Mathematics for AI

Georg Regensburger

Institut for Algebra
Johannes Kepler University Linz

Version: January 27, 2021

Inhalt

Chapter 1

Functions and relations

1.1 Functions

Many questions in mathematics and its applications can be formulated and studied in terms of functions. In this section, we first give the abstract set-theoretic definition of a function and illustrate it with various examples and classes of functions. In doing so, we outline some of the main topics of the lecture. We also study basic properties of arbitrary functions and discuss the composition of functions and inverse functions.

A function associates to each element of a set X a single element of a set Y. Intuitively, one can think of a function as a procedure that takes as an input an element from X and gives as an output an element from Y. For example, in calculus, we study functions from the real numbers to the real numbers and these functions are often given by formulas like

$$f(x) = x^2, \quad f(x) = \sqrt{x}, \quad f(x) = \mathrm{e}^x, \quad \dots.$$

To formalize what it means for abstract sets that a function maps an $x \in X$ to a $y \in Y$, we consider the pair

$$(x, y) \in X \times Y$$

and we identify a function with the set of all such pairs.

Definition 1. Let X and Y be sets. A **function** from X to Y, denoted by

$$f \colon X \to Y,$$

is a subset

$$f \subseteq X \times Y,$$

such that, for each $x \in X$, there is a unique $y \in Y$ with $(x, y) \in f$.

For each $x \in X$, the unique element $y \in Y$ with $(x, y) \in f$ is called the **value** of f at the **argument** x. We write $y = f(x)$ or

$$f \colon x \mapsto y$$

and say that f **maps** x **to** y.

The set X is called the **domain** of f and Y the **codomain** of f. The words **map** or **mapping** are also used for functions.

Using quantifiers, the *existence* of a y for each x reads as

$$\forall\, x \in X\ \exists\, y \in Y : (x, y) \in f$$

and the *uniqueness* condition as

$$\forall\, x \in X\ \forall\, y_1, y_2 \in Y : (x, y_1) \in f \wedge (x, y_2) \in f \implies y_1 = y_2.$$

We also note that two functions

$$f\colon X \to Y \quad \text{and} \quad g\colon X' \to Y'$$

are *equal* if they have the same domain and codomain, that is, $X = X'$, $Y = Y'$, and

$$\forall\, x \in X\colon f(x) = g(x)$$

As a first example, we consider the identity function that always returns the same value that was used as its argument.

Definition 2. Let X be a set. The **identity function** on X is defined as

$$\mathrm{id}_X\colon X \to X, \quad x \mapsto x.$$

For sets X and Y, we can also consider the set of all functions from X to Y. We first need to introduce the set of all subsets of a given set.

Definition 3. Let X be a set. The **power set** of X is defined as

$$\mathcal{P}(X) = \{Y \mid Y \subseteq X\}.$$

Example. For $X = \{1, 2\}$,

$$\mathcal{P}(X) = \{\emptyset, \{1\}, \{2\}, \{1, 2\}\}.$$

Definition 4. Let X, Y be sets. We denote the **set of all functions** from X to Y by

$$Y^X = \{f \in \mathcal{P}(X \times Y) \mid f\colon X \to Y\}.$$

A main topic of the course will be to study **real functions**, that is, functions

$$f\colon D \to \mathbb{R},$$

whose domain $D \subseteq \mathbb{R}$ is a subset of the reals containing an interval and whose codomain is the real numbers.

Example.

1. **Linear functions**
$$f\colon \mathbb{R} \to \mathbb{R}, \quad x \mapsto ax,$$
where $a \in \mathbb{R}$. These include the **zero function** $x \mapsto 0$ for $a = 0$ as a special case.

2. Adding a constant, we obtain **affine functions**
$$f\colon \mathbb{R} \to \mathbb{R}, \quad x \mapsto ax + b,$$
where $a, b \in \mathbb{R}$. These include linear ($b = 0$) and **constant functions** ($a = 0$) as special cases.

3. More generally, we have **polynomial functions**
$$p\colon \mathbb{R} \to \mathbb{R}, \quad p(x) = a_0 + a_1 x + \cdots + a_n x^n,$$
where $a_0, a_1, \ldots, a_n \in \mathbb{R}$.

4. **Rational functions** are function $f\colon D \to \mathbb{R}$ that can be written in the form
$$f(x) = \frac{p(x)}{q(x)},$$
where p and q are polynomial functions and q is not the zero function. The domain D of f is the set of all x for which the denominator is not zero
$$D = \{x \in \mathbb{R} \mid q(x) \neq 0\}.$$
For example,
$$f\colon \mathbb{R} \setminus \{0\} \to \mathbb{R}, \quad f(x) = \frac{1}{x}.$$

5. Additional **elementary functions** are, for example, **root functions**,
$$f\colon D \to \mathbb{R}, \quad x \mapsto \sqrt[n]{x} = x^{\frac{1}{n}},$$
where $D = [0, \infty)$ and $n \geq 2$,

6. **trigonometric functions** like $\sin(x)$ and $\cos(x)$,

7. or the **exponential function**
$$a^x = \mathrm{e}^{x \ln a}$$
with base a, where $a > 0$ and $a \neq 1$. The corresponding **logarithm**
$$\log_a x = \frac{\ln x}{\ln a}$$
to base a with domain $(0, \infty)$ is the inverse of the exponential function, that is,
$$y = a^x \iff x = \log_a y.$$

8. Sometimes functions are defined by case distinctions. Important examples are the **absolute value function**

$$|\cdot|\colon \mathbb{R} \to \mathbb{R}, \quad |x| = \begin{cases} x, & \text{if } x \geq 0, \\ -x, & \text{if } x < 0, \end{cases}$$

9. or the **sign function**

$$\operatorname{sgn}\colon \mathbb{R} \to \mathbb{R}, \quad \operatorname{sgn}(x) = \begin{cases} 1, & x > 0, \\ 0, & x = 0, \\ -1, & x < 0. \end{cases}$$

For real functions, we will study important properties like *continuity* and *differentiability*. We will see how the *derivative* allows us to approximate a differentiable function by a linear function, or, more generally, how we can approximate a *smooth function* by a polynomial function. We will consider (local) *minima* and *maxima* of differentiable functions as a tool to solve *optimization problems* in applications. We will also discuss *antiderivatives* (*indefinite integrals*) and *definite integrals* of real functions.

For defining continuity and differentiability, we will study the notion of *convergence* of a sequence of real numbers or real vectors to a *limit*. Formally, a sequence in a set X is just a function from the natural numbers to X. However, we think of a sequence as prescribing a zeroth, first, second, ..., nth element in X.

Definition 5. Let X be a set. A **sequence in** X is a function

$$a\colon \mathbb{N} \to X.$$

For $n \in \mathbb{N}$, we write $a_n = a(n)$ for the n**th element** of a sequence and

$$(a_n)_{n\in\mathbb{N}} = (a_n)_{n=0}^{\infty} = (a_0, a_1, a_2, \ldots)$$

for the sequence a.

Note that, in this lecture, we define the natural numbers to include zero and so sequences start from 0. Often, one considers also sequences starting at 1. To distinguish the two cases, we use the notation

$$(a_n)_{n\geq 0} = (a_0, a_1, a_2, \ldots) \quad \text{and} \quad (a_n)_{n\geq 1} = (a_1, a_2, a_3, \ldots).$$

Example.

1. For an arbitrary non-empty set X and an $x \in X$, we can consider the **constant sequence** $a\colon n \mapsto x$ in X, that is, a is the sequence

$$(x, x, x, \ldots).$$

2. The real sequence $a_n = 1/n$ for $n \geq 1$ is

$$(a_n)_{n\geq 1} = (1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \ldots).$$

3. The **Fibonacci numbers** F_n form a sequence such that each number is the sum of the two preceding ones, starting from 0 and 1. So they are defined by the recurrence

$$F_{n+2} = F_n + F_{n+1}$$

for all $n \in \mathbb{N}$ with the initial conditions $F_0 = 0$ and $F_1 = 1$. Hence the beginning of this sequence is

$$(0, 1, 1, 2, 3, 5, 8, 13, 21, \ldots).$$

To specify a function $f\colon X \to Y$ between finite sets, we can also list all pairs

$$(x, f(x)) \in X \times Y.$$

Example. Let $X = \{1, 2, 3\}$ and $Y = \{a, b, c\}$. Which of the following subsets of $X \times Y$ defines a function and, if not, which condition is violated? Recall that a function has to assign to each $x \in X$ a unique $y \in Y$.

1. $\{(1, a), (2, c), (3, b)\}$,
2. $\{(1, a), (2, b), (3, b)\}$,
3. $\{(1, a), (3, b)\}$,
4. $\{(1, a), (2, b), (2, c), (3, a)\}$.

So far we considered only real functions depending on one real variable, that is, the domain is a subset $D \subseteq \mathbb{R}$. Such functions are also called **univariate functions**. In the abstract definition of a function, the domain X can also be subset of a cartesian product and then its elements consist of tuples. For example, $D \subseteq \mathbb{R}^2$ with elements (x, y), or, more generally,

$$D \subseteq \mathbb{R}^n$$

with elements being n**-tuples**

$$(x_1, \ldots, x_n).$$

Such functions, depending on several (real) variables, are called **multivariate functions**. If the codomain of a function is the real numbers, one speaks of a **real-valued function**. So a multivariate real-valued function is a function

$$f\colon D \to \mathbb{R}, \quad (x_1, \ldots, x_n) \mapsto f(x_1, \ldots, x_n).$$

Note that by convention one omits the parentheses surrounding tuples, whereas the general notation for functions would be $f((x_1, \ldots, x_n))$.
To get a first impression, we look at some simple examples of **bivariate functions**

$$(x, y) \mapsto f(x, y),$$

depending on two variables, which can be visualized in $\mathbb{R}^3$.

Example.

1. Linear functions
$$f\colon \mathbb{R}^2 \to \mathbb{R}, \quad f(x,y) = ax + by,$$
where $a, b \in \mathbb{R}$. Note that such a linear function is determined by the vector
$$(a,b) \in \mathbb{R}^2.$$
2. An example of a multivariate polynomial function is
$$f\colon \mathbb{R}^2 \to \mathbb{R}, \quad f(x,y) = x^2 + y^2,$$
3. and a function involving a trigonometric real function
$$f\colon \mathbb{R}^2 \to \mathbb{R}, \quad f(x,y) = \sin(x^2 + y^2).$$

If the codomain of a function is $\mathbb{R}^m$ with $m > 1$, one speaks of a **vector-valued function**. We look again at some examples with codomain $\mathbb{R}^2$. In this case, one can only visualize the values of a function
$$f\colon D \subseteq \mathbb{R} \to \mathbb{R}^2, \quad t \mapsto (f_1(t), f_2(t))$$
as a curve in the plane but not functions $f\colon \mathbb{R}^2 \to \mathbb{R}^2$.

Example.

1. The function
$$f\colon \mathbb{R} \to \mathbb{R}^2, \quad f(t) = (t^2, t^3)$$
parametrizes a cusp and the function
$$f\colon [0, 2\pi) \to \mathbb{R}^2, \quad f(t) = (\sin(t), \cos(t))$$
parametrizes the unit circle.
2. Linear functions
$$f\colon \mathbb{R}^2 \to \mathbb{R}^2, \quad (x,y) \mapsto (ax + by, cx + dy),$$
where $a, b, c, d \in \mathbb{R}$. Note that such a linear function is determined by the matrix
$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathbb{R}^{2\times 2}.$$
3. As a more complicated example, consider the function
$$f\colon \mathbb{R}^2 \to \mathbb{R}^2, \quad (x,y) \mapsto (xy, \sin(x^2 + y^2)).$$

Multivariate functions are crucial for modeling many problems in applications and of course also in machine learning and artificial intelligence. In the lecture, we will study *linear multivariate functions* (*linear maps*) and the corresponding *matrices*, which are central objects in *linear algebra*. Linear maps are also the basic tool for developing *multivariate calculus*. For real functions, the derivative at a point is a real number. We will see that for *multivariate real-valued functions*, the derivative at a point is a *vector*, and, for *vector-valued functions*, it is a *matrix* and we will discuss how we can approximate a differentiable function by the corresponding linear map. For studying optimization problems, we will study the second derivative of real-valued functions and the tools needed from linear algebra to analyze the corresponding *quadratic forms*.

In the following, we discuss some basic properties and notions for arbitrary functions. Given a function

$$f\colon X \to Y \quad \text{and} \quad y \in Y,$$

we can ask whether there exists an $x \in X$ such that

$$f(x) = y.$$

If there exists *at most one* such $x \in X$, the function is called *injective*. If we can find for every $y \in Y$ *at least one* such $x \in X$, the function is called *surjective*. It is called *bijective* if for every $y \in Y$ there *exists exactly one* $x \in X$ such that $f(x) = y$.

Definition 6. Let $f\colon X \to Y$ be a function.

1. f is **injective** (an **injection**) if

$$\forall\, x_1, x_2 \in X : f(x_1) = f(x_2) \implies x_1 = x_2.$$

2. f is **surjective** (a **surjection**) if

$$\forall\, y \in Y\ \exists\, x \in X : f(x) = y.$$

3. f is **bijective** (a **bijection**) if it is injective and surjective.

Example.

1. The identity function id_X on a set X is bijective.

2. The function $f\colon \mathbb{R} \to \mathbb{R}$, $f(x) = 2x + 1$ is bijective.

   To prove that f is injective, let $x_1, x_2 \in \mathbb{R}$ such that $f(x_1) = f(x_2)$, that is,

$$2x_1 + 1 = 2x_2 + 1.$$

   Hence $2x_1 = 2x_2$ and $x_1 = x_2$. We conclude that f is injective.

   To prove that f is surjective, let $y \in \mathbb{R}$. We have to find an $x \in \mathbb{R}$ such that

$$f(x) = 2x + 1 = y.$$

   Solving for x, we obtain

$$x = \frac{1}{2}y - \frac{1}{2}.$$

   We verify that $f(x) = 2(\frac{1}{2}y - \frac{1}{2}) + 1 = y$ and conclude that f is surjective.

3. The function

$$f\colon \mathbb{R} \to \mathbb{R}, \quad x \mapsto x^2$$

   is neither injective nor surjective.

   It is not injective, since, for example, $f(2) = f(-2) = 4$. It is not surjective, since, $f(x) \geq 0$ for all $x \in \mathbb{R}$ and, in particular, $f(x) \neq -1$ for all $x \in \mathbb{R}$ but -1 is an element of the codomain of f.

By the contrapositive, a function is *injective* if and only if (shortened as iff) it *maps distinct arguments* of the domain to *distinct values*, that is,

$$\forall\, x_1, x_2 \in X : x_1 \neq x_2 \implies f(x_1) \neq f(x_2).$$

However, usually the definition of injectivity above is easier to verify. Note that a function being injective does not depend on the codomain.

Proving surjectivity is usually more difficult because one has to find an explicit solution for equation $f(x) = y$ for every element y in the codomain, like in the example above, or, one has to prove by other methods that such a solution exists. A function being surjective obviously depends on the codomain.

By *restricting* the domain/codomain of a function to suitable subsets, one can obtain injective/surjective functions. We first look at some examples and then we introduce the corresponding definitions.

Example.

1. The function
$$f\colon [0,\infty) \to \mathbb{R}, \quad x \mapsto x^2$$
is injective (but not surjective).

2. Using the square root $\sqrt{x}$ of a nonnegative real number $x \in \mathbb{R}$, we see that
$$f\colon \mathbb{R} \to [0,\infty), \quad x \mapsto x^2$$
is surjective (but not injective).

3. The function
$$f\colon [0,\infty) \to [0,\infty), \quad x \mapsto x^2$$
is bijective.

Definition 7. Let $f\colon X \to Y$ be a function and $S \subseteq X$ a subset.

1. The **restriction of f to S** is the function
$$f|_S\colon S \to Y, \quad x \mapsto f(x).$$

2. The **(direct) image of S under f** is defined as
$$f(S) = \{\, y \in Y \mid \exists\, x \in S : f(x) = y \,\} \subseteq Y.$$

3. The **image of f** (or **range of f**) is defined as
$$\operatorname{im} f = f(X).$$

Note that if S is a proper subset of X, then the restriction

$$f|_S \neq f,$$

since they have different domains. By definition, we have that f is *surjective* iff

$$\operatorname{im} f = Y.$$

Hence, if we restrict the codomain of a function to its image, we obtain a surjective function.

The image of a subset under a function associates to every subset of the domain a subset of the codomain. Conversely, the inverse image associates to every subset of the codomain a subset of the domain

Definition 8. Let $f\colon X \to Y$ be a function and $W \subseteq Y$ a subset. The **inverse image** (or **preimage**) **of** W **under** f is defined as

$$f^{-1}(W) = \{x \in X \mid f(x) \in W\} \subseteq X.$$

Note that a function is *injective* iff for each $y \in Y$ the inverse image

$$f^{-1}(\{y\}) \subseteq X$$

consists of at most one element.

Using the notion of bijectivity, we can also formalize when two sets have the same number of elements and one uses the following terminology.

Definition 9. We say that a set X is **finite** if for some $n \in \mathbb{N}$, $n \geq 1$, there exists a bijection

$$f\colon X \to \{1, \ldots, n\}.$$

In this case, we denote the number of elements (or the **cardinality**) of X by

$$|X| = n.$$

The empty set $\emptyset$ is considered finite and we define $|\emptyset| = 0$.

Definition 10. Let X be a set. We say

1. X is **infinite** if it is not finite.
2. X is **countable** if it is finite or there exists a bijection $f\colon \mathbb{N} \to X$.
3. X is **uncountable** if it is not countable.

Example. One can prove that

1. the set of rational numbers $\mathbb{Q}$ is countable.
2. the set of real numbers $\mathbb{R}$ is uncountable.

For *finite* sets X and Y with the same number of elements,

$$|X| = |Y|,$$

it is intuitively clear that a function $f\colon X \to Y$

$$\text{is } \textit{injective} \text{ iff it is } \textit{surjective} \text{ iff it is } \textit{bijective}.$$

To prove this fact one needs a formal definition of the natural numbers. However, this equivalence does not hold for *infinite* sets.

Example. The function

$$f\colon \mathbb{N} \to \mathbb{N}, \quad x \mapsto x + 1,$$

is injective but not surjective, since $0 \notin \operatorname{im} f$.

Bijective functions from a set onto itself are also called *permutations*. The term *permutation* is mainly used for *finite sets* and one uses the following notation.

Definition 11. Let $X = \{1, 2, \ldots, n\}$ with $n \in \mathbb{N}$, $n \geq 1$. A **permutation** of X is a bijective function $\pi\colon X \to X$ and one writes

$$\pi = \begin{pmatrix} 1 & 2 & \cdots & n \\ \pi(1) & \pi(2) & \cdots & \pi(n) \end{pmatrix}.$$

The **set of all permutations** of X is denoted by S_n.

Example. The permutation

$$\pi = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 4 & 1 & 3 & 5 & 2 \end{pmatrix}$$

maps $1 \mapsto 4$, $2 \mapsto 1$, $3 \mapsto 3$, $4 \mapsto 5$, $5 \mapsto 2$.

There are

$$|S_n| = n! = \prod_{k=1}^{n} k$$

permutations, which can be proved by induction.

Let X and Y be finite sets. One can also prove by induction the following identities for the cardinality of the *cartesian product*, the *set of all functions* from X to Y, and the *power set* of X:

$$|X \times Y| = |X||Y|, \qquad |Y^X| = |Y|^{|X|}, \qquad |\mathcal{P}(X)| = 2^{|X|}.$$

In the remainder of this section, we discuss the *composition* of functions. If the *codomain* and the *domain* of two functions

$$f\colon X \to Y \quad \textit{and} \quad g\colon Y \to Z$$

match, we can compose the two functions to a new function by first evaluating f on x and then evaluating g on $f(x)$. In particular, composition allows to construct more complex functions from simpler ones.

Definition 12. Let $f\colon X \to Y$ and $g\colon Y \to Z$ be functions. The **composition of g and f**, denoted by $g \circ f$, is the function

$$g \circ f\colon X \to Z, \quad x \mapsto g(f(x)).$$

Example. The elementary function

$$\mathrm{e}^{\sin x}$$

is obtained by composing $g(x) = \mathrm{e}^x$ with $f(x) = \sin x$. The *chain rule* in calculus is a formula for computing the derivative $(g \circ f)'$ in terms to g' and f':

$$(g \circ f)' = g'(f(x))f'(x).$$

We will study the chain rule for univariate and multivariate functions later in detail.

First, we show that *composition* of functions *preserves injectivity/surjectivity*. In particular, it follows that the *composition of bijective functions* is *bijective*.

Theorem 1. The composition of injective/surjective functions is injective/surjective.

Proof. Let $f\colon X \to Y$ and $g\colon Y \to Z$ be injective functions. We prove that $g \circ f$ is injective. Let $x_1, x_2 \in X$ such that $(g \circ f)(x_1) = (g \circ f)(x_2)$, that is,

$$g(f(x_1)) = g(f(x_2)).$$

Since g is injective, it follows that $f(x_1) = f(x_2)$ and since f is injective, it follows that $x_1 = x_2$. Hence $g \circ f$ is injective.
The statement about the composition of surjective functions is left as an exercises. ∎

Note that for functions $f, g\colon X \to X$ on a set X both compositions

$$f \circ g \quad \text{and} \quad g \circ f$$

are defined but, in general, they are not equal. So composition of functions is *not commutative*.

