Jacobs University Bremen
School of Engineering and Science
Götz Pfander

Fall Term 2007,
Spring Term 2008

Analysis I,II

Version (May 15, 2008)

Contents

Preface

This script contains all the theorems and definitions, but only a few examples covered in Analysis I, II in the academic year 2007/2008.

Most proofs have been omitted from this script. With the exception of two or three theorems, all statements have been proven in either the script, in class, or in the homeworks.

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

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.*

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 $B_1 = f(f^{-1}(B_1))$.

DEFINITION 1.8. 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 1.9. 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 $[x]$.

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

DEFINITION 1.10. 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 1.11. 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

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

Example 1.12. 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. The set of equivalence classes is the group $\mathbb{Z}_n$ of n elements with addition given by

$$[n] + [m] = [n + m].$$

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.)

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

Definition 1.13. 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}.

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

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

Examples 1.14.

i. 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

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

is a ring[1], called the *ring of integers* . 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.[2] 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, $7 - 3 = 10 - 6$

[1]Since we shall not use any rings in this course, we omit a definition of *rings*. Please consult a textbook.

[2]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".

ii. 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[3], 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 $r = \dfrac{z}{m} = [(z,m)] \in \mathbb{Q}$.

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

$$\mathbb{N} \underset{\neq}{\subseteq} \mathbb{Z} \underset{\neq}{\subseteq} \mathbb{Q}.$$

Definition 1.15. A *field* is a set F on which two binary laws of composition, *addition* '+' and *multiplication* '$\cdot$' are defined with

(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 $x, y, z \in F$.

Definition 1.16. 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 $(\widetilde{x}, x) \in O$.

All orders discussed in Examples 1.17 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.

Note that the order on $\mathbb{N}$ which we mention in Examples 1.17.i can be easily defined using elementary set theory.

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

[3]Fields will be defined shortly

Examples 1.17.

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

ii. The relation $O_{\mathbb{Z}} = \Big\{\Big([(n,m)],[(\widetilde{n},\widetilde{m})]\Big) \in \mathbb{Z} \times \mathbb{Z} : \quad n+\widetilde{m} \le \widetilde{n}+m\Big\}$ extends the order on $\mathbb{N}$ to the integers $\mathbb{Z}$.

iii. The relation $O_{\mathbb{Q}} = \Big\{\Big([(z,m)],[(\widetilde{z},\widetilde{m})]\Big) \in \mathbb{Q} \times \mathbb{Q} : \quad z \cdot \widetilde{m} \le \widetilde{z} \cdot m\Big\}$ extends the order on $\mathbb{Z}$ to the rational numbers $\mathbb{Q}$.

In the following we shall simply write $r \le \widetilde{r}$ if $(r,\widetilde{r}) \in O_{\mathbb{Q}}$.

Definition 1.18. A field F is called *ordered* if

(O1) There exists an order '$\le$' 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) $x \le y$ implies $x + z \le y + z$ for all $x, y, z \in F$ and if $x, y > 0$ then $x \cdot y > 0$.

Definition 1.19. 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 1.20. The set of rational numbers $\mathbb{Q}$ together with the two binary operations addition and multiplication defined in Examples 1.14.ii and the order given in Examples 1.17.iii is an archimedean ordered field.

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

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 1.22. 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.

EXAMPLES 1.23. $\{q \in \mathbb{Q} : q < 2\}|\{q \in \mathbb{Q} : q \geq 2\}$ and
$\{q \in \mathbb{Q} : q < 0 \text{ or } q^2 < 2\}|\{q \in \mathbb{Q} : q \geq 0 \text{ and } q^2 > 2\}$ are cuts, but
$\{q \in \mathbb{Q} : q \leq 2\}|\{q \in \mathbb{Q} : q > 2\}$, $\{q \in \mathbb{Q} : q^2 \leq 2\}|\{q \in \mathbb{Q} : q^2 > 2\}$ and
$\{q \in \mathbb{Q} : q < 2\}|\{q \in \mathbb{Q} : q \geq 3\}$ are not.

DEFINITION 1.24. Dedekind-cuts in $\mathbb{Q}$ are called *real numbers*, the set of all real numbers is denoted by $\mathbb{R}$.

REMARK 1.25. 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 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 1.26. Let X be a linearly ordered set, $S \subseteq X$. $M \in X$ 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 1.27. (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).

Definition 1.28. 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 1.29. 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 1.30. 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

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

By definition we have $n_F + m_F = (n+m)_F$. The injectivity of this mapping follows from an inductive argument using $n_F + 1_F >_F n^* + 0_F$. Let us also note that implies that the order on $\mathbb{N}$ is preserved under 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

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

(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 $x < \frac{u_F + 1_F}{m_F} < y$, 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. $\square$

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 $x, y, a, b, \alpha, \beta, \ldots$.

THEOREM 1.31. 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 1.32. NESTED INTERVAL PROPERTY.
For $n \in \mathbb{N}$, 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 $\bigcap\limits_{n\in\mathbb{N}} I_n \neq \emptyset$.

DEFINITION 1.33. 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 1.34. A set X is *countable* if there is a surjective function (sequence) $a\colon \mathbb{N} \longrightarrow X, n \mapsto a(n)$.

THEOREM 1.35. If the sets $A_n \subset X$, $n \in \mathbb{N}$, are countable, then $\bigcup_{n\in\mathbb{N}} A_n$ is countable.

COROLLARY 1.36. $\mathbb{Q}$ is countable.

THEOREM 1.37. The set containing all sequences with values in $\{0, 1, 2, \ldots, n\}$, $n \geq 1$, is not countable.

THEOREM 1.38. $\mathbb{R}$ is not countable.

1.4. Complex numbers

We shall now define the complex number system.

DEFINITION 1.39. The cartesian product $\mathbb{R} \times \mathbb{R}$ together with the binary operations

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

form a field with additive neutral element $(0,0)$ and multiplicative neutral element $(1,0)$ which is called the *field of complex numbers*.

THEOREM 1.40. 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 $a, b \in \mathbb{R}$

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

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

PROPOSITION 1.41. For $i := (0,1)$, we have $i^2 = (-1,0)$, 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 1.42. 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 $|z| = \sqrt{a^2+b^2}$.

PROPOSITION 1.43. For all $z = a + bi,\ w = c + di\ \in \mathbb{C}$ with $a, b, c, d \in \mathbb{R}$ we have

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

REMARK 1.44. 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 spaces of real and complex numbers. In these special cases, the distance between two numbers x and y is $|x-y|$.

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

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

If $(x_n)_{n\in\mathbb{N}}$ converges to x_0 in $\mathbb{R}$ 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 $x_0 \in \mathbb{R}$ 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

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

If (x_n) converges to x_0 in (X, d) 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 $x_0 \in X$ 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

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

and

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

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 $\mathbb{C}$.

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 $x_0 = \widetilde{x_0}$.

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.

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. Monotonic sequences converge in $\mathbb{R}$ if and only if they are bounded.

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

ii. $a_n b_n \xrightarrow{n\to\infty} a_0 b_0$, and

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

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 $\mathbb{C}$ with $a_n \leq b_n$ for all $n \in \mathbb{N}$, then $a_0 \leq b_0$.

Examples 2.14.

i. For $p > 0$ we have $\lim\limits_{n\to\infty} \dfrac{1}{n^p} = 0$.

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

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

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

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

Definition 2.15. 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 2.16. 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 2.17. Every subsequence $(s_{n_k})_k$ of a convergent sequence $(s_n)_n$ in (X, d) converges to the same limit as $(s_n)_n$.

Example 2.18. 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 $\mathbb{R}$.

Theorem 2.19. Bolzano–Weierstrass Theorem. Every bounded sequence $(s_n)_n$ in $\mathbb{R}$ has a converging subsequence.

2.2. The extended real number system, $\limsup$ and $\liminf$

Definition 2.20. 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

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

then we write $\lim_{n\to\infty} x_n = \infty$, 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

$$x_n \leq -M \qquad \text{for all naturals } n \geq N,$$

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

Proposition 2.21. 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.

Proposition 2.22. Let (x_n) be a sequence of real numbers. Then

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

is not empty.

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).

Examples 2.24.

