ANALYSIS I AND II
TU DARMSTADT, 2006/07

KARSTEN GROSSE-BRAUCKMANN

Contents

References

The following textbooks cover the material of this class:

[K] Königsberger: Analysis I+II, Springer (the approach is similar to ours)

[F] O. Forster: Analysis I+II, Vieweg 1976, 1999 (very concise)

[F3] O. Forster: Analysis III, Vieweg

[Sp] M. Spivak: Calculus, Benjamin, New York 1967 (contains only one-variable calculus – is fun to read)

[Hi] S. Hildebrandt: Analysis I+II, Springer 2003 (a broad presentation, with an interesting introduction of the exponential and trigonometric functions via their differential equations)

[CJ] Courant, John: Introduction to calculus and analysis, Springer reprint (a classic, written by experts with lots of experience)

[AE] Amann/Escher: Analysis I+II

[J] K. Jänich: Mathematik 1+2, Geschrieben für Physiker, Springer 2001/02 (Geared towards physicists, this is definitely worth a look for mathematicians, too: Chapter 1 to 17 contain all you need to know about first year Analysis and Linear algebra – without any formalities! Only thereafter the technical exposition starts and goes as far as the Stokes theorem.)

Further reading on selected topics:

[D] O. Deiser: Reelle Zahlen, Springer 2007 (A well-written book which contains all you want to know about real numbers, and much more! Worth reading is in particular the concise 10 page account on the history of analysis, p.61–70.)

[E] Ebbinghaus, Hermes, Hirzebruch, Koecher, Mainzer, Neukirch, Prestel, Remmert: Numbers, Springer 1991 (Beware: This book is written for an advanced readership. Chapter 1-3 cover our chapter on numbers)

[GKP] R. Graham, D. Knuth, O. Patashnik: Concrete Mathematics, A foundation for computer science, Addison-Wesley 1989,1994. (D. Knuth, one of the leading computer scientists, provides lots if interesting material.)

[J] G. A. Jennings: Modern Geometry with applications, Springer 1994 (The book explains Euclidean, spherical, and projective geometry. Recommended in particular for teacher's degree students.)

[Sa] H. Sagan: Space-filling curves, Springer 93 (Contains beautiful examples of curves and mappings with unexpected properties, both in illustration and theory.)

Introduction

These lecture notes cover the first year analysis of a class for students of mathematics and physics, taught in 2006/07. They are a revision of lectures notes from Bonn in 2000/01 and Darmstadt in 2002/03. For this reason, I kept the English language from the previous version although classes were hold in German.

There are two main ideas of first year university analysis, reaching beyond high school calculus. First, limits are always quantified. They appear as error bounds or remainder term estimates. This defines limits of sequences and series, as well as continuity, differentiability or the (Riemann) integral of functions. Second, differentiation is understood as linear approximation. The first fact makes inequalities or estimates a main tool of analysis. The second fact brings linear algebra into the game of multidimensional calculus.

The presentation is always rigorous (I hope!). While ideas follow intuition, the reasoning is strict and makes only use of the explicit stated assumptions. To learn such a rigorous reasoning certainly is the most important goal of the course. Let us now describe the contents.

We start with a chapter on numbers. We introduce real and complex numbers by their properties which allow us to do calculus: they are complete fields, and $\mathbb{R}$ is ordered while $\mathbb{C}$ has a modulus.

The next Chapter covers sequences and series, in particular Cauchy sequences. The existence of the real numbers in terms of Cauchy sequences of rational numbers is sketched, but we do not explain the real numbers as equivalence classes of Cauchy sequences in depth.

In the third Chapter, we discuss continuous functions and introduce the exponential and trigonometric functions. Continuity is introduced via the limit test, and augmented with the ε-δ test. The Weierstrass theorem on the maximum then follows. The chapter is concluded with a section on the trigonometric functions, introduced by their power series, the number π, and a discussion of the complex logarithm.

The fourth chapter deals with differentiation and integration in one variable. The derivative is introduced in terms of the difference quotient and characterized as a linear approximation. The Riemann integral follows, along with a section on the fundamental theorem. Finally, we discuss power series and Taylor's theorem.

The second term started with the fifth chapter, on multi-dimensional space $\mathbb{R}^n$. We introduce the notions of openness, closedness, and compactness in the generality of metric spaces.

In the sixth chapter we introduce differentiability in several variables as linear approximability. Then we disucss partial and directional derivatives. The discussion of extrema is based on the Taylor series.

The seventh chapter contains the three big theorems of differentiation in several variables: The inverse mapping theorem, the implicit mapping theorem, and the theorem on extrema under constraints. Our perspective is from the point of view of solving equations. We describe the tangent and normal space to an implicitely defined submanifold. As a direct conxequence we obtain the theorem on extrema under constraints.

The final short chapter introduces multi-dimensional integration. For mathematics students, integration is covered more thoroughly in the fourth term. Thus the goal here is to present the main properties of the integral: These are Fubini's theorem and the change of variables formula. It goes without saying that the space of integrable functions introduced, namely continuous functions with compact support, is too narrow for most applications.

Whenever feasible, the single variable exposition applies to the complex case: Sequences, series, functions, differentiability, and the Riemann integral are discussed in their complex version. This is to some extent necessary anyway: For instance the complex exponential function must be introduced. In other cases it simplifies, for instance when differentiating and integrating the trigonometric functions.

My main goal in these notes has been to introduce abstraction carefully, always accompanied with examples. Thus new concepts are often only introduced alongside with their first application. To give but one example, uniform continuity is not dealt with in the chapter on continuity; instead it is covered in the section on integration. At this place it is employed to verify that continuous functions are integrable.

Some of the material is improved in comparison to the presentation in class. Material not presented in class appears in small print. These notes contain an index which may be useful to check at which places a theorem gets eventually applied. I thank all students whose comments helped me to improve these notes.

Darmstadt, July 07 Karsten Grosse-Brauckmann

Part 1. Numbers

1. Vorlesung, Dienstag, 17.10.06

We will review the number system: natural, integer, rational, real, and complex numbers. These number spaces can be introduced systematically so that each space arises as an equivalence class of products of the previous space. This leaves only the natural numbers without foundation. The process is described in detail in the book *Numbers* by Ebbinghaus and others, together with a beautiful account of the historical development.

Our goal here is more modest: We will merely illustrate what the main problems are. Thereby, we focus on the natural and the real numbers. These are the most important spaces, and the derivation *[Herleitung]* of the other number spaces from these is not hard.

0.1. **Mappings.** For a *set [Menge]* X we use the following notation. We write $x \in X$ if x is an element of X. If X, Y are sets, then we can take

$$X \cap Y := \{a : a \in X \text{ and } a \in Y\}, \qquad X \cup Y := \{a : a \in X \text{ or } a \in Y\},$$
$$X \setminus Y := \{a : a \in X \text{ but not } a \in Y\}.$$

For the latter we do not assume that $Y \subset X$. Moreover, we write $X \subset Y$ if all elements of X are also elements of Y; this is true, in particular, in case $X = Y$. Note that instead sometimes the symbol $X \subseteq Y$ is used to denote containment, while $X \subset Y$ is reserved for strict containment ($X \neq Y$).

We also asssume the notion of a *mapping* or *map [Abbildung]* f from a set X to a set Y, denoted

$$f\colon X \to Y, \qquad x \mapsto f(x).$$

It means that to each element $x \in X$, we assign an element $f(x) \in Y$. Maps between number spaces are usually called *functions.* If $f(x) = y$ we call y the *image [Bild]* of x, whereas x is called a *preimage [Urbild]* of y.

A mapping f is *injective* when $f(x_1) = f(x_2)$ implies $x_1 = x_2$, that is, each element $y \in Y$ has at most one preimage. The mapping is *surjective* if each $y \in Y$ has a preimage $x \in X$, that is, for each element $y \in Y$ there is at least one preimage. Finally, a mapping is *bijective* if it is injective and surjective.

Examples. 1. $f\colon \mathbb{R} \to \mathbb{R}$, $f(x) := x^3$ is bijective.
2. $f\colon \mathbb{R} \to \mathbb{R}$, $f(x) := x^2$ is neither injective (since $3^2 = (-3)^2$) nor surjective (-1 is not attained *[angenommen]*).

When we have two mappings,

$$f\colon X \to Y, \qquad g\colon Y \to Z$$

$\circ$we can *compose* [*verknüpfen*] them to "g after f" [*"g nach f"*]

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

Example. $\sqrt{x+4}$ is the composition of the root after the map which adds 4.

1. The natural numbers $\mathbb{N}$

1.1. **The Peano axioms for the natural numbers.** The following properties of the natural numbers are familiar:

- We can enumerate them as $1, 2, 3, \dots$ and
- we can add and multiply (but not always subtract and divide).

In the following subsections we will see that the first property is basic, while the second property can be derived from the first.

Every child learns the natural numbers quickly. Nevertheless, our first property cannot be derived from anything else and hence [*daher*] must be postulated. Mathematicians use the name *axiom* for such postulates:

Axiom 1 (Peano 1889). *There is a set $\mathbb{N}$, called the* natural numbers [*natürliche Zahlen*] *which contains a distinguished element* 1, *together with a successor map* $s\colon \mathbb{N} \to \mathbb{N}$ *satisfying:*
(P1) s is injective,
(P2) $1 \notin s(\mathbb{N})$,
(P3) If a subset $M \subset \mathbb{N}$ contains 1 and is mapped into itself by s (i.e., $s(M) \subset M$) then $M = \mathbb{N}$.

In the language of cardinality (see Sec. 4 below) (P1) and (P2) say that $\mathbb{N}$ is an infinite set, while (P3) says that $\mathbb{N}$ is countable.

Remarks. 1. Which property fails for the positive reals $\{x \in \mathbb{R} : x > 0\}$ with $s(x) := x + 1$?
2. Which property fails for the three-element-set $\big\{P = s(R), Q = s(P), R = s(Q)\big\}$?

So formally the count of the natural numbers is

$$1, \quad s(1), \quad s\big(s(1)\big), \quad s\big(s(s(1))\big), \quad \dots .$$

Instead of these clumsy expressions we will use the common notation

$$\mathbb{N} = \{1, 2, 3, \dots\}$$

where $2 := s(1)$, $3 := s(2) = s\big(s(1)\big)$, etc. Saying "notation" here means that it does *not* matter what we actually use to abreviate. Equally fine for our purposes would be roman notation, $I := 1$, $II := s(I)$, $III := s(II)$, $IV := s(III)$ $\dots$, or binary notation $1, 10, 11, 100, \dots$, etc. The natural numbers are unique up to such relabelling.

1.2. **The principle of induction.** Formally, induction does the following:

Theorem 2. *Let $E(n)$ be a property formulated for each natural number $n \in \mathbb{N}$. Suppose*
(i) (base case [Induktionsbeginn]) $E(1)$ *is true, and*
(ii) (induction step [Induktionsschritt]) *For all $n \in \mathbb{N}$ holds: If $E(n)$ is true then $E(n+1)$ is true.*
Then $E(n)$ is true for all natural numbers $n \in \mathbb{N}$.

Proof. Let $M \subset \mathbb{N}$ be the subset of natural numbers for which $E(n)$ holds. By (i), M contains 1. The step (ii) extends the definition from M to $s(M) \subset M$. Thus by (P3) we have $M = \mathbb{N}$. □

Let us give an example (it uses addition, to be defined only below):

Proposition 3. *For each $n \in \mathbb{N}$*

$$1 + 2 + \ldots + n = \frac{n(n+1)}{2}. \tag{1}$$

Proof. Denote with $E(n)$ the fact that formula (1) holds for $n \in \mathbb{N}$.

We need to show the base case $E(1)$. Clearly, $1 = \frac{1 \cdot 2}{2}$ holds.

Induction step: The assumption $E(n)$ is $1 + 2 + \ldots + n = \frac{n(n+1)}{2}$. We deduce $E(n+1)$:

$$\begin{aligned} 1 + 2 + \ldots + (n{+}1) &= (1 + 2 + \ldots + n) + (n+1) \\ &\overset{\text{by } E(n)}{=} \frac{n(n+1)}{2} + (n+1) = \frac{(n+2)(n+1)}{2}. \end{aligned}$$

□

Using the sum-sign (to be defined below), we can rewrite (1) as

$$\sum_{k=1}^{n} k = \frac{n(n+1)}{2}.$$

A nice humouros application of induction is the *interesting number paradox*, see `en.wikipedia.org/wiki/`

1.3. **Recursive definitions.** Let us motivate first by an example:

For each $n \in \mathbb{N}$ we can define the *factorial [Fakultät]* in two ways. First we could write:

$$n! := 1 \cdot \ldots \cdot (n-1) \cdot n$$

To programme this formula on a computer, one would instead employ a recursive definition:

$$1! := 1, \qquad (n+1)! := (n+1)\, n!$$

To be more formal, let us say that a *recursive definition* is the definition of a quantity $E(n)$ in terms of

- an *initialization* $E(1)$, and
- a *step*, that is a definition of $E(n+1)$ in terms of $E(n)$.

To see this defines $E(n)$ for all $n \in \mathbb{N}$, only replace the word "true" by "defined" in Thm. 2 and its proof. When a recursive definition is programmed on a computer, the step becomes a loop, which must be executed n times in order to compute $E(n)$ from $E(0)$.

Whenever we use *dots* in a mathematical definition, this is to be understood as a recursive definition. Recursive definitions are abundant in mathematics:

Examples. 1. Addition and multiplication of natural numbers. For instance, multiplication has a recursive definition in terms of addition:

$$a \cdot 1 := a \quad \text{and} \quad a \cdot (n+1) := a \cdot n + a \quad \text{for } a, n \in \mathbb{N}$$

2. The sum sign also has a recursive definition (state it!), which in terms of dots reads

$$\sum_{k=1}^{n} a_k := a_1 + \ldots + a_n.$$

3. For $n \in \mathbb{N}$ let a^n denote the n-th *power [Potenz]* $a \cdot \ldots \cdot a$ (n factors) of a. State the recursive definition! For now a must be natural, but lateron we will use the same notation with a real or complex.

1.4. **Some combinatorics.** We extend the natural numbers with zero to

$$\mathbb{N}_0 := \{0\} \cup \mathbb{N} = \{0, 1, 2, \ldots\}.$$

The following convention will be useful. Whenever we consider a product over zero factors, we let it be 1 by definition. For instance, we set $0! := 1$, and $a^0 := 1$ for each $a \in \mathbb{N}$ (or $\mathbb{R}$).

Let a set X of $n \in \mathbb{N}_0$ elements be given. For instance, the elements could be n distinguishable balls in a container. Now we consider various ways of picking elements (balls).

1. We pick $k \in \mathbb{N}_0$ balls from the container, returning each one before picking the next. Thus each time we pick there are n choices; this gives altogether n^k different results. This is the number of elements in $X^k = X \times \ldots \times X$ (k factors). For instance, there are 10^k different natural numbers with at most k digits *[Ziffern]*.

`2. Vorlesung, Donnerstag, 19.10.06` ________________________________

2. (*Permutations*) We pick $0 \leq k \leq n$ balls in order *[Reihenfolge]*, this time without returning them. There is

$$\underbrace{n \cdot (n-1) \cdots (n-k+1)}_{k \text{ factors}} = \frac{n!}{(n-k)!}$$

different results. The particular case $k = n$ corresponds to the $n!$ different ways of ordering a given set of n elements.

3. We pick $0 \leq k \leq n$ balls, without putting them back. We want to know the number of different results irrespective of order, that is, we ask: How many different sub*sets* does X have? We have seen there are $n \cdot (n-1) \cdots (n-k+1)$ *ordered* sets of k (different) elements. But exactly $k!$ of them must consist of the same k elements, since by (*ii*) we know that k elements can be ordered in $k!$ different ways. So the number of unordered sets of k elements must be

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

Here the *binomial coefficient [Binomialkoeffizient]* $\binom{n}{k}$, "n choose k" *["n über k"]*, is defined for $0 \leq k \leq n$. From the definition it is immediate that $\binom{n}{k} = \binom{n}{n-k}$. It follows from a divisibility consideration or from Lemma 4 below that $\binom{n}{k} \in \mathbb{N}$.

The following fact about binomial coefficients will be proven in the problems:

Lemma 4. *We have* $\binom{n+1}{k+1} = \binom{n}{k} + \binom{n}{k+1}$ *for all* $n \in \mathbb{N}$ *and* $0 \leq k \leq n-1$.

The formula tells us that binomial coefficients also could be defined recursively using addition only, as in the *Pascal triangle* (which in fact the Chinese mathematician Chu Shih-Chieh had already presented in 1303!)

$$\begin{array}{ccccccccc} & & & & \binom{0}{0} & & & & \\ & & & \binom{1}{0} & & \binom{1}{1} & & & \\ & & \binom{2}{0} & & \binom{2}{1} & & \binom{2}{2} & & \\ & \binom{3}{0} & & \binom{3}{1} & & \binom{3}{2} & & \binom{3}{3} & \\ \binom{4}{0} & & \binom{4}{1} & & \binom{4}{2} & & \binom{4}{3} & & \binom{4}{4} \\ & & & & \cdots & & & & \end{array} \quad = \quad \begin{array}{ccccccccc} & & & & 1 & & & & \\ & & & 1 & & 1 & & & \\ & & 1 & & 2 & & 1 & & \\ & 1 & & 3 & & 3 & & 1 & \\ 1 & & 4 & & 6 & & 4 & & 1 \\ & & & & \cdots & & & & \end{array}$$

2. The real numbers $\mathbb{R}$

The reals are all too common from measurements. Nevertheless, mathematicians consider their existence less obvious and state it in terms of properties:

Theorem 5. *There is a nonempty set of* real numbers $\mathbb{R}$ *with the following three properties: It is a field, it is ordered, and complete.*

In the following subsections, we will explain the properties. Only in Sect. 2.4 below, we will sketch how the existence is established in terms of Cauchy sequences. The real numbers are unique in a sense explained for instance in the book by Ebbinghaus and others, [E, p.50].

2.1. **Groups.** A moment's thought shows that addition and multiplication of numbers are subject to similar laws. In fact, the following four are sufficient for each:

Definition. A set G with a map or *composition*

$$\circ\colon G \times G \to G \qquad (x, y) \mapsto x \circ y$$

is called a *group* if for all $x, y, z \in G$ holds:
1. *Associative law [*Assoziativitätsgesetz*]*: $x \circ (y \circ z) = (x \circ y) \circ z$
2. *Neutral element [*Neutrales Element*]*: There is an element $e \in G$, such that $e \circ x = x$ holds for all $x \in G$.
3. *Inverse element [*Inverses*]*: For each $x \in G$ there is an inverse element, denoted $x^{-1} \in G$, such that $x \circ x^{-1} = e$.

If in addition the following law holds, it is called an *Abelian* or *commutative group [Abelsche oder kommutative Gruppe]*:
4. *Commutative law [*Kommutativitätsgesetz*]*: $x \circ y = y \circ x$

We will use the shorthand notation $(G, \circ)$ for "a set G with composition $\circ$".

To see an example of a composition which is neither commutative nor associative consider division:

$$4 : 5 \neq 5 : 4, \qquad (1 : 2) : 3 = \frac{1}{6} \ \neq \ 1 : (2 : 3) = \frac{3}{2}.$$

Examples of Abelian groups. 1. The *integers [*ganze Zahlen*]*

$$\mathbb{Z} := \{\dots, -2, -1, 0, 1, 2, \dots\}$$

with addition $\circ := +$. The neutral element is $e = 0$, the inverse $x^{-1} := -x$.

2. $(\mathbb{R}, +)$ and $(\mathbb{R} \setminus \{0\}, \cdot)$ (check!).

3. Consider the two-element-set $\{O, I\}$. We think of O as "even *[gerade]* integer" and I as "odd *[ungerade]* integer" and accordingly define compositions $\oplus$ or $\otimes$ by

$$\begin{aligned} &O \oplus O = O, \quad O \oplus I = I, \quad I \oplus O = I, \quad I \oplus I = O, \\ &O \odot O = O, \quad O \odot I = O, \quad I \odot O = O, \quad I \odot I = I. \end{aligned} \tag{2}$$

It turns out that $(\{O, I\}, \oplus)$ is an Abelian group, while $(\{O, I\}, \odot)$ is not (why?).

4. The set of the rotations by 0, 120, and 240 degrees (details?).

On the other hand, the natural numbers $\mathbb{N}_0$ with addition do not form an Abelian group, as inverse elements do not exist: $3 + x = 0$ does not have a solution in $\mathbb{N}_0$.

All the rules of addition can be derived from the group laws:

Example. In $(\mathbb{R}, +)$ we have $-(-x) = x$.
Proof: By definition, $-(-x)$ is the inverse element of $-x$, that is, $0 = (-x) + \big(-(-x)\big)$. We add x to this equation and apply the group rules:

$$\begin{aligned} x &\overset{\text{neutral el.}}{=} x + 0 \;=\; x + \big((-x) + (-(-x))\big) \\ &\overset{\text{ass.law}}{=} \big(x + (-x)\big) + \big(-(-x)\big) \overset{\text{inverse}}{=} 0 + \big(-(-x)\big) \overset{\text{neutral el.}}{=} -(-x). \end{aligned} \tag{3}$$

Subtraction $x - y$ is a shorthand notation for $x + (-y)$, likewise *division* $\frac{x}{y}$ stands for $x \cdot y^{-1}$. Note that mathematicians only consider addition and multiplication the basic operations.

2.2. **Fields.** We now introduce a formal structure which presents a minimal set of rules for addition and multiplication.

Definition. A set $F = (F, +, \cdot)$ is called a *field [Körper]*, if the following holds:
(i) *Addition*: $(F, +)$ is an Abelian group, with a neutral element 0.
(ii) *Multiplication*: $(F \setminus \{0\}, \cdot)$ is an Abelian group with a neutral element 1.
(iii) The *distributive law [Distributivgesetz]* $r \cdot (x + y) = (x + y) \cdot r = r \cdot x + r \cdot y$ holds for all $r, x, y \in F$.

Examples of fields: 1. $\mathbb{R}$ and $\mathbb{C}$ (to be introduced later)

2. The smallest field has two elements and is $\mathbb{F}_2 := \big(\{O, I\}, \oplus, \odot\big)$ as in (2).

3. The *rational numbers [rationale Zahlen]*, also called *fractions [Brüche]*,

$$\mathbb{Q} := \left\{ \frac{p}{q} : p \in \mathbb{Z}, q \in \mathbb{N} \right\},$$

where $\frac{p}{q}$ stands for $p \cdot q^{-1}$. Addition and multiplication is defined in terms of the wellknown rules for fractions.

4. $\{x + y\sqrt{2} : x, y \in \mathbb{Q}\}$ with $+, \cdot$ as for $\mathbb{R}$.

On the other hand, $(\mathbb{Z}, +, \cdot)$ is not a field (why?).

All the well-known arithmetic rules can be derived from the properties of a field:

Example (standard binomial formula):

$$\begin{aligned}(x+y)^2 &= (x+y)(x+y) \overset{\text{distr. law}}{=} x(x+y) + y(x+y)\\ &\overset{\text{distr. law}}{=} x^2 + xy + yx + y^2 \overset{\text{comm. law (mult)}}{=} x^2 + xy + xy + y^2\\ &\overset{\text{neutr. el. (mult)}}{=} x^2 + 1xy + 1xy + y^2 \overset{\text{distr. law}}{=} x^2 + (1+1)yx + y^2 = x^2 + 2xy + y^2\end{aligned}$$

When we generalize to arbitrary powers, we encounter the binomial coefficients:

Theorem 6 (Binomial theorem)**.** *For any $n \in \mathbb{N}_0$ and for any $x, y \in F$ there holds*

$$(x+y)^n = \sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k}. \tag{4}$$

Remember that the right hand side denotes

$$\binom{n}{0} y^n + \binom{n}{1} x y^{n-1} + \ldots + \binom{n}{n-1} x^{n-1} y + \binom{n}{n} x^n.$$

To digest the following proof, consider the case $(x+y)^3 = (x+y)(x+y)(x+y) = a_0x^3 + a_1x^2y + a_2xy^2 + a_3y^3$. The coefficient a_1 of the term x^2y corresponds to the number of choices of picking x from two of the three parentheses and y from remaining one; there are $a_1 = 3$ such choices. Generalizing this thought gives:

Proof. For $n = 0$ we need to check that $1 = \binom{0}{0} \cdot 1 \cdot 1$, which is ok. For $n \geq 1$ let us apply the distributive and commutative law to the product

$$\underbrace{(x+y) \cdot \ldots \cdot (x+y)}_{n \text{ factors}} = \sum_{k=0}^{n} a_k \, x^{n-k} y^k.$$

The coefficient a_k arises by picking y from k parentheses, and x from the remaining $n-k$ parentheses; more precisely, it agrees with the number of such choices. In Sect. 1.4,3 we saw there are altogether $\binom{n}{k}$ ways to select k of the n parantheses. Hence $a_k = \binom{n}{k}$ which establishes (4). $\square$

`3. Vorlesung, Dienstag, 24.10.06` ___

It is a good exercise to prove (4) by induction as well. Setting $x = y = 1$ in (4) we obtain:

Corollary 7. *For all $n \in \mathbb{N}_0$ we have $\sum_{k=0}^{n} \binom{n}{k} = 2^n$.*

In view of Sect. 1.4,3 this says a set of n elements has 2^n different (unordered) subsets. This count includes the empty set and the entire set (for which values of k do they arise?)

Proposition 8. *For any $n \in \mathbb{N}_0$ and $x \neq 1$ in a field F we have the* geometric sum *formula*

$$1 + x + x^2 + x^3 + \ldots + x^n = \frac{1 - x^{n+1}}{1 - x}. \tag{5}$$

Proof. We calculate, employing a *telescope sum*:

$$(1 - x)\big(1 + x + \ldots + x^n\big) = (1 - x) + (x - x^2) + \ldots + (x^n - x^{n+1}) = 1 - x^{n+1}$$

Division by $1 - x \neq 0$ gives the claim. $\square$

2.3. **Ordering.** We analyse the inequality relation "$<$".

Definition. A set X with a relation $<$ is called *totally ordered* if
• (*transitivity*) $x < y$ and $y < z$ implies $x < z$, and
• (*trichotomy*) for all $x, y \in X$ precisely one of the three relations $x < y$ or $y < x$ or $x = y$ holds.

For any order we write:

$$y > x \text{ for } x < y; \qquad \text{and} \qquad x \leq y \text{ or } y \geq x \quad \text{for} \quad \big(x < y \text{ or } x = y\big). \tag{6}$$

Definition. A field F is *ordered* if it is totally ordered as a set and the ordering is compatible with addition and multiplication in the following sense:

$$x < y \quad \Longleftrightarrow \quad 0 < y - x, \quad \text{and} \tag{7}$$

$$0 < x, y \quad \text{implies} \quad 0 < xy. \tag{8}$$

Examples. 1. $\mathbb{Q}$, $\mathbb{R}$ are ordered fields.
2. It can be shown that neither $\mathbb{F}_2$ nor $\mathbb{C}$ can be ordered.

We want to mention a few consequences of an ordering without being exhaustive. These hold for any ordered field ($x, y \in F$).

1. *Adding to an inequality*:

$$x < y \quad \Longleftrightarrow \quad x + a < y + a \quad \text{for all } a \in F \tag{9}$$

Proof: By (7), $x < y \iff 0 < y - x = y + a - (x + a) \iff x + a < y + a$. $\square$

To draw a consequence, consider $0 < x, y$. From (9) follows $y < x + y$. But $0 < y$ and transitivity yield $0 < x + y$, so that

$$0 < x, y \quad \text{implies} \quad 0 < x + y.$$

2. *Multiplying an inequality*: Let $x < y$. Then $ax < ay$ in case $a > 0$, and $ax > ay$ in case $a < 0$.
Proof: $x < y \iff 0 < y - x$.
First case, $a > 0$. Then from (8) we conclude $0 < a(y-x) = ay - ax \iff ax < ay$.
Second case, $a < 0$. Then (7) (with $y = 0$) gives $-a > 0$. that is, $-a$ is positive. Applying the first case, we conclude $(-a)x < (-a)y$. But adding $ax + ay$ to both sides preserves the inequality by (9), and thus $ay < ax$. □

3. If $x \neq 0$ then $x^2 = xx > 0$ (proof: exercise); in particular $1 > 0$.
4. $x \geq y$ and $x \leq y \Longrightarrow x = y$.
5. If $x \geq y > 0$ multiply with the positive number $\frac{1}{xy}$ to obtain $\frac{1}{x} \leq \frac{1}{y}$.

For simplicity, let us state the following directly for $\mathbb{R}$:

Proposition 9 (Bernoulli's Inequality)**.** *Let $n \in \mathbb{N}$. Then for any $x \in \mathbb{R}$ with $x \geq -1$ we have*

$$(1+x)^n \geq 1 + nx. \tag{10}$$

For the special case $x \geq 0$ this is simple to see: The binomial formula (4) gives

$$(1+x)^n = 1 + nx + \binom{n}{2}x^2 + \ldots + x^n.$$

Since $x^2 > 0$, $x^3 > 0$, etc., the inequality is obvious. Now we give a general argument:

Proof. By induction: $n = 1$ ok.
Step $n \mapsto n+1$: We use $(1+x)^n \geq 1 + nx$. By 2., multiplication with $1 + x \geq 0$ preserves the inequality, so that $(1+x)^{n+1} \geq (1+nx)(1+x)$. Hence

$$(1+x)^{n+1} \geq 1 + (n+1)x + nx^2 \overset{3.}{\geq} 1 + (n+1)x.$$

□

2.4. **Modulus.**

Definition. On a field F, a *modulus* or *absolute value [(Absolut-)Betrag]* is a map $|.|\colon F \to \mathbb{R}$ which has the following properties for all $x, y \in F$:
1. $|x| \geq 0$, and $|x| = 0$ if and only if $x = 0$,
2. $|xy| = |x||y|$,
3. *triangle inequality [Dreiecks-Ungleichung]* $|x+y| \leq |x| + |y|$

On $\mathbb{R}$, such as on any ordered field, we can define a modulus by setting

$$|x| := \begin{cases} x & \text{for} \quad x \geq 0, \\ -x & \text{for} \quad x < 0. \end{cases} \tag{11}$$

Immediate from this definition is the following property:
4. $|x| \geq 0$ and $-|x| \leq x \leq |x|$.

Case distinctions show that (11) satisfies properties 1. to 3. For example, let us prove the triangle inequality 3.:
• Case $x + y \geq 0$: Then using 4. we obtain $|x + y| = x + y \leq |x| + |y|$.
• Case $x + y < 0$: From 4. we always have $-x \leq |x|$ and $-y \leq |y|$. So we obtain $|x + y| = -x - y \leq |x| + |y|$.

The name triangle inequality will become clear when we deal with the vector case. It implies, for any field with a modulus:
5. *Inverse triangle inequality [verschärfte Dreiecks-Ungleichung]* $\big||x| - |y|\big| \leq |x \pm y|$.

2.5. **Bounds and maxima of sets.** Before we explain the completeness of $\mathbb{R}$ we introduce two useful notions (think of the cases $X = \mathbb{Q}$ or $\mathbb{R}$).

Definition. A subset A of a totally ordered set X is called *bounded [beschränkt]*, if there is $C \in X$ with

$$|x| \leq C \qquad \text{for all } x \in A.$$

More specifically, A is

bounded from $\left\{ {\textit{above} \atop \textit{below}} \right\} \Leftrightarrow$ there is an $\left\{ {\textit{upper} \atop \textit{lower}} \right\}$ *bound* $b \in X$ such that $\left\{ {x \leq b \atop x \geq b} \right\}$ for each $x \in A$.

[A ist nach oben/unten beschränkt mit oberer/unterer Schranke b].

Examples. 1. $\{\frac{1}{n} \colon n \in \mathbb{N}\}$ is bounded: an upper bound is 1, and a lower bound is 0.
2. $\mathbb{N}$ is bounded below by 1. As we will show, it does not have an upper bound, and so is not bounded.
3. Any finite subset $A = \{x_1, \ldots, x_n\} \subset \mathbb{R}$ is bounded by $C := \max\{|x_1|, \ldots, |x_n|\}$. Here, the *maximum* of $n \in \mathbb{N}$ numbers is defined recursively. For $n = 1$, we set $\max\{x\} := x$. For $n = 2$ we set

$$\max\{x, y\} := \begin{cases} x & \text{for } x \geq y, \\ y & \text{for } x < y; \end{cases} \tag{12}$$

equivalently, $\max\{x, y\} := \frac{1}{2}\big(x + y + |y - x|\big)$ (see problems). The maximum of $n \geq 3$ numbers is

$$\max\{x_1, \ldots, x_n\} := \max\big\{ \max\{x_1, \ldots, x_{n-1}\}, x_n \big\}.$$

There are similar definitions for the *minimum* min.

Let us now define maximum and minimum for arbitrary subsets of $\mathbb{R}$:

Definition. Let A be a subset of a totally ordered set X. A number $m \in A$ is called the $\left\{ \substack{maximum \\ minimum} \right\}$ of A, if $\left\{ \substack{x \le m \\ x \ge m} \right\}$ for all $x \in A$.

Don't forget that we require the maximum or minimum is an element of the set A.

Examples. 1. The two sets $\{\frac{1}{n} \colon n \in \mathbb{N}\}$ and $\{\frac{1}{1+x^2} : x \in \mathbb{R}\}$ have 1 as a maximum, but from Cor. 13 it will follow they do not have a minimum.

2. Let $a \le b$. The closed interval $[a, b] := \{x \in \mathbb{R} \colon a \le x \le b\}$ has the maximum b. Indeed, $b \in [a, b]$ and $x \le b$ for all $x \in [a, b]$.

3. The open interval $(a, b) := \{x \in \mathbb{R} \colon a < x < b\}$ (defined for $a < b$) does not have a maximum. To see this, suppose $x \in (a, b)$ were a maximum. Then $x < b$ which gives us $x < \frac{b+x}{2} < b$. But that means the number $\frac{b+x}{2} \in (a, b)$ is not a bound. Hence x is not the maximum.

Problem. Show that each subset $A \neq \emptyset$ of $\mathbb{N}$ contains a minimal element. (Hint: If not, then $A \cap \{1, \ldots, n\} = \emptyset \quad$ for all $n \in \mathbb{N}$.)

2.6. **Completeness of $\mathbb{R}$.** Completeness certainly is the subtlest of the defining properties of the real numbers. Roughly, completeness means that the real lines has no holes. Before giving the precise definition, let us consider an example of a non-complete space: the rational numbers $\mathbb{Q}$.

Proposition 10. *The equation $x^2 = 2$ has no solution $x \in \mathbb{Q}$.*

Proof. On the contrary, we assume $2 = (\frac{m}{n})^2$ where $m \in \mathbb{Z}$, $n \in \mathbb{N}$, that is

$$2n^2 = m^2. \tag{13}$$

The Euclidean algorithm (see linear algebra?) shows that each natural number admits a unique prime factor decomposition; we use this fact now. We may assume that the fraction $p = \frac{m}{n}$ is a reduced representation, that is, m and n have no common prime factors.

The left hand side of (13) contains the prime factor 2. Therefore the right hand side must contain it, and so m must be even. But then m^2 is divisible by 4, and so $2n^2$ is divisible by 4, which means that n must be even. But we assumed that m and n have no common prime factors, and therefore we have derived a contradiction from (13); that is, (13) cannot hold. $\square$

`4. Vorlesung, Donnerstag, 26.10.06`

The most generally applicable characterization of completeness is in terms of Cauchy sequences (see Sect. 2.3 below). It is, however, somewhat abstract, and so we first introduce a more obvious characterization in terms of the supremum. This characterization is, however, limited to ordered fields, and so will not cover the complex numbers. A further characterization in terms of nested intervals will be discussed in Sect. 1.5 below.

More generally, we define optimal bounds, no matter if they are element of the sets under consideration:

Definition. Given a subset X of an ordered field F the number $s \in F$ is called the *supremum* if s is the *least upper bound [kleinste obere Schranke]*, that is,
(i) s is an upper bound for X, but
(ii) no $t < s$ is upper bound for X.
We write $\sup X := s$. Similarly, the *infimum*, denoted $\inf X$, is the *greatest lower bound.*

Examples. 1. $\inf\{\frac{1}{n} \colon n \in \mathbb{N}\} = 0$ and $\sup\{\frac{1}{n} \colon n \in \mathbb{N}\} = 1$ (check (i) and (ii)!)
2. Let $b \in \mathbb{R}$. Then the sets $(-\infty, b) := \{x \in \mathbb{R} : x < b\}$ and $(-\infty, b] := \{x \in \mathbb{R} : x \leq b\}$ have b as supremum, but they do not posess an infimum.
3. The empty set does not have a supremum nor an infimum (why?).

We can now state what it means for $\mathbb{R}$ to be complete:

Definition. An ordered field F is *complete [vollständig]* if each nonempty subset $X \subset F$ which is bounded above has a supremum $\sup X \in F$.

Let us give two applications of the completeness of $\mathbb{R}$. First we show the equation *[Gleichung]* $x^2 = 2$ is solvable *[lösbar]* within the reals.

Theorem 11. *For each $a > 0$ there is a unique number $x \in (0, \infty)$ with $a = x^2$. We write $\sqrt{a} := x$.*

Proof. Let us first show that x is unique. So suppose two different numbers $0 < x < \tilde{x}$ satisfy $x^2 = \tilde{x}^2 = a$. But then to $x^2 < x\tilde{x} < \tilde{x}^2$ which gives a contradiction. Therefore x is unique. (Where have we used trichotomy?)

Consider now the set $X := \{y \in \mathbb{R} \colon y \geq 0, y^2 < a\}$. Then X is non-empty, since $0 \in X$. Furthermore, X is bounded by $1 + a$. Indeed, if $y > 1 + a$ then in particular $y > 1$ and so $y^2 > y > 1 + a > a$, so that $y \notin X$.

We let $x := \sup X$ and show $x^2 = a$ by ruling out $x^2 < a$ and $x^2 > a$. Let us deal with the case $x^2 < a$. Suppose that ε is a number in $(0,1)$; then $\varepsilon \cdot \varepsilon < \varepsilon \cdot 1 = \varepsilon$. Using this, together with the upper bound $(1+a)$ for X, we find

$$(x+\varepsilon)^2 = x^2 + 2\varepsilon x + \varepsilon^2 < x^2 + 2\varepsilon(1+a) + \varepsilon = x^2 + \varepsilon(3+2a).$$

Hence for $\varepsilon \in (0,1)$ we have

$$\text{if} \quad \varepsilon < \frac{a-x^2}{3+2a} \qquad \text{then} \quad (x+\varepsilon)^2 < x^2 + \varepsilon(3+2a) < a.$$

Choosing such an ε gives $x + \varepsilon \in X$, contradicting the fact that x is an upper bound for X. Similarly, $x^2 > a$ leads to a contradiction with x being the least bound; we skip the details. $\square$

Problem. Extend the proof as to establish (i) the existence of n-th roots of positive numbers and (ii) the existence of odd roots of negative numbers.

Remark. Our proof shows that no number different from $\sqrt{a}$ can be the supremum of the set X. Specifically, in case $a = 2$, we showed before $\sqrt{2} \notin \mathbb{Q}$, and so the set $\{x \in \mathbb{Q} \colon x^2 < 2\}$ does not have a supremum in $\mathbb{Q}$. Thus $\mathbb{Q}$ is not complete.

2.7. **Archimedean property.** Our second application of the existence of the supremum is the Archimedean property. We formulate this useful consequence directly for $\mathbb{R}$:

Proposition 12 (Archimedean property)**.** *For each real number $a \in \mathbb{R}$ there is a natural number $n \in \mathbb{N}$ with $a < n$.*

Proof. Indirectly: Suppose, there exists $a \in \mathbb{R}$ with $n \le a$ for all $n \in \mathbb{N}$. Then the set $\mathbb{N}$ would be bounded above and have a supremum $b = \sup \mathbb{N}$. Since $b - 1 < b$ we have that $b-1$ is not an upper bound. Thus there must exist $n \in \mathbb{N}$ with $b-1 < n$, that is, $b < n+1$. The latter equation contradicts the fact that b is a bound, contradiction. $\square$

We will often employ an equivalent version of the Archimedean property:

Corollary 13. *For each real number $\varepsilon > 0$ there is an $n \in \mathbb{N}$ with $\frac{1}{n} < \varepsilon$.*

Proof. According to the Archimedean property there is $n \in \mathbb{N}$ with $n > \frac{1}{\varepsilon} \in \mathbb{R}$; the result follows. $\square$

Problems. 1. Prove $\inf\{\frac{1}{n} : n \in \mathbb{N}\} = 0$.

2. Each bounded subset $A \neq \emptyset$ of $\mathbb{N}$ contains a maximal element.

3. For each $x, y > 0$ there exists $n \in \mathbb{N}$ such that $nx > y$.

Remark. It was the property of Problem 3 which Archimede stated originally in geometric terms: Any long distance $y > 0$ can be exceeded by some multiple of a short distance $x > 0$.

`5. Vorlesung, Dienstag, 31.10.06` ______________________________

3. Complex numbers

As nice as they may be, the real numbers $\mathbb{R}$ have the drawback that equations such as $x^2 = -1$ cannot be solved – recall that $x^2 \geq 0$ for any $x \in \mathbb{R}$! If we give up the ordering property, we can solve such equations.

Remark. Why do we care about a solution of $x^2 = -1$? Historically, the problem was to decide about the solution formulas for quadratic or cubic equations: What does $\sqrt{b^2 - 4ac}$ denote in case the number under the root is not positive?

Definition. The set of *complex numbers [komplexe Zahlen]* $\mathbb{C}$ consists of the pairs of real numbers $\mathbb{R} \times \mathbb{R}$, together with the compositions

$$(x,y) + (u,v) = (x+u, y+v), \qquad (x,y) \cdot (u,v) := (xu - yv,\, xv + yu). \tag{14}$$

Addition is simply addition of vectors in the plane $\mathbb{R}^2 = \mathbb{R} \times \mathbb{R}$. Let us now give a nice geometric interpretation of the less obvious product. As we shall prove only later (Thm. III.24), each vector $(x,y) \in \mathbb{R}^2$ has a *polar representation* $(x,y) = (r\cos\varphi, r\sin\varphi)$ with angle or *argument* $\varphi \in \mathbb{R}$ (not unique) and length or *modulus*

$$|z| := \sqrt{x^2 + y^2} \geq 0.$$

Inserting polar representations into the product and using the addition theorems for sin and cos, it becomes transparent that the product (14) of two complex numbers in $\mathbb{R}^2$ satisfies:
- the length of the product is the product of the lengths, and
- the angle of the product is the sum of the angles.

Examples. $(0,1)^2 = (-1,0)$ and $(1,1)^2 = (0,2)$. Verify the geometric interpretation for these two products!

Theorem 14. *The complex numbers $\mathbb{C}$ form a field.*

Proof. The proof of the field properties is by calculation (exercise!). Here, let us only mention that $(0,0)$ is the neutral element of addition, and $(1,0)$ is the neutral element of multiplication; moreover the multiplicative inverse of (x,y) is $\left(\frac{x}{x^2+y^2}, -\frac{y}{x^2+y^2}\right)$ (check!). $\square$

Remarks. 1. It can be shown that the field $\mathbb{C}$ has no ordering (problems).
2. In Sect. 2.3 below we will see the complex numbers are complete in the sense of Cauchy sequences.

Another name for the complex numbers is the *imaginary* numbers. It hints at the fact that the reals had seemed perfectly appropriate for real-life measurements, and so the complex numbers were considered an artefact (read Euler's mistaken view on [E, p.59]!). However, nowadays mathematicians, scientists, and engineers consider complex numbers an equally natural and useful concept.

Writing $i := (0, 1)$ for the *imaginary unit* i, it is customary to express

$$z = (x, y) = (1, 0)x + (0, 1)y = x + iy.$$

The product law (14) then becomes easy to memorize: We take standard products of real numbers subject to the additional relation $i^2 = -1$. For instance, to calculate the second example,

$$(1 + i)^2 = 1 + 2i + i^2 = 2i.$$

We call the components $x = \operatorname{Re} z$ the *real part [Realteil]* of z, and $y = \operatorname{Im} z$ the *imaginary part [Imaginärteil]*. The notation $z = x + iy$ incorporates the identification of the real numbers $\mathbb{R}$ with the subset $\mathbb{R} \times \{0\}$. (Check that addition and multiplication on $\mathbb{R}$ and $\mathbb{R} \times \{0\} \subset \mathbb{C}$ coincides!)

Recall the properties for a modulus from Sect. 2.4:
1. $|z| > 0$ for all $z \neq 0$,
2. $|zw| = |z||w|$,
3. *triangle inequalities*: $|z + w| \leq |z| + |w|$ and $\big||z| - |w|\big| \leq |z \pm w|$

We leave it as an exercise to verify that for all $z, w \in \mathbb{C}$ these properties are satisfied by $|z| := \sqrt{\operatorname{Re}^2 z + \operatorname{Im}^2 z}$. Note that 2. is precisely the length law in our geometric interpretation of the product, and the triangle inequality can now be seen as an inequality on the edgelengths of a triangle with vertices 0, v, w. There is a further property to mention for $\mathbb{C}$:
4. $|\operatorname{Re} z| \leq |z|$ and $|\operatorname{Im} z| \leq |z|$.

It is useful to define the number $\overline{z} := x - iy$, the *conjugate [Konjugierte]* of z. Geometrically conjugation is reflection about the real axis. We have the following properties (the proof is an exercise):

$$\overline{\overline{z}} = z, \quad |z|^2 = z\overline{z}, \quad \overline{z + w} = \overline{z} + \overline{w}, \quad \overline{zw} = \overline{z}\,\overline{w}, \quad \operatorname{Re} z = \frac{z + \overline{z}}{2}, \quad \operatorname{Im} z = \frac{z - \overline{z}}{2i}$$

Let us finally discuss the solvability of equations. We start with

$$z^n = 1 \qquad \text{for } z \in \mathbb{C},\ n \in \mathbb{N}. \tag{15}$$

By 2. any given z with $|z| = 1$ satisfies $|z^n| = |z|^n = 1$. Taking into account the angle law as well, we see there are n distinct *[verschieden]* solutions of (15): These are the numbers $z_0 = 1, z_1, z_2, \ldots, z_{n-1}$ on the unit circle with the angles $0, \frac{1\cdot 2\pi}{n}, \frac{2\cdot 2\pi}{n}, \ldots, \frac{(n-1)\cdot 2\pi}{n}$. They are called the *n-th roots of unity [n-te Einheitswurzeln]*.

More generally, the *fundamental theorem of algebra* asserts that any *polynomial*

$$p(z) := a_0 + a_1 z + a_2 z^2 + \ldots + a_n z^n \qquad \text{with } a_0, a_1, \ldots, a_n \in \mathbb{C} \text{ and } a_n \neq 0$$

has n zeros. That is, we can find a factorization

$$p(z) = c(z - z_0) \cdot \ldots \cdot (z - z_{n-1}) \qquad \text{with } c, z_0, \ldots, z_{n-1} \in \mathbb{C},$$

where the zeros $z_0, \ldots, z_{n-1}$ need not be distinct.

4. Cardinality

Two sets X and Y are said to be *equipotent [gleichmächtig]* or of the same *cardinality* if there exists a bijection $f\colon X \to Y$. If a set X has the same cardinality as $\{1, \ldots, n\}$ for $n \in \mathbb{N}_0$ then we say *X has n elements* or X is *finite [endlich]*; otherwise, X is *infinite [unendlich]*.

We call an infinite set which is in bijection to $\mathbb{N}$ *countable [abzählbar]*.

Proposition 15. *The integers $\mathbb{Z}$ are countable.*

Proof. We count them as $0,\ 1, -1,\ 2, -2,\ 3, -3, \ldots$, that is, we set $f(n) := \frac{n}{2}$ for n even and $f(n) := \frac{1-n}{2}$ for n odd. □

We can rephrase the Proposition to say that $\mathbb{Z}$ is bijective to its proper subset $\mathbb{N}$! This property can be shown to characterize infinite sets. Similarly:

Proposition 16. *The rational numbers $\mathbb{Q}$ are countable.*

Proof. The table

$$\begin{array}{ccccccc}
0 & 1 & -1 & 2 & -2 & 3 & \ldots \\
\frac{0}{2} & \frac{1}{2} & -\frac{1}{2} & \frac{2}{2} & -\frac{2}{2} & \frac{3}{2} & \ldots \\
\frac{0}{3} & \frac{1}{3} & -\frac{1}{3} & \frac{2}{3} & -\frac{2}{3} & \frac{3}{3} & \ldots \\
\frac{0}{4} & \frac{1}{4} & -\frac{1}{4} & \frac{2}{4} & -\frac{2}{4} & \frac{3}{4} & \ldots \\
 & \ldots & & & & &
\end{array}$$

contains all rational numbers; in fact each one occurs infinitely often. We can count them using diagonal paths; we make sure we count a further rational number only when it is new (that is, we ignore fractions which coincide after reduction). □

The real numbers are not countable as we shall see later.

Let us mention a further result. If X is an arbitrary set, then its *power set* $\mathcal{P}(X)$ is the set of all subsets of X. For instance, if X is finite and has n elements, then $\mathcal{P}(X)$ has 2^n elements (why?); in particular X and $\mathcal{P}(X)$ are not equipotent. The result we would like to mention is that also an infinite set X and its power set $\mathcal{P}(X)$ are not equipotent.

We refer to [D, Chapter 2] for a discussion of cardinality.

Summary

We introduced numbers in terms of their properties: For the natural numbers this is the successor property, for the reals these are the arithmetic, ordering and completeness properties.

That the natural numbers exist is a philosophical problem – they were given by God, according to Kronecker. But all the other numbers can be derived step by step from the natural numbers. For the real numbers, we will indicate how this is done in the next section.

Along the way, we have introduced important concepts which we will apply allover analysis. A key technique is induction. We also mentioned computational results such as the binomial theorem or the geometric sum formula. It could be said that analysis is the computation with inequalities. In this sense notions like boundedness, maximum and supremum are crucial. The computation with inequalities and moduli, (in particular the triangle inequalities) will be ubiquitous in the secquel.

Part 2. Sequences and Series

1. Sequences

Sequences are a basic tool of analysis: Continuity, differentiability and many other concepts are formulated in terms of sequences.

1.1. Convergence of sequences.

Definition. A *sequence [Folge]* is a map

$$a\colon \mathbb{N} \to \mathbb{C}, \quad n \mapsto a_n.$$

If all $a_n \in \mathbb{R}$ we call the sequence *real.*

We will always denote the mapping a by the symbol $(a_n)_{n\in\mathbb{N}}$ or (a_n). Only when we say this explicitely, we will allow for the domain $\mathbb{N}_0$ as well.

A sequence can be defined by its mapping law, such as $a_n = n^2$, or by enumeration $1, 4, 9, 16, \dots$ (if the mapping law is obvious). It can also have a recursive definition:

$$a_1 := 1, \quad a_2 := 1, \quad \text{and} \quad a_n := a_{n-1} + a_{n-2} \quad \text{for } n \geq 3,$$

defines the *Fibonacci sequence* $1, 1, 2, 3, 5, 8, 13, 21, \dots$ (see problems).

Definition. A sequence (a_n) is said to *converge* or *tend to* $a \in \mathbb{C}$ *[Folge konvergiert oder strebt gegen a]* if the following holds:

(1) For each $\varepsilon > 0$ there exists $N = N(\varepsilon) \in \mathbb{N}$ such that $|a_n - a| < \varepsilon$ for all $n \geq N$.

We call a the *limit [Grenzwert]* of (a_n) and write

$$\lim_{n\to\infty} a_n = a \qquad \text{or} \qquad a_n \to a \text{ for } n \to \infty.$$

If (a_n) converges to $a = 0$ it is called a *null sequence [Nullfolge]*. We say (a_n) *diverges* if it does not converge.

We consider ε as an error bound *[Fehlerschranke]*; the elements $a_N, a_{N+1}, a_{N+2}, \dots$ (which is all but finitely many!) must beat the given error bound. Rephrased in geometric language, these elements are contained in the ε-*ball* about a,

$$B_\varepsilon(a) := \{z \in \mathbb{C} : |z - a| < \varepsilon\}.$$

By writing $N = N(\varepsilon)$ we indicate that the choice of N depends on ε. The magnitude of $N(\varepsilon)$ can be viewed as to measure the speed of convergence: If N must be large, the convergence is slow.

Using the symbols (or *quantors*) $\exists$ for "there exists" and $\forall$ for "for all" we can write (1) as

$$\forall \varepsilon > 0 \quad \exists N = N(\varepsilon) \in \mathbb{N} : |a_n - a| < \varepsilon \quad \forall n \geq N.$$

Examples. 1. The constant sequence $a_n = a \in \mathbb{C}$ converges to a: For any $\varepsilon > 0$ pick $N = 1$.

2. $(\frac{1}{n})$ is a null sequence.

Proof: Let $\varepsilon > 0$ be given. We want to determine an index N such that $|\frac{1}{n} - 0| < \varepsilon$ for all $n \geq N$. According to the Archimedean property Proposition 12 we can pick $N \in \mathbb{N}$ with $\frac{1}{N} < \varepsilon$. Then

$$\left|\frac{1}{n} - 0\right| = \frac{1}{n} \leq \frac{1}{N} < \varepsilon \quad \text{for all } n \geq N,$$

that is, (1) holds. Any larger choice of N would work equally well. □

3. Let $z \in \mathbb{C}$ with $|z| < 1$. The *geometric sequence* $z^0 = 1, z, z^2, z^3, \ldots$ is null.

If we knew logarithms we could give a straightforward proof based on the fact that $\varepsilon > |z|^N \geq |z|^{N+1} \geq \ldots$ holds for $N > \log \varepsilon / \log |z|$. Instead, we have to refer to elementary arguments in the following.

Proof: In case $z = 0$ we choose $N := 2$.

Otherwise $0 < |z| < 1$ and we can write $|\frac{1}{z}| =: 1 + \delta$ with some $0 < \delta$. The estimate

$$\frac{1}{|z^n|} = \left|\frac{1}{z}\right|^n = (1+\delta)^n \overset{\text{Bernoulli}}{\geq} 1 + n\delta > n\delta$$

gives

$$|z^n| < \frac{1}{n\delta}. \tag{2}$$

We now determine N by the requirement $\frac{1}{n\delta} < \varepsilon$ for all $n \geq N$. According to the Archimedean property Proposition 12 we can choose $N \in \mathbb{N}$ with $\frac{1}{N} < \varepsilon\delta$. Then indeed we have for all $n \geq N$

$$|z^n| \overset{(2)}{<} \frac{1}{n\delta} \overset{n \geq N}{\leq} \frac{1}{N} \cdot \frac{1}{\delta} \overset{\text{choice of } N}{<} \varepsilon\delta \cdot \frac{1}{\delta} = \varepsilon,$$

as required for (1). □

`6. Vorlesung, Donnerstag, 2.11.06`

---

4. $a_n = 2n$ diverges: Suppose a was the limit. Using the Archimedean property we find $N \in \mathbb{N}$ with $a < N < a_n$. Consequently, $a_n - a \geq a_N - a > 2$ for all $n \geq N + 1$, contradicting convergence with $\varepsilon = 1$.

5. $a_n := (-1)^n$ diverges. Suppose a is the limit. Then $a \neq +1$ or $a \neq -1$. In the first case take $\varepsilon := \frac{1}{2}|a - 1| > 0$. Then for all n even we have $|a_n - a| = |1 - a| > \varepsilon$. Thus N as required does not exist. Similar for the case $a \neq -1$.

The last two examples display different kinds of the divergence of real sequences, which we would like to distinguish:

Definition. A real sequence (a_n) *diverges to [divergiert bestimmt gegen]* $\left\{ {\infty \atop -\infty} \right\}$, if for each $C \in \mathbb{R}$ there is an index $N \in \mathbb{N}$ with $\left\{ {a_n > C \atop a_n < C} \right\}$ for all $n \geq N$.

Remark. The use of the symbol ∞ for limits is purely symbolic: ∞ is not a number, and expressions like $\infty + 1$ are not allowed.

1.2. **Properties of convergent sequences.** The following property of the limit must be derived from its definition:

Theorem 1. *The limit of a convergent sequence is uniquely determined.*

Proof. Let us give an indirect proof, that is, we derive a contradiction from assuming that a sequence (a_n) has the two different limits $a \neq b$. The idea is to find ε-balls about a and b which are disjoint. We set $\varepsilon := \frac{|a-b|}{2} > 0$. By convergence, we find $N', N'' \in \mathbb{N}$ with

$$|a_n - a| < \varepsilon \;\text{ for }\; n \geq N' \qquad \text{and} \qquad |a_n - b| < \varepsilon \;\text{ for }\; n \geq N''.$$

Both inequalities hold when $n \geq N := \max\{N', N''\}$, and so we conclude for these n

$$2\varepsilon = |a - b| = \big|(a - a_n) + (a_n - b)\big| \overset{\Delta-\text{inequ.}}{\leq} |a - a_n| + |a_n - b| < 2\varepsilon.$$

This contradiction shows that $a \neq b$ cannot hold. $\square$

Definition. A sequence $(a_n)_{n\in\mathbb{N}}$ is *bounded* if its range $X := \{a_1, a_2, a_3, \ldots\}$ is a bounded set.

Examples. 1. For the sequence $1, -1, 1, -1, \ldots$ the set X consists of exactly two elements, $X = \{1, -1\}$. Hence the sequence is bounded ($C = 1$).
2. The sequence $(\frac{1}{n})_{n\in\mathbb{N}}$ is bounded since $|\frac{1}{n}| \leq 1 =: C$.

Lemma 2. *A convergent sequence is bounded.*

Note that the converse is not true (see Example 1).

The idea of the proof is that a suitable "tail" of the sequence is within distance 1 of the limit, and so only finitely many elements remain to be considered.

Proof. Let $\lim a_n = a$. Then there is $N \in \mathbb{N}$ such that for all $n \geq N$ we have $|a_n - a| < 1$ and thus

$$|a_n| = \big|a + (a_n - a)\big| \overset{\Delta\text{-inequ.}}{\leq} |a| + |a_n - a| < |a| + 1.$$

So if we set $C := \max\{|a_1|, \ldots, |a_{N-1}|, |a| + 1\}$ then $|a_n| \leq C$ for all $n \in \mathbb{N}$. $\square$

Example. Geometric sequence $a_n = z^n$ for $|z| > 1$. In this case, we can write $|z| = 1 + \delta$ with $\delta > 0$. Then $|z|^n = (1+\delta)^n > 1 + n\delta$ and so (z^n) is not bounded. It follows that z^n is divergent for $|z| > 1$. The sequence is also divergent when $|z| = 1$ and $z \neq 1$, but this fact does not follow from the lemma.

1.3. **Limit theorems.** Convergence is preserved under many operations.

Proposition 3. *Let (a_n) and (b_n) be sequences which converge; let moreover $c \in \mathbb{C}$. Then also $(a_n + b_n)$, (ca_n), $(a_n b_n)$, and $(|a_n|)$ converge to the following limits:*

$$\text{(3)} \qquad \begin{array}{llll} (i) & \lim_{n\to\infty} (a_n + b_n) = \lim_{n\to\infty} a_n + \lim_{n\to\infty} b_n, & (ii) & \lim_{n\to\infty} (ca_n) = c \lim_{n\to\infty} a_n, \\ (iii) & \lim_{n\to\infty} (a_n b_n) = \lim_{n\to\infty} a_n \lim_{n\to\infty} b_n, & (iv) & \lim_{n\to\infty} |a_n| = \Big| \lim_{n\to\infty} a_n \Big|. \end{array}$$

Rephrased in the language of linear algebra, the space of convergent sequences is a vector space. By (i) and (ii) the limit is a linear map from this space to $\mathbb{C}$ (that is, a functional). Does the same hold for divergent sequences, or bounded sequences?

Proof. Let a and b be the limits of (a_n) and (b_n), respectively; moreover, let $\varepsilon > 0$ be an arbitrary number.

(i) Since (a_n) and (b_n) converge, we find $N', N'' \in \mathbb{N}$ with

$$|a_n - a| < \frac{\varepsilon}{2} \quad \text{for } n \geq N' \quad \text{and} \quad |b_n - b| < \frac{\varepsilon}{2} \quad \text{for } n \geq N''.$$

Both of these two inequalities hold when $n \geq N := \max\{N', N''\}$, and so

$$\big|(a_n + b_n) - (a+b)\big| \overset{\Delta\text{-inequ}}{\leq} |a_n - a| + |b_n - b| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon \qquad \text{for all } n \geq N.$$

Thus $(a_n + b_n)$ converges to $a + b$.

(iii) Since (a_n) converges, Lemma 2 gives $|a_n| \leq C$ where $C > 0$ depends on the sequence (a_n) but not on ε. We may also assume $|b| < C$. Therefore for all n:

$$|a_n b_n - ab| = \big|a_n(b_n - b) + b(a_n - a)\big| \leq |a_n||b_n - b| + |b||a_n - a| \leq C\big(|b_n - b| + |a_n - a|\big)$$

We now choose N in order for this to become less than ε for large n: By the convergence of (a_n) and (b_n) we have

$$|a_n - a| < \frac{\varepsilon}{2C} \quad \text{for } n > N' \quad \text{and} \quad |b_n - b| < \frac{\varepsilon}{2C} \quad \text{for } n > N''.$$

Setting $N := \max\{N', N''\}$ gives altogether

$$|a_n b_n - ab| \leq C\Big(\frac{\varepsilon}{2C} + \frac{\varepsilon}{2C}\Big) = \varepsilon \qquad \text{for all } n \geq N.$$

(ii) follows from (iii) by setting $b_n := c$. We leave the proof of (iv) as an exercise. $\square$

Theorem 4. *If (a_n) is a sequence with $a_n \neq 0$ for all n and $\lim_{n\to\infty} a_n \neq 0$, then*

$$\lim_{n\to\infty} \frac{1}{a_n} = \frac{1}{\lim_{n\to\infty} a_n}.$$

Letting $a := \lim_{n\to\infty} a_n$, again our ansatz is to express the error in terms of $a_n - a$:

$$\left|\frac{1}{a_n} - \frac{1}{a}\right| = \frac{|a_n - a|}{|a_n||a|}. \tag{4}$$

Using the convergence of (a_n), we can estimate the numerator *[Zähler]*. In the denominator *[Nenner]*, while $|a|$ is a fixed number causing no problems, the factor $|a_n|$ is a dangerous term: We need to show that $|a_n|$ cannot approach 0, in which case the fraction (4) would be impossible to bound. We formulate a bound on $\frac{1}{|a_n|}$ as a separate statement.

Lemma 5. *If (a_n) converges to $a \neq 0$ then there is $M \in \mathbb{N}$ such that $|a_n| > \frac{|a|}{2}$ for all $n \geq M$.*

In particular, $a_n = 0$ can only be true for finitely many indices n.

Proof. By assumption $\varepsilon := \frac{1}{2}|a|$ is positive, and since a_n converges we can choose M such that $|a_n - a| < \frac{1}{2}|a|$ for all $n \geq M$. Consequently,

$$|a_n| = |a + (a_n - a)| \geq |a| - |a_n - a| > |a| - \frac{1}{2}|a| = \frac{1}{2}|a| \quad \text{for all } n \geq M. \qquad \square$$

Proof of the theorem. Given $\varepsilon > 0$, we choose $N' \in \mathbb{N}$ with

$$|a - a_n| < \frac{1}{2}\varepsilon|a|^2 \qquad \text{for all } n \geq N'.$$

Moreover, we also invoke the result of the Lemma, and obtain for $N := \max\{M, N'\}$ that

$$\left|\frac{1}{a_n} - \frac{1}{a}\right| = \frac{|a_n - a|}{|a_n||a|} < \frac{\frac{1}{2}\varepsilon|a|^2}{(\frac{1}{2}|a|)|a|} = \varepsilon \quad \text{for all } n \geq N. \qquad \square$$

Note that Thm. 4 combined with Prop. 3 shows that $\frac{a_n}{b_n} = a_n \cdot \frac{1}{b_n}$ converges, provided (b_n) satisfies the appropriate assumptions.

Example. We can apply the above rules (which ones, precisely?) to conclude

$$\lim_{n\to\infty} \frac{3n|n^2 - 17|}{(n+1)^3} = \lim_{n\to\infty} \frac{3|1 - \frac{17}{n^2}|}{(1 + \frac{1}{n})^3} = 3.$$

`7. Vorlesung, Dienstag, 7.11.06` ______________________________

Let us finally relate real and complex sequences:

Proposition 6. *A sequence of complex numbers (a_n) converges if and only if its real and imaginary parts $(\operatorname{Re} a_n)$ and $(\operatorname{Im} a_n)$ converge. In that case, $\lim a_n = \lim \operatorname{Re} a_n + i \lim \operatorname{Im} a_n$. In particular, a limit of a real sequence is real.*

Proof. Suppose (a_n) converges to a. Then $|\operatorname{Re} a_n - \operatorname{Re} a| = |\operatorname{Re}(a_n - a)| \le |a_n - a| \to 0$ and so $\operatorname{Re} a_n \to \operatorname{Re} a$. Similarly for $\operatorname{Im} a_n$. Conversely, if the sequences $\operatorname{Re} a_n$ and $\operatorname{Im} a_n$ converge, then by Prop. 3(i),(ii) also $\operatorname{Re} a_n + i \operatorname{Im} a_n = a_n$ converges. $\square$

1.4. **Real sequences.**

Definition. A real sequence (a_n) is *monotonically* $\left\{ \begin{smallmatrix} \text{increasing} \\ \text{decreasing} \end{smallmatrix} \right\}$ if $\left\{ \begin{smallmatrix} a_1 \le a_2 \le a_3 \le \dots \\ a_1 \ge a_2 \ge a_3 \ge \dots \end{smallmatrix} \right\}$.
[monoton wachsend/fallend]

Examples. 1. $\left(\frac{1}{n}\right)$ is monotonically decreasing, $\left(-\frac{1}{n}\right)$ is monotonically increasing.
2. The constant sequence $17, 17, 17, \dots$ is monotonically increasing and decreasing.

Theorem 7. *Let (a_n) be a monotone sequence. Then (a_n) converges if and only if it is bounded.*

Example. The sequence $1,\ 1+\frac{1}{2},\ 1+\frac{1}{2}+\frac{1}{3},\ \dots$ is monotonically increasing. It is not immediately apparent whether the sequence is bounded and thus, according to the theorem, convergent. Try to let a computer figure out if the sequence is bounded!

Proof. If a_n is convergent, then it is bounded by Lemma 2.

We prove the converse. Consider, for instance, the increasing case. The completeness of $\mathbb{R}$ implies that the bounded set $\{a_n : n \in \mathbb{N}\}$ has a supremum a. We now prove $\lim_{n\to\infty} a_n = a$.

Let $\varepsilon > 0$. Since a is the least upper bound, the number $a - \varepsilon$ no longer is an upper bound and thus we can find $N \in \mathbb{N}$ with $a_N > a - \varepsilon$. Consequently, for all $n \ge N$,

$$a - \varepsilon \quad < \quad a_N \quad \overset{\text{monotonicity}}{\le} \quad a_n \qquad \Rightarrow \qquad a - a_n < \varepsilon.$$

But a is the supremum of the a_n, that is, $a_n \le a$. Thus $|a - a_n| = a - a_n < \varepsilon$. $\square$

The previous proof employed the supremum in an essential way. In fact, the Theorem is equivalent to the completeness of the real numbers (see Subsection 2.3 below).

