Inleiding Analyse in Meer Variabelen

E.P. van den Ban, A. del Pino Gómez

Foreword

These lecture notes were developed by Erik van den Ban in 2020 for the (at the time) newly created course *Inleiding Analyse in Meer Variabelen*. They were originally written in Dutch and were designed to follow naturally the lecture notes of the course *Inleiding Analyse* and to connect seamlessly to the contents of the course *Functies en Reeksen*.

The text has since been updated by Álvaro del Pino. Even though the overall structure and learning objectives have not changed, the current version (2025) was designed to address the comments brought up in student evaluations along the years. Namely:

- Each chapter now includes a greater detail of discussion in order to motivate the material better. More highlight is given to the main results and ideas.
- More examples have been added.
- Two chapters have been added to include necessary background on metric spaces (Chapter 1) and linear algebra (Chapter 2).
- Some of the terminology (Chapters 6 and 7) has been updated and slightly expanded to emphasise the key ideas and connect better to later courses.

We are looking forward to your comments to keep improving the lecture notes.

What will you learn in this course?

The overarching learning objective of the course *Inleiding Analyse in Meer Variabelen* is for you to become familiar with the analysis of multivariate functions $\mathbb{R}^n \to \mathbb{R}^m$.

In Chapter 2 we get familiar with the simplest examples of functions on $\mathbb{R}^n$: the polynomials. These include the linear functions and, as you will see, this chapter is largely about linear algebra. We expect you to be familiar with large parts of it (Sections 2.1, 2.2, 2.3, and some of 2.5) but you will also encounter new material.

In Chapter 1 we discuss metric spaces. The idea is that we should be very comfortable with the continuity of functions before we look into their differentiability. The theory of continuity can be found in Section 1.1; this is material that you have seen before in *Inleiding Analyse*. The rest of the chapter will be new to you: it discusses paths in metric spaces and how these can be used to study the "shape" of a metric space. These ideas play a fundamental role in Chapter 7 and will appear in many later courses, including *Functies en Reeksen*, *Topologie en Meetkunde*, and *Differentieerbare variëteiten*.

The theory of multivariate differentiation is developed in Chapters 3 and 4. The key insight behind it is the concept of *linearisation*: Given a function $f : \mathbb{R}^n \to \mathbb{R}^m$, the total derivative $Df(x)$ is the linear function that best approximates it at the point x. This leads naturally (eventually) to the concept of k-order Taylor polynomial ("the polynomial of order k that best approximates f at x"). Along the way we develop many useful results, including the chain rule, the mean value theorem, and the commutativity of partial derivatives. An important motto to keep in mind is that our arguments are usually a combination of linear algebra and using results from analysis in one variable.

In Chapter 6 we encounter the inverse function theorem. This is a major result that comes with a crucial message (that appears in many forms across all of Mathematics): if $U, V \subset \mathbb{R}^n$ are two opens related by a change of coordinates $\phi : U \to V$, we can think of U and V as two incarnations of the same object, up a change of perspective. This can be very useful, since the change of coordinates ϕ can make some things simpler in V compared to U (and viceversa). The inverse function theorem will allow us to construct such coordinate changes. This tool will allow us to initiate the theory of submanifolds $N \subset \mathbb{R}^n$; these can be thought as "the nicest closed subsets of $\mathbb{R}^n$". Our main application will be in constrained optimisation: we will develop the method of Lagrange multipliers, which will allow us to find the extrema of a function along a submanifold. (Sub)manifolds will reappear in *Inleiding Topologie*, *Analyse in meer variabelen* and *Differentieerbare variëteiten*.

This wraps up the theory of differentiation. We will also look into the theory of integration in multiple variables. The integration of continuous functions over hypercubes of $\mathbb{R}^n$ is developed in Chapter 5. This sets the foundations for the theory of the multivariate Riemann integral. You will encounter a more complete treatment in *Analyse in meer variabelen*, *Differentieerbare variëteiten*, and *Maat en integratie*.

It turns out that in higher dimensions one can consider other types of integrals. In Chapter 7 we will develop the integration of covector fields along paths. This will allow us to take a first look at the extremely deep and surprising relation between Analysis and Topology. Namely, we will be able to relate the shape of an open $U \subset \mathbb{R}^n$ to the analytic properties of its covector fields. This idea is central in Mathematics. In the follow-up course *Functies en Reeksen* it will be crucial in the study of holomorphic functions, leading to Cauchy's integral theorem. In Physics you will encounter it in the study of fields and their potentials. In *Analyse in meer variabelen* and *Differentieerbare variëteiten* it will be generalised to higher differential forms, eventually leading to the construction of de Rham cohomology, an important invariant of manifolds.

The lecture notes also include two "extra" chapters. These are not meant to be covered in the course. They are meant to serve as a reference for later courses. Chapter 8 develops some of the basic theory of series, which you will see in detail in *Functies en Reeksen*. Chapter 9 formalises the theory of improper integrals, which you have seen informally already.

Throughout the lecture notes we will use blue boxes to highlight important examples and computations.

We will use red boxes to highlight the central results of the course.

Notation

As in the lecture notes *Inleiding Analyse*, we use the following notation, which may differ from the notation used in the course *Bewijzen in de Wiskunde*.

We write $\mathbb{N} = \{0, 1, 2, \ldots\}$; thus we also consider 0 as a natural number. Furthermore, we write $\mathbb{N}^*$ or $\mathbb{Z}_+$ for $\mathbb{N} \setminus \{0\}$.

If A and B are sets, then $A \subset B$ means that every element of A also belongs to B. In *Bewijzen in de Wiskunde*, the notation $A \subseteq B$ was customary. Moreover, we use the notation $A \subsetneq B$ to express $A \subset B$ and $A \neq B$.

For two points $a, b \in \mathbb{R}^n$, we define the *closed line segment* $[a, b]$ in $\mathbb{R}^n$ with endpoints a and b as the set of points

$$[a, b] := \{a + t(b - a) \ | \ 0 \leq t \leq 1\}.$$

Such an interval consists of a single point if $a = b$. Similarly, we write

$$(a, b) := [a, b] \setminus \{a, b\} = \{a + t(b - a) \mid 0 < t < 1\}$$

for the *open line segment*. Do observe that these intervals are not meant to be oriented, so there is no preferred start or end point. In particular $[a, b] = [b, a]$. We use these definitions still in the one-dimensional case $n = 1$.

Finally, we also use the notation $a := b$ to indicate that "a is *defined* to be equal to b". This differs from $=$, which is an equality that we establish via computation, not definitionally.

Inhoudsopgave

1 Metric spaces

The goal of this course is to get familiar with the theory of differentiation and integration of functions $\mathbb{R}^n \to \mathbb{R}^m$. As such, we will mostly work with functions that are differentiable at least once. However, as we go along, we will often face the following situation: before we show a function is differentiable, we will need to prove that it is continuous.

Continuity is thus an important theme for us. It turns out that continuity is best discussed in the more general setting of metric spaces, which is what this chapter is about. In this chapter we will:

- Recall the sum/product/composition rules for continuity.
- Discuss the idea of uniform continuity. It will be important in Chapter 5.
- Develop the theory of paths/loops and their homotopies.

The first two items were worked out in detail in the course *Inleiding Analyse*, so we will just recall some key facts. Regarding the third item: the main idea is that we are trying to understand a bit more about the *shape* of metric spaces. To do so, we "test" our metric space using paths and loops. You will learn much more about this in *Inleiding Topologie* and *Topologie en Meetkunde*. Our main motivation in this course is the following: the shape of a domain $U \subset \mathbb{R}^n$ has important consequences regarding the properties of its functions (Chapter 3) and covector fields (Chapter 7).

1.1 Continuity

The contents of this section appeared already in *Inleiding Analyse*.

Definition 1.1. A **metric space** is pair consisting of a set X and a distance function $d : X \times X \to [0, \infty)$ satisfying

- $d(x, y) = 0$ if and only if $x = y$.
- Symmetry: $d(x, y) = d(y, x)$.
- Triangle inequality: $d(x, z) \leq d(x, y) + d(y, z)$. △

Using the distance we can define the **open ball** of center $x \in X$ and radius $r > 0$:

$$B(x, r) := \{x' \in X \mid d(x', x) < r\}$$

and the corresponding **closed ball**:

$$\bar{B}(x, r) := \{x' \in X \mid d(x', x) \leq r\}.$$

More generally, we will also use the notation:

Definition 1.2. Let X be a metric space. A subset $U \subset X$ is **open** if for each $x \in U$ there is $\delta > 0$ so that $B(x, \delta) \subset U$. A **neighbourhood** of a point $x \in X$ is an open containing it. A subset $K \subset X$ is **closed** if its complement $K^c := X \setminus K$ is open. △

More generally, a point $a \in V \subset X$ is internal if there is some $\delta > 0$ such that the open ball $B(a, \delta) \subset X$ is contained in V. The collection of all such points of V is called the interior of V and is denoted by $\mathrm{inw}(V)$. The subset V is open if the equality $V = \mathrm{inw}(V)$ holds.

1.1.1 Continuity

Distance functions also allow us to discuss continuity. Namely:

Definition 1.3. A function $f : (X, d_X) \to (Y, d_Y)$ between metric spaces is **continuous** at a point $x \in X$ if

$$\lim_{x' \to x} d_Y(f(x'), f(x)) = 0.$$

A function is continuous if it is continuous at all points. △

This definition can be spelled out in various equivalent ways. Given $x \in X$...

- and $\varepsilon > 0$, there is $\delta > 0$ such that for all $x' \in X$ it holds that
$$d_X(x, x') < \delta \quad \Longrightarrow \quad d_Y(f(x), f(x')) < \varepsilon.$$
- and $\varepsilon > 0$, there is $\delta > 0$ such that $f(B(x, \delta)) \subset B(f(x), \varepsilon)$.
- and a neighbourhood $V \subset Y$ of $f(x)$ there is $\delta > 0$ such that $f(B(x, \delta)) \subset V$.
- and a neighbourhood $V \subset Y$ of $f(x)$ there is a neighbourhood U of x such that $f(U) \subset V$.

Example 1.4. In this course we will focus on the metric spaces $(\mathbb{R}^n, d_{\text{std}})$, where d_{std} is the usual Euclidean distance

$$d_{\text{std}}(x, y) := ||x - y|| = \left(\sum_{i=1}^{n} (x_i - y_i)^2\right)^{1/2}.$$

An important remark is that very often we will consider open subspaces $U \subset \mathbb{R}^n$, which are metric spaces as well once we restrict d_{std}. For simplicity we will henceforth write $\mathbb{R}^n$ and implicitly assume it is endowed with d_{std}. △

The following statements will be our main tools to check whether a map is continuous:

Proposition 1.5. *Fix a metric space (X, d_X). Suppose $f, g : X \to \mathbb{R}^n$ and $h : X \to \mathbb{R}$ are continuous. Then:*

- *Sum rule: $f + g : X \to \mathbb{R}^n$ is also continuous.*
- *Product rule: $hf : X \to \mathbb{R}^n$ is also continuous.*
- *Quotient rule: If h is nowhere zero then $1/h : X \to \mathbb{R} \setminus \{0\}$ is also continuous.*

Suppose moreover that we have metric spaces (Y, d_Y) and (Z, d_Z) and continuous functions $\phi : X \to Y$ and $\psi : Y \to Z$. Then:

- *Composition rule: $\psi \circ \phi : X \to Z$ is also continuous.*

If you want to work out the proof for yourself, you may want to establish the following lemma first:

Lemma 1.6. *Let (X, d) be a metric space. A function $f = (f_1, \cdots, f_n) : X \to \mathbb{R}^n$ is continuous if and only if each entry f_n is continuous.*

In particular, given two continuous functions $g, h : X \to \mathbb{R}$ it follows that $(g, h) : X \to \mathbb{R}^2$ is also continuous.

An important result, which will use frequently, is that:

Lemma 1.7. *Consider metric spaces (X, d_X) and (Y, d_Y), a continuous map $f : X \to Y$, and closed subset $Z \subset Y$. Then the preimage $f^{-1}(Z)$ is a closed subset of X.*

A particularly relevant example is the case of the **level set** $f^{-1}(y)$ associated to a point $y \in Y$. The study of level sets of differentiable functions is an important topic, which will be one of the focuses of Chapter 6.

1.1.2 Uniform continuity

Suppose $f : (X, d_X) \to (Y, d_Y)$ is a continuous function. What this means is that, for a given $x \in X$, if we want to land ε-close to its image $f(x)$, we should pick points $x' \in X$ that are δ-close to x. That is, the quantity δ not only depends on ε, it also depends on x. We can therefore denote it as $\delta(\varepsilon, x)$ and think of it as a function $(0, \infty) \times X \to (0, \infty)$.

We can ask ourselves the following question: "Fix $\varepsilon > 0$. Is there a $\delta > 0$ that works for this ε and all x?" This is the same as asking whether we can choose the function $\delta(\varepsilon, x)$ to be independent of x. This is not always the case:

Example 1.8. Consider the function $f : \mathbb{R} \to \mathbb{R}$ given by $f(x) = x^2$. Since $f'(x) = 2x$, the function f has larger slope as x grows; i.e. it is becoming steeper as $x \to \infty$. This means that $\delta(\varepsilon, x)$ will get smaller the larger x is. Let us work it out. We want $|f(x + h) - f(x)| < \varepsilon$ and we have to find δ so that $|(x + h) - x| = |h| < \delta$ implies it. We bound:

$$|f(x+h) - f(x)| = |(x+h)^2 - x^2| = |2xh + h^2| \leq 2|x||h| + |h|^2 \leq 2|x|\delta + \delta^2,$$

which tells us that we want both terms smaller than $\varepsilon/2$. Bounding the second implies $\delta < \sqrt{\varepsilon/2}$, which does not depend on x. However, bounding the first one implies $\delta < \varepsilon/(4|x|)$, which indeed goes to zero as $|x|$ grows. △

The reason why we couldn't find a δ that worked for all x at once (for a given ε) was because f was becoming more and more steep as we went to infinity in $\mathbb{R}$. However, what if we focus on metric spaces that "don't have an infinity", so functions cannot become arbitrarily steep?

Definition 1.9. A metric space (X, d) is **(sequentially) compact** if every sequence $\{x_i\}_{i=0}^{\infty}$ in X has a convergent subsequence. △

You should imagine this as saying that X has no infinity, since sequences cannot escape. The most important examples are given by the *Bolzano-Weierstrass theorem*:

Proposition 1.10. *Suppose $X \subset \mathbb{R}^n$ is a closed and bounded subset. Then X is a sequentially compact metric space (with the standard Euclidean distance).*

Finally, we define:

Definition 1.11. A function $f : (X, d_X) \to (Y, d_Y)$ between metric spaces is **uniformly continuous** if for every $\varepsilon > 0$ there is some $\delta > 0$, depending only on ε, such that for all $x, x' \in X$:

$$d_X(x', x) < \delta \implies d_Y(f(x'), f(x)) < \varepsilon. \qquad \triangle$$

In Chapter 5 we will need the following variation. The idea is that sometimes we are working with functions f defined over a non-compact metric space (for instance, $\mathbb{R}^n$) but we focus on the behaviour of f close to a subspace A that is compact:

Definition 1.12. Fix metric spaces (X, d_X) and (Y, d_Y), as well as a subspace $A \subset X$, and a function $f : X \to Y$. We will say that f is **uniformly continuous along** A if for every $\varepsilon > 0$ there is some $\delta > 0$, depending only on ε, such that for all $x \in X$ and all $a \in A$ we have:

$$d_X(x, a) < \delta \implies d_Y(f(x), f(a)) < \varepsilon. \qquad \triangle$$

Which allows us to establish:

Proposition 1.13. *Consider metric spaces (X, d_x) and (Y, d_Y), a sequentially compact subspace $A \subset X$, and a continuous function $f : X \to Y$. Then f is uniformly continuous along A.*

Proof. You have seen this statement in *Analyse* for the special case $A = X$. The proof of the general case is very similar.

We assume that the conclusion does not hold and will show that this leads to a contradiction. The negation of the conclusion gives that there exists $\varepsilon > 0$ such that for every $\delta = 1/j > 0$ there exists $a_j \in A$ and $x_j \in X$ with $d_X(x_j, a_j) < 1/j$ which do not satisfy the estimate $d_Y(f(x_j), f(a_j)) < \varepsilon$, i.e., for which $d_Y(f(x_j), f(a_j)) \geq \varepsilon$.

We have thus a sequence $\{a_j\}_{j=1}^{\infty}$ in A and another sequence $\{x_j\}_{j=1}^{\infty}$ in X. Since A is compact, we deduce that $\{a_j\}_{j=1}^{\infty}$ has a convergent subsequence $\{a_{j_i}\}_{i=1}^{\infty}$ with limit $a_\infty \in A$. This means that $\lim_{i\to\infty} d_X(a_{j_i}, a_\infty) = 0$. Using the triangular inequality we deduce:

$$d_X(x_{j_i}, a_\infty) \leq d_X(x_{j_i}, a_{j_i}) + d_X(a_{j_i}, a_\infty) \leq 1/j_i + d_X(a_{j_i}, a_\infty) \to 0$$

as $i \to \infty$. We can also use the reverse triangular inequality and the continuity of f to show:

$$d_Y(f(x_{j_i}), f(a_\infty)) \geq d_Y(f(x_{j_i}), f(a_{j_i})) - d_Y(f(a_{j_i}), f(a_\infty)) \geq \varepsilon - d_Y(f(a_{j_i}), f(a_\infty)) \to \varepsilon$$

as $i \to \infty$. However, this is a contradiction with the continuity of f: we have found a sequence of points $\{x_{j_i}\}$ in X approaching a_∞, but whose values $f(x_{j_i})$ do not approach $f(a_\infty)$. $\square$

Example 1.14. We saw that $f : \mathbb{R} \to \mathbb{R}$ given by $f(x) = x^2$ is not uniformly continuous. However, we can consider the subspace $[0, 1] \subset \mathbb{R}$, which is bounded and closed and therefore compact by Proposition 1.10. According to Proposition 1.13, given ε we can find a δ that works for all $x \in [0, 1]$. Indeed, we computed that we needed $\delta < \varepsilon/(4|x|)$, whose minimum (for $x \in [0, 1]$) is $\varepsilon/4$. $\triangle$

1.1.3 Neighbourhoods of uniform radius

In later chapters we will need the following technical result:

Lemma 1.15. *Let (X, d) be a metric space, and let $U \subset X$ be open. Let $A \subset U$. If A is sequentially compact, then there exists a $\delta > 0$ such that for all $a \in A$ we have $B(a; \delta) \subset U$.*

You should think of this as yet another example of uniform continuity. Given a point a inside an open U we can always find some δ, depending on a, so the δ-ball centered at a is contained in U. The point is whether this δ can be chosen uniformly (i.e independently of a as the point a ranges over the open subset A).

Proof. We give a proof by contradiction. Suppose that no such $\delta > 0$ exists. Then, for every integer $n \geq 1$, there exists $a_n \in A$ such that the ball $B(a_n; 1/n)$ is not contained in U, i.e., it contains a point $y_n \in X \setminus U$. By construction, $d(a_n, y_n) < 1/n$ for all n.

By the sequential compactness of A, there exists a subsequence a_{n_j} with a limit $a \in A$. From $d(y_{n_j}, a_{n_j}) < \frac{1}{n_j}$ and the triangular inequality it follows that:

$$d(y_{n_j}, a) \leq d(y_{n_j}, a_{n_j}) + d(a_{n_j}, a) \to 0$$

as $j \to \infty$. It follows that $y_{n_j} \to a$. Therefore, a is a limit point of the set $X \setminus U$. Since $X \setminus U$ is closed we deduce that $a \in X \setminus U$, which is a contradiction. □

1.2 Paths

In this section we explore what paths and loops tell us about the shape of a metric space. This will be relevant in Chapters 3 and Chapter 7.

Definition 1.16. Suppose $[a, b] \subset \mathbb{R}$ is an interval and X is a metric space. A continuous map $\gamma : [a, b] \to X$ is said to be a **path** or **curve**. The point $\gamma(a) \in X$ is the initial point of γ. We say that $\gamma(b)$ is the endpoint. △

It is convenient to think of $t \in [a, b]$ as the *time variable* of the curve γ. The image $\gamma(t) \in X$ can then be interpreted as the *position* of the trajectory γ at time t. The curve γ gives thus the position as a function of time. We can also say that γ describes a *motion* in the space X.

When $X = \mathbb{R}^n$ we can also consider differentiable paths; see Definition 3.50.

1.2.1 Operations on paths

Given a path, we can run it at a different speed:

Definition 1.17. Let $\gamma : [a, b] \to X$ be a path. A **reparametrisation** of γ is another path $\nu : [c, d] \to X$ satisfying

$$\nu = \gamma \circ \rho,$$

where $\rho : [c, d] \to [a, b]$ is a continuous, increasing bijection. △

You should think of ρ as a change of coordinates telling us that γ and ν are pretty much the same curve, just expressed differently.

We can also run a curve in the opposite direction:

Definition 1.18. Given a continuous curve $\gamma : [0, 1] \to X$ we define its **reverse** to be the curve $\bar{\gamma} : [0, 1] \to X$ defined by the expression $\bar{\gamma}(t) := \gamma(1 - t)$. △

It is immediate that the starting point of γ is the final point of $\bar{\gamma}$ and viceversa.

If a curve finishes at a point $x \in X$ and another curve starts at x, we can run one first and then the other:

Definition 1.19. Suppose $\gamma_0, \gamma_1 : [0,1] \to X$ are continuous curves. We say that they are concatenable if $\gamma_0(1) = \gamma_1(0)$. We then define their **concatenation** $\gamma_1 \cdot \gamma_0 : [0,1] \to X$ using the expression

$$\gamma_1 \cdot \gamma_0(t) = \begin{cases} \gamma_0(2t) & \text{als} \quad t \in [0, \frac{1}{2}], \\ \gamma_1(2t-1) & \text{als} \quad t \in [\frac{1}{2}, 1]. \end{cases}$$ △

The reader should verify that the resulting curve is indeed continuous.

1.2.2 Path-connectedness

You may recall the following definition from *Inleiding Analyse*:

Definition 1.20. A metric space X is **path-connected** if, given any two points $p, q \in X$, we can find a continuous path $\gamma : [0,1] \to X$ with initial point $\gamma(0) = p$ and endpoint $\gamma(1) = q$. △

When studying the shape of X, being path-connected is one of the first properties you could check. A main example is the following:

Lemma 1.21. *$\mathbb{R}^n$ is path-connected.*

Proof. Given points $x, y \in \mathbb{R}^n$ we can consider the straight line segment $[x, y] \subset \mathbb{R}^n$. It can be expressed as the image of the curve $\gamma : [0,1] \to \mathbb{R}^n$ given by $\gamma(t) = x + t(y - x)$. For completeness, let us spell out why it is continuous. Its ith component is $\gamma_i(t) = x_i + t(y_i - x_i)$, which is a linear function $[0,1] \to \mathbb{R}$ so, as seen in *Inleiding Analyse*, γ_i is continuous. The result then follows from Lemma 1.6. □

The exact same proof shows:

Lemma 1.22. *Suppose $U \subset \mathbb{R}^n$ is **convex**. I.e. for any two points $x, y \in U$ it holds that the straight segment $[x, y]$ is in U. $\mathbb{R}^n$ is path-connected. Then U is path-connected.*

In order to tackle more complicated examples, we will use the following auxiliary lemma, which follows from the constructions of paths we introduced earlier:

Proposition 1.23. *Let X be a metric space. Then the following properties hold for any points $x, y, z \in X$:*

- *Reflexivity: There is a path from x to x.*
- *Symmetry: Given a path from x to y, there is a path from y to x.*
- *Transivity: Given a path from x to y and a path from y to z, there is a path from x to z.*

I.e. being connected by a path is an equivalence relation.

Proof. The constant path $\gamma(t) = x$ goes from x to x. If γ goes from x to y then the reverse path $\bar{\gamma}$ (Definition 1.18) goes from y to x. Lastly, if γ goes from x to y and ν from y to z, their concatenation $\nu \cdot \gamma$ (Definition 1.19) goes from x to z. □

Notation 1.24. We will write $\pi_0(X)$ for the set of equivalence classes described by Proposition 1.23. Equivalently, $\pi_0(X)$ is the number of path-connected pieces into which you can decompose X. △

Proposition 1.23 can be used to address the following more involved example:

Lemma 1.25. *Suppose $U \subset \mathbb{R}^n$ is **star-shaped**. I.e. there exists a point $x_0 \in U$ such that for every $x \in U$ it holds that the straight segment $[x, x_0]$ is in U. Then U is path-connected.*

Proof. By assumption, there is a path from x_0 to any other x. Using transitivity (Proposition 1.23) we deduce that there are paths between any two points (namely, by passing via x_0). □

Exercise 1.26. Consider a finite collection of points $K \subset \mathbb{R}^2$. Show that $\mathbb{R}^2 \setminus K$ is path-connected. △

In $\mathbb{R}$ it is possible for us to fully characterise which subsets are path-connected:

Lemma 1.27. *Let $V \subset \mathbb{R}$ be a subset. The following statements are equivalent:*

(a) *V is path-connected;*

(b) *V is an interval.*

Proof. We recall that an interval in $\mathbb{R}$ can be characterized as a subset $I \subset \mathbb{R}$ with the property that for every pair of points $a, b \in I$ with $a < b$ we have $[a, b] \subset I$. In particular, a single point is considered to be an interval.

Assume (a) holds. Let $p, q \in V$ with $p < q$. Then there exists a curve $\gamma : [0, 1] \to V$ with $\gamma(0) = p$ and $\gamma(1) = q$. By the intermediate value theorem for continuous functions, $\gamma([0, 1]) \supset [p, q]$, hence $[p, q] \subset V$. From this we see that V is an interval.

Now assume (b) holds. Then for every pair of points $p, q \in V$ with $p < q$, the segment $[p, q]$ is contained in V. The mapping $\gamma : [0, 1] \to [p, q], t \mapsto p + t(q - p)$ is continuous. This shows that (a) holds. □

I.e. the easiest example of a non-path-connected metric space is a collection of disjoint intervals in $\mathbb{R}$.

1.2.3 Locally constant functions

Proposition 1.23 is our main tool to prove that a metric space X is path-connected. However, how do we prove that it is not? In light of Lemma 1.27, we are particularly interested in the case where X is a subset of $\mathbb{R}^n$ with $n > 1$. It turns out that the answer has to do with the study of *locally constant* functions $X \to Y$, as we will see now.

Definition 1.28. Suppose X and Y are metric spaces. A function $f : X \to Y$ is **locally constant** if for each $a \in X$ there exists a neighbourhood $V \ni a$ such that $f|_V$ is constant. △

A natural question then is whether every locally constant function $X \to Y$ is in fact constant globally. The following lemma handles the case of an interval, which will be the key to handle general spaces X:

Lemma 1.29. *A locally constant function $f : [a, b] \subset \mathbb{R} \to Y$ is constant.*

Proof. Zij $y := f(0) \in Y$. Consider the level set $S := f^{-1}(y) \subset [a, b]$. According to Lemma 1.7, it is closed, since f is continuous. At the same time, it is also open, since f is locally constant.

Observe that S contains y and is thus non-empty. Assume for contradiction that $S \neq [a, b]$. Its complement $S^c = [a, b] \setminus S$ is thus non-empty and bounded, so it has an infimum s_0. There are now two cases. If $s_0 \in S$, every point close to s_0 is also in S, because S is open; this contradicts that s_0 was the infimum. If $s_0 \in S^c$, then every nearby point is in S^c, since S^c is open (because S is closed); this contradicts that s_0 was the infimum. □

Now we can test locally constant functions with paths:

Proposition 1.30. *Let X and Y be metric spaces. Let $f : X \to Y$ be a locally constant function. If X is path-connected, then f is constant.*

Proof. Let p, q be a pair of points in X. Then there exists a continuous curve $\gamma : [0, 1] \to X$ with $\gamma(0) = p$ and $\gamma(1) = q$. We claim that $g := f \circ \gamma : [0, 1] \to Y$ is locally constant. From this, Lemma 1.29 implies that $f(p) = g(0) = g(1) = f(q)$, so f is constant.

The proof of the claim is as follows. Let $t_0 \in [0, 1]$. Since f is locally constant, there exists a $\delta > 0$ such that f is constant on the ball $B(\gamma(t_0), \delta)$. From the continuity of γ, there exists an open neighborhood J of t_0 in $[0, 1]$ such that $\gamma(J) \subset B(\gamma(t_0), \delta)$. It follows that $g = f \circ \gamma$ is constant on J and thus locally constant. □

We now spell out some consequences regarding the shape of metric spaces.

Corollary 1.31. *Suppose (X, d) is a path-connected metric space. Suppose $Y \subset X$ is both closed and open. Then Y is either the emptyset or X itself.*

Proof. X is the disjoint union of Y and its complement Y^c. Consider then the function $\chi : X \to \mathbb{R}$ that is identically zero over Y and identically one over Y^c. We claim that this function is locally constant. Indeed, given a point $y \in Y$, we see that Y itself is a neighbourhood with $f|_Y$ constant. The same is true for Y^c. According to Proposition 1.30, χ is constant, so it must attain only one value. This means that either Y or Y^c must be empty. □

Corollary 1.32. *Suppose $X \subset \mathbb{R}^n$ is the disjoint union of two non-empty opens. Then X is not path-connected.*

Proof. Denote the opens that form X as $\{X_i\}_{i=1,2}$. Define $\chi : X \to \mathbb{R}$ to have value i on X_i. This is a locally constant function, since each X_i is open. However, it is clearly not constant, so X is not path-connected (Proposition 1.30). □

Corollary 1.33. *Suppose $X \subset \mathbb{R}^n$ is the disjoint union of two non-empty closed subsets. Then X is not path-connected.*

Proof. Let X_1 and X_2 be the two claimed closed subsets. For contradiction, suppose that X is path-connected and take a path $\gamma : [0, 1] \to X$ starting at X_1 and finishing at X_2. This partitions $[0, 1]$ into the closed subsets $J_i = \gamma^{-1}(X_i)$, which are closed (Lemma 1.7). Since J_1 and J_2 are the complement of each other, they are also open. This contradicts 1.31. □

1.3 Homotopies of paths and loops

Up until now we have studied whether a metric space X may be decomposable into path-connected pieces. We will now study whether X may have a "hole" somewhere. The example you should keep in mind is the following: $\mathbb{R}^2$ and $\mathbb{R}^2 \setminus \{0\}$ are both path-connected, but the latter has a hole in the middle.

In order to detect said hole, we will do the following: Imagine that there is a closed piece of string (a closed curve) that goes around, i.e. it is "tied" to it. It will remain tied regardless of any pushing and pulling (i.e. even if we deform the curve). Thus: in order to detect that the hole is there, what we actually show is that closed curves are getting "stuck" somewhere.

This idea is fundamental. It is the focus in *Topologie en Meetkunde* (where it is studied in an algebraic/topological way). In this course we will, naturally, study it from an analytical perspective. We will see that we can use covector fields to see whether a closed curve is tied to something (because, in some sense, the covector field itself is tied as well!). However, this has to wait until Chapter 7. In this chapter we just introduce the relevant terminology. These ideas will be revisited, from the perspective of complex analysis, in *Functies en reeksen*. You will also encounter them if you study Physics: the presence of holes in a space has important consequences in how fields in it behave.

1.3.1 Homotopies

The following concept encapsulates the idea of deforming a curve:

Definition 1.34. Fix a metric space X. Let $\gamma_0, \gamma_1 : [a,b] \to X$ be a pair of continuous paths. A **homotopy** from γ_0 to γ_1 is a continuous function

$$\Gamma : [a,b] \times [0,1] \to X$$

satisfying for each $t \in [a,b]$:

$$\Gamma(t,0) = \gamma_0(t) \quad \text{and} \quad \Gamma(t,1) = \gamma_1(t). \qquad \triangle$$

In this situation we write $\gamma_s : [a,b] \to X$ for the path $t \mapsto \Gamma(t,s)$. As such, a homotopy can be understood as a continuous deformation starting at the curve γ_0 and finishing γ_1, passing through the intermediate curves γ_s.

Observe that this definition is not so useful yet. A homotopy allows us to deform a curve letting their endpoints move freely. I.e. we cannot tie it to anything. This is formalised by the following exercise:

Exercise 1.35. Fix a metric space X and let $\gamma : [a,b] \to X$ be a path. Show that γ is homotopic to the constant path ν with value $\gamma(a)$. Hint: Show first that the identity path $\iota : [a,b] \to [a,b]$ is homotopic to the constant path with value a. $\triangle$

Because of this, we should constrain the endpoints of the curves. This can be done in two ways. The first option reads:

Definition 1.36. A curve $\gamma[a,b] \to X$ is **closed** if $\gamma(a) = \gamma(b)$. A closed curve is also called a **loop**. $\triangle$

The second situation is that:

Definition 1.37. Let $\gamma_0, \gamma_1 : [a, b] \to X$ be continuous curves. The curves γ_0 and γ_1 are said to have the same endpoints if:

$$\gamma_0(a) = \gamma_1(a) \quad \text{and} \quad \gamma_0(b) = \gamma_1(b). \qquad \triangle$$

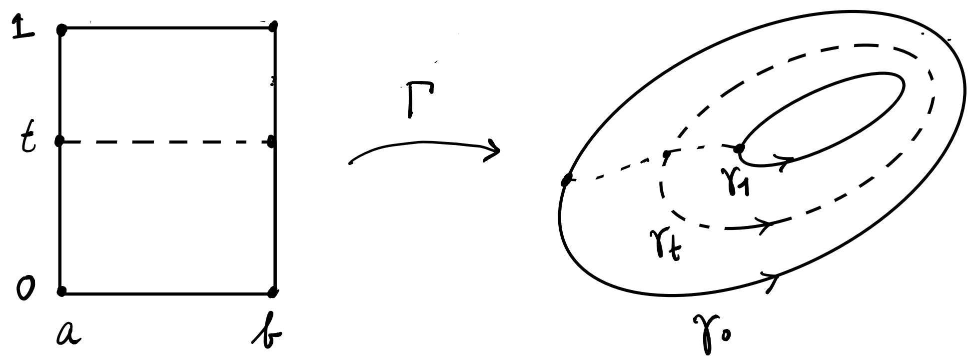
It may seem that this is not very helpful if we want to wrap our curves around a hole. However, you should imagine the following situation: in $\mathbb{R}^2 \setminus \{0\}$ imagine γ_0 going from $(1, 0)$ to $(-1, 0)$ by passing above the hole. Similarly, imagine γ_1 going from $(1, 0)$ to $(-1, 0)$ but passing below. Neither of them surrounds the hole per se, but in Chapter 7 we will nonetheless detect that they are "on a different side". One could do more complicated things: i.e. consider γ_2 that goes from $(1, 0)$ to $(-1, 0)$ but makes a whole loop around zero along the way.

What we can do now is deform curves keeping these constraints:

Definition 1.38. labeldef:loopHomotopy Suppose $\gamma_0, \gamma_1 : [a, b] \to U$ are loops. Let $\Gamma : [a, b] \times [0, 1] \to U$ be a homotopy from γ_0 to γ_1. We say that Γ is a **homotopy of loops** if for each $s \in [0, 1]$ we have that:

$$\Gamma(a, s) = \Gamma(b, s).$$

Said differently, we require each curve $\gamma_s = \Gamma(-, s)$ to be a loop. $\triangle$



Figuur 1: A homotopy of loops (i.e. closed curves).

Regarding the second situation:

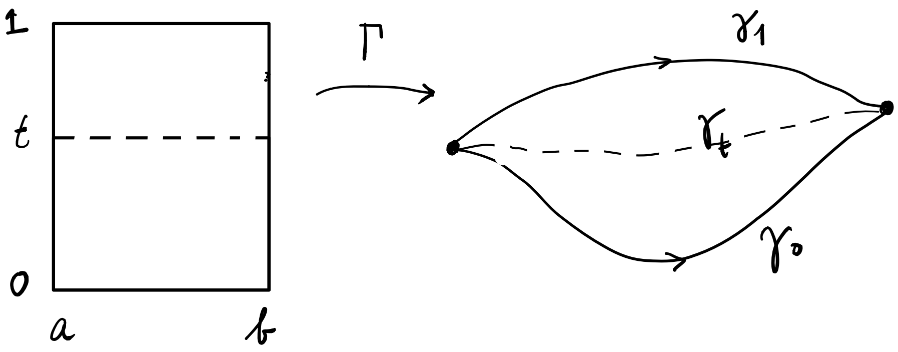
Definition 1.39. Suppose $\gamma_0, \gamma_1 : [a, b] \to U$ are curves with the same endpoints. Let $\Gamma : [a, b] \times [0, 1] \to U$ be a homotopy from γ_0 to γ_1. We say that Γ is a **homotopy relative endpoints** is for each $s \in [0, 1]$ we have that:

$$\Gamma(a, s) = \gamma_s(a) = \gamma_0(a) \quad \textit{and} \quad \Gamma(b, s) = \gamma_s(b) = \gamma_0(b). \tag{1.1}$$

That is, all the curves throughout the homotopy have the same endpoints. $\triangle$

1.3.2 Simply-connected metric spaces

A particularly important situation is the following:



Figuur 2: A homotopy relative endpoints.

Definition 1.40. Let (X, d) be a metric space. A closed curve $\gamma : [a, b] \to X$ is **nullhomotopic** or **contractible** if there is a homotopy of loops $\Gamma : [a, b] \times [0, 1] \to X$ from γ to a constant curve. △

Informally, this means that γ was not tied to anything, since we were able to push it down to a point. If we have a metric space in which we cannot tie curves at all, we say that:

Definition 1.41. A metric space (X, d) is **simply-connected** if the following conditions hold:

(a) X is path-connected.

(b) Each loop $\gamma : [a, b] \to X$ is nullhomotopic. △

Our main example reads:

Proposition 1.42. *Every star-shaped subset $U \subset \mathbb{R}^n$ is simply-connected. In particular, $\mathbb{R}^m$ itself, balls, cubes, and other convex subsets are simply-connected.*

Proof. Recall that U is path-connected, as shown in Lemma 1.25.

Pick a point $x_0 \in U$ with the property $[x, x_0] \subset U$ for all $x \in U$. Define a map $G : U \times [0, 1] \to U$ using the formula:

$$G(x, s) = (1 - s)x + sx_0.$$

This function is continuous and contracts the whole of U to a single point, since $G(x, 0) = x$ and $G(x, 1) = x_0$ for all $x \in U$. One then says that U is a *contractible space*. This is enough to show that every loop $\gamma : [a, b] \to U$ is contractible. Indeeed, the map $(t, s) \mapsto G(\gamma(x), s)$ is a homotopy $\Gamma : [a, b] \times [0, 1] \to U$ of loops with $\gamma_0 = \gamma$ and γ_1 identically constant with value x_0. □

Proving that a subset $U \subset \mathbb{R}^n$ is not simply-connected is much more difficult, and requires the machinery of Chapter 7. The main application there will be to show that $\mathbb{R}^2 \setminus \{0\}$ is indeed not simply-connected (Theorem 7.8).

Exercise 1.43. Prove that the following conditions are equivalent for a metric space (X, d):

- X is simply-connected.

- Any two loops $\gamma, \nu : [a, b] \to X$ are homotopic to each other as loops. △

The following tells us that, in a simply-connected space, there is a unique way, up to homotopy, of going from one point to another.

Proposition 1.44. *Suppose (X, d) is simply-connected. Then any two curves $\gamma_0, \gamma_1 : [a, b] \to X$ with the same endpoints (e.g. starting at some $x = \gamma_0(a) = \gamma_1(a)$ and finishing at some $y = \gamma_0(b) = \gamma_1(b)$) are homotopic relative endpoints.*

Proof. Consider the concatenation $\bar{\gamma}_1 \cdot \gamma_0$. This is a loop that is nullhomotopic, since X is simply-connected. Let Γ be the corresponding nullhomotopy of loops. Consider the curves $\eta_0, \eta_1 : [0, 1] \to [0, 1]^2$ given by $\eta_0(t) = (t/2, 0)$ and $\eta_1(t) = (1 - t/2, 0)$. By construction, $\Gamma \circ \eta_i = \gamma_i$. We also see that the upper side of the square yields the constant curve $t \mapsto \Gamma(t, 1)$. The sides $s \mapsto \Gamma(0, s)$ and $s \mapsto \Gamma(1, s)$ are constant as well, with value the initial point x.

We can use Γ to construct a homotopy relative endpoints now. To homotope γ_0, we homotope η_0 instead. Consider the rectangle $R_{s_0} = \{s \leq s_0, t \leq 1/2\}$. Its lower side is η_0. Let us denote the other three sides as the curve ν_{s_0}. We can orient the curve so that it has the same endpoints as η_0. By construction, R_{s_0} provides a homotopy relative endpoints between η_0 and ν_{s_0}. In particular, $\gamma_0 = \Gamma \circ \eta_0$ is homotopic relative endpoints to $\Gamma \circ \nu_1$.

In the same manner, consider the rectangles $S_{s_0} = \{s \leq s_0, t \geq 1/2\}$. The lower side is η_1 and the other three sides define a curve β_{s_0} with the same endpoints. It follows that γ_1 is homotopic relative endpoints to $\Gamma \circ \beta_1$. The argument now concludes by noting that $\Gamma \circ \beta_1$ and $\Gamma \circ \nu_1$ are the same curve. □

Do note that, in the proof above, we did not need to explicitly exhibit a homotopy relative endpoints between η_0 and ν_{s_0}. Since both had the same endpoints and a rectangle is convex, Proposition 1.42 implies that a homotopy relative endpoints exists.

2 Linear algebra

The goal of this course is to study multivariate functions. I.e. functions $\mathbb{R}^n \to \mathbb{R}^m$, which have thus n inputs and m outputs. To do so, it is best that we first get familiar with the "easiest" examples thereof. These are, in order, the constant functions, the linear functions, and the polynomial functions.

The goals of this chapter are to:

- Review the linear algebra that we will use during the rest of the course. A lot of it you have seen already in earlier courses.
- Explain how to write and manipulate multivariate polynomial expressions.
- Explain how changes of basis can help to simplify an expression. You are already familiar with this in the linear setting (where you know it as diagonalisation), but we will be particularly interested in the bilinear setting (e.g. the study of second order polynomials).

This last point is extremely important and will reappear many times in this course. Changing coordinates allows us to "find the right perspective" so the situation under study simplifies.

In this chapter we will use blue boxes not just for examples, but also to spell out how different abstract linear algebra concepts look in terms of matrices.

2.1 Vector spaces and linear maps

The main objects of study in linear algebra are **vector spaces** V which, in this course, will always be finite-dimensional and given over the real numbers $\mathbb{R}$.

Example 2.1. Our favourite family of vector spaces are the Euclidean spaces $\mathbb{R}^n$, as n ranges over the natural numbers. We will write e_i for for the ith coordinate vector:

$$\begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \in \mathbb{R}^n;$$

i.e. the only non-zero entry is the ith one. Together, they form the standard basis $\{e_1, \cdots, e_n\}$ of $\mathbb{R}^n$. We will stick to the following convention: vectors will always be written as *colums*. This is important when we perform calculations. $\triangle$

With that said, columns take some space. Because of this, we will sometimes write $(a_1, \ldots, a_n)$ for the column vector

$$\begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}.$$

The use of commas indicate that it is meant to be a column. Row vectors will be written without commas.

Given two vector spaces V and W, we can study linear maps $A : V \to W$. Recall that a map is **linear** if

$$A(v + \lambda v') = A(v) + \lambda A(v')$$

for every two vectors $v, v' \in V$ and every scalar $\lambda \in \mathbb{R}$. That is, the map A is compatible with the operations (vector addition and scalar multiplication) in V and W.

Notation 2.2. Consider two vector spaces V and W. We write $\mathrm{Lin}(V, W)$ for the set of linear maps from V to W. △

A particularly important example of linear map is the following:

Definition 2.3. A linear map $B : V \to W$ is a (linear) **isomorphism** if there is another linear map $C : W \to V$ such that:

- $C \circ B : V \to V$ is the identity in V.
- $B \circ C : W \to W$ is the identity in W.

If this is the case, we say that V and W are **isomorphic**. △

Example 2.4. Elements in $A \in \mathrm{Lin}(\mathbb{R}^n, \mathbb{R}^m)$ can be thought as n-times-m matrices (i.e. a matrix with n colums and m rows):

$$\begin{pmatrix} a_{11} & \cdots & a_{n1} \\ & \vdots & \\ a_{1m} & \cdots & a_{nm} \end{pmatrix}.$$

Applying the linear map A to a vector $v \in \mathbb{R}^n$ will then amount to performing the usual multiplication Av between a matrix (on the left) and a column vector (on the right, with n entries), producing a new column vector (now with m entries). In particular, note that $A(e_i)$ is the ith column of A.

A linear map $A \in \mathrm{Lin}(\mathbb{R}^n, \mathbb{R}^n)$ is an isomorphism if and only if the corresponding matrix is invertible, which is equivalent to its determinant being non-zero. △

2.1.1 Choosing a basis

V and W being isomorphic means that they are pretty much the same vector space, just expressed differently. In particular, since they are "the same", we can transfer data from one to the other. For instance, consider the following construction. Fix a linear map $A : V \to V$ and an isomorphism $B : W \to V$. Then $B^{-1} \circ A \circ B : W \to W$ is also a linear map. In fact, it is the copy of A, under the isomorphism B. This is precisely what change of basis is for a matrix, as we now explain.

First, recall that working with arbitrary vector spaces is no different than working with $\mathbb{R}^n$:

Lemma 2.5. *Every real vector space V of dimension n is isomorphic to $\mathbb{R}^n$.*

However, you should remember that the identification between V and $\mathbb{R}^n$ is not unique. Namely:

Lemma 2.6. *Suppose V is a real vector space of dimension n. Given a basis $\{v_1, \cdots, v_n\}$ of V we can define a linear isomorphism $A : \mathbb{R}^n \to V$ by setting $A(e_i) = v_i$. Equivalently, A sends the column vector*

$$\begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} \in \mathbb{R}^n$$

to $\sum_i a_i v_i \in V$.

Conversely, given an isomorphism $B : \mathbb{R}^n \to V$ we can define a basis $\{B(e_1), \cdots, B(e_n)\}$ of V.

These two processes are inverses to each other and define a bijective correspondence between bases of V and isomorphisms $\mathbb{R}^n \to V$.

To summarise: given any V you can say "let me take a basis " $\{v_1, \cdots, v_n\}$. This identifies V with $\mathbb{R}^n$, allowing you to write elements of v as columns. In particular, the ith basis element v_i will be represented by the ith coordinate vector e_i.

You can also write linear maps $A \in \mathrm{Lin}(V, V)$ as matrices. Indeed, if you fix an isomorphism $B : \mathbb{R}^n \to V$, then $B^{-1} \circ A \circ B \in \mathrm{Lin}(\mathbb{R}^n, \mathbb{R}^n)$ can be seen as a matrix. However, do not forget that this depends on B (i.e. on the basis chosen on V)!

Example 2.7. Bases are important even if we work in $\mathbb{R}^n$. Namely, consider a basis $\{v_1, \cdots, v_n\}$ in $\mathbb{R}^n$, possibly different from the standard one. Suppose we consider a linear map $A \in \mathrm{Lin}(\mathbb{R}^n, \mathbb{R}^n)$ that in the standard basis is represented by a matrix M. What is the matrix M' that represents A in the new basis? (Recall the meaning of this: the ith column of M' should be the vector $A(v_i)$, expressed in the basis $\{v_1, \cdots, v_n\}$).

We proceed as follows. We can use the vectors in the basis $\{v_1, \cdots, v_n\}$ to define a matrix

$$B := (v_1\, v_2\, \cdots\, v_n)$$

whose columns are the vectors v_i. This B is a linear map $\mathbb{R}^n \to \mathbb{R}^n$ that sends e_i to v_i. Then, M' is the matrix $B^{-1}MB$ (i.e. you apply the change of basis, you apply M, you apply the change back).

To summarise: M and $M' = B^{-1}MB$ both represent the same linear map A. M represents A in the standard basis and M' represents it in the basis given by the columns of B. △

This change of basis mechanism is what allows us to *diagonalise*, which we recall further in Subsection 2.3.

2.1.2 The vector space of linear maps

Linear maps between two different vector spaces can also be represented by matrices. Naturally, doing so still requires choosing bases in both:

Lemma 2.8. *Consider two vector spaces V and W of dimensions n and m, respectively. Then, $\mathrm{Lin}(V, W)$ is a real vector space of dimension nm. Moreover, choosing bases for V and W identifies $\mathrm{Lin}(V, W)$ with the space of n-times-m matrices.*

Proof. Given linear maps $A, B : V \to W$ and a scalar $\lambda \in \mathbb{R}$ you can verify that $v \mapsto A(v)+\lambda B(v)$ is a linear map from V to W, which we can naturally call $A + \lambda B$. This shows that $\mathrm{Lin}(V, W)$ is a vector space. Suppose we have now chosen bases $\{v_1, \cdots, v_n\}$ of V and $\{w_1, \cdots, w_m\}$ of W. Given $A \in \mathrm{Lin}(V, W)$ we can then consider the matrix whose (i, j) entry is $\langle A(v_i), w_j\rangle$; here the braces denote the usual scalar product. It is left to the reader to define the inverse process (i.e. to define an element in $\mathrm{Lin}(V, W)$ from an n-times-m matrix, using the given bases). □

Otherwise said: fixing bases gives us isomorphisms $\phi : \mathbb{R}^n \to V$ and $\psi : W \to \mathbb{R}^m$. As such, to any $A \in \mathrm{Lin}(V, W)$ we can associate $\psi \circ A \circ \phi : \mathbb{R}^n \to \mathbb{R}^m$, which we can naturally see as a matrix.

Exercise 2.9. Show that the vector space $\mathrm{Lin}(\mathbb{R}, V)$ is isomorphic (canonically, without choosing a basis) to V. △

Exercise 2.10. Consider vector spaces V and W and fix a vector $v \in V$. Consider the function $e_v : \mathrm{Lin}(V, W) \to W$ given by $A \mapsto A(v)$; we call it the *evaluation* at v. Show that e_v is a linear map. △

2.2 Kernels, rank, nullity

Given a linear map $A : V \to W$ you should also be familiar with the **kernel**:

$$\ker(A) := \{v \in V \mid A(v) = 0\}$$

and the **image** $\mathrm{im}(A)$. Observe that $\ker(A)$ is a **vector subspace** of V and $\mathrm{im}(A)$ is a subspace of W. They satisfy the so-called *rank-nullity theorem*:

Lemma 2.11. *Given a linear map $A : V \to W$, it holds that:*

$$\dim(V) = \dim(\ker(A)) + \dim(\mathrm{im}(A)).$$

The number $\dim(\ker(A))$ is called the **nullity** of A. Moreover:

Given a linear map $A : \mathbb{R}^n \to \mathbb{R}^m$ we think of it as a n-times-m matrix. Let us write $w_1, \cdots, w_n$ for its columns. These are vectors in $\mathbb{R}^m$ and, concretely, w_i is the image of $e_i \in \mathbb{R}^n$ under A. It follows that the image of A is spanned by the $w_1, \cdots, w_n$.

The **rank** of A is the number of linearly independent columns (or rows). As such, it is the dimension of the image of A. To compute it, the usual strategy is to find the dimension k of the largest minor of A with non-zero determinant. Recall that a **minor** of dimension k is a k-by-k submatrix; you choose it by picking the square matrix determined by the choice of k rows and k columns.

The kernel of A is spanned by the vectors $v \in \mathbb{R}^n$ with $Av = 0$. We will discuss more about computing the kernel in Subsection 2.4.3.

The summary of the previous discussion is that:

Proposition 2.12. *Consider $A \in \mathrm{Lin}(V, W)$ and let M be a matrix representing it (by choosing bases). Then the following numbers are the same:*

- $\dim(\mathrm{im}(A))$,
- *the rank of M,*
- *the dimension k of the largest minor of M with non-zero determinant.*

Example 2.13. Consider the matrix

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 0 \\ 3 & 3 & 3 \end{pmatrix},$$

which represents a linear map $\mathbb{R}^3 \to \mathbb{R}^3$. The span of its columns is the image $\mathrm{im}(A)$. We ask ourselves what is the dimension of said image (i.e. the rank of A).

You can verify that $\det(A) = 0$, meaning that the rank of A is less than 3; you can also see this by noting that the third row is the sum of the first two. We claim that the rank is in fact 2. For this, we need to find either (1) two linearly independent columns, (2) two linearly independent rows, or (3) a 2-by-2 minor with non-zero determinant. To see (1) just note that the first two columns are not proportional to each other. To see (2), note that the same is true for the first two rows. For (3), consider the 2-by-2 minor in the upper left. Having seen this, we deduce that the first two columns span $\mathrm{im}(A)$, which is thus a 2-dimensional vector space.

Using Lemma 2.11 we also deduce that the kernel is 1-dimensional, but it would be best to describe it explicitly, finding a vector (x, y, z) spanning it. By definition, it must satisfy the system of equations:

$$\begin{cases} x + 2y + 3z = 0 \\ 2x + y = 0 \end{cases}$$

that the first two rows define. We have that $y = -2x$, so the first equation simplifies to $0 = x - 4x + 3z = -3x + 3z$, yielding $x = z$. As such, the kernel contains exactly the vectors of the form $(x, -2x, x)$. To choose a basis for the kernel we just need to pick one concrete vector; $(1, -2, 1)$ does the job. △

2.3 Diagonalisation of linear maps

An important motto of this course is that one should change coordinates/basis whenever it is convenient. A fundamental example you are already familiar with is:

Definition 2.14. Let $A : V \to V$ be a linear map. An **eigenvector** of V is an element $v \in V$ such that $A(v) = \lambda v$, for some number $\lambda \in \mathbb{R}$, called the **eigenvalue**.

A basis $\{v_1, \cdots, v_n\}$ of V **diagonalises** A if it consists of eigenvectors. △

Equivalently: if we write A in an eigenvector basis, the matrix that represents it is diagonal, and its entries are the eigenvalues.

You should recall that not all matrices diagonalise over the real numbers. However, diagonalisation can be achieved in the following particular situation, described by the so-called *spectral theorem*:

Proposition 2.15. *Suppose $A : \mathbb{R}^n \to \mathbb{R}^n$ is a linear map represented by a symmetric matrix. Then A can be diagonalised using a orthonormal basis.*

Let us spell out what this means. An n-times-n matrix O is **orthogonal** if its columns form a orthonormal basis (i.e. they have length 1 and are orthogonal to each other). This is the same as $O^tO = id$, meaning that $O^t = O^{-1}$.

Diagonalising A using a orthonormal basis means that there is an orthogonal matrix O such that

$$O^{-1}AO = \begin{pmatrix} A_- & 0 & 0 \\ 0 & A_0 & 0 \\ 0 & 0 & A_+ \end{pmatrix},$$

where A_- is a diagonal matrix with negative numbers in the diagonal, A_0 is the zero matrix, and A_+ is diagonal with positive numbers in the diagonal.

These three matrices are all square matrices. If we denote their sizes by k_-, k_0, and k_+ we have that $n = k_- + k_0 + k_+$. Then, the first k_- columns of O span a vector subspace V_-, the middle k_0 columns span a subspace V_0, and the last k_+ columns span a subspace V_+. This is splitting $\mathbb{R}^n = V_- \times V_0 \times V_+$ into three pieces. According to the diagonalisation above, V_0 is the kernel of A. Furthermore, A acts by dilating negatively on V_- and by dilating positively on V_+.

2.3.1 Something that is not exactly diagonalisation

Observe that diagonalisation is the answer to the following question: "Given $A \in \text{Lin}(V, V)$, can I find a basis of V in which A looks as simple as possible?"

