A_i be a subset of S. We define the intersection $\bigcap_{i\in\varnothing} A_i$ (which would otherwise be undefined, since $\varnothing$ is the empty set) to mean the set S. (Thus, $\bigcap_{i\in I} A_i$ is defined for any subset I of G, not just for nonempty subsets I.) Then,

$$\left|S\setminus\left(\bigcup_{i\in G} A_i\right)\right| = \sum_{I\subseteq G} (-1)^{|I|}\left|\bigcap_{i\in I} A_i\right|.$$

Theorem 3.42 and Theorem 3.43 are commonly known as the *principle of inclusion and exclusion*, or as the *Sylvester sieve formula*. They are not hard to prove (see, e.g., [Galvin17, §16] for two proofs, and [Sagan19, proof of (2.5)] for yet another). Rather than proving them directly, we shall however generalize them and prove the generalization, from which they are easily obtained as particular cases.

We generalize these theorems in two steps. First, we observe that the $S\setminus\left(\bigcup_{i\in G} A_i\right)$ in Theorem 3.43 is simply the set of all elements of S that belong to none of the subsets A_i (for $i \in G$). In other words,

$$S\setminus\left(\bigcup_{i\in G} A_i\right) = \{s\in S \mid \text{the number of } i\in G \text{ satisfying } s\in A_i \text{ equals } 0\}.$$

We can replace the "0" here by any number k, and ask for the size of the resulting set. The answer is given by the following result of Charles Jordan (see [Comtet74, §4.8, Theorem A] and [DanRot78] for fairly complicated proofs):

Theorem 3.44. Let S be a finite set. Let G be a finite set. For each $i \in G$, let A_i be a subset of S. We define the intersection $\bigcap\limits_{i \in \varnothing} A_i$ (which would otherwise be undefined, since $\varnothing$ is the empty set) to mean the set S. (Thus, $\bigcap\limits_{i \in I} A_i$ is defined for any subset I of G, not just for nonempty subsets I.)

Let $k \in \mathbb{N}$. Let

$$S_k = \{ s \in S \mid \text{the number of } i \in G \text{ satisfying } s \in A_i \text{ equals } k \} .$$

(In other words, S_k is the set of all elements of S that belong to exactly k of the subsets A_i.) Then,

$$|S_k| = \sum_{I \subseteq G} (-1)^{|I|-k} \binom{|I|}{k} \left| \bigcap_{i \in I} A_i \right| .$$

A different generalization of Theorem 3.43 (closely related to the *Bonferroni inequalities*, for which see [Galvin17, §17, problem (2)]) explores what happens when the sum on the right hand side of the formula is restricted to only those subsets I of G whose size doesn't surpass a given integer m:

Theorem 3.45. Let S be a finite set. Let G be a finite set. For each $i \in G$, let A_i be a subset of S. We define the intersection $\bigcap\limits_{i \in \varnothing} A_i$ (which would otherwise be undefined, since $\varnothing$ is the empty set) to mean the set S. (Thus, $\bigcap\limits_{i \in I} A_i$ is defined for any subset I of G, not just for nonempty subsets I.)

Let $m \in \mathbb{N}$. For each $s \in S$, let $c(s)$ denote the number of $i \in G$ satisfying $s \in A_i$. Then,

$$(-1)^m \sum_{s \in S} \binom{c(s) - 1}{m} = \sum_{\substack{I \subseteq G; \\ |I| \leq m}} (-1)^{|I|} \left| \bigcap_{i \in I} A_i \right| .$$

Finally, Theorem 3.44 and Theorem 3.45 can be merged into one common general principle:

Theorem 3.46. Let S be a finite set. Let G be a finite set. For each $i \in G$, let A_i be a subset of S. We define the intersection $\bigcap\limits_{i \in \varnothing} A_i$ (which would otherwise be undefined, since $\varnothing$ is the empty set) to mean the set S. (Thus, $\bigcap\limits_{i \in I} A_i$ is defined for any subset I of G, not just for nonempty subsets I.)

Let $k \in \mathbb{N}$ and $m \in \mathbb{N}$ be such that $m \geq k$. For each $s \in S$, let $c(s)$ denote the number of $i \in G$ satisfying $s \in A_i$. Then,

$$(-1)^m \sum_{s \in S} \binom{c(s)}{k} \binom{c(s)-k-1}{m-k} = \sum_{\substack{I \subseteq G; \\ |I| \leq m}} (-1)^{|I|} \binom{|I|}{k} \left| \bigcap_{i \in I} A_i \right| .$$

As we said, we shall first prove Theorem 3.46, and then derive all the preceding theorems in this section from it. The proof of Theorem 3.46 will rely on several ingredients, the first of which is the following simple identity:

Lemma 3.47. Let $n \in \mathbb{N}$ and $k \in \mathbb{N}$. Let $m \in \{k, k+1, k+2, \ldots\}$. Then,

$$\sum_{r=0}^{m} (-1)^r \binom{n}{r} \binom{r}{k} = (-1)^m \binom{n}{k} \binom{n-k-1}{m-k}.$$

Exercise 3.11. Prove Lemma 3.47.

Next, we introduce a simple yet immensely helpful notation that will facilitate our proof:

Definition 3.48. If $\mathcal{A}$ is any logical statement, then we define an integer $[\mathcal{A}] \in \{0,1\}$ by

$$[\mathcal{A}] = \begin{cases} 1, & \text{if } \mathcal{A} \text{ is true;} \\ 0, & \text{if } \mathcal{A} \text{ is false} \end{cases}.$$

For example, $[1+1=2] = 1$ (since $1+1=2$ is true), whereas $[1+1=1] = 0$ (since $1+1=1$ is false).

If $\mathcal{A}$ is any logical statement, then the integer $[\mathcal{A}]$ is known as the *truth value* of $\mathcal{A}$. The notation $[\mathcal{A}]$ is known as the *Iverson bracket notation*.

Clearly, if $\mathcal{A}$ and $\mathcal{B}$ are two equivalent logical statements, then $[\mathcal{A}] = [\mathcal{B}]$. This and a few other useful properties of the Iverson bracket notation are collected in the following exercise:

Exercise 3.12. Prove the following rules for truth values:

(a) If $\mathcal{A}$ and $\mathcal{B}$ are two equivalent logical statements, then $[\mathcal{A}] = [\mathcal{B}]$.

(b) If $\mathcal{A}$ is any logical statement, then $[\text{not } \mathcal{A}] = 1 - [\mathcal{A}]$.

(c) If $\mathcal{A}$ and $\mathcal{B}$ are two logical statements, then $[\mathcal{A} \wedge \mathcal{B}] = [\mathcal{A}][\mathcal{B}]$.

(d) If $\mathcal{A}$ and $\mathcal{B}$ are two logical statements, then $[\mathcal{A} \vee \mathcal{B}] = [\mathcal{A}] + [\mathcal{B}] - [\mathcal{A}][\mathcal{B}]$.

(e) If $\mathcal{A}$, $\mathcal{B}$ and $\mathcal{C}$ are three logical statements, then

$$[\mathcal{A} \vee \mathcal{B} \vee \mathcal{C}] = [\mathcal{A}] + [\mathcal{B}] + [\mathcal{C}] - [\mathcal{A}][\mathcal{B}] - [\mathcal{A}][\mathcal{C}] - [\mathcal{B}][\mathcal{C}] + [\mathcal{A}][\mathcal{B}][\mathcal{C}].$$

The Iverson bracket helps us rewrite the cardinality of a set as a sum:

Lemma 3.49. Let S be a finite set. Let T be a subset of S. Then,

$$|T| = \sum_{s\in S} [s \in T] .$$

Proof of Lemma 3.49. We have

$$\begin{aligned}
\sum_{s\in S}[s\in T] &= \underbrace{\sum_{\substack{s\in S;\\ s\in T}}}_{\substack{=\sum_{s\in T}\\ \text{(since } T \text{ is a}\\ \text{subset of } S\text{)}}} \underbrace{[s\in T]}_{\substack{=1\\ \text{(since } s\in T \text{ is true)}}} + \sum_{\substack{s\in S;\\ s\notin T}} \underbrace{[s\in T]}_{\substack{=0\\ \text{(since } s\in T \text{ is false}\\ \text{(since } s\notin T\text{))}}} \\
&\qquad \left(\begin{array}{c}\text{since each } s\in S \text{ satisfies}\\ \text{either } s\in T \text{ or } s\notin T \text{ (but not both)}\end{array}\right) \\
&= \underbrace{\sum_{s\in T} 1}_{=|T|\cdot 1} + \underbrace{\sum_{\substack{s\in S;\\ s\notin T}} 0}_{=0} = |T|\cdot 1 + 0 = |T| .
\end{aligned}$$

This proves Lemma 3.49. □

We can now state the main precursor to Theorem 3.46:

Lemma 3.50. Let S be a finite set. Let G be a finite set. For each $i \in G$, let A_i be a subset of S. We define the intersection $\bigcap_{i\in\varnothing} A_i$ (which would otherwise be undefined, since $\varnothing$ is the empty set) to mean the set S. (Thus, $\bigcap_{i\in I} A_i$ is defined for any subset I of G, not just for nonempty subsets I.)

Let $k \in \mathbb{N}$ and $m \in \mathbb{N}$ be such that $m \geq k$.

Let $s \in S$. Let $c(s)$ denote the number of $i \in G$ satisfying $s \in A_i$. Then,

$$\sum_{\substack{I\subseteq G;\\ |I|\leq m}} (-1)^{|I|} \binom{|I|}{k} \left[s \in \bigcap_{i\in I} A_i \right] = (-1)^m \binom{c(s)}{k} \binom{c(s)-k-1}{m-k}.$$

Proof of Lemma 3.50. From $m \geq k$ and $m \in \mathbb{N}$, we obtain $m \in \{k, k+1, k+2, \ldots\}$.

Define a subset C of G by

$$C = \{ i \in G \mid s \in A_i \} .$$

Thus,

$$\begin{aligned}
|C| &= |\{ i \in G \mid s \in A_i \}| \\
&= (\text{the number of } i \in G \text{ satisfying } s \in A_i) = c(s)
\end{aligned}$$

(since $c(s)$ was defined as the number of $i \in G$ satisfying $s \in A_i$). In other words, C is a $c(s)$-element set. Hence, for each $r \in \mathbb{N}$, Proposition 3.12 (applied to $c(s)$, r and C instead of m, n and S) shows that

$$\dbinom{c(s)}{r} \text{ is the number of all } r\text{-element subsets of } C. \tag{283}$$

Let I be a subset of G. We have the following equivalence:

$$\left(s \in \bigcap_{i \in I} A_i\right) \iff (s \in A_i \text{ for all } i \in I) \tag{284}$$

[130].

But if $i \in I$, then we have the equivalence

$$(i \in C) \iff (s \in A_i) \tag{285}$$

(since $C = \{i \in G \mid s \in A_i\}$).

Hence, the equivalence (284) becomes

$$\left(s \in \bigcap_{i \in I} A_i\right) \iff \left(\underbrace{s \in A_i}_{\substack{\iff (i \in C) \\ \text{(by (285))}}} \text{ for all } i \in I\right)$$
$$\iff (i \in C \text{ for all } i \in I) \iff (I \subseteq C).$$

In other words, the two statements $\left(s \in \bigcap_{i \in I} A_i\right)$ and $(I \subseteq C)$ are equivalent. Hence, Exercise 3.12 **(a)** (applied to $\mathcal{A} = \left(s \in \bigcap_{i \in I} A_i\right)$ and $\mathcal{B} = (I \subseteq C)$) shows that

$$\left[s \in \bigcap_{i \in I} A_i\right] = [I \subseteq C]. \tag{286}$$

[130] *Proof of (284):* If I is nonempty, then the equivalence (284) follows immediately from the equality

$$\bigcap_{i \in I} A_i = \{x \mid x \in A_i \text{ for all } i \in I\}$$

(which is the definition of the intersection $\bigcap_{i \in I} A_i$). Thus, for the rest of this proof, we WLOG assume that I is not nonempty.

Hence, the set I is empty. In other words, $I = \varnothing$. Hence, $\bigcap_{i \in I} A_i = \bigcap_{i \in \varnothing} A_i = S$. Thus, $s \in S = \bigcap_{i \in I} A_i$. Hence, the statement $\left(s \in \bigcap_{i \in I} A_i\right)$ is true.

Also, there exist no $i \in I$ (since the set I is empty). Hence, the statement $(s \in A_i \text{ for all } i \in I)$ is vacuously true.

Thus, the statements $\left(s \in \bigcap_{i \in I} A_i\right)$ and $(s \in A_i \text{ for all } i \in I)$ are both true, and therefore equivalent. This proves the equivalence (284).

Now, forget that we fixed I. We thus have proven the equality (286) for every subset I of G.

Now,

$$\begin{aligned}
&\sum_{\substack{I\subseteq G;\\ |I|\leq m}} (-1)^{|I|} \binom{|I|}{k} \underbrace{\left[s \in \bigcap_{i\in I} A_i\right]}_{\substack{=[I\subseteq C]\\ \text{(by (286))}}} \\
&= \sum_{\substack{I\subseteq G;\\ |I|\leq m}} (-1)^{|I|} \binom{|I|}{k} [I \subseteq C] \\
&= \underbrace{\sum_{\substack{I\subseteq G;\\ |I|\leq m;\\ I\subseteq C}}}_{\substack{=\sum_{\substack{I\subseteq G;\\ I\subseteq C;\\ |I|\leq m}} = \sum_{\substack{I\subseteq C;\\ |I|\leq m}} \\ \text{(since } C \text{ is a subset of } G)}} (-1)^{|I|} \binom{|I|}{k} \underbrace{[I \subseteq C]}_{\substack{=1\\ \text{(since } I\subseteq C)}} + \sum_{\substack{I\subseteq G;\\ |I|\leq m;\\ \text{not } I\subseteq C}} (-1)^{|I|} \binom{|I|}{k} \underbrace{[I \subseteq C]}_{\substack{=0\\ \text{(since we don't have } I\subseteq C)}} \\
&\qquad \left(\begin{array}{c}\text{since each subset } I \text{ of } G \text{ satisfies either } I \subseteq C\\ \text{or } (\text{not } I \subseteq C) \text{ (but not both)}\end{array}\right) \\
&= \sum_{\substack{I\subseteq C;\\ |I|\leq m}} (-1)^{|I|} \binom{|I|}{k} + \underbrace{\sum_{\substack{I\subseteq G;\\ |I|\leq m;\\ \text{not } I\subseteq C}} (-1)^{|I|} \binom{|I|}{k} 0}_{=0} = \underbrace{\sum_{\substack{I\subseteq C;\\ |I|\leq m}}}_{=\sum_{r=0}^{m} \sum_{\substack{I\subseteq C;\\ |I|=r}}} (-1)^{|I|} \binom{|I|}{k} \\
&= \sum_{r=0}^{m} \sum_{\substack{I\subseteq C;\\ |I|=r}} \underbrace{(-1)^{|I|}}_{\substack{=(-1)^r\\ \text{(since } |I|=r)}} \underbrace{\binom{|I|}{k}}_{\substack{=\binom{r}{k}\\ \text{(since } |I|=r)}}
\end{aligned}$$

$$= \sum_{r=0}^{m} \underbrace{\sum_{\substack{I\subseteq C;\\ |I|=r}} (-1)^r \binom{r}{k}}_{=(\text{the number of all subsets } I \text{ of } C \text{ satisfying } |I|=r)(-1)^r \binom{r}{k}}$$

$$= \sum_{r=0}^{m} \underbrace{(\text{the number of all subsets } I \text{ of } C \text{ satisfying } |I|=r)}_{\substack{=(\text{the number of all } r\text{-element subsets of } C)=\binom{c(s)}{r} \\ (\text{by } (283))}} (-1)^r \binom{r}{k}$$

$$= \sum_{r=0}^{m} \binom{c(s)}{r} (-1)^r \binom{r}{k} = \sum_{r=0}^{m} (-1)^r \binom{c(s)}{r} \binom{r}{k} = (-1)^m \binom{c(s)}{k} \binom{c(s)-k-1}{m-k}$$
$$(\text{by Lemma 3.47 (applied to } n = c(s))).$$

This proves Lemma 3.50. □

We now easily obtain Theorem 3.46:

Proof of Theorem 3.46. For each subset I of G, the intersection $\bigcap_{i\in I} A_i$ is a subset of S [131]. Hence, for each subset I of G, we obtain

$$\left|\bigcap_{i\in I} A_i\right| = \sum_{s\in S} \left[s \in \bigcap_{i\in I} A_i\right] \qquad (287)$$

[131] *Proof.* Let I be a subset of G. We must show that the intersection $\bigcap_{i\in I} A_i$ is a subset of S.

If I is nonempty, then this is clear (because each of the sets A_i is a subset of S). Hence, for the rest of this proof, we can WLOG assume that I is not nonempty. Assume this.

The set I is empty (since I is not nonempty). Hence, $I = \varnothing$. Thus, $\bigcap_{i\in I} A_i = \bigcap_{i\in\varnothing} A_i = S$ (since we defined $\bigcap_{i\in\varnothing} A_i$ to be S). Hence, $\bigcap_{i\in\varnothing} A_i$ is a subset of S. Qed.

(by Lemma 3.49 (applied to $T = \bigcap_{i\in I} A_i$)). Hence,

$$\begin{aligned}
&\sum_{\substack{I\subseteq G;\\ |I|\leq m}} (-1)^{|I|} \dbinom{|I|}{k} \underbrace{\left|\bigcap_{i\in I} A_i\right|}_{\substack{=\sum_{s\in S}\left[s\in\bigcap_{i\in I} A_i\right] \\ \text{(by (287))}}} \\
&= \sum_{\substack{I\subseteq G;\\ |I|\leq m}} (-1)^{|I|} \dbinom{|I|}{k} \sum_{s\in S} \left[s\in \bigcap_{i\in I} A_i\right] = \underbrace{\sum_{\substack{I\subseteq G;\\ |I|\leq m}} \sum_{s\in S}}_{=\sum_{s\in S}\sum_{\substack{I\subseteq G;\\ |I|\leq m}}} (-1)^{|I|} \dbinom{|I|}{k} \left[s\in \bigcap_{i\in I} A_i\right] \\
&= \sum_{s\in S} \underbrace{\sum_{\substack{I\subseteq G;\\ |I|\leq m}} (-1)^{|I|} \dbinom{|I|}{k} \left[s\in \bigcap_{i\in I} A_i\right]}_{\substack{=(-1)^m \binom{c(s)}{k}\binom{c(s)-k-1}{m-k} \\ \text{(by Lemma 3.50)}}} \\
&= \sum_{s\in S} (-1)^m \dbinom{c(s)}{k} \dbinom{c(s)-k-1}{m-k} = (-1)^m \sum_{s\in S} \dbinom{c(s)}{k} \dbinom{c(s)-k-1}{m-k}.
\end{aligned}$$

This proves Theorem 3.46. □

Having proven Theorem 3.46, we can now easily derive the other (less general) versions of the inclusion-exclusion principle:

Exercise 3.13. Prove Theorem 3.42, Theorem 3.43, Theorem 3.44 and Theorem 3.45.

3.6. Additional exercises

This section contains some further exercises. These will not be used in the rest of the notes, and they can be skipped at will[132]. I provide solutions to only a few of them.

Exercise 3.14. Find a different proof of Proposition 3.32 **(f)** that uses a double-counting argument (i.e., counting some combinatorial objects in two different ways, and then concluding that the results are equal).

[**Hint:** How many $(x+y+1)$-element subsets does the set $\{1,2,\ldots,n+1\}$ have? Now, for a given $k\in\{0,1,\ldots,n\}$, how many $(x+y+1)$-element subsets whose $(x+1)$-th smallest element is $k+1$ does the set $\{1,2,\ldots,n+1\}$ have?]

[132] The same, of course, can be said for many of the standard exercises.

Exercise 3.15. Let $n \in \mathbb{N}$ and $k \in \mathbb{N}$ be fixed. Show that the number of all k-tuples $(a_1, a_2, \ldots, a_k) \in \mathbb{N}^k$ satisfying $a_1 + a_2 + \cdots + a_k = n$ equals $\dbinom{n+k-1}{n}$.

Remark 3.51. Exercise 3.15 can be restated in terms of multisets. Namely, let $n \in \mathbb{N}$ and $k \in \mathbb{N}$ be fixed. Also, fix a k-element set K. Then, the number of n-element multisets whose elements all belong to K is $\dbinom{n+k-1}{n}$. Indeed, we can WLOG assume that $K = \{1, 2, \ldots, k\}$ (otherwise, just relabel the elements of K); then, the multisets whose elements all belong to K are in bijection with the k-tuples $(a_1, a_2, \ldots, a_k) \in \mathbb{N}^k$. The bijection sends a multiset M to the k-tuple $(m_1(M), m_2(M), \ldots, m_k(M))$, where each $m_i(M)$ is the multiplicity of the element i in M. The size of a multiset M corresponds to the sum $a_1 + a_2 + \cdots + a_k$ of the entries of the resulting k-tuple; thus, we get a bijection between

- the n-element multisets whose elements all belong to K

and

- the k-tuples $(a_1, a_2, \ldots, a_k) \in \mathbb{N}^k$ satisfying $a_1 + a_2 + \cdots + a_k = n$.

As a consequence, Exercise 3.15 shows that the number of the former multisets is $\dbinom{n+k-1}{n}$.

Similarly, we can reinterpret the classical combinatorial interpretation of $\dbinom{k}{n}$ (as the number of n-element subsets of $\{1, 2, \ldots, k\}$) as follows: The number of all k-tuples $(a_1, a_2, \ldots, a_k) \in \{0, 1\}^k$ satisfying $a_1 + a_2 + \cdots + a_k = n$ equals $\dbinom{k}{n}$.

See [Galvin17, §13] and [Loehr11, §1.11] for more about multisets.

Exercise 3.16. Let n and a be two integers with $n \geq a \geq 1$. Prove that

$$\sum_{k=a}^{n} \frac{(-1)^k}{k} \binom{n-a}{k-a} = \frac{(-1)^a}{a \dbinom{n}{a}}.$$

Exercise 3.17. Let $n \in \mathbb{N}$ and $k \in \mathbb{N}$. Prove that

$$\sum_{u=0}^{k} \binom{n+u-1}{u} \binom{n}{k-2u} = \binom{n+k-1}{k}.$$

Here, $\dbinom{a}{b}$ is defined to be 0 when $b < 0$.

Exercise 3.17 is solved in [Grinbe16a].

Exercise 3.18. Let $N \in \mathbb{N}$. The *binomial transform* of a finite sequence $(f_0, f_1, \ldots, f_N) \in \mathbb{Z}^{N+1}$ is defined to be the sequence $(g_0, g_1, \ldots, g_N)$ defined by

$$g_n = \sum_{i=0}^{n} (-1)^i \binom{n}{i} f_i \qquad \text{for every } n \in \{0, 1, \ldots, N\}.$$

(a) Let $(f_0, f_1, \ldots, f_N) \in \mathbb{Z}^{N+1}$ be a finite sequence of integers. Let $(g_0, g_1, \ldots, g_N)$ be the binomial transform of $(f_0, f_1, \ldots, f_N)$. Show that $(f_0, f_1, \ldots, f_N)$ is, in turn, the binomial transform of $(g_0, g_1, \ldots, g_N)$.

(b) Find the binomial transform of the sequence $(1, 1, \ldots, 1)$.

(c) For any given $a \in \mathbb{N}$, find the binomial transform of the sequence $\left(\binom{0}{a}, \binom{1}{a}, \ldots, \binom{N}{a}\right)$.

(d) For any given $q \in \mathbb{Z}$, find the binomial transform of the sequence $(q^0, q^1, \ldots, q^N)$.

(e) Find the binomial transform of the sequence $(1, 0, 1, 0, 1, 0, \ldots)$ (this ends with 1 if N is even, and with 0 if N is odd).

(f) Let $B : \mathbb{Z}^{N+1} \to \mathbb{Z}^{N+1}$ be the map which sends every sequence $(f_0, f_1, \ldots, f_N) \in \mathbb{Z}^{N+1}$ to its binomial transform $(g_0, g_1, \ldots, g_N) \in \mathbb{Z}^{N+1}$. Thus, part **(a)** of this exercise states that $B^2 = \operatorname{id}$.

On the other hand, let $W : \mathbb{Z}^{N+1} \to \mathbb{Z}^{N+1}$ be the map which sends every sequence $(f_0, f_1, \ldots, f_N) \in \mathbb{Z}^{N+1}$ to $\left((-1)^N f_N, (-1)^N f_{N-1}, \ldots, (-1)^N f_0\right) \in \mathbb{Z}^{N+1}$. It is rather clear that $W^2 = \operatorname{id}$.

Show that, furthermore, $B \circ W \circ B = W \circ B \circ W$ and $(B \circ W)^3 = \operatorname{id}$.

Exercise 3.19. Let $n \in \mathbb{N}$. Prove that

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

[**Hint:** How does the left hand side grow when n is replaced by $n+1$?]

Exercise 3.19 is taken from [AndFen04, Example 3.7].

Exercise 3.20. Let $n \in \mathbb{N}$.

(a) Prove that

$$\sum_{k=0}^{n} \frac{(-1)^k}{\binom{n}{k}} = 2 \cdot \frac{n+1}{n+2} [n \text{ is even}].$$

(Here, we are using the Iverson bracket notation, as in Definition 3.48; thus, $[n \text{ is even}]$ is 1 if n is even and 0 otherwise.)

(b) Prove that

$$\sum_{k=0}^{n} \frac{1}{\binom{n}{k}} = \frac{n+1}{2^{n+1}} \sum_{k=1}^{n+1} \frac{2^k}{k}.$$

[**Hint:** Show that $\frac{1}{\binom{n}{k}} = \left(\frac{1}{\binom{n+1}{k}} + \frac{1}{\binom{n+1}{k+1}} \right) \frac{n+1}{n+2}$ for each $k \in \{0, 1, \ldots, n\}$.]

Exercise 3.20 **(a)** is [KurLis78, (8)]. Exercise 3.20 **(b)** is [KurLis78, (9)] and [AndFen04, Example 3.9] and [AndDos12, Lemma 3.14] and part of [Rocket81, Theorem 1], and also appears with proof in `https://math.stackexchange.com/a/481686/` (where it is used to show that $\lim_{n\to\infty} \sum_{k=0}^{n} \frac{1}{\binom{n}{k}} = 2$).

Exercise 3.21. For any $n \in \mathbb{N}$ and $m \in \mathbb{N}$, define a polynomial $Z_{m,n} \in \mathbb{Z}[X]$ by

$$Z_{m,n} = \sum_{k=0}^{n} (-1)^k \binom{n}{k} \left(X^{n-k} - 1 \right)^m .$$

Show that $Z_{m,n} = Z_{n,m}$ for any $n \in \mathbb{N}$ and $m \in \mathbb{N}$.

Exercise 3.22. Let $n \in \mathbb{N}$. Prove

$$\sum_{k=0}^{n} (-1)^k \binom{X}{k} \binom{X}{n-k} = \begin{cases} (-1)^{n/2} \binom{X}{n/2}, & \text{if } n \text{ is even;} \\ 0, & \text{if } n \text{ is odd} \end{cases}$$

(an identity between polynomials in $\mathbb{Q}[X]$).

[**Hint:** It is enough to prove this when X is replaced by a nonnegative integer r (why?). Now that you have gotten rid of polynomials, introduce new polynomials. Namely, compute the coefficient of X^n in $(1+X)^r (1-X)^r$. Compare with the coefficient of X^n in $\left(1 - X^2\right)^r$.]

Exercise 3.23. Let $n \in \mathbb{N}$.

(a) Prove that

$$\sum_{k=0}^{n} \binom{2k}{k} \binom{2(n-k)}{n-k} = 4^n.$$

(b) Prove that

$$\sum_{k=0}^{n} (-1)^k \binom{2k}{k} \binom{2(n-k)}{n-k} = \begin{cases} 2^n \binom{n}{n/2}, & \text{if } n \text{ is even;} \\ 0, & \text{if } n \text{ is odd} \end{cases}.$$

[**Hint:** Recall Exercise 3.2 **(b)**.]

Exercise 3.24. Let m be a positive integer. Prove the following:

(a) The binomial coefficient $\dbinom{2m}{m}$ is even.

(b) If m is odd and satisfies $m > 1$, then the binomial coefficient $\dbinom{2m-1}{m-1}$ is even.

(c) If m is odd and satisfies $m > 1$, then $\dbinom{2m}{m} \equiv 0 \mod 4$.

Exercise 3.25. For any $m \in \mathbb{N}$ and $n \in \mathbb{N}$, define a rational number $T(m,n)$ by

$$T(m,n) = \frac{(2m)!\,(2n)!}{m!n!\,(m+n)!}.$$

Prove the following facts:

(a) We have $4T(m,n) = T(m+1,n) + T(m,n+1)$ for every $m \in \mathbb{N}$ and $n \in \mathbb{N}$.

(b) We have $T(m,n) \in \mathbb{N}$ for every $m \in \mathbb{N}$ and $n \in \mathbb{N}$.

(c) If $m \in \mathbb{N}$ and $n \in \mathbb{N}$ are such that $(m,n) \neq (0,0)$, then the integer $T(m,n)$ is even.

(d) If $m \in \mathbb{N}$ and $n \in \mathbb{N}$ are such that $m+n$ is odd and $m+n > 1$, then $4 \mid T(m,n)$.

(e) We have $T(m,0) = \dbinom{2m}{m}$ for every $m \in \mathbb{N}$.

(f) We have $T(m,n) = \dfrac{\dbinom{2m}{m}\dbinom{2n}{n}}{\dbinom{m+n}{m}}$ for every $m \in \mathbb{N}$ and $n \in \mathbb{N}$.

(g) We have $T(m,n) = T(n,m)$ for every $m \in \mathbb{N}$ and $n \in \mathbb{N}$.

(h) Let $m \in \mathbb{N}$ and $n \in \mathbb{N}$. Let $p = \min\{m,n\}$. Then,

$$\sum_{k=-p}^{p} (-1)^k \binom{m+n}{m+k}\binom{m+n}{n+k} = \binom{m+n}{m}.$$

(i) Let $m \in \mathbb{N}$ and $n \in \mathbb{N}$. Let $p = \min\{m,n\}$. Then,

$$T(m,n) = \sum_{k=-p}^{p} (-1)^k \binom{2m}{m+k}\binom{2n}{n-k}.$$

Remark 3.52. The numbers $T(m,n)$ introduced in Exercise 3.25 are the so-called *super-Catalan numbers*; much has been written about them (e.g., [Gessel92] and [AleGhe14]). Exercise 3.25 **(b)** suggests that these numbers count something, but no one has so far discovered what. Exercise 3.25 **(i)** is a result of von Szily (1894); see [Gessel92, (29)]. Exercise 3.25 **(b)** is a result of Eugène Catalan (1874), and has also been posed as Problem 3 of the International Mathematical Olympiad 1972. Parts of Exercise 3.25 are also discussed on the thread `https://artofproblemsolving.com/community/c6h1553916s1_supercatalan_numbers` .

The following exercise is a variation on (238):

Exercise 3.26. Let a and b be two integers such that $b \neq 0$. Let $n \in \mathbb{N}$. Show that there exists some $N \in \mathbb{N}$ such that $b^N \dbinom{a/b}{n} \in \mathbb{Z}$.

[**Hint:** I am not aware of a combinatorial solution to this exercise! (I.e., I don't know what the numbers $b^N \dbinom{a/b}{n}$ count, even when they are nonnegative.) All solutions that I know use some (elementary) number theory. For the probably slickest (although unmotivated) solution, basic modular arithmetic suffices; here is a roadmap: First, show that if b and c are integers such that $c > 0$, then there exists an $s \in \mathbb{Z}$ such that $b^{c-1} \equiv sb^c \mod c$ [133]. Apply this to $c = n!$ and conclude that $b^{n!}(a/b - i) \equiv b^{n!}(sa - i) \mod n!$ for every $i \in \mathbb{Z}$. Now use $\dbinom{sa}{n} \in \mathbb{Z}$.]

Exercise 3.27. **(a)** If x and y are two real numbers such that $x + y = 1$, and if $n \in \mathbb{N}$ and $m \in \mathbb{N}$, then prove that

$$x^{m+1} \sum_{k=0}^{n} \binom{m+k}{k} y^k + y^{n+1} \sum_{k=0}^{m} \binom{n+k}{k} x^k = 1.$$

(b) Let $n \in \mathbb{N}$. Prove that

$$\sum_{k=0}^{n} \binom{n+k}{k} \frac{1}{2^k} = 2^n.$$

Remark 3.53. Exercise 3.27 **(a)** is Problem 7 from the IMO Shortlist 1975. It is also closely related to the Daubechies identity [Zeilbe93] (indeed, the first equality in [Zeilbe93] follows by applying it to p, $1-p$, $n-1$ and $n-1$ instead of n and m). Exercise 3.27 **(b)** is a fairly well-known identity for binomial coefficients (see, e.g., [GrKnPa94, (5.20)]).

[133] To prove this, argue that at least two of $b^0, b^1, \ldots, b^c$ are congruent modulo c.

4. Recurrent sequences

4.1. Basics

Two of the most famous integer sequences defined recursively are the Fibonacci sequence and the Lucas sequence:

- The *Fibonacci sequence* is the sequence $(f_0, f_1, f_2, \ldots)$ of integers which is defined recursively by $f_0 = 0$, $f_1 = 1$, and $f_n = f_{n-1} + f_{n-2}$ for all $n \geq 2$. We have already introduced this sequence in Example 2.25. Its first terms are

  $$f_0 = 0, \quad f_1 = 1, \quad f_2 = 1, \quad f_3 = 2, \quad f_4 = 3, \quad f_5 = 5,$$
  $$f_6 = 8, \quad f_7 = 13, \quad f_8 = 21, \quad f_9 = 34, \quad f_{10} = 55,$$
  $$f_{11} = 89, \quad f_{12} = 144, \quad f_{13} = 233.$$

  (Some authors[134] prefer to start the sequence at f_1 rather than f_0; of course, the recursive definition then needs to be modified to require $f_2 = 1$ instead of $f_0 = 0$.)

- The *Lucas sequence* is the sequence $(\ell_0, \ell_1, \ell_2, \ldots)$ of integers which is defined recursively by $\ell_0 = 2$, $\ell_1 = 1$, and $\ell_n = \ell_{n-1} + \ell_{n-2}$ for all $n \geq 2$. Its first terms are

  $$\ell_0 = 2, \quad \ell_1 = 1, \quad \ell_2 = 3, \quad \ell_3 = 4, \quad \ell_4 = 7, \quad \ell_5 = 11,$$
  $$\ell_6 = 18, \quad \ell_7 = 29, \quad \ell_8 = 47, \quad \ell_9 = 76, \quad \ell_{10} = 123,$$
  $$\ell_{11} = 199, \quad \ell_{12} = 322, \quad \ell_{13} = 521.$$

A lot of papers and even books have been written about these two sequences, the relations between them, and the identities that hold for their terms.[135] One of their most striking properties is that they can be computed explicitly, albeit using irrational numbers. In fact, the *Binet formula* says that the n-th Fibonacci number f_n can be computed by

$$f_n = \frac{1}{\sqrt{5}}\varphi^n - \frac{1}{\sqrt{5}}\psi^n, \tag{288}$$

where $\varphi = \dfrac{1+\sqrt{5}}{2}$ and $\psi = \dfrac{1-\sqrt{5}}{2}$ are the two solutions of the quadratic equation $X^2 - X - 1 = 0$. (The number φ is known as the *golden ratio*; the number ψ can be obtained from it by $\psi = 1 - \varphi = -1/\varphi$.) A similar formula, using the very same numbers φ and ψ, exists for the Lucas numbers:

$$\ell_n = \varphi^n + \psi^n. \tag{289}$$

[134] such as Vorobiev in his book [Vorobi02]

[135] See https://oeis.org/A000045 and https://oeis.org/A000032 for an overview of their properties. The book [Vorobi02] is a readable introduction to the Fibonacci sequence, which also surveys a lot of other mathematics (elementary number theory, continued fractions, and even some geometry) along the way. Another introduction to the Fibonacci sequence is [CamFon07].

Remark 4.1. How easy is it to compute f_n and ℓ_n using the formulas (288) and (289)?

This is a nontrivial question. Indeed, if you are careless, you may find them rather useless. For instance, if you try to compute f_n using the formula (288) and using approximate values for the irrational numbers φ and ψ, then you might end up with a wrong value for f_n, because the error in the approximate value for φ propagates when you take φ to the n-th power. (And for high enough n, the error will become larger than 1, so you will not be able to get the correct value by rounding.) The greater n is, the more precise you need a value for φ to approximate f_n this way. Thus, approximating φ is not a good way to compute f_n. (Actually, the opposite is true: You can use (288) to approximate φ by computing Fibonacci numbers. Namely, it is easy to show that $\varphi = \lim\limits_{n\to\infty} \dfrac{f_n}{f_{n-1}}$.)

A better approach to using (288) is to work with the exact values of φ and ψ. To do so, you need to know how to add, subtract, multiply and divide real numbers of the form $a + b\sqrt{5}$ with $a, b \in \mathbb{Q}$ without ever using approximations. (Clearly, φ, ψ and $\sqrt{5}$ all have this form.) There are rules for this, which are simple to check:

$$\begin{aligned}
\left(a+b\sqrt{5}\right)+\left(c+d\sqrt{5}\right) &= (a+c)+(b+d)\sqrt{5};\\
\left(a+b\sqrt{5}\right)-\left(c+d\sqrt{5}\right) &= (a-c)+(b-d)\sqrt{5};\\
\left(a+b\sqrt{5}\right)\cdot\left(c+d\sqrt{5}\right) &= (ac+5bd)+(bc+ad)\sqrt{5};\\
\frac{a+b\sqrt{5}}{c+d\sqrt{5}} &= \frac{(ac-5bd)+(bc-ad)\sqrt{5}}{c^2-5d^2} \qquad \text{for } (c,d)\neq(0,0).
\end{aligned}$$

(The last rule is an instance of "rationalizing the denominator".) These rules give you a way to exactly compute things like φ^n, $\dfrac{1}{\sqrt{5}}\varphi^n$, ψ^n and $\dfrac{1}{\sqrt{5}}\psi^n$, and thus also f_n and ℓ_n. If you use exponentiation by squaring to compute n-th powers, this actually becomes a fast algorithm (a lot faster than just computing f_n and ℓ_n using the recurrence). So, yes, (288) and (289) are useful.

We shall now study a generalization of both the Fibonacci and the Lucas sequences, and generalize (288) and (289) to a broader class of sequences.

Definition 4.2. If a and b are two complex numbers, then a sequence $(x_0, x_1, x_2, \ldots)$ of complex numbers will be called *(a,b)-recurrent* if every $n \geq 2$ satisfies

$$x_n = ax_{n-1} + bx_{n-2}.$$

So, the Fibonacci sequence and the Lucas sequence are $(1,1)$-recurrent. An (a,b)-recurrent sequence $(x_0, x_1, x_2, \ldots)$ is fully determined by the four values a, b, x_0

and x_1, and can be constructed for any choice of these four values. Here are some further examples of (a, b)-recurrent sequences:

- The sequence $(x_0, x_1, x_2, \ldots)$ in Theorem 2.26 is (a, b)-recurrent (by its very definition).

- A sequence $(x_0, x_1, x_2, \ldots)$ is $(2, -1)$-recurrent if and only if every $n \geq 2$ satisfies $x_n = 2x_{n-1} - x_{n-2}$. In other words, a sequence $(x_0, x_1, x_2, \ldots)$ is $(2, -1)$-recurrent if and only if every $n \geq 2$ satisfies $x_n - x_{n-1} = x_{n-1} - x_{n-2}$. In other words, a sequence $(x_0, x_1, x_2, \ldots)$ is $(2, -1)$-recurrent if and only if $x_1 - x_0 = x_2 - x_1 = x_3 - x_2 = \cdots$. In other words, the $(2, -1)$-recurrent sequences are precisely the arithmetic progressions.

- Geometric progressions are also (a, b)-recurrent for appropriate a and b. Namely, any geometric progression $\left(u, uq, uq^2, uq^3, \ldots\right)$ is $(q, 0)$-recurrent, since every $n \geq 2$ satisfies $uq^n = q \cdot uq^{n-1} + 0 \cdot uq^{n-2}$. However, not every $(q, 0)$-recurrent sequence $(x_0, x_1, x_2, \ldots)$ is a geometric progression (since the condition $x_n = qx_{n-1} + 0x_{n-2}$ for all $n \geq 2$ says nothing about x_0, and thus x_0 can be arbitrary).

- A sequence $(x_0, x_1, x_2, \ldots)$ is $(0, 1)$-recurrent if and only if every $n \geq 2$ satisfies $x_n = x_{n-2}$. In other words, a sequence $(x_0, x_1, x_2, \ldots)$ is $(0, 1)$-recurrent if and only if it has the form $(u, v, u, v, u, v, \ldots)$ for two complex numbers u and v.

- A sequence $(x_0, x_1, x_2, \ldots)$ is $(1, 0)$-recurrent if and only if every $n \geq 2$ satisfies $x_n = x_{n-1}$. In other words, a sequence $(x_0, x_1, x_2, \ldots)$ is $(1, 0)$-recurrent if and only if it has the form $(u, v, v, v, v, \ldots)$ for two complex numbers u and v. Notice that u is not required to be equal to v, because we never claimed that $x_n = x_{n-1}$ holds for $n = 1$.

- A sequence $(x_0, x_1, x_2, \ldots)$ is $(1, -1)$-recurrent if and only if every $n \geq 2$ satisfies $x_n = x_{n-1} - x_{n-2}$. Curiously, it turns out that every such sequence is 6-periodic (i.e., it satisfies $x_{n+6} = x_n$ for every $n \in \mathbb{N}$), because every $n \in \mathbb{N}$ satisfies

$$\begin{aligned} x_{n+6} &= \underbrace{x_{n+5}}_{=x_{n+4}-x_{n+3}} - x_{n+4} = \left(x_{n+4} - x_{n+3}\right) - x_{n+4} = - \underbrace{x_{n+3}}_{=x_{n+2}-x_{n+1}} \\ &= -\left(\underbrace{x_{n+2}}_{=x_{n+1}-x_n} - x_{n+1}\right) = -\left(x_{n+1} - x_n - x_{n+1}\right) = x_n. \end{aligned}$$

  More precisely, a sequence $(x_0, x_1, x_2, \ldots)$ is $(1, -1)$-recurrent if and only if it has the form $(u, v, v - u, -u, -v, u - v, \ldots)$ (where the "..." stands for "repeat the preceding 6 values over and over" here) for two complex numbers u and v.

- The above three examples notwithstanding, most (a,b)-recurrent sequences of course are not periodic. However, here is another example which provides a great supply of non-periodic (a,b)-recurrent sequences and, at the same time, explains why we get so many periodic ones: If α is any angle, then the sequences
$$\begin{aligned}&(\sin(0\alpha), \sin(1\alpha), \sin(2\alpha), \ldots) \qquad \text{and}\\ &(\cos(0\alpha), \cos(1\alpha), \cos(2\alpha), \ldots)\end{aligned}$$
are $(2\cos\alpha, -1)$-recurrent. More generally, if α and β are two angles, then the sequence
$$(\sin(\beta+0\alpha), \sin(\beta+1\alpha), \sin(\beta+2\alpha), \ldots)$$
is $(2\cos\alpha, -1)$-recurrent[136]. When $\alpha \in 2\pi\mathbb{Q}$ (that is, $\alpha = 2\pi r$ for some $r \in \mathbb{Q}$), this sequence is periodic.

4.2. Explicit formulas (à la Binet)

Now, we can get an explicit formula (similar to (288) and (289)) for every term of an (a,b)-recurrent sequence (in terms of a, b, x_0 and x_1) in the case when $a^2+4b \neq 0$. Here is how this works:

Remark 4.3. Let a and b be complex numbers such that $a^2+4b \neq 0$. Let $(x_0, x_1, x_2, \ldots)$ be an (a,b)-recurrent sequence. We want to construct an explicit formula for each x_n in terms of x_0, x_1, a and b.

[136] *Proof.* Let α and β be two angles. We need to show that the sequence $(\sin(\beta+0\alpha), \sin(\beta+1\alpha), \sin(\beta+2\alpha), \ldots)$ is $(2\cos\alpha, -1)$-recurrent. In other words, we need to prove that
$$\sin(\beta+n\alpha) = 2\cos\alpha\sin(\beta+(n-1)\alpha) + (-1)\sin(\beta+(n-2)\alpha)$$
for every $n \geq 2$. So fix $n \geq 2$.

One of the well-known trigonometric identities states that $\sin x + \sin y = 2\sin\frac{x+y}{2}\cos\frac{x-y}{2}$ for any two angles x and y. Applying this to $x = \beta+n\alpha$ and $y = \beta+(n-2)\alpha$, we obtain
$$\begin{aligned}\sin(\beta+n\alpha)+\sin(\beta+(n-2)\alpha) &= 2\sin\underbrace{\frac{(\beta+n\alpha)+(\beta+(n-2)\alpha)}{2}}_{=\beta+(n-1)\alpha}\cos\underbrace{\frac{(\beta+n\alpha)-(\beta+(n-2)\alpha)}{2}}_{=\alpha}\\ &= 2\sin(\beta+(n-1)\alpha)\cos\alpha = 2\cos\alpha\sin(\beta+(n-1)\alpha).\end{aligned}$$

Hence,
$$\begin{aligned}\sin(\beta+n\alpha) &= 2\cos\alpha\sin(\beta+(n-1)\alpha) - \sin(\beta+(n-2)\alpha)\\ &= 2\cos\alpha\sin(\beta+(n-1)\alpha) + (-1)\sin(\beta+(n-2)\alpha),\end{aligned}$$

qed.

To do so, we let q_+ and q_- be the two solutions of the quadratic equation $X^2 - aX - b = 0$, namely

$$q_+ = \frac{a + \sqrt{a^2 + 4b}}{2} \qquad \text{and} \qquad q_- = \frac{a - \sqrt{a^2 + 4b}}{2}.$$

We notice that $q_+ \neq q_-$ (since $a^2 + 4b \neq 0$). It is easy to see that the sequences $\left(1, q_+, q_+^2, q_+^3, \ldots\right)$ and $\left(1, q_-, q_-^2, q_-^3, \ldots\right)$ are (a, b)-recurrent. As a consequence, for any two complex numbers λ_+ and λ_-, the sequence

$$\left(\lambda_+ + \lambda_-, \lambda_+ q_+ + \lambda_- q_-, \lambda_+ q_+^2 + \lambda_- q_-^2, \ldots\right)$$

(the n-th term of this sequence, with n starting at 0, is $\lambda_+ q_+^n + \lambda_- q_-^n$) must also be (a, b)-recurrent (check this!). We denote this sequence by L_{λ_+, λ_-}.

We now need to find two complex numbers λ_+ and λ_- such that this sequence L_{λ_+, λ_-} is our sequence $(x_0, x_1, x_2, \ldots)$. In order to do so, we only need to ensure that $\lambda_+ + \lambda_- = x_0$ and $\lambda_+ q_+ + \lambda_- q_- = x_1$ (because once this holds, it will follow that the sequences L_{λ_+, λ_-} and $(x_0, x_1, x_2, \ldots)$ have the same first two terms; and this will yield that these two sequences are identical, because two (a, b)-recurrent sequences with the same first two terms must be identical). That is, we need to solve the system of linear equations

$$\begin{cases} \lambda_+ + \lambda_- = x_0; \\ \lambda_+ q_+ + \lambda_- q_- = x_1 \end{cases} \qquad \text{in the unknowns } \lambda_+ \text{ and } \lambda_-.$$

Thanks to $q_+ \neq q_-$, this system has a unique solution:

$$\lambda_+ = \frac{x_1 - q_- x_0}{q_+ - q_-}; \qquad \lambda_- = \frac{q_+ x_0 - x_1}{q_+ - q_-}.$$

Thus, if we set (λ_+, λ_-) to be this solution, then $(x_0, x_1, x_2, \ldots) = L_{\lambda_+, \lambda_-}$, so that

$$x_n = \lambda_+ q_+^n + \lambda_- q_-^n \tag{290}$$

for every nonnegative integer n. This is an explicit formula, at least if the square roots do not disturb you. When $x_0 = 0$ and $x_1 = a = b = 1$, you get the famous Binet formula (288) for the Fibonacci sequence.

In the next exercise you will see what happens if the $a^2 + 4b \neq 0$ condition does not hold.

Exercise 4.1. Let a and b be complex numbers such that $a^2 + 4b = 0$. Consider an (a, b)-recurrent sequence $(x_0, x_1, x_2, \ldots)$. Find an explicit formula for each x_n in terms of x_0, x_1, a and b.

[**Note:** The polynomial $X^2 - aX - b$ has a double root here. Unlike the case of two distinct roots studied above, you won't see any radicals here. The explicit formula really deserves the name "explicit".]

Remark 4.3 and Exercise 4.1, combined, solve the problem of finding an explicit formula for any term of an (a,b)-recurrent sequence when a and b are complex numbers, at least if you don't mind having square roots in your formula. Similar tactics can be used to find explicit forms for the more general case of sequences satisfying "homogeneous linear recurrences with constant coefficients"[137], although instead of square roots you will now need roots of higher-degree polynomials. (See [LeLeMe16, §22.3.2 ("Solving Homogeneous Linear Recurrences")] for an outline of this; see also [Heffer20, Topic "Linear Recurrences"] for a linear-algebraic introduction.)

4.3. Further results

Here are some more exercises from the theory of recurrent sequences. I am not going particularly deep here, but we may encounter generalizations later.

First, an example: If we "split" the Fibonacci sequence

$$(f_0, f_1, f_2, \ldots) = (0, 1, 1, 2, 3, 5, 8, \ldots)$$

into two subsequences

$$(f_0, f_2, f_4, \ldots) = (0, 1, 3, 8, 21, \ldots) \qquad \text{and} \qquad (f_1, f_3, f_5, \ldots) = (1, 2, 5, 13, \ldots)$$

(each of which contains every other Fibonacci number), then it turns out that each of these two subsequences is $(3, -1)$-recurrent[138]. This is rather easy to prove, but one can always ask for generalizations: What happens if we start with an arbitrary (a,b)-recurrent sequence, instead of the Fibonacci numbers? What happens if we split it into three, four or more subsequences? The answer is rather nice:

Exercise 4.2. Let a and b be complex numbers. Let $(x_0, x_1, x_2, \ldots)$ be an (a,b)-recurrent sequence.

(a) Prove that the sequences $(x_0, x_2, x_4, \ldots)$ and $(x_1, x_3, x_5, \ldots)$ are (c,d)-recurrent for some complex numbers c and d. Find these c and d.

(b) Prove that the sequences $(x_0, x_3, x_6, \ldots)$, $(x_1, x_4, x_7, \ldots)$ and $(x_2, x_5, x_8, \ldots)$ are (c,d)-recurrent for some (other) complex numbers c and d.

(c) For every nonnegative integers N and K, prove that the sequence $(x_K, x_{N+K}, x_{2N+K}, x_{3N+K}, \ldots)$ is (c,d)-recurrent for some complex numbers c and d which depend only on N, a and b (but not on K or x_0 or x_1).

[137] These are sequences $(x_0, x_1, x_2, \ldots)$ which satisfy

$$(x_n = c_1 x_{n-1} + c_2 x_{n-2} + \cdots + c_k x_{n-k} \qquad \text{for all } n \geq k)$$

for a fixed $k \in \mathbb{N}$ and a fixed k-tuple $(c_1, c_2, \ldots, c_k)$ of complex numbers. When $k = 2$, these are the (c_1, c_2)-recurrent sequences.

[138] In other words, we have $f_{2n} = 3f_{2(n-1)} + (-1) f_{2(n-2)}$ and $f_{2n+1} = 3f_{2(n-1)+1} + (-1) f_{2(n-2)+1}$ for every $n \geq 2$.

The next exercise gives a combinatorial interpretation of the Fibonacci numbers:

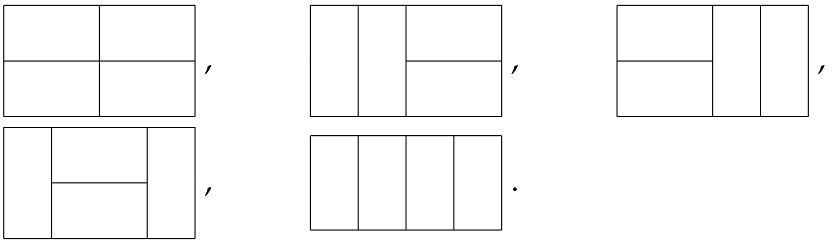
Exercise 4.3. Recall that the Fibonacci numbers $f_0, f_1, f_2, \ldots$ are defined recursively by $f_0 = 0$, $f_1 = 1$ and $f_n = f_{n-1} + f_{n-2}$ for all $n \geq 2$. For every positive integer n, show that f_n is the number of subsets I of $\{1, 2, \ldots, n-2\}$ such that no two elements of I are consecutive (i.e., there exists no $i \in \mathbb{Z}$ such that both i and $i+1$ belong to I). For instance, for $n = 5$, these subsets are $\varnothing$, $\{1\}$, $\{2\}$, $\{3\}$ and $\{1, 3\}$.

Notice that $\{1, 2, \ldots, -1\}$ is to be understood as the empty set (since there are no integers x satisfying $1 \leq x \leq -1$). (So Exercise 4.3, applied to $n = 1$, says that f_1 is the number of subsets I of the empty set such that no two elements of I are consecutive. This is correct, because the empty set has only one subset, which of course is empty and thus has no consecutive elements; and the Fibonacci number f_1 is precisely 1.)

Remark 4.4. Exercise 4.3 is equivalent to another known combinatorial interpretation of the Fibonacci numbers.

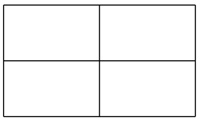
Namely, let n be a positive integer. Consider a rectangular table of dimensions $2 \times (n-1)$ (that is, with 2 rows and $n-1$ columns). How many ways are there to subdivide this table into dominos? (A *domino* means a set of two adjacent boxes.)

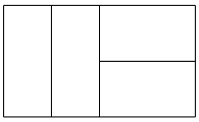
For $n = 5$, there are 5 ways:

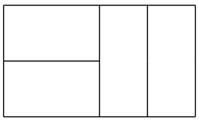


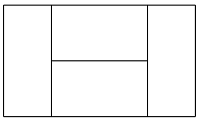
In the general case, there are f_n ways. Why?

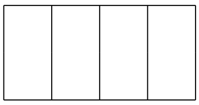
As promised, this result is equivalent to Exercise 4.3. Let us see why. Let P be a way to subdivide the table into dominos. We say that a *horizontal domino* is a domino which consists of two adjacent boxes in the same row; similarly, we define a vertical domino. It is easy to see that (in the subdivision P) each column of the table is covered either by a single vertical domino, or by two horizontal dominos (in which case either both of them "begin" in this column, or both of them "end" in this column). Let $J(P)$ be the set of all $i \in \{1, 2, \ldots, n-1\}$ such that the i-th column of the table is covered by two horizontal dominos, both of which "begin" in this column. For instance,

$J\Bigg($  $\Bigg) = \{1,3\}\,;$

$J\Bigg($  $\Bigg) = \{3\}\,;$

$J\Bigg($  $\Bigg) = \{1\}\,;$

$J\Bigg($  $\Bigg) = \{2\}\,;$

$J\Bigg($  $\Bigg) = \varnothing.$

It is easy to see that the set $J(P)$ is a subset of $\{1,2,\ldots,n-2\}$ containing no two consecutive integers. Moreover, this set $J(P)$ uniquely determines P, and for every subset I of $\{1,2,\ldots,n-2\}$ containing no two consecutive integers, there exists some way P to subdivide the table into dominos such that $J(P) = I$.

Hence, the number of all ways to subdivide the table into dominos equals the number of all subsets I of $\{1,2,\ldots,n-2\}$ containing no two consecutive integers. Exercise 4.3 says that this latter number is f_n; therefore, so is the former number.

(I have made this remark because I found it instructive. If you merely want a proof that the number of all ways to subdivide the table into dominos equals f_n, then I guess it is easier to just prove it by induction without taking the detour through Exercise 4.3. This proof is sketched in [GrKnPa94, §7.1], followed by an informal yet insightful discussion of "infinite sums of dominos" and various related ideas.)

Either Exercise 4.3 or Remark 4.4 can be used to prove properties of Fibonacci numbers in a combinatorial way; see [BenQui04] for some examples of such proofs.

Here is another formula for certain recursive sequences, coming out of a recent paper on cluster algebras[139]:

[139] Specifically, Exercise 4.4 is part of [LeeSch11, Definition 1], but I have reindexed the sequence and fixed the missing upper bound in the sum.

Exercise 4.4. Let $r \in \mathbb{Z}$. Define a sequence $(c_0, c_1, c_2, \ldots)$ of integers recursively by $c_0 = 0$, $c_1 = 1$ and $c_n = rc_{n-1} - c_{n-2}$ for all $n \geq 2$. Show that

$$c_n = \sum_{i=0}^{n-1} (-1)^i \binom{n-1-i}{i} r^{n-1-2i} \qquad (291)$$

for every $n \in \mathbb{N}$. Here, we use the following convention: Any expression of the form $a \cdot b$, where a is 0, has to be interpreted as 0, even if b is undefined.[140]

4.4. Additional exercises

This section contains some further exercises. As the earlier "additional exercises", these will not be relied on in the rest of this text, and solutions will not be provided.

Exercise 4.5. Let q and r be two complex numbers. Prove that the sequence $\left(q^0 - r^0, q^1 - r^1, q^2 - r^2, \ldots\right)$ is (a,b)-recurrent for two appropriately chosen a and b. Find these a and b.

Exercise 4.6. Let φ be the golden ratio (i.e., the real number $\dfrac{1+\sqrt{5}}{2}$). Let $(f_0, f_1, f_2, \ldots)$ be the Fibonacci sequence.

(a) Show that $f_{n+1} - \varphi f_n = \dfrac{1}{\sqrt{5}}\psi^n$ for every $n \in \mathbb{N}$, where $\psi = \dfrac{1-\sqrt{5}}{2}$. (Notice that $\psi = \dfrac{1-\sqrt{5}}{2} \approx -0.618$ lies between -1 and 0, and thus the powers ψ^n converge to 0 as $n \to \infty$. So $f_{n+1} - \varphi f_n \to 0$ as $n \to \infty$, and consequently $\dfrac{f_{n+1}}{f_n} \to \varphi$ as well.)

(b) Show that

$$f_n = \operatorname{round}\left(\frac{1}{\sqrt{5}}\varphi^n\right) \qquad \text{for every } n \in \mathbb{N}.$$

Here, if x is a real number, then $\operatorname{round} x$ denotes the integer closest to x (where, in case of a tie, we take the higher of the two candidates[141]).

[140] The purpose of this convention is to make sure that the right hand side of (291) is well-defined, even though the expression r^{n-1-2i} that appears in it might be undefined (it will be undefined when $r = 0$ and $n - 1 - 2i < 0$).

Of course, the downside of this convention is that we might not have $a \cdot b = b \cdot a$ (because $a \cdot b$ might be well-defined while $b \cdot a$ is not, or vice versa).

[141] This does not really matter in our situation, because $\dfrac{1}{\sqrt{5}}\varphi^n$ will never be a half-integer.

Exercise 4.7. Let $(f_0, f_1, f_2, \ldots)$ be the Fibonacci sequence. A set I of integers is said to be *lacunar* if no two elements of I are consecutive (i.e., there exists no $i \in I$ such that $i + 1 \in I$). Show that, for every $n \in \mathbb{N}$, there exists a unique lacunar subset S of $\{2, 3, 4, \ldots\}$ such that $n = \sum\limits_{s \in S} f_s$.

(For example, if $n = 17$, then $S = \{2, 4, 7\}$, because $17 = 1 + 3 + 13 = f_2 + f_4 + f_7$.)

Remark 4.5. The representation of n in the form $n = \sum\limits_{s \in S} f_s$ in Exercise 4.7 is known as the *Zeckendorf representation* of n. It has a number of interesting properties and trivia related to it; for example, there is a rule of thumb for converting miles into kilometers that uses it. It can also be used to define a curious "Fibonacci multiplication" operation on nonnegative integers [Knuth88].

Exercise 4.8. Let $(f_0, f_1, f_2, \ldots)$ be the Fibonacci sequence.

(a) Prove the identities

$$\begin{aligned} 1f_n &= f_n && \text{for all } n \geq 0; \\ 2f_n &= f_{n-2} + f_{n+1} && \text{for all } n \geq 2; \\ 3f_n &= f_{n-2} + f_{n+2} && \text{for all } n \geq 2; \\ 4f_n &= f_{n-2} + f_n + f_{n+2} && \text{for all } n \geq 2. \end{aligned}$$

(b) Notice that the right hand sides of these identities have a specific form: they are sums of f_{n+t} for t ranging over a lacunar subset of $\mathbb{Z}$. (See Exercise 4.7 for the definition of "lacunar".) Try to find similar identities for $5f_n$ and $6f_n$.

(c) Prove that such identities exist in general. More precisely, prove the following: Let T be a finite set, and a_t be an integer for every $t \in T$. Then, there exists a unique finite lacunar subset S of $\mathbb{Z}$ such that

$$\sum_{t \in T} f_{n+a_t} = \sum_{s \in S} f_{n+s} \qquad \text{for every } n \in \mathbb{Z} \text{ which}$$
$$\text{satisfies } n \geq \max\left(\{-a_t \mid t \in T\} \cup \{-s \mid s \in S\}\right).$$

(The condition $n \geq \max\left(\{-a_t \mid t \in T\} \cup \{-s \mid s \in S\}\right)$ merely ensures that all the f_{n+a_t} and f_{n+s} are well-defined.)

Remark 4.6. Exercise 4.8 **(c)** is [Grinbe11, Theorem 1.4]. It is also a consequence of [CamFon07, Lemma 6.2] (applied to $k = 2$). I'd be delighted to see other proofs!

Similarly I am highly interested in analogues of Exercises 4.7 and 4.8 for other (a, b)-recurrent sequences (e.g., Lucas numbers).

Exercise 4.9. **(a)** Let $(f_0, f_1, f_2, \ldots)$ be the Fibonacci sequence. For every $n \in \mathbb{N}$ and $k \in \mathbb{N}$ satisfying $0 \leq k \leq n$, define a rational number $\dbinom{n}{k}_F$ by

$$\binom{n}{k}_F = \frac{f_n f_{n-1} \cdots f_{n-k+1}}{f_k f_{k-1} \cdots f_1}.$$

This is called the (n,k)-th *Fibonomial coefficient* (in analogy to the binomial coefficient $\dbinom{n}{k} = \dfrac{n(n-1)\cdots(n-k+1)}{k(k-1)\cdots 1}$).

Show that $\dbinom{n}{k}_F$ is an integer.

(b) Try to extend as many identities for binomial coefficients as you can to Fibonomial coefficients.

(c) Generalize to (a,b)-recurrent sequences with arbitrary a and b.

5. Permutations

This chapter is devoted to permutations. We first recall how they are defined.

5.1. Permutations and the symmetric group

Definition 5.1. First, let us stipulate, once and for all, how we define the composition of two maps: If X, Y and Z are three sets, and if $\alpha : X \to Y$ and $\beta : Y \to Z$ are two maps, then $\beta \circ \alpha$ denotes the map from X to Z which sends every $x \in X$ to $\beta(\alpha(x))$. This map $\beta \circ \alpha$ is called the *composition* of β and α (and is sometimes abbreviated as $\beta\alpha$). This is the classical notation for composition of maps, and the reason why I am so explicitly reminding you of it is that some people (e.g., Herstein in [Herstei75]) use a different convention that conflicts with it: They write maps "on the right" (i.e., they denote the image of an element $x \in X$ under the map $\alpha : X \to Y$ by x^α or $x\alpha$ instead of $\alpha(x)$), and they define composition "the other way round" (i.e., they write $\alpha \circ \beta$ for what we call $\beta \circ \alpha$). They have reasons for what they are doing, but I shall use the classical notation because most of the literature agrees with it.

Definition 5.2. Let us also recall what it means for two maps to be *inverse*.

Let X and Y be two sets. Two maps $f : X \to Y$ and $g : Y \to X$ are said to be *mutually inverse* if they satisfy $g \circ f = \mathrm{id}_X$ and $f \circ g = \mathrm{id}_Y$. (In other words, two maps $f : X \to Y$ and $g : Y \to X$ are mutually inverse if and only if every $x \in X$ satisfies $g(f(x)) = x$ and every $y \in Y$ satisfies $f(g(y)) = y$.)

Let $f : X \to Y$ be a map. If there exists a map $g : Y \to X$ such that f and g are mutually inverse, then this map g is unique (this is easy to check) and is called the *inverse* of f and denoted by f^{-1}. In this case, the map f is said to be *invertible*. It is easy to see that if g is the inverse of f, then f is the inverse of g.

It is well-known that a map $f : X \to Y$ is invertible if and only if f is bijective (i.e., both injective and surjective). The words "invertible" and "bijective" are thus synonyms (at least when used for a map between two sets – in other situations, they can be rather different). Nevertheless, both of them are commonly used, often by the same authors (since they convey slightly different mental images).

A bijective map is also called a *bijection* or a *1-to-1 correspondence* (or a *one-to-one correspondence*). When there is a bijection from X to Y, one says that the elements of X are *in bijection with* (or *in one-to-one correspondence with*) the elements of Y. It is well-known that two sets X and Y have the same cardinality if and only if there exists a bijection from X to Y. (This is precisely Theorem 1.2.)

Definition 5.3. A *permutation* of a set X means a bijection from X to X. The permutations of a given set X can be composed (i.e., if α and β are two permutations of X, then so is $\alpha \circ \beta$) and have inverses (which, again, are permutations of X). More precisely:

- If α and β are two permutations of a given set X, then the composition $\alpha \circ \beta$ is again a permutation of X.
- Any three permutations α, β and γ of X satisfy $(\alpha \circ \beta) \circ \gamma = \alpha \circ (\beta \circ \gamma)$. (This holds, more generally, for arbitrary maps which can be composed.)
- The identity map $\operatorname{id} : X \to X$ (this is the map which sends every element $x \in X$ to itself) is a permutation of X; it is also called the *identity permutation*. Every permutation α of X satisfies $\operatorname{id} \circ \alpha = \alpha$ and $\alpha \circ \operatorname{id} = \alpha$. (Again, this can be generalized to arbitrary maps.)
- For every permutation α of X, the inverse map α^{-1} is well-defined and is again a permutation of X. We have $\alpha \circ \alpha^{-1} = \operatorname{id}$ and $\alpha^{-1} \circ \alpha = \operatorname{id}$.

In the lingo of algebraists, these four properties show that the set of all permutations of X is a group whose binary operation is composition, and whose neutral element is the identity permutation $\operatorname{id} : X \to X$. This group is known as the *symmetric group of the set* X. (We will define the notion of a group later, in Definition 6.116; thus you might not understand the preceding two sentences at this point. If you do not care about groups, you should just remember that the symmetric group of X is the set of all permutations of X.)

Remark 5.4. Some authors define a permutation of a finite set X to mean a list of all elements of X, each occurring exactly once. This is **not** the meaning that the word "permutation" has in these notes! It is a different notion which, for historical reasons, has been called "permutation" as well. On the Wikipedia page for "permutation", the two notions are called "active" and "passive", respectively: An "active" permutation of X means a bijection from X to X (that is, a permutation of X in our meaning of this word), whereas a "passive" permutation of X means a list of all elements of X, each occurring exactly once. For example, if $X = \{\text{"cat"}, \text{"dog"}, \text{"archaeopteryx"}\}$, then the map

$$\begin{aligned} \text{"cat"} &\mapsto \text{"archaeopteryx"}, \\ \text{"archaeopteryx"} &\mapsto \text{"dog"}, \\ \text{"dog"} &\mapsto \text{"cat"} \end{aligned}$$

is an "active" permutation of X, whereas the list ("dog", "cat", "archaeopteryx") is a "passive" permutation of X.

When X is the set $\{1, 2, \ldots, n\}$ for some $n \in \mathbb{N}$, then it is possible to equate each "active" permutation of X with a "passive" permutation of X (namely, its one-line notation, defined below). More generally, this can be done when X comes with a fixed total order. In general, if X is a finite set, then the number of "active" permutations of X equals the number of "passive" permutations of X (and both numbers equal $|X|!$), but until you fix some ordering of the elements of X, there is no "natural" way to match the "passive" permutations with the "active" ones. (And when X is infinite, the notion of a "passive" permutation is not even well-defined.)

To reiterate: For us, the word "permutation" shall always mean an "active" permutation!

Recall that $\mathbb{N} = \{0, 1, 2, \ldots\}$.

Definition 5.5. Let $n \in \mathbb{N}$.

Let S_n be the symmetric group of the set $\{1, 2, \ldots, n\}$. This is the set of all permutations of the set $\{1, 2, \ldots, n\}$. It contains the identity permutation $\operatorname{id} \in S_n$ which sends every $i \in \{1, 2, \ldots, n\}$ to i.

A well-known fact states that for every $n \in \mathbb{N}$, the size of the symmetric group S_n is $|S_n| = n!$ (that is, there are exactly $n!$ permutations of $\{1, 2, \ldots, n\}$). (One proof of this fact – not the simplest – is given in the proof of Corollary **??** below.)

We will often write a permutation $\sigma \in S_n$ as the list $(\sigma(1), \sigma(2), \ldots, \sigma(n))$ of its values. This is known as the *one-line notation* for permutations (because it is a single-rowed list, as opposed to e.g. the two-line notation which is a two-rowed table).[142] For instance, the permutation in S_3 which sends 1 to 2, 2 to 1 and 3 to 3 is written $(2, 1, 3)$ in one-line notation.

The exact relation between lists and permutations is given by the following simple fact:

[142] Combinatorialists often omit the parentheses and the commas (i.e., they just write $\sigma(1)\sigma(2)\cdots\sigma(n)$, hoping that noone will mistake this for a product), since there is unfortunately another notation for permutations (the *cycle notation*) which also writes them as lists (actually, lists of lists) but where the lists have a different meaning.

Proposition 5.6. Let $n \in \mathbb{N}$. Let $[n] = \{1, 2, \ldots, n\}$.

(a) If $\sigma \in S_n$, then each element of $[n]$ appears exactly once in the list $(\sigma(1), \sigma(2), \ldots, \sigma(n))$.

(b) If $(p_1, p_2, \ldots, p_n)$ is a list of elements of $[n]$ such that each element of $[n]$ appears exactly once in this list $(p_1, p_2, \ldots, p_n)$, then there exists a unique permutation $\sigma \in S_n$ such that $(p_1, p_2, \ldots, p_n) = (\sigma(1), \sigma(2), \ldots, \sigma(n))$.

(c) Let $k \in \{0, 1, \ldots, n\}$. If $(p_1, p_2, \ldots, p_k)$ is a list of some elements of $[n]$ such that $p_1, p_2, \ldots, p_k$ are distinct, then there exists a permutation $\sigma \in S_n$ such that $(p_1, p_2, \ldots, p_k) = (\sigma(1), \sigma(2), \ldots, \sigma(k))$.

At this point, let us clarify what we mean by "distinct": Several objects $u_1, u_2, \ldots, u_k$ are said to be *distinct* if every $i \in \{1, 2, \ldots, k\}$ and $j \in \{1, 2, \ldots, k\}$ satisfying $i \neq j$ satisfy $u_i \neq u_j$. (Some people call this "pairwise distinct".) So, for example, the numbers $2, 1, 6$ are distinct, but the numbers $6, 1, 6$ are not (although 6 and 1 are distinct). Instead of saying that some objects $u_1, u_2, \ldots, u_k$ are distinct, we can also say that "the list $(u_1, u_2, \ldots, u_k)$ has no repetitions"[143].

Remark 5.7. The σ in Proposition 5.6 **(b)** is uniquely determined, but the σ in Proposition 5.6 **(c)** is not (in general). More precisely, in Proposition 5.6 **(c)**, there are $(n-k)!$ possible choices of σ that work. (This is easy to check.)

Proof of Proposition 5.6. Proposition 5.6 is a basic fact and its proof is simple. I am going to present the proof in great detail, but you are not missing much if you skip it for its obviousness (just make sure you know **why** it is obvious).

Recall that S_n is the set of all permutations of the set $\{1, 2, \ldots, n\}$. In other words, S_n is the set of all permutations of the set $[n]$ (since $\{1, 2, \ldots, n\} = [n]$).

(a) Let $\sigma \in S_n$. Let $i \in [n]$.

We have $\sigma \in S_n$. In other words, σ is a permutation of $[n]$ (since S_n is the set of all permutations of the set $[n]$). In other words, σ is a bijective map $[n] \to [n]$. Hence, σ is both surjective and injective.

Now, we make the following two observations:

- The number i appears in the list $(\sigma(1), \sigma(2), \ldots, \sigma(n))$ [144].
- The number i appears at most once in the list $(\sigma(1), \sigma(2), \ldots, \sigma(n))$ [145].

[143] A repetition just means an element which occurs more than once in the list. It does not matter whether the occurrences are at consecutive positions or not.

[144] *Proof.* The map σ is surjective. Hence, there exists some $j \in [n]$ such that $i = \sigma(j)$. In other words, the number i appears in the list $(\sigma(1), \sigma(2), \ldots, \sigma(n))$. Qed.

[145] *Proof.* Let us assume the contrary (for the sake of contradiction). Thus, i appears more than once in the list $(\sigma(1), \sigma(2), \ldots, \sigma(n))$. In other words, i appears at least twice in this list. In other words, there exist two distinct elements p and q of $[n]$ such that $\sigma(p) = i$ and $\sigma(q) = i$. Consider these p and q.

We have $p \neq q$ (since p and q are distinct), so that $\sigma(p) \neq \sigma(q)$ (since σ is injective). This contradicts $\sigma(p) = i = \sigma(q)$. This contradiction proves that our assumption was wrong, qed.

Combining these two observations, we conclude that the number i appears exactly once in the list $(\sigma(1), \sigma(2), \ldots, \sigma(n))$.

Let us now forget that we fixed i. We thus have shown that if $i \in [n]$, then i appears exactly once in the list $(\sigma(1), \sigma(2), \ldots, \sigma(n))$. In other words, each element of $[n]$ appears exactly once in the list $(\sigma(1), \sigma(2), \ldots, \sigma(n))$. This proves Proposition 5.6 **(a)**.

(b) Let $(p_1, p_2, \ldots, p_n)$ be a list of elements of $[n]$ such that each element of $[n]$ appears exactly once in this list $(p_1, p_2, \ldots, p_n)$.

We have $p_i \in [n]$ for every $i \in [n]$ (since $(p_1, p_2, \ldots, p_n)$ is a list of elements of $[n]$).

We define a map $\tau : [n] \to [n]$ by setting

$$(\tau(i) = p_i \qquad \text{for every } i \in [n]). \tag{292}$$

(This is well-defined, because we have $p_i \in [n]$ for every $i \in [n]$.) The map τ is injective[146] and surjective[147]. Hence, the map τ is bijective. In other words, τ is a permutation of $[n]$ (since τ is a map $[n] \to [n]$). In other words, $\tau \in S_n$ (since S_n is the set of all permutations of the set $[n]$). Clearly, $(\tau(1), \tau(2), \ldots, \tau(n)) = (p_1, p_2, \ldots, p_n)$ (because of (292)), so that $(p_1, p_2, \ldots, p_n) = (\tau(1), \tau(2), \ldots, \tau(n))$.

Hence, there exists a permutation $\sigma \in S_n$ such that $(p_1, p_2, \ldots, p_n) = (\sigma(1), \sigma(2), \ldots, \sigma(n))$ (namely, $\sigma = \tau$). Moreover, there exists **at most one** such permutation[148]. Combining the claims of the previous two

[146] *Proof.* Let u and v be two elements of $[n]$ such that $\tau(u) = \tau(v)$. We shall show that $u = v$.

Indeed, we assume the contrary (for the sake of contradiction). Thus, $u \neq v$.

The definition of $\tau(u)$ shows that $\tau(u) = p_u$. But we also have $\tau(u) = \tau(v) = p_v$ (by the definition of $\tau(v)$). Now, the element $\tau(u)$ of $[n]$ appears (at least) twice in the list $(p_1, p_2, \ldots, p_n)$: once at the u-th position (since $\tau(u) = p_u$), and again at the v-th position (since $\tau(u) = p_v$). (And these are two distinct positions, because $u \neq v$.)

But let us recall that each element of $[n]$ appears exactly once in this list $(p_1, p_2, \ldots, p_n)$. Hence, no element of $[n]$ appears more than once in the list $(p_1, p_2, \ldots, p_n)$. In particular, $\tau(u)$ cannot appear more than once in this list $(p_1, p_2, \ldots, p_n)$. This contradicts the fact that $\tau(u)$ appears twice in the list $(p_1, p_2, \ldots, p_n)$.

This contradiction shows that our assumption was wrong. Hence, $u = v$ is proven.

Now, let us forget that we fixed u and v. We thus have proven that if u and v are two elements of $[n]$ such that $\tau(u) = \tau(v)$, then $u = v$. In other words, the map τ is injective. Qed.

[147] *Proof.* Let $u \in [n]$. Each element of $[n]$ appears exactly once in the list $(p_1, p_2, \ldots, p_n)$. Applying this to the element u of $[n]$, we conclude that u appears exactly once in the list $(p_1, p_2, \ldots, p_n)$. In other words, there exists exactly one $i \in [n]$ such that $u = p_i$. Consider this i. The definition of τ yields $\tau(i) = p_i$. Compared with $u = p_i$, this yields $\tau(i) = u$.

Hence, there exists a $j \in [n]$ such that $\tau(j) = u$ (namely, $j = i$).

Let us now forget that we fixed u. We thus have proven that for every $u \in [n]$, there exists a $j \in [n]$ such that $\tau(j) = u$. In other words, the map τ is surjective. Qed.

[148] *Proof.* Let σ_1 and σ_2 be two permutations $\sigma \in S_n$ such that $(p_1, p_2, \ldots, p_n) = (\sigma(1), \sigma(2), \ldots, \sigma(n))$. Thus, σ_1 is a permutation in S_n such that $(p_1, p_2, \ldots, p_n) = (\sigma_1(1), \sigma_1(2), \ldots, \sigma_1(n))$, and σ_2 is a permutation in S_n such that $(p_1, p_2, \ldots, p_n) = (\sigma_2(1), \sigma_2(2), \ldots, \sigma_2(n))$.

We have $(\sigma_1(1), \sigma_1(2), \ldots, \sigma_1(n)) = (p_1, p_2, \ldots, p_n) = (\sigma_2(1), \sigma_2(2), \ldots, \sigma_2(n))$. In other words, every $i \in [n]$ satisfies $\sigma_1(i) = \sigma_2(i)$. In other words, $\sigma_1 = \sigma_2$.

sentences, we conclude that there exists a unique permutation $\sigma \in S_n$ such that $(p_1, p_2, \ldots, p_n) = (\sigma(1), \sigma(2), \ldots, \sigma(n))$. This proves Proposition 5.6 **(b)**.

(c) Let $(p_1, p_2, \ldots, p_k)$ be a list of some elements of $[n]$ such that $p_1, p_2, \ldots, p_k$ are distinct. Thus, the list $(p_1, p_2, \ldots, p_k)$ contains k of the n elements of $[n]$ (because $p_1, p_2, \ldots, p_k$ are distinct). Let $q_1, q_2, \ldots, q_{n-k}$ be the remaining $n-k$ elements of $[n]$ (listed in any arbitrary order, with no repetition). Then, $(p_1, p_2, \ldots, p_k, q_1, q_2, \ldots, q_{n-k})$ is a list of all n elements of $[n]$, with no repetitions[149]. In other words, each element of $[n]$ appears exactly once in this list $(p_1, p_2, \ldots, p_k, q_1, q_2, \ldots, q_{n-k})$ (and each entry in this list is an element of $[n]$). Hence, we can apply Proposition 5.6 **(b)** to $(p_1, p_2, \ldots, p_k, q_1, q_2, \ldots, q_{n-k})$ instead of $(p_1, p_2, \ldots, p_n)$. As a consequence, we conclude that there exists a unique permutation $\sigma \in S_n$ such that $(p_1, p_2, \ldots, p_k, q_1, q_2, \ldots, q_{n-k}) = (\sigma(1), \sigma(2), \ldots, \sigma(n))$. Let τ be this σ.

Thus, $\tau \in S_n$ is a permutation such that

$$(p_1, p_2, \ldots, p_k, q_1, q_2, \ldots, q_{n-k}) = (\tau(1), \tau(2), \ldots, \tau(n)).$$

Now,

$$\begin{aligned}
&(p_1, p_2, \ldots, p_k) \\
&= \left(\text{the list of the first } k \text{ entries of the list } \underbrace{(p_1, p_2, \ldots, p_k, q_1, q_2, \ldots, q_{n-k})}_{=(\tau(1),\tau(2),\ldots,\tau(n))}\right) \\
&= (\text{the list of the first } k \text{ entries of the list } (\tau(1), \tau(2), \ldots, \tau(n))) \\
&= (\tau(1), \tau(2), \ldots, \tau(k)).
\end{aligned}$$

Hence, there exists a permutation $\sigma \in S_n$ such that $(p_1, p_2, \ldots, p_k) = (\sigma(1), \sigma(2), \ldots, \sigma(k))$ (namely, $\sigma = \tau$). This proves Proposition 5.6 **(c)**. □

5.2. Inversions, lengths and the permutations $s_i \in S_n$

Let us now forget that we fixed σ_1 and σ_2. We thus have shown that if σ_1 and σ_2 are two permutations $\sigma \in S_n$ such that $(p_1, p_2, \ldots, p_n) = (\sigma(1), \sigma(2), \ldots, \sigma(n))$, then $\sigma_1 = \sigma_2$. In other words, any two permutations $\sigma \in S_n$ such that $(p_1, p_2, \ldots, p_n) = (\sigma(1), \sigma(2), \ldots, \sigma(n))$ must be equal to each other. In other words, there exists **at most one** permutation $\sigma \in S_n$ such that $(p_1, p_2, \ldots, p_n) = (\sigma(1), \sigma(2), \ldots, \sigma(n))$. Qed.

[149] It has no repetitions because:

- there are no repetitions among $p_1, p_2, \ldots, p_k$;
- there are no repetitions among $q_1, q_2, \ldots, q_{n-k}$;
- the two lists $(p_1, p_2, \ldots, p_k)$ and $(q_1, q_2, \ldots, q_{n-k})$ have no elements in common (because we defined $q_1, q_2, \ldots, q_{n-k}$ to be the "remaining" $n-k$ elements of $[n]$, where "remaining" means "not contained in the list $(p_1, p_2, \ldots, p_k)$").

Definition 5.8. Let $n \in \mathbb{N}$. For each $i \in \{1, 2, \ldots, n-1\}$, let s_i be the permutation in S_n that swaps i with $i+1$ but leaves all other numbers unchanged. Formally speaking, s_i is the permutation in S_n given by

$$\left(s_i(k) = \begin{cases} i+1, & \text{if } k = i; \\ i, & \text{if } k = i+1; \\ k, & \text{if } k \notin \{i, i+1\} \end{cases} \qquad \text{for all } k \in \{1, 2, \ldots, n\} \right).$$

Thus, in one-line notation

$$s_i = (1, 2, \ldots, i-1, i+1, i, i+2, \ldots, n).$$

Notice that $s_i^2 = \operatorname{id}$ for every $i \in \{1, 2, \ldots, n-1\}$. (Here, we are using the notation α^2 for $\alpha \circ \alpha$, where α is a permutation in S_n.)

Exercise 5.1. Let $n \in \mathbb{N}$.

(a) Show that $s_i \circ s_{i+1} \circ s_i = s_{i+1} \circ s_i \circ s_{i+1}$ for all $i \in \{1, 2, \ldots, n-2\}$.

(b) Show that every permutation $\sigma \in S_n$ can be written as a composition of several permutations of the form s_k (with $k \in \{1, 2, \ldots, n-1\}$). For example, if $n = 3$, then the permutation[150] $(3, 1, 2)$ in S_3 can be written as the composition $s_2 \circ s_1$, while the permutation $(3, 2, 1)$ in S_3 can be written as the composition $s_1 \circ s_2 \circ s_1$ or also as the composition $s_2 \circ s_1 \circ s_2$.

[**Hint:** If you do not immediately see why this works, consider reading further.]

(c) Let w_0 denote the permutation in S_n which sends each $k \in \{1, 2, \ldots, n\}$ to $n + 1 - k$. (In one-line notation, this w_0 is written as $(n, n-1, \ldots, 1)$.) Find an **explicit** way to write w_0 as a composition of several permutations of the form s_i (with $i \in \{1, 2, \ldots, n-1\}$).

Remark 5.9. Symmetric groups appear in almost all parts of mathematics; unsurprisingly, there is no universally accepted notation for them. We are using the notation S_n for the n-th symmetric group; other common notations for it are $\mathfrak{S}_n$, Σ_n and $\operatorname{Sym}(n)$. The permutations that we call $s_1, s_2, \ldots, s_{n-1}$ are often called $\sigma_1, \sigma_2, \ldots, \sigma_{n-1}$. As already mentioned in Definition 5.1, some people write the composition of maps "backwards", which causes their $\sigma \circ \tau$ to be our $\tau \circ \sigma$, etc.. (Sadly, most authors are so sure that their notation is standard that they never bother to define it.)

In the language of group theory, the statement of Exercise 5.1 **(b)** says (or, more precisely, yields) that the permutations $s_1, s_2, \ldots, s_{n-1}$ generate the group S_n.

[150] Recall that we are writing permutations in one-line notation. Thus, "the permutation $(3, 1, 2)$ in S_3" means the permutation $\sigma \in S_3$ satisfying $(\sigma(1), \sigma(2), \sigma(3)) = (3, 1, 2)$.

Definition 5.10. Let $n \in \mathbb{N}$. Let $\sigma \in S_n$ be a permutation.

(a) An *inversion* of σ means a pair (i, j) of integers satisfying $1 \leq i < j \leq n$ and $\sigma(i) > \sigma(j)$. For instance, the inversions of the permutation $(3, 1, 2)$ (again, shown here in one-line notation) in S_3 are $(1, 2)$ and $(1, 3)$ (because $3 > 1$ and $3 > 2$), while the only inversion of the permutation $(1, 3, 2)$ in S_3 is $(2, 3)$ (since $3 > 2$).

(b) The *length* of σ means the number of inversions of σ. This length is denoted by $\ell(\sigma)$; it is a nonnegative integer.

If $n \in \mathbb{N}$, then any $\sigma \in S_n$ satisfies $0 \leq \ell(\sigma) \leq \binom{n}{2}$ (since the number of inversions of σ is clearly no larger than the total number of pairs (i, j) of integers satisfying $1 \leq i < j \leq n$; but the latter number is $\binom{n}{2}$). The only permutation in S_n having length 0 is the identity permutation $\operatorname{id} = (1, 2, \ldots, n) \in S_n$ [151].

Exercise 5.2. Let $n \in \mathbb{N}$.

(a) Show that every permutation $\sigma \in S_n$ and every $k \in \{1, 2, \ldots, n-1\}$ satisfy

$$\ell(\sigma \circ s_k) = \begin{cases} \ell(\sigma) + 1, & \text{if } \sigma(k) < \sigma(k+1); \\ \ell(\sigma) - 1, & \text{if } \sigma(k) > \sigma(k+1) \end{cases} \tag{293}$$

and

$$\ell(s_k \circ \sigma) = \begin{cases} \ell(\sigma) + 1, & \text{if } \sigma^{-1}(k) < \sigma^{-1}(k+1); \\ \ell(\sigma) - 1, & \text{if } \sigma^{-1}(k) > \sigma^{-1}(k+1) \end{cases}. \tag{294}$$

(b) Show that any two permutations σ and τ in S_n satisfy $\ell(\sigma \circ \tau) \equiv \ell(\sigma) + \ell(\tau) \bmod 2$.

(c) Show that any two permutations σ and τ in S_n satisfy $\ell(\sigma \circ \tau) \leq \ell(\sigma) + \ell(\tau)$.

(d) If $\sigma \in S_n$ is a permutation satisfying $\sigma(1) \leq \sigma(2) \leq \cdots \leq \sigma(n)$, then show that $\sigma = \operatorname{id}$.

(e) Let $\sigma \in S_n$. Show that σ can be written as a composition of $\ell(\sigma)$ permutations of the form s_k (with $k \in \{1, 2, \ldots, n-1\}$).

(f) Let $\sigma \in S_n$. Then, show that $\ell(\sigma) = \ell(\sigma^{-1})$.

(g) Let $\sigma \in S_n$. Show that $\ell(\sigma)$ is the smallest $N \in \mathbb{N}$ such that σ can be written as a composition of N permutations of the form s_k (with $k \in \{1, 2, \ldots, n-1\}$).

Example 5.11. Let us justify Exercise 5.2 **(a)** on an example. The solution to Exercise 5.2 **(a)** given below is essentially a (tiresome) formalization of the ideas seen in this example.

Let $n = 5$, $k = 3$ and $\sigma = (4, 2, 1, 5, 3)$ (written in one-line notation). Then, $\sigma \circ s_k = (4, 2, 5, 1, 3)$; this is the permutation obtained by swapping the k-th and

[151] The fact that the identity permutation $\operatorname{id} \in S_n$ has length $\ell(\operatorname{id}) = 0$ is trivial. The fact that it is the only one such permutation is easy (it essentially follows from Exercise 5.2 **(d)**).

the $(k+1)$-th entry of σ (where the word "entry" refers to the one-line notation). On the other hand, $s_k \circ \sigma = (3,2,1,5,4)$; this is the permutation obtained by swapping the entry k with the entry $k+1$ of σ. Mind the difference between these two operations.

The inversions of $\sigma = (4,2,1,5,3)$ are $(1,2)$, $(1,3)$, $(1,5)$, $(2,3)$ and $(4,5)$. These are the pairs (i,j) of positions such that i is before j (that is, $i < j$) but the i-th entry of σ is larger than the j-th entry of σ (that is, $\sigma(i) > \sigma(j)$). In other words, these are the pairs of positions at which the entries of σ are out of order. On the other hand, the inversions of $s_k \circ \sigma = (3,2,1,5,4)$ are $(1,2)$, $(1,3)$, $(2,3)$ and $(4,5)$. These are precisely the inversions of σ except for $(1,5)$. This is no surprise: In fact, $s_k \circ \sigma$ is obtained from σ by swapping the entry k with the entry $k+1$, and this operation clearly preserves all inversions other than the one that is directly being turned around (i.e., the inversion (i,j) where $\{\sigma(i), \sigma(j)\} = \{k, k+1\}$; in our case, this is the inversion $(1,5)$). In general, when $\sigma^{-1}(k) > \sigma^{-1}(k+1)$ (that is, when k appears further left than $k+1$ in the one-line notation of σ), the inversions of $s_k \circ \sigma$ are the inversions of σ except for $\left(\sigma^{-1}(k+1), \sigma^{-1}(k)\right)$. Therefore, in this case, the number of inversions of $s_k \circ \sigma$ equals the number of inversions of σ plus 1. That is, in this case, $\ell(s_k \circ \sigma) = \ell(\sigma) + 1$. When $\sigma^{-1}(k) < \sigma^{-1}(k+1)$, a similar argument shows $\ell(s_k \circ \sigma) = \ell(\sigma) - 1$. This explains why (294) holds (although formalizing this argument will be tedious).

The inversions of $\sigma \circ s_k = (4,2,5,1,3)$ are $(1,2)$, $(1,4)$, $(1,5)$, $(2,4)$, $(3,4)$ and $(3,5)$. Unlike the inversions of $s_k \circ \sigma$, these are not directly related to the inversions of σ, so the argument in the previous paragraph does not prove (293). However, instead of considering inversions of σ, one can consider inversions of σ^{-1}. These are even more intuitive: They are the pairs of integers (i,j) with $1 \leq i < j \leq n$ such that i appears further right than j in the one-line notation of σ. For instance, the inversions of σ^{-1} are $(1,2)$, $(1,4)$, $(2,4)$, $(3,4)$ and $(3,5)$, whereas the inversions of $(\sigma \circ s_k)^{-1}$ are all of these and also $(1,5)$. But there is no need to repeat our proof of (294); it is easier to deduce (293) from (294) by applying (294) to σ^{-1} instead of σ and appealing to Exercise 5.2 **(f)**. (Again, see the solution below for the details.)

Notice that Exercise 5.2 **(e)** immediately yields Exercise 5.1 **(b)**.

Remark 5.12. When $n = 0$ or $n = 1$, we have $\{1, 2, \ldots, n-1\} = \varnothing$. Hence, Exercise 5.1 **(e)** looks strange in the case when $n = 0$ or $n = 1$, because in this case, there are no permutations of the form s_k to begin with. Nevertheless, it is correct. Indeed, when $n = 0$ or $n = 1$, there is only one permutation $\sigma \in S_n$, namely the identity permutation id, and it has length $\ell(\sigma) = \ell(\text{id}) = 0$. Thus, in this case, Exercise 5.1 **(e)** claims that id can be written as a composition of 0 permutations of the form s_k (with $k \in \{1, 2, \ldots, n-1\}$). This is true: Even from an empty set we can always pick 0 elements; and the composition of 0 permutations will be id.

Remark 5.13. The word "length" for $\ell(\sigma)$ can be confusing: It does not refer to the length of the n-tuple $(\sigma(1), \sigma(2), \ldots, \sigma(n))$ (which is n). The reason why it is called "length" is Exercise 5.2 **(g)**: it says that $\ell(\sigma)$ is the smallest number of permutations of the form s_k which can be multiplied to give σ; thus, it is the smallest possible length of an expression of σ as a product of s_k's.

The use of the word "length", unfortunately, is not standard across literature. Some authors call "Coxeter length" what we call "length", and use the word "length" itself for a different notion.

Exercise 5.3. Let $n \in \mathbb{N}$. Let $\sigma \in S_n$. In Exercise 5.1 **(b)**, we have seen that σ can be written as a composition of several permutations of the form s_k (with $k \in \{1, 2, \ldots, n-1\}$). Usually there will be several ways to do so (for instance, $\mathrm{id} = s_1 \circ s_1 = s_2 \circ s_2 = \cdots = s_{n-1} \circ s_{n-1}$). Show that, whichever of these ways we take, the number of permutations composed will be congruent to $\ell(\sigma)$ modulo 2.

5.3. The sign of a permutation

Definition 5.14. Let $n \in \mathbb{N}$.

(a) We define the *sign* of a permutation $\sigma \in S_n$ as the integer $(-1)^{\ell(\sigma)}$. We denote this sign by $(-1)^{\sigma}$ or $\operatorname{sign} \sigma$ or $\operatorname{sgn} \sigma$.

(b) We say that a permutation σ is *even* if its sign is 1 (that is, if $\ell(\sigma)$ is even), and *odd* if its sign is -1 (that is, if $\ell(\sigma)$ is odd).

Signs of permutations have the following properties:

Proposition 5.15. Let $n \in \mathbb{N}$.

(a) The sign of the identity permutation $\mathrm{id} \in S_n$ is $(-1)^{\mathrm{id}} = 1$. In other words, $\mathrm{id} \in S_n$ is even.

(b) For every $k \in \{1, 2, \ldots, n-1\}$, the sign of the permutation $s_k \in S_n$ is $(-1)^{s_k} = -1$.

(c) If σ and τ are two permutations in S_n, then $(-1)^{\sigma \circ \tau} = (-1)^{\sigma} \cdot (-1)^{\tau}$.

(d) If $\sigma \in S_n$, then $(-1)^{\sigma^{-1}} = (-1)^{\sigma}$. (Here and in the following, the expression "$(-1)^{\sigma^{-1}}$" should be read as "$(-1)^{\left(\sigma^{-1}\right)}$", not as "$\left((-1)^{\sigma}\right)^{-1}$"; this is similar to Convention 2.58 (although σ is not a number).)

Proof of Proposition 5.15. **(a)** The identity permutation id satisfies $\ell(\mathrm{id}) = 0$ [152]. Now, the definition of $(-1)^{\mathrm{id}}$ yields $(-1)^{\mathrm{id}} = (-1)^{\ell(\mathrm{id})} = 1$ (since $\ell(\mathrm{id}) = 0$). In other words, $\mathrm{id} \in S_n$ is even. This proves Proposition 5.15 **(a)**.

[152] *Proof.* An inversion of id is the same as a pair (i, j) of integers satisfying $1 \leq i < j \leq n$ and $\mathrm{id}(i) > \mathrm{id}(j)$ (by the definition of "inversion"). Thus, if (i, j) is an inversion of id, then $1 \leq i < j \leq n$ and $\mathrm{id}(i) > \mathrm{id}(j)$; but this leads to a contradiction (since $\mathrm{id}(i) > \mathrm{id}(j)$ contradicts $\mathrm{id}(i) = i < j = \mathrm{id}(j)$). Hence, we obtain a contradiction for each inversion of id. Thus, there are no inversions of id. But $\ell(\mathrm{id})$ is defined as the number of inversions of id. Hence, $\ell(\mathrm{id}) = (\text{the number of inversions of id}) = 0$ (since there are no inversions of id).

(b) Let $k \in \{1, 2, \ldots, n-1\}$. Applying (293) to $\sigma = \operatorname{id}$, we obtain

$$\begin{aligned} \ell(\operatorname{id} \circ s_k) &= \begin{cases} \ell(\operatorname{id}) + 1, & \text{if } \operatorname{id}(k) < \operatorname{id}(k+1); \\ \ell(\operatorname{id}) - 1, & \text{if } \operatorname{id}(k) > \operatorname{id}(k+1) \end{cases} \\ &= \underbrace{\ell(\operatorname{id})}_{=0} + 1 \qquad (\text{since } \operatorname{id}(k) = k < k+1 = \operatorname{id}(k+1)) \\ &= 1. \end{aligned}$$

This rewrites as $\ell(s_k) = 1$ (since $\operatorname{id} \circ s_k = s_k$). Now, the definition of $(-1)^{s_k}$ yields $(-1)^{s_k} = (-1)^{\ell(s_k)} = -1$ (since $\ell(s_k) = 1$). This proves Proposition 5.15 **(b)**.

(c) Let $\sigma \in S_n$ and $\tau \in S_n$. Exercise 5.2 **(b)** yields $\ell(\sigma \circ \tau) \equiv \ell(\sigma) + \ell(\tau) \bmod 2$, so that $(-1)^{\ell(\sigma\circ\tau)} = (-1)^{\ell(\sigma)+\ell(\tau)} = (-1)^{\ell(\sigma)} \cdot (-1)^{\ell(\tau)}$. But the definition of the sign of a permutation yields $(-1)^{\sigma\circ\tau} = (-1)^{\ell(\sigma\circ\tau)}$ and $(-1)^{\sigma} = (-1)^{\ell(\sigma)}$ and $(-1)^{\tau} = (-1)^{\ell(\tau)}$. Hence, $(-1)^{\sigma\circ\tau} = (-1)^{\ell(\sigma\circ\tau)} = \underbrace{(-1)^{\ell(\sigma)}}_{=(-1)^{\sigma}} \cdot \underbrace{(-1)^{\ell(\tau)}}_{=(-1)^{\tau}} = (-1)^{\sigma} \cdot (-1)^{\tau}$. This proves Proposition 5.15 **(c)**.

(d) Let $\sigma \in S_n$. The definition of $(-1)^{\sigma^{-1}}$ yields $(-1)^{\sigma^{-1}} = (-1)^{\ell(\sigma^{-1})}$. But Exercise 5.2 **(f)** says that $\ell(\sigma) = \ell(\sigma^{-1})$. The definition of $(-1)^{\sigma}$ yields $(-1)^{\sigma} = (-1)^{\ell(\sigma)} = (-1)^{\ell(\sigma^{-1})}$ (since $\ell(\sigma) = \ell(\sigma^{-1})$). Compared with $(-1)^{\sigma^{-1}} = (-1)^{\ell(\sigma^{-1})}$, this yields $(-1)^{\sigma^{-1}} = (-1)^{\sigma}$. This proves Proposition 5.15 **(d)**. □

If you are familiar with some basic concepts of abstract algebra, then you will immediately notice that parts **(a)** and **(c)** of Proposition 5.15 can be summarized as the statement that "sign is a group homomorphism from the group S_n to the multiplicative group $\{1, -1\}$". In this statement, "sign" means the map from S_n to $\{1, -1\}$ which sends every permutation σ to its sign $(-1)^{\sigma} = (-1)^{\ell(\sigma)}$, and the "multiplicative group $\{1, -1\}$" means the group $\{1, -1\}$ whose binary operation is multiplication.

We have defined the sign of a permutation $\sigma \in S_n$. More generally, it is possible to define the sign of a permutation of an arbitrary finite set X, even though the length of such a permutation is not defined![153]

[153] How does it work? If X is a finite set, then we can always find a bijection $\phi : X \to \{1, 2, \ldots, n\}$ for some $n \in \mathbb{N}$. (Constructing such a bijection is tantamount to writing down a list of all elements of X, with no duplicates.) Given such a bijection ϕ, we can define the sign of any permutation σ of X as follows:

$$(-1)^{\sigma} = (-1)^{\phi\circ\sigma\circ\phi^{-1}}. \qquad (295)$$

Here, the right hand side is well-defined because $\phi \circ \sigma \circ \phi^{-1}$ is a permutation of $\{1, 2, \ldots, n\}$. What is not immediately obvious is that this sign is independent on the choice of ϕ, and that it is a group homomorphism to $\{1, -1\}$ (that is, we have $(-1)^{\operatorname{id}} = 1$ and $(-1)^{\sigma\circ\tau} = (-1)^{\sigma} \cdot (-1)^{\tau}$). We will prove these facts further below (in Exercise 5.12).

Exercise 5.4. Let $n \geq 2$. Show that the number of even permutations in S_n is $n!/2$, and the number of odd permutations in S_n is also $n!/2$.

The sign of a permutation is used in the combinatorial definition of the determinant. Let us briefly show this definition now; we shall return to it later (in Chapter 6) to study it in much more detail.

Definition 5.16. Let $n \in \mathbb{N}$. Let $A = \left(a_{i,j}\right)_{1\leq i\leq n,\ 1\leq j\leq n}$ be an $n \times n$-matrix (say, with complex entries, although this does not matter much – it suffices that the entries can be added and multiplied and the axioms of associativity, distributivity, commutativity, unity etc. hold). The *determinant* $\det A$ of A is defined as

$$\sum_{\sigma \in S_n} (-1)^{\sigma} a_{1,\sigma(1)} a_{2,\sigma(2)} \cdots a_{n,\sigma(n)}. \tag{296}$$

Let me try to describe the sum (296) in slightly more visual terms: The sum (296) has $n!$ addends, each of which has the form "$(-1)^{\sigma}$ times a product". The product has n factors, which are entries of A, and are chosen in such a way that there is exactly one entry taken from each row and exactly one from each column. Which precise entries are taken depends on σ: namely, for each i, we take the $\sigma(i)$-th entry from the i-th row.

Convince yourself that the classical formulas

$$\det\left(\begin{array}{c} a \end{array}\right) = a;$$

$$\det\left(\begin{array}{cc} a & b \\ c & d \end{array}\right) = ad - bc;$$

$$\det\left(\begin{array}{ccc} a & b & c \\ d & e & f \\ g & h & i \end{array}\right) = aei + bfg + cdh - ahf - bdi - ceg$$

are particular cases of (296). Whenever $n \geq 2$, the sum in (296) contains precisely $n!/2$ plus signs and $n!/2$ minus signs (because of Exercise 5.4).

Definition 5.16 is merely one of several equivalent definitions of the determinant. You will probably see two of them in an average linear algebra class. Each of them has its own advantages and drawbacks. Definition 5.16 is the most direct, assuming that one knows about the sign of a permutation.

5.4. Infinite permutations

(This section is optional; it explores some technical material which is useful in combinatorics, but is not necessary for what follows. I advise the reader to skip it at the first read.)

We have introduced the notion of a permutation of an arbitrary set; but so far, we have only studied permutations of finite sets. In this section (which is tangential to our project; probably nothing from this section will be used ever after), let me discuss permutations of the infinite set $\{1,2,3,\ldots\}$. (A lot of what I say below can be easily adapted to the sets $\mathbb{N}$ and $\mathbb{Z}$ as well.)

We recall that a permutation of a set X means a bijection from X to X.

Let S_∞ be the symmetric group of the set $\{1,2,3,\ldots\}$. This is the set of all permutations of $\{1,2,3,\ldots\}$. It contains the identity permutation $\operatorname{id} \in S_\infty$ which sends every $i \in \{1,2,3,\ldots\}$ to i. The set S_∞ is uncountable[154].

We shall try to study S_∞ similarly to how we studied S_n for $n \in \mathbb{N}$. However, we soon will notice that the analogy between S_∞ and S_n will break down.[155] To amend this, we shall define a subset $S_{(\infty)}$ of S_∞ (mind the parentheses around the "∞") which is smaller and more wieldy, and indeed shares many of the properties of the finite symmetric group S_n.