We need some order preserving properties of real sequences:

Proposition 8. *Suppose $a_n \to a$ and $b_n \to b$ are convergent real sequences.*
(i) *If $a_n \le b_n$ for all $n \in \mathbb{N}$ then $a \le b$.*
(ii) *If $a_n \le x_n \le b_n$ for all $n \in \mathbb{N}$ and $a = b$ then (x_n) converges to $a = b$ as well.*

Proof. Note that

$$(5)\qquad \begin{aligned} |a_n - a| < \varepsilon \quad &\Rightarrow \quad a - a_n < \varepsilon \quad \Rightarrow \quad a - \varepsilon < a_n \\ |b_n - b| < \varepsilon \quad &\Rightarrow \quad b_n - b < \varepsilon \quad \Rightarrow \quad b_n < b + \varepsilon \end{aligned}$$

In case (i) using this in conjunction with $a_n \leq b_n$, we find

$$(6)\qquad a - \varepsilon < b + \varepsilon \qquad \text{for all } \varepsilon > 0.$$

But (6) implies $a \leq b$, which we prove indirectly: Were $a > b$, then we use $\varepsilon := \frac{a-b}{2}$ in two times (6); this gives $a + b = 2a - (a - b) < 2b + (a - b) = b + a$, a contradiction.

In case (ii), we conclude from (5) that for all $\varepsilon > 0$ there is $N \in \mathbb{N}$ such that

$$a - \varepsilon \leq x_n \leq b + \varepsilon = a + \varepsilon \quad \Rightarrow \quad |x_n - a| < \varepsilon,$$

thereby proving $x_n \to a$. □

Note, however, that $\frac{1}{n} > 0$ does not imply $\lim_{n\to\infty} \frac{1}{n} > 0$. Thus for convergent sequeences, the strong inequality $a_n < b_n$ for all n implies only the weak inequality $a = \lim a_n \leq b = \lim b_n$!

1.5. **Nested intervals.** We discuss what is perhaps the most transparant characterization of the completeness of $\mathbb{R}$.

An *interval* I is a subset of $\mathbb{R}$ such that for each pair of points $x < y \in I$ the set of intermediate points $\{\xi : x < \xi < y\}$ is also contained in I.

Examples. Open intervals (a, b) and closed intervals $[a, b]$. Likewise *halfopen intervals* such as $(a, b]$ and $[a, b)$ (for $a < b$). In all these cases, we call $b - a$ the *length* of the interval. Moroever, $\mathbb{R}$ and $(a, \infty) = \{x \in \mathbb{R} \mid a < x\}$ etc. are intervals.

Let us now define:

Definition. A sequence $(I_n)_{n\in\mathbb{N}}$ of closed intervals $I_n = [a_n, b_n]$ is called an *interval nesting [Intervallschachtelung]* provided the following holds:
(i) $I_n \supset I_{n+1}$ for all $n \in \mathbb{N}$.
(ii) The interval length $|I_n| = b_n - a_n$ converges to 0.

We can formulate the completeness of $\mathbb{R}$ in terms of interval nestings:

Theorem 9. *For each interval nesting (I_n) there exists a unique real number x, which is contained in all intervals I_n.*

Examples. 1. Given the decimal expansion of π, we can define nested intervals $I_1 = [3,4]$, $I_2 = [3,1;3,2]$, $I_3 = [3,14;3,15]$, etc. They contain π as the only point in common.
2. The same intervals, taken as subsets of $\mathbb{Q}$, do not contain a point in common.
3. Does a sequence of nested open intervals always contain a point in common?

Proof. The sequence (a_n) is monotone increasing and bounded by, say, b_1. By Thm. 7 it converges to some number a. Similarly, (b_n) converges to b. We can conclude $a = b$ by applying Prop. 3:

$$|b-a| = |\lim b_n - \lim a_n| = |\lim(b_n - a_n)| = \lim |b_n - a_n| = 0.$$

By monotonicity, $a_n \le a = b \le b_n$, and so certainly $a \in I_n$ for all n. On the other hand, suppose that also c lies in all I_n. Then $a_n \le c \le b_n$. By Prop. 8(ii) this implies $a = \lim a_n \le c = \lim b_n = a$, that is, $a = c$. □

We will not show the converse, which is nevertheless true: The interval nesting property of $\mathbb{R}$ implies that each bounded set has a supremum.

We can use nested intervals to show a surprising fact:

Theorem 10 (Cantor 1878)**.** *The real numbers $\mathbb{R}$ are not countable.*

Proof. (indirectly) Suppose $x_0, x_1, x_2, x_3, \dots$ is an enumeration of the reals. We will construct nested intervals $(I_n)_{n\in\mathbb{N}_0}$ such that $x_n \notin I_n$ for all n. By Thm. 9 there exists a number $x \in \bigcap_{k\in\mathbb{N}} I_k$. But x_n cannot be contained in $\bigcap_{k\in\mathbb{N}} I_k$, and so x_n must be missing in the enumeration.

Let us now define recursively the intervals I_n which have length $(\frac{1}{3})^n$:
• $I_0 := [x_0+1, x_0+2] \not\ni x_0$.
• For $n \in \mathbb{N}$ we subdivide I_{n-1} into three closed intervals J, K, L of equal length. We define

$$I_n := \begin{cases} J & \text{if } x_n \in L, \\ L & \text{if } x_n \notin L. \end{cases}$$

This implies $x_n \notin I_n$. (A partition into just two closed subintervals would not work as two subintervals both contain the midpoint.) □

A proof which gives more insight in cardinality shows that the power set $P(\mathbb{N})$ has the same cardinaltiy as $\mathbb{R}$, see [D, p.78].

2. Cauchy sequences and completeness

Cauchy sequences are a useful tool to decide about convergence. Their real importance lies, however, in the fact that they give rise to a notion of completeness: While the supremum refers to the ordering, a Cauchy sequence can still be defined for spaces with a modulus. Thus the Cauchy sequence characterization of completeness applies, for instance, to the set of complex numbers. It has the power to generalizes to many other situations, such as function spaces in place of numbers.

2.1. **Cauchy sequences.** The definition of convergence of a sequence (a_n) is in terms of its limit a. Can convergence be characterized without refering to the limit? Cauchy found a smart criterion:

`8. Vorlesung, Donnerstag, 9.11.06` ______________________________

Definition (Cauchy 1821)**.** A sequence (a_n) is a *Cauchy* (or *fundamental*) *sequence* if

(7) for each $\varepsilon > 0$ there is $N = N(\varepsilon) \in \mathbb{N}$ with $|a_n - a_m| < \varepsilon$ for all $n, m \geq N$.

To abbreviate, we will denote this often as $|a_n - a_m| \to 0$ as $n, m \to \infty$.

Examples. 1. $\frac{1}{n}$ is Cauchy. Indeed, for any $\varepsilon > 0$ let us choose $N > \frac{2}{\varepsilon}$. Then for $n, m \geq N$

$$\left|\frac{1}{n} - \frac{1}{m}\right| \overset{\Delta\text{-inequ.}}{\leq} \frac{1}{n} + \frac{1}{m} < \frac{2}{N} = \varepsilon,$$

as required.

2. $(-1)^n$ is not Cauchy: We have $|a_n - a_{n+1}| = 2$, so for $\varepsilon < 2$ no N can satisfy (7).

3. Problem: Show that a decimal expansion is Cauchy.

These examples indicate that in $\mathbb{R}$ or $\mathbb{C}$ convergent sequences are Cauchy and vice versa. One of these directions is simple to prove:

Proposition 11. *If a sequence converges then it is a Cauchy sequence.*

Proof. We generalize the proof of Example 1. If $a_n \to a$ then for each $\varepsilon > 0$ there is $N \in \mathbb{N}$ such that

$$|a_n - a| < \frac{\varepsilon}{2} \quad \text{for all } n \geq N.$$

Hence we have for all $n, m \geq N$ that

$$|a_n - a_m| = \big|(a_n - a) - (a_m - a)\big| \overset{\Delta\text{-inequ.}}{\leq} |a_n - a| + |a_m - a| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

□

To prove the less obvious converse, we first need to introduce the Bolzano-Weierstrass theorem. The key technical problem is: Which method lets us construct the unknown limit of the sequence?

2.2. **The Bolzano-Weierstrass theorem.** This theorem is the most subtle fact we discuss in the present part. In order to state it, we need a definition first.

Definition. Let $(a_n)_{n\in\mathbb{N}}$ be a sequence, and $(n_k)_{k\in\mathbb{N}}$ be a strictly increasing sequence of natural numbers, that is, $n_1 < n_2 < n_3 < \ldots$. Then the sequence

$$a_{n_1}, a_{n_2}, a_{n_3}, \ldots = (a_{n_k})_{k\in\mathbb{N}}$$

is called a *subsequence [Teilfolge]* of (a_n).

Examples. 1. The sequence $(a_n)_{n\in\mathbb{N}} = 0, 2, 0, 2, 0, 2, \ldots$ has constant subsequences: Namely the odd index sequence $(n_k)_{k\in\mathbb{N}} = (2k-1)_{k\in\mathbb{N}} = 1, 3, 5, \ldots$ defines the subsequence $0, 0, 0, \ldots$, while the even numbers $(n_k) = 2, 4, 6, \ldots$ yield the subsequence $2, 2, 2, \ldots$.

2. Even numbers, prime numbers, and squares are subsequences of the natural numbers (describe $(n_k)_{k\in\mathbb{N}}$!).

3. If a seqence (a_n) converges to a, then each subsequence (a_{n_k}) converges to a as well. To see this, note that $n_k \geq k$, and so if $|a_n - a| < \varepsilon$ for all $n \geq N$ then certainly $|a_{n_k} - a| < \varepsilon$ for all $k \geq N$. We will use this fact in the following proof.

We know that a convergent sequence is bounded. Conversely, a bounded sequence need not converge, as $a_n := (-1)^n$ shows. Still, for such a sequence we can make the following convergence statement:

Theorem 12 (Bolzano-Weierstrass). *Each bounded (real or complex) sequence contains a convergent subsequence.*

Examples. 1. For the sequence $a_n := (-1)^n$ there are subsequences with different limits: For even indices $n_k = 2k$ we have $a_{n_k} \to 1$ as $k \to \infty$; for odd indices $n_k = 2k-1$ we have $a_{n_k} \to -1$.

2. Show by counterexample that this statement fails for $\mathbb{Q}$.

Proof. Let us first consider the case that the sequence (x_n) is real. By assumption, $|x_n| \leq C$. We proceed in two steps:
1. We determine a "possible limit a" for a subsequence.
2. We construct a subsequence converging to the chosen value a.

1. We claim there is a sequence of nested intervals $\big(I_n = [a_n, b_n]\big)$ for $n \in \mathbb{N}$ subject to:
(i) Each interval I_n has length $|b_n - a_n| = \frac{2C}{2^{n-1}}$.

(ii) Each interval I_n contains infinitely many x_k, that is, $x_k \in I_n$ for infinitely many indices k.

To achieve this, we use *interval bisection [Intervallhalbierung]*, defined recursively:
• $I_1 = [a_1, b_1] := [-C, C]$. This interval satisfies (i) and (ii).
• Given I_n with (i) and (ii), we consider the bisection $I_n = [a_n, \frac{a_n+b_n}{2}] \cup [\frac{a_n+b_n}{2}, b_n]$. Since I_n contains infinitely many x_k, one of the intervals on the right will do so as well (perhaps, both). We pick such an interval for I_{n+1}. Properties (i) and (ii) hold.

According to the interval nesting property, Thm. 9, there is some number $a \in \mathbb{R}$ contained in all intervals.

2. Let us now choose a subsequence $(x_{n_k})_{k\in\mathbb{N}}$. Again we define recursively:
• We pick $x_{n_1} := x_1 \in I_1$.
• Given $x_{n_k} \in I_k$, we want to choose $x_{n_{k+1}} \in I_{k+1}$. Since I_{k+1} contains x_n for infinitely many indices n, in particular we can pick an index $n_{k+1} > n_k$ such that $x_{n_{k+1}} \in I_{k+1}$.

Now we show $x_{n_k} \to a$. Indeed, as $x_{n_k}, a \in I_k$,

$$|x_{n_k} - a| \leq |b_k - a_k| = \frac{2C}{2^{k-1}} \to 0 \qquad \text{as } k \to \infty.$$

Finally, let us consider the case of a complex sequence (z_n). We will apply the real result first to the real part $\operatorname{Re} z_n$, then to the imaginary part $\operatorname{Im} z_n$. Note $|\operatorname{Re} z_n| \leq |z_n|$, so that $(\operatorname{Re} z_n)_{n\in\mathbb{N}}$ is a bounded real sequence. Thus the previous proof gives us a convergent subsequence $(\operatorname{Re} z_{n_k})_{k\in\mathbb{N}}$. Now consider the corresponding sequence of imaginary parts, $(\operatorname{Im} z_{n_k})_{k\in\mathbb{N}}$, which is again a bounded real sequence. To this sequence we can apply the real statement once more. We obtain a further convergent subsequence, $(\operatorname{Im} z_{n_{k_\ell}})_{\ell\in\mathbb{N}}$. For this choice of subsequence, both real and imaginary part converge. By Prop. 8, the sum $(z_{n_{k_\ell}} = \operatorname{Re} z_{n_{k_\ell}} + i \operatorname{Im} z_{n_{k_\ell}})_{\ell\in\mathbb{N}}$ also converges; this is a subsequence as desired. □

`9. Vorlesung, Dienstag, 14.11.06 (Ü 4)` ________________________________

2.3. **Cauchy sequences converge.**

Theorem 13. *A sequence of (complex or real) numbers is a Cauchy sequence if and only if it converges.*

Note, however, that the Cauchy sequence of rational numbers

$$3, \quad 3.1, \quad 3.14, \quad 3.141, \quad 3.1415, \quad \ldots$$

does not converge in $\mathbb{Q}$.

Proof. In view of Prop. 11 we only need to show that for a Cauchy sequence there exists an $a \in \mathbb{C}$ such that $a_n \to a$.

We first show (a_n) is bounded. Indeed, for $n, m \geq N$ we have $|a_n - a_m| < 1$ and consequently $|a_n| < |a_N| + 1$. Thus $C := \max\{|a_1|, \ldots, |a_{N-1}|, |a_N| + 1\}$ is a bound.

By the Bolzano-Weierstrass theorem, $(a_n)_{n\in\mathbb{N}}$ has a subsequence $(a_{n_k})_{k\in\mathbb{N}}$ which converges to $a \in \mathbb{C}$. We claim that the entire sequence converges to a. Let $\varepsilon > 0$ and choose N' with $|a_n - a_m| < \frac{\varepsilon}{2}$ for $n, m > N'$. Moreover, choose an index $n_k \geq N'$ with $|a_{n_k} - a| < \frac{\varepsilon}{2}$. Then, as desired,

$$|a_n - a| \leq \left|a_n - a_{n_k}\right| + \left|a_{n_k} - a\right| < \varepsilon \qquad \text{for all } n > N'.$$

Finally note that if the sequence is real, then the limit a is real by Proposition 6. ☐

The previous theorem ultimately rests on the completeness of the real numbers. Indeed, an inspection of the proofs we have given indicates that each of the following statements for $\mathbb{R}$ was used to prove the next:

0. existence of the supremum of bounded sets (def. of completeness)
1. bounded monotone sequences converge, Thm. 7
2. nested intervals contain a common point, Thm. 9
3. Bolzano-Weierstrass Theorem
4. Cauchy sequences converge.

Convince yourself that none of these properties holds for $\mathbb{Q}$!

These properties are not equivalent, however. Let 2', 3', 4' be the above properties 2, 3, 4 augmented with "and the Archimedean property holds". Then for $\mathbb{R}$ (or any ordered field) it can be shown that 0 $\Leftrightarrow$ 1 $\Leftrightarrow$ 2' $\Leftrightarrow$ 3' $\Leftrightarrow$ 4'. Put in another way, for any ordered field with the Archimedean property, our original statements 0. to 4. are equivalent.

Since 3. and, in particular, 4. can be generalized to other sets besides $\mathbb{R}$, let us define:

Definition. A field F with modulus is called *complete* if each sequence $(a_n) \in F$ which is Cauchy, $|a_n - a_m| \to 0$ as $n, m \to \infty$, converges to some number $a \in F$.

Our new definition of completeness is advantageous as it works also for many further spaces encountered in mathematics. By Thm. 13 we certainly have:

Corollary 14. *The set of complex numbers $\mathbb{C}$ is complete.*

2.4. **Outline: Existence of the real numbers.** In Thm. I.5 of Part 1 we introduced the real numbers $\mathbb{R}$ by claiming they satisfy certain properties. It could be that we have asked for too much, in which case the real numbers would simply not exist. Thus the task is to construct the real numbers $\mathbb{R}$ from the rational numbers $\mathbb{Q}$.

The idea is to characterize a real number as a Cauchy sequence of rational numbers converging to it. But there are many such sequences. To define the set of all sequences converging to the same number, let us say (without refering to the limit!):

Definition. We call two Cauchy sequences of rational numbers *equivalent* if their difference is a null sequence, that is,

$$(a_n) \sim (b_n) \quad :\Longleftrightarrow \quad a_n - b_n \to 0 \text{ as } n \to \infty.$$

Examples. 1. The two sequences $0, 0, 0, \ldots$ and $1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \ldots$ are equivalent, but they are not equivalent to $1, 1, 1, \ldots$.

2. The two sequences $(a_n) = 3,\ 3.1,\ 3.14,\ 3.141\ \ldots$ and $(b_n) = 4,\ 3.2,\ 3.15,\ 3.142\ \ldots$ are equivalent, as their difference $1,\ 0.1,\ 0.01,\ 0.001\ \ldots$ is null. The same holds for the boundaries of arbitrary nested intervals $I_n = [a_n, b_n]$.

3. All rational null sequences are equivalent.

4. Two sequences which differ in finitely many elements are equivalent.

It is not hard to show that $\sim$ is an *equivalence class/relation [Äquivalenzrelation]*, that is, the following three properties hold for all x, y, z. (i) *Reflexivity:* $x \sim x$; (ii) *symmetry:* $x \sim y$ implies $y \sim x$; (iii) *transitivity:* if $x \sim y$ and $y \sim z$ then $x \sim z$ (see tutorial). Equivalence relations will be studied more closely in linear algebra (notion of a quotient vector space). The three properties make sets of equivalent elements, called *(equivalence) classes*, well-defined. Here, the sets are sets of Cauchy sequences whose pairwise difference are null, and each class is a real number:

Definition (Cantor 1883)**.** A *real number* [*reelle Zahl*] is a set of equivalent Cauchy sequences of rational numbers with respect to $\sim$.

Examples. 1. The set of Cauchy sequences equivalent to $0, 0, 0, \ldots$, that is, the set of rational null sequences, is defined to be the real number 0; it will turn out as the neutral element of addition of $\mathbb{R}$.

2. Similarly, for $q \in \mathbb{Q}$ the set of rational Cauchy sequences equivalent to the constant sequence $q, q, q, \ldots$ defines the rational number $q \in \mathbb{R}$.

3. The real number π denotes the set of all sequences equivalent to $3, 3.1, 3.14, \ldots$.

The arithmetic operations $+, \cdot$, and the order $<$, can then be defined on the classes of rational Cauchy sequences, that is, on the real numbers. It can then be proven they satisfy the properties of Thm. I.5. We skip the somewhat lengthy details here (see, for instance, [E]).

Remark. There are other ways to define the real numbers, namely *Dedekind cuts [Dedekindsche Schnitte]* of rational numbers, or nested intervals of rational numbers; see [D] or [E]. These definitions do not have the power to generalize.

Let us summarize the number system. The existence of $\mathbb{N}$ had to be postulated as an axiom. The integers $\mathbb{Z}$ and $\mathbb{Q}$ are constructed as pairs of natural or integer numbers. Cauchy sequences of rational numbers define the real numbers $\mathbb{R}$. Pairs of real numbers give the complex numbers $\mathbb{C}$.

3. SERIES

A series is another name for an infinite sum. Later we shall introduce many functions as infinite sums: the exponential function, trigonometric functions, etc. Thus we want to investigate series in general.

3.1. **Partial sums and convergence.** The most prominent example of a series is perhaps the exponential function $\exp(z) = 1 + z + \frac{z^2}{2!} + \frac{z^3}{3!} + \ldots$. For each $z \in \mathbb{C}$, we regard it the limit of the sequence (s_n) of numbers $s_1 = 1$, $s_2 = 1 + z$, $s_3 = 1 + z + \frac{z^2}{2!}$, ... Similarly in general:

Definition. (i) Let $(a_n)_{n\in\mathbb{N}}$ be a (complex) sequence. Then a *series [Reihe]* is the sequence $(s_n)_{n\in\mathbb{N}}$ of *partial sums*

$$s_n := a_1 + \ldots + a_n.$$

Usually we write $\sum_{n=1}^{\infty} a_n$ for the sequence $(s_n)_{n\in\mathbb{N}}$, and call a_n its *terms [Summanden]*.
(ii) In case the series (s_n) converges to $s \in \mathbb{C}$ we write

$$\sum_{n=1}^{\infty} a_n := \lim_{n\to\infty} \sum_{k=1}^{n} a_k = \lim_{n\to\infty} s_n = s.$$

Note. In the convergent case, the notation $\sum_{n=1}^{\infty} a_n$ has two different meanings:
- The sequence of partial sums $(a_1 + \ldots + a_n)_{n\in\mathbb{N}}$,
- a number $s \in \mathbb{C}$, namely the limit of the partial sums; it is also called the *value [Wert]* of the series.

Examples. 1. *Decimal expansion*: $3.14\ldots = 3 + \frac{1}{10} + \frac{4}{100} + \ldots$. We will study these series in more detail below.

2. We claim $\sum_{n=1}^{\infty} \frac{1}{n(n+1)} = 1$, that is, we claim for the partial sums

$$s_n := \frac{1}{1\cdot 2} + \frac{1}{2\cdot 3} + \ldots + \frac{1}{n(n+1)} \quad \to \quad 1 \quad \text{as} \quad n \to \infty.$$

Proof: Writing

$$\frac{1}{n(n+1)} = \frac{-(n^2-1)+n^2}{n(n+1)} = -\frac{n-1}{n} + \frac{n}{n+1}, \qquad \text{for } n \in \mathbb{N},$$

we see we can apply a telescope sum trick:

$$\begin{aligned} s_n &= \Big(-0 + \frac{1}{2}\Big) + \Big(-\frac{1}{2} + \frac{2}{3}\Big) + \Big(-\frac{2}{3} + \frac{3}{4}\Big) + \ldots + \Big(-\frac{n-1}{n} + \frac{n}{n+1}\Big) \\ &= -0 + \frac{n}{n+1} = \frac{1}{1+\frac{1}{n}} \quad \to \quad 1 \qquad \text{as } n \to \infty. \end{aligned}$$

If we are careless, we can easily run into contradictions:

$$0 = (1-1) + (1-1) + \ldots = 1 + (-1+1) + (-1+1) + \ldots = 1$$

In naive language, infinite sums are not associative. Thus only manipulations stipulated by the limit theorems for sequences are admissable.

`10. Vorlesung, Donnerstag, 16.11.06 (T5)`

The space $\{(a_n) : \sum a_n \text{ is convergent }\}$ is a vector space, on which the value of the series is a linear functional. This follows from the limit theorems for sequences, applied to series.

A series in $\mathbb{C}$ is an infinite sum of vectors a_n in the complex plane (picture it with arrows!). If the sum converges, then the length of these vectors must be a null sequence:

Theorem 15. *If $\sum_{n=1}^{\infty} a_n$ converges then $a_n \to 0$ as $n \to \infty$.*

Proof. We have $a_n = s_n - s_{n-1}$ for $n \geq 2$ and thus, using $s_n = \sum_{k=1}^{n} a_k \to s$,

$$\lim_{n\to\infty} a_n = \lim_{n\to\infty} (s_n - s_{n-1}) = \lim_{n\to\infty} s_n - \lim_{n\to\infty} s_{n-1} = s - s = 0.$$

□

Does the converse of the theorem hold? This is not the case:

Example. The *harmonic series*

$$1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \ldots.$$

has terms $\frac{1}{n}$ forming a null sequence. Nevertheless, the partial sums are not bounded. Indeed, for a subsequence,

$$\begin{aligned} s_{2^n} &= 1 + \frac{1}{2} + \frac{1}{3} + \ldots + \frac{1}{2^n} \\ &= 1 + \frac{1}{2} + \underbrace{\Big(\frac{1}{3} + \frac{1}{4}\Big)}_{\geq 1/2} + \underbrace{\Big(\frac{1}{5} + \ldots + \frac{1}{8}\Big)}_{\geq 1/2} + \ldots + \underbrace{\Big(\frac{1}{2^{n-1}+1} + \ldots + \frac{1}{2^n}\Big)}_{\geq 1/2} \\ &\geq 1 + \frac{n}{2} \to \infty. \end{aligned}$$

This unboundedness means the harmonic series cannot converge, a result known to the French scientist and bishop Nikolaus von Oresme in the 14th century. Moreover (s_n) is increasing, and hence our argument shows that $\sum \frac{1}{n}$ diverges to infinity; as for sequences we denote this symbolically by $\sum \frac{1}{n} = \infty$. (Give a proof for this fact using the Cauchy test!)

The most important series will turn out to be the following:

Theorem 16. *Let $z \in \mathbb{C}$. The* geometric series $1+z+z^2+z^3+\ldots$ *converges for all $|z| < 1$ to*

$$\sum_{n=0}^{\infty} z^n = \frac{1}{1-z},$$

while for $|z| \geq 1$ the series diverges.

Proof. The geometric sum I(5) gives

$$s_n = 1 + z + z^2 + \ldots + z^n = \frac{1 - z^{n+1}}{1-z} \qquad \text{for } z \neq 1. \tag{8}$$

When $|z| < 1$ we see from Sect. 1.1, ex. 2., that $z^n \to 0$ as $n \to \infty$; hence $\lim s_n = \frac{1}{1-z}$.

For $|z| \geq 1$ also $|z^n| = |z|^n \geq 1$, and so (z^n) is not a null sequence. Theorem 15 gives that $\sum z^n$ cannot converge. ☐

Example. $|\frac{i}{2}| < 1$ and hence

$$\sum_{n=0}^{\infty} \Big(\frac{i}{2}\Big)^n = 1 + \frac{i}{2} - \frac{1}{4} - \frac{i}{8} + \ldots = \frac{1}{1 - \frac{i}{2}} = \frac{1 + \frac{i}{2}}{1 - \frac{i^2}{4}} = \frac{1 + \frac{i}{2}}{\frac{5}{4}} = \frac{4}{5} + \frac{2}{5} i.$$

3.2. **Series of real numbers.** There are two useful tests for convergence of real series. The first one can deal with series whose sign alternates:

Theorem 17 (Leibniz). *Let $(a_n)_{n \geq 0}$ be a monotone decreasing null sequence, $a_0 \geq a_1 \geq a_2 \geq \ldots \geq 0$. Then the* alternating sum $\sum_{n=0}^{\infty} (-1)^n a_n$ *converges.*

Example. The *alternating harmonic series*

$$1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} \pm \dots.$$

converges. As we shall see later (Sect. IV 5.5), the limit is $\log 2$.

Proof. The idea is to see the alternating series defines an interval nesting whose common point is the limit.

To see this, consider odd and even partial sums,

$$A_n := s_{2n+1} = a_0 - a_1 + \ldots + a_{2n} - a_{2n+1} \quad \text{and} \quad B_n := s_{2n} = a_0 - a_1 + \ldots + a_{2n},$$

where $n \in \mathbb{N}_0$. The equations

$$A_n = A_{n-1} + \underbrace{a_{2n} - a_{2n+1}}_{\geq 0}, \quad B_n = B_{n-1} \underbrace{-a_{2n-1} + a_{2n}}_{\leq 0} = A_n + \underbrace{a_{2n+1}}_{\searrow 0} \qquad \text{for } n \in \mathbb{N}$$

prove the following facts: (A_n) increases monotonously, (B_n) decreases monotonously, $A_n \leq B_n$, and $(B_n - A_n)$ is a null sequence.

Hence $[A_n, B_n]$ is a sequence of nested intervals, containing a common point s, and so

$$s = \lim_{n\to\infty} A_n = \lim_{n\to\infty} B_n = \lim_{n\to\infty} s_n = \sum_{n=0}^{\infty} (-1)^n a_n. \qquad \square$$

A second test applies to real series whose terms all have the same sign:

Theorem 18. *A series $\sum_{n=1}^{\infty} a_n$ with $a_n \geq 0$ converges if and only if its partial sums are bounded.*

Proof. The assumption $a_n \geq 0$ means that the sequence of partial sums (s_n) is increasing. Thus the statement follows from Theorem 7. $\square$

The boundedness criterion can be used for a comparison test for convergence:

Corollary 19 (Majorization of real series)**.** *Suppose $(x_n)_{n\in\mathbb{N}}$ is a real sequence for which there exists a convergent series $\sum_{n=1}^{\infty} a_n$ of real numbers $a_n \geq 0$ with*

$$0 \leq x_n \leq a_n \qquad \text{for all } n \in \mathbb{N}.$$

Then $\sum_{n=1}^{\infty} x_n$ also converges and $\sum_{n=1}^{\infty} x_n \leq \sum_{n=1}^{\infty} a_n$.

We call a_n a *majorant* of x_n.

`11. Vorlesung, Dienstag, 21.11.06 (Ü 5)` ______

Proof. We consider partial sums. By the theorem, $\sum_{k=1}^{n} a_k \le C$ for some $C \in \mathbb{R}$ and so

$$0 \le \sum_{k=1}^{n} x_k \le \sum_{k=1}^{n} a_k \le C. \tag{9}$$

But applying the theorem once again, we see that $\sum x_k$ must converge.

To prove the stated inequality, we also must argue in terms of partial sums: The inequality for the partial sums (9) is preserved in the limit by Prop. 8(i). $\square$

Problem. (Minorization) Suppose for a real series $\sum a_n$ there exists a sequence (x_n) with $a_n \ge x_n \ge 0$ such that $\sum x_n$ is divergent. Prove that $\sum a_n$ diverges as well.

Example. Let $s \in \mathbb{N}$, $s \ge 2$. Then we claim

$$\zeta(s) := \sum_{n=1}^{\infty} \frac{1}{n^s} = 1 + \frac{1}{2^s} + \frac{1}{3^s} + \dots$$

converges. Indeed, first of all

$$\frac{1}{n^2} \le \frac{2}{n(n+1)} \quad \text{for all } n \in \mathbb{N},$$

and $\sum_{n=1}^{\infty} \frac{2}{n(n+1)}$ is convergent (see the first example in 3.2). Thus $\sum_{n=1}^{\infty} \frac{1}{n^2}$ is majorized by a convergent series and therefore, by the corollary, converges itself. Moreover,

$$\frac{1}{n^s} \le \frac{1}{n^2} \qquad \text{for all } n \in \mathbb{N},$$

and so applying the corollary once again we obtain that $\sum_{n=1}^{\infty} \frac{1}{n^s}$ converges for $s \ge 2$.

Remark. The *zeta-function* $\zeta(s)$ can be defined not only for $s = 2, 3, \dots$, but for most $s \in \mathbb{C}$. It is a famous function in mathematics, which surprisingly is related to many questions about prime numbers. One of the big open problems in mathematics is the *Riemann hypothesis*, stated by Riemann in 1859, that the zeros $\zeta(z) = 0$ for which z has positive real part all lie on the line $\operatorname{Re} z = \frac{1}{2}$. It is among a list of 7 problems for which a price of a million dollar has been set out for a solution, see: `www.claymath.org/millennium/Riemann_Hypothesis/`

3.3. **Decimal expansions.** In antiquity, the only numbers that could be represented arithmetically were rational numbers $\mathbb{Q}$ or *proportions*. Geometry was considered superior to algebra as it could deal with "all" numbers. Since its invention in medieval time, decimal representations have changed the view of mathematics. Nowadays many people believe that real numbers and decimal representations are identical, so that the nonuniqueness of the type $1 = 0.999\dots$ poses a problem. This problem is easy to resolve once decimal expansions are regarded as series.

Let us first deal with real numbers which are not negative.

Definition. A *decimal expansion* is a series

$$\sum_{n=0}^{\infty} \frac{d_n}{10^n} = d_0 + \frac{d_1}{10} + \frac{d_2}{10^2} + \ldots,$$

with $d_0 \in \mathbb{N}_0$ and digits $d_n \in \{0, \ldots, 9\}$ for $n \geq 1$.

It is a nice exercise to show that the partial sums of a decimal expansion are Cauchy.

Then each decimal expansion defines some real number, and for each real number there is at least one decimal representation:

Lemma 20. (i) *Each decimal expansion* $\sum_{n=0}^{\infty} \frac{d_n}{10^n}$ *converges to a number*

$$x = \sum_{n=0}^{\infty} \frac{d_n}{10^n} \in [0, \infty). \tag{10}$$

(ii) *For each real number* $x \in [0, \infty)$ *there exists* $d_0 \in \mathbb{N}_0$ *and a sequence* $d_n \in \{0, \ldots, 9\}$ *for* $n \geq 1$ *such that* (10) *holds.*

Proof. (i) We can majorize:

$$0 \leq \sum_{n=0}^{\infty} \frac{d_n}{10^n} \leq d_0 + \sum_{n=1}^{\infty} \frac{9}{10^n} \overset{I(5)}{=} d_0 + \frac{9}{10} \frac{1}{1 - \frac{1}{10}} = d_0 + 1$$

(Our estimate says that $0.99\ldots9$, with n digits, is indeed less than 1). Cor. 19 gives the claim.

(ii) To find x, let us define the digits d_n by an interval nesting: There exists $d_0 \in \mathbb{N}$ with $d_0 \leq x < d_0 + 1$. Then we define recursively: Suppose $d_1, \ldots, d_n$ are constructed, such that

$$a_n := d_0 + \frac{d_1}{10} + \ldots + \frac{d_n}{10^n} \leq x < d_0 + \frac{d_1}{10} + \ldots + \frac{d_n + 1}{10^n} =: b_n \tag{11}$$

Then subdivide the interval $I_n = [a_n, b_n)$ into the 10 halfopen disjoint intervals

$$[a_n, a_n + \frac{1}{10^{n+1}}),\ [a_n + \frac{1}{10^{n+1}}), a_n + \frac{2}{10^{n+1}}),\ \ldots [a_n + \frac{9}{10^{n+1}}), b_n),$$

whose union gives I_n. One of these ten intervals contains x; call it $I_{n+1} = [a_n, b_n)$; this constructs d_{n+1} such that (11) holds. But $a_1 \leq \ldots \leq a_n \leq x < b_n \leq \ldots \leq b_1$ and so By the interval nesting property, $\bigcap [a_n, b_n]$ contains x, and so $\lim_{n\to\infty} \sum_{k=0}^{n} \frac{d_n}{10^n} = x$, meaning that (10) holds. $\square$

Problems. 1. State the common decimal expansion for negative numbers – why have we not included this case into our statement?
2. Give several examples of decimal expansions which define the same real number x. Which of these expansions does our algorithm (proof of part (ii)) generate?

A periodic decimal expansion is, up to an additive constant, a geometric series; it always defines a rational number. For example,

$$2.\overline{34} := 2.343434\ldots = 2 + \frac{34}{10^2} + \frac{34}{10^4} + \frac{34}{10^6} + \cdots = 2 + \frac{34}{100}\left(1 + \frac{1}{100} + \frac{1}{100^2} + \ldots\right)$$
$$= 2 + \frac{34}{100} \cdot \frac{1}{1 - \frac{1}{100}} = 2 + \frac{34}{100} \cdot \frac{100}{99} = 2 + \frac{34}{99} = \frac{232}{99}.$$

Problem. Prove that conversely each rational number gives a finite or periodic decimal expansion.

3.4. **Absolute convergence and comparison tests.** We wish to extend the comparison test Corollary 19 to complex series.

Definition. A series $\sum_{n=1}^{\infty} a_n$ is *absolutely convergent* if $\sum_{n=1}^{\infty} |a_n|$ converges.

Examples. 1. $\sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n}$ is not absolutely convergent as $\sum_{n=1}^{\infty} \frac{1}{n}$ diverges.
2. Let $z \in \mathbb{C}$. Then $\sum_{n=0}^{\infty} z^n$ converges absolutely for $|z| < 1$, since then $\sum_{j=0}^{n} |z|^j = \frac{1-|z|^{n+1}}{1-|z|} \to \frac{1}{1-|z|}$ as $n \to \infty$. This means that for a geometric series, convergence and absolute convergence are equivalent.

First we show that absolute convergence implies standard convergence:

Theorem 21. *Let $\sum_{n=1}^{\infty} a_n$ be absolutely convergent, where $a_n \in \mathbb{C}$. Then (i) the series itself, $\sum_{n=1}^{\infty} a_n$, converges and (ii) the* triangle inequality *holds:*

$$\left|\sum_{n=1}^{\infty} a_n\right| \le \sum_{n=1}^{\infty} |a_n| \tag{12}$$

Proof. (*i*) To understand the idea, note that the limit of $\sum a_n$ is not computable from the limit $\sum |a_n|$. Hence we must avoid mentioning the limit. Thus we employ the Cauchy test (7) which reads, when applied to a series,

$$|s_n - s_k| = \left|\sum_{j=1}^{n} a_j - \sum_{i=1}^{k} a_j\right| = \left|a_{k+1} + \ldots + a_n\right| \to 0 \qquad \text{as } n > k \to \infty.$$

By assumption, the series of lengths $\sum_{n=1}^{\infty} |a_n|$ converges, and so by Thm. 13 it is Cauchy. Thus we have

$$\left|a_{k+1} + \ldots + a_n\right| \overset{\Delta\text{-inequ.}}{\le} |a_{k+1}| + \ldots + |a_n| \to 0 \qquad \text{as } n > k \to \infty$$

But that means the Cauchy test for $\sum |a_n|$ implies the Cauchy test for $\sum a_n$. By Thm. 13 again, this implies $\sum a_n$ converges. (Note we have applied the completeness of $\mathbb{C}$.)

(*ii*) Upon induction, the standard triangle inequality gives $\left|\sum_{j=1}^{n} a_j\right| \le \sum_{j=1}^{n} |a_j|$. By part (*i*), both sides of the inequality converge as $n \to \infty$. By Prop. 8(*ii*) the inequality is preserved in the limit. $\square$

Remark. A *rearrangement of a series [Umordnung]* is a change in the order of summation. For absolutely convergent series, the limit remains unchanged upon rearrangement. For convergent series which are not absolutely convergent (*conditionally convergent series*) the limit surprisingly can change, however. Let us rephrase this fact by saying that the commutative law is not automatic for convergent series. (See problems).

Since absolute values are non-negative we can test for absolute convergence by appealing to Theorem 18:

Corollary 22 (Majorization of series)**.** *Suppose that for a series $\sum_{n=1}^{\infty} a_n$ there exists a convergent real series $\sum_{n=1}^{\infty} x_n$ (called* majorant*) which satisfies $|a_n| \le x_n$ for all $n \in \mathbb{N}$. Then $\sum_{n=1}^{\infty} a_n$ converges absolutely.*

The geometric series $1 + q + q^2 + \ldots$ converges for real $q \in (-1, 1)$. Using it as a majorant gives:

Theorem 23 (Ratio test *[Quotientenkriterium]*)**.** *Let $\sum_{n=0}^{\infty} a_n$ be a series with $a_n \neq 0$ for all $n \ge N \in \mathbb{N}_0$. Suppose there is a number $0 < q < 1$, such that*

$$\frac{|a_{n+1}|}{|a_n|} \le q \qquad \text{for all } n \ge N. \tag{13}$$

Then the series $\sum_{n=0}^{\infty} a_n$ converges absolutely.

Proof. We consider the case $N = 0$ first. By induction,

$$|a_n| \le q^n |a_0| \qquad \text{for all } n \ge 0. \tag{14}$$

Thus $q^n|a_0|$ is a majorant for a_n, which is convergent due to

$$\sum_{n=0}^{\infty} |a_0| q^n = |a_0| \sum_{n=0}^{\infty} q^n \overset{\text{Thm.16}}{=} \frac{|a_0|}{1-q}.$$

According to Cor. 22 this implies that $\sum_{n=1}^{\infty} |a_n|$ is convergent.

Likewise, for the general case $N \in \mathbb{N}$ we majorize with the geometric series starting at the N-th term. We obtain convergence with

$$\sum_{n=0}^{\infty} |a_n| \;=\; |a_0| + \ldots + |a_{N-1}| + \sum_{n=N}^{\infty} |a_n| \;\le\; |a_0| + \ldots + |a_{N-1}| + \frac{|a_N|}{1-q}.$$

$\square$

Example. Consider the series
$$\sum_{n=1}^{\infty} n^2 \Big(\frac{1-i}{3+4i}\Big)^n.$$
So if a_n denotes the n-th term, we have a ratio
$$\frac{|a_{n+1}|}{|a_n|} = \frac{(n+1)^2}{n^2} \cdot \frac{|1-i|}{|3+4i|} = \frac{(n+1)^2}{n^2} \cdot \underbrace{\frac{\sqrt{2}}{5}}_{<1/3} < \frac{(n+1)^2}{3n^2} \overset{(*)}{<} \frac{2}{3} =: q \qquad \text{for } n \geq 3.$$
At $(*)$ we used that $(n+1)^2 = n^2 + 2n + 1 < 2n^2 \iff 2n+1 < n^2$ holds provided $n \geq 3$. Hence the series passes the ratio test with $q = \frac{2}{3}$ and $N := 3$.

Remarks. 1. (Warning:) If the ratio only satisfies $\frac{|a_{n+1}|}{|a_n|} < 1$ then convergence cannot be asserted(!). To see this, consider the harmonic series $\sum_{n=1}^{\infty} \frac{1}{n}$. It gives rise to a ratio $\frac{1/(n+1)}{1/n} = \frac{n}{n+1} < 1$ but is divergent.
2. The ratio test is sufficient for the convergence but not necessary: The series $\sum \frac{1}{n^2}$ is convergent, but it does not satisfy the ratio test.

Problems. 1. Show that each of the following tests also implies that $\sum a_n$ converges absolutely:
(i) *Limit version of ratio test*: $\frac{|a_{n+1}|}{|a_n|} \to \tilde{q}$ with $0 \leq \tilde{q} < 1$ (hint: set $q := \frac{1}{2}(1+\tilde{q})$).
(ii) *Root test*: $\sqrt[n]{|a_n|} < q$ for $0 < q < 1$ (show (14)!).
2. Show the root test is more general then the ratio test. Think about reordering a geometric series!

`12. Vorlesung, Donnerstag, 23.11.06 (T6)` ______________________________

SUMMARY

Sequences have a rigorous definition of convergence: It is quantitative and avoids mentioning inifinity as such. There are three important facts which we formulated for sequences of real numbers – each of them is in fact equivalent to their completeness: First, monotone sequences converge if and only if they are bounded. Second, interval nestings contain precisely one point. Third, for bounded sequences the Bolzano-Weierstrass guarantees the convergence of subsequences.

Then the notion of *Cauchy sequence* was introduced. In complete spaces, the Cauchy test guarantees the convergence of a sequence without refering to its limit. We confirmed this for $\mathbb{R}$ and $\mathbb{C}$ by applying the Bolzano-Weierstrass theorem. The notion of Cauchy sequence is significant in two respects:
1. Our original notion of the completeness of $\mathbb{R}$ was in terms of the supremum. We

extended it to more general spaces by requiring that all Cauchy sequences converges. The latter definition of completeness is not only applicable to $\mathbb{C}$, but to many function spaces encountered in mathematics and physics.
2. The real numbers can be defined in terms of Cauchy sequences. This important definition was only sketched in our lectures, but it is a definition which provides the foundation for all of analysis.

Series were another topic. The notion of convergenge was defined in terms of the convergence of the sequence of partial sums. Nevertheless, there are a number of surprising facts about sequences. For instance, the commutative rule does not hold for the terms of a series (the value of the infinite sum can change upon rearrangement); also, as discussed in the next section, when we wish to multiply two series it is not clear we can apply the distributive law. The key notion, which saves all these desired laws, is *absolute convergence.*

There are just two or three types of particular series whose convergence properties each student should memorize:
1. The *geometric series* $1 + z + z^2 + \dots$ converges and absolutely converges for $|z| < 1$ to the number $\frac{1}{1-z}$; it diverges for $|z| \geq 1$.
2. The series $\frac{1}{1^k} + \frac{1}{2^k} + \frac{1}{3^k} + \dots$ is divergent for $k = 1$ (harmonic series), it converges for $k \geq 2$ as was seen by majorization with the telescope sum $\sum \frac{1}{n(n+1)}$. In fact, by Cauchy condensation (see tutorial), it can be seen to converge for all real $k > 1$; nevertheless, the value cannot be computed without developing further theory first.
3. *Alternating real series*, which monotonically decrease in modulus, converge (Leibniz).

The key tool to prove convergence of series is majorisation. As a special case, we reformulated majorization with the geometric series as the *ratio test.*

Part 3. Functions: Continuity and special functions

Functions are the primary object of study in analysis. We consider here functions of one variable, which in real life often arise as time dependent data, such as height of a free fall (Galilei). temperature, velocity, or stock prices.

Up to the 19th century, a function was what could be stated in terms of an explicit description, like x^2, exp, sin. But more and more functions arose in terms of rather implicit descriptions, and so a more general notion, going back to Dirchlet in the 1850's, is:

Definition. Let D be a subset of $\mathbb{C}$. A *function* on D is a mapping $f\colon D \to \mathbb{C}$. We call

$$\Gamma(f) := \big\{\big(z, f(z)\big) \colon z \in D\big\} \subset D \times \mathbb{C}$$

the *graph* of f. A particular case are *real functions* $f\colon D \to \mathbb{R}$ where $D \subset \mathbb{R}$.

Examples. 1. The function $p(z) = a_0 + a_1 z + \ldots + a_n z^n$ is called a *polynomial [Polynom]*. If $a_n \neq 0$ we call n the *degree* of p.

2. The *floor function [Gauß-Klammer]*

$$\lfloor . \rfloor \colon \mathbb{R} \to \mathbb{R}, \qquad \lfloor x \rfloor := \sup\{n \in \mathbb{Z} \colon n \leq x\}.$$

Sketch the graph!

Only $\mathbb{R}$ can be ordered. Real functions which respect the ordering are particularly useful:

Definition. Let $D \subset \mathbb{R}$, and $f\colon D \to \mathbb{R}$ a function. Then

$$f \text{ is } \textit{(monotonously)} \left\{ \begin{smallmatrix} \textit{increasing} \\ \textit{decreasing} \end{smallmatrix} \right\} \text{ if } \left\{ \begin{smallmatrix} f(x) \leq f(y) \\ f(x) \geq f(y) \end{smallmatrix} \right\} \text{ for each pair } x < y \text{ in } D.$$

When the inequality is strict, the function is *strictly* monotone *[streng monoton]*.

Examples. 1. $f\colon [0, \infty) = \{x \in \mathbb{R} \colon x \geq 0\} \to \mathbb{R}$, $f(x) = x^2$ is strictly increasing, since $0 \leq x < y$ implies $x^2 \leq xy < y^2$.

2. Similarly, x^n is an increasing function from $\mathbb{R}_0^+$ to $\mathbb{R}$ for each $n \in \mathbb{N}$; for n odd, this holds on all of $\mathbb{R}$.

A function $f\colon D \to E := f(D)$ is called invertible, if there is an *inverse function [Umkehrfunktion]* $f^{-1}\colon E \to D$ with $f \circ f^{-1} = \mathrm{id}_E$ and $f^{-1} \circ f = \mathrm{id}_D$. A strictly monotone function $f\colon D \to \mathbb{R}$ is injective, and hence has an inverse.

1. The exponential function

We want to introduce the most famous mathematical function which is not elementary. We will work directly in the complex setting; if you feel uncomfortable with this choice, assume on a first reading that z is a real number.

1.1. **The exponential series.** For $z \in \mathbb{C}$ let us define the *exponential series* (Newton 1669)

$$\exp(z) := 1 + z + \frac{z^2}{2!} + \frac{z^3}{3!} + \ldots = \sum_{n=0}^{\infty} \frac{z^n}{n!}.$$

Note that $\exp(x) \in \mathbb{R}$ for real x; indeed, in that case all partial sums have vanishing imaginary part.

By the next theorem, the complex exponential series is a function from $\mathbb{C}$ to $\mathbb{C}$, which we will later use to introduce the trigonometric functions $\sin, \cos, \ldots$, as well as the number π.

Theorem 1. *The series* $\exp(z)$ *converges absolutely for each* $z \in \mathbb{C}$*, that is,* $\exp$ *defines a function from* $\mathbb{C}$ *to* $\mathbb{C}$*.*

Proof. For $z = 0$ there is nothing to show. Now fix $z \neq 0$. The n-th term is $a_n := \frac{z^n}{n!}$ and so the ratio test gives

$$\frac{|a_{n+1}|}{|a_n|} = \frac{|z^{n+1}|}{(n+1)!} \cdot \frac{n!}{|z^n|} = \frac{|z|}{n+1}.$$

Hence if we choose $N \in \mathbb{N}$ depending on z such that $|z| \leq \frac{N}{2}$, we obtain for all $n \geq N$

$$\frac{|a_{n+1}|}{|a_n|} \leq \frac{N}{2(n+1)} \leq \frac{1}{2}\frac{n}{n+1} < \frac{1}{2}.$$

Therefore, the exponential series passes the ratio test with $q = \frac{1}{2}$. □

Particular values include $\exp(0) = 1$ and the *Euler number* $\exp(1) = 2.71828\ldots =: e$ (Euler 1736).

Many properties of the exponential function will be derived from the following error bound on the partial sums of $\exp(z)$:

Theorem 2 (Remainder term estimate *[Restgliedabsch"atzung]*)**.** *For* $n \in \mathbb{N}_0$*, let* $R_n(z)$ *be the series defined by*

$$R_n(z) := \frac{1}{n!}z^n + \frac{1}{(n+1)!}z^{n+1} + \ldots = \sum_{k=n}^{\infty} \frac{z^k}{k!}. \tag{1}$$

Then

$$\left|R_n(z)\right| \leq 2\frac{|z|^n}{n!} \qquad \textit{for} \qquad |z| \leq \frac{n+1}{2}.$$

We can write $\exp(z) = s_{n-1}(z) + R_n(z)$. Here s_{n-1} is a polynomial, namely the $(n-1)$-st partial sum, and $R_n(z)$ measures the error of approximating exp by this polynomial.

Proof. We apply the infinite triangle-inequality II(12) to the series (1), and factor out $\frac{|z|^n}{n!}$:

$$\big|R_n(z)\big| \le \frac{|z|^n}{n!}\left[1 + \frac{|z|}{n+1} + \frac{|z|^2}{(n+1)(n+2)} + \dots\right] \le \frac{|z|^n}{n!}\left[1 + \frac{|z|}{n+1} + \Big(\frac{|z|}{n+1}\Big)^2 + \dots\right]$$

Let us now apply the comparison test with the geometric series, this time in a quanitative way: Using our assumption $|z| \le \frac{n+1}{2}$ we can majorize $[\dots]$ with the geometric series $1 + \frac{1}{2} + \frac{1}{4} + \dots = 2$ (that is, we invoke Corollary II.19). This establishes the desired error estimate. $\square$

1.2. **Growth laws.** We want to explain the real exponential function in an off-hand way using derivatives. The two simplest growth laws for a quantity $f(t)$ which depends on the time t are:

1. *Linear growth.* Here f could be the diameter of a tree, the height of sediment deposits, etc. In these cases, the change of the quantity f in time is proportional to a constant, $f' = c$ with a constant $c \in \mathbb{R}$. Thus the growth law is affine linear: $f(t) = ct + a$ where $a, c \in \mathbb{R}$.

2. *Exponential growth.* Consider the size of population, an investment growing with fixed interest rates, or radioactive material (decaying with radiation). Here the quantity changes proportional to its size, that is, $f' = cf$ or $\frac{f'}{f} = c$ for some constant $c \in \mathbb{R}$. We claim the growth law is exponential, $f(t) = a\exp(ct)$ with $a, c \in \mathbb{R}$. Indeed, differentiating the exponential series term by term we obtain

$$\big(\exp(x)\big)' = \left(\sum_{n=0}^{\infty} \frac{1}{n!}x^n\right)' \overset{(*)}{=} \sum_{n=0}^{\infty} \Big(\frac{1}{n!}x^n\Big)' = \sum_{n=1}^{\infty} \frac{n}{n!}x^{n-1} = \sum_{n=1}^{\infty} \frac{1}{(n-1)!}x^{n-1} = \exp(x),$$

and similarly $a\exp(cx)' = c\big(a\exp(cx)\big)$. Finite sums can be differentiated termwise. However, for infinite sums, the step $(*)$ needs justification, which we will provide only later.

Are there any other solutions to the differential equation $f' = cf$? We will show later that all solutions are of the type $f(x) := a\exp(cx)$. Let us show here only that f cannot be polynomial. So let $p(x) = a_0 + \dots + a_n x^n$. Note that if p has degree $n \ge 1$, then p' has degree $n-1$. Thus $p' = cp$ (with $c \ne 0$) can hold only if $p \equiv 0$.

Consequently, the modelling of the simple growth law $f' = cf$ requires the study of an infinite series $\sum \frac{1}{n!}x^n$. This is a non-elementary function in the sense that it can only evaluated by approximation.

So the completion process leading from rational to real numbers has an analogue when dealing with functions: We must "complete" the space of polynomials to the class of infinite series $\lim_{n\to\infty} p_n(x) = \sum_{n=0}^{\infty} a_n x^n$ if we want to solve natural differential equations.

`13. Vorlesung, Dienstag, 28.11.06 (Ü 6)` ___

1.3. **The functional equation for** exp**.** Our goal is to assert that the function exp has the properties of an exponential function. What are these? By definition, we have $a^{17} \cdot a^3 = a^{17+3}$ or, in general, $a^k a^l = a^{k+l}$ for $k, l \in \mathbb{N}$. Similarly, we will show $\exp(z)\exp(w) = \exp(z+w)$; this will then allow us to write $\exp(z) = e^z$.

Since $\exp(z)\exp(w)$ is a product of two series, we will now consider the product of two (complex) series in general. When we multiply finite sums,

$$(a_0 + \ldots + a_k)(b_0 + \ldots + b_l),$$

we obtain $(k+1)\cdot(l+1)$ terms; when we sum them, their order is irrelevant. However, when we multiply two series $\sum_{k=0}^{\infty} a_k \ \sum_{l=0}^{\infty} b_l$ we need to add the infinitely many terms listed in the following table:

$$\begin{array}{lllll} a_0b_0 & a_0b_1 & a_0b_2 & a_0b_3 & \ldots \\ a_1b_0 & a_1b_1 & a_1b_2 & \ldots & \\ a_2b_0 & a_2b_1 & \ldots & & \\ a_3b_0 & \ldots & & & \\ \ldots & & & & \end{array}$$

Remember from the alternating harmonic series that the order of summation does matter when dealing with infinitely many terms. Thus in general the order of summing of the terms will lead to different values (limits). For absolutely convergent series, however, the value is independent of the order of summation.

We will not prove this general fact, but consider a particular order which is convenient: We enumerate all terms along diagonal paths. That is, we first take diagonal sums,

$$c_n := a_n b_0 + \ldots + a_0 b_n \qquad \text{for } n \in \mathbb{N}, \tag{2}$$

and then wish to sum all c_n. Assuming absolute convergence, this particular order of summation gives the series product:

Lemma 3 (Cauchy product of series)**.** *Let $\sum_{n=0}^{\infty} a_n$ and $\sum_{n=0}^{\infty} b_n$ be absolutely convergent complex series, and c_n as in* (2)*. Then* (*i*) *$\sum_{n=0}^{\infty} c_n$ is also absolutely convergent, and* (*ii*)

$$\sum_{n=0}^{\infty} c_n = \sum_{n=0}^{\infty} a_n \ \sum_{n=0}^{\infty} b_n. \tag{3}$$

Without the assumption of absolute convergence this statement is false: In general the infinite sum over all terms $a_i b_j$ will depend on the order of summation chosen!

Proof. (*ii*) We denote the partial sums and their known limits by

$$s_n := a_0 + \ldots + a_n \to s, \qquad t_n := b_0 + \ldots + b_n \to t, \qquad u_n := c_0 + \ldots + c_n.$$

Let us first show (3), that is, $\sum c_n$ converges to st. This means we want to prove that $st - u_n \to 0$ as $n \to \infty$. By Prop. II.3(*iii*), the product of the sequences s_n and t_n converges, that is, $s_n t_n \to st$.Thus it suffices to show $s_n t_n - u_n \to 0$ as $n \to \infty$. Written in this form, the claim involves only finite sums. We investigate them now.

The expression $s_n t_n - u_n$ contains the products $a_i b_j$ for certain pairs of indices. To be explicit, let

$$\square_n := \{(i,j) \in \mathbb{N}_0^2 \colon 0 \leq i,j \leq n\}$$

be a square of indices, and

$$\Delta_n := \{(i,j) \in \mathbb{N}_0^2 \colon i + j \leq n\}$$

be a triangle contained in $\square_n$. We can now write

$$s_n t_n - u_n = \sum_{(i,j)\in\square_n} a_i b_j - \sum_{(i,j)\in\Delta_n} a_i b_j = \sum_{(i,j)\in\square_n\setminus\Delta_n} a_i b_j.$$

Since $\square_{\lfloor n/2\rfloor} \subset \Delta_n$ we have that $\square_n \setminus \Delta_n \subset \square_n \setminus \square_{\lfloor n/2\rfloor}$ for all $n \in \mathbb{N}_0$. This gives

$$\left|s_n t_n - u_n\right| = \left|\sum_{(i,j)\in\square_n\setminus\Delta_n} a_i b_j\right| \overset{\Delta\text{-inequ.}}{\leq} \sum_{(i,j)\in\square_n\setminus\Delta_n} |a_i||b_j| \leq \sum_{(i,j)\in\square_n\setminus\square_{\lfloor n/2\rfloor}} |a_i||b_j|.$$

Let us now use the assumption that $\sum a_n$ and $\sum b_n$ are absolutely convergent. This means the partial sums

$$A_n := |a_0| + \ldots + |a_n| \qquad \text{and} \qquad B_n := |b_0| + \ldots + |b_n|$$

are convergent sequences. Consequently, the product sequence $(A_n B_n)_{n\in\mathbb{N}_0}$ is also convergent. In particular, it is a Cauchy sequence (Prop. II.11), that is,

$$\left|A_n B_n - A_{\lfloor n/2\rfloor} B_{\lfloor n/2\rfloor}\right| \to 0 \qquad \text{as } n \to \infty.$$

The claim follows from combining our results:

$$\left|s_n t_n - u_n\right| \leq \sum_{(i,j)\in\square_n\setminus\square_{\lfloor n/2\rfloor}} |a_i||b_j| = \left|A_n B_n - A_{\lfloor n/2\rfloor} B_{\lfloor n/2\rfloor}\right| \to 0 \qquad \text{as } n \to \infty$$

(i) It remains to show that $\sum c_n$ is absolutely convergent. Note first that the triangle inequality gives

$$|c_n| \leq \sum_{k=0}^{n} |a_k||b_{n-k}| =: d_n.$$

But the series $\sum_{n=0}^{\infty} d_n$ is convergent, with

$$\sum_{n=0}^{\infty} d_n = \sum_{n=0}^{\infty} \Big(\sum_{k=0}^{n} |a_k||b_{n-k}| \Big) = \sum_{n=0}^{\infty} |a_n| \sum_{n=0}^{\infty} |b_n|;$$

this follows from applying the proven Cauchy product formula (3) to the two real series $\sum |a_n|$ and $\sum |b_n|$.

Consequently, $\sum d_n$ is a convergent series, majorizing $\sum |c_n|$. By Corollary II.22 the series $\sum c_n$ converges absolutely. □

Let us now apply the lemma to the exponential series.

Theorem 4 (Functional equation). *For all $z, w \in \mathbb{C}$ we have*

$$\exp(z + w) = \exp(z)\exp(w). \tag{4}$$

Proof. We wish to compute the right hand side of (4) as a Cauchy product: If we set $\sum a_n := \sum \frac{z^n}{n!}$ and $\sum b_n := \sum \frac{w^n}{n!}$ then their Cauchy product has the terms

$$c_n := \sum_{k=0}^{n} \frac{z^k}{k!} \cdot \frac{w^{n-k}}{(n-k)!} = \frac{1}{n!} \sum_{k=0}^{n} \binom{n}{k} z^k w^{n-k} \overset{\mathrm{I}(4)}{=} \frac{1}{n!}(z + w)^n \quad \text{for all } n \in \mathbb{N}_0,$$

using the Binomial Theorem.

The exponential series is absolutely convergent (Thm. 1), and so we can apply the lemma to compute the Cauchy product

$$\exp(z)\exp(w) \overset{(3)}{=} \sum_{n=0}^{\infty} c_n = \sum_{n=0}^{\infty} \frac{1}{n!}(z + w)^n = \exp(z + w).$$

□

In particular we conclude $\exp(z)\exp(-z) = \exp(z - z) = \exp(0) = 1$, that is,

$$\exp(-z) = \frac{1}{\exp(z)}. \tag{5}$$

This has the following consequence for exp evaluated on real numbers:

Corollary 5. (i) $\exp(x) > 0$ *for all* $x \in \mathbb{R}$*, and* $\exp(x) \geq 1$ *for* $x \geq 0$.
(ii) $\exp\colon \mathbb{R} \to \mathbb{R}$ *is strictly increasing.*
(iii) *For* $k \in \mathbb{Z}$ *holds* $\exp(k) = e^k$*, where* $e \in \mathbb{R}$ *is the Euler number.*

Here, by definition $z^0 := 1$ for $z \in \mathbb{C}$ and $z^{-k} := \frac{1}{z^k}$ for $z \in \mathbb{C} \setminus \{0\}$ and $k \in \mathbb{N}$.

`14. Vorlesung, Donnerstag, 28.11.06 (T 7)` ______________________________

Proof. (i) In case $x \geq 0$ we verify $\exp(x) = 1 + x + \frac{1}{2!}x^2 + \ldots \geq 1$. Using (5), this gives $\exp(-x) = \frac{1}{\exp(x)} > 0$.

(ii) For $x < y$ we use (i) to show

$$\exp(y) = \exp(y - x + x) \overset{(4)}{=} \underbrace{\exp(y - x)}_{>1} \underbrace{\exp(x)}_{>0} > \exp(x).$$

(iii) For $k = 0$, by definition $\exp(0) = e^0 = 1$ holds. Let now $k \in \mathbb{N}$. Using $\exp(1) = e$ and (4) gives

$$\exp(k) = \exp(1 + \ldots + 1) \overset{(4)}{=} \exp(1) \cdot \ldots \cdot \exp(1) = \exp(1)^k = e^k;$$

again, formally this is by induction. Finally,

$$\exp(-k) \overset{(5)}{=} \frac{1}{\exp(k)} = \frac{1}{e^k} = e^{-k}.$$

□

Property (iii) says that $\exp(x)$ interpolates the natural powers of e. Hence it makes sense to call exp the *exponential* function.

We want to take the inverse function of $\exp\colon \mathbb{R} \to \mathbb{R}$, the logarithm log. An inverse function is defined on the range of the original function. But why does the range of exp consist of all positive real numbers? We need to study the continuity of exp to assert that its range has no gaps.

2. Continuity

Legend has it that Galilei (1564-1642) studied free fall in an experiment by dropping two different balls from the leaning tower of Pisa. Aristotle had claimed that the time t for free fall depends on the mass of the falling body. To Galilei is attributed the claim that t is solely a function of the height h, that is, $t = f(h)$. For the height of the tower of Pisa, $h = 54m$, the time of free fall is $3.3sec$. This data may perhaps not be accurately known but we assume: *If the height is close to* $54m$ *then also falling time is close to* $f(54m) = 3.3sec$. This is the idea of continuity, which is fundamental for laws and experiments in the sciences: *Natura non saltat* – nature does not jump. Only on these grounds, the reproduction of experiments can become meaningful.

Historically, the significance of continuity was discovered only in the 19th century – long after the introduction of differentiation and integration. Before, continuity had been assumed tacitly. Find a real life example where similarly we assume that data at one time (or place) are meaningful for data at a nearby time (or place).

2.1. **Definition in terms of the limit.** If not specified otherwise, we will consider functions $f\colon D \to \mathbb{C}$ defined on arbitrary subsets $D \subset \mathbb{C}$. For $a \in D$ let us introduce the notation

$$\lim_{z\to a} f(z) = c \quad :\Longleftrightarrow \quad \lim_{n\to\infty} f(z_n) = c \text{ for each sequence } z_n \in D \text{ with } z_n \to a. \tag{6}$$

Definition (Bolzano 1817)**.** A function $f\colon D \to \mathbb{C}$ is *continuous [stetig] at* $a \in D$, if

$$\lim_{z\to a} f(z) = f(a) \tag{7}$$

The map f is *continuous*, if f is continuous in each $a \in D$.

Examples. 1. *Constant functions*, $f(z) = c$ for all $z \in D$; a shorthand notation is to write $f \equiv c$. Continuity is obvious: $\lim_{z\to a} f(z) = c = f(a)$

2. The *identity mapping* $z \mapsto z$, is continuous: $\lim_{z\to a} z = a$.

3. The floor function $f\colon \mathbb{R} \to \mathbb{R}$, $f(x) = \lfloor x \rfloor$, is not continuous at $a \in \mathbb{Z}$. Indeed, it is enough to find a particular sequence, such that (7) fails. We take $x_n := a - \frac{1}{n} \to a$:

$$\lim_{n\to\infty} \left\lfloor a - \frac{1}{n} \right\rfloor = \lim_{n\to\infty} (a-1) = a - 1 \quad \neq \quad \lfloor a \rfloor = a$$

On $\mathbb{R} \setminus \mathbb{Z}$ the function is continuous (show that!).

4. $f(x) := 1$ for $x \in \mathbb{Q}$ and $f(x) := 0$ for $x \in \mathbb{R} \setminus \mathbb{Q}$ is nowhere continuous, but $xf(x)$ is continuous at 0 (see problems).

5a) $\exp\colon \mathbb{C} \to \mathbb{C}$ is continuous at 0.

Proof: The remainder term estimate, Thm. 2, says for $n = 1$ that $R_1(z) = z + \frac{z^2}{2!} + \ldots = \exp(z) - 1$ is subject to

$$|R_1(z)| \le 2|z| \qquad \text{when} \qquad |z| \le 1.$$

For each null sequence $(z_n)_{n\in\mathbb{N}}$ there exists $N \in \mathbb{N}$, such that the assumption $|z_n| \le 1$ of the remainder term estimate holds for all $n \ge N$; therefore

$$|R_1(z_n)| = \big|\exp(z_n) - 1\big| \le 2|z_n| \to 0 \qquad \text{as } n \to \infty. \tag{8}$$

We conclude the continuity of exp in 0:

$$\lim_{z\to 0} \exp(z) = 1 = \exp(0),$$

b) $\exp\colon \mathbb{C} \to \mathbb{C}$ is continuous.

