Jacobs University Bremen
School of Engineering and Science
Götz E. Pfander

Analysis I and II

Version July 13, 2015

This script contains all the theorems and definitions, but only few examples, covered in Analysis I in the academic year 2013. The material of Analysis I is contained in Sections 1–4, the material of Analysis II in Sections 5–8.

Contents

1 Numbers **4**

1.1 Sets, relations and functions 4

1.2 Groups, fields, the integers and the rational numbers 7

1.3 Real numbers 10

1.4 Complex numbers 17

2 Convergence of sequences in metric spaces and numeric series **19**

2.1 Sequences in metric spaces 19

2.2 The extended real number system, lim sup and lim inf 24

2.3 Cauchy sequences and complete metric spaces 27

2.4 Real and complex series 28

3 Topology and continuity **37**

3.1 Continuous functions 37

3.2 Topological spaces 43

3.3 Compact sets 47

3.4 Connected sets 53

3.5 Sequences of functions, uniform convergence 56

4 Differentiation **58**

4.1 Central results 58

4.2 Taylor series 63

4.3 The exponential function and friends 65

4.4 Fixed point theorems and approximative methods 66

5 Integration **68**

5.1 The Riemann integral 68

5.2 Integration and differentiation 72

5.3 Improper Riemann integral 73

5.4 Infinite dimensional vector spaces and orthonormal bases 74

5.5 Fourier series 78

6 Multivariable Calculus **81**
6.1 Some facts from linear algebra 81
6.2 Curves 82
6.3 Derivatives of multivariable functions 83
6.4 Higher derivatives 90
6.5 Taylor's theorem and applications 93
6.6 Implicit functions and the inverse function theorem 97

7 Integration on $\mathbb{R}^d$ **104**
7.1 Essentials 104
7.2 Jordan content 108
7.3 Change of variables 109
7.4 Multivariate improper integrals 112
7.5 The Gamma function 113

8 Ordinary differential equations **114**

1. NUMBERS

1.1. Sets, relations and functions

Definition 1.1. The *cartesian product* $X_1 \times X_2 \times \ldots \times X_n$ of the n sets $X_1, X_2, \ldots, X_n$ is the set of all (ordered) *n-tupels* $(x_1, x_2, \ldots, x_n)$ with $x_1 \in X_1, x_2 \in X_2, \ldots, x_n \in X_n$. That is,

$$X_1 \times X_2 \times \ldots \times X_n := \{(x_1, x_2, \ldots, x_n) : x_1 \in X_1,\ x_2 \in X_2, \ldots,\ x_n \in X_n\}.$$

Note that $A \times \emptyset = \emptyset \times A = \emptyset$, and $A \times B = B \times A$ if and only if $A = B$ or $A = \emptyset$ or $B = \emptyset$.

Examples 1.2. i. $\{1,2,3\} \times \{7,12\} = \{(1,7),(2,7),(3,7),(1,12),(2,12),(3,12)\}$

ii. $\{7,12\} \times \{1,2,3\} = \{(7,1),(7,2),(7,3),(12,1),(12,2),(12,3)\}$

iii. $\{7,12\} \times \{\} = \{\}$

iv. $\{7,12\} \times \{1,2\} \times \{a,b\}$
$= \{(7,1,a),(7,2,a),(12,1,a),(12,2,a),(7,1,b),(7,2,b),(12,1,b),(12,2,b)\}$

Definition 1.3. Any subset R of the cartesian product $X \times Y$ of two sets X and Y, that is, $R \subset X \times Y$, is called *relation* between X and Y. If $X = Y$ we say that $R \subset X \times X$ is a relation on X.

$\mathcal{D}(R) = \mathcal{D}_R = \{x \in X :$ there exists $y \in Y$ with $(x,y) \in R\}$ is called *domain of* R, and
$\mathcal{R}(R) = \mathcal{R}_R = \{y \in Y :$ there exists $x \in X$ with $(x,y) \in R\}$ is called *range of* R.

Definition 1.4. Let X and Y be sets. A *function* (or *mapping*) $f : X \longrightarrow Y$ is a rule that associates to **every** element in $x \in X$ an element $f(x) \in Y$. X is called *domain* of f and is denoted by $\mathcal{D}_f$.

For $A \subseteq X$ and $B \subseteq Y$ we set

$$f(A) = \{y \in Y : \quad \text{there exists } x \in A \text{ with } f(x) = y\}$$

and

$$f^{-1}(B) = \{x \in X : \quad \text{there exists } y \in B \text{ with } f(x) = y\}.$$

The *range* of f is given by $\mathcal{R}_f = f(X)$. The *graph* of f is the relation $\Gamma_f = \{(x,y) \in X \times Y : f(x) = y\}$ between X and Y.

The function f is *injective* (*one–to–one*) if $f(x) = f(\widetilde{x})$ implies $x = \widetilde{x}$, and f is *surjective* (*onto*) if $\mathcal{R}_f = Y$. If f is surjective and injective, we call f *bijective* We refer to an injective map also as *embedding*.

Remark 1.5. Note that the distinction between a function and its graph is done for psychological reasons only. A strictly axiomatic introduction of analysis is based on set theory and functions are simply defined as certain subsets of $X \times Y$.

Proposition 1.6. *A relation $\Gamma \subset X \times Y$ is the graph of a function $f : \mathcal{D}_\Gamma \longrightarrow Y$, if and only if $(x,y),(x,\widetilde{y}) \in \Gamma$ implies $y = \widetilde{y}$ for all $x \in X$ and $y, \widetilde{y} \in Y$. In this case we have $\mathcal{R}_f = \mathcal{R}_{\Gamma_f}$ and $\mathcal{D}_f = \mathcal{D}_{\Gamma_f}$.*

Theorem 1.7. *Given a function $f : X \longrightarrow Y$ and sets $A_i \subset X$, $i \in \mathbb{N}$, and $B_i \subset Y$, $i \in \mathbb{N}$, we have*

i. $A_1 \subseteq A_2$ *implies* $f(A_1) \subseteq f(A_2)$

ii. $B_1 \subseteq B_2$ *implies* $f^{-1}(B_1) \subseteq f^{-1}(B_2)$

iii. $A_1 \subseteq f^{-1}(f(A_1))$ *and* $B_1 \supseteq f(f^{-1}(B_1))$

iv. $f(\bigcup_{i=1}^{\infty} A_i) = \bigcup_{i=1}^{\infty} f(A_i)$ *and* $f(\bigcap_{i=1}^{\infty} A_i) \subseteq \bigcap_{i=1}^{\infty} f(A_i)$

If f is injective we have in addition $A_1 = f^{-1}(f(A_1))$ and $f(\bigcap_{i=1}^{\infty} A_i) = \bigcap_{i=1}^{\infty} f(A_i)$ and if f is surjective $B_1 = f(f^{-1}(B_1))$.

Proof. We shall only proof that $A_1 \subseteq f^{-1}(f(A_1))$ and give an example where equality fails. For $x \in A_1$ we have $y = f(x) \in f(A_1)$ and $x \in f^{-1}(\{y\}) \subseteq f^{-1}(f(A_1))$. Now, consider $X = \{0, 1\}$ and $Y = \{0\}$. The only function $f : X \longrightarrow Y$ that exists, namely, the function $f(0) = f(1) = 0$ does the job since for $A_1 = \{0\}$ we have $f^{-1}(f(A_1)) = \{0, 1\} \neq \{0\}$. □

Remark 1.8. Concerning the proof above. In mathematics, we aim for generality when stating and proving theorems. We aim for simplicity when providing counterexamples. The simpler, the better.

Definition 1.9. A relation R on X is called

i. *reflexive* if for all $x \in X$ we have $(x, x) \in R$,

ii. *transitive* if $(x, \widetilde{x}) \in R$ and $(\widetilde{x}, \widetilde{\widetilde{x}}) \in R$ implies $(x, \widetilde{\widetilde{x}}) \in R$,

iii. *symmetric* if $(x, \widetilde{x}) \in R$ implies $(\widetilde{x}, x) \in R$, and

iv. *antisymmetric* if $(x, \widetilde{x}) \in R$ and $(\widetilde{x}, x) \in R$ implies $x = \widetilde{x}$.

Definition 1.10. A reflexive, symmetric, and transitive relation R on X is called *equivalence relation*. If R is an equivalence relation we shall write $x \sim \widetilde{x}$ if $(x, \widetilde{x}) \in \mathbb{R}$ and call x and $\widetilde{x}$ *equivalent* with respect to R.

$[x] = \{\widetilde{x} \in X : \quad (x, \widetilde{x}) \in R\}$ is called *equivalence class* of x, and any $\widetilde{x} \in [x]$ is called *representative* of $[x]$.

Example 1.11. Fix $n \in \mathbb{N}$ and set $X = \mathbb{Z}$. The relation

$$R_{\mathbb{Z}_n} = \{(k, m) \in \mathbb{Z} \times \mathbb{Z} : \quad k - m = l \cdot n \text{ for some } l \in \mathbb{Z}\}$$

is an equivalence relation [check reflexivity, transitivity, and symmetry]. The set of equivalence classes is the group $\mathbb{Z}_n$ of n elements with addition given by

$$[k] + [m] = [k + m].$$

To see this, you would have to check whether addition is well defined and you need to check all group properties (which are discussed in detail below).

To see that addition on $\mathbb{Z}_n$ is well defined, we need to show that the sum of two equivalence classes $A = [k]$ and $B = [m]$ does not depend on the chosen representatives k and m. Namely, given $k, \widetilde{k}, m, \widetilde{m} \in \mathbb{Z}$ with $k \sim \widetilde{k}$ and $m \sim \widetilde{m}$, that is, $[k] = A = [\widetilde{k}]$ and $[m] = B = [\widetilde{m}]$. We have to show that $k + m \sim \widetilde{k} + \widetilde{m}$, that is, $[k + m] = A + B = [\widetilde{k} + \widetilde{m}]$, to see that addition is independent of the representatives we choose from A and B.

But this follows from the fact that $k \sim \widetilde{k}$ and $m \sim \widetilde{m}$ imply that $k - \widetilde{k}$ and $m - \widetilde{m}$ are multiples of n, hence their sum $k - \widetilde{k} + m - \widetilde{m} = k + m - (\widetilde{k} + \widetilde{m})$ is a multiple of n which shows $k + m \sim \widetilde{k} + \widetilde{m}$.

The concept of a partition of a set helps to understand equivalence classes and their equivalence relations.

Definition 1.12. A family of sets $\{M_i : \ i \in I\}$ is a partition of the set $M \neq \emptyset$, if

i. $\emptyset \neq M_i \subset M$ for $i \in I$,

ii. $i \neq j$ implies $M_i \cap M_j = \emptyset$ for $i, j \in I$, and

iii. $\bigcup_{i \in I} M_i = M$.

Theorem 1.13. *For a set $M \neq \emptyset$ we have:*

i. The distinct equivalence classes of an equivalence relation on M form a partition on M.

ii. A partition $\{M_i : \ i \in I\}$ on M induces an equivalence relation on M via

$$a \sim b \quad \text{if and only if} \quad a, b \in M_{i_0} \text{ for some } i_0 \in I.$$

Proof. There is not much to be proven here. For example, given distinct equivalence classes $[x_i]$, $i \in I$ with I being an arbitrary index set. If $x \in [x_i] \cap [x_j]$, then $x \sim x_i$ and $x \sim x_j$. By symmetry, we have also $x_i \sim x$, so transitivity implies $x_i \sim x_j$. Hence $[x_i] = [x_j]$, which contradicts that we considered distinct equivalence classes. □

Definition 1.14. A relation O on X is called *order* on X if O is reflexive, transitive, and antisymmetric. The order O is called *linear* if for all $x, \widetilde{x} \in X$ either $(x, \widetilde{x}) \in O$ or $(\widetilde{x}, x) \in O$.

Example 1.15. The relation $O_{\mathbb{N}} = \{(n, m) \in \mathbb{N} \times \mathbb{N} : \quad n \leq m\}$ is a linear order on $\mathbb{N}$.

Note that the natural order on $\mathbb{N}$ can be easily defined using elementary set theory.

1.2. Groups, fields, the integers and the rational numbers

Definition 1.16. A *group* is a set G, together with a binary law of composition $\mu : G \times G \longrightarrow G$ which satisfies the axioms G1, G2, and G3 given below. We shall write $xy := \mu(x, y)$.

(G1) *Associativity*: $(xy)z = x(yz)$ for all $x, y, z \in G$.

(G2) *Identity*: There exists an element $e \in G$ called *identity* such that $xe = ex = x$ for all $x \in G$.

(G3) *Inverses*: To each element $x \in G$ exists an element $y \in G$ called *inverse* of x with $xy = yx = e$. The inverse to x is denoted by x^{-1}, in the case of additively (using for the group composition law the symbol "+") written abelian groups as $-x$.

A group is called *abelian* if μ is commutative, that is, if we have

(C) $xy = yx$ for all $x, y \in G$.

Examples 1.17. i. Addition $+ : \mathbb{N} \times \mathbb{N} \longrightarrow \mathbb{N},\ (n, m) \mapsto n + m$ on the natural numbers $\mathbb{N} = \{1, 2, 3, \ldots\}$ is associative. But $\mathbb{N}$ contains no neutral element. Hence, $(\mathbb{N}, +)$ is not a group.

ii. Addition $+ : \mathbb{N}_0 \times \mathbb{N}_0 \longrightarrow \mathbb{N}_0),\ (n, m) \mapsto n + m$ on the natural numbers with zero, $\mathbb{N}_0 = \{0, 1, 2, 3, \ldots\}$ is associative and 0 is a neutral element. But no element in $\mathbb{N}_0$ other than 0 has an inverse element! Hence, $(\mathbb{N}_0, +)$ is not a group as well.

iii. Let $X = \mathbb{N} \times \mathbb{N}$ and define

$$R_{\mathbb{Z}} = \{((n, m), (\widetilde{n}, \widetilde{m})) \in (\mathbb{N} \times \mathbb{N}) \times (\mathbb{N} \times \mathbb{N}) : \quad n + \widetilde{m} = \widetilde{n} + m\}.$$

$R_{\mathbb{Z}}$ is an equivalence relation. The set of equivalence classes $\mathbb{Z} := \{[(n, m)]\}$ equipped with [1]

$$[(n, m)] +_{\mathbb{Z}} [(\widetilde{n}, \widetilde{m})] = [(n + \widetilde{n}, m + \widetilde{m})]$$

is an Abelian group with neutral element $[(1, 1)]$ and inverses $-[(n, m)] = [(m, n)]$.

Note that we can also define a multiplication on $\mathbb{Z}$, namely[2]

$$[(n, m)] \cdot_{\mathbb{Z}} [(\widetilde{n}, \widetilde{m})] = [(n \cdot \widetilde{n} + m \cdot \widetilde{m}, n \cdot \widetilde{m} + m \cdot \widetilde{n})].$$

In fact, $\mathbb{Z}$ equipped with the addition above and the product below forms a so-called commutative ring, called *ring of integers* . Since we shall not use any rings in this course, we omit a definition of *rings.*

[1] Why is the following definition well defined, that is, independent of the representatives of the equivalence classes?

[2] To see that this makes sense, recall that $[(n, m)]$ is just a clumsy way of writing the integer $n - m$ without using "−". We know that better $(n - m)(\widetilde{n} - \widetilde{m}) = n\widetilde{n} - n\widetilde{m} - \widetilde{m}n + n\widetilde{n}$, which, avoiding "−" is simply $[(n \cdot \widetilde{n} + m \cdot \widetilde{m}, n \cdot \widetilde{m} + m \cdot \widetilde{n})]$. So this is how we came up with a definition of multiplication on $\mathbb{Z}$, and in maths a good guess / good intuition is worth half the money. It still remains to show that this is meaningful, for example, that this map satisfies associativity.

We can embed (map injectively) the naturals into this ring of equivalence classes via

$$i : \mathbb{N} \longrightarrow \mathbb{Z}, \quad n \mapsto n^* := [(n+1,1)].$$

This mapping is nice, since it respects addition and multiplication on $\mathbb{N}$, that is,

$$i(n+\widetilde{n}) = i(n) +_{\mathbb{Z}} i(\widetilde{n}), \text{ and } i(n \cdot \widetilde{n}) = i(n) \cdot_Z i(\widetilde{n})$$

Hence, using an appropriate equivalence relation on $\mathbb{N} \times \mathbb{N}$, we have created a ring of equivalence classes which can be identified with the set of integers.[3] In the following, we will not make a distinction between a natural number n and its integer counterpart n^*. We shall use the common short hand notation $z = n - m = [(n,m)] \in \mathbb{Z}$. Note that $[(7,3)] = [(10,6)]$, since $7+6 = 3+10$, that is, $7-3 = 10-6$.

Note that the relation $O_{\mathbb{Z}} = \left\{ \Big([(n,m)], [(\widetilde{n},\widetilde{m})] \Big) \in \mathbb{Z} \times \mathbb{Z} : \quad n + \widetilde{m} \leq \widetilde{n} + m \right\}$ extends the order on $\mathbb{N}$ to the integers $\mathbb{Z}$.

iv. Let $X = \mathbb{Z} \times \mathbb{N}$ and define

$$R_{\mathbb{Q}} = \left\{ \big((z,m), (\widetilde{z},\widetilde{m}) \big) \in (\mathbb{Z} \times \mathbb{N}) \times (\mathbb{Z} \times \mathbb{N}) : \quad z \cdot \widetilde{m} = \widetilde{z} \cdot m \right\}.$$

$R_{\mathbb{Q}}$ is an equivalence relation. The set of equivalence classes $\{[(z,m)]\}$ equipped with

- $[(z,m)] +_{\mathbb{Q}} [(\widetilde{z},\widetilde{m})] = [(z \cdot_{\mathbb{Z}} \widetilde{m} + \widetilde{z} \cdot_{\mathbb{Z}} m, m \cdot_{\mathbb{Z}} \widetilde{m})]$
- $[(z,m)] \cdot_{\mathbb{Q}} [(\widetilde{z},\widetilde{m})] = [(z \cdot_{\mathbb{Z}} \widetilde{z}, m \cdot_{\mathbb{Z}} \widetilde{m})]$

is a field[4], called the field of *rational numbers*. Again, we can embed the integers in a natural way by setting

$$i : \mathbb{Z} \longrightarrow \mathbb{Q}, \quad z \mapsto z^* := [(z,1)].$$

This embedding respects multiplication and addition, hence, we consider $\mathbb{Z}$ as a subring of the ring (field) of equivalence classes we just defined. The field we defined is the field of rational numbers. From now on, we shall use them the way we are used to. Certainly, we shall write $r = \dfrac{z}{m} = [(z,m)] \in \mathbb{Q}$.

The relation $O_{\mathbb{Q}} = \left\{ \Big([(z,m)], [(\widetilde{z},\widetilde{m})] \Big) \in \mathbb{Q} \times \mathbb{Q} : \quad z \cdot \widetilde{m} \leq \widetilde{z} \cdot m \right\}$ extends the order on $\mathbb{Z}$ to the rational numbers $\mathbb{Q}$. In the following we shall simply write $r \leq \widetilde{r}$ if $(r,\widetilde{r}) \in O_{\mathbb{Q}}$.

Starting from the natural numbers we have created the integers, from those we have created the rationals. Since the embeddings are canonical, we shall ignore its formalism and simply take

$$\mathbb{N} \subsetneq \mathbb{Z} \subsetneq \mathbb{Q}.$$

Definition 1.18. A *field* is a set F on which two binary laws of composition, *addition* '+' and *multiplication* '·' are defined with

[3] We only assume *a-priori* knowledge of the naturals. Similar to the attitude of Leopold Kronecker, 1823-1891, who supposedly said "God made the integers; all else is the work of man".

[4] Fields will be defined shortly.

(F1) $(F,+)$ is an abelian group. We shall denote the identiy of $(F,+)$ as 0.

(F2) $(F\backslash\{0\},\cdot)$ is an abelian group. The identity of $(F\backslash\{0\},\cdot)$ is denoted by 1.

(F3) The *distributive law* holds, that is, $(x+y)\cdot z = xz + yz$ for all $x, y, z \in F$.

All orders discussed in Examples and 1.15 and **??** are those orders on $\mathbb{N}$, $\mathbb{Z}$, and $\mathbb{Q}$ which you are familiar with. In our attempt of presenting a self–contained constructive approach to introduce the real numbers, we include the formal definitions below.

These definitions are not very enlightening and they will not play a crucial part throughout the remainder of Analysis 1.

Definition 1.19. A field F is called *ordered* if

(O1) There exists an order '$\leq$' on F.

(O2) The order is linear, that is, for all $x, y \in F$ either $x < y$ or $x > y$ or $x = y$.

(O3) $x \leq y$ implies $x + z \leq y + z$ for all $x, y, z \in F$ and if $x, y > 0$ then $x \cdot y > 0$.

Definition 1.20. An ordered field F is called *archimedean* if for all $x, y \in F$, $x, y > 0$, exists $n \in \mathbb{N}$ with

$$nx := \underbrace{x + x + \ldots + x}_{n-times} > y.$$

Theorem 1.21. *The set of rational numbers $\mathbb{Q}$ together with the two binary operations addition and multiplication and the order defined in Examples 1.17.iv is an archimedean ordered field.*

Proof. We have discussed the binary operations addition and multiplication by themselves earlier. To see that both "go hand in hand", that is, that the distributive law holds, observe that with $k, p, u \in \mathbb{Z}$ and $l, q, v \in \mathbb{N}$ we have

$$\frac{k}{l}\left(\frac{p}{q}+\frac{u}{v}\right) = \frac{k}{l}\left(\frac{pv+uq}{qv}\right) = \left(\frac{kpv+kuq}{lqv}\right) = \left(\frac{kplv+kulq}{lqlv}\right) = \frac{kp}{lq}+\frac{ku}{lv} = \frac{k}{l}\frac{p}{q}+\frac{k}{l}\frac{u}{v}.$$

It is easily checked that the previously introduced order on $\mathbb{Q}$ makes $\mathbb{Q}$ an ordered field.

It remains to show that the order on the field is archimedean. To this end, let $x = \frac{k}{l} > 0$ and $y = \frac{p}{q} > 0$. Set $n = l(p+1)$ and observe that

$$nx = l(p+1)\frac{k}{l} = (p+1)k > pk \geq \frac{p}{q} = y. \quad \square$$

1.3. Real numbers

Given a right angled, isosceles triangle with two sides of length 1. What is the length l of the third side?

According to Phythagoras, we have $l^2 = 1^2 + 1^2 = 1 + 1 = 2$. We shall write $l = \sqrt{2}$.

Theorem 1.22. $\sqrt{2} \notin \mathbb{Q}$, *that is, there exists no* $m \in \mathbb{Z}$ *and* $n \in \mathbb{N}$ *with* $\left(\frac{m}{n}\right)^2 = 2$.

Proof. Assume $\left(\frac{m}{n}\right)^2 = 2$ with $\frac{m}{n}$ in lowest terms. Then $m^2 = 2n^2$. Hence, m^2 and therefore m is even, say $m = 2k$. But then $4k^2 = 2n^2$ and n even. This contradicts that $\frac{m}{n}$ is in lowest terms. □

We conclude that there exist line segments with non rational length. Can we define a set $S \supseteq \mathbb{Q}$ containing all "lengths", and to which we can extend all arithmetic properties of $\mathbb{Q}$? Yes, we can!

Definition 1.23. A *Dedekind–cut* $A|B$ in $\mathbb{Q}$ is a pair of subsets A, B of $\mathbb{Q}$ with

i. $A \cup B = \mathbb{Q}$, $A \neq \emptyset$ and $B \neq \emptyset$, $A \cap B = \emptyset$,

ii. for all $a \in A$ and $b \in B$ we have $a < b$, that is, $a \leq b$ and $a \neq b$, and

iii. A contains no largest element.

For $p \in \mathbb{Q}$ we call
$$p^\star := \{q \in \mathbb{Q} :\ q < p\}|\{q \in \mathbb{Q} :\ q \geq p\}.$$
a rational cut in $\mathbb{Q}$.

Dedekind-cuts in $\mathbb{Q}$ are called *real numbers*, the set of all real numbers is denoted by $\mathbb{R}$.

Examples 1.24. The pairs of subsets

$$\{q \in \mathbb{Q} : q < 2\}|\{q \in \mathbb{Q} : q \geq 2\}$$
$$\{q \in \mathbb{Q} : q < 0 \text{ or } q^2 < 2\}|\{q \in \mathbb{Q} : q \geq 0 \text{ and } q^2 > 2\}$$

are Dedekind cuts, but

$$\{q \in \mathbb{Q} : q \leq 2\}|\{q \in \mathbb{Q} : q > 2\}$$
$$\{q \in \mathbb{Q} : q^2 \leq 2\}|\{q \in \mathbb{Q} : q^2 > 2\}$$
$$\{q \in \mathbb{Q} : q < 2\}|\{q \in \mathbb{Q} : q \geq 3\}$$

are not.

To see that indeed, the set of real numbers does not share the shortcoming of $\mathbb{Q}$ of having holes, we define define the following.

Definition 1.25. Let X be a linearly ordered set, $S \subseteq X$ be not empty. $M \in X$ is an *upper* [resp. *lower*] *bound* of S, if for each $s \in S$ we have $s \leq M$ [resp. $s \geq M$]. If there is an upper [resp. lower] bound $M \in X$, then we call S *bounded above* [resp. *bounded below*].

$M_0 \in X$ is called the *least upper bound* or *supremum* [resp. *greatest lower bound* or *infimum*] of $S \subseteq X$ if for all upper [lower] bounds $M \in X$ we have $M_0 \leq M$ [resp. $M_0 \geq M$]. The least upper bound [resp. greatest lower bound] of the set S is denoted by $\sup S$ [resp. $\inf S$].

Definition 1.26. (LUP) An ordered set X has the *least upper bound property* if any nonempty subset S of X which is bounded above has a least upper bound (in X).

Example 1.27. The set of rational numbers $\mathbb{Q}$ does not have the least upper bound property. For example, the set $\{q \in \mathbb{Q} : q^2 < 2\}$ is bounded, for example, by 2, but has no least upper bound in $\mathbb{Q}$.

Theorem 1.28. *The set $\mathbb{R}$, that is, the set of Dedekind cuts in $\mathbb{Q}$, with the linear* **order** *'$\leq$' defined by*

$$A|B \leq C|D \quad \textit{if} \quad A \subseteq C$$

has the least upper bound property

Proof. Let $S = \{A_i|B_i,\ i \in I\}$ be a non empty set of Dedekind cuts in $\mathbb{Q}$ which is bounded by $C|D$. Set $A = \bigcup_{i \in I} A_i$ and $B = \mathbb{Q} \setminus A$. First, observe that $A|B$ is indeed a Dedekind cut:

i. $A \cup B = \mathbb{Q}$ by definition.

ii. $A \neq \emptyset$ since S is not empty.

iii. $B \neq \emptyset$ since $D \subset B$.

iv. $A \cap B = \emptyset$ by definition.

v. For all $a \in A$ and $b \in B$ we have $a < b$ since $a \in A_i$ for some i and $b \in B$ implies $b \notin A_i$ for all i.

vi. If A would contain a largest element a, any set A_i with $a \in A_i$ would obviously also contain a largest element.

Clearly, $A_i \subseteq A$ implies $A_i|B_i \leq A|B$ for all $i \in I$, so $A|B$ is an upper bound of S. Suppose there would exist an upper bound $E|F$ of S with $E|F < A|B$. Then $A_i \subseteq E$ for all $i \in I$ and, hence, $A = \bigcup_{i \in I} A_i \subseteq E \subsetneq A$, a contradiction. □

Remark 1.29. We can embed rational numbers in $\mathbb{R}$ via

$$p \mapsto p^\star := \{q \in \mathbb{Q} :\ q < p\}|\{q \in \mathbb{Q} :\ q \geq p\}.$$

A cut of the form $p^\star := \{q \in \mathbb{Q} :\ q < p\}|\{q \in \mathbb{Q} :\ q \geq p\}$, $p \in \mathbb{Q}$ is called rational cut in $\mathbb{Q}$. The embeddings discussed so far are $\mathbb{N} \hookrightarrow \mathbb{Z} \hookrightarrow \mathbb{Q} \hookrightarrow \mathbb{R}$. Since $\hookrightarrow$ denotes injective maps which respect algebraic properties, we shall omit later the * notation and identify elements in the domain with the corresponding elements in the range. That is, we shall write

$$\mathbb{N} \subseteq \mathbb{Z} \subseteq \mathbb{Q} \subseteq \mathbb{R}.$$

At this point of time, we have not defined any algebraic operations on $\mathbb{R}$ (the set of Dedekind cuts in $\mathbb{Q}$), but we will do this shortly.

Definition 1.30. On $\mathbb{R}$, that is, on the set of Dedekind cuts in $\mathbb{Q}$, we define:

i. A linear **order** '$\leq$' on $\mathbb{R}$ via $A|B \leq C|D$ if $A \subseteq C$.

ii. For $x = A|B,\ y = C|D \in \mathbb{R}$ we set

$$E := \{e \in \mathbb{Q} : \quad \text{there exists } a \in A \text{ and } c \in C \text{ with } e = a + c\}, \qquad F := \mathbb{Q} \backslash E$$

and define **addition** on $\mathbb{R}$ via

$$x + y = A|B + C|D := E|F.$$

Further we set $-x = A^-|B^-$, with $A^- = \big\{-b, b \in B \backslash \{\text{smallest element of } B \text{ (if it exists)}\}\big\}$ and $B^- = \mathbb{Q} \backslash A^-$.
(Note that $-(-x) = x$, that $x + (-x) = 0^*$ for all $x \in \mathbb{R}$, that $x \geq 0$ if and only if $-x \leq 0$, and that $q^* + \widetilde{q}^* = (q + \widetilde{q})^*$ and $(-q)^* = -q^*$ for all $q, \widetilde{q} \in \mathbb{Q}$.)

iii. For $x = A|B \geq 0^*,\ y = C|D \geq 0^* \in \mathbb{R}$ we set

$$G := \{e \in \mathbb{Q} : \quad e \leq 0 \text{ or there exists } a > 0 \in A \text{ and } c > 0 \in C \text{ with } e = a{\cdot}c\}, \qquad H := \mathbb{Q} \backslash G$$

and define the **product**

$$x \cdot y = A|B \cdot C|D := G|H.$$

If $x \geq 0$ and $y < 0$ set $x \cdot y = -(x \cdot (-y))$, if $x < 0$ and $y \geq 0$ set $x \cdot y = -((-x) \cdot y)$, and if $x < 0$ and $y < 0$ set $x \cdot y = (-x) \cdot (-y)$. Hence, we have (well) defined **multiplication**

$$\cdot : \mathbb{R} \times \mathbb{R} \longrightarrow \mathbb{R}, \qquad (x, y) \mapsto x \cdot y$$

(Note that $q^* \cdot \widetilde{q}^* = (q\widetilde{q})^*$ for all $q, \widetilde{q} \in \mathbb{Q}$.)

Theorem 1.31. *The set of Dedekind cuts in $\mathbb{Q}$ denoted by $\mathbb{R}$ together with the order, the two binary operations addition and multiplication defined above is an archimedean ordered field which satisfies the least upper bound property.*

Theorem 1.32. Uniqueness of the real number system. *$\mathbb{R}$ is unique in the following sense: Let F be an archimedean ordered field which has the least upper bound property. Then there exists a bijective mapping $u : F \longrightarrow \mathbb{R}$ which preserves addition, multiplication and order.*

Proof. (Sketch) Let F be an archimedean ordered field with the least upper bound property. First note that $1_F >_F 0_F$ since $1_F \neq_F 0_F$ and if $1_F <_F 0_F$ we get $-1_F >_F 0_F$ by (O3) and $1_F = (-1_F)(-1_F) >_F 0_F$ by (O3), a contradiction to (O1). Further, observe that $\mathbb{N}$ can be embedded into F via

$$i : \mathbb{N} \longrightarrow F, \quad n \mapsto n_F = \underbrace{1_F + 1_F + \ldots + 1_F}_{n-times}.$$

By definition we have $n_F + m_F = (n+m)_F$. The injectivity of this mapping follows from an inductive argument using $n_F + 1_F >_F n_F + 0_F$. Let us also note that implies that the order on $\mathbb{N}$ is preserved under the embedding i, a very important fact as we shall see later. Further, all $n_F > 0_F$ have an inverse element with respect to addition in F and we may extend i injectively to $\mathbb{Z}$ by setting $n \mapsto -(-n)_F$ for $n < 0$. We can show that $n_F + m_F = (n+m)_F$ still holds, now for all $n, m \in \mathbb{Z}$. Note that (F1) together with (O3) on F implies that $-1_F <_F 0$, since else,we would have $-1_F >_F 0_F$ and $0_F >_F 1_F$.

Further, we can use the same strategy to extend i to cover al rational numbers by setting

$$i : \mathbb{Q} \longrightarrow F, \quad \frac{n}{m} \mapsto \frac{n_F}{m_F} = n_F \cdot m_F^{-1}.$$

(To detail this proof, we would have to show that i is well defined, that is, that the image of q under i does not depend on the particular representation of q as fraction of integer and natural number.)

Note that, again, we have $0 < \frac{n}{m} < \frac{\widetilde{n}}{\widetilde{m}}$ if and only if $0_F <_F \frac{n_F}{m_F} <_F \frac{\widetilde{n}_F}{\widetilde{m}_F}$ due to (O3) since else $n_F \cdot_F \widetilde{m}_F > \widetilde{n}_F \cdot_F m_F$. Further $q_F + r_F = (q+r)_F$ and $q_F \cdot_F r_F = (q \cdot_F r)_F$ holds for all $q, r \in \mathbb{Q}$.

After having observed that any ordered field contains a copy of $\mathbb{Q}$ as an ordered subfield, we can proceed to define the "uniqueness" map u:

$$u : F \longrightarrow \mathbb{R}, \quad x \mapsto A_x | B_x = \{q \in \mathbb{Q} : \quad q_F <_F x\} | \{q \in \mathbb{Q} : \quad q_F \geq_F x\}.$$

It remains to show that u is well defined (are these elements on the right really Dedekind cuts?), it preserves addition, multiplication, and order, and that u is bijective. Note that we still have not used the fact that the order on F is archimedean and that F has the least upper bound property.

So let us first look whether the map is well defined. Clearly $A_x \cap B_x = \emptyset$ and $A_x \cup B_x = \mathbb{Q}$. If $x >_F 0_F$ we have $0 \in A_x$ and $B_x \neq \emptyset$ since the archimedean property implies the existence of $n \in \mathbb{N}$ such that

$$n_F = \underbrace{1_F +_F 1_F +_F \ldots +_F 1_F}_{n-times} > x$$

and therefore $n_F \in B_x$. If $x \leq_F 0_F$ we get $B_x \neq \emptyset$ cheaply and we can use a similar argument as above to show that $A_x \neq \emptyset$.

Transitivity shows that for $a \in A_x$ and $b \in B_x$ we have $a_F < x \leq b_F$ and therefore $a \leq b$.

To show that A_x has no largest element, we need to show the following fact, which we shall repeatedly use not only in this proof.

Claim: Let F be an archimedean ordered field which has the least upper bound property and let $x, y \in F$. If $x < y$, then exists $q \in \mathbb{Q}$ such that $x < q_F < y$.

Proof of the claim: Fix $x, y \in F$ with $x < y$. Then $y - x > 0$ and therefore $(y-x)^{-1} > 0$. Pick $m_F > (y-x)^{-1} > 0$. Set $u = \sup\{n \in \mathbb{Z} : \frac{n_F}{m_F} \leq x\}$. Then $x < \frac{u_F +_F 1_F}{m_F} < y$, since $\frac{u_F +_F 1_F}{m_F} > y$ would imply $\frac{u_F +_F 1_F}{m_F} > y > x \geq \frac{u_F}{m_F}$ and $\frac{1_F}{m_F} = \frac{u_F +_F 1_F}{m_F} - \frac{u_F}{m_F} > y - x > \frac{1}{m_F}$, a contradiction.

The set A_x has no largest element, since for any q_F, $(q \in \mathbb{Q})$ in A_x we can find $\widetilde{q}_F$, $(\widetilde{q} \in \mathbb{Q})$ with $x > \widetilde{q}_F > q_F$.

We have shown that $A_x | B_x \in \mathbb{R}$, let us now check surjectivity of u. Let $A|B$ be any cut in $\mathbb{Q}$. Set $A_F = \{q_F \in F : q \in A\}$ and $x_{A|B} = \sup A_F$ which exists due to the l.u.b. property of F. It is easy to see that $u(x_{A|B}) = A_x | B_x = A|B$.

Injectivity follows from the claim proven above (why?). The mapping u preserves multiplication and addition since it does fulfill these properties on $\mathbb{Q}$ and due to the definition of $\mathbb{R}$ and u. □

That's it for Dedekind cuts, we are done. From now on, we will think of real numbers as elements on the real line, its elements are denoted with letters such as $x, y, a, b, \alpha, \beta, \ldots$. Using the order $\leq$ on $\mathbb{R}$ as well as $<$ defined by $a < b$ if $a \leq b$ and $a \neq b$, we can define closed intervals $[a, b] = \{c \in \mathbb{R} : a \leq c \leq b\}$, $[a, \infty) = \{c \in \mathbb{R} : a \leq c\}$, $(-\infty, b] = \{c \in \mathbb{R} : c \leq b\}$, half closed intervals $(a, b] = \{c \in \mathbb{R} : a < c \leq b\}$, $[a, b) = \{c \in \mathbb{R} : a \leq c < b\}$ and open intervals $(a, b) = \{c \in \mathbb{R} : a < c < b\}$, $(a, \infty) = \{c \in \mathbb{R} : a < c < b\}$, $(-\infty, b) = \{c \in \mathbb{R} : c < b\}$ with $a \leq b \in \mathbb{R}$. Also, $\mathbb{R} = (-\infty, \infty)$ can be considered an interval. The terminology "open" and "closed" intervals will become apparent when we study metric and topological spaces.

Theorem 1.33. *For every real number $x > 0$ and $n \in \mathbb{N}$ exists exactly one real number $y > 0$ with $y^n = x$. This y is called n–th root of x and is denoted by $x^{\frac{1}{n}}$ or $\sqrt[n]{x}$.*

Theorem 1.34. Nested Interval Property. *Let $I_n = [a_n, b_n] = \{x \in \mathbb{R} : a_n \leq x \leq b_n\} \subset \mathbb{R}$ be closed intervals with $I_n \supseteq I_{n+1}$ for all $n \in \mathbb{N}$. Then $\bigcap_{n \in \mathbb{N}} I_n \neq \emptyset$.*

Proof. The set of left endpoints $\{a_n : n \in \mathbb{N}\}$ is not empty and is bounded by, for example, b_1 (but also by all other b_n as we shall use below). Hence, $\{a_n : n \in \mathbb{N}\}$ has a least upper bound $\alpha = \sup\{a_n : n \in \mathbb{N}\} = \sup_{n \in \mathbb{N}} a_n$. The fact that α is an upper bound of the set implies $a_n \leq \alpha$ for all n. The fact that it is the least upper bound also shows that $\alpha \leq b_n$ for all n as, given one b_n with $b_n < \alpha$, we would have found a smaller lower bound, a contradiction. We conclude that $a_n \leq \alpha \leq b_n$ for all n, that is, $\alpha \in \bigcap_{n \in \mathbb{N}} [a_n, b_n]$. □

Definition 1.35. A *sequence* a in a set X is a function $a \colon \mathbb{N} \longrightarrow X, n \mapsto a(n)$. Note that by convention we shall write a_n instead of $a(n)$, and a is often denoted by $(a_n)_{n \in \mathbb{N}}$ or $\{a_n\}_{n \in \mathbb{N}}$. Do not confuse the sequence $a = (a_n)_{n \in \mathbb{N}} = \{a_n\}_{n \in \mathbb{N}}$ with the set $\{a_n,\ n \in \mathbb{N}\} = \mathcal{R}_a$.

Definition 1.36. A set X is *countable* if there is a surjective function (sequence) $a \colon \mathbb{N} \longrightarrow X, n \mapsto a(n)$.

Example 1.37.
  i. Finite sets are countable. For example, consider $X = \{0, 1, 2\}$ and define $a : \mathbb{N} \longrightarrow X$ by $a_1 = 0$, $a_2 = 1$, $a_n = 2$ for all $n \geq 3$.

ii. The integers $\mathbb{Z}$ are countable. Consider $a : \mathbb{N} \longrightarrow \mathbb{Z}$ given by $a_1 = 0$, $a_2 = -1$, $a_3 = 1$, $a_4 = -2$, $a_5 = 2$, and so on. That is

$$a_n = \begin{cases} \frac{n-1}{2}, & \text{if } n \text{ is odd;} \\ -\frac{n}{2}, & \text{if } n \text{ is even.} \end{cases}$$

Remark 1.38. Some authors define a set to be countable if there exists as a bijective function (sequence) $a\colon \mathbb{N} \longrightarrow X, n \mapsto a(n)$. Then, different from this lecture, finite sets are not countable! Be aware of both definitions of countability when reading textbooks.

Theorem 1.39. *If the sets $A_m \subset X$, $m \in \mathbb{N}$, are countable, then $\bigcup_{m\in\mathbb{N}} A_m$ is countable.*

Proof. Let $a^m : \mathbb{N} \longrightarrow A_m$ be surjective, that is, $\{a^m_n,\ n \in \mathbb{N}\} = A_m$ for $m \in N$. The elements of $\bigcup_{m\in\mathbb{N}} A_m$ appear in the table

$$\begin{array}{cccccccc}
a^1_1 & a^1_2 & a^1_3 & a^1_4 & a^1_5 & a^1_6 & a^1_7 & \dots \\
a^2_1 & a^2_2 & a^2_3 & a^2_4 & a^2_5 & a^2_6 & a^2_7 & \dots \\
a^3_1 & a^3_2 & a^3_3 & a^3_4 & a^3_5 & a^3_6 & a^3_7 & \dots \\
a^4_1 & a^4_2 & a^4_3 & a^4_4 & a^4_5 & a^4_6 & a^4_7 & \dots \\
a^5_1 & a^5_2 & a^5_3 & a^5_4 & a^5_5 & a^5_6 & a^5_7 & \dots \\
a^6_1 & a^6_2 & a^6_3 & a^6_4 & a^6_5 & a^6_6 & a^6_7 & \dots \\
a^7_1 & a^7_2 & a^7_3 & a^7_4 & a^7_5 & a^7_6 & a^7_7 & \dots \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots
\end{array}$$

and, enumerating diagonally, that is,

$$b_1 = a^1_1,\ b_2 = a^2_1,\ b_3 = a^1_2,\ b_4 = a^3_1,\ b_5 = a^2_2,\ b_6 = a^1_3,\ b_7 = a^4_1,\ b_8 = a^3_2,\ b_9 = a^2_3, \dots$$

defines a surjective map $b : \mathbb{N} \longrightarrow \bigcup_{m\in\mathbb{N}} A_m$. □

Corollary 1.40. $\mathbb{Q}$ *is countable.*

Proof. Clearly, $\mathbb{Z}$ is countable (see the example below), and so are the sets $\frac{1}{m}\mathbb{Z} = \{\frac{n}{m},\ n \in Z\}$ for $m \in \mathbb{N}$. Clearly, $\mathbb{Q} = \bigcup_{m\in\mathbb{N}} \frac{1}{m}\mathbb{Z}$, so $\mathbb{Q}$ is countable by Theorem 1.39. □

Theorem 1.41. *The set containing all sequences with values in $\{0, 1, 2, \dots, n\}$, $n \geq 1$, is not countable.*

Proof. Assume that the set X of sequences with values in $\{0, 1, 2, \dots, n\}$ are countable. Then exists a surjective map $b : \mathbb{N} \longrightarrow X$, $m \mapsto b^m \in X$, so $X = \{b^m,\ m \in \mathbb{N}\}$. These countably many sequences $b^m : \mathbb{N} \longrightarrow \{0, 1, 2, \ldots, n\}$ in the following way

$$\begin{array}{cccccccccc} b^1 = & b^1_1, & b^1_2, & b^1_3, & b^1_4, & b^1_5, & b^1_6, & b^1_7, & \ldots \\ b^2 = & b^2_1, & b^2_2, & b^2_3, & b^2_4, & b^2_5, & b^2_6, & b^2_7, & \ldots \\ b^3 = & b^3_1, & b^3_2, & b^3_3, & b^3_4, & b^3_5, & b^3_6, & b^3_7, & \ldots \\ b^4 = & b^4_1, & b^4_2, & b^4_3, & b^4_4, & b^4_5, & b^4_6, & b^4_7, & \ldots \\ b^5 = & b^5_1, & b^5_2, & b^5_3, & b^5_4, & b^5_5, & b^5_6, & b^5_7, & \ldots \\ b^6 = & b^6_1, & b^6_2, & b^6_3, & b^6_4, & b^6_5, & b^6_6, & b^6_7, & \ldots \\ b^7 = & b^7_1, & b^7_2, & b^7_3, & b^7_4, & b^7_5, & b^7_6, & b^7_7, & \ldots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{array}$$

Define the sequence $b^0 : \mathbb{N} \longrightarrow \{0, 1, 2, \ldots, n\}$ by $b^0_n = 1$ if $b^n_n = 0$ and $b^0_n = 0$ if $b^n_n \neq 0$. This sequence differs from all sequences in b^m (in the m-th entry) and is therefore does not appear in our list. (Sequences are identical if and only if all entries are equal.) Hence, our list was not exhaustive, a contradiction. □

Theorem 1.42. $\mathbb{R}$ *is not countable.*

Proof. We shall embed the space of sequences in $\{0, 1, 2, \ldots, 8\}$ in $\mathbb{R}$ by mapping a sequence $b : \mathbb{N} \longrightarrow \{0, 1, 2, \ldots, 8\}$ to

$$i(b) = \sup\{b_1 \cdot 10^{-1} + b_2 \cdot 10^{-2} + b_3 \cdot 10^{-3} + \ldots + b_N \cdot 10^{-N},\ N \in \mathbb{N}\}.$$

It is easily seen that this map is injective. (Note that we did not consider $b : \mathbb{N} \longrightarrow \{0, 1, 2, \ldots, 8, 9\}$ as then, for example, the sequences $1, 0, 0, 0, \ldots$ and $0, 9, 9, 9, 9, \ldots$ would map to the same real number.)

Clearly, if $\mathbb{R}$ were countable, so would be any subset of $\mathbb{R}$, for example, the image under the injective map i. But by injectivity, this would imply that also its domain is countable, a contradiction. □

1.4. Complex numbers

We shall now define the complex number system.

Definition 1.43. The cartesian product $\mathbb{R} \times \mathbb{R}$ together with the binary operations

$$\begin{array}{rcll} + & : & (\mathbb{R}\times\mathbb{R})\times(\mathbb{R}\times\mathbb{R}) \longrightarrow \mathbb{R}\times\mathbb{R}, & ((a,b),(c,d)) \mapsto (a+c, b+d) \\ \cdot & : & (\mathbb{R}\times\mathbb{R})\times(\mathbb{R}\times\mathbb{R}) \longrightarrow \mathbb{R}\times\mathbb{R}, & ((a,b),(c,d)) \mapsto (ac-bd, ad+bc) \end{array}$$

form a field with additive neutral element $(0,0)$ and multiplicative neutral element $(1,0)$ which is called the *field of complex numbers*. It is denoted by $\mathbb{C}$.

Theorem 1.44. *The map $G : \mathbb{R} \longrightarrow \mathbb{C}$, $a \mapsto (a,0)$ is an embedding of the real numbers into the complex numbers, that is, G is injective and we have for all $a, b \in \mathbb{R}$*

$$G(a+b) = G(a) + G(b) \quad \textit{and} \quad G(ab) = G(a)\cdot G(b).$$

Hence, we can consider $\mathbb{R}$ as a subfield of $\mathbb{C}$.

Proof. This result follows from the definition of addition and multiplication on $\mathbb{C}$. □

Remark 1.45. For $i := (0,1)$, we have $i^2 = (0,1)\cdot(0,1) = (0-1, 0+0) = (-1,0)$, and for $a, b \in \mathbb{R}$ we have $G(a) + G(b)\cdot i = (a,b)$. From now on we shall consider $\mathbb{R}$ as a subfield of $\mathbb{C}$ and drop the embedding G in our description of complex numbers. Hence, we shall write $a + bi = (a,b) \in \mathbb{C}$.

Definition 1.46. For $z = a + bi \in \mathbb{C}$ with $a, b \in \mathbb{R}$ we shall call $a = \Re(z) \in \mathbb{R}$ the *real part* of z and $b = \Im(z) \in \mathbb{R}$ the *imaginary part* of z. The *conjugate* of z is $\overline{z} = a - bi$ and the *absolute value* of z is $|z| = \sqrt{a^2+b^2}$. The *argument* of $z \neq 0$ is $\arg(z) = z \cdot |z|^{-1}$, so $z = |z|\arg(z)$.

Proposition 1.47. *For all $z = a + bi$, $w = c + di \in \mathbb{C}$ with $a, b, c, d \in \mathbb{R}$ we have*

$$\begin{array}{rcl} \Re(z+w) & = & \Re(z) + \Re(w) \\ \Im(z+w) & = & \Im(z) + \Im(w) \\ |\Re(z)| & \leq & |z| \\ |\Im(z)| & \leq & |z| \\ \overline{z+w} & = & \overline{z} + \overline{w} \\ \overline{zw} & = & \overline{z}\,\overline{w} \\ z\overline{z} & = & |z|^2 \\ z + \overline{z} & = & 2\Re(z) \\ z - \overline{z} & = & 2i\Im(z) \\ |z||w| & = & |zw| \\ |z| + |w| & \geq & |z+w| \\ z^{-1} & = & \dfrac{1}{|z|^2}\overline{z} \quad \textit{if } z \neq 0. \end{array}$$

Proof. All the results follow from computation. As example, we shall prove $|z| + |w| \geq |z + w|$. To this end, we compute

$$\begin{aligned}|z+w|^2 &= (z+w)\overline{(z+w)} = (z+w)(\overline{z}+\overline{w}) = z\overline{z} + w\overline{z} + z\overline{w} + w\overline{w} = |z|^2 + 2\Re(z\overline{w}) + |w|^2 \\ &\leq |z|^2 + 2|z\overline{w}| + |w|^2 = |z|^2 + 2|z|\,|\overline{w}| + |w|^2 = |z|^2 + 2|z|\,|w| + |w|^2 = (|z| + |w|)^2.\end{aligned}$$

It remains to argue that if $a, b \geq 0$ satisfy $a^2 \geq b^2 > 0$, then $a \geq b$. Assume $b > a$. Since $\mathbb{R}$ is an ordered field, we conclude that $b - a > 0$ and $b + a > 0$. Using again that $\mathbb{R}$ is an ordered field, we conclude that

$$0 < (b-a)(b+a) = b^2 - a^2$$

and, hence, $b^2 > a^2$, a contradiction. □

Remark 1.48. A more geometrical treatise of complex numbers is contained in the homework.

2. CONVERGENCE OF SEQUENCES IN METRIC SPACES AND NUMERIC SERIES

The goal of this section is to discuss real and complex valued sequences and series. Many results concerning real and complex sequences hold in a more general setup, that is, in metric spaces. In order to avoid the repetition of arguments, we shall phrase some results in the metric space setup, nevertheless, at this point of time it might be best to think of only two metric spaces, that is, the space of real and the space of complex numbers. In these special cases, the distance between two numbers x and y is $d(x, y) = |x - y|$.

2.1. Sequences in metric spaces

Definition 2.1. A set X together with a binary function $d : X \times X \longrightarrow \mathbb{R}$ is a *metric space* with *metric* d if d satisfies

i. $d(x, \widetilde{x}) > 0$ if $x \neq \widetilde{x}$ and $d(x, x) = 0$ for all $x \in X$,

ii. $d(x, \widetilde{x}) = d(\widetilde{x}, x)$ for all $x, \widetilde{x} \in X$,

iii. $d(x, \widetilde{\widetilde{x}}) \leq d(x, \widetilde{x}) + d(\widetilde{x}, \widetilde{\widetilde{x}})$ for all $x, \widetilde{x}, \widetilde{\widetilde{x}} \in X$.

The function d is called *metric* or *distance function* on the set X and we shall denote a metric space by (X, d) or simply by X if it is well understood which metric d on X is being considered.