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

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

LEMMA 2.25. 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 $x_n \leq s$ for all $n \geq N$. If $s < \liminf_{n\to\infty} x_n$, then exists $N \in \mathbb{N}$ such that $x_n \geq s$ for all $n \geq N$.

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 $x_n \leq s$ for all $n \geq N$. The second assertion follows verbatim.

If $s = \infty$, then $s_n \leq s = \infty$ for all $n \geq 1$.

We have $s > \limsup_{n\to\infty} x_n \geq -\infty$, and, hence, we can turn our attention to the remaining case $s \in \mathbb{R}$. Suppose that for any $N \in \mathbb{N}$ there exists an index $n_N \in \mathbb{N}$ such that $x_{n_N} > s$. In this case, we can pick n_1 such that $x_{n_1} > s$, then $n_2 > n_1$ with $x_{n_2} > s$, and, inductively $n_{k+1} > n_k$, $k \in \mathbb{N}$.

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 $y \geq s$ since $x_{n_k} \geq s$ for all $k \in \mathbb{N}$. The fact that $y \in E_{(x_n)_n}$ implies $\limsup_{n\to\infty} x_n \geq y \geq s > \limsup_{n\to\infty} x_n$, which is nonsense. Contradiction! □

THEOREM 2.26. Let (x_n) be a sequence in $\mathbb{R}$. Then for $x_0 \in \mathbb{R}^*$ we have $\lim_{n\to\infty} x_n = x_0$ if and only if $\liminf_{n\to\infty} x_n = \limsup_{n\to\infty} x_n = x_0$.

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

Let us now assume $\liminf_{n\to\infty} x_n = \limsup_{n\to\infty} x_n = x_0$ with $x_0 \in \mathbb{R}$. Fix $\epsilon > 0$ and use Lemma 2.25 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 $\lim_{n\to\infty} x_n = x_0$.

Let us assume $\liminf_{n\to\infty} x_n = \limsup_{n\to\infty} x_n = +\infty$. Lemma 2.25 implies that for all $M < \infty$ exists $N \in \mathbb{N}$ with $x_n > M$ for $n \geq N$. This gives $\lim_{n\to\infty} x_n = \infty$.

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

2.3. Cauchy sequences and complete metric spaces

DEFINITION 2.27. 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 2.28. Any converging sequence in a metric space is a Cauchy sequence.

PROPOSITION 2.29. Any Cauchy sequence in a metric space is bounded.

DEFINITION 2.30. A metric space (X, d) is called complete if all Cauchy sequences in X converge in X.

REMARK 2.31. 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 $\mathbb{R} \setminus \{0\}$, since if it would converge to say $\alpha \in \mathbb{R} \setminus \{0\}$, then it is easy to see that for any $\epsilon > 0$ there would exist some N_ϵ such that

$$|\alpha - 0| \leq |\alpha - x_n| + |0 - x_n| < \epsilon + \epsilon = 2\epsilon.$$

Hence $|\alpha - 0| \leq 2\epsilon$ for all $\epsilon > 0$ and therefore $|\alpha - 0| = 0$ and $\alpha = 0$, a contradiction to $\alpha \in \mathbb{R} \setminus \{0\}$.

PROPOSITION 2.32. 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 $x_n \xrightarrow{n \to \infty} x_0$.

THEOREM 2.33. $\mathbb{R}$ and $\mathbb{C}$ are complete.

2.4. Real and complex series

DEFINITION 2.34. 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 convergence).

EXAMPLE 2.35. 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 2.36. Set $e = \sum_{n=0}^{\infty} \frac{1}{n!} \in \mathbb{R}$.

REMARK 2.37. 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 and therefore the series $\sum_{n=0}^{\infty} \frac{1}{n!}$ converges.

THEOREM 2.38. $\lim_{n\to\infty}\left(1+\frac{1}{n}\right)^n = e$

THEOREM 2.39. e is irrational.

THEOREM 2.40. 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

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

Proposition 2.41. If $\sum_{n=1}^{\infty} a_n$ converges in $\mathbb{C}$ then $a_n \xrightarrow{n\to\infty} 0$.

Theorem 2.42. 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.

Corollary 2.43. 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 2.44. 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 2.45. 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 2.46.

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 $\sum_{n=1}^{\infty} c_{\pi_x(n)} = x$.

Example 2.47. Take $S = \sum_{n=1}^{\infty} (-1)^n \frac{1}{n} \neq 0$. Consider:

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

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

Theorem 2.49. 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.

Proposition 2.50. For $p \in \mathbb{R}$ we have $\sum_{n=1}^{\infty} \frac{1}{n^p}$ converges if and only if $p > 1$.

Theorem 2.51. 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. Here, we shall only show iii.

We have $\limsup_{n\to\infty} \sqrt[n]{\frac{1}{n}} = \frac{1}{\lim_{n\to\infty} \sqrt[n]{n}} = 1$ 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 2.52. 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.

Examples 2.53.

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

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

Definition 2.54. 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 2.55. 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 $|z| = R_{(c_n)}$.

Remark 2.56. 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 a diagonal summation of the product:

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

Theorem 2.57. 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 $\sum_{n=0}^{\infty} c_n = A \cdot B$.

Corollary 2.58. For $z, w \in \mathbb{C}$ we have $\exp(z + w) = \exp(z) \exp(w)$.

Corollary 2.59. For $x \in \mathbb{R}$ we have $\exp(x) = e^x$.

Motivated by this corollary, we shall write e^z for $\exp(z)$ for any $z \in \mathbb{C}$.

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 $x_0 = 2$.

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 3.61.)

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) 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 $d_X(x_0, x) < \delta$.

DEFINITION 3.5. 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

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

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

THEOREM 3.6. Let (X, d_X) and (Y, d_Y) be metric spaces, and let $f : X \to Y$, and $x_0 \in X$. f is continuous at x_0 if and only if for all $\varepsilon > 0$ exists $\delta > 0$ s.t. $f(B_\delta(x_0)) \subseteq B_\varepsilon(f(x_0))$.

THEOREM 3.7. CAUCHY–SCHWARZ INEQUALITY.
Let $a_1, \ldots, a_n, b_1, \ldots, b_n \in \mathbb{C}$. Then

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

EXAMPLES 3.8. Examples of metrics d_0, d_1, d_2, d_∞ on $\mathbb{R}^n$. Describe respective balls.

THEOREM 3.9. 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 $i, j \in \{1, 2, \infty\}$.

REMARK 3.10. 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 3.11. 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*.

THEOREM 3.12. Let (X, d) be a metric space, then open balls are (metric-) open.

PROPOSITION 3.13. 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 $(\mathbb{R}^n, d_2)$.

THEOREM 3.14. $f : (X, d_X) \to (Y, d_Y)$ is continuous on X if and only if $f^{-1}(U)$ is open in (X, d_X) for all U open in (Y, dy).

THEOREM 3.15. Let $\{U_i, i \in I\}$ be a family of (metric-) open sets in (X, d). Then

i. $U_i \cap U_j$ is open in (X, d) for any $i, j \in I$,

ii. $\bigcup_{i \in I} U_i$ is open in (X, d), and

iii. $\emptyset, X$ are open.

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 3.16. 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. $B \subset A$ 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. $B \subset A$ 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. $B \subset A$ 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 3.15 provides all properties of metric spaces needed to extend the concept of continuous maps on metric spaces to maps between more general spaces. We shall use these properties to define topological spaces.

DEFINITION 3.17. Let X be any set and let $\mathcal{T}$ be a collection of subsets of X with

i. $\bigcup_{U\in \mathcal{S}} U \in \mathcal{T}$ if $\mathcal{S} \subseteq \mathcal{T}$

ii. $\bigcap_{U\in \mathcal{S}} U \in \mathcal{T}$ if $\mathcal{S} \subseteq \mathcal{T}$ with $\mathcal{S}$ is a finite set

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

Then we call $\mathcal{T}$ a *topology* on the *topological space* X, the members U of $\mathcal{T}$ are called *(topology-) open* .

EXAMPLE 3.18.

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 3.15). 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, $S = \bigcup_{x\in S}\{x\}$.

REMARK 3.19. Recall that, using those properties of (metric-) open sets in a metric space (X, d) that the concept of continuity is based on, 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 3.16.)

