Preparatory course for the M. Sc. in Mathematical Physics

Paul Gondolf

paul.gondolf@student.uni-tuebingen.de

Fachbereich Mathematik
Universität Tübingen

logo_uni_tuebingen.png

August 30, 2025

Acknowledgements

This preparatory course follows the course of Prof Dr Stefan Teufel, which he gave as an introductory course in the year 2020. He kindly allowed me to follow the structure as well as the content. I rephrased some parts and did slight modifications, however, kept the main structure and topics. Let it be noted that these notes are for private use only and should not be made public. There might be inconsistencies, typos and errors contained and if you come across those I would be happy and thankful for corrections as well as comments that help improve the script.

Contents

Chapter 1

Topological, metric, and normed spaces

Definition 1.1 (Sequence)**.**
A sequence $(x_n) = (x_1, x_2, x_3, \ldots)$ in a set X is a map:

$$x : \mathbb{N} \to X, \qquad n \mapsto x_n$$

Definition 1.2 (Convergence of a sequence)**.**
A sequence (x_n) in $\mathbb{R}$ converges to $x \in \mathbb{R}$, if

$$\forall \varepsilon > 0 \; \exists N_\varepsilon \in \mathbb{N} \; \forall n \geq N_\varepsilon : \quad |x_n - x| < \varepsilon$$

(if for every $\varepsilon > 0$ it holds 'eventually' that $|x_n - x| < \varepsilon$)
A sequence (x_n) in $\mathbb{R}^k$ converges to $x \in \mathbb{R}^k$, if

$$\forall \varepsilon > 0 \; \exists N_\varepsilon \in \mathbb{N} \; \forall n \geq N_\varepsilon : \quad \|x_n - x\| < \varepsilon$$

A sequence (x_n) in a metric space converges to x in the metric space:

$$\forall \varepsilon > 0 \; \exists N_\varepsilon \in \mathbb{N} \; \forall n \geq N_\varepsilon : \quad d(x_n, x) < \varepsilon,$$

where $d(\cdot, \cdot)$ is called the metric of the space.
A sequence (x_n) in a topological space converges to x in the topological space, if for every neighbourhood U of x eventually $x_n \in U$.

Definition 1.3 (Norm and normed space)**.**
Let V be a vector space either over $\mathbb{R}$ or $\mathbb{C}$. A *norm* $\|\cdot\|$ on V is a map:

$$\|\cdot\| : V \to [0, \infty), \qquad x \mapsto \|x\|$$

with the properties:

1. $\|x\| = 0 \quad \Leftrightarrow \quad x = 0$
2. $\forall x \in V, \; \lambda \in \mathbb{K} : \qquad \|\lambda x\| = |\lambda| \cdot \|x\|$

3. $\forall x, y \in V: \quad \|x + y\| \leq \|x\| + \|y\|$

The pair $(V, \|\cdot\|)$ is called a *normed space*.

Example 1.4. 1. On $V = \mathbb{R}^n$ or $\mathbb{C}^n$ the following maps are norms:

$$\|x\|_2 = \sqrt{|x_1|^2 + |x_2|^2 + \ldots + |x_n|^2} \qquad \text{euclidean norm}$$
$$\|x\|_\infty = \max\{|x_1|, \ldots, |x_n|\} \qquad \text{maximum norm}$$
$$\|x\|_1 = |x_1| + |x_2| + \ldots + |x_n| \qquad \text{1-norm}$$

or, more generally, for $p \in [1, \infty)$, we obtain:

$$\|x\|_p = \Big(\sum_{j=1}^{n} |x_j|^p\Big)^{\frac{1}{p}} \qquad \text{p-norm}$$

2. Let X a set, $(Y, \|\cdot\|_Y)$ a normed space, and

$$V := \{f : X \to Y \mid \sup_{x \in X} \|f(x)\|_Y < \infty\}.$$

Then $\|f\|_\infty = \sup_{x \in X} \|f(x)\|_Y$ is a norm on V.

Exercise 1.1. Show Item 2 of Example 1.4.

Definition 1.5 (Metric and metric space)**.**
Let X be a set. A *metric* d on X is a map:

$$d : X \times X \to [0, \infty)$$

with the following properties:

1. $d(x, y) = 0 \quad \Leftrightarrow \quad x = y$
2. Symmetry: $\forall x, y \in X: \quad d(x, y) = d(y, x)$
3. Triangle inequality: $\forall x, y, z \in X: \quad d(x, z) \leq d(x, y) + d(y, z)$

The pair (X, d) is called a *metric space*

Example 1.6. 1. Let $(V, \|\cdot\|)$ be a normed space. Then $d : V \times V \to [0, \infty)$, $(x, y) \mapsto d(x, y) := \|x - y\|$ defines a metric on V.

2. Let X be as set. The discrete metric on X is defined by:

$$d(x, y) = \begin{cases} 0 & \text{if } x = y \\ 1 & \text{otherwise} \end{cases}$$

3. The euclidean unit sphere $S^2 := \{x \in \mathbb{R}^3 \mid \|x\|_2 = 1\}$ with the metric:

$$d(x, y) := \arccos(\langle x, y|x, y\rangle)$$

is a metric space.

Definition 1.7 (Open sets in a metric space)**.**
Let (X, d) be a metric space.

1. For $x_0 \in X$ and $r > 0$ the set:

$$B_r(x_0) := \{x \in X \mid d(x, x_0) < r\}$$

is called the *open ball*, with the radius r and the centre x_0.

2. A subset of $U \subset X$ is called a *neighbourhood* of $x_0 \in X$, if U contains an open ball around x_0, i.e.

$$\exists r > 0 : \quad B_r(x_0) \subset U.$$

Then x_0 is called an *interior point* of U.

3. A subset $U \subset X$ is called *open*, if it contains only interior points, i.e.

$$\forall x \in U \exists r > 0 : \quad B_r(x) \subset U.$$

Example 1.8. 1. Let (X, d) be a metric space. Then for any $x \in X$ and $r > 0$ the set $B_r(x)$ is open. (Exercise 2: Prove this!)

2. Let X be equipped with the discrete metric. Then any subset $U \subseteq X$ is open: $B_{\frac{1}{2}} = \{x\} \quad \forall x \in X$

Exercise 1.2. Show Item 2 of Example 1.8.

Proposition 1.9. *Let (X, d) be a metric space. Then:*

1. *$\emptyset$ and X are open.*

2. *If $U, V \subset X$ are open, then also $U \cap V$ is open.*

3. *If $U_i \subset X$ is open for all $i \in I$, then also $\bigcup_{i \in I} U_i$ is open.*

Exercise 1.3. Proof Proposition 1.9.

Item 2 implies that intersections of finitely many open sets are open. However, this does not hold for infinite intersections: Let $U_n = (-\frac{1}{n}, \frac{1}{n}) \subset \mathbb{R}$, $n \in \mathbb{N}$. Then $\bigcap_{n \in \mathbb{N}} U_n = \{0\}$ is not open.

Definition 1.10 (Closed set)**.**
A subset $A \subset X$ of a metric space is *closed*, if its complement is open, i.e. $A^C = \{x \in X \mid x \notin A\}$ is open.

Example 1.11.
1. Let $X = \mathbb{R}$ and $a, b \in \mathbb{R}$ with $a < b$. Then $[a, b]$, $[a, \infty)$ are closed, but $[a, b)$ is neither open nor closed.
2. For any metric space (X, d) the sets $\emptyset$ and X are open and closed.

Definition 1.12 (Topology)**.**
Let X be a set. A *topology* $\mathcal{T}$ on X is a collection $\mathcal{T} \subset \mathcal{P}(X)$ of subsets of X with the following properties:

1. $\emptyset, X \in \mathcal{T}$
2. $U, V \in \mathcal{T} \quad \Rightarrow \quad U \cap V \in \mathcal{T}$
3. $U_i \in \mathcal{T}$ for all $i \in I \quad \Rightarrow \quad \bigcup_{i \in I} U_i \in \mathcal{T}$

The sets $U \in \mathcal{T}$ are called the *open sets* and $(X, \mathcal{T})$ is called a *topological space*. $A \subset X$ is *closed*, if $A^C \in \mathcal{T}$. $U \subset X$ is called a *neighbourhood* of a $x_0 \in X$ (and x_0 is an *interior point* of U), if:

$$\exists O \in \mathcal{T} \quad \text{with} \quad x_0 \in O \subset U.$$

Example 1.13. According to Proposition 1.9, the open sets in a metric space form a topology.