A different, but similar problem, is the following: "Given $A \in \text{Lin}(V, W)$, can I find bases of V and W in which A looks as simple as possible?" The following lemma tells us that the answer is always yes:

Lemma 2.16. *Ler V and W be vector spaces of dimensions n and m, respectively. Let $A : V \to W$ be a linear map of rank $k \leq m$ and nullity $n - k \leq n$. Then, there are bases $\{v_1, \cdots, v_n\}$ of V and $\{w_1, \cdots, w_m\}$ of W such that $A(v_i) = w_i$ if $i = 1, \cdots, k$ and $A(v_i) = 0$ otherwise.*

Equivalently, the matrix M representing A in these bases has an identity k-times-k submatrix in the upper left corner and is everywhere else zero.

You should observe that, even when $V = W$, Lemma 2.16 is different from diagonalisation. Indeed, given $A \in \text{Lin}(V, V)$, the lemma will produce two different bases for V, one when we think of V as the domain of A and another when we see it as the target of A.

2.4 The dual vector space

Given vector spaces V and W, we have seen that $\mathrm{Lin}(V, W)$ is another vector space (Lemma 2.8). The following particular case is extremely important:

Definition 2.17. Let V be a vector space. Its **dual vector space** is

$$V^* := \mathrm{Lin}(V, \mathbb{R}).$$

An element of V^* is said to be a **covector**. △

The name "dual" refers to the fact that, by definition, you can plug in vectors $v \in V$ into covectors $\alpha \in V^*$ to yield numbers $\alpha(v) \in \mathbb{R}$. The "co" in "covector" is a common prefix to indicate that this object is dual to something else (to a vector, in this case).

Recall that we write vectors in $\mathbb{R}^n$ as columns. This was important so things work well with matrix multiplication. Now this is particularly relevant: a covector $\alpha \in (\mathbb{R}^n)^* = \mathrm{Lin}(\mathbb{R}^n, \mathbb{R})$ is an n-times-1 matrix, i.e. a *row vector*. We readily see that the matrix multiplication of a covector against a vector yields a number:

$$\alpha(v) = (\alpha_1 \cdots \alpha_n) \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} = \sum_i \alpha_i v_i.$$

Much like in $\mathbb{R}^n$ we have the ith coordinate vector e_i, in $(\mathbb{R}^n)^*$ we have the ith coordinate covector e_i^*. This is the row with a 1 in the ith entry and all other entries zero. By definition:

$$e_i^*(v) = v_i,$$

i.e. e_i^* is the linear map that sends a vector to its ith coordinate.

Example 2.18. Suppose $V = \mathbb{R}^3$. The covector $e_1^* \in V^* = \mathrm{Lin}(\mathbb{R}^3, \mathbb{R})$ is simply the linear function $e_1^* : \mathbb{R}^3 \to \mathbb{R}$ given by:

$$(x, y, z) \mapsto x.$$

In particular, it is zero over the plane $\langle e_2, e_3 \rangle$ (i.e. the (y, z)-plane), which is its kernel. △

The previous story applies to general vector spaces V of dimension n. Indeed, once we choose a basis $\{v_1, \cdots, v_n\}$ of V, we can consider the so-called **dual basis** $\{v_1^*, \cdots, v_n^*\}$ of V^* defined as follows: $v_i^*(v_j) = 1$ if $i = j$ and is otherwise zero. If you use the basis $\{v_1, \cdots, v_n\}$ to identify $\mathbb{R}^n \to V$, so e_i goes to v_i, you will also identify $(\mathbb{R}^n)^*$ with V^*. Namely, v_i^* will be identified with e_i^*.

Exercise 2.19. Suppose that $V = \mathbb{R}^n$ and consider the usual inner product in $\mathbb{R}^n$. Given a vector $v \in V$ we can define its dual to be the covector $v^* \in V^*$ given by the expression $v^*(w) := \langle w, v \rangle$. Show that the map $v \in V \mapsto v^* \in V^*$ is an isomorphism. Hint: check that it sends the standard basis to its dual basis. △

Exercise 2.20. Suppose that $v \in V$ satisfies $\alpha(v) = 0$ for all $\alpha \in V^*$. Show that v is zero. Hint: take a basis and use the dual basis. △

Exercise 2.21. Suppose $A : V \to W$ is a linear map. Consider the *dual* map $A^* : W^* \to V^*$ given by $\alpha \in W^* \mapsto \alpha \circ A \in V^*$. Show that A^* is linear. △

2.4.1 The dual of the dual

We have seen that the process of plugging a vector $v \in V$ (a column) into a covector $\alpha \in V^*$ (a row), is rather symmetric. Indeed, it amounts to multiplying them as matrices. As such, we could imagine that it is α that is plugged into v instead. This is formalised by the following statement:

Lemma 2.22. *Let V be an n-dimensional vector space. Given a vector $v \in V$, we can define $\psi_v \in (V^*)^* = \mathrm{Lin}(V^*, \mathbb{R})$ to be the linear map $\alpha \in V^* \mapsto \alpha(v) \in \mathbb{R}$.*

The map $\psi : V \to (V^*)^*$ given by $v \mapsto \psi_v$ is a linear isomorphism.

Proof. It was shown in Exercise 2.10 that ψ_v is indeed a linear map.

To see that ψ is injective we must show that ψ_v is zero if and only if v is zero. The if direction is clear. For the only if, observe that $\psi_v = 0$ means that $\psi_v(\alpha) = \alpha(v) = 0$ for all α in V^*. This implies that v is zero (Exercise 2.20). We have shown that ψ is injective, so its kernel has dimension zero. According to the rank-nullity theorem (Lemma 2.11), this means that the image has dimension n. Since $(V^*)^*$ has dimension n as well, the image is everything, so ψ is an isomorphism. □

I.e. a vector is the same as a cocovector. At this point the following classic joke is relevant: "A mathematician turns coffee into theorems. A comathematician turns cotheorems into ffee."

Exercise 2.23. Consider a vector space V, its dual V^*, and the dual of the dual $(V^*)^*$. Fixing a basis $\{v_1, \cdots, v_n\}$ of V gives us the dual basis $\{v_1^*, \cdots, v_n^*\}$ of V^* and the dual dual basis $\{v_1^{**}, \cdots, v_n^{**}\}$ of $(V^*)^*$. Show that the isomorphism described in Lemma 2.22 sends v_i to v_i^{**}. (This provides an alternate proof of the lemma). △

Example 2.24. Consider $\mathbb{R}^3$ with coordinates (x, y, z). Write $(x^*\ y^*\ z^*)$ for the dual coordinates in $(\mathbb{R}^3)^*$, meaning that each covector in $(\mathbb{R}^3)^*$ is expressed as $x^*e_1^* + y^*e_2^* + z^*e_3^*$. Then, the vector $e_1 \in \mathbb{R}^3$ can be seen as the linear function $e_1 : (\mathbb{R}^3)^* \to \mathbb{R}$ satisfying $e_1(x^*\ y^*\ z^*) = x^*$. That is, e_1 sends e_1^* to 1 and e_2^* and e_3^* to zero. △

2.4.2 The annihilator

This idea of duality becomes even clearer if we look at subspaces. Namely:

Definition 2.25. Let V be a vector space and let V^* be its dual. Fix a subspace $W \subset V$. Then its **annihilator** is the subset:

$$\mathrm{Ann}(W) := \{\alpha \in V^* \mid \alpha(w) = 0, \forall w \in W\}.$$ △

Observe that this works both ways, thanks to Lemma 2.22. Given a subspace $Z \subset V^*$ we can also consider its annihilator, which is a subspace of V:

$$\mathrm{Ann}(Z) := \{v \in V \mid \alpha(v) = 0, \forall \alpha \in Z\}.$$

Example 2.26. Consider the subspace $W \subset \mathbb{R}^3$ spanned by e_1. It is 1-dimensional, i.e. a line. By construction, $e_1^*(e_1) = 1$, so e_1^* is not in the annihilator. This means that $\mathrm{Ann}(W)$ has dimension at most 2. Then we recall: $e_2^*(e_1) = e_3^*(e_1) = 0$, showing that $\text{Ann}(W)$ has indeed dimension 2, being spanned by e_2^* and e_3^*.

Conversely, we could have started with $Z = \text{Ann}(W) = \langle e_2^*, e_3^* \rangle \subset (\mathbb{R}^3)^*$. Then $W = \text{Ann}(Z)$. $\triangle$

In general:

Proposition 2.27. *Fix a vector space V of dimension n and let W be a k-dimensional subspace. Then:*

- $\text{Ann}(W)$ *is a subspace of V^* of dimension $n - k$.*
- $\text{Ann}(\text{Ann}(W)) = W$.

This defines a 1-to-1 correspondence between k-dimensional subspaces of V and $(n-k)$-dimensional subspaces of V^*.

Proposition 2.27 follows from the following very nice idea, which once again shows that changing basis can make the situation more transparent:

Lemma 2.28. *Fix a vector space V of dimension n and let W be a k-dimensional subspace. Then, there is a basis $\{v_1, \cdots, v_n\}$ of V in which W is spanned by $\{v_1, \cdots, v_k\}$ and $\text{Ann}(W)$ is spanned by $\{v_{k+1}^*, \cdots, v_n^*\}$.*

2.4.3 Linear equations and the dual

This may all seem very abstract, but you should simply think of elements of V^* as linear equations on V.

Example 2.29. Suppose $V = \mathbb{R}^3$ and consider the equation

$$2x + y + 3z = 0.$$

Our goal is to solve the equation. I.e. we want to find the subspace $A \subset V$ of vectors (x, y, z) that satisfy it. To do this, we can simply solve for y, yielding $y = -2x - 3z$. This means that A contains the vectors of the form $(x, -2x - 3z, z)$. You can readily see that A has dimension 2, and is spanned by $(1, -2, 0)$ and $(0, -3, 1)$, for instance. This makes sense: for each linearly independent equation we impose, the dimension of the subspace of solutions decreases by one.

This process can be understood from the perspective of V^* as follows. First, we write the equation in matrix form:

$$(2\,1\,3) \begin{pmatrix} x \\ y \\ z \end{pmatrix} = 0.$$

This means that the equation is fully described by the row vector $\alpha_1 = (2\,1\,3)$, i.e. a covector! Observe that our space of solutions A is, by definition, the set of those vectors $v \in V$ such that $\alpha_1(v) = 0$.

Consider then B, the 1-dimensional subspace of V^* spanned by α_1. Observe that all elements α in B are multiples of α_1. This means that if $v \in A$ then $\alpha(v) = 0$, for every $\alpha \in B$. We have thus shown that $A = \mathrm{Ann}(B)$. In particular, since the dimension of B is 1, Proposition 2.27 implies that the dimension of A is 2. △

More generally, the story is the following, which follows immediately from the definition of annihilator:

Lemma 2.30. *Consider a system of k linear equations in $\mathbb{R}^n$:*

$$\begin{cases} a_{11}x_1 + & \cdots & +a_{n1}x_n & = 0 \\ & \vdots & & \\ a_{1k}x_1 + & \cdots & +a_{nk}x_n & = 0 \end{cases}$$

and write it in matrix form as:

$$Mv = \begin{pmatrix} a_{11} & \cdots & a_{n1} \\ & \vdots & \\ a_{1k} & \cdots & a_{nk} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = 0.$$

Then:

- *The columns of M are vectors in $\mathbb{R}^k$. They span the image.*
- *The kernel of M is the space of solutions $A \subset \mathbb{R}^n$.*
- *The rows of M are covectors in $(\mathbb{R}^n)^*$. They span a "subspace of equations" $B \subset (\mathbb{R}^n)^*$.*
- *$A = \mathrm{Ann}(B)$.*

That is, the solutions are dual to the equations they solve!

Example 2.31. Let us use Lemmas 2.30 and 2.28 to solve the equation of Example 2.29. This was quite easy to do "by hand", but the following is a general procedure. It basically says the following: any system of linear equations can be solved by inverting a matrix.

The idea is that we want to change basis to make the equation look simple. To do so, we pick $\alpha_2 = (0\,1\,0)$ and $\alpha_3 = (0\,0\,1)$ which, together with $\alpha_1 = (2\,1\,3)$, form a basis of V^*. This means that the matrix of equations M, whose only row is α_1, can be extended to a matrix:

$$L := \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{pmatrix}.$$

What we have to do next is find the basis of V dual to $\{\alpha_1, \alpha_2, \alpha_3\}$. By definition, the dual basis will be the columns of the matrix K that satisfies $LK = id$, i.e. K is the inverse of L, namely:

$$K := \begin{pmatrix} 1/2 & -1/2 & -3/2 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Let us call its columns v_1, v_2, and v_3. By definition of dual basis we have that $\alpha_1(v_2) = \alpha_1(v_3) = 0$. E.g. v_2 and v_3 solve the equation and therefore span A, the space of solutions. △

Here is another example:

Example 2.32. Suppose $V = \mathbb{R}^4$, fix coordinates (x, y, z, w), and consider the system of equations

$$\begin{cases} 2x + y + 3z = 0 \\ z + 2w = 0 \end{cases}$$

Rewritten in matrix form:

$$\begin{pmatrix} 2 & 1 & 3 & 0 \\ 0 & 0 & 1 & 2 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \\ w \end{pmatrix} = 0.$$

We can see that the two equations (i.e. the two rows) are linearly independent (for instance, look at the 2-by-2 minor in the middle, which has determinant 1). We can call these two covectors α_1 and α_2; they span the subspace of equations $B \subset V^*$, which is 2-dimensional. We deduce (Proposition 2.27) that the subspace of solutions $A = \mathrm{Ann}(B) \subset V$ has dimension $2 = 4 - \dim(B)$ as well. △

2.5 Bilinear algebra

Consider the following familiar example:

Given an n-by-n matrix A, we can consider the function $\mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ defined by:

$$(v, w) \mapsto w^t A v = (w_1 \cdots w_1) A \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix},$$

where v and w are (column) vectors in $\mathbb{R}^n$ (but w gets transposed to become a row vector). For instance, we could have taken A to be the identity matrix, in which case the map is simply the usual scalar product:

$$(v, w) \mapsto \sum_{i=1}^{n} v_i w_i.$$

You can readily see that this function is not linear, even though it is represented by a matrix.

This is the most important example of a *bilinear map*, which is what we are exploring in this subsection. We will see that bilinear maps are closely related to second order polynomials.

2.5.1 Bilinear maps

Definition 2.33. Let V and W be linear spaces. A function $A : V \times V \to W$ is **bilinear** if $v \mapsto A(v, v')$ is linear for each $v' \in V$ and $v' \mapsto A(v, v')$ is linear for each $v \in V$. △

It turns out that our starting example is pretty much general. When $V = \mathbb{R}^n$ and $W = \mathbb{R}$, we have that:

Lemma 2.34. *Let $A : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ be a bilinear map. Then A can be written as:*

$$A(x, y) = y^t A x = \sum_{i,j=1,\cdots,n} a_{i,j} x_i y_j,$$

i.e. the coefficients of A form an n-by-n matrix $(a_{i,j})$.

When V and W are arbitrary vector spaces we can still take bases to identify V with $\mathbb{R}^n$ and W with $\mathbb{R}^m$, in order to obtain a very similar description.

Example 2.35. A typical example for us will be the case $V = \mathbb{R}^2$ and $W = \mathbb{R}$. Then a bilinear map can be written as a 2-by-2 matrix:

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

which can be applied to a pair of vectors as follows:

$$w^t A v = (w_1\, w_2) \begin{pmatrix} a_{11} & a_{21} \\ a_{12} & a_{22} \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = a_{11}v_1w_1 + a_{21}v_2w_1 + a_{12}v_1w_2 + a_{22}v_2w_2.$$

You see the bilinearity in the fact that each term in the right-most expression contains a coefficient a_{ij}, a single entry from v, and a single entry from w. △

2.5.2 The space of bilinear maps

Definition 2.36. Suppose V and W are vector spaces. We let $\mathrm{Lin}^2(V, W)$ be the set of bilinear maps $V \times V \to W$. △

As in the linear case, we have:

Lemma 2.37. $\mathrm{Lin}^2(V, W)$ *is a vector space.*

Proof. Given bilinear maps A and B and a scalar $\lambda \in \mathbb{R}$ we can define $A + \lambda B$ to be the map

$$(v, v') \mapsto A(v, v') + \lambda B(v, v')$$

which is readily seen to be bilinear. □

In the concrete case $V = \mathbb{R}^n$ and $W = \mathbb{R}$, we have that $\mathrm{Lin}^2(V, W)$ is the space of n-by-n matrices. It is a vector space if we use addition and scalar multiplication entry by entry, as we have seen before.

2.5.3 Symmetry and quadratic forms

A bilinear form has two inputs, but nothing stops us from plugging the same vector twice:

Definition 2.38. Let V be a vector space. Suppose $A : V \times V \to \mathbb{R}$ is a bilinear map. The associated **quadratic form** is the map $v \mapsto A(v, v)$. △

Quadratic forms can be represented by matrices as well (basically, by the matrix that represents the starting bilinear map). We can see it in the following concrete example:

Example 2.39. The bilinear map $A : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$ given by the 2-by-2 matrix:

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

yields the quadratic form $Q : \mathbb{R} \to \mathbb{R}$ given by:

$$Q(v) = A(v, v) = v^t A v = (v_1\, v_2) \begin{pmatrix} a_{11} & a_{21} \\ a_{12} & a_{22} \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = a_{11} v_1^2 + (a_{21} + a_{12}) v_1 v_2 + a_{22} v_2^2,$$

which is a polynomial of order two.

You may observe that Q is determined by three coefficients: a_{11}, $a_{21} + a_{12}$, and a_{22}. In particular, observe that the concrete values of a_{21} and a_{12} are not important for Q, only their sum. △

The example tells us that different bilinear maps can define the same quadratic form. However:

Definition 2.40. Let V and W be vector spaces. Suppose $A : V \times V \to W$ is a bilinear map. We say that it is **symmetric** if $A(v, v') = A(v', v)$ for every v and v' in V. △

Suppose $A : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ is bilinear and symmetric. Then the matrix representing A is symmetric.

Lemma 2.41. *Suppose V is a vector space of dimension n. Then, there is a bijective correspondence between bilinear symmetric maps $V \times V \to \mathbb{R}$ and quadratic forms. Moreover, if we fix a basis, the two are also in correspondence with n-by-n symmetric matrices.*

Proof. By definition, to each bilinear form we can associate a quadratic form. It is sufficient if we show that this is bijective when we restrict to symmetric bilinear forms. This is most easily done by choosing a basis first, which identifies V with $\mathbb{R}^n$. Once that is done, any bilinear map A can be written as $A(x, y) = \sum_{i,j=1,\cdots,n} a_{i,j} x_i y_j$. It is symmetric if $a_{i,j} = a_{j,i}$. Then, the quadratic form associated to it is

$$Q(x) = \sum_{1 \leq i \leq n} x_i^2 + \sum_{1 \leq i < j \leq n} 2 a_{i,j} x_i x_j,$$

where we have separated the diagonal terms from the rest. Therefore, the coefficients that describe Q uniquely recover A. □

2.5.4 The classification of quadratic forms

An important conclusion of the proof of Lemma 2.41 is that:

Every quadratic form $Q : \mathbb{R}^n \to \mathbb{R}$ can be written in the form

$$Q(x) = \sum_{1 \leq i \leq n} x_i^2 + \sum_{1 \leq i < j \leq n} 2 a_{i,j} x_i x_j.$$

I.e. a quadratic form is a polynomial of order 2 that is *pure* (i.e. its constant and linear terms are zero). Quadratic forms are thus the easiest functions after constant and linear.

As such, we want to understand them a bit more. First, note that every quadratic form satisfies $Q(0) = 0$, i.e. it vanishes at the origin of $\mathbb{R}^n$. The following tells us a bit more about the vanishing locus of Q:

Definition 2.42. Let $A : V \times V \to \mathbb{R}$ be a symmetric bilinear map with associated quadratic form $Q : V \to \mathbb{R}$ (Lemma 2.41). We say that A and Q are **non-degenerate** if for every $v \in V$ there is some $w \in V$ such that $A(v, w) \neq 0$. Otherwise we say that they are **degenerate**.

Moreover, there are three types of non-degeneracy. We say that A and Q are...

- **positive definite** if $Q(v) \geq 0$ for all $v \in V$, and $Q(v) = 0$ implies $v = 0$.
- **negative definite** if $-Q$ is positive definite.
- **indefinite**, otherwise. △

Example 2.43. The prototypical examples of quadratic forms in $\mathbb{R}^2$ are the following. First, both the zero matrix and

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$$

are degenerate (but the zero matrix is "more degenerate" than A, since A has rank 1). The identity matrix id is positive definite. Its negative $-id$ is negative definite. Lastly,

$$B = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

is indefinite. △

Suppose we have a symmetric matrix representing a quadratic form Q. You can ask yourself how we can tell whether Q is degenerate or not, or whether it is definite. This is explained by the following result, *Sylvester's law of inertia*:

Proposition 2.44. *Suppose $Q : \mathbb{R}^n \to \mathbb{R}$ is a quadratic form. Then Q is non-degenerate if and only if it has non-zero determinant.*

Moreover, the following are equivalent:

- *Q is positive definite.*
- *The eigenvalues of Q are all positive.*
- *For each $i = 1, \cdots, n$, the i-times-i upper left minor of Q is positive if i is even and is negative if i is odd.*

The concrete case that will be most important for us is the following:

Proposition 2.45. *Suppose $Q : \mathbb{R}^2 \to \mathbb{R}$ is a quadratic form represented by a symmetric 2-by-2 matrix. Then:*

- *If $\det(Q) = 0$, then Q is degenerate.*
- *If $\det(Q) < 0$, then Q is indefinite. It has one positive and one negative eigenvalues.*
- *If $\det(Q) > 0$ and $\operatorname{tr}(Q) > 0$, then Q is positive definite. Both eigenvalues are positive.*
- *If $\det(Q) > 0$ and $\operatorname{tr}(Q), 0$, then Q is negative definite. Both eigenvalues are negative.*

Before we get to the proof, recall that the **trace** of a square matrix is the sum of its diagonal entries. This number is invariant under change of basis, so it agrees with the sum of eigenvalues (if the matrix diagonalises).

Proof. From Proposition 2.15 we know that Q has an orthonormal basis v_1, v_2 of eigenvectors. Let λ_1 and λ_2 be the corresponding eigenvalues. Then Q is positive (resp. negative) definite if and only if λ_1 and λ_2 are positive (resp. negative). Thus Q is definite if and only if $\det(Q) = \lambda_1\lambda_2 > 0$. Otherwise it is indefinite (one eigenvalue of each sign and thus $\det(Q) < 0$) or degenerate (at least one eigenvalue is zero and thus $\det(Q) = 0$).

Moreover, the trace of Q is independent of the choice of basis, so $\mathrm{tr}(Q) = \lambda_1 + \lambda_2$. If the trace is positive, both eigenvalues are positive and Q is positive definite. If the trace is negative, Q is negative definite. Do note that the trace is zero if and only if at least one eigenvalue is zero (i.e. Q is degenerate). □

The diagonalisation argument used in the proof in fact generalises to all dimensions to provide the classification of quadratic forms:

Proposition 2.46. *Suppose $Q : V \to \mathbb{R}$ is a quadratic form. Then, there is a basis of V in which M, the matrix representing Q, is of the form:*

$$\begin{pmatrix} -\mathrm{Id} & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \mathrm{Id} \end{pmatrix}.$$

In particular, V can be decomposed into three subspaces $V = V_- \times V_0 \times V_+$ so that Q is negative definite over V_-, zero over V_0, and positive definite over V_+ and, moreover, vectors in different subspaces do not interact with each other via Q.

Remark 2.47. Proposition 2.46 could be called the "diagonalisation of quadratic forms"; we explain it now. Keep in mind that this is different from the diagonalisation of linear maps (Proposition 2.15) and different from the change of bases of Lemma 2.16.

Instead of looking at the quadratic form $Q : V \to \mathbb{R}$, consider the corresponding $A : V \times V \to \mathbb{R}$ that is bilinear and symmetric. If we were given a linear map $C \in \mathrm{Lin}(\mathbb{R}^n, V)$ the only composition we could possibly do is:

$$A'(v, w) = A(C(v), C(w)),$$

which is now a bilinear form $\mathbb{R}^n \to \mathbb{R}$; it is still symmetric.

Now, how does that look in coordinates? Suppose that $V = \mathbb{R}^n$. Then we can think of Q, A, and C as matrices; note that $Q = A$ is symmetric. Then:

$$w^t A' v = A'(v, w) = A(C(v), C(w)) = (C(w))^t A C(v) = w^t C^t A C v.$$

That is, A' is the matrix $C^t AC$. This is different from the base change for linear maps $B : \mathbb{R}^n \to \mathbb{R}^n$, which would instead yield $C^{-1}BC$.

How does this connect with Proposition 2.46? Given Q, we want a basis of $\mathbb{R}^n$ in which Q looks simpler. If we choose a basis $\{c_1, \cdots, c_n\}$, we can assemble its vectors as colums of a square matrix C. You should think of C as a linear map $\mathbb{R}^n \to \mathbb{R}^n$ that sends the standard basis $\{e_1, \cdots, e_n\}$ to the new basis $\{c_1, \cdots, c_n\}$. Then, writing Q in the basis $\{c_1, \cdots, c_n\}$ is precisely computing $C^t QC$. △

2.5.5 Inner products, norms, distances

There are two concrete examples of bilinear maps and quadratic forms that you are probably most familiar with:

Definition 2.48. A bilinear map $g : V \times V \to \mathbb{R}$ is said to be an **inner/scalar product** if it is symmetric and positive definite. △

Example 2.49. The standard Euclidean inner product in $\mathbb{R}^n$ is defined as:

$$(v, w) \mapsto \langle v, w \rangle = \sum_{i=1}^{n} v_i w_i = (w_1, \cdots, w_n)\mathrm{Id} \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}.$$

I.e., as a bilinear form it is represented by the identity matrix Id. △

According to Proposition 2.46 this example is, up to change of basis, completely general:

Lemma 2.50. *Let V be a vector space with inner product g. Then there is a basis V in which g is represented by the identity matrix. Identically, there is a vector space isomorphism $\mathbb{R}^n \to V$ that identifies g with the standard Euclidean inner product.*

Given an inner product $g : V \times V \to \mathbb{R}$ we can consider the associated quadratic form $v \mapsto g(v, v)$ and take the square root $l(v) = g(v, v)^{1/2}$. The function l is an example of:

Definition 2.51. A function $l : V \to \mathbb{R}$ is a **norm** if it satisfies the following axioms:

- Positive-definiteness: $l(0) = 0$ is zero and $l(v) > 0$ otherwise.
- Homogeneity: $l(\lambda v) = |\lambda| l(v)$ for every $v \in V$ and every $\lambda \in \mathbb{R}$.
- Triangular inequality: $l(v + w) \leq l(v) + l(w)$ for every $v, w \in V$. △

Example 2.52. The usual Euclidean inner product in $\mathbb{R}^n$ yields the Euclidean norm:

$$v \mapsto ||v|| = \langle v, v \rangle^{1/2} = \left(\sum_{i=1}^{n} v_i^2 \right)^{1/2}.$$

Which in turn allows us to define the Euclidean distance:

$$d(v, w) := ||v - w|| = \left(\sum_{i=1}^{n} (v_i - w_i)^2 \right)^{1/2}.$$

I.e. the norm tells us how large a vector is, i.e. how far it is from the origin. The distance then allows us to measure how far apart two vectors are from one another. △

This is a general phenomenon:

Definition 2.53. Given a norm $l : V \to \mathbb{R}$, its associated **distance** is the map $V \times V \to \mathbb{R}$ given by

$$(v, w) \mapsto l(v - w).$$ △

2.6 Polynomial functions

In Lemma 2.41 we saw that quadratic forms $\mathbb{R}^n \to \mathbb{R}$ are the same as polynomials of (pure) order 2. Let us look now into the general theory of polynomials of multiple variables.

Definition 2.54. A function $A : \mathbb{R}^n \to \mathbb{R}$ is a **polynomial of pure order** r if it can be written as:

$$A(v) = \sum_{i_1,\cdots,i_r=1,\cdots,n} a_{i_1,\cdots,i_r} v_{i_1} \cdots v_{i_r}.$$

A function $\mathbb{R}^n \to \mathbb{R}^m$ is polynomial of pure order r if each component is. △

Observe that constant functions, linear functions, and quadratic forms are polynomials of pure order 0, 1, and 2, respectively.

Example 2.55. Consider the map $A : \mathbb{R}^2 \to \mathbb{R}$ given by the expression:

$$A((v_1, v_2)) = v_1^3 - v_2^3.$$

This is a polynomial of pure order 3. The non-zero coefficients are $a_{1,1,1} = 1$ and $a_{2,2,2} = -1$.

Consider instead $B : \mathbb{R}^2 \to \mathbb{R}$ given by:

$$B((v_1, v_2)) = v_1 v_2^2.$$

Now we have some choices. We could take $a_{1,2,2} = 1$ and all other coefficients zero. We could also take $a_{1,2,2} = 1/2$ and $a_{2,1,2} = 1/2$ and all others zero. △

To have a unique expression, it is more convenient to adopt the following convention. This is exactly the same as the correspondence between quadratic forms and symmetric bilinear forms (Lemma 2.41):

Lemma 2.56. *Each polynomial $A : \mathbb{R}^n \to \mathbb{R}$ of pure order r can be uniquely written as:*

$$A(v) = \sum_{i_1,\cdots,i_r=1,\cdots,n} a_{i_1,\cdots,i_r} v_{i_1} \cdots v_{i_r}$$

if we assume that the coefficients satisfy $a_{i_1,\cdots,i_j,\cdots,i_k,\cdots,i_r} = a_{i_1,\cdots,i_k,\cdots,i_j,\cdots,i_r}$, for every i and j.

That is, if you were to assemble the coefficients into a n-times-n-times-...-times-n “hypermatrix” with r dimensions (this is called a *tensor*), it would be symmetric. See Subsection 2.6.3.

Example 2.57. Going back to the example $B : \mathbb{R}^2 \to \mathbb{R}$ with

$$B((v_1, v_2)) = v_1 v_2^2$$

we see that the symmetric choice is to take $a_{1,2,2} = a_{2,1,2} = a_{2,2,1} = 1/3$, and all other coefficients zero. △

If we add up pure polynomials of different orders, we obtain:

Definition 2.58. A function $f : \mathbb{R}^n \to \mathbb{R}^m$ is **polynomial** of order r if it is a sum of polynomials of pure orders $i \leq r$. △

2.6.1 The linear space of polynomials

We already saw that $\mathrm{Lin}(V, W)$ and $\mathrm{Lin}^2(V, W)$ are both linear spaces. We now prove the analogous fact for:

Definition 2.59. We write $\mathrm{Pol}^r(\mathbb{R}^n, \mathbb{R}^m)$ and $\mathrm{Pol}^{\leq r}(\mathbb{R}^n, \mathbb{R}^m)$ for the sets of polynomials $\mathbb{R}^n \to \mathbb{R}^m$ of pure order r and order r, respectively. We also denote

$$\mathrm{Pol}(\mathbb{R}^n, \mathbb{R}^m) := \cup_{r=0}^{\infty} \mathrm{Pol}^{\leq r}(\mathbb{R}^n, \mathbb{R}^m)$$

for the set of polynomials of arbitrary order. △

The following is left to the reader:

Lemma 2.60. $\mathrm{Pol}^r(\mathbb{R}^n, \mathbb{R}^m)$*,* $\mathrm{Pol}^{\leq r}(\mathbb{R}^n, \mathbb{R}^m)$*, and* $\mathrm{Pol}(\mathbb{R}^n, \mathbb{R}^m)$ *are vector spaces, for every non-negative integer* r*. The first two are finite dimensional. Moreover, the inclusions*

$$\mathrm{Pol}^r(\mathbb{R}^n, \mathbb{R}^m) \to \mathrm{Pol}^{\leq r}(\mathbb{R}^n, \mathbb{R}^m) \to \mathrm{Pol}^{\leq r+1}(\mathbb{R}^n, \mathbb{R}^m) \to \mathrm{Pol}(\mathbb{R}^n, \mathbb{R}^m)$$

are linear maps.

Given a function $f : \mathbb{R}^n \to \mathbb{R}^m$ and a point $a \in \mathbb{R}^n$, we can define the *translated function* $T_a(f) : \mathbb{R}^n \to \mathbb{R}^m$ using the expression

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

We then have:

Proposition 2.61. *Suppose* $f \in \mathrm{Pol}(\mathbb{R}^n, \mathbb{R}^m)$*. Then* $T_a(f)$ *is also polynomial.*

Moreover, the function $T_a : \mathrm{Pol}(\mathbb{R}^n, \mathbb{R}^m) \to \mathrm{Pol}(\mathbb{R}^n, \mathbb{R}^m)$ *is a linear isomorphism*[1]*. Its inverse is the opposite translation* $(T_a)^{-1} = T_{-a}$*.*

Proof. Before we begin, it is helpful to make a number of auxiliary remarks. First of all, you should verify the following identities:

a. $T_a(g + h) = T_a(g) + T_a(h)$,

b. $T_a(gh) = T_a(g)T_a(h)$,

which hold for all functions $g, h : \mathbb{R}^n \to \mathbb{R}^m$. Secondly:

c. A sum of polynomials is also a polynomial. The same is true for the product.

d. A function $f = (f_1, \cdots, f_m) : \mathbb{R}^n \to \mathbb{R}^m$ is polynomial if and only if each component $f_i : \mathbb{R}^n \to \mathbb{R}$ is polynomial.

Now we prove the first claim. According to item (d) it is enough if we prove it for $m = 1$. If f is a constant function, we have that $T_a(f) = f$ is also constant.

[1] The proof in fact shows something stronger: the set of polynomials is a ring (it has an addition and a multiplication) and T_a is a ring homomorphism.

Suppose next that $f : \mathbb{R}^n \to \mathbb{R}$ is linear. Every linear function can be written, as a linear combination, using the dual basis $e_i^* \in (\mathbb{R}^n)^*$. This fact, together with items (a) and (c), implies that it is enough if we show that each $T_a(e_i^*)$ is polynomial. By definition, $e_i^*(v) = v_i$, so its translate is

$$T_a(e_i^*)(x) = e_i^*(x + a) = x_i + a_i.$$

which is a first order polynomial again (whose linear part is the same as before, but now has a constant term). This concludes the linear case.

Suppose now that $f : \mathbb{R}^n \to \mathbb{R}$ is an arbitrary polynomial. As such, it can be written as a sum of monomials of the form $v \mapsto a_{i_1,\cdots,i_r} v_{i_1} \cdots v_{i_r}$. Each monomial is in turn a product of linear functions, whose translations we already know are polynomials as well. Facts (a), (b), and (c) then imply that $T_a(f)$ is polynomial, as we wanted to prove.

To establish the second claim observe that items (a) and (b) imply that T_a is indeed linear. It is easy to check that T_{-a} is the inverse. □

The previous proof communicates a very important idea: when proving statements about polynomials it is very helpful to see them as a sum of products of linear functions.

2.6.2 Multi-indices

In Lemma 2.56 we saw that the coefficients describing a pure polynomial are unique if we ask that they are symmetric. We now discuss another way of writing down polynomials, which will become handy.

Definition 2.62. Een rijtje $\alpha = (\alpha_1, \ldots, \alpha_n)$ van niet-negatieve gehele getallen wordt een *multi-index* genoemd. Het getal

$$|\alpha| := \sum_{j=1}^{n} \alpha_j$$

heet de *orde* van de multi-index α. △

A multi-index is very useful in order to condense expressions involving many indices. For instance, given a multi-index $\alpha = (\alpha_1, \cdots, \alpha_n)$, we can write a monomial as:

$$v \in \mathbb{R}^n \mapsto v^\alpha := \prod_{i=1}^{n} v_j^{\alpha_j} \in \mathbb{R};$$

which is a pure polynomial of order $|\alpha|$. A more general pure polynomial of order r is then written as:

$$v \mapsto \sum_{|\alpha|=r} a_\alpha v^\alpha \qquad (2.1)$$

where each $a_\alpha \in \mathbb{R}$ is the coefficient in front of the monomial v^α.

Example 2.63. Once again we consider the example $B : \mathbb{R}^2 \to \mathbb{R}$ given by

$$B(v) = B((v_1, v_2)) = v_1 v_2^2 = v^{(1,2)}.$$

Here $v^{(1,2)}$ should be read as "v to the power of the multi-index $(1, 2)$", meaning that you take v_1 once and v_2 to the power of two, and you multiply them. The coefficient in front is called

$a_{(1,2)} = 1$, and all others (namely, $a_{(3,0)}$, $a_{(2,1)}$, and $a_{(0,3)}$) are zero. △

Exercise 2.64. Let $v \in \mathbb{R}^n$ be a vector and α be a multi-index. Establish the inequality:

$$\|v^\alpha\| \leq \|v\|^{|\alpha|}.$$ △

Each pure polynomial can be written, uniquely, in the form described in Equation 2.1. This uniqueness means that the monomials v^α, as α ranges over all multi-indices of length r, form a basis of $\mathrm{Pol}^r(\mathbb{R}^n, \mathbb{R}^m)$. This allows us to deduce:

Corollary 2.65. $\mathrm{Pol}^r(\mathbb{R}^n, \mathbb{R}^m)$ *has dimension* $\binom{r+n-1}{n-1}$.

Proof. There are as many α with $|\alpha| = r$ as manners of distributing r items in n boxes. This is the same as choosing $n-1$ separators in a collection of size $r+n-1$, since that determines n blocks whose sizes add up to r. □

2.6.3 Extra: multilinear maps

This subsubsection is optional, you do not need to study it. Nonetheless, it may help giving further context to the theory of polynomials that we have seen. The main idea is the following: just like pure second order polynomials (i.e. quadratic forms) are in correspondence with symmetric bilinear maps (Lemma 2.41), pure r-order polynomials are in correspondence with symmetric multilinear maps of order r.

Definition 2.66. Fix a positive integer r and let V and W be linear spaces. A function $A : V^r \to W$ is **multilinear** (of order r) if: For every i, the map $v_i \in V \mapsto A(v_1, \cdots, v_i, \cdots, v_r) \in W$ is linear whenever we fix all the other v_j, $j \neq i$. △

Note that here each v_i is a vector in V, not a coefficient.

Definition 2.67. We write $\mathrm{Lin}^r(V, W)$ for the set of multilinear maps $V^r \to W$. △

Example 2.68. Suppose $V = \mathbb{R}^2$ and $W = \mathbb{R}$. Then, the map $A : V \times V \times V \to W$ defined by:

$$A((a_1, a_2), (b_1, b_2), (c_1, c_2)) = a_1 b_1 c_1 + a_2 b_2 c_2$$

is trilinear (i.e. multilinear of order 3). If you plug in the same vector three times you get

$$A((a_1, a_2), (a_1, a_2), (a_1, a_2)) = a_1^3 + a_2^3,$$

which is polynomial of pure order three. △

The previous example is a concrete instance of:

Definition 2.69. Suppose $A : V^r \to W$ is a multilinear map. We say that $A \in \mathrm{Lin}^r(V, W)$ is **symmetric** if

$$A(v_1, \cdots, v_i, \cdots, v_j, \cdots, v_r) = A(v_1, \cdots, v_j, \cdots, v_i, \cdots, v_r)$$

for every i and j. △

Any multilinear map can be expressed as follows (which should remind you of how we write pure polynomials):

Lemma 2.70. *Suppose $A \in \mathrm{Lin}^r(\mathbb{R}^n, \mathbb{R})$ is a multilinear map. Then:*

$$A(v_1, \cdots, v_r) = A\left(\begin{pmatrix}(v_1)_1\\ \vdots \\ (v_1)_n\end{pmatrix}, \cdots, \begin{pmatrix}(v_r)_1\\ \vdots \\ (v_r)_n\end{pmatrix}\right) = \sum_{j_1,\cdots,j_r=1,\cdots,n} a_{j_1,\cdots,j_r}(v_1)_{j_1}(v_2)_{j_2}\cdots(v_r)_{j_r}.$$

Moreover, A is symmetric if and only if $a_{j_1,\cdots,j_i,\cdots,j_k,\cdots,j_r} = a_{j_1,\cdots,j_k,\cdots,j_i,\cdots,j_r}$ for every i and k between 1 and r.

2.7 Basic analysis of polynomial functions

To wrap up the chapter, we are going to start doing some Analysis, focusing on polynomial functions. We have not yet developed the theory of multivariate differentiation (this is the goal of Chapters 3 and 4), so we will focus solely on continuity (using the results of Chapter 1) and some key estimates.

2.7.1 Continuity

Proposition 2.71. *A polynomial $f : \mathbb{R}^n \to \mathbb{R}^m$ is a continuous function.*

Proof. We use the same idea as in the proof of Proposition 2.61. First, f is continuous if and only if each of its components $f_i : \mathbb{R}^n \to \mathbb{R}$ is continuous (Lemma 1.6), so we just need to worry about the case $m = 1$. Morever, a polynomial $f : \mathbb{R}^n \to \mathbb{R}$ is a sum of pure polynomials so, according to the sum rule for continuity (Proposition 1.5), we just need to show that a pure polynomial is continuous.

In turn, each pure polynomial is a sum of terms of the form

$$v \mapsto a_{i_1,\cdots,i_r} v_{i_1} \cdots v_{i_r},$$

a so-called *monomial*. Applying the sum rule again, we just need to show monomials are continuous. However, aach monomial consists of a scalar $a_{i_1,\cdots,i_r} \in \mathbb{R}$ multiplied by the functions $v_{i_j}^* : v \mapsto v_{i_j}$, each of which is linear. According to the product rule for continuity (Proposition 1.5) it is enough if we show that linear functions are continuous.

In fact, it is enough if we show that the functions in the dual basis $v_i^* : v \mapsto v_i$ are continuous. But this is clear, since

$$|v_i^*(v) - v_i^*(v')| = |v_i - v_i'| = |(v - v')_i| \leq ||v - v'||,$$

which shows

$$\lim_{v \to v'} |v_i^*(v) - v_i^*(v')| = 0.$$

□

In future chapters we will see that polynomials are in fact smooth, as you would expect.

In proving continuity we needed to use the euclidean norm and the estimate $|x_i| \leq ||x||$. In the next subsections we will study this norm further and develop a series of estimates that will allow us to prove statements like Proposition 2.71 much more easily.

2.7.2 Norms for linear maps

Consider vector spaces V and W of dimensions n and m, respectively. As we discussed in Lemma 2.8, $\mathrm{Lin}(V, W)$ is itself a vector space of dimension nm. This means that we can consider inner products and norms in $\mathrm{Lin}(V, W)$. This is actually quite natural: if we apply a linear map A to a vector v, we want to relate $||Av||$ to $||v||$, and this should involve the "size of A", which should be some sort of norm[2] $||A||$.

To keep things grounded, we will focus on the concrete case of the n-times-p matrices. Observe that $\mathrm{Lin}(\mathbb{R}^n, \mathbb{R}^p)$ can be identified with $\mathbb{R}^{np}$ (by forming a long column out of the columns of a given matrix), which allows us to consider the usual Euclidean norm:

Definition 2.72. The **norm** $\|L\|$ of a linear map $L : \mathbb{R}^n \to \mathbb{R}^p$ is defined to be the expression:

$$\|L\| := \left(\sum_{i=1}^{p} \sum_{j=1}^{n} L_{ij}^2 \right)^{1/2},$$

where the L_{ij} denote the matrix coefficients of L. △

Naturally, we can also consider the associated distance. Note moreover that one can define a scalar product for matrices, but this will not come into play in this course. By construction, we have the following statement:

Corollary 2.73. $\mathrm{Lin}(\mathbb{R}^n, \mathbb{R}^p)$*, endowed with the norm from Definition 2.72, is isomorphic to* $\mathbb{R}^{np}$ *endowed with the usual Euclidean norm.*

2.7.3 The Cauchy–Schwarz inequality

You are probably familiar with the *Cauchy-Schwarz* inequality:

Proposition 2.74. *Let V be a vector space with a scalar product $\langle\, ,\, \rangle$. Then:*

$$|\langle v, w\rangle| \leq ||v||.||w||$$

with the equality holding if and only if w and v are proportional.

It turns out that this statement can be generalised to the case in which we have a linear map and a vector, or two linear maps:

Proposition 2.75. *Consider a linear map $L : \mathbb{R}^n \to \mathbb{R}^p$ and a vector $v \in \mathbb{R}^n$. Then the following holds:*

$$\|L\, v\| \leq \|L\|\, \|v\|. \tag{2.2}$$

Suppose we have a second linear map $M : \mathbb{R}^p \to \mathbb{R}^q$. Dan geldt voor de samenstelling $M \circ L : \mathbb{R}^n \to \mathbb{R}^q$ dat

$$\|M \circ L\| \leq \|M\|\, \|L\|. \tag{2.3}$$

[2]Of course, one can choose different norms, but we will not explore this in this course, you will have to wait for *Functional analysis*.

Proof. Let us address the first case first. The idea is that we apply the usual Cauchy–Schwarz (Proposition 2.74) one row of L at a time (equivalently, one entry of the result $L(v)$ at a time). I.e for each $1 \leq i \leq p$ we have that:

$$((L\,v)_i)^2 = \left(\sum_{j=1}^{n} L_{ij}\, v_j \right)^2 \leq \left(\sum_{j=1}^{n} (L_{ij})^2 \right) \|v\|^2.$$

Summation over i yields $\|L\,v\|^2 \leq \|L\|^2\, \|v\|^2$, which implies the claimed Equation (2.2).

The second statement is analogous. We apply the Cauchy–Schwarz inequality to each entry of the resulting matrix $M \circ L$:

$$((M\,L)_{hj})^2 = \left(\sum_{i=1}^{p} M_{hi}\, L_{ij} \right)^2 \leq \left(\sum_{i=1}^{p} (M_{hi})^2 \right) \left(\sum_{i=1}^{p} (L_{ij})^2 \right).$$

The argument concludes by summing over h and j. □

Using the Cauchy-Schwarz inequality we recover the continuity of linear maps:

Corollary 2.76. *Let $A : \mathbb{R}^n \to \mathbb{R}^p$ be a linear map. Then A is continuous.*

Proof. We must show that $||A(v) - A(w)||$ goes to zero as $||v - w||$ goes to zero. But this is immediate once we use linearity and Proposition 2.3:

$$||A(v) - A(w)|| \leq ||A||.||v - w||.$$ □

Example 2.77. Suppose $A \in \mathrm{Lin}(\mathbb{R}^n, \mathbb{R}^n)$ can be diagonalised with eigenvectors v_i of eigenvalue λ_i, with $i = 1, \cdots, n$. Then we see that, for each i:

$$|\lambda_i|.||v_i|| = ||A(v_i)|| \leq ||A||.||v_i||$$

so we deduce that the size of the eigenvalues is bounded above by $||A||$. △

2.7.4 Norm of a bilinear map

The next step is clear. We have studied how the norm interacts with linear maps. If we are given a bilinear form $A : V \times V \to \mathbb{R}$ we could similarly try to relate $|A(v, v')|$ to $||v||$ and $||v'||$; this should involve some quantity $||A||$. We define it as follows, in the case of Euclidean space:

Definition 2.78. Suppose $A : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ is a bilinear form. Express it as a square matrix $(a_{ij})_{i,j=1}^n$. Then we define its **norm** to be

$$||A|| := \left(\sum_{i,j=1}^{n} a_{ij}^2 \right)^{1/2}.$$

Moreover, given a quadratic form Q defined by the bilinear form A, we can define the norm of Q to be $||A||$. △

As you can see this is very similar to what we did in the linear case. We are simply identifying bilinear forms with the matrices that represent them, which identifies $\mathrm{Lin}^2(\mathbb{R}^n, \mathbb{R})$ with $\mathbb{R}^{n^2}$ and thus allows us to use the standard euclidean norm.

We then have the following version of Cauchy-Schwarz:

Lemma 2.79. *Suppose $A : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ is a bilinear form. Then the following holds for any two vectors $v, v' \in V$:*

$$|A(v, v')| \leq ||A||.||v||.||v'||.$$

Proof. Write $V = \mathbb{R}^n$ for notational convenience. Let a_{ij} be the coefficients of A. Observe that $v' \mapsto A(v, v')$ is a covector, which we could call $\alpha \in V^*$. Observe that its entries are

$$\alpha_j = \sum_{i,j} a_{ij} v_i$$

so that

$$A(v, v') = \sum_{i,j} a_{ij} v_i v'_j = \sum_j \alpha_j v'_j.$$

Having introduced this notation, the matrix version of Cauchy-Schwarz (Proposition 2.75) tells us that:

$$||\alpha|| = ||A(v, -)|| \leq ||A||.||v||.$$

We can then apply the usual Cauchy–Schwarz (Proposition 2.74):

$$|A(v, v')| = |\sum_j \alpha_j v'_j| \leq ||\alpha||.||v'|| \leq ||A||.||v||.||v'||.$$ □

Which can be used to deduce:

Corollary 2.80. *Let $B : \mathbb{R}^n \times \mathbb{R}^m \to R^p$ be a bilinear map. Then B is continuous.*

Proof. As above, we bound

$$||B(v - v', w - w')|| \leq ||B||.||v - v'||.||w - w'||,$$

which shows that $||B(v - v', w - w')||$ goes to zero as $||v - v'||$ and $||w - w'||$ do, which is what we had to show. □

Remark 2.81. We claim that the evaluation map $\mathrm{Lin}(\mathbb{R}^n, \mathbb{R}^p) \times \mathbb{R}^n \to \mathbb{R}^p$ given by $(L, v) \mapsto Lv$ is continuous. You can verify that this map is bilinear, which allows us to apply Corollary 2.80. Nonetheless, let us spell out how the proof goes, so you see how Cauchy-Schwarz for matrices is used in practice.

Consider vectors $v,\ v_0 \in \mathbb{R}^n$ and linear maps $L,\ L_0 \in \mathrm{Lin}(\mathbb{R}^n, \mathbb{R}^p)$. Using Equation (2.2) we deduce that

$$\begin{aligned} Lv - L_0 v_0 &= L\,v - L\,v_0 + L\,v_0 - L_0\,v_0 \\ &= L\,(v - v_0) + (L - L_0)\,v_0 \\ &= L_0\,(v - v_0) + (L - L_0)\,v_0 + (L - L_0)\,(v - v_0), \end{aligned}$$

implying that

$$\|Lv - L_0 v_0\| \leq \|L_0\|\,\|v - v_0\| + \|L - L_0\|\,\|v_0\| + \|L - L_0\|\,\|v - v_0\|, \tag{2.4}$$

which implies continuity. △

Exercise 2.82. Using similar argumetns you can show that matrix composition $(L, M) \mapsto M \circ L$ defines a continuous, bilinear map $\mathrm{Lin}(\mathbb{R}^n, \mathbb{R}^p) \times \mathrm{Lin}(\mathbb{R}^p, \mathbb{R}^q) \to \mathrm{Lin}(\mathbb{R}^n, \mathbb{R}^q)$. △

2.7.5 The size of a quadratic form

Lemma 2.79 and Corollary 2.80 imply:

Corollary 2.83. *Suppose $Q : \mathbb{R}^n \to \mathbb{R}$ is a quadratic form. Then Q is continuous and satisfies:*

$$|Q(v)| \leq ||Q||.||v||^2.$$

Proof. If B is the symmetric bilinear form associated to Q, we can write $Q = B \circ \Delta$ with $\Delta : \mathbb{R}^n \to \mathbb{R}^n \times \mathbb{R}^n$ the map $v \mapsto (v, v)$. Since B is bilinear and Δ is linear, earlier statements tell us that they are continuous. As such, so is their composition. □

This tells us that $|Q|$ is bounded above by the square of the usual euclidean norm (up to the factor $||Q||$). The following result tells us that that this works both ways if Q is itself an inner product:

Proposition 2.84. *Suppose $Q : \mathbb{R}^n \to \mathbb{R}$ is a positive definite quadratic form (i.e. an inner product). Then there is a constant $C > 0$ such that*

$$Q(v) \geq C||v||^2.$$

Proof I (using linear algebra). The idea is that we can switch the roles of Q and the standard inner product if we change basis, so the result will follow from Corollary 2.83.

More formally: We think of Q as a symmetric matrix. According to Proposition 2.15, there is an orthonormal basis of eigenvectors $\{v_1, \cdots, v_n\}$ with eigenvalues λ_i that diagonalise Q. Moreover, since Q is positive definite, we have that all these eigenvalues are positive.

Given any $v \in \mathbb{R}^n$ we can express it as $v = \sum_i a_i v_i$ using said basis. Then, using bilinearity we see that:

$$Q(v) = v^t Q v = \sum_{i=1}^{n} \lambda_i ||v_i||^2 = \sum_i \lambda i > 0.$$

Where we have used that the v_i form an orthonormal basis and that the eigenvalues are positive. □

Proof II (using analysis). Observe that what we are showing is that there is a *uniform* estimate for the growth of Q, that works for all vectors v. We can therefore argue as we did in Proposition 1.13.

Suppose that the statement is not true, for contradiction. Then for each positive integer k there is a vector $v_k \in R^n$ such that

$$Q(v_k) < \frac{1}{k}||v_k||^2.$$

Do note that this implies that $v_k \neq 0$, so we can instead consider the vectors $u_k := v_k/||v_k||$, which are contained in the unit sphere $\mathcal{S} = \{v \in \mathbb{R}^n \mid \|v\| = 1\}$. Using bilinearity we see that:

$$Q(u_k) < \frac{1}{k}.$$

So we have a sequence $\{u_k\}_{k=1}^{\infty}$ of points in $\mathcal{S}$ whose values $Q(u_k) \to 0$ as k goes to infinity. Observe now that $\mathcal{S}$ is a closed and bounded subset of $\mathbb{R}^n$. This means (Proposition 1.10) that $\{u_k\}_{k=1}^{\infty}$ has a subsequence that converges to some $u_\infty \in \mathcal{S}$. Since Q is continuous (Corollary 2.83) we deduce that $Q(u_\infty) = 0$. This contradicts the fact that Q is positive definite. □

The analytical proof more generally shows that *any two norms on a finite dimensional vector space are equivalent* (i.e. they bound each other up to some constants).

The proof of the following fact is left to you. You may want to review Example 2.43 and Proposition 2.46:

Exercise 2.85. Let $Q : V \to \mathbb{R}$ be a quadratic form. Then:

- If Q is positive definite, it has a global minimum at the origin. It has no maximum.
- If Q is negative definite, it has a global maximum at the origin. It has no minimum.
- If Q is indefinite, its zero level set contains points other than the origin. It has no global maximum nor minimum.
- If Q is degenerate, its zero level set contains points other than the origin. It may or may not have global maxima/minima. △

The theme of critical points/maxima/minima will be explored in depth in later chapters.

2.7.6 The size of a polynomial

In analogy with earlier statements we deduce:

Lemma 2.86. *Suppose $P : \mathbb{R}^n \to \mathbb{R}^m$ is a polynomial of order r. Then P is continuous and there is a constant $C > 0$ such that:*

$$|P(v)| \leq C||v||^r.$$

Proof. Given a monomial $M : v \in V \mapsto v_{i_1}v_{i_2}\cdots v_{i_k} \in \mathbb{R}$ we can estimate $|M(v)| < ||v||^k$. The claim follows since P is a sum of monomials of order at most r. This estimate shows continuity and is left for the reader. □

3 Partial, directional, and total derivatives

In this chapter we finally begin our study of differentiation of functions $f : \mathbb{R}^n \to \mathbb{R}^m$. Assuming that one can indeed differentiate (not all functions are differentiable afterall, see Example 3.28!), we will:

- Define the (total) derivative of f as a map $Df : \mathbb{R}^n \to \mathrm{Lin}(\mathbb{R}^n, \mathbb{R}^m)$ (Definition 3.13). I.e. at each point $x \in \mathbb{R}^n$ we get a linear function $Df(x) \in \mathrm{Lin}(\mathbb{R}^n, \mathbb{R}^m)$ that is the best linear approximation of f at x.
- See that the columns of $Df(x)$ are given by the partial derivatives $D_i f(x)$ along the coordinate directions and prove (Theorem 3.29) that differentiability can be checked fully in terms of partial derivatives.
- Establish the sum, product, quotient rules (Proposition 3.46) and the chain rule (Theorem 3.48) for differentiation, which are our main tools to show that a function is differentiable (and to compute what the derivative is in many cases).
- Introduce the first order Taylor polynomial $P_x^1(f)$ (Definition 3.23). This is the first order polynomial that best approximates f at x.
- State and prove two versions of the mean value theorem (Theorem 3.40 and Proposition 3.45). Mean value results are one of the main technical ingredients behind many of the important results of the course. In this chapter we will use them to prove the chain rule. In later chapters we will establish other versions (Theorem 6.8) in order to deduce further applications.