By virtue of Theorem 3.14 we can extend the concept of continuous maps to general topological spaces:

DEFINITION 3.20. 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 3.21. 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. For $U \in \mathcal{S}$ we have $g^{-1}(U) \in \mathcal{F}$ since g is continuous and $(g \circ f)^{-1}(U) = f^{-1}(g^{-1}(U)) \in \mathcal{T}$ since f is continuous. Hence $g \circ f$ 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 3.22. 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 3.23. The topological spaces $(X, \mathcal{T})$ and $(Y, \mathcal{F})$ are called *homeomorph* if there exists a homeomorphism $f : X \longrightarrow Y$.

DEFINITION 3.24. 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 3.25. 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 topological space $(X, \mathcal{T}_d)$.

EXAMPLE 3.26. 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 $\left(f^{-1}(\cos(2\pi - \frac{1}{n}) + i \sin(2\pi - \frac{1}{n}))\right)_n = \left(2\pi - \frac{1}{n}\right)_n$ 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 3.27. In the following table we shall consider sequences in $\mathbb{R}$ where $\mathbb{R}$ is equipped with different topologies.

	$\mathcal{T}_{d_0} = \mathcal{P}(\mathbb{R})$	$\mathcal{T} = \{\emptyset, \mathbb{R}\}$	$\mathcal{T}_{d_2}$
$x_n = 1, \ \forall n \in \mathbb{N}$	$\lim_{n\to\infty} x_n = 1$	$\lim_{n\to\infty} x_n = x$ for any $x \in \mathbb{R}$	$\lim_{n\to\infty} x_n = 1$
$y_n = \frac{1}{n}, \ \forall n \in \mathbb{N}$	(y_n) does not converge	$\lim_{n\to\infty} y_n = y$ for any $y \in \mathbb{R}$	$\lim_{n\to\infty} y_n = 0$
$z_n = n, \ \forall n \in \mathbb{N}$	(z_n) does not converge	$\lim_{n\to\infty} z_n = z$ for any $z \in \mathbb{R}$	(z_n) does not converge
$u_n = (1 + \frac{1}{n})^n$	(u_n) does not converge	$\lim_{n\to\infty} u_n = u$ for any $u \in \mathbb{R}$	$\lim_{n\to\infty} (1 + \frac{1}{n})^n = e$

The ambivalence in column $\mathcal{T} = \{\emptyset, \mathbb{R}\}$ is only possible since the topology is 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.)

DEFINITION 3.28. A subset A of a topological space $(X, \mathcal{T})$ is called *closed* if $A^C = X \setminus A \in \mathcal{T}$, that is if A^C, the complement of A, is open.

THEOREM 3.29. Let (X, d) be a metric space, then A is closed in $(X, \mathcal{T}_d)$ if and only if given any sequence (x_n) in A with $x_n \to x_0 \in X$ then automatically $x_0 \in A$.

REMARK 3.30. The characterization of closed sets in metric spaces in Theorem 3.29 does not hold in general topological space.

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

THEOREM 3.31. 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 $d(x_0, x) < \delta$ implies $d(f(x_0), f(x)) < \varepsilon$.

ii. For all $\varepsilon > 0$ exists some $\delta > 0$ such that $f(B_\delta(x_0)) \subseteq B_\varepsilon(f(x_0))$.

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

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

THEOREM 3.32. 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 on X, that is for all $x_0 \in X$ and for all $\varepsilon > 0$ exists some $\delta > 0$ such that $d(x_0, x) < \delta$ implies $d(f(x_0), f(x)) < \varepsilon$.

ii. For all $x_0 \in X$ and for all $\varepsilon > 0$ exists some $\delta > 0$ such that $f(B_\delta(x_0)) \subseteq B_\varepsilon(f(x_0))$.

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

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

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

DEFINITION 3.33. 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 $\bar{A} = \bigcap_{\substack{C \supseteq A \\ C closed}} C$.

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

iv. A' denotes the *set of all cluster points*, that is
$A' = \{x_0 \in X \text{ s.t. there exists a sequence } (x_n) \text{ in } A \text{ with } \lim_{n\to\infty} x_n = x_0 \text{ and } x_n \neq x_0\}$.

3.3. Compactness

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 3.34. 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 3.35. 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 3.36.

i. Any finite set is compact.

ii. The set $\{\frac{1}{n} : n \in \mathbb{N}\}$ is not compact.

iii. The set $\{\frac{1}{n} : n \in \mathbb{N}\} \cup \{0\}$ is compact.

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 3.37. 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 3.38. 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.

Theorem 3.39. Any compact set A in (X,d) is bounded, that is, compact sets are bounded.

Theorem 3.40. Any infinite subset B of a compact set A in (X,d) has at least one cluster point in A.

Theorem 3.41. Any compact set is closed.

Theorem 3.41 combines with Theorem 3.39 to the statement that compact sets are closed and bounded. Does the converse hold? It would be nice, we would get a criterium for compactness which is easy to check. Sadly, the converse does not hold in general (see Remark 3.49, 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 3.42. 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 *Lebesgue number* for the covering $\mathfrak{U}$ of A.

Lemma 3.43. 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 $\mathfrak{U} = (U_i)_{i\in I}$ of X without a Lebesgue-number, that is for all $n \in \mathbb{N}$ we can choose some $x_n \in X$ such that for all $B_{\frac{1}{n}}(x_n) \not\subseteq U_i$ for all $i \in I$.

Since A is sequential compact, we can extract a convergent subsequence $(x_{n_k})_k$ of (x_n) and set $x_0 := \lim_k x_{n_k} \in X$. Since $\mathfrak{U}$ is a covering, we have $x_0 \in U_{i_0}$ for some $i_0 \in I$. Since U_{i_0} is open, there is an $n \in \mathbb{N}$ such that $B_{\frac{1}{n}}(x_0) \subseteq U_{i_0}$.

Pick $K \in \mathbb{N}$ such that $K \geq 2n$ and $d(x_{n_K}, x_0) < \frac{1}{2n}$. We have $B_{\frac{1}{n_K}}(x_{n_K}) \subseteq B_{\frac{1}{n}}(x_0)$ since $d(x, x_{n_{2n}}) < \frac{1}{2n}$ implies $d(x_0, x) < d(x_0, x_{n_K}) + d(x_{n_K} - x) < \frac{1}{2n} + \frac{1}{2n} = \frac{1}{n}$..

We conclude that $B_{\frac{1}{n_K}}(x_{n_K}) \subseteq B_{1/n}(x_0) \subseteq U_{i_0}$, a contradiction. □

Now we shall provide the main result of this chapter.

Theorem 3.44. Let (X,d) be a metric space. A set $A \subseteq X$ is sequentially compact if and only if it 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 compact sets whose diameter tends to zero is a single point $x_0 \in A$ (check!), we get by construction, $x_{n_k} \to x_0$. Since A is closed, we have $x \in A$.

Let us now suppose that A is sequentially compact. Let $\mathcal{U} = (U_i)_{i \in I}$ be an arbitrary open cover. We want to show that $\mathcal{U}$ admits a finite subcover. By Lemma 3.43, this cover has a Lebesgue-number $\lambda > 0$: every $x \in X$ has an $i = i(x)$ such that $B_\lambda(x) \subseteq U_{i(x)}$.

Choose any $x_1 \in X$. Then either $U_1 := U_{i(x_1)}$ covers X and we are done. Otherwise choose any $x_2 \in X \setminus U_1$ and set $U_2 := U_{i(x_2)}$. Again, either $U_1 \cup U_2$ already covers X and we are done, or we can choose $x_3 \in X \setminus (U_1 \cup U_2)$ 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 3.45. For $a \leq b$ we have $[a, b]$ is compact in $\mathbb{R}$. (Recall, if not specified we let $d = d_2$ in $\mathbb{R}^n$.)

Lemma 3.46. 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 $(\mathbb{R}^{n+m}, d_k)$ does 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 3.47. Any set of the form $[a_1, b_1] \times [a_2, b_2] \times \ldots \times [a_n, b_n] \subset \mathbb{R}^n$ are compact.

Proof. Proof by induction using Lemma 3.46. □

Theorem 3.48. (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 3.47. Since A is therefore a closed subset of a compact set, we have A compact by Theorem 3.38. □

Remark 3.49. 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

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

contains no finite $\mathcal{U}$–subcovering of A.