Example. Let $f, g\colon \mathbb{R} \to \mathbb{R}$ be given by $f(x) = x + 1$, $g(x) = x^2$. Then,

$$(f \circ g)(x) = x^2 + 1 \quad \text{and} \quad (g \circ f)(x) = (x + 1)^2,$$

and hence

$$f \circ g \neq g \circ f.$$

However, composition of functions is always *associative*.

Lemma 2. Let $f\colon X \to Y$, $g\colon Y \to Z$, and $h\colon Z \to W$ be functions. Then,

$$(h \circ g) \circ f = h \circ (g \circ f).$$

Proof. By definition of the composition, the two functions

$$(h \circ g) \circ f \quad \text{and} \quad h \circ (g \circ f)$$

have the same domain X and codomain W. Let $x \in X$. We have

$$((h \circ g) \circ f)(x) = (h \circ g)(f(x)) = h(g(f(x)))$$

and

$$(h \circ (g \circ f))(x) = h((g \circ f)(x)) = h(g(f(x)))$$

and hence the two functions are equal. ∎

So we need not write parentheses and the notation $h \circ g \circ f$ is unambiguous.

The next lemma shows that composition with the identity function behaves like multiplication by 1 for numbers.

Lemma 3. Let $f\colon X \to Y$ be a function. Then,

$$\mathrm{id}_Y \circ f = f = f \circ \mathrm{id}_X\,.$$

Proof. All three functions have the same domain X and codomain Y. Let $x \in X$. We have

$$(\mathrm{id}_Y \circ f)(x) = \mathrm{id}_Y(f(x)) = f(x) \quad \text{and} \quad (f \circ \mathrm{id}_X)(x) = f(\mathrm{id}_X(x)) = f(x)$$

and so the three functions are equal. ∎

In the following, we show that a function is bijective iff it is invertible in the following sense.

Definition 13. A function $f\colon X \to Y$ is called **invertible** if there exists a function $g\colon Y \to X$ such that

$$g \circ f = \mathrm{id}_X \quad \text{and} \quad f \circ g = \mathrm{id}_Y\,.$$

Such a function g is uniquely determined. If f is invertible, we call g the **inverse** of f and write

$$f^{-1} = g.$$

Example.

1. The inverse of
$$f\colon [0,\infty) \to [0,\infty), \quad f(x) = x^2$$
is
$$f^{-1}\colon [0,\infty) \to [0,\infty), \quad f^{-1}(x) = \sqrt{x}.$$

2. The inverse of the permutation
$$\pi = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 4 & 1 & 3 & 5 & 2 \end{pmatrix}$$
is obtained by reversing the arrows $1 \leftarrow\!\shortmid 4$, $2 \leftarrow\!\shortmid 1$, $3 \leftarrow\!\shortmid 3$, $4 \leftarrow\!\shortmid 5$, $5 \leftarrow\!\shortmid 2$, hence
$$\pi^{-1} = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 5 & 3 & 1 & 4 \end{pmatrix}.$$

Note that the *same notation* f^{-1} is used for two different things: for the *inverse image* of a set under a function and for the *inverse* of a function. From the context it is usually clear what we mean when we write f^{-1}.

In fact, the definition above contains also the statement that g is *uniquely determined*, which we still have to prove. We formulate this statement as a separate lemma. It is a typical example of a uniqueness proof in algebra.

Lemma 4. If $f\colon X \to Y$, $g\colon Y \to X$, and $\tilde{g}\colon Y \to X$ are functions such that
$$g \circ f = \mathrm{id}_X \quad \text{and} \quad f \circ \tilde{g} = \mathrm{id}_Y,$$
then $g = \tilde{g}$.

Proof. Using Lemma 3 twice and associativity, Lemma 2, we obtain
$$g = g \circ \mathrm{id}_Y = g \circ (f \circ \tilde{g}) = (g \circ f) \circ \tilde{g} = \mathrm{id}_X \circ \tilde{g} = \tilde{g}.$$
■

Note also that if $f\colon X \to Y$ is an invertible function with inverse $f^{-1}\colon Y \to X$, we have by definition
$$f^{-1} \circ f = \mathrm{id}_X \quad \text{and} \quad f \circ f^{-1} = \mathrm{id}_Y,$$
that is,
$$\forall\, x \in X : f^{-1}(f(x)) = x \quad \text{and} \quad \forall\, y \in Y : f(f^{-1}(y)) = y.$$
Hence for all $x \in X$ and $y \in Y$, we have
$$f(x) = y \quad \Longleftrightarrow \quad f^{-1}(y) = x.$$

As an exercise, we will show the following lemma.

Lemma 5. If $f\colon X \to Y$ and $g\colon Y \to X$ are functions such that

$$g \circ f = \mathrm{id}_X,$$

then f is injective and g is surjective.

Theorem 6. A function $f\colon X \to Y$ is invertible iff it is bijective.

Proof. "⇒" For the inverse function $f^{-1}\colon Y \to X$, we have

$$f^{-1} \circ f = \mathrm{id}_X \quad \text{and} \quad f \circ f^{-1} = \mathrm{id}_Y\,.$$

By Lemma 5, it follows that f is injective and surjective.
"⇐" Since f is bijective, for every $y \in Y$, there exists exactly one $x \in X$ such that $f(x) = y$. Hence, we can define a function $g\colon Y \to X$ by

$$g(y) = x \iff f(x) = y.$$

By this definition, for all $x \in X$, we have $g(f(x)) = x$, that is, $g \circ f = \mathrm{id}_X$. Let $y \in Y$ and $x \in X$ such that $f(x) = y$. Then,

$$f(g(y)) = f(x) = y.$$

Hence, $f \circ g = \mathrm{id}_Y$ and g is the inverse of f. ∎

By Theorem 1, we know that the composition of bijective functions is bijective, hence we also know that *composition preserves invertibility*.

In particular, we know that the set of all permutations S_n is *closed under composition*, that is, for all $\pi, \sigma \in S_n$:

$$\pi \circ \sigma \in S_n.$$

Hence, composition defines a function

$$S_n \times S_n \to S_n, \quad (\pi, \sigma) \mapsto \pi \circ \sigma.$$

Such a function is called a **binary operation**. We know the following *three identities* for this operation:
It is **associative**, that is, for all $\pi, \sigma, \tau \in S_n$:

$$(\pi \circ \sigma) \circ \tau = \pi \circ (\sigma \circ \tau).$$

The *identity permutation* $\mathrm{id} \in S_n$ is the **identity element** for this operation, that is, for all $\pi \in S_n$:

$$\pi \circ \mathrm{id} = \mathrm{id} \circ \pi = \pi.$$

Every permutation $\pi \in S_n$ has an **inverse** $\pi^{-1} \in S_n$ satisfying

$$\pi \circ \pi^{-1} = \pi^{-1} \circ \pi = \mathrm{id}\,.$$

In abstract algebra, a set with a *binary operation* satisfying these three axioms (*associativity*, *identity*, and *invertibility*) is called a **group** and the set of permutations S_n are very important examples of finite groups.

Finally, we show that the inverse of the composite of two functions is the composition of the inverses in *reverse order*. This is true in general for inverses in algebra with the analogous proof.

Theorem 7. If $f\colon X \to Y$ and $g\colon Y \to Z$ are invertible functions, then $g \circ f\colon X \to Z$ is invertible and

$$(g \circ f)^{-1} = f^{-1} \circ g^{-1}.$$

Proof. Using associativity of composition, Lemma 2, Definition 13, and Lemma 3, we verify directly that

$$(f^{-1} \circ g^{-1}) \circ (g \circ f) = f^{-1} \circ (g^{-1} \circ g) \circ f = f^{-1} \circ \mathrm{id}_Y \circ f = f^{-1} \circ f = \mathrm{id}_X\,.$$

Analogously, one verifies that

$$(g \circ f) \circ (f^{-1} \circ g^{-1}) = \mathrm{id}_Z\,.$$

Hence, by definition, $g \circ f$ is invertible with inverse $f^{-1} \circ g^{-1}$. ∎

1.2 Equivalence relations and orders

Relations are used in mathematics and computer science to model a variety of different concepts. In this section, we discuss equivalence and order relations along with several important examples. To specify that two elements x, y of a given set X are in a certain relation to each other, we consider the pair

$$(x, y) \in X \times X$$

and we identify a relation with the set of all such pairs.

Definition 14. Let X be a set. A (binary) **relation** on X is a subset

$$R \subseteq X \times X.$$

Instead of $(x, y) \in R$, we write

$$xRy$$

and say that x **is in relation** R **to** y.

The equality relation satisfies the following properties for all x, y, z:

$$\begin{gathered} x = x, \\ x = y \implies y = x, \\ x = y \text{ and } y = z \implies x = z. \end{gathered}$$

If an arbitrary relation has all three properties (*reflexivity, symmetry, and transitivity*), it is called an equivalence relation.

Definition 15. A relation $\sim$ on a set X is

1. **reflexive** if for all $x \in X$,
$$x \sim x,$$
2. **symmetric** if for all $x, y \in X$,
$$x \sim y \implies y \sim x,$$
3. **transitive** if for all $x, y, z \in X$,
$$x \sim y \text{ and } y \sim z \implies x \sim z.$$

An **equivalence relation** is a relation that is reflexive, symmetric, and transitive.

Equivalence relations are almost everywhere in mathematics and appear in many applications. For example, if you compute with rational numbers or modulo an integer, you compute with equivalence classes, as we will see below.

Example.

1. The equality relation is given by the subset
$$\{(x, x) \in X \times X \mid x \in X\}.$$
To which function on X corresponds this subset?
2. If we consider the equivalence relation defined by the whole cartesian product $X \times X$, all elements of X are equivalent to each other, that is, $x \sim y$, for all $x, y \in X$.
3. For $a, b \in \mathbb{Z}$, we say that b **divides** a, if
$$a = qb$$
for some $q \in \mathbb{Z}$ and we write
$$b \mid a.$$
In this case, b is called a **divisor** of a and a is called a **multiple** b.
Divisibility on $\mathbb{Z}$ is a *reflexive* and *transitive* relation but it is *not symmetric*.
Moreover, one can immediately verify that, $b \mid 0$,
$$b \mid a \implies b \mid -a,$$
and, if an integer c divides two integers a and b, it divides their sum, that is,
$$c \mid a \text{ and } c \mid b \implies c \mid (a + b).$$

4. Given an $m \in \mathbb{N}$, we say that two integers $a, b \in \mathbb{Z}$ are **congruent modulo** m, if

$$m \mid (a - b)$$

and we write

$$a \equiv b \mod m.$$

We denote the corresponding **congruence relation** on $\mathbb{Z}$ by $\equiv_m$. We prove that it is an equivalence relation using the properties of divisibility above.

Reflexivity: Let $a \in \mathbb{Z}$. Since $m \mid (a - a) = 0$, we have $a \equiv_m a$.

Symmetry: Let $a, b \in \mathbb{Z}$ such that $a \equiv_m b$, that is, $m \mid (a - b)$. Then,

$$m \mid -(a - b) = b - a$$

and so $b \equiv_m a$.

Transitivity: Let $a, b, c \in \mathbb{Z}$ such that $a \equiv_m b$ and $b \equiv_m c$, that is, $m \mid (a - b)$ and $m \mid (b - c)$. Then,

$$m \mid ((a - b) + (b - c)) = a - c$$

and so $a \equiv_m c$.

Usually, one assumes $m \geq 2$, because $a \equiv b \mod 0$ iff $a = b$ and $a \equiv b \mod 1$ holds for all $a, b \in \mathbb{Z}$.

5. Let

$$Q = \mathbb{Z} \times \mathbb{Z} \setminus \{0\}$$

and define a relation $\sim$ on Q by *cross-multiplication*:

$$(a, b) \sim (c, d) \quad \text{if } ad = bc.$$

We leave it as an exercise to verify that $\sim$ is an equivalence relation. For verifying transitivity, one needs that one *can cancel a nonzero* integer, that is, for $a, b, c \in \mathbb{Z}$ and $a \neq 0$,

$$ab = ac \implies b = c.$$

Collecting all equivalent elements in one set, we obtain an *equivalence class*.

Definition 16. Let $\sim$ be an equivalence relation on a set X. If $x \in X$, we define the **equivalence class** of x by

$$[x]_\sim = \{y \in X \mid y \sim x\} \subseteq X.$$

If the equivalence relation is fixed, we also write $[x]$ instead of $[x]_\sim$.
Each element $y \in [x]$ is called a **representative** of the equivalence class $[x]$.

Note that, by reflexivity, $x \sim x$, and hence

$$x \in [x].$$

So an equivalence class is *never the empty set* and x is a representative of $[x]$.

We discuss the equivalence classes arising from the equivalence relations above.

Example.

1. For the equality relation, the equivalence class of an $x \in X$ consists of the **singleton** $\{x\}$, that is,
$$[x] = \{x\}.$$

2. For $\sim = X \times X$, we have for all $x \in X$,
$$[x] = X.$$

3. By definition of the congruence relation, we have
$$b \equiv_m a \iff b = a + qm \quad \text{for some } q \in \mathbb{Z}.$$
Hence the equivalence class (or **congruence class**) of an $a \in \mathbb{Z}$ is given by
$$[a] = \{b \in \mathbb{Z} \mid b = a + qm \text{ with } q \in \mathbb{Z}\}.$$
For example, with the congruence $\equiv_3$ modulo 3, we have
$$\begin{aligned}[0] &= \{\dots, -3, 0, 3, 6, \dots\} \\ [1] &= \{\dots, -2, 1, 4, 7, \dots\} \\ [2] &= \{\dots, -1, 2, 5, 8, \dots\}\end{aligned}$$
and
$$[0] = [3], \quad [1] = [4], \quad [2] = [5], \dots.$$
For $m \geq 2$, one usually considers the representatives
$$[0], [1], \dots, [m-1]$$
for the m different equivalence classes.

4. For $(a, b) \in Q$ with Q from above, the equivalence class under cross-multiplication is
$$[(a, b)] = \{(c, d) \mid ad = bc\}.$$
If we write
$$a/b = [(a, b)],$$
this equivalence class is precisely the fraction usually denoted by $\frac{a}{b}$. For example, the pairs
$$(1, 2) \neq (2, 4),$$
but
$$[(1, 2)] = [(2, 4)],$$
that is, $\frac{1}{2} = \frac{2}{4}$.

The following theorem allows us to replace an equivalence relation by an equality sign, if we replace elements by equivalence classes.

Theorem 8. Let $\sim$ be an equivalence relation on a set X and $x, y \in X$. Then,

$$x \sim y \iff [x] = [y] \iff [x] \cap [y] \neq \emptyset.$$

Proof. We show the implications $A \Rightarrow B \Rightarrow C \Rightarrow A$ for the three statements.
"$A \Rightarrow B$" Assume $x \sim y$. Let $z \in [x]$. By definition, $z \sim x$ and so transitivity gives $z \sim y$, that is, $z \in [y]$. Hence

$$[x] \subseteq [y].$$

By symmetry, we also have $y \sim x$, which gives the reverse inclusion $[y] \subseteq [x]$. Thus,

$$[x] = [y].$$

"$B \Rightarrow C$" follows immediately, since $x \in [x] = [y]$.
"$C \Rightarrow A$" Let $z \in [x] \cap [y]$. Hence, $z \sim x$ and $z \sim y$. By symmetry, $x \sim z$ and, by transitivity, it follows that

$$x \sim y.$$

Hence all three statements are equivalent. ∎

By the contrapositive, the second equivalence can also be stated as

$$[x] \neq [y] \iff [x] \cap [y] = \emptyset.$$

Recall that for every $x \in X$, we also have $x \in [x]$. Hence equivalence classes provide a division of X into *nonempty subsets*, any two of which are disjoint, that is, they are **pairwise disjoint**. This is called a **partition** of X.

Example. For $\equiv_2$, the two equivalence classes

$$[0] = \{\ldots, -2, 0, 2, \ldots\} \quad \text{and} \quad [1] = \{\ldots, -1, 1, 3, \ldots\}$$

are the sets of even and odd numbers that form a partition of the integers $\mathbb{Z}$.

Given an equivalence relation, we can also consider the set of all equivalence classes and the function that maps each element to its equivalence class.

Definition 17. Let $\sim$ be an equivalence relation on a set X. The set of all equivalence classes, denoted by

$$X/\!\sim = \{[x] \mid x \in X\},$$

is called the **quotient set of X by $\sim$**. The function

$$\pi \colon X \to X/\!\sim, \quad x \mapsto [x]$$

is called the **canonical** or **natural map** from X to $X/\!\sim$.

Formally, the quotient set

$$X/\!\sim \subseteq \mathcal{P}(X)$$

is a subset of the powerset of X. However, we usually think of *equivalence classes as objects* themselves. Clearly, the *canonical map* is always *surjective*. It is *not injective*, unless the relation on X is equality.

Example.

1. For the congruence relation $\equiv_m$, we denote the quotient set by

$$\mathbb{Z}_m = \mathbb{Z}/\equiv_m.$$

   For $m \geq 1$, it consists of m different equivalence classes

$$\mathbb{Z}_m = \{[0], [1], \ldots, [m-1]\}$$

   and the canonical map

$$\pi \colon \mathbb{Z} \to \mathbb{Z}_m, \quad \pi(a) = [a]$$

   maps every integer to its corresponding equivalence class modulo m.

2. For the equivalence relation $\sim$ defined by cross-multiplication, we can identify the quotient set $Q/\sim$ with the set of rational numbers $\mathbb{Q}$.

In the following, we discuss how to define *functions on quotient sets* by considering functions that have the same value on all elements of any equivalence class.

Definition 18. Let $\sim$ be an equivalence relation on a set X. A function $f \colon X \to Y$ is **invariant under** $\sim$ if for all $x, y \in X$,

$$x \sim y \implies f(x) = f(y).$$

Using an invariant function, we can assign a unique element to each equivalence class by applying it to any of its representatives. More formally, we have the following statement.

Theorem 9. Let $\sim$ be an equivalence relation on a set X. If a function $f \colon X \to Y$ is invariant under $\sim$, then

$$\bar{f} = \{([x], f(x)) \mid x \in X\} \subseteq X/\sim \times Y$$

is a function $\bar{f} \colon X/\sim \to Y$ with

$$\bar{f}([x]) = f(x).$$

Proof. By definition, we have to show that for each equivalence class $C \in X/\sim$, there exists a unique element $y \in Y$ such that $(C, y) \in \bar{f}$.
For proving existence, let $C \in X/\sim$. Then, $C = [x]$ for some $x \in X$. For $y = f(x) \in Y$,

$$(C, y) = ([x], f(x)) \in \bar{f}.$$

For proving uniqueness, let $(C, y_1), (C, y_2) \in \bar{f}$. Let $x_1, x_2 \in X$ such that

$$(C, y_1) = ([x_1], f(x_1)) \quad \text{and} \quad (C, y_2) = ([x_2], f(x_2)).$$

Since $C = [x_1] = [x_2]$, we have $x_1 \sim x_2$. Since f is invariant under $\sim$, it follows that

$$y_1 = f(x_1) = f(x_2) = y_2.$$

■

In this case, one also says that the function

$$\bar{f}\colon X/\!\sim \to Y, \quad \bar{f}([x]) = f(x)$$

is **well-defined**, or that the invariant function $f\colon X \to Y$ **induces** by

$$\bar{f}([x]) = f(x)$$

a well-defined function $\bar{f}\colon X/\!\sim \to Y$. In summary, for defining functions on a quotient set $X/\!\sim$, we just need functions $f\colon X \to Y$ that are invariant under $\sim$.

Example.

1. For the congruence relation $\equiv_m$, the remainder of the division by m is an invariant function. In more detail, we know from *division with remainder* that given integers a and d with $d \neq 0$, there exist unique integers q and r such that

   $$a = qd + r \quad \text{and} \quad 0 \leq r < |d|.$$

   In this case, one calls, respectively, a the **dividend**, d the **divisor**, q the **quotient**, and r the **remainder**. We write

   $$\operatorname{quo}(a, d) = q \quad \text{and} \quad \operatorname{rem}(a, d) = r$$

   for the quotient and remainder of the division of a by d. Note that these depend on the sign of a and d. For example, $7 = 2 \cdot 3 + 1$ and $-7 = -3 \cdot 3 + 2$.

   Let $m \in \mathbb{N}$ with $m \geq 1$. We show that

   $$f\colon \mathbb{Z} \to \mathbb{Z}, \quad x \mapsto \operatorname{rem}(x, m),$$

   is invariant under $\equiv_m$. Let $a, b \in \mathbb{Z}$ such that $a \equiv_m b$, that is, $m \mid (a - b)$. Let

   $$a = qm + r \quad \text{and} \quad b = q'm + r' \quad \text{with } 0 \leq r < m, \quad 0 \leq r' < m.$$

   We have to show that $r = \operatorname{rem}(a, m) = \operatorname{rem}(b, m) = r'$. We have

   $$a - b = (q - q')m + r - r' \quad \text{and} \quad |r - r'| < m.$$

   Hence, in particular, $m \mid (r - r')$. Assuming that $r - r' \neq 0$ this implies $|r - r'| \geq m$, contradicting $|r - r'| < m$. Hence $r - r' = 0$, that is, $r = r'$.