3.1 Partial derivatives

In the following, U is an open subset of $\mathbb{R}^n$ and $a \in U$ a point. Let $f : U \to \mathbb{R}$ be a real-valued function. Let $1 \leq j \leq n$. If we fix the coordinates with index different from j, and let the jth coordinate vary freely, then we obtain the real-valued function

$$\varphi\colon t \mapsto f(a_1, \dots, a_{j-1}, t, a_{j+1}, \dots, a_n) \tag{3.1}$$

of a single real variable t.

> We have just encountered one of the central ideas of the course. Whenever we are trying to prove/define/construct something in the multivariate setting, we should ask ourselves: *Is it possible to reduce the proof/construction to the case of one variable?*

The function φ is defined on the following subset of $\mathbb{R}$:

$$I_j(a) := \{t \in \mathbb{R} \mid (a_1, \dots, a_{j-1}, t, a_{j+1}, \dots, a_n) \in U\}.$$

It can be seen as follows that $I_j(a)$ contains an open interval around a_j: Since U is open, there exists a $\delta > 0$ such that $B(a;\delta) \subset U$. For $t \in (a_j - \delta, a_j + \delta)$ we then have

$$(a_1, \dots, a_{j-1}, t, a_{j+1}, \dots, a_n) \in B(a;\delta),$$

and therefore $(a_j - \delta, a_j + \delta) \subset I_j(a)$.

Since φ is defined on an open interval around a_j, the question of differentiability of the function φ makes sense. This is the reason why we assumed that U is open.

Definition 3.1. The function f is called **partially differentiable** with respect to the j-th variable at the point $a \in U$ if the function φ defined by (3.1) is differentiable at the point $t = a_j$.

If this is the case, then the derivative of the function φ at the point $t = a_j$ is called the **partial derivative** of f with respect to the j-th variable at the point a and is denoted by

$$D_j f(a) := \varphi'(a_j) = \frac{d}{dt} f(a_1, \ldots, a_{j-1}, t, a_{j+1}, \ldots, a_n)\bigg|_{t=a_j}. \qquad \triangle$$

Example 3.2. Let $f : \mathbb{R}^2 \to \mathbb{R}$ be defined by $f(x, y) = e^{xy}y + x$. Then, according to the above definition, the function f is at every point $a = (a_1, a_2) \in \mathbb{R}^2$ partially differentiable with respect to both the first and the second variables. Moreover:

$$D_1 f(a_1, a_2) = \frac{d}{dt}(e^{ta_2}a_2 + t)\bigg|_{t=a_1} = (e^{ta_2}a_2^2 + 1)\big|_{t=a_1} = e^{a_1 a_2}a_2^2 + 1. \qquad \triangle$$

Remark 3.3. In the above, it was required that the domain U of f was open. For a more general $V \subset \mathbb{R}^n$, we can consider a function $f : V \to \mathbb{R}$ and a point $a \in \operatorname{inw}(V)$. The function f is then called partially differentiable at a if the restriction of f to the interior of V is partially differentiable there. $\triangle$

We return to the situation where $U \subset \mathbb{R}^n$ is open.

Definition 3.4. A function $f : U \to \mathbb{R}$ is called partially differentiable with respect to the j-th variable if it is partially differentiable with respect to the j-th variable at every point $a \in U$.

In that case we can consider the partial derivative $D_j f(a)$ as a function of $a \in U$. This function is called the partial derivative of f with respect to the j-th variable and is denoted by:

$$D_j f : U \to \mathbb{R}, \quad x \mapsto D_j f(x). \qquad \triangle$$

Remark 3.5. In the literature, the notation $\partial_j f$ is used sometimes instead of $D_j f$. If we denote the coordinates in $U \subset \mathbb{R}^n$ by the variables $x = (x_1, \ldots, x_n)$, one can also write

$$\frac{\partial f(x)}{\partial x_j}.$$

If one uses other variables (for example (x, y) in the plane), then one may write analogous expressions with them. In larger formulas one often it is sometimes handy to write

$$\frac{\partial}{\partial x_j} f(x)$$

or even $\partial_j f(x)$. The symbol ∂ was introduced around 1840 by Jacobi.

The partial derivative $D_j f(a)$ of f at a point a may be denoted by

$$\frac{\partial f(x)}{\partial x_j}\bigg|_{x=a} = \frac{\partial f}{\partial x_j}(a).$$

This notation can be confusing. For example, for a function $f(y, t)$ we have that

$$\frac{\partial f}{\partial y}(0, 0)$$

denotes the partial derivative with respect to the first variable at $(0,0)$, but a careless reader may assume it was the second.

This can become more complicated with expressions such as $\frac{\partial f}{\partial x}(x,x)$, which can mean:

$$\left.\frac{\partial f(x,y)}{\partial x}\right|_{y=x} = (D_1 f)(x,x),$$

or

$$\left.\frac{\partial f(y,x)}{\partial x}\right|_{y=x} = (D_2 f)(x,x),$$

or $\frac{d}{dx}f(x,x)$. This illustrates that it is important to choose notations so that the meaning remains clear. Keep this in mind when writing your exam! △

Example 3.6. In the situation of Example 3.2 we have

$$\frac{\partial}{\partial x}f(x,y) = \frac{\partial}{\partial x}(e^{xy}y + x) = e^{xy}y^2 + 1,$$
$$\frac{\partial}{\partial y}f(x,y) = \frac{\partial}{\partial y}(e^{xy}y + x) = e^{xy}(xy + 1). \qquad \triangle$$

Since partial differentiation is in fact differentiation of a function of a single real variable (namely the j-th coordinate), the following calculation rules are a direct consequence of the corresponding familiar rules for functions of a single variable.

Proposition 3.7. *Let f, g be two functions $U \subset \mathbb{R}^n \to \mathbb{R}$, and let $a \in U$. If f and g are partially differentiable with respect to the j-th variable at a, then $f+g$ and $f\,g$ are as well and we have*

$$D_j(f+g)(a) = D_j f(a) + D_j g(a), \tag{3.2}$$
$$D_j(f\,g)(a) = D_j f(a) g(a) + f(a) D_j g(a). \tag{3.3}$$

If moreover $f(a) \neq 0$, then the function $1/f$ is partially differentiable at a with respect to the j-th variable and we have

$$D_j\left(\frac{1}{f}\right)(a) = -\frac{1}{f(a)^2} D_j f(a). \tag{3.4}$$

3.1.1 Multiple outputs

One can also speak about partial differentiability of vector-valued functions $f : U \subset \mathbb{R}^n \to \mathbb{R}^p$. For such a function we denote its *components* by f_i for $1 \leq i \leq p$. Namely:

$$f(x) = \begin{pmatrix} f_1(x) \\ \vdots \\ f_p(x) \end{pmatrix}.$$

Based on the analogous lemma for functions of a single variable, the following holds:

Lemma 3.8. *Let $U \subset \mathbb{R}^n$ open, $f : U \to \mathbb{R}^p$, and $a \in U$. Let $1 \leq j \leq n$. Then the following statements are equivalent.*

(a) *The function f is partially differentiable with respect to the j-th variable at the point a.*

(b) *For every $1 \leq i \leq p$ the component f_i is partially differentiable with respect to the j-th variable at the point a.*

If either of the above conditions are satisfied, then we have

$$D_j f(a) = \begin{pmatrix} D_j f_1(a) \\ \vdots \\ D_j f_p(a) \end{pmatrix} \in \mathbb{R}^p.$$

In particular, if $f : U \to \mathbb{R}^p$ is everywhere partially differentiable in the jth direction, we will be able to consider the function

$$D_j f : U \to \mathbb{R}^p.$$

3.2 Directional derivatives

The concept of a partial derivative can be seen as a special case of the concept of a directional derivative. We define the latter concept as follows.

Definition 3.9. Let $U \subset \mathbb{R}^n$ be an open subset and $f : U \to \mathbb{R}^p$ a function. Fix a point $a \in U$ and a direction $v \in \mathbb{R}^n$. The mapping f is called **directionally differentiable** at the point a in the direction v if the function $t \mapsto f(a + tv)$ is differentiable at $t = 0$.

The derivative

$$D_v f(a) := \frac{d}{dt} f(a + tv)\Big|_{t=0}$$

is in that case called the **directional derivative** of f at the point a in the direction v.

If f is directionally differentiable in direction v at all points we can then consider the function $D_v f : U \to \mathbb{R}^p$. △

Observe that $t \mapsto f(a + tv)$ is once again a function of one variable. It should be thought as the restriction of f to the line passing via a with direction v.

Remark 3.10. We note that the directional differentiability of f at a in the direction v is equivalent to the existence of the limit

$$\lim_{t \to 0} \frac{f(a + tv) - f(a)}{t}.$$

If this limit exists, its value equals the directional derivative $D_v f(a)$. △

For the directional derivative the principle of componentwise differentiation (Lemma 3.8) also holds:

Lemma 3.11. *Let $a \in U$ and $v \in \mathbb{R}^n$. Let $f : U \to \mathbb{R}^p$. Then the following statements are equivalent:*

(a) *the function f is directionally differentiable at a in the direction v;*

(b) *for every $1 \leq i \leq p$ the component function $f_i : U \to \mathbb{R}$ is directionally differentiable at a in the direction v.*

If conditions (a) and (b) are satisfied, then

$$D_v f(a) = \begin{pmatrix} D_v f_1(a) \\ \vdots \\ D_v f_p(a) \end{pmatrix}.$$

Partial differentiation can be regarded as directional differentiation in specific directions:

Lemma 3.12. *Let e_j be the j-th standard basis vector in $\mathbb{R}^n$. Then the following statements are equivalent.*

(a) *The partial derivative $D_j f(a)$ exists.*

(b) *The function f is directionally differentiable at a in the direction e_j.*

Moreover, in the case where (a) and (b) are true, we have

$$D_j f(a) = D_{e_j} f(a). \tag{3.5}$$

Proof. There exists a $\delta > 0$ such that $B(a;\delta) \subset U$. We introduce the open interval $I := (a_j - \delta, a_j + \delta)$ and define the function $\varphi : I \to \mathbb{R}^p$ by

$$\varphi(s) = f(a_1, \dots, a_{j-1}, s, a_{j+1}, \dots, a_n).$$

For the (ordinary) derivative of φ with respect to the variable s it follows from the chain rule for (ordinary) differentiation that φ is differentiable at a_j if and only if the function $t \mapsto \varphi(a_j + t)$ is differentiable at 0. The first statement is by definition equivalent to (a). The second statement is equivalent to (b) because for all $t \in (-\delta, \delta)$ we have $\varphi(a_j + t) = f(a + te_j)$. Moreover, by the chain rule for ordinary differentiation, if either (a) or (b) are true we then have

$$D_j f(a) = \varphi'(a_j) = \frac{d}{dt}\varphi(a_j + t)\bigg|_{t=0} = \frac{d}{dt} f(a + te_j)\bigg|_{t=0} = D_{e_j} f(a). \qquad \square$$

3.3 The total derivative

The partial and directional derivatives give us information about the function f, but only one direction at a time. Our next goal is to define an object encapsulating the "slope" of f, at a given point a, in all directions at once. In analogy with the one variable case, we expect such an object to be the best linear approximation to f at a.

Definition 3.13. Consider an open subset $U \subset \mathbb{R}^n$ and a function $f : U \to \mathbb{R}^p$. Fix a point $a \in U$. The function f is said to be **(totally) differentiable at** a if there exists a linear map $A : \mathbb{R}^n \to \mathbb{R}^p$ such that

$$\lim_{h \to 0} \frac{\|f(a+h) - f(a) - A(h)\|}{\|h\|} = 0. \tag{3.6}$$

The linear map A is called the **total derivative** of f at a. We denote it by $Df(a)$. The matrix representing it with respect to the standard basis of $\mathbb{R}^n$ is called the **Jacobian** of f at a. △

We have to show that this concept is indeed well-defined. This is accomplished in the following lemma, which shows that the directional derivatives uniquely determine the linear map $A : \mathbb{R}^n \to \mathbb{R}^p$ appearing in Equation (3.6).

Lemma 3.14. *Suppose the function f is differentiable at a and let A satisfy (3.6). Then, for every $v \in \mathbb{R}^n$, the function f is directionally differentiable at a in the direction v. The corresponding directional derivative is given by*

$$D_v f(a) = A(v).$$

In particular, A is uniquely determined.

Proof. The statement is clear for $v = 0$. We therefore assume that $v \neq 0$. By substituting $h = tv$ into (3.6) we find that

$$\lim_{t \to 0} \frac{\|f(a + tv) - f(a) - A(tv)\|}{|t|\|v\|} = 0.$$

From the linearity of A it follows that $A(tv) = tA(v)$, so also

$$\lim_{t \to 0} \|v\|^{-1} \left\| \frac{f(a + tv) - f(a)}{t} - A(v) \right\| = 0$$

and we conclude that

$$\lim_{t \to 0} \left(\frac{f(a + tv) - f(a)}{t} - A(v) \right) = 0,$$

so f is directionally differentiable at a in the direction v and the claimed formula holds. □

Corollary 3.15. *If f is totally differentiable at a, then it is directionally differentiable for all directions v. Moreover, the formula*

$$Df(a)(v) = D_v f(a) \tag{3.7}$$

holds.

Example 3.16. Let $U \subset \mathbb{R}^n$ be open and $f : U \to \mathbb{R}^p$ a map such that there exists a linear map $T : \mathbb{R}^n \to \mathbb{R}^p$ with $f = T|_U$. Then f is totally differentiable at every point $a \in U$, with total derivative $Df(a) = T$. Indeed, Definition 3.13 applies with $A = T$. △

Example 3.17. We consider the map $f : \mathbb{R}^n \to \mathbb{R}, x \mapsto \langle x, x \rangle$. Let $a \in \mathbb{R}^n$. Then for $h \in \mathbb{R}^n$ we have

$$f(a + h) - f(a) = \langle a + h, a + h \rangle - \langle a, a \rangle = 2\langle a, h \rangle + \langle h, h \rangle.$$

Let A be the linear map $\mathbb{R}^n \to \mathbb{R}$ given by $A(h) = 2\langle a, h \rangle$. You should verify that this map is indeed linear. Then

$$\|f(a + h) - f(a) - A(h)\| = \|h\|^2$$

so (3.6) holds. We conclude that f is totally differentiable at the point a with total derivative $Df(a) : \mathbb{R}^n \to \mathbb{R}$ given by $Df(a)(h) = 2\langle a, h \rangle$. It follows that f is also directionally differentiable at a in every direction $v \in \mathbb{R}^n$, with directional derivative

$$D_v f(a) = 2\langle a, v \rangle.$$

We can also derive this directly. Indeed, due to the bilinearity of the inner product we have $f(a + tv) = \langle a, a \rangle + 2t\langle a, v \rangle + t^2\langle v, v \rangle$, for $t \in \mathbb{R}$. This expression is differentiable with respect to t with derivative

$$\frac{d}{dt} f(a + tv) = 2\langle a, v \rangle + 2t\langle v, v \rangle.$$

Substituting $t = 0$ yields $D_v f(a) = 2\langle a, v \rangle$. △

According to Lemma 3.12 and Corollary 3.15, a totally differentiable function is also partially differentiable:

Corollary 3.18. *Let $f : U \to \mathbb{R}^p$ be totally differentiable at a. Its Jacobi matrix at a is given by*

$$Df(a)_{ij} = D_j f_i(a) \qquad (1 \le j \le n, 1 \le i \le p).$$

Proof. The derivative $Df(a)$ is a linear map $\mathbb{R}^n \to \mathbb{R}^p$ and can thus be interpreted as a matrix with p rows and n columns. The element in the i-th row and the j-th column is denoted by $Df(a)_{ij}$. As we know from linear algebra, this element is given by

$$Df(a)_{ij} = (Df(a)e_j)_i = (D_{e_j} f(a))_i,$$

by Corollary 3.15. Applying (3.5) and Lemma 3.8 we find that

$$Df(a)_{ij} = (D_j f(a))_i = D_j f_i(a). \qquad \square$$

Exercise 3.19. Consider an open $U \subset \mathbb{R}^n$ and a function $f : U \to \mathbb{R}^m$. Show that f is totally differentiable at $a \in U$ if and only if each component $f_i : U \to \mathbb{R}$ is totally differentiable at a. If that is the case, prove that the total derivative $Df(a)$ has the covectors $Df_i(a)$ as its rows. △

3.3.1 The total derivative as a function

Notation 3.20. Consider an open $U \subset \mathbb{R}^n$ and a function $f : U \to \mathbb{R}^p$. If f is totally differentiable at all $a \in U$ we will write $Df : U \to \mathrm{Lin}(\mathbb{R}^n, \mathbb{R}^p)$ for the function $a \mapsto Df(a)$. △

If $p = 1$, the total derivative $Df : U \to \mathrm{Lin}(\mathbb{R}^n, \mathbb{R})$ takes values in the dual space, so $Df(a)$ is a covector (i.e. a row). We will not make much use of it, but it is common in the literature to transpose it and see it as a (column) vector instead:

Definition 3.21. Suppose $f : U \to \mathbb{R}$ is totally differentiable at $a \in U$. The **gradient** of f at a is the vector $\mathrm{grad}(f)(a) := Df(a)^t = (D_1 f(a), \cdots, D_n f(a))$. If it exists at all points, we regard it as a function $\mathrm{grad}(f) : U \to \mathbb{R}^n$. △

3.3.2 The first order Taylor polynomial

Given a continuous function $f : U \to \mathbb{R}^p$ and a point $a \in U$, the constant function that best approximates f at a is $x \mapsto f(a)$. In this case, a "good zero order approximation" means that $\lim_{h\to 0} f(a + h) - f(a) = 0$, i.e. the definition of continuity at a.

The following observation tells us that f is differentiable at a if and only if there is a first order polynomial that is a "good first order approximation" around a. The meaning of this is explained in the following proposition, which follows immediately from the definition of total derivative. (Do note that it is more restrictive than being a "good zero order approximation").

Proposition 3.22. *The following are equivalent for a function $f : U \to \mathbb{R}^p$:*

- *f is differentiable at $a \in U$ with total derivative $Df(a) \in \mathrm{Lin}(\mathbb{R}^n, \mathbb{R}^p)$.*
- *The function $R(h) = f(a + h) - [f(a) + Df(a)(h)]$ satisfies $\lim_{h\to 0} \frac{\|R(h)\|}{\|h\|} = 0$.*

Lemma 3.14 implies that there is a unique first order polynomial approximating f well:

Definition 3.23. Suppose $f : U \subset \mathbb{R}^n \to \mathbb{R}^p$ is differentiable at a. The function $P_a^1(f) : \mathbb{R}^n \to \mathbb{R}^p$ given by $P_a^1(f)(x) := f(a) + Df(a)(x - a)$ is called the **first order Taylor polynomial** of f at a. The function

$$R(h) := f(a + h) - P_a^1(f)(a + h) = f(a + h) - [f(a) + Df(a)(h)]$$

is called the **remainder**. △

That is, the remainder measures the error in approximating f by $P_a^1(f)$.

Example 3.24. We consider the function $f : \mathbb{R}^2 \to \mathbb{R}$ defined by $f(x) = x_1x_2$. Let $a \in \mathbb{R}^2$ be a given fixed point. Then for all $h \in \mathbb{R}^2$ we have

$$f(a + h) - f(a) = (a_1 + h_1)(a_2 + h_2) - a_1a_2 = a_2h_1 + a_1h_2 + h_1h_2 = A(h) + R(h),$$

with $A(h) = (a_2 \quad a_1)(h_1, h_2)^t$ and $R(h) = h_1h_2$. (The important step is to 'split off' the appropriate linear mapping A.) The defined mapping A is linear, and for R we have that $|R(h)| \leq |h_1||h_2| \leq \|h\|^2$, hence

$$\lim_{h \to 0} \frac{|R(h)|}{\|h\|} = 0.$$

We see thus that the mapping f is totally differentiable at a, with derivative $Df(a) : \mathbb{R}^2 \to \mathbb{R}$ given by the row-matrix $A = (a_2 \quad a_1)$. By Corollary 3.18 it follows that

$$D_1f(a) = a_2 \qquad \text{and} \quad D_2f(a) = a_1.$$

This can also be directly derived from the rules for partial differentiation. The first Taylor polynomial is then

$$P_a^1(f)(a + h) = a_1a_2 + (a_2h_1 + a_1h_2),$$

which is seen to indeed be of first order on h. △

3.3.3 The total derivative in one variable

We now compare the newly introduced concept of total derivative in the case $n = 1$ with the ordinary derivative.

Lemma 3.25. *Let $I \subset \mathbb{R}$ be an open interval and $f : I \to \mathbb{R}^p$ a function. Let $a \in I$. Then the following two statements are equivalent.*

(a) *The function f is differentiable at a in the sense of* Inleiding Analyse.

(b) *The function f is totally differentiable at a.*

If f is differentiable at a, then the relation between the two derivatives is given by

$$f'(a) = Df(a)(1). \tag{3.8}$$

Proof. First assume that (b) holds. Then it follows from Corollary 3.18 that f is partially differentiable. Since there is only one variable, f is simply differentiable with respect to that variable. This implies (a).

Now assume conversely that (a) holds, i.e. that f is differentiable at a with derivative $f'(a) \in \mathbb{R}^p$. Then

$$\lim_{h\to 0}\left(\frac{f(a+h)-f(a)}{h} - f'(a)\right) = 0.$$

It follows that

$$f(a+h) - f(a) = f'(a)h + R(h)$$

with $\lim_{h\to 0} h^{-1}R(h) = 0$ hence also

$$\lim_{h\to 0}\frac{\|R(h)\|}{|h|} = \lim_{h\to 0}\left\|\frac{R(h)}{h}\right\| = 0.$$

The mapping $A : \mathbb{R} \to \mathbb{R}^p$ defined by $A(h) = hf'(a)$ is linear. From Proposition 3.22 it now follows that f is totally differentiable at a with derivative $Df(a) = A$. We conclude that (b) holds, and furthermore that $Df(a)(1) = A(1) = f'(a)$. □

Remark 3.26. The relation (3.8) can also be understood as follows. The linear mapping $Df(a) : \mathbb{R} \to \mathbb{R}^p$ has as Jacobian matrix the column

$$Df(a) = (f_1'(a), \dots, f_p'(a)) \in \mathbb{R}^p.$$

If we let $Df(a)$ act on the point $1 \in \mathbb{R}$ (which can be seen as the standard basis vector e_1 of $\mathbb{R}$) then we find that

$$Df(a)(1) = f'(a). \qquad \triangle$$

3.3.4 Total differentiability implies continuity

We can now use the first order Taylor polynomial (Proposition 3.22) and the estimates for linear maps developed in Subsection 2.7 to obtain the following very important consequence.

Proposition 3.27. *Fix an open subset $U \subset \mathbb{R}^n$, a function $f : U \to \mathbb{R}^p$, and a point $a \in U$. Suppose that f is totally differentiable at a. Then f is continuous at the point a.*

Proof. Consider the remainder at a:

$$R(h) := f(a+h) - [f(a) + Df(a)(h)], \qquad (h \in -a + U).$$

The idea of the proof is that the Taylor polynomial is continuous (by polynomiality, according to Proposition 2.71) and f differs from it by R, which is also continuous (at $h = 0$).

Indeed, for all $h \in -a + U$ we have that:

$$\begin{aligned} \|f(a+h) - f(a)\| &= \|Df(a)(h) + R(h)\| \\ &\le \|Df(a)(h)\| + \|R(h)\| \\ &\le \|Df(a)\|\,\|h\| + \|R(h)\|. \end{aligned} \tag{3.9}$$

From the definition of differentiability it follows that $\|h\|^{-1}\|R(h)\|$ has limit 0 as $h \to 0$. This means that $\|R(h)\|$ itself has limit 0 as $h \to 0$. In detail: First, note that $R(0) = 0$. Secondly, there exists a $\delta > 0$ such that $a + B(0;\delta) = B(a;\delta) \subset U$ and such that for all $h \in B(0;\delta)$ we have $\|h\|^{-1}\|R(h)\| \le 1$. From this it follows that $\|R(h)\| \le \|h\|$, so the limit is indeed zero.

If we now go back to Equation (3.9), we see that for all $h \in B(0;\delta)$ we have

$$\|f(a+h) - f(a)\| \le (\|Df(a)\| + 1)\|h\|,$$

showing that $\|f(a+h) - f(a)\| \to 0$ as $h \to 0$, proving continuity at a. □

In contrast, the following (important!) example shows that partial differentiability does not imply continuity (and therefore it does not imply total differentiability):

Example 3.28. Define $f : \mathbb{R}^2 \to \mathbb{R}$ (Figure 3.3.4) by

$$f(x, y) := \frac{xy^2}{x^2 + y^4}$$

if $(x, y) \neq (0, 0)$ and by $f(0, 0) := 0$. Then, f is partially differentiable with continuous partial derivatives on $\mathbb{R}^2 \setminus \{(0, 0)\}$.

Take $v = (a, b) \in \mathbb{R}^2$ with $a \neq 0$. Then

$$\frac{f(t.v) - f(0)}{t} = \frac{t^3 ab^2}{t(t^2 a^2 + t^4 b^4)} = \frac{ab^2}{a^2 + t^2 b^4} \longrightarrow \frac{ab^2}{a^2} = \frac{b^2}{a}$$

as $t \neq 0$ and $t \to 0$. On the other hand, if $b \neq 0$ and $t \neq 0$, then $\frac{f(t.0,t.b) - f(0,0)}{t} = 0$. Thus, at the point $(0, 0)$, the function f is directionally differentiable in all directions, with $D_v f(0,\, 0) = b^2/a$ if $v = (a,\, b)$ and $a \neq 0$, and $D_v f(0,\, 0) = 0$ if $v = (0,\, b)$. In particular, f is partially differentiable at $(0, 0)$ with partial derivatives $D_1 f(0, 0) = D_2 f(0, 0) = 0$.

It is clear that the directional derivative $D_v f(0, 0)$ does not depend linearly on the direction vector $v = (a, b)$. In view of Remark 3.15 we conclude that f cannot be totally differentiable at the point $(0, 0)$.

It is perhaps surprising that the function f, despite the existence of the partial derivatives at $(0, 0)$, is not continuous at that point. We see this as follows. For every $c \in \mathbb{R}$ and $y \neq 0$ we have

$$f(cy^2, y) = \frac{cy^4}{c^2 y^4 + y^4} = \frac{c}{c^2 + 1}.$$

Suppose that the function f is continuous at $(0, 0)$. As $y \to 0$, we have that $(cy^2, y) \to (0, 0)$, so the continuity of f at the point $(0, 0)$ would imply that $f(cy^2, y)$ converges to 0 as $y \to 0$. However, if $c \neq 0$ then $f(cy^2, y)$ is equal to the constant $c/\big(c^2 + 1\big) \neq 0$ for every $y \neq 0$. Contradiction.

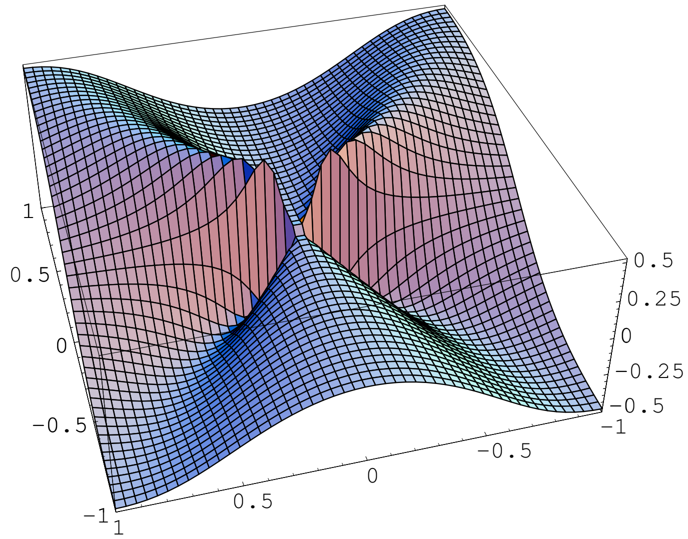
In summary, this is an example of a function that is differentiable in each variable at every point, but is not continuous at the origin. △

3.3.5 Continuous differentiability

The following result provides the most commonly used criterion to conclude the total differentiability of functions. It is one of the most important results of the chapter.

Theorem 3.29. *Let $U \subset \mathbb{R}^n$ be an open set, $f : U \to \mathbb{R}^p$ a function, and $a \in U$ a point. Suppose that f is partially differentiable and all its partial derivatives $D_j f : U \to \mathbb{R}^p$ are continuous at a. Then f is totally differentiable at a.*

You should compare this to Example 3.28. There we saw that partial differentiability does *not* imply total differentiability. The theorem says that *continuous partial differentiability* does.



Figuur 3: $f(x, y) = x\,y^2/\left(x^2 + y^4\right)$ for $-1 < x < 1$, $-1 < y < 1$.

We can also consider:

Definition 3.30. Consider an open subset $U \subset \mathbb{R}^n$ and a function $f : U \to \mathbb{R}^p$. We say that f is **continuously differentiable**, or simply C^1, if it is totally differentiable everywhere and $Df : U \to \operatorname{Lin}(\mathbb{R}^n, \mathbb{R}^p)$ is continuous. △

Theorem 3.29 immediately implies that:

Corollary 3.31. *Suppose $f : U \subset \mathbb{R}^n \to \mathbb{R}^p$ is partially differentiable and all its partial derivatives $D_j f : U \to \mathbb{R}^p$ are continuous. Then f is C^1.*

The proof of Theorem 3.29 requires some preparation. We will build towards it and provide a proof in Subsection 3.5.

3.4 Growth of functions

The total derivative measures the slope of a function in all directions. As such, it gives us an idea of how much this function may grow as we move to a nearby point. We explore this idea in the upcoming paragraphs.

3.4.1 Extrema

The simplest situation one could encounter is for the total derivative to vanish at a point a. This means that the function f has zero slope in all directions, so its value does not change very much around a:

Definition 3.32. Let the function $f : U \to \mathbb{R}$ be totally differentiable at $a \in U$. We say that $a \in U$ is **stationary** or **critical** if $Df(a) = 0$. △

In many situations we are interested in *local extrema* of functions of several variables. According to the *variational principle*, these are critical points:

Proposition 3.33. *Let $U \subset \mathbb{R}^n$ be open, $a \in U$, and $f : U \to \mathbb{R}$ directionally differentiable at $a \in U$ along the direction $v \in \mathbb{R}^n$. If f has a local maximum (resp. minimum) at a then $D_v f(a) = 0$.*
In particular, if f is totally differentiable at a, it holds that $Df(a) = 0$.

Proof. The assumption implies that the function of one variable $\phi(t) := f(a + tv)$ has a local maximum/minimum at the point $t = 0$. From the theory of differentiable functions of one variable it is known that this implies

$$D_v f(a) = \phi'(0) = 0. \qquad \square$$

Example 3.34. Recall the standard quadratic forms in $\mathbb{R}^2$ from Example 2.43. In Exercise 2.85 we characterised their extrema. We now revisit them from the point of view of differentiation.

Let $f(x, y) = x^2 + y^2 : \mathbb{R}^2 \to \mathbb{R}$. Then f has a minimum at $(0, 0)$, hence $D(f)(0, 0) = 0$. This can also be verified by computing the partial derivatives:

$$\partial f(x, y)/\partial x = 2x = 0 \quad \text{when} \quad x = 0,$$

$$\partial f(x, y)/\partial y = 2y = 0 \quad \text{when} \quad y = 0.$$

Now take $f(x, y) = x^2 - y^2$. Then

$$Df(x, y) = (2x \quad -2y).$$

So also in this case $D(f)(0, 0) = 0$, so $(0, 0)$ is a stationary point of f. However, f does not have a local minimum at $(0, 0)$ because arbitrarily close to $(0, 0)$ there are points (x, y) with $f(x, y) < 0 = f(0, 0)$, for example take $x = 0$ and $y \neq 0$. On the other hand, by considering points (x, y) with $y = 0$ and $x \neq 0$, we see that f also does not have a local maximum at $(0, 0)$. A stationary point of f in which f has neither a local minimum nor a local maximum is also called a *saddle point* of f. △

3.4.2 Locally constant functions

We studied locally constant functions, in general metric spaces, in Subsection 1.2.3. Those results can be refined if we restrict ourselves to differentiable functions.

From *Inleiding Analyse* it is known that for a differentiable function $f : [a, b] \to \mathbb{R}^p$ of one variable it holds: if $f' = 0$ then f is constant. You can use this fact to then show:

Exercise 3.35. Let for each $1 \leq j \leq n$ an open interval $I_j = (a_j, b_j)$ with $a_j < b_j$ be given. Then $V = I_1 \times \cdots \times I_n$ is an open subset of $\mathbb{R}^n$. If $f : V \to \mathbb{R}^p$ is a partially differentiable function with $D_j f = 0$ for all $1 \leq j \leq n$, then f is constant. △

Which implies:

Exercise 3.36. Consider an arbitrary open $U \subset \mathbb{R}^n$. Let $f : U \to \mathbb{R}^p$ be a partially differentiable function with $D_j f = 0$ for all $1 \leq j \leq n$. Then f is locally constant. △

Applying Proposition 1.30 we thus deduce:

Proposition 3.37. *Let $f : U \to \mathbb{R}^p$ be a partially differentiable function with $D_j f = 0$ for all $1 \leq j \leq n$. If X is path-connected, then f is constant.*

Observe that this result is an elementary case of our main Theorem 3.29. Namely, we assume that f has partial derivatives that are zero (and are thus continuous) and we deduce that f is constant (so is in particular totally differentiable).

Example 3.38. Suppose $f : \mathbb{R}^2 \to \mathbb{R}$ satisfies $Df(x, y) = (xe^{xy} \quad ye^{xy})$. Then f is necessarily $g(x, y) = e^{xy}$ plus a constant. The reason is that $Df = Dg$ so $D(f - g) = 0$, which implies (Proposition 3.37) that $f - g$ is constant. △

3.4.3 The mean value theorem

The most important result relating the growth of a differentiable function to the size of its total derivative is the mean value theorem. Observe that it is a *global* statement: knowing the slope everywhere allows us to estimate the size of the function everywhere. You can think of it as a generalisation of Proposition 3.37, which is the slope zero case.

Recall the case of one variable:

Lemma 3.39. *Let $I \subset \mathbb{R}$ be an interval and $\varphi : I \to \mathbb{R}$ a differentiable function. Then for all $a, b \in I$ there exists a $c \in [a, b]$ such that*

$$\varphi(b) - \varphi(a) = \varphi'(c)(b - a). \tag{3.10}$$

Proof. For $a < b$ this result is a consequence of the mean value theorem proven in the notes of *Inleiding Analyse*. The identity (3.10) then even holds for a $c \in (a, b)$. For $a = b$ the result is evident. For $a > b$ the result is a consequence of the mean value theorem applied to the interval $[b, a]$. □

This result can then be applied to deduce the standard *multivariate mean value theorem*:

Theorem 3.40. *Let $U \subset \mathbb{R}^n$ be open and $f : U \to \mathbb{R}$ a function that is totally differentiable everywhere. Then, for any two points $p, q \in U$ with $[p, q] \subset U$ there exists a point $r \in [p, q]$ such that*

$$f(q) - f(p) = Df(r)(q - p).$$

Proof. Write $v := q - p$ and consider the function of one variable $\varphi : [0, 1] \to \mathbb{R}$ defined by

$$\varphi(t) := f(p + tv).$$

This will allow us to reduce to the case of one variable.

We first claim that φ is differentiable everywhere. Indeed, fix a time $t_0 \in [0, 1]$. Given $t \in -t_0 + [0, 1]$ we have that

$$\varphi(t_0 + t) = f(p + t_0 v + tv).$$

Since f is totally differentiable at $p + t_0 v \in [p, q] \subset U$, it follows that f is also directionally differentiable at that point in the direction v, with directional derivative $D_v f(p + t_0 v) = Df(p + t_0 v)(v)$. We conclude that

$$\varphi'(t_0) = \lim_{t \to 0} \frac{\varphi(t_0 + t) - \varphi(t_0)}{t} = Df(p + t_0 v)(v).$$

The final step is to apply the mean value theorem for ordinary differentiation. It implies that there exists a $c \in [0,1]$ such that

$$f(q) - f(p) = \varphi(1) - \varphi(0) = \varphi'(c) = Df(p + cv)(v).$$

This gives the result with $r = p + cv \in [p,q]$. □

This result can be used to bound the growth of a differentiable function along a segment:

Corollary 3.41. *Suppose $U \subset \mathbb{R}^n$ is an open subset and $f : U \to \mathbb{R}$ is a totally differentiable function. We suppose further that there is $M \geq 0$ such that $\|Df(x)\| \leq M$ for all $x \in U$. Then, for any two points $p, q \in U$ with $[p,q] \subset U$ it holds that:*

$$|f(q) - f(p)| \leq M\|q - p\|.$$

Proof. According to Theorem 3.40, there is an $r \in [p,q]$ so that:

$$|f(q) - f(p)| = |Df(r)(q - p)| \leq \|Df(r)\|\|q - p\| \leq M\|q - p\|. \qquad \square$$

Exercise 3.42. Show that the same statement is not necessarily true if we do not assume that $[p,q] \subset U$. △

You should note that Theorem 3.40 deals with functions with values in $\mathbb{R}$, not a higher euclidean space. The same holds for Corollary 3.41. Nonetheless, one can generalise Corollary 3.41 to functions $U \to \mathbb{R}^p$:

Exercise 3.43. Let $U \subset \mathbb{R}^n$ be open and $f : U \to \mathbb{R}^p$ a function that is totally differentiable everywhere. Suppose further that $a, b \in U$ and that $[a,b] \subset U$. Finally suppose that $M > 0$ and that $\|Df(x)\| \leq M$ for all $x \in [a,b]$.

(a) Show that for all $v \in \mathbb{R}^p$ it holds that

$$|\langle f(b) - f(a), v\rangle| \leq M\|v\|\|b - a\|.$$

Hint: consider the function $F : U \to \mathbb{R}$ given by $x \mapsto \langle f(x), v\rangle$. It measures the size of $f(x)$ in the direction of v.

(b) Show that

$$\|f(b) - f(a)\| \leq M\|b - a\|.$$

Hint: apply (a) with a suitable choice of $v \in \mathbb{R}^p$. △

3.5 The proof of Theorem 3.29

Our goal now is to establish Theorem 3.29, which shows total differentiability once we have continuous partial derivatives. The strategy is the following: we will show that, in the presence of continuous partial derivatives, it is possible to deduce a partial-derivative-version of the mean value theorem (Proposition 3.45). Once we have that, the theorem will follow.

We first show that having one partial derivative gives us a directional version of the mean value theorem:

Lemma 3.44. *Consider an open subset $U \subset \mathbb{R}^n$, a function $f : U \to \mathbb{R}$ that is partially differentiable on U in direction j, and a pair of points $p, q \in U$ satisfying $[p, q] \subset U$.*

Suppose further that $p - q$ is a multiple of e_j (identically, $p_i = q_i$ for each $i \neq j$). Then there exists an $r \in [p, q]$ such that

$$f(q) - f(p) = D_j f(r) \cdot (q_j - p_j). \tag{3.11}$$

Proof. We consider the function $\varphi : [p_j, q_j] \to \mathbb{R}$ defined by

$$\varphi(t) = f(p_1, \dots, p_{j-1}, t, p_{j+1}, \dots, p_n).$$

Then the function φ is differentiable, $\varphi(p_j) = f(p)$, and $\varphi(q_j) = f(q)$. By the mean value theorem for ordinary differentiability there exists a $c \in [p_j, q_j]$ such that

$$f(q) - f(p) = \varphi(q_j) - \varphi(p_j) = \varphi'(c)(q_j - p_j) = D_j f(r)(q_j - p_j),$$

where $r := (p_1, \dots, p_{j-1}, c, p_{j+1}, \dots p_n)$. □

This lemma can be applied one direction at a time to establish the promised mean value theorem:

Proposition 3.45. *Let $U \subset \mathbb{R}^n$ be open and $a \in U$. Let $f : U \to \mathbb{R}$ be a function that is partially differentiable and, moreover, for each $1 \leq j \leq n$, the partial derivative $D_j f : U \to \mathbb{R}$ is continuous at a.*

Then there is an open ball $B \subset U$ with center a and a function $L : B \to \mathrm{Lin}(\mathbb{R}^n, \mathbb{R})$ such that

$$f(x) - f(a) = L(x)(x - a) = \begin{pmatrix} L_1(x) & \cdots & L_n(x) \end{pmatrix}(x - a) = \sum_{j=1}^{n} L_j(x)(x_j - a_j)$$

and $\lim_{x \to a} L_j(x) = D_j f(a)$ for every $1 \leq j \leq n$ and $x \in B$.

You should keep in mind is that we do not know yet whether f is totally differentiable at a (that is what we are aiming to show). However, it if was, the covector $\begin{pmatrix} D_1 f(a) & \cdots & D_n f(a) \end{pmatrix}$ would be the total derivative. What this result says is that we can construct a family of covectors $L : B \to \mathrm{Lin}(\mathbb{R}^n, \mathbb{R})$ that interact with the function f as the actual total derivative would.

Proof. First observe that the open ball $B(a, \delta)$ is contained in U if δ is sufficiently small, because U is open; this yields the claimed B. Fix $x \in B$. The idea of the proof is to decompose $f(x) - f(a)$ into a sum of differences of function values, where each time only one of the variables is varied. We will then be able to apply Proposition 3.45 to each term.

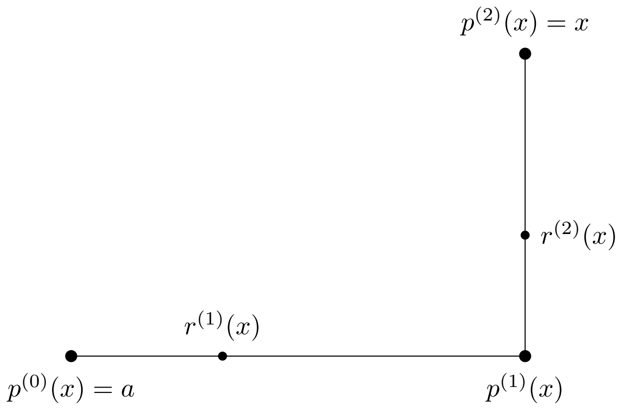
Define a sequence of points by setting $p^{(0)}(x) := a$ and, for each $1 \leq j \leq n$:

$$p^{(j)}(x) := (x_1, \, \dots, x_j, \, a_{j+1}, \, \dots, \, a_n). \tag{3.12}$$

For $j = n$ we interpret this so that $p^{(n)}(x) = x$. The successive points $p^{(j-1)}(x)$ and $p^{(j)}(x)$ differ only in the j-th coordinate. The connecting line segments $[p^{(j-1)}(x), p^{(j)}(x)]$ together form a piecewise path from a to x with only one coordinate varying in each segment.

We claim that the points $p^{(j)}(x)$ lie in the ball $B = B(a, \delta) \subset U$. Indeed:

$$\|p^{(j)}(x) - a\| = \left(\sum_{k=1}^{j} (x_k - a_k)^2 \right)^{1/2} \leq \|x - a\| \leq \delta.$$



Figuur 4: The piecewise path from a to x in the planar case ($n = 2$).

As such, the same holds for the connecting line segments, since B is convex.

We then write the difference $f(x) - f(a)$ as

$$f(x) - f(a) = f(p^{(n)}(x)) - f(p^{(0)}(x)) = \sum_{j=1}^{n} (f(p^{(j)}(x)) - f(p^{(j-1)}(x)). \tag{3.13}$$

We now focus on rewriting the j-th term. The points $p^{(j-1)}(x)$ and $p^{(j)}(x)$ differ solely in the j-th coordinate. Concretely:

$$p^{(j)}(x)_j - p^{(j-1)}(x)_j = (x_j - a_j).$$

Since the line segment $[p^{(j-1)}(x), p^{(j)}(x)]$ lies entirely in U, Proposition 3.45 implies that there exists an intermediate point $r^{(j)}(x) \in [p^{(j-1)}(x), p^{(j)}(x)]$ satisfying

$$f(p^{(j)}(x)) - f(p^{(j-1)}(x)) = D_j f(r^{(j)}(x))(x_j - a_j). \tag{3.14}$$

Plugging this into Equation (3.13) it follows that

$$f(x) - f(a) = \sum_{j=1}^{n} D_j f(r^{(j)}(x))(x_j - a_j).$$

This leads us to define the functions $L_j : B \to \mathbb{R}$ using the expressions

$$L_j(x) := D_j f(r^{(j)}(x)). \tag{3.15}$$

We put them together as the covector valued function $L : B \to \mathrm{Lin}(\mathbb{R}^n, \mathbb{R})$, yielding:

$$f(x) - f(a) = \sum_{j=1}^{n} L_j(x) \cdot (x_j - a_j). = L(x)(x - a).$$

It remains to study the behaviour of $L(x)$ as $x \to 0$. Since $\|r^{(j)}(x) - a\| \leq \|x - a\|$, it follows that $r^{(j)}(x) \to a$ as $x \to a$. On the other hand, the partial derivative $D_j f$ is continuous at a. It therefore follows that

$$L_j(x) = D_j f(r^{(j)}(x)) \to D_j f(a) \qquad (x \to a).$$ □

Without further ado:

Proof of Theorem 3.29. It is enough to establish the result for the case $f : U \to \mathbb{R}$, i.e. f takes values in $\mathbb{R}$. The general case is then a consequence of Exercise 3.19.

Let $\delta > 0$ such that $B = B(a, \delta) \subset U$. We use the mean value theorem of Proposition 3.45 to deduce that there is some $L : B \to \mathrm{Lin}(\mathbb{R}^n, \mathbb{R})$ such that, for each $h \in B(0, \delta)$ (thus $a + h \in B$):

$$f(a+h) - f(a) = L(a+h)(h) = \sum_{j=1}^{n} L_j(a+h)h_j.$$

We define a linear map $A : \mathbb{R}^n \to \mathbb{R}$ by

$$A(h) := \sum_{j=1}^{n} D_j f(a) h_j = (D_1 f(a) \quad \cdots \quad D_n f(a))(h)$$

and we consider the remainder:

$$R(h) := f(a+h) - f(a) - A(h).$$

Then for $h \in B(0, \delta)$ we have

$$R(h) = L(a+h)(h) - A(h).$$

Using linearity and Cauchy-Schwarz (Proposition 2.75) we deduce that:

$$\frac{|R(h)|}{||h||} = \frac{||L(a+h) - A||.||h||}{||h||} = ||L(a+h) - A||$$

which goes to zero as h goes to zero, according to Proposition 3.45. Proposition 3.22 then implies that f is differentiable at a with total derivative A. □

3.6 Computation rules for the total derivative

In practice, functions are often given to us as combinations of simpler functions. The following results allow us to compute their total derivative.

3.6.1 The sum, product, and quotient rules for the total derivative

Proposition 3.46. *Fix an open $U \subset \mathbb{R}^n$. Let $f, g : U \to \mathbb{R}$ be totally differentiable at the point $a \in U$. Then $f + g$ and fg are totally differentiable at a and*

$$D(f+g)(a) = Df(a) + Dg(a), \tag{3.16}$$

$$D(fg)(a) = g(a)Df(a) + f(a)Dg(a). \tag{3.17}$$

Moreover, if $f(a) \neq 0$, then $1/f$ is totally differentiable at a and

$$D(1/f)(a) = -\frac{Df(a)}{f(a)^2}. \tag{3.18}$$

Note that, in products, we always place scalars before the linear maps; this is the usual order. Writing them in the opposite order could suggest erroneously that one thinks of $Df(a)$ as a number, instead of as a linear map $\mathbb{R}^n \to \mathbb{R}$.

Remark 3.47. If f and g are continuously differentiable on $U \subset \mathbb{R}^n$, then Proposition 3.46 follows from Proposition 3.7 and Theorem 3.29. Alternatively, Proposition 3.46 can be proved directly from Definition 3.13.

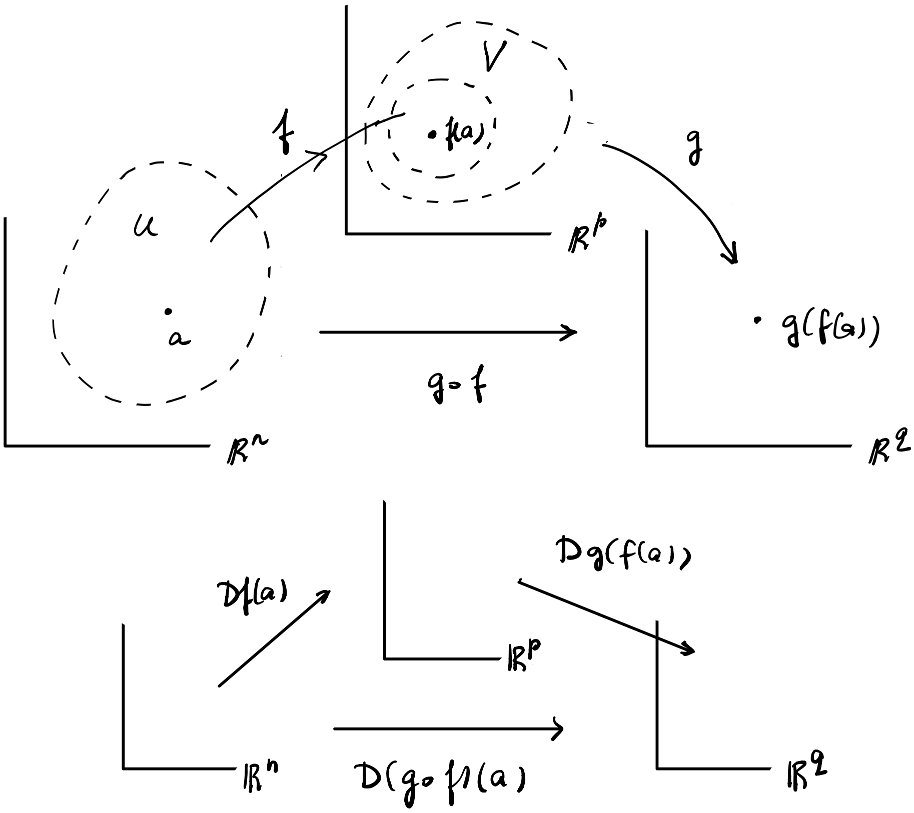
The computation rules for totally differentiable vector-valued functions follow by applying Proposition 3.46 to each coordinate function. △

3.6.2 The chain rule

The following is the *chain rule for total derivatives*, one of the main results of this chapter.

Theorem 3.48. *Let U be an open subset of $\mathbb{R}^n$ and V an open subset of $\mathbb{R}^p$. Let $f : U \to V$ be a map that is totally differentiable at the point $a \in U$. Let $g : V \to \mathbb{R}^q$ be totally differentiable at $f(a) \in V$. Then, the composition $g \circ f : U \to \mathbb{R}^q$ is totally differentiable at a. Moreover:*

$$D(g \circ f)(a) = Dg(f(a)) \circ Df(a). \tag{3.19}$$



Figuur 5: The chain rule.

Proof. The strategy of the proof is very similar to how we established Theorem 3.29. We will apply the "weak" mean value theorem (Proposition 3.45) to f and g. The results can be combined to produce a "mean value theorem" for the composition. This will be enough to invoke the characterisation of differentiability in terms of the first order Taylor polynomial (Proposition 3.22), showing that $g \circ f$ is differentiable.

Write $b = f(a)$. According to Proposition 3.45, there are balls $B_U \subset U$ and $B_V \subset V$, containing a and b, respectively, and functions $L : B_U \to \text{Lin}(\mathbb{R}^n, \mathbb{R}^p)$ and $H : B_V \to \text{Lin}(\mathbb{R}^p, \mathbb{R}^q)$ such that:

$$\begin{aligned} f(a+u) - f(a) &= L(a+u)(u) \\ g(b+v) - g(b) &= H(b+v)(v) \end{aligned}$$

that moreover satisfy $\lim_{u\to 0} L(a+u) = Df(a)$ and $\lim_{v\to 0} H(b+v) = Dg(b)$.

We therefore obtain the following mean value result:

$$g(f(a+u)) - g(f(a)) = H(f(a+u))(f(a+u) - f(a)) = H(f(a+u))(L(a+u)(u));$$

do note that in the formula above we set $v = f(a+u) - f(a)$. The limits above tell us that

$$lim_{u\to 0} H(f(a+u)) \circ L(a+u) = Dg(f(a)) \circ Df(a);$$

we also invoked the continuity of f.

Now the proof is complete, using Proposition 3.22, if we show that the remainder:

$$\begin{aligned} ||R(u)|| &= ||g \circ f(a+u) - [g \circ f(a) + Dg(f(a)) \circ Df(a)(u)]|| \\ &= ||H(f(a+u))(L(a+u)(u)) - Dg(f(a)) \circ Df(a)(u)|| \\ &\leq ||H(f(a+u)) \circ L(a+u) - Dg(f(a)) \circ Df(a)|| \cdot ||u|| \end{aligned}$$

satisfies $\lim_{u\to 0} ||R(u)||/||u|| = 0$, which follows from the limit we computed above. □

Remark 3.49. Equation (3.19) expresses $D(g\circ f)(a)$ as the matrix multiplication of $Dg(f(a))$ with $Df(a)$. If you spell out what this means at the level of matrix coefficients, we get the following expression for the chain rule of the jth directional derivative:

$$D_j(g \circ f)(a) = \sum_{i=1}^{p} D_i g(f(a)) D_j f_i(a).$$

You will often encounter it in classical notation as well:

$$\frac{\partial (g \circ f)_h}{\partial x_j}(a) = \sum_{i=1}^{p} \frac{\partial g_h}{\partial y_i}(f(a)) \frac{\partial f_i}{\partial x_j}(a) \tag{3.20}$$

for $1 \leq h \leq q$ and $1 \leq j \leq n$. Here we are spelling it out for the hth entry of $g \circ f$. Here $x = (x_1, \ldots, x_n)$ denotes the variables in $\mathbb{R}^n$ and $y = (y_1, \ldots, y_p)$ the variables in $\mathbb{R}^p$. The way to remember it is as follows: in order to differentiate $g \circ f$ with respect to x_j, you differentiate g with respect to each variable y_i and in turn you differentiate each $y_i = f_i(x)$ with respect to x_j.