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 3.50. A compact metric space (X,d) is complete.

Theorem 3.51. 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).

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

Definition 3.52. 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. $d_Y(f(x), f(y)) < \varepsilon$ 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 3.53. Any uniformly continuous function $f : (X,d_X) \longrightarrow (Y,d_Y)$ is continuous.

Example 3.54.

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 3.55. 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. (See homework problem 11.2.)

3.4. Connectedness

Again, we constrain ourselves to metric spaces.

Definition 3.56. A metric space (X, d) is *connected* if X and $\emptyset$ are the only *clopen*, that is, open and closed, subsets of X.

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 $\emptyset = U \cap V$.

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

Proposition 3.57. A metric space (X, d) is connected if and only if there exists no separation of X.

The most important result of this section is fairly elementary:

Theorem 3.58. If (X, d) is connected and $f : (X, d_X) \longrightarrow (Y, d_Y)$ is continuous, then $\mathcal{R}_f = f(X)$ is connected.

Remark 3.59. 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 3.60. 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 a (possibly unbounded) interval.

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

Corollary 3.61. Let $f : [a, b] \longrightarrow \mathbb{R}$ be continuous. Then exists $c, d \in \mathbb{R}$ with $f([a, b]) = [c, d]$.

THEOREM 3.62. (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 3.63. 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.

EXAMPLE 3.64. 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 3.63 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 3.65. 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 3.66. 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.

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

Definition 3.67. 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,

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

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 $C(X)$. Note that (X, d) being compact implies that all continuous functions defined on X are bounded and we have $C(X) \subseteq B(X)$, and, therefore, $C(X)$ inherits the metric d_∞ from $B(X)$.

Definition 3.68. Let (X, d_X) and (Y, d_Y) be metric spaces and let $f_n : \ X \longrightarrow Y$, $n \in \mathbb{N}$ be a sequence of functions mapping X to Y.

The sequence $(f_n)_{n\in\mathbb{N}}$ *converges pointwise* to $f_0 : \ X \longrightarrow Y$, if $\lim\limits_{n\to\infty} f_n(x) = f_0(x)$ for all $x \in X$, that is, if $\lim\limits_{n\to\infty} d_Y(f_n(x), f_0(x))$ for all $x \in X$.

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

$$d_Y(f_n(x), f_0(x)) < \epsilon \quad \text{ for all } x \in X \text{ and for all } n \geq N.$$

That is

$$\lim_{n\to\infty} \sup \{ d_Y(f_n(x), f_0(x)) : \ x \in X \} = 0.$$

Proposition 3.69. The sequence (f_n) converges in $(B(X), d_\infty)$ to f_0 if and only if (f_n) converges to $f_0 : X \longrightarrow \mathbb{C}$ uniformly.

Theorem 3.70. Let (f_n) be a sequence of continuous functions in $(B(X), d_\infty)$ which converges to f_0. Then f_0 is continuous and for any sequence (x_k) in X with $\lim\limits_{k\to\infty} x_k = x_0$ we have

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

Corollary 3.71. If (X, d) is compact, then $C(X)$ is a closed subspace of $B(X)$.

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, by Theorem 3.70 we have $C(X)$ closed in $(B(X), d_\infty)$. □

Theorem 3.72. Let (X, d) be a compact metric space. Then $(C(X), d_\infty)$ is a complete metric space.

4. DIFFERENTIATION

4.1. Central results

In this section, we shall discuss derivatives of real valued functions defined on subsets of $\mathbb{R}$. Our main objective is to illuminate the interplay of continuity and differentiability.

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 $\displaystyle\lim_{h\to 0} \frac{f(x_0 + h) - f(x_0)}{h} = L.$

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.

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

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 $\left(\frac{f}{g}\right)'(x_0) = \dfrac{f'(x_0)g(x_0) - f(x_0)g'(x_0)}{g(x_0)^2}$.

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 $f(x_0) \in V$. Then $g \circ f$ is differentiable at x_0 and we have $(g \circ f)'(x_0) = g'(f(x_0))f'(x_0)$.

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 on U. If there exists a maximum [resp. minimum] of f at c, then $f'(c) = 0$.

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 $c \in (a, b)$ such that $f'(c) = 0$.

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 $c \in (a, b)$ such that $f'(c) = \frac{f(b)-f(a)}{b-a}$.

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 "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 $x \in (a, b)$.

Example 4.16. Discussion of x^n, $n \in \mathbb{N}_0$, including the remark 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 $\phi : [c, d] \to \mathbb{R}$ be the inverse function of f. If f is differentiable at $x_0 \in (a, b)$ with $f'(x_0) \neq 0$, then ϕ is differentiable at $y_0 = f(x_0) \in (c, d)$ and $\phi'(y_0) = \dfrac{1}{f'(x_0)} = \dfrac{1}{f'(\phi(y_0))}$.

Definition 4.18. 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 4.19. 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 $g(x) \to -\infty$.

4.2. Taylor series

DEFINITION 4.20. 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 A has the property that all its members are cluster points, that is, $A \subseteq A'$. For example, we could consider $C^2([0,1])$.

REMARK 4.21. 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 4.22. 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

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

REMARK 4.23. 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 are small.

COROLLARY 4.24. 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 4.25. If for $f \in C^\infty(a,b)$ there exists $M > 0$ with $|f^{(n)}(\xi)| \leq M$ 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 4.26. For $f \in C^\infty(a,b)$ and $x_0 \in (a,b)$, call the formal power series

$$T_{f,x_0}(x) = \sum_{k=0}^{\infty} \frac{f^{(k)}(x_0)}{k!}(x - x_0)^k, \quad x \in (a,b)$$

Taylor series of f at x_0.

Remark 4.27.

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 4.28. 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 4.29. 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 4.28. □

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 4.30. 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 4.31. 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 4.32.

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 4.33. 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 4.34. We define the *sine* function $\sin : \mathbb{R} \longrightarrow \mathbb{R}$ by setting $\sin(x) = \operatorname{Im} \exp(ix)$ for $x \in \mathbb{R}$ and the *cosine* function $\cos : \mathbb{R} \longrightarrow \mathbb{R}$ by setting $\cos(x) = \operatorname{Re} \exp(ix)$ 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 4.35.

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 4.36. $\exp : \mathbb{C} \longrightarrow \mathbb{C}$ 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 $z \in \mathbb{C}$. $\square$

4.4. Fixed point theorems and approximative methods

Definition 4.37. An element $x_0 \in X$ is called *fixed point* of $f : X \longrightarrow X$ if $f(x_0) = x_0$.

Definition 4.38. A *contraction* on a metric space (X, d) is a map $f : X \longrightarrow X$ such that for some constant k, $0 \leq k < 1$, we have

$$d(f(x), f(y)) \leq k\, d(x, y) \text{ for all } x, y \in X\,.$$

Proposition 4.39. Contractions are uniformly continuous mappings.

Theorem 4.40. 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 $x_0 \in X$, and for any choice of $x_1 \in X$, the sequence (x_n) defined by

$$x_1, x_2 = f(x_1), x_3 = f(x_2) = f(f(x_1)) = f \circ f(x_1), \ldots, x_{n+1} = f(x_n), \ldots\,,$$

converges to x_0. Moreover, we have

$$d(x_n, x_0) \leq \frac{k}{1-k}\; d(x_k, x_{k-1}) \leq \frac{k^{n-1}}{1-k}\; d(x_2, x_1).$$

Theorem 4.41. Newton's Method.[4] 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 4.42. 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:

$$|\xi - x_n| \leq \frac{q}{1-q}|x_n - x_{n-1}| \leq \frac{q^n}{1-q}|x_1 - x_0|.$$

[4]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 $P = \{x_0, x_1, \ldots, x_N\}$ such that $a = x_0 < x_1 < x_2 < \ldots < x_{N-1} < x_N = b$. 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 $T = \{t_1, t_2, \ldots, t_N\}$ with $a = x_0 \le t_1 \le x_1 \le t_2 \le x_2 \le \ldots \le x_{N-1} \le t_N \le x_N = b$.

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 $P = \{x_0, x_1, \ldots, x_N\}$ with $\operatorname{mesh} P < \delta$ and $T = \{t_1, t_2, \ldots, t_N\}$.