Proof: Suppose that $z_n \to a$ as $n \to \infty$. We use a) together with the functional equation (4):

$$\lim_{n\to\infty} \exp(z_n) \overset{z_n = a + z_n - a}{=} \lim_{n\to\infty} \big(\exp(a)\exp(z_n - a)\big) = \exp(a) \lim_{n\to\infty} \exp(\underbrace{z_n - a}_{\to 0}) \overset{\text{a)}}{=} \exp(a).$$

Remark. (*i*) Given a continuous function $f\colon D \to \mathbb{C}$ we call $\tilde{f}\colon E \to \mathbb{C}$ a *continuous extension*, if $E \supset D$, $\tilde{f}$ is continuous, and $\tilde{f}(z) = f(z)$ for $z \in D$. For the following kind of points the extension is unique, provided it exists.

(*ii*) An *accumulation point [Häufungspunkt]* of a set D is a point a such that a sequence $x_k \in D \setminus \{a\}$ with $x_k \to a$ exists.

Examples: 1. 0 is the only accumulation point of $D := \{\frac{1}{n} : n \in \mathbb{N}\}$.

2. $[a, b]$ is the set of accumulation points of (a, b).

3. $\mathbb{R}$ is the set of accumulation points of $\mathbb{Q}$.

The notation (6), as well as the notion of continuity, extends from D to the set of its accumulation points. For example, at the accumulation point 0 of $D = (0, 1)$, the function $f(x) := \sin\frac{1}{x}$ is not continuous, while $g(x) := x \sin\frac{1}{x}$ is.

2.2. **Operations which preserve continuity.** Various properties of continuous functions are direct consequences of the similar properties of limits:

Theorem 6. *Let $\lambda \in \mathbb{C}$. If $f, g\colon D \to \mathbb{C}$ are functions which are continuous at $a \in D$, then also $f + g$, fg, λf, $|f|$ are. The same holds for the quotient $\frac{f}{g}$ provided $g\colon D \to \mathbb{C} \setminus \{0\}$.*

`15. Vorlesung, Dienstag, 5.12.06 (Testklausur)` ______________________

Proof. The claimed properties follow from the respective laws for limits (Prop. II.3 and Thm. 4). As an example, let us show the first claim. Suppose that $z_n \to a$ as $n \to \infty$. Then

$$f(a) + g(a) = \lim_{n\to\infty} f(z_n) + \lim_{n\to\infty} g(z_n) \overset{\text{Prop. II.3}}{=} \lim_{n\to\infty} \big(f(z_n) + g(z_n)\big),$$

or $\big(f + g\big)(a) = \lim_{n\to\infty}(f + g)(z_n)$. Thus $f + g$ is continuous at a. $\square$

Examples. 1. $z \mapsto z^n$ as a map from $\mathbb{C}$ to $\mathbb{C}$ is continuous for $n \in \mathbb{N}$. Indeed, the identity $z \mapsto z$ is continuous, and by the theorem also $z \mapsto z \cdot z$, etc. (induction).

2 *Polynomials* $p(z) = a_0 + a_1 z + \ldots + a_n z^n$, with $a_k \in \mathbb{C}$, are continuous on $\mathbb{C}$: Indeed, z^k is continuous by 1. By the theorem, each product $a_k \cdot z^k$ is continuous. So is the sum $\sum_{k=0}^{n} a_k z^k$, again by the theorem.

Applying Prop. II.6 we obtain:

Theorem 7. *A function* $f = \operatorname{Re} f + i \operatorname{Im} f \colon D \to \mathbb{C}$ *is continuous if and only if* $\operatorname{Re} f$ *and* $\operatorname{Im} f$ *are continuous.*

Finally, we show that continuity is preserved under composition. This will allow us to derive the continuity of complicated functions from the known continuity of simple functions.

Theorem 8. *Let* $D, E \subset \mathbb{C}$*. If* $f \colon D \to E$ *is continuous at* a *and* $g \colon E \to \mathbb{C}$ *is continuous at* $f(a)$*, then* $g \circ f \colon D \to \mathbb{C}$ *is also continuous at* a*.*

Proof. Let $z_n \in D$ be an arbitrary sequence which converges to a. Then,

$$g\big(f(a)\big) = g\big(f(\lim_{n\to\infty} z_n)\big) \overset{f \text{ cts.}}{=} g\big(\lim_{n\to\infty} f(z_n)\big) \overset{g \text{ cts.}}{=} \lim_{n\to\infty} g\big(f(z_n)\big),$$

that is, $(g \circ f)(a) = \lim_{n\to\infty} (g \circ f)(z_n)$. ☐

Example. Given a point $p \in \mathbb{C}$, the function $z \mapsto |z - p|$ measures the distance to p. It is continuous on $\mathbb{C}$ as we can consider it a composition of $g(y) := |y|$ with $f(x) := x - p$.

2.3. **The ε-δ-test.** This test is quantitative, that is, in terms of an error bound. It has great theoretical significance, see, e.g., the discussion of uniform continuity below.

Theorem 9 (ε-δ-test, Heine 1872)**.** *Let* $D \subset \mathbb{C}$ *and* $f \colon D \to \mathbb{C}$*. Then* f *is continuous at* $a \in D$*, if and only if the following holds: For each* $\varepsilon > 0$ *there is* $\delta = \delta(a, \varepsilon) > 0$*, such that*

$$|f(z) - f(a)| < \varepsilon \qquad \text{for all } z \in D \text{ with } |z - a| < \delta. \tag{9}$$

Let us visualize (9) in the complex plane: Each ε-ball $B_\varepsilon\big(f(a)\big)$ about $f(a)$ contains the entire image of a sufficiently small δ-ball about a. So all points in $B_\delta(a) \cap D$ meet the error bound ε for $f(a)$. In the case of real functions, sketch the condition in the graph!

Before supplying the important proof, let us give examples on how to apply the ε-δ test to a given function. The strategy is as follows:

- To prove continuity, we must guess $\delta(a, \varepsilon)$, for instance by inspecting the graph. Then we estimate $|f(z) - f(a)| < \varepsilon$ by expressing this difference in terms of $|z - a|$; this can also lead to the guess of $\delta(a, \varepsilon)$ rightaway.
- If we want to prove f is discontinuous at a then all we need to do is find some $\varepsilon > 0$ (again, we may inspect the graph) for which no $\delta > 0$ will work. That is, for each $\delta > 0$ we need to exhibit one $z \in B_\delta(a)$ with $|f(z) - f(a)| > \varepsilon$.

Examples. 1. $f \colon \mathbb{C} \to \mathbb{R}$, $f(z) := 3|z|$. From the graph, we expect that for any $\varepsilon > 0$ the number $\delta := \frac{1}{3}\varepsilon$ will satisfy (9). Indeed, if $|z - a| < \delta$ then, as desired,

$$|f(z) - f(a)| = 3\big||z| - |a|\big| \overset{\text{inverse } \Delta\text{-inequ.}}{\leq} 3\big|z - a\big| < 3\frac{\varepsilon}{3} = \varepsilon.$$

2. Let the *sign function [Signumfunktion]* $\operatorname{sgn}\colon \mathbb{R} \to \mathbb{R}$ be defined by $\operatorname{sgn}(x) := 1$ for $x > 0$, $\operatorname{sgn}(x) := -1$ for $x < 0$ and $\operatorname{sgn}(0) := 0$. Then sgn does not satisfy (9) at $a = 0$. Indeed, for $\varepsilon := \frac{1}{2}$, whatever $\delta > 0$ is, we can choose $x := \frac{\delta}{2} \in (-\delta, \delta)$ to see $|f(x) - f(0)| = 1 \not< \varepsilon$.

3. Let $f\colon \mathbb{C} \to \mathbb{C}$, $f(z) := z^2$. Restricted to the reals, this function has a graph which gets steeper as $|a|$ becomes larger. Hence we expect that $\delta \to 0$ as $a \to \infty$. To compute the exact dependence, assume $|z - a| < \delta$. Then

$$|z^2 - a^2| = |z - a|\,|\underbrace{z + a}_{z-a+2a}| \overset{\Delta\text{-inequ.}}{\leq} |z - a|\big(|z - a| + |2a|\big) < \delta(\delta + |2a|)$$

We want to determine $\delta(\varepsilon, a)$ such that this becomes less than ε. Assuming $\delta \leq 1$ we find

$$\delta(\delta + |2a|) \leq \delta(1 + |2a|) \overset{!}{<} \varepsilon \qquad \Leftrightarrow \qquad \delta < \frac{\varepsilon}{1 + 2|a|}.$$

Alltogether we obtain that for $\delta := \min\big(1, \frac{\varepsilon}{1+2|a|}\big)$ we have $|z^2 - a^2| < \varepsilon$.

The last example is not at all straightforward. To understand why this is so, note that the ε-δ-test is a quantitative measure of the continuity of f at a, while the limit test is merely qualitative. Hence for concrete functions, the ε-δ test is usually harder to check.

Proof. "$\Leftarrow$": We assume the ε-δ-condition at $a \in D$. We need to show that for any given sequence $z_n \to a$ in the domain, the image sequence satisfies $f(z_n) \to f(a)$. Let $\varepsilon > 0$ be arbitrary, and pick $\delta = \delta(a, \varepsilon)$ from (9). Since $z_n \to a$ we can choose $N = N(\delta(a, \varepsilon)) \in \mathbb{N}$ such that $|z_n - a| < \delta$ for all $n \geq N$. But (9) then implies $\big|f(z_n) - f(a)\big| < \varepsilon$ for all $n \geq N$. As ε was arbitrary, this gives $f(z_n) \to f(a)$, as desired.

"$\Rightarrow$". Assume the limit test holds. We show that for each $\varepsilon > 0$ there is $\delta > 0$ with (9), applying an indirect argument.

Suppose for a particular error bound $\varepsilon > 0$ there were no $\delta > 0$ satisfying condition (9). In particular, (9) could not be satisfied with any $\delta = \frac{1}{n}$, where $n \in \mathbb{N}$. Thus there exist $z_n \in D$ with $|z_n - a| < \frac{1}{n}$, so that $|f(z_n) - f(a)| \geq \varepsilon$. Therefore we have $z_n \to a$ but not $f(z_n) \to f(a)$, contradicting the limit test. $\square$

A property as simple as important can easily be derived from the ε-δ-test:

Corollary 10. *Let $f\colon D \to \mathbb{C}$ be continuous at $a \in D$ with $f(a) \neq 0$. Then there is $\delta > 0$ such that $f(z) \neq 0$ for all $z \in B_\delta(a)$.*

Proof. For $\varepsilon := |f(a)|$ let us choose $\delta = \delta(\varepsilon, a)$ according to the ε-δ-condition. Then for all z with $|z - a| < \delta$ we verify:

$$|f(z)| = \big|f(a) + f(z) - f(a)\big| \overset{\text{inv. }\Delta\text{-inequ.}}{\geq} |f(a)| - \underbrace{|f(z) - f(a)|}_{<\varepsilon = |f(a)|} > 0$$

□

`16. Vorlesung, Donnerstag, 7.12.06 (T 8)` ———

3. Properties of continuous real functions

In the next two sections we consider real functions $f\colon I \to \mathbb{R}$. Here I will always denote an interval (see II.1.5).

3.1. **Intermediate value theorem.** A common explanation of continuity of real functions is that their graph can be drawn in one turn. This interpretation can run into problems: Continuity is also defined for domains such as $\mathbb{Q}$ or $\mathbb{Z}$. But even if the domain is an interval, the drawing can take infinite time! But the main content of the explanation is, perhaps, that the range does not leave any "gaps" if the domain is an interval—that is indeed true:

Theorem 11 (Intermediate value theorem, Bolzano 1817)**.** *Let f: $[a, b] \to \mathbb{R}$ be continuous.*
(i) If $f(a) < 0$ and $f(b) > 0$ (or vice versa), then there exists $x \in (a, b)$ with $f(x) = 0$, called a zero *[Nullstelle]* of *f.*
(ii) Similarly, each value c in between $f(a)$ and $f(b)$ is attained.

In general, x will not be unique. Note also that it the completeness of $\mathbb{R}$ which is essential for this theorem to hold: The function $f\colon \mathbb{Q} \to \mathbb{Q}$, $f(x) := x^2 - 2$ has no zero! (Where does the following proof fail for that case?)

Proof. (i) We locate x by the method of interval bisection (compare with the proof of the Bolzano-Weierstrass Thm. II.12).

We claim there are nested intervals $(I_n := [a_n, b_n])_{n\in\mathbb{N}}$ with the property $f(a_n) < 0$ and $f(b_n) \geq 0$. To do this, we set $[a_1, b_1] := [a, b]$ and, recursively for $n \geq 2$,

$$I_{n+1} := \begin{cases} [a_n, \frac{a_n+b_n}{2}] & \text{if } f(\frac{a_n+b_n}{2}) \geq 0, \\ [\frac{a_n+b_n}{2}, b_n] & \text{if } f(\frac{a_n+b_n}{2}) < 0. \end{cases}$$

Since (I_n) are nested intervals, there exists $x \in I_n$ for all n (Thm. II.9). From continuity and the fact that convergence respects weak inequalities (Prop. II.8) we conclude

$$f(x) = \lim_{n\to\infty} f(a_n) \leq 0 \quad \text{and} \quad f(x) = \lim_{n\to\infty} f(b_n) \geq 0,$$

and so $f(x) = 0$ must hold.

(ii) Similar by replacing 0 with c in part (i). □

Corollary 12. (i) $\exp\colon \mathbb{R} \to (0,\infty)$ *is bijective.*
(ii) *For $n \in \mathbb{N}$ the power $x^n\colon [0,\infty) \to [0,\infty)$ is bijective.*

Proof. (i) We need to show that for any $y > 0$ there is x such that $\exp(x) = y$. Note first that

$$e > 1 \qquad \Rightarrow \qquad \lim_{j\to\infty} e^j = \infty \quad \text{and} \quad \lim_{j\to\infty} e^{-j} = \lim_{j\to\infty} \left(\frac{1}{e}\right)^j = 0.$$

Hence there exist numbers $k, n \in \mathbb{Z}$ with $\exp(k) = e^k < y < e^n = \exp(n)$. The function $\exp\colon [k,n] \to \mathbb{R}$ is continuous, and so the intermediate value theorem gives that the value y is attained. Moreover, by the strict monotonicity of exp asserted in Corollary 5(ii), exp is also injective.

(ii) Similar arguments apply. The polynomial x^n is continuous (first example in 2.2). Moreover, $0^n = 0$ and $j^n \to \infty$ as $j \to \infty$. Again the intermediate value theorem proves the surjectivity of $x^n\colon [0,\infty) \to [0,\infty)$, while the strict monotonictity (see example in 1.1) proves injectivity. □

The following consequence of the intermediate value theorem will be used when we discuss changes of variable.

Theorem 13. *Let I, J be intervals, $f\colon I \to J$ bijective and continuous. Then f is strictly monotone.*

Proof. Consider three points $x < y < z$ contained in I. By assumption, these points have distinct images under f. If f is not montone, the value $f(y)$ is not between $f(x)$ and $f(z)$. For instance we could have the case

$$f(x) < f(z) < f(y).$$

But then the intermediate value theorem would give a $\xi \in (x,y)$ (hence $\xi \neq z$) with $f(\xi) = f(z)$, contradicting injectivity. Similar in the three remaining cases. □

3.2. **Minima and maxima.** When functions come up in real life, an important issue is to determine their maxima and minima. In real life the goal may be to minimize time or effort, to maximize money, or to optimize a geometry.

We can distinguish two problems:
1. Do extrema exist, that is, do functions attain their supremum or infimum?
2. How do we locate extrema? Differentiation will be the key tool, see Part IV.

Here, we will study the first point. Extrema do not always exist:

Examples. 1. $\frac{1}{1+x^2}$ has maximum 1 at 0, but does not take a minimum over $\mathbb{R}$.
2. $\frac{1}{x}$ does neither have a maximum nor a minimum on $(0,1)$.

On closed (bounded) intervals, however, we have:

Theorem 14 (Maximum value theorem). *A continuous function $f\colon [a,b] \to \mathbb{R}$ takes a* maximum *and* minimum. *That is, there are $x_{\min}, x_{\max} \in [a,b]$ with*

$$f(x_{\min}) \le f(x) \le f(x_{\max}) \qquad \text{for all } x \in [a,b].$$

In particular f is bounded, *that is, there is $C \in \mathbb{R}$ with $|f(x)| < C$ for all $x \in [a,b]$.*

Example. By definition, a constant function takes its maximum and minimum at all points of its domain, that is, any point of the domain can serve as $x_{\max}$ and $x_{\min}$.

Proof. (i) Let

$$B := \sup\big\{f(x)\colon x \in [a,b]\big\} \in \mathbb{R} \cup \{\infty\},$$

where we formally write $B = \infty$ in case the range of f is not bounded above. There is a sequence $x_n \in [a,b]$ with $f(x_n) \to B$. In general, $(x_n)_{n\in\mathbb{N}}$ need not converge (consider the example that f is constant). But (x_n) is bounded, and so by the theorem of Bolzano-Weierstrass it contains a subsequence $(x_{\nu_k})_{k\in\mathbb{N}}$ which converges to some point $x_{\max}$. This point is contained in $[a,b]$ by Prop. 8(i).

Now we use the continuity of f which gives

$$f(x_{\max}) = \lim_{k\to\infty} f(x_{\nu_k}) = B.$$

Since $f(x_{\max}) \in \mathbb{R}$, we can conclude $B \neq \infty$.

Similarly for $x_{\min}$. $\square$

4. Logarithm and general powers

We introduce logarithm and general roots as inverses of the real functions exp and x^n.

4.1. **The logarithm.** By Corollary 12(i), the function $\exp\colon \mathbb{R} \to (0,\infty)$ has an inverse function, $\log := \exp^{-1}\colon \mathbb{R}^+ \to \mathbb{R}$, that is, $y = \exp(x)$ if and only if $x = \log y$. It is called the *(natural) logarithm.* To see the logarithm is continuous, we prove a general statement, which once again applies the $\varepsilon - \delta$-test:

Theorem 15. *Let $f\colon (a,b) \to \mathbb{R}$ be strictly monotone and continuous. Then the inverse function $f^{-1}\colon f\big((a,b)\big) \to \mathbb{R}$ is strictly monotone and continuous as well.*

Example. If the domain is not an interval, the statement no longer holds: Find a function which is continuous and strictly monotone and defined on the union of two intervals $(a, b] \cup (c, d)$, such that the inverse function is *not* continuous!

Proof. Monotonocity is obvious (check!).

To prove f^{-1} is continuous, we use the ε-δ-test. Let $x \in (a, b)$ and $\varepsilon > 0$ be given. Note that it is sufficient to verify the ε-δ-test for small $\varepsilon > 0$. Here, we assume that ε is small enough so that $I_\varepsilon(x) := (x - \varepsilon, x + \varepsilon) \subset (a, b)$.

Since f is continuous the intermediate value theorem implies that $f\big((x - \varepsilon, x + \varepsilon)\big)$ is an interval. In the strictly increasing case, this must be the interval $\big(f(x - \varepsilon), f(x + \varepsilon)\big)$.

We want to find $\delta > 0$ such that f^{-1} maps $I_\delta(f(x))$ into $I_\varepsilon(x)$. Drawing a graph shows that $\delta := \min\big(|f(x + \varepsilon) - f(x)|, |f(x) - f(x - \varepsilon)|\big) > 0$ should work. Indeed, then $I_\delta\big(f(x)\big) \subset f\big(I_\varepsilon(x)\big)$ and so, as desired,

$$f^{-1}\big(I_\delta(f(x))\big) \subset f^{-1}\big(f(I_\varepsilon(x))\big) \overset{f \text{ bij.}}{=} I_\varepsilon(x).$$

□

`17. Vorlesung, Dienstag, 12.12.06 (Ü 7)` ______________________

Theorem 16. *The logarithm* $\log\colon (0, \infty) \to \mathbb{R}$ *is continuous, strictly increasing, bijective, and satisfies the functional equation*

$$\log(xy) = \log x + \log y \qquad \textit{for all } x, y \in (0, \infty). \tag{10}$$

Proof. It remains to prove the functional equation by applying the functional equation for exp:

$$\exp\big(\log(xy)\big) = xy = \exp(\log x)\exp(\log y) \overset{(4)}{=} \exp(\log x + \log y)$$

Injectivity of exp implies our claim. □

4.2. **Powers.** Like for exp we can also combine Corollary 12(ii) with Theorem 15 to obtain:

Theorem 17. *For each* $n \in \mathbb{N}$*, the power* $f\colon [0, \infty) \to [0, \infty)$*,* $f(x) = x^n$*, has a strictly monotone and continuous inverse function, denoted* $\sqrt[n]{.}\colon [0, \infty) \to \mathbb{R}$*.*

Our goal is to extend the definition of the power function to real and complex exponents. To extend to rational exponents is easy by composition:

Definition. For each $q = \frac{k}{n} \in \mathbb{Q}$ where $k \in \mathbb{Z}$, $n \in \mathbb{N}$ let

$$.^q\colon (0, \infty) \to \mathbb{R}, \qquad x^q := \sqrt[n]{x^k}.$$

Then the function $x \mapsto x^q$ is continuous according to the composition rule Thm. 7.

We change our point of view and now consider the power function a^q for fixed basis $a \in (0, \infty)$, but with variable exponent $q \in \mathbb{Q}$. A possible approach would be to assert that $q \mapsto a^q$ is continuous. Thus the a^x could be defined as the unique continuous function on $\mathbb{R}$ which agrees with a^x for $x \in \mathbb{Q}$.

Instead, we now pursue a less intuitive approach which is, however, technically simple, works for complex exponents as well, and introduces a formula which is worth having when differentiating and integrating. For each $a > 0$ we define a function

$$f_a \colon \mathbb{C} \to \mathbb{C}, \quad \text{with} \quad f_a(z) := \exp(z \log a).$$

Once we have established the rules for powers (see end of section), this is an obvious formula, as indeed $a^z = (e^{\log a})^z = e^{(\log a)z}$.

Our function f_a extends the rational powers to complex exponents:

Theorem 18. *The function $f_a \colon \mathbb{C} \to \mathbb{C}$ is continuous. Its restriction to $\mathbb{Q}$ satisfies*

$$f_a(q) = a^q \qquad \textit{for all } q \in \mathbb{Q}.$$

Proof. f_a is continuous as the composition $z \mapsto z \log a \mapsto \exp(z \log a)$.

f_a satisfies the same functional equation as exp:

$$f_a(z+w) = \exp\big((z+w)\log a\big) \overset{(4)}{=} \exp\big(z \log a\big) \exp\big(w \log a\big) = f_a(z) f_a(w). \tag{11}$$

Therefore, for $k \in \mathbb{N}$ we have $f_a(k) = f_a(1 + \ldots + 1) \overset{(11)}{=} \big(f_a(1)\big)^k = \big(\exp(1 \log a)\big)^k = a^k$. Using $f_a(0) = 1$ and (11) we also obtain $f_a(-z) = \frac{1}{f_a(z)}$. In conclusion, the result is

$$f_a(k) = a^k \qquad \text{for all } k \in \mathbb{Z}. \tag{12}$$

Writing $q = \frac{k}{n}$ with $k \in \mathbb{Z}$ and $n \in \mathbb{N}$ we conclude

$$\big(f_a(q)\big)^n = \left(f_a\Big(\frac{k}{n}\Big)\right)^n \overset{(11)}{=} f_a\Big(\frac{k}{n} + \ldots + \frac{k}{n}\Big) = f_a(k) \overset{(12)}{=} a^k.$$

As an application of Thm. 17 we can take the n-th root of this equation to establish the claim $f_a(q) = \sqrt[n]{a^k} = a^{\frac{k}{n}} = a^q$. □

The theorem asserts that the following definition is well-defined, that is, it is unambiguous for the previously defined case of rational powers:

Definition. The *general power* is defined by

$$a^z := \exp(z \log a) \qquad \text{for } z \in \mathbb{C},\ a > 0.$$

When we specialize to $a := e = \exp(1)$ we obtain

$$e^z = \exp(z \log e) = \exp(z) \qquad \text{for all } z \in \mathbb{C},$$

and thereby justify the familiar notation for the exponential function. We will use it from now on.

Let us finally list the arithmetic rules for powers, which hold for complex z just as they do for rational values $(a, b > 0)$:

$$a^z a^w \overset{(11)}{=} a^{z+w},$$

$$a^z b^z = e^{z \log a} e^{z \log b} \overset{(4)}{=} e^{z(\log a + \log b)} \overset{(10)}{=} e^{z \log(ab)} = (ab)^z,$$

$$\left(\frac{1}{a}\right)^z = e^{z \log(1/a)} = e^{-z \log a} = a^{-z}$$

Problem. Show $(a^x)^w = a^{xw}$. Why do we need to require $x \in \mathbb{R}$, $w \in \mathbb{C}$?

4.3. **Growth rates of** exp **and** log**.** We need the notion of a limit of a function at ∞. We write $\lim_{x\to\infty} f(x) = c \in \mathbb{R} \cup \{\infty\} \cup \{-\infty\}$, if (i) there is a sequence $x_n \in D \subset \mathbb{R}$ which diverges to ∞, and (ii) if $\lim_{n\to\infty} f(x_n) = c$ for all such sequences. Similarly we define $\lim_{x\to-\infty} f(x)$.

Examples. 1. $\lim_{x\to\infty} e^x = \infty$ and $\lim_{x\to-\infty} e^x = 0$ can be derived from the values of exp on $\mathbb{Z}$ and its monotonicity. (Compare with the proof of Corollary 12.)

2. exp increases stronger than any power (see problems):

$$\lim_{x\to\infty} \frac{e^x}{x^n} = \infty \qquad \text{for each } n \in \mathbb{N} \tag{13}$$

3. log increases weaker than any power:

$$\lim_{x\to\infty} \frac{\log x}{x^a} = 0 \qquad \text{for each } a > 0 \tag{14}$$

Proof:

$$\frac{\log x}{x^a} = \frac{\log x}{e^{a \log x}} = \frac{1}{\frac{e^{a \log x}}{\log x}} = \frac{\frac{1}{a}}{\frac{e^{a \log x}}{a \log x}}$$

Let us now take the limit $x \to \infty$. Then also $\log x \to \infty$, as log is increasing with range $\mathbb{R}$. Therefore, also $y := a \log x \to \infty$ and

$$\lim_{x\to\infty} \frac{\log x}{x^a} = \lim_{y\to\infty} \frac{\frac{1}{a}}{\frac{e^y}{y}} = \frac{1}{a} \lim_{y\to\infty} \frac{1}{\frac{e^y}{y}}.$$

By (13) the denominator diverges to ∞, and so the fraction converges to 0 (check!)

4.4. **Landau-symbols o and O.** We write

$$f(x) = o\big(h(x)\big) \quad \text{as } x \to \infty \qquad :\Longleftrightarrow \qquad \lim_{x\to\infty} \frac{f(x)}{h(x)} = 0.$$

Examples. 1. For each $a > 0$ we have $\log x = o(x^a)$ by (14).
2. The limit (13) is equivalent to $\lim_{x\to\infty} \frac{x^k}{e^x} = 0$ and therefore $x^k = o(e^x)$ for all $k \in \mathbb{N}$.

Moreover, we write

$$f(x) = O\big(h(x)\big) \quad \text{as } x \to \infty \qquad :\Longleftrightarrow \qquad \frac{f(x)}{h(x)} \text{ stays bounded as } x \to \infty,$$

i.e., there exist $C, x_0 \in \mathbb{R}$ with $\left|\frac{f(x)}{h(x)}\right| \le C$ for all $x > x_0$.

Examples. 1. $p(x) := a_0 + a_1 x + \ldots + a_n x^n = O(x^n)$ (check!).
2. If $f(x) = o\big(h(x)\big)$ then certainly $f(x) = O\big(h(x)\big)$.

Notation such as $O(x)$ and $o(x^5)$ does not denote functions: It is only short hand for the limit properties of functions.

For Computer Science, $O(n \log n)$ is important: This is the runtime of an algorithm which sorts a list with n entries (to alphabetical order, according to size, etc.). Note that $17n \log n + 1000n = O(n \log n)$, that is, for *fixed* n the Landau-notation does not imply any information on the actual size of a quantity. Only the asymptotic growth order as $n \to \infty$ is prescribed.

More generally, Landau symbols are in use for any limit $x \to a$:

Example. $e^x = 1 + O(x)$ as $x \to 0$ by (8).

`18. Vorlesung, Donnerstag, 14.12.06 (T 9)` ______________________

5. Trigonometric functions

There are different ways to introduce the functions sine and cosine:

1. *Geometry*: As the edge lengths $\sin x, \cos x, 1$ of a right-angled triangle with angle x. In this sense, the functions are triangle-measuring or *trigonometric*.

2. *Dynamics*: In terms of a differential equation: Consider a circular motion about the origin in the plane. The vector function $t \mapsto f(t)$ then satisfies a motion law like $f'' = -f$. This differential equation has $t \mapsto (\cos t, \sin t)$ as a vector-valued solution.

3. *Analysis*: We will introduce the trigonometric functions by their series representations. The rigorous verification that these series satisfy the properties 1. and 2. can be given once differentiation and integration have been introduced.

5.1. **Sine and Cosine.**

Definition. For $z \in \mathbb{C}$ we define the series *sine [Sinus]* and *cosine [Cosinus]* by

$$\cos z := 1 - \frac{z^2}{2!} + \frac{z^4}{4!} \mp \ldots = \sum_{n=0}^{\infty} (-1)^n \frac{z^{2n}}{(2n)!}, \tag{15}$$

$$\sin z := z - \frac{z^3}{3!} + \frac{z^5}{5!} \mp \ldots = \sum_{n=0}^{\infty} (-1)^n \frac{z^{2n+1}}{(2n+1)!}. \tag{16}$$

Assuming for a moment we are allowed to differentiate a series $\sum_{n=0}^{\infty} a_n z^n$ termwise. The differential equation $f'' = -f$ then gives a recursive formula for the a_n. This formula shows that the solutions are linear combinations of the series of sin and cos (check!).

Note that for $x \in \mathbb{R}$ also $\cos x, \sin x \in \mathbb{R}$.

Each of the series sin and cos contains every other term of the exponential series, up to sign. Thus they are closely linked to exp:

Theorem 19. (i) *The series* $\cos z$ *and* $\sin z$ *converge absolutely for all* $z \in \mathbb{C}$.
(ii) *Cosine is an* even function *and sine is* odd*, that is,*

$$\cos(-z) = \cos z, \qquad \sin(-z) = -\sin z \qquad \textit{for all } z \in \mathbb{C}.$$

(iii) $e^{iz} = \cos z + i \sin z$ (Euler formula).
(iv)
$$\cos z = \frac{e^{iz} + e^{-iz}}{2}, \qquad \sin z = \frac{e^{iz} - e^{-iz}}{2i}.$$

(v) *Cosine and sine are continuous functions on* $\mathbb{C}$.

Proof. (i) This can be verified using the ratio test, or using majorization with the series $\exp(|z|)$ (check!).

(ii) This holds for the partial sums, hence for their limits.

(iii) Since $i^2 = -1$, $i^3 = -i$, $i^4 = 1$, etc., we obtain

$$\exp(iz) = 1 + iz - \frac{z^2}{2!} - i\frac{z^3}{3!} + \frac{z^4}{4!} + i\frac{z^5}{5!} + \ldots.$$

Thus partial sums of $\exp(iz)$ agree with the partial sums of $\cos z + i \sin z$. By convergence of all three series we can take the limit, which is the Euler formula.

(iv) Using (ii) the Euler formula gives $e^{-iz} = \cos z - i \sin z$. Adding or subtracting the Euler formula from this gives the result.

(v) Since exp is continuous, the right hand sides of (iii) are also continuous. □

The Euler formula says geometrically that the complex vector $e^{it} = \cos t + i \sin t$ for $t \in \mathbb{R}$ gives the hypothenuse of a right-angled triangle with cathetes $\cos t$ and $\sin t$. Let us show that the hypothenuse has length 1, that is, e^{it} is an element of the unit circle $\{|z| = 1\} \subset \mathbb{C}$. To see this, note that $\exp(\overline{z}) = \overline{\exp(z)}$ implies

$$\left|e^{it}\right|^2 = e^{it}\overline{e^{it}} = e^{it}e^{-it} = 1 \qquad \text{for all } t \in \mathbb{R}.$$

Remark. Let us now justify that the angle of the triangle is t. We claim that $e^{it} = \cos t + i \sin t$ is the point with angle $t \in \mathbb{R}$ on the unit circle. Let us use complex differentiation (to be justified later). Then the velocity vector is $(e^{it})' = ie^{it}$ and so has modulus $|ie^{it}| = |i||e^{it}| = 1$. This means e^{it} travels with unit speed on the unit circle. An integration (also to be justified later) then gives that the arc-length on the unit circle from $e^0 = 1$ to e^{it} must be t. Thus the point e^{it} encloses the angle t with the real axis, measured in radians.

Proposition 20. *Sine and cosine satisfy the* addition theorems

$$\cos(z + w) = \cos z \cos w - \sin z \sin w, \qquad \sin(z + w) = \cos z \sin w + \sin z \cos w \tag{17}$$

for each $z, w \in \mathbb{C}$. In particular, we have for all $z \in \mathbb{C}$

$$1 = \cos^2 z + \sin^2 z, \qquad \cos(2z) = \cos^2 z - \sin^2 z, \qquad \sin(2z) = 2 \cos z \sin z. \tag{18}$$

Proof. Let us use odd- and evenness of sine and cosine:

$$\begin{aligned} e^{\pm i(z+w)} &= e^{\pm iz}e^{\pm iw} = (\cos z \pm i \sin w)(\cos z \pm i \sin w) \\ &= \cos z \cos w - \sin z \sin w \ \pm\ i\big(\cos z \sin w + \sin z \cos w\big) \end{aligned}$$

Representing $2\cos(z + w) = e^{i(z+w)} + e^{-i(z+w)}$ and $2i \sin(z + w) = e^{i(z+w)} - e^{-i(z+w)}$, we find the addition theorems (17).

Setting $z = -w$ in the cosine formula gives the Pythagorean rule in (18), setting $z = w$ gives the double angle formulas. ☐

Exercises: 1. Interpret $e^z = e^{x+iy} = e^x(\cos y + i \sin y)$ in terms of polar coordinates: What is the polar angle, what the modulus of e^z? Given this interpretation, explain why the functional equation is a consequence.
2. Derive formulas for $\cos(3x)$ and $\sin(3x)$.
3. Determine real and imaginary part of $\cos z = \cos(x + iy)$ and $\sin z = \sin(x + iy)$.

5.2. **The number π.** We take a route as quick as ungeometric and will define the circle number π as two times the first zero of $\cos\colon [0, \infty) \to \mathbb{R}$. At this point, we do not yet know that cos and sin have any zeros (apart from $\sin 0 = \operatorname{Im}(e^0) = 0$). The following will help us to locate the desired zeros.

Lemma 21. *For $x \in (0,2]$ we have the inclusions*

$$1 - \frac{x^2}{2} < \cos x < 1 - \frac{x^2}{2} + \frac{x^4}{24}, \tag{19}$$

$$0 < x - \frac{x^3}{6} < \sin x < x. \tag{20}$$

Proof. For real numbers x, the series $\sin x$ and $\cos x$ alternate. Recall the Leibniz test Thm. II.17 and its proof: If $a_n > 0$ is a decreasing null sequence then the partial sums form nested intervals with common limit $s := a_0 - a_1 + a_2 - a_3 \pm \dots$

$$a_0 - a_1 \pm \dots - a_{2n+1} < s < a_0 - a_1 \pm \dots + a_{2m} \qquad \text{for all } n, m \geq 0 \tag{21}$$

On the positive reals, the series sine (16) is the alternating sum of the terms $a_n := \frac{x^{2n+1}}{(2n+1)!}$. For $x \in (0,2]$ they are decreasing as

$$\frac{x^{2n+3}}{(2n+3)!} \cdot \frac{(2n+1)!}{x^{2n+1}} = \frac{x^2}{(2n+3)(2n+2)} < 1 \qquad \text{for all } n \geq 0.$$

We apply (21) in the form $a_0 - a_1 < s < a_0$ to give (20); the first inequality follows from $6x - x^3 = x(6 - x^2) > 0$ for $0 < x \leq 2$.

For cosine (15) we have $a_n := \frac{x^{2n}}{(2n)!}$. Provided that $0 < x \leq 2$ these terms decrease, starting at the second term:

$$\frac{x^{2n+2}}{(2n+2)!} \cdot \frac{(2n)!}{x^{2n}} = \frac{x^2}{(2n+1)(2n+2)} < 1 \qquad \text{for all } n \geq 1.$$

This time, we apply (21) in the form $a_0 - a_1 < s < a_0 - a_1 + a_2$ to obtain (19). $\square$

Theorem 22. *Cosine has exactly one zero in the interval* $[0,2]$.

Proof. By (19) we have $\cos 0 = 1 > 0$ and $\cos 2 < 1 - \frac{4}{2} + \frac{16}{24} = 1 - 2 + \frac{2}{3} = -\frac{1}{3}$. Now cos is a continuous function (Thm. 19(*iv*)), and so the Intermediate Value Theorem 11 establishes the existence of a zero.

It remains to show uniqueness. We show cosine is strictly decreasing on $[0,2]$. For $x, y \in \mathbb{R}$ let us write $u := \frac{y+x}{2}$, $v := \frac{y-x}{2}$. Then

$$\begin{aligned}
\cos x - \cos y \;&=\; \cos(u - v) - \cos(u + v) \\
&\overset{(17)}{=} \cos u \underbrace{\cos(-v)}_{\cos v} - \sin u \underbrace{\sin(-v)}_{-\sin v} - \cos u \cos v + \sin u \sin v \\
&=\; 2 \sin u \sin v = 2 \sin \frac{y+x}{2} \sin \frac{y-x}{2}.
\end{aligned} \tag{22}$$

Let us now specialize to $0 \leq x < y \leq 2$. Then $0 < \frac{y \pm x}{2} \leq 2$ and so (20) gives $\sin \frac{y \pm x}{2} > 0$. This proves $\cos x > \cos y$, as desired. $\square$

Definition (Baltzer 1875)**.** We denote two times the zero of cos on $[0,2]$ by π.

`19. Vorlesung, Dienstag, 19.12.06 (Ü 8)` ______________________

Remarks. 1. From a computational point of view, this definition is impractical. Efficient ways to compute π are discussed in Section 8.11 (S. 132/33) of [K], along with further references.

2. Using integration, we will later verify the classical definitions of π: The unit disk has area π and circumference 2π.

3. The number π is irrational (Lambert 1762). It is in fact *transcendental* (Lindemann 1882), that is, π is not the zero of any polynomial $p(z) = \sum_{k=0}^{n} a_k z^k$ with rational (or integer) coefficients a_k. For a proof, see Hardy/Wright: Einführung in die Zahlentheorie, S.195-200. As shown in classes on algebra, the significance of this result is that it proved a conjecture that had been open for about two-thousand years: *Given a disk, it is impossible to construct a square of equal area, using compasses [Zirkel] and ruler alone.* Much of medieval mathematics was concerned with the problem to measure the lengths of curved arcs or the areas of regions bounded by curved arcs; the first was called the *rectifiability* problem, the second the *quadrature* problem. Thus the conjecture is cited as the impossibility of a quadrature of the circle.

4. In the book *Numbers* [E, English p.129/130, or German S.103/104], Remmert gives an interesting account on how Landau was critizised by Nazi mathematicians for publicizing Baltzer's definition of π.

5.3. **Periodicity of sine, cosine,** exp**.** Let us now draw consequences of the identity $\cos\frac{\pi}{2} = 0$.

(i) We claim that $\sin\frac{\pi}{2} = 1$. Indeed, sin is positive on $(0,2]$ and $1 = \cos^2\frac{\pi}{2} + \sin^2\frac{\pi}{2} = \sin^2\frac{\pi}{2}$.

(ii) Therefore, using the functional equation:

$$e^{in\frac{\pi}{2}} = \left(e^{i\frac{\pi}{2}}\right)^n = \left(\cos\frac{\pi}{2} + i\sin\frac{\pi}{2}\right)^n \overset{(i)}{=} i^n, \qquad n \in \mathbb{Z},$$

and in particular,

$$e^{\pi i} = -1, \qquad e^{2\pi i} = 1. \tag{23}$$

The first formula in (23) relates four famous numbers and, arguably, can be considered the most beautiful formula of mathematics. We collect consequences of (23), which are not at all obvious from the power series definitions of sine and cosine:

Proposition 23. (i) *We have the following periodicities for all $z \in \mathbb{C}$:*

$$
\begin{aligned}
e^{z+2\pi i} &= e^z, \qquad \cos(z+2\pi) = \cos z, \qquad \sin(z+2\pi) = \sin z\\
e^{z+\pi i} &= -e^z, \qquad \cos(z+\pi) = -\cos z, \qquad \sin(z+\pi) = -\sin z\\
\cos\left(\frac{\pi}{2}-z\right) &= \sin z, \qquad \cos(\pi - z) = -\cos z, \qquad \sin(\pi - z) = \sin z
\end{aligned}
$$

(ii) *The zeros of the functions* $\sin, \cos\colon \mathbb{C} \to \mathbb{C}$ *are:*

$$\big\{z \in \mathbb{C}\colon\ \cos z = 0\big\} = \big\{k\pi + \frac{\pi}{2}\colon k \in \mathbb{Z}\big\}, \qquad \big\{z \in \mathbb{C}\colon\ \sin z = 0\big\} = \big\{k\pi\colon k \in \mathbb{Z}\big\} = \pi\mathbb{Z}$$

(iii) $e^z = 1$ *only holds for* $z \in 2\pi i\mathbb{Z} = \{\dots, -2\pi i, 0, 2\pi i, \dots\}$.

Proof. (i) Using addition theorems, the Euler formula, and (23) we obtain, for example:

$$
\begin{aligned}
\cos(z+2\pi) &\overset{(17)}{=} \cos z \underbrace{\cos 2\pi}_{=\operatorname{Re} e^{2\pi i}=1} - \sin z \underbrace{\sin 2\pi}_{=\operatorname{Im} e^{2\pi i}=0} = \cos z\\
\cos\left(\frac{\pi}{2}-z\right) &\overset{(17)}{=} \cos\frac{\pi}{2}\cos z + \sin\frac{\pi}{2}\sin z = \sin z,\\
\cos(\pi - z) = -\cos(-z) &\overset{\cos\text{ even}}{=} -\cos z
\end{aligned}
$$

(ii) By (i), it is enough to know cosine on the interval $[0, \frac{\pi}{2}]$ in order to deduce all values of sine and cosine on the reals. So the claim restricted to the reals follows from Thm. 22. But all zeros of sine and cosine are real: This follows from

$$
\begin{aligned}
\cos z = \cos(x+iy) &\overset{(17)}{=} \cos x \cos(iy) - \sin x \sin(iy)\\
&\overset{\text{Thm.19}(iii)}{=} \cos x \frac{1}{2}\underbrace{(e^{-y}+e^{y})}_{>0} + i\frac{1}{2}\sin x(e^{-y}-e^{y}).
\end{aligned}
$$

Using the hyperbolic functions, the result can also be written as $\cos x \cosh y - i \sin x \sinh y$. Similarly for sine.

(iii) To study the solutions of $e^z = 1$ let us write

$$1 = e^{x+iy} = e^x\big(\cos y + i \sin y\big).$$

The modulus of this equation is $1 = |e^x|$ which gives $x = 0$. Taking the imaginary part implies $y \in \pi\mathbb{Z}$. But odd multiples of π yield $\cos y = -1$ and so $e^{iy} = -1$, while even multiples satisfy the equation. $\square$

5.4. **Invertibility of sine, cosine,** exp. Being periodic functions, the real functions sine, cosine, and the complex exponential cannot be invertible. Nevertheless, after restricting these functions to suitable subdomains we can guarantee injectivity and hence invertibility.

We claim that

$$\cos\colon [0,\pi] \to [-1,1] \qquad \text{and} \qquad \sin\colon \Big[-\frac{\pi}{2},\frac{\pi}{2}\Big] \to [-1,1] \tag{24}$$

are bijective. Each function is continuous with boundary values ± 1. By the intermediate value theorem each function is surjective. Moreover, cosine is strictly decreasing on $[0,\frac{\pi}{2}]$ and, by virtue of $\cos(\pi - x) = -\cos x$, in fact strictly decreasing on all of $[0,\pi]$; similarly for sine. This shows injectivity.

The bijective restrictions (24) of sine and cosine have inverse functions, called *principal branches [Hauptzweige]*,

$$\arcsin\colon [-1,1] \to \Big[-\frac{\pi}{2},\frac{\pi}{2}\Big], \qquad \arccos\colon [-1,1] \to [0,\pi].$$

We can now make one of our claims about complex numbers rigorous:

Theorem 24 (Polar coordinates)**.** *Each complex number z has a representation*

$$z = re^{i\varphi}, \qquad \text{where } r = |z| \text{ and } \varphi \in \mathbb{R}.$$

If $z \neq 0$ then φ is unique up to addition of numbers in $2\pi\mathbb{Z}$. In particular, the map $\mathbb{R} \to \{z \in \mathbb{C}\colon |z| = 1\}$, $\varphi \mapsto e^{i\varphi}$, is surjective.

We call (r,φ) *polar coordinates* of z. They represent z in terms of its modulus and a non-unique *argument* φ.

Proof. Let $z \neq 0$ and write $\frac{z}{|z|} = \xi + i\eta$ where $\xi := \frac{1}{|z|}\operatorname{Re} z$, $\eta := \frac{1}{|z|}\operatorname{Im} z \in \mathbb{R}$. Then $\xi^2 + \eta^2 = 1$. To produce an argument in between $-\pi$ and π, let us set:

$$\varphi\colon \mathbb{C} \to (-\pi,\pi], \qquad \varphi(z) := \begin{cases} \arcsin\eta, & \text{for } \xi \geq 0, \\ \pi - \arcsin\eta, & \text{for } \xi < 0 \text{ and } \eta \geq 0, \\ -\pi - \arcsin\eta, & \text{for } \xi < 0 \text{ and } \eta < 0. \end{cases} \tag{25}$$

Consider, for instance, the first case. Then $\sin\varphi = \eta$ with $\varphi \in [-\frac{\pi}{2},\frac{\pi}{2}]$ and so $\cos\varphi \geq 0$. From $\cos^2\varphi + \sin^2\varphi = 1$ we conclude $\cos\varphi = \xi$. Thus, as desired,

$$\xi + i\eta = \cos\varphi + i\sin\varphi = e^{i\varphi}.$$

Uniqueness follows from Prop. 23(*iii*). The other cases are similar. □

We are now in a position to study the invertibility of exp as a complex function. Since $\exp(x+iy) = e^x(\cos y + i \sin y)$ we can describe the geometry of the exponential mapping as follows. It maps the open strip *[Streifen]* with $-\frac{\pi}{2} < \operatorname{Im} z < \frac{\pi}{2}$ onto the slit *[geschlitzt]* domain $\mathbb{C} \setminus (-\infty, 0]$. The vertical lines $\operatorname{Re} z = a$ map to circles of radius e^a. The horizontal lines $\operatorname{Im} z = b$ map to radial rays, making an angle b with the positive real axis. In particular the two horizontal boundary lines bounding the strip map onto the negative real axis: The line $\operatorname{Im} z = \frac{\pi}{2}$ corresponds to a limit with $\operatorname{Im} z > 0$, while $\operatorname{Im} z = -\frac{\pi}{2}$ corresponds $\operatorname{Im} z < 0$.

`20. Vorlesung, Donnerstag, 21.12.06 (T 10)` ________________________________

Theorem 25. (i) $\exp\colon \mathbb{C} \to \mathbb{C} \setminus \{0\}$ *is surjective and* $2\pi i$*-periodic, that is,*

$$\exp(z+w) = \exp(z) \qquad \Longleftrightarrow \qquad w \in 2\pi i \mathbb{Z}. \tag{26}$$

(ii) $\exp\colon \{z \in \mathbb{C}\colon -\pi < \operatorname{Im} z < \pi\} \to \mathbb{C} \setminus (-\infty, 0]$ *has a continuous inverse function, the* principal branch of the complex logarithm*,*

$$\log\colon \mathbb{C} \setminus (-\infty, 0] \to \{z \in \mathbb{C}\colon -\pi < \operatorname{Im} z < \pi\},$$

such that $\log z = \log|z| + i \arg z$ *where* $\arg z := \varphi(z)$ *from* (25).

Proof. (i) To see surjectivity, use the polar coordinates of Thm. 24 to represent an arbitrary $z \neq 0$ as $z = re^{i\varphi} = e^{\log r + i\varphi}$. Let us show (26): Since $|\exp z| = \exp(\operatorname{Re} z) \neq 0$, we can conclude from $\exp(z+w) = \exp z \exp w \overset{!}{=} \exp z$ that $\exp w = 1$; Prop. 23 gives the claim.

(ii) By part (i), for each $a \in \mathbb{R}$ the function exp is injective on the strip $\{z \in \mathbb{C}\colon a < \operatorname{Im} z \leq a + 2\pi\}$ and can be inverted there. Moreover, polar coordinates give $e^z = e^{\log|z| + i\varphi}$, with $-\pi < \varphi(z) < \pi$. This means $\varphi(z) = \arg z$, which is a continuous function on $\mathbb{C} \setminus (-\infty, 0]$. □

SUMMARY

We introduced the concept of continuity. We used the limit test as a definition and proved it to be equivalent to the ε-δ test. One of the main goals of our class is to understand the meaning and to be able to apply these tests.

For real functions defined on closed intervals, continuity has consequences such as the intermediate value theorem, and the existence of extrema. These theorems depend in fact on the completeness of the real numbers: they fail when the functions are restricted to rational numbers (consider the rational polynomials $x^2 - 2$, and $\frac{1}{3}x^3 - 2x$, respectively).

We also introduced the most important examples of some non-elementary functions: We defined the functions $\exp, \sin, \cos$ in terms of their power series; we also mentioned the differential equations they satisfy. We used continuity to invert $\exp$ on the reals and found the real logarithm. The logarithm allows us to define general powers. For $\sin, \cos$ as real functions, and $\exp\colon \mathbb{C} \to \mathbb{C}$ we introduced inverse functions by restricting the domain of the original functions.

Part 4. Differentiation and integration in one variable

Our exposition goes backwards in time: We have covered the number system (late 19th century) and continuity of functions (early 19th century). With the present chapter we discuss differentiation (late 17th century) before integration (early 17th century).

Here, we cover the theory for one variable. For several variables, differentiation will be covered in the second term, and integration in the fourth term.

1. DIFFERENTIATION

Historically, two different problems lead to the notion of differentiability.

1. Dynamics: The velocity problem.
Suppose that at time t a body is at the position $x(t) \in \mathbb{R}$. The average speed of the body, with respect to two times $s \neq t$, is $v_{s,t} := \frac{x(s)-x(t)}{s-t}$. Problem: Can we also attribute a speed to a single time t? In 1664–66, Newton solved this problem with the *difference quotient* which is the limit

$$v(t) := \lim_{s \to t} v_{s,t} = \lim_{s \to t} \frac{x(s) - x(t)}{s - t}.$$

2. Geometric optics: The problem of tangents.
When light is reflected at a planar mirror *[ebener Spiegel]*, the angle of incidence and the angle of reflection are congruent. But what happens when the mirror is curved? The idea developed by Leibniz is that the reflection of a curved surface at a point p coincides with a mirror in tangent position at p. Problem: Does a given curve have a tangent? To compute the tangents for the case of graphs, Leibniz developed his theory of differentiation, independently of Newton, in 1672–76. In what follows we will characterize the tangent line to the graph of a function as the graph of the *best linear approximation to the function.*

As our first task we will show that, surprisingly, the two mathematical descriptions, namely the difference quotient and the best linear approximation to a function, coincide.

We will develop the basic theory of differentiation in the complex setting, with emphasis on the real situation. Thus within the present section, D will denote an arbitrary subset of $\mathbb{C}$.

1.1. Limit of the difference quotient.

Definition. A function $f : D \to \mathbb{C}$ is *differentiable [differenzierbar] at* $a \in D$ if the limit

$$f'(a) := \lim_{z \to a,\, z \in D \setminus \{a\}} \frac{f(z) - f(a)}{z - a} \tag{1}$$

exists. We call $f'(a)$ the *derivative [Ableitung] of f in a*. The function f is *differentiable*, if $f'(a)$ exists for all $a \in D$.

While the notation $f'(a)$ is due to Newton, the notation $\frac{df}{dx}(a) := f'(a)$ or, occasionally, $\frac{df}{dx}\big|_a$ is due to Leibniz. We can rewrite (1) as

$$f'(z) = \lim_{h\to 0,\ h\neq 0,\ z+h\in D} \frac{f(z+h)-f(z)}{h}. \tag{2}$$

When writing the limits (1) or (2) we wish to omit the condition $z \neq a, z \in D$ or $h \neq 0, z+h \in D$ So, for the sake of simplifying notation, from now on we agree to write $\lim_{z\to a} g(z) = y$ for a function g only provided that:
• there is a sequence z_n with $z_n \to a$ in the domain of g, that is, a is an accumulation point of the domain, and
• for all such sequences $\lim_{n\to\infty} g(z_n) = y$.

Let us give a geometric interpretation of the derivative $f'(x)$. Consider the graph of f. The *secant (line) [Sekante]* passing through two distinct points $\big(x, f(x)\big)$ and $\big(x+h, f(x+h)\big)$ has a slope *[Steigung]* which is the *difference quotient [Differenzenquotient]* $\frac{f(x+h)-f(x)}{h}$. By definition, if f is differentiable at x, we can take the limit as $h \to 0$. The limit of the secants becomes a *tangent (line) [Tangente]* through $\big(x, f(x)\big)$ with slope $f'(x)$

Examples. (We consider the domain $D = \mathbb{C}$.)
1. Linear functions $f(z) = cz$ with $c \in \mathbb{C}$:

$$f'(a) = \lim_{z\to a} \frac{f(z)-f(a)}{z-a} = \lim_{z\to a} \frac{cz-ca}{z-a} = c$$

2. $f(z) = z^n$ with $n \in \mathbb{N}$:

$$f'(a) = \lim_{z\to a} \frac{z^n - a^n}{z-a} = \lim_{z\to a} \left(z^{n-1} + z^{n-2}a + \ldots + a^{n-1}\right) = na^{n-1} \tag{3}$$

3. (Counterexample) The function $z \mapsto |z|$ is not differentiable at 0: Indeed, the difference quotient does not have a limit as, for instance,

$$\lim_{h\to 0, h>0} \frac{|h|-|0|}{h} = 1 \quad \neq \quad \lim_{h\to 0, h<0} \frac{|h|-|0|}{h} = -1.$$

4. For the exponential function,

$$\begin{aligned} \exp'(z) &= \lim_{h\to 0} \frac{\exp(z+h)-\exp(z)}{h} \overset{\text{fct.equn.}}{=} \lim_{h\to 0} \exp(z)\frac{\exp(h)-1}{h} \\ &= \exp(z) \lim_{h\to 0} \frac{\exp(h)-1}{h} \overset{!}{=} \exp(z). \end{aligned}$$

It remains to verify the last step. The remainder term estimate Theorem III.2 gives

$$\exp h = 1 + h + R_2(h) \qquad \text{with} \quad |R_2(h)| \le |h|^2 \quad \text{for } |h| \le \frac{3}{2}.$$

Hence after dividing by $h \neq 0$ we obtain for these h

$$\left| \frac{\exp h - 1}{h} - 1 \right| = \left| \frac{R_2(h)}{h} \right| \le |h| \to 0 \qquad \text{as } h \to 0.$$

We conclude $\lim_{h\to 0} \frac{\exp(h)-1}{h}$ exists and equals 1.

Problem. Verify $\cos'(z) = -\sin z$ and $\sin'(z) = \cos z$ by by the following calculation:

$$\begin{aligned} \cos'(z) &= \lim_{h\to 0} \frac{\cos(z+h) - \cos z}{h} = \lim_{h\to 0} \frac{-2\sin\frac{h}{2}\sin(z+\frac{h}{2})}{h} \\ &= \underbrace{-\lim_{h\to 0} \frac{\sin\frac{h}{2}}{\frac{h}{2}}}_{=1} \lim_{h\to 0} \sin\left(z + \frac{h}{2}\right) = -\sin z \end{aligned}$$

Justify this calculation using III(20) and (22), among other facts. (See also (14) below.)

Note that termwise differentiation of the series representations of exp, sin and cos yield the same results; however, this method will only be justified in Subsection 5.5 below.

1.2. **Functions with linear approximation are differentiable.** Recall the geometric interpretation of the derivative. We are interested in how well the tangent line approximates f near x. If we approximate f by some affine linear function L we are left with some error $r_x(h) := f(x+h) - L(x+h)$. If we take for L specifically the affine linear function $L(x+h) := f(x) + f'(x)h$ which desribes the tangent of the graph, we expect that the error is better than linear near x. This statement turns out to be equivalent to differentiability:

Theorem 1. *$f\colon D \to \mathbb{C}$ is differentiable at $z \in D$ if and only if there exists $c \in \mathbb{C}$ such that*

$$f(z+h) = f(z) + ch + r_z(h) \qquad \textit{with} \quad \lim_{h\to 0} \frac{r_z(h)}{h} = 0. \tag{4}$$

Moreover, then $c = f'(z)$.

`21. Vorlesung, Dienstag, 9.1.07 (Ü 10)` ______

We call r_z the *remainder term [Restglied]*; r_z is defined on the set $\{h \in \mathbb{C} : z + h \in D\}$.

Proof. We divide (4) by $h \neq 0$ to obtain

$$\frac{r_z(h)}{h} = \frac{f(z+h) - f(z)}{h} - c. \tag{5}$$

"⇒" Since f is differentiable at z, for $c := f'(z)$ the right hand side of (5) has limit 0 as $h \to 0$, as desired.

"⇐" Conversely, if (4) holds, the left hand side of (5) has limit 0 for $h \to 0$, and so $\lim\limits_{h\to 0} \frac{f(z+h)-f(z)}{h}$ exists and equals c. □

One particular reason for the significance of (4) is that it will generalize to several variables, unlike (1).

Using Landau symbols, (4) reads $r_x(h) = o(|h|)$. Thus we can also write (4) as

$$f(z+h) = f(z) + f'(z)h + o\big(|h|\big) \qquad \text{as } h \to 0.$$

However, choosing $c \neq f'(z)$ in (4) will lead to an error which is only $O(|h|)$, that is, $\frac{r_z(h)}{h} \to (c - f'(z)) \neq 0$.

Since $r_z(h)$ decays faster than linearly near z we have in particular

$$0 = 0 \cdot 0 = \lim_{h\to 0} h \cdot \lim_{h\to 0} \frac{r_z(h)}{h} = \lim_{h\to 0} r_z(h).$$

This gives:

Corollary 2. *If $f\colon D \to \mathbb{C}$ is differentiable at $z \in D$, then f is continuous at z.*

Proof. Since $\lim_{h\to 0} r_z(h) = 0$ we conclude from (4) that

$$\lim_{h\to 0} f(z+h) = \lim_{h\to 0} \Big(f(z) + f'(z)\underbrace{h}_{\to 0} + \underbrace{r_h(z)}_{\to 0}\Big) = f(z).$$

□

On the other hand, in the tutorial we will construct a continuous functions which is not differentiable at any point of its domain!

1.3. **Rules for differentiation.** In this subsection, we discuss rules which allow us to compute the derivative of composed functions.

Proposition 3. *Let $f, g\colon D \to \mathbb{C}$ be differentiable at $z \in \mathbb{C}$ and $\lambda \in \mathbb{C}$. Then λf, $f+g$ and fg are differentiable at $z \in \mathbb{C}$ as well, and*

$$(\lambda f)'(z) = \lambda f'(z), \qquad (f+g)'(z) = f'(z) + g'(z) \qquad \textit{(linearity)}, \tag{6}$$

$$(fg)'(z) = f'(z)g(z) + f(z)g'(z) \qquad \textit{(product law)}. \tag{7}$$

If, moreover, $g(z) \neq 0$, then also

$$\left(\frac{f}{g}\right)'(z) = \frac{f'(z)g(z) - f(z)g'(z)}{g^2(z)} \qquad \textit{(quotient law)}. \tag{8}$$

Proof. Linearity (6) follows from the linearity of the limit of sequences (Prop. II.3).

For the product law (7) we write

$$\frac{f(z+h)g(z+h)-f(z)g(z)}{h}=\frac{f(z+h)-f(z)}{h}g(z+h)+f(z)\frac{g(z+h)-g(z)}{h}.$$

The product law follows from taking the limit $h\to 0$; note that $g(z+h)\to g(z)$ by Corollary 2.

Let us now prove the quotient law (8): Corollary 2 gives that g is continuous at z, and so we can apply Corollary III.10 to yield $\delta>0$ with $g(z+h)\neq 0$ whenever $|h|<\delta$. For such h let us write

$$\begin{aligned}\frac{\frac{f(z+h)}{g(z+h)}-\frac{f(z)}{g(z)}}{h}&=\frac{1}{g(z+h)g(z)}\,\frac{f(z+h)g(z)-f(z)g(z+h)}{h}\\&=\frac{1}{g(z+h)g(z)}\left(\frac{f(z+h)-f(z)}{h}g(z)-f(z)\frac{g(z+h)-g(z)}{h}\right).\end{aligned}$$

The limit as $h\to 0$ is precisely (8). □

Examples. 1. Linearity and $(z^k)'=kz^{k-1}$ (see Ex. 2 of Sect. 1.1) yields that polynomials are differentiable.

2. Let $n\in\mathbb{N}$. Applying the quotient law to $\frac{f}{g}=\frac{1}{z^n}$ we find, using (3),

$$\left(\frac{1}{z^k}\right)'=\frac{-nz^{n-1}}{z^{2n}}=-nz^{-n-1}\qquad\text{for } z\neq 0.$$

Altogether we obtain

$$\frac{d}{dz}z^n=nz^{n-1}\qquad\text{for all } n\in\mathbb{Z}\text{ (assuming } z\neq 0\text{ in case } n\leq 0). \tag{9}$$

3. Using the derivatives of sine and cosine, see (14) below, we find the derivative of $\tan\colon\mathbb{C}\setminus\{k\pi+\frac{\pi}{2}\mid k\in\mathbb{Z}\}\to\mathbb{C}$, $\tan z:=\frac{\sin z}{\cos z}$:

$$(\tan z)'=\left(\frac{\sin z}{\cos z}\right)'=\frac{(\sin z)'\cos z-\sin z(\cos z)'}{\cos^2 z}=\frac{\cos^2 z+\sin^2 z}{\cos^2 z}=\frac{1}{\cos^2 z} \tag{10}$$

Most important is the rule for the composition of two maps:

Theorem 4 (Chain Rule *[Kettenregel]*)**.** *Let $D,E\subset\mathbb{C}$ and $f\colon D\to E$, $g\colon E\to\mathbb{C}$. Suppose f is differentiable at $a\in D$, and g at $b:=f(a)\in E$. Then $g\circ f\colon D\to\mathbb{C}$ is differentiable at a with*

$$\big(g\circ f\big)'(a)=g'\big(f(a)\big)f'(a).$$

In case f is injective the chain rule has a simple proof: The difference quotient for $f\circ g$, extended with $f(z)-f(a)\neq 0$, gives

$$\frac{g\big(f(z)\big)-g\big(f(a)\big)}{z-a}=\frac{g\big(f(z)\big)-g\big(f(a)\big)}{f(z)-f(a)}\,\frac{f(z)-f(a)}{z-a},\qquad z\neq a. \tag{11}$$

Taking the limit $z \to a$, whereby using continuity of f at a, establishes the chain rule. In general, however, $f(z) = f(a)$ may hold for $z \neq a$. (Let us remark that physicists apply the same reasoning directly to the limit case, by writing it with "differentials": $\frac{dg}{dz} = \frac{dg}{df} \cdot \frac{df}{dz}$.)

We will modify (11) by using a third test for differentiability. A function is differentiable at a point precisely if the difference quotient extends as a continuous function to that point:

Lemma 5. *A function $g\colon E \to \mathbb{C}$ is differentiable at $b \in E$ if and only if there is a function*

$$d\colon E \to \mathbb{C}, \qquad g(y) - g(b) = d(y)(y - b); \tag{12}$$

which is continuous at b; if so then $d(b) := g'(b)$.

Proof. If g is differentiable, set $d(y) := \frac{g(y)-g(b)}{y-b}$ for $y \neq b$; since the limit exists, d extends continuously with $d(b) := g'(b)$ and (12) holds. Conversely, if d as in (12) is continuous then the difference quotient has a limit, which gives differentiability of g at b. □

Proof of the Chain Rule. For $z \neq a$, let us multiply (12) with $\frac{1}{z-a}$ and set $y := f(z)$, $b := f(a)$. This gives

$$\frac{g\big(f(z)\big) - g\big(f(a)\big)}{z - a} = d\big(f(z)\big)\frac{f(z) - f(a)}{z - a} \qquad \text{for all } z \in D \setminus \{a\}. \tag{13}$$

Note that unlike (11) this is still valid when $f(z) = f(a)$.

Consider now the limit of (13) for $z \to a$. Certainly the difference quotient of f tends to $f'(a)$. Moreover, differentiability implies continuity, by Corollary 2. Thus $y = f(z) \to f(a) = b$, and by continuity of d, also $\lim_{z\to a} d\big(f(z)\big) = d(b) = g'(b)$. We conclude that the right hand side of (13) has the limit $g'(b)f'(a)$. This establishes the existence of the limit $(g \circ f)'(a)$, and the thus the chain rule. □

Examples. 1. $f(z) = a^z$ for $a > 0$ has the derivative

$$f'(z) = \big(\exp(z \log a)\big)' = \exp(z \log c) \log a = a^z \log a.$$

2. We have

$$(\cos z)' = -\sin z \qquad \text{and} \qquad (\sin z)' = \cos z. \tag{14}$$

since, for instance,

$$(\cos z)' = \frac{1}{2}\big(e^{iz} + e^{-iz}\big)' \overset{\text{Chain R.}}{=} \frac{1}{2}\big(ie^{iz} - ie^{-iz}\big) = \frac{1}{2i}\big(-e^{iz} + e^{-iz}\big) = -\sin z.$$

`22. Vorlesung, Donnerstag, 11.1.07 (T 11)`

If a function f has an inverse function g, then the graph of g is the graph of f, reflected at the middiagonal. Now, if a tangent to f has slope c, its mirror image is a tangent to g

with slope $\frac{1}{c}$. So the derivative of the inverse function is the reciprocal of the derivative of the original function, at the appropriate point. This geometric fact is easy to check by the chain rule:

$$g\big(f(z)\big) = z \quad \overset{\text{Chain Rule}}{\Rightarrow} \quad g'\big(f(z)\big)f'(z) = 1 \quad \Rightarrow \quad g'\big(f(z)\big) = \frac{1}{f'(z)}$$

This calculation, however, assumes the fact that the inverse function g is differentiable at $f(z)$. The following theorem establishes this as a fact:

Theorem 6 (Differentiation of the inverse function)**.** *Suppose $f\colon D \to \mathbb{C}$ is differentiable at $a \in D$ with $f'(a) \neq 0$. If f has a continuous inverse function $g\colon f(D) \to \mathbb{C}$ then g is differentiable at $b := f(a)$ with*

$$g'(b) = \frac{1}{f'(a)} = \frac{1}{f'\big(g(b)\big)}. \tag{15}$$

In particular, a real strictly monotone function f, defined on an interval will have a continuous inverse g, by Thm. III.15.