Let us provide a proof. The first matrix $D(g \circ f)(a)$ has entries $D_j(g \circ f)_h(a)$. The second matrix $Dg(f(a))$ has entries $D_i g_h(f(a))$. The third matrix $Df(a)$ has entries $D_j f_i(a)$. Here $1 \leq h \leq q$, $1 \leq i \leq p$ and $1 \leq j \leq n$. Applying matrix multiplication we get the claimed equations. Note that $(g \circ f)(x) = (g_1(f(x)), \ldots, g_q(f(x)))$, so $(g \circ f)_h = g_h \circ f$ for all $1 \leq h \leq q$. △

3.6.3 The chain rule along a curve

An important special case of the chain rule arises when a function is differentiated "along a curve". We already saw continuous curves in Definition 1.16, but from now on we will mostly focus on the case of:

Definition 3.50. Suppose $[a,b] \subset \mathbb{R}$ is an interval and $U \subset \mathbb{R}^n$ is an open subset. If the curve $\gamma : [a,b] \to U$ is differentiable we can consider its **velocity** $\gamma' : [a,b] \to \mathbb{R}^n$. △

We think of $\gamma'(t) \in \mathbb{R}^n$ as the direction of motion for γ at time t. In applications we will be mostly concerned with the case in which γ is C^1, so γ' is continuous and thus itself a path.

Consider now a function $g : U \to \mathbb{R}^p$. The composition $g \circ \gamma : I \to \mathbb{R}^p$ is also called the *function g along the curve γ* or simply the restriction of g to γ. By differentiating g along γ we mean differentiating the composition $g \circ \gamma$.

Lemma 3.51. *Let $U \subset \mathbb{R}^n$ be open and $I \subset \mathbb{R}$ be an interval. Let $\gamma : I \to U$ be a curve in U and $g : U \to \mathbb{R}^p$ a function. If γ is differentiable at a point $t_0 \in I$ and g is differentiable at $\gamma(t_0)$, then $g \circ \gamma : I \to \mathbb{R}^p$ is differentiable at t_0 and it holds that*

$$(g \circ \gamma)'(t_0) = Dg(\gamma(t_0))(\gamma'(t_0)) = D_{\gamma'(t_0)} g(\gamma(t_0)). \tag{3.21}$$

I.e. $(g \circ \gamma)'(t_0)$ is the directional derivative of g along $\gamma'(t_0)$ at $\gamma(t_0)$.

Proof. From the chain rule for the total derivative (Theorem 3.48) it follows that $g \circ \gamma$ is totally differentiable at t_0, with derivative

$$D(g \circ \gamma)(t_0) = Dg(\gamma(t_0)) \circ D\gamma(t_0).$$

Applying both sides to the element $e_1 = 1 \in \mathbb{R}$ now leads, in view of Lemma 3.25, to the formula (3.21). □

In the case $p = 1$, g is a scalar function and the formula (3.21) can then be rewritten as

$$(g \circ \gamma)'(t_0) = \sum_{j=1}^{n} \frac{\partial g}{\partial x_j}(\gamma(t_0))\gamma_j'(t_0). \tag{3.22}$$

As an example of how these computations look in practice, consider the following observations

Lemma 3.52. *Let $I \subset \mathbb{R}$ be an open interval, $t_0 \in I$ and $n \geq 1$. Suppose the function $h : I^n \to \mathbb{R}$ is differentiable at $(t_0, \ldots, t_0)$. Then*

$$\frac{d}{dt} h(t, t, \ldots, t)\Big|_{t=t_0} = \sum_{j=1}^{n} \frac{d}{dt} h(t_0, \ldots, \overset{(j)}{t}, \ldots, t_0)\Big|_{t=t_0}.$$

Proof. The map $\gamma : I \to U$ given by $t \mapsto (t, t, \ldots, t)$ is differentiable, with constant derivative

$$\gamma'(t) = (1, 1, \ldots, 1), \qquad t \in I.$$

From (3.21) and (3.22), with h instead of g, we deduce that

$$\frac{d}{dt} h(t, t, \ldots, t)\Big|_{t=t_0} = Dh(\gamma(t_0))(\gamma'(t_0)) = \sum_{j=1}^{n} D_j h(t_0, \ldots, t_0).$$

The proof is completed by noting that

$$D_j h(t_0, \ldots, t_0) = \frac{d}{dt} h(t_0, \ldots, \overset{(j)}{t}, \ldots, t_0)\bigg|_{t=t_0}. \qquad \square$$

Remark 3.53. Conversely, the chain rule for total differentiation (Remark 3.49) can be derived from the above sum rule. In the setting of Theorem 3.48, the components of the composed function $g \circ f$ are given by

$$(g \circ f)_h(x) = g_h\left(f_1\left(x_1, \ldots, x_j, \ldots, x_n\right), \ldots, f_p\left(x_1, \ldots, x_j, \ldots, x_n\right)\right).$$

The partial derivative at a with respect to the j-th variable is now given by

$$\frac{\partial g_h(f(x))}{\partial x_j}\bigg|_{x=a} = \frac{d}{dt}\bigg|_{t=a_j} g_h\left(f_1\left(a_1, \ldots, \overset{(j)}{t} \ldots, a_n\right), \ldots, f_p\left(a_1, \ldots, \overset{(j)}{t}, \ldots, a_n\right)\right)$$

By applying the sum rule (with $t_0 = a_j$) we find

$$\frac{\partial g_h(f(x))}{\partial x_j}\bigg|_{x=a} = \sum_{i=1}^{n} \frac{d}{dt}\bigg|_{t=a_j} g_h\left(f_1\left(a\right), \ldots, f_i\left(a_1, \ldots, \overset{(j)}{t}, \ldots, a_n\right), \ldots, f_p\left(a\right)\right).$$

By applying the one-variable chain rule we see that the i-th term of the above sum is equal to

$$\frac{\partial g_h(y)}{\partial y_i}\bigg|_{y=f(a)} \frac{\partial f_i(x)}{\partial x_j}\bigg|_{x=a}.$$

This again leads to the formula (3.20) and shows that the sum rule is equivalent to the chain rule. The above also provides another way of looking at the chain rule. △

4 Higher derivatives

In the previous chapter we encountered C^1 functions $f : U \subset \mathbb{R}^n \to \mathbb{R}^m$, i.e. functions such that $Df : U \to \mathrm{Lin}(\mathbb{R}^n, \mathbb{R}^m)$ is continuous. Naturally, if Df is differentiable, we can in turn consider its (first order) partial derivatives, which will be the second order partial derivatives of f itself. One may then continue taking further derivatives.

Studying these higher derivatives is the purpose of this chapter. Concretely, we will:

- Show that the order of differentiation does not matter, if f is sufficiently nicely behaved (Theorem 4.5).
- State the sum, product, and composition rules for higher derivatives (Propositions 4.3 and 4.4).
- Introduce the higher order total derivatives and show that they are homogeneous polynomials (Proposition 4.10).
- Introduce higher order Taylor polynomials and prove the Taylor formula with remainder (Theorem 4.20).
- Give a sufficient criterion guaranteeing that a critical point is a local maximum/minimum, in terms of second derivatives (Theorem 4.29).

4.1 Higher partial derivatives

Let us fix an open subset $U \subset \mathbb{R}^n$. Inductively on the regularity k, we define:

Definition 4.1. A function $f : U \subset \mathbb{R}^n \to \mathbb{R}^p$ is k **times continuously differentiable** if f is C^1 and the total derivative $Df : U \to \mathrm{Lin}(\mathbb{R}^n, \mathbb{R}^p)$ is a C^{k-1} function. One also says that f is a C^k function. △

Observe that Df is made out of the columns $D_j f : U \to \mathbb{R}^p$, as j ranges from 1 to n. This means that Df is C^{k-1} if and only if each of the $D_j f$ is C^{k-1}. Dually, Df has the covectors Df_i as its rows, so Df being C^{k-1} is equivalent to all the Df_i being C^{k-1}.

The set of all C^k functions $U \to \mathbb{R}^p$ is denoted by $C^k(U, \mathbb{R}^p)$. Furthermore, one writes $C^k(U)$ for $C^k(U, \mathbb{R})$. Note that for all $1 \leq l \leq k$ it holds that $C^k(U, \mathbb{R}^p) \subset C^l(U, \mathbb{R}^p)$.

Definition 4.2. A function $f : U \to \mathbb{R}^p$ is said to be **infinitely differentiable**, or **smooth**, or C^∞ if f is a C^k function for every $k \in \mathbb{N}$. △

The space of smooth functions is denoted by $C^\infty(U, \mathbb{R}^p)$. Note that $C^\infty(U, \mathbb{R}^p) = \cap_{k \in \mathbb{N}} C^k(U, \mathbb{R}^p)$.

For $f \in C^k(U, \mathbb{R}^p)$ and any choice of indices $j_1, \ldots, j_k$ with $1 \leq j_i \leq n$, we define the *iterated* or *mixed partial derivative* inductively by

$$D_{j_k} D_{j_{k-1}} \ldots D_{j_1} f := D_{j_k} \left(D_{j_{k-1}} \ldots D_{j_1} f \right).$$

Classically, this is also denoted as

$$\frac{\partial^k f(x)}{\partial x_{j_k} \partial x_{j_{k-1}} \ldots \partial x_{j_1}}.$$

By repeated application of Proposition 3.7 it follows that:

Proposition 4.3. *If $f, g \in C^k(U, \mathbb{R}^p)$, then $f + g$ is also C^k. In particular, $C^k(U, \mathbb{R}^p)$, equipped with pointwise addition and scalar multiplication, is a real vector space.*
Consider then $h \in C^k(U, \mathbb{R})$. Then fh is also C^k. If moreover $h(x) \neq 0$ for every $x \in U$ then $1/h \in C^k(U, \mathbb{R}^p)$.

By the same reasoning, but using instead the chain rule (Theorem 3.48):

Proposition 4.4. *Fix $U \subset \mathbb{R}^n$ and $V \subset \mathbb{R}^p$ open sets and $f \in C^k(U, V)$ and $g \in C^k(V, \mathbb{R}^q)$. Then $g \circ f \in C^k(U, \mathbb{R}^q)$.*

Both statements are left to the reader. They follow by applying induction on k.

4.2 Switching in the order of differentiation

The following is a crucial result about the behaviour of higher partial derivatives. It tells us that it does not matter in which order we differentiate, only how many times we differentiate in each variable:

Theorem 4.5. *Suppose $U \subset \mathbb{R}^n$ is an open subset. Fix a function $f \in C^k(U, \mathbb{R}^p)$, a collection of indices $j_1, \dots, j_l$ with $l \leq k$ and $1 \leq j_i \leq n$, and a permutation σ of the set $\{1, \dots, l\}$. Then:*

$$D_{j_l} \dots D_{j_1} f = D_{j_{\sigma(l)}} \dots D_{j_{\sigma(1)}} f.$$

Before we get to the proof, let us discuss some of its consequences. Since the order of differentiation does not matter for a C^k-function, it makes sense to perform derivatives respecting the order of the variables. This means that each expression $D_{j_l} \dots D_{j_1} f$, with $l \leq k$, can be uniquely written as

$$D_1^{\alpha_1} \dots D_n^{\alpha_n} f = \frac{\partial^l f}{\partial x_1^{\alpha_1} \dots \partial x_n^{\alpha_n}}, \tag{4.1}$$

where α_j is the number of times that the differentiation D_j appears in the sequence $D_{j_l} \dots D_{j_1}$. Equivalently:

$$\alpha_j = \#\{i \in \{1, \dots, l\} \mid j_i = j\}.$$

In particular $\alpha_j = 0$ means that the differentiation D_j does not appear in the sequence. Do note that $l = \alpha_1 + \dots + \alpha_n$ must hold.

You can thus see that $\alpha = (\alpha_1, \dots, \alpha_n)$ is a multi-index (Definition 2.62) of order l. We will henceforth also notate (4.1) as:

$$D^\alpha f = \frac{\partial^{|\alpha|} f}{\partial x^\alpha}$$

and we will say that D^α is a *differential operator* of order l.

Now we address the proof of Theorem 4.5. It can be reduced, via induction, to studying the case in which there are two variables and we differentiate once in each.

Proposition 4.6. *Let $V \subset \mathbb{R}^2$ be an open subset, and let $f : V \to \mathbb{R}$ be a partially differentiable function. Let $a \in V$, and suppose the following conditions hold:*

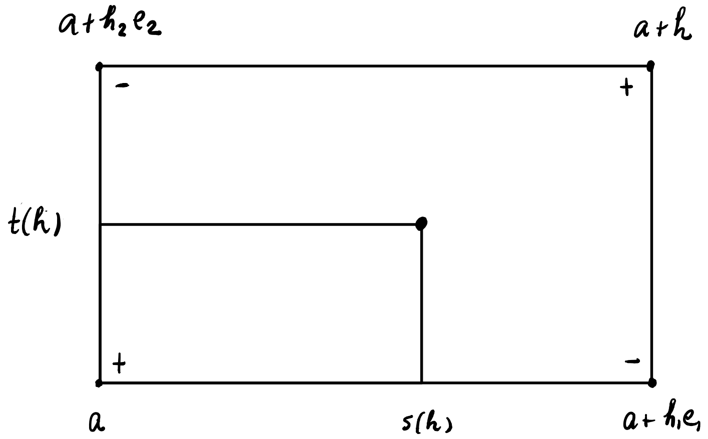
(a) *$D_1 f$ is partially differentiable with respect to the second variable;*

(b) *$D_2 f$ is partially differentiable with respect to the first variable;*

(c) *The mixed partial derivatives $D_2 D_1 f$ and $D_1 D_2 f$ are continuous at a.*

Then

$$D_1 D_2 f(a) = D_2 D_1 f(a). \tag{4.2}$$



Figuur 6: Figure for the proof of Proposition 4.6.

Proof. Since V is open, there exists $\delta > 0$ such that $B(a; 2\delta) \subset V$. For $h = (h_1, h_2) \in \mathbb{R}^2$ with $|h_1|, |h_2| < \delta$, we have $a + h \in B(a; 2\delta) \subset V$. For such h with nonzero components h_1, h_2, define:

$$Q(h) = \frac{f(a+h) - f(a + h_1 e_1) - f(a + h_2 e_2) + f(a)}{h_1 h_2}, \tag{4.3}$$

where $a + h_1 e_1 = (a_1 + h_1, a_2)$ and $a + h_2 e_2 = (a_1, a_2 + h_2)$.

The idea of the proof is to show that Q satisfies:

$$\lim_{h \to 0} Q(h) = D_2 D_1 f(a). \tag{4.4}$$

Moreover, you should notice that the first and second variables in the definition of Q play symmetric roles. The assumptions (a)-(c) are also symmetric in the two variables. As such, the same proof will also apply with the order of the partial derivatives swapped. Hence:

$$\lim_{h \to 0} Q(h) = D_1 D_2 f(a). \tag{4.5}$$

By the uniqueness of limits, (4.4) and (4.5) together will imply (4.2).

Let us establish (4.4). For $h_2 \neq 0$, introduce the auxiliary function $v_{h_2} : (a_1 - \delta, a_1 + \delta) \to \mathbb{R}$ given by

$$v_{h_2}(s) := \frac{f(s, a_2 + h_2) - f(s, a_2)}{h_2}.$$

It is easy to check that for $0 < |h_1|, |h_2| < \delta$:

$$Q(h) = \frac{v_{h_2}(a_1 + h_1) - v_{h_2}(a_1)}{h_1}.$$

The function v_{h_2} is differentiable with derivative

$$v'_{h_2}(s) = \frac{D_1 f(s, a_2 + h_2) - D_1 f(s, a_2)}{h_2}.$$

Applying the mean value theorem in one variable, there exists a number $s(h)$ between a_1 and $a_1 + h_1$ such that

$$Q(h) = v'_{h_2}(s(h)) = \frac{D_1 f(s(h), a_2 + h_2) - D_1 f(s(h), a_2)}{h_2}.$$

Applying the mean value theorem again to the differentiable function

$$\varphi : (a_2 - \delta, a_2 + \delta) \to \mathbb{R}, \qquad t \mapsto D_1 f(s(h), t),$$

we obtain a number $t(h)$ between a_2 and $a_2 + h_2$ such that

$$Q(h) = \varphi'(t(h)) = D_2 D_1 f(s(h), t(h)). \tag{4.6}$$

From the above, we have

$$\|(s(h), t(h)) - (a_1, a_2)\| \leq |s(h) - a_1| + |t(h) - a_2| \leq |h_1| + |h_2|,$$

which implies that $\lim_{h\to 0}(s(h), t(h)) = (a_1, a_2) = a$. Combining this with the continuity of $D_2 D_1 f$ at a, the substitution rule for limits applied to (4.6) yields (4.4). □

Proof of Theorem 4.5. Since every permutation can be written as a composition of adjacent transpositions, it suffices to prove this result for $l = 2$ and σ being the adjacent transposition (12). The desired formula then becomes

$$D_{j_2} D_{j_1} f = D_{j_1} D_{j_2} f.$$

This equality is evident if $j_1 = j_2$. Therefore, we may assume $1 \leq j_1 < j_2 \leq n$. Fix $a \in U$ and choose $\delta > 0$ such that $B(a; \delta) \subset U$. Let $B(\delta)$ denote the open ball in $\mathbb{R}^2$ with center (a_{j_1}, a_{j_2}) and radius δ. Define $\varphi : B(\delta) \to \mathbb{R}^p$ by

$$\varphi(s, t) := f(a_1, \ldots, a_{j_1 - 1}, s, a_{j_1 + 1}, \ldots, a_{j_2 - 1}, t, a_{j_2 + 1}, \ldots a_n).$$

Then φ is a C^2 function on $B(\delta)$ while

$$\begin{aligned} D_1 D_2 \varphi(a_{j_1}, a_{j_2}) &= D_{j_1} D_{j_2} f(a), \\ D_2 D_1 \varphi(a_{j_1}, a_{j_2}) &= D_{j_2} D_{j_1} f(a). \end{aligned}$$

The mixed partial derivatives on the left-hand side are equal by Proposition 4.6. This yields the desired identity. □

4.3 Higher order total derivatives

In first order, we defined partial derivatives, then we observed that one can, more generally, consider directional derivatives and, lastly, we introduced the total derivative. The expression $Df(x)(v) = D_v f(x)$ told us that the total derivative $Df(x)$ is the unique linear map that collects all directional derivatives.

We can proceed in the exact same way in higher order.

4.3.1 Higher directional derivatives

Definition 4.7. Consider an open $U \subset \mathbb{R}^n$, a vector $v \in \mathbb{R}^n$, and a C^k function $f : U \to \mathbb{R}^p$. Its *kth directional derivative* in direction v is defined as:

$$(D_v)^k f := D_v(D_v(\cdots D_v f)) : U \to \mathbb{R}^p. \qquad \triangle$$

The study of these higher directional derivatives reduces to one variable, as the following lemma shows:

Lemma 4.8. *Consider an open $U \subset \mathbb{R}^n$, a point $a \in U$, and a vector $v \in \mathbb{R}^n$. Suppose that $[a, a+v] \subset U$. Then for all $f \in C^k(U, \mathbb{R}^p)$, the function $\varphi : [0,1] \to \mathbb{R}^p$ defined by $t \mapsto f(a+tv)$ is C^k. Moreover:*

$$\frac{d^k\varphi}{dt^k}(t) = (D_v)^k f(a+tv). \tag{4.7}$$

Proof. We prove this by induction on k. For $k = 0$ the result is clear. Let $k \geq 1$ and assume that the result has already been proven for strictly smaller values of k. Then φ is a C^{k-1} function, and

$$\varphi^{(k-1)}(t) = F(a+tv) \quad \text{with} \quad F := (D_v)^{k-1} f.$$

Now $F : U \to \mathbb{R}^p$ is a C^1 function on U. By the definition of the directional derivative, $\varphi^{(k-1)}$ is differentiable on $[0,1]$ with derivative

$$\frac{d}{dt}\varphi^{(k-1)}(t) = \frac{d}{ds}F(a+(t+s)v)\Big|_{s=0} = D_v F(a+tv).$$

It follows that $\varphi \in C^k([0,1], \mathbb{R}^p)$ and that (4.7) holds. □

4.3.2 Higher order total derivatives

Now we can ask ourselves whether the higher directional derivatives $(D_v)^k f$, for varying v, can be packaged in a single object. The higher total derivatives are defined to be the unique polynomial maps that do this:

Definition 4.9. Let $U \subset \mathbb{R}^n$ be an open subset. Consider a function $f \in C^k(U, \mathbb{R}^p)$. Its **$k$th order total derivative** is the function $D^k f : U \to \mathrm{Pol}^k(\mathbb{R}^n, \mathbb{R}^p)$ defined by:

$$D^k f(x)(v) := (D_v)^k f(x). \qquad \triangle$$

The claim that $D^k f$ takes values in pure polynomials follows once we spell out its definition:

Proposition 4.10. *Let $U \subset \mathbb{R}^n$ be open and $f \in C^k(U, \mathbb{R}^p)$. Then for every direction $v \in \mathbb{R}^n$ we have*

$$D^k f(x)(v) = \sum_{i_1=1}^{n} \cdots \sum_{i_k=1}^{n} D_{i_1} \cdots D_{i_k} f(x) v_{i_1} \cdots v_{i_k}.$$

In particular, $D^k f(x)$ is a pure polynomial of order k.

Proof. For every function $g \in C^1(U, \mathbb{R}^p)$ and point $y \in U$, the directional derivative reads:

$$D_v g(y) = Dg(y)(v) = \sum_{j=1}^{n} v_j D_j g(y).$$

By applying this repeatedly we find that

$$(D_v)^k f = (v_1 D_1 + \cdots v_n D_n)^k f = \sum_{i_1=1}^{n} \cdots \sum_{i_k=1}^{n} v_{i_1} D_{i_1} \cdots v_{i_k} D_{i_k} f. \qquad \square$$

This formula can be rewritten in terms of multi-indices $\alpha = (\alpha_1, \cdots, \alpha_n)$. For this, we need the notation $\alpha! := \alpha_1! \cdots \alpha_n!$. Then:

Lemma 4.11. *Let $U \subset \mathbb{R}^n$ be open and $f \in C^k(U, \mathbb{R}^p)$. Then for every point $x \in U$ and vector $v \in \mathbb{R}^n$ we have*

$$D^k f(x)(v) = \sum_{|\alpha|=k} \frac{k!}{\alpha!} D^\alpha f(x) v^\alpha.$$

Proof. This result is purely combinatorial. It amounts to applying Proposition 4.10 and counting how many times each multi-index $\alpha = (\alpha_1, \cdots, \alpha_n)$ shows up. Namely, we want to count how many choices $i_1, \cdots, i_k$ there are such that $D_{i_1} \cdots D_{i_k} = D^\alpha = D_1^{\alpha_1} \cdots D_n^{\alpha_n}$. We claim that it is $\frac{k!}{\alpha!}$.

Write $\mathcal{F}$ for the collection of tuples $I = (i_1, \cdots, i_k)$ of indices associated to the multi-index α. We consider the action of the permutation group S_k on $\mathcal{F}$ given by

$$\sigma \cdot I := I \circ \sigma^{-1} \qquad (I \in \mathcal{F}, \sigma \in S_k).$$

Observe that any two elements in $\mathcal{F}$ are related by some σ. Write $I_0 \in \mathcal{F}$ for the unique tuple in which the i_j appear ordered (i.e. it begins with α_1 ones, then α_2 twos, and so on). The stabilizer S_α of I_0 is the subgroup $S_{\alpha_1} \times \cdots \times S_{\alpha_n} \hookrightarrow S_k$. It follows that its size is $\#S_\alpha = \alpha!$. By the stabilizer-subgroup theorem from group theory, the map $S_k \to \mathcal{F}$ given by $\sigma \mapsto \sigma \cdot I_0$ induces a bijection $S_k/S_\alpha \simeq \mathcal{F}$, concluding the proof. $\square$

Remark 4.12. In the above, the number of elements of $\mathcal{F}$ can also be found as follows. An element I of $\mathcal{F}$ can be described by n successive choices. First, we choose a subset of α_1 elements from $\{1, \cdots, k\}$; there are $\binom{k}{\alpha_1} := k!/\alpha_1!(k-\alpha_1)!$ choices. Next, we choose a subset of α_2 elements from $\{1, \cdots, k-\alpha_1\}$; the number of choices is $\binom{k-\alpha_1}{\alpha_2}$. The l-th choice is that of α_l elements from the collection $\{1, \cdots, k-\alpha_1-\cdots-\alpha_{l-1}\}$. In total, the number of possible choices is

$$\begin{aligned}
&\binom{k}{\alpha_1}\binom{k-\alpha_1}{\alpha_2}\cdots\binom{k-(\alpha_1+\cdots+\alpha_{n-1})}{\alpha_n} \\
&= \frac{k!}{\alpha_1!(k-\alpha_1)!}\frac{(k-\alpha_1)!}{\alpha_2!(k-\alpha_1-\alpha_2)!}\cdots\frac{(k-(\alpha_1+\cdots+\alpha_{n-1}))!}{\alpha_n!0!} \\
&= \frac{k!}{\alpha_1!\ldots\alpha_n!} = \frac{k!}{\alpha!}.
\end{aligned}$$

$\triangle$

Lemma 4.8 can also be spelled out in terms of multi-indices:

Corollary 4.13. *Let $U \subset \mathbb{R}^n$ be open and convex, and $f : U \to \mathbb{R}^p$ a C^k-function. Then for any pair $a, x \in U$, the function $\varphi : [0, 1] \to \mathbb{R}^p$ defined by*

$$\varphi(t) = f(a + t(x - a)) \tag{4.8}$$

is C^k with k-th order derivative given by

$$\frac{1}{k!}\frac{d^k\varphi}{dt^k}(t) = \sum_{|\alpha|=k} \frac{1}{\alpha!} D^\alpha f(a + t(x - a)) \cdot (x - a)^\alpha \tag{4.9}$$

for $0 \le t \le 1$.

Proof. Write $v = x - a$. Then $[a, a + v] \subset U$ and $\varphi(t) = f(a + tv)$. According to Lemmas 4.8 and 4.11, φ is a C^k function, and

$$\frac{1}{k!}\frac{d^k\varphi}{dt^k}(t) = \sum_{|\alpha|=k} \frac{1}{\alpha!} D^\alpha f(a + tv) \cdot v^\alpha.$$

By substituting $v = x - a$, we obtain (4.9). □

4.3.3 Hessian and symmetry

We saw in Proposition 4.10 that the kth total derivative $D^k f(x)$ at a point x is a pure polynomial of order k. In particular, $D^2 f(x)$ is a pure polynomial of second order, i.e., a quadratic form. This means (Lemma 2.41) that we can think of it as a bilinear map. In the particular case of functions with values in $\mathbb{R}$, it has a special name and we can see it as a square matrix:

Definition 4.14. Let $U \subset \mathbb{R}^n$ be open and $f : U \to \mathbb{R}$ a C^2 function. Then for every $a \in U$, the **Hessian** of f at a is defined as the following $n \times n$ matrix:

$$H_f(a) = (D_j D_i f(a))_{1 \le i,j \le n}.$$

△

From Theorem 4.5 it follows that:

Corollary 4.15. *The Hessian $H_f(a)$ is a symmetric matrix.*

We will often write

$$v \mapsto v^t H_f(a) v = \langle H_f(a)v, v\rangle = D^2 f(a)(v)$$

to indicate that we are seeing the matrix $H_f(a)$ as a quadratic form with input v (or as a bilinear form with v plugged in twice).

Example 4.16. Consider the function $f : \mathbb{R}^2 \to \mathbb{R}$ defined by

$$f(x, y) = x^2 + xy + y^2 + x + y.$$

We have $Df(x, y) = (2x + y + 1, 2y + x + 1)$. From this, by computing the second order derivatives, we get

$$H_f(x, y) = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix},$$

for all $x, y \in \mathbb{R}^2$. △

Example 4.17. Consider the function $f : \mathbb{R}^2 \to \mathbb{R}$ defined by

$$f(x, y) = x^2 + y^2 + e^{xy}.$$

By direct computation, we find

$$D_1 D_1 f(x, y) = 2 + y^2 e^{xy}, \quad D_2 D_2 f(x, y) = 2 + x^2 e^{xy}, \quad D_1 D_2 f(x, y) = e^{xy}(1 + xy).$$

Thus,

$$H_f(x, y) = \begin{pmatrix} 2 + y^2 e^{xy} & e^{xy}(1 + xy) \\ e^{xy}(1 + xy) & 2 + x^2 e^{xy} \end{pmatrix},$$

(note that $D_1 D_2 f = D_2 D_1 f$). △

4.4 Higher order Taylor polynomials

In Subsection 3.3.2 we developed the theory of the first order Taylor polynomial. We can work out the higher order case using higher total derivatives.

Definition 4.18. Let $U \subset \mathbb{R}^n$ be an open subset. Given a function $f \in C^k(U, \mathbb{R}^p)$ and a point $a \in U$ we define the **order-k Taylor polynomial** $P_a^k(f) \in \mathrm{Pol}^{\leq k}(\mathbb{R}^n, \mathbb{R}^p)$ using the expression:

$$P_a^k(f)(a + v) := \sum_{0 \leq l \leq k} \frac{1}{l!} D^l f(a)(v) = \sum_{|\alpha| \leq k} \frac{1}{\alpha!} D^\alpha f(a) \cdot v^\alpha.$$

The corresponding remainder term $R \in C^k(U, \mathbb{R}^p)$ is therefore given by the expression:

$$R(x) = f(x) - P_a^k(f)(x).$$ △

Do note that we are defining the remainder R as a function of x. This is a (non-important) difference compared to the first order case (Definition 3.22), where R was expressed as a function of h, with $x = a + h$.

Our goal in this subsection is to establish the Taylor formula with remainder. It will tell us that $P_a^k(f)$ is the best kth order polynomial approximation of f at a. Before we do so, and with this idea in mind, we observe that polynomial functions interact with Taylor polynomials as one may expect:

Lemma 4.19. *Let $P : \mathbb{R}^n \to \mathbb{R}^p$ be a polynomial of order at least k. Fix a point $a \in \mathbb{R}^n$. Then:*

$$P(x) = \sum_{|\alpha| \leq k} \frac{1}{\alpha!} D^\alpha p(a) \cdot (x - a)^\alpha.$$

That is, P is its own Taylor polynomial $P_a^k(f)$.

Proof. From the fact that P is polynomial it follows that the translation $T_a p : v \mapsto p(a + v)$ is also polynomial (Proposition 2.61). Hence, there exist coefficients $c_\alpha \in \mathbb{R}$ such that

$$p(a + v) = \sum_{|\alpha| \leq k} c_\alpha \cdot v^\alpha.$$

By substituting v with $x - a$ we deduce that

$$p(x) = \sum_{|\alpha| \leq k} c_\alpha \cdot (x - a)^\alpha.$$

We write p_α for the monomial $x \mapsto (x-a)^\alpha$. If β is a multi-index we note that

$$D^\beta p_\alpha\Big|_{x=a} = \begin{cases} 0 & \text{if} \quad \alpha \neq \beta \\ \beta! & \text{if} \quad \alpha = \beta \end{cases}$$

This is easy to see by writing $p_\alpha(x) = (x-a)^\alpha$ as the product of factors $q_j(x) = (x_j - a_j)^{\alpha_j}$ and by writing D^β as a composition of $D_j^{\beta_j}$. Then it turns out that

$$D^\beta p_\alpha = \prod_{j=1}^n D_j^{\beta_j}(q_j) = \prod_{j=1}^n \delta_{\beta_j \alpha_j} = \delta_{\alpha\beta},$$

where the last expression denotes the Kronecker symbol. From this remark it follows that

$$D^\beta p(a) = \beta!\, c_\beta,$$

for each $|\beta| \leq k$ and thus the statement. □

We can now formulate and prove the *kth order Taylor theorem with remainder.*

Theorem 4.20. *Let $U \subset \mathbb{R}^n$ be an open convex set and $a \in U$ a point. Let $f \in C^{k+1}(U)$. Then, for each $x \in U$, the remainder in the expression*

$$f(x) = P_a^k(f)(x) + R(x),$$

can be computed as

$$R(x) = \sum_{|\alpha|=k+1} \frac{1}{\alpha!} D^\alpha f(\xi_x) \cdot (x-a)^\alpha, \tag{4.10}$$

for some $\xi_x \in [a, x]$. Equivalently:

$$R(x) = \frac{1}{(k+1)!} D^{k+1} f(\xi_x)(x-a). \tag{4.11}$$

Proof. Define $\varphi : [0,1] \to \mathbb{R}$ by $\varphi(t) := f(a+tv)$ with $v = x - a$. According to Corollary 4.13, the function φ belongs to $C^{k+1}([0,1])$. By the Taylor theorem with remainder for functions of one variable, for every $t \in [0,1]$ we have

$$\varphi(t) = \sum_{j=0}^k \frac{1}{j!} \varphi^{(j)}(0) t^j + \rho(t),$$

where the remainder is given by

$$\rho(t) = \frac{1}{(k+1)!} \varphi^{(k+1)}(\tau) t^{k+1},$$

for some $\tau \in [0,t]$. Using Corollary 4.13 and taking $t = 1$ we find that

$$f(x) = f(a + (x-a)) = p(x) + R(x),$$

where

$$p(x) = \sum_{j=0}^k \sum_{|\alpha|=j} \frac{1}{\alpha!} D^\alpha f(a) \cdot (x-a)^\alpha = P_a^k(f)(x)$$

and

$$R(x) = \frac{1}{(k+1)!} \varphi^{(k+1)}(a + \tau_x(x-a))$$

for some $\tau_x \in [0,1]$. This yields the claim with $\xi_x := a + \tau_x(x-a) \in [a,x]$. □

Do note that the theorem is phrased for functions with values in $\mathbb{R}$. However, it can be applied component by component to deduce:

Proposition 4.21. *Let $U \subset \mathbb{R}^n$ be an open convex set and $a \in U$. Let $f \in C^{k+1}(U, \mathbb{R}^p)$. Then, for each $x \in B(a;\delta) \subset U$, the remainder in the expression*

$$f(x) = P_a^k(f)(x) + R(x)$$

satisfies

$$|R(x)| \leq M\|x-a\|^{k+1},$$

for some constants $\delta > 0$ and $M > 0$ such that $B(a;\delta) \subset U$.

Proof. Write $f = (f_1, \cdots, f_p)$ for the components of f. Apply Theorem 4.20 to each of them, so the remainder R_i of f_i is computed at an intermediate point $\tau_x^i \in [a, x]$.

Choose $\delta > 0$ such that $\bar{B}(a;\delta) \subset U$. The functions $D^\alpha f_i$ are continuous on the closed and bounded set $\bar{B}(a;\delta)$. Therefore, there exists a $C > 0$ such that for each α with $|\alpha| = k+1$ and all i we have $|D^\alpha f_i| \leq C$ on $\bar{B}(a;\delta)$. For $x \in B(a;\delta)$ we have $\tau_x^i \in [a, x] \subset B(a;\delta)$, so

$$|R_i(x)| \leq \sum_{|\alpha|=k+1} \frac{1}{\alpha!} C\,|(x-\alpha)^\alpha|.$$

As such, it is enough if we set

$$M := pC \sum_{|\alpha|=k+1} \frac{1}{\alpha!}.$$

□

These results can be particularised to the cases of low order. For first order, the Hessian enters the computation of the remainder:

Corollary 4.22 (First order Taylor formula with remainder). *Let $U \subset \mathbb{R}^n$ be a convex open set and $a \in U$. Let $f : U \to \mathbb{R}$ be a C^2 function. Then, for every $x \in U$, there exists a $\xi = \xi_x \in (a, x)$ such that*

$$f(x) = f(a) + Df(a)(x-a) + (x-a)^t H_f(\xi)(x-a). \tag{4.12}$$

For second order we have, analogously:

Corollary 4.23. *Let $U \subset \mathbb{R}^n$ be a convex open set and $a \in U$. Let $f : U \to \mathbb{R}$ be a C^3 function. Then, for every $x \in U$, we have*

$$f(x) = f(a) + Df(a)(x-a) + (x-a)^t H_f(a)(x-a) + R(x), \tag{4.13}$$

where there exists a $0 < \tau_x < 1$ such that

$$R(x) = \frac{1}{3!}\frac{d^3}{dt^3} f(a + t(x-a))\Big|_{t=\tau_x}. \tag{4.14}$$

In particular:

$$\lim_{x\to a} \|x-a\|^{-2} R(x) = 0.$$

We close this subsection with our original claim: the Taylor polynomial is the (unique!) best approximation of order k.

Exercise 4.24. Fix a function $f \in C^k(U)$. Prove that the Taylor polynomial $P_a^k(f)$ is the unique polynomial $p \in \mathrm{Pol}^{\leq k}(\mathbb{R}^n)$ satisfying the property

$$||f(x) - p(x)|| \leq C.||x - a||^{k+1}$$

for all $x \in B(a, \delta) \subset U$, for some $C > 0$ and $\delta > 0$. △

Hint: By construction, $P_a^k(f)$ is the unique polynomial $p \in \mathrm{Pol}^{\leq k}(\mathbb{R}^n)$ satisfying

$$D^\alpha f(a) = D^\alpha(p)(a)$$

for each multi-index with $|\alpha| \leq k$. As such, every other polynomial q differs from f in some derivative. You can use this to find a direction $(x - a)$ in which the difference $||f(x) - q(x)||$ has order $||x - a||^i$, with $i \leq k$.

4.5 Critical points and Hessian

We close this chapter with a discussion of critical points (recall Definition 3.32). Concretely, we will relate their nature to the behaviour of the Hessian. This is analogous to the corresponding phenomenon in one variable, where non-vanishing second derivatives imply that a critical point is a local maximum/minimum.

4.5.1 Reminder: the case of one-variable

Lemma 4.25. *Let $I \subset \mathbb{R}$ be an interval and $f : I \to \mathbb{R}$ a differentiable function. Then the following hold:*

(a) *f is non-decreasing if and only if $f'(x) \geq 0$ for all $x \in I$,*

(b) *f is strictly monotonically increasing if $f' > 0$ on $\mathrm{inw}(I)$,*

(c) *f is non-increasing if and only if $f'(x) \leq 0$ for all $x \in I$,*

(d) *f is strictly monotonically decreasing if $f' < 0$ on $\mathrm{inw}(I)$.*

Proof. Since (c) and (d) follow by applying (a) and (b) to $-f$, we can restrict ourselves to (a) and (b). We start with (a).

Assume that f is monotonically increasing on I and let $a \in I$. First, assume that a is not a right endpoint of the interval. Then there exists $\delta > 0$ such that $[a, a + \delta) \subset I$. Then for all $0 < h < \delta$, $a + h \in I$, so $f(a + h) \geq f(a)$, hence $[f(a + h) - f(a)]/h \geq 0$. Taking the limit as $h \downarrow 0$ gives $f'(a) \geq 0$.

If a is a right endpoint, then it is not a left endpoint, so there exists $\delta > 0$ such that $(a - \delta, a) \subset I$. Then for $-\delta < h < 0$, $f(a + h) - f(a) \leq 0$, so $[f(a + h) - f(a)]/h \geq 0$. Taking the limit as $h \uparrow 0$ now gives $f'(a) \geq 0$.

Now assume that $f' \geq 0$ on I and let $x, y \in I$ with $x < y$. Then by the mean value theorem, $f(y) - f(x) = f'(\xi)(y - x)$ for some $\xi \in [x, y] \subset I$. Since $f'(\xi) \geq 0$ and $y - x > 0$, it follows that $f(y) - f(x) \geq 0$. Hence f is non-decreasing. This proves (a).

We prove (b). Let the condition hold and let $x, y \in I$ with $x < y$. By the mean value theorem, there exists $\xi \in (x, y)$ such that $f(y) - f(x) = f'(\xi)(y - x)$. Since $f'(\xi) > 0$, we find that $f(y) - f(x) > 0$. Hence f is strictly monotonically increasing. □

Lemma 4.26. *Let $f : (a,b) \to \mathbb{R}$ be a twice continuously differentiable function. Fix a point $c \in (a,b)$ with $f'(c) = 0$ and $f''(c) \neq 0$. Then the following holds:*

(a) *If $f''(c) > 0$, then f has a local minimum at c;*

(b) *If $f''(c) < 0$, then f has a local maximum at c.*

In any case, the function f has a local extremum at c.

Proof. We show (a). Claim (b) follows by applying (a) to the function $-f$.

From the continuity of f'' it follows that there exists $\delta > 0$ with $I := (c-\delta, c+\delta) \subset (a,b)$ such that $f''(x) > 0$ for all $x \in I$. By applying Lemma 4.25 (b) to f' instead of f, it follows that f' is strictly monotonically increasing on I. In particular, $f' < 0$ on $(c-\delta, c)$ and $f' > 0$ on $(c, c+\delta)$. By applying Lemma 4.25 (d) and (b), we conclude that f is strictly monotonically decreasing on $(c-\delta, c]$ and strictly monotonically increasing on $[c, c+\delta)$. Hence $f(x) > f(c)$ for $x \in (c-\delta, c)$ and for $x \in (c, c+\delta)$. This proves (a). □

Remark 4.27. From the above proof it follows the stronger statement that the local extremum c is isolated in the sense that there exists $\delta > 0$ such that $I := (c-\delta, c+\delta) \subset (a,b)$ and $f|_I$ has no other local extremum than c. △

We give an alternative proof of Lemma 4.26 using the Taylor theorem with remainder. We will follow a similar approach for the case of multiple variables.

Proof of Lemma 4.26. We first show (a). Let $x \in (a,b)$. Then, by the Taylor theorem with remainder, there exists $\xi = \xi_x$ between c and x such that

$$f(x) = f(c) + \frac{1}{2} f''(\xi)(x-c)^2.$$

Here we have used $f'(c) = 0$, so the first order term is zero. From the continuity of f'' it follows that there exists $\delta > 0$ such that $(c-\delta, c+\delta) \subset (a,b)$ and for all $\xi \in (c-\delta, c+\delta)$ we have $f''(\xi) > 0$. For $x \in (c-\delta, c+\delta)$ with $x \neq c$, it follows that

$$f(x) - f(c) = \frac{1}{2} f''(\xi_x)(x-c)^2 > 0.$$

Hence, f has a local minimum at c. Claim (b) follows by applying (a) to the function $-f$. □

4.5.2 The multivariate case

The first part of the proof goes exactly the same in the multivariate case:

Lemma 4.28. *Let $U \subset \mathbb{R}^n$ be open and $f : U \to \mathbb{R}$ a C^2 function. Let $a \in U$ and assume that the Hessian $H_f(a)$ is positive definite. Then there exist constants $m > 0$ and $\delta > 0$ such that $B(a;\delta) \subset U$ and for all $x \in B(a;\delta)$ and all $v \in \mathbb{R}^n$:*

$$\langle H_f(x)v, v\rangle \geq m\|v\|^2.$$

Proof. First, compare this lemma to Proposition 2.84. The idea is that, since $H_f(a)$ is positive definite, $H_f(x)$ is also positive definite for all $x \in B(a;\delta)$ if δ is sufficiently small. Then, the bound of Proposition 2.84 applies to all such $H_f(x)$ at once.

Instead of invoking Proposition 2.84, it is easier to do the proof from scratch, following the same steps as in the analytic proof of Proposition 2.84. First, note that there is some $\delta_0 > 0$ such that $B(a;\delta_0) \subset U$. Suppose then that the claim is false. Then for every $\delta < \delta_0$ and every integer $k \geq 1$, there exists $x \in B(a;\delta)$ and $v \in \mathbb{R}^n$ such that

$$\langle H_f(x)v, v\rangle < \frac{1}{k}\|v\|^2.$$

Note that this inequality implies $v \neq 0$, so by replacing v with $v/\|v\|$ we see that such a v can also be found in the unit sphere $S = \{y \in \mathbb{R}^n \mid \|y\| = 1\}$.

By taking $\delta = \frac{1}{k}$, we find $x_k \in B(a;\frac{1}{k})$ and $v_k \in S$ such that

$$\langle H_f(x_k)v_k, v_k\rangle < \frac{1}{k}.$$

The sequence (v_k) is contained in the set S, which is closed and bounded in $\mathbb{R}^n$, hence sequentially compact. Therefore, (v_k) has a convergent subsequence (v_{k_j}) with a limit $v \in S$. In particular, $v \neq 0$. We have $x_{k_j} \to a$ as $j \to \infty$. By continuity of the matrix elements of H_f, $H_f(a) = \lim_{j\to\infty} H_f(x_{k_j})$. Hence,

$$\langle H_f(a)v, v\rangle = \lim_{j\to\infty} \left\langle H_f(x_{k_j})v_{k_j}, v_{k_j}\right\rangle \leq \lim_{k\to\infty} \frac{1}{k} = 0,$$

a contradiction with the fact that $H_f(a)$ was positive definite. □

Finally, we get to the result:

Theorem 4.29. *Let $U \subset \mathbb{R}^n$ be open and $f \in C^2(U)$. Let $a \in U$ be a critical point of f and assume that the Hessian $H_f(a)$ is positive (resp. negative) definite. Then there exist constants $\delta > 0$ and $c > 0$ such that $B(a;\delta) \subset U$ and for all $x \in B(a;\delta)$:*

$$f(x) \geq f(a) + c\|x-a\|^2 \qquad (\text{resp. } f(x) \leq f(a) - c\|x-a\|^2).$$

In particular, f has a local minimum (resp. maximum) at a.

Proof. It suffices to prove this in the case where $H_f(a)$ is positive definite (replacing f by $-f$ gives the other case).

Since a is a critical point, $Df(a) = 0$. By the previous lemma, there exists $\delta > 0$ such that $B(a;\delta) \subset U$ and $H_f(x)$ is positive definite for all $x \in B(a;\delta)$. For $x \in B(a;\delta)$, $[a,x] \subset B(a;\delta)$. By the Taylor formula with remainder (4.12) we get:

$$f(x) - f(a) = (x-a)^t H_f(a + \tau_x(x-a))(x-a) \geq \frac{m}{2}\|x-a\|^2.$$ □

Example 4.30. Consider the function f from Example 4.16. From the computed total derivative, (x,y) is a stationary point of f if and only if $2x + y + 1 = 0$ and $x + 2y + 1 = 0$. These equations have the unique solution $(1/3, 1/3)$. Hence this is the only stationary point of f. For the Hessian:

$$H = H_f(\tfrac{1}{3},\tfrac{1}{3}) = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}.$$

Since $\det H = 4 - 1 = 3 > 0$ and $H_{11} = 2 > 0$, H is positive definite (Lemma 2.45). By Propositie 4.29, f has a local minimum at $(1/3, 1/3)$. △

Example 4.31. Consider the function $f(x,y) = x^2 + y^2 + e^{xy} : \mathbb{R}^2 \to \mathbb{R}$ from Voorbeeld 4.17. It is easy to check that $(0,0)$ is a stationary point of f. Recall the previously found formula for the Hessian:

$$H = H_f(0,0) = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}.$$

Since this Hessian is positive definite, Proposition 4.29 implies that f attains a local minimum at $(0,0)$. △

Exercise 4.32. Show that in the example above, $(0,0)$ is the only local extremum. △

Exercise 4.33. Let $U \subset \mathbb{R}^2$ be open and $f \in C^2(U)$. Assume that $a \in U$ is a critical point of f, and that $\det H_f(a) < 0$. Show that f has no local extremum at a. **Hint:** Use Proposition 2.45 to find two directions, one along which f increases, and another one along which it decreases. △

5 Introduction to the multivariate Riemann integral

We have developed the theory of differentiation in multiple variables, so the natural next step is to look into the theory of integration. It turns out that this is a subtle story and in this course we will only be able to develop part of it. Before we state the main results, let us provide a general discussion so you are aware of the ideas that come into play (both in this and in later courses).

5.1 Opening discussion

Consider an open $U \subset \mathbb{R}^n$ and the usual coordinates $(x_1, \cdots, x_n)$. Given a function $f : U \to \mathbb{R}$ it is natural for us to consider the integral

$$\int_U f(x) dx_1 \cdots dx_n,$$

which we would want to define so that it is the "$(n+1)$-dimensional volume" under the graph of f (much like the usual integral in one variable computes the area under the graph).

It turns out that in higher dimensions this is not the only type of integral you can consider. In Chapter 7 we will study instead *line integrals*; i.e. integrals of *covector fields* along paths in U. The way to think about it is that $f(x)dx_1 \cdots dx_n$ is a gadget that can be integrated over the whole of U, which is n-dimensional, whereas a covector field is a gadget that integrates over paths, that are 1-dimensional. There are other intermediate objects, called *differential k-forms*, that can be integrated over k-dimensional objects in U (so-called *submanifolds*, to be introduced in Chapter 6). We will not study differential forms in general in this course; you will see them in *Analyse in meer variabelen* and/or *Differentieerbare varieteiten*.

Let us go back to integrals of functions, which is the focus of this chapter. The question is how to define such a thing. A reasonable approach is to repeat what we did for the Riemann integral in one variable. We can cover U by n-dimensional hypercubes and consider "step" functions that are constant on each hypercube. Integrating those is easy (just take the hypervolume of each hypercube, scale it by the value of the function there, and add it all up). This will allow us to define an "upper Riemann sum" (consider all such functions, larger than f, and take the infimum of their integrals) and a "lower Riemann sum" (consider all such functions, now smaller than f, and take the supremum of integrals). If these two agree, we say that the result is

$$\int_U f(x) dx_1 \cdots dx_n.$$

Of course, now one would have to prove that this always exists under some conditions (for instance, if f is continuous).

In this chapter we will follow a different approach, that will save us a lot of work:

Definition 5.1. Consider an open $U \subset \mathbb{R}^n$, a hypercube $C = \prod_{i=1}^n [a_i, b_i] \subset U$, and a continuous function $f : C \to \mathbb{R}^p$. The **(Riemann) integral** of f over C is

$$\int_C f(x) dx_1 \cdots dx_n := \int_{a_n}^{b_n} \cdots \int_{a_1}^{b_1} f(x_1, \cdots, x_n) dx_1 \cdots dx_n. \qquad \triangle$$

That is, we can integrate "one variable at a time". This reduces the computation of higher-dimensional integrals to computing n integrals of one variable. Our goal in this chapter is to the prove the following results:

- We will prove (Theorem 5.12) that integrating a continuous function $f(x_1, \cdots, x_n)$ with respect to the first variable yields a function $F(x_2, \cdots, x_n) = \int_{a_1}^{b_1} f(x_1, \cdots, x_n)dx_1$ that is still continuous. This means that we can keep integrating with respect to the other variables and thus shows that Definition 5.1 is actually well-defined.
- We will prove that the order of integration does not matter (Theorem 5.25). That is, we could have started integrating with the last variable and got the same result.
- We will relate integration to differentation, establishing the so-called "differentiation under the integral sign" theorem (Theorem 5.21).
- Before we get to these more involved results we will review the theory of vector valued integrals for functions of one variable (Subsection 5.2).

There are a couple of things that we will not do and that you will encounter in later courses. First, we will only learn about integrating over hypercubes. This is not a great loss of generality, in fact. In *Analyse in meer variabelen* and/or *Differentieerbare varieteiten* you will see that one can use this to handle general opens U by expressing any f as a sum of functions f_i that are supported on hypercubes (i.e. there is hypercube C_i such that $f_i = 0$ outside).

Secondly, we will only develop the theory of integration for continuous functions. In one variable you are probably familiar with the fact that Riemann-integrable functions need not be continuous. This is of course also true in higher dimensions. In fact, there is an alternative theory of integration, known as the Lebesgue integral, that allows for the integration of an even larger class of functions. You will see it in detail in *Maat en Integratie* or *Functionaal analyse*.

5.2 Vector-valued integrals

You are already familiar with integrating functions $\mathbb{R} \to \mathbb{R}$, i.e. with one input and one output. We will now discuss the Riemann-integral of vector-valued functions $f : [a, b] \subset \mathbb{R} \to \mathbb{R}^p$ of one variable.

To do so, we will work with the components of $f = (f_1, \cdots, f_p)$. The main message is that, in order to integrate f, we simply integrate each f_j separately. This follows from the fact that limits, derivatives, and sums can be performed component-wise.

Definition 5.2. A function $f : [a, b] \to \mathbb{R}^p$ is called *Riemann-integrable* if for each $1 \leq j \leq p$ the component $f_j : [a, b] \to \mathbb{R}$ is Riemann-integrable. The *Riemann integral* of f is defined as the unique vector $I(f)$ in $\mathbb{R}^p$ whose j-th component is

$$I(f)_j = \int_a^b f_j(x)dx \qquad (1 \leq j \leq p).$$

We denote this vector by

$$I(f) = \int_a^b f(x)dx. \qquad \triangle$$

The collection of Riemann-integrable functions $[a, b] \to \mathbb{R}^p$ is denoted by $\mathcal{R}([a, b], \mathbb{R}^p)$. Equipped with pointwise addition and scalar multiplication, this is a linear space over $\mathbb{R}$.

Lemma 5.3. *The Riemann integral* $I : f \mapsto I(f)$ *is a linear map* $\mathcal{R}([a, b], \mathbb{R}^p) \to \mathbb{R}^p$.

Proof. Let $f, g \in \mathcal{R}([a,b],\mathbb{R}^p)$ and $\lambda \in \mathbb{R}$. Then the j-th component of $f+\lambda g$ is $(f+\lambda g)_j = f_j + \lambda g_j$. By the usual rules of Riemann integration for scalar functions, this component is Riemann-integrable. Hence $f+\lambda g$ is Riemann-integrable and

$$I(f+\lambda g)_j = \int_a^b [f_j(x) + \lambda g_j(x)]\, dx = I(f)_j + \lambda I(g)_j.$$

This equality for each j implies the linearity of $I : \mathcal{R}([a,b],\mathbb{R}^p) \to \mathbb{R}^p$. □

Definition 5.4. $F : [a,b] \to \mathbb{R}^p$ is a *primitive* of $f : [a,b] \to \mathbb{R}^p$ if F is differentiable with derivative f. △

The following is the vector-valued analogue of the fundamental theorem of calculus:

Theorem 5.5. *Every continuous function $f : [a,b] \to \mathbb{R}^p$ is Riemann-integrable.*

If $f : [a,b] \to \mathbb{R}^p$ is continuous, then the function $G : [a,b] \to \mathbb{R}^p$ defined by

$$G(x) := \int_a^x f(t)dt$$

is a primitive of f.

If $F : [a,b] \to \mathbb{R}^p$ is a primitive of f, then

$$\int_a^b f(t)dt = F(b) - F(a). \tag{5.1}$$

Proof. If f is continuous, all components $f_j : [a,b] \to \mathbb{R}$ are continuous and thus Riemann-integrable. The j-th component of G is

$$G_j(x) = \int_a^x f_j(t)\, dt.$$

Hence G_j is differentiable with derivative f_j, so G is differentiable with derivative f.

Finally, let F be a primitive of f. Then each component F_j is a primitive of f_j, so the usual fundamental theorem of calculus in one variable applies:

$$\int_a^b f_j(t)dt = F_j(b) - F_j(a).$$ □

5.2.1 Vector-valued functions as curves

It is convenient to keep the following geometric picture in mind. We think of $f : [a,b] \to \mathbb{R}^p$ as a choice of velocity vector $f(t)$ for each $t \in [a,b]$. It tells us in which direction we move at time t. A primitive $F(t) = x_0 + \int_a^t f(s)ds$ is thus a curve in $\mathbb{R}^p$, starting at $F(a) = x_0$ and with velocity $f(t)$ at time t. If you draw F, you can draw $f(t)$ as a little tangent vector at $F(t)$.

The following lemma is useful: if we take a primitive F and apply a linear map A, the resulting curve $A \circ F$ has $A \circ f$ as velocity.