   In other words, $a \equiv_m b$ implies $\operatorname{rem}(a, m) = \operatorname{rem}(b, m)$. Conversely, if $\operatorname{rem}(a, m) = \operatorname{rem}(b, m)$, then $a - b = (q - q')m$ and $a \equiv_m b$. Hence, we have the equivalences

   $$a \equiv_m b \iff [a] = [b] \iff \operatorname{rem}(a, m) = \operatorname{rem}(b, m).$$

   Therefore, we can test whether two equivalence classes are equal by computing the remainder of the division of representatives by m.

2. As an exercise, we show that every function $f\colon X \to Y$ gives rise to an equivalence relation $\sim_f$ on X by defining

   $$x \sim_f y \iff f(x) = f(y).$$

   By definition, $\sim_f$ is invariant under f and one can verify that the induced function $\bar{f}\colon X/\!\sim_f \to Y$ with $\bar{f}([x]) = f(x)$ is injective.

Addition and multiplication in $\mathbb{Z}$ is invariant under congruences in the following sense.

Lemma 10. Let $m \in \mathbb{N}$. If $a, b, a', b' \in \mathbb{Z}$ such that $a \equiv_m a'$ and $b \equiv_m b'$, then

$$a + b \equiv_m a' + b' \quad \text{and} \quad ab \equiv_m a'b'.$$

Proof. Let $a, b, a', b' \in \mathbb{Z}$ such that $a \equiv_m a'$ and $b \equiv_m b'$, that is,

$$m \mid (a - a') \quad \text{and} \quad m \mid (b - b').$$

Then m divides the sum of the differences

$$m \mid (a - a' + b - b') = (a + b) - (a' + b').$$

Hence, $a + b \equiv_m a' + b'$. For the product, we have to show that $m \mid (ab - a'b')$, this follows from the identity

$$ab - a'b' = (ab - ab') + (ab' - a'b') = a(b - b') + (a - a')b'.$$

Hence, $ab \equiv_m a'b'$. ∎

By Theorem 9, the two binary operations of addition and multiplication on the integers

$$+ \colon \mathbb{Z} \times \mathbb{Z} \to \mathbb{Z}, \quad (a, b) \mapsto a + b \quad \text{and} \quad \cdot \colon \mathbb{Z} \times \mathbb{Z} \to \mathbb{Z}, \quad (a, b) \mapsto a \cdot b$$

induce a well-defined addition and multiplication on the quotient set by adding and multiplying representatives

$$+ \colon \mathbb{Z}_m, \times \mathbb{Z}_m \to \mathbb{Z}_m, \quad ([a], [b]) \mapsto [a + b] \quad \text{and} \quad \cdot \colon \mathbb{Z}_m, \times \mathbb{Z}_m \to \mathbb{Z}_m, \quad ([a], [b]) \mapsto [a \cdot b].$$

Note that we use the same symbols for both operations in $\mathbb{Z}$ and $\mathbb{Z}_m$.

Example.

1. In $\mathbb{Z}_5$,
$$[2] + [3] = [2 + 3] = [5] = [0], \qquad [2] \cdot [2] = [2 \cdot 2] = [4].$$
2. In $\mathbb{Z}_4$,
$$[2] + [3] = [2 + 3] = [5] = [1], \qquad [2] \cdot [2] = [2 \cdot 2] = [4] = [0].$$

Recall that addition and multiplication on $\mathbb{Z}$ satisfy the following properties. *Addition* is **associative**, **commutative**, 0 is the **additive identity**, and there exists an **additive inverse** for every integer, that is, for all $a, b, c \in \mathbb{Z}$,

1. $(a + b) + c = a + (b + c)$,
2. $a + b = b + a$,
3. $a + 0 = a$,
4. $a + (-a) = 0$.

In other words, $\mathbb{Z}$ under plus is a commutative (or **abelian**) group. *Multiplication* is also *associative* and *commutative* and 1 is the **multiplicative identity**, that is, for all $a, b, c \in \mathbb{Z}$,

5. $(ab)c = a(bc)$,
6. $ab = ba$,
7. $1a = a1 = a$.

Finally, multiplication is **distributive** with respect to addition, that is, for all $a, b, c \in \mathbb{Z}$,

8. $a(b + c) = ab + ac$ and $(b + c)a = ba + ca$.

In abstract algebra, a set R with *two binary operations*

$$+\colon R \times R \to R \quad \text{and} \quad \cdot\colon R \times R \to R$$

and elements $0, 1 \in R$ satisfying (1.)–(8.) for all $a, b, c \in R$ is called a **commutative ring**. Thus the set $\mathbb{Z}$ of all integers is a commutative ring.

Note that in (6.) *multiplication* is required to be *commutative*. If we on do not require (6.), one calls R a **ring**. We will see in the next section that *square matrices* of a fixed size form a ring that is *noncommutative*.

The *additive identity* 0 in a ring R is called **zero** or **zero element** of R and the *multiplicative identity* 1 is called **one** or **unit element** of R. Note that in any ring, there can only be one *unit element*, for if $1'$ is a is another one, then we have

$$1' = 1' \cdot 1 = 1.$$

Similarly, the *zero element* is uniquely determined. As for inverse functions (Lemma 4), one sees that the additive inverse $-a$ of an element $a \in R$ is unique.

Example.

1. $\mathbb{Q}$ and $\mathbb{R}$ are commutative rings.
2. $\mathbb{N}$ is not a commutative ring. One property does not hold, which one?
3. We can add and multiply *real-valued functions* just by adding and multiplying their values in $\mathbb{R}$. By these **pointwise operations**, we obtain new rings, where the elements are functions.

   For example, the set $\mathbb{R}^{\mathbb{R}}$ of all real functions $f\colon \mathbb{R} \to \mathbb{R}$ forms a commutative ring with addition and multiplication of $f, g \in \mathbb{R}^{\mathbb{R}}$ defined by

   $$(f + g)(x) = f(x) + g(x) \quad \text{and} \quad (f \cdot g)(x) = f(x) \cdot g(x).$$

   The zero element is the zero function $\mathbf{0}\colon x \mapsto 0$ and the unit element is the constant function $\mathbf{1}\colon x \mapsto 1$. The additive inverse is given by $(-f)(x) = -f(x)$.

The properties (1.)–(8.) can be shown by using the corresponding identities in $\mathbb{R}$. We verify, for example, associativity (1.). For $x \in \mathbb{R}$, we have

$$\begin{aligned}((f+g)+h)(x) &= (f+g)(x)+h(x) = (f(x)+g(x))+h(x)\\ &= f(x)+(g(x)+h(x)) = f(x)+(g+h)(x) = (f+(g+h))(x),\end{aligned}$$

and hence $(f+g)+h = f+(g+h)$. Similarly, for showing (7.), we have for $x \in \mathbb{R}$,

$$(\mathbf{1}\cdot f)(x) = \mathbf{1}(x)\cdot f(x) = 1\cdot f(x) = f(x),$$

and hence $\mathbf{1}\cdot f = f$.

Note that in this construction, we only use that the functions map to the real numbers. Hence, for any nonempty set X, the set $\mathbb{R}^X$ of functions $f\colon X \to \mathbb{R}$ with pointwise operations forms a commutative ring.

In particular, we can *add and multiply sequences* in $\mathbb{R}$ *pointwise*, which reads in this case as

$$(a_n)_{n\in\mathbb{N}} + (b_n)_{n\in\mathbb{N}} = (a_n+b_n)_{n\in\mathbb{N}} \quad \text{and} \quad (a_n)_{n\in\mathbb{N}}\cdot(b_n)_{n\in\mathbb{N}} = (a_n\cdot b_n)_{n\in\mathbb{N}}.$$

Since addition and multiplication in $\mathbb{Z}_m$ is defined by adding and multiplying representatives, all properties (1.)–(8.) carry over from $\mathbb{Z}$ to $\mathbb{Z}_m$. So we obtain a whole family of finite commutative rings.

Theorem 11. For $m \in \mathbb{N}$, the set $\mathbb{Z}_m$ with addition and multiplication for $[a],[b] \in \mathbb{Z}_m$ defined by

$$[a]+[b] = [a+b] \quad \text{and} \quad [a]\cdot[b] = [a\cdot b]$$

is a commutative ring with the zero element $[0]$, the unit element $[1]$, and the additive inverse given by $-[a] = [-a]$.

Proof. We prove only (1.), (4.), and (7.), leaving the others as exercises.

1.: We have

$$([a]+[b])+[c] = [a+b]+[c] = [(a+b)+c]$$

and

$$[a]+([b]+[c]) = [a]+[b+c] = [a+(b+c)].$$

Both equivalence classes are equal, since associativity (1.) holds in $\mathbb{Z}$.

4.: Since $-a$ is the additive inverse of a in $\mathbb{Z}$ (4.), we have

$$[a]+(-[a]) = [a]+[-a] = [a+(-a)] = [0].$$

8.: From distributivity in $\mathbb{Z}$ (8.), it follows that

$$[a]([b]+[c]) = [a][b+c] = [a(b+c)] = [ab+ac] = [ab]+[ac] = [a][b]+[a][c].$$

■

Example.

1. The *addition* and *multiplication tables* for $\mathbb{Z}_2$ are:

$+$	$[0]$	$[1]$
$[0]$	$[0]$	$[1]$
$[1]$	$[1]$	$[0]$

$\cdot$	$[0]$	$[1]$
$[0]$	$[0]$	$[0]$
$[1]$	$[0]$	$[1]$

2. For $\mathbb{Z}_3$, they are:

$+$	$[0]$	$[1]$	$[2]$
$[0]$	$[0]$	$[1]$	$[2]$
$[1]$	$[1]$	$[2]$	$[0]$
$[2]$	$[2]$	$[0]$	$[1]$

$\cdot$	$[0]$	$[1]$	$[2]$
$[0]$	$[0]$	$[0]$	$[0]$
$[1]$	$[0]$	$[1]$	$[2]$
$[2]$	$[0]$	$[2]$	$[1]$

3. For $\mathbb{Z}_4$, they are:

$+$	$[0]$	$[1]$	$[2]$	$[3]$
$[0]$	$[0]$	$[1]$	$[2]$	$[3]$
$[1]$	$[1]$	$[2]$	$[3]$	$[0]$
$[2]$	$[2]$	$[3]$	$[0]$	$[1]$
$[3]$	$[3]$	$[0]$	$[1]$	$[2]$

$\cdot$	$[0]$	$[1]$	$[2]$	$[3]$
$[0]$	$[0]$	$[0]$	$[0]$	$[0]$
$[1]$	$[0]$	$[1]$	$[2]$	$[3]$
$[2]$	$[0]$	$[2]$	$[0]$	$[2]$
$[3]$	$[0]$	$[3]$	$[2]$	$[1]$

4. For completeness, we also note that, $\mathbb{Z}_0 = \mathbb{Z}$ and $\mathbb{Z}_1$ is the ring consisting of a single element.

We discuss some notions and additional properties of an arbitrary ring R. For $a, b \in R$, the **subtraction** is defined by

$$a - b = a + (-b).$$

The familiar identities for multiplication by the zero element $0 \in R$ and the negative of the unit element $-1 \in R$, follow from the defining identities above. In particular, we have for all $a \in R$,

$$0a = a0 = 0. \tag{1.1}$$

Because, by (3.) and (8.),

$$0a = (0 + 0)a = 0a + 0a.$$

Hence by subtraction $0a = 0$. Similarly, it follows that $a0 = 0$. Moreover, for all $a \in R$,

$$(-1)a = a(-1) = -a. \tag{1.2}$$

Because, by (1.1), (3.), and (8.),

$$0 = 0a = (1 + (-1))a = a + (-1)a.$$

Hence $(-1)a$ is the unique additive inverse of a, that is, $(-1)a = -a$. Similarly, it follows that $a(-1) = -a$.

Definition 19. Let R be a ring. An element $a \in R$ is called **invertible** or a **unit** if there exists an element $b \in R$ such that

$$ba = ab = 1.$$

Such an element b is uniquely determined. If a is invertible, we call b the **(multiplicative) inverse** of a and write

$$b = a^{-1}.$$

Note the analogy with Definition 13 of an invertible function. As in Lemma 4, one sees that the multiplicative *inverse* is *uniquely determined*. As in Theorem 7, one sees that the product of invertible elements $a, b \in R$ is invertible and

$$(ab)^{-1} = b^{-1}a^{-1}.$$

Example.

1. In any ring, by (7.) and (1.2),

   $$1 \cdot 1 = 1 \quad \text{and} \quad (-1) \cdot (-1) = -(-1) = 1,$$

   hence 1 and -1 are always invertible.

2. In $\mathbb{Z}$, apart from 1 and -1, no integer has a multiplicative inverse.

3. In $\mathbb{Z}_3$, $[2] \cdot [2] = [1]$, hence $[2]^{-1} = [2]$.

4. In $\mathbb{Z}_4$, $[3] \cdot [3] = [1]$, hence $[3]^{-1} = [3]$. However, using the multiplication table, we can check that $[2]$ is not invertible.

5. In $\mathbb{Z}_5$,

   $$[2] \cdot [3] = [1] \quad \text{and} \quad [4] \cdot [4] = [1],$$

   hence $[2]^{-1} = [3]$, $[3]^{-1} = [2]$, and $[4]^{-1} = [4]$. So, as in $\mathbb{Z}_2$ and $\mathbb{Z}_3$, every nonzero element is invertible.

6. In $\mathbb{Q}$ and $\mathbb{R}$, every $a \neq 0$ has an inverse a^{-1}.

7. If in a ring R the zero element $0 \in R$ is invertible, then, by (1.1),

   $$1 = 0 \cdot 0^{-1} = 0.$$

   However, this can only be the case when R consists of a single element, that is, $R = \{0\}$ is the **zero ring**. Assume that $1 = 0$, then for any $a \in R$,

   $$a = 1a = 0a = 0.$$

   Incidentally, this explains also why dividing by 0 is forbidden. Often, the zero ring is excluded from considerations as a degenerate case by assuming $1 \neq 0$.

In view of the last example, the most we can ask for is that every nonzero element is invertible.

Definition 20. A **field** F is a commutative ring with $1 \neq 0$ in which every nonzero element is invertible, that is, for all $a \in F \setminus \{0\}$, there exists $a^{-1} \in F$ with $a^{-1}a = 1$.

For a, b in a field with $b \neq 0$, the **division** is defined by

$$a/b = ab^{-1}.$$

Example.

1. $\mathbb{Q}$ and $\mathbb{R}$ are infinite fields.
2. $\mathbb{Z}_2$ is the smallest field.
3. $\mathbb{Z}_3$, $\mathbb{Z}_5$ are other small finite fields.

We show that $\mathbb{Z}$ modulo a prime number is always a finite field.

Theorem 12. Let $m \in \mathbb{N}$ with $m \geq 2$. The commutative ring $\mathbb{Z}_m$ is a field iff m is a prime number.

Before proving the theorem, we note that in any field F the product of any two nonzero elements is nonzero or, equivalently, for all $a, b \in F$,

$$ab = 0 \implies a = 0 \text{ or } b = 0. \tag{1.3}$$

Assume that $ab = 0$ and $a \neq 0$. Multiplying both sides by a^{-1} gives $a^{-1}ab = a^{-1}0 = 0$ and hence $1b = b = 0$.

A commutative ring R with $1 \neq 0$ satisfying (1.3) for all $a, b \in R$ is a called an **integral domain**. We note that in an integral domain R one *can cancel a nonzero* element, that is, for $a, b, c \in R$ and $a \neq 0$,

$$ab = ac \implies b = c.$$

If $ab = ac$, then $a(b-c) = 0$. Since $a \neq 0$ and R is an integral domain, we have $b - c = 0$, that is, $b = c$.

Example.

1. Every field is an integral domain.
2. $\mathbb{Z}$ is an integral domain that is not a field.
3. If $m \geq 2$ is not a prime, then $\mathbb{Z}_m$ is not an integral domain.

   In this case, m is a **composite number**, that is, $m = ab$ with $m > a, b \geq 2$. Hence, in $\mathbb{Z}_m$, we have $[a] \neq 0$ and $[b] \neq 0$, but

   $$[a][b] = [ab] = [m] = [0].$$

   For example, $6 = 2 \cdot 3$ and in $\mathbb{Z}_6$, we have $[2][3] = [6] = [0]$.

In the proof of Theorem 12, we also use *Euclid's lemma*, which states that if a prime p divides a product ab of two integers it must divide a or b, that is,

$$p \mid ab \implies p \mid a \text{ or } p \mid b.$$

Proof. (Theorem 12) "⇒" We prove the contrapositive. From the last example above, we know that if m is not a prime, $\mathbb{Z}_m$ is not an integral domain, in particular, it cannot be a field.

"⇐" Let m be a prime. In $\mathbb{Z}_m$, clearly $[1] \neq [0]$. We have to show that every nonzero $[a] \in \mathbb{Z}_m$ is invertible. First, we show that $\mathbb{Z}_m$ is an integral domain. Assume that $[a], [b] \in \mathbb{Z}_m$ such that $[a][b] = [ab] = [0]$. Hence

$$m \mid (ab - 0) = ab.$$

By Euclid's lemma, we have $m \mid a$ or $m \mid b$, that is, $[a] = [0]$ or $[b] = [0]$. Hence $\mathbb{Z}_m$ is an integral domain.

Let now $[a] \in \mathbb{Z}_m$ and $[a] \neq [0]$. We consider the function

$$f \colon \mathbb{Z}_m \to \mathbb{Z}_m, \quad [b] \mapsto [a][b]$$

and show that it is injective. Let $[b_1], [b_2] \in \mathbb{Z}_m$ such that

$$f([b_1]) = [a][b_1] = [a][b_2] = f([b_2]).$$

Since $\mathbb{Z}_m$ is an integral domain and $[a] \neq 0$, we can cancel $[a]$, which yields $[b_1] = [b_2]$. Hence f is injective. Since $\mathbb{Z}_m$ is finite, this implies that f is also surjective. In particular, there exists a $[b] \in \mathbb{Z}_m$ such that

$$f([b]) = [a][b] = [1],$$

that is, $[a]$ is invertible and $[b] = [a]^{-1}$. ∎

For a prime p, we know by Theorem 12 that $\mathbb{Z}_p$ is a field. In particular, we know that for every nonzero $[a] \in \mathbb{Z}_p$, there exists an inverse $[a]^{-1} \in \mathbb{Z}_p$. For small p, we can easily compute the full multiplication table. For a given $[a] \in \mathbb{Z}_p$, the inverse can be efficiently computed with the *extended Euclidean algorithm*, which is based on successive division with remainder.

We also outline the construction of the field of rational numbers based on the equivalence relation defined by cross-multiplication, where

$$Q = \mathbb{Z} \times \mathbb{Z} \setminus \{0\}, \quad \text{and} \quad (a, b) \sim (c, d) \quad \text{if } ad = bc.$$

Moreover, we denote the equivalence classes $[(a, b)]$ by a/b and the quotient set by

$$\mathbf{Q} = Q/\!\sim.$$

One can verify that the binary operations on Q corresponding to the usual addition and multiplication of rational numbers are invariant under $\sim$. Hence the operations

$$a/b + c/d = (ad + cb)/(bd) \quad \text{and} \quad a/b \cdot c/d = (ac)/(bd) \tag{1.4}$$

on $\mathbf{Q}$ are well-defined. Note that here we also need that the product of nonzero integers is nonzero which holds since $\mathbb{Z}$ is an integral domain. With this addition and multiplication $\mathbf{Q}$ is a field that can be identified with the rational numbers $\mathbb{Q}$.

Theorem 13. The set $\mathbf{Q}$ with addition and multiplication defined by (1.4) is a field, with the zero element $0/1$, the unit element $1/1$. For $a/b \in \mathbf{Q}$, the additive inverse is given by $-a/b = (-a)/b$ and if a/b is nonzero, the multiplicative inverse is given by $(a/b)^{-1} = b/a$.

Proof. One can verify the properties (1.)–(8.) of a commutative ring based on the corresponding identities in $\mathbb{Z}$. For example, we verify the existence of an additive inverse.

4.: We have

$$a/b + (-a)/b = (ab + (-a)b)/b^2 = 0/b^2 = 0/1.$$

Clearly, $1/1 \neq 0/1$ and it remains to prove the existence of an inverse for every nonzero element. Let $a/b \in \mathbf{Q}$ and $a/b \neq 0/1$, that is, $a = a1 \neq 0b = 0$. Therefore, $b/a \in \mathbf{Q}$ and

$$b/a \cdot a/b = (ba)/(ab) = 1/1,$$

and hence $(a/b)^{-1} = b/a$ ∎

For the rest of the section, we discuss order relations on a set, which enable us to compare different elements. Note that the standard order $\leq$ on $\mathbb{Z}$ satisfies the following properties for all $a, b, c \in \mathbb{Z}$,

$$\begin{aligned} a &\leq a, \\ a \leq b \text{ and } b \leq a &\implies a = b, \\ a \leq b \text{ and } b \leq c &\implies a \leq c. \end{aligned}$$

Moreover, any two integers $a, b \in \mathbb{Z}$ are *comparable*, that is,

$$a \leq b \text{ or } b \leq a.$$

Definition 21. A relation $\leq$ on a set X is

1. **antisymmetric** if for all $x, y \in X$,

$$x \leq y \text{ and } y \leq x \implies x = y$$

2. **total** or **linear** if for all $x, y \in X$,

$$x \leq y \text{ or } y \leq x.$$

A **partial order** is a relation that is reflexive, antisymmetric, and transitive. A set with a partial order is called a **partially ordered set** or **poset**. For a partial order $\leq$, we write $x < y$ to mean $x \leq y$ but $x \neq y$.