Examples 2.2. i. The set of real numbers $\mathbb{R}$ with metric $d_2(x, y) = |x - y|$ is a metric space. If no other metric is explicitly mentioned, we shall always consider $\mathbb{R}$ to be equipped with the *euclidean metric* d_2.

ii. The set of complex numbers $\mathbb{C}$ with metric $d_2(x, y) = |x-y| = \sqrt{(\mathrm{Re}(x - y))^2 + (\mathrm{Im}(x - y))^2}$ is a metric space. If no other metric is explicitly mentioned, we shall always consider $\mathbb{C}$ to be equipped with the d_2 metric.

iii. Given any set X, we can define a metric on X via

$$d_0(x, y) = \begin{cases} 0 & \text{if } x = y; \\ 1 & \text{else} \end{cases} \qquad \text{for } x, y \in X.$$

This metric is called *discrete metric* on X.

Definition 2.3. A sequence $(x_n)_{n\in\mathbb{N}}$ in $\mathbb{R}$ is said to *converge* to $x_0 \in \mathbb{R}$ if for all $\varepsilon > 0$ exists $N \in \mathbb{N}$ such that

$$|x_n - x_0| < \varepsilon \quad \text{for all } n \geq N.$$

If $(x_n)_{n\in\mathbb{N}}$ converges to x_0 in $\mathbb{R}$ we write $\lim_{n\to\infty} x_n = x_0$, or $x_n \xrightarrow{n\to\infty} x_0$, or simply $x_n \longrightarrow x_0$. The element $x_0 \in \mathbb{R}$ is called *limit* of $(x_n)_{n\in\mathbb{N}}$ in $\mathbb{R}$.

Definition 2.4. A sequence $(x_n)_{n\in\mathbb{N}}$ in a metric space (X,d) is said to *converge* to $x_0 \in X$ if for all $\varepsilon > 0$ (that is, $\varepsilon \in \mathbb{R}$ with $\varepsilon >_{\mathbb{R}} 0_{\mathbb{R}}$) exists $N \in \mathbb{N}$ such that

$$d(x_n, x_0) < \varepsilon \quad \text{for all } n \geq N.$$

If (x_n) converges to x_0 in (X,d) we write $\lim\limits_{n\to\infty} x_n = x_0$, or $x_n \xrightarrow{n\to\infty} x_0$, or simply $x_n \longrightarrow x_0$. The element $x_0 \in X$ is called *limit* of (x_n) in (X,d).

Examples 2.5.

i. The sequence $(\frac{1}{n})_{n\in\mathbb{N}}$ in $(\mathbb{R}, d_2)$ converges to $0 \in \mathbb{R}$.

ii. The sequence $(\frac{1}{n})_{n\in\mathbb{N}}$ in $(\mathbb{R}, d_0)$ does not converge to any $x_0 \in \mathbb{R}$, since for any $x_0 \in \mathbb{R}$ we have $d_0(x_0, x_n) < \frac{1}{2}$ for at most one index $n \in \mathbb{N}$.

Proposition 2.6. *A sequence $(z_n)_n$ in $\mathbb{C}$ converges in $(\mathbb{C}, d_2)$ (or simply in $\mathbb{C}$) if and only if*

$$Re(z_n) \xrightarrow{n\to\infty} Re(z_0) \text{ in } \mathbb{R}$$

and

$$Im(z_n) \xrightarrow{n\to\infty} Im(z_0) \text{ in } \mathbb{R}.$$

That is, sequences converge in $\mathbb{C}$ if and only if both, real and imaginary part converge in $\mathbb{R}$. Therefore, a real valued sequence converges in $\mathbb{R}$ if and only if it converges in $\mathbb{C}$.

Proof. Suppose $z_n \longrightarrow z_0$. We will show that $\Re(z_n) \xrightarrow{n\to\infty} \mathrm{Re}(z_0)$ in $\mathbb{R}$. Fix $\epsilon > 0$ and choose $N_\epsilon \in \mathbb{N}$ with $|z_n - z_0| < \epsilon$ for all $n \geq N_\epsilon$. But then

$$|\Re(z_n) - \Re(z_0)| = |\Re(z_n - z_0)| \leq |z_n - z_0| < \epsilon, \quad n \geq N_\epsilon,$$

so $\Re(z_n) \xrightarrow{n\to\infty} \mathrm{Re}(z_0)$ in $\mathbb{R}$. Replacing Re by Im in the above shows the convergence of the imaginary part.

Now, we suppose $\Re(z_n) \xrightarrow{n\to\infty} \mathrm{Re}(z_0)$ and $\Im(z_n) \xrightarrow{n\to\infty} \Im(z_0)$. Fix $\epsilon > 0$ and choose N_r so that $|\Re(z_n) - \Re(z_0)| < \frac{\epsilon}{2}$ for all $n \geq N_r$ and N_i with $|\Im(z_n) - \Im(z_0)| < \frac{\epsilon}{2}$ for all $n \geq N_i$. Set $N_\epsilon = \max\{N_r, N_i\}$. Then

$$\begin{aligned}
|z_n - z_0| &= |\Re(z_n) + i\Im(z_n) - (\Re(z_0) + i\Im(z_0))| \\
&= |(\Re(z_n) - \Re(z_0) + i(\Im(z_n) - \Im(z_0))| \\
&\leq |\Re(z_n) - \Re(z_0)| + |i(\Im(z_n) - \Im(z_0))| \\
&= |\Re(z_n) - \Re(z_0)| + |\Im(z_n) - \Im(z_0)| \\
&< \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon, \quad n \geq N_\epsilon.
\end{aligned}$$

□

Theorem 2.7. *The limit of a converging sequence in a metric space (X,d) is unique, that is, if $x_n \xrightarrow{n\to\infty} x_0 \in X$ and $x_n \xrightarrow{n\to\infty} \widetilde{x_0} \in X$, then $x_0 = \widetilde{x_0}$.*

Proof. Let $x_n \xrightarrow{n\to\infty} x_0$ and $x_n \xrightarrow{n\to\infty} \widetilde{x}_0$. Fix $\epsilon > 0$ and choose N_ϵ so that $|x_n - x_0| < \frac{\epsilon}{2}$ and $|x_n - \widetilde{x}_0| < \frac{\epsilon}{2}$ for all $n \geq N_\epsilon$. Then

$$0 \leq d(x_0, \widetilde{x}_0) \leq d(x_0, x_{N_\epsilon}) + d(x_{N_\epsilon}, \widetilde{x}_0) < \epsilon.$$

We conclude that $d(x_0, \widetilde{x}_0) < \epsilon$ for all $\epsilon > 0$, hence, $d(x_0, \widetilde{x}_0) = 0$ and $x_0 = \widetilde{x}_0$ follows. □

Definition 2.8. A subset S in a metric space (X, d) is called *bounded* if there is $x_0 \in X$ and $M \in \mathbb{R}^+$ such that $d(x_0, x) \leq M$ for all $x \in S$.

A sequence (x_n) is *bounded* in (X, d) if its range $\{x_n : n \in \mathbb{N}\}$ is a bounded set in (X, d).

Theorem 2.9. *Every converging sequence (x_n) in a metric space (X, d) is bounded.*

Proof. As (x_n) is converging, there exists $x_0 \in X$ and $N \in \mathbb{N}$ with $d(x_n, x_0) < 1$ for all $n \geq N$. Set $M = \max\{d(x_1, x_0), d(x_2, x_0), d(x_3, x_0), \ldots, d(x_{N-1}, x_0)\} + 1$. Clearly, we have $d(x_0, x_n) < M < \infty$. for all $n \in N$. □

Definition 2.10. A sequence (x_n) in $\mathbb{R}$ is

i. *monotonically increasing* if $x_n \leq x_{n+1}$ for all $n \in \mathbb{N}$,

ii. *strictly monotonically increasing* if $x_n < x_{n+1}$ for all $n \in \mathbb{N}$,

iii. *monotonically decreasing* if $x_n \geq x_{n+1}$ for all $n \in \mathbb{N}$, and

iv. *strictly monotonically decreasing* if $x_n > x_{n+1}$ for all $n \in \mathbb{N}$.

A sequence is called monotone if it is either monotonically increasing or decreasing.

Theorem 2.11. *Monotone sequences converge in $\mathbb{R}$ if and only if they are bounded.*

Proof. Converging sequences are bounded, so all that remains to be shown is that a sequence which is monotonic and bounded does converge.

Let (x_n) be a monotonically increasing, bounded sequence. The set $\{x_n,\ n \in \mathbb{N}\} \neq \emptyset$ is bounded and $\mathbb{R}$ has the least upper bound property, so $x_0 = \sup\{x_n,\ n \in \mathbb{N}\}$ exits. We claim that $x_n \xrightarrow{n\to\infty} x_0$. To this end, fix $\epsilon > 0$. As x_0 is the least upper bound of $\{x_n,\ n \in \mathbb{N}\}$, $x_0 - \epsilon$ is not an upper bound of $\{x_n,\ n \in \mathbb{N}\}$ and there exists some $N_\epsilon > 0$ with $x_{N_\epsilon} > x_0 - \epsilon$. Now, for $n \geq N_\epsilon$ we have

$$x_0 - \epsilon < x_{N_\epsilon} \leq x_n \leq x_0 < x_0 + \epsilon.$$

Substracting x_0, we get $-\epsilon < x_n - x_0 < \epsilon$ which is $|x_n - x_0| < \epsilon$ for all $n \geq N_\epsilon$. □

Theorem 2.12. Algebraic Limit Theorem. *If $a_n \xrightarrow{n\to\infty} a_0$ and $b_n \xrightarrow{n\to\infty} b_0$ in $\mathbb{C}$. Then*

i. $ca_n \xrightarrow{n\to\infty} ca_0, \quad c \in \mathbb{C}$,

ii. $(a_n + b_n) \xrightarrow{n\to\infty} a_0 + b_0$,

iii. $a_n b_n \xrightarrow{n\to\infty} a_0 b_0$, *and* *iv.* $\frac{1}{a_n} \xrightarrow{n\to\infty} \frac{1}{a_0}$ *if* $a_0, a_n \neq 0$ *for* $n \in \mathbb{N}$.

Proof. Statement i is simple, statement ii is proven similarly to Proposition 2.6. For statement iii use the boundedness of converging sequences to choose M such that $|a_n| \leq M$ and $|b_0| \leq M$ for all $n \in \mathbb{N}$. Fix $\epsilon > 0$ and N_ϵ with $|a_n - a_0| < \frac{\epsilon}{2M}$ and $|b_n - b_0| < \frac{\epsilon}{2M}$ for all $n \geq N_\epsilon$. Then

$$\begin{aligned}|a_nb_n - a_0b_0| &= |a_nb_n - a_nb_0 + a_nb_0 - a_0b_0| \leq |a_nb_n - a_nb_0| + |a_nb_0 - a_0b_0| \\ &= |a_n||b_n - b_0| + |b_0||a_n - a_0| \leq M\frac{\epsilon}{2M} + M\frac{\epsilon}{2M} = \epsilon, \quad n \geq N_\epsilon.\end{aligned}$$

Let us now show statement iv. As $a_n \xrightarrow{n\to\infty} a_0$ and $|a_0|/2 > 0$, there exists N such that $|a_n - a_0| \leq |a_0|/2$ for $n \geq N$, so $a_n \geq |a_0|/2$ for $n \geq N$.

Fix $\epsilon > 0$ and pick $N_\epsilon \in \mathbb{N}$ with $N_\epsilon \geq N$ and $|a_n - a_0| < \epsilon|a_0|^2/w$ for all $n \geq N_\epsilon$. Then

$$\left|\frac{1}{a_n} - \frac{1}{a_0}\right| = \left|\frac{a_0 - a_n}{a_na_0}\right| = |a_0 - a_n|\frac{1}{|a_n||a_0|} < \frac{\epsilon|a_0|^2}{2}\frac{2}{|a_0|}\frac{1}{|a_0|} = \epsilon, \quad n \geq N_\epsilon. \quad \square$$

Theorem 2.13. Order Limit Theorem. *If* $a_n \xrightarrow{n\to\infty} a_0$ *and* $b_n \xrightarrow{n\to\infty} b_0$ *in* $\mathbb{R}$ *with* $a_n \leq b_n$ *for all* $n \in \mathbb{N}$*, then* $a_0 \leq b_0$.

Proof. Assume $b_0 < a_0$. Then $\frac{a_0 - b_0}{2} > 0$ and there exists N with $|a_n - a_0| < \frac{a_0 - b_0}{2}$ and $|b_n - b_0| < \frac{a_0 + b_0}{2}$ for all $n \geq N$. But this implies in particular

$$b_N < b_0 + \frac{a_0 - b_0}{2} = \frac{a_0 + b_0}{2} = a_0 - \frac{a_0 - b_0}{2} < a_N,$$

a contradiction to $a_n \leq b_n$ for all $n \in \mathbb{N}$. $\square$

Theorem 2.14. Squeezing Theorem. *If* $a_n \leq b_n \leq c_n$ *for all* $n \in \mathbb{N}$ *and* $a_n \xrightarrow{n\to\infty} a_0$ *and* $c_n \xrightarrow{n\to\infty} a_0$ *in* $\mathbb{R}$*, then* (b_n) *converges with* $b_n \xrightarrow{n\to\infty} a_0$.

Proof. Homework. $\square$

Examples 2.15.

i. For $p > 0$ we have $\lim_{n\to\infty} \frac{1}{n^p} = 0$.

ii. For $p > 0$ we have $\lim_{n\to\infty} \sqrt[n]{p} = 1$.

iii. $\lim_{n\to\infty} \sqrt[n]{n} = 1$.

iv. For $p > 0$ and $\alpha \in \mathbb{R}$ we have $\lim_{n\to\infty} \frac{n^\alpha}{(1+p)^n} = 0$.

v. If $x \in \mathbb{C}$ with $|x| < 1$, then $\lim_{n\to\infty} x^n = 0$.

Proof. These examples are taken from Rudin's book *Principles of Mathematical Analysis.* We shall only discuss *ii* here.

The result is obvious for $p = 1$. For $p > 1$ we have $\sqrt[n]{p} > 1$ (since else $p \leq 1$) and $x_n = \sqrt[n]{p} - 1 > 0$. We have

$$p = (x_n + 1)^n = \sum_{k=0}^{n} \binom{n}{k} x_n^k \, 1^{n-k} \geq 1 + n x_n,$$

where we used the non negativity of the summands in the binomial formula. We conclude that $0 \leq x_n \leq \frac{p-1}{n}$. Clearly, with $a_n = 0$ for $n \in \mathbb{N}$ and $b_n = \frac{p-1}{n}$ we have $a_n \xrightarrow{n\to\infty} 0$ and $b_n \xrightarrow{n\to\infty} 0$, so the squeezing theorem implies $x_n \xrightarrow{n\to\infty} 0$. Using the algebraic limit theorem with the sequence $c_n = 1$, we obtain that $\sqrt[n]{p} = x_n + 1 \xrightarrow{n\to\infty} 0 + 1 = 1$. □

Definition 2.16. Let (x_n) be a sequence in (X, d) and let $n_1 < n_2 < n_3 < \ldots$ be a strictly increasing sequence of natural numbers. Then $(x_{n_k})_{k\in\mathbb{N}}$ is called *subsequence* of (x_n).

Example 2.17. Given the sequence $1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \frac{1}{5}, \ldots$, we have $\frac{1}{2}, \frac{1}{4}, \frac{1}{6}, \ldots$ is a subsequence of $1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \frac{1}{5}, \ldots$, but $1, \frac{1}{2}, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \frac{1}{5} \ldots$ and $\frac{1}{2}, 1, \frac{1}{4}, \frac{1}{3}, \frac{1}{6}, \frac{1}{5} \ldots$ are not. In general, $(x_{n_k})_{k\in\mathbb{N}}$ with $x_{n_k} = x_{2k}$ is a subsequence of (x_n).

Theorem 2.18. *Every subsequence $(s_{n_k})_k$ of a convergent sequence $(s_n)_n$ in (X, d) converges to the same limit as $(s_n)_n$.*

Proof. Fix $\epsilon > 0$ and choose N_ϵ with $d(s_n, s_0) < \epsilon$ for all $n \geq N$. Then $d(s_{n_k}, s_0) < \epsilon$ for all $k \geq N$ as $n_k \geq k \geq N$. □

Example 2.19. The sequence $\frac{1}{2}, \dfrac{1}{2+\frac{1}{2}}, \dfrac{1}{2+\frac{1}{2+\frac{1}{2}}}, \dfrac{1}{2+\frac{1}{2+\frac{1}{2+\frac{1}{2}}}}, \ldots$, converges to $\sqrt{2} - 1$ in $\mathbb{R}$.

Proof. Boundedness: The sequence is given by $x_1 = \frac{1}{2}$ and $x_{n+1} = \frac{1}{2+x_n}$, $n \geq 1$, so clearly $x_n \in (0, \frac{1}{2}]$, that is, (x_n) is bounded. In fact, $x_n \in (0, \frac{1}{2}]$ implies that $x_n \geq \frac{1}{2+\frac{1}{2}} = \frac{2}{5}$, so $x_n \in [\frac{2}{5}, \frac{1}{2}]$ for all $n \in \mathbb{N}$.

Monotonicity: If the sequence was monotone, then boundedness would imply convergence. But looking at the first terms, $\frac{1}{2} = 0.5$, $\frac{2}{5} = 0.4$, $\frac{5}{12} = 0.41\overline{6}$, $\frac{12}{29} = 0.4137...$, $\frac{29}{70} = 0.4142...$, we realize that it is not monotone.

But, $x_1 \geq x_3 \geq x_5$ and $x_2 \leq x_4 \leq x_6$. We claim that $x_{2k+3} \leq x_{2k+1}$ and $x_{2k+2} \geq x_{2k}$ for all $k \geq 0$, that is, the subsequence with odd indices is monotonically decreasing and the subsequence with even induces is monotonically increasing. This follows from an inductive argument. We already realized that $x_1 \geq x_3$. The fact that

$$x_{2k+4} - x_{2k+2} = \frac{1}{2 + x_{2k+3}} - \frac{1}{2 + x_{2k+1}} = \frac{2 + x_{2k+1} - (2 + x_{2k+3})}{(2 + x_{2k+3})(2 + x_{2k+1})} = \frac{x_{2k+1} - x_{2k+3}}{(2 + x_{2k+3})(2 + x_{2k+1})}$$

then confirms that $x_4 \geq x_2$ and

$$x_{2k+1} - x_{2k+3} = \frac{1}{2 + x_{2k}} - \frac{1}{2 + x_{2k+2}} = \frac{2 + x_{2k+2} - (2 + x_{2k})}{(2 + x_{2k})(2 + x_{2k+2})} = \frac{x_{2k+2} - x_{2k}}{(2 + x_{2k})(2 + x_{2k+2})},$$

implies $x_5 \leq x_3$. Using the two formulas above in alternating fashion we obtain the claimed monotonicity.

Convergence and limit: We conclude that the subsequences $\{x_{2k+1}\}_{k\in\mathbb{N}}$ and $\{x_{2k}\}_{k\in\mathbb{N}}$ converge to, say x_{odd} and x_{even} in $[\frac{2}{5}, \frac{1}{2}]$ respectively. The algebraic limit theorem implies that

$$x_{even} = \lim_{k\to\infty} x_{2k+2} = \lim_{k\to\infty} \frac{1}{2 + \frac{1}{2+x_{2k}}} = \frac{1}{2 + \frac{1}{2+\lim_{k\to\infty} x_{2k}}} = \frac{1}{2 + \frac{1}{2+x_{even}}} = \frac{2 + x_{even}}{5 + 2x_{even}},$$

so x_{even} solves the quadratic equation $x^2+2x-1 = 0$, that is, $(x-1)^2 = 2$. This equation has two real solutions, namely $\sqrt{2}-1$ and $-\sqrt{2}-1$. As $x_{even} \in [\frac{2}{5}, \frac{1}{2}]$ we conclude that $x_{even} = \sqrt{2}-1$. The same arguments show that $x_{odd} = \sqrt{2}-1$.

Now, it is easy to show that if $\lim_{k\to\infty} x_{2k} = \lim_{k\to\infty} x_{2k+1}$ then (x_n) converges to the same limit. □

Theorem 2.20. Bolzano–Weierstrass Theorem. *Every bounded sequence $(s_n)_n$ in $\mathbb{R}$ has a converging subsequence.*

Proof. As $(s_n)_n$ is bounded, there exists $M > 0$ such that $s_n \in I_0 = [-M, M]$ for all $n \in \mathbb{N}$. Then, either $s_n \in [-M, 0]$ for infinitely many $n \in \mathbb{N}$ or $s_n \in [-M, 0]$ for infinitely many $n \in \mathbb{N}$. Let I_1 be one of the two intervals that is met infinitely often, and $n_1 \in \mathbb{N}$ so that $s_{n_1} \in I_1$. We split I_1 into two intervals of length $M/2$ and choose I_2 to be one the two intervals, one that is met infinitely often by (s_n). Pick $n_2 > n_1$ so that $s_{n_2} \in I_2$. Continue the process.

As the intervals I_n are nested, we have $A = \bigcap_{n\in\mathbb{N}} I_n \neq \emptyset$. Pick $s_0 \in A$. We claim that $s_{n_k} \xrightarrow{k\to\infty} s_0$. To this end, pick $\epsilon > 0$ and choose K so that $2^{-K+1}M < \epsilon$. For $k \geq K$ we have $s_{n_k} \in I_K$ which also contains s_0. As any two points in I_K have distance at most $2^{-K+1}M$, this implies $|s_{n_k} - s_0| \leq 2^{-K+1}M < \epsilon$ for $k > K$. (Note that the uniqueness of limits implies that indeed $A = \{s_0\}$.) □

2.2. The extended real number system, lim sup and lim inf

Definition 2.21. The *extended real number system* is the linear ordered set $\mathbb{R}^* = \mathbb{R} \cup \{+\infty, -\infty\}$ with $-\infty <_{\mathbb{R}^*} x <_{\mathbb{R}^*} y <_{\mathbb{R}^*} +\infty$ for all $x <_{\mathbb{R}} y$ in $\mathbb{R}$.

Note that the field structure on $\mathbb{R}$ cannot be extended (in a meaningful way) to $\mathbb{R}^*$. Nevertheless, it is customary to set

$$\begin{aligned} x + (+\infty) &= +\infty \quad \text{for } x \in \mathbb{R}, \\ x + (-\infty) = x - (+\infty) &= -\infty \quad \text{for } x \in \mathbb{R}, \text{ and} \\ \frac{x}{+\infty} = \frac{x}{-\infty} &= 0 \quad \text{for } x \in \mathbb{R}. \end{aligned}$$

If $x > 0$ we set $x \cdot (+\infty) = +\infty$, $x \cdot (-\infty) = -\infty$, if $x < 0$ then $x \cdot (+\infty) = -\infty$ and $x \cdot (-\infty) = +\infty$.

Further, if for all $M \in \mathbb{R}^+$ there exists $N \in \mathbb{N}$ such that

$$x_n \geq M \quad \text{for all naturals } n \geq N,$$

then we write $\lim\limits_{n\to\infty} x_n = \infty$, or $x_n \xrightarrow{n\to\infty} \infty$, or simply $x_n \longrightarrow \infty$. Correspondingly, if for all $M \in \mathbb{R}^+$ there exists $N \in \mathbb{N}$ such that

$$x_n \le -M \quad \text{for all naturals } n \ge N,$$

then we write $\lim\limits_{n\to\infty} x_n = -\infty$, or $x_n \xrightarrow{n\to\infty} -\infty$, or $x_n \longrightarrow -\infty$.

Proposition 2.22. *The linearly ordered set $\mathbb{R}^*$ has the least upper bound property. Since in addition every subset of $\mathbb{R}^*$ is bounded above by ∞, each non–empty subset of $\mathbb{R}^*$ has a least upper bound.*

Definition 2.23. Let (x_n) be a sequence of real numbers. Set

$$E_{(x_n)} = \{x_0 \in \mathbb{R}^* : \text{ there exists a subsequence } (x_{n_k}) \text{ of } (x_n) \text{ with } x_{n_k} \xrightarrow{k\to\infty} x_0\} \subseteq \mathbb{R}^*$$

and define

$$\begin{aligned} \limsup_{n\to\infty} x_n &= \sup E_{(x_n)} = l.u.b.E_{(x_n)} \quad \in \mathbb{R}^*, \text{ and} \\ \liminf_{n\to\infty} x_n &= \inf E_{(x_n)} = -\, l.u.b.(-E_{(x_n)}) \quad \in \mathbb{R}^*. \end{aligned}$$

Any $x_0 \in E_{(x_n)} \cap \mathbb{R}$ is called *limit point* of the real valued sequence (x_n).

Remark 2.24. $\limsup_{n\to\infty} x_n$ and $\liminf_{n\to\infty} x_n$ are well defined since the set of limit points $E_{(x_n)}$ is never empty. Indeed, if its range $\{x_n,\ n \in \mathbb{N}\}$ is bounded, then exists a subsequence converging to some $x_0 \in \mathbb{R}$. If not, exists $(x_{n_k})_k$ with $x_{n_k} \xrightarrow{k\to\infty} +\infty$ or $x_{n_k} \xrightarrow{k\to\infty} -\infty$

Examples 2.25.

i. Choose (x_n) such that $\{x_n, n \in \mathbb{N}\} = \mathbb{Q}$. Then $\limsup\limits_{n\to\infty} x_n = +\infty$ and $\liminf\limits_{n\to\infty} x_n = -\infty$.

ii. Let $x_n = (-1)^n \left(1 + \frac{1}{n}\right)$ for $n \in \mathbb{N}$. Then $\limsup\limits_{n\to\infty} x_n = +1$ and $\liminf\limits_{n\to\infty} x_n = -1$.

Lemma 2.26. *Let (x_n) be a sequence in $\mathbb{R}$ and $s \in \mathbb{R}$. If $s > \limsup_{n\to\infty} x_n$, then exists $N \in \mathbb{N}$ such that $x_n \le s$ for all $n \ge N$. If $s < \liminf_{n\to\infty} x_n$, then exists $N \in \mathbb{N}$ such that $x_n \ge s$ for all $n \ge N$.*

Proof. Fix (x_n) and $s \in \mathbb{R}$ with $s > \limsup_{n\to\infty} x_n$. We shall show that there exists $N \in \mathbb{N}$ such that $x_n \le s$ for all $n \ge N$. The second assertion follows verbatim.

Suppose that for any $N \in \mathbb{N}$ there exists an index $n_N \ge N$ such that $x_{n_N} > s$. In this case, we can pick n_1 such that $x_{n_1} > s$, then $n_2 \ge n_1 + 1$ with $x_{n_2} > s$, and, inductively $n_{k+1} \ge n_k + 1$, $k \in \mathbb{N}$.

Since (x_{n_k}) is a subsequence of (x_n) and, therefore, any subsequence of (x_{n_k}) is also a subsequence of (x_n), we have $E_{(x_{n_k})_k} \subseteq E_{(x_n)_n}$. Pick $y \in E_{(x_{n_k})_k} \neq \emptyset$ and observe that an application of the order limit theorem to subsequences of $(x_{n_k})_k$ implies $y \ge s$ since $x_{n_k} \ge s$ for all $k \in \mathbb{N}$. The fact that $y \in E_{(x_n)_n}$ implies $\limsup_{n\to\infty} x_n \ge y \ge s > \limsup_{n\to\infty} x_n$, which is nonsense. Contradiction! $\square$

Remark 2.27. A frequent MISTAKE is to assume that if $s \geq \limsup_{n\to\infty} x_n$, then exists N such that $x_n \leq s$ for all $n \geq N$. For example, choose $x_n = \frac{1}{n}$. Then $\limsup_{n\to\infty} x_n = \sup\{0\} = 0 \leq 0$ but $\frac{1}{n} > 0$ for all $n \in \mathbb{N}$.

Theorem 2.28. ORDER LIMIT THEOREM FOR $\limsup$ AND $\liminf$. *If $a_n \leq b_n$ for all $n \in \mathbb{N}$, then $\limsup_{n\to\infty} a_n \leq \limsup_{n\to\infty} b_n$ and $\liminf_{n\to\infty} a_n \leq \liminf_{n\to\infty} b_n$.*

Proof. If $\limsup_{n\to\infty} a_n > \limsup_{n\to\infty} b_n$, then $\limsup_{n\to\infty} a_n > \alpha > \limsup_{n\to\infty} b_n$ for some $\alpha \in \mathbb{R}$. But then exists N such that $b_n \leq \alpha$ for all $n \geq N$. This implies that all limit points of (b_n) are bounded above by α, in particular $\limsup_{n\to\infty} b_n \leq \alpha$, a contradiction. The proof for $\liminf$ is analogous. □

Example 2.29. Let (a_n) be a sequence in $\mathbb{R}$ and set $b_n = \frac{a_1+a_2+\ldots+a_n}{n}$. Then $\limsup_{n\to\infty} b_n \leq \limsup_{n\to\infty} a_n$.

Proof by contradiction. Assume $\limsup_{n\to\infty} b_n > \alpha > \limsup_{n\to\infty} a_n$. Then Lemma 2.26 implies that for some $N \in \mathbb{N}$ we have $a_n \leq \alpha$ for all $n \geq N$. Then, for $n \geq N+1$,

$$\begin{aligned} b_n &= \frac{a_1+a_2+\ldots+a_n}{n} = \frac{a_1+a_2+\ldots+a_N}{n} + \frac{a_{N+1}+a_{N+2}+\ldots+a_n}{n} \\ &\leq \frac{a_1+a_2+\ldots+a_N}{n} + \frac{n-N}{n}\alpha = \frac{a_1}{n} + \frac{a_2}{n} + \ldots + \frac{a_N}{n} + \big(1-\frac{N}{n}\big)\alpha = c_n. \end{aligned}$$

Hence,

$$\limsup_{n\to\infty} b_n \leq \limsup_{n\to\infty} c_n = \lim_{n\to\infty} c_n = \alpha,$$

a contradiction.

Theorem 2.30. *Let (x_n) be a sequence in $\mathbb{R}$. Then for $x_0 \in \mathbb{R}^*$ we have $\lim\limits_{n\to\infty} x_n = x_0$ if and only if $\liminf\limits_{n\to\infty} x_n = \limsup\limits_{n\to\infty} x_n = x_0$.*

Proof. Let us first assume $\lim\limits_{n\to\infty} x_n = x_0 \in \mathbb{R}^*$. Then $E_{(x_n)_n} = \{x_0\}$ and therefore $\liminf\limits_{n\to\infty} x_n = \limsup\limits_{n\to\infty} x_n = x_0$.

Let us now assume $\liminf\limits_{n\to\infty} x_n = \limsup\limits_{n\to\infty} x_n = x_0$ with $x_0 \in \mathbb{R}$. Fix $\epsilon > 0$ and use Lemma 2.26 to obtain $N \in \mathbb{N}$ such

$$x_0 - \epsilon < \liminf_{n\to\infty} x_n - \frac{\epsilon}{2} \leq \quad x_n \quad \leq \limsup_{n\to\infty} x_n + \frac{\epsilon}{2} < x_0 + \epsilon \qquad \text{for all } n \geq N.$$

Since $\epsilon > 0$ was chosen arbitrarily, we have that (x_n) converges and $\lim\limits_{n\to\infty} x_n = x_0$.

Let us assume $\liminf\limits_{n\to\infty} x_n = \limsup\limits_{n\to\infty} x_n = +\infty$. Lemma 2.26 implies that for all $M < \infty$ exists $N \in \mathbb{N}$ with $x_n > M$ for $n \geq N$. This gives $\lim\limits_{n\to\infty} x_n = \infty$.

The case $\liminf\limits_{n\to\infty} x_n = \limsup\limits_{n\to\infty} x_n = -\infty$ can be treated in the same way as the case $\liminf\limits_{n\to\infty} x_n = \limsup\limits_{n\to\infty} x_n = +\infty$. □

2.3. Cauchy sequences and complete metric spaces

Definition 2.31. A sequence (x_n) in a metric space (X,d) is called *Cauchy sequence* if for all $\varepsilon > 0$ there exists $N \in \mathbb{N}$ such that $d(x_n, x_m) < \varepsilon$ for all $n, m \geq N$.

Proposition 2.32. *Every converging sequence in a metric space is a Cauchy sequence.*

Proof. Left to the reader. □

Proposition 2.33. *Every Cauchy sequence in a metric space is bounded.*

Proof. Pick $N \in \mathbb{N}$ so that $d(x_n, x_m) < 1$ for all $n > m$. Then,

$$d(x_N, x_n) \leq M := \max\{1, d(x_N, x_1), d(x_N, x_1), \dots, d(x_N, x_{N-1})\} < \infty, \quad n \in \mathbb{N}. \quad \square$$

Definition 2.34. A metric space (X,d) is called complete if all Cauchy sequences in X converge in X.

Remark 2.35. Not every metric space is complete. For example, consider the punctured real line $\mathbb{R} \setminus \{0\}$ with $d(x,y) = |x-y|$. The sequence $a_n = \frac{1}{n}$ is Cauchy in $\mathbb{R} \setminus \{0\}$ with $d(x,y) = |x-y|$ since for fixed $\epsilon > 0$ we can pick $N > \frac{1}{\epsilon}$ and get

$$d(x_n, x_m) = |x_n - x_m| = \left|\frac{1}{n} - \frac{1}{m}\right| = \left|\frac{m-n}{mn}\right| < \frac{1}{\max\{n,m\}} \leq \frac{1}{N} < \epsilon$$

for all $n, m \geq N$. Nevertheless, (a_n) does not converge in $\mathbb{R} \setminus \{0\}$. Indeed, if it would converge to say $\alpha \in \mathbb{R} \setminus \{0\}$, then it would also converge to α in $\mathbb{R}$. But, clearly, $\lim \frac{1}{n} = 0$ in $\mathbb{R}$, so by uniqueness of limits in metric spaces we conclude $\alpha = 0$, a contradiction.

Proposition 2.36. *Let (X,d) be a metric space and (x_n) be a Cauchy sequence with a converging subsequence, that is there exists (x_{n_k}) with $x_{n_k} \xrightarrow{k\to\infty} x_0$. Then $x_n \xrightarrow{n\to\infty} x_0$.*

Proof. Fix $\epsilon > 0$ and pick $N_\epsilon \in \mathbb{N}$ with $d(x_n, x_m) < \epsilon/2$ for all $n, m \geq N$. Choose $K \in \mathbb{N}$ with $n_K \geq N_\epsilon$ and $d(x_{n_K}, x_0) < \epsilon/2$. Then,

$$d(x_n, x_0) \leq d(x_n, x_{n_K}) + d(x_{n_K}, x_0) \leq \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon, \quad n \geq N. \quad \square$$

Theorem 2.37. $\mathbb{R}$ *and* $\mathbb{C}$ *are complete.*

Proof. Let (x_n) be a Cauchy sequence in $\mathbb{R}$. (x_n) is bounded, hence exists a converging subsequence. By the above, this implies that (x_n) converges. So $\mathbb{R}$ is complete.

If (x_n) is Cauchy in $\mathbb{C}$, then $(\Re x_n)$ and $(\Im x_n)$ are Cauchy in $\mathbb{R}$. By completeness of $\mathbb{R}$ it follows that $(\Re x_n)$ converges to some α and $(\Im x_n)$ to some β, but then (x_n) converges to $\alpha + i\beta$. □

2.4. Real and complex series

Definition 2.38. Let (a_n) be a sequence in $\mathbb{C}$. We call the expression $\sum_{n=1}^{\infty} a_n$ *infinite series* in $\mathbb{C}$. Further, $S_N = a_1 + a_2 + \ldots + a_N = \sum_{n=1}^{N} a_n$ is called the *N–th partial sum* of $\sum_{n=1}^{\infty} a_n$.

If the sequence $(S_N)_{N\in\mathbb{N}}$ of partial sums converges, we set $\sum_{n=1}^{\infty} a_n = \lim_{N\to\infty} S_N$. (Be aware of the abuse of notation: $\sum_{n=1}^{\infty} a_n$ denotes a series as well as the limit of its partial sums in case of its convergence).

Example 2.39. Let $a \in \mathbb{C}$ with $|a| < 1$. Then $S_N = \sum_{n=0}^{N} a^n = \frac{a^{N+1}-1}{a-1}$ and $\sum_{n=0}^{\infty} a^n = \frac{1}{1-a}$.

Definition 2.40. Set $e = \sum_{n=0}^{\infty} \frac{1}{n!} \in \mathbb{R}$.

Remark 2.41. e is well defined:

$$\begin{aligned} S_N = \sum_{n=0}^{N} \frac{1}{n!} &= 1 + 1 + \frac{1}{2} + \frac{1}{2}\cdot\frac{1}{3} + \frac{1}{2}\cdot\frac{1}{3}\cdot\frac{1}{4} + \ldots + \frac{1}{N!} \\ &< 1 + 1 + \frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \ldots + \frac{1}{2^{N-1}} \\ &< 1 + \left(\sum_{n=0}^{\infty}\left(\frac{1}{2}\right)^n\right) = 1 + \frac{1}{1-\frac{1}{2}} = 3 \end{aligned}$$

Hence, (S_N) is bounded. Since (S_N) is also monotone, the sequence of partial sums converges which is the defining property for the series $\sum_{n=0}^{\infty} \frac{1}{n!}$ to converge.

Theorem 2.42. $\lim_{n\to\infty}\left(1+\frac{1}{n}\right)^n = e.$

Proof. We compute

$$\begin{aligned}
t_N = \left(1+\frac{1}{N}\right)^N &= \sum_{n=0}^{N} \binom{N}{n} \frac{1}{N}^n \\
&= \binom{N}{0}\frac{1}{N}^0 + \binom{N}{1}\frac{1}{N} + \binom{N}{2}\frac{1}{N}^2 + \binom{N}{3}\frac{1}{N}^3 + \ldots + \binom{N}{N}\frac{1}{N}^N \\
&= 1 + \frac{N}{N} + \frac{N(N-1)}{2!}\frac{1}{N}^2 + \frac{N(N-1)(N-3)}{3!}\frac{1}{N}^3 + \ldots + \frac{N(N-1)(N-3)\cdot\ldots\cdot 1}{N!}\frac{1}{N}^N \\
&= 1 + 1 + \frac{1}{2!}\frac{N-1}{N} + \frac{1}{3!}\frac{N-1}{N}\frac{N-2}{N} + \ldots + \frac{1}{N!}\frac{N-1}{N}\frac{N-2}{N}\cdot\ldots\cdot\frac{1}{N} \\
&= 1 + 1 + \frac{1}{2!}(1-\frac{1}{N}) + \frac{1}{3!}(1-\frac{1}{N})(1-\frac{2}{N}) + \ldots + \frac{1}{N!}(1-\frac{1}{N})(1-\frac{2}{N})\cdot\ldots\cdot(1-\frac{N-1}{N}) \\
&\leq 1 + 1 + \frac{1}{2!} + \frac{1}{3!} + \ldots + \frac{1}{N!} = \sum_{n=0}^{N}\frac{1}{n!} = S_N.
\end{aligned}$$

We conclude that $\limsup_{N\to\infty} t_N \leq \limsup_{N\to\infty} S_N = \lim_{N\to\infty} S_N = e$.

Clearly, truncation shows that for all $M \leq N$ we have

$$t_N \geq 1 + 1 + \frac{1}{2!}(1-\frac{1}{N}) + \frac{1}{3!}(1-\frac{1}{N})(1-\frac{2}{N}) + \ldots + \frac{1}{M!}(1-\frac{1}{N})(1-\frac{2}{N})\cdot\ldots\cdot(1-\frac{N-M}{N}).$$

which leads to

$$\begin{aligned}
\liminf_{N\to\infty} t_N \geq \liminf_{N\to\infty} 1 + 1 + \frac{1}{2!}(1-\frac{1}{N}) + \frac{1}{3!}(1-\frac{1}{N})(1-\frac{2}{N}) + \\
\ldots + \frac{1}{M!}(1-\frac{1}{N})(1-\frac{2}{N})\cdot\ldots\cdot(1-\frac{N-M}{N}) \\
= 1 + 1 + \frac{1}{2!} + \frac{1}{3!} + \ldots + \frac{1}{M!} = S_M.
\end{aligned}$$

As this holds for all M, we have $\liminf_{N\to\infty} t_N \geq \lim_{M\to\infty} S_M = e$. □

Theorem 2.43. *e is irrational.*

Proof. See Rudin. □

Theorem 2.44. Cauchy Criterion. *The complex series $\sum_{n=1}^{\infty} a_n$ converges in $\mathbb{C}$ if and only if for all $\varepsilon > 0$ there exists $N \in \mathbb{N}$ such that*

$$\left|\sum_{n=k}^{m} a_n\right| < \varepsilon \quad \textit{for all } m \geq k \geq N.$$

Proof. With $S_N = \sum_{n=1}^{N} a_n$, we have $\sum_{n=k}^{m} a_n = S_m - S_{k-1}$, so the Cauchy criterion establishes that the partial sums form a Cauchy sequence. Completeness of $\mathbb{C}$ then guarantees convergence of the partial sums. □

Proposition 2.45. *If* $\sum_{n=1}^{\infty} a_n$ *converges in* $\mathbb{C}$ *then* $a_n \xrightarrow{n\to\infty} 0$.

Proof. This follows from the Cauchy Criterion. (Why?) $\square$

Theorem 2.46. Dominated Convergence Theorem (DCT). *Let* (a_n) *be a sequence in* $\mathbb{C}$.

i. If there is a real valued, non–negative sequence (b_n) *with* $\sum_{n=1}^{\infty} b_n$ *converges and* $|a_n| \leq b_n$ *for all* $n \geq N_0, n \in \mathbb{N}$, *then* $\sum_{n=1}^{\infty} a_n$ *converges.*

ii. If $a_n \geq b_n > 0$ *for* $n \geq N_0, n \in \mathbb{N}$ *and if* $\sum_{n=1}^{\infty} b_n$ *diverges, then* $\sum_{n=1}^{\infty} a_n$ *diverges.*

Proof. Part *i.* follows from Cauchy Criterion since

$$\left|\sum_{n=k}^{m} a_n\right| \leq \sum_{n=k}^{m} |a_n| \leq \sum_{n=k}^{m} b_n = \left|\sum_{n=k}^{m} b_n\right|.$$

Part *ii.* follows from part *i.* as convergence of $\sum_{n=1}^{\infty} a_n$ would imply convergence of $\sum_{n=1}^{\infty} b_n$. $\square$

As a direct consequence, we have the following.

Corollary 2.47. *Let* (a_n) *be a sequence in* $\mathbb{C}$. *If* $\sum_{n=1}^{\infty} |a_n|$ *converges, so does* $\sum_{n=1}^{\infty} a_n$.

Definition 2.48. A complex valued series $\sum_{n=1}^{\infty} a_n$ with $\sum_{n=1}^{\infty} |a_n|$ converges, is called *absolutely convergent*. If $\sum_{n=1}^{\infty} a_n$ converges, but $\sum_{n=1}^{\infty} |a_n|$ does not converge, the we call $\sum_{n=1}^{\infty} a_n$ *conditionally convergent*.

Definition 2.49. Let (c_n) be a sequence of complex numbers and let $\pi : \mathbb{N} \to \mathbb{N}$ be bijective. Then we call the series $\sum_{n=1}^{\infty} c_{\pi(n)}$ a *rearrangement* of the series $\sum_{n=1}^{\infty} c_n$.

Theorem 2.50. *i. If* $\sum_{n=1}^{\infty} c_n$ *converges absolutely, then any rearrangement* $\sum_{n=1}^{\infty} c_{\pi(n)}$ *converges absolutely to the same limit, that is* $\sum_{n=1}^{\infty} c_{\pi(n)} = \sum_{n=1}^{\infty} c_n$ *for any bijective* $\pi : \mathbb{N} \longrightarrow \mathbb{N}$.

ii. If $(c_n)_n$ is real and if $\sum_{n=1}^{\infty} c_n$ converges conditionally, then for any $x \in \mathbb{R}$ exists bijective $\pi_x : \mathbb{N} \to \mathbb{N}$ such that $\sum_{n=1}^{\infty} c_{\pi_x(n)} = x$.

Proof. We shall only prove the first part, the second part is assigned as homework problem.

Fix $\epsilon > 0$. Since $\sum_{n=1}^{\infty} |c_n|$ converges, the Cauchy criterion provides us with some $M \in \mathbb{N}$ such that $\sum_{n=m}^{k} |c_n| < \frac{\epsilon}{3}$ for all $k, m \geq M$. Hence, $\sum_{n=M}^{\infty} |c_n| < \frac{\epsilon}{2}$. Now, let $N = \max\{n \in \mathbb{N} : \pi(n) < M\}$, so

$$\{1, 2, \ldots, M-1\} \subseteq \{\pi(n),\ n = 1, 2, \ldots, N\}$$

and observe that for for $p, q \geq N$, we have

$$\sum_{n=p}^{q} |c_{\pi(n)}| \leq \sum_{n=N}^{\infty} |c_{\pi(n)}| \leq \sum_{n=M}^{\infty} |c_n| < \frac{\epsilon}{2} < \epsilon$$

which, by means of the Cauchy criterion, shows absolute convergence of $\sum_{n=1}^{\infty} c_{\pi(n)}$.

To see that indeed $\sum_{n=1}^{\infty} c_{\pi(n)}$ converges to $c = \sum_{n=1}^{\infty} c_n$, we first observe that, clearly,

$$|c - \sum_{n=1}^{M-1} c_n| = \lim_{R\to\infty} |\sum_{n=1}^{R} c_n - \sum_{n=1}^{M-1} c_n| = \lim_{R\to\infty} |\sum_{n=M}^{R} c_n| \leq \lim_{R\to\infty} \sum_{n=M}^{R} |c_n| = \sum_{n=M}^{\infty} |c_n| < \frac{\epsilon}{2}.$$

We now compute for $p > N$,

$$\begin{aligned} |c - \sum_{n=1}^{p} c_{\pi(n)}| &= |c - \sum_{n=1}^{M-1} c_n| + |\sum_{n=1}^{M-1} c_n - \sum_{n=1}^{p} c_{\pi(n)}| \\ &< \frac{\epsilon}{2} + |\sum_{n\in A} c_n| \leq \frac{\epsilon}{2} + \sum_{n=M}^{\infty} |c_n| < \frac{\epsilon}{2} + \frac{\epsilon}{2} \end{aligned}$$

where

$$A = \{\pi(n),\ n = 1, 2, \ldots, p\} \setminus \{1, 2, \ldots, M-1\}. \quad \square$$

Example 2.51. Take $S = \sum_{n=1}^{\infty}(-1)^n \frac{1}{n} \neq 0$. Consider:

$$\begin{array}{rrcl} & S & = & -1 + \frac{1}{2} - \frac{1}{3} + \frac{1}{4} - \frac{1}{5} + \frac{1}{6} - \frac{1}{7} + \frac{1}{8} - \ldots \leq -\frac{1}{2} \\ + & \frac{1}{2}S & = & \quad -\frac{1}{2} \qquad + \frac{1}{4} \qquad - \frac{1}{6} + \ldots + \frac{1}{8} \\ --- & & - & ------------------------- \\ = & \frac{3}{2}S & = & -1 + 0 - \frac{1}{3} + \frac{1}{2} - \frac{1}{5} + 0 - \frac{1}{7} + \frac{1}{4} + \ldots \\ \text{but} & \frac{3}{2}S & \neq & -1 + \frac{1}{2} - \frac{1}{3} + \frac{1}{4} - \frac{1}{5} + \frac{1}{6} - \frac{1}{7} + \frac{1}{8} - \ldots = S \end{array}$$

since $S \neq 0$. Hence, $\sum_{n=1}^{\infty}(-1)^n \frac{1}{n}$ converges conditionally.

The following criterion is helpful to prove convergence of series which do not converge absolutely.

Theorem 2.52. Leibniz Criterion for Alternating Series. *Let (a_n) be a decreasing sequence of positive real numbers with $a_n \longrightarrow 0$. Then $\sum_{n=0}^{\infty}(-1)^n a_n$ converges.*

Proof. The sequence of partial sums

$$\begin{aligned} S_{2N+1} &= a_0 - a_1 + a_2 - a_3 + \ldots + a_{2N-2} - a_{2N-1} + a_{2N} - a_{2N+1} \\ &= (a_0 - a_1) + (a_2 - a_3) + \ldots + (a_{2N-2} - a_{2N-1}) + (a_{2N} - a_{2N+1}) \\ &= a_0 - (a_1 - a_2) - (a_3 - a_4) - \ldots - (a_{2N-1} - a_{2N}) - a_{2N+1} \end{aligned}$$

is monotonically increasing and bounded above by a_0. Hence S_{2N+1} converges to some $s \in \mathbb{R}$. The algebraic limit theorem implies that $S_{2N} = S_{2N+1} + a_{2N+1}$ converges to $s + 0 = s$. As the summands with even index and the summands of odd index both converge to the same limit, we conclude thate S_N converges to that limit s too. □

Theorem 2.53. Cauchy Condensation Theorem. *Suppose $a_1 \geq a_2 \geq \ldots \geq 0$. Then $\sum_{n=1}^{\infty} a_n$ converges if and only if $\sum_{k=1}^{\infty} 2^k a_{2^k}$ converges.*

Proof. See homework. □

Proposition 2.54. *For $p \in \mathbb{R}$ we have $\sum_{n=1}^{\infty} \frac{1}{n^p}$ converges if and only if $p > 1$.*

Proof. The sequence of summands decreases to 0, so we can use the Cauchy Condensation Theorem to establish convergence of $\sum_{n=1}^{\infty} \frac{1}{n^p}$ by observing that

$$\sum_{k=1}^{\infty} 2^k \frac{1}{(2^k)^p} = \sum_{k=1}^{\infty} 2^{k(1-p)}$$

converges if and only if $p > 1$. □

Theorem 2.55. Root Test. *Given a complex series* $\sum a_n$*, set* $\alpha = \limsup_{n\to\infty} \sqrt[n]{|a_n|}$.

i. If $\alpha < 1$*, then* $\sum a_n$ *converges absolutely.*

ii. If $\alpha > 1$*, then* $\sum a_n$ *diverges.*

iii. If $\alpha = 1$*, then* $\sum a_n$ *might converge or diverge.*

Proof. This follows from the dominated convergence theorem. Indeed, if $\alpha = \limsup_{n\to\infty} \sqrt[n]{|a_n|} < \beta < 1$, then Lemma 2.26 implies that for some $N \in \mathbb{N}$, we have $\sqrt[n]{|a_n|} \leq \beta$ for all $n \geq N$. But then $|a_n| \leq \beta^n$ for $n \geq N$ and convergence of $\sum \beta^n$, $|\beta| < 1$ establishes *i.*

If $\alpha > 1$, then $|a_n| > 1$ for infinitely many $n \in \mathbb{N}$, so a_n does not converge to 0, a necessary condition for $\sum a_n$ to converge.

For *iii.*, note that $\limsup_{n\to\infty} \sqrt[n]{\frac{1}{n}} = \frac{1}{\lim_{n\to\infty} \sqrt[n]{n}} = 1$ but $\sum \frac{1}{n}$ does not converge.

On the other hand, $\limsup_{n\to\infty} \sqrt[n]{\frac{1}{n^2}} = 1$ and $\sum \frac{1}{n^2}$ does converge. □

Theorem 2.56. Ratio Test. *Let* $\sum_{n=1}^{\infty} a_n$ *be a series of complex numbers.*

i. If $\limsup_{n\to\infty} \left|\frac{a_{n+1}}{a_n}\right| < 1$*, then* $\sum_{n=1}^{\infty} a_n$ *converges absolutely.*

ii. If there is $N \in \mathbb{N}$ *with* $\left|\frac{a_{n+1}}{a_n}\right| \geq 1$ *for all* $n > N$*, then* $\sum_{n=1}^{\infty} a_n$ *diverges.*

Proof. If $\limsup_{n\to\infty} \left|\frac{a_{n+1}}{a_n}\right| < \beta < 1$, then for some $N \in \mathbb{N}$, we have $|\frac{a_{n+1}}{a_n}| \leq \beta$, and, hence, $|a_{n+1}| \leq \beta |a_n|$ for $n \geq N$. This implies that $|a_n| \leq |a_N| \beta^{N-n}$ for $n \geq N$ and we can apply again the dominated convergence theorem.