We define $S_{(\infty)}$ as follows:

$$S_{(\infty)} = \{\sigma \in S_\infty \mid \sigma(i) = i \text{ for all but finitely many } i \in \{1,2,3,\ldots\}\}. \qquad (297)$$

Let us first explain what "all but finitely many $i \in \{1,2,3,\ldots\}$" means:

Definition 5.17. Let I be a set. Let $\mathcal{A}(i)$ be a statement for every $i \in I$. Then, we say that "$\mathcal{A}(i)$ for all but finitely many $i \in I$" if and only if there exists some finite subset J of I such that every $i \in I \setminus J$ satisfies $\mathcal{A}(i)$. [156]

Thus, for a permutation $\sigma \in S_\infty$, we have the following equivalence of statements:

$$\begin{aligned}
&(\sigma(i) = i \text{ for all but finitely many } i \in \{1,2,3,\ldots\}) \\
\iff &(\text{there exists some finite subset } J \text{ of } \{1,2,3,\ldots\} \text{ such that} \\
&\quad \text{every } i \in \{1,2,3,\ldots\} \setminus J \text{ satisfies } \sigma(i) = i) \\
\iff &(\text{there exists some finite subset } J \text{ of } \{1,2,3,\ldots\} \text{ such that} \\
&\quad \text{the only } i \in \{1,2,3,\ldots\} \text{ that satisfy } \sigma(i) \neq i \text{ are elements of } J) \\
\iff &(\text{the set of all } i \in \{1,2,3,\ldots\} \text{ that satisfy } \sigma(i) \neq i \text{ is} \\
&\quad \text{contained in some finite subset } J \text{ of } \{1,2,3,\ldots\}) \\
\iff &(\text{there are only finitely many } i \in \{1,2,3,\ldots\} \text{ that satisfy } \sigma(i) \neq i).
\end{aligned}$$

[154] More generally, while a finite set of size n has $n!$ permutations, an infinite set X has uncountably many permutations (even if X is countable).

[155] The uncountability of S_∞ is the first hint that S_∞ is "too large" a set to be a good analogue of the finite set S_n.

[156] Thus, the statement "$\mathcal{A}(i)$ for all but finitely many $i \in I$" can be restated as "$\mathcal{A}(i)$ holds for all $i \in I$, apart from finitely many exceptions" or as "there are only finitely many $i \in I$ which do not satisfy $\mathcal{A}(i)$". I prefer the first wording, because it makes the most sense in constructive logic.

Caution: Do not confuse the words "all but finitely many $i \in I$" in this definition with the words "infinitely many $i \in I$". For instance, it is true that n is even for infinitely many $n \in \mathbb{Z}$, but it is not true that n is even for all but finitely many $n \in \mathbb{Z}$. Conversely, it is true that $n > 1$ for all but finitely many $n \in \{1,2\}$ (because the only $n \in \{1,2\}$ which does not satisfy $n > 1$ is 1), but it is not true that $n > 1$ for infinitely many $n \in \{1,2\}$ (because there are no infinitely many $n \in \{1,2\}$ to begin with).

You will encounter the "all but finitely many" formulation often in abstract algebra. (Some people abbreviate it as "almost all", but this abbreviation means other things as well.)

Hence, (297) rewrites as follows:

$$S_{(\infty)} = \{\sigma \in S_\infty \mid \text{ there are only finitely many } i \in \{1,2,3,\ldots\} \text{ that satisfy } \sigma(i) \neq i\}.$$

Example 5.18. Here is an example of a permutation which is in S_∞ but not in $S_{(\infty)}$: Let τ be the permutation of $\{1,2,3,\ldots\}$ given by

$$(\tau(1), \tau(2), \tau(3), \tau(4), \tau(5), \tau(6), \ldots) = (2,1,4,3,6,5,\ldots).$$

(It adds 1 to every odd positive integer, and subtracts 1 from every even positive integer.) Then, $\tau \in S_\infty$ but $\tau \notin S_{(\infty)}$.

On the other hand, let us show some examples of permutations in $S_{(\infty)}$. For each $i \in \{1,2,3,\ldots\}$, let s_i be the permutation in S_∞ that swaps i with $i+1$ but leaves all other numbers unchanged. (This is similar to the permutation s_i in S_n that was defined earlier. We have taken the liberty to re-use the name s_i, hoping that no confusion will arise.)

Again, we have $s_i^2 = \text{id}$ for every $i \in \{1,2,3,\ldots\}$ (where α^2 means $\alpha \circ \alpha$ for any $\alpha \in S_\infty$).

Proposition 5.19. We have $s_k \in S_{(\infty)}$ for every $k \in \{1,2,3,\ldots\}$.

Proof of Proposition 5.19. Let $k \in \{1,2,3,\ldots\}$. The permutation s_k has been defined as the permutation in S_∞ that swaps k with $k+1$ but leaves all other numbers unchanged. In other words, it satisfies $s_k(k) = k+1$, $s_k(k+1) = k$ and

$$s_k(i) = i \qquad \text{for every } i \in \{1,2,3,\ldots\} \text{ such that } i \notin \{k,k+1\}. \tag{298}$$

Now, every $i \in \{1,2,3,\ldots\} \setminus \{k,k+1\}$ satisfies $s_k(i) = i$ [157]. Hence, there exists some finite subset J of $\{1,2,3,\ldots\}$ such that every $i \in \{1,2,3,\ldots\} \setminus J$ satisfies $s_k(i) = i$ (namely, $J = \{k,k+1\}$). In other words, $s_k(i) = i$ for all but finitely many $i \in \{1,2,3,\ldots\}$.

Thus, s_k is an element of S_∞ satisfying $s_k(i) = i$ for all but finitely many $i \in \{1,2,3,\ldots\}$. Hence,

$$s_k \in \{\sigma \in S_\infty \mid \sigma(i) = i \text{ for all but finitely many } i \in \{1,2,3,\ldots\}\} = S_{(\infty)}.$$

This proves Proposition 5.19. □

Permutations can be composed and inverted, leading to new permutations. Let us first see that the same is true for elements of $S_{(\infty)}$:

[157] *Proof.* Let $i \in \{1,2,3,\ldots\} \setminus \{k,k+1\}$. Thus, $i \in \{1,2,3,\ldots\}$ and $i \notin \{k,k+1\}$. Hence, (298) shows that $s_k(i) = i$, qed.

Proposition 5.20. **(a)** The identity permutation $\operatorname{id} \in S_\infty$ of $\{1, 2, 3, \ldots\}$ satisfies $\operatorname{id} \in S_{(\infty)}$.

(b) For every $\sigma \in S_{(\infty)}$ and $\tau \in S_{(\infty)}$, we have $\sigma \circ \tau \in S_{(\infty)}$.

(c) For every $\sigma \in S_{(\infty)}$, we have $\sigma^{-1} \in S_{(\infty)}$.

Proof of Proposition 5.20. We have defined $S_{(\infty)}$ as the set of all $\sigma \in S_\infty$ such that $\sigma(i) = i$ for all but finitely many $i \in \{1, 2, 3, \ldots\}$. In other words, $S_{(\infty)}$ is the set of all $\sigma \in S_\infty$ such that there exists a finite subset K of $\{1, 2, 3, \ldots\}$ such that (every $i \in \{1, 2, 3, \ldots\} \setminus K$ satisfies $\sigma(i) = i$). As a consequence, we have the following two facts:

- If K is a finite subset of $\{1, 2, 3, \ldots\}$, and if $\gamma \in S_\infty$ is a permutation such that
$$\left(\text{every } i \in \{1, 2, 3, \ldots\} \setminus K \text{ satisfies } \gamma(i) = i\right), \tag{299}$$
then
$$\gamma \in S_{(\infty)}. \tag{300}$$

- If $\gamma \in S_{(\infty)}$, then
$$\left(\begin{array}{c} \text{there exists some finite subset } K \text{ of } \{1, 2, 3, \ldots\} \\ \text{such that every } i \in \{1, 2, 3, \ldots\} \setminus K \text{ satisfies } \gamma(i) = i \end{array}\right). \tag{301}$$

We can now step to the actual proof of Proposition 5.20.

(a) Every $i \in \{1, 2, 3, \ldots\} \setminus \varnothing$ satisfies $\operatorname{id}(i) = i$. Thus, (300) (applied to $K = \varnothing$ and $\gamma = \operatorname{id}$) yields $\operatorname{id} \in S_{(\infty)}$. This proves Proposition 5.20 **(a)**.

(b) Let $\sigma \in S_{(\infty)}$ and $\tau \in S_{(\infty)}$.

From (301) (applied to $\gamma = \sigma$), we conclude that there exists some finite subset K of $\{1, 2, 3, \ldots\}$ such that every $i \in \{1, 2, 3, \ldots\} \setminus K$ satisfies $\sigma(i) = i$. Let us denote this K by J_1. Thus, J_1 is a finite subset of $\{1, 2, 3, \ldots\}$, and
$$\text{every } i \in \{1, 2, 3, \ldots\} \setminus J_1 \text{ satisfies } \sigma(i) = i. \tag{302}$$

From (301) (applied to $\gamma = \tau$), we conclude that there exists some finite subset K of $\{1, 2, 3, \ldots\}$ such that every $i \in \{1, 2, 3, \ldots\} \setminus K$ satisfies $\tau(i) = i$. Let us denote this K by J_2. Thus, J_2 is a finite subset of $\{1, 2, 3, \ldots\}$, and
$$\text{every } i \in \{1, 2, 3, \ldots\} \setminus J_2 \text{ satisfies } \tau(i) = i. \tag{303}$$

The sets J_1 and J_2 are finite. Hence, their union $J_1 \cup J_2$ is finite. Moreover,
$$\text{every } i \in \{1, 2, 3, \ldots\} \setminus (J_1 \cup J_2) \text{ satisfies } (\sigma \circ \tau)(i) = i$$

[158]. Therefore, (300) (applied to $K = J_1 \cup J_2$ and $\gamma = \sigma \circ \tau$) yields $\sigma \circ \tau \in S_{(\infty)}$. This proves Proposition 5.20 **(b)**.

(c) Let $\sigma \in S_{(\infty)}$.

From (301) (applied to $\gamma = \sigma$), we conclude that there exists some finite subset K of $\{1, 2, 3, \ldots\}$ such that every $i \in \{1, 2, 3, \ldots\} \setminus K$ satisfies $\sigma(i) = i$. Consider this K. Thus, K is a finite subset of $\{1, 2, 3, \ldots\}$, and

$$\text{every } i \in \{1, 2, 3, \ldots\} \setminus K \text{ satisfies } \sigma(i) = i. \tag{304}$$

Now,

$$\text{every } i \in \{1, 2, 3, \ldots\} \setminus K \text{ satisfies } \sigma^{-1}(i) = i$$

[159]. Therefore, (300) (applied to $\gamma = \sigma^{-1}$) yields $\sigma^{-1} \in S_{(\infty)}$. This proves Proposition 5.20 **(c)**. □

In the language of group theorists, Proposition 5.20 show that $S_{(\infty)}$ is a subgroup of the group S_∞. The elements of $S_{(\infty)}$ are called the *finitary permutations of* $\{1, 2, 3, \ldots\}$, and $S_{(\infty)}$ is called the *finitary symmetric group of* $\{1, 2, 3, \ldots\}$.

We now have the following analogue of Exercise 5.1 (without its part **(c)**):

Exercise 5.5. **(a)** Show that $s_i \circ s_{i+1} \circ s_i = s_{i+1} \circ s_i \circ s_{i+1}$ for all $i \in \{1, 2, 3, \ldots\}$.

(b) Show that every permutation $\sigma \in S_{(\infty)}$ can be written as a composition of several permutations of the form s_k (with $k \in \{1, 2, 3, \ldots\}$).

Remark 5.21. In the language of group theory, the statement of Exercise 5.5 **(b)** says (or, more precisely, yields) that the permutations $s_1, s_2, s_3, \ldots$ generate the group $S_{(\infty)}$.

If $\sigma \in S_\infty$ is a permutation, then an *inversion* of σ means a pair (i, j) of integers satisfying $1 \leq i < j$ and $\sigma(i) > \sigma(j)$. This definition of an inversion is, of course, analogous to the definition of an inversion of a $\sigma \in S_n$; thus we could try to define the length of a $\sigma \in S_\infty$. However, here we run into troubles: A permutation $\sigma \in S_\infty$ might have infinitely many inversions!

It is here that we really need to restrict ourselves to $S_{(\infty)}$. This indeed helps:

Proposition 5.22. Let $\sigma \in S_{(\infty)}$. Then:

(a) There exists some $N \in \{1, 2, 3, \ldots\}$ such that every integer $i > N$ satisfies $\sigma(i) = i$.

(b) There are only finitely many inversions of σ.

[158] *Proof.* Let $i \in \{1, 2, 3, \ldots\} \setminus (J_1 \cup J_2)$. Thus, $i \in \{1, 2, 3, \ldots\}$ and $i \notin J_1 \cup J_2$.

We have $i \notin J_1 \cup J_2$ and thus $i \notin J_1$ (since $J_1 \subseteq J_1 \cup J_2$). Hence, $i \in \{1, 2, 3, \ldots\} \setminus J_1$. Similarly, $i \in \{1, 2, 3, \ldots\} \setminus J_2$. Thus, (303) yields $\tau(i) = i$. Hence, $(\sigma \circ \tau)(i) = \sigma\left(\underbrace{\tau(i)}_{=i}\right) = \sigma(i) = i$ (by (302)), qed.

[159] *Proof.* Let $i \in \{1, 2, 3, \ldots\} \setminus K$. Thus, $\sigma(i) = i$ (according to (304)), so that $\sigma^{-1}(i) = i$, qed.

Proof of Proposition 5.22. **(a)** We can apply (301) to $\gamma = \sigma$. As a consequence, we obtain that there exists some finite subset K of $\{1, 2, 3, \ldots\}$ such that

$$\text{every } i \in \{1, 2, 3, \ldots\} \setminus K \text{ satisfies } \sigma(i) = i. \qquad (305)$$

Consider this K.

The set K is finite. Hence, the set $K \cup \{1\}$ is finite; this set is also nonempty (since it contains 1) and a subset of $\{1, 2, 3, \ldots\}$. Therefore, this set $K \cup \{1\}$ has a greatest element (since every finite nonempty subset of $\{1, 2, 3, \ldots\}$ has a greatest element). Let n be this greatest element. Clearly, $n \in K \cup \{1\} \subseteq \{1, 2, 3, \ldots\}$, so that $n > 0$.

Every $j \in K \cup \{1\}$ satisfies

$$j \leq n \qquad (306)$$

(since n is the greatest element of $K \cup \{1\}$). Now, let i be an integer such that $i > n$. Then, $i > n > 0$, so that i is a positive integer. If we had $i \in K$, then we would have $i \in K \subseteq K \cup \{1\}$ and thus $i \leq n$ (by (306), applied to $j = i$), which would contradict $i > n$. Hence, we cannot have $i \in K$. We thus have $i \notin K$. Since $i \in \{1, 2, 3, \ldots\}$, this shows that $i \in \{1, 2, 3, \ldots\} \setminus K$. Thus, $\sigma(i) = i$ (by (305)).

Let us now forget that we fixed i. We thus have shown that every integer $i > n$ satisfies $\sigma(i) = i$. Hence, Proposition 5.22 **(a)** holds (we can take $N = n$).

(b) Proposition 5.22 **(a)** shows that there exists some $N \in \{1, 2, 3, \ldots\}$ such that

$$\text{every integer } i > N \text{ satisfies } \sigma(i) = i. \qquad (307)$$

Consider such an N. We shall now show that

$$\text{every inversion of } \sigma \text{ is an element of } \{1, 2, \ldots, N\}^2.$$

In fact, let c be an inversion of σ. We will show that c is an element of $\{1, 2, \ldots, N\}^2$.

We know that c is an inversion of σ. In other words, c is a pair (i, j) of integers satisfying $1 \leq i < j$ and $\sigma(i) > \sigma(j)$ (by the definition of an "inversion of σ"). Consider this (i, j). We then have $i \leq N$ [160] and $j \leq N$ [161]. Consequently, $(i, j) \in \{1, 2, \ldots, N\}^2$. Hence, $c = (i, j) \in \{1, 2, \ldots, N\}^2$.

Now, let us forget that we fixed c. We thus have shown that if c is an inversion of σ, then c is an element of $\{1, 2, \ldots, N\}^2$. In other words, every inversion of σ is an element of $\{1, 2, \ldots, N\}^2$. Thus, there are only finitely many inversions of σ (since there are only finitely many elements of $\{1, 2, \ldots, N\}^2$). Proposition 5.22 **(b)** is thus proven. □

[160] *Proof.* Assume the contrary. Thus, $i > N$. Hence, (307) shows that $\sigma(i) = i$. Also, $i < j$, so that $j > i > N$. Hence, (307) (applied to j instead of i) shows that $\sigma(j) = j$. Thus, $\sigma(i) = i < j = \sigma(j)$. This contradicts $\sigma(i) > \sigma(j)$. This contradiction shows that our assumption was wrong, qed.

[161] *Proof.* Assume the contrary. Thus, $j > N$. Hence, (307) (applied to j instead of i) shows that $\sigma(j) = j$. Now, $\sigma(i) > \sigma(j) = j > N$. Therefore, (307) (applied to $\sigma(i)$ instead of i) yields $\sigma(\sigma(i)) = \sigma(i)$. But σ is a permutation, and thus an injective map. Hence, from $\sigma(\sigma(i)) = \sigma(i)$, we obtain $\sigma(i) = i$. Thus, $\sigma(i) = i < j = \sigma(j)$. This contradicts $\sigma(i) > \sigma(j)$. This contradiction shows that our assumption was wrong, qed.

Actually, Proposition 5.22 **(b)** has a converse: If a permutation $\sigma \in S_\infty$ has only finitely many inversions, then σ belongs to $S_{(\infty)}$. This is easy to prove; but we won't use this.

If $\sigma \in S_{(\infty)}$ is a permutation, then the *length* of σ means the number of inversions of σ. This is well-defined, because there are only finitely many inversions of σ (by Proposition 5.22 **(b)**). The length of σ is denoted by $\ell(\sigma)$; it is a nonnegative integer. The only permutation having length 0 is the identity permutation $\operatorname{id} \in S_\infty$.

We have the following analogue of Exercise 5.2:

Exercise 5.6. (a) Show that every permutation $\sigma \in S_{(\infty)}$ and every $k \in \{1, 2, 3, \ldots\}$ satisfy

$$\ell(\sigma \circ s_k) = \begin{cases} \ell(\sigma) + 1, & \text{if } \sigma(k) < \sigma(k+1); \\ \ell(\sigma) - 1, & \text{if } \sigma(k) > \sigma(k+1) \end{cases} \tag{308}$$

and

$$\ell(s_k \circ \sigma) = \begin{cases} \ell(\sigma) + 1, & \text{if } \sigma^{-1}(k) < \sigma^{-1}(k+1); \\ \ell(\sigma) - 1, & \text{if } \sigma^{-1}(k) > \sigma^{-1}(k+1) \end{cases}. \tag{309}$$

(b) Show that any two permutations σ and τ in $S_{(\infty)}$ satisfy $\ell(\sigma \circ \tau) \equiv \ell(\sigma) + \ell(\tau) \bmod 2$.

(c) Show that any two permutations σ and τ in $S_{(\infty)}$ satisfy $\ell(\sigma \circ \tau) \leq \ell(\sigma) + \ell(\tau)$.

(d) If $\sigma \in S_{(\infty)}$ is a permutation satisfying $\sigma(1) \leq \sigma(2) \leq \sigma(3) \leq \cdots$, then show that $\sigma = \operatorname{id}$.

(e) Let $\sigma \in S_{(\infty)}$. Show that σ can be written as a composition of $\ell(\sigma)$ permutations of the form s_k (with $k \in \{1, 2, 3, \ldots\}$).

(f) Let $\sigma \in S_{(\infty)}$. Then, show that $\ell(\sigma) = \ell(\sigma^{-1})$.

(g) Let $\sigma \in S_{(\infty)}$. Show that $\ell(\sigma)$ is the smallest $N \in \mathbb{N}$ such that σ can be written as a composition of N permutations of the form s_k (with $k \in \{1, 2, 3, \ldots\}$).

We also have an analogue of Exercise 5.3:

Exercise 5.7. Let $\sigma \in S_{(\infty)}$. In Exercise 5.5 **(b)**, we have seen that σ can be written as a composition of several permutations of the form s_k (with $k \in \{1, 2, 3, \ldots\}$). Usually there will be several ways to do so (for instance, $\operatorname{id} = s_1 \circ s_1 = s_2 \circ s_2 = s_3 \circ s_3 = \cdots$). Show that, whichever of these ways we take, the number of permutations composed will be congruent to $\ell(\sigma)$ modulo 2.

Having defined the length of a permutation $\sigma \in S_{(\infty)}$, we can now define the sign of such a permutation. Again, we mimic the definition of the sign of a $\sigma \in S_n$:

Definition 5.23. We define the *sign* of a permutation $\sigma \in S_{(\infty)}$ as the integer $(-1)^{\ell(\sigma)}$. We denote this sign by $(-1)^\sigma$ or $\operatorname{sign} \sigma$ or $\operatorname{sgn} \sigma$. We say that a permutation σ is *even* if its sign is 1 (that is, if $\ell(\sigma)$ is even), and *odd* if its sign is -1 (that is, if $\ell(\sigma)$ is odd).

Signs of permutations have the following properties:

Proposition 5.24. **(a)** The sign of the identity permutation $\operatorname{id} \in S_{(\infty)}$ is $(-1)^{\operatorname{id}} = 1$. In other words, $\operatorname{id} \in S_{(\infty)}$ is even.

(b) For every $k \in \{1,2,3,\ldots\}$, the sign of the permutation $s_k \in S_{(\infty)}$ is $(-1)^{s_k} = -1$.

(c) If σ and τ are two permutations in $S_{(\infty)}$, then $(-1)^{\sigma \circ \tau} = (-1)^{\sigma} \cdot (-1)^{\tau}$.

(d) If $\sigma \in S_{(\infty)}$, then $(-1)^{\sigma^{-1}} = (-1)^{\sigma}$.

The proof of Proposition 5.24 is analogous to the proof of Proposition 5.15.

Remark 5.25. We have defined the sign of a permutation $\sigma \in S_{(\infty)}$. No such notion exists for permutations $\sigma \in S_\infty$. In fact, one can show that if an element λ_σ of $\{1,-1\}$ is chosen for each $\sigma \in S_\infty$ in such a way that every two permutations $\sigma, \tau \in S_\infty$ satisfy $\lambda_{\sigma \circ \tau} = \lambda_\sigma \cdot \lambda_\tau$, then all of the λ_σ are 1. (Indeed, this follows from a result of Oystein Ore; see `http://mathoverflow.net/questions/54371` .)

Remark 5.26. For every $n \in \mathbb{N}$ and every $\sigma \in S_n$, we can define a permutation $\sigma_{(\infty)} \in S_{(\infty)}$ by setting

$$\sigma_{(\infty)}(i) = \begin{cases} \sigma(i), & \text{if } i \leq n; \\ i, & \text{if } i > n \end{cases} \qquad \text{for all } i \in \{1,2,3,\ldots\}.$$

Essentially, $\sigma_{(\infty)}$ is the permutation σ extended to the set of all positive integers in the laziest possible way: It just sends each $i > n$ to itself.

For every $n \in \mathbb{N}$, there is an injective map $\mathbf{i}_n : S_n \to S_{(\infty)}$ defined as follows:

$$\mathbf{i}_n(\sigma) = \sigma_{(\infty)} \qquad \text{for every } \sigma \in S_n.$$

This map $\mathbf{i}_n$ is an example of what algebraists call a *group homomorphism*: It satisfies

$$\begin{aligned} \mathbf{i}_n(\operatorname{id}) &= \operatorname{id}; \\ \mathbf{i}_n(\sigma \circ \tau) &= \mathbf{i}_n(\sigma) \circ \mathbf{i}_n(\tau) \qquad \text{for all } \sigma, \tau \in S_n; \\ \mathbf{i}_n\left(\sigma^{-1}\right) &= (\mathbf{i}_n(\sigma))^{-1} \qquad \text{for all } \sigma \in S_n. \end{aligned}$$

(This is all easy to check.) Thus, we can consider the image $\mathbf{i}_n(S_n)$ of S_n under this map as a "copy" of S_n which is "just as good as the original" (i.e., composition in this copy behaves in the same way as composition in the original). It is easy to characterize this copy inside $S_{(\infty)}$: Namely,

$$\mathbf{i}_n(S_n) = \left\{\sigma \in S_{(\infty)} \mid \sigma(i) = i \text{ for all } i > n\right\}.$$

Using Proposition 5.22 **(a)**, it is easy to check that $S_{(\infty)} = \bigcup_{n\in\mathbb{N}} \mathbf{i}_n(S_n) = \mathbf{i}_0(S_0) \cup \mathbf{i}_1(S_1) \cup \mathbf{i}_2(S_2) \cup \cdots$. Therefore, many properties of $S_{(\infty)}$ can be derived from analogous properties of S_n for finite n. For example, using this tactic, we could easily derive Exercise 5.6 from Exercise 5.2, and derive Exercise 5.7 from Exercise 5.3. (However, we can just as well solve Exercises 5.6 and 5.7 by copying the solutions of Exercises 5.2 and 5.3 and making the necessary changes; this is how I solve these exercises further below.)

5.5. More on lengths of permutations

Let us summarize some of what we have learnt about permutations. We have defined the length $\ell(\sigma)$ and the inversions of a permutation $\sigma \in S_n$, where n is a nonnegative integer. We recall the basic properties of these objects:

- For each $k \in \{1,2,\ldots,n-1\}$, we defined s_k to be the permutation in S_n that swaps k with $k+1$ but leaves all other numbers unchanged. These permutations satisfy $s_i^2 = \operatorname{id}$ for every $i \in \{1,2,\ldots,n-1\}$ and

$$s_i \circ s_{i+1} \circ s_i = s_{i+1} \circ s_i \circ s_{i+1} \qquad \text{for all } i \in \{1,2,\ldots,n-2\} \tag{310}$$

(according to Exercise 5.1 **(a)**). Also, it is easy to check that

$$s_i \circ s_j = s_j \circ s_i \qquad \text{for all } i,j \in \{1,2,\ldots,n-1\} \text{ with } |i-j| > 1. \tag{311}$$

- An *inversion* of a permutation $\sigma \in S_n$ means a pair (i,j) of integers satisfying $1 \leq i < j \leq n$ and $\sigma(i) > \sigma(j)$. The *length* $\ell(\sigma)$ of a permutation $\sigma \in S_n$ is the number of inversions of σ.

- Any two permutations $\sigma \in S_n$ and $\tau \in S_n$ satisfy

$$\ell(\sigma \circ \tau) \equiv \ell(\sigma) + \ell(\tau) \bmod 2 \tag{312}$$

(by Exercise 5.2 **(b)**) and

$$\ell(\sigma \circ \tau) \leq \ell(\sigma) + \ell(\tau) \tag{313}$$

(by Exercise 5.2 **(c)**).

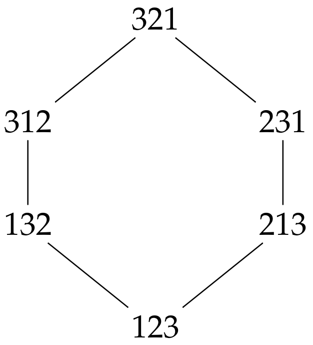
- If $\sigma \in S_n$, then $\ell(\sigma) = \ell\left(\sigma^{-1}\right)$ (according to Exercise 5.2 **(f)**).

- If $\sigma \in S_n$, then $\ell(\sigma)$ is the smallest $N \in \mathbb{N}$ such that σ can be written as a composition of N permutations of the form s_k (with $k \in \{1,2,\ldots,n-1\}$). (This follows from Exercise 5.2 **(g)**.)

By now, we know almost all about the s_k's and about the lengths of permutations that is necessary for studying determinants. ("Almost" because Exercise 5.10 below will also be useful.) I shall now sketch some more advanced properties of these things, partly as a curiosity, partly to further your intuition; none of these properties shall be used further below. The rest of Section 5.5 shall rely on some notions we have not introduced in these notes; in particular, we will use the concepts of undirected graphs ([LeLeMe16, Chapter 12]), directed graphs ([LeLeMe16, Chapter 10]) and (briefly) polytopes (see, e.g., [AigZie14, Chapter 10]).

First, here is a way to visualize lengths of permutations using graph theory:

Fix $n \in \mathbb{N}$. We define the *n-th right Bruhat graph* to be the (undirected) graph whose vertices are the permutations $\sigma \in S_n$, and whose edges are defined as follows: Two vertices $\sigma \in S_n$ and $\tau \in S_n$ are adjacent if and only if there exists a $k \in \{1, 2, \ldots, n-1\}$ such that $\sigma = \tau \circ s_k$. (This condition is clearly symmetric in σ and τ: If $\sigma = \tau \circ s_k$, then $\tau = \sigma \circ s_k$.) For instance, the 3-rd right Bruhat graph looks as follows:



,

where we are writing permutations in one-line notation (and omitting parentheses and commas). The 4-th right Bruhat graph can be seen on Wikipedia.[162]

These graphs have lots of properties. There is a canonical way to direct their edges: The edge between σ and τ is directed towards the vertex with the larger length. (The lengths of σ and τ always differ by 1 if there is an edge between σ and τ.) This way, the n-th right Bruhat graph is an acyclic directed graph. It therefore defines a partially ordered set, called the *right permutohedron order*[163] on S_n, whose elements are the permutations $\sigma \in S_n$ and whose order relation is defined as follows: We have $\sigma \leq \tau$ if and only if there is a directed path from σ to τ in the directed n-th right Bruhat graph. If you know the (combinatorial) notion of a lattice, you might notice that this right permutohedron order is a lattice.

The word "permutohedron" in "permutohedron order" hints at what might be its least expected property: The n-th Bruhat graph can be viewed as the graph formed by the vertices and the edges of a certain polytope in n-dimensional space $\mathbb{R}^n$. This polytope – called the *n-th permutohedron*[164] – is the convex hull of the

[162] Don't omit the word "right" in "right Bruhat graph"; else it means a different graph with more edges.

[163] also known as the *right weak order* or *right weak Bruhat order* (but, again, do not omit the words "right" and "weak")

[164] Some spell it "*permutahedron*" instead. The word is of relatively recent origin (1969).

points $\left(\sigma\left(1\right),\sigma\left(2\right),\ldots,\sigma\left(n\right)\right)$ for $\sigma \in S_n$. These points are its vertices; however, its vertex $\left(\sigma\left(1\right),\sigma\left(2\right),\ldots,\sigma\left(n\right)\right)$ corresponds to the vertex σ^{-1} (not σ) of the n-th Bruhat graph. Feel free to roam its Wikipedia page for other (combinatorial and geometric) curiosities.

The notion of a length fits perfectly into this whole picture. For instance, the length $\ell\left(\sigma\right)$ of a permutation σ is the smallest length of a path from $\operatorname{id} \in S_n$ to σ on the n-th right Bruhat graph (and this holds no matter whether the graph is considered to be directed or not). For the undirected Bruhat graphs, we have something more general:

Exercise 5.8. Let $n \in \mathbb{N}$. Let $\sigma \in S_n$ and $\tau \in S_n$. Show that $\ell\left(\sigma^{-1} \circ \tau\right)$ is the smallest length of a path between σ and τ on the (undirected) n-th right Bruhat graph.

(Recall that the length of a path in a graph is defined as the number of edges along this path.)

How many permutations in S_n have a given length? The number is not easy to compute directly; however, its generating function is nice. (See [LeLeMe16, Chapter 16] for the notion of generating functions.) Namely,

$$\sum_{w \in S_n} q^{\ell(w)} = \left(1+q\right)\left(1+q+q^2\right)\cdots\left(1+q+q^2+\cdots+q^{n-1}\right)$$

(where q is an indeterminate). This equality (with q renamed as x) is Corollary 5.53 (which is proven below, in the solution to Exercise 5.18). Another proof can be found in [Stanle11, Corollary 1.3.13] (but notice that Stanley denotes S_n by $\mathfrak{S}_n$, and $\ell\left(w\right)$ by $\operatorname{inv}\left(w\right)$).

Remark 5.27. Much more can be said. Let me briefly mention (without proof) two other related results.

We can ask ourselves in what different ways a permutation can be written as a composition of N permutations of the form s_k. For instance, the permutation $w_0 \in S_3$ which sends 1, 2 and 3 to 3, 2 and 1, respectively (that is, $w_0 = (3,2,1)$ in one-line notation) can be written as a product of three s_k's in the two forms

$$w_0 = s_1 \circ s_2 \circ s_1, \qquad w_0 = s_2 \circ s_1 \circ s_2, \tag{314}$$

but can also be written as a product of five s_k's (e.g., as $w_0 = s_1 \circ s_2 \circ s_1 \circ s_2 \circ s_2$) or seven s_k's or nine s_k's, etc. Are the different representations of w_0 related?

Clearly, the two representations in (314) are connected to each other by the equality $s_1 \circ s_2 \circ s_1 = s_2 \circ s_1 \circ s_2$, which is a particular case of (310). Also, the representation $w_0 = s_1 \circ s_2 \circ s_1 \circ s_2 \circ s_2$ reduces to $w_0 = s_1 \circ s_2 \circ s_1$ by "cancelling" the two adjacent s_2's at the end (recall that $s_i \circ s_i = s_i^2 = \operatorname{id}$ for every i).

Interestingly, this generalizes. Let $n \in \mathbb{N}$ and $\sigma \in S_n$. A *reduced expression* for σ will mean a representation of σ as a composition of $\ell\left(\sigma\right)$ permutations of the form s_k. (As we know, less than $\ell(\sigma)$ such permutations do not suffice; thus the name "reduced".) Then, (one of the many versions of) *Matsumoto's theorem* states that any two reduced expressions of σ can be obtained from one another by a rewriting process, each step of which is either an application of (310) (i.e., you pick an "$s_i \circ s_{i+1} \circ s_i$" in the expression and replace it by "$s_{i+1} \circ s_i \circ s_{i+1}$", or vice versa) or an application of (311) (i.e., you pick an "$s_i \circ s_j$" with $|i-j| > 1$ and replace it by "$s_j \circ s_i$", or vice versa). For instance, for $n = 4$ and $\sigma = (4,3,1,2,5)$ (in one-line notation), the two reduced expressions $\sigma = s_1 \circ s_2 \circ s_3 \circ s_1 \circ s_2$ and $\sigma = s_2 \circ s_3 \circ s_1 \circ s_2 \circ s_3$ can be obtained from one another by the following rewriting process:

$$\begin{aligned} s_1 \circ s_2 \circ \underbrace{s_3 \circ s_1}_{\substack{=s_1 \circ s_3 \\ \text{(by (311))}}} \circ s_2 &= \underbrace{s_1 \circ s_2 \circ s_1}_{\substack{=s_2 \circ s_1 \circ s_2 \\ \text{(by (310))}}} \circ s_3 \circ s_2 = s_2 \circ s_1 \circ \underbrace{s_2 \circ s_3 \circ s_2}_{\substack{=s_3 \circ s_2 \circ s_3 \\ \text{(by (310))}}} \\ &= s_2 \circ \underbrace{s_1 \circ s_3}_{\substack{=s_3 \circ s_1 \\ \text{(by (311))}}} \circ s_2 \circ s_3 = s_2 \circ s_3 \circ s_1 \circ s_2 \circ s_3. \end{aligned}$$

See, e.g., Williamson's thesis [Willia03, Corollary 1.2.3] or Knutson's notes [Knutso12, §2.3] for a proof of this fact. (Knutson, instead of saying that "$\sigma = s_{k_1} \circ s_{k_2} \circ \cdots \circ s_{k_p}$ is a reduced expression for σ", says that "$k_1 k_2 \cdots k_p$ is a reduced word for σ".)

Something subtler holds for "non-reduced" expressions. Namely, if we have a representation of σ as a composition of some number of permutations of the form s_k (not necessarily $\ell(\sigma)$ of them), then we can transform it into a reduced expression by a rewriting process which consists of applications of (310) and (311) as before and also of cancellation steps (i.e., picking an "$s_i \circ s_i$" in the expression and removing it). This follows from [LLPT95, Chapter SYM, Proposition (2.6)][165], and can also easily be derived from [Willia03, Corollary 1.2.3 and Corollary 1.1.6].

This all is stated and proven in greater generality in good books on Coxeter groups, such as [BjoBre05]. We won't need these results in the following, but they are an example of what one can see if one looks at permutations closely.

5.6. More on signs of permutations

In Section 5.3, we have defined the sign $(-1)^{\sigma} = \operatorname{sign} \sigma = \operatorname{sgn} \sigma$ of a permutation σ. We recall the most important facts about it:

[165] What the authors of [LLPT95] call a "presentation" of a permutation $\sigma \in S_n$ is a finite list $\left(s_{k_1}, s_{k_2}, \ldots, s_{k_p}\right)$ of elements of $\{s_1, s_2, \ldots, s_{n-1}\}$ satisfying $\sigma = s_{k_1} \circ s_{k_2} \circ \cdots \circ s_{k_p}$. What the authors of [LLPT95] call a "minimal presentation" of σ is what we call a reduced expression of σ.

- We have $(-1)^{\sigma} = (-1)^{\ell(\sigma)}$ for every $\sigma \in S_n$. (This is the definition of $(-1)^{\sigma}$.) Thus, for every $\sigma \in S_n$, we have $(-1)^{\sigma} = (-1)^{\ell(\sigma)} \in \{1, -1\}$.
- The permutation $\sigma \in S_n$ is said to be *even* if $(-1)^{\sigma} = 1$, and is said to be *odd* if $(-1)^{\sigma} = -1$.
- The sign of the identity permutation $\mathrm{id} \in S_n$ is $(-1)^{\mathrm{id}} = 1$.
- For every $k \in \{1, 2, \ldots, n-1\}$, the sign of the permutation $s_k \in S_n$ is $(-1)^{s_k} = -1$.
- If σ and τ are two permutations in S_n, then
$$(-1)^{\sigma \circ \tau} = (-1)^{\sigma} \cdot (-1)^{\tau} . \tag{315}$$
- If $\sigma \in S_n$, then
$$(-1)^{\sigma^{-1}} = (-1)^{\sigma} . \tag{316}$$

A simple consequence of the above facts is the following proposition:

Proposition 5.28. Let $n \in \mathbb{N}$ and $k \in \mathbb{N}$. Let $\sigma_1, \sigma_2, \ldots, \sigma_k$ be k permutations in S_n. Then,
$$(-1)^{\sigma_1 \circ \sigma_2 \circ \cdots \circ \sigma_k} = (-1)^{\sigma_1} \cdot (-1)^{\sigma_2} \cdot \cdots \cdot (-1)^{\sigma_k} . \tag{317}$$

Proof of Proposition 5.28. Straightforward induction over k. The induction base (i.e., the case when $k = 0$) follows from the fact that $(-1)^{\mathrm{id}} = 1$ (since the composition of 0 permutations is id). The induction step is easily done using (315). □

Let us introduce another notation:

Definition 5.29. Let $n \in \mathbb{N}$. Let i and j be two distinct elements of $\{1, 2, \ldots, n\}$. We let $t_{i,j}$ be the permutation in S_n which swaps i with j while leaving all other elements of $\{1, 2, \ldots, n\}$ unchanged. Such a permutation is called a *transposition* (and is often denoted by (i, j) in literature; but we prefer not to do so, since it is too similar to one-line notation).

Notice that the permutations $s_1, s_2, \ldots, s_{n-1}$ are transpositions (namely, $s_i = t_{i,i+1}$ for every $i \in \{1, 2, \ldots, n-1\}$), but they are not the only transpositions (when $n \geq 3$).

For the next exercise, we need one further definition, which extends Definition 5.29:

Definition 5.30. Let $n \in \mathbb{N}$. Let i and j be two elements of $\{1, 2, \ldots, n\}$. We define a permutation $t_{i,j} \in S_n$ as follows:

- If $i \neq j$, then the permutation $t_{i,j}$ has already been defined in Definition 5.29.
- If $i = j$, then we define the permutation $t_{i,j}$ to be the identity $\mathrm{id} \in S_n$.

Exercise 5.9. Whenever m is an integer, we shall use the notation $[m]$ for the set $\{1, 2, \ldots, m\}$.

Let $n \in \mathbb{N}$. Let $\sigma \in S_n$. Prove that there is a unique n-tuple $(i_1, i_2, \ldots, i_n) \in [1] \times [2] \times \cdots \times [n]$ such that

$$\sigma = t_{1,i_1} \circ t_{2,i_2} \circ \cdots \circ t_{n,i_n}.$$

Example 5.31. For this example, set $n = 4$, and let $\sigma \in S_4$ be the permutation that sends $1, 2, 3, 4$ to $3, 1, 4, 2$. Then, Exercise 5.9 claims that there is a unique 4-tuple $(i_1, i_2, i_3, i_4) \in [1] \times [2] \times [3] \times [4]$ such that $\sigma = t_{1,i_1} \circ t_{2,i_2} \circ t_{3,i_3} \circ t_{4,i_4}$.

This 4-tuple can easily be found: it is $(1, 1, 1, 3)$. In fact, we have $\sigma = t_{1,1} \circ t_{2,1} \circ t_{3,1} \circ t_{4,3}$.

Exercise 5.10. Let $n \in \mathbb{N}$. Let i and j be two distinct elements of $\{1, 2, \ldots, n\}$.

(a) Find $\ell\left(t_{i,j}\right)$.

(b) Show that $(-1)^{t_{i,j}} = -1$.

Exercise 5.11. Let $n \in \mathbb{N}$. Let w_0 denote the permutation in S_n which sends each $k \in \{1, 2, \ldots, n\}$ to $n + 1 - k$. Compute $\ell(w_0)$ and $(-1)^{w_0}$.

Exercise 5.12. Let X be a finite set. We want to define the sign of any permutation of X. (We have sketched this definition before (see (295)), but now we shall do it in detail.)

Fix a bijection $\phi : X \to \{1, 2, \ldots, n\}$ for some $n \in \mathbb{N}$. (Such a bijection always exists. Indeed, constructing such a bijection is tantamount to writing down a list of all elements of X, with no duplicates.) For every permutation σ of X, set

$$(-1)_{\phi}^{\sigma} = (-1)^{\phi \circ \sigma \circ \phi^{-1}}.$$

Here, the right hand side is well-defined because $\phi \circ \sigma \circ \phi^{-1}$ is a permutation of $\{1, 2, \ldots, n\}$.

(a) Prove that $(-1)_{\phi}^{\sigma}$ depends only on the permutation σ of X, but not on the bijection ϕ. (In other words, for a given σ, any two different choices of ϕ will lead to the same $(-1)_{\phi}^{\sigma}$.)

This allows us to define the *sign* of the permutation σ to be $(-1)_{\phi}^{\sigma}$ for any choice of bijection $\phi : X \to \{1, 2, \ldots, n\}$. We denote this sign simply by $(-1)^{\sigma}$. (When $X = \{1, 2, \ldots, n\}$, then this sign is clearly the same as the sign $(-1)^{\sigma}$ we defined before, because we can pick the bijection $\phi = \mathrm{id}$.)

(b) Show that the permutation $\mathrm{id} : X \to X$ satisfies $(-1)^{\mathrm{id}} = 1$.

(c) Show that $(-1)^{\sigma \circ \tau} = (-1)^{\sigma} \cdot (-1)^{\tau}$ for any two permutations σ and τ of X.

Remark 5.32. A sufficiently pedantic reader might have noticed that the definition of $(-1)^{\sigma}$ in Exercise 5.12 is not completely kosher. In fact, the set X may be $\{1, 2, \ldots, n\}$ for some $n \in \mathbb{N}$; in this case, σ is an element of S_n, and thus the sign $(-1)^{\sigma}$ has already been defined in Definition 5.14. Thus, in this case, we are defining the notation $(-1)^{\sigma}$ a second time in Exercise 5.12. Woe to us if this second definition yields a different number than the first!

Fortunately, it does not. The definition of $(-1)^{\sigma}$ in Exercise 5.12 does not conflict with the original meaning of $(-1)^{\sigma}$ as defined in Definition 5.14. Indeed, in order to prove this, we temporarily rename the number $(-1)^{\sigma}$ defined in Exercise 5.12 as $(-1)^{\sigma}_{\text{new}}$ (in order to ensure that we don't confuse it with the number $(-1)^{\sigma}$ defined in Definition 5.14). Now, consider the situation of Exercise 5.12, and assume that $X = \{1, 2, \ldots, n\}$. We must then prove that $(-1)^{\sigma}_{\text{new}} = (-1)^{\sigma}$. But the definition of $(-1)^{\sigma}_{\text{new}}$ in Exercise 5.12 says that $(-1)^{\sigma}_{\text{new}} = (-1)^{\sigma}_{\phi}$, where ϕ is any bijection $X \to \{1, 2, \ldots, n\}$. We can apply this to $\phi = \text{id}$ (because clearly, id is a bijection $X \to \{1, 2, \ldots, n\}$), and thus obtain $(-1)^{\sigma}_{\text{new}} = (-1)^{\sigma}_{\text{id}}$. But the definition of $(-1)^{\sigma}_{\text{id}}$ yields $(-1)^{\sigma}_{\text{id}} = (-1)^{\text{id} \circ \sigma \circ \text{id}^{-1}} = (-1)^{\sigma}$ (since $\text{id} \circ \sigma \circ \underbrace{\text{id}^{-1}}_{=\text{id}} = \sigma$). Thus, $(-1)^{\sigma}_{\text{new}} = (-1)^{\sigma}_{\text{id}} = (-1)^{\sigma}$. This is precisely what we wanted to prove. Thus, we have shown that the definition of $(-1)^{\sigma}$ in Exercise 5.12 does not conflict with the original meaning of $(-1)^{\sigma}$ as defined in Definition 5.14.

Remark 5.33. Let $n \in \mathbb{N}$. Recall that a *transposition* in S_n means a permutation of the form $t_{i,j}$, where i and j are two distinct elements of $\{1, 2, \ldots, n\}$. Therefore, if τ is a transposition in S_n, then

$$(-1)^{\tau} = -1. \tag{318}$$

(In fact, if τ is a transposition in S_n, then τ can be written in the form $\tau = t_{i,j}$ for two distinct elements i and j of $\{1, 2, \ldots, n\}$; and therefore, for these two elements i and j, we have $(-1)^{\tau} = (-1)^{t_{i,j}} = -1$ (according to Exercise 5.10 **(b)**). This proves (318).)

Now, let $\sigma \in S_n$ be any permutation. Assume that σ is written in the form $\sigma = \tau_1 \circ \tau_2 \circ \cdots \circ \tau_k$ for some transpositions $\tau_1, \tau_2, \ldots, \tau_k$ in S_n. Then,

$$\begin{aligned} (-1)^{\sigma} = (-1)^{\tau_1 \circ \tau_2 \circ \cdots \circ \tau_k} &= \underbrace{(-1)^{\tau_1}}_{\substack{=-1 \\ (\text{by } (318))}} \cdot \underbrace{(-1)^{\tau_2}}_{\substack{=-1 \\ (\text{by } (318))}} \cdot \cdots \cdot \underbrace{(-1)^{\tau_k}}_{\substack{=-1 \\ (\text{by } (318))}} \\ &\qquad (\text{by (317), applied to } \sigma_i = \tau_i) \\ &= \underbrace{(-1) \cdot (-1) \cdot \cdots \cdot (-1)}_{k \text{ factors}} = (-1)^k . \end{aligned} \tag{319}$$

Since many permutations can be written as products of transpositions in a simple way, this formula gives a useful method for computing signs.

Remark 5.34. Let $n \in \mathbb{N}$. It is not hard to prove that

$$(-1)^{\sigma} = \prod_{1 \leq i < j \leq n} \frac{\sigma(i) - \sigma(j)}{i - j} \qquad \text{for every } \sigma \in S_n. \tag{320}$$

(Of course, it is no easier to compute $(-1)^{\sigma}$ using this seemingly explicit formula than by counting inversions.)

We shall prove (320) in Exercise 5.13 **(c)**.

Remark 5.35. The sign of a permutation is also called its *signum* or its *signature*. Different authors define the sign of a permutation σ in different ways. Some (e.g., Hefferon in [Heffer20, Chapter Four, Definition 4.7], or Strickland in [Strick13, Definition B.4]) define it as we do; others (e.g., Conrad in [Conrad1], or Máté in [Mate14], or Hoffman and Kunze in [HofKun71, p. 152]) define it using (319); yet others define it using something called the *cycle decomposition* of a permutation; some even define it using (320), or using a similar ratio of two polynomials. However, it is not hard to check that all of these definitions are equivalent. (We already know that the first two of them are equivalent.)

Exercise 5.13. Let $n \in \mathbb{N}$. Let $\sigma \in S_n$.

(a) If $x_1, x_2, \ldots, x_n$ are n elements of $\mathbb{C}$, then prove that

$$\prod_{1 \leq i < j \leq n} \left(x_{\sigma(i)} - x_{\sigma(j)} \right) = (-1)^{\sigma} \cdot \prod_{1 \leq i < j \leq n} \left(x_i - x_j \right).$$

(b) More generally: For every $(i, j) \in \{1, 2, \ldots, n\}^2$, let $a_{(i,j)}$ be an element of $\mathbb{C}$. Assume that

$$a_{(j,i)} = -a_{(i,j)} \qquad \text{for every } (i, j) \in \{1, 2, \ldots, n\}^2. \tag{321}$$

Prove that

$$\prod_{1 \leq i < j \leq n} a_{(\sigma(i), \sigma(j))} = (-1)^{\sigma} \cdot \prod_{1 \leq i < j \leq n} a_{(i,j)}.$$

(c) Prove (320).

(d) Use Exercise 5.13 **(a)** to give a new solution to Exercise 5.2 **(b)**.

The next exercise relies on the notion of "the list of all elements of S in increasing order (with no repetitions)", where S is a finite set of integers. This notion means exactly what it says; it was rigorously defined in Definition 2.50.

Exercise 5.14. Let $n \in \mathbb{N}$. Let I be a subset of $\{1, 2, \ldots, n\}$. Let $k = |I|$. Let $(a_1, a_2, \ldots, a_k)$ be the list of all elements of I in increasing order (with no repetitions). Let $(b_1, b_2, \ldots, b_{n-k})$ be the list of all elements of $\{1, 2, \ldots, n\} \setminus I$ in increasing order (with no repetitions). Let $\alpha \in S_k$ and $\beta \in S_{n-k}$. Prove the following:

(a) There exists a unique $\sigma \in S_n$ satisfying

$$(\sigma(1), \sigma(2), \ldots, \sigma(n)) = \left(a_{\alpha(1)}, a_{\alpha(2)}, \ldots, a_{\alpha(k)}, b_{\beta(1)}, b_{\beta(2)}, \ldots, b_{\beta(n-k)}\right).$$

Denote this σ by $\sigma_{I,\alpha,\beta}$.

(b) Let $\sum I$ denote the sum of all elements of I. (Thus, $\sum I = \sum\limits_{i \in I} i$.) We have

$$\ell\left(\sigma_{I,\alpha,\beta}\right) = \ell(\alpha) + \ell(\beta) + \sum I - (1 + 2 + \cdots + k)$$

and

$$(-1)^{\sigma_{I,\alpha,\beta}} = (-1)^{\alpha} \cdot (-1)^{\beta} \cdot (-1)^{\sum I - (1+2+\cdots+k)}.$$

(c) Forget that we fixed α and β. We thus have defined an element $\sigma_{I,\alpha,\beta} \in S_n$ for every $\alpha \in S_k$ and every $\beta \in S_{n-k}$. The map

$$S_k \times S_{n-k} \to \{\tau \in S_n \mid \tau(\{1, 2, \ldots, k\}) = I\},$$
$$(\alpha, \beta) \mapsto \sigma_{I,\alpha,\beta}$$

is well-defined and a bijection.

We can define transpositions not only in the symmetric group S_n, but also more generally for arbitrary sets X:

Definition 5.36. Let X be a set. Let i and j be two distinct elements of X. We let $t_{i,j}$ be the permutation of X which swaps i with j while leaving all other elements of X unchanged. Such a permutation is called a *transposition* of X.

Clearly, Definition 5.36 is a generalization of Definition 5.29.

Exercise 5.15. Let X be a finite set. Recall that if σ is any permutation of X, then the sign $(-1)^{\sigma}$ of σ is well-defined (by Exercise 5.12). Prove the following:

(a) For any two distinct elements i and j of X, we have $(-1)^{t_{i,j}} = -1$.

(b) Any permutation of X can be written as a composition of finitely many transpositions of X.

(c) Let σ be a permutation of X. Assume that σ can be written as a composition of k transpositions of X. Then, $(-1)^{\sigma} = (-1)^{k}$.

5.7. Cycles

Next, we shall discuss another specific class of permutations: the *cycles*.

Definition 5.37. Let $n \in \mathbb{N}$. Let $[n] = \{1, 2, \ldots, n\}$.

Let $k \in \{1, 2, \ldots, n\}$. Let $i_1, i_2, \ldots, i_k$ be k distinct elements of $[n]$. We define $\operatorname{cyc}_{i_1,i_2,\ldots,i_k}$ to be the permutation in S_n which sends $i_1, i_2, \ldots, i_k$ to $i_2, i_3, \ldots, i_k, i_1,$

respectively, while leaving all other elements of $[n]$ fixed. In other words, we define $\operatorname{cyc}_{i_1,i_2,\ldots,i_k}$ to be the permutation in S_n given by

$$\left(\operatorname{cyc}_{i_1,i_2,\ldots,i_k}(p) = \begin{cases} i_{j+1}, & \text{if } p = i_j \text{ for some } j \in \{1,2,\ldots,k\}; \\ p, & \text{otherwise} \end{cases} \qquad \text{for every } p \in [n] \right),$$

where i_{k+1} means i_1.

(Again, the notation $\operatorname{cyc}_{i_1,i_2,\ldots,i_k}$ conceals the parameter n, which will hopefully not cause any confusion.)

A permutation of the form $\operatorname{cyc}_{i_1,i_2,\ldots,i_k}$ is said to be a *k-cycle* (or sometimes just a *cycle*, or a *cyclic permutation*). Of course, the name stems from the fact that it "cycles" through the elements $i_1, i_2, \ldots, i_k$ (by sending each of them to the next one and the last one back to the first) and leaves all other elements unchanged.

Example 5.38. Let $n \in \mathbb{N}$. The following facts follow easily from Definition 5.37:

(a) For every $i \in \{1,2,\ldots,n\}$, we have $\operatorname{cyc}_i = \operatorname{id}$. In other words, any 1-cycle is the identity permutation id.

(b) If i and j are two distinct elements of $\{1,2,\ldots,n\}$, then $\operatorname{cyc}_{i,j} = t_{i,j}$. (See Definition 5.29 for the definition of $t_{i,j}$.)

(c) If $k \in \{1,2,\ldots,n-1\}$, then $\operatorname{cyc}_{k,k+1} = s_k$.

(d) If $n = 5$, then $\operatorname{cyc}_{2,5,3}$ is the permutation which sends 1 to 1, 2 to 5, 3 to 2, 4 to 4, and 5 to 3. (In other words, it is the permutation which is $(1,5,2,4,3)$ in one-line notation.)

(e) If $k \in \{1,2,\ldots,n\}$, and if $i_1, i_2, \ldots, i_k$ are k pairwise distinct elements of $[n]$, then

$$\operatorname{cyc}_{i_1,i_2,\ldots,i_k} = \operatorname{cyc}_{i_2,i_3,\ldots,i_k,i_1} = \operatorname{cyc}_{i_3,i_4,\ldots,i_k,i_1,i_2} = \cdots = \operatorname{cyc}_{i_k,i_1,i_2,\ldots,i_{k-1}}.$$

(In less formal words: The k-cycle $\operatorname{cyc}_{i_1,i_2,\ldots,i_k}$ does not change when we cyclically rotate the list $(i_1, i_2, \ldots, i_k)$.)

Remark 5.39. What we called $\operatorname{cyc}_{i_1,i_2,\ldots,i_k}$ in Definition 5.37 is commonly denoted by $(i_1, i_2, \ldots, i_k)$ in the literature. But this latter notation $(i_1, i_2, \ldots, i_k)$ would clash with one-line notation for permutations (the cycle $\operatorname{cyc}_{1,2,3} \in S_3$ is not the same as the permutation which is $(1,2,3)$ in one-line notation) and also with the standard notation for k-tuples. This is why we prefer to use the notation $\operatorname{cyc}_{i_1,i_2,\ldots,i_k}$. (That said, we are not going to use k-cycles very often.)

Any k-cycle is a composition of $k-1$ transpositions, as the following exercise shows:

Exercise 5.16. Let $n \in \mathbb{N}$. Let $[n] = \{1, 2, \ldots, n\}$. Let $k \in \{1, 2, \ldots, n\}$. Let $i_1, i_2, \ldots, i_k$ be k distinct elements of $[n]$. Prove that

$$\operatorname{cyc}_{i_1, i_2, \ldots, i_k} = t_{i_1, i_2} \circ t_{i_2, i_3} \circ \cdots \circ t_{i_{k-1}, i_k}.$$

(We are using Definition 5.29 here.)

The following exercise gathers some further properties of cycles. Parts **(a)** and **(d)** and, to a lesser extent, **(b)** are fairly important and you should make sure you know how to solve them. The significantly more difficult part **(c)** is more of a curiosity with an interesting proof (I have not found an application of it so far; skip it if you do not want to spend time on what is essentially a contest problem).

Exercise 5.17. Let $n \in \mathbb{N}$. Let $[n] = \{1, 2, \ldots, n\}$. Let $k \in \{1, 2, \ldots, n\}$.

(a) For every $\sigma \in S_n$ and every k distinct elements $i_1, i_2, \ldots, i_k$ of $[n]$, prove that

$$\sigma \circ \operatorname{cyc}_{i_1, i_2, \ldots, i_k} \circ \sigma^{-1} = \operatorname{cyc}_{\sigma(i_1), \sigma(i_2), \ldots, \sigma(i_k)}.$$

(b) For every $p \in \{0, 1, \ldots, n-k\}$, prove that

$$\ell\left(\operatorname{cyc}_{p+1, p+2, \ldots, p+k}\right) = k - 1.$$

(c) For every k distinct elements $i_1, i_2, \ldots, i_k$ of $[n]$, prove that

$$\ell\left(\operatorname{cyc}_{i_1, i_2, \ldots, i_k}\right) \geq k - 1.$$

(d) For every k distinct elements $i_1, i_2, \ldots, i_k$ of $[n]$, prove that

$$(-1)^{\operatorname{cyc}_{i_1, i_2, \ldots, i_k}} = (-1)^{k-1}.$$

Remark 5.40. Exercise 5.17 **(d)** shows that every k-cycle in S_n has sign $(-1)^{k-1}$. However, the length of a k-cycle need not be $k-1$. Exercise 5.17 **(c)** shows that this length is always $\geq k-1$, but it can take other values as well. For instance, in S_4, the length of the 3-cycle $\operatorname{cyc}_{1,4,3}$ is 4. (Another example are the transpositions $t_{i,j}$ from Definition 5.29; these are 2-cycles but can have length > 1.)

I don't know a simple way to describe when equality holds in Exercise 5.17 **(c)**. It holds whenever $i_1, i_2, \ldots, i_k$ are consecutive integers (due to Exercise 5.17 **(b)**), but also in some other cases; for example, the 4-cycle $\operatorname{cyc}_{1,3,4,2}$ in S_4 has length 3.

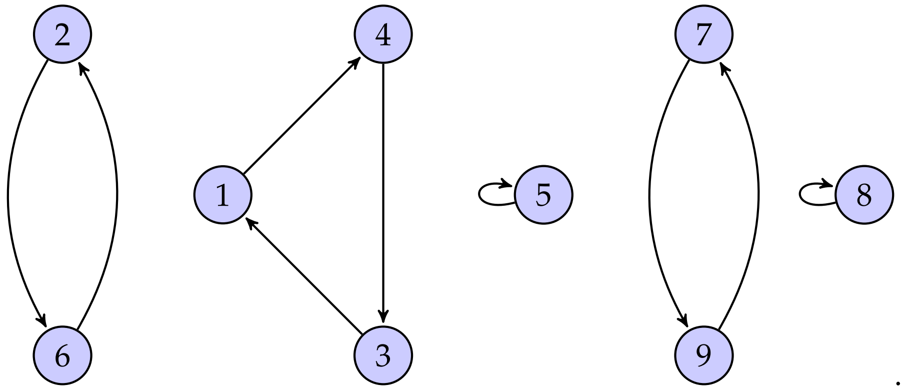
Remark 5.41. The main reason why cycles are useful is that, essentially, every permutation can be "decomposed" into cycles. We shall not use this fact, but since it is generally important, let us briefly explain what it means. (You will probably learn more about it in any standard course on abstract algebra.)

Fix $n \in \mathbb{N}$. Let $[n] = \{1, 2, \ldots, n\}$. Two cycles α and β in S_n are said to be *disjoint* if they can be written as $\alpha = \operatorname{cyc}_{i_1, i_2, \ldots, i_k}$ and $\beta = \operatorname{cyc}_{j_1, j_2, \ldots, j_\ell}$ for $k + \ell$ distinct elements $i_1, i_2, \ldots, i_k, j_1, j_2, \ldots, j_\ell$ of $[n]$. For example, the two cycles $\operatorname{cyc}_{1,3}$ and $\operatorname{cyc}_{2,6,7}$ in S_8 are disjoint, but the two cycles $\operatorname{cyc}_{1,4}$ and $\operatorname{cyc}_{2,4}$ are not. It is easy to see that any two disjoint cycles α and β commute (i.e., satisfy $\alpha \circ \beta = \beta \circ \alpha$). Therefore, when you see a composition $\alpha_1 \circ \alpha_2 \circ \cdots \circ \alpha_p$ of several pairwise disjoint cycles, you can reorder its factors arbitrarily without changing the result (for example, $\alpha_3 \circ \alpha_1 \circ \alpha_4 \circ \alpha_2 = \alpha_1 \circ \alpha_2 \circ \alpha_3 \circ \alpha_4$ if $p = 4$).

Now, the fact I am talking about says the following: Every permutation in S_n can be written as a composition of several pairwise disjoint cycles. For example, let $n = 9$, and let $\sigma \in S_9$ be the permutation which is written $(4, 6, 1, 3, 5, 2, 9, 8, 7)$ in one-line notation (i.e., we have $\sigma(1) = 4$, $\sigma(2) = 6$, etc.). Then, σ can be written as a composition of several pairwise disjoint cycles as follows:

$$\sigma = \operatorname{cyc}_{1,4,3} \circ \operatorname{cyc}_{7,9} \circ \operatorname{cyc}_{2,6}. \qquad (322)$$

Indeed, here is how such a decomposition can be found: Let us draw a directed graph whose vertices are $1, 2, \ldots, n$, and which has an arc $i \to \sigma(i)$ for every $i \in [n]$. (Thus, it has n arcs altogether; some of them can be loops.) For our permutation $\sigma \in S_9$, this graph looks as follows:



. (323)

Obviously, at each vertex i of this graph, exactly one arc begins (namely, the arc $i \to \sigma(i)$). Moreover, since σ is invertible, it is also clear that at each vertex i of this graph, exactly one arc ends (namely, the arc $\sigma^{-1}(i) \to i$). Due to the way we constructed this graph, it is clear that it completely describes our permutation σ: Namely, if we want to find $\sigma(i)$ for a given $i \in [n]$, we should just locate the vertex i on the graph, and follow the arc that begins at this vertex; the endpoint of this arc will be $\sigma(i)$.

Now, a look at this graph reveals five directed cycles (in the sense of "paths which end at the same vertex at which they begin", not yet in the sense of "cyclic permutations"). The first one passes through the vertices 2 and 6; the second passes through the vertices 3, 1 and 4; the third, through the vertex 5 (it is what is called a "trivial cycle"), and so on. To each of these cycles we can assign a cyclic permutation in S_n: namely, if the cycle passes through the vertices $i_1, i_2, \ldots, i_k$ (in this order, and with no repetitions), then we assign to it the cyclic permutation $\operatorname{cyc}_{i_1,i_2,\ldots,i_k} \in S_n$. The cyclic permutations assigned to all five directed cycles are pairwise disjoint, and their composition is

$$\operatorname{cyc}_{2,6} \circ \operatorname{cyc}_{3,1,4} \circ \operatorname{cyc}_5 \circ \operatorname{cyc}_{7,9} \circ \operatorname{cyc}_8 .$$

But this composition must be σ (because if we apply this composition to an element $i \in [n]$, then we obtain the "next vertex after i" on the directed cycle which passes through i; but due to how we constructed our graph, this "next vertex" will be precisely $\sigma(i)$). Hence, we have

$$\sigma = \operatorname{cyc}_{2,6} \circ \operatorname{cyc}_{3,1,4} \circ \operatorname{cyc}_5 \circ \operatorname{cyc}_{7,9} \circ \operatorname{cyc}_8 . \qquad (324)$$

Thus, we have found a way to write σ as a composition of several pairwise disjoint cycles. We can rewrite (and even simplify) this representation a bit: Namely, we can simplify (324) by removing the factors cyc_5 and cyc_8 (because both of these factors equal id); thus we obtain $\sigma = \operatorname{cyc}_{2,6} \circ \operatorname{cyc}_{3,1,4} \circ \operatorname{cyc}_{7,9}$. We can furthermore swap $\operatorname{cyc}_{2,6}$ with $\operatorname{cyc}_{3,1,4}$ (since disjoint cycles commute), therefore obtaining $\sigma = \operatorname{cyc}_{3,1,4} \circ \operatorname{cyc}_{2,6} \circ \operatorname{cyc}_{7,9}$. Next, we can swap $\operatorname{cyc}_{2,6}$ with $\operatorname{cyc}_{7,9}$, obtaining $\sigma = \operatorname{cyc}_{3,1,4} \circ \operatorname{cyc}_{7,9} \circ \operatorname{cyc}_{2,6}$. Finally, we can rewrite $\operatorname{cyc}_{3,1,4}$ as $\operatorname{cyc}_{1,4,3}$, and we obtain (322).

In general, for every $n \in \mathbb{N}$, every permutation $\sigma \in S_n$ can be represented as a composition of several pairwise disjoint cycles (which can be found by drawing a directed graph as in our example above). This representation is not literally unique, because we can modify it by:

- adding or removing trivial factors (i.e., factors of the form $\operatorname{cyc}_i = \operatorname{id}$);
- swapping different cycles;
- rewriting $\operatorname{cyc}_{i_1,i_2,\ldots,i_k}$ as $\operatorname{cyc}_{i_2,i_3,\ldots,i_k,i_1}$.

However, it is unique **up to these modifications**; in other words, any two representations of σ as a composition of several pairwise disjoint cycles can be transformed into one another by such modifications.

The proofs of all these statements are fairly easy. (One does have to check certain things, e.g., that the directed graph really consists of disjoint directed cycles. For a complete proof, see [Goodma15, Theorem 1.5.3] or [Bourba74, Chapter I, §5.7, Proposition 7] or [Sagan19, §1.9, proof of Theorem 1.5.1] or various other texts on algebra.)

Representing a permutation $\sigma \in S_n$ as a composition of several pairwise disjoint cycles can be done very quickly, and thus gives a quick way to find $(-1)^\sigma$ (because Exercise 5.17 **(d)** tells us how to find the sign of a k-cycle). This is significantly faster than counting inversions of σ.

5.8. The Lehmer code

In this short section, we shall introduce the *Lehmer code* of a permutation. Throughout Section 5.8, we will use the following notations:

Definition 5.42. **(a)** Whenever m is an integer, we shall use the notation $[m]$ for the set $\{1, 2, \ldots, m\}$. (This is an empty set when $m \leq 0$.)

(b) Whenever m is an integer, we shall use the notation $[m]_0$ for the set $\{0, 1, \ldots, m\}$. (This is an empty set when $m < 0$.)

Definition 5.43. Let $n \in \mathbb{N}$. We consider n to be fixed throughout Section 5.8.

Let H denote the set $[n-1]_0 \times [n-2]_0 \times \cdots \times [n-n]_0$.

Definition 5.44. Let $\sigma \in S_n$ and $i \in [n]$. Then, $\ell_i(\sigma)$ shall denote the number of all $j \in \{i+1, i+2, \ldots, n\}$ such that $\sigma(i) > \sigma(j)$.

Example 5.45. For this example, set $n = 5$, and let $\sigma \in S_5$ be the permutation that sends $1, 2, 3, 4, 5$ to $4, 3, 2, 1, 5$. Then, $\ell_2(\sigma)$ is the number of all $j \in \{3, 4, 5\}$ such that $\sigma(2) > \sigma(j)$. These j are 3 and 4 (because $\sigma(2) > \sigma(3)$ and $\sigma(2) > \sigma(4)$ but not $\sigma(2) > \sigma(5)$); therefore, $\ell_2(\sigma) = 2$. Similarly, $\ell_1(\sigma) = 3$, $\ell_3(\sigma) = 1$, $\ell_4(\sigma) = 0$ and $\ell_5(\sigma) = 0$.

The following two facts are almost trivial:[166]

Proposition 5.46. Let $\sigma \in S_n$. Then, $\ell(\sigma) = \ell_1(\sigma) + \ell_2(\sigma) + \cdots + \ell_n(\sigma)$.

Proposition 5.47. Let $\sigma \in S_n$. Then, $(\ell_1(\sigma), \ell_2(\sigma), \ldots, \ell_n(\sigma)) \in H$.

The following simple lemma gives two equivalent ways to define $\ell_i(\sigma)$ for $\sigma \in S_n$ and $i \in [n]$:

Lemma 5.48. Let $\sigma \in S_n$ and $i \in [n]$. Then:

(a) We have $\ell_i(\sigma) = |[\sigma(i) - 1] \setminus \sigma([i])|$.

(b) We have $\ell_i(\sigma) = |[\sigma(i) - 1] \setminus \sigma([i-1])|$.

(c) We have $\sigma(i) \leq i + \ell_i(\sigma)$.

Before we state the next proposition, we introduce another notation:

Definition 5.49. Let $(a_1, a_2, \ldots, a_n)$ and $(b_1, b_2, \ldots, b_n)$ be two n-tuples of integers. We say that $(a_1, a_2, \ldots, a_n) <_{\text{lex}} (b_1, b_2, \ldots, b_n)$ if and only if there exists some $k \in [n]$ such that $a_k \neq b_k$, and the **smallest** such k satisfies $a_k < b_k$.

[166] See Exercise 5.18 below for the proofs of all the following results.

For example, $(4,1,2,5) <_{\text{lex}} (4,1,3,0)$ and $(1,1,0,1) <_{\text{lex}} (2,0,0,0)$. The relation $<_{\text{lex}}$ is usually pronounced "is lexicographically smaller than"; the word "lexicographic" comes from the idea that if numbers were letters, then a "word" $a_1 a_2 \cdots a_n$ would appear earlier in a dictionary than $b_1 b_2 \cdots b_n$ if and only if $(a_1, a_2, \ldots, a_n) <_{\text{lex}} (b_1, b_2, \ldots, b_n)$.

Proposition 5.50. Let $\sigma \in S_n$ and $\tau \in S_n$ be such that

$$(\sigma(1), \sigma(2), \ldots, \sigma(n)) <_{\text{lex}} (\tau(1), \tau(2), \ldots, \tau(n)).$$

Then,

$$(\ell_1(\sigma), \ell_2(\sigma), \ldots, \ell_n(\sigma)) <_{\text{lex}} (\ell_1(\tau), \ell_2(\tau), \ldots, \ell_n(\tau)).$$

We can now define the Lehmer code:

Definition 5.51. Define the map $L : S_n \to H$ by

$$(L(\sigma) = (\ell_1(\sigma), \ell_2(\sigma), \ldots, \ell_n(\sigma)) \qquad \text{for each } \sigma \in S_n).$$

(This is well-defined because of Proposition 5.47.)

If $\sigma \in S_n$ is any permutation, then $L(\sigma) = (\ell_1(\sigma), \ell_2(\sigma), \ldots, \ell_n(\sigma))$ is called the *Lehmer code* of σ.

Theorem 5.52. The map $L : S_n \to H$ is a bijection.

Using this theorem and Proposition 5.46, we can easily show the following:

Corollary 5.53. We have

$$\sum_{w \in S_n} x^{\ell(w)} = (1+x)\left(1+x+x^2\right)\cdots\left(1+x+x^2+\cdots+x^{n-1}\right)$$

(an equality between polynomials in x). (The right hand side of this equality should be understood as the empty product when $n \leq 1$.)

Exercise 5.18. Prove Proposition 5.46, Proposition 5.47, Lemma 5.48, Proposition 5.50, Theorem 5.52 and Corollary 5.53.

See [Manive01, §2.1] and [Kerber99, §11.3] for further properties of permutations related to the Lehmer code. (In particular, [Manive01, proof of Proposition 2.1.2] and [Kerber99, Corollary 11.3.5] give two different ways of reconstructing a permutation from its Lehmer code; moreover, [Kerber99, Corollary 11.3.5] shows how the Lehmer code of a permutation $\sigma \in S_n$ leads to a specific representation of σ as a product of some of the $s_1, s_2, \ldots, s_{n-1}$.)

Exercise 5.19. Let $n \in \mathbb{N}$. Let $\sigma \in S_n$ and $\tau \in S_n$. We shall use the notation from Definition 3.48.

(a) Prove that each $i \in [n]$ satisfies

$$\begin{aligned} & \ell_{\tau(i)}(\sigma) + \ell_i(\tau) - \ell_i(\sigma \circ \tau) \\ & = \sum_{j \in [n]} [j > i] [\tau(i) > \tau(j)] [\sigma(\tau(j)) > \sigma(\tau(i))] \\ & \qquad + \sum_{j \in [n]} [i > j] [\tau(j) > \tau(i)] [\sigma(\tau(i)) > \sigma(\tau(j))] . \end{aligned}$$

(b) Prove that

$$\ell(\sigma) + \ell(\tau) - \ell(\sigma \circ \tau) = 2 \sum_{i \in [n]} \sum_{j \in [n]} [j > i] [\tau(i) > \tau(j)] [\sigma(\tau(j)) > \sigma(\tau(i))] .$$

(c) Give a new solution to Exercise 5.2 **(a)**.

(d) Give a new solution to Exercise 5.2 **(b)**.

(e) Give a new solution to Exercise 5.2 **(c)**.

Exercise 5.20. Let $n \in \mathbb{N}$ and $\sigma \in S_n$. Let i and j be two elements of $[n]$ such that $i < j$ and $\sigma(i) > \sigma(j)$. Let Q be the set of all $k \in \{i+1, i+2, \ldots, j-1\}$ satisfying $\sigma(i) > \sigma(k) > \sigma(j)$. Prove that

$$\ell\left(\sigma \circ t_{i,j}\right) = \ell(\sigma) - 2|Q| - 1.$$

The following exercise shows an explicit way of expressing every permutation $\sigma \in S_n$ as a product of $\ell(\sigma)$ many simple transpositions (i.e., transpositions of the form s_i with $i \in \{1, 2, \ldots, n-1\}$):

Exercise 5.21. Let $n \in \mathbb{N}$ and $\sigma \in S_n$. If u and v are any two elements of $[n]$ such that $u \leq v$, then we define a permutation $c_{u,v} \in S_n$ by

$$c_{u,v} = \operatorname{cyc}_{v, v-1, v-2, \ldots, u} .$$

For each $i \in [n]$, we define a permutation $a_i \in S_n$ by

$$a_i = c_{i, i + \ell_i(\sigma)} .$$

(a) Prove that a_i is well-defined for each $i \in [n]$.

(b) Prove that each $i \in [n]$ satisfies $a_i = s_{i'-1} \circ s_{i'-2} \circ \cdots \circ s_i$, where $i' = i + \ell_i(\sigma)$.

(c) Prove that $\sigma = a_1 \circ a_2 \circ \cdots \circ a_n$.

(d) Solve Exercise 5.2 **(e)** again.

(e) Solve Exercise 5.1 **(c)** again.

5.9. Extending permutations

In this short section, we shall discuss a simple yet useful concept: that of extending a permutation of a set Y to a larger set X (where "larger" means that $Y \subseteq X$). The following notations will be used throughout this section:

Definition 5.54. Let X be a set. Then, S_X denotes the set of all permutations of X.

Definition 5.55. Let X be a set. Let Y be a subset of X. For every map $\sigma : Y \to Y$, we define a map $\sigma^{(Y \to X)} : X \to X$ by

$$\left(\sigma^{(Y \to X)} (x) = \begin{cases} \sigma(x), & \text{if } x \in Y; \\ x, & \text{if } x \notin Y \end{cases} \qquad \text{for every } x \in X \right).$$

(This map $\sigma^{(Y \to X)}$ is indeed well-defined, according to Proposition 5.56 below.)

The latter of these two definitions relies on the following lemma:

Proposition 5.56. Let X be a set. Let Y be a subset of X. Let $\sigma : Y \to Y$ be a map. Then, the map $\sigma^{(Y \to X)}$ in Definition 5.55 is well-defined.

Proposition 5.56 is easy to prove; its proof is part of Exercise 5.22 further below.

The idea behind the definition of $\sigma^{(Y \to X)}$ in Definition 5.55 is simple: $\sigma^{(Y \to X)}$ is just the most straightforward way of extending $\sigma : Y \to Y$ to a map from X to X (namely, by letting it keep every element of $X \setminus Y$ unchanged).

Example 5.57. **(a)** If $X = \{1, 2, 3, 4, 5, 6, 7\}$ and $Y = \{1, 2, 3, 4\}$, and if $\sigma \in S_Y = S_4$ is the permutation whose one-line notation is $(4, 1, 3, 2)$, then $\sigma^{(Y \to X)} \in S_X = S_7$ is the permutation whose one-line notation is $(4, 1, 3, 2, 5, 6, 7)$.

(b) More generally, if $X = \{1, 2, \ldots, n\}$ and $Y = \{1, 2, \ldots, m\}$ for two nonnegative integers n and m satisfying $n \geq m$, and if $\sigma \in S_Y = S_m$ is any permutation, then the permutation $\sigma^{(Y \to X)} \in S_X = S_n$ has one-line notation $(\sigma(1), \sigma(2), \ldots, \sigma(m), m+1, m+2, \ldots, n)$.

(c) If $X = \{1, 2, 3, \ldots\}$ and $Y = \{1, 2, \ldots, n\}$ for some $n \in \mathbb{N}$, and if $\sigma \in S_Y = S_n$ is any permutation, then the permutation $\sigma^{(Y \to X)} \in S_X = S_\infty$ is precisely the permutation $\sigma_{(\infty)}$ defined in Remark 5.26.

Here are some further properties of the operation that transforms σ into $\sigma^{(Y \to X)}$:

Proposition 5.58. Let X be a set. Let Y be a subset of X.

(a) If $\alpha : Y \to Y$ and $\beta : Y \to Y$ are two maps, then

$$(\alpha \circ \beta)^{(Y \to X)} = \alpha^{(Y \to X)} \circ \beta^{(Y \to X)}.$$

(b) The map $\operatorname{id}_Y : Y \to Y$ satisfies

$$\left(\operatorname{id}_Y\right)^{(Y \to X)} = \operatorname{id}_X .$$

(c) Every permutation $\sigma \in S_Y$ satisfies $\sigma^{(Y \to X)} \in S_X$ and

$$\left(\sigma^{-1}\right)^{(Y \to X)} = \left(\sigma^{(Y \to X)}\right)^{-1} .$$

(d) We have

$$\left\{\delta^{(Y \to X)} \mid \delta \in S_Y\right\} = \left\{\tau \in S_X \mid \tau\left(z\right) = z \text{ for every } z \in X \setminus Y\right\} .$$

(e) The map

$$S_Y \to \left\{\tau \in S_X \mid \tau\left(z\right) = z \text{ for every } z \in X \setminus Y\right\} ,$$
$$\delta \mapsto \delta^{(Y \to X)}$$

is well-defined and bijective.

Proposition 5.59. Let X be a set. Let Y be a subset of X. Let Z be a subset of Y. Let $\sigma : Z \to Z$ be any map. Then,

$$\left(\sigma^{(Z \to Y)}\right)^{(Y \to X)} = \sigma^{(Z \to X)} .$$

Proposition 5.60. Let X be a set. Let Y be a subset of X. Let $\alpha : Y \to Y$ be a map. Let $\beta : X \setminus Y \to X \setminus Y$ be a map. Then,

$$\alpha^{(Y \to X)} \circ \beta^{(X \setminus Y \to X)} = \beta^{(X \setminus Y \to X)} \circ \alpha^{(Y \to X)} .$$

The above propositions are fairly straightforward; again, see Exercise 5.22 for their proofs. Interestingly, we can use these simple facts to prove the following nontrivial theorem:

Theorem 5.61. Let X be a finite set. Let $\pi \in S_X$. Then, there exists a $\sigma \in S_X$ such that $\sigma \circ \pi \circ \sigma^{-1} = \pi^{-1}$.

In fact, we can show (by induction on $\left|X\right|$) the following more general fact:

Proposition 5.62. Let X be a finite set. Let $x \in X$. Let $\pi \in S_X$. Then, there exists a $\sigma \in \left\{\delta^{(X \setminus \{x\} \to X)} \mid \delta \in S_{X \setminus \{x\}}\right\}$ such that $\sigma \circ \pi \circ \sigma^{-1} = \pi^{-1}$.

Exercise 5.22. Prove Proposition 5.56, Proposition 5.58, Proposition 5.59, Proposition 5.60, Proposition 5.62 and Theorem 5.61.

Theorem 5.61 is a known fact, and it is commonly obtained as part of the study of conjugacy in symmetric groups. If π_1 and π_2 are two permutations of a set X, then π_1 is said to be *conjugate* to π_2 if and only if there exists some $\sigma \in S_X$ such that $\sigma \circ \pi_1 \circ \sigma^{-1} = \pi_2$. Thus, Theorem 5.61 says that every permutation π of a finite set X is conjugate to its inverse π^{-1}. Standard proofs of this theorem[167] tend to derive it from the fact that two permutations π_1 and π_2 of a finite set X are conjugate to one another[168] if and only if they have the same "cycle type" (see [Conrad3, Theorem 5.7] for what this means and for a proof).

5.10. Additional exercises

Permutations and symmetric groups are a staple of combinatorics; there are countless results involving them. For an example, Bóna's book [Bona22], as well as significant parts of Stanley's [Stanle11] and [Stanle01] are devoted to them. In this section, I shall only give a haphazard selection of exercises, which are not relevant to the rest of these notes (thus can be skipped at will). I am not planning to provide solutions for all of them.

Exercise 5.23. Let $n \in \mathbb{N}$. Let $\sigma \in S_n$. Let $a_1, a_2, \ldots, a_n$ be any n numbers. (Here, "number" means "real number" or "complex number" or "rational number", as you prefer; this makes no difference.) Prove that

$$\sum_{\substack{1 \leq i < j \leq n; \\ \sigma(i) > \sigma(j)}} \left(a_j - a_i\right) = \sum_{i=1}^{n} a_i \left(i - \sigma\left(i\right)\right).$$

[Here, the summation sign "$\sum_{\substack{1 \leq i < j \leq n; \\ \sigma(i) > \sigma(j)}}$" means "$\sum_{\substack{(i,j) \in \{1,2,\ldots,n\}^2; \\ i<j \text{ and } \sigma(i) > \sigma(j)}}$"; this is a sum over all inversions of σ.]

Exercise 5.24. Let $n \in \mathbb{N}$. Let $\pi \in S_n$.

(a) Prove that

$$\sum_{\substack{1 \leq i < j \leq n; \\ \pi(i) > \pi(j)}} \left(\pi\left(j\right) - \pi\left(i\right)\right) = \sum_{\substack{1 \leq i < j \leq n; \\ \pi(i) > \pi(j)}} \left(i - j\right).$$

[167] which, incidentally, also holds for infinite sets X, provided that one believes in the Axiom of Choice

[168] It is easy to see that being conjugate is a symmetric relation: If a permutation π_1 is conjugate to a permutation π_2, then π_2 is, in turn, conjugate to π_1.

[Here, the summation sign "$\sum\limits_{\substack{1\leq i<j\leq n;\\ \pi(i)>\pi(j)}}$" means "$\sum\limits_{\substack{(i,j)\in\{1,2,\ldots,n\}^2;\\ i<j \text{ and } \pi(i)>\pi(j)}}$"; this is a sum over all inversions of π.]

(b) Prove that

$$\sum_{\substack{1\leq i<j\leq n;\\ \pi(i)<\pi(j)}} \left(\pi\left(j\right)-\pi\left(i\right)\right) = \sum_{\substack{1\leq i<j\leq n;\\ \pi(i)<\pi(j)}} \left(j-i\right).$$

[Here, the summation sign "$\sum\limits_{\substack{1\leq i<j\leq n;\\ \pi(i)<\pi(j)}}$" means "$\sum\limits_{\substack{(i,j)\in\{1,2,\ldots,n\}^2;\\ i<j \text{ and } \pi(i)<\pi(j)}}$".]

Exercise 5.24 is [SacUlf11, Proposition 2.4].

Exercise 5.25. Whenever m is an integer, we shall use the notation $[m]$ for the set $\{1,2,\ldots,m\}$. Also, recall Definition 5.30.

Let $n\in\mathbb{N}$. Let $\sigma\in S_n$. Exercise 5.9 shows that there is a unique n-tuple $(i_1,i_2,\ldots,i_n)\in[1]\times[2]\times\cdots\times[n]$ such that

$$\sigma = t_{1,i_1}\circ t_{2,i_2}\circ\cdots\circ t_{n,i_n}.$$

Consider this $(i_1,i_2,\ldots,i_n)$.

For each $k\in\{0,1,\ldots,n\}$, we define a permutation $\sigma_k\in S_n$ by $\sigma_k = t_{1,i_1}\circ t_{2,i_2}\circ\cdots\circ t_{k,i_k}$.

For each $k\in[n]$, we let $m_k=\sigma_k\left(k\right)$.

(a) Prove that $\sigma_k\left(i\right)=i$ for each $i\in[n]$ and each $k\in\{0,1,\ldots,i-1\}$.

(b) Prove that $m_k\in[k]$ for all $k\in[n]$.

(c) Prove that $\sigma_k\left(i_k\right)=k$ for all $k\in[n]$.

(d) Prove that $\sigma_k=t_{k,m_k}\circ\sigma_{k-1}$ for all $k\in[n]$.

(e) Show that $\sigma^{-1}=t_{1,m_1}\circ t_{2,m_2}\circ\cdots\circ t_{n,m_n}$.

(f) Let $x_1,x_2,\ldots,x_n,y_1,y_2,\ldots,y_n$ be any $2n$ numbers (e.g., rational numbers or real numbers or complex numbers). Prove that

$$\sum_{k=1}^n x_ky_k-\sum_{k=1}^n x_ky_{\sigma(k)}=\sum_{k=1}^n\left(x_{i_k}-x_k\right)\left(y_{m_k}-y_k\right).$$

(g) Now assume that the numbers $x_1,x_2,\ldots,x_n,y_1,y_2,\ldots,y_n$ are real and satisfy $x_1\geq x_2\geq\cdots\geq x_n$ and $y_1\geq y_2\geq\cdots\geq y_n$. Conclude that

$$\sum_{k=1}^n x_ky_k\geq\sum_{k=1}^n x_ky_{\sigma(k)}.$$

Remark 5.63. The claim of Exercise 5.25 **(g)** is known as the *rearrangement inequality*. It has several simple proofs (see, e.g., its Wikipedia page); the approach suggested by Exercise 5.25 is probably the most complicated, but it has the advantage of giving an "explicit" formula for the difference between the two sides (in Exercise 5.25 **(f)**).

Exercise 5.26. Let $n \in \mathbb{N}$. Let $d = \operatorname{lcm}(1,2,\ldots,n)$. (Here, "lcm" stands for the least common multiple of several integers: Thus, $\operatorname{lcm}(1,2,\ldots,n)$ is the smallest positive integer that is divisible by $1,2,\ldots,n$.)

(a) Show that $\pi^d = \operatorname{id}$ for every $\pi \in S_n$.

(b) Let k be an integer such that every $\pi \in S_n$ satisfies $\pi^k = \operatorname{id}$. Show that $d \mid k$.

Exercise 5.27. Let U and V be two finite sets. Let σ be a permutation of U. Let τ be a permutation of V. We define a permutation $\sigma \times \tau$ of the set $U \times V$ by setting

$$(\sigma \times \tau)(a,b) = (\sigma(a), \tau(b)) \qquad \text{for every } (a,b) \in U \times V.$$

(a) Prove that $\sigma \times \tau$ is a well-defined permutation.

(b) Prove that $\sigma \times \tau = (\sigma \times \operatorname{id}_V) \circ (\operatorname{id}_U \times \tau)$.

(c) Prove that $(-1)^{\sigma \times \tau} = \left((-1)^{\sigma}\right)^{|V|} \left((-1)^{\tau}\right)^{|U|}$. (See Exercise 5.12 for the definition of the signs $(-1)^{\sigma \times \tau}$, $(-1)^{\sigma}$ and $(-1)^{\tau}$ appearing here.)

Exercise 5.28. Let $n \in \mathbb{N}$. Let $[n]$ denote the set $\{1,2,\ldots,n\}$. For each $\sigma \in S_n$, define an integer $h(\sigma)$ by

$$h(\sigma) = \sum_{i \in [n]} |\sigma(i) - i|.$$

Let $\sigma \in S_n$.

(a) Prove that

$$h(\sigma) = 2 \sum_{\substack{i \in [n]; \\ \sigma(i) > i}} (\sigma(i) - i) = 2 \sum_{\substack{i \in [n]; \\ \sigma(i) < i}} (i - \sigma(i)).$$

(b) Prove that $h(\sigma \circ \tau) \leq h(\sigma) + h(\tau)$ for any $\tau \in S_n$.

(c) Prove that $h(s_k \circ \sigma) \leq h(\sigma) + 2$ for each $k \in \{1,2,\ldots,n-1\}$.

(d) Prove that

$$h(\sigma)/2 \leq \ell(\sigma) \leq h(\sigma).$$

[**Hint:** The second inequality in part **(d)** is tricky. One way to proceed is by classifying all inversions (i,j) of σ into two types: *Type-I inversions* are those that satisfy $\sigma(i) < j$, whereas *Type-II inversions* are those that satisfy $\sigma(i) \geq j$. Prove that the number of Type-I inversions is $\leq \sum_{\substack{j \in [n]; \\ \sigma(j) < j}} (j - \sigma(j) - 1) \leq \sum_{\substack{i \in [n]; \\ \sigma(i) < i}} (i - \sigma(i))$, whereas the number of Type-II inversions is $\leq \sum\limits_{\substack{i\in[n];\\ \sigma(i)>i}} (\sigma(i) - i)$. Add these together to obtain an upper bound on $\ell(\sigma)$.]

Exercise 5.28 is a result of Diaconis and Graham [DiaGra77, (3.5)][169]. The integer $h(\sigma)$ defined in Exercise 5.28 is called *Spearman's disarray* or *total displacement* of σ. A related concept (the depth of a permutation) has been studied by Petersen and Tenner [PetTen14].

The next two exercises concern the inversions of a permutation. They use the following definition:

Definition 5.64. Let $n \in \mathbb{N}$. For every $\sigma \in S_n$, we let $\operatorname{Inv}\sigma$ denote the set of all inversions of σ.

Exercise 5.2 **(c)** shows that any $n \in \mathbb{N}$ and any two permutations σ and τ in S_n satisfy the inequality $\ell(\sigma \circ \tau) \leq \ell(\sigma) + \ell(\tau)$. In the following exercise, we will see when this inequality becomes an equality:

Exercise 5.29. Let $n \in \mathbb{N}$. Let $\sigma \in S_n$ and $\tau \in S_n$.

(a) Prove that $\ell(\sigma \circ \tau) = \ell(\sigma) + \ell(\tau)$ holds if and only if $\operatorname{Inv}\tau \subseteq \operatorname{Inv}(\sigma \circ \tau)$.

(b) Prove that $\ell(\sigma \circ \tau) = \ell(\sigma) + \ell(\tau)$ holds if and only if $\operatorname{Inv}(\sigma^{-1}) \subseteq \operatorname{Inv}(\tau^{-1} \circ \sigma^{-1})$.

(c) Prove that $\operatorname{Inv}\sigma \subseteq \operatorname{Inv}\tau$ holds if and only if $\ell(\tau) = \ell(\tau \circ \sigma^{-1}) + \ell(\sigma)$.

(d) Prove that if $\operatorname{Inv}\sigma = \operatorname{Inv}\tau$, then $\sigma = \tau$.

(e) Prove that $\ell(\sigma \circ \tau) = \ell(\sigma) + \ell(\tau)$ holds if and only if $(\operatorname{Inv}\sigma) \cap (\operatorname{Inv}(\tau^{-1})) = \varnothing$.

Exercise 5.29 **(d)** shows that if two permutations in S_n have the same set of inversions, then they are equal. In other words, a permutation in S_n is uniquely determined by its set of inversions. The next exercise shows what set of inversions a permutation can have:

Exercise 5.30. Let $n \in \mathbb{N}$. Let $G = \{(i,j) \in \mathbb{Z}^2 \mid 1 \leq i < j \leq n\}$.

A subset U of G is said to be *transitive* if every $a, b, c \in \{1, 2, \ldots, n\}$ satisfying $(a,b) \in U$ and $(b,c) \in U$ also satisfy $(a,c) \in U$.

A subset U of G is said to be *inversive* if there exists a $\sigma \in S_n$ such that $U = \operatorname{Inv}\sigma$.

Let U be a subset of G. Prove that U is inversive if and only if both U and $G \setminus U$ are transitive.

[169] Note that their notations are different; what they call $I(\pi)$ and $D(\pi)$ would be called $\ell(\pi)$ and $h(\pi)$ (respectively) in my terminology.

6. An introduction to determinants

In this chapter, we will define and study determinants in a combinatorial way (in the spirit of Hefferon's book [Heffer20], Gill Williamson's notes [Willia18, Chapter 3], Laue's notes [Laue15] and Zeilberger's paper [Zeilbe85][170]). Nowadays, students usually learn about determinants in the context of linear algebra, after having made the acquaintance of vector spaces, matrices, linear transformations, Gaussian elimination etc.; this approach to determinants (which I like to call the "linear-algebraic approach") has certain advantages and certain disadvantages compared to our combinatorial approach[171].

[170] My notes differ from these sources in the following:

- Hefferon's book [Heffer20] is an introductory textbook for a first course in Linear Algebra, and so treats rather little of the theory of determinants (far less than what we do). It is, however, a good introduction into the "other part" of linear algebra (i.e., the theory of vector spaces and linear maps), and puts determinants into the context of that other part, which makes some of their properties appear less mysterious. (Like many introductory textbooks, it only discusses matrices over fields, not over commutative rings; it also uses more handwaving in the proofs.)
- Zeilberger's paper [Zeilbe85] mostly proves advanced results (apart from its Section 5, which proves our Theorem 6.23). I would recommend reading it after reading this chapter.
- Laue's notes [Laue15] are a brief introduction to determinants that prove the main results in just 14 pages (although at the cost of terser writing and stronger assumptions on the reader's preknowledge). If you read these notes, make sure to pay attention to the "Prerequisites and some Terminology" section, as it explains the (unusual) notations used in these notes.
- Gill Williamson's [Willia18, Chapter 3] probably comes the closest to what I am doing below (and is highly recommended, not least because it goes much further into various interesting directions!). My notes are more elementary and more detailed in what they do.

Other references treating determinants in a combinatorial way are Day's [Day16, Chapter 6], Herstein's [Herstei75, §6.9], Strickland's [Strick13, §12 and Appendix B], Mate's [Mate14], Walker's [Walker87, §5.4], and Pinkham's [Pinkha15, Chapter 11] (but they all limit themselves to the basics).

[171] Its main advantage is that it gives more motivation and context. However, the other (combinatorial) approach requires less preknowledge and involves fewer technical subtleties (for example, it defines the determinant directly by an explicit formula, while the linear-algebraic approach defines it implicitly by a list of conditions which happen to determine it uniquely), which is why I have chosen it.

Examples of texts that introduce determinants via the linear-algebraic approach are [BirMac99, Chapter IX], [GalQua22, Chapter 7], [Goodma15, §8.3], [HofKun71, Chapter 5] and [Hunger03, §VII.3].

Artin, in [Artin10, Chapter 1], takes a particularly quick approach to determinants over a field (although it is quick at the cost of generality: for example, the proof he gives for [Artin10, Theorem 1.4.9] does not generalize to matrices over commutative rings). Axler's [Axler15, Chapter 10B] gives a singularly horrible treatment of determinants – defining them only for real and complex matrices and in a way that utterly hides their combinatorial structure.

We shall study determinants of matrices over *commutative rings*.[172] First, let us define what these words ("commutative ring", "matrix" and "determinant") mean.

6.1. Commutative rings

We begin by defining the concept of commutative rings, and exploring some examples for it. This is only a brief introduction; much more can be found in any text on abstract algebra (e.g., [Artin10], [Goodma15], [Hunger14], [Hunger03], [Herstei75] or [ZarSam67]).

Definition 6.1. If $\mathbb{K}$ is a set, then a *binary operation* on $\mathbb{K}$ means a map from $\mathbb{K} \times \mathbb{K}$ to $\mathbb{K}$. (In other words, it means a function which takes two elements of $\mathbb{K}$ as input, and returns an element of $\mathbb{K}$ as output.) For instance, the map from $\mathbb{Z} \times \mathbb{Z}$ to $\mathbb{Z}$ which sends every pair $(a, b) \in \mathbb{Z} \times \mathbb{Z}$ to $3a - b$ is a binary operation on $\mathbb{Z}$.

Sometimes, a binary operation f on a set $\mathbb{K}$ will be *written infix*. This means that the image of $(a, b) \in \mathbb{K} \times \mathbb{K}$ under f will be denoted by afb instead of $f(a, b)$. For instance, the binary operation $+$ on the set $\mathbb{Z}$ (which sends a pair (a, b) of integers to their sum $a + b$) is commonly written infix, because one writes $a + b$ and not $+(a, b)$ for the sum of a and b.

Definition 6.2. A *commutative ring* means a set $\mathbb{K}$ endowed with

- two binary operations called "addition" and "multiplication", and denoted by $+$ and $\cdot$, respectively, and both written infix[173], and
- two elements called $0_{\mathbb{K}}$ and $1_{\mathbb{K}}$

such that the following axioms are satisfied:

- *Commutativity of addition:* We have $a + b = b + a$ for all $a \in \mathbb{K}$ and $b \in \mathbb{K}$.
- *Commutativity of multiplication:* We have $ab = ba$ for all $a \in \mathbb{K}$ and $b \in \mathbb{K}$. Here and in the following, the expression "ab" is shorthand for "$a \cdot b$" (as is usual for products of numbers).
- *Associativity of addition:* We have $a + (b + c) = (a + b) + c$ for all $a \in \mathbb{K}$, $b \in \mathbb{K}$ and $c \in \mathbb{K}$.
- *Associativity of multiplication:* We have $a(bc) = (ab)c$ for all $a \in \mathbb{K}$, $b \in \mathbb{K}$ and $c \in \mathbb{K}$.

[172] This is a rather general setup, which includes determinants of matrices with real entries, of matrices with complex entries, of matrices with polynomial entries, and many other situations. One benefit of working combinatorially is that studying determinants in this general setup is no more difficult than studying them in more restricted settings.

- *Neutrality of* 0*:* We have $a + 0_{\mathbb{K}} = 0_{\mathbb{K}} + a = a$ for all $a \in \mathbb{K}$.
- *Existence of additive inverses:* For every $a \in \mathbb{K}$, there exists an element $a' \in \mathbb{K}$ such that $a + a' = a' + a = 0_{\mathbb{K}}$. This a' is commonly denoted by $-a$ and called the *additive inverse* of a. (It is easy to check that it is unique.)
- *Unitality (a.k.a. neutrality of* 1*):* We have $1_{\mathbb{K}} a = a 1_{\mathbb{K}} = a$ for all $a \in \mathbb{K}$.
- *Annihilation:* We have $0_{\mathbb{K}} a = a 0_{\mathbb{K}} = 0_{\mathbb{K}}$ for all $a \in \mathbb{K}$.
- *Distributivity:* We have $a(b + c) = ab + ac$ and $(a + b)c = ac + bc$ for all $a \in \mathbb{K}$, $b \in \mathbb{K}$ and $c \in \mathbb{K}$. Here and in the following, we are following the usual convention ("PEMDAS") that multiplication-like operations have higher precedence than addition-like operations; thus, the expression "$ab + ac$" must be understood as "$(ab) + (ac)$" (and not, for example, as "$a(b + a)c$").

(Some of these axioms are redundant, in the sense that they can be derived from others. For instance, the equality $(a + b)c = ac + bc$ can be derived from the axiom $a(b + c) = ab + ac$ using commutativity of multiplication. Also, annihilation follows from the other axioms[174]. The reasons why we have chosen these axioms and not fewer (or more, or others) are somewhat a matter of taste. For example, I like to explicitly require annihilation, because it is an important axiom in the definition of a *semiring*, where it no longer follows from the others.)

Definition 6.3. As we have seen in Definition 6.2, a commutative ring consists of a set $\mathbb{K}$, two binary operations on this set named $+$ and $\cdot$, and two elements of this set named 0 and 1. Thus, formally speaking, we should encode a commutative ring as the 5-tuple $(\mathbb{K}, +, \cdot, 0_{\mathbb{K}}, 1_{\mathbb{K}})$. Sometimes we will actually do so; but most of the time, we will refer to the commutative ring just as the "commutative ring $\mathbb{K}$", hoping that the other four entries of the 5-tuple (namely, $+$, $\cdot$, $0_{\mathbb{K}}$ and $1_{\mathbb{K}}$) are clear from the context. This kind of abbreviation is commonplace in mathematics; it is called "*pars pro toto*" (because we are referring to a large

[173] i.e., we write $a + b$ for the image of $(a, b) \in \mathbb{K} \times \mathbb{K}$ under the binary operation called "addition", and we write $a \cdot b$ for the image of $(a, b) \in \mathbb{K} \times \mathbb{K}$ under the binary operation called "multiplication"

[174] In fact, let $a \in \mathbb{K}$. Distributivity yields $(0_{\mathbb{K}} + 0_{\mathbb{K}}) a = 0_{\mathbb{K}} a + 0_{\mathbb{K}} a$, so that $0_{\mathbb{K}} a + 0_{\mathbb{K}} a = \underbrace{(0_{\mathbb{K}} + 0_{\mathbb{K}})}_{\substack{= 0_{\mathbb{K}} \\ \text{(by neutrality of } 0_{\mathbb{K}})}} a = 0_{\mathbb{K}} a$. Adding $-(0_{\mathbb{K}} a)$ on the left, we obtain $-(0_{\mathbb{K}} a) + (0_{\mathbb{K}} a + 0_{\mathbb{K}} a) = -(0_{\mathbb{K}} a) + 0_{\mathbb{K}} a$. But $-(0_{\mathbb{K}} a) + 0_{\mathbb{K}} a = 0_{\mathbb{K}}$ (by the definition of $-(0_{\mathbb{K}} a)$), and associativity of addition shows that $-(0_{\mathbb{K}} a) + (0_{\mathbb{K}} a + 0_{\mathbb{K}} a) = \underbrace{(-(0_{\mathbb{K}} a) + 0_{\mathbb{K}} a)}_{= 0_{\mathbb{K}}} + 0_{\mathbb{K}} a = 0_{\mathbb{K}} + 0_{\mathbb{K}} a = 0_{\mathbb{K}} a$ (by neutrality of $0_{\mathbb{K}}$), so that $0_{\mathbb{K}} a = -(0_{\mathbb{K}} a) + (0_{\mathbb{K}} a + 0_{\mathbb{K}} a) = -(0_{\mathbb{K}} a) + 0_{\mathbb{K}} a = 0_{\mathbb{K}}$. Thus, $0_{\mathbb{K}} a = 0_{\mathbb{K}}$ is proven. Similarly one can show $a 0_{\mathbb{K}} = 0_{\mathbb{K}}$. Therefore, annihilation follows from the other axioms.

structure by the same symbol as for a small part of it, and hoping that the rest can be inferred from the context). It is an example of what is called "abuse of notation".

The elements $0_{\mathbb{K}}$ and $1_{\mathbb{K}}$ of a commutative ring $\mathbb{K}$ are called the *zero* and the *unity*[175] of $\mathbb{K}$. They are usually denoted by 0 and 1 (without the subscript $\mathbb{K}$) when this can cause no confusion (and, unfortunately, often also when it can). They are not always identical with the actual integers 0 and 1.

The binary operations $+$ and $\cdot$ in Definition 6.2 are also usually not identical with the binary operations $+$ and $\cdot$ on the set of integers, and are denoted by $+_{\mathbb{K}}$ and $\cdot_{\mathbb{K}}$ when confusion can arise.