Proof. Let $y_n \in f(D) \setminus \{b\}$ be an arbitrary sequence with $y_n \to b$. Since g is continuous at b, we have $z_n := g(y_n) \to g(b) = a$. Also, g is injective and so $y_n \neq b$ implies $z_n \neq a$. Moreover, by differentiability of f at a, there is a sequence $z_n \to a$, and so some sequence $y_n = f(z_n)$ as above does indeed exist. Using all these facts, we can write

$$g'(b) = \lim_{n\to\infty} \frac{g(y_n) - g(b)}{y_n - b} = \lim_{n\to\infty} \frac{z_n - a}{f(z_n) - f(a)} = \frac{1}{\lim\limits_{n\to\infty} \frac{f(z_n)-f(a)}{z_n - a}} = \frac{1}{f'(a)}.$$

In particular, the limit $g'(b)$ exists. $\square$

We have verified what physicists like to phrase as: If $y(x)$ has derivative $\frac{dy}{dx}$ then $x(y)$ has derivative $\frac{dx}{dy} = 1/\frac{dy}{dx}$.

Examples. 1. The real function $f(x) = \exp x$ has the inverse $g(y) = \log y$ for $y > 0$ with

$$(\log y)' = \frac{1}{\exp'(\log y)} = \frac{1}{\exp(\log y)} = \frac{1}{y}.$$

Consequently, for $a \in \mathbb{C}$ and $x > 0$ we establish the following generalization of (9):

$$\frac{d}{dx}x^a = \big(\exp(a\log x)\big)' \overset{\text{Chain Rule}}{=} \exp(a\log x)a\underbrace{(\log x)'}_{=1/x} = x^a\, a\frac{1}{x} = ax^{a-1} \tag{16}$$

2. As shown in some problem, the function $f\colon (-\frac{\pi}{2}, \frac{\pi}{2}) \to \mathbb{R}$, $f(x) := \tan x$ has an inverse function $g(y) = \arctan y$. We claim that $f = \tan$ satisfies the differential equation $f' = 1 + f^2$. Indeed, as in (10),

$$(\tan x)' = \frac{\cos^2 x + \sin^2 x}{\cos^2 x} = 1 + \tan^2 x.$$

Therefore, we can compute the derivative of arctan to be

$$(\arctan y)' = \frac{1}{\tan'(\arctan y)} = \frac{1}{1 + \tan^2(\arctan y)} = \frac{1}{1 + y^2}.$$

It is surprising that the transcendent functions log and arctan have derivatives which are elementary!

3. (Counterexample) The differentiable function $.^3\colon \mathbb{R} \to \mathbb{R}$ is invertible. However, its inverse $\sqrt[3]{.}$ is not differentiable at 0: the horizontal tangent to x^3 at $a = 0$ implies a vertical tangent of $\sqrt[3]{y}$ at $b = 0$. Thus in Theorem 6, the assumption $f'(a) \neq 0$ is essential!

2. Extrema of real functions

We now specialize to the case of real functions. A common task is to locate their maxima and minima. The main method to do this is to exhibit zeros of the derivative: the derivative can be computed by the rules of the previous section.

2.1. **Local extrema: necessary conditions.**

Zeros of the derivative will detect extrema of the following kind:

Definition. A function $f\colon D \to \mathbb{R}$ takes a *local* $\left\{ \substack{\textit{maximum} \\ \textit{minimum}} \right\}$ *[lokales* $\left\{ \substack{\textit{Maximum} \\ \textit{Minimum}} \right\}$*]* at $a \in D$ if there is $\varepsilon > 0$ such that

$$(a - \varepsilon, a + \varepsilon) \subset D \qquad \text{and} \qquad \left\{ \substack{f(a) \geq f(x) \\ f(a) \leq f(x)} \right\} \text{ for all } x \in \big(a - \varepsilon, a + \varepsilon\big).$$

An *extremum [Extremum]* is a minimum or maximum. If the inequality is strict for $x \neq a$ the local extremum is called *strict [strikt]*.

Note that a local extremum is not necessarily a (global) extremum. Note, moreover, that extrema at the boundary are not called local.

Examples. 1. While the function $f\colon \mathbb{R} \to \mathbb{R}$, $f(x) = x^4 - x^2 = x^2(x^2 - 1)$ does not take a maximum (since $f(x) \to \infty$ as $|x| \to \infty$), it takes a local maximum at 0.

2. The function $f\colon [0, 1] \to \mathbb{R}$, $f(x) := x$ does not take a local maximum, but it has a maximum. Likewise for minimum.

Over closed and open intervals, the behaviour of functions is very different:
• On bounded closed intervals, continuous functions take a maximum, see the maximum value theorem Thm. III.14.
• Over open intervals, functions may not have any extrema at all; the identity on $(-1,1)$ is an example. However, if they have, then these extrema are also local extrema.

Theorem 7 (Necessary condition for extrema)**.** *Suppose the function $f\colon D \to \mathbb{R}$ is differentiable at x. If f takes a local extremum at x, then $f'(x) = 0$.*

Proof. Without loss of generality we consider the case of a maximum at x. Working on a sufficiently small ε-neighbourhood of x, the one-sided limits must exist and satisfy

$$f'(x) = \lim_{h\to 0,\, h<0} \frac{f(x+h)-f(x)}{h} \geq 0 \qquad \text{(numerator and denominator } \leq 0\text{)},$$
$$f'(x) = \lim_{h\to 0,\, h>0} \frac{f(x+h)-f(x)}{h} \leq 0 \qquad \text{(numerator } \leq 0\text{, denominator } > 0\text{)}.$$

Consequently, $f'(x) = 0$. □

Conversely, however, if $f'(x) = 0$, then x need not be a local extremum: Indeed, $f(x) = x^3$ does not have a local extremum at 0. Hence $f'(x) = 0$ is not sufficient for an extremum. We will come back to sufficient conditions after digressing to an important general result.

2.2. **Mean Value Theorem of differentiation.** If a function takes the same value at two different points, then part (i) of the following theorem asserts that there exists a local extremum in between.

Theorem 8. *Let $f\colon [a,b] \to \mathbb{R}$ be continuous on $[a,b]$ and differentiable on (a,b) .*
(i) [Rolle 1690 (for polynomials)] *If $f(a) = f(b)$, then there is $\xi \in (a,b)$ with $f'(\xi) = 0$.*
(ii) [Mean Value Theorem (MVT) *[Mittelwertsatz]*, Lagrange 1797] *There is $\xi \in (a,b)$ with*

$$f'(\xi) = \frac{f(b)-f(a)}{b-a}.$$

Proof. (i) For f constant we can pick any $\xi \in (a,b)$ and are done. In general, the continuous function f takes a maximum and minimum on the closed interval $[a,b]$, by Thm. III.14. Assuming that f is not constant, one of these must be different from $f(a) = f(b)$. Hence there is a local extremum $f(\xi) \neq f(a) = f(b)$, with $\xi \in (a,b)$. From Theorem 7 we conclude the claim.

(*ii*) The function $F(x) := f(x) - \frac{f(b)-f(a)}{b-a}(x-a)$ has boundary values $F(a) = f(a) = F(b)$. Applying (*i*) to F then yields $\xi \in (a,b)$ with

$$0 = F'(\xi) = f'(\xi) - \frac{f(b)-f(a)}{b-a}.$$

□

`23. Vorlesung, Dienstag, 16.1.07 (Ü 11)` ________________________________

The Mean Value Theorem has great significance for the developement of the theory. Indeed, it provides a direct way to deduce information about a function f from a property of f'. The ambiguity arising from the point of evaluation ξ being at an unknown location within (a,b) will present no drawback. As an example, let us impose a bound *[Schranke]* on f' and infer a bound on f:

Theorem 9. *Let $f\colon [a,b] \to \mathbb{R}$ be continuous, and differentiable on (a,b). Moreover, suppose there are numbers $m, M \in \mathbb{R}$ satisfying*

$$m \le f'(\xi) \le M \qquad \text{for all } \xi \in (a,b).$$

Then we have the inclusion

$$m(y-x) \le f(y) - f(x) \le M(y-x) \qquad \text{for all } x, y \in [a,b] \text{ with } x \le y. \tag{17}$$

Similarly, strict inequality implies strict inequality.

Proof. The upper bound follows from $\frac{f(y)-f(x)}{y-x} \overset{\text{MVT}}{=} f'(\xi) \le M$; similar for the lower bound.

□

Example. For a symmetric bound $|f'(\xi)| \le L$ for all $\xi \in (a,b)$ the resulting bound (17) is called a *Lipschitz-bound*

$$|f(x) - f(y)| \le L|x-y| \qquad \text{for all } x, y \in [a,b]. \tag{18}$$

Functions which satisfy (18) are continuous (see tutorial); they are called *Lipschitz-continuous.*

When $L = 0$ the bound (18) gives for real-valued (and hence also for complex-valued) functions:

Corollary 10. *Let $f\colon [a,b] \to \mathbb{C}$ be continuous, and differentiable on (a,b). If $f'(x) = 0$ for all $x \in (a,b)$, then f is constant.*

An application of the corollary is to prove uniqueness of solutions to differential equations. For example, we prove the only functions which satisfy $f' = f$ are multiples of exp:

Proposition 11. *Let $a \in \mathbb{R}$ and suppose $f\colon \mathbb{R} \to \mathbb{R}$ satisfies the differential equation*

$$f'(x) = af(x) \qquad \text{for all } x \in \mathbb{R}.$$

Then

$$f(x) = ce^{ax},$$

where $c := f(0)$.

Proof. We claim the auxiliary function $F(x) := f(x)e^{-ax}$ is constant. Indeed,

$$F'(x) = f'(x)e^{-ax} - af(x)e^{-ax} = \big(\underbrace{f'(x) - af(x)}_{=0}\big)e^{-ax} = 0,$$

and so by the corollary, F must be constant. Thus $F(x) = F(0) = f(0)e^0 = c$ which gives $f(x) = F(x)e^{ax} = ce^{ax}$. □

Outlook. Many natural laws take the form of an *ordinary differential equation [gewöhnliche Differentialgleichung]*: Given $\Phi\colon \mathbb{R}^2 \to \mathbb{R}$ we want to find differentiable functions $f\colon \mathbb{R} \to \mathbb{R}$ solving the differential equation $f'(x) = \Phi\big(x, f(x)\big)$. In the above case of $\Phi(x,y) = ay$, the solution f is explicit and unique for given $f(0)$, by the proposition. In general, the solutions will not be explicit, but still there are theorems on existence and uniqueness. There is a class in the third term on this topic.

Problem. Let $s, c\colon \mathbb{R} \to \mathbb{R}$ be two functions with $s' = c$ and $c' = -s$, which have the particular values $s(0) = 0$ and $c(0) = 1$. Show that $s = \sin$ and $c = \cos$.

2.3. **Local extrema: sufficient conditions.** Theorem 9 states that a function can be bounded in terms of its derivative. Here, we consider the special case $m = 0$ or $M = 0$. Then the theorem predicts the function is monotone; We will apply this to extrema below.

Theorem 12. *Let $f\colon [a,b] \to \mathbb{R}$ be continuous, and differentiable on (a,b).*
(i) Then $\left\{ {f'\geq 0 \atop f'\leq 0} \right\}$ on (a,b) if and only if f is monotonically $\left\{ {\text{increasing} \atop \text{decreasing}} \right\}$ on $[a,b]$.
(ii) Moreover, if $\left\{ {f'>0 \atop f'<0} \right\}$ on (a,b) then f is strictly $\left\{ {\text{increasing} \atop \text{decreasing}} \right\}$ on $[a,b]$.

The converse of (ii) is false: $f(x) = x^3$ is strictly increasing but $f'(x) > 0$ is violated at 0.

Proof. The two implications "⇒" are immediate from Theorem 9 with $\left\{ {m=0 \atop M=0} \right\}$.

"⇐" in (i): Suppose that $x \in (a,b)$ is arbitrary. Then, for example,

$$f \text{ monotonically increasing} \quad \Leftrightarrow \quad \frac{f(y) - f(x)}{y - x} \geq 0 \quad \forall y \in (x,b) \quad \overset{y \to x}{\Rightarrow} \quad f'(x) \geq 0.$$

□

Note that the proof does not give the converse of (*ii*): While the difference quotient may be strictly larger than 0, its limit does merely satisfy $f'(x) \geq 0$.

A simple but useful sufficient condition for extrema is:

Lemma 13. *Let $f\colon (a,b) \to \mathbb{R}$ be differentiable and $x_0 \in (a,b)$. If*

$$\left\{ \begin{smallmatrix} f'(x)\geq 0 \\ f'(x)\leq 0 \end{smallmatrix} \right\} \text{ for } x < x_0, \quad \text{and} \quad \left\{ \begin{smallmatrix} f'(x)\leq 0 \\ f'(x)\geq 0 \end{smallmatrix} \right\} \text{ for } x > x_0, \tag{19}$$

then f takes a (global) $\left\{ \begin{smallmatrix} maximum \\ minimum \end{smallmatrix} \right\}$ at x_0. Furthermore, if both inequalities in (19) *are strict, then so is the extremum.*

Proof. Let us consider the top case. By Theorem 12 the function f is monotonically increasing on $[a, x_0]$, and monotonically decreasing on $[x_0, b]$. Thus $f(x_0)$ is precisely the maximum of f. □

Example. We consider $f(x) := x^x$ on $(0, \infty)$. Using Chain rule and product law we find

$$f'(x) = \big(\exp(x \log x)\big)' = \exp(x \log x)\big(1 \log x + x\frac{1}{x}\big) = x^x\big(\log x + 1\big).$$

We know a zero of the derivative: $\log \frac{1}{e} = -1$. Indeed, from the functional equation III.(10) follows $0 = \log(x\frac{1}{x}) = \log x + \log \frac{1}{x}$, and so $0 = \log e + \log \frac{1}{e}$. Moreover, we claim $\log x + 1$ is strictly monotone on $(0, \infty)$. Indeed, its derivative $\frac{1}{x}$ is positive on $(0, \infty)$, and the claim follows from Thm. 12. Since $x^x > 0$ for all $x > 0$ we can conclude that $f'(x) < 0$ for $0 < x < \frac{1}{e}$, and $f'(x) > 0$ for $x > \frac{1}{e}$. That is, the function $f(x) = x^x$ satisfies the conditions of Lemma 13, bottom. Consequently, x^x has a unique strict global minimum at $x = \frac{1}{e}$ with value $1/e^{1/e} = 1/\sqrt[e]{e}$.

In order to discuss a sufficient condition which requires only knowledge on f and its derivatives at a single point, let us define recursively the *n-th derivative*, for $n \in \mathbb{N}$, by setting $f^{(n)} := \big(f^{(n-1)}\big)'$, $f^{(1)} := f'$. We also write $f''(x) := (f')'(x)$ for the second derivative.

Theorem 14. *Let $f\colon (a,b) \to \mathbb{R}$ be twice continuously differentiable (that is, f'' exists and is continuous). If, for some $x_0 \in (a,b)$,*

$$f'(x_0) = 0 \text{ and } \left\{ \begin{smallmatrix} f''(x_0)<0 \\ f''(x_0)>0 \end{smallmatrix} \right\},$$

then f attains a strict local $\left\{ \begin{smallmatrix} maximum \\ minimum \end{smallmatrix} \right\}$ at x_0.

Note that this condition is not necessary for an extremum: x^4 takes a strict minimum at 0, however, $f''(0) = 0$, and so f does not conform to the condition.

Proof. Let us consider the bottom case $0 < f''(x_0)$. Cor. III.10 gives that the continuous function f'' is still positive in a small neighbourhood of x_0; that is, there is $\delta > 0$ such that $f''(\xi) > 0$ for all $\xi \in (x_0 - \delta, x_0 + \delta)$.

Let us now consider $x \neq x_0$ arbitrary within this interval. The Mean Value Theorem, applied to $f' \colon [x_0 - \delta, x_0 + \delta] \to \mathbb{R}$, then gives for each such x:

$$\frac{f'(x)}{x - x_0} = \frac{f'(x) - f'(x_0)}{x - x_0} \overset{\text{MVT}}{=} f''(\xi) \quad \text{for some } \xi \in (x_0 - \delta, x_0 + \delta).$$

But $f''(\xi) > 0$ and so $f'(x)$ and $x - x_0$ have the same sign, meaning that f' changes sign at x_0. Thus we are in the bottom case of (19) in Lemma 13, and so x_0 is a strict minimum on $(x_0 - \delta, x_0 + \delta)$. We conclude x_0 is a strict local minimum. $\square$

Example. Let us apply the theorem to our previous example $f(x) = x^x$. Certainly, we have $f(\frac{1}{e}) = 0$. Moreover,

$$f''(x) = \big(x^x(\log x + 1)\big)' = x^x(\log x + 1)^2 + x^x \frac{1}{x};$$

at $x = \frac{1}{e}$, the first term vanishes, while the second is positive. Hence $f''(\frac{1}{e}) > 0$, and we are in the bottom case of Thm. 14. Therefore, f takes a strict local minimum at $\frac{1}{e}$. Note that the application of Lemma 13 gave a stronger statement, namely that the minimum is global.

3. Integration

In antiquity, Archimede determined the volume of special bodies such as the cone, sphere, and cylinder. To calculate areas or volumes in general is the main task of integration. The first attempt for a systematic treatment of integration goes back to Cavalieri in the 17th century.

The integral of a function of one variable is the oriented area content bounded by the graph. Two questions arise:
- For which functions can we declare the integral?
- How do we compute integrals?

The answer to the second question will be deferred until Section 4: The Fundamental Theorem of Calculus will turn out to be crucial.

The first question has less practical impact; in fact, all functions of daily life are integrable. It is, however, an interesting mathematical problem. We will approach it as follows. We set off with step functions, for which integration is obvious, and then use a limit process to extend the integral to a large class of functions. Suprisingly, this class is not explicit, and so, in a second step, we will show that, for instance, continuous functions belong to this class.

3.1. **Step functions.** A function $\varphi\colon [a,b] \to \mathbb{R}$ is a *step function [Treppenfunktion]*, if there is a *partition [Zerlegung]* $a = x_0 < x_1 < \ldots < x_{n-1} < x_n = b$ of $[a,b]$, such that φ is constant on each interval (x_{k-1}, x_k) for $k = 1, \ldots, n$. Note that the number n of steps is finite, and we do not constrain the values $\varphi(x_k)$.

Let us denote the set of step functions on $[a,b]$ by $S[a,b]$. The sole point of introducing these functions is that their integrals are obvious, as we know the area of a rectangle:

Definition (Integral of step functions)**.** Let $\varphi \in S[a,b]$ with $\varphi(x) = c_k$ on (x_{k-1}, x_k) for $k = 1, \ldots, n$. Then we set

$$\int_a^b \varphi(x)\,dx := \sum_{k=1}^{n} c_k(x_k - x_{k-1}). \tag{20}$$

We also admit $a = b$, in which case the sum is empty and $\int_a^a \varphi(x)\,dx := 0$.

The same step function can be described with respect to many different partitions, for instance, we can always include additional support points into a given partition. Then $\int_a^b \varphi(x)\,dx$ remains invariant:

• For just one additional support point, this is seen as follows: If $\varphi(x) := c$ on $[a,b]$ and $\xi \in (a,b)$, then

$$c(b-a) = c(\xi - a) + c(b - \xi), \tag{21}$$

just as rectangle areas add.

• For the general case, if X is a partition $a = x_0 < x_1 < \ldots < x_i = b$ and Y is $a = y_0 < y_1 < \ldots < y_j = b$ then their union forms a partition Z of form $a = z_0 < z_1 < \ldots < z_k = b$, having $k \le i + j$ points. Appealing to (21), we see that the sums with respect to X and Z are equal, and so are the sums with respect to Y and Z. Consequently, the sums for X and Y are also equal, which means the integral is well-defined.

Using the union of two partitions, we also see that the sum of two step functions is once again a step function, and so is a scalar multiple. Therefore, the set of step functions $S[a,b]$ forms a vector space. On this vector space, the integral is a linear functional.

Proposition 15. *Let $\varphi, \psi \in S[a,b]$ and $\lambda \in \mathbb{R}$, then:*
(i) $\int_a^b \lambda\varphi + \psi\,dx = \lambda \int_a^b \varphi\,dx + \int_a^b \psi\,dx$ *(linearity)*
(ii) For $a \le \xi \le b$ *we have* $\int_a^\xi f + \int_\xi^b f = \int_a^b f$.
(iii) $\varphi \le \psi \implies \int_a^b \varphi\,dx \le \int_a^b \psi\,dx$.

In (*iii*), the notation $\varphi \le \psi$ is shorthand for $\varphi(x) \le \psi(x)$ for all $x \in [a,b]$. For any linear functional, property (*ii*) is called *monotonicity [Monotonie]*.

3.2. **The Riemann integral.**

Definition (Lower and upper integral *[Unter- und Oberintegral]*)**.** Suppose $f\colon [a,b] \to \mathbb{R}$ is an arbitrary bounded function. Then we set

$$L := \left\{ \int_a^b \varphi(x)\,dx : \varphi \in S[a,b],\ \varphi \le f \right\}, \quad U := \inf \left\{ \int_a^b \varphi(x)\,dx : \varphi \in S[a,b],\ \varphi \ge f \right\}.$$

$$\underline{\int_a^b} f(x)\,dx := \sup L, \qquad \overline{\int_a^b} f(x)\,dx := \inf U$$

Since we assume $|f| \le C$ the set U contains the constant step function C and is nonempty. Moreover, U is bounded from below by $-C(b-a)$ and so $\inf U$ exists. Likewise for L.

Due to monotonicity it is immediate that $\underline{\int_a^b} f(x)\,dx \le \overline{\int_a^b} f(x)\,dx$.

Examples. 1. For $\varphi \in S[a,b]$ we have $\underline{\int_a^b}\varphi(x)\,dx = \overline{\int_a^b}\varphi(x)\,dx = \int_a^b \varphi(x)\,dx$ (why?).

2. Let $\chi_{\mathbb{Q}}\colon [0,1] \to \mathbb{R}$ be the *characteristic function of* $\mathbb{Q}$ with $f(x) = 1$ for $x \in \mathbb{Q}$, and 0 otherwise. Since the rational numbers are dense in the irrational ones, each step function $\varphi \ge f$ satisfies $\varphi \ge 1$ (except, perhaps, at the partition points), and so $\overline{\int_a^b}\chi_{\mathbb{Q}}(x)\,dx = 1$. Similarly, $\underline{\int_a^b}\chi_{\mathbb{Q}}(x)\,dx = 0$.

For the second example, the "area" of the graph is a doubtful quantity: Is it 0, 1, or any intermediate value? However, when upper and lower integral coincide, these should represent "the" area:

Definition (Riemann 1854)**.** (*i*) A bounded function $f\colon [a,b] \to \mathbb{R}$ is *(Riemann) integrable [(Riemann)-integrierbar]*, if $\underline{\int_a^b} f(x)\,dx = \overline{\int_a^b} f(x)\,dx$. In that case we write

$$\int_a^b f(x)\,dx := \underline{\int_a^b} f(x)\,dx = \overline{\int_a^b} f(x)\,dx.$$

(*ii*) We call $f\colon [a,b] \to \mathbb{C}$ *(Riemann) integrable* if $\operatorname{Re} f$ and $\operatorname{Im} f$ are Riemann integrable; then we set $\int_a^b f(x)\,dx := \int_a^b \operatorname{Re} f(x)\,dx + i\int_a^b \operatorname{Im} f(x)\,dx$.

The definition leaves open how to assert that a function is integrable. We will show below that functions which are continuous or monotone are integrable. The following reformulation of the integrability definition will be our test for integrability:

`24. Vorlesung, Donnerstag, 18.1.07 (T 12)` ______

Proposition 16. *A function $f\colon [a,b] \to \mathbb{R}$ is integrable if and only if there are sequences of step functions $\varphi_n^-, \varphi_n^+ \in S[a,b]$ with*

$$\varphi_n^- \le f \le \varphi_n^+ \qquad \textit{and} \qquad \int_a^b \varphi_\cdot^+(x)\,dx - \int_a^b \varphi_\cdot^-(x)\,dx \to 0.$$

For the proof, note that if f is integrable then there exist step functions in U and L, for each of the integral converges to the value $\int f$. Conversely, if the sequences exist then we must have equality in the inequality $\sup L \le \inf U$.

It is not too practical but nevertheless not impossible to compute integrals in terms of sequences of step functions $(\varphi_n^\pm)$. Effectively, this approximates an integral by the sums (20); for this reason they are called *Riemann sums [Riemannsche Summen]*.

Example. $f(x) = x$ is integrable with $\int_0^b x\,dx = \frac{1}{2}b^2$.
To define a sequence $\varphi_n^\pm(x) \in S[0,b]$ we use an equidistant partition of $[0,b]$ into $n \in \mathbb{N}$ intervals, with partition points

$$x_k^n := \frac{k}{n}b \qquad \text{for} \qquad k = 0,\dots,n.$$

We set $\varphi_n^\pm$ equal to the value of f on the left, respectively right, endpoint of each interval:

$$\varphi_n^-(x) := x_{k-1}^n, \qquad \varphi_n^+(x) := x_k^n, \qquad \text{if } x \in [x_{k-1}^n, x_k^n) \text{ for some } k = 1,\dots,n.$$

Moreover we set $\varphi_n^\pm(b) := f(b)$. Then $\varphi_n^-(x) \le x \le \varphi_n^+(x)$, and

$$\int_0^b \varphi_n^+(x)\,dx = \sum_{k=1}^n x_k^n \frac{b}{n} = \sum_{k=1}^n \left(\frac{k}{n}b\right)\frac{b}{n} = \frac{b^2}{n^2}\sum_{k=1}^n k = \frac{b^2}{n^2}\frac{n(n+1)}{2} = \frac{b^2}{2}\frac{n+1}{n} \quad \to \quad \frac{b^2}{2}.$$

Similarly, $\int_0^b \varphi_n^- = \frac{b^2}{n^2}\sum_{k=0}^{n-1} k = \frac{b^2}{2}\frac{n-1}{n} \to \frac{b^2}{2}$. Since the two limits agree, we have that $f(x) = x$ is integrable with $\int_0^b x\,dx = \frac{1}{2}b^2$.

Problem. Compute similarly $\int_0^b \cos t\,dt = \sin b$ for $0 < b < \pi$.

We can generalize the proof of the example to show:

Theorem 17. *Each monotone function $f\colon [a,b] \to \mathbb{R}$ is integrable.*

Proof. We partition $[a,b]$ equidistantly into k intervals, that is, we subdivide at

$$x_k^n := a + \frac{k}{n}(b-a), \qquad k = 0,\dots,n.$$

Let us consider the increasing case. We set

$$\varphi_n^-(x) := f(x_{k-1}^n), \quad \varphi_n^+(x) := f(x_k^n) \quad \text{if } x \in [x_{k-1}^n, x_k^n) \text{ for } k = 1,\dots,n,$$

as well as $\varphi_n^\pm(b) := f(b)$. By the monotonicity of f,

$$\varphi_n^- \le f \le \varphi_n^+.$$

Moreover,

$$\int_a^b \varphi_n^+(x)\,dx - \int_a^b \varphi_n^-(x)\,dx = \sum_{k=1}^n f(x_k^n)\underbrace{(x_k^n - x_{k-1}^n)}_{(b-a)/n} - \sum_{k=1}^n f(x_{k-1}^n)\underbrace{(x_k^n - x_{k-1}^n)}_{(b-a)/n}$$

$$= \frac{b-a}{n}\big(f(x_n^n) - f(x_0^n)\big) = \frac{1}{n}(b-a)\big(f(b)-f(a)\big) \;\to\; 0$$

This verifies the integration test Proposition 16. □

Outlook. There are other integration theories which can integrate more functions. The most well-known is the *Lebesgue integral*, which we will introduce to integrate functions of several variables, in the fourth term. For instance, the characteristic function of $\mathbb{Q}$ is still Lebesgue integrable, with integral 0.

3.3. **Uniform continuity and the integrability of continuous functions.** Are continuous functions integrable? As for monotone functions, we will consider an equidistant partition into n intervals. Then the obvious idea is to define step functions by taking the smallest, respectively largest, value of a given function f over each partition interval. To satisfy the integrability test we must verify the following property: Over subintervals of length $\frac{b-a}{n}$, the difference of the largest and the smallest value of f approaches zero as $n \to \infty$; for a sequence of intervals containing a fixed point this follows from continuity, but we are interested to verify this for all subintervals at the time.

Recall that a mapping $f\colon D \to \mathbb{C}$ is continuous at a point a if for all error margins $\varepsilon > 0$ there is $\delta(a,\varepsilon)$ such that $|f(z) - f(a)| < \varepsilon$ for all z with $|z-a| < \delta$. In the following definition we require that δ can be chosen independently of a, that is, the error bound ε on the values of the function is met over any δ-ball contained in D:

Definition. A mapping $f\colon D \to \mathbb{C}$, for $D \subset \mathbb{C}$ is *uniformly [gleichmäßig] continuous on D*, if for each $\varepsilon > 0$ there is a $\delta = \delta(\varepsilon) > 0$, such that

$$\big|f(z) - f(a)\big| < \varepsilon \quad \text{for all } x, a \in D \text{ with } |z-a| < \delta. \tag{22}$$

Examples. 1. The function $z \mapsto 17z$ is uniformly continuous over any domain $D \subset \mathbb{C}$, since $\delta := \frac{\varepsilon}{17}$ satisfies (22).

2. The function $f(x) = \sin\frac{1}{x}$ on $(0,1]$ is continuous, but not uniformly continuous. Indeed, no matter how small δ is chosen, there is a pair of points $x_\pm \in (0,\delta)$ with $f(x_\pm) = \pm 1$ (specify $x_\pm$!). Hence, for $\varepsilon := 1$ no $\delta > 0$ can satisfy (22).

3. $f(x) := x^2$ is not uniformly continuous on all of $\mathbb{R}$. The problem arises as $x \to \pm\infty$, where the graph becomes arbitrarily steep. Indeed, two points $a < x := a + \delta$ have

$$\big|f(a) - f(x)\big| = a^2 - (a+\delta)^2 = 2a\delta + \delta^2 > 2a\delta.$$

This cannot be less than ε independently of a, so that (22) is violated.

4. The function $\frac{1}{x}$ is not uniformly continuous on $(0, 1]$. Like in the previous example, the problem arises as $x \to 0$. (For the explicit calculation see [F], p.103.)

Problems. 1. Do the necessary calculations to make the above examples rigorous.

2. Lipschitz-continuous functions, for which (18) holds, are uniformly continuous.

3. Show (conversely) that $\sqrt[3]{\cdot}\colon [-1, 1] \to \mathbb{R}$ is uniformly continuous, but not Lipschitz.

It is not by chance that our counterexamples to uniform continuity are defined on intervals which are unbounded (as in 3.), or not closed (2. and 4.):

Theorem 18. *Each continuous function $f\colon [a, b] \to \mathbb{C}$ is uniformly continuous.*

Note that the functions of the examples 2. and 4. are uniformly continuous on each interval $[\gamma, 1]$, with $\gamma > 0$; however, they do not have a continuous extension to the interval $[0, 1]$.

Proof. Indirectly. We suppose f is not uniformly continuous. Then, for some $\varepsilon > 0$ no $\delta > 0$ will satisfy (22). In particular, for this ε, the choice $\delta := \frac{1}{n}$ will not satisfy (22) for any $n \in \mathbb{N}$. Thus there are pairs of points $x_n, a_n \in [a, b]$ violating (22), that is,

$$\big|f(x_n) - f(a_n)\big| \geq \varepsilon \quad \text{for } |x_n - a_n| < \frac{1}{n}. \tag{23}$$

Since $[a, b]$ is bounded, the Theorem of Bolzano-Weierstrass (Thm. II.12) allows us to pick a convergent subsequence $(x_{n_k})_{k\in\mathbb{N}}$ of (x_n), with $x_{n_k} \to x \in \mathbb{R}$ as $k \to \infty$. By Prop. 8, from $a \leq x_n \leq b$ we can conclude $x \in [a, b]$. We claim that also $a_{n_k} \to x$. Indeed, $|x_n - a_n| < \frac{1}{n}$ and so

$$|x - a_{n_k}| \leq \underbrace{|x - x_{n_k}|}_{\to 0} + \underbrace{|x_{n_k} - a_{n_k}|}_{\leq 1/n_k \to 0} \to 0 \qquad \text{as } k \to \infty.$$

But since f is continuous at the point x we have $|f(x_{n_k}) - f(a_{n_k})| \to |f(x) - f(x)| = 0$, in contradiction with (23). $\square$

Problem. To understand the proof better, check where precisely it fails when applied to Examples 2. to 4.

We can now return to our integrability problem.

Theorem 19. *Each continuous function $f\colon [a, b] \to \mathbb{C}$ is integrable.*

Proof. Suppose f is real-valued. For each $\varepsilon > 0$ we find two step functions (φ^-), (φ^+) which sandwich f in the sense $\varphi^- \le f \le \varphi^+$ and satisfy

$$0 \le \varphi_n^+(x) - \varphi_n^-(x) \le \frac{\varepsilon}{b-a} \quad \text{for all } x \in [a,b]. \tag{24}$$

Assuming (24) gives

$$\int_a^b \varphi^+(x)\,dx - \int_a^b \varphi^-(x)\,dx = \int_a^b \varphi^+(x) - \varphi^-(x)\,dx \overset{\text{monotonicity}}{\le} \int_a^b \frac{\varepsilon}{b-a}\,dx = \varepsilon.$$

Repeating this argument for a sequence $\varepsilon = \varepsilon_n \to 0$, we see that the integrability test Proposition 16 is satisfied.

Since f is uniformly continuous, there is $\delta = \delta(\varepsilon) > 0$ such that $|f(x) - f(y)| < \frac{\varepsilon}{b-a}$ for $|x-y| < \delta$. We choose $n = n(\varepsilon) \in \mathbb{N}$ such that $\frac{b-a}{n} < \delta$.

To define the step functions, let us partition $[a,b]$ equidistantly into n intervals. That is, we choose support points

$$x_k := a + \frac{k}{n}(b-a), \qquad k = 0, \dots, n,$$

and intervals $I(k) = [x_{k-1}, x_k)$ for $k = 1, \dots, n$. We consider bounds for f over $I(k)$, namely $m_k := \inf\{f(x) : x \in I(k)\}$ and $M_k := \sup\{f(x) : x \in I(k)\}$. We set

$$\varphi^-(x) := m_k, \quad \varphi^+(x) := M_k \qquad \text{for } x \in I(k) \text{ and } k = 1, \dots, n,$$

as well as $\varphi^\pm(b) := f(b)$. Since each interval $I(k)$ has length $\frac{b-a}{n} < \delta$ we have that $f(x) - f(y) < \frac{\varepsilon}{b-a}$ for any pair of points $x, y \in I(k)$. Thus also

$$M_k - m_k = \sup_{x \in I(k)} f(x) - \inf_{y \in I(k)} f(y) = \sup_{x,y \in I(k)} \big(f(x) - f(y)\big) \le \frac{\varepsilon}{b-a},$$

which establishes (24).

If f is complex-valued, apply the preceding proof to $\operatorname{Re} f$ and $\operatorname{Im} f$. $\square$

`25. Vorlesung, Dienstag, 23.1.07 (Ü 12)` ______________________________

3.4. **Mean Value Theorem of Integration.** For each interval $[a,b]$, the Riemann-integrable functions form an (infinite-dimensional) vector space, and the integral is a linear monotone functional on this vector space:

Proposition 20. *Let $f, g\colon [a,b] \to \mathbb{C}$ be integrable and $\lambda \in \mathbb{C}$. Then also $\lambda f + g$ is integrable and:*
(i) $\int_a^b \lambda f + g\,dx = \lambda \int_a^b f\,dx + \int_a^b g\,dx$ *(linearity),*
(ii) For $a \le \xi \le b$ we have $\int_a^\xi f + \int_\xi^b f = \int_a^b f$.

(iii) If f, g are real valued, then $f \le g$ implies $\int_a^b f\,dx \le \int_a^b g\,dx$ (monotonicity).
(iv) We have $\left|\int_a^b f(x)\,dx\right| \le \int_a^b |f(x)|\,dx$.

Proof. (i) Suppose f, g, λ are real. By Proposition 16 there are sequences of step functions $\varphi_n^\pm, \psi_n^\pm \in S[a,b]$ with

$$\varphi_n^- \le f \le \varphi_n^+ \qquad \text{and} \qquad \int \varphi_n^+ - \int \varphi_n^- \to 0,$$
$$\psi_n^- \le g \le \psi_n^+ \qquad \text{and} \qquad \int \psi_n^+ - \int \psi_n^- \to 0.$$

Consider the case $\lambda \ge 0$. We multiply the first equation with λ, and add it to the second:

$$\lambda\varphi_n^- + \psi_n^- \le \lambda f + g \le \lambda\varphi_n^+ + \psi_n^+ \quad \text{and} \quad \int \lambda\varphi^+ + \psi^+ - \int \lambda\varphi^- + \psi^- \to 0;$$

here we invoked Prop. (15)(i) and the the limit laws for sequences. Thus the sequences of step functions $(\lambda\varphi^\pm + \psi^\pm)$ show the integrability of $\lambda f + g$ in view of the test Proposition 16. In case $\lambda < 0$, similarly with $(\lambda\varphi^\mp - \psi^\pm)$. For f, g, λ complex, apply the real result to real and imaginary part of $\lambda f + g$.

(ii) This holds for step functions and therefore for each integrable function.

(iii) Let $\varphi^- \in S[a,b]$ with $\varphi^- \le f$. Then also $\varphi^- \le g$. Thus the lower integral for g is taken over a larger set of step functions than the lower integral for f. Passing to the suprema gives $\underline{\int_a^b} f \le \underline{\int_a^b} g$. Since f and g are integrable, this means $\int_a^b f \le \int_a^b g$.

(iv) Let us first give the proof for the case f is real valued. Then $\pm f \le |f|$ and monotonicity implies $\pm \int f \le \int |f|$, as desired. We leave the somewhat tedius proof that $|f|$ is integrable to the reader.

We now extend the real case to complex valued f. Let $t \in \mathbb{R}$ be arbitrary. Then, using the real result, we find

$$\int_a^b |f(x)|dx = \int_a^b |e^{it} f(x)|dx \ge \int_a^b \operatorname{Re}\left(e^{it} f(x)\right) dx = \operatorname{Re}\left(e^{it} \int_a^b f(x)\,dx\right).$$

In the particular case $\int_a^b f(x)\,dx = 0$ this proves the claim. Else, for $t := -\arg\left(\int_a^b f(x)\,dx\right)$ we have $e^{it} \int_a^b f(x)\,dx > 0$, which gives

$$\operatorname{Re}\left(e^{it} \int_a^b f(x)\,dx\right) = \left|\int_a^b f(x)\,dx\right|.$$

□

The value $\frac{1}{b-a} \int_a^b f(x)\,dx$ can be interpreted as an average value of f over the interval $[a,b]$. If f is continuous, this average is attained:

Theorem 21 (Mean Value Theorem of Integration)**.** *Let $f\colon [a,b] \to \mathbb{R}$ be continuous, and $a < b$.*
(i) Then there is $\xi \in [a,b]$, such that

$$\frac{1}{b-a}\int_a^b f(x)\,dx = f(\xi).$$

(ii) (Generalized form) *If furthermore $g\colon [a,b] \to [0,\infty)$ (or $g\colon [a,b] \to (-\infty,0]$) is continuous, then there exists $\xi \in [a,b]$ with*

$$\int_a^b f(x)g(x)\,dx = f(\xi)\int_a^b g(x)\,dx.$$

Proof. We set

$$m := \min\{f(x) : x \in [a,b]\}, \qquad M := \max\{f(x) : x \in [a,b]\}$$

(i) The constant step functions m, M sandwich f, that is, $m \le f \le M$. Monotonicity as stated in Proposition 20 (ii) gives

$$m = \frac{1}{b-a}\int_a^b m\,dx \overset{\text{mon.}}{\le} \frac{1}{b-a}\int_a^b f(x)\,dx \le \frac{1}{b-a}\int_a^b M\,dx = M.$$

By the Intermediate Value Theorem III.11 the continuous function $f(x)$ attains each value inbetween m and M. So in particular it attains the value $\frac{1}{b-a}\int_a^b f(x)\,dx$.

(ii) Consider the case $g \ge 0$. Then $mg \le fg \le Mg$ implies

$$m\int_a^b g(x)\,dx \le \int_a^b f(x)g(x)\,dx \le M\int_a^b g(x)\,dx.$$

Thus the claim follows by applying the IVT to the continuous function $x \mapsto f(x)\int_a^b g(t)\,dt$. In case $g \le 0$ we have $Mg \le fg \le mg$, and the same conclusion holds. $\square$

Problem. Which special choice of g in (ii) gives (i)? Moreover, show that (i) doesn't hold when f is not assumed to be continuous, and that (ii) can fail when g takes both signs.

4. The link between integration and differentiation

We can now state the most important theorem of one-variable calculus. It has many applications, for instance it serves us to compute integrals.

4.1. **The Fundamental Theorem of Calculus.** A differentiable function $F\colon [a,b] \to \mathbb{C}$ is called a *primitive* or *antiderivative [Stammfunktion]* of $f\colon [a,b] \to \mathbb{C}$, if $F' = f$.

Examples. (i) For $f(x) = x^2$ the function $F(x) = \frac{1}{3}x^3$ is a primitive.
(ii) For $f(x) = e^{ix}$ the function $F(x) = -ie^{ix}$ is a primitive.

Often, the explicit form of a primitive can only be guessed. Nevertheless it always exists for f continuous:

Proposition 22. *Let* $f\colon [a,b] \to \mathbb{C}$ *be continuous. Then the* indefinite integral *[unbestimmtes Integral]*

$$I(x) := \int_a^x f(t)\,dt$$

gives a differentiable function $I\colon [a,b] \to \mathbb{C}$*. Moreover,* I *is a primitive of* f*, that is,* $I'(x) = f(x)$*.*

Problem. If f is merely integrable then I is only Lipschitz.

Proof. We first suppose f is real-valued and compute the difference quotient of $I(x)$. Suppose $a \le x < b$. Then for sufficiently small $h > 0$ we have $x + h < b$. For such h follows

$$\frac{I(x+h) - I(x)}{h} = \frac{1}{h}\Big(\int_a^{x+h} f(t)\,dt - \int_a^x f(t)\,dt\Big) = \frac{1}{h}\int_x^{x+h} f(t)\,dt. \tag{25}$$

Moreover, by the Mean Value Theorem of Integration, Thm. 21, there exists $\xi_h \in [x, x+h]$ with

$$\frac{1}{h}\int_x^{x+h} f(t)\,dt = f(\xi_h).$$

Now as $h \to 0$ we have $\lim_{h\to 0} \xi_h = x$, and so the limit of (25) exists:

$$I'(x) = \lim_{h\to 0} f(\xi_h) \overset{f \text{ continuous}}{=} f(x)$$

In case $a < x \le b$ we can similarly consider $h < 0$ and proceed as before: Then $\frac{I(x+h)-I(x)}{h} = -\frac{1}{h}\int_{x+h}^x f(t)\,dt = -\frac{1}{h}|h| f(\xi_h) = f(\xi_h)$ for some $\xi_h \in [x+h, x]$ which again implies $I'(x) = f(x)$. Finally, if f is complex valued, we apply the above to the real and imaginary parts. □

Let us rephrase the statement, which presents, perhaps, the most important fact of calculus. The equation $\frac{d}{dx}\int_a^x f(t)\,dt = f(x)$ means that indefinite integration and differentiation are inverse operations, cancelling one another. This is not at all clear from the definition of integral and derivative!

For f constant, $f(x) \equiv c$, this is immediate to see: $I(x) = (x-a)c$ and so $I'(x) = c$. The same formula for $I(x)$ holds up to an arbitrarily small error, when f more generally is a continuous function. So we might expect the equation $I'(x) = f(x)$.

Obviously, when F is a primitive of f, then so is $F + c$ for c constant. Conversely, any two primitives $F, G\colon [a,b] \to \mathbb{C}$ of the same function f satisfy

$$(F-G)' = F' - G' = f - f = 0;$$

Corollary 10 implies that $F - G$ is constant. That is, a primitive of f is well-defined up to a constant. Making use of this property we see that the integral of f can be computed using any of its primitives F:

Theorem 23 (Fundamental theorem). *Suppose a continuous function $f\colon [a,b] \to \mathbb{C}$ has a primitive $F\colon [a,b] \to \mathbb{C}$. Then*

$$\int_a^b f(x)\,dx = F(b) - F(a).$$

Proof. By Prop. 22, the function $I(x) := \int_a^x f(t)\,dt$ is a primitive of f. Hence $F(x) - I(x)$ is constant, say equal to $c \in \mathbb{C}$, and

$$\int_a^b f(x)\,dx = I(b) - \underbrace{I(a)}_{=0} = \big(F(b) - c\big) - \big(F(a) - c\big) = F(b) - F(a). \qquad \square$$

The Fundamental Theorem allows us to integrate most functions introduced so far. It will be convenient to write $F(x)\big|_a^b := F(b) - F(a)$.

Examples. From the examples for differentiation, the following is immediate:
1.
$$\int_a^b e^x\,dx = e^x\Big|_a^b, \qquad \int_a^b e^{ix}\,dx = -ie^{ix}\Big|_a^b.$$
Invoking the Euler formula and taking real and imaginary parts of the second integral or (14) we find
$$\int_a^b \cos x\,dx = \sin x\Big|_a^b, \qquad \int_a^b \sin x\,dx = -\cos x\Big|_a^b.$$
Moreover,
$$\int_a^b \frac{1}{1+x^2}\,dx = \arctan x\Big|_a^b,$$
and, provided $[a,b]$ does not contain a zero of cosine,
$$\int_a^b \frac{1}{\cos^2 x}\,dx = \tan x\Big|_a^b.$$

2. We calculated $\big(x^{n+1}\big)' = (n+1)x^n$ for three cases: (i) $x \in \mathbb{R}$ and $n \in \mathbb{N}_0$, (ii) $x \neq 0$ and $n \in \mathbb{Z}$, (iii) $x > 0$ and $n \in \mathbb{R}$. Thus for $n \neq -1$ we deduce

$$\int_a^b x^n\,dx = \frac{1}{n+1}x^{n+1}\Big|_a^b, \tag{26}$$

for (i) $a, b \in \mathbb{R}$, $n \in \mathbb{N}_0$, (ii) $0 \notin [a,b]$, $n \in \mathbb{Z}\setminus\{-1\}$, (iii) $0 < a \leq b$, $n \in \mathbb{R}\setminus\{-1\}$. Thanks to the linearity of the integral this formula suffices to integrate polynomials.
The remaining case $n = -1$ is settled using $(\log x)' = \frac{1}{x}$ for $x > 0$:

$$\int_a^b \frac{1}{x}\,dx = \log x\Big|_a^b, \qquad \text{for } 0 < a < b;$$

similarly, when $x < 0$ we have $\big(\log(-x)\big)' = \frac{1}{-x}(-1) = \frac{1}{x}$ (Chain Rule) and so $\int_a^b \frac{1}{x}\,dx = \log(-x)|_a^b$ when $a < b < 0$.

`26. Vorlesung, Donnerstag, 25.1.07 (T 13)` ______________________________

4.2. **Rules for integration.** Each law of differentiation yields a law for integration, via the Fundamental Theorem.

Let us call a function *continuously differentiable [stetig differenzierbar]* if its derivative is continuous. Such a function is continuous itself according to Corollary 2.

We consider the product law first.

Theorem 24 (Integration by parts)**.** *If $f, g\colon [a,b] \to \mathbb{C}$ are continuously differentiable, then*

$$\int_a^b f'(x)g(x)\,dx = f(x)g(x)\Big|_a^b - \int_a^b f(x)g'(x)\,dx.$$

Note the two integrals on the right hand side exist in view of our assumptions on f, g.

Proof. The function $h := fg$ can be differentiated using product law: $h' = f'g + fg'$. In particular, h' is continuous, and so

$$\int_a^b f'(x)g(x)\,dx + \int_a^b f(x)g'(x)\,dx = \int_a^b h'(x)\,dx \overset{\text{Fund'l Thm.}}{=} h(x)\Big|_a^b = f(x)g(x)\Big|_a^b.$$

□

Examples. 1. To integrate $\log x$, we set $f(x) := x$ and $g(x) := \log x$.

$$\int_a^b \log x\,dx = x\log x\Big|_a^b - \int_a^b x\frac{1}{x}\,dx = x\log x - x\Big|_a^b \qquad \text{for } 0 < a < b$$

This can as well be tried for for any function with known derivative. For instance, for $\arctan x =: g(x)$ we obtain

$$\int_a^b \arctan x\,dx = x\arctan x\Big|_a^b - \int_a^b x\frac{1}{1+x^2}\,dx = x\arctan x - \frac{1}{2}\log(1+x^2)\Big|_a^b.$$

Indeed, according to the Chain Rule, $\big(\log(1+x^2)\big)' = \frac{1}{1+x^2}\,2x$.

3. *Claim*:

$$\int_{-\pi/2}^{\pi/2} \sin^2 x\,dx = \int_{-\pi/2}^{\pi/2} \cos^2 x\,dx = \frac{\pi}{2}. \tag{27}$$

Proof: With $f(x) := -\cos x$ and $g(x) := \sin x$ we obtain on the one hand

$$\int_{-\pi/2}^{\pi/2} \sin^2 x\,dx = -\cos x \sin x\Big|_{-\pi/2}^{\pi/2} + \int_{-\pi/2}^{\pi/2} \cos^2 x\,dx = \int_{-\pi/2}^{\pi/2} \cos^2 x\,dx;$$

on the other hand, $\sin^2 x + \cos^2 x = 1$, and so

$$\int_{-\pi/2}^{\pi/2} \sin^2 x\,dx + \int_{-\pi/2}^{\pi/2} \cos^2 x\,dx = \int_{-\pi/2}^{\pi/2} 1\,dx = \pi.$$

We now discuss the Chain Rule. Let us first introduce some more notation. Suppose $F, f = F' \colon [a,b] \to \mathbb{R}$ and $x, y \in [a,b]$. Then the Fundamental Theorem gives $F(y) - F(x) = \int_x^y f(t)\,dt$. The same formula will hold for $x > y$ as well provided we set

$$\int_x^y f(t)\,dt := -\int_y^x f(t)\,dt. \tag{28}$$

Theorem 25 (Substitution)**.** *Let $f \colon [\alpha,\beta] \to \mathbb{C}$ be continuous and $\varphi \colon [a,b] \to [\alpha,\beta]$ be continuously differentiable. Then*

$$\int_a^b f\big(\varphi(t)\big)\varphi'(t)\,dt = \int_{\varphi(a)}^{\varphi(b)} f(x)\,dx. \tag{29}$$

If, moreover, φ is invertible between the above intervals, then

$$\int_\alpha^\beta f(x)\,dx = \int_{\varphi^{-1}(\alpha)}^{\varphi^{-1}(\beta)} f\big(\varphi(t)\big)\varphi'(t)\,dt. \tag{30}$$

Physicists like to write (30) in a form which makes it easy to memorize: Setting $x = x(t)$ gives $\int_{x_1}^{x_2} f(x)\,dx = \int_{t_1}^{t_2} f\big(x(t)\big)\frac{dx}{dt}\,dt$. This notation is problematic as it mixes the names of functions with variables.

Proof. Let $F \colon [\alpha,\beta] \to \mathbb{R}$ be a primitive of f. According to the Chain Rule,

$$\big(F\circ\varphi\big)'(t) = F'\big(\varphi(t)\big)\varphi'(t) = f\big(\varphi(t)\big)\varphi'(t),$$

and so (29) follows from

$$\int_a^b f\big(\varphi(t)\big)\varphi'(t)\,dt \overset{\text{Fund'l Thm.}}{=} \big(F\circ\varphi\big)(t)\Big|_a^b = F\big(\varphi(b)\big) - F\big(\varphi(a)\big) \overset{\text{Fund'l Thm.}}{=} \int_{\varphi(a)}^{\varphi(b)} f(x)\,dx.$$

Recall from Thm. III.13 that a continuous function which is bijective between intervals must be strictly monotone. Thus, for φ increasing, we have $\alpha = \varphi(a)$ and $\beta = \varphi(b)$, so that (30) is immediate from (29). In the decreasing case, we have $\beta = \varphi(a)$ and $\alpha = \varphi(b)$, and (30) follows with endpoints exchanged. But applying (28) to both sides we establish (the negative of) (30). $\square$

Examples. 1. Integration is invariant under translation in the domain: For $c \in \mathbb{R}$,

$$\int_a^b f(\underbrace{t+c}_{\varphi(t)})\,dt \overset{(29)}{=} \int_{\varphi(a)=a+c}^{\varphi(b)=b+c} f(x)\,dx \qquad (\varphi'(t)=1).$$

2. For $c \in \mathbb{R}$ and $\varphi(t) := ct$ we have

$$\int_a^b f(ct)c\,dt \overset{(29)}{=} \int_{ca}^{cb} f(x)\,dx \qquad \overset{c\neq 0}{\Longrightarrow} \qquad \int_a^b f(ct)\,dt = \frac{1}{c}\int_{ca}^{cb} f(x)\,dx.$$

3. For $f(x) = \frac{1}{x}$ and $\varphi(t) > 0$ continuously differentiable, we find

$$\int_a^b \frac{1}{\varphi(t)}\varphi'(t)\,dt \overset{(29)}{=} \int_{\varphi(a)}^{\varphi(b)} \frac{1}{x}\,dx = \log x\Big|_{\varphi(a)}^{\varphi(b)} = \log\big(\varphi(t)\big)\Big|_a^b.$$

(Conversely, it is straightforward that the *logarithmic derivative* of φ is $(\log\varphi)' = \frac{\varphi'}{\varphi}$.) As an application, we integrate tan over $[a,b] \subset (-\frac{\pi}{2}, \frac{\pi}{2})$:

$$\int_a^b \tan t\,dt = -\int_a^b \frac{-\sin t}{\cos t}\,dt = -\log\cos t\Big|_a^b$$

4. Let us now discuss a classical problem: the area of the unit disk. The area of the upper half disk is the integral $\int_{-1}^1 \sqrt{1-x^2}\,dx$. We want to substitute x by $\varphi(t) := \sin t$ in order to take advantage of the identity $\sin^2 t + \cos^2 t = 1$. Note that $\varphi\colon [-\frac{\pi}{2}, \frac{\pi}{2}] \to [-1,1]$ is continuously differentiable and invertible. Substitution gives

$$\int_{-1}^1 \sqrt{1-x^2}\,dx \overset{(30)}{=} \int_{\varphi^{-1}(-1)}^{\varphi^{-1}(1)} \underbrace{\sqrt{1-\sin^2 t}}_{\sqrt{\cos^2 t}}\,\underbrace{(\sin t)'}_{\cos t}\,dt = \int_{-\pi/2}^{\pi/2} \cos^2 t\,dt \overset{(27)}{=} \frac{\pi}{2}.$$

Here, we used the fact $\cos t \geq 0$ for $t \in [-\frac{\pi}{2}, \frac{\pi}{2}]$. Thus the unit disk has area π.

Improper integrals [uneigentliche Integrale] of the type $\int_a^\infty f(x)\,dx := \lim_{b\to\infty}\int_a^b f(x)\,dx$ will be discussed in the problem session.

Remarks. 1. Unlike for the rules of differentiation, which can be applied mechanically, the application of the rules of integration can require experience and skill. Thus symbolic integration is a useful feature of mathematical software. The predicted integrals are straightforward to check using the rules for differentiation.

2. Complicated composed functions can always be differentiated using the rules of differentiation; however, to integrate them, the rules of integration can be insufficient. For instance, the functions $\frac{1}{\log x}$, e^{-x^2}, $\frac{\sin x}{x}$ cannot be integrated in terms of "elementary" functions. Thus integration is a limit process which can be used to create new functions.

5. Integration and differentiation of sequences of functions

5.1. **Pointwise convergence.** The exponential and trigonometric functions were introduced by series. We want to consider such sequences in general, thereby returning to the complex setting:

Definition. (i) A sequence of functions $f_n\colon D \to \mathbb{C}$ (for $D \subset \mathbb{C}$) *converges (pointwise)* to $f\colon D \to \mathbb{C}$, if $f_n(z) \to f(z)$ for each $z \in D$.
(ii) In particular, a (complex) *power series [Potenzreihe]* $f(z) = \sum_{k=0}^{\infty} c_k z^k$ is the pointwise limit of the sequence of partial sums $f_n(z) := \sum_{k=0}^{n} c_k z^k$ (for $z \in \mathbb{C}$).

Taylor's formula, discussed in the next section, provides a tool to represent an arbitrary function by a power series.

Example. The function $\frac{\sin z}{z}$ has a series representation, obtained from multiplying the series for $\sin z$ with the number $\frac{1}{z}$:

$$\frac{\sin z}{z} = \frac{1}{z}\Big(z - \frac{z^3}{3!} + \frac{z^5}{5!} \mp \dots\Big) = 1 - \frac{z^2}{3!} + \frac{z^4}{5!} \mp \dots, \qquad z \neq 0. \tag{31}$$

Suppose we are given a function by a power series. If we wish to differentiate or integrate it, then the obvious idea is to do this term by term. This will be a particularly useful method for integration, since functions like $\frac{\sin t}{t}$ do not have an elementary integral, as pointed out before. But we could still integrate termwise

$$\int_0^x \frac{\sin t}{t}\,dt = \int_0^x 1 - \frac{t^2}{3!} + \frac{t^4}{5!} \mp \dots\,dt \overset{?}{=} x - \frac{x^3}{3\cdot 3!} + \frac{x^5}{5\cdot 5!} - \frac{x^7}{7\cdot 7!} \pm \dots \tag{32}$$

(*integral sine*), provided we can justify this for infinite sums. In fact, it is not clear that the function we consider is integrable: Is the series in (31) continuous at 0? We will ultimately confirm this, and show that we also can take limits termwise,

$$\lim_{x\to 0} \frac{\sin x}{x} = 1 - \frac{0^2}{3!} + \frac{0^4}{5!} \mp \dots = 1.$$

More generally, given a sequence of functions $(f_n)_{n\in\mathbb{N}}$ with pointwise limit $f(x) := \lim f_n(x)$, we are interested if the following holds:

$$\begin{aligned} &f_n \text{ continuous} &&\overset{?}{\Rightarrow} && f \text{ continuous ,} \\ (33)\qquad &f_n \text{ integrable} &&\overset{?}{\Rightarrow} && f \text{ integrable with } \lim_{n\to\infty}\int_a^b f_n(x)\,dx = \int_a^b f(x)\,dx, \\ &f_n \text{ differentiable} &&\overset{?}{\Rightarrow} && f \text{ differentiable with } \lim_{n\to\infty}(f_n)'(x) = f'(x) \end{aligned}$$

The goal of the present section is to establish these properties when f is a power series with partial sums f_n. Note that for arbitrary sequences $(f_n)_{n\in\mathbb{N}}$ the properties (33) fail:

Counterexamples:
1. The functions $f_n(x) := x^n \colon [0,1] \to \mathbb{R}$ have as a limit the function f with $f(1) := 1$, and $f(x) := \lim_{n\to\infty} f_n(x) = 0$ otherwise.
All f_n are continuous and differentiable; however, at $x = 1$ the limit f is neither continuous nor differentiable.

2. Let us consider $g_n(x) := \frac{1}{n}\sin(nx) \colon [0, 2\pi] \to \mathbb{R}$. Since $|g_n(x)| \le \frac{1}{n}$ the limit function is $g(x) = 0$. All the g_n are differentiable, and so is g. Still, at $x = 0$ we have

$$\lim_{n\to\infty} g_n'(0) \overset{\text{Chain Rule}}{=} \lim_{n\to\infty} \frac{1}{n} n \cos 0 = 1 \quad \ne \quad g'(0) = 0.$$

3. Let $h_n(x) \colon [0,2] \to \mathbb{R}$ be defined by

$$h_n(x) := \begin{cases} n^2 x & \text{for } x \in [0, \frac{1}{n}], \\ 2n - n^2 x & \text{for } x \in [\frac{1}{n}, \frac{2}{n}], \\ 0 & \text{otherwise.} \end{cases}$$

We claim the limit vanishes, that is, $h(x) := \lim_{n\to\infty} h_n(x) = 0$. Indeed, for $x = 0$ we have $0 = h_n(0) = h(0)$. Consider now an arbitrary $x \in (0,2]$. For all $n \in \mathbb{N}$ with $n > \frac{2}{x} \iff x > \frac{2}{n}$ we have $h_n(x) = 0$. But that means $h_n(x) = 0$ for all $n > \frac{2}{x}$, and so $\lim_{n\to\infty} h_n(x) = 0$, proving the claim. Now h_n and h are integrable but, nevertheless,

$$\lim_{n\to\infty} \int_0^2 h_n(x)\,dx = 1 \quad \ne \quad \int_0^2 h(x)\,dx = 0.$$

`27. Vorlesung, Dienstag, 30.1.07 (Ü 14)` ———————————

5.2. **Uniform convergence.** To study these problems, the following terminology will be useful:

Definition. The *supremum norm [Supremumsnorm]* for $f\colon D \to \mathbb{C}$ bounded is

$$\|f\|_D := \sup\big\{|f(z)| : z \in D\big\} \in [0,\infty);$$

We set $\|f\|_D := \infty$ if f is unbounded.

Examples. $\|\sin x\|_{\mathbb{R}} = 1$, $\|z^2\|_{\mathbb{C}} = \infty$, $\|e^{iz}\|_{\mathbb{S}^1} = 1$.

Here *norm* refers to a mapping $\|.\|$ of a set X into $[0,\infty)$ which has the three properties (i) $\|f\| = 0$ only for $f = 0$, (ii) $\|\lambda f\| = |\lambda|\|f\|$ for all $\lambda \in \mathbb{C}$, and (iii) $\|f+g\| \le \|f\| + \|g\|$; these must hold for all $f, g \in X$. That is, a norm is a generalized modulus. If X is the set of bounded functions on D then $\|f\|_D$ is indeed a norm.

We will resolve our convergence problems by demanding convergence in a stronger sense:

Definition. A sequence of functions $f_n\colon D \to \mathbb{C}$ (for $D \subset \mathbb{C}$) is *uniformly convergent [gleichmäßig konvergent]* to $f\colon D \to \mathbb{C}$, if the sequence of supremum norms $\|f - f_n\|_D$ is null. Equivalently, for each $\varepsilon > 0$ there exists $N = N(\varepsilon) \in \mathbb{N}$ with

$$\big|f(z) - f_n(z)\big| < \varepsilon \qquad \text{for all } n \ge N \text{ and all } z \in D. \tag{34}$$

Unlike for pointwise convergence, in (34) for given ε the index N is independent of z. That is, the speed of convergence is independent of z. Thus the terminology "uniform". Note that uniform convergence implies pointwise convergence.

For the graph of a real function, uniform convergence means that the f_n are contained in a strip of (vertical) width $\pm\varepsilon$ about f, provided $n \ge N$.

Previous examples continued: Only g_n converges uniformly. Indeed:

$$\|f_n - f\|_{[0,1]} = \sup\big\{|x^n| : x \in [0,1]\big\} = 1, \qquad \|g_n - g\|_{[0,2\pi]} = \frac{1}{n} \to 0, \qquad \|h_n - h\|_{[0,2]} = n.$$

We conclude that uniform convergence cannot suffice to allow the exchange of a limit with differentiation.

5.3. **Continuity, integral, and derivative under uniform convergence.** Continuity is preserved under uniform convergence:

Theorem 26 (Weierstrass, 1861)**.** *Let $f_n\colon D \to \mathbb{C}$ be a sequence of continuous functions which converge uniformly to $f\colon D \to \mathbb{C}$. Then f is continuous as well.*

Proof. Let $a \in D$. To verify continuity of f at a we show: For each $\varepsilon > 0$ there is $\delta > 0$ such that

$$\big|f(z) - f(a)\big| < \varepsilon \quad \text{for all } z \in D \text{ with } |z - a| < \delta.$$

Since (f_n) converges uniformly we find $N \in \mathbb{N}$ such that for all $n \geq N$

$$\big|f(\xi) - f_n(\xi)\big| < \frac{\varepsilon}{3} \qquad \text{for all } \xi \in D;$$

we need this for $n = N$. Moreover, since f_N is continuous we can pick $\delta > 0$ with

$$\big|f_N(z) - f_N(a)\big| < \frac{\varepsilon}{3} \quad \text{for all } z \in D \text{ with } |z-a| < \delta.$$

Altogether, we obtain for all $z \in D$ with $|z-a| < \delta$:

$$\big|f(z) - f(a)\big| \leq \big|f(z) - f_N(z)\big| + \big|f_N(z) - f_N(a)\big| + \big|f_N(a) - f(a)\big| < \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon$$

□

Assuming uniform convergence, integration and limit can be exchanged:

Theorem 27. *Suppose the sequence $f_n\colon [a,b] \to \mathbb{C}$, $n \in \mathbb{N}$, of continuous functions converges uniformly to $f\colon [a,b] \to \mathbb{C}$. Then*

$$\lim_{n\to\infty} \int_a^b f_n(x)\,dx = \int_a^b f(x)\,dx.$$

Proof. Consider the real case. The limit function f is continuous by Theorem 26, and therefore integrable. Using the triangle inequality for integrals, we find

$$\left|\int_a^b f(x)\,dx - \int_a^b f_n(x)\,dx\right| \overset{\text{Prop.20}(iv)}{\leq} \int_a^b \big|f(x) - f_n(x)\big|\,dx \leq (b-a)\|f - f_n\|_{[a,b]}.$$

Since $\|f - f_n\|_{[a,b]} \to 0$ this proves the convergence of the sequence of numbers $\int_a^b f_n(x)\,dx$ to $\int_a^b f(x)\,dx$. For f complex apply this argument to real and imaginary part separately. □

For differentiation, a similar statement cannot hold, as is evident from the sequence of functions $g_n(x) = \frac{1}{n}\sin(nx)$ (Example 2). To derive a statement on the exchangability of differentiation with convergence, we will use the previous theorem on the derivative level, and then integrate the result invoking the Fundamental Theorem. This way, we must impose hypotheses stronger than for the previous two theorems; still, we will be able to verify them for power series.

Theorem 28. *Suppose the sequence of functions $f_n\colon B_r(z_0) \to \mathbb{C}$ is continuously differentiable and satisfies the following:*

- *f_n converges pointwise to $f\colon B_r(z_0) \to \mathbb{C}$, and*
- *f_n' converges uniformly.*

Then f is differentiable, and

$$f'(z) = \lim_{n\to\infty} f_n'(z) \qquad \text{for all } z \in B_r(z_0).$$

Our example $g_n = \frac{1}{n}\sin(nx)$ violatang the conclusion, does indeed not satisfy the assumptions of the theorem: The derivative sequence $g_n'(x) = \cos(nx)$ does not converge uniformly.