Definition 1.14 (Relative topology)**.**
$(X, \mathcal{T})$ a topological space and $Y \subset X$ a subset of X. Then $\mathcal{T}|_Y = \{O \cap Y \mid O \subset \mathcal{T}\}$ is a topology on Y, the *subspace* or *relative topology*. The elements $U \subset \mathcal{T}|_Y$ are called *relative open sets*.

Example 1.15.
1. $X = \mathbb{R}$, $Y = [0, 1]$. Then $Y \in \mathcal{T}|_Y$, i.e. Y is relatively open in itself. Also $[0, \frac{1}{2}) \subset Y$ is relatively open, since $[0, \frac{1}{2}) = (-\frac{1}{2}, \frac{1}{2}) \cap Y$
2. (X, d) a metric space and $Y \subset X$ a subset. Then $(Y, d|_Y)$ is a metric space.

$$d|_Y : Y \times Y \to [0, \infty], \quad (y_1, y_2) \mapsto d|_Y(y_1, y_2) = d(y_1, y_2)$$

3. $(V, \|\cdot\|)$ normed space and $U \subset V$ a vector subspace. Then $(U, \|\cdot\||_U)$ is a normed space.

Definition 1.16 (Interior, closure, boundary)**.**
Let $(X, \mathcal{T})$ be a topological space and $Y \subset X$.

1. The set $\mathring{Y} = \bigcup_{\substack{U \in \mathcal{T} \\ U \subset Y}} U$ is called the *interior* of Y.

2. The set $\overline{Y} = \bigcap_{\substack{U \in \mathcal{T} \\ U \subset Y^C}} U^C$ is called the *closure* of Y.

3. The set $\partial Y = \overline{Y} \backslash \overset{\circ}{Y}$ is called the *boundary* of Y.

Proposition 1.17. *1.* $\overset{\circ}{Y} \subset Y \subset \overline{Y}$.

2. $\overset{\circ}{Y}$ *is the largest open set contained in* Y*.*

3. Y *is open* $\Leftrightarrow Y = \overset{\circ}{Y}$*.*

4. $\overline{Y}$ *is the smallest closed set containing* Y*.*

5. Y *is closed* $\Leftrightarrow Y = \overline{Y}$*.*

6. $(\overset{\circ}{Y})^C = \overline{(Y^C)}$ *and* $\left(Y^C\right)^\circ = \left(\overline{Y}\right)^C$*.*[1]

Proposition 1.18. *Let* $(X, \mathcal{T})$ *be a topological space and* $Y \subset X$*. Then:*

1. $\overset{\circ}{Y}$ *is the set of interior points of* Y*.*

2. $x \in \partial Y \quad \Leftrightarrow \quad$ *for any neighbourhood* U *of* x $U \cap Y \neq \emptyset$ *and* $U \cap Y^C \neq \emptyset$*.*

3. $\overset{\circ}{Y} = Y \backslash \partial Y$ *and* $\overline{Y} = Y \cup \partial Y$

Example 1.19. 1. For $Y = [a, b) \subset \mathbb{R}$ we have $\overset{\circ}{Y} = (a, b)$, $\overline{Y} = [a, b]$, $\partial Y = \{a, b\}$.

2. For $\mathbb{Q} \subset \mathbb{R}$ we have $\overset{\circ}{\mathbb{Q}} = \emptyset$, $\overline{\mathbb{Q}} = \mathbb{R}$, $\partial \mathbb{Q} = \mathbb{R}$.

Definition 1.20 (Convergence in a topological space)**.**
Let X be a topological space. A sequence (x_n) in X *converges* to $a \in X$ and we write:

$$\lim_{n \to \infty} x_n = a,$$

if for any neighbourhood U of the point a there exists $N \in \mathbb{N}$ such that $x_n \in U$ for all $n \geq N$.

Remark 1.21. In general convergence points are not unique: On any set X with the indiscrete topology

$$\mathcal{T} = \{\emptyset, X\}$$

any sequence (x_n) in X converges to every point in X!

[1] This can be proven using de Morgan's laws: For a family of sets $(A_i)_{i \in I}$ it holds

(a) $\left(\bigcap_{i \in I} A_i\right)^C = \bigcup_{i \in I} A_i^C$,

(b) $\left(\bigcup_{i \in I} A_i\right)^C = \bigcap_{i \in I} A_i^C$.

Definition 1.22 (Hausdorff spaces)**.**
A topological space $(X, \mathcal{T})$ is *Hausdorff*, if:

$$\forall x, y \in X, x \neq y, \exists U, V \in \mathcal{T}: \quad x \in U, y \in V, \quad U \cap Y = \emptyset\,.$$

Remark 1.23.
1. In Hausdorff spaces, sequences have at most one limit.
2. Metric spaces are Hausdorff: For $x, y \in X$, $x \neq y$, we have $d(x, y) = r > 0$ and $B_{\frac{r}{2}}(x) \cap B_{\frac{r}{2}}(y) = \emptyset$.

Definition 1.24 (Cluster point)**.**
A point $a \in X$ is called a *cluster point* of a sequence (x_n) if any neighbourhood U of a contains infinitely many elements of (x_n).

Proposition 1.25. *Let X be a metric space and $A \subset X$. Then*

$$a \in \overline{A} \quad \Leftrightarrow \quad \exists (x_n) \text{ in } A\text{: } \lim_{n\to\infty} x_n = a$$

Remark 1.26. The implication $\Leftarrow$ of Proposition 1.25 holds also in topological spaces.

Definition 1.27 (Cauchy sequence)**.**
A sequence (x_n) in a metric space X is called *Cauchy sequence*, if

$$\forall \varepsilon > 0 \exists N \in \mathbb{N} \forall n, m \geq N: \quad d(x_n, x_m) < \varepsilon$$

Proposition 1.28. *Every convergent sequence in a metric space is also a Cauchy sequence.*

Exercise 1.4. Proof Proposition 1.28.

Definition 1.29 (Complete metric space, Banach space)**.**
1. A metric space X is called *complete* if every Cauchy sequence in X converges.
2. A complete normed space (the norm induces the metric) is called a *Banach space*.

Example 1.30.
1. $\mathbb{R}, \mathbb{C}, \mathbb{R}^n, \mathbb{C}^n$ (with e.g. the Euclidean norm) are Banach spaces.
2. $(\mathbb{Q}, |\cdot|)$ is not complete, as there exists a Cauchy sequence $(x_n) \subset \mathbb{Q}$ with $\lim_{n\to\infty} x_n = \sqrt{2}$ but $\sqrt{2} \notin \mathbb{Q}$.

Chapter 2

Continuity, compact sets, connected sets

Definition 2.1 (Continuity and sequential continuity)**.**
Let X, Y be topological spaces, $f : X \to Y$ a map, and $a \in X$

1. We say that f is *sequentially continuous* at a, if for a sequence (x_n), $\lim_{n\to\infty} x_n = a$ implies that
$$\lim_{n\to\infty} f(x_n) = f(a).$$

2. We say that f is *continuous* at a, if
$$\forall U \in \mathcal{U}(f(a))\, \exists V \in \mathcal{U}(a) : \quad f(V) \subset U.$$ [1]

If a function is (sequentially) continuous at all points $a \in X$, then we say that f is *(sequentially) continuous on X*.

Proposition 2.2. *If $f : X \to Y$ is continuous at $a \in X$, then f is also sequentially continuous at a.*

Exercise 2.1. Proof Proposition 2.2.