For *ii.* observe that that the condition implies $|a_n| \geq |a_N| > 0$ for all $n \geq N$, so $\lim a_n = 0$ is violated once more. □

Examples 2.57. i. Let $a_n = \frac{1}{n}$. Then $\limsup_{n\to\infty} \frac{a_{n+1}}{a_n} = \limsup_{n\to\infty} \frac{n}{n+1} = \lim_{n\to\infty} \frac{n}{n+1} = 1$, but the series $\sum_{n=1}^{\infty} \frac{1}{n}$ does not converge.

ii. Let $b_n = \frac{1}{n^2}$. Then $\limsup_{n\to\infty} \frac{b_{n+1}}{b_n} = \limsup_{n\to\infty} \frac{n^2}{(n+1)^2} = \lim_{n\to\infty} \frac{n^2}{(n+1)^2} = 1$ and $\sum_{n=1}^{\infty} \frac{1}{n^2}$ does converge.

Definition 2.58. The series $\sum_{n=0}^{\infty} c_n z^n$ is called a *power series* with coefficients $c_n \in \mathbb{C}$, $n \in \mathbb{N}$. For $\alpha = \limsup_{n\to\infty} \sqrt[n]{|c_n|} \in [0,\infty] \subset \mathbb{R}^*$ we call

$$R_{(c_n)} = \begin{cases} \frac{1}{\alpha} & \text{if } \alpha \in (0,\infty); \\ \infty & \text{if } \alpha = 0; \\ 0 & \text{if } \alpha = \infty \end{cases}$$

the *radius of convergence* of the power series $\sum_{n=0}^{\infty} c_n z^n$.

Theorem 2.59. *The series $\sum_{n=0}^{\infty} c_n z^n$ converges if $|z| < R_{(c_n)}$ and diverges if $|z| > R_{(c_n)}$, and $\sum_{n=0}^{\infty} c_n z^n$ may or may not converge for $z \in \mathbb{C}$ with $|z| = R_{(c_n)}$.*

Proof. For fixed $z \in \mathbb{C}$ we apply the root criterion to $\sum a_n = \sum c_n z^n$. Clearly,

$$\beta = \limsup_{n\to\infty} \sqrt[n]{|c_n z^n|} = \limsup_{n\to\infty} |z| \sqrt[n]{|c_n|} = |z| \limsup_{n\to\infty} \sqrt[n]{|c_n|},$$

so $\beta < 1$ if and only if $|z| < R_{(c_n)}$. The result now follows from the root test. □

Remark 2.60. It is easy to see that a series of the form $\sum_{n=0}^{\infty} c_n (z-z_0)^n$ converges if $|z-z_0| < R_{(c_n)}$ and diverges if $|z-z_0| > R_{(c_n)}$, a fact which is relevant when discussing Taylor series of a function f at a point $z_0 \in \mathbb{R}$. (See Section 4.)

We conclude this section with a brief discussion of the exponential function $\exp(z) = \sum_{n=0}^{\infty} \frac{z^n}{n!}$, $z \in \mathbb{C}$. To derive the functional equation $\exp(z+w) = \exp(z)\exp(w)$ we use theorem discussing the product of two series. This theorem is based on diagonal summation of the terms in

$$\begin{array}{rcccccccccc} (a_0+a_1+a_2+\dots)\cdot(b_0+b_1+b_2+\dots) = & & a_0b_0 & + & a_0b_1 & + & a_0b_2 & + & a_0b_3 & + & \dots, \\ & + & a_1b_0 & + & a_1b_1 & + & a_1b_2 & + & a_1b_3 & + & \dots \\ & + & a_2b_0 & + & a_2b_1 & + & a_2b_2 & + & a_2b_3 & + & \dots \\ & + & a_3b_0 & + & a_3b_1 & + & a_3b_2 & + & a_3b_3 & + & \dots \\ & & \vdots & & \vdots & & \vdots & & \vdots & & \end{array}$$

$$= c_0 + c_1 + c_2 + c_3 + \dots$$

that is, with $c_n = \sum_{k=0}^n a_k b_{n-k}$, we have that c_0 is the sum of red terms, c_1 is the sum of blue terms, c_2 is the sum of green terms, c_3 is the sum of magenta terms, and so on. For the partial sums, we can resort the (finitely many) summands to obtain

$$\sum_{n=0}^{N} c_n = \sum_{n=0}^{N} \sum_{k=0}^{n} a_k b_{n-k} = \sum_{k=0}^{N} a_k \sum_{n=0}^{N-k} b_n = \sum_{k=0}^{N} d_k.$$

For $N = 3$, we first sum red, then blue, then green, then magenta in the equation below:

$$\sum_{n=0}^{3} c_n = d_0 + d_1 + d_2 + d_3 = \begin{array}{ccccccc} & a_0 b_0 & + & a_0 b_1 & + & a_0 b_2 & + & a_3 b_3 \\ + & a_1 b_0 & + & a_1 b_1 & + & a_1 b_2 & & \\ + & a_2 b_0 & + & a_2 b_1 & & & & \\ + & a_3 b_0 & & & & & & \end{array},$$

Theorem 2.61. Product of series. *Let (a_n) and (b_n) be complex sequences with $\sum_{n=0}^\infty a_n = A$ converges absolutely, and $\sum_{n=0}^\infty b_n = B$. For $c_n = \sum_{k=0}^n a_k b_{n-k}$, $n \in \mathbb{N}_0$ we have $\sum_{n=0}^\infty c_n = A \cdot B$.*

Proof. Fix $\epsilon > 0$. Set $C = \sum_{k=0}^\infty |a_k| < \infty$ and $D = \sup\{|B - \sum_{k=0}^N b_k|,\ N \in \mathbb{N}\} < \infty$. Pick $N_\epsilon \in \mathbb{N}$ so that for all $N \geq N_\epsilon$ we have

$$\Big|A - \sum_{k=0}^{N} a_k\Big| < \frac{\epsilon}{3B}, \quad \Big|B - \sum_{\ell=0}^{\lfloor \frac{N}{2} \rfloor} b_\ell\Big| < \frac{\epsilon}{3C}, \quad \Big|C - \sum_{n=0}^{\lfloor \frac{N}{2} \rfloor} |a_n|\Big| < \frac{\epsilon}{3D},$$

where $\lfloor x \rfloor$ denotes the largest integer smaller than x.

For $N \geq N_\epsilon$ we compute

$$\begin{aligned}
\Big|AB - \sum_{n=0}^{N} c_n\Big| &= \Big|AB - \sum_{n=0}^{N}\sum_{k=0}^{N} a_k b_{n-k}\Big| \\
&= \Big|AB - \sum_{k=0}^{N} a_k \sum_{n=0}^{N-k} b_n\Big| \\
&= \Big|AB - B\sum_{k=0}^{N} a_k + B\sum_{k=0}^{N} a_k - \sum_{k=0}^{N} a_k \sum_{n=0}^{N-k} b_n\Big| \\
&\leq |B|\Big|A - \sum_{k=0}^{N} a_k\Big| + \Big|\sum_{k=0}^{N} a_k \Big(B - \sum_{n=0}^{N-k} b_n\Big)\Big| \\
&< |B|\frac{\epsilon}{3B} + \sum_{k=0}^{N} |a_k|\Big|B - \sum_{n=0}^{N-k} b_n\Big| \\
&\leq \frac{\epsilon}{3} + \sum_{k=0}^{\lfloor N/2 \rfloor} |a_k|\Big|B - \sum_{n=0}^{N-k} b_n\Big| + \sum_{k=\lfloor N/2 \rfloor + 1}^{N} |a_k|\Big|B - \sum_{n=0}^{N-k} b_n\Big| \\
&< \frac{\epsilon}{3} + C\frac{\epsilon}{3C} + \frac{\epsilon}{3D}D = \epsilon.
\end{aligned}$$

□

Remark 2.62. Note that the hypotheses in Theorem 2.61 do not imply that $\sum_{n=0}^{\infty} c_n$ converges absolutely. In fact, $(a_n) = 1, 0, 0, \ldots$ and $(b_n) = ((-1)^n \frac{1}{n})$ satisfy the hypothesis, but $c_n = \sum_{k=0}^{n} a_k b_{n-k} = b_n$ implies that (c_n) does not converge absolutely.

Corollary 2.63. *For $z, w \in \mathbb{C}$ we have $\exp(z + w) = \exp(z)\exp(w)$.*

Proof. Let $(a_n) = (\frac{z^n}{n!})$ and $(b_n) = (\frac{w^n}{n!})$, so $\sum_{n=0}^{\infty} a_n$ and $\sum_{n=0}^{\infty} b_n$ converge absolutely. Since

$$c_n = \sum_{k=0}^{n} a_k b_{n-k} = \sum_{k=0}^{n} \frac{z^k}{k!} \frac{w^{n-k}}{(n-k)!} = \frac{1}{n!} \sum_{k=0}^{n} \frac{n!}{(n-k)!k!} z^k w^{n-k} = \frac{1}{n!} \sum_{k=0}^{n} \binom{n}{k} z^k w^{n-k} = \frac{(z+w)^n}{n!},$$

we can apply Theorem 2.61 to obtain

$$\exp(z)\exp(w) = \Big(\sum_{n=0}^{\infty} a_n\Big)\Big(\sum_{n=0}^{\infty} b_n\Big) = \sum_{n=0}^{\infty} c_n = \exp(z + w). \qquad \square$$

Corollary 2.64. *For $x \in \mathbb{Q}$ we have $\exp(x) = e^x$.*

Proof. This is clear for $x = 1$. The equality then follows for $x = 2 = 1 + 1$ since

$$e^2 = e \cdot e = \exp(1) \cdot \exp(1) = \exp(1 + 1) = \exp(2).$$

Similarly, we obtain $\exp(x) = e^x$ for $x \in \mathbb{N}$. Since, clearly, $\exp(-1)$ is the (unique) muliplicative inverse of $e = \exp(1)$, we have $e^{-1} = \exp(-1)$ and $\exp(x) = e^x$ follows for $x \in \mathbb{Z}$.

For $x = \frac{1}{n}$, $n \in \mathbb{N}$, we observe that $\exp(\frac{1}{n}) > 0$ and $\exp(\frac{1}{n})^n = \exp(\frac{1}{n} + \ldots + \frac{1}{n}) = \exp(1) = e$, so $\exp(\frac{1}{n})$ is the (unique) n-th non-negative root of e. But then also for $x = \frac{m}{n}$, $m, n \in \mathbb{N}$, we obtain that

$$\Big(\exp\big(\frac{m}{n}\big)\Big)^n = \exp\big(\frac{m}{n} + \ldots + \frac{m}{n}\big) = \exp(m) = e^m$$

and the result follows. Clearly, using $e^{-1} = \exp(-1)$ allows us to extend this argument to negative $x \in \mathbb{Q}$. $\square$

We shall show later that $\exp(x) = e^x$ holds for all $x \in \mathbb{R}$. Motivated by this, we shall then write e^z for $\exp(z)$ for any $z \in \mathbb{C}$.

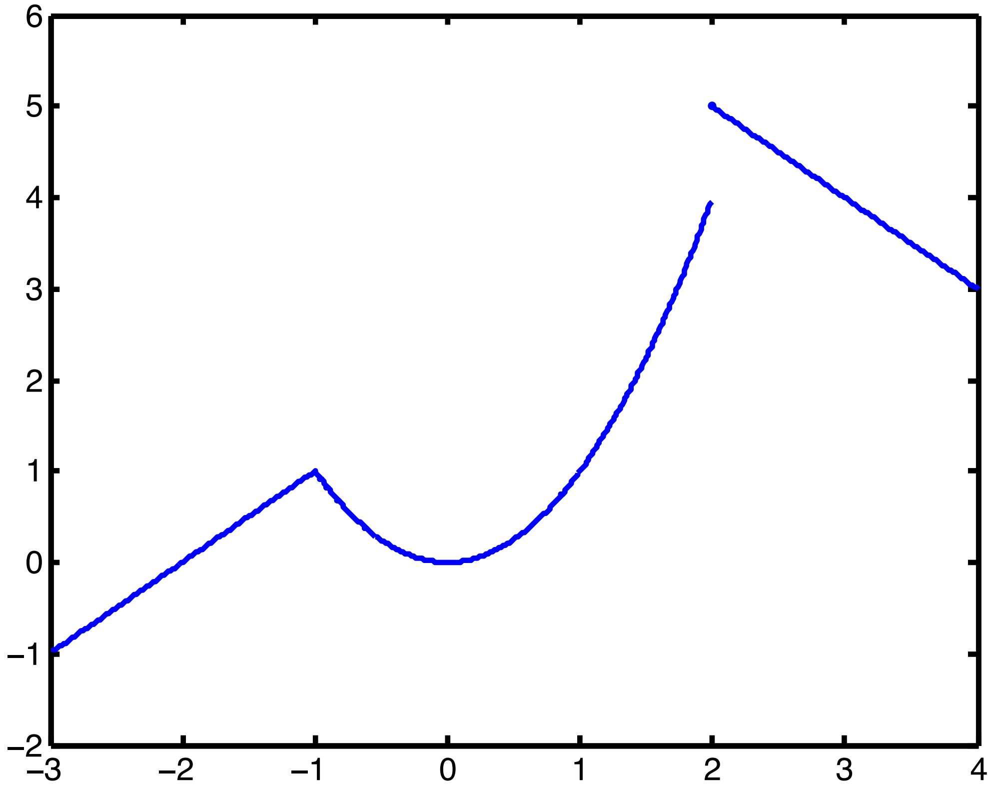

Figure 1. A graph of the function f from Example 3.2.

3. TOPOLOGY AND CONTINUITY

3.1. Continuous functions

Definition 3.1. A function $f : \mathbb{R} \to \mathbb{R}$ is *continuous* at $x_0 \in \mathbb{R}$ if for all $\varepsilon > 0$ exists $\delta > 0$ s.t. $|f(x) - f(x_0)| < \varepsilon$ if $|x - x_0| < \delta$.

Example 3.2. The function

$$f : \mathbb{R} \longrightarrow \mathbb{R},\ x \mapsto \begin{cases} x+2, & \text{if } x \leq -1; \\ x^2, & \text{if } -1 < x < 2; \\ -x+7, & \text{if } 2 \leq x. \end{cases}$$

is continuous at any point x_0 in $\mathbb{R} \setminus \{2\}$ and discontinuous at $x_0 = 2$. See Figure 3.1 for a graph of this function.

Remark 3.3. Continuous functions have some remarkable properties. Most prominently, the intermediate value theorem and the maximum value theorem for real valued functions defined on $\mathbb{R}$ state that given a continuous function $f : [a, b] \longrightarrow \mathbb{R}$ then exists $c, d \in \mathbb{R}$, such that $f([a, b]) = [c, d]$. (See Corollary 3.60.)

This theorem can be generalized to metric spaces: If X is a *compact* and *connected* metric space, and $f : X \longrightarrow Y$ is *continuous*, then $f(X)$ is *compact* and *connected*. In case of $Y = \mathbb{R}$

we get immediately $f(X) = [c, d]$ for some $c, d \in \mathbb{R}$ since closed intervals are the only subsets of $\mathbb{R}$ which are both, *compact* and *connected*. Well, we need some new vocabulary.

Definition 3.4. Let (X, d_X) be a metric space, $x_0 \in X$, and $r \in \mathbb{R}^+$. The *open* [respectively *closed*] *ball* in X of center x_0 and radius r is the set

$$\begin{aligned} B_r(x_0) &= \{x \in X : d_X(x, x_0) < r\} \subseteq X \\ \text{[resp. } B_r^{closed} &= \{x \in X : d_X(x, x_0) \leq r\}] \end{aligned}$$

We shall also refer to the open ball $B_r(x_0)$ as *r-neighborhood* of x_0.

Definition 3.5. Let (X, d_X) and (Y, d_Y) be metric spaces. A function $f : X \to Y$ is *continuous* at $x_0 \in X$, if for all $\varepsilon \in \mathbb{R} > 0$ exists $\delta > 0$ s.t. $d_Y(f(x), f(x_0)) < \varepsilon$ if $d_X(x, x_0) < \delta$, that is, $f(B_\delta(x_0)) \subseteq B_\varepsilon(f(x_0))$.

Examples 3.6.

i. The most important metric on $\mathbb{R}^n$ respectively $\mathbb{C}^n$ is given by Euclidean distance, namely, $d_2(x, y) = \sqrt{\sum_{i=1}^n |x_i - y_i|^2}$.

ii. Almost equally important is the 1 metric on $\mathbb{R}^n$ respectively $\mathbb{C}^n$ is given by $d_1(x, y) = \sum_{i=1}^n |x_i - y_i|$.

iii. Same holds for the ∞ metric on $\mathbb{R}^n$ respectively $\mathbb{C}^n$ is given by $d_\infty(x, y) = \max_{i=1,\ldots,n} |x_i - y_i|$.

iv. A bit obscure example is the discrete metric on a set X (e.g., $\mathbb{R}^n$ or $\mathbb{C}^n$) given bey $d_0(x, y) = 1$ if $x \neq y$ and $d_0(x, x) = 0$.

See Figure 3.1 illustrating respective balls of radius 1.

To compare balls with respect to different metrics, the following is useful.

Theorem 3.7. Cauchy–Schwarz Inequality.
Let $a_1, \ldots, a_n, b_1, \ldots, b_n \in \mathbb{C}$. Then

$$\left| \sum_{i=1}^n a_i \overline{b_i} \right|^2 \leq \sum_{i=1}^n |a_i|^2 \sum_{i=1}^n |b_i|^2.$$

In Hilbert space terminology, this reads that the modulus of the inner product of two vectors in a Hilbert spaces does not exceed the product of their norms. The proof given below does not generalize to this setting but establishes the fact for the Hilbert space $\mathbb{C}^n$ with Euclidean inner product. (Obviously, you are not expected to know yet what a Hilbert space is.)

Proof. If $a_1 = a_2 = \ldots = a_n = 0$, then the result follows trivially. Else, $\sum_{i=1}^n |a_i|^2 > 0$ and for

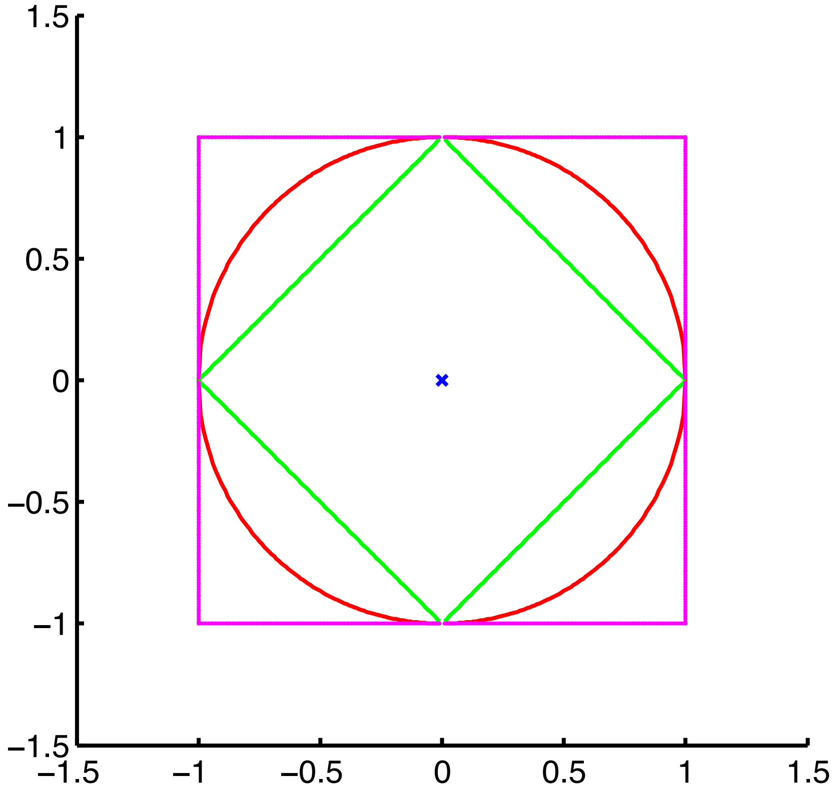

Figure 2. The boundaries of balls of radius one with respect to the d_1 metric (green), the d_2 metric (red), and the d_∞ metric (magenta). The ball of radius one with respect to the discrete metric d_0 is $\{0\}$ and shown in blue.

$x \in \mathbb{R}$, we compute

$$0 \le \sum_{i=1}^{n}(|a_i|x+|b_i|)^2 = \sum_{i=1}^{n}|a_i|^2x^2+2|a_i||b_i|x+|b_i|^2$$
$$= \Big(\sum_{i=1}^{n}|a_i|^2\Big)x^2+2\Big(\sum_{i=1}^{n}|a_i||b_i|\Big)x+\Big(\sum_{i=1}^{n}|b_i|^2\Big)$$
$$= \Big(\sum_{i=1}^{n}|a_i|^2\Big)\Big(x+\frac{\sum_{i=1}^{n}|a_i||b_i|}{\sum_{i=1}^{n}|a_i|^2}\Big)^2-\frac{\Big(\sum_{i=1}^{n}|a_i||b_i|\Big)^2}{\Big(\sum_{i=1}^{n}|a_i|^2\Big)^2}+\Big(\sum_{i=1}^{n}|b_i|^2\Big).$$

Using the triangular inequality for the first inequality below and setting $x = -\frac{\sum_{i=1}^{n}|a_i||b_i|}{\sum_{i=1}^{n}|a_i|^2}$ to obtain the second inequality, we have

$$\Big|\sum_{i=1}^{n}a_i\overline{b_i}\Big|^2 \le \Big(\sum_{i=1}^{n}|a_i||b_i|\Big)^2 \le \Big(\sum_{i=1}^{n}|a_i|^2\Big)\Big(\sum_{i=1}^{n}|b_i|^2\Big). \qquad \square$$

Theorem 3.8. *If $f:(\mathbb{R}^n,d_{i_0})\to(\mathbb{R}^m,d_{j_0})$ is continuous at $x_0\in\mathbb{R}^n$ for some $i_0,j_0\in\{1,2,\infty\}$, then $f:(\mathbb{R}^n,d_i)\to(\mathbb{R}^m,d_j)$ for any $i,j\in\{1,2,\infty\}$.*

Proof. We shall first prove that

$$B_\epsilon^\infty(x) \supseteq B_\epsilon^2(x) \supseteq B_\epsilon^1(x) \supseteq B_{\epsilon/\sqrt{m}}^2(x) \supseteq B_{\epsilon/m}^\infty(x). \tag{1}$$

See Figure 3.1 for an illustration of this inclusion in the case of $n=m=2$.

Using Cauchy Schwarz, we observe that for any $x,y\in\mathbb{R}^m$ we have

$$0 < \max_{i=1,\dots,m}|x_i-y_i| \le \sqrt{\sum_{i=1}^{m}|x_i-y_i|^2} \le \sum_{i=1}^{m}|x_i-y_i| = \sum_{i=1}^{m}|x_i-y_i|\cdot 1$$
$$\le \sqrt{\Big(\sum_{i=1}^{m}|x_i-y_i|^2\Big)\Big(\sum_{i=1}^{m}1^2\Big)} = \sqrt{m}\sqrt{\sum_{i=1}^{m}|x_i-y_i|^2} \le \sqrt{m}\sqrt{\sum_{i=1}^{m}\big(\max_{i=1,\dots,m}|x_i-y_i|\big)^2}$$
$$= m\max_{i=1,\dots,m}|x_i-y_i|,$$

that is, $d_\infty(x,y)\le d_2(x,y)\le d_1(x,y)\le\sqrt{m}\,d_2(x,y)\le m\,d_\infty(x,y)$. Now, $d_2(x,y)<\epsilon$ implies $d_\infty(x,y)<\epsilon$ and, hence, $B_\epsilon^2(x)\subseteq B_\epsilon^\infty(x)$. Similarly, $d_2(x,y)<\epsilon$ implies $d_1(x,y)<\sqrt{m}\,\epsilon$, and, hence, $B_\epsilon^2(x)\subseteq B_{\sqrt{m}\epsilon}^\infty(x)$. These arguments provide (1).

Fix $i,j\in\{0,2,\infty\}$ and $\epsilon>0$. As $f:(\mathbb{R}^n,d_{i_0})\to(\mathbb{R}^m,d_{j_0})$ is continuous at $x_0\in\mathbb{R}^n$, there exists $r>0$ such that $f(B_r^{i_0}(x_0))\subseteq B_{\epsilon/m}^{j_0}(f(x_0))$ and set $\delta=\frac{r}{n}>0$. Then

$$f(B_\delta^i(x_0))\subseteq f(B_{n\delta}^{i_0}(x_0)) = f(B_r^{i_0}(x_0))\subseteq B_{\epsilon/m}^{j_0}(f(x_0))\subseteq B_\epsilon^j(f(x_0)). \qquad \square$$

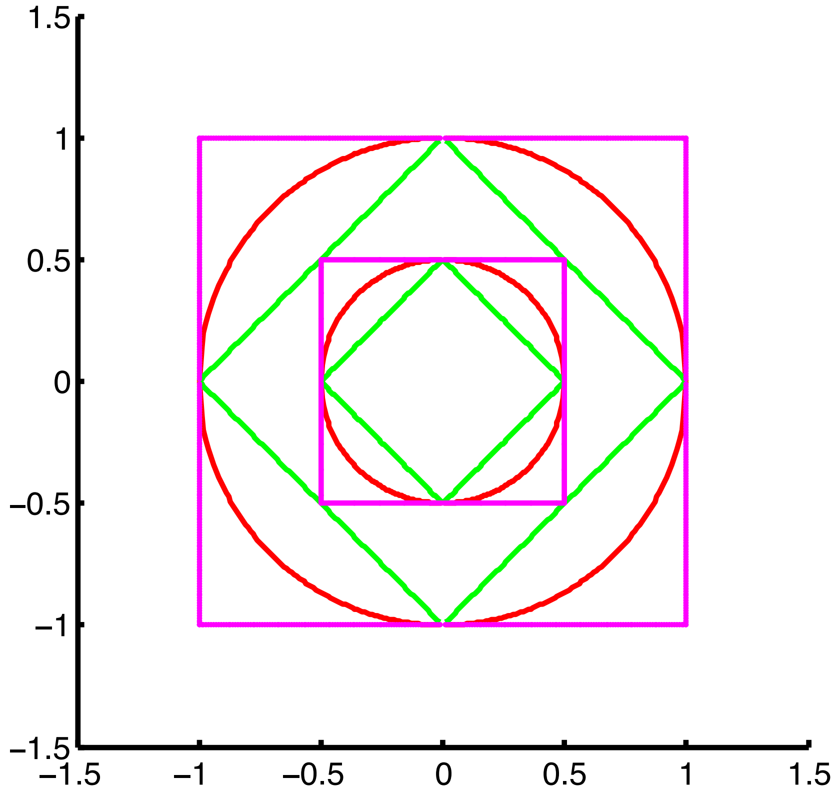

Figure 3. Illustration of inclusion of balls of different metrics which proves Theorem 3.8 and Theorem 3.12.

Remark 3.9. Obviously, continuity does depend on the metric of choice. Nevertheless, different metrics (not all) lead to the same concept of continuity. We shall now extract the essence of continuous functions between metric spaces which will lead to a whole new class of spaces, namely topological spaces.

Definition 3.10. Let (X, d) be a metric space. $U \subseteq X$ is called *(metric-) open* if for each $x_0 \in U$ exists $\varepsilon > 0$ s.t. $B_\varepsilon(x_0) \subseteq U$. A set $A \subseteq X$ is called *(metric-) closed* if its complement A^c is (metric-) open.

We should check consistency of our vocabulary. We did define *open balls* before defining *open sets* after all, so open balls better be open sets.

Theorem 3.11. *Let (X, d) be a metric space, then open balls are (metric-) open.*

Proof. For the open ball $B_\delta(y_0)$ choose $x_0 \in B_\delta(y_0)$. Set $\epsilon = \delta - d(x_0, y_0) > 0$ and observe for $x \in B_\epsilon(x_0)$ we have $d(x, y_0) \leq d(x, x_0) + d(x_0, y_0) < \delta - d(x_0, y_0) + d(x_0, y_0) = \delta$ due the triangular inequality. Hence, $B_\epsilon(x_0) \subseteq B_\delta(y_0)$. Such ϵ exists for each x_0, hence, $B_\delta(y_0)$ is open. □

Proposition 3.12. *U is open in $(\mathbb{R}^n, d_\infty)$ if and only if U is open in $(\mathbb{R}^n, d_1)$ if and only if U is open in $(\mathbb{R}^n, d_2)$.*

Proof. This follows again from $B^i_{\epsilon/\sqrt{n}}(x_0) \subseteq B^i_\epsilon(x_0)$ for all $i, j \in \{1, 2, \infty\}$. □

Theorem 3.13. *A function $f : (X, d_X) \to (Y, d_Y)$ is continuous on X (that is, continuous at all $x_0 \in X$) if and only if $f^{-1}(U)$ is open in (X, d_X) for all U open in (Y, d_Y).*

Proof. "⇒" Let U be open in (Y, d_Y). To show that the pre image $f^{-1}(U)$ is open in (X, d_X), choose $x_0 \in f^{-1}(U)$. Then $f(x_0) \in U$, and, as U open, exists $\epsilon > 0$ such that $B_\epsilon(f(x_0)) \subset U$. Now, use continuity of f at x_0 to choose $\delta > 0$ with $f(B_\delta(x_0)) \subseteq B_\epsilon(f(x_0))$.

"⇐" The obtain the reverse inequality, just note that for $\epsilon > 0$ and $x_0 \in X$, we have $f^{-1}(B_\epsilon(f(x_0))$ is an open set containing x_0. Hence, there exists $\delta > 0$ such that $B_\delta(x_0) \subseteq f^{-1}(B_\epsilon(f(x_0)))$, that is, $f(B_\delta(x_0)) \subseteq B_\epsilon(f(x_0)) \subseteq U$, but this is again just $B_\delta(x_0) \subseteq f^{-1}(U)$. □

Theorem 3.14. *Let $\mathfrak{U}$ be a family of (metric-) open sets in (X, d). Then*

i. $\emptyset, X$ *are open,*

ii. $\bigcup_{U \in \mathfrak{U}} U$ *is open in* (X, d)*, and*

iii. $U \cap V$ *is open in* (X, d) *for any* $U, V \in \mathfrak{U}$*.*

Proof. i. There exists no $x_0 \in \emptyset$, hence, there is nothing to check to show that $\emptyset$ is open. As any $B_\epsilon(x_0) \subseteq X$, in particular, for $x_0 \in X$, we have $B_1(x_0) \subseteq X$.

ii. Let $x_0 \in \bigcup_{U \in \mathfrak{U}} U$. Then $x_0 \in U_0$ for some $U_0 \in \mathfrak{U}$. As U_0 is open, there exists $\delta > 0$ such that $B_\delta(x_0) \subseteq U_0 \bigcup_{U \in \mathfrak{U}} U$.

iii. Let $x_0 \in U \cap V$. As U and V are open, there exists $\delta_U > 0$ and $\delta_V > 0$ with $B_{\delta_U}(x_0) \subseteq U$ and $B_{\delta_V}(x_0) \subseteq V$. But then, setting $\delta = \min\{\delta_U, \delta_V\}$ implies $B_\delta(x_0) \subseteq U \cap V$.

□

Let us now provide a very important and useful result for the understanding of open sets in subspaces of metric spaces. This result will be used extensively when discussing connected subsets of metric spaces.

Theorem 3.15. INHERITANCE PRINCIPLE. *Let (X, d_X) be a metric space and $A \subseteq X$. Then (A, d_A) becomes a metric space when setting $d_A = d_X|_{A \times A}$, that is, $d_A(a, b) = d_X(a, b)$ for $a, b \in A$. Further, the following hold:*

i. $B \subseteq A$ is open in (A, d_A) if and only there exists $\widetilde{B}$ open in (X, d_X) such that $B = A \cap \widetilde{B}$.

ii. $B \subseteq A$ is closed in (A, d_A) if and only there exists $\widetilde{B}$ closed in (X, d_X) such that $B = A \cap \widetilde{B}$.

iii. $B \subseteq A$ is clopen (closed and open) in (A, d_A) **if** *there exists $\widetilde{B}$ clopen in (X, d_X) such that $B = A \cap \widetilde{B}$.*

3.2. Topological spaces

Theorem 3.14 provides all properties of metric spaces needed to extend the concept of continuous maps on metric spaces to maps between more general spaces, namely, topological spaces.

Definition 3.16. Let X be any set and let $\mathcal{T}$ be a collection of subsets of X which satisfies

i. $X, \emptyset \in \mathcal{T}$,

ii. $\bigcup_{i \in I} U_i \in \mathcal{T}$ whenever $U_i \in \mathcal{T}$, $i \in I$, and

iii. $U \cap V \in \mathcal{T}$ if $U, V \in \mathcal{T}$.

Then we call $\mathcal{T}$ a *topology* on the *topological space* X, the members U of $\mathcal{T}$ are called *(topology-) open*. A set $A \subseteq X$ is called *closed* if $A^c = X \setminus A \in \mathcal{T}$, that is, if A is the complement of an open set.

Example 3.17. i. Any set X becomes a topological space when choosing the trivial topology $\mathcal{T} = \{\emptyset, X\}$. This topology is also called *indiscrete* topology.

ii. Any set X becomes a topological space when choosing as topology the powerset of X, that is, $\mathcal{T} = \mathcal{P}(X)$. This topology is also called *discrete* topology .

iii. The metric open sets in a metric space (X, d) form a topology on X (see Theorem 3.14). This topology is *induced* by the metric d and we denote it by $\mathcal{T}_d$.

iv. Note that for any set X and discrete metric d_0 on X, (ii) and (iii) lead to the same topology, that is, $\mathcal{T}_{d_0} = \mathcal{P}(X)$. This is easy to see since in (X, d_0) (d_0 denotes the discrete metric) we have that $B_1(x) = \{x\}$ for any $x \in X$. Hence, all singletons (sets with only one element) are open and any $S \in \mathcal{P}(X)$ is open since it can be written as union of open sets, for example, $S = \bigcup_{x \in S} \{x\}$.

v. It is not difficult to construct (maybe meaningless) topologies on any set. For example, the set $\mathcal{T} = \{\emptyset, [0, 1], \mathbb{R}\}$ defines a topology on $\mathbb{R}$. By definition, $[0, 1]$ is open, $(0, 1)$ and $[1, 2]$ are not since they are not listed in $\mathcal{T}$.

Remark 3.18. Recall that, using properties of (metric-) open sets in a metric space (X, d), we introduced a new family of spaces which is custom made to study continuous maps.

Many properties of metric induced topologies now serve as defining properties when dealing with general topological spaces. For example, given a topological space $(X, \mathcal{T})$ and a subset A in X, we can equip A with the so called relative topology $\mathcal{T}_A = \{A \cap U : \; U \in \mathcal{T}\}$ to obtain a topological space $(A, \mathcal{T}_A)$. (Compare to the inheritance principle, Theorem 3.15.)

By virtue of Theorem 3.13 we can extend the concept of continuous maps to general topological spaces:

Definition 3.19. Let $(X, \mathcal{T})$, $(Y, \mathcal{F})$ be topological spaces. A function $f : X \to Y$ is called *continuous* if $f^{-1}(V) \in \mathcal{T}$ for all $V \in \mathcal{F}$.

Theorem 3.20. *Let $(X, \mathcal{T})$, $(Y, \mathcal{F})$, and $(Z, \mathcal{S})$ be topological spaces and $f : X \longrightarrow Y$ and $g : Y \longrightarrow Z$ be continuous. Then $g \circ f : X \longrightarrow Z$, $x \mapsto g \circ f(x) = g(f(x))$ is continuous.*

Proof. Let $U \in \mathcal{S}$ and observe that

$$
\begin{aligned}
(g \circ f)^{-1}(U) &= \{x \in X : \; g(f(x)) \in U\} = \{x \in X : \; f(x) \in g^{-1}(U)\} \\
&= \{x \in X : \; x \in f^{-1}(g^{-1}(U))\} = f^{-1}(g^{-1}(U)).
\end{aligned}
$$

As g is continuous, we have $g^{-1}(U) \in \mathcal{F}$ and, f continuous implies then

$(g \circ f)^{-1}(U) = f^{-1}(g^{-1}(U)) \in \mathcal{T}$. As we chose arbitrary $U \in \mathcal{S}$, we conclude that $g \circ f$ continuous. □

In the mathematical discipline topology, one studies whether two topological spaces X and Y have “identical topologies”, that is, whether there exists a continuous, bijective map which maps open sets to open sets (that is, f^{-1} (which exists and is defined on all of Y since f is bijective) is continuous as well).

Definition 3.21. If $f : X \longrightarrow Y$ is bijective and continuous, and if the function $f^{-1} : Y \longrightarrow X$ is continuous as well then we call f a *homeomorphism*.

Definition 3.22. The topological spaces $(X, \mathcal{T})$ and $(Y, \mathcal{F})$ are called *homeomorph* if there exists a homeomorphism $f : X \longrightarrow Y$.

Definition 3.23. A sequence (x_n) in the topological space $(X, \mathcal{T})$ *converges* to x_0 in $(X, \mathcal{T})$, if for all $U \in \mathcal{T}$ with $x_0 \in U$ there exists $N_0 \in \mathbb{N}$ s.t. $x_n \in U$ if $n \geq N_0$.

Our back is covered:

Theorem 3.24. *A sequence (x_n) converges to x_0 in the metric space (X, d) if and only if x_n converges to x_0 in the metric induced topological space $(X, \mathcal{T}_d)$.*

Proof. "$\Rightarrow$" Let $x_n \to x_0$ in (X, d), that is, X considered as metric space. Fix $U \in \mathcal{T}$ with $x_0 \in U$. As U open, there exists $\epsilon > 0$ with $B_\epsilon(x_0) \subseteq U$. Now, $x_n \to x_0$ in (X, d) implies that for some $N \in \mathbb{N}$ we have $x_n \in B_\epsilon(x_0) \subseteq U$ for all $n \geq N$, and $x_n \to x_0$ in $(X, \mathcal{T}_d)$ is shown.

"$\Leftarrow$" Now, let $x_n \to x_0$ in $(X, \mathcal{T}_d)$, that is, X considered as topological space. Fix $\epsilon > 0$. As $B_\epsilon(x_0) \in \mathcal{T}_d$ and $x_n \to x_0$ in $(X, \mathcal{T}_d)$, there exists $N \in \mathbb{N}$ such that $x_n \in B_\epsilon(x_0)$ for all $n \geq N$, and $x_n \to x_0$ in the metric space (X, d) is shown. □

Example 3.25. The function

$$f : [0, 2\pi) \longrightarrow \mathcal{R}_f = \{z \in \mathbb{C} : |z| = 1\} \subset \mathbb{C}, \quad x \mapsto \cos(x) + i \sin(x),$$

is continuous, 1-1, surjective, and continuous, but f^{-1} is not continuous at $1 = \cos(0) + i \sin(0)$. Hence, f is not a homeomorphism. (We shall define cos and sin in Section 4.3. At this point of time, we only assume High–School knowledge of trigonometric functions.)

To see this, observe that $\lim_{n\to\infty} \cos(2\pi - \frac{1}{n}) + i \sin(2\pi - \frac{1}{n}) = 1$, but its image under f^{-1} is the sequence $\big(f^{-1}(\cos(2\pi - \frac{1}{n}) + i \sin(2\pi - \frac{1}{n}))\big)_n = \big(2\pi - \frac{1}{n}\big)_n$ which does not converge in $[0, 2\pi)$

In fact, we shall see later that $[0, 2\pi)$ and $\mathcal{R}_f = \{z \in \mathbb{C} : |z| = 1\}$ are not homeomorphic, that is, there exist no homeomorphism $f : [0, 2\pi) \longrightarrow \{z \in \mathbb{C} : |z| = 1\}$.

Example 3.26. In the following table we shall consider sequences in $\mathbb{R}$ where $\mathbb{R}$ is equipped with different topologies.

	$\mathcal{T}_{d_0} = \mathcal{P}(\mathbb{R})$	$\mathcal{T} = \{\emptyset, \mathbb{R}\}$	$\mathcal{T} = \{\emptyset, [0,1], \mathbb{R}\}$	$\mathcal{T}_{d_2}$
$x_n = 1$	$\lim_{n\to\infty} x_n = 1$	$\forall x \in \mathbb{R}$: $\lim_{n\to\infty} x_n = x$	$\lim_{n\to\infty} x_n = 1$	$\lim_{n\to\infty} x_n = 1$
$y_n = \frac{1}{n}$	not convergent	$\forall y \in \mathbb{R}$: $\lim_{n\to\infty} y_n = y$	$\forall y \in \mathbb{R}$: $\lim_{n\to\infty} y_n = y$	$\lim_{n\to\infty} y_n = 0$
$z_n = -\frac{1}{n}$	not convergent	$\forall z \in \mathbb{R}$: $\lim_{n\to\infty} z_n = z$	$\forall z \in \mathbb{R} \setminus [0,1]$: $\lim_{n\to\infty} z_n = z$	$\lim_{n\to\infty} z_n = 0$
$u_n = n$	not convergent	$\forall\, u \in \mathbb{R}$: $\lim_{n\to\infty} u_n = u$	$\forall u \in \mathbb{R} \setminus [0,1]$: $\lim_{n\to\infty} u_n = u$	not convergent
$v_n = (1 + \frac{1}{n})^{-n}$	not convergent	$\lim_{n\to\infty} v_n = v$ for all $v \in \mathbb{R}$	$\forall v \in \mathbb{R}$: $\lim_{n\to\infty} v_n = v$	$\lim_{n\to\infty} v_n = \frac{1}{e}$

The ambivalence in columns $\mathcal{T} = \{\emptyset, \mathbb{R}\}$ and $\mathcal{T} = \{\emptyset, [0,1], \mathbb{R}\}$ are only possible since these topologies are not induced by a metric on $\mathbb{R}$. (We have shown earlier that a sequence in a metric space can only converge to one point.)

Theorem 3.27. *Let (X, d) be a metric space, then A is closed in $(X, \mathcal{T}_d)$ if and only if for any sequence (x_n) in A with $x_n \to x_0 \in X$ we have $x_0 \in A$.*

Proof. Let A be closed and $x_n \to x_0$ with (x_n) in A. If $x_0 \in A^c$, which is open, then exists $N \in \mathbb{N}$ with $x_n \in A^c$ for $n \geq N$, so in particular, $x_N \in A^c$, a contradiction to (x_n) in A.

Conversely, assume that A has the property that given any sequence (x_n) in A with $x_n \to x_0 \in X$ then automatically $x_0 \in A$. Assume A not closed, that is, A^c not open. Then exists x_0 in A^c with $B_\epsilon(x_0) \cap A \neq \emptyset$ for all $\epsilon > 0$. In particular, for $n \in \mathbb{N}$ we can pick $x_n \in B_{1/n}(x_0) \cap A$ and observe that (x_n) in A with $x_n \to x_0 \in X$ but $x_0 \notin A$, a contradiction. □

Remark 3.28. The characterization of closed sets in metric spaces in Theorem 3.27 does not hold in general topological space. In general topological spaces we need to replace the concept of converging sequence with converging filters.

Continuity at a point $x_0 \in X$ can be described in numerous ways.

Theorem 3.29. *Let $(X, d_X), (Y, d_Y)$ be metric spaces, $x_0 \in X$, and $f : X \longrightarrow Y$. The following are equivalent:*

i. The function f is continuous at x_0, that is, for all $\varepsilon > 0$ exists some $\delta > 0$ such that $f(B_\delta(x_0)) \subseteq B_\varepsilon(f(x_0))$.

ii. For all sequences (x_n) in X with $\lim_{n\to\infty} x_n = x_0$ we have $\lim_{n\to\infty} f(x_n) = f(x_0)$.

iii. For all open sets U in Y with $f(x_0) \in U$ exists V open in X with $x_0 \in V$ and $f(V) \subseteq U$.

Proof. *i* "⇒" *ii*. Fix $\epsilon > 0$. To find N with $f(x_n) \in B_\varepsilon(f(x_0))$ for all $n \geq N$, we first choose δ with $f(B_\delta(x_0)) \subseteq B_\varepsilon(f(x_0))$ and then $N = N_\delta$ with $x_n \in B_\delta(x_0)$ for all $n \geq N$. By construction, $f(x_n) \in B_\varepsilon(f(x_0))$ for all $n \geq N$.

ii "⇒" *iii*. Fix U open in Y with $f(x_0) \in U$. As U is open, we have $B_\varepsilon(f(x_0)) \subseteq U$ for some $\epsilon > 0$. Then, *ii* delivers $\delta > 0$ with $V = B_\delta(x_0)$ satisfies $f(V) \subseteq B_\varepsilon(f(x_0)) \subseteq U$.

iii "⇒" *i*. Fix $\epsilon > 0$. For open $U = B_\varepsilon(f(x_0))$ exists V open in X with $x_0 \in V$ and $f(V) \subseteq U$. But as V is open and contains x_0, there is $\delta > 0$ such that $B_\delta(x_0) \subseteq V$, hence, $f(B_\delta(x_0)) \subseteq B_\varepsilon(f(x_0))$. □

Theorem 3.30. *Let $(X, d_X), (Y, d_Y)$ be metric spaces and $f : X \longrightarrow Y$. The following are equivalent:*

i. The function f is continuous at every $x_0 \in X$, that is, for all $x_0 \in X$ and all $\varepsilon > 0$ exists some $\delta = \delta_{x_0,\epsilon} > 0$ such that $f(B_\delta(x_0)) \subseteq B_\varepsilon(f(x_0))$.

ii. For all sequences (x_n) that converge in X we have $\lim_{n\to\infty} f(x_n) = f(\lim_{n\to\infty} x_n)$.

iii. For all open sets U in Y we have $f^{-1}(U)$ is open in X.

iv. For all closed sets A in Y we have $f^{-1}(A)$ is closed in X.

Proof. Equivalence of *i* and *ii* follows from Theorem 3.29 as both statements describe continuity at each point x_0. The equivalence to *iii* is Theorem 3.13 Statements *iii* and *iv* are equivalent since $f^{-1}(A^c) = \big(f^{-1}(A)\big)^c$. □

Definition 3.31. Let $(X, \mathcal{T})$ be a topological space and let $A \subseteq X$.

i. The *interior* A° of A is given by $A^\circ = \bigcup_{\substack{U \subseteq A \\ U \in \mathcal{T}}} U$.

ii. The *closure* $\bar{A}$ of A is given by $\bar{A} = \bigcap_{\substack{C \supseteq A \\ C \text{ closed}}} C$. The set A is *dense* in X if $\overline{A} = X$.

iii. The *boundary* ∂A of A is given by $\partial A = \bar{A} \cap \overline{A^C}$.

iv. A point $x_0 \in X$ is called *cluster point* of A, if there exists a sequence (x_n) in A with $x_n \neq x_0$ and $\lim_{n\to\infty} x_n = x_0$. We denote by A' the *set of all cluster points* of A, that is

$$A' = \{x_0 \in X : \text{ there exists a sequence } (x_n) \text{ in } A \text{ with } x_n \neq x_0 \text{ and } \lim_{n\to\infty} x_n = x_0\}.$$

3.3. Compact sets

Even though the concept of compact and connected sets and spaces are of topological nature, we shall restrict our treatise to metric spaces (which certainly are just a special breed of topological spaces.)

Definition 3.32. Let A be a subset of a metric space (X, d) and let $\mathcal{U}$ and $\mathcal{V}$ be collections of subsets of X.

i. The family $\mathcal{U}$ is a *covering of* A if $A \subseteq \bigcup_{U \in \mathcal{U}} U$.

ii. The family $\mathcal{V}$ is a $\mathcal{U}$*-subcovering of* A if $\mathcal{V} \subseteq \mathcal{U}$ and $A \subseteq \bigcup_{U \in \mathcal{V}} U$.

iii. A family of sets $\mathcal{U}$ is called *open* if all $U \in \mathcal{U}$ are open.

iv. The family $\mathcal{U}$ is *finite* if $\mathcal{U}$ consists of finitely many sets (which in turn might contain infinitely many elements of X.)

Definition 3.33. A subset A of a metric space (X, d) is called *(covering-) compact* if **every** open cover $\mathcal{U}$ of A contains a finite $\mathcal{U}$-subcover $\mathcal{V}$.

Examples 3.34. i. Any finite set is compact. Indeed, if $\mathcal{U}$ is an open cover of a finite set $\{x_1, \ldots, x_N\}$, then we can choose $\mathcal{V} = \{U_1, \ldots, U_N\} \subseteq \mathcal{U}$ with $x_i \in U_i$ for $i = 1, \ldots, N$.

ii. The set $\{\frac{1}{n} : n \in \mathbb{N}\}$ is not compact. To see this, note that $\mathcal{U} = \{U_n = (\frac{1}{n+1}, \frac{1}{n-1}),\ n \in \mathbb{N}\}$ is an open cover of $\{\frac{1}{n} : n \in \mathbb{N}\}$ which has no finite $\mathcal{U}$-subcover of A. Indeed, the only $\mathcal{U}$-subcover of A is $\mathcal{U}$ itself which is not finite.

iii. The set $\{\frac{1}{n} : n \in \mathbb{N}\} \cup \{0\}$ is compact. Let $\mathcal{U}$ be an arbitrary open cover of $\{\frac{1}{n} : n \in \mathbb{N}\} \cup \{0\}$. Choose $U_0 \in \mathcal{U}$ with $0 \in U_0$. As U_0 is open and $\lim_{n\to\infty} \frac{1}{n} = 0$, there exists $N \in \mathbb{N}$ such that $\frac{1}{n} \in U_0$ for all $n \geq N$. For $n = 1, \ldots, N-1$, pick $U_n \in \mathcal{U}$ such that $\frac{1}{n} \in U_n$. Then, clearly, $\mathcal{V} = \{U_0, U_1, \ldots, U_{N-1}\}$ is a finite $\mathcal{U}$-subcover of $\{\frac{1}{n} : n \in \mathbb{N}\} \cup \{0\}$.

iv. In general, let (x_n) be a converging sequence in the metric space (X, d). Then $\{x_n : n \in \mathbb{N}\} \cup \{\lim_{n\to\infty} x_n\}$ is compact.

v. The open interval $(0,1) \subset \mathbb{R}$ is not compact in $(\mathbb{R}, d_2)$, since $\mathcal{U} = \{(\frac{1}{n}, 1)\}$ is an open cover of $(0,1)$ which contains no finite $\mathcal{U}$-subcover.

Definition 3.35. A subset A in the metric space (X, d) is *sequentially compact* if any sequence (a_n) in A has a subsequence (a_{n_k}) with $\lim_{k\to\infty} a_{n_k} = a_0$ and $a_0 \in A$.

One of the main goals of this section is to prove that in metric spaces sequentially compactness and covering compactness are the same, that is, a set A is sequentially compact if and only if A is covering compact. Be aware that this theorem does not hold in general topological spaces.

Before proving this theorem, we shall discuss some consequences of compactness.

Theorem 3.36. *Let (X, d) be a metric space and $A \subseteq X$ be compact. If $B \subset A$ is closed in X, then B is compact. Shortly: closed subsets of compact sets are compact.*

Proof. Let $\mathcal{U}$ be an open cover of B. As B^c is open too, $\widetilde{\mathcal{U}} = \mathcal{U} \cup \{B^c\}$ is an open cover of A. Since A is compact, there exists a finite $\widetilde{\mathcal{U}}$-subcover $\widetilde{\mathcal{V}}$ of A. Clearly, B^c does not contribute to covering B, so $\mathcal{V} = \widetilde{\mathcal{V}} \setminus \{B^c\}$ is a finite $\mathcal{U}$-subcover of B and the result is proven. □

Theorem 3.37. *Any compact set A in (X, d) is bounded, that is, compact sets are bounded.*

Proof. Pick $x_0 \in A$. Clearly $\mathcal{U} = \{B_n(x_0),\ n \in \mathbb{N}\}$ is an open cover of X and, hence, also of $A \subseteq X$. As A is compact, there exist $n_1 < n_2 < \ldots < n_K \in \mathbb{N}$ such that $A \subseteq B_{n_1}(x_0) \cup B_{n_2}(x_0) \cup \ldots \cup B_{n_K}(x_0) = B_{n_K}(x_0)$, so A is bounded. □

Theorem 3.38. *Any infinite subset B of a compact set A in (X, d) has at least one cluster point in A.*

Proof. Suppose that there exists no cluster point of B in A. Hence, for all $x \in A$ exists $\epsilon(x) > 0$ such that $B_{\epsilon(x)}(x)$ contains only finitely many points from B. Clearly, $\mathcal{U} = \{B_{\epsilon(x)}(x),\ x \in A\}$ is an open cover of A, hence, there exists a finite $\mathcal{U}$-subcover $\mathcal{V} = \{B_{\epsilon(x_i)}(x_i),\ i = 1, \ldots, N\}$. Since $B \subseteq A$, B is covered by $\mathcal{V}$, but, on the other hand, each of the finitely many $V \in \mathcal{V}$ covers only finitely many points in B, a contradiction to B is infinite. □

Theorem 3.39. *Compact sets are closed.*

Proof. Theorem 3.27 implies that if the compact set A is not closed, then exists a sequence (x_n) in A with $\lim_{n\to\infty} x_n = x_0 \in A^c$. Clearly, $x_n \neq x_0$ and it is not hard to see that the range of the sequence $\{x_n,\ n \in \mathbb{N}\}$ is an infinite set in A. Also, $\lim_{n\to\infty} x_n = x_0$ implies $x_0 \in A^c$ is the only cluster point of $\{x_n,\ n \in \mathbb{N}\}$ in X, so $\{x_n,\ n \in \mathbb{N}\}$ has no cluster point in A, a contradiction to A compact (Theorem 3.38). □

Theorem 3.39 combines with Theorem 3.37 to the statement that compact sets are closed and bounded. Does the converse hold? This would be nice since it would give us a criterium for compactness that is easily checked. Sadly, the converse does not hold in general (see Remark 3.48, but it does hold in euclidean space, that is, $\mathbb{R}^n$.