Proof. For $z \in B_r(z_0)$ choose $h \in \mathbb{C}$ such that $z+h \in B_r(z_0)$. By fundamental theorem and chain rule

$$f_n(z+h) - f_n(z) = h\int_0^1 f_n'(z+th)\,dt.$$

As $n \to \infty$, this equation has the limit

$$f(z+h) - f(z) = h\lim_{n\to\infty}\int_0^1 f_n'(z+th)\,dt. \overset{\text{Thm.27}}{=} \int_0^1 g(z+th)\,dt.$$

We divide by h to obtain

$$f'(z) = \lim_{h\to 0}\frac{f(z+h)-f(z)}{h} = \lim_{h\to 0}\int_0^1 g(z+th)\,dt$$

We claim that the right hand side is $g(z)$, which proves the theorem. To prove the claim, we use the continuity of g: For any $\varepsilon > 0$ there is $\rho > 0$ such that $|g(z+h) - g(z)| < \varepsilon$ for $|h| < \rho$. In conjunction with the standard integral estimate Prop. 20(*iv*) this gives

$$\Big|\int_0^1 g(z+th)\,dt - g(z)\Big| = \Big|\int_0^1 g(z+th) - g(z)\,dt\Big| \le \int_0^1 \big|g(z+th) - g(z)\big|\,dt \le \varepsilon.$$

Thus $\lim_{h\to 0}$ of this expression vanishes indeed. □

5.4. **The radius of convergence of a power series.** Our goal is to show that power series can be integrated and differentiated term by term; that is, we want to apply the results of the previous subsection to partial sums

$$f_n(z) := \sum_{k=0}^n c_k z^k$$

of a power series. But is a power series is uniformly convergent?

Example. For $z \in B_1(0) = \{|z| < 1\} \subset \mathbb{C}$ the standard geometric series $f(z) := \sum z^k$ is convergent, with limit $\sum_{k=0}^\infty z^k = \frac{1}{1-z}$. Unfortunately, the convergence is not uniform on $B_1(0)$. To see this, consider

$$\begin{aligned} d(n) :&= \big\|f(z) - f_n(z)\big\|_{B_1(0)} = \sup\Big\{\Big|\sum_{k=0}^\infty z^k - \sum_{k=0}^{n-1} z^k\Big|,\ z \in B_1(0)\Big\} \\ &= \sup\Big\{|z^n + z^{n+1} + \dots|,\ z \in B_1(0)\Big\}. \end{aligned}$$

To disprove uniform convergence we need to show that $d(n) \not\to 0$. But for any n, the number $z(n) := \sqrt[n]{\frac{1}{2}} \in B_1(0)$ satisfies $z(n)^n = \frac{1}{2}$. Moreover, any higher power of $z(n)$ is positive, and so $d(n) \geq \frac{1}{2}$. Consequently, the standard geometric series does *not* converge uniformly on $B_1(0)$.

`28. Vorlesung, Donnerstag, 1.2.07 (T 14)` ______________________________

We need a test for uniform convergence. Majorization, working uniformly in z, provides such a test:

Lemma 29 (Weierstrass)**.** *Let $P\colon D \to \mathbb{C}$ be a power series, $P(z) = \sum_{n=0}^{\infty} c_n z^n$. Suppose there exists a convergent series $\sum_{n=0}^{\infty} a_n$ of real numbers with*

$$|c_n z^n| \leq a_n \qquad \text{for all } z \in D \subset \mathbb{C}.$$

Then $\sum_{n=0}^{\infty} c_n z^n$ converges uniformly on D, that is,

$$\Big\| \sum_{k=n}^{\infty} c_k z^k \Big\|_D \to 0 \quad \text{as } n \to \infty.$$

Example. We saw above that for the geometric series $\sum_{k=0}^{\infty} z^k$ the assumption cannot hold for $D = B_1(0)$. However, for any $0 \leq \rho < 1$ the terms $a_n := \rho^n$ give a majorant on the ball $D = \overline{B_\rho}(0)$. Consequently, the geometric series converges uniformly on each closed ball $\overline{B_\rho}(0)$ contained in $B_1(0)$.

Proof. According to the majorant test for series, Cor. II.22, $\lim_{n\to\infty} \sum_{k=0}^{n} c_k z^k$ converges absolutely, for each $z \in D$.

To show uniform convergence, let $\varepsilon > 0$. Since $\sum_{k=0}^{\infty} a_k$ converges, there is $N = N(\varepsilon) \in \mathbb{N}$ with

$$\sum_{k=n+1}^{\infty} a_k < \varepsilon \qquad \text{for all } n \geq N.$$

Therefore, invoking the triangle inequality for series, we obtain for all $z \in D$

$$\Big| P(z) - \sum_{k=0}^{n} c_k z^k \Big| = \Big| \sum_{k=n+1}^{\infty} c_k z^k \Big| \overset{\text{II(12)}}{\leq} \sum_{k=n+1}^{\infty} |c_k z^k| \leq \sum_{k=n+1}^{\infty} a_k < \varepsilon \qquad \text{for all } n \geq N.$$

Taking the supremum over D, we find $\big\| P(z) - \sum_{k=0}^{n} c_k z^k \big\|_D \to 0$, that is, the series converges uniformly. ☐

We now want to investigate where a power series converges.

Definition. Let $P(z) = \sum_{n=0}^{\infty} c_n z^n$ be a complex power series. Then the number

$$R := \sup\{r \geq 0 : (c_n r^n)_{n\in\mathbb{N}_0} \text{ is a bounded sequence}\} \in [0, \infty]$$

is called the *radius of convergence [Konvergenzradius]* of P.

Examples. 1. The standard geometric series $\sum_{n=0}^{\infty} z^n = 1$ has radius of convergence $R = 1$.
2. For the exponential series, $R = \infty$. Indeed, for each $z \in \mathbb{C}$, the ratio test proves $\frac{1}{n!} z^n \to 0$.
3. For $c_n := n^n$ we have $R = 0$. Indeed, $(nr)^n \to \infty$ for each $r > 0$.

As indicated by the name, a power series converges on a disk having the radius of convergence R (we set $B_\infty(0) := \mathbb{C}$):

Theorem 30. *Let the power series $P(z) = \sum_{n=0}^{\infty} c_n z^n$ have radius of convergence R. Then*
(i) $P(z)$ diverges for all z with $|z| > R$, and
(ii) For each $0 < r < R$, the series $P(z)$ converges uniformly on $\overline{B_r}(0) := \{z \in \mathbb{C} : |z| \le r\}$.
(iii) $P(z)$ converges (absolutely) for $z \in B_R(0) = \{z \in \mathbb{C} : |z| < R\}$.

Remarks. 1. For z with $|z| = R$ the theorem does not assert anything. In fact, at points on the boundary of the ball, the power series may or may not converge.
2. By (*iii*) the series converges uniformly on each subset $\overline{B_r}(0) \subset B_R(0)$, but it does not necessarily do so on their union, $\bigcup_{0<r<R} \overline{B_r}(0) = B_R(0)$. Similarly, a power series which converges on all of $\mathbb{C}$ converges uniformly on each ball $\overline{B_r}(0) \subset \mathbb{C}$, but not necessarily over all of $\mathbb{C}$. (Show this for exp!)

Proof. (*i*) It is sufficient to show: $P(z)$ convergent $\Rightarrow |z| \le R$. If $P(z)$ is convergent, its terms $c_n z^n$ form a null sequence (Thm. II.15). Thus $c_n z^n$ is bounded and equivalently $c_n |z|^n$ is bounded. We conclude $|z| \le R$.

(*ii*) If $R = 0$, there is nothing to show. So consider $R > 0$. Let us pick ρ with $r < \rho < R$. Then, by definition of R, there is a $C \in \mathbb{R}$ such that

$$|c_n \rho^n| \le C \qquad \text{for all } n \in \mathbb{N}_0.$$

(This would not hold with R in place of ρ – why?) We conclude

$$|c_n z^n| = \left|c_n \rho^n\right| \left|\frac{z^n}{\rho^n}\right| \le C\Big(\frac{r}{\rho}\Big)^n \quad \text{for all } z \in \overline{B_r}(0).$$

Therefore,

$$|c_n z^n| \le C q^n \qquad \text{where } q := \frac{r}{\rho} < 1,$$

i.e., the series $C \sum q^n$ is a convergent majorant on $\overline{B_r}(0)$. Using Lemma 29, this implies uniform convergence of $\sum_{n=0}^{\infty} c_n z^n$ over $\overline{B_r}(0)$. (Would this proof work with $\rho := r$?)

(*iii*) This follows from (*ii*) by choosing $r := |z| < R$. $\square$

The proof underlines the significance of the geometric series: In order to decide where a power series converges, all we need to know is where the geometric series converges!

Let us state some common formulas for the radius of convergence (without proof):

$$R = \frac{1}{\limsup \sqrt[n]{|c_n|}} \text{ (Cauchy 1821, Hadamard 1892)}, \qquad R = \frac{1}{\lim \frac{|c_{n+1}|}{|c_n|}} \text{ (Euler).}$$

These formulas are valid under the convention that $\frac{1}{0} := \infty$ and $\frac{1}{\infty} := 0$. For Euler's formula to hold we assume that the limit exists. For the Cauchy-Hadamard formula we define for a real sequence (a_n) the *limit superior* $\limsup a_n := \lim_{n\to\infty} \sup\{a_n, a_{n+1}, \ldots\}$, that is, it is the largest accumulation point of the set $\{a_n : n \in \mathbb{N}\}$. The proof of these formulas is left as an exercise.

5.5. **Continuity, integration and differentiation of power series.** Now that uniform convergence has been asserted for power series, we can apply our theorems for sequences of functions to power series.

Corollary 31. *Let $P(z) := \sum_{n=0}^{\infty} c_n z^n$ be a power series with radius of convergence R.*
(i) Then $P(z)$ is a continuous function on $B_R(0) \subset \mathbb{C}$.
(ii) P can be integrated termwise over any subinterval $[a, b] \subset (-R, R)$,

$$\int_a^b \sum_{n=0}^{\infty} c_n x^n \, dx = \sum_{n=0}^{\infty} \frac{c_n}{n+1} x^{n+1}\Big|_a^b.$$

Proof. (i) Consider an arbitrary $z \in B_R(0)$ and pick ρ with $|z| < \rho < R$. The polynomials $f_n(z) = \sum_{k=0}^{n} c_k z^k$ are continuous and converge uniformly on $\overline{B_\rho}(0)$ for each $0 < \rho < R$. By Thm. 26 also $P(z) = \lim_{n\to\infty} f_n(z)$ is continuous at $z \in \overline{B_\rho}(0)$.

(ii) Using the rule (26) for integrating polynomials we obtain

$$\int_a^b \sum_{k=0}^{\infty} c_k x^k \, dx \overset{\text{by def.}}{=} \int_a^b \lim_{n\to\infty} \sum_{k=0}^{n} c_k x^k \, dx$$

$$\overset{\text{Thm.27}}{=} \lim_{n\to\infty} \int_a^b \sum_{k=0}^{n} c_k x^k \, dx = \lim_{n\to\infty} \sum_{k=0}^{n} \frac{c_k}{k+1} x^{k+1}\Big|_a^b = \sum_{k=0}^{\infty} \frac{c_k}{k+1} x^{k+1}\Big|_a^b. \qquad \square$$

Remark. Abel's Theorem [Abelscher Grenzwertsatz] says that if $P(z)$ converges at a point z_0 with $|z_0| = R$, then P is continuous on $B_R(0) \cup \{z_0\}$, and the convergence is actually uniform on the segment $\{tz_0 : 0 \le t \le 1\}$. See [F] S.232/33. So the lack of uniform continuity of a power series is tied to its unboundedness.

Examples. 1. (*Integral sine*) The power series (31) for $\frac{\sin z}{z}$ converges pointwise at each $z \in \mathbb{C}$. Therefore the series converges uniformly on each ball $\overline{B_r}(0)$. This justifies (32); and it proves continuity of the right hand side of (31).

2. (*Logarithm-series*, Mercator 1668) We derive the power series for log. The geometric series gives $\frac{1}{1+t} = \sum_{n=0}^{\infty}(-t)^n$, with radius of convergence $R = 1$; in particular, the convergence is uniform on $t \in \big[-|x|,|x|\big]$ for each $|x| < 1$. That gives

$$\begin{aligned} \log(1+x) = \log(1+t)\Big|_0^x &= \int_0^x \frac{1}{1+t}\,dt \overset{\text{geom. series}}{=} \int_0^x \sum_{n=0}^{\infty}(-t)^n\,dt \\ &\overset{\text{Thm.27}}{=} \sum_{n=0}^{\infty}\Big((-1)^n \int_0^x t^n\,dt\Big) = \sum_{n=0}^{\infty}(-1)^n\frac{x^{n+1}}{n+1} \qquad (35) \\ &= x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} \pm \dots \qquad \text{for each } |x| < 1. \end{aligned}$$

`29. Vorlesung, Dienstag, 6.2.07 (Ü 15)`

We have the following:

• For $|x| < 1$ the expansion (35) is valid, and so $R \ge 1$.

• For $x = -1$ the right hand side is the harmonic series $-1 - \frac{1}{2} - \frac{1}{3} - \dots$, which diverges. By Thm. 30 this shows $R \le 1$, and so $R = 1$.

• For $x = 1$ the right hand side is the alternating harmonic series. Thus the series converges (due to the Leibniz test). The left hand side is continuous at $x = 1$, and the right hand side is continuous at $x = 1$ by Abel's Theorem or by a direct argument (see [K], p.118); consequently (35) is valid at $x = 1$, that is, $\log 2 = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} \pm \dots$.

• For $x > 1$ only the left hand side of (35) is defined, the right hand side is divergent!

3. *Arc-tangent-series* and a series representation for π: See problems

Let us now discuss the termwise differentiability of power series.

Lemma 32. *Suppose the series $P(z) := \sum_{n=0}^{\infty} c_n z^n$ has radius of convergence $R = R_P$. Then R is also the radius of convergence of the formally differentiated series $Q(z) := \sum_{n=1}^{\infty} n c_n z^{n-1}$.*

Proof. Let R_Q be the radius of convergence of Q. First we show $R_Q \le R_P$. If $0 < r < R_Q$ then $(nc_n r^{n-1})$ is a bounded sequence, and so is the sequence $(nc_n r^n)$. Hence $(c_n r^n)$ is bounded which implies $r \le R_P$. Taking the limit $r \to R_Q$ gives $R_Q \le R_P$.

Second we show $R_Q \ge R_P$. We claim that for $0 < q < 1$ the series

$$\sum_{n=1}^{\infty} nq^{n-1}$$

converges. Indeed, we can check this using the limit version of the ratio test:

$$\lim_{n\to\infty} \frac{(n+1)q^n}{nq^{n-1}} = \lim_{n\to\infty} q\frac{n+1}{n} = q < 1$$

Consider now $r < R_P$, and pick ρ with $r < \rho < R_P$. By assumption, $c_n\rho^n$ is a bounded sequence. Hence, $\frac{1}{\rho}c_n\rho^n$ is also bounded, say by C, and consequently

$$nc_n r^{n-1} = \frac{1}{\rho}c_n\rho^n\, n\frac{r^{n-1}}{\rho^{n-1}} \leq Cnq^{n-1},$$

where $q := \frac{r}{\rho} < 1$. Hence, taking the claim into account, we have determined a convergent majorant of the series Q. Thus $Q(z)$ converges for all $r < R_P$ and so $R_Q \geq R_P$. $\square$

Using the lemma, we can apply Theorem 28 to power series:

Theorem 33 (Termwise differentiation of power series)**.** *Suppose $P(z) := \sum_{n=0}^{\infty} c_n z^n$ has radius of convergence R. Then*

$$P'(z) = \Big(\sum_{n=0}^{\infty} c_n z^n\Big)' = \sum_{n=1}^{\infty} nc_n z^{n-1} \qquad \textit{for } z \in B_R(0),$$

and the series for P' also has radius of convergence R.

Proof. By the lemma, the derivative series $Q(z) = \sum nc_n z^{n-1}$ has radius of convergence R as well. By Theorem 30, converges uniformly on any closed ball $\overline{B_r} \subset B_R$. Hence the assumptions for Thm. 28 hold for $f_n(z) = \sum_{k=0}^{n} c_k z^k$ over $\overline{B_r}$, and so

$$P'(z) = \Big(\lim_{n\to\infty}\sum_{k=0}^{n} c_k z^k\Big)' \stackrel{\text{Thm. 28}}{=} \lim_{n\to\infty}\Big(\sum_{k=0}^{n} c_k z^k\Big)' = \lim_{n\to\infty}\sum_{k=1}^{n} kc_k z^{k-1}. \qquad \square$$

We would like to mention an important result for power series. If power series have the same values, the coefficients agree:

Theorem 34 (Identity theorem *[Identitätssatz]*)**.** *Let $f(z) = \sum_{n=0}^{\infty} b_n z^n$ and $g(z) = \sum_{n=0}^{\infty} c_n z^n$ be power series. Suppose they converge on a sequence $z_n \to 0$ and satisfy $f(z_n) = g(z_n)$. Then $b_n = c_n$ for all $n \in \mathbb{N}_0$.*

6. Taylor's formula

Can every function be represented by a power series? If so, we can differentiate and integrate the series termwise. In particular, we can plug a power series into a differential equation and obtain a solution by determining the coefficients of the series.

Let us first illustrate on the example of a polynomial how a power series can be obtained. We can recover the coefficients of $f(x) = a_0 + a_1x + \ldots + a_nx^n$ as derivatives at 0:

$$f(0) = a_0, \quad f'(0) = a_1, \quad f''(0) = 2a_2, \quad \ldots, \quad f^{(n)}(0) = n!\, a_n,$$

the higher derivatives vanish. Thus we can write the polynomial in terms of its derivatives at 0:

$$f(h) = f(0) + f'(0)h + \frac{f''(0)}{2}h^2 + \ldots + \frac{f^{(n)}(0)}{n!}h^n \tag{36}$$

This formula is also useful to rewrite a polynomial of the form $\sum b_k(x-x_0)^k$ as $\sum a_k x^k$.

For f a general function, the sequence of nonzero derivatives is usually infinite. Hence we obtain a representation for f which is no longer a polynomial, but a power series.

Definition. Let I be an interval, and $f\colon I \to \mathbb{R}$ be a function which has arbitrary many derivatives at $x \in I$. Then for $n \in \mathbb{N}_0$ we define the n-th *Taylor polynomial* by

$$T_x^n f(h) := f(x) + f'(x)h + \ldots + \frac{f^{(n)}(x)}{n!}h^n.$$

The *Taylor series [Taylor-Reihe]* is the limiting power series

$$T_x f(h) := \lim_{n\to\infty} T_x^n f(h) = \sum_{n=0}^{\infty} \frac{f^{(n)}(x)}{n!}h^n.$$

In particular, $T_x^1 f(h) = f(x) + f'(x)h$ coincides with the linear approximation to f at x. Note that our notation is in terms of the base point x and an increment h. If instead f is evaluated at $\xi := x + h$ the Taylor polynomial has to be evaluated at $h = \xi - x$ (compare with books!).

We saw in (36) that $T_x f$ reproduces a polynomial f in the sense

$$f(x+h) = T_x f(h) \qquad \text{for all } x, x+h \text{ in the domain of } f. \tag{37}$$

Similarly, it is clear that $T_x f$ reproduces a power series such as exp or sin. Thus we expect (37) to hold in general. Unfortunately, this need not be true:

Counterexamples: 1. Let $f(x) := \exp(-\frac{1}{x^2})$ for $x \neq 0$ and $f(0) := 0$. Then f is infinitely often differentiable with $f^{(n)}(0) = 0$ (see problems). So its Taylor series is $T_0 f \equiv 0$. But f is nonzero (except at $x = 0$) and so the identity $f(h) = T_0 f(h)$ only holds at $h = 0$.

2. The logarithm has the series (35),

$$\log(1+h) = h - \frac{h^2}{2} + \frac{h^3}{3} \pm \ldots.$$

It is not hard to check by induction that the right hand side represents $(T_1 \log)(h)$ (this also follows from Thm. 34). The left hand side is defined all $h > -1$; nevertheless the logarithm series on the right diverges for $h > 1$! Therefore, the identity $\log(1+h) = (T_1 \log)(h)$ only holds for $1 + h \in (0, 2]$, that is, on a proper subset of the domain $(0, \infty]$ of f!

To deal with these problems, we will follow the standard approach of analysis: Rewrite the problem as an error estimate. So suppose $f\colon I \to \mathbb{R}$ has at least $n-1$ derivatives at $x \in I$. Then we define a *remainder term [Restglied]*

$$R_n\colon \{h : x+h \in I\} \to \mathbb{R}, \qquad R_n(h) := f(x+h) - T_x^{n-1} f(h), \quad \text{for } n \in \mathbb{N}.$$

Checking (37) then becomes equivalent to showing $R_n(h) \to 0$ as $n \to \infty$.

`30. Vorlesung, Donnerstag, 8.2.07 (T 15)`

Lemma 35 (Taylor formula with integral remainder term)**.** *Suppose that for $n \in \mathbb{N}$ the function $f\colon I \to \mathbb{R}$ is n-times continuously differentiable, and let $x, x+h \in I$. Then the n-th remainder term has the representation*

$$R_n(h) = \frac{1}{(n-1)!} \int_x^{x+h} f^{(n)}(t)\,(x+h-t)^{n-1}\,dt. \tag{38}$$

Check that the integral is defined!

Proof. By induction: The Fundamental Theorem $f(x+h) - f(x) = \int_x^{x+h} f'(t)\,dt$, together with $T_x^0(h) = f(x)$, yields the base case $n = 1$.

Step $n \to n+1$: With R_n given as in (38), we need to show

$$R_{n+1}(h) = f(x+h) - T_x^n f(h) = f(x+h) - T_x^{n-1} f(h) - \frac{1}{n!} f^{(n)} h^n = R_n(h) - \frac{1}{n!} f^{(n)} h^n.$$

But indeed, we can derive this from (38) for n:

$$\begin{aligned}
R_n(h) &\overset{(38)}{=} -\int_x^{x+h} f^{(n)}(t) \frac{d}{dt} \frac{(x+h-t)^n}{n!}\,dt \\
&\overset{\text{int. by parts}}{=} -f^{(n)}(t) \frac{(x+h-t)^n}{n!}\bigg|_{t=x}^{t=x+h} + \frac{1}{n!} \int_x^{x+h} \left(\frac{d}{dt} f^{(n)}(t)\right) (x+h-t)^n\,dt \\
&= \frac{f^{(n)}(x)}{n!} h^n + \underbrace{\frac{1}{n!} \int_x^{x+h} f^{(n+1)}(t)\,(x+h-t)^n\,dt}_{=R_{n+1}(h)}
\end{aligned}$$

□

Remark. The proof tells us that the n-th Taylor polynomial $T^n f$ is precisely the result of applying the first the fundamental theorem to f, followed by n integration by parts.

Often the following form of the remainder term is more useful:

Theorem 36 (Taylor formula with remainder term in Lagrangian form)**.** *Under the assumptions of the lemma, there exists ξ between $x \in I$ and $x+h \in I$ with*

$$R_n(h) = \frac{f^{(n)}(\xi)}{n!} h^n.$$

Proof. We apply the generalized Mean Value Theorem of Integration: There exists ξ between x and $x+h$ with

$$R_n(h) \overset{(38)}{=} \int_x^{x+h} f^n(t)\frac{(x+h-t)^{n-1}}{(n-1)!}\,dt \overset{\text{Thm. } 21(ii)}{=} f^{(n)}(\xi)\int_x^{x+h} \frac{(x+h-t)^{n-1}}{(n-1)!}\,dt$$
$$= -f^{(n)}(\xi)\frac{(x+h-t)^n}{n!}\Big|_{t=x}^{t=x+h} = \frac{f^{(n)}(\xi)}{n!}h^n.$$
□

Examples. 1. (*Convexity [Konvexität]*) A twice continuously differentiable function $f\colon I \to \mathbb{R}$, defined on an interval I, is called *convex* if $f'' \geq 0$. Then

$$f(x+h) \geq f(x) + f'(x)h \qquad \text{for all } h \text{ with } x+h \in I,$$

that is, f lies above its tangent.
Proof: For some $\xi \in (x, x+h)$ we have

$$f(x+h) - \big(f(x) + f'(x)h\big) = R_2(h) \overset{\text{Thm. } 36}{=} \frac{1}{2}f''(\xi)h^2 \geq 0.$$

2. (*Binomial series*, Newton 1665) For $a \in \mathbb{C}$ consider the a-th power $f\colon (0,\infty) \to \mathbb{R}$, $f(x) = x^a$. To calculate its derivatives let us first define *generalized binomial coefficients*,

$$\binom{a}{n} := \frac{a(a-1)\cdots(a-n+1)}{1\cdot 2\cdot \ldots \cdot n}, \qquad \text{for } n \in \mathbb{N},\ a \in \mathbb{C},$$

and $\binom{a}{0} := 1$. For instance, $\binom{1/2}{1} = \frac{1}{2}$ and $\binom{1/2}{2} = \frac{1/2\cdot(-1/2)}{2} = -\frac{1}{8}$. Then, by (16),

$$\frac{1}{n!}f^{(n)}(x) = \frac{1}{n!}a(a-1)\cdots(a-n+1)x^{a-n} = \binom{a}{n}x^{a-n}$$

and, in particular, $\frac{1}{n!}f^{(n)}(1) = \binom{a}{n}$. So the Taylor series taken at $x=1$ is

$$T_1 f(h) = \sum_{n=0}^{\infty}\binom{a}{n}h^n.$$

By the ratio test in the limit version, the series converges for $|h| < 1$. Indeed,

$$\left|\frac{\binom{a}{n+1}h^{n+1}}{\binom{a}{n}h^n}\right| = \left|\frac{a-n}{n+1}\right||h| \to |h| \qquad \text{as } n \to \infty,$$

which means the limit is less than 1.

Now for each $|h| < 1$ we need to show that $R_n(h) \to 0$ as $n \to \infty$. We skip the proof and refer to [F], p.236. This gives $T_1 f(h) = f(1+h)$ for these h, and so

$$(1+h)^a = \sum_{n=0}^{\infty}\binom{a}{n}h^n \qquad \text{for } |h| < 1,\ a \in \mathbb{C}.$$

For the particular case $a = \frac{1}{2}$ we find

$$\sqrt{1+h} = 1 + \frac{1}{2}h - \frac{1}{8}h^2 + \frac{1}{16}h^3 - \frac{5}{128}h^4 \pm \dots .$$

When $|h| < 1$ we can use Thm. 36 to estimate the error. Even simpler is an error estimate for $0 < h < 1$: Then the series alternates, and so the Leibniz test works. For instance, when approximating the binomial series linearly, we obtain the remainder term estimate

$$\Big|\sqrt{1+h} - \big(1 + \frac{h}{2}\big)\Big| = |R_2(h)| \le \frac{1}{8}h^2.$$

Remark. We have not given a satisfactory answer to the question where a Taylor series coincides with a function. The tools of complex analysis will give the complete answer: Let a function $f\colon D \to \mathbb{C}$ be given which is (complex) differentiable on some domain $D \subset \mathbb{C}$. Let moreover $z \in D$. Then over each ball $B_r(z)$ which is entirely contained in D, the Taylor series $T_z f(h)$ converges and coincides with $f(z+h)$. As an example, take $f(z) = \frac{1}{1+z^2}$. This function is defined on $\mathbb{C} \setminus \{\pm i\}$. Hence the Taylor series of f at $z = 0$ has radius of convergence 1 and coincides with f on $B_1(0)$. Note that the behaviour of the real function becomes only transparanent after transition to the complex picture.

Summary

We first introduced differentiation. We realized that the existence of the limit of the difference quotient is equivalent to the existence of a good linear approximation (i.e., with an error of from $o(h)$). We derived rules of differentiation which let us calculate the derivatives for all explicitly given functions. Then we discussed extremals of real valued functions. We saw that at an extremal x the derivative has a zero, $f'(x) = 0$. (by a monotonicity argument). The converse holds, for instance, if f' changes sign at x, in particular if $f''(x) \neq 0$. Each student should memorize the derivatives of x^n or x^a, exp, sin, cos, log, and perhaps of tan, arctan.

Second we introduced integration. A function is Riemann integrable if it admits a two-sided approximation by step functions, with the same limiting integrals. Note that the two-sided approximation makes use of the order of $\mathbb{R}$! Using equidistant partitions, we found that monotone functions are integrable; we introduced the concept of uniform continuity to see why continuous funcitons are integrable. The most essential fact about integration is the fundamental theorem: It allows to compute integrals by guessing primitives, and it allows to derive the (somewhat unsatisfactory) rules of integration from the rules of differentiation. Integrability is preserved under uniform convergence; this result depends on the trivial estimate $\int_a^b f \le (b-a)\|f\|_{[a,b]}$. It is essential to know the primitives of functions such as x^n, x^a, $1/x$, exp, cos, sin, and perhaps of $1/\cos^2$ and $\frac{1}{1+x^2}$. It is also worth knowing the basic improper integrals: $\int_0^1 \frac{1}{x^s}dx$ exists for $0 < s < 1$ but is divergent for $s \geq 1$, while $\int_1^\infty \frac{1}{x^s}dx$ exists for $s > 1$ but is divergent for $0 < s \leq 1$.

Finally, we studied power series $P(z) = \sum c_n z^n$. For the number $R := \sup\{r > 0 : (c_n r^n) \text{is bounded}\}$, we proved that P converges in $B_R(0)$, and diverges on $\mathbb{C} \setminus \overline{B_R}(0)$. It converges uniformly on $B_r(0)$ for any $r < R$. These results were achieved by majorization and minorization with the geometric series. Since continuity is preserved under uniform convergence of a sequence of functions (the proof employs an $\varepsilon/3$'s argument), power series are continuous on B_R. They can be integrated and differentiated termwise. The latter result uses the stability of the integral under uniform convergence in an application on the derivative level. Finally, given a function we exhibit a power series, called Taylor series, which represents the function in many cases, but unfortunately not always.

Part 5. Sequences and continuous functions in multidimensional space

31. Lecture, Tuesday, 17. April 07

We wish to generalize the concepts familiar to us in one variable to the case of several variables, that is, from $\mathbb{R}$ to $\mathbb{R}^n$. We will consider sequences in $\mathbb{R}^n$ and continuous functions between $\mathbb{R}^n$ and $\mathbb{R}^m$.

The generalization of the Bolzano-Weierstrass theorem and the theorem of the maximum will be the key topics of the present part. A closer inspection of these assertions for $\mathbb{R}^n$ leads to a new notion, namely compactness. It also explains at a deeper level why the two theorems hold for one variable.

Much of the present section is formulated in greater generality than just for the space $\mathbb{R}^n$. We will go to normed vector spaces and metric spaces. These more abstract spaces have important applications in mathematics which we cannot appropriately represent here. Besides, by defining only as much structure as is needed to carry our arguments out makes these arguments more transparent.

1. The Euclidean vector space $\mathbb{R}^n$ and its generalizations

1.1. **Euclidean scalar product and norm on $\mathbb{R}^n$.** We consider the *n-dimensional space* $\mathbb{R}^n = \mathbb{R} \times \ldots \times \mathbb{R}$ with n factors (state the recursive definition!). Its elements are the *points* $x = (x_1, \ldots, x_n)^\top$; we call the x_i *components*. We ususally regard vectors as columns, and so invoke the transposed sign $\cdot^\top$ when writing rows. Defining addition componentwise makes $(\mathbb{R}^n, +)$ an additive Abelian group. Thus it is justified to write 0 for the neutral element $(0, \ldots, 0)^\top$. Moreover, multiplication with a scalar $\lambda \in \mathbb{R}$ can be defined, $\lambda x := (\lambda x_1, \ldots, \lambda x_n)^\top$. These two operations endow $\mathbb{R}^n$ with the structure of a *vector space.* Accordingly, we regard $x \in \mathbb{R}^n$ a vector as well.

On $\mathbb{R}^n$, we will usually work with the *Euclidean scalar product*

$$(1) \qquad \langle .,. \rangle \colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}, \qquad \langle x, y \rangle := x_1 y_1 + \ldots + x_n y_n = \sum_{k=1}^n x_k y_k,$$

also denoted as a *dot product* $x \cdot y$. It is symmetric, $\langle x, y \rangle = \langle y, x \rangle$, and it is *linear*:

$$(2) \qquad \langle \lambda x + y, z \rangle = \lambda \langle x, z \rangle + \langle y, z \rangle \qquad \text{for all } \lambda \in \mathbb{R};$$

consequently it is also linear in the second entry. Two vectors x, y are *orthogonal* when $\langle x, y \rangle = 0$; in particular, the vector $x = 0$ is orthogonal to any other vector.

When $n \geq 2$, there is no ordering on $\mathbb{R}^n$ compatible with addition. As for $\mathbb{R}^2 = \mathbb{C}$, we define a *length* or *Euclidean norm* $\|.\| \colon \mathbb{R}^n \to \mathbb{R}$ by

$$\|x\| := \sqrt{\langle x, x\rangle} = \sqrt{x_1^2 + \ldots + x_n^2}. \tag{3}$$

We will verify below that $\|.\|$ is a norm. Using linearity of the scalar product alone, we can derive the general *Pythagorean theorem*

$$\|x \pm y\|^2 = \big\langle x \pm y, x \pm y\big\rangle = \|x\|^2 \pm 2\langle x, y\rangle + \|y\|^2. \tag{4}$$

A geometric interpretation of the scalar product is $\langle x, y\rangle = \cos\alpha \,\|x\|\,\|y\|$, where α is the angle enclosed by x and y, measured in the plane $\operatorname{span}\{x, y\}$. Let us make a somewhat weaker assertion:

Lemma 1 (Schwarz inequality *[Cauchy-Schwarz-Ungleichung]*). *For all $x, y \in \mathbb{R}^n$ we have*

$$\big|\langle x, y\rangle\big| \leq \|x\|\,\|y\|. \tag{5}$$

Moreover, equality holds precisely when x and y are linearly dependent.

Proof. The inequality is correct if $x = 0$ or $y = 0$. Else, let us consider the unit vectors $\frac{x}{\|x\|}$ and $\frac{y}{\|y\|}$. Replacing x and y in (4) by these vectors, we find

$$0 \leq 1 \pm 2\Big\langle \frac{x}{\|x\|}, \frac{y}{\|y\|}\Big\rangle + 1 \quad\Leftrightarrow\quad \pm\Big\langle \frac{x}{\|x\|}, \frac{y}{\|y\|}\Big\rangle \leq 1 \quad\Leftrightarrow\quad \pm\langle x, y\rangle \leq \|x\|\,\|y\|,$$

which agrees with (5).

In particular, inequality (5) is strict unless $\|\frac{x}{\|x\|} \pm \frac{y}{\|y\|}\| = 0$ holds for one sign $\pm$, that is, unless x and y are linearly dependent. $\square$

The Schwarz inequality is the key to proving that $\|.\|$ is a norm:

Theorem 2. $\|.\|$ *is a* norm *on $\mathbb{R}^n$, that is, $\|.\| \colon \mathbb{R}^n \to [0, \infty)$ satisfies:*
(i) positivity: $\|x\| = 0 \Leftrightarrow x = 0$,
(ii) homogeneity: $\|\lambda x\| = |\lambda|\|x\|$ *for all $x \in \mathbb{R}^n$, $\lambda \in \mathbb{R}$,*
(iii) triangle inequality *or* subadditivity: $\|x + y\| \leq \|x\| + \|y\|$ *for $x, y \in \mathbb{R}^n$.*

Proof. (i) and (ii) are immediate from (3). The triangle inequality (iii) follows from the Schwarz inequality: Indeed,

$$\|x + y\|^2 \overset{(4)}{=} \|x\|^2 + 2\langle x, y\rangle + \|y\|^2 \overset{(5)}{\leq} \|x\|^2 + 2\|x\|\|y\| + \|y\|^2 = \big(\|x\| + \|y\|\big)^2.$$

$\square$

As for the Schwarz inequality, if x and y are linearly independent then the triangle inequality is strict.

Let us derive from (*iii*) one more property of the norm, the *inverse triangle inequality [verschärfte Dreiecksungleichung]*

$$\big|\|x\| - \|y\|\big| \leq \|x \pm y\|. \tag{6}$$

To prove it, note that $\|x+y-y\| \leq \|x+y\| + \|y\|$ gives $\|x\| - \|y\| \leq \|x+y\|$. Exchanging the letters x and y yields $\|y\| - \|x\| \leq \|x+y\|$. Taken together, this is the inverse triangle inequality with "+". Exchanging y by $-y$ yields the version with the other sign "−".

1.2. **Lengths and normed vector spaces.** In the previous subsection we used only the linearity of the scalar product (2), –not the particular representation (1) of the Euclidean scalar product– to derive for $\|.\| := \sqrt{\langle .,. \rangle}$ the Schwarz inequality $\langle x, y\rangle \leq \|x\|\|y\|$ and further the norm properties of $\|.\|$. Consequently, these properties hold for any vector space with a scalar product.

Many important norms, however, are not induced by a scalar product. Indeed, on $\mathbb{R}^n$, the following are also norms:

$$\|(x_1, \ldots, x_n)^\top\|_p := \sqrt[p]{|x_1|^p + \ldots + |x_n|^p} \qquad \text{for } p \in [1, \infty),$$
$$\|(x_1, \ldots, x_n)^\top\|_\infty := \max\{|x_1|, \ldots, |x_n|\}$$

The first is called the *p-norm*, the second *maximum norm*. Note that the 2-norm agrees with the Euclidean norm. It is interesting to compare the unit balls $\{x \in \mathbb{R}^n : \|x\|_p < 1\}$.

Remark. While the properties (*i*) and (*ii*) are obvious for above norms (check!), the triangle inequality is harder: The fact

$$\|x+y\|_p \leq \|x\|_p + \|y\|_p \qquad \text{for all } p \in [1, \infty] \quad \text{and } x, y \in \mathbb{R}^n$$

is called *Minkowski's inequality* (it fails for $p < 1$). For the cases $p = 1$ and $p = \infty$, Minkowski's inequality is straightforward, for the remaining cases it follows from *Hölder's inequality*

$$\langle x, y\rangle \leq \|x\|_p\|y\|_q, \qquad \text{for all } x, y \in \mathbb{R}^n \quad \text{and } p, q \in (1, \infty) \text{ with } \frac{1}{p} + \frac{1}{q} = 1.$$

(Proof: problems?)

Two norms $\|.\|_1$ and $\|.\|_2$ on a vector space V are called *equivalent* if there are constants $c, C > 0$ such that

$$c\|x\|_1 \leq \|x\|_2 \leq C\|x\|_1 \qquad \text{for all } x \in V.$$

Check that this relation is an equivalence relation.

Proposition 3. *Euclidean and maximum norm are equivalent on* $\mathbb{R}^n$,

$$\|x\|_\infty \le \|x\| \le \sqrt{n}\|x\|_\infty \qquad \text{for all } x \in \mathbb{R}^n. \tag{7}$$

Proof. For any $1 \le k \le n$ we have

$$x_k^2 \le x_1^2 + \ldots + x_n^2 \le n \max_{1\le i\le n} x_i^2 \le n\big(\max_{1\le i\le n} |x_i|\big)^2 \qquad \text{for all } x \in \mathbb{R}^n.$$

Taking the root gives

$$|x_k| \le \|x\| \le \sqrt{n}\|x\|_\infty \qquad \text{for each } 1 \le k \le n.$$

But there is an index k, for which the left hand side agrees with $\|x\|_\infty$ so that (7) is established. $\square$

From (7) we conclude that $\{\|x\|_\infty < 1\} \supset \{\|x\|_2 < 1\} \supset \{\|x\|_\infty < \frac{1}{\sqrt{n}}\}$. That is, the cube with edgelength 2 contains the unit ball, which contains in turn the cube of edgelength $\frac{2}{\sqrt{n}}$.

Later we will see that any two norms on $\mathbb{R}^n$ are equivalent (see Thm. 20).

`32. Lecture, Thursday, 19. April 07` ________________________________ `T 1`

1.3. **Distances and metric spaces.** Our goal is the definition of convergence and continuity for $\mathbb{R}^n$. For the one-dimensional case of $\mathbb{R}$ (or $\mathbb{C}$), these notions only involve the distance $|x-y|$ of two points x, y. In several dimensions, the distance *[Abstand]* on $\mathbb{R}^n$ is the function

$$d\colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}, \qquad d(x,y) := \|x-y\| = \sqrt{\sum_{k=1}^n (x_k - y_k)^2}.$$

More generally, $d(x,y) := \|x-y\|$ defines or *induces* a distance on any normed vector space. But it is worth generalizing one step further. We will single out the key properties of the distance function. These are precisely the properties we have used in dimension one, and we will use them in several dimensions. Moreover, we want to attribute a name to spaces which admit a distance function, no matter if induced by a norm or not:

Definition. A *metric space [metrischer Raum]* (X,d) is a set X together with a mapping $d\colon X \times X \to \mathbb{R}$, called *metric [Metrik]*, satisfying the following properties (for all $x, y, z \in X$):
(i) (positivity) $d(x,y) \ge 0$ and $d(x,y) = 0$ if and only if $x = y$,
(ii) (symmetry) $d(x,y) = d(y,x)$,
(iii) (triangle inequality) $d(x,z) \le d(x,y) + d(y,z)$.

If d is induced by a norm $\|.\|$, then (i)-(iii) hold. For instance, the triangle inequality for d follows from the one for $\|.\|$ (check!).

Examples. 1. (Manhattan metric) Let X be a standard ruled grid (orthogonal streets). Take for $d(x, y)$ the minimal length of a curve, running on the grid which joins x with y. This is the distance which matters to the New York taxi driver.

2. (Surfaces) Let X be the surface of your coffee mug (or any other surface) and take for $d(x, y)$ the minimum of the length of curves running within X from x to y (to be defined rigorously only later).

3. (Function spaces) Let $X = \{f : [a, b] \to \mathbb{R}^n : f_i \text{ continuous for } i = 1, \dots, n\}$, also denoted with $C^0([a, b], \mathbb{R}^n)$. Then $d(f, g) := \sup\big\{|f(x) - g(x)| : x \in [a, b]\big\}$ is a metric. What fails (i) when we replace $[a, b]$ by a general space or (ii) consider general, not necessary continuous, functions?

4. (Discrete metric) Any set X with $d(x, y) = 0$ for $x = y$ and $d(x, y) = 1$ for $x \neq y$.

5. Let $f : [0, \infty) \to [0, \infty)$ be a function with (i) $f(0) = 0$, (ii) $f(t) > 0$ for $t > 0$, and (iii) (concavity) $f\big(\sigma s + (1 - \sigma)t\big) \geq \sigma f(s) + (1 - \sigma)f(t)$ for $t > 0$ and $0 \leq \sigma \leq 1$. Show that $(\mathbb{R}, d)$ with $d(x, y) := f(|x - y|)$ is a metric space.

6. Let (X, d) be a metric space and $A \subset X$ be a subset. Then (A, d) again is a metric space, In particular, any subset of $\mathbb{R}^n$ is a metric space by itself.

As in the complex plane, the metric allows us to define open balls,

$$B_r(a) := \{x \in X : d(x, a) < r\}, \qquad a \in X, r > 0. \tag{8}$$

More precisely, $B_r(a)$ is the *open distance ball about a of radius r*. Likewise, *closed balls* are given by $\overline{B}_r(a) = \{x \in X : d(x, a) \leq r\}$. Note that these definitions agree with those for $\mathbb{R}$ or $\mathbb{C}$ endowed with the euclidean distance $d(x, y) = |x - y|$.

Example. For $X = C^0([a, b], \mathbb{R}^n)$, the sin function lies in the 2-ball of the 0-function, but not in the 1-ball.

Boundedness of subsets in X can be defined in terms of the distance:

Definition. A set $Y \subset \mathbb{R}^n$ or $\subset X$ is *bounded* if there is $r > 0$, such that $Y \subset B_r(0)$.

Boundedness may not mean a lot in arbitrary metric spaces: For instance, all subsets of a set X with the discrete metric (Example 4) are bounded. Thus it makes sense to use boundedness only for the case $(\mathbb{R}^n, \|.\|)$.

1.4. **Open sets.** In the one variable case, suppose a function is defined on an open interval of $\mathbb{R}$. Then an extremum can be detected as a zeros of the derivative. The same fails in the case of the domain a closed interval. To arrive at similar statements for functions defined on $\mathbb{R}^n$ let us generalize open intervals:

Definition. (Hausdorff 1914) Let X be $\mathbb{R}^n$ or any other metric space. A subset $U \subset X$ is *open [offen]* if for each $x \in U$ there is an open ball $B_r(x)$ with $r > 0$, such that $B_r(x) \subset U$.

Examples. 1. Any ball $B_R(a)$ is open. To see this, pick $x \in B_R(a)$ and set $r := d(x, a) < R$. We claim $B_{R-r}(x) \subset B_R(a)$, that is, if $y \in B_{R-r}(x)$ then $y \in B_R(a)$. Indeed:

$$y \in B_{R-r}(x) \quad \Rightarrow \quad d(y, x) < R - r \quad \Rightarrow \quad d(y, a) \leq d(y, x) + d(x, a) < (R - r) + r = R.$$

2. $(0, 1] \subset \mathbb{R}$ is not open: For the point 1, there is no ball $B_r(1) \subset (0, 1]$.

The following notion is useful.

Definition. Let $a \in X$. A set $Y \subset X$ is called a *neighbourhood* of a if (i) $a \in Y$ and (ii) there exists some radius $r > 0$ such that $B_r(a) \subset Y$.

Examples. The interval $(0, 1] \subset \mathbb{R}$ is a neighbourhood of each point $0 < y < 1$, but not of $y = 1$.

An open set is characterized as a set which is a neighbourhood for all the points it contains. We will later see that this makes open sets useful as domains for differentiable functions: If $f : U \to X$ and U is open then each point in U can be approached by a sequence from any direction.

We have the following properties of open sets, valid for $\mathbb{R}^n$ or each other metric space:

Theorem 4. *For any metric space (X, d) the following holds:*
(i) $\emptyset$ and X are open.
(ii) Consider any (index) set I, and a family of open sets $\{U_i : i \in I\}$. Then their union $U := \bigcup_{i \in I} U_i$ is also open.

Proof. (i) For the empty set, there is nothing to be shown, while for $X \ni x$ the ball $B_1(x) \subset X$ satisfies the definition.

(ii) Pick an arbitrary $x \in U = \bigcup_{i \in I} U_i$. Then there is an index $i \in I$ such that $x \in U_i$. Since U_i is open there is a ball $B_r(x) \subset U_i$. But then also $B_r(x) \subset \bigcup_{i \in I} U_i$, which shows that U is open. $\square$

Example. Count the rational numbers as $\mathbb{Q} = \{q_k : k \in \mathbb{N}\}$. Then for any $\varepsilon > 0$ the set

$$U := \bigcup_{k\in\mathbb{N}} \left(q_k - \frac{\varepsilon}{2^k},\ q_k + \frac{\varepsilon}{2^k} \right) \tag{9}$$

is a union of open intervals. Hence it is an open set containing $\mathbb{Q}$. Since the intervals need not be disjoint the total length of U satisfies an inequality, namely

$$L(U) \le \sum_{k\in\mathbb{N}} \frac{2\varepsilon}{2^k} = 2\varepsilon.$$

This is an indication that the set of rational numbers is very small: It has zero measure. That is, the propability of a real number to be rational is zero.

Problem. Prove that any finite intersection of open sets $U_1 \cap \ldots \cap U_k$ is open. Show also by counterexample that the same need not be true for infinite intersections.

For those who like some further abstraction, let us quickly introduce the most general concept of space where analysis can work. This is done by converting Thm. 4 into a definition: A *topological space* is a set X together with an assignment of subsets $\mathcal{U} = \{U_k \subset X : k \in I\}$, called the open sets. For $\mathcal{U}$ the properties (i) and (ii) of the theorem are required, that is, it contains X and $\emptyset$ and has the property that an arbitrary system of open sets is still open, $\bigcup_{j\in J} U_j \in \mathcal{U}$. A physical model of a topological space is a body made from perfect rubber: Distances make no sense any longer as we can tear or squeeze the body, but still we know which sets form a neighbourhood of points. This suffices to define convergence or continuity.

`33. Lecture, Tuesday, 24. April 07` ———————————— `Ü 1`

2. Sequences, closed and compact sets

2.1. **Convergence and completeness.** We can define convergence in terms of the distance of a metric space:

Definition. Let $(x_k)_{k\in\mathbb{N}}$ be a sequence of points x_k in a metric space X (perhaps $X = \mathbb{R}^n$).
(i) It is *convergent* if there is a *limit* $a \in X$ with

$$d(x_k, a) \to 0 \qquad \text{as } k \to \infty;$$

we write $x_k \to a$ or $\lim_{k\to\infty} x_k = a$.
(ii) It is a *Cauchy* sequence if

$$d(x_k, x_m) \to 0 \qquad \text{as } k, m \to \infty.$$

(iii) It is *divergent* if it is not convergent.

Example. Consider $(\mathbb{R}^n, d)$, where d is the discrete metric with $d(x,x) = 1$ and $d(x,y) = 1$ for $x \neq y$. Then the sequence $(1/n)$ diverges. If a sequence (x_n) converges to $a \in \mathbb{R}$, then there is $N \in \mathbb{N}$ such that $x_n = a$ for all $n \geq N$.

Inserting the definition of real convergence, we see that a sequence $x_k \in X$ converges if for each $\varepsilon > 0$ there is $N \in \mathbb{N}$ such that $x_k \in B_\varepsilon(a)$ for all $k \geq N$.

In an arbitrary metric space, a Cauchy sequence need not converge (e.g., $\mathbb{Q}$, $\mathbb{Q}^n$). Thus let us define:

Definition. A metric space (X, d) is *complete* if all Cauchy sequences converge.

2.2. **Sequences in $\mathbb{R}^n$.** We call sequences in $\mathbb{R}^n$ *vector sequences* and always use the distance function $d(x,y) = \|x - y\|$. The theory of vector sequences reduces to the case of scalar sequences thanks to the following assertion:

Theorem 5. *(i) A sequence (x_k) of points in $\mathbb{R}^n$ converges to $a \in \mathbb{R}^n$ if and only if for each $i = 1, \ldots, n$ the i-th component sequence $(x_{ki})_{k\in\mathbb{N}}$ converges as a real sequence,*

$$x_{ki} \to a_i \qquad \text{as} \quad k \to \infty.$$

(ii) Similarly, (x_k) is Cauchy if and only if all component sequences are Cauchy.

We have verified this before for the case $n = 2$ of complex sequences $z_k = x_k + iy_k$: They converge if and only if real and imaginary part converge; similarly for Cauchy.

Proof. (i) Using the norm equivalence (7) we find that for each k

$$\|x_k - a\|_\infty \leq \|x_k - a\| = d(x_k, a) \leq \sqrt{n}\|x_k - a\|_\infty.$$

Thus

$$d(x_k, a) \to 0 \quad \Leftrightarrow \quad \|x_k - a\|_\infty \to 0$$

and the right hand side is equivalent to $|x_{ki} - a_i| \to 0$ for all i.

(ii) Once again, $\|x_k - x_m\|_\infty \to 0 \Leftrightarrow \|x_k - x_m\| \to 0$. $\square$

Let now derive a fact from the previous theorem, which is familiar to us from the case $\mathbb{C} = \mathbb{R}^2$.

Corollary 6 (Completeness of $\mathbb{R}^n$)**.** *Each Cauchy sequence in $\mathbb{R}^n$ converges.*

Proof. Let (x_k) be a Cauchy sequence. By Theorem 5 all component sequences $(x_{ki})_{k\in\mathbb{N}}$ are Cauchy, for $i = 1, \dots, n$. By completeness of $\mathbb{R}$, Theorem II.13, $(x_{ki})_{k\in\mathbb{N}}$ converges to some $a_i \in \mathbb{R}$ as $k \to \infty$. Applying Theorem 5 once again gives that (x_k) converges to $a := (a_1, \dots, a_n)$. □

Corollary 7 (Bolzano-Weierstrass in $\mathbb{R}^n$)**.** *Each bounded sequence in $\mathbb{R}^n$ contains a convergent subsequence.*

Proof. Let (x_k) be a sequence with $\|x_k\| \le C$. Consider the first component sequence: It is bounded, $|x_{k1}| \le C$, and so the Bolzano-Weierstrass Theorem II.12 allows us to extract from $(x_k)_{k\in\mathbb{N}}$ a subsequence $(x_{k_\nu})_{\nu\in\mathbb{N}}$ with convergent first component: $x_{k_\nu 1} \to a_1$.

Consider the second component of this subsequence. It is also bounded, $|x_{k_\nu 2}| \le C$, and so from the first subsequence (x_{k_ν}) we can extract a further subsequence (x_{l_ν}), such that $x_{l_\nu 2} \to a_2$. The first components of the subsequence still converge, $x_{l_\mu 1} \to a_1$.

We repeat this argument for all components, extracting subsequences altogether n-times. This gives a subsequence $(x_{m_\nu})_{\nu\in\mathbb{N}}$ which converges in all components, to a point $a := (a_1, \dots, a_n)$. □

Another consequence of the theorem is that the limit theorems generalize to the vector case: If (x_k), (y_k) are convergent and $\lambda \in \mathbb{R}$, then also $(x_k + y_k)$, (λx_k), $\|x_k\|$ are convergent. To see the last claim, note that if $x_k \to a$ then also

$$\|x_k\| = \left\|(x_{k1}, \dots x_{kn})^\top\right\| = \sqrt{x_{k1}^2 + \dots + x_{kn}^2} \overset{\sqrt{\cdot}\text{ cts.}}{\longrightarrow} \sqrt{a_1^2 + \dots + a_n^2} = \|a\|. \tag{10}$$

Let us include here a simple statement which we will often use.

Lemma 8. *Let (x_k) be a bounded sequence in $\mathbb{R}^n$ and (λ_k) be a null sequence in $\mathbb{R}$. Then $\lambda_k x_k \to 0$.*

Proof. Since $\|x_k\| \le C$ for some $C \in \mathbb{R}$ we find

$$\|\lambda_k x_k\| = |\lambda_k| \|x_k\| \le \lambda_k C \to 0.$$

□

From Thm. 5 it is immediate that properties familiar from real sequences extend to the vector case: A convergent sequence has a unique limit, it is bounded, and it is Cauchy.

But these properties are also valid for sequences in arbitrary metric spaces X, as the proof given in Analysis I generalizes. Let us verify this on the example of the uniqueness of the limit: Suppose a, b are two limits of the sequence $(x_k) \in X$. Then

$$0 \le d(a,b) \le d(a, x_k) + d(b, x_k) \to 0 \quad \Rightarrow \quad d(a,b) = 0 \quad \overset{d\text{ positive}}{\Rightarrow} \quad a = b.$$

2.3. **Closed sets.** We generalize closed intervals to arbitrary metric spaces X (or $\mathbb{R}^n$):

Definition. (Cantor 1884) A subset $A \subset X$ is *closed [abgeschlossen]* if it has the following property: For each sequence $x_k \in A$ which converges to some $a \in X$ the limit a is also an element of A.

Note that closedness of a set A depends on the ambient space X, and that closed sets may also be open.

Examples. 1. $[a, b] \subset \mathbb{R}$ is closed: If $a \leq x_k \leq b$ and $x_k \to x$ then also $a \leq x \leq b$.

2. The interval $(0, 1] \subset \mathbb{R}$ is not closed: The sequence $(\frac{1}{k})_{k \in \mathbb{N}}$ has the limit 0 which is not contained in $(0, 1]$.

3. $\mathbb{R}$ is closed, similarly any metric space X (check!).

4. $\overline{B_r}(y)$ is closed: Consider $\overline{B_r}(y) \ni (x_k) \to a$. By convergence, for each $\varepsilon > 0$ there is $N(\varepsilon)$ with $d(x_k, a) < \varepsilon$ for all $k \geq N(\varepsilon)$. Thus $d(y, a) \leq d(y, x_k) + d(x_k, a) \leq r + \varepsilon$. But ε was arbitrary and so $d(a, y) \leq r$, that is, $a \in \overline{B_r}(y)$.

5. The rational numbers are not closed in $\mathbb{R}$: Any irrational number can be represented as the limit of a rational sequence.

By definition, any closed subset $Y \subset \mathbb{R}^n$ is a complete metric space (with standard distance).

Closed and open sets are complementary:

Theorem 9. *A subset $A \subset X$ is closed if and only if its* complement $A^c := X \setminus A$ *is open.*

For a topological space, this theorem becomes the definition of closed sets.

Proof. "$\Rightarrow$". Indirectly, to be shown: "A^c not open $\Rightarrow$ A not closed".

If A^c is not open then there is $a \in \mathbb{R}^n \setminus A$ such that for each ball $B_r(a)$ contains a point of A. This means $d(x_k, a) \leq \frac{1}{k} \to 0$. That is, $A \ni x_k \to a \notin A$, meaning that A is not closed.

`34. Lecture, Thursday, 26. April 07` ________________________________ `T 2`

"$\Leftarrow$": Indirectly, to be shown: "A not closed $\Rightarrow$ A^c not open".

If A is not closed, then we have a sequence $A \ni x_k \to a \notin A$. As $d(x_k, a) \to 0$, we can find for each $r > 0$ an index k, such that the ball $B_r(a)$ contains the sequence element $x_k \in A$. Thus $a \in A^c$ does not admit a ball $B_r(a) \subset A^c$ and so A^c is not open. $\square$

The theorem allows us to assert that some interesting subsets A of metric spaces X are closed:

Examples. 6. Any finite set $A := \{x_1, \dots, x_k\}$ is closed. Indeed, for $x \in A^c$ let us set $r(x) := \min\{d(x_i, x) : 1 \le i \le n\} > 0$, and so $B_{r(x)}(x) \subset A^c$, meaning A is open.

7. Suppose $x_k \to a$. We claim $A := \{x_k : k \in \mathbb{N}\} \cup \{a\}$ is closed. Indeed, for $x \in A^c$ set $\rho(x) := \frac{1}{2} d(x, a) > 0$. By convergence, there is $N \in \mathbb{N}$ such that only $x_1, \dots, x_N \notin B_\rho(a)$. Thus $r(x) := \min\big(\rho(x), d(x_1, x), \dots, d(x_N, x)\big) > 0$ and $B_{r(x)}(x) \cap A = \emptyset$, as desired.

Corollary 10. *In any metric space X, the closed sets satisfy:*
(i) X and $\emptyset$ are closed.
(ii) Arbitrary intersections $\bigcap_{i \in I} A_i$ of closed sets A_i are again closed.
(iii) Finite unions of closed sets are closed.

Proof. By the Theorem, it is sufficient to show that the complements of the sets in question are open. But this follows from Thm. 4 and the subsequent problem. For instance, for (i) note that $X^c = \emptyset$ is open by Thm. $4(i)$, which gives X is closed. Likewise for (ii), $\big(\bigcap_{i \in I} A_i\big)^c = \bigcup_{i \in I} A_i^c$ (check!), and for A_i^c open this set is again open by Thm. $4(ii)$. $\square$

Problems. 1. Find an example of an infinite union of closed sets which is not closed.
2. Show the Cantor set is closed.

There are two further notions worth introducing:

Definition. For $Y \subset X$ we define the following sets: The *closure [Abschluss]*

$$\overline{Y} := \big\{x \in X : \exists (x_k) \in Y \text{ such that } \lim_{k \to \infty} x_k = x\big\} = \big\{x \in X : B_\varepsilon(x) \cap X \neq \emptyset \; \forall \varepsilon > 0\big\}$$

and the *boundary [Rand]*

$$\partial Y = \big\{x \in X : \forall \varepsilon > 0 \; \exists y, z \in B_r(x) \text{ with } y \in Y \text{ and } z \in Y^c\big\}.$$

Examples. 1. $(0, 1] \subset \mathbb{R}$ has closure $[0, 1]$ and boundary $\{0\} \cup \{1\}$.
2. $\partial B_r(a) = \{x \in X : d(x, a) = r\}$.
3. $\mathbb{Q} \subset \mathbb{R}$ has $\overline{Q} = \mathbb{R}$ and $\partial \mathbb{Q} = \mathbb{R}$.
4. $Y := \mathbb{R} \times \{0\} \subset \mathbb{R}^2$ has closure Y and boundary Y.

There are many equivalent characterizations of closure and boundary: For instance, the closure is the smallest closed set containing Y, or the intersection of all closed supersets of Y. See problems.

2.4. **Compact subsets of $\mathbb{R}^n$.** We are interested to exhibit subsets Y of $\mathbb{R}^n$ such that the following two problems can be solved:

1. Given a sequence in Y, does it contain a subsequence converging to a limit in Y?
2. Given a continuous function $f\colon Y \to \mathbb{R}$, does it attain a maximum?

For $\mathbb{R}^n$, the answer to both questions is in the affirmative provided the set Y is closed and bounded. Indeed, for 1., boundedness of Y gives the existence of a convergent subsequence by the Bolzano-Weierstrass theorem, Cor. 7; and closedness gives that the subsequence has a limit within Y. We will study 2. only later.

Since such sets are significant, they deserve their own name:

Definition. A subset $K \subset \mathbb{R}^n$ is *compact [kompakt]* if it is closed and bounded.

Example. 1. A finite subset $\{x_1, \dots, x_k\} \subset \mathbb{R}^n$ is compact.
2. A convergent sequence together with its limit is compact (see previous subsection).
3. A closed ball is compact.
4. On open ball in $\mathbb{R}^n$ is not compact.
5. $\mathbb{R}^n$ is not compact.

To give but one indication for the significance of compact sets, let us show that compact subsets of the real line contain their supremum and infimum:

Proposition 11. *A compact subset $K \subset \mathbb{R}$ contains a maximum and minimum.*

Proof. Since K is bounded and $\mathbb{R}$ is complete, the supremum $s := \sup K$ exists. Let $x_k \in K$ be a sequence with $x_k \to s$. Since K is closed, $s = \lim_{k\to\infty} x_k \in K$. Similarly for the infimum. $\square$

We now introduce an important characterization of compactness. It is often simple to apply, and it will be the property to use when introducing compactness in metric spaces.

To state it, let us define an *open covering [offene Überdeckung]* of a subset $Y \subset X$ as a family $(U_i)_{i\in I}$ of open sets $U_i \subset X$, such that $\bigcup_{i\in I} U_i \supset Y$. We do not make any requirements on the index set I.

Examples. 1. The sets $\{U_i := (\frac{1}{i+2}, \frac{1}{i}) : i \in \mathbb{N}\}$ form an open covering of the interval $(0, \frac{1}{2}]$.
2. The rational numbers $\mathbb{Q} \subset \mathbb{R}$ have the open covering (9).

Definition. A set $Y \subset X$ has the *Heine-Borel covering property*, if for each open covering $(U_i)_{i\in I}$ of Y there exists a finite subcovering, that is, there are finitely many indices $i_1, \dots, i_k$, such that

$$Y \subset U_{i_1} \cup \dots \cup U_{i_k}.$$

Example. 1. The set $\{\frac{1}{k} : k \in \mathbb{N}\}$ does not have the Heine-Borel covering property (consider the covering of Ex. 1 above).

2. $\{\frac{1}{k} : k \in \mathbb{N}\} \cup \{0\}$ has the covering property. Let us prove this fact. If $(U_i)_{i\in I}$ is the open covering, then there must exist an index N such that $\{0\} \in U_N$. But U_N is open and hence $(-\varepsilon, \varepsilon) \subset U_N$ for some ε with $\frac{1}{\varepsilon} \in \mathbb{N}$. Thus also $\{\frac{1}{k} : k > \varepsilon\} \subset U_N$. But the remaining set $M := \{1, \frac{1}{2}, \ldots, \varepsilon\}$ is finite. Thus M is contained in finitely many U_i's, whose union, together with U_N, forms the desired finite cover.

Theorem 12 (Heine-Borel)**.** *A subset $K \subset \mathbb{R}^n$ is compact if and only if it has the Heine-Borel covering property.*

Proof. "$\Leftarrow$" Let us show boundedness first. Fix some point $a \in \mathbb{R}^n$. As $\bigcup_{i\in\mathbb{N}} B_i(a) = \mathbb{R}^n$, the sets $(B_i(a))_{i\in\mathbb{N}}$ form a finite open covering of K. By the Heine-Borel covering property, finitely many of these sets cover, $K \subset B_{i_1}(a) \cup \ldots \cup B_{i_k}(a)$. Thus for $N := \max\{i_1, \ldots, i_k\}$ we have $K \subset B_N(a)$. This means K is bounded.

To see K is closed, let $y \in K^c$ be arbitrary. Set for $n \in \mathbb{N}$

$$U_i := \Big\{x \in X : d(x, y) > \frac{1}{i}\Big\}.$$

Then U_i is open (why?) and

$$\bigcup_{i=1}^{\infty} U_i = \mathbb{R}^n \setminus \{y\} \supset K.$$

Once again, $K \subset U_{i_1} \cup \ldots \cup U_{i_k}$, and so for $N := \max\{i_1, \ldots, i_k\}$ we find $K \subset U_N$. Thus $B_{1/N}(y) \cap K = \emptyset$. Since y is arbitrary this means K^c is open, hence K is closed by Thm. 9.

"$\Rightarrow$" Let $(U_i)_{i\in I}$ be a covering of K. Indirectly, we assume that K is not covered by finitely many of these sets. In the following proof we will use nested cubes to derive a contradiction; this is a natural a generalization of interval bisection to several variables.

Since K is bounded, there is a cube W_0 containing K. Suppose W_0 has edgelength ℓ. Let us subdivide W_0 into 2^n cubes having edglength $\ell/2$. At least one of these cubes, call it W_1, must have an intersection $W_1 \cap K$ which cannot be covered with finitely many U_i's.

We iterate this reasoning to find a sequence of nested cubes $W_0 \supset W_1 \supset \ldots \supset W_k \supset \ldots$ such that W_k has edgelength $\ell/2^k$. Each cube W_k has the property $(*)$ that $W_k \cap K$ cannot covered with finitely many of the U_i.

Now choose a sequence of points $x_k \in W_k \cap K$ for $k \in \mathbb{N}$. By construction, this sequence is Cauchy and hence has a limit $a \in \mathbb{R}^n$. Since K is closed, the limit a is contained in K. By assumption, the point a must be contained in some set U_N of the covering. But U_N is open, hence contains a ball $B_r(a)$. This ball also contains all but finitely many cubes W_k. That is, there is an index M, such that U_N contains the cubes $W_M, W_{M+1}, \ldots$, contradicting the fact that they satisfy property $(*)$. Thus our assumption is false. $\square$

`35. Lecture, Thursday, 3. May 07 ______________________________ Ü 2, T 3`

2.5. **Compact subsets of metric spaces.** Here the situation is different from $\mathbb{R}^n$, since boundedness is not a significant property for metric spaces. In fact, for closed and bounded subsets of metric spaces, sequences need not have a convergent subsequence, functions need not take their extrema, and the Heine-Borel theorem is no longer true:

Example. Consider the metric space $(\mathbb{R}, d)$, where $d(x,y) := \frac{|x-y|}{1+|x-y|}$. As shown in Ex. 5 of 1.3 and Problem 1, this is a metric. Note that d is bounded, and so the set $\mathbb{R}$ becomes a bounded subset of $(\mathbb{R}, d)$; it is also closed. But the sequence $(n)_{n\in\mathbb{N}}$ does not contain a convergent subsequence, the function $f(x) = d(x, 0)$ does not attain its supremum, and $\{(n-1, n+1) : n \in \mathbb{Z}\}$ is an open cover of $\mathbb{R}$ which does not contain a finite subcovering.