The set $\mathbb{K}$ is called the *underlying set* of the commutative ring $\mathbb{K}$. Let us again remind ourselves that the underlying set of a commutative ring $\mathbb{K}$ is just a part of the data of $\mathbb{K}$.

Here are some examples and non-examples of rings:[176]

- The sets $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$ and $\mathbb{C}$ (endowed with the usual addition, the usual multiplication, the usual 0 and the usual 1) are commutative rings. (Notice that existence of **multiplicative** inverses is not required[177]!)

- The set $\mathbb{N}$ of nonnegative integers (again endowed with the usual addition, the usual multiplication, the usual 0 and the usual 1) is **not** a commutative ring. It fails the existence of additive inverses. (Of course, negative numbers exist, but this does not count because they don't lie in $\mathbb{N}$.)

- We can define a commutative ring $\mathbb{Z}'$ as follows: We define a binary operation $\widetilde{\times}$ on $\mathbb{Z}$ (written infix) by

$$\left(a \widetilde{\times} b = -ab \qquad \text{for all } (a,b) \in \mathbb{Z} \times \mathbb{Z}\right).$$

  Now, let $\mathbb{Z}'$ be the **set** $\mathbb{Z}$, endowed with the usual addition $+$ and the (unusual) multiplication $\widetilde{\times}$, with the zero $0_{\mathbb{Z}'} = 0$ and with the unity $1_{\mathbb{Z}'} = -1$.

[175] Some people say "unit" instead of "unity", but other people use the word "unit" for something different, which makes every use of this word a potential pitfall.

[176] The following list of examples is long, and some of these examples rely on knowledge that you might not have yet. As usual with examples, you need not understand them all. When I say that Laurent polynomial rings are examples of commutative rings, I do not assume that you know what Laurent polynomials are; I merely want to ensure that, **if** you have already encountered Laurent polynomials, then you get to know that they form a commutative ring.

[177] A *multiplicative inverse* of an element $a \in \mathbb{K}$ means an element $a' \in \mathbb{K}$ such that $aa' = a'a = 1_{\mathbb{K}}$. (This is analogous to an additive inverse, except that addition is replaced by multiplication, and $0_{\mathbb{K}}$ is replaced by $1_{\mathbb{K}}$.) In a commutative ring, every element is required to have an additive inverse (by the definition of a commutative ring), but not every element is guaranteed to have a multiplicative inverse. (For instance, 2 has no multiplicative inverse in $\mathbb{Z}$, and 0 has no multiplicative inverse in any of the rings $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$ and $\mathbb{C}$.)

We shall study multiplicative inverses more thoroughly in Section 6.8 (where we will just call them "inverses").

It is easy to check that $\mathbb{Z}'$ is a commutative ring[178]; it is an example of a commutative ring whose unity is clearly **not** equal to the integer 1 (which is why it is important to never omit the subscript $\mathbb{Z}'$ in $1_{\mathbb{Z}'}$ here).

That said, $\mathbb{Z}'$ is not a very interesting ring: It is essentially "a copy of $\mathbb{Z}$, except that every integer n has been renamed as $-n$". To formalize this intuition, we would need to introduce the notion of a *ring isomorphism*, which we don't want to do right here; but the idea is that the bijection

$$\varphi : \mathbb{Z} \to \mathbb{Z}', \qquad n \mapsto -n$$

satisfies

$$\begin{aligned} \varphi(a+b) &= \varphi(a) + \varphi(b) && \text{for all } (a,b) \in \mathbb{Z} \times \mathbb{Z}; \\ \varphi(a \cdot b) &= \varphi(a) \widetilde{\times} \varphi(b) && \text{for all } (a,b) \in \mathbb{Z} \times \mathbb{Z}; \\ \varphi(0) &= 0_{\mathbb{Z}'}; \\ \varphi(1) &= 1_{\mathbb{Z}'}, \end{aligned}$$

and thus the ring $\mathbb{Z}'$ can be viewed as the ring $\mathbb{Z}$ with its elements "relabelled" using this bijection.

- The polynomial rings $\mathbb{Z}[x]$, $\mathbb{Q}[a,b]$, $\mathbb{C}[z_1, z_2, \ldots, z_n]$ are commutative rings. Laurent polynomial rings are also commutative rings. (Do not worry if you have not seen these rings yet.)

- The set of all functions $\mathbb{Q} \to \mathbb{Q}$ is a commutative ring, where addition and multiplication are defined pointwise (i.e., addition is defined by $(f+g)(x) = f(x) + g(x)$ for all $x \in \mathbb{Q}$, and multiplication is defined by $(fg)(x) = f(x) \cdot g(x)$ for all $x \in \mathbb{Q}$), where the zero is the "constant-0" function (sending every $x \in \mathbb{Q}$ to 0), and where the unity is the "constant-1" function (sending every $x \in \mathbb{Q}$ to 1). Of course, the same construction works if we consider functions $\mathbb{R} \to \mathbb{C}$, or functions $\mathbb{C} \to \mathbb{Q}$, or functions $\mathbb{N} \to \mathbb{Q}$, instead of functions $\mathbb{Q} \to \mathbb{Q}$. [179]

- The set $\mathbb{S}$ of all real numbers of the form $a + b\sqrt{5}$ with $a, b \in \mathbb{Q}$ (endowed with the usual notions of "addition" and "multiplication" defined on $\mathbb{R}$) is a commutative ring[180].

[178] Notice that we have named this new commutative ring $\mathbb{Z}'$, not $\mathbb{Z}$ (despite having $\mathbb{Z}' = \mathbb{Z}$ as sets). The reason is that if we had named it $\mathbb{Z}$, then we could no longer speak of "the commutative ring $\mathbb{Z}$" without being ambiguous (we would have to specify every time whether we mean the usual multiplication or the unusual one).

[179] But not if we consider functions $\mathbb{Q} \to \mathbb{N}$; such functions might fail the existence of additive inverses.

Generally, if X is any set and $\mathbb{K}$ is any commutative ring, then the set of all functions $X \to \mathbb{K}$ is a commutative ring, where addition and multiplication are defined pointwise, where the zero is the "constant-$0_{\mathbb{K}}$" function, and where the unity is the "constant-$1_{\mathbb{K}}$" function.

[180] To prove this, we argue as follows:

- We could define a different ring structure on the set $\mathbb{S}$ (that is, a commutative ring which, as a set, is identical with $\mathbb{S}$, but has a different choice of operations) as follows: We define a binary operation $*$ on $\mathbb{S}$ by setting
$$\left(a+b\sqrt{5}\right) * \left(c+d\sqrt{5}\right) = ac+bd\sqrt{5} \qquad \text{for all } (a,b) \in \mathbb{Q}\times\mathbb{Q} \text{ and } (c,d) \in \mathbb{Q}\times\mathbb{Q}.$$
[181] Now, let $\mathbb{S}'$ be the set $\mathbb{S}$, endowed with the usual addition $+$ and the (unusual) multiplication $*$, with the zero $0_{\mathbb{S}'} = 0$ and with the unity $1_{\mathbb{S}'} = 1+\sqrt{5}$ (not the integer 1). It is easy to check that $\mathbb{S}'$ is a commutative ring[182]. The **sets** $\mathbb{S}$ and $\mathbb{S}'$ are identical, but the **commutative rings** $\mathbb{S}$ and $\mathbb{S}'$ are not[183]: For example, the ring $\mathbb{S}'$ has two nonzero elements whose product is 0 (namely, $1 * \sqrt{5} = 0$), whereas the ring $\mathbb{S}$ has no such things. This shows that not only do we have $\mathbb{S}' \neq \mathbb{S}$ as commutative rings, but there is also no way to regard $\mathbb{S}'$ as "a copy of $\mathbb{S}$ with its elements renamed" (in the same way as we have regarded $\mathbb{Z}'$ as "a copy of $\mathbb{Z}$ with its elements renamed"). This example should stress the point that a commutative ring $\mathbb{K}$ is not just a set; it is a set endowed with two operations ($+$ and $\cdot$) and two elements ($0_{\mathbb{K}}$ and $1_{\mathbb{K}}$), and these operations and elements are no less important than the set.

- The set $\mathbb{S}_3$ of all real numbers of the form $a+b\sqrt[3]{5}$ with $a,b \in \mathbb{Q}$ (endowed with the usual addition, the usual multiplication, the usual 0 and the usual 1) is **not** a commutative ring. Indeed, multiplication is not a binary operation on this set $\mathbb{S}_3$: It does not always send two elements of $\mathbb{S}_3$ to an element of $\mathbb{S}_3$. For instance, $\left(1+1\sqrt[3]{5}\right)\left(1+1\sqrt[3]{5}\right) = 1+2\sqrt[3]{5}+\left(\sqrt[3]{5}\right)^2$ is not in $\mathbb{S}_3$.

- The set of all 2×2-matrices over $\mathbb{Q}$ is **not** a commutative ring, because commutativity of multiplication does not hold for this set. (In general, $AB \neq BA$ for matrices.)

- If you like the empty set, you will enjoy the *zero ring*. This is the commutative ring which is defined as the one-element set $\{0\}$, with zero and unity both

– Addition and multiplication are indeed two binary operations on $\mathbb{S}$. This is because the sum of two elements of $\mathbb{S}$ is an element of $\mathbb{S}$ (namely, $\left(a+b\sqrt{5}\right)+\left(c+d\sqrt{5}\right) = (a+c)+(b+d)\sqrt{5}$), and so is their product (namely, $\left(a+b\sqrt{5}\right)\cdot\left(c+d\sqrt{5}\right) = (ac+5bd)+(bc+ad)\sqrt{5}$).

– All axioms of a commutative ring are satisfied for $\mathbb{S}$, except maybe the existence of additive inverses. This is simply because the addition and the multiplication in $\mathbb{S}$ are "inherited" from $\mathbb{R}$, and clearly all these axioms come with the inheritance.

– Existence of additive inverses also holds in $\mathbb{S}$, because the additive inverse of $a+b\sqrt{5}$ is $(-a)+(-b)\sqrt{5}$.

[181] This is well-defined, because every element of $\mathbb{S}$ can be written in the form $a+b\sqrt{5}$ for a **unique** pair $(a,b) \in \mathbb{Q}\times\mathbb{Q}$. This is a consequence of the irrationality of $\sqrt{5}$.

[182] Again, we do not call it $\mathbb{S}$, in order to be able to distinguish between different ring structures.

[183] Keep in mind that, due to our "pars pro toto" notation, "commutative ring $\mathbb{S}$" means more than "set $\mathbb{S}$".

being 0 (nobody said that they have to be distinct!), with addition given by $0 + 0 = 0$ and with multiplication given by $0 \cdot 0 = 0$. Of course, it is not an empty set[184], but it plays a similar role in the world of commutative rings as the empty set does in the world of sets: It carries no information itself, but things would break if it were to be excluded[185].

Notice that the zero and the unity of the zero ring are identical, i.e., we have $0_{\mathbb{K}} = 1_{\mathbb{K}}$. This shows why it is dangerous to omit the subscripts and just denote the zero and the unity by 0 and 1; in fact, you don't want to rewrite the equality $0_{\mathbb{K}} = 1_{\mathbb{K}}$ as "$0 = 1$"! (Most algebraists make a compromise between wanting to omit the subscripts and having to clarify what 0 and 1 mean: They say that "$0 = 1$ in $\mathbb{K}$" to mean "$0_{\mathbb{K}} = 1_{\mathbb{K}}$".)

Generally, a *trivial ring* is defined to be a commutative ring containing only one element (which then necessarily is both the zero and the unity of this ring). The addition and the multiplication of a trivial ring are uniquely determined (since there is only one possible value that a sum or a product could take). Every trivial ring can be viewed as the zero ring with its element 0 relabelled.[186]

- In set theory, the *symmetric difference* of two sets A and B is defined to be the set $(A \cup B) \setminus (A \cap B) = (A \setminus B) \cup (B \setminus A)$. This symmetric difference is denoted by $A \bigtriangleup B$. Now, let S be any set. Let $\mathcal{P}(S)$ denote the powerset of S (that is, the set of all subsets of S). It is easy to check that the following ten properties hold:

$$\begin{gathered}
A \bigtriangleup B = B \bigtriangleup A \qquad \text{for any sets } A \text{ and } B; \\
A \cap B = B \cap A \qquad \text{for any sets } A \text{ and } B; \\
(A \bigtriangleup B) \bigtriangleup C = A \bigtriangleup (B \bigtriangleup C) \qquad \text{for any sets } A, B \text{ and } C; \\
(A \cap B) \cap C = A \cap (B \cap C) \qquad \text{for any sets } A, B \text{ and } C; \\
A \bigtriangleup \varnothing = \varnothing \bigtriangleup A = A \qquad \text{for any set } A; \\
A \bigtriangleup A = \varnothing \qquad \text{for any set } A; \\
A \cap S = S \cap A = A \qquad \text{for any subset } A \text{ of } S; \\
\varnothing \cap A = A \cap \varnothing = \varnothing \qquad \text{for any set } A; \\
A \cap (B \bigtriangleup C) = (A \cap B) \bigtriangleup (A \cap C) \qquad \text{for any sets } A, B \text{ and } C; \\
(A \bigtriangleup B) \cap C = (A \cap C) \bigtriangleup (B \cap C) \qquad \text{for any sets } A, B \text{ and } C.
\end{gathered}$$

Therefore, $\mathcal{P}(S)$ becomes a commutative ring, where the addition is defined to be the operation $\bigtriangleup$, the multiplication is defined to be the operation $\cap$, the

[184] A commutative ring cannot be empty, as it contains at least one element (namely, 0).

[185] Some authors **do** prohibit the zero ring from being a commutative ring (by requiring every commutative ring to satisfy $0 \neq 1$). I think most of them run into difficulties from this decision sooner or later.

[186] In more formal terms, the preceding statement would say that "every trivial ring is isomorphic to the zero ring".

zero is defined to be the set $\varnothing$, and the unity is defined to be the set S. [187]

The commutative ring $\mathcal{P}(S)$ has the property that $a \cdot a = a$ for every $a \in \mathcal{P}(S)$. (This simply means that $A \cap A = A$ for every $A \subseteq S$.) Commutative rings that have this property are called *Boolean rings*. (Of course, $\mathcal{P}(S)$ is the eponymic example for a Boolean ring; but there are also others.)

- For every positive integer n, the residue classes of integers modulo n form a commutative ring, which is called $\mathbb{Z}/n\mathbb{Z}$ or $\mathbb{Z}_n$ (depending on the author). This ring has n elements (often called "integers modulo n"). When n is a composite number (e.g., $n = 6$), this ring has the property that products of nonzero[188] elements can be zero (e.g., we have $2 \cdot 3 \equiv 0 \operatorname{mod} 6$); this means that there is no way to define division by all nonzero elements in this ring (even if we are allowed to create fractions). Notice that $\mathbb{Z}/1\mathbb{Z}$ is a trivial ring.

  We notice that if n is a positive integer, and if $\mathbb{K}$ is the commutative ring $\mathbb{Z}/n\mathbb{Z}$, then $\underbrace{1_{\mathbb{K}} + 1_{\mathbb{K}} + \cdots + 1_{\mathbb{K}}}_{n \text{ times}} = 0_{\mathbb{K}}$ (because the left hand side of this equality is the residue class of n modulo n, while the right hand side is the residue class of 0 modulo n, and these two residue classes are clearly equal).

- Let us try to define "division by zero". So, we introduce a new symbol ∞, and we try to extend the addition on $\mathbb{Q}$ to the set $\mathbb{Q} \cup \{\infty\}$ by setting $a + \infty = \infty$ for all $a \in \mathbb{Q} \cup \{\infty\}$. We might also try to extend the multiplication in some way, and perhaps to add some more elements (such as another symbol $-\infty$ to serve as the product $(-1)\infty$). I claim that (whatever we do with the multiplication, and whatever new elements we add) we do not get a commutative ring. Indeed, assume the contrary. Thus, there exists a commutative ring $\mathbb{W}$ which contains $\mathbb{Q} \cup \{\infty\}$ as a subset, and which has $a + \infty = \infty$ for all $a \in \mathbb{Q}$. Thus, in $\mathbb{W}$, we have $1 + \infty = \infty = 0 + \infty$. Adding $(-1)\infty$ to both sides of this equality, we obtain $1 + \infty + (-1)\infty = 0 + \infty + (-1)\infty$, so that $1 = 0$ [189]; but this is absurd. Hence, we have found a contradiction. This is why "division by zero is impossible": One can define objects that behave like "infinity" (and they **are** useful), but they break various standard rules such as the axioms of a commutative ring. In contrast to this, adding a "number" i satisfying $i^2 = -1$ to the real numbers is harmless: The complex numbers $\mathbb{C}$ are still a commutative ring.

[187] The ten properties listed above show that the axioms of a commutative ring are satisfied for $(\mathcal{P}(S), \triangle, \cap, \varnothing, S)$. In particular, the sixth property shows that every subset A of S has an additive inverse – namely, itself. Of course, it is unusual for an element of a commutative ring to be its own additive inverse, but in this example it happens all the time!

[188] An element a of a commutative ring $\mathbb{K}$ is said to be *nonzero* if $a \neq 0_{\mathbb{K}}$. (This is not the same as saying that a is not the integer 0, because the integer 0 might not be $0_{\mathbb{K}}$.)

[189] because $\infty + (-1)\infty = 1\infty + (-1)\infty = \underbrace{(1 + (-1))}_{=0}\infty = 0\infty = 0$

- Here is an "almost-ring" beloved to many combinatorialists: the *max-plus semiring* $\mathbb{T}$ (also called the *tropical semiring*[190]). We create a new symbol $-\infty$, and we set $\mathbb{T} = \mathbb{Z} \cup \{-\infty\}$ as sets, but we do **not** "inherit" the addition and the multiplication from $\mathbb{Z}$. Instead, we denote the "addition" and "multiplication" operations on $\mathbb{Z}$ by $+_{\mathbb{Z}}$ and $\cdot_{\mathbb{Z}}$, and we define two new "addition" and "multiplication" operations $+_{\mathbb{T}}$ and $\cdot_{\mathbb{T}}$ on $\mathbb{T}$ as follows:

  $$a +_{\mathbb{T}} b = \max\{a, b\};$$
  $$a \cdot_{\mathbb{T}} b = a +_{\mathbb{Z}} b.$$

  (Here, we set $\max\{-\infty, n\} = \max\{n, -\infty\} = n$ and $(-\infty) +_{\mathbb{Z}} n = n +_{\mathbb{Z}} (-\infty) = -\infty$ for every $n \in \mathbb{T}$.)

  It turns out that the set $\mathbb{T}$ endowed with the two operations $+_{\mathbb{T}}$ and $\cdot_{\mathbb{T}}$, the zero $0_{\mathbb{T}} = -\infty$ and the unity $1_{\mathbb{T}} = 0$ comes rather close to being a commutative ring. It satisfies all axioms of a commutative ring except for the existence of additive inverses. Such a structure is called a *semiring*. Other examples of semirings are $\mathbb{N}$ and a reasonably defined $\mathbb{N} \cup \{\infty\}$ (with $0\infty = 0$ and $a\infty = \infty$ for all $a > 0$).

If $\mathbb{K}$ is a commutative ring, then we can define a subtraction in $\mathbb{K}$, even though we have not required a subtraction operation as part of the definition of a commutative ring $\mathbb{K}$. Namely, the *subtraction* of a commutative ring $\mathbb{K}$ is the binary operation $-$ on $\mathbb{K}$ (again written infix) defined as follows: For every $a \in \mathbb{K}$ and $b \in \mathbb{K}$, set $a - b = a + b'$, where b' is the additive inverse of b. It is not hard to check that $a - b$ is the unique element c of $\mathbb{K}$ satisfying $a = b + c$; thus, subtraction is "the undoing of addition" just as in the classical situation of integers. Again, the notation $-$ for the subtraction of $\mathbb{K}$ is denoted by $-_{\mathbb{K}}$ whenever a confusion with the subtraction of integers could arise.

Whenever a is an element of a commutative ring $\mathbb{K}$, we write $-a$ for the additive inverse of a. This is the same as $0_{\mathbb{K}} - a$.

The intuition for commutative rings is essentially that all computations that can be performed with the operations $+$, $-$ and $\cdot$ on integers can be similarly made in any commutative ring. For instance, if $a_1, a_2, \ldots, a_n$ are n elements of a commutative ring, then the sum $a_1 + a_2 + \cdots + a_n$ is well-defined, and can be computed by adding the elements $a_1, a_2, \ldots, a_n$ to each other in any order[191]. More generally: If S is a finite set, if $\mathbb{K}$ is a commutative ring, and if $(a_s)_{s \in S}$ is a $\mathbb{K}$-valued S-family[192], then the sum $\sum\limits_{s \in S} a_s$ is defined in the same way as finite sums of numbers were defined in Section 1.4 (but with $\mathbb{A}$ replaced by $\mathbb{K}$, of course[193]); this definition is

[190] Caution: Both of these names mean many other things as well.

[191] For instance, we can compute the sum $a + b + c + d$ of four elements a, b, c, d in many ways: For example, we can first add a and b, then add c and d, and finally add the two results; alternatively, we can first add a and b, then add d to the result, then add c to the result. In a commutative ring, all such ways lead to the same result.

[192] See Definition 2.107 for the definition of this notion.

[193] and, consequently, 0 replaced by $0_{\mathbb{K}}$

still legitimate[194], and these finite sums of elements of $\mathbb{K}$ satisfy the same properties as finite sums of numbers (see Section 1.4 for these properties). All this can be proven in the same way as it was proven for numbers (in Section 2.14 and Section 1.4). The same holds for finite products. Furthermore, if n is an integer and a is an element of a commutative ring $\mathbb{K}$, then we define an element na of $\mathbb{K}$ by

$$na = \begin{cases} \underbrace{a + a + \cdots + a}_{n \text{ addends}}, & \text{if } n \geq 0; \\ -\left(\underbrace{a + a + \cdots + a}_{-n \text{ addends}}\right), & \text{if } n < 0 \end{cases}.$$

[195]

If n is a nonnegative integer and a is an element of a commutative ring $\mathbb{K}$, then a^n is a well-defined element of $\mathbb{K}$ (namely, $a^n = \underbrace{a \cdot a \cdot \cdots \cdot a}_{n \text{ factors}}$). In particular, applying this definition to $n = 0$, we obtain

$$a^0 = \underbrace{a \cdot a \cdot \cdots \cdot a}_{0 \text{ factors}} = (\text{empty product}) = 1 \qquad \text{for each } a \in \mathbb{K}.$$

[194] i.e., the result does not depend on the choice of t in (1)

[195] Notice that this definition of na is **not** a particular case of the product of two elements of $\mathbb{K}$, because n is not an element of $\mathbb{K}$.

The following identities hold:

$$(n+m)\,a = na + ma \qquad \text{for } a \in \mathbb{K} \text{ and } n, m \in \mathbb{Z}; \tag{325}$$
$$n\,(a+b) = na + nb \qquad \text{for } a, b \in \mathbb{K} \text{ and } n \in \mathbb{Z}; \tag{326}$$
$$-\,(a+b) = (-a) + (-b) \qquad \text{for } a, b \in \mathbb{K}; \tag{327}$$
$$1a = a \qquad \text{for } a \in \mathbb{K}; \tag{328}$$
$$0a = 0_{\mathbb{K}} \qquad \text{for } a \in \mathbb{K} \tag{329}$$
$$(\text{here, the ``0'' on the left hand side means the integer } 0)\,;$$
$$(-1)\,a = -a \qquad \text{for } a \in \mathbb{K}; \tag{330}$$
$$-\,(-a) = a \qquad \text{for } a \in \mathbb{K}; \tag{331}$$
$$-\,(ab) = (-a)\,b = a\,(-b) \qquad \text{for } a, b \in \mathbb{K}; \tag{332}$$
$$-\,(na) = (-n)\,a = n\,(-a) \qquad \text{for } a \in \mathbb{K} \text{ and } n \in \mathbb{Z}; \tag{333}$$
$$n\,(ab) = (na)\,b = a\,(nb) \qquad \text{for } a, b \in \mathbb{K} \text{ and } n \in \mathbb{Z}; \tag{334}$$
$$(nm)\,a = n\,(ma) \qquad \text{for } a \in \mathbb{K} \text{ and } n, m \in \mathbb{Z}; \tag{335}$$
$$n0_{\mathbb{K}} = 0_{\mathbb{K}} \qquad \text{for } n \in \mathbb{Z};$$
$$1^n = 1 \qquad \text{for } n \in \mathbb{N};$$
$$0^n = \begin{cases} 0, & \text{if } n > 0; \\ 1, & \text{if } n = 0 \end{cases} \qquad \text{for } n \in \mathbb{N}; \tag{336}$$
$$a^{n+m} = a^n a^m \qquad \text{for } a \in \mathbb{K} \text{ and } n, m \in \mathbb{N}; \tag{337}$$
$$a^{nm} = (a^n)^m \qquad \text{for } a \in \mathbb{K} \text{ and } n, m \in \mathbb{N};$$
$$(ab)^n = a^n b^n \qquad \text{for } a, b \in \mathbb{K} \text{ and } n \in \mathbb{N}. \tag{338}$$

Here, we are using the standard notations $+$, $\cdot$, 0 and 1 for the addition, the multiplication, the zero and the unity of $\mathbb{K}$, because confusion (e.g., confusion of the 0 with the integer 0) is rather unlikely.[196] We shall keep doing so in the following, apart from situations where confusion can realistically occur.[197]

The identities listed above are not hard to prove. Indeed, they are generalizations of well-known identities holding for rational numbers; and some of them (for example, (337) and (338)) can be proved in exactly the same way as those identities

[196] For instance, in the statement "$-\,(a+b) = (-a) + (-b)$ for $a, b \in \mathbb{K}$", it is clear that the $+$ can only stand for the addition of $\mathbb{K}$ and not (say) for the addition of integers (since a, b, $-a$ and $-b$ are elements of $\mathbb{K}$, not (generally) integers). The only statement whose meaning is ambiguous is "$0^n = \begin{cases} 0, & \text{if } n > 0; \\ 1, & \text{if } n = 0 \end{cases}$ for $n \in \mathbb{N}$". In this statement, the "0" in "$n > 0$" and the "0" in "$n = 0$" clearly mean the integer 0 (since they are being compared with the integer n), but the other two appearances of "0" and the "1" are ambiguous. I hope that the context makes it clear enough that they mean the zero and the unity of $\mathbb{K}$ (and not the integers 0 and 1), because otherwise this equality would not be a statement about $\mathbb{K}$ at all.

[197] Notice that the equalities (334) and (335) are **not** particular cases of the associativity of multiplication which we required to hold for $\mathbb{K}$. Indeed, the latter associativity says that $a\,(bc) = (ab)\,c$ for all $a \in \mathbb{K}$, $b \in \mathbb{K}$ and $c \in \mathbb{K}$. But in (334) and (335), the n is an integer, not an element of $\mathbb{K}$.

for rational numbers.[198]

If a and b are two elements of a commutative ring $\mathbb{K}$, then the expression "$-ab$" appears ambiguous, since it can be interpreted either as "$-(ab)$" or as "$(-a)b$". But (332) shows that these two interpretations yield the same result; thus, we can write this expression "$-ab$" without fearing ambiguity. Similarly, if $n \in \mathbb{Z}$ and $a, b \in \mathbb{K}$, then the expression "nab" is unambiguous, because (334) shows that the two possible ways to interpret it (namely, as "$n(ab)$" and as "$(na)b$") yield the same result. Similarly, if $n, m \in \mathbb{Z}$ and $a \in \mathbb{K}$, then the expression "nma" is unambiguous, because of (335).

Furthermore, finite sums such as $\sum\limits_{s \in S} a_s$ (where S is a finite set, and $a_s \in \mathbb{K}$ for every $s \in S$), and finite products such as $\prod\limits_{s \in S} a_s$ (where S is a finite set, and $a_s \in \mathbb{K}$ for every $s \in S$) are defined whenever $\mathbb{K}$ is a commutative ring. Again, the definition is the same as for numbers, and these sums and products behave as they do for numbers.[199] For example, Exercise 5.13 still holds if we replace "$\mathbb{C}$" by "$\mathbb{K}$" in it (and the same solution proves it) whenever $\mathbb{K}$ is a commutative ring. From the fact that finite sums and finite products of elements of $\mathbb{K}$ are well-defined, we can also conclude that expressions such as "$a_1 + a_2 + \cdots + a_k$" and "$a_1 a_2 \cdots a_k$" (where $a_1, a_2, \ldots, a_k$ are finitely many elements of $\mathbb{K}$) are well-defined.

Various identities that hold for numbers also hold for elements of arbitrary commutative rings. For example, an analogue of the binomial formula (Proposition 3.21) holds: If $\mathbb{K}$ is a commutative ring, then

$$(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k} \qquad \text{for } a, b \in \mathbb{K} \text{ and } n \in \mathbb{N}. \tag{339}$$

(We can obtain a proof of (339) by re-reading the solution to Exercise 3.6, while replacing every "x" by an "a" and replacing every "y" by a "b". Another proof of (339) is given in the solution to Exercise 6.1 **(b)**.)

Remark 6.4. The notion of a "commutative ring" is not fully standardized; there exist several competing definitions:

For some people, a "commutative ring" is **not** endowed with an element 1 (although it **can** have such an element), and, consequently, does not have to satisfy the unitality axiom. According to their definition, for example, the set

$$\{\ldots, -4, -2, 0, 2, 4, \ldots\} = \{2n \mid n \in \mathbb{Z}\} = (\text{the set of all even integers})$$

[198] For example, it is well-known that

$$(ab)^n = a^n b^n \qquad \text{for any } a, b \in \mathbb{Q} \text{ and } n \in \mathbb{N}.$$

This can be easily proven by induction on n, using the commutativity and associativity rules for multiplication of rational numbers and the fact that $1 \cdot 1 = 1$. The same argument can be used to prove (338). The only change required is replacing every appearance of "$\mathbb{Q}$" by "$\mathbb{K}$".

[199] Of course, empty sums of elements of $\mathbb{K}$ are defined to equal $0_{\mathbb{K}}$, and empty products of elements of $\mathbb{K}$ are defined to equal $1_{\mathbb{K}}$.

is a commutative ring (with the usual addition and multiplication). (In contrast, our definition of a "commutative ring" does not accept this set as a commutative ring, because it does not contain any element which would fill the role of 1.) These people tend to use the notation "commutative ring with unity" (or "commutative ring with 1") to mean a commutative ring which is endowed with a 1 and satisfies the unitality axiom (i.e., what we call a "commutative ring").

On the other hand, there are authors who use the word "ring" for what we call "commutative ring". These are mostly the authors who work with commutative rings all the time and find the name "commutative ring" too long.

When you are reading about rings, it is important to know which meaning of "ring" the author is subscribing to. (Often this can be inferred from the examples given.)

Exercise 6.1. Let $\mathbb{K}$ be a commutative ring. For every $n \in \mathbb{N}$, let $[n]$ denote the set $\{1, 2, \ldots, n\}$.

(a) Let $n \in \mathbb{N}$. Let $a_1, a_2, \ldots, a_n$ be n elements of $\mathbb{K}$. Let $b_1, b_2, \ldots, b_n$ be n further elements of $\mathbb{K}$. Prove that

$$\prod_{i=1}^{n} (a_i + b_i) = \sum_{I \subseteq [n]} \left(\prod_{i \in I} a_i \right) \left(\prod_{i \in [n] \setminus I} b_i \right).$$

(Here, as usual, the summation sign $\sum_{I \subseteq [n]}$ means $\sum_{I \in \mathcal{P}([n])}$, where $\mathcal{P}([n])$ denotes the powerset of $[n]$.)

(b) Use Exercise 6.1 to give a new proof of (339).

Exercise 6.2. For each $m \in \mathbb{N}$ and $(k_1, k_2, \ldots, k_m) \in \mathbb{N}^m$, let us define a positive integer $\mathbf{m}(k_1, k_2, \ldots, k_m)$ by $\mathbf{m}(k_1, k_2, \ldots, k_m) = \dfrac{(k_1 + k_2 + \cdots + k_m)!}{k_1! k_2! \cdots k_m!}$. (This is indeed a positive integer, because Exercise 3.1 says so.)

Let $\mathbb{K}$ be a commutative ring. Let $m \in \mathbb{N}$. Let $a_1, a_2, \ldots, a_m$ be m elements of $\mathbb{K}$. Let $n \in \mathbb{N}$. Prove that

$$(a_1 + a_2 + \cdots + a_m)^n = \sum_{\substack{(k_1, k_2, \ldots, k_m) \in \mathbb{N}^m; \\ k_1 + k_2 + \cdots + k_m = n}} \mathbf{m}(k_1, k_2, \ldots, k_m) \prod_{i=1}^{m} a_i^{k_i}.$$

(This is called the *multinomial formula*.)

6.2. Matrices

We have briefly defined determinants in Definition 5.16, but we haven't done much with them. This will be amended now. But let us first recall the definitions of basic notions in matrix algebra.

In the following, we fix a commutative ring $\mathbb{K}$. The elements of $\mathbb{K}$ will be called *scalars* (to distinguish them from *vectors* and *matrices*, which we will soon discuss, and which are structures containing several elements of $\mathbb{K}$).

If you feel uncomfortable with commutative rings, you are free to think that $\mathbb{K} = \mathbb{Q}$ or $\mathbb{K} = \mathbb{C}$ in the following; but everything I am doing works for any commutative ring unless stated otherwise.

Given two nonnegative integers n and m, an *$n \times m$-matrix* (or, more precisely, *$n \times m$-matrix over* $\mathbb{K}$) means a rectangular table with n rows and m columns whose entries are elements of $\mathbb{K}$. [200] For instance, when $\mathbb{K} = \mathbb{Q}$, the table $\left(\begin{array}{ccc} 1 & -2/5 & 4 \\ 1/3 & -1/2 & 0 \end{array}\right)$ is a 2×3-matrix. A *matrix* simply means an $n \times m$-matrix for some $n \in \mathbb{N}$ and $m \in \mathbb{N}$. These n and m are said to be the *dimensions* of the matrix.

If A is an $n \times m$-matrix, and if $i \in \{1, 2, \ldots, n\}$ and $j \in \{1, 2, \ldots, m\}$, then the *$(i, j)$-th entry of A* means the entry of A in row i and column j. For instance, the $(1, 2)$-th entry of the matrix $\left(\begin{array}{ccc} 1 & -2/5 & 4 \\ 1/3 & -1/2 & 0 \end{array}\right)$ is $-2/5$.

If $n \in \mathbb{N}$ and $m \in \mathbb{N}$, and if we are given an element $a_{i,j} \in \mathbb{K}$ for every $(i, j) \in \{1, 2, \ldots, n\} \times \{1, 2, \ldots, m\}$, then we use the notation $\left(a_{i,j}\right)_{1 \leq i \leq n,\ 1 \leq j \leq m}$ for the $n \times m$-matrix whose (i, j)-th entry is $a_{i,j}$ for all $(i, j) \in \{1, 2, \ldots, n\} \times \{1, 2, \ldots, m\}$. Thus,

$$\left(a_{i,j}\right)_{1 \leq i \leq n,\ 1 \leq j \leq m} = \left(\begin{array}{cccc} a_{1,1} & a_{1,2} & \cdots & a_{1,m} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n,1} & a_{n,2} & \cdots & a_{n,m} \end{array}\right).$$

The letters i and j are not set in stone; they are bound variables like the k in "$\sum\limits_{k=1}^{n} k$". Thus, you are free to write $\left(a_{x,y}\right)_{1 \leq x \leq n,\ 1 \leq y \leq m}$ or $\left(a_{j,i}\right)_{1 \leq j \leq n,\ 1 \leq i \leq m}$ instead of $\left(a_{i,j}\right)_{1 \leq i \leq n,\ 1 \leq j \leq m}$ (and we will use this freedom eventually).[201]

Matrices can be added if they share the same dimensions: If n and m are two nonnegative integers, and if $A = \left(a_{i,j}\right)_{1 \leq i \leq n,\ 1 \leq j \leq m}$ and $B = \left(b_{i,j}\right)_{1 \leq i \leq n,\ 1 \leq j \leq m}$ are two $n \times m$-matrices, then $A + B$ means the $n \times m$-matrix $\left(a_{i,j} + b_{i,j}\right)_{1 \leq i \leq n,\ 1 \leq j \leq m}$. Thus,

[200] Formally speaking, this means that an $n \times m$-matrix is a map from $\{1, 2, \ldots, n\} \times \{1, 2, \ldots, m\}$ to $\mathbb{K}$. We represent such a map as a rectangular table by writing the image of $(i, j) \in \{1, 2, \ldots, n\} \times \{1, 2, \ldots, m\}$ into the cell in the i-th row and the j-th column.

Thus, the notion of an $n \times m$-matrix is closely akin to what we called an "$n \times m$-table of elements of $\mathbb{K}$" in Definition 2.110. The main difference between these two notions is that an $n \times m$-matrix "knows" $\mathbb{K}$, whereas an $n \times m$-table does not (i.e., two $n \times m$-matrices that have the same entries in the same positions but are defined using different commutative rings $\mathbb{K}$ are considered different, but two such $n \times m$-tables are considered identical).

[201] Many authors love to abbreviate "$a_{i,j}$" by "a_{ij}" (hoping that the reader will not mistake the subscript "ij" for a product or (in the case where i and j are single-digit numbers) for a two-digit number). The only advantage of this abbreviation that I am aware of is that it saves you a comma; I do not understand why it is so popular. But you should be aware of it in case you are reading other texts.

matrices are added "entry by entry"; for example, $\begin{pmatrix} a & b & c \\ d & e & f \end{pmatrix} + \begin{pmatrix} a' & b' & c' \\ d' & e' & f' \end{pmatrix} =$ $\begin{pmatrix} a+a' & b+b' & c+c' \\ d+d' & e+e' & f+f' \end{pmatrix}$. Similarly, subtraction is defined: If $A = \left(a_{i,j}\right)_{1\leq i\leq n,\ 1\leq j\leq m}$ and $B = \left(b_{i,j}\right)_{1\leq i\leq n,\ 1\leq j\leq m}$, then $A - B = \left(a_{i,j} - b_{i,j}\right)_{1\leq i\leq n,\ 1\leq j\leq m}$.

Similarly, one can define the product of a scalar $\lambda \in \mathbb{K}$ with a matrix A: If $\lambda \in \mathbb{K}$ is a scalar, and if $A = \left(a_{i,j}\right)_{1\leq i\leq n,\ 1\leq j\leq m}$ is an $n \times m$-matrix, then λA means the $n \times m$-matrix $\left(\lambda a_{i,j}\right)_{1\leq i\leq n,\ 1\leq j\leq m}$.

Defining the product of two matrices is trickier. Matrices are **not** multiplied "entry by entry"; this would not be a very interesting definition. Instead, their product is defined as follows: If n, m and ℓ are three nonnegative integers, then the product AB of an $n \times m$-matrix $A = \left(a_{i,j}\right)_{1\leq i\leq n,\ 1\leq j\leq m}$ with an $m \times \ell$-matrix $B = \left(b_{i,j}\right)_{1\leq i\leq m,\ 1\leq j\leq \ell}$ means the $n \times \ell$-matrix

$$\left(\sum_{k=1}^{m} a_{i,k} b_{k,j}\right)_{1\leq i\leq n,\ 1\leq j\leq \ell}.$$

This definition looks somewhat counterintuitive, so let me comment on it. First of all, for AB to be defined, A and B are **not** required to have the same dimensions; instead, A must have as many columns as B has rows. The resulting matrix AB then has as many rows as A and as many columns as B. Every entry of AB is a sum of products of an entry of A with an entry of B (not a single such product). More precisely, the (i,j)-th entry of AB is a sum of products of an entry in the i-th row of A with the respective entry in the j-th column of B. For example,

$$\begin{pmatrix} a & b & c \\ d & e & f \end{pmatrix} \begin{pmatrix} a' & d' & g' \\ b' & e' & h' \\ c' & f' & i' \end{pmatrix} = \begin{pmatrix} aa'+bb'+cc' & ad'+be'+cf' & ag'+bh'+ci' \\ da'+eb'+fc' & dd'+ee'+ff' & dg'+eh'+fi' \end{pmatrix}.$$

The multiplication of matrices is not commutative! It is easy to find examples of two matrices A and B for which the products AB and BA are distinct, or one of them is well-defined but the other is not[202].

For given $n \in \mathbb{N}$ and $m \in \mathbb{N}$, we define the $n \times m$ *zero matrix* to be the $n \times m$-matrix whose all entries are 0 (that is, the $n \times m$-matrix $(0)_{1\leq i\leq n,\ 1\leq j\leq m}$). We denote this matrix by $0_{n\times m}$. If A is any $n \times m$-matrix, then the $n \times m$-matrix $-A$ is defined to be $0_{n\times m} - A$.

A sum $\sum\limits_{i\in I} A_i$ of finitely many matrices A_i is defined in the same way as a sum of numbers or of elements of a commutative ring[203]. However, a product $\prod\limits_{i\in I} A_i$ of

[202] This happens if A has as many columns as B has rows, but B does not have as many columns as A has rows.

[203] with the caveat that an empty sum of $n \times m$-matrices is not the number 0, but the $n \times m$-matrix $0_{n,m}$

finitely many matrices A_i (in general) cannot be defined, because the result would depend on the order of multiplication.

For every $n \in \mathbb{N}$, we let I_n denote the $n \times n$-matrix $\left(\delta_{i,j}\right)_{1 \leq i \leq n,\ 1 \leq j \leq n}$, where $\delta_{i,j}$ is defined to be $\begin{cases} 1, & \text{if } i = j; \\ 0, & \text{if } i \neq j \end{cases}$. [204] This matrix I_n looks as follows:

$$I_n = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}.$$

It has the property that $I_n B = B$ for every $m \in \mathbb{N}$ and every $n \times m$-matrix B; also, $AI_n = A$ for every $k \in \mathbb{N}$ and every $k \times n$-matrix A. (Proving this is a good way to check that you understand how matrices are multiplied.[205]) The matrix I_n is called the *$n \times n$ identity matrix*. (Some call it E_n or just I, when the value of n is clear from the context.)

Matrix multiplication is associative: If $n, m, k, \ell \in \mathbb{N}$, and if A is an $n \times m$-matrix, B is an $m \times k$-matrix, and C is a $k \times \ell$-matrix, then $A(BC) = (AB)C$. The proof of this is straightforward using our definition of products of matrices[206]. This associativity allows us to write products like ABC without parentheses. By induction, we can see that longer products such as $A_1 A_2 \cdots A_k$ for arbitrary $k \in \mathbb{N}$ can also be bracketed at will, because all bracketings lead to the same result (e.g., for four matrices A, B, C and D, we have $A(B(CD)) = A((BC)D) = (AB)(CD) = (A(BC))D = ((AB)C)D$, provided that the dimensions of the matrices are appropriate to make sense of the products). We define an empty product of $n \times n$-matrices to be the $n \times n$ identity matrix I_n.

For every $n \times n$-matrix A and every $k \in \mathbb{N}$, we can thus define an $n \times n$-matrix A^k by $A^k = \underbrace{AA \cdots A}_{k \text{ factors}}$. In particular, $A^0 = I_n$ (since we defined an empty product of $n \times n$-matrices to be I_n).

Further properties of matrix multiplication are easy to state and to prove:

- For every $n \in \mathbb{N}$, $m \in \mathbb{N}$, $k \in \mathbb{N}$ and $\lambda \in \mathbb{K}$, every $n \times m$-matrix A and every $m \times k$-matrix B, we have $\lambda(AB) = (\lambda A)B = A(\lambda B)$. (This allows us to write λAB for each of the matrices $\lambda(AB)$, $(\lambda A)B$ and $A(\lambda B)$.)

[204] Here, 0 and 1 mean the zero and the unity of $\mathbb{K}$ (which may and may not be the integers 0 and 1).

[205] See [Grinbe16b, §2.12] for a detailed proof of the equality $AI_n = A$. (Interpret the word "number" in [Grinbe16b, §2.12] as "element of $\mathbb{K}$".) The proof of $I_n B = B$ is rather similar.

[206] Check that $A(BC)$ and $(AB)C$ both are equal to the matrix $\left(\sum_{u=1}^{m} \sum_{v=1}^{k} a_{i,u} b_{u,v} c_{v,j}\right)_{1 \leq i \leq n,\ 1 \leq j \leq \ell}$. For details of this proof, see [Grinbe16b, §2.9]. (Interpret the word "number" in [Grinbe16b, §2.9] as "element of $\mathbb{K}$".)

- For every $n \in \mathbb{N}$, $m \in \mathbb{N}$ and $k \in \mathbb{N}$, every two $n \times m$-matrices A and B, and every $m \times k$-matrix C, we have $(A+B)\,C = AC + BC$.

- For every $n \in \mathbb{N}$, $m \in \mathbb{N}$ and $k \in \mathbb{N}$, every $n \times m$-matrix A, and every two $m \times k$-matrices B and C, we have $A\,(B+C) = AB + AC$.

- For every $n \in \mathbb{N}$, $m \in \mathbb{N}$, $\lambda \in \mathbb{K}$ and $\mu \in \mathbb{K}$, and every $n \times m$-matrix A, we have $\lambda\,(\mu A) = (\lambda\mu)\,A$. (This allows us to write $\lambda\mu A$ for both $\lambda\,(\mu A)$ and $(\lambda\mu)\,A$.)

For given $n \in \mathbb{N}$ and $m \in \mathbb{N}$, we let $\mathbb{K}^{n\times m}$ denote the set of all $n \times m$-matrices. (This is one of the two standard notations for this set; the other is $\mathrm{M}_{n,m}(\mathbb{K})$.)

A *square matrix* is a matrix which has as many rows as it has columns; in other words, a square matrix is an $n \times n$-matrix for some $n \in \mathbb{N}$. If $A = (a_{i,j})_{1\leq i\leq n,\ 1\leq j\leq n}$ is a square matrix, then the n-tuple $(a_{1,1}, a_{2,2}, \ldots, a_{n,n})$ is called the *diagonal* of A. (Some authors abbreviate $(a_{i,j})_{1\leq i\leq n,\ 1\leq j\leq n}$ by $(a_{i,j})_{1\leq i,j\leq n}$; this notation has some mild potential for confusion, though[207].) The entries of the diagonal of A are called the *diagonal entries* of A. (Some authors like to say "main diagonal" instead of "diagonal".)

For a given $n \in \mathbb{N}$, the product of two $n \times n$-matrices is always well-defined, and is an $n \times n$-matrix again. The set $\mathbb{K}^{n\times n}$ satisfies all the axioms of a commutative ring except for commutativity of multiplication. This makes it into what is commonly called a *noncommutative ring*[208]. We shall study noncommutative rings later (in Section 6.17).

6.3. Determinants

Square matrices have determinants. Let us recall how determinants are defined:

Definition 6.5. Let $n \in \mathbb{N}$. Let $A = (a_{i,j})_{1\leq i\leq n,\ 1\leq j\leq n}$ be an $n \times n$-matrix. The *determinant* $\det A$ of A is defined as

$$\sum_{\sigma\in S_n} (-1)^{\sigma}\, a_{1,\sigma(1)} a_{2,\sigma(2)} \cdots a_{n,\sigma(n)}. \tag{340}$$

[207] The comma between "i" and "j" in "$1 \leq i,j \leq n$" can be understood either to separate i from j, or to separate the inequality $1 \leq i$ from the inequality $j \leq n$. I remember seeing this ambiguity causing a real misunderstanding.

[208] A *noncommutative ring* is defined in the same way as we defined a commutative ring, except for the fact that commutativity of multiplication is removed from the list of axioms. (The words "noncommutative ring" do not imply that commutativity of multiplication must be false for this ring; they merely say that commutativity of multiplication is **not required** to hold for it. For example, the noncommutative ring $\mathbb{K}^{n\times n}$ is actually commutative when $n \leq 1$ or when $\mathbb{K}$ is a trivial ring.)

Instead of saying "noncommutative ring", many algebraists just say "ring". We shall, however, keep the word "noncommutative" in order to avoid confusion.

In other words,

$$\det A = \sum_{\sigma \in S_n} (-1)^{\sigma} \underbrace{a_{1,\sigma(1)} a_{2,\sigma(2)} \cdots a_{n,\sigma(n)}}_{= \prod_{i=1}^{n} a_{i,\sigma(i)}} \tag{341}$$

$$= \sum_{\sigma \in S_n} (-1)^{\sigma} \prod_{i=1}^{n} a_{i,\sigma(i)}. \tag{342}$$

For example, the determinant of a 1×1-matrix $\left(a_{1,1} \right)$ is

$$\det \left(a_{1,1} \right) = \sum_{\sigma \in S_1} (-1)^{\sigma} a_{1,\sigma(1)} = \underbrace{(-1)^{\operatorname{id}}}_{=1} \underbrace{a_{1,\operatorname{id}(1)}}_{=a_{1,1}} \qquad \left(\text{since the only permutation } \sigma \in S_1 \text{ is } \operatorname{id}\right)$$
$$= a_{1,1}. \tag{343}$$

The determinant of a 2×2-matrix $\begin{pmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{pmatrix}$ is

$$\det \begin{pmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{pmatrix} = \sum_{\sigma \in S_2} (-1)^{\sigma} a_{1,\sigma(1)} a_{2,\sigma(2)}$$
$$= \underbrace{(-1)^{\operatorname{id}}}_{=1} \underbrace{a_{1,\operatorname{id}(1)}}_{=a_{1,1}} \underbrace{a_{2,\operatorname{id}(2)}}_{=a_{2,2}} + \underbrace{(-1)^{s_1}}_{=-1} \underbrace{a_{1,s_1(1)}}_{=a_{1,2}} \underbrace{a_{2,s_1(2)}}_{=a_{2,1}} \qquad \left(\text{since the only permutations } \sigma \in S_2 \text{ are } \operatorname{id} \text{ and } s_1\right)$$
$$= a_{1,1} a_{2,2} - a_{1,2} a_{2,1}.$$

Similarly, for a 3×3-matrix, the formula is

$$\det \begin{pmatrix} a_{1,1} & a_{1,2} & a_{1,3} \\ a_{2,1} & a_{2,2} & a_{2,3} \\ a_{3,1} & a_{3,2} & a_{3,3} \end{pmatrix} = a_{1,1} a_{2,2} a_{3,3} + a_{1,2} a_{2,3} a_{3,1} + a_{1,3} a_{2,1} a_{3,2}$$
$$- a_{1,1} a_{2,3} a_{3,2} - a_{1,2} a_{2,1} a_{3,3} - a_{1,3} a_{2,2} a_{3,1}. \tag{344}$$

Also, the determinant of the 0×0-matrix is 1 [209]. (This might sound like hair-

[209] In more details:

There is only one 0×0-matrix; it has no rows and no columns and no entries. According to (342), its determinant is

$$\sum_{\sigma \in S_0} (-1)^{\sigma} \underbrace{\prod_{i=1}^{0} a_{i,\sigma(i)}}_{=(\text{empty product})=1} = \sum_{\sigma \in S_0} (-1)^{\sigma} = (-1)^{\operatorname{id}} \qquad \left(\text{since the only } \sigma \in S_0 \text{ is } \operatorname{id}\right)$$
$$= 1.$$

splitting, but being able to work with 0×0-matrices simplifies some proofs by induction, because it allows one to take $n = 0$ as an induction base.)

The equality (342) (or, equivalently, (341)) is known as the *Leibniz formula*. Out of several known ways to define the determinant, it is probably the most direct. In practice, however, computing a determinant using (342) quickly becomes impractical when n is high (since the sum has $n!$ terms). In most situations that occur both in mathematics and in applications, determinants can be computed in various simpler ways.

Some authors write $|A|$ instead of $\det A$ for the determinant of a square matrix A. I do not like this notation, as it clashes (in the case of 1×1-matrices) with the notation $|a|$ for the absolute value of a real number a.

Here is a first example of a determinant which ends up very simple:

Example 6.6. Let $n \in \mathbb{N}$. Let $x_1, x_2, \ldots, x_n$ be n elements of $\mathbb{K}$, and let $y_1, y_2, \ldots, y_n$ be n further elements of $\mathbb{K}$. Let A be the $n \times n$-matrix $\left(x_i y_j\right)_{1 \leq i \leq n,\ 1 \leq j \leq n}$. What is $\det A$?

For $n = 0$, we have $\det A = 1$ (since the 0×0-matrix has determinant 1).

For $n = 1$, we have $A = \left(\begin{array}{c} x_1 y_1 \end{array} \right)$ and thus $\det A = x_1 y_1$.

For $n = 2$, we have $A = \left(\begin{array}{cc} x_1 y_1 & x_1 y_2 \\ x_2 y_1 & x_2 y_2 \end{array} \right)$ and thus $\det A = \left(x_1 y_1\right) \left(x_2 y_2\right) - \left(x_1 y_2\right) \left(x_2 y_1\right) = 0$.

What do you expect for greater values of n ? The pattern might not be clear at this point yet, but if you compute further examples, you will realize that $\det A = 0$ also holds for $n = 3$, for $n = 4$, for $n = 5$... This suggests that $\det A = 0$ for every $n \geq 2$. How to prove this?

Let $n \geq 2$. Then, (341) (applied to $a_{i,j} = x_i y_j$) yields

$$\begin{aligned}
\det A &= \sum_{\sigma \in S_n} (-1)^\sigma \underbrace{\left(x_1 y_{\sigma(1)}\right) \left(x_2 y_{\sigma(2)}\right) \cdots \left(x_n y_{\sigma(n)}\right)}_{= \left(x_1 x_2 \cdots x_n\right) \left(y_{\sigma(1)} y_{\sigma(2)} \cdots y_{\sigma(n)}\right)} \\
&= \sum_{\sigma \in S_n} (-1)^\sigma \left(x_1 x_2 \cdots x_n\right) \underbrace{\left(y_{\sigma(1)} y_{\sigma(2)} \cdots y_{\sigma(n)}\right)}_{\substack{= y_1 y_2 \cdots y_n \\ \text{(since } \sigma \text{ is a permutation)}}} \\
&= \sum_{\sigma \in S_n} (-1)^\sigma \left(x_1 x_2 \cdots x_n\right) \left(y_1 y_2 \cdots y_n\right) \\
&= \left(\sum_{\sigma \in S_n} (-1)^\sigma \right) \left(x_1 x_2 \cdots x_n\right) \left(y_1 y_2 \cdots y_n\right) . \qquad (345)
\end{aligned}$$

Now, every $\sigma \in S_n$ is either even or odd (but not both), and thus we have

$$\begin{aligned}
&\sum_{\sigma \in S_n} (-1)^\sigma \\
&= \sum_{\substack{\sigma \in S_n; \\ \sigma \text{ is even}}} \underbrace{(-1)^\sigma}_{\substack{=1 \\ (\text{since } \sigma \text{ is even})}} + \sum_{\substack{\sigma \in S_n; \\ \sigma \text{ is odd}}} \underbrace{(-1)^\sigma}_{\substack{=-1 \\ (\text{since } \sigma \text{ is odd})}} \\
&= \underbrace{\sum_{\substack{\sigma \in S_n; \\ \sigma \text{ is even}}} 1}_{=(\text{the number of even permutations } \sigma \in S_n) \cdot 1} + \underbrace{\sum_{\substack{\sigma \in S_n; \\ \sigma \text{ is odd}}} (-1)}_{=(\text{the number of odd permutations } \sigma \in S_n) \cdot (-1)} \\
&= \underbrace{(\text{the number of even permutations } \sigma \in S_n)}_{\substack{=n!/2 \\ (\text{by Exercise 5.4})}} \cdot 1 \\
&\qquad + \underbrace{(\text{the number of odd permutations } \sigma \in S_n)}_{\substack{=n!/2 \\ (\text{by Exercise 5.4})}} \cdot (-1) \\
&= (n!/2) \cdot 1 + (n!/2) \cdot (-1) = 0.
\end{aligned}$$

Hence, (345) becomes $\det A = \underbrace{\left(\sum_{\sigma \in S_n} (-1)^\sigma \right)}_{=0} (x_1 x_2 \cdots x_n) (y_1 y_2 \cdots y_n) = 0$, as we wanted to prove.

We will eventually learn a simpler way to prove this.

Example 6.7. Here is an example similar to Example 6.6, but subtler.

Let $n \in \mathbb{N}$. Let $x_1, x_2, \ldots, x_n$ be n elements of $\mathbb{K}$, and let $y_1, y_2, \ldots, y_n$ be n further elements of $\mathbb{K}$. Let A be the $n \times n$-matrix $\left(x_i + y_j\right)_{1 \leq i \leq n,\ 1 \leq j \leq n}$. What is $\det A$?

For $n = 0$, we have $\det A = 1$ again.

For $n = 1$, we have $A = \begin{pmatrix} x_1 + y_1 \end{pmatrix}$ and thus $\det A = x_1 + y_1$.

For $n = 2$, we have $A = \begin{pmatrix} x_1 + y_1 & x_1 + y_2 \\ x_2 + y_1 & x_2 + y_2 \end{pmatrix}$ and thus $\det A = (x_1 + y_1)(x_2 + y_2) - (x_1 + y_2)(x_2 + y_1) = -(y_1 - y_2)(x_1 - x_2)$.

However, it turns out that for every $n \geq 3$, we again have $\det A = 0$. This is harder to prove than the similar claim in Example 6.6. We will eventually see how to do it easily, but as for now let me outline a direct proof. (I am being rather telegraphic here; do not worry if you do not understand the following argument, as there will be easier and more detailed proofs below.)

From (341), we obtain

$$\det A = \sum_{\sigma \in S_n} (-1)^\sigma \left(x_1 + y_{\sigma(1)}\right) \left(x_2 + y_{\sigma(2)}\right) \cdots \left(x_n + y_{\sigma(n)}\right). \tag{346}$$

If we expand the product $\left(x_1 + y_{\sigma(1)}\right)\left(x_2 + y_{\sigma(2)}\right)\cdots\left(x_n + y_{\sigma(n)}\right)$, we obtain a sum of 2^n terms:

$$\begin{aligned}&\left(x_1 + y_{\sigma(1)}\right)\left(x_2 + y_{\sigma(2)}\right)\cdots\left(x_n + y_{\sigma(n)}\right)\\&=\sum_{I\subseteq[n]}\left(\prod_{i\in I}x_i\right)\left(\prod_{i\in[n]\setminus I}y_{\sigma(i)}\right)\end{aligned} \tag{347}$$

(where $[n]$ means the set $\{1,2,\ldots,n\}$). (To obtain a fully rigorous proof of (347), apply Exercise 6.1 **(a)** to $a_i = x_i$ and $b_i = y_{\sigma(i)}$.) Thus, (346) becomes

$$\begin{aligned}\det A &= \sum_{\sigma\in S_n}(-1)^\sigma \underbrace{\left(x_1 + y_{\sigma(1)}\right)\left(x_2 + y_{\sigma(2)}\right)\cdots\left(x_n + y_{\sigma(n)}\right)}_{\substack{=\sum_{I\subseteq[n]}\left(\prod_{i\in I}x_i\right)\left(\prod_{i\in[n]\setminus I}y_{\sigma(i)}\right)\\ \text{(by (347))}}}\\&=\sum_{\sigma\in S_n}(-1)^\sigma\sum_{I\subseteq[n]}\left(\prod_{i\in I}x_i\right)\left(\prod_{i\in[n]\setminus I}y_{\sigma(i)}\right)\\&=\sum_{I\subseteq[n]}\sum_{\sigma\in S_n}(-1)^\sigma\left(\prod_{i\in I}x_i\right)\left(\prod_{i\in[n]\setminus I}y_{\sigma(i)}\right).\end{aligned}$$

We want to prove that this is 0. In order to do so, it clearly suffices to show that every $I \subseteq [n]$ satisfies

$$\sum_{\sigma\in S_n}(-1)^\sigma\left(\prod_{i\in I}x_i\right)\left(\prod_{i\in[n]\setminus I}y_{\sigma(i)}\right)=0. \tag{348}$$

So let us fix $I \subseteq [n]$, and try to prove (348). We must be in one of the following two cases:

Case 1: The set $[n] \setminus I$ has at least two elements. In this case, let us pick two distinct elements a and b of this set, and split the set S_n into disjoint two-element subsets by pairing up every even permutation $\sigma \in S_n$ with the odd permutation $\sigma \circ t_{a,b}$ (where $t_{a,b}$ is as defined in Definition 5.29). The addends on the left hand side of (348) corresponding to two permutations paired up cancel out each other (because the products $\prod_{i\in[n]\setminus I} y_{\sigma(i)}$ and $\prod_{i\in[n]\setminus I} y_{\left(\sigma\circ t_{a,b}\right)(i)}$ differ only in the order of their factors), and thus the whole left hand side of (348) is 0. Thus, (348) is proven in Case 1.

Case 2: The set $[n] \setminus I$ has at most one element. In this case, the set I has at least two elements (it is here that we use $n \geq 3$). Pick two distinct elements c and d of I, and split the set S_n into disjoint two-element subsets by pairing up every even permutation $\sigma \in S_n$ with the odd permutation $\sigma \circ t_{c,d}$. Again, the addends on the left hand side of (348) corresponding to two permutations paired up cancel out each other (because the products $\prod\limits_{i \in [n] \setminus I} y_{\sigma(i)}$ and $\prod\limits_{i \in [n] \setminus I} y_{\left(\sigma \circ t_{c,d}\right)(i)}$ are identical), and thus the whole left hand side of (348) is 0. This proves (348) in Case 2.

We thus have proven (348) in both cases. So $\det A = 0$ is proven. This was a tricky argument, and shows the limits of the usefulness of (341).

We shall now discuss basic properties of the determinant.

Exercise 6.3. Let $A = \left(a_{i,j}\right)_{1 \leq i \leq n,\ 1 \leq j \leq n}$ be an $n \times n$-matrix. Assume that $a_{i,j} = 0$ for every $(i,j) \in \{1,2,\ldots,n\}^2$ satisfying $i < j$. Show that

$$\det A = a_{1,1} a_{2,2} \cdots a_{n,n}.$$

Definition 6.8. An $n \times n$-matrix A satisfying the assumption of Exercise 6.3 is said to be *lower-triangular* (because its entries above the diagonal are 0, and thus its nonzero entries are concentrated in the triangular region southwest of the diagonal). Exercise 6.3 thus says that the determinant of a lower-triangular matrix is the product of its diagonal entries. For instance, $\det \begin{pmatrix} a & 0 & 0 \\ b & c & 0 \\ d & e & f \end{pmatrix} = acf$.

Example 6.9. Let $n \in \mathbb{N}$. The $n \times n$ identity matrix I_n is lower-triangular, and its diagonal entries are $1, 1, \ldots, 1$. Hence, Exercise 6.3 shows that its determinant is $\det(I_n) = 1 \cdot 1 \cdot \cdots \cdot 1 = 1$.

Definition 6.10. The *transpose* of a matrix $A = \left(a_{i,j}\right)_{1 \leq i \leq n,\ 1 \leq j \leq m}$ is defined to be the matrix $\left(a_{j,i}\right)_{1 \leq i \leq m,\ 1 \leq j \leq n}$. It is denoted by A^T. For instance, $\begin{pmatrix} 1 & 2 & -1 \\ 4 & 0 & 1 \end{pmatrix}^T = \begin{pmatrix} 1 & 4 \\ 2 & 0 \\ -1 & 1 \end{pmatrix}$.

Remark 6.11. Various other notations for the transpose of a matrix A exist in the literature. Some of them are A^t (with a lower case t) and ${}^T A$ and ${}^t A$.

Exercise 6.4. Let $n \in \mathbb{N}$. Let A be an $n \times n$-matrix. Show that $\det\left(A^T\right) = \det A$.

The transpose of a lower-triangular $n \times n$-matrix is an upper-triangular $n \times n$-matrix (i.e., an $n \times n$-matrix whose entries below the diagonal are 0). Thus, combining Exercise 6.3 with Exercise 6.4, we see that the determinant of an upper-triangular matrix is the product of its diagonal entries.

The following exercise presents five fundamental (and simple) properties of transposes:

Exercise 6.5. Prove the following:

(a) If u, v and w are three nonnegative integers, if P is a $u \times v$-matrix, and if Q is a $v \times w$-matrix, then
$$(PQ)^T = Q^T P^T. \tag{349}$$

(b) Every $u \in \mathbb{N}$ satisfies
$$(I_u)^T = I_u. \tag{350}$$

(c) If u and v are two nonnegative integers, if P is a $u \times v$-matrix, and if $\lambda \in \mathbb{K}$, then
$$(\lambda P)^T = \lambda P^T. \tag{351}$$

(d) If u and v are two nonnegative integers, and if P and Q are two $u \times v$-matrices, then
$$(P+Q)^T = P^T + Q^T.$$

(e) If u and v are two nonnegative integers, and if P is a $u \times v$-matrix, then
$$\left(P^T\right)^T = P. \tag{352}$$

Here is yet another simple property of determinants that follows directly from their definition:

Proposition 6.12. Let $n \in \mathbb{N}$ and $\lambda \in \mathbb{K}$. Let A be an $n \times n$-matrix. Then, $\det(\lambda A) = \lambda^n \det A$.

Proof of Proposition 6.12. Write A in the form $A = \left(a_{i,j}\right)_{1 \leq i \leq n,\ 1 \leq j \leq n}$. Thus, $\lambda A = \left(\lambda a_{i,j}\right)_{1 \leq i \leq n,\ 1 \leq j \leq n}$ (by the definition of λA). Hence, (342) (applied to λA and $\lambda a_{i,j}$

instead of A and $a_{i,j}$) yields

$$\begin{aligned}\det(\lambda A) &= \sum_{\sigma \in S_n} (-1)^{\sigma} \underbrace{\prod_{i=1}^{n} \left(\lambda a_{i,\sigma(i)} \right)}_{= \lambda^n \prod\limits_{i=1}^{n} a_{i,\sigma(i)}} = \sum_{\sigma \in S_n} (-1)^{\sigma} \lambda^n \prod_{i=1}^{n} a_{i,\sigma(i)} \\ &= \lambda^n \underbrace{\sum_{\sigma \in S_n} (-1)^{\sigma} \prod_{i=1}^{n} a_{i,\sigma(i)}}_{\substack{= \det A \\ \text{(by (342))}}} = \lambda^n \det A.\end{aligned}$$

Proposition 6.12 is thus proven. □

Exercise 6.6. Let $a, b, c, d, e, f, g, h, i, j, k, \ell, m, n, o, p$ be elements of $\mathbb{K}$.

(a) Find a simple formula for the determinant

$$\det \begin{pmatrix} a & b & c & d \\ \ell & 0 & 0 & e \\ k & 0 & 0 & f \\ j & i & h & g \end{pmatrix}.$$

(b) Find a simple formula for the determinant

$$\det \begin{pmatrix} a & b & c & d & e \\ f & 0 & 0 & 0 & g \\ h & 0 & 0 & 0 & i \\ j & 0 & 0 & 0 & k \\ \ell & m & n & o & p \end{pmatrix}.$$

(Do not mistake the "o" for a "0".)

[**Hint:** Part **(b)** is simpler than part **(a)**.]

In the next exercises, we shall talk about rows and columns; let us first make some pedantic remarks about these notions.

If $n \in \mathbb{N}$, then an $n \times 1$-matrix is said to be a *column vector* with n entries[210], whereas a $1 \times n$-matrix is said to be a *row vector* with n entries. Column vectors and row vectors store exactly the same kind of data (namely, n elements of $\mathbb{K}$), so you might wonder why I make a difference between them (and also why I distinguish them from n-tuples of elements of $\mathbb{K}$, which also contain precisely the same kind of data). The reason for this is that column vectors and row vectors behave differently under matrix multiplication: For example,

$$\begin{pmatrix} a \\ b \end{pmatrix} \begin{pmatrix} c & d \end{pmatrix} = \begin{pmatrix} ac & ad \\ bc & bd \end{pmatrix}$$

[210] It is also called a *column vector* of size n.

is not the same as

$$\left(\begin{array}{cc} a & b \end{array}\right)\left(\begin{array}{c} c \\ d \end{array}\right)=\left(\begin{array}{c} ac+bd \end{array}\right).$$

If we would identify column vectors with row vectors, then this would cause contradictions.

The reason to distinguish between row vectors and n-tuples is subtler: We have defined row vectors only for a commutative ring $\mathbb{K}$, whereas n-tuples can be made out of elements of any set. As a consequence, the sum of two row vectors is well-defined (since row vectors are matrices and thus can be added entry by entry), whereas the sum of two n-tuples is not. Similarly, we can take the product λv of an element $\lambda \in \mathbb{K}$ with a row vector v (by multiplying every entry of v by λ), but such a thing does not make sense for general n-tuples. These differences between row vectors and n-tuples, however, cause no clashes of notation if we use the same notations for both types of object. Thus, we are often going to identify a row vector $\left(\begin{array}{cccc} a_1 & a_2 & \cdots & a_n \end{array}\right)$ with the n-tuple $(a_1, a_2, \ldots, a_n) \in \mathbb{K}^n$. Thus, $\mathbb{K}^n$ becomes the set of all row vectors with n entries.[211]

The column vectors with n entries are in 1-to-1 correspondence with the row vectors with n entries, and this correspondence is given by taking the transpose: The column vector v corresponds to the row vector v^T, and conversely, the row vector w corresponds to the column vector w^T. In particular, every column vector $\left(\begin{array}{c} a_1 \\ a_2 \\ \vdots \\ a_n \end{array}\right)$ can be rewritten in the form $\left(\begin{array}{cccc} a_1 & a_2 & \cdots & a_n \end{array}\right)^T = (a_1, a_2, \ldots, a_n)^T$. We shall often write it in the latter form, just because it takes up less space on paper.

The rows of a matrix are row vectors; the columns of a matrix are column vectors. Thus, terms like "the sum of two rows of a matrix A" or "-3 times a column of a matrix A" make sense: Rows and columns are vectors, and thus can be added (when they have the same number of entries) and multiplied by elements of $\mathbb{K}$.

Let $n \in \mathbb{N}$ and $j \in \{1, 2, \ldots, n\}$. If v is a column vector with n entries (that is, an $n \times 1$-matrix), then the *j-th entry of v* means the $(j, 1)$-th entry of v. If v is a row vector with n entries (that is, a $1 \times n$-matrix), then the *j-th entry of v* means the $(1, j)$-th entry of v. For example, the 2-nd entry of the row vector $\left(\begin{array}{ccc} a & b & c \end{array}\right)$ is b.

Exercise 6.7. Let $n \in \mathbb{N}$. Let A be an $n \times n$-matrix. Prove the following:

(a) If B is an $n \times n$-matrix obtained from A by swapping two rows, then $\det B = -\det A$. ("Swapping two rows" means "swapping two distinct rows", of course.)

(b) If B is an $n \times n$-matrix obtained from A by swapping two columns, then $\det B = -\det A$.

[211] Some algebraists, instead, identify column vectors with n-tuples, so that $\mathbb{K}^n$ is then the set of all column vectors with n entries. This is a valid convention as well, but one must be careful not to use it simultaneously with our convention (i.e., with the convention that row vectors are identified with n-tuples); this is why we will not use it.

(c) If a row of A consists of zeroes, then $\det A = 0$.

(d) If a column of A consists of zeroes, then $\det A = 0$.

(e) If A has two equal rows, then $\det A = 0$.

(f) If A has two equal columns, then $\det A = 0$.

(g) Let $\lambda \in \mathbb{K}$ and $k \in \{1, 2, \ldots, n\}$. If B is the $n \times n$-matrix obtained from A by multiplying the k-th row by λ (that is, multiplying every entry of the k-th row by λ), then $\det B = \lambda \det A$.

(h) Let $\lambda \in \mathbb{K}$ and $k \in \{1, 2, \ldots, n\}$. If B is the $n \times n$-matrix obtained from A by multiplying the k-th column by λ, then $\det B = \lambda \det A$.

(i) Let $k \in \{1, 2, \ldots, n\}$. Let A' be an $n \times n$-matrix whose rows equal the corresponding rows of A except (perhaps) the k-th row. Let B be the $n \times n$-matrix obtained from A by adding the k-th row of A' to the k-th row of A (that is, by adding every entry of the k-th row of A' to the corresponding entry of the k-th row of A). Then, $\det B = \det A + \det A'$.

(j) Let $k \in \{1, 2, \ldots, n\}$. Let A' be an $n \times n$-matrix whose columns equal the corresponding columns of A except (perhaps) the k-th column. Let B be the $n \times n$-matrix obtained from A by adding the k-th column of A' to the k-th column of A. Then, $\det B = \det A + \det A'$.

Example 6.13. Let us show examples for several parts of Exercise 6.7 (especially, for Exercise 6.7 **(i)**, which has a somewhat daunting statement).

(a) Exercise 6.7 **(a)** yields (among other things) that

$$\det \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix} = -\det \begin{pmatrix} g & h & i \\ d & e & f \\ a & b & c \end{pmatrix}$$

for any $a, b, c, d, e, f, g, h, i \in \mathbb{K}$.

(c) Exercise 6.7 **(c)** yields (among other things) that

$$\det \begin{pmatrix} a & b & c \\ 0 & 0 & 0 \\ d & e & f \end{pmatrix} = 0$$

for any $a, b, c, d, e, f \in \mathbb{K}$.

(e) Exercise 6.7 **(e)** yields (among other things) that

$$\det \begin{pmatrix} a & b & c \\ d & e & f \\ d & e & f \end{pmatrix} = 0$$

for any $a, b, c, d, e, f \in \mathbb{K}$.

(g) Exercise 6.7 **(g)** (applied to $n = 3$ and $k = 2$) yields that

$$\det \begin{pmatrix} a & b & c \\ \lambda d & \lambda e & \lambda f \\ g & h & i \end{pmatrix} = \lambda \det \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix}$$

for any $a, b, c, d, e, f \in \mathbb{K}$ and $\lambda \in \mathbb{K}$.

(i) Set $n = 3$ and $k = 2$. Set $A = \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix}$. Then, a matrix A' satisfying the conditions of Exercise 6.7 **(i)** has the form $A' = \begin{pmatrix} a & b & c \\ d' & e' & f' \\ g & h & i \end{pmatrix}$. For such a matrix A', we obtain $B = \begin{pmatrix} a & b & c \\ d+d' & e+e' & f+f' \\ g & h & i \end{pmatrix}$. Exercise 6.7 **(i)** then claims that $\det B = \det A + \det A'$.

Parts **(a)**, **(c)**, **(e)**, **(g)** and **(i)** of Exercise 6.7 are often united under the slogan "the determinant of a matrix is multilinear and alternating in its rows"[212]. Similarly, parts **(b)**, **(d)**, **(f)**, **(h)** and **(j)** are combined under the slogan "the determinant of a matrix is multilinear and alternating in its columns". Many texts on linear algebra (for example, [HofKun71]) use these properties as the **definition** of the determinant[213]; this is a valid approach, but I prefer to use Definition 6.5 instead, since it is more explicit.

Exercise 6.8. Let $n \in \mathbb{N}$. Let A be an $n \times n$-matrix. Prove the following:

(a) If we add a scalar multiple of a row of A to another row of A, then the determinant of A does not change. (A *scalar multiple* of a row vector v means a row vector of the form λv, where $\lambda \in \mathbb{K}$.)

(b) If we add a scalar multiple of a column of A to another column of A, then the determinant of A does not change. (A *scalar multiple* of a column vector v means a column vector of the form λv, where $\lambda \in \mathbb{K}$.)

Example 6.14. Let us visualize Exercise 6.8 **(a)**. Set $n = 3$ and $A = \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix}$. If we add -2 times the second row of A to the first row of A, then we obtain the matrix $\begin{pmatrix} a + (-2)\,d & b + (-2)\,e & c + (-2)\,f \\ d & e & f \\ g & h & i \end{pmatrix}$. Exercise 6.8 **(a)** now claims that this new matrix has the same determinant as A (because -2 times the second

[212] Specifically, parts **(c)**, **(g)** and **(i)** say that it is "multilinear", while parts **(a)** and **(e)** are responsible for the "alternating".

[213] More precisely, they define a *determinant function* to be a function $F : \mathbb{K}^{n \times n} \to \mathbb{K}$ which is multilinear and alternating in the rows of a matrix (i.e., which satisfies parts **(a)**, **(c)**, **(e)**, **(g)** and **(i)** of Exercise 6.7 if every appearance of "det" is replaced by "F" in this Exercise) and which satisfies $F(I_n) = 1$. Then, they show that there is (for each $n \in \mathbb{N}$) exactly one determinant function $F : \mathbb{K}^{n \times n} \to \mathbb{K}$. They then denote this function by det. This is a rather slick definition of a determinant, but it has the downside that it requires showing that there is exactly one determinant function (which is often not easier than our approach).

row of A is a scalar multiple of the second row of A).

Notice the word "another" in Exercise 6.8. Adding a scalar multiple of a row of A to **the same** row of A will likely change the determinant.

Remark 6.15. Exercise 6.8 lets us prove the claim of Example 6.7 in a much simpler way.

Namely, let n and $x_1, x_2, \ldots, x_n$ and $y_1, y_2, \ldots, y_n$ and A be as in Example 6.7. Assume that $n \geq 3$. We want to show that $\det A = 0$.

The matrix A has at least three rows (since $n \geq 3$), and looks as follows:

$$A = \begin{pmatrix} x_1 + y_1 & x_1 + y_2 & x_1 + y_3 & x_1 + y_4 & \cdots & x_1 + y_n \\ x_2 + y_1 & x_2 + y_2 & x_2 + y_3 & x_2 + y_4 & \cdots & x_2 + y_n \\ x_3 + y_1 & x_3 + y_2 & x_3 + y_3 & x_3 + y_4 & \cdots & x_3 + y_n \\ x_4 + y_1 & x_4 + y_2 & x_4 + y_3 & x_4 + y_4 & \cdots & x_4 + y_n \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ x_n + y_1 & x_n + y_2 & x_n + y_3 & x_n + y_4 & \cdots & x_n + y_n \end{pmatrix}$$

(where the presence of terms like x_4 and y_4 does not mean that the variables x_4 and y_4 exist, in the same way as one can write "$x_1, x_2, \ldots, x_k$" even if $k = 1$ or $k = 0$). Thus, if we subtract the first row of A from the second row of A, then we obtain the matrix

$$A' = \begin{pmatrix} x_1 + y_1 & x_1 + y_2 & x_1 + y_3 & x_1 + y_4 & \cdots & x_1 + y_n \\ x_2 - x_1 & x_2 - x_1 & x_2 - x_1 & x_2 - x_1 & \cdots & x_2 - x_1 \\ x_3 + y_1 & x_3 + y_2 & x_3 + y_3 & x_3 + y_4 & \cdots & x_3 + y_n \\ x_4 + y_1 & x_4 + y_2 & x_4 + y_3 & x_4 + y_4 & \cdots & x_4 + y_n \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ x_n + y_1 & x_n + y_2 & x_n + y_3 & x_n + y_4 & \cdots & x_n + y_n \end{pmatrix}$$

(because $\left(x_2 + y_j\right) - \left(x_1 + y_j\right) = x_2 - x_1$ for every j). The transformation we just did (subtracting a row from another row) does not change the determinant of the matrix (by Exercise 6.8 **(a)**, because subtracting a row from another row is tantamount to adding the (-1)-multiple of the former row to the latter), and thus we have $\det A' = \det A$.

We notice that each entry of the second row of A' equals $x_2 - x_1$.

Next, we subtract the first row of A' from the third row of A', and obtain the matrix

$$A'' = \begin{pmatrix} x_1 + y_1 & x_1 + y_2 & x_1 + y_3 & x_1 + y_4 & \cdots & x_1 + y_n \\ x_2 - x_1 & x_2 - x_1 & x_2 - x_1 & x_2 - x_1 & \cdots & x_2 - x_1 \\ x_3 - x_1 & x_3 - x_1 & x_3 - x_1 & x_3 - x_1 & \cdots & x_3 - x_1 \\ x_4 + y_1 & x_4 + y_2 & x_4 + y_3 & x_4 + y_4 & \cdots & x_4 + y_n \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ x_n + y_1 & x_n + y_2 & x_n + y_3 & x_n + y_4 & \cdots & x_n + y_n \end{pmatrix}.$$

Again, the determinant is unchanged (because of Exercise 6.8 **(a)**), so we have $\det A'' = \det A' = \det A$.

We notice that each entry of the second row of A'' equals $x_2 - x_1$ (indeed, these entries have been copied over from A'), and that each entry of the third row of A'' equals $x_3 - x_1$.

Next, we subtract the first column of A'' from each of the other columns of A''. This gives us the matrix

$$A''' = \begin{pmatrix} x_1 + y_1 & y_2 - y_1 & y_3 - y_1 & y_4 - y_1 & \cdots & y_n - y_1 \\ x_2 - x_1 & 0 & 0 & 0 & \cdots & 0 \\ x_3 - x_1 & 0 & 0 & 0 & \cdots & 0 \\ x_4 + y_1 & y_2 - y_1 & y_3 - y_1 & y_4 - y_1 & \cdots & y_n - y_1 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ x_n + y_1 & y_2 - y_1 & y_3 - y_1 & y_4 - y_1 & \cdots & y_n - y_1 \end{pmatrix}. \qquad (353)$$

This step, again, has not changed the determinant (because Exercise 6.8 **(b)** shows that subtracting a column from another column does not change the determinant, and what we did was doing $n-1$ such transformations). Thus, $\det A''' = \det A'' = \det A$.

Now, let us write the matrix A''' in the form $A''' = \left(a'''_{i,j}\right)_{1\leq i\leq n,\ 1\leq j\leq n}$. (Thus, $a'''_{i,j}$ is the (i,j)-th entry of A''' for every (i,j).) Then, (341) (applied to A''' instead of A) yields

$$\det A''' = \sum_{\sigma\in S_n} (-1)^\sigma a'''_{1,\sigma(1)} a'''_{2,\sigma(2)} \cdots a'''_{n,\sigma(n)}. \qquad (354)$$

I claim that

$$a'''_{1,\sigma(1)} a'''_{2,\sigma(2)} \cdots a'''_{n,\sigma(n)} = 0 \qquad \text{for every } \sigma \in S_n. \qquad (355)$$

[*Proof of (355):* Let $\sigma \in S_n$. Then, σ is injective, and thus $\sigma(2) \neq \sigma(3)$. Therefore, at least one of the integers $\sigma(2)$ and $\sigma(3)$ must be $\neq 1$ (because otherwise, we would have $\sigma(2) = 1 = \sigma(3)$, contradicting $\sigma(2) \neq \sigma(3)$). We WLOG assume that $\sigma(2) \neq 1$. But a look at (353) reveals that all entries of the second row of A''' are zero except for the first entry. Thus, $a'''_{2,j} = 0$ for every $j \neq 1$. Applied to $j = \sigma(2)$, this yields $a'''_{2,\sigma(2)} = 0$ (since $\sigma(2) \neq 1$). Hence, $a'''_{1,\sigma(1)} a'''_{2,\sigma(2)} \cdots a'''_{n,\sigma(n)} = 0$ (because if 0 appears as a factor in a product, then the whole product must be 0). This proves (355).]

Now, (354) becomes

$$\det A''' = \sum_{\sigma\in S_n} (-1)^\sigma \underbrace{a'''_{1,\sigma(1)} a'''_{2,\sigma(2)} \cdots a'''_{n,\sigma(n)}}_{\substack{=0 \\ \text{(by (355))}}} = \sum_{\sigma\in S_n} (-1)^\sigma 0 = 0.$$

Compared with $\det A''' = \det A$, this yields $\det A = 0$. Thus, $\det A = 0$ is proven again.

Remark 6.16. Here is another example for the use of Exercise 6.8.

Let $n \in \mathbb{N}$. Let $x_1, x_2, \ldots, x_n$ be n elements of $\mathbb{K}$. Let A be the matrix $\left(x_{\max\{i,j\}}\right)_{1\leq i\leq n,\ 1\leq j\leq n}$. (Recall that $\max S$ denotes the greatest element of a nonempty set S.)

For example, if $n = 4$, then

$$A = \begin{pmatrix} x_1 & x_2 & x_3 & x_4 \\ x_2 & x_2 & x_3 & x_4 \\ x_3 & x_3 & x_3 & x_4 \\ x_4 & x_4 & x_4 & x_4 \end{pmatrix}.$$

We want to find $\det A$. First, let us subtract the first row of A from each of the other rows of A. Thus we obtain a new matrix A'. The determinant has not changed (according to Exercise 6.8 **(a)**); i.e., we have $\det A' = \det A$. Here is how A' looks like in the case when $n = 4$:

$$A' = \begin{pmatrix} x_1 & x_2 & x_3 & x_4 \\ x_2 - x_1 & 0 & 0 & 0 \\ x_3 - x_1 & x_3 - x_2 & 0 & 0 \\ x_4 - x_1 & x_4 - x_2 & x_4 - x_3 & 0 \end{pmatrix}. \qquad (356)$$

Notice the many zeroes; zeroes are useful when computing determinants. To generalize the pattern we see on (356), we write the matrix A' in the form $A' = \left(a'_{i,j}\right)_{1\leq i\leq n,\ 1\leq j\leq n}$ (so that $a'_{i,j}$ is the (i,j)-th entry of A' for every (i,j)). Then, for every $(i,j) \in \{1,2,\ldots,n\}^2$, we have

$$a'_{i,j} = \begin{cases} x_{\max\{i,j\}}, & \text{if } i = 1; \\ x_{\max\{i,j\}} - x_{\max\{1,j\}}, & \text{if } i > 1 \end{cases} \qquad (357)$$

(since we obtained the matrix A' by subtracting the first row of A from each of the other rows of A). Hence, for every $(i,j) \in \{1,2,\ldots,n\}^2$ satisfying $1 < i \leq j$, we have

$$a'_{i,j} = x_{\max\{i,j\}} - x_{\max\{1,j\}} = x_j - x_j \qquad \left(\begin{array}{c}\text{since } \max\{i,j\} = j \text{ (because } i \leq j) \\ \text{and } \max\{1,j\} = j \text{ (because } 1 < j)\end{array}\right)$$
$$= 0. \qquad (358)$$

This is the general explanation for the six 0's in (356). We notice also that the first row of the matrix A' is $(x_1, x_2, \ldots, x_n)$.

Now, we want to transform A' further. Namely, we first swap the first row with the second row; then we swap the second row (which used to be the first row) with the third row; then, the third row with the fourth row, and so on, until we finally swap the $(n-1)$-th row with the n-th row. As a result of these $n-1$

swaps, the first row has moved all the way down to the bottom, past all the other rows. We denote the resulting matrix by A''. For instance, if $n = 4$, then

$$A'' = \begin{pmatrix} x_2 - x_1 & 0 & 0 & 0 \\ x_3 - x_1 & x_3 - x_2 & 0 & 0 \\ x_4 - x_1 & x_4 - x_2 & x_4 - x_3 & 0 \\ x_1 & x_2 & x_3 & x_4 \end{pmatrix}. \tag{359}$$

This is a lower-triangular matrix. To see that this holds in the general case, we write the matrix A'' in the form $A'' = \left(a''_{i,j}\right)_{1\leq i\leq n,\ 1\leq j\leq n}$ (so that $a''_{i,j}$ is the (i,j)-th entry of A'' for every (i,j)). Then, for every $(i,j) \in \{1,2,\ldots,n\}^2$, we have

$$a''_{i,j} = \begin{cases} a'_{i+1,j'} & \text{if } i < n; \\ a'_{1,j'} & \text{if } i = n \end{cases} \tag{360}$$

(because the first row of A' has become the n-th row of A'', whereas every other row has moved up one step). In particular, for every $(i,j) \in \{1,2,\ldots,n\}^2$ satisfying $1 \leq i < j \leq n$, we have

$$\begin{aligned} a''_{i,j} &= \begin{cases} a'_{i+1,j'} & \text{if } i < n; \\ a'_{1,j'} & \text{if } i = n \end{cases} = a'_{i+1,j} \qquad (\text{since } i < j \leq n) \\ &= 0 \qquad \left(\begin{array}{c} \text{by (358), applied to } i+1 \text{ instead of } i \\ (\text{because } i < j \text{ yields } i+1 \leq j) \end{array}\right). \end{aligned}$$

This shows that A'' is indeed lower-triangular. Hence, Exercise 6.3 (applied to A'' and $a''_{i,j}$ instead of A and $a_{i,j}$) shows that $\det A'' = a''_{1,1}a''_{2,2}\cdots a''_{n,n}$.

Using (360) and (357), it is easy to see that every $i \in \{1,2,\ldots,n\}$ satisfies

$$a''_{i,i} = \begin{cases} x_{i+1} - x_i, & \text{if } i < n; \\ x_n, & \text{if } i = n \end{cases}. \tag{361}$$

(This is precisely the pattern you would guess from the diagonal entries in (359).) Now, multiplying the equalities (361) for all $i \in \{1,2,\ldots,n\}$, we obtain $a''_{1,1}a''_{2,2}\cdots a''_{n,n} = (x_2 - x_1)(x_3 - x_2)\cdots(x_n - x_{n-1})x_n$. Thus,

$$\det A'' = a''_{1,1}a''_{2,2}\cdots a''_{n,n} = (x_2 - x_1)(x_3 - x_2)\cdots(x_n - x_{n-1})x_n. \tag{362}$$

But we want $\det A$, not $\det A''$. First, let us find $\det A'$. Recall that A'' was obtained from A' by swapping rows, repeatedly – namely, $n-1$ times. Every time we swap two rows in a matrix, its determinant gets multiplied by -1 (because of Exercise 6.7 **(a)**). Hence, $n-1$ such swaps cause the determinant to be multiplied by $(-1)^{n-1}$. Since A'' was obtained from A' by $n-1$ such swaps, we thus conclude that $\det A'' = (-1)^{n-1} \det A'$, so that

$$\det A' = \underbrace{\frac{1}{(-1)^{n-1}}}_{=(-1)^{n-1}} \underbrace{\det A''}_{=(x_2-x_1)(x_3-x_2)\cdots(x_n-x_{n-1})x_n}$$
$$= (-1)^{n-1} (x_2 - x_1)(x_3 - x_2) \cdots (x_n - x_{n-1}) x_n.$$

Finally, recall that $\det A' = \det A$, so that

$$\det A = \det A' = (-1)^{n-1} (x_2 - x_1)(x_3 - x_2) \cdots (x_n - x_{n-1}) x_n.$$

6.4. $\det(AB)$

Next, a lemma that will come handy in a more important proof:

Lemma 6.17. Let $n \in \mathbb{N}$. Let $[n]$ denote the set $\{1, 2, \ldots, n\}$. Let $\kappa : [n] \to [n]$ be a map. Let $B = (b_{i,j})_{1 \leq i \leq n,\ 1 \leq j \leq n}$ be an $n \times n$-matrix. Let B_κ be the $n \times n$-matrix $\left(b_{\kappa(i),j}\right)_{1 \leq i \leq n,\ 1 \leq j \leq n}$.

(a) If $\kappa \in S_n$, then $\det(B_\kappa) = (-1)^\kappa \cdot \det B$.

(b) If $\kappa \notin S_n$, then $\det(B_\kappa) = 0$.

Remark 6.18. Lemma 6.17 **(a)** simply says that if we permute the rows of a square matrix, then its determinant gets multiplied by the sign of the permutation used. For instance, let $n = 3$ and $B = \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix}$. If κ is the permutation $(2, 3, 1)$ (in one-line notation), then $B_\kappa = \begin{pmatrix} d & e & f \\ g & h & i \\ a & b & c \end{pmatrix}$, and Lemma 6.17 **(a)** says that $\det(B_\kappa) = \underbrace{(-1)^\kappa}_{=1} \cdot \det B = \det B$.

Of course, a similar result holds for permutations of columns.

Remark 6.19. Exercise 6.7 **(a)** is a particular case of Lemma 6.17 **(a)**. Indeed, if B is an $n \times n$-matrix obtained from A by swapping the u-th and the v-th row (where u and v are two distinct elements of $\{1, 2, \ldots, n\}$), then $B = \left(a_{t_{u,v}(i),j}\right)_{1 \leq i \leq n,\ 1 \leq j \leq n}$ (where A is written in the form $A = (a_{i,j})_{1 \leq i \leq n,\ 1 \leq j \leq n}$).

Proof of Lemma 6.17. Recall that S_n is the set of all permutations of $\{1,2,\ldots,n\}$. In other words, S_n is the set of all permutations of $[n]$ (since $[n] = \{1,2,\ldots,n\}$). In other words, S_n is the set of all bijective maps $[n] \to [n]$.

(a) Assume that $\kappa \in S_n$. We define a map $\Phi : S_n \to S_n$ by

$$\Phi(\sigma) = \sigma \circ \kappa \qquad \text{for every } \sigma \in S_n.$$

We also define a map $\Psi : S_n \to S_n$ by

$$\Psi(\sigma) = \sigma \circ \kappa^{-1} \qquad \text{for every } \sigma \in S_n.$$

The maps Φ and Ψ are mutually inverse[214]. Hence, the map Φ is a bijection.

We have $B = (b_{i,j})_{1\leq i\leq n,\ 1\leq j\leq n}$. Hence, (342) (applied to B and $b_{i,j}$ instead of A and $a_{i,j}$) yields

$$\det B = \sum_{\sigma\in S_n} (-1)^\sigma \underbrace{\prod_{i=1}^{n}}_{=\prod_{i\in[n]}} b_{i,\sigma(i)} = \sum_{\sigma\in S_n} (-1)^\sigma \prod_{i\in[n]} b_{i,\sigma(i)}. \tag{363}$$

Now, $B_\kappa = \left(b_{\kappa(i),j}\right)_{1\leq i\leq n,\ 1\leq j\leq n}$. Hence, (342) (applied to B_κ and $b_{\kappa(i),j}$ instead of A and $a_{i,j}$) yields

$$\begin{aligned}\det(B_\kappa) &= \sum_{\sigma\in S_n} (-1)^\sigma \underbrace{\prod_{i=1}^{n}}_{=\prod_{i\in[n]}} b_{\kappa(i),\sigma(i)} = \sum_{\sigma\in S_n} (-1)^\sigma \prod_{i\in[n]} b_{\kappa(i),\sigma(i)} \\ &= \sum_{\sigma\in S_n} (-1)^{\Phi(\sigma)} \prod_{i\in[n]} b_{\kappa(i),(\Phi(\sigma))(i)} \end{aligned} \tag{364}$$

(here, we have substituted $\Phi(\sigma)$ for σ in the sum, since Φ is a bijection).

But every $\sigma \in S_n$ satisfies $(-1)^{\Phi(\sigma)} = (-1)^\kappa \cdot (-1)^\sigma$ [215] and $\prod_{i\in[n]} b_{\kappa(i),(\Phi(\sigma))(i)} =$

[214] *Proof.* Every $\sigma \in S_n$ satisfies

$$\begin{aligned}(\Psi\circ\Phi)(\sigma) &= \Psi\left(\underbrace{\Phi(\sigma)}_{=\sigma\circ\kappa}\right) = \Psi(\sigma\circ\kappa) = \sigma\circ\underbrace{\kappa\circ\kappa^{-1}}_{=\mathrm{id}} \qquad \text{(by the definition of } \Psi) \\ &= \sigma = \mathrm{id}(\sigma).\end{aligned}$$

Thus, $\Psi\circ\Phi = \mathrm{id}$. Similarly, $\Phi\circ\Psi = \mathrm{id}$. Combined with $\Psi\circ\Phi = \mathrm{id}$, this yields that the maps Φ and Ψ are mutually inverse, qed.

[215] *Proof.* Let $\sigma \in S_n$. Then, $\Phi(\sigma) = \sigma\circ\kappa$, so that

$$\begin{aligned}(-1)^{\Phi(\sigma)} &= (-1)^{\sigma\circ\kappa} = (-1)^\sigma \cdot (-1)^\kappa \qquad \text{(by (315), applied to } \tau = \kappa) \\ &= (-1)^\kappa \cdot (-1)^\sigma,\end{aligned}$$

qed.

$\prod_{i\in[n]} b_{i,\sigma(i)}$ [216]. Thus, (364) becomes

$$\begin{aligned}\det\left(B_\kappa\right) &= \sum_{\sigma\in S_n} \underbrace{(-1)^{\Phi(\sigma)}}_{=(-1)^\kappa\cdot(-1)^\sigma} \underbrace{\prod_{i\in[n]} b_{\kappa(i),(\Phi(\sigma))(i)}}_{=\prod_{i\in[n]} b_{i,\sigma(i)}} = \sum_{\sigma\in S_n} (-1)^\kappa \cdot (-1)^\sigma \prod_{i\in[n]} b_{i,\sigma(i)} \\ &= (-1)^\kappa \cdot \underbrace{\sum_{\sigma\in S_n} (-1)^\sigma \prod_{i\in[n]} b_{i,\sigma(i)}}_{\substack{=\det B \\ \text{(by (363))}}} = (-1)^\kappa \cdot \det B.\end{aligned}$$

This proves Lemma 6.17 **(a)**.

(b) Assume that $\kappa \notin S_n$.

The following fact is well-known: If U is a finite set, then every injective map $U \to U$ is bijective[217]. We can apply this to $U = [n]$, and thus conclude that every injective map $[n] \to [n]$ is bijective. Therefore, if the map $\kappa : [n] \to [n]$ were injective, then κ would be bijective and therefore would be an element of S_n (since S_n is the set of all bijective maps $[n] \to [n]$); but this would contradict the fact that $\kappa \notin S_n$. Hence, the map $\kappa : [n] \to [n]$ cannot be injective. Therefore, there exist two distinct elements a and b of $[n]$ such that $\kappa(a) = \kappa(b)$. Consider these a and b.

Thus, a and b are two distinct elements of $[n] = \{1, 2, \ldots, n\}$. Hence, a transposition $t_{a,b} \in S_n$ is defined (see Definition 5.29 for the definition). This transposition satisfies $\kappa \circ t_{a,b} = \kappa$ [218]. Exercise 5.10 **(b)** (applied to $i = a$ and $j = b$) yields $(-1)^{t_{a,b}} = -1$.

[216] *Proof.* Let $\sigma \in S_n$. We have $\Phi(\sigma) = \sigma \circ \kappa$. Thus, for every $i \in [n]$, we have $(\Phi(\sigma))(i) = (\sigma \circ \kappa)(i) = \sigma(\kappa(i))$. Hence, $\prod_{i\in[n]} b_{\kappa(i),(\Phi(\sigma))(i)} = \prod_{i\in[n]} b_{\kappa(i),\sigma(\kappa(i))}$.

But $\kappa \in S_n$. In other words, κ is a permutation of the set $\{1, 2, \ldots, n\} = [n]$, hence a bijection from $[n]$ to $[n]$. Therefore, we can substitute $\kappa(i)$ for i in the product $\prod_{i\in[n]} b_{i,\sigma(i)}$. We thus obtain $\prod_{i\in[n]} b_{i,\sigma(i)} = \prod_{i\in[n]} b_{\kappa(i),\sigma(\kappa(i))}$. Comparing this with $\prod_{i\in[n]} b_{\kappa(i),(\Phi(\sigma))(i)} = \prod_{i\in[n]} b_{\kappa(i),\sigma(\kappa(i))}$, we obtain $\prod_{i\in[n]} b_{\kappa(i),(\Phi(\sigma))(i)} = \prod_{i\in[n]} b_{i,\sigma(i)}$, qed.

[217] *Proof.* Let U be a finite set, and let f be an injective map $U \to U$. We must show that f is bijective.

Since f is injective, we have $|f(U)| = |U|$. Thus, $f(U)$ is a subset of U which has size $|U|$. But the only such subset is U itself (since U is a finite set). Therefore, $f(U)$ must be U itself. In other words, the map f is surjective. Hence, f is bijective (since f is injective and surjective), qed.

[218] *Proof.* We are going to show that every $i \in [n]$ satisfies $(\kappa \circ t_{a,b})(i) = \kappa(i)$.

So let $i \in [n]$. We shall show that $(\kappa \circ t_{a,b})(i) = \kappa(i)$.

The definition of $t_{a,b}$ shows that $t_{a,b}$ is the permutation in S_n which swaps a with b while leaving all other elements of $\{1, 2, \ldots, n\}$ unchanged. In other words, we have $t_{a,b}(a) = b$, and $t_{a,b}(b) = a$, and $t_{a,b}(j) = j$ for every $j \in [n] \setminus \{a, b\}$.

Now, we have $i \in [n]$. Thus, we are in one of the following three cases:

Case 1: We have $i = a$.

Case 2: We have $i = b$.

Case 3: We have $i \in [n] \setminus \{a, b\}$.

Let A_n be the set of all even permutations in S_n. Let C_n be the set of all odd permutations in S_n.

We have $\sigma \circ t_{a,b} \in C_n$ for every $\sigma \in A_n$ [219]. Hence, we can define a map $\Phi : A_n \to C_n$ by

$$\Phi(\sigma) = \sigma \circ t_{a,b} \qquad \text{for every } \sigma \in A_n.$$

Consider this map Φ. Furthermore, we have $\sigma \circ (t_{a,b})^{-1} \in A_n$ for every $\sigma \in C_n$ [220]. Thus, we can define a map $\Psi : C_n \to A_n$ by

$$\Psi(\sigma) = \sigma \circ (t_{a,b})^{-1} \qquad \text{for every } \sigma \in C_n.$$

Consider this map Ψ.

Let us first consider Case 1. In this case, we have $i = a$, so that $(\kappa \circ t_{a,b})\left(\underbrace{i}_{=a}\right) = (\kappa \circ t_{a,b})(a) = \kappa\left(\underbrace{t_{a,b}(a)}_{=b}\right) = \kappa(b)$. Compared with $\kappa\left(\underbrace{i}_{=a}\right) = \kappa(a) = \kappa(b)$, this yields $(\kappa \circ t_{a,b})(i) = \kappa(i)$. Thus, $(\kappa \circ t_{a,b})(i) = \kappa(i)$ is proven in Case 1.

Let us next consider Case 2. In this case, we have $i = b$, so that $(\kappa \circ t_{a,b})\left(\underbrace{i}_{=b}\right) = (\kappa \circ t_{a,b})(b) = \kappa\left(\underbrace{t_{a,b}(b)}_{=a}\right) = \kappa(a) = \kappa(b)$. Compared with $\kappa\left(\underbrace{i}_{=b}\right) = \kappa(b)$, this yields $(\kappa \circ t_{a,b})(i) = \kappa(i)$. Thus, $(\kappa \circ t_{a,b})(i) = \kappa(i)$ is proven in Case 2.

Let us finally consider Case 3. In this case, we have $i \in [n] \setminus \{a,b\}$. Hence, $t_{a,b}(i) = i$ (since $t_{a,b}(j) = j$ for every $j \in [n] \setminus \{a,b\}$). Therefore, $(\kappa \circ t_{a,b})(i) = \kappa\left(\underbrace{t_{a,b}(i)}_{=i}\right) = \kappa(i)$. Thus, $(\kappa \circ t_{a,b})(i) = \kappa(i)$ is proven in Case 3.

We now have shown $(\kappa \circ t_{a,b})(i) = \kappa(i)$ in each of the three Cases 1, 2 and 3. Hence, $(\kappa \circ t_{a,b})(i) = \kappa(i)$ always holds.

Now, let us forget that we fixed i. We thus have shown that $(\kappa \circ t_{a,b})(i) = \kappa(i)$ for every $i \in [n]$. In other words, $\kappa \circ t_{a,b} = \kappa$, qed.

[219] *Proof.* Let $\sigma \in A_n$. Then, σ is an even permutation in S_n (since A_n is the set of all even permutations in S_n). Hence, $(-1)^\sigma = 1$. Now, (315) (applied to $\tau = t_{a,b}$) yields $(-1)^{\sigma \circ t_{a,b}} = \underbrace{(-1)^\sigma}_{=1} \cdot \underbrace{(-1)^{t_{a,b}}}_{=-1} = -1$. Thus, the permutation $\sigma \circ t_{a,b}$ is odd. Hence, $\sigma \circ t_{a,b}$ is an odd permutation in S_n. In other words, $\sigma \circ t_{a,b} \in C_n$ (since C_n is the set of all odd permutations in S_n), qed.

[220] *Proof.* Let $\sigma \in C_n$. Then, σ is an odd permutation in S_n (since C_n is the set of all odd permutations in S_n). Hence, $(-1)^\sigma = -1$.

Applying (316) to $t_{a,b}$ instead of σ, we obtain $(-1)^{(t_{a,b})^{-1}} = (-1)^{t_{a,b}} = -1$. Now, (315) (applied to $\tau = (t_{a,b})^{-1}$) yields $(-1)^{\sigma \circ (t_{a,b})^{-1}} = \underbrace{(-1)^\sigma}_{=-1} \cdot \underbrace{(-1)^{(t_{a,b})^{-1}}}_{=-1} = (-1) \cdot (-1) = 1$. Thus, the permutation $\sigma \circ (t_{a,b})^{-1}$ is even. Hence, $\sigma \circ (t_{a,b})^{-1}$ is an even permutation in S_n. In other words, $\sigma \circ (t_{a,b})^{-1} \in A_n$ (since A_n is the set of all even permutations in S_n), qed.