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 $T' = \{t_1, \ldots, t_{n-1}, s_n, t_{n+1}, \ldots t_N\ )\}$ 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 $\displaystyle\int_a^b f(x)\,dx \le \int_a^b g(x)\,dx$.

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

$$L(f, P) = \sum_{n=1}^{N} m_n(x_n - x_{n-1}), \quad m_n = \inf\{f(x) : \ x \in [x_{n-1}, x_n]\} = \inf\, f([x_{n-1}, x_n]) \quad \text{for } n = 1, \dots N$$

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, $\displaystyle\int_a^b f(x)\, dx = I(f)$.

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 $x \in \mathbb{R}\backslash\{\dots, -4\pi, -2\pi, 0, 2\pi, 4\pi, \dots\}$ 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 $\int_b^a f(x)\, dx$. 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

$$\operatorname{osc}_{x_0}(f) = \limsup_{x\to x_0} f(x) - \liminf_{x\to x_0} f(x) = \lim_{n\to\infty} \left(\sup f([x_0 - \tfrac{1}{n}, x_0 + \tfrac{1}{n}]) - \inf f([x_0 - \tfrac{1}{n}, x_0 + \tfrac{1}{n}]) \right)$$

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

$$\int_a^b f(y)\,dx = \int_c^d f(g(x))\,g'(x)\,dx.$$

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 $\lim\limits_{c \nearrow b} \int_a^c f(x)\,dx$ exists[5], 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 $\lim\limits_{c_2 \nearrow c} \int_c^{c_2} f(x)\,dx$ and $\lim\limits_{c_1 \searrow a} \int_{c_1}^{c} f(x)\,dx$ 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. $\lim\limits_{R \to \infty} \int_R^{-R} x\,dx$ 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 $\sum\limits_{n=1}^{\infty} f(n)$ converges if and only if the improper integral $\int_1^{\infty} f(x)\,dx$ exists.

PROPOSITION 5.39. WALLIS' PRODUCT. $\quad \dfrac{\pi}{2} = \prod\limits_{n=1}^{\infty} \dfrac{4n^2}{4n^2 - 1}.$

[5] We write $\lim\limits_{c \nearrow b}$ for $\lim\limits_{\substack{c \to b \\ c < b}}$ and $\lim\limits_{c \searrow a}$ for $\lim\limits_{\substack{c \to a \\ c > a}}$.

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 $\langle \cdot, \dot{\rangle}$ 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 $\langle f, g \rangle = \int_0^1 f(x)\overline{g(x)}\, dx$,
- the corresponding norm $\|f\|_2 = \sqrt{\langle f, f \rangle} = \sqrt{\int_0^1 |f(x)|^2 \, dx}$,
- and the corresponding metric $d_2(f, g) = \|f - g\|_2 = \sqrt{\int_0^1 |f(x) - g(x)|^2 \, dx}$.

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 $f = \lim\limits_{N\to\infty} S(f,N)$ with convergence in $(\mathcal{R}'([0,1]), d_2)$.

Theorem 5.66. For any $f \in \mathcal{R}'([0,1])$ we have $f = \lim\limits_{N\to\infty} S(f,N)$ 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

$$\begin{aligned}
f(x) &= \mathrm{Re} f(x) = \mathrm{Re} \sum_{k=-\infty}^{\infty} \widehat{f}(k) e_k(x) \\
&= \sum_{k=-\infty}^{\infty} \mathrm{Re}\big((\mathrm{Re}\widehat{f}(k) + i\mathrm{Im}\widehat{f}(k))(\cos 2\pi kx + i \sin 2\pi kx)\big) \\
&= \sum_{k=-\infty}^{\infty} \big(\mathrm{Re}\widehat{f}(k) \cos 2\pi kx - \mathrm{Im}\widehat{f}(k) \sin 2\pi kx\big) \\
&= \mathrm{Re}\widehat{f}(0) + \sum_{k=1}^{\infty} \big(\mathrm{Re}\widehat{f}(k) \cos 2\pi kx - \mathrm{Im}\widehat{f}(k) \sin 2\pi kx\big) \\
&\quad\quad + \big(\mathrm{Re}\widehat{f}(-k) \cos(-2\pi kx) - \mathrm{Im}\widehat{f}(k) \sin(-2\pi xn)\big) \\
&= \mathrm{Re}\widehat{f}(0) + \sum_{k=1}^{\infty} \big(\mathrm{Re}\widehat{f}(k) + \mathrm{Re}\widehat{f}(-k)\big) \cos 2\pi kx - \big(\mathrm{Im}\widehat{f}(k) - \mathrm{Im}\widehat{f}(-k)\big) \sin 2\pi kx.
\end{aligned}$$

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

$$\begin{aligned} d(L(x), L(y)) &= \|L(x) - L(y)\|_W = \|L(x - y)\|_W = \|L(\sum_{k=1}^n (x_k - y_k) e_k)\|_W \\ &= \|\sum_{k=1}^n (x_k - y_k) L(e_k)\|_W \ \le \sum_{k=1}^n |x_k - y_k| \, \|L(e_k)\|_W \\ &\le \sum_{k=1}^n \|x - y\| \, \|L(e_k)\|_W < n\delta M = \epsilon \end{aligned}$$

□

We can now use analysis to show

Theorem 6.2. Let W be a 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 $\|L\|_{\mathcal{L}} = \sup\left\{\frac{\|L(x)\|}{\|x\|} : x \in \mathbb{R}^n \setminus \{0\}\right\}$.

Further, for all $x \in \mathbb{R}^n$ we have $\|L(x)\|_{\mathbb{R}^m} \le \|L\|_{\mathcal{L}} \|x\|_{\mathbb{R}^n}$.

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. □

6.2. Curves

Curves are functions mapping intervals in $\mathbb{R}$ into $\mathbb{R}^m$, that is

$$f = \begin{pmatrix} f_1 \\ f_2 \\ \vdots \\ f_m \end{pmatrix} : \ I \longrightarrow \ \mathbb{R}^m, \qquad t \mapsto \begin{pmatrix} f_1(t) \\ f_2(t) \\ \vdots \\ f_m(t) \end{pmatrix}. \tag{1}$$

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 1, we set $f' = (f_1', f_2', \dots, f_m')^T$.

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 $P = \{x_0, \dots, x_N\}$ 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

$$\|\gamma'(t)\|_2^2 = \left(\frac{d\gamma_1}{dt}(t)\right)^2 + \dots + \left(\frac{d\gamma_m}{dt}(t)\right)^2 > 0\,.$$

where $\gamma'(t) = \left(\dfrac{d\gamma_i}{dt}(t)\right)_{i=1,\dots,m} = \left(\dfrac{d\gamma_1}{dt}(t), \dots, \dfrac{d\gamma_m}{dt}(t)\right)^T.$

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 $\Lambda(\gamma) = \displaystyle\int_a^b \|\gamma'(t)\|_2\, dt$.

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.

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 $f : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ is a linear map, we have $(Df)_a = f$ for all $a \in \mathbb{R}^n$.

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

$$(Df)_a(x) = \lim_{t \to 0} \frac{1}{t}\big(f(a+tx) - f(a)\big) \quad \text{for any } x \in \mathbb{R}^n.$$

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.

Note that any function $f : U \longrightarrow \mathbb{R}^m$, $U \subset \mathbb{R}^n$, can be represented by m real valued functions $f_1, f_2, \ldots, f_m$ via $f(x) = (f_1(x), f_2(x), \ldots, f_m(x))$, $x \in U$. 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 $f = (f_1, f_2, \ldots, f_m) : U \longrightarrow \mathbb{R}^m$ 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

$$\begin{pmatrix} \frac{\partial f_1}{\partial x_1}(a) & \frac{\partial f_1}{\partial x_2}(a) & \cdots & \frac{\partial f_1}{\partial x_n}(a) \\ \frac{\partial f_2}{\partial x_1}(a) & \frac{\partial f_2}{\partial x_1}(a) & \cdots & \frac{\partial f_i}{\partial x_n}(a) \\ \vdots & \vdots & & \vdots \\ \frac{\partial f_m}{\partial x_1}(a) & \frac{\partial f_m}{\partial x_2}(a) & \cdots & \frac{\partial f_m}{\partial x_n}(a) \end{pmatrix},$$

as *Jacobian matrix* of f at a.