Proposition 2.3 (ε-δ-continuity in metric spaces)**.** *A function $f : X \to Y$ between metric spaces X, Y is continuous at $a \in X$, if and only if*
$$\forall \varepsilon > 0 \exists \delta > 0 : \quad f(B_\delta(a)) \subset B_\varepsilon(f(a))$$

Exercise 2.2. Proof Proposition 2.3.

Proposition 2.4. *A function $f : X \to Y$ between metric spaces X, Y is continuous at $a \in X$, if and only if it is sequentially continuous at a.*

Proof. $\Rightarrow$ Proposition 2.2

[1] $\mathcal{U}(x)$ is the set of all neighbourhoods of the point x.

$\Leftarrow$ (by contraposition $A \Rightarrow B \;\Leftrightarrow\; \neg B \Rightarrow \neg A$)
Assume that f is not continuous at a, i.e.

$$\exists \varepsilon > 0 \forall \delta > 0 : \quad f(B_\delta(a)) \not\subset B_\varepsilon(f(a)).$$

For $\delta = \frac{1}{n}$ choose $x_n \in B_\delta(a) \backslash f^{-1}(B_\varepsilon(f(a))) \neq \emptyset$. Then $\lim_{n\to\infty} x_n = a$, but $f(x_n) \notin B_\varepsilon(f(a)) \forall n \Rightarrow f$ is not sequentially continuous.

□

Theorem 2.5. *Let X, Y be topological spaces. A map $f : X \to Y$ is continuous (on X), if the preimage $f^{-1}(O) \subset X$ of any open set $O \subset Y$ is open.*

Example 2.6.

1. In a metric space (X, d) the distance function to a point $b \in X$,

$$d_b : X \to [0, \infty), \quad x \mapsto d_b(x) := d(x, b)$$

is continuous.[2]

2. In a normed space $(V, \|\cdot\|)$ the norm:

$$\|\cdot\| : V \to [0, \infty),$$

addition:

$$+ : V \times V \to V, \quad (x, y) \mapsto x + y,$$

and multiplication by scalars:

$$\cdot : \mathbb{K} \times V \to V, \quad (\lambda, v) \mapsto \lambda \cdot v$$

are all continuous.

3. The composition of continuous functions is continuous. If $f : X \to Y$ and $g : Y \to Z$ are continuous then also $g \circ f : X \to Z$ is continuous.

Proof. Let $O \subset Z$ be open $\overset{g \text{ cont.}}{\Rightarrow} g^{-1}(O) \subset Y$ is open $\overset{f \text{ cont.}}{\Rightarrow} f^{-1}\left(g^{-1}(O)\right) = (g \circ f)^{-1}(O) \subset X$ is open. $\Rightarrow g \circ f$ cont. □

Remark 2.7.

1. Let (X, d_X) and (Y, d_Y) be metric spaces. Then a metric on $X \times Y$ is for example

$$d((x_1, y_1), (x_2, y_2)) := (d_x(x_1, x_2)^p + d_y(y_1, y_2)^p)^{1/p} \quad 1 \le p < \infty$$

2. Let $(X, \mathcal{T}_X)$ and $(Y, \mathcal{T}_Y)$ topological space. Then the (product) topology on $X \times Y$ is generated by

$$\{O_1 \times O_2 \;:\; O_1 \in \mathcal{T}_X, O_2 \in \mathcal{T}_Y\}$$

also called *bose topology*.

[2] Also $d : X \times X \to [0, \infty)$ is continuous using a suitable metric on $X \times X$. For the definition of this metric, see Remark 2.7.

3. Let $(X_i, \mathcal{T}_i)$, $i \in I$, be topological spaces. Then the product topology on $\prod_{i \in I} X_i$ is generated by

$$\Big\{ \prod_{i \in I} O_i \ : \ O_i \in \mathcal{T}_i \text{ and } O_i \neq X_i \text{ only for finitely many } i \in I \Big\}.$$

Definition 2.8 (Lipschitz continuity)**.**
Let X, Y be metric spaces. A function $f : X \to Y$ is called *Lipschitz-continuous*, if there exists $L \geq 0$ such that

$$\forall x_1, x_2 \in X : \quad d_Y(f(x_1), f(x_2)) \leq L \cdot d_X(x_1, x_2).$$

Then L is called a *Lipschitz-constant* for f. If f has a Lipschitz-constant $L < 1$, then f is called *contraction*.

Example 2.9.

1. $f(x) = ax + b$ is Lipschitz continuous with $L = a$.
2. $f \in C^1(\mathbb{R})$ then $L = \sup_{x \in \mathbb{R}} |f'(x)|$.
3. $f(x) = x^2$ is continuous but not Lipschitz continuous.
4. $f(x) = \sqrt{|x|}$ is continuous but not Lipschitz continuous, as its derivative around 0 diverges.

Definition 2.10 (Homeomorphic functions, isometries and isometric isomorphisms)**.**

1. Two topological spaces X, Y are *homeomorphic* if there exists a bicontinuous bijection
$$f : X \to Y \quad \text{a } \textit{homeomorphism}$$
2. A map $f : X \to Y$ between metric spaces is an *isometry*, if
$$\forall x_1, x_2 \in X : \quad d_Y(f(x_1), f(x_2)) = d_x(x_1, x_2)\,.$$
X and Y are *isometric*, if there exists a bijective isometry $f : X \to Y$.
3. Two normed spaces V and W are *isometrically isomorphic*, if there exists a linear bijection (isomorphism) $A : V \to W$ such that
$$\forall v \in V : \quad \|Av\|_W = \|v\|_V\,.$$

Example 2.11. The isometries of Euclidean space $(\mathbb{R}^n, d_2)$ are translations, rotations and reflections and compositions thereof (euclidean group).

Definition 2.12 (Pointwise and uniform convergence)**.**
Let X be a set, Y a metric space and

$$f_n : X \to Y,\ n \in \mathbb{N} \quad \text{and} \quad f : X \to Y$$

both functions.

1. We say that f_n *converges pointwise* to f, if
$$\forall x \in X : \quad \lim_{n\to\infty} d_Y(f_n(x), f(x)) = 0. \quad \Leftrightarrow \quad \lim_{n\to\infty} f_n(x) = f(x)$$

2. We say that f_n *converges uniformly* to f, if
$$\lim_{n\to\infty} \sup_{x\in X} d_Y(f_n(x), f(x)) = 0$$
If $(Y, \|\cdot\|)$ is a normed space, then $f_n \to f$ uniformly, if and only if
$$\lim_{n\to\infty} \|f_n - f\|_\infty = 0$$

Example 2.13. $f_n : [0,1] \to [0,1]$, $x \mapsto f_n(x) = x^n$, then pointwise

$$f_n(x) \overset{n\to\infty}{\longrightarrow} f(x) = \begin{cases} 0 & \text{for } x < 1 \\ 1 & \text{for } x = 1 \end{cases}.$$

However, (f_n) does not converge uniformly to f since $\sup_{x\in[0,1]} |f_n(x) - f(x)| = 1$

Definition 2.14 (Uniform limits of continuous functions are continuous)**.**
Let $(X, \mathcal{T})$ a topological and (Y, d) a metric space. Let $f_n : X \to Y$ be a sequence of continuous functions and let $f_n \to f$ uniformly. Then f is continuous.

Corollary 2.15. *Let X be a topological space, $(Y, \|\cdot\|_Y)$ a complete normed space and $C_b(X, Y)$ the space of continuous bounded functions equipped with the $\|\cdot\|_\infty$-norm.*

$$\|f\|_\infty = \sup_{x\in X} \|f(x)\|_Y$$

Then the normed space $(C_b(X, Y), \|\cdot\|_\infty)$ is complete.