To prove the main result of this chapter, we need to introduce the concept of a Lebesgue number.

Definition 3.40. Let $\mathfrak{U}$ be a covering of a set A in the metric space (X, d). Any number $\lambda > 0$ with the property that for all $a \in A$ exists $U \in \mathfrak{U}$ such that $B_\lambda(a) \subseteq U$ is called a *Lebesgue number* for the covering $\mathfrak{U}$ of A.

Lemma 3.41. *Let $\mathfrak{U}$ be an open covering of a sequentially compact set A in the metric space (X, d). Then exists a Lebesgue number $\lambda > 0$ for the covering $\mathfrak{U}$ of A.*

Proof. Assume there is an open cover $\mathfrak{U}$ of A for which there exists no Lebesgue number, that is for all $n \in \mathbb{N}$ we can choose some $a_n \in A$ such that for all $B_{\frac{1}{n}}(a_n) \nsubseteq U$ for all $U \in \mathfrak{U}$.

Since A is sequential compact, we can extract a convergent subsequence $(a_{n_k})_k$ of (a_n) and set $a_0 := \lim_k a_{n_k} \in A$. Since $\mathfrak{U}$ is a covering, we have $a_0 \in U_0$ for some $U_0 \in \mathfrak{U}$. Since U_0 is open, there is an $n \in \mathbb{N}$ such that $B_{\frac{1}{n}}(a_0) \subseteq U_0$.

Pick $K \in \mathbb{N}$ such that $K \geq 2n$ and $d(a_{n_K}, a_0) < \frac{1}{2n}$. We have $B_{\frac{1}{n_K}}(a_{n_K}) \subseteq B_{\frac{1}{n}}(a_0)$ since $d(a, a_{n_{2n}}) < \frac{1}{2n}$ implies $d(a_0, a) < d(a_0, a_{n_K}) + d(a_{n_K}, a) < \frac{1}{2n} + \frac{1}{2n} = \frac{1}{n}$.

We conclude that $B_{\frac{1}{n_K}}(a_{n_K}) \subseteq B_{1/n}(a_0) \subseteq U_0$, a contradiction to $B_{\frac{1}{n_K}}(a_{n_K})$ not being subset of any $U \in \mathfrak{U}$. □

Now we can prove the main result of this chapter.

Theorem 3.42. *Let (X, d) be a metric space. The set $A \subseteq X$ is sequentially compact if and only if A is covering compact.*

Proof. Suppose A is covering compact. Let (x_n) be an arbitrary sequence in A. We have to find a convergent subsequence.

Cover A with balls of radius 1. Since (by covering-compactness) finitely many of them suffice, we throw away all but finitely many of them. Now among the remaining finitely many balls there has to be at least one ball containing x_n for infinitely many values of n. Let us call this ball B_1. Let n_1 be an index such that x_{n_1} is contained in B_1.

Now we do the same thing again: cover the set $\overline{B}_1 \cap A$, which is a covering-compact set, with (finitely many!) balls of radius $\frac{1}{2}$; one of them, which we call B_2, must have the property that $B_2 \cap B_1$ is visited infinitely often by the sequence. Choose $n_2 > n_1$ such that $x_{n_2} \in B_2 \cap B_1$. Now continue with $\overline{B}_2$ and radius $\frac{1}{4}$ to construct B_3 and n_3 and continue the process.

Set $C_n = \overline{\bigcap_{k=1}^n B_k} \cap A$ and observe that sequence $X \supseteq C_1 \supseteq C_2 \supseteq \ldots$. Since the nested intersection of non-empty compact sets whose diameter tends to zero is a single point x_0 (check!), we get by construction, $x_{n_k} \to x_0$. Since A is closed, we have $x_0 \in A$.

Let us now suppose that A is sequentially compact. Let $\mathfrak{U}$ be an arbitrary open cover. We want to show that $\mathfrak{U}$ admits a finite subcover. By Lemma 3.41, this cover has a Lebesgue-number $\lambda > 0$: for every $x \in X$ exists $U_x \in \mathfrak{U}$ such that $B_\lambda(x) \subseteq U_x$.

Choose any $x_1 \in X$. Then either U_{x_1} covers X and we are done. Otherwise choose any $x_2 \in X \setminus U_{x_1}$. Again, either $U_{x_1} \cup U_{x_2}$ already covers X and we are done, or we can choose $x_3 \in X \setminus (U_{x_1} \cup U_{x_2})$ and so on. Either X is covered after a finite number of steps, or this construction produces an infinite sequence (x_n) in X. However, this sequence has no convergent subsequence, because for all $m \neq n$, $d(x_m, x_n) \geq \lambda$. Hence this case is impossible. □

Lemma 3.43. *For $a \leq b$ we have $[a, b]$ is compact in $\mathbb{R}$. (Recall, if not specified otherwise, we let $d = d_2$ in $\mathbb{R}^n$, so in case of $\mathbb{R}$, $d(x, y) = |x - y|$.)*

Proof. The Bolzano Weierstrass Theorem (Theorem 2.20) states that $[a, b]$ is sequentially compact. □

Lemma 3.44. *Let A be compact in $(\mathbb{R}^n, d_i)$ and B be compact in $(\mathbb{R}^m, d_j)$, $i, j \in \{1, 2, \infty\}$. Then $A \times B$ is compact in $(\mathbb{R}^{n+m}, d_k)$, $k = 1, 2, \infty$.*

Proof. Since the topology on $(\mathbb{R}^n, d_i)$, $(\mathbb{R}^m, d_j)$ and $(\mathbb{R}^{n+m}, d_k)$ do not depend on $i, j, k \in \{1, 2, \infty\}$, we may assume that $i = j = k = 1$.

For $\big((x_n, y_n)\big)_{n\in\mathbb{N}}$ we have $\lim\limits_{n\to\infty} (x_n, y_n) = (x_0, y_0)$ in $(\mathbb{R}^{n+m}, d_1)$ if and only if $\lim\limits_{n\to\infty} x_n = x_0$ $(\mathbb{R}^n, d_1)$ and $\lim\limits_{n\to\infty} y_n = y_0$ and $(\mathbb{R}^m, d_1)$, since $d_1\big((x_n, y_n), (x_0, y_0)\big) = d_1\big(x_n, x_0\big) + d_1\big(y_n, y_0\big)$

Let $\big((a_n, b_n)\big)_{n\in N}$ be a sequence in $A \times B$. We shall construct a subsequence of $\big((a_n, b_n)\big)_{n\in N}$ which converges in $A \times B$.

Using sequential compactness of A, we choose a subsequence $(a_{n_k})_{k\in\mathbb{N}}$ of $\big(a_n\big)_{n\in N}$ which converges to $a_0 \in A$. Similarly, we pick a subsequence $(b_{n_{k_l}})_{l\in\mathbb{N}}$ of $(b_{n_k})_{k\in\mathbb{N}}$ which converges to $b_0 \in B$. The subsequence $\big((a_{n_{k_l}}, b_{n_{k_l}})\big)_{l\in\mathbb{N}}$ of $\big((a_n, b_n)\big)_{n\in N}$ obviously converges to $(a_0, b_0) \in A \times B$. □

Theorem 3.45. *Any set of the form $[a_1, b_1] \times [a_2, b_2] \times \ldots \times [a_n, b_n] \subset \mathbb{R}^n$ is compact.*

Proof. Proof by induction using Lemma 3.44. □

Theorem 3.46. (Heine–Borel) *Consider the metric space $\mathbb{R}^n$ equipped with one of the standard metrics d_1, d_2 or d_∞. Any $A \subset \mathbb{R}^n$ is compact if and only if A is closed and bounded.*

Proof. If A is bounded it is contained in some set of the form $[a_1, b_1] \times [a_2, b_2] \times \ldots \times [a_n, b_n] \subset \mathbb{R}^n$ which is compact by Theorem 3.45. Since A is therefore a closed subset of a compact set, we have A compact by Theorem 3.36. □

Remark 3.47. Heine-Borel does not hold in generic metric spaces. The easiest example is the metric space $\mathbb{Q}$ with distance metric inherited from $\mathbb{R}$, so $d(2, \frac{9}{2}) = \frac{5}{2}$. Clearly, the set A of rationals of absolute value less than or equal to 3 is closed and bounded in $\mathbb{Q}$, but not compact. Indeed, the sequence $\{(1 + \frac{1}{n})^n\}_{n\in\mathbb{N}}$ is in A but has no convergent subsequence in A (it approaches e which is not rational).

Even if we "kill the gaps" by requiring the metric space to be complete, Heine-Borel is not generally applicable. For example, it does not hold in the infinite dimensional metric space $\ell^2(\mathbb{N})$ which consists of those complex valued sequences with $\|\{c_n\}_{n\in\mathbb{N}}\|_2 = \sqrt{\sum_{n=1}^{\infty} |c_n|^2} < \infty$, equipped with metric

$$d(\{c_n\}, \{d_n\}) = \|\{c_n\} - \{d_n\}\|_2 = \sqrt{\sum_{n=1}^{\infty} |c_n - d_n|^2}.$$

Indeed, the closed unit ball in $\ell^2(\mathbb{N})$ is closed and bounded, but not compact. To see this, simply observe that the sequence of sequences $\{e^k\}_{k\in\mathbb{N}}$ with $e^k_n = 1$ if $k = n$ and 0 else satisfies $d(e^k, e^\ell) = \sqrt{2}$ if $k \neq \ell$, and, hence, $\{e^k\}_{k\in\mathbb{N}}$ has no convergent subsequences.

In case of complete metric spaces, this problem can be fixed by replacing *bounded* by a stronger concept, namely, *totally bounded*. A set A is called totally bounded if for each $\epsilon > 0$ there exists a finite set of points $\{x_1, \ldots, x_N\}$ in A with $A \subseteq \bigcup_{n=1}^{N} B_\epsilon(x_n)$.

We claim that a subset of a complete metric space is compact if and only if it is totally bounded and closed. The one direction is trivial, compactness implies clearly totally boundedness, and we proved above that it implies that the set is closed. Now, let assume that a set A in a complete metric space is totally bounded and closed. Given a sequence, we construct a subsequence as in Theorem 3.42. The completeness of the metric space implies that the intersection of nested sets whose diameter tends to zero contains exactly one point x_0. Since A is closed, and the original sequence was in A, we have that $x_0 \in A$. Hence, the sequence has a subsequence that converges in A, so we showed that the set is sequentially compact, which is equivalent to covering compact in metric spaces.

Remark 3.48. The continuous functions

$$f_n : [0,1] \longrightarrow \mathbb{R}, \quad x \mapsto \begin{cases} 1, & \text{for } x \leq \frac{1}{n+1} \\ -n(n+1)x + n + 1, & \text{for } \frac{1}{n+1} < x \leq \frac{1}{n} \\ 0, & \text{for } \frac{1}{n} < x \leq 1 \end{cases}$$

in $C([0,1])$ have the properties $d(f_n, f_m) = 1$ if $n \neq m$ and $d(f_n, 0) = 1$. The set $A = \{f_n, \quad n \in \mathbb{N}\} \subset B_2(0)$ is bounded in $C([0,1])$ and closed, since any convergent sequence in A converges to a limit in A (there are no convergent sequences in A). But A is not compact, since the open covering

$$\mathfrak{U} = \{B_{\frac{1}{2}}(f_n)\}$$

contains no finite $\mathfrak{U}$–subcovering of A. See Figure 3.3 for an illustration of this example.

As additional example let us consider $\mathbb{R}$ with the discrete metric and $A = (0,1)$, or $\mathbb{R}^n$ with the metric $\widetilde{d_2} : (x,y) \mapsto \frac{d_2(x,y)}{1+d_2(x,y)}$ and $A = \mathbb{R}^n$. In both cases A is bounded and closed but not compact.

Theorem 3.49. *A compact metric space (X,d) is complete.*

Proof. Let (x_n) be Cauchy in X. Since X is sequentially compact, (x_n) has a subsequence that converges in X. But then, (x_n) converges in X (to the same limit) by Proposition 2.36. □

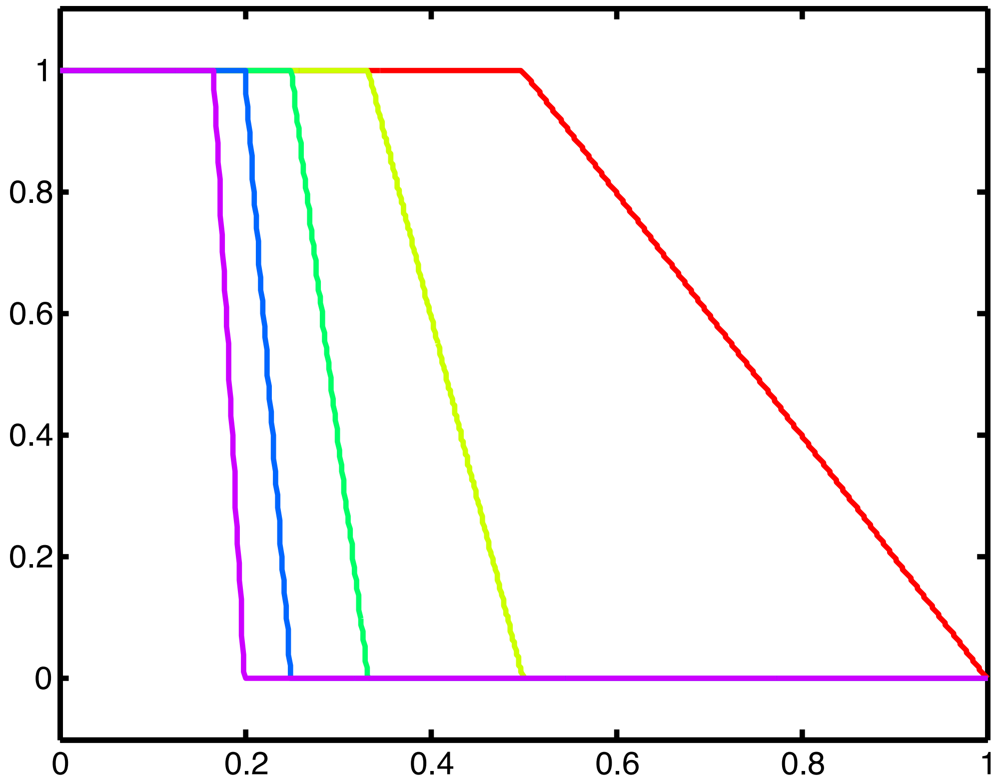

Figure 4. The functions f_1 (red), f_2 (yellowish green), f_3 (green), f_4 (blue), and f_5 (purple) from Remark 3.48.

Theorem 3.50. *Let* (X, d_X) *be compact, and* $f : (X, d_X) \longrightarrow (Y, d_Y)$ *be continuous. Then* $\mathcal{R}_f = f(X)$ *is compact in* (Y, d_Y).

Proof. We shall show that every sequence in $\mathcal{R}_f$ has a subsequence that converges in $\mathcal{R}_f$. Let (y_n) be a sequence in $\mathcal{R}_f$. Choose x_n with $f(x_n) = y_n$. The sequence (x_n) has a subsequence (x_{n_k}) with $\lim_{k\to\infty} x_{n_k} x_0$ by compactness of X, continuity then implies that its image (y_{n_k}) converges. □

To appreciate compactness some more, let us visit a *strong* form of continuity.

Definition 3.51. Let (X, d_X) and (Y, d_Y) be metric spaces. A function $f : X \longrightarrow Y$ is *uniformly continuous* on X, if for all $\varepsilon \in \mathbb{R} > 0$ exists $\delta > 0$ s.t. $d_Y(f(x), f(\widetilde{x})) < \varepsilon$ for all $x, \widetilde{x}$ with $d_X(x, \widetilde{x}) < \delta$.

This is obviously equivalent to $\forall \varepsilon > 0 \; \exists \; \delta > 0$ s.t. $\forall x \in X \;\; f(B_\delta(x)) \subseteq B_\varepsilon(f(x))$.

Proposition 3.52. *Any uniformly continuous function* $f : (X, d_X) \longrightarrow (Y, d_Y)$ *is continuous.*

Example 3.53. i. $f : \mathbb{R} \longrightarrow \mathbb{R}, \; x \mapsto 2x$ is uniformly continuous.

ii. $f : \mathbb{R}^+ \longrightarrow \mathbb{R}, \; x \mapsto \frac{1}{x}$ is continuous but not uniformly continuous.

Theorem 3.54. *Any continuous function defined on compact metric spaces is uniformly continuous. That is, given a compact metric space* (X, d_X) *and continuous* $f : (X, d_X) \longrightarrow (Y, d_Y)$, *then* f *is uniformly continuous as well.*

3.4. Connected sets

While dealing with a topological concept, we consider only metric spaces.

Definition 3.55. A *separation* of a metric space (X, d) is a pair of nonempty open subsets $U, V \subset X$ with $X = U \cup V$ and $\emptyset = U \cap V$. (Note that then U and V are also complements of open sets, that is, closed sets.)

A metric space (X, d) is *connected* if there exists no separation of X, that is, if X and $\emptyset$ are the only subsets of X that are both, open and closed.

A subset A of the metric space (X, d) is connected if the metric space $(A, d|_{A\times A})$ is connected.

Example 3.56. *i.* The set $[0, 1] \cup (3, 4]$ is a subset of $\mathbb{R}$ which is not connected. To see this, we have to find a subset A of the metric space $([0, 1] \cup (3, 4], d_{\mathbb{R}})$ which is open, closed, not empty, and, not all of $[0, 1] \cup (3, 4]$. But indeed, $[0, 1]$ is such a set. At first sight, it may be surprising that the set $[0, 1]$ is open, and, it is indeed not open as a subset of $\mathbb{R}$. But here, we need to show that $[0, 1]$ is an open subset of the metric space $([0, 1] \cup (3, 4], d_{\mathbb{R}})$. To see this, you can use the inheritance principle, i.e., it satisfies to find an open set U in $\mathbb{R}$ such that $[0, 1] = U \cap ([0, 1] \cup (3, 4])$. Clearly, the set $U = (-1, 2)$ has this property.

ii. The set $[0, 1]$ is a connected subset of $\mathbb{R}$. First, observe that if A is a closed subset of $([0, 1], d_{\mathbb{R}})$, then A is also closed in $\mathbb{R}$. Indeed, the inheritance principle states that for A closed in $([0,1], d_{\mathbb{R}})$ exists $\widetilde{A}$ closed in $\mathbb{R}$ with $\widetilde{A} \cap [0,1] = A$ which is closed as intersection of closed sets.

Now, assume that $A \subsetneq [0,1]$ is closed and open in $([0,1], d_{\mathbb{R}})$, and non-empty. Without loss of generality, $1 \notin A$ (else, replace A by A^c). Let $x_0 = \sup A \in \mathbb{R}$. As the discussion above showed that A is not only closed in $([0,1], d_{\mathbb{R}})$, but also in $(\mathbb{R}, d_{\mathbb{R}})$, we have $x_0 \in A$ (Choosing $x_n \in A$ with $x_n \to x_0$ in $\mathbb{R}$, A closed in $\mathbb{R}$ implies $x_0 \in A$.) This implies in particular, $x_0 < 1$. As A is also open in the metric space $([0,1], d_{\mathbb{R}})$, there exists $\epsilon > 0$ such that if $x \in [0,1]$ satisfies $|x - x_0| < \epsilon$, then $x \in A$. This clearly contradicts that $x_0 = \sup A \in \mathbb{R}$ since $x_0 + \min\{\frac{\epsilon}{2}, \frac{1-x_0}{2}\} \in A$.

The most important result of this section is fairly elementary:

Theorem 3.57. *If (X, d_X) is connected and $f : (X, d_X) \longrightarrow (Y, d_Y)$ is continuous, then $\mathcal{R}_f = f(X)$ is connected.*

Proof. We have to show that $(f(X), d_Y)$ is a connected metric space. To this end, first observe that the continuity of $f : (X, d_X) \longrightarrow (Y, d_Y)$ implies the continuity of $\widetilde{f} : (X, d_X) \longrightarrow (f(X), d_Y)$, $\widetilde{f}(x) = f(x)$ for $x \in X$. Indeed, the inheritance principle states that doe U open in $f(X)$ exists $\widetilde{U}$ open in Y such that $U = \widetilde{U} \cap f(X)$. Then $\widetilde{f}^{-1}(U) = f^{-1}(\widetilde{U})$ is open in X by continuity of f.

We proof the result by contraposition. If $f(X)$ is not connected, then exists a nontrivial (neither $f(X)$, nor $\emptyset$) clopen subset A in $(f(X), d_Y)$. By continuity of $\widetilde{f}$, $f^{-1}(A)$ is clopen in (X, d_X), and, clearly, $f^{-1}(A) \neq X, \emptyset$, that is, X is not connected. □

Remark 3.58. Using the fact that images of compacts under continuous transformations are compact and that images of connected sets under continuous transformations are connected, we can easily see that none of the sets

i. $[0,1] \subset \mathbb{R}$.

ii. $[0,1) \subset \mathbb{R}$.

iii. $S^1 = \{z \in \mathbb{C} : \ |z| = 1\}$ in $\mathbb{C}$.

iv. The 8 set $S^1 \cup \{z \in \mathbb{C} : \ |z - 2i| = 1\}$ in $\mathbb{C}$.

is homeomorphic to another set in the list.

Theorem 3.59. *Let us consider the real line $\mathbb{R}$ with metric d_1, d_2, and d_∞. The following are equivalent:*

i. The set $A \subset \mathbb{R}$ is connected.

ii. For any $a, b \in A \subset \mathbb{R}$ and any $c \in \mathbb{R}$ with with $a < c < b$ we have $c \in A$.

iii. The set $A \subset \mathbb{R}$ is an (possibly unbounded) interval.

That is, connected sets in $\mathbb{R}$ are exactly the intervals.

Proof. $i \Rightarrow ii$. If there exist $a, b, c \in \mathbb{R}$ with $a < c < b$ and $a, b \in A$ but $c \notin A$, then $A \cap (-\infty, c)$ and $B \cap (c, \infty)$ are two nonempty open sets in $(A, d_{\mathbb{R}})$ whose union is A.

$ii \Rightarrow i$. Proof by contradiction. Let U, V form a separation of $(A, d_{\mathbb{R}})$. Choose $u \in U \subseteq A$ and $v \in V \subseteq A$. Without loss of generality, $u < v$. Property ii. implies that then $[u, v] \subseteq A$. Also, the inheritance principle implies that $U \cap [u, v]$ and $V \cap [u, v]$ are a separation of $([u, v], d_{\mathbb{R}})$. But as seen for $u = 0$ and $v = 1$ in Example 3.56.ii, $[u, v]$ is connected.

$ii \Leftrightarrow iii$. This is obvious. □

Corollary 3.60. *Let $f : [a, b] \longrightarrow \mathbb{R}$ be continuous. Then exists $c, d \in \mathbb{R}$ with $f([a, b]) = [c, d]$.*

Proof. As $[a, b]$ is compact and connected, so is its image $f([a, b])$. But the only sets in $\mathbb{R}$ that are compact and connected are closed intervals. □

Theorem 3.61. (Intermediate value theorem) *Let (X, d) be connected and $f : X \longrightarrow \mathbb{R}$ be continuous. Given any x_1, x_2 in X and $c \in \mathbb{R}$ with $f(x_1) < c < f(x_2)$, then exists $x \in X$ with $f(x) = c$.*

Theorem 3.62. *Let S_i, $i \in I$ be a family of connected sets in a metric space (X, d). If $\bigcap_{i \in I} S_i \neq \emptyset$, then $\bigcup_{i \in I} S_i$ is connected.*

Proof. Homework. □

Example 3.63. Open and closed balls in $(\mathbb{R}^n, d_i)$, $i = 1, 2, \infty$ are connected. To see this, let A be an open or closed ball in $(\mathbb{R}^n, d_{i_0})$, for some $i_0 \in \{1, 2, \infty\}$. For $x \in A$ consider

$$f_x : [0, 1] \longrightarrow \mathbb{R}^n, \quad t \mapsto tx + (1 - t)x_0.$$

The functions f_x are continuous and their ranges $\mathcal{R}_{f_x}$ are therefore connected. The result follows from Theorem 3.62 since

$$A = \bigcup_{x \in A} \mathcal{R}_{f_x} \quad \text{and} \quad \bigcap_{x \in A} \mathcal{R}_{f_x} = \{x_0\} \neq \emptyset.$$

Definition 3.64. A metric space (X, d) is called *totally disconnected* if for each $x \in X$ and $\epsilon > 0$ exists a clopen set A in X with $x \in A \subseteq B_\epsilon(x)$.

Example 3.65. Cantor's middle third set is an uncountable set which is totally disconnected.

3.5. Sequences of functions, uniform convergence

In this section we shall discuss in detail the metric space $C(X)$ of continuous, complex valued functions defined on a compact metric space X. You may want to read X to be a compact interval in $\mathbb{R}$ with the usual metric. The functions considered are complex valued, but $\mathbb{C}$ (with its usual metric) can be easily replaced with another metric space (Y, d_Y). (For Theorem3.72 we need (Y, d_Y) to be complete.)

The metric on $C(X)$ has been discussed in numerous homework problems.

Definition 3.66. Let (X, d_X) be a metric space and let $B(X)$ be the set of all bounded, complex valued functions on X, that is,

$$B(X) = \big\{f : X \longrightarrow \mathbb{C} : \text{ for } f \text{ exists } M \in \mathbb{R}^+ \text{ such that } |f(x)| \leq M \text{ for all } x \in X\big\}.$$

On $B(X)$ we can define the metric

$$d_\infty(f, g) = \sup\{|f(x) - g(x)| : \ x \in X\}.$$

The set of continuous, complex valued functions on X is denoted by $C(X)$.

Definition 3.67. Let (X, d_X) be a metric spaces and let $f_n : \ X \longrightarrow \mathbb{C}$, $n \in \mathbb{N}$ be a sequence of functions mapping X to Y.

The sequence $(f_n)_{n\in\mathbb{N}}$ *converges pointwise* to $f_0 : \ X \longrightarrow \mathbb{C}$, if $\lim\limits_{n\to\infty} f_n(x) = f_0(x)$ for all $x \in X$, that is, if for all $x \in X$ and $\epsilon > 0$ exists $N = N(x, \epsilon) \in \mathbb{N}$ such that $|f_n(x) - f_0(x)| \leq \epsilon$ for all $n \geq N$.

The sequence $(f_n)_{n\in\mathbb{N}}$ *converges uniformly* to $f_0 : \ X \longrightarrow Y$, if for all $\epsilon > 0$ exists $N = N(\epsilon) \in \mathbb{N}$ such that

$$|f_n(x) - f_0(x)| < \epsilon \qquad \text{for all } x \in X \text{ and for all } n \geq N.$$

That is

$$d_\infty(f_n, f_0) = \sup\big\{|f_n(x) - f_0(x)| : \ x \in X\big\} \stackrel{n\to\infty}{\longrightarrow} 0.$$

Remark 3.68. Note that if (f_n) converges uniformly to f_0, then (f_n) converges also pointwise to f_0.

Example 3.69. Let us revisit the functions given in Remark 3.48 and shown in Figure 3.3. The functions

$$f_n : [0,1] \longrightarrow \mathbb{R}, \qquad x \mapsto \begin{cases} 1, & \text{for } x \leq \frac{1}{n+1} \\ -n(n+1)x + n + 1, & \text{for } \frac{1}{n+1} < x \leq \frac{1}{n} \\ 0, & \text{for } \frac{1}{n} < x \leq 1 \end{cases}$$

are continuous functions in $B([0,1])$ that converge pointwise to the discontinuous function

$$f_0 : [0,1] \longrightarrow \mathbb{R}, \qquad x \mapsto \begin{cases} 1, & \text{for } x = 0 \\ 0, & \text{for } 0 < x \leq 1 \end{cases}$$

As $d_\infty(f_n, f_m) = 1$ if $n \neq m$, the sequence (f_n) is not Cauchy in $(B([0,1]), d_\infty)$ and does therefore not converge in $(B([0,1]), d_\infty)$ (convergent sequences are always Cauchy). Hence f_n does not converge uniformly to f_0.

Theorem 3.70. *$C(X) \cap B(X)$ is a closed set in $(B(X), d_\infty)$, hence, if (f_n) is a sequence of continuous functions in $(B(X), d_\infty)$ which converges uniformly to f_0, then f_0 is continuous on X.*

Proof. Fix $x_0 \in X$ and $\epsilon > 0$. Uniform convergence of (f_n) provides us with $N \in \mathbb{N}$ such that $d_\infty(f_n, f_0) = \sup\left\{d_Y(f_n(x), f_0(x)) : \ x \in X\right\} < \epsilon/3$ for all $n \geq N$. Now, use continuity of f_N to pick $\delta > 0$ with $f_N(B_\delta(x_0)) \subseteq B_{\epsilon/3}(f_N(x_0))$. Then

$$|f_0(x_0) - f_0(x)| \leq |f_0(x_0) - f_N(x_0)| + |f_N(x_0) - f_N(x)| + |f_N(x) - f_0(x)| < \epsilon/3 + \epsilon/3 + \epsilon/3 = \epsilon$$

for all $x \in B_\delta(x_0)$, that is, $f_0(B_\delta(x_0)) \subseteq B_\epsilon(f_0(x_0))$, so f_0 is continuous. □

Remark 3.71. The result above allows us to interchange limits in the following setting. Let (x_k) in X with $\lim\limits_{k\to\infty} x_k = x_0$. Then

$$\lim_{k\to\infty}\lim_{n\to\infty} f_n(x_k) = \lim_{k\to\infty} f_0(x_k) \stackrel{!}{=} f_0(x_0) = \lim_{n\to\infty} f_n(x_0) = \lim_{n\to\infty}\lim_{k\to\infty} f_n(x_k).$$

Quite a few results in analysis require to interchange limits as in $\lim_{k\to\infty}\lim_{n\to\infty} f_n(x_k) = \lim_{n\to\infty}\lim_{k\to\infty} f_n(x_k)$. But this always requires careful justification. For example,

$$\lim_{k\to\infty}\lim_{n\to\infty}\frac{k}{k+n} = \lim_{k\to\infty} 0 = 0 \neq 1 = \lim_{n\to\infty} 1 = \lim_{n\to\infty}\lim_{k\to\infty}\frac{1}{1+\frac{n}{k}} = \lim_{n\to\infty}\lim_{k\to\infty}\frac{k}{k+n}.$$

Theorem 3.72. *The metric spaces $(B(X), d_\infty)$ and $(C(X) \cap B(X), d_\infty)$ are complete.*

Proof. Let (f_n) be a Cauchy sequence in $(B(X), d_\infty)$. Fix $\epsilon > 0$ and $x_0 \in X$. Choose $N \in N$ such that $d_\infty(f_n, f_m) < \epsilon/2$ for all $n, m \geq N$. Then, in particular

$$|f_n(x_0) - f_m(x_0)| \leq d_\infty(f_n, f_m) < \epsilon/2 \text{ for all } n, m \geq N, \tag{2}$$

and, completeness of $\mathbb{C}$ implies that there exists y_0 with $\lim_{n\to\infty} f_n(x_0) = y_0$.

This argument works for any $x_0 \in X$, hence, we assign to any $x = x_0$ a $y = y_0$, thereby defining a function f_0 on X with $f_0(x_0) = y_0$ and f_n converges pointwise to f_0.

For uniform convergence, simply observe that, for $n \geq N$,

$$d_\infty(f_n, f_0) = \sup_{x\in X}|f_n(x) - f_0(x)| = \sup_{x\in X}\lim_{m\to\infty}|f_n(x) - f_m(x)| \leq \epsilon/2 < \epsilon.$$

Any closed subspace of a complete metric space is complete, hence, completeness of $(C(X) \cap B(X), d_\infty)$ follows from Theorem 3.70. □

Corollary 3.73. *If (X, d) is compact, then $(C(X), d_\infty)$ is a complete metric space.*

Proof. Since (X, d) is compact we have $f(X)$ is compact and therefore bounded for any continuous $f : X \longrightarrow \mathbb{C}$. Hence $C(X) \subseteq B(X)$ and $(C(X), d_\infty) = (C(X) \cap B(X), d_\infty)$ is complete. □

4. DIFFERENTIATION

4.1. Central results

In this section, we shall discuss derivatives of real valued functions defined on subsets of $\mathbb{R}$. One of our objectives is to illuminate the interplay of continuity and differentiability in analysis.

To define derivatives of real valued functions, we shall analyze so–called difference quotients. The discussion of such requires the following definition of functional limits.

Definition 4.1. Let (X, d_X) and (Y, d_Y) be metric spaces and let f map X to Y. If x is a cluster point in X, we write $f(x) \to y_0$ as $x \to x_0$ or $\lim_{x\to x_0} f(x) = y_0$ if $y_0 \in Y$ and if for any $\epsilon > 0$ exists $\delta > 0$ such that $d_Y(f(x), y_0) < \epsilon$ whenever $0 < d_X(x, x_0) < \delta$. The point $y_0 \in Y$ is called functional limit of f as x approaches x_0.

Remark 4.2. If we restrict ourselves to cluster points, we could rephrase previous results using functional limits. For example., we have:

i. If x is a cluster point in (X, d_X), then $\lim_{x\to x_0} f(x) = y_0$ if and only if for all sequences (x_n) in X with $x_n \neq x_0$, $n \in \mathbb{N}$, we have $\lim_{n\to\infty} f(x_n) = y_0$.

ii. Let (X, d_X) and (Y, d_Y) be metric spaces, let f map X to Y, and let x be a cluster point in (X, d_X). Then $\lim_{x\to x_0} f(x) = f(x_0)$ if and only if f is continuous at x_0.

iii. For U open in $\mathbb{R}$ we have $U' \supset U$, hence, the restriction to cluster points will not play a role in the following discussion of derivatives. By the way, any set A in a metric space (X, d) with $A = A'$ is called *perfect*.

Definition 4.3. Let $A \subseteq \mathbb{R}$ and $f : A \longrightarrow \mathbb{R}$. We say that f is *differentiable* at a cluster point x_0 in A, that is, at $x_0 \in A \cap A'$, if

$$\lim_{x\to x_0} \frac{f(x) - f(x_0)}{x - x_0} = L$$

for some $L \in \mathbb{R}$. In this case L is called derivative of f at x_0 and we write $f'(x_0) = L$. If $A \subseteq A'$ and f is differentiable at x for all $x \in A$, then we call f differentiable on A.

Further, we have that $f'(x_0) = L$ if and only if $\lim_{h\to 0} \dfrac{f(x_0 + h) - f(x_0)}{h} = L.$

In order to avoid "cluster point" disclaimers, we shall mostly restrict ourselves to consider open sets U as domains of differentiable functions. Open subsets of $\mathbb{R}$ have the property that all its elements are cluster points.

Proposition 4.4. *For* $\exp : \mathbb{R} \longrightarrow \mathbb{R}$, $x \mapsto \sum_{n=0}^{\infty} \frac{x^n}{n!}$, *we have* $\exp'(x) = \exp(x)$.

Proof. For $x_0 \in \mathbb{R}$,

$$\begin{aligned}\frac{\exp(x_0+h)-\exp(x_0)}{h} &\overset{(1)}{=} \frac{\exp(x_0)\exp(h)-\exp(x_0)}{h} = \exp(x_0)\frac{\exp(h)-1}{h}\\ &= \exp(x_0)\frac{\sum_{n=0}^{\infty}\frac{h^n}{n!}-1}{h} = \exp(x_0)\frac{1+h+\frac{h^2}{2}+\frac{h^3}{3}+\ldots-1}{h}\\ &\overset{(2)}{=} \exp(x_0)\frac{h+\frac{h^2}{2}+\frac{h^3}{3}+\ldots}{h} \overset{(3)}{=} \exp(x_0)\big(1+\frac{h}{2}+\frac{h^2}{3}+\ldots\big)\\ &\overset{(4)}{=} \exp(x_0)\big(1+h\big(\frac{1}{2}+\frac{h}{3}+\ldots\big)\big).\end{aligned}$$

Note that (1) follows from Corollary 2.63 and (2), (3), and (4) from the algebraic limit theorem applied to the partial sums $\sum_{n=0}^{N}\frac{h^n}{n!}$. As $\frac{1}{2}+\frac{h}{3}+\ldots$ is easily seen to converge in $\mathbb{R}$ for $|h|\leq 1$, we have invoking again the algebraic limit theorem, $\lim_{h\to 0}\big(1+h\big(\frac{1}{2}+\frac{h}{3}+\ldots\big)\big)=1$. □

Differentiable functions are continuous:

Theorem 4.5. *For U open in $\mathbb{R}$ and $f: U \longrightarrow \mathbb{R}$ differentiable at $x_0 \in U$ we have f continuous at x_0.*

Theorem 4.6. (Sum, product, and quotient rule) *Let U be open in $\mathbb{R}$ and $f, g: U \longrightarrow \mathbb{R}$ be differentiable at $x_0 \in U$. Then*

i. $f+g$ is differentiable at x_0 and $(f+g)'(x_0) = f'(x_0)+g'(x_0)$.

ii. fg is differentiable at x_0 and $(fg)'(x_0) = f'(x_0)g(x_0)+f(x_0)g'(x_0)$.

iii. If $g(x_0)\neq 0$, then $\frac{f}{g}$ is differentiable at x_0 and $\left(\frac{f}{g}\right)'(x_0) = \frac{f'(x_0)g(x_0)-f(x_0)g'(x_0)}{g(x_0)^2}$.

Proof. These are direct consequences of the algebraic limit theorem. These are applicable to functional limits, as if x_0 is a cluster point in the domain of a function F, then $\lim_{x\to x_0}F(x)=\alpha$ if and only if $\lim_{n\to\infty}F(x_n)=\alpha$ for all $x_n \longrightarrow x_0$ with $x_n \neq x_0$.

To show *iii*, first consider the case $f(x)=1$ for all $x\in U$ and then upgrade your result by using *ii*. □

Theorem 4.7. (Chain rule) *Let U, V be open in $\mathbb{R}$ and $f: U \longrightarrow V$ be differentiable at $x_0 \in U$ and $g: V \longrightarrow \mathbb{R}$ be differentiable at $y_0 = f(x_0) \in V$. Then $g\circ f$ is differentiable at x_0 and we have $(g\circ f)'(x_0) = g'(f(x_0))f'(x_0)$.*

Proof. Let

$$\widetilde{g}(y) = \begin{cases}\frac{g(y)-g(y_0)}{y-y_0}, & \text{if } y\in V,\ y\neq y_0\\ g'(y_0), & \text{if } y=y_0.\end{cases}$$

We have $g(y)-g(y_0)=\widetilde{g}(y)(y-y_0)$ for all $y\in V$ and differentiability of g at y_0 implies that $\lim_{y\to y_0}\widetilde{g}(y)=g'(y_0)$. We have

$$\frac{g(f(x))-g(f(x_0))}{x-x_0}=\frac{\widetilde{g}(f(x))(f(x)-f(x_0))}{x-x_0}=\widetilde{g}(f(x))\frac{f(x)-f(x_0)}{x-x_0}.$$

Since f is differentiable at x_0, the second factor converges as $x\to x_0$. As f is also continuous at x_0 and $\widetilde{g}$ is continuous at $f(x_0)$, the algebraic limit theorem implies

$$\lim_{x\to x_0}\frac{g(f(x))-g(f(x_0))}{x-x_0}=\lim_{x\to x_0}\widetilde{g}(f(x))\lim_{x\to x_0}\frac{f(x)-f(x_0)}{x-x_0}=g'(f(x_0))f'(x_0). \qquad \square$$

Examples 4.8. For $n=0,1,2,3$ set $f_n(x)=\begin{cases} x^n\sin\left(\frac{1}{x}\right) & \text{if } x\neq 0\\ 0 & \text{if } x=0\end{cases}$. Note that f_n, $n=0,1,2,3$, is continuous and differentiable on $\mathbb{R}\setminus\{0\}$, and its derivative is a continuous function on $\mathbb{R}\setminus\{0\}$.

i. f_0 is not continuous at 0.

ii. f_1 is continuous at 0 but not differentiable at 0.

iii. f_2 is differentiable at 0, and, hence, on $\mathbb{R}$, but its derivative f_2' is not continuous at 0.

iv. f_3 is again differentiable on $\mathbb{R}$ and its derivative f_3' is continuous on $\mathbb{R}$.

Theorem 4.9. Interior extremum theorem. *Let $U\subset\mathbb{R}$ be open and $f:U\longrightarrow\mathbb{R}$ be differentiable at $x_0\in U$. If there exists a maximum [resp. minimum] of f at x_0, then $f'(x_0)=0$.*

Proof. Assume $f(x_0)\geq f(x)$ for all $x\in U$. Choose $x_n\longrightarrow x_0$, $x_n<x_0$ for all $n\in\mathbb{N}$. Then $\frac{f(x_n)-f(x_0)}{x_n-x_0}\geq 0$ for all $n\in\mathbb{N}$, and, by the algebraic limit theorem

$$f'(x_0)=\lim_{n\to\infty}\frac{f(x_n)-f(x_0)}{x_n-x_0}\geq 0.$$

Similarly, choosing $\widetilde{x}_n\longrightarrow x_0$, $\widetilde{x}_n>x_0$ for all $n\in\mathbb{N}$, we obtain $\frac{f(\widetilde{x}_n)-f(x_0)}{\widetilde{x}_n-x_0}\leq 0$ for all $n\in\mathbb{N}$ which implies $f'(x_0)\leq 0$, and, in combination with $f'(x_0)\geq 0$ we obtain $f'(x_0)=0$.

For $f(x_0)\leq f(x)$ for all $x\in U$, apply the previous case to $\widetilde{f}(x)=-f(x)$. $\square$

Theorem 4.10. Rolle's theorem. *Let $b>a$, and $f:[a,b]\longrightarrow\mathbb{R}$ be continuous and differentiable on (a,b). If $f(a)=f(b)$, then exists $x_0\in(a,b)$ such that $f'(x_0)=0$.*

Proof. Corollary 3.60 states that $f([a,b])=[\alpha,\beta]$ for some $\alpha,\beta\in\mathbb{R}$. If $\alpha=\beta$, then $f(x)=\alpha$ for all $x\in(a,b)$ and $f'(x)=0$ for all $x\in(a,b)$. Else, $\beta>\alpha$, and at most one of the bounds can be attained by $f(a)=f(b)$. Hence, either exists $x_0\in(a,b)$ with $f(x_0)=\alpha$ or $f(x_0)=\beta$. In either cases, the Interior Extremum Theorem implies $f'(x_0)=0$. $\square$

Theorem 4.11. Mean value theorem. *Let $b>a$, and $f:[a,b]\longrightarrow\mathbb{R}$ be continuous and differentiable on (a,b). Then exists $x_0\in(a,b)$ such that $f'(x_0)=\frac{f(b)-f(a)}{b-a}$.*

Proof. This follows from applying Rolle's theorem to the function $F(x) = f(x) - \frac{f(b)-f(a)}{b-a}(x-a)$. (You have to check whether F satisfies the hypothesis in Rolle's theorem and then see that, indeed, the x_0 supplied by Rolle satisfies $f'(x_0) = \frac{f(b)-f(a)}{b-a}$) ☐

Theorem 4.12. Generalized Mean Value Theorem. *Let $b > a$, and $f, g : [a, b] \longrightarrow \mathbb{R}$ be continuous and differentiable on (a, b). Then exists $c \in (a, b)$ such that $(g(b) - g(a))f'(c) = (f(b) - f(a))g'(c)$.*

Proof. Apply Rolle's theorem to $h(x) = (g(b) - g(a))f(x) = (f(b) - f(a))g(x)$, $x \in [a, b]$. ☐

We have seen that not all functions which are differentiable on an open interval have continuous derivatives. Nevertheless, they do not have so-called "jump–discontinuities":

Theorem 4.13. Darboux's theorem. *Let $f : (a, b) \longrightarrow \mathbb{R}$ be differentiable. Then the function $f' : (a, b) \longrightarrow \mathbb{R}$ has the intermediate value property, that is, for $u, v \in (a, b)$ and $\xi \in \mathbb{R}$ with $f'(u) < \xi < f'(v)$ exists $c \in (\min\{u, v\}, \max\{u, v\})$ with $f'(c) = \xi$.*

Definition 4.14. A function $f : A \longrightarrow \mathbb{R}$ is

i. *monotonically increasing*, or simply *increasing*, if $f(x) \leq f(y)$ for all $x, y \in A$, with $x < y$

ii. *strictly monotonically increasing*, or simply *strictly increasing*, if $f(x) < f(y)$ for all $x, y \in A$, with $x < y$

iii. *monotonically decreasing*, or simply *decreasing*, if $f(x) \geq f(y)$ for all $x, y \in A$, with $x < y$, and

iv. *strictly monotonically decreasing*, or simply *strictly decreasing*, if $f(x) > f(y)$ for all $x, y \in A$, with $x < y$.

A function is called *monotone* if it is either monotonically increasing or decreasing, and *strictly monotone* if it is either strictly increasing or strictly decreasing.

Theorem 4.15. *Let $f : (a, b) \longrightarrow \mathbb{R}$ be differentiable. Then f is*

i. monotonically increasing if and only if $f'(x) \geq 0$ for all $x \in (a, b)$, and

ii. monotonically decreasing if and only if $f'(x) \leq 0$ for all $x \in (a, b)$.

Proof. We shall show i, ii then follows from applying i to $\widetilde{f}(x) = -f(x)$.

If f is monotonically increasing, then $\frac{f(x)-f(x_0)}{x-x_0} \geq 0$ for all x, x_0 in (a, b) and $f'(x) \geq 0$ for all $x \in (a, b)$ follows.

The Mean Value Theorem implies that If $f'(x_0) < 0$ for some $x_0 \in (a, b)$, then exists $\alpha < \beta$ with $\frac{f(\beta)-f(\alpha)}{\beta-\alpha} = f'(x_0) < 0$, so f is not monotonically increasing. ☐

Example 4.16. Note that $f(x) = x^3$ is strictly increasing on $\mathbb{R}$ but $f'(0) = 0$.

Theorem 4.17. *Let $f : [a,b] \longrightarrow \mathbb{R}$ be continuous and strictly monotone. Let $[c,d] = f([a,b])$ and $g : [c,d] \to \mathbb{R}$ be the inverse function of f.*

i. The inverse function g is continuous on $[c,d]$.

ii. If f is differentiable at $x_0 \in (a,b)$ with $f'(x_0) \neq 0$, then g is differentiable at $y_0 = f(x_0) \in (c,d)$ and $g'(y_0) = \dfrac{1}{f'(x_0)} = \dfrac{1}{f'(g(y_0))}$.

Proof. Let us consider the case that f is strictly increasing.

First, we show that continuity of f at $x_0 \in (a,b)$ implies that g is continuous at $y_0 = f(x_0)$. Indeed, for $\epsilon > 0$ choose $x_1 \in (x_0 - \epsilon, x_0) \cap (a,b)$ and $x_2 \in (x_0, x_0 + \epsilon) \cap (a,b)$, observe that monotonicity implies that for some small δ, we have $B_\delta(y_0) \subseteq (f(x_1), f(x_2))$. But then, $g(B_\delta(y_0)) \subseteq g((f(x_1), f(x_2)) = (x_1, x_2) \subseteq B_\epsilon(x_0)$.

Choose $y_n \neq y_0$ with $y_n \to y_0$. Then $x_n = g(y_n) \to g(y_0) = x_0$ with $x_n \neq x_0$ and

$$\frac{g(y_n) - g(y_0)}{y_n - y_0} = \frac{1}{\frac{f(x_n) - f(x_0)}{x_n - x_0}} \longrightarrow \frac{1}{f'(x_0)} = \frac{1}{f'(g(y_0))}. \quad \square$$

Example 4.18. The function $\exp(x) = \sum_{k=0}^{\infty} \frac{x^k}{k!}$ is differentiable, hence, continuous on $\mathbb{R}$ and clearly strictly increasing on $[1, \infty]$. Then, $\exp(-x) = \exp(x)^{-1}$ shows that exp is strictly monotone on all of $\mathbb{R}$. As $\lim_{x \to \infty} \exp(x) = \infty$ (see Definition 4.19 below) we also obtain $\lim_{x \to -\infty} \exp(x) = \lim_{x \to \infty} \exp(x)^{-1} = 0$ and $\exp : \mathbb{R} \longrightarrow \mathbb{R}^+$ is bijective. The inverse function of $\exp(x)$ is called *natural logarithm* and is denoted by $\ln : \mathbb{R}^+ \longrightarrow \mathbb{R}$. Theorem 4.17 implies that

$$\ln'(x) = \frac{1}{\exp'(\ln(x))} = \frac{1}{\exp(\ln(x))} = \frac{1}{x}.$$

Definition 4.19. Infinite limits and limits at infinity. Let $f : A \longrightarrow \mathbb{R}$, $A \subseteq \mathbb{R}$ and let $x_0, L \in \mathbb{R}^* = \mathbb{R} \cup \{+\infty, -\infty\}$. For $\epsilon > 0$, we call $(\frac{1}{\epsilon}, \infty)$ an ϵ-neighborhood of ∞ and $(-\infty, -\frac{1}{\epsilon})$ an ϵ-neighborhood of $-\infty$.

Further, we say that $f(x) \to L$ as $x \to a$ or $f(x)$ approaches L as x approaches x_0, if for all $\epsilon > 0$ exists a $\delta > 0$ with

$$\left.\begin{array}{ll} x_0 \in A' \subset \mathbb{R}: & 0 < |x - x_0| < \delta \\ \text{or } x_0 = \infty: & x_0 > \frac{1}{\delta} \\ \text{or } x_0 = -\infty: & x_0 < -\frac{1}{\delta} \end{array}\right\} \text{ with } x \in A \text{ implies } \left\{\begin{array}{ll} f(x) \in B_\epsilon(x_0), & \text{if } L \in \mathbb{R}; \\ f(x) \in (\frac{1}{\epsilon}, \infty) & \text{if } L = \infty; \\ f(x) \in (-\infty, -\frac{1}{\epsilon}), & \text{if } L = -\infty. \end{array}\right.$$

Theorem 4.20. L'Hospital's rule *Suppose that f and g are real valued differentiable functions defined on (a,b) where $a \in \mathbb{R} \cup \{-\infty\}$ and $b \in \mathbb{R} \cup \{\infty\}$, $g'(x) \neq 0$ on (a,b), and $\dfrac{f'(x)}{g'(x)} \to L \in \mathbb{R}^*$ as $x \to a$.*

If $f(x) \to 0$ and $g(x) \to 0$ as $x \to a$, or if $g(x) \to \infty$ as $x \to a$, **then** *$\dfrac{f(x)}{g(x)} \to L \in \mathbb{R}^*$ as $x \to a$.*

An analogous statement holds of course if $x \to b$ or if $g(x) \to -\infty$.