If $m = 1$, then we call the Jacobian matrix *gradient* and denote it by grad$f$ 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 $J_f(x) = \det \left(\frac{\partial f_i}{\partial x_j}\right)_{i,j=1,\dots n}$ is called *Jacobian determinant* (or simply *Jacobian*).

The *divergence* of f is the function $\operatorname{div} f = \sum_{i=1}^{n} \frac{\partial f_i}{\partial x_i}$.

If $m = n = 3$, then we call the vector field $\operatorname{rot} f = \left(\frac{\partial f_3}{\partial x_2} - \frac{\partial f_2}{\partial x_3}, \frac{\partial f_1}{\partial x_3} - \frac{\partial f_3}{\partial x_1}, \frac{\partial f_2}{\partial x_1} - \frac{\partial f_1}{\partial x_2}\right)$ *rotation* of f.

Remark 6.15. If $f : U \longrightarrow \mathbb{R}$, $U \subseteq \mathbb{R}^n$ open, is differentiable at $a \in U$ 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.

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

$$[(Df)_a]_{ij} = \frac{\partial f_i(p)}{\partial x_j}(a)$$

The existence of partial derivatives at a does not imply differentiability of f at a. (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 $D(g \circ f)_a = D(g)_{f(a)} \circ D(f)_a$.

Easy but very important is

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 $\pi_i \circ (Df)_x = (Df_i)x$.

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.

We shall not prove not 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

$$[a, x_1] = \{(1 - t) \cdot a + t \cdot x_1 : t \in [0, 1]\} \subset \mathbb{R}^n$$

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!

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 $i = 1, \dots, n$ and $j = 1, \dots m$, 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 $f : U \longrightarrow \mathbb{R}^m$, $U \subseteq \mathbb{R}$. 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 $(Df)_a = T_{a,a}$.

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

$$F'(y) = \int_a^b \frac{\partial f}{\partial y}(x,y)\, dx, \quad y \in (c,d)\,.$$

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

$$f(x) = f(a) + (Df)_a(x-a) + R(x) \quad \text{with} \lim_{x \to a} \frac{1}{\|x-a\|} R(x) = 0.$$

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 $(D^2 f)_a \in \mathcal{L}(\mathbb{R}^n, \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)) =: \mathcal{L}^2(\mathbb{R}^n, \mathbb{R}^m)$. In addition, we can consider $(D^2 f)_a$ as bilinear form, that is, we can write $(D^2 f)_a : \mathbb{R}^n \times \mathbb{R}^n \longrightarrow \mathbb{R}^m$.

If $(D^2 f)_x$ is defined for all x in a neighborhood of $a \in U$, and if $D^2 f$ is differentiable at $a \in U$, then we call $(D^3 f)_a = (D(D^2 f))_a$ *third derivative* of f at a. Note that $(D^3 f)_a \in \mathcal{L}(\mathbb{R}^n, \mathcal{L}(\mathbb{R}^n, \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m))) =: \mathcal{L}^3(\mathbb{R}^n, \mathbb{R}^m)$.

...

If $(D^r f)_x$ is defined for all x in a neighborhood of $a \in U$, and if $D^r f$ is differentiable at $a \in U$, then we call $(D^{r+1} f)_a = (D(D^2 f))_a$ 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 $D^r f : \mathbb{R}^n \longrightarrow \mathcal{L}^r(\mathbb{R}^n, \mathbb{R}^m)$ 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 $1 \le r \le \infty$. A bijective function $f : U \longrightarrow V$ is called C^r*–diffeomorphism* if $f \in C^r(U, \mathbb{R}^m)$ and $f^{-1} \in C^r(V, \mathbb{R}^m)$.

Definition 6.29. For $f = (f_1, \dots, f_m) : U \longrightarrow \mathbb{R}^m$, $U \subseteq \mathbb{R}^n$ open, we refer to $\dfrac{\partial^2 f_k}{\partial x_i \partial x_j}(a) \in \mathbb{R}$, $i, j = 1, \dots n$, $k = 1, \dots, m$ 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.

Theorem 6.30. Let $f = (f_1, \dots, f_m) : U \longrightarrow \mathbb{R}^m$, $U \subseteq \mathbb{R}^n$ open. If $(D^2 f)_a$ exists, then for $k = 1, \dots, n$, $(D^2 f_k)_a$ exists, the Hessians $\mathrm{Hess} f_k(a)$, $k = 1, \dots, m$ exist and

$$(D^2 f_k)_a(e_i, e_j) = \frac{\partial^2 f_k}{\partial x_i \partial x_j}(a), \quad i, j = 1, \dots, n.$$

Also, if all second partials of f exist on U and if they are all continuous at a, then Df is differentiable at a.

Theorem 6.31. Let $f : U \longrightarrow \mathbb{R}^m$, $U \subseteq \mathbb{R}^n$ open. If $(D^2 f)_a$ exists for $a \in U$, then

$$(D^2 f)_a(u, v) = (D^2 f)_a(v, u).$$

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 $\nabla f(a) = 0$.

DEFINITION 6.35. Let $A \in \mathbb{R}^{m \times n}$ be symmetric.

i. A is called *positive definite* if $\langle \xi, A\xi \rangle > 0$ for all $\xi \in \mathbb{R}^n \setminus \{0\}$.

ii. A is called *negative definite* if $\langle \xi, A\xi \rangle < 0$ for all $\xi \in \mathbb{R}^n \setminus \{0\}$.

iii. A is called *positive semidefinite* if $\langle \xi, A\xi \rangle \geq 0$ for all $\xi \in \mathbb{R}^n \setminus \{0\}$.

iv. A is called *negative semidefinite* if $\langle \xi, A\xi \rangle \leq 0$ for all $\xi \in \mathbb{R}^n \setminus \{0\}$.

v. A is called *indefinite* if $\langle \xi, A\xi \rangle > 0$ and $\langle \eta, A\eta \rangle < 0$ for some $\xi, \eta \in \mathbb{R}^n$.

LEMMA 6.36. 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 6.37. 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 $(s_1, \ldots, s_n)$ such that each s_i is a non-negative integer, and $s_i + \ldots + s_n \leq m - 1$, and the remainder r satisfies

$$\lim_{x \to 0} \frac{r(x)}{\|x\|^{m-1}} = 0.$$

THEOREM 6.38. 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

$$f(x) = f(a) + (Df)_a(x - a) + \frac{1}{2}(Df)_a(x - a, x - a) + R(x) \text{ with } \lim_{x \to a} \frac{1}{\|x - a\|^2} R(x) = 0.$$

THEOREM 6.39. Let $f : U \longrightarrow \mathbb{R}$, $U \subseteq \mathbb{R}^n$ open, $f \in C^2(U)$ be given with $\nabla f(a) = 0$ 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.)

6.6. Implicit functions and the inverse function theorem

DEFINITION 6.40. 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

$$(2)\ \ f(x,y) = \begin{pmatrix} f_1(x_1,\dots,x_n,y_1,\dots,y_m) \\ f_2(x_1,\dots,x_n,y_1,\dots,y_m) \\ \vdots \\ f_m(x_1,\dots,x_n,y_1,\dots,y_m) \end{pmatrix} = \begin{pmatrix} z_1 \\ z_2 \\ \vdots \\ z_m \end{pmatrix} = c \ \text{ for } (x,y) \in B_\epsilon(a,b),\ \epsilon > 0.$$

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 (2) holds if and only if

$$y = g(x) \quad \text{ for some } x \in B_\epsilon(a),$$

that is, all solutions to (2) 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 (2).

THEOREM 6.41. 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 $f \in C^r(U)$, $1 \le r \le \infty$. If $f(a,b) = c \in \mathbb{R}^m$ and $B = \frac{\partial f}{\partial y}(a,b) = \left(\frac{\partial f_i}{\partial y_j}(a,b)\right)_{i,j=1,\dots m}$ is invertible, then exists $\epsilon > 0$ and $g : B_\epsilon \longrightarrow \mathbb{R}^m$, $g \in C^r(B_\epsilon(a))$ with

$$\{(x,y) \in B_\epsilon(a,b) : \ f(x,y) = c\} = \{(x, g(x)), \quad x \in B_\epsilon(a)\} = \Gamma_g.$$

COROLLARY 6.42. 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

$$\frac{\partial g}{\partial x}(x) = -\left(\frac{\partial f}{\partial y}(x, g(x))\right)^{-1} \frac{\partial f}{\partial x}(x, g(x)) \quad \text{for } x \in B_\epsilon(a)\,.$$

Proof. See proof of the Implicit Function Theorem. □

REMARK 6.43. Note that 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 (f^{-1} is well defined on (c,d)),
- homeomorph (f^{-1} is continuous on (c,d)),
- diffeomorph (f^{-1} is continuously differentiable on (c,d)),

and we have $(f^{-1})'(y) = \dfrac{1}{f'(f^{-1}(y))}$ for $y \in (c,d)$.