Definition 2.16 (Open cover and finite subcover)**.**
Let $(X, \mathcal{T})$ be a topological space and $Y \subset X$. A family $(U_i)_{i\in I}$ of open sets, $U_i \in \mathcal{T}\ \forall i \in I$, is called an *open cover* of Y, if

$$Y \subset \bigcup_{i\in I} U_i$$

A set $K \subset X$ is called *compact*, if any open cover $(U_i)_{i\in I}$ of K admits a finite subcover, i.e. there exists $i_1, \ldots, i_n \in I$ such that:

$$K \subset \bigcup_{i=i_1,\ldots,i_n} U_i$$

Example 2.17. 1. Every finite subset $K = \{x_1, \dots, x_n\}$ of a topological space is compact.

2. $(0,1] \subset \mathbb{R}$ is not a compact set. The open cover $(0,1] \subset \bigcup_{n=2}^{\infty} (\frac{1}{n}, 2)$ admits no finite subcover.

Theorem 2.18 (Bolzano-Weierstraß)**.** *Let $K \subset X$ be compact. Then any sequence in K has a cluster point in K.*

Exercise 2.3. Proof Theorem 2.18.

Remark 2.19. In metric spaces also the converse is true.

Proposition 2.20. *Let $f : X \to Y$ be a continuous function and $K \subset X$ a compact set. Then also $f(K) \subset Y$ is compact.*

Proposition 2.21. *1. Let X be a topological space and $K \subset X$ compact. Then any close subset $A \subset K$ is also compact.*

2. If X is a Hausdorff space and K compact, then K is closed.

Definition 2.22 (Bounded sets and the diameter of a set)**.**
Let X be a metric space.

1. A subset $B \subset X$ is *bounded*, if

$$\exists C \in \mathbb{R} \forall x, y \in B : \quad d(x,y) \leq C$$

2. The *diameter* of the set $Y \subset X$ is

$$\operatorname{diam}(Y) = \sup\{d(x,y) \mid x, y \in Y\} \in [0, \infty) \cup \{\infty\}$$

Theorem 2.23.
Let X be a metric space and $K \subset X$ compact. The K is bounded and closed.

Theorem 2.24 (Heine-Borel)**.** *A subset K of a finite-dimensional normed space is compact if it is bounded and closed.*

Theorem 2.25 (Weierstraß)**.** *Let $f : K \to \mathbb{R}$ be a continuous function and K compact. Then f is bounded ($f(K) \subset \mathbb{R}$ is bounded) and attains its maximum and its minimum.*

Definition 2.26 (Equicontinuity)**.**
Let X, Y be metric spaces and $A \subset C(X,Y)$. Then the set A is called *equicontinuous* at $x \in X$, if

$$\forall \varepsilon > 0 \exists \delta > 0 \, \forall f \in A : \quad f(B_\delta(x)) \subset B_\varepsilon(f(x)) \, .$$

Theorem 2.27 (Arzela-Ascoli)**.** *Let X be a compact metric space and consider $C(X,\mathcal{C})$ equipped with the $\|\cdot\|_\infty$-norm. A subset $K \subset C(X,\mathbb{C})$ is compact, if and only if it is closed, bounded pointwise (i.e. $\forall x \in X$:*

$$\sup_{f\in K} |f(x)| < \infty)$$

and equicontinuous.

Definition 2.28 (Connected, disconnected and path connected spaces)**.**
Let X be a topological space. If X is the union of two disjoint, open, non-empty sets, then X is *disconnected*, otherwise *connected*.
X is *path-connected*, if any two points $x_0, x_1 \in X$ can be connected by a continuous path, i.e. there exists

$$\gamma : [0,1] \to X$$

continuous, with $\gamma(0) = x_0$ and $\gamma(1) = x_1$.

Proposition 2.29. *If X is path-connected then X is connected.*

Proposition 2.30. *Let O be an open subset of a normed space. Then O is connected, if and only if it is path connected.*

Proposition 2.31. *Let $f : X \to Y$ be continuous and $A \subset X$ (path) connected. Then also $f(A) \subset Y$ is (path) connected.*

Definition 2.32 (Domain)**.**
A non-empty, open, connected subset $D \subset X$ of a topological space X is called *domain*.

Definition 2.33 (Bounded functions)**.**
A function $f : X \to Y$ with X a set and (Y,d) a metric space, is called bounded, if and only if $f(X) \subset Y$ is bounded.

Definition 2.34 (Bounded linear maps and their norms)**.**
A linear map $A : V \to W$ between normed spaces is called bounded, if $A(B_1(0))$ is *bounded*, i.e.

$$\exists C \in \mathbb{R} \forall x \in V : \quad \|Ax\|_W \leq C\|x\|_V.$$

The smallest such constant C is called the *operator norm* of A, i.e.

$$\|A\|_{op} := \sup\{\|Ax\|_W \mid x \in \overline{B_1(0)}\}$$

The space of bounded linear maps $V \to W$ is denoted by

$$\mathcal{L}(V,W) \text{ or } \mathcal{B}(V,W)$$

and $\|\cdot\|_{op}$ is a norm on $\mathcal{L}(V,W)$.

Remark 2.35.

1. If $A \in \mathcal{L}(V,W)$ we have for all $x \in V$

$$\|Ax\|_W \leq \|A\|_{op} \cdot \|x\|_V$$

2. If $(W, \|\cdot\|_W)$ is a Banach space, then $(\mathcal{L}(V,W), \|\cdot\|_{op})$ is also complete.
3. If $\dim V < \infty$, then all linear maps $V \to W$ are bounded.

Chapter 3

Differential calculus

Remark 3.1. Recall that for a function $f : \mathbb{R} \supset D \to \mathbb{R}$

$$f'(x_0) = \lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0} = \lim_{h \to 0} \frac{f(x_0 + h) - f(x_0)}{h}$$

This can also be formulated with sequences:

$$\forall (x_n) \subset D \backslash \{x_0\}, \lim_{n \to \infty} x_n = x_0 : \quad \lim_{n \to \infty} \frac{f(x_n) - f(x_0)}{x_n - x_0}$$

Definition 3.2 (Partial derivative)**.**
Let $n \in \mathbb{N}$, $D \subset \mathbb{R}^n$ open, $(W, \|\cdot\|)$ a normed space. For $x \in D$ and $j \in \{1, \dots, n\}$ a function $f : D \to W$ is called *partially differentiable* in the j^{th} coordinate direction at x, if the limit:

$$\lim_{h \to 0} \frac{f(x + he_j) - f(x)}{h} = \lim_{h \to 0} \frac{f(x_1, x_2, \dots, x_j + h, \dots, x_n) - f(x_1, \dots, x_n)}{h}$$

exists. One writes:

$$\frac{\partial f}{\partial x_j}(x) = \partial_j f(x) := \lim_{h \to 0} \frac{f(x + he_j) - f(x)}{h}$$

and calls the vector $\partial_j f(x) \in W$ the j^{th} partial derivative at x.
If f is partially differentiable and the partial derivatives $\partial_j f : D \to W$ are continuous functions, then f is called continuously partially differentiable. The vector space of the *continuous partially differentiable* functions on $D \subset \mathbb{R}^n$ is denoted by $C^1(D, W)$. The *gradient* of f at x is

$$\nabla f(x) := (\partial_1 f(x), \dots, \partial_n f(x)) \in W^n$$

Definition 3.3 (Vector field)**.**
A map $f : \mathbb{R}^n \supset D \to \mathbb{R}^n$ is called a *vector field*.

Example 3.4. The gradient $\nabla f : D \to \mathbb{R}^n$ of a function $f : \mathbb{R}^n \supset D \to \mathbb{R}$ defines a vector field.

Definition 3.5 (Higher order partial derivatives)**.** A function $f : \mathbb{R}^n \supset D \to W$ (with W being a $\mathbb{K}$ vector space) is called *r-times continuously partially differentiable*, if for all $j = (j_1, \dots, j_r)$, $j_i \in \{1, \dots, n\}$

1. f is c.p.d.
2. $\partial_{j_1} f$ is c.p.d.
3. $\partial_{j_2}\partial_{j_1}$ is c.p.d.

   $\vdots$

4. $\partial_{j_r} \dots \partial_{j_1} f$ is continuous

The $\mathbb{K}$ vector space of r-times c.p.d. functions is denoted by $C^r(D, W)$.