Proof. The proof of this result is a bit technical. The main idea is the following. The generalized mean value theorem (Theorem 4.12) implies that for any x, y exists $c_{x,y}$ between x and y such that $\frac{f(x)-f(y)}{g(x)-g(y)} = \frac{f'(c_{x,y})}{g'(c_{x,y})}$ As $x, y \to a$, automatically $c_{x,y} \to a$. Now, letting $y \to a$ faster than x, we can use $f(y) \to 0$ and $g(y) \to 0$ to observe that then $\frac{f(x)-f(y)}{g(x)-g(y)}$ behaves as $\frac{f(x)}{g(x)}$. □

4.2. Taylor series

Definition 4.21. HIGHER DERIVATIVES. For $r \in \mathbb{N}$ we say that $f : U \longrightarrow \mathbb{R}$, U open in $\mathbb{R}$, has an *n-th derivative* at x_0 if $f^{(0)} = f$, $f^{(1)} = f'$, $f^{(2)} = f''$, $\ldots$, $f^{(n-1)} = f^{(n-2)\,\prime}$ are defined on $(x_0 - \epsilon, x_0 + \epsilon)$ for some $\epsilon > 0$ and $f^{(n-1)}$ is differentiable at x_0.

If f has an n-th derivative on U, that is, f has an n-th derivative at x_0 for all $x_0 \in U$, and if $f^{(n)} = f^{(n-1)\,\prime}$ is continuous on U, then we write $f \in C^n(U)$. If $f \in C^n(U)$ for all $n \in \mathbb{N}$, then we write $f \in C^\infty(U)$ and say f is called *smooth*.

Certainly, we shall also write $C^n(A)$ or $C^\infty(A)$ if the requirement on A to be open is weakend to the property that all its members are cluster points, that is, $A \subseteq A'$. For example, one frequently considers $C^2([0,1])$.

Remark 4.22. Note that the notation described above is in accordance to the symbol $C^0(U) = C(U)$ of continuous functions on U.

If U is an interval, for example $U = (a, b)$ we shall write $C^n(a, b)$ rather than $C^n\big((a, b)\big)$.

Theorem 4.23. TAYLOR'S THEOREM. *Given $f : (a, b) \longrightarrow \mathbb{R}$ and $n \in \mathbb{N}$ with $f \in C^{n-1}(a, b)$ and $f^{(n)}$ defined (but not necessarily continuous) on (a, b). For x_0 in (a, b) define the $n-1$-th degree* Taylor polynomial *as*

$$P_{f,x_0}(x) = \sum_{k=0}^{n-1} \frac{f^{(k)}(x_0)}{k!}(x - x_0)^k, \quad x \in (a, b).$$

For any $x \in (a, b)$ exists a ξ_x between x_0 and x such that

$$f(x) = P_{f,x_0}(x) + \frac{f^{(n)}(\xi_x)}{n!}(x - x_0)^n.$$

Proof. Let x and x_0 be fixed. Choose $M \in \mathbb{R}$ with $f(x) - P_{f,x_0}(x) = M(x - x_0)^n$. We have to show that for some ξ_x between x_0 and x we have $\frac{f^{(n)}(\xi_x)}{n!} = M$.

□

Remark 4.24. Taylor's Theorem is used to compute approximate values of functions by means of evaluating polynomials.

For example, if $|f^{(n)}(\xi)| < M$ for all ξ between x and x_0, then we have

$$|f(x) - P_{f,x_0}(x)| = \left| \frac{f^{(n)}(\xi_x)}{n!}(x - x_0)^n \right| \leq \frac{M}{n!}|x - x_0|^n$$

For x being close to x_0 the right hand side, and, therefore, the approximation error, is small.

Corollary 4.25. *If $f \in C^n(a,b)$ with $f^{(n)}(\xi) = 0$ for all $\xi \in (a,b)$, then f is a polynomial of degree at most $(n-1)$.*

Corollary 4.26. *If for $f \in C^\infty(a,b)$ there exists $M > 0$ with $|f^{(n)}(\xi)| \le M$ for all $\xi \in (a,b)$ and $n \in \mathbb{N}$, then for any $x_0 \in (a,b)$, we have*

$$f(x) = \sum_{k=0}^{\infty} \frac{f^{(k)}(x_0)}{k!}(x-x_0)^k, \quad x \in (a,b).$$

Definition 4.27. For $f \in C^\infty(a,b)$ and $x_0 \in (a,b)$, the formal power series

$$T_{f,x_0}(x) = \sum_{k=0}^{\infty} \frac{f^{(k)}(x_0)}{k!}(x-x_0)^k, \quad x \in (a,b)$$

is the *Taylor series* of f at x_0.

Remark 4.28.

i. The radius of convergence of a Taylor series is not necessarily larger than 0.

ii. Even if the Taylor series of a function converges, it might not converge to the function. For example, consider $f(x) = \begin{cases} e^{-\frac{1}{x}^2}, & \text{for } x \neq 0 \\ 0, & \text{else.} \end{cases}$ satisfies $f \in C^\infty(\mathbb{R})$, $f^{(n)}(0) = 0$ for $n \in \mathbb{N}$ and, therefore, $T_{f,0}$ has radius of convergence $R = \infty$ and $T_{f,0}(x) = 0 \neq f(x)$ for $x \neq 0$.

Theorem 4.29. *Assume that (f_n) is a sequence of functions which are differentiable on (c,d), and let $[a,b] \subset (c,d)$. If $\sum_{n=1}^{\infty} f_n(x)$ converges at some $x_0 \in [a,b]$ and $\sum_{n=1}^{\infty} f_n'(x)$ converges uniformly on $[a,b]$, then $\sum_{n=1}^{\infty} f_n(x)$ converges to a differentiable function, and*

$$\left(\sum_{n=1}^{\infty} f_n(x)\right)' = \sum_{n=1}^{\infty} f_n'(x).$$

Proof. Use Theorem 3.70. □

Proposition 4.30. *If $f(x) = \sum_{k=0}^{\infty} c_k(x-x_0)^k$ for $x \in (a,b)$, then $f \in C^\infty(a,b)$ and $f^{(k)}(x) = c_k\, k!$ for $k \in \mathbb{N}$. Further, we have $f'(x) = \sum_{k=1}^{\infty} c_k\, k(x-x_0)^{k-1}$ for $x \in (a,b)$, that is, we can differentiate the series of functions f term by term.*

Proof. Use Theorem 4.29. □

4.3. The exponential function and friends

The following theorem lists important facts regarding the exponential function $\exp(z) = \sum_{k=0}^{\infty} \frac{z^k}{k!}$, $z \in \mathbb{C}$, some of which we stated and proved earlier.

Theorem 4.31. The Exponential Function.

i. $\exp(z) = \sum_{k=0}^{\infty} \frac{z^k}{k!}$ *converges absolutely for* $z \in \mathbb{C}$.

ii. $\exp(z+w) = \exp(z)\exp(w)$ *for* $z, w \in \mathbb{C}$.

iii. $\exp(x) = \exp(1)^x = e^x$ *for* $x \in \mathbb{R}$.

iv. $\exp'(x) = \exp(x)$ *for* $x \in \mathbb{R}$.

v. $\exp(x) > 0$ *for* $x \in \mathbb{R}$ *and* exp *is strictly monotonically increasing.*

vi. $\exp(x) \to \infty$ *as* $x \to \infty$ *and* $\exp(x) \to 0$ *as* $x \to -\infty$.

vii. $\exp : \mathbb{R} \to \mathbb{R}^+$ *is bijective.*

viii. $\frac{x^n}{\exp(x)} \to 0$ *as* $x \to \infty$ *for all* $n \in \mathbb{N}$.

Definition 4.32. The inverse function of $\exp : \mathbb{R} \longrightarrow \mathbb{R}^+$ is called natural logarithm and is denoted by $\log : \mathbb{R}^+ \longrightarrow \mathbb{R}$.

Proposition 4.33. *i.* $\log(xy) = \log(x) + \log(y)$ *for* $x, y \in \mathbb{R}^+$.

ii. The natural logarithm is a differentiable function with $\log'(x) = \frac{1}{x}$ *for* $x \in \mathbb{R}^+$.

iii. For $x > 0$ *we have* $x^a = \exp(a\log(x)) = e^{a\log(x)}$ *and* $f : \mathbb{R}^+ \longrightarrow \mathbb{R}$, $x \mapsto x^a$ *is differentiable with* $f'(x) = ax^{a-1}$.

iv. For $a > 0$ *we have again* $a^x = \exp(x\log(a)) = e^{x\log(a)}$ *and* $g : \mathbb{R} \longrightarrow \mathbb{R}$, $x \mapsto a^x$ *is differentiable with* $g'(x) = a^x\log(a)$.

Proof. ii. Use Theorem 4.17, iii. and iv. by chain rule. □

Definition 4.34. For $a > 0$, the function of $g(x) : \mathbb{R} \longrightarrow \mathbb{R}^+$, $x \mapsto a^x$ is bijective and its inverse is called logarithm to base a. We shall denote g^{-1} by $\log_a : \mathbb{R}^+ \longrightarrow \mathbb{R}$.

After discussing the behavior of the restriction of the function $\exp : \mathbb{C} \longrightarrow \mathbb{C}$ to the real axis $\mathbb{R}$, that is, $\exp : \mathbb{R} \longrightarrow \mathbb{C}$, we shall now consider its restriction to the imaginary axis $i\mathbb{R} \subset \mathbb{C}$. Once we described its properties, we fully understand $\exp : \mathbb{C} \longrightarrow \mathbb{C}$ since $\exp(a+bi) = \exp(a)\exp(bi)$ for $a, b \in \mathbb{R}$.

We shall study $\exp : i\mathbb{R} \longrightarrow \mathbb{C}$ by studying its real and imaginary part.

Definition 4.35. We define the *sine* function $\sin : \mathbb{R} \longrightarrow \mathbb{R}$ by setting $\sin(x) = \Im \exp(ix)$ for $x \in \mathbb{R}$ and the *cosine* function $\cos : \mathbb{R} \longrightarrow \mathbb{R}$ by setting $\cos(x) = \Re \exp(ix)$ for $x \in \mathbb{R}$.

For convenience, we shall write $\cos x$ for $\cos(x)$, $\sin x$ for $\sin(x)$, $\cos^n x$ for $(\cos(x))^n$, and $\sin^n x$ for $(\sin(x))^n$, for $x \in \mathbb{R}$ and $n \in \mathbb{N}$.

Theorem 4.36. *i.* $\sin x = \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k+1}}{(2k+1)!}$ *for* $x \in \mathbb{R}$.

ii. $\cos x = \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k}}{(2k)!}$ *for* $x \in \mathbb{R}$.

iii. $\sin' = \cos$ *and* $\cos' = -\sin$.

iv. $\sin^2 x + \cos^2 x = 1$ *for* $x \in \mathbb{R}$.

v. $\cos$ *and* $\sin$ *are* 2π*–periodic, that is,* $\sin(x + 2\pi) = \sin x$, $\cos(x + 2\pi) = \cos x$, *where* $\frac{\pi}{2}$ *is the smallest* $x > 0$ *such that* $\cos x = 0$.

Corollary 4.37. $\exp : \mathbb{C} \longrightarrow \mathbb{C}$ *is* $2\pi i$*–periodic.*

Proof. $\exp(z + 2\pi i) = \exp(z) \exp(2\pi i) = \exp(z)(\cos(2\pi) + i \sin(2\pi)) = \exp(z)$ for $z \in \mathbb{C}$. □

4.4. Fixed point theorems and approximative methods

Definition 4.38. An element $x_0 \in X$ is called *fixed point* of $f : X \longrightarrow X$ if $f(x_0) = x_0$.

Definition 4.39. A *contraction* on a metric space (X, d) is a map $f : X \longrightarrow X$ such that for some constant k, $0 \leq \theta < 1$, we have

$$d(f(x), f(y)) \leq \theta\, d(x, y) \text{ for all } x, y \in X\,.$$

Proposition 4.40. *Contractions are uniformly continuous mappings.*

Theorem 4.41. Banach Fixed Point Theorem. *If* $f : X \longrightarrow X$ *is a contraction on a complete metric space* (X, d)*, then exists a unique fixed point* $x^* \in X$*, and for any choice of* $x_0 \in X$*, the sequence* (x_n) *defined by*

$$x_0, x_1 = f(x_0), x_2 = f(x_1), x_3 = f(x_2) = f(f(x_1)) = f \circ f(x_1), \ldots, x_{n+1} = f(x_n), \ldots\,,$$

converges to x^**. Moreover, we have*

$$d(x_n, x^*) \leq \frac{\theta}{1-\theta}\, d(x_n, x_{n-1}) \leq \frac{\theta^{n-1}}{1-\theta}\, d(x_1, x_0).$$

Proof. First note that if x^* and $x^\star$ are both fixed points, then

$$0 \leq d(x^*, x^\star) = d(f(x^*), f(x^\star)) \leq \theta d(x^*, x^\star),$$

and $0 \leq \theta < 1$ implies $d(x^*, x^\star) = 0$ so $x^* = x^\star$ and uniqueness is shown.

To show that the constructed sequence (x_n) converges, we shall establish that it is Cauchy. Indeed, with $m \geq n$, we have

$$\begin{aligned} d(x_m, x_n) &= d(f(x_{m-1}), f(x_{n-1})) \leq \theta d(x_{m-1}, x_{n-1}) \leq \ldots \leq \theta^n d(x_{m-n}, x_0) \\ &\leq \theta^n \big(d(x_{m-n}, x_{m-n-1}) + d(x_{m-n-1}, x_{m-n-2}) + \ldots + d(x_2, x_1) + d(x_1, x_0)\big) \\ &\leq \theta^n \big(\theta^{m-n} d(x_1, x_0) + \theta^{m-n} d(x_{m-n-2}, x_{m-n-3}) + \ldots + \theta d(x_1, x_0) + d(x_1, x_0)\big) \\ &= \theta^n d(x_1, x_0) \sum_{k=0}^{m-n} \theta^k \leq \theta^n d(x_1, x_0) \frac{1}{1-\theta}, \end{aligned}$$

which implies that (x_n) is Cauchy since for fixed ϵ just choose N so that $\theta^N d(x_1, x_0) < (1-\theta)\epsilon$.

Completeness of X now ensures that (x_n) converges to some x^*. Letting $m \to \infty$ above shows that $d(x_n, x_*) \leq \frac{\theta^n}{1-\theta} d(x_1, x_0)$. Moreover, a contradiction is easily seen to be continuous. Hence,

$$f(x^*) = f(\lim_{n\to\infty} x_n) = \lim_{n\to\infty} f(x_n) = \lim_{n\to\infty} x_{n+1} = x^*,$$

and x^* is a fixed point.

□

Theorem 4.42. Newton's Method.[5] *Let f be continuous on $[a, b]$ and twice differentiable on (a, b). If $f(a) \leq 0$, $f(b) > 0$, $f'(x) \geq \delta > 0$ and $0 \leq f''(x) \leq M$ for $x \in (a, b)$, then exists a unique point $\xi \in (a, b)$ with $f(\xi) = 0$. Moreover, for any x_1 with $f(x_1) > 0$ the sequence recursively defined by*

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}$$

converges to ξ and we have

$$|x_{n+1} - \xi| \leq \frac{2\delta}{M} \left| \frac{M}{2\delta}(x_1 - \xi) \right|^{2n}.$$

Theorem 4.43. *Let $f : [a, b] \to \mathbb{R}$, $a < b$, be a differentiable function with $f([a, b]) \subset [a, b]$ and let $q < 1$ such that $|f'(x)| \leq q, \forall x \in D$. For $x_1 \in [a, b]$ set $x_n = f(x_{n-1})$ for $n \geq 1$. Then the sequence (x_n) converges to the unique solution $\xi \in D$ of the equation $f(\xi) = \xi$ and the following inequalities holds:*

$$|\xi - x_n| \leq \frac{q}{1-q} |x_n - x_{n-1}| \leq \frac{q^n}{1-q} |x_1 - x_0|.$$

[5] We shall only give one of the many cases/versions of Newton's method.

5. INTEGRATION

5.1. The Riemann integral

Definition 5.1. A *partition* P of the closed interval $[a,b] \subset \mathbb{R}$, $a < b$, is a finite set $P = \{x_0, x_1, \dots, x_N\}$ such that $a = x_0 < x_1 < x_2 < \dots < x_{N-1} < x_N = b$. The *mesh* or *width* of P is given by $\operatorname{mesh} P = \max\{x_n - x_{n-1} : n = 1 : N\}$. A *sampling set* T associated to P is a set $T = \{t_1, t_2, \dots, t_N\}$ with $a = x_0 \le t_1 \le x_1 \le t_2 \le x_2 \le \dots \le x_{N-1} \le t_N \le x_N = b$.

The *Riemann sum* $R(f,P,T)$ of $f : [a,b] \longrightarrow \mathbb{R}$ corresponding to P,T is

$$R(f,P,T) = \sum_{n=1}^{N} f(t_n)\,(x_n - x_{n-1}).$$

Definition 5.2. A function $f : [a,b] \longrightarrow \mathbb{R}$ is *Riemann integrable* if there exists a real number I_f such that for all $\varepsilon > 0$ exists a $\delta > 0$ with the property that for any partition P with $\operatorname{mesh} P < \delta$ and any corresponding sampling set T we have $|R(f,P,T) - I_f| < \epsilon$.

In this case, we call the number I_f *Riemann integral* of f on $[a,b]$ and we write

$$\int_a^b f(x)\,dx = I_f.$$

The set of all Riemann integrable functions on $[a,b]$ is denoted by $\mathcal{R}([a,b])$.

Theorem 5.3. *Any Riemann integrable function is bounded.*

Proof. Let us assume that $f : [a,b] \longrightarrow \mathbb{R}$ is unbounded and Riemann integrable. For $\epsilon = 1$ choose $\delta > 0$ so that $|R(f,P,T) - I_f| < 1$ for all partitions P with $\operatorname{mesh} P < \delta$ and any corresponding sampling set T. Fix such $P = \{x_0, x_1, \dots, x_N\}$ with $\operatorname{mesh} P < \delta$ and $T = \{t_1, t_2, \dots, t_N\}$.

Since f is unbounded on $[a,b]$ there exists n such that f is unbounded on $[x_{n-1}, x_n]$. However small $x_n - x_{n-1}$ is, we can find $s_n \in [x_{n-1}, x_n]$ with $|f(s_n)\,(x_n - x_{n-1}) - f(t_n)\,(x_n - x_{n-1})| > 1000$. We set $T' = \{t_1, \dots, t_{n-1}, s_n, t_{n+1}, \dots t_N\ )\}$ and conclude that

$$1000 < |R(f,P,T) - R(f,P,T')| \le |R(f,P,T) - I_f| + |I_f - R(f,P,T')| < 1 + 1 = 2,$$

a contradiction. □

Theorem 5.4. Linearity of the Riemann integral. *The set $\mathcal{R}([a,b])$ of Riemann integrable functions on $[a,b]$ is a real vector space and the map*

$$\int_a^b : \mathcal{R}([a,b]) \longrightarrow \mathbb{R} \quad f \mapsto \int_a^b f(x)\,dx$$

is linear.

Theorem 5.5. Monotonicity of the integral. *If $f,g \in \mathcal{R}([a,b])$ and $f(x) \le g(x)$ for all $x \in \mathbb{R}$, then $\int_a^b f(x)\,dx \le \int_a^b g(x)\,dx$.*

Definition 5.6. For a bounded function $f : [a, b] \longrightarrow \mathbb{R}$ and a partition $P = \{x_0, x_1, \dots x_N\}$, we call

$$L(f, P) = \sum_{n=1}^{N} m_n(x_n - x_{n-1}), \quad m_n = \inf\{f(x) : \ x \in [x_{n-1}, x_n]\} = \inf\ f([x_{n-1}, x_n])$$
$$\text{for } n = 1, \dots N$$

lower sum of f with respect to P, and

$$U(f, P) = \sum_{n=1}^{N} M_n(x_n - x_{n-1}), \quad M_n = \sup\ f([x_{n-1}, x_n]) \quad \text{for } n = 1, \dots N$$

upper sum of f with respect to P.

Definition 5.7. The *lower [resp. upper] integral* of a bounded function $f : [a, b] \longrightarrow \mathbb{R}$ is $\underline{I}(f) = \sup_P L(f, P)$ [resp. $\overline{I}(f) = \inf_P U(f, P)$.

If $\underline{I}(f) = \overline{I}(f)$, then we call f *Darboux integrable* on $[a, b]$ and $I(f) = \underline{I}(f) = \overline{I}(f)$ the *Darboux integral* of f on $[a, b]$.

Definition 5.8. A partition P' of $[a, b]$ *refines* the partition P of $[a, b]$ if $P' \supset P$. P' is called *refinement* of P.

A partition P' of $[a, b]$ which refines simultaneously two partitions P_1 and P_2 is called *common refinement* of P_1 and P_2.

Lemma 5.9. Refinement Principle. *If P' refines P on $[a, b]$ and $f : [a, b] \longrightarrow \mathbb{R}$ is bounded, then $L(f, P) \le L(f, P') \le U(f, P') \le U(f, P)$.*

Lemma 5.10. *A bounded function $f : [a, b] \longrightarrow \mathbb{R}$ is Darboux integrable if and only if*

for all $\varepsilon > 0$ exists a partition P_ε such that $U(f, P_\varepsilon) - L(f, P_\varepsilon) < \varepsilon$.

Theorem 5.11. *A function $f : [a, b] \longrightarrow \mathbb{R}$ is Darboux integrable if and only if f is Riemann inegrable. Further, $\int_a^b f(x)\, dx = I(f)$.*

Proof. Let us first assume f is Riemann integrable. Fix $\epsilon > 0$ and choose $\delta > 0$ so that $|I_f - R(f, P, T)| < \frac{\epsilon}{3}$ for all P with width less than δ. Fix such P

Note that
$$U(f, P) = \sup_T R(f, P, T) \quad \text{and} \quad L(f, P) = \inf_T R(f, P, T)$$
and, hence $|I_f - U(f, P)| \le \frac{\epsilon}{3}$ and $|I_f - L(f, P)| \le \frac{\epsilon}{3}$, and, therefore

$$|U(f, P) - L(f, P)| \le |I_f - U(f, P)| + |I_f - L(f, P)| \le \frac{\epsilon}{3} + \frac{\epsilon}{3} < \epsilon,$$

so $P_\epsilon = P$ does the job.

Now assume f is Darboux integrable. Fix $\epsilon > 0$ and choose $P_\epsilon = \{x_0, \ldots, x_N\}$ so that $U(f, P_\epsilon) - L(f, P_\epsilon) < \frac{\epsilon}{2}$. Let M be chosen with $M \geq |f(x)|$ for all $x \in [a, b]$ and set $\delta = \frac{\epsilon}{4MN}$.

Let $P = \{y_0, y_1, \ldots, y_{\widetilde{N}}\}$ be a partition of width δ and T a respective sampling set. Note that

$$\sum_{n=1}^{N} m_n \chi_{[x_{n-1} - x_n]})(x) \leq \sum_{n=1}^{\widetilde{N}} f(t_n)\chi_{[y_{n-1}-y_n]})(x) \leq \sum_{n=1}^{N} M_n \chi_{[x_{n-1}-x_n]})(x), \quad x \in A,$$

where A contains all points in $[a, b]$ with the possible exception of the union of N intervals of size less than or equal to δ. Hence, with $I(f)$ denoting the Darboux integral, we have

$$I(f) - R(f, P, T) < I(f) - L(f, P_\epsilon) + 2MN\delta < \frac{\epsilon}{2} + \frac{\epsilon}{2}$$

and

$$R(f, P, T) - I(f) < U(f, P_\epsilon) - I(f) + 2MN\delta < \frac{\epsilon}{2} + \frac{\epsilon}{2}.$$

□

Theorem 5.12. Riemann Integrability Criterion. *A bounded function is Riemann integrable if and only if*

for all $\varepsilon > 0$ exists a partition P_ε such that $U(f, P_\varepsilon) - L(f, P_\varepsilon) < \varepsilon$.

Corollary 5.13. *$C([a, b]) \subset \mathcal{R}([a, b])$, that is, continuous functions are Riemann integrable.*

Lemma 5.14. *For $x \in \mathbb{R} \setminus \{\ldots, -4\pi, -2\pi, 0, 2\pi, 4\pi, \ldots\}$ we have* $\frac{1}{2} + \sum_{n=1}^{N} \cos nx = \frac{\sin(N + \frac{1}{2})x}{2 \sin \frac{1}{2}x}$.

Example 5.15. An application of Lemma 5.14 shows that for $a < 0$ we have $\int_a^0 \cos x \, dx = \sin a$.

Theorem 5.16. *Let $a < b < c$ and $f : [a, c] \longrightarrow \mathbb{R}$. Then $f \in \mathcal{R}([a, c])$ if and only if $f|_{[a,b]} \in \mathcal{R}([a, b])$ and $f|_{[b,c]} \in \mathcal{R}([b, c])$. If these conditions are satisfied, we have*

$$\int_a^c f(x)\, dx = \int_a^b f(x)\, dx + \int_b^c f(x)\, dx.$$

In light of Theorem 5.16 we shall now extend the definition of Riemann integral.

Definition 5.17. For $a < b$ and $f \in \mathcal{R}([a, b])$, we set $\int_a^a f(x)\, dx = 0$ and $\int_b^a f(x)\, dx = -\int_a^b f(x)\, dx$.

The conscientious reader should check which ones of the properties of the Riemann integral depend on $a < b$ in $\int_a^b f(x)\, dx$. For example, linearity does not, while monotonicity does.

Definition 5.18. A set $Z \subset \mathbb{R}$ is called *zero set* if for all $\varepsilon > 0$ exists a countable covering of open intervals (a_n, b_n), $n \in \mathbb{N}$, that is, $\bigcup_{n=1}^{\infty}(a_n, b_n)$, with $\sum_{n=1}^{\infty}(b_n - a_n) < \varepsilon$.

Example 5.19.

i. Finite sets are zero sets.

ii. $\mathbb{Q}$ is a zero set.

iii. Subsets of zero sets are zero sets.

iv. if Z_r, $r \in \mathbb{N}$ are zero sets, then $Z = \bigcup_{r=1}^{\infty} Z_r$ is a zero set.

Definition 5.20. Let $f : [a, b] \longrightarrow \mathbb{R}$ and $x_0 \in [a, b]$. The *oscillation* of f at x_0 is given by

$$\operatorname{osc}_{x_0}(f) = \limsup_{x \to x_0} f(x) - \liminf_{x \to x_0} f(x) = \lim_{n \to \infty} \big(\sup f([x_0 - \tfrac{1}{n}, x_0 + \tfrac{1}{n}]) - \inf f([x_0 - \tfrac{1}{n}, x_0 + \tfrac{1}{n}]) \big)$$

with the obvious adjustments if $x_0 = a$ or $x_0 = b$.

Lemma 5.21. *The function $f : [a, b] \longrightarrow \mathbb{R}$ is continuous at $x_0 \in [a, b]$ if and only if* $\operatorname{osc}_{x_0}(f) = 0$.

Theorem 5.22. Riemann–Lebesgue Theorem. *A function $f : [a, b] \longrightarrow \mathbb{R}$ is Riemann integrable if and only if f is bounded and its set of discontinuities*

$$D = \{x \in [a, b] : \ f \text{ is not continuous at } x\}$$

is a zero set.

Corollary 5.23. *The product of Riemann integrable functions is Riemann integrable, that is, if $f, g \in \mathcal{R}([a, b])$, then $fg \in \mathcal{R}([a, b])$.*

Corollary 5.24. *If $f : [a, b] \longrightarrow [c, d]$ is Riemann integrable and $\Phi : [c, d] \longrightarrow \mathbb{R}$ is continuous, then $\Phi \circ f \in \mathcal{R}([a, b])$.*

Corollary 5.25. *If $f \in \mathcal{R}([a, b])$, then $|f| \in \mathcal{R}([a, b])$.*

Theorem 5.26. *Every monotone function $f : [a, b] \longrightarrow \mathbb{R}$ is Riemann integrable.*

Theorem 5.27. Mean Value Theorem For Integrals. *Let $f, \varphi : [a, b] \longrightarrow \mathbb{R}$ be continuous functions and $\varphi \geq 0$. Then there exists $\xi \in [a, b]$ such that*

$$\int_a^b f(x)\varphi(x)dx = f(\xi) \int_a^b \varphi(x)dx.$$

Theorem 5.28. *Let $(f_n)_{n\in\mathbb{N}}$ be a sequence of Riemann integrable functions on $[a, b]$. If $f_n \longrightarrow f$ uniformly on $[a, b]$, then f is Riemann integrable and*

$$\lim_{n\to\infty} \int_a^b f_n(x)\, dx = \int_a^b f(x)\, dx.$$

5.2. Integration and differentiation

Definition 5.29. For $f \in \mathcal{R}([a,b])$ we call the function $F(x) = \displaystyle\int_a^x f(y)\,dy$, $x \in [a,b]$, the *indefinite integral* of f on $[a,b]$.

Theorem 5.30. Fundamental Theorem of Calculus I. *Let $f \in \mathcal{R}([a,b])$ and $F(x) = \displaystyle\int_a^x f(y)\,dy$. Then $F(x)$ is continuous and $F'(x) = f(x)$ for all $x \in (a,b)$ at which f is continuous.*

Remark 5.31. In calculus, the symbol $\int f(y)\,dy$ is often referred to as the *indefinite integral* of a function f which is continuous on its domain. In fact, in this context, the indefinite integral represents the set of functions satisfying $F'(x) = f(x)$. As $F'(x) = G'(x)$ on connected sets if and only if $F(x) = G(x) + C$ for $C \in \mathbb{R}$, one commonly writes $\int f(y)\,dy = F(x) + C$. For example, $\int \sin(y)\,dy = \cos(x) + C$, or, abusing notation even more, $\int \sin(x)\,dx = \cos(x) + C$.

Theorem 5.32. Fundamental Theorem of Calculus II. *If $f \in \mathcal{R}([a,b])$ and $F \in C([a,b])$ is given with $F'(x) = f(x)$ for all $x \in (a,b)$, then $\displaystyle\int_a^b f(x)\,dx = F(b) - F(a)$.*

Theorem 5.33. Integration by Parts. *Suppose $F, G : [a,b] \longrightarrow \mathbb{R}$ are continuous and differentiable on (a,b), $f, g \in \mathcal{R}([a,b])$ with $F'(x) = f(x)$ and $G'(x) = g(x)$ on (a,b). Then*

$$\int_a^b F(x)g(x)\,dx = F(b)G(b) - F(a)G(a) - \int_a^b f(x)G(x)\,dx.$$

Theorem 5.34. Integration by Substitution. *Let $f \in \mathcal{R}([a,b])$ and assume $g : [c,d] \longrightarrow [a,b]$ bijective and continuously differentiable with $g'(x) > 0$ for $x \in [a,b]$, then*

$$\int_a^b f(y)\,dy = \int_c^d f(g(x))\,g'(x)\,dx.$$

5.3. Improper Riemann integral

Example 5.35. Calculate the length of a circle using "improper" integrals.

Definition 5.36. Let $f : [a, b) \longrightarrow \mathbb{R}$, $b \in \mathbb{R} \cup \{+\infty\}$ be Riemann integrable on $[a, c]$ for any $a < c < b$. If $\lim_{c \nearrow b} \int_a^c f(x)\, dx$ exists[6], then it is called *improper Riemann integral* and we extend our use of notation and write

$$\int_a^b f(x)\, dx = \lim_{c \nearrow b} \int_a^c f(x)\, dx.$$

Under similar conditions on $f : (a, b] \longrightarrow \mathbb{R}$, $a \in \mathbb{R} \cup \{-\infty\}$, we define the *improper integral* by setting

$$\int_a^b f(x)\, dx = \lim_{c \searrow a} \int_c^b f(x)\, dx.$$

If $f : (a, b) \longrightarrow \mathbb{R}$, $b \in \mathbb{R} \cup \{+\infty\}$ is Riemann integrable on $[c_1, c_2]$ for any $a < c_1 < c_2 < b$, and if $\lim_{c_2 \nearrow c} \int_c^{c_2} f(x)\, dx$ and $\lim_{c_1 \searrow a} \int_{c_1}^c f(x)\, dx$ exist for some $c \in (a, b)$, then we set

$$\int_a^b f(x)\, dx = \lim_{c_2 \nearrow b} \int_c^{c_2} f(x)\, dx + \lim_{c_1 \searrow a} \int_{c_1}^c f(x)\, dx.$$

Remark 5.37. $\lim_{R \to \infty} \int_R^{-R} x\, dx$ exists, but, by definition, $\int_{-\infty}^{\infty} x\, dx$ does not exist as improper integral.

Theorem 5.38. Integral Criterion for Sums. *Let $f : [1, \infty) \longrightarrow \mathbb{R}^+ \cup \{0\}$ be a monotonic decreasing function.*

The series $\sum_{n=1}^{\infty} f(n)$ converges if and only if the improper integral $\int_1^{\infty} f(x)\, dx$ exists.

Proposition 5.39. Wallis' Product. $\frac{\pi}{2} = \prod_{n=1}^{\infty} \frac{4n^2}{4n^2 - 1}.$

[6] We write $\lim_{c \nearrow b}$ for $\lim_{\substack{c \to b \\ c < b}}$ and $\lim_{c \searrow a}$ for $\lim_{\substack{c \to a \\ c > a}}$.

5.4. Infinite dimensional vector spaces and orthonormal bases

Definition 5.40. A *vector space* over the field $K = \mathbb{C}$ or $K = \mathbb{R}$ is a set V with an addition $+ : V \times V \to V$ and a *scalar multiplication* $\cdot : K \times V \to V$ which satisfy:

i. $(V, +)$ is a commutative group. The neutral element is denoted by 0 (not to be confused with the scalar $0 \in K$).

ii. For $v, w \in V$ and $r, s \in K$, we have

   - $r \cdot (v + w) = (r \cdot v) + (r \cdot w)$;
   - $(r + s) \cdot v = (r \cdot v) + (s \cdot v)$;
   - $(rs) \cdot v = r \cdot (s \cdot v)$;
   - $1 \cdot v = v$.

   As customary with multiplication in fields, the symbol "$\cdot$" for scalar multiplication is often omitted.

Definition 5.41. A *norm* on the vector space V over $K = \mathbb{R}$ or $K = \mathbb{C}$ is a function $\|\cdot\| : V \to \mathbb{R}$ with

i. $\|v\| > 0$ if $v \neq 0$ and $\|0\| = 0$;

ii. $\|rv\| = |r|\|v\|$ for $r \in K$ and $v \in V$;

iii. $\|v + w\| \leq \|v\| + \|w\|$ for $v, w \in V$.

A vector space with norm is called *normed vector space* .

Remark 5.42. A norm $\| \cdot \|$ on a vector space V over $K = \mathbb{R}$ or $K = \mathbb{C}$ induces the metric $d : V \times V \longrightarrow \mathbb{R}, (v, w) \mapsto d(v, w) = \|v - w\|$ on V.

Definition 5.43. A normed vector space V over $K = \mathbb{R}$ or $K = \mathbb{C}$ which is a complete metric space with respect to the metric $d : V \times V \longrightarrow \mathbb{R}, (v, w) \mapsto d(v, w) = \|v - w\|$ is called *Banach space.*

Definition 5.44. An *inner product* (*scalar product*) on the vector space V over $K = \mathbb{R}$ or $K = \mathbb{C}$ is a binary function $\langle \cdot, \cdot \rangle : V \times V \to K$ which satisfies

i. $\langle v, v \rangle > 0$ if $v \neq 0$;

ii. $\langle u + v, w \rangle = \langle u, w \rangle + \langle v, w \rangle$ for $u, v, w \in V$ and $\langle rv, w \rangle = r\langle v, w \rangle$ for $r \in K$ and $v, w \in V$.

iii. $\langle u, v + w \rangle = \langle u, v \rangle + \langle u, w \rangle$ for $u, v, w \in V$ and $\langle v, rw \rangle = \overline{r}\langle v, w \rangle$ for $r \in K$ and $v, w \in V$.

iv. $\langle v, w \rangle = \overline{\langle w, v \rangle}$ for $v, w \in V$.

A vector space with inner product is called *inner product space.*

Remark 5.45. An inner product $\langle \cdot, \rangle$ on a vector space V over $K = \mathbb{R}$ or $K = \mathbb{C}$ induces the norm $\| \cdot \| : V \to \mathbb{R}, v \mapsto \sqrt{\langle v, v \rangle}$ and therefore a metric $d : V \times V \longrightarrow \mathbb{R}, (v, w) \mapsto d(v, w) = \|v - w\| = \sqrt{\langle v - w, v - w \rangle}$ on V.

Inner product and induced norm satisfy the Cauchy–Schwarz inequality $|\langle v, w \rangle| \leq \|v\| \, \|w\|$ for $v, w \in V$. A special case of this inequality has been given as Theorem 3.7.

Definition 5.46. An inner product vector space V over $K = \mathbb{R}$ or $K = \mathbb{C}$ which is a complete metric space with respect to the metric $d : V \times V \longrightarrow \mathbb{R}, (v, w) \mapsto d(v, w) = \|v - w\| = \sqrt{\langle v - w, v - w \rangle}$ is called *Hilbert space.*

Definition 5.47. A set $B \subset V$ is a *basis* of the vector space V over the field $K = \mathbb{C}$ or $K = \mathbb{R}$ if any $v \in V$ can be represented as a linear combination of the elements in B and if no subset $B' \subsetneqq B$ has this property.

If B is a basis of V consisting of $N \in \mathbb{N}$ elements, then we call N the dimension of V and write $\dim V = N$. If V has not a finite basis, then we call V infinite dimensional.

Definition 5.48. A family of vectors $\mathcal{O}$ in an inner product space is called *orthogonal* , if $\langle v, w \rangle = 0$ for $v, w \in \mathcal{O}$, $v \neq w$. If in addition $\langle v, v \rangle = 1$ for $v \in \mathcal{O}$, then we call $\mathcal{O}$ an *orthonormal system (ONS)*

Remark 5.49. A family $\mathcal{O}$ of orthogonal vectors in any inner product space is linear independent, as $\sum_{n=1}^{N} a_n v_n = 0$, $\{v_n\}$ orthogonal, implies

$$0 = \langle \sum a_n v_n, v_m \rangle = \sum a_n \langle v_n, v_m \rangle = a_m$$

for all $m = 1, \ldots N$. Consequently, any family $\mathcal{O}$ of N orthogonal vectors in an N dimensional vector space is a basis.

Example 5.50. Consider the space K^n of vectors with n entries in $K = \mathbb{R}$ or $K = \mathbb{C}$. Then

- $\langle (v_1, v_2, \ldots, v_n), (w_1, w_2, \ldots, w_n) \rangle = \sum_{k=1}^{n} v_k \overline{w_k}$ defines an inner product on K^n;
- $\|(v_1, v_2, \ldots, v_n)\| = \|(v_1, v_2, \ldots, v_n)\|_2 = \sqrt{\sum_{k=1}^{n} |v_k|^2}$ is the norm induced by $\langle \cdot, \cdot \rangle$;
- $d_2((v_1, v_2, \ldots, v_n), (w_1, w_2, \ldots, w_n)) = \sqrt{\sum_{k=1}^{n} |v_k - w_k|^2}$ is the metric induced by $\| \cdot \|_2$;
- (K^n, d_2) is complete, K^n is therefore a Hilbert space;
- $\{e_1 = (1, 0, 0, \ldots, 0, 0), e_2 = (0, 1, 0, \ldots, 0), \ldots, e_{n-1} = (0, 0, 0, \ldots, 1, 0), e_n = (0, 0, 0, \ldots, 0, 1)\}$ is a basis of K^n and an ONS in K^n;
- K^n has dimension n.

As K^n can be considered as functions mapping the finite index set $\{1, 2, \ldots, n\}$ into K, it is natural to consider as vector space sets of functions mapping infinite sets S into K. The following example discusses the case $S = \mathbb{N}$.

Example 5.51. We consider the space of square summable sequences $l^2(\mathbb{N}) = \{(v_k)_{k \in \mathbb{N}} : \sum_{k=1}^{\infty} |v_k|^2 < \infty\}$.

- $\langle (v_1, v_2, v_3, \ldots), (w_1, w_2, w_3, \ldots) \rangle = \sum_{k=1}^{\infty} v_k \overline{w_k}$ defines an inner product on $l^2(\mathbb{N})$;
- $\|(v_1, v_2, v_3, \ldots)\| = \|(v_1, v_2, v_3, \ldots)\|_2 = \sqrt{\sum_{k=1}^{\infty} |v_k|^2}$ is the norm on $l^2(\mathbb{N})$ induced by $\langle \cdot, \cdot \rangle$;
- $d_2((v_1, v_2, v_3, \ldots), (w_1, w_2, w_3, \ldots)) = \sqrt{\sum_{k=1}^{\infty} |v_k - w_k|^2}$ is the metric induced by $\| \cdot \|_2$;
- $(l^2(\mathbb{N}), d_2)$ is complete, $l^2(\mathbb{N})$ is therefore a Hilbert space;
- $\{e_1 = (1, 0, 0, 0, \ldots), e_2 = (0, 1, 0, 0, \ldots), e_3 = (0, 0, 1, 0, \ldots), \ldots\}$ is an ONS in K^n;
- $l^2(\mathbb{N})$ is an infinite dimensional vector space.

Remark 5.52. In Example 5.50 we have for any $v \in K^n$,

$$\begin{aligned} v &= (v_1, \ldots, v_n) \\ &= v_1(1, 0, 0, \ldots, 0, 0) + v_2(0, 1, 0, \ldots, 0) + \ldots + v_{n-1}(0, 0, 0, \ldots, 1, 0) + v_n(0, 0, 0, \ldots, 0, 1) \\ &= \sum_{k=1}^{n} \langle v, e_k \rangle e_k. \end{aligned}$$

This is a consequence of the fact that $\{e_k\}$ form a basis and an ONS.

Actually, whenever $\{\varphi_n\}_{n=1,\ldots,K}$ is an orthonormal system in $\mathbb{C}^N$, with $K = N$, then then $\{\varphi_n\}_{n=1,\ldots,N}$ is a so–called orthonormal basis of $\mathbb{C}^N$, and for any $v \in \mathbb{C}^N$, we have $v = \sum_{n=1}^{N} \langle v, \varphi_n \rangle \, \varphi_n$.

In the infinitely dimensional vector space $l^2(\mathbb{N})$ which is discussed in Example 5.51, the infinite set $\{e_1 = (1, 0, 0, 0, \ldots), e_2 = (0, 1, 0, 0, \ldots), e_3 = (0, 0, 1, 0, \ldots), \ldots\}$ is not a basis (in the sense of linear algebra) of $l^2(\mathbb{N})$. For example, note that the sequence $\{\frac{1}{k}\} \in l^2(\mathbb{N})$ cannot be written as a finite linear combination of vectors in $\{e_1 = (1, 0, 0, 0, \ldots), e_2 = (0, 1, 0, 0, \ldots), e_3 = (0, 0, 1, 0, \ldots), \ldots\}$. But we do have that

$$\sum_{k=1}^{N} \langle \{\frac{1}{k}\}, e_k \rangle e_k = \sum_{k=1}^{N} \frac{1}{k} e_k = (1. \frac{1}{2}, \frac{1}{3}, \ldots, \frac{1}{N-1}, \frac{1}{N}, 0, 0, \ldots) \longrightarrow (1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \frac{1}{5}, \ldots) = \{\frac{1}{k}\}.$$

with convergence in the d_2 metric which is induced by the inner product on $l^2(\mathbb{N})$.

Theorem 5.53. *If $\{\phi_n\}_{n\in\mathbb{N}}$ is an orthonormal system in the inner product space V, then for any $v \in V$ we have*

$$\|f - \sum_{n=1}^{N} \langle v, \phi_n \rangle \, \phi_n \, \|_2^2 = \|v\|^2 - \sum_{n=1}^{N} |\langle v, \phi_n \rangle|^2.$$

Hence, for $v \in V$, we have $v = \sum_{n=1}^{\infty} \langle v, \phi_n \rangle \, \phi_n$ in V (with convergence in the metric topology induced by the inner product on V) if and only if $\|f\|^2 = \sum_{n=1}^{\infty} |\langle v, \phi_n \rangle|^2$.

Corollary 5.54. Bessel Inequality. *If $\{\phi_n\}_{n\in\mathbb{N}}$ is an orthonormal system in the inner product space V, then for any $v \in V$ we have*

$$\|v\|^2 \geq \sum_{n=1}^{\infty} |\langle v, \phi_n\rangle|^2.$$

Definition 5.55. If $\{\phi_n\}_{n\in\mathbb{N}}$ is an orthonormal system in the inner product space V and if for all $v \in V$ we have $v = \sum_{n=1}^{\infty} \langle v, \phi_n\rangle \, \phi_n$ in V, then we call $\{\phi_n\}_{n\in\mathbb{Z}}$ *orthonormal basis* of V.

Remark 5.56. The concept of othornomal bases in infinite dimensional inner product spaces generalizes the concept of orthonromal bases in finite dimensional vector spaces. As not every basis in finite dimensions is orthonormal, you may expect that there are also more general concepts of bases in infinite dimensional spaces. The most common generalizations of bases in finite dimensions are so–called unconditional bases, Riesz bases, and Schauder bases.

Let us now consider Riemann integrable functions mapping the infinite set $[0,1]$ into K with $K = \mathbb{R}$ or $K = \mathbb{C}$. To cover both cases, we need to first define what a complex valued Riemann integrable function is.

Definition 5.57. Integrals of complex valued functions. If for $f : [a,b] \mapsto \mathbb{C}$, the real valued functions $Re(f), Im(f)$ satisfy $Re(f), Im(f) \in \mathcal{R}([a,b])$, then we set $\int_a^b f(x)dx = \int_a^b Re(f(x))\, dx + i \int_a^b Im(f(x))\, dx$ and say that f is Riemann integrable.

This extends the definition of real valued Riemann integrable functions and from now on, the set $\mathcal{R}([a,b])$ denotes the complex vector space of complex valued and Riemann integrable functions.

Example 5.58. On the complex vector space $\mathcal{R}([0,1])$ we define an equivalence relation by $f \sim g$ if $\int_0^1 |f(x) - g(x)|^2\, dx = 0$. The set of equivalence classes w.r.t. $\sim$ is denoted by $\mathcal{R}'([0,1])$. We shall abuse notation by identifying the equivalence class $[f]$ with its representative f.

On the complex vector space $\mathcal{R}'([0,1])$ we define as in Example **??**

- the inner product $\langle f, g\rangle = \displaystyle\int_0^1 f(x)\overline{g(x)}\, dx$,
- the corresponding norm $\|f\|_2 = \sqrt{\langle f, f\rangle} = \sqrt{\displaystyle\int_0^1 |f(x)|^2\, dx}$,
- and the corresponding metric $d_2(f,g) = \|f - g\|_2 = \sqrt{\displaystyle\int_0^1 |f(x) - g(x)|^2\, dx}$.

Note that the metric space $(\mathcal{R}'([0,1]), d_2)$ is not complete. To see this, consider, for example, the Cauchy sequence given by the functions $f_n(x) = x^{-\frac{1}{2}}$ for $x > 1/n$ and $f_n(x) = 0$ for $x \leq 1/n$. Also, note that d_2 fails to be a metric on $\mathcal{R}([0,1]$ as $d_2(\chi_{\{1/2\}}, 0) = 0$. In $\mathcal{R}'([0,1]$ we have $\chi_{\{1/2\}} = 0$ while in $\mathcal{R}([0,1]$ we have $\chi_{\{1/2\}} \neq 0$.

5.5. Fourier series

In this section, we will consider the infinite dimensional inner product space $\mathcal{R}'([0,1])$. We will show that the Fourier system $\{e^{2\pi inx}\}_{n\in\mathbb{Z}}$ is an orthonormal basis of $\mathcal{R}'([0,1])$. Other than in the previous section, we shall consider the integers $\mathbb{Z}$ as index set. Sums of the form $\sum_{k=-\infty}^{\infty}\cdots$ stand for the limit $\lim_{N} \to \infty \sum_{-N}^{N} \ldots$.

Remark 5.59. A family of functions $\{\phi_n\}_{n\in\mathbb{Z}} \subset \mathcal{R}'([0,1])$ is orthogonal, if $\langle \phi_n, \overline{\phi_m}\rangle = \int_0^1 \phi_n(x)\overline{\phi_m(x)}\,dx = 0$ for $n \neq m$. If in addition $\|\phi_n\|_2 = \sqrt{\int_0^1 |\phi_n(x)|^2\,dx}$ for $n \in \mathbb{Z}$, then $\{\phi_n\}_{n\in\mathbb{Z}}$ is an ONS.

Definition 5.60. Let $e_n \in C^\infty(\mathbb{R})$ be given by $e_n(x) = e^{2\pi inx}$, $x \in \mathbb{R}$.

i. Let $c_n \in \mathbb{C}$ for $n \in \mathbb{Z}$. A function $f : [0,1] \longrightarrow \mathbb{C}$ with $f(x) = \sum_{n=-N}^{N} c_n e^{2\pi inx}$ is called *trigonometric polynomial.*

   The formal expression $\sum_{n=-\infty}^{\infty} c_n e_n = \sum_{n=-\infty}^{\infty} c_n e^{2\pi inx}$ is called *trigonometric series*, and $\sum_{n=-\infty}^{\infty} c_n e^{2\pi inx}$ converges pointwise or uniformly to a function g if $\lim_{N\to\infty} \sum_{n=-N}^{N} c_n e^{2\pi inx}$ converges pointwise or uniformly. Similarly, if X is a vector space of functions on $[0,1]$ with $\{e_n\}_{n\in\mathbb{Z}} \subset X$ and if d_X is a metric on X, then $\sum_{n=-\infty}^{\infty} c_n e_n$ converges to g in X if

   $$\lim_{N\to\infty} d_X\left(\sum_{n=-N}^{N} c_n e_n\,,\ g\right) = 0.$$

ii. For $f \in \mathcal{R}'([0,1])$ and $n \in \mathbb{Z}$ we call the complex number $\widehat{f}(n) = \int_0^1 f(x)e^{-2\pi inx}\,dx$ the $n-th$ *Fourier coefficient* of f.

iii. The trigonometric series $\sum_{n=-\infty}^{\infty} \widehat{f}(n)e^{2\pi inx}$ is called *Fourier series* of f and we write $f \sim \sum_{n=-\infty}^{\infty} \widehat{f}(n)e^{2\pi inx}$. For the partial sums of a Fourier series, we write $S(f,N) = \sum_{n=-N}^{N} \widehat{f}(n)e_n$.

Proposition 5.61. *The family $\{e_n\}_{n\in\mathbb{Z}}$ is an orthonormal system in $\mathcal{R}'([0,1])$.*

Corollary 5.62. *(Riemann–Lebesgue Lemma) For $f \in \mathcal{R}'([0,1])$, we have $|\widehat{f}(n)| \longrightarrow 0$ as $|n| \to \infty$.*

Remark 5.63. The question arises whether for f in the infinitely dimensional vector space $\mathcal{R}'([0,1])$ we have

$$f = \lim_{N\to\infty} \sum_{n=-N}^{N} \langle f, e_n\rangle\, e_n = \lim_{N\to\infty} \sum_{n=-N}^{N} \widehat{f}(n)\, e_n = \lim_{N\to\infty} S(f,N)$$

with convergence in $(\mathcal{R}'([0,1]), d_2)$.

Note that the ONS $\{e_{2n}\}_{n\in\mathbb{Z}}$ contains also infinitely many orthonormal elements, but

$$e_1 \neq 0 = \lim_{N\to\infty} \sum_{n=-N}^{N} 0\, e_{2n} = \lim_{N\to\infty} \sum_{n=-N}^{N} \langle e_1, e_{2n}\rangle\, e_{2n}.$$

We conclude that not every infinite ONS behaves as an ONB in a finite dimensional vector space. The question remains whether the full family $\{e_n\}_{n\in\mathbb{Z}}$ contains "sufficiently" many elements to do the trick.

Further, if $\{e_n\}_{n\in\mathbb{Z}}$ does the trick do we have automatically

$$f(x) = \lim_{N\to\infty} \sum_{n=-N}^{N} \widehat{f}(n)\, e^{2\pi i n x}$$

for all $x \in [0,1]$, that is, do we have pointwise convergence?

Lemma 5.64. *For $x \in [0,1]$, we have* $\displaystyle\sum_{n=1}^{\infty} \frac{\cos 2\pi n x}{n^2} = \pi^2 \left(\frac{(2x-1)^2}{4} - \frac{1}{12} \right)$.