This theorem does not extend to higher dimensions, for example, consider $f(x) = (\cos(x), \sin(x))$, $x \in \mathbb{R}$.

Theorem 6.44. Inverse Function Theorem. Let $U, V \subseteq \mathbb{R}^n$ be open and $f : U \longrightarrow V$ be given with $f \in C^r(U)$, $1 \leq r \leq \infty$. If $(Df)_a \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^n)$ is an isomorphism, then exists an open neighborhood $U' \subseteq U$ of a and an an open neighborhood $V' \subseteq V$ of $f(a)$ such that $f : U' \longrightarrow V'$ is a C^r–diffeomorphism.

Theorem 6.45. Lagrange Multipliers. Let $f, h \in C^1(U)$, $U \subseteq \mathbb{R}^n$ open, be given with $\nabla f(a) = (Df)_a \neq 0$ and $h(a) > h(x)$ for all $x \in B_\epsilon(a) \cap \{x : \ f(x) = 0\}$, $\epsilon > 0$. Then $\nabla f(a) = \lambda \nabla g(a)$ 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 $U \subseteq \mathbb{R}^{n-1} \times \mathbb{R}$. The implicit function theorem provides a function g. Differentiation of the equation $f(x_1, \ldots, x_{n-1}, g(x_1, \ldots, x_{n-1})) = 0$ and the function $H(x_1, \ldots, x_{n-1}) = h(x_1, \ldots, x_{n-1}, g(x_1, \ldots, x_{n-1}))$ at $(a_1, \ldots, a_{n-1})$ provides two equations which combine to proof the result. □

7. INTEGRATION ON $\mathbb{R}^D$

We shall focus on the case $d = 2$, that is, we shall discuss integrals $\displaystyle\int_A f(x)\,dx$ 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 $\mathrm{mesh}(G) = \max\limits_{R_{ij}\in G} \mathrm{diam}\, R_{ij} < \delta$.

The number I is called *Riemann integral* of f on R and is denoted by $I = \displaystyle\int f = \int_R f(x, y)\, d(x, y)$, 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 $\displaystyle\underline{\int} f = \sup\{L(f, G),\ G \text{ grid in } R\}$, and the *upper integral* of f is $\displaystyle\overline{\int} f = \inf\{U(f, G),\ G \text{ grid in } R\}$

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 $\underline{\int} f = \overline{\int} f$.

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 discontinuouities of f is a zeroset.

THEOREM 7.7. FUBINI'S THEOREM. Let $f : R = [a, b] \times [c, d] \longrightarrow \mathbb{R}$ be Riemann integrable.

i. The functions

$$\underline{F} : [c, d] \longrightarrow \mathbb{R}, \quad \underline{F}(y) = \underline{\int_a^b} f(x, y)\, dx \quad \text{and} \quad \overline{F} : [c, d] \longrightarrow \mathbb{R}, \quad \overline{F}(y) = \underline{\int_a^b} f(x, y)\, dx$$

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 $f(\cdot, y)$ 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

$$\int_c^d \int_a^b f(x, y)\, dx\, dy = \int_R f = \int_a^b \int_c^d f(x, y)\, dy\, dx.$$

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 $n = 1, 2, \ldots$, 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 $E_i \subseteq \mathbb{R}^n$, $i = 1, \ldots, \infty$, 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 $\{S_k\}_{k=1,\ldots,N}$ 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.

THEOREM 7.15. CHANGE OF VARIABLES I. If $\varphi : D_t \longrightarrow D_x$, $D_t, D_x \subseteq \mathbb{R}^n$ open and bounded, is a diffeomorphism, $f \in \mathcal{R}(D_x)$ and $\operatorname{supp} f$ is a compact subset of D_x. Then $f \circ \varphi \cdot |\det \varphi'| \in \mathcal{R}(D_t)$ and

$$(3) \qquad \int_{D_x} f = \int_{D_x=\varphi(D_t)} f(x)\,dx = \int_{D_t} f(\varphi(t))|\det \varphi'(t)|\,dt = \int_{D_t} f \circ \varphi |\det \varphi'|.$$

Here, φ' denotes the Jacobian matrix of φ and, hence, $\det \varphi'$ is the Jacobian J_φ.

The proof of this theorem is quite involved and we shall have to provide some additional vocabulary and proof first some special cases.

DEFINITION 7.16. For $D \subseteq \mathbb{R}^n$ and $f : D \longrightarrow \mathbb{R}$, we define the *support of* f as $\operatorname{supp} f = \{x \in D : \ f(x) \neq 0\}$.

LEMMA 7.17. If $\varphi : D_t \longrightarrow D_x$, $D_t, D_x \subseteq \mathbb{R}^n$ open and bounded, is a diffeomorphism, then $\varphi(E_t)$ is a zeroset whenever $E_t \subseteq D_t$ is a zeroset.

LEMMA 7.18.

i. If I_t, I_x are open and bounded intervals in $\mathbb{R}$ and $\varphi : I_t \longrightarrow I_x$ is a diffeomorphism, then $f \circ \varphi \cdot |\varphi'| \in \mathcal{R}(I_t)$ whenever $f \in \mathcal{R}(I_x)$ and $\int_{I_x} f(x)\,dx = \int_{I_t} f \circ \varphi(t)|\varphi'(t)|dt$.
ii. If U_t, U_x are open and bounded sets in $\mathbb{R}$ and $\varphi : U_t \longrightarrow U_x$ is a diffeomorphism, then $f \circ \varphi \cdot |\varphi'| \in \mathcal{R}(U_t)$ whenever $f \in \mathcal{R}(U_x)$ and $\int_{U_x} f(x)\,dx = \int_{U_t} f \circ \varphi(t)|\varphi'(t)|dt$.

LEMMA 7.19. Equality (3) holds if $\varphi : D_t \longrightarrow D_x$ satisfies

$$x_1 = \varphi(t_1, \ldots, t_n) = t_1, \ \ \ldots \ , \ x_{n-1} = \varphi(t_1, \ldots, t_n) = t_{n-1}, \ x_n = \varphi^n(t_1, \ldots, t_n),$$

that is, if only x_n depends on $(t_1, \ldots, t_n)$ in a non–trivial manner.

LEMMA 7.20. If equality (3) holds for $\psi : D_\tau \longrightarrow D_t$ and $\varphi : D_t \longrightarrow D_x$, then equality (3) holds for $\varphi \circ \psi : D_\tau \longrightarrow D_x$.

THEOREM 7.21. CHANGE OF VARIABLES II. Let $D_t, D_x \subseteq \mathbb{R}^n$ be Jordan domains, $S_t \subseteq D_t$ and $S_x \subseteq D_x$ be zerosets such that $D_t \setminus S_t$ and $D_x \setminus S_x$ are open and $\varphi : D_t \setminus S_t \longrightarrow D_x \setminus S_x$, is a diffeomorphism. Then for all $f \in \mathcal{R}(D_x)$ we have $f \circ \varphi \cdot |\det \varphi'| \in \mathcal{R}(D_t \setminus S_t)$ and

$$\int_{D_x} f = \int_{D_x=\varphi(D_t)} f(x)\,dx = \int_{D_t} f(\varphi(t))|\det \varphi'(t)|\,dt = \int_{D_t\setminus S_t} f \circ \varphi |\det \varphi'|..$$

If in addition, $|\det \varphi'|$ can be extedned to D_t as a bounded function, then

$$\int_{D_x} f = \int_{D_x=\varphi(D_t)} f(x)\,dx = \int_{D_t} f(\varphi(t))|\det \varphi'(t)|\,dt = \int_{D_t} f \circ \varphi |\det \varphi'|.$$

7.4. Multivariate improper integrals

Definition 7.22. 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 $\operatorname{vol}(B_r(0) \setminus C_k) \to 0$ as $k \to \infty$ for all $r > 0$.