A **total order** or **linear order** is a partial order that is total. A set with a total order is called a **totally ordered set**.

Example.

1. $\leq$ on $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$, or $\mathbb{R}$ is a total order.
2. $\subseteq$ on the power set $\mathcal{P}(X)$ of a set X is a partial order that is not total if X contains more than one element. If $x, y \in X$ are distinct elements, then we have neither $\{x\} \subseteq \{y\}$ nor $\{y\} \subseteq \{x\}$.
3. Divisibility $\mid$ on $\mathbb{N}$ is a partial order that is not total. For example, $2 \nmid 3$ and $3 \nmid 2$.
4. On $\mathbb{R} \times \mathbb{R}$, we can define a partial order by comparing vectors componentwise,

$$(x_1, y_1) \leq (x_2, y_2), \text{ if } x_1 \leq x_2 \text{ and } y_1 \leq y_2.$$

   For example, $(1, 2) \leq (2, 3)$ but we have neither $(1, 2) \leq (2, 1)$ nor $(2, 1) \leq (1, 2)$.

Definition 22. In a poset X, an element $g \in X$ is a **greatest element** if for all $x \in X$,

$$x \leq g,$$

and an element $m \in X$ is a **maximal element** if for all $x \in X$,

$$m \leq x \implies m = x.$$

In words, a greatest element surpasses all other elements, and a maximal element is surpassed by none. In a *totally ordered set*, these two concepts are *equivalent*, which explains why often they are not distinguished. However, in a poset these notions are distinct: a greatest element is always maximal, but not conversely. A greatest element, if it exists, is unique. If g_1 and g_2 are greatest elements, then $g_2 \leq g_1$ and $g_1 \leq g_2$, hence, by antisymmetry, $g_1 = g_2$.

Example.

1. In $P = \{\emptyset, \{x\}, \{y\}, \{x, y\}\}$ with $\subseteq$, the greatest element is $\{x, y\}$.
2. In $P = \{\emptyset, \{x\}, \{y\}\}$ with $\subseteq$, there are two maximal elements $\{x\}$ and $\{y\}$ but there is no greatest element.
3. In $\mathbb{Z}$ with $\leq$, there is no greatest element.

Least and minimal elements are defined analogously. A $u \in X$ is a **least element** if $u \leq x$ for all $x \in X$ and $v \in X$ is a **minimal element** if $x \leq v$ implies $x = v$.

Example. In $\mathbb{N}$ with $\leq$, zero is the least element but there is no greatest element.

In a poset, if $x \leq y$, we write

$$\min(x, y) = x \quad \text{and} \quad \max(x, y) = y.$$

In a partially ordered set, we can also formalize when a subset is *bounded* and we introduce the corresponding terminology. In this lecture, we will use these concepts mainly for subsets of $\mathbb{R}$ with the standard order $\leq$.

Definition 23. Let X be a poset with $\leq$ and $A \subseteq X$ a subset. An element $b \in X$ is an **upper bound for** A (or A **is bounded from above by** b) if

$$\forall\, x \in A : x \leq b.$$

We say that A is **bounded from above** if there exists an upper bound for A. A **lower bound** and **bounded from below** is defined analogously.
We say that A is **bounded** if there exists an upper and a lower bound for A.

Example. We consider the following subsets of $\mathbb{R}$.

1. $(-\infty, 0)$ has upper bound 0 but no lower bound.
2. $[0, 1]$ is bounded from above by 1 and is bounded from below by 0.
3. $[0, 1)$ is bounded from above by 1 and is bounded from below by 0.

Note that in these examples, we always consider the *least upper bound*.

Definition 24. Let X be a poset with $\leq$ and $A \subseteq X$ a subset. An upper bound b of A is called **supremum** (or **least upper bound**) of A if for all upper bounds b' of A,

$$b \leq b'.$$

The **infimum** (or **greatest lower bound**) of A is defined analogously.
If the supremum/infimum of A exists, it is uniquely determined, and we denote it by

$$\sup A \quad \text{and} \quad \inf A,$$

respectively.

Example.

1. $\sup [0, 1] = 1$ and $\inf [0, 1] = 0$.
2. $\sup (0, 1) = 1$ and $\inf (0, 1) = 0$.
3. For $A = \{1/n \mid n \in \mathbb{N}, n \geq 1\}$, we have $\sup A = 1$ and $\inf A = 0$.

Note that the supremum of a set A, assuming it exists, does not necessarily belong to it. If it does, it is a greatest element of A. Conversely, if A has a greatest element, it is the supremum of A.

We also note that the supremum of a set may fail to exist even if the set is bounded from above. Consider

$$A = \{x \in \mathbb{Q} \mid x^2 \leq 2\}.$$

Clearly, A is bounded from above by, for example, 2. If we consider A as a subset of $\mathbb{R}$, its supremum is $\sqrt{2}$. However, if we consider A as a subset of $\mathbb{Q}$, it has no supremum,

since $\sqrt{2} \notin \mathbb{Q}$, but $\sqrt{2}$ can be approximated arbitrarily well from below by a rational number.
In this example, the problem is that $\sqrt{2}$ is one of the "gaps" in $\mathbb{Q}$. The fundamental difference between $\mathbb{Q}$ and $\mathbb{R}$ is that the real numbers have no such "missing points". This can be formalized by stating that the real numbers satisfy the

Completeness axiom. Every nonempty subset A of $\mathbb{R}$ that is bounded from above has a least upper bound, that is, there exists $b \in \mathbb{R}$ such that $b = \sup A$.

The completeness axiom is crucial in analysis for proving statements about the existence of real numbers with certain properties. As a first application, we show that the natural numbers $\mathbb{N}$ are not bounded from above in $\mathbb{R}$. This is called the **Archimedean property** of $\mathbb{R}$. Although this property seems obvious at first, its proof depends on the completeness axiom.

Theorem 14. For every $x \in \mathbb{R}$, there exists an $n \in \mathbb{N}$ such that $n > x$.

Proof. Assume, for contradiction, that $\mathbb{N}$ is bounded from above in $\mathbb{R}$. By the completeness axiom, it would have a least upper bound, say $b = \sup \mathbb{N}$. Since b is the least upper bound, $b - 1$ is not an upper bound of $\mathbb{N}$. Hence there exists an $n \in \mathbb{N}$ such that $n > b - 1$. But then

$$n + 1 > b.$$

Since $n + 1 \in \mathbb{N}$, this contradicts b being an upper bound of $\mathbb{N}$. ∎

As an exercise, we show that the real sequence $(1/n)_{n \geq 1}$ gets arbitrarily close to 0 in the following sense.

Corollary 15. For every $\varepsilon > 0$, there exists an $n \in \mathbb{N}$ such that $1/n < \varepsilon$.

Using the Archimedean property $\mathbb{R}$, one can also prove that between any two real numbers there is a rational number. One says that $\mathbb{Q}$ is *dense* in $\mathbb{R}$.

Finally, we prove the familiar properties for working with inequalities. For example, multiplication by positive/negative numbers preserves/reverses inequalities or no square is negative. In fact, these properties hold in any *ordered field*, which is a field with a total order that is compatible with addition and multiplication.

Definition 25. A field F with a total order $\leq$ is an **ordered field** if for all $x, y, z \in F$,

$$x < y \implies x + z < y + z,$$

and

$$0 < x \text{ and } 0 < y \implies 0 < xy.$$

An $x \in F$ is called **positive** if $x > 0$ and **negative** if $x < 0$. It is called **nonnegative** if $x \geq 0$ and **nonpositive** if $x \leq 0$.

In words, in an ordered field, we can add/subtract the same element to both sides of an inequality and the product of two positive elements is positive.

Example. The fields $\mathbb{Q}$ and $\mathbb{R}$ with the standard order $\leq$ are ordered fields.

Theorem 16. Let F with $\leq$ be an ordered field. Then, for all $x, y, z \in F$,

1. $x > 0 \iff -x < 0$,
2. $x < y \iff -x > -y$,
3. $y < z$ and $x > 0 \implies xy < xz$,
4. $y < z$ and $x < 0 \implies xy > xz$,
5. $x \neq 0 \implies x^2 > 0$, in particular, $1 > 0$,
6. $x > 0 \iff x^{-1} > 0$.
7. $0 < x < y \implies x^{-1} > y^{-1} > 0$.

Proof. We leave 2. and 5. as an exercise.
1.: If $x > 0$, then

$$0 = x - x > 0 - x = -x.$$

If $-x < 0$, then $-x + x = 0 < 0 + x = x$.
3.: Since $y < z$, we have $0 = y - y < z - y$. Since $0 < x$, this implies

$$0 < x(z - y) = xz - xy.$$

Adding xy on both sides, gives $xy < xz$.
4.: If $x < 0$, then we have $-x > 0$ by 1. By 3., it follows that $-xy < -xz$ and by 2., we conclude $xy > xz$.
6.: By 5., $(x^{-1})^2 > 0$ and $x^2 > 0$. The implications "$\Rightarrow$" and "$\Leftarrow$" follow by multiplying the inequalities with $(x^{-1})^2$ and x^2, respectively.
7. Since $x > 0$ and $y > 0$, we have $xy > 0$ and also $x^{-1} > 0$ and $y^{-1} > 0$ by 6. Hence,

$$(xy)^{-1} = y^{-1}x^{-1} > 0$$

by 6. By 3., multiplying $x < y$ by $y^{-1}x^{-1}$ gives

$$y^{-1} = (y^{-1}x^{-1})x < (y^{-1}x^{-1})y = x^{-1}.$$

■

Chapter 2

Vectors, matrices, and linear maps

2.1 A two-dimensional introduction

In this section, we introduce and explain the main concepts for vectors, matrices and the corresponding linear maps in a two-dimensional setting where we do not have to struggle with indices. Everything generalizes in a straightforward manner to vectors in n-dimensional space and we give the details in the next section.

We write vectors as **column vectors**, that is, a vector $v \in \mathbb{R}^2$ is written as

$$v = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$$

with **components** $v_1, v_2 \in \mathbb{R}$. Recall that we *add two vectors*

$$v = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} \quad \text{and} \quad w = \begin{pmatrix} w_1 \\ w_2 \end{pmatrix}$$

componentwise

$$v + w = \begin{pmatrix} v_1 + w_1 \\ v_2 + w_2 \end{pmatrix}.$$

We *multiply a vector by a number* $\alpha \in \mathbb{R}$ (a **scalar**) by multiplying each component

$$\alpha v = \begin{pmatrix} \alpha v_1 \\ \alpha v_2 \end{pmatrix}.$$

Combining these two operations, we obtain **linear combinations**

$$\alpha v + \beta w = \begin{pmatrix} \alpha v_1 + \beta w_1 \\ \alpha v_2 + \beta w_2 \end{pmatrix}$$

of the vectors v, w with **coefficients** α, β.

Example.

$$2 \begin{pmatrix} 1 \\ 2 \end{pmatrix} + 3 \begin{pmatrix} 3 \\ 1 \end{pmatrix} = \begin{pmatrix} 11 \\ 7 \end{pmatrix}.$$

Every vector

$$v = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$$

can be *written uniquely* as a linear combination of the **standard unit vectors**

$$e_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad \text{and} \quad e_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

since

$$v_1 e_1 + v_2 e_2 = \begin{pmatrix} v_1 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ v_2 \end{pmatrix} = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}.$$

Example.

$$\begin{pmatrix} 11 \\ 7 \end{pmatrix} = 11 \begin{pmatrix} 1 \\ 0 \end{pmatrix} + 7 \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

From the corresponding properties in $\mathbb{R}$, it follows immediately that addition of vectors is *associative* and *commutative*, that is, for all $u, v, w \in \mathbb{R}^2$,

$$v + w = w + v \quad \text{and} \quad (u + v) + w = u + (v + w).$$

For the **zero vector** and the **opposite vector**,

$$0 = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad \text{and} \quad -v = \begin{pmatrix} -v_1 \\ -v_2 \end{pmatrix},$$

we have

$$v + 0 = 0 + v = v \quad \text{and} \quad v + (-v) = 0$$

Hence, the zero vector is the *additive identity* and the opposite vector is the *additive inverse*. In short, *vectors with addition* form an *abelian group*.

Using the corresponding identities in $\mathbb{R}$, one can verify that the **scalar multiplication** of a vector by a scalar satisfies the following properties. For all $\alpha, \beta \in \mathbb{R}$ and $v, w \in \mathbb{R}^2$,

1. $\alpha(v + w) = \alpha v + \alpha w$,
2. $(\alpha + \beta)v = \alpha v + \beta v$,
3. $(\alpha\beta)v = \alpha(\beta v)$,
4. $1v = v$.

In words, the scalar multiplication is *distributive* with respect to *vector addition* and *addition of scalars*, respectively. Moreover, it is *associative* with respect to *multiplication of scalars* and multiplication by 1 behaves as usual.

The **scalar**, **inner** or **dot product** of two vectors gives a number and it is defined by

$$v \cdot w = v_1 w_1 + v_2 w_2.$$

Example.

$$\begin{pmatrix}1\\2\end{pmatrix}\cdot\begin{pmatrix}3\\1\end{pmatrix}=1\cdot 3+2\cdot 1=5.$$

Note that, while the result of the *scalar multiplication* of a scalar and a vector is a *vector*, the result of the *scalar product* of two vectors is a *scalar*.

Two vectors are called **orthogonal** if their scalar product is zero.

Example.

$$e_1\cdot e_2=\begin{pmatrix}1\\0\end{pmatrix}\cdot\begin{pmatrix}0\\1\end{pmatrix}=1\cdot 0+0\cdot 1=0$$

Computing the *dot product* of a vector with the *standard unit vectors* gives the *components* of this vector,

$$e_1\cdot v=\begin{pmatrix}1\\0\end{pmatrix}\cdot\begin{pmatrix}v_1\\v_2\end{pmatrix}=1\cdot v_1+0\cdot v_2=v_1$$

and

$$e_2\cdot v=\begin{pmatrix}0\\1\end{pmatrix}\cdot\begin{pmatrix}v_1\\v_2\end{pmatrix}=0\cdot v_1+1\cdot v_2=v_2.$$

As an exercise, we verify that the dot product satisfies the following properties. For all $u,v,w\in\mathbb{R}^2$ and all $\alpha\in\mathbb{R}$,

1. $v\cdot w=w\cdot v$,
2. $(u+v)\cdot w=u\cdot w+v\cdot w$,
3. $(\alpha v)\cdot w=\alpha(v\cdot w)=v\cdot(\alpha w)$.

In words, the dot product is *commutative*, *distributive*, and *associative* with respect to *multiplication by scalars*.

For representing linear maps by matrices, we define how to *multiply a matrix by a vector*.

A real 2×2 **matrix** is given by an array

$$A=\begin{pmatrix}a_{11}&a_{12}\\a_{21}&a_{22}\end{pmatrix}$$

with **entries** $a_{ij}\in\mathbb{R}$, $i=1,2$, $j=1,2$. It has two **rows** and two **columns** denoted by

$$a_1=(a_{11},a_{12}),\ a_2=(a_{21},a_{22})\quad\text{and}\quad a^1=\begin{pmatrix}a_{11}\\a_{21}\end{pmatrix},\ a^2=\begin{pmatrix}a_{12}\\a_{22}\end{pmatrix},$$

respectively.
We denote the set of all 2×2 matrices by $\mathbb{R}^{2\times 2}$.

Example. The rows and columns of the matrix

$$A = \begin{pmatrix} 1 & 3 \\ 2 & 1 \end{pmatrix}$$

are

$$a_1 = (1,3),\ a_2 = (2,1) \quad \text{and} \quad a^1 = \begin{pmatrix} 1 \\ 2 \end{pmatrix},\ a^2 = \begin{pmatrix} 3 \\ 1 \end{pmatrix}.$$

Given two vectors

$$v = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} \quad \text{and} \quad w = \begin{pmatrix} w_1 \\ w_2 \end{pmatrix},$$

we denote by

$$(v, w) = \begin{pmatrix} v_1 & w_1 \\ v_2 & w_2 \end{pmatrix}$$

the matrix with columns v and w. Analogously, given two **row vectors**

$$c = (c_1, c_2) \quad \text{and} \quad d = (d_1, d_2),$$

we denote by

$$\begin{pmatrix} c \\ d \end{pmatrix} = \begin{pmatrix} c_1 & c_2 \\ d_1 & d_2 \end{pmatrix}$$

the matrix with rows c and d.

The **matrix-vector multiplication** of a matrix by a vector,

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix}, \quad v = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix},$$

gives a vector that is denoted by

$$Av.$$

Its components are defined by the *dot product* of the *rows* of A with the *vector* v,

$$Av = \begin{pmatrix} a_1 \cdot v \\ a_2 \cdot v \end{pmatrix} = \begin{pmatrix} a_{11}v_1 + a_{12}v_2 \\ a_{21}v_1 + a_{22}v_2 \end{pmatrix}.$$

Example.

$$\begin{pmatrix} 1 & 3 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} 2 \\ 3 \end{pmatrix} = \begin{pmatrix} 1 \cdot 2 + 3 \cdot 3 \\ 2 \cdot 2 + 1 \cdot 3 \end{pmatrix} = \begin{pmatrix} 11 \\ 7 \end{pmatrix}.$$

The matrix-vector multiplication Av can also be interpreted as computing *linear combinations of the columns* of A with *coefficients* given by the *components* of v. For

$$A = (a^1, a^2),$$

we have

$$v_1 a^1 + v_2 a^2 = \begin{pmatrix} v_1 a_{11} \\ v_1 a_{21} \end{pmatrix} + \begin{pmatrix} v_2 a_{12} \\ v_2 a_{22} \end{pmatrix} = \begin{pmatrix} a_{11}v_1 + a_{12}v_2 \\ a_{21}v_1 + a_{22}v_2 \end{pmatrix} = Av.$$

Example.

$$\begin{pmatrix} 1 & 3 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} 2 \\ 3 \end{pmatrix} = 2 \begin{pmatrix} 1 \\ 2 \end{pmatrix} + 3 \begin{pmatrix} 3 \\ 1 \end{pmatrix} = \begin{pmatrix} 11 \\ 7 \end{pmatrix}.$$

By the properties of the scalar multiplication and vector addition, it follows that the matrix-vector multiplication is *linear* in the following sense.

For all $A \in \mathbb{R}^{2\times 2}$, all $v, w \in \mathbb{R}^2$, and all $\alpha \in \mathbb{R}$,

$$A(v+w) = Av + Aw \quad \text{and} \quad A(\alpha v) = \alpha A v.$$

We have

$$\begin{aligned} A(v+w) &= (v_1 + w_1)a^1 + (v_2 + w_2)a^2 \\ &= v_1 a^1 + w_1 a^1 + v_2 a^2 + w_2 a^2 \\ &= (v_1 a^1 + v_2 a^2) + (w_1 a^1 + w_2 a^2) \\ &= Av + Aw \end{aligned}$$

and

$$\begin{aligned} A(\alpha v) &= (\alpha v_1)a^1 + (\alpha v_2)a^2 \\ &= \alpha(v_1 a^1) + \alpha(v_2 a^2) \\ &= \alpha(v_1 a^1 + v_2 a^2) = \alpha A v. \end{aligned}$$

Every *matrix* $A \in \mathbb{R}^{2\times 2}$ gives rise to a *function*

$$h_A \colon \mathbb{R}^2 \to \mathbb{R}^2, \quad v \mapsto Av.$$

By the linearity of the matrix-vector multiplication, it follows that h_A is *linear* in the following sense. For all $v, w \in \mathbb{R}^2$ and $\alpha \in \mathbb{R}$,

$$h_A(v+w) = h_A(v) + h_A(w) \quad \text{and} \quad h_A(\alpha v) = \alpha h_A(v).$$

We have

$$h_A(v+w) = A(v+w) = Av + Aw = h_A(v) + h_A(w)$$

and

$$h_A(\alpha v) = A(\alpha v) = \alpha A v = \alpha h_A(v).$$

Example.

1. For $A = \begin{pmatrix} 1 & 3 \\ 2 & 1 \end{pmatrix}$ and $v = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} \in \mathbb{R}^2$, we have

$$h_A(v) = \begin{pmatrix} 1 & 3 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} v_1 + 3v_2 \\ 2v_1 + v_2 \end{pmatrix}.$$

2. For the **identity matrix**
$$I_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},$$
we obtain the *identity function* on $\mathbb{R}^2$
$$h_{I_2}(v) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}.$$

3. For the **zero matrix**
$$0 = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix},$$
we obtain the *zero map* $v \mapsto 0$ on $\mathbb{R}^2$,
$$h_0(v) = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$
Note that we use the same symbol 0 to denote the *zero matrix*, the *zero element*, and the *zero vector*, respectively.

4. For $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = (a^1, a^2)$ and the standard unit vectors e_1, e_2, we have
$$h_A(e_1) = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} a_{11} \\ a_{21} \end{pmatrix} = a^1$$
and
$$h_A(e_2) = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} a_{12} \\ a_{22} \end{pmatrix} = a^2,$$
that is, we obtain the corresponding columns of A.