Lemma 5.65. *Let $f \in \mathcal{R}'([0,1])$ be a real valued step function, that is, for some partition $P = \{x_0, x_1, \ldots, x_R\}$ of $[0,1]$, f is constant on the intervals $[x_r, x_{r+1})$, $r = 0, \ldots R-1$, then $f = \lim_{N\to\infty} S(f,N)$ with convergence in $(\mathcal{R}'([0,1]), d_2)$.*

Theorem 5.66. *For any $f \in \mathcal{R}'([0,1])$ we have $f = \lim_{N\to\infty} S(f,N)$ in $(\mathcal{R}'([0,1]), d_2)$, and, hence, $\{e_n\}_{n\in\mathbb{Z}}$ is an orthonormal basis in $\mathcal{R}'([0,1])$.*

Proposition 5.67. *If $f \in C([0,1])$ with $\widehat{f}(n) = 0$ for all $n \in \mathbb{Z}$, then $f(x) = 0$ for all $x \in [0,1]$.*

Proposition 5.68. *If $f \in C([0,1])$ with $\displaystyle\sum_{n=-\infty}^{+\infty} |\widehat{f}(n)|$ convergent, then $S(f,N) \longrightarrow f$ uniformly (and therefore pointwise).*

Remark 5.69. Theorem 5.66 naturally applies to real valued function $f \in \mathcal{R}'([0,1])$, that is for a Riemann integrable function $f : [0,1] \longrightarrow \mathbb{R}$, we have

$$f = \sum_{k=-\infty}^{\infty} \widehat{f}(k) e_k = \sum_{k=-\infty}^{\infty} \langle f, e_k\rangle e_k\,.$$

But this kind of expansion is not very intriguing as it involves expanding a real valued function as a series of complex valued functions with complex coefficients, knowing that in the end, the imaginary contribution will cancel out. Using the fact that f is real valued, we compute

$$\begin{aligned}
f(x) &= \mathrm{Re} f(x) = \mathrm{Re} \sum_{k=-\infty}^{\infty} \widehat{f}(k) e_k(x) \\
&= \sum_{k=-\infty}^{\infty} \mathrm{Re}\big((\mathrm{Re}\widehat{f}(k) + i\mathrm{Im}\widehat{f}(k))(\cos 2\pi kx + i \sin 2\pi kx)\big) \\
&= \sum_{k=-\infty}^{\infty} \big(\mathrm{Re}\widehat{f}(k) \cos 2\pi kx - \mathrm{Im}\widehat{f}(k) \sin 2\pi kx\big) \\
&= \mathrm{Re}\widehat{f}(0) + \sum_{k=1}^{\infty} \big(\mathrm{Re}\widehat{f}(k) \cos 2\pi kx - \mathrm{Im}\widehat{f}(k) \sin 2\pi kx\big) \\
&\qquad + \big(\mathrm{Re}\widehat{f}(-k) \cos(-2\pi kx) - \mathrm{Im}\widehat{f}(k) \sin(-2\pi xn)\big) \\
&= \mathrm{Re}\widehat{f}(0) + \sum_{k=1}^{\infty} \big(\mathrm{Re}\widehat{f}(k) + \mathrm{Re}\widehat{f}(-k)\big) \cos 2\pi kx - \big(\mathrm{Im}\widehat{f}(k) - \mathrm{Im}\widehat{f}(-k)\big) \sin 2\pi kx.
\end{aligned}$$

Further, using again the fact that f is real valued, we obtain $\mathrm{Re}\widehat{f}(0) = \mathrm{Re} \int_0^1 f(x)\,dx = \int_0^1 f(x)\,dx = \widehat{f}(0)$,

$$\mathrm{Re}\widehat{f}(-k) = \mathrm{Re}\langle f, \cos(-2\pi kx) + i \sin(-2\pi kx)\rangle = \langle f, \cos 2\pi kx\rangle = \mathrm{Re}\widehat{f}(k)$$

and

$\mathrm{Im}\widehat{f}(-k) = \mathrm{Im}\langle f, \cos(-2\pi kx) + i \sin(-2\pi kx)\rangle = \mathrm{Im} - i\langle f, -\sin 2\pi kx\rangle = \langle f, \sin 2\pi kx\rangle = -\mathrm{Im}\widehat{f}(k)$.

Consequently,

$$\begin{aligned}
f(x) &= \widehat{f}(0) + \sum_{k=1}^{\infty} 2\mathrm{Re}\widehat{f}(k) \cos 2\pi kx - 2\mathrm{Im}\widehat{f}(k) \sin 2\pi kx \\
&= a_0 + \sum_{k=1}^{\infty} a_k \sqrt{2} \cos 2\pi kx + b_k \sqrt{2} \sin 2\pi kx.
\end{aligned}$$

with real valued coefficients

$$\begin{aligned}
a_0 &= \langle f, \cos(2\pi \cdot 0(\cdot))\rangle = \int_0^1 f(x)\,dx \\
a_k &= \sqrt{2}\langle f, \cos(2\pi k(\cdot))\rangle = \int_0^1 f(x)\sqrt{2} \cos(2\pi kx)\,dx,\ k \in \mathbb{N}, \\
b_k &= \sqrt{2}\langle f, \sin(2\pi k(\cdot))\rangle = \int_0^1 f(x)\sqrt{2} \sin(2\pi kx)\,dx, k \in \mathbb{N}.
\end{aligned}$$

In fact, it is easy to see that the real valued functions in $\{1, \sqrt{2} \cos 2\pi kx, \sqrt{2} \sin 2\pi kx\}$ form an ONS in the space $\mathcal{R}'([0,1])$ of complex valued functions. Further, it is easy to deduce from the computations above that $\{1, \sqrt{2} \cos 2\pi kx, \sqrt{2} \sin 2\pi kx\}$ is an ONB for $\mathcal{R}'([0,1])$ and that in general $a_0 = \widehat{f}(0)$, and $a_k = \frac{1}{\sqrt{2}}\big(\widehat{f}(k) + \widehat{f}(-k)\big)$ and $b_k = \frac{i}{\sqrt{2}}\big(\widehat{f}(k) - \widehat{f}(-k)\big)$ for $k \in \mathbb{N}$.

6. MULTIVARIABLE CALCULUS

6.1. Some facts from linear algebra

We shall assume familiarity with basic linear algebra, that is, with concepts such as finite dimensional vector spaces, linear independence, basis, linear transformations, matrices, determinants and norms. The real vector spaces which will be of interest are Euclidean space $\mathbb{R}^n$ equipped with the Euclidean norm $\|x\| = \|x\|_2 = \sqrt{x_1^2 + x_2^2 + \ldots + x_n^2}$, $x = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$ and metric $d_2(x, y) = \|x - y\|_2$, $x, y \in \mathbb{R}^n$, and spaces of linear operators mapping one finite dimensional space into another one. $\{e_1, e_2, \ldots, e_n\}$ denotes the Euclidean basis of $\mathbb{R}^n$.

If $L : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ is linear, we write $L \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ and we denote by $[L] \in \mathbb{R}^{m \times n}$ the matrix representing L with respect to the Euclidean bases of $\mathbb{R}^n$ and $\mathbb{R}^m$.

Let's bring analysis to the table.

Theorem 6.1. *Any linear transformation $L : \mathbb{R}^n \longrightarrow W$, W is a vector space with norm $\|\cdot\|_W$, is uniformly continuous.*

Proof. Obviously, $L = 0$, that is, the linear transformation mapping all of $\mathbb{R}^n$ to $0 \in W$, is uniformly continuous. For $L \neq 0$, fix $\epsilon > 0$. Set $M = \max\{\|L(e_1)\|, \|L(e_2)\|, \ldots, \|L(e_n)\|\}$ and $\delta = \frac{\epsilon}{Mn}$. Note that $M > 0$ since else $L = 0$.

Fix $x, y \in \mathbb{R}^n$ such that $d_2(x, y) = \|x - y\|_2 < \delta$ and observe that

$$\begin{aligned} d(L(x), L(y)) &= \|L(x) - L(y)\|_W = \|L(x - y)\|_W = \|L(\sum_{k=1}^n (x_k - y_k) e_k)\|_W \\ &= \|\sum_{k=1}^n (x_k - y_k) L(e_k)\|_W \;\leq \sum_{k=1}^n |x_k - y_k| \, \|L(e_k)\|_W \\ &\leq \sum_{k=1}^n \|x - y\| \, \|L(e_k)\|_W < n\delta M = \epsilon \end{aligned}$$

□

We can now use analysis to show

Theorem 6.2. *Let W be a normed vector space of dimension $m \in \mathbb{N}$. The set $\mathcal{L}(\mathbb{R}^n, W) = \{L : \mathbb{R}^n \mapsto W,\ L \text{ linear}\}$ is a vector space of dimension $n \cdot m$ with operator norm $\|L\|_{\mathcal{L}} = \sup\left\{\frac{\|L(x)\|_W}{\|x\|} : x \in \mathbb{R}^n \setminus \{0\}\right\}$.*

Further, for all $x \in \mathbb{R}^n$ we have $\|L(x)\|_W \leq \|L\|_{\mathcal{L}} \|x\|_{\mathbb{R}^n}$.

Proof. To show that $\mathcal{L}(\mathbb{R}^n, W)$ is a linear space of dimension $n \cdot m$ is easy. Now we shall show that $\left\{\frac{\|L(x)\|}{\|x\|} : x \in \mathbb{R}^n \setminus \{0\}\right\}$ is bounded, and therefore $\|\cdot\| : \mathcal{L}(\mathbb{R}^n, W) \longrightarrow \mathbb{R}$ is well defined. First, observe that $S = \{\|x\| = 1\}$ is closed in R^n since $\|\cdot\|_2 : \mathbb{R}^n \longrightarrow \mathbb{R}$ is continuous, $\{1\} \subset \mathbb{R}$ is closed and $S = (\|\cdot\|_2)^{-1}(\{1\})$. Together with the fact that $S \subset \mathbb{R}^n$ is also bounded we get that S is compact (Heine-Borel theorem). The set $\left\{\frac{\|L(x)\|}{\|x\|} : x \in \mathbb{R}^n \setminus \{0\}\right\} =$

$\{\|L(x)\| : x \in \mathbb{R}^n, \|x\| = 1\}$ is the image of a compact set under a continuous function $\|\cdot\| \circ L$ and therefore compact and hence, bounded. (We could have chosen a direct proof using the same inequality presented in Theorem 1, but I enjoyed arguments from Analysis I.)

To show the norm properties is easy, I leave it to you. $\square$

6.2. Curves

Curves are functions mapping intervals in $\mathbb{R}$ into $\mathbb{R}^m$, that is

$$f = \begin{pmatrix} f_1 \\ f_2 \\ \vdots \\ f_m \end{pmatrix} : \ I \longrightarrow \ \mathbb{R}^m, \qquad t \mapsto \begin{pmatrix} f_1(t) \\ f_2(t) \\ \vdots \\ f_m(t) \end{pmatrix}. \tag{3}$$

Curves are vector valued functions, or, equivalently, a curve is a vector of functions each depending of one real variable. This makes them fairly easy objects to study on the basis of single variable calculus: the derivative of a curve is given componentwise, that is, for f given in 3, we set $f' = (f_1', f_2', \ldots, f_m')^T$.

Definition 6.3. A continuous mapping $\gamma : I \longrightarrow \mathbb{R}^m$, $I \subseteq \mathbb{R}$ is an interval, is called a *curve* in $\mathbb{R}^m$ or a curve on I in $\mathbb{R}^m$.

If γ is one-to-one, γ is called an *arc*, if $I = [a, b]$ and $\gamma(a) = \gamma(b)$, then γ is a *closed curve.*

Definition 6.4. For a partition $P = \{x_0, \ldots, x_N\}$ of $[a, b]$ and a curve $\gamma : [a, b] \longrightarrow \mathbb{R}^m$, we set

$$\Lambda(P, \gamma) = \sum_{n=1}^{N} \|\gamma(x_n) - \gamma(x_{n-1})\|_2.$$

The *length of a curve* on the interval I is $\Lambda(\gamma) = \sup\{\Lambda(P, \gamma) : \ P \text{ partitions } I\}$.

If $\Lambda(\gamma) < \infty$, then we call γ *rectifiable.*

Definition 6.5. Two curves $\gamma_1 : I_1 \longrightarrow \mathbb{R}^m$ and $\gamma_2 : I_2 \longrightarrow \mathbb{R}^m$ are called *equivalent* if for some bijective and continuous map $\beta : I_1 \longrightarrow I_2$ we have $\gamma_1 = \gamma_2 \circ \beta$.

Definition 6.6. A curve is *regular* if $\gamma \in C^1(I)$ and if for all $t \in I$ we have

$$\|\gamma'(t)\|_2^2 = \left(\frac{d\gamma_1}{dt}(t)\right)^2 + \ldots + \left(\frac{d\gamma_m}{dt}(t)\right)^2 > 0\,.$$

where $\gamma'(t) = \left(\frac{d\gamma_i}{dt}(t)\right)_{i=1,\ldots,m} = \left(\frac{d\gamma_1}{dt}(t), \ldots, \frac{d\gamma_m}{dt}(t)\right)^T$.

Remark 6.7. For a regular $\gamma : I \longrightarrow \mathbb{R}^m$, $\tau(t) = \dfrac{\gamma'(t)}{\|\gamma'(t)\|_2}$ is the unit vector of the tangent line of γ at t and is independent of the parametrization γ. Also, $\left\|\left(\frac{d\gamma_i}{dt}(t)\right)_i\right\|_2$is considered the speed of γ at t.

Theorem 6.8. *If $\gamma \in C^1[a, b]$, then γ is rectifiable and $\Lambda(\gamma) = \int_a^b \|\gamma'(t)\|_2 \, dt$.*

6.3. Derivatives of multivariable functions

Definition 6.9. Let $U \subset \mathbb{R}^n$ be open. The function $f : U \longrightarrow \mathbb{R}^m$ is *differentiable* at $a \in U$ with derivative $(Df)_a \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ if

$$f(x) = f(a) + (Df)_a(x-a) + R(x), \quad \text{with} \quad \lim_{x \to a} \frac{1}{\|x-a\|} R(x) = 0.$$

The linear map $(Df)_a$ is called *total derivative*, or simply derivative, of f at a.

We shall see below, that the total derivative of a function at a point is unique, that is, we can talk about *the* derivative of f at a and not about a derivative of f at a.

Be aware of the fact that for each a where f is differentiable, $(Df)_a$ is a linear map and not a number as in the one dimensional case. Further if $f : U \longrightarrow \mathbb{R}^m$ is differentiable for all $a \in U$, we obtain a function $Df : U \longrightarrow \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$, $x \mapsto (Df)_x$. Later, we will extend the definition given above by replacing $\mathbb{R}^m$ by any normed finite dimensional vector space. Since $\mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ is such a space, we will be able to pose the question whether Df is differentiable as well, that is, exists a second derivative $D^2 f$?

Example 6.10. If $A : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ is a linear map, we have $(DA)_a = A$ for all $a \in \mathbb{R}^n$.

Indeed, comparing $A(x) = A(a) + A(x-a)$ with $A(x) = A(a) + (DA)_a(x-a) + R(x)$ implies that $(DA)_a = A$ and $R(x) = 0 \in \mathbb{R}^m$ (where clearly $\lim_{x\to a} \frac{1}{\|x-a\|} R(x) = \lim_{x \to a} \frac{0}{\|x-a\|} = 0$.

Theorem 6.11. *If $f : U \longrightarrow \mathbb{R}^m$, $U \subset \mathbb{R}^n$ open, is differentiable at $a \in U$, we can determine the action of $(Df)_a$ according to the limit formula*

$$(Df)_a(x) = \lim_{t \to 0} \frac{1}{t}\big(f(a+tx) - f(a)\big) \quad \textit{for any } x \in \mathbb{R}^n.$$

Proof. With $x_t = a + tx$, we have

$$\begin{aligned} \frac{1}{t}\big(f(a+tx) - f(a)\big) &= \frac{1}{t}\big(f(x_t) - f(a)\big) = \frac{1}{t}\big((Df)_a(x_t - a) + R(x_t)\big) \\ &= \frac{1}{t}\big((Df)_a(tx) + R(x_t)\big) = \frac{t}{t}(Df)_a(x) + \frac{1}{t}R(x_t)\big) \\ &= (Df)_a(x) + \frac{1}{\|tx\|}R(x_t) = (Df)_a(x) + \frac{1}{\|x_t - a\|}R(x_t). \end{aligned}$$

The result follows by the differentiability of f, that is, $\frac{1}{\|x_t - a\|}R(x_t) \to 0$ since $\lim_{t\to\infty} x_t = a$. □

Definition 6.12. If $f : U \longrightarrow \mathbb{R}^m$, $U \subset \mathbb{R}^n$ open, $a \in U$, and $\|x\| = 1$, then we call the limit, if it exists, $\lim_{t\to 0} \frac{1}{t}\big(f(a+tx) - f(a)\big)$ *directional derivative* at a in the direction x.

Recall the simple observation that a vector valued function is a vector of scalar valued functions, that is, any function $f : U \longrightarrow \mathbb{R}^m$, $U \subset \mathbb{R}^n$, can be represented by m real valued functions $f_1, f_2, \dots, f_m$ via $f(x) = (f_1(x), f_2(x), \dots, f_m(x))$, $x \in U$. Fixing $n-1$ components of x, we obtain a real valued function defined on an open subset of $\mathbb{R}$. The derivative of this function is a partial derivative.

Definition 6.13. The $(i\,j)^{\text{th}}$ *partial derivative* of $f = (f_1, f_2, \dots, f_m) : U \longrightarrow \mathbb{R}^m$ at a is the limit, if it exists,

$$(D_j f_i)(a) = \frac{\partial f_i}{\partial x_j}(a) = \lim_{t\to 0} \frac{f_i(a + te_j) - f_i(a)}{t} \in \mathbb{R}.$$

If all partial derivatives of f exist at $a \in U$, then we refer to the matrix of partials, that is, to

$$\begin{pmatrix} \frac{\partial f_1}{\partial x_1}(a) & \frac{\partial f_1}{\partial x_2}(a) & \cdots & \frac{\partial f_1}{\partial x_n}(a) \\ \frac{\partial f_2}{\partial x_1}(a) & \frac{\partial f_2}{\partial x_2}(a) & \cdots & \frac{\partial f_i}{\partial x_n}(a) \\ \vdots & \vdots & & \vdots \\ \frac{\partial f_m}{\partial x_1}(a) & \frac{\partial f_m}{\partial x_2}(a) & \cdots & \frac{\partial f_m}{\partial x_n}(a) \end{pmatrix},$$

as *Jacobian matrix* of f at a.

If $m = 1$, then we call the Jacobian matrix *gradient* and denote it by $\text{grad} f(x)$ or $\nabla f(x)$.

Definition 6.14. For $f : U \longrightarrow \mathbb{R}^n$, $U \subset \mathbb{R}^n$, that is, $m = n$, then f is called a *vector field* and $J_f(x) = \det \left(\frac{\partial f_i}{\partial x_j}\right)_{i,j=1,\dots n}(x)$ is called *Jacobian determinant* (or simply *Jacobian*).

The *divergence* of f is the function $\text{div} f = \sum_{i=1}^{n} \frac{\partial f_i}{\partial x_i}$.

If $m = n = 3$, then we call the vector field $\text{rot} f = \left(\frac{\partial f_3}{\partial x_2} - \frac{\partial f_2}{\partial x_3}, \frac{\partial f_1}{\partial x_3} - \frac{\partial f_3}{\partial x_1}, \frac{\partial f_2}{\partial x_1} - \frac{\partial f_1}{\partial x_2}\right)$ *rotation* of f.

Remark 6.15. If $f : U \longrightarrow \mathbb{R}$, $U \subseteq \mathbb{R}^n$ open, is differentiable at $a \in U$ and $\|x\|_2 = 1$, then

$$(Df)_a x = \nabla f(a) \cdot x = \langle \nabla f(a)^T, x \rangle = \cos\theta \|\nabla f(a)\|_2,$$

where θ is the angle between a and x. Hence, $\nabla f(a)^T$ is the direction of steepest slope (ascent) of f at a.

Theorem 6.16. *If $f : U \longrightarrow \mathbb{R}^m$, $U \subset \mathbb{R}^n$ open, is differentiable at $a \in U$, then it is continuous at a.*

Proof. We estimate for $x, a \in U$,

$$\begin{aligned} 0 \le \|f(x) - f(a)\| &= \|(Df)_a(x-a) + R(x)\| \le \|(Df)_a(x-a)\| + \|R(x)\| \\ &\le \|(Df)_a\|_{\mathcal{L}} \|x - a\| + \|R(x)\|, \end{aligned}$$

where we used Theorem 6.2. Since $\|(Df)_a\|_{\mathcal{L}}\|x-a\| \to 0$ and $\|R(x)\| \to 0$ as $x \to a$, we can apply the squeezing theorem to obtain $\|f(x) - f(a)\| \to 0$ as $x \to a$ which is another way of saying $f(x) \to f(a)$ as $x \to a$. □

Theorem 6.17. *Existence of the total derivative of $f : U \longrightarrow \mathbb{R}^m$, $U \subset \mathbb{R}^n$ open, implies the existence of the partial derivatives, and we have*

$$[(Df)_a]_{ij} = \frac{\partial f_i}{\partial x_j}(a)$$

Proof. Theorem 6.11 ensures that

$$(Df)_a(e_j) = \lim_{t\to 0} \frac{1}{t}\big(f(a+te_j) - f(a)\big),$$

with convergence in $\mathbb{R}^m$. Since a sequence of vectors $\{v_n\}$ converges to v_0 in $\mathbb{R}^m$ if and only if each component $[v_n]_i$ converges to respective $[v_0]_i$, $i = 1, \ldots, m$, we have

$$[(Df)_a]_{ij} = [(Df)_a(e_j)]_i = \lim_{t\to 0} \frac{1}{t}\big[f(a+te_j) - f(a)\big]_i = \lim_{t\to 0} \frac{f_i(a+te_j) - f_i(a)}{t} = \frac{\partial f_i}{\partial x_j}(a). \quad \square$$

It is important to note that Theorem 6.17 does not feature an if and only if statement. Indeed, it is possible that a function f has partial derivatives at a but that differentiability of f at a is not given. (See homework problems.)

Nevertheless, in the following we will state central results which allows us to calculate derivatives without having to use the definition of derivatives.

For example, the following rules will help us to calculate derivatives.

Theorem 6.18. *Let $U \subset \mathbb{R}^n$ and $V \subset \mathbb{R}^m$ be open.*

i. If $f : U \longrightarrow \mathbb{R}^m$ is constant, we have $(Df)_a = 0$ for all $a \in U$.

ii. If $f : U \longrightarrow \mathbb{R}^m$ is linear, we have $(Df)_a = f$ for all $a \in U$.

iii. If $f, g : U \longrightarrow \mathbb{R}^m$ are differentiable at $a \in U$, so is $c\,f + d\,g$, $c, d \in \mathbb{R}$, with $D(c\,f + d\,g)_a = c\,(Df)_a + d\,(Dg)_a$.

iv. If $f : U \longrightarrow V$ is differentiable at $a \in U$ and if $g : V \longrightarrow \mathbb{R}^k$ is differentiable at $f(a) \in V$, then $g \circ f : U \longrightarrow \mathbb{R}^k$ is differentiable with $D(g \circ f)_a = D(g)_{f(a)} \circ D(f)_a$.

Proof. i. Just observe that with $Df = 0$, we obtain $R = 0$, so

$$\lim_{x\to a} \frac{1}{\|x - a\|} R(x) = 0$$

holds trivially.

ii. Again, after "guessing" $Df = f$, we simply observe $R = 0$.

iii. Here, we make the educated guess $D(c\,f + d\,g)_a = c\,(Df)_a + d\,(Dg)_a$ and realize that $R_{c\,f+d\,g} = cR_f + dR_g$.

iv. To prove this chain rule is the only involved case, it was discussed in class and can be found in the literature. $\square$

Easy but very important is the following application of the chain rule.

Theorem 6.19. *A function*

$$f : U \longrightarrow \mathbb{R}^m,\ (x_1, x_1, \dots x_n) \mapsto (f_1(x_1, x_1, \dots x_n), f_2(x_1, x_1, \dots x_n), \dots, f_m(x_1, x_1, \dots x_n)),$$

$U \subset \mathbb{R}^n$ open is differentiable at x if and only if $f_i : U \longrightarrow \mathbb{R}$ is differentiable for $i = 1, \dots, m$, and, moreover $\pi_i \circ (Df)_x = (Df_i)x$.

Proof. Homework Problem. □

Theorem 6.20. *All partial derivatives of $f : U \longrightarrow \mathbb{R}^m$, $U \subset \mathbb{R}^n$ open, exist and are continuous on U if and only if f is differentiable on U and $(Df) : U \longrightarrow \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ is continuous.*

Proof. If f is differentiable on U, then the partial derivatives exist on U. Now, if f has a continuous derivative (Df) at $a \in U$, then for $i = 1, \dots, m$ and $j = 1, \dots, n$, we have

$$\begin{aligned} 0 \le \Big|\frac{\partial f_i}{\partial x_j}(x) - \frac{\partial f_i}{\partial x_j}(a)\Big| &= |[(Df)_x e_j]_i - [(Df)_a e_j]_i| = |[((Df)_x - (Df)_a)e_j]_i| \\ &\le \|((Df)_x - (Df)_a)e_j\| \le \|(Df)_x - (Df)_a\|_{\mathcal{L}}\, \|e_j\| = \|(Df)_x - (Df)_a\|_{\mathcal{L}} \end{aligned}$$

where continuity of (Df) at a implies that $\|(Df)_x - (Df)_a\|_{\mathcal{L}}$ goes to zero as $x \to a$. Hence, $\frac{\partial f_i}{\partial x_j}(x) \to \frac{\partial f_i}{\partial x_j}(a)$ as $x \to a$.

Now, the challenging part. We assume that all partials of f exist on U and are continuous on U. With J_f denoting the Jacobian of f, we will show that for $a \in U$,

$$(4) \qquad R(x) = f(x) - f(a) - J_f(x - a) \quad \text{satisfies} \quad \frac{R(x)}{\|x - a\|} \to 0 \quad \text{as} \quad x \to a.$$

We conclude that the total derivative exists at each $a \in U$ and equals the Jacobian matrix. Clearly, the continuous dependence of the partials on a implies the continuous dependence of the Jacobian and therefore, of the total derivative.

To show (4), that is, total differentiability of f with derivative the Jacobian, we fix $\epsilon > 0$. Continuity of the partials allows us to choose $\delta > 0$ so that

$$\Big|\frac{\partial f_i}{\partial x_j}(x) - \frac{\partial f_i}{\partial x_j}(a)\Big| < \epsilon/\sqrt{mn} \quad \text{whenever} \quad \|x - a\| < \delta.$$

By possibly reducing δ, we can also assume that $B_\delta(a) \subseteq U$.

Now, we fix x with $\|x - a\| < \delta$. Choose coefficients v_j with $x = a + \sum_{j=1}^n v_j e_j$ and set $x_k = a + \sum_{j=1}^k v_j e_j$ for $k = 1, \dots, n$ and $x_0 = a$.

For i, j fixed, define $g(t) = f_i(x_{j-1} + tv_j)$, $t \in [0, 1]$, which is well defined since the line segments $x_{j-1} + tv_j$ are contained in U. We shall now apply the Mean Value Theorem, Theorem 4.11, to conclude

$$f_i(x_j)-f_i(x_{j-1})=g(1)-g(0)=g'(t_{ij})=\lim_{s\to\infty}\frac{f_i(x_{j-1}+(t_{ij}+s)v_je_j)-f_i(x_{j-1}+t_{ij}v_je_j)}{s}$$
$$=v_j\lim_{s\to\infty}\frac{f_i(x_{j-1}+t_{ij}v_je_j+sv_je_j)-f_i(x_{j-1}+t_{ij}v_je_j)}{sv_j}$$
$$=v_j\lim_{\widetilde{s}\to\infty}\frac{f_i(x_{j-1}+t_{ij}v_je_j+\widetilde{s}e_j)-f_i(x_{j-1}+t_{ij}v_je_j)}{\widetilde{s}}$$
$$=v_j\lim_{\widetilde{s}\to\infty}\frac{f_i(x_{j-1}+t_{ij}v_j+\widetilde{s})-f_i(x_{j-1}+t_{ij}v_j)}{\widetilde{s}}=\frac{\partial f_i}{\partial x_j}(x_{j-1}+t_{ij}v_j)v_j. \tag{5}$$

We set $x_{ij}=x_{j-1}+t_{ij}v_j\in B_\delta(a)$ and compute

$$|[R(x)]_i|=\big|[f(x)-f(a)-J_f(x-a)]_i\big|=\Big|\sum_{j=1}^n f_i(x_j)-f_i(x_{j-1})-\sum_{j=1}^n\frac{\partial f_i}{\partial x_j}(a)(x-a)_j\Big|$$
$$=\Big|\sum_{j=1}^n\frac{\partial f_i}{\partial x_j}(x_{ij})v_j-\sum_{j=1}^n\frac{\partial f_i}{\partial x_j}(a)v_j\Big|$$
$$=\Big|\Big\langle\Big(\frac{\partial f_i}{\partial x_1}(x_{i1})-\frac{\partial f_i}{\partial x_1}(a)),\dots,\frac{\partial f_i}{\partial x_n}(x_{in})-\frac{\partial f_i}{\partial x_n}(a)\Big)^T,(v_1,\dots,v_n)^T\Big\rangle\Big|$$
$$\le\Big\|\Big(\frac{\partial f_i}{\partial x_1}(x_{i1})-\frac{\partial f_i}{\partial x_1}(a)),\dots,\frac{\partial f_i}{\partial x_n}(x_{in})-\frac{\partial f_i}{\partial x_n}(a)\Big)^T\Big\|\cdot\|(v_1,\dots,v_n)^T\|$$
$$\le\sqrt{n}\,\frac{\epsilon}{\sqrt{mn}}\|x-a\|=\frac{\epsilon}{\sqrt{m}}\|x-a\|.$$

Consequently,

$$\|R(x)\|=\sqrt{\sum_{i=1}^m|[R(x)]_i|^2}\le\sqrt{\sum_{i=1}^m\frac{\epsilon^2}{m}\|x-a\|^2}=\epsilon\|x-a\|.$$

We conclude that for every $\epsilon>0$, there exists $\delta>0$ with $\|x-a\|<\delta$ implies

$$\frac{\|R(x)\|}{\|x-a\|}<\epsilon,$$

so (4) follows. □

We shall not prove nor use the following beautiful rule, which generalizes the product rule in the 1-D case.

Theorem 6.21. LEIBNIZ RULE. *Let $\boldsymbol{\beta}:\mathbb{R}^k\times\mathbb{R}^l\mapsto\mathbb{R}^m$ be bilinear, that is, for all fixed $b\in\mathbb{R}^k$, the function $\boldsymbol{\beta}(b,\cdot):\mathbb{R}^l\longrightarrow\mathbb{R}^m$ is linear and for all fixed $c\in\mathbb{R}^l$, the function $\boldsymbol{\beta}(\cdot,c):\mathbb{R}^k\longrightarrow\mathbb{R}^m$ is linear. Let U be open in $\mathbb{R}^n$ and let $f:U\longrightarrow\mathbb{R}^k$ and if $g:U\longrightarrow\mathbb{R}^l$ be both differentiable at $a\in U$. Then*

$$\boldsymbol{\beta}(f,g):\mathbb{R}^n\longrightarrow\mathbb{R}^m,\quad x\mapsto\boldsymbol{\beta}\big(f(x),g(x)\big)$$

is differentiable at a with derivative defined by

$$\big(D\boldsymbol{\beta}(f,g)\big)_a(x)=\boldsymbol{\beta}\big((Df)_a(x),g(a)\big)+\boldsymbol{\beta}\big(f(a),(Dg)_a(x)\big).$$

Theorem 6.22. Multivariate Mean Value Theorem. *If $f : U \longrightarrow \mathbb{R}^m$, $U \subset \mathbb{R}^n$ open, is differentiable on U and the line segment*

$$[a, x_1] = \{a + t(x_1 - a) : t \in [0, 1]\} \subset \mathbb{R}^n$$

is contained in U, then

$$\|f(x_1) - f(a)\| \leq \sup\{\|(Df)_x\|_{\mathcal{L}} : x \in U\} \ \|x_1 - a\|.$$

Note that other than in the 1-D case we do not obtain an equality, neither of function values, nor of respective norms!

Proof. Let $g(t) = \langle f(x_1) - f(a), f\big(a + t(x_1 - a)\big)\rangle$, $t \in [0, 1]$. We shall apply the Mean Value Theorem for scalar valued functions on $\mathbb{R}$ as in (5). We obtain the existence of t_0 such that $g(1) - g(0) = g'(t_0)$, hence,

$$\begin{aligned}
\|f(x_1) - f(a)\|^2 &= \langle f(x_1) - f(a), f(x_1) - f(a)\rangle = \langle f(x_1) - f(a), f(x_1)\rangle - \langle f(x_1) - f(a), f(a)\rangle \\
&= g(1) - g(0) = g'(t_0) = \langle f(x_1) - f(a), (Df)_{a+t_0(x_1-a)}(x_1 - a)\rangle \\
&\leq \|f(x_1) - f(a)\| \, \|(Df)_{a+t_0(x_1-a)}(x_1 - a)\| \\
&\leq \|f(x_1) - f(a)\| \, \|(Df)_{a+t_0(x_1-a)}\|_{op} \, \|x_1 - a\|.
\end{aligned}$$

The result follows. □

Definition 6.23. If

$$F : [a, b] \longrightarrow \mathbb{R}^{m\times n} = \mathrm{Mat}_{m\times n}(\mathbb{R}), \quad t \mapsto \begin{pmatrix} F_{11}(t) & \cdots & F_{1n}(t) \\ \vdots & & \vdots \\ F_{m1}(t) & \cdots & F_{mn}(t) \end{pmatrix}$$

satisfies $F_{ij} \in \mathcal{R}[a, b]$ for $i = 1, \ldots, n$ and $j = 1, \ldots m$, then we say F is *Riemann integrable*, $F \in \mathcal{R}[a, b]$ with integral

$$\int_a^b F(t)\, dt = \begin{pmatrix} \int_a^b F_{11}(t)\, dt & \cdots & \int_a^b F_{1n}(t)\, dt \\ \vdots & & \vdots \\ \int_a^b F_{m1}(t)\, dt & \cdots & \int_a^b F_{mn}(t)\, dt \end{pmatrix} \in \mathbb{R}^{m\times n}.$$

Further, if $F : [a, b] \longrightarrow \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ is continuous, then $\displaystyle\int_a^b F(t)\, dt$ denotes the linear operator given by

$$\left[\int_a^b F(t)\, dt\right] = \int_a^b [F(t)] \ dt$$

Theorem 6.24. C^1 – Mean Value Theorem. *If $f : U \longrightarrow \mathbb{R}^m$, $U \subseteq \mathbb{R}^n$ open, satisfies $f \in C^1(U)$, and if $[a, b] \subseteq U$, then*

$$f(b) - f(a) = T(b - a) \ \text{where}\ T = \int_0^1 (Df)_{a+t(b-a)}\, dt\,.$$

Conversely, if there is a family of linear maps $T_{a,b} \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ such that $f(b) - f(a) = T_{a,b}(b-a)$ for all a, b with $[a, b] \subseteq U$ and if $T_{a,b}$ depends continuously on a and b, then $f \in C^1(U)$ and $(Df)_a = T_{a,a}$.

Proof. We shall only proof the first part. Once again, we use the auxiliary function $g(t) = f_i(a + t(b - a))$. Clearly, the function is continuously differentiable, hence, the fundamental theorem of calculus is applicable, and, by applying the chainrule, we have

$$f_i(b) - f_i(a) = g(1) - g(0) = \int_0^1 g'(t)\, dt = \int_0^1 (Df_i)_{a+t(b-a)}(b - a)\, dt.$$

Stacking up gives us

$$f(b) - f(a) = \int_0^1 (Df)_{a+t(b-a)}(b - a)\, dt. \qquad \square$$

Theorem 6.25. *Assume that* $f : [a, b] \times (c, d) \longrightarrow \mathbb{R}$ *is continuous and* $\frac{\partial f}{\partial y}$ *exists and is continuous on* $[a, b] \times (c, d)$*. Then* $F(y) = \int_a^b f(x, y)\, dx$ *is* $C^1(c, d)$*, and*

$$F'(y) = \int_a^b \frac{\partial f}{\partial y}(x, y)\, dx, \quad y \in (c, d)\,.$$

Proof. Fix $y_0 \in (c, d)$ and $\beta > 0$ with $y_0 \in [c + \beta, d - \beta]$. The function $\frac{\partial f}{\partial y}$ is continuos on the compact set $[a, b] \times [c + \beta, d - \beta]$, hence, $\frac{\partial f}{\partial y}$ is uniformly continuous.

For $\epsilon > 0$ fixed, we can find δ with

$$\Big|\frac{\partial f}{\partial y}(x, y) - \frac{\partial f}{\partial y}(\widetilde{x}, \widetilde{y})\Big| < \frac{\epsilon}{b - a}, \quad \text{if} \quad \|(x, y) - (\widetilde{x}, \widetilde{y})\| < \delta.$$

For $h \neq 0$ with $|h| < \delta$, we compute

$$\begin{aligned}
\Big|\frac{F(y_0 + h) - F(y_0)}{h} - \int_a^b \frac{\partial f}{\partial y}(x, y_0)\, dx\Big| &= \Big|\frac{1}{h}\int_a^b f(x, y_0 + h) - f(x, y_0)\, dx - \int_a^b \frac{\partial f}{\partial y}(x, y_0)\, dx\Big| \\
&= \Big|\frac{1}{h}\int_a^b \int_0^1 \frac{\partial f}{\partial y}(x, y_0 + th)h\, dt\, dx - \int_a^b \int_0^1 \frac{\partial f}{\partial y}(x, y_0)\, dt\, dx\Big| \\
&= \Big|\int_a^b \int_0^1 \frac{\partial f}{\partial y}(x, y_0 + th) - \frac{\partial f}{\partial y}(x, y_0)\, dt\, dx\Big| \\
&\leq \int_a^b \int_0^1 \Big|\frac{\partial f}{\partial y}(x, y_0 + th) - \frac{\partial f}{\partial y}(x, y_0)\Big|\, dt\, dx \\
&< \int_a^b \int_0^1 \frac{\epsilon}{b - a}\, dt\, dx = \epsilon,
\end{aligned}$$

where we used that $\|(x, y_0 + th) - (x, y_0)\| = |th| < \delta$ by assumption on h and $t \in [0, 1]$. It is easily checked that F' is continuous. $\square$

6.4. Higher derivatives

Definition 6.26. Let $U \subseteq \mathbb{R}^n$ be open and $f : U \longrightarrow W$, where W is a finite-dimensional normed vector space. Then f is differentiable at a with derivative $(Df)_a \in \mathcal{L}(\mathbb{R}^n, W)$, if

$$f(x) = f(a) + (Df)_a(x-a) + R(x) \quad \text{with} \quad \lim_{x \to a} \frac{1}{\|x-a\|} R(x) = 0.$$

Definition 6.27. Let $U \subseteq \mathbb{R}^n$ be open and $f : U \longrightarrow \mathbb{R}^m$. If $(Df)_x$ is defined for all x in a neighborhood of $a \in U$, and if Df is differentiable at $a \in U$, then we call $(D^2 f)_a = (D(Df))_a$ *second derivative* of f at a. Note that $(D^2f)_a \in \mathcal{L}(\mathbb{R}^n, \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)) =: \mathcal{L}^2(\mathbb{R}^n, \mathbb{R}^m)$. In addition, we can consider $(D^2f)_a$ as bilinear form, that is, we can write $(D^2f)_a : \mathbb{R}^n \times \mathbb{R}^n \longrightarrow \mathbb{R}^m$.

If $(D^2f)_x$ is defined for all x in a neighborhood of $a \in U$, and if D^2f is differentiable at $a \in U$, then we call $(D^3f)_a = (D(D^2f))_a$ *third derivative* of f at a. Note that $(D^3f)_a \in \mathcal{L}(\mathbb{R}^n, \mathcal{L}(\mathbb{R}^n, \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m))) =: \mathcal{L}^3(\mathbb{R}^n, \mathbb{R}^m)$.

...

If $(D^rf)_x$ is defined for all x in a neighborhood of $a \in U$, and if D^rf is differentiable at $a \in U$, then we call $(D^{r+1}f)_a = (D(D^rf))_a$ the $r+1$-st derivative of f at a.

The function f is of class $C^r(U, \mathbb{R}^m)$ if the r-th derivative of f exists at each $a \in U \subset \mathbb{R}^n$ and if $D^rf : \mathbb{R}^n \longrightarrow \mathcal{L}^r(\mathbb{R}^n, \mathbb{R}^m)$ is continuous. The function f is *smooth* if $f \in C^r(U, \mathbb{R}^m)$ for all $r \in \mathbb{N}$.

If $m = 1$, then we simply write $C^r(U)$ for $C^r(U, \mathbb{R})$, $r \in \mathbb{N} \cup \{\infty\}$.

Definition 6.28. Let $U, V \subseteq \mathbb{R}^n$ be open and $1 \leq r \leq \infty$. A bijective function $f : U \longrightarrow V$ is called C^r*–diffeomorphism* if $f \in C^r(U, \mathbb{R}^n)$ and $f^{-1} \in C^r(V, \mathbb{R}^n)$.

Theorem 6.29. *Let $f : U \longrightarrow \mathbb{R}^m$, $U \subseteq \mathbb{R}^n$ open. If $(D^2f)_a$ exists for $a \in U$, then*

$$(D^2f)_a(u,v) = (D^2f)_a(v,u).$$

Proof. Clearly, it suffices to assume that $m = 1$. Now, for $a, u, v \in \mathbb{R}^n$ fixed, we shall show the interesting fact that

$$(D^2f)_a(u,v) = \lim_{t \to 0} \frac{f(a+tu+tv) - f(a+tu) - f(a+tv) + f(a)}{t^2}. \tag{6}$$

The right hand side in (6) is symmetric in u and v, hence, the left hand side is too, therefore proving the result.

To show (6), we define again an auxiliary function, namely

$$g(s) = f(a+tu+stv) - f(a+stv),$$

so $g(0) = f(a+tu) - f(a)$ and $g(1) = f(a+tu+tv) - f(a+tv)$. By the one dimensional Mean Value Theorem, there exists $\theta_t \in [0,1]$ with the property

$$\begin{aligned}
\frac{1}{t^2}\big(f(a+tu+tv) - f(a+tv) - f(a+tu) + f(a)\big) &= \frac{1}{t^2}\big(g(1)-g(0)\big) = \frac{1}{t^2}g'(\theta_t) \\
&= \frac{1}{t^2}\big((Df)_{a+tu+\theta_t tv}(tv) - (Df)_{a+\theta_t tv}(tv)\big) \\
&= \frac{t}{t^2}\big((Df)_{a+tu+\theta_t tv}(v) - (Df)_{a+\theta_t tv}(v)\big) \\
&= \frac{1}{t}\big((Df)_a(v) + (D^2f)_a(tu+\theta_t tv)(v) + R_f(a+tu+\theta_t tv) \\
&\quad - (Df)_a(v) - (D^2f)_a(\theta_t tv)(v)\big) - R_f(a+\theta_t tv)\big) \\
&= \frac{1}{t}\big((D^2f)_a(tu+\theta_t tv)(v) - (D^2f)_a(\theta_t tv)(v) \\
&\quad + R_f(a+tu+\theta_t tv) - R_f(a+\theta_t tv)\big) \\
&= \frac{t}{t}\big((D^2f)_a(u+\theta_t v)(v) - (D^2f)_a(\theta_t v)(v)\big) \\
&\quad + \frac{1}{t}\big(R_f(a+tu+\theta_t tv) - R_f(a+\theta_t tv)\big) \\
&= (D^2f)_a(u)(v) + (D^2f)_a(\theta_t v)(v) - (D^2f)_a(\theta_t v)(v) \\
&\quad + \frac{1}{t}\big(R_f(a+tu+\theta_t tv) - R_f(a+\theta_t tv)\big) \\
&= (D^2f)_a(u)(v) + \frac{1}{t}\big(R_f(a+tu+\theta_t tv) - R_f(a+\theta_t tv)\big).
\end{aligned}$$

where we used that differentiability of $Df : U \to \mathcal{L}(\mathbb{R}^n, \mathbb{R})$ at a implies that

$$(Df)_y(\cdot) = (Df)_a(\cdot) + (D^2f)_a(y-x)(\cdot) + R_f(y)(\cdot),$$

with $R_f(y)(\cdot)/\|y-x\| \to 0$ in operator norm (or any other norm) as $y \to x$. The result follows by observing that, for example

$$0 \leq \Big|\frac{R_f(a+tu+\theta_t tv)}{t}\Big| = \frac{|R_f(a+tu+\theta_t tv)|}{|t|}\frac{\|tu+\theta_t tv\|}{\|tu+\theta_t tv\|} = \frac{|R_f(a+tu+\theta_t tv)|}{\|tu+\theta_t tv\|}\,\|u+\theta_t v\| \to 0$$

as $t \to 0$, since $\theta_t \in (0,1)$ implies $\|u+\theta_t v\|$ is bounded. □

Definition 6.30. For $f = (f_1, \ldots, f_m) : U \longrightarrow \mathbb{R}^m$, $U \subseteq \mathbb{R}^n$ open, we refer to $\dfrac{\partial^2 f_k}{\partial x_i \partial x_j}(a) \in \mathbb{R}$, $i, j = 1, \ldots n$, $k = 1, \ldots, m$ as *second partials* of f at $a \in U$.

For $m = 1$, that is, $f : U \longrightarrow \mathbb{R}$, we refer to the matrix $\mathrm{Hess} f(a) = [\frac{\partial^2 f_k}{\partial x_i \partial x_j}(a)]_{i,j} \in \mathbb{R}^{n\times n}$ as *Hessian matrix* of f at a.

We obtain the following corollaries to Theorem 6.29

Corollary 6.31. *Let $f = (f_1, \ldots, f_m) : U \longrightarrow \mathbb{R}^m$, $U \subseteq \mathbb{R}^n$ open. If $(D^2 f)_a$ exists, then for $k = 1, \ldots, n$, $(D^2 f_k)_a$ exists, the Hessians $\mathrm{Hess} f_k(a)$, $k = 1, \ldots, m$ exist and*

$$(D^2 f_k)_a(e_i, e_j) = \frac{\partial^2 f_k}{\partial x_i \partial x_j}(a), \quad i, j = 1, \ldots, n.$$

Also, if all second partials of f exist on U and if they are all continuous at a, then Df is differentiable at a.

Proof. This is mainly a vocabulary test. The result follows from the results on the first derivative of a function, which is a matrix valued function. □

Corollary 6.32. *For $f : U \longrightarrow \mathbb{R}^m$, $U \subseteq \mathbb{R}^n$ open, such that $(D^2 f)_a$ exists for $a \in U$, then*

$$\frac{\partial^2 f_k}{\partial x_i \partial x_j}(a) = \frac{\partial^2 f_k}{\partial x_j \partial x_i}(a), \qquad \text{for } i, j = 1, \ldots, n \text{ and } k = 1, \ldots, m,$$

in short, the Hessians of differentiable functions are symmetric.

6.5. Taylor's theorem and applications

Definition 6.33. Let $U \subseteq \mathbb{R}^n$ and $f : U \longrightarrow \mathbb{R}$. Then $f(a)$ is called a *maximum* or *global maximum* attained at a [resp. *minimum*] if $f(a) \geq f(x)$ [resp. $f(a) \leq f(x)$] for all $x \in U$. In either case, $f(a)$ is called *extremum* of f on U.

If $f(a) \geq f(x)$ [resp. $f(a) \leq f(x)$] for all $x \in B_\epsilon(a)$ for some $\epsilon > 0$, then f has a *local maximum* [resp. *local minimum*] $f(a)$ at a. Also, $f(a)$ is called *local extremum* of f.

Theorem 6.34. *Let $U \subseteq \mathbb{R}^n$ open, $f : U \longrightarrow \mathbb{R}$ with $\frac{\partial f}{\partial x_i}$ exists for $i = 1, \ldots, n$. If $f(a)$ is a local maximum or a local minimum of f, then $\nabla f(a) = 0$.*

Proof. This is a simple application of single variable calculus. For $i = 1, \ldots, n$, define the function $g_i(t) = f(a + te_i)$. Since f has a local maximum (or minimum) at a, g_i has one at 0. Hence, $0 = g'(0) = (Df)_a(e_i) = \frac{\partial f}{\partial x_i}(a)$. □

Definition 6.35. Let $A \in \mathbb{R}^{m \times n}$ be symmetric.

i. A is called *positive definite* if $\langle A\xi, \xi \rangle > 0$ for all $\xi \in \mathbb{R}^n \setminus \{0\}$.

ii. A is called *negative definite* if $\langle A\xi, \xi \rangle < 0$ for all $\xi \in \mathbb{R}^n \setminus \{0\}$.

iii. A is called *positive semidefinite* if $\langle A\xi, \xi \rangle \geq 0$ for all $\xi \in \mathbb{R}^n \setminus \{0\}$.

iv. A is called *negative semidefinite* if $\langle A\xi, \xi \rangle \leq 0$ for all $\xi \in \mathbb{R}^n \setminus \{0\}$.

v. A is called *indefinite* if $\langle A\xi, \xi \rangle > 0$ and $\langle A\eta, \eta \rangle < 0$ for some $\xi, \eta \in \mathbb{R}^n$.

Remark 6.36. Since A is symmetric, there exists an orthonormal basis $\{u_i\}$ of $\mathbb{R}^n$ of eigenvectors, so $Au_i = \lambda_i u_i$ for some $\lambda_i \in \mathbb{R}$. For $\xi = \sum_{i=1}^n c_i u_i$, we have

$$\langle A\xi, \xi \rangle = \langle A \sum_{i=1}^n c_i u_i, \sum_{j=1}^n c_j u_j \rangle = \sum_{i=1}^n \sum_{j=1}^n c_i \bar{c}_j \langle \lambda_i u_i, u_j \rangle = \sum_{i=1}^n \sum_{j=1}^n c_i \bar{c}_j \lambda_i \langle u_i, u_j \rangle = \sum_{i=1}^n |c_i|^2 \lambda_i,$$

so A is positive definite if all eigenvalues are positive, negative definite if all eigenvalues are negative and indefinite if some eigenvalues are positive and some are negative.

Lemma 6.37. *The matrix* $A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} \in \mathbb{R}^{n \times n}$ *is positive definite if and only if* $\det \begin{pmatrix} a_{11} & \cdots & a_{1k} \\ \vdots & & \vdots \\ a_{k1} & \cdots & a_{kk} \end{pmatrix} > 0$ *for $k = 1, \ldots, n$.*

Theorem 6.38. Multivariate Taylor's Theorem I. *Let U be an open subset of $\mathbb{R}^n$ and $f : U \longrightarrow \mathbb{R}$ with $f \in C^m(E)$. Fix $a \in E$, and suppose $[a, x] \subseteq U$. Then*

$$f(x) = \sum \frac{(D_1^{s_1} \ldots D_n^{s_n} f)(a)}{s_1! \ldots s_n!} (x_1 - a_1)^{s_1} \ldots (x_n - a_n)^{s_n} + r(x) \qquad D_k = \frac{\partial}{\partial x_k}$$

where the summation extends over all ordered n-tuples $(s_1, ..., s_n)$ *such that each* s_i *is a non-negative integer, and* $s_i + ... + s_n \leq m-1$*, and the remainder r satisfies*

$$\lim_{x \to 0} \frac{r(x)}{\|x\|^{m-1}} = 0.$$

We shall only use and prove the following version of Taylor's theorem.