Definition 3.6. Let $D \subset \mathbb{R}^n$, $g \in C^1(D, \mathbb{R}^n)$, $f \in C^2(D, \mathbb{R})$. Then:

$$\operatorname{div} g : D \to \mathbb{R}, \qquad x \mapsto \operatorname{div} g(x) = \sum_{j=1}^{n} \frac{\partial g_j}{\partial x_j}(x)$$

is called the *divergence* of g,

$$\operatorname{curl} g : D \to \mathbb{R}^n, \qquad x \mapsto \operatorname{curl} g(x) = \begin{pmatrix} \partial_2 g_3(x) - \partial_3 g_2(x) \\ \partial_3 g_1(x) - \partial_1 g_3(x) \\ \partial_1 g_2(x) - \partial_2 g_1(x) \end{pmatrix}$$

for $n = 3$ is called the *curl* of g, and:

$$\Delta f : D \to \mathbb{R}, \qquad x \mapsto \Delta f(x) = \operatorname{div}(\nabla f)(x) = \sum_{j=1}^{n} \frac{\partial^2 f}{\partial x_j^2}(x)$$

is called the *Laplace* of f.

Example 3.7. Let $g : \mathbb{R}^n \to \mathbb{R}^n$, $\quad x \mapsto g(x) = x$ $(g = id)$ and $f : \mathbb{R}^n \backslash \{0\} \to \mathbb{R}$, $x \mapsto \|x\|_2$. Compute $\operatorname{div} g$ and Δf.

Theorem 3.8 (Schwarz)**.**
Let $f \in C^2(D, W)$, $D \subset \mathbb{R}^n$. Then $\forall x \in D$, $\quad j, i \in \{1, \dots, n\}$

$$\partial_j \partial_i f(x) = \partial_i \partial_j f(x)\,.$$

Corollary 3.9. *Let $D \subset \mathbb{R}^3$, $f \in C^2(D, \mathbb{R})$ and $g \in C^2(D, \mathbb{R}^3)$. Then* $\operatorname{curl}(\nabla f) = 0$ *and* $\operatorname{div}(\operatorname{curl} g) = 0$.

Definition 3.10 (Directional derivative)**.** Let V be a real, normed space, W a normed space, $D \subset V$ open and $f : D \to W$. Then the *directional derivative* of f at $x \in D$ in the direction $v \in V$ is defined

$$\partial_v f(x) = \lim_{h\to 0} \frac{f(x+hv) - f(x)}{h} = \frac{d}{dh} f(x+hv)\big|_{h=0},$$

if the limit exists.

Example 3.11. For $f : \mathbb{R}^n \to W$ we have $\partial_{e_j} f(x) = \partial_j f(x)$.

3.1 The derivative as linear approximation

For $f : \mathbb{R} \to W$ differentiability at $x_0 \in \mathbb{R}$ means

$$\lim_{x\to x_0} \left(\frac{f(x) - f(x_0)}{x - x_0} - f'(x_0) \right) = \lim_{x\to x_0} \frac{\varphi(x, x_0)}{x - x_0}$$

where $\varphi(x, x_0) = f(x) - f(x_0) - f'(x_0)(x - x_0)$ or, after reshuffling

$$f(x) = f(x_0) + f'(x_0)(x - x_0) + \varphi(x, x_0)$$

where

$$\varphi(x, x_0) = o(|x - x_0|) \quad :\Leftrightarrow \quad \lim_{x\to x_0} \frac{|\varphi(x, x_0)|}{|x - x_0|} = 0$$

The map $\mathbb{R} \to W$, $\quad x \mapsto f'(x_0) \cdot x$ is $\mathbb{R}$-linear and the map $\mathbb{R} \to W$, $\quad x \mapsto f(x_0) + f'(x_0)(x - x_0)$ is affine-$\mathbb{R}$-linear. Hence, we think of $f(x_0) + f'(x_0)(x - x_0)$ as the (affine) linear approximation to f near x_0.

Definition 3.12 (Total derivative)**.**
Let V be finite-dimensional real vector space, W a normed space, $G \subset V$ open and $f : G \to W$. We call f differentiable at $x_0 \in G$, if there exists an $\mathbb{R}$-linear-map $A : V \to W$ such that:[1]

$$\lim_{x\to x_0} \frac{f(x) - f(x_0) - A(x - x_0)}{\|x - x_0\|_V} = 0$$

Then A is uniquely determined by the above equation, is denoted by $Df|_{x_0}$, and called the total derivative or the differential of f at x_0.
If $f : G \to W$ is differentiable at all $x \in G$, then f is called *differentiable on G* and:

$$Df : G \to \mathcal{L}(V, W), \qquad x \mapsto Df|_x$$

is a function on G taking values in the linear maps from V to W.

[1] Here $\|\cdot\|_V$ is any norm on V, since V is a finite-dimensional vector space

Exercise 3.1. Show that the $\mathbb{R}$-linear map A in Definition 3.12 is unique.

Remark 3.13. $f : G \to W$ is differentiable at $x_0 \Rightarrow f(x) = f(x_0) + Df|_{x_0}(x - x_0) + o(\|x - x_0\|_V)$

Example 3.14. Let $L : V \to W$ be a $\mathbb{R}$-linear map. Then

$$L(x) = L(x_0 + x - x_0) = L(x_0) + L(x - x_0) = L(x_0) + DL|_{x_0}(x - x_0)$$

and hence $DL|_{x_0} = L$.

Theorem 3.15 (Jacobi matrix)**.**
Let $G \subset \mathbb{R}^n$ and let $f : G \to \mathbb{K}^m$ be differentiable at a point $x_0 \in G$. Then

$$(Df|_{x_0})_{ij} = \frac{\partial f_i}{\partial x_j}(x_0)$$

or, more explicitly

$$Df|_{x_0} = \begin{pmatrix} \partial_1 f_1(x_0) & \dots & \partial_n f_1(x_0) \\ \vdots & & \vdots \\ \partial_1 f_m(x_0) & \dots & \partial_n f_m(x_0) \end{pmatrix} = \begin{pmatrix} \nabla f_1(x_0) \\ \vdots \\ \nabla f_m(x_0) \end{pmatrix}$$

the Jacobian matrix.

Theorem 3.16. *Let $G \subset \mathbb{R}^n$ open and $f \in C^1(G, \mathbb{K}^m)$*[2]*. Then f is differentiable.*

$$\begin{array}{ccccc} \text{cont. part. diff.} & \Rightarrow & \text{differentiable} & \Rightarrow & \text{part. diff.} \\ & & \Downarrow & & \\ & & \text{continuous} & & \end{array}$$

None of the implications holds in the reversed direction! But cont. part. diff. $\Leftrightarrow$ differentiable with continuous derivative.

Theorem 3.17 (Chain rule)**.** *Let U, V be finite dimensional real vector spaces, W a normed space, $G \subset U$, $H \subset V$ open and $g : G \to V$, $f : H \to W$ maps with $g(G) \subset H$, i.e.*

$$U \supset G \xrightarrow{g} H \subset V \xrightarrow{f} W$$

If g is differentiable at $x \in G$ and f is differentiable at $g(x) \in H$, then $f \circ g : G \to W$ is differentiable at x and

$$D(f \circ g)|_x = Df|_{g(x)} \cdot Dg|_x$$

[2]continuously partially differentiable

Corollary 3.18. *For a function $f \in C^1(G, W)$, $G \subset V$, $x_0 \in G$, and $v \in V$ we have*

$$\partial_v f(x_0) = Df|_{x_0} v$$

Exercise 3.2. Proof Corollary 3.18.