(We could have simplified our life a bit by noticing that $\left(t_{a,b}\right)^{-1} = t_{a,b}$, so that the maps Φ and Ψ are given by the same formula, albeit defined on different domains. But I wanted to demonstrate a use of (316).)

The maps Φ and Ψ are mutually inverse[221]. Hence, the map Ψ is a bijection. Moreover, every $\sigma \in C_n$ satisfies

$$\prod_{i\in[n]} b_{\kappa(i),(\Psi(\sigma))(i)} = \prod_{i\in[n]} b_{\kappa(i),\sigma(i)}. \tag{365}$$

222

We have $B_\kappa = \left(b_{\kappa(i),j}\right)_{1\leq i\leq n,\ 1\leq j\leq n}$. Hence, (342) (applied to B_κ and $b_{\kappa(i),j}$ instead

[221] *Proof.* Every $\sigma \in A_n$ satisfies

$$\begin{aligned}\left(\Psi\circ\Phi\right)(\sigma) &= \Psi\left(\underbrace{\Phi(\sigma)}_{=\sigma\circ\tau_{a,b}}\right) = \Psi\left(\sigma\circ\tau_{a,b}\right) = \sigma\circ\underbrace{\tau_{a,b}\circ\left(\tau_{a,b}\right)^{-1}}_{=\mathrm{id}} \qquad \text{(by the definition of } \Psi\text{)}\\ &= \sigma = \mathrm{id}(\sigma).\end{aligned}$$

Thus, $\Psi\circ\Phi = \mathrm{id}$. Similarly, $\Phi\circ\Psi = \mathrm{id}$. Combined with $\Psi\circ\Phi = \mathrm{id}$, this yields that the maps Φ and Ψ are mutually inverse, qed.

[222] *Proof of (365):* Let $\sigma \in C_n$. The map $t_{a,b}$ is a permutation of $[n]$, thus a bijection $[n]\to[n]$. Hence, we can substitute $t_{a,b}(i)$ for i in the product $\prod\limits_{i\in[n]} b_{\kappa(i),(\Psi(\sigma))(i)}$. Thus we obtain

$$\prod_{i\in[n]} b_{\kappa(i),(\Psi(\sigma))(i)} = \prod_{i\in[n]} b_{\kappa\left(t_{a,b}(i)\right),(\Psi(\sigma))\left(t_{a,b}(i)\right)} = \prod_{i\in[n]} b_{\kappa(i),\sigma(i)}$$

(since every $i\in[n]$ satisfies $\kappa\left(t_{a,b}(i)\right) = \underbrace{\left(\kappa\circ t_{a,b}\right)}_{=\kappa}(i) = \kappa(i)$ and

$$\underbrace{(\Psi(\sigma))}_{=\sigma\circ\left(t_{a,b}\right)^{-1}}\left(t_{a,b}(i)\right) = \left(\sigma\circ\left(t_{a,b}\right)^{-1}\right)\left(t_{a,b}(i)\right) = \sigma\left(\underbrace{\left(t_{a,b}\right)^{-1}\left(t_{a,b}(i)\right)}_{=i}\right) = \sigma(i)$$

). This proves (365).

of A and $a_{i,j}$) yields

$$\begin{aligned}
\det\left(B_\kappa\right) &= \sum_{\sigma\in S_n} (-1)^\sigma \underbrace{\prod_{i=1}^{n}}_{=\prod_{i\in[n]}} b_{\kappa(i),\sigma(i)} = \sum_{\sigma\in S_n} (-1)^\sigma \prod_{i\in[n]} b_{\kappa(i),\sigma(i)} \\
&= \underbrace{\sum_{\substack{\sigma\in S_n;\\ \sigma \text{ is even}}}}_{\substack{=\sum_{\sigma\in A_n}\\ \text{(since } A_n \text{ is the}\\ \text{set of all even}\\ \text{permutations}\\ \text{in } S_n)}} \underbrace{(-1)^\sigma}_{\substack{=1\\ \text{(since } \sigma \text{ is even)}}} \prod_{i\in[n]} b_{\kappa(i),\sigma(i)} + \underbrace{\sum_{\substack{\sigma\in S_n;\\ \sigma \text{ is odd}}}}_{\substack{=\sum_{\sigma\in C_n}\\ \text{(since } C_n \text{ is the}\\ \text{set of all odd}\\ \text{permutations}\\ \text{in } S_n)}} \underbrace{(-1)^\sigma}_{\substack{=-1\\ \text{(since } \sigma \text{ is odd)}}} \prod_{i\in[n]} b_{\kappa(i),\sigma(i)} \\
&\qquad \left(\begin{array}{c}\text{since every permutation } \sigma\in S_n \text{ is}\\ \text{either even or odd, but not both}\end{array}\right) \\
&= \sum_{\sigma\in A_n} \prod_{i\in[n]} b_{\kappa(i),\sigma(i)} + \sum_{\sigma\in C_n} (-1) \prod_{i\in[n]} b_{\kappa(i),\sigma(i)} \\
&= \sum_{\sigma\in A_n} \prod_{i\in[n]} b_{\kappa(i),\sigma(i)} - \sum_{\sigma\in C_n} \prod_{i\in[n]} b_{\kappa(i),\sigma(i)} = 0,
\end{aligned}$$

since

$$\begin{aligned}
&\sum_{\sigma\in A_n} \prod_{i\in[n]} b_{\kappa(i),\sigma(i)} \\
&= \sum_{\sigma\in C_n} \underbrace{\prod_{i\in[n]} b_{\kappa(i),(\Psi(\sigma))(i)}}_{\substack{=\prod_{i\in[n]} b_{\kappa(i),\sigma(i)}\\ \text{(by (365))}}} \\
&\qquad \left(\text{here, we have substituted } \Psi\left(\sigma\right) \text{ for } \sigma, \text{ since the map } \Psi \text{ is a bijection}\right) \\
&= \sum_{\sigma\in C_n} \prod_{i\in[n]} b_{\kappa(i),\sigma(i)}.
\end{aligned}$$

This proves Lemma 6.17 **(b)**. □

Now let us state a basic formula for products of sums in a commutative ring:

Lemma 6.20. For every $n\in\mathbb{N}$, let $[n]$ denote the set $\{1,2,\ldots,n\}$.

Let $n\in\mathbb{N}$. For every $i\in[n]$, let $p_{i,1},p_{i,2},\ldots,p_{i,m_i}$ be finitely many elements of $\mathbb{K}$. Then,

$$\prod_{i=1}^{n}\sum_{k=1}^{m_i} p_{i,k} = \sum_{(k_1,k_2,\ldots,k_n)\in[m_1]\times[m_2]\times\cdots\times[m_n]} \prod_{i=1}^{n} p_{i,k_i}.$$

(**Pedantic remark:** If $n=0$, then the Cartesian product $[m_1]\times[m_2]\times\cdots\times[m_n]$ has no factors; it is what is called an *empty Cartesian product*. It is understood to be a 1-element set, and its single element is the 0-tuple $()$ (also known as the empty list).)

I tend to refer to Lemma 6.20 as the *product rule* (since it is related to the product rule for joint probabilities); I think it has no really widespread name. However, it is a fundamental algebraic fact that is used very often and tacitly (I suspect that most mathematicians have never thought of it as being a theorem that needs to be proven). The idea behind Lemma 6.20 is that if you expand the product

$$\begin{aligned}
&\prod_{i=1}^{n} \sum_{k=1}^{m_i} p_{i,k} \\
&= \prod_{i=1}^{n} \left(p_{i,1} + p_{i,2} + \cdots + p_{i,m_i}\right) \\
&= \left(p_{1,1} + p_{1,2} + \cdots + p_{1,m_1}\right) \left(p_{2,1} + p_{2,2} + \cdots + p_{2,m_2}\right) \cdots \left(p_{n,1} + p_{n,2} + \cdots + p_{n,m_n}\right),
\end{aligned}$$

then you get a sum of $m_1 m_2 \cdots m_n$ terms, each of which has the form

$$p_{1,k_1} p_{2,k_2} \cdots p_{n,k_n} = \prod_{i=1}^{n} p_{i,k_i}$$

for some $(k_1, k_2, \ldots, k_n) \in [m_1] \times [m_2] \times \cdots \times [m_n]$. (More precisely, it is the sum of all such terms.) A formal proof of Lemma 6.20 could be obtained by induction over n using the distributivity axiom[223]. For the details (if you care about them), see the solution to the following exercise:

Exercise 6.9. Prove Lemma 6.20.

Remark 6.21. Lemma 6.20 can be regarded as a generalization of Exercise 6.1 **(a)**. Indeed, let me sketch how Exercise 6.1 **(a)** can be derived from Lemma 6.20:

Let n, $(a_1, a_2, \ldots, a_n)$ and $(b_1, b_2, \ldots, b_n)$ be as in Exercise 6.1 **(a)**. For every $i \in [n]$, set $m_i = 2$, $p_{i,1} = a_i$ and $p_{i,2} = b_i$. Then, Lemma 6.20 yields

$$\begin{aligned}
\prod_{i=1}^{n} (a_i + b_i) &= \underbrace{\sum_{(k_1,k_2,\ldots,k_n) \in \underbrace{[2] \times [2] \times \cdots \times [2]}_{n \text{ factors}}}}_{= \sum_{(k_1,k_2,\ldots,k_n) \in [2]^n}} \underbrace{\prod_{i=1}^{n} p_{i,k_i}}_{= \left(\prod_{\substack{i \in [n]; \\ k_i = 1}} a_i\right) \left(\prod_{\substack{i \in [n]; \\ k_i = 2}} b_i\right)} \\
&= \sum_{(k_1,k_2,\ldots,k_n) \in [2]^n} \left(\prod_{\substack{i \in [n]; \\ k_i = 1}} a_i\right) \left(\prod_{\substack{i \in [n]; \\ k_i = 2}} b_i\right).
\end{aligned} \tag{366}$$

[223] and the observation that the n-tuples $(k_1, k_2, \ldots, k_n) \in [m_1] \times [m_2] \times \cdots \times [m_n]$ are in bijection with the pairs $((k_1, k_2, \ldots, k_{n-1}), k_n)$ of an $(n-1)$-tuple $(k_1, k_2, \ldots, k_{n-1}) \in [m_1] \times [m_2] \times \cdots \times [m_{n-1}]$ and an element $k_n \in [m_n]$

But there is a bijection between the set $[2]^n$ and the powerset $\mathcal{P}([n])$ of $[n]$: Namely, to every n-tuple $(k_1, k_2, \ldots, k_n) \in [2]^n$, we can assign the set $\{i \in [n] \mid k_i = 1\} \in \mathcal{P}([n])$. It is easy to see that this assignment really is a bijection $[2]^n \to \mathcal{P}([n])$, and that it furthermore has the property that every n-tuple $(k_1, k_2, \ldots, k_n) \in [2]^n$ satisfies

$$\left(\prod_{\substack{i\in[n];\\ k_i=1}} a_i\right)\left(\prod_{\substack{i\in[n];\\ k_i=2}} b_i\right) = \left(\prod_{i\in I} a_i\right)\left(\prod_{i\in[n]\setminus I} b_i\right),$$

where I is the image of $(k_1, k_2, \ldots, k_n)$ under this bijection. Hence,

$$\sum_{(k_1,k_2,\ldots,k_n)\in[2]^n} \left(\prod_{\substack{i\in[n];\\ k_i=1}} a_i\right)\left(\prod_{\substack{i\in[n];\\ k_i=2}} b_i\right) = \sum_{I\subseteq[n]} \left(\prod_{i\in I} a_i\right)\left(\prod_{i\in[n]\setminus I} b_i\right).$$

Hence, (366) rewrites as

$$\prod_{i=1}^{n} (a_i + b_i) = \sum_{I\subseteq[n]} \left(\prod_{i\in I} a_i\right)\left(\prod_{i\in[n]\setminus I} b_i\right).$$

But this is precisely the claim of Exercise 6.1 **(a)**.

We shall use a corollary of Lemma 6.20:

Lemma 6.22. For every $n \in \mathbb{N}$, let $[n]$ denote the set $\{1, 2, \ldots, n\}$.

Let $n \in \mathbb{N}$ and $m \in \mathbb{N}$. For every $i \in [n]$, let $p_{i,1}, p_{i,2}, \ldots, p_{i,m}$ be m elements of $\mathbb{K}$. Then,

$$\prod_{i=1}^{n} \sum_{k=1}^{m} p_{i,k} = \sum_{\kappa:[n]\to[m]} \prod_{i=1}^{n} p_{i,\kappa(i)}.$$

Proof of Lemma 6.22. For the sake of completeness, let us give this proof.

Lemma 6.20 (applied to $m_i = m$ for every $i \in [n]$) yields

$$\prod_{i=1}^{n} \sum_{k=1}^{m} p_{i,k} = \sum_{(k_1,k_2,\ldots,k_n)\in\underbrace{[m]\times[m]\times\cdots\times[m]}_{n \text{ factors}}} \prod_{i=1}^{n} p_{i,k_i}. \qquad (367)$$

Let $\operatorname{Map}([n],[m])$ denote the set of all functions from $[n]$ to $[m]$. Now, let Φ be the map from $\operatorname{Map}([n],[m])$ to $\underbrace{[m]\times[m]\times\cdots\times[m]}_{n\text{ factors}}$ given by

$$\Phi(\kappa)=(\kappa(1),\kappa(2),\ldots,\kappa(n)) \qquad \text{for every } \kappa\in\operatorname{Map}([n],[m]).$$

So the map Φ takes a function κ from $[n]$ to $[m]$, and outputs the list $(\kappa(1),\kappa(2),\ldots,\kappa(n))$ of all its values. Clearly, the map Φ is injective (since a function $\kappa\in\operatorname{Map}([n],[m])$ can be reconstructed from the list $(\kappa(1),\kappa(2),\ldots,\kappa(n))=\Phi(\kappa)$) and surjective (since every list of n elements of $[m]$ is the list of values of some function $\kappa\in\operatorname{Map}([n],[m])$). Thus, Φ is bijective. Therefore, we can substitute $\Phi(\kappa)$ for $(k_1,k_2,\ldots,k_n)$ in the sum $\sum\limits_{(k_1,k_2,\ldots,k_n)\in\underbrace{[m]\times[m]\times\cdots\times[m]}_{n\text{ factors}}}\prod\limits_{i=1}^{n}p_{i,k_i}$. In other words, we can substitute $(\kappa(1),\kappa(2),\ldots,\kappa(n))$ for $(k_1,k_2,\ldots,k_n)$ in this sum (since $\Phi(\kappa)=(\kappa(1),\kappa(2),\ldots,\kappa(n))$ for each $\kappa\in\operatorname{Map}([n],[m])$). We thus obtain

$$\sum_{(k_1,k_2,\ldots,k_n)\in\underbrace{[m]\times[m]\times\cdots\times[m]}_{n\text{ factors}}}\prod_{i=1}^{n}p_{i,k_i}=\underbrace{\sum_{\kappa\in\operatorname{Map}([n],[m])}}_{=\sum_{\kappa:[n]\to[m]}}\prod_{i=1}^{n}p_{i,\kappa(i)}$$
$$=\sum_{\kappa:[n]\to[m]}\prod_{i=1}^{n}p_{i,\kappa(i)}.$$

Thus, (367) becomes

$$\prod_{i=1}^{n}\sum_{k=1}^{m}p_{i,k}=\sum_{(k_1,k_2,\ldots,k_n)\in\underbrace{[m]\times[m]\times\cdots\times[m]}_{n\text{ factors}}}\prod_{i=1}^{n}p_{i,k_i}=\sum_{\kappa:[n]\to[m]}\prod_{i=1}^{n}p_{i,\kappa(i)}.$$

Lemma 6.22 is proven. □

Now we are ready to prove what is probably the most important property of determinants, known as the *multiplicativity of the determinant*[224]:

Theorem 6.23. Let $n\in\mathbb{N}$. Let A and B be two $n\times n$-matrices. Then,

$$\det(AB)=\det A\cdot\det B.$$

[224] Theorem 6.23 appears, e.g., in [HofKun71, §5.3, Theorem 3], in [Laue15, proof of Theorem 5.7], in [Strick13, Theorem B.17], in [Mate14, "Multiplications of determinants", Lemma], in [Pinkha15, Theorem 11.4.2], in [GalQua22, Proposition 7.9], in [Zeilbe85, §5], and in [Loehr11, Theorem 9.54].

Proof of Theorem 6.23. Write A and B in the forms $A = \left(a_{i,j}\right)_{1\leq i\leq n,\ 1\leq j\leq n}$ and $B = \left(b_{i,j}\right)_{1\leq i\leq n,\ 1\leq j\leq n}$. The definition of AB thus yields $AB = \left(\sum\limits_{k=1}^{n} a_{i,k}b_{k,j}\right)_{1\leq i\leq n,\ 1\leq j\leq n}$. Therefore, (342) (applied to AB and $\sum\limits_{k=1}^{n} a_{i,k}b_{k,j}$ instead of A and $a_{i,j}$) yields

$$\begin{aligned}
\det\left(AB\right) &= \sum_{\sigma\in S_n} (-1)^{\sigma} \underbrace{\prod_{i=1}^{n}\left(\sum_{k=1}^{n} a_{i,k}b_{k,\sigma(i)}\right)}_{\substack{=\sum\limits_{\kappa:[n]\to[n]}\prod\limits_{i=1}^{n}\left(a_{i,\kappa(i)}b_{\kappa(i),\sigma(i)}\right) \\ \text{(by Lemma 6.22, applied to } m=n \\ \text{and } p_{i,k}=a_{i,k}b_{k,\sigma(i)})}} \\
&= \sum_{\sigma\in S_n} (-1)^{\sigma} \sum_{\kappa:[n]\to[n]} \underbrace{\prod_{i=1}^{n}\left(a_{i,\kappa(i)}b_{\kappa(i),\sigma(i)}\right)}_{=\left(\prod\limits_{i=1}^{n} a_{i,\kappa(i)}\right)\left(\prod\limits_{i=1}^{n} b_{\kappa(i),\sigma(i)}\right)} \\
&= \sum_{\sigma\in S_n} (-1)^{\sigma} \sum_{\kappa:[n]\to[n]} \left(\prod_{i=1}^{n} a_{i,\kappa(i)}\right)\left(\prod_{i=1}^{n} b_{\kappa(i),\sigma(i)}\right) \\
&= \sum_{\kappa:[n]\to[n]} \left(\prod_{i=1}^{n} a_{i,\kappa(i)}\right)\left(\sum_{\sigma\in S_n} (-1)^{\sigma} \prod_{i=1}^{n} b_{\kappa(i),\sigma(i)}\right). \qquad (368)
\end{aligned}$$

Now, for every $\kappa : [n] \to [n]$, we let B_κ be the $n \times n$-matrix $\left(b_{\kappa(i),j}\right)_{1\leq i\leq n,\ 1\leq j\leq n}$. Then, for every $\kappa : [n] \to [n]$, the equality (342) (applied to B_κ and $b_{\kappa(i),j}$ instead of A and $a_{i,j}$) yields

$$\det\left(B_\kappa\right) = \sum_{\sigma\in S_n} (-1)^{\sigma} \prod_{i=1}^{n} b_{\kappa(i),\sigma(i)}. \qquad (369)$$

Thus, (368) becomes

$$\begin{aligned}
\det(AB) &= \sum_{\kappa:[n]\to[n]} \left(\prod_{i=1}^n a_{i,\kappa(i)}\right) \underbrace{\left(\sum_{\sigma\in S_n} (-1)^\sigma \prod_{i=1}^n b_{\kappa(i),\sigma(i)}\right)}_{\substack{=\det(B_\kappa) \\ \text{(by (369))}}} \\
&= \sum_{\kappa:[n]\to[n]} \left(\prod_{i=1}^n a_{i,\kappa(i)}\right) \det(B_\kappa) \\
&= \underbrace{\sum_{\substack{\kappa:[n]\to[n];\\ \kappa\in S_n}}}_{\substack{=\sum_{\kappa\in S_n} \\ \text{(since every } \kappa\in S_n \text{ automatically} \\ \text{is a map } [n]\to[n])}} \left(\prod_{i=1}^n a_{i,\kappa(i)}\right) \underbrace{\det(B_\kappa)}_{\substack{=(-1)^\kappa\cdot\det B \\ \text{(by Lemma 6.17 (a))}}} \\
&\qquad + \sum_{\substack{\kappa:[n]\to[n];\\ \kappa\notin S_n}} \left(\prod_{i=1}^n a_{i,\kappa(i)}\right) \underbrace{\det(B_\kappa)}_{\substack{=0 \\ \text{(by Lemma 6.17 (b))}}} \\
&= \sum_{\kappa\in S_n} \left(\prod_{i=1}^n a_{i,\kappa(i)}\right) (-1)^\kappa \cdot \det B + \underbrace{\sum_{\substack{\kappa:[n]\to[n];\\ \kappa\notin S_n}} \left(\prod_{i=1}^n a_{i,\kappa(i)}\right) 0}_{=0} \\
&= \sum_{\kappa\in S_n} \left(\prod_{i=1}^n a_{i,\kappa(i)}\right) (-1)^\kappa \cdot \det B = \sum_{\sigma\in S_n} \left(\prod_{i=1}^n a_{i,\sigma(i)}\right) (-1)^\sigma \cdot \det B \\
&\qquad \text{(here, we renamed the summation index } \kappa \text{ as } \sigma\text{)} \\
&= \underbrace{\left(\sum_{\sigma\in S_n} (-1)^\sigma \prod_{i=1}^n a_{i,\sigma(i)}\right)}_{\substack{=\det A \\ \text{(by (342))}}} \cdot \det B = \det A \cdot \det B.
\end{aligned}$$

This proves Theorem 6.23. □

> **Remark 6.24.** The analogue of Theorem 6.23 with addition instead of multiplication does not hold. If A and B are two $n \times n$-matrices for some $n \in \mathbb{N}$, then $\det(A+B)$ does usually **not** equal $\det A + \det B$. (However, there is a formula for $\det(A+B)$ in terms of determinants of *submatrices* of A and B; see Theorem 6.160 below for this.)

We shall now show several applications of Theorem 6.23. First, a simple corollary:

Corollary 6.25. Let $n \in \mathbb{N}$.

(a) If $B_1, B_2, \ldots, B_k$ are finitely many $n \times n$-matrices, then $\det(B_1 B_2 \cdots B_k) = \prod_{i=1}^{k} \det(B_i)$.

(b) If B is any $n \times n$-matrix, and $k \in \mathbb{N}$, then $\det(B^k) = (\det B)^k$.

Proof of Corollary 6.25. Corollary 6.25 easily follows from Theorem 6.23 by induction over k. (The induction base, $k = 0$, relies on the fact that the product of 0 matrices is I_n and has determinant $\det(I_n) = 1$.) We leave the details to the reader. □

Example 6.26. Recall that the Fibonacci sequence is the sequence $(f_0, f_1, f_2, \ldots)$ of integers which is defined recursively by $f_0 = 0$, $f_1 = 1$, and $f_n = f_{n-1} + f_{n-2}$ for all $n \geq 2$. We shall prove that

$$f_{n+1} f_{n-1} - f_n^2 = (-1)^n \qquad \text{for every positive integer } n. \tag{370}$$

(This is a classical fact known as the *Cassini identity* and easy to prove by induction, but we shall prove it differently to illustrate the use of determinants.)

Let B be the 2×2-matrix $\begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}$ (over the ring $\mathbb{Z}$). It is easy to see that $\det B = -1$. However, for every positive integer n, we have

$$B^n = \begin{pmatrix} f_{n+1} & f_n \\ f_n & f_{n-1} \end{pmatrix}. \tag{371}$$

Indeed, (371) can be easily proven by induction over n: For $n = 1$ it is clear by inspection; if it holds for $n = N$, then for $n = N + 1$ it follows from

$$\begin{aligned} B^{N+1} &= \underbrace{B^N}_{\substack{= \begin{pmatrix} f_{N+1} & f_N \\ f_N & f_{N-1} \end{pmatrix} \\ \text{(by the induction hypothesis)}}} \underbrace{B}_{= \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}} = \begin{pmatrix} f_{N+1} & f_N \\ f_N & f_{N-1} \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \\ &= \begin{pmatrix} f_{N+1} \cdot 1 + f_N \cdot 1 & f_{N+1} \cdot 1 + f_N \cdot 0 \\ f_N \cdot 1 + f_{N-1} \cdot 1 & f_N \cdot 1 + f_{N-1} \cdot 0 \end{pmatrix} \\ &\qquad \text{(by the definition of a product of two matrices)} \\ &= \begin{pmatrix} f_{N+1} + f_N & f_{N+1} \\ f_N + f_{N-1} & f_N \end{pmatrix} = \begin{pmatrix} f_{N+2} & f_{N+1} \\ f_{N+1} & f_N \end{pmatrix} \end{aligned}$$

(since $f_{N+1} + f_N = f_{N+2}$ and $f_N + f_{N-1} = f_{N+1}$).

Now, let n be a positive integer. Then, (371) yields

$$\det(B^n) = \det \begin{pmatrix} f_{n+1} & f_n \\ f_n & f_{n-1} \end{pmatrix} = f_{n+1} f_{n-1} - f_n^2.$$

On the other hand, Corollary 6.25 **(b)** (applied to $k = n$) yields $\det(B^n) = \left(\underbrace{\det B}_{=-1}\right)^n = (-1)^n$. Hence, $f_{n+1}f_{n-1} - f_n^2 = \det(B^n) = (-1)^n$. This proves (370).

We can generalize (370) as follows:

Exercise 6.10. Let a and b be two complex numbers. Let $(x_0, x_1, x_2, \ldots)$ be a sequence of complex numbers such that every $n \geq 2$ satisfies

$$x_n = ax_{n-1} + bx_{n-2}. \tag{372}$$

(We called such sequences "(a,b)-recurrent" in Definition 4.2.) Let $k \in \mathbb{N}$. Prove that

$$x_{n+1}x_{n-k-1} - x_n x_{n-k} = (-b)^{n-k-1}(x_{k+2}x_0 - x_{k+1}x_1). \tag{373}$$

for every integer $n > k$.

We notice that (370) can be obtained by applying (373) to $a = 1$, $b = 1$, $x_i = f_i$ and $k = 0$. Thus, (373) is a generalization of (370). Notice that you could have easily come up with the identity (373) by trying to generalize the proof of (370) we gave; in contrast, it is not that straightforward to guess the general formula (373) from the classical proof of (370) by induction. So the proof of (370) using determinants has at least the advantage of pointing to a generalization.

Example 6.27. Let $n \in \mathbb{N}$. Let $x_1, x_2, \ldots, x_n$ be n elements of $\mathbb{K}$, and let $y_1, y_2, \ldots, y_n$ be n further elements of $\mathbb{K}$. Let A be the $n \times n$-matrix $(x_i y_j)_{1 \leq i \leq n,\ 1 \leq j \leq n}$. In Example 6.6, we have shown that $\det A = 0$ if $n \geq 2$. We can now prove this in a simpler way.

Namely, let $n \geq 2$. Define an $n \times n$-matrix B by $B = \begin{pmatrix} x_1 & 0 & 0 & \cdots & 0 \\ x_2 & 0 & 0 & \cdots & 0 \\ x_3 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ x_n & 0 & 0 & \cdots & 0 \end{pmatrix}$. (Thus, the first column of B is $(x_1, x_2, \ldots, x_n)^T$, while all other columns are filled with zeroes.) Define an $n \times n$-matrix C by $C = \begin{pmatrix} y_1 & y_2 & y_3 & \cdots & y_n \\ 0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 0 \end{pmatrix}$. (Thus, the first row of C is $(y_1, y_2, \ldots, y_n)$, while all other rows are filled with zeroes.)

The second row of C consists of zeroes (and this second row indeed exists, because $n \geq 2$). Thus, Exercise 6.7 **(c)** (applied to C instead of A) yields $\det C = 0$.

Similarly, using Exercise 6.7 **(d)**, we can show that $\det B = 0$. Now, Theorem 6.23 (applied to B and C instead of A and B) yields $\det(BC) = \det B \cdot \underbrace{\det C}_{=0} = 0$. But what is BC ?

Write B in the form $B = (b_{i,j})_{1 \leq i \leq n,\ 1 \leq j \leq n}$, and write C in the form $C = (c_{i,j})_{1 \leq i \leq n,\ 1 \leq j \leq n}$. Then, the definition of BC yields

$$BC = \left(\sum_{k=1}^{n} b_{i,k} c_{k,j} \right)_{1 \leq i \leq n,\ 1 \leq j \leq n}.$$

Therefore, for every $(i,j) \in \{1,2,\ldots,n\}^2$, the (i,j)-th entry of the matrix BC is

$$\sum_{k=1}^{n} b_{i,k} c_{k,j} = \underbrace{b_{i,1}}_{=x_i} \underbrace{c_{1,j}}_{=y_j} + \sum_{k=2}^{n} \underbrace{b_{i,k}}_{=0} \underbrace{c_{k,j}}_{=0} = x_i y_j + \underbrace{\sum_{k=2}^{n} 0 \cdot 0}_{=0} = x_i y_j.$$

But this is the same as the (i,j)-th entry of the matrix A. Thus, every entry of BC equals the corresponding entry of A. Hence, $BC = A$, so that $\det(BC) = \det A$. Thus, $\det A = \det(BC) = 0$, just as we wanted to show.

Example 6.28. Let $n \in \mathbb{N}$. Let $x_1, x_2, \ldots, x_n$ be n elements of $\mathbb{K}$, and let $y_1, y_2, \ldots, y_n$ be n further elements of $\mathbb{K}$. Let A be the $n \times n$-matrix $(x_i + y_j)_{1 \leq i \leq n,\ 1 \leq j \leq n}$. In Example 6.7, we have shown that $\det A = 0$ if $n \geq 3$.

We can now prove this in a simpler way. The argument is similar to Example 6.27, and so I will be very brief:

Let $n \geq 3$. Define an $n \times n$-matrix B by $B = \begin{pmatrix} x_1 & 1 & 0 & \cdots & 0 \\ x_2 & 1 & 0 & \cdots & 0 \\ x_3 & 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ x_n & 1 & 0 & \cdots & 0 \end{pmatrix}$. (Thus, the first column of B is $(x_1, x_2, \ldots, x_n)^T$, the second column is $(1, 1, \ldots, 1)^T$, while all other columns are filled with zeroes.) Define an $n \times n$-matrix C by $C = \begin{pmatrix} 1 & 1 & 1 & \cdots & 1 \\ y_1 & y_2 & y_3 & \cdots & y_n \\ 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 0 \end{pmatrix}$. (Thus, the first row of C is $(1, 1, \ldots, 1)$, the second row is $(y_1, y_2, \ldots, y_n)$, while all other rows are filled with zeroes.) It is now easy to show that $BC = A$ (check this!), but both $\det B$ and $\det C$ are 0 (due to having a column or a row filled with zeroes). Thus, again, we obtain $\det A = 0$.

Exercise 6.11. Let $n \in \mathbb{N}$. Let A be the $n \times n$-matrix $\left(\dbinom{i+j-2}{i-1} \right)_{1 \leq i \leq n,\ 1 \leq j \leq n} = \begin{pmatrix} \dbinom{0}{0} & \dbinom{1}{0} & \cdots & \dbinom{n-1}{0} \\ \dbinom{1}{1} & \dbinom{2}{1} & \cdots & \dbinom{n}{1} \\ \vdots & \vdots & \ddots & \vdots \\ \dbinom{n-1}{n-1} & \dbinom{n}{n-1} & \cdots & \dbinom{2n-2}{n-1} \end{pmatrix}$. (This matrix is a piece of Pascal's triangle "rotated by 45°". For example, for $n = 4$, we have $A = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 4 \\ 1 & 3 & 6 & 10 \\ 1 & 4 & 10 & 20 \end{pmatrix}$.)

Show that $\det A = 1$.

The matrix A in Exercise 6.11 is one of the so-called *Pascal matrices*; see [EdeStr04] for an enlightening exposition of some of its properties (but beware of the fact that the very first page reveals a significant part of the solution of Exercise 6.11).

Remark 6.29. There exists a more general notion of a matrix, in which the rows and the columns are indexed not necessarily by integers from 1 to n (for some $n \in \mathbb{N}$), but rather by arbitrary objects. For instance, this more general notion allows us to speak of a matrix with two rows labelled "spam" and "eggs", and with three columns labelled 0, 3 and ∞. (It thus has 6 entries, such as the ("spam", 3)-th entry or the ("eggs", ∞)-th entry.) This notion of matrices is more general and more flexible than the one used above (e.g., it allows for infinite matrices), although it has some drawbacks (e.g., notions such as "lower-triangular" are not defined per se, because there might be no canonical way to order the rows and the columns; also, infinite matrices cannot always be multiplied). We might want to define the determinant of such a matrix. Of course, this only makes sense when the rows of the matrix are indexed by the same objects as its columns (this essentially says that the matrix is a "square matrix" in a reasonably general sense). So, let X be a set, and A be a "generalized matrix" whose rows and columns are both indexed by the elements of X. We want to define $\det A$. We assume that X is finite (indeed, while $\det A$ sometimes makes sense for infinite X, this only happens under some rather restrictive conditions). Then, we can define $\det A$ by

$$\det A = \sum_{\sigma \in S_X} (-1)^{\sigma} \prod_{i \in X} a_{i,\sigma(i)},$$

where S_X denotes the set of all permutations of X. This relies on a definition of $(-1)^{\sigma}$ for every $\sigma \in S_X$; fortunately, we have provided such a definition in Exercise 5.12.

We shall see more about determinants later. So far we have barely scratched the surface. Huge collections of problems and examples on the computation of determinants can be found in [Prasol94] and [Kratte99] (and, if you can be bothered with 100-years-old notation and level of rigor, in Muir's five-volume book series [Muir30] – one of the most comprehensive collections of "forgotten tales" in mathematics[225]).

Let us finish this section with a brief remark on the geometrical use of determinants.

Remark 6.30. Let us consider the Euclidean plane $\mathbb{R}^2$ with its Cartesian coordinate system and its origin 0. If $A = (x_A, y_A)$ and $B = (x_B, y_B)$ are two points on $\mathbb{R}^2$, then the area of the triangle $0AB$ is $\frac{1}{2}\left|\det\begin{pmatrix} x_A & x_B \\ y_A & y_B \end{pmatrix}\right|$. The absolute value here reflects the fact that determinants can be negative, while areas must always be ≥ 0 (although they can be 0 when 0, A and B are collinear); however, it makes working with areas somewhat awkward. This can be circumvented by the notion of a *signed area*. (The signed area of a triangle ABC is its regular area if the triangle is "directed clockwise", and otherwise it is the negative of its area.) The signed area of the triangle $0AB$ is $\frac{1}{2}\det\begin{pmatrix} x_A & x_B \\ y_A & y_B \end{pmatrix}$. (Here, 0 stands for the origin; i.e., "the triangle $0AB$" means the triangle with vertices at the origin, at A and at B.)

If $A = (x_A, y_A)$, $B = (x_B, y_B)$ and $C = (x_C, y_C)$ are three points in $\mathbb{R}^2$, then the signed area of triangle ABC is $\frac{1}{2}\det\begin{pmatrix} x_A & x_B & x_C \\ y_A & y_B & y_C \\ 1 & 1 & 1 \end{pmatrix}$.

Similar formulas hold for tetrahedra: If $A = (x_A, y_A, z_A)$, $B = (x_B, y_B, z_B)$ and $C = (x_C, y_C, z_C)$ are three points in $\mathbb{R}^3$, then the signed volume of the tetrahedron $0ABC$ is $\frac{1}{6}\det\begin{pmatrix} x_A & x_B & x_C \\ y_A & y_B & y_C \\ z_A & z_B & z_C \end{pmatrix}$. (Again, take the absolute value for the non-signed volume.) There is a 4 × 4 determinant formula for the signed volume of a general tetrahedron $ABCD$.

[225] In this series, Muir endeavors to summarize every paper that had been written about determinants until the year 1920. Several of these papers contain results that have fallen into oblivion, and not always justly so; Muir's summaries are thus a goldmine of interesting material. However, his notation is antiquated and his exposition is often extremely unintelligible (e.g., complicated identities are often presented by showing an example and hoping that the reader will correctly guess the pattern; others are stated in verbose sentences spanning multiple lines); very few proofs are given.

Three other classical British texts on determinants are Muir's and Metzler's [MuiMet60], Turnbull's [Turnbu29] and Aitken's [Aitken56]; these texts (particularly the first two) contain a wealth of remarkable results, many of which are barely remembered today. Unfortunately, their clarity and their level of rigor leave much to be desired by modern standards. The two-volume treatise [Vodick50] and [Vodick51] by Vodicka (in Czech) might be one of the most modern texts in this classical tradition.

More generally, formulas like this hold in the n-dimensional space $\mathbb{R}^n$. Indeed, the notion of a triangle in $\mathbb{R}^2$ and the notion of a tetrahedron in $\mathbb{R}^3$ can be generalized to a notion of a *simplex* in $\mathbb{R}^n$. The signed volume of such simplices can then be defined using determinants. Then, one can extend this definition to arbitrary polytopes (roughly speaking, convex bodies with flat bounding faces, as opposed to "round" ones like spheres). While the general definition of the volume of a convex body uses analysis (specifically, integrals and Lebesgue measure), such a purely algebraic definition has its own advantages.

6.5. The Cauchy-Binet formula

This section is devoted to the Cauchy-Binet formula: a generalization of Theorem 6.23 which is less well-known than the latter, but still comes useful. This formula appears in the literature in various forms; we follow the one on PlanetMath (although we use different notations).

First, we introduce a notation for "picking out some rows of a matrix and throwing away the rest" (and also the analogous thing for columns):

Definition 6.31. Let $n \in \mathbb{N}$ and $m \in \mathbb{N}$. Let $A = \left(a_{i,j}\right)_{1\leq i\leq n,\ 1\leq j\leq m}$ be an $n \times m$-matrix.

(a) If $i_1, i_2, \ldots, i_u$ are some elements of $\{1, 2, \ldots, n\}$, then we let $\operatorname{rows}_{i_1,i_2,\ldots,i_u} A$ denote the $u \times m$-matrix $\left(a_{i_x,j}\right)_{1\leq x\leq u,\ 1\leq j\leq m}$. For instance, if $A = \begin{pmatrix} a & a' & a'' \\ b & b' & b'' \\ c & c' & c'' \\ d & d' & d'' \end{pmatrix}$, then $\operatorname{rows}_{3,1,4} A = \begin{pmatrix} c & c' & c'' \\ a & a' & a'' \\ d & d' & d'' \end{pmatrix}$. For every $p \in \{1, 2, \ldots, u\}$, we have

$$\begin{aligned} & \left(\text{the } p\text{-th row of } \operatorname{rows}_{i_1,i_2,\ldots,i_u} A\right) \\ &= \left(a_{i_p,1}, a_{i_p,2}, \ldots, a_{i_p,m}\right) \qquad \left(\text{since } \operatorname{rows}_{i_1,i_2,\ldots,i_u} A = \left(a_{i_x,j}\right)_{1\leq x\leq u,\ 1\leq j\leq m}\right) \\ &= \left(\text{the } i_p\text{-th row of } A\right) \qquad \left(\text{since } A = \left(a_{i,j}\right)_{1\leq i\leq n,\ 1\leq j\leq m}\right). \end{aligned} \tag{374}$$

Thus, $\operatorname{rows}_{i_1,i_2,\ldots,i_u} A$ is the $u \times m$-matrix whose rows (from top to bottom) are the rows labelled $i_1, i_2, \ldots, i_u$ of the matrix A.

(b) If $j_1, j_2, \ldots, j_v$ are some elements of $\{1, 2, \ldots, m\}$, then we let $\operatorname{cols}_{j_1,j_2,\ldots,j_v} A$ denote the $n \times v$-matrix $\left(a_{i,j_y}\right)_{1\leq i\leq n,\ 1\leq y\leq v}$. For instance, if $A = \begin{pmatrix} a & a' & a'' \\ b & b' & b'' \\ c & c' & c'' \end{pmatrix}$, then $\operatorname{cols}_{3,2} A = \begin{pmatrix} a'' & a' \\ b'' & b' \\ c'' & c' \end{pmatrix}$. For every $q \in \{1, 2, \ldots, v\}$, we have

$$\begin{aligned} & \left(\text{the } q\text{-th column of } \operatorname{cols}_{j_1, j_2, \ldots, j_v} A\right) \\ & = \begin{pmatrix} a_{1,j_q} \\ a_{2,j_q} \\ \vdots \\ a_{n,j_q} \end{pmatrix} \qquad \left(\text{since } \operatorname{cols}_{j_1, j_2, \ldots, j_v} A = \left(a_{i,j_y}\right)_{1 \leq i \leq n,\ 1 \leq y \leq v}\right) \\ & = \left(\text{the } j_q\text{-th column of } A\right) \qquad \left(\text{since } A = \left(a_{i,j}\right)_{1 \leq i \leq n,\ 1 \leq j \leq m}\right). \end{aligned} \tag{375}$$

Thus, $\operatorname{cols}_{j_1, j_2, \ldots, j_v} A$ is the $n \times v$-matrix whose columns (from left to right) are the columns labelled $j_1, j_2, \ldots, j_v$ of the matrix A.

Now we can state the *Cauchy-Binet formula*:

Theorem 6.32. Let $n \in \mathbb{N}$ and $m \in \mathbb{N}$. Let A be an $n \times m$-matrix, and let B be an $m \times n$-matrix. Then,

$$\det(AB) = \sum_{1 \leq g_1 < g_2 < \cdots < g_n \leq m} \det\left(\operatorname{cols}_{g_1, g_2, \ldots, g_n} A\right) \cdot \det\left(\operatorname{rows}_{g_1, g_2, \ldots, g_n} B\right). \tag{376}$$

Remark 6.33. The summation sign $\sum\limits_{1 \leq g_1 < g_2 < \cdots < g_n \leq m}$ in (376) is an abbreviation for

$$\sum_{\substack{(g_1, g_2, \ldots, g_n) \in \{1, 2, \ldots, m\}^n; \\ g_1 < g_2 < \cdots < g_n}}. \tag{377}$$

In particular, if $n = 0$, then it signifies a summation over all 0-tuples of elements of $\{1, 2, \ldots, m\}$ (because in this case, the chain of inequalities $g_1 < g_2 < \cdots < g_n$ is a tautology); such a sum always has exactly one addend (because there is exactly one 0-tuple).

When both n and m equal 0, then the notation $\sum\limits_{1 \leq g_1 < g_2 < \cdots < g_n \leq m}$ is slightly confusing: It appears to mean an empty summation (because $1 \leq m$ does not hold). But as we said, we mean this notation to be an abbreviation for (377), which signifies a sum with exactly one addend. But this is enough pedantry for now; for $n > 0$, the notation $\sum\limits_{1 \leq g_1 < g_2 < \cdots < g_n \leq m}$ fortunately means exactly what it seems to mean.

We shall soon give a detailed proof of Theorem 6.32; see [AigZie14, Chapter 32, Theorem] for a different proof[226].

[226] Note that the formulation of Theorem 6.32 in [AigZie14, Chapter 32, Theorem] is slightly different:

Before we prove Theorem 6.32, let us give some examples for its use. First, here is a simple fact:

Lemma 6.34. Let $n \in \mathbb{N}$.

(a) There exists exactly one n-tuple $(g_1, g_2, \ldots, g_n) \in \{1, 2, \ldots, n\}^n$ satisfying $g_1 < g_2 < \cdots < g_n$, namely the n-tuple $(1, 2, \ldots, n)$.

(b) Let $m \in \mathbb{N}$ be such that $m < n$. Then, there exists no n-tuple $(g_1, g_2, \ldots, g_n) \in \{1, 2, \ldots, m\}^n$ satisfying $g_1 < g_2 < \cdots < g_n$.

As for its intuitive meaning, Lemma 6.34 can be viewed as a "pigeonhole principle" for strictly increasing sequences: Part **(b)** says (roughly speaking) that there is no way to squeeze a strictly increasing sequence $(g_1, g_2, \ldots, g_n)$ of n numbers into the set $\{1, 2, \ldots, m\}$ when $m < n$; part **(a)** says (again, informally) that the only such sequence for $m = n$ is $(1, 2, \ldots, n)$.

Exercise 6.12. Give a formal proof of Lemma 6.34. (Do not bother doing this if you do not particularly care about formal proofs and find Lemma 6.34 obvious enough.)

Example 6.35. Let $n \in \mathbb{N}$. Let A and B be two $n \times n$-matrices. It is easy to check that $\operatorname{cols}_{1,2,\ldots,n} A = A$ and $\operatorname{rows}_{1,2,\ldots,n} B = B$. Now, Theorem 6.32 (applied to $m = n$) yields

$$\det(AB) = \sum_{1 \leq g_1 < g_2 < \cdots < g_n \leq n} \det\left(\operatorname{cols}_{g_1, g_2, \ldots, g_n} A\right) \cdot \det\left(\operatorname{rows}_{g_1, g_2, \ldots, g_n} B\right). \qquad (379)$$

But Lemma 6.34 **(a)** yields that there exists exactly one n-tuple $(g_1, g_2, \ldots, g_n) \in \{1, 2, \ldots, n\}^n$ satisfying $g_1 < g_2 < \cdots < g_n$, namely the n-tuple $(1, 2, \ldots, n)$. Hence, the sum on the right hand side of (379) has exactly one addend: namely,

In our notations, it says that if A is an $n \times m$-matrix and if B is an $m \times n$-matrix, then

$$\det(AB) = \sum_{\substack{\mathcal{Z} \subseteq \{1,2,\ldots,m\};\\ |\mathcal{Z}| = n}} \det(\operatorname{cols}_{\mathcal{Z}} A) \cdot \det(\operatorname{rows}_{\mathcal{Z}} B), \qquad (378)$$

where the matrices $\operatorname{cols}_{\mathcal{Z}} A$ and $\operatorname{rows}_{\mathcal{Z}} B$ (for $\mathcal{Z}$ being a subset of $\{1, 2, \ldots, m\}$) are defined as follows: Write the subset $\mathcal{Z}$ in the form $\{z_1, z_2, \ldots, z_k\}$ with $z_1 < z_2 < \cdots < z_k$, and set $\operatorname{cols}_{\mathcal{Z}} A = \operatorname{cols}_{z_1, z_2, \ldots, z_k} A$ and $\operatorname{rows}_{\mathcal{Z}} B = \operatorname{rows}_{z_1, z_2, \ldots, z_k} B$. (Apart from this, [AigZie14, Chapter 32, Theorem] also requires $n \leq m$; but this requirement is useless.)

The equalities (376) and (378) are equivalent, because the n-tuples $(g_1, g_2, \ldots, g_n) \in \{1, 2, \ldots, m\}^n$ satisfying $g_1 < g_2 < \cdots < g_n$ are in a bijection with the subsets $\mathcal{Z}$ of $\{1, 2, \ldots, m\}$ satisfying $|\mathcal{Z}| = n$. (This bijection sends an n-tuple $(g_1, g_2, \ldots, g_n)$ to the subset $\{g_1, g_2, \ldots, g_n\}$.)

The proof of (378) given in [AigZie14, Chapter 32] uses the *Lindström-Gessel-Viennot lemma* (which it calls the "lemma of Gessel-Viennot") and is highly worth reading.

A closely related combinatorial proof of Theorem 6.32 appears in [Zeng93, §2]. Two other proofs (one of which is more or less the one we shall give below) are sketched in [Prasol94, Theorem 2.3 and Problem 28.7]. Yet another proof (using characteristic polynomials and block matrices) can be found in [LLPT95, (APP.1.2)].

the addend for $(g_1, g_2, \ldots, g_n) = (1, 2, \ldots, n)$. Therefore, this sum simplifies as follows:

$$\begin{aligned}
&\sum_{1 \leq g_1 < g_2 < \cdots < g_n \leq n} \det\left(\operatorname{cols}_{g_1, g_2, \ldots, g_n} A\right) \cdot \det\left(\operatorname{rows}_{g_1, g_2, \ldots, g_n} B\right) \\
&= \det\left(\underbrace{\operatorname{cols}_{1,2,\ldots,n} A}_{=A}\right) \cdot \det\left(\underbrace{\operatorname{rows}_{1,2,\ldots,n} B}_{=B}\right) = \det A \cdot \det B.
\end{aligned}$$

Hence, (379) becomes

$$\begin{aligned}
\det(AB) &= \sum_{1 \leq g_1 < g_2 < \cdots < g_n \leq n} \det\left(\operatorname{cols}_{g_1, g_2, \ldots, g_n} A\right) \cdot \det\left(\operatorname{rows}_{g_1, g_2, \ldots, g_n} B\right) \\
&= \det A \cdot \det B.
\end{aligned}$$

This, of course, is the statement of Theorem 6.23. Hence, Theorem 6.23 is a particular case of Theorem 6.32.

Example 6.36. Let $n \in \mathbb{N}$ and $m \in \mathbb{N}$ be such that $m < n$. Thus, Lemma 6.34 **(b)** shows that there exists no n-tuple $(g_1, g_2, \ldots, g_n) \in \{1, 2, \ldots, m\}^n$ satisfying $g_1 < g_2 < \cdots < g_n$.

Now, let A be an $n \times m$-matrix, and let B be an $m \times n$-matrix. Then, Theorem 6.32 yields

$$\begin{aligned}
\det(AB) &= \sum_{1 \leq g_1 < g_2 < \cdots < g_n \leq m} \det\left(\operatorname{cols}_{g_1, g_2, \ldots, g_n} A\right) \cdot \det\left(\operatorname{rows}_{g_1, g_2, \ldots, g_n} B\right) \\
&= (\text{empty sum}) \\
&\qquad \left(\begin{array}{c}\text{since there exists no } n\text{-tuple } (g_1, g_2, \ldots, g_n) \in \{1, 2, \ldots, m\}^n \\ \text{satisfying } g_1 < g_2 < \cdots < g_n\end{array}\right) \\
&= 0. \qquad (380)
\end{aligned}$$

This identity allows us to compute $\det A$ in Example 6.27 in a simpler way: Instead of defining two $n \times n$-matrices B and C by $B = \begin{pmatrix} x_1 & 0 & 0 & \cdots & 0 \\ x_2 & 0 & 0 & \cdots & 0 \\ x_3 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ x_n & 0 & 0 & \cdots & 0 \end{pmatrix}$ and $C = \begin{pmatrix} y_1 & y_2 & y_3 & \cdots & y_n \\ 0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 0 \end{pmatrix}$, it suffices to define an $n \times 1$-matrix B' by $B' = (x_1, x_2, \ldots, x_n)^T$ and a $1 \times n$-matrix C' by $C' = (y_1, y_2, \ldots, y_n)$, and argue that $A = B'C'$. (We leave the details to the reader.) Similarly, Example 6.28 could be dealt with.

Remark 6.37. The equality (380) can also be derived from Theorem 6.23. Indeed, let $n \in \mathbb{N}$ and $m \in \mathbb{N}$ be such that $m < n$. Let A be an $n \times m$-matrix, and let B be an $m \times n$-matrix. Notice that $n - m > 0$ (since $m < n$). Let A' be the $n \times n$-matrix obtained from A by appending $n - m$ new columns to the right of A and filling these columns with zeroes. (For example, if $n = 4$ and $m = 2$ and $A = \begin{pmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \\ a_{3,1} & a_{3,2} \\ a_{4,1} & a_{4,2} \end{pmatrix}$, then $A' = \begin{pmatrix} a_{1,1} & a_{1,2} & 0 & 0 \\ a_{2,1} & a_{2,2} & 0 & 0 \\ a_{3,1} & a_{3,2} & 0 & 0 \\ a_{4,1} & a_{4,2} & 0 & 0 \end{pmatrix}$.) Also, let B' be the $n \times n$-matrix obtained from B by appending $n - m$ new rows to the bottom of B and filling these rows with zeroes. (For example, if $n = 4$ and $m = 2$ and $B = \begin{pmatrix} b_{1,1} & b_{1,2} & b_{1,3} & b_{1,4} \\ b_{2,1} & b_{2,2} & b_{2,3} & b_{2,4} \end{pmatrix}$, then $B' = \begin{pmatrix} b_{1,1} & b_{1,2} & b_{1,3} & b_{1,4} \\ b_{2,1} & b_{2,2} & b_{2,3} & b_{2,4} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$.) Then, it is easy to check that $AB = A'B'$ (in fact, just compare corresponding entries of AB and $A'B'$). But recall that $n - m > 0$. Hence, the matrix A' has a column consisting of zeroes (namely, its last column). Thus, Exercise 6.7 **(d)** (applied to A' instead of A) shows that $\det(A') = 0$. Now,

$$\begin{aligned} \det\left(\underbrace{AB}_{=A'B'}\right) &= \det(A'B') = \underbrace{\det(A')}_{=0} \cdot \det(B') \\ &\qquad \left(\text{by Theorem 6.23, applied to } A' \text{ and } B' \text{ instead of } A \text{ and } B\right) \\ &= 0. \end{aligned}$$

Thus, (380) is proven again.

Example 6.38. Let us see what Theorem 6.32 says for $n = 1$. Indeed, let $m \in \mathbb{N}$; let $A = (a_1, a_2, \ldots, a_m)$ be a $1 \times m$-matrix (i.e., a row vector with m entries), and let $B = (b_1, b_2, \ldots, b_m)^T$ be an $m \times 1$-matrix (i.e., a column vector with m entries). Then, AB is the 1×1-matrix $\left(\sum_{k=1}^{m} a_k b_k\right)$. Thus,

$$\det(AB) = \det\left(\sum_{k=1}^{m} a_k b_k\right) = \sum_{k=1}^{m} a_k b_k \qquad (\text{by (343)}). \tag{381}$$

What would we obtain if we tried to compute $\det(AB)$ using Theorem 6.32?

Theorem 6.32 (applied to $n = 1$) yields

$$\begin{aligned}
\det(AB) &= \underbrace{\sum_{1 \leq g_1 \leq m}}_{=\sum_{g_1=1}^{m}} \det\left(\underbrace{\operatorname{cols}_{g_1} A}_{=\left(\begin{array}{c} a_{g_1} \end{array}\right)}\right) \cdot \det\left(\underbrace{\operatorname{rows}_{g_1} B}_{=\left(\begin{array}{c} b_{g_1} \end{array}\right)}\right) \\
&= \sum_{g_1=1}^{m} \underbrace{\det\left(\begin{array}{c} a_{g_1} \end{array}\right)}_{\substack{=a_{g_1} \\ \text{(by (343))}}} \cdot \underbrace{\det\left(\begin{array}{c} b_{g_1} \end{array}\right)}_{\substack{=b_{g_1} \\ \text{(by (343))}}} = \sum_{g_1=1}^{m} a_{g_1} \cdot b_{g_1}.
\end{aligned}$$

This is, of course, the same result as that of (381) (with the summation index k renamed as g_1). So we did not gain any interesting insight from applying Theorem 6.32 to $n = 1$.

Example 6.39. Let us try a slightly less trivial case. Indeed, let $m \in \mathbb{N}$; let $A = \left(\begin{array}{cccc} a_1 & a_2 & \cdots & a_m \\ a_1' & a_2' & \cdots & a_m' \end{array}\right)$ be a $2 \times m$-matrix, and let $B = \left(\begin{array}{cc} b_1 & b_1' \\ b_2 & b_2' \\ \vdots & \vdots \\ b_m & b_m' \end{array}\right)$ be an $m \times 2$-matrix. Then, AB is the 2×2-matrix $\left(\begin{array}{cc} \sum\limits_{k=1}^{m} a_k b_k & \sum\limits_{k=1}^{m} a_k b_k' \\ \sum\limits_{k=1}^{m} a_k' b_k & \sum\limits_{k=1}^{m} a_k' b_k' \end{array}\right)$. Hence,

$$\begin{aligned}
\det(AB) &= \det\left(\begin{array}{cc} \sum\limits_{k=1}^{m} a_k b_k & \sum\limits_{k=1}^{m} a_k b_k' \\ \sum\limits_{k=1}^{m} a_k' b_k & \sum\limits_{k=1}^{m} a_k' b_k' \end{array}\right) \\
&= \left(\sum_{k=1}^{m} a_k b_k\right)\left(\sum_{k=1}^{m} a_k' b_k'\right) - \left(\sum_{k=1}^{m} a_k' b_k\right)\left(\sum_{k=1}^{m} a_k b_k'\right). \qquad (382)
\end{aligned}$$

On the other hand, Theorem 6.32 (now applied to $n = 2$) yields

$$\begin{aligned}
\det(AB) &= \sum_{1\leq g_1<g_2\leq m} \det\left(\operatorname{cols}_{g_1,g_2} A\right)\cdot\det\left(\operatorname{rows}_{g_1,g_2} B\right) \\
&= \sum_{1\leq i<j\leq m} \det\left(\underbrace{\operatorname{cols}_{i,j} A}_{=\begin{pmatrix} a_i & a_j \\ a'_i & a'_j \end{pmatrix}}\right)\cdot\det\left(\underbrace{\operatorname{rows}_{i,j} B}_{=\begin{pmatrix} b_i & b'_i \\ b_j & b'_j \end{pmatrix}}\right) \\
&\qquad \left(\begin{array}{c}\text{here, we renamed the summation indices } g_1 \text{ and } g_2 \\ \text{as } i \text{ and } j\text{, since double subscripts are annoying}\end{array}\right) \\
&= \sum_{1\leq i<j\leq m} \underbrace{\det\begin{pmatrix} a_i & a_j \\ a'_i & a'_j \end{pmatrix}}_{=a_ia'_j-a_ja'_i}\cdot\underbrace{\det\begin{pmatrix} b_i & b'_i \\ b_j & b'_j \end{pmatrix}}_{=b_ib'_j-b_jb'_i} \\
&= \sum_{1\leq i<j\leq m} \left(a_ia'_j-a_ja'_i\right)\cdot\left(b_ib'_j-b_jb'_i\right).
\end{aligned}$$

Compared with (382), this yields

$$\begin{aligned}
&\left(\sum_{k=1}^m a_kb_k\right)\left(\sum_{k=1}^m a'_kb'_k\right)-\left(\sum_{k=1}^m a'_kb_k\right)\left(\sum_{k=1}^m a_kb'_k\right) \\
&= \sum_{1\leq i<j\leq m} \left(a_ia'_j-a_ja'_i\right)\cdot\left(b_ib'_j-b_jb'_i\right). \qquad (383)
\end{aligned}$$

This identity is called the *Binet-Cauchy identity* (I am not kidding – look it up on the Wikipedia). It is fairly easy to prove by direct computation; thus, using Theorem 6.32 to prove it was quite an overkill. However, (383) might not be very easy to come up with, whereas deriving it from Theorem 6.32 is straightforward. (And Theorem 6.32 is easier to memorize than (383).)

Here is a neat application of (383): If $a_1, a_2, \ldots, a_m$ and $a'_1, a'_2, \ldots, a'_m$ are real numbers, then (383) (applied to $b_k = a_k$ and $b'_k = a'_k$) yields

$$\begin{aligned}
&\left(\sum_{k=1}^m a_ka_k\right)\left(\sum_{k=1}^m a'_ka'_k\right)-\left(\sum_{k=1}^m a'_ka_k\right)\left(\sum_{k=1}^m a_ka'_k\right) \\
&= \sum_{1\leq i<j\leq m} \underbrace{\left(a_ia'_j-a_ja'_i\right)\cdot\left(a_ia'_j-a_ja'_i\right)}_{=\left(a_ia'_j-a_ja'_i\right)^2\geq 0} \geq \sum_{1\leq i<j\leq m} 0 = 0,
\end{aligned}$$

so that

$$\left(\sum_{k=1}^m a_ka_k\right)\left(\sum_{k=1}^m a'_ka'_k\right) \geq \left(\sum_{k=1}^m a'_ka_k\right)\left(\sum_{k=1}^m a_ka'_k\right).$$

In other words,

$$\left(\sum_{k=1}^{m} a_k^2\right) \left(\sum_{k=1}^{m} \left(a'_k\right)^2\right) \geq \left(\sum_{k=1}^{m} a_k a'_k\right)^2 .$$

This is the famous Cauchy-Schwarz inequality.

Let us now prepare for the proof of Theorem 6.32. First comes a fact which should be fairly clear:

Proposition 6.40. Let $n \in \mathbb{N}$. Let $a_1, a_2, \ldots, a_n$ be n integers.

(a) There exists a permutation $\sigma \in S_n$ such that $a_{\sigma(1)} \leq a_{\sigma(2)} \leq \cdots \leq a_{\sigma(n)}$.

(b) If $\sigma \in S_n$ is such that $a_{\sigma(1)} \leq a_{\sigma(2)} \leq \cdots \leq a_{\sigma(n)}$, then, for every $i \in \{1, 2, \ldots, n\}$, the value $a_{\sigma(i)}$ depends only on $a_1, a_2, \ldots, a_n$ and i (but not on σ).

(c) Assume that the integers $a_1, a_2, \ldots, a_n$ are distinct. Then, there is a **unique** permutation $\sigma \in S_n$ such that $a_{\sigma(1)} < a_{\sigma(2)} < \cdots < a_{\sigma(n)}$.

Let me explain why this proposition should be intuitively obvious.[227] Proposition 6.40 **(a)** says that any list $(a_1, a_2, \ldots, a_n)$ of n integers can be sorted in weakly increasing order by means of a permutation $\sigma \in S_n$. Proposition 6.40 **(b)** says that the result of this sorting process is independent of how the sorting happened (although the permutation σ will sometimes be non-unique). Proposition 6.40 **(c)** says that if the integers $a_1, a_2, \ldots, a_n$ are distinct, then the permutation $\sigma \in S_n$ which sorts the list $(a_1, a_2, \ldots, a_n)$ in increasing order is uniquely determined as well. We required $a_1, a_2, \ldots, a_n$ to be n integers for the sake of simplicity, but we could just as well have required them to be elements of any *totally ordered set* (i.e., any set with a less-than relation satisfying some standard axioms).

The next fact looks slightly scary, but is still rather simple:

Lemma 6.41. For every $n \in \mathbb{N}$, let $[n]$ denote the set $\{1, 2, \ldots, n\}$.

Let $n \in \mathbb{N}$ and $m \in \mathbb{N}$. We let $\mathbf{E}$ be the subset

$$\left\{(k_1, k_2, \ldots, k_n) \in [m]^n \mid \text{the integers } k_1, k_2, \ldots, k_n \text{ are distinct}\right\}$$

of $[m]^n$. We let $\mathbf{I}$ be the subset

$$\left\{(k_1, k_2, \ldots, k_n) \in [m]^n \mid k_1 < k_2 < \cdots < k_n\right\}$$

of $[m]^n$. Then, the map

$$\begin{aligned} \mathbf{I} \times S_n &\to \mathbf{E}, \\ ((g_1, g_2, \ldots, g_n), \sigma) &\mapsto \left(g_{\sigma(1)}, g_{\sigma(2)}, \ldots, g_{\sigma(n)}\right) \end{aligned}$$

is well-defined and is a bijection.

[227] See the solution of Exercise 6.13 further below for a formal proof of this proposition.

The intuition for Lemma 6.41 is that every n-tuple of distinct elements of $\{1,2,\ldots,m\}$ can be represented uniquely as a permuted version of a strictly increasing[228] n-tuple of elements of $\{1,2,\ldots,m\}$, and therefore, specifying an n-tuple of distinct elements of $\{1,2,\ldots,m\}$ is tantamount to specifying a strictly increasing n-tuple of elements of $\{1,2,\ldots,m\}$ and a permutation $\sigma \in S_n$ which says how this n-tuple is to be permuted.[229] This is not a formal proof, but this should explain why Lemma 6.41 is usually applied throughout mathematics without even mentioning it as a statement. If desired, a formal proof of Lemma 6.41 can be obtained using Proposition 6.40.[230]

Exercise 6.13. Prove Proposition 6.40 and Lemma 6.41. (Ignore this exercise if you find these two facts sufficiently obvious and are uninterested in the details of their proofs.)

Before we return to Theorem 6.32, let me make a digression about sorting:

Exercise 6.14. Let $n \in \mathbb{N}$ and $m \in \mathbb{N}$ be such that $n \geq m$. Let $a_1, a_2, \ldots, a_n$ be n integers. Let $b_1, b_2, \ldots, b_m$ be m integers. Assume that

$$a_i \leq b_i \qquad \text{for every } i \in \{1,2,\ldots,m\}. \tag{384}$$

Let $\sigma \in S_n$ be such that $a_{\sigma(1)} \leq a_{\sigma(2)} \leq \cdots \leq a_{\sigma(n)}$. Let $\tau \in S_m$ be such that $b_{\tau(1)} \leq b_{\tau(2)} \leq \cdots \leq b_{\tau(m)}$. Then,

$$a_{\sigma(i)} \leq b_{\tau(i)} \qquad \text{for every } i \in \{1,2,\ldots,m\}.$$

Remark 6.42. Loosely speaking, Exercise 6.14 says the following: If two lists $(a_1, a_2, \ldots, a_n)$ and $(b_1, b_2, \ldots, b_m)$ of integers have the property that each entry of the first list is $\leq$ to the corresponding entry of the second list (as long as the latter is well-defined), then this property still holds after both lists are sorted in increasing order, provided that we have $n \geq m$ (that is, the first list is at least as long as the second list).

A consequence of Exercise 6.14 is the following curious fact, known as the "non-messing-up phenomenon" ([Tenner04, Theorem 1] and [GalKar71, Example 1]): If we start with a matrix filled with integers, then sort the entries of each

[228] An n-tuple $(k_1, k_2, \ldots, k_n)$ is said to be *strictly increasing* if and only if $k_1 < k_2 < \cdots < k_n$.

[229] For instance, the 4-tuple $(4,1,6,2)$ of distinct elements of $\{1,2,\ldots,7\}$ can be specified by specifying the strictly increasing 4-tuple $(1,2,4,6)$ (which is its sorted version) and the permutation $\pi \in S_4$ which sends $1,2,3,4$ to $3,1,4,2$, respectively (that is, $\pi = (3,1,4,2)$ in one-line notation). In the terminology of Lemma 6.41, the map

$$\mathbf{I} \times S_n \to \mathbf{E},$$
$$((g_1, g_2, \ldots, g_n), \sigma) \mapsto \left(g_{\sigma(1)}, g_{\sigma(2)}, \ldots, g_{\sigma(n)}\right)$$

sends $((1,2,4,6), \pi)$ to $(4,1,6,2)$.

[230] Again, see the solution of Exercise 6.13 further below for such a proof.

row of the matrix in increasing order, and then sort the entries of each column of the resulting matrix in increasing order, then the final matrix still has sorted rows (i.e., the entries of each row are still sorted). That is, the sorting of the columns did not "mess up" the sortedness of the rows. For example, if we start with the matrix $\begin{pmatrix} 1 & 3 & 2 & 5 \\ 2 & 1 & 4 & 2 \\ 3 & 1 & 6 & 0 \end{pmatrix}$, then sorting the entries of each row gives us the matrix $\begin{pmatrix} 1 & 2 & 3 & 5 \\ 1 & 2 & 2 & 4 \\ 0 & 1 & 3 & 6 \end{pmatrix}$, and then sorting the entries of each column results in the matrix $\begin{pmatrix} 0 & 1 & 2 & 4 \\ 1 & 2 & 3 & 5 \\ 1 & 2 & 3 & 6 \end{pmatrix}$. The rows of this matrix are still sorted, as the "non-messing-up phenomenon" predicts. To prove this phenomenon in general, it suffices to show that any entry in the resulting matrix is $\leq$ to the entry directly below it (assuming that the latter exists); but this follows easily from Exercise 6.14.

We are now ready to prove Theorem 6.32.

Proof of Theorem 6.32. We shall use the notations of Lemma 6.41.

Write the $n \times m$-matrix A as $A = \left(a_{i,j}\right)_{1 \leq i \leq n,\ 1 \leq j \leq m}$. Write the $m \times n$-matrix B as $B = \left(b_{i,j}\right)_{1 \leq i \leq m,\ 1 \leq j \leq n}$. The definition of AB thus yields $AB = \left(\sum\limits_{k=1}^{m} a_{i,k} b_{k,j}\right)_{1 \leq i \leq n,\ 1 \leq j \leq n}$.

Therefore, (342) (applied to AB and $\sum\limits_{k=1}^{m} a_{i,k} b_{k,j}$ instead of A and $a_{i,j}$) yields

$$\det(AB) = \sum_{\sigma \in S_n} (-1)^{\sigma} \prod_{i=1}^{n} \left(\sum_{k=1}^{m} a_{i,k} b_{k,\sigma(i)} \right). \tag{385}$$

But for every $\sigma \in S_n$, we have

$$\begin{aligned} \prod_{i=1}^{n} \left(\sum_{k=1}^{m} a_{i,k} b_{k,\sigma(i)} \right) &= \sum_{(k_1,k_2,\ldots,k_n) \in [m]^n} \underbrace{\prod_{i=1}^{n} \left(a_{i,k_i} b_{k_i,\sigma(i)} \right)}_{= \left(\prod\limits_{i=1}^{n} a_{i,k_i} \right) \left(\prod\limits_{i=1}^{n} b_{k_i,\sigma(i)} \right)} \\ &\qquad \left(\text{by Lemma 6.20, applied to } m_i = n \text{ and } p_{i,k} = a_{i,k} b_{k,\sigma(i)} \right) \\ &= \sum_{(k_1,k_2,\ldots,k_n) \in [m]^n} \left(\prod_{i=1}^{n} a_{i,k_i} \right) \left(\prod_{i=1}^{n} b_{k_i,\sigma(i)} \right). \end{aligned}$$

Hence, (385) rewrites as

$$\begin{aligned}
\det(AB) &= \sum_{\sigma\in S_n}(-1)^{\sigma}\sum_{(k_1,k_2,\ldots,k_n)\in[m]^n}\left(\prod_{i=1}^n a_{i,k_i}\right)\left(\prod_{i=1}^n b_{k_i,\sigma(i)}\right)\\
&= \underbrace{\sum_{\sigma\in S_n}\sum_{(k_1,k_2,\ldots,k_n)\in[m]^n}}_{=\sum_{(k_1,k_2,\ldots,k_n)\in[m]^n}\sum_{\sigma\in S_n}}(-1)^{\sigma}\left(\prod_{i=1}^n a_{i,k_i}\right)\left(\prod_{i=1}^n b_{k_i,\sigma(i)}\right)\\
&= \sum_{(k_1,k_2,\ldots,k_n)\in[m]^n}\sum_{\sigma\in S_n}(-1)^{\sigma}\left(\prod_{i=1}^n a_{i,k_i}\right)\left(\prod_{i=1}^n b_{k_i,\sigma(i)}\right)\\
&= \sum_{(k_1,k_2,\ldots,k_n)\in[m]^n}\left(\prod_{i=1}^n a_{i,k_i}\right)\left(\sum_{\sigma\in S_n}(-1)^{\sigma}\prod_{i=1}^n b_{k_i,\sigma(i)}\right).
\end{aligned} \tag{386}$$

But every $(k_1,k_2,\ldots,k_n)\in[m]^n$ satisfies

$$\sum_{\sigma\in S_n}(-1)^{\sigma}\prod_{i=1}^n b_{k_i,\sigma(i)} = \det\left(\operatorname{rows}_{k_1,k_2,\ldots,k_n} B\right) \tag{387}$$

[231]. Hence, (386) becomes

$$\begin{aligned}
\det(AB) &= \sum_{(k_1,k_2,\ldots,k_n)\in[m]^n}\left(\prod_{i=1}^n a_{i,k_i}\right)\underbrace{\left(\sum_{\sigma\in S_n}(-1)^{\sigma}\prod_{i=1}^n b_{k_i,\sigma(i)}\right)}_{=\det\left(\operatorname{rows}_{k_1,k_2,\ldots,k_n} B\right)}\\
&= \sum_{(k_1,k_2,\ldots,k_n)\in[m]^n}\left(\prod_{i=1}^n a_{i,k_i}\right)\det\left(\operatorname{rows}_{k_1,k_2,\ldots,k_n} B\right).
\end{aligned} \tag{388}$$

But for every $(k_1,k_2,\ldots,k_n)\in[m]^n$ satisfying $(k_1,k_2,\ldots,k_n)\notin\mathbf{E}$, we have

$$\det\left(\operatorname{rows}_{k_1,k_2,\ldots,k_n} B\right) = 0 \tag{389}$$

[231] *Proof.* Let $(k_1,k_2,\ldots,k_n)\in[m]^n$. Recall that $B=\left(b_{i,j}\right)_{1\le i\le m,\ 1\le j\le n}$. Hence, the definition of $\operatorname{rows}_{k_1,k_2,\ldots,k_n} B$ gives us

$$\operatorname{rows}_{k_1,k_2,\ldots,k_n} B = \left(b_{k_x,j}\right)_{1\le x\le n,\ 1\le j\le n} = \left(b_{k_i,j}\right)_{1\le i\le n,\ 1\le j\le n}$$

(here, we renamed the index x as i). Hence, (342) (applied to $\operatorname{rows}_{k_1,k_2,\ldots,k_n} B$ and $b_{k_i,j}$ instead of A and $a_{i,j}$) yields

$$\det\left(\operatorname{rows}_{k_1,k_2,\ldots,k_n} B\right) = \sum_{\sigma\in S_n}(-1)^{\sigma}\prod_{i=1}^n b_{k_i,\sigma(i)},$$

qed.

[232]. Therefore, in the sum on the right hand side of (388), all the addends corresponding to $(k_1, k_2, \ldots, k_n) \in [m]^n$ satisfying $(k_1, k_2, \ldots, k_n) \notin \mathbf{E}$ evaluate to 0. We can therefore remove all these addends from the sum. The remaining addends are those corresponding to $(k_1, k_2, \ldots, k_n) \in \mathbf{E}$. Therefore, (388) becomes

$$\det(AB) = \sum_{(k_1,k_2,\ldots,k_n)\in\mathbf{E}} \left(\prod_{i=1}^n a_{i,k_i}\right) \det\left(\operatorname{rows}_{k_1,k_2,\ldots,k_n} B\right). \tag{390}$$

On the other hand, Lemma 6.41 yields that the map

$$\mathbf{I} \times S_n \to \mathbf{E},$$
$$((g_1, g_2, \ldots, g_n), \sigma) \mapsto \left(g_{\sigma(1)}, g_{\sigma(2)}, \ldots, g_{\sigma(n)}\right)$$

is well-defined and is a bijection. Hence, we can substitute $\left(g_{\sigma(1)}, g_{\sigma(2)}, \ldots, g_{\sigma(n)}\right)$ for $(k_1, k_2, \ldots, k_n)$ in the sum on the right hand side of (390). We thus obtain

$$\begin{aligned}&\sum_{(k_1,k_2,\ldots,k_n)\in\mathbf{E}} \left(\prod_{i=1}^n a_{i,k_i}\right) \det\left(\operatorname{rows}_{k_1,k_2,\ldots,k_n} B\right)\\ &= \sum_{((g_1,g_2,\ldots,g_n),\sigma)\in\mathbf{I}\times S_n} \left(\prod_{i=1}^n a_{i,g_{\sigma(i)}}\right) \det\left(\operatorname{rows}_{g_{\sigma(1)},g_{\sigma(2)},\ldots,g_{\sigma(n)}} B\right).\end{aligned}$$

Thus, (390) becomes

$$\begin{aligned}\det(AB) &= \sum_{(k_1,k_2,\ldots,k_n)\in\mathbf{E}} \left(\prod_{i=1}^n a_{i,k_i}\right) \det\left(\operatorname{rows}_{k_1,k_2,\ldots,k_n} B\right)\\ &= \sum_{((g_1,g_2,\ldots,g_n),\sigma)\in\mathbf{I}\times S_n} \left(\prod_{i=1}^n a_{i,g_{\sigma(i)}}\right) \det\left(\operatorname{rows}_{g_{\sigma(1)},g_{\sigma(2)},\ldots,g_{\sigma(n)}} B\right).\end{aligned} \tag{391}$$

But every $(k_1, k_2, \ldots, k_n) \in [m]^n$ and every $\sigma \in S_n$ satisfy

$$\det\left(\operatorname{rows}_{k_{\sigma(1)},k_{\sigma(2)},\ldots,k_{\sigma(n)}} B\right) = (-1)^\sigma \cdot \det\left(\operatorname{rows}_{k_1,k_2,\ldots,k_n} B\right) \tag{392}$$

[232] *Proof of (389):* Let $(k_1, k_2, \ldots, k_n) \in [m]^n$ be such that $(k_1, k_2, \ldots, k_n) \notin \mathbf{E}$. Then, the integers $k_1, k_2, \ldots, k_n$ are not distinct (because $\mathbf{E}$ is the set of all n-tuples in $[m]^n$ whose entries are distinct). Thus, there exist two distinct elements p and q of $[n]$ such that $k_p = k_q$. Consider these p and q. But $\operatorname{rows}_{k_1,k_2,\ldots,k_n} B$ is the $n \times n$-matrix whose rows (from top to bottom) are the rows labelled $k_1, k_2, \ldots, k_n$ of the matrix B. Since $k_p = k_q$, this shows that the p-th row and the q-th row of the matrix $\operatorname{rows}_{k_1,k_2,\ldots,k_n} B$ are equal. Hence, the matrix $\operatorname{rows}_{k_1,k_2,\ldots,k_n} B$ has two equal rows (since p and q are distinct). Therefore, Exercise 6.7 **(e)** (applied to $\operatorname{rows}_{k_1,k_2,\ldots,k_n} B$ instead of A) yields $\det\left(\operatorname{rows}_{k_1,k_2,\ldots,k_n} B\right) = 0$, qed.

[233]. Hence, (391) becomes

$$\begin{aligned}
&\det(AB) \\
&= \underbrace{\sum_{((g_1,g_2,\ldots,g_n),\sigma)\in\mathbf{I}\times S_n}}_{=\sum_{(g_1,g_2,\ldots,g_n)\in\mathbf{I}}\sum_{\sigma\in S_n}} \left(\prod_{i=1}^n a_{i,g_{\sigma(i)}}\right) \underbrace{\det\left(\operatorname{rows}_{g_{\sigma(1)},g_{\sigma(2)},\ldots,g_{\sigma(n)}} B\right)}_{\substack{=(-1)^\sigma\cdot\det\left(\operatorname{rows}_{g_1,g_2,\ldots,g_n} B\right) \\ \text{(by (392), applied to } k_i=g_i\text{)}}} \\
&= \sum_{(g_1,g_2,\ldots,g_n)\in\mathbf{I}} \sum_{\sigma\in S_n} \left(\prod_{i=1}^n a_{i,g_{\sigma(i)}}\right) (-1)^\sigma \cdot \det\left(\operatorname{rows}_{g_1,g_2,\ldots,g_n} B\right) \\
&= \sum_{(g_1,g_2,\ldots,g_n)\in\mathbf{I}} \left(\sum_{\sigma\in S_n} \left(\prod_{i=1}^n a_{i,g_{\sigma(i)}}\right) (-1)^\sigma\right) \cdot \det\left(\operatorname{rows}_{g_1,g_2,\ldots,g_n} B\right). \qquad (393)
\end{aligned}$$

But every $(g_1, g_2, \ldots, g_n) \in \mathbf{I}$ satisfies $\sum_{\sigma\in S_n} \left(\prod_{i=1}^n a_{i,g_{\sigma(i)}}\right) (-1)^\sigma = \det\left(\operatorname{cols}_{g_1,g_2,\ldots,g_n} A\right)$

[233] *Proof of (392):* Let $(k_1, k_2, \ldots, k_n) \in [m]^n$ and $\sigma \in S_n$. We have $\operatorname{rows}_{k_1,k_2,\ldots,k_n} B = \left(b_{k_i,j}\right)_{1\leq i\leq n,\ 1\leq j\leq n}$ (as we have seen in one of the previous footnotes) and $\operatorname{rows}_{k_{\sigma(1)},k_{\sigma(2)},\ldots,k_{\sigma(n)}} B = \left(b_{k_{\sigma(i)},j}\right)_{1\leq i\leq n,\ 1\leq j\leq n}$ (for similar reasons). Hence, we can apply Lemma 6.17 **(a)** to σ, $\operatorname{rows}_{k_1,k_2,\ldots,k_n} B$, $b_{k_i,j}$ and $\operatorname{rows}_{k_{\sigma(1)},k_{\sigma(2)},\ldots,k_{\sigma(n)}} B$ instead of κ, B, $b_{i,j}$ and B_κ. As a consequence, we obtain

$$\det\left(\operatorname{rows}_{k_{\sigma(1)},k_{\sigma(2)},\ldots,k_{\sigma(n)}} B\right) = (-1)^\sigma \cdot \det\left(\operatorname{rows}_{k_1,k_2,\ldots,k_n} B\right).$$

This proves (392).

[234]. Hence, (393) becomes

$$\begin{aligned}
&\det(AB) \\
&= \sum_{(g_1,g_2,\ldots,g_n)\in\mathbf{I}} \underbrace{\left(\sum_{\sigma\in S_n}\left(\prod_{i=1}^{n} a_{i,g_{\sigma(i)}}\right)(-1)^{\sigma}\right)}_{=\det\left(\operatorname{cols}_{g_1,g_2,\ldots,g_n} A\right)} \cdot \det\left(\operatorname{rows}_{g_1,g_2,\ldots,g_n} B\right) \\
&= \sum_{(g_1,g_2,\ldots,g_n)\in\mathbf{I}} \det\left(\operatorname{cols}_{g_1,g_2,\ldots,g_n} A\right)\cdot\det\left(\operatorname{rows}_{g_1,g_2,\ldots,g_n} B\right). \qquad (394)
\end{aligned}$$

Finally, we recall that $\mathbf{I}$ was defined as the set

$$\left\{(k_1,k_2,\ldots,k_n)\in[m]^n \mid k_1<k_2<\cdots<k_n\right\}.$$

Thus, summing over all $(g_1,g_2,\ldots,g_n)\in\mathbf{I}$ means the same as summing over all $(g_1,g_2,\ldots,g_n)\in[m]^n$ satisfying $g_1<g_2<\cdots<g_n$. In other words,

$$\sum_{(g_1,g_2,\ldots,g_n)\in\mathbf{I}} = \sum_{\substack{(g_1,g_2,\ldots,g_n)\in[m]^n;\\ g_1<g_2<\cdots<g_n}} = \sum_{1\leq g_1<g_2<\cdots<g_n\leq m}$$

(an equality between summation signs – hopefully its meaning is obvious). Hence, (394) becomes

$$\det(AB) = \sum_{1\leq g_1<g_2<\cdots<g_n\leq m} \det\left(\operatorname{cols}_{g_1,g_2,\ldots,g_n} A\right)\cdot\det\left(\operatorname{rows}_{g_1,g_2,\ldots,g_n} B\right).$$

This proves Theorem 6.32. □

6.6. Prelude to Laplace expansion

Next we shall show a fact which will allow us to compute some determinants by induction:

[234] *Proof.* Let $(g_1,g_2,\ldots,g_n)\in\mathbf{I}$. We have $A=(a_{i,j})_{1\leq i\leq n,\ 1\leq j\leq m}$. Thus, the definition of $\operatorname{cols}_{g_1,g_2,\ldots,g_n} A$ yields

$$\operatorname{cols}_{g_1,g_2,\ldots,g_n} A = \left(a_{i,g_y}\right)_{1\leq i\leq n,\ 1\leq y\leq n} = \left(a_{i,g_j}\right)_{1\leq i\leq n,\ 1\leq j\leq n}$$

(here, we renamed the index y as j). Hence, (342) (applied to $\operatorname{cols}_{g_1,g_2,\ldots,g_n} A$ and a_{i,g_j} instead of A and $a_{i,j}$) yields

$$\det\left(\operatorname{cols}_{g_1,g_2,\ldots,g_n} A\right) = \sum_{\sigma\in S_n}(-1)^{\sigma}\prod_{i=1}^{n} a_{i,g_{\sigma(i)}} = \sum_{\sigma\in S_n}\left(\prod_{i=1}^{n} a_{i,g_{\sigma(i)}}\right)(-1)^{\sigma},$$

qed.

Theorem 6.43. Let n be a positive integer. Let $A = \left(a_{i,j}\right)_{1\leq i\leq n,\ 1\leq j\leq n}$ be an $n \times n$-matrix. Assume that

$$a_{n,j} = 0 \qquad \text{for every } j \in \{1,2,\ldots,n-1\}. \tag{395}$$

Then, $\det A = a_{n,n} \cdot \det\left(\left(a_{i,j}\right)_{1\leq i\leq n-1,\ 1\leq j\leq n-1}\right)$.

The assumption (395) says that the last row of the matrix A consists entirely of zeroes, apart from its last entry $a_{n,n}$ (which may and may not be 0). Theorem 6.43 states that, under this assumption, the determinant can be obtained by multiplying this last entry $a_{n,n}$ with the determinant of the $(n-1)\times(n-1)$-matrix obtained by removing both the last row and the last column from A. For example, for $n = 3$, Theorem 6.43 states that

$$\det\begin{pmatrix} a & b & c \\ d & e & f \\ 0 & 0 & g \end{pmatrix} = g \det\begin{pmatrix} a & b \\ d & e \end{pmatrix}.$$

Theorem 6.43 is a particular case of *Laplace expansion*, which is a general recursive formula for the determinants that we will encounter further below. But Theorem 6.43 already has noticeable applications of its own, which is why I have chosen to start with this particular case.

The proof of Theorem 6.43 essentially relies on the following fact:

Lemma 6.44. Let n be a positive integer. Let $\left(a_{i,j}\right)_{1\leq i\leq n-1,\ 1\leq j\leq n-1}$ be an $(n-1)\times(n-1)$-matrix. Then,

$$\sum_{\substack{\sigma\in S_n;\\ \sigma(n)=n}} (-1)^{\sigma} \prod_{i=1}^{n-1} a_{i,\sigma(i)} = \det\left(\left(a_{i,j}\right)_{1\leq i\leq n-1,\ 1\leq j\leq n-1}\right).$$

Proof of Lemma 6.44. We define a subset T of S_n by

$$T = \{\tau \in S_n \mid \tau(n) = n\}.$$

(In other words, T is the set of all $\tau \in S_n$ such that if we write τ in one-line notation, then τ ends with an n.)

Now, we shall construct two mutually inverse maps between S_{n-1} and T.

For every $\sigma \in S_{n-1}$, we define a map $\widehat{\sigma} : \{1,2,\ldots,n\} \to \{1,2,\ldots,n\}$ by setting

$$\widehat{\sigma}(i) = \begin{cases} \sigma(i), & \text{if } i < n; \\ n, & \text{if } i = n \end{cases} \qquad \text{for every } i \in \{1,2,\ldots,n\}.$$

[235] It is straightforward to see that this map $\widehat{\sigma}$ is well-defined and belongs to T. Thus, we can define a map $\Phi : S_{n-1} \to T$ by setting

$$\Phi\left(\sigma\right) = \widehat{\sigma} \qquad \text{for every } \sigma \in S_{n-1}.$$

Loosely speaking, for every $\sigma \in S_{n-1}$, the permutation $\Phi\left(\sigma\right) = \widehat{\sigma} \in T$ is obtained by writing σ in one-line notation and appending n on its right. For example, if $n = 4$ and if $\sigma \in S_3$ is the permutation that is written as $\left(2,3,1\right)$ in one-line notation, then $\Phi\left(\sigma\right) = \widehat{\sigma}$ is the permutation that is written as $\left(2,3,1,4\right)$ in one-line notation.

On the other hand, for every $\gamma \in T$, we define a map $\overline{\gamma} : \left\{1,2,\ldots,n-1\right\} \to \left\{1,2,\ldots,n-1\right\}$ by setting

$$\overline{\gamma}\left(i\right) = \gamma\left(i\right) \qquad \text{for every } i \in \left\{1,2,\ldots,n-1\right\}.$$

It is straightforward to see that this map $\overline{\gamma}$ is well-defined and belongs to S_{n-1}. Hence, we can define a map $\Psi : T \to S_{n-1}$ by setting

$$\Psi\left(\gamma\right) = \overline{\gamma} \qquad \text{for every } \gamma \in T.$$

Loosely speaking, for every $\gamma \in T$, the permutation $\Psi\left(\gamma\right) = \overline{\gamma} \in S_{n-1}$ is obtained by writing γ in one-line notation and removing the n (which is the rightmost entry in the one-line notation, because $\gamma\left(n\right) = n$). For example, if $n = 4$ and if $\gamma \in S_4$ is the permutation that is written as $\left(2,3,1,4\right)$ in one-line notation, then $\Psi\left(\gamma\right) = \overline{\gamma}$ is the permutation that is written as $\left(2,3,1\right)$ in one-line notation.

The maps Φ and Ψ are mutually inverse.[236] Thus, the map Φ is a bijection.

It is fairly easy to see that every $\sigma \in S_{n-1}$ satisfies

$$\left(-1\right)^{\widehat{\sigma}} = \left(-1\right)^{\sigma} \tag{396}$$

[237] and

$$\prod_{i=1}^{n-1} a_{i,\widehat{\sigma}(i)} = \prod_{i=1}^{n-1} a_{i,\sigma(i)} \tag{397}$$

[235] Note that if we use Definition 5.55, then this map $\widehat{\sigma}$ is exactly the map $\sigma^{\left(\left\{1,2,\ldots,n-1\right\}\to\left\{1,2,\ldots,n\right\}\right)}$.

[236] This should be clear enough from the descriptions we gave using one-line notation. A formal proof is straightforward.

[237] *Proof of (396):* Let $\sigma \in S_{n-1}$. We want to prove that $\left(-1\right)^{\widehat{\sigma}} = \left(-1\right)^{\sigma}$. It is clearly sufficient to show that $\ell\left(\widehat{\sigma}\right) = \ell\left(\sigma\right)$ (because $\left(-1\right)^{\widehat{\sigma}} = \left(-1\right)^{\ell\left(\widehat{\sigma}\right)}$ and $\left(-1\right)^{\sigma} = \left(-1\right)^{\ell\left(\sigma\right)}$). In order to do so, it is sufficient to show that the inversions of $\widehat{\sigma}$ are precisely the inversions of σ (because $\ell\left(\widehat{\sigma}\right)$ is the number of inversions of $\widehat{\sigma}$, whereas $\ell\left(\sigma\right)$ is the number of inversions of σ).

If $\left(i,j\right)$ is an inversion of σ, then $\left(i,j\right)$ is an inversion of $\widehat{\sigma}$ (because if $\left(i,j\right)$ is an inversion of σ, then both i and j are $< n$, and thus the definition of $\widehat{\sigma}$ yields $\widehat{\sigma}\left(i\right) = \sigma\left(i\right)$ and $\widehat{\sigma}\left(j\right) = \sigma\left(j\right)$). In other words, every inversion of σ is an inversion of $\widehat{\sigma}$.

On the other hand, let $\left(u,v\right)$ be an inversion of $\widehat{\sigma}$. We shall prove that $\left(u,v\right)$ is an inversion of σ.

Indeed, $\left(u,v\right)$ is an inversion of $\widehat{\sigma}$. In other words, $\left(u,v\right)$ is a pair of integers satisfying $1 \leq u < v \leq n$ and $\widehat{\sigma}\left(u\right) > \widehat{\sigma}\left(v\right)$.

If we had $v = n$, then we would have $\widehat{\sigma}\left(u\right) > \widehat{\sigma}\left(\underbrace{v}_{=n}\right) = \widehat{\sigma}\left(n\right) = n$ (by the definition of

[238].

Now,

$$\begin{aligned}
&\underbrace{\sum_{\substack{\sigma\in S_n;\\ \sigma(n)=n}}}_{\substack{=\sum_{\sigma\in\{\tau\in S_n\ \mid\ \tau(n)=n\}}=\sum_{\sigma\in T}\\ \text{(since }\{\tau\in S_n\ \mid\ \tau(n)=n\}=T\text{)}}} (-1)^{\sigma}\prod_{i=1}^{n-1}a_{i,\sigma(i)} \\
&= \sum_{\sigma\in T}(-1)^{\sigma}\prod_{i=1}^{n-1}a_{i,\sigma(i)} = \sum_{\sigma\in S_{n-1}}\underbrace{(-1)^{\Phi(\sigma)}\prod_{i=1}^{n-1}a_{i,(\Phi(\sigma))(i)}}_{\substack{=(-1)^{\widehat{\sigma}}\prod_{i=1}^{n-1}a_{i,\widehat{\sigma}(i)}\\ \text{(since }\Phi(\sigma)=\widehat{\sigma}\text{)}}} \\
&\qquad\left(\begin{array}{c}\text{here, we have substituted }\Phi(\sigma)\text{ for }\sigma\text{ in the sum,}\\ \text{since the map }\Phi : S_{n-1}\to T\text{ is a bijection}\end{array}\right) \\
&= \sum_{\sigma\in S_{n-1}}\underbrace{(-1)^{\widehat{\sigma}}}_{\substack{=(-1)^{\sigma}\\ \text{(by (396))}}}\underbrace{\prod_{i=1}^{n-1}a_{i,\widehat{\sigma}(i)}}_{\substack{=\prod_{i=1}^{n-1}a_{i,\sigma(i)}\\ \text{(by (397))}}} = \sum_{\sigma\in S_{n-1}}(-1)^{\sigma}\prod_{i=1}^{n-1}a_{i,\sigma(i)}.
\end{aligned}$$

Compared with

$$\det\left(\left(a_{i,j}\right)_{1\leq i\leq n-1,\ 1\leq j\leq n-1}\right) = \sum_{\sigma\in S_{n-1}}(-1)^{\sigma}\prod_{i=1}^{n-1}a_{i,\sigma(i)}$$
$$\left(\begin{array}{c}\text{by (342), applied to } n-1 \text{ and}\\ \left(a_{i,j}\right)_{1\leq i\leq n-1,\ 1\leq j\leq n-1}\text{ instead of } n \text{ and } A\end{array}\right),$$

this yields

$$\sum_{\substack{\sigma\in S_n;\\ \sigma(n)=n}}(-1)^{\sigma}\prod_{i=1}^{n-1}a_{i,\sigma(i)} = \det\left(\left(a_{i,j}\right)_{1\leq i\leq n-1,\ 1\leq j\leq n-1}\right).$$

$\widehat{\sigma}$), which would contradict $\widehat{\sigma}(u)\in\{1,2,\ldots,n\}$. Thus, we cannot have $v=n$. We therefore have $v<n$, so that $v\leq n-1$. Now, $1\leq u<v\leq n-1$. Thus, both $\sigma(u)$ and $\sigma(v)$ are well-defined. The definition of $\widehat{\sigma}$ yields $\widehat{\sigma}(u)=\sigma(u)$ (since $u\leq n-1<n$) and $\widehat{\sigma}(v)=\sigma(v)$ (since $v\leq n-1<n$), so that $\sigma(u)=\widehat{\sigma}(u)>\widehat{\sigma}(v)=\sigma(v)$. Thus, (u,v) is a pair of integers satisfying $1\leq u<v\leq n-1$ and $\sigma(u)>\sigma(v)$. In other words, (u,v) is an inversion of σ.

We thus have shown that every inversion of $\widehat{\sigma}$ is an inversion of σ. Combining this with the fact that every inversion of σ is an inversion of $\widehat{\sigma}$, we thus conclude that the inversions of $\widehat{\sigma}$ are precisely the inversions of σ. As we have already said, this finishes the proof of (396).

[238] *Proof of (397):* Let $\sigma\in S_{n-1}$. The definition of $\widehat{\sigma}$ yields $\widehat{\sigma}(i)=\sigma(i)$ for every $i\in\{1,2,\ldots,n-1\}$. Thus, $a_{i,\widehat{\sigma}(i)}=a_{i,\sigma(i)}$ for every $i\in\{1,2,\ldots,n-1\}$. Hence, $\prod_{i=1}^{n-1}\underbrace{a_{i,\widehat{\sigma}(i)}}_{=a_{i,\sigma(i)}}=\prod_{i=1}^{n-1}a_{i,\sigma(i)}$, qed.

This proves Lemma 6.44. □

Proof of Theorem 6.43. Every permutation $\sigma \in S_n$ satisfying $\sigma(n) \neq n$ satisfies

$$a_{n,\sigma(n)} = 0 \qquad (398)$$

[239].

From (342), we obtain

$$\begin{aligned}
\det A &= \sum_{\sigma \in S_n} (-1)^{\sigma} \underbrace{\prod_{i=1}^{n} a_{i,\sigma(i)}}_{=\left(\prod_{i=1}^{n-1} a_{i,\sigma(i)}\right) a_{n,\sigma(n)}} = \sum_{\sigma \in S_n} (-1)^{\sigma} \left(\prod_{i=1}^{n-1} a_{i,\sigma(i)}\right) a_{n,\sigma(n)} \\
&= \sum_{\substack{\sigma \in S_n; \\ \sigma(n)=n}} (-1)^{\sigma} \left(\prod_{i=1}^{n-1} a_{i,\sigma(i)}\right) \underbrace{a_{n,\sigma(n)}}_{\substack{=a_{n,n} \\ \text{(since } \sigma(n)=n\text{)}}} + \sum_{\substack{\sigma \in S_n; \\ \sigma(n)\neq n}} (-1)^{\sigma} \left(\prod_{i=1}^{n-1} a_{i,\sigma(i)}\right) \underbrace{a_{n,\sigma(n)}}_{\substack{=0 \\ \text{(by (398))}}} \\
&\qquad \left(\begin{array}{c} \text{since every } \sigma \in S_n \text{ satisfies} \\ \text{either } \sigma(n) = n \text{ or } \sigma(n) \neq n \text{ (but not both)} \end{array}\right) \\
&= \sum_{\substack{\sigma \in S_n; \\ \sigma(n)=n}} (-1)^{\sigma} \left(\prod_{i=1}^{n-1} a_{i,\sigma(i)}\right) a_{n,n} + \underbrace{\sum_{\substack{\sigma \in S_n; \\ \sigma(n)\neq n}} (-1)^{\sigma} \left(\prod_{i=1}^{n-1} a_{i,\sigma(i)}\right) 0}_{=0} \\
&= \sum_{\substack{\sigma \in S_n; \\ \sigma(n)=n}} (-1)^{\sigma} \left(\prod_{i=1}^{n-1} a_{i,\sigma(i)}\right) a_{n,n} = a_{n,n} \cdot \underbrace{\sum_{\substack{\sigma \in S_n; \\ \sigma(n)=n}} (-1)^{\sigma} \prod_{i=1}^{n-1} a_{i,\sigma(i)}}_{\substack{=\det\left(\left(a_{i,j}\right)_{1\leq i\leq n-1,\ 1\leq j\leq n-1}\right) \\ \text{(by Lemma 6.44)}}} \\
&= a_{n,n} \cdot \det\left(\left(a_{i,j}\right)_{1\leq i\leq n-1,\ 1\leq j\leq n-1}\right).
\end{aligned}$$

This proves Theorem 6.43. □

Let us finally state an analogue of Theorem 6.43 in which the last column (rather than the last row) is required to consist mostly of zeroes:

Corollary 6.45. Let n be a positive integer. Let $A = \left(a_{i,j}\right)_{1\leq i\leq n,\ 1\leq j\leq n}$ be an $n \times n$-matrix. Assume that

$$a_{i,n} = 0 \qquad \text{for every } i \in \{1, 2, \ldots, n-1\}. \qquad (399)$$

Then, $\det A = a_{n,n} \cdot \det\left(\left(a_{i,j}\right)_{1\leq i\leq n-1,\ 1\leq j\leq n-1}\right)$.

[239] *Proof of (398):* Let $\sigma \in S_n$ be a permutation satisfying $\sigma(n) \neq n$. Since $\sigma(n) \in \{1, 2, \ldots, n\}$ and $\sigma(n) \neq n$, we have $\sigma(n) \in \{1, 2, \ldots, n\} \setminus \{n\} = \{1, 2, \ldots, n-1\}$. Hence, (395) (applied to $j = \sigma(n)$) shows that $a_{n,\sigma(n)} = 0$, qed.

Proof of Corollary 6.45. We have $n-1 \in \mathbb{N}$ (since n is a positive integer).

We have $A = \left(a_{i,j}\right)_{1\leq i\leq n,\ 1\leq j\leq n}$, and thus $A^T = \left(a_{j,i}\right)_{1\leq i\leq n,\ 1\leq j\leq n}$ (by the definition of the transpose matrix A^T). Also, for every $j \in \{1, 2, \ldots, n-1\}$, we have $a_{j,n} = 0$ (by (399), applied to $i = j$). Thus, Theorem 6.43 (applied to A^T and $a_{j,i}$ instead of A and $a_{i,j}$) yields

$$\det\left(A^T\right) = a_{n,n} \cdot \det\left(\left(a_{j,i}\right)_{1\leq i\leq n-1,\ 1\leq j\leq n-1}\right). \tag{400}$$

But Exercise 6.4 shows that $\det\left(A^T\right) = \det A$. Thus, $\det A = \det\left(A^T\right)$. Also, the definition of the transpose of a matrix shows that $\left(\left(a_{i,j}\right)_{1\leq i\leq n-1,\ 1\leq j\leq n-1}\right)^T = \left(a_{j,i}\right)_{1\leq i\leq n-1,\ 1\leq j\leq n-1}$. Thus,

$$\det\left(\left(\left(a_{i,j}\right)_{1\leq i\leq n-1,\ 1\leq j\leq n-1}\right)^T\right) = \det\left(\left(a_{j,i}\right)_{1\leq i\leq n-1,\ 1\leq j\leq n-1}\right).$$

Comparing this with

$$\det\left(\left(\left(a_{i,j}\right)_{1\leq i\leq n-1,\ 1\leq j\leq n-1}\right)^T\right) = \det\left(\left(a_{i,j}\right)_{1\leq i\leq n-1,\ 1\leq j\leq n-1}\right)$$

(by Exercise 6.4, applied to $n-1$ and $\left(a_{i,j}\right)_{1\leq i\leq n-1,\ 1\leq j\leq n-1}$ instead of n and A), we obtain

$$\det\left(\left(a_{j,i}\right)_{1\leq i\leq n-1,\ 1\leq j\leq n-1}\right) = \det\left(\left(a_{i,j}\right)_{1\leq i\leq n-1,\ 1\leq j\leq n-1}\right).$$

Now,

$$\begin{aligned}\det A = \det\left(A^T\right) &= a_{n,n} \cdot \underbrace{\det\left(\left(a_{j,i}\right)_{1\leq i\leq n-1,\ 1\leq j\leq n-1}\right)}_{=\det\left(\left(a_{i,j}\right)_{1\leq i\leq n-1,\ 1\leq j\leq n-1}\right)} \qquad \text{(by (400))} \\ &= a_{n,n} \cdot \det\left(\left(a_{i,j}\right)_{1\leq i\leq n-1,\ 1\leq j\leq n-1}\right).\end{aligned}$$

This proves Corollary 6.45. □

6.7. The Vandermonde determinant

6.7.1. The statement

An example for an application of Theorem 6.43 is the famous *Vandermonde determinant*:

Theorem 6.46. Let $n \in \mathbb{N}$. Let $x_1, x_2, \ldots, x_n$ be n elements of $\mathbb{K}$. Then:

(a) We have

$$\det\left(\left(x_i^{n-j}\right)_{1\leq i\leq n,\ 1\leq j\leq n}\right) = \prod_{1\leq i<j\leq n} \left(x_i - x_j\right).$$

(b) We have

$$\det\left(\left(x_j^{n-i}\right)_{1\leq i\leq n,\ 1\leq j\leq n}\right)=\prod_{1\leq i<j\leq n}\left(x_i-x_j\right).$$

(c) We have

$$\det\left(\left(x_i^{j-1}\right)_{1\leq i\leq n,\ 1\leq j\leq n}\right)=\prod_{1\leq j<i\leq n}\left(x_i-x_j\right).$$

(d) We have

$$\det\left(\left(x_j^{i-1}\right)_{1\leq i\leq n,\ 1\leq j\leq n}\right)=\prod_{1\leq j<i\leq n}\left(x_i-x_j\right).$$

Remark 6.47. For $n=4$, the four matrices appearing in Theorem 6.46 are

$$\left(x_i^{n-j}\right)_{1\leq i\leq n,\ 1\leq j\leq n}=\begin{pmatrix} x_1^3 & x_1^2 & x_1 & 1 \\ x_2^3 & x_2^2 & x_2 & 1 \\ x_3^3 & x_3^2 & x_3 & 1 \\ x_4^3 & x_4^2 & x_4 & 1 \end{pmatrix},$$

$$\left(x_j^{n-i}\right)_{1\leq i\leq n,\ 1\leq j\leq n}=\begin{pmatrix} x_1^3 & x_2^3 & x_3^3 & x_4^3 \\ x_1^2 & x_2^2 & x_3^2 & x_4^2 \\ x_1 & x_2 & x_3 & x_4 \\ 1 & 1 & 1 & 1 \end{pmatrix},$$

$$\left(x_i^{j-1}\right)_{1\leq i\leq n,\ 1\leq j\leq n}=\begin{pmatrix} 1 & x_1 & x_1^2 & x_1^3 \\ 1 & x_2 & x_2^2 & x_2^3 \\ 1 & x_3 & x_3^2 & x_3^3 \\ 1 & x_4 & x_4^2 & x_4^3 \end{pmatrix},$$

$$\left(x_j^{i-1}\right)_{1\leq i\leq n,\ 1\leq j\leq n}=\begin{pmatrix} 1 & 1 & 1 & 1 \\ x_1 & x_2 & x_3 & x_4 \\ x_1^2 & x_2^2 & x_3^2 & x_4^2 \\ x_1^3 & x_2^3 & x_3^3 & x_4^3 \end{pmatrix}.$$

It is clear that the second of these four matrices is the transpose of the first; the fourth is the transpose of the third; and the fourth is obtained from the second by rearranging the rows in opposite order. Thus, the four parts of Theorem 6.46 are rather easily seen to be equivalent. (We shall prove part **(a)** and derive the others from it.) Nevertheless it is useful to have seen them all.