Lemma 5.6. *Let $A : \mathbb{R}^p \to \mathbb{R}^q$ be a linear map and $f \in \mathcal{R}([a,b],\mathbb{R}^p)$. Then $A \circ f : [a,b] \to \mathbb{R}^q$ is Riemann-integrable, and*

$$A\left(\int_a^b f(x)dx\right) = \int_a^b A(f(x))dx.$$

Proof. For each $1 \le i \le q$, the i-th component of $A \circ f$ is

$$(A \circ f)_i = \sum_{j=1}^{p} A_{ij} f_j.$$

The result follows by repeated application of Lemma 5.3. □

Exercise 5.7. Instead of a linear map A, consider a differentiable function $\phi \in C^1(\mathbb{R}^p, \mathbb{R}^q)$ applied to a primitive $F : [a, b] \to \mathbb{R}^p$ of the velocity curve f. According to Lemma 3.51, the velocity of $\phi \circ F$ at time t is $D\phi(F(t))(f(t))$. Check that Lemma 5.6 is a special case. △

Another useful property is the triangle inequality for Riemann integrals. Before discussing it, consider the following notion:

Definition 5.8. Let $F : [a, b] \to \mathbb{R}^p$ be a C^1 curve. The **length** of F is:

$$\operatorname{len}(F) := \int_a^b \|F'(t)\| dt.$$ △

Note that $\|F'(t)\|$ is the speed at time t, so integrating it gives the total length of the curve.

We also need the following auxiliary lemma:

Lemma 5.9. *Let $v \in \mathbb{R}^p$. Then*

$$\|v\| = \max\{\langle v, w\rangle \mid w \in \mathbb{R}^p,\ \|w\| = 1\}.$$

Proof. First, $\|v\| = \langle v, w\rangle$ if $w = v/\|v\|$, giving the upper bound. The opposite bound is the Cauchy-Schwarz inequality (Proposition 2.74): $\langle v, w\rangle \le \|v\|\,\|w\| = \|v\|$ if $\|w\| = 1$. □

Proposition 5.10 (Triangle inequality for the integral)**.** *Let $f : [a, b] \to \mathbb{R}^p$ be continuous. Then the norm $\|f\| : [a, b] \to \mathbb{R}$ defined by $t \mapsto \|f(t)\|$ is continuous. Moreover:*

$$\left\| \int_a^b f(t)dt \right\| \le \int_a^b \|f(t)\|\, dt. \tag{5.2}$$

Rephrasing in terms of curves, write $F(t) = \int_a^t f(s)ds$ as the primitive of f starting at $F(a) = 0$. Then $F(b) = \int_a^b f(t)\, dt$ is the endpoint of F, and the left hand side of (5.2) is $\|F(b)\|$, its distance from the origin. The right side is the total length of F, so we must show $\|F(b)\| \le \operatorname{len}(F)$.

Proof. From $\|f(t)\|^2 = \sum_{j=1}^p f_j(t)^2$, the function $\|f\|^2$ is continuous on $[a, b]$. Since $y \mapsto \sqrt{y}$ is continuous on $[0, \infty)$,

$$\|f\| : t \mapsto \sqrt{\|f(t)\|^2}$$

is continuous. Denote the integral by $I(f)$, which is $F(b)$ in our notation.

Take $w \in \mathbb{R}^p$ with $\|w\| = 1$. The map $A : v \mapsto \langle v, w\rangle$ is linear. By Lemma 5.6, $A \circ f$ is Riemann-integrable, and

$$\langle I(f), w\rangle = A\left(\int_a^b f(t)dt\right) = \int_a^b A(f(x))dx = \int_a^b \langle f(x), w\rangle dx.$$

The left side measures the total displacement in the w-direction, the right side displays it as the integral of the infinitesimal displacement along w.

By Cauchy-Schwarz, using that $||w|| = 1$:

$$|\langle f(x), w\rangle| \leq \|f(x)\|\|w\| = \|f(x)\|.$$

Which plugged into the integral yields:

$$|\langle I(f), w\rangle| \leq \int_a^b \|f(x)\|\, dx.$$

Since this holds for all w, Lemma 5.9 applies and we deduce that (5.2) holds. □

Remark 5.11. One can generalize this to Riemann-integrable f. Then $\|f\|$ is also Riemann-integrable and the inequality remains valid. The proof is omitted here. △

5.3 Integrals in a single direction

We now begin developing the theory of the Riemann integral in multiple variables. Our goal in this subsection is to discuss the integration of a multivariate function with respect to a single variable.

Consider an open $V \subset \mathbb{R}^n$, an interval $[a, b] \subset \mathbb{R}$, and a function $f : V \times [a, b] \to \mathbb{R}^p$. For each $x \in V$ we can consider the function $f_x : [a, b] \to \mathbb{R}^p$ given by $t \mapsto f(x, t)$. In this manner we are thinking of f as a function of one variable $t \in [a, b]$ and n *parameters* $(x_1, \ldots, x_n)$.

If we assume that each f_x is Riemann-integrable then we can define the primitive in the t-direction:

$$F(x, t) := \int_a^t f_x(s)ds = \int_a^t f(x, s)ds, \tag{5.3}$$

as well as the total Riemann integral in the t-direction:

$$G(x) := \int_a^b f_x(t)\,\mathrm{d}t = \int_a^b f(x,\, t)\,\mathrm{d}t \tag{5.4}$$

The former is a function $F : V \times [a, b] \to \mathbb{R}^p$, while the latter is $G : V \to \mathbb{R}^p$.

The following statement says that if f depends continuously on all its variables $(x_1, \ldots, x_n, t)$, then the integral along $t \in [a, b]$ also depends continuously on all variables, including the parameters $(x_1, \ldots, x_n)$.

Theorem 5.12. *Consider an open set $V \subset \mathbb{R}^n$ and an interval $[a, b] \in \mathbb{R}$. Suppose that the function $f : V \times [a, b] \subset \mathbb{R}^{n+1} \to \mathbb{R}^p$ is continuous. Then:*

- *The function $F : V \times [a, b] \to \mathbb{R}^p$ defined by Equation (5.3) is continuous. Morever, $D_t F = f$.*
- *The function $G : V \to \mathbb{R}^p$ defined by Equation (5.4) is continuous.*

Proof. Observe that $D_t F = f$ follows from the fundamental theorem of calculus. Moreover, $G(x) = F(x, b)$, so G being continuous follows from F being continuous. It remains to show F continuous. The key ingredient behind the argument is the notion of uniform continuity (Subsection 1.1.2) and the fact that $[a, b]$ is closed and bounded and therefore sequentially compact (Proposition 1.10).

Indeed, fix $x_0 \in V$ and thus the interval $K = \{x_0\} \times [a, b] \subset V \times [a, b]$. Compactness implies that $\|f|_K\|$ achieves a maximum, which we can call M. Moreover, Proposition 1.13 implies that f is uniformly continuous over K. In particular, this means that for every $\eta > 0$ there is some $\delta > 0$ such that for all $x, x_0 \in V$ and all $t, t_0 \in [a, b]$ satisfying $\|x - x_0\|, \|t - t_0\| \leq \delta$ it holds that

$$\|f(x,t) - f(x_0, t_0)\| \leq \eta.$$

This leads to the estimate:

$$\begin{aligned}
\|F(x,t) - F(x_0,t_0)\| &= \left\| \int_a^t (f(x,s) - f(x_0,s))\, ds + \int_{t_0}^t f(x_0,s) ds \right\| \\
&= \left\| \int_a^t (f(x,s) - f(x_0,s))\, ds \right\| + \left\| \int_{t_0}^t f(x_0,s) ds \right\| \\
&\leq \int_a^b \|f(x,s) - f(x_0,s)\|\, ds + M\|t_0 - t\| \\
&\leq \eta(b-a) + M\delta.
\end{aligned}$$

They key idea is that we separated the estimate into two integrals. One of them we can bound by uniform continuity of f. The other one we can bound using that f is bounded. Thus: if we want $\eta(b-a) + M\delta \leq \varepsilon$ we can take $\eta \leq \varepsilon/(2(b-a))$ and δ accordingly and additionally satisfying $\delta \leq \varepsilon/(2M)$. This establishes continuity. □

Corollary 5.13. *The Riemann integral defined in Definition 5.1 is well-defined.*

Proof. According to Theorem 5.12, the function $G_1(x_2, \cdots, x_n) = \int_{a_1}^{b_1} f(x_1, \cdots, x_n) dx_1$ is continuous and in particular Riemann integrable. As such, we can integrate it with respect to x_2, yielding a function $G_2(x_3, \cdots, x_n)$ that is continuous as well. Doing this n times yields the integral over the whole hypercube:

$$\int_C f(x) dx_1 \cdots dx_n = \int_{a_n}^{b_n} \cdots \int_{a_1}^{b_1} f(x_1, \cdots, x_n) dx_1 \cdots dx_n. \qquad \square$$

The continuity of the function G from Theorem 5.12 means that G is continuous at every point $x_0 \in V$. This in turn means that

$$\lim_{x \to x_0} G(x) = G(x_0).$$

If we substitute the definition of G above, and use that f is continuous in (x_0, t) for every $t \in [a, b]$, so that $f(x,t) \to f(x_0, t)$ as $x \to x_0$, we see that:

Corollary 5.14. *Suppose that f is continuous as a function of all its variables. Then:*

$$\lim_{x \to x_0} \int_a^b f(x,t) dt = \int_a^b f(x_0, t) dt = \int_a^b \left(\lim_{x \to x_0} f(x,t) \right) dt, \tag{5.5}$$

Formula (5.5) states that we may "swap limits and integrals", but it is very important to keep in mind that this switching is not always allowed. For instance:

Example 5.15. Consider the function $f(x,t) = xe^{-xt}$ defined over $A = \{t \geq 0\} \subset \mathbb{R}^2$. You can verify that it is continuous, since it is a composition and product of continuous functions. We can now consider $G(x) := \int_0^\infty f(x,t) dt$. Do observe that we are now integrating over $[0, \infty)$, which is not covered by Theorem 5.12. In particular, it could happen that the integral is not defined. However, in this case the integral does exist for each x. You can see that $G(0) = 0$ and $G(x) = \int_0^\infty xe^{-xt} dt = -e^{-xt}|_0^\infty = 1$ for all other x. That is, even though the integral does exist, the resulting function of x is not continuous. △

Nonetheless, Corollary 5.14 can be very useful for computing integrals and limits explicitly, as the following examples show.

Example 5.16. We consider the function $f : \mathbb{R} \times [-1, 1] \to \mathbb{R}$ given by $f(x, t) = e^{xt}$. Applying the Corollary with $V = \mathbb{R}$, $p = 1$, and $[a, b] = [-1, 1]$, we see that the function $F : \mathbb{R} \to \mathbb{R}$ defined by

$$F(x) = \int_{-1}^{1} e^{xt} dt$$

is continuous. If $x \neq 0$, the integrand has $t \mapsto e^{xt}/x$ as primitive, from which it follows that $F(x) = (e^x - e^{-x})/x$. On the other hand, for $x = 0$ the integrand is constant 1, and we see that $F(0) = 2$. The continuity of F gives $F(x) \to F(0) = 2$ as $x \to 0$. Of course, we could also derive this using l'Hôpital's theorem, see the notes Introduction to Analysis. △

Example 5.17 (Beta-function)**.** The *Euler Beta-function* is the function of two real variables p, q, defined by

$$B(p, q) := \int_0^1 t^{p-1} (1 - t)^{q-1} dt. \tag{5.6}$$

The integrand

$$b(p, q, t) := t^{p-1}(1 - t)^{q-1}$$

is a continuous function of $(p, q, t) \in [1, \infty) \times [1, \infty) \times [0, 1]$. By Theorem 5.12, it follows that B is continuous on $[1, \infty) \times [1, \infty)$. △

Example 5.18 (Gamma-function)**.** The above result is not directly applicable to functions defined via so-called improper integrals. For example, consider the Euler Gamma-function $\Gamma : (0, \infty) \to \mathbb{R}$, defined by

$$\Gamma(x) = \int_0^{\infty} t^{x-1} e^{-t} dt. \tag{5.7}$$

This is a function of the form $F(x) = \int_0^\infty f(x, t)dt$ with $f(x, t) = t^{x-1}e^{-t}$. There are two issues here. First, the interval of integration is unbounded above. Second, the function $f_x : t \mapsto f(x, t)$ is not defined at 0 for $0 < x < 1$. A general theory showing that the Gamma-function is continuous, and even C^∞, on the interval $(0, \infty)$ is explained in Chapter 9. △

5.4 Intermezzo: divided differences

Given a function $f : V \times [a, b] \to \mathbb{R}^p$ and its integral $G : V \to \mathbb{R}^p$, as in (5.4), we want to understand how derivatives of G relate to those of f. We will do this in Subsection 5.5. Before we get there we discuss an auxiliary result. Even if auxiliary, it is very interesting by itself and quite helpful to understand geometrically the process of differentiation.

First, consider the following simplified setting:

Lemma 5.19. *Fix a function $f : \mathbb{R} \to \mathbb{R}$ with $f(0) = 0$ and differentiable at 0. Then, the function $q : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$*

$$q(x, t) := \begin{cases} \frac{f(xt)}{t} & \text{if} \quad t \neq 0 \\ f'(0)x & \text{if} \quad t = 0. \end{cases}$$

is continuous.

Proof. q is defined in terms of f by composing with $(x,t) \mapsto xt$ and dividing by t. As such, it is clearly continuous if $t \neq 0$. We therefore must show continuity at each $(x_0, 0)$. By differentiability of f we know that $\lim_{h\to 0} \frac{f(h)}{h} = f'(0)$, so we can set $h = xt$ to deduce:

$$\lim_{(x,t)\to(x_0,0)} q(x,t) = \lim_{(x,t)\to(x_0,0)} x\frac{f(xt)}{tx} = x_0 \lim_{(x,t)\to(x_0,0)} \frac{f(xt)}{tx} = x_0 f'(0)$$

showing continuity at $(x_0, 0)$. □

The following geometric interpretation is helpful: Think of t as a zoom parameter that helps us look more closely at f close to $x = 0$. Namely, if t is small, $x \mapsto xt$ amounts to changing coordinates to look at points closer to 0. Then $\frac{f(xt)}{t}$ can be seen as the result of taking these small points xt and scaling up their values $f(xt)$. I.e. We zoom in simultaneously in the domain and the target of f. This sequence of "zoomed-in" maps has a limit: the derivative $f'(0)$, which can be seen as a linear map $\mathbb{R} \to \mathbb{R}$.

This lemma can be generalised to the multivariate setting. We do it in a directional manner: we zoom in the first variable, leaving the rest as parameters. Other versions, handling all variables at once, also exist, but they are beyond the scope of the course.

Proposition 5.20. *Let I be an interval in $\mathbb{R}$ and V an open subset of $\mathbb{R}^n$. Let $f : I \times V \to \mathbb{R}$ be differentiable with respect to the first variable, and assume that the function $D_1 f$ is continuous on $I \times V$. Define the function $q : I \times I \times V \to \mathbb{R}$ by*

$$q(t,s,x) := \begin{cases} \frac{f(t,x)-f(s,x)}{t-s} & \text{if } \quad s \neq t \\ D_1 f(s,x) & \text{if } \quad t = s. \end{cases}$$

Then q is continuous on $I \times I \times V$.

Proof. According to the fundamental theorem of calculus applied to the first variable:

$$\begin{aligned} f(t,x) - f(s,x) &= \int_s^t D_1 f(b,x)\, db \\ &= \int_0^1 \frac{d}{da} f(s + a(t-s), x)\, da \\ &= \int_0^1 D_1 f(s + a(t-s), x)\,(t-s)\, da \\ &= (t-s) \int_0^1 D_1 f(s + a(t-s), x)\, da. \end{aligned}$$

In the second equality we changed variable, setting $b = s + a(t - s)$. In the third equality we applied the chain rule Lemma 3.51. From this it follows that

$$q(t,s,x) = \int_0^1 D_1 f(s + a(t-s), x)\, da \tag{5.8}$$

for every $t \neq s$. Formula (5.8) is also valid when $s = t$, because in that case the integrand on the right-hand side is equal to $D_1 f(s,x)$ for every a. Now apply Theorem 5.12 with (t,s,x) as parameters to conclude that q is continuous. □

5.5 Differentiation under the integral sign

We can now use Proposition 5.20 in the proof of the following theorem about differentiation under the integral sign. We state it first for two variables. The general case is stated below as Corollary 5.22.

Theorem 5.21. *Consider intervals $I = [a, b]$ and $J \subset \mathbb{R}$. Fix coordinates (t, x) in the product $I \times J$. Let a function $f : I \times J \to \mathbb{R}$ be given, and assume that the following conditions hold:*

(a) *For every $x \in J$, the function $t \mapsto f(t, x)$ is Riemann-integrable over I.*

(b) *The function f is partially differentiable with respect to the x-variable and $D_x f$ is continuous on $I \times J$.*

Then, the function $G : J \to \mathbb{R}^n$ defined by

$$G(x) := \int_a^b f(t, x)dt$$

is a differentiable function of x that satisfies:

$$G'(x) = \int_a^b D_x f(t, x)dt. \tag{5.9}$$

Suppose moreover that f is continuous. Then the function $F : I \times J \to \mathbb{R}^n$ defined by

$$F(t, x) := \int_a^t f(s, x)ds$$

is differentiable and

$$D_x F(t, x) = \int_a^t D_x f(s, x)ds.$$

The formula (5.9) says that we may "swap differentiation and integration", in the sense that the derivative with respect to x of the integral over t is equal to the integral over t of the derivative with respect to x. The idea of the proof is to relate differentiation to continuity (using Proposition 5.20) and use the fact that integration preserves continuity (Theorem 5.12).

Proof. We first prove the statement for G. The case of F is similar and left to the reader. As in Proposition 5.20, we define a function $q : I \times J \times J \to \mathbb{R}$ from f by setting $q(t, x, x') := (f(t, x) - f(t, x'))/(x - x')$ for $x \neq x'$ and $q(t, x, x) := D_x f(t, x)$. This function is continuous according to the lemma, so we can integrate it using Theorem 5.12, yielding

$$Q(x, x') := \int_a^b q(t, x, x')dt,$$

which is a continuous function $J \times J \to \mathbb{R}$. Moreover, by the definition of q:

$$Q(x, x') = \begin{cases} \frac{G(x)-G(x')}{x-x'} & \text{if} \quad x \neq x' \\ \int_a^b D_x f(t, x)dt & \text{if} \quad x = x'. \end{cases}$$

From the continuity of Q, it follows that:

$$\lim_{x' \to x} \frac{G(x) - G(x')}{x - x'} = \int_a^b (D_x f)(t, x)dt,$$

which is exactly Equation 5.9.

The proof for F is very similar and left to the reader. The key difference is that we assume continuity of f so that $D_t F = f$ is continuous. This, together with the continuity of $D_x F$, implies that F is continuously partially differentiable and thus totally differentiable (Theorem 3.29). □

This result can now be applied repeatedly for higher-order derivatives:

Corollary 5.22. *Let V be an open subset of $\mathbb{R}^n$ and $[a,b] \subset \mathbb{R}$ an interval. Let $f : [a,b] \times V \to \mathbb{R}$ be a function that is k times partially differentiable with respect to the last n variables. Assume further that the partial derivatives $D_{j(l)} \cdots D_{j(1)} f$ are continuous, for all collections of indices $(j(1), \ldots, j(l)) \in \mathbb{N}^l$, with $0 \leq l \leq k$.*

Consider then the functions $F(t,x) = \int_a^t f(s,x)ds$ and $G(x) = \int_a^b f(t,x)dt$. Then F is k-times differentiable in x and satisfies

$$D_{j(l)} \cdots D_{j(1)} F(t,x) = \int_a^t D_{j(l)} \cdots D_{j(1)} f(s,x)ds.$$

Similarly, G is C^k and satisfies:

$$D_{j(l)} \cdots D_{j(1)} G(x) = \int_a^b D_{j(l)} \cdots D_{j(1)} f(t,x)dt.$$

Proof. This is proven by induction over k, using Theorem 5.21 in the induction step. □

Using this result we can establish an analogue of Proposition 5.20 involving higher derivatives:

Corollary 5.23. *Let I be an interval in $\mathbb{R}$, $k \in \mathbb{Z}_{\geq 0}$, and $f \in C^{k+1}(I, \mathbb{R})$. Define*

$$q(t,s) = \begin{cases} \frac{f(t)-f(s)}{t-s} & \text{if} \quad t \neq s \\ f'(t) & \text{if} \quad t = s. \end{cases}$$

Then $q \in C^k(I \times I, \mathbb{R})$.

Proof. The differentiability statement is most interesting at the points (t,t), because on the set $\{(t,s) \mid t \neq s\}$, the function $q(t,s)$ is already C^{k+1}. According to Proposition 5.20, $q(t,s)$ is continuous. Formula (5.8) gives

$$q(t,s) = \int_0^1 f'(s + a(t-s))da,$$

where the integrand is a C^k function of the variables (t,s). Applying Corollary 5.22 gives $q \in C^k$. □

We conclude this subsection with an application:

Example 5.24. Consider the function $f : \mathbb{R} \to \mathbb{R}$ given by $f(x) = \sin(x)$. Since f is smooth, Corollary 5.23 tells us that the corresponding function $q : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ is in $C^\infty(\mathbb{R} \times \mathbb{R})$. The function $\sigma : x \mapsto q(x,0)$, $\mathbb{R} \to \mathbb{R}$ is given by

$$\sigma(x) = \begin{cases} (\sin x)/x & \text{if} \quad x \neq 0, \\ 1 & \text{if} \quad x = 0. \end{cases}$$

We conclude that $\sigma \in C^\infty([0,1])$. △

5.6 Switching the order of integration

In the context of "switching theorems", we give the following result. It states that the order of integration in the definition of the Riemann integral (Definition 5.1) does not matter:

Theorem 5.25. *Consider a continuous function $f : [a, b] \times [c, d] \to \mathbb{R}$. Then:*

$$\int_c^d \left(\int_a^b f(t, s)dt \right) ds = \int_a^b \left(\int_c^d f(t, s)ds \right) dt. \tag{5.10}$$

Proof. Integrating in t we can produce a primitive:

$$\varphi(t, s) := \int_a^t f(x, s)dx.$$

Note that $\varphi(a, s) = 0$ for every s. The fundamental theorem of calculus tells us that $t \mapsto \varphi(t, s)$ is differentiable with derivative

$$D_t\varphi(t, s) = f(t, s),$$

which is a continuous function of both variables. Integrating along s then yields a total integral:

$$\Phi(t) := \int_c^d \varphi(t, s)ds = \int_c^d \int_a^t f(x, s)dxs.$$

Note that $\Phi(a) = 0$. Observe that $\varphi(t, s)$ computes the area under the graph of f along the $[a, t] \times \{s\}$ segment and thus $\Phi(t)$ computes the volume under the graph of f over the rectangle $[a, t] \times [c, d]$.

Theorem 5.21 gives that the function Φ is differentiable on $[a, b]$, with derivative

$$\Phi'(t) := \int_c^d D_t\varphi(t, s)ds = \int_c^d f(t, s)ds.$$

Integrating this over $t \in [a, b]$ now gives

$$\begin{aligned} \int_a^b \left(\int_c^d f(t, s)ds \right) dt &= \int_a^b \Phi'(t)dt \\ &= \Phi(b) - \Phi(a) = \Phi(b) \\ &= \int_c^d \left(\int_a^b f(t, s)dt \right) ds. \end{aligned}$$

Concluding the proof. □

It follows that we can also write the identity (5.25) without brackets:

$$\int_c^d \int_a^b f(t,\ s)\,\mathrm{d}t\,\mathrm{d}s = \int_a^b \int_c^d f(t,\ s)\,\mathrm{d}s\,\mathrm{d}t,$$

since from the order of the integral signs and of ds and dt it is clear in which order the integrations are to be taken.

It is interesting that Theorem 5.21 can also be derived from Theorem 5.25.

Proof of Theorem 5.21 assuming Theorem 5.25. Assume f is a function as in Theorem 5.21. Let $c \in I$. For every $x \in I$ with $x > c$, it follows, by the fundamental theorem of calculus, that for every $t \in [a, b]$:

$$f(x,t) - f(c,t) = \int_c^x D_1 f(s,t) ds.$$

Thus

$$\begin{aligned} \int_a^b f(x,\, t)\,\mathrm{d}t &= \int_a^b \left(f(c,\, t) + \int_c^x D_1 f(s,\, t)\,\mathrm{d}s \right) \mathrm{d}t \\ &= \int_a^b f(c,\, t)\,dt + \int_a^b \int_c^x D_1 f(s,\, t)\,\mathrm{d}s\,\mathrm{d}t \\ &= \int_a^b f(c,\, t)\,\mathrm{d}t + \int_c^x \int_a^b D_1 f(s,\, t)\,\mathrm{d}t\ \mathrm{d}s. \end{aligned}$$

In the third equality we have applied Theorem 5.25 to $D_1 f$, which is continuous by assumption. Observe that, on the right-hand side, the variable x appears as the upper bound of an integration interval. As such, the right-hand side is differentiable with respect to x. The same is thus true for the left-hand side, with derivative equal to

$$\frac{d}{dx} \int_a^b f(x,\, t)\,\mathrm{d}t = \int_a^b D_1 f(x,\, t)\,\mathrm{d}t. \qquad \square$$

Remark 5.26. In the courses *Analyse in Meer Variabelen* and *Maat en Integratie* the switching of the integration order (*Fubini's theorem*) is further developed. This result holds for a class of functions of several variables that is much larger than the class of continuous functions. △

6 The inverse function theorem and its applications

The previous chapters have built the elementary machinery of differentiation and integration. In this chapter and the next we look into more advanced results. The goal in this chapter is to state and prove the inverse function theorem, and to work out some of its most important consequences.

An idea that you are familiar with, and that we reviewed in Chapter 2, is that it is convenient to change basis when studying linear maps. The reason is simple: the standard basis $\{e_1, \cdots, e_n\}$ of $\mathbb{R}^n$ is not important by itself, so we should feel free to use a different one if that simplifies computations.

This reasoning applies across all Mathematics, not just to Linear Algebra. In our context:

Definition 6.1. Let U and V be open subsets of $\mathbb{R}^n$. A bijection $f : U \to V$ is a C^1**-diffeomorphism** if both f and its inverse $f^{-1} : V \to U$ are continuously differentiable. U and V are said to be *diffeomorphic*. △

This says that we should think of U and V as basically the same object, seen from two different perspectives. The map f is a change of coordinates relating the two.

Example 6.2. Consider a diffeomorphism $\phi : U \to V$ and an open $U' \subset U$. Then $\phi|_{U'} : U' \to \phi(U')$ is also a diffeomorphism. △

Example 6.3. An invertible linear map $A : \mathbb{R}^n \to \mathbb{R}^n$ is in particular a diffeomorphism. Consider for instance $A : \mathbb{R}^2 \to \mathbb{R}^2$ given by $A(x, y) = (2x, y)$. We can then restrict it to the unit open ball $B \subset \mathbb{R}^2$, yielding a diffeomorphism

$$B^2 \to A(B^2) = \{(x, y) \in \mathbb{R}^2 \mid (x/2)^2 + y^2 < 1\}$$

between the ball and the ellipsoid. △

Our goals in this chapter are:

- To establish the inverse function theorem, in its local (Theorem 6.7) and global (Theorem 6.12) versions. They tell us that, in order to show that a funtion is a diffeomorphism, we do not need to explicitly compute the inverse. As you can imagine, this simplifies things considerably.
- To look into some of the typical examples of diffeomorphisms (Subsection 6.2).
- To start the study of (k-dimensional) submanifolds (Definition 6.27). These are the "nicest" subsets of $\mathbb{R}^n$, in the sense that, up to a local change of coordinates, they resemble the linear subspace $\mathbb{R}^k \subset \mathbb{R}^n$. The existence of such a change of coordinates is established in the regular value theorem (Theorem 6.34).
- To study the theory of Lagrange multipliers (Theorem 6.37), which deals with the critical points of functions $f|_N : N \to \mathbb{R}$, where $f : U \to \mathbb{R}$ is a C^1-function defined over an open and $N \subset U$ is a submanifold.

These topics will reappear in many places. (Sub)manifolds are central objects in Analysis, Geometry, and Topology. The theory of Lagrange multipliers is the theory of optimisation (finding critical points/maxima/minima) under constraints (when restricting to subsets) and, as such, it appears everywhere in theory and applications.

6.1 The inverse function theorem

6.1.1 The one variable case

The proof of the multivariate inverse function theorem requires a number of non-trivial ideas. For convenience, let us first recall the inverse function theorem for functions of one variable, including its proof.

Lemma 6.4. *Let $\mathcal{O} \subset \mathbb{R}$ be an open set and $f : \mathcal{O} \to \mathbb{R}$ a C^1-function. Let $a \in \mathcal{O}$ and suppose that $f'(a) \neq 0$. Then there exist open neighborhoods U of a in $\mathcal{O}$ and V of $f(a)$ in $\mathbb{R}$ such that $f|_U$ maps the neighborhood U bijectively onto V, while the inverse $g := (f|_U)^{-1}$ is a C^1-map. Moreover*

$$g'(f(x)) = 1/f'(x) \qquad \text{for all } x \in U.$$

Proof. From $f'(a) \neq 0$ it follows that $f'(a) > 0$ or $f'(a) < 0$. We treat the first case; the second one is similar.

The first step is showing that there exists a set-theoretical inverse once we restrict f to a sufficiently small neighbourhood of f. By continuity of f', there exists $\delta > 0$ such that $(a - 2\delta, a + 2\delta) \subset \mathcal{O}$ and $f'(x) > 0$ for all $x \in (a - 2\delta, a + 2\delta)$. By Lemma 4.25 it follows that f is strictly increasing on $[a - \delta, a + \delta]$. In particular, it is injective. Hence f maps the interval $U := (a - \delta, a + \delta)$ onto the interval $V := (f(a - \delta), f(a + \delta))$. Since $f : U \to V$ is bijective by construction, it has an inverse $g : V \to U$.

For latter use do observe that taking $\delta \to 0$ makes U shrink to a and V shrink to $f(a)$, by continuity of f.

The second step is showing that g is differentiable at every point $y_0 \in V$. To do so, we will make use of the mean value theorem. Write $x_0 = g(y_0)$. For each $y \in V$ write $x = g(y)$ and note that one has

$$y - y_0 = f(x) - f(x_0) = f'(\xi_x)(x - x_0).$$

for some $\xi_x \in [x_0, x]$. Our observation shows that $y \to y_0$ implies that $x \to x_0$ and thus $\xi_x \to x_0$. Since $\xi_x \in U$ we have $f'(\xi_x) > 0$, so we can divide:

$$g(y) - g(y_0) = x - x_0 = f'(\xi_x)^{-1}(y - y_0).$$

Since f' is continuous it follows that:

$$g'(y_0) = \lim_{y \to y_0} \frac{g(y) - g(y_0)}{y - y_0} = \lim_{y \to y_0} 1/f'(\xi_x) = 1/f'(x_0). \qquad \square$$

Example 6.5. In *Inleiding Analyse* the logarithmic function $\log : (0\,, \infty) \to \mathbb{R}$ was introduced by means of the Riemann integral:

$$\log x := \int_0^x \frac{1}{t} dt.$$

From the fundamental theorem of calculus it followed that $f := \log$ is a C^1-function with derivative $f'(x) = \frac{1}{x}$. It was then shown that f is a strictly monotone bijection from $(0, \infty)$ onto $\mathbb{R}$. The exponential function $g : \mathbb{R} \to (0, \infty)$ given by $y \mapsto e^y$ was introduced as the inverse of log. Using the lemma above it followed that g is continuously differentiable with derivative

$$g'(\log x) = 1/f'(x) = x = g(\log x),$$

hence $g'(y) = g(y)$ for all $y \in \mathbb{R}$. $\triangle$

6.1.2 Statement of the local inverse function theorem

The following is an elementary but important result. It gives us a necessary criterion, in terms of the total derivative, for a function f to be a diffeomorphism.

Proposition 6.6. *Fix a C^1 diffeomorphism $f : U \to V$ and a point $a \in U$. Then, the total derivative $Df(a) : \mathbb{R}^n \to \mathbb{R}^n$ is an invertible linear map and, moreover:*

$$D(f^{-1})(f(a)) = Df(a)^{-1}. \tag{6.1}$$

Proof. We have $f^{-1} \circ f = \mathrm{id}_U$. As such, the chain rule (Theorem 3.48) combined with Example 3.16, implies that

$$D(f^{-1})(f(a)) \circ Df(a) = D(\mathrm{id}_U)(a) = \mathrm{Id}$$

holds for every $a \in U$, where Id denotes the identity matrix. We thus see that the linear map $Df(a) : \mathbb{R}^n \to \mathbb{R}^n$ is invertible with inverse $D(f^{-1})(f(a))$. It follows that it must agree with $Df(a)^{-1}$ due to the uniqueness of the inverse. □

The inverse function theorem provides a very powerful converse of this result.

Theorem 6.7 (Inverse function theorem, local version)**.** *Let $\mathcal{O} \subset \mathbb{R}^n$ be open, $f : \mathcal{O} \to \mathbb{R}^n$ a C^1-map, and $a \in \mathcal{O}$ a point. If the linear map $Df(a) : \mathbb{R}^n \to \mathbb{R}^n$ is invertible, then there exists an open neighborhood U of a in $\mathcal{O}$ such that*

(a) *$f(U)$ is open in $\mathbb{R}^n$,*

(b) *the map $f|_U$ is a C^1 diffeomorphism $U \to f(U)$.*

The proof of this result is quite involved and in fact requires some non-trivial ideas. We will spend the rest of this subsection building towards it. Nonetheless, the argument follows the same logic as in the one variable case: constructing a set-theoretical inverse, locally, and then showing that the inverse is indeed C^1.

6.1.3 Step I: Injectivity via an improved mean value theorem

Our first goal is to show that $Df(a)$ being injective implies that f is locally injective (i.e. injective in a sufficiently small neighbourhood of a). This is the content of Corollary 6.9. To establish it, we will need the following:

Theorem 6.8 (Mean Value Theorem 2.0)**.** *Let $U \subset \mathbb{R}^n$ be a convex open subset and $f : U \to \mathbb{R}^p$ a C^1-map. Then there exists a continuous map $L : U \times U \to \mathrm{Lin}(\mathbb{R}^n, \mathbb{R}^p)$ such that*

$$f(y) - f(x) = L(x,y)(y-x), \qquad x, y \in U,$$

and

$$L(x,x) = Df(x), \qquad x \in U.$$

Proof. This proof contains no new ideas. It follows the exact same argument as in the original mean value Theorem 3.40.

Consider $x, y \in U$. By convexity, the line segment $[x,y]$ lies entirely in U; we parametrise it as $c = c_{x,y} : [0,1] \to U$ using the formula $t \mapsto x + t(y-x)$. According to Lemma 3.51, the derivative with respect to t of the map $f \circ c_{x,y} : [0,1] \to \mathbb{R}^p$ reads:

$$\frac{d}{dt}(f \circ c)(t) = Df(c(t))c'(t) = Df(x + t(y-x))(y-x).$$

By Theorem 5.5 it follows that

$$\begin{aligned} f(y) - f(x) &= f \circ c(1) - f \circ c(0) = \int_0^1 (f \circ c)'(t)dt \\ &= \int_0^1 Df(x + t(y-x))(y-x)dt = L(x,y)(y-x) \end{aligned}$$

with $L(x,y)$ the map $\mathbb{R}^n \to \mathbb{R}^p$ given by

$$L(x,y)(v) := \int_0^1 Df(x + t(y-x))v dt.$$

By linearity of the $\mathbb{R}^p$-valued Riemann integration, it is clear that $L(x,y) \in \mathrm{Lin}(\mathbb{R}^n, \mathbb{R}^p)$ for all $x, y \in \mathbb{R}^n$.

The matrix-valued function $[0,1] \times U \times U \to \mathrm{Lin}(\mathbb{R}^n, \mathbb{R}^p)$ given by $(t,x,y) \mapsto Df(x + (t(y - x))$ is continuous. $L(x,y)$ is obtained from it by integrating coefficient-by-coefficient, so it is also continuous, according to Theorem 5.12.

For the second part let $x \in U$. Consider the remainder

$$R(h) := f(x+h) - [f(x) + L(x,x)(h)] = [L(x, x+h) - L(x,x)](h).$$

with $h \in -x + U$. Then

$$\|h\|^{-1}\|R(h)\| \leq \|L(x, x+h) - L(x,x)\| \to 0 \qquad (h \to 0),$$

by continuity of L at (x,x). From this we conclude that f is totally differentiable at x with total derivative equal to $Df(x) = L(x,x)$. □

We go back to the proof of Theorem 6.7. As such, and for the rest of the subsection, we assume that $f : U \to \mathbb{R}^n$ satisfies the assumptions of the theorem.

Corollary 6.9. *Let $a \in U$ and assume that $Df(a) : \mathbb{R}^n \to \mathbb{R}^n$ is an invertible linear map. Then there exists $\delta > 0$ such that $B(a;\delta) \subset U$ and such that the restriction of f to $B(a;\delta)$ is injective.*

Proof. By possibly shrinking U to an open ball we may assume that U is convex. Let $L : U \times U \to \mathrm{Lin}(\mathbb{R}^n, \mathbb{R}^n)$ be as in Theorem 6.8. All matrix coefficients $L_{ij}(x,y) = \langle e_i, L(x,y)e_j \rangle$ are continuous. Since $\det L(x,y)$ is a polynomial in these matrix coefficients, it follows by the substitution rule that $(x,y) \mapsto \det L(x,y)$ is continuous. Since $L(a,a) = Df(a)$ is invertible and has thus non-zero determinant, there exists $\delta' > 0$ such that for all $(x,y) \in \mathbb{R}^n \times \mathbb{R}^n$ with $\|(x,y) - (a,a)\| < \delta'$ we have

$$|\det L(x,y)| > \tfrac{1}{2}|\det Df(a)| > 0.$$

In particular, the linear map $L(x,y)$ is invertible for such (x,y). Choose $\delta > 0$ such that $x, y \in B(a;\delta) \implies \|(x,y) - (a,a)\| < \delta'$. Then for all $x, y \in B(a;\delta)$ we have

$$f(y) - f(x) = 0 \implies L(x,y)(y-x) = 0 \implies y - x = 0,$$

so f is injective on $B(a;\delta)$. □

6.1.4 Step II: Surjectivity

Similarly, we now show that f is locally surjective (i.e. its image contains a neighbourhood of $f(a)$). We will need the following auxiliary lemma:

Lemma 6.10. *Let $U \subset \mathbb{R}^n$ and $f : U \to \mathbb{R}^p$ a C^1-function. Let $b \in \mathbb{R}^p$ and define $G : U \to \mathbb{R}$ by*

$$G(x) := \langle f(x) - b, f(x) - b \rangle.$$

Then G is a C^1-function. At any $a \in U$ its derivative $DG(a) : \mathbb{R}^n \to \mathbb{R}$ is given by

$$v \mapsto 2\langle Df(a)v, f(a) - b \rangle.$$

Proof. By the rules of partial differentiation G is a C^1-function, and for each $a \in U$ its jth partial derivative reads:

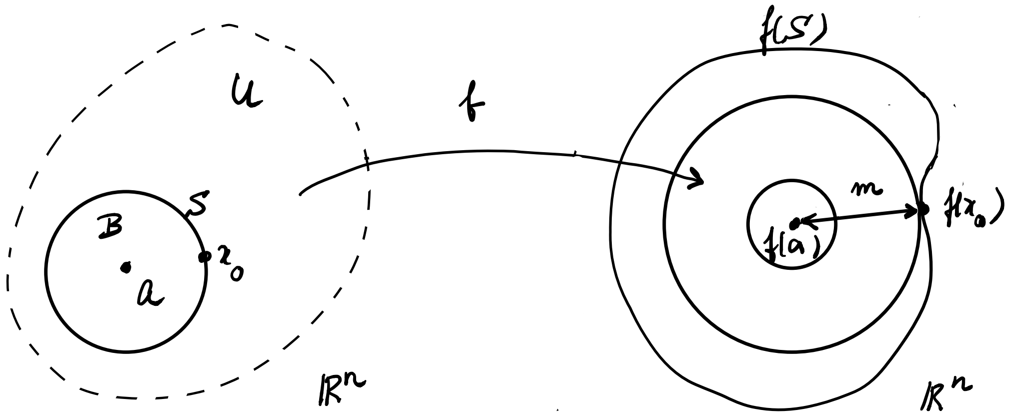
$$D_jG(a) = 2\langle D_jf(a), f(a) - b \rangle.$$

It follows that G is totally differentiable and its derivative $DG(a) : \mathbb{R}^n \to \mathbb{R}$ is given by

$$DG(a)(v) = \sum_{j=1}^{n} v_j D_jG(a) = 2\left\langle \sum_{j=1}^{n} v_j D_jf(a), f(a) - b \right\rangle = 2\langle Df(a)v, f(a) - b \rangle. \qquad \square$$

Thus:

Lemma 6.11. *Let $a \in U$ and assume that $Df(a) : \mathbb{R}^n \to \mathbb{R}^n$ is an invertible linear map. Then there exists $r > 0$ such that $f(U) \supset B(f(a); r)$.*



Figuur 7: Proof of Lemma 6.11.

Proof. According to Corollary 6.9, by shrinking U if necessary, we may assume that $f : U \to \mathbb{R}^n$ is injective.

Since f is C^1, the component-wise partial derivatives $D_jf_i : U \to \mathbb{R}$ are continuous. These can be combined to produce the determinant map $U \to \mathbb{R}$ using the formula $x \mapsto \det Df(x)$; by the product and sum rules it is continuous. By invertibility $|\det Df(a)| > 0$, so there exists $\delta_1 > 0$ such that $B(a;\delta_1) \subset U$ and

$$x \in B(a;\delta_1) \implies |\det Df(x)| \geq \tfrac{1}{2}|\det Df(a)| > 0.$$

Choose $0 < \delta < \delta_1$ and write B for the open ball $B(a;\delta)$ and $\bar{B}$ for its closure.

The spherical boundary $S := \bar{B} \setminus B$ is closed and bounded in $\mathbb{R}^n$, hence compact. The function

$$d : S \to \mathbb{R}, \quad x \mapsto \|f(x) - f(a)\|$$

is continuous, so it attains a minimum m at some $x_0 \in S$. Then $x_0 \neq a$, and by injectivity of f on $\bar{B}$ we have $f(x_0) \neq f(a)$, so $m > 0$. Choose $r = m/2$. We will show that this r works, i.e. $B(f(a);r) \subset f(\bar{B}) \subset f(U)$.

Given $b \in B(f(a);r)$ we consider the continuous function $g : \bar{B} \to \mathbb{R}$ given by $g(x) := \|f(x) - b\|$. I.e. we are measuring how close the image of $f|_{\bar{B}}$ is from b. What we want to show is that $g(x) = 0$ for some x, since that will show $b \in f(\bar{B})$.

Since g is continuous and $\bar{B}$ is compact, g attains a minimum, say at x_1. Using minimality we deduce $g(x_1) \leq g(a) = \|f(a) - b\| < r$. If $x_1 \in S$ then the triangular inequality tells us:

$$g(x_1) = \|f(x_1) - b\| \geq \|f(x_1) - f(a)\| - \|f(a) - b\| \geq m - r = r,$$

a contradiction. It follows that x_1 must be in the interior $\bar{B} \setminus S = B$.

Consider then the function $G : B \to \mathbb{R}$ given by

$$G(x) := g^2(x) = \langle f(x) - b, f(x) - b \rangle.$$

Thanks to the squaring, it is differentiable. x_1 is still a minimum for G. Since B is open, the variational principle of Lemma 3.33 implies that $DG(x_1) = 0$. According to the auxiliary Lemma 6.10:

$$0 = DG(x_1)(v) = 2\langle Df(x_1)v, f(x_1) - b \rangle.$$

We conclude that $\langle Df(x_1)v, f(x_1) - b \rangle = 0$ for all $v \in \mathbb{R}^n$. Since $Df(x_1)$ is invertible it follows that $\langle w, f(x_1) - b \rangle = 0$ for all $w \in \mathbb{R}^n$. Hence $f(x_1) = b$. Thus $b \in f(B)$. We conclude that $B(f(a);r) \subset f(B) \subset f(U)$. □

Recall that an extremum, in an open subset, is a critical point (Lemma 3.33). This is not necessarily true in a closed set, like $\bar{B}$, which is why we had to verify that $x_1 \notin S$. In a closed set it could happen that a maximum/minimum appears at the boundary and is non-critical (i.e. it has non-zero total derivative). An easy example is an increasing function $[0,1] \to \mathbb{R}$. In that case 0 is a minimum and 1 is a maximum, but they do not need to be critical points. In Section 6.4 we will explore how such extrema can be found.

6.1.5 Step III: Proving the inverse is C^1

Proof of Theorem 6.7. We first summarise the result up to now. Using Corollary 6.9 and Lemma 6.11 we can choose $U := B(a;\delta)$ with $\delta > 0$ small enough so that $f|_U$ is injective and $Df(x)$ invertible for all $x \in U$. Moreover, for each $x \in U$ and each open neighbourhood V of x in U there exists $r_{x,V} > 0$ such that $f(V) \supset B(f(x); r_{x,V})$. Thus $f(x)$ is an interior point of $f(V)$. We conclude that $f(V)$ is open in $\mathbb{R}^n$. In particular $f(U)$ is open in $\mathbb{R}^n$.

By injectivity of f on U we have that $f|_U : U \to f(U)$ is bijective. We denote its inverse $f(U) \to U$ by g. We claim that g is continuous. Indeed, given any open $V \subset U$ we have $g^{-1}(V) = f(V)$, which is also open. We have thus shown that f is locally invertible by a continuous function.

The second and final step is to show that g is also C^1. Let $x_0 \in U$ and write $y_0 := f(x_0)$ for its image. Let $L : U \times U \to \mathrm{Lin}(\mathbb{R}^n, \mathbb{R}^n)$ be defined as in Theorem 6.8. The idea at this point is very similar to what we did for Theorems 3.29 and 3.48 (and also for the one-variable case). We will prove a mean value theorem for g (using L) and from it, it will follow that $g \in C^1$.

The map $\det \circ L : U \times U \to \mathbb{R}$ is continuous, and nonzero at (a, a). Thus by choosing $\delta > 0$ smaller if necessary we may assume that $\det L(x, y)$ is nonzero for all $x, y \in U$. It follows that $L(x, y) : \mathbb{R}^n \to \mathbb{R}^n$ is invertible. By Cramer's formula for the inverse of a matrix it now follows that the map

$$(x, y) \mapsto L(x, y)^{-1}, \qquad U \times U \to \mathrm{Lin}(\mathbb{R}^n, \mathbb{R}^n)$$

is continuous. Let $y \in f(U)$ and write $x := g(y)$. Then

$$y - y_0 = f(x) - f(x_0) = L(x, x_0)(x - x_0) = L(x, x_0)(g(y) - g(y_0))$$

using invertibility we deduce:

$$g(y) - g(y_0) = L(x, x_0)^{-1}(y - y_0).$$

This yields the remainder formula:

$$R(y) := g(y) - [g(y_0) - L(x_0, x_0)^{-1}(y - y_0)] = [L(x, x_0)^{-1} - L(x_0, x_0)^{-1}](y - y_0)$$

which satisfies

$$\frac{||R(y)||}{||y - y_0||} = ||L(g(y), g(y_0))^{-1} - L(g(y), g(y_0))^{-1}|| \to 0$$

as $y \to y_0$. This uses the continuity of L and g. We have thus shown that g is differentiable at y_0 with derivative

$$Dg(y_0) = L(x_0, x_0)^{-1} = Df(x_0)^{-1}.$$

By Cramer's formula we see that $x \mapsto Df(x)^{-1}$ is continuous $U \to \mathrm{Lin}(\mathbb{R}^n, \mathbb{R}^n)$, and by the substitution rule it follows that Dg is also a continuous function $f(U) \to \mathrm{Lin}(\mathbb{R}^n, \mathbb{R}^n)$. As such, g is a C^1-map $f(U) \to U$ and the proof is complete. □

6.1.6 The global inverse function theorem

The inverse function theorem is often applied in the following form:

Theorem 6.12 (Inverse function theorem, global version)**.** *Let $U \subset \mathbb{R}^n$ be an open subset and $f : U \to \mathbb{R}^n$ a C^1-map such that f is injective on U and $Df(x)$ is invertible for every $x \in U$. Then $V := f(U)$ is open in $\mathbb{R}^n$ and $f : U \to V$ is a diffeomorphism.*

Proof. From the assumptions it follows that $f : U \to V$ is bijective. We denote its inverse by $g : V \to U$.

It is sufficient to show that V is open and that g is a C^1-map. Let $b \in V$ be arbitrary and let $a := g(b) \in U$. Note $f(a) = b$. According to the local version of the inverse function Theorem 6.7, there is an open neighborhood U_0 of a in U such that $V_0 := f(U_0)$ is open and $f|_{U_0}$ is a diffeomorphism from U_0 onto $V_0 := f(U_0)$. From this it follows that b is an interior point of V_0 and hence of V. Since b was arbitrary we conclude that V is open. Morever, this means that $f|_{U_0}$ has an inverse $h : V_0 \to U_0$ which is C^1. It is clear that $h = g|_{V_0}$, by uniqueness of the inverse. Thus g is a C^1-map on the neighborhood V_0 of b. Since $b \in V$ was arbitrary we conclude that g is a C^1-map $V \to U$. □

Remark 6.13. In the above, the derivative of f^{-1} at each point $b \in V$ is given by

$$D(f^{-1})(b) = Df(f^{-1}(b))^{-1}.$$

This is a direct consequence of Proposition 6.6. △

6.1.7 Just a change of perspective

You should think of a C^1-diffeomorphism $\phi : U \subset \mathbb{R}^n \to V \subset \mathbb{R}^n$ as a change of coordinates that preserves all the "C^1 information" of U and V. Let us remark that one can define in a straightforward manner also higher C^k-diffeomorphisms, which will then preserve the C^k information. One can also consider the corresponding C^0 notion:

Exercise 6.14. Suppose X and Y are metric spaces. Suppose $\phi : X \to Y$ is a *homeomorphism* (i.e. a continuous map with continuous inverse). Let $f : Y \to \mathbb{R}$ be a continuous function. Observe that $f \circ \phi$ is also continuous. Show that x is a (local) maximum of $f \circ \phi$ if and only if $\phi(x)$ is a (local) maximum of f. Same for minima. △

Observe that a C^1-diffeomorphism is in particular a homeomorphism. Homeomorphisms preserve all the "C^0 information" so, in particular, they preserve (local) maxima/minima. This is because being a (local) maximum is something that only depends on values of functions, not derivatives.

Critical points are defined in terms of the derivative and are thus C^1 information. They are preserved by C^1-diffeomorphisms:

Exercise 6.15. Suppose $\phi : U \to V$ is a C^1-diffeomorphism. Let $f : V \to \mathbb{R}$ be a C^1 function.

- Show that $f \circ \phi : U \to \mathbb{R}$ is C^1.
- Consider points $x \in U$ and $\phi(x) \in V$. Show that x is critical for $f \circ \phi$ if and only if $\phi(x)$ is critical for f. △

In particular:

Exercise 6.16. Suppose $\phi : U \to V$ is a C^1-diffeomorphism. Consider the vector spaces $C^1(U, \mathbb{R})$ and $C^1(V, \mathbb{R})$. Show that $f \mapsto f \circ \phi$ is a linear isomorphism $C^1(V, \mathbb{R}) \to C^1(U, \mathbb{R})$. This map is called the **pullback of functions**. △

It is instructive to see that this result is truly sharp. I.e. a homeomorphism need not preserve C^1 information:

Exercise 6.17. Consider the functions $\phi, f : \mathbb{R} \to \mathbb{R}$ given by $\phi(x) = x^3$ and $f(x) = x$.

- Show that ϕ is a homeomorphism (Hint: you can explicitly write an inverse and see that it is continuous).

- Show that ϕ is not a C^1-diffeomorphism.
- Show that f has no critical points but $f \circ \phi$ does.
- Show that $f \circ \phi^{-1}$ is not even differentiable.

I.e. composing with ϕ has introduced new critical points and composing with ϕ^{-1} replaces f by a function that is not even C^1. △

In the same manner, a C^1-diffemorphism need not preserve C^2 information:

Example 6.18. Consider the function $g : \mathbb{R} \to \mathbb{R}$ given by $g(x) = 1 + |x|$. This function is continuous and strictly positive. As such, its primitive $\phi(x) = \int_0^x g(t)dt$ is C^1 and strictly increasing (and thus injective). In fact, ϕ is seen to be a C^1-diffeomorphism because it is also surjective (note that $\phi'(x) \geq 0$). Explicitly, $\phi(x) = x \pm x^2$ depending on the sign of x.

However, ϕ'' has a discontinuity at 0, so ϕ is not C^2. In particular, if we take $f : \mathbb{R} \to \mathbb{R}$ given by $f(x) = x$, which is C^2, we see that $f \circ \phi = \phi$ is not C^2. △

6.2 Examples

The inverse function theorem provides us with a tool to describe open subsets of $\mathbb{R}^n$ using (differentiable) coordinates other than the standard Euclidean ones.

> The following examples, even if classic, explain the usual strategy to handle diffeomorphisms. I recommend that you read them carefully in order to do similar exercises.

6.2.1 Polar coordinates

As a first example, we consider polar coordinates. Let $F : (0, \infty) \times \mathbb{R} \to \mathbb{R}^2$ be defined by

$$F(r, \phi) := (r \cos \phi, r \sin \phi).$$

Then F is a C^1 mapping, with Jacobian matrix

$$DF(r, \phi) = \begin{pmatrix} \cos \phi & -r \sin \phi \\ \sin \phi & r \cos \phi \end{pmatrix}$$

The determinant of this matrix is given by

$$\det DF(r, \phi) = r \cos^2 \phi + r \sin^2 \phi = r.$$

We therefore see that $DF(r, \phi)$ is invertible for all $r > 0$ and $\phi \in \mathbb{R}$. According to the local inverse function Theorem 6.7, F is a *local* diffeomorphism. I.e. for each (r, ϕ) there is some neighbourhood U such that $F|_U : U \to F(U)$ is a diffeomorphism.

Now we ask ourselves whether F is a *global* diffeomorphism. The answer is *no*: F is not surjective (it misses zero) and is also not injective. Indeed, $F(r, \phi) = F(r, \phi + 2k\pi)$ for every integer k. That is, over the whole of $\mathbb{R}^2 \setminus \{0\}$ polar coordinates are not truly coordinates, since the angle of a point is not well-defined.

Still, suppose $U \subset (0, \infty) \times \mathbb{R}$ is an open subset on which F is injective. Then $V := F(U)$ is open, by the global inverse function Theorem 6.12, and $F : U \to V$ is a diffeomorphism. The question then is: *how large can we pick such an open U?*

A typical choice is $U = \{\phi \in (0, 2\pi)\}$. Its image $V = F(U)$ is $\mathbb{R}^2 \setminus \{y = 0, x \geq 0\}$. Let $G : V \to U$ be the inverse; then the components G_1, G_2 of G are called the polar coordinates on V. Traditionally, these are denoted by $G_1(x, y) = r(x, y)$ and $G_2(x, y) = \phi(x, y)$. Note that these are indeed global coordinates on V.

The partial derivatives of $r(x, y)$ and $\phi(x, y)$ with respect to x and y can be found as follows. By Proposition 6.6, we have

$$DG(F(r, \phi)) = DF(r, \phi)^{-1} = \frac{1}{r} \begin{pmatrix} r\cos\phi & r\sin\phi \\ -\sin\phi & \cos\phi \end{pmatrix},$$

so

$$DG_1(F(r, \phi)) = (\cos\phi, \sin\phi)^t, \quad \text{and} \quad DG_2(F(r, \phi)) = r^{-1}(-\sin\phi, \cos\phi)^t.$$

It follows that

$$\left(\frac{\partial r}{\partial x}, \frac{\partial r}{\partial y}\right) = \|(x, y)\|^{-1}(x, y), \qquad \left(\frac{\partial \phi}{\partial x}, \frac{\partial \phi}{\partial y}\right) = \|(x, y)\|^{-2}(-y, x), \quad ((x, y) \in V).$$

The following shows that U is *maximal* in the sense that we cannot take a larger open over which F is still a diffeomorphism.