We now want to state the Taylor theorem. To do so we, however, have to understand higher-order differentials. For $f : G \to W$, with $G \subset V$ an open subset, the differential Df is a map

$$Df : G \to \mathcal{L}(V, W).$$

Thus the second derivative $D(Df)$ is a map

$$D(Df) : G \to \mathcal{L}(V, \mathcal{L}(V, W)) \cong \mathcal{L}_2(V \times V, W)$$

($=$ bilinear maps $V \times V \to W$) and the k^{th} derivative:

$$D^k f : G \to \mathcal{L}_k(\underbrace{V \times \ldots \times V}_{k\text{-times}}, W)$$

Theorem 3.19 (Taylor)**.** *Let $G \subset V$ open, $x_0 \in G$, and $\delta > 0$ such that $B_\delta(x_0) \subset G$. Then for any function $f \in C^k(G, W)$ and $x \in B_\delta(x_0)$*

$$\begin{aligned} f(x) = f(x_0) + Df|_{x_0}(x - x_0) + \frac{1}{2} D^2 f|_{x_0}(x - x_0, x - x_0) \\ + \frac{1}{k!} D^k f|_{x_0}(x - x_0, \ldots, x - x_0) + o(\|x - x_0\|_V^k) \end{aligned}$$

Definition 3.20 ((Strict) local maximum)**.** Let X be a topological space and $f : X \to \mathbb{R}$. A point $x_0 \in X$ is called a (strict) local maximum of f, if

$$\exists U \subset \mathcal{U}(x_0) : \forall x \in U \backslash \{x_0\} : \quad f(x) \underset{<}{\leq} f(x_0)$$

The definition for a local minimum is analogous.

Theorem 3.21. *Let $G \subset V$ and $f \in C^1(G, \mathbb{R})$ have a local extremum at $x_0 \in G$. Then $Df|_{x_0} = 0$.*

Theorem 3.22. *Let $G \subset V$ and $f \in C^2(G, \mathbb{R})$ and $x_0 \in G$ such that $Df|_{x_0} = 0$.*

1. *If $D^2 f|_{x_0}(h, h) > 0$ $\forall h \in V \backslash \{0\}$, then f has a strict local minimum at x_0*
2. *$D^2 f|_{x_0}(h, h) < 0$ $\forall h \in V \backslash \{0\}$, then f has a stric local maximum at x_0*
3. *If $D^2 f|_{x_0}$ is indefinite, then f has no local extremum at x_0*

Theorem 3.23 (Mean value theorem in one dimension)**.** *For a function $f : [a, b] \to \mathbb{R}$ continuous and differentiable on (a, b). Then $\exists x_0 \in (a, b)$:*

$$f'(x_0) = \frac{f(b) - f(a)}{b - a}$$

Theorem 3.24. *Let $G \subset \mathbb{R}^n$ and $f : G \to \mathbb{K}^n$ be continuously differentiable. Let $\gamma : [a, b] \to G$ continuously differentiable. Then*

$$f(\gamma(b)) - f(\gamma(a)) = \int_a^b \underbrace{Df|_{\gamma(t)}}_{\mathbb{K}^n \leftarrow \mathbb{R}^n} \underbrace{\gamma'(t)}_{\in \mathbb{R}^n} \, dt$$

Theorem 3.25 (Mean value theorem)**.** *$G \subset \mathbb{R}^n$, $f \in C^1(G, \mathbb{K}^m)$. Let $x \in \mathbb{R}^n$ such that $\{x + th \mid t \in [0, 1]\} \subset G$. Then*

$$f(x + h) - f(x) = \left(\int_0^1 Df|_{x+th} \, dt \right) \cdot h \, .$$

Corollary 3.26. *The setup is as in the Theorem 3.25. Then*

$$\|f(x) - f(y)\| \le \underbrace{\left\| \int_0^1 Df|_{x+th} \, dt \right\|_{op}}_{\sup\limits_{z \in \overline{xy}} \|Df|_z\|} \cdot \|x - y\|$$

For $f : \mathbb{R}^n \to \mathbb{R}$ we obtain again

$$f(y) - f(x) = Df|_z \cdot (y - x)$$

Definition 3.27 (Equivalence of norms)**.**
Two norms $\|\cdot\|_a$ and $\|\cdot\|_b$ on a vector space V are equivalent, if $\exists c, C > 0 \forall x \in V$:

$$c\|x\|_a \le \|x\|_b \le C\|x\|_a \, .$$

Theorem 3.28. *On finite dimensional vector spaces, all norms are equivalent.*

Theorem 3.29. *All finite dimensional normed spaces are complete (Banach spaces).*

Definition 3.30 (Frechet derivative)**.** Let X and Y be Banach spaces and $G \subset X$ open. A map $f : G \to Y$ is differentiable at $x \in G$, if there exists a *continuous* linear map $A : X \to Y$ such that

$$f(x + h) = f(x) + Ah + o(\|h\|_X)$$

for h in a neighbourhood of $0 \in X$. The notation $A = Df|_x$ remains.

Example 3.31. $X = C^2([0, T], \mathbb{R}^n)$. An element $x \in X$ is a map $x : [0, T] \to \mathbb{R}^n$. We can equip this space with a norm

$$\|x\|_X = \|x\|_\infty + \|\dot{x}\|_\infty + \|\ddot{x}\|_\infty$$

an turn it into a Banach space with respect to that norm. The action is given by

$$S : X \to \mathbb{R} \quad x \mapsto S(x) = \int_0^T L(x(t), \dot{x}(t))\, dt\,,$$

where $L : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}, \quad (q, v) \mapsto L(q, v) \in C^2$ (e.g. $L(q, v) = \frac{1}{2}m\|v\|^2 - V(q)$).
We compute the derivative $DS|_x$ of S: $x, h \in X$

$$\begin{aligned}
S(x+h) &= \int_0^T L\left(x(t) + h(t), \dot{x}(t) + \dot{h}(t)\right)\, dt \\
&= \int_0^T \left(L(x(t), \dot{x}(t)) + \left\{D_qL|_{(x(t),\dot{x}(t))} \cdot h(t) + D_vL|_{(x(t),\dot{x}(t))} \cdot \dot{h}(t)\right\}\right) + o(\|h\|_X^2) \\
&= S(x) + \underbrace{\int_0^T \left(D_qL|_{(x(t),\dot{x}(t))} \cdot h(t) + D_vL|_{(x(t),\dot{x}(t))} \cdot \dot{h}(t)\right)}_{DS|_x \cdot h} + o(\|h\|_X^2) \\
&= S(x) + D_vL|_{(x(T),\dot{x}(T))} \cdot h(T) - D_vL|_{(x(0),\dot{x}(0))} \\
&+ \int_0^T \left(D_qL|_{(x(t),\dot{x}(t))} - \left(\frac{d}{dt} D_vL|_{(x(t),\dot{x}(t))}\right)\right) h(t)\, dt + o(\|h\|_X^2)
\end{aligned}$$

for $h \in X$ such that $h(0) = h(T) = 0$

$$DS|_x h = 0 \quad \Leftrightarrow \quad D_qL|_{(x(t),\dot{x}(t))} - \frac{d}{dt} D_vL|_{(x(t),\dot{x}(t))} = 0$$

the Euler-Lagrange equation.

Definition 3.32 (Dense subspace)**.**
A subset $A \subset X$ of a topological space $(X, \mathcal{T})$ is called dense, if and only if $\overline{A} = X$.

Remark 3.33. If $A \subset X$ is dense, then $A \cap O \neq \emptyset$ for any $O \in \mathcal{T} \backslash \{\emptyset\}$

Proposition 3.34. *Let $f, g : X \to Y$ be continuous functions, Y a Hausdorff space, and A a dense set. Then*

$$f|_A = g|_A \quad \Rightarrow \quad f = g$$

Exercise 3.3. Proof Proposition 3.34.