Theorem 6.46 is a classical result (known as the *Vandermonde determinant*, although it is unclear whether it has been proven by Vandermonde): Almost all texts on linear algebra mention it (or, rather, at least one of its four parts), although some only prove it in lesser generality. It is a fundamental result that has various applications to abstract algebra, number theory, coding theory, combinatorics and numerical mathematics.

Theorem 6.46 has many known proofs[240]. In this section, I will show two of these proofs. Another proof (of Theorem 6.46 **(c)** only; but as I said above, the other parts are easily seen to be equivalent) can be found in Section **??** or in [Grinbe10, Theorem 1]. Before I get to the proofs, let me show yet another down-to-earth example of Theorem 6.46:

Example 6.48. Let $x, y, z \in \mathbb{K}$. Let $A = \begin{pmatrix} 1 & x & x^2 \\ 1 & y & y^2 \\ 1 & z & z^2 \end{pmatrix}$. Then, (344) shows that

$$\begin{aligned} \det A &= 1yz^2 + xy^2 \cdot 1 + x^2 \cdot 1z - 1y^2z - x \cdot 1z^2 - x^2y \cdot 1 \\ &= yz^2 + xy^2 + x^2z - y^2z - xz^2 - x^2y \\ &= yz\left(z-y\right) + zx\left(x-z\right) + xy\left(y-x\right). \end{aligned} \tag{401}$$

On the other hand, Theorem 6.46 **(c)** (applied to $n = 3$, $x_1 = x$, $x_2 = y$ and $x_3 = z$) yields $\det A = (y-x)(z-x)(z-y)$. Compared with (401), this yields

$$(y-x)(z-x)(z-y) = yz(z-y) + zx(x-z) + xy(y-x). \tag{402}$$

You might have encountered this curious identity as a trick of use in contest problems. When x, y, z are three distinct complex numbers, we can divide (402) by $(y-x)(z-x)(z-y)$, and obtain

$$1 = \frac{yz}{(y-x)(z-x)} + \frac{zx}{(z-y)(x-y)} + \frac{xy}{(x-z)(y-z)}.$$

6.7.2. A proof by induction

We now approach the proofs of Theorem 6.46. The first proof has the advantage of demonstrating how Theorem 6.43 can be used (together with induction) in computing determinants. Before we embark on this proof, let us see what happens to the determinant of an arbitrary square matrix if we rearrange the rows in opposite order:

[240] For four combinatorial proofs, see [Gessel79], [Aigner07, §5.3], [Loehr11, §12.9] and [BenDre07]. (Specifically, [Gessel79] and [BenDre07] prove Theorem 6.46 **(c)**, whereas [Aigner07, §5.3] and [Loehr11, §12.9] prove Theorem 6.46 **(b)**. But as we will see, the four parts of Theorem 6.46 are easily seen to be equivalent to each other.) Gessel's proof from [Gessel79] is also explained in more detail in [Grinbe22].

Lemma 6.49. Let $n \in \mathbb{N}$. Let $\left(a_{i,j}\right)_{1\leq i\leq n,\ 1\leq j\leq n}$ be an $n \times n$-matrix. Then,

$$\det\left(\left(a_{n+1-i,j}\right)_{1\leq i\leq n,\ 1\leq j\leq n}\right) = (-1)^{n(n-1)/2} \det\left(\left(a_{i,j}\right)_{1\leq i\leq n,\ 1\leq j\leq n}\right).$$

Proof of Lemma 6.49. Let $[n]$ denote the set $\{1,2,\ldots,n\}$. Define a permutation w_0 in S_n as in Exercise 5.11. In the solution of Exercise 5.11, we have shown that $(-1)^{w_0} = (-1)^{n(n-1)/2}$.

Now, we can apply Lemma 6.17 **(a)** to $\left(a_{i,j}\right)_{1\leq i\leq n,\ 1\leq j\leq n}$, w_0 and $\left(a_{w_0(i),j}\right)_{1\leq i\leq n,\ 1\leq j\leq n}$ instead of B, κ and B_κ. As a result, we obtain

$$\begin{aligned}\det\left(\left(a_{w_0(i),j}\right)_{1\leq i\leq n,\ 1\leq j\leq n}\right) &= \underbrace{(-1)^{w_0}}_{=(-1)^{n(n-1)/2}} \cdot \det\left(\left(a_{i,j}\right)_{1\leq i\leq n,\ 1\leq j\leq n}\right)\\ &= (-1)^{n(n-1)/2} \det\left(\left(a_{i,j}\right)_{1\leq i\leq n,\ 1\leq j\leq n}\right). \qquad (403)\end{aligned}$$

But $w_0(i) = n+1-i$ for every $i \in \{1,2,\ldots,n\}$ (by the definition of w_0). Thus, (403) rewrites as $\det\left(\left(a_{n+1-i,j}\right)_{1\leq i\leq n,\ 1\leq j\leq n}\right) = (-1)^{n(n-1)/2} \det\left(\left(a_{i,j}\right)_{1\leq i\leq n,\ 1\leq j\leq n}\right)$. This proves Lemma 6.49. □

First proof of Theorem 6.46. **(a)** For every $u \in \{0,1,\ldots,n\}$, let A_u be the $u \times u$-matrix $\left(x_i^{u-j}\right)_{1\leq i\leq u,\ 1\leq j\leq u}$.

Now, let us show that

$$\det(A_u) = \prod_{1\leq i<j\leq u} (x_i - x_j) \qquad (404)$$

for every $u \in \{0,1,\ldots,n\}$.

[*Proof of (404):* We will prove (404) by induction over u:

Induction base: The matrix A_0 is a 0×0-matrix and thus has determinant $\det(A_0) = 1$. On the other hand, the product $\prod_{1\leq i<j\leq 0} (x_i - x_j)$ is an empty product (i.e., a product of 0 elements of $\mathbb{K}$) and thus equals 1 as well. Hence, both $\det(A_0)$ and $\prod_{1\leq i<j\leq 0} (x_i - x_j)$ equal 1. Thus, $\det(A_0) = \prod_{1\leq i<j\leq 0} (x_i - x_j)$. In other words, (404) holds for $u = 0$. The induction base is thus complete.

Induction step: Let $U \in \{1,2,\ldots,n\}$. Assume that (404) holds for $u = U-1$. We need to prove that (404) holds for $u = U$.

Recall that $A_U = \left(x_i^{U-j}\right)_{1\leq i\leq U,\ 1\leq j\leq U}$ (by the definition of A_U).

For every $(i,j) \in \{1,2,\ldots,U\}^2$, define $b_{i,j} \in \mathbb{K}$ by

$$b_{i,j} = \begin{cases} x_i^{U-j} - x_U x_i^{U-j-1}, & \text{if } j < U; \\ 1, & \text{if } j = U \end{cases}.$$

Let B be the $U \times U$-matrix $\left(b_{i,j}\right)_{1\leq i\leq U,\ 1\leq j\leq U}$. For example, if $U = 4$, then

$$A_U = \begin{pmatrix} x_1^3 & x_1^2 & x_1 & 1 \\ x_2^3 & x_2^2 & x_2 & 1 \\ x_3^3 & x_3^2 & x_3 & 1 \\ x_4^3 & x_4^2 & x_4 & 1 \end{pmatrix} \qquad \text{and}$$

$$B = \begin{pmatrix} x_1^3 - x_4x_1^2 & x_1^2 - x_4x_1 & x_1 - x_4 & 1 \\ x_2^3 - x_4x_2^2 & x_2^2 - x_4x_2 & x_2 - x_4 & 1 \\ x_3^3 - x_4x_3^2 & x_3^2 - x_4x_3 & x_3 - x_4 & 1 \\ x_4^3 - x_4x_4^2 & x_4^2 - x_4x_4 & x_4 - x_4 & 1 \end{pmatrix}$$
$$= \begin{pmatrix} x_1^3 - x_4x_1^2 & x_1^2 - x_4x_1 & x_1 - x_4 & 1 \\ x_2^3 - x_4x_2^2 & x_2^2 - x_4x_2 & x_2 - x_4 & 1 \\ x_3^3 - x_4x_3^2 & x_3^2 - x_4x_3 & x_3 - x_4 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

We claim that $\det B = \det\left(A_U\right)$. Indeed, here are two ways to prove this:

First proof of $\det B = \det\left(A_U\right)$*:* Exercise 6.8 **(b)** shows that the determinant of a $U \times U$-matrix does not change if we subtract a multiple of one of its columns from another column. Now, let us subtract x_U times the 2-nd column of A_U from the 1-st column, then subtract x_U times the 3-rd column of the resulting matrix from the 2-nd column, and so on, all the way until we finally subtract x_U times the U-th column of the matrix from the $(U-1)$-st column[241]. The resulting matrix is B (according to our definition of B). Thus, $\det B = \det\left(A_U\right)$ (since our subtractions never change the determinant). This proves $\det B = \det\left(A_U\right)$.

Second proof of $\det B = \det\left(A_U\right)$*:* Here is another way to prove that $\det B = \det\left(A_U\right)$, with some less handwaving.

For every $(i,j) \in \{1,2,\ldots,U\}^2$, we define $c_{i,j} \in \mathbb{K}$ by

$$c_{i,j} = \begin{cases} 1, & \text{if } i = j; \\ -x_U, & \text{if } i = j+1; \\ 0, & \text{otherwise} \end{cases}.$$

Let C be the $U \times U$-matrix $\left(c_{i,j}\right)_{1\leq i\leq U,\ 1\leq j\leq U}$.

For example, if $U = 4$, then

$$C = \begin{pmatrix} 1 & 0 & 0 & 0 \\ -x_4 & 1 & 0 & 0 \\ 0 & -x_4 & 1 & 0 \\ 0 & 0 & -x_4 & 1 \end{pmatrix}.$$

[241] So, all in all, we subtract the x_U-multiple of each column from its neighbor to its left, but the order in which we are doing it (namely, from left to right) is important: It means that the column we are subtracting is unchanged from A_U. (If we would be doing these subtractions from right to left instead, then the columns to be subtracting would be changed by the preceding steps.)

The matrix C is lower-triangular, and thus Exercise 6.3 shows that its determinant is $\det C = \underbrace{c_{1,1}}_{=1} \underbrace{c_{2,2}}_{=1} \cdots \underbrace{c_{U,U}}_{=1} = 1$.

On the other hand, it is easy to see that $B = A_U C$ (check this!). Thus, Theorem 6.23 yields $\det B = \det(A_U) \cdot \underbrace{\det C}_{=1} = \det(A_U)$. So we have proven $\det B = \det(A_U)$ again.

[*Remark:* It is instructive to compare the two proofs of $\det B = \det(A_U)$ given above. They are close kin, although they might look different at first. In the first proof, we argued that B can be obtained from A_U by subtracting multiples of some columns from others; in the second, we argued that $B = A_U C$ for a specific lower-triangular matrix C. But a look at the matrix C makes it clear that multiplying a $U \times U$-matrix with C on the right (i.e., transforming a $U \times U$-matrix X into the matrix XC) is tantamount to subtracting multiples of some columns from others, in the way we did it to A_U to obtain B. So the main difference between the two proofs is that the first proof used a step-by-step procedure to obtain B from A_U, whereas the second proof obtained B from A_U by a single-step operation (namely, multiplication by a matrix C).]

Next, we observe that for every $j \in \{1, 2, \ldots, U-1\}$, we have

$$\begin{aligned} b_{U,j} &= \begin{cases} x_U^{U-j} - x_U x_U^{U-j-1}, & \text{if } j < U; \\ 1, & \text{if } j = U \end{cases} \qquad (\text{by the definition of } b_{U,j}) \\ &= x_U^{U-j} - \underbrace{x_U x_U^{U-j-1}}_{= x_U^{(U-j-1)+1} = x_U^{U-j}} \qquad (\text{since } j < U \text{ (since } j \in \{1, 2, \ldots, U-1\})) \\ &= x_U^{U-j} - x_U^{U-j} = 0. \end{aligned}$$

Hence, Theorem 6.43 (applied to U, B and $b_{i,j}$ instead of n, A and $a_{i,j}$) yields

$$\det B = b_{U,U} \cdot \det\left(\left(b_{i,j}\right)_{1 \leq i \leq U-1,\ 1 \leq j \leq U-1} \right). \tag{405}$$

Let B' denote the $(U-1) \times (U-1)$-matrix $\left(b_{i,j}\right)_{1 \leq i \leq U-1,\ 1 \leq j \leq U-1}$.

The definition of $b_{U,U}$ yields

$$\begin{aligned} b_{U,U} &= \begin{cases} x_U^{U-U} - x_U x_U^{U-U-1}, & \text{if } U < U; \\ 1, & \text{if } U = U \end{cases} \qquad (\text{by the definition of } b_{U,U}) \\ &= 1 \qquad (\text{since } U = U). \end{aligned}$$

Thus, (405) becomes

$$\det B = \underbrace{b_{U,U}}_{=1} \cdot \det \left(\underbrace{\left(b_{i,j}\right)_{1 \leq i \leq U-1,\ 1 \leq j \leq U-1}}_{=B'} \right) = \det\left(B'\right).$$

Compared with $\det B = \det(A_U)$, this yields

$$\det(A_U) = \det(B'). \tag{406}$$

Now, let us take a closer look at B'. Indeed, every $(i,j) \in \{1,2,\ldots,U-1\}^2$ satisfies

$$\begin{aligned} b_{i,j} &= \begin{cases} x_i^{U-j} - x_U x_i^{U-j-1}, & \text{if } j < U; \\ 1, & \text{if } j = U \end{cases} \qquad \left(\text{by the definition of } b_{i,j}\right) \\ &= \underbrace{x_i^{U-j}}_{=x_i x_i^{U-j-1}} - x_U x_i^{U-j-1} \qquad \left(\begin{array}{c} \text{since } j < U \text{ (since } j \in \{1,2,\ldots,U-1\} \\ \text{(since } (i,j) \in \{1,2,\ldots,U-1\}^2\text{))} \end{array}\right) \\ &= x_i x_i^{U-j-1} - x_U x_i^{U-j-1} = (x_i - x_U) \underbrace{x_i^{U-j-1}}_{=x_i^{(U-1)-j}} = (x_i - x_U) x_i^{(U-1)-j}. \end{aligned} \tag{407}$$

Hence,

$$B' = \left(\underbrace{b_{i,j}}_{\substack{=(x_i - x_U)x_i^{(U-1)-j} \\ \text{(by (407))}}} \right)_{1 \le i \le U-1,\ 1 \le j \le U-1} = \left((x_i - x_U) x_i^{(U-1)-j} \right)_{1 \le i \le U-1,\ 1 \le j \le U-1}. \tag{408}$$

On the other hand, the definition of A_{U-1} yields

$$A_{U-1} = \left(x_i^{(U-1)-j} \right)_{1 \le i \le U-1,\ 1 \le j \le U-1}. \tag{409}$$

Now, we claim that

$$\det(B') = \det(A_{U-1}) \cdot \prod_{i=1}^{U-1} (x_i - x_U). \tag{410}$$

Indeed, here are two ways to prove this:

First proof of (410): Comparing the formulas (408) and (409), we see that the matrix B' is obtained from the matrix A_{U-1} by multiplying the first row by $x_1 - x_U$, the second row by $x_2 - x_U$, and so on, and finally the $(U-1)$-st row by $x_{U-1} - x_U$. But every time we multiply a row of a $(U-1) \times (U-1)$-matrix by some scalar $\lambda \in \mathbb{K}$, the determinant of the matrix gets multiplied by λ (because of Exercise 6.7 **(g)**). Hence, the determinant of B' is obtained from that of A_{U-1} by first multiplying by $x_1 - x_U$, then multiplying by $x_2 - x_U$, and so on, and finally multiplying with $x_{U-1} - x_U$. In other words,

$$\det(B') = \det(A_{U-1}) \cdot \prod_{i=1}^{U-1} (x_i - x_U).$$

This proves (410).

Second proof of (410): For every $(i,j) \in \{1,2,\ldots,U-1\}^2$, we define $d_{i,j} \in \mathbb{K}$ by

$$d_{i,j} = \begin{cases} x_i - x_U, & \text{if } i = j; \\ 0, & \text{otherwise} \end{cases}.$$

Let D be the $(U-1) \times (U-1)$-matrix $\left(d_{i,j}\right)_{1 \leq i \leq U-1,\ 1 \leq j \leq U-1}$.

For example, if $U = 4$, then

$$D = \begin{pmatrix} x_1 - x_4 & 0 & 0 \\ 0 & x_2 - x_4 & 0 \\ 0 & 0 & x_3 - x_4 \end{pmatrix}.$$

The matrix D is lower-triangular (actually, diagonal[242]), and thus Exercise 6.3 shows that its determinant is

$$\det D = (x_1 - x_U)(x_2 - x_U) \cdots (x_{U-1} - x_U) = \prod_{i=1}^{U-1} (x_i - x_U).$$

On the other hand, it is easy to see that $B' = DA_{U-1}$ (check this!). Thus, Theorem 6.23 yields

$$\det(B') = \det D \cdot \det(A_{U-1}) = \det(A_{U-1}) \cdot \underbrace{\det D}_{= \prod_{i=1}^{U-1} (x_i - x_U)} = \det(A_{U-1}) \cdot \prod_{i=1}^{U-1} (x_i - x_U).$$

Thus, (410) is proven again.

[*Remark:* Again, our two proofs of (410) are closely related: the first one reveals B' as the result of a step-by-step process applied to A_{U-1}, while the second shows how B' can be obtained from A_{U-1} by a single multiplication. However, here (in contrast to the proofs of $\det B = \det(A_U)$), the step-by-step process involves transforming rows (not columns), and the multiplication is a multiplication from the left (we have $B' = DA_{U-1}$, not $B' = A_{U-1}D$).]

Now, (406) becomes

$$\det(A_U) = \det(B') = \det(A_{U-1}) \cdot \prod_{i=1}^{U-1} (x_i - x_U). \tag{411}$$

But we have assumed that (404) holds for $u = U - 1$. In other words,

$$\det(A_{U-1}) = \underbrace{\prod_{1 \leq i < j \leq U-1}}_{= \prod_{j=1}^{U-1} \prod_{i=1}^{j-1}} (x_i - x_j) = \prod_{j=1}^{U-1} \prod_{i=1}^{j-1} (x_i - x_j).$$

[242] A square matrix $E = \left(e_{i,j}\right)_{1 \leq i \leq n,\ 1 \leq j \leq n}$ is said to be *diagonal* if every $(i,j) \in \{1,2,\ldots,n\}^2$ satisfying $i \neq j$ satisfies $e_{i,j} = 0$. In other words, a square matrix is said to be *diagonal* if it is both upper-triangular and lower-triangular.

Hence, (411) yields

$$\begin{aligned}\det\left(A_U\right) &= \underbrace{\det\left(A_{U-1}\right)}_{=\prod_{j=1}^{U-1}\prod_{i=1}^{j-1}\left(x_i-x_j\right)} \cdot \prod_{i=1}^{U-1}\left(x_i-x_U\right) \\ &= \left(\prod_{j=1}^{U-1}\prod_{i=1}^{j-1}\left(x_i-x_j\right)\right)\cdot\prod_{i=1}^{U-1}\left(x_i-x_U\right).\end{aligned}$$

Compared with

$$\underbrace{\prod_{1\leq i<j\leq U}}_{=\prod_{j=1}^{U}\prod_{i=1}^{j-1}}\left(x_i-x_j\right) = \prod_{j=1}^{U}\prod_{i=1}^{j-1}\left(x_i-x_j\right) = \left(\prod_{j=1}^{U-1}\prod_{i=1}^{j-1}\left(x_i-x_j\right)\right)\cdot\prod_{i=1}^{U-1}\left(x_i-x_U\right)$$

$$\left(\text{here, we have split off the factor for } j=U \text{ from the product}\right),$$

this yields $\det\left(A_U\right) = \prod_{1\leq i<j\leq U}\left(x_i-x_j\right)$. In other words, (404) holds for $u=U$. This completes the induction step.

Now, (404) is proven by induction.]

Hence, we can apply (404) to $u=n$. As the result, we obtain $\det\left(A_n\right) = \prod_{1\leq i<j\leq n}\left(x_i-x_j\right)$. Since $A_n = \left(x_i^{n-j}\right)_{1\leq i\leq n,\ 1\leq j\leq n}$ (by the definition of A_n), this rewrites as $\det\left(\left(x_i^{n-j}\right)_{1\leq i\leq n,\ 1\leq j\leq n}\right) = \prod_{1\leq i<j\leq n}\left(x_i-x_j\right)$. This proves Theorem 6.46 **(a)**.

(b) The definition of the transpose of a matrix yields $\left(\left(x_j^{n-i}\right)_{1\leq i\leq n,\ 1\leq j\leq n}\right)^T = \left(x_i^{n-j}\right)_{1\leq i\leq n,\ 1\leq j\leq n}$. Hence,

$$\det\left(\underbrace{\left(\left(x_j^{n-i}\right)_{1\leq i\leq n,\ 1\leq j\leq n}\right)^T}_{=\left(x_i^{n-j}\right)_{1\leq i\leq n,\ 1\leq j\leq n}}\right) = \det\left(\left(x_i^{n-j}\right)_{1\leq i\leq n,\ 1\leq j\leq n}\right) = \prod_{1\leq i<j\leq n}\left(x_i-x_j\right)$$

(by Theorem 6.46 **(a)**). Compared with

$$\det\left(\left(\left(x_j^{n-i}\right)_{1\leq i\leq n,\ 1\leq j\leq n}\right)^T\right) = \det\left(\left(x_j^{n-i}\right)_{1\leq i\leq n,\ 1\leq j\leq n}\right)$$

(by Exercise 6.4, applied to $A = \left(x_j^{n-i}\right)_{1\leq i\leq n,\ 1\leq j\leq n}$), this yields

$$\det\left(\left(x_j^{n-i}\right)_{1\leq i\leq n,\ 1\leq j\leq n}\right) = \prod_{1\leq i<j\leq n}\left(x_i - x_j\right).$$

This proves Theorem 6.46 **(b)**.

(d) Applying Lemma 6.49 to $a_{i,j} = x_j^{n-i}$, we obtain

$$\det\left(\left(x_j^{n-(n+1-i)}\right)_{1\leq i\leq n,\ 1\leq j\leq n}\right) = (-1)^{n(n-1)/2} \underbrace{\det\left(\left(x_j^{n-i}\right)_{1\leq i\leq n,\ 1\leq j\leq n}\right)}_{\substack{=\prod_{1\leq i<j\leq n}\left(x_i-x_j\right) \\ \text{(by Theorem 6.46 (b))}}}$$
$$= (-1)^{n(n-1)/2} \prod_{1\leq i<j\leq n}\left(x_i - x_j\right).$$

This rewrites as

$$\det\left(\left(x_j^{i-1}\right)_{1\leq i\leq n,\ 1\leq j\leq n}\right) = (-1)^{n(n-1)/2} \prod_{1\leq i<j\leq n}\left(x_i - x_j\right) \qquad (412)$$

(since every $(i,j) \in \{1,2,\ldots,n\}^2$ satisfies $x_j^{n-(n+1-i)} = x_j^{i-1}$).

Now, in the solution to Exercise 5.11, we have shown that the number of all pairs (i,j) of integers satisfying $1 \leq i < j \leq n$ is $n(n-1)/2$. In other words,

$$\left(\text{the number of all } (i,j) \in \{1,2,\ldots,n\}^2 \text{ such that } i<j\right) = n(n-1)/2. \qquad (413)$$

Now,

$$\prod_{1\leq j<i\leq n}\left(x_i - x_j\right) = \prod_{1\leq i<j\leq n}\underbrace{\left(x_j - x_i\right)}_{=(-1)\left(x_i-x_j\right)} \qquad \left(\begin{array}{c}\text{here, we renamed the index } (j,i) \\ \text{as } (i,j) \text{ in the product}\end{array}\right)$$
$$= \prod_{1\leq i<j\leq n}\left((-1)\left(x_i - x_j\right)\right)$$
$$= \underbrace{(-1)^{\left(\text{the number of all } (i,j)\in\{1,2,\ldots,n\}^2 \text{ such that } i<j\right)}}_{\substack{=(-1)^{n(n-1)/2} \\ \text{(by (413))}}} \prod_{1\leq i<j\leq n}\left(x_i - x_j\right)$$
$$= (-1)^{n(n-1)/2} \prod_{1\leq i<j\leq n}\left(x_i - x_j\right).$$

Compared with (412), this yields $\det\left(\left(x_j^{i-1}\right)_{1\leq i\leq n,\ 1\leq j\leq n}\right) = \prod_{1\leq j<i\leq n}\left(x_i - x_j\right)$. This proves Theorem 6.46 **(d)**.

(c) We can derive Theorem 6.46 **(c)** from Theorem 6.46 **(d)** in the same way as we derived part **(b)** from **(a)**. □

6.7.3. A proof by factoring the matrix

Next, I shall outline another proof of Theorem 6.46, which proceeds by writing the matrix $\left(x_j^{i-1}\right)_{1\leq i\leq n,\ 1\leq j\leq n}$ as a product of a lower-triangular matrix with an upper-triangular matrix. The idea of this proof appears in [OruPhi00, Theorem 2.1], [GohKol96, Theorem 2] and [OlvSha18, proof of Lemma 5.16] (although the first two of these three sources use slightly different arguments, and the third gives only a hint of the proof).

We will need several lemmas for the proof. The proofs of these lemmas are relegated to the solution of Exercise 6.15.

We begin with a definition that will be used throughout Subsection 6.7.3:

Definition 6.50. Let $k \in \mathbb{Z}$ and $n \in \mathbb{N}$. Let $x_1, x_2, \ldots, x_n$ be n elements of $\mathbb{K}$. Then, we define an element $h_k\left(x_1, x_2, \ldots, x_n\right) \in \mathbb{K}$ by

$$h_k\left(x_1, x_2, \ldots, x_n\right) = \sum_{\substack{(a_1,a_2,\ldots,a_n)\in\mathbb{N}^n;\\ a_1+a_2+\cdots+a_n=k}} x_1^{a_1} x_2^{a_2} \cdots x_n^{a_n}.$$

(Note that the sum on the right hand side of this equality is finite, because only finitely many $\left(a_1, a_2, \ldots, a_n\right) \in \mathbb{N}^n$ satisfy $a_1 + a_2 + \cdots + a_n = k$.)

The element $h_k\left(x_1, x_2, \ldots, x_n\right)$ defined in Definition 6.50 is often called the *k-th complete homogeneous function of the n elements $x_1, x_2, \ldots, x_n$* (although, more often, this notion is reserved for a different, more abstract object, of whom $h_k\left(x_1, x_2, \ldots, x_n\right)$ is just an evaluation). Let us see some examples:

Example 6.51. **(a)** If $n \in \mathbb{N}$, and if $x_1, x_2, \ldots, x_n$ are n elements of $\mathbb{K}$, then

$$h_1\left(x_1, x_2, \ldots, x_n\right) = x_1 + x_2 + \cdots + x_n$$

and

$$h_2\left(x_1, x_2, \ldots, x_n\right) = \left(x_1^2 + x_2^2 + \cdots + x_n^2\right) + \sum_{1\leq i<j\leq n} x_i x_j.$$

For example, for $n = 3$, we obtain $h_2\left(x_1, x_2, x_3\right) = \left(x_1^2 + x_2^2 + x_3^2\right) + \left(x_1x_2 + x_1x_3 + x_2x_3\right)$.

(b) If x and y are two elements of $\mathbb{K}$, then

$$h_k\left(x, y\right) = \sum_{\substack{(a_1,a_2)\in\mathbb{N}^2;\\ a_1+a_2=k}} x^{a_1} y^{a_2} = x^k y^0 + x^{k-1} y^1 + \cdots + x^0 y^k$$

for every $k \in \mathbb{N}$.

(c) If $x \in \mathbb{K}$, then $h_k\left(x\right) = x^k$ for every $k \in \mathbb{N}$.

Lemma 6.52. Let $n \in \mathbb{N}$. Let $x_1, x_2, \ldots, x_n$ be n elements of $\mathbb{K}$.

(a) We have $h_k(x_1, x_2, \ldots, x_n) = 0$ for every negative integer k.

(b) We have $h_0(x_1, x_2, \ldots, x_n) = 1$.

As we have said, all lemmas in Subsection 6.7.3 will be proven in the solution to Exercise 6.15.

Three further lemmas on the $h_k(x_1, x_2, \ldots, x_n)$ will be of use:

Lemma 6.53. Let k be a positive integer. Let $x_1, x_2, \ldots, x_k$ be k elements of $\mathbb{K}$. Let $q \in \mathbb{Z}$. Then,

$$h_q(x_1, x_2, \ldots, x_k) = \sum_{r=0}^{q} x_k^r h_{q-r}(x_1, x_2, \ldots, x_{k-1}).$$

Lemma 6.54. Let k be a positive integer. Let $x_1, x_2, \ldots, x_k$ be k elements of $\mathbb{K}$. Let $q \in \mathbb{Z}$. Then,

$$h_q(x_1, x_2, \ldots, x_k) = h_q(x_1, x_2, \ldots, x_{k-1}) + x_k h_{q-1}(x_1, x_2, \ldots, x_k).$$

Lemma 6.55. Let i be a positive integer. Let $x_1, x_2, \ldots, x_i$ be i elements of $\mathbb{K}$. Let $u \in \mathbb{K}$. Then,

$$\sum_{k=1}^{i} h_{i-k}(x_1, x_2, \ldots, x_k) \prod_{p=1}^{k-1} (u - x_p) = u^{i-1}.$$

Next, let us introduce two matrices:

Lemma 6.56. Let $n \in \mathbb{N}$. Let $x_1, x_2, \ldots, x_n$ be n elements of $\mathbb{K}$. Define an $n \times n$-matrix $U \in \mathbb{K}^{n \times n}$ by

$$U = \left(\prod_{p=1}^{i-1} (x_j - x_p) \right)_{1 \leq i \leq n,\ 1 \leq j \leq n}.$$

Then, $\det U = \prod_{1 \leq j < i \leq n} (x_i - x_j)$.

Example 6.57. If $n = 4$, then the matrix U defined in Lemma 6.56 looks as follows:

$$U = \left(\prod_{p=1}^{i-1} (x_j - x_p) \right)_{1 \leq i \leq 4,\ 1 \leq j \leq 4}$$
$$= \begin{pmatrix} \prod_{p=1}^{0} (x_1 - x_p) & \prod_{p=1}^{0} (x_2 - x_p) & \prod_{p=1}^{0} (x_3 - x_p) & \prod_{p=1}^{0} (x_4 - x_p) \\ \prod_{p=1}^{1} (x_1 - x_p) & \prod_{p=1}^{1} (x_2 - x_p) & \prod_{p=1}^{1} (x_3 - x_p) & \prod_{p=1}^{1} (x_4 - x_p) \\ \prod_{p=1}^{2} (x_1 - x_p) & \prod_{p=1}^{2} (x_2 - x_p) & \prod_{p=1}^{2} (x_3 - x_p) & \prod_{p=1}^{2} (x_4 - x_p) \\ \prod_{p=1}^{3} (x_1 - x_p) & \prod_{p=1}^{3} (x_2 - x_p) & \prod_{p=1}^{3} (x_3 - x_p) & \prod_{p=1}^{3} (x_4 - x_p) \end{pmatrix}$$
$$= \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & x_2 - x_1 & x_3 - x_1 & x_4 - x_1 \\ 0 & 0 & (x_3 - x_1)(x_3 - x_2) & (x_4 - x_1)(x_4 - x_2) \\ 0 & 0 & 0 & (x_4 - x_1)(x_4 - x_2)(x_4 - x_3) \end{pmatrix}.$$

(Here, we have used the fact that if $i > j$, then the product $\prod_{p=1}^{i-1} (x_j - x_p)$ contains the factor $x_j - x_j = 0$ and thus equals 0.)

Lemma 6.58. Let $n \in \mathbb{N}$. Let $x_1, x_2, \ldots, x_n$ be n elements of $\mathbb{K}$. Define an $n \times n$-matrix $L \in \mathbb{K}^{n \times n}$ by

$$L = (h_{i-j}(x_1, x_2, \ldots, x_j))_{1 \leq i \leq n,\ 1 \leq j \leq n}.$$

Then, $\det L = 1$.

Example 6.59. If $n = 4$, then the matrix L defined in Lemma 6.58 looks as follows:

$$L = (h_{i-j}(x_1, x_2, \ldots, x_j))_{1 \leq i \leq n,\ 1 \leq j \leq n}$$
$$= \begin{pmatrix} h_0(x_1) & h_{-1}(x_1, x_2) & h_{-2}(x_1, x_2, x_3) & h_{-3}(x_1, x_2, x_3, x_4) \\ h_1(x_1) & h_0(x_1, x_2) & h_{-1}(x_1, x_2, x_3) & h_{-2}(x_1, x_2, x_3, x_4) \\ h_2(x_2) & h_1(x_1, x_2) & h_0(x_1, x_2, x_3) & h_{-1}(x_1, x_2, x_3, x_4) \\ h_3(x_3) & h_2(x_1, x_2) & h_1(x_1, x_2, x_3) & h_0(x_1, x_2, x_3, x_4) \end{pmatrix}$$
$$= \begin{pmatrix} 1 & 0 & 0 & 0 \\ x_1 & 1 & 0 & 0 \\ x_1^2 & x_1 + x_2 & 1 & 0 \\ x_1^3 & x_1^2 + x_1 x_2 + x_2^2 & x_1 + x_2 + x_3 & 1 \end{pmatrix}.$$

(The fact that the diagonal entries are 1 is a consequence of Lemma 6.52 **(b)**, and the fact that the entries above the diagonal are 0 is a consequence of Lemma 6.52 **(a)**.)

Lemma 6.60. Let $n \in \mathbb{N}$. Let $x_1, x_2, \ldots, x_n$ be n elements of $\mathbb{K}$. Let L be the $n \times n$-matrix defined in Lemma 6.58. Let U be the $n \times n$-matrix defined in Lemma 6.56. Then,

$$\left(x_j^{i-1}\right)_{1\leq i\leq n,\ 1\leq j\leq n} = LU.$$

Exercise 6.15. Prove Lemma 6.52, Lemma 6.53, Lemma 6.54, Lemma 6.55, Lemma 6.56, Lemma 6.58 and Lemma 6.60.

Now, we can prove Theorem 6.46 again:

Second proof of Theorem 6.46. **(d)** Let L be the $n \times n$-matrix defined in Lemma 6.58. Let U be the $n \times n$-matrix defined in Lemma 6.56. Then, Lemma 6.60 yields $\left(x_j^{i-1}\right)_{1\leq i\leq n,\ 1\leq j\leq n} = LU$. Hence,

$$\begin{aligned}\det\left(\underbrace{\left(x_j^{i-1}\right)_{1\leq i\leq n,\ 1\leq j\leq n}}_{=LU}\right) &= \det(LU) = \underbrace{\det L}_{\substack{=1 \\ \text{(by Lemma 6.58)}}} \cdot \underbrace{\det U}_{\substack{=\prod\limits_{1\leq j<i\leq n}(x_i-x_j) \\ \text{(by Lemma 6.56)}}} \\ &\qquad \left(\begin{array}{c}\text{by Theorem 6.23, applied to } L \text{ and } U \\ \text{instead of } A \text{ and } B\end{array}\right) \\ &= 1 \cdot \prod_{1\leq j<i\leq n}(x_i - x_j) = \prod_{1\leq j<i\leq n}(x_i - x_j).\end{aligned}$$

This proves Theorem 6.46 **(d)**.

Now, it remains to prove parts **(a)**, **(b)** and **(c)** of Theorem 6.46. This is fairly easy: Back in our First proof of Theorem 6.46, we have derived parts **(b)**, **(d)** and **(c)** from part **(a)**. By essentially the same arguments (sometimes done in reverse), we can derive parts **(a)**, **(b)** and **(c)** from part **(d)**. (We need to use the fact that $\left((-1)^{n(n-1)/2}\right)^2 = 1$.) □

6.7.4. Remarks and variations

Remark 6.61. One consequence of Theorem 6.46 is a new solution to Exercise 5.13 **(a)**:

Namely, let $n \in \mathbb{N}$ and $\sigma \in S_n$. Let $x_1, x_2, \ldots, x_n$ be n elements of $\mathbb{C}$ (or of any commutative ring). Then, Theorem 6.46 **(a)** yields

$$\det\left(\left(x_i^{n-j}\right)_{1\leq i\leq n,\ 1\leq j\leq n}\right) = \prod_{1\leq i<j\leq n}(x_i - x_j).$$

On the other hand, Theorem 6.46 **(a)** (applied to $x_{\sigma(1)}, x_{\sigma(2)}, \ldots, x_{\sigma(n)}$ instead of $x_1, x_2, \ldots, x_n$) yields

$$\det\left(\left(x_{\sigma(i)}^{n-j}\right)_{1\leq i\leq n,\ 1\leq j\leq n}\right) = \prod_{1\leq i<j\leq n}\left(x_{\sigma(i)} - x_{\sigma(j)}\right). \tag{414}$$

But Lemma 6.17 **(a)** (applied to $B = \left(x_i^{n-j}\right)_{1\leq i\leq n,\ 1\leq j\leq n}$, $\kappa = \sigma$ and $B_\kappa = \left(x_{\sigma(i)}^{n-j}\right)_{1\leq i\leq n,\ 1\leq j\leq n}$) yields

$$\det\left(\left(x_{\sigma(i)}^{n-j}\right)_{1\leq i\leq n,\ 1\leq j\leq n}\right) = (-1)^\sigma \cdot \underbrace{\det\left(\left(x_i^{n-j}\right)_{1\leq i\leq n,\ 1\leq j\leq n}\right)}_{=\prod_{1\leq i<j\leq n}(x_i - x_j)}$$
$$= (-1)^\sigma \cdot \prod_{1\leq i<j\leq n}\left(x_i - x_j\right).$$

Compared with (414), this yields

$$\prod_{1\leq i<j\leq n}\left(x_{\sigma(i)} - x_{\sigma(j)}\right) = (-1)^\sigma \cdot \prod_{1\leq i<j\leq n}\left(x_i - x_j\right).$$

Thus, Exercise 5.13 **(a)** is solved. However, Exercise 5.13 **(b)** cannot be solved this way.

Exercise 6.16. Let n be a positive integer. Let $x_1, x_2, \ldots, x_n$ be n elements of $\mathbb{K}$. Prove that

$$\det\left(\left(\begin{cases} x_i^{n-j}, & \text{if } j > 1; \\ x_i^n, & \text{if } j = 1 \end{cases}\right)_{1\leq i\leq n,\ 1\leq j\leq n}\right) = (x_1 + x_2 + \cdots + x_n) \prod_{1\leq i<j\leq n}\left(x_i - x_j\right).$$

(For example, when $n = 4$, this states that

$$\det\begin{pmatrix} x_1^4 & x_1^2 & x_1 & 1 \\ x_2^4 & x_2^2 & x_2 & 1 \\ x_3^4 & x_3^2 & x_3 & 1 \\ x_4^4 & x_4^2 & x_4 & 1 \end{pmatrix}$$
$$= (x_1 + x_2 + x_3 + x_4)(x_1 - x_2)(x_1 - x_3)(x_1 - x_4)(x_2 - x_3)(x_2 - x_4)(x_3 - x_4).$$

)

Remark 6.62. We can try to generalize Vandermonde's determinant. Namely, let $n \in \mathbb{N}$. Let $x_1, x_2, \ldots, x_n$ be n elements of $\mathbb{K}$. Let $a_1, a_2, \ldots, a_n$ be n nonnegative integers. Let A be the $n \times n$-matrix

$$\left(x_i^{a_j}\right)_{1\leq i\leq n,\ 1\leq j\leq n} = \begin{pmatrix} x_1^{a_1} & x_1^{a_2} & \cdots & x_1^{a_n} \\ x_2^{a_1} & x_2^{a_2} & \cdots & x_2^{a_n} \\ \vdots & \vdots & \ddots & \vdots \\ x_n^{a_1} & x_n^{a_2} & \cdots & x_n^{a_n} \end{pmatrix}.$$

What can we say about $\det A$?

Theorem 6.46 says that if $(a_1, a_2, \ldots, a_n) = (n-1, n-2, \ldots, 0)$, then $\det A = \prod_{1\leq i<j\leq n} (x_i - x_j)$.

Exercise 6.16 says that if $n > 0$ and $(a_1, a_2, \ldots, a_n) = (n, n-2, n-3, \ldots, 0)$, then $\det A = (x_1 + x_2 + \cdots + x_n) \prod_{1\leq i<j\leq n} (x_i - x_j)$.

This suggests a general pattern: We would suspect that for every $(a_1, a_2, \ldots, a_n)$, there is a polynomial $P_{(a_1,a_2,\ldots,a_n)}$ in n indeterminates $X_1, X_2, \ldots, X_n$ such that

$$\det A = P_{(a_1,a_2,\ldots,a_n)} (x_1, x_2, \ldots, x_n) \cdot \prod_{1\leq i<j\leq n} (x_i - x_j).$$

It turns out that this is true. Moreover, this polynomial $P_{(a_1,a_2,\ldots,a_n)}$ is:

- zero if two of $a_1, a_2, \ldots, a_n$ are equal;
- homogeneous of degree $a_1 + a_2 + \cdots + a_n - \dbinom{n}{2}$;
- symmetric in $X_1, X_2, \ldots, X_n$.

For example,

$$\begin{aligned} P_{(n-1,n-2,\ldots,0)} &= 1; \\ P_{(n,n-2,n-3,\ldots,0)} &= \sum_{i=1}^{n} X_i = X_1 + X_2 + \cdots + X_n; \\ P_{(n,n-1,\ldots,n-k+1,n-k-1,n-k-2,\ldots,0)} &= \sum_{1\leq i_1<i_2<\cdots<i_k\leq n} X_{i_1} X_{i_2} \cdots X_{i_k} \\ &\qquad \text{for every } k \in \{0, 1, \ldots, n\}; \\ P_{(n+1,n-2,n-3,\ldots,0)} &= \sum_{1\leq i\leq j\leq n} X_i X_j; \\ P_{(n+1,n-1,n-3,n-4,\ldots,0)} &= \sum_{1\leq i<j\leq n} \left(X_i^2 X_j + X_i X_j^2\right) + 2 \sum_{1\leq i<j<k\leq n} X_i X_j X_k. \end{aligned}$$

But this polynomial $P_{(a_1,a_2,...,a_n)}$ can actually be described rather explicitly for general $(a_1, a_2, \ldots, a_n)$; it is a so-called *Schur polynomial* (at least when $a_1 > a_2 > \cdots > a_n$; otherwise it is either zero or ± a Schur polynomial). See [Stembr02, The Bi-Alternant Formula], [Stanle01, Theorem 7.15.1], [Leeuwe06] or [Grinbe21, Theorem 7.3.11] for the details. (Notice that [Leeuwe06] uses the notation $\varepsilon(\sigma)$ for the sign of a permutation σ.) The theory of Schur polynomials shows, in particular, that all coefficients of the polynomial $P_{(a_1,a_2,...,a_n)}$ have equal sign (which is positive if $a_1 > a_2 > \cdots > a_n$).

Remark 6.63. There are plenty other variations on the Vandermonde determinant. For instance, one can try replacing the powers x_i^{j-1} by binomial coefficients $\dbinom{x_i}{j-1}$ in Theorem 6.46 **(c)**, at least when these binomial coefficients are well-defined (e.g., when the $x_1, x_2, \ldots, x_n$ are complex numbers). The result is rather nice: If $x_1, x_2, \ldots, x_n$ are any n complex numbers, then

$$\prod_{1 \leq i < j \leq n} \frac{x_i - x_j}{i - j} = \det \left(\left(\dbinom{x_i}{j-1} \right)_{1 \leq i \leq n,\ 1 \leq j \leq n} \right).$$

(This equality is proved, e.g., in [Grinbe10, Corollary 11] and in [AndDos10, §9, Example 5], and follows easily from Exercise 6.62; it also appeared in [FadSom72, exercise 269].) This equality has the surprising consequence that, whenever $x_1, x_2, \ldots, x_n$ are n integers, the product $\prod\limits_{1 \leq i < j \leq n} \dfrac{x_i - x_j}{i - j}$ is an integer as well (because it is the determinant of a matrix whose entries are integers). This is a nontrivial result! (A more elementary proof appears in [AndDos10, §3, Example 8]. See also [Sury95] for a proof using cyclotomic polynomials, and [Bharga00, Theorem 3] for a placement of this result in a more general context.)

Another "secret integer" (i.e., rational number which turns out to be an integer for non-obvious reasons) is

$$\frac{H(a) H(b) H(c) H(a+b+c)}{H(b+c) H(c+a) H(a+b)}, \tag{415}$$

where a, b, c are three nonnegative integers, and where $H(n)$ (for $n \in \mathbb{N}$) denotes the *hyperfactorial* of n, defined by

$$H(n) = \prod_{k=0}^{n-1} k! = 0! \cdot 1! \cdot \cdots \cdot (n-1)!.$$

I am aware of two proofs of the fact that (415) gives an integer for every $a, b, c \in \mathbb{N}$: One proof is combinatorial, and argues that (415) is the number of *plane partitions inside an $a \times b \times c$-box* (see [Stanle01, last equality in §7.21] for a proof), or, equivalently, the number of *rhombus tilings of a hexagon with sidelengths a, b, c, a, b, c* (see [Eisenk99] for a precise statement). Another proof (see

[Grinbe10, Theorem 0]) exhibits (415) as the determinant of a matrix, again using the Vandermonde determinant!

For some more exercises related to Vandermonde determinants, see [Prasol94, Chapter 1, problems 1.12–1.22]. Here comes one of them ([Kratte05, Lemma 9 and Lemma 10]):

Exercise 6.17. Let n be a positive integer. Let $x_1, x_2, \ldots, x_n$ be n elements of $\mathbb{K}$. Let $y_1, y_2, \ldots, y_n$ be n elements of $\mathbb{K}$.

(a) For every $m \in \{0, 1, \ldots, n-2\}$, prove that

$$\det\left(\left((x_i + y_j)^m\right)_{1\leq i\leq n,\ 1\leq j\leq n}\right) = 0.$$

(b) Prove that

$$\begin{aligned} &\det\left(\left((x_i + y_j)^{n-1}\right)_{1\leq i\leq n,\ 1\leq j\leq n}\right) \\ &= \left(\prod_{k=0}^{n-1}\binom{n-1}{k}\right)\left(\prod_{1\leq i<j\leq n}(x_i - x_j)\right)\left(\prod_{1\leq i<j\leq n}(y_j - y_i)\right). \end{aligned}$$

[Hint: Use the binomial theorem.]

(c) Let $(p_0, p_1, \ldots, p_{n-1}) \in \mathbb{K}^n$ be an n-tuple of elements of $\mathbb{K}$. Let $P(X) \in \mathbb{K}[X]$ be the polynomial $\sum\limits_{k=0}^{n-1} p_k X^k$. Prove that

$$\begin{aligned} &\det\left(\left(P(x_i + y_j)\right)_{1\leq i\leq n,\ 1\leq j\leq n}\right) \\ &= p_{n-1}^n\left(\prod_{k=0}^{n-1}\binom{n-1}{k}\right)\left(\prod_{1\leq i<j\leq n}(x_i - x_j)\right)\left(\prod_{1\leq i<j\leq n}(y_j - y_i)\right). \end{aligned}$$

(d) Let $(p_0, p_1, \ldots, p_{n-1}) \in \mathbb{K}^n$ be an n-tuple of elements of $\mathbb{K}$. Let $P(X) \in \mathbb{K}[X]$ be the polynomial $\sum\limits_{k=0}^{n-1} p_k X^k$. Prove that

$$\begin{aligned} &\det\left(\left(P(x_i y_j)\right)_{1\leq i\leq n,\ 1\leq j\leq n}\right) \\ &= \left(\prod_{k=0}^{n-1} p_k\right)\left(\prod_{1\leq i<j\leq n}(x_i - x_j)\right)\left(\prod_{1\leq i<j\leq n}(y_i - y_j)\right). \end{aligned}$$

Notice how Exercise 6.17 **(a)** generalizes Example 6.7 (for $n \geq 3$).

6.8. Invertible elements in commutative rings, and fields

We shall now interrupt our study of determinants for a moment. Let us define the notion of inverses in $\mathbb{K}$. (Recall that $\mathbb{K}$ is a commutative ring.)

Definition 6.64. Let $a \in \mathbb{K}$. Then, an element $b \in \mathbb{K}$ is said to be an *inverse* of a if it satisfies $ab = 1$ and $ba = 1$.

Of course, the two conditions $ab = 1$ and $ba = 1$ in Definition 6.64 are equivalent, since $ab = ba$ for every $a \in \mathbb{K}$ and $b \in \mathbb{K}$. Nevertheless, we have given both conditions, because this way the similarity between the inverse of an element of $\mathbb{K}$ and the inverse of a map becomes particularly clear.

For example, the element 1 of $\mathbb{Z}$ is its own inverse (since $1 \cdot 1 = 1$), and the element -1 of $\mathbb{Z}$ is its own inverse as well (since $(-1) \cdot (-1) = 1$). These elements 1 and -1 are the only elements of $\mathbb{Z}$ which have an inverse in $\mathbb{Z}$. However, in the larger commutative ring $\mathbb{Q}$, every nonzero element a has an inverse (namely, $\frac{1}{a}$).

Proposition 6.65. Let $a \in \mathbb{K}$. Then, there exists at most one inverse of a in $\mathbb{K}$.

Proof of Proposition 6.65. Let b and b' be any two inverses of a in $\mathbb{K}$. Since b is an inverse of a in $\mathbb{K}$, we have $ab = 1$ and $ba = 1$ (by the definition of an "inverse of a"). Since b' is an inverse of a in $\mathbb{K}$, we have $ab' = 1$ and $b'a = 1$ (by the definition of an "inverse of a"). Now, comparing $b \underbrace{ab'}_{=1} = b$ with $\underbrace{ba}_{=1} b' = b'$, we obtain $b = b'$.

Let us now forget that we fixed b and b'. We thus have shown that if b and b' are two inverses of a in $\mathbb{K}$, then $b = b'$. In other words, any two inverses of a in $\mathbb{K}$ are equal. In other words, there exists at most one inverse of a in $\mathbb{K}$. This proves Proposition 6.65. □

Definition 6.66. (a) An element $a \in \mathbb{K}$ is said to be *invertible* (or, more precisely, *invertible in* $\mathbb{K}$) if and only if there exists an inverse of a in $\mathbb{K}$. In this case, this inverse of a is unique (by Proposition 6.65), and thus will be called *the inverse of* a and denoted by a^{-1}.

(b) It is clear that the unity 1 of $\mathbb{K}$ is invertible (having inverse 1). Also, the product of any two invertible elements a and b of $\mathbb{K}$ is again invertible (having inverse $(ab)^{-1} = a^{-1}b^{-1}$).

(c) If a and b are two elements of $\mathbb{K}$ such that a is invertible (in $\mathbb{K}$), then we write $\frac{b}{a}$ (or b/a) for the product ba^{-1}. These fractions behave just as fractions of integers behave: For example, if a, b, c, d are four elements of $\mathbb{K}$ such that a and c are invertible, then $\frac{b}{a} + \frac{d}{c} = \frac{bc + da}{ac}$ and $\frac{b}{a} \cdot \frac{d}{c} = \frac{bd}{ac}$ (and the product ac is indeed invertible, so that the fractions $\frac{bc + da}{ac}$ and $\frac{bd}{ac}$ actually make sense).

Of course, the meaning of the word "invertible" depends on the ring $\mathbb{K}$. For example, the integer 2 is invertible in $\mathbb{Q}$ (because $\dfrac{1}{2}$ is an inverse of 2 in $\mathbb{Q}$), but not invertible in $\mathbb{Z}$ (since it has no inverse in $\mathbb{Z}$). Thus, it is important to say "invertible in $\mathbb{K}$" unless the context makes it clear what $\mathbb{K}$ is.

One can usually work with invertible elements in commutative rings in the same way as one works with nonzero rational numbers. For example, if a is an invertible element of $\mathbb{K}$, then we can define a^n not only for all $n \in \mathbb{N}$, but also for all $n \in \mathbb{Z}$ (by setting $a^n = \left(a^{-1}\right)^{-n}$ for all negative integers n). Of course, when $n = -1$, this is consistent with our notation a^{-1} for the inverse of a.

Next, we define the notion of a *field*[243].

Definition 6.67. A commutative ring $\mathbb{K}$ is said to be a *field* if it satisfies the following two properties:

- We have $0_{\mathbb{K}} \neq 1_{\mathbb{K}}$ (that is, $\mathbb{K}$ is not a trivial ring).
- Every element of $\mathbb{K}$ is either zero or invertible.

For example, $\mathbb{Q}$, $\mathbb{R}$ and $\mathbb{C}$ are fields, whereas polynomial rings such as $\mathbb{Q}[x]$ or $\mathbb{R}[a, b]$ are not fields[244]. For n being a positive integer, the ring $\mathbb{Z}/n\mathbb{Z}$ (that is, the ring of residue classes of integers modulo n) is a field if and only if n is a prime number.

Linear algebra (i.e., the study of matrices and linear transformations) becomes much easier (in many aspects) when $\mathbb{K}$ is a field[245]. This is one of the main reasons why most courses on linear algebra work over fields only (or begin by working over fields and only later move to the generality of commutative rings). In these

[243] We are going to use the following simple fact: A commutative ring $\mathbb{K}$ is a trivial ring if and only if $0_{\mathbb{K}} = 1_{\mathbb{K}}$.

Proof. Assume that $\mathbb{K}$ is a trivial ring. Thus, $\mathbb{K}$ has only one element. Hence, both $0_{\mathbb{K}}$ and $1_{\mathbb{K}}$ have to equal this one element. Therefore, $0_{\mathbb{K}} = 1_{\mathbb{K}}$.

Now, forget that we assumed that $\mathbb{K}$ is a trivial ring. We thus have proven that

$$\text{if } \mathbb{K} \text{ is a trivial ring, then } 0_{\mathbb{K}} = 1_{\mathbb{K}}. \tag{416}$$

Conversely, assume that $0_{\mathbb{K}} = 1_{\mathbb{K}}$. Then, every $a \in \mathbb{K}$ satisfies $a = a \cdot \underbrace{1_{\mathbb{K}}}_{=0_{\mathbb{K}}} = a \cdot 0_{\mathbb{K}} = 0_{\mathbb{K}} \in \{0_{\mathbb{K}}\}$. In other words, $\mathbb{K} \subseteq \{0_{\mathbb{K}}\}$. Combining this with $\{0_{\mathbb{K}}\} \subseteq \mathbb{K}$, we obtain $\mathbb{K} = \{0_{\mathbb{K}}\}$. Hence, $\mathbb{K}$ has only one element. In other words, $\mathbb{K}$ is a trivial ring.

Now, forget that we assumed that $0_{\mathbb{K}} = 1_{\mathbb{K}}$. We thus have proven that

$$\text{if } 0_{\mathbb{K}} = 1_{\mathbb{K}}, \text{ then } \mathbb{K} \text{ is a trivial ring.}$$

Combining this with (416), we conclude that $\mathbb{K}$ is a trivial ring if and only if $0_{\mathbb{K}} = 1_{\mathbb{K}}$.

[244] For example, the polynomial x is not invertible in $\mathbb{Q}[x]$.

[245] Many properties of a matrix over a field (such as its rank) are not even well-defined over an arbitrary commutative ring.

notes we are almost completely limiting ourselves to the parts of matrix theory which work over any commutative ring. Nevertheless, let us comment on how determinants can be computed fast when $\mathbb{K}$ is a field.

Remark 6.68. Assume that $\mathbb{K}$ is a field. If A is an $n \times n$-matrix over $\mathbb{K}$, then the determinant of A can be computed using (341)... but in practice, you probably do not **want** to compute it this way, since the right hand side of (341) contains a sum of $n!$ terms.

It turns out that there is an algorithm to compute $\det A$, which is (usually) a lot faster. It is a version of the Gaussian elimination algorithm commonly used for solving systems of linear equations.

Let us illustrate it on an example: Set

$$n = 4, \qquad \mathbb{K} = \mathbb{Q} \qquad \text{and } A = \begin{pmatrix} 1 & 2 & 3 & 0 \\ 0 & -1 & 0 & 2 \\ 2 & 4 & -2 & 3 \\ 5 & 1 & 3 & 5 \end{pmatrix}.$$

We want to find $\det A$.

Exercise 6.8 **(b)** shows that if we add a scalar multiple of a column of a matrix to another column of this matrix, then the determinant of the matrix does not change. Now, by adding appropriate scalar multiples of the fourth column of A to the first three columns of A, we can make sure that the first three entries of the fourth row of A become zero: Namely, we have to

- add (-1) times the fourth column of A to the first column of A;
- add $(-1/5)$ times the fourth column of A to the second column of A;
- add $(-3/5)$ times the fourth column of A to the third column of A.

These additions can be performed in any order, since none of them "interacts" with any other (more precisely, none of them uses any entries that another of them changes). As we know, none of these additions changes the determinant of the matrix.

Having performed these three additions, we end up with the matrix

$$A' = \begin{pmatrix} 1 & 2 & 3 & 0 \\ -2 & -7/5 & -6/5 & 2 \\ -1 & 17/5 & -19/5 & 3 \\ 0 & 0 & 0 & 5 \end{pmatrix}. \qquad (417)$$

We have $\det(A') = \det A$ (because A' was obtained from A by three operations which do not change the determinant). Moreover, the fourth row of A' contains only one nonzero entry – namely, its last entry. In other words, if we write A'

in the form $A' = \left(a'_{i,j}\right)_{1\leq i\leq 4,\ 1\leq j\leq 4}$, then $a'_{4,j} = 0$ for every $j \in \{1,2,3\}$. Thus, Theorem 6.43 (applied to 4, A' and $a'_{i,j}$ instead of n, A and $a_{i,j}$) shows that

$$\det\left(A'\right) = \underbrace{a'_{4,4}}_{=5} \cdot \det\left(\underbrace{\left(a'_{i,j}\right)_{1\leq i\leq 3,\ 1\leq j\leq 3}}_{=\begin{pmatrix} 1 & 2 & 3 \\ -2 & -7/5 & -6/5 \\ -1 & 17/5 & -19/5 \end{pmatrix}}\right)$$
$$= 5 \cdot \det\begin{pmatrix} 1 & 2 & 3 \\ -2 & -7/5 & -6/5 \\ -1 & 17/5 & -19/5 \end{pmatrix}.$$

Comparing this with $\det\left(A'\right) = \det A$, we obtain

$$\det A = 5 \cdot \det\begin{pmatrix} 1 & 2 & 3 \\ -2 & -7/5 & -6/5 \\ -1 & 17/5 & -19/5 \end{pmatrix}.$$

Thus, we have reduced the problem of computing $\det A$ (the determinant of a 4×4-matrix) to the problem of computing $\det\begin{pmatrix} 1 & 2 & 3 \\ -2 & -7/5 & -6/5 \\ -1 & 17/5 & -19/5 \end{pmatrix}$ (the determinant of a 3×3-matrix). Likewise, we can try to reduce the latter problem to the computation of the determinant of a 2×2-matrix, and then further to the computation of the determinant of a 1×1-matrix. (In our example, we obtain $\det A = -140$ at the end.)

This looks like a viable algorithm (which is, furthermore, fairly fast: essentially as fast as Gaussian elimination). But does it always work? It turns out that it **almost** always works. There are cases in which it can get "stuck", and it needs to be modified to deal with these cases.

Namely, what can happen is that the (n,n)-th entry of the matrix A could be 0. Again, let us observe this on an example: Set $n = 4$ and $A = \begin{pmatrix} 1 & 2 & 3 & 0 \\ 0 & -1 & 0 & 2 \\ 2 & 4 & -2 & 3 \\ 5 & 1 & 3 & 0 \end{pmatrix}$. Then, we cannot turn the first three entries of the fourth row of A into zeroes by adding appropriate multiples of the fourth column to the first three columns. (Whatever multiples we add, the fourth row stays unchanged.) However, we can now swap the second row of A with the fourth row. This operation produces the matrix $B = \begin{pmatrix} 1 & 2 & 3 & 0 \\ 5 & 1 & 3 & 0 \\ 2 & 4 & -2 & 3 \\ 0 & -1 & 0 & 2 \end{pmatrix}$, which satisfies $\det B = -\det A$ (by Exercise 6.7 **(a)**). Thus, it suffices to compute $\det B$; and this can be done as above.

The reason why we swapped the second row of A with the fourth row is that the last entry of the second row of A was nonzero. In general, we need to find a $k \in \{1, 2, \ldots, n\}$ such that the last entry of the k-th row of A is nonzero, and swap the k-th row of A with the n-th row. But what if no such k exists? In this case, we need another way to compute $\det A$. It turns out that this is very easy: If there is no $k \in \{1, 2, \ldots, n\}$ such that the last entry of the k-th row of A is nonzero, then the last column of A consists of zeroes, and thus Exercise 6.7 **(d)** shows that $\det A = 0$.

When $\mathbb{K}$ is not a field, this algorithm breaks (or, at least, **can** break). Indeed, it relies on the fact that the (n, n)-th entry of the matrix A is either zero or invertible. Over a commutative ring $\mathbb{K}$, it might be neither. For example, if we had tried to work with $\mathbb{K} = \mathbb{Z}$ (instead of $\mathbb{K} = \mathbb{Q}$) in our above example, then we would not be able to add $(-1/5)$ times the fourth column of A to the second column of A (because $-1/5 \notin \mathbb{Z} = \mathbb{K}$). Fortunately, of course, $\mathbb{Z}$ is a subset of $\mathbb{Q}$ (and its operations $+$ and $\cdot$ are consistent with those of $\mathbb{Q}$), so that we can just perform the whole algorithm over $\mathbb{Q}$ instead of $\mathbb{Z}$. However, we aren't always in luck: Some commutative rings $\mathbb{K}$ cannot be "embedded" into fields in the way $\mathbb{Z}$ is embedded into $\mathbb{Q}$. (For instance, $\mathbb{Z}/4\mathbb{Z}$ cannot be embedded into a field.)

Nevertheless, there **are** reasonably fast algorithms for computing determinants over any commutative ring; see [Rote01, §2].

6.9. The Cauchy determinant

Now, we can state another classical formula for a determinant: the *Cauchy determinant*. In one of its many forms, it says the following:

Exercise 6.18. Let $n \in \mathbb{N}$. Let $x_1, x_2, \ldots, x_n$ be n elements of $\mathbb{K}$. Let $y_1, y_2, \ldots, y_n$ be n elements of $\mathbb{K}$. Assume that $x_i + y_j$ is invertible in $\mathbb{K}$ for every $(i, j) \in \{1, 2, \ldots, n\}^2$. Then, prove that

$$\det\left(\left(\frac{1}{x_i + y_j}\right)_{1 \leq i \leq n,\ 1 \leq j \leq n}\right) = \frac{\prod\limits_{1 \leq i < j \leq n} \left(\left(x_i - x_j\right)\left(y_i - y_j\right)\right)}{\prod\limits_{(i,j) \in \{1,2,\ldots,n\}^2} \left(x_i + y_j\right)}.$$

There is a different version of the Cauchy determinant floating around in literature; it differs from Exercise 6.18 in that each "$x_i + y_j$" is replaced by "$x_i - y_j$", and in that "$y_i - y_j$" is replaced by "$y_j - y_i$". Of course, this version is nothing else than the result of applying Exercise 6.18 to $-y_1, -y_2, \ldots, -y_n$ instead of $y_1, y_2, \ldots, y_n$.

Exercise 6.19. Let n be a positive integer. Let $\left(a_{i,j}\right)_{1\leq i\leq n,\ 1\leq j\leq n}$ be an $n\times n$-matrix such that $a_{n,n}$ is invertible (in $\mathbb{K}$). Prove that

$$\begin{aligned}&\det\left(\left(a_{i,j}a_{n,n}-a_{i,n}a_{n,j}\right)_{1\leq i\leq n-1,\ 1\leq j\leq n-1}\right)\\&=a_{n,n}^{n-2}\cdot\det\left(\left(a_{i,j}\right)_{1\leq i\leq n,\ 1\leq j\leq n}\right).\end{aligned}\tag{418}$$

Exercise 6.19 is known as the *Chio pivotal condensation theorem*[246].

Remark 6.69. Exercise 6.19 gives a way to reduce the computation of an $n\times n$-determinant (the one on the right hand side of (418)) to the computation of an $(n-1)\times(n-1)$-determinant (the one on the left hand side), provided that $a_{n,n}$ is invertible. If this reminds you of Remark 6.68, you are thinking right...

Remark 6.70. Exercise 6.19 holds even without the assumption that $a_{n,n}$ be invertible, as long as we assume (instead) that $n\geq 2$. (If we don't assume that $n\geq 2$, then the $a_{n,n}^{n-2}$ on the right hand side of (418) will not be defined for non-invertible $a_{n,n}$.) Proving this is beyond these notes, though. (A proof of this generalized version of Exercise 6.19 can be found in [KarZha16]. It can also be obtained as a particular case of [BerBru08, (4)][247].)

6.10. Further determinant equalities

Next, let us provide an assortment of other exercises on determinants. Hundreds of exercises (ranging from easy to challenging) on the properties and evaluations of determinants can be found in [FadSom72, Chapter 2], and some more in [Prasol94, Chapter I]; in comparison, our selection is rather small.

Exercise 6.20. Let $n\in\mathbb{N}$. Let $\left(a_{i,j}\right)_{1\leq i\leq n,\ 1\leq j\leq n}$ be an $n\times n$-matrix. Let $b_1, b_2, \ldots, b_n$ be n elements of $\mathbb{K}$. Prove that

$$\sum_{k=1}^{n}\det\left(\left(a_{i,j}b_i^{\delta_{j,k}}\right)_{1\leq i\leq n,\ 1\leq j\leq n}\right)=\left(b_1+b_2+\cdots+b_n\right)\det\left(\left(a_{i,j}\right)_{1\leq i\leq n,\ 1\leq j\leq n}\right),$$

where $\delta_{j,k}$ means the nonnegative integer $\begin{cases}1, & \text{if } j=k;\\ 0, & \text{if } j\neq k\end{cases}$. Equivalently (in more

[246] See [Heinig11, footnote 2] and [Abeles14, §2] for some hints about its history. A variant of the formula (singling out the 1-st row and the 1-st column instead of the n-th row and the n-th column) appears in [Heffer20, Chapter Four, Topic "Chio's Method"].

[247] In more detail: If we apply [BerBru08, (4)] to $k=n-1$, then the right hand side is precisely $\det\left(\left(a_{i,j}a_{n,n}-a_{i,n}a_{n,j}\right)_{1\leq i\leq n-1,\ 1\leq j\leq n-1}\right)$, and so the formula becomes (418).

reader-friendly terms): Prove that

$$\begin{aligned}
&\det \begin{pmatrix} a_{1,1}b_1 & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1}b_2 & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n,1}b_n & a_{n,2} & \cdots & a_{n,n} \end{pmatrix} + \det \begin{pmatrix} a_{1,1} & a_{1,2}b_1 & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2}b_2 & \cdots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n,1} & a_{n,2}b_n & \cdots & a_{n,n} \end{pmatrix} \\
&\qquad + \cdots + \det \begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n}b_1 \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n}b_2 \\ \vdots & \vdots & \ddots & \vdots \\ a_{n,1} & a_{n,2} & \cdots & a_{n,n}b_n \end{pmatrix} \\
&= (b_1 + b_2 + \cdots + b_n) \det \begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n,1} & a_{n,2} & \cdots & a_{n,n} \end{pmatrix}.
\end{aligned}$$

Exercise 6.21. Let $n \in \mathbb{N}$. Let $a_1, a_2, \ldots, a_n$ be n elements of $\mathbb{K}$. Let $x \in \mathbb{K}$. Prove that

$$\det \begin{pmatrix} x & a_1 & a_2 & \cdots & a_{n-1} & a_n \\ a_1 & x & a_2 & \cdots & a_{n-1} & a_n \\ a_1 & a_2 & x & \cdots & a_{n-1} & a_n \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ a_1 & a_2 & a_3 & \cdots & x & a_n \\ a_1 & a_2 & a_3 & \cdots & a_n & x \end{pmatrix} = \left(x + \sum_{i=1}^{n} a_i \right) \prod_{i=1}^{n} (x - a_i).$$

Exercise 6.22. Let $n > 1$ be an integer. Let $a_1, a_2, \ldots, a_n$ be n elements of $\mathbb{K}$. Let $b_1, b_2, \ldots, b_n$ be n elements of $\mathbb{K}$. Let A be the $n \times n$-matrix

$$\begin{aligned}
&\left(\begin{cases} a_j, & \text{if } i = j; \\ b_j, & \text{if } i \equiv j + 1 \bmod n; \\ 0, & \text{otherwise} \end{cases} \right)_{1 \leq i \leq n,\ 1 \leq j \leq n} \\
&= \begin{pmatrix} a_1 & 0 & 0 & \cdots & 0 & b_n \\ b_1 & a_2 & 0 & \cdots & 0 & 0 \\ 0 & b_2 & a_3 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & a_{n-1} & 0 \\ 0 & 0 & 0 & \cdots & b_{n-1} & a_n \end{pmatrix}.
\end{aligned}$$

Prove that

$$\det A = a_1 a_2 \cdots a_n + (-1)^{n-1} b_1 b_2 \cdots b_n.$$

Remark 6.71. If we replace "$i \equiv j + 1 \mod n$" by "$i \equiv j + 2 \mod n$" in Exercise 6.22, then the pattern can break. For instance, for $n = 4$ we have

$$\det \begin{pmatrix} a_1 & 0 & b_3 & 0 \\ 0 & a_2 & 0 & b_4 \\ b_1 & 0 & a_3 & 0 \\ 0 & b_2 & 0 & a_4 \end{pmatrix} = (a_2a_4 - b_2b_4)(a_1a_3 - b_1b_3),$$

which is not of the form $a_1a_2a_3a_4 \pm b_1b_2b_3b_4$ anymore. Can you guess for which $d \in \{1, 2, \ldots, n-1\}$ we can replace "$i \equiv j + 1 \mod n$" by "$i \equiv j + d \mod n$" in Exercise 6.22 and still get a formula of the form $\det A = a_1a_2 \cdots a_n \pm b_1b_2 \cdots b_n$? (The answer to this question requires a little bit of elementary number theory – namely, the concept of "coprimality". See [HanKra00, Proposition 6] for the answer.)

6.11. Alternating matrices

Our next two exercises will concern two special classes of matrices: the *antisymmetric* and the *alternating matrices*. Let us first define these classes:

Definition 6.72. Let $n \in \mathbb{N}$. Let $A = (a_{i,j})_{1 \leq i \leq n,\ 1 \leq j \leq n}$ be an $n \times n$-matrix.

(a) The matrix A is said to be *antisymmetric* if and only if $A^T = -A$. (Recall that A^T is defined as in Definition 6.10.)

(b) The matrix A is said to be *alternating* if and only if it satisfies $A^T = -A$ and ($a_{i,i} = 0$ for all $i \in \{1, 2, \ldots, n\}$).

Example 6.73. A 1×1-matrix is alternating if and only if it is the zero matrix $0_{1\times 1} = \begin{pmatrix} 0 \end{pmatrix}$.

A 2×2-matrix is alternating if and only if it has the form $\begin{pmatrix} 0 & a \\ -a & 0 \end{pmatrix}$ for some $a \in \mathbb{K}$.

A 3×3-matrix is alternating if and only if it has the form $\begin{pmatrix} 0 & a & b \\ -a & 0 & c \\ -b & -c & 0 \end{pmatrix}$ for some $a, b, c \in \mathbb{K}$.

Visually speaking, an $n \times n$-matrix is alternating if and only if its diagonal entries are 0 and its entries below the diagonal are the negatives of their "mirror-image" entries above the diagonal.

Remark 6.74. Clearly, any alternating matrix is antisymmetric. It is easy to see that an $n \times n$-matrix $A = (a_{i,j})_{1 \leq i \leq n,\ 1 \leq j \leq n}$ is antisymmetric if and only if every $(i, j) \in \{1, 2, \ldots, n\}^2$ satisfies $a_{i,j} = -a_{j,i}$. Thus, if $A = (a_{i,j})_{1 \leq i \leq n,\ 1 \leq j \leq n}$ is antisymmetric, then every $i \in \{1, 2, \ldots, n\}$ satisfies $a_{i,i} = -a_{i,i}$ and thus $2a_{i,i} = 0$. If

$\mathbb{K}$ is one of the rings $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$ and $\mathbb{C}$, then we can cancel 2 from this last equality, and conclude that every antisymmetric $n \times n$-matrix A is alternating. However, there are commutative rings $\mathbb{K}$ for which this does not hold (for example, the ring $\mathbb{Z}/2\mathbb{Z}$ of integers modulo 2).

Antisymmetric matrices are also known as *skew-symmetric* matrices.

Exercise 6.23. Let $n \in \mathbb{N}$ and $m \in \mathbb{N}$. Let A be an alternating $n \times n$-matrix. Let S be an $n \times m$-matrix. Prove that the $m \times m$-matrix $S^T AS$ is alternating.

Exercise 6.24. Let $n \in \mathbb{N}$ be odd. Let A be an $n \times n$-matrix. Prove the following:

(a) If A is antisymmetric, then $2 \det A = 0$.

(b) If A is alternating, then $\det A = 0$.

Remark 6.75. If $\mathbb{K}$ is one of the rings $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$ and $\mathbb{C}$, then Exercise 6.24 **(b)** follows from Exercise 6.24 **(a)** (because any alternating matrix is antisymmetric, and because we can cancel 2 from the equality $2 \det A = 0$). However, this quick way of solving Exercise 6.24 **(b)** does not work for general $\mathbb{K}$.

Remark 6.76. Exercise 6.24 **(b)** provides a really simple formula for $\det A$ when A is an alternating $n \times n$-matrix for **odd** n. One might wonder what can be said about $\det A$ when A is an alternating $n \times n$-matrix for **even** n. The answer is far less simple, but more interesting: It turns that $\det A$ is the square of a certain element of $\mathbb{K}$, called the *Pfaffian* of A. See [Conrad2, (5.5)] for a short introduction into the Pfaffian (although at a less elementary level than these notes); see [BruRys91, §9.5] and [Loehr11, §12.12][248] for a more combinatorial treatment of the Pfaffian (and an application to matchings of graphs!). For example, the Pfaffian of an alternating 4×4-matrix $A = \begin{pmatrix} 0 & a & b & c \\ -a & 0 & d & e \\ -b & -d & 0 & f \\ -c & -e & -f & 0 \end{pmatrix}$ is $af - be + cd$, and it is indeed easy to check that this matrix satisfies $\det A = (af - be + cd)^2$.

6.12. Laplace expansion

We shall now state Laplace expansion in full. We begin with an example:

[248] Beware that Loehr, in [Loehr11, §12.12], seems to work only in the setting where 2 is cancellable in the ring $\mathbb{K}$ (that is, where $2a = 0$ for an element $a \in \mathbb{K}$ implies $a = 0$). Thus, Loehr does not have to distinguish between antisymmetric and alternating matrices (he calls them "skew-symmetric matrices" instead). His arguments, however, can easily be adapted to the general case.

Example 6.77. Let $A = \left(a_{i,j}\right)_{1\leq i\leq 3,\ 1\leq j\leq 3}$ be a 3×3-matrix. From (344), we obtain

$$\det A = a_{1,1}a_{2,2}a_{3,3} + a_{1,2}a_{2,3}a_{3,1} + a_{1,3}a_{2,1}a_{3,2} - a_{1,1}a_{2,3}a_{3,2} - a_{1,2}a_{2,1}a_{3,3} - a_{1,3}a_{2,2}a_{3,1}. \tag{419}$$

On the right hand side of this equality, we have six terms, each of which contains either $a_{2,1}$ or $a_{2,2}$ or $a_{2,3}$. Let us combine the two terms containing $a_{2,1}$ and factor out $a_{2,1}$, then do the same with the two terms containing $a_{2,2}$, and with the two terms containing $a_{2,3}$. As a result, (419) becomes

$$\begin{aligned}
&\det A \\
&= a_{1,1}a_{2,2}a_{3,3} + a_{1,2}a_{2,3}a_{3,1} + a_{1,3}a_{2,1}a_{3,2} - a_{1,1}a_{2,3}a_{3,2} - a_{1,2}a_{2,1}a_{3,3} - a_{1,3}a_{2,2}a_{3,1} \\
&= a_{2,1}\underbrace{\left(a_{1,3}a_{3,2} - a_{1,2}a_{3,3}\right)}_{=\det\begin{pmatrix} a_{1,3} & a_{1,2} \\ a_{3,3} & a_{3,2} \end{pmatrix}} + a_{2,2}\underbrace{\left(a_{1,1}a_{3,3} - a_{1,3}a_{3,1}\right)}_{=\det\begin{pmatrix} a_{1,1} & a_{1,3} \\ a_{3,1} & a_{3,3} \end{pmatrix}} + a_{2,3}\underbrace{\left(a_{1,2}a_{3,1} - a_{1,1}a_{3,2}\right)}_{=\det\begin{pmatrix} a_{1,2} & a_{1,1} \\ a_{3,2} & a_{3,1} \end{pmatrix}} \\
&= a_{2,1}\det\begin{pmatrix} a_{1,3} & a_{1,2} \\ a_{3,3} & a_{3,2} \end{pmatrix} + a_{2,2}\det\begin{pmatrix} a_{1,1} & a_{1,3} \\ a_{3,1} & a_{3,3} \end{pmatrix} + a_{2,3}\det\begin{pmatrix} a_{1,2} & a_{1,1} \\ a_{3,2} & a_{3,1} \end{pmatrix}.
\end{aligned} \tag{420}$$

This is a nice formula with an obvious pattern: The right hand side can be rewritten as $\sum\limits_{q=1}^{3} a_{2,q}\det\left(B_{2,q}\right)$, where $B_{2,q} = \begin{pmatrix} a_{1,q+2} & a_{1,q+1} \\ a_{3,q+2} & a_{3,q+1} \end{pmatrix}$ (where we set $a_{i,4} = a_{i,1}$ and $a_{i,5} = a_{i,2}$ for all $i \in \{1,2,3\}$). Notice the cyclic symmetry (with respect to the index of the column) in this formula! Unfortunately, in this exact form, the formula does not generalize to bigger matrices (or even to smaller: the analogue for a 2×2-matrix would be $\det\begin{pmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{pmatrix} = -a_{2,1}a_{1,2} + a_{2,2}a_{1,1}$, which has a minus sign unlike $\sum\limits_{q=1}^{3} a_{2,q}\det\left(B_{2,q}\right)$).

However, we can slightly modify our formula, sacrificing the cyclic symmetry but making it generalize. Namely, let us rewrite $a_{1,3}a_{3,2} - a_{1,2}a_{3,3}$ as

$-\left(a_{1,2}a_{3,3}-a_{1,3}a_{3,2}\right)$ and $a_{1,2}a_{3,1}-a_{1,1}a_{3,2}$ as $-\left(a_{1,1}a_{3,2}-a_{1,2}a_{3,1}\right)$; we thus obtain

$$\begin{aligned}
&\det A\\
&=a_{2,1}\underbrace{\left(a_{1,3}a_{3,2}-a_{1,2}a_{3,3}\right)}_{=-\left(a_{1,2}a_{3,3}-a_{1,3}a_{3,2}\right)}+a_{2,2}\left(a_{1,1}a_{3,3}-a_{1,3}a_{3,1}\right)+a_{2,3}\underbrace{\left(a_{1,2}a_{3,1}-a_{1,1}a_{3,2}\right)}_{=-\left(a_{1,1}a_{3,2}-a_{1,2}a_{3,1}\right)}\\
&=-a_{2,1}\underbrace{\left(a_{1,2}a_{3,3}-a_{1,3}a_{3,2}\right)}_{=\det\begin{pmatrix}a_{1,2}&a_{1,3}\\a_{3,2}&a_{3,3}\end{pmatrix}}+a_{2,2}\underbrace{\left(a_{1,1}a_{3,3}-a_{1,3}a_{3,1}\right)}_{=\det\begin{pmatrix}a_{1,1}&a_{1,3}\\a_{3,1}&a_{3,3}\end{pmatrix}}-a_{2,3}\underbrace{\left(a_{1,1}a_{3,2}-a_{1,2}a_{3,1}\right)}_{=\det\begin{pmatrix}a_{1,1}&a_{1,2}\\a_{3,1}&a_{3,2}\end{pmatrix}}\\
&=-a_{2,1}\det\begin{pmatrix}a_{1,2}&a_{1,3}\\a_{3,2}&a_{3,3}\end{pmatrix}+a_{2,2}\det\begin{pmatrix}a_{1,1}&a_{1,3}\\a_{3,1}&a_{3,3}\end{pmatrix}-a_{2,3}\det\begin{pmatrix}a_{1,1}&a_{1,2}\\a_{3,1}&a_{3,2}\end{pmatrix}\\
&=\sum_{q=1}^{3}(-1)^{q}a_{2,q}\det\left(C_{2,q}\right), \qquad (421)
\end{aligned}$$

where $C_{2,q}$ means the matrix obtained from A by crossing out the 2-nd row and the q-th column. This formula (unlike (420)) involves powers of -1, but it can be generalized.

How? First, we notice that we can find a similar formula by factoring out $a_{1,1}, a_{1,2}, a_{1,3}$ (instead of $a_{2,1}, a_{2,2}, a_{2,3}$); this formula will be

$$\det A=\sum_{q=1}^{3}(-1)^{q-1}a_{1,q}\det\left(C_{1,q}\right),$$

where $C_{1,q}$ means the matrix obtained from A by crossing out the 1-st row and the q-th column. This formula, and (421), suggest the following generalization: If $A=\left(a_{i,j}\right)_{1\leq i\leq n,\ 1\leq j\leq n}$ is an $n\times n$-matrix, and if $p\in\{1,2,\ldots,n\}$, then

$$\det A=\sum_{q=1}^{n}(-1)^{p+q}a_{p,q}\det\left(C_{p,q}\right), \qquad (422)$$

where $C_{p,q}$ means the matrix obtained from A by crossing out the p-th row and the q-th column. (The only part of this formula which is not easy to guess is $(-1)^{p+q}$; you might need to compute several particular cases to guess this pattern. Of course, you could also have guessed $(-1)^{p-q}$ or $(-1)^{q-p}$ instead, because $(-1)^{p+q}=(-1)^{p-q}=(-1)^{q-p}$.)

The formula (422) is what is usually called the Laplace expansion with respect to the p-th row. We will prove it below (Theorem 6.82 **(a)**), and we will also prove an analogous "Laplace expansion with respect to the q-th column" (Theorem 6.82 **(b)**).

Let us first define a notation:

Definition 6.78. Let $n \in \mathbb{N}$ and $m \in \mathbb{N}$. Let $A = \left(a_{i,j}\right)_{1 \leq i \leq n,\ 1 \leq j \leq m}$ be an $n \times m$-matrix. Let $i_1, i_2, \ldots, i_u$ be some elements of $\{1, 2, \ldots, n\}$; let $j_1, j_2, \ldots, j_v$ be some elements of $\{1, 2, \ldots, m\}$. Then, we define $\operatorname{sub}_{i_1,i_2,\ldots,i_u}^{j_1,j_2,\ldots,j_v} A$ to be the $u \times v$-matrix $\left(a_{i_x, j_y}\right)_{1 \leq x \leq u,\ 1 \leq y \leq v}$.

When $i_1 < i_2 < \cdots < i_u$ and $j_1 < j_2 < \cdots < j_v$, the matrix $\operatorname{sub}_{i_1,i_2,\ldots,i_u}^{j_1,j_2,\ldots,j_v} A$ can be obtained from A by crossing out all rows other than the i_1-th, the i_2-th, etc., the i_u-th row and crossing out all columns other than the j_1-th, the j_2-th, etc., the j_v-th column. Thus, in this case, $\operatorname{sub}_{i_1,i_2,\ldots,i_u}^{j_1,j_2,\ldots,j_v} A$ is called a *submatrix* of A.

For example, if $n = 3$, $m = 4$ and $A = \begin{pmatrix} a & b & c & d \\ e & f & g & h \\ i & j & k & \ell \end{pmatrix}$, then $\operatorname{sub}_{1,3}^{2,3,4} A = \begin{pmatrix} b & c & d \\ j & k & \ell \end{pmatrix}$ (this is a submatrix of A) and $\operatorname{sub}_{2,3}^{3,1,1} A = \begin{pmatrix} g & e & e \\ k & i & i \end{pmatrix}$ (this is not, in general, a submatrix of A).

The following properties follow trivially from the definitions:

Proposition 6.79. Let $n \in \mathbb{N}$ and $m \in \mathbb{N}$. Let A be an $n \times m$-matrix. Recall the notations introduced in Definition 6.31.

(a) We have $\operatorname{sub}_{1,2,\ldots,n}^{1,2,\ldots,m} A = A$.

(b) If $i_1, i_2, \ldots, i_u$ are some elements of $\{1, 2, \ldots, n\}$, then

$$\operatorname{rows}_{i_1,i_2,\ldots,i_u} A = \operatorname{sub}_{i_1,i_2,\ldots,i_u}^{1,2,\ldots,m} A.$$

(c) If $j_1, j_2, \ldots, j_v$ are some elements of $\{1, 2, \ldots, m\}$, then

$$\operatorname{cols}_{j_1,j_2,\ldots,j_v} A = \operatorname{sub}_{1,2,\ldots,n}^{j_1,j_2,\ldots,j_v} A.$$

(d) Let $i_1, i_2, \ldots, i_u$ be some elements of $\{1, 2, \ldots, n\}$; let $j_1, j_2, \ldots, j_v$ be some elements of $\{1, 2, \ldots, m\}$. Then,

$$\operatorname{sub}_{i_1,i_2,\ldots,i_u}^{j_1,j_2,\ldots,j_v} A = \operatorname{rows}_{i_1,i_2,\ldots,i_u}\left(\operatorname{cols}_{j_1,j_2,\ldots,j_v} A\right) = \operatorname{cols}_{j_1,j_2,\ldots,j_v}\left(\operatorname{rows}_{i_1,i_2,\ldots,i_u} A\right).$$

(e) Let $i_1, i_2, \ldots, i_u$ be some elements of $\{1, 2, \ldots, n\}$; let $j_1, j_2, \ldots, j_v$ be some elements of $\{1, 2, \ldots, m\}$. Then,

$$\left(\operatorname{sub}_{i_1,i_2,\ldots,i_u}^{j_1,j_2,\ldots,j_v} A\right)^T = \operatorname{sub}_{j_1,j_2,\ldots,j_v}^{i_1,i_2,\ldots,i_u}\left(A^T\right).$$

Definition 6.80. Let $n \in \mathbb{N}$. Let $a_1, a_2, \ldots, a_n$ be n objects. Let $i \in \{1, 2, \ldots, n\}$. Then, $(a_1, a_2, \ldots, \widehat{a_i}, \ldots, a_n)$ shall mean the list $(a_1, a_2, \ldots, a_{i-1}, a_{i+1}, a_{i+2}, \ldots, a_n)$

(that is, the list $(a_1, a_2, \ldots, a_n)$ with its i-th entry removed). (Thus, the "hat" over the a_i means that this a_i is being omitted from the list.)

For example, $\left(1^2, 2^2, \ldots, \widehat{5^2}, \ldots, 8^2\right) = \left(1^2, 2^2, 3^2, 4^2, 6^2, 7^2, 8^2\right)$.

Definition 6.81. Let $n \in \mathbb{N}$ and $m \in \mathbb{N}$. Let A be an $n \times m$-matrix. For every $i \in \{1, 2, \ldots, n\}$ and $j \in \{1, 2, \ldots, m\}$, we let $A_{\sim i, \sim j}$ be the $(n-1) \times (m-1)$-matrix $\operatorname{sub}^{1,2,\ldots,\widehat{j},\ldots,m}_{1,2,\ldots,\widehat{i},\ldots,n} A$. (Thus, $A_{\sim i, \sim j}$ is the matrix obtained from A by crossing out the i-th row and the j-th column.)

For example, if $n = m = 3$ and $A = \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix}$, then $A_{\sim 1, \sim 2} = \begin{pmatrix} d & f \\ g & i \end{pmatrix}$ and $A_{\sim 3, \sim 2} = \begin{pmatrix} a & c \\ d & f \end{pmatrix}$.

The notation $A_{\sim i, \sim j}$ introduced in Definition 6.81 is not very standard; but there does not seem to be a standard one[249].

Now we can finally state Laplace expansion:

Theorem 6.82. Let $n \in \mathbb{N}$. Let $A = \left(a_{i,j}\right)_{1 \leq i \leq n,\ 1 \leq j \leq n}$ be an $n \times n$-matrix.

(a) For every $p \in \{1, 2, \ldots, n\}$, we have

$$\det A = \sum_{q=1}^{n} (-1)^{p+q} a_{p,q} \det\left(A_{\sim p, \sim q}\right).$$

(b) For every $q \in \{1, 2, \ldots, n\}$, we have

$$\det A = \sum_{p=1}^{n} (-1)^{p+q} a_{p,q} \det\left(A_{\sim p, \sim q}\right).$$

Theorem 6.82 **(a)** is known as the *Laplace expansion along the p-th row* (or *Laplace expansion with respect to the p-th row*), whereas Theorem 6.82 **(b)** is known as the *Laplace expansion along the q-th column* (or *Laplace expansion with respect to the q-th column*). Notice that Theorem 6.82 **(a)** is equivalent to the formula (422), because the $A_{\sim p, \sim q}$ in Theorem 6.82 **(a)** is precisely what we called $C_{p,q}$ in (422).

We prepare the field for the proof of Theorem 6.82 with a few lemmas.

Lemma 6.83. For every $n \in \mathbb{N}$, let $[n]$ denote the set $\{1, 2, \ldots, n\}$.

Let $n \in \mathbb{N}$. For every $p \in [n]$, we define a permutation $g_p \in S_n$ by $g_p = \operatorname{cyc}_{p,p+1,\ldots,n}$ (where we are using the notations of Definition 5.37).

[249] For example, Gill Williamson uses the notation $A(i \mid j)$ in [Willia18, Chapter 3].

(a) We have $\left(g_p(1), g_p(2), \ldots, g_p(n-1)\right) = (1, 2, \ldots, \widehat{p}, \ldots, n)$ for every $p \in [n]$.

(b) We have $(-1)^{g_p} = (-1)^{n-p}$ for every $p \in [n]$.

(c) Let $p \in [n]$. We define a map

$$g'_p : [n-1] \to [n] \setminus \{p\}$$

by

$$\left(g'_p(i) = g_p(i) \qquad \text{for every } i \in [n-1]\right).$$

This map g'_p is well-defined and bijective.

(d) Let $p \in [n]$ and $q \in [n]$. We define a map

$$T : \{\tau \in S_n \mid \tau(n) = n\} \to \{\tau \in S_n \mid \tau(p) = q\}$$

by

$$\left(T(\sigma) = g_q \circ \sigma \circ \left(g_p\right)^{-1} \qquad \text{for every } \sigma \in \{\tau \in S_n \mid \tau(n) = n\}\right).$$

Then, this map T is well-defined and bijective.

Proof of Lemma 6.83. **(a)** This is trivial.

(b) Let $p \in [n]$. Exercise 5.17 **(d)** (applied to $k = p+1$ and $(i_1, i_2, \ldots, i_k) = (p, p+1, \ldots, n)$) yields

$$(-1)^{\operatorname{cyc}_{p,p+1,\ldots,n}} = (-1)^{n-(p+1)-1} = (-1)^{n-p-2} = (-1)^{n-p}.$$

Now, $g_p = \operatorname{cyc}_{p,p+1,\ldots,n}$, so that $(-1)^{g_p} = (-1)^{\operatorname{cyc}_{p,p+1,\ldots,n}} = (-1)^{n-p}$. This proves Lemma 6.83 **(b)**.

(c) We have $g_p(n) = p$ (since $g_p = \operatorname{cyc}_{p,p+1,\ldots,n}$). Also, g_p is injective (since g_p is a permutation). Therefore, for every $i \in [n-1]$, we have

$$g_p(i) \neq g_p(n) = p, \qquad \left(\text{since } i \neq n \text{ (because } i \in [n-1]\text{) and since } g_p \text{ is injective}\right)$$

so that $g_p(i) \in [n] \setminus \{p\}$. This shows that the map g'_p is well-defined.

To prove that g'_p is bijective, we can construct its inverse. Indeed, for every $i \in [n] \setminus \{p\}$, we have

$$\left(g_p\right)^{-1}(i) \neq n \qquad \left(\text{since } i \neq p = g_p(n)\right)$$

and thus $\left(g_p\right)^{-1}(i) \in [n-1]$. Hence, we can define a map $h : [n] \setminus \{p\} \to [n-1]$ by

$$\left(h(i) = \left(g_p\right)^{-1}(i) \qquad \text{for every } i \in [n] \setminus \{p\}\right).$$

It is straightforward to check that the maps g'_p and h are mutually inverse. Thus, g'_p is bijective. Lemma 6.83 **(c)** is thus proven.

(d) We have $g_p(n) = p$ (since $g_p = \operatorname{cyc}_{p,p+1,\ldots,n}$) and $g_q(n) = q$ (similarly). Hence, $(g_p)^{-1}(p) = n$ (since $g_p(n) = p$) and $(g_q)^{-1}(q) = n$ (since $g_q(n) = q$).

For every $\sigma \in \{\tau \in S_n \mid \tau(n) = n\}$, we have $\sigma(n) = n$ and thus

$$\left(g_q \circ \sigma \circ (g_p)^{-1}\right)(p) = g_q\left(\sigma\left(\underbrace{(g_p)^{-1}(p)}_{=n}\right)\right) = g_q\left(\underbrace{\sigma(n)}_{=n}\right) = g_q(n) = q$$

and therefore $g_q \circ \sigma \circ (g_p)^{-1} \in \{\tau \in S_n \mid \tau(p) = q\}$. Thus, the map T is well-defined.

We can also define a map

$$Q : \{\tau \in S_n \mid \tau(p) = q\} \to \{\tau \in S_n \mid \tau(n) = n\}$$

by

$$\left(Q(\sigma) = (g_q)^{-1} \circ \sigma \circ g_p \qquad \text{for every } \sigma \in \{\tau \in S_n \mid \tau(p) = q\}\right).$$

The well-definedness of Q can be checked similarly to how we proved the well-definedness of T. It is straightforward to verify that the maps Q and T are mutually inverse. Thus, T is bijective. This completes the proof of Lemma 6.83 **(d)**. □

Our next step towards the proof of Theorem 6.82 is the following lemma:

Lemma 6.84. Let $n \in \mathbb{N}$. Let $A = (a_{i,j})_{1 \leq i \leq n,\ 1 \leq j \leq n}$ be an $n \times n$-matrix. Let $p \in \{1, 2, \ldots, n\}$ and $q \in \{1, 2, \ldots, n\}$. Then,

$$\sum_{\substack{\sigma \in S_n; \\ \sigma(p) = q}} (-1)^{\sigma} \prod_{\substack{i \in \{1,2,\ldots,n\}; \\ i \neq p}} a_{i,\sigma(i)} = (-1)^{p+q} \det\left(A_{\sim p, \sim q}\right).$$

Proof of Lemma 6.84. Let us use all notations introduced in Lemma 6.83.

We have $p \in \{1, 2, \ldots, n\} = [n]$. Hence, g_p is well-defined. Similarly, g_q is well-defined. We have

$$\left(g_p(1), g_p(2), \ldots, g_p(n-1)\right) = \left(1, 2, \ldots, \widehat{p}, \ldots, n\right) \tag{423}$$

(by Lemma 6.83 **(a)**) and

$$\left(g_q(1), g_q(2), \ldots, g_q(n-1)\right) = \left(1, 2, \ldots, \widehat{q}, \ldots, n\right) \tag{424}$$

(by Lemma 6.83 **(a)**, applied to q instead of p). Now, the definition of $A_{\sim p,\sim q}$ yields

$$\begin{aligned}
A_{\sim p,\sim q} &= \operatorname{sub}_{1,2,\ldots,\widehat{p},\ldots,n}^{1,2,\ldots,\widehat{q},\ldots,n} A = \operatorname{sub}_{g_p(1),g_p(2),\ldots,g_p(n-1)}^{g_q(1),g_q(2),\ldots,g_q(n-1)} A \qquad \text{(by (423) and (424))} \\
&= \left(a_{g_p(x),g_q(y)}\right)_{1\leq x\leq n-1,\ 1\leq y\leq n-1} \\
&\qquad \left(\text{by the definition of } \operatorname{sub}_{g_p(1),g_p(2),\ldots,g_p(n-1)}^{g_q(1),g_q(2),\ldots,g_q(n-1)} A\right) \\
&= \left(a_{g_p(i),g_q(j)}\right)_{1\leq i\leq n-1,\ 1\leq j\leq n-1} \\
&\qquad \left(\text{here, we renamed the index } (x,y) \text{ as } (i,j)\right).
\end{aligned} \tag{425}$$

Also, $[n]$ is nonempty (since $p \in [n]$), and thus we have $n > 0$.

Now, let us recall the map $T : \{\tau \in S_n \mid \tau(n) = n\} \to \{\tau \in S_n \mid \tau(p) = q\}$ defined in Lemma 6.83 **(d)**. Lemma 6.83 **(d)** says that this map T is well-defined and bijective. Every $\sigma \in \{\tau \in S_n \mid \tau(n) = n\}$ satisfies

$$(-1)^{T(\sigma)} = (-1)^{p+q} \cdot (-1)^{\sigma} \tag{426}$$

[250] and

$$\prod_{\substack{i\in\{1,2,\ldots,n\};\\ i\neq p}} a_{i,(T(\sigma))(i)} = \prod_{i=1}^{n-1} a_{g_p(i),g_q(\sigma(i))} \qquad (427)$$

[251].

[250] *Proof of (426):* Let $\sigma \in \{\tau \in S_n \mid \tau(n) = n\}$. Applying Lemma 6.83 **(b)** to q instead of p, we obtain $(-1)^{g_q} = (-1)^{n-q} = (-1)^{n+q}$ (since $n - q \equiv n + q \mod 2$).

The definition of $T(\sigma)$ yields $T(\sigma) = g_q \circ \sigma \circ (g_p)^{-1}$. Thus,

$$\underbrace{T(\sigma)}_{=g_q\circ\sigma\circ(g_p)^{-1}} \circ g_p = g_q \circ \sigma \circ \underbrace{(g_p)^{-1} \circ g_p}_{=\mathrm{id}} = g_q \circ \sigma,$$

so that

$$\begin{aligned}(-1)^{T(\sigma)\circ g_p} &= (-1)^{g_q\circ\sigma} = \underbrace{(-1)^{g_q}}_{=(-1)^{n+q}} \cdot (-1)^{\sigma} \qquad \left(\text{by (315), applied to } g_q \text{ and } \sigma \text{ instead of } \sigma \text{ and } \tau\right)\\ &= (-1)^{n+q} \cdot (-1)^{\sigma}.\end{aligned}$$

Compared with

$$\begin{aligned}(-1)^{T(\sigma)\circ g_p} &= (-1)^{T(\sigma)} \cdot \underbrace{(-1)^{g_p}}_{\substack{=(-1)^{n-p}\\ \text{(by Lemma 6.83 (b))}}} \qquad \left(\text{by (315), applied to } T(\sigma) \text{ and } g_p \text{ instead of } \sigma \text{ and } \tau\right)\\ &= (-1)^{T(\sigma)} \cdot (-1)^{n-p},\end{aligned}$$

this yields

$$(-1)^{T(\sigma)} \cdot (-1)^{n-p} = (-1)^{n+q} \cdot (-1)^{\sigma}.$$

We can divide both sides of this equality by $(-1)^{n-p}$ (since $(-1)^{n-p} \in \{1, -1\}$ is clearly an invertible integer), and thus we obtain

$$(-1)^{T(\sigma)} = \frac{(-1)^{n+q} \cdot (-1)^{\sigma}}{(-1)^{n-p}} = \underbrace{\frac{(-1)^{n+q}}{(-1)^{n-p}}}_{\substack{=(-1)^{(n+q)-(n-p)}=(-1)^{p+q}\\ \text{(since } (n+q)-(n-p)=p+q)}} \cdot (-1)^{\sigma} = (-1)^{p+q} \cdot (-1)^{\sigma}.$$

This proves (426).

[251] *Proof of (427):* Let $\sigma \in \{\tau \in S_n \mid \tau(n) = n\}$. Let us recall the map $g'_p : [n-1] \to [n] \setminus \{p\}$ introduced in Lemma 6.83 **(c)**. Lemma 6.83 **(c)** says that this map g'_p is well-defined and bijective. In other words, g'_p is a bijection.

Let $i \in [n-1]$. Then, $g'_p(i) = g_p(i)$ (by the definition of g'_p). Also, the definition of T yields $T(\sigma) = g_q \circ \sigma \circ (g_p)^{-1}$, so that

$$\left(\underbrace{T(\sigma)}_{=g_q\circ\sigma\circ(g_p)^{-1}}\right)\left(\underbrace{g'_p(i)}_{=g_p(i)}\right) = \left(g_q \circ \sigma \circ (g_p)^{-1}\right)(g_p(i)) = g_q\left(\sigma\left(\underbrace{(g_p)^{-1}(g_p(i))}_{=i}\right)\right) = g_q(\sigma(i)).$$

From $g'_p(i) = g_p(i)$ and $(T(\sigma))\left(g'_p(i)\right) = g_q(\sigma(i))$, we obtain

$$a_{g'_p(i),(T(\sigma))\left(g'_p(i)\right)} = a_{g_p(i),g_q(\sigma(i))}. \qquad (428)$$

Now, let us forget that we fixed i. We thus have proven (428) for every $i \in [n-1]$. But now, we have

$$\begin{aligned}
&\underbrace{\prod_{\substack{i\in\{1,2,\ldots,n\};\\ i\neq p}}}_{\substack{=\prod_{\substack{i\in[n];\\ i\neq p}}\\ \text{(since }\{1,2,\ldots,n\}=[n])}} a_{i,(T(\sigma))(i)}\\
&= \underbrace{\prod_{\substack{i\in[n];\\ i\neq p}}}_{=\prod_{i\in[n]\setminus\{p\}}} a_{i,(T(\sigma))(i)} = \prod_{i\in[n]\setminus\{p\}} a_{i,(T(\sigma))(i)} = \underbrace{\prod_{i\in[n-1]}}_{=\prod_{i=1}^{n-1}} \underbrace{a_{g'_p(i),(T(\sigma))\left(g'_p(i)\right)}}_{\substack{=a_{g_p(i),g_q(\sigma(i))}\\ \text{(by (428))}}}\\
&\qquad \left(\begin{array}{c}\text{here, we have substituted } g'_p(i) \text{ for } i\text{, since}\\ g'_p:[n-1]\to[n]\setminus\{p\} \text{ is a bijection}\end{array}\right)\\
&= \prod_{i=1}^{n-1} a_{g_p(i),g_q(\sigma(i))}.
\end{aligned}$$

This proves (427).