Exercise 6.19. Prove that there is no open U', strictly containing $U = \{\phi \in (0, 2\pi)\}$, such that $F|_{U'}$ is a C^1-diffeomorphism with its image. **Hint:** Show that $F|_{U'}$ would not be injective. △

In fact, you can also show that:

Exercise 6.20. Prove that it is impossible to choose the open subset $U' \subset (0, \infty) \times \mathbb{R}$ such that $F|_{U'}$ is a diffeomorphism with image $\mathbb{R}^2 \setminus \{0\}$. △

In our choice of U above we decided to miss, via F, the points with angle 0. Of course, we could have made some other choice. Consider the unit circle $S := \{(x, y) \mid x^2 + y^2 = 1\} \subset \mathbb{R}^2$ and pick an angle $\sigma \in S$. Then there exists $\phi_0 \in \mathbb{R}$ such that $\sigma = (\cos\phi_0, \sin\phi_0)$. Define the half-line $L(\sigma) := \{r\sigma \mid r \geq 0\}$ in $\mathbb{R}^2$. Define the open neighborhood

$$U(\phi_0) := (0, \infty) \times (\phi_0, \phi_0 + 2\pi).$$

Then it is easy to verify that F maps the set $U(\phi_0)$ bijectively onto $\mathbb{R}^2 \setminus L(\phi_0)$, thus defining a C^1 diffeomorphism. This $U(\phi_0)$ is also maximal.

Exercise 6.21. Prove that there are other opens $W \subset (0, \infty) \times \mathbb{R}$ with $F|_W$ a C^1-diffeomorphism with its image, that are also maximal in the sense above. **Hint:** Consider a curve of the form $(r, \phi(r))$. △

6.2.2 Cylindrical coordinates

We consider the mapping $C : (0, \infty) \times \mathbb{R}^2 \to \mathbb{R}^3$ defined by

$$C(r, \phi, z) := (r\cos\phi, r\sin\phi, z) = (F(r, \phi), z),$$

with F as in the previous example. Using a similar reasoning as above, C is a diffeomorphism from the set $U(\phi_0) \times \mathbb{R}$ to $(\mathbb{R}^2 \setminus L(\phi_0)) \times \mathbb{R} = \mathbb{R}^3 \setminus L(\phi_0) \times \mathbb{R}$, the complement of a closed half-plane in $\mathbb{R}^3$.

6.2.3 Spherical coordinates

Every point $(x, y, z) \in \mathbb{R}^3 \setminus \{0\}$ can be uniquely written as $r \cdot \sigma$ with σ a point on the unit sphere S in $\mathbb{R}^3$ and $r > 0$. Note that

$$r = \|(x, y, z)\| = \sqrt{x^2 + y^2 + z^2}.$$

If $\sigma = (x, y, z)$ is a point on the unit sphere, and $\rho = \sqrt{x^2 + y^2}$, then (ρ, z) is a point on the unit circle in $\mathbb{R}^2$, with a nonnegative first coordinate. It follows that $(\rho, z) = (\cos\theta, \sin\theta)$ for a unique $\theta \in [-\pi/2, \pi/2]$. Furthermore, $\rho^{-1} \cdot (x, y)$ is a point on the unit circle in $\mathbb{R}^2$, so there exists a unique $-\pi < \phi \leq \pi$ such that $(x, y) = \rho \cdot (\cos\phi, \sin\phi)$. We conclude that

$$\sigma = (\cos\phi\cos\theta, \sin\phi\cos\theta, \sin\theta),$$

for a unique point $(\phi, \theta) \in \mathbb{R}^2$ with $-\pi < \phi \leq \pi$ and $-\pi/2 \leq \theta \leq \pi/2$.

Motivated by the above, we consider the mapping $\Phi : \mathbb{R}^3 \to \mathbb{R}^3$ defined by

$$\Phi(r, \phi, \theta) := r(\cos\phi\cos\theta, \sin\phi\cos\theta, \sin\theta).$$

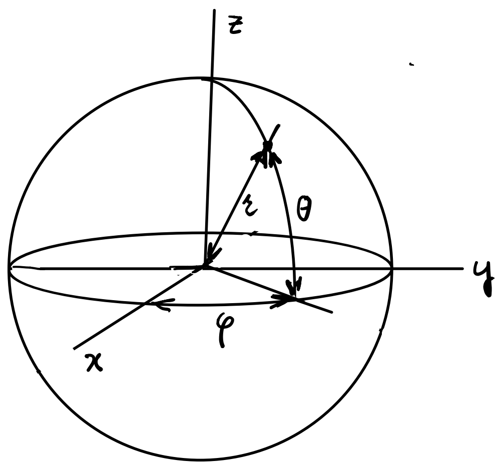
This is a C^1 mapping whose Jacobian is given by

$$\det D\Phi(r, \phi, \theta) = r^2 \cos\theta.$$

(Check this!) The mapping Φ takes $[0, \infty) \times [-\pi, \pi] \times [-\pi/2, \pi/2]$ surjectively onto $\mathbb{R}^3$ and is injective on the open set

$$U := (0, \infty) \times (-\pi, \pi) \times (-\pi/2, \pi/2),$$

on which the Jacobian is also nonzero. By the global inverse function Theorem 6.12, the image $V := \Phi(U)$ is open in $\mathbb{R}^3$ and $\Phi : U \to V$ is a C^1 diffeomorphism. The inverse $\Psi : V \to U$ of Φ is a C^1 mapping. The components of Ψ are also called the spherical coordinates on V. The traditional notation for these components is (r, ϕ, θ). Note that $\mathbb{R}^3 \setminus V$ equals the half-plane consisting of points $(x, y, z) \in \mathbb{R}^3$ with $y = 0$ and $x \leq 0$.



Figuur 8: Spherical coordinates

6.2.4 Another example

Let $f : \mathbb{R}^2 \to \mathbb{R}^2$ be the function $f(x,y) := (x(1-x), y^2)$. Given $p = (1,1) \in \mathbb{R}^2$, we ask ourselves to determine the largest open $U \subset \mathbb{R}^2$ satisfying:

- U contains p,
- U is path-connected,
- $f|_U : U \to f(U)$ is a C^1-diffeomorphism.

First, let us compute

$$Df(x,y) = \begin{pmatrix} 1-2x & 0 \\ 0 & 2y \end{pmatrix},$$

which has determinant $\det Df(x,y) = (1-2x)2y$. It vanishes on the subset $\Sigma := \{x = 1/2\} \cup \{y = 0\}$. According to Proposition 6.6, the U we are interested in must be disjoint from Σ.

Observe now that $\mathbb{R}^2 \setminus \Sigma$ is partitioned into four pieces $U_{\pm,\pm} := \{\pm x > \pm 1/2, \pm y > 0\}$. The point p belongs to $U_{+,+}$.

We claim that f is injective over each $U_{\pm,\pm}$. Indeed, both components $f = (f_1, f_2)$ are functions of a single variable that have non-zero derivative in the complement of Σ, so both are monotone and thus injective. We conclude, using the global inverse function Theorem 6.12, that F restricted to each $U_{\pm,\pm}$ is a diffeomorphism with its image. You can verify that this image is the same for the four opens, namely $\{x < 1/4, y > 0\}$.

We also claim that $U_{\pm,\pm}$ is convex and thus path-connected (Lemma 1.22). Convexity can be seen from the fact hat each $U_{\pm,\pm}$ is a product of two intervals, each of which is convex. Alternatively (and this is the more general way to handle this) you can also check that $ta + (1-t)b$ satisfies the conditions $\pm x > \pm 1/2$ and $\pm y > 0$ if a and b do, for each $t \in [0,1]$.

We claim that the desired open is $U_{+,+}$. It remains to show maximality, i.e. that we cannot find a larger open with these properties. Indeed, suppose a larger open U would exist. Then it would some additional point q. This point cannot be in Σ, according to Proposition 6.6. It cannot be in one of the other $U_{\pm,\pm}$ either due to injectity: indeed, all of them have the same image, namely $\{x < 1/4, y > 0\}$.

6.3 Submanifolds and regular values

Let $U \subset \mathbb{R}^n$ be open. A C^1 function $g = (g_1, \dots, g_k) : U \to \mathbb{R}^k$ determines a subset

$$N := g^{-1}(\{0\}) = \{x \in U \mid g_1(x) = g_2(x) = \dots = g_k(x) = 0\}.$$

Since g is continuous and $\{0\}$ is a closed subset of $\mathbb{R}^k$, N is a closed subset of U (Lemma 1.7).

Our goal in Subsection 6.4 will be to determine the extrema of $f|_N$, where $f : U \to \mathbb{R}$ is a C^1 function. This application of the inverse function theorem is known as the *Lagrange multipliers method*, a widely used tool to determine extrema under such boundary conditions.

However, it is difficult to determine the extrema of $f|_N$, unless N is a "nice" subset. The prototypical example of a nice subset would be a vector subspace $N \subset \mathbb{R}^n$ of dimension $n-k$. Let us review how this concrete case follows into the setup described above. According to Subsection 2.4, the annihilator $N^\perp \subset (\mathbb{R}^n)^*$ is a subspace of the dual of dimension k. By taking a basis we then obtain k covectors $\alpha_1, \cdots, \alpha_k$ spanning $N^\perp$. These covectors can be seen as a system of k linearly independent equations whose solution set is N. Identically, these k covectors form a linear function $g : \mathbb{R}^n \to \mathbb{R}^k$ whose kernel is $N = g^{-1}(0)$.

Observe that in this concrete case it is easy to understand $f|_N$. Since N is a vector subspace, we can identify it with $\mathbb{R}^{n-k}$ and thus use the usual theory of Subsections 3.4.1 and 4.5.

We now generalise these ideas. We need to introduce some assumptions on g to guarantee N being nice.

Definition 6.22. Suppose $g : U \to \mathbb{R}^k$ is C^1. A point $x \in U$ is said to be **critical** or **singular** if $Dg(x) : \mathbb{R}^n \to \mathbb{R}^k$ is not surjective. △

Recall (Subsection 2.2) that $Dg(x)$ is surjective if and only if its rows are all linearly independent. This is also equivalent to finding a k-by-k minor with non-zero determinant.

Exercise 6.23. Verify that, for a function $g : U \to \mathbb{R}$, this is equivalent to the standard Definition 3.32. △

Definition 6.24. Suppose $g : U \to \mathbb{R}^k$ is C^1. Fix a point $y \in \mathbb{R}^k$. The level set $N = g^{-1}(y) \subset U$ is said to the **regular** if it contains no critical points. Otherwise it is **singular**. We say that N has **dimension** $n - k$.

The **tangent space** of N at $x \in N$ is the kernel $T_xN := (Dg(x))^{-1}(0)$. △

The intuition for this definition is the following: $Dg(x)$ being surjective means that it defines a linearly independent system of linear equations, so $(Dg(x))^{-1}(0)$ is a vector space of dimension $n-k$. We can think of g as a system of equations, but these equations need not be linear anymore. Nonetheless, at the point x they are best approximated by the linear system $Dg(x)$. As such, the solution set $(Dg(x))^{-1}(0)$ is the best linear approximation at $x \in N$ to the solution set $N = g^{-1}(y)$.

Example 6.25. Consider $g : \mathbb{R}^2 \to \mathbb{R}$ given by $(x, y) \mapsto x^2 + y^2$. The differential $Dg(x, y) = (2x2y)$ vanishes only at the origin, so all level sets, except 0, are regular. In particular, $N = g^{-1}(1)$, the unit circle, is a regular level set. Observe furthermore that, at a point $(x, y) \in N$, the tangent space $T_{(x,y)}N = \ker(Dg(x, y))$ consists of the vectors orthogonal to (x, y), as expected. N has dimension 1. △

You can similarly verify that the unit sphere in $\mathbb{R}^n$ is a regular level set of dimension $n - 1$.

Example 6.26. If $N = g^{-1}(y)$ is regular then $U \cap N$ is also a regular level set, since you can simply restrict $g|_U$. The dimension and the tangent spaces remain the same. △

One can consider the following more general notion:

Definition 6.27. A subset $N \subset \mathbb{R}^n$ is a **submanifold** if for each point $p \in N$ you can find an open U such that $U \cap N$ is a regular level set (for some function). N has dimension $n - k$ if each $U \cap N$ does. △

The **tangent space** T_pN of a submanifold N at a point $p \in N$ is then defined by taking a function g representing N locally as a regular level set and applying Definition 6.24. One must that this is independent of the chosen g and thus well-defined. This will be one of the main corollaries (Corollary 6.38) of the theory of Lagrange multipliers.

Lemma 6.28. *$N \subset \mathbb{R}^n$ is a submanifold if and only if for each point $p \in N$ you can find an open U and a function $g : U \to \mathbb{R}^k$ with $N \cap U = g^{-1}(y)$ such that $Dg(p)$ surjective.*

Proof. $Dg(p)$ being surjective means that it has a k-by-k minor with non-zero determinant. Since Dg is continuous and the determinant is a continuous function, the same is true for all x in a sufficiently small ball $B(p, \varepsilon)$ around p. It follows that $N \cap B(p, \varepsilon)$ is a regular level set. □

The theory of (sub)manifolds is developed in more detail in the courses *Analyse in meer variabelen* and *Differentieerbare varieteiten*.

6.3.1 Examples

Example 6.29. In $\mathbb{R}^3$, let us consider the functions $f(x, y, z) := x^2+y^2-z$ and $g(x, y, z) := z - 1$, as well as the function $(f, g) : \mathbb{R}^3 \to \mathbb{R}^2$ that they form together.

We ask ourselves whether $A := f^{-1}(0)$, $B := g^{-1}(0)$ and $C := (f, g)^{-1}(0) = A \cap B$ are submanifolds of $\mathbb{R}^3$. We claim that the answer is yes, because all of them are regular level sets.

We compute $Df(x, y, z) = (2x \quad 2y \quad -1)$, $Dg(x, y, z) = (0 \quad 0 \quad 1)$, and thus:

$$D(f, g)(x, y, z) = \begin{pmatrix} 2x & 2y & -1 \\ 0 & 0 & 1 \end{pmatrix}.$$

We observe that Df and Dg are never zero. This means that all level sets of f and g are regular, including A and B. We deduce that A and B are submanifolds of dimension $2 = 3 - 1$.

Next, we check whether (f, g) has critical points. This amounts to finding those (x, y, z) such that the matrix $D(f, g)(x, y, z)$ has rank 1 (note that it can never have rank zero since both rows are non-vanishing). Having rank 1 means that the two rows are proportional to each other. I.e. there is some λ such that $Df = \lambda Dg$. Looking at the third coefficient we see that necessarily $\lambda = -1$. This implies that $x = y = 0$. I.e. we find that the singular points of (f, g) ocurr along the subset

$$\Sigma := \{x = y = 0\} = \{(0, 0, z)\},$$

i.e. the z-axis. This means that a level set of (f, g) is regular if and only if it is disjoint from Σ. We see that: $(f, g)(0, 0, z) = (-z, z - 1) \neq 0$. I.e. Σ and C are indeed disjoint, so C is a regular level set and thus a submanifold of dimension $1 = 3 - 2$.

For completeness, do note that A is the graph of $z(x, y) = x^2 + y^2$, a paraboloid. Graphs are always submanifolds. Similarly, B is the graph of $z(x, y) = 1$, i.e. it is the horizontal plane at height $z = 1$. Their intersection C is therefore the unit circle at height $z = 1$. △

Example 6.30. In $\mathbb{R}^3$, consider the function $f(x, y, z) := x^2 + y^2 - z^2$. We ask ourselves whether $A := f^{-1}(0)$ is a regular level set and thus a submanifold. We compute $Df(x, y, z) = (2x \quad 2y \quad -2z)$. This means that f has a single critical point, namely $(0, 0, 0)$. This point is precisely in the level set A, so A is not a regular level set.

Nonetheless, A is a submanifold almost everywhere. Namely, consider the open $U := \mathbb{R}^3 \backslash \{0\}$. Then $A \cap U = (f|_U)^{-1}(0)$ is a regular level set and thus a submanifold of dimension 2.

Drawing A shows exactly this: A is a cone and $(0,0,0)$ is precisely the cone point in which the two sheets of the cone come together. At that point A does not resemble a linear subspace so it is visibly not a submanifold at that point (see the upcoming Theorem 6.34). △

The procedure described in the previous exercise is general:

Exercise 6.31. Let $f : U \to \mathbb{R}^k$ be a C^1 function and let $A = f^{-1}(y)$ be a level set, not necessarily regular. Suppose $\Sigma \subset U$ is the set of critical points of f. Show that Σ is closed. Deduce that $A \cap V$ is a regular level subset of the open $V := U \setminus \Sigma$. △

6.3.2 Submanifolds and change of coordinates

In Subsection 6.1.7 we saw that C^1-diffeomorphisms preserve the C^1 information of functions. The same is true for submanifolds:

Lemma 6.32. *Let $W \subset \mathbb{R}^n$ be an open containing a submanifold N. Given a C^1 diffeomorphism $f : W \to f(W)$, the image $f(N)$ is also a submanifold.*

Proof. We must show that $f(N)$ is a regular level set close to each $x \in f(N)$. Since N is a submanifold we can consider $a = f^{-1}(x) \in N$ and find an open U containing a such that $N \cap U = g^{-1}(y)$ is a regular level set, with $g : U \to \mathbb{R}^k$ some C^1 function.

Since f is a C^1-diffeomorphism we have that $f(U)$ is an open containing x. We also deduce that $g \circ f^{-1}$ is C^1. We see that $f(N \cap U) = g \circ f^{-1}(y)$, but we have to show that it is a regular level set. Using the chain rule (Theorem 3.48) we see that $D(g \circ f^{-1})(x) = Dg(a) \circ D(f^{-1})(x)$ is surjective because $Dg(a)$ is surjective and $D(f^{-1})(x)$ is invertible. This proves the claim. □

We think of the tangent space T_aN as the best linear, first order approximation of N at a. As such, it is C^1 information, so it should be transformed nicely under a C^1-diffeomorphism. This is indeed the case, since: $T_aN = \ker(Dg(a))$ and $T_x(f(N)) = \ker(Dg(a) \circ D(f^{-1})(x))$ are related by the linear isomorphism $D(f)(a)|_{T_aN} : T_aN \to T_x(f(N))$.

Lemma 6.33. *Suppose $N = g^{-1}(y)$ is a regular level set of $g : U \to \mathbb{R}^k$. Suppose $h : \mathbb{R}^k \to \mathbb{R}^k$ is a diffeomorphism. Then N is also the regular level set $(h \circ g)^{-1}(h(y))$.*

Proof. $N = (h \circ g)^{-1}(h(y))$ follows from the fact that h is a bijection. That it is regular follows once more from the chain rule (Theorem 3.48), since $D(h \circ g)(x) = Dh(g(x)) \circ Dg(x)$ and $Dg(x)$ is surjective and $Dh(g(x))$ is bijective. □

Observe also that $\ker(Dg(x)) = \ker(Dh(g(x)) \circ Dg(x))$, since h is an isomorphism. This was to be expected, since both describe the tangent space T_xN.

6.3.3 The regular value theorem

We now prove our main result about submanifolds, which says that, in suitable local coordinates, every submanifold of dimension $n-k$ is equivalent to $\mathbb{R}^{n-k}$. Remember the classic example: the unit sphere in $\mathbb{R}^3$ is a two-dimensional manifold and, as such, locally, it resembles the plane $\mathbb{R}^2$. However, globally they are very different. I.e. the Earth may seem flat, if we look close to us, but this is not true globally!

Theorem 6.34. *Suppose $N \subset \mathbb{R}^n$ is a submanifold of dimension $n-k$. Fix a point $a \in N$. Then, there are opens $U \ni a$ and V in $\mathbb{R}^n$ and a diffeomorphism $\phi : V \to U$ such that*

- $\phi(\{0\} \times \mathbb{R}^{n-k} \cap V) = N \cap U$.
- $\phi(0) = a$.
- *$D\phi(0)$ restricts to an isomorphism $\{0\} \times \mathbb{R}^{n-k} \to T_a N$.*

Proof. The goal is to construct ϕ via the inverse function theorem. Consider an open U' containing a such that $N \cap U'$ is the regular level set $g^{-1}(0)$ of a function $g : U' \to \mathbb{R}^k$. In particular, we have that $Dg(a) : \mathbb{R}^n \to \mathbb{R}^k$ is surjective. Our first goal is to use linear transformations to simplify how N and g look at a.

According to Lemma 2.16 we can find linear changes of basis $A : \mathbb{R}^n \to \mathbb{R}^n$ and $B : \mathbb{R}^k \to \mathbb{R}^k$ so that $B \circ Dg(a) \circ A$ has a initial k-by-k block that is the identity matrix and its last $n-k$ columns are zero. Observe that $B \circ g$ still has $N \cap U'$ as zero level set, since B is invertible. We can then consider the change of coodinates $\psi(x) = a + A(x) : \mathbb{R}^n \to \mathbb{R}^n$. Observe that it sends 0 to a. The function

$$h = B \circ g \circ \psi : \psi^{-1}(U') \to \mathbb{R}^k$$

satisfies $Dh(0) = B \circ Dg(a) \circ A$. We deduce (Lemmas 6.32 and 6.33) that $M = h^{-1}(0) = \psi^{-1}(N)$ is a regular level set with $T_0 M = \{0\} \times \mathbb{R}^{n-k}$. That is, we have simplified the situation to the case where the tangent space looks like $\{0\} \times \mathbb{R}^{n-k}$. Now we will change coordinates further to make M look like it.

Consider the function $H : \psi^{-1}(U') \to \mathbb{R}^n$ given by the formula

$$H(x) = (h(x), \langle e_{k+1}, x\rangle, \cdots, \langle e_n, x\rangle).$$

The logic is that the first k outputs are the function h, whose zero level set is M. The last $(n-k)$ outputs are the usual coordinates in $\{0\} \times \mathbb{R}^{n-k}$. That is, we are trying to use the coordinates in $T_0 M$ to produce coordinates in M itself. By construction, $DH(0)$ is the identity so the local inverse function theorem applies (Theorem 6.7), telling us that $H : W \subset \psi^{-1}(U') \to V = H(W)$ is a diffeomorphism if we restrict to a small neighbourhood W of 0.

Now we check. By construction, $H(x) \in \mathbb{R}^{n-k}$ if and only if $h(x) = 0$ if and only if $x \in M$. We deduce that $H(M \cap W) = \mathbb{R}^{n-k} \cap V$. As such, the proof is complete if we set $U = \psi(W)$ and $\phi = \psi \circ H^{-1}$. □

6.4 Lagrange multipliers

We can now define critical points for the restriction of a function to a submanifold. It resembles the definition in the usual case:

Definition 6.35. Let $U \subset \mathbb{R}^n$ be an open, $N \subset U$ a submanifold, and $f : U \to \mathbb{R}$ a C^1-function. A point $a \in N$ is *stationary* or *critical* for $f|_N$ if $Df(a)|_{T_a N} : T_a N \to \mathbb{R}$ is zero. △

We can use linear algebra to reformulate Definition 6.35 in various helpful ways:

Lemma 6.36. *Consider an open $U \subset \mathbb{R}^n$, a regular level set $N = g^{-1}(y) \subset U$, and a C^1-function $f : U \to \mathbb{R}$. The following conditions are all equivalent:*

(a) *a is a critical point of $f|_N$.*

(b) *There is a unique collection of* ***Lagrange multipliers*** *$\lambda_1, \ldots, \lambda_k \in \mathbb{R}$ such that:*

$$Df(a) = \lambda_1 Dg_1(a) + \cdots + \lambda_k Dg_k(a). \tag{6.2}$$

(c)

$$\ker Df(a) \supset \ker Dg(a). \tag{6.3}$$

Proof. Recall Proposition 2.27. The point a is stationary if and only if $Df(a)|_{T_aN} = 0$, meaning that $Df(a)$ belongs to the annihilator of T_aN. Note that the annihilator is spanned by the rows $\{Dg_j(a)\}$ by construction. Condition (b) means that $Df(a)$ belongs to the linear span of the $\{Dg_j(a)\}$. Condition (c) means that f vanishes on the annihilator of $Dg(a)$, which is T_aN. All conditions are thus equivalent. □

The main result of this subsection reads:

Theorem 6.37. *Consider an open $U \subset \mathbb{R}^n$, a C^1-function $g : U \to \mathbb{R}^k$ with regular level set $N = g^{-1}(y)$, and a C^1-function $f : U \to \mathbb{R}$. Suppose that the restriction $f|_N : N \to \mathbb{R}$ has a (local) maximum or minimum $a \in N$. Then a is a critical point of $f|_N$.*

Proof. We will prove it in two steps.

Step I: The simplified case of $\mathbb{R}^{n-k}$. Consider first the following simplified setting: Suppose g is given by the coordinate functions $g_j(x) = x_j : U \to \mathbb{R}$ with $1 \le j \le k$, so $N = U \cap (\{0\} \times \mathbb{R}^{n-k})$. Suppose moreover that $f|_N$ has a local maximum/minimum at the origin.

These assumptions imply that we can write $N = \{0\} \times U'$, with U' a neighbourhood of zero 0 in $\mathbb{R}^{n-k}$. Furthermore, we can restrict f to $F : U' \to \mathbb{R}$ by setting $x' \mapsto f(0, x')$. This is a C^1-function, defined over an Euclidean open, that also has a local maximum/minimum at the origin. According to Lemma 3.33 it then follows that $DF(0) = 0$ so $D_j f(0) = 0$ voor all $j > k$.

This tells us that the annihilator of $\{0\} \times \mathbb{R}^{n-k}$ is spanned by the $\{Dg_j(0)\}$ and that $Df(0)$ belongs to it. It follows that the latter is a linear combination of the former, so Equation (6.2) follows and thus the claim (Lemma 6.36).

Step II: End of the proof by changing coordinates. We now address the general case. We let N, g, and f be arbitrary, as in the statement of the theorem.

Given a point $a \in N$ we can find a diffeomorphism $\phi : V \to U$ such that:

- $\phi(\{0\} \times \mathbb{R}^{n-k} \cap V) = N \cap U$,
- $\phi(0) = a$,
- $D\phi(0)$ is a linear isomorphism between the tangent spaces $\{0\} \times \mathbb{R}^{n-k}$ and $\ker(Dg(a))$.

Suppose that $a \in N$ is a local extremum of $f|_N$. Then 0 is a local extremum of $f \circ \phi|_{\{0\} \times \mathbb{R}^{n-k} \cap V}$ and Step I applies, showing that $D(f \circ \phi)(0) = Df(a) \circ D\phi(0)$ is in the annihilator of $\{0\} \times \mathbb{R}^{n-k}$. Since

$D\phi(0)$ is a linear isomorphism, this is equivalent to $Df(a)$ being in the annihilator of $D\phi(0)(\{0\} \times \mathbb{R}^{n-k}) = \ker(Dg(a))$. According to Lemma 6.36 this concludes the proof. □

An important consequence is that the tangent space T_xN of a submanifold N at x is well-defined, since it does not depend on the concrete function used to present it as a regular level set:

Corollary 6.38. *Suppose $f, g : U \subset \mathbb{R}^n \to \mathbb{R}^k$ are C^1-functions and $N = f^{-1}(0) = g^{-1}(0)$ is a regular level set for both. Then $\ker(Df(x)) = \ker(Dg(x))$ for every $x \in N$.*

Proof. Recall Subsection 2.4.3. Fix a point $x \in N$. Since N is a regular level set, $\ker(Df(x)) \subset \mathbb{R}^n$ is a $(n-k)$-dimensional subspace and its annihilator is the k-dimensional subspace $\ker(Df(x))^\perp \subset (\mathbb{R}^n)^*$ spanned by the covectors $\{Df_j(x)\}$. The same is true for g: the $\{Dg_i(x)\}$ span the k-dimensional annihilator $\ker(Dg(x))^\perp$ of $\ker(Dg(x))$.

Now, since $g|_N = 0$, all the points in N are critical for the functions $\{g_i\}$. According to Theorem 6.37, this means that each $Dg_i(x)$ belongs to the span of the $\{Df_j(x)\}$. I.e. $\ker(Dg(x))^\perp \subset \ker(Df(x))^\perp$. Since both have dimension k, equality holds. The claim then follows by taking annihilators (Proposition 2.27). □

We wrap up this chapter with an application:

Example 6.39. Consider the circle $C \subset \mathbb{R}^3$ consisting of the points (x, y, z) with $z = 0$ en $x^2 + y^2 = 1$. We first show that this is a submanifold of dimension 1. Consider indeed the function $g : \mathbb{R}^3 \to \mathbb{R}^2$ with entries $g_1, g_2 \in C^1(\mathbb{R}^3)$ defined as

$$g_1(x, y, z) = z, \qquad g_2(x, y, z) = x^2 + y^2 - 1.$$

We observe that $C = g^{-1}(0) = \{(x, y, x) \in \mathbb{R}^3 \mid g_1(x, y, z) = g_2(x, y, z) = 0\}$ is a regular level set, since the pair

$$Dg_1(x, y, z) = (0 \quad 0 \quad 1), \qquad Dg_2(x, y, z) = (2x \quad 2y \quad 0)$$

is linearly independent in the locus $\{x, y \neq 0\}$ (i.e. away from the z-axis).

We now consider the following problem: fix the point $p = (1, -2, 3)$ and compute which points from C are closest to p. This means that we are interested in the global minima of $f|_C$, where $f \in C^1(\mathbb{R}^3)$ is the distance function

$$f(x, y, z) = \|(x, y, z) - p\|^2 = (x-1)^2 + (y+2)^2 + (z-3)^2.$$

Since C is closed and bounded, we know that $f|_C$ must have at least one global maximum and one global minimum. To find them we use the Lagrange multipliers method.

We compute

$$Df(x, y, z) = (2(x-1) \quad 2(y+2) \quad 2(z-3)).$$

According to Lemma 6.36 and Theorem 6.37, a point $(x, y, z) \in C$ is critical if and only if the Lagrange multiplier equation

$$Df(x, y, z) = \lambda Dg_1(x, y, z) + \mu Dg_2(x, y, z)$$

has a solution, for some $\lambda, \mu \in \mathbb{R}$. In our concrete case it reads

$$(2(x-1) \quad 2(y+2) \quad 2(z-3)) = \lambda(0 \quad 0 \quad 1) + \mu(2x \quad 2y \quad 0).$$

We can simplify this a bit; indeed, the constraint $(x, y, z) \in C$ means that $z = 0$ and $x^2 + y^2 = 1$, so the equation simplifies as:

$$(2(x-1) \quad 2(y+2) \quad -6) = \lambda(0 \quad 0 \quad 1) + \mu(2x \quad 2y \quad 0)$$

which implies that $\lambda = -6$. We then have the equations

$$(x-1) = \mu x, \qquad y + 2 = \mu y, \qquad x^2 + y^2 = 1,$$

which can be rewritten as

$$x((1-\mu) = 1, \qquad y(\mu - 1) = 2, \qquad x^2 + y^2 = 1,$$

and then as

$$x(1-\mu) = 1, \qquad y = -2x, \qquad x^2 + y^2 = 1.$$

Which gives the solutions:

$$(x', y', \mu') = (\frac{1}{5}\sqrt{5}, -\frac{2}{5}\sqrt{5}, 1 + \sqrt{5}), \qquad (x'', y'', \mu'') = (-\frac{1}{5}\sqrt{5}, \frac{2}{5}\sqrt{5}, 1 - \sqrt{5}).$$

It must be the case that one of them is the maximum and the other the minimum. We can compute and see that $f(x', y', 0) < f(x'', y'', 0)$ so $(x', y', 0)$ is the global minimum and $(x'', y'', 0)$ the global maximum. △

7 Line integrals

In Chapter 5 we studied the integration of multivariate functions $f : U \subset \mathbb{R}^n \to \mathbb{R}$ over n-dimensional hypercubes $C \subset U$. In Subsection 5.1 we remarked that there are other notions of integration: there are so-called *differential k-forms* that can be integrated over k-dimensional submanifolds of U.

This is a topic you will see fully worked out in upcoming courses. For now, we will look solely into the case $k = 1$. We recommend that you review the contents of Subsection 1.2.

7.1 Covector fields

Definition 7.1. Consider an open subset $U \subset \mathbb{R}^n$. A *covector field* on U is a function $\alpha : U \to \mathrm{Lin}(\mathbb{R}^n, \mathbb{R})$. △

That is, a covector field assigns a covector $\alpha(x) \in (\mathbb{R}^n)^*$ to each point $x \in U$. We can then speak of covector fields that are continuous or C^k. Covector fields are also known as *differential 1-forms*. We will see below that they are the objects that can be naturally integrated along curves.

The most important example is the following:

Definition 7.2. Suppose $f : U \subset \mathbb{R}^n \to \mathbb{R}$ is a C^k function. Then $Df : U \to \mathrm{Lin}(\mathbb{R}^n, \mathbb{R})$ is a C^{k-1} covector field. We say that:

- The function f is a **primitive** or **potential** of Df.
- The vector field Df is **exact**. △

This motivates us to ask the following question: *Does every continuous covector field have a potential?* In one variable the answer is yes, according to the fundamental theorem of calculus. However, the answer is *no* if $n \geq 2$, as the following observation shows:

Lemma 7.3. *Suppose $\alpha : U \to \mathrm{Lin}(\mathbb{R}^n, \mathbb{R})$ is a C^1 covector field with C^2 potential $f : U \to \mathbb{R}$. Then it holds, for each i and j, that*

$$D_j\alpha_i = D_jD_if = D_iD_jf = D_i\alpha_j.$$

Identically, if you write $D\alpha : U \to \mathrm{Lin}(\mathbb{R}^n, \mathrm{Lin}(\mathbb{R}^n, \mathbb{R}))$ as a square matrix, it holds that it is a symmetric matrix.

Proof. The proof is already in the statement. Observe that $\alpha = Df$ being C^1 implies automatically that f is C^2. Then, according to Theorem 4.5, we can switch the order of differentiation when considering D_jD_if. The claim follows. □

This concept is so important that it deserves a name:

Definition 7.4. A C^1 covector field $\alpha : U \to \mathrm{Lin}(\mathbb{R}^n, \mathbb{R})$ is **closed** if $D_j\alpha_i = D_i\alpha_j$ for every i and j. △

The lemma above can then be stated as:

Proposition 7.5. *A covector field that is exact is also closed.*

7.1.1 Statement of the main results

This leads us to the main question of this chapter: *Is every closed covector field in U exact?* The answer is *yes* if U is simply-connected:

Theorem 7.6. *Let U be a simply-connected open. Then any closed covector field in U is exact.*

From this result we obtain a criterion to detect whether an open U is simply-connected:

Corollary 7.7. *Suppose U admits a closed but non-exact covector field. Then U is not simply-connected.*

Which we will use to show that:

Theorem 7.8. $\mathbb{R}^2 \setminus \{0\}$ *admits a closed but non-exact covector field. In particular, it is not simply-connected.*

These results are incredibly remarkable: they say that the shape of U is closely related to the analysis of functions and covector fields on U.

Proving these results will take the rest of the chapter. The path towards a proof is natural: being simply-connected is about loops in U. These interact with covector fields precisely because covector fields are the objects that naturally integrate over curves. In particular, we will:

- Define the *line integral* $\int_\gamma \alpha$ *of a covector field* α *along a path* γ (Subsection 7.2).
- See that the line integral only depends on the endpoints of γ if α is exact (Lemma 7.12). In particular, it is zero if γ is a loop.
- State and prove the main technical statement of the chapter, Theorem 7.19, which explains how ine integrals of closed covector fields behave under homotopy.

7.2 Line integrals

As claimed earlier:

Definition 7.9. Fix an open $U \subset \mathbb{R}^n$ and a continuous covector field $\alpha : U \to \mathrm{Lin}(\mathbb{R}^n, \mathbb{R})$. Suppose $\gamma : [a, b] \to U$ is a C^1 curve. The **line integral** of α along γ is defined by the expression:

$$\int_\gamma \alpha := \int_a^b \alpha(\gamma(t))(\gamma'(t))dt. \qquad \triangle$$

Recall that $\gamma'(t) \in \mathbb{R}^n$ is a vector, e.g. a column. $\alpha(\gamma(t))$ is a covector, e.g. a row. We can matrix multiply the two, yielding a number:

$$\alpha(\gamma(t))(\gamma'(t)) = \sum_{j=1}^{n} \alpha_j(\gamma(t))\gamma_j'(t).$$

You should interpret it as "the length of $\gamma'(t)$ according to $\alpha(\gamma(t))$". Equivalently, as a function of t, you can think of it as "the restriction of α to γ". This means that the integral $\int_\gamma \alpha$ should be thought as "the length of γ according to α". Refer to Corollary 7.25 for a bit more discussion.

Note moreover that the line integral is indeed well-defined. First, the velocity γ' is continuous because $\gamma \in C^1$. Secondly, according to Proposition 1.5, the function $t \mapsto \alpha(\gamma(t))(\gamma'(t))$ is continuous. It follows that it is Riemann integrable.

7.2.1 The line integral is intrinsic

The following result says that the line integral is invariant under reparametrisation (recall Definition 1.17):

Proposition 7.10. *Let α be a continuous covector field in U, $\gamma : [a, b] \to U$ a C^1 curve, and $\nu = \gamma \circ \rho : [c, d] \to U$ a reparametrisation that is C^1 (meaning that the change of coordinates $\rho : [c, d] \to [a, b]$ is a C^1-diffeomorphism). Then:*

$$\int_\nu \alpha = \int_\gamma \alpha.$$

Proof. According to the chain rule and using linearity we deduce that

$$\nu'(t) = D\gamma(\rho(t))(\rho'(t)) = \rho'(t)\gamma'(\rho(t)).$$

As such:

$$\begin{aligned} \int_\nu \alpha &= \int_c^d \alpha(\nu(t))(\nu'(t))dt \\ &= \int_c^d \alpha(\gamma(\rho(t)))(\rho'(t)\gamma'(\rho(t)))dt \\ &= \int_c^d \alpha(\gamma(\rho(t)))(\gamma'(\rho(t)))\rho'(t)dt \\ &= \int_a^b \alpha(\gamma(s))(\gamma'(s))ds = \int_\gamma \alpha. \end{aligned}$$

In the last step we used the substitution $s = \rho(t)$ which implies $ds = \rho'(t)dt$. □

This proof is *tautological*: we defined covector fields they way we did precisely so that they interact well with the substitution rule for the integral. Because of this, they define a notion of integration that is *intrinsic*, i.e. independent of how we parametrise curves.

7.2.2 The line integral in the exact case

First let us remark:

Lemma 7.11. *Suppose $U \subset \mathbb{R}^n$ is a path-connected open and $\alpha : U \to \mathrm{Lin}(\mathbb{R}^n, \mathbb{R})$ is an exact covector field with primitives f and g. Then $f = g + C$, with C a constant.*

Proof. By definition $Df = \alpha = Dg$ so $D(f - g) = 0$, meaning that $f - g$ is locally constant (Proposition 3.37) and thus constant, since U is path-connected. □

The following tells us that computing integrals of exact fields is very easy:

Lemma 7.12. *Suppose $U \subset \mathbb{R}^n$ is an open, $\alpha : U \to \mathrm{Lin}(\mathbb{R}^n, \mathbb{R})$ is an exact covector field with primitive $f \in C^1(U)$, and $\gamma : [a,b] \to U$ is a C^1 curve. Then:*

$$\int_\gamma \alpha = f(\gamma(b)) - f(\gamma(a)).$$

In particular, the integral does not depend on γ, only on its endpoints.

Proof. Observe that, according to the chain rule, the integrand satisfies:

$$Df(\gamma(t))(\gamma'(t)) = D(f \circ \gamma)(t) = (f \circ \gamma)'(t)$$

As such, we can apply the fundamental theorem of calculus to see that:

$$\int_\gamma \alpha = \int_a^b (f \circ \gamma)'(t)dt = f(\gamma(b)) - f(\gamma(a)).$$ □

In particular, since a loop has the same beginning and endpoints:

Corollary 7.13. *Suppose $\gamma : [0,1] \to U$ is a C^1-loop. Then $\int_\gamma Df = 0$.*

7.2.3 Piecewise curves

We argued above that γ being a C^1 curve was important in order for the line integral to be well-defined. In practice (and in some of the arguments below), one can relax this assumption as follows:

Definition 7.14. A continuous curve $\gamma : [a,b] \to U$ is said to be *piecewise* C^1 if there is a partition $a = a_0 < \ldots < a_N = b$ of the interval $[a,b]$ so that the restriction $\gamma|_{[a_{j-1},a_j]}$ to each subinterval is C^1. △

The integral of a piecewise curve is then defined as:

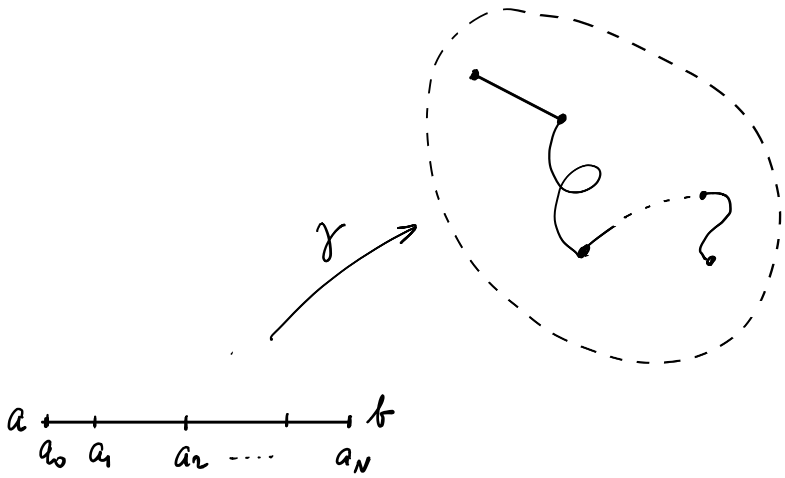
Definition 7.15. Suppose $\gamma : [a,b] \to U$ is a piecewise C^1 curve, with associated partition $a = a_0 < \ldots < a_N = b$. Then the line integral of a C^0 covector field $\alpha : U \to \mathrm{Lin}(\mathbb{R}^n, \mathbb{R})$ is defined by:

$$\int_\gamma \alpha := \sum_{j=1}^{n} \int_{\gamma|_{[a_{j-1},a_j]}} \alpha.$$ △

A couple of remarks are in order. First, observe that one can choose different partitions by subdividing $[a,b]$ into smaller intervals. However, $\int_\gamma \alpha$ does not depend on the choice of partition, due to the additivity of the integral. Secondly, this definition extends Definition 7.9 in the sense that every C^1 curve is in particular piecewise C^1 and both notions compute the same integral in that case.

7.2.4 Operations on curves

In Subsection 1.2.1 we discussed three operations on paths: concatenation, reversing, and reparametrisation. We have already seen how the latter behaves with the line integral (Proposition 7.10). We now discuss the other two.



Figuur 9: Stuksgewijze C^1 kromme.

Lemma 7.16. *Suppose* $\gamma, \nu : [0,1] \to U \subset \mathbb{R}^n$ *are* C^1 *curves with* $\gamma(1) = \nu(0)$. *Suppose* $\alpha : U \to \operatorname{Lin}(\mathbb{R}^n, \mathbb{R})$ *is a continuous covector field. Then the line integral of the concatenation satisfies:*

$$\int_{\nu\cdot\gamma} \alpha = \int_\nu \alpha + \int_\gamma \alpha.$$

Proof. First observe that $\int_{\nu\cdot\gamma} \alpha$ is well-defined because $\nu\cdot\gamma$ is piecewise C^1. Moreover, $(\nu\cdot\gamma)|_{[0,1/2]}$ is a reparametrisation of γ, so both have the same integral (Proposition 7.10). Similarly, $(\nu\cdot\gamma)|_{[1/2,1]}$ is a reparametrisation of ν. The claim follows by splitting the integral into two, as in Definition 7.15. □

Regarding the reverse curve:

Lemma 7.17. *Suppose* $\gamma : [0,1] \to U \subset \mathbb{R}^n$ *is a* C^1 *curve and* $\alpha : U \to \operatorname{Lin}(\mathbb{R}^n, \mathbb{R})$ *is a continuous covector field. Then:*

$$\int_{\bar\gamma} \alpha = -\int_\gamma \alpha.$$

Proof. We argue as in Proposition 7.10. Since $\bar\gamma'(t) = -\gamma'(1-t)$, we can apply the substitution $s = 1-t$ and thus $ds = -dt$:

$$\begin{aligned}
\int_{\bar\gamma} \alpha &= \int_0^1 \alpha(\gamma(1-t))(-\gamma'(1-t))dt \\
&= \int_1^0 \alpha(\gamma(s))(\gamma'(s))ds \\
&= -\int_0^1 \alpha(\gamma(s))(\gamma'(s))ds \\
&= -\int_\gamma \alpha,
\end{aligned}$$

where we used that reversing the direction of integration reverses the value of the integral. □

7.3 Proof of the main results

7.3.1 The Poincaré lemma

We first establish Theorem 7.6 in the particular case in which U is convex. This is by itself an important result, known as the **Poincaré lemma** for differential 1-forms. It is also the first ingredient towards a proof of the general case.

Theorem 7.18. *Let $U \subset \mathbb{R}^n$ be convex. Then every closed covector field in U is exact.*

Proof. Fix a closed covector field $\alpha : U \to \mathrm{Lin}(\mathbb{R}^n, \mathbb{R})$. Note that it is at least C^1, since it is closed. We want to construct a potential $f \in C^2(U, \mathbb{R})$. The crucial idea is that we can construct f via integration using Lemma 7.12. We will also use induction on n.

First, the base case is $n = 1$. In this situation we know that every covector field is exact, since we can simply integrate to yield a primitive.

For the inductive step we assume that the statement is true for all dimensions $m < n$. Fix then a constant C such that the hyperplane $\{x_n = C\}$ intersects U in a non-empty set; call it H. We can then consider the restriction $\alpha|_H : H \to \mathrm{Lin}(\mathbb{R}^{n-1}, \mathbb{R})$ given by $\alpha|_H = (\alpha_1 \quad \cdots \quad \alpha_{n-1})$. This is now a C^1 covector field in one less variable, and is still closed. Since H is convex, we deduce that $\alpha|_H$ is exact. Let us denote its potential by $g : H \to \mathbb{R}$.

We now define $f : U \to \mathbb{R}$, the primitive of α, using g. First, for each point $x \in H$ we set $f(x) := g(x)$. Given any other point $x \in U$ we can write it as $(y, x_n) = (x_1, \cdots, x_{n-1}, x_n)$ and consider the point $(y, C) \in H$ to which it projects. The two points are connected by the vertical curve $\gamma_y(t) := (y, t)$ with $t \in [C, x_n]$. According to Lemma 7.12, the desired equality $Df = \alpha$ forces that:

$$\begin{aligned} f(y, x_n) &= f(y, C) + \int_{\gamma_y} \alpha \\ &= g(y, C) + \int_C^{x_n} \alpha(y, t)(e_n)dt \\ &= g(y, C) + \int_C^{x_n} \alpha_n(y, t)dt, \end{aligned}$$

where we used that $\gamma_y'(t) = e_n$.

Lastly, we verify that this f is indeed a potential of α. Indeed, for each $i < n$:

$$D_i f(y, x_n) = D_i g(y, C) + \int_C^{x_n} (D_i\alpha_n)(y, t)dt = \alpha_i(y, C) + \int_C^{x_n} (D_n\alpha_i)(y, t)dt = \alpha_i(y, x_n).$$

In the first equality we used Theorem 5.21, the switching of integration and differentiation, using that α is C^1 and that we are integrating over a closed interval. In the middle inequality we used the closedness of α to switch the indices n and i. It remains to check the claim for the index $i = n$, which is immediate from the fundamental theorem of calculus:

$$D_n f(y, x_n) = D_n \int_C^{x_n} \alpha_n(y, t)dt = \alpha_n(y, x_n),$$

since the term $g(y, C)$ does not depend on the variable x_n. This finishes the proof. □

7.3.2 The main technical ingredient

Recall that all covector fields are closed and exact in one-variable. This means that, along a path, we can always find a potential. Similarly, a square is a convex subset, so Theorem 7.18 applies, allowing us to deduce that closed covector fields have a potential over the square. Since a homotopy is a map that has the square as domain, this reasoning suggests that:

Theorem 7.19. *Fix an open subset $U \subset \mathbb{R}^n$ and a closed covector field $\alpha : U \to \mathrm{Lin}(\mathbb{R}^n, \mathbb{R})$. Fix moreover curves $\gamma_0, \gamma_1 : [a, b] \to U$ and a homotopy $\Gamma : [a, b] \times [0, 1] \to U$ between the two. Suppose that one of the following two assumptions holds:*

(a) *Γ is a homotopy relative endpoints.*

(b) *Γ is a homotopy of loops.*

Then:

$$\int_{\gamma_0} \alpha = \int_{\gamma_1} \alpha. \tag{7.1}$$

This result is often informally called the "invariance of the line integral under homotopy", but you should keep in mind that the invariance only holds under the assumptions of the theorem.

An immediate corollary is the following:

Corollary 7.20. *Fix an open subset $U \subset \mathbb{R}^n$ and a C^1, closed covector field $\alpha : U \to \mathrm{Lin}(\mathbb{R}^n, \mathbb{R})$. Suppose γ is a nullhomotopic loop. Then $\int_\gamma \alpha = 0$.*

Proof. By assumption γ is homotopic as a loop to a constant loop η. Then Theorem 7.19 implies that $\int_\gamma \alpha = \int_\eta \alpha$, but the later is zero because $\eta'(t) = 0$ for all t. □

The proof of Theorem 7.19 amounts to formalising the heuristic arguments above. Something that should catch your attention is that the curves γ_i are just required to be continuous! Indeed, defining the line integral of α over a continuous curve will be a crucial part of the proof. We postpone this for now, since it will take the rest of the chapter (Section 7.4). Instead, let us see first how this result implies Theorem 7.6.

7.3.3 Proof of Theorem 7.6

Proof of Theorem 7.6. Fix a C^1 closed covector field $\alpha : U \to \mathrm{Lin}(\mathbb{R}^n, \mathbb{R})$, as well as a point $x \in U$. We now define a primitive $f : U \to \mathbb{R}$ of α. We do so as in Theorem 7.18. Given any other $y \in U$ we pick a path γ_y connecting x to y (which exists because U is path-connected) and we set:

$$f(y) := \int_{\gamma_y} \alpha.$$

Recall that this equality is forced by Lemma 7.12.

First we observe that f is in fact well-defined, i.e. it does not depend on the choice of γ_y. Indeed, given any other path ν connecting x to y, we can find a homotopy Γ between the two using the fact that U is simply-connected (Proposition 1.44). Theorem 7.19 then tells us that $\int_{\gamma_y} \alpha = \int_\nu \alpha$, since α is closed.

It remains to show that f is a potential for α. To see this, take a ball V containing y. Given any $z \in V$ we can consider a path η_z connecting y to z. Then the concatenation $\eta_z \cdot \gamma_y$ is a path from x to z, so it follows (Lemma 7.16) that:

$$f(z) = \int_{\gamma_z} \alpha = \int_{\gamma_y} \alpha + \int_{\eta_z} \alpha = f(y) + \int_{\eta_z} \alpha. \tag{7.2}$$

Since V is convex, Theorem 7.18 applies, producing a potential g_V for $\alpha|_V$. According to Lemma 7.12, this function must also satisfy (7.2). It follows that g_V and $f|_V$ differ by a constant, so the latter is also a potential. Since this is true for all y, the proof is complete. □

Theorems 7.19 and 7.6 together then imply that:

Corollary 7.21. *Suppose $U \subset \mathbb{R}^n$ is a simply-connected open and α is a closed covector field on U. Then:*

(a) *All loops $\gamma : [a, b] \to U$ satisfy:*

$$\int_\gamma \alpha = 0.$$

(b) *All pairs of curves $\gamma_1, \gamma_2 : [a, b] \to U$ with equal endpoints satisfy:*

$$\int_{\gamma_1} \alpha = \int_{\gamma_2} \alpha.$$

7.3.4 Proof of Theorem 7.8

Proof of Theorem 7.8. According to Theorem 7.6 we simply need to find a covector field α in $U = \mathbb{R}^2 \setminus 0$ that is closed but not exact. In light of Corollary 7.21 this amounts to finding a loop $\gamma : [a, b] \to \mathbb{R}^2 \setminus 0$ such that $\int_\gamma \alpha \neq 0$.

We consider the C^1 covector field

$$\alpha(x, y) := \frac{1}{x^2 + y^2}(-y \quad x)$$

on U. A direct computation yields

$$D_2\alpha_1(x, y) = -\frac{1}{x^2 + y^2} - \frac{2y^2}{(x^2 + y^2)^2} = \frac{y^2 - x^2}{(x^2 + y^2)^2}.$$

In a similar way one finds

$$D_1\alpha_2(x, y) = \frac{y^2 - x^2}{(x^2 + y^2)^2}.$$

Which shows that α is indeed closed.

Now we consider the closed curve $\gamma : [0, 2\pi] \to U$ defined by $\gamma(t) := (\cos t, \sin t)$. This is the usual parametrisation of the unit circle in $\mathbb{R}^2$ at unit speed. The curve γ is C^∞ with derivative $\gamma'(t) = (-\sin t, \cos t)$. Now we have $\alpha(\gamma(t)) = (-\sin t, \cos t)$ and $\alpha(\gamma(t))(\gamma'(t)) = \cos^2 t + \sin^2 t = 1$. As such, the line integral reads:

$$\int_\gamma \alpha = \int_0^{2\pi} 1 dt = 2\pi \neq 0.$$

It follows that U is not simply-connected (Corollary 7.21), γ is not nullhomotopic (Corollary 7.20), and α is not exact (Corollary 7.13). □

Exercise 7.22. Consider $U = \mathbb{R}^2 \setminus \{p, q\}$, where $p, q \in \mathbb{R}^2$ are some arbitrary points. Show that U is path-connected but it is not simply-connected. △

7.3.5 Higher euclidean spaces with punctures

In *Topologie en Meetkunde* you will learn the tools to prove that $U := \mathbb{R}^n \setminus \{0\}$ is simply-connected if $n \geq 3$. For now, just assume that this is the case.

Example 7.23. Consider the covector field $\alpha : U \to \mathbb{R}^n$ defined by

$$\alpha(x) := -\frac{x}{\|x\|^3}.$$

This covector field satisfies $\|\alpha(x)\| = \|x\|^{-2}$ and, for $n = 3$, is known from Newton's theory of gravitation. For $i \neq j$ we have

$$\begin{aligned} D_i\alpha_j(x) &= -\frac{\partial}{\partial x_i}\left(x_j(\|x\|^2)^{-3/2}\right) \\ &= -\frac{3}{2}x_j(\|x\|^2)^{-5/2}\frac{\partial}{\partial x_i}(\|x\|^2) \\ &= -3x_jx_i\|x\|^{-5}. \end{aligned}$$

It follows that α is closed. By Theorem 7.6, it follows that it has a potential $f : \mathbb{R}^n \setminus \{0\} \to \mathbb{R}$. △

One can take this a step further and compute a explicit potential:

Exercise 7.24. Let α be as in Example 7.23 and let $f : U = \mathbb{R}^n \setminus \{0\} \to \mathbb{R}$ be the (unique) potential satisfying $f(e_1) = 0$. Let S denote the unit sphere in $\mathbb{R}^n$.

(a) Show that $f(re_1) = \frac{1}{r} - 1$, for all $r > 0$.

(b) Show that for every orthogonal linear transformation $A : \mathbb{R}^n \to \mathbb{R}^n$ and every C^1 curve $\gamma : [0, 1] \to U$ we have

$$\int_\gamma \alpha = \int_{A\circ\gamma} \alpha.$$

(c) Show that for any pair of vectors $y, z \in S$ with $y, z \perp e_1$ there exists an orthogonal transformation $A : \mathbb{R}^n \to \mathbb{R}^n$ such that $Ae_1 = e_1$ and $Ay = z$.