5. For $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathbb{R}^{2\times 2}$ and $\begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2$, we have
$$h_A(\begin{pmatrix} x \\ y \end{pmatrix}) = \begin{pmatrix} ax + by \\ cx + dy \end{pmatrix}.$$

In the terminology of the introductory examples in Section 1.1, h_A is a *multivariate vector-valued function*. We have seen above that is also linear. More generally, we say that a function
$$h\colon \mathbb{R}^2 \to \mathbb{R}^2$$
is a **linear map** if for all $v, w \in \mathbb{R}^2$ and $\alpha \in \mathbb{R}$,
$$h(v + w) = h(v) + h(w) \quad \text{and} \quad h(\alpha v) = \alpha h(v).$$
In words, a *linear map* is *compatible* with *addition* and *scalar multiplication* of vectors.

As discussed above, every matrix A gives rise to a linear map h_A given by

$$v \mapsto Av.$$

We call h_A the **linear map associated to the matrix** A. Conversely, we describe in the following how *every linear map* can be *represented by a matrix*.

We first note that a linear map

$$h\colon \mathbb{R}^2 \to \mathbb{R}^2$$

is *determined* by its *values* on the *standard unit vectors*. Recall that every vector

$$v = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$$

can be written uniquely as a linear combination of the standard unit vectors,

$$v = v_1 e_1 + v_2 e_2.$$

By linearity, we have

$$h(v) = h(v_1 e_1 + v_2 e_2) = v_1 h(e_1) + v_2 h(e_2).$$

So $h(v)$ is given by the linear combination of the vectors $h(e_1)$, $h(e_2)$ with coefficients v_1 and v_2. The matrix

$$M_h = \big(h(e_1), h(e_2)\big) \in \mathbb{R}^{2\times 2}$$

with the columns $h(e_1), h(e_2)$ given by the values of the standard unit vectors is the one, we are looking for. We have

$$M_h v = \big(h(e_1), h(e_2)\big) \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = v_1 h(e_1) + v_2 h(e_2) = h(v)$$

for all $v \in \mathbb{R}^2$. Therefore, h is the linear map associated to the matrix M_h. We call M_h the **transformation matrix of** h. In formulas,

$$h = h_{M_h}$$

and $h(v) = M_h v$ for all $v \in \mathbb{R}^2$.

Conversely, if h_A is the linear map associated to a matrix

$$A = (a^1, a^2),$$

then its transformation matrix is A since

$$(h_A(e_1), h_A(e_2)) = (a^1, a^2) = A.$$

that is,

$$M_{h_A} = A.$$

In summary, every *linear map* corresponds to a *matrix* and *vice versa*.

Example.

1. The transformation matrix of the *identity function* $\mathrm{id}_{\mathbb{R}^2}$ is the *identity matrix* I_2,

$$M_{\mathrm{id}_{\mathbb{R}^2}} = I_2.$$

2. The transformation matrix of the linear map

$$h\colon \mathbb{R}^2 \to \mathbb{R}^2, \quad (x, y) \mapsto (ax + by, cx + dy),$$

where $a, b, c, d \in \mathbb{R}$, is

$$M_h = \begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

We can *add two functions* $f, g\colon \mathbb{R}^2 \to \mathbb{R}^2$ *pointwise* by adding their values in $\mathbb{R}^2$,

$$(f + g)(v) = f(v) + g(v).$$

We can *multiply a function* $h\colon \mathbb{R}^2 \to \mathbb{R}^2$ *by a scalar* $\alpha \in \mathbb{R}$ by multiplying its value in $\mathbb{R}^2$,

$$(\alpha h)(v) = \alpha h(v).$$

If f, g, h are *linear maps*, we again obtain *linear maps*. We have for $v, w \in \mathbb{R}^2$,

$$\begin{aligned} (f + g)(v + w) &= f(v + w) + g(v + w) \\ &= f(v) + f(w) + g(v) + g(w) \\ &= f(v) + g(v) + f(w) + g(w) \\ &= (f + g)(v) + (f + g)(w) \end{aligned} \tag{2.1}$$

and for $\alpha \in \mathbb{R}$,

$$\begin{aligned} (f + g)(\alpha v) &= f(\alpha v) + g(\alpha v) \\ &= \alpha f(v) + \alpha g(v) \\ &= \alpha(f(v) + g(v)) \\ &= \alpha(f + g)(v). \end{aligned} \tag{2.2}$$

Similarly, one can verify that the product of a linear map by a scalar is a linear map. In short, the *sum/scalar multiple* of *linear maps* is *linear*.

If we define the *addition* and *scalar multiplication* of *matrices entrywise*, it corresponds to addition and scalar multiplication of linear maps. In more detail, for

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}, \quad B = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix}$$

and $\alpha \in \mathbb{R}$, we define

$$A + B = \begin{pmatrix} a_{11} + b_{11} & a_{12} + b_{12} \\ a_{21} + b_{21} & a_{22} + b_{22} \end{pmatrix} \quad \text{and} \quad \alpha A = \begin{pmatrix} \alpha a_{11} & \alpha a_{12} \\ \alpha a_{21} & \alpha a_{22} \end{pmatrix}.$$

With these definitions the *transformation matrix* of the *sum of linear maps* is the *sum* of the corresponding *transformation matrices* and analogously for the scalar multiplication of linear maps and matrices. In formulas,

$$M_{f+g} = M_f + M_g \quad \text{and} \quad M_{\alpha h} = \alpha M_h.$$

Next, we consider the *composition of linear maps*, which leads to the definition of the *matrix product*. As an exercises, we prove that the *composition* of *linear maps* is *linear*. In the following, we determine the corresponding transformation matrices.

Let

$$f\colon \mathbb{R}^2 \to \mathbb{R}^2 \quad \text{and} \quad g\colon \mathbb{R}^2 \to \mathbb{R}^2$$

be two linear maps with transformation matrices

$$M_f \quad \text{and} \quad M_g.$$

To determine the columns of the transformation matrix

$$M_{g\circ f},$$

of the composition $g \circ f$, we compute its values on the standard unit vectors. Recall that for all $v \in \mathbb{R}^2$,

$$g(v) = M_g v \quad \text{and} \quad M_f = (f(e_1), f(e_2)).$$

Hence

$$(g \circ f)(e_1) = g(f(e_1)) = M_g f(e_1) \quad \text{and} \quad (g \circ f)(e_2) = g(f(e_2)) = M_g f(e_2)$$

and

$$M_{g\circ f} = (M_g f(e_1), M_g f(e_2)).$$

So for the transformation matrix of the composition, we have to compute the matrix-vector product of the transformation Matrix M_g with the columns of M_f.

Accordingly, we define the **matrix product** of a matrix a

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix}$$

with rows a_1, a_2 with a matrix

$$B = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} = (b^1, b^2)$$

with columns b^1, b^2 by

$$A \cdot B = (Ab^1, Ab^2).$$

Then, by definition, the *transformation matrix* of a *composition* of *linear maps* is the *product* of the *transformation matrices*. In formulas,

$$M_{g\circ f} = M_g \cdot M_f.$$

If we write the matrix-vector products in terms of the dot product, we obtain

$$C = A \cdot B = \begin{pmatrix} a_1 \cdot b^1 & a_1 \cdot b^2 \\ a_2 \cdot b^1 & a_2 \cdot b^2 \end{pmatrix},$$

that is, the *entry*

$$c_{ik} = a_i \cdot b^k$$

of the *product* is obtained by computing the *dot product* of the ith *row* of A with the kth *column* of B. In terms of the entries of the matrices, we have

$$c_{ik} = a_{i1}b_{1k} + a_{i2}b_{2k}$$

for $i = 1, 2$, $k = 1, 2$.

Example. For

$$A = \begin{pmatrix} 1 & 3 \\ 2 & 1 \end{pmatrix}, \quad B = \begin{pmatrix} 2 & 1 \\ 3 & 4 \end{pmatrix},$$

we have

$$AB = \begin{pmatrix} 1 \cdot 2 + 3 \cdot 3 & 1 \cdot 1 + 3 \cdot 4 \\ 2 \cdot 2 + 1 \cdot 3 & 2 \cdot 1 + 1 \cdot 4 \end{pmatrix} = \begin{pmatrix} 11 & 13 \\ 7 & 6 \end{pmatrix}.$$

By the *correspondence* between *linear maps and matrices*, properties of functions *carry over* to *matrices*. We first discuss the *composition* of linear functions and the *matrix product* and then the *sum* of *vector-valued* functions and *matrix addition*.

As function composition, matrix multiplication is, in general, *not commutative*.

Example. For $A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ and $B = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$, we have

$$AB = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad BA = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

By Lemma 2, the *matrix product* is *associative*, that is, for all $A, B, C \in \mathbb{R}^{2\times 2}$,

$$(AB)C = A(BC).$$

By Lemma 3, multiplication with the *identity matrix* behaves like multiplying by 1,

$$I_2 A = A I_2 = A.$$

Both properties can also be proved directly using the definition of the matrix product.

One can verify that the *pointwise addition* of *vector-valued* functions $f\colon X \to \mathbb{R}^2$ from a nonempty set X to $\mathbb{R}^2$ is *associative* and *commutative*, since vector addition is associative and commutative. Similarly, from the corresponding properties of vector addition, it follows that the *zero function* $x \mapsto 0$ is the *additive identity* and the *additive inverse* of a function f is given by

$$(-f)(x) = -f(x).$$

In short, the set of vector-valued functions $f\colon X \to \mathbb{R}^2$ with the addition of $f, g\colon X \to \mathbb{R}^2$ defined by

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

form an abelian group.

If we consider the special case of linear maps $h\colon \mathbb{R}^2 \to \mathbb{R}^2$, this implies that the set of 2×2 matrices form an abelian group. So for all $A, B, C \in \mathbb{R}^{2\times 2}$, we have

$$(A+B)+C = A+(B+C) \quad \text{and} \quad A+B = B+A.$$

For the zero matrix $0 \in \mathbb{R}^{2\times 2}$ and

$$-A = \begin{pmatrix} -a_{11} & -a_{12} \\ -a_{21} & -a_{22} \end{pmatrix},$$

we have

$$A+0 = A \quad \text{and} \quad A+(-A) = 0.$$

Using the entrywise definition of matrix addition, these properties can also be verified directly.

Composition of linear maps is also *distributive* with respect to the addition of linear maps, that is, for all linear maps $f, g, h\colon \mathbb{R}^2 \to \mathbb{R}^2$,

$$f \circ (g+h) = f \circ g + f \circ h.$$

We have for all $v \in \mathbb{R}^2$,

$$\begin{aligned} (f \circ (g+h))(v) &= f((g+h)(v)) \\ &= f(g(v) + h(v)) \\ &= f(g(v)) + f(h(v)) \\ &= (f \circ g)(v) + (f \circ h)(v). \end{aligned} \tag{2.3}$$

Note that for the third equality, we use the linearity of f. Similarly, one can verify that distributivity from the right holds for arbitrary functions, that is, for all $f, g, h\colon \mathbb{R}^2 \to \mathbb{R}^2$,

$$(g+h) \circ f = g \circ f + h \circ f.$$

In terms of matrices, we have for all $A, B, C \in \mathbb{R}^{2\times 2}$,

$$A(B+C) = AB + AC \quad \text{and} \quad (B+C)A = BA + CA.$$

In summary, the set of 2×2 matrices $\mathbb{R}^{2\times 2}$ with entrywise addition and matrix multiplication form a *noncommutative ring*.

Finally, we consider *invertible linear maps*, which lead to the definition of *invertible matrices*. Let

$$f\colon \mathbb{R}^2 \to \mathbb{R}^2$$

be an invertible linear map with inverse $f^{-1}\colon \mathbb{R}^2 \to \mathbb{R}^2$, that is,

$$f^{-1} \circ f = \mathrm{id}_{\mathbb{R}^2} \quad \text{and} \quad f \circ f^{-1} = \mathrm{id}_{\mathbb{R}^2}.$$

Recall that the transformation matrix of a composition is the product of the transformation matrices and that the transformation matrix of the identity map is the identity matrix. Hence, for the corresponding transformation matrices, we have

$$M_{f^{-1}} \cdot M_f = I_2 \quad \text{and} \quad M_f \cdot M_{f^{-1}} = I_2.$$

Accordingly, a matrix $A \in \mathbb{R}^{2\times 2}$ is called **invertible** or **nonsingular** if there exists a matrix $B \in \mathbb{R}^{2\times 2}$ such that

$$BA = AB = I_2.$$

In other words, A is invertible if it is an invertible element in the ring of 2×2 matrices. If A is invertible, its inverse is uniquely determined and we write

$$A^{-1} = B.$$

Example.

1. A **diagonal matrix**
$$\operatorname{diag}(d_1, d_2) = \begin{pmatrix} d_1 & 0 \\ 0 & d_2 \end{pmatrix}$$
with $d_1, d_2 \in \mathbb{R}$ is invertible iff d_1 and d_2 are nonzero. Its inverse is given by
$$\operatorname{diag}(d_1, d_2)^{-1} = \operatorname{diag}(d_1^{-1}, d_2^{-1}).$$

2. $A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ is not invertible. For $B = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$, we have $AB = 0$ and therefore A cannot be invertible.
If A were invertible with inverse A^{-1}, then from $AB = 0$ it follows that
$$A^{-1}AB = A^{-1}0 = 0,$$
and hence $I_2B = B = 0$, a contradiction.

3. As an exercise, we prove that a matrix $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ is *invertible* iff its **determinant**
$$\det(A) = ad - bc$$
is *nonzero*. For proving this, we consider the **adjugate** of A defined by
$$\operatorname{adj}(A) = \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$
By a direct computation, one sees that
$$\operatorname{adj}(A)A = A\operatorname{adj}(A) = \operatorname{diag}(\det(A), \det(A)) = \det(A)I_2.$$
If $\det(A)$ is nonzero, then the inverse of A is given by
$$A^{-1} = \frac{1}{\det(A)} \operatorname{adj}(A)$$
So a *matrix* is *invertible* iff its *determinant* is *invertible*.

So far, we only considered 2×2 matrices and we will consider the general case of $m \times n$ matrices in the next section. In terms of matrices, we can interpret a column vector

$$v = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$$

also as a 2×1 matrix. We define its **transpose** by the corresponding row vector interpreted as a 1×2 matrix and we denote it by

$$v^\top = \begin{pmatrix} v_1 & v_2 \end{pmatrix}.$$

Conversely, we define the **transpose** of a 1×2 matrix

$$c = \begin{pmatrix} c_1 & c_2 \end{pmatrix}$$

by the corresponding column vector interpreted as a 2×1 matrix

$$c^\top = \begin{pmatrix} c_1 \\ c_2 \end{pmatrix}.$$

With this convention, the *dot product* of two vectors $v, w \in \mathbb{R}^2$ can be interpreted as the *matrix product* of the 1×2 matrix $v^\top$ with the 2×1 matrix w, whose result is 1×1 matrix given by the dot product of the first row with the first column. In formulas,

$$v^\top w = v \cdot w.$$

Example. For $v = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$, $w = \begin{pmatrix} 3 \\ 1 \end{pmatrix}$, we have

$$v^\top w = \begin{pmatrix} 1 & 2 \end{pmatrix} \begin{pmatrix} 3 \\ 1 \end{pmatrix} = 1 \cdot 3 + 2 \cdot 1 = 5.$$

More generally, we define the **transpose** of a matrix by *switching* its *row* and *column indices*. For a 2×2 matrix

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix},$$

its transpose is defined by

$$A^\top = \begin{pmatrix} a_{11} & a_{21} \\ a_{12} & a_{22} \end{pmatrix}.$$

So the *columns* of a matrix A are the *rows* of its transpose $A^\top$ and the *rows* of a A are the *columns* of $A^\top$, respectively.

Example. The transpose of $A = \begin{pmatrix} 1 & 3 \\ 2 & 1 \end{pmatrix}$ is

$$A^\top = \begin{pmatrix} 1 & 2 \\ 3 & 1 \end{pmatrix}$$

Clearly, transposing twice gives the original matrix, that is,

$$(A^\top)^\top = A,$$

and the transpose of the identity matrix is the identity matrix

$$I_2^\top = I_2.$$

More generally, a matrix is called **symmetric** if

$$A^\top = A.$$

A 2×2 matrix $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$ is symmetric iff $a_{12} = a_{21}$.

One can also verify directly that taking the *transpose reverses the order* of the matrix multiplication, that is,

$$(AB)^\top = B^\top A^\top$$

Hence, for an *invertible* matrix A with inverse A^{-1}, we obtain from

$$A^{-1}A = I_2 \quad \text{and} \quad AA^{-1} = I_2$$

by taking the transpose

$$A^\top (A^{-1})^\top = I_2 \quad \text{and} \quad (A^{-1})^\top A^\top = I_2.$$

Therefore, $A^\top$ is invertible with inverse

$$(A^\top)^{-1} = (A^{-1})^\top.$$

Finally, a direct computation shows that the *matrix-vector product*, the *dot product*, and the *transpose*, are connected by the following identity

$$Av \cdot w = v \cdot A^\top w.$$

2.2 The n-dimensional case

We discuss the generalizations of the notions for vectors, matrices, and linear maps from $\mathbb{R}^2$ in the previous section to the n-dimensional space. We note that we did not use any special properties of the real numbers other than being a field so we consider vector, matrices, and linear maps over an arbitrary field. In fact, the results can also be formulated for vectors and matrices with entries in an arbitrary commutative ring and, for example, integer matrices appear in many applications. However, for simplicity, we

restrict ourselves to vectors and matrices with entries in a field. We also refer to the previous section for further explanations and examples.

Throughout this section, F denotes a *field*. We write vectors as **column vectors**, that is, a vector $v \in F^n$ is written as

$$v = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}$$

with **components** $v_1, v_2, \ldots, v_n \in F$. We define the **addition** and **scalar multiplication** of *vectors*

$$+\colon F^n \times F^n \to F^n \quad \text{and} \quad \cdot\colon F \times F^n \to F^n$$

componentwise by

$$\begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} + \begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix} \mapsto \begin{pmatrix} v_1 + w_1 \\ \vdots \\ v_n + w_n \end{pmatrix} \quad \text{and} \quad \alpha \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} \mapsto \begin{pmatrix} \alpha v_1 \\ \vdots \\ \alpha v_n \end{pmatrix}.$$

The **zero vector** and the **opposite vector** are defined by

$$0 = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix} \quad \text{and} \quad -v = \begin{pmatrix} -v_1 \\ \vdots \\ -v_n \end{pmatrix}.$$

Using the corresponding properties for the components of the vectors in F, one can verify the following properties of the vector addition and scalar multiplication.

Theorem 17. The set of all vectors F^n with componentwise addition is an abelian group where the zero vector is the additive identity and the opposite vector is the additive inverse.
The scalar multiplication satisfies for all $\alpha, \beta \in F$ and all $v, w \in \mathbb{F}^n$,

1. $\alpha(v + w) = \alpha v + \alpha w$,
2. $(\alpha + \beta)v = \alpha v + \beta v$,
3. $(\alpha\beta)v = \alpha(\beta v)$,
4. $1v = v$.

Definition 26. If $v_1, \ldots, v_m \in F^n$ and $\alpha_1, \ldots, \alpha_m \in F$, we call

$$\alpha_1 v_1 + \cdots + \alpha_m v_m$$

a **linear combination** of $v_1, \ldots, v_m$ with **coefficients** $\alpha_1, \ldots, \alpha_m$.

Every vector $v \in F^n$ can be *written uniquely* as a linear combination of the **standard unit vectors** in F^n,

$$e_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \; e_2 = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \quad \ldots \quad, \; e_n = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix},$$

since

$$\sum_{i=1}^{n} v_i e_i = v_1 e_1 + \cdots + v_n e_n = v.$$

Example.

$$\begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \end{pmatrix} = 1 \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} + 2 \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} + 3 \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} + 4 \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}.$$

Definition 27. The **scalar**, **inner** or **dot product**

$$\cdot \colon F^n \times F^n \to F$$

is defined by

$$v \cdot w = \sum_{i=1}^{n} v_i w_i = v_1 w_1 + \cdots + v_n w_n.$$

Two vectors $v, w \in F^n$ are called **orthogonal** if $v \cdot w = 0$.

Example. The standard unit vectors are pairwise orthogonal, that is, for $i, j = 1, \ldots, n$ and $i \neq j$,

$$e_i \cdot e_j = 0.$$

Moreover, for $i = 1, \ldots, n$,

$$e_i \cdot e_i = 1.$$

Using the **Kronecker delta** defined by

$$\delta_{ij} = \begin{cases} 1, & \text{if } i = j, \\ 0, & \text{otherwise,} \end{cases}$$

we can write this compactly as

$$e_i \cdot e_j = \delta_{ij}$$

for $i, j = 1, \ldots, n$.

One can easily verify that the dot product satisfies the following properties.

Theorem 18. For all $u, v, w \in F^n$ and all $\alpha \in F$,

1. $v \cdot w = w \cdot v$,
2. $(u + v) \cdot w = u \cdot w + v \cdot w$,
3. $(\alpha v) \cdot w = \alpha(v \cdot w) = v \cdot (\alpha w)$.

In words, the dot product is *commutative* and (2.) and (3.) can also be stated by saying that the dot product is *linear* in *each factor*.

Computing the *dot product* of a vector with the *standard unit vectors* gives the *components* of this vector. We write a vector $v \in F^n$ as a linear combination of the standard unit vectors

$$v = \sum_{i=1}^{n} v_i e_i.$$

By the linearity of the dot product in the first factor, we have for all $j = 1, \ldots, n$,

$$v \cdot e_j = (\sum_{i=1}^{n} v_i e_i) \cdot e_j = \sum_{i=1}^{n} v_i (e_i \cdot e_j) = \sum_{i=1}^{n} v_i \delta_{ij} = v_j.$$

For representing linear maps by matrices, we define how to *multiply a matrix by a vector*.