These problems are resolved by turning the covering property into a definition:

Definition. A subset K of a metric space (X, d) is called *compact* if it has the Heine-Borel covering property.

This means no change for subsets of $\mathbb{R}^n$, thanks to the Heine-Borel Theorem. But our assertions 1. and 2., stated at the beginning of the preceding subsection, will then become true for compact subsets of metric (or topological) spaces. Let us first confirm assertion 1:

Theorem 13 (Bolzano-Weierstrass theorem)**.** *Let K be a compact subset of a metric space (X, d), and $(x_k)_{k\in\mathbb{N}}$ be a sequence of points in K. Then (x_k) has a convergent subsequence with limit $a \in K$.*

Proof. Suppose no subsequence converges to a limit $a \in K$. Then for each $a \in K$ we find some radius $r(a)$ such that $U(a) := B_{r(a)}(a) \cap \{x_k : k \in \mathbb{N}\}$ is finite. (Indeed, else we find for each $i \in \mathbb{N}$ an index $k(i)$ such that $x_{k(i)} \in B_{1/i}(a)$; selecting an increasing sequence of indices from $(k(i))_{i\in\mathbb{N}}$ gives a subsequence convergent to a.)

Let us now use the Heine-Borel covering property: Since $K \subset \bigcup_{a\in K} U(a)$ we can select finitely many points $a_1, \ldots, a_n \in K$, such that $K \subset U(a_1) \cup \ldots \cup U(a_n)$. But the right hand side contains only finitely many sequence points, which is a contradiction. $\square$

3. Continuous mappings in several variables

3.1. **Continuity in terms of the limit.** To define continuity, we apply our concept of the limit, working in the generality of metric spaces:

Definition. Let X, Y be metric spaces (in particular, they may be subsets of $\mathbb{R}^n$, $\mathbb{R}^m$). A mapping $f\colon X \to Y$ is *continuous [stetig] at* $a \in X$, if

$$\lim_{k\to\infty} f(x_k) = f(a) \qquad \text{for each sequence } x_k \in X \text{ with } \lim_{k\to\infty} x_k = a. \tag{11}$$

f is *continuous*, if f is continuous in each $a \in X$.

Examples. 1. *(Linear functionals)*: For given $a = (a_1, \dots, a_n)^\top \in \mathbb{R}^n$ consider the linear map $L\colon \mathbb{R}^n \to \mathbb{R}$,

$$L(x) = a_1x_1 + \ldots + a_nx_n = \langle a, x\rangle.$$

Invoking the Schwarz inequality gives

$$\big|L(x+h) - L(x)\big| \overset{L \text{ linear}}{=} |L(h)| = |\langle a, h\rangle| \overset{(5)}{\leq} \|a\|\|h\|. \tag{12}$$

We conclude that for $h \to 0$ also $L(x+h) \to L(x)$, implying that L is continuous.

2. Any norm $\|.\|\colon \mathbb{R}^n \to \mathbb{R}$ is continuous (see (10) for the case of the Euclidean norm). Using the inverse triangle inequality, we verify for $h \to 0$

$$\big|\|x+h\| - \|x\|\big| \overset{(6)}{\leq} \|h\| \to 0 \qquad \Rightarrow \qquad \|x+h\| \to \|x\|.$$

For the case of $\mathbb{R}^n$, we saw that convergence is equivalent to componentwise convergence, see Thm. 5. Consequently:

Theorem 14. *A map $f\colon X \to \mathbb{R}^m$ is continuous at $a \in X$ if and only if all component functions $f_1\colon X \to \mathbb{R}$, $\dots f_m\colon X \to \mathbb{R}$ are continuous at a.*

Example. 1. Linear maps $L\colon \mathbb{R}^n \to \mathbb{R}^m$, $L(x) = Ax$ where A is a $m \times n$-matrix, are continuous: Each component of $L = (L_1, \dots, L_m)^\top$ is a linear functional, which is continuous.

2. Addition

$$+\colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n, \qquad (x, y) \mapsto x + y$$

is continuous: If $(x_k, y_k) \to (a, b)$ then indeed $\lim(x_k + y_k) = \lim x_k + \lim y_k = a + b$.

3. Similarly: Multiplication from $\mathbb{R}^2$ to $\mathbb{R}$ (or $\mathbb{C}^2 \to \mathbb{C}$) is continuous.

Theorem 15. *Let X, Y, Z be metric spaces (or subsets of Euclidean spaces). If $f\colon X \to Y$ is continuous at a and $g\colon Y \to Z$ is continuous at the point $f(a)$, then $g \circ f\colon X \to Z$ is continuous at a.*

Proof. Copying the proof from one variable we find, as desired,

$$x_k \to a \quad \Rightarrow \quad g\big(f(a)\big) = g\big(f(\lim_{k\to\infty} x_k)\big) \overset{f \text{ cts.}}{=} g\big(\lim_{k\to\infty} f(x_k)\big) \overset{g \text{ cts.}}{=} \lim_{k\to\infty} g\big(f(x_k)\big). \qquad \square$$

Examples. 1. Suppose $f, g \colon X \to \mathbb{R}^m$ are continuous and $\lambda \in \mathbb{R}$. Then $f + g$ is continuous as the composition of the addition function after (f, g), and λf is the composition of f with a linear map.
2. For $f \colon X \to \mathbb{R}^n$ continuous also $\|f\|$ is continuous: Again this is the composition of $\|.\|$ after f.

3.2. The ε-δ-test.

Theorem 16 (ε-δ-test)**.** *Let X, Y be metric spaces; perhaps $X \subset \mathbb{R}^n$, $Y = \mathbb{R}^m$. Then f is continuous at $a \in X$ if and only if the following holds: For each $\varepsilon > 0$ there is $\delta = \delta(a, \varepsilon) > 0$, such that*

$$d\big(f(x), f(a)\big) < \varepsilon \qquad \text{for all } x \in X \text{ with } d(x, a) < \delta. \tag{13}$$

As for complex maps this means that each ε-ball $B_\varepsilon\big(f(a)\big)$ about $f(a)$ contains the entire image of some δ-ball about a.

Example. Let $p \in \mathbb{R}^n$ and $f \colon \mathbb{R}^n \to \mathbb{R}$, $f(x) := \|x - p\|$. The graph is a cone with unit slope, and so we expect that $\delta := \varepsilon$ works. Indeed,

$$|f(x) - f(a)| = \big|\|x - p\| - \|a - p\|\big| \overset{\text{inverse } \Delta\text{-inequ.}}{\leq} \big\|x - p - (a - p)\big\| = \|x - a\|,$$

so that $\|x - a\| < \delta$ verifies $|f(x) - f(a)| < \varepsilon$.

Proof. The proof carries over literally from the one-dimensional case.

"$\Leftarrow$": We assume the ε-δ-test at $a \in X$. We need to show that if $x_k \to a$ then $f(x_k) \to f(a)$. Let $\varepsilon > 0$ be arbitrary, and pick $\delta = \delta(a, \varepsilon)$ by (13). Since $x_k \to a$ we can choose $N = N(\delta) \in \mathbb{N}$ such that $d(x_k, a) < \delta$ for all $k \geq N$. But (13) then implies $d\big(f(x_k), f(a)\big) < \varepsilon$ for all $k \geq N$. As ε was arbitrary, this gives $f(x_k) \to f(a)$, as desired.

"$\Rightarrow$". Indirectly: Suppose for a particular error bound $\varepsilon > 0$ there were no $\delta > 0$ satisfying the condition (13). In particular, (13) could not be satisfied with any $\delta = \frac{1}{k}$, where $k \in \mathbb{N}$. Thus there exist $x_k \in X$ with $d(x_k, a) < \frac{1}{k}$, so that $d\big(f(x_k), f(a)\big) \geq \varepsilon$. Therefore we have $x_k \to a$ but not $f(x_k) \to f(a)$, contradicting the limit test. $\square$

Again we have the corollary:

Corollary 17. *Let X be a metric space and $f\colon X \to \mathbb{R}^m$ be continuous. If $f(a) \neq 0$ then there is $\delta > 0$ such that $f(x) \neq 0$ for all $x \in B_\delta(a)$.*

Proof. For $\varepsilon := \|f(a)\|$ let us choose δ satisfying the ε-δ-test of continuity.

Then for all x with $d(x,a) < \delta$ we verify:

$$\|f(x)\| = \|f(a) + f(x) - f(a)\| \overset{\text{inv. }\Delta\text{-inequ.}}{\geq} \|f(a)\| - \underbrace{\|f(x) - f(a)\|}_{<\varepsilon = \|f(a)\|} > 0.$$

□

3.3. **The topological characterization in terms of open sets.** There is third characterization of continuity, which we will need when discussing the inverse mapping theorem. At the same time, it suggests how to define continuity for the most general setup, namely for maps between topological spaces.

The characterization is in terms of open sets. Note that the continuous image of an open set is not necessarily open: The interval $(-1,1)$ maps under the continuous function x^2 onto the interval $[0,1)$, which is not open. But for inverse images we have:

Theorem 18. *Let X, Y be metric spaces (perhaps $X \subset \mathbb{R}^n$, $Y = \mathbb{R}^m$) and let $f\colon D \to Y$ for $D \subset X$ open. Then f is continuous if and only if for each open set $U \subset Y$ the preimage $f^{-1}(U)$ is open in X.*

Examples. 1. Consider the continuous map from $\mathbb{R} \to \mathbb{R}$, $x \mapsto x^2$. The preimage of the open interval $(1,4)$ is $(-2,-1) \cup (1,2)$, which is indeed open.
2. Consider the discontinuous map $f(x) := 0$ for $x \leq 0$ and $f(x) := 1$ for $x > 0$. The preimage of $(-\frac{1}{2}, \frac{1}{2})$ is $(-\infty, 0]$, which is not open.

Proof. "$\Leftarrow$": Let $y = f(x)$. We want to assert the ε-δ-condition at x. Let $\varepsilon > 0$. Consider the open subset $U := B_\varepsilon(y)$ of Y. By assumption its preimage is open again. Hence the preimage contains an open ball about each of its points. In particular, it contains some ball $B_\delta(x)$. Hence (13) holds.

"$\Rightarrow$": Let f be continuous and $x \in f^{-1}(U)$. We need to show that some ball about x is entirely contained in the preimage $f^{-1}(U)$.

Consider first the case $D = X$. Since U is open, we can choose $\varepsilon > 0$ so that $B_\varepsilon\big(f(x)\big) \subset U$. Then the ε-δ-test gives $\delta > 0$ such that the ball $B_\delta(x) = \{\xi \in X : d(x,\xi) < \delta\} \subset X$ maps into $B_\varepsilon\big(f(x)\big) \subset U$. Thus $B_\delta(x) \subset f^{-1}(U)$, showing that $f^{-1}(U)$ is open.

In case D is a proper subset of X, we only know that $B_\delta(x) \cap D$ maps into $B_\varepsilon\big(f(x)\big)$. But $B_\delta(x) \cap D$ is the intersection of open sets, hence itself open. Thus there exists $0 < r \leq \delta$, such that $B_r(x) \subset B_\delta(x) \cap D$. Hence $B_r(x) \subset f^{-1}(U)$, proving openness in this case as well. $\square$

`36. Lecture, Tuesday, 8. May 07` ______________________________ `Ü 3`

3.4. **Minima and maxima of continuous functions.** Once again we ask: Does a function $f\colon X \to \mathbb{R}$ attain its supremum and infimum? In one dimension, we saw that for X an interval $[a,b] \subset \mathbb{R}$ this is true, while for open intervals or $X = \mathbb{R}$ it may fail (examples?). To generalize to arbitrary dimension, we replace closed bounded intervals with compact sets:

Theorem 19. *Let X, Y be metric spaces (or $\mathbb{R}^n, \mathbb{R}^m$ respectively), and suppose $K \subset X$ is a nonempty compact set.*
(i) [Weierstrass' Hauptlehrsatz, 1861] *If $f\colon K \to \mathbb{R}$ is continuous then* f attains a maximum and minimum over K, *that is, there are points $a, b \in K$ such that*

$$f(a) \le f(x) \le f(b) \qquad \text{for all } x \in K.$$

(ii) *If $f\colon K \to Y$ is continuous then its range $f(K)$ is compact.*

The proof is a nice combination of the Heine-Borel covering property with the topological characterization of continuity:

Proof. (ii) Let $(V_i)_{i\in I}$ be an open covering of $f(K) \subset Y$. By continuity, $U_i := f^{-1}(V_i)$ is open, and so $(U_i)_{i\in I}$ forms an open covering of K. The Heine-Borel covering property lets us select finitely many indices such that $K \subset U_{i_1} \cup \ldots \cup U_{i_k}$. But then, as desired,

$$f(K) \subset f\big(U_{i_1}\big) \cup \ldots \cup f\big(U_{i_k}\big) \subset V_{i_1} \cup \ldots \cup V_{i_k},$$

using the fact that $f\big(f^{-1}(V)\big) \subset V$ for any set V.

(i) By (ii), the set $f(K) \subset \mathbb{R}$ is compact. By Prop. 11, the compact set $f(K)$ contains maximum and minimum. $\square$

Problems. 1. For $f\colon X \to Y \supset V$ prove $f\big(f^{-1}(V)\big) \subset V$. Can the inclusion be strict? Second, determine a relationship between $U \subset X$ and $f^{-1}\big(f(U)\big)$ and prove it.
2. Check that part (i) of the theorem can be proven as in one dimension (see Thm. I.14) by applying, this time, the generalized Bolzano-Weierstrass Thm. 13. If Y happens to be a normed vector space then part (ii) follows from (i), by considering the real valued function $\|.\|$ on K.

We want to give two applications. First:

Example. Let $K \subset \mathbb{R}^n$ be compact and $a \in \mathbb{R}^n \setminus K$. Then K contains a point closest to a: Indeed, $x \mapsto d(x,a)$ attains a minimum on some point $x_0 \in K$, that is, $d(x,a) \geq d(x_0,a)$ for all $x \in K$. Show by counterexample that our statement fails for open sets.

Our second application is a theorem:

Theorem 20. *All norms on $\mathbb{R}^n$ are equivalent.*

Proof. Let $\|.\|_1$ and $\|.\|_2$ be two norms on $\mathbb{R}^n$, $n \geq 1$. Since norm equivalence is an equivalence relation, it suffices to show that $\|.\|_1$ is equivalent to the Euclidean norm $\|.\|$. That is, we need to find constants $c, C \in (0,\infty)$ such that $c\|x\| \leq \|x\|_1 \leq C\|x\|$ holds for all $x \in \mathbb{R}^n$.

Consider the *unit sphere* $\mathbb{S}^{n-1} := \{\xi \in \mathbb{R}^n : \|\xi\| = 1\}$ of dimension $(n-1) \in \mathbb{N}_0$. The unit sphere is compact: Clearly it is bounded. Moreover, for each convergent sequence $\mathbb{S}^{n-1} \ni (x_k) \to a \in \mathbb{R}^n$, the continuity of $\|.\|$ implies that the limit satisfies $\|a\| = 1$. So indeed $a \in \mathbb{S}^{n-1}$, proving closedness.

Now consider the function

$$f\colon \mathbb{S}^{n-1} \to (0,\infty), \qquad f(\xi) := \|\xi\|_1 = \frac{\|\xi\|_1}{\|\xi\|}.$$

By Example 2 of Sect. 3.1 the function f is continuous and so takes a minimum c and a maximum C on $\mathbb{S}^{n-1}$. Since c, C are attained by f, they must be positive. Consequently, for $\xi \in \mathbb{S}^{n-1}$,

$$c \leq \frac{\|\xi\|_1}{\|\xi\|} \leq C \qquad \Rightarrow \qquad c\|\xi\| \leq \|\xi\|_1 \leq C\|\xi\|. \tag{14}$$

Now consider arbitrary $x = \|x\|\frac{x}{\|x\|} \in \mathbb{R}^n \setminus \{0\}$. We consider (14) for $\xi = \frac{x}{\|x\|} \in \mathbb{S}^{n-1}$ and multiply with $\|x\| > 0$ to obtain

$$c\|x\|\left\|\frac{x}{\|x\|}\right\| \leq \|x\|\left\|\frac{x}{\|x\|}\right\|_1 \leq C\|x\|\left\|\frac{x}{\|x\|}\right\| \quad \overset{\text{norm linear}}{\Longleftrightarrow} \quad c\|x\| \leq \|x\|_1 \leq C\|x\|.$$

Since the last equation is trivial for $x = 0$ our claim is proved. $\square$

The proof can be summarized by saying that linearity of the norm implies that a norm on $\mathbb{R}^n$ is determined by its values on the compact set $\mathbb{S}^{n-1}$.

The theorem no longer holds for infinite dimensional normed vector spaces. Such vector spaces, if complete (Cauchy sequences converge), are called *Banach spaces.* They are the main topic of functional analysis.

Problem. Find to inequivalent norms on the infinite dimensional vector space
$\ell^2 := \{(a_k)_{k\in\mathbb{N}} \in \mathbb{C}\colon \sum_{k=0}^{\infty} |a_k|^2 \text{ converges}\}$.

4. Curves

After some rather formal sections let us now discuss a concrete geometric topic.

4.1. **Continuous and differentiable curves.** To start with an example, the circle $\mathbb{S}^1$ is on the one hand the image of the mapping $[0, 2\pi] \ni t \mapsto e^{it} \in \mathbb{C}$, and on the other hand the level set $\{x \in \mathbb{R}^2 : \|x\| - 1 = 0\}$ of the Euclidean norm $\|.\|$. Likewise, an arbitrary curve can be described in two ways:
• The *parametric description* of a curve is a map $t \mapsto c(t)$, that is the curve consists of the points $\{(t, c(t)) : t \in I\}$.
• The *implicit description* is by the level set of a function. In case of a curve in the plane $\mathbb{R}^2$, the curve is written $\{(x, y) \in \mathbb{R}^2 : f(x, y) = 0\}$ for some function $f \colon \mathbb{R}^2 \to \mathbb{R}$.
We will deal with the problem of how to transit between the two descriptions only later.

The first description is advantageous for us since it allows us to apply mappings more directly. Hence we define:

Definition. (i) A *(parameterized) curve [(parametrisierte) Kurve]* is a continuous map $c \colon I \to \mathbb{R}^m$ of an interval $I \subset \mathbb{R}$.
(ii) The curve c is *differentiable* if all its component functions $c(t) = \big(c_1(t), \ldots, c_m(t)\big)^\top$ are continuously differentiable; the vector $c'(t) = \big(c_1'(t), \ldots, c_m'(t)\big)^\top$ is called *tangent vector [Tangentialvektor]*.

Our definition of $c'(t)$ can again be regarded as the limit of the secant directions:

$$\begin{aligned} c'(t) &:= \big(c_1'(t), \ldots, c_m'(t)\big)^\top \\ &= \Big(\lim_{s\to 0} \frac{c_1(t+s) - c_1(t)}{s}, \cdots, \lim_{s\to 0} \frac{c_m(t+s) - c_m(t)}{s}\Big)^\top = \lim_{s\to 0} \frac{c(t+s) - c(t)}{s} \end{aligned}$$

In physics, a parameterized curve describes the motion of a point or particle. At time t, the particle is at location $c(t)$ and has velocity vector *[Geschwindigkeitsvektor]* $c'(t)$; its velocity is $\|c'(t)\|$.

Often we are only interested in the *path* [*Weg*] $\{c(t) : t \in I\}$ which is a subset of $\mathbb{R}^m$. Given a path we could say that a curve describes a path together with a specific parametrization. To use a concrete picture, the path corresponds to railway track, while the parametrized curve corresponds to a timetable for a train.

Examples of curves. 1. A straight line $t \mapsto p + ta$ for $p, a \in \mathbb{R}^m$.
2. *Neil's parabola [Neilsche Parabel]* $c(t) = (t^2, t^3)^\top$ for $t \in \mathbb{R}$. The tangent vector is

$c'(t) = (2t, 3t^2)^\top$; at the tip *[Spitze]* $c(0) = 0$ the curve turns backwards, and the velocity vanishes, $\|c'(0)\| = 0$.

While most curves we encounter are differentiable, there are interesting pathological counterexamples.

Theorem 21 (Peano 1890)**.** *There exists a (continuous) curve* $c\colon [0,1] \to [0,1]\times[0,1] \subset \mathbb{R}^2$ *which is surjective onto the unit square.*

The Peano curves indicate that continuity is a rather weak assumption. For instance, the notion of dimension is not preserved under continuous maps. For the proof see, e.g., the beautiful book of Sagan [S].

It is also interesting to discuss this result with respect to cardinality: Since there is a bijection of some subset of the unit interval $[0,1]$ to the unit square, the sets $[0,1]$ and $[0,1]^2$ are equipotent. Similarly $[0,1]$ is equipotent to $[0,1]^n$ for $n \in \mathbb{N}$.

Problem. If we only want to show $[0,1]$ and $[0,1]^2$ are equipotent, (discontinuous) bijections suffice. Use decimal expansions to exhibit an explicit bijection of the unit interval to the unit square.

`37. Lecture, Thursday, 10. May 07 ________________________ T 4`

4.2. **Length.** To define the length of a curve we give a continuous version of the formula "distance travelled = velocity · time":

Definition. Let $c\colon [a,b] \to \mathbb{R}^m$ be differentiable. Then the *length [Länge]* of c is

$$L(c) := \int_a^b \|c'(t)\|\, dt. \tag{15}$$

Examples. 1. Consider the *circle [Kreis]* of radius $r > 0$ with parameterization

$$c\colon [0,2\pi] \to \mathbb{C} = \mathbb{R}^2, \qquad c(t) = re^{it} = (r\cos t, r\sin t)^\top.$$

It has velocity $\|c'(t)\| = r|e^{it}| = r$. Thus the length is $L(c) = \int_0^{2\pi} \|c'(t)\|\, dt = 2\pi r$. This assertion concludes a long chain of arguments: We had defined $\frac{\pi}{2}$ as the first positive zero of the function $\cos t = \operatorname{Re} e^{it}$. Only now have we verified that this agrees with the ancient definition of π as the ratio between diameter *[Durchmesser]* and circumference *[Umfang]* of a circle.

2. An *ellipse [Ellipse]* can be parametrized as $c\colon [0,2\pi] \to \mathbb{R}^2$, $c(t) = \big(a\cos t, b\sin t\big)^\top$, where $a, b > 0$. Since $\|c'(t)\| = \sqrt{a^2\sin^2 t + b^2\cos^2 t}$ the circumference of the ellipse is

$$L = \int_0^{2\pi} \sqrt{a^2\sin^2 t + b^2\cos^2 t}\, dt.$$

This *elliptic integral* has no elementary solution; traditionally, elliptic integrals were tabulated. Let us mention a few important facts about ellipses (check or see, e.g. [J, p.87/88]; problems?):
(i) An ellipse is the image of a circle under some linear map (which one?).
(ii) Any linear image of a circle or of an ellipse is again an ellipse (possibly degenerate with a or $b = 0$). This can be seen by a principal axis transformation.
(iii) The ellipse is a conic section, that is, the result of an intersection of a cone with a plane.
(iv) It is also the locus of points with distance sum to the two foci constant.
(v) Light emitted from one focus gets focused at the other focus.

3. A graph $c(t) = \big(t, f(t)\big)$, where $f\colon [a,b] \to \mathbb{R}$ is continuously differentiable, has tangent vector $c' = (1, f')$ and length $L(c) = \int_a^b \sqrt{1 + (f'(t))^2}\, dt$.

Let us state two properties of the length. First, the chain rule shows that our definition does not depend on the parameterization chosen: If $\varphi\colon [a,b] \to [\alpha,\beta]$ is differentiable with differentiable inverse, then the reparameterized curve $\tilde{c} := c \circ \varphi$ has the same length,

$$L(\tilde{c}) = \int_\alpha^\beta \big\|(c\circ\varphi)'(t)\big\|\, dt \overset{\text{chain rule}}{=} \int_\alpha^\beta \big\|c'\big(\varphi(t)\big)\varphi'(t)\big\|\, dt \overset{\text{substitution}}{=} \int_a^b \big\|c'(s)\big\|\, ds = L(c).$$

Second, consider a *motion* B of $\mathbb{R}^m$, that is, $B(x) = Ax + T$, where A is an orthogonal matrix, and $T \in \mathbb{R}^m$ represents a translation. It can be shown that motions are the only diffeomorphisms of $\mathbb{R}^m$ which preserve length. Thus $A \in \mathsf{O}(m)$ preserves the norm, $\|Av\| = \|v\|$, and so

$$L(B\circ c) = \int_a^b \big\|(Ac+T)'(t)\big\|\, dt = \int_a^b \big\|Ac'(t)\big\|\, dt = \int_a^b \big\|c'(t)\big\|\, dt = L(c).$$

Thus the length of a curve is motion invariant, or, expressed in physics language, it does dependent on the choice of coordinate system.

Like differentiation, we declare integration componentwise by setting, for $f\colon [a,b] \to \mathbb{R}^m$ continuous,

$$\int_a^b f(t)\, dt := \Big(\int_a^b f_1(t)\, dt, \cdots, \int_a^b f_m(t)\, dt\Big)^\top \in \mathbb{R}^m.$$

We can now state a continuous version of the triangle inequality:

Proposition 22 (Integral estimate)**.** *If $f\colon [a,b] \to \mathbb{R}^n$ is continuous then*

$$\Big\| \int_a^b f(t)\, dt \Big\| \le \int_a^b \big\|f(t)\big\|\, dt. \tag{16}$$

Proof. Define the vector $I := \int_a^b f(t)\,dt$. Clearly (16) holds for $I = 0$. Otherwise:

$$\begin{aligned}\|I\|^2 = \langle I, I\rangle &= \Big\langle \int_a^b f(t)\,dt,\ I\Big\rangle = \int_a^b f_1(t)\,dt\cdot I_1 + \ldots + \int_a^b f_m(t)\,dt\cdot I_m \\ &= \int_a^b \big(f_1(t)I_1 + \ldots + f_m(t)I_m\big)\,dt = \int_a^b \big\langle f(t), I\big\rangle\,dt \\ &\overset{\text{Schwarz}}{\leq} \int_a^b \|f(t)\|\,\|I\|\,dt = \|I\| \int_a^b \|f(t)\|\,dt.\end{aligned}$$

Dividing by $\|I\|$ gives (16). □

As an application, let us invoke the Fundamental Theorem of calculus componentwise to each component:

$$L(c) = \int_a^b \|c'(t)\|\,dt \geq \left\| \int_a^b c'(t)\,dt \right\| \overset{\text{Fund.Thm.}}{=} \big\|c(b) - c(a)\big\| \tag{17}$$

Thus the length of any differentiable curve is at least the distance between its endpoints. As the latter is precisely the length of a straight line, we find that no curve is shorter than a straight line.

Remark. We can apply this insight to polygons P inscribed in a differentiable curve c: By (17), any such P has shorter (or equal) length than the curve, $L(P) \leq L(c)$. In case c is differentiable, the supremum of $L(P)$ can be shown to agree with $L(c)$ (problems?). This idea can be turned into a definition of length for curves which are not necessarily differentiable: If the length of all inscribed polygons has a supremum, then we call the curve *rectifiable [rektifizierbar]*. There are many curves which are not differentiable but still are rectifiable (examples?). The Peano curve provides an example of a non-rectifiable curve.

Outlook. The theory of curves and surfaces in space is called *differential geometry*; I will teach a class on this topic in your fourth term.

Summary

In the theory of differentiability, which we will study next, our main concern will be the space $\mathbb{R}^n$. Nevertheless, in the present section we introduced some more general spaces, which are significant for several areas of higher mathematics. The space $\mathbb{R}^n$ with the Euclidean norm $\|x\| := \sqrt{\sum x_i^2}$ presents a special case of a *normed vector space*. When the dimension is infinite, such vector spaces are the basic object of functional analysis. In turn, a normed vector space becomes a *metric space* by setting $d(x, y) := \|x - y\|$; this is the starting point of topology.

In the generality of metric spaces, the notions of open, closed, and compact sets were introduced: Open sets contain a neighbourhood of each point, closed sets have open complements, and compact sets have the Heine-Borel covering property. The most intricate notion here is compactness. In order for the Bolzano-Weierstrass theorem and the theorem of the maximum to remain valid in arbitrary metric spaces, we need to define compact sets by the covering property; specifically for $\mathbb{R}^n$, compact subsets agree with closed and bounded subsets.

Mathematicians love to present a concept in larger generality than directly necessary. The reason is that by introducing a concept like openness or continuity in a structure that only has the ingredients necessary for the concept (here: metric spaces), the arguments to be used become most transparent. The same could be said about linear algebra.

Convergence of sequences and continuity of functions have definitions that generalize directly from the case of $\mathbb{R}$ or $\mathbb{C}$ to $\mathbb{R}^n$ or metric spaces. There is a further test for continuity in terms of open sets. For the case of a vector target $\mathbb{R}^n$ the notions of convergence and continuity are equivalent to componentwise convergence or continuity.

A curve is differentiable or integrable if all its components have the respective property. Only the notion of differentiability is necessary to define the length of a curve. Curves in the plane, in space, or in $\mathbb{R}^n$, will play an important role for multi-variable calculus: By restricting a function of several variables to a curve, we obtain a single-variable function, and thus we can use the one-variable tools for its investigation.

Part 6. Differentiation in several variables

We study differentiation for mappings between higher dimensional spaces. We have seen that the one-variable tools apply directly to the case of curves with their vector valued range. However, it requires new concepts to differentiate when the domain has several dimensions. The idea of differentiation as linear approximation has the power to generalize to this setting. This means we approximate functions of several variables by linear maps, and so linear algebra plays an essential role in the present part.

We introduce several notions of differentiability. Then we derive differentiation rules, the most important one being the chain rule. As an application of differentiability, we discuss extrema of scalar valued functions alongside with a generalized Taylor formula.

For this part, always $n, m \in \mathbb{N}$ and U denotes an open subset of $\mathbb{R}^n$. We write vectors as columns, and use the transposed sign $\cdot^\top$. We employ the standard basis of $\mathbb{R}^n$ throughout, and denote it with $e_1 := (1, 0, \dots, 0)^\top$, …, $e_n := (0, \dots, 0, 1)^\top$.

0.1. **Visualisation of maps.** Maps $f\colon X \to \mathbb{R}^m$ with $X \subset \mathbb{R}^n$ can be visualized with one of the following methods:

1. The *graph* of f is the set

$$\Gamma(f) := \big\{\big(x, f(x)\big) : x \in X\big\} \subset X \times \mathbb{R}^m \subset \mathbb{R}^{n+m}.$$

For functions of planar domains ($n = 2$, $m = 1$) the graph is a subset of $\mathbb{R}^3$.
Describe the following geometric objects as the graphs of suitable functions: cone, hemisphere, half cylinder (with horizontal axis), a plane with normal $(1, 1, 1)$ through the origin.

2. *Level set* representation: The levels $N_y := \{x \in X : f(x) = y\}$ can be drawn into the domain X for various values of y. This is the familiar way to code heights on maps.

3. *Vector fields* for the case $n = m$: We draw the domain and attach to each point x the image vector $f(x)$. Which function $f\colon \mathbb{R}^2 \to \mathbb{R}^2$ generates a radial vector field, which one a curl?

4. *Grid models [Gitternetzmodelle]* are particular useful to draw complex functions $f\colon \mathbb{C} \to \mathbb{C}$, which cannot be represented in $\mathbb{R}^3$ using graphs or level sets. A grid is chosen in the domain (it may be axis-parallel or polar), and its image is drawn in the target. Check the case of a linear map $f\colon \mathbb{R}^2 \to \mathbb{R}^2$: How does the image grid look like? In particular, how does the rank show up? What is the area of the image of $[0, 1]^2$?

`38. Lecture, Tuesday, 15. May 07 ______________________ "U 4`

1. Differentiability for several variables

1.1. **The differential.** For the case of curves we could define differentiation and integration componentwise. The situation is very different when we go to functions whose preimage has several dimensions.

Let us recall the definition of differentiability for functions of one variable, $f\colon \mathbb{R} \to \mathbb{R}$:

$$f'(x) = \lim_{h\to 0} \frac{f(x+h) - f(x)}{h}$$

In case x and h are vectors, it is impossible to use the same formula: We cannot divide by a vector. Instead we will generalize the characterization of differentiability as linear approximability, Thm. II.1: A differentiable function satisfies

$$f(x+h) = f(x) + f'(x)h + r_x(h) \qquad \text{with} \quad \lim_{h\to 0} \frac{r_x(h)}{h} = 0. \tag{1}$$

For fixed x, the second term $h \mapsto L_x(h) := f'(x)h$ is a linear map $L\colon \mathbb{R} \to \mathbb{R}$, given by multiplication with the constant $f'(x) \in \mathbb{R}$. The affine mapping $x + h \mapsto f(x) + L_x(h)$ represents the tangent line to the graph of f at x.

In the case of a differentiable function of one variable, (1) requires that there is a good choice of $L = L_x$, such that the approximation of $f(x+h)$ by $f(x) + L_x(h)$ leaves an error decaying less than linear. This is literally our definition of differentiability for mappings of several variables:

Definition. A mapping $f\colon U \to \mathbb{R}^m$ is *differentiable at* $x \in U$ if there is a linear mapping $L_x\colon \mathbb{R}^n \to \mathbb{R}^m$ such that for all h with $x + h \in U$,

$$f(x+h) = f(x) + L_x(h) + r_x(h) \qquad \text{with} \quad \lim_{h\to 0} \frac{r_x(h)}{\|h\|} = 0. \tag{2}$$

Here $r_x\colon \{h \in \mathbb{R}^n : x+h \in U\} \to \mathbb{R}^m$ is called the *remainder term*. If (2) holds we call the linear map $L = L_x$ the *differential* of f at x, and write $df_x(h) := L_x(h)$ or $df(h) := L(h)$ for short. Finally, f is *differentiable* if the above holds for all $x \in U$.

It is not hard to show that the differential is uniquely defined (problems).

How can we visualize this definition?

• When $f\colon U \to \mathbb{R}$ is a scalar-valued function on $U \subset \mathbb{R}^n$, the graph of f is contained in $\mathbb{R}^{n+1}$. The image of the affine linear map $h \mapsto f(x) + L(h)$ is an affine hyperplane in $\mathbb{R}^{n+1}$ passing through the point $f(x)$. In particular, for $n = 2$ we can visualize the image as the *tangent plane* to the graph in $\mathbb{R}^3$.

• In the grid model, the affine grid of the mapping $h \mapsto f(x) + L_x(h)$ approximates the curved (non-linear) grid of $x \mapsto f(x)$ up to an error which decays better than linear.

We will discuss efficient means to calculate the differential in the next subsection. Still, we can employ our definition to deal with the simplest mappings:

Examples. 1. *Affine linear maps* $f\colon \mathbb{R}^n \to \mathbb{R}^m$, $f(x) := Ax+b$, where A is an $m\times n$-matrix and $b \in \mathbb{R}^m$. For sure, we expect that this mapping is its own affine approximation, with vanishing error. To check this, write

$$f(x+h) = A\cdot(x+h)+b = f(x)+Ah.$$

The linear term $L_x(h) = Ah$ must be the differential of f; it does not depend on the value of x. Hence (2) holds with a vanishing remainder term $r_x \equiv 0$. Note the analogy to the one-dimensional case: For $f\colon \mathbb{R}\to\mathbb{R}$, $f(x) = cx$ we have $f'(x) = c$, which we can deduce from the equation $c(x+h) = cx + ch$.

2. We consider the quadratic function $f\colon \mathbb{R}^2 \to \mathbb{R}$, $f(x,y) := x^2+y^2$. Its graph in $\mathbb{R}^3$ is the *paraboloid of revolution [Rotationsparaboloid]*. Denoting $h := (k,l)$ we find

$$f(x+k,\, y+l) = x^2+2xk+k^2 \;+\; y^2+2yl+l^2 = f(x,y) + 2\left\langle \begin{pmatrix} x\\ y\end{pmatrix}, \begin{pmatrix} k\\ l\end{pmatrix}\right\rangle + \left\| \begin{pmatrix} k\\ l\end{pmatrix}\right\|^2.$$

The scalar product is a linear functional, and so we set

$$L_{(x,y)}(k,l) :=\; 2\left\langle \begin{pmatrix} x\\ y\end{pmatrix}, \begin{pmatrix} k\\ l\end{pmatrix}\right\rangle \quad\Rightarrow\quad r_{(x,y)}(k,l) = \left\| \begin{pmatrix} k\\ l\end{pmatrix}\right\|^2.$$

Then indeed the decay condition (2) holds: $\frac{r_{(x,y)}(k,l)}{\|(k,l)\|} = \left\|\left(\begin{smallmatrix} k\\ l\end{smallmatrix}\right)\right\| \to 0$ as $h = \left(\begin{smallmatrix} k\\ l\end{smallmatrix}\right) \to 0$. This proves that f is differentiable with $df_{(x,y)} = L_{(x,y)}$. For instance, at $(x,y) = (0,0)$ the graph has a horizontal tangent plane, and the larger (x,y) gets (in modulus) the steeper the tangent plane becomes. Compare our calculation with the case of one variable!

3. (Problem:) Generalize the previous calculation to *quadratic forms*

$$Q\colon \mathbb{R}^n \to \mathbb{R}, \qquad Q(x) := x^\top A x \quad \Big(= \langle x, Ax\rangle = \sum_{i,j=1}^{n} a_{ij}x_i x_j \Big),$$

where A is a $n\times n$ matrix. Show that Q is continuous and compute the differential.

In Thm V.5 we asserted that the convergence of vector sequences is equivalent to componentwise convergence; similarly, by Thm. V.14 a vector valued function is continuous if all its components are continuous. Similarly so for differentiability:

Theorem 1. *A vector valued mapping $f\colon U \to \mathbb{R}^m$ is differentiable at $x \in U$ if and only if all its (scalar valued) component functions $f_1,\dots,f_m$ are differentiable at x.*

Proof. Note that the vector sequence $\frac{r(h)}{\|h\|}$ is null if and only if its component sequences are null (Thm. V.5). Hence for "$\Rightarrow$" we can restrict

$$f(x+h) = f(x) + df_x(h) + r_x(h) \qquad \text{with} \quad \lim_{h\to 0} \frac{r_x(h)}{\|h\|} = 0$$

to each of the m components. Conversely, for "$\Leftarrow$" we collect the m component differentials df_k to form the linear map df. Likewise for the remainder terms $r_k(h)$. Validity of the component equations then implies the vector equation. $\square$

If f is differentiable, then as $h \to 0$ we have $r_x(h) \to 0$. Consequently, $f(x+h) \to f(x)$, and so, as for one variable:

Theorem 2. *A mapping f which is differentiable at x is also continuous at x.*

1.2. **Partial derivatives and the Jacobian.** Just as Riemann sums are not useful to compute integrals, so is the definition of the differential df not adapted to computation. In the present section, we will restrict a function of a multidimensional domain to lines parallel to the coordinate directions, in order to use standard derivatives to compute df.

Definition. Let $f = \big(f_1, \dots, f_m\big) \colon U \to \mathbb{R}^m$ be a mapping. Then we call

$$\frac{\partial f_i}{\partial x_j}(x) := \partial_j f_i(x) := \frac{d}{dt} f_i(x + te_j)\Big|_{t=0} = \lim_{t\to 0} \frac{f_i(x+te_j) - f_i(x)}{t},$$

where $i \in \{1, \dots, m\}$ and $j \in \{1, \dots, n\}$, the j-th *partial derivative [partielle Ableitung]* of f_i at x. If all partial derivatives exist at all points of the domain, we call f *partially differentiable.*

Examples. 1. $f(x,y) := 1 + xy^3$. Then $\frac{\partial f}{\partial x}(x,y) = y^3$ and $\frac{\partial f}{\partial y}(x,y) = 3xy^2$.
2. Consider $f \colon \mathbb{C} \to \mathbb{C}$ with $f(z) = z^2 = (x+iy)^2 = x^2 - y^2 + 2ixy = \left(\begin{smallmatrix} x^2-y^2 \\ 2xy \end{smallmatrix}\right)$. It has four partial derivatives. Employing vector notation, these are:

$$\frac{\partial f}{\partial x} = \begin{pmatrix} 2x \\ 2y \end{pmatrix}, \qquad \frac{\partial f}{\partial y} = \begin{pmatrix} -2y \\ 2x \end{pmatrix}$$

Not alone for notational reasons it is convenient to collect the mn partial derivatives into a matrix or, in the special case of a scalar-valued function, into a vector:

Definition. Suppose $f \colon U \to \mathbb{R}^m$ is partially differentiable. Then for each $x \in U$ the *Jacobian [Jacobi-Matrix]* of f at x is the $m \times n$-matrix

$$J_f(x) := \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(x) & \cdots & \frac{\partial f_1}{\partial x_n}(x) \\ \vdots & & \vdots \\ \frac{\partial f_m}{\partial x_1}(x) & \cdots & \frac{\partial f_m}{\partial x_n}(x) \end{pmatrix}.$$

In the scalar valued case $f\colon U \to \mathbb{R}$, the Jacobian has only one row; then we call its transpose

$$\operatorname{grad} f\colon U \to \mathbb{R}^n, \qquad \operatorname{grad} f(x) := \Big(\frac{\partial f}{\partial x_1}(x), \cdots, \frac{\partial f}{\partial x_n}(x)\Big)^\top$$

the *gradient* of f at x; another notation is $\nabla f(x)$ (read: *nabla* f at x).

In one variable, $f'(x)$ is a number, and $f(x) + f'(x)h$ is the affine linear mapping approximating $f(x+h)$. Similarly, in several variables the Jacobian is matrix of numbers and, as we shall see, $J_f(x)h$ is the linear mapping approximating $f(x+h) - f(x)$. So it is fair to extend the notation f' to the Jacobian, as some authors do.

Examples. 1. For the first example above, $\operatorname{grad} f(x,y) = \left(\begin{smallmatrix} y^3 \\ 3xy^2 \end{smallmatrix}\right)$ and for the second, $J_f(x,y) = \left(\begin{smallmatrix} 2x & -2y \\ 2y & 2x \end{smallmatrix}\right)$.

2. For $f(x) := \|x\|^2 = x_1^2 + \ldots + x_n^2$ we have $\operatorname{grad} f(x) = (2x_1, \ldots, 2x_n)^\top = 2x$.

`39. Lecture, Tuesday, 22. May 07 ____________________ T 5, "U 5`

3. *Polar coordinates* $P\colon \mathbb{R}^2 \to \mathbb{R}^2$ with $P(r,\varphi) := \left(\begin{smallmatrix} P_1(r,\varphi) \\ P_2(r,\varphi) \end{smallmatrix}\right) := \left(\begin{smallmatrix} r\cos\varphi \\ r\sin\varphi \end{smallmatrix}\right)$ have the Jacobian

$$J_P(r,\varphi) = \begin{pmatrix} \frac{\partial P_1}{\partial r} & \frac{\partial P_1}{\partial \varphi} \\ \frac{\partial P_2}{\partial r} & \frac{\partial P_2}{\partial \varphi} \end{pmatrix}(r,\varphi) = \begin{pmatrix} \cos\varphi & -r\sin\varphi \\ \sin\varphi & r\cos\varphi \end{pmatrix}.$$

Visualize P using a grid model.

The matrix representing the differential as a linear map turns out to be precisely the Jacobian. That is, to compute the differential, we only need to compute partial derivatives.

Theorem 3. *If $f\colon U \to \mathbb{R}^m$ is differentiable at x then f is partially differentiable at x. In particular, the Jacobian represents the matrix of the differential,*

$$df_x(h) = J_f(x) \cdot h \qquad \text{for all } h \in \mathbb{R}^n. \tag{3}$$

In the scalar valued case, this specializes to $df_x(h) = \langle \operatorname{grad} f(x), h\rangle$.

Proof. We will establish for $j = 1, \ldots, n$ that

$$f(x + te_j) = f(x) + \frac{\partial f}{\partial x_j}(x)\, t + r_x(t) \qquad \text{with} \quad \lim_{t\to 0} \frac{r_x(t)}{t} = 0. \tag{4}$$

We can then specialize this vector equation to each of its m components. By the one-dimensional characterization of differentiability by linear approximability (Thm. V.1), this proves that $t \mapsto f_i(x + te_j)$ has a derivative at $t = 0$, that is, all partial derivatives $\frac{\partial f_i}{\partial x_j}(x)$ exist.

Our strategy to verify (4) is to specialize the definition of differentiability (2) to a coordinate direction. This gives us a function of one variable,

$$t \mapsto f(x+te_j) \overset{(2)}{=} f(x) + df_x(te_j) + r_x(te_j) \overset{df_x \text{ linear}}{=} f(x) + t\,df_x(e_j) + r_x(te_j);$$

since U is open, the equation is defined for all sufficiently small $|t|$. Moreover, (2) gives

$$\frac{1}{\|te_j\|} r_x(te_j) = \frac{1}{|t|} r_x(te_j) \to 0 \qquad \text{as} \quad \|te_j\| = |t| \to 0.$$

Thus, setting $\frac{\partial f}{\partial x_j}(x) := df_x(e_j)$ we have verified the remainder term decay (4).

It remains to prove (3). The linear mapping df_x is determined by its values on the standard basis, and so

$$\begin{aligned} df_x(h) &= df_x\big(h_1e_1 + \ldots + h_ne_n\big) \overset{df \text{ linear}}{=} h_1 df_x(e_1) + \ldots + h_n df_x(e_n) \\ &= \frac{\partial f}{\partial x_1}(x)\, h_1 + \ldots + \frac{\partial f}{\partial x_n}(x)\, h_n = \Big(\frac{\partial f}{\partial x_1}(x), \ldots, \frac{\partial f}{\partial x_n}(x)\Big) \cdot h = J_f \cdot h. \end{aligned}$$

□

Let us now study the converse of the preceding theorem: If f has all partial derivatives, is it then also differentiable? This is not generally true! Indeed, the existence of the Jacobian J_f alone will not imply the that the remainder term $r(h) = f(x+h) - f(x) - J_f(x) \cdot h$ has sublinear decay (2):

Counterexample. Let $f\colon \mathbb{R}^2 \to \mathbb{R}$ be defined by

$$f(x,y) := \frac{x^2y}{x^2+y^2} \text{ for } (x,y) \neq 0 \quad \text{and} \quad f(0,0) := 0.$$

Then $f(tx,ty) = tf(x,y)$ for all $t \in \mathbb{R}$, that is, the graph of f consists of lines through the origin.

1st claim: f has partial derivatives with respect to x and y.

Away from 0, they can be computed using the differentiation rules. Moreover, $f(x,0) \equiv 0$ for all $x \in \mathbb{R}$, and $f(0,y) \equiv 0$ for all $y \in \mathbb{R}$. Therefore, the partial derivatives also exist at $(0,0)$, namely

(5) $$\frac{\partial f}{\partial x}(0,0) = \frac{\partial f}{\partial y}(0,0) = 0.$$

2nd claim: f is not differentiable at $(0,0)$.

Let us assume the contrary. Then, by Theorem 3 and (5), we have $df_{(0,0)} \equiv 0$. But $f(t,t) = \frac{t}{2}$ for $t \in \mathbb{R}$, and so

$$r_{(0,0)}(t,t) = f(t,t) - f(0,0) - df_{(0,0)}(t,t) = \frac{t}{2} \quad \Longrightarrow \quad \frac{r_{(0,0)}(t,t)}{\|(t,t)\|} = \frac{t}{2\sqrt{2}\,|t|} \in \{\pm\frac{1}{2\sqrt{2}}\}.$$

Hence (2) is not satisfied, showing that f cannot be differentiable at $(0,0)$.

If, however, the Jacobian $J_f(x)$ depends continuously on x, the mapping f is differentiable:

Theorem 4. *Assume $f\colon U \to \mathbb{R}^m$ is* continuously (partially) differentiable*, that is, all partial dervatives $\frac{\partial f_i}{\partial x_j}$ are continuous functions on U. Then f is differentiable with differential $df_x(h) = J_f(x)\cdot h$.*

Proof. Let us first assume that f is scalar valued, $f\colon U \to \mathbb{R}$. We set $df_x(h) = J_f(x)h$ and need to verify that the remainder term $r(h) = f(x+h) - f(x) - J_f(x)h$ has sublinear decay. To do this, we represent $f(x+h) - f(x)$ by differences over coordinate directions, to which we can apply the Mean Value Theorem of Differentiation.

For $x \in U$ and $h \in \mathbb{R}^n$, consider the $n+1$ points

$$x^{(1)} := x, \quad x^{(2)} := x + (h_1, 0, \dots, 0) \quad x^{(3)} := x + (h_1, h_2, 0, \dots, 0), \quad \dots, \quad x^{(n+1)} := x + h.$$

Since U is open some ball $B_r(x)$ is contained in U. Thus assuming $\|h\| < r$, the ball $B_r(x)$ contains the above points, as well as the linear segments between successive points. Let us now construct n numbers $\xi_j \in [-|h_j|, |h_j|]$ for $j = 1, \dots, n$, satisfying

$$f(x^{(2)}) - f(x^{(1)}) = \frac{\partial f}{\partial x_1}(x^{(1)} + \xi_1 e_1)\, h_1, \quad \dots, \quad f(x^{(n+1)}) - f(x^{(n)}) = \frac{\partial f}{\partial x_n}(x^{(n)} + \xi_n e_n)\, h_n.$$

In case $h_j \neq 0$, the existence of ξ_j is guaranteed by the Mean Value Theorem IV.8, applied to $t \mapsto f(x^{(j)} + te_j)$ on the appropriate interval. In case $h_j = 0$ we have $x^{(j)} = x^{(j+1)}$ and so the respective equation trivially holds for $\xi_j := 0$.

Summing, we obtain the telescope sum

$$\begin{aligned} f(x+h) - f(x) &= f\big(x^{(n)}\big) - f\big(x^{(n-1)}\big) + \ \dots \ + f\big(x^{(1)}\big) - f\big(x^{(0)}\big) \\ &= \frac{\partial f}{\partial x_n}\big(x^{(n)} + \xi_n e_n\big)\, h_n + \dots + \frac{\partial f}{\partial x_1}\big(x^{(1)} + \xi_1 e_1\big)\, h_1. \end{aligned}$$

Hence setting $df_x(h) := J_f(x)\cdot h$, the remainder term has the representation

$$(6) \quad \frac{r_x(h)}{\|h\|} = \frac{1}{\|h\|}\Big(f(x+h) - f(x) - J_f(x)\cdot h\Big) = \sum_{j=1}^{n}\Big(\frac{\partial f}{\partial x_j}(x^{(j)} + \xi_j e_j) - \frac{\partial f}{\partial x_j}(x)\Big)\frac{h_j}{\|h\|}.$$

Note that $\frac{h_j}{\|h\|}$ is bounded, namely $\frac{|h_j|}{\|h\|} \leq 1$.

Let us now consider the limit of (6). As $h \to 0$ we have $\xi_j \to 0$ and $x^{(j)} \to x$. Thus our assumption that the partial derivatives are continuous at x implies

$$\lim_{h\to 0}\Big(\frac{\partial f}{\partial x_j}(x^{(j)} + \xi_j e_j) - \frac{\partial f}{\partial x_j}(x)\Big) = 0.$$

This means, as $h \to 0$, each term under the sum in (6) is of the type "null sequence times bounded sequence". By Lemma V.8 the sum is a null sequence, and so $\lim_{h\to 0} \frac{r_x(h)}{\|h\|} = 0$. This means f is indeed differentiable with differential (3).

In case f is vector valued, we apply our proof to each component; in view of Thm. 1 we conclude f is differentiable. $\square$

The converse of the theorem does not hold: In fact, already in one dimension the function $x^2 \sin\frac{1}{x}$ (and 0 at 0) is differentiable, but is not continuously differentiable at 0 (check!).

Thus, we can summarize the previous two theorems to say

continuously partially differentiable $\Rightarrow$ differentiable $\Rightarrow$ partially differentiable.

We want to list rules for differentiation:

Theorem 5. *Assume $f, g\colon U \to \mathbb{R}^m$ are differentiable. Then the differentiation is linear, that is, at each $x \in U$ we have, for $\lambda \in \mathbb{R}$,*

$$d(\lambda f + g)_x = \lambda df_x + dg_x.$$

Moreover, in the scalar-valued case, $m = 1$, also the product fg is differentiable, and likewise the reciprocal of f, provided $f(x) \neq 0$:

$$\operatorname{grad}(fg) = (\operatorname{grad} f)g + f \operatorname{grad} g \qquad \textit{(Product Rule)}, \qquad \operatorname{grad}\frac{1}{f} = -\frac{\operatorname{grad} f}{f^2}. \tag{7}$$

Proof. We will only give a proof under the additional assumption that f, g have continuous partial derivatives. To prove the Product Rule we apply its one-variable counterpart to obtain

$$\begin{aligned}\operatorname{grad}(fg) &= \Big(\frac{\partial}{\partial x_1}(fg), \dots, \frac{\partial}{\partial x_n}(fg)\Big)^\top \\ &= \Big(\frac{\partial f}{\partial x_1} g + f\frac{\partial g}{\partial x_1}, \dots, \frac{\partial f}{\partial x_n} g + f\frac{\partial g}{\partial x_n}\Big)^\top = (\operatorname{grad} f)g + f \operatorname{grad} g\end{aligned}$$

Since this is continuous, it actually represents the differential of fg. Similarly for the other rules. $\square$

Problems. 1. Give the proof of the theorem in the general case, that is, without refering to partial derivatives: You need to show that the remainder terms have the required decay.

2. Differentiate $\frac{f}{g}$ and $\langle f, g\rangle$.

`40. Lecture, Thursday, 24. May 07 ____________________ T 6`

1.3. **The Chain Rule.** The chain rule presents the only rule of differentiation which is not an immediate consequence of the theory of one variable. Suppose two differentiable maps f, g can be composed. What is the differential of $f \circ g$, i.e., what is the linear approximation to $f \circ g$?

Let us look at the linear case first. Suppose $f(x) = Ax$ for a matrix A and $g(x) = Bx$ for some matrix B. Then the composed linear map satisfies $(f \circ g)(x) = A(Bx) = (AB)x$, that is, its matrix is the product matrix $(AB)_{ij} = \sum_k a_{ik} b_{kj}$.

Now let us go on to the derivative level. Linear maps are reproduced, and so $df_y(v) = Av$, $dg_x(w) = Bw$ hold independently of x, y. Likewise, the composed map must have the differential $d(f \circ g)_x = AB$. Consequently the Jacobian of the composition agrees with the matrix product of the Jacobians.

This result continues to hold in the nonlinear case:

Theorem 6 (Chain Rule)**.** *Suppose $U \subset \mathbb{R}^n$ and $V \subset \mathbb{R}^m$ are open, and let $g\colon U \to V$ and $f\colon V \to \mathbb{R}^\ell$. If g is differentiable at x, and f at $y = g(x)$, then the composed map $f \circ g$ is differentiable at x with differential*

$$d(f \circ g)_x = df_y \circ dg_x.$$

That is, the Jacobian of $f \circ g$ is given by the product of Jacobians,

$$J_{f \circ g}(x) = J_f(y) \cdot J_g(x).$$

The chain rule can be taken as evidence of how powerful the idea of linearization is. If g is linearized by dg, and f by df, then we the composition $df \circ dg$ is once again linear, and so we could expect rightaway that it is the linearization of $f \circ g$.

Before we give the proof, let us study examples for the chain rule.

Examples. 1. Let $\mathbb{R}^2 \xrightarrow{g} \mathbb{R}^3 \xrightarrow{f} \mathbb{R}$ with $g(x) := (x_1 x_2,\, x_1^2,\, x_2)$ and $f(y) := y_1 + y_2^2 + y_3^3$. Then the Jacobian of $f \circ g\colon \mathbb{R}^2 \to \mathbb{R}$ is

$$J_{f \circ g}(x) = J_f\big(g(x)\big) \cdot J_g(x) = \Big(1,\, \underbrace{2g_2(x)}_{=2x_1^2},\, \underbrace{3g_3(x)^2}_{=3x_2^2}\Big) \cdot \begin{pmatrix} x_2 & x_1 \\ 2x_1 & 0 \\ 0 & 1 \end{pmatrix} = \Big(x_2 + 4x_1^3,\; x_1 + 3x_2^2\Big).$$

2. Often, the two composed functions are not explicit. For instance, one seeks the x-derivative of some function $x \mapsto f\big(x, y, h(x, y)\big)$. We must view this as the composition $f \circ g$ where $g(x, y) := \big(x, y, h(x, y)\big)$. Then we are interested in the first partial derivative

of the composition $\partial_1(f\circ g)$. For that end, only the first column of the 3×2-Jacobian of g is relevant. We obtain

$$\partial_1(f\circ g) = (\partial_1 f, \partial_2 f, \partial_3 f)\cdot \begin{pmatrix} 1 \\ 0 \\ \partial_1 h \end{pmatrix} = \partial_1 f + \partial_3 f\,\partial_1 h. \tag{8}$$

Slightly abusing notation, the left hand side is often written as $\frac{df}{dx}$, which is a notation that avoids refering to g.

The matrix product constituting the chain rule can also be written in summation form:

$$\frac{\partial (f\circ g)_i}{\partial x_j} = \sum_{k=1}^m \frac{\partial f_i}{\partial y_k}\circ g\,\frac{\partial g_k}{\partial x_j}. \qquad \text{for } 1\le i\le \ell,\ \ 1\le j\le n. \tag{9}$$

Physicists like to set $y := g(x)$ which they use not only for the function $y(x)$ but also as a symbol for the variables of f. Then they write

$$\frac{\partial f_i}{\partial x_j} = \sum_{k=1}^m \frac{\partial f_i}{\partial y_k}\frac{\partial y_k}{\partial x_j} \qquad \text{or} \qquad \frac{\partial f_i}{\partial x_j} = \frac{\partial f_i}{\partial y_k}\frac{\partial y_k}{\partial x_j},$$

where in the last version the *summation convention* is in force: sum over repeated indices. Notationally, this way the chain rule appears as an expansion of a fraction, similar to the case of one variable.

Example. Consider $t\mapsto f(th)$ where $f\colon \mathbb{R}^n\to\mathbb{R}$ and $h\in\mathbb{R}^n$. Let us employ (9) to compute the second derivative:

$$\frac{d^2}{dt^2}f(th) = \Big(\sum_{k=1}^n \partial_i f(th)\,h_k\Big)' = \sum_{k=1}^n\sum_{\ell=1}^n \partial_\ell\partial_k f(th)\,h_k h_\ell$$

For the proof of the chain rule, it is convenient to introduce:

Definition. Let $L\colon\mathbb{R}^m\to\mathbb{R}^n$ be a linear map, $L(x) = A\cdot x$ for A an $m\times n$-matrix. Then

$$\|L\| := \|A\| := \sup\big\{\|L(x)\| = \|A(x)\| : \|x\|\le 1\big\}. \tag{10}$$

is called the *operator norm* of L or the *matrix norm* of A.

By Thm. V.19 the function $\|Ax\|$ takes a maximum on the compact subset $\overline{B_1}(0)\ni x$. Therefore $\|A\|$ is well-defined for each matrix A.

For $x\neq 0$ we have

$$\frac{1}{\|x\|}\|A\cdot x\| \overset{A \text{ linear}}{=} \Big\|A\cdot\frac{x}{\|x\|}\Big\| \le \|A\|.$$

This gives the important estimate for a matrix product

$$\|A\cdot x\|\le \|A\|\|x\|. \tag{11}$$

Problems. 1. Show $\|A\| \leq mn \max\{|a_{ij}| : 1 \leq i \leq m,\ 1 \leq j \leq n\}$. This verifies that $\|A\|$ is well-defined by direct calculation.
2. Check that $\|A\|$ is indeed a norm on the space of matrices.
3. If A has only one row or one column, then check the norm of A agrees with the standard norm of the row or column vector, respectively.

Proof. For $d(f \circ g) := df \circ dg$ we need to verify the remainder term estimate

$$f\big(g(x+h)\big) = f\big(g(x)\big) + df_y\big(dg_x(h)\big) + R(h) \qquad \text{with} \quad \lim_{h\to 0} \frac{R(h)}{\|h\|} = 0. \tag{12}$$

We must use our differentiability assumptions on f and g, which we rewrite as

$$g(x+h) = g(x) + dg_x(h) + \|h\| r_1(h) \qquad \text{with} \quad \lim_{h\to 0} r_1(h) = 0, \tag{13}$$

$$f(y+k) = f(y) + df_y(k) + \|k\| r_2(k) \qquad \text{with} \quad \lim_{k\to 0} r_2(k) = 0. \tag{14}$$

Using the shorthand notation

$$k(h) := dg_x(h) + \|h\| r_1(h)$$

we find the following representation

$$\begin{aligned} f\big(g(x+h)\big) &\overset{(13)}{=} f\Big(g(x) + dg_x(h) + \|h\| r_1(h)\Big) \\ &\overset{(14)}{=} f\big(g(x)\big) + df_y\Big(dg_x(h) + \|h\| r_1(h)\Big) + \|k(h)\| r_2\big(k(h)\big) \\ &\overset{df \text{ linear}}{=} f\big(g(x)\big) + df_y\big(dg_x(h)\big) + \underbrace{\|h\| df_y\big(r_1(h)\big)}_{=:I(h)} + \underbrace{\|k(h)\| r_2\big(k(h)\big)}_{=:II(h)}. \end{aligned}$$

Comparing this equation with (12) we find that $R(h) = I(h) + II(h)$. We now complete the proof by verifying the remainder term estimate $\lim_{h\to 0} \frac{R(h)}{\|h\|} = 0$ seperately for I and II.

For I, we need to show $\lim_{h\to 0} df_y\big(r_1(h)\big) = 0$. But, as a linear mapping, df_y is continuous and thus

$$\lim_{h\to 0} df_y\big(r_1(h)\big) \overset{df \text{ cts.}}{=} df_y\big(\lim_{h\to 0} r_1(h)\big) \overset{(14)}{=} df_y(0) = 0.$$

For II, let us first invoke the matrix norm estimate (11) on dg_x, to obtain

$$\big\|k(h)\big\| \leq \|dg_x(h)\| + \|h\| \|r_1(h)\| \overset{(11)}{\leq} \|h\| \big(\|dg_x\| + \|r_1(h)\|\big).$$

We conclude $\lim_{h\to 0} \|k(h)\| = 0$ and moreover

$$\lim_{h\to 0} \frac{\|k(h)\|}{\|h\|} \|r_2(k(h))\| \leq \lim_{h\to 0} \Bigg(\Big(\|dg_x\| + \|r_1(h)\|\Big) \big\|r_2(k(h))\big\|\Bigg). \tag{15}$$

Since $\|r_1(h)\| \to 0$, we have $\|r_1(h)\| \leq 1$ for h small enough. Thus $\|dg_x\| + \|r_1(h)\|$ is bounded. Moreover, since $\|k(h)\| \to 0$, also $\|r_2\big(k(h)\big)\| \to 0$. Thus (15) involves the product of a bounded sequence with a null sequence. By Lemma V.8 the limit (15) vanishes. □

`41. Lecture, Tuesday, 29. May 07`

1.4. **Directional derivative. Meaning of differential and gradient.** More general than partial derivatives are derivatives in an arbitrary direction of space:

Definition. A mapping $f\colon U \to \mathbb{R}^m$ has a *directional derivative [Richtungsableitung]* in direction $v \in \mathbb{R}^n$ at the point $x \in U$ if

$$D_v f(x) := \frac{d}{dt} f(x+tv)\Big|_{t=0} = \lim_{t\to 0} \frac{f(x+tv)-f(x)}{t} \in \mathbb{R}^m$$

exists.

Examples. 1. The linear function $f(x,y) := x$ has directional derivative

$$D_v f(x,y) = \frac{d}{dt} f(x+tv)_{t=0} = \frac{d}{dt}\big[x + tv_1\big]_{t=0} = v_1.$$

In particular, if we consider a unit vector $v(\alpha) := (\cos\alpha, \sin\alpha)$ then $D_{v(\alpha)} f(x,y) = \cos\alpha$. Hence the directional derivative vanishes when $\alpha = \pm\frac{\pi}{2}$, and it is maximal for $\alpha = 0$.

2. The directional derivatives with respect to the vectors of a standard basis $v = e_j$ are precisely the partial derivatives, $D_{e_j} f(x) = \frac{\partial f}{\partial x_j}(x)$.

We can interpret $df_x(h)$ as the directional derivative of f in direction h:

Theorem 7. *If f is differentiable at x, then for each $v \in \mathbb{R}^n$*

$$D_v f(x) = J_f(x)\cdot v = df_x(v).$$

Proof. Setting $c(t) = x + tv$ we can write $f(x+tv) = (f\circ c)(t)$. Here, c has the constant tangent $c' \equiv v$; this column vector is also the Jacobi matrix J_c. Hence the Chain Rule gives

$$D_v f(x) = (f\circ c)'(0) \overset{\text{Chain Rule}}{=} J_f\big(c(0)\big)\cdot c'(0) = J_f\big(c(0)\big)\cdot v.$$

□

The theorem gives rise to the following interpretation of the differential. Consider an arbitrary differentiable curve $c(t)$ in the domain U; its tangent vector is $c'(t)$. Then the image curve $f\big(c(t)\big)$ has the tangent vector

$$\frac{d}{dt}(f\circ c)(t) = df_{c(t)}\big(c'(t)\big).$$

That is, the differential maps the tangent vector of a curve to the tangent vector of the image curve. Or, in more condensed form: The differential df is the unique and linear map which maps tangent vectors to tangent vectors.

Like the partial derivatives they generalize, directional derivatives can still exist when f fails to be differentiable. The example of Sect. 1.2 serves as an example for this fact as well.

Let us now endow the gradient of function $f\colon U \to \mathbb{R}$ with a concrete meaning. Recall that a pair of vectors $v, w \in \mathbb{R}^n \setminus \{0\}$ encloses an angle $\angle(v,w) \in [0,\pi]$ provided $\langle v, w\rangle = \|v\|\|w\| \cos\angle(v,w)$. Thus for a unit direction v with $\|v\| = 1$,

$$D_v f(x) = df_x(v) = \langle \operatorname{grad} f(x), v\rangle = \|\operatorname{grad} f(x)\| \cdot \cos \angle\big(v, \operatorname{grad} f(x)\big),$$

where we assume that $v, \operatorname{grad} f \neq 0$. Now let us vary v over $\mathbb{S}^{n-1} = \{v : \|v\| = 1\}$:
• The maximal value of $D_v f(x)$ is attained for $v = \frac{\operatorname{grad} f(x)}{\|\operatorname{grad} f(x)\|}$. That is, the gradient points in the direction of the steepest ascent *[Anstieg]* of f at x.
• The minimal value of $D_v f(x)$ is attained for the negative, $v = -\frac{\operatorname{grad} f(x)}{\|\operatorname{grad} f(x)\|}$, that is, $-\operatorname{grad} f$ points to the steepest descent *[Abstieg]*.
• Any direction $v \perp \operatorname{grad} f$ will have vanishing directional derivative, that is, f is constant in first order.
Let us add that the length of the gradient vector, $\|\operatorname{grad} f(x)\|$, is a measure of the steepness.