(d) Show that for all $y, z \in S \cap e_1^\perp$ we have $f(y) = f(z)$.

(e) Show that for all $x \in S$ we have $f(x) = 1$.

(f) Show that $f(x) = \|x\|^{-1} - 1$ for $x \in U$.

(g) Show in two ways that $g : x \mapsto \|x\|^{-1}$, $U \to \mathbb{R}$ defines a potential of α. First, by using item (e). Secondly, by direct computation. △

7.4 Proof of Theorem 7.19

We dedicate the rest of the chapter to proving Theorem 7.19. The idea of the proof is the following: Given two paths γ_0 and γ_1, homotopic relative endpoints via a homotopy Γ, we "restrict the closed covector field α to Γ". Since the domain of Γ is a square C (and thus convex), Theorem 7.18 tells us that α is exact along Γ. It follows that the integral of α along the boundary of Γ is zero (Corollary 7.13). However, the boundary is the concatenation of γ_0, the constant curve with value $\gamma_0(1)$, the reverse $\bar{\gamma}_1$, and the constant curve with value $\gamma_0(0)$. As such, we can split the integral into four pieces (as in Definition 7.15), two of which are zero, the two others being $\int_{\gamma_0} \alpha$ and $-\int_{\gamma_1} \alpha$. The claim then follows.

There are a couple of issues with this argument. First, we have to understand what it means to restrict α to Γ. A second more difficult issue is that the theory developed in Subsection 1.2 deals with continuous paths and homotopies (since we were working in the setting of metric spaces). However, the theory of line integrals requires us to work in the C^1 setting. One can then proceed in two ways:

1. Repeat all the theory of Subsection 1.2 in C^1 regularity, and show that all key notions agree (e.g. that an open U is simply-connected if and only if every C^1 loop is nullhomotopic via a C^1 homotopy of loops).
2. Alternatively, extend the theory of line integrals of closed covector fields to the setting of continuous curves.

Both approaches are valid and contain interesting ideas. In these notes we will pursue approach (2), which is slightly less technical.

7.4.1 Line integral along continuous curves

First consider the following consequence of Lemma 7.12:

Corollary 7.25. *Suppose $U \subset \mathbb{R}^n$ is an open, $\alpha : U \to \mathrm{Lin}(\mathbb{R}^n, \mathbb{R})$ is an exact covector field with primitive $f \in C^1(U)$, and $\gamma : [a,b] \to U$ is a C^1 curve. Then:*

$$\int_a^t \alpha(\gamma(s))(\gamma'(s))ds = f(\gamma(t)) - f(\gamma(a)).$$

This result tells us that we should think of $\alpha(\gamma(s))(\gamma'(s))$ as the restriction of α to γ. It therefore becomes a covector field in one variable, so it can be integrated to find a primitive $t \mapsto \int_a^t \alpha(\gamma(s))(\gamma'(s))ds$ along the curve. If a potential f already exists in U, then the two agree along γ.

We can generalise this to continuous curves as long as α is closed:

Definition 7.26. Suppose $U \subset \mathbb{R}^n$ is an open subset and α is a closed covector field. Consider a continuous path $\gamma : [a,b] \to U$. A *primitive* of α along γ is a function $\varphi : [a,b] \to \mathbb{R}$ satisfying the following property: For every $t_0 \in [a,b]$ there is a potential f of α, defined over some ball contaning $\gamma(t_0)$, such that $\varphi(t) = f(\gamma(t))$ for every t in a sufficiently small neighbourhood of t_0 in $[a,b]$. △

Lemma 7.27. *Fix a closed covector field α on U. Let $\gamma : [a,b] \to U$ be a continuous curve. Then α has exactly one primitive φ along γ satisfying $\varphi(a) = 0$. Other primitives along γ are given by $\varphi + c$, with c a constant.*

Proof. We first consider the case $\alpha = 0$. Then the zero function is a potential along γ. Consider some other potential $\varphi : [a, b] \to \mathbb{R}$ along γ. Given $t_0 \in [a, b]$, there exists a ball neighbourhood B of $\gamma(t_0)$ and a primitive $f \in C^2(B)$ such that $f \circ \gamma = \varphi$ on an open neighbourhood of t_0. According to Lemma 7.11, f is constant on B. Hence φ is constant on an open neighbourhood of t_0 in $[a, b]$. Thus φ is locally constant on $[a, b]$ and by Lemma 1.29 we conclude that φ is constant.

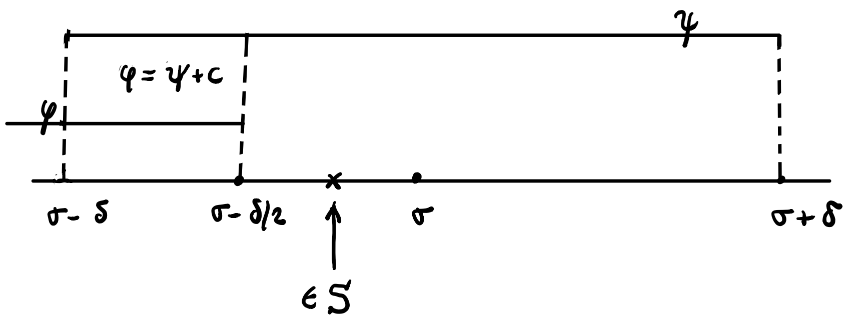
Consider now the general case in which α is a closed covector field on U. We first prove that the difference of any two primitives is constant. Suppose φ_1 and φ_2 are two primitives along γ. Then $\varphi := \varphi_1 - \varphi_2$ is a primitive of the zero covector field along γ. From the first part of the proof we know that it is constant.

Finally, we prove the existence of a primitive. We define S as the set of $s \in [a, b]$ such that α has a primitive along $\gamma|_{[a,s]}$. This set is bounded above by b, and contains a, so it is nonempty. By the least upper bound property, S has a supremum $\sigma := \sup S \leq b$. We will show that $\sigma = b$. Suppose not, then $\sigma < b$.

Choose an open ball neighbourhood B of $\gamma(\sigma)$ in U. By the continuity of γ at σ, there exists an open neighbourhood J of σ in $[a, b]$ such that $\gamma(J) \subset B$. By Theorem 7.18, α has a primitive f on B. We distinguish two cases, namely $\sigma = a$ and $\sigma > a$.

First suppose $\sigma = a$. There exists a $\delta > 0$ such that $[a, a + \delta] \subset J$. Then $\varphi : t \mapsto f(\gamma(t))$ is a primitive of v along $\gamma|_{[a,a+\delta]}$. Since $a + \delta > \sigma$ we get a contradiction.

Now suppose $a < \sigma < b$. Fix $\delta > 0$ such that $[\sigma - \delta, \sigma + \delta] \subset J$. The map $\psi : t \mapsto f(\gamma(t))$ is now a primitive of v along $\gamma|_{[\sigma-\delta,\sigma+\delta]}$. From $\sigma - \delta/2 < \sigma$ it follows that $[\sigma - \delta, \sigma] \subset S$, so α has a primitive φ along $\gamma|_{[a,\sigma-\delta/2]}$. See Figure 10.



Figuur 10: Continuation of primitive.

The restrictions of ψ and φ to $I := [\sigma - \delta, \sigma - \delta/2]$ are both primitives of α along $\gamma|_I$. From what we proved earlier, there exists a constant $c \in \mathbb{R}$ such that $\varphi = \psi + c$ on I. It follows that the function $\varphi : [a, \sigma - \delta/2] \to \mathbb{R}$ can be extended to a primitive on $[a, \sigma + \delta]$, by defining it as $\varphi := \psi + c$ on the interval $[\sigma - \delta/2, \sigma + \delta]$. It follows that $\sigma + \delta \in S$, a contradiction. We conclude that $\sigma = b$, which completes the proof. □

The characterization of all primitives in Lemma 7.27 makes the following definition possible.

Definition 7.28. Let $\gamma : [a, b] \to U$ be a continuous curve, and let $\alpha : U \to \mathbb{R}^n$ be a closed cvector field. We define the integral of α along γ by

$$\int_\gamma^* \alpha := \varphi(b) - \varphi(a),$$

with φ a primitive of α along γ. △

First, observe that the definition does not depend on the choice of primitive. Secondly, observe that this defines the integral only for closed covector fields. In general it is not possible to define the line integral of an arbitrary covector field along a continuous curve. Lastly, you should think of this definition as being auxiliary. Our goal in this subsection is to show (Lemma 7.32) that this extends the usual definition of line integral along a C^1-curve. Once we have shown that we can get rid of the star superscript. In order to show that this recovers the usual notion, we will show that it satisfies the same properties that the usual line integral does.

We first show that the star-integral is invariant under continuous reparametrizations of γ:

Lemma 7.29. *Let α be a closed covector field on U, $\gamma : [a, b] \to U$ a continuous curve, and ρ a continuous reparametrization of γ. Then:*

$$\int_\rho \alpha = \int_\gamma \alpha.$$

Proof. There exists a continuous change of coordinates $\tau : [c, d] \to [a, b]$ with $\tau(c) = a$ and $\tau(d) = b$ such that $\rho = \gamma \circ \tau$. Let φ be a primitive of α along γ. We will show that $\varphi \circ \tau$ is a primitive of α along ρ.

Let $s_0 \in [c, d]$ and write $t_0 := \tau(s_0) \in [a, b]$. There exists an open ball B in U centered at $\gamma(t_0) = \rho(s_0)$ and a primitive f of α on B such that

$$\varphi(t) = f(\gamma(t)), \qquad (t \in I),$$

on a neighbourhood I of $t_0 = \tau(s_0)$ in $[a, b]$. By continuity of τ it follows that there exists an open neighbourhood J of s_0 in $[c, d]$ such that $\tau(J) \subset I$. Now for all $s \in J$ we have $\tau(s) \in I$, hence

$$f(\rho(s)) = f(\gamma(\tau(s))) = \varphi(\tau(s)) = \varphi \circ \tau(s).$$

According to the definition, $\varphi \circ \tau$ is thus a primitive of α along ρ. Consequently,

$$\int_\rho \alpha = \varphi \circ \tau(d) - \varphi \circ \tau(c) = \varphi(b) - \varphi(a) = \int_\gamma \alpha. \qquad \square$$

The star-integral also behaves well under concatenation:

Lemma 7.30. *Let α be a closed covector field and $\gamma : [a, b] \to U$ a continuous curve. Let $c \in (a, b)$, and define $\gamma_1 := \gamma|_{[a,c]}$ and $\gamma_2 := \gamma|_{[c,b]}$. Then*

$$\int_\gamma^* \alpha = \int_{\gamma_1}^* \alpha + \int_{\gamma_2}^* \alpha.$$

Proof. Let φ be a potential of α along γ. Then the integral in the left-hand side equals $\varphi(b) - \varphi(a)$. Moreover, $\varphi|_{[a,c]}$ is a potential of α along γ_1 and $\varphi|_{[c,b]}$ is a potential of α along γ_2. Hence the sum of the integrals in the right-hand side equals

$$[\varphi(c) - \varphi(a)] + [\varphi(b) - \varphi(c)] = \varphi(b) - \varphi(a). \qquad \square$$

And also with respect to reversing:

Lemma 7.31. *Let α be a closed covector field and $\gamma : [a, b] \to U$ a continuous curve. If f is a primitive of α along γ then $\bar{f}(t) := f(1-t)$ is a primitive along the reverse $\bar{\gamma}$. As such:*

$$\int_{\bar{\gamma}}^{*} \alpha = -\int_{\gamma}^{*} \alpha.$$

We finally show that, for C^1 curves, the star-integral is just the usual integral:

Lemma 7.32. *Let $\gamma : [a, b] \to U$ be a C^1 curve, and α a closed covector field. Then*

$$\int_{\gamma}^{*} \alpha = \int_{\gamma} \alpha.$$

Proof. We first work under the assumption that $\gamma([a, b])$ is contained in an open ball $B \subset U$. Then, by Theorem 7.6, α has a primitive f on B. Moreover, $f \circ \gamma$ is a primitive of α along γ. Combining Definition 7.28 and Lemma 7.12 yields the result:

$$\int_{\gamma}^{*} \alpha = f(\gamma(b)) - f(\gamma(a)) = \int_{a}^{b} \alpha.$$

The idea in the general case is to split the curve into pieces, each of which has its image contained in a ball within U. This goes as follows. From compactness of $[a, b]$ and continuity of γ it follows that $\gamma([a, b])$ is compact in U. Then, according to Lemma 1.15, there exists an $\varepsilon > 0$ such that for every $t \in [a, b]$ we have $B(\gamma(t); \varepsilon) \subset U$.

Again by compactness and continuity (Propositions 1.10 and 1.13) it follows that $\gamma : [a, b] \to U$ is uniformly continuous. Hence there exists a $\delta > 0$ such that for all $t_1, t_2 \in [a, b]$ it holds that

$$|t_1 - t_2| < \delta \implies \|\gamma(t_1) - \gamma(t_2)\| < \varepsilon.$$

We can then choose a partition $a = a_0 < a_1 < \ldots < a_N = b$ of the interval $[a, b]$ such that $a_j - a_{j-1} < \delta$ for all $1 \leq j \leq N$. As such, for all $1 \leq j \leq N$ we have that $\gamma([a_{j-1}, a_j]) \subset B(\gamma(a_j); \varepsilon)$.

Write $\gamma_j := \gamma|_{[a_{j-1}, a_j]}$ for each $1 \leq j \leq N$. It follows from the first part of the proof that

$$\int_{\gamma_j}^{*} \alpha = \int_{\gamma_j} \alpha.$$

Summing over $j = 1, \ldots, N$ and repeatedly applying Lemma 7.30 yields the result:

$$\int_{\gamma}^{*} \alpha = \sum_{j=1}^{N} \int_{\gamma_j}^{*} \alpha = \sum_{j=1}^{N} \int_{\gamma_j} \alpha = \int_{\gamma} \alpha. \qquad \square$$

The exact same reasoning implies the slightly more general claim:

Corollary 7.33. *Suppose $\gamma : [a, b] \to U$ is piecewise C^1 and α is a closed covector field. Then:*

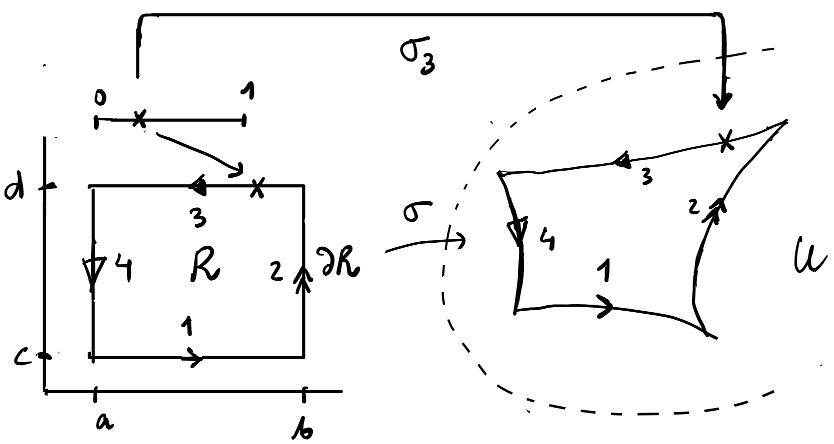
$$\int_{\gamma}^{*} \alpha = \int_{\gamma} \alpha.$$

7.4.2 Covector fields along homotopies

Our next step is to discuss homotopies of (continuous) curves and how they interact with closed covector fields. Recall that a homotopy of curves is parametrised by a rectangle. As such, let $R = [a,b] \times [c,d]$ be a rectangle in $\mathbb{R}^2$. We denote the boundary of R in $\mathbb{R}^2$ by ∂R. This boundary consists of four line segments. Suppose a continuous mapping $\sigma : \partial R \to U$ is given. This is not a homotopy, only its boundary. We introduce the following notation for the four continuous curves associated to the four boundary segments:

$$\sigma_1 : [0,1] \to U,\ t \mapsto \sigma(a + t(b-a), c), \quad \sigma_2 : [0,1] \to U,\ t \mapsto \sigma(b, c + t(d-c)),$$
$$\sigma_3 : [0,1] \to U,\ t \mapsto \sigma(b + t(a-b), d), \quad \sigma_4 : [0,1] \to U,\ t \mapsto \sigma(a, d + t(c-d)).$$

Observe that these segments are being transversed in a counterclockwise manner. See Figure 11.



Figuur 11: The four boundary curves σ_j defined by a homotopy σ.

If we now fix a closed covector field α on $U \subset \mathbb{R}^n$, we can consider its line integral along σ. It splits as the sum of the line integrals along the four aforementioned segments:

$$\int_\sigma \alpha := \sum_{j=1}^{4} \int_{\sigma_j} \alpha. \tag{7.3}$$

The following is the analogue, for continuous curves, of Corollary 7.13:

Lemma 7.34. *Let $\sigma : \partial R \to U$ be continuous and let α be a exact covector field on U. Then:*

$$\int_\sigma v(x) \cdot dx = 0.$$

Proof. Let f be a primitive of α. For every $1 \le j \le 4$ the function $f \circ \sigma_j$ is a primitive of α along σ_j. Therefore, the right-hand side of (7.3) equals

$$\sum_{j=1}^{4} [f(\sigma_j(1)) - f(\sigma_j(0))]. \tag{7.4}$$

From the fact that $\sigma_j(1) = \sigma_{j+1}(0)$ for $j = 1, 2, 3$, and $\sigma_4(1) = \sigma_1(0)$ it follows that the sum in (7.4) is equal to zero. □

We now introduce homotopies into the discussion. We will say that $\sigma : \partial R \to U$ **extends** to R if there is a homotopy $\Gamma : R \to U$ such that $\sigma = \Gamma|_{\partial R}$. Then:

Proposition 7.35. *Let $\sigma : \partial R \to U$ be continuous and α a closed covector field on U. If σ extends to R then*

$$\int_\sigma \alpha = 0. \tag{7.5}$$

Observe that this result pretty much establishes Theorem 7.19 already. The idea behind the proof is to partition R into small rectangles and apply Lemma 7.34 to each of them. This is a key idea in the study of homotopies and you will encounter it again in the proof of the theorem of van Kampen, the main result in *Topologie en Meetkunde*.

Indeed, let $a = s_0 < s_1 < \ldots < s_p = b$ be a partition of $[a, b]$ and $c = t_0 < t_1 < \ldots < t_q = d$ a partition of $[c, d]$. Write

$$R_{jk} = [s_{j-1}, s_j] \times [t_{k-1}, t_k],$$

so R is the union of the rectangles R_{jk} for $1 \le j \le p$ and $1 \le k \le q$.

Lemma 7.36. *Let $\Gamma : R \to U$ be a homotopy and α a closed covector field on U For every partition of R into subrectangles as above, we have:*

$$\int_{\Gamma|_{\partial R}} \alpha = \sum_{k=1}^{q} \sum_{j=1}^{p} \int_{\Gamma|_{\partial R_{jk}}} \alpha.$$

Proof. We first prove the lemma in the situation where the number of rectangles R_{ij} equals two, i.e. $pq = 2$. By swapping the roles of the coordinates we may restrict to the case $p = 2$ and $q = 1$. Then $R = R_{11} \cup R_{21}$ with $R_{11} \cap R_{21} = \{s_1\} \times [c, d]$. We write $\sigma = \Gamma|_{\partial R}$ for the boundary of the big rectangle, $\sigma^L = \Gamma|_{\partial R_{11}}$ for the boundary of the left rectangle, and $\sigma^R = \Gamma|_{\partial R_{21}}$ for the boundary for the right one. It holds that:

$$\int_{\Gamma|_{\partial R_{11}}} = \int_{\sigma^L} = \sum_{j=1}^{4} \int_{\sigma_j^L}, \tag{7.6}$$

where we have dropped the integrand α for convenince and split the integral along the boundary into a sum of integrals along each of its four sides. Similarly:

$$\int_{\Gamma|_{\partial R_{21}}} = \int_{\sigma^R} = \sum_{j=1}^{4} \int_{\sigma_j^R}. \tag{7.7}$$

We can see that:

$$\sigma_1 = \sigma_1^L \cdot \sigma_1^R, \qquad \sigma_3 = \sigma_3^R \cdot \sigma_3^L$$

and:

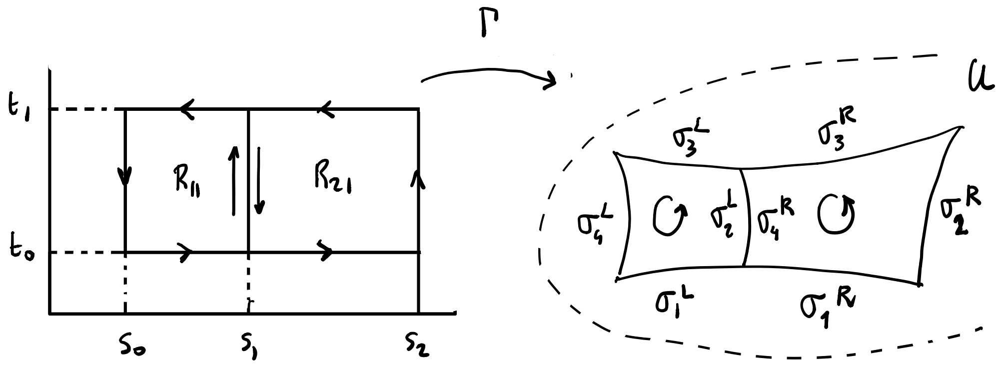
$$\sigma_2 = \sigma_2^R, \qquad \sigma_4 = \sigma_4^L.$$

Adding (7.6) and (7.7), and taking into account Lemma 7.16, we obtain the sum

$$\sum_{j=1}^{4} \int_{\sigma_j} + \int_{\sigma_2^L} + \int_{\sigma_4^R}.$$

Since $\sigma_4^R = \overline{(\sigma_2^L)}$, the last two integrals in this sum cancel each other due to Lemma 7.17. Thus we find:

$$\int_{\Gamma|_{\partial R_{11}}} + \int_{\Gamma|_{\partial R_{21}}} = \sum_{j=1}^{4} \int_{\sigma_j} = \int_{\Gamma|_{\partial R}}.$$



Figuur 12: Proof of Lemma 7.36.

We now treat the general case. For $1 \leq k \leq q$ we write $R(k) = [a,b] \times [t_{k-1}, t_k]$. By repeated application of the special case treated above it follows that

$$\int_{\Gamma|_{\partial R(k)}} \alpha = \sum_{j=1}^{p} \int_{\Gamma|_{\partial R_{jk}}} \alpha.$$

Summing over k and again repeatedly applying the special case we obtain the desired identity. □

We are now ready for the proof of Proposition 7.35.

Proof of Proposition 7.35. Since $R := [a,b] \times [c,d]$ is closed and bounded in $\mathbb{R}^2$ and therefore sequentially compact, it follows that $\Gamma(R)$ is a sequentially compact subset of U. From this, by Lemma 1.15, it follows that there exists $\varepsilon > 0$ such that for every $x \in \Gamma(R)$ we have $B(x;\varepsilon) \subset U$.

From the sequential compactness of R it follows that Γ is uniformly continuous on its domain R. Thus there exists $\delta > 0$ such that for all $\xi, \eta \in R$ with $\|\xi - \eta\| < \delta$ it holds that $\|\Gamma(\xi) - \Gamma(\eta)\| < \varepsilon$. We now choose a partition $a = s_0 < s_1 < \ldots s_p = b$ and a partition $c = t_0 < t_1 < \ldots < t_q$ such that $s_j - s_{j-1} < \delta/2$ and $t_k - t_{k-1} < \delta/2$ for all $1 \leq j \leq p$ and $1 \leq k \leq q$. Let $R_{jk} := [s_{j-1}, s_j] \times [t_{k-1}, t_k]$. Then for all $\xi, \eta \in R_{jk}$ we have $\|\xi - \eta\| < \delta$ hence $\|\Gamma(\xi) - \Gamma(\eta)\| < \varepsilon$. It follows that

$$\Gamma(R_{jk}) \subset B(\Gamma(s_j, t_k); \varepsilon) \subset U.$$

By Theorem 7.6, α is exact on $B(\Gamma(s_j, t_k); \varepsilon)$. By Corollary 7.13 it follows that

$$\int_{\partial \Gamma_{jk}} \alpha = 0.$$

for all $1 \leq j \leq p$ and $1 \leq k \leq q$. The desired result now follows by summing over j and k and applying Lemma 7.36. □

7.4.3 Homotopy invariance of the line integral

We can finally establish our main technical result:

Proof Proof of Theorem 7.19. Let $\Gamma : [a, b] \times [0, 1] \to U$ be a homotopy between γ_0 and γ_1; denote $R = [a, b] \times [0, 1]$. Further write $\sigma = \Gamma|_{\partial R}$ for the boundary curve. The segment σ_1 is a reparametrization of γ_0 and $\overline{\sigma_3}$ a reparametrization of γ_1. Hence, with Proposition 7.35 we find that

$$0 = \int_{\Gamma|_{\partial R}} = \int_{\gamma_0} + \int_{\sigma_2} - \int_{\gamma_1} + \int_{\sigma_4}; \tag{7.8}$$

again, we have omitted the integrand α in the notation. In item (a), σ_2 and σ_4 are constant curves, so the corresponding line integrals are zero. From this the claim follows.

We continue with item (b). In this case, for every $t \in [0, 1]$ we have

$$\sigma_2(t) = \Gamma(b, t) = \Gamma(a, t) = \sigma_4(1 - t),$$

so the curves σ_2 and σ_4 are each other's reverse. From this it follows that the corresponding integrals in (7.8) cancel. This concludes the proof. □

7.5 Line integrals of vector fields along curves

Vectors $v \in \mathbb{R}^n$ and covectors $\alpha \in (\mathbb{R}^n)^*$ are dual to each other, and we think of $\alpha(v)$ as the length of v as witnessed by α. However, the standard inner product in $\mathbb{R}^n$ allows us to see each vector v as the covector $v^* = \langle v, - \rangle$. In down to earth terms this is the usual transposing of a column vector v to see it as a row vector $v^t = v^*$.

What this means is that all the theory we have developed for covector fields can be rephrased in terms of vector fields:

Definition 7.37. Let $U \subset \mathbb{R}^n$ be an open. A vector field is a function $v : U \to \mathbb{R}^n$. △

We can once again talk about vector fields being continuous or C^k. By taking the inner product we then see that the transpose v^t is a covector field in U. In this manner we obtain a bijective correspondence between vector fields and covector fields.

In particular, for each $f \in C^1(U, \mathbb{R})$, the covector field Df is identified with the gradient vector field $\mathrm{grad} f$. The function f is once again called the potential of $\mathrm{grad} f$. We can also focus on those v such that Dv is symmetric, which we call being *rotation-free*. These correspond to closed covector fields. Lastly, we can define line integrals of vector fields v along curves γ as

$$\int_\gamma v = \int_\gamma v^* = \int \langle v(\gamma(t)), \gamma'(t) \rangle dt.$$

It follows that all the results we established in this chapter have a counterpart in the language of vector fields. For instance, the counterpart of Theorem 7.6 states that if U is simply-connected every rotation-free vector field has a potential.

There is nothing to gain by translating all the results to the setting of vector fields. However, you should be aware that many texts use this alternative terminology (particularly older texts in Analysis).

You should also be aware that vector fields represent direction fields (i.e. at each point x, the vector $v(x)$ can be thought as a little arrow indicating a direction motion). As such, vector fields are very interesting on their own and play an important role in Analysis, Dynamics, and Geometry. Their theory will be developed in later courses.

8 Extra: Reeksen

In deze paragraaf behandelen we de basis van de theorie van de complexe reeksen. Op de theorie in deze paragraaf zal later voortgebouwd worden in de cursus *Functies en Reeksen*.

8.1 Reeksen in $\mathbb{C}$

Ons standpunt zal zijn dat het lichaam $\mathbb{C}$ hetzelfde is als $\mathbb{R}^2$ voorzien van de complexe vermenigvuldiging. Hierbij is $x + iy$ een notatie voor $(x, y) \in \mathbb{R}^2$, die er toe dient om aan te geven dat we (x, y) willen zien als complex getal. In deze notatie wordt de complexe vermenigvuldiging gegeven door:

$$(x + iy)(u + iv) = (xu - yv) + i(xv + yu),$$

voor $(x, y) \in \mathbb{R}^2, (u, v) \in \mathbb{R}^2$. Deze formule volgt uit de gebruikelijke rekenregels voor optelling en vermenigvuldiging, aangevuld met de regel dat $i \cdot i = -1$. We kunnen deze rekenregel ook in de $\mathbb{R}^2$-notatie invoeren door

$$(x, y)(u, v) = ((xu - yv), (xv + yu)). \tag{8.1}$$

Het is gemakkelijk in te zien dat deze vermenigvuldiging commutatief is. Als we vervolgens afspreken dat we $\mathbb{R}$ zien als deel van $\mathbb{R}^2$ via de injectieve reëel lineaire afbeelding $x \mapsto (x, 0)$ en dat i een notatie is voor $(0, 1)$, dan krijgen we bekende complexe notatie terug uit

$$(x, y) = (x, 0) + (0, 1)(y, 0) = x + yi.$$

Het gemak van deze identificaties is dat we gegeven definities voor $\mathbb{R}^2$ direct kunnen vertalen naar $\mathbb{C}$. In het bijzonder komt de modulus $|\cdot|$ op $\mathbb{C}$ overeen met de norm $\|\cdot\|$ op $\mathbb{R}^2$. Immers

$$|x + iy| = \sqrt{x^2 + y^2} = \|(x, y)\|.$$

De Euclidische metriek op $\mathbb{R}^2$ wordt in de complexe notatie beschreven door

$$d(z, w) := |z - w|, \qquad (z, w \in \mathbb{C}).$$

Op deze manier wordt het limietbegrip zinvol voor rijen in $\mathbb{C}$. De te verwachten bijbehorende rekenregels volgen gemakkelijk uit de overeenkomstige rekenregels voor rijen in $\mathbb{R}^2$.

Laat $(a_k)_{k\geq 0}$ een rij complexe getallen zijn. We gebruiken de notatie $\sum_{k\geq 0} a_k$ om aan te geven dat we de intentie hebben om de elementen a_k van de gegeven rij te sommeren. Voor $n \geq 0$ definiëren we de n-de partiële som van de reeks door

$$A_n := \sum_{k=0}^{n} a_k \tag{8.2}$$

Remark 8.1. Merk op dat de *reeks* $\sum_{k\geq 0} a_k$ iets anders is dan de *rij* $(a_k)_{k\geq 0}$. Als we de reeks als formeel wiskundige object willen introduceren, dan kunnen we dit beter doen door de reeks te definiëren als de rij $(A_n)_{n\geq 0}$ van partiële sommen. △

Definition 8.2. Laat $(a_k)_{k\geq 0}$ een rij complexe getallen zijn. De *reeks* $\sum_{k\geq 0} a_k$ heet *convergent* indien de n-de partiële sommen A_n, gedefinieerd door (8.2), een convergente rij in $\mathbb{C}$ vormen. In dat geval schrijven we

$$\sum_{k=0}^{\infty} a_k := \lim_{n\to\infty} A_n.$$

Dit getal heet de *som van de reeks*. △

Remark 8.3. Voor een gegeven geheel getal $p \geq 1$ kan men ook een rij $(a_k)_{k\geq p}$ beschouwen en de bijbhorende reeks $\sum_{k\geq p} a_k$. Dit is terug te voeren op het bovenstaande door de reeks $\sum_{k\geq 0} a_{p+k}$ te beschouwen. Aldus zien we dat convergentie van de reeks equivalent is met convergentie van de rij $(A_n)_{n\geq p}$ van partiële sommen, gedefinieerd door $A_n := \sum_{k=p}^{n} a_k$. In geval van convergentie schrijven we dan

$$\sum_{k=p}^{\infty} a_k := \lim_{n\to\infty} A_n. \qquad \triangle$$

Lemma 8.4. *Laat $a_k \geq 0$, voor $k \in \mathbb{N}$. De reeks $\sum_{k\geq 0} a_k$ is convergent dan en slechts dan als de bijbehorende rij (A_n) van partiële sommen naar boven begrensd is. In geval van convergentie is*

$$\sum_{k=0}^{\infty} a_k = \sup\{A_n \mid n \geq 0\}. \tag{8.3}$$

Proof. De partiële sommen zijn reëel en voldoen aan

$$A_{n+1} = A_n + a_n \geq A_n.$$

De rij van partiële sommen is dus monotoon stijgend. In de cursus *Inleiding Analyse* hebben we gezien dat een dergelijke rij convergent is dan en slechts dan als hij naar boven begrensd is. Bovendien geldt in geval van convergentie dat

$$\lim_{n\to\infty} A_n = \sup\{A_n \mid n \geq 0\}.$$

Hieruit volgt (8.3). □

8.2 Convergentie kenmerken

Lemma 8.5. *Laat $a_k \in \mathbb{C}$, voor $k \in \mathbb{N}$. Dan geldt:*

$$\sum_{k\geq 0} a_k \ \text{convergent} \quad \Longrightarrow \quad \lim_{n\to\infty} a_n = 0. \tag{8.4}$$

Proof. Schrijf $A_n = \sum_{k=0}^{n} a_k$. Dan heeft de complexe rij $(A_n)_{n\geq 0}$ een limiet die we noteren met A. Dus $\lim_{n\to\infty} A_n = A$. Hieruit volgt dat ook $\lim_{n\to\infty} A_{n-1} = A$. Anderzijds geldt $a_n = A_n - A_{n-1}$. Met de somregel voor limieten leiden we nu af dat

$$\lim_{n\to\infty} a_n = \lim_{n\to\infty} (A_n - A_{n-1}) = A - A = 0. \qquad \square$$

Remark 8.6 (Waarschuwing)**.** Het omgekeerde van de bewering (8.4) is in het algemeen niet waar. Dit blijkt bijvoorbeeld uit het volgende lemma. △

Lemma 8.7. *De* harmonische *reeks $\sum_{k\geq 1} \frac{1}{k}$ is divergent.*

Proof. We beschouwen de rij A_n van partiële sommen en merken op dat voor $m \geq 1$ geldt:

$$\begin{aligned} A_{2^m} &= A_{2^{m-1}} + \frac{1}{2^{m-1}+1} + \cdots + \frac{1}{2^{m-1}+2^{m-1}} \\ &\geq A_{2^{m-1}} + 2^{m-1}\frac{1}{2^m} = A_{2^{m-1}} + \frac{1}{2}. \end{aligned}$$

Met inductie volgt hieruit dat

$$A_{2^m} \geq \frac{m+1}{2}, \qquad (m \geq 0).$$

De rij (A_n) is in $\mathbb{R}$ niet naar boven begrensd, en daarom niet convergent. We concluderen dat de harmonische reeks divergent is. □

Algemener geldt het volgende resultaat. Voor $s = 1$ geeft dit een ander bewijs van het bovenstaande lemma.

Lemma 8.8. *Zij $s \in \mathbb{R}$. De reeks*

$$\sum_{k\geq 1} \frac{1}{k^s} \tag{8.5}$$

is convergent dan en slechts dan als $s > 1$.

Proof. Als $s \leq 0$, dan geldt niet dat $1/k^s \to 0$ voor $k \to \infty$. Wegens Lemma 8.5 is de reeks dan niet convergent. We mogen ons daarom beperken tot het geval dat $s > 0$. In dit geval is de functie $f : [1, \infty[\to \mathbb{R}, x \mapsto 1/x^s$ continu, monotoon dalend en niet-negatief op $[1, \infty]$. Wegens het onderstaande lemma is de reeks (8.8) convergent dan en slechts dan als de integraal $I_n := \int_1^{n+1} f(x)\, dx$ een limiet heeft voor $n \to \infty$.

Met de fundamentaalstelling voor de integraalrekening zien we dat

$$I_n = \left.\frac{x^{1-s}}{1-s}\right|_1^{n+1} = \frac{(n+1)^{1-s} - 1}{1-s}.$$

Hieruit blijkt dat de rij (I_n) divergent is voor $s < 1$ terwijl voor $s > 1$ geldt $I_n \to 1/(s-1)$, $(n \to \infty)$. Voor $s = 1$ geldt

$$I_n = \log(n+1) \to \infty, \quad (n \to \infty),$$

dus in dit geval is de rij (I_n) divergent. □

Lemma 8.9 (Vergelijking reeks en integraal)**.** *Zij $f : [1, \infty[\to [0, \infty[$ een monotoon dalende functie zo dat voor iedere $N \in \mathbb{N}$ de functie f Riemann-integreerbaar is over $[1, N]$. Schrijf*

$$A_n := \sum_{k=1}^{n} f(k), \quad I_n := \int_1^{n+1} f(x)\, dx \quad \text{en} \quad R_n = A_n - I_n.$$

De rij $(R_n)_{n\geq 1}$ is convergent met een in $[0, f(1)]$ gelegen limiet. In het bijzonder is de rij $(A_n)_{\geq 1}$ convergent dan en slechts dan als de rij $(I_n)_{n\geq 1}$ dat is.

Proof. Zij V_n de verdeling van het interval $[1, n+1]$ in stukken van lengte 1. Dan wordt de bovensom van de integraal van f over $[1, n+1]$ ten aanzien van deze verdeling gegeven door

$$\overline{S}(f, V_n) = \sum_{k=2}^{n+1} f(k-1) = A_n,$$

Hierdoor gemotiveerd vinden we

$$R_n = \sum_{k=1}^{n} \left[f(k) - \int_k^{k+1} f(x)\, dx \right] = \sum_{k=1}^{n} r_k$$

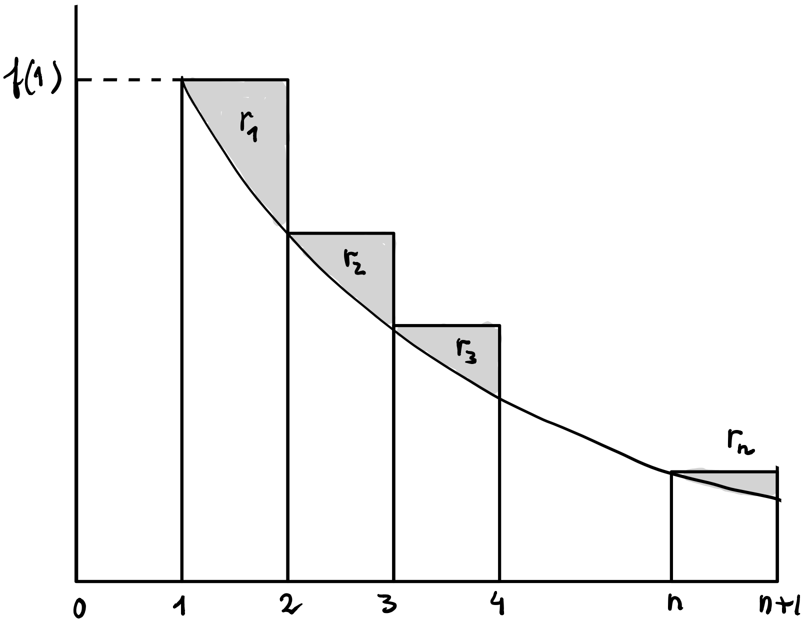
waarin

$$r_k = \int_k^{k+1} [f(k) - f(x)]\, dx,$$

zie Figuur 13.

Uit de monotonie van f volgt dat

$$0 \leq r_k \leq f(k) - f(k+1), \qquad (k \geq 1).$$



Figuur 13: $R_n = \sum_{k=1}^{n} r_k$.

Hieruit volgt dat de rij (R_n) monotoon stijgend is. Door sommatie over k volgt dat $0 \leq R_n \leq f(1) - f(n+1) \leq f(1)$. De rij (R_n) is daarom convergent met een in $[0, f(1)]$ gelegen limiet, die we noteren met R. Is de rij (I_n) convergent met limiet I dan volgt met de somregel voor limieten dat de rij (A_n) convergent is met limiet $R + I$. Is de rij (A_n) convergent met limiet A, dan volgt wederom met de somregel dat de rij (I_n) convergent is met limiet $A - R$. $\square$

Exercise 8.10. Laat zien dat de limiet

$$\gg := \lim_{n\to\infty} \left(\sum_{k=1}^{n} \frac{1}{k} - \log n \right)$$

bestaat, en een positief geheel getal $\gg > 0$ definieert. Het is niet bekend of deze constante van Euler–Mascheroni irrationaal is. $\triangle$

Exercise 8.11 (Rekenregels)**.** Laat $\sum_{k\geq 0} a_k$ en $\sum_{k\geq 0} b_k$ een tweetal convergente complexe reeksen zijn. Dan is ook de reeks $\sum_k (a_k + b_k)$ convergent, terwijl

$$\sum_{k=0}^{\infty} (a_k + b_k) = \sum_{k=0}^{\infty} a_k + \sum_{k=0}^{\infty} b_k$$

Zij $\lambda \in \mathbb{C}$. Dan is ook de reeks $\sum_{k\geq 0} \lambda a_k$ convergent, en er geldt:

$$\lambda \sum_{k=0}^{\infty} a_k = \sum_{k=0}^{\infty} \lambda a_k. \qquad \triangle$$

Ter voorbereiding op de theorie van de reeksen geven we nog het volgende resultaat.

Lemma 8.12. *Zij (a_k) een rij complexe getallen zo dat de reeks $\sum_{k\geq 0} a_k$ convergent is. Dan is voor iedere $n \in \mathbb{N}$ de reeks $\sum_{k\geq n} a_k$ convergent. Bovendien geldt*

$$\lim_{n\to\infty} \sum_{k=n}^{\infty} a_k = 0.$$

Proof. Voor alle $m \geq n$ geldt $\sum_{k=0}^{m} a_k = \sum_{k=0}^{n-1} a_k + \sum_{k=n}^{m} a_k$. Door de limiet voor $m \to \infty$ te nemen blijkt hieruit dat de genoemde reeks $\sum_{k\geq n} a_k$ convergent is, terwijl

$$\sum_{k=n}^{\infty} a_k = \sum_{k=0}^{\infty} a_k - \sum_{k=0}^{n-1} a_k.$$

Door de limiet voor $n \to \infty$ te nemen leiden we hieruit met de somregel voor limieten af dat

$$\lim_{n\to\infty} \sum_{k=n}^{\infty} a_k = \sum_{k=0}^{\infty} a_k - \lim_{n\to\infty} \sum_{k=0}^{n-1} a_k = 0.$$ □

Example 8.13 (Meetkundige reeks). Zij $z \in \mathbb{C}$. De reeks

$$\sum_{k\geq 0} z^k \tag{8.6}$$

staat bekend als de *meetkundige reeks* met *reden* z. Hierbij dient z^0 gelezen te worden als 1, ook als $z = 0$.

Zij $S_n = \sum_{k=0}^{n} r^k$ de n-de partiële som van de reeks. Dan geldt:

$$zS_n - S_n = \sum_{k=1}^{n+1} z^k - \sum_{k=0}^{n} z^k = z^{n+1} - 1,$$

dus als $z \neq 1$, dan is

$$S_n = \frac{1 - r^{n+1}}{1 - r}.$$

Als $z = 1$ dan is $S_n = n + 1$. △

Lemma 8.14. *Zij $z \in \mathbb{C}$. De meetkundige reeks (8.6) convergeert dan en slechts dan als $|z| < 1$. Voor $|z| < 1$ geldt dat*

$$\sum_{k=0}^{\infty} z^k = \frac{1}{1 - z}.$$

Proof. Voor $|z| \geq 1$ geldt $|z^k| \geq 1$, dus niet $\lim_{k\to\infty} z^k = 0$. Hieruit volgt wegens Lemma 8.5 dat de meetkundige reeks divergeert.

Veronderstel nu dat $|z| < 1$. Dan geldt dat $|z^{n+1}| = |z|^{n+1} \to 0$ voor $n \to \infty$, dus

$$\lim_{n\to\infty} S_n = \frac{1 - \lim z^{n+1}}{1 - z} = \frac{1}{1 - z}.$$

Hieruit volgt het gestelde. □

Definition 8.15 (Absolute convergentie). Een complexe reeks $\sum_{k\geq 0} a_k$ heet *absoluut convergent* indien de reeks $\sum_{k\geq 0} |a_k|$ convergent is. △

Lemma 8.16. *Laat $(a_k)_{k\geq 0}$ een rij complexe getallen zijn. Indien de reeks $\sum_{k\geq 0} a_k$ absoluut convergent is, dan is hij ook convergent, en er geldt:*

$$\left|\sum_{k=0}^{\infty} a_k\right| \leq \sum_{k=0}^{\infty} |a_k|. \tag{8.7}$$

Proof. Uit de absolute convergentie volgt dat de rij $B_n := \sum_{k=0}^{n} |a_k|$ convergent is, dus Cauchy. Zij $\geq> 0$ dan bestaat er een N zo dat voor alle $q \geq p \geq N$ geldt $|B_p - B_q| <\geq$. Zij A_n de n-de partiële som van de reeks $\sum_{k\geq 0} a_k$. Dan geldt voor alle $q \geq p \geq N$ dat

$$|A_q - A_p| = |\sum_{k=p+1}^{q} a_k| \leq \sum_{k=p+1}^{q} |a_k| = |B_q - B_p| <\geq .$$

Dus (A_n) is een Cauchy-rij, en aangezien $\mathbb{C} \simeq \mathbb{R}^2$ volledig is, concluderen we dat deze rij van partiële sommen convergeert. De reeks $\sum_{k\geq 0} a_k$ convergeert dus.

Voor alle $n \geq 0$ geldt dat

$$\left|\sum_{k=0}^{n} a_k\right| \leq \sum_{k=0}^{n} |a_k|.$$

Nemen we de limiet voor $n \to \infty$, dan concluderen we dat (8.7) geldt. □

Om voor de hand liggende redenen vatten we de uitspraak dat de reeks $\sum_{k\geq 0} a_k$ absoluut convergeert soms ook samen in de formule

$$\sum_{k=0}^{\infty} |a_k| < \infty.$$

Exercise 8.17. Is $\sum_{k\geq 0} a_k$ een absoluut convergente complexe reeks, dan zijn ook beide reeksen

$$\sum_{j\geq 0} a_{2j} \quad \text{en} \quad \sum_{j\geq 0} a_{2j+1}$$

absoluut convergent, terwijl

$$\sum_{k=0}^{\infty} a_k = \sum_{j=0}^{\infty} a_{2j} + \sum_{j=0}^{\infty} a_{2j+1}$$ △

Het volgende resultaat wordt zeer vaak gebruikt om de convergentie van reeksen aan te tonen.

Theorem 8.18 (Majorantiekenmerk voor convergentie). *Laat (a_k) een complexe rij zijn, en (t_k) een reële rij, terwijl er een $C > 0$ bestaat zo dat*

$$|a_k| \leq Ct_k \qquad (\forall k \geq 0).$$

Indien $\sum_k t_k$ convergeert, dan convergeert de reeks $\sum_k a_k$ absoluut, en er geldt dat

$$\sum_{k=0}^{\infty} |a_k| \leq C \sum_{k=0}^{\infty} t_k. \tag{8.8}$$

Proof. We noteren de n-de partiële som van de reeks $\sum_{k\geq 0} t_k$ met T_n, en die van $\sum_{k\geq 0} |a_k|$ met B_n. Uit het gegeven volgt dat $B_n \leq CT_n$ voor alle n. Veronderstel dat de reeks $\sum_k t_k$ convergent is, dan is de rij (T_n) convergent, dus begrensd. Er is dus een $M > 0$ zo dat $T_n \leq M$. De rij (B_n) is dus begrensd door CM. Uit

$$B_{n+1} = B_n + |a_{n+1}| \geq B_n$$

volgt dat de rij (B_n) monotoon stijgend en naar boven begrensd is, dus convergent. We concluderen dat de reeks $\sum_k |a_k|$ convergent is. Voor alle $n \geq 0$ geldt

$$\sum_{k=0}^{n} |a_k| \leq C \sum_{k=0}^{n} t_k.$$

De ongelijkheid (8.8) volgt hieruit door limietovergang voor $n \to \infty$. □

Door het majorantie criterium te combineren met kennis over de meetkundige reeks leiden we het volgende af.

Lemma 8.19 (Quotiëntkenmerk)**.** *Zij (a_n) een rij in $\mathbb{C}$ zo dat*

$$\lim_{k\to\infty} \left| \frac{a_{k+1}}{a_k} \right| = L$$

(in het bijzonder veronderstellen we het bestaan van de limiet).

Als $|L| < 1$, dan is de reeks

$$\sum_{k\geq 0} a_k \tag{8.9}$$

absoluut convergent. Als $|L| > 1$, dan is de reeks divergent.

Proof. Veronderstel eerst dat $L < 1$. Kies $\geq > 0$ zo dat $L+ \geq < 1$. Dan is er een N zo dat voor $k \geq N$ geldt dat

$$\frac{|a_{k+1}|}{|a_k|} \leq L+ \geq .$$

Voor $k \geq N$ geldt daarom

$$|a_k| \leq (L+ \geq)^{k-N} |a_N| \leq C(L+ \geq)^k$$

waarbij $C = (L+ \geq)^{-N}|a_N|$. Aangezien de meetkundige reeks $\sum_k (L+ \geq)^k$ convergent is (Lemma 8.14) volgt nu met Stelling 8.18 dat de reeks (8.9) absoluut convergeert. □

Example 8.20. Als in de setting van het bovenstaande lemma geldt dat $L = 1$, dan kan de reeks zowel convergeren als divergeren. Nemen we $a_k = k^{-s}$, met $s > 0$, dan geldt dat $L = 1$, terwijl de reeks

$$\sum_{k\geq 1} a_k$$

volgens Lemma 8.8 convergeert voor $s > 1$ en divergeert voor $s \leq 1$. △

Example 8.21 (Complexe e-macht)**.** We beschouwen, voor $z \in \mathbb{C}$, de reeks

$$\sum_{k\geq 0} \frac{z^k}{k!}.$$

Schrijven we $a_k = z^k/k!$ dan zien we dat

$$\left| \frac{a_{k+1}}{a_k} \right| = \frac{|z|}{k+1} \to L = 0$$

voor $k \to \infty$. Hieruit volgt dat de reeks convergeert, voor elke $z \in \mathbb{C}$. We definiëren de complexe e-macht door

$$e^z = \sum_{k=0}^{\infty} \frac{z^k}{k!}. \tag{8.10}$$

Wegens een eerder in de cursus *Inleiding Analyse* gegeven toepassing van de stelling van Taylor met rest komt deze e-macht voor reële z overeen met de bekende reële e-macht.

Door $z = iy$ met $y \in \mathbb{R}$ in te vullen in (8.10) en de reeks te splitsen als in Opmerking 8.17 vinden we dat

$$\begin{aligned} e^{iy} &= \sum_{k \geq 0} i^{2k} \frac{y^{2k}}{(2k)!} + \sum_{k \geq 0} i^{2k+1} \frac{y^{2k+1}}{(2k+1)!} \\ &= \sum_{k \geq 0} (-1)^k \frac{y^{2k}}{(2k)!} + i \sum_{k \geq 0} (-1)^k \frac{y^{2k+1}}{(2k+1)!}. \end{aligned}$$

Hieruit volgt de bekende formule

$$e^{iy} = \cos y + i \sin y.$$

△

9 Extra: Oneigenlijke integralen

In deze paragraaf zullen we het begrip *oneigenlijke integraal* invoeren. Daarna zullen we oneigenlijke integralen met een parameter beschouwen, zodat we in het bijzonder het gedrag van de integraal voor de Gamma-functie zullen kunnen analyseren, zie Voorbeeld 5.18.

Het begrip oneigenlijke Riemann-integraal is een verruiming van het begrip Riemann-integraal van gesloten en begrensde intervallen naar willekeurige intervallen.

Example 9.1. Als eerste motiverende voorbeeld beschouwen we de integraal

$$\int_0^\infty e^{-x}\,dx.$$

De functie $f : [0,\infty[\to \mathbb{R},\ x \mapsto e^{-x}$ is continu, en dus Riemann-integreerbaar over ieder gesloten en begrensd interval van de vorm $[0,\beta]$, met $0 \le \beta < \infty$. Met de hoofdstelling van de integraalrekening vinden we

$$\int_0^\beta e^{-x} = \left[-e^{-x}\right]_0^\beta = 1 - e^{-\beta}.$$

Hieraan zien we dat

$$\lim_{\beta\to\infty}\int_0^\beta e^{-x}\,dx = 1.$$

We zeggen ook wel dat $x \mapsto e^{-x}$ oneigenlijk Riemann-integreerbaar is over $[0,\infty[$, met als oneigenlijke integraal

$$\int_0^\infty e^{-x}\,dx = 1. \qquad \triangle$$

Example 9.2. Als tweede motiverend voorbeeld beschouwen we de functie $f :]0,1] \to \mathbb{R}$ gedefinieerd door

$$f(x) = \frac{1}{\sqrt{x}}.$$

Deze functie is niet begrensd op $]0,1]$, dus kan niet opgevat worden als Riemann-integreerbare functie op $[0,1]$ (door hem een willekeurige waarde in 0 toe te kennen). Hij is continu, dus Riemann-integreerbaar op $[\alpha,1]$ voor iedere $\alpha \in]0,1]$. Bovendien volgt met de hoofdstelling van de integraalrekening dat

$$\int_\alpha^1 f(x)\,dx = [2\sqrt{x}]_\alpha^1 = 2 - 2\sqrt{\alpha}.$$

Hieruit volgt dat

$$\lim_{\alpha\downarrow 0}\int_\alpha^1 f(x)\,dx = 2.$$

In dit geval zeggen we dat $f : x \mapsto 1/\sqrt{x}$ oneigenlijk Riemann-integreerbaar is over $]0,1]$, en we schrijven

$$\int_0^1 f(x)\,dx = 2. \qquad \triangle$$

Example 9.3. We beschouwen de functie $f : I =]0,\infty[\to \mathbb{R}$ gedefinieerd door

$$f(x) = \frac{1}{e^x\sqrt{x}}, \qquad (0 < x < 1).$$

In dit geval is f continu, dus Riemann-integreerbaar over ieder segment $[\alpha,\beta] \subset]0,\infty[$. Het ligt voor de hand te zeggen dat f oneigenlijk Riemann-integreerbaar is over $]0,\infty[$ indien de integraal

$$\int_\alpha^\beta f(x)\, dx$$

een limiet heeft voor $\alpha \downarrow 0$ en $\beta \to \infty$. Daarmee bedoelen we dat er een $S \in \mathbb{R}$ bestaat zo dat voor iedere $\geq < 0$ elementen $\alpha_0, \beta_0 \in I$ bestaan zo dat voor alle $\alpha, \beta \in I$ geldt

$$\alpha \leq \alpha_0, \beta \geq \beta_0 \implies \left| \int_\alpha^\beta f(x)\, dx - S \right| < \geq .$$

Verderop zullen we zien dat zo'n S in dit geval bestaat en uniek is, en dan schrijven we

$$\int_0^\infty \frac{1}{e^x \sqrt{x}}\, dx = S. \qquad \triangle$$

Geïnspireerd door de bovenstaande Voorbeelden 9.1, 9.2 en 9.3 zullen we een definitie opstellen van oneigenlijke integreerbaarheid voor een functie f met als domein een niet-leeg interval $I \subset \mathbb{R}$. Om dit in algemeenheid te kunnen doen gebruiken we de bekende karakterisering van een interval uit de cursus 'Inleiding Analyse'.

Karakterisering interval. Een interval is een deelverzameling I van $\mathbb{R}$ met de eigenschap dat voor alle $\alpha, \beta \in I$ met $\alpha < \beta$ geldt $[\alpha, \beta] \subset I$. De nu volgende definitie dient ertoe het bestaan van de Riemann-integraal van een functie $I \to \mathbb{R}$ over een deelsegment van I te garanderen.