Now,

$$\begin{aligned}
&\underbrace{\sum_{\substack{\sigma\in S_n;\\ \sigma(p)=q}}}_{=\sum_{\sigma\in\{\tau\in S_n\ \mid\ \tau(p)=q\}}} (-1)^{\sigma} \prod_{\substack{i\in\{1,2,\ldots,n\};\\ i\neq p}} a_{i,\sigma(i)} \\
&= \sum_{\sigma\in\{\tau\in S_n\ \mid\ \tau(p)=q\}} (-1)^{\sigma} \prod_{\substack{i\in\{1,2,\ldots,n\};\\ i\neq p}} a_{i,\sigma(i)} \\
&= \underbrace{\sum_{\sigma\in\{\tau\in S_n\ \mid\ \tau(n)=n\}}}_{=\sum_{\substack{\sigma\in S_n;\\ \sigma(n)=n}}} \underbrace{(-1)^{T(\sigma)}}_{\substack{=(-1)^{p+q}\cdot(-1)^{\sigma}\\ \text{(by (426))}}} \underbrace{\prod_{\substack{i\in\{1,2,\ldots,n\};\\ i\neq p}} a_{i,(T(\sigma))(i)}}_{\substack{=\prod_{i=1}^{n-1} a_{g_p(i),g_q(\sigma(i))}\\ \text{(by (427))}}} \\
&\qquad \left(\begin{array}{c}\text{here, we have substituted } T\left(\sigma\right) \text{ for } \sigma \text{ in the sum,}\\ \text{since the map } T:\left\{\tau\in S_n\ \mid\ \tau\left(n\right)=n\right\}\to\left\{\tau\in S_n\ \mid\ \tau\left(p\right)=q\right\}\\ \text{is a bijection}\end{array}\right) \\
&= \sum_{\substack{\sigma\in S_n;\\ \sigma(n)=n}} (-1)^{p+q}\cdot(-1)^{\sigma} \prod_{i=1}^{n-1} a_{g_p(i),g_q(\sigma(i))} \\
&= (-1)^{p+q} \underbrace{\sum_{\substack{\sigma\in S_n;\\ \sigma(n)=n}} (-1)^{\sigma} \prod_{i=1}^{n-1} a_{g_p(i),g_q(\sigma(i))}}_{\substack{=\det\left(\left(a_{g_p(i),g_q(j)}\right)_{1\leq i\leq n-1,\ 1\leq j\leq n-1}\right)\\ \text{(by Lemma 6.44, applied to } a_{g_p(i),g_q(j)} \text{ instead of } a_{i,j}\text{)}}} \\
&= (-1)^{p+q}\det\left(\underbrace{\left(a_{g_p(i),g_q(j)}\right)_{1\leq i\leq n-1,\ 1\leq j\leq n-1}}_{\substack{=A_{\sim p,\sim q}\\ \text{(by (425))}}}\right) = (-1)^{p+q}\det\left(A_{\sim p,\sim q}\right).
\end{aligned}$$

This proves Lemma 6.84. □

Now, we can finally prove Theorem 6.82:

Proof of Theorem 6.82. **(a)** Let $p \in \{1, 2, \ldots, n\}$. From (342), we obtain

$$\begin{aligned}
\det A &= \sum_{\sigma \in S_n} (-1)^{\sigma} \prod_{i=1}^{n} a_{i,\sigma(i)} \\
&= \sum_{q \in \{1,2,\ldots,n\}} \sum_{\substack{\sigma \in S_n;\\ \sigma(p)=q}} (-1)^{\sigma} \underbrace{\prod_{i=1}^{n}}_{= \prod_{i \in \{1,2,\ldots,n\}}} a_{i,\sigma(i)} \\
&\qquad \left(\begin{array}{c}\text{because for every } \sigma \in S_n \text{, there exists} \\ \text{exactly one } q \in \{1, 2, \ldots, n\} \text{ satisfying } \sigma(p) = q\end{array}\right) \\
&= \sum_{q \in \{1,2,\ldots,n\}} \sum_{\substack{\sigma \in S_n;\\ \sigma(p)=q}} (-1)^{\sigma} \underbrace{\prod_{i \in \{1,2,\ldots,n\}} a_{i,\sigma(i)}}_{\substack{= a_{p,\sigma(p)} \prod_{\substack{i \in \{1,2,\ldots,n\};\\ i \neq p}} a_{i,\sigma(i)} \\ \text{(here, we have split off the factor for } i=p \text{ from the product)}}} \\
&= \underbrace{\sum_{q \in \{1,2,\ldots,n\}}}_{= \sum_{q=1}^{n}} \sum_{\substack{\sigma \in S_n;\\ \sigma(p)=q}} (-1)^{\sigma} \underbrace{a_{p,\sigma(p)}}_{\substack{= a_{p,q}\\ \text{(since } \sigma(p)=q\text{)}}} \prod_{\substack{i \in \{1,2,\ldots,n\};\\ i \neq p}} a_{i,\sigma(i)} \\
&= \sum_{q=1}^{n} \underbrace{\sum_{\substack{\sigma \in S_n;\\ \sigma(p)=q}} (-1)^{\sigma} a_{p,q} \prod_{\substack{i \in \{1,2,\ldots,n\};\\ i \neq p}} a_{i,\sigma(i)}}_{= a_{p,q} \sum_{\substack{\sigma \in S_n;\\ \sigma(p)=q}} (-1)^{\sigma} \prod_{\substack{i \in \{1,2,\ldots,n\};\\ i \neq p}} a_{i,\sigma(i)}} \\
&= \sum_{q=1}^{n} a_{p,q} \underbrace{\sum_{\substack{\sigma \in S_n;\\ \sigma(p)=q}} (-1)^{\sigma} \prod_{\substack{i \in \{1,2,\ldots,n\};\\ i \neq p}} a_{i,\sigma(i)}}_{\substack{= (-1)^{p+q} \det\left(A_{\sim p, \sim q}\right) \\ \text{(by Lemma 6.84)}}} \\
&= \sum_{q=1}^{n} \underbrace{a_{p,q} (-1)^{p+q}}_{= (-1)^{p+q} a_{p,q}} \det\left(A_{\sim p, \sim q}\right) = \sum_{q=1}^{n} (-1)^{p+q} a_{p,q} \det\left(A_{\sim p, \sim q}\right).
\end{aligned}$$

This proves Theorem 6.82 **(a)**.

(b) Let $q \in \{1, 2, \ldots, n\}$. From (342), we obtain

$$\begin{aligned}
\det A &= \sum_{\sigma\in S_n} (-1)^\sigma \prod_{i=1}^n a_{i,\sigma(i)} \\
&= \sum_{p\in\{1,2,\ldots,n\}} \sum_{\substack{\sigma\in S_n;\\ \sigma^{-1}(q)=p}} (-1)^\sigma \underbrace{\prod_{i=1}^n}_{=\prod_{i\in\{1,2,\ldots,n\}}} a_{i,\sigma(i)} \\
&\qquad \left(\begin{array}{c}\text{because for every } \sigma\in S_n \text{, there exists}\\ \text{exactly one } p\in\{1,2,\ldots,n\} \text{ satisfying } \sigma^{-1}(q)=p\end{array}\right) \\
&= \sum_{p\in\{1,2,\ldots,n\}} \underbrace{\sum_{\substack{\sigma\in S_n;\\ \sigma^{-1}(q)=p}}}_{\substack{=\sum_{\substack{\sigma\in S_n;\\ \sigma(p)=q}}\\ \text{(because for any } \sigma\in S_n\text{,}\\ \text{the statement } \left(\sigma^{-1}(q)=p\right)\\ \text{is equivalent to the}\\ \text{statement } (\sigma(p)=q))}} (-1)^\sigma \underbrace{\prod_{i\in\{1,2,\ldots,n\}} a_{i,\sigma(i)}}_{\substack{=a_{p,\sigma(p)} \prod_{\substack{i\in\{1,2,\ldots,n\};\\ i\neq p}} a_{i,\sigma(i)}\\ \text{(here, we have split off the}\\ \text{factor for } i=p \text{ from the product)}}} \\
&= \underbrace{\sum_{p\in\{1,2,\ldots,n\}}}_{=\sum_{p=1}^n} \sum_{\substack{\sigma\in S_n;\\ \sigma(p)=q}} (-1)^\sigma \underbrace{a_{p,\sigma(p)}}_{\substack{=a_{p,q}\\ \text{(since } \sigma(p)=q\text{)}}} \prod_{\substack{i\in\{1,2,\ldots,n\};\\ i\neq p}} a_{i,\sigma(i)} \\
&= \sum_{p=1}^n \underbrace{\sum_{\substack{\sigma\in S_n;\\ \sigma(p)=q}} (-1)^\sigma a_{p,q} \prod_{\substack{i\in\{1,2,\ldots,n\};\\ i\neq p}} a_{i,\sigma(i)}}_{=a_{p,q} \sum_{\substack{\sigma\in S_n;\\ \sigma(p)=q}} (-1)^\sigma \prod_{\substack{i\in\{1,2,\ldots,n\};\\ i\neq p}} a_{i,\sigma(i)}} \\
&= \sum_{p=1}^n a_{p,q} \underbrace{\sum_{\substack{\sigma\in S_n;\\ \sigma(p)=q}} (-1)^\sigma \prod_{\substack{i\in\{1,2,\ldots,n\};\\ i\neq p}} a_{i,\sigma(i)}}_{\substack{=(-1)^{p+q}\det\left(A_{\sim p,\sim q}\right)\\ \text{(by Lemma 6.84)}}} \\
&= \sum_{p=1}^n \underbrace{a_{p,q}(-1)^{p+q}}_{=(-1)^{p+q}a_{p,q}} \det\left(A_{\sim p,\sim q}\right) = \sum_{p=1}^n (-1)^{p+q} a_{p,q} \det\left(A_{\sim p,\sim q}\right).
\end{aligned}$$

This proves Theorem 6.82 **(b)**. □

Let me make three simple observations (which can easily be checked by the reader):

- Theorem 6.82 **(b)** could be (alternatively) proven using Theorem 6.82 **(a)** (applied to A^T and $a_{q,p}$ instead of A and $a_{p,q}$) and Exercise 6.4.
- Theorem 6.43 is a particular case of Theorem 6.82 **(a)**.
- Corollary 6.45 is a particular case of Theorem 6.82 **(b)**.

Remark 6.85. Some books use Laplace expansion to define the notion of a determinant. For example, one can define the determinant of a square matrix recursively, by setting the determinant of the 0×0-matrix to be 1, and defining the determinant of an $n \times n$-matrix $A = \left(a_{i,j}\right)_{1\leq i\leq n,\ 1\leq j\leq n}$ (with $n > 0$) to be $\sum\limits_{q=1}^{n} \left(-1\right)^{1+q} a_{1,q} \det\left(A_{\sim 1,\sim q}\right)$ (assuming that determinants of $(n-1) \times (n-1)$-matrices such as $A_{\sim 1,\sim q}$ are already defined). Of course, this leads to the same notion of determinant as the one we are using, because of Theorem 6.82 **(a)**.

6.13. Tridiagonal determinants

In this section, we shall study the so-called *tridiagonal matrices*: a class of matrices whose all entries are zero everywhere except in the "direct proximity" of the diagonal (more specifically: on the diagonal and "one level below and one level above"). We shall find recursive formulas for the determinants of these matrices. These formulas are a simple example of an application of Laplace expansion, but also interesting in their own right.

Definition 6.86. Let $n \in \mathbb{N}$. Let $a_1, a_2, \ldots, a_n$ be n elements of $\mathbb{K}$. Let $b_1, b_2, \ldots, b_{n-1}$ be $n-1$ elements of $\mathbb{K}$ (where we take the position that "-1 elements of $\mathbb{K}$" means "no elements of $\mathbb{K}$"). Let $c_1, c_2, \ldots, c_{n-1}$ be $n-1$ elements of $\mathbb{K}$. We now set

$$A = \begin{pmatrix} a_1 & b_1 & 0 & \cdots & 0 & 0 & 0 \\ c_1 & a_2 & b_2 & \cdots & 0 & 0 & 0 \\ 0 & c_2 & a_3 & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & a_{n-2} & b_{n-2} & 0 \\ 0 & 0 & 0 & \cdots & c_{n-2} & a_{n-1} & b_{n-1} \\ 0 & 0 & 0 & \cdots & 0 & c_{n-1} & a_n \end{pmatrix}.$$

(More formally,

$$A = \left(\begin{cases} a_i, & \text{if } i = j; \\ b_i, & \text{if } i = j-1; \\ c_j, & \text{if } i = j+1; \\ 0, & \text{otherwise} \end{cases} \right)_{1\leq i\leq n,\ 1\leq j\leq n}.$$

)

The matrix A is called a *tridiagonal matrix*.

We shall keep the notations n, $a_1, a_2, \ldots, a_n$, $b_1, b_2, \ldots, b_{n-1}$, $c_1, c_2, \ldots, c_{n-1}$ and A fixed for the rest of Section 6.13.

Playing around with small examples, one soon notices that the determinants of tridiagonal matrices are too complicated to have neat explicit formulas in full generality. For $n \in \{0, 1, 2, 3\}$, the determinants look as follows:

$$\det A = \det(\text{the } 0 \times 0\text{-matrix}) = 1 \qquad \text{if } n = 0;$$
$$\det A = \det \begin{pmatrix} a_1 \end{pmatrix} = a_1 \qquad \text{if } n = 1;$$
$$\det A = \det \begin{pmatrix} a_1 & b_1 \\ c_1 & a_2 \end{pmatrix} = a_1 a_2 - b_1 c_1 \qquad \text{if } n = 2;$$
$$\det A = \det \begin{pmatrix} a_1 & b_1 & 0 \\ c_1 & a_2 & b_2 \\ 0 & c_2 & a_3 \end{pmatrix} = a_1 a_2 a_3 - a_1 b_2 c_2 - a_3 b_1 c_1 \qquad \text{if } n = 3.$$

(And these formulas get more complicated the larger n becomes.) However, the many zeroes present in a tridiagonal matrix make it easy to find a recursive formula for its determinant using Laplace expansion:

Proposition 6.87. For every two elements x and y of $\{0, 1, \ldots, n\}$ satisfying $x \leq y$, we let $A_{x,y}$ be the $(y - x) \times (y - x)$-matrix

$$\begin{pmatrix} a_{x+1} & b_{x+1} & 0 & \cdots & 0 & 0 & 0 \\ c_{x+1} & a_{x+2} & b_{x+2} & \cdots & 0 & 0 & 0 \\ 0 & c_{x+2} & a_{x+3} & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & a_{y-2} & b_{y-2} & 0 \\ 0 & 0 & 0 & \cdots & c_{y-2} & a_{y-1} & b_{y-1} \\ 0 & 0 & 0 & \cdots & 0 & c_{y-1} & a_y \end{pmatrix} = \operatorname{sub}_{x+1,x+2,\ldots,y}^{x+1,x+2,\ldots,y} A.$$

(a) We have $\det(A_{x,x}) = 1$ for every $x \in \{0, 1, \ldots, n\}$.

(b) We have $\det(A_{x,x+1}) = a_{x+1}$ for every $x \in \{0, 1, \ldots, n-1\}$.

(c) For every $x \in \{0, 1, \ldots, n\}$ and $y \in \{0, 1, \ldots, n\}$ satisfying $x \leq y - 2$, we have

$$\det(A_{x,y}) = a_y \det(A_{x,y-1}) - b_{y-1} c_{y-1} \det(A_{x,y-2}).$$

(d) For every $x \in \{0, 1, \ldots, n\}$ and $y \in \{0, 1, \ldots, n\}$ satisfying $x \leq y - 2$, we have

$$\det(A_{x,y}) = a_{x+1} \det(A_{x+1,y}) - b_{x+1} c_{x+1} \det(A_{x+2,y}).$$

(e) We have $A = A_{0,n}$.

Proof of Proposition 6.87. **(e)** The definition of $A_{0,n}$ yields

$$A_{0,n} = \begin{pmatrix} a_1 & b_1 & 0 & \cdots & 0 & 0 & 0 \\ c_1 & a_2 & b_2 & \cdots & 0 & 0 & 0 \\ 0 & c_2 & a_3 & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & a_{n-2} & b_{n-2} & 0 \\ 0 & 0 & 0 & \cdots & c_{n-2} & a_{n-1} & b_{n-1} \\ 0 & 0 & 0 & \cdots & 0 & c_{n-1} & a_n \end{pmatrix} = A.$$

This proves Proposition 6.87 **(e)**.

(a) Let $x \in \{0, 1, \ldots, n\}$. Then, $A_{x,x}$ is an $(x-x) \times (x-x)$-matrix, thus a 0×0-matrix. Hence, its determinant is $\det(A_{x,x}) = 1$. This proves Proposition 6.87 **(a)**.

(b) Let $x \in \{0, 1, \ldots, n-1\}$. The definition of $A_{x,x+1}$ shows that $A_{x,x+1}$ is the 1×1-matrix $\left(a_{x+1} \right)$. Hence, $\det(A_{x,x+1}) = \det\left(a_{x+1} \right) = a_{x+1}$. This proves Proposition 6.87 **(b)**.

(c) Let $x \in \{0, 1, \ldots, n\}$ and $y \in \{0, 1, \ldots, n\}$ be such that $x \leq y-2$. We have

$$A_{x,y} = \begin{pmatrix} a_{x+1} & b_{x+1} & 0 & \cdots & 0 & 0 & 0 \\ c_{x+1} & a_{x+2} & b_{x+2} & \cdots & 0 & 0 & 0 \\ 0 & c_{x+2} & a_{x+3} & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & a_{y-2} & b_{y-2} & 0 \\ 0 & 0 & 0 & \cdots & c_{y-2} & a_{y-1} & b_{y-1} \\ 0 & 0 & 0 & \cdots & 0 & c_{y-1} & a_y \end{pmatrix}. \qquad (429)$$

This is a $(y-x) \times (y-x)$-matrix. If we cross out its $(y-x)$-th row (i.e., its last row) and its $(y-x)$-th column (i.e., its last column), then we obtain $A_{x,y-1}$. In other words, $\left(A_{x,y}\right)_{\sim(y-x),\sim(y-x)} = A_{x,y-1}$.

Let us write the matrix $A_{x,y}$ in the form $A_{x,y} = \left(u_{i,j}\right)_{1 \leq i \leq y-x,\ 1 \leq j \leq y-x}$. Thus,

$$\begin{aligned} & \left(u_{y-x,1}, u_{y-x,2}, \ldots, u_{y-x,y-x}\right) \\ & = \left(\text{the last row of the matrix } A_{x,y}\right) = \left(0, 0, \ldots, 0, c_{y-1}, a_y\right). \end{aligned}$$

In other words, we have

$$\begin{aligned} & \left(u_{y-x,q} = 0 \qquad \text{for every } q \in \{1, 2, \ldots, y-x-2\}\right), \qquad (430) \\ & u_{y-x,y-x-1} = c_{y-1}, \qquad \text{and} \\ & u_{y-x,y-x} = a_y. \end{aligned}$$

Now, Laplace expansion along the $(y-x)$-th row (or, more precisely, Theorem

6.82 **(a)**, applied to $y-x$, $A_{x,y}$, $u_{i,j}$ and $y-x$ instead of n, A, $a_{i,j}$ and p) yields

$$\begin{aligned}
\det\left(A_{x,y}\right) &= \sum_{q=1}^{y-x} (-1)^{(y-x)+q} u_{y-x,q} \det\left(\left(A_{x,y}\right)_{\sim(y-x),\sim q}\right) \\
&= \sum_{q=1}^{y-x-2} (-1)^{(y-x)+q} \underbrace{u_{y-x,q}}_{\substack{=0 \\ \text{(by (430))}}} \det\left(\left(A_{x,y}\right)_{\sim(y-x),\sim q}\right) \\
&\qquad + \underbrace{(-1)^{(y-x)+(y-x-1)}}_{=-1} \underbrace{u_{y-x,y-x-1}}_{=c_{y-1}} \det\left(\left(A_{x,y}\right)_{\sim(y-x),\sim(y-x-1)}\right) \\
&\qquad + \underbrace{(-1)^{(y-x)+(y-x)}}_{=1} \underbrace{u_{y-x,y-x}}_{=a_y} \det\left(\underbrace{\left(A_{x,y}\right)_{\sim(y-x),\sim(y-x)}}_{=A_{x,y-1}}\right) \\
&\qquad \left(\text{since } y-x \geq 2 \text{ (since } x \leq y-2\text{)}\right) \\
&= \underbrace{\sum_{q=1}^{y-x-2} (-1)^{(y-x)+q} 0 \det\left(\left(A_{x,y}\right)_{\sim(y-x),\sim q}\right)}_{=0} \\
&\qquad - c_{y-1} \det\left(\left(A_{x,y}\right)_{\sim(y-x),\sim(y-x-1)}\right) + a_y \det\left(A_{x,y-1}\right) \\
&= -c_{y-1} \det\left(\left(A_{x,y}\right)_{\sim(y-x),\sim(y-x-1)}\right) + a_y \det\left(A_{x,y-1}\right). \qquad (431)
\end{aligned}$$

Now, let $B = \left(A_{x,y}\right)_{\sim(y-x),\sim(y-x-1)}$. Thus, (431) becomes

$$\begin{aligned}
\det\left(A_{x,y}\right) &= -c_{y-1} \det\left(\underbrace{\left(A_{x,y}\right)_{\sim(y-x),\sim(y-x-1)}}_{=B}\right) + a_y \det\left(A_{x,y-1}\right) \\
&= -c_{y-1} \det B + a_y \det\left(A_{x,y-1}\right). \qquad (432)
\end{aligned}$$

Now,

$$\begin{aligned}
B &= \left(A_{x,y}\right)_{\sim(y-x),\sim(y-x-1)} \\
&= \begin{pmatrix}
a_{x+1} & b_{x+1} & 0 & \cdots & 0 & 0 & 0 \\
c_{x+1} & a_{x+2} & b_{x+2} & \cdots & 0 & 0 & 0 \\
0 & c_{x+2} & a_{x+3} & \cdots & 0 & 0 & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\
0 & 0 & 0 & \cdots & a_{y-3} & b_{y-3} & 0 \\
0 & 0 & 0 & \cdots & c_{y-3} & a_{y-2} & 0 \\
0 & 0 & 0 & \cdots & 0 & c_{y-2} & b_{y-1}
\end{pmatrix} \qquad (\text{because of (429)}).
\end{aligned}$$

(433)

Now, let us write the matrix B in the form $B = \left(v_{i,j}\right)_{1\leq i\leq y-x-1,\ 1\leq j\leq y-x-1}$. Thus,

$$\begin{aligned}&\left(v_{1,y-x-1}, v_{2,y-x-1}, \ldots, v_{y-x-1,y-x-1}\right)^T \\ &= \left(\text{the last column of the matrix } B\right) = \left(0, 0, \ldots, 0, b_{y-1}\right)^T\end{aligned}$$

(because of (433)). In other words, we have

$$\begin{aligned}&\left(v_{p,y-x-1} = 0 \qquad \text{for every } p \in \{1, 2, \ldots, y-x-2\}\right), \qquad \text{and} \\ &v_{y-x-1,y-x-1} = b_{y-1}.\end{aligned} \tag{434}$$

Now, Laplace expansion along the $(y-x-1)$-th column (or, more precisely, Theorem 6.82 **(b)**, applied to $y-x-1$, B, $v_{i,j}$ and $y-x-1$ instead of n, A, $a_{i,j}$ and q) yields

$$\begin{aligned}\det B &= \sum_{p=1}^{y-x-1} (-1)^{p+(y-x-1)} v_{p,y-x-1} \det\left(B_{\sim p,\sim(y-x-1)}\right) \\ &= \sum_{p=1}^{y-x-2} (-1)^{p+(y-x-1)} \underbrace{v_{p,y-x-1}}_{\substack{=0 \\ \text{(by (434))}}} \det\left(B_{\sim p,\sim(y-x-1)}\right) \\ &\qquad + \underbrace{(-1)^{(y-x-1)+(y-x-1)}}_{=1} \underbrace{v_{y-x-1,y-x-1}}_{=b_{y-1}} \det\left(B_{\sim(y-x-1),\sim(y-x-1)}\right) \\ &\qquad \left(\text{since } y-x-1 \geq 1 \text{ (since } x \leq y-2)\right) \\ &= \underbrace{\sum_{p=1}^{y-x-2} (-1)^{p+(y-x-1)} 0 \det\left(B_{\sim p,\sim(y-x-1)}\right)}_{=0} + b_{y-1} \det\left(B_{\sim(y-x-1),\sim(y-x-1)}\right) \\ &= b_{y-1} \det\left(B_{\sim(y-x-1),\sim(y-x-1)}\right).\end{aligned} \tag{435}$$

Finally, a look at (433) reveals that

$$B_{\sim(y-x-1),\sim(y-x-1)} = \begin{pmatrix} a_{x+1} & b_{x+1} & 0 & \cdots & 0 & 0 & 0 \\ c_{x+1} & a_{x+2} & b_{x+2} & \cdots & 0 & 0 & 0 \\ 0 & c_{x+2} & a_{x+3} & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & a_{y-4} & b_{y-4} & 0 \\ 0 & 0 & 0 & \cdots & c_{y-4} & a_{y-3} & b_{y-3} \\ 0 & 0 & 0 & \cdots & 0 & c_{y-3} & a_{y-2} \end{pmatrix} = A_{x,y-2}.$$

Hence, (435) becomes

$$\det B = b_{y-1} \det\left(\underbrace{B_{\sim(y-x-1),\sim(y-x-1)}}_{=A_{x,y-2}}\right) = b_{y-1} \det\left(A_{x,y-2}\right).$$

Therefore, (432) becomes

$$\begin{aligned}\det\left(A_{x,y}\right) &= -c_{y-1}\underbrace{\det B}_{=b_{y-1}\det\left(A_{x,y-2}\right)} + a_y \det\left(A_{x,y-1}\right)\\ &= -c_{y-1}b_{y-1}\det\left(A_{x,y-2}\right) + a_y\det\left(A_{x,y-1}\right)\\ &= a_y\det\left(A_{x,y-1}\right) - b_{y-1}c_{y-1}\det\left(A_{x,y-2}\right).\end{aligned}$$

This proves Proposition 6.87 **(c)**.

(d) The proof of Proposition 6.87 **(d)** is similar to the proof of Proposition 6.87 **(c)**. The main difference is that we now have to perform Laplace expansion along the 1-st row (instead of the $(y-x)$-th row) and then Laplace expansion along the 1-st column (instead of the $(y-x-1)$-th column). □

Proposition 6.87 gives us two fast recursive algorithms to compute $\det A$:

The first algorithm proceeds by recursively computing $\det\left(A_{0,m}\right)$ for every $m \in \{0,1,\ldots,n\}$. This is done using Proposition 6.87 **(a)** (for $m = 0$), Proposition 6.87 **(b)** (for $m = 1$) and Proposition 6.87 **(c)** (to find $\det\left(A_{0,m}\right)$ for $m \geq 2$ in terms of $\det\left(A_{0,m-1}\right)$ and $\det\left(A_{0,m-2}\right)$). The final value $\det\left(A_{0,n}\right)$ is $\det A$ (by Proposition 6.87 **(e)**).

The second algorithm proceeds by recursively computing $\det\left(A_{m,n}\right)$ for every $m \in \{0,1,\ldots,n\}$. This recursion goes backwards: We start with $m = n$ (where we use Proposition 6.87 **(a)**), then turn to $m = n-1$ (using Proposition 6.87 **(b)**), and then go further and further down (using Proposition 6.87 **(d)** to compute $\det\left(A_{m,n}\right)$ in terms of $\det\left(A_{m+1,n}\right)$ and $\det\left(A_{m+2,n}\right)$).

So we have two different recursive algorithms leading to one and the same result. Whenever you have such a thing, you can package up the equivalence of the two algorithms as an exercise, and try to make it less easy by covering up the actual goal of the algorithms (in our case, computing $\det A$). In our case, this leads to the following exercise:

Exercise 6.25. Let $n \in \mathbb{N}$. Let $a_1, a_2, \ldots, a_n$ be n elements of $\mathbb{K}$. Let $b_1, b_2, \ldots, b_{n-1}$ be $n-1$ elements of $\mathbb{K}$.

Define a sequence $\left(u_0, u_1, \ldots, u_n\right)$ of elements of $\mathbb{K}$ recursively by setting $u_0 = 1$, $u_1 = a_1$ and

$$u_i = a_i u_{i-1} - b_{i-1} u_{i-2} \qquad \text{for every } i \in \{2,3,\ldots,n\}.$$

Define a sequence $\left(v_0, v_1, \ldots, v_n\right)$ of elements of $\mathbb{K}$ recursively by setting $v_0 = 1$, $v_1 = a_n$ and

$$v_i = a_{n-i+1} v_{i-1} - b_{n-i+1} v_{i-2} \qquad \text{for every } i \in \{2,3,\ldots,n\}.$$

Prove that $u_n = v_n$.

This exercise generalizes IMO Shortlist 2013 problem A1[252].

Our recursive algorithms for computing $\det A$ also yield another observation: The determinant $\det A$ depends not on the $2(n-1)$ elements $b_1, b_2, \ldots, b_{n-1}, c_1, c_2, \ldots, c_{n-1}$ but only on the products $b_1c_1, b_2c_2, \ldots, b_{n-1}c_{n-1}$.

Exercise 6.26. Define $A_{x,y}$ as in Proposition 6.87. Prove that

$$\frac{\det A}{\det\left(A_{1,n}\right)} = a_1 - \cfrac{b_1c_1}{a_2 - \cfrac{b_2c_2}{a_3 - \cfrac{b_3c_3}{a_4 - \ddots - \cfrac{b_{n-2}c_{n-2}}{a_{n-1} - \cfrac{b_{n-1}c_{n-1}}{a_n}}}}},$$

provided that all denominators in this equality are invertible.

Exercise 6.27. Assume that $a_i = 1$ for all $i \in \{1, 2, \ldots, n\}$. Also, assume that $b_i = 1$ and $c_i = -1$ for all $i \in \{1, 2, \ldots, n-1\}$. Let $(f_0, f_1, f_2, \ldots)$ be the Fibonacci sequence (defined as in Chapter 4). Show that $\det A = f_{n+1}$.

Remark 6.88. Consider once again the Fibonacci sequence $(f_0, f_1, f_2, \ldots)$ (defined as in Chapter 4). Let n be a positive integer. Combining the results of Exercise

[252] I have a suspicion that IMO Shortlist 2009 problem C3 also can be viewed as an equality between two recursive ways to compute a determinant; but this determinant seems to be harder to find (I don't think it can be obtained from Proposition 6.87).

6.26 and Exercise 6.27 (the details are left to the reader), we obtain the equality

$$\begin{aligned}\frac{f_{n+1}}{f_n} &= 1 - \cfrac{1\left(-1\right)}{1 - \cfrac{1\left(-1\right)}{1 - \cfrac{1\left(-1\right)}{1- \cfrac{\ddots}{-\cfrac{1\left(-1\right)}{1-\cfrac{1\left(-1\right)}{1}}}}}} \qquad \left(\text{with } n-1 \text{ fractions in total}\right) \\ &= 1 + \cfrac{1}{1 + \cfrac{1}{1 + \cfrac{1}{1+ \cfrac{\ddots}{+\cfrac{1}{1+\cfrac{1}{1}}}}}} \qquad \left(\text{with } n-1 \text{ fractions in total}\right).\end{aligned}$$

If you know some trivia about the golden ratio, you might recognize this as a part of the continued fraction for the golden ratio φ. The whole continued fraction for φ is

$$\varphi = 1 + \cfrac{1}{1 + \cfrac{1}{1 + \cfrac{1}{1+ \cfrac{}{\ddots}}}} \qquad \left(\text{with infinitely many fractions}\right).$$

This hints at the fact that $\lim\limits_{n\to\infty} \dfrac{f_{n+1}}{f_n} = \varphi$. (This is easy to prove without continued fractions, of course.)

6.14. On block-triangular matrices

Definition 6.89. Let n, n', m and m' be four nonnegative integers.

Let $A = \left(a_{i,j}\right)_{1\leq i\leq n,\ 1\leq j\leq m}$ be an $n \times m$-matrix.

Let $B = \left(b_{i,j}\right)_{1\leq i\leq n,\ 1\leq j\leq m'}$ be an $n \times m'$-matrix.

Let $C = \left(c_{i,j}\right)_{1\leq i\leq n',\ 1\leq j\leq m}$ be an $n' \times m$-matrix.

Let $D = \left(d_{i,j}\right)_{1\leq i\leq n',\ 1\leq j\leq m'}$ be an $n' \times m'$-matrix.

Then, $\begin{pmatrix} A & B \\ C & D \end{pmatrix}$ will mean the $(n+n') \times (m+m')$-matrix

$$\begin{pmatrix}
a_{1,1} & a_{1,2} & \cdots & a_{1,m} & b_{1,1} & b_{1,2} & \cdots & b_{1,m'} \\
a_{2,1} & a_{2,2} & \cdots & a_{2,m} & b_{2,1} & b_{2,2} & \cdots & b_{2,m'} \\
\vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\
a_{n,1} & a_{n,2} & \cdots & a_{n,m} & b_{n,1} & b_{n,2} & \cdots & b_{n,m'} \\
c_{1,1} & c_{1,2} & \cdots & c_{1,m} & d_{1,1} & d_{1,2} & \cdots & d_{1,m'} \\
c_{2,1} & c_{2,2} & \cdots & c_{2,m} & d_{2,1} & d_{2,2} & \cdots & d_{2,m'} \\
\vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\
c_{n',1} & c_{n',2} & \cdots & c_{n',m} & d_{n',1} & d_{n',2} & \cdots & d_{n',m'}
\end{pmatrix}.$$

(Formally speaking, this means that

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix} = \left(\begin{cases} a_{i,j}, & \text{if } i \leq n \text{ and } j \leq m; \\ b_{i,j-m}, & \text{if } i \leq n \text{ and } j > m; \\ c_{i-n,j}, & \text{if } i > n \text{ and } j \leq m; \\ d_{i-n,j-m}, & \text{if } i > n \text{ and } j > m \end{cases} \right)_{1 \leq i \leq n+n',\ 1 \leq j \leq m+m'} . \qquad (436)$$

Less formally, we can say that $\begin{pmatrix} A & B \\ C & D \end{pmatrix}$ is the matrix obtained by gluing the matrices A, B, C and D to form one big $(n+n') \times (m+m')$-matrix, where the right border of A is glued together with the left border of B, the bottom border of A is glued together with the top border of C, etc.)

Do not get fooled by the notation $\begin{pmatrix} A & B \\ C & D \end{pmatrix}$: It is (in general) not a 2×2-matrix, but an $(n+n') \times (m+m')$-matrix, and its entries are not A, B, C and D but the entries of A, B, C and D.

Example 6.90. If $A = \begin{pmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{pmatrix}$, $B = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}$, $C = \begin{pmatrix} c_1 & c_2 \end{pmatrix}$ and $D = \begin{pmatrix} d \end{pmatrix}$, then $\begin{pmatrix} A & B \\ C & D \end{pmatrix} = \begin{pmatrix} a_{1,1} & a_{1,2} & b_1 \\ a_{2,1} & a_{2,2} & b_2 \\ c_1 & c_2 & d \end{pmatrix}$.

The notation $\begin{pmatrix} A & B \\ C & D \end{pmatrix}$ introduced in Definition 6.89 is a particular case of a more general notation – the *block-matrix construction* – for gluing together multiple matrices with matching dimensions[253]. We shall only need the particular case that is Definition 6.89, however.

Definition 6.91. Let $n \in \mathbb{N}$ and $m \in \mathbb{N}$. Recall that $\mathbb{K}^{n\times m}$ is the set of all $n \times m$-matrices.

We use $0_{n\times m}$ (or sometimes just 0) to denote the $n \times m$ *zero matrix*. (As we recall, this is the $n \times m$-matrix whose all entries are 0; in other words, this is the $n \times m$-matrix $(0)_{1\leq i\leq n,\ 1\leq j\leq m}$.)

Exercise 6.28. Let n, n', m, m', ℓ and ℓ' be six nonnegative integers. Let $A \in \mathbb{K}^{n\times m}$, $B \in \mathbb{K}^{n\times m'}$, $C \in \mathbb{K}^{n'\times m}$, $D \in \mathbb{K}^{n'\times m'}$, $A' \in \mathbb{K}^{m\times \ell}$, $B' \in \mathbb{K}^{m\times \ell'}$, $C' \in \mathbb{K}^{m'\times \ell}$ and $D' \in \mathbb{K}^{m'\times \ell'}$. Then, prove that

$$\left(\begin{array}{cc} A & B \\ C & D \end{array}\right)\left(\begin{array}{cc} A' & B' \\ C' & D' \end{array}\right) = \left(\begin{array}{cc} AA'+BC' & AB'+BD' \\ CA'+DC' & CB'+DD' \end{array}\right).$$

Remark 6.92. The intuitive meaning of Exercise 6.28 is that the product of two matrices in "block-matrix notation" can be computed by applying the usual multiplication rule "on the level of blocks", without having to fall back to multiplying single entries. However, when applying Exercise 6.28, do not forget to check that its conditions are satisfied. Let me give an example and a non-example:

Example: If $A = \left(\begin{array}{c} a_1 \\ a_2 \\ a_3 \end{array}\right)$, $B = \left(\begin{array}{cc} b_{1,1} & b_{1,2} \\ b_{2,1} & b_{2,2} \\ b_{3,1} & b_{3,2} \end{array}\right)$, $C = \left(\begin{array}{c} c \end{array}\right)$, $D = \left(\begin{array}{cc} d_1 & d_2 \end{array}\right)$, $A' = \left(\begin{array}{cc} a'_1 & a'_2 \end{array}\right)$, $B' = \left(\begin{array}{cc} b'_1 & b'_2 \end{array}\right)$, $C' = \left(\begin{array}{cc} c'_{1,1} & c'_{1,2} \\ c'_{2,1} & c'_{2,2} \end{array}\right)$ and $D' = \left(\begin{array}{cc} d'_{1,1} & d'_{1,2} \\ d'_{2,1} & d'_{2,2} \end{array}\right)$,

[253] This construction defines an $(n_1+n_2+\cdots+n_x) \times (m_1+m_2+\cdots+m_y)$-matrix

$$\left(\begin{array}{cccc} A_{1,1} & A_{1,2} & \cdots & A_{1,y} \\ A_{2,1} & A_{2,2} & \cdots & A_{2,y} \\ \vdots & \vdots & \ddots & \vdots \\ A_{x,1} & A_{x,2} & \cdots & A_{x,y} \end{array}\right) \tag{437}$$

whenever you have given two nonnegative integers x and y, an x-tuple $(n_1, n_2, \ldots, n_x) \in \mathbb{N}^x$, a y-tuple $(m_1, m_2, \ldots, m_y) \in \mathbb{N}^y$, and an $n_i \times m_j$-matrix $A_{i,j}$ for every $i \in \{1,2,\ldots,x\}$ and every $j \in \{1,2,\ldots,y\}$. I guess you can guess the definition of this matrix. So you start with an "$x \times y$-matrix of matrices" and glue them together to an $(n_1+n_2+\cdots+n_x) \times (m_1+m_2+\cdots+m_y)$-matrix (provided that the dimensions of these matrices allow them to be glued – e.g., you cannot glue a 2×3-matrix to a 4×6-matrix along its right border, nor on any other border).

It is called "block-matrix construction" because the original matrices $A_{i,j}$ appear as "blocks" in the big matrix (437). Most authors define block matrices to be matrices which are "partitioned" into blocks as in (437); this is essentially our construction in reverse: Instead of gluing several "small" matrices into a big one, they study big matrices partitioned into many small matrices. Of course, the properties of their "block matrices" are equivalent to those of our "block-matrix construction".

then Exercise 6.28 can be applied (with $n = 3$, $n' = 1$, $m = 1$, $m' = 2$, $\ell = 2$ and $\ell' = 2$), and thus we obtain

$$\left(\begin{array}{cc} A & B \\ C & D \end{array}\right)\left(\begin{array}{cc} A' & B' \\ C' & D' \end{array}\right) = \left(\begin{array}{cc} AA' + BC' & AB' + BD' \\ CA' + DC' & CB' + DD' \end{array}\right).$$

Non-example: If $A = \left(\begin{array}{c} a_1 \\ a_2 \end{array}\right)$, $B = \left(\begin{array}{cc} b_{1,1} & b_{1,2} \\ b_{2,1} & b_{2,2} \end{array}\right)$, $C = \left(\begin{array}{c} c_1 \\ c_2 \end{array}\right)$, $D = \left(\begin{array}{cc} d_{1,1} & d_{1,2} \\ d_{2,1} & d_{2,2} \end{array}\right)$, $A' = \left(\begin{array}{cc} a'_{1,1} & a'_{1,2} \\ a'_{2,1} & a'_{2,2} \end{array}\right)$, $B' = \left(\begin{array}{cc} b'_{1,1} & b'_{1,2} \\ b'_{2,1} & b'_{2,2} \end{array}\right)$, $C' = \left(\begin{array}{cc} c'_1 & c'_2 \end{array}\right)$ and $D' = \left(\begin{array}{cc} d'_1 & d'_2 \end{array}\right)$, then Exercise 6.28 cannot be applied, because there exist no $n, m, \ell \in \mathbb{N}$ such that $A \in \mathbb{K}^{n\times m}$ and $A' \in \mathbb{K}^{m\times \ell}$. (Indeed, the number of columns of A does not equal the number of rows of A', but these numbers would both have to be m.) The matrices $\left(\begin{array}{cc} A & B \\ C & D \end{array}\right)$ and $\left(\begin{array}{cc} A' & B' \\ C' & D' \end{array}\right)$ still exist in this case, and can even be multiplied, but their product is not given by a simple formula such as the one in Exercise 6.28. Thus, beware of seeing Exercise 6.28 as a panacea for multiplying matrices blockwise.

Exercise 6.29. Let $n \in \mathbb{N}$ and $m \in \mathbb{N}$. Let A be an $n \times n$-matrix. Let B be an $n \times m$-matrix. Let D be an $m \times m$-matrix. Prove that

$$\det \left(\begin{array}{cc} A & B \\ 0_{m\times n} & D \end{array}\right) = \det A \cdot \det D.$$

Example 6.93. Exercise 6.29 (applied to $n = 2$ and $m = 3$) yields

$$\det \left(\begin{array}{ccccc} a_{1,1} & a_{1,2} & b_{1,1} & b_{1,2} & b_{1,3} \\ a_{2,1} & a_{2,2} & b_{2,1} & b_{2,2} & b_{2,3} \\ 0 & 0 & c_{1,1} & c_{1,2} & c_{1,3} \\ 0 & 0 & c_{2,1} & c_{2,2} & c_{2,3} \\ 0 & 0 & c_{3,1} & c_{3,2} & c_{3,3} \end{array}\right) = \det \left(\begin{array}{cc} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{array}\right) \cdot \det \left(\begin{array}{ccc} c_{1,1} & c_{1,2} & c_{1,3} \\ c_{2,1} & c_{2,2} & c_{2,3} \\ c_{3,1} & c_{3,2} & c_{3,3} \end{array}\right).$$

Remark 6.94. Not every determinant of the form $\det \left(\begin{array}{cc} A & B \\ 0_{m\times n} & D \end{array}\right)$ can be computed using Exercise 6.29. In fact, Exercise 6.29 requires A to be an $n \times n$-matrix and D to be an $m \times m$-matrix; thus, both A and D have to be square matrices in order for Exercise 6.29 to be applicable. For instance, Exercise 6.29 cannot be applied to compute $\det \left(\begin{array}{ccc} a_1 & b_{1,1} & b_{1,2} \\ a_2 & b_{2,1} & b_{2,2} \\ 0 & c_1 & c_2 \end{array}\right)$.

Remark 6.95. You might wonder whether Exercise 6.29 generalizes to a formula for $\det \begin{pmatrix} A & B \\ C & D \end{pmatrix}$ when $A \in \mathbb{K}^{n\times n}$, $B \in \mathbb{K}^{n\times m}$, $C \in \mathbb{K}^{m\times n}$ and $D \in \mathbb{K}^{m\times m}$. The general answer is "No". However, when D is invertible, there exists such a formula (the Schur complement formula shown in Exercise 6.36 below). Curiously, there is also a formula for the case when $n = m$ and $CD = DC$ (see [Silves00, Theorem 3]).

We notice that Exercise 6.29 allows us to solve Exercise 6.6 in a new way.

An analogue of Exercise 6.29 exists in which the $0_{m\times n}$ in the lower-left part of the matrix is replaced by a $0_{n\times m}$ in the upper-right part:

Exercise 6.30. Let $n \in \mathbb{N}$ and $m \in \mathbb{N}$. Let A be an $n \times n$-matrix. Let C be an $m \times n$-matrix. Let D be an $m \times m$-matrix. Prove that

$$\det \begin{pmatrix} A & 0_{n\times m} \\ C & D \end{pmatrix} = \det A \cdot \det D.$$

Exercise 6.31. (a) Compute the determinant of the 7×7-matrix

$$\begin{pmatrix} a & 0 & 0 & 0 & 0 & 0 & b \\ 0 & a' & 0 & 0 & 0 & b' & 0 \\ 0 & 0 & a'' & 0 & b'' & 0 & 0 \\ 0 & 0 & 0 & e & 0 & 0 & 0 \\ 0 & 0 & c'' & 0 & d'' & 0 & 0 \\ 0 & c' & 0 & 0 & 0 & d' & 0 \\ c & 0 & 0 & 0 & 0 & 0 & d \end{pmatrix},$$

where $a, a', a'', b, b', b'', c, c', c'', d, d', d'', e$ are elements of $\mathbb{K}$.

(b) Compute the determinant of the 6×6-matrix

$$\begin{pmatrix} a & 0 & 0 & \ell & 0 & 0 \\ 0 & b & 0 & 0 & m & 0 \\ 0 & 0 & c & 0 & 0 & n \\ g & 0 & 0 & d & 0 & 0 \\ 0 & h & 0 & 0 & e & 0 \\ 0 & 0 & k & 0 & 0 & f \end{pmatrix},$$

where $a, b, c, d, e, f, g, h, k, \ell, m, n$ are elements of $\mathbb{K}$.

Exercise 6.32. Invent and solve an exercise on computing determinants.

6.15. The adjugate matrix

We start this section with a variation on Theorem 6.82:

Proposition 6.96. Let $n \in \mathbb{N}$. Let $A = \left(a_{i,j}\right)_{1\leq i\leq n,\ 1\leq j\leq n}$ be an $n \times n$-matrix. Let $r \in \{1, 2, \ldots, n\}$.

(a) For every $p \in \{1, 2, \ldots, n\}$ satisfying $p \neq r$, we have

$$0 = \sum_{q=1}^{n} (-1)^{p+q} a_{r,q} \det\left(A_{\sim p,\sim q}\right).$$

(b) For every $q \in \{1, 2, \ldots, n\}$ satisfying $q \neq r$, we have

$$0 = \sum_{p=1}^{n} (-1)^{p+q} a_{p,r} \det\left(A_{\sim p,\sim q}\right).$$

Proof of Proposition 6.96. **(a)** Let $p \in \{1, 2, \ldots, n\}$ be such that $p \neq r$.

Let C be the $n \times n$-matrix obtained from A by replacing the p-th row of A by the r-th row of A. Thus, the p-th and the r-th rows of C are equal. Therefore, the matrix C has two equal rows (since $p \neq r$). Hence, $\det C = 0$ (by Exercise 6.7 **(e)**, applied to C instead of A).

Let us write the $n \times n$-matrix C in the form $C = \left(c_{i,j}\right)_{1\leq i\leq n,\ 1\leq j\leq n}$.

The p-th row of C equals the r-th row of A (by the construction of C). In other words,

$$c_{p,q} = a_{r,q} \qquad \text{for every } q \in \{1, 2, \ldots, n\}. \tag{438}$$

On the other hand, the matrix C equals the matrix A in all rows but the p-th one (again, by the construction of C). Hence, if we cross out the p-th rows in both C and A, then the matrices C and A become equal. Therefore,

$$C_{\sim p,\sim q} = A_{\sim p,\sim q} \qquad \text{for every } q \in \{1, 2, \ldots, n\} \tag{439}$$

(because the construction of $C_{\sim p,\sim q}$ from C involves crossing out the p-th row, and so does the construction of $A_{\sim p,\sim q}$ from A).

Now, $\det C = 0$, so that

$$\begin{aligned} 0 = \det C &= \sum_{q=1}^{n} (-1)^{p+q} \underbrace{c_{p,q}}_{\substack{=a_{r,q} \\ \text{(by (438))}}} \det\left(\underbrace{C_{\sim p,\sim q}}_{\substack{=A_{\sim p,\sim q} \\ \text{(by (439))}}}\right) \\ &\qquad \left(\text{by Theorem 6.82 \textbf{(a)}, applied to } C \text{ and } c_{i,j} \text{ instead of } A \text{ and } a_{i,j}\right) \\ &= \sum_{q=1}^{n} (-1)^{p+q} a_{r,q} \det\left(A_{\sim p,\sim q}\right). \end{aligned}$$

This proves Proposition 6.96 **(a)**.

(b) This proof is rather similar to the proof of Proposition 6.96 **(a)**, except that rows are now replaced by columns. We leave the details to the reader. □

We now can define the "adjugate" of a matrix:

Definition 6.97. Let $n \in \mathbb{N}$. Let A be an $n \times n$-matrix. We define a new $n \times n$-matrix $\operatorname{adj} A$ by

$$\operatorname{adj} A = \left((-1)^{i+j} \det \left(A_{\sim j, \sim i} \right) \right)_{1 \leq i \leq n,\ 1 \leq j \leq n}.$$

This matrix $\operatorname{adj} A$ is called the *adjugate* of the matrix A. (Some authors call it the "adjunct" or "adjoint" or "classical adjoint" of A instead. However, beware of the word "adjoint": It means too many different things; in particular it has a second meaning for a matrix.)

The appearance of $A_{\sim j, \sim i}$ (not $A_{\sim i, \sim j}$) in Definition 6.97 might be surprising, but it is not a mistake. We will soon see what it is good for.

There is also a related notion, namely that of a "cofactor matrix". The *cofactor matrix* of an $n \times n$-matrix A is defined to be $\left((-1)^{i+j} \det \left(A_{\sim i, \sim j} \right) \right)_{1 \leq i \leq n,\ 1 \leq j \leq n}$. This is, of course, the transpose $(\operatorname{adj} A)^T$ of $\operatorname{adj} A$. The entries of this matrix are called the *cofactors* of A.

Example 6.98. The adjugate of the 0×0-matrix is the 0×0-matrix.

The adjugate of a 1×1-matrix $\left(\begin{array}{c} a \end{array} \right)$ is $\operatorname{adj} \left(\begin{array}{c} a \end{array} \right) = \left(\begin{array}{c} 1 \end{array} \right)$. (Yes, this shows that all 1×1-matrices have the same adjugate.)

The adjugate of a 2×2-matrix $\left(\begin{array}{cc} a & b \\ c & d \end{array} \right)$ is

$$\operatorname{adj} \left(\begin{array}{cc} a & b \\ c & d \end{array} \right) = \left(\begin{array}{cc} d & -b \\ -c & a \end{array} \right).$$

The adjugate of a 3×3-matrix $\left(\begin{array}{ccc} a & b & c \\ d & e & f \\ g & h & i \end{array} \right)$ is

$$\operatorname{adj} \left(\begin{array}{ccc} a & b & c \\ d & e & f \\ g & h & i \end{array} \right) = \left(\begin{array}{ccc} ei - fh & ch - bi & bf - ce \\ fg - di & ai - cg & cd - af \\ dh - ge & bg - ah & ae - bd \end{array} \right).$$

Proposition 6.99. Let $n \in \mathbb{N}$. Let A be an $n \times n$-matrix. Then, $\operatorname{adj} \left(A^T \right) = (\operatorname{adj} A)^T$.

Proof of Proposition 6.99. Let $i \in \{1, 2, \ldots, n\}$ and $j \in \{1, 2, \ldots, n\}$.

From $i \in \{1, 2, \ldots, n\}$, we obtain $1 \leq i \leq n$, so that $n \geq 1$ and thus $n - 1 \in \mathbb{N}$.

The definition of $A_{\sim i,\sim j}$ yields $A_{\sim i,\sim j} = \operatorname{sub}^{1,2,\ldots,\widehat{j},\ldots,n}_{1,2,\ldots,\widehat{i},\ldots,n} A$. But the definition of $\left(A^T\right)_{\sim j,\sim i}$ yields

$$\left(A^T\right)_{\sim j,\sim i} = \operatorname{sub}^{1,2,\ldots,\widehat{i},\ldots,n}_{1,2,\ldots,\widehat{j},\ldots,n} \left(A^T\right). \tag{440}$$

On the other hand, Proposition 6.79 **(e)** (applied to $m = n$, $u = n-1$, $v = n-1$, $(i_1, i_2, \ldots, i_u) = \left(1, 2, \ldots, \widehat{i}, \ldots, n\right)$ and $(j_1, j_2, \ldots, j_v) = \left(1, 2, \ldots, \widehat{j}, \ldots, n\right)$) yields $\left(\operatorname{sub}^{1,2,\ldots,\widehat{j},\ldots,n}_{1,2,\ldots,\widehat{i},\ldots,n} A\right)^T = \operatorname{sub}^{1,2,\ldots,\widehat{i},\ldots,n}_{1,2,\ldots,\widehat{j},\ldots,n} \left(A^T\right)$. Compared with (440), this yields

$$\left(A^T\right)_{\sim j,\sim i} = \left(\underbrace{\operatorname{sub}^{1,2,\ldots,\widehat{j},\ldots,n}_{1,2,\ldots,\widehat{i},\ldots,n} A}_{=A_{\sim i,\sim j}}\right)^T = \left(A_{\sim i,\sim j}\right)^T.$$

Hence,

$$\det\left(\underbrace{\left(A^T\right)_{\sim j,\sim i}}_{=\left(A_{\sim i,\sim j}\right)^T}\right) = \det\left(\left(A_{\sim i,\sim j}\right)^T\right) = \det\left(A_{\sim i,\sim j}\right) \tag{441}$$

(by Exercise 6.4, applied to $n-1$ and $A_{\sim i,\sim j}$ instead of n and A).

Let us now forget that we fixed i and j. We thus have shown that (441) holds for every $i \in \{1, 2, \ldots, n\}$ and $j \in \{1, 2, \ldots, n\}$.

Now, $\operatorname{adj} A = \left((-1)^{i+j} \det\left(A_{\sim j,\sim i}\right)\right)_{1\leq i\leq n,\ 1\leq j\leq n}$, and thus the definition of the transpose of a matrix shows that

$$(\operatorname{adj} A)^T = \left(\underbrace{(-1)^{j+i}}_{=(-1)^{i+j}} \det\left(A_{\sim i,\sim j}\right)\right)_{1\leq i\leq n,\ 1\leq j\leq n} = \left((-1)^{i+j} \det\left(A_{\sim i,\sim j}\right)\right)_{1\leq i\leq n,\ 1\leq j\leq n}.$$

Compared with

$$\operatorname{adj}\left(A^T\right) = \left((-1)^{i+j} \underbrace{\det\left(\left(A^T\right)_{\sim j,\sim i}\right)}_{\substack{=\det\left(A_{\sim i,\sim j}\right) \\ \text{(by (441))}}}\right)_{1\leq i\leq n,\ 1\leq j\leq n}$$
$$\left(\text{by the definition of } \operatorname{adj}\left(A^T\right)\right)$$
$$= \left((-1)^{i+j} \det\left(A_{\sim i,\sim j}\right)\right)_{1\leq i\leq n,\ 1\leq j\leq n},$$

this yields $\operatorname{adj}\left(A^T\right) = (\operatorname{adj} A)^T$. This proves Proposition 6.99. □

The most important property of adjugates, however, is the following fact:

Theorem 6.100. Let $n \in \mathbb{N}$. Let A be an $n \times n$-matrix. Then,

$$A \cdot \operatorname{adj} A = \operatorname{adj} A \cdot A = \det A \cdot I_n.$$

(Recall that I_n denotes the $n \times n$ identity matrix. Expressions such as $\operatorname{adj} A \cdot A$ and $\det A \cdot I_n$ have to be understood as $(\operatorname{adj} A) \cdot A$ and $(\det A) \cdot I_n$, respectively.)

Example 6.101. Recall that the adjugate of a 2×2-matrix is given by the formula $\operatorname{adj}\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$. Thus, Theorem 6.100 (applied to $n = 2$) yields

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \cdot \begin{pmatrix} d & -b \\ -c & a \end{pmatrix} = \begin{pmatrix} d & -b \\ -c & a \end{pmatrix} \cdot \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \det \begin{pmatrix} a & b \\ c & d \end{pmatrix} \cdot I_2.$$

(Of course, $\det \begin{pmatrix} a & b \\ c & d \end{pmatrix} \cdot I_2 = (ad - bc) \cdot I_2 = \begin{pmatrix} ad - bc & 0 \\ 0 & ad - bc \end{pmatrix}$.)

Proof of Theorem 6.100. For any two objects i and j, we define $\delta_{i,j}$ to be the element $\begin{cases} 1, & \text{if } i = j; \\ 0, & \text{if } i \neq j \end{cases}$ of $\mathbb{K}$. Then, $I_n = \left(\delta_{i,j}\right)_{1 \leq i \leq n,\ 1 \leq j \leq n}$ (by the definition of I_n), and thus

$$\det A \cdot \underbrace{I_n}_{=\left(\delta_{i,j}\right)_{1 \leq i \leq n,\ 1 \leq j \leq n}} = \det A \cdot \left(\delta_{i,j}\right)_{1 \leq i \leq n,\ 1 \leq j \leq n} = \left(\det A \cdot \delta_{i,j}\right)_{1 \leq i \leq n,\ 1 \leq j \leq n}. \tag{442}$$

On the other hand, let us write the matrix A in the form $A = \left(a_{i,j}\right)_{1 \leq i \leq n,\ 1 \leq j \leq n}$.

Then, the definition of the product of two matrices shows that

$$\begin{aligned}
& A \cdot \operatorname{adj} A \\
& = \left(\sum_{k=1}^{n} a_{i,k} (-1)^{k+j} \det \left(A_{\sim j, \sim k} \right) \right)_{1 \leq i \leq n,\ 1 \leq j \leq n} \\
& \qquad \left(\begin{array}{c} \text{since } A = \left(a_{i,j} \right)_{1 \leq i \leq n,\ 1 \leq j \leq n} \\ \text{and } \operatorname{adj} A = \left((-1)^{i+j} \det \left(A_{\sim j, \sim i} \right) \right)_{1 \leq i \leq n,\ 1 \leq j \leq n} \end{array} \right) \\
& = \left(\sum_{q=1}^{n} \underbrace{a_{i,q} (-1)^{q+j}}_{=(-1)^{q+j} a_{i,q}} \det \left(A_{\sim j, \sim q} \right) \right)_{1 \leq i \leq n,\ 1 \leq j \leq n} \\
& \qquad \text{(here, we renamed the summation index } k \text{ as } q\text{)} \\
& = \left(\sum_{q=1}^{n} (-1)^{q+j} a_{i,q} \det \left(A_{\sim j, \sim q} \right) \right)_{1 \leq i \leq n,\ 1 \leq j \leq n} . \qquad (443)
\end{aligned}$$

Now, we claim that

$$\sum_{q=1}^{n} (-1)^{q+j} a_{i,q} \det \left(A_{\sim j, \sim q} \right) = \det A \cdot \delta_{i,j} \qquad (444)$$

for any $(i, j) \in \{1, 2, \ldots, n\}^2$.

[*Proof of (444):* Fix $(i, j) \in \{1, 2, \ldots, n\}^2$. We are in one of the following two cases:

Case 1: We have $i = j$.

Case 2: We have $i \neq j$.

Let us consider Case 1 first. In this case, we have $i = j$. Hence, $\delta_{i,j} = 1$. Now, Theorem 6.82 **(a)** (applied to $p = i$) yields

$$\det A = \sum_{q=1}^{n} \underbrace{(-1)^{i+q}}_{\substack{=(-1)^{q+i}=(-1)^{q+j} \\ \text{(since } i=j\text{)}}} a_{i,q} \det \left(\underbrace{A_{\sim i, \sim q}}_{\substack{=A_{\sim j, \sim q} \\ \text{(since } i=j\text{)}}} \right) = \sum_{q=1}^{n} (-1)^{q+j} a_{i,q} \det \left(A_{\sim j, \sim q} \right).$$

In view of $\det A \cdot \underbrace{\delta_{i,j}}_{=1} = \det A$, this rewrites as

$$\det A \cdot \delta_{i,j} = \sum_{q=1}^{n} (-1)^{q+j} a_{i,q} \det \left(A_{\sim j, \sim q} \right).$$

Thus, (444) is proven in Case 1.

Let us next consider Case 2. In this case, we have $i \neq j$. Hence, $\delta_{i,j} = 0$ and $j \neq i$. Now, Proposition 6.96 **(a)** (applied to $p = j$ and $r = i$) yields

$$0 = \sum_{q=1}^{n} \underbrace{(-1)^{j+q}}_{=(-1)^{q+j}} a_{i,q} \det\left(A_{\sim j,\sim q}\right) = \sum_{q=1}^{n} (-1)^{q+j} a_{i,q} \det\left(A_{\sim j,\sim q}\right).$$

In view of $\det A \cdot \underbrace{\delta_{i,j}}_{=0} = 0$, this rewrites as

$$\det A \cdot \delta_{i,j} = \sum_{q=1}^{n} (-1)^{q+j} a_{i,q} \det\left(A_{\sim j,\sim q}\right).$$

Thus, (444) is proven in Case 2.

We have now proven (444) in each of the two Cases 1 and 2. Thus, (444) is proven.]

Now, (443) becomes

$$A \cdot \operatorname{adj} A = \left(\underbrace{\sum_{q=1}^{n} (-1)^{q+j} a_{i,q} \det\left(A_{\sim j,\sim q}\right)}_{\substack{=\det A \cdot \delta_{i,j} \\ \text{(by (444))}}} \right)_{1\leq i\leq n,\ 1\leq j\leq n}$$
$$= \left(\det A \cdot \delta_{i,j}\right)_{1\leq i\leq n,\ 1\leq j\leq n} = \det A \cdot I_n \qquad (445)$$

(by (442)).

It now remains to prove that $\operatorname{adj} A \cdot A = \det A \cdot I_n$. One way to do this is by mimicking the above proof using Theorem 6.82 **(b)** and Proposition 6.96 **(b)** instead of Theorem 6.82 **(a)** and Proposition 6.96 **(a)**. However, here is a slicker proof:

Let us forget that we fixed A. We thus have shown that (445) holds for every $n \times n$-matrix A.

Now, let A be any $n \times n$-matrix. Then, we can apply (445) to A^T instead of A. We thus obtain

$$A^T \cdot \operatorname{adj}\left(A^T\right) = \underbrace{\det\left(A^T\right)}_{\substack{=\det A \\ \text{(by Exercise 6.4)}}} \cdot I_n = \det A \cdot I_n. \qquad (446)$$

Now, (349) (applied to $u = n$, $v = n$, $w = n$, $P = \operatorname{adj} A$ and $Q = A$) shows that

$$(\operatorname{adj} A \cdot A)^T = A^T \cdot \underbrace{(\operatorname{adj} A)^T}_{\substack{=\operatorname{adj}\left(A^T\right) \\ \text{(by Proposition 6.99)}}} = A^T \cdot \operatorname{adj}\left(A^T\right) = \det A \cdot I_n \qquad \text{(by (446))}.$$

Hence,

$$\Big(\underbrace{(\operatorname{adj} A \cdot A)^T}_{=\det A \cdot I_n}\Big)^T = (\det A \cdot I_n)^T = \det A \cdot \underbrace{(I_n)^T}_{\substack{=I_n \\ \text{(by (350), applied to } u=n)}}$$

$$\text{(by (351), applied to } u = n,\ v = n,\ P = I_n \text{ and } \lambda = \det A)$$

$$= \det A \cdot I_n .$$

Compared with

$$\left((\operatorname{adj} A \cdot A)^T\right)^T = \operatorname{adj} A \cdot A \qquad \text{(by (352), applied to } u = n,\ v = n \text{ and } P = \operatorname{adj} A \cdot A\text{)},$$

this yields $\operatorname{adj} A \cdot A = \det A \cdot I_n$. Combined with (445), this yields

$$A \cdot \operatorname{adj} A = \operatorname{adj} A \cdot A = \det A \cdot I_n .$$

This proves Theorem 6.100. □

The following is a simple consequence of Theorem 6.100:

Corollary 6.102. Let $n \in \mathbb{N}$. Let A be an $n \times n$-matrix. Let v be a column vector with n entries. If $Av = 0_{n\times 1}$, then $\det A \cdot v = 0_{n\times 1}$.

(Recall that $0_{n\times 1}$ denotes the $n \times 1$ zero matrix, i.e., the column vector with n entries whose all entries are 0.)

Proof of Corollary 6.102. Assume that $Av = 0_{n\times 1}$. It is easy to see that every $m \in \mathbb{N}$ and every $n \times m$-matrix B satisfy $I_n B = B$. Applying this to $m = 1$ and $B = v$, we obtain $I_n v = v$.

It is also easy to see that every $m \in \mathbb{N}$ and every $m \times n$-matrix B satisfy $B \cdot 0_{n\times 1} = 0_{m\times 1}$. Applying this to $m = n$ and $B = \operatorname{adj} A$, we obtain $\operatorname{adj} A \cdot 0_{n\times 1} = 0_{n\times 1}$.

Now, Theorem 6.100 yields $\operatorname{adj} A \cdot A = \det A \cdot I_n$. Hence,

$$\underbrace{(\operatorname{adj} A \cdot A)}_{=\det A \cdot I_n} v = (\det A \cdot I_n)\, v = \det A \cdot \underbrace{(I_n v)}_{=v} = \det A \cdot v.$$

Compared to

$$(\operatorname{adj} A \cdot A)\, v = \operatorname{adj} A \cdot \underbrace{(Av)}_{=0_{n\times 1}} \qquad \text{(since matrix multiplication is associative)}$$

$$= \operatorname{adj} A \cdot 0_{n\times 1} = 0_{n\times 1},$$

this yields $\det A \cdot v = 0_{n\times 1}$. This proves Corollary 6.102. □

Exercise 6.33. Let $n \in \mathbb{N}$. Let A and B be two $n \times n$-matrices. Prove that

$$\operatorname{adj}(AB) = \operatorname{adj} B \cdot \operatorname{adj} A.$$

Let me end this section with another application of Proposition 6.96:

Exercise 6.34. Let $n \in \mathbb{N}$. For every n elements $y_1, y_2, \ldots, y_n$ of $\mathbb{K}$, we define an element $V(y_1, y_2, \ldots, y_n)$ of $\mathbb{K}$ by

$$V(y_1, y_2, \ldots, y_n) = \prod_{1 \leq i < j \leq n} (y_i - y_j).$$

Let $x_1, x_2, \ldots, x_n$ be n elements of $\mathbb{K}$. Let $t \in \mathbb{K}$. Prove that

$$\sum_{k=1}^{n} x_k V(x_1, x_2, \ldots, x_{k-1}, x_k + t, x_{k+1}, x_{k+2}, \ldots, x_n) = \left(\binom{n}{2} t + \sum_{k=1}^{n} x_k\right) V(x_1, x_2, \ldots, x_n).$$

[**Hint:** Use Theorem 6.46, Laplace expansion, Proposition 6.96 and the binomial formula.]

Exercise 6.34 is part of [Fulton97, §4.3, Exercise 10].

6.16. Inverting matrices

We now will study inverses of matrices. We begin with a definition:

Definition 6.103. Let $n \in \mathbb{N}$ and $m \in \mathbb{N}$. Let A be an $n \times m$-matrix.

(a) A *left inverse* of A means an $m \times n$-matrix L such that $LA = I_m$. We say that the matrix A is *left-invertible* if and only if a left inverse of A exists.

(b) A *right inverse* of A means an $m \times n$-matrix R such that $AR = I_n$. We say that the matrix A is *right-invertible* if and only if a right inverse of A exists.

(c) An *inverse* of A (or *two-sided inverse* of A) means an $m \times n$-matrix B such that $BA = I_m$ and $AB = I_n$. We say that the matrix A is *invertible* if and only if an inverse of A exists.

The notions "left-invertible", "right-invertible" and "invertible" depend on the ring $\mathbb{K}$. We shall therefore speak of "left-invertible over $\mathbb{K}$", "right-invertible over $\mathbb{K}$" and "invertible over $\mathbb{K}$" whenever the context does not unambiguously determine $\mathbb{K}$.

The notions of "left inverse", "right inverse" and "inverse" are not interchangeable (unlike for elements in a commutative ring). We shall soon see in what cases they are identical; but first, let us give a few examples.

Example 6.104. For this example, set $\mathbb{K} = \mathbb{Z}$.

Let P be the 1×2-matrix $\begin{pmatrix} 1 & 2 \end{pmatrix}$. The matrix P is right-invertible. For instance, $\begin{pmatrix} -1 \\ 1 \end{pmatrix}$ and $\begin{pmatrix} 3 \\ -1 \end{pmatrix}$ are two right inverses of P (because $P\begin{pmatrix} -1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \end{pmatrix} = I_1$ and $P\begin{pmatrix} 3 \\ -1 \end{pmatrix} = \begin{pmatrix} 1 \end{pmatrix} = I_1$). This example shows that the right inverse of a matrix is not always unique.

The 2×1-matrix $P^T = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$ is left-invertible. The left inverses of P^T are the transposes of the right inverses of P.

The matrix P is not left-invertible; the matrix P^T is not right-invertible.

Let Q be the 2×2-matrix $\begin{pmatrix} 1 & -1 \\ 3 & -2 \end{pmatrix}$. The matrix Q is invertible. Its inverse is $\begin{pmatrix} -2 & 1 \\ -3 & 1 \end{pmatrix}$ (since $\begin{pmatrix} -2 & 1 \\ -3 & 1 \end{pmatrix} Q = I_2$ and $Q\begin{pmatrix} -2 & 1 \\ -3 & 1 \end{pmatrix} = I_2$). It is not hard to see that this is its only inverse.

Let R be the 2×2-matrix $\begin{pmatrix} 1 & 2 \\ 2 & -1 \end{pmatrix}$. It can be seen that this matrix is not invertible **as a matrix over** $\mathbb{Z}$. On the other hand, if we consider it as a matrix over $\mathbb{K} = \mathbb{Q}$ instead, then it is invertible, with inverse $\begin{pmatrix} 1/5 & 2/5 \\ 2/5 & -1/5 \end{pmatrix}$.

Of course, any inverse of a matrix A is automatically both a left inverse of A and a right inverse of A. Thus, an invertible matrix A is automatically both left-invertible and right-invertible.

The following simple fact is an analogue of Proposition 6.65:

Proposition 6.105. Let $n \in \mathbb{N}$ and $m \in \mathbb{N}$. Let A be an $n \times m$-matrix. Let L be a left inverse of A. Let R be a right inverse of A.

(a) We have $L = R$.

(b) The matrix A is invertible, and $L = R$ is an inverse of A.

Proof of Proposition 6.105. We know that L is a left inverse of A. In other words, L is an $m \times n$-matrix such that $LA = I_m$ (by the definition of a "left inverse").

We know that R is a right inverse of A. In other words, R is an $m \times n$-matrix such that $AR = I_n$ (by the definition of a "right inverse").

Now, recall that $I_m G = G$ for every $k \in \mathbb{N}$ and every $m \times k$-matrix G. Applying this to $k = n$ and $G = R$, we obtain $I_m R = R$.

Also, recall that $GI_n = G$ for every $k \in \mathbb{N}$ and every $k \times n$-matrix G. Applying this to $k = m$ and $G = L$, we obtain $LI_n = L$. Thus, $L = L \underbrace{I_n}_{=AR} = \underbrace{LA}_{=I_m} R = I_m R = R$. This proves Proposition 6.105 **(a)**.

(b) We have $LA = I_m$ and $A \underbrace{L}_{=R} = AR = I_n$. Thus, L is an $m \times n$-matrix such that $LA = I_m$ and $AL = I_n$. In other words, L is an inverse of A (by the definition of an "inverse"). Thus, $L = R$ is an inverse of A (since $L = R$). This proves Proposition 6.105 **(b)**. □

Corollary 6.106. Let $n \in \mathbb{N}$ and $m \in \mathbb{N}$. Let A be an $n \times m$-matrix.

(a) If A is left-invertible and right-invertible, then A is invertible.

(b) If A is invertible, then there exists exactly one inverse of A.

Proof of Corollary 6.106. **(a)** Assume that A is left-invertible and right-invertible. Thus, A has a left inverse L (since A is left-invertible). Consider this L. Also, A has a right inverse R (since A is right-invertible). Consider this R. Proposition 6.105 **(b)** yields that the matrix A is invertible, and $L = R$ is an inverse of A. Corollary 6.106 **(a)** is proven.

(b) Assume that A is invertible. Let B and B' be any two inverses of A. Since B is an inverse of A, we know that B is an $m \times n$-matrix such that $BA = I_m$ and $AB = I_n$ (by the definition of an "inverse"). Thus, in particular, B is an $m \times n$-matrix such that $BA = I_m$. In other words, B is a left inverse of A. Since B' is an inverse of A, we know that B' is an $m \times n$-matrix such that $B'A = I_m$ and $AB' = I_n$ (by the definition of an "inverse"). Thus, in particular, B' is an $m \times n$-matrix such that $AB' = I_n$. In other words, B' is a right inverse of A. Now, Proposition 6.105 **(a)** (applied to $L = B$ and $R = B'$) shows that $B = B'$.

Let us now forget that we fixed B and B'. We thus have shown that if B and B' are two inverses of A, then $B = B'$. In other words, any two inverses of A are equal. In other words, there exists at most one inverse of A. Since we also know that there exists at least one inverse of A (since A is invertible), we thus conclude that there exists exactly one inverse of A. This proves Corollary 6.106 **(b)**. □

Definition 6.107. Let $n \in \mathbb{N}$ and $m \in \mathbb{N}$. Let A be an invertible $n \times m$-matrix. Corollary 6.106 **(b)** shows that there exists exactly one inverse of A. Thus, we can speak of "*the inverse of* A". We denote this inverse by A^{-1}.

In contrast to Definition 6.66, we do **not** define the notation B/A for two matrices B and A for which A is invertible. In fact, the trouble with such a notation would be its ambiguity: should it mean BA^{-1} or $A^{-1}B$? (In general, BA^{-1} and $A^{-1}B$ are not the same.) Some authors do write B/A for the matrices BA^{-1} and $A^{-1}B$ when these matrices are equal; but we shall not have a reason to do so.

Remark 6.108. Let $n \in \mathbb{N}$ and $m \in \mathbb{N}$. Let A be an invertible $n \times m$-matrix. Then, the inverse A^{-1} of A is an $m \times n$-matrix and satisfies $AA^{-1} = I_n$ and $A^{-1}A = I_m$. This follows from the definition of the inverse of A; we are just stating it once again, because it will later be used without mention.

Example 6.104 (and your experiences with a linear algebra class, if you have taken one) suggest the conjecture that only square matrices can be invertible. Indeed, this is **almost** true. There is a stupid counterexample: If $\mathbb{K}$ is a trivial ring, then every matrix over $\mathbb{K}$ is invertible[254]. It turns out that this is the only case where nonsquare matrices can be invertible. Indeed, we have the following:

Theorem 6.109. Let $n \in \mathbb{N}$ and $m \in \mathbb{N}$. Let A be an $n \times m$-matrix.

(a) If A is left-invertible and if $n < m$, then $\mathbb{K}$ is a trivial ring.

(b) If A is right-invertible and if $n > m$, then $\mathbb{K}$ is a trivial ring.

(c) If A is invertible and if $n \neq m$, then $\mathbb{K}$ is a trivial ring.

Proof of Theorem 6.109. **(a)** Assume that A is left-invertible, and that $n < m$.

The matrix A has a left inverse L (since it is left-invertible). Consider this L.

We know that L is a left inverse of A. In other words, L is an $m \times n$-matrix such that $LA = I_m$ (by the definition of a "left inverse"). But (380) (applied to m, n, L and A instead of n, m, A and B) yields $\det(LA) = 0$ (since $n < m$). Thus, $0 = \det\left(\underbrace{LA}_{=I_m}\right) = \det(I_m) = 1$. Of course, the 0 and the 1 in this equality mean the elements $0_{\mathbb{K}}$ and $1_{\mathbb{K}}$ of $\mathbb{K}$ (rather than the integers 0 and 1); thus, it rewrites as $0_{\mathbb{K}} = 1_{\mathbb{K}}$. In other words, $\mathbb{K}$ is a trivial ring. This proves Theorem 6.109 **(a)**.

(b) Assume that A is right-invertible, and that $n > m$.

The matrix A has a right inverse R (since it is right-invertible). Consider this R.

We know that R is a right inverse of A. In other words, R is an $m \times n$-matrix such that $AR = I_n$ (by the definition of a "right inverse"). But (380) (applied to $B = R$) yields $\det(AR) = 0$ (since $m < n$). Thus, $0 = \det\left(\underbrace{AR}_{=I_n}\right) = \det(I_n) = 1$. Of course, the 0 and the 1 in this equality mean the elements $0_{\mathbb{K}}$ and $1_{\mathbb{K}}$ of $\mathbb{K}$ (rather than the integers 0 and 1); thus, it rewrites as $0_{\mathbb{K}} = 1_{\mathbb{K}}$. In other words, $\mathbb{K}$ is a trivial ring. This proves Theorem 6.109 **(b)**.

(c) Assume that A is invertible, and that $n \neq m$. Since $n \neq m$, we must be in one of the following two cases:

Case 1: We have $n < m$.

Case 2: We have $n > m$.

Let us first consider Case 1. In this case, we have $n < m$. Now, A is invertible, and thus left-invertible (since every invertible matrix is left-invertible). Hence, $\mathbb{K}$ is a trivial ring (according to Theorem 6.109 **(a)**). Thus, Theorem 6.109 **(c)** is proven in Case 1.

[254] For example, the 1×2-matrix $\begin{pmatrix} 0_{\mathbb{K}} & 0_{\mathbb{K}} \end{pmatrix}$ over a trivial ring $\mathbb{K}$ is invertible, having inverse $\begin{pmatrix} 0_{\mathbb{K}} \\ 0_{\mathbb{K}} \end{pmatrix}$. If you don't believe me, just check that

$$\begin{pmatrix} 0_{\mathbb{K}} \\ 0_{\mathbb{K}} \end{pmatrix} \begin{pmatrix} 0_{\mathbb{K}} & 0_{\mathbb{K}} \end{pmatrix} = \begin{pmatrix} 0_{\mathbb{K}} & 0_{\mathbb{K}} \\ 0_{\mathbb{K}} & 0_{\mathbb{K}} \end{pmatrix} = \begin{pmatrix} 1_{\mathbb{K}} & 0_{\mathbb{K}} \\ 0_{\mathbb{K}} & 1_{\mathbb{K}} \end{pmatrix} \qquad (\text{since } 0_{\mathbb{K}} = 1_{\mathbb{K}})$$
$$= I_2$$

and $\begin{pmatrix} 0_{\mathbb{K}} & 0_{\mathbb{K}} \end{pmatrix} \begin{pmatrix} 0_{\mathbb{K}} \\ 0_{\mathbb{K}} \end{pmatrix} = \begin{pmatrix} 0_{\mathbb{K}} \end{pmatrix} = \begin{pmatrix} 1_{\mathbb{K}} \end{pmatrix} = I_1$.

Let us now consider Case 2. In this case, we have $n > m$. Now, A is invertible, and thus right-invertible (since every invertible matrix is right-invertible). Hence, $\mathbb{K}$ is a trivial ring (according to Theorem 6.109 **(b)**). Thus, Theorem 6.109 **(c)** is proven in Case 2.

We have thus proven Theorem 6.109 **(c)** in both Cases 1 and 2. Thus, Theorem 6.109 **(c)** always holds. □

Theorem 6.109 **(c)** says that the question whether a matrix is invertible is only interesting for square matrices, unless the ring $\mathbb{K}$ is given so inexplicitly that we do not know whether it is trivial or not[255]. Let us now study the invertibility of a square matrix. Here, the determinant turns out to be highly useful:

Theorem 6.110. Let $n \in \mathbb{N}$. Let A be an $n \times n$-matrix.

(a) The matrix A is invertible if and only if the element $\det A$ of $\mathbb{K}$ is invertible (in $\mathbb{K}$).

(b) If $\det A$ is invertible, then the inverse of A is $A^{-1} = \dfrac{1}{\det A} \cdot \operatorname{adj} A$.

When $\mathbb{K}$ is a field, the invertible elements of $\mathbb{K}$ are precisely the nonzero elements of $\mathbb{K}$. Thus, when $\mathbb{K}$ is a field, the statement of Theorem 6.110 **(a)** can be rewritten as "The matrix A is invertible if and only if $\det A \neq 0$"; this is a cornerstone of linear algebra. But our statement of Theorem 6.110 **(a)** works for an arbitrary commutative ring $\mathbb{K}$. In particular, it works for $\mathbb{K} = \mathbb{Z}$. Here is a consequence:

Corollary 6.111. Let $n \in \mathbb{N}$. Let $A \in \mathbb{Z}^{n \times n}$ be an $n \times n$-matrix over $\mathbb{Z}$. Then, the matrix A is invertible if and only if $\det A \in \{1, -1\}$.

Proof of Corollary 6.111. If g is an integer, then g is invertible (in $\mathbb{Z}$) if and only if $g \in \{1, -1\}$. In other words, for every integer g, we have the following equivalence:

$$(g \text{ is invertible (in } \mathbb{Z})) \iff (g \in \{1, -1\}) . \tag{447}$$

Now, Theorem 6.110 **(a)** (applied to $\mathbb{K} = \mathbb{Z}$) yields that the matrix A is invertible if and only if the element $\det A$ of $\mathbb{Z}$ is invertible (in $\mathbb{Z}$). Thus, we have the following chain of equivalences:

$$\begin{aligned} & (\text{the matrix } A \text{ is invertible}) \\ \iff & (\det A \text{ is invertible (in } \mathbb{Z})) \iff (\det A \in \{1, -1\}) \\ & (\text{by (447), applied to } g = \det A) . \end{aligned}$$

This proves Corollary 6.111. □

[255] This actually happens rather often in algebra! For example, rings are often defined by "generators and relations" (such as "the ring with commuting generators a, b, c subject to the relations $a^2 + b^2 = c^2$ and $ab = c$"). Sometimes the relations force the ring to become trivial (for instance, the ring with generator a and relations $a = 1$ and $a^2 = 2$ is clearly the trivial ring, because in this ring we have $2 = a^2 = 1^2 = 1$). Often this is not clear a-priori, and theorems such as Theorem 6.109 can be used to show this. The triviality of a ring can be a nontrivial statement! (Richman makes this point in [Richma88].)

Notice that Theorem 6.110 **(b)** yields an explicit way to compute the inverse of a square matrix A (provided that we can compute determinants and the inverse of $\det A$). This is not the fastest way (at least not when $\mathbb{K}$ is a field), but it is useful for various theoretical purposes.

Proof of Theorem 6.110. **(a)** $\Longrightarrow$:[256] Assume that the matrix A is invertible. In other words, an inverse B of A exists. Consider such a B.

The matrix B is an inverse of A. In other words, B is an $n \times n$-matrix such that $BA = I_n$ and $AB = I_n$ (by the definition of an "inverse"). Theorem 6.23 yields $\det(AB) = \det A \cdot \det B$, so that $\det A \cdot \det B = \det\left(\underbrace{AB}_{=I_n}\right) = \det(I_n) = 1$. Of course, we also have $\det B \cdot \det A = \det A \cdot \det B = 1$. Thus, $\det B$ is an inverse of $\det A$ in $\mathbb{K}$. Therefore, the element $\det A$ is invertible (in $\mathbb{K}$). This proves the $\Longrightarrow$ direction of Theorem 6.110 **(a)**.

$\Longleftarrow$: Assume that the element $\det A$ is invertible (in $\mathbb{K}$). Thus, its inverse $\dfrac{1}{\det A}$ exists. Theorem 6.100 yields

$$A \cdot \operatorname{adj} A = \operatorname{adj} A \cdot A = \det A \cdot I_n.$$

Now, define an $n \times n$-matrix B by $B = \dfrac{1}{\det A} \cdot \operatorname{adj} A$. Then,

$$A \underbrace{B}_{=\frac{1}{\det A}\cdot \operatorname{adj} A} = A \cdot \left(\frac{1}{\det A} \cdot \operatorname{adj} A\right) = \frac{1}{\det A} \cdot \underbrace{A \cdot \operatorname{adj} A}_{=\det A \cdot I_n} = \underbrace{\frac{1}{\det A} \cdot \det A}_{=1} \cdot I_n = I_n$$

and

$$\underbrace{B}_{=\frac{1}{\det A}\cdot \operatorname{adj} A} A = \frac{1}{\det A} \cdot \underbrace{\operatorname{adj} A \cdot A}_{=\det A \cdot I_n} = \underbrace{\frac{1}{\det A} \cdot \det A}_{=1} \cdot I_n = I_n.$$

Thus, B is an $n \times n$-matrix such that $BA = I_n$ and $AB = I_n$. In other words, B is an inverse of A (by the definition of an "inverse"). Thus, an inverse of A exists; in other words, the matrix A is invertible. This proves the $\Longleftarrow$ direction of Theorem 6.110 **(a)**.

[256] In case you don't know what the notation "$\Longrightarrow$:" here means:

Theorem 6.110 **(a)** is an "if and only if" assertion. In other words, it asserts that $\mathcal{U} \Longleftrightarrow \mathcal{V}$ for two statements $\mathcal{U}$ and $\mathcal{V}$. (In our case, $\mathcal{U}$ is the statement "the matrix A is invertible", and $\mathcal{V}$ is the statement "the element $\det A$ of $\mathbb{K}$ is invertible (in $\mathbb{K}$)".) In order to prove a statement of the form $\mathcal{U} \Longleftrightarrow \mathcal{V}$, it is sufficient to prove the implications $\mathcal{U} \Longrightarrow \mathcal{V}$ and $\mathcal{U} \Longleftarrow \mathcal{V}$. Usually, these two implications are proven separately (although not always; for instance, in the proof of Corollary 6.111, we have used a chain of equivalences to prove $\mathcal{U} \Longleftrightarrow \mathcal{V}$ directly). When writing such a proof, one often uses the abbreviations "$\Longrightarrow$:" and "$\Longleftarrow$:" for "Here comes the proof of the implication $\mathcal{U} \Longrightarrow \mathcal{V}$:" and "Here comes the proof of the implication $\mathcal{U} \Longleftarrow \mathcal{V}$:", respectively.

We have now proven both directions of Theorem 6.110 **(a)**. Theorem 6.110 **(a)** is thus proven.

(b) Assume that $\det A$ is invertible. Thus, its inverse $\dfrac{1}{\det A}$ exists. We define an $n \times n$-matrix B by $B = \dfrac{1}{\det A} \cdot \operatorname{adj} A$. Then, B is an inverse of A [257]. In other words, B is **the** inverse of A. In other words, $B = A^{-1}$. Hence, $A^{-1} = B = \dfrac{1}{\det A} \cdot \operatorname{adj} A$. This proves Theorem 6.110 **(b)**. □

Corollary 6.112. Let $n \in \mathbb{N}$. Let A and B be two $n \times n$-matrices such that $AB = I_n$.

(a) We have $BA = I_n$.

(b) The matrix A is invertible, and the matrix B is the inverse of A.

Proof of Corollary 6.112. Theorem 6.23 yields $\det(AB) = \det A \cdot \det B$, so that $\det A \cdot \det B = \det \left(\underbrace{AB}_{=I_n} \right) = \det(I_n) = 1$. Of course, we also have $\det B \cdot \det A = \det A \cdot \det B = 1$. Thus, $\det B$ is an inverse of $\det A$ in $\mathbb{K}$. Therefore, the element $\det A$ is invertible (in $\mathbb{K}$). Therefore, the matrix A is invertible (according to the $\Longleftarrow$ direction of Theorem 6.110 **(b)**). Thus, the inverse of A exists. Let C be this inverse. Thus, C is a left inverse of A (since every inverse of A is a left inverse of A).

The matrix B is an $n \times n$-matrix satisfying $AB = I_n$. In other words, B is a right inverse of A. On the other hand, C is a left inverse of A. Hence, Proposition 6.105 **(a)** (applied to $L = C$ and $R = B$) yields $C = B$. Hence, the matrix B is the inverse of A (since the matrix C is the inverse of A). Thus, Corollary 6.112 **(b)** is proven.

Since B is the inverse of A, we have $BA = I_n$ and $AB = I_n$ (by the definition of an "inverse"). This proves Corollary 6.112 **(a)**. □

Remark 6.113. Corollary 6.112 is **not** obvious! Matrix multiplication, in general, is not commutative (we have $AB \neq BA$ more often than not), and there is no reason to expect that $AB = I_n$ implies $BA = I_n$. The fact that this is nevertheless true for square matrices took us quite some work to prove (we needed, among other things, the notion of an adjugate). This fact would **not** hold for rectangular matrices. Nor does it hold for "infinite square matrices": Without wanting to go into the details of how products of infinite matrices are defined, I invite you to check that the two infinite matrices $A = \begin{pmatrix} 0 & 1 & 0 & \cdots \\ 0 & 0 & 1 & \cdots \\ 0 & 0 & 0 & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}$ and $B = A^T =$

[257] We have shown this in our proof of the $\Longleftarrow$ direction of Theorem 6.110 **(a)**.

$\begin{pmatrix} 0 & 0 & 0 & \cdots \\ 1 & 0 & 0 & \cdots \\ 0 & 1 & 0 & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}$ satisfy $AB = I_\infty$ but $BA \neq I_\infty$. This makes Corollary 6.112 **(a)** all the more interesting.

Here are some more exercises involving matrices in "block-matrix form":

Exercise 6.35. Let $n \in \mathbb{N}$ and $m \in \mathbb{N}$. Let $A \in \mathbb{K}^{n\times n}$, $B \in \mathbb{K}^{n\times m}$, $C \in \mathbb{K}^{m\times n}$ and $D \in \mathbb{K}^{m\times m}$. Furthermore, let $W \in \mathbb{K}^{m\times m}$ and $V \in \mathbb{K}^{m\times n}$ be such that $VA = -WC$. Prove that

$$\det W \cdot \det \begin{pmatrix} A & B \\ C & D \end{pmatrix} = \det A \cdot \det (VB + WD).$$

[**Hint:** Use Exercise 6.28 to simplify the product $\begin{pmatrix} I_n & 0_{n\times m} \\ V & W \end{pmatrix} \begin{pmatrix} A & B \\ C & D \end{pmatrix}$; then, take determinants.]

Exercise 6.35 can often be used to compute the determinant of a matrix given in block-matrix form (i.e., determinants of the form $\det \begin{pmatrix} A & B \\ C & D \end{pmatrix}$) by only computing determinants of smaller matrices (such as W, A and $VB + WD$). It falls short of providing a general method for computing such determinants[258], but it is one of the most general facts about them. The next two exercises are two special cases of Exercise 6.35:

Exercise 6.36. Let $n \in \mathbb{N}$ and $m \in \mathbb{N}$. Let $A \in \mathbb{K}^{n\times n}$, $B \in \mathbb{K}^{n\times m}$, $C \in \mathbb{K}^{m\times n}$ and $D \in \mathbb{K}^{m\times m}$ be such that the matrix A is invertible. Prove that

$$\det \begin{pmatrix} A & B \\ C & D \end{pmatrix} = \det A \cdot \det \left(D - CA^{-1}B\right).$$

Exercise 6.36 is known as the *Schur complement formula* (or, at least, it is one of several formulas sharing this name); and the matrix $D - CA^{-1}B$ appearing on its right hand side is known as the *Schur complement* of the block A in the matrix $\begin{pmatrix} A & B \\ C & D \end{pmatrix}$.

[258] Indeed, it only gives a formula for $\det W \cdot \det \begin{pmatrix} A & B \\ C & D \end{pmatrix}$, not for $\det \begin{pmatrix} A & B \\ C & D \end{pmatrix}$. If $\det W$ is invertible, then it allows for computing $\det \begin{pmatrix} A & B \\ C & D \end{pmatrix}$; but Exercise 6.35 gives no hint on how to find matrices W and V such that $\det W$ is invertible and such that $VA = -WC$. (Actually, such matrices do not always exist!)

Exercise 6.37. Let $n \in \mathbb{N}$ and $m \in \mathbb{N}$. Let $A \in \mathbb{K}^{n \times n}$, $B \in \mathbb{K}^{n \times m}$, $C \in \mathbb{K}^{m \times n}$ and $D \in \mathbb{K}^{m \times m}$. Let $A' \in \mathbb{K}^{n \times n}$, $B' \in \mathbb{K}^{n \times m}$, $C' \in \mathbb{K}^{m \times n}$ and $D' \in \mathbb{K}^{m \times m}$. Assume that the matrix $\left(\begin{array}{cc} A & B \\ C & D \end{array}\right)$ is invertible, and that its inverse is the matrix $\left(\begin{array}{cc} A' & B' \\ C' & D' \end{array}\right)$. Prove that

$$\det A = \det \left(\begin{array}{cc} A & B \\ C & D \end{array}\right) \cdot \det \left(D'\right).$$

Exercise 6.37 can be rewritten in the following more handy form:

Exercise 6.38. We shall use the notations introduced in Definition 6.78.

Let $n \in \mathbb{N}$. Let $A \in \mathbb{K}^{n \times n}$ be an invertible matrix. Let $k \in \{0, 1, \ldots, n\}$. Prove that

$$\det \left(\operatorname{sub}_{1,2,\ldots,k}^{1,2,\ldots,k} A\right) = \det A \cdot \det \left(\operatorname{sub}_{k+1,k+2,\ldots,n}^{k+1,k+2,\ldots,n} \left(A^{-1}\right)\right).$$

Exercise 6.38 is a particular case of the so-called *Jacobi complementary minor theorem* (Exercise 6.56 further below).

6.17. Noncommutative rings

I think that here is a good place to introduce two other basic notions from algebra: that of a noncommutative ring, and that of a group.

Definition 6.114. The notion of a *noncommutative ring* is defined in the same way as we have defined a commutative ring (in Definition 6.2), except that we no longer require the "Commutativity of multiplication" axiom.

As I have already said, the word "noncommutative" (in "noncommutative ring") does not mean that commutativity of multiplication has to be false in this ring; it only means that commutativity of multiplication is not required. Thus, every commutative ring is a noncommutative ring. Therefore, each of the examples of a commutative ring given in Section 6.1 is also an example of a noncommutative ring. Of course, it is more interesting to see some examples of noncommutative rings which actually fail to obey commutativity of multiplication. Here are some of these examples:

- If $n \in \mathbb{N}$ and if $\mathbb{K}$ is a commutative ring, then the set $\mathbb{K}^{n \times n}$ of matrices becomes a noncommutative ring (when endowed with the addition and multiplication of matrices, with the zero $0_{n \times n}$ and with the unity I_n). This is actually a commutative ring when $\mathbb{K}$ is trivial or when $n \leq 1$, but in all "interesting" cases it is not commutative.

- If you have heard of the quaternions, you should realize that they form a noncommutative ring.

- Given a commutative ring $\mathbb{K}$ and n distinct symbols $X_1, X_2, \ldots, X_n$, we can define a *ring of polynomials in the* ***noncommutative*** *variables* $X_1, X_2, \ldots, X_n$ over $\mathbb{K}$. We do not want to go into the details of its definition at this point, but let us just mention some examples of its elements: For instance, the ring of polynomials in the noncommutative variables X and Y over $\mathbb{Q}$ contains elements such as $1 + \dfrac{2}{3}X$, $X^2 + \dfrac{3}{2}Y - 7XY + YX$, $2XY$, $2YX$ and $5X^2Y - 6XYX + 7Y^2X$ (and of course, the elements XY and YX are not equal).

- If $n \in \mathbb{N}$ and if $\mathbb{K}$ is a commutative ring, then the set of all lower-triangular $n \times n$-matrices over $\mathbb{K}$ becomes a noncommutative ring (with addition, multiplication, zero and unity defined in the same way as in $\mathbb{K}^{n \times n}$). This is because the sum and the product of any two lower-triangular $n \times n$-matrices over $\mathbb{K}$ are again lower-triangular[259], and because the matrices $0_{n \times n}$ and I_n are lower-triangular.

- In contrast, the set of all invertible 2×2-matrices over $\mathbb{K}$ is **not** a noncommutative ring (for example, because the sum of the two invertible matrices I_2 and $-I_2$ is not invertible[260]).

- If $\mathbb{K}$ is a commutative ring, then the set of all 3×3-matrices (over $\mathbb{K}$) of the form $\begin{pmatrix} a & b & c \\ 0 & d & 0 \\ 0 & 0 & f \end{pmatrix}$ (with $a, b, c, d, f \in \mathbb{K}$) is a noncommutative ring (again, with the same addition, multiplication, zero and unity as for $\mathbb{K}^{n \times n}$). [261]

- On the other hand, if $\mathbb{K}$ is a commutative ring, then the set of all 3×3-matrices (over $\mathbb{K}$) of the form $\begin{pmatrix} a & b & 0 \\ 0 & c & d \\ 0 & 0 & f \end{pmatrix}$ (with $a, b, c, d, f \in \mathbb{K}$) is **not** a noncommutative ring (unless $\mathbb{K}$ is trivial), because products of matrices in

[259] Check this! (For the sum, it is clear, but for the product, it is an instructive exercise.)

[260] unless the ring $\mathbb{K}$ is trivial

[261] To check this, one needs to prove that the matrices $0_{3 \times 3}$ and I_3 have this form, and that the sum and the product of any two matrices of this form is again a matrix of this form. All of this is clear, except for the claim about the product. The latter claim follows from the computation

$$\begin{pmatrix} a & b & c \\ 0 & d & 0 \\ 0 & 0 & f \end{pmatrix} \begin{pmatrix} a' & b' & c' \\ 0 & d' & 0 \\ 0 & 0 & f' \end{pmatrix} = \begin{pmatrix} aa' & bd' + ab' & cf' + ac' \\ 0 & dd' & 0 \\ 0 & 0 & ff' \end{pmatrix}.$$

this set are not always in this set[262].

For the rest of this section, we let $\mathbb{L}$ be a **noncommutative** ring. What can we do with elements of $\mathbb{L}$? We can do some of the things that we can do with a commutative ring, but not all of them. For example, we can still define the sum $a_1 + a_2 + \cdots + a_n$ and the product $a_1 a_2 \cdots a_n$ of n elements of a noncommutative ring. But we cannot arbitrarily reorder the factors of a product and expect to always get the same result! (With a sum, we can do this.) We can still define na for any $n \in \mathbb{Z}$ and $a \in \mathbb{L}$ (in the same way as we defined na for $n \in \mathbb{Z}$ and $a \in \mathbb{K}$ when $\mathbb{K}$ was a commutative ring). We can still define a^n for any $n \in \mathbb{N}$ and $a \in \mathbb{L}$ (again, in the same fashion as for commutative rings). The identities (327), (331), (332), (333), (334), (335), (336) and (337) still hold when the commutative ring $\mathbb{K}$ is replaced by the noncommutative ring $\mathbb{L}$; but the identities (338) and (339) may not (although they **do** hold if we additionally assume that $ab = ba$). Finite sums such as $\sum\limits_{s \in S} a_s$ (where S is a finite set, and $a_s \in \mathbb{L}$ for every $s \in S$) are well-defined, but finite products such as $\prod\limits_{s \in S} a_s$ are not (unless we specify the order in which their factors are to be multiplied).

We can define matrices over $\mathbb{L}$ in the same way as we have defined matrices over $\mathbb{K}$. We can even define the determinant of a square matrix over $\mathbb{L}$ using the formula (341); however, this determinant lacks many of the important properties that determinants over $\mathbb{K}$ have (for instance, it satisfies neither Exercise 6.4 nor Theorem 6.23), and is therefore usually not studied.[263]

We define the notion of an *inverse* of an element $a \in \mathbb{L}$; in order to do so, we simply replace $\mathbb{K}$ by $\mathbb{L}$ in Definition 6.64. (Now it suddenly matters that we required both $ab = 1$ and $ba = 1$ in Definition 6.64.) Proposition 6.65 still holds (and its proof still works) when $\mathbb{K}$ is replaced by $\mathbb{L}$.

We define the notion of an *invertible element* of $\mathbb{L}$; in order to do so, we simply replace $\mathbb{K}$ by $\mathbb{L}$ in Definition 6.66 **(a)**. We cannot directly replace $\mathbb{K}$ by $\mathbb{L}$ in Definition 6.66 **(b)**, because for two invertible elements a and b of $\mathbb{L}$ we do not necessarily have $(ab)^{-1} = a^{-1}b^{-1}$; but something very similar holds (namely, $(ab)^{-1} = b^{-1}a^{-1}$). Trying to generalize Definition 6.66 **(c)** to noncommutative rings is rather hopeless: In general, we cannot bring a "noncommutative fraction" of the form $ba^{-1} + dc^{-1}$ to a "common denominator".

Example 6.115. Let $\mathbb{K}$ be a commutative ring. Let $n \in \mathbb{N}$. As we know, $\mathbb{K}^{n \times n}$ is a noncommutative ring. The invertible elements of this ring are exactly the invertible $n \times n$-matrices. (To see this, just compare the definition of an invert-

[262] Indeed, $\begin{pmatrix} a & b & 0 \\ 0 & c & d \\ 0 & 0 & f \end{pmatrix} \begin{pmatrix} a' & b' & 0 \\ 0 & c' & d' \\ 0 & 0 & f' \end{pmatrix} = \begin{pmatrix} aa' & ab' + bc' & bd' \\ 0 & cc' & cd' + df' \\ 0 & 0 & ff' \end{pmatrix}$ can have $bd' \neq 0$.

[263] Some algebraists have come up with subtler notions of determinants for matrices over noncommutative rings. But I don't want to go in that direction here.

ible element of $\mathbb{K}^{n\times n}$ with the definition of an invertible $n \times n$-matrix. These definitions are clearly equivalent.)

6.18. Groups, and the group of units

Let me finally define the notion of a *group*.

Definition 6.116. A *group* means a set G endowed with

- a binary operation called "multiplication" (or "composition", or just "binary operation"), and denoted by $\cdot$, and written infix, and
- an element called 1_G (or e_G)

such that the following axioms are satisfied:

- *Associativity:* We have $a(bc) = (ab)c$ for all $a \in G$, $b \in G$ and $c \in G$. Here and in the following, the expression "ab" is shorthand for "$a \cdot b$" (as is usual for products of numbers).
- *Neutrality of* 1*:* We have $a1_G = 1_G a = a$ for all $a \in G$.
- *Existence of inverses:* For every $a \in G$, there exists an element $a' \in G$ such that $aa' = a'a = 1_G$. This a' is commonly denoted by a^{-1} and called the *inverse* of a. (It is easy to check that it is unique.)

Definition 6.117. The element 1_G of a group G is denoted the *neutral element* (or the *identity*) of G.

The binary operation $\cdot$ in Definition 6.116 is usually not identical with the binary operation $\cdot$ on the set of integers, and is denoted by $\cdot_G$ when confusion can arise.

The definition of a group has similarities with that of a noncommutative ring. Viewed from a distance, it may look as if a noncommutative ring would "consist" of two groups with the same underlying set. This is not quite correct, though, because the multiplication in a nontrivial ring does not satisfy the "existence of inverses" axiom. But it is true that there are two groups in every noncommutative ring:

Proposition 6.118. Let $\mathbb{L}$ be a noncommutative ring.

(a) The set $\mathbb{L}$, endowed with the **addition** $+_{\mathbb{L}}$ (as multiplication) and the element $0_{\mathbb{L}}$ (as neutral element), is a group. This group is called the *additive group* of $\mathbb{L}$, and denoted by $\mathbb{L}^{+}$.

(b) Let $\mathbb{L}^{\times}$ denote the set of all invertible elements of $\mathbb{L}$. Then, the product of two elements of $\mathbb{L}^{\times}$ again belongs to $\mathbb{L}^{\times}$. Thus, we can define a binary operation $\cdot_{\mathbb{L}^{\times}}$ on the set $\mathbb{L}^{\times}$ (written infix) by

$$a \cdot_{\mathbb{L}^{\times}} b = ab \qquad \text{for all } a \in \mathbb{L}^{\times} \text{ and } b \in \mathbb{L}^{\times}.$$

The set $\mathbb{L}^{\times}$, endowed with the multiplication $\cdot_{\mathbb{L}^{\times}}$ (as multiplication) and the element $1_{\mathbb{L}}$ (as neutral element), is a group. This group is called the *group of units* of $\mathbb{L}$.

Proof of Proposition 6.118. **(a)** The addition $+_{\mathbb{L}}$ is clearly a binary operation on $\mathbb{L}$, and the element $0_{\mathbb{L}}$ is clearly an element of $\mathbb{L}$. The three axioms in Definition 6.116 are clearly satisfied for the binary operation $+_{\mathbb{L}}$ and the element $0_{\mathbb{L}}$ [264]. Therefore, the set $\mathbb{L}$, endowed with the addition $+_{\mathbb{L}}$ (as multiplication) and the element $0_{\mathbb{L}}$ (as neutral element), is a group. This proves Proposition 6.118 **(a)**.

(b) If $a \in \mathbb{L}^{\times}$ and $b \in \mathbb{L}^{\times}$, then $ab \in \mathbb{L}^{\times}$ [265]. In other words, the product of two elements of $\mathbb{L}^{\times}$ again belongs to $\mathbb{L}^{\times}$. Thus, we can define a binary operation $\cdot_{\mathbb{L}^{\times}}$ on the set $\mathbb{L}^{\times}$ (written infix) by

$$a \cdot_{\mathbb{L}^{\times}} b = ab \qquad \text{for all } a \in \mathbb{L}^{\times} \text{ and } b \in \mathbb{L}^{\times}.$$

Also, $1_{\mathbb{L}}$ is an invertible element of $\mathbb{L}$ (indeed, its inverse is $1_{\mathbb{L}}$), and thus an element of $\mathbb{L}^{\times}$.