Chapter 4

Implicit functions and ordinary differential equations

4.1 Implicit function theorem

We ask the question of when it is possible to smoothly parameterise the level sets of a function.

$$F : \underbrace{\mathbb{R}^n \times \mathbb{R}^m}_{\mathbb{R}^{n+m}} \to \mathbb{R}^m, \qquad (x,y) \mapsto F(x,y)$$

If F is "smooth", then we expect the level sets to be n-dimensional submanifolds of the domain $\mathbb{R}^{n+m}$, i.e. sets that locally looks like the graph of a smooth function $g : \mathbb{R}^n \to \mathbb{R}^m$

Theorem 4.1 (Implicite). *Let $G \subset \mathbb{R}^{n+m}$ be open, $F \in C^1(G, \mathbb{R}^m)$, and*

$$N := \{(x,y) \in G \mid F(x,y) = 0\}.$$

If for $(a,b) \in N$ it holds that the matrix:

$$D_y F|_{(a,b)} = \begin{pmatrix} \frac{\partial F_1}{\partial y_1} & \cdots & \frac{\partial F_1}{\partial y_m} \\ \vdots & & \vdots \\ \frac{\partial F_m}{\partial y_1} & \cdots & \frac{\partial F_m}{\partial y_m} \end{pmatrix} (a,b)$$

is invertible, then there exists open neighbourhoods $U_1 \subset \mathbb{R}^n$ of a and $U_2 \in \mathbb{R}^m$ of b with $U_1 \times U_2 \subset G$ and a function $g \in C^1(U_1, U_2)$ such that

$$N \cap (U_1 \times U_2) = \mathrm{graph}(g)$$

(more explicit: $\forall (x,y) \in U_1 \times U_2 : F(x,y) = 0 \quad \Leftrightarrow \quad g(x) = y$)
($\Leftrightarrow$ one can solve $F(x,y) = g(x)$ locally for y, $F(x, g(x)) = 0$)
Moreover,

$$Dg|_x = -\left(D_y F|_{(x,g(x))}\right)^{-1} \cdot D_x F|_{(x,g(x))}$$

Definition 4.2. Let $G, H \subset \mathbb{R}^n$ be open. A map $C^1(G, H)$ is called a *diffeomorphism*, if it is bijective and also the inverse $f^{-1} \in C^1(H, G)$.

Exercise 4.1.

1. Prove that the differential $Df|_x : \mathbb{R}^n \to \mathbb{R}^n$ of a diffeomorphism f is always invertible for all $x \in G$.
2. Give an example of a bijection $f \in C^1$ such that f^{-1} is not continuously differentiable.

Theorem 4.3 (Inverse function theorem)**.** *Let $G \subset \mathbb{R}^n$ be open and $f \in C^1(G, \mathbb{R}^n)$. If for $a \in G$ it holds that $Df|_a$ is invertible then there exists an open neighbourhood U of a such that $f|_U : U \to f(U) \subset \mathbb{R}^n$ is a diffeomorphism.*

Definition 4.4 (Local extremum under constraint)**.**
Let $G \subset \mathbb{R}^n$ be open and $f, h \in C^1(G, \mathbb{R})$. Let $N = \{x \in G \mid h(x) = 0\}$ and $a \in N$. We say that f has a *local extremum* (maximum or minimum) *at the point a under the constraint h = 0* if $f|_N$ has a local extremum at a.

Theorem 4.5 (Necessary condition for local extremum under constraint)**.** *Let G, f, h, N as above. If $a \in N$ is a regular point of h (i.e. $\nabla h|_a \neq 0$) and a local extremum of f under the constraint $h = 0$, then there exists $\lambda \in \mathbb{R}$ such that:*

$$\nabla f(a) = \lambda \nabla h(a) \tag{4.1}$$

with λ being the Lagrange parameter.

Theorem 4.6 (Sufficient condition for local extremum under constraint)**.** *Let $G \subset \mathbb{R}^n$ be open, $f, h \in C^2(G, \mathbb{R})$. Let for $a \in N$ the necessary condition Eq. (4.1) be satisfied, i.e. there exists $\lambda \in \mathbb{R}$ such that $\nabla F(a) = \nabla(f - \lambda h)(a) = 0$, then:*

1. *If $D^2F|_a(v, v) > 0$ for all $v \in \mathbb{R}^n \backslash \{0\}$, $Dh|_A v = 0$, then f has a strict local minimum at a under the constraint $h = 0$.*
2. *If $D^2F|_a(v, v) < 0$ for all $v \in \mathbb{R}^n \backslash \{0\}$, $Dh|_A v = 0$, then f has a strict local maximum at a under the constraint $h = 0$.*
3. *If $D^2F|_a(v, v)0$ is indefinite, then f has no local extremum at x_0.*

Remark 4.7. If $h : G \subset \mathbb{R}^n \to \mathbb{R}^k$, then $N = \{h = 0\}$ is a $n - k$-dimensional submanifold. In this case, the necessary condition for extremum under constraint N becomes

$$\begin{aligned} & \nabla f(a) \in \operatorname{span}\{\nabla h_1(a), \nabla h_2(a), \ldots, \nabla h_k(a)\} \\ \Leftrightarrow \quad & \exists \lambda \in \mathbb{R}^k : \quad \nabla(f - \lambda \cdot h)|_a = 0 \quad \Leftrightarrow \quad \nabla f = \lambda_1 \nabla h_1 + \ldots + \lambda_k \nabla h_k \,. \end{aligned}$$

4.2 Ordinary differential equations

Definition 4.8 (Autonomous first order ODE)**.** Let $G \subset \mathbb{R}^n$ open, $v \in C(G, \mathbb{R}^n)$ a continuous vector field, and $I \subset \mathbb{R}$ an open interval containing $0 \in \mathbb{R}$. A curve $\gamma \in C^1(I, G)$ is a solution of the *autonomous first order ODE*

$$\gamma' = v(\gamma) \text{ with } \textit{initial datum } x_0 \in G\,,$$

if $\gamma' = v \circ \gamma$, i.e. $\gamma'(t) = v(\gamma(t))\ \forall t \in I$, and $\gamma(0) = x_0$. In this context, γ is also called an *integral curve* of v.

Exercise 4.2. Determine and draw some integral curves for the vector fields

$$\begin{aligned} v :&\mathbb{R}^2 \to \mathbb{R}^2, \quad (x,y) \mapsto v(x,y) = \begin{pmatrix} -x \\ y \end{pmatrix}\,, \\ w :&\mathbb{R}^2 \to \mathbb{R}^2, \quad (x,y) \mapsto w(x,y) = \begin{pmatrix} x \\ y \end{pmatrix}\,. \end{aligned}$$

Definition 4.9 (Autonomous m'th-order ODE)**.** Let $m \in \mathbb{N}$. An *autonomous n^{th} order ODE* on a domain $D \subset \mathbb{R}^n$ is given by a continuous function

$$f : D \times \underbrace{\mathbb{R}^n \times \ldots \times \mathbb{R}^n}_{(m-1)\text{-copies}} \to \mathbb{R}^n$$

and the equation

$$\gamma^{(m)} = f(\gamma, \gamma', \gamma'', \ldots, \gamma^{(m-1)})\,.$$

Now $\gamma \in C^m(I, D)$ is called *a solution* with initial datum $(x_0, y_1, \ldots, y_{m-1})$, if

$$\gamma^{(m)}(t) = f(\gamma(t), \gamma'(t), \ldots, \gamma^{(m-1)}(t)) \quad \forall t \in I$$

and $\gamma(0) = x_0$ and $\gamma^{(j)}(0) = y_j\ \forall j = 1, \ldots, m-1$.

Definition 4.10 (Non-autonomous first order ODE)**.** Let $J \subset \mathbb{R}$ be an open interval. A continuous map

$$v : J \times D \to \mathbb{R}^n, \quad (t, x) \mapsto v(t, x)$$

is called a *time-dependent vector field*. The ODE

$$\gamma' = v(t, \gamma)\,,$$

is called a *non-autonomous first order ODE*. If $I \subset J$ is an open subinterval, $t_0 \in I$, $x_0 \in D$ then $\gamma : I \to D$ is a solution with initial value x_0 for initial time t_0, if

$$\gamma'(t) = v(t, \gamma(t)) \qquad \forall t \in I$$

and $\gamma(t_0) = x_0$.

Remark 4.11. Non-autonomous first-order and autonomous m'th-order ODEs all reduce to autonomous first-order ODEs.