Definition 9.4. (Lokaal Riemann-integreerbaar) Laat $I \subset \mathbb{R}$ een niet-leeg interval zijn. Een functie $f : I \to \mathbb{R}$ heet *lokaal Riemann-integreerbaar* indien voor alle $\alpha, \beta \in I$ met $\alpha < \beta$ geldt dat de beperking $f|_{[\alpha,\beta]}$ Riemann-integreerbaar is over $[\alpha, \beta]$. $\triangle$

Remark 9.5. We merken op dat een continue functie $f : I \to \mathbb{R}$ lokaal Riemann-integreerbaar is. $\triangle$

De setting van Definitie 9.4 garandeert het bestaan van de Riemann integraal $\int_a^\beta f(x)\, dx$, voor iedere $\beta \geq a$. Hiermee wordt de volgende definitie zinvol.

Definition 9.6. Laat $I \subset \mathbb{R}$ een niet-leeg interval zijn en veronderstel dat $f : I \to \mathbb{R}$ lokaal Riemann-integreerbaar is. Laat $S \in \mathbb{R}$; dan betekent

$$\lim_{[\alpha,\beta]\nearrow I} \int_\alpha^\beta f(x)\, dx = S \tag{9.11}$$

dat er voor elke $\geq > 0$ een gesloten en begrensd interval $I_0 \subset I$ bestaat met de volgende eigenschap. Voor alle $\alpha, \beta \in \mathbb{R}$ met $\alpha < \beta$ geldt

$$I_0 \subset [\alpha, \beta] \subset I \quad \implies \quad \left| \int_\alpha^\beta f(x)\, dx - S \right| < \geq . \tag{9.12}$$

$\triangle$

De hierboven geïntroduceerde limiet is uniek bepaald.

Lemma 9.7. *Laat $f : I \to \mathbb{R}$ lokaal Riemann-integreerbaar zijn en veronderstel dat $S, S' \in \mathbb{R}$ en*

$$\lim_{[\alpha,\beta]\nearrow I} \int_\alpha^\beta f(x)\, dx = S \quad \text{en} \quad \lim_{[\alpha,\beta]\nearrow I} \int_\alpha^\beta f(x)\, dx = S'.$$

Dan is $S = S'$.

Proof. Kies I_0 en I_0' als in (9.11) voor respectievelijk S en S'. Zij $[\alpha, \beta] \subset I$ een gesloten en begrensd interval dat zowel I_0 als I_0' bevat. Dan volgt dat

$$|S - S'| \leq \left| S - \int_\alpha^\beta f(x)\, dx \right| + \left| \int_\alpha^\beta f(x)\, dx - S' \right| < 2 \geq .$$

Dit geldt voor elke $\geq > 0$, dus $S = S'$. □

Definition 9.8. (Oneigenlijk Riemann-integreerbaar) Zij $I \subset \mathbb{R}$ een niet-leeg interval. Een functie $f : I \to \mathbb{R}$ heet *oneigenlijk Riemann-integreerbaar* over I indien het volgende geldt

(a) De functie f is lokaal Riemann-integreerbaar.

(b) Er bestaat een (noodzakelijkerwijs uniek) getal $S \in \mathbb{R}$ zo dat (9.11) geldt.

Is $f : I \to \mathbb{R}$ oneigenlijk Riemann-integreerbaar, dan noemen we het unieke getal S uit (b) de *oneigenlijke Riemann-integraal* van f over I, notatie

$$\int_I f(x)\, dx := S = \lim_{[\alpha,\beta] \nearrow I} \int_\alpha^\beta f(x)\, dx. \qquad \triangle$$

Remark 9.9. Is $f : I \to \mathbb{R}$ lokaal Riemann-integreerbaar, dan zeggen we in plaats van 'f is oneigenlijk Riemann-integreerbaar over I' ook wel dat de integraal $\int_I f(x)\, dx$ *convergeert.* Is f niet oneigenlijk Riemann-integreerbaar, dan zeggen we ook wel dat de integraal *divergeert.* △

De oneigenlijke Riemann-integreerbaarheid is voor een niet-negatieve functie te karakteriseren op een manier die sterke overeenkomst vertoont met Lemma 8.4

Lemma 9.10. *Zij $I \subset \mathbb{R}$ een niet-leeg interval, en $f : I \to \mathbb{R}$ een lokaal Riemann-integreerbare functie met $f(x) \geq 0$ voor alle $x \in I$. Dan zijn de volgende uitspraken equivalent.*

(a) *f is oneigenlijk Riemann-integreerbaar over I.*

(b) *Er is een $M > 0$ zo dat voor alle $\alpha, \beta \in \mathbb{R}$ met $\alpha < \beta$ geldt*

$$[\alpha, \beta] \subset I \implies \int_\alpha^\beta f(x)\, dx \leq M.$$

Is aan (a) en (b) voldaan, dan is

$$\int_I f(x)\, dx = \sup_{[\alpha,\beta] \subset I} \int_\alpha^\beta f(x)\, dx.$$

Proof. We beginnen met de opmerking dat voor ieder tweetal segmenten $I_1 = [\alpha_1, \alpha_1] \subset I$ en $I_2 = [\alpha_2, \beta_2] \subset I$ met $I_1 \subset I_2$ geldt dat

$$\int_{I_2} f(x)\, dx = \int_{\alpha_2}^{\alpha_1} f(x)\, dx + \int_{\alpha_1}^{\beta_1} f(x)\, dx + \int_{\beta_1}^{\beta_2} f(x)\, dx \geq \int_{I_1} f(x)\, dx.$$

Veronderstel nu eerst dat (a) geldt, en laat $S := \int_I f(x)\, dx$.

Dan is er een segment $I_0 \subset I$ zo dat voor $a < b$ met $I_0 \subset [a, b] \subset I$ geldt $|\int_a^b f(x)\, dx - S| < 1$. Uit dit laatste volgt

$$\int_a^b f(x)\, dx < S + 1. \tag{9.13}$$

Zij nu $\alpha < \beta$ zo dat $[\alpha, \beta] \subset I$. Zij $a = \min(I_0 \cup \{\alpha\})$ en $b = \max(I_0, \{\beta\})$. Dan geldt $I_0 \subset [a, b] \subset I$, dus (9.13). Uit $a \leq \alpha \leq \beta \leq b$ volgt wegens het eerste deel van het bewijs dat

$$\int_\alpha^\beta f(x)\, dx \leq \int_a^b f(x)\, dx < S + 1.$$

Dus (b) geldt met $M = S + 1$.

Veronderstel omgekeerd dat (b) geldt. Dan is de collectie

$$V := \left\{ \int_\alpha^\beta f(x)\, dx \mid \alpha < \beta, [\alpha, \beta] \subset I \right\}$$

een niet-lege deelverzameling van $\mathbb{R}$ die naar boven begrensd is door M. Derhalve heeft V een kleinste bovengrens

$$S = \sup V.$$

Zij $\geq > 0$. Dan is $S - \geq$ geen bovengrens van V dus er is een segment $I_0 \subset I$ zo dat $\int_{I_0} f(x)\, dx > S - \geq$. Zij $\alpha < \beta$ zo dat $I_0 \subset [\alpha, \beta] \subset I$, dan volgt wegens het eerste deel van het bewijs dat

$$S - \geq < \int_{I_0} f(x)\, dx \leq \int_\alpha^\beta f(x)\, dx \leq S.$$

Hieruit volgt (a), terwijl

$$\int_I f(x)\, dx = S = \sup_{[\alpha,\beta]\subset I} \int_\alpha^\beta f(x)\, dx.$$

□

Example 9.11. We beschouwen nogmaals de functie $f : I =]0, \infty[\to \mathbb{R}$ uit Voorbeeld 9.3 gegeven door

$$f(x) = \frac{1}{e^x \sqrt{x}}.$$

Deze functie is continu, dus lokaal Riemann-integreerbaar op I, terwijl $f(x) > 0$ voor alle $x \in I$. Voor $0 < x \leq 1$ geldt $f(x) \leq e/\sqrt{x}$, dus voor $0 < \alpha \leq 1$ geldt:

$$\int_\alpha^1 f(x)\, dx \leq e \int_\alpha^1 \frac{1}{\sqrt{x}} = e(2 - 2\sqrt{\alpha}) < 2e.$$

Voor $x \geq 1$ geldt $f(x) \leq e^{-x}$, dus voor $1 \leq \beta < \infty$ geldt:

$$\int_1^\beta f(x)\, dx \leq \int_1^\beta e^{-x}\, dx = 1 - e^{-\beta} < 1.$$

Hieruit leiden we gemakkelijk af dat voor alle $0 < \alpha < \beta < \infty$ geldt dat

$$\int_\alpha^\beta f(x)\, dx \leq 2e + 1.$$

Wegens het bovenstaande lemma is f daarom oneigenlijk Riemann-integreerbaar op I. △

Uit het volgende resultaat blijkt dat oneigenlijke Riemann-integreerbaarheid voor gesloten en begrensde intervallen samenvalt met Riemann-integreerbaarheid.

Lemma 9.12. *Zij $I = [a, b] \subset \mathbb{R}$ een gesloten en begrensd interval, met $a < b$. Dan zijn de volgende twee uitspraken equivalent.*

(a) *De functie f is Riemann-integreerbaar op I.*

(b) *De functie f is oneigenlijk Riemann-integreerbaar op I.*

Is aan (een van de) eisen (a) en (b) voldaan, dan is

$$\int_I f(x)\,dx = \int_a^b f(x)\,dx. \tag{9.14}$$

Proof. Uit (b) volgt per definitie dat f lokaal Riemann-integreerbaar is op $[a, b]$, dus Riemann-integreerbaar op $[a, b]$.

Veronderstel dat (a) geldt. Dan is f ook lokaal Riemann-integreerbaar. Kies $I_0 = [a, b]$, dan geldt voor alle $\alpha < \beta$ met $I_0 \subset [\alpha, \beta] \subset I$ dat $I_0 = [\alpha, \beta] = I$, dus ook

$$\left| \int_\alpha^\beta f(x)\,dx - \int_a^b f(x)\,dx \right| = 0 <\geq$$

voor iedere $\geq> 0$. Hieruit volgt dat (b) geldt en dat bovendien (9.14). □

Remark 9.13. Uit de cursus 'Inleiding Analyse' weten we dat ieder niet-leeg interval $I \subset \mathbb{R}$ één van de volgende vormen heeft:

(a) $I = [a, b]$ met $-\infty < a < b < \infty$,

(b) $I = [a, b\,[$ met $-\infty < a < b \leq \infty$,

(c) $I = \,]\,a, b]$ met $-\infty \leq a < b < \infty$,

(d) $I = \,]\,a, b\,[$ met $-\infty \leq a < b \leq \infty$.

In al deze gevallen noemen we a en b de grenzen van het interval en schrijven we ook

$$\int_a^b f(x)\,dx := \int_I f(x)\,dx = \lim_{[\alpha,\beta]\nearrow I} f(x)\,dx.$$

Bovendien hanteren we de conventie dat

$$\int_b^a f(x)\,dx := -\int_a^b f(x)\,dx.$$

Wegens Lemma 9.12 is deze notatie in overeenstemming met de reeds gebruikte notatie voor de eigenlijke Riemann integraal. △

In het vervolg veronderstellen we dat $I \subset \mathbb{R}$ een interval is, dat niet gesloten en begrensd is. Zo'n interval heeft dus één van de in Opmerking 9.13 genoemde vormen (b)-(c). We zullen de definitie van oneigenlijke Riemann-integreerbaarheid in elk van deze gevallen apart onderzoeken.

Lemma 9.14. *Veronderstel dat $I \subset \mathbb{R}$ een interval van de vorm $I = [a, b[$ is, met $a < b \leq \infty$. Veronderstel nu dat $f : I \to \mathbb{R}$ een lokaal Riemann-integreerbare functie is. Dan zijn de volgende beweringen equivalent.*

(a) *De functie f is oneigenlijk Riemann-integreerbaar over het interval $[a, b[$.*

(b) *De limiet*

$$\lim_{\beta \uparrow b} \int_a^\beta f(x)\, dx$$

bestaat.

Indien (a) en (b), dan geldt

$$\int_a^b f(x)\, dx = \lim_{\beta \uparrow b} \int_a^\beta f(x)\, dx. \tag{9.15}$$

In de bovenstaande situatie noteren we de integraal in het vervolg ook met

$$\int_a^b f(x)\, dx = \lim_{\beta \uparrow b} \int_a^\beta f(x)\, dx.$$

Proof. Veronderstel eerst dat (a) geldt. Zij $\geq> 0$ Kies een gesloten en begrensd interval $I_0 = [a_0, b_0] \subset [a, b[$ zo dat (9.12) geldt. Dan geldt voor alle $\beta \in]b_0, b[$ dat $I_0 \subset [a, \beta] \subset I$, dus

$$\left| \int_a^\beta f(x)\, dx - \int_I f(x)\, dx \right| <\geq .$$

Hieruit blijkt dat (b) geldt, met limiet gelijk aan $\int_I f(x)\, dx$.

Veronderstel omgekeerd dat (b) geldt, en zij S de waarde van de limiet. Zij $\geq> 0$. Dan is er een b_0 met $a < b_0 < b$ zo dat voor alle $\beta \in [b_0, b[$ geldt

$$\left| \int_a^\beta f(x)\, dx - S \right| <\geq .$$

Zij $\alpha, \beta \in \mathbb{R}$ zo dat $\alpha < \beta$ en $I_0 \subset [\alpha, \beta] \subset I$. Dan geldt dat $\alpha = a$ en $b_0 \leq \beta < b$, en de bovenstaande schatting geldt met $a = \alpha$. Hieraan zien we dat

$$S = \lim_{[\alpha,\beta] \nearrow I} \int_\alpha^\beta f(x)\, dx.$$

We concluderen dat (a) geldt, en (9.15). □

Corollary 9.15. *Zij $-\infty < a < b \leq \infty$ en zij $f : [a, b[\to \mathbb{C}$ lokaal Riemann-integreerbaar. Zij $a \leq c < b$. Dan zijn de volgende uitspraken equivalent.*

(a) *De functie f is oneigenlijk Riemann-integreerbaar over $[a, b[$.*

(b) *De functie f is oneigenlijk Riemann-integreerbaar over $[c, b[$.*

Als (a) en (b) gelden, dan geldt bovendien dat

$$\int_a^b f(x)\, dx = \int_a^c f(x)\, dx + \int_c^b f(x)\, dx$$

Proof. Voor alle $\beta > c$ geldt

$$\int_a^\beta f(x)\, dx = \int_a^c f(x)\, dx + \int_c^\beta f(x)\, dx$$

De uitspraken volgen hieruit door de limiet voor $\beta \uparrow b$ te nemen. □

Example 9.16. We beschouwen de functie $f : x \mapsto x^s$ op $I = [1, \infty[$, met $s \in \mathbb{R}$ een constante, ongelijk aan -1. Deze functie is continu, dus lokaal Riemann-integreerbaar. Voor $\beta > 1$ geldt dat

$$\int_1^\beta f(x)\, dx = \frac{x^{s+1}}{s+1}\Big|_1^\beta = \frac{\beta^{s+1} - 1}{s+1}. \tag{9.16}$$

De laatste uitdrukking heeft een limiet voor $\beta \uparrow \infty$ dan en slechts dan als $s + 1 < 0$. In dit geval is de functie f oneigenlijk Riemann integreerbaar over $[1, \infty[$, met als oneigenlijke integraal de limiet:

$$\int_1^\infty x^s\, dx = \lim_{\beta\to\infty} \frac{\beta^{s+1} - 1}{s+1} = -\frac{1}{s+1}, \qquad (s < -1).$$

De uitdrukking (9.16) heeft geen limiet voor $s > -1$, ofwel, de integraal divergeert in dat geval.

Tenslotte beschouwen we ook nog het geval dat $s = -1$. Dan heeft $f(x) = 1/x$ de functie $\log x$ als primitieve, en dus heeft

$$\int_1^\beta \frac{1}{x}\, dx = \log \beta$$

geen limiet voor $\beta \to \infty$. De bijbehorende integraal $\int_1^\beta x^{-1}\, dx$ is dan ook divergent. Samenvattend concluderen we dat het onderstaande lemma geldt. △

Lemma 9.17. *Zij $s \in \mathbb{R}$. Dan convergeert de oneigenlijke Riemann-integraal*

$$\int_1^\infty x^s\, dx \tag{9.17}$$

dan en slechts dan als $s < -1$. In dat geval is de waarde van de integraal gelijk aan $1/(-s-1)$.

Soortgelijke beschouwingen als hier boven leiden tot een andere karakterisering van oneigenlijke Riemann-integreerbaarheid op intervallen als in Opmerking 9.13 (c), dus $I =]a, b]$ met $-\infty \le a < b < \infty$.

Een interessant voorbeeld wordt gegeven door het onderstaande lemma.

Lemma 9.18. *Zij $s \in \mathbb{R}$. De oneigenlijke integraal*

$$\int_0^1 x^s\, dx$$

is convergent dan en slechts dan als $s > -1$. In dat geval is de oneigenlijke integraal gelijk aan $1/(s+1)$.

Proof. De functie $f : x \mapsto x^s$ is continu op het interval $I =]0, 1]$, dus Riemann-integreerbaar op ieder deelinterval $[\alpha, 1] \subset I$. We veronderstellen eerst dat $s \neq -1$. Dan is $(s+1)^{-1}x^{s+1}$ primitieve van f, dus

$$\int_\alpha^1 x^s\, dx = \frac{1}{s+1} - \frac{\alpha^{s+1}}{s+1}$$

voor alle $0 < \alpha < 1$. We zien dat de limiet voor $\alpha \downarrow 0$ bestaat dan en slechts dan als $s > -1$. In dat geval geldt

$$\int_0^1 x^s = \frac{1}{s+1}.$$

We beschouwen tenslotte het geval dat $s = -1$. Dan heeft f de functie log als primitieve op I, zodat

$$\int_\alpha^1 x^{-1}\, dx = -\log \alpha.$$

Deze uitdrukking heeft geen limiet voor $\alpha \downarrow 0$, zodat de bijbehorende oneigenlijke integraal divergent is. Het lemma volgt. $\square$

Om het geval (d) van Opmerking 9.13 te begrijpen, waarin sprake is van een tweezijdig open interval $I = \,]a, b[$, met $-\infty \leq a < b \leq \infty$, hebben we de volgende karakterisering van oneigenlijke integreerbaarheid nodig.

Lemma 9.19 (Cauchy criterium). *Zij $I \subset \mathbb{R}$ een niet-leeg interval en zij $f : I \to \mathbb{R}$ lokaal Riemann-integreerbaar. Dan zijn de volgende twee uitspraken equivalent.*

(a) *De functie f is oneigenlijk Riemann-integreerbaar over I.*

(b) *Voor elke $\geq > 0$ bestaat een gesloten en begrensd interval $I_0 \subset I$ zo dat voor alle gesloten en begrensde intervallen J_1, J_2 met $I_0 \subset J_j \subset I$ geldt dat*

$$\left| \int_{J_1} f(x)\, dx - \int_{J_2} f(x)\, dx \right| < \geq .$$

Proof. Veronderstel eerst dat (a) geldt. Zij $\geq > 0$. Dan is er een gesloten en begrensd interval $I_0 \subset I$ zo dat voor ieder gesloten en begrensd interval $[\alpha, \beta]$ met $I_0 \subset [\alpha, \beta] \subset I$ geldt

$$\left| \int_\alpha^\beta f(x)\, dx - \int_I f(x)\, dx \right| < \geq /2.$$

Zijn J_1 en J_2 gesloten en begrensde intervallen als in (b), dan geldt dat

$$\begin{aligned} &\left| \int_{J_1} f(x)\, dx - \int_{J_2} f(x)\, dx \right| \\ &\quad \leq \left| \int_{J_1} f(x)\, dx - \int_I f(x)\, dx \right| + \left| \int_I f(x)\, dx - \int_{J_2} f(x)\, dx \right| < \geq . \end{aligned}$$

Hieruit volgt (b).

We veronderstellen nu dat (b) geldt. Er bestaat een rij $J(n) = [\alpha_n, \beta_n]$ van gesloten en begrensde intervallen zo dat $J(n) \subset J(n+1) \subset I$ en zo dat

$$\bigcup_{k \geq 0} J(k) = I.$$

We zullen laten zien dat de integraalwaarden

$$S_n := \int_{J(n)} f(x)\, dx$$

een Cauchy-rij in $\mathbb{R}$ vormen. Laat $\geq > 0$. Dan is er een gesloten en begrensd interval $[\alpha, \beta] \subset I$ als in (b). Er bestaan $N_1, N_2 \in \mathbb{N}$ zodat $\alpha \in J(N_1)$ en $\beta \in J(N_2)$. Zij $N = \max(N_1, N_2)$, dan geldt $[\alpha, \beta] \subset J(N)$. Voor $p, q > N$ geldt $[a, b] \subset J(p)$ en $[\alpha, \beta] \subset J(q)$ dus $|S_p - S_q| < \geq$. De rij (S_n) is dus inderdaad Cauchy in $\mathbb{R}$. Wegens de volledigheid van $\mathbb{R}$ bestaat $S = \lim_{n \to \infty} S_n$.

We tonen tenslotte aan dat

$$\lim_{[\alpha,\beta] \nearrow I} \int_\alpha^\beta f(x)\, dx = S. \tag{9.18}$$

Laat daartoe $\geq > 0$ gegeven zijn en zij $[\alpha, \beta]$ als in (b). Er is een $N \in \mathbb{N}$ zo dat voor alle $n \geq N$ geldt dat $[\beta, \beta] \subset J(n)$ en $|S_n - S| < \geq$. Zij J een gesloten begrend interval met $[\alpha, \beta] \subset J \subset I$. Dan geldt voor $n \geq N$ dat

$$\left| \int_J f(x)\, dx - S \right| \leq \left| \int_J f(x)\, dx - \int_{J_n} f(x)\, dx \right| + \left| \int_{J_n} f(x)\, dx - S \right| < 2 \geq .$$

Hieruit volgt inderdaad (9.18). We concluderen dat f oneigenlijk integreerbaar is over I. □

Uit het volgende lemma blijkt dat oneigenlijke integreerbaarheid over een interval van de vorm (d) uit Opmerking 9.13 herleid kan worden tot de twee reeds behandelde gevallen (b) en (c).

Lemma 9.20. *Zij* $I =]a, b[$, *met* $-\infty \leq a < b \leq \infty$ *en zij* $f : I \to \mathbb{R}$ *lokaal Riemann-integreerbaar. Laat voorts* $c \in I$. *Dan zijn de volgende twee uitspraken equivalent.*

(a) *De functie* f *is oneigenlijk Riemann-integreerbaar over* I.

(b) *De functie* f *is oneigenlijk Riemann-integreerbaar over zowel* $]a, c]$ *als* $[c, b[$.

Indien (a) en (b) gelden, dan is

$$\int_I f(x)\, dx = \int_a^c f(x)\, dx + \int_c^b f(x)\, dx. \tag{9.19}$$

Proof. We veronderstellen eerst dat (a) geldt. Zij $\geq > 0$. Dan is er wegens Lemma 9.19 een gesloten en begrensd interval $I_0 = [a_0, b_0]$ in I zo dat voor alle gesloten en begrensde intervalleen J_1, J_2 met $I_0 \subset J_j \subset I$ geldt dat

$$\left| \int_{J_1} f(x)\, dx - \int_{J_2} f(x)\, dx \right| < \geq .$$

Als we I_0 vervangen door een groter interval, dan blijft deze uitspraak geldig. We mogen daarom aannemen dat $a_0 < c < b_0$. Veronderstel nu dat twee gesloten en begrensde intervallen J_j^+ gegeven zijn met $[c, b_0] \subset J_j^+ \subset [c, b[$. Definieer $J_j = [a_0, c] \cup J_j^+$. Dan zijn J_j, voor $j = 1, 2$, gesloten en begrensde intervallen met $I_0 \subset J_j \subset I$. Bovendien geldt

$$\int_{J_j} f(x)\, dx = \int_{a_0}^c f(x)\, dx + \int_{J_j^+} f(x)\, dx.$$

Hieruit volgt dat

$$\left| \int_{J_1^+} f(x)\, dx - \int_{J_2^+} f(x)\, dx \right| = \left| \int_{J_1} f(x)\, dx - \int_{J_2} f(x)\, dx \right| < \geq .$$

We concluderen met behulp van Lemma 9.19 dat f oneigenlijk integreerbaar is over $[c, b[$. Op soortgelijke wijze zien we dat f oneigenlijk integreerbaar is over $]a, c]$. Dus (b) geldt.

Veronderstel nu dat (b) geldt. Zij $\geq > 0$. Dan is er een gesloten en begrensd interval $[c, b_0]$ zo dat voor elk gesloten en begrensd interval $[c, \beta]$ met $I_0^+ \subset [c, \beta] \subset [c, b[$ geldt dat

$$\left| \int_c^\beta f(x)\, dx - \int_c^b f(x)\, dx \right| < \frac{\geq}{2}.$$

Evenzo is er een gesloten en begrensd interval $[a_0, c] \subset]a, c]$ zo dat voor elk interval $[\alpha, c]$ met $[a_0, c] \subset [\alpha, c] \subset]a, c]$ geldt dat

$$\left| \int_\alpha^c f(x)\, dx - \int_a^c f(x)\, dx \right| < \frac{\geq}{2}.$$

Zij $I_0 = [a_0, b_0]$. En zij $[\alpha, \beta]$ zo dat $[a_0, b_0] \subset [\alpha, \beta] \subset]a, b[$. Dan geldt $[c, b_0] \subset [c, \beta] \subset [c, b[$ en $[a_0, c] \subset [a_0, c] \subset]a_0, c]$. Uit de twee bovenstaande schattingen volgt nu met behulp van de driehoeksongelijkheid dat

$$\left| \int_\alpha^\beta f(x)\, dx - \left(\int_a^c f(x)\, dx - \int_c^b f(x)\, dx \right) \right| < \geq .$$

Hieruit concluderen we met Definitie 9.8 dat f oneigenlijk integreerbaar is over $]a, b[$ en bovendien dat

$$\int_I f(x)\, dx = \int_a^c f(x)\, dx + \int_c^b f(x)\, dx.$$

□

De integraal in het linkerlid van (9.19) schrijven we in het vervolg ook als $\int_a^b f(x)\, dx$.

Definition 9.21. (Absoluut integreerbaar) Zij $I \subset \mathbb{R}$ een niet-leeg interval. Een lokaal Riemann-integreerbare functie $f : I \to \mathbb{R}$ heet absoluut oneigenlijk Riemann-integreerbaar indien de functie $|f| : x \mapsto |f(x)|, I \to \mathbb{R}$, oneigenlijk Riemann-integreerbaar is. △

Remark 9.22. Merk op dat in de bovenstaande definitie de functie $|f|$ lokaal Riemann-integreerbaar is. Voorts wordt de convergentie van de integraal $\int_I |f(x)|\, dx$ wel genoteerd met

$$\int_I |f(x)|\, dx < \infty.$$ △

Het volgende resultaat is analoog aan Lemma 8.16.

Lemma 9.23. *Laat $I \subset \mathbb{R}$ een niet-leeg interval zijn met grenzen $-\infty \leq a < b \leq \infty$, en veronderstel dat $f : I \to \mathbb{R}$ lokaal Riemann-integreerbaar is. Indien f absoluut oneigenlijk integreerbaar is, dan is f oneigenlijk integreerbaar, en*

$$\left| \int_a^b f(x)\, dx \right| \leq \int_a^b |f(x)|\, dx.$$

Proof. Zij $\geq > 0$. Dan is er een gesloten en begrensd interval $I_0 \subset I$ zo dat voor elk tweetal gesloten en begrensde intervallen J_1, J_2 met $I_0 \subset J_j \subset I$ geldt

$$\left| \int_{J_1} |f(x)|\, dx - \int_{J_2} |f(x)|\, dx \right| < \geq /2.$$

In het bijzonder volgt hieruit voor dergelijke intervallen dat

$$\int_{J_j \setminus I_0} |f(x)|\, dx = \left| \int_{J_j} |f(x)|\, dx - \int_{I_0} |f(x)|\, dx \right| < \geq /2, \qquad (j = 1, 2). \tag{9.20}$$

Strikt genomen is $J_j \setminus I_0$ de vereniging van een of twee begrensde intervallen, waarvan de afsluitingen tot I behoren. Met de integraal over $J_j \setminus I_0$ wordt de som van de Riemann-integralen over deze afsluitingen bedoeld.

Uit (9.20) leiden we af dat voor dergelijke intervallen J_1, J_2 geldt dat

$$\begin{aligned} \left| \int_{J_1} f(x)\, dx - \int_{J_2} f(x)\, dx \right| &= \left| \int_{J_1 \setminus I_0} f(x)\, dx - \int_{J_2 \setminus I_0} f(x)\, dx \right| \\ &\leq \int_{J_1 \setminus I_0} |f(x)|\, dx + \int_{J_2 \setminus I_0} |f(x)|\, dx \\ &< \ \geq /2 + \geq /2 = \geq . \end{aligned}$$

We concluderen dat $f : I \to \mathbb{R}$ voldoet aan conditie (b) van Lemma 9.19. Dus f is Riemann-integreerbaar over I. Voor alle $\alpha < \beta$ met $[\alpha, \beta] \subset I$ geldt wegens de driehoeksongelijkheid voor Riemann integralen dat

$$\left| \int_\alpha^\beta f(x)\, dx \right| \leq \int_\alpha^\beta |f(x)|\, dx.$$

Hieruit volgt (9.21) door limietovergang voor $[\alpha, \beta] \nearrow I$. □

Theorem 9.24 (Majorantiekenmerk voor integreerbaarheid)**.** *Laat $I \subset \mathbb{R}$ een niet-leeg interval zijn met grenzen $-\infty \leq a < b \leq \infty$, en veronderstel dat $f, g : I \to \mathbb{R}$ lokaal Riemann-integreerbaar zijn, $C > 0$ en dat voor alle $x \in I$ geldt:*

$$|f(x)| \leq C g(x)$$

Indien g oneigenlijk Riemann-integreerbaar is op I, dan is f dat ook, en er geldt bovendien dat

$$\left| \int_a^b f(x)\, dx \right| \leq C \int_a^b g(x)\, dx. \tag{9.21}$$

Remark 9.25. Dit resultaat kan opgevat worden als het analogon van het eerdere majorantie-kenmerk voor reeksen, zie Stelling 8.18. △

Proof. Uit de voorwaarden blijkt dat $g \geq 0$ en dat voor elk segment $[\alpha, \beta] \subset I$ geldt

$$\int_\alpha^\beta |f(x)|\, dx \leq C \int_\alpha^\beta g(x)\, dx \leq \int_I g(x)\, dx.$$

Met Lemma 9.10 volgt hieruit dat $|f|$ oneigenlijk integreerbaar is en dat

$$\int_a^b |f(x)|\, dx \leq C \int_a^b g(x)\, dx.$$

Het bewijs wordt voltooid door toepassing van Lemma 9.23. □

Example 9.26. (Gamma-functie) We beschouwen wederom de volgende integraal voor de Gamma-functie, zie ook Voorbeeld 5.18,

$$\Gamma(x) := \int_0^\infty t^{x-1} e^{-t}\, dt, \quad x > 0. \tag{9.22}$$

Als $0 < x < 1$, dan gaat de integrand naar oneindig als $t \downarrow 0$, dus dan moeten we ook bij de ondergrens $t = 0$ de integraal als een oneigenlijke integraal opvatten.

We zullen nu met behulp van het majorantie-criterium aantonen dat de integraal voor de Gamma-functie convergeert. Daartoe verdelen we het interval $]0, \infty[$ in de stukken $]0, 1]$ en $[1, \infty[$.

Voor $t \in]0, 1]$ geldt dat $|t^{x-1}e^{-t}| \leq t^{x-1}$ en $\int_0^1 t^{x-1}\, dt$ convergeert, dus ook

$$\int_0^1 t^{x-1}e^{-t}\, dt \tag{9.23}$$

convergeert.

We beschouwen nu het deel van de integraal over $[1, \infty[$. Zij $N \in \mathbb{N}, N > x - 1$. Dan geldt voor $t \geq 1$ dat $t^{x-1}e^{-t} \leq t^N e^{-t}$. Uit $\lim_{t\to\infty} t^N e^{-t/2} = 0$ volgt het bestaan van een constante $C > 0$ zo dat

$$t^N e^{-t} \leq Ce^{-t/2}, \qquad (t \geq 1).$$

Omdat de integraal $\int_1^\infty e^{-t/2}\, dt$ convergent is, concluderen we nu dat

$$\int_1^\infty t^{x-1}\, e^{-t}\, dt \tag{9.24}$$

convergent is.

Uit de convergentie van (9.23) en (9.24) concluderen we tenslotte dat de integraal (9.22) convergent is voor alle $x > 0$.

Men kan aantonen dat de Gamma-functie niet op een algebraïsche manier in termen van de bekende functies is uit te drukken. △

Example 9.27. (Bèta-functie) We beschouwen opnieuw de *Bèta-functie van Euler* uit Voorbeeld 5.17. Dit is de functie van twee reële variabelen p, q, gedefinieerd door

$$B(p,\, q) := \int_0^1 t^{p-1}\, (1-t)^{q-1}\, \mathrm{d}t. \tag{9.25}$$

Deze functie is, net als de Gamma-functie, niet op een algebraïsche manier in termen van bekende functies uit te drukken.

De gegeven integraal voor $B(p, q)$ convergeert voor $p, q > 0$. Dit is als volgt in te zien. Voor genoemde p, q is de functie

$$f : t \mapsto t^{p-1}(1-t)^{q-1}$$

continu dus lokaal Riemann-integreerbaar op het interval $]0, 1[$. We splitsen dit interval in twee delen, namelijk $]0, \frac{1}{2}]$ and $[\frac{1}{2}, 1[$ en behandelen de bijbehorende integralen afzonderlijk.

De functie $t \mapsto (1-t)^{q-1}$ is continu op $[0, \frac{1}{2}]$, dus begrensd door een constante $C > 0$. Voor $0 < t \leq \frac{1}{2}$ geldt daarom dat

$$|f(t)| \leq Ct^{p-1}.$$

De functie in het rechterlid van deze uitdrukking is oneigenlijk Riemann-integreerbaar over $]0, \frac{1}{2}]$ wegens Lemma 9.18. Hieruit volgt de convergentie van de integraal van f over $]0, \frac{1}{2}]$. De functie $t \mapsto t^{p-1}$ is continu op $[\frac{1}{2}, 1]$ dus begrensd door een constante $C' > 0$. Voor $\frac{1}{2} \leq t < 1$ geldt daarom dat

$$|f(t)| \leq C'(1-t)^{q-1}.$$

De functie in het rechterlid van deze uitdrukking is oneigenlijk Riemann-integreerbaar over $[\frac{1}{2}, 1[$, wegens Lemma 9.18 (pas de substitutieregel toe om dit in te zien). We concluderen dat f oneigenlijk integreerbaar is over $]\frac{1}{2}, 1]$. △

Uit het majorantiekenmerk voor de convergentie van oneigenlijke integralen volgt het eveneens gemakkelijk hanteerbare limietkenmerk.

Corollary 9.28 (Limietkenmerk voor integreerbaarheid). *Laat I een interval van de vorm $[c, b[$ zijn, met $-\infty < c < b \leq \infty$. Veronderstel voorts dat $f, g : I \to \mathbb{R}$ lokaal Riemann-integreerbare functies zijn, terwijl $g > 0$ op I en*

$$\lim_{x \uparrow b} \frac{|f(x)|}{g(x)} = L \in [0, \infty[.$$

Als g oneigenlijk integreerbaar is op I, dan is f dat ook.

Proof. Er bestaat een $\beta > 0$ zo dat $||f(x)|/g(x) - L| < 1$ voor alle $x \in [\beta, b[$. Hieruit volgt dat $|f(x)| \leq (L+1)g(x)$ voor al dergelijke x. De functie $(L+1)g(x)$ is oneigenlijk integreerbaar over I, dus ook over $[\beta, b[$, en wegens het majorantiekenmerk volgt dat f oneigenlijk integreerbaar is over $[\beta, b[$. Hieruit volgt dat f oneigenlijk integreerbaar is over I. □

Remark 9.29. Uiteraard geldt een soortgelijk limietkenmerk voor lokaal integreerbare functies op een interval van de vorm $I =]a, c]$, met $-\infty \leq a < c < \infty$. △

Ook voor oneigenlijke integralen geldt een verwisselingsstelling met limieten. We bewijzen eerst twee technische resultaten. Daaruit leiden we dan een dominantie-kenmerk af dat in de praktijk vaak goed werkt.

Lemma 9.30. *Laat I een niet-leeg interval zijn met grenzen $-\infty \leq a < b \leq \infty$. Laat $V \subset \mathbb{R}^n$ zijn en $f : V \times I \to \mathbb{R}$ een continue functie. Veronderstel verder dat de volgene voorwaarden vervuld zijn.*

(a) *Voor elke $x \in V$ is de functie $t \mapsto f(x, t)$ oneigenlijk integreerbaar over I.*

(b) *Voor iedere $\geq> 0$ bestaat er een gesloten en begrensd interval $[\alpha, \beta] \subset I$ zo dat voor alle $x \in V$ geldt dat:*

$$\left| \int_a^b f(x,t)\, dt - \int_\alpha^\beta f(x,t)\, dt \right| <\geq \tag{9.26}$$

Dan is de functie $F : V \to \mathbb{R}$ gedefinieerd door

$$F(x) = \int_a^b f(x,t)\, dt$$

continu.

Proof. Laat $x_0 \in V$. Dan is het voldoende de continuïteit van F in het punt x_0 aan te tonen. Voor $\alpha, \beta \in I$ met $\alpha < \beta$ definiëren we

$$F_\alpha^\beta : x \mapsto \int_\alpha^\beta f(x,t)\, dt.$$

Zij nu $\geq> 0$, dan volgt uit de hypothese dat er $\alpha, \beta \in I$ bestaan met $\alpha < \beta$, zo dat

$$|F(x) - F_\alpha^\beta(x)| <\geq /3,$$

voor alle $x \in V$. Uit Stelling 5.12 volgt dat de functie F_α^β continu is op V, dus in het bijzonder in x_0. Er bestaat dus een $\delta > 0$ zo dat voor alle $x \in B(x_0; \delta)$ geldt dat

$$|F_\alpha^\beta(x) - F_\alpha^\beta(x_0)| <\geq /3.$$

We merken nu op dat voor alle $x \in B(x_0; \delta)$ geldt dat

$$\begin{aligned} |F(x) - F(x_0)| &\leq |F(x) - F_\alpha^\beta(x)| + |F_\alpha^\beta(x) - F_\alpha^\beta(x_0)| + |F_\alpha^\beta(x_0) - F(x_0)| \\ &< \geq /3 + \geq /3 + \geq /3 = \geq . \end{aligned}$$

Hiermee is de continuïteit van F in x_0 aangetoond. □

Ook het volgende lemma zal nuttig blijken. Is I een niet-leeg interval, en $c \in I$ dan definiëren we de volgende deelintervallen van I,

$$I_{\leq c} := \{x \in I \mid x \leq c\}, \quad \text{en} \quad I_{\geq c} := \{x \in I \mid x \geq c\}.$$

Lemma 9.31. *Zij I een niet-leeg interval en $a, b \in I$ met $a < b$. Dan geldt voor elke oneigenlijk Riemann-integreerbare functie $f : I \to \mathbb{R}$ dat f oneigenlijk integreerbaar is over $I_{\leq a}$ en over $I_{\geq b}$, terwijl*

$$\int_I f(x)\, dx = \int_{I_{\leq a}} f(x)\, dx + \int_a^b f(x)\, dx + \int_{I_{\geq b}} f(x)\, dx.$$

Proof. Door toepassen van Lemma 9.20 en Gevolg 9.15 vinden we dat

$$\begin{aligned} \int_I f(x)\, dx &= \int_{I_{\leq a}} f(x)\, dx + \int_{I_{\geq a}} f(x)\, dx \\ &= \int_{I_{\leq a}} f(x)\, dx + \int_a^b f(x)\, dx + \int_{I_{\geq b}} f(x)\, dx. \end{aligned}$$

□

Uit het bovenstaande leiden we het volgende practisch goed toepasbare principe van gedomineerde continuïteit af.

Theorem 9.32 (Gedomineerde continuïteit)**.** *Laat $I \subset \mathbb{R}$ een niet-leeg interval zijn met grenzen $-\infty \leq a < b \leq \infty$. Zij $V \subset \mathbb{R}^n$ en $f : V \times I \to \mathbb{R}$ een continue functie. Veronderstel verder dat er een oneigenlijk Riemann-integreerbare functie $g : I \to \mathbb{R}$ bestaat zo dat*

$$|f(x,t)| \leq g(t) \qquad \text{voor alle} \quad (x,t) \in V \times I.$$

Dan is de functie $F : V \to \mathbb{R}$ gedefinieerd door

$$F(x) = \int_a^b f(x,t)\, dt$$

continu.

Proof. We zullen laten zien dat de voorwaarden van Lemma 9.30 vervuld zijn. Zij $x \in V$. Dan is de functie $f_x : t \mapsto f(x,t)$, $I \to \mathbb{R}$ continu, dus lokaal Riemann-integreerbaar, terwijl $|f_x| \leq g$ op I. Dus f_x is oneigenlijk integreerbaar wegens Stelling 9.24. Hiermee is voorwaarde (a) aangetoond. Zij $\geq > 0$ en zij $c \in I$. Uit de oneigenlijke Riemann-integreerbaarheid van g volgt het bestaan van $\alpha, \beta \in I$ met $\alpha < \beta$ zo dat

$$\left| \int_I g(t)\, dx - \int_\alpha^\beta g(t)\, dt \right| < \geq .$$

Hieruit volgt voor alle $x \in V$ dat

$$\begin{aligned}
\left|\int_I f(x,t)\, dx - \int_\alpha^\beta f(x,t)\, dt\right| &= \left|\int_{I_{\leq\alpha}} f(x,t)\, dt + \int_{I_{\geq\beta}} f(x,t)\, dt\right| \\
&\leq \int_{I_{\leq\alpha}} |f(x,t)|\, dt + \int_{I_{\geq\beta}} |f(x,t)|\, dt \\
&\leq \int_{I_{\leq\alpha}} g(t)\, dx + \int_{I_{\geq\beta}} g(t)\, dt \\
&= \int_I g(t)\, dt - \int_\alpha^\beta g(t)\, dt <\geq .
\end{aligned}$$

Hieruit volgt de ongelijkheid (9.26) waaruit blijkt dat voorwaarde (b) van Lemma 9.30 vervuld is. □

Remark 9.33. Het idee van de voorwaarde in Stelling 9.32 is dat $t \mapsto f(x,t)$ gedomineerd wordt door de oneigenlijk integreerbare (niet-negatieve) functie $t \mapsto g(t)$, met uniformiteit in de parameter $x \in V$. Dit dwingt de voorwaarden van Lemma 9.30 af. △

Example 9.34. **(Gamma-functie)** We passen het bovenstaande toe op de Gamma-functie

$$\Gamma(x) = \int_0^\infty t^{x-1}\, e^{-t}\, dt, \qquad (x > 0).$$

Zij $0 < a < b$ en $X =]a,b[$. Dan geldt voor alle $t \in]0,1]$ dat $t^{x-1} = e^{(x-1)\log t} \leq t^{a-1}$. De functie $f(x,t) = t^{x-1}e^{-t}$ is continu op $]a,b[\times]0,1]$ en voor alle $(x,t) \in X\times]0,1]$ geldt dat $|f(x,t)| \leq g(t) := t^{a-1}e^{-t}$, terwijl g oneigenlijk integreerbaar is, dus

$$F_0 : x \mapsto \int_0^1 t^{x-1}\, e^{-t}\, dt$$

definieert een continue functie op X.

Anderzijds is f ook continu op $]a,b[\times[1,\infty[$, terwijl op deze verzameling een majorantie van de vorm $|f(t,x)| \leq t^{b-1}e^{-t}$ bestaat. De laatste functie is weer oneigenlijk integreerbaar op $]1,\infty[$, dus

$$F_1 : x \mapsto \int_1^\infty t^{x-1}\, e^{-t}\, dt$$

definieert een continue functie op $]a,b[$. Hieruit volgt dat $\Gamma = F_0 + F_1$ continu is op $]a,b[$. Aangezien a, b willekeurig waren volgt dat Γ continu is op $]0,\infty[$. △

Remark 9.35. We merken op dat

$$\Gamma(1) = \int_0^\infty e^{-t}\, dt = \lim_{R\to\infty} \left[-e^{-t}\right]_0^R = 1.$$

Zij $x > 0$, dan volgt uit het bovenstaande dat

$$\Gamma(x+1) = \lim_{R\to\infty} \int_0^R t^x\, e^{-t}\, dt.$$

De integraal is met behulp van partiële integratie als volgt te herschrijven:

$$\begin{aligned}
\int_0^R t^x\, e^{-t}\, dt &= -\int_0^R t^x \frac{d}{dt} e^{-t}\, dt \\
&= \left[-t^x\, e^{-t}\right]_0^R + x\int_0^R t^{x-1}\, e^{-t}\, dt.
\end{aligned}$$

De laatste integraal is convergent. Omdat $x > 0$ is, geldt $t^x|_{t=0} = 0$. Tevens geldt $R^x\, e^{-R} \to 0$ voor $R \to \infty$. Door de limiet voor $R \to \infty$ te nemen concluderen we daarom dat

$$\Gamma(x+1) = x\int_0^\infty t^{x-1}\, e^{-t}\, dt$$

dus

$$\Gamma(x+1) = x\Gamma(x), \quad (x > 0).$$

Passen we dit toe met $x = n-1$, $n \in \mathbb{Z}_+$, dan vinden we met inductie dat

$$\Gamma(n) = (n-1)!\,\Gamma(1) = (n-1)!.$$

Anders gezegd, de Gamma-functie $x \mapsto \Gamma(x)$ levert een continue uitbreiding tot de positieve reële as van de faculteitsfunctie $n \mapsto (n-1)!$, waarbij de laatste functie alleen voor de gehele positieve getallen n is gedefinieerd. △

Example 9.36. **(Béta-functie)** We passen het bovenstaande toe op de Bèta-functie

$$B(p,q) := \int_0^1 t^{p-1}(1-t)^{q-1}\, dt \tag{9.27}$$

De integrand is continu als functie van (p,q,t), voor $p,q > 0$ en $0 < t < 1$. Fixeer $p_0, q_0 > 0$. Dan geldt voor alle $p \geq p_0$, $q \geq q_0$ en $t \in \,]0,1[$ dat

$$0 \leq t^{p-1}(1-t)^{q-1} \leq t^{p_0-1}(1-t)^{q_0-1}.$$

Zoals we eerder in Voorbeeld 9.27 zagen is de functie in het rechterlid oneigenlijk integreerbaar over $]\,0,1[$. Met Stelling 9.32 concluderen we nu dat B continu is op $[p_0,\infty[\,\times\,[q_0,\infty[$. Dit geldt voor iedere $p_0, q_0 > 0$. Dus B is continu op de verzameling $]\,0,\infty[\,\times\,]\,0,\infty[$. △

We zien aan deze voorbeelden dat uniforme majorantie vaak gemakkelijker is toe te passen na splitsing van de oneigenlijke integratie in integraties over intervallen die minstens een der eindpunten bevatten, zodat men zich alleen op het gedrag van de integrand naar het overgebleven eindpunt hoeft te concentreren.

Er is ook een versie van differentiatie onder het integraalteken voor oneigenlijke integralen. Ook dit gaat weer in termen van een geschikte uniforme dominantie.

Theorem 9.37. **(Differentiatie onder het integraalteken)** *Zij $X \subset \mathbb{R}$ een open interval en I een niet-leeg interval met grenzen $-\infty \leq a < b \leq \infty$. Zij verder $f : X \times I \to \mathbb{R}$ een continue functie die voldoet aan de volgende eigenschappen.*

(a) *voor alle $x \in X$ is de functie $f_x : t \mapsto f(x,t)$ oneigenlijk Riemann-integreerbaar over I;*

(b) *de functie f is partieel differentieerbaar naar de eerste variabele, $D_1 f$ is continu op $X \times I$ en er is een oneigenlijk Riemann-integreerbare functie $g : I \to \mathbb{R}$ zo dat*

$$|D_1 f(x,t)| \leq g(t) \qquad \text{voor alle} \quad (x,t) \in X \times I.$$

Dan is de functie $F : X \to \mathbb{R}$ gedefinieerd door

$$F(x) = \int_a^b f(x,t)\, dt$$

(continu) differentieerbaar op X en er geldt dat

$$F'(x) = \int_a^b D_1 f(x,t)\, dt. \tag{9.28}$$

Proof. Zij $x_0 \in X$. We zullen de differentieerbaarheid van F in x_0 aantonen. Hiertoe definiëren we de functie $q : X \times I \to \mathbb{R}$ door

$$q(x,t) = \frac{f(x,t) - f(x_0,t)}{x - x_0}, \qquad (x \in X \setminus \{x_0\}, \quad t \in I),$$

en

$$q(x_0,t) = D_1 f(x_0,t), \qquad (t \in I).$$

De functie q is continu op $X \times I$ wegens Lemma 5.20. We zullen laten zien dat voor alle $x \in X$ en $t \in I$ geldt dat

$$|q(x,t)| \leq g(t). \tag{9.29}$$

Voor $x = x_0$ volgt dit uit de voorwaarde (b). Laat $(x,t) \in (X \setminus \{x_0\}) \times I$. Dan geldt vanwege de middelwaardestelling toegepast op de eerste variabele van f dat er een tussen x_0 en x gelegen $\xi = \xi(x,t)$ bestaat zo dat $q(x,t) = D_1 f(\xi,t)$. De schatting (9.29) volgt nu ook uit voorwaarde (b).

Wegens het majorantiekenmerk is de functie $q : t \mapsto q(x,t)$ oneigenlijk Riemann-integreerbaar over I, voor elke $x \in X$. Wegens Stelling 9.32 is de functie $Q : X \to \mathbb{R}$ gedefinieerd door

$$Q(x) = \int_a^b q(x,t)\, dt$$

continu op X, dus in het bijzonder in x_0. Uit de definities volgt direct dat

$$F(x) - F(x_0) = Q(x)(x - x_0)$$

voor alle $x \in X \setminus \{x_0\}$. En uiteraard is de bewering ook geldig voor $x = x_0$. Omdat Q continu is in x_0 leiden we hieruit af dat F differentieerbaar is in x_0, en dat de afgeleide gegeven wordt door

$$F'(x_0) = Q(x_0) = \int_a^b D_1 f(x_0,t)\, dt.$$

Hieruit volgt dat F differentieerbaar is op X. Uit de formule (9.28) volgt door toepassing van Stelling 9.32 dat de afgeleide continu is. □

Example 9.38. (Gamma-functie) We tonen aan dat de Gamma-functie willekeurig vaak differentieerbaar is op $]\,0,\infty\,[$, terwijl

$$\Gamma^{(k)}(x) = \int_0^\infty (\log t)^k\, t^{x-1}\, e^{-t}\, dt, \qquad (k \in \mathbb{N},\ x > 0).$$

De Gamma-functie is daarmee een gladde uitbreiding tot de positieve reële as van de faculteitsfunctie $(n-1)!$, $n \in \mathbb{Z}_{>0}$. We schrijven $f_k(x,t)$ voor de integrand.

Zij $\geq > 0$ willekeurig. Dan is

$$\lim_{t \downarrow 0} \ (\log t)^k\, t^{\geq} = 0,$$

dus er bestaat een constante $C_\geq > 0$ zo dat $|\log t|^k \leq C_\geq t^{-\geq}$ voor alle $t \in]\,0,1]$. Dit geeft een schatting van het type

$$|f_k(x,t)| \leq C_\geq t^{x-1-\geq}, \qquad (0 < t \leq 1).$$

Hierbij kunnen we $\geq > 0$ kiezen met $\geq < x$, zodat de dominerende functie $t \mapsto C_\geq t^{x-1-\geq}$ oneigenlijk integreerbaar is op het interval $]\,0,1]$. Hieruit volgt de convergentie van $\int_0^1 f_k(x,t)\, dt$.

Voor de integratie over $[1,\infty\,[$ merken we op dat

$$\lim_{t \to \infty} (\log t)^k t^N e^{-t/2} = 0$$

voor alle $k, N \in \mathbb{N}$. Hieruit volgt dat er een $C_k > 0$ bestaat zo dat

$$|f_k(x,t)| \leq C_k e^{-t/2} \qquad (t \geq 1).$$

Hieruit volgt de convergentie van $\int_1^\infty f_k(x,t)\, dt$.

Laat nu $0 < a < b$ zijn, en veronderstel dat $k \in \mathbb{N}$. Dan geldt voor alle $x \in\,]\,a, b\,[$ dat

$$|f_k(x,t)| \leq |f_k(a,t)|, \qquad (0 < t \leq 1),$$

en dat

$$|f_k(x,t)| \leq |f_k(b,t)|, \qquad (t \geq 1).$$

Voor alle $k \in \mathbb{N}, x > 0, t > 0$ geldt dat

$$\frac{\partial}{\partial x} f_k(x,t) = f_{k+1}(x,t).$$

Het resultaat volgt nu met inductie naar k, door toepassing van Stelling 9.37. △

Example 9.39. **(Bèta-functie)** We beschouwen nogmaals de Bèta-functie van Euler, zie (9.27), waarvoor we nu de sterkere uitspraak zullen bewijzen dat hij willekeurig vaak differentieerbaar is op $]0,\infty[\times]0,\infty[$ terwijl voor alle $k, l \in \mathbb{Z}_{\geq 0}$ geldt dat

$$\frac{\partial^{k+l} B(p,\,q)}{\partial p^k\, \partial q^l} = \int_0^1 (\log t)^k\, t^{p-1}\, (\log(1-t))^l\, (1-t)^{q-1}\, \mathrm{d}t. \tag{9.30}$$

De continuïteit, van de integrand als functie van $(p,\,q,\,t) \in]1,\,\infty[\times]1,\,\infty[\times]0,\,1[$ is evident. Als functie van t is de integrand dus lokaal Riemann integreerbaar op $]0,1[$. Zij nu $p_0, q_0 > 0$. Dan geldt voor $p > 2p_0$ en $q > 2q_0$ dat

$$|(\log t)^k\, t^{p-1}\, (\log(1-t))^l\, (1-t)^{q-1}| \leq \psi(t)\, t^{p_0-1} t^{q_0-1} \tag{9.31}$$

met

$$\psi(t) := (\log t)^k\, t^{p_0}\, (\log(1-t))^l\, (1-t)^{q_0}.$$

Deze functie is continu voortzetbaar tot $[0,1]$, omdat

$$\lim_{t\downarrow 0} \psi(t) = 0 \quad \text{en} \quad \lim_{t\uparrow 1} \psi(t) = 0.$$

Hieruit volgt dat er een $M > 0$ bestaat zo dat $|\psi(t)| \leq M$ voor alle $0 < t < 1$. We concluderen dat de functie in het rechterlid van (9.31) op $]\,0,1\,[$ gemajoreerd kan worden door de functie

$$t \mapsto M t^{p_0-1}(1-t)^{q_0-1},$$

die absoluut convergent is op $]\,0,1\,[$, wegens Voorbeeld 9.36.

Door herhaald Stelling 9.37 toe te passen op de variabelen p en q concluderen we dat de functie B willekeurig vaak differentieerbaar is op $]2p_0,\infty[\times]2q_0,\infty[$, met partiële afgeleiden die gegeven worden door (9.30). Aangezien dit geldt voor alle $p_0, q_0 > 0$ zien we dat B willekeurig vaak differentieerbaar is op $]0,\infty[\times]0,\infty[$ met de gegeven partiële afgeleiden. △

Index