Definition 28. Let $m, n \in \mathbb{N}$ with $m, n \geq 1$. An $m \times n$ **matrix** (m by n matrix) over a field F is given by an array

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}$$

with **entries** $a_{ij} \in F$, $i = 1, \ldots, m$, $j = 1, \ldots, n$. We also write

$$A = (a_{ij})_{i=1,j=1}^{m,n} \quad \text{or} \quad A = (a_{ij})$$

to denote an $m \times n$ matrix and we write $F^{m \times n}$ for the set of all $m \times n$-matrices over F. A $m \times n$ matrix $A = (a_{ij})$ has m **rows** and n **columns** and we denote the ith row and the jth column by

$$a_i = (a_{i1}, \ldots, a_{in}) \in F^{1 \times n} \quad \text{and} \quad a^j = \begin{pmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{pmatrix} \in F^{m \times 1} = F^m,$$

respectively.

Example. The rows and columns of the matrix

$$A = \begin{pmatrix} 1 & 3 & -1 \\ 2 & 1 & 4 \end{pmatrix}$$

are

$$a_1 = (1, 3, -1),\ a_2 = (2, 1, 4) \quad \text{and} \quad a^1 = \begin{pmatrix} 1 \\ 2 \end{pmatrix},\ a^2 = \begin{pmatrix} 3 \\ 1 \end{pmatrix},\ a^3 = \begin{pmatrix} -1 \\ 4 \end{pmatrix}.$$

Given n (column) vectors $v_1, \ldots, v_n \in F^m$, we denote by

$$\big(v_1, \ldots, v_n\big) \in F^{m \times n}$$

the matrix with *columns* $v_1, \dots, v_n$. Given m **row vectors** $c_1, \dots, c_m \in F^{1\times n}$, we denote by

$$\begin{pmatrix} c_1 \\ \vdots \\ c_m \end{pmatrix} \in F^{m\times n}$$

the matrix with *rows* $c_1, \dots, c_m$.

Definition 29. The **matrix-vector multiplication** of a matrix $A \in F^{m\times n}$ by a vector in $v \in F^n$,

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}, \quad v = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix},$$

gives a vector in F^m whose components are defined by the dot product of the rows of A with the vector v,

$$Av = \begin{pmatrix} a_1 \cdot v \\ \vdots \\ a_m \cdot v \end{pmatrix} = \begin{pmatrix} a_{11}v_1 + \cdots + a_{1n}v_n \\ \vdots \\ a_{m1}v_1 + \cdots + a_{mn}v_n \end{pmatrix} \in F^m.$$

Example.

$$\begin{pmatrix} 1 & 3 & -1 \\ 2 & 1 & 4 \end{pmatrix} \begin{pmatrix} 2 \\ 3 \\ 1 \end{pmatrix} = \begin{pmatrix} 1\cdot 2 + 3\cdot 3 - 1\cdot 1 \\ 2\cdot 2 + 1\cdot 3 + 4\cdot 1 \end{pmatrix} = \begin{pmatrix} 10 \\ 11 \end{pmatrix}.$$

The matrix-vector multiplication Av can also be interpreted as computing *linear combinations of the columns* of A with *coefficients* given by the *components* of v. For

$$A = (a^1, \dots, a^n),$$

we have

$$v_1 a^1 + \cdots + v_n a^n = \begin{pmatrix} v_1 a_{11} \\ \vdots \\ v_1 a_{m1} \end{pmatrix} + \cdots + \begin{pmatrix} v_n a_{1n} \\ \vdots \\ v_n a_{mn} \end{pmatrix} = \begin{pmatrix} a_{11}v_1 + \cdots + a_{1n}v_n \\ \vdots \\ a_{m1}v_1 + \cdots + a_{mn}v_n \end{pmatrix} = Av.$$

Example.

$$\begin{pmatrix} 1 & 3 & -1 \\ 2 & 1 & 4 \end{pmatrix} \begin{pmatrix} 2 \\ 3 \\ 1 \end{pmatrix} = 2\begin{pmatrix} 1 \\ 2 \end{pmatrix} + 3\begin{pmatrix} 3 \\ 1 \end{pmatrix} + 1\begin{pmatrix} -1 \\ 4 \end{pmatrix} = \begin{pmatrix} 10 \\ 11 \end{pmatrix}.$$

By the properties of the scalar multiplication and vector addition, one can verify analogously to the two-dimensional case that the matrix-vector multiplication is *linear* in the following sense. For all $A \in F^{m\times n}$, all $v, w \in \mathbb{F}^n$, and all $\alpha \in F$,

$$A(v + w) = Av + Aw \quad \text{and} \quad A(\alpha v) = \alpha Av.$$

Definition 30. We say that a function

$$h\colon F^n \to F^m$$

is a **linear map** if for all $v, w \in F^n$ and all $\alpha \in F$,

$$h(v+w) = h(v) + h(w) \quad \text{and} \quad h(\alpha v) = \alpha h(v).$$

In words, a *linear map* is *compatible* with *addition* and *scalar multiplication* of vectors.

Every *matrix* $A \in F^{m\times n}$ gives rise to a *function*

$$h_A\colon F^n \to F^m, \quad v \mapsto Av.$$

By the linearity of the matrix-vector multiplication, it follows that h_A is a linear map from F^n to F^m and we call h_A the **linear map associated to the matrix** A.

Example.

1. For $A = \begin{pmatrix} 1 & 3 & -1 \\ 2 & 1 & 4 \end{pmatrix}$, we have $h_A\colon F^3 \to F^2$ and

$$h_A(v) = \begin{pmatrix} 1 & 3 & -1 \\ 2 & 1 & 4 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} = \begin{pmatrix} v_1 + 3v_2 - v_3 \\ 2v_1 + v_2 + 4v_3 \end{pmatrix}.$$

2. For the **identity matrix**

$$I_n = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & 1 \end{pmatrix} \in F^{n\times n},$$

   we obtain the *identity function* $\mathrm{id}_{F^n}\colon F^n \to F^n$, $v \mapsto v$, that is,

$$h_{I_n} = \mathrm{id}_{F^n}.$$

3. For the **zero matrix**

$$0 = \begin{pmatrix} 0 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & 0 \end{pmatrix} \in F^{m\times n},$$

   we obtain the *zero map*

$$h_0\colon F^n \to F^m, \quad v \mapsto 0.$$

4. For
$$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} = (a^1, \ldots, a^n) \in F^{m\times n},$$
and the the standard unit vectors $e_1, \ldots, e_n \in F^n$, we have for $i = 1, \ldots, n$,
$$h_A(e_i) = a^i,$$
that is, we obtain the corresponding columns of A.

5. For a row vector $c = (c_1, \ldots, c_n) \in F^{1\times n}$, we obtain a linear map that assigns to each vector $v \in F^n$ an element in F given by the dot product with c,
$$h_c \colon F^n \to F, \quad v \mapsto c_1v_1 + \cdots + c_nv_n = c \cdot v.$$
Linear maps whose codomain is F are also called **linear functionals**.

In the following, we describe how *every linear map* can be *represented by a matrix*. First, we note that a linear map
$$h \colon F^n \to F^m$$
is *determined* by its *values* on the *standard unit vectors*. Recall that every vector $v \in F^n$ can be written uniquely as a linear combination of the standard unit vectors in F^n,
$$v = \sum_{i=1}^{n} v_i e_i.$$
By the linearity of h, we have
$$h(v) = h(\sum_{i=1}^{n} v_i e_i) = \sum_{i=1}^{n} v_i h(e_i). \tag{2.4}$$
So $h(v)$ is given by the linear combination of the vectors $h(e_1), \ldots, h(e_n) \in F^m$ with coefficients $v_1, \ldots, v_n \in F$.

Definition 31. If $h \colon F^n \to F^m$ is a linear map, its **transformation matrix** is defined by
$$M_h = \big(h(e_1), \ldots, h(e_n)\big) \in F^{m\times n}.$$

Example.

1. The transformation matrix of the *identity function* id_{F^n} is the *identity matrix* I_n,
$$M_{\mathrm{id}_{F^n}} = I_n.$$

2. The transformation matrix of the linear map
$$h \colon F^3 \to F^2, \quad (x, y, z) \mapsto (ax + by + cz, dx + ey + fz),$$
where $a, b, c, d, e, f \in F$, is
$$M_h = \begin{pmatrix} a & b & c \\ d & e & f \end{pmatrix}.$$

The next theorem shows that every *linear map* corresponds to a *matrix* and *vice versa*.

Theorem 19. Let

$$h\colon F^n \to F^m$$

be a linear map and M_h be its transformation matrix. Then, h is the linear map associated to the matrix M_h, that is,

$$h = h_{M_h}.$$

If h_A is the linear map associated to a matrix

$$A = (a^1, \dots, a^n) \in F^{m\times n},$$

then its transformation matrix is A, that is,

$$M_{h_A} = A.$$

Proof. Let $v \in F^n$. Using Equation (2.4), we have

$$h_{M_h}(v) = M_h v = \big(h(e_1), \dots, h(e_n)\big) \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} = \sum_{i=1}^{n} v_i h(e_i) = h(v),$$

and hence, $h = h_{M_h}$.
We also have

$$(h_A(e_1), \dots, h_A(e_n)) = (a^1, \dots, a^n) = A,$$

and hence, $M_{h_A} = A$. ∎

We can *add two functions*

$$f, g\colon X \to F^m$$

from a nonempty set X to F^m *pointwise* by adding their values in F^m,

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

for $x \in X$. Note that we can only *add functions* with the *same domain and codomain*. Similarly, we can *multiply a function by a scalar* $\alpha \in F$ by multiplying its value in F^m,

$$(\alpha f)(x) = \alpha f(x).$$

Considering the special case $X = F^n$, we show that the *sum/scalar multiple* of *linear maps* is *linear*.

Theorem 20. If $f, g\colon F^n \to F^m$ are linear maps, then

$$f + g\colon F^n \to F^m \quad \text{and} \quad \alpha f\colon F^n \to F^m$$

for $\alpha \in F$ are linear maps.

Proof. The sum $f + g$ is a linear map, since Equations (2.1) and (2.2) hold for all $v, w \in F^n$ and all $\alpha \in F$.
Similarly, one can verify that αf is a linear map. ∎

We define the **addition** and **scalar multiplication** of *matrices*

$$+\colon F^{m\times n} \times F^{m\times n} \to F^{m\times n} \quad \text{and} \quad \cdot\colon F \times F^{m\times n} \to F^{m\times n}$$

entrywise by

$$(a_{ij}) + (b_{ij}) \mapsto (a_{ij} + b_{ij}) \quad \text{and} \quad \alpha(a_{ij}) \mapsto (\alpha a_{ij}).$$

Note that we can only *add matrices* of the *same size*.

Example. For

$$A = \begin{pmatrix} 1 & 3 & -1 \\ 2 & 1 & 4 \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} 2 & -1 & 1 \\ 1 & 3 & -2 \end{pmatrix},$$

we have

$$A + B = \begin{pmatrix} 3 & 2 & 0 \\ 3 & 4 & 2 \end{pmatrix} \quad \text{and} \quad 2A = \begin{pmatrix} 2 & 6 & -2 \\ 4 & 2 & 8 \end{pmatrix}.$$

One can verify that with these definitions the *transformation matrix* of the *sum of linear maps* is the *sum* of the corresponding *transformation matrices* and analogously for the scalar multiplication of linear maps and matrices. In formulas,

$$M_{f+g} = M_f + M_g \quad \text{and} \quad M_{\alpha h} = \alpha M_h.$$

The *composition of linear maps* leads to the definition of the *matrix product*. As an exercises, we proved that the *composition* of *linear maps* is *linear*. The next statement gives the the corresponding transformation matrices.

Theorem 21. Let

$$f\colon F^r \to F^n \quad \text{and} \quad g\colon F^n \to F^m$$

be two linear maps. Then, the transformation matrix of the composition

$$g \circ f\colon F^r \to F^m$$

is given by

$$M_{g\circ f} = (M_g f(e_1), \ldots, M_g f(e_r)) \in F^{m\times r}.$$

Proof. For $i = 1, \ldots, r$ and the ith standard unit vector $e_i \in F^r$, we have

$$(g \circ f)(e_i) = g(f(e_i)) = M_g f(e_i),$$

which is by definition the ith column of $M_{g \circ f}$. ∎

So for the transformation matrix

$$M_{g \circ f} \in F^{m \times r}$$

of the composition $g \circ f$, we have to compute the matrix-vector product of the transformation matrix

$$M_g \in F^{m \times n}$$

with the columns of transformation matrix

$$M_f = \big(f(e_1), \ldots, f(e_r)\big) \in F^{n \times r}.$$

Accordingly, we define the *matrix product* of an $m \times n$ matrix A with an $n \times r$ matrix B. In particular, the product of two matrices is defined only when number of *columns* of A *equals* the number of *rows* of B. The product AB has as many rows as A and as many columns as B. This relation between columns and rows is easily remembered when it is written as

$$\underset{m \times n}{A} \; \underset{n \times r}{B} = \underset{m \times r}{C}.$$

Definition 32. The **matrix product** of an $m \times n$ matrix

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} = \begin{pmatrix} a_1 \\ \vdots \\ a_m \end{pmatrix} \in F^{m \times n}$$

with rows $a_1, \ldots, a_m$ with an $n \times r$ matrix

$$B = \begin{pmatrix} b_{11} & \cdots & b_{1r} \\ \vdots & \ddots & \vdots \\ b_{n1} & \cdots & b_{nr} \end{pmatrix} = (b^1, \ldots, b^r) \in F^{n \times r}$$

with columns $b^1, \ldots, b^r$ is defined as the $m \times r$ matrix given by

$$A \cdot B = (Ab^1, \ldots, Ab^r) \in F^{m \times r}.$$

In terms of the dot product, we have

$$C = A \cdot B = \begin{pmatrix} a_1 \cdot b^1 & \cdots & a_1 \cdot b^r \\ \vdots & \ddots & \vdots \\ a_m \cdot b^1 & \cdots & a_m \cdot b^r \end{pmatrix},$$

that is, the *entry*

$$c_{ik} = a_i \cdot b^k$$

of the *product* is obtained by computing the *dot product* of the ith *row* of A with the kth *column* of B. In terms of the entries of the matrices, we have

$$c_{ik} = \sum_{j=1}^{n} a_{ij} b_{jk}$$

for $i = 1, \ldots, m$, $k = 1, \ldots, r$.

By definition, the *transformation matrix* of a *composition* of *linear maps* is the *product* of the *transformation matrices*. In formulas,

$$M_{g \circ f} = M_g \cdot M_f.$$

Example. For

$$A = \begin{pmatrix} 1 & 3 & -1 \\ 2 & 1 & 4 \end{pmatrix}, \quad B = \begin{pmatrix} 2 & 1 \\ 3 & 4 \\ 1 & 2 \end{pmatrix},$$

we have

$$AB = \begin{pmatrix} 1 \cdot 2 + 3 \cdot 3 - 1 \cdot 1 & 1 \cdot 1 + 3 \cdot 4 - 1 \cdot 2 \\ 2 \cdot 2 + 1 \cdot 3 + 4 \cdot 1 & 2 \cdot 1 + 1 \cdot 4 + 4 \cdot 2 \end{pmatrix} = \begin{pmatrix} 10 & 11 \\ 11 & 14 \end{pmatrix}.$$

Properties of functions *carry over* to *matrices* by the *correspondence* between *linear maps and matrices*.

By Lemma 2, the *matrix product* is *associative* and by Lemma 3, multiplication with the *identity matrix* behaves like multiplying by 1. Note that the *sizes* of the *matrices* have to *match* for the the matrix product to be defined.

Theorem 22. For $A \in F^{m \times n}, B \in F^{n \times r}, C \in F^{r \times s}$, we have

$$(AB)C = A(BC)$$

and

$$I_m A = A \quad \text{and} \quad A I_n = A.$$

As in the two-dimensional case, one sees that the set of *vector-valued functions*

$$f \colon X \to F^m$$

from a nonempty set X to F^m with the *pointwise addition* of $f, g \colon X \to F^m$ defined by

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

form an *abelian group*. Considering the special case of linear maps

$$h \colon F^n \to F^m,$$

we see that this holds in particular for $m \times n$ matrices.

Theorem 23. The set of all $m \times n$ matrices $F^{m\times n}$ with entrywise addition is an abelian group.

Composition of linear maps is also *distributive* with respect to the addition of linear maps, that is, for all linear maps

$$f\colon F^n \to F^m \quad \text{and} \quad g, h\colon F^r \to F^n,$$

we have

$$f \circ (g + h) = f \circ g + f \circ h,$$

since Equation (2.3) holds for all $v \in F^r$. Similarly, one can verify that distributivity from the right holds for arbitrary functions, that is, for all

$$g, h\colon F^n \to F^m \quad \text{and} \quad f\colon F^r \to F^n,$$

we have

$$(g + h) \circ f = g \circ f + h \circ f.$$

In terms of matrices, we obtain the following statements.

Theorem 24. For all $A \in F^{m\times n}$ and all $B, C \in F^{n\times r}$,

$$A(B + C) = AB + AC.$$

For all $B, C \in F^{m\times n}$ and all $A \in F^{n\times r}$,

$$(B + C)A = BA + CA.$$

As an immediate consequence of Theorem 22, 23, and 24, we obtain the following corollary.

Corollary 25. The set of $n \times n$ matrices $F^{n\times n}$ with entrywise addition and matrix multiplication is a noncommutative ring.

In the following, we consider *invertible linear maps*, which lead to the definition of *invertible matrices*. Let

$$f\colon F^n \to F^n$$

be an invertible linear map with inverse $f^{-1}\colon F^n \to F^n$, that is,

$$f^{-1} \circ f = \mathrm{id}_{F^n} \quad \text{and} \quad f \circ f^{-1} = \mathrm{id}_{F^n}.$$

Since the transformation matrix of a composition is the product of the transformation matrices and the transformation matrix of the identity map is the identity matrix, we have

$$M_{f^{-1}} \cdot M_f = I_n \quad \text{and} \quad M_f \cdot M_{f^{-1}} = I_n.$$

Accordingly, we define when an $n \times n$ matrix is invertible.

Definition 33. A matrix $A \in F^{n\times n}$ is called **invertible** or **nonsingular** if there exists a matrix $B \in F^{n\times n}$ such that

$$BA = AB = I_n.$$

Such a matrix B is uniquely determined. If A is invertible, we call B the **inverse** or **inverse matrix** of A and write

$$A^{-1} = B.$$

In other words, A is invertible if it is an invertible element in the ring of $n \times n$ matrices, see Definition 13. In particular, we know that the *product of invertible matrices* $A, B \in F^{n\times n}$ is *invertible* and

$$(AB)^{-1} = B^{-1}A^{-1}.$$

Example. A **diagonal matrix**

$$\operatorname{diag}(d_1, \dots, d_n) = \begin{pmatrix} d_1 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & d_n \end{pmatrix}$$

with $d_1, \dots, d_n \in F$ is invertible iff all d_i's are nonzero. In this case, its inverse is given by

$$\operatorname{diag}(d_1, \dots, d_n)^{-1} = \operatorname{diag}(d_1^{-1}, \dots, d_n^{-1}).$$

This follows from the fact that for the product of two diagonal matrices, we have

$$\operatorname{diag}(a_1, \dots, a_n) \operatorname{diag}(b_1, \dots, b_n) = \operatorname{diag}(a_1b_1, \dots, a_nb_n).$$

We discuss the *determinant* and *adjugate matrix* of an $n \times n$ matrix, which gives an explicit formula for the inverse matrix, in a later section.

Finally, we study the *transpose* of an $m \times n$ matrix. It is the $n \times m$ matrix obtained by *switching row* and *column indices*.

Definition 34. For a matrix $A = (a_{ij}) \in F^{m\times n}$, its **transpose** is defined as the matrix

$$A^\top = (a'_{ij}) \in F^{n\times m},$$

where $a'_{ij} = a_{ji}$, that is,

$$A^\top = \begin{pmatrix} a_{11} & \cdots & a_{m1} \\ \vdots & \ddots & \vdots \\ a_{1n} & \cdots & a_{mn} \end{pmatrix}.$$

So the *columns* of a matrix A are the *rows* of its transpose $A^\top$ and the *rows* of a A are the *columns* of $A^\top$, respectively.

Example. The transpose of

$$A = \begin{pmatrix} 1 & 3 & -1 \\ 2 & 1 & 4 \end{pmatrix}$$

is

$$A^\top = \begin{pmatrix} 1 & 2 \\ 3 & 1 \\ -1 & 4 \end{pmatrix}.$$

In particular, we note that the transpose of a *column vector*

$$v = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} \in F^n,$$

interpreted as an $n \times 1$ matrix, is the $1 \times n$ matrix

$$v^\top = \begin{pmatrix} v_1 & \cdots & v_n \end{pmatrix},$$

which can be interpreted as a *row vector*. Conversely, the transpose of a row vector

$$c = (c_1, \ldots, c_n),$$

interpreted as a $1 \times n$ matrix, is the $n \times 1$ matrix

$$c^\top = \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix},$$

which can be interpreted as a *column vector*.