Theorem 6.39. Multivariate Taylor's Theorem II. *For* $f : U \longrightarrow \mathbb{R}^m$*,* $U \subseteq \mathbb{R}^n$ *open,* $f \in C^2(U)$*, we have*

$$(7) \quad f(x) = f(a) + (Df)_a(x-a) + \frac{1}{2}(Df)_a(x-a, x-a) + R(x) \text{ with } \lim_{x \to a} \frac{1}{\|x-a\|^2} R(x) = 0.$$

Proof. Clearly, the punchline of (7) is that the remainder satisfies $\frac{1}{\|x-a\|^2}R(x) = 0$. To check this, it suffices to consider x close to a, so we first pick $\epsilon > 0$ with $B_{2\epsilon}(a) \subseteq U$. For $x \in B_\epsilon(a)$, define the auxiliary function $g(t) = f(a + t(x-a))$, $t \in (-2, 2)$. g is the composition of the twice continuously differentiable function f and the smooth function $t \mapsto a + t(x-a)$. Hence, g is $C^2(-2, 2)$ and by Theorem 4.23 exists $\theta \in (0, 1)$ with

$$g(1) = g(0) + g'(0)(1-0) + \frac{1}{2}g''(\theta)(1-0)^2.$$

Now, $g(0) = f(a)$. By the chain rule, we have

$$\begin{aligned} g'(t) &= \frac{d}{dt}f(a + t(x-a)) = (Df)_{a+t(x-a)} \circ (x-a) = (Df)_{a+t(x-a)}(x-a) \\ &= \nabla f(a + t(x-a))\,(x-a) = \sum_{j=1}^{n}(x_j - a_j)\frac{\partial}{\partial x_j}f(a + t(x-a)) \end{aligned}$$

and $g'(0) = (Df)_a(x-a) = (Df)_a(x-a)$. Further,

$$\begin{aligned} g''(t) = \frac{d}{dt}(Df)_{a+t(x-a)}(x-a) &= \sum_{j=1}^{n}(x_j - a_j)\frac{d}{dt}\frac{\partial}{\partial x_j}f(a + t(x-a)) \\ &= \sum_{j=1}^{n}(x_j - a_j)(D\frac{\partial}{\partial x_j}f)_{a+t(x-a))} \circ (x-a) \\ &= \sum_{j=1}^{n}(x_j - a_j)(\nabla\frac{\partial}{\partial x_j}f)(a + t(x-a)))\,(x-a) \\ &= \sum_{j=1}^{n}(x_j - a_j)\sum_{i=1}^{n}(x_i - a_i)\frac{\partial}{\partial x_i \partial x_j}f(a + t(x-a))) \\ &= \langle Hf_{a+t(x-a)}(x-a), (x-a)\rangle, \end{aligned}$$

where $Hf_y = \left(\frac{\partial}{\partial x_i \partial x_j} f(y)\right)_{i,j}$ is the Hessian matrix of f at y, which is $(D^2 f)_y$ expressed as matrix with respect to the Euclidean orthonormal basis. We have

$$\begin{aligned}
f(x) &= g(1) = g(0) + g'(0) + \frac{1}{2} g''(\theta) \\
&= f(a) + (Df)_a(x-a) + \frac{1}{2}(D^2 f)_{a+\theta(x-a)}(x-a, x-a) \\
&= f(a) + (Df)_a(x-a) + \frac{1}{2}(D^2 f)_a(x-a, x-a) \\
&\quad + \Big(\frac{1}{2}(D^2 f)_{a+\theta(x-a)}(x-a, x-a) - \frac{1}{2}(D^2 f)_a(x-a, x-a)\Big) \\
&= f(a) + (Df)_a(x-a) + \frac{1}{2}(D^2 f)_a(x-a, x-a) + \varphi(x)
\end{aligned}$$

where

$$\begin{aligned}
\frac{1}{\|x-a\|^2}|\varphi(x)| &= \frac{1}{2\|x-a\|^2}|\langle\big(Hf_{a+\theta(x-a)} - Hf_a\big)x-a,\, x-a\rangle| \\
&\leq \frac{1}{2\|x-a\|^2}\|\big(Hf_{a+\theta(x-a)} - Hf_a\big)x-a\|\,\|x-a\| \\
&\leq \frac{1}{2\|x-a\|^2}\|Hf_{a+\theta(x-a)} - Hf_a\|_{op(2)}\|x-a\|\,\|x-a\| \\
&\leq \frac{1}{2}\|Hf_{a+\theta(x-a)} - Hf_a\|_{op(2)},
\end{aligned}$$

where continuity of the second derivative implies that the right hand side goes to zero as $x \to a$, and so does $\frac{1}{\|x-a\|^2}\varphi(x)$.

□

Theorem 6.40. *Let $f : U \longrightarrow \mathbb{R}$, $U \subseteq \mathbb{R}^n$ open, $f \in C^2(U)$ be given with $(Df)_a = \nabla f(a) = 0$ for some $a \in U$.*

i. If $[(D^2 f)_a]$ is positive definite, then f has a local minimum at a, in fact, for some $\epsilon > 0$ we have $f(x) > f(a)$ for all $x \in B_\epsilon(a)$.

ii. If $[(D^2 f)_a]$ is negative definite, then f has a local maximum at a, and in fact, for some $\epsilon > 0$ we have $f(x) < f(a)$ for all $x \in B_\epsilon(a)$.

iii. If $[(D^2 f)_a]$ is indefinite, then f has neither a local minimum or maximum at a. (In this case, we speak of a saddle point or pass of f at a.)

Proof. The proof of this result relies on Theorem 6.39.

The function $F : \xi \mapsto \langle[(D^2 f)_a]\xi, \xi\rangle$ is continuous and assumes its minima on the compact set $S = \{\xi \in \mathbb{R}^n, \|\xi\| = 1\}$, that is, there exists ξ_0 with $\beta = \min\{F(\xi),\ \xi \in S\} = F(\xi_0)$. Since $[(D^2 f)_a]$ is positive definite, we have $\beta > 0$. Moreover, note that for all $\xi \in \mathbb{R}^n \setminus \{0\}$, we have

$$F(\xi) = \langle[(D^2 f)_a]\xi, \xi\rangle = \langle[(D^2 f)_a]\frac{\xi}{\|\xi\|}, \frac{\xi}{\|\xi\|}\rangle\|\xi\|^2 \geq \beta\|\xi\|^2.$$

Since $\varphi(x)/\|x-a\|^2 \to 0$ as $x \to a$, there exists $\delta > 0$ with $|\varphi(x)|/\|x-a\|^2 < \frac{\beta}{2}$ if $\|x-a\| < \delta$. For $x \in B_\delta(a)$ we compute

$$\begin{aligned} f(x) &= f(a) + (Df)_a(x-a) + \frac{1}{2}(D^2 f)_a(x-a, x-a) + \varphi(x) \\ &= f(a) + \frac{1}{2}F(x-a) + \varphi(x) \\ &> f(a) + \frac{1}{2}\beta\|x-a\|^2 - \frac{\beta}{2}\|x-a\|^2 = f(a), \end{aligned}$$

so $f(a)$ is a local maximum.

Note that the Hessian of f is negative definite if and only if the Hessian of $-f$ is positive definite, so the second part of the theorem follows from the first part.

Finally, we assume that there exits $\xi_0 > 0$ with $F(\xi_0) > 0$ and ξ_1 with $F(\xi_1) < 0$. By normalizing the ξ_i, we can assume $\|\xi_0\| = 1 = \|\xi_1\|$. We have for t with $B_t(a) \subseteq U$ and $\frac{|\varphi(a+t\xi_0)|}{\|a+t\xi_0-a\|^2} < \frac{1}{2}F(\xi_0)$ that

$$\begin{aligned} f(a+t\xi_0) &= f(a) + \frac{1}{2}(D^2 f)_a(t\xi_0, t\xi_0) + \varphi(a+t\xi_0) \\ &= f(a) + \frac{t^2}{2}F(\xi_0) - \|a+t\xi_0-a\|^2\frac{1}{2}F(\xi_0) \\ &> f(a). \end{aligned}$$

The same argument delivers $f(a+t\xi_1) < f(a)$ for t sufficiently small. Hence, we neither have a local minima, nor a local maxima. $\square$

6.6. Implicit functions and the inverse function theorem

Definition 6.41. Let $U \subseteq \mathbb{R}^n \times \mathbb{R}^m = \mathbb{R}^{n+m}$ be open and $f : U \longrightarrow \mathbb{R}^m$. For $(a, b) \in R^n \times \mathbb{R}^m$ and $f(a, b) = c \in \mathbb{R}^m$ we shall try to solve the system of not necessarily linear equations

$$(8)\ \ f(x,y) = \begin{pmatrix} f_1(x_1, \ldots, x_n, y_1, \ldots, y_m) \\ f_2(x_1, \ldots, x_n, y_1, \ldots, y_m) \\ \vdots \\ f_m(x_1, \ldots, x_n, y_1, \ldots, y_m) \end{pmatrix} = \begin{pmatrix} z_1 \\ z_2 \\ \vdots \\ z_m \end{pmatrix} = c \quad \text{for } (x, y) \in B_\epsilon(a, b),\ \epsilon > 0.$$

Similar to the linear case, we expect that if we fix x, then exists exactly one y solving the equation. In fact, in many cases there is $g : B_\epsilon(a) \longrightarrow \mathbb{R}^m$ such that (8) holds if and only if

$$y = g(x) \qquad \text{for some } x \in B_\epsilon(a),$$

that is, all solutions to (8) in $B_\epsilon(a, b)$ are given by $f(x, g(x)) = c$, $x \in B_\epsilon(a)$.

Then g is the *implicit function* defined by (8).

Theorem 6.42. Implicit Function Theorem. *Let $U \subseteq \mathbb{R}^n \times \mathbb{R}^m$ be open and $f : U \longrightarrow \mathbb{R}^m$ be given with $f \in C^r(U)$, $1 \leq r \leq \infty$. If $f(a, b) = c \in \mathbb{R}^m$ and $B = \frac{\partial f}{\partial y}(a, b) = \left(\frac{\partial f_i}{\partial y_j}(a, b)\right)_{i,j=1,\ldots m}$ is invertible, then exists $V \subseteq \mathbb{R}^n$ and $W \subseteq \mathbb{R}^m$ open with $V \times W \subseteq U$ and $g : V \longrightarrow W$, $g \in C^r(V)$ with*

$$\{(x, y) \in V \times W :\ f(x, y) = c\} = \{(x, g(x)), \quad x \in V\} = \Gamma_g.$$

Proof. Since translating F and adding a constant to F does not impact the differential properties of F, we can assume without loss of generality that $(a, b) = (0, 0) \in \mathbb{R}^n \times \mathbb{R}^m$ and $c = 0$. Note that in the proof, we make use of three norms, the standard Euclidean norm on $\mathbb{R}^m$ or $\mathbb{R}^n$ is denoted by $\|\cdot\|$, the operator norm for linear maps on Euclidean space is given by $\|L\|_{op(2)} = \sup\{\|Lx\|,\ \|x\| = 1\}$, and we use the supremums norm on continuous functions mapping a compact set $K \subseteq \mathbb{R}^n$ to $\mathbb{R}^m$, namely $\|\psi\|_\infty = \sup\{\|\psi(x)\|,\ x \in K\}$.

By assumption, $B = \frac{\partial F}{\partial y}(0, 0)$ is invertible, so we can define

$$G(x, y) = y - B^{-1}F(x, y).$$

Note that $G(x, y) = y$ if and only if $F(x, y) = 0$, and we indeed made the transition from F to G in order to apply the Banach Fixed Point Theorem to show that for each x close to 0 there exists a unique y with $G(x, y) = y$. Then we shall call $y = g(x)$.

To this end, note first that a simple application of rules for derivatives, we have for $(x, y) \in U$,

$$\frac{\partial G}{\partial y}(x, y) = \mathrm{Id}_{m\times m} - B^{-1}\frac{\partial F}{\partial y}(x, y) \in \mathcal{L}(\mathbb{R}^n \times \mathbb{R}^m, \mathbb{R}^m),$$

and, in particular, $\frac{\partial G}{\partial y}(0,0) = \mathrm{Id} - B^{-1}B = 0$. By assumption, $\frac{\partial F}{\partial y}(x,y)$ is continuous, hence, $\frac{\partial G}{\partial y}(x,y)$ is continuous, and we can find r, s with

$$\left\|\frac{\partial G}{\partial y}(x,y)\right\|_{op(2)} \leq \frac{1}{2}, \quad x \in B_s(0),\ y \in \overline{B_r(0)}, \quad B_s(0)\times\overline{B_r(0)} \subseteq U \subseteq \mathbb{R}^n \times \mathbb{R}^m. \tag{9}$$

The Mean Value Theorem, Theorem 6.22, implies immediately

$$\|G(x,y) - G(x,z)\| \leq \frac{1}{2}\|y-z\|, \quad y,z \in \overline{B_r(0)}. \tag{10}$$

To apply the Banach Fixed Point Theorem, Theorem 4.41, we need to ensure that $G(x,\cdot) : A \longrightarrow A$ for some closed set A in the complete metric space $\mathbb{R}^n$. To this end, observe that by continuity of G there exists $0 < t < s$ with

$$\|G(x,0)\| \leq \frac{r}{2}, \quad x \in \overline{B_t(0)}.$$

With $z = 0 \in \overline{B_r(0)}$ in (10), we conclude

$$\|G(x,y)\| \leq \frac{1}{2}\|y-0\| + \|G(x,0)\| \leq \frac{r}{2} + \frac{r}{2} = r, \quad y \in \overline{B_r(0)}, \tag{11}$$

so indeed, for each $x \in \overline{B_t(0)}$, the function $G(x,\cdot)$ is a contraction on the closed subset $A = \overline{B_r(0)}$ of the complete metric space $\mathbb{R}^m$.

The Banach Fixed Point theorem therefore guarantees for $x \in \overline{B_t(0)}$ a unique $y \in \overline{B_r(0)}$ with $G(x,y) = y$, so $F(x,y) = 0$. This defines a function $g : \overline{B_t(0)} \longrightarrow \overline{B_r(0)}$. Note that for the unique solution at $x = 0$ is already known, namely, $y = 0$, so $g(0) = 0$, and g passes through the origin.

To see that the function g defined above is continuous, we shall apply the Banach Fixed Point Theorem now to function $\mathcal{G}$ defined on a closed subset $\mathcal{A}$ of the complete metric space of continuous functions $C(\overline{B_t(0)}, \mathbb{R}^m)$ which is equipped with norm $\|\psi\|_\infty = \sup\{\|\psi(x)\|_2,\ x \in \overline{B_t(0)}\}$. (This is well defined since compactness of $\overline{B_t(0)}$ implies that all functions in $C(\overline{B_t(0)}, \mathbb{R}^m)$ are bounded.)

We set $\mathcal{A} = \{\psi \in C(\overline{B_t(0)}, \mathbb{R}^m) \text{ with } \|\psi\|_\infty \leq r\}$ and define

$$\mathcal{G} : \mathcal{A} \longrightarrow \mathcal{A}, \quad \psi(\cdot) \mapsto G(\cdot, \psi(\cdot)),$$

so $\mathcal{G}\psi$ is defined by $(\mathcal{G}\psi)(x) = G(x,\psi(x))$. It is crucial to observe that $\mathcal{A}$ is closed and $\mathcal{G}$ is well defined, that is $\mathcal{G}\psi \in \mathcal{A}$ for $\psi \in \mathcal{A}$. The latter follows from G being continuous and (11), that is, $\|\psi\|_\infty \leq r$ implies $\|\mathcal{G}\psi\|_\infty \leq r$. Now, $\mathcal{G}$ is indeed a contraction on $\mathcal{A}$ since.

$$\|\mathcal{G}\psi - \mathcal{G}\phi\|_\infty = \sup_{x\in\overline{B_t(0)}} \|G(x,\psi(x)) - G(x,\phi(x))\| \leq \sup_{x\in\overline{B_t(0)}} \frac{1}{2}\|\psi(x) - \phi(x)\| = \frac{1}{2}\|\psi - \phi\|_\infty.$$

The Banach Fixed Point Theorem now implies the existence of a unique function $\phi \in \mathcal{A} \subseteq C(\overline{B_t(0)}, \mathbb{R}^m)$ with $G(x,\phi(x)) = (G\phi)(x) = \phi(x)$. Since for each $x \in \overline{B_t(0)}$ there exists exactly one y with $G(x,y) = y$ by our first application of the Banach Fixed Point Theorem, we must have $\phi(x) = y = g(x)$, so g is continuous on $\overline{B_t(0)}$.

In the next step, we shall show that there exists $0 < \epsilon < t$ so that g is differentiable on $B_\epsilon(0)$, that is, for $u \in B_\epsilon(0)$, there exists a linear operator $(Dg)_u \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ with

$$g(x) = g(u) + (Dg)_u(x-u) + R_g(x), \qquad \frac{R_g(x)}{\|x-u\|} \to 0 \text{ as } x \to u.$$

To start things off, we compute the one and only candidate for $(Dg)_u$. We have $P(x) = F(x, g(x)) = 0$ for $x \in B_t(0)$, so P is constant. If g is differentiable at u, then the differentiability of F together with the chain rule implies the equality of linear maps

$$\begin{aligned} 0 = (DP)_u = (DF)_{(u,g(u))} \circ \begin{bmatrix} \mathrm{Id}_{n\times n} \\ (Dg)_u \end{bmatrix} &= \left[\frac{\partial F}{\partial x}(u,g(u)), \frac{\partial F}{\partial y}(u,g(u))\right] \circ \begin{bmatrix} \mathrm{Id}_{n\times n} \\ (Dg)_u \end{bmatrix} \\ &= \frac{\partial F}{\partial x}(u,g(u)) + \frac{\partial F}{\partial y}(u,g(u)) \circ (Dg)_u. \end{aligned}$$

Note that F being C^1 implies that all partials depend continuously on (x,y), and hence, $\frac{\partial F}{\partial y}(u,g(u))$ depends continuously on u. Invertibility of $\frac{\partial F}{\partial y}(u,g(u))$ is characterized by its determinant being non-zero. Since the determinant depends continuously on $\frac{\partial F}{\partial y}(u,g(u))$, we have that there exists $0 < \epsilon < t$ with $\frac{\partial F}{\partial y}(u,g(u))$ is invertible for $u \in B_\epsilon(0)$.

For $u \in B_\epsilon(0)$, we must have by the computation above

$$(Dg)_u = -\left(\frac{\partial F}{\partial y}(u,g(u))\right)^{-1} \circ \frac{\partial F}{\partial x}(u,g(u)). \tag{12}$$

The proof that g is differentiable at u with $(Dg)_u$ satisfying (12) under the assumption that $\frac{\partial F}{\partial y}(u,g(u))$ is invertible does not depend on u and we shall assume for convenience that $u = 0$.

Set $A = \frac{\partial F}{\partial x}(0,0)$ and $B = \frac{\partial F}{\partial y}(0,0)$. The differentiability of F implies that

$$\begin{aligned} F(x,y) &= F(0,0) + (DF)_{(0,0)}\big((x,y) - (0,0)\big) + R_F(x,y) \\ &= Ax + By + R_F(x,y), \qquad \frac{R_F(x,y)}{\|(x,y)\|} \to 0 \text{ as } (x,y) \to (0,0). \end{aligned}$$

Inserting for y the function $g(x)$, this gives

$$0 = F(x, g(x)) = Ax + Bg(x) + R_F(x, g(x)),$$

and resorting as before, we have

$$g(x) = -B^{-1}Ax + B^{-1}R_F(x, g(x)).$$

For differentiability of g at 0, it remains to show $R_g(x) = B^{-1}R_F(x, g(x))$ satisfies $\frac{R_g(x)}{\|x\|} \to 0$ as $x \to 0$.

To this end, we shall first show that there exists $K > 0$ such that $\|g(x)\| \le K\|x\|$ for x small. We set $\alpha = \|B^{-1}A\|_{op(2)}$ and $\beta = \|B^{-1}\|_{op(2)}$. Choose $0 < \mu < t, r$ so that $\|R_F(x,y)\| \le \frac{1}{2\beta}(\|x\| + \|y\|)$ whenever $\|(x,y)\| < \mu$. By continuity of g, we can pick $0 < \delta \le \mu$ so that $\|g(x)\| < \mu$ whenever $\|x\| < \delta$. Now, we compute

$$\|g(x)\| \le \| - B^{-1}A\|_{op(2)}\|x\| + \|B^{-1}\|_{op(2)}\|R_F(x, g(x))\| \le \alpha\|x\| + \frac{1}{2}(\|x\| + \|g(x)\|)$$

which, by subtracting $\frac{1}{2}\|g(x)\|$ on both sides, implies

$$\|g(x)\| \le (2\alpha + 1)\|x\| = K\|x\|$$

for $x \in B_\delta(0)$.

For $R_g(x)$,

$$\begin{aligned} 0 \le \frac{\|R_g(x)\|}{\|x\|} = \frac{\|B^{-1}R_F(x, g(x))\|}{\|x\|} &\le \beta \frac{\|R_F(x, g(x))\|}{\frac{1}{K+1}\|x\| + \frac{K}{K+1}\|x\|} \\ &\le \beta \frac{\|R_F(x, g(x))\|}{\frac{1}{K+1}\|x\| + \frac{1}{K+1}\|g(x)\|} \\ &\le (K+1)\beta \frac{\|R_F(x, g(x))\|}{\|(x, g(x))\|} \to 0 \quad \text{as} \quad x \to 0, \end{aligned}$$

where we used $g(x) \to 0$ as $x \to 0$ and hence $\|(x, g(x))\| \to 0$ as $x \to 0$.

Recall, we first showed the existence of g on $B_s(0)$, then continuity of g on $B_t(0)$, $t \le s$, and finally, differentiability of g on $B_\epsilon(0)$. It still remains to observe that the derivative Dg is continuous on $B_\epsilon(0)$, so $g \in C^1(B_\epsilon(0))$. But this follows directly from the continuity of $\left(\frac{\partial F}{\partial y}(u, g(u))\right)^{-1}$ and $\frac{\partial F}{\partial x}(u, g(u))$ by means of (12). (Here, we use also that if $B(u)$ is a family of invertible matrices depending continuously on u, then $(B(u))^{-1}$ depends also continuously on u. For example, this follows from Cramer's rule.)

Similarly, if $f \in C^r(U)$, $r \ge 0$, then $\frac{\partial F}{\partial y}(u, g(u))$ and $\frac{\partial F}{\partial x}(u, g(u))$ is $C^{r-1}(B_\epsilon(0)$. Arguing with Cramer's rule as above, we obtain that $\left(\frac{\partial F}{\partial y}(u, g(u))\right)^{-1}$ is $C^{r-1}(B_\epsilon(0))$, so the right hand side of (12) is in $C^{r-1}(B_\epsilon(0))$, that implies for the left hand side $Dg \in C^{r-1}(B_\epsilon(0))$ and by definition $g \in C^r(B_\epsilon(0))$. □

Note that in the proof we constructed $V = B_\epsilon(a)$ and $W = B_r(f(a))$. In many applications, we seek the largest possible set V as domain for the implicit function (and a large W where we have unique solutions of $F(x, y) = c$.

Corollary 6.43. *Let $U \subseteq \mathbb{R}^n \times \mathbb{R}^m$ be open and $f : U \longrightarrow \mathbb{R}^m$ be given with $f \in C^1(U)$. Let $(a, b) \in U$ and $B = \frac{\partial f}{\partial y}(a, b)$ be invertible and let and $g : B_\epsilon \longrightarrow \mathbb{R}^m$, $g \in C^1(B_\epsilon(a))$, be the function implicitly defined by $f(x, y) = f(a, b)$. Then*

$$\frac{\partial g}{\partial x}(x) = -\left(\frac{\partial f}{\partial y}(x, g(x))\right)^{-1} \frac{\partial f}{\partial x}(x, g(x)) \quad \textit{for } x \in B_\epsilon(a)\,.$$

Proof. See proof of the Implicit Function Theorem. □

Remark 6.44. Note that in the one dimensional case, for a surjective and differentiable function $f : (a, b) \longrightarrow (c, d)$ with f' continuous and $f'(x) \neq 0$ for all $x \in (a, b)$, we "get for free" that f is

- injective since f is strictly increasing or decreasing. So f^{-1} is well defined on (c, d);
- homeomorph, that is, bijective, continuous, and with continuous inverse f^{-1} on (c, d);
- diffeomorph, that is, bijective with f and f^{-1} differentiable on (a, b) respectively (c, d). In addition, $(f^{-1})'$ is continuous as f' is;

and we have $(f^{-1})'(y) = \dfrac{1}{f'(f^{-1}(y))}$ for $y \in (c, d)$.

This result does not extend to higher dimensions, for example, consider $f : \mathbb{R} \to \mathbb{R}^2$, $x \mapsto (\cos(x), \sin(x))$.

Theorem 6.45. Inverse Function Theorem. *Let $W \subseteq \mathbb{R}^n$ be open and $f : W \longrightarrow \mathbb{R}^n$ be given with $f \in C^r(W)$, $1 \leq r \leq \infty$. If $(Df)_\beta \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^n)$, $\beta \in W$ is an isomorphism, then exists an open neighborhood $\widetilde{W}$ of β and an an open neighborhood $\widetilde{V}$ of $f(\beta)$ such that $f : \widetilde{W} \longrightarrow \widetilde{V}$ is a C^r–diffeomorphism with*

$$\frac{\partial f^{-1}}{\partial x} = -\left(\frac{\partial f}{\partial y}\right)^{-1} \tag{13}$$

on $\widetilde{V}$.

Proof. Not surprisingly, we give a proof based on the Implicit Function Theorem. To this end, we define the auxiliary function

$$F : U = \mathbb{R}^n \times W \longrightarrow \mathbb{R}^n, \quad F(x, y) = x - f(y),$$

and observe that with $a = f(\beta)$, we have $F(a, \beta) = 0$. Moreover, $f \in C^r(W)$ implies $F \in C^r(\mathbb{R}^n \times W)$ and, using the notation from the implicit function theorem, we have $\frac{\partial F}{\partial y}(a, \beta) = -(Df)_\beta$ which is invertible by hypothesis. Using the Implicit Function Theorem, we learn of the existence of $V', W' \subseteq \mathbb{R}^n$ open and $g \in C^r(V')$ with

$$0 = F(x, g(x)) = x - f(g(x)),$$

where for $x \in V'$, $y = g(x)$ is the unique point in W' solving $0 = F(x, y) = x - f(y)$, that is, $f(y) = x$. So $g : V' \longrightarrow W'$, but note that g is not necessarily onto.

Since f is continuous and V' open with $f(\beta) \in V'$, there exists an open set $\widetilde{W} \subseteq W'$ with $\beta \in \widetilde{W}$ and $f(\widetilde{W}) \subseteq V'$. Now, $\widetilde{V} = f(\widetilde{W}) = g^{-1}(\widetilde{W})$ is open and the result is proven. □

A common problem in optimization is to maximize a real valued objective function based on one or multiple constraints. That is, we may be confronted with the problem

$$\text{(14)} \qquad \text{Maximize } \ h(x) \ \text{ with } \ x \in K \subseteq \mathbb{R}^n \ \text{ subject to } \ f(x) = c.$$

The following result validates the Lagrange multiplier method which you may encountered in calculus. As is generally the case in maximization problems, the result gives a necessary condition on $a \in K$ to be a local maximum for the problem (14), namely,

$$h(a) > h(x) \ \text{ for } \ x \in B_\epsilon(a) \cap K \text{ with } \ f(x) = c.$$

In practice, we shall first find all x that satisfy the hypothesis of the following result, and then analyze them individually to see which x of the candidates solves (14).

Note that the result requires that both, objective and constrain function are continuously differentiable.

Theorem 6.46. Lagrange Multipliers. *Let $f, h \in C^1(U)$, $U \subseteq \mathbb{R}^n$ open, be given with $h(a) > h(x)$ for all $x \in B_\epsilon(a) \cap \{x : \ f(x) = 0\}$, $\epsilon > 0$. If $\nabla f(a) = (Df)_a \neq 0$, then we can conclude that $\nabla f(a) = \lambda \nabla g(a)$ for some $\lambda \in \mathbb{R}$. Such λ is called Lagrange Multiplier.*

Proof. W.l.o.g., assume that $\frac{\partial f}{\partial x_n} \neq 0$ and consider $U \subseteq \mathbb{R}^{n-1} \times \mathbb{R}$ (else, permute the x_i). In preparation to applying the Implicit Function Theorem, we rename $y = x_n$, $x' = (x_1, \ldots, x_{n-1})$, $b = a_n$, and $a' = (a_1, \ldots, a_{n-1})$, and observe that by hypothesis the linear map on $\mathbb{R}$ given by

$$B = \frac{\partial f}{\partial y}(a', b) = \frac{\partial f}{\partial x_n}(a)$$

is invertible. Hence, the implicit function theorem provides a neighborhood $B_\epsilon(a')$ and a function $g : B_{\epsilon(a')} \to \mathbb{R}$ with

$$0 = F(x_1, \ldots, x_{n-1}) = f(x_1, \ldots, x_{n-1}, g(x_1, \ldots, x_{n-1})) = 0, \quad (x_1, \ldots, x_{n-1}) \in B_\epsilon(a').$$

After defining the function $\varphi : B_\epsilon(a') \to \mathbb{R}^m$ by

$$\varphi(x_1, \ldots, x_{n-1}) = (x_1, \ldots, x_{n-1}, g(x_1, \ldots, x_{n-1})),$$

we first observe that

$$[(D\varphi)_{x'}] = \begin{pmatrix} 1 & 0 & 0 & \ldots & 0 & 0 \\ 0 & 1 & 0 & \ldots & 0 & 0 \\ 0 & 0 & 1 & \ldots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \ldots & 1 & 0 \\ 0 & 0 & 0 & \ldots & 0 & 1 \\ \frac{\partial \varphi}{\partial x_1}(x') & \frac{\partial \varphi}{\partial x_2}(x') & \frac{\partial \varphi}{\partial x_3}(x') & \ldots & \frac{\partial \varphi}{\partial x_{n-2}}(x') & \frac{\partial \varphi}{\partial x_{n-1}}(x') \end{pmatrix} \in \mathbb{R}^{n \times (n-1)}$$

As F is constant, we can compute in a neighborhood of a',

$$0 = \nabla F(x') = [(DF)_{x'}] = [(D(f \circ \varphi))_{x'}] = [(Df)_{\varphi(x')} \circ (D\varphi)_{x'}] = [(Df)_{\varphi(x')}] \cdot [(D\varphi)_{x'}]$$

$$= \left(\frac{\partial f}{\partial x_1}(x', g(x')), \dots, \frac{\partial f}{\partial x_n}(x', g(x'))\right) \begin{pmatrix} 1 & 0 & 0 & \dots & 0 & 0 \\ 0 & 1 & 0 & \dots & 0 & 0 \\ 0 & 0 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 1 & 0 \\ 0 & 0 & 0 & \dots & 0 & 1 \\ \frac{\partial \varphi}{\partial x_1}(x') & \frac{\partial \varphi}{\partial x_2}(x') & \frac{\partial \varphi}{\partial x_3}(x') & \dots & \frac{\partial \varphi}{\partial x_{n-2}}(x') & \frac{\partial \varphi}{\partial x_{n-1}}(x') \end{pmatrix}$$

$$= \left(\frac{\partial f}{\partial x_1}(x', g(x')) + \frac{\partial f}{\partial x_n}(x', g(x'))\frac{\partial \varphi}{\partial x_1}(x'), \dots, \frac{\partial f}{\partial x_{n-1}}(x', g(x')) + \frac{\partial f}{\partial x_n}(x', g(x'))\frac{\partial \varphi}{\partial x_{n-1}}(x')\right),$$

and

$$\frac{\frac{\partial f}{\partial x_i}(a', g(a'))}{\frac{\partial f}{\partial x_n}(a', g(a'))} = -\frac{\partial \varphi}{\partial x_i}(a'), \quad i = 1, \dots, n-1. \tag{15}$$

Note that since h has at a' a local extremum subject to $f(a) = 0$, we obtain that

$$0 = \nabla H(a') = [(Dh)_{\varphi(a')}] \cdot [(D\varphi)_{a'}]$$
$$= \left(\frac{\partial h}{\partial x_1}(a', g(a')) + \frac{\partial h}{\partial x_n}(a', g(a'))\frac{\partial \varphi}{\partial x_1}(a'), \dots, \frac{\partial h}{\partial x_{n-1}}(x', g(x')) + \frac{\partial h}{\partial x_n}(a', g(a'))\frac{\partial \varphi}{\partial x_{n-1}}(a')\right).$$

and

$$\frac{\frac{\partial h}{\partial x_i}(a', g(a'))}{\frac{\partial h}{\partial x_n}(a', g(a'))} = -\frac{\partial \varphi}{\partial x_i}(a'), \quad i = 1, \dots, n-1.$$

Inserting (15) gives

$$\frac{\partial h}{\partial x_i}(a', g(a')) = \frac{\frac{\partial f}{\partial x_i}(a', g(a'))}{\frac{\partial f}{\partial x_n}(a', g(a'))}\frac{\partial h}{\partial x_n}(a', g(a')) \quad i = 1, \dots, n-1.$$

and the result follows with

$$\lambda = \frac{\frac{\partial h}{\partial x_n}(a)}{\frac{\partial f}{\partial x_n}(a)}.$$

□

7. INTEGRATION ON $\mathbb{R}^D$

We shall focus on the case $d = 2$, that is, we shall discuss integrals $\int_A f(x)\,dx$ where $A \subseteq \mathbb{R}^2$ and $f : A \longrightarrow \mathbb{R}$. This way we avoid some notational difficulties. Generalizations to higher dimensions are straightforward.

7.1. Essentials

Definition 7.1. Consider a rectangle (interval) $[a, b] \times [c, d] \subset \mathbb{R}^2$, $-\infty < a < b < \infty$ and $-\infty < c < d < \infty$ and partitions $P = \{a = x_0, x_1, \ldots, x_{m-1}, x_m = b\}$ and $Q = \{c = y_0, y_1, \ldots, y_{n-1}, y_n = d\}$ of $[a, b]$ and $[c, d]$ respectively. Then

$$G = \{R_{ij} = [x_{i-1}, x_i] \times [y_{j-1}, y_j],\ i = 1, \ldots, m,\ j = 1, \ldots, n\}$$

is a *grid* of rectangles in R. For a sample set

$$S = \{(s_{ij}, t_{ij}) \in R_{ij},\ i = 1, \ldots, m,\ j = 1, \ldots, n\}$$

and $f : R \longrightarrow \mathbb{R}$ we define the *Riemann sum*

$$\mathcal{R}(f, G, S) = \sum_{i=1}^{m} \sum_{j=1}^{n} f(s_{ij}, t_{ij})|R_{ij}|,$$

where $|R_{ij}| = (x_i - x_{i-1})(y_j - y_{j-1})$ is the area of the rectangle R_{ij}, $i = 1, \ldots, m,\ j = 1, \ldots, n$.

Note that in higher dimensions, we shall call a set $[a_1, b_1] \times [a_2, b_2] \times \ldots \times, [a_n, b_n]$ generalized rectangles or simply intervals.

Definition 7.2. A function $f : R \longrightarrow \mathbb{R}$ is *Riemann integrable* if for some number $I \in \mathbb{R}$ such that for all $\epsilon > 0$ exists a $\delta > 0$ such that $|I - \mathcal{R}(f, G, S)| < \epsilon$ whenever $\operatorname{mesh}(G) = \max_{R_{ij} \in G} \operatorname{diam} R_{ij} < \delta$.

The number I is called *Riemann integral* of f on R and is denoted by $I = \int f = \int_R f(x, y)\,d(x, y)$, or as $\int_R f(x)\,dx$ where x is considered to be a variable in $\mathbb{R}^2$.

The space of all Riemann integrable functions on R is denoted by $\mathcal{R}(R) = \{f : R \longrightarrow \mathbb{R},\quad f \text{ is Riemann integrable }\}$.

Definition 7.3. The lower and upper sums of a bounded function f with respect to the grid G are

$$L(f, G) = \sum_{i=1}^{m} \sum_{j=1}^{n} m_{ij}|R_{ij}| \quad \text{and} \quad U(f, G) = \sum_{i=1}^{m} \sum_{j=1}^{n} M_{ij}|R_{ij}|,$$

where $m_{ij} = \inf f(R_{ij})$ and $M_{ij} = \sup f(R_{ij})$, $i = 1, \ldots, m,\ j = 1, \ldots, n$.

The *lower integral* of f is $\underline{\int} f = \sup\{L(f, G),\ G \text{ grid in } R\}$, and the *upper integral* of f is $\overline{\int} f = \inf\{U(f, G),\ G \text{ grid in } R\}$

Theorem 7.4. *Let R be a rectangle in $\mathbb{R}^2$.*

i. If $f : R \longrightarrow \mathbb{R}$ is Riemann integrable, then f is bounded.

ii. $\mathcal{R}(R)$ is a vector space.

iii. A constant function $f : R \longrightarrow \mathbb{R}$, $(x, y) \mapsto k$, $k \in \mathbb{R}$, is Riemann integrable and $\int_R f(x, y)d(x, y) = k|R|$.

iv. For $f, g \in \mathcal{R}(R)$ with $f \leq g$ we have $\int f \leq \int g$.

v. For $f : R \longrightarrow \mathbb{R}$ bounded, we have $\underline{\int} f \leq \overline{\int} f$.

vi. A bounded function $f : R \longrightarrow \mathbb{R}$ is Riemann integrable if and only if $\underline{\int} f = \overline{\int} f$.

Proof. The proofs are identical to the respective proofs in the one dimensional setting. □

Definition 7.5. A set $Z \subset \mathbb{R}^2$ is a *zeroset* if for all $\epsilon > 0$ there exists a countable family of open rectangles $\{S_k\}_{k\in\mathbb{N}}$ such that $Z \subseteq \bigcup_{k=1}^{\infty} S_k$ and $\sum_{k=1}^{\infty} |S_k| < \epsilon$.

Theorem 7.6. Multivariable Riemann–Lebesgue theorem. *A bounded function $f : R \longrightarrow \mathbb{R}$ is Riemann integrable if and only if the set of discontinuities of f is a zeroset.*

Proof. The proof is analogous to the proof in the one dimensional setting. □

Now, the first result that is multivariate in nature.

Theorem 7.7. Fubini's theorem. *Let $f : R = [a, b] \times [c, d] \longrightarrow \mathbb{R}$ be Riemann integrable.*

i. The functions

$$\underline{F} : [c, d] \longrightarrow \mathbb{R}, \quad \underline{F}(y) = \underline{\int_a^b} f(x, y)\, dx \quad \text{and} \quad \overline{F} : [c, d] \longrightarrow \mathbb{R}, \quad \overline{F}(y) = \overline{\int_a^b} f(x, y)\, dx$$

are Riemann integrable on $[c, d]$ and we have

$$\int_c^d \underline{\int_a^b} f(x, y)\, dx\, dy = \int_c^d \underline{F}(y)\, dy = \int_R f = \int_c^d \overline{F}(y)\, dy = \int_c^d \overline{\int_a^b} f(x, y)\, dx\, dy$$

ii. There exists a zeroset $Y \subseteq [c, d]$ such that the y-section $f_y(\cdot) = f(\cdot, y)$ is Riemann integrable on $[a, b]$ for all $y \in [c, d] \setminus Y$. We set

$$\widetilde{f}(x, y) = \begin{cases} f(x, y), & \text{if } y \notin Y \\ \underline{F}(y), & \text{if } y \in Y \end{cases}$$

and obtain

$$\int_R f = \int_R \widetilde{f} = \int_c^d \int_a^b \widetilde{f}(x,y)\,dx\,dy.$$

Since $\widetilde{f}(x,y) = f(x,y)$ *on* $R \setminus Z$ *where* Z *is a zeroset, it is customary not to distinguish between* f *and* $\widetilde{f}$*. Hence, we shall simply write*

$$\int_c^d \int_a^b f(x,y)\,dx\,dy = \int_R f = \int_a^b \int_c^d f(x,y)\,dy\,dx.$$

Proof. i. Let us first fix partitions $P = \{a = x_0, x_1, \ldots, x_{m-1}, x_m = b\}$ and $Q = \{c = y_0, y_1, \ldots, y_{n-1}, y_n = d\}$ of $[a,b]$ and $[c,d]$ respectively, as well as the corresponding grid $G = \{R_{ij} = [x_{i-1}, x_i] \times [y_{j-1}, y_j],\ i = 1, \ldots, m,\ j = 1, \ldots, n\}$ in R.

Observe that for $\widetilde{y} \in [y_{j-1}, y_j]$ fixed and $m_{ij} = \inf f(R_{ij})$ we have

$$\begin{aligned}
\sum_i m_{ij}(x_i - x_{i-1}) &= \sum_i \inf\left\{f(x,y),\ x \in [x_{i-1}, x_i],\ y \in [y_{j-1}, x_j]\right\}(x_i - x_{i-1}) \\
&\leq \sum_i \inf\left\{f(x,\widetilde{y}),\ x \in [x_{i-1}, x_i]\right\}(x_i - x_{i-1}) \\
&= L(f_{\widetilde{y}}, P) \\
&\leq \underline{\int_a^b} f(x,\widetilde{y})\,dx = \underline{F}(\widetilde{y}).
\end{aligned}$$

As this holds for each $\widetilde{y} \in [y_{j-1}, y_j]$, we have also

$$\sum_i m_{ij}(x_i - x_{i-1}) \leq \inf\{\underline{F}(\widetilde{y}),\ \widetilde{y} \in [y_{j-1}, y_j]\} = m_j(\underline{F}),$$

and

$$L(f,G) = \sum_{ij} m_{ij}(x_i - x_{i-1})(y_j - y_{j-1}) \leq \sum_j m_j(\underline{F})\,(y_j - y_{j-1}) = L(\underline{F}, Q).$$

Similarly, we can show $U(\overline{F}, Q) \leq U(f,G)$ and summarize

$$L(f,G) \leq L(\underline{F}, Q) \leq U(\underline{F}, Q) \leq U(\overline{F}, Q) \leq U(f,G),$$

and, taking infimum and supremum and using integrability of f,

$$\begin{aligned}
\int f = \inf_G L(f,G) &\leq \inf_G L(\underline{F}, Q) = \inf_Q L(\underline{F}, Q) \leq \inf_Q U(\underline{F}, Q) \\
&\leq \sup_Q U(\underline{F}, Q) \leq \sup_Q U(\overline{F}, Q) = \sup_G U(\overline{F}, Q) \leq \sup_G U(f,G) = \int f.
\end{aligned}$$

We conclude equality between all terms, in particular,

$$\underline{\int_c^d} \underline{F}(y)\,dy = \inf_Q L(\underline{F}, Q) = \sup_Q U(\underline{F}, Q) = \overline{\int_c^d} \underline{F}(y)\,dy = \int f,$$

so $\underline{F}$ is integrable with $\int_c^d \underline{F}(y)\,dy = \int f$. Similarly, we obtain that $\overline{F}$ is integrable with $\int_c^d \overline{F}(y)\,dy = \int f$.

ii. Clearly, $\overline{F} \geq \underline{F}$, and by *i.*, we have

$$0 = \int f - \int f = \int_c^d \overline{F}(y) - \underline{F}(y)\,dy.$$

So $\overline{F} - \underline{F}$ is a non-negative function that integrates to zero, it is not hard to see that the function then needs to be constant zero outside a zerozet Y. That is,

$$\overline{\int_a^b} f(x,y)\,dx = \overline{F}(y) = \underline{F}(y) = \underline{\int_a^b} f(x,y)\,dx, \quad y \in [c,d] \setminus Y.$$

But equality of upper and lower integral implies Riemann integrability, so the y-section $f_y(\cdot) = f(\cdot, y)$ is integrable for every $y \in [c,d] \setminus Y$.

The conclusions on $\widetilde{f}$ we leave for a homework problem. □

7.2. Jordan content

Integration is one mean to measure the size (area, volume) of a set in $\mathbb{R}^n$, for example, $\int_0^1 x\,dx$ measures the size of an isoceles right triangle. The ultimate goal would be to assign a size to all sets in $\mathbb{R}^n$. Certainly, thats not hard: lets just say all sets have size 0. Nevertheless, once you require the "measure" to follow some rudimentary ideas, we face problems.

Theorem 7.8. *For $n = 1, 2, \dots$, there exists no function $\mu : \mathcal{P}(\mathbb{R}^n) \longrightarrow [0, \infty] \subseteq \mathbb{R}^*$ such that*

i. size adds up: $\mu(\bigcup_{i=1}^\infty E_i) = \sum_{i=1}^\infty \mu(E_i)$ if all $E_i \subseteq \mathbb{R}^n$, $i = 1, \dots, \infty$, are disjoint,

ii. size is translation invariant: $\mu(E + a) = \mu(E)$ for all $E \subseteq \mathbb{R}^n$, $a \in \mathbb{R}^n$, where $E + a = \{e + a,\ e \in E\}$, and

iii. size is nontrivial, that is, normalized by $\mu([0, 1] \times \ldots \times [0, 1]) = 1$.

To circumvent these problems without losing any of the three key properties listed above, one chooses to work only with some "nice" subsets of $\mathbb{R}^n$. The biggest breakthrough in "measure theory" was the classification of a large class of sets, so called Lebesgue measurable sets $\mathcal{L} \subsetneq \mathcal{P}(\mathbb{R}^n)$, for which a measure, that is, the Lebesgue measure, can be defined to satisfy the three properties listed above. Here, we shall be even more restrictive and only discuss those sets which have a so–called Jordan content, that is, Jordan domains.

Definition 7.9. A bounded subset in $D \subseteq \mathbb{R}^n$ is said to be a *Jordan domain* if its boundary ∂D is a zeroset.

Lemma 7.10. *Let D be a Jordan domain and $f : D \longrightarrow \mathbb{R}$ be bounded and continuous. Let R a rectangle containing D. Then $\widetilde{f} : R \longrightarrow \mathbb{R}$ with $\widetilde{f}(x) = \begin{cases} f(x), & \text{for } x \in D, \\ 0, & \text{for } x \in R \setminus D. \end{cases}$ is Riemann integrable, that is, $\widetilde{f} \in \mathcal{R}(R)$.*

Definition 7.11. Let D be a Jordan domain and R a rectangle containing D. For a continuous function $f : D \longrightarrow \mathbb{R}$ we set $\widetilde{f}(x) = \begin{cases} f(x), & \text{for } x \in D, \\ 0, & \text{for } x \notin D, \end{cases}$, $x \in R$ and define the Riemann integral of f on D as $\displaystyle\int_D f = \int_R \widetilde{f}$.

(Note that $\displaystyle\int_D f$ is well defined, that is, does not depend on the choice of R.)

Definition 7.12. Let D be a Jordan domain. The *Jordan content* (or volume) $\operatorname{vol} D$ of D is given by $\operatorname{vol} D = \displaystyle\int_D 1$.

Remark 7.13. Note that for a bounded set D, we have that D is a Jordan domain if and only if ∂D is a zeroset which holds if and only if for all $\epsilon > 0$ there exists a **finite** family of open rectangles $\{S_k\}_{k=1,\dots,N}$ such that $Z \subseteq \bigcup_{k=1}^N S_k$ and $\sum_{k=1}^N |S_k| < \epsilon$, which holds if and only if ∂D is a *Jordan zeroset* , that is, ∂D is a Jordan domain with $\operatorname{vol} \partial D = 0$.

These equivalences hold since ∂D is bounded and closed, and, hence, compact. Certainly, there are sets which are zerosets, but not Jordan zerosets, that is, consider $[0, 1] \cap \mathbb{Q} \subset \mathbb{R}$.

Remark 7.14. If D is a Jordan domain, then for any rectangle $R \supseteq D$, we have

$$\underline{\int_R} \chi(x)\,dx = \text{vol } D = \overline{\int_R} \chi(x)\,dx\,.$$

This fact is often used to define Jordan content via the equality of *inner Jordan content* (left hand side) and *outer Jordan content* (right hand side).

7.3. Change of variables

In this section we shall prove change of variable formulas in great generality.

Proposition 7.15. *Let S be a Jordan domain and $T \in \mathbb{R}^{2\times 2}$ be invertible. Then $T(S)$ is a Jordan domain and* $\operatorname{vol}(T(S)) = |\det T| \operatorname{vol}(S)$.

Proof. A linear algebra result states that every $T \in \mathbb{R}^{2\times 2}$ can be written as product of matrices of the form $\begin{pmatrix} 1 & 0 \\ 0 & \alpha \end{pmatrix}$, $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$, and $\begin{pmatrix} 1 & \beta \\ 0 & 1 \end{pmatrix}$ where $\alpha, \beta \neq 0$. For each matrix E of this form, it is easily computed that for any rectangle R we have $\operatorname{vol}(E(R)) = |\det E| \operatorname{vol}(R)$.

We shall now show that $\operatorname{vol}(E(S)) = |\det E| \operatorname{vol}(S)$ whenever S is a Jordan domain. To this end, fix S and then $\epsilon > 0$. Using integrability of χ_S, there exists a grid G with $U(\chi_S, G) - L(\chi_S, G) \leq \epsilon$, hence,

$$\begin{aligned} \operatorname{vol}(S) - \epsilon = \int \chi_S - \epsilon = \underline{\int} \chi_S - \epsilon \leq L(\chi_S, G) = \sum_{R \subset S} \operatorname{vol}(R) \\ \leq \sum_{R \cap S \neq \emptyset} \operatorname{vol}(R) = U(\chi_S, G) \leq \operatorname{vol}(S) + \epsilon. \end{aligned}$$

Clearly,

$$\sum_{R \subset S} \chi_{R^\circ} \leq \chi_S \leq \sum_{R \cap S \neq \emptyset} \chi_R$$

and, after deforming with E,

$$\sum_{R \subset S} \chi_{E(R)^\circ} \leq \chi_{E(S)} \leq \sum_{R \cap S \neq \emptyset} \chi_{E(R)}.$$

Integrating and using monotonicity of the integral leads to

$$\begin{aligned} |\det(E)|(\operatorname{vol}(S) - \epsilon) &\leq |\det(E)| \sum_{R \subset S} \operatorname{vol}(R) \\ &= \sum_{R \subset S} \operatorname{vol}(E(R)) = \sum_{R \subset S} \int \chi_{E(R)^\circ} = \sum_{R \subset S} \underline{\int} \chi_{E(R)^\circ} \\ &\leq \underline{\int} \chi_{E(S)} \leq \overline{\int} \chi_{E(S)} \leq \sum_{R \cap S \neq \emptyset} \operatorname{vol}(E(R)) \\ &= |\det(E)| \sum_{R \cap S \neq \emptyset} \operatorname{vol}(R) \leq |\det(E)|(\operatorname{vol}(S) + \epsilon). \end{aligned}$$

Since ϵ can be chosen arbitrarily small, we conclude

$$|\det(E)|\,\mathrm{vol}(S) = \underline{\int} \chi_{E(S)} = \overline{\int} \chi_{E(S)} \leq |\det(E)|\,\mathrm{vol}(S),$$

which implies integrability of $\chi_{E(S)}$ and $\mathrm{vol}(E(S)) = |\det E|\,\mathrm{vol}(S)$.

For generic $T \in \mathbb{R}^{2\times 2}$, we write $T = E_1 E_2 \dots E_n$ and observe that

$$\begin{aligned}\mathrm{vol}(T(S)) &= \mathrm{vol}(E_1E_2\dots E_n(S)) = |\det E_1|\,\mathrm{vol}(E_2\dots E_n(S))\\ &= \dots = |\det E_1||\det E_2|\dots|\det E_n|\,\mathrm{vol}(S)|\det E|\,\mathrm{vol}(S)\\ &= |\det E_1E_2\dots E_n|\,\mathrm{vol}(S) = |\det T|\,\mathrm{vol}(S).\end{aligned}$$ □

Lemma 7.16. *Let $\psi : U \longrightarrow \mathbb{R}^2$, $U \subseteq \mathbb{R}^2$ open, be a C^1 function with $0 \in U$ and $\psi(0) = 0$. If $\|(D\psi)_u - I\|_{op(p)} \leq \epsilon$ for all $u \in U$, then for all r such that $B^p_r(0) = \{x : \|x\|_p < r\} \subseteq U$ we have $\psi(B^p_r(0)) \subseteq B^p_{r(1+\epsilon)}(0)$.*

Proof. By the C_1 Mean Value Theorem, we have

$$\begin{aligned}\|\psi(u)\|_p = \|\psi(u)-\psi(0)\|_p &= \Big\|\Big(\int_0^1 (D\psi)_{tu}dt\Big)(u-0)\Big\|_p = \Big\|\Big(\Big(\int_0^1 (D\psi)_{tu} - I dt\Big)u\Big) + u\Big\|_p\\ &\leq \Big\|\Big(\Big(\int_0^1 (D\psi)_{tu} - I dt\Big)\Big\|_{op(p)}\|u\|_p + \|u\|_p\\ &\leq \Big(\Big(\int_0^1 \big\|(D\psi)_{tu} - I\big\|_{op(p)}dt\Big)\Big\|\|u\| + \|u\| \leq (1+\epsilon)\|u\|_p.\end{aligned}$$ □

Lemma 7.17. *If $h : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ is Lipschitz, that is, for some $L \geq 0$, $\|h(x) - h(y)\| \leq L\|x - y\|$ for all $x, y \in \mathbb{R}^2$, then* $\mathrm{vol}(h(Z)) = 0$ *whenever* $\mathrm{vol}(Z) = 0$.