For $f\colon U \to \mathbb{R}$ differentiable, consider the *level set [Niveaumenge]*

$$N_y := \{x \in U : f(x) = y\}.$$

Suppose that $c\colon I \to N_y \subset U$ is a differentiable curve. Then $f \circ c \equiv y$ and so

$$0 = \frac{d}{dt}(f \circ c) \overset{\text{Chain Rule}}{=} \big\langle \operatorname{grad} f(c(t)), c'(t)\big\rangle \qquad \Longleftrightarrow \qquad \operatorname{grad} f\big(c(t)\big) \perp c'(t).$$

Since this holds for any differentiable curve running through N_y, we can say that the gradient vector of f is perpendicular to the level sets N_y.

1.5. **Mean value theorem and a bound in terms of the differential.** The technically most important result in the theory of differentiation in one variable is the mean value theorem. Thus it presents no surprise that we will eventually need the same statement in several variables. We will apply it to prove the inverse mapping theorem, in Section 1 below.

Let us introduce some terminology first. Given a pair of points $a, b \in \mathbb{R}^n$, we call the straight connection $\overline{ab} = \{x \in \mathbb{R}^n : x = a + t(b-a) \text{ with } t \in (0,1)\}$ a *segment [Strecke]*. A subset $M \subset \mathbb{R}^n$ is called *convex* if for each $a, b \in M$ also $\overline{ab} \subset M$.

Proposition 8 (Mean Value Theorem). *Let $f\colon U \to \mathbb{R}$ be differentiable. Moreover, suppose that for $a, b \in U$ the entire segment $\overline{ab}$ is contained in U. Then there is $p \in \overline{ab}$ such that*

$$f(b) - f(a) = \big\langle \operatorname{grad} f(p),\, b - a \big\rangle.$$

Proof. Consider the curve $c(t) := a + t(b-a)$ for $0 \le t \le 1$. We apply the Mean Value Theorem of Differentiation (Thm. IV.8) to $F := f \circ c\colon [0,1] \to \mathbb{R}$ to find $\xi \in (0,1)$ with

$$f(b) - f(a) = F(1) - F(0) \overset{\text{MVT}}{=} \frac{dF}{dt}(\xi) = \big(f \circ c\big)'(\xi) \overset{\text{Chain R.}}{=} \big\langle \operatorname{grad} f\big(c(\xi)\big), c'(\xi) \big\rangle.$$

Setting $p := c(\xi)$ and noting $c'(t) \equiv b - a$, the result follows. □

For vector-valued target, examples indicate that the obvious generalization $f(b) - f(a) = df_p(b-a)$ cannot hold in general (problems?). But still, as in one dimension (Thm. IV.9), we can employ the Mean Value Theorem to derive a bound on a function in terms of its Jacobian *[Schrankensatz]*:

Proposition 9. *Suppose $f\colon U \to \mathbb{R}^m$ is differentiable, with uniformly bounded operator norm $\|df_x\| \le C$ for some $C \ge 0$ and all $x \in U$. Then, for all a, b such that $\overline{ab} \subset U$,*

$$\|f(b) - f(a)\| \le \sqrt{n}\, C \|b - a\|. \tag{16}$$

Note that over compact subsets of U, some bound on the operator norm is automatic.

Proof. Let us first consider the case of a scalar function, $m = 1$. We apply the Mean Value Theorem to find $p \in \overline{ab}$ such that the following holds:

$$|f(b) - f(a)| = \big|\big\langle \operatorname{grad} f(p),\, b - a \big\rangle\big| \overset{\text{Schwarz}}{\le} \big\|\operatorname{grad} f(p)\big\| \|b - a\| \tag{17}$$

Now we consider general m. Let us first note that the j-th column of the Jacobi matrix is $df_x(e_j)$. Invoking the operator norm gives

$$\sum_{i=1}^m \left|\frac{\partial f_i}{\partial x_j}(x)\right|^2 = \|df_x(e_j)\|^2 \le \|df_x\|^2 \le C^2. \tag{18}$$

We now apply the Mean Value Theorem to each component to obtain m points $p_i \in \overline{ab}$ such that the following holds:

$$\begin{aligned}\|f(b) - f(a)\|^2 &= \sum_{i=1}^m \big(f_i(b) - f_i(a)\big)^2 \overset{(17)}{\le} \|b-a\|^2 \sum_{i=1}^m \big\|\operatorname{grad} f_i(p_i)\big\|^2 \\ &\le \|b-a\|^2 \sum_{j=1}^n \sum_{i=1}^m \left|\frac{\partial f_i}{\partial x_j}(p_i)\right|^2 \overset{(18)}{\le} \|b-a\|^2\, n\, C^2\end{aligned}$$

Taking the root yields the desired result. □

2. Extrema of scalar valued functions

Sufficient and necessary conditions for extrema in one variable are, for instance, the following:

$$f'(x) = 0,\ f''(x) < 0 \quad \Rightarrow \quad f \text{ attains a local max at } x \quad \Rightarrow \quad f'(x) = 0,\ f''(x) \le 0$$

Our derivation of these conditions depended heavily on monotonicity arguments, that is, we appealed to the order of the domain.

Let us indicate this for the sufficient condition, that is the left of the above implications. We still have $f''(x+h) < 0$ for $|h|$ small, and thus $h \mapsto f'(x+h)$ is monotonically decreasing in a neighbourhood of x. Since $f'(x) = 0$ it follows that f' changes sign at x. Therefore, f increases montonically up to x, and decreases thereafter, so that x must be a maximum. Recall the details from Part IV, Thm. 12, Lemma 13, and Thm. 14!

The monotonicity arguments cannot generalize to the case of several variables. Hence we will follow a different track. Let us indicate our approach in the one variable case for the sufficient condition. By the Taylor formula, $f(x+h) = f(x) + f'(x)h + \frac{1}{2}f''(x)h^2 + o(h^2)$. At a critical point x, the right hand side becomes $f(x)+\frac{1}{2}f''(x)h^2+o(h^2)$. Hence $f''(x) < 0$ implies $f(x) > f(x+h)$ for $|h| \neq 0$ sufficiently small, and so x must be a maximum.

In order to generalize the sufficient condition $f''(x) < 0$ to several variables, we will first discuss higher derivatives. Then we will state Taylor's formula for several variables, and employ it to generalize the argument we outlined.

`42. Lecture, Thursday, 31. May 07` ______________________________ `T 7`

2.1. **Higher derivatives and the Theorem of Schwarz.** A function $f\colon U \to \mathbb{R}$ is *twice partially differentiable* at $a \in U$, if

$$\partial_i \partial_j f(a) := \frac{\partial^2 f}{\partial x_i \partial x_j}(a) := \frac{\partial}{\partial x_i}\Big(\frac{\partial f}{\partial x_j}\Big)(a)$$

exists for all $1 \le i, j \le n$. Recursively, we define *partial derivatives of k-th order*:

$$\partial_{i_k} \cdots \partial_{i_1} f(a) := \frac{\partial^k f}{\partial x_{i_k} \cdots \partial x_{i_1}}(a) := \frac{\partial}{\partial x_{i_k}}\Big(\frac{\partial}{\partial x_{i_{k-1}}} \cdots \frac{\partial}{\partial x_{i_1}} f\Big)(a)$$

Let $k \in \mathbb{N}_0$. If $f\colon U \to \mathbb{R}^m$ is k-times partially differentiable such that all k-th partial derivatives are continuous, we write $f \in C^k(U, \mathbb{R}^m)$; we also say f is a C^k function.

There are n^k partial derivatives of k-th order. But many of these coincide, as the order of differentiation is irrelevant:

Theorem 10 (Schwarz)**.** *Suppose $f \in C^2(U, \mathbb{R})$, that is, $f\colon U \to \mathbb{R}$ has continuous second partial derivatives. Then*

$$\partial_i \partial_j f(a) = \partial_j \partial_i f(a) \qquad \text{at each point } a \in U.$$

Examples. 1. If $f(x, y) = x^k y^l$, then

$$\partial_1 f = k x^{k-1} y^l, \qquad \text{and} \quad \partial_2 \partial_1 f = kl x^{k-1} y^{l-1},$$

as well as

$$\partial_2 f = l x^k y^{l-1} \qquad \text{and} \quad \partial_1 \partial_2 f = kl x^{k-1} y^{l-1}.$$

By linearity, the Schwarz theorem holds for polynomials.

2. On the other hand, for the function $f(x, y) := xy \frac{x^2 - y^2}{x^2 + y^2}$ for $(x, y) \neq 0$ and $f(0, 0) := 0$ we have $\partial_1 \partial_2 f(0, 0) \neq \partial_2 \partial_1 f(0, 0)$, where $\partial^2 f$ is dicontinuous. This function indicates that the continuity assumption is necessary for the Schwarz theorem to hold. (Problems?)

Let us also give an interesting application of the Schwarz lemma: We claim that $(2y, x)$ is a vector field on $\mathbb{R}^2$, such that no $f\colon \mathbb{R}^2 \to \mathbb{R}$ with $\operatorname{grad} f(x, y) = (2y, x)$ exists (such an f is called *potential*). Indeed, if this were so, then $\partial_1 \partial_2 f = \partial_1 x = 1$, but on the other hand $\partial_2 \partial_1 f = \partial_2 (2y) = 2$, contradicting the Schwarz lemma.

Proof. There is no loss of generality in assuming $n = 2$ and $a = 0$. Since U is open, the rectangle $[0, x] \times [0, y]$ is contained in U for sufficiently small $x > 0$ and $y > 0$. We claim that there are points (ξ, η) and $(\tilde{\xi}, \tilde{\eta})$ in the rectangle, that is,

$$0 < \xi, \tilde{\xi} < x \qquad \text{as well as} \qquad 0 < \eta, \tilde{\eta} < y, \tag{19}$$

such that

$$\frac{1}{xy} \big(f(x, y) - f(x, 0) - f(0, y) + f(0, 0) \big) = \partial_2 \partial_1 f(\xi, \eta) = \partial_1 \partial_2 f(\tilde{\xi}, \tilde{\eta}). \tag{20}$$

Let us now apply the claim to a sequence of points $(x, y) \to 0$, where $x > 0$ and $y > 0$. From (19) follows $(\xi, \eta) \to 0$ and $(\tilde{\xi}, \tilde{\eta}) \to 0$. Since $\partial_1 \partial_2 f$ and $\partial_2 \partial_1 f$ are continuous functions at 0, the two expressions on the right hand side of (20) approach the common limit $\partial_2 \partial_1 f(0, 0) = \partial_1 \partial_2 f(0, 0)$, as desired.

We use the Mean Value Theorem of Differentiation for one variable (MVT) to establish the claim, that is, to exhibit (ξ, η) and $(\tilde{\xi}, \tilde{\eta})$. Considering y as fixed, the MVT gives the existence of ξ as in (19) with

$$\underbrace{f(x, y) - f(x, 0)}_{=: F_y(x)} \underbrace{- f(0, y) + f(0, 0)}_{= -F_y(0)} = x F_y'(\xi).$$

A further application of the MVT then gives η as in (19) with

$$F_y'(\xi) = \partial_1 \big(f(\xi, y) - f(\xi, 0) \big) = \partial_1 f(\xi, y) - \partial_1 f(\xi, 0) = y \, \partial_2 (\partial_1 f)(\xi, \eta).$$

This establishes the first equality in (20).

Working in the converse order, we obtain the existence of $(\tilde{\xi}, \tilde{\eta})$ subject to (19):

$$\begin{aligned}\underbrace{f(x,y) - f(0,y)}_{=:G_x(y)}\underbrace{-f(x,0) + f(0,0)}_{=-G_x(0)} &= y\, G'_x(\tilde{\eta}) \\ &= y\big(\partial_2 f(x,\tilde{\eta}) - \partial_2 f(0,\tilde{\eta})\big) = xy\partial_1(\partial_2 f)(\tilde{\xi},\tilde{\eta}).\end{aligned}$$

This gives the second equation in (20). □

2.2. **Taylor's formula.** Taylor's formula serves to approximate a function by polynomials. In most applications, the approximation includes the linear or the quadratic terms, but does not go beyond. The coefficients of the Taylor series are determined as partial derivatives of the function, similar to the one-variable case.

Consider $f\colon U \to \mathbb{R}$ which is k-times continously differentiable. Fix $x \in U$ and $h \in \mathbb{R}^n$, such that $x + th \in U$ for all $t \in [0,1]$. By openness of U this holds for h sufficiently small. Now we apply the one-dimensional Taylor formula to

$$F\colon [0,1] \to \mathbb{R}, \qquad F(t) := f(x+th),$$

that is, to the restriction of f to the segment $\overline{x, x+h}$. By the Lagrangian form of Taylor's Formula for one variable (Thm. IV.36) there exists $\xi \in [0,1]$ with

$$F(t) = \sum_{j=0}^{k-1} \frac{F^{(j)}(0)}{j!} t^j + \frac{F^{(k)}(\xi)}{k!} t^k. \tag{21}$$

This expresses F in terms of j-th directional derivatives.

We now use the summation form (9) of the Chain rule to calculate the j-th derivatives of $F^{(j)}(t)$ in (21). First,

$$F'(t) = \frac{d}{dt} f(x+th) = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(x+th) \frac{\partial (x+th)_i}{\partial t} = \sum_{i=1}^{n} \partial_i f(x+th)\, h_i,$$

which agrees with $\langle \operatorname{grad} f(x+th),\, h\rangle$. Similarly,

$$F''(t) = \frac{d}{dt}\Big(\sum_{i=1}^{n} \partial_i f(x+th)\, h_i\Big) = \sum_{j=1}^{n}\sum_{i=1}^{n} \partial_j \partial_i f(x+th)\, h_i h_j. \tag{22}$$

Continuing, we can inductively express the k-th directional derivative of f in terms of partial derivatives:

$$\begin{aligned} F^{(k)}(t) &= \frac{d}{dt}F^{(k-1)}(t) = \frac{d}{dt}\Bigg(\sum_{i_1=1}^{n}\cdots\sum_{i_{k-1}=1}^{n}\partial_{i_1}\ldots\partial_{i_{k-1}}f(x+th)\,h_{i_1}\ldots h_{i_{k-1}}\Bigg) \\ &\overset{\text{Chain Rule}}{=} \sum_{i_1=1}^{n}\cdots\sum_{i_k=1}^{n}\partial_{i_1}\ldots\partial_{i_k}f(x+th)\,h_{i_1}\ldots h_{i_k} \end{aligned}$$

To rewrite this formula, let us introduce a more condensed notation. For $k \in \mathbb{N}$, a vector of indices

$$I = (i_1, \ldots, i_k) \in \{1, \ldots, n\}^k$$

is called a *multiindex*; we usually abreviate $I = i_1 \ldots i_k$. Then we let

$$\partial_I f := \partial_{i_1}\cdots\partial_{i_k} f, \qquad h_I := h_{i_1}\cdot\ldots\cdot h_{i_k}, \qquad \text{and} \qquad |I| := k.$$

Examples. 1. The index $I = (2, 1, 2)$ of order $|I| = 3$ can be used to denote the product of h-components $h_I = h_2h_1h_2 = h_1h_2^2$ and the third partials $\partial_I f = \partial_{212}f = \partial_2\partial_1\partial_2 f$.

2. We can write the binomial formula as

$$(x_1 + x_2)^2 = \sum_{|I|=2} a_I x_I,$$

where $a_{11} := a_{12} := a_{21} := a_{22} := 1$. The coefficients a_{12} and a_{21} are not uniquely determined, only $a_{12} + a_{21} = 2$ must hold. Similarly, any polynomial of degree k in $x = (x_1, \ldots, x_n)$ can be written as $\sum_{|I|\le k} a_I x_I$ for suitable coefficients a_I.

For the following, we also wish to include the case of order $k = 0$: Then $I = \emptyset$ and $|I| = 0$, and we set the product over zero components of h equal to 1, and the zeroth derivative to be the function itself.

Using multiindices, we can rewrite, for instance, $F''(t) = \sum_{I\in\{1,\ldots,n\}^2} \partial_I f(x + th)\, h_I$ or, more generally, the j-th directional derivative of F in direction h as

$$F^{(j)}(t) = \sum_{|I|=j} \partial_I f(x + th)\, h_I, \qquad j \in \mathbb{N}_0. \tag{23}$$

We now set $t = 1$ in (21), and insert our result into (23). This proves the following expansion for $F(1) = f(x + h)$:

Theorem 11. *Suppose $f : U \to \mathbb{R}$ has k continuous partial derivatives ($k \in \mathbb{N}$). Moreover, for $h \in \mathbb{R}^n$ let the segment $\{x + th : 0 \le t \le 1\} \subset U$. Then there exists $\xi \in [0, 1]$ with*

$$f(x + h) = \sum_{0\le|I|\le k-1} \frac{\partial_I f(x)}{|I|!}\, h_I + \sum_{|I|=k} \frac{\partial_I f(x + \xi h)}{k!}\, h_I. \tag{24}$$

We denote the first sum by $T_x^{k-1}(h)$ and call it the *Taylor polynomial of degree* $k-1$ *for* f *with respect to* x. The Taylor polynomial T_x^k is the sum over all possible derivatives of order up to k, each multiplied with the appropriate product of h-components. The sum $\sum_{0\le|I|\le k}$ contains one term with $|I|=0$ (namely for $I=\emptyset$), n terms of order $|I|=1$, n^2 derivatives of order $|I|=2$, etc., up to n^k terms of order $|I|=k$. For instance, when $n=2$,

$$T_x^3(h) = f(x) + \Big(\partial_1 f(x)\,h_1 + \partial_2 f(x)\,h_2\Big) + \frac{1}{2!}\Big(\partial_{11}f(x)\,h_1^2 + \underline{\partial_{12}f(x)\,h_1h_2 + \partial_{21}f(x)\,h_2h_1}$$
$$+ \partial_{22}f(x)\,h_2^2\Big) + \frac{1}{3!}\Big(\underbrace{\partial_{111}f(x)\,h_1^3 + \partial_{112}f(x)\,h_1^2h_2 + \ldots + \partial_{222}f(x)\,h_2^3}_{2^3=8\text{ terms}}\Big).$$

Note that the two underlined terms agree by the Theorem of Schwarz; thus we can replace them by $2\,\partial_{12}f(x)\,h_1h_2$. Similarly, for third order, we can reduce the 8 terms to 4, etc. This means that our Taylor formula contains more terms than is necessary. To eliminate the redundant terms in (24), however, some further notation must be introduced. Only for such a form, the coefficients of the Taylor polynomial will be unique. Compare with the literature!

`43. Lecture, Tuesday, 5. June 07 ______________ Testklausur, "U 6`

To discuss extremals we will approximate a function by its quadratic Taylor polynomial $T_x^2(h)$. It will be convenient to have the remainder term estimated as follows:

Corollary 12 (Qualitative Taylor formula)**.** *Suppose* $f\in C^k(U,\mathbb{R}^m)$ *and* $B_r(x)\subset U$. *Then, for* h *with* $\|h\|<r$,

$$f(x+h) = \sum_{0\le|I|\le k}\frac{\partial_I f(x)}{|I|!}\,h_I + o(\|h\|^k). \tag{25}$$

According to the definition of the Landau symbol this is equivalent to

$$0 \;=\; \lim_{h\to0}\frac{1}{\|h\|^k}\Big(f(x+h) - \sum_{0\le|I|\le k}\frac{\partial_I f(x)}{|I|!}h_I\Big).$$

Proof. In case $m=1$, the theorem gives

$$f(x+h) - \sum_{0\le|I|\le k-1}\frac{\partial_I f(x)}{|I|!}\,h_I = \sum_{|I|=k}\frac{\partial_I f(x+\xi h)}{k!}\,h_I$$
$$\Rightarrow\quad \frac{1}{\|h\|^k}\Big(f(x+h) - \sum_{0\le|I|\le k}\frac{\partial_I f(x)}{|I|!}\,h_I\Big) = \sum_{|I|=k}\underbrace{\frac{\partial_I f(x+\xi h)-\partial_I f(x)}{k!}}_{\to0\text{ as }h\to0\text{ (continuity)}}\,\frac{h_I}{\|h\|^k}.$$

As $h \to 0$, the right hand has terms which are products of the type null sequence times bounded sequence. Thus their (finite) sum is once again a null sequence. Indeed, the second term is bounded:

$$\frac{|h_I|}{\|h\|^k} = \frac{|h_{i_1}| \cdots |h_{i_k}|}{\|h\|^k} = \frac{|h_{i_1}|}{\|h\|} \cdots \frac{|h_{i_k}|}{\|h\|} \le 1$$

This proves (25) for $m = 1$. Consequently, for f vector valued, (25) holds in each component, that is, it holds for f. □

Let us explicitely state the qualitative Taylor formula for small k.

• $k = 1$: For f continuously differentiable, the corollary gives

$$f(x+h) = f(x) + \big\langle \operatorname{grad} f(x), h \big\rangle + o(\|h\|),$$

which coincides with the definition of differentiability (2).

• $k = 2$: To write the quadratic term of the Taylor polynomial T_x^2 as a matrix product, we introduce the *Hessian [Hesse-Matrix]*

$$\operatorname{hess} f(x) := \big(\partial_i \partial_j f(x)\big)_{1 \le i,j \le n}.$$

Other common notations for the Hessian are $H(f)$, $D^2 f$, $d^2 f$. The Theorem of Schwarz implies that a Hessian with continuous entries is a symmetric matrix. Using the Hessian, we can write

$$f(x+h) = f(x) + \big\langle \operatorname{grad} f(x), h \big\rangle + \frac{1}{2} h^\top \operatorname{hess} f(x)\, h + o(\|h\|^2). \tag{26}$$

• $k = 0$: The statement of the corollary extends to this case (but not its proof). Each function for with continuous 0-th derivative, namely each continuous function, satisfies indeed

$$f(x+h) = f(x) + o(1).$$

2.3. **Extremals.** While the existence of extremals follows from the theorem of the maximum, let us now discuss how we locate extremals by the tools of differential calculus. As for one variable, this will only detect extremals of the following kind:

Definition. Let $X \subset \mathbb{R}^n$. Then a function $f\colon X \to \mathbb{R}$ attains a *local* $\left\{ \begin{smallmatrix} \textit{maximum} \\ \textit{minimum} \end{smallmatrix} \right\}$ at $x \in X$ if there is $r > 0$ with $B_r(x) \subset X$ such that $\left\{ \begin{smallmatrix} f(x) \ge f(x+h) \\ f(x) \le f(x+h) \end{smallmatrix} \right\}$ holds for all $\|h\| < r$.
If the inequality is strict for $h \ne 0$ then we say the local extremum is *strict* (or *isolated*).

Examples. Let $f\colon \mathbb{R}^2 \to \mathbb{R}$.
1. $f(x,y) := x^2 + y^2$ has a strict minimum at 0, likewise $f(x,y) := x^4 + y^4$.
2. $f(x,y) := x^2$ has a local minimum at each point in $\{0\} \times \mathbb{R}$.

3. These functions restricted to the domain $B_1(0)$ do not attain a maximum (only a supremum). Over $\overline{B_1}(0)$ they attain a maximum, but it is not local.

We now discuss extrema on open subsets $U \subset \mathbb{R}^n$.

Theorem 13 (Necessary condition for extrema)**.** *If $f\colon U \to \mathbb{R}$ is partially differentiable at x and attains a local extremum at x, then*

$$\operatorname{grad} f(x) = 0,$$

that is, $\partial_1 f(x) = \ldots = \partial_n f(x) = 0$.

Points $x \in U$ with $\operatorname{grad} f(x) = 0$ are called *critical points [kritische Punkte]* of f.

Proof. The function $g_j(t) := f(x+te_j)$ is defined for $t \in (-r, r)$. Since g_j has an extremum at $t = 0$ (restriction preserves extrema), we must have $0 = g_j'(0) = \partial_j f(x)$. □

Let us return to the Taylor expansion (26). At an extremum x, we have $\operatorname{grad} f(x) = 0$. Thus the quadratic term governs the local behaviour of f near x. To obtain sufficient conditions for extrema, we must impose sign conditions on the Hessian:

Definition. A symmetric $n \times n$-matrix A, or the quadratic form $Q(x) = x^\top Ax$, is called
(i) $\left\{ {\textit{positive} \atop \textit{negative}} \right\}$ *definite*, if $x^\top A\, x \left\{ {> \atop <} \right\} 0$ for all $x \in \mathbb{R}^n \setminus \{0\}$,
(ii) $\left\{ {\textit{positive} \atop \textit{negative}} \right\}$ *semidefinite* if $x^\top A\, x \left\{ {\geq \atop \leq} \right\} 0$ for all $x \in \mathbb{R}^n$, and
(iii) *indefinite* if $x^\top A\, x$ attains both positive and negative values.

Examples. $\left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right)$ is positive definite, $\left(\begin{smallmatrix} -1 & 0 \\ 0 & -1 \end{smallmatrix}\right)$ is negative definite, $\left(\begin{smallmatrix} 1 & 0 \\ 0 & 0 \end{smallmatrix}\right)$ is positive semidefinite, and $\left(\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right)$ is indefinite.

Let us relate definiteness of matrices to signs of eigenvalues. One of the most important assertions of Linear Algebra is the principal axis theorem, see Thm. 13 below. It tells us that (real) symmetric matrices can be diagonalized. That is, for A symmetric there is an orthonormal basis $v_1, \ldots, v_n$ such that writing any vector in the form $v = y_1 v_1 + \ldots + y_n v_n$ the quadratic form becomes

$$v^\top A\, v = \lambda_1 y_1^2 + \ldots + \lambda_n y_n^2.$$

Here, the numbers $\lambda_i \in \mathbb{R}$ are the *eigenvalues* of A, and the basis $v_1, \ldots, v_n$ of *eigenvectors* can be obtained by rotation from the standard basis $e_1, \ldots, e_n$.

This leads to the following description of definiteness in terms of eigenvalues:

Proposition 14. *Let A be a symmetric matrix. Then:*
A $\left\{ {\textit{positive} \atop \textit{negative}} \right\}$ definite $\iff$ all eigenvalues of A are $\left\{ {> \atop <} \right\} 0$.
A indefinite $\iff$ A has eigenvalues > 0 and < 0 as well.

Let us specialize the previous result to dimension $n = 2$. Then $\det A = \lambda_1\lambda_2$ and so:
A indefinite $\iff \det A < 0$.
A definite $\iff \det A > 0$. The sign of $\operatorname{trace} A = \lambda_1 + \lambda_2$ tells us whether A is positive or negative definite.

In the Taylor formula, the second order term involves the Hessian. Let us now require definiteness conditions which assign the same sign to this term for all $h \in \mathbb{R}^n$:

Theorem 15 (Sufficient condition for extrema)**.** *Let $f\colon U \to \mathbb{R}$ be twice continuously differentiable. Suppose $x \in U$ is a critical point for f, that is,* $\operatorname{grad} f(x) = 0$.
(i) If $\operatorname{hess} f(x)$ *is* $\left\{ \substack{\textit{positive}\\ \textit{negative}} \right\}$ *definite, then x is strict local* $\left\{ \substack{\textit{minimum}\\ \textit{maximum}} \right\}$.
(ii) If $\operatorname{hess} f(x)$ *is indefinite, then x cannot be an extremum of f.*

`44. Lecture, Tuesday, 12. June 07 ________________ T 8, "U 7`

Case (i) of the theorem generalizes the one-dimensional statement $f'(x) = 0$, $f''(x) > 0$ $\Rightarrow$ x minimum. Case (ii) is genuine multidimensional, and the proof will show that f restricted to lines in some direction has a maximum at x, while to others a minimum; thus we call x a *saddle point.*

In case of a semidefinite matrix the theorem does not assert anything, just as in one variable the conditions $f'(x) = 0$, $f''(x) \geq 0$ can hold at x a saddle point, minimum, or maximum (give examples!).

Example. We want to locate the extrema of the function

$$f(x, y) := x(x-1)^2 - 2y^2.$$

1. To determine critical points, we compute

$$\operatorname{grad} f(x,y) = \begin{pmatrix} (x-1)^2 + 2x(x-1) \\ -4y \end{pmatrix} = \begin{pmatrix} (3x-1)(x-1) \\ -4y \end{pmatrix}$$

Thus $p := (\frac{1}{3}, 0)$ and $q := (1, 0)$ are critical.
2. We calculate the Hessian:

$$\operatorname{hess} f(x,y) = \begin{pmatrix} 6x-4 & 0 \\ 0 & -4 \end{pmatrix} \quad \Rightarrow \quad \operatorname{hess} f(p) = \begin{pmatrix} -2 & 0 \\ 0 & -4 \end{pmatrix}, \quad \operatorname{hess} f(q) = \begin{pmatrix} 2 & 0 \\ 0 & -4 \end{pmatrix}$$

3. We determine the sign of the eigenvalues. For our example we can read them off without further calculation:
- At p the Hessian is negative definite, and p is a strict local maximum, while
- at q it is indefinite so that q is a saddle point.

Proof. (i) Let us consider the case that $A := \operatorname{hess} f(x)$ is positive definite. The negative definite case is similar.

We want to use the qualitative Taylor formula (26),

$$f(x+h) = f(x) + \langle \underbrace{\operatorname{grad} f(x)}_{=0}, h \rangle + \frac{1}{2} h^\top A\, h + R(h) \qquad \text{with} \quad \lim_{h\to 0} \frac{R(h)}{\|h\|^2} = 0.$$

We claim that there are numbers $m > 0$, $r > 0$, such that

$$(27) \qquad (a) \quad h^\top A h \geq m\|h\|^2 \qquad \text{and} \qquad (b) \quad |R(h)| \leq \frac{m}{4}\|h\|^2$$

for all h with $\|h\| < r$. For these h we then obtain

$$f(x+h) \geq f(x) + \frac{m}{2}\|h\|^2 - \frac{m}{4}\|h\|^2 = f(x) + \frac{m}{4}\|h\|^2.$$

As $m > 0$ this implies $f(x+h) > f(x)$ for all $h \neq 0$ with $\|h\| < r$. Hence x is a strict local minimum, as desired. (Note: From definiteness alone it is immediate that (27) holds with $m \geq 0$. This, however, is not good enough to compensate the error term $R(h)$.)

To show (a), we need to estimate $h^\top A h$ from below. We use a beautiful trick. Let us consider this quadratic form first restricted to the unit sphere $\mathbb{S}^{n-1} = \{h \in \mathbb{R}^n : \|h\| = 1\}$. It is closed and bounded, hence compact. Moreover, $h \mapsto h^\top A h = \sum a_{ij} h_i h_j$ is continuous. By Theorem IV.13, the function $h^\top A h$ takes a minimum, that is,

$$(28) \qquad \exists\ m \in \mathbb{R}, h_0 \in \mathbb{S}^{n-1} \quad \text{such that} \quad h^\top A h \geq h_0^\top A h_0 =: m \qquad \text{for all } h \in \mathbb{S}^{n-1}.$$

Moreover, $m > 0$ since $A = \operatorname{hess} f(x)$ is positive definite. Our claim (a) is obvious for $h = 0$. Else, $\frac{h}{\|h\|}$ is a unit vector and so (a) follows from (28):

$$h^\top A h = \|h\|^2 \frac{h^\top}{\|h\|} A \frac{h}{\|h\|} \geq \|h\|^2 m \qquad \text{for all } h \in \mathbb{R}^n$$

Note that invoking the principal axis theorem, it becomes clear that m is precisely the smallest eigenvalue of $\operatorname{hess} f(x)$.

Let us now pick $r > 0$ to prove (b). We have $\lim_{h\to 0} \frac{R(h)}{\|h\|^2} = 0$. Thus for each $\varepsilon > 0$ there is $r = r(\varepsilon) > 0$, such that $\frac{|R(h)|}{\|h\|^2} < \varepsilon$ holds for all $h \in B_r(0)$. (If not, there is some sequence $h_k \to 0$ with $\frac{|R(h_k)|}{\|h_k\|^2} \geq \varepsilon$, contradicting convergence to 0.) Picking r for $\varepsilon := \frac{m}{4}$ small enough for $B_r(x) \subset U$ to hold, implies (b)

(ii) By assumption there are $v, w \in \mathbb{R}^n$ with $v^\top \operatorname{hess} f(x)\, v > 0$ and $w^\top \operatorname{hess} f(x)\, w < 0$. Let us consider the restrictions $g(t) := f(x+tv)$ and $h(t) := f(x+tw)$, defined on intervals about 0. Since x is critical,

$$g'(0) = \langle \operatorname{grad} f(x), v \rangle = 0 \qquad \text{and} \qquad h'(0) = \langle \operatorname{grad} f(x), v \rangle = 0.$$

Second directional derivatives of f are given by the quadratic form associated to the Hessian, see (22):

$$g''(0) = v^\top \mathrm{hess}\, f(x)\, v > 0 \qquad \text{and} \qquad h''(0) = w^\top \mathrm{hess}\, f(x)\, w < 0$$

By the sufficient condition for extrema in one variable we conclude: g has a strict local minimum at 0, and h has a strict local maximum. Thus each neighbourhood of x contains points whose images under f are (strictly) smaller or larger than $f(x)$. This means that x is not a local extremum for f. $\square$

Part (ii) of the proof emphasizes once again the meaning of the second order term in the Taylor series. Consider the restriction of f to a line through x in direction v. The graph of this restriction is given by the intersection of a vertical plane with the entire graph of f. The second derivative of this restriction at x agrees with $\frac{1}{2}v^\top \mathrm{hess}\, f(x)\, v$.

SUMMARY

We introduced differentiability as linear approximability. The dimension of the range is insignificant, as we can view the functions componentwiser. However, higher dimensions of the domain necessitates the theory we are describing.

There are two geometric interpretations of the differential:
1. The differential $df_x(v)$ is the directional derivative of the function f in direction v. In fact, we reduced higher dimensional derivative to the one-variable case by considering the function $t \mapsto f(x + tv)$. This picture is particularly useful for the scalar valued case, when it is possible to view the two-dimensional graph of $t \mapsto f(x + tv)$ as the section of the $(n + 1)$-dimensional graph of f with the plane spanned by v and e_{n+1}.
2. We can view the differential as the map which sends tangent vectors $v = c'(0)$ to curves c in the domain to the vectors tangent to the image curve, $df(v) = \frac{d}{dt}(f \circ c)(t)|_{t=0}$ (chain rule!).

Usually, the differential is computed via partial derivatives: The partial derivatives give the matrix representation of the differential, called the Jacobian. This works provided the partial derivatives are continuous, that is, for C^1-functions.

The most remarkable rule for differentiation is the chain rule, which is genuine multi-dimensional: The composition of the linearisations is precisely the linearisation of the composed map!

We derived Taylor's formula in several variables from the one-dimensional Taylor formula, by restricting the function to a line in the domain. Mostly, it is sufficient to know the Taylor expansion up to second order terms. It involves the gradient and Hessian of a function. At a critical point, the Hessian has one sign (independent of the direction) in the definite case. This knowledge is sufficient to distinguish maxima from minima or saddle points; similar to the case of one dimension, this discussion is incomplete in case the Hessian has zero eigenvalues.

Part 7. Nonlinear equations

A main task of mathematics is to solve equations, that is, to determine solutions x of equations $f(x) = y$. We consider here the case that f maps from $\mathbb{R}^{n+k}$ to $\mathbb{R}^n$ or subsets thereof. We assume $k \geq -n$ and so have a system of $n \in \mathbb{N}$ equations in $n + k \geq 0$ unknowns.

As an example, let us consider the particular case that f is linear and $y = 0$. Then the set of all solutions $M \ni x$ forms a linear subspace. It can be represented explicitely: M has a basis $v_i \in \mathbb{R}^{n+k}$, $i = 1, \dots, \dim \ker f$, that is, all solutions are parameterized by the map

$$\mathbb{R}^k \ni (a_1, \dots, a_k) \mapsto \sum_{i=1}^{k} a_i v_i \in M.$$

In the non-linear case we can no longer hope for an explicit representation of the solution set.

Nevertheless we can make assertions for the solvability of non-linear equations $f(x) = y$. As in the linear case we distinguish depending on the number of equations and unknowns:
• For n equations in $0 \leq n + k < n$ unknowns, the system is overdetermined: We do not expect a solution at all and disregard this case.
• For n equations in n unknowns, the *Inverse Mapping Theorem* says: Given a solution $f(x_0) = y_0$, for each y near y_0 there is a unique x near x_0, such that $f(x) = y$.
• For n equations in $n + k > n$ unknowns, the system is underdetermined. For each y, we expect many solutions and the task is to parameterize the space of solutions. The *Implicit Mapping Theorem* does this locally in terms of graphs.

1. The inverse mapping theorem

The inverse mapping theorem is one of the major theorems in calculus. We will first illustrate for which kind of statement we can hope. Then we will discuss auxiliary results, which are also interesting in their own right. Together with the results of Section VI 1.5, they will lead to a proof of the inverse mapping theorem.

1.1. **Global and local invertibility.** In the single variable case, how would you check whether a function $f\colon [a, b] \to \mathbb{R}$ is invertible? The easiest is to verify $f'(x) \neq 0$ for all $x \in [a, b]$; if so, f is strictly monotone and f is invertible. For several dimensions, the situation is more involved, as we shall see next.

Recall that a map $f\colon X \to Y$ is *invertible* if it has an inverse $f^{-1}\colon Y \to X$, which inverse from left and right, $f^{-1} \circ f = \mathrm{id}_X$, $f \circ f^{-1} = \mathrm{id}_Y$. From linear algebra we know that a linear map $L\colon \mathbb{R}^n \to \mathbb{R}^m$ can have rank at most $\min\{n,m\}$ and so:

$$\text{(1)} \qquad L \text{ invertible} \qquad \Leftrightarrow \qquad \operatorname{rank} L = n = m.$$

In order to invert a nonlinear map, it is therefore reasonable to concentrate on the case $f\colon U \subset \mathbb{R}^n \to \mathbb{R}^n$ and require that its linear approximation is invertible, that is,

$$\text{(2)} \qquad df_x \quad \text{has rank } n \text{ for all } x \in U.$$

This is the natural generalization of the condition $f'(x) \neq 0$ familiar from the one-variable case.

45. Lecture, Thursday, 14. June 07 ______________________ T 9

Condition (2) is not sufficient for invertibility, as we want to show on the example of the polar coordinates map $P\colon (0,\infty) \times \mathbb{R} \to \mathbb{R}^2 \setminus \{0\}$, $P(r,\varphi) = (r\cos\varphi, r\sin\varphi)$ with

$$\det J_P(r,\varphi) = \det \begin{pmatrix} \cos\varphi & -r\sin\varphi \\ \sin\varphi & r\cos\varphi \end{pmatrix} = r \neq 0 \quad \Rightarrow \quad \operatorname{rank} df_x = 2.$$

Hence (2) is satisfied. Nevertheless:

• P is not *(globally) invertible*: $P(r,\varphi) = P(r,\varphi + 2k\pi)$ for all $k \in \mathbb{Z}$. Indeed, "polar coordinates of $(x,y) \in \mathbb{R}^2$" are given as a preimage $(r,\varphi) \in P^{-1}(x,y)$. Here the polar angle φ is only defined up to multiples of 2π. It is the reason why we introduced P and not P^{-1} as the polar coordinate

• P is *locally invertible*, that is, invertible on sufficiently small subsets of its domain. Indeed, the restriction of P to a strip of height 2π, like

$$P \ : \ \Omega := \mathbb{R}_+ \times \left[-\frac{\pi}{2}, \frac{3}{2}\pi\right) \ \to \ \mathbb{R}^2 \setminus \{0\},$$

is bijective. Unfortunately, however, such inverses $P^{-1}\colon \mathbb{R}^2 \to \Omega$ is not continuous (and so in particular not differentiable).

Problem. Restrict P to the strip $\mathbb{R}^+ \times (-\frac{\pi}{2}, \frac{\pi}{2})$ of width π. Check that the inverse has the explicit form $P^{-1}(x,y) := \left(\sqrt{x^2+y^2},\ \arctan\frac{y}{x}\right)$.

Suppose we consider maps between spaces of equal dimension. It is the statement of the inverse mapping theorem that a mapping is locally invertible provided its Jacobian has the maximal rank (2).

1.2. **The differential of the inverse mapping.** To distinguish different forms of bijective maps we define:

Definition. Suppose $f\colon U \to V$, for $U \subset \mathbb{R}^n$, $V \subset \mathbb{R}^m$.
(i) f is a *homeomorphism [Homöomorphismus]* if f is bijective, f is continuous, and the inverse f^{-1} is continuous.
(ii) A homeomorphism f is called a *diffeomorphism [Diffeomorphismus]* if f and f^{-1} are continuously differentiable.

Thus we have the following implications: diffeomorphism $\Rightarrow$ homeomorphism $\Rightarrow$ bijection.

Examples. 1. arctan is a diffeomorphism between $\mathbb{R}$ and the open interval $(-\frac{\pi}{2}, \frac{\pi}{2})$.
2. $f(x) := \frac{x}{\|x\|}\arctan\|x\|$ for $x \neq 0$, and $f(0) := 0$ is a diffeomorphism between $\mathbb{R}^n$ and $B^n_{\pi/2}(0)$.
3. $x \mapsto x$ for $x \in (0,1]$ and $x \mapsto x-1$ for $x \in (2,3)$ is a bijection onto its image $(0,2)$ but not a homeomorphism.
4. x^3 is a homeomorphism of $\mathbb{R}$, but not a diffeomorphism.
5. There is no homeomorphism from $\mathbb{S}^1$ to $\mathbb{R}$. Indeed, any continuous map from the compact set $\mathbb{S}^1$ must have a compact image, and hence cannot be surjective onto $\mathbb{R}$.

Outlook. Homeomorphisms define an equivalence relation on the set of subsets of $\mathbb{R}^n$. They are classified in *topology.* One can say that topology studies the properties of objects which persist when these are made from (ideal) rubber: arbitrary deformations of the objects are allowed, but tearing them apart is forbidden. Fork and plate are then in the same equivalence class (there is a homeomorphism from fork to plate, considered as open sets), but a mug is in a different class, due to its handle. Literally homeomorphism means "same shape", from greek $\mu o\rho\varphi\eta$ = form, shape, appearence, and $o\mu o\sigma$ stands for alike, similar.

Proposition 1. *Suppose $U \subset \mathbb{R}^n$, $V \subset \mathbb{R}^m$ are open and $f\colon U \to V$ is a diffeomorphism with inverse $g\colon V \to U$. Then*
(i) $n = m$, and
(ii) at each $x \in U$ and $y := f(x)$ the differentials are invertible linear maps with

$$dg_y = (df_x)^{-1} \quad \text{or} \quad J_g(y) = \big(J_f(x)\big)^{-1}. \tag{3}$$

Proof. We use the Chain Rule to differentiate $\mathrm{id}_U = g \circ f$:

$$\mathrm{id}_{\mathbb{R}^n} = d(\mathrm{id}_U)_x = d(g \circ f)_x \overset{\text{Chain R.}}{=} dg_{f(x)} \circ df_x. \tag{4}$$

Similarly, $\mathrm{id}_{\mathbb{R}^m} = df_x \circ dg_{f(x)}$. Thus $dg_{f(x)}$ is invertible with (left and right) inverse df_x. The linear algebra fact (1) gives (i), while (ii) follows directly from (4). $\square$

Thus diffeomorphisms preserve dimension. Homeomorphisms also do that, but that is harder to prove. However, bijections (and even continuous surjections) do *not* preserve dimension; examples such as the Peano curve show that.

As we saw in Example 1 above, a continuous bijection does not necessarily have a continuous inverse. But if it has, and the original mapping is differentiable, then the differentiability of the inverse is automatic:

Proposition 2. *Suppose for $U, V \subset \mathbb{R}^n$ open the mapping $f\colon U \to V$ is a homeomorphism, which is continuously differentiable. If for all $x \in U$ the differential df_x is invertible, that is,* $\det J_f(x) \neq 0$, *then f is a diffeomorphism.*

Proof. For each $x \in U$, we need to show the differentiability of the inverse mapping at $f(x)$. We consider only the case

$$x = 0, \quad f(0) = 0, \quad df_0 = \mathrm{id}\,. \tag{5}$$

We translate in domain and range so that the first two assumptions hold. Then, to satisfy the last one, we consider the linear map $L := df_0^{-1}$ which by assumption is an isomorphism. We now consider by $\tilde{f} := L \circ f$. The Chain Rule gives $d\tilde{f}_0 = d(L \circ f)_0 = df_0^{-1} \circ df_0 = \mathrm{id}$, so that the third condition holds for $\tilde{f}$. Moreover, f is a homeomorphism, or a diffeomorphism, if and only if $\tilde{f}$ is. Thus our assumption means no loss of generality.

For f satisfying (5), differentiability implies

$$f(h) = h + r(h) \qquad \text{with} \quad \lim_{h \to 0} \frac{r(h)}{\|h\|} = 0. \tag{6}$$

For the inverse homeomorphism $g := f^{-1}$, assumption (5) implies $g(0) = 0$. Moreover, if it is differentiable then, in view of Prop. 1, its differential at 0 is also id. Thus we need to show

$$g(k) = k + R(k) \qquad \text{with} \quad \lim_{k \to 0} \frac{R(k)}{\|k\|} = 0. \tag{7}$$

Since f is bijective, each small k can be written as $k = f(h)$, or equivalently, $h = g(k)$.

We now claim there is $\delta > 0$ such that for $k = f(h)$

$$\frac{\|R(k)\|}{\|k\|} \overset{(*)}{\leq} 2\frac{\big\|r\big(g(k)\big)\big\|}{\|g(k)\|} = 2\frac{\|r(h)\|}{\|h\|} \quad \text{for all } k \in B_\delta(0) \setminus \{0\}, \tag{8}$$

where $(*)$ remains to be shown. Then if $k \to 0$, by continuity, also $h = g(k) \to 0$ and so (8) proves the desired remainder term estimate (7) for g.

We now prove $(*)$ separately for nominator and denominator. For the nominator, plugging $h = g(k)$ into (6) gives $k = g(k) + r\big(g(k)\big)$. Comparing with (7) yields, as desired,

$$R(k) = -r\big(g(k)\big).$$

Now we come to the denominator. By (6) there is $\rho > 0$ such that $\|r(h)\| \leq \frac{1}{2}\|h\|$ for all $h \in B_\rho(0)$. Moreover, since g is continuous, we can pick $\delta > 0$ such that $\|g(k)\| < \rho$ for all $k \in B_\delta(0)$; by openness we may assume δ is small enough for $B_\delta(0) \subset V$. Consequently,

$$\big\|R(k)\big\| \leq \frac{1}{2}\big\|g(k)\big\| \quad \text{for all } k \in B_\delta(0).$$

Using this in $g(k) = k + R(k)$ gives us $\|g(k)\| \leq \|k\| + \frac{1}{2}\|g(k)\|$, i.e.,

$$\big\|g(k)\big\| \leq 2\|k\| \quad \text{for all } k \in B_\delta(0).$$

This proves $(*)$ in (8) and thus g is differentiable.

`46. Lecture, Tuesday, 19. June 07` ———————————————— `"U 8`

It remains to show that g is continuously differentiable. This can be seen from the formula for the inverse (3), as follows. By assumption, $x \mapsto df_x$ is continuous and so is the composition $y \mapsto df_{g(y)}$. By (3), the inverse of this differential gives the desired differential, $(df_{g(y)})^{-1} = dg_y$. Thus the proof will be completed by invoking a linear algebra fact: If $A \in \mathsf{GL}(n)$ then $A \mapsto A^{-1}$ is a continuous map from $\mathbb{R}^{n^2}$ to $\mathbb{R}^{n^2}$.

Indeed, the inverse can be represented as $A^{-1} = \frac{1}{\det A} A_{\text{adj}}$. Here the *adjoint* matrix of A is given by $(A_{\text{adj}})_{ij} := (-1)^{i+j} \det X_{ij}(A)$, where $X_{ij}(A)$ is the $(n-1) \times (n-1)$ matrix obtained from A by crossing out the i-th row and the j-th column. The formula can be obtained as a byproduct of Cramer's rule. It may be impractical for computation, but it shows that A^{-1} continuously depends on the entries of A. □

1.3. **The contraction mapping principle.** We present a powerful theorem with many applications. Nevertheless its proof is simple.

Definition. Let (X, d) be a non-empty metric space. A mapping $f\colon X \to X$ is called a *contraction [Kontraktion]* if there is a number $0 \leq \lambda < 1$ such that

$$d\big(f(x), f(y)\big) \leq \lambda d(x, y) \qquad \text{for all } x, y \in X. \tag{9}$$

More generally, if (9) holds with $\lambda \in [0, \infty)$ then f is called a *Lipschitz continuous*. A Lipschitz continuous map is always continuous (check!), in particular contractions are continuous.

Example. An endomorphism of $\mathbb{R}^n$, represented by a $n \times n$-matrix A, such that $\|A\| =: \lambda < 1$, is a contraction. Indeed, $\|A(x-y)\| \le \lambda \|x-y\|$ by (11).

Recall that a metric space is *complete* if each Cauchy sequence converges.

Example. Any closed subset $X \subset \mathbb{R}^n$ with standard metric $d(x,y) := \|x-y\|$ is complete: Consider a Cauchy sequence (x_k) in X. As a sequence in $\mathbb{R}^n$ we thus have convergence, $x_k \to x \in \mathbb{R}^n$. Closedness then gives $x \in X$.

Theorem 3 (Contraction mapping theorem [*Banachscher Fixpunktsatz*])**.** *Let X be a complete metric space, and $f\colon X \to X$ be a contraction. Then f has a unique* fixed point*, that is, there is exactly one point $a \in X$ such that $f(a) = a$.*

A nice interpretation of the theorem is as follows: Take a map of the area you live in, and place it on the floor. Then there is exactly one point on the floor which is exactly underneath its image on the map.

In case $\lambda = 0$ the map $f\colon X \to X$ is constant, that is, $f(x) = a$. In particular $f(a) = a$ and a is the unique fixpoint.

If we drop the clompleteness assumption on X the contraction mapping theorem will fail: To see this, consider $f\colon (0,1) \to (0,1)$, $f(x) := \frac{1}{2}x$.

Proof. We pick an arbitrary point $a_0 \in X$, and define recursively a sequence (a_k) by setting

$$a_1 := f(a_0), \quad a_2 := f(a_1), \quad a_3 := f(a_2), \quad \dots .$$

We claim that (a_k) is Cauchy. We have $d(a_{k+1}, a_k) = d\big(f(a_k), f(a_{k-1})\big) \le \lambda d(a_k, a_{k-1})$ and therefore, by induction,

$$d(a_{k+1}, a_k) \le \lambda^k d(a_1, a_0).$$

Using the triangle inequality and the previous formula gives, for $k, m \in \mathbb{N}$,

$$\begin{aligned} d(a_{m+k}, a_k) &\le d(a_{m+k}, a_{m+k-1}) + \ldots + d(a_{k+1}, a_k) \\ &\le \big(\lambda^{m-1} + \ldots + 1\big)\lambda^k d(a_1, a_0) \overset{\text{geom. sum}}{\le} \lambda^k \frac{1}{1-\lambda} d(a_1, a_0). \end{aligned}$$

Since $\lambda^k \to 0$ this proves the claim.

Since X is complete, the Cauchy sequence (a_k) converges to some $a \in X$. Then

$$\begin{aligned} d\big(a, f(a)\big) &\overset{\Delta-\text{inequ.}}{\le} d(a, a_{k+1}) + d\big(f(a_k), f(a)\big) \\ &\le d(a, a_{k+1}) + \lambda d\big(a_k, a\big) \to 0 \quad \text{as } k \to \infty, \end{aligned}$$

and so in fact $d\big(a, f(a)\big) = 0$, implying that a is a fixed point.

Finally, to show uniqueness of a, suppose that b is also a fixed point. Then $d(a,b) = d\big(f(a), f(b)\big) \le \lambda d(a,b)$. Noting that $\lambda < 1$, this gives $d(a,b) = 0$, that is, $a = b$. $\square$

In the applications, often X is a closed ball. This may be a ball in $\mathbb{R}^n$ or in any complete normed verctor space (Banach space), for instance in a function space with suitable norm. Indeed, you will see such an application in the class on ordinary differential equations.

Example. Let $A \in \mathsf{M}(n)$ represent a linear map, $\mathrm{Id} \in \mathsf{M}(n)$ be the unit matrix, and $y \in \mathbb{R}^n$. In the following we will construct solutions of $(\mathrm{Id} - A)x = y$. This equation can be regarded as a fixed point problem: Indeed,

$$(\mathrm{Id} - A)x = y \quad \Leftrightarrow \quad x = Ax + y =: f_y(x) \tag{10}$$

Under the assumption $\|A\| < 1$ the map f_y is a contraction:

$$\|f_y(x) - f_y(y)\| = \|Ax - Ay\| \le \|A\| \|x - y\|$$

The contraction mapping theorem yields a unique fixed point $x = Ax + y$, and hence a solution $y := (\mathrm{Id} - A)x$ of (10); in particular $\mathrm{Id} - A$ is invertible. Note that the identity is certainly invertible, and so we have convinced ourselves that if $\|A\| < 1$ the perturbation of the identity $\mathrm{Id} - A$ is still invertible.
Moreover, the proof of the contraction mapping theorem gives us a useful explicit formula for x. Indeed, taking y for the initial point, we consider the sequence

$$f_y^0(y) = y, \qquad f_y^1(y) = Ay + y, \qquad f_y^2(y) = A(Ay + y) + y = A^2 y + Ay + y, \quad \dots$$

so that for general $k \in \mathbb{N}$ (by induction) $f_y^k(y) = \sum_{i=0}^k A^i y$. We conclude that for $\|A\| < 1$ the fixed point has the representation

$$x = \lim_{k \to \infty} f_y^k(y) = \sum_{i=0}^{\infty} A^i y,$$

where the series converges. We have constructed an explicit inverse of $\mathrm{Id} - A$, in form of the *Neumann series* of A:

$$\|A\| < 1 \quad \Rightarrow \quad (\mathrm{Id} - A)^{-1} = \sum_{i=0}^{\infty} A^i$$

For $n = 1$, this is the geometric series. Our derivation also works in infinite dimensional complete normed vector spaces (Banach spaces) and has various applications.

`47. Lecture, Thursday, 21. June 07` ——————————————— `T 10`

1.4. **Local invertibility.** We can now show that a mapping with invertible linearization is locally invertible.

Theorem 4 (Inverse mapping theorem (IMT) *[Umkehrsatz]*)**.** *Let $f: U \to \mathbb{R}^n$ be continuously differentiable, and $a \in U$. Suppose that df_a is invertible, i.e.,* $\det J_f(a) \neq 0$*. Then a has an open neighbourhood $V \subset U$, such that $f: V \to f(V)$ is a diffeomorphism, that is,*
(i) $f: V \to \mathbb{R}^n$ is injective,
(ii) $f(V)$ is open,
(iii) f has a inverse $f^{-1} \in C^1\big(f(V), V\big)$.

Before giving the proof, let us comment on the statement.

Remarks. 1. Regarding the assertion of inverse mapping theorem:
• It is a perturbation result: For a given solution $f(a) = b$ we can find a solution $f(x) = y$ such that x is near a and y near b.
• It is an existence statement. It does not give a recipe on how to compute the solution.
• There is no claim on the problem of global invertibility. In fact, I'm not aware of a useful general statement about this case. Mostly this problem is decided on a case to case basis.

2. Regarding the assumption $\det J_f(a) \neq 0$:
• The assumption is not necessary for f to be locally invertible. (Cf. the one-dimensional invertible example $f(x) := x^3$ with $f'(0) = 0$.) It is, however, necessary for f to have a local inverse which is differentiable (by Prop. 1).
• The assumption is only required at the point a. But automatically it is also valid in some neighbourhood of a, since for any C^1-map, $x \mapsto \det J_f(x)$ is continuous, and hence still nonzero in some neighbourhood of a (by Cor. V.17).

Example. Consider the mapping $f: \mathbb{R}^2 \to \mathbb{R}^2$, $f(x, y) := (x^2, y)$. This map folds the left halfplane of $\mathbb{R}^2$ over the right halfplane. Additionally, both copies are subject to a stretch in the x-direction. At which points is f locally invertible? Points $(0, y)$ do not possess a neighbourhood V as in the theorem: A neighbourhood will contain the points $(-\varepsilon, y) \neq (\varepsilon, y)$ for some $\varepsilon > 0$, which take the same image. From the geometric description it follows that Note that indeed the Jacobian $J_f(x, y) = \left(\begin{smallmatrix} 2x & 0 \\ 0 & 1 \end{smallmatrix}\right)$ has determinant $2x$, which is nonzero for all (x, y) with $x \neq 0$. Thus the inverse mapping theorem asserts that for (x, y) with $x \neq 0$ there is a radius $r > 0$ such that the restriction $f: B_r(x, y) \to f\big(B_r(x, y)\big)$ becomes invertible. Do you see how large $r = r(x, y)$ can be at most for this example?

Proof. Let us first consider the case

$$a = 0, \quad f(a) = 0, \quad df_a = \mathrm{id}\,. \tag{11}$$

To solve the equation $f(x) = y$ let us introduce the mapping

$$F_y(x) := x - f(x) + y, \qquad x, y \in U.$$

Then we are interested in locating fixed points of F_y. Indeed,

$$f(x) = y \qquad \Longleftrightarrow \qquad x \text{ is a fixed point of } F_y;$$

We will exhibit such fixed points by applying the contraction mapping principle.

To do this, we determine $r > 0$ with $\overline{B_{2r}} = \overline{B_{2r}}(0) \subset U$, such that for all $y \in B_r(0)$
(a) the mappings F_y map $\overline{B_{2r}}$ into itself, and
(b) F_y has the contraction property (9) for any pair of points $x_1, x_2 \in \overline{B_{2r}}$.

The mapping $x \mapsto \|\operatorname{id} - df_x\|$ (operator norm) is continuous and vanishes at 0 by (11). By the ε-δ definition of continuity we can determine $r > 0$ such that

$$\|\operatorname{id} - df_x\| = \|(\operatorname{id} - df_x) - (\operatorname{id} - df_0)\| \leq \frac{1}{2\sqrt{n}} \qquad \text{for each } x \in \overline{B_{2r}}.$$

This gives a derivative bound,

$$\big\|(dF_y)_x\big\| = \|\operatorname{id} - df_x\| \leq \frac{1}{2\sqrt{n}} \qquad \text{for each } x \in \overline{B_{2r}},\ y \in U,$$

which in view of Proposition VI.9 implies a bound on the difference of values of F_y,

$$\text{(12)} \quad \big\|F_y(x_1) - F_y(x_2)\big\| \overset{\text{Pr.VI.9}}{\leq} \|(dF_y)_x\|\sqrt{n}\|x_1 - x_2\| \leq \frac{1}{2}\|x_1 - x_2\| \quad \forall x_1, x_2 \in \overline{B_{2r}},\ y \in U.$$

This verifies the contraction property (b). To verify (a), let us now restrict to $y \in B_r$. Then for $x \in \overline{B_{2r}}$ we indeed have

$$\|F_y(x)\| = \|x - f(x) + y\| \leq \|\underbrace{x - f(x)}_{=F_0(x)}\| + \|y\| = \|F_0(x) - \underbrace{F_0(0)}_{=0}\| + \|y\| \overset{(12)}{\leq} \frac{1}{2}\|x\| + \|y\| < 2r.$$

For arbitrary $y \in B_r$, the contraction mapping principle gives a unique $x \in \overline{B_{2r}}$ with $f(x) = y \in B_r$. That is, points $y \in B_r$ have a unique preimage $f^{-1}(y) = x$ in $\overline{B_{2r}}$.

We set $V := f^{-1}(B_r) \cap \overline{B_{2r}}$ and have the following:
- For $y \in B_r$ the preimage $f^{-1}(y) \cap V$ is unique. Hence $f\colon V \to B_r$ is injective, proving (i).
- $f(V) = B_r$ is open as asserted in (ii).
- The set V is the preimage of the open set B_r under the continuous map f. By Thm. V.18, the set V is open. Since $0 \in V$, the set V is an open neighbourhood of $a = 0$.

Let us now show that $g := f^{-1}$ is continuous. Since $F_0(x) = x - f(x)$,

$$x_2 - x_1 = F_0(x_2) - F_0(x_1) + \big(f(x_2) - f(x_1)\big)$$

we find, using (12),

$$\|x_2 - x_1\| \le \frac{1}{2}\|x_2 - x_1\| + \|f(x_2) - f(x_1)\|$$

and so, for $y_i := f(x_i) \Leftrightarrow g(y_i) = x_i$,

$$\|g(y_2) - g(y_1)\| \le 2\|y_2 - y_1\|.$$

Therefore, g is Lipschitz, and so continuous. The more difficult fact that g is a C^1-map follows from Proposition 2. We have established (*iii*).

Let us finally show that imposing (11) means no loss of generality; we reason as in Prop. 2. Given f as in the theorem, then $(df_a)^{-1}$ exists and $\tilde f(x) := df_a^{-1}\big(f(x+a) - f(a)\big)$ satisfies (11). Our proof gives that $\tilde f$ is a diffeomorphism from, say, $\tilde V \ni 0$ to $\tilde f(\tilde V) \ni 0$. But this means that $f(x) = df_a \cdot \tilde f(x-a) + \tilde f(a)$ is a diffeomorphism from $V := df_a(\tilde V + a)$ to $f(V)$. $\square$

`48. Lecture, Tuesday, 26. June 07 ________________ "U 9`

2. Implicitly defined mappings

2.1. **Implicit function theorem.** We investigate underdetermined systems of equations $f(\zeta) = z$ where $f\colon \mathbb{R}^{n+k} \to \mathbb{R}^k$ and z is given. Here, the number of unknows $n+k$ is larger than the number of equations k. The main task is to find a way to parameterize the set of solutions. We expect the solution space to be n-dimensional with codimension k in $\mathbb{R}^{n+k}$.

It is sufficient to consider the case $z = 0$, since else we can subtract z from f. Moreover, we write $\zeta = (x, y) \in \mathbb{R}^n \times \mathbb{R}^k$ with the goal to represent the solutions as a graph of the form $\big(x_1, \dots, x_n, g(x_1, \dots, x_n)\big)$. This means that we consider the first n coordinates as independent variables, while the last k are dependent ones (i.e., determined by our equation $f(x, y) = 0$). Let us demonstrate this on the linear case.

Examples. 1. For $n = 2$, $k = 1$ consider the specific linear map $f(x_1, x_2, y) := x_1 + x_2 + y$. The solution set is a plane and has the desired graph representation, namely

$$\big\{(x_1, x_2, y) \in \mathbb{R}^3 : f(x_1, x_2, y) = 0\big\} = \big\{(x_1, x_2, y) : y = g(x_1, x_2) := -x_1 - x_2\big\}.$$

2. Consider now $f\colon \mathbb{R}^{n+k} \to \mathbb{R}^k$ linear. By linearity we can decompose

$$f(x, y) = f_X(x) + f_Y(y), \quad \text{where} \quad f_X\colon \mathbb{R}^n \to \mathbb{R}^k, \;\; f_Y\colon \mathbb{R}^k \to \mathbb{R}^k.$$

Therefore: The equation $0 = f(x, y) = f_X(x) + f_Y(y)$ has solutions of type $(x, y(x)) \Leftrightarrow$ $y(x) = -f_Y^{-1} f_X(x) \Leftrightarrow$

$$f_Y \text{ is invertible} \quad \Leftrightarrow \quad \det f_Y \ne 0.$$

In the general, non-linear case we also wish to represent the solution set as an n-dimensional graph of some mapping $g\colon \mathbb{R}^n \to \mathbb{R}^k$. However, we can do this only locally:

Example. For $f(x,y) := x^2 + y^2 - 1$ the solution set $\{f(x,y) = 0\}$ is the unit circle. We need two graphs, namely the closed upper and lower semicircle, $y = g_\pm(x) := \pm\sqrt{1-x^2}$, each defined on the open interval $(-1,1)$. In a neighbourhood of the two boundary points $(x,y) = (\pm 1, 0)$ it is not true that the solution set is a graph of the type $y(x)$.

We can now solve $f(x,y) = 0$ in the form $y = g(x)$ on a small product set $X \times Y$ provided its linearization $0 = df(x,y) = df_X(x) + df_Y(y)$ at the point (a,b) is uniquely solvable for y:

Theorem 5 (Implicit function theorem *[Satz für implizite Funktionen]*).
Let $\Omega \subset \mathbb{R}^n \times \mathbb{R}^k = \mathbb{R}^{n+k}$ be open and $f\colon \Omega \to \mathbb{R}^k$, $(x,y) \mapsto f(x,y)$ be continuously differentiable. Moreover, let $(a,b) \in \Omega$ be a point with

$$f(a,b) = 0 \qquad \textit{and} \qquad \det\begin{pmatrix} \frac{\partial f_1}{\partial y_1}(a,b) & \dots & \frac{\partial f_1}{\partial y_k}(a,b) \\ \vdots & & \vdots \\ \frac{\partial f_k}{\partial y_1}(a,b) & \dots & \frac{\partial f_k}{\partial y_k}(a,b) \end{pmatrix} \neq 0. \tag{13}$$

Then there are neighbourhoods $X \subset \mathbb{R}^n$ of a and $Y \subset \mathbb{R}^k$ of b with $X \times Y \subset \Omega$, as well as a continuously differentiable mapping $g\colon X \to Y$ with

$$f(x,y) = 0 \textit{ for } (x,y) \in X \times Y \qquad \Longleftrightarrow \qquad y = g(x) \textit{ for } x \in X. \tag{14}$$

Remarks. 1. Terminology: With (14) we *solve the equation $f(x,y) = 0$ for y in terms of x*. We also say the equation $f(x,y) = 0$ *defines $y = g(x)$ implicitely.*

2. Meaning of the determinant condition in (13):
a) The condition prohibits the graph to become vertical. To see this, consider the case $k = n = 1$. We can only write the level set of f as a C^1-graph $y = g(x)$ provided the graph is not vertical, i.e., not pointing into the y-direction. This means, we want the gradient $\operatorname{grad} f(a,b) = \left(\frac{\partial f}{\partial x}(a,b), \frac{\partial f}{\partial y}(a,b)\right)^\top$ to be non-horizontal: Recall that $\operatorname{grad} f$ is perpendicular to the level set $\{f(x) = 0\}$. But the gradient condition means that $\frac{\partial f}{\partial y}(a,b) \neq 0$, which is (13). – Similarly, in arbitrary dimensions the determinant condition means that no bad "vertical" y_i-direction is tangent to the level set at (a,b).
b) The condition prohibits singular points. For example, the 0-level set of $f(x,y) := x^2 - y^2$ consists of the union of the two diagonals in $\mathbb{R}^2$. The origin is a singular point: Two one-dimensional branches meet and a graph representation is impossible. But the gradient, $\operatorname{grad} f(x,y) = 2(x,-y)$, vanishes precisely at $(0,0)$, so that (13) is violated.