With this convention, the *dot product* of two vectors

$$v, w \in F^n$$

is the *matrix product* of the $1 \times n$ matrix $v^\top$ (a *row vector*) with the $n \times 1$ matrix w (a *column vector*). The result is a 1×1 matrix, which we identify with the corresponding element in F. In formulas,

$$v^\top w = v \cdot w.$$

Example. For

$$v = \begin{pmatrix} 1 \\ 3 \\ -1 \end{pmatrix}, \quad w = \begin{pmatrix} 2 \\ 3 \\ 1 \end{pmatrix},$$

we have

$$v^\top w = \begin{pmatrix} 1 & 3 & -1 \end{pmatrix} \begin{pmatrix} 2 \\ 3 \\ 1 \end{pmatrix} = 1 \cdot 2 + 3 \cdot 3 - 1 \cdot 1 = 10.$$

Note that if we compute the matrix product of the $n \times 1$ matrix v (a *column vector*) with the $1 \times n$ matrix $w^\top$ (a *row vector*), we obtain an $n \times n$ matrix. More generally, the *outer product* of two column vectors of different size is defined as follows.

Definition 35. For $u \in F^m$ and $v \in F^n$, the **outer product** is the $m \times n$ matrix defined by

$$uv^\top = \begin{pmatrix} u_1v_1 & \cdots & u_1v_n \\ \vdots & \ddots & \vdots \\ u_mv_1 & \cdots & u_mv_n \end{pmatrix}.$$

So the outer product is obtained by multiplying each component of u with each component of v. Note that *each column* of the outer product $uv^\top$ is *a multiple* of u and each *row* is *a multiple* of v, respectively.

Example. For

$$u = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \quad \text{and} \quad v = \begin{pmatrix} 4 \\ 5 \end{pmatrix}$$

the outer product is

$$uv^\top = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \begin{pmatrix} 4 & 5 \end{pmatrix} = \begin{pmatrix} 4 & 5 \\ 8 & 10 \\ 12 & 15 \end{pmatrix}.$$

Clearly, transposing a matrix A twice gives the original matrix, that is,

$$(A^\top)^\top = A,$$

and the transpose of the identity matrix is the identity matrix

$$I_n^\top = I_n.$$

More generally, we have the following important class of $n \times n$ matrices.

Definition 36. A matrix $A \in F^{n \times n}$ is called **symmetric** if

$$A^\top = A.$$

One can also verify that taking the *transpose reverses the order* of the matrix multiplication. As in the previous section, one also sees that the transpose of an invertible matrix is invertible and that its inverse satisfies the identity in the following statement.

Theorem 26. For all $A \in F^{m \times n}$ and $B \in F^{n \times r}$, we have

$$(AB)^\top = B^\top A^\top.$$

If $A \in F^{n \times n}$ is invertible, then $A^\top$ is invertible with inverse

$$(A^\top)^{-1} = (A^{-1})^\top.$$

Finally, as an exercise, we show that the *matrix-vector product*, the *dot product*, and the *transpose*, are connected by the following identity

$$(Av) \cdot w = v \cdot (A^\top w).$$

Chapter 3

Convergence and continuity

A fundamental notion in analysis is the *convergence* of a *sequence* of objects to a *limit*. Recall from Definition 5 that a sequence in a set X is a function $a\colon \mathbb{N} \to X$. We write $a_n = a(n)$ for the nth element of a sequence and

$$(a_n)_{n\in\mathbb{N}} = (a_0, a_1, a_2, \ldots)$$

for the sequence a. Convergence formalizes the concept that a sequence gets as close as we want to its limit if we go far enough out in the sequence. Convergence and limits are used to define continuity, derivatives, and integrals. In applications, sequences are, for example, used to analyze iterative methods that generate a sequence of approximate solutions for a class of problems.

3.1 Real sequences and series

In this section, we discuss sequences of real numbers. In the following, we write (a_n) for a sequence $(a_n)_{n\in\mathbb{N}}$ in $\mathbb{R}$. We also use the properties of inequalities from Definition 25 and Theorem 16 without explicitly mentioning them.

To measure the distance between real numbers, we consider the absolute value. As an exercise, we proved the following properties of the absolute value function.

Lemma 27. For all $x, y \in \mathbb{R}$,

1. $|x| \geq 0$,
2. $|x| = 0 \iff x = 0$,
3. $|x| \geq x$,
4. $|x \cdot y| = |x| \cdot |y|$.

Note that from 4. with $y = -1$, it follows in particular that

$$|-x| = |x|.$$

For two real numbers a and b, we think of

$$|b - a| = |a - b|$$

as the *distance* between the points a and b on the number line. For a given real number $a \in \mathbb{R}$ and a positive $\varepsilon > 0$, the distance of an $x \in \mathbb{R}$ to a is less than ε if

$$|x - a| < \varepsilon$$

The set of all such $x \in \mathbb{R}$ is called an ε-neighborhood of a.

Definition 37. For $a \in \mathbb{R}$ and $\varepsilon > 0$, the set

$$U_\varepsilon(a) = \{x \in \mathbb{R} \mid |x - a| < \varepsilon\}$$

is called an **ε-neighborhood** of a.

In other words, $U_\varepsilon(a)$ is the open interval, centered at a, with radius ε, that is,

$$U_\varepsilon(a) = (a - \varepsilon, a + \varepsilon).$$

We also note that if two real numbers a and b are *arbitrarily close* in the sense that for every $\varepsilon > 0$,

$$|a - b| < \varepsilon,$$

then they are *equal* $a = b$. Assume, for contradiction, $a \neq b$. For $\varepsilon = |a - b| > 0$, we have

$$\varepsilon = |a - b| < \varepsilon.$$

An important property of the absolute value is the so-called **triangle inequality**.

Theorem 28. For all $x, y \in \mathbb{R}$,

$$|x + y| \leq |x| + |y|.$$

Proof. From $x \leq |x|$ and $y \leq |y|$, it follows that

$$x + y \leq |x| + y \quad \text{and} \quad |x| + y \leq |x| + |y|.$$

By transitivity, we have

$$x + y \leq |x| + |y|.$$

From $-x \leq |-x| = |x|$ and $-y \leq |-y| = |y|$, it follows analogously that

$$-(x + y) = -x - y \leq |x| + |y|.$$

Hence, $|x + y| \leq |x| + |y|$. ∎

The triangle inequality is often employed in the following way, which also explains its name. Given $a, b, c \in \mathbb{R}$, we obviously have

$$|a - b| = |(a - c) + (c - b)|.$$

By the triangle inequality, it follows that

$$|(a-c)+(c-b)| \leq |a-c|+|c-b|,$$

and hence

$$|a-b| \leq |a-c|+|c-b|.$$

Interpreting the absolute values as distances, the inequality states that the distance from a to b is less or equal to the distance from a to c plus the distance from c to b.

As an exercise, we prove the following consequence of the triangle inequality.

Corollary 29. For all $x, y \in \mathbb{R}$,

$$||x|-|y|| \leq |x-y|.$$

For investigating the behavior of some sequences, we also recall **Bernoulli's inequality**, which we discussed as an example for a proof by induction. It approximates powers of $(1+x)$ and states that for a real number $x \geq -1$, we have for all $n \in \mathbb{N}$,

$$(1+x)^n \geq 1+nx.$$

As an application, we prove that the powers of a real number

$$b > 1$$

grow *arbitrarily large* in the following sense.

Theorem 30. Let $b > 1$. Then, for every $K > 0$, there exists an $n \in \mathbb{N}$ such that $b^n > K$.

Proof. Let $K > 0$ and let $x = b-1$. Then, $x > 0$ and, by Bernoulli's inequality, we have

$$b^n = (1+x)^n \geq 1+nx.$$

By the Archimedean property (Theorem 14) of the real numbers, there exists an $n \in \mathbb{N}$ such that

$$n > \frac{K-1}{x}.$$

Hence, $nx > K-1$ and $b^n > K$. ∎

As a consequence, we obtain that powers of a real number

$$0 < b < 1$$

get arbitrarily close to 0, which we prove as an exercise.

Corollary 31. Let $0 < b < 1$. Then, for every $\varepsilon > 0$, there exists an $n \in \mathbb{N}$ such that $b^n < \varepsilon$.

The following definition formalizes what it means that a convergent real sequence gets as close as we want to its limit if we go far enough out in the sequence.

Definition 38. A real sequence (a_n) **converges** to a real number a if for every $\varepsilon > 0$, there exists an $N \in \mathbb{N}$ such that for all $n \in \mathbb{N}$, $n \geq N$ implies that

$$|a_n - a| < \varepsilon.$$

If (a_n) converges to a, then a is uniquely determined and it is called the **limit** of the sequence (a_n). In this case, we write

$$\lim_{n\to\infty} a_n = a, \quad \lim a_n = a, \quad \text{or} \quad (a_n) \to a.$$

A real sequence is called **convergent** if it converges to some real number. Otherwise, it is called **divergent**.

Note that the value of N depends in general on the choice of ε. The smaller we choose ε, the larger N may have to be.

In terms of ε-neighborhoods, convergence of a real sequence

$$(a_n) \to a$$

states that, for every $\varepsilon > 0$, there exists an $N \in \mathbb{N}$ such that

$$a_n \in U_\varepsilon(a), \quad \text{for all } n \geq N.$$

In other words, every ε-neighborhood of a contains *all but finitely many elements* of (a_n), in particular, all except for possibly $a_0, \ldots, a_{N-1}$. In this case, one also says that the sequence (a_n) is **eventually** in $U_\varepsilon(a)$ or

$$a_n \in U_\varepsilon(a)$$

for **sufficiently large** n.

Example.

1. For $c \in \mathbb{R}$, the constant sequence with $a_n = c$ for $n \in \mathbb{N}$, that is, the sequence

   $$(c, c, c, \ldots)$$

   obviously converges to c. To prove this formally, let $\varepsilon > 0$ be arbitrary. Choose $N = 0$. Then, for all $n \geq N$, we have

   $$|a_n - c| = |c - c| = 0 < \varepsilon.$$

   Hence,

   $$(c, c, c, \ldots) \to c.$$

2. The sequence $(1/n)_{n\geq 1}$ converges to 0, that is,

   $$\lim_{n\to\infty} \frac{1}{n} = 0.$$

Let $\varepsilon > 0$ be arbitrary. By Corollary 15 (a consequence of the Archimedean property), there exists an $N \in \mathbb{N}$ such that $1/N < \varepsilon$. Then, for all $n \geq N > 0$, we have

$$0 < \frac{1}{n} \leq \frac{1}{N} < \varepsilon,$$

and hence

$$\left|\frac{1}{n} - 0\right| = \frac{1}{n} < \varepsilon.$$

A sequence that converges to zero is also called a **null sequence**.

3. As an exercise, we prove that

$$\lim_{n\to\infty} \frac{n}{n+1} = 1.$$

4. The sequence $a_n = (-1)^n$ for $n \in \mathbb{N}$, that is, the sequence

$$(1, -1, 1, -1, \ldots)$$

is *divergent*. Intuitively, this is clear since the distance between consecutive elements of the sequence is 2, that is, for all $n \in \mathbb{N}$,

$$|a_{n+1} - a_n| = 2.$$

Hence, the sequence cannot get arbitrarily close to a single real number.

For a formal proof, assume, for contradiction, that (a_n) converges to some real number a. Then, for $\varepsilon = 1$, there exists an $N \in \mathbb{N}$ such that for all $n \geq N$, we have

$$|a_n - a| < 1.$$

For all $n \geq N$, by the triangle inequality, it follows that

$$2 = |a_{n+1} - a_n| = |(a_{n+1} - a) + (a - a_n)| \leq |a_{n+1} - a| + |a - a_n| < 1 + 1 = 2,$$

that is, $2 < 2$, a contradiction.

The definition of convergence contains also the statement that the *limit* of a convergent sequence is *uniquely determined*, which we still have to prove.

Theorem 32. Let (a_n) be a real sequence. If $(a_n) \to a$ and $(a_n) \to b$, then $a = b$.

Proof. We prove $a = b$ by showing that for every $\varepsilon > 0$, we have $|a - b| < \varepsilon$. Let $\varepsilon > 0$. Since $(a_n) \to a$, there exists an $N_1 \in \mathbb{N}$, such that

$$|a_n - a| < \frac{\varepsilon}{2}, \quad \text{for } n \geq N_1.$$

Since $(a_n) \to b$, there exists an $N_2 \in \mathbb{N}$, such that

$$|a_n - b| < \frac{\varepsilon}{2}, \quad \text{for } n \geq N_2.$$

Therefore, we have by the triangle inequality for $n \geq \max(N_1, N_2)$,

$$|a - b| = |a - a_n + a_n - b| \leq |a - a_n| + |a_n - b| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

■.

Another important property of convergent sequences is that they are always *bounded*.

Definition 39. A real sequence (a_n) is **bounded** if its range

$$\{a_n \mid n \in \mathbb{N}\} \subseteq \mathbb{R}$$

is a bounded set, that is, if there exists $M \geq 0$ such that

$$|a_n| \leq M$$

for all $n \in \mathbb{N}$.

Geometrically, this means that we can find an interval $[-M, M]$ that contains all elements of the sequence (a_n), that is,

$$a_n \in [-M, M]$$

for all $n \in \mathbb{N}$.

Theorem 33. Every convergent sequence is bounded.

Proof. Let $(a_n) \to a$ be a convergent real sequence. Then, by choosing $\varepsilon = 1$, we know that there exists an $N \in \mathbb{N}$ such that

$$|a_n - a| < 1, \quad \text{for } n \geq N.$$

Hence, by the triangle inequality, we have

$$|a_n| = |a + (a_n - a)| \leq |a| + |a_n - a| \leq |a| + 1$$

for $n \geq N$. If we let

$$M = \max(|a_0|, \ldots, |a_{N-1}|, |a| + 1),$$

we have

$$|a_n| \leq M, \quad \text{for all } n \in \mathbb{N},$$

so (a_n) is bounded.

The converse of Theorem 33 is false. For example, the sequence

$$(1, -1, 1, -1, \ldots)$$

is bounded with $M = 1$ but it is divergent, as discussed above.

For every $x \in \mathbb{R}$, we can consider the sequence of powers of x,

$$(x^n)_{n \in \mathbb{N}} = (1, x, x^2, x^3, \ldots),$$

which is a sequence that appears in many different contexts. Its behaviour depends on the (absolute) value x.

Theorem 34. Let $x \in \mathbb{R}$.

1. If $|x| < 1$, then
$$\lim_{n \to \infty} x^n = 0.$$
2. If $|x| > 1$, then (x^n) is unbounded.

Proof. If $|x| < 1$, by Corollary 31, there exists an $N \in \mathbb{N}$ such that

$$|x|^N < \varepsilon.$$

Then, for all $n \geq N$, we have

$$|x^n - 0| = |x^n| = |x|^n \leq |x|^N < \varepsilon.$$

Hence, the sequence (x^n) converges to zero.
If $|x| > 1$, it follows from Theorem 30 that (x^n) is unbounded. ∎

For $|x| > 1$, it follows, in particular, that (x^n) is *divergent* since it is *unbounded*. Finally, we consider the two special cases

$$x = 1 \quad \text{and} \quad x = -1.$$

If $x = 1$, then $x^n = 1$ for all $n \in \mathbb{N}$, that is, it is the constant sequence 1 with limit 1. If $x = -1$, then $x^n = (-1)^n$ for all $n \in \mathbb{N}$, which diverges as discussed above.

Recall that we can *add and multiply real sequences elementwise*,

$$(a_n) + (b_n) = (a_n + b_n) \quad \text{and} \quad (a_n)(b_n) = (a_n b_n),$$

which is a special case of the pointwise sum and product of real-valued functions. The next theorem tells us that the *sum and product* of *convergent sequences* is *convergent* and that these operations are *compatible* with *taking limits*.

Theorem 35. Let (a_n) and (b_n) two convergent real sequences. Then, $(a_n + b_n)$ and $(a_n b_n)$ are convergent and we have

$$\lim_{n\to\infty}(a_n + b_n) = \lim_{n\to\infty} a_n + \lim_{n\to\infty} b_n$$

and

$$\lim_{n\to\infty}(a_n b_n) = (\lim_{n\to\infty} a_n)(\lim_{n\to\infty} b_n).$$

Proof. Let

$$a = \lim_{n\to\infty} a_n \quad \text{and} \quad b = \lim_{n\to\infty} b_n.$$

To show that

$$\lim_{n\to\infty}(a_n + b_n) = a + b,$$

we need to make the difference $|a_n + b_n - (a + b)|$ arbitrarily small. By the triangle inequality, we have

$$|a_n + b_n - (a + b)| = |(a_n - a) + (b_n - b)| \leq |a_n - a| + |b_n - b|.$$

Now, let $\varepsilon > 0$. Since $(a_n) \to a$ and $(b_n) \to b$, there exist $N_1, N_2 \in \mathbb{N}$ such that

$$|a_n - a| < \frac{\varepsilon}{2}, \quad \text{for } n \geq N_1 \quad \text{and} \quad |b_n - b| < \frac{\varepsilon}{2}, \quad \text{for } n \geq N_2.$$

If we let $N = \max(N_1, N_2)$, then $n \geq N$ implies that

$$|a_n + b_n - (a + b)| \leq |a_n - a| + |b_n - b| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

Hence, $\lim(a_n + b_n) = a + b$.
To show that

$$\lim_{n\to\infty} (a_n b_n) = ab,$$

we use the inequality

$$\begin{aligned} |a_n b_n - ab| &= |(a_n b_n - a_n b) + (a_n b - ab)| \\ &\leq |a_n b_n - a_n b| + |a_n b - ab| \\ &= |a_n||b_n - b| + |b||a_n - a|. \end{aligned}$$

By Theorem 33, we know that (a_n) is bounded since it is convergent, that is, there exists an $M_1 > 0$ such that

$$|a_n| \leq M_1$$

for all $n \in \mathbb{N}$. For $M = \max(M_1, |b|)$, we obtain the inequality

$$|a_n b_n - ab| \leq M|b_n - b| + M|a_n - a|.$$

Now, let $\varepsilon > 0$. Then, there exist $N_1, N_2 \in \mathbb{N}$ such that

$$|a_n - a| < \frac{\varepsilon}{2M}, \quad \text{for } n \geq N_1 \quad \text{and} \quad |b_n - b| < \frac{\varepsilon}{2M}, \quad \text{for } n \geq N_2.$$

If we let $N = \max(N_1, N_2)$, then $n \geq N$ implies that

$$|a_n b_n - ab| < M\frac{\varepsilon}{2M} + M\frac{\varepsilon}{2M} = \varepsilon.$$

Hence, $\lim a_n b_n = ab$. ∎

We can also multiply a sequence by a real number (a scalar) $\alpha \in \mathbb{R}$ elementwise,

$$\alpha(a_n) = (\alpha a_n).$$

We obtain the same result by multiplying the constant sequence (α) with (a_n) since

$$(\alpha)(a_n) = (\alpha a_n).$$

Hence, we get as a consequence from the previous theorem that also *multiplication by a scalar* is *compatible* with *taking the limit.*

Corollary 36. Let $\alpha \in \mathbb{R}$ and let (a_n) be a convergent real sequence. Then, (αa_n) is convergent and we have

$$\lim_{n\to\infty} (\alpha a_n) = \alpha \lim_{n\to\infty} a_n.$$

Hence, we know in particular that a *linear combination*

$$\alpha(a_n) + \beta(b_n) = (\alpha a_n + \beta b_n)$$

of *convergent sequences* (a_n) and (b_n) with $\alpha, \beta \in \mathbb{R}$ is *convergent* and we have

$$\lim_{n\to\infty} (\alpha a_n + \beta b_n) = \alpha \lim_{n\to\infty} a_n + \beta \lim_{n\to\infty} b_n. \tag{3.1}$$

Hence, taking the limit is *linear*.

As an exercise, we prove that by taking a suitable linear combination it follows that two convergent sequences (a_n) and (b_n) have the *same limit* iff their difference

$$(a_n - b_n)$$

is a *null sequence*.

For a sequence (b_n) with $b_n \neq 0$ for all $n \in \mathbb{N}$, we can also *invert* its elements and consider the sequence

$$(1/b_n).$$

If a sequence (b_n) is convergent and its limit is *nonzero*, then its elements are nonzero for *sufficiently large* n and *inverting* is *compatible* with *taking the limit*.

Theorem 37. Let (b_n) be a convergent sequence with

$$\lim_{n\to\infty} b_n = b \neq 0.$$

Then, there exists an $n_0 \in \mathbb{N}$ such that $b_n \neq 0$ for all $n \geq n_0$, the sequence $(1/b_n)_{n\geq n_0}$ is convergent and we have

$$\lim_{n\to\infty} \left(\frac{1}{b_n}\right)_{n\geq n_0} = \frac{1}{\lim b_n} = \frac{1}{b}.$$

Proof. Since $\lim b_n = b \neq 0$, we know, by choosing $\varepsilon = |b|/2$, that there exists an $n_0 \in \mathbb{N}$ such that

$$|b_n - b| < \frac{|b|}{2}, \quad \text{for } n \geq n_0.$$

Hence, for $n \geq n_0$,

$$|b_n| = |b - (b - b_n)| \geq |b| - |b - b_n| > \frac{|b|}{2}.$$

In particular, we have $b_n \neq 0$ for all $n \geq n_0$.
Now, let $\varepsilon > 0$. Then, there exists an $N_1 \in \mathbb{N}$ such that

$$|b_n - b| < \frac{\varepsilon |b|^2}{2}, \quad \text{for } n \geq N_1.$$

If we let $N = \max(N_1, n_0)$, then $n \geq N$ implies that

$$\left|\frac{1}{b_n} - \frac{1}{b}\right| = \frac{1}{|b_n||b|}|b - b_n| < \frac{2}{|b|^2}\frac{\varepsilon|b|^2}{2} = \varepsilon.$$

Hence, $\lim(1/b_n)_{n\geq n_0} = 1/b$. ∎

For an arbitrary sequence (a_n) and a sequence (b_n) with $b_n \neq 0$ for all $n \in \mathbb{N}$, we can write the *sequence of quotients*

$$(a_n/b_n)$$

as the product

$$(a_n)(1/b_n) = (a_n/b_n).$$

Hence, if (a_n) is convergent and $(b_n) \to b$ with $b \neq 0$, we obtain

$$\lim_{n\to\infty} \frac{a_n}{b_n} = \frac{\lim a_n}{\lim b_n}. \tag{3.2}$$

In summary, *algebraic operations* on sequences *preserve convergence* and are *compatible* with *taking limits*. This is a powerful tool for *proving convergence* and *computing limits*, without the need to directly use the formal definition, by decomposing a sequence into simpler parts for which convergence and limits are known.