Proof. The proof is straight forward. Fix ϵ and choose a cover of open rectangles R_k with $\sum \mathrm{vol}(R_k) \leq \epsilon/(2L^2)$. Then observe that the Lipschitz condition allows one to cover $h(S_k)$ with a rectangle $\widetilde{S}_k$ with $\mathrm{vol}(\widetilde{S}_k) \leq 2L^2\,\mathrm{vol}(S_k)$. The result follows. □

Theorem 7.18. Change of variables. *Let $U, W \subseteq \mathbb{R}^2$ open and $\varphi : U \longrightarrow W$ be a C^1 diffeomorphism. For $f : W \longrightarrow \mathbb{R}$ Riemann integrable on a rectangle R in U we have*

$$\int_R f \circ \varphi\ |\det D\varphi| = \int_{\varphi(R)} f.$$

Proof. Note that $\varphi(R)$ is a bounded (in fact, a compact) set, so there exists some $\widetilde{R}$ with $\varphi(R) \subseteq \widetilde{R}$ and the right hand side above is defined to be $\int_{\varphi(R)} f = \int_{\widetilde{R}} \chi_{\varphi(R)} f$. With D denoting the set of discontinuities of f, observe that the set of discontinuities of $\chi_{\varphi(R)} f$ is the union of D and the boundary set $\partial\varphi(R)$. We have $\partial\varphi(R) = \varphi(\partial R)$ since φ is an homeomorphism. The C^1 Mean Value Theorem implies that φ is Lipschitz, hence, $\varphi(\partial R)$ is a zero set. We conclude that $\chi_{\varphi(R)} f$ is Riemann integrable on $\widetilde{R}$ and the RHS is well defined.

Since $D\varphi$ is continuous, so is $|\det D\varphi|$. The set of discontinuities of $f \circ \varphi \,|\det D\varphi|$ in R is therefore $\varphi^{-1}D$, which is a zero set since φ^{-1} is Lipschitz on R and D is a zeroset. So the LHS is well defined as well.

It remains to show that LHS=RHS. To this end, fix $\epsilon > 0$. Let $G = \{R_{ij}\}$ be a grid on R of mesh $r > 0$. (We will determine the value of r depending on ϵ later.) With z_{ij} being the center of R_{ij}, we consider the Taylor approximation T_{ij} of φ at z_{ij}, that is,

$$T_{ij} = \varphi(z_{ij}) + (D\varphi)_{z_{ij}}(z - z_{ij}).$$

Note that $S_{ij} = T_{ij}^{-1} \circ \varphi$ satisfies $S_{ij}z_{ij} = z_{ij}$ and $(DS_{ij})_{z_{ij}} = I$. If we choose r small enough, then for all R_{ij} we have

$$S_{ij}R_{ij} \subseteq (1+\epsilon)R_{ij},$$

and, applying the same argument to $\widetilde{S}_{ij} = \varphi^{-1} \circ T_{ij}$ we get further (with r possibly made even smaller) that

$$\widetilde{S}_{ij}(1+\epsilon)^{-1}R_{ij} \subseteq R_{ij}.$$

Together, this gives

$$T_{ij}(1+\epsilon)^{-1}R_{ij} \subseteq \varphi(R_{ij}) \subseteq T_{ij}(1+\epsilon)^{-1}R_{ij},$$

and

$$|\det T_{ij}|(1+\epsilon)^{-1}\operatorname{vol}(R_{ij})\operatorname{vol}(R_{ij}) \leq \operatorname{vol}\varphi(R_{ij}) \leq |\det T_{ij}|(1+\epsilon)^{-1}\operatorname{vol}(R_{ij}).$$

Setting $J_{ij} = |\det T_{ij}| = |\det(D\varphi)_{z_{ij}}|$ we obtain the area estimate

$$\frac{J_{ij}\operatorname{vol}(R_{ij})}{(1+\epsilon)^2} \leq \operatorname{vol}\varphi(R_{ij}) \leq J_{ij}\operatorname{vol}(R_{ij})(1+\epsilon)^2$$

and

$$\frac{1}{(1+\epsilon)^2} \leq \frac{\operatorname{vol}\varphi(R_{ij})}{J_{ij}\operatorname{vol}(R_{ij})} \leq (1+\epsilon)^2.$$

A simple computation then shows that

$$|\operatorname{vol}\varphi(R_{ij}) - J_{ij}\operatorname{vol}(R_{ij})| \leq 16\epsilon J_{ij}\operatorname{vol}(R_{ij}).$$

Finally, with m_{ij} and M_{ij} denoting the infinum and supremum of $f \circ \varphi$ on R_{ij}, we have

$$\sum m_{ij}\operatorname{vol}(\varphi(R_{ij})) = \int \sum m_{ij}\chi_{\varphi(R_{ij})^\circ} \leq \int_{\varphi(R)} f \leq \int \sum M_{ij}\chi_{\varphi(R_{ij})} = \sum M_{ij}\operatorname{vol}(\varphi(R_{ij}))$$

and

$$\sum m_{ij}J_{ij}\operatorname{vol}(R_{ij}) - 16\epsilon J_{ij}\operatorname{vol}(R_{ij})m_{ij} \leq \int_{\varphi(R)} f \leq \sum M_{ij}J_{ij}\operatorname{vol}(R_{ij}) + 16\epsilon J_{ij}\operatorname{vol}(R_{ij})M_{ij}.$$

Clearly, we have also

$$\sum m_{ij}J_{ij}\operatorname{vol}(R_{ij}) \leq \int_R f \circ \varphi \,|\det D\varphi| \leq \sum M_{ij}J_{ij}\operatorname{vol}(R_{ij})$$

and the difference of the left and the right term can be made arbitrarily small. Moreover, we have

$$\sum 16\epsilon J_{ij}\operatorname{vol}(R_{ij})m_{ij},\ \sum 16\epsilon J_{ij}\operatorname{vol}(R_{ij})M_{ij} \leq 16\epsilon J\operatorname{vol}(R)M,$$

where $J = \max\{|\det(D\varphi)_z|,\ z \in R\}$ and $M = \max\{|f(z)|,\ z \in R\}$, the latter is independent of the grid and can be made arbitrarily small by choosing ϵ small. □

7.4. Multivariate improper integrals

Definition 7.19. A sequence $\{C_k\}_{k\in\mathbb{N}}$ of Jordan domains is *exhaustive* if $C_k \subseteq C_{k+1}$ for all $k \in \mathbb{N}$ and $\mathrm{vol}(B_r(0) \setminus C_k) \to 0$ as $k \to \infty$ for all $r > 0$.

Definition 7.20. The not necessarily bounded set M has Jordan content $K \in [0, \infty]$ if for some exhaustive sequence $\{C_k\}$, the sets $C_k \cap M$, $k = 1, 2, 3, \ldots$, are Jordan domains with $\mathrm{vol}(C_k \cap M) \to K$ as $k \to \infty$.

Definition 7.21. We say that f is *improper (absolutely) Riemann integrable* on $M \subseteq \mathbb{R}^n$ and write $f \in \mathcal{R}(M)$, if there is an exhaustive sequence $\{C_k\}$ and $L \in \mathbb{R}^+$ such that $\int_{C_k\cap M} |f(x)|\,dx < \infty$. Then $\lim\limits_{k\to\infty} \displaystyle\int_{C_k\cap M} f(x)\,dx$ converges in $\mathbb{R}$ and we call

$$\int_M f = \lim_{k\to\infty} \int_{C_k\cap M} f(x)\,dx\,.$$

improper Riemann integral of f on M.

Certainly, one must show that all this is well defined and does not depend on the choice of the $\{C_k\}$. We leave the prove of this to the conscientious (and hopefully conscious but possibly contentious) reader.

Example 7.22. $\displaystyle\int_{\mathbb{R}^2} e^{-(x^2+y^2)} d(x, y) = \pi$ and therefore $\displaystyle\int_{-\infty}^{\infty} e^{-x^2} dx = \sqrt{\pi}$.

7.5. The Gamma function

Definition 7.23. The *Gamma function* $\Gamma : \mathbb{R}^+ \longrightarrow \mathbb{R}^+$ is given by $\Gamma(x) = \displaystyle\int_0^\infty t^{x-1}e^{-t}\,dt$.

Definition 7.24. A function $f : (a, b) \longrightarrow \mathbb{R}$ is *convex* if for all $x, y \in (a, b)$ with $x < y$ and all $0 < \lambda < 1$ we have

$$f\big(\lambda x + (1-\lambda)y\big) \leq \lambda f(x) + (1-\lambda)f(y).$$

Lemma 7.25. Hölder's inequality *Let $f, g : \mathbb{R} \longrightarrow \mathbb{C}$ be bounded with $f, g \in \mathcal{R}(\mathbb{R})$ and $p, q > 1$ with $\frac{1}{p} + \frac{1}{q} = 1$. Then*

$$\left|\int_{-\infty}^{\infty} f(x)g(x)\,dx\right| \leq \left(\int_{-\infty}^{\infty} |f(x)|^p\,dx\right)^{\frac{1}{p}} \left(\int_{-\infty}^{\infty} |g(x)|^q\,dx\right)^{\frac{1}{q}}$$

For $p = q = 2$ this is a special case of the Cauchy–Schwarz inequality .

Theorem 7.26. *The Gamma function satisfies*

i. the functional equation $f(x+1) = xf(x)$ for $x \in (0, \infty)$,

ii. $f(n+1) = n!$, and

iii. f is convex.

Moreover, the Gamma function is the only positive function satisfying i, ii, iii , that is, if f is any function satisfying i, ii, ii, then $f(x) = \Gamma(x)$ for all $x \in (0, \infty)$.

Definition 7.27. The *Beta function* $B : \mathbb{R}^+ \times \mathbb{R}^+ \longrightarrow \mathbb{R}^+$ is given by

$$B(x, y) = \int_0^\infty t^{x-1}(1-t)^{y-1}\,dt.$$

Lemma 7.28. $B(x, y) = \dfrac{\Gamma(x)\Gamma(y)}{\Gamma(x+y)}$.

Proposition 7.29. Stirling's formula. $\displaystyle\lim_{x\to\infty} \frac{\Gamma(x+1)}{\left(\frac{x}{e}\right)^x \sqrt{2\pi x}} = 1$, *and, in particular,*

$$\lim_{n\to\infty} \frac{n!}{\left(\frac{n}{e}\right)^n \sqrt{2\pi n}} = 1$$

8. ORDINARY DIFFERENTIAL EQUATIONS

This chapter follows closely Chapter II of the book Analysis 2 authored by Otto Forster.

Definition 8.1. Let $G \subseteq \mathbb{R} \times \mathbb{R}$ and $f : G \longrightarrow \mathbb{R}$, $(x, y) \mapsto f(x, y)$ be a continuous function. The formal expression

$$y' = f(x, y) \tag{16}$$

is called *first order differential equations.*

A solution to (16) is a function $\varphi : I \longrightarrow \mathbb{R}$, I being an interval, such that the graph Γ_φ of φ satisfies $\Gamma_\varphi \subseteq G$ and

$$\varphi'(x) = f(x, \varphi(x)) \quad \text{for all } x \in I.$$

Remark 8.2. The set of solutions to a first order differential equation is commonly visualized through a slope field. To obtain the slope field (direction field) of $y' = f(x, y)$, $f : G \to \mathbb{R}$, a set of points $\{(x_i, y_j)\} \subset G$ — normally placed on a regular grid — is chosen and at each point (x_i, y_j) a small line parallel to the vector $(1, f(x_i, y_j))$ is drawn. This line indicates that any solution φ passing through (x_i, y_j) has the slope $\varphi'(x_i) = y'(x_i) = f(x_i, y_j) = f(x_i, y_j)/1$ at x_i.

Definition 8.3. Let $G \subset \mathbb{R} \times \mathbb{R}^n$ and $f : G \longrightarrow \mathbb{R}^n$, $(x, y) \mapsto f(x, y)$, be a continuous function. The formal expression

$$y' = f(x, y) \tag{17}$$

is called a *system of n first order differential equations.*

A solution to (17) is a function $\varphi : I \longrightarrow \mathbb{R}^n$, I being an interval, such that the graph $\Gamma_\varphi \subseteq G$ and

$$\varphi'(x) = \begin{pmatrix} \varphi_1'(x) \\ \varphi_2'(x) \\ \vdots \\ \varphi_n'(x) \end{pmatrix} = \begin{pmatrix} f_1\big(x, \varphi_1(x), \varphi_2(x), \dots \varphi_n(x)\big) \\ f_2\big(x, \varphi_1(x), \varphi_2(x), \dots \varphi_n(x)\big) \\ \vdots \\ f_n\big(x, \varphi_1(x), \varphi_2(x), \dots \varphi_n(x)\big) \end{pmatrix} = f(x, \varphi(x)) \quad \text{for all } x \in I.$$

Definition 8.4. Let I, J be open intervals and $f : I \to \mathbb{R}$, $g : J \to \mathbb{R}$ continuous with $g(y) \neq 0$ for $y \in J$. Then we refer to the differential equation $y' = f(x)g(y)$ as *separable differential equation.*

Theorem 8.5. *Let I, J be open intervals and $f : I \to \mathbb{R}$, $g : J \to \mathbb{R}$ continuous with $g(y) \neq 0$ for $y \in J$ and let $(x_0, y_0) \in I \times J$. Further, assume $g(I) \subseteq J$.*

Then exists a unique solution $\varphi : I \to \mathbb{R}$ of the separable differential equation $y' = f(x)g(y)$ satisfying $\varphi(x_0) = y_0$. Moreover, the solution φ satisfies the equation $G(\varphi(x)) = F(x)$, $x \in I$, with

$$F(x) = \int_{x_0}^{x} f(t)\, dt, \quad \text{and} \quad G(x) = \int_{y_0}^{y} \frac{dt}{g(t)}.$$

Definition 8.6. Let I be an interval and $a, b : I \to \mathbb{R}$ continuous. Then we refer to the differential equation $y' = a(x)y + b(x)$ as *linear differential equation*. If $b(x) = 0$, then the linear differential equation is called *homogeneous*, else, it is called *inhomogeneous*.

Theorem 8.7. *Let I be an interval and $a, b : I \to \mathbb{R}$ continuous.*

i. The homogeneous linear differential equation $y' = a(x)y$ has a unique solution $\varphi_c : I \to \mathbb{R}$ satisfying $\varphi_c(x_0) = c$, $x_0 \in I$, namely

$$\varphi_c(x) = c \exp\big(\int_{x_0}^{x} a(t)\,dt\big).$$

ii. The inhomogeneous linear differential equation $y' = a(x)y + b(x)$ has a unique solution $\psi : I \to \mathbb{R}$ satisfying $\psi(x_0) = c$, $x_0 \in I$, namely

$$\varphi(x) = \varphi_1(x)\big(c + \int_{x_0}^{x} \frac{b(t)}{\varphi_1(t)}\,dt\big).$$

Definition 8.8. Let J be an interval and $f : J \to \mathbb{R}$ continuous. For $G = \{(x, y) \in \mathbb{R}\backslash\{0\} \times \mathbb{R} : \frac{y}{x} \in J\}$, we refer to

$$y' = f(\frac{y}{x}), \quad (x, y) \in G$$

as *homogeneous differential equation*.

Theorem 8.9. *Let J be an interval, $f : J \to \mathbb{R}$, $G = \{(x, y) \in \mathbb{R}\backslash\{0\} \times \mathbb{R} : \frac{y}{x} \in J\}$, $(x_0, y_0) \in G$ and $\varphi, \psi : I \to \mathbb{R}$ with $\psi(x) = \frac{\varphi(x)}{x}$ for $x \in I$. Then, φ solves $y' = f(\frac{y}{x})$, $\varphi(x_0) = y_0$ if and only if ψ solves $z' = \frac{1}{x}(f(z) - z)$, $\psi(x_0) = \frac{y_0}{x_0}$.*

Definition 8.10. Let $G \subset \mathbb{R} \times \mathbb{R}^n$ and $f : G \longrightarrow \mathbb{R}$, $(x, y) \mapsto f(x, y)$ be a continuous function. The formal expression

$$y^{(n)} = f(x, y, y', y'', \ldots, y^{(n-1)}) \tag{18}$$

is called *n-th order differential equations*.

A solution to (18) is a function $\varphi : I \longrightarrow \mathbb{R}$, I being an interval, such that the graph set $\Gamma_\varphi^{(n-1)} = \{(x, \varphi(x), \varphi'(x), \varphi''(x), \ldots, \varphi^{(n-1)}(x)) : \ x \in I\}$ satisfies $\Gamma_\varphi^{(n-1)} \subseteq G$ and

$$\varphi^{(n)}(x) = f(x, \varphi(x), \varphi'(x), \varphi''(x), \ldots, \varphi^{(n-1)}(x)) \quad \text{for all } x \in I.$$

Remark 8.11. Note that any n-th order differential equation can be solved by reducing it first to a system of first order differential equations. In fact, given the n-th order differential equation

$$y^{(n)} = f(x, y, y', y'', \ldots, y^{(n-1)}), \tag{19}$$

we define $F : G \to \mathbb{R}$, $G \subseteq \mathbb{R} \times \mathbb{R}^n$, and the respective system of first order linear equations $z' = F(x, z)$ by

$$z' = \begin{pmatrix} z_0' \\ z_1' \\ z_2' \\ \vdots \\ z_{n-2}' \\ z_{n-1}' \end{pmatrix} = \begin{pmatrix} z_1 \\ z_2 \\ z_3 \\ \vdots \\ z_{n-1} \\ f(x, z_0, z_1, \ldots, z_{n-1}) \end{pmatrix} = \begin{pmatrix} F_1\big(x, z_0, z_1, \ldots z_{n-1}\big) \\ F_2\big(x, z_0, z_1, \ldots z_{n-1}\big) \\ F_3\big(x, z_0, z_1, \ldots z_{n-1}\big) \\ \vdots \\ F_{n-1}\big(x, z_0, z_1, \ldots z_{n-1}\big) \\ F_n\big(x, z_0, z_1, \ldots z_{n-1}\big) \end{pmatrix} = F(x, z).$$

It is easily seen that any solution φ of $z' = F(x,z)$ on an interval I satisfies

$$\varphi_1'(x) = \varphi_2(x),\ \varphi_2'(x) = \varphi_3(x), \dots, \varphi_{n-1}'(x) = \varphi_n(x),\ \varphi_n'(x) = f(x, \varphi(x))$$

and, cutting out the middle men,

$$\varphi_1^{(n)} = f(x, \varphi(x)) = f(x, \varphi_1(x), \varphi_2(x), \dots, \varphi_n(x)) = f(x, \varphi_1(x), \varphi_1'(x), \dots, \varphi^{(n-1)}(x)),$$

that is, φ_1 is a solution of $y = f(x, y, y', \dots, y^{(n-1)})$ on the interval I.

Definition 8.12. Let $G \subset \mathbb{R} \times \mathbb{R}^n$. A function $f : G \longrightarrow \mathbb{R}^k$ satisfies a *Lipschitz condition* (in y) with *Lipschitz constant* $L \geq 0$ if

$$\|f(x,y) - f(x,\widetilde{y})\| \leq L\|y - \widetilde{y}\| \qquad \text{for all } x \in \mathbb{R}, y, \widetilde{y} \in \mathbb{R}^n.$$

The function f satisfies a *local Lipschitz condition* in G if for all $(x_0, y_0) \in G$ there exists an $\epsilon > 0$ such that $f|_{G \cap B_\epsilon(x_0,y_0)}$ satisfies a Lipschitz condition.

Note that in case of systems of first order linear differential equations we have $k = n$, while in case of a single n-th order differential equation we have $k = 1$.

Proposition 8.13. *Let $G \subset \mathbb{R} \times \mathbb{R}^n$ be open and $f : G \longrightarrow \mathbb{R}^n$ is continuously differentiable. Then f satisfies a local Lipschitz condition on G.*

Proof. Fix (x_0, y_0) in G. Since G open, exists $r > 0$ such that $K = B_r(x_0) \times B_r(y_0) \subseteq G$ where $B_r(x_0) = \{x : \ \|x - x_0\| \leq r\} \subseteq \mathbb{R}$ and $B_r(y_0) = \{y : \ \|y - y_0\| \leq r\} \subseteq \mathbb{R}^n$. As K is compact and Df is continuous, there exists L such that $\|Df_{(x,y)}\|_{\mathcal{L}} \leq L$ for all $(x,y) \in K$. The Mean Value Theorem, Theorem 6.22, then implies that for $(x,y), (x,\widetilde{y}) \in K$, we have

$$\|f(x,y) - f(x,\widetilde{y})\| \leq \sup\{\|(Df)_{(x,y)}\|_{\mathcal{L}} : (x,y) \in [(x,y),(x,\widetilde{y})]\} \leq L\|(x,y) - (x,\widetilde{y})\| = L\|y - \widetilde{y}\|.$$

□

Lemma 8.14. *Let $G \subset \mathbb{R} \times \mathbb{R}^n$ be open and $f : G \longrightarrow \mathbb{R}^n$ be continuous. A continuous function $\varphi : I \to \mathbb{R}^n$ solves $y' = f(x,y)$ with $\varphi(x_0) = y_0$ on G if and only if φ solves the integral equation*

$$\varphi(x) = y_0 + \int_{x_0}^{x} f(t, \varphi(t))\, dt, \qquad x \in I. \tag{20}$$

Proof. If $\varphi : I \to \mathbb{R}^n$ solves (20), then continuity of the map $t \mapsto f(t, \varphi(t))$ allows us to differentiate both sides of (20) and we obtain

$$\varphi'(x) = \frac{d}{dx}\varphi(x) = \frac{d}{dx}y_0 + \int_{x_0}^{x} f(t,\varphi(t))\, dt = f(x, \varphi(x)), \quad x \in I.$$

On the other hand, if $\varphi'(x) = f(x, \varphi(x))$ for $x \in I$ and $\varphi(x_0) = y_0$, then integration gives

$$\int_{x_0}^{x} f(t, \varphi(t))\, dt = \int_{x_0}^{x} \varphi'(t)\, dt = \varphi(x) - \varphi(x_0) = \varphi(x) - y_0,$$

so (20) is satisfied. □

Theorem 8.15. Uniqueness Theorem. *Let $G \subset \mathbb{R} \times \mathbb{R}^n$ and let $f : G \longrightarrow \mathbb{R}^n$ be continuous and satisfy a local Lipschitz condition. If $\varphi, \psi : I \longrightarrow \mathbb{R}^n$, I being an interval, are two solutions to the system of differential equations $y' = f(x, y)$ with $\varphi(x_0) = \psi(x_0)$ for some $x_0 \in I$, then $\varphi(x) = \psi(x)$ for all $x \in I$.*

Proof. Using the local Lipschitz condition of f, we can find $\delta > 0$ and $L > 0$ with

$$\|f(x, y) - f(x, \widetilde{y})\| \leq L\|y - \widetilde{y}\| \quad \text{for all } (x, y) \in B_\delta(x_0, y_0) \cap G.$$

Choosing $0 < \epsilon \leq \delta$ with $\|\varphi(x) - \varphi(x_0)\|, \|\psi(x) - \psi(x_0)\| < \epsilon$ for $\|x - x_0\| < \delta$, we obtain

$$\|f(x, \varphi(x)) - f(x, \psi(x))\| \leq L\|\varphi(x) - \psi(x)\| \quad \text{for all } x \in B_\delta(x_0) \cap I.$$

Now, set $\gamma = \min\{\frac{1}{2L}, \delta\}$ and

$$M = \sup_{x \in B_\gamma(x_0) \cap I} \|\varphi(x) - \psi(x)\| \leq \|\varphi(x_0)\| + 2\epsilon < \infty.$$

It follows

$$\begin{aligned}\|\varphi(x) - \psi(x)\| &= \left\| \int_{x_0}^{x} f(t, \varphi(t)) - f(t, \psi(t))\, dt \right\| \leq \int_{x_0}^{x} \|f(t, \varphi(t)) - f(t, \psi(t))\|\, dt \\ &\leq L \int_{x_0}^{x} \|\varphi(t) - \psi(t)\|\, dt \leq L\frac{1}{2L}M = \frac{M}{2}, \quad x \in B_\gamma(x_0) \cap I,\end{aligned}$$

which implies $M \leq \frac{M}{2}$, that is, $M = 0$.

Next, set $x_1 = \sup\{x : \ x > x_0 \text{ and } \psi(t) = \varphi(t) \text{ on } [x_0, x]\}$. If there exists $\widetilde{x} > x_1$ in I, then we can obtain the local argument above to contradict the construction of x_1. We use the same kind of argument, namely, $x_2 = \inf\{x : \ x > x_0 \text{ and } \psi(t) = \varphi(t) \text{ on } [x, x_0]\}$ to show that $\varphi(x) = \psi(x)$ on all of I. □

Example 8.16. Note that the differential equation $y' = y^{2/3}$, $y(0) = 0$, has multiple solutions, for example, $y = 0$ and $y(x) = \frac{1}{27}x^3$. Moreover, we can define

$$y = \begin{cases} 0 \text{ for } x \in [-a, a]; \\ \frac{1}{27}(x - a)^3, \text{ else }. \end{cases}$$

Theorem 8.17. Picard–Lindelöf Existence Theorem. *Let $G \subseteq \mathbb{R} \times \mathbb{R}^n$ be open and $f : G \longrightarrow \mathbb{R}^n$ be continuous and satisfy a local Lipschitz condition. Then exists for any point $(x_0, y_0) \in G$ an $\epsilon > 0$ and a solution $\varphi : [x_0 - \epsilon, x_0 + \epsilon] \longrightarrow \mathbb{R}^n$ to the differential equation $y' = f(x, y)$ which satisfies $\varphi(a) = \epsilon$.*

Proof. We will apply the Banach Fixed Point Theorem, Theorem 4.41 to the operator

$$T : \varphi \mapsto T\varphi, \quad T\varphi(x) = y_0 + \int_{x_0}^{x} f(t, \varphi(t))\, dt.$$

To do this, we need to choose a complete metric space A, show that T maps A to A, and, that T is a contraction on A.

For ϵ chosen later, recall that $C([x_0-\epsilon, x_0+\epsilon], \mathbb{R}^n)$ is a Banach space with norm $\|\varphi\|_\infty = \max\{\|\varphi(x)\|\}$. The set

$$A_{r,\epsilon} = \{\varphi \in C([x_0-\epsilon, x_0+\epsilon], \mathbb{R}^n), \|\varphi - y_0\|_\infty \leq r\}$$

is a closed subset of $C([x_0-\epsilon, x_0+\epsilon], \mathbb{R}^n)$, therefore, a complete metric space.

We choose $\delta, r > 0$ with the property that $\overline{B_\delta(x_0) \times B_r(y_0)} \subseteq G$ and f is Lipschitz on $\overline{B_\delta(x_0) \times B_r(y_0)}$ with Lipschitz constant L. Let M satisfy

$$\|f(x,y)\| \leq M, \quad (x,y) \in \text{ the compact set } \overline{B_\delta(x_0) \times B_r(y_0)}.$$

For $\epsilon = \min\{\delta, \frac{r}{M}, \frac{1}{2L}\}$ and $\varphi \in A_{r,\epsilon}$ we compute

$$\begin{aligned}\|T\varphi(x) - y_0\| &= \Big\| \int_{x_0}^{x} f(t, \varphi(t))\, dt \Big\| \leq \int_{x_0}^{x} \|f(t,\varphi(t))\|\, dt \\ &\leq |x - x_0| M \leq r, \quad x \in [x_0-\epsilon, x_0+\epsilon],\end{aligned}$$

and $T\varphi \in A_{r,\epsilon}$. Hence, T is well defined on $A_{r,\epsilon}$.

To see that T is a contraction on $A_{r,\epsilon}$, we compute similarly

$$\begin{aligned}\|T\varphi(x) - T\psi(x)\| &= \Big\| \int_{x_0}^{x} f(t,\varphi(t)) - f(t,\psi(t))\, dt \Big\| \leq \int_{x_0}^{x} \|f(t,\varphi(t)) - f(t,\psi(t))\|\, dt \\ &\leq \int_{x_0}^{x} L\|\varphi(t))\psi(t)\|\, dt \leq |x-x_0| L \|\varphi - \psi(t)\|_\infty \leq \frac{1}{2}\|\varphi - \psi(t)\|_\infty.\end{aligned}$$

□

We now combine Theorem 8.15 and Theorem 8.17 with Remark 8.11 to obtain

Theorem 8.18. *Let $G \subseteq \mathbb{R} \times \mathbb{R}^n$ be open and let $f : G \to \mathbb{R}$ be continuous and satisfy a local Lipschitz condition.*

i. If $\varphi, \psi : I \mapsto \mathbb{R}$ are two solutions to the differential equation $y^{(n)} = f(x, y, y', y'', \ldots, y^{(n-1)})$ with

$$\varphi(a) = \psi(a),\ \varphi'(a) = \psi'(a),\ \varphi''(a) = \psi''(a),\ \varphi^{(n-1)}(a) = \psi^{(n-1)}(a)$$

for some $a \in I$, then $\varphi(x) = \psi(x)$ for all $x \in I$.

ii. For any given $(a, c_0, \ldots, c_{(n-1)}) \in G$ exists $\epsilon > 0$ and $\varphi : [a-\epsilon, a+\epsilon] \to \mathbb{R}$, such that $\varphi^{(n)}(x) = f(x, \varphi(x), \varphi'(x), \varphi''(x), \ldots, \varphi^{(n-1)}(x))$ for all $x \in [a-\epsilon, a+\epsilon]$ and

$$\varphi(a) = c_0,\ \varphi'(a) = c_1,\ \varphi''(a) = c_2,\ \ldots\ , \varphi^{(n-1)}(a) = c_{(n-1)}.$$

Definition 8.19. Let I be an interval and

$$A = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nn} \end{pmatrix} : I \to \mathbb{R}^{n\times n}$$

be a continuous mapping into $\mathbb{R}^{n\times n}$ equipped with the operator norm with respect to the $\|\cdot\|_2$ norm on $\mathbb{R}^n$ (or, equivalently, any other norm on $\mathbb{R}^{n\times n}$). Further, let

$$b = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} : I \to \mathbb{R}^n$$

be a continuous mapping into $\mathbb{R}^n$ equipped with the $\|\cdot\|_2$ norm. (Note that A, respectively b, is continuous in the sense described above if and only if all a_{ij} are continuous real valued functions, respectively if all b_i are continuous real valued functions.)

i. The system of differential equations $y' = A(x)y$ is called *homogeneous system of linear equations.*

ii. The system of differential equations $y' = A(x)y + b(x)$, $b \neq 0$, is called *inhomogeneous system of linear equations.*

Theorem 8.20. *Let I be an interval and $A : I \to \mathbb{R}^{n\times n}$ and $b : I \to \mathbb{R}^n$ be continuous. Then exists to each $x_0 \in I$ and $c \in \mathbb{R}^n$ exactly one solution $\varphi : I \to \mathbb{R}^n$ to the differential equation $y' = A(x)y + b(x)$ which satisfies $\varphi(x_0) = y_0$.*

Proof. Note that uniqueness follows from the Uniqueness Theorem, Theorem 8.15, once we establish the simple fact that $f(x, y) = A(x)y$ satisfies a local Lipschitz condition. Existence in a neighborhood of x_0 then follows from the Existence Theorem, Theorem 8.17. The real "news" of this theorem is that we obtain existence on all of I.

Let $J \subseteq I$ be compact. We shall show that $f(x, y) = A(x)y+b(x)$ satisfies a global Lipschitz condition in y on J. To this end, note that continuity of $A(x)$ implies that there exists L_J with $\|A(x)\|_{op(2)} \leq L_J$ for all $x \in J$.[7] We now compute

$$\|f(x, y) - f(x, \widetilde{y})\| = \|A(x)y - A(x)\widetilde{y}\| \leq \|A(x)\|_{op(2)}\|y - \widetilde{y}\| \leq L_J\|y - \widetilde{y}\|$$

Motivated by the use of Theorem 4.41 in the proof of Theorem 8.17, we define a sequence of functions on I as follows. The function $\varphi_0(x) = y_0$ is constant, then

$$\varphi_{k+1}(x) = y_0 + \int_{x_0}^{x} f(t, \varphi_k(t))\, dt = y_0 + \int_{x_0}^{x} A(t)\varphi_k(t) + b(t)\, dt.$$

[7]Recall that the operator norm $\|\cdot\|_{op(2)}$ is defined by $\|A\|_{op(2)} \sup_{y\in\mathbb{R}^n\setminus\{0\}} \frac{\|Ay\|_2}{\|y\|_2}$. Consequently $\|Ay\|_2 \leq \|A\|_{op(2)}\|y\|_2$, a fact that is used in this proof. Note that in this section $\|\cdot\| = \|\cdot\|_2$.

We shall show that for J compact with $x_0 \in J$, the sequence $\varphi_k|_J$ is an uniformly convergent sequence of functions on J. Indeed, note that A, b, φ_0 is continuous on J, we obtain that φ_1 is differentiable on J, and hence continuous. Hence, $K_J = \sup_{x\in J} \|\varphi_1(x) - \varphi_0(x)\| < \infty$.

Claim. $\|\varphi_{k+1}(x) - \varphi_k(x)\| \leq \dfrac{K_J L_J^k |x - x_0|^k}{k!}, \quad x \in J.$

The base case holds by definition of K_J. For the induction step, we compute

$$\begin{aligned}
\|\varphi_{k+1}(x) - \varphi_k(x)\| &= \Big\| \int_{x_0}^x A(t)\big(\varphi_k(t) - \varphi_{k-1}(t)\big)\, dt \Big\| \leq \int_{x_0}^x \|A(t)\|_{op(2)} \big\|\varphi_k(t) - \varphi_{k-1}(t)\big\|\, dt \\
&\leq \int_{x_0}^x L_J \frac{K_J L_J^{k-1} |t - x_0|^{k-1}}{(k-1)!} dt = \Big| \int_{x_0}^x L_J \frac{K_J L_J^{k-1} (t - x_0)^{k-1}}{(k-1)!} dt \Big| \\
&= \frac{K_J L_J^k |x - x_0|^k}{k!}.
\end{aligned}$$

The claim implies that the the series φ_k converges uniformly on J. Indeed, since $C(J, \mathbb{R}^n)$ with norm $\|\cdot\|_\infty$ is a complete metric space, it suffices to show that (φ_k) is a Cauchy sequence. To this end, set $M = \sup_{x\in J} |x - x_0|$. Now, for $\epsilon > 0$ fixed, choose N such that

$$\sum_{\ell=N}^{\infty} \frac{K_J (L_J M)^k}{k!} < \epsilon.$$

This is possible since $\sum_{\ell=0}^{\infty} \frac{K_J (L_J M)^k}{k!}$ is a convergent series, in fact, $\sum_{\ell=N}^{\infty} \frac{K_J (L_J M)^k}{k!} = K_J e^{L_J M}$.

For $m > n > N$, we compute

$$\begin{aligned}
\|\varphi_m - \varphi_n\|_\infty &= \sup_{x\in J} \|\varphi_m(x) - \varphi_n(x)\| \leq \sup_{x\in J} \Big\| \sum_{\ell=n}^{m-1} \varphi_{\ell+1}(x) - \varphi_\ell(x) \Big\| \\
&\leq \sum_{\ell=n}^{m-1} \sup_{x\in J} \|\varphi_{\ell+1}(x) - \varphi_\ell(x)\| \leq \sum_{\ell=n}^{m-1} \sup_{x\in J} \frac{K_J L_J^k |x - x_0|^k}{k!} \\
&\leq \sum_{\ell=n}^{m-1} \frac{K_J (L_J M)^k}{k!} \leq \sum_{\ell=n}^{\infty} \frac{K_J (L_J M)^k}{k!} < \epsilon.
\end{aligned}$$

We conclude that (φ_k) converges uniformly on J to some function ψ

Now, uniform convergence on the compact set J allows us to take limits on both sides of

$$\varphi_{k+1}(x) = y_0 + \int_{x_0}^x A(t)\varphi_k(t) + b(t)\, dt$$

to obtain

$$\psi(x) = y_0 + \int_{x_0}^x A(t)\psi(t) + b(t)\, dt,$$

differentiating both sides shows that ψ solves our ODE on J. Since J was chosen arbitrarily, we obtain the envisioned result. □

Theorem 8.21. *Let I be an interval and $A : I \to \mathbb{R}^{n\times n}$ be continuous. The set $L_H = \{\varphi : I \mapsto \mathbb{R}^n : \varphi'(x) = A(x)\varphi(x), x \in I\}$ is an n-dimensional vector space over $\mathbb{R}$.*

For $\varphi_1, \varphi_2, \ldots, \varphi_k \in L_H$, the following are equivalent

i. $\varphi_1, \varphi_2, \ldots, \varphi_k$ *are linearly independent functions;*

ii. $\varphi_1(x_0), \varphi_2(x_0), \ldots, \varphi_k(x_0)$ *are linearly independent vectors for some* $x_0 \in I$.

iii. $\varphi_1(x), \varphi_2(x), \ldots, \varphi_k(x)$ *are linearly independent vectors for each* $x \in I$.

Proof. First, it is easily checked that L_H is a vector space. For example, if $\varphi'(x) = A(x)\varphi(x)$ and $\psi'(x) = A(x)\psi(x)$ on I, then for $\alpha, \beta \in \mathbb{R}$, we have

$$(\alpha\varphi + \beta\psi)'(x) = \alpha\varphi'(x) + \beta\psi'(x) = \alpha A(x)\varphi(x) + \beta A(x)\psi(x) = A(x)(\alpha\varphi(x) + \beta\psi(x)).$$

For the equivalence, note that trivially *iii.* $\Longrightarrow$ *ii.* $\Longrightarrow$ *i.* For *i.* $\Longrightarrow$ *iii*, assume that $\varphi_1, \varphi_2, \ldots, \varphi_k$ are linearly independent functions but $\varphi_1(x_0), \varphi_2(x_0), \ldots, \varphi_k(x_0)$ are NOT linearly independent for some x_0. Then exist coefficients $\alpha_1, \ldots, \alpha_k$ with $\sum_{i=1}^k \alpha_i\varphi_i(x_0) = 0$. Now, $\sum_{i=1}^k \alpha_i\varphi_i$ solves the ODE $y' = A(x)y$ with $\varphi(x_0) = 0$, and so does the constant function $\psi(x) = 0$. By uniqueness, we have the function equality $\sum_{i=1}^k \alpha_i\varphi_i = \psi = 0$ and the φ_i are linearly dependent.

It remains to show that L_H is an n-dimensional vector space. To this end, fix $x_0 \in I$ and use the existence result to obtain solutions φ_j of $y' = A(x)y$ with $\varphi(x_0) = e_j$, where e_j denotes the j-th element of the Euclidean basis of $\mathbb{R}^n$. Now, *ii* above is satisfied, and, hence, *i*, so we found n linearly independent solutions φ_j in L_H, implying that the dimension of L_H is greater than or equal to n.

If the dimension of L_H would be strictly greater n, then we could consider $n+1$ linearly independent solutions $\psi_1, \ldots, \psi_{n+1}$. Evaluating at some point $x_0 \in I$, we obtain using $i \Longrightarrow ii$ exactly $n+1$ linearly independent vectors $\psi_1(x_0), \ldots, \psi_{n+1}(x_0)$ in $\mathbb{R}^n$, a contradiction. □

Definition 8.22. A basis of L_H given in Theorem 8.21 is called *fundamental system of solutions* to the system of differential equation $y' = A(x)y$.

Remark 8.23. Let $\varphi_1 = (\varphi_{11}, \varphi_{21}, \ldots, \varphi_{n1})^T$, $\varphi_2 = (\varphi_{12}, \varphi_{22}, \ldots, \varphi_{n2})^T$, …, $\varphi_n = (\varphi_{1n}, \varphi_{2n}, \ldots, \varphi_{nn})^T$ be a fundamental system of solutions of $y' = A(x)y$. We set

$$\Phi = \begin{pmatrix} \varphi_{11} & \varphi_{12} & \cdots & \varphi_{1n} \\ \varphi_{21} & \varphi_{22} & \cdots & \varphi_{2n} \\ \vdots & \vdots & & \vdots \\ \varphi_{n1} & \varphi_{n2} & \cdots & \varphi_{nn} \end{pmatrix}$$

and observe that any solution φ to $y' = A(x)y$ can be written as $\varphi = \Phi c$ for appropriate $d \in \mathbb{R}^n$.

Considering the initial value problem $y' = A(x)y$, $\varphi(x_0) = c \in \mathbb{R}^n$, we note that for the solution φ, $c = \varphi(x_0) = \Phi(x_0)d$. According to Theorem 8.21, we have that for any $x_0 \in I$, the matrix

$$\Phi(x_0) = \begin{pmatrix} \varphi_{11}(x_0) & \varphi_{12}(x_0) & \dots & \varphi_{1n}(x_0) \\ \varphi_{21}(x_0) & \varphi_{22}(x_0) & \dots & \varphi_{2n}(x_0) \\ \vdots & \vdots & & \vdots \\ \varphi_{n1}(x_0) & \varphi_{n2}(x_0) & \dots & \varphi_{nn}(x_0) \end{pmatrix}$$

is invertible, hence, we have $d = \Phi(x_0)^{-1}c$.

Theorem 8.24. *Let I be an interval and $A : I \to \mathbb{R}^{n\times n}$, $b : I \to \mathbb{R}^n$ be continuous. Let $L_H = \{\varphi : I \mapsto \mathbb{R}^n : \varphi'(x) = A(x)\varphi(x), x \in I\}$ and $L_I = \{\psi : I \mapsto \mathbb{R}^n : \ \psi'(x) = A(x)\psi(x) + b(x), x \in I\}$. For any $\psi_0 \in L_I$, we have $L_I = \psi_0 + L_H$.*

Proof. If $\psi \in L_I$, then $\varphi = \psi - \psi_0 \in L_H$ since

$$\varphi'(x) = \psi'(x) - \psi_0'(x) = A(x)\psi(x) + b(x) - A(x)\psi_0(x) - b(x) = A(x)(\psi(x) - \psi_0(x)) = A(x)\varphi(x),$$

so $\psi \in \psi_0 + L_H$. Similarlly, $\varphi \in L_H$ implies $\varphi + \psi_0 \in L_I$. □

Theorem 8.25. *Let I be an interval and $A : I \to \mathbb{R}^{n\times n}$, $b : I \to \mathbb{R}^n$ be continuous. Let $\varphi_1 = (\varphi_{11}, \varphi_{21}, \dots, \varphi_{n1})^T$, $\varphi_2 = (\varphi_{12}, \varphi_{22}, \dots, \varphi_{n2})^T$, ..., $\varphi_n = (\varphi_{1n}, \varphi_{2n}, \dots, \varphi_{nn})^T$ be a fundamental system of solutions of $y' = A(x)y$. For*

$$\Phi = \begin{pmatrix} \varphi_{11} & \varphi_{12} & \dots & \varphi_{1n} \\ \varphi_{21} & \varphi_{22} & \dots & \varphi_{2n} \\ \vdots & \vdots & & \vdots \\ \varphi_{n1} & \varphi_{n2} & \dots & \varphi_{nn} \end{pmatrix},$$

a solution $\psi : I \to \mathbb{R}^n$ to $y' = A(x)y + b(x)$ is given by $\psi(x) = \Phi(x)u(x)$ with

$$u(x) = \int_{x_0}^{x} \Phi(t)^{-1}b(t)\,dt + C.$$

Proof. We have $\psi(x) = \Phi(x)u(x) = \Phi(x)\big(\int_{x_0}^{x} \Phi(t)^{-1}b(t)\,dt + C\big)$. Differentiating gives

$$\begin{aligned} \psi'(x) &= \frac{d}{dx}\Phi(x)u(x) = \Phi'(x)u(x) + \Phi(x)u'(x) \\ &= A(x)\Phi(x)u(x) + \Phi(x)\Phi(x)^{-1}b(x) = A(x)\Psi(x) + b(x). \end{aligned}$$ □

Example 8.26. Solve $y_1' = -y_2$, $y_2' = y_1 + x$.

Index

C^1 – mean value theorem, 88
C^r–diffeomorphism, 90

C^r, 90
Riemann integral, 68

abelian group, 7
absolute value of a complex number, 17
absolutely convergent series, 30
addition in a field, 8
Algebraic Limit Theorem, 21
antisymmetric relation, 5
arc, 82
archimedean property, 9
argument of a complex number, 17
associativity, 7

Banach Fixed Point Theorem, 66
Banach space, 74
basis of a vector space, 75
Bessel Inequality, 77
Beta function, 113
bijective, 4
Bolzano–Weierstrass Theorem, 24
boundary of a set, 47
bounded above, 11
bounded below, 11
bounded sequences in metric spaces, 21
bounded sets in metric spaces, 21

Cantor set, 55
Cartesian product, 4
Cauchy Condensation Theorem, 32
Cauchy criterion, 29
Cauchy sequences, 27
Cauchy–Schwarz Inequality, 113
Cauchy-Schwarz Inequality, 38, 75
Chain Rule, 59
Change of variable formula , 110
closed ball, 38
closed sets, 42, 43
closure of a set, 47
cluster points of a set, 47
common refinement, 69
commutative group, 7
compact set, 47
complete metric spaces, 27
complex numbers, 17
conditionally convergent series, 30
conjugate of a complex number, 17
connected set, 53
continuous functions on $\mathbb{R}$, 37
continuous functions on metric spaces, 38
continuous functions on topological spaces, 44
contraction, 66
convergence in the extended real number system, 24
convergence in topological spaces, 45
converging sequences in metric spaces, 20
converging sequences in the reals, 19
convex function, 113
cosine function, 66
countable sets, 14
covering compact set, 47
covering of a set, 47
curve, 82

Darboux integral, 69
Darboux's Theorem., 61
decreasing function, 61
decreasing sequences, 21
Dedekind–cut, 10
dense subset, 47
derivative, multivariable, 83
differentiable functions, 58
differential equations, 114
differentiation of integrands, 89
direction field, 114
directional derivative, 83
discrete metric, 19
discrete topology, 43
distance function, 19
distributive law, 9
divergence of a vector field, 84
domain of a function, 4
domain of a relation, 4
Dominated Convergence Theorem, 30

e, 28

embedding, 4
equivalence class, 5
equivalence relation, 5
equivalent curves, 82
equivalent elements, 5
euclidean metric, 19
exhaustive, 112
exponential function, 65
extended real number system, 24
extremum, 93

field, 8
finite cover, 47
first order differential equation, 114
fixed point, 66
Fourier coefficient, 78
function, 4
functional limits, 58
fundamental system of solutions, 121
Fundamental Theorem of Calculus I, 72
Fundamental Theorem of Calculus II, 72

Gamma function, 113
Generalized Mean Value Theorem, 61
global maximum, 93
global minimum, 93
gradient, 84
graph of a function, 4
greatest lower bound, 11
grid of rectangles, 104
group, 7

Hölder's inequality, 113
Heine–Borel Theorem, 50
Hessian matrix, 91
higher derivatives, 63, 90
Hilbert space, 75
homeomorph topological spaces, 45
homeomorphism, 44
homogeneous differential equation, 115
homogeneous linear differential equation, 115
homogeneous system of linear equations, 119

identity element, 7
imaginary part of a complex number, 17
implicit function, 97
Implicit Function Theorem, 97
improper Riemann integral, 112
improper Riemann integral, 73
increasing function, 61
increasing sequences, 21
indefinite, 93
indefinite integral, 72
indiscrete topology, 43
infimum, 11
infinite limits, 24, 62
infinite series, 28
Inheritance Principle, 43
inhomogeneous linear differential equation, 115
inhomogeneous system of linear equations, 119
injective, 4
inner Jordan content, 109
inner product, 74
inner product space, 74
integers, 7
Integral Criterion for Sums, 73
integrals of complex valued functions, 77
Integration by Parts, 72
Integration by Substitution, 72
Interior Extremum Theorem, 60
interior of a set, 47
intervals, 55
inverse element, 7
Inverse Function Theorem, 101

Jacobian determinant, 84
Jacobian matrix, 84
Jordan content, 108, 112
Jordan domain, 108
Jordan zeroset, 108

L'Hospital's Rule., 62
Lagrange Multiplier, 102
least upper bound, 11
least upper bound property, 11
Lebesgue number, 49
Leibniz Criterion for Alternating Series, 32
Leibniz Rule, 87
length of a curve, 82
limit inferior (liminf), 25
limit point of a real valued sequence, 25
limit superior (limsup), 25

limits at infinity, 62
linear differential equation, 115
linear Order, 6
Lipschitz condition, 116
Lipschitz constant, 116
local minimum, 93
local Lipschitz condition, 116
local maximum, 93
logarithm, 62
logarithm naturales, 65
logarithm to base a, 65
lower bound, 11
lower integral, 69, 104
lower sum, 69

mapping, 4
Mean Value Theorem, 60
Mean Value Theorem for Integrals, 71
Mean Value Theorem, multivariate, 88
mesh of a partition, 68
metric, 19
metric space, 19
metric space of bounded functions, 56
metric space of continuous functions, 57
monotone function, 61
monotone sequences, 21
multiplication in a field, 8

n–th derivative, 63
n–th root, 14
n-th order differential equation, 115
n-tupel, 4
natural logarithm, 65
negative definite, 93
negative semidefinite, 93
neighborhood, 38
neighborhood of $\pm\infty$, 62
nested interval property, 14
Newton's Method, 67
norm, 74
normed vector space, 74

one–to–one, 4
onto, 4
open ball, 38
open cover, 47
open sets, 42, 43
Order Limit Theorem, 22, 26
order relation, 6
ordered field, 9
ordinary differential equations, 114
orthogonal family, 75
orthonormal system (ONS), 75
oscillation, 71
othonormal basis (ONB), 77
outer Jordan content, 109

partial derivative, 84
partial sums, 28
partition of a set, 6
partition of an interval, 68
Picard–Lindelöf Existence Theorem, 117
pointwise convergence of sequences of functions, 56
positive definite, 93
positive semidefinite, 93
power series, 34
product of series, 35
Product Rule, 59

Quotient Rule, 59

radius of convergence, 34, 64
range of a function, 4
range of a relation, 4
Ratio Test, 33
rational numbers, 8
real numbers, 10
real part of a complex number, 17
rearrangement of a series, 30
rectifiable curve, 82
refinement, 69
Refinement Principle, 69
reflexive relation, 5
regular curve, 82
relation, 4
representative of an equivalence class, 5
Riemann Lebesgue theorem, 105
Riemann Integrability Criterion, 70
Riemann integral, 88, 104
Riemann sum, 68, 104
Riemann–Lebesgue Theorem, 71

Rolle's Theorem, 60
Root Test, 33
rotation of a vector field, 84

saddle point, 95
sampling set, 68
scalar multiplication, 74
scalar product, 74
second partials, 91
separable differential equation, 114
separation of a metric space, 53
sequences, 14
sequentially compact set, 48
series, 28
sine function, 66
slope field, 114
smooth function, 90
smooth functions, 63
Squeezing Theorem, 22
Stirling's formula, 113
strictly decreasing function, 61
strictly decreasing sequences, 21
strictly increasing function, 61
strictly increasing sequences, 21
strictly monotone function, 61
subcoverings, 47
subsequence of a sequence, 23
Sum Rule, 59
supremum, 11
surjective, 4
symmetric relation, 5

Taylor polynomial, 63
Taylor series, 64
Taylor's Theorem, 63, 94
topological space, 43
topology on X, 43
total derivative, 83
totally bounded, 51
totally disconnected sets, 55
transitive relation, 5
trigonometric polynomial, 78
trigonometric series, 78
trivial topology, 43

uniform continuity, 53
uniform convergence of sequences of functions, 56
uniqueness of the real number system, 12
upper bound, 11
upper integral, 69, 104
upper sum, 69

vector field, 84
vector space, 74
volume, 108

Wallis' Product, 73
width of a partition, 68

zero set, 71
zeroset in $\mathbb{R}^2$, 105