Definition 7.23. 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 $\operatorname{vol}(C_k \cap M) \to K$ as $k \to \infty$.

Definition 7.24. 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 7.25. $\displaystyle\int_{\mathbb{R}^2} e^{-(x^2+y^2)} d(x, y) = \pi$ and therefore $\displaystyle\int_{-\infty}^{\infty} e^{-x^2} dx = \sqrt{\pi}$.

7.5. The Gamma function

DEFINITION 7.26. 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 7.27. 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 7.28. 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 7.29. 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 7.30. 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 7.31. $B(x, y) = \dfrac{\Gamma(x)\Gamma(y)}{\Gamma(x + y)}$.

PROPOSITION 7.32. 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,

$\displaystyle\lim_{n\to\infty} \frac{n!}{\left(\frac{n}{e}\right)^n \sqrt{2\pi n}} = 1$

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

$$y' = f(x, y) \tag{4}$$

is called *first order differential equations.*

A solution to (4) 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

$$y' = f(x, y) \tag{5}$$

is called a *system of n first order differential equations.*

A solution to (5) 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 $G \subset \mathbb{R} \times \mathbb{R}^n$ and $f : G \longrightarrow \mathbb{R}$, $(x, y) \mapsto f(x, y)$ be a continuous function. The formal expression

$$y^{(n)} = f(x, y, y', y'', \dots, y^{(n-1)}) \tag{6}$$

is called *n-th order differential equations.*

A solution to (6) is a function $\varphi : I \longrightarrow \mathbb{R}$, I being an interval, such that the graph set $\Gamma_\varphi^{(n-1)} = \{(x, \varphi(x), \varphi'(x), \varphi''(x), \dots, \varphi^{(n-1)}(x)) :\ x \in I\}$ satisfies $\Gamma_\varphi^{(n-1)} \subseteq G$ and

$$\varphi^{(n)}(x) = f(x, \varphi(x), \varphi'(x), \varphi''(x), \dots, \varphi^{(n-1)}(x)) \quad \text{for all } x \in I.$$

REMARK 8.5. 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

$$y^{(n)} = f(x, y, y', y'', \dots, y^{(n-1)}), \tag{7}$$

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

$$z' = \begin{pmatrix} z'_0 \\ z'_1 \\ z'_2 \\ \vdots \\ z'_{n-2} \\ z'_{n-1} \end{pmatrix} = \begin{pmatrix} z_1 \\ z_2 \\ z_3 \\ \vdots \\ z_{n-1} \\ f(x, z_0, z_1, \dots, z_{n-1}) \end{pmatrix} = \begin{pmatrix} F_1\big(x, z_0, z_1, \dots z_{n-1}\big) \\ F_2\big(x, z_0, z_1, \dots z_{n-1}\big) \\ F_3\big(x, z_0, z_1, \dots z_{n-1}\big) \\ \vdots \\ F_{n-1}\big(x, z_0, z_1, \dots z_{n-1}\big) \\ F_n\big(x, z_0, z_1, \dots z_{n-1}\big) \end{pmatrix} = F(x, z).$$

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 8.6. Let $G \subset \mathbb{R} \times \mathbb{R}^n$. A function $f : G \longrightarrow \mathbb{R}^k$ satisfies a *Lipschitz condition* with *Lipschitz constant* $L \geq 0$ if

$$\|f(x, y) - f(x, \widetilde{y})\| \leq L\|y - \widetilde{y}\| \quad \text{for all } x \in \mathbb{R}, y, \widetilde{y} \in \mathbb{R}^n.$$

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 $k = 1$.

LEMMA 8.7. Let $G \subset \mathbb{R} \times \mathbb{R}^n$ be open and for $f : G \longrightarrow \mathbb{R}^n$ exist all partial derivatives. Then f satisfies a local Lipschitz condition on G.

THEOREM 8.8. 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 $x \in I$.

THEOREM 8.9. 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

$(a, c) \in G$ an $\epsilon > 0$ and a solution $\varphi : [a-\epsilon, a+\epsilon] \longrightarrow \mathbb{R}^n$ to the differential equation $y' = f(x, y)$ which satisfies $\varphi(a) = \epsilon$.

We now combine Theorem 8.8 and Theorem 8.9 with Remark 8.5 to obtain

THEOREM 8.10. 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

$$\varphi(a) = c_0,\ \varphi'(a) = c_1,\ \varphi''(a) = c_2,\ \varphi^{(n-1)}(a) = c_{(n-1)}.$$

DEFINITION 8.11. 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 8.12. 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 8.13. 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 8.14. 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 8.15. Let J be an interval and $f : J \to \mathbb{R}$ continuous. For $G = \{(x, y) \in \mathbb{R} \setminus \{0\} \times \mathbb{R} : \frac{y}{x} \in J\}$, we refer to

$$y' = f(\frac{y}{x}), \quad (x, y) \in G$$

as *homogeneous differential equation.*

Theorem 8.16. Let J be an interval, $f : J \to \mathbb{R}$, $G = \{(x, y) \in \mathbb{R} \setminus \{0\} \times \mathbb{R} : \frac{y}{x} \in J\}$, $(x_0, y_0) \in G$ 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 8.17. 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 8.18. 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 $\varphi(x_0) = c$.

Theorem 8.19. 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 $\varphi_1, \varphi_2, \dots, \varphi_k \in L_H$, the following are equivalent

i. $\varphi_1, \varphi_2, \dots, \varphi_k$ are linearly independent functions;

ii. $\varphi_1(x_0), \varphi_2(x_0), \dots, \varphi_k(x_0)$ are linearly independent vectors for some $x_0 \in I$.

iii. $\varphi_1(x), \varphi_2(x), \dots, \varphi_k(x)$ are linearly independent vectors for each $x \in I$.

DEFINITION 8.20. A basis of L_H given in Theorem 8.19 is called *fundamental system of solutions* to the system of differential equation $y' = A(x)y$.

REMARK 8.21. Let $\varphi_1 = (\varphi_{11}, \varphi_{21}, \ldots, \varphi_{n1})^T, \varphi_2 = (\varphi_{12}, \varphi_{22}, \ldots, \varphi_{n2})^T, \ldots, \varphi_n = (\varphi_{1n}, \varphi_{2n}, \ldots, \varphi_{nn})^T$ 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 8.19, we have that for any $x_0 \in I$, the matrix

$$\Phi(x_0) = \begin{pmatrix} \varphi_{11}(x_0) & \varphi_{12}(x_0) & \ldots & \varphi_{1n}(x_0) \\ \varphi_{21}(x_0) & \varphi_{22}(x_0) & \ldots & \varphi_{2n}(x_0) \\ \vdots & \vdots & & \vdots \\ \varphi_{n1}(x_0) & \varphi_{n2}(x_0) & \ldots & \varphi_{nn}(x_0) \end{pmatrix}$$

is invertible, hence, we have $d = \Phi(x_0)^{-1}c$.

THEOREM 8.22. 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 $\psi \in L_I$, we have $L_I = \psi + L_H$.

THEOREM 8.23. Let I be an interval and $A : I \to \mathbb{R}^{n\times n}$, $b : I \to \mathbb{R}^n$ be continuous. Let $\varphi_1 = (\varphi_{11}, \varphi_{21}, \ldots, \varphi_{n1})^T$, $\varphi_2 = (\varphi_{12}, \varphi_{22}, \ldots, \varphi_{n2})^T$, …, $\varphi_n = (\varphi_{1n}, \varphi_{2n}, \ldots, \varphi_{nn})^T$ be a fundamental system of solutions of $y' = A(x)y$. For

$$\Phi = \begin{pmatrix} \varphi_{11} & \varphi_{12} & \cdots & \varphi_{1n} \\ \varphi_{21} & \varphi_{22} & \cdots & \varphi_{2n} \\ \vdots & \vdots & & \vdots \\ \varphi_{n1} & \varphi_{n2} & \cdots & \varphi_{nn} \end{pmatrix},$$

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

$$u(x) = \int_{x_0}^{x} \Phi(t)^{-1} b(t)\, dt + C.$$

Index