Now, we need to prove that the set $\mathbb{L}^{\times}$, endowed with the multiplication $\cdot_{\mathbb{L}^{\times}}$ (as multiplication) and the element $1_{\mathbb{L}}$ (as neutral element), is a group. In order to do so, we need to check that the "associativity", "neutrality of 1" and "existence of inverses" axioms are satisfied.

The "associativity" axiom follows from the "associativity of multiplication" axiom in the definition of a noncommutative ring. The "neutrality of 1" axiom follows from the "unitality" axiom in the definition of a noncommutative ring. It thus remains to prove that the "existence of inverses" axiom holds.

[264] In fact, they boil down to the "associativity of addition", "neutrality of 0" and "existence of additive inverses" axioms in the definition of a noncommutative ring.

[265] *Proof.* Let $a \in \mathbb{L}^{\times}$ and $b \in \mathbb{L}^{\times}$. We have $a \in \mathbb{L}^{\times}$; in other words, a is an invertible element of $\mathbb{L}$ (because $\mathbb{L}^{\times}$ is the set of all invertible elements of $\mathbb{L}$). Thus, the inverse a^{-1} of a is well-defined. Similarly, the inverse b^{-1} of b is well-defined. Now, since we have

$$\left(b^{-1}a^{-1}\right)(ab) = b^{-1}\underbrace{a^{-1}a}_{=1_{\mathbb{L}}}b = b^{-1}b = 1_{\mathbb{L}}$$

and

$$(ab)\left(b^{-1}a^{-1}\right) = a\underbrace{bb^{-1}}_{=1_{\mathbb{L}}}a^{-1} = aa^{-1} = 1_{\mathbb{L}},$$

we see that the element $b^{-1}a^{-1}$ of $\mathbb{L}$ is an inverse of ab. Thus, the element ab has an inverse. In other words, ab is invertible. In other words, $ab \in \mathbb{L}^{\times}$ (since $\mathbb{L}^{\times}$ is the set of all invertible elements of $\mathbb{L}$), qed.

Thus, let $a \in \mathbb{L}^{\times}$. We need to show that there exists an $a' \in \mathbb{L}^{\times}$ such that $a \cdot_{\mathbb{L}^{\times}} a' = a' \cdot_{\mathbb{L}^{\times}} a = 1_{\mathbb{L}}$ (since $1_{\mathbb{L}}$ is the neutral element of $\mathbb{L}^{\times}$).

We know that a is an invertible element of $\mathbb{L}$ (since $a \in \mathbb{L}^{\times}$); it thus has an inverse a^{-1}. Now, a itself is an inverse of a^{-1} (since $aa^{-1} = 1_{\mathbb{L}}$ and $a^{-1}a = 1_{\mathbb{L}}$), and thus the element a^{-1} of $\mathbb{L}$ has an inverse. In other words, a^{-1} is invertible, so that $a^{-1} \in \mathbb{L}^{\times}$. The definition of the operation $\cdot_{\mathbb{L}^{\times}}$ shows that $a \cdot_{\mathbb{L}^{\times}} a^{-1} = aa^{-1} = 1_{\mathbb{L}}$ and that $a^{-1} \cdot_{\mathbb{L}^{\times}} a = a^{-1}a = 1_{\mathbb{L}}$. Hence, there exists an $a' \in \mathbb{L}^{\times}$ such that $a \cdot_{\mathbb{L}^{\times}} a' = a' \cdot_{\mathbb{L}^{\times}} a = 1_{\mathbb{L}}$ (namely, $a' = a^{-1}$). Thus we have proven that the "existence of inverses" axiom holds. The proof of Proposition 6.118 **(b)** is thus complete. □

We now have a plentitude of examples of groups: For every noncommutative ring $\mathbb{L}$, we have the two groups $\mathbb{L}^{+}$ and $\mathbb{L}^{\times}$ defined in Proposition 6.118. Another example, for every set X, is the symmetric group of X (endowed with the composition of permutations as multiplication, and the identity permutation $\operatorname{id} : X \to X$ as the neutral element). (Many other examples can be found in textbooks on algebra, such as [Artin10] or [Goodma15]. Groups also naturally appear in the analysis of puzzles such as Rubik's cube; this is explained in various sources such as [Mulhol16], [Bump02] and [Joyner08], which can also be read as introductions to groups.)

Remark 6.119. Throwing all notational ballast aside, we can restate Proposition 6.118 **(b)** as follows: The set of all invertible elements of a noncommutative ring $\mathbb{L}$ is a group (where the binary operation is multiplication). We can apply this to the case where $\mathbb{L} = \mathbb{K}^{n \times n}$ for a commutative ring $\mathbb{K}$ and an integer $n \in \mathbb{N}$. Thus, we obtain that the set of all invertible elements of $\mathbb{K}^{n \times n}$ is a group. Since we know that the invertible elements of $\mathbb{K}^{n \times n}$ are exactly the invertible $n \times n$-matrices (by Example 6.115), we thus have shown that the set of all invertible $n \times n$-matrices is a group. This group is commonly denoted by $\operatorname{GL}_n(\mathbb{K})$; it is called the *general linear group of degree n* over $\mathbb{K}$.

6.19. Cramer's rule

Let us return to the classical properties of determinants. We have already proven many, but here is one more: It is an application of determinants to solving systems of linear equations.

Theorem 6.120. Let $n \in \mathbb{N}$. Let A be an $n \times n$-matrix. Let $b = (b_1, b_2, \ldots, b_n)^T$ be a column vector with n entries (that is, an $n \times 1$-matrix).[266]

For every $j \in \{1, 2, \ldots, n\}$, let $A_j^{\#}$ be the $n \times n$-matrix obtained from A by replacing the j-th column of A with the vector b.

(a) We have $A \cdot \left(\det\left(A_1^{\#}\right), \det\left(A_2^{\#}\right), \ldots, \det\left(A_n^{\#}\right)\right)^T = \det A \cdot b$.

(b) Assume that the matrix A is invertible. Then,

$$A^{-1}b = \left(\frac{\det\left(A_1^{\#}\right)}{\det A}, \frac{\det\left(A_2^{\#}\right)}{\det A}, \ldots, \frac{\det\left(A_n^{\#}\right)}{\det A}\right)^T.$$

Theorem 6.120 (or either part of it) is known as *Cramer's rule*.

Remark 6.121. A system of n linear equations in n variables $x_1, x_2, \ldots, x_n$ can be written in the form $Ax = b$, where A is a fixed $n \times n$-matrix and b is a column vector with n entries (and where x is the column vector $(x_1, x_2, \ldots, x_n)^T$ containing all the variables). When the matrix A is invertible, it thus has a unique solution: namely, $x = A^{-1}b$ (just multiply the equation $Ax = b$ from the left with A^{-1} to see this), and this solution can be computed using Theorem 6.120. This looks nice, but isn't actually all that useful for solving systems of linear equations: For one thing, this does not immediately help us solve systems of fewer or more than n equations in n variables; and even in the case of exactly n equations, the matrix A coming from a system of linear equations will not always be invertible (and in the more interesting cases, it will not be). For another thing, at least when $\mathbb{K}$ is a field, there are faster ways to solve a system of linear equations than anything that involves computing $n+1$ determinants of $n \times n$-matrices. Theorem 6.120 nevertheless turns out to be useful in proofs.

Proof of Theorem 6.120. **(a)** Fix $j \in \{1, 2, \ldots, n\}$. Let $C = A_j^{\#}$. Thus, $C = A_j^{\#}$ is the $n \times n$-matrix obtained from A by replacing the j-th column of A with the vector b. In particular, the j-th column of C is the vector b. In other words, we have

$$c_{p,j} = b_p \qquad \text{for every } p \in \{1, 2, \ldots, n\}. \tag{448}$$

Furthermore, the matrix C is equal to the matrix A in all columns but its j-th column (because it is obtained from A by replacing the j-th column of A with the vector b). Thus, if we cross out the j-th columns in both matrices C and A, then these two matrices become equal. Consequently,

$$C_{\sim p, \sim j} = A_{\sim p, \sim j} \qquad \text{for every } p \in \{1, 2, \ldots, n\} \tag{449}$$

(because the matrices $C_{\sim p, \sim j}$ and $A_{\sim p, \sim j}$ are obtained by crossing out the p-th row

[266] The reader should keep in mind that $(b_1, b_2, \ldots, b_n)^T$ is just a space-saving way to write $\begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}$.

and the j-th column in the matrices C and A, respectively). Now,

$$\det\left(\underbrace{A_j^{\#}}_{=C}\right) = \det C = \sum_{p=1}^{n} (-1)^{p+j} \underbrace{c_{p,j}}_{\substack{=b_p \\ \text{(by (448))}}} \det\left(\underbrace{C_{\sim p,\sim j}}_{\substack{=A_{\sim p,\sim j} \\ \text{(by (449))}}}\right)$$

$$\left(\begin{array}{c}\text{by Theorem 6.82 } \textbf{(b)}\text{, applied} \\ \text{to } C,\ c_{i,j} \text{ and } j \text{ instead of } A,\ a_{i,j} \text{ and } q\end{array}\right)$$

$$= \sum_{p=1}^{n} (-1)^{p+j} b_p \det\left(A_{\sim p,\sim j}\right). \tag{450}$$

Let us now forget that we fixed j. We thus have proven (450) for every $j \in \{1,2,\ldots,n\}$. Now, fix $i \in \{1,2,\ldots,n\}$. Then, for every $p \in \{1,2,\ldots,n\}$ satisfying $p \neq i$, we have

$$\sum_{q=1}^{n} b_p (-1)^{p+q} a_{i,q} \det\left(A_{\sim p,\sim q}\right) = 0 \tag{451}$$

[267]. Also, we have

$$\sum_{q=1}^{n} b_i (-1)^{i+q} a_{i,q} \det\left(A_{\sim i,\sim q}\right) = \det A \cdot b_i \tag{453}$$

[267] *Proof of (451):* Let $p \in \{1,2,\ldots,n\}$ be such that $p \neq i$. Hence, Proposition 6.96 **(a)** (applied to $r = i$) shows that

$$0 = \sum_{q=1}^{n} (-1)^{p+q} a_{i,q} \det\left(A_{\sim p,\sim q}\right). \tag{452}$$

Now,

$$\sum_{q=1}^{n} b_p (-1)^{p+q} a_{i,q} \det\left(A_{\sim p,\sim q}\right) = b_p \underbrace{\sum_{q=1}^{n} (-1)^{p+q} a_{i,q} \det\left(A_{\sim p,\sim q}\right)}_{\substack{=0 \\ \text{(by (452))}}} = 0.$$

Thus, (451) is proven.

[268]. Now,

$$\begin{aligned}
&\sum_{k=1}^{n} a_{i,k} \underbrace{\det\left(A_k^{\#}\right)}_{\substack{=\sum\limits_{p=1}^{n}(-1)^{p+k} b_p \det\left(A_{\sim p,\sim k}\right) \\ \text{(by (450), applied to } j=k\text{)}}} \\
&= \sum_{k=1}^{n} a_{i,k} \sum_{p=1}^{n} (-1)^{p+k} b_p \det\left(A_{\sim p,\sim k}\right) = \sum_{q=1}^{n} a_{i,q} \sum_{p=1}^{n} (-1)^{p+q} b_p \det\left(A_{\sim p,\sim q}\right) \\
&\qquad \text{(here, we renamed the summation index } k \text{ as } q \text{ in the first sum)} \\
&= \underbrace{\sum_{q=1}^{n} \sum_{p=1}^{n}}_{\substack{=\sum\limits_{p=1}^{n}\sum\limits_{q=1}^{n} \\ =\sum\limits_{p\in\{1,2,\ldots,n\}}\sum\limits_{q=1}^{n}}} \underbrace{a_{i,q} (-1)^{p+q} b_p}_{=b_p(-1)^{p+q} a_{i,q}} \det\left(A_{\sim p,\sim q}\right) \\
&= \sum_{p\in\{1,2,\ldots,n\}} \sum_{q=1}^{n} b_p (-1)^{p+q} a_{i,q} \det\left(A_{\sim p,\sim q}\right) \\
&= \sum_{\substack{p\in\{1,2,\ldots,n\}; \\ p\neq i}} \underbrace{\sum_{q=1}^{n} b_p (-1)^{p+q} a_{i,q} \det\left(A_{\sim p,\sim q}\right)}_{\substack{=0 \\ \text{(by (451))}}} + \underbrace{\sum_{q=1}^{n} b_i (-1)^{i+q} a_{i,q} \det\left(A_{\sim i,\sim q}\right)}_{\substack{=\det A\cdot b_i \\ \text{(by (453))}}} \\
&\qquad \text{(here, we have split off the addend for } p=i \text{ from the sum)} \\
&= \underbrace{\sum_{\substack{p\in\{1,2,\ldots,n\}; \\ p\neq i}} 0}_{=0} + \det A \cdot b_i = \det A \cdot b_i. \qquad (455)
\end{aligned}$$

Now, let us forget that we fixed i. We thus have proven (455) for every $i \in$

[268] *Proof of (453):* Applying Theorem 6.82 **(a)** to $p = i$, we obtain

$$\det A = \sum_{q=1}^{n} (-1)^{i+q} a_{i,q} \det\left(A_{\sim i,\sim q}\right). \qquad (454)$$

Now,

$$\sum_{q=1}^{n} b_i (-1)^{i+q} a_{i,q} \det\left(A_{\sim i,\sim q}\right) = b_i \underbrace{\sum_{q=1}^{n} (-1)^{i+q} a_{i,q} \det\left(A_{\sim i,\sim q}\right)}_{\substack{=\det A \\ \text{(by (454))}}} = b_i \det A = \det A \cdot b_i.$$

This proves (453).

$\{1, 2, \ldots, n\}$. Now, let d be the vector $\left(\det\left(A_1^{\#}\right), \det\left(A_2^{\#}\right), \ldots, \det\left(A_n^{\#}\right)\right)^T$. Thus,

$$d = \left(\det\left(A_1^{\#}\right), \det\left(A_2^{\#}\right), \ldots, \det\left(A_n^{\#}\right)\right)^T = \begin{pmatrix} \det\left(A_1^{\#}\right) \\ \det\left(A_2^{\#}\right) \\ \vdots \\ \det\left(A_n^{\#}\right) \end{pmatrix}$$
$$= \left(\det\left(A_i^{\#}\right)\right)_{1 \leq i \leq n,\ 1 \leq j \leq 1}.$$

The definition of the product of two matrices shows that

$$A \cdot d = \left(\underbrace{\sum_{k=1}^{n} a_{i,k} \det\left(A_k^{\#}\right)}_{\substack{= \det A \cdot b_i \\ \text{(by (455))}}} \right)_{1 \leq i \leq n,\ 1 \leq j \leq 1}$$
$$\left(\text{since } A = \left(a_{i,j}\right)_{1 \leq i \leq n,\ 1 \leq j \leq n} \text{ and } d = \left(\det\left(A_i^{\#}\right)\right)_{1 \leq i \leq n,\ 1 \leq j \leq 1}\right)$$
$$= \left(\det A \cdot b_i\right)_{1 \leq i \leq n,\ 1 \leq j \leq 1} = \left(\det A \cdot b_1, \det A \cdot b_2, \ldots, \det A \cdot b_n\right)^T.$$

Comparing this with

$$\det A \cdot \underbrace{b}_{= (b_1, b_2, \ldots, b_n)^T} = \det A \cdot \left(b_1, b_2, \ldots, b_n\right)^T = \left(\det A \cdot b_1, \det A \cdot b_2, \ldots, \det A \cdot b_n\right)^T,$$

we obtain $A \cdot d = \det A \cdot b$. Since $d = \left(\det\left(A_1^{\#}\right), \det\left(A_2^{\#}\right), \ldots, \det\left(A_n^{\#}\right)\right)^T$, we can rewrite this as $A \cdot \left(\det\left(A_1^{\#}\right), \det\left(A_2^{\#}\right), \ldots, \det\left(A_n^{\#}\right)\right)^T = \det A \cdot b$. This proves Theorem 6.120 **(a)**.

(b) Theorem 6.110 **(a)** shows that the matrix A is invertible if and only if the element $\det A$ of $\mathbb{K}$ is invertible (in $\mathbb{K}$). Hence, the element $\det A$ of $\mathbb{K}$ is invertible (since the matrix A is invertible). Thus, $\dfrac{1}{\det A}$ is well-defined. Clearly,

$\underbrace{\frac{1}{\det A} \cdot \det A}_{=1} \cdot b = b$, so that

$$\begin{aligned} b &= \frac{1}{\det A} \cdot \underbrace{\det A \cdot b}_{\substack{=A\cdot\left(\det\left(A_1^{\#}\right),\det\left(A_2^{\#}\right),\ldots,\det\left(A_n^{\#}\right)\right)^T \\ \text{(by Theorem 6.120 (a))}}} \\ &= \frac{1}{\det A} \cdot A \cdot \left(\det\left(A_1^{\#}\right), \det\left(A_2^{\#}\right), \ldots, \det\left(A_n^{\#}\right)\right)^T \\ &= A \cdot \underbrace{\left(\frac{1}{\det A} \cdot \left(\det\left(A_1^{\#}\right), \det\left(A_2^{\#}\right), \ldots, \det\left(A_n^{\#}\right)\right)^T\right)}_{\substack{=\left(\frac{1}{\det A}\det\left(A_1^{\#}\right),\frac{1}{\det A}\det\left(A_2^{\#}\right),\ldots,\frac{1}{\det A}\det\left(A_n^{\#}\right)\right)^T \\ =\left(\frac{\det\left(A_1^{\#}\right)}{\det A},\frac{\det\left(A_2^{\#}\right)}{\det A},\ldots,\frac{\det\left(A_n^{\#}\right)}{\det A}\right)^T}} \\ &= A \cdot \left(\frac{\det\left(A_1^{\#}\right)}{\det A}, \frac{\det\left(A_2^{\#}\right)}{\det A}, \ldots, \frac{\det\left(A_n^{\#}\right)}{\det A}\right)^T. \end{aligned}$$

Therefore,

$$\begin{aligned} &A^{-1} \underbrace{b}_{=A\cdot\left(\frac{\det\left(A_1^{\#}\right)}{\det A},\frac{\det\left(A_2^{\#}\right)}{\det A},\ldots,\frac{\det\left(A_n^{\#}\right)}{\det A}\right)^T} \\ &= \underbrace{A^{-1}A}_{=I_n} \cdot \left(\frac{\det\left(A_1^{\#}\right)}{\det A}, \frac{\det\left(A_2^{\#}\right)}{\det A}, \ldots, \frac{\det\left(A_n^{\#}\right)}{\det A}\right)^T \\ &= I_n \cdot \left(\frac{\det\left(A_1^{\#}\right)}{\det A}, \frac{\det\left(A_2^{\#}\right)}{\det A}, \ldots, \frac{\det\left(A_n^{\#}\right)}{\det A}\right)^T = \left(\frac{\det\left(A_1^{\#}\right)}{\det A}, \frac{\det\left(A_2^{\#}\right)}{\det A}, \ldots, \frac{\det\left(A_n^{\#}\right)}{\det A}\right)^T. \end{aligned}$$

This proves Theorem 6.120 **(b)**. □

6.20. The Desnanot-Jacobi identity

We now move towards more exotic places. In this section[269], we shall prove the *Desnanot-Jacobi identity*, also known as *Lewis Carroll identity*[270]. We will need some

[269] which, unfortunately, has become more technical and tedious than it promised to be – for which I apologize

[270] See [Bresso99, §3.5] for the history of this identity (as well as for a proof different from ours, and for an application). In a nutshell: Desnanot discovered it in 1819; Jacobi proved it in 1833 (and again in 1841); in 1866, Charles Lutwidge Dodgson (better known as Lewis Carroll, although his mathematical works were not printed under this pen name) popularized it by publishing an algorithm for evaluating determinants that made heavy use of this identity.

notations to state this identity in the generality I want; but first I shall state the two best known particular cases (from which the general version can actually be easily derived, although this is not the way I will take):[271]

Proposition 6.122. Let $n \in \mathbb{N}$ be such that $n \geq 2$. Let $A = \left(a_{i,j}\right)_{1\leq i\leq n,\ 1\leq j\leq n}$ be an $n \times n$-matrix. Let A' be the $(n-2) \times (n-2)$-matrix $\left(a_{i+1,j+1}\right)_{1\leq i\leq n-2,\ 1\leq j\leq n-2}$. (In other words, A' is what remains of the matrix A when we remove the first row, the last row, the first column and the last column.) Then,

$$\begin{aligned}
&\det A \cdot \det\left(A'\right) \\
&= \det\left(A_{\sim 1,\sim 1}\right) \cdot \det\left(A_{\sim n,\sim n}\right) - \det\left(A_{\sim 1,\sim n}\right) \cdot \det\left(A_{\sim n,\sim 1}\right).
\end{aligned}$$

Proposition 6.123. Let $n \in \mathbb{N}$ be such that $n \geq 2$. Let $A = \left(a_{i,j}\right)_{1\leq i\leq n,\ 1\leq j\leq n}$ be an $n \times n$-matrix. Let $\widetilde{A}$ be the $(n-2) \times (n-2)$-matrix $\left(a_{i+2,j+2}\right)_{1\leq i\leq n-2,\ 1\leq j\leq n-2}$. (In other words, $\widetilde{A}$ is what remains of the matrix A when we remove the first two rows and the first two columns.) Then,

$$\begin{aligned}
&\det A \cdot \det \widetilde{A} \\
&= \det\left(A_{\sim 1,\sim 1}\right) \cdot \det\left(A_{\sim 2,\sim 2}\right) - \det\left(A_{\sim 1,\sim 2}\right) \cdot \det\left(A_{\sim 2,\sim 1}\right).
\end{aligned}$$

Example 6.124. For this example, set $n = 4$ and $A = \begin{pmatrix} a_1 & b_1 & c_1 & d_1 \\ a_2 & b_2 & c_2 & d_2 \\ a_3 & b_3 & c_3 & d_3 \\ a_4 & b_4 & c_4 & d_4 \end{pmatrix}$. Then, Proposition 6.122 says that

$$\begin{aligned}
&\det \begin{pmatrix} a_1 & b_1 & c_1 & d_1 \\ a_2 & b_2 & c_2 & d_2 \\ a_3 & b_3 & c_3 & d_3 \\ a_4 & b_4 & c_4 & d_4 \end{pmatrix} \cdot \det \begin{pmatrix} b_2 & c_2 \\ b_3 & c_3 \end{pmatrix} \\
&= \det \begin{pmatrix} b_2 & c_2 & d_2 \\ b_3 & c_3 & d_3 \\ b_4 & c_4 & d_4 \end{pmatrix} \cdot \det \begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{pmatrix} \\
&\qquad - \det \begin{pmatrix} a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \\ a_4 & b_4 & c_4 \end{pmatrix} \cdot \det \begin{pmatrix} b_1 & c_1 & d_1 \\ b_2 & c_2 & d_2 \\ b_3 & c_3 & d_3 \end{pmatrix}.
\end{aligned}$$

[271] We shall use the notations of Definition 6.81 throughout this section.

Meanwhile, Proposition 6.123 says that

$$\det \begin{pmatrix} a_1 & b_1 & c_1 & d_1 \\ a_2 & b_2 & c_2 & d_2 \\ a_3 & b_3 & c_3 & d_3 \\ a_4 & b_4 & c_4 & d_4 \end{pmatrix} \cdot \det \begin{pmatrix} c_3 & d_3 \\ c_4 & d_4 \end{pmatrix}$$
$$= \det \begin{pmatrix} b_2 & c_2 & d_2 \\ b_3 & c_3 & d_3 \\ b_4 & c_4 & d_4 \end{pmatrix} \cdot \det \begin{pmatrix} a_1 & c_1 & d_1 \\ a_3 & c_3 & d_3 \\ a_4 & c_4 & d_4 \end{pmatrix}$$
$$- \det \begin{pmatrix} a_2 & c_2 & d_2 \\ a_3 & c_3 & d_3 \\ a_4 & c_4 & d_4 \end{pmatrix} \cdot \det \begin{pmatrix} b_1 & c_1 & d_1 \\ b_3 & c_3 & d_3 \\ b_4 & c_4 & d_4 \end{pmatrix} .$$

Proposition 6.122 occurs (for instance) in [Zeilbe98, *(Alice)*], in [Bresso99, Theorem 3.12], in [Willia15, Example 3.3] and in [Kratte99, Proposition 10] (without a proof, but with a brief list of applications). Proposition 6.123 occurs (among other places) in [BerBru08, (1)] (with a generalization). The reader can easily see that Proposition 6.123 is equivalent to Proposition 6.122[272]; we shall prove a generalization of both.

Let me now introduce some notations:[273]

Definition 6.125. Let $n \in \mathbb{N}$. Let r and s be two elements of $\{1, 2, \ldots, n\}$ such that $r < s$. Then, $(1, 2, \ldots, \widehat{r}, \ldots, \widehat{s}, \ldots, n)$ will denote the $(n-2)$-tuple

$$\left(\underbrace{1, 2, \ldots, r-1}_{\substack{\text{all integers} \\ \text{from } 1 \text{ to } r-1}}, \underbrace{r+1, r+2, \ldots, s-1}_{\substack{\text{all integers} \\ \text{from } r+1 \text{ to } s-1}}, \underbrace{s+1, s+2, \ldots, n}_{\substack{\text{all integers} \\ \text{from } s+1 \text{ to } n}} \right).$$

In other words, $(1, 2, \ldots, \widehat{r}, \ldots, \widehat{s}, \ldots, n)$ will denote the result of removing the entries r and s from the n-tuple $(1, 2, \ldots, n)$.

We can now state a more general version of the Desnanot-Jacobi identity:

Theorem 6.126. Let $n \in \mathbb{N}$ be such that $n \geq 2$. Let p, q, u and v be four elements of $\{1, 2, \ldots, n\}$ such that $p < q$ and $u < v$. Let A be an $n \times n$-matrix. Then,

$$\det A \cdot \det \left(\operatorname{sub}_{1,2,\ldots,\widehat{p},\ldots,\widehat{q},\ldots,n}^{1,2,\ldots,\widehat{u},\ldots,\widehat{v},\ldots,n} A \right)$$
$$= \det \left(A_{\sim p, \sim u} \right) \cdot \det \left(A_{\sim q, \sim v} \right) - \det \left(A_{\sim p, \sim v} \right) \cdot \det \left(A_{\sim q, \sim u} \right) .$$

[272] Indeed, one is easily obtained from the other by swapping the 2-nd and the n-th rows of the matrix A and swapping the 2-nd and the n-th columns of the matrix A, applying parts **(a)** and **(b)** of Exercise 6.7 and checking that all signs cancel.

[273] Recall that we are using the notations of Definition 6.31, of Definition 6.80, and of Definition 6.81.

Example 6.127. If we set $n = 3$, $p = 1$, $q = 2$, $u = 1$, $v = 3$ and $A = \begin{pmatrix} a & a' & a'' \\ b & b' & b'' \\ c & c' & c'' \end{pmatrix}$, then Theorem 6.126 says that

$$\det \begin{pmatrix} a & a' & a'' \\ b & b' & b'' \\ c & c' & c'' \end{pmatrix} \det \begin{pmatrix} c' \end{pmatrix}$$
$$= \det \begin{pmatrix} b' & b'' \\ c' & c'' \end{pmatrix} \cdot \det \begin{pmatrix} a & a' \\ c & c' \end{pmatrix} - \det \begin{pmatrix} b & b' \\ c & c' \end{pmatrix} \cdot \det \begin{pmatrix} a' & a'' \\ c' & c'' \end{pmatrix}.$$

Before we prove this theorem, let me introduce some more notations:

Definition 6.128. Let $n \in \mathbb{N}$ and $m \in \mathbb{N}$. Let $A \in \mathbb{K}^{n \times m}$ be an $n \times m$-matrix.

(a) For every $u \in \{1, 2, \ldots, n\}$, we let $A_{u,\bullet}$ be the u-th row of the matrix A. This $A_{u,\bullet}$ is thus a row vector with m entries, i.e., a $1 \times m$-matrix.

(b) For every $v \in \{1, 2, \ldots, m\}$, we let $A_{\bullet,v}$ be the v-th column of the matrix A. This $A_{\bullet,v}$ is thus a column vector with n entries, i.e., an $n \times 1$-matrix.

(c) For every $u \in \{1, 2, \ldots, n\}$, we set $A_{\sim u,\bullet} = \operatorname{rows}_{1,2,\ldots,\widehat{u},\ldots,n} A$. This $A_{\sim u,\bullet}$ is thus an $(n-1) \times m$-matrix. (In more intuitive terms, the definition of $A_{\sim u,\bullet}$ rewrites as follows: $A_{\sim u,\bullet}$ is the matrix obtained from the matrix A by removing the u-th row.)

(d) For every $v \in \{1, 2, \ldots, m\}$, we set $A_{\bullet,\sim v} = \operatorname{cols}_{1,2,\ldots,\widehat{v},\ldots,m} A$. This $A_{\bullet,\sim v}$ is thus an $n \times (m-1)$-matrix. (In more intuitive terms, the definition of $A_{\bullet,\sim v}$ rewrites as follows: $A_{\bullet,\sim v}$ is the matrix obtained from the matrix A by removing the v-th column.)

Example 6.129. If $n = 3$, $m = 4$ and $A = \begin{pmatrix} a & b & c & d \\ a' & b' & c' & d' \\ a'' & b'' & c'' & d'' \end{pmatrix}$, then

$$A_{2,\bullet} = \begin{pmatrix} a' & b' & c' & d' \end{pmatrix}, \qquad A_{\bullet,2} = \begin{pmatrix} b \\ b' \\ b'' \end{pmatrix},$$
$$A_{\sim 2,\bullet} = \begin{pmatrix} a & b & c & d \\ a'' & b'' & c'' & d'' \end{pmatrix}, \qquad A_{\bullet,\sim 2} = \begin{pmatrix} a & c & d \\ a' & c' & d' \\ a'' & c'' & d'' \end{pmatrix}.$$

Here are some simple properties of the notations introduced in Definition 6.128:

Proposition 6.130. Let $n \in \mathbb{N}$ and $m \in \mathbb{N}$. Let $A \in \mathbb{K}^{n \times m}$ be an $n \times m$-matrix.

(a) For every $u \in \{1, 2, \ldots, n\}$, we have

$$A_{u,\bullet} = (\text{the } u\text{-th row of the matrix } A) = \operatorname{rows}_u A.$$

(Here, the notation $\operatorname{rows}_u A$ is a particular case of Definition 6.31 **(a)**.)

(b) For every $v \in \{1, 2, \ldots, m\}$, we have

$$A_{\bullet,v} = (\text{the } v\text{-th column of the matrix } A) = \operatorname{cols}_v A.$$

(c) For every $u \in \{1, 2, \ldots, n\}$ and $v \in \{1, 2, \ldots, m\}$, we have

$$\left(A_{\bullet,\sim v}\right)_{\sim u,\bullet} = \left(A_{\sim u,\bullet}\right)_{\bullet,\sim v} = A_{\sim u,\sim v}.$$

(d) For every $v \in \{1, 2, \ldots, m\}$ and $w \in \{1, 2, \ldots, v-1\}$, we have $\left(A_{\bullet,\sim v}\right)_{\bullet,w} = A_{\bullet,w}$.

(e) For every $v \in \{1, 2, \ldots, m\}$ and $w \in \{v, v+1, \ldots, m-1\}$, we have $\left(A_{\bullet,\sim v}\right)_{\bullet,w} = A_{\bullet,w+1}$.

(f) For every $u \in \{1, 2, \ldots, n\}$ and $w \in \{1, 2, \ldots, u-1\}$, we have $\left(A_{\sim u,\bullet}\right)_{w,\bullet} = A_{w,\bullet}$.

(g) For every $u \in \{1, 2, \ldots, n\}$ and $w \in \{u, u+1, \ldots, n-1\}$, we have $\left(A_{\sim u,\bullet}\right)_{w,\bullet} = A_{w+1,\bullet}$.

(h) For every $v \in \{1, 2, \ldots, m\}$ and $w \in \{1, 2, \ldots, v-1\}$, we have

$$\left(A_{\bullet,\sim v}\right)_{\bullet,\sim w} = \operatorname{cols}_{1,2,\ldots,\widehat{w},\ldots,\widehat{v},\ldots,m} A.$$

(i) For every $v \in \{1, 2, \ldots, m\}$ and $w \in \{v, v+1, \ldots, m-1\}$, we have

$$\left(A_{\bullet,\sim v}\right)_{\bullet,\sim w} = \operatorname{cols}_{1,2,\ldots,\widehat{v},\ldots,\widehat{w+1},\ldots,m} A.$$

(j) For every $u \in \{1, 2, \ldots, n\}$ and $w \in \{1, 2, \ldots, u-1\}$, we have

$$\left(A_{\sim u,\bullet}\right)_{\sim w,\bullet} = \operatorname{rows}_{1,2,\ldots,\widehat{w},\ldots,\widehat{u},\ldots,n} A.$$

(k) For every $u \in \{1, 2, \ldots, n\}$ and $w \in \{u, u+1, \ldots, n-1\}$, we have

$$\left(A_{\sim u,\bullet}\right)_{\sim w,\bullet} = \operatorname{rows}_{1,2,\ldots,\widehat{u},\ldots,\widehat{w+1},\ldots,n} A.$$

(l) For every $v \in \{1, 2, \ldots, n\}$, $u \in \{1, 2, \ldots, n\}$ and $q \in \{1, 2, \ldots, m\}$ satisfying $u < v$, we have

$$\left(A_{\sim v,\bullet}\right)_{\sim u,\sim q} = \operatorname{rows}_{1,2,\ldots,\widehat{u},\ldots,\widehat{v},\ldots,n}\left(A_{\bullet,\sim q}\right).$$

Proposition 6.131. Let $n \in \mathbb{N}$. Let u and v be two elements of $\{1, 2, \ldots, n\}$ such that $u < v$. Then, $\left(\left(I_n\right)_{\bullet,u}\right)_{\sim v,\bullet} = \left(I_{n-1}\right)_{\bullet,u}$. (Recall that I_m denotes the $m \times m$

identity matrix for each $m \in \mathbb{N}$.)

Exercise 6.39. **(a)** Prove Proposition 6.130 and Proposition 6.131.

(b) Derive Proposition 6.123 and Proposition 6.122 from Theorem 6.126.

Here comes another piece of notation:

Definition 6.132. Let $n \in \mathbb{N}$ and $m \in \mathbb{N}$. Let $A \in \mathbb{K}^{n \times m}$ be an $n \times m$-matrix. Let $v \in \mathbb{K}^{n \times 1}$ be a column vector with n entries. Then, $(A \mid v)$ will denote the $n \times (m+1)$-matrix whose $m+1$ columns are $A_{\bullet,1}, A_{\bullet,2}, \ldots, A_{\bullet,m}, v$ (from left to right). (Informally speaking, $(A \mid v)$ is the matrix obtained when the column vector v is "attached" to A at the right edge.)

Example 6.133. We have $\left(\begin{pmatrix} a & b \\ c & d \end{pmatrix} \mid \begin{pmatrix} p \\ q \end{pmatrix} \right) = \begin{pmatrix} a & b & p \\ c & d & q \end{pmatrix}$.

The following properties of the notation introduced in Definition 6.132 are not too hard to see (see the solution of Exercise 6.40 for their proofs), and will be used below:

Proposition 6.134. Let $n \in \mathbb{N}$ and $m \in \mathbb{N}$. Let $A \in \mathbb{K}^{n \times m}$ be an $n \times m$-matrix. Let $v \in \mathbb{K}^{n \times 1}$ be a column vector with n entries.

(a) Every $q \in \{1, 2, \ldots, m\}$ satisfies $(A \mid v)_{\bullet,q} = A_{\bullet,q}$.

(b) We have $(A \mid v)_{\bullet,m+1} = v$.

(c) Every $q \in \{1, 2, \ldots, m\}$ satisfies $(A \mid v)_{\bullet,\sim q} = \left(A_{\bullet,\sim q} \mid v\right)$.

(d) We have $(A \mid v)_{\bullet,\sim(m+1)} = A$.

(e) We have $(A \mid v)_{\sim p,\bullet} = \left(A_{\sim p,\bullet} \mid v_{\sim p,\bullet}\right)$ for every $p \in \{1, 2, \ldots, n\}$.

(f) We have $(A \mid v)_{\sim p,\sim(m+1)} = A_{\sim p,\bullet}$ for every $p \in \{1, 2, \ldots, n\}$.

Proposition 6.135. Let n be a positive integer. Let $A \in \mathbb{K}^{n \times (n-1)}$.

(a) For every $v = (v_1, v_2, \ldots, v_n)^T \in \mathbb{K}^{n \times 1}$, we have

$$\det (A \mid v) = \sum_{i=1}^{n} (-1)^{n+i} v_i \det \left(A_{\sim i,\bullet}\right).$$

(b) For every $p \in \{1, 2, \ldots, n\}$, we have

$$\det \left(A \mid (I_n)_{\bullet,p}\right) = (-1)^{n+p} \det \left(A_{\sim p,\bullet}\right).$$

(Notice that $(I_n)_{\bullet,p}$ is the p-th column of the $n \times n$ identity matrix, i.e., the column vector $\left(\underbrace{0, 0, \ldots, 0}_{p-1 \text{ zeroes}}, 1, \underbrace{0, 0, \ldots, 0}_{n-p \text{ zeroes}} \right)^T$.)

Proposition 6.136. Let $n \in \mathbb{N}$. Let $A \in \mathbb{K}^{n \times n}$.

(a) If $n > 0$, then $\left(A_{\bullet,\sim n} \mid A_{\bullet,n}\right) = A$.

(b) For every $q \in \{1, 2, \ldots, n\}$, we have $\det\left(A_{\bullet,\sim q} \mid A_{\bullet,q}\right) = (-1)^{n+q} \det A$.

(c) If r and q are two elements of $\{1, 2, \ldots, n\}$ satisfying $r \neq q$, then $\det\left(A_{\bullet,\sim q} \mid A_{\bullet,r}\right) = 0$.

(d) For every $p \in \{1, 2, \ldots, n\}$ and $q \in \{1, 2, \ldots, n\}$, we have $\det\left(A_{\bullet,\sim q} \mid (I_n)_{\bullet,p}\right) = (-1)^{n+p} \det\left(A_{\sim p,\sim q}\right)$.

(e) If u and v are two elements of $\{1, 2, \ldots, n\}$ satisfying $u < v$, and if r is an element of $\{1, 2, \ldots, n-1\}$ satisfying $r \neq u$, then $\det\left(A_{\bullet,\sim u} \mid (A_{\bullet,\sim v})_{\bullet,r}\right) = 0$.

(f) If u and v are two elements of $\{1, 2, \ldots, n\}$ satisfying $u < v$, then $(-1)^u \det\left(A_{\bullet,\sim u} \mid (A_{\bullet,\sim v})_{\bullet,u}\right) = (-1)^n \det A$.

Exercise 6.40. Prove Proposition 6.134, Proposition 6.135 and Proposition 6.136.

Now, we can state a slightly more interesting identity:

Proposition 6.137. Let n be a positive integer. Let $A \in \mathbb{K}^{n \times (n-1)}$ and $C \in \mathbb{K}^{n \times n}$. Let $v \in \{1, 2, \ldots, n\}$. Then,

$$\det\left(A_{\sim v,\bullet}\right) \det C = \sum_{q=1}^{n} (-1)^{n+q} \det\left(A \mid C_{\bullet,q}\right) \det\left(C_{\sim v,\sim q}\right).$$

Example 6.138. If we set $n = 3$, $A = \begin{pmatrix} a & a' \\ b & b' \\ c & c' \end{pmatrix}$, $C = \begin{pmatrix} x & x' & x'' \\ y & y' & y'' \\ z & z' & z'' \end{pmatrix}$ and $v = 2$, then Proposition 6.137 states that

$$\begin{aligned} &\det\begin{pmatrix} a & a' \\ c & c' \end{pmatrix} \det\begin{pmatrix} x & x' & x'' \\ y & y' & y'' \\ z & z' & z'' \end{pmatrix} \\ &= \det\begin{pmatrix} a & a' & x \\ b & b' & y \\ c & c' & z \end{pmatrix} \det\begin{pmatrix} x' & x'' \\ z' & z'' \end{pmatrix} - \det\begin{pmatrix} a & a' & x' \\ b & b' & y' \\ c & c' & z' \end{pmatrix} \det\begin{pmatrix} x & x'' \\ z & z'' \end{pmatrix} \\ &\quad + \det\begin{pmatrix} a & a' & x'' \\ b & b' & y'' \\ c & c' & z'' \end{pmatrix} \det\begin{pmatrix} x & x' \\ z & z' \end{pmatrix}. \end{aligned}$$

Proof of Proposition 6.137. Write the $n \times n$-matrix C in the form $C = \left(c_{i,j}\right)_{1 \leq i \leq n,\ 1 \leq j \leq n}$.

Fix $q \in \{1,2,\ldots,n\}$. Then, $C_{\bullet,q}$ is the q-th column of the matrix C (by the definition of $C_{\bullet,q}$). Thus,

$$\begin{aligned} C_{\bullet,q} &= (\text{the } q\text{-th column of the matrix } C) \\ &= \begin{pmatrix} c_{1,q} \\ c_{2,q} \\ \vdots \\ c_{n,q} \end{pmatrix} \qquad \left(\text{since } C = \left(c_{i,j}\right)_{1\leq i\leq n,\ 1\leq j\leq n}\right) \\ &= \left(c_{1,q}, c_{2,q}, \ldots, c_{n,q}\right)^T . \end{aligned}$$

Now, Proposition 6.135 **(a)** (applied to $C_{\bullet,q}$ and $c_{i,q}$ instead of v and v_i) yields

$$\begin{aligned} \det\left(A \mid C_{\bullet,q}\right) &= \sum_{i=1}^{n} (-1)^{n+i} c_{i,q} \det\left(A_{\sim i,\bullet}\right) \\ &= \sum_{p=1}^{n} (-1)^{n+p} c_{p,q} \det\left(A_{\sim p,\bullet}\right) \end{aligned} \tag{456}$$

(here, we have renamed the summation index i as p).

Now, forget that we fixed q. We thus have proven (456) for each $q \in \{1,2,\ldots,n\}$.

We have

$$\sum_{q=1}^{n} (-1)^{v+q} c_{v,q} \det\left(C_{\sim v,\sim q}\right) = \det C \tag{457}$$

[274].

On the other hand, every $p \in \{1,2,\ldots,n\}$ satisfying $p \neq v$ satisfies

$$\sum_{q=1}^{n} (-1)^{p+q} c_{p,q} \det\left(C_{\sim v,\sim q}\right) = 0 \tag{459}$$

[275].

[274] *Proof of (457):* Theorem 6.82 (applied to C, $c_{i,j}$ and v instead of A, $a_{i,j}$ and p) yields

$$\det C = \sum_{q=1}^{n} (-1)^{v+q} c_{v,q} \det\left(C_{\sim v,\sim q}\right). \tag{458}$$

This proves (457).

[275] *Proof of (459):* Let $p \in \{1,2,\ldots,n\}$ be such that $p \neq v$. Thus, $v \neq p$. Hence, Proposition 6.96 **(a)** (applied to C, $c_{i,j}$, p and v instead of A, $a_{i,j}$, r and p) yields

$$0 = \sum_{q=1}^{n} (-1)^{v+q} c_{p,q} \det\left(C_{\sim v,\sim q}\right). \tag{460}$$

Now, every $q \in \{1, 2, \ldots, n\}$ satisfies

$$\begin{aligned}(-1)^{p+q} &= (-1)^{(p-v)+(v+q)} \qquad (\text{since } p+q = (p-v)+(v+q)) \\ &= (-1)^{p-v} (-1)^{v+q}.\end{aligned}$$

Hence,

$$\begin{aligned}&\sum_{q=1}^{n} \underbrace{(-1)^{p+q}}_{=(-1)^{p-v}(-1)^{v+q}} c_{p,q} \det\left(C_{\sim v,\sim q}\right) \\ &= \sum_{q=1}^{n} (-1)^{p-v} (-1)^{v+q} c_{p,q} \det\left(C_{\sim v,\sim q}\right) = (-1)^{p-v} \underbrace{\sum_{q=1}^{n} (-1)^{v+q} c_{p,q} \det\left(C_{\sim v,\sim q}\right)}_{\substack{=0 \\ (\text{by } (460))}} = 0.\end{aligned}$$

This proves (459).

Now,

$$\begin{aligned}
&\sum_{q=1}^{n} (-1)^{n+q} \underbrace{\det\left(A \mid C_{\bullet,q}\right)}_{\substack{=\sum\limits_{p=1}^{n} (-1)^{n+p} c_{p,q} \det\left(A_{\sim p,\bullet}\right) \\ \text{(by (456))}}} \det\left(C_{\sim v,\sim q}\right) \\
&= \sum_{q=1}^{n} (-1)^{n+q} \left(\sum_{p=1}^{n} (-1)^{n+p} c_{p,q} \det\left(A_{\sim p,\bullet}\right) \right) \det\left(C_{\sim v,\sim q}\right) \\
&= \underbrace{\sum_{q=1}^{n} \sum_{p=1}^{n}}_{=\sum\limits_{p=1}^{n} \sum\limits_{q=1}^{n}} \underbrace{(-1)^{n+q} (-1)^{n+p}}_{\substack{=(-1)^{(n+q)+(n+p)}=(-1)^{p+q} \\ \text{(since } (n+q)+(n+p)=2n+p+q\equiv p+q \operatorname{mod} 2\text{)}}} c_{p,q} \det\left(A_{\sim p,\bullet}\right) \det\left(C_{\sim v,\sim q}\right) \\
&= \underbrace{\sum_{p=1}^{n}}_{=\sum\limits_{p\in\{1,2,\ldots,n\}}} \underbrace{\sum_{q=1}^{n} (-1)^{p+q} c_{p,q} \det\left(A_{\sim p,\bullet}\right) \det\left(C_{\sim v,\sim q}\right)}_{=\det\left(A_{\sim p,\bullet}\right) \sum\limits_{q=1}^{n} (-1)^{p+q} c_{p,q} \det\left(C_{\sim v,\sim q}\right)} \\
&= \sum_{p\in\{1,2,\ldots,n\}} \det\left(A_{\sim p,\bullet}\right) \sum_{q=1}^{n} (-1)^{p+q} c_{p,q} \det\left(C_{\sim v,\sim q}\right) \\
&= \det\left(A_{\sim v,\bullet}\right) \underbrace{\sum_{q=1}^{n} (-1)^{v+q} c_{v,q} \det\left(C_{\sim v,\sim q}\right)}_{\substack{=\det C \\ \text{(by (457))}}} \\
&\qquad + \sum_{\substack{p\in\{1,2,\ldots,n\}; \\ p\neq v}} \det\left(A_{\sim p,\bullet}\right) \underbrace{\sum_{q=1}^{n} (-1)^{p+q} c_{p,q} \det\left(C_{\sim v,\sim q}\right)}_{\substack{=0 \\ \text{(by (459))}}} \\
&\qquad \left(\begin{array}{c} \text{here, we have split off the addend for } p=v \text{ from the sum,} \\ \text{since } v\in\{1,2,\ldots,n\} \end{array} \right) \\
&= \det\left(A_{\sim v,\bullet}\right) \det C + \underbrace{\sum_{\substack{p\in\{1,2,\ldots,n\}; \\ p\neq v}} \det\left(A_{\sim p,\bullet}\right) 0}_{=0} = \det\left(A_{\sim v,\bullet}\right) \det C.
\end{aligned}$$

This proves Proposition 6.137. □

Now, let us show a purely technical lemma (gathering a few equalities for easy access in a proof further down):

Lemma 6.139. Let n be a positive integer. Let $B \in \mathbb{K}^{n \times (n-1)}$. Let u and v be two elements of $\{1, 2, \ldots, n\}$ such that $u < v$.

Consider the vector $(I_n)_{\bullet,u} \in \mathbb{K}^{n \times 1}$. (This is the u-th column of the identity matrix I_n. [276])

Define an $n \times n$-matrix $C \in \mathbb{K}^{n \times n}$ by

$$C = \left(B \mid (I_n)_{\bullet,u}\right).$$

Then, the following holds:

(a) We have

$$\det\left(C_{\sim v, \sim q}\right) = -\left(-1\right)^{n+u} \det\left(\operatorname{rows}_{1,2,\ldots,\widehat{u},\ldots,\widehat{v},\ldots,n}\left(B_{\bullet,\sim q}\right)\right)$$

for every $q \in \{1, 2, \ldots, n-1\}$.

(b) We have

$$\left(-1\right)^{n+q} \det\left(C_{\sim v, \sim q}\right) = -\left(-1\right)^{q+u} \det\left(\operatorname{rows}_{1,2,\ldots,\widehat{u},\ldots,\widehat{v},\ldots,n}\left(B_{\bullet,\sim q}\right)\right) \qquad (461)$$

for every $q \in \{1, 2, \ldots, n-1\}$.

(c) We have

$$C_{\sim v, \sim n} = B_{\sim v, \bullet}. \qquad (462)$$

(d) We have

$$C_{\bullet,q} = B_{\bullet,q} \qquad (463)$$

for every $q \in \{1, 2, \ldots, n-1\}$.

(e) Any $A \in \mathbb{K}^{n \times (n-1)}$ satisfies

$$\det\left(A \mid C_{\bullet,n}\right) = \left(-1\right)^{n+u} \det\left(A_{\sim u, \bullet}\right). \qquad (464)$$

(f) We have

$$\det C = \left(-1\right)^{n+u} \det\left(B_{\sim u, \bullet}\right). \qquad (465)$$

Proof of Lemma 6.139. We have $n - 1 \in \mathbb{N}$ (since n is a positive integer). Also, $u < v \leq n$ (since $v \in \{1, 2, \ldots, n\}$) and thus $u \leq n - 1$ (since u and n are integers). Combining this with $u \geq 1$ (since $u \in \{1, 2, \ldots, n\}$), we obtain $u \in \{1, 2, \ldots, n-1\}$.

(a) Let $q \in \{1, 2, \ldots, n-1\}$. From $C = \left(B \mid (I_n)_{\bullet,u}\right)$, we obtain

$$C_{\bullet,\sim q} = \left(B \mid (I_n)_{\bullet,u}\right)_{\bullet,\sim q} = \left(B_{\bullet,\sim q} \mid (I_n)_{\bullet,u}\right)$$

[276] Explicitly,

$$(I_n)_{\bullet,u} = \left(\underbrace{0, 0, \ldots, 0}_{u-1 \text{ zeroes}}, 1, \underbrace{0, 0, \ldots, 0}_{n-u \text{ zeroes}}\right)^T.$$

(by Proposition 6.134 **(c)**, applied to $n-1$, B and $(I_n)_{\bullet,u}$ instead of m, A and v). But Proposition 6.130 **(c)** (applied to n, n, C, v and q instead of n, m, A, u and v) yields $\left(C_{\bullet,\sim q}\right)_{\sim v,\bullet} = (C_{\sim v,\bullet})_{\bullet,\sim q} = C_{\sim v,\sim q}$. Hence,

$$\begin{aligned} C_{\sim v,\sim q} &= \left(\underbrace{C_{\bullet,\sim q}}_{=\left(B_{\bullet,\sim q} \mid (I_n)_{\bullet,u}\right)} \right)_{\sim v,\bullet} = \left(B_{\bullet,\sim q} \mid (I_n)_{\bullet,u}\right)_{\sim v,\bullet} \\ &= \left(\left(B_{\bullet,\sim q}\right)_{\sim v,\bullet} \mid \left((I_n)_{\bullet,u}\right)_{\sim v,\bullet} \right) \qquad (466) \end{aligned}$$

(by Proposition 6.134 **(e)**, applied to n, $n-2$, $B_{\bullet,\sim q}$, $(I_n)_{\bullet,u}$ and v instead of n, m, A, v and p).

On the other hand, Proposition 6.130 **(c)** (applied to n, $n-1$, B, v and q instead of n, m, A, u and v) yields $\left(B_{\bullet,\sim q}\right)_{\sim v,\bullet} = (B_{\sim v,\bullet})_{\bullet,\sim q} = B_{\sim v,\sim q}$. Now, (466) becomes

$$C_{\sim v,\sim q} = \left(\underbrace{\left(B_{\bullet,\sim q}\right)_{\sim v,\bullet}}_{=(B_{\sim v,\bullet})_{\bullet,\sim q}} \mid \underbrace{\left((I_n)_{\bullet,u}\right)_{\sim v,\bullet}}_{\substack{=(I_{n-1})_{\bullet,u} \\ \text{(by Proposition 6.131)}}} \right) = \left((B_{\sim v,\bullet})_{\bullet,\sim q} \mid (I_{n-1})_{\bullet,u} \right).$$

Hence,

$$\begin{aligned} \det \underbrace{\left(C_{\sim v,\sim q}\right)}_{=\left((B_{\sim v,\bullet})_{\bullet,\sim q} \mid (I_{n-1})_{\bullet,u}\right)} &= \det \left((B_{\sim v,\bullet})_{\bullet,\sim q} \mid (I_{n-1})_{\bullet,u} \right) \\ &= (-1)^{(n-1)+u} \det \left((B_{\sim v,\bullet})_{\sim u,\sim q} \right) \end{aligned}$$

(by Proposition 6.136 **(d)**, applied to $n-1$, $B_{\sim v,\bullet}$ and u instead of n, A and p). Thus,

$$\begin{aligned} \det \left(C_{\sim v,\sim q}\right) &= \underbrace{(-1)^{(n-1)+u}}_{\substack{=(-1)^{n+u+1} \\ \text{(since } (n-1)+u=(n+u+1)-2 \\ \equiv n+u+1 \bmod 2)}} \det \left(\underbrace{(B_{\sim v,\bullet})_{\sim u,\sim q}}_{\substack{=\operatorname{rows}_{1,2,\ldots,\widehat{u},\ldots,\widehat{v},\ldots,n}\left(B_{\bullet,\sim q}\right) \\ \text{(by Proposition 6.130 } \textbf{(l)}\text{,} \\ \text{applied to } B \text{ and } n-1 \text{ instead of } A \text{ and } m)}} \right) \\ &= \underbrace{(-1)^{n+u+1}}_{=-(-1)^{n+u}} \det \left(\operatorname{rows}_{1,2,\ldots,\widehat{u},\ldots,\widehat{v},\ldots,n} \left(B_{\bullet,\sim q}\right) \right) \\ &= -(-1)^{n+u} \det \left(\operatorname{rows}_{1,2,\ldots,\widehat{u},\ldots,\widehat{v},\ldots,n} \left(B_{\bullet,\sim q}\right) \right). \end{aligned}$$

This proves Lemma 6.139 **(a)**.

(b) Let $q \in \{1, 2, \ldots, n-1\}$. Then,

$$\begin{aligned}
&(-1)^{n+q} \underbrace{\det\left(C_{\sim v, \sim q}\right)}_{\substack{= -(-1)^{n+u} \det\left(\operatorname{rows}_{1,2,\ldots,\widehat{u},\ldots,\widehat{v},\ldots,n}\left(B_{\bullet,\sim q}\right)\right) \\ \text{(by Lemma 6.139 (a))}}} \\
&= (-1)^{n+q} \left(-(-1)^{n+u} \det\left(\operatorname{rows}_{1,2,\ldots,\widehat{u},\ldots,\widehat{v},\ldots,n}\left(B_{\bullet,\sim q}\right)\right)\right) \\
&= - \underbrace{(-1)^{n+q}(-1)^{n+u}}_{\substack{=(-1)^{(n+q)+(n+u)}=(-1)^{q+u} \\ \text{(since } (n+q)+(n+u)=2n+q+u \equiv q+u \operatorname{mod} 2)}} \det\left(\operatorname{rows}_{1,2,\ldots,\widehat{u},\ldots,\widehat{v},\ldots,n}\left(B_{\bullet,\sim q}\right)\right) \\
&= -(-1)^{q+u} \det\left(\operatorname{rows}_{1,2,\ldots,\widehat{u},\ldots,\widehat{v},\ldots,n}\left(B_{\bullet,\sim q}\right)\right).
\end{aligned}$$

This proves Lemma 6.139 **(b)**.

(c) We have $C = \left(B \mid (I_n)_{\bullet,u}\right)$. Thus,

$$\begin{aligned}
C_{\sim v,\sim n} &= \left(B \mid (I_n)_{\bullet,u}\right)_{\sim v,\sim n} = \left(B \mid (I_n)_{\bullet,u}\right)_{\sim v,\sim((n-1)+1)} \qquad (\text{since } n = (n-1)+1) \\
&= B_{\sim v,\bullet} \qquad \left(\begin{array}{c}\text{by Proposition 6.134 (f), applied to} \\ n-1,\ B \text{ and } (I_n)_{\bullet,u} \text{ instead of } m,\ A \text{ and } v\end{array}\right).
\end{aligned}$$

This proves Lemma 6.139 **(c)**.

(d) Let $q \in \{1, 2, \ldots, n-1\}$. From $C = \left(B \mid (I_n)_{\bullet,u}\right)$, we obtain

$$C_{\bullet,q} = \left(B \mid (I_n)_{\bullet,u}\right)_{\bullet,q} = B_{\bullet,q}$$

(by Proposition 6.134 **(a)**, applied to $n-1$, B and $(I_n)_{\bullet,u}$ instead of m, A and v). This proves Lemma 6.139 **(d)**.

(e) Let $A \in \mathbb{K}^{(n-1)\times n}$. Proposition 6.134 **(b)** (applied to $n-1$, B and $(I_n)_{\bullet,u}$ instead of m, A and v) yields $\left(B \mid (I_n)_{\bullet,u}\right)_{\bullet,(n-1)+1} = (I_n)_{\bullet,u}$. This rewrites as $\left(B \mid (I_n)_{\bullet,u}\right)_{\bullet,n} = (I_n)_{\bullet,u}$ (since $(n-1)+1 = n$). Now, $C = \left(B \mid (I_n)_{\bullet,u}\right)$, so that

$$C_{\bullet,n} = \left(B \mid (I_n)_{\bullet,u}\right)_{\bullet,n} = (I_n)_{\bullet,u}.$$

Hence,

$$\det\left(A \mid \underbrace{C_{\bullet,n}}_{=(I_n)_{\bullet,u}}\right) = \det\left(A \mid (I_n)_{\bullet,u}\right) = (-1)^{n+u} \det\left(A_{\sim u,\bullet}\right)$$

(by Proposition 6.135 **(b)**, applied to $p = u$). This proves Lemma 6.139 **(e)**.

(f) We have

$$\det \underbrace{C}_{=\left(B\mid(I_n)_{\bullet,u}\right)} = \det\left(B \mid (I_n)_{\bullet,u}\right) = (-1)^{n+u} \det\left(B_{\sim u,\bullet}\right)$$

(by Proposition 6.135 **(b)**, applied to B and u instead of A and p). This proves Lemma 6.139 **(f)**. □

Next, we claim the following:

Proposition 6.140. Let n be a positive integer. Let $A \in \mathbb{K}^{n\times(n-1)}$ and $B \in \mathbb{K}^{n\times(n-1)}$. Let u and v be two elements of $\{1, 2, \ldots, n\}$ such that $u < v$. Then,

$$\begin{aligned}&\sum_{r=1}^{n-1} (-1)^r \det\left(A \mid B_{\bullet,r}\right) \det\left(\operatorname{rows}_{1,2,\ldots,\widehat{u},\ldots,\widehat{v},\ldots,n}\left(B_{\bullet,\sim r}\right)\right)\\ &= (-1)^n \left(\det\left(A_{\sim u,\bullet}\right)\det\left(B_{\sim v,\bullet}\right) - \det\left(A_{\sim v,\bullet}\right)\det\left(B_{\sim u,\bullet}\right)\right).\end{aligned}$$

Example 6.141. If we set $n = 3$, $A = \begin{pmatrix} a & a' \\ b & b' \\ c & c' \end{pmatrix}$, $B = \begin{pmatrix} x & x' \\ y & y' \\ z & z' \end{pmatrix}$, $u = 1$ and $v = 2$, then Proposition 6.140 claims that

$$\begin{aligned}&-\det\begin{pmatrix} a & a' & x \\ b & b' & y \\ c & c' & z \end{pmatrix}\det\begin{pmatrix} z' \end{pmatrix} + \det\begin{pmatrix} a & a' & x' \\ b & b' & y' \\ c & c' & z' \end{pmatrix}\det\begin{pmatrix} z \end{pmatrix}\\ &= (-1)^3\left(\det\begin{pmatrix} b & b' \\ c & c' \end{pmatrix}\det\begin{pmatrix} x & x' \\ z & z' \end{pmatrix} - \det\begin{pmatrix} a & a' \\ c & c' \end{pmatrix}\det\begin{pmatrix} y & y' \\ z & z' \end{pmatrix}\right).\end{aligned}$$

Proof of Proposition 6.140. We have $n - 1 \in \mathbb{N}$ (since n is a positive integer).

Define an $n \times n$-matrix $C \in \mathbb{K}^{n\times n}$ as in Lemma 6.139.

Proposition 6.137 yields

$$\begin{aligned}
&\det\left(A_{\sim v,\bullet}\right)\det C \\
&= \sum_{q=1}^{n} \underbrace{(-1)^{n+q}\det\left(A \mid C_{\bullet,q}\right)}_{=\det\left(A\mid C_{\bullet,q}\right)(-1)^{n+q}} \det\left(C_{\sim v,\sim q}\right) = \sum_{q=1}^{n} \det\left(A \mid C_{\bullet,q}\right)(-1)^{n+q}\det\left(C_{\sim v,\sim q}\right) \\
&= \sum_{q=1}^{n-1} \det\left(A \mid \underbrace{C_{\bullet,q}}_{\substack{=B_{\bullet,q} \\ \text{(by (463))}}}\right) \underbrace{(-1)^{n+q}\det\left(C_{\sim v,\sim q}\right)}_{\substack{=-(-1)^{q+u}\det\left(\operatorname{rows}_{1,2,\ldots,\widehat{u},\ldots,\widehat{v},\ldots,n}\left(B_{\bullet,\sim q}\right)\right) \\ \text{(by (461))}}} \\
&\qquad + \underbrace{\det\left(A \mid C_{\bullet,n}\right)}_{\substack{=(-1)^{n+u}\det\left(A_{\sim u,\bullet}\right) \\ \text{(by (464))}}} \underbrace{(-1)^{n+n}}_{\substack{=1 \\ \text{(since } n+n=2n \text{ is even)}}} \det\left(\underbrace{C_{\sim v,\sim n}}_{\substack{=B_{\sim v,\bullet} \\ \text{(by (462))}}}\right) \\
&\qquad \left(\text{here, we have split off the addend for } q=n \text{ from the sum}\right) \\
&= \underbrace{\sum_{q=1}^{n-1} \det\left(A \mid B_{\bullet,q}\right)\left(-(-1)^{q+u}\det\left(\operatorname{rows}_{1,2,\ldots,\widehat{u},\ldots,\widehat{v},\ldots,n}\left(B_{\bullet,\sim q}\right)\right)\right)}_{=-\sum_{q=1}^{n-1}(-1)^{q+u}\det\left(A\mid B_{\bullet,q}\right)\det\left(\operatorname{rows}_{1,2,\ldots,\widehat{u},\ldots,\widehat{v},\ldots,n}\left(B_{\bullet,\sim q}\right)\right)} \\
&\qquad + (-1)^{n+u}\det\left(A_{\sim u,\bullet}\right)\det\left(B_{\sim v,\bullet}\right) \\
&= -\sum_{q=1}^{n-1}(-1)^{q+u}\det\left(A \mid B_{\bullet,q}\right)\det\left(\operatorname{rows}_{1,2,\ldots,\widehat{u},\ldots,\widehat{v},\ldots,n}\left(B_{\bullet,\sim q}\right)\right) \\
&\qquad + (-1)^{n+u}\det\left(A_{\sim u,\bullet}\right)\det\left(B_{\sim v,\bullet}\right).
\end{aligned}$$

Adding $\sum\limits_{q=1}^{n-1}(-1)^{q+u}\det\left(A \mid B_{\bullet,q}\right)\det\left(\operatorname{rows}_{1,2,\ldots,\widehat{u},\ldots,\widehat{v},\ldots,n}\left(B_{\bullet,\sim q}\right)\right)$ to both sides of this equality, we obtain

$$\begin{aligned}
&\sum_{q=1}^{n-1}(-1)^{q+u}\det\left(A \mid B_{\bullet,q}\right)\det\left(\operatorname{rows}_{1,2,\ldots,\widehat{u},\ldots,\widehat{v},\ldots,n}\left(B_{\bullet,\sim q}\right)\right) + \det\left(A_{\sim v,\bullet}\right)\det C \\
&= (-1)^{n+u}\det\left(A_{\sim u,\bullet}\right)\det\left(B_{\sim v,\bullet}\right).
\end{aligned}$$

Subtracting $\det(A_{\sim v,\bullet})\det C$ from both sides of this equality, we find

$$\begin{aligned}
&\sum_{q=1}^{n-1}(-1)^{q+u}\det\left(A\mid B_{\bullet,q}\right)\det\left(\operatorname{rows}_{1,2,\ldots,\widehat{u},\ldots,\widehat{v},\ldots,n}\left(B_{\bullet,\sim q}\right)\right)\\
&=(-1)^{n+u}\det(A_{\sim u,\bullet})\det(B_{\sim v,\bullet})-\det(A_{\sim v,\bullet})\underbrace{\det C}_{\substack{=(-1)^{n+u}\det(B_{\sim u,\bullet})\\ \text{(by (465))}}}\\
&=(-1)^{n+u}\det(A_{\sim u,\bullet})\det(B_{\sim v,\bullet})-\underbrace{\det(A_{\sim v,\bullet})(-1)^{n+u}}_{=(-1)^{n+u}\det(A_{\sim v,\bullet})}\det(B_{\sim u,\bullet})\\
&=(-1)^{n+u}\det(A_{\sim u,\bullet})\det(B_{\sim v,\bullet})-(-1)^{n+u}\det(A_{\sim v,\bullet})\det(B_{\sim u,\bullet})\\
&=(-1)^{n+u}\left(\det(A_{\sim u,\bullet})\det(B_{\sim v,\bullet})-\det(A_{\sim v,\bullet})\det(B_{\sim u,\bullet})\right).
\end{aligned}$$

Multiplying both sides of this equality by $(-1)^u$, we obtain

$$\begin{aligned}
&(-1)^u\sum_{q=1}^{n-1}(-1)^{q+u}\det\left(A\mid B_{\bullet,q}\right)\det\left(\operatorname{rows}_{1,2,\ldots,\widehat{u},\ldots,\widehat{v},\ldots,n}\left(B_{\bullet,\sim q}\right)\right)\\
&=\underbrace{(-1)^u(-1)^{n+u}}_{\substack{=(-1)^{u+(n+u)}=(-1)^n\\ \text{(since } u+(n+u)=2u+n\equiv n\bmod 2)}}\left(\det(A_{\sim u,\bullet})\det(B_{\sim v,\bullet})-\det(A_{\sim v,\bullet})\det(B_{\sim u,\bullet})\right)\\
&=(-1)^n\left(\det(A_{\sim u,\bullet})\det(B_{\sim v,\bullet})-\det(A_{\sim v,\bullet})\det(B_{\sim u,\bullet})\right),
\end{aligned}$$

so that

$$\begin{aligned}
&(-1)^n\left(\det(A_{\sim u,\bullet})\det(B_{\sim v,\bullet})-\det(A_{\sim v,\bullet})\det(B_{\sim u,\bullet})\right)\\
&=(-1)^u\sum_{q=1}^{n-1}(-1)^{q+u}\det\left(A\mid B_{\bullet,q}\right)\det\left(\operatorname{rows}_{1,2,\ldots,\widehat{u},\ldots,\widehat{v},\ldots,n}\left(B_{\bullet,\sim q}\right)\right)\\
&=\sum_{q=1}^{n-1}\underbrace{(-1)^u(-1)^{q+u}}_{\substack{=(-1)^{u+(q+u)}=(-1)^q\\ \text{(since } u+(q+u)=2u+q\equiv q\bmod 2)}}\det\left(A\mid B_{\bullet,q}\right)\det\left(\operatorname{rows}_{1,2,\ldots,\widehat{u},\ldots,\widehat{v},\ldots,n}\left(B_{\bullet,\sim q}\right)\right)\\
&=\sum_{q=1}^{n-1}(-1)^q\det\left(A\mid B_{\bullet,q}\right)\det\left(\operatorname{rows}_{1,2,\ldots,\widehat{u},\ldots,\widehat{v},\ldots,n}\left(B_{\bullet,\sim q}\right)\right)\\
&=\sum_{r=1}^{n-1}(-1)^r\det(A\mid B_{\bullet,r})\det(\operatorname{rows}_{1,2,\ldots,\widehat{u},\ldots,\widehat{v},\ldots,n}(B_{\bullet,\sim r}))
\end{aligned}$$

(here, we have renamed the summation index q as r). This proves Proposition 6.140. □

Now, we can finally prove Theorem 6.126:

Proof of Theorem 6.126. We have $v \in \{1,2,\ldots,n\}$ and thus $v \leq n$. Hence, $u < v \leq n$, so that $u \leq n-1$ (since u and n are integers). Also, $u \in \{1,2,\ldots,n\}$, so that $1 \leq u$. Combining $1 \leq u$ with $u \leq n-1$, we obtain $u \in \{1,2,\ldots,n-1\}$.

Proposition 6.79 **(d)** (applied to n, $n-2$, $(1,2,\ldots,\widehat{p},\ldots,\widehat{q},\ldots,n)$, $n-2$ and $(1,2,\ldots,\widehat{u},\ldots,\widehat{v},\ldots,n)$ instead of m, u, $(i_1,i_2,\ldots,i_u)$, v and $(j_1,j_2,\ldots,j_v)$) yields

$$\begin{aligned}\operatorname{sub}_{1,2,\ldots,\widehat{p},\ldots,\widehat{q},\ldots,n}^{1,2,\ldots,\widehat{u},\ldots,\widehat{v},\ldots,n} A &= \operatorname{rows}_{1,2,\ldots,\widehat{p},\ldots,\widehat{q},\ldots,n}\left(\operatorname{cols}_{1,2,\ldots,\widehat{u},\ldots,\widehat{v},\ldots,n} A\right) \\ &= \operatorname{cols}_{1,2,\ldots,\widehat{u},\ldots,\widehat{v},\ldots,n}\left(\operatorname{rows}_{1,2,\ldots,\widehat{p},\ldots,\widehat{q},\ldots,n} A\right).\end{aligned} \tag{467}$$

We have $u < v$, so that $u \leq v-1$ (since u and v are integers). Combining this with $1 \leq u$, we obtain $u \in \{1,2,\ldots,v-1\}$. Thus, Proposition 6.130 **(h)** (applied to $m = n$ and $w = u$) yields $\left(A_{\bullet,\sim v}\right)_{\bullet,\sim u} = \operatorname{cols}_{1,2,\ldots,\widehat{u},\ldots,\widehat{v},\ldots,n} A$. Thus,

$$\begin{aligned}&\operatorname{rows}_{1,2,\ldots,\widehat{p},\ldots,\widehat{q},\ldots,n}\left(\underbrace{\left(A_{\bullet,\sim v}\right)_{\bullet,\sim u}}_{=\operatorname{cols}_{1,2,\ldots,\widehat{u},\ldots,\widehat{v},\ldots,n} A}\right) \\ &= \operatorname{rows}_{1,2,\ldots,\widehat{p},\ldots,\widehat{q},\ldots,n}\left(\operatorname{cols}_{1,2,\ldots,\widehat{u},\ldots,\widehat{v},\ldots,n} A\right) \\ &= \operatorname{sub}_{1,2,\ldots,\widehat{p},\ldots,\widehat{q},\ldots,n}^{1,2,\ldots,\widehat{u},\ldots,\widehat{v},\ldots,n} A \qquad \text{(by (467))}.\end{aligned} \tag{468}$$

Proposition 6.130 **(c)** (applied to n, p and u instead of m, u and v) yields $\left(A_{\bullet,\sim u}\right)_{\sim p,\bullet} = \left(A_{\sim p,\bullet}\right)_{\bullet,\sim u} = A_{\sim p,\sim u}$. Similarly, $\left(A_{\bullet,\sim v}\right)_{\sim p,\bullet} = \left(A_{\sim p,\bullet}\right)_{\bullet,\sim v} = A_{\sim p,\sim v}$ and $\left(A_{\bullet,\sim u}\right)_{\sim q,\bullet} = \left(A_{\sim q,\bullet}\right)_{\bullet,\sim u} = A_{\sim q,\sim u}$ and $\left(A_{\bullet,\sim v}\right)_{\sim q,\bullet} = \left(A_{\sim q,\bullet}\right)_{\bullet,\sim v} = A_{\sim q,\sim v}$.

The integer n is positive (since $n \geq 2$). Thus, Proposition 6.140 (applied to $A_{\bullet,\sim u}$, $A_{\bullet,\sim v}$, p and q instead of A, B, u and v) yields

$$\begin{aligned}&\sum_{r=1}^{n-1} (-1)^r \det\left(A_{\bullet,\sim u} \mid \left(A_{\bullet,\sim v}\right)_{\bullet,r}\right) \det\left(\operatorname{rows}_{1,2,\ldots,\widehat{p},\ldots,\widehat{q},\ldots,n}\left(\left(A_{\bullet,\sim v}\right)_{\bullet,\sim r}\right)\right) \\ &= (-1)^n \left(\det\left(\underbrace{\left(A_{\bullet,\sim u}\right)_{\sim p,\bullet}}_{=A_{\sim p,\sim u}}\right) \det\left(\underbrace{\left(A_{\bullet,\sim v}\right)_{\sim q,\bullet}}_{=A_{\sim q,\sim v}}\right)\right. \\ &\qquad\qquad \left. - \det\left(\underbrace{\left(A_{\bullet,\sim u}\right)_{\sim q,\bullet}}_{=A_{\sim q,\sim u}}\right) \det\left(\underbrace{\left(A_{\bullet,\sim v}\right)_{\sim p,\bullet}}_{=A_{\sim p,\sim v}}\right)\right) \\ &= (-1)^n \left(\det\left(A_{\sim p,\sim u}\right) \det\left(A_{\sim q,\sim v}\right) - \det\left(A_{\sim q,\sim u}\right) \det\left(A_{\sim p,\sim v}\right)\right).\end{aligned}$$

Hence,

$$\begin{aligned}
&(-1)^n \left(\det\left(A_{\sim p,\sim u}\right)\det\left(A_{\sim q,\sim v}\right)-\det\left(A_{\sim q,\sim u}\right)\det\left(A_{\sim p,\sim v}\right)\right)\\
&=\underbrace{\sum_{r=1}^{n-1}}_{=\sum_{r\in\{1,2,\ldots,n-1\}}} (-1)^r \det\left(A_{\bullet,\sim u}\mid\left(A_{\bullet,\sim v}\right)_{\bullet,r}\right)\det\left(\operatorname{rows}_{1,2,\ldots,\widehat{p},\ldots,\widehat{q},\ldots,n}\left(\left(A_{\bullet,\sim v}\right)_{\bullet,\sim r}\right)\right)\\
&=\sum_{r\in\{1,2,\ldots,n-1\}} (-1)^r \det\left(A_{\bullet,\sim u}\mid\left(A_{\bullet,\sim v}\right)_{\bullet,r}\right)\det\left(\operatorname{rows}_{1,2,\ldots,\widehat{p},\ldots,\widehat{q},\ldots,n}\left(\left(A_{\bullet,\sim v}\right)_{\bullet,\sim r}\right)\right)\\
&=\sum_{\substack{r\in\{1,2,\ldots,n-1\};\\ r\neq u}} (-1)^r \underbrace{\det\left(A_{\bullet,\sim u}\mid\left(A_{\bullet,\sim v}\right)_{\bullet,r}\right)}_{\substack{=0\\ \text{(by Proposition 6.136 }\textbf{(e)}\text{)}}}\det\left(\operatorname{rows}_{1,2,\ldots,\widehat{p},\ldots,\widehat{q},\ldots,n}\left(\left(A_{\bullet,\sim v}\right)_{\bullet,\sim r}\right)\right)\\
&\qquad+\underbrace{(-1)^u\det\left(A_{\bullet,\sim u}\mid\left(A_{\bullet,\sim v}\right)_{\bullet,u}\right)}_{\substack{=(-1)^n\det A\\ \text{(by Proposition 6.136 }\textbf{(f)}\text{)}}}\det\left(\underbrace{\operatorname{rows}_{1,2,\ldots,\widehat{p},\ldots,\widehat{q},\ldots,n}\left(\left(A_{\bullet,\sim v}\right)_{\bullet,\sim u}\right)}_{\substack{=\operatorname{sub}^{1,2,\ldots,\widehat{u},\ldots,\widehat{v},\ldots,n}_{1,2,\ldots,\widehat{p},\ldots,\widehat{q},\ldots,n}A\\ \text{(by (468))}}}\right)\\
&\qquad\left(\begin{array}{c}\text{here, we have split off the addend for } r=u \text{ from the sum,}\\ \text{since } u\in\{1,2,\ldots,n-1\}\end{array}\right)\\
&=\underbrace{\sum_{\substack{r\in\{1,2,\ldots,n-1\};\\ r\neq u}} (-1)^r\, 0\det\left(\operatorname{rows}_{1,2,\ldots,\widehat{p},\ldots,\widehat{q},\ldots,n}\left(\left(A_{\bullet,\sim v}\right)_{\bullet,\sim r}\right)\right)}_{=0}\\
&\qquad+(-1)^n\det A\cdot\det\left(\operatorname{sub}^{1,2,\ldots,\widehat{u},\ldots,\widehat{v},\ldots,n}_{1,2,\ldots,\widehat{p},\ldots,\widehat{q},\ldots,n}A\right)\\
&=(-1)^n\det A\cdot\det\left(\operatorname{sub}^{1,2,\ldots,\widehat{u},\ldots,\widehat{v},\ldots,n}_{1,2,\ldots,\widehat{p},\ldots,\widehat{q},\ldots,n}A\right).
\end{aligned}$$

Multiplying both sides of this equality by $(-1)^n$, we obtain

$$\begin{aligned}
&(-1)^n(-1)^n\left(\det\left(A_{\sim p,\sim u}\right)\det\left(A_{\sim q,\sim v}\right)-\det\left(A_{\sim q,\sim u}\right)\det\left(A_{\sim p,\sim v}\right)\right)\\
&=\underbrace{(-1)^n(-1)^n}_{\substack{=(-1)^{n+n}=1\\ \text{(since } n+n=2n \text{ is even)}}}\det A\cdot\det\left(\operatorname{sub}^{1,2,\ldots,\widehat{u},\ldots,\widehat{v},\ldots,n}_{1,2,\ldots,\widehat{p},\ldots,\widehat{q},\ldots,n}A\right)\\
&=\det A\cdot\det\left(\operatorname{sub}^{1,2,\ldots,\widehat{u},\ldots,\widehat{v},\ldots,n}_{1,2,\ldots,\widehat{p},\ldots,\widehat{q},\ldots,n}A\right).
\end{aligned}$$

Thus,

$$\begin{aligned}
&\det A \cdot \det\left(\operatorname{sub}_{1,2,\ldots,\widehat{p},\ldots,\widehat{q},\ldots,n}^{1,2,\ldots,\widehat{u},\ldots,\widehat{v},\ldots,n} A\right) \\
&= \underbrace{(-1)^n (-1)^n}_{\substack{=(-1)^{n+n}=1 \\ \text{(since } n+n=2n \text{ is even)}}} \left(\det\left(A_{\sim p,\sim u}\right) \det\left(A_{\sim q,\sim v}\right) - \det\left(A_{\sim q,\sim u}\right) \det\left(A_{\sim p,\sim v}\right)\right) \\
&= \det\left(A_{\sim p,\sim u}\right) \det\left(A_{\sim q,\sim v}\right) - \det\left(A_{\sim q,\sim u}\right) \det\left(A_{\sim p,\sim v}\right).
\end{aligned}$$

This proves Theorem 6.126. □

Now that Theorem 6.126 is proven, we conclude that Proposition 6.123 and Proposition 6.122 hold as well (because in Exercise 6.39 **(b)**, these two propositions have been derived from Theorem 6.126).

Exercise 6.41. Let n be a positive integer. Let $B \in \mathbb{K}^{n\times(n-1)}$. Fix $q \in \{1,2,\ldots,n-1\}$. For every $x \in \{1,2,\ldots,n\}$, set

$$\alpha_x = \det\left(B_{\sim x,\bullet}\right).$$

For every two elements x and y of $\{1,2,\ldots,n\}$ satisfying $x < y$, set

$$\beta_{x,y} = \det\left(\operatorname{rows}_{1,2,\ldots,\widehat{x},\ldots,\widehat{y},\ldots,n}\left(B_{\bullet,\sim q}\right)\right).$$

(Note that this depends on q, but we do not mention q in the notation because q is fixed.)

Let u, v and w be three elements of $\{1,2,\ldots,n\}$ such that $u < v < w$. Thus, $\beta_{u,v}$, $\beta_{v,w}$ and $\beta_{u,w}$ are well-defined elements of $\mathbb{K}$. Prove that

$$\alpha_u\beta_{v,w} + \alpha_w\beta_{u,v} = \alpha_v\beta_{u,w}.$$

Example 6.142. If we set $n = 4$, $B = \begin{pmatrix} a & a' & a'' \\ b & b' & b'' \\ c & c' & c'' \\ d & d' & d'' \end{pmatrix}$, $q = 3$, $u = 1$, $v = 2$ and $w = 3$, then Exercise 6.41 says that

$$\begin{aligned}
&\det\begin{pmatrix} b & b' & b'' \\ c & c' & c'' \\ d & d' & d'' \end{pmatrix} \cdot \det\begin{pmatrix} a & a' \\ d & d' \end{pmatrix} + \det\begin{pmatrix} a & a' & a'' \\ b & b' & b'' \\ d & d' & d'' \end{pmatrix} \cdot \det\begin{pmatrix} c & c' \\ d & d' \end{pmatrix} \\
&= \det\begin{pmatrix} a & a' & a'' \\ c & c' & c'' \\ d & d' & d'' \end{pmatrix} \cdot \det\begin{pmatrix} b & b' \\ d & d' \end{pmatrix}.
\end{aligned}$$

Remark 6.143. Exercise 6.41 appears in [KenWil14, proof of Theorem 9], where it (or, rather, a certain transformation of determinant expressions that relies on it) is called the "jaw move".

Exercise 6.42. Let $n \in \mathbb{N}$. Let A be an alternating $n \times n$-matrix. (See Definition 6.72 **(b)** for what this means.) Let S be any $n \times n$-matrix. Prove that each entry of the matrix $(\operatorname{adj} S)^T A (\operatorname{adj} S)$ is a multiple of $\det S$.

6.21. The Plücker relation

The following section is devoted to the *Plücker relations*, or, rather, one of the many things that tend to carry this name in the literature[277]. The proofs will be fairly short, since we did much of the necessary work in Section 6.20 already.

We shall use the notations of Definition 6.128 throughout this section.

We begin with the following identity:

Proposition 6.144. Let n be a positive integer. Let $B \in \mathbb{K}^{n\times(n-1)}$. Then:

(a) We have

$$\sum_{r=1}^{n} (-1)^r \det(B_{\sim r,\bullet}) B_{r,\bullet} = 0_{1\times(n-1)}.$$

(Recall that the product $\det(B_{\sim r,\bullet}) B_{r,\bullet}$ in this equality is the product of the scalar $\det(B_{\sim r,\bullet}) \in \mathbb{K}$ with the row vector $B_{r,\bullet} \in \mathbb{K}^{1\times(n-1)}$; as all such products, it is computed entrywise, i.e., by the formula $\lambda(a_1, a_2, \ldots, a_{n-1}) = (\lambda a_1, \lambda a_2, \ldots, \lambda a_{n-1})$.)

(b) Write the matrix B in the form $B = (b_{i,j})_{1\le i\le n,\ 1\le j\le n-1}$. Then,

$$\sum_{r=1}^{n} (-1)^r \det(B_{\sim r,\bullet}) b_{r,q} = 0$$

for every $q \in \{1, 2, \ldots, n-1\}$.

Note that part **(a)** of Proposition 6.144 claims an equality between two row vectors (indeed, $B_{r,\bullet}$ is a row vector with $n-1$ entries for each $r \in \{1, 2, \ldots, n\}$), whereas part **(b)** claims an equality between two elements of $\mathbb{K}$ (for each $q \in \{1, 2, \ldots, n-1\}$). That said, the two parts are essentially restatements of one another, and we will derive part **(a)** from part **(b)** soon enough. Let us first illustrate Proposition 6.144 on an example:

[277] Most of the relevant literature, unfortunately, is not very elementary, as the Plücker relations are at their most useful in the algebraic geometry of the Grassmannian and of flag varieties ("Schubert calculus"). See [KleLak72], [Jacobs10, §3.4] and [Fulton97, §9.1] for expositions (all three, however, well above the level of the present notes).

Example 6.145. For this example, set $n = 4$ and $B = \begin{pmatrix} a & a' & a'' \\ b & b' & b'' \\ c & c' & c'' \\ d & d' & d'' \end{pmatrix}$. Then, Proposition 6.144 **(a)** says that

$$\begin{aligned} & -\det \begin{pmatrix} b & b' & b'' \\ c & c' & c'' \\ d & d' & d'' \end{pmatrix} \cdot \begin{pmatrix} a & a' & a'' \end{pmatrix} + \det \begin{pmatrix} a & a' & a'' \\ c & c' & c'' \\ d & d' & d'' \end{pmatrix} \cdot \begin{pmatrix} b & b' & b'' \end{pmatrix} \\ & \qquad - \det \begin{pmatrix} a & a' & a'' \\ b & b' & b'' \\ d & d' & d'' \end{pmatrix} \cdot \begin{pmatrix} c & c' & c'' \end{pmatrix} + \det \begin{pmatrix} a & a' & a'' \\ b & b' & b'' \\ c & c' & c'' \end{pmatrix} \cdot \begin{pmatrix} d & d' & d'' \end{pmatrix} \\ & = 0_{1\times 3}. \end{aligned}$$

Proposition 6.144 **(b)** (applied to $q = 3$) yields

$$\begin{aligned} & -\det \begin{pmatrix} b & b' & b'' \\ c & c' & c'' \\ d & d' & d'' \end{pmatrix} \cdot a'' + \det \begin{pmatrix} a & a' & a'' \\ c & c' & c'' \\ d & d' & d'' \end{pmatrix} \cdot b'' \\ & \qquad - \det \begin{pmatrix} a & a' & a'' \\ b & b' & b'' \\ d & d' & d'' \end{pmatrix} \cdot c'' + \det \begin{pmatrix} a & a' & a'' \\ b & b' & b'' \\ c & c' & c'' \end{pmatrix} \cdot d'' \\ & = 0. \end{aligned}$$

Proof of Proposition 6.144. **(b)** Let $q \in \{1, 2, \ldots, n-1\}$.

We shall use the notation introduced in Definition 6.132. We know that B is an $n \times (n-1)$-matrix (since $B \in \mathbb{K}^{n\times(n-1)}$). Thus, $\left(B \mid B_{\bullet,q}\right)$ is an $n \times n$-matrix. This $n \times n$-matrix $\left(B \mid B_{\bullet,q}\right)$ is defined as the $n \times ((n-1)+1)$-matrix whose columns are $B_{\bullet,1}, B_{\bullet,2}, \ldots, B_{\bullet,n-1}, B_{\bullet,q}$; thus, it has two equal columns (indeed, the column vector $B_{\bullet,q}$ appears twice among the columns $B_{\bullet,1}, B_{\bullet,2}, \ldots, B_{\bullet,n-1}, B_{\bullet,q}$). Thus, Exercise 6.7 **(f)** (applied to $A = \left(B \mid B_{\bullet,q}\right)$) shows that $\det\left(B \mid B_{\bullet,q}\right) = 0$.

But $B_{\bullet,q}$ is the q-th column of the matrix B (by the definition of $B_{\bullet,q}$). Thus,

$$\begin{aligned} B_{\bullet,q} &= (\text{the } q\text{-th column of the matrix } B) \\ &= \begin{pmatrix} b_{1,q} \\ b_{2,q} \\ \vdots \\ b_{n,q} \end{pmatrix} \qquad \left(\text{since } B = \left(b_{i,j}\right)_{1\le i\le n,\ 1\le j\le n-1}\right) \\ &= \left(b_{1,q}, b_{2,q}, \ldots, b_{n,q}\right)^T. \end{aligned}$$

Hence, Proposition 6.135 **(a)** (applied to B, $B_{\bullet,q}$ and $b_{i,q}$ instead of A, v and v_i) yields

$$\det\left(B \mid B_{\bullet,q}\right) = \sum_{i=1}^{n} (-1)^{n+i} b_{i,q} \det\left(B_{\sim i,\bullet}\right).$$

Comparing this with $\det\left(B \mid B_{\bullet,q}\right) = 0$, we obtain

$$0 = \sum_{i=1}^{n} (-1)^{n+i} b_{i,q} \det\left(B_{\sim i,\bullet}\right).$$

Multiplying both sides of this equality by $(-1)^n$, we obtain

$$\begin{aligned} 0 &= (-1)^n \sum_{i=1}^{n} (-1)^{n+i} b_{i,q} \det\left(B_{\sim i,\bullet}\right) \\ &= \sum_{i=1}^{n} \underbrace{(-1)^n (-1)^{n+i}}_{\substack{=(-1)^{n+(n+i)}=(-1)^i \\ \text{(since } n+(n+i)=2n+i\equiv i \bmod 2)}} \underbrace{b_{i,q} \det\left(B_{\sim i,\bullet}\right)}_{=\det\left(B_{\sim i,\bullet}\right) b_{i,q}} \\ &= \sum_{i=1}^{n} (-1)^i \det\left(B_{\sim i,\bullet}\right) b_{i,q} = \sum_{r=1}^{n} (-1)^r \det\left(B_{\sim r,\bullet}\right) b_{r,q} \end{aligned}$$

(here, we renamed the summation index i as r). This proves Proposition 6.144 **(b)**.

(a) Write the matrix B in the form $B = \left(b_{i,j}\right)_{1\leq i\leq n,\ 1\leq j\leq n-1}$. For every $r \in \{1, 2, \ldots, n\}$, we have

$$\begin{aligned} B_{r,\bullet} &= (\text{the } r\text{-th row of the matrix } B) \\ &\qquad \left(\begin{array}{c}\text{since } B_{r,\bullet} \text{ is the } r\text{-th row of the matrix } B \\ \text{(by the definition of } B_{r,\bullet}\text{)}\end{array}\right) \\ &= \left(b_{r,1}, b_{r,2}, \ldots, b_{r,n-1}\right) \qquad \left(\text{since } B = \left(b_{i,j}\right)_{1\leq i\leq n,\ 1\leq j\leq n-1}\right) \\ &= \left(b_{r,j}\right)_{1\leq i\leq 1,\ 1\leq j\leq n-1}. \end{aligned}$$

Thus,

$$\begin{aligned} &\sum_{r=1}^{n} (-1)^r \det\left(B_{\sim r,\bullet}\right) \underbrace{B_{r,\bullet}}_{=\left(b_{r,j}\right)_{1\leq i\leq 1,\ 1\leq j\leq n-1}} \\ &= \sum_{r=1}^{n} \underbrace{(-1)^r \det\left(B_{\sim r,\bullet}\right) \left(b_{r,j}\right)_{1\leq i\leq 1,\ 1\leq j\leq n-1}}_{=\left((-1)^r \det\left(B_{\sim r,\bullet}\right) b_{r,j}\right)_{1\leq i\leq 1,\ 1\leq j\leq n-1}} = \sum_{r=1}^{n} \left((-1)^r \det\left(B_{\sim r,\bullet}\right) b_{r,j}\right)_{1\leq i\leq 1,\ 1\leq j\leq n-1} \\ &= \left(\underbrace{\sum_{r=1}^{n} (-1)^r \det\left(B_{\sim r,\bullet}\right) b_{r,j}}_{\substack{=0 \\ \text{(by Proposition 6.144 \textbf{(b)}, applied to } q=j)}}\right)_{1\leq i\leq 1,\ 1\leq j\leq n-1} = (0)_{1\leq i\leq 1,\ 1\leq j\leq n-1} = 0_{1\times(n-1)}. \end{aligned}$$

This proves Proposition 6.144 **(a)**. □

Let us next state a variant of Proposition 6.144 where rows are replaced by columns (and n is renamed as $n+1$):

Proposition 6.146. Let $n \in \mathbb{N}$. Let $B \in \mathbb{K}^{n\times(n+1)}$. Then:

(a) We have

$$\sum_{r=1}^{n+1} (-1)^r \det\left(B_{\bullet,\sim r}\right) B_{\bullet,r} = 0_{n\times 1}.$$

(b) Write the matrix B in the form $B = \left(b_{i,j}\right)_{1\leq i\leq n,\ 1\leq j\leq n+1}$. Then,

$$\sum_{r=1}^{n+1} (-1)^r \det\left(B_{\bullet,\sim r}\right) b_{q,r} = 0$$

for every $q \in \{1,2,\ldots,n\}$.

Example 6.147. For this example, set $n = 2$ and $B = \begin{pmatrix} a & a' & a'' \\ b & b' & b'' \end{pmatrix}$. Then, Proposition 6.146 **(a)** says that

$$-\det\begin{pmatrix} a' & a'' \\ b' & b'' \end{pmatrix}\cdot\begin{pmatrix} a \\ b \end{pmatrix} + \det\begin{pmatrix} a & a'' \\ b & b'' \end{pmatrix}\cdot\begin{pmatrix} a' \\ b' \end{pmatrix} - \det\begin{pmatrix} a & a' \\ b & b' \end{pmatrix}\cdot\begin{pmatrix} a'' \\ b'' \end{pmatrix}$$
$$= 0_{2\times 1}.$$

Proposition 6.146 **(b)** (applied to $q = 2$) yields

$$-\det\begin{pmatrix} a' & a'' \\ b' & b'' \end{pmatrix}\cdot b + \det\begin{pmatrix} a & a'' \\ b & b'' \end{pmatrix}\cdot b' - \det\begin{pmatrix} a & a' \\ b & b' \end{pmatrix}\cdot b'' = 0.$$

We shall obtain Proposition 6.146 by applying Proposition 6.144 to $n+1$ and B^T [278] instead of n and B. For this, we shall need a really simple lemma:

Lemma 6.148. Let $n \in \mathbb{N}$ and $m \in \mathbb{N}$. Let $r \in \{1,2,\ldots,m\}$. Let $B \in \mathbb{K}^{n\times m}$. Then, $\left(B^T\right)_{\sim r,\bullet} = \left(B_{\bullet,\sim r}\right)^T$.

Proof of Lemma 6.148. Taking the transpose of a matrix turns all its columns into rows. Thus, if we remove the r-th **column** from B and then take the transpose of the resulting matrix, then we obtain the same matrix as if we first take the transpose of B and then remove the r-th **row** from it. Translating this statement into formulas, we obtain precisely $\left(B_{\bullet,\sim r}\right)^T = \left(B^T\right)_{\sim r,\bullet}$. Thus, Lemma 6.148 is proven.[279] □

[278] Recall that B^T denotes the transpose of the matrix B (see Definition 6.10).

[279] A more formal proof of this could be given using Proposition 6.79 **(e)**.

Proof of Proposition 6.146. **(b)** Let $q \in \{1, 2, \ldots, n\}$. Thus, $q \in \{1, 2, \ldots, n\} = \{1, 2, \ldots, (n+1) - 1\}$ (since $n = (n+1) - 1$).

We have $B = (b_{i,j})_{1 \leq i \leq n,\ 1 \leq j \leq n+1}$. Thus, the definition of B^T yields

$$B^T = (b_{j,i})_{1 \leq i \leq n+1,\ 1 \leq j \leq n} = (b_{j,i})_{1 \leq i \leq n+1,\ 1 \leq j \leq (n+1)-1}$$

(since $n = (n+1) - 1$). Also, $B^T = (b_{j,i})_{1 \leq i \leq n+1,\ 1 \leq j \leq (n+1)-1} \in \mathbb{K}^{(n+1) \times ((n+1)-1)}$. Thus, Proposition 6.144 **(b)** (applied to $n+1$, B^T and $b_{j,i}$ instead of n, B and $b_{i,j}$) yields

$$\sum_{r=1}^{n+1} (-1)^r \det\left(\left(B^T\right)_{\sim r, \bullet}\right) b_{q,r} = 0. \tag{469}$$

But every $r \in \{1, 2, \ldots, n+1\}$ satisfies

$$\det\left(\left(B^T\right)_{\sim r, \bullet}\right) = \det\left(B_{\bullet, \sim r}\right) \tag{470}$$

[280]. Hence, (469) yields

$$0 = \sum_{r=1}^{n+1} (-1)^r \underbrace{\det\left(\left(B^T\right)_{\sim r, \bullet}\right)}_{\substack{=\det(B_{\bullet,\sim r}) \\ \text{(by (470))}}} b_{q,r} = \sum_{r=1}^{n+1} (-1)^r \det\left(B_{\bullet, \sim r}\right) b_{q,r}.$$

This proves Proposition 6.146 **(b)**.

(a) Write the matrix B in the form $B = (b_{i,j})_{1 \leq i \leq n,\ 1 \leq j \leq n+1}$. For every $r \in \{1, 2, \ldots, n+1\}$, we have

$$\begin{aligned} B_{\bullet, r} &= (\text{the } r\text{-th column of the matrix } B) \\ &\qquad \left(\begin{array}{c} \text{since } B_{\bullet,r} \text{ is the } r\text{-th column of the matrix } B \\ \text{(by the definition of } B_{\bullet,r}\text{)} \end{array}\right) \\ &= \begin{pmatrix} b_{1,r} \\ b_{2,r} \\ \vdots \\ b_{n,r} \end{pmatrix} \qquad \left(\text{since } B = (b_{i,j})_{1 \leq i \leq n,\ 1 \leq j \leq n+1}\right) \\ &= (b_{i,r})_{1 \leq i \leq n,\ 1 \leq j \leq 1}. \end{aligned}$$

[280] *Proof of (470):* Let $r \in \{1, 2, \ldots, n+1\}$. Then, $B_{\bullet, \sim r} \in \mathbb{K}^{n \times n}$ (since $B \in \mathbb{K}^{n \times (n+1)}$). In other words, $B_{\bullet, \sim r}$ is an $n \times n$-matrix. Thus, Exercise 6.4 (applied to $A = B_{\bullet, \sim r}$) yields $\det\left((B_{\bullet, \sim r})^T\right) = \det(B_{\bullet, \sim r})$.

But Lemma 6.148 (applied to $m = n+1$) yields $\left(B^T\right)_{\sim r, \bullet} = (B_{\bullet, \sim r})^T$. Thus, $\det\left(\underbrace{\left(B^T\right)_{\sim r, \bullet}}_{=(B_{\bullet,\sim r})^T}\right) = \det\left((B_{\bullet, \sim r})^T\right) = \det(B_{\bullet, \sim r})$. This proves (470).

Thus,

$$\begin{aligned}
&\sum_{r=1}^{n+1} (-1)^r \det(B_{\bullet,\sim r}) \underbrace{B_{\bullet,r}}_{=(b_{i,r})_{1\leq i\leq n,\ 1\leq j\leq 1}} \\
&= \sum_{r=1}^{n+1} \underbrace{(-1)^r \det(B_{\bullet,\sim r}) (b_{i,r})_{1\leq i\leq n,\ 1\leq j\leq 1}}_{=((-1)^r \det(B_{\bullet,\sim r}) b_{i,r})_{1\leq i\leq n,\ 1\leq j\leq 1}} = \sum_{r=1}^{n+1} \left((-1)^r \det(B_{\bullet,\sim r}) b_{i,r}\right)_{1\leq i\leq n,\ 1\leq j\leq 1} \\
&= \left(\underbrace{\sum_{r=1}^{n+1} (-1)^r \det(B_{\bullet,\sim r}) b_{i,r}}_{\substack{=0 \\ \text{(by Proposition 6.146 } \textbf{(b)}\text{, applied to } q=i\text{)}}} \right)_{1\leq i\leq n,\ 1\leq j\leq 1} = (0)_{1\leq i\leq n,\ 1\leq j\leq 1} = 0_{n\times 1}.
\end{aligned}$$

This proves Proposition 6.146 **(a)**. □

Remark 6.149. Proposition 6.146 **(a)** can be viewed as a restatement of Cramer's rule (Theorem 6.120 **(a)**). More precisely, it is easy to derive one of these two facts from the other (although neither of the two is difficult to prove to begin with). Let us sketch one direction of this argument: namely, let us derive Theorem 6.120 **(a)** from Proposition 6.146 **(a)**.

Indeed, let n, A, $b = (b_1, b_2, \ldots, b_n)^T$ and $A_j^\#$ be as in Theorem 6.120 **(a)**. We want to prove that $A \cdot \left(\det\left(A_1^\#\right), \det\left(A_2^\#\right), \ldots, \det\left(A_n^\#\right)\right)^T = \det A \cdot b$.

Let $B = (A \mid b)$ (using the notations of Definition 6.132); this is an $n \times (n+1)$-matrix.

Fix $r \in \{1, 2, \ldots, n\}$. The matrix $B_{\bullet,\sim r}$ differs from the matrix $A_r^\#$ only in the order of its columns: More precisely,

- the matrix $B_{\bullet,\sim r}$ is obtained from the matrix A by removing the r-th column and attaching the column vector b to the right edge, whereas
- the matrix $A_r^\#$ is obtained from the matrix A by replacing the r-th column by the column vector b.

Thus, the matrix $B_{\bullet,\sim r}$ can be obtained from the matrix $A_r^\#$ by first swapping the r-th and $(r+1)$-th columns, then swapping the $(r+1)$-th and $(r+2)$-th columns, etc., until finally swapping the $(n-1)$-th and n-th columns. Each of these swaps multiplies the determinant by -1 (by Exercise 6.7 **(b)**); thus, our sequence of swaps multiplies the determinant by $(-1)^{n-r} = (-1)^{n+r}$. Hence,

$$\det(B_{\bullet,\sim r}) = (-1)^{n+r} \det\left(A_r^\#\right). \qquad (471)$$

Now, forget that we fixed r. It is easy to see that every $(v_1, v_2, \ldots, v_n)^T \in \mathbb{K}^{1\times n}$ satisfies

$$A \cdot (v_1, v_2, \ldots, v_n)^T = \sum_{r=1}^{n} v_r A_{\bullet,r}.$$

Applying this to $(v_1, v_2, \ldots, v_n)^T = \left(\det\left(A_1^{\#}\right), \det\left(A_2^{\#}\right), \ldots, \det\left(A_n^{\#}\right)\right)^T$, we obtain

$$A \cdot \left(\det\left(A_1^{\#}\right), \det\left(A_2^{\#}\right), \ldots, \det\left(A_n^{\#}\right)\right)^T = \sum_{r=1}^{n} \det\left(A_r^{\#}\right) A_{\bullet,r}. \tag{472}$$

But Proposition 6.146 **(a)** yields

$$\begin{aligned}
0_{n\times 1} &= \sum_{r=1}^{n+1} (-1)^r \det\left(B_{\bullet,\sim r}\right) B_{\bullet,r} \\
&= \sum_{r=1}^{n} (-1)^r \underbrace{\det\left(B_{\bullet,\sim r}\right)}_{\substack{=(-1)^{n+r}\det\left(A_r^{\#}\right) \\ \text{(by (471))}}} \underbrace{B_{\bullet,r}}_{\substack{=A_{\bullet,r} \\ \text{(since } B=(A|b) \text{ and } r\leq n)}} \\
&\qquad + (-1)^{n+1} \det\left(\underbrace{B_{\bullet,\sim(n+1)}}_{\substack{=A \\ \text{(since } B=(A|b))}}\right) \underbrace{B_{\bullet,n+1}}_{\substack{=b \\ \text{(since } B=(A|b))}} \\
&= \sum_{r=1}^{n} \underbrace{(-1)^r (-1)^{n+r}}_{=(-1)^n} \det\left(A_r^{\#}\right) A_{\bullet,r} + \underbrace{(-1)^{n+1}}_{=-(-1)^n} \det A \cdot b \\
&= \sum_{r=1}^{n} (-1)^n \det\left(A_r^{\#}\right) A_{\bullet,r} - (-1)^n \det A \cdot b \\
&= (-1)^n \left(\sum_{r=1}^{n} \det\left(A_r^{\#}\right) A_{\bullet,r} - \det A \cdot b\right).
\end{aligned}$$

Multiplying both sides of this equality by $(-1)^n$, we obtain

$$\begin{aligned}
0_{n\times 1} &= \underbrace{(-1)^n (-1)^n}_{=1} \left(\sum_{r=1}^{n} \det\left(A_r^{\#}\right) A_{\bullet,r} - \det A \cdot b\right) \\
&= \sum_{r=1}^{n} \det\left(A_r^{\#}\right) A_{\bullet,r} - \det A \cdot b.
\end{aligned}$$

Hence,

$$\det A \cdot b = \sum_{r=1}^{n} \det\left(A_r^{\#}\right) A_{\bullet,r} = A \cdot \left(\det\left(A_1^{\#}\right), \det\left(A_2^{\#}\right), \ldots, \det\left(A_n^{\#}\right)\right)^T$$

(by (472)). Thus, we have derived Theorem 6.120 **(a)** from Proposition 6.146 **(a)**. Essentially the same argument (but read backwards) can be used to derive Proposition 6.146 **(a)** from Theorem 6.120 **(a)**.