3. The assumption (13) is not necessary for the conclusion of the implicit function theorem to hold. For example, the function $f(x,y) = (x^2 + y^2 - 1)^2$ has the unit circle as its zero

level set. However, the condition (13) does not hold at *any* point of the level set. At a point p of the unit circle, the Taylor series of f starts with the second order term, and so $\det J_f^Y(p) = 0$. Thus the implicit function theorem will not apply.

Proof. We will reduce the implicit function theorem to the inverse mapping theorem. To do this, we consider

$$F\colon \Omega \to \mathbb{R}^{n+k}, \qquad F(x,y) := \big(x, f(x,y)\big).$$

It maps points (x,y) in the zero level set to points $(x,0)$. We are interested in the inverse image $F^{-1}(x,0) = (x,y)$: its second component represents the desired function $g(x)$.

To check the inverse mapping theorem applies, let us write the Jacobian of f in terms of block matrices $J_f = (J_f^X \mid J_f^Y)$. Then the Jacobian of F is a $(n+k)\times(n+k)$-matrix with block representation

$$J_F(a,b) = \begin{pmatrix} \mathrm{Id}_n & 0 \\ J_f^X(a,b) & J_f^Y(a,b) \end{pmatrix}.$$

The 0 block on the top right implies

$$\det J_F(a,b) = \underbrace{\det \mathrm{Id}_n}_{=1} \det J_f^Y(a,b) \overset{(13)}{\neq} 0.$$

Thus we can apply the inverse mapping theorem to F at (a,b). It gives us neighbourhoods $U \subset \Omega$ of (a,b) and $V \subset \mathbb{R}^{n+k}$ of $F(a,b) = (a,0)$, such that $F\colon U \to V$ has a continuously differentiable inverse $G\colon V \to U$. It follows from

$$V \ni\ (x,y) = F\big(G(x,y)\big) = \Big(G_1(x,y), \ldots, G_n(x,y),\ f\big(G(x,y)\big)\Big)$$

that the first n components of G are the identity. Thus we can write $G(x,y) = \big(x, \varphi(x,y)\big)$ where $\varphi := f \circ G\colon V \to \mathbb{R}^k$. In particular, as a composition, φ is continuously differentiable.

For $(x,y) \in U$ we have

$$f(x,y) = 0 \iff F(x,y) = (x,0) \iff (x,y) = G(x,0) = \big(x, \varphi(x,0)\big) \iff y = \varphi(x,0).$$

That is, the function $g(x) := \varphi(x,0)\colon \{x \in \mathbb{R}^n : (x,0) \in V\} \to \mathbb{R}^k$, which is continuously differentiable, is the desired resolution of the equation $f(x,y) = 0$.

To conclude the proof, we must specify a neighbourhood of (a,b) in the form $X \times Y \subset U$. Clearly, upon restriction the last displayed equation is preserved, so that g also describes the zero set of f in the set $X \times Y$. But U is an open neighbourhood of (a,b) and thus the product of sufficiently small neighbourhoods of a and b is still contained in U. Moreover, g is continuous, so that a sufficiently small neighbourhood X of a is indeed mapped onto a small neighbourhood $Y := g(X)$ of b. □

The function g has the same differentiability as f. To see this, write once again $J_f = (J_f^X | J_f^Y)$ and differentiate $f\big(x, g(x)\big) = 0$. This gives $J_f^X \operatorname{Id}_n + J_f^Y J_g = 0$. Since J_f^Y is invertible we obtain

$$dg = -(df_Y)^{-1} \circ df_X \quad \text{or} \quad J_g = -(J_f^Y)^{-1} \cdot J_f^X$$

n particular, if f is a C^α-function then so is g.

`49. Lecture, Thursday, 28. June 07` ——— `T 11`

2.2. **Application: Zeros of polynomials.** We want to give an example on how the implicit function theorem is typically applied. An important issue for many problems is the continuous dependence of solutions on parameters. Clearly, any naturally arising problem should have this property (why?). Let us discuss this problem on the example of zeros of polynomials.

Examples. 1. Consider a polynomial of second order,

$$t \mapsto f(p, q, t) := t^2 + pt + q.$$

We are interested how zeros t of f depend on the coefficients p, q. The implicit function theorem says there is a local resolution $t = t(p, q)$ of the equation $f(p, q, t) = 0$, provided

$$0 \neq \frac{\partial f}{\partial t}(p, q, t) = 2t + p. \tag{15}$$

To understand the meaning of (15), note that $f(q, p, t) = \big(t - t_1(p, q)\big)\big(t - t_2(p, q)\big)$ where the zeros are represented by

$$t_{1,2}(p, q) := -\frac{p}{2} \pm \sqrt{\frac{p^2}{4} - q}.$$

We conclude

$$2t + p = 0 \quad \Leftrightarrow \quad \frac{p^2}{4} - q = 0 \quad \Leftrightarrow \quad t_1 = t_2 \quad \Leftrightarrow \quad \text{zero is double.}$$

Over $U := \{(p, q) \in \mathbb{R}^2 \colon \frac{p^2}{4} - q > 0\}$ the zero set $(p, q, t) \in \mathbb{R}^3$ is given by two surface branches which are folded together along a curve over the parabola ∂U. Note that over ∂U the zeros are double, while over $\mathbb{R}^2 \setminus \overline{U}$ there are no zeros.

2. Now consider a polynomial of degree n,

$$f(x, t) = t^n + x_{n-1} t^{n-1} + \ldots + x_1 t + x_0,$$

where $x := (x_0, \ldots, x_{n-1})$. Again we are interested in zeros $t = t(x)$ of $f(x, t) = 0$. The implicit function theorem requires $\partial_t f(x, t_0) \neq 0$ in order to obtain a resolution $t(x)$ of the zero set in a neighbourhood of a given zero t_0. The condition is equivalent to t_0

being simple: Indeed, write $f(x,t) = (t - t_0)p(t)$ with $p(t)$ a polynomial. The derivative $f'(t_0) = p(t_0)$ is nonzero if $p(t_0) \neq 0$, that is, if t_0 is a simple zero of f. Thus the implicit function theorem yields the conclusion: Simple zeros of a polynomial depend differentiably on its coefficients.

3. Submanifolds

Submanifolds are sets which locally look like $\mathbb{R}^n$, in a sense made precise by the local graph representation of the implicit function theorem. However, the focus is now on the global structure of these objects.

Many problems lend themselves to a description by n-dimensional submanifolds of $\mathbb{R}^{n+k}$. The name submanifold indicates that these are but a special case of manifolds. Indeed, there is a way to define manifolds abstractly, without being contained in an ambient space. Much of 20th century mathematics deals with the study of manifolds, a concept envisioned by Riemann as early as 1854. Manifolds play also a crucial rôle in physics, where they are often considered as constraints; their dimension is the number of *degrees of freedom*.

Examples. 1.a) A segment (or rod) of fixed length $\ell > 0$ can be described by the coordinates x and y of its endpoints. The set M of all configurations is therefore $(x, y) \in \mathbb{R}^6$ with the constraint $\sum_{i=1}^{3}(x_i - y_i)^2 - \ell^2 = 0$.
b) To get a more explicit description of M, let us take $d := y - x$ with $\|d\| = \ell$. Then all configurations are given by coordinates $(x, d) \in \mathbb{R}^3 \times \mathbb{S}^2_\ell \subset \mathbb{R}^6$.

2. What would a similar description of the space of triangles in the plane be? Or of lines in the plane?

Another way to think of a manifold is in terms of geometry. Introduced rigorously, curves or surfaces in space are one- or two-dimensional submanifolds of $\mathbb{R}^3$. The notion of an n-dimensional submanifold in $\mathbb{R}^{n+k}$ generalizes these examples to arbitrary dimenion n and codimension k.

3.1. Sets which locally look like $\mathbb{R}^n$.

Definition. Let $n \in \mathbb{N}$, $k \in \mathbb{N}_0$, and $\alpha \in \mathbb{N}$. A set $M \subset \mathbb{R}^{n+k}$ is an *n-dimensional C^α-submanifold [Untermannigfaltigkeit]* if and only if for each $p \in M$ there exists the following:
- an open neighbourhood $V \subset \mathbb{R}^{n+k}$ of p, an open set $U \subset \mathbb{R}^{n+k}$, and
- a C^α-diffeomorphism $F\colon U \to V$ such that

$$F\big(U \cap \mathbb{R}^n \times \{0\}\big) = V \cap M. \tag{16}$$

We call F a *chart [Karte]* for M. An *atlas* of M is a family of charts $\{F_i\colon U_i \to \mathbb{R}^{n+k}\colon i \in I\}$ which covers M, that is, $M \subset \bigcup_{i\in I} F_i(U_i)$.

We call k the *codimension* of M in $\mathbb{R}^{n+k}$.

Trivial examples of submanifolds are:

Examples. 1. Any open set $U \subset \mathbb{R}^n$ is an n-dimensional submanifold (with codimension $k = 0$), for instance the unit ball in $\mathbb{R}^n$ or $\mathbb{R}^n \setminus \{0\}$. (Set $V = U$ and $F = \mathrm{id}$)

2. Any n-dimensional subvectorspace of $\mathbb{R}^{n+k}$: For instance, for $\mathbb{R}^n \times \{0\}$ take $U = V = \mathbb{R}^{n+k}$ and $F = \mathrm{id}$.

A graph is a simple example of a submanifold:

Lemma 6. *Suppose $g \in C^\alpha(D, \mathbb{R}^k)$ where $D \subset \mathbb{R}^n$ is open. Then the* graph

$$\Gamma(g) = \big\{(x, g(x))\colon x \in D\big\} \subset D \times \mathbb{R}^k$$

is an n-dimensional C^α-submanifold.

Proof. Consider the mapping

$$F\colon U := D \times \mathbb{R}^k \to V := D \times \mathbb{R}^k, \qquad F(x, y) := \big(x, g(x) + y\big). \tag{17}$$

The map F is a diffeomorphism with inverse $F^{-1}(x, y) = \big(x, y - g(x)\big)$. An atlas of the graph is given by the single chart F. $\square$

We note that a submanifold can be characterized as a set which admits a covering with graphs:

Proposition 7. *A set $M \subset \mathbb{R}^{n+k}$ is an n-dimensional submanifold, if and only if at each point p the coordinates can be relabelled such that the following holds: There is an open neighbourhood $V = X \times Y \subset \mathbb{R}^{n+k}$ with $X \subset \mathbb{R}^n$, $Y \subset \mathbb{R}^k$ and there exists $g \in C^\alpha(X, Y)$ with*

$$M \cap X \times Y = \{(x, g(x)) \in \mathbb{R}^{n+k} : x \in X\}.$$

Proof. While "$\Leftarrow$" is direct consequence of the lemma, the direction "$\Rightarrow$" will follow from the implicit description of submanifolds, Thm. 8. $\square$

Example. Spheres $\mathbb{S}^n \subset \mathbb{R}^{n+1}$ are n-dimensional submanifolds. To see that we decompose the sphere into $2(n + 1)$ graphs, namely the open hemispheres

$$H_i^\pm := \{x \in \mathbb{S}^n\colon\ \pm\langle x, e_i\rangle > 0\}.$$

Each of these hemispheres is a graph of type $g(x) = \pm\sqrt{1 - |x|^2}$ over the n-dimensional disk $x \in D := B_1^n(0)$ for n out of the $n + 1$ coordinates. The $2(n + 1)$ hemispheres $H_i^\pm$

cover $\mathbb{S}^n$ and so by Prop. 7 the sphere $\mathbb{S}^n$ is submanifold. Our atlas consists of $2(n+1)$ charts, each as in (17).

Submanifolds have the following properties:
• They are C^α-differentiable and hence have no edges, vertices or tips.
• They do not have self-intersections: The union of the x- and y-axis is not a submanifold.
• They keep a distance to themselves: The union $\{0\} \times (-1,1) \cup (0,1) \times \{0\}$ of two open intervals is not a submanifold, and neither is the union of the y-axis with the graph of $\sin \frac{1}{x}$ for $x > 0$.
• They do not contain their "boundaries": The open ball $B_r(0)$ is a submanifold, but the closed ball $\overline{B_r}(0)$ is not.
• They need not be connected.

It means a little work to prove that some of the counterexamples we mentioned in the previous list are in fact not submanifolds. This work will simplify considerably when we have the implicit definition at hand, given in the next subsection.

Remark. Let us mention a recently solved famous problem in topology. It can be accurately presented in our language. All we need in addition is the notion of *simply connectivity*: A set X is simply connected if each closed curve can be contracted to a point; the contraction is via a continuous one-parameter family of curves. For a submanifold M, it is important to note that the entire family of curves must stay in M. For instance, $\mathbb{R}^n$ or for $\mathbb{S}^{n-1}$ are simply connected, but a torus of revolution (or a surface with more "holes") is not. It turns out that that the only two-dimensional submanifold which is simply connected is $\mathbb{S}^2$, in the sense that any other submanifold is homeomorphic to $\mathbb{S}^2$. Here the codimension k may be arbitrary. In 1904, Poincaré, formulated the same assertion for dimension $n = 3$ as a problem: *Is any three-dimensional submanifold, which is simply connected, homeomorphic to* $\mathbb{S}^3$*?* He commented, *Mais cette question nous entrainerait trop loin.* Indeed, it took a hundred years till the Russian mathematician Perelman, building on ideas of the American Hamilton, gave a proof of the conjecture. Its validity is accepted in the mathematical community and Perelman was awarded the Fields medal, the mathematician's equivalent of the Nobel price. Nevertheless, he did not appear in the ceremony in Madrid 2006 to accept the price. The Poincaré conjecture belongs to the seven millenium problems, for whose published solution a prize of a million dollar has been set out, see: `www.claymath.org/millennium/Poincare_Conjecture/`

`50. Lecture, Tuesday, 3. July 07` ________________________________ `"U 10`

3.2. **Implicitly defined submanifolds.** Level sets of a single mapping provide a distinguished example of submanifolds. The implicit function theorem assumes a rank condition on the Jacobian at one point, see (13). In order to require this condition for an entire submanifold, let us define:

Definition. Let $f \in C^1(U, \mathbb{R}^k)$, where $U \subset \mathbb{R}^{n+k}$ and $k, n \in \mathbb{N}$.
(i) $x \in U$ is a *critical point*, if $\operatorname{rank} df_x < k$.
(ii) $y \in \mathbb{R}^k$ is a *regular value* if $f^{-1}(y)$ contains no critical points. Else, y is a *critical value.*

Let us note that a critical value may also have preimages which are not critical. Also, a value which is not attained is regular by definition(!).

Examples. 1. For $x \mapsto \|x\|^2$, all $x = 0$ is the only critical point, and $y = 0$ is the only critical value.

2. For a constant map $f(x) = c$ all points x in the domain are critical, but all values except for c are regular.
3. Take a map of the earth and consider the height function. A height is regular, if there are no summits, sinks, or passes with that particular height. For regular heights, the respective level set (height lines) contains no crossings or isolated points.

As a direct consequence of the implicit function theorem, level sets for regular values are submanifolds:

Theorem 8. *$M \subset \mathbb{R}^{n+k}$ is an n-dimensional C^α-submanifold if and only if there exists for each $p \in M$ an open neighbourhood V and $h \in C^\alpha(V, \mathbb{R}^k)$ with regular value $q \in \mathbb{R}^k$ such that $M = h^{-1}(q)$.*

Proof. "$\Leftarrow$". Pick $p \in h^{-1}(q)$. Since $\operatorname{rank} dh_p = k$, the Jacobian $J_h(p)$ contains k linearly independent columns. Relabel the coordinates in V so that these become the last k coordinates. Then the implicit function theorem applies to give a neighbourhood $V := X \times Y$ of p such that $h^{-1}(q) \cap X \times Y$ is graph of the form $(x, g(x))$. Thus Lemma 6 gives a chart containing p, and hence M is submanifold.

"$\Rightarrow$": For $p \in M$, let F be a chart and consider the diffeomorphism F^{-1}, mapping $M \cap V$ into $\mathbb{R}^n \times \{0\}$. We project F^{-1} to $\mathbb{R}^k$, that is, we consider the last k components of F^{-1}: The map $h \colon V \to \mathbb{R}^k$ with $h_1 := (F^{-1})_{n+1}, \ldots, h_k := (F^{-1})_{n+k}$ maps M to 0.

Note that F^{-1} is a diffeomorphism and so $\operatorname{rank} d(F^{-1}) = n + k$, that is, all $n + k$ component vectors of $d(F^{-1})$ are linearly independent. In particular dh has rank k. ☐

Examples. 1. The spheres $\mathbb{S}^n$ are the preimage of the regular value 1 of the function $\|x\|^2 \colon \mathbb{R}^{n+1} \to \mathbb{R}$ and therefore submanifolds.

2. Take $f \colon \mathbb{R}^2 \to \mathbb{R}$, $f(x, y) = x^2 - y^2$ with $\operatorname{grad} f(x, y) = 2(x, y)$. Thus $(0, 0)$ is the only

critical point, and $t = 0$ the only critical value. The preimages are submanifolds for $t \neq 0$, namely hyperbolas such as $\{(x, y) : y = \pm\sqrt{x^2 - t}\}$. For the singular case $t = 0$, the level set is the union of the two diagonals (not a submanifold).

3. Consider a quadratic form $x \mapsto Q(x) := x^\top A x$ where A is symmetric. Suppose $A \in \mathsf{GL}(n)$. Then $\operatorname{grad} Q(x) = 2Ax$ is nonzero for $x \neq 0$, that is, all values $y \neq 0$ are regular. Thus $Q^{-1}(y)$ for $y \neq 0$ is an n-dimensional submanifold of $\mathbb{R}^{n+1}$. For $A = \mathrm{Id}_n$, the submanifolds are spheres of radius $\sqrt{y}$, for A diagonal with positive entries they are ellipsoids, etc.

4. We regard $n \times n$-matrices $\mathsf{M}(n)$ as points of $\mathbb{R}^{n^2}$. We claim that the *orthogonal group*

$$\mathsf{O}(n) = \{A \in \mathsf{M}(n) : A^\top A = \mathrm{Id}\},$$

is a submanifold with dimension $N := \frac{1}{2}n(n-1)$ in the space $\mathsf{M}(n) = \mathbb{R}^{n^2} = \mathbb{R}^{N+K}$ (its codimension is $K := \frac{1}{2}n(n+1)$). To see this, we will assert that $\mathsf{O}(n)$ is defined implicitly: It is the level set of the function $f(A) = A^\top A$, for which Id is a regular value. To make this rigorous, consider the space of *symmetric matrices*

$$\mathsf{Sym}(n) := \{A \in \mathsf{M}(n) : A^\top = A\} = \mathbb{R}^{n(n+1)/2}$$

and the map

$$f\colon \mathsf{M}(n) = \mathbb{R}^{n^2} = \mathbb{R}^{N+K} \to \mathsf{Sym}(n) = \mathbb{R}^{n(n+1)/2} = \mathbb{R}^K, \qquad A \mapsto A^\top A;$$

indeed $f(A) \in \mathsf{Sym}(n)$ thanks to $(A^\top A)^\top = A^\top A^{\top\top} = A^\top A$. Now $\mathsf{O}(n) = f^{-1}(\mathrm{Id})$. To apply Thm. 8, we need to compute the differential of f and show that $\mathrm{Id} \in \mathsf{Sym}(n)$ is a regular value. To compute df, pick $H \in \mathbb{R}^{n^2}$ and calculate

$$f(A + tH) = (A + tH)^\top (A + tH) = A^\top A + t(H^\top A + A^\top H) + t^2 H^\top H.$$

Thus

$$df_A(H) = H^\top A + A^\top H.$$

To show that Id is a regular value, we prove that for $A \in \mathsf{O}(n)$ the mapping $H \mapsto df_A(H)$ is surjective. Indeed, for $B \in \mathsf{Sym}(n)$ arbitrary we have

$$df_A\Big(\frac{1}{2}AB\Big) = \frac{1}{2}B^\top A^\top A + \frac{1}{2}A^\top AB = B.$$

Hence Thm. 8 asserts that $\mathsf{O}(n)$ is a submanifold of $\mathbb{R}^{n^2}$; its dimension is

$$\dim \mathsf{M}(n) - \dim \mathsf{Sym}(n) = (N + K) - K = N = \frac{1}{2}n(n-1).$$

3.3. **Tangent and normal space to a submanifold.**

Definition. Let $M \subset \mathbb{R}^{n+k}$ be an n-dimensional submanifold and $p \in M$.
(i) A vector $v \in \mathbb{R}^{n+k}$ is a *tangent vector to M at p*, if for some $\varepsilon > 0$ there is a differentiable curve $c\colon (-\varepsilon, \varepsilon) \to M$ with $c(0) = p$ and $c'(0) = v$. The *tangent space* $T_p(M)$ is the set of all tangent vectors to M at p.
(ii) The *normal space* is the set

$$N_pM := T_pM^\perp = \bigl\{\nu \in \mathbb{R}^{n+k} : \langle \nu, w\rangle = 0 \text{ for all } w \in T_p(M)\bigr\}.$$

Its elements are *normal vectors to M at p.*

Thus at each point $p \in M$ we have the orthogonal decomposition of subvectorspaces

$$T_pM \oplus N_pM = \mathbb{R}^{n+k}.$$

Before giving examples, let us specify dimensions:

Proposition 9. *Let p be a point of a manifold M, contained F. Then*

$$T_pM = dF_x\bigl(\mathbb{R}^n \times \{0\}\bigr) \subset \mathbb{R}^{n+k}, \qquad \text{where } x := F^{-1}(p). \tag{18}$$

Consequently, $\dim T_pM = n$ *and* $\dim N_pM = k$.

Proof. To prove (18) let $v \in \mathbb{R}^n \times \{0\}$ and $c(t) := x + tv$ be a curve in $\mathbb{R}^n \times \{0\}$. The image curve $C(t) := F\bigl(c(t)\bigr)$ in M is defined for small $|t|$ and gives us a tangent vector $C' = dF(c') = dF(v)$. This implies $dF_x\bigl(\mathbb{R}^n \times \{0\}\bigr) \subset T_pM$. We can turn this argument around: If C is an arbitrary curves in M, then its preimage $F^{-1}(C)$ is a curve in $\mathbb{R}^n \times \{0\}$. Hence any tangent vector in T_pM arises in the form $dF_x(v)$, and so $dF_x\bigl(\mathbb{R}^n \times \{0\}\bigr) \supset T_pM$.

Let us now verify the claimed dimensions. The differential dF_x has rank $n + k$ and so $\ker dF_x = \{0\}$. Thus the image of the n-dimensional space $\mathbb{R}^n \times \{0\}$ under dF_x is again n-dimensional, as claimed. By definition, N_pM is the orthogonal complement of N_pM, and hence it has dimension $(n + k) - n = k$ as claimed. □

Examples of tangent and normal spaces: 1. *Subspaces* $M := \mathbb{R}^n \times \{0\} \subset \mathbb{R}^{n+k} = \{(x, y)\}$. In this case, T_pM and N_pM are independent of p: $T_p\mathbb{R}^n = \{(x, 0)\} = M$ and $N_p\mathbb{R}^n = \{(0, y)\}$. What is the description for a subspace in general position?

2. *Spheres* $\mathbb{S}^n \subset \mathbb{R}^{n+1}$: For $p \in \mathbb{S}^n$ consider the n-dimensional subspace

$$p^\perp := \{v \in \mathbb{R}^{n+1} : \langle v, p\rangle = 0\}.$$

If $v \in p^\perp$ then $c(t) := \cos\big(\|v\|t\big)p + \sin\big(\|v\|t\big)\frac{v}{\|v\|}$ is a great circle to $\mathbb{S}^n$ through $c(0) = p$ with tangent vector $c'(0) = v$ (chain rule). Thus $p^\perp \subset T_pM$. But the left hand side is n-dimensional, and so is the right hand side by Prop. 9. Hence

$$T_p\mathbb{S}^n = p^\perp = \{v \in \mathbb{R}^{n+1} : v \perp p\} \quad \text{and} \quad N_p\mathbb{S}^n = \{sp : s \in \mathbb{R}\}.$$

Given a submanifold in implicit description, what are the tangent and normal vectors?

Theorem 10. *Suppose that an n-dimensional submanifold of $M \subset \mathbb{R}^{n+k}$ is given locally by the level $\{p \in U \subset \mathbb{R}^{n+k} : f(p) = 0\}$ of a function $f \in C^1(U, \mathbb{R}^k)$ with 0 a regular value and $U \subset \mathbb{R}^{n+k}$ open. Then*

$$T_pM = \ker df_p \quad \textit{and} \quad N_pM = \operatorname{span}\big\{\operatorname{grad} f_1(p), \dots, \operatorname{grad} f_k(p)\big\}. \tag{19}$$

Proof. Let $v \in T_pM$. There is $c(t)$ with $c(0) = p$ and $c'(0) = v$. Then $f \circ c \equiv 0$, and so by the chain rule $df_{c(0)}(v) = 0$ or $v \in \ker df_p$. This proves

$$T_pM \subset \ker df_p.$$

But both sides are n-dimensional: For T_pM this follows from Prop. 9, for $\ker df_p$ since $\operatorname{rank} df_p = k$. Hence the spaces agree.

`51. Lecture, Thursday, 5. July 07` ________________________________ `T 12`

To exhibit N_pM, regard the matrix product J_fv as a scalar product with the rows of the Jacobian:

$$J_fv = \begin{pmatrix} (\operatorname{grad} f_1)^\top \\ \vdots \\ (\operatorname{grad} f_k)^\top \end{pmatrix} v = \begin{pmatrix} \langle \operatorname{grad} f_1, v\rangle \\ \vdots \\ \langle \operatorname{grad} f_k, v\rangle \end{pmatrix}$$

Consequently, $\ker J_f(p)$ is orthogonal to each of the k row vectors of $J_f(p)$. By the regular value assumption the rows are linearly independent and hence they form the orthogonal complement of T_pM, that is, they span the k-dimensional space N_pM. $\square$

Let us consider the special case of graphs

$$M = \big\{\big(x, g(x)\big) : x \in U\big\} \qquad \text{where} \qquad g \in C^\alpha(U, \mathbb{R}).$$

Then M can be considered an implicitely defined submanifold since

$$h \in C^\alpha(U \times \mathbb{R}, \mathbb{R}), \qquad h(x, y) := y - g(x)$$

has 0 as a regular value, due to

$$\operatorname{grad} h(x, y) = \begin{pmatrix} -\operatorname{grad} g(x) \\ 1 \end{pmatrix} \neq 0. \tag{20}$$

Thus $M = h^{-1}(0)$ is a submanifold. By (19), the normal space is spanned by (20); a unit normal (the upper one) would be

$$\nu\big(x, g(x)\big) = \frac{1}{\sqrt{1 + \|\operatorname{grad} g(x)\|^2}} \begin{pmatrix} -\operatorname{grad} g(x) \\ 1 \end{pmatrix}.$$

The tangent space can now be constructed as the orthogonal complement to $\operatorname{grad} h(x, y)$ as

$$T_{(x,g(x))}M = \operatorname{span}\left\{ \begin{pmatrix} e_1 \\ \partial_1 g(x) \end{pmatrix}, \dots, \begin{pmatrix} e_n \\ \partial_n g(x) \end{pmatrix} \right\}.$$

Examples. 1. Consider the paraboloid of revolution $M = \{(x, y, x^2 + y^2)), (x, y) \in \mathbb{R}^2\}$. The normal and tangent spaces in $p = (x, y, x^2 + y^2) \in M$ are

$$N_pM = \operatorname{span}\left\{ \begin{pmatrix} -2x \\ -2y \\ 1 \end{pmatrix} \right\}, \quad T_pM = \operatorname{span}\left\{ \begin{pmatrix} 1 \\ 0 \\ 2x \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 2y \end{pmatrix} \right\}.$$

2. In the case of codimension $k = 1$, a *unit normal* is a continuous mapping $\nu\colon M \to \mathbb{S}^{n-1}$ such that $\nu(p) \in N_pM$. This will not always exist, however: A *Möbius strip* M is not *orientable*, meaning that there is no globally defined unit normal on the surface. This implies, for instance, that it cannot be described as the level set of a single function. Indeed, if such a function existed, then $M = f^{-1}(0)$ had the unit normal $\operatorname{grad} f / \|\operatorname{grad} f\|$. Another way to view this example is to consider the sign of f to the two sides of an implicity defined submanifold $M^n \subset \mathbb{R}^{n+1}$: It is positive to one side of M and negative to the other. But for the Möbius strip, there is only one side, contradiction. This example indicates that it makes sense to admit for an implicit description of a submanifold more than one function.

4. Constrained extrema

Often one seeks the extrema of a function $f(x)$ under a *constraint [Nebenbedingung]* such as $x \in \mathbb{S}^{n-1}$ or more generally on a submanifold M. The following necessary condition for an extremum usually serves to determine it:

Theorem 11. *Let M be a submanifold and suppose*

$$f \in C^1(U, \mathbb{R}) \text{ attains a } \left\{ \substack{\text{maximum} \\ \text{minimum}} \right\} \text{ at } p \in M \quad :\Leftrightarrow \quad f(p) \left\{ \substack{\geq \\ \leq} \right\} f(x) \quad \textit{for all } x \in M.$$

Then $\operatorname{grad} f(p) \in N_pM$*, that is, the tangential part of* $\operatorname{grad} f(p)$ *vanishes.*

This is a necessary condition for extrema. In fact, a point with vanishing tangential part, $(\operatorname{grad} f(p))^{\tan} = 0$, could be called a *critical point on M*, and this may be a maximum, minimum or a saddle point of the constrained problem.

Proof. To prove $\operatorname{grad} f(p) \in N_pM$ amounts to showing that $\operatorname{grad} f(p)$ is perpendicular to each $v \in T_pM$. By definition, there exists a differentiable curve $c(t)$ through $p = c(0)$ with tangent $c'(0) = v$. But since $t \mapsto f \circ c(t)$ attains an extremum at $t = 0$, we have

$$0 = \frac{d}{dt} f\big(c(t)\big)\Big|_{t=0} = \big\langle \operatorname{grad} f\big(c(0)\big), c'(0)\big\rangle = \big\langle \operatorname{grad} f(p), v\big\rangle.$$

Thus $v \perp T_pM$, as desired. □

A normal vector, such as $\operatorname{grad} f(p) \in N_pM$ can be written as a linear combination of a basis for the normal space at N_pM. In Thm. 10 we gave with (19) a basis representation for the normal space of an implicitely defined submanifold. It yields the following useful statement:

Corollary 12 (Lagrange multipliers, Euler 1744, Lagrange 1788)**.** *Let $U \subset \mathbb{R}^{n+k}$ is open and $f \in C^1(U, \mathbb{R})$. Suppose f attains an extremum at p over a submanifold M which is given implicitely,*

$$M := \{x \in U : h(x) = 0\} \qquad \text{where } h \in C^1(U, \mathbb{R}^k) \text{ and } 0 \text{ is a regular value of } h.$$

Then there exist k numbers $\lambda_1, \dots, \lambda_k \in \mathbb{R}$, called Lagrange multipliers*, such that*

$$\operatorname{grad} f(p) = \lambda_1 \operatorname{grad} h_1(p) + \ldots + \lambda_k \operatorname{grad} h_k(p). \tag{21}$$

In the codimension 1 case, (21) says $\operatorname{grad} f(p) = \lambda \operatorname{grad} g(p)$. Suppose that $\operatorname{grad} f(p) \neq 0$. Then the maximal level $N := \{x \in U : f(x) = f(p)\}$ of f is also a submanifold of codimension 1 in a neighbourhood of p. Since the gradients represent normal vectors, the Lagrange mulitplier rule says that M and N touch tangentially.

Remarks. 1. A typical application of the Corollary is as follows. Suppose we know that a function takes a maximum on M. For instance, the theorem of the maximum may be applied to the closure $\overline{M}$, and it can be shown that the maximum is in fact a point of M. If, on the other hand, there exists precisely one p satisfying (21) then this must be the maximum.

2. When extrema over closed sets are seeked, the methods of the Lagrange multipliers can be combined with the methods to locate interior extrema (Sect. V.2.3).

Let us exemplify the previous remarks on a simple case:

Example. We determine all extrema of $f(x,y) := \frac{1}{2}x^2 - y$ on the closed unit disk $\overline{B_1}(0) := \{(x,y) \in \mathbb{R}^2 : x^2 + y^2 \le 1\}$. Clearly, $\operatorname{grad} f(x,y) = (x,-1)$ and so f has no critical points. Thus there are no extrema in the interior of B. It remains to determine the extrema on the boundary $M = \mathbb{S}^1 := \{(x,y) \in \mathbb{R}^2 : h(x,y) := x^2 + y^2 - 1 = 0\}$, where 0 is a regular value of h. Hence, at an extremum $(x,y) \in M$ of f, by (21) the vectors $\operatorname{grad} f(x,y)$ and $\operatorname{grad} h(x,y)$ are parallel, and so

$$0 = \det\Big(\operatorname{grad} f, \operatorname{grad} h\Big)(x,y) = \det\begin{pmatrix} x & 2x \\ -1 & 2y \end{pmatrix} = 2xy + 2x \iff x = 0 \text{ or } y = -1.$$

This condition holds for $(x,y) \in \mathbb{S}^1$ precisely at $(x,y) = (0,\pm 1)$. On the other hand, $M = \mathbb{S}^1$ is compact, and there must be at least two extrema (Thm. V.19), namely a maximum and a minumum. This identifies the two points $(0,\pm1)$ as minimum and maximum. Consequently, $f(0,1) = -1$ is the minimum of f over $\overline{B_1}(0)$ and $f(0,-1) = 1$, so that $(0,1)$ is the minimum of f and $(0,-1)$ the maximum. These are all extrema of f on In this case we can verify the result explicitely, since the levels of f are the sets $N_c := \{y = \frac{1}{2}x^2 - c\}$ (check!).

`52. Lecture, Tuesday, 10. July 07` ________________ `"U 11`

4.1. **Application: Principal axis transformation.** We apply the theorem on extrema under a constraint to construct eigenvalues. Let $A \in \mathsf{Sym}(n)$ be a symmetric $n \times n$-matrix, and $Q(x) = x^\top Ax$ be the associated quadratic form on $\mathbb{R}^n$. We have $\operatorname{grad} Q(x) = 2Ax$. We want to study extrema of Q over the submanifold $\mathbb{S}^{n-1} = \{x \in \mathbb{R}^n : h(x) = \|x\|^2 - 1 = 0\}$, where $\operatorname{grad} h(x) = 2x$.

Since $\mathbb{S}^{n-1}$ is compact and Q continuous, the maximum value theorem, Thm. V.19, gives us an extremum $v \in \mathbb{S}^{n-1}$ of Q. Then the theorem on constrained extrema yields $\lambda \in \mathbb{R}$ with

$$\operatorname{grad} Q(v) = \lambda \operatorname{grad} h(v) \qquad \Leftrightarrow \qquad 2Av = 2\lambda v,$$

that is, we have established an eigenvalue equation $Av = \lambda v$ for the matrix A. Thus we have established the existence of one real eigenvalue λ, with eigenvector v for the matrix $A \in \mathsf{Sym}(n)$!

Let us now show more generally:

Theorem 13 (Principal axis transformation/Hauptachsentransformation).
Let $A \in \mathsf{Sym}(n)$ be a symmetric $n \times n$-matrix.
(i) There is an orthonormal system of n eigenvectors $v_1, \ldots, v_n$ for A.

(ii) The respective eigenvalues $\lambda_1 \geq \ldots \geq \lambda_n$ are real. Moreover, λ_k is the maximum of $Q(x) := x^\top Ax$ over

$$M_k := \mathbb{S}^{n-1} \cap \operatorname{span}\{v_1, \ldots, v_{k-1}\}^\perp \qquad \text{for each } 2 \leq k \leq n,$$

or over all of $\mathbb{S}^{n-1}$ for $k = 1$.

Proof. By induction. We constructed $v_1 := v$ and $\lambda_1 := \lambda$ above, which is the base case. For the step $1 \leq k-1 \mapsto k \leq n$, we suppose that $v_1, \ldots, v_{k-1}$ are pairwise orthogonal eigenvectors to eigenvalues $\lambda_1, \ldots, \lambda_{k-1}$. The space M_k is compact, and hence by the maximum value theorem, Q takes a maximum at v_k over M_k.

To represent M_k implicitely as $h^{-1}(0)$, we impose the k constraints

$$h_1(x) := \langle v_1, x\rangle = 0, \quad \ldots, \quad h_{k-1}(x) := \langle v_{k-1}, x\rangle = 0$$
$$h_k(x) := \|x\|^2 - 1 = 0.$$

The Jacobian of $h := (h_1, \ldots, h_k) \colon \mathbb{R}^n \to \mathbb{R}^k$ is

$$J_h(x) = \begin{pmatrix} (\operatorname{grad} h_1)^\top \\ \cdots \\ (\operatorname{grad} h_{k-1})^\top \\ (\operatorname{grad} h_k)^\top \end{pmatrix} = \begin{pmatrix} v_1^\top \\ \cdots \\ v_{k-1}^\top \\ 2x^\top \end{pmatrix}.$$

For $x \in M_k$, all rows are orthogonal. Therefore $0 \in \mathbb{R}^k$ is a regular value of h and $M_k = h^{-1}(0)$ is a submanifold.

We may apply Thm. 11 on constrained extrema. By (21) there exist Lagrange multipliers $\mu_1, \ldots, \mu_{k-1}, \lambda_k$ with $\operatorname{grad} Q(v_k) = \sum_{i=1}^{k-1} \mu_i \operatorname{grad} h_i(v_k) + \lambda_k \operatorname{grad} h_k(v_k)$, that is,

$$2Av_{k+1} = \mu_1 v_1 + \ldots + \mu_{k-1} v_{k-1} + \lambda_k 2 v_k. \tag{22}$$

Let now $i \in \{1, \ldots, k-1\}$ arbitrary. Due to the symmetry of A,

$$\langle Av_k, v_i\rangle = v_k^\top A^\top v_i = v_k^\top A v_i = \langle v_k, Av_i\rangle = \langle v_k, \mu_i v_i\rangle = 0.$$

Hence

$$\mu_i = 2\lambda_k \langle v_k, v_i\rangle + \mu_i \overset{(22)}{=} 2\langle Av_k, v_i\rangle = 0$$

and (22) reduces to the desired eigenvalue equation $Av_k = \lambda_k v_k$.

Since $Q(v_k) = v_k^\top A v_k = \lambda_k \|v_k\|^2 = \lambda_k$ we see that the value of the maximum is λ_k. But $M_k \subsetneq M_{k-1}$, and so the maximum λ_k over the smaller set M_k is less or equal than the maximum λ_{k-1} over the larger set M_{k-1}. $\square$

Summary

The topic of nonlinear equations comprises various topics. In the case of mappings from n-dimensional into n-dimensional space, the inverse mapping theorem yields the local invertibility of a map under the condition that its linearisation is invertible. From the facts being used in the proof the contraction mapping theorem is perhaps the most important ingredient.

For more unknows than equations, the theory becomes more delicate. We want to analyse the level sets of a mapping $f\colon \mathbb{R}^{n+k} \to \mathbb{R}^k$, which are normally n-dimensional. We first studied the local case. As an application of the inverse mapping theorem, the implicit mapping theorem yields a graph representation of the level set, in the neighbourhood of any point for which $\operatorname{rank} df_p = k$.

When the last condition holds over the entire level set of f, it can be covered with graphs in suitable directions. We saw that such a set is regular in the following sense: It looks like a bent image of $\mathbb{R}^n$, as in the definition of a submanifold.

There are two aspects of submanifolds:
• *Locally* a submanifold is a nice regular set. Thus sufficiently small pieces can be described as bent Euclidean space, graph, or implicitly. Due to the local regularity, normal and tangent space become well-defined.
• *Globally* we do not want to impose any restrictions. Indeed, usually the main problem is to understand what the space looks like globally.

We used the analysis of the normal space to analyse the extrema of functions on submanifolds. This has many applications, as boundary sets of domains of functions can usually be described as a submanifold or the union of these.

Part 8. Integration in $\mathbb{R}^n$

`53. Lecture, Thursday 12. July 07` ________________ `T 13`

Like in one dimension, we do not expect that every "wild" function of several variables is integrable. Moreover, there are sets "wild" enough such that not even a constant function is integrable. It is not easy to exhibit these examples explicitely, but since the 1900's it is clear that many paradoxa emerge, once it is assumed that the integration problems would be solvable for arbitrary functions and domains.

There are, perhaps, three different integrals to consider:
• The *Lebesgue integral* is the most general integral. Under rather weak assumptions, limits of integrable functions are integrable once again. To arrive at such a mathematically pleasing integral takes, however, a larger effort. We will go this way in the fourth term.
• A *Riemann integral* can also defined in several variables. It is sufficient for most purposes of calculation. It still needs some effort for definition, and so exceeds the resources we can provide here.
• As a direct consequence of the one-dimensional integral, we will define in the present section the integration for *continuous functions with compact support*. This is not good enough for many desirable calculations, but it serves to demonstrate the main properties of the integral. Moreover, since continuous functions with compact support are "dense" within the Lebesgue integrable functions, the properties can be shown to extend.

Another important integration topic is path integration. This will be covered in the class on complex analysis.

1. Iterated volume integrals

We can use the one-dimensional Riemann-integral to calculate integrals over multi-dimensional domains. As domains we will only consider the product of intervals:

Definition. (i) Let $n \in \mathbb{N}$ and $a = (a_1, \dots, a_n), b = (b_1, \dots, b_n) \in \mathbb{R}^n$. We call the set

$$[a, b] := [a_1, b_1] \times \cdots \times [a_n, b_n] \subset \mathbb{R}^n$$

a *rectangular box [Quader]*. Its volume is $\operatorname{vol}([a, b]) := (b_1 - a_1) \cdot \ldots \cdot (b_n - a_n)$.
(ii) Suppose $f\colon [a, b] \to \mathbb{R}$ is continuous. Then the *(iterated Riemann-)integral* of f is defined by

$$\int_{[a,b]} f(x)\, dx := \int_{a_n}^{b_n} \Big(\cdots \Big(\int_{a_1}^{b_1} f(x_1, \dots, x_n)\, dx_1 \Big) \dots \Big) dx_n. \tag{1}$$

Example.

$$\int_{[0,1]^2} x^2 + y \, dx \, dy = \int_0^1 \left(\int_0^1 x^2 + y \, dx \right) dy = \int_0^1 \left[\frac{1}{3} x^3 + yx \right]_{x=0}^{x=1} dy$$
$$= \int_0^1 \frac{1}{3} + y \, dy = \frac{1}{3} y + \frac{1}{2} y^2 \Big|_0^1 = \frac{5}{6}$$

Interpretation 1. We can regard (1) as the $(n+1)$-dimensional oriented volume content under the graph. Specifically we can set for $f \colon [a,b] \to [0,\infty)$

$$\operatorname{vol}\{(x,y) : x \in [a,b], 0 \le y \le f(x)\} := \int_{[a,b]} f(x) \, dx$$

2. If $f \colon [a,b] \to [0,\infty)$ represents a mass density then $\int_{[a,b]} f(x) \, dx$ is its total mass.
3. Then $\frac{1}{\operatorname{vol}([a,b])} \int_{[a,b]} f(x) \, dx$ is the mean value *[Mittelwert]* of f over $[a,b]$.

Definition (1) poses two problems:
1. Is the integral (1) well-defined, that is, the result of the innermost integral gives a continous function of the parameter $(x_2, \ldots, x_n)$, and hence can be integrated again; etc.? This will be answered in the next subsection.
2. Is the integral independent of the order of definition? This will be answered in 1.3.

1.1. **Parameter dependent integrals.** To show that the result of a single integration gives a continuous function which can be integrated further, we will establish uniform continuity of the function to be integrated.

So let us first generalize uniform continuity to metric spaces.

Definition. Let X, Y be metric spaces. A mapping $f \colon X \to Y$ is *uniformly continuous on* X, if for each $\varepsilon > 0$ there is a $\delta = \delta(\varepsilon) > 0$, such that

$$\text{if } x, a \in X \text{ satisfy } d(x,a) < \delta \quad \text{then} \quad d\big(f(x), f(a)\big) < \varepsilon. \tag{2}$$

We know that on normed spaces the function $(x,y) \mapsto \|x-y\|$ is continuous (why?). Let us now show the same for metric spaces: If (X,d) is a metric space then $d \colon X \times X \to [0,\infty)$ is continuous, in the sense that

$$d(x_n, x) \to 0, \quad d(y_n, y) \to 0 \quad \Rightarrow \quad d(x_n, y_n) \to d(x,y).$$

The following estimate verifies this claim:

$$\big|d(x_n, y_n) - d(x,y)\big| \le \Big|\big(d(x_n, x) + d(x,y) + d(y, y_n)\big) - d(x,y)\Big| = \big|d(x_n, x) + d(y, y_n)\big| \to 0$$

In Thm.IV.18 we saw that real valued functions from a closed bounded interval are uniformly continuous. Let us now generalize this fact to say that continuous maps of compact spaces are uniformly continuous.

Theorem 1. *Suppose X and Y are metric spaces, where X is compact. Then each continuous mapping, $f\colon X \to Y$, is uniformly continuous.*

Proof. Suppose on the contrary that f is not uniformly continuous. Then, for some $\varepsilon > 0$ no $\delta > 0$ will satisfy (2). In particular, for this ε no $k \in \mathbb{N}$ will give a $\delta := \frac{1}{k}$ satisfying (2). Thus there are pairs of points $x_k, a_k \in X$ violating (2), that is,

$$d\big(f(x_k), f(a_k)\big) \geq \varepsilon \quad \text{for } d(x_k, a_k) < \frac{1}{k}. \tag{3}$$

Since X is compact, the Theorem of Bolzano-Weierstrass (Thm. V.13) allows us to pick a convergent subsequence $(x_{k_j})_{j\in\mathbb{N}}$ of (x_k), with $x_{k_j} \to x \in X$ as $j \to \infty$. We claim that also $a_{k_j} \to x$. Indeed, $d(x_k, a_k) < \frac{1}{k}$ and so

$$d(x, a_{k_j}) \leq \underbrace{d(x, x_{k_j})}_{\to 0} + \underbrace{d(x_{k_j}, a_{k_j})}_{\leq 1/k_j \to 0} \to 0.$$

But since f is continuous at the point x we have $f(x_{k_j}), f(a_{k_j}) \to f(x)$. The continuity of d implies $d\big(f(x_{k_j}), f(a_{k_j})\big) \to \big(f(x), f(x)\big) = 0$, in contradiction with (3). $\square$

We can now assert the continuous dependence of integrals on parameters.

Lemma 2. *Let $Y \subset \mathbb{R}^{n-1}$ be open, $a, b \in \mathbb{R}$ and $f\colon [a,b] \times Y \to \mathbb{R}$ be continuous. Then the function $F\colon Y \to \mathbb{R}$, $F(y) := \int_a^b f(x,y)\,dx$ is continuous.*

Proof. Pick $y \in Y$, and let $\overline{B_r(y)} \subset Y$. Now let $\varepsilon > 0$. We show there is $\delta = \delta(y) > 0$ such that if $\|h\| < \delta$ and $y + h \in Y_0$ then $\big|F(y+h) - F(y)\big| < \varepsilon$.

By the theorem, f is uniformly continuous on $[a,b] \times \overline{B_r(y)}$. Hence for points in this set there is $0 < \delta = \delta(y) \leq r$ with

$$\big\|(x', y') - (x, y)\big\| < \delta \quad \Rightarrow \quad \big\|f(x', y') - f(x, y)\big\| < \frac{\varepsilon}{b-a}.$$

Specializing to $x' = x$ and $y' = y + h$ where $\|h\| < \delta$ gives

$$\big|F(y+h) - F(y)\big| = \Big|\int_a^b f(x, y+h) - f(x,y)\,dx\Big| \leq \int_a^b \big|f(x, y+h) - f(x,y)\big|\,dx < \varepsilon.$$

$\square$

1.2. **Differentiation under the integral.** Our definition (1) raises another question: Is the value of the iterated integral independent of the order of integration? Since we want to invoke the fundamental theorem later, we wish to differentiate a parameter-dependent integral with respect to the parameter. This is a common problem in many other contexts, and so we formulate our result as a theorem:

Theorem 3 (Differentiation under the integral)**.** *Let* $f\colon [a,b]\times[c,d] \to \mathbb{R}$, $(x,y)\mapsto f(x,y)$, *be continuously partially differentiable with respect to* y*. Then the map* $F\colon [c,d] \to \mathbb{R}$, $F(y) := \int_a^b f(x,y)\,dx$, *is also continuously differentiable and its derivative satisfies*

$$\frac{dF}{dy}(y) = \int_a^b \frac{\partial f}{\partial y}(x,y)\,dx. \tag{4}$$

Proof. For $y_k \to y_0$ with $y_k \neq y_0$ let us justify the following calculation:

$$\begin{aligned}\frac{dF}{dy}(y_0) &= \lim_{k\to\infty} \frac{F(y_k)-F(y_0)}{y_k-y_0} = \lim_{k\to\infty} \int_a^b \underbrace{\frac{f(x,y_k)-f(x,y_0)}{y_k-y_0}}_{=:d_k(x)}\,dx \\ &\overset{!}{=} \int_a^b \lim_{k\to\infty} d_k(x)\,dx = \int_a^b \frac{\partial f}{\partial y}(x,y_0)\,dx.\end{aligned}$$

To do that, we need to assert

$$d_k(x) \to \frac{\partial f}{\partial y}(x,y_0) \qquad \text{uniformly on } [a,b]\ni x.$$

Then the equality marked with "!" is justified by the fact that integration and limit are interchangeable. Moreover, due to the assumed continuity of $\frac{\partial f}{\partial y}(x,y)$ we can then conclude from Lemma 2 that $\frac{dF}{dy}$ is continuous.

`54. Lecture, Tuesday 17. July 07` ——————————————— `"U 12`

We now prove the claim. By assumption, $(x,y)\mapsto \frac{\partial f}{\partial y}(x,y)$ is continuous and therefore uniformly continuous on $[a,b]\times[c,d]$. Thus for given $\varepsilon>0$ there is $\delta>0$, such that for all $y,y'\in[c,d]$ with $|y-y'|<\delta$ we have

$$\left|\frac{\partial f}{\partial y}(x,y) - \frac{\partial f}{\partial y}(x,y')\right| < \varepsilon, \qquad \text{for all } x\in[a,b].$$

We now invoke the mean value theorem of differentiation, Thm IV.8: There exists $\eta_k = \eta_k(x)$ between y_0 and y_k, such that

$$d_k(x) = \frac{\partial f}{\partial y}(x,\eta_k).$$

If k_0 is chosen such that $|y_k-y_0|<\delta$ for all $k\geq k_0$, then also $|\eta_k-y_0|<\delta$ and hence

$$\left|\frac{\partial f}{\partial y}(x,y_0) - d_k(x)\right| = \left|\frac{\partial f}{\partial y}(x,y_0) - \frac{\partial f}{\partial y}(x,\eta_k)\right| < \varepsilon,$$

independently of $x\in[a,b]$. $\square$

1.3. **Fubini's theorem.**

Theorem 4 (Fubini)**.** *(i) If $f\colon [a,b]\times[c,d]\to\mathbb{R}$ is continuous, then the double integral is independent of the order of integration,*

$$\int_c^d \left(\int_a^b f(x,y)\,dx\right)dy = \int_a^b \left(\int_c^d f(x,y)\,dy\right)dx.$$

(ii) The value of the n-fold integral (1) *is independent of the order of integration,*

Proof. (i) Let us recall the fundamental theorem: We have $\frac{d}{dx}\int_c^y f(t)\,dt = f(y)$ for f continuous and $F(d)-F(c)=\int_c^d F'(t)\,dt$ for F' continuous, see Thms. IV.22 and 23.

The map

$$F\colon [c,d]\to\mathbb{R}, \qquad F(y) := \int_a^b \left(\int_c^y f(x,t)\,dt\right)dx.$$

is differentiable by Thm. 3 with

$$\frac{dF}{dy}(y) = \int_a^b \left(\frac{d}{dy}\int_c^y f(x,t)\,dt\right)dx \overset{\text{Fund.thm.}}{=} \int_a^b f(x,y)\,dx.$$

Moreover, using $F(c)=0$,

$$\int_a^b \left(\int_c^d f(x,y)\,dy\right)dx = F(d) \overset{\text{Fund. thm.}}{=} \int_c^d \frac{dF}{dy}(y)\,dy = \int_c^d \left(\int_a^b f(x,y)\,dx\right)dy.$$

(ii) The proof of (i) gives directly that two subsequent integrations in an n-fold integral can be interchanged. But these transpositions generate arbitrary permutations of the order. □

1.4. **The integral for continuous functions with compact support.** Recall that a set $X\subset\mathbb{R}^n$ has the *closure [Abschluss]*

$$\overline{X} := \{x\in\mathbb{R}^n : \exists\,(x_k)\in X \text{ with } x_k\to x\} = \{x\in\mathbb{R}^n : B_\varepsilon(x)\cap X\neq\emptyset\ \forall\varepsilon>0\}$$

Equivalently, $\overline{X}$ is the intersection of all closed sets which contain X.

The *support [Träger]* of a function $f\colon\mathbb{R}^n\to\mathbb{R}$ is the set

$$\operatorname{supp} f := \overline{\{x\in\mathbb{R}^n : f(x)\neq 0\}} \subset \mathbb{R}^n.$$

Examples. 1. $f\colon\mathbb{R}\to\mathbb{R}$ $f(x) := \max\{0, x^2(1-x^2)\}$ has the support $\operatorname{supp} f = [-1,1]$.
2. $\chi_{\mathbb{Q}}\colon\mathbb{R}\to\mathbb{R}$ with $\chi_{\mathbb{Q}}(x) := 1$ for $x\in\mathbb{Q}$ and else $=0$ has $\operatorname{supp}\chi_Q = \mathbb{R}$.

Problem. Let $\varepsilon>0$ and $\alpha\in\mathbb{N}$. Construct a function $f\in C^\alpha(\mathbb{R}^n,\mathbb{R})$ with $\operatorname{supp} f\subset B_1(0)$ and $f(x)=1$ on $B_{1-\varepsilon}(0)$.

Definition. (i) A function $f \in C^0(\mathbb{R}^n)$ with $\operatorname{supp} f$ compact is called a *continuous function with compact support*; we write

$$C_c^0(\mathbb{R}^n) := \{f \colon U \to \mathbb{R} : \operatorname{supp} f \subset U \text{ compact } \}.$$

(ii) For $f \in C_c^0(\mathbb{R}^n)$, we can declare an integral by

$$\int_{\mathbb{R}^n} f(x)\,dx := \int_{[a,b]} f(x)\,dx,$$

where $[a,b]$ is a suffienctly large box containing $\operatorname{supp} f$.

It is rather restrictive to consider only continuous functions with compact support for integration. Nevertheless, in the 4th term course on integration we will see that the very general Lebesgue-integrable functions can be approximated well by C_c^0-functions. Indeed, the C_c^0-functions will turn out to be dense in the space of integrable functions. Thus properties of the integral over C_c^0 will extend immediately to properties of the integral over more general functions.

`55. Lecture, Thursday 19. July 07` ______________________________

The following properties are well-known for the one-dimensional Riemann integral. They are preserved by iterated integration, and hence we get:

Theorem 5. *The integral over $C_c^0(\mathbb{R}^n)$ has the following poperties:*
(i) $f \mapsto \int_{\mathbb{R}^n} f(x)\,dx$ is linear.
(ii) Monotonicity: $f \le g \Rightarrow \int_{\mathbb{R}^n} f(x)\,dx \le \int_{\mathbb{R}^n} g(x)$.
(iii) Translation invariance: $\int_{\mathbb{R}^n} f(x-a)\,dx = \int_{\mathbb{R}^n} f(x)\,dx$ for all $a \in \mathbb{R}^n$.

It can be shown that these properties determine the integral up to a multiplicative constant (Proof: [F3], p.4-11).

It is often useful to define a continuous function f with compact support in an open set $U \subset \mathbb{R}^n$ as a a function $f \in C^0(U)$ with $\operatorname{supp} f \subset U$. To understand this notion and to define an integral we need the following fact. Consider U open and $K \subset U$ compact. Then there is an ε-neighbourhood

$$K_\varepsilon := \{y \in \mathbb{R}^n : \|x - y\| < \varepsilon \text{ for some } x \in K\}$$

with $\varepsilon > 0$ which is entirely contained in U (exercise).

We can apply this fact to $f \in C_c^0(U)$ and $K := \operatorname{supp} f$ to see that f has boundary values 0 on U in the sense that $\lim_{x\to y} f(x) = 0$ whenever $y \in \partial U$. Consequently, if we extend $f \in C_c^0(U)$ with 0 beyond U, that is, we set $\tilde f(x) := f(x)$ for $x \in U$ and $\tilde f(x) := 0$ else, then $\tilde{f} \in C_c^0(\mathbb{R}^n)$. This fact allows us to define $\int_U f(x)\,dx := \int_{\mathbb{R}^n} \tilde{f}(x)\,dx$. Note that if U is not open then the resulting extension would in general not be continuous.

1.5. **The integral under a linear change of variables.** We expect that the integral over $C_c^0(\mathbb{R}^n)$ is invariant also under rotation. However, this is not obvious to see. In the following, we will investigate the behaviour of the integral more generally under linear transformation. This will explain the behaviour for general (non-linear) changes of variables.

Let A be an $n \times n$-matrix of a linear map $x \mapsto Ax$. Linear Algebra explains the determinant $\det A$ usually as the change of oriented volume under the linear map. Taking this explanation for granted, the formula for the change of variables is no surprise:

Theorem 6. *If $f \in C_c(\mathbb{R}^n)$ and $A \in \mathsf{GL}_n(\mathbb{R})$ then*

$$\int_{\mathbb{R}^n} f(Ax)\,|\det A|\,dx = \int_{\mathbb{R}^n} f(y)\,dy. \tag{5}$$

Note that $|\det A|$ is a number which we can factor out of the integral.

Examples. 1. For $A \in \mathsf{O}(n)$ orthogonal, the integral stays invariant.
2. One-dimensional substitution gives

$$\int_a^b f\big(\varphi(x)\big)\varphi'(x)\,dx = \int_{\varphi(a)}^{\varphi(b)} f(y)\,dy = \pm \int_{\min(\varphi(a),\varphi(b))}^{\max(\varphi(a),\varphi(b))} f(y)\,dy,$$

where the sign is "+" for φ monotonically increasing, and "−" for φ monotonically decrasing. Placing this sign in the left integral, we obtain the one-dimensional case of (5)

$$\int_{\mathbb{R}} f\big(\varphi(x)\big)\big|\varphi'(x)\big|\,dx = \int_{\mathbb{R}} f(y)\,dy \qquad \text{for } f \in C_c^0([a,b]).$$

In particular, for $\varphi(x) = \lambda x$ with $\lambda \neq 0$ this gives

$$\int_{\mathbb{R}} f(\lambda x)|\lambda|\,dx = \int_{\mathbb{R}} f(y)\,dy. \tag{6}$$

Proof. Recall Gaussian elimination from linear algebra: Each $A \in \mathsf{GL}(n)$ can be multiplied with a finite number of elementary matrices, such that the product becomes the unit matrix Id_n. The product of the elementary matrices then represents the inverse of A, as

$$T_k \cdots T_1 A = \mathrm{Id}_n \qquad \Longleftrightarrow \qquad A^{-1} = T_1 \cdots T_k.$$

Let us apply this to A^{-1} in place of A. The result is a representation $A = T_1 \cdots T_k$. The elementary matrices T_i can be chosen from just two types,

$$M(\lambda, i) := \begin{pmatrix} 1 & & & & & \\ & \ddots & & & 0 & \\ & & 1 & & & \\ & & & \lambda & & \\ & & & & 1 & \\ & 0 & & & & \ddots & \\ & & & & & & 1 \end{pmatrix}, \qquad P(\lambda, i, j) := \begin{pmatrix} 1 & & & & & \\ & \ddots & & & 0 & \\ & & 1 & & & \\ & & \vdots & \ddots & & \\ & & \lambda & \cdots & 1 & \\ & 0 & & & & \ddots & \\ & & & & & & 1 \end{pmatrix},$$

where $\lambda \neq 0$ is located in the (i,i), or (i,j) position, respectively.

It is sufficient to prove the following three facts:
1. The change of variables formula holds for linear map $M = M(\lambda, i)$.
2. It holds for $P = P(\lambda, i, j)$.
3. If it holds for two matrices A and B then it holds for their product AB.

Induction then proves the formula for $A = T_1 \cdots T_k$.

1. Invoking Fubini's theorem, let us first integrate over i-th component:

$$\int_{\mathbb{R}} f\big(M(\lambda,i)x\big)\big|\det(M(\lambda,i))\big|\,dx_i$$
$$= \int_{\mathbb{R}} f(x_1, \dots, \lambda x_i, \dots, x_n)|\lambda|\,dx_i \overset{(6)}{=} \int_{\mathbb{R}} f(x_1, \dots, x_i, \dots x_n)\,dx_i.$$

Subsequent integration over the remaining components gives the result.

2. Since $\det P(\lambda, i, j) = 1$ we need to show

$$\int_{\mathbb{R}^n} f(Px)\,dx = \int_{\mathbb{R}^n} f(x_1, \dots, x_{i-1}, x_i + \lambda x_j, x_{i+1}, \dots, x_n)\,dx \overset{!}{=} \int_{\mathbb{R}^n} f(x_1, \dots, x_n)\,dx.$$

By Fubini's theorem we can integrate over the variables not equal to i, j first. We consider the result as a function $\tilde{f} \colon \mathbb{R}^2 \to \mathbb{R}$. It remains to show that this function satisfies

$$\int_{\mathbb{R}} \int_{\mathbb{R}} \tilde{f}(x + \lambda y, y)\,dx\,dy = \int_{\mathbb{R}} \int_{\mathbb{R}} \tilde{f}(x, y)\,dx\,dy;$$

geometrically this is shear invariance of the integral. But for each y,

$$\int_{\mathbb{R}} \tilde{f}(x + \lambda y, y)\,dx = \int_{\mathbb{R}} \tilde{f}(x, y)\,dx$$

by the translation invariance of the one-dimensional integral.

3. Using the product law for determinants we obtain, as desired:

$$\int_{\mathbb{R}^n} f(ABx)\big|\det(AB)\big|\,dx = \int_{\mathbb{R}^n} f\big(A(Bx)\big)|\det A||\det B|\,dx$$
$$\overset{\text{step 1 or 2}}{=} \int_{\mathbb{R}^n} f\big(A(y)\big)|\det A|\,dy \overset{\text{step 1 or 2}}{=} \int_{\mathbb{R}^n} f(x)\,dx.$$

$\square$

We can now verify that the determinant measures volume distortion. Let us use the notation $\chi_X(x)$ for the *characteristic function* which is 1 for $x \in X$ and 0 for $x \notin X$. Now consider the cube $W := [0,1]^n$ under a linear map $A \in \mathsf{GL}_n(\mathbb{R})$. Note that $\chi_{A(W)}$ is 1 provided $x \in A(W) \iff A^{-1}x \in W$, that is, $\chi_{A(W)}(x) = \chi_W(A^{-1}x)$. Hence we find

$$\begin{aligned} \operatorname{vol}\big(A(W)\big) &:= \int_{\mathbb{R}^n} \chi_{A(W)}(x)\,dx = |\det A| \int_{\mathbb{R}^n} \chi_W(A^{-1}x)|\det A^{-1}|\,dx \\ &\overset{(5)}{=} |\det A| \int_{\mathbb{R}^n} \chi_W(y)\,dy = |\det A| \operatorname{vol}(W) = |\det A|. \end{aligned} \tag{7}$$

Clearly, χ_W is not a function in $C_c^0(\mathbb{R}^n)$ and so our computation is beyond the scope of the integral we have introduced. But working with continuous functions and inequalities we would get the same result. The calculation will work directly with the Lebesgue integral.

1.6. **Change of variables for continuous functions with compact support.** Let us state the nonlinear version of the change of variables formula.

Theorem 7 (Change of variables *[Transformationsformel]*)**.** *Let $U, V \subset \mathbb{R}^n$ be open and $\varphi\colon U \to V$ be a C^1-diffeomorphism. Then for each $f \in C_c^0(V)$*

$$\int_U f\big(\varphi(x)\big)\,|\det d\varphi_x|\,dx = \int_V f(y)\,dy.$$

The proof of change of variables formula is lengthy. We will not present it here, but refer to Forster for a proof [F 3, S.16–21].

The strategy of the proof is the following. In the special case of an affine linear map $\varphi(x) = Ax + b$ the proof follows from our statements on translation invariance and the linear case of the change of variables formula.

1. Subdivide the domain U into sufficiently small pieces, that is, use a covering of $U \subset \bigcup_{i\in I} U_i$ with finitely many cubes of edgelength δ.
2. In each cube, use an affine linear approximation to φ, that is, write $\varphi(x+h) := \varphi(x) + d\varphi_x(h) + r_x(h)$ where $\|h\| < \delta$.
3. Note that the change of variables formula holds for the affine approximation $\varphi(x) + d\varphi_x(h)$. Thus the total error ε is given by the sum of the integrals over the remainder term $r_x(h)$ for each cube. Since the error is quadratic in h, the sum of the errors will converge to zero when $\delta \to 0$.

Example. As an application we will calculate the indefinite integral $\int_{\mathbb{R}} e^{-x^2}\,dx$, which is not elementary. To do that, we will integrate the seemingly more complicated function $e^{-x^2-y^2}$

over $\mathbb{R}^2$. We introduce polar coordinates,

$$P\colon U := (0,\infty)\times(0,2\pi) \to V := \mathbb{R}^2 \setminus [0,\infty)\times\{0\}, \qquad (r,\varphi) \mapsto \begin{pmatrix} r\cos\varphi \\ r\sin\varphi \end{pmatrix}.$$

Then $\det dP = r$ and so

$$\begin{aligned}\int_{\mathbb{R}^2} e^{-x^2-y^2}\,dxdy &= \int_V e^{-x^2-y^2}\,dxdy = \int_U e^{-r^2\cos^2\varphi + r^2\sin^2\varphi} r\,drd\varphi \\ &= \int_U e^{-r^2} r\,drd\varphi = \int_0^{2\pi}\int_0^\infty e^{-r^2} r\,drd\varphi = 2\pi\Big[\frac{-1}{2}e^{-r^2}\Big]_0^\infty = \pi\end{aligned}$$

On the other hand,

$$\int_{\mathbb{R}^2} e^{-x^2-y^2}\,dxdy = \int_{\mathbb{R}}\int_{\mathbb{R}} e^{-x^2}e^{-y^2}\,dxdy = \int_{\mathbb{R}} e^{-x^2}\,dx \int_{\mathbb{R}} e^{-y^2}\,dy = \Big(\int_{\mathbb{R}} e^{-x^2}\,dx\Big)^2.$$

Consequently,

$$\int_{\mathbb{R}} e^{-x^2}\,dx = \sqrt{\pi}.$$

Clearly, this integration is once again beyond the class $C_c^0(\mathbb{R}^2)$: our functions do not have compact support in U or V. Nevertheless all steps can be justified, either by approximation or using the Lebesgue integral.

`End of Lecture` ______________________________

Index