Example.

1. For the sequences
$$a_n = \frac{n}{n+1},$$
we have for $n \geq 1$,
$$a_n = \frac{1}{1+1/n}.$$
Since $\lim 1 = 1$ and $\lim 1/n = 0$, we know by Theorem 35 that $\lim(1 + 1/n) = 1$. Hence, by Theorem 37,
$$\lim_{n\to\infty} \frac{1}{1+1/n} = \frac{1}{\lim(1+1/n)} = 1/1 = 1.$$

2. For the sequence
$$a_n = \frac{42n^2 + 3n}{n^2 - 2},$$
we have for $n \geq 1$,
$$a_n = \frac{42 + 3/n}{1 - 2/n^2}.$$
Since $\lim 1/n = 0$, we know by Theorem 35 that $\lim 1/n^2 = 0$. By Equation (3.1), we know
$$\lim(42 + 3/n) = 42 \quad \text{and} \quad \lim(1 - 2/n^2) = 1.$$
Finally, by Equation (3.2), we obtain
$$\lim_{n\to\infty} \frac{42n^2 + 3n}{n^2 - 2} = \frac{\lim(42 + 3/n)}{\lim(1 - 2/n^2)} = \frac{42}{1} = 42.$$

Another useful fact is that the order relation $\leq$ is *preserved* when *taking limits*.

Theorem 38. Let (a_n) be a convergent real sequences. If $a_n \geq 0$ for all $n \in \mathbb{N}$, then $\lim a_n \geq 0$.

Proof. Assume, for contradiction, that $a = \lim a_n < 0$. We know, by choosing $\varepsilon = |a|$, that there exists an $N \in \mathbb{N}$ such that

$$|a_n - a| < |a|, \quad \text{for } n \geq N.$$

In particular, we would have $a_N \in (a - |a|, a + |a|) = (2a, 0)$, since $|a| = -a$. Hence, $a_N < 0$, a contradiction. ∎

Note that *taking limits* does *not preserve strict inequalities*, that is, if $a_n > 0$ for all $n \in \mathbb{N}$, it does not follow that $\lim a_n > 0$.

Example. For $(1/n)_{n\geq 1}$, we have $1/n > 0$ for all $n \geq 1$. However, $\lim 1/n = 0$.

By considering the difference $(b_n - a_n)$ between two convergent sequences (a_n) and (b_n) and constant sequences, we obtain with Theorem 35 the following consequences, which we prove as an exercise.

Corollary 39. Let (a_n) and (b_n) be two convergent real sequences and let $K, M \in \mathbb{R}$.

1. If $a_n \leq b_n$ for all $n \in \mathbb{N}$, then $\lim a_n \leq \lim b_n$.
2. If $a_n \leq M$ for all $n \in \mathbb{N}$, then $\lim a_n \leq M$.
3. If $K \leq a_n$ for all $n \in \mathbb{N}$, then $K \leq \lim a_n$.

In the following, we define when a sequence *tends to infinity/minus infinity*, that is, when it *eventually gets arbitrarily large/small*.

Definition 40. A real sequence (a_n) **tends to infinity** (or **diverges to infinity**) if for every $K \in \mathbb{R}$, there exists an $N \in \mathbb{N}$ such that for all $n \in \mathbb{N}$, $n \geq N$ implies that

$$a_n > K.$$

In this case, we write

$$\lim_{n\to\infty} a_n = \infty, \quad \lim a_n = \infty, \quad \text{or} \quad (a_n) \to \infty.$$

A real sequence (a_n) **tends to minus infinity** (or **diverges to minus infinity**) if $(-a_n)$ tends to infinity. In this case, we write

$$\lim_{n\to\infty} a_n = -\infty, \quad \lim a_n = -\infty, \quad \text{or} \quad (a_n) \to -\infty.$$

Note that a sequence tending (*diverging*) to infinity/minus infinity is in particular *unbounded* and therefore it is *divergent*. However, an unbounded sequence does not necessarily tend to infinity or minus infinity.

Example.

1. $\lim_{n\to\infty} n = \infty$.
2. For a real number $x > 1$, it follows from Theorem 30 that
$$\lim_{n\to\infty} x^n = \infty.$$
3. $\lim_{n\to\infty} -2^n = -\infty$.
4. The sequence
$$((-1)^n n) = (0, -1, 2, -3, 4, \ldots)$$
is unbounded but does not tend to infinity or minus infinity.

The next two statements show the relationship between *infinite limits* and *zero limits*. If a sequence *tends to infinity or minus infinity*, the sequence of inverses *converges to zero* and if a *positive/negative sequence converges to zero*, the sequence of inverses tends to *infinity/minus infinity*.

Theorem 40. If a real sequence (a_n) tends to infinity or minus infinity, there exists an $n_0 \in \mathbb{N}$ such that $a_n \neq 0$ for all $n \geq n_0$ and we have
$$\lim_{n\to\infty} \frac{1}{a_n} = 0.$$

Proof. Let $\lim a_n = \infty$. We know, by choosing $K = 0$, that there exists an $n_0 \in \mathbb{N}$ such that $a_n > 0$ for $n \geq n_0$. In particular, we have $a_n \neq 0$ for all $n \geq n_0$.
It remains to show that $\lim 1/a_n = 0$. Let $\varepsilon > 0$. Since $\lim a_n = \infty$, there exists an $N \in \mathbb{N}$ such that
$$a_n > 1/\varepsilon, \quad \text{for } n \geq N.$$
Hence, we have $1/a_n < \varepsilon$ for all $n \geq N$.
The case $\lim a_n = -\infty$ follows by considering the sequence $(-a_n)$ with $\lim(-a_n) = \infty$. ∎

Similarly, one can also prove the statement about null sequences.

Theorem 41. If (a_n) is a null sequence with
$$a_n > 0, \quad \text{for } n \in \mathbb{N}$$
(respectively, $a_n < 0$ for $n \in \mathbb{N}$), then
$$\lim_{n\to\infty} \frac{1}{a_n} = \infty$$
(respectively, $\lim 1/a_n = -\infty$).

Closely related to the notion of a sequence is that of an *(infinite) series*. Given a real sequence (a_n), we can informally think of series as an *infinite sum*
$$a_0 + a_1 + a_2 + \cdots$$
that is obtained by adding up the elements of the sequence (a_n). More formally, we have the following definition.

Definition 41. Let (a_n) be a real sequence. The sequence (s_n) of **partial sums** given by

$$s_n = \sum_{k=0}^{n} a_k = a_0 + a_1 + \cdots + a_n$$

is called a **series** with **terms** a_n and is denoted by

$$\sum_{k=0}^{\infty} a_k.$$

A series is called **convergent** if the sequence of partial sums is convergent. In this case, we write

$$\sum_{k=0}^{\infty} a_k = \lim_{n\to\infty} s_n$$

and call the limit of the partial sums the **sum** of the series. A series that is not convergent is called **divergent**.

Observe that the notation

$$\sum_{k=0}^{\infty} a_k$$

is used both for the *series itself* and for *its sum* if it is convergent.

We also note that the sequence (a_n) of terms of a series $\sum_{n=0}^{\infty} a_k$ can be obtained from the partial sums (s_n) by the identities

$$a_0 = s_0 \quad \text{and} \quad a_n = s_n - s_{n-1}, \quad \text{for } n \geq 1. \tag{3.3}$$

From this point of view, series are just another way to look at sequences and any statement about series can be formulated in terms of sequences and vice versa. However, in many situations series arise naturally and statements about series are often more transparent than their reformulations in terms of sequences of partial sums. A prominent example is the **geometric series** that one obtains by summing up powers of a real number.

Theorem 42. Let $x \in \mathbb{R}$. If $|x| < 1$, the series $\sum_{k=0}^{\infty} x^k$ converges and we have

$$\sum_{k=0}^{\infty} x^k = \frac{1}{1-x}.$$

Proof. We consider the nth partial sum

$$s_n = \sum_{k=0}^{n} x^k$$

and note that

$$s_n(1-x) = \sum_{k=0}^{n} x^k - \sum_{k=0}^{n} x^{k+1} = \sum_{k=0}^{n} x^k - \sum_{k=1}^{n+1} x^k = 1 - x^{n+1}.$$

Since $x \neq 1$, this yields

$$s_n = \frac{1 - x^{n+1}}{1 - x}.$$

Since $|x| < 1$, we know by Theorem 34 that $\lim x^n = \lim x^{n+1} = 0$. Hence,

$$\sum_{k=0}^{\infty} x^k = \lim_{n\to\infty} s_n = \lim_{n\to\infty} \frac{1 - x^{n+1}}{1 - x} = \frac{1 - \lim x^{n+1}}{1 - x} = \frac{1}{1 - x}.$$

■

Example.

1. For $x = 1/2$, we have

$$\sum_{k=0}^{\infty} \frac{1}{2^k} = 1 + \frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \frac{1}{16} \cdots = \frac{1}{1 - 1/2} = 2.$$

2. For $x = -1/2$, we have

$$\sum_{k=0}^{\infty} \frac{1}{(-2)^k} = 1 - \frac{1}{2} + \frac{1}{4} - \frac{1}{8} + \frac{1}{16} \mp \cdots = \frac{1}{1 + 1/2} = \frac{2}{3}.$$

If $|x| \geq 1$, then the geometric series

$$\sum_{k=0}^{\infty} x^k$$

is *divergent*. This follows from the next theorem stating that a *necessary condition* for a *series* to be *convergent* is that the *sequence of its terms is a null sequence*.

Theorem 43. If a series $\sum_{k=0}^{\infty} a_k$ is convergent, then $\lim a_n = 0$.

Proof. By definition, $\sum_{k=0}^{\infty} a_k$ converges if the sequence of partial sums (s_n) converges. Let $(s_n) \to s$. By Equation (3.3), we have $a_n = s_n - s_{n-1}$ for $n \geq 1$. By Theorem 35, it follows that (a_n) is convergent and

$$\lim a_n = \lim s_n - \lim s_{n-1} = s - s = 0.$$

■

An important example of a *divergent series* whose terms converge to zero is the **harmonic series**

$$\sum_{n=1}^{\infty} \frac{1}{n}.$$

In fact, the sequence of *partial sums tend to infinity*. For showing this, we consider the partial sums of the first 2^m terms of the series for $m \geq 1$. We have

$$\begin{aligned} s_{2^m} = \sum_{n=1}^{2^m} \frac{1}{n} &= 1 + \frac{1}{2} + \sum_{i=1}^{m-1} \left(\sum_{n=2^i+1}^{2^{i+1}} \frac{1}{n} \right) \\ &= 1 + \frac{1}{2} + \left(\frac{1}{3} + \frac{1}{4} \right) \\ &\quad + \left(\frac{1}{5} + \frac{1}{6} + \frac{1}{7} + \frac{1}{8} \right) + \cdots + \left(\frac{1}{2^{m-1} + 1} + \cdots + \frac{1}{2^m} \right). \end{aligned}$$

Since

$$\left(\frac{1}{3}+\frac{1}{4}\right) \geq 2\,\frac{1}{4}, \quad \left(\frac{1}{5}+\frac{1}{6}+\frac{1}{7}+\frac{1}{8}\right) \geq 4\,\frac{1}{8}, \quad \ldots, \quad \left(\frac{1}{2^{m-1}+1}+\cdots+\frac{1}{2^m}\right) \geq 2^{m-1}\,\frac{1}{2^m},$$

the sum in each parenthesis is greater than $1/2$, and we have the following estimate

$$s_{2^m} \geq 1+\frac{m}{2}.$$

Hence,

$$\sum_{n=1}^{\infty}\frac{1}{n} = \lim_{n\to\infty} s_n = \infty.$$

We note, however, that the harmonic series diverges very slowly to infinity.

Recall form Equation (3.1) that linear combinations of convergent sequences are convergent. Using this fact for sequences of partial sums of series, we obtain the following statement.

Corollary 44. Let $\sum_{k=0}^{\infty} a_k$ and $\sum_{k=0}^{\infty} b_k$ be two convergent series and $\alpha, \beta \in \mathbb{R}$. Then, $\sum_{k=0}^{\infty}(\alpha a_k + \beta b_k)$ is a convergent series and we have

$$\sum_{k=0}^{\infty}(\alpha a_k + \beta b_k) = \alpha \sum_{k=0}^{\infty} a_k + \beta \sum_{k=0}^{\infty} b_k.$$

So far, we have seen several results that allow us to show that a sequence or series converges. However, all of these results require that we already *know the limit* of the given sequence or of its building blocks. In the following, we discuss criteria that enable us to *prove convergence* of real sequences or series *without knowing the limit* in advance. In fact, many quantities in analysis are defined in terms of a limit of a sequence or series.

First, we discuss *monotone sequences*, which appear in many contexts.

Definition 42. A real sequence (a_n) is **increasing** if $a_n \leq a_{n+1}$ for all $n \in \mathbb{N}$ and **decreasing** if $a_n \geq a_{n+1}$ for all $n \in \mathbb{N}$. A sequence is **monotone** if it is either increasing or decreasing.

The terms *nondecreasing* and *nonincreasing* are also used instead of *increasing* and *decreasing*.

Example.

1. (n) is increasing and $(-n)$ is decreasing, respectively, and both sequences are unbounded.
2. $(n/(n+1))$ is increasing and $(1/n)_{n\geq 1}$ is decreasing, respectively, and both sequences are bounded.
3. A sequence is both increasing and decreasing iff it is constant.

The *Monotone Convergence Theorem* states that for proving convergence of a monotone sequence, we only have to show that it is bounded. In particular, a *bounded increasing* sequence converges to the *supremum* and a *bounded decreasing* sequence to the infimum of its elements, respectively.

Theorem 45. A monotone real sequence is convergent iff it is bounded.
In particular, let (a_n) be a bounded monotone real sequence. If (a_n) is increasing,

$$\lim_{n\to\infty} a_n = \sup\{a_n \mid n \in \mathbb{N}\}$$

and if (a_n) is decreasing,

$$\lim_{n\to\infty} a_n = \inf\{a_n \mid n \in \mathbb{N}\}.$$

Proof. A convergent sequence is necessarily bounded by Theorem 33.
For the converse implication, let (a_n) be a bounded increasing real sequence. By definition (a_n) is bounded if its range

$$A = \{a_n \mid n \in \mathbb{N}\} \subseteq \mathbb{R}$$

is bounded and by the Completeness axiom of the real numbers its supremum $a = \sup A$ exists. We show that

$$\lim a_n = a.$$

Let $\varepsilon > 0$. Then, $a - \varepsilon < a$ is not an upper bound for A and hence there exists an $N \in \mathbb{N}$ such that $a - \varepsilon < a_N$. Since (a_n) is increasing and a is an upper bound for A, we have

$$a - \varepsilon < a_N \leq a_n \leq a, \quad \text{for } n \geq N.$$

In particular, $|a_n - a| \leq \varepsilon$ for all $n \geq N$. Hence, $\lim a_n = a$.
Similarly, one shows that a bounded decreasing real sequence (a_n) converges to $\inf A$. ∎

As an application, we prove that every positive real number $a > 0$ has a unique **positive kth root** for every integer $k \geq 1$, that is, there exists a unique positive real number $r > 0$ such that

$$r^k = a.$$

In this case, one writes

$$r = \sqrt[k]{a}.$$

Uniqueness of a positive kth root follows, since $0 < s < r$ implies $0 < s^k < r^k$.

To prove the *existence* of a positive kth root for a given $a > 0$, we consider the real sequence (x_n) defined by the recurrence

$$x_{n+1} = \frac{1}{k}\left((k-1)x_n + \frac{a}{x_n^{k-1}}\right) = x_n\left(1 + \frac{a - x_n^k}{kx_n^k}\right) \tag{3.4}$$

for $n \in \mathbb{N}$ with $x_0 = a + 1$. Note that in an exercise at the beginning of the lecture, we considered this recursion for $k = 2$ and $a = 2$.

One can prove, by induction, using Bernoulli's inequality that for all $n \in \mathbb{N}$,

$$x_n > 0, \quad x_n > x_{n+1}, \quad x_n^k > a.$$

In particular, this shows that (x_n) is a decreasing sequence bounded from below by 0. By Theorem 45, it converges and we can use the recursion (3.4) to determine the limit

$$x = \lim x_n.$$

We have

$$kx_{n+1}x_n^{k-1} = (k-1)x_n^k + a.$$

By Theorem 35, we can take the limit on both sides and obtain

$$k \lim x_{n+1} (\lim x_n)^{k-1} = (k-1)(\lim x_n)^k + a.$$

Hence,

$$kx^k = (k-1)x^k + a,$$

Therefore, $x^k = a$, that is, $x = \sqrt[k]{a}$.

Note that the recursion (3.4) gives us also an iterative method to approximate the positive kth root of an $a > 0$. We will learn later in the lecture how find this recursion as a special case of *Newton's method*, which is a very efficient method to approximate roots of a differentiable real function.

List of Symbols

$f\colon X \to Y$	function f from X to Y	3
$f\colon x \mapsto y$	f maps x to y	3
id_X	identity function on a set X	4
$\mathcal{P}(X)$	power set of a set X	4
Y^X	set of all functions from X to Y	4
$\lvert x \rvert$	absolute value of $x \in \mathbb{R}$	5
$\operatorname{sgn}(x)$	sign of $x \in \mathbb{R}$	6
$(a)_{n\in\mathbb{N}}$	sequence $(a_0, a_1, a_2, \ldots)$	6
$(a_n)_{n\geq 0}$	sequence $(a_0, a_1, a_2, \ldots)$	6
$(a_n)_{n\geq 1}$	sequence $(a_1, a_2, a_3, \ldots)$	6
$f\lvert_S$	restriction of f to a set S	10
$f(S)$	image of subset S under f	10
$\operatorname{im} f$	image of f	10
$f^{-1}(W)$	inverse image of subset W under f	11
$\lvert X \rvert$	number of elements in a finite set	11
S_n	set of permutations of $\{1, 2, \ldots, n\}$	12
$n!$	factorial	12
$g \circ f$	composition of functions g and f	13
f^{-1}	inverse of a function f	14
xRy	x is in relation R to y	17
$b \mid a$	b divides a	18
$a \equiv b \mod m$	a is congruent b modulo m	18
$\equiv_m$	congruence relation modulo m	18
$[x]_\sim$	equivalence class of x for $\sim$	19
$[x]$	equivalence class of x	19
$X/\!\sim$	quotient set of X by $\sim$	21
$\mathbb{Z}_m$	$\mathbb{Z}$ modulo m	21
$\operatorname{quo}(a, d)$	quotient of the division of a by d	23
$\operatorname{rem}(a, d)$	remainder of the division of a by d	23
0	zero element (additive identity) of a ring	25
1	unit element (multiplicative identity) of a ring	25
a^{-1}	multiplicative inverse of an invertible element a	27

$\sup A$	supremum (least upper bound) of a set A	33
$\inf A$	infimum (greatest lower bound) of a set A	33
A^{-1}	inverse matrix of an invertible matrix A	47
$\det(A)$	determinant of a matrix A	47
$\mathrm{adj}(A)$	adjugate matrix of a matrix A	47
$v \cdot w$	dot product of vectors v and w	51
$F^{m\times n}$	set of $m \times n$ matrices over a field F	52
h_A	linear map associated to a matrix A	54
I_n	$n \times n$ identity matrix	54
M_h	transformation matrix of a linear map h	55
A^{-1}	inverse matrix of an invertible matrix A	60
$A^\top$	transpose of a matrix A	61
$v^\top$	transpose of a vector v	62
$U_\varepsilon(a)$	ε-neighborhood of a	65
$\lim\limits_{n\to\infty} a_n = a$	limit of a sequence (a_n) is a	66
$(a_n) \to a$	limit of a sequence (a_n) is a	66
$\lim\limits_{n\to\infty} a_n = \infty$	a sequence (a_n) tends to infinity	74
$(a_n) \to \infty$	a sequence (a_n) tends to infinity	74
$\lim\limits_{n\to\infty} a_n = -\infty$	a sequence (a_n) tends to minus infinity	74
$(a_n) \to -\infty$	a sequence (a_n) tends to minus infinity	74
$\sum\limits_{k=0}^{\infty} a_k$	series and its sum if it converges	76
$\sqrt[k]{a}$	positive kth root of $a > 0$	79

Index