Now, we can easily prove the *Plücker identity*:

Theorem 6.150. Let n be a positive integer. Let $A \in \mathbb{K}^{n\times(n-1)}$ and $B \in \mathbb{K}^{n\times(n+1)}$. Then,

$$\sum_{r=1}^{n+1} (-1)^r \det\left(A \mid B_{\bullet,r}\right) \det\left(B_{\bullet,\sim r}\right) = 0$$

(where we are using the notations from Definition 6.132 and from Definition 6.128).

Example 6.151. If $n = 3$, $A = \begin{pmatrix} a & a' \\ b & b' \\ c & c' \end{pmatrix}$ and $B = \begin{pmatrix} x_1 & x_2 & x_3 & x_4 \\ y_1 & y_2 & y_3 & y_4 \\ z_1 & z_2 & z_3 & z_4 \end{pmatrix}$, then Theorem 6.150 says that

$$\begin{aligned}
& - \det \begin{pmatrix} a & a' & x_1 \\ b & b' & y_1 \\ c & c' & z_1 \end{pmatrix} \det \begin{pmatrix} x_2 & x_3 & x_4 \\ y_2 & y_3 & y_4 \\ z_2 & z_3 & z_4 \end{pmatrix} \\
& \qquad + \det \begin{pmatrix} a & a' & x_2 \\ b & b' & y_2 \\ c & c' & z_2 \end{pmatrix} \det \begin{pmatrix} x_1 & x_3 & x_4 \\ y_1 & y_3 & y_4 \\ z_1 & z_3 & z_4 \end{pmatrix} \\
& \qquad - \det \begin{pmatrix} a & a' & x_3 \\ b & b' & y_3 \\ c & c' & z_3 \end{pmatrix} \det \begin{pmatrix} x_1 & x_2 & x_4 \\ y_1 & y_2 & y_4 \\ z_1 & z_2 & z_4 \end{pmatrix} \\
& \qquad + \det \begin{pmatrix} a & a' & x_4 \\ b & b' & y_4 \\ c & c' & z_4 \end{pmatrix} \det \begin{pmatrix} x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \\ z_1 & z_2 & z_3 \end{pmatrix} \\
& = 0.
\end{aligned}$$

Proof of Theorem 6.150. Write the matrix B in the form $B = \left(b_{i,j}\right)_{1\leq i\leq n,\ 1\leq j\leq n+1}$.

Let $r \in \{1, 2, \ldots, n+1\}$. Then,

$$\begin{aligned} B_{\bullet,r} &= (\text{the } r\text{-th column of the matrix } B) \\ &\qquad \left(\begin{array}{c}\text{since } B_{\bullet,r} \text{ is the } r\text{-th column of the matrix } B \\ \text{(by the definition of } B_{\bullet,r}\text{)}\end{array}\right) \\ &= \begin{pmatrix} b_{1,r} \\ b_{2,r} \\ \vdots \\ b_{n,r} \end{pmatrix} \qquad \left(\text{since } B = \left(b_{i,j}\right)_{1\leq i\leq n,\ 1\leq j\leq n+1}\right) \\ &= \left(b_{1,r}, b_{2,r}, \ldots, b_{n,r}\right)^T . \end{aligned}$$

Hence, Proposition 6.135 **(a)** (applied to $B_{\bullet,r}$ and $b_{i,r}$ instead of v and v_i) shows that

$$\det\left(A \mid B_{\bullet,r}\right) = \sum_{i=1}^{n} (-1)^{n+i} b_{i,r} \det\left(A_{\sim i,\bullet}\right). \qquad (473)$$

Now, forget that we fixed r. We thus have proven (473) for each $r \in \{1, 2, \ldots, n+1\}$. Now,

$$\begin{aligned} &\sum_{r=1}^{n+1} (-1)^r \underbrace{\det\left(A \mid B_{\bullet,r}\right)}_{\substack{=\sum_{i=1}^{n}(-1)^{n+i} b_{i,r}\det\left(A_{\sim i,\bullet}\right) \\ \text{(by (473))}}} \det\left(B_{\bullet,\sim r}\right) \\ &= \sum_{r=1}^{n+1} (-1)^r \left(\sum_{i=1}^{n} (-1)^{n+i} b_{i,r} \det\left(A_{\sim i,\bullet}\right)\right) \det\left(B_{\bullet,\sim r}\right) \\ &= \underbrace{\sum_{r=1}^{n+1} \sum_{i=1}^{n}}_{=\sum_{i=1}^{n}\sum_{r=1}^{n+1}} \underbrace{(-1)^r (-1)^{n+i} b_{i,r} \det\left(A_{\sim i,\bullet}\right) \det\left(B_{\bullet,\sim r}\right)}_{=(-1)^r \det\left(B_{\bullet,\sim r}\right) b_{i,r} \cdot (-1)^{n+i} \det\left(A_{\sim i,\bullet}\right)} \\ &= \sum_{i=1}^{n} \sum_{r=1}^{n+1} (-1)^r \det\left(B_{\bullet,\sim r}\right) b_{i,r} \cdot (-1)^{n+i} \det\left(A_{\sim i,\bullet}\right) \\ &= \sum_{i=1}^{n} \underbrace{\left(\sum_{r=1}^{n+1} (-1)^r \det\left(B_{\bullet,\sim r}\right) b_{i,r}\right)}_{\substack{=0 \\ \text{(by Proposition 6.146 \textbf{(b)},} \\ \text{applied to } q=i\text{)}}} (-1)^{n+i} \det\left(A_{\sim i,\bullet}\right) \\ &= \sum_{i=1}^{n} 0 (-1)^{n+i} \det\left(A_{\sim i,\bullet}\right) = 0. \end{aligned}$$

This proves Theorem 6.150. □

Remark 6.152. Theorem 6.150 (at least in the case when $\mathbb{K}$ is a field) is essentially equivalent to [Fulton97, §9.1, Exercise 1], to [KleLak72, (QR)], to [Jacobs10, Theorem 3.4.11 (the "necessary" part)], and to [Lampe13, Proposition 3.3.2 (the "only if" part)].

Exercise 6.43. Use Theorem 6.150 to give a new proof of Proposition 6.137.

6.22. Laplace expansion in multiple rows/columns

In this section, we shall see a (somewhat unwieldy, but classical and important) generalization of Theorem 6.82. First, we shall need some notations:

Definition 6.153. Throughout Section 6.22, we shall use the following notations:

- If I is a finite set of integers, then $\sum I$ shall denote the sum of all elements of I. (Thus, $\sum I = \sum\limits_{i \in I} i$.)
- If I is a finite set of integers, then $w(I)$ shall denote the list of all elements of I in increasing order (with no repetitions). (See Definition 2.50 for the formal definition of this list.) (For example, $w(\{3,4,8\}) = (3,4,8)$.)

We shall also use the notation introduced in Definition 6.78. If n, m, A, $(i_1, i_2, \ldots, i_u)$ and $(j_1, j_2, \ldots, j_v)$ are as in Definition 6.78, then we shall use the notation $\operatorname{sub}_{(i_1,i_2,\ldots,i_u)}^{(j_1,j_2,\ldots,j_v)} A$ as a synonym for $\operatorname{sub}_{i_1,i_2,\ldots,i_u}^{j_1,j_2,\ldots,j_v} A$.

A consequence of this definition is that if A is an $n \times m$-matrix, and if U is a subset of $\{1,2,\ldots,n\}$, and if V is a subset of $\{1,2,\ldots,m\}$, then $\operatorname{sub}_{w(U)}^{w(V)} A$ is a well-defined $|U| \times |V|$-matrix[281] (actually, a submatrix of A).

Example 6.154. If $n = 3$ and $m = 4$ and $A = \begin{pmatrix} a & b & c & d \\ a' & b' & c' & d' \\ a'' & b'' & c'' & d'' \end{pmatrix}$, then $\operatorname{sub}_{w(\{2,3\})}^{w(\{1,3,4\})} A = \begin{pmatrix} a' & c' & d' \\ a'' & c'' & d'' \end{pmatrix}$.

The following fact is obvious from the definition of $w(I)$:

Proposition 6.155. Let I be a finite set of integers. Then, $w(I)$ is an $|I|$-tuple of elements of I.

[281] because $w(U)$ is a list of $|U|$ elements of $\{1,2,\ldots,n\}$, and because $w(V)$ is a list of $|V|$ elements of $\{1,2,\ldots,m\}$

Theorem 6.156. Let $n \in \mathbb{N}$. Let $A \in \mathbb{K}^{n \times n}$. For any subset I of $\{1, 2, \ldots, n\}$, we let $\widetilde{I}$ denote the complement $\{1, 2, \ldots, n\} \setminus I$ of I. (For instance, if $n = 4$ and $I = \{1, 4\}$, then $\widetilde{I} = \{2, 3\}$.)

(a) For every subset P of $\{1, 2, \ldots, n\}$, we have

$$\det A = \sum_{\substack{Q \subseteq \{1,2,\ldots,n\}; \\ |Q| = |P|}} (-1)^{\sum P + \sum Q} \det \left(\operatorname{sub}_{w(P)}^{w(Q)} A \right) \det \left(\operatorname{sub}_{w\left(\widetilde{P}\right)}^{w\left(\widetilde{Q}\right)} A \right).$$

(b) For every subset Q of $\{1, 2, \ldots, n\}$, we have

$$\det A = \sum_{\substack{P \subseteq \{1,2,\ldots,n\}; \\ |P| = |Q|}} (-1)^{\sum P + \sum Q} \det \left(\operatorname{sub}_{w(P)}^{w(Q)} A \right) \det \left(\operatorname{sub}_{w\left(\widetilde{P}\right)}^{w\left(\widetilde{Q}\right)} A \right).$$

Theorem 6.156 is actually a generalization of Theorem 6.82, known as "Laplace expansion in multiple rows (resp. columns)". It appears (for example) in [Willia18, Theorem 3.61] and in [Prasol94, Theorem 2.4.1].[282] Theorem 6.82 **(a)** can be recovered from Theorem 6.156 **(a)** by setting $P = \{p\}$; similarly for the part **(b)**.

Example 6.157. Let us see what Theorem 6.156 **(a)** says in a simple case. For this example, set $n = 4$ and $A = \left(a_{i,j}\right)_{1 \leq i \leq 4,\ 1 \leq j \leq 4}$. Also, set $P = \{1, 4\} \subseteq \{1, 2, 3, 4\}$; thus, $w(P) = (1, 4)$, $\sum P = 1 + 4 = 5$, $|P| = 2$, $\widetilde{P} = \{2, 3\}$ and $w\left(\widetilde{P}\right) = (2, 3)$.

[282] Of course, parts **(a)** and **(b)** of Theorem 6.156 are easily seen to be equivalent; thus, many authors confine themselves to only stating one of them. For example, Theorem 6.156 is [CaSoSp12, Lemma A.1 (f)].

Now, Theorem 6.156 **(a)** says that

$$\begin{aligned}
&\det A\\
&= \sum_{\substack{Q\subseteq\{1,2,3,4\};\\ |Q|=2}} (-1)^{5+\sum Q} \det\left(\operatorname{sub}_{1,4}^{w(Q)} A\right) \det\left(\operatorname{sub}_{2,3}^{w(\widetilde{Q})} A\right)\\
&= (-1)^{5+(1+2)} \det\left(\operatorname{sub}_{1,4}^{1,2} A\right) \det\left(\operatorname{sub}_{2,3}^{3,4} A\right)\\
&\qquad + (-1)^{5+(1+3)} \det\left(\operatorname{sub}_{1,4}^{1,3} A\right) \det\left(\operatorname{sub}_{2,3}^{2,4} A\right)\\
&\qquad + (-1)^{5+(1+4)} \det\left(\operatorname{sub}_{1,4}^{1,4} A\right) \det\left(\operatorname{sub}_{2,3}^{2,3} A\right)\\
&\qquad + (-1)^{5+(2+3)} \det\left(\operatorname{sub}_{1,4}^{2,3} A\right) \det\left(\operatorname{sub}_{2,3}^{1,4} A\right)\\
&\qquad + (-1)^{5+(2+4)} \det\left(\operatorname{sub}_{1,4}^{2,4} A\right) \det\left(\operatorname{sub}_{2,3}^{1,3} A\right)\\
&\qquad + (-1)^{5+(3+4)} \det\left(\operatorname{sub}_{1,4}^{3,4} A\right) \det\left(\operatorname{sub}_{2,3}^{1,2} A\right)\\
&= \det\begin{pmatrix} a_{1,1} & a_{1,2} \\ a_{4,1} & a_{4,2} \end{pmatrix} \det\begin{pmatrix} a_{2,3} & a_{2,4} \\ a_{3,3} & a_{3,4} \end{pmatrix}\\
&\qquad - \det\begin{pmatrix} a_{1,1} & a_{1,3} \\ a_{4,1} & a_{4,3} \end{pmatrix} \det\begin{pmatrix} a_{2,2} & a_{2,4} \\ a_{3,2} & a_{3,4} \end{pmatrix}\\
&\qquad + \det\begin{pmatrix} a_{1,1} & a_{1,4} \\ a_{4,1} & a_{4,4} \end{pmatrix} \det\begin{pmatrix} a_{2,2} & a_{2,3} \\ a_{3,2} & a_{3,3} \end{pmatrix}\\
&\qquad + \det\begin{pmatrix} a_{1,2} & a_{1,3} \\ a_{4,2} & a_{4,3} \end{pmatrix} \det\begin{pmatrix} a_{2,1} & a_{2,4} \\ a_{3,1} & a_{3,4} \end{pmatrix}\\
&\qquad - \det\begin{pmatrix} a_{1,2} & a_{1,4} \\ a_{4,2} & a_{4,4} \end{pmatrix} \det\begin{pmatrix} a_{2,1} & a_{2,3} \\ a_{3,1} & a_{3,3} \end{pmatrix}\\
&\qquad + \det\begin{pmatrix} a_{1,3} & a_{1,4} \\ a_{4,3} & a_{4,4} \end{pmatrix} \det\begin{pmatrix} a_{2,1} & a_{2,2} \\ a_{3,1} & a_{3,2} \end{pmatrix}.
\end{aligned}$$

The following lemma will play a crucial role in our proof of Theorem 6.156 (similar to the role that Lemma 6.84 played in our proof of Theorem 6.82):

Lemma 6.158. Let $n \in \mathbb{N}$. For any subset I of $\{1,2,\ldots,n\}$, we let $\widetilde{I}$ denote the complement $\{1,2,\ldots,n\} \setminus I$ of I.

Let $A = \left(a_{i,j}\right)_{1\leq i\leq n,\ 1\leq j\leq n}$ and $B = \left(b_{i,j}\right)_{1\leq i\leq n,\ 1\leq j\leq n}$ be two $n \times n$-matrices. Let

P and Q be two subsets of $\{1,2,\ldots,n\}$ such that $|P| = |Q|$. Then,

$$\sum_{\substack{\sigma \in S_n;\\ \sigma(P)=Q}} (-1)^{\sigma} \left(\prod_{i\in P} a_{i,\sigma(i)}\right) \left(\prod_{i\in\widetilde{P}} b_{i,\sigma(i)}\right)$$
$$= (-1)^{\sum P + \sum Q} \det\left(\operatorname{sub}_{w(P)}^{w(Q)} A\right) \det\left(\operatorname{sub}_{w(\widetilde{P})}^{w(\widetilde{Q})} B\right).$$

The proof of Lemma 6.158 is similar (in its spirit) to the proof of Lemma 6.84, but it requires a lot more bookkeeping (if one wants to make it rigorous). This proof shall be given in the solution to Exercise 6.44:

Exercise 6.44. Prove Lemma 6.158 and Theorem 6.156.

[Hint: First, prove Lemma 6.158 in the case when $P = \{1,2,\ldots,k\}$ for some $k \in \{0,1,\ldots,n\}$; in order to do so, use the bijection from Exercise 5.14 **(c)** (applied to $I = Q$). Then, derive the general case of Lemma 6.158 by permuting the rows of the matrices. Finally, prove Theorem 6.156.]

The following exercise generalizes Proposition 6.96 in the same way as Theorem 6.156 generalizes Theorem 6.82:

Exercise 6.45. Let $n \in \mathbb{N}$. Let $A \in \mathbb{K}^{n\times n}$. For any subset I of $\{1,2,\ldots,n\}$, we let $\widetilde{I}$ denote the complement $\{1,2,\ldots,n\} \setminus I$ of I. Let R be a subset of $\{1,2,\ldots,n\}$. Prove the following:

(a) For every subset P of $\{1,2,\ldots,n\}$ satisfying $|P| = |R|$ and $P \neq R$, we have

$$0 = \sum_{\substack{Q\subseteq\{1,2,\ldots,n\};\\ |Q|=|P|}} (-1)^{\sum P + \sum Q} \det\left(\operatorname{sub}_{w(R)}^{w(Q)} A\right) \det\left(\operatorname{sub}_{w(\widetilde{P})}^{w(\widetilde{Q})} A\right).$$

(b) For every subset Q of $\{1,2,\ldots,n\}$ satisfying $|Q| = |R|$ and $Q \neq R$, we have

$$0 = \sum_{\substack{P\subseteq\{1,2,\ldots,n\};\\ |P|=|Q|}} (-1)^{\sum P + \sum Q} \det\left(\operatorname{sub}_{w(P)}^{w(R)} A\right) \det\left(\operatorname{sub}_{w(\widetilde{P})}^{w(\widetilde{Q})} A\right).$$

Exercise 6.45 can be generalized to non-square matrices:

Exercise 6.46. Let $n \in \mathbb{N}$ and $m \in \mathbb{N}$. For any subset I of $\{1,2,\ldots,n\}$, we let $\widetilde{I}$ denote the complement $\{1,2,\ldots,n\} \setminus I$ of I. Let J and K be two subsets of $\{1,2,\ldots,m\}$ satisfying $|J| + |K| = n$ and $J \cap K \neq \varnothing$. Prove the following:

(a) For every $A \in \mathbb{K}^{m\times n}$, we have

$$0 = \sum_{\substack{Q\subseteq\{1,2,\ldots,n\};\\ |Q|=|J|}} (-1)^{\sum Q} \det\left(\operatorname{sub}_{w(J)}^{w(Q)} A\right) \det\left(\operatorname{sub}_{w(K)}^{w(\widetilde{Q})} A\right).$$

(b) For every $A \in \mathbb{K}^{n\times m}$, we have

$$0 = \sum_{\substack{P\subseteq\{1,2,\ldots,n\};\\ |P|=|J|}} (-1)^{\sum P} \det\left(\operatorname{sub}_{w(P)}^{w(J)} A\right) \det\left(\operatorname{sub}_{w(\widetilde{P})}^{w(K)} A\right).$$

(Exercise 6.45 can be obtained as a particular case of Exercise 6.46; we leave the details to the reader.)

The following exercise gives a first application of Theorem 6.156 (though it can also be solved with more elementary methods):

Exercise 6.47. Let $n \in \mathbb{N}$. Let P and Q be two subsets of $\{1,2,\ldots,n\}$. Let $A = \left(a_{i,j}\right)_{1\leq i\leq n,\ 1\leq j\leq n} \in \mathbb{K}^{n\times n}$ be an $n \times n$-matrix such that

$$\text{every } i \in P \text{ and } j \in Q \text{ satisfy } a_{i,j} = 0. \tag{474}$$

For any subset I of $\{1,2,\ldots,n\}$, we let $\widetilde{I}$ denote the complement $\{1,2,\ldots,n\} \setminus I$ of I. (For instance, if $n = 4$ and $I = \{1,4\}$, then $\widetilde{I} = \{2,3\}$.)

Prove the following:

(a) If $|P| + |Q| > n$, then $\det A = 0$.

(b) If $|P| + |Q| = n$, then

$$\det A = (-1)^{\sum P + \sum \widetilde{Q}} \det\left(\operatorname{sub}_{w(P)}^{w(\widetilde{Q})} A\right) \det\left(\operatorname{sub}_{w(\widetilde{P})}^{w(Q)} A\right).$$

Example 6.159. **(a)** Applying Exercise 6.47 **(a)** to $n = 5$, $P = \{1,3,5\}$ and $Q = \{2,3,4\}$, we see that

$$\det\begin{pmatrix} a & 0 & 0 & 0 & b \\ c & d & e & f & g \\ h & 0 & 0 & 0 & i \\ j & k & \ell & m & n \\ o & 0 & 0 & 0 & p \end{pmatrix} = 0$$

(don't mistake the letter "o" for a zero) for any $a,b,c,d,e,f,g,h,i,j,k,\ell,m,n,o,p \in \mathbb{K}$ (since the $n \times n$-matrices $A = \left(a_{i,j}\right)_{1\leq i\leq n,\ 1\leq j\leq n}$ satisfying (474) are precisely the matrices of the form $\begin{pmatrix} a & 0 & 0 & 0 & b \\ c & d & e & f & g \\ h & 0 & 0 & 0 & i \\ j & k & \ell & m & n \\ o & 0 & 0 & 0 & p \end{pmatrix}$).

(b) Applying Exercise 6.47 **(a)** to $n = 5$, $P = \{2,3,4\}$ and $Q = \{2,3,4\}$, we see that

$$\det \begin{pmatrix} a & b & c & d & e \\ f & 0 & 0 & 0 & g \\ h & 0 & 0 & 0 & i \\ j & 0 & 0 & 0 & k \\ \ell & m & n & o & p \end{pmatrix} = 0$$

(don't mistake the letter "o" for a zero) for any $a, b, c, d, e, f, g, h, i, j, k, \ell, m, n, o, p \in \mathbb{K}$. This is precisely the claim of Exercise 6.6 **(b)**.

(c) Applying Exercise 6.47 **(b)** to $n = 4$, $P = \{2, 3\}$ and $Q = \{2, 3\}$, we see that

$$\det \begin{pmatrix} a & b & c & d \\ \ell & 0 & 0 & e \\ k & 0 & 0 & f \\ j & i & h & g \end{pmatrix} = \det \begin{pmatrix} \ell & e \\ k & f \end{pmatrix} \cdot \det \begin{pmatrix} b & c \\ i & h \end{pmatrix}$$

for all $a, b, c, d, e, f, g, h, i, j, k, \ell \in \mathbb{K}$. This solves Exercise 6.6 **(a)**.

(d) It is not hard to derive Exercise 6.29 by applying Exercise 6.47 **(b)** to $n + m$, $\begin{pmatrix} A & 0_{n\times m} \\ C & D \end{pmatrix}$, $\{1, 2, \ldots, n\}$ and $\{n+1, n+2, \ldots, n+m\}$ instead of n, A, P and Q. Similarly, Exercise 6.30 can be derived from Exercise 6.47 **(b)** as well.

6.23. $\det(A + B)$

As Theorem 6.23 shows, the determinant of the product AB of two square matrices can be easily and neatly expressed through the determinants of A and B. In contrast, the determinant of a sum $A + B$ of two square matrices cannot be expressed in such a way[283]. There is, however, a formula for $\det(A + B)$ in terms of the determinants of submatrices of A and B. While it is rather unwieldy (a far cry from the elegance of Theorem 6.23), it is nevertheless useful sometimes; let us now show it:

Theorem 6.160. Let $n \in \mathbb{N}$. For any subset I of $\{1, 2, \ldots, n\}$, we let $\widetilde{I}$ denote the complement $\{1, 2, \ldots, n\} \setminus I$ of I. (For instance, if $n = 4$ and $I = \{1, 4\}$, then $\widetilde{I} = \{2, 3\}$.) Let us use the notations introduced in Definition 6.78 and in Definition 6.153.

Let A and B be two $n \times n$-matrices. Then,

$$\det(A + B) = \sum_{P \subseteq \{1,2,\ldots,n\}} \sum_{\substack{Q \subseteq \{1,2,\ldots,n\}; \\ |P| = |Q|}} (-1)^{\sum P + \sum Q} \det\left(\operatorname{sub}_{w(P)}^{w(Q)} A\right) \det\left(\operatorname{sub}_{w(\widetilde{P})}^{w(\widetilde{Q})} B\right).$$

[283] It is easy to find two 2×2-matrices A_1 and B_1 and two other 2×2-matrices A_2 and B_2 such that $\det(A_1) = \det(A_2)$ and $\det(B_1) = \det(B_2)$ but $\det(A_1 + B_1) \neq \det(A_2 + B_2)$. This shows that $\det(A + B)$ cannot generally be computed from $\det A$ and $\det B$.

Example 6.161. For this example, set $n = 2$, $A = \left(a_{i,j}\right)_{1\leq i\leq 2,\ 1\leq j\leq 2}$ and $B = \left(b_{i,j}\right)_{1\leq i\leq 2,\ 1\leq j\leq 2}$. Then, Theorem 6.160 says that

$$\begin{aligned}
&\det\left(A+B\right) \\
&= \sum_{P\subseteq\{1,2\}} \sum_{\substack{Q\subseteq\{1,2\};\\ |P|=|Q|}} (-1)^{\sum P+\sum Q} \det\left(\operatorname{sub}_{w(P)}^{w(Q)} A\right) \det\left(\operatorname{sub}_{w\left(\widetilde{P}\right)}^{w\left(\widetilde{Q}\right)} B\right) \\
&= (-1)^{\sum\varnothing+\sum\varnothing} \det \underbrace{\left(\operatorname{sub} A\right)}_{\substack{\text{this is the}\\ 0\times 0\text{-matrix}}} \det\left(\operatorname{sub}_{1,2}^{1,2} B\right) \\
&\qquad + (-1)^{\sum\{1\}+\sum\{1\}} \det\left(\operatorname{sub}_1^1 A\right) \det\left(\operatorname{sub}_2^2 B\right) \\
&\qquad + (-1)^{\sum\{1\}+\sum\{2\}} \det\left(\operatorname{sub}_1^2 A\right) \det\left(\operatorname{sub}_2^1 B\right) \\
&\qquad + (-1)^{\sum\{2\}+\sum\{1\}} \det\left(\operatorname{sub}_2^1 A\right) \det\left(\operatorname{sub}_1^2 B\right) \\
&\qquad + (-1)^{\sum\{2\}+\sum\{2\}} \det\left(\operatorname{sub}_2^2 A\right) \det\left(\operatorname{sub}_1^1 B\right) \\
&\qquad + (-1)^{\sum\{1,2\}+\sum\{1,2\}} \det\left(\operatorname{sub}_{1,2}^{1,2} A\right) \det \underbrace{\left(\operatorname{sub} B\right)}_{\substack{\text{this is the}\\ 0\times 0\text{-matrix}}} \\
&= \underbrace{\det\left(\text{the } 0\times 0\text{-matrix}\right)}_{=1} \det\begin{pmatrix} b_{1,1} & b_{1,2} \\ b_{2,1} & b_{2,2} \end{pmatrix} + \det\begin{pmatrix} a_{1,1} \end{pmatrix} \det\begin{pmatrix} b_{2,2} \end{pmatrix} \\
&\qquad - \det\begin{pmatrix} a_{1,2} \end{pmatrix} \det\begin{pmatrix} b_{2,1} \end{pmatrix} - \det\begin{pmatrix} a_{2,1} \end{pmatrix} \det\begin{pmatrix} b_{1,2} \end{pmatrix} \\
&\qquad + \det\begin{pmatrix} a_{2,2} \end{pmatrix} \det\begin{pmatrix} b_{1,1} \end{pmatrix} + \det\begin{pmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{pmatrix} \underbrace{\det\left(\text{the } 0\times 0\text{-matrix}\right)}_{=1} \\
&= \det\begin{pmatrix} b_{1,1} & b_{1,2} \\ b_{2,1} & b_{2,2} \end{pmatrix} + a_{1,1}b_{2,2} - a_{1,2}b_{2,1} - a_{2,1}b_{1,2} + a_{2,2}b_{1,1} + \det\begin{pmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{pmatrix}.
\end{aligned}$$

Exercise 6.48. Prove Theorem 6.160.
[**Hint:** Use Lemma 6.158.]

Theorem 6.160 takes a simpler form in the particular case when the matrix B is diagonal (i.e., has all entries outside of its diagonal equal to 0):

Corollary 6.162. Let $n \in \mathbb{N}$. For every two objects i and j, define $\delta_{i,j} \in \mathbb{K}$ by $\delta_{i,j} = \begin{cases} 1, & \text{if } i = j; \\ 0, & \text{if } i \neq j \end{cases}.$

Let A be an $n \times n$-matrix. Let $d_1, d_2, \ldots, d_n$ be n elements of $\mathbb{K}$. Let D be the $n \times n$-matrix $\left(d_i \delta_{i,j}\right)_{1 \leq i \leq n,\ 1 \leq j \leq n}$. Then,

$$\det\left(A + D\right) = \sum_{P \subseteq \{1,2,\ldots,n\}} \det\left(\operatorname{sub}_{w(P)}^{w(P)} A\right) \prod_{i \in \{1,2,\ldots,n\} \setminus P} d_i .$$

This corollary can easily be derived from Theorem 6.160 using the following fact:

Lemma 6.163. Let $n \in \mathbb{N}$. For every two objects i and j, define $\delta_{i,j} \in \mathbb{K}$ by $\delta_{i,j} = \begin{cases} 1, & \text{if } i = j; \\ 0, & \text{if } i \neq j \end{cases}$.

Let $d_1, d_2, \ldots, d_n$ be n elements of $\mathbb{K}$. Let D be the $n \times n$-matrix $\left(d_i \delta_{i,j}\right)_{1 \leq i \leq n,\ 1 \leq j \leq n}$. Let P and Q be two subsets of $\{1, 2, \ldots, n\}$ such that $|P| = |Q|$. Then,

$$\det\left(\operatorname{sub}_{w(P)}^{w(Q)} D\right) = \delta_{P,Q} \prod_{i \in P} d_i .$$

Proving Corollary 6.162 and Lemma 6.163 in detail is part of Exercise 6.49 further below.

A particular case of Corollary 6.162 is the following fact:

Corollary 6.164. Let $n \in \mathbb{N}$. Let A be an $n \times n$-matrix. Let $x \in \mathbb{K}$. Then,

$$\det\left(A + xI_n\right) = \sum_{P \subseteq \{1,2,\ldots,n\}} \det\left(\operatorname{sub}_{w(P)}^{w(P)} A\right) x^{n-|P|} \qquad (475)$$

$$= \sum_{k=0}^{n} \left(\sum_{\substack{P \subseteq \{1,2,\ldots,n\}; \\ |P| = n-k}} \det\left(\operatorname{sub}_{w(P)}^{w(P)} A\right) \right) x^k . \qquad (476)$$

Exercise 6.49. Prove Corollary 6.162, Lemma 6.163 and Corollary 6.164.

Remark 6.165. Let $n \in \mathbb{N}$. Let A be an $n \times n$-matrix over the commutative ring $\mathbb{K}$. Consider the commutative ring $\mathbb{K}[X]$ of polynomials in the indeterminate X over $\mathbb{K}$ (that is, polynomials in the indeterminate X with coefficients lying in $\mathbb{K}$). We can then regard A as a matrix over the ring $\mathbb{K}[X]$ as well (because every element of $\mathbb{K}$ can be viewed as a constant polynomial in $\mathbb{K}[X]$).

Consider the $n \times n$-matrix $A + XI_n$ over the commutative ring $\mathbb{K}[X]$. (For example, if $n = 2$ and $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, then $A + XI_n = \begin{pmatrix} a & b \\ c & d \end{pmatrix} + X \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} =$

$\begin{pmatrix} a+X & b \\ c & d+X \end{pmatrix}$. In general, the matrix $A + XI_n$ is obtained from A by adding an X to each diagonal entry.)

The determinant $\det(A + XI_n)$ is a polynomial in $\mathbb{K}[X]$. (For instance, for $n = 2$ and $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, we have

$$\begin{aligned} \det(A + XI_n) &= \det \begin{pmatrix} a+X & b \\ c & d+X \end{pmatrix} = (a+X)(d+X) - bc \\ &= X^2 + (a+d)X + (ad - bc). \end{aligned}$$

)

This polynomial $\det(A + XI_n)$ is a highly important object; it is a close relative of what is called the *characteristic polynomial* of A. (More precisely, the characteristic polynomial of A is either $\det(XI_n - A)$ or $\det(A - XI_n)$, depending on the conventions that one is using; thus, the polynomial $\det(A + XI_n)$ is either the characteristic polynomial of $-A$ or $(-1)^n$ times this characteristic polynomial.) For more about the characteristic polynomial, see [Artin10, Section 4.5] or [Heffer20, Chapter Five, Section II, §3] (or various other texts on linear algebra).

Using Corollary 6.164, we can explicitly compute the coefficients of the polynomial $\det(A + XI_n)$. In fact, (476) (applied to $\mathbb{K}[X]$ and X instead of $\mathbb{K}$ and x) yields

$$\det(A + XI_n) = \sum_{k=0}^{n} \left(\sum_{\substack{P \subseteq \{1,2,\ldots,n\}; \\ |P| = n-k}} \det\left(\operatorname{sub}_{w(P)}^{w(P)} A\right) \right) X^k.$$

Hence, for every $k \in \{0, 1, \ldots, n\}$, the coefficient of X^k in the polynomial $\det(A + XI_n)$ is

$$\sum_{\substack{P \subseteq \{1,2,\ldots,n\}; \\ |P| = n-k}} \det\left(\operatorname{sub}_{w(P)}^{w(P)} A\right).$$

In particular:

- The coefficient of X^n in the polynomial $\det(A + XI_n)$ is

$$\begin{aligned} &\sum_{\substack{P \subseteq \{1,2,\ldots,n\}; \\ |P| = n-n}} \det\left(\operatorname{sub}_{w(P)}^{w(P)} A\right) \\ &= \det \underbrace{\left(\operatorname{sub}_{w(\varnothing)}^{w(\varnothing)} A\right)}_{=(\text{the } 0\times 0\text{-matrix})} \qquad \left(\begin{array}{c} \text{since the only subset } P \text{ of } \{1,2,\ldots,n\} \\ \text{satisfying } |P| = n-n \text{ is the empty set } \varnothing \end{array}\right) \\ &= \det(\text{the } 0 \times 0\text{-matrix}) = 1. \end{aligned}$$

- The coefficient of X^0 in the polynomial $\det(A + XI_n)$ is

$$\begin{aligned}
&\sum_{\substack{P\subseteq\{1,2,\ldots,n\};\\ |P|=n-0}} \det\left(\operatorname{sub}_{w(P)}^{w(P)} A\right)\\
&= \det \underbrace{\left(\operatorname{sub}_{w(\{1,2,\ldots,n\})}^{w(\{1,2,\ldots,n\})} A\right)}_{=\operatorname{sub}_{1,2,\ldots,n}^{1,2,\ldots,n} A = A}\\
&\qquad \left(\begin{array}{c}\text{since the only subset } P \text{ of } \{1,2,\ldots,n\}\\ \text{satisfying } |P| = n-0 \text{ is the set } \{1,2,\ldots,n\}\end{array}\right)\\
&= \det A.
\end{aligned}$$

- Write the matrix A as $A = (a_{i,j})_{1\leq i\leq n,\ 1\leq j\leq n}$. Then, the coefficient of X^{n-1} in the polynomial $\det(A + XI_n)$ is

$$\begin{aligned}
&\sum_{\substack{P\subseteq\{1,2,\ldots,n\};\\ |P|=n-(n-1)}} \det\left(\operatorname{sub}_{w(P)}^{w(P)} A\right)\\
&= \sum_{\substack{P\subseteq\{1,2,\ldots,n\};\\ |P|=1}} \det\left(\operatorname{sub}_{w(P)}^{w(P)} A\right) = \sum_{k=1}^{n} \det \underbrace{\left(\operatorname{sub}_{w(\{k\})}^{w(\{k\})} A\right)}_{\substack{=\operatorname{sub}_k^k A = \left(\begin{array}{c} a_{k,k}\end{array}\right)\\ \text{(this is a } 1\times 1\text{-matrix)}}}\\
&\qquad \left(\begin{array}{c}\text{since the subsets } P \text{ of } \{1,2,\ldots,n\}\\ \text{satisfying } |P| = 1 \text{ are the sets } \{1\},\{2\},\ldots,\{n\}\end{array}\right)\\
&= \sum_{k=1}^{n} \underbrace{\det\left(\begin{array}{c} a_{k,k}\end{array}\right)}_{=a_{k,k}} = \sum_{k=1}^{n} a_{k,k}.
\end{aligned}$$

  In other words, this coefficient is the sum of all diagonal entries of A. This sum is called the *trace* of A, and is denoted by $\operatorname{Tr} A$.

6.24. Some alternating-sum formulas

The next few exercises don't all involve determinants; what they have in common is that they contain alternating sums (i.e., sums where the addend contains a power of -1).

Exercise 6.50. For every $n \in \mathbb{N}$, let $[n]$ denote the set $\{1, 2, \ldots, n\}$.

Let $\mathbb{L}$ be a noncommutative ring. (Keep in mind that our definition of a "noncommutative ring" includes all commutative rings.)

Let $n \in \mathbb{N}$. The summation sign $\sum\limits_{I \subseteq [n]}$ shall mean $\sum\limits_{I \in \mathcal{P}([n])}$, where $\mathcal{P}([n])$ denotes the powerset of $[n]$.

Let $v_1, v_2, \ldots, v_n$ be n elements of $\mathbb{L}$.

(a) Prove that

$$\sum_{I \subseteq [n]} (-1)^{n-|I|} \left(\sum_{i \in I} v_i \right)^m = \sum_{\substack{f:[m] \to [n]; \\ f \text{ is surjective}}} v_{f(1)} v_{f(2)} \cdots v_{f(m)}$$

for each $m \in \mathbb{N}$.

(b) Prove that

$$\sum_{I \subseteq [n]} (-1)^{n-|I|} \left(\sum_{i \in I} v_i \right)^m = 0$$

for each $m \in \{0, 1, \ldots, n-1\}$.

(c) Prove that

$$\sum_{I \subseteq [n]} (-1)^{n-|I|} \left(\sum_{i \in I} v_i \right)^n = \sum_{\sigma \in S_n} v_{\sigma(1)} v_{\sigma(2)} \cdots v_{\sigma(n)}.$$

(d) Now, assume that $\mathbb{L}$ is a **commutative** ring. Prove that

$$\sum_{I \subseteq [n]} (-1)^{n-|I|} \left(\sum_{i \in I} v_i \right)^n = n! \cdot v_1 v_2 \cdots v_n.$$

[**Hint:** First, generalize Lemma 6.22 to the case of a noncommutative ring $\mathbb{K}$.]

Exercise 6.51. For every $n \in \mathbb{N}$, let $[n]$ denote the set $\{1, 2, \ldots, n\}$.

Let $\mathbb{L}$ be a noncommutative ring. (Keep in mind that our definition of a "noncommutative ring" includes all commutative rings.)

Let $n \in \mathbb{N}$. The summation sign $\sum\limits_{I \subseteq [n]}$ shall mean $\sum\limits_{I \in \mathcal{P}([n])}$, where $\mathcal{P}([n])$ denotes the powerset of $[n]$.

Let $v_1, v_2, \ldots, v_n$ be n elements of $\mathbb{L}$.

(a) Prove that

$$\sum_{I \subseteq [n]} (-1)^{n-|I|} \left(w + \sum_{i \in I} v_i \right)^m = 0$$

for each $m \in \{0, 1, \ldots, n-1\}$ and $w \in \mathbb{L}$.

(b) Prove that

$$\sum_{I\subseteq[n]} (-1)^{n-|I|} \left(w+\sum_{i\in I} v_i\right)^n = \sum_{\sigma\in S_n} v_{\sigma(1)}v_{\sigma(2)}\cdots v_{\sigma(n)}$$

for every $w \in \mathbb{L}$.

(c) Prove that

$$\sum_{I\subseteq[n]} (-1)^{n-|I|} \left(\sum_{i\in I} v_i - \sum_{i\in[n]\setminus I} v_i\right)^m = 0$$

for each $m \in \{0,1,\ldots,n-1\}$.

(d) Prove that

$$\sum_{I\subseteq[n]} (-1)^{n-|I|} \left(\sum_{i\in I} v_i - \sum_{i\in[n]\setminus I} v_i\right)^n = 2^n \sum_{\sigma\in S_n} v_{\sigma(1)}v_{\sigma(2)}\cdots v_{\sigma(n)}.$$

Note that parts **(b)** and **(c)** of Exercise 6.50 are special cases of parts **(a)** and **(b)** of Exercise 6.51 (obtained by setting $w=0$).

The following exercise generalizes Exercise 6.50:

Exercise 6.52. For every $n \in \mathbb{N}$, let $[n]$ denote the set $\{1,2,\ldots,n\}$.

Let $\mathbb{L}$ be a noncommutative ring. (Keep in mind that our definition of a "noncommutative ring" includes all commutative rings.)

Let G be a finite set. Let H be a subset of G. Let $n \in \mathbb{N}$.

For each $i \in G$ and $j \in [n]$, let $b_{i,j}$ be an element of $\mathbb{L}$. For each $j \in [n]$ and each subset I of G, we define an element $b_{I,j} \in \mathbb{L}$ by $b_{I,j} = \sum_{i\in I} b_{i,j}$.

(a) Prove that

$$\sum_{\substack{I\subseteq G;\\ H\subseteq I}} (-1)^{|I|} b_{I,1}b_{I,2}\cdots b_{I,n} = (-1)^{|G|} \sum_{\substack{f:[n]\to G;\\ G\setminus H\subseteq f([n])}} b_{f(1),1}b_{f(2),2}\cdots b_{f(n),n}.$$

(b) If $n < |G\setminus H|$, then prove that

$$\sum_{\substack{I\subseteq G;\\ H\subseteq I}} (-1)^{|I|} b_{I,1}b_{I,2}\cdots b_{I,n} = 0.$$

(c) If $n = |G|$, then prove that

$$\sum_{I\subseteq G} (-1)^{|G\setminus I|} b_{I,1}b_{I,2}\cdots b_{I,n} = \sum_{\substack{f:[n]\to G\\ \text{is bijective}}} b_{f(1),1}b_{f(2),2}\cdots b_{f(n),n}.$$

Remark 6.166. Let $\mathbb{K}$ be a commutative ring. Let $n \in \mathbb{N}$. Let $A = \left(a_{i,j}\right)_{1\leq i\leq n,\ 1\leq j\leq n} \in \mathbb{K}^{n\times n}$ be an $n\times n$-matrix. Then, the *permanent* $\operatorname{per} A$ of A is defined to be the element

$$\sum_{\sigma\in S_n} a_{1,\sigma(1)}a_{2,\sigma(2)}\cdots a_{n,\sigma(n)}$$

of $\mathbb{K}$. The concept of a permanent is thus similar to the concept of a determinant (the only difference is that the factor $(-1)^\sigma$ is missing from the definition of the permanent); however, it has far fewer interesting properties. One of the properties that it does have is the so-called *Ryser formula* (see, e.g., [Comtet74, §4.9, [9e]]), which says that

$$\operatorname{per} A = (-1)^n \sum_{I\subseteq\{1,2,\ldots,n\}} (-1)^{|I|} \prod_{j=1}^{n} \sum_{i\in I} a_{i,j}.$$

The reader is invited to check that this formula follows from Exercise 6.52 **(c)** (applied to $\mathbb{L} = \mathbb{K}$, $G = \{1,2,\ldots,n\}$, $H = \varnothing$ and $b_{i,j} = a_{i,j}$).

Exercise 6.53. Let $n \in \mathbb{N}$. Let G be a finite set such that $n < |G|$. For each $i \in G$, let $A_i \in \mathbb{K}^{n\times n}$ be an $n\times n$-matrix. Prove that

$$\sum_{I\subseteq G} (-1)^{|I|} \det\left(\sum_{i\in I} A_i\right) = 0.$$

Exercise 6.54. Let $n \in \mathbb{N}$. Let $A = \left(a_{i,j}\right)_{1\leq i\leq n,\ 1\leq j\leq n}$ be an $n\times n$-matrix.

(a) Prove that

$$\sum_{\sigma\in S_n} (-1)^\sigma \left(\sum_{i=1}^{n} a_{i,\sigma(i)}\right)^k = 0$$

for each $k \in \{0,1,\ldots,n-2\}$.

(b) Assume that $n \geq 1$. Prove that

$$\sum_{\sigma\in S_n} (-1)^\sigma \left(\sum_{i=1}^{n} a_{i,\sigma(i)}\right)^{n-1} = (n-1)! \cdot \sum_{p=1}^{n}\sum_{q=1}^{n} (-1)^{p+q} \det\left(A_{\sim p,\sim q}\right).$$

(Here, we are using the notation introduced in Definition 6.81.)

Exercise 6.55. Let $n \in \mathbb{N}$. Let $a_1, a_2, \ldots, a_n$ be n elements of $\mathbb{K}$. Let $b_1, b_2, \ldots, b_n$ be n elements of $\mathbb{K}$.

Let $m = \dbinom{n}{2}$.

(a) Prove that

$$\sum_{\sigma \in S_n} (-1)^\sigma \left(\sum_{i=1}^n a_i b_{\sigma(i)} \right)^k = 0$$

for each $k \in \{0, 1, \ldots, m-1\}$.

(b) Prove that

$$\sum_{\sigma \in S_n} (-1)^\sigma \left(\sum_{i=1}^n a_i b_{\sigma(i)} \right)^m = \mathbf{m}(0, 1, \ldots, n-1) \cdot \prod_{1 \leq i < j \leq n} \left(\left(a_i - a_j\right) \left(b_i - b_j\right) \right).$$

Here, we are using the notation introduced in Exercise 6.2.

(c) Let $k \in \mathbb{N}$. Prove that

$$\begin{aligned} & \sum_{\sigma \in S_n} (-1)^\sigma \left(\sum_{i=1}^n a_i b_{\sigma(i)} \right)^k \\ & = \sum_{\substack{(g_1, g_2, \ldots, g_n) \in \mathbb{N}^n; \\ g_1 < g_2 < \cdots < g_n; \\ g_1 + g_2 + \cdots + g_n = k}} \mathbf{m}(g_1, g_2, \ldots, g_n) \\ & \qquad \cdot \det \left(\left(a_i^{g_j} \right)_{1 \leq i \leq n,\ 1 \leq j \leq n} \right) \cdot \det \left(\left(b_i^{g_j} \right)_{1 \leq i \leq n,\ 1 \leq j \leq n} \right). \end{aligned}$$

Here, we are using the notation introduced in Exercise 6.2.

Note that Exercise 6.55 generalizes [AndDos10, Exercise 12.13] and [AndDos12, §12.1, Problem 1].

6.25. Additional exercises

Here are a few more additional exercises, with no importance to the rest of the text (and mostly no solutions given).

Exercise 6.56. Let $n \in \mathbb{N}$. For any subset I of $\{1,2,\ldots,n\}$, we let $\widetilde{I}$ denote the complement $\{1,2,\ldots,n\} \setminus I$ of I. (For instance, if $n = 4$ and $I = \{1,4\}$, then $\widetilde{I} = \{2,3\}$.) Let us use the notations introduced in Definition 6.78 and in Definition 6.153.

Let $A \in \mathbb{K}^{n\times n}$ be an invertible matrix. Let P and Q be two subsets of $\{1,2,\ldots,n\}$ satisfying $|P| = |Q|$. Prove that

$$\det\left(\operatorname{sub}_{w(P)}^{w(Q)} A\right) = (-1)^{\sum P + \sum Q} \det A \cdot \det\left(\operatorname{sub}_{w\left(\widetilde{Q}\right)}^{w\left(\widetilde{P}\right)} \left(A^{-1}\right)\right). \qquad (477)$$

[**Hint:** Apply Exercise 6.38 to a matrix obtained from A by permuting the rows and permuting the columns.]

Remark 6.167. The claim of Exercise 6.56 is the so-called *Jacobi complementary minor theorem*. It appears, for example, in [Lalond96, (1)] and in [CaSoSp12, Lemma A.1 (e)], and is used rather often when working with determinants (for example, it is used in [LLPT95, Chapter SYM, proof of Proposition (7.5) (5)] and many times in [CaSoSp12]).

The determinant of a submatrix of a matrix A is called a *minor* of A. Thus, in the equality (477), the determinant $\det\left(\operatorname{sub}_{w(P)}^{w(Q)} A\right)$ on the left hand side is a minor of A, whereas the determinant $\det\left(\operatorname{sub}_{w\left(\widetilde{Q}\right)}^{w\left(\widetilde{P}\right)} \left(A^{-1}\right)\right)$ on the right hand side is a minor of A^{-1}. Thus, roughly speaking, the equality (477) says that any minor of A equals a certain minor of A^{-1} times $\det A$ times a certain sign.

It is instructive to check the particular case of (477) obtained when both P and Q are sets of cardinality $n-1$ (so that $\widetilde{P}$ and $\widetilde{Q}$ are 1-element sets). This particular case turns out to be the statement of Theorem 6.110 **(b)** in disguise.

Exercise 6.38 is the particular case of Exercise 6.56 obtained when $P = \{1,2,\ldots,k\}$ and $Q = \{1,2,\ldots,k\}$.

Exercise 6.57. Let n and k be positive integers such that $k \leq n$. Let $A \in \mathbb{K}^{n\times(n-k)}$ and $B \in \mathbb{K}^{n\times(n+k)}$.

Let us use the notations from Definition 6.128. For any subset I of $\{1,2,\ldots,n+k\}$, we introduce the following five notations:

- Let $\sum I$ denote the sum of all elements of I. (Thus, $\sum I = \sum\limits_{i\in I} i$.)
- Let $w(I)$ denote the list of all elements of I in increasing order (with no repetitions). (See Definition 2.50 for the formal definition of this list.) (For example, $w(\{3,4,8\}) = (3,4,8)$.)
- Let $(A \mid B_{\bullet,I})$ denote the $n \times (n-k+|I|)$-matrix whose columns are $\underbrace{A_{\bullet,1}, A_{\bullet,2}, \ldots, A_{\bullet,n-k}}_{\text{the columns of } A}, B_{\bullet,i_1}, B_{\bullet,i_2}, \ldots, B_{\bullet,i_\ell}$ (from left to right), where

$(i_1, i_2, \ldots, i_\ell) = w(I)$.

- Let $B_{\bullet,\sim I}$ denote the $n \times (n + k - |I|)$-matrix whose columns are $B_{\bullet,j_1}, B_{\bullet,j_2}, \ldots, B_{\bullet,j_h}$ (from left to right), where $(j_1, j_2, \ldots, j_h) = w(\{1, 2, \ldots, n+k\} \setminus I)$.
  (Using the notations of Definition 6.31, we can rewrite this definition as $B_{\bullet,\sim I} = \operatorname{cols}_{j_1,j_2,\ldots,j_h} B$, where $(j_1, j_2, \ldots, j_h) = w(\{1, 2, \ldots, n+k\} \setminus I)$.)

Then, prove that

$$\sum_{\substack{I \subseteq \{1,2,\ldots,n+k\};\\ |I|=k}} (-1)^{\sum I + (1+2+\cdots+k)} \det(A \mid B_{\bullet,I}) \det(B_{\bullet,\sim I}) = 0.$$

(Note that this generalizes Theorem 6.150; indeed, the latter theorem is the particular case for $k = 1$.)

Exercise 6.58. Recall that the binomial coefficients satisfy the recurrence relation (234), which (visually) says that every entry of Pascal's triangle is the sum of the two entries left-above it and right-above it.

Let us now define a variation of Pascal's triangle as follows: Define a nonnegative integer $\dbinom{m}{n}_D$ for every $m \in \mathbb{N}$ and $n \in \mathbb{N}$ recursively as follows:

- Set $\dbinom{0}{n}_D = \begin{cases} 1, & \text{if } n = 0; \\ 0, & \text{if } n > 0 \end{cases}$ for every $n \in \mathbb{N}$.
- For every $m \in \mathbb{Z}$ and $n \in \mathbb{Z}$, set $\dbinom{m}{n}_D = 0$ if either m or n is negative.
- For every positive integer m and every $n \in \mathbb{N}$, set
$$\binom{m}{n}_D = \binom{m-1}{n-1}_D + \binom{m-1}{n}_D + \binom{m-2}{n-1}_D.$$

(Thus, if we lay these $\dbinom{m}{n}_D$ out in the same way as the binomial coefficients $\dbinom{m}{n}$ in Pascal's triangle, then every entry is the sum of the three entries left-above it, right-above it, and straight above it.)

The integers $\dbinom{m}{n}_D$ are known as the Delannoy numbers.

(a) Show that

$$\binom{n+m}{n}_D = \sum_{i=0}^{n} \binom{n}{i} \binom{m+i}{n} = \sum_{i=0}^{n} \binom{n}{i} \binom{m}{i} 2^i$$

for every $n \in \mathbb{N}$ and $m \in \mathbb{N}$. (The second equality sign here is a consequence of Proposition 3.39 **(e)**.)

(b) Let $n \in \mathbb{N}$. Let A be the $n \times n$-matrix $\left(\binom{i+j-2}{i-1}_D \right)_{1 \leq i \leq n,\ 1 \leq j \leq n}$ (an analogue of the matrix A from Exercise 6.11). Show that

$$\det A = 2^{n(n-1)/2}.$$

The next exercise is known under the (arguably not very distinctive) name of "matrix determinant lemma":

Exercise 6.59. Let $n \in \mathbb{N}$. Let u be a column vector with n entries, and let v be a row vector with n entries. (Thus, uv is an $n \times n$-matrix, whereas vu is a 1×1-matrix.) Let A be an $n \times n$-matrix. Prove that

$$\det (A + uv) = \det A + v (\operatorname{adj} A) u$$

(where we regard the 1×1-matrix $v (\operatorname{adj} A) u$ as an element of $\mathbb{K}$).

The next exercise relies on Definition 6.89.

Exercise 6.60. Let $n \in \mathbb{N}$. Let $u \in \mathbb{K}^{n \times 1}$ be a column vector with n entries, and let $v \in \mathbb{K}^{1 \times n}$ be a row vector with n entries. (Thus, uv is an $n \times n$-matrix, whereas vu is a 1×1-matrix.) Let $h \in \mathbb{K}$. Let H be the 1×1-matrix $\begin{pmatrix} h \end{pmatrix} \in \mathbb{K}^{1 \times 1}$.

(a) Prove that every $n \times n$-matrix $A \in \mathbb{K}^{n \times n}$ satisfies

$$\det \begin{pmatrix} A & u \\ v & H \end{pmatrix} = h \det A - v (\operatorname{adj} A) u$$

(where we regard the 1×1-matrix $v (\operatorname{adj} A) u$ as an element of $\mathbb{K}$).

(b) Write the vector u in the form $u = (u_1, u_2, \ldots, u_n)^T$. Write the vector v in the form $v = (v_1, v_2, \ldots, v_n)$.

For every two objects i and j, define $\delta_{i,j} \in \mathbb{K}$ by $\delta_{i,j} = \begin{cases} 1, & \text{if } i = j; \\ 0, & \text{if } i \neq j \end{cases}$.

Let $d_1, d_2, \ldots, d_n$ be n elements of $\mathbb{K}$. Let D be the $n \times n$-matrix $\left(d_i \delta_{i,j} \right)_{1 \leq i \leq n,\ 1 \leq j \leq n}$. Prove that

$$\det \begin{pmatrix} D & u \\ v & H \end{pmatrix} = h \cdot (d_1 d_2 \cdots d_n) - \sum_{i=1}^{n} u_i v_i \prod_{\substack{j \in \{1,2,\ldots,n\}; \\ j \neq i}} d_j.$$

Example 6.168. Let $n = 3$. Then, Exercise 6.60 **(b)** states that

$$\det \begin{pmatrix} d_1 & 0 & 0 & u_1 \\ 0 & d_2 & 0 & u_2 \\ 0 & 0 & d_3 & u_3 \\ v_1 & v_2 & v_3 & h \end{pmatrix} = h \cdot (d_1 d_2 d_3) - \sum_{i=1}^{3} u_i v_i \prod_{\substack{j \in \{1,2,3\}; \\ j \neq i}} d_j$$
$$= h d_1 d_2 d_3 - (u_1 v_1 d_2 d_3 + u_2 v_2 d_1 d_3 + u_3 v_3 d_1 d_2)$$

for any ten elements $u_1, u_2, u_3, v_1, v_2, v_3, d_1, d_2, d_3, h$ of $\mathbb{K}$.

Exercise 6.61. Let $P = \sum_{k=0}^{d} p_k X^k$ and $Q = \sum_{k=0}^{e} q_k X^k$ be two polynomials over $\mathbb{K}$ (where $p_0, p_1, \ldots, p_d \in \mathbb{K}$ and $q_0, q_1, \ldots, q_e \in \mathbb{K}$ are their coefficients) such that $d + e > 0$. Define a $(d+e) \times (d+e)$-matrix A as follows:

- For every $k \in \{1, 2, \ldots, e\}$, the k-th row of A is

$$\left(\underbrace{0, 0, \ldots, 0}_{k-1 \text{ zeroes}}, p_d, p_{d-1}, \ldots, p_1, p_0, \underbrace{0, 0, \ldots, 0}_{e-k \text{ zeroes}} \right).$$

- For every $k \in \{1, 2, \ldots, d\}$, the $(e+k)$-th row of A is

$$\left(\underbrace{0, 0, \ldots, 0}_{k-1 \text{ zeroes}}, q_e, q_{e-1}, \ldots, q_1, q_0, \underbrace{0, 0, \ldots, 0}_{d-k \text{ zeroes}} \right).$$

(For example, if $d = 4$ and $e = 3$, then

$$A = \begin{pmatrix} p_4 & p_3 & p_2 & p_1 & p_0 & 0 & 0 \\ 0 & p_4 & p_3 & p_2 & p_1 & p_0 & 0 \\ 0 & 0 & p_4 & p_3 & p_2 & p_1 & p_0 \\ q_3 & q_2 & q_1 & q_0 & 0 & 0 & 0 \\ 0 & q_3 & q_2 & q_1 & q_0 & 0 & 0 \\ 0 & 0 & q_3 & q_2 & q_1 & q_0 & 0 \\ 0 & 0 & 0 & q_3 & q_2 & q_1 & q_0 \end{pmatrix}.$$

)

Assume that the polynomials P and Q have a common root z (that is, there exists a $z \in \mathbb{K}$ such that $P(z) = 0$ and $Q(z) = 0$). Show that $\det A = 0$.

[**Hint:** Find a column vector v with $d + e$ entries satisfying $Av = 0_{(d+e) \times 1}$; then apply Corollary 6.102.]

Remark 6.169. The matrix A in Exercise 6.61 is called the *Sylvester matrix* of the polynomials P and Q (for degrees d and e); its determinant $\det A$ is known as their *resultant* (at least when d and e are actually the degrees of P and Q). According to the exercise, the condition $\det A = 0$ is necessary for P and Q to have a common root. In the general case, the converse does not hold: For one, you can always force $\det A$ to be 0 by taking $d > \deg P$ and $e > \deg Q$ (so $p_d = 0$ and $q_e = 0$, and thus the 1-st column of A consists of zeroes). More importantly, the resultant of the two polynomials $X^3 - 1$ and $X^2 + X + 1$ is 0, but they only have common roots in $\mathbb{C}$, not in $\mathbb{R}$. Thus, there is more to common roots than just the vanishing of a determinant.

However, if $\mathbb{K}$ is an algebraically closed field (I won't go into the details of what this means, but an example of such a field is $\mathbb{C}$), and if $d = \deg P$ and $e = \deg Q$, then the polynomials P and Q have a common root **if and only if** their resultant is 0.

The next two exercises can be seen as variations on the Vandermonde determinant. The first one is [Kratte99, Proposition 1] (and also appears in [Grinbe10, Theorem 2]):

Exercise 6.62. Let $n \in \mathbb{N}$. Let $a_1, a_2, \ldots, a_n \in \mathbb{K}$. Furthermore, for each $j \in \{1, 2, \ldots, n\}$, let $P_j \in \mathbb{K}[X]$ be a polynomial such that

$$\deg\left(P_j\right) \leq j - 1.$$

(In particular, $\deg(P_1) \leq 1 - 1 = 0$, so that the polynomial P_1 is constant.)

For each $j \in \{1, 2, \ldots, n\}$, let c_j be the coefficient of X^{j-1} in the polynomial P_j. Prove that

$$\det\left(\left(P_j\left(a_i\right)\right)_{1\leq i\leq n,\ 1\leq j\leq n}\right) = \left(\prod_{j=1}^{n} c_j\right) \cdot \prod_{1\leq j<i\leq n} \left(a_i - a_j\right).$$

The next exercise is [Kratte99, Lemma 6] (with the notations changed):

Exercise 6.63. Let $n \in \mathbb{N}$. Let $a_1, a_2, \ldots, a_n \in \mathbb{K}$ and $b_2, b_3, \ldots, b_n \in \mathbb{K}$. For each $j \in \{1, 2, \ldots, n\}$, we define a polynomial $Q_j \in \mathbb{K}[X]$ by setting

$$Q_j = \left(X - b_{j+1}\right)\left(X - b_{j+2}\right)\cdots\left(X - b_n\right).$$

(In particular, $Q_n = (\text{empty product}) = 1$.) Furthermore, for each $j \in \{1, 2, \ldots, n\}$, let $P_j \in \mathbb{K}[X]$ be a polynomial such that

$$\deg\left(P_j\right) \leq j - 1.$$

(In particular, $\deg(P_1) \leq 1 - 1 = 0$, so that the polynomial P_1 is constant.)

Prove that

$$\det\left(\left(P_j(a_i)\,Q_j(a_i)\right)_{1\leq i\leq n,\ 1\leq j\leq n}\right)=\left(\prod_{j=1}^{n}P_j(b_j)\right)\cdot\prod_{1\leq i<j\leq n}\left(a_i-a_j\right).$$

[**Hint:** For each $j\in\{1,2,\ldots,n\}$, construct j elements $c_{j,1},c_{j,2},\ldots,c_{j,j}$ of $\mathbb{K}$ satisfying $P_jQ_j=\sum\limits_{k=1}^{j}c_{j,k}Q_k$ and $c_{j,j}=P_j(b_j)$.]

Exercise 6.63 has many applications:[284]

Exercise 6.64. Let $n\in\mathbb{N}$. Let $a_1,a_2,\ldots,a_n\in\mathbb{K}$ and $b_2,b_3,\ldots,b_n\in\mathbb{K}$ and $c_1,c_2,\ldots,c_{n-1}\in\mathbb{K}$.

(a) Prove that

$$\det\left(\left(\left(\prod_{u=1}^{j-1}(a_i-c_u)\right)\left(\prod_{u=j+1}^{n}(a_i-b_u)\right)\right)_{1\leq i\leq n,\ 1\leq j\leq n}\right)$$
$$=\left(\prod_{1\leq i<j\leq n}(b_j-c_i)\right)\cdot\prod_{1\leq i<j\leq n}\left(a_i-a_j\right).$$

(b) Prove that

$$\det\left(\left(\left(\prod_{u=1}^{j-1}(a_i+c_u)\right)\left(\prod_{u=j+1}^{n}(a_i+b_u)\right)\right)_{1\leq i\leq n,\ 1\leq j\leq n}\right)$$
$$=\left(\prod_{1\leq i<j\leq n}(c_i-b_j)\right)\cdot\prod_{1\leq i<j\leq n}\left(a_i-a_j\right).$$

(c) Use this to solve Exercise 6.18 again.

The next exercise (more precisely, its part **(a)**) is a recent result of Fraser and Yeats [FraYea21, Theorem 2.1.4] (slightly extended):

Exercise 6.65. (a) Let $n\in\mathbb{N}$. Let $A\in\mathbb{K}^{n\times n}$ be an $n\times n$-matrix. Let $k\in\mathbb{N}$. Let $u_1,u_2,\ldots,u_k$ be k elements of $\{1,2,\ldots,n\}$. Let $v_1,v_2,\ldots,v_k$ be k further elements of $\{1,2,\ldots,n\}$. Let $U=\{u_1,u_2,\ldots,u_k\}$ and $V=\{v_1,v_2,\ldots,v_k\}$. Let

[284] Exercise 6.64 **(b)** appears in [Kratte99, Lemma 3], in [Gorin21, Lemma 2.7], in [Knuth1, §1.2.3, exercise 47] and in [Bresso99, Theorem 2.9]. (In all of these sources, it appears with modified notations. For instance, [Kratte99, Lemma 3] is Exercise 6.64 **(b)** for $a_i=X_i$, $b_i=A_i$ and $c_i=B_{i+1}$.)

$t \in \{1, 2, \ldots, n\} \setminus V$. Prove that

$$\begin{aligned} & \det A \cdot \det \left(\operatorname{sub}_{u_1, u_2, \ldots, u_k}^{v_1, v_2, \ldots, v_k} A \right) \\ & = \sum_{s \in \{1, 2, \ldots, n\} \setminus U} (-1)^{s+t} \det \left(A_{\sim s, \sim t} \right) \det \left(\operatorname{sub}_{s, u_1, u_2, \ldots, u_k}^{t, v_1, v_2, \ldots, v_k} A \right). \end{aligned}$$

(Here, we are using the notations of Definition 6.78 and of Definition 6.81.)

(b) Use this to give a new proof of Proposition 6.123.

For the next two exercises, we need a notation for exchanging certain columns between two matrices:

Definition 6.170. Let $n \in \mathbb{N}$ and $m \in \mathbb{N}$ and $p \in \mathbb{N}$. Let $A \in \mathbb{K}^{n \times m}$ and $B \in \mathbb{K}^{n \times p}$ be two matrices. Let J be a subset of $\{1, 2, \ldots, m\}$, and let K be a subset of $\{1, 2, \ldots, p\}$ such that $|J| = |K|$. Let $j_1, j_2, \ldots, j_q$ be all elements of J, listed in increasing order (with no repetitions). Thus, $q = |J| = |K|$. Hence, the set K has exactly q elements. Let $k_1, k_2, \ldots, k_q$ be all elements of K, listed in increasing order (with no repetitions). (This is well-defined, since K has exactly q elements.) Then, we let $\left(\begin{array}{c} A \leftarrow B \\ J \leftarrow K \end{array} \right)$ denote the $n \times m$-matrix obtained from A by replacing the j_1-st, j_2-nd, ..., j_q-th columns of A by the k_1-st, k_2-nd, ..., k_q-th columns of B, respectively. In other words, if we write the matrices A and B as $A = \left(a_{x,y} \right)_{1 \leq x \leq n,\ 1 \leq y \leq m}$ and $B = \left(b_{x,y} \right)_{1 \leq x \leq n,\ 1 \leq y \leq p}$, then $\left(\begin{array}{c} A \leftarrow B \\ J \leftarrow K \end{array} \right)$ is defined to be the $n \times m$-matrix $\left(c_{x,y} \right)_{1 \leq x \leq n,\ 1 \leq y \leq m}$, where we let

$$c_{x,y} = \begin{cases} a_{x,y}, & \text{if } y \notin J; \\ b_{x,k_i}, & \text{if } y = j_i \text{ for some } i \in \{1, 2, \ldots, q\} \end{cases} \qquad (478)$$
$$\text{for all } x \in \{1, 2, \ldots, n\} \text{ and } y \in \{1, 2, \ldots, m\}.$$

Example 6.171. Let $A = \begin{pmatrix} a & b & c & d \\ a' & b' & c' & d' \\ a'' & b'' & c'' & d'' \end{pmatrix}$ and $B = \begin{pmatrix} e & f & g \\ e' & f' & g' \\ e'' & f'' & g'' \end{pmatrix}$ and $J = \{1, 3\}$ and $K = \{2, 3\}$. Then,

$$\left(\begin{array}{c} A \leftarrow B \\ J \leftarrow K \end{array} \right) = \begin{pmatrix} f & b & g & d \\ f' & b' & g' & d' \\ f'' & b'' & g'' & d'' \end{pmatrix} \qquad \text{and}$$

$$\left(\begin{array}{c} B \leftarrow A \\ K \leftarrow J \end{array} \right) = \begin{pmatrix} e & a & c \\ e' & a' & c' \\ e'' & a'' & c'' \end{pmatrix}.$$

The next exercise is known as *Sylvester's Lemma* (appearing, e.g., in [Fulton97, Chapter 8, Lemma 2] or [Muir30, §137]):

Exercise 6.66. Let $n \in \mathbb{N}$. Let $A \in \mathbb{K}^{n \times n}$ and $B \in \mathbb{K}^{n \times n}$ be two $n \times n$-matrices, and let J be a subset of $\{1, 2, \ldots, n\}$. Prove that

$$\det A \cdot \det B = \sum_{\substack{K \subseteq \{1,2,\ldots,n\}; \\ |K| = |J|}} \det \begin{pmatrix} A \leftarrow B \\ J \leftarrow K \end{pmatrix} \det \begin{pmatrix} B \leftarrow A \\ K \leftarrow J \end{pmatrix}.$$

The next exercise is essentially [Muir30, §319]:

Exercise 6.67. Let $n \in \mathbb{N}$. Let $A \in \mathbb{K}^{n \times n}$ and $B \in \mathbb{K}^{n \times n}$ be two $n \times n$-matrices, and let $r \in \mathbb{N}$. Prove that

$$\sum_{\substack{J \subseteq \{1,2,\ldots,n\}; \\ |J| = r}} \det \begin{pmatrix} A \leftarrow B \\ J \leftarrow J \end{pmatrix} = \sum_{\substack{J \subseteq \{1,2,\ldots,n\}; \\ |J| = r}} \det \begin{pmatrix} A^T \leftarrow B^T \\ J \leftarrow J \end{pmatrix}.$$

The next exercise generalizes [Muir30, §231][285] as well as [AlAmra97, §VI.4, ASSERTION][286]:

Exercise 6.68. Let $n \in \mathbb{N}$. Let $k \in \{0, 1, \ldots, n\}$. Let $A \in \mathbb{K}^{k \times n}$ and $B \in \mathbb{K}^{n \times (n-k)}$ be two matrices satisfying $AB = 0_{k \times (n-k)}$. (Recall that $0_{k \times (n-k)}$ means the $k \times (n-k)$ zero matrix.) Let P and Q be two subsets of $\{1, 2, \ldots, n\}$ such that $|P| = |Q| = k$.

We shall use the following notations:

- If I is a finite set of integers, then $w(I)$ shall denote the list of all elements of I in increasing order (with no repetitions). (See Definition 2.50 for the formal definition of this list.) (For example, $w(\{3, 4, 8\}) = (3, 4, 8)$.)
- For any subset I of $\{1, 2, \ldots, n\}$, we let $\widetilde{I}$ denote the complement $\{1, 2, \ldots, n\} \setminus I$ of I. (For instance, if $n = 4$ and $I = \{1, 4\}$, then $\widetilde{I} = \{2, 3\}$.)
- If $i_1, i_2, \ldots, i_u$ are some elements of $\{1, 2, \ldots, n\}$, then $\operatorname{cols}_{(i_1, i_2, \ldots, i_u)} A$ shall denote the matrix $\operatorname{cols}_{i_1, i_2, \ldots, i_u} A$.

[285] Muir, in [Muir30, §231], requires $\mathbb{K}$ to be the field $\mathbb{R}$ of real numbers, and (rather surprisingly) his proof actually uses this (extraneous) requirement.

[286] To wit, [AlAmra97, §VI.4, ASSERTION] can be obtained by applying Exercise 6.68 to $k = 1$, $A = (U_{i,j})_{1 \leq i \leq n-1,\ 1 \leq j \leq n}$, $B = (X_i)_{1 \leq i \leq n,\ 1 \leq j \leq 1}$, $P = \{1, 2, \ldots, n\} \setminus \{i\}$ and $Q = \{1, 2, \ldots, n\} \setminus \{j\}$ and $\mathbb{K} = k[U_{pq}, X_t] \diagup (f_1, f_2, \ldots, f_{n-1})$.

- If $i_1, i_2, \ldots, i_u$ are some elements of $\{1, 2, \ldots, n\}$, then $\operatorname{rows}_{(i_1, i_2, \ldots, i_u)} B$ shall denote the matrix $\operatorname{rows}_{i_1, i_2, \ldots, i_u} B$.

Prove that

$$\begin{aligned}
&(-1)^{\sum \widetilde{Q}} \det\left(\operatorname{cols}_{w(P)} A\right) \cdot \det\left(\operatorname{rows}_{w(\widetilde{Q})} B\right) \\
&= (-1)^{\sum \widetilde{P}} \det\left(\operatorname{cols}_{w(Q)} A\right) \cdot \det\left(\operatorname{rows}_{w(\widetilde{P})} B\right).
\end{aligned}$$

Example 6.172. Let us illustrate Exercise 6.68 on an example. Set $n = 5$ and $k = 2$, and let

$$A = \begin{pmatrix} a & b & c & d & e \\ a' & b' & c' & d' & e' \end{pmatrix} \qquad \text{and} \qquad B = \begin{pmatrix} x & x' & x'' \\ y & y' & y'' \\ z & z' & z'' \\ u & u' & u'' \\ v & v' & v'' \end{pmatrix}$$

be two matrices satisfying $AB = 0_{2\times 3}$. Let $P = \{1, 3\}$ and $Q = \{3, 5\}$. Then, Exercise 6.68 claims that

$$\begin{aligned}
&(-1)^{1+2+4} \det\left(\operatorname{cols}_{1,3} A\right) \cdot \det\left(\operatorname{rows}_{1,2,4} B\right) \\
&= (-1)^{2+4+5} \det\left(\operatorname{cols}_{3,5} A\right) \cdot \det\left(\operatorname{rows}_{2,4,5} B\right).
\end{aligned}$$

In other words, it claims that

$$\begin{aligned}
&(-1)^{1+2+4} \det\begin{pmatrix} a & c \\ a' & c' \end{pmatrix} \cdot \det\begin{pmatrix} x & x' & x'' \\ y & y' & y'' \\ u & u' & u'' \end{pmatrix} \\
&= (-1)^{2+4+5} \det\begin{pmatrix} c & e \\ c' & e' \end{pmatrix} \cdot \det\begin{pmatrix} y & y' & y'' \\ u & u' & u'' \\ v & v' & v'' \end{pmatrix}.
\end{aligned}$$

It is not at all straightforward to derive this equality from the condition $AB = 0_{2\times 3}$.

The next exercise (taken from [Knuth1, §1.2.3, Exercise 44]) states some deeper properties of the *Cauchy matrix* $\left(\dfrac{1}{x_i + y_j}\right)_{1\leq i\leq n,\ 1\leq j\leq n}$ whose determinant we computed in Exercise 6.18:

Exercise 6.69. Let $n \in \mathbb{N}$. Let $x_1, x_2, \ldots, x_n$ be n elements of $\mathbb{K}$. Let $y_1, y_2, \ldots, y_n$ be n elements of $\mathbb{K}$. Assume that $x_i + y_j$ is invertible in $\mathbb{K}$ for every $(i, j) \in \{1, 2, \ldots, n\}^2$. Let C be the matrix $\left(\dfrac{1}{x_i + y_j} \right)_{1 \leq i \leq n,\ 1 \leq j \leq n}$. Prove the following:

(a) The sum of all entries of the adjugate matrix $\operatorname{adj} C$ is

$$\left(\sum_{k=1}^{n} x_k + \sum_{k=1}^{n} y_k \right) \cdot \frac{\prod\limits_{1 \leq i < j \leq n} \left(\left(x_i - x_j \right) \left(y_i - y_j \right) \right)}{\prod\limits_{(i,j) \in \{1,2,\ldots,n\}^2} \left(x_i + y_j \right)}.$$

(b) If the matrix C is invertible, then the sum of all entries of its inverse C^{-1} is

$$\sum_{k=1}^{n} x_k + \sum_{k=1}^{n} y_k.$$

(c) Let $u \in \mathbb{K}^n$ be the column vector whose all n entries equal 1. Let D be the $(n+1) \times (n+1)$-matrix $\begin{pmatrix} C & u \\ u^T & 0_{1 \times 1} \end{pmatrix}$ (where we are using the notation from Definition 6.89). Then,

$$\det D = - \left(\sum_{k=1}^{n} x_k + \sum_{k=1}^{n} y_k \right) \cdot \frac{\prod\limits_{1 \leq i < j \leq n} \left(\left(x_i - x_j \right) \left(y_i - y_j \right) \right)}{\prod\limits_{(i,j) \in \{1,2,\ldots,n\}^2} \left(x_i + y_j \right)}.$$

[**Hint:** To prove part **(a)**, first show that

$$\sum_{i=1}^{n} \sum_{j=1}^{n} (-1)^{i+j} a_{i,j} \left(x_i + y_j \right) \det \left(A_{\sim i, \sim j} \right) = \left(\sum_{k=1}^{n} x_k + \sum_{k=1}^{n} y_k \right) \cdot \det A$$

for any $n \times n$-matrix $A = \left(a_{i,j} \right)_{1 \leq i \leq n,\ 1 \leq j \leq n}$.]

Example 6.173. If $n = 3$, then the matrix D in Exercise 6.69 **(c)** is

$$\begin{pmatrix} \dfrac{1}{x_1 + y_1} & \dfrac{1}{x_1 + y_2} & \dfrac{1}{x_1 + y_3} & 1 \\ \dfrac{1}{x_2 + y_1} & \dfrac{1}{x_2 + y_2} & \dfrac{1}{x_2 + y_3} & 1 \\ \dfrac{1}{x_3 + y_1} & \dfrac{1}{x_3 + y_2} & \dfrac{1}{x_3 + y_3} & 1 \\ 1 & 1 & 1 & 0 \end{pmatrix}.$$

References

[Abeles14] Francine F. Abeles, *Chiò's and Dodgson's determinantal identities*, Linear Algebra and its Applications, Volume 454, 1 August 2014, pp. 130–137.

[Aigner07] Martin Aigner, *A Course in Enumeration*, Graduate Texts in Mathematics #238, Springer 2007.

[AigZie14] Martin Aigner, Günter M. Ziegler, *Proofs from the Book*, 6th edition, Springer 2018.

[Aitken56] A. C. Aitken, *Determinants and Matrices*, 9th edition, Oliver and Boyd 1956.

[AlAmra97] Abdallah Al-Amrani, *An Introduction to TRAGHEITSFORMEN*, 1997. `http://hal.archives-ouvertes.fr/hal-00912907`

[AleGhe14] Emily Allen, Irina Gheorghiciuc, *A weighted interpretation for the super Catalan numbers*, Journal of Integer Sequences, Vol. 17 (2014), Article 14.10.7.
A preprint is arXiv:1403.5246v2.

[AmaEsc05] Herbert Amann, Joachim Escher, *Analysis I*, translated from the German by Gary Brookfield, Birkhäuser 2005.

[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.

[AndFen04] Titu Andreescu, Zuming Feng, *A Path to Combinatorics for Undergraduates: Counting Strategies*, Springer 2004.

[Artin10] Michael Artin, *Algebra*, 2nd edition, Pearson 2010.

[Axler15] Sheldon Axler, *Linear Algebra Done Right*, 3rd edition, Springer 2015.

[BenDre07] Arthur T. Benjamin and Gregory P. Dresden, *A Combinatorial Proof of Vandermonde's Determinant*, The American Mathematical Monthly, Vol. 114, No. 4 (Apr., 2007), pp. 338–341.
(Also available at `http://scholarship.claremont.edu/hmc_fac_pub/524/` .)

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

[BenQui04] Arthur T. Benjamin and Jennifer J. Quinn, *Proofs that Really Count: The Magic of Fibonacci Numbers and More*, Mathematical Adventures for Students and Amateurs, (David F. Hayes and Tatiana Shubin, editors), Spectrum Series of MAA, pp. 83–98, 2004.

[BenQui08] Arthur T. Benjamin and Jennifer J. Quinn, *An Alternate Approach to Alternating Sums: A Method to DIE for*, The College Mathematics Journal, Volume 39, Number 3, May 2008, pp. 191-202(12).

[BerBru08] Adam Berliner and Richard A. Brualdi, *A combinatorial proof of the Dodgson/Muir determinantal identity*, International Journal of Information and Systems Sciences, Volume 4 (2008), Number 1, pp. 1–7.

[Bergma15] George M. Bergman, *A Companion to Lang's Algebra*, website (2015).
`https://math.berkeley.edu/~gbergman/.C.to.L/`

[Bharga00] Manjul Bhargava, *The Factorial Function and Generalizations*, The American Mathematical Monthly, Vol. **107**, No. 9 (Nov., 2000), pp. 783–799.

[BirMac99] Saunders Mac Lane, Garrett Birkhoff, *Algebra*, 3rd edition, AMS Chelsea Publishing 1999.

[BjoBre05] Anders Björner, Francesco Brenti, *Combinatorics of Coxeter Groups*, Graduate Texts in Mathematics #231, Springer 2005.
See `https://www.mat.uniroma2.it/~brenti/correct.ps` for errata.

[Bona22] Miklós Bóna, *Combinatorics of Permutations*, 3rd edition, CRC Press 2022.

[Bourba74] Nicolas Bourbaki, *Algebra I, Chapters 1-3*, Hermann 1974.
`https://archive.org/details/ElementsOfMathematics-AlgebraPart1/`

[Bresso99] David M. Bressoud, *Proofs and Confirmations: The Story of the Alternating Sign Matrix Conjecture*, Cambridge University Press 1999.
See `https://www.macalester.edu/~bressoud/books/PnC/PnCcorrect.html` for errata.

[BruRys91] Richard A. Brualdi, Herbert J. Ryser, *Combinatorial Matrix Theory*, Cambridge University Press 1991.

[Bump02] Daniel Bump, *Mathematics of the Rubik's Cube*, lecture notes (in 2 versions).
`http://sporadic.stanford.edu/bump/match/rubik.html`

[CamFon07] Peter J. Cameron, Dima G. Fon-Der-Flaass, *Fibonacci notes*, 1 August 2007.
`http://www.maths.qmul.ac.uk/~pjc/comb/fibo.pdf`

[CaSoSp12] Sergio Caracciolo, Alan D. Sokal, Andrea Sportiello, *Algebraic/combinatorial proofs of Cayley-type identities for derivatives of determinants and pfaffians*, Advances in Applied Mathematics 50, pp. 474–594 (2013). Also appears as arXiv preprint `arXiv:1105.6270v2`.

[Clemen22] Philippe Clément, *From natural numbers to prime fields and finite fields,* arXiv:2209.01069v1.

[Comtet74] Louis Comtet, *Advanced Combinatorics: The Art of Finite and Infinite Expansions*, D. Reidel Publishing Company, 1974.

[Conrad1] Keith Conrad, *Sign of permutations*,
`https://kconrad.math.uconn.edu/blurbs/grouptheory/sign.pdf` .

[Conrad2] Keith Conrad, *Bilinear forms*,
`https://kconrad.math.uconn.edu/blurbs/linmultialg/bilinearform.pdf`

[Conrad3] Keith Conrad, *Conjugation in a group*.
`https://kconrad.math.uconn.edu/blurbs/grouptheory/conjclass.pdf`

[DanRot78] Ottavio D'Antona, Gian-Carlo Rota, *Two Rings Connected with the Inclusion-Exclusion Principle*, Journal of Combinatorial Theory, Series A, Volume 24, Issue 3, May 1978, pp. 395–402.

[daSilv12] Patrick Da Silva, *Polynomial in* $\mathbb{Q}[x]$ *sending integers to integers?, math.stackexchange answer #108318*.

[Day16] Martin V. Day, *An Introduction to Proofs and the Mathematical Vernacular*, 7 December 2016.
`https://web.archive.org/web/20180521211821/https://www.math.vt.edu/people/day/ProofsBook/IPaMV.pdf` .

[DiaGra77] Persi Diaconis, R. L. Graham, *Spearman's Footrule as a Measure of Disarray*, Journal of the Royal Statistical Society, Series B **39**, No. 2 (1977), pp. 262–268.
`https://statweb.stanford.edu/~cgates/PERSI/papers/77_04_spearmans.pdf`

[dilemi17] Dilemian et al, *math.stackexchange post #2455428 ("An identity involving binomial coefficients and rational functions")*, `https://math.stackexchange.com/q/2455428` .

[DouNie19] Darrin Doud and Pace P. Nielsen, *A Transition to Advanced Mathematics*, 10 June 2019.
`https://math.byu.edu/~pace/Transition-ereader.pdf`

[EdeStr04] Alan Edelman and Gilbert Strang, *Pascal Matrices*, American Mathematical Monthly, Vol. 111, No. 3 (March 2004), pp. 189–197.

[Eisenk99] Theresia Eisenkölbl, *Rhombus Tilings of a Hexagon with Three Fixed Border Tiles*, arXiv:math/9712261v2. Published in: J. Combin. Theory Ser. A 88 (1999), pp. 368–378.

[Elman19] Richard Elman, *Lectures on Abstract Algebra*, 25 September 2021.
`https://www.math.ucla.edu/~rse/algebra_book.pdf`

[FadSom72] D. Faddeev, I. Sominsky, *Problems in Higher Algebra*, Mir 1972.

[Fische01] Hanspeter Fischer, *Linear recurrence relations with constant coefficients*, 9 April 2001.
`http://matematicas.uam.es/~mavi.melian/CURSO_15_16/web_Discreta/recurrence.pdf`

[FomZel01] Sergey Fomin, Andrei Zelevinsky, *Cluster Algebras I: Foundations*, Journal of the American Mathematical Society, Volume 15, Number 2, pp. 497–529.
`https://doi.org/10.1090/S0894-0347-01-00385-X`

[FomZel02] Sergey Fomin, Andrei Zelevinsky, *The Laurent Phenomenon*, Advances in Applied Mathematics **28** (2002), pp. 119–144.

[FoWiZe16] Sergey Fomin, Lauren Williams, Andrei Zelevinsky, *Introduction to Cluster Algebras. Chapters 1-3*, arXiv:1608.05735v2.

[FraYea21] Melanie Fraser, Karen Yeats, *Column expansion identities and quadratic spanning forest identities*, arXiv:2109.13401v3.

[Fulton97] William Fulton, *Young Tableaux: With Applications to Representation Theory and Geometry*, Cambridge University Press 1997.

[Gale98] David Gale, *Tracking the automatic ant and other mathematical explorations*, Springer 1998.

[GalKar71] David Gale, Richard M. Karp, *A phenomenon in the theory of sorting*, Journal of Computer and System Sciences, Volume 6, Issue 2, April 1972, pp. 103–115.

[GalQua22] Jean Gallier and Jocelyn Quaintance, *Algebra, Topology, Differential Calculus, and Optimization Theory For Computer Science and Machine Learning*, 18 March 2022.
`https://www.cis.upenn.edu/~jean/gbooks/geomath.html`

[Galvin17] David Galvin, *Basic discrete mathematics*, 13 December 2017.
`http://www-users.math.umn.edu/~dgrinber/comb/60610lectures2017-Galvin.pdf`
(The URL might change, and the text may get updated. In order to reliably obtain the version of 13 December 2017, you can use the archive.org Wayback Machine: `https://web.archive.org/web/20180205122609/http://www-users.math.umn.edu/~dgrinber/comb/60610lectures2017-Galvin.pdf` .)

[Gessel79] Ira Gessel, *Tournaments and Vandermonde's Determinant*, Journal of Graph Theory, Vol. 3 (1979), pp. 305–307.

[Gessel92] Ira M. Gessel, *Super Ballot Numbers*, Journal of Symbolic Computation, Volume 14, Issues 2–3, August-September 1992, pp. 179–194.
`https://doi.org/10.1016/0747-7171(92)90034-2`

[GohKol96] Israel Gohberg, Israel Koltracht, *Triangular factors of Cauchy and Vandermonde matrices*, Integr. Equat. Oper. Th. Vol. 26 (1996), pp. 46–59.

[Goodma15] Frederick M. Goodman, *Algebra: Abstract and Concrete*, edition 2.6, 1 May 2015.
`http://homepage.math.uiowa.edu/~goodman/algebrabook.dir/book.2.6.pdf` .

[Gorin21] Vadim Gorin, *Lectures on random lozenge tilings*, 10 February 2021, to be published at Cambridge University Press.
`https://people.math.wisc.edu/~vadicgor/Random_tilings.pdf`

[Gould10] H. W. Gould, *Combinatorial Identities: Table I: Intermediate Techniques for Summing Finite Series*, Edited and Compiled by Jocelyn Quaintance, May 3, 2010.

[Grinbe09] Darij Grinberg, *Solution to Problem 19.9 from "Problems from the Book"*.
`http://www.cip.ifi.lmu.de/~grinberg/solutions.html`

[Grinbe10] Darij Grinberg, *A hyperfactorial divisibility*, version of 27 July 2015.
`http://www.cip.ifi.lmu.de/~grinberg/`

[Grinbe11] Darij Grinberg, *Zeckendorf family identities generalized*, arXiv preprint `arXiv:1103.4507v2`.

[Grinbe16a] Darij Grinberg, *4th QEDMO, Problem 13 with solution.*

[Grinbe16b] Darij Grinberg, *Notes on linear algebra*, version of 26 September 2019.
`https://github.com/darijgr/lina`

[Grinbe21] Darij Grinberg, *An Introduction to Algebraic Combinatorics [Math 701, Spring 2021 lecture notes]*, 3 May 2022.
`https://www.cip.ifi.lmu.de/~grinberg/t/21s/lecs.pdf`

[Grinbe22] Darij Grinberg, *UMN, Spring 2017, Math 5707: Lecture 8 (Vandermonde determinant using tournaments)*, 10 May 2022.
`https://www.cip.ifi.lmu.de/~grinberg/t/17s/5707lec8.pdf`

[GriRei20] Darij Grinberg, Victor Reiner, *Hopf algebras in Combinatorics*, version of 27 July 2020, `arXiv:1409.8356v7`.
See also `http://www.cip.ifi.lmu.de/~grinberg/algebra/HopfComb-sols.pdf` for a version that gets updated.

[GrKnPa94] Ronald L. Graham, Donald E. Knuth, Oren Patashnik, *Concrete Mathematics, Second Edition*, Addison-Wesley 1994.
See `https://www-cs-faculty.stanford.edu/~knuth/gkp.html` for errata.

[Hammac15] Richard Hammack, *Book of Proof*, 3rd edition 2018.
`https://www.people.vcu.edu/~rhammack/BookOfProof/`

[HanKra00] Guo-Niu Han and Christian Krattenthaler, *Rectangular Scott-type permanents*, Séminaire Lotharingien Combin. **43** (2000), Article B43g, 25 pp.

[Heffer20] Jim Hefferon, *Linear Algebra*, 4th edition, 8 May 2020,
`http://joshua.smcvt.edu/linearalgebra/` .

[Heinig11] Peter Christian Heinig, *Chio Condensation and Random Sign Matrices*, arXiv:1103.2717v3.

[Herstei75] I. N. Herstein, *Topics in Algebra*, 2nd edition, John Wiley & Sons, 1975.

[HofKun71] Kenneth Hoffman, Ray Kunze, *Linear algebra*, 2nd edition, Prentice-Hall 1971.
See `https://math.stackexchange.com/questions/437253/` for errata.

[Hunger03] Thomas W. Hungerford, *Algebra*, 12th printing, Springer 2003.

[Hunger14] Thomas W. Hungerford, *Abstract Algebra: An Introduction*, 3rd edition, Brooks/Cole 2014.

[Jacobs10] Nathan Jacobson, *Finite-Dimensional Division Algebras Over Fields*, corrected 2nd printing, Springer 2010.

[Joyce17] David Joyce, *Introduction to Modern Algebra*, 5 December 2017.
`https://mathcs.clarku.edu/~djoyce/ma225/algebra.pdf`

[Joyner08] W. D. Joyner, *Mathematics of the Rubik's cube*, 19 August 2008.
`https://web.archive.org/web/20160304122348/http://www.permutationpuzzles.org/rubik/webnotes/` (link to the PDF at the bottom).

[KarZha16] Karthik Karnik, Anya Zhang, *Combinatorial proof of Chio Pivotal Condensation*, 25 May 2016.

[KenWil14] Richard W. Kenyon, David B. Wilson, *The space of circular planar electrical networks*, arXiv:1411.7425v3.

[Kerber99] Adalbert Kerber, *Applied Finite Group Actions*, 2nd edition, Springer 1999.

[KleLak72] S. L. Kleiman and Dan Laksov, *Schubert Calculus*, The American Mathematical Monthly, Vol. 79, No. 10, (Dec., 1972), pp. 1061–1082.

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

[Knuth1] Donald Ervin Knuth, *The Art of Computer Programming, volume 1: Fundamental Algorithms*, 3rd edition, Addison–Wesley 1997.
See `https://www-cs-faculty.stanford.edu/~knuth/taocp.html` for errata.

[Knuth88] Donald E. Knuth, *Fibonacci Multiplication*, Appl. Math. Lett., Vol. 1, No. 1, pp. 57–60, 1988.

[Knutso12] Allen Knutson, *Schubert polynomials and Symmetric Functions*, lecture notes, July 28, 2012.

[Kratte05] Christian Krattenthaler, *Advanced Determinant Calculus: A Complement*, Linear Algebra Appl. **411** (2005), pp. 68–166, arXiv:math/0503507v2.

[Kratte99] Christian Krattenthaler, *Advanced Determinant Calculus*, Séminaire Lotharingien Combin. 42 (1999) (The Andrews Festschrift), paper B42q, 67 pp., arXiv:math/9902004v3.

[KurLis78] L. Kurlyandchik, A. Lisitskiy, *Summy i proizvedeniya ("Sums and products")*, Kvant 1978, issue 10, pp. 31–37.
`http://kvant.mccme.ru/1978/10/summy_i_proizvedeniya.htm`

[LLPT95] D. Laksov, A. Lascoux, P. Pragacz, and A. Thorup, *The LLPT Notes*, edited by A. Thorup, 28 March 2018,
`http://web.math.ku.dk/noter/filer/sympol.pdf` .

[Lalond96] Pierre Lalonde, *A non-commutative version of Jacobi's equality on the cofactors of a matrix*, Discrete Mathematics 158 (1996), pp. 161–172.
`https://doi.org/10.1016/0012-365X(95)00040-4` .

[Lampe13] Philipp Lampe, *Cluster algebras*, 4 December 2013,
`http://www.math.uni-bielefeld.de/~lampe/teaching/cluster/cluster.pdf` .

[Lang02] Serge Lang, *Algebra*, Revised Third Edition, Graduate Texts in Mathematics #211, Springer 2002.

[Laue15] Hartmut Laue, *Determinants*, version 17 May 2015,
`http://www.math.uni-kiel.de/algebra/laue/homepagetexte/det.pdf` .

[LeeSch11] Kyungyong Lee, Ralf Schiffler, *A Combinatorial Formula for Rank* 2 *Cluster Variables*, Journal of Algebraic Combinatorics **37** (2013), Issue 1, pp. 67–85.
A preprint is available at `http://arxiv.org/abs/1106.0952v3` .

[LeeSch13] Kyungyong Lee, Ralf Schiffler, *Positivity for cluster algebras*, Annals of Mathematics (second series) **182**, No. 1 (July, 2015), pp. 73–125.
A preprint is available at `http://arxiv.org/abs/1306.2415v3` .

[Leeuwe06] Marc A. A. van Leeuwen, *Schur functions and alternating sums*, Electronic Journal of Combinatorics Vol 11(2) A5 (2006), also available as arXiv:math.CO/0602357.

[LeLeMe16] Eric Lehman, F. Thomson Leighton, Albert R. Meyer, *Mathematics for Computer Science*, revised Tuesday 6th June 2018,
`https://courses.csail.mit.edu/6.042/spring18/mcs.pdf` .

[Loehr11] Nicholas A. Loehr, *Bijective Combinatorics*, Chapman & Hall/CRC 2011.

[Manive01] Laurent Manivel, *Symmetric Functions, Schubert Polynomials and Degeneracy Loci*, SMF/AMS Texts and Monographs #6, American Mathematical Society 2001.

[Marsh13] Robert J. Marsh, *Lecture Notes on Cluster Algebras*, European Mathematical Society 2013.
See `http://www1.maths.leeds.ac.uk/~marsh/errata.pdf` for a list of errata.

[Mate14] Attila Maté, *Determinants*, version 1 October 2017.
`http://www.sci.brooklyn.cuny.edu/~mate/misc/determinants.pdf`

[Muir30] Thomas Muir, *The theory of determinants in the historical order of development*, 5 volumes (1906–1930), later reprinted by Dover.
`http://www-igm.univ-mlv.fr/~al/`

[MuiMet60] Thomas Muir, *A Treatise on the Theory of Determinants*, revised and enlarged by William H. Metzler, Dover 1960.

[Mulhol16] Jamie Mulholland, *Permutation Puzzles: A Mathematical Perspective*,
`http://www.sfu.ca/~jtmulhol/math302/`

[MusPro07] Gregg Musiker, James Propp, *Combinatorial Interpretations for Rank-Two Cluster Algebras of Affine Type*, The Electronic Journal of Combinatorics **14** (2007), #R15.
`http://www.combinatorics.org/ojs/index.php/eljc/article/view/v14i1r15`

[Newste19] Clive Newstead, *An Infinite Descent into Pure Mathematics*, version 0.4 (1 January 2020).
`https://infinitedescent.xyz`

[NouYam02] Masatoshi Noumi, Yasuhiko Yamada, *Tropical Robinson-Schensted-Knuth correspondence and birational Weyl group actions*, arXiv:math-ph/0203030v2.

[OlvSha18] Peter J. Olver, Chehrzad Shakiban, *Applied Linear Algebra*, 2nd edition, Springer 2018.
`https://doi.org/10.1007/978-3-319-91041-3`
See `http://www.math.umn.edu/~olver/ala.html` for errata.

[OruPhi00] Halil Oruç, George M. Phillips, *Explicit factorization of the Vandermonde matrix*, Linear Algebra and its Applications 315 (2000), pp. 113–123.

[PetTen14] T. Kyle Petersen, Bridget Eileen Tenner, *The depth of a permutation*, Journal of Combinatorics **6** (2015), Number 1–2, pp. 145–178, arXiv:1202.4765v3.

[Pinkha15] Henry C. Pinkham, *Linear Algebra*, draft of a textbook, version 10 July 2015.
`https://www.math.columbia.edu/~pinkham/HCP_LinearAlgebra.pdf`

[Prasol94] Viktor V. Prasolov, *Problems and Theorems in Linear Algebra*, Translations of Mathematical Monographs, vol. #134, AMS 1994.
See `https://staff.math.su.se/mleites/books/prasolov-1994-problems.pdf` for a preprint.
See also `https://drive.google.com/open?id=0B2UfTLwpN9okblBJbGx0ZXc4Rm8` for a newer edition in Russian.

[Richma88] Fred Richman, *Nontrivial uses of trivial rings*, Proceedings of the American Mathematical Society, Volume 103, Number 4, August 1988, pp. 1012–1014.

[Riorda68] John Riordan, *Combinatorial Identities*, Robert E. Krieger Publishing Company, Huntington, NY 1979.

[Rocket81] Andrew M. Rockett, *Sums of the Inverses of Binomial Coefficients*, The Fibonacci Quarterly 19 (1981), issue 5, pp. 433–437.
`https://www.fq.math.ca/Scanned/19-5/rockett.pdf`

[Rote01] Günter Rote, *Division-Free Algorithms for the Determinant and the Pfaffian: Algebraic and Combinatorial Approaches*, Computational Discrete Mathematics, Lecture Notes in Computer Science, Volume 2122, 2001, pp. 119–135.
`https://web.archive.org/web/20161005071023/http://www.inf.fu-berlin.de/groups/ag-ti/people/rote/Papers/pdf/Division-free+algorithms.pdf`

[Rotman15] Joseph J. Rotman, *Advanced Modern Algebra: Part 1*, Graduate Studies in Mathematics #165, 3rd edition, AMS 2015.

[SacUlf11] Joshua Sack, Henning Úlfarsson, *Refined inversion statistics on permutations*, The Electronic Journal of Combinatorics **19** (2012), #P29. See also arXiv:1106.1995v2 for a preprint version.

[Sagan19] Bruce Sagan, *Combinatorics: The Art of Counting*, Graduate Studies in Mathematics **210**, 21 September 2020.
`https://users.math.msu.edu/users/bsagan/Books/Aoc/final.pdf`

[Schmit04] William R. Schmitt, *Incidence Hopf algebras*, Journal of Pure and Applied Algebra **96** (1994), pp. 299–330, `https://doi.org/10.1016/0022-4049(94)90105-8` .

[Silves00] John R. Silvester, *Determinants of Block Matrices*, The Mathematical Gazette, Vol. 84, No. 501 (Nov., 2000), pp. 460–467.

[Spivey12a] Mike Spivey, *math.stackexchange post #180440 ("Alternating sum of squares of binomial coefficients")*, `https://math.stackexchange.com/a/180440/` .

[Spivey12b] Mike Spivey, *Combinatorial Interpretation of the Alternating Convolution of the Central Binomial Coefficients*, `https://mikespivey.wordpress.com/2012/03/16/altconvcentralbinom/` .

[Stanle11] Richard Stanley, *Enumerative Combinatorics, volume 1*, Second edition, Cambridge University Press 2012.
See `http://math.mit.edu/~rstan/ec/` for errata and for a preprint version (from 2011).

[Stanle01] Richard Stanley, *Enumerative Combinatorics, volume 2*, First edition, Cambridge University Press 2001.
See `http://math.mit.edu/~rstan/ec/` for errata.

[Stembr02] John R. Stembridge, *A Concise Proof of the Littlewood-Richardson Rule*, The Electronic Journal of Combinatorics, Volume 9 (2002), Note #N5.

[Strick13] Neil Strickland, *MAS201 Linear Mathematics for Applications*, lecture notes, 11 February 2020.
`https://neilstrickland.github.io/linear_maths/`

[Sury95] B. Sury, *An Integral Polynomial*, Mathematics Magazine **68**, No. 2 (Apr., 1995), pp. 134–135.

[Sved84] Marta Sved, *Counting and Recounting: The Aftermath*, The Mathematical Intelligencer **6** (1984), no. 2, pp. 44–45.

[Swanso20] Irena Swanson, *Introduction to Analysis with Complex Numbers*, 25 July 2020.
`https://web.archive.org/web/20201012174324/http://people.reed.edu/~iswanson/analysisconstructR.pdf`

[Tenner04] Bridget Eileen Tenner, *A Non-Messing-Up Phenomenon for Posets*, Annals of Combinatorics 11 (2007), pp. 101–114.
A preprint is available as arXiv:math/0404396.

[Turnbu29] H. W. Turnbull, *The Theory of Determinants, Matrices, and Invariants*, Blackie & Son Ltd, 1929.
`https://archive.org/details/in.ernet.dli.2015.4951`

[Vellem06] Daniel J. Velleman, *How to Prove It*, 2nd edition, Cambridge University Press 2006.

[Vodick50] Václav Vodička, *Determinanty a matice v theorii a praxi. Část první*, Prague 1950.
`https://dml.cz/handle/10338.dmlcz/403265`

[Vodick51] Václav Vodička, *Determinanty a matice v theorii a praxi. Část druhá*, Prague 1951.
`https://dml.cz/handle/10338.dmlcz/403282`

[Vorobi02] Nicolai N. Vorobiev, *Fibonacci Numbers*, Translated from the Russian by Mircea Martin, Springer 2002 (translation of the 6th Russian edition).

[Walker87] Elbert A. Walker, *Introduction to Abstract Algebra*, Random House/Birkhauser, New York, 1987.

[Willia03] Geordie Williamson, *Mind your P and Q-symbols: Why the Kazhdan-Lusztig basis of the Hecke algebra of type A is cellular*, Honours thesis at the University of Sydney, October 2003.

[Willia15] Stanley Gill Williamson, *The common-submatrix Laplace expansion*, arXiv:1505.05486v1.

[Willia18] Stanley Gill Williamson, *Matrix Canonical Forms: notational skills and proof techniques*, 28 April 2018.

[ZarSam67] Oscar Zariski, Pierre Samuel, *Commutative Algebra, volume 1*, D. van Nostrand Company, 6th printing 1967.

[Zeilbe85] Doron Zeilberger, *A combinatorial approach to matrix algebra*, Discrete Mathematics 56 (1985), pp. 61–72.

[Zeilbe93] Doron Zeilberger, *On an identity of Daubechies*, The American Mathematical Monthly, Vol. 100, No. 5 (May, 1993), p. 487.
`http://sites.math.rutgers.edu/~zeilberg/mamarim/mamarimhtml/wavelet.html`

[Zeilbe98] Doron Zeilberger, *Dodgson's Determinant-Evaluation Rule proved by Two-Timing Men and Women*, The Electronic Journal of Combinatorics, vol. 4, issue 2 (1997) (The Wilf Festschrift volume), R22.
`http://www.combinatorics.org/ojs/index.php/eljc/article/view/v4i2r22`
Also available as arXiv:math/9808079v1.
`http://arxiv.org/abs/math/9808079v1`

[Zeng93] Jian Zeng, *A bijective proof of Muir's identity and the Cauchy-Binet formula*, Linear algebra and its applications **184** (1993), pp. 79–82.