Definition 4.12 (Local and global Lipschitz condition)**.** Let $U \subset \mathbb{R} \times \mathbb{R}^n$ and $v \in C(U, \mathbb{R}^n)$

1. We say that v satisfies a *Lipschitz condition*, if there exists $L \geq 0$ such that

$$\forall (t,x),(t,y) \in U: \quad \|v(t,x) - v(t,y)\| \leq L\|x - y\|$$

2. We say that v satisfies a *local Lipschitz condition*, if every $(t,x) \in U$ admits a neighbourhood $V \subset U$ such that $v|_V$ satisfies a Lipschitz condition.

Theorem 4.13 (Picard-Lindelöf)**.** *Let $U \subset \mathbb{R} \times \mathbb{R}^n$ be a domain and let $v \in C(U, \mathbb{R}^n)$ satisfy a local Lipschitz condition.*

1. *Local existence: For any $(t_0, x_0) \in U$ there exists $\delta > 0$ and a curve $\gamma \in C^1((t_0 - \delta, t_0 + \delta), \mathbb{R}^n)$ that is a solution of $\gamma' = v(t, \gamma)$ with initial datum $\gamma(t_0) = x_0$.*

2. *Uniqueness: If $J \subset \mathbb{R}$ is an interval with $t_0 \in J$ and $\tilde{\gamma} : J \to \mathbb{R}^n$ solves $\gamma' = v(t, \gamma)$ with $\tilde{\gamma}(t_0) = x_0$, then*

$$\tilde{\gamma}(t) = \gamma(t) \qquad \forall t \in J \cap (t_0 - \delta, t_0 + \delta)\,.$$

Definition 4.14 (Maximal solution)**.** Let $v \in C(J \times G, \mathbb{R}^n)$ satisfy a local Lipschitz condition. A solution $\gamma : I \to G$ of $\gamma' = v(t, \gamma)$ is called maximal solution, if the following holds: If $I \subset \tilde{I} \subset J$ and $\tilde{\gamma} : \tilde{I} \to G$ is a solution of $\gamma' = v(t, x)$ with $\tilde{\gamma}|_I = \gamma$, then $\tilde{I} = I$.

Corollary 4.15. *Under the conditions of the Picard-Lindelöf-theorem, there exists for any initial value a unique maximal solution.*

Theorem 4.16. *Let $J = (j_-, j_+) \subset \mathbb{R}$, $G \subset \mathbb{R}^n$ a domain, and $v \in C(J \times G, \mathbb{R}^n)$ satisfy a local Lipschitz condition. Let $\gamma : (t_-(t_0, x_0), t_+(t_0, x_0)) \to G$ be the unique maximal solution of $\gamma' = v(t, x)$ for the initial value $(t_0, x_0) \in J \times G$. If $t_+(t_0, x_0) < j_+$, then for any compact $\subset G$ there exists $0 < \tau_K < t_+(t_0, x_0)$ such that*

$$\gamma(t) \notin K \qquad \forall t \in (\tau_K, t_+(t_0, x_0))\,.$$

Definition 4.17. A locally Lipschitz vector field $v \in C(G, \mathbb{R}^n)$ is *complete*, if there exists a global solution $\gamma_{x_0} \in C^1(\mathbb{R}, G)$ of $\gamma' = v(\gamma)$ with $\gamma_{x_0}(0) = x_0$ for any initial value $x_0 \in G$. The associated *flow* is:

$$\Phi : \mathbb{R} \times G \to G, \qquad (t,x) \mapsto \Phi(t,x) = \gamma_x(t)$$

and

$$\gamma_t : G \to G, \qquad x \mapsto \Phi_t(x) = \Phi(t,x)$$

is called the *flow map* at time t. It satisfies

$$\Phi_t \circ \Phi_s = \Phi_{t+s} \qquad \forall t, s \in \mathbb{R}$$

i.e.

$$\mathbb{R} \to \mathrm{Bij}(G \to G), \quad t \mapsto \Phi_t$$

is a groups action of $(\mathbb{R}, +)$ on the set G.

Theorem 4.18. *If v satisfies a local Lipschitz condition and is complete, then the corresponding flow maps $\Phi_t : G \to G$ are continuous. If $v \in C^1$, then the flow maps $\Phi_t : G \to G$ are also C^1.*

4.3 Linear ordinary differential equations

Definition 4.19 (Non-autonomous homogeneous linear system)**.** Let $J \subset \mathbb{R}$ open interval, $A : J \to \mathcal{L}(\mathbb{R}^n)$ is continuous.

1. The ODE
$$\gamma' = A(t)\gamma \qquad (v(\gamma) = A(t)\gamma)$$
is called a non-autonomous, homogeneous, linear system.

2. If $b : J \to \mathbb{R}^n$ is continuous, then
$$\gamma' = A(t)\gamma + b(t)$$
is called a non-autonomous, inhomogeneous, linear ODE.

Example 4.20. In the homogeneous autonomous case

$$\gamma' = A\gamma$$

the unique global solution with initial datum $x_o \in \mathbb{R}^n$ is

$$\gamma(t) = e^{At}x_0$$

where $e^{At} = \sum\limits_{n=0}^{\infty} \frac{t^n A^n}{n!}$.

Theorem 4.21. *$J \subset \mathbb{R}$ open, $A : J \to \mathcal{L}(\mathbb{R}^n)$ and $b : J \to \mathbb{R}^n$ continuous. Then for very $t_0 \in J$ and $x_0 \in \mathbb{R}^n$ there exists a unique maximal solution $\gamma : J \to \mathbb{R}^n$ of the ODE*

$$\gamma' = A(t)\gamma + b(t), \qquad \textit{with} \qquad \gamma(t_0) = x_0\,.$$

Lemma 4.22 (Grönwall)**.** *Let* $a < b$ *and* $u : [a, b] \to [0, \infty)$ *continuous. Assume* $\exists L, C \geq 0$ *such that for* $t \in [a, b]$*:*

$$u(t) \leq C + L \int_a^t u(s)\, ds\,.$$

Then

$$u(t) \leq Ce^{L(t-a)}\,.$$

Theorem 4.23 (The propagator of a non-autonomous, homogeneous linear system)**.** $J \subset \mathbb{R}$ *open,* $A : J \to \mathcal{L}(\mathbb{R}^n)$ *continuous constituting a non-autonomous, homogeneous linear system. For fixed* $t_0 \in J$ *we define*

$$\Phi_t : \mathbb{R}^n \to \mathbb{R}^n, \quad x_0 \mapsto \gamma_{x_0}(t) \quad \forall t \in J \tag{4.2}$$

with $\gamma_{x_0}(t_0) = x_0$ *and call it the flow map (Eq. (4.2)) or the propagator.*

Theorem 4.24. $\Phi_t : \mathbb{R}^n \to \mathbb{R}^n$ *from Eq. (4.2) is a linear isomorphism.*

We hence get that the solutions $\{\gamma \in C^1(J, \mathbb{R}^n) \mid \gamma' = A(t)\gamma\}$ with A as in Theorem 4.23, form a n-dimensional subspace of $C^1(J, \mathbb{R}^n)$.

Theorem 4.25 (Variation of constants)**.** *Let* $\Phi_t : \mathbb{R}^n \to \mathbb{R}^n$ *be the propagator of a homogeneous linear system* $\gamma' = A(t)\gamma$ *and* $b : J \to \mathbb{R}^n$ *continuous. Then the solution of the inhomogeneous equation:*

$$\gamma' = A(t)\gamma + b(t) \qquad \textit{with} \qquad \gamma(t_0) = x_0$$

is

$$\gamma(t) = \Phi_t\Big(x_0 + \int_{t_0}^t \Phi_s^{-1} b(s)\, ds\Big)\,.$$

This approach is called the variation of constants.

Chapter 5

Measure and integration theory

Remark 5.1. 1. Idea of the Riemann Integral: Approximate f by "stair functions", i.e. decompose the domain into intervals (rectangles, cubes, ...) and use

$$g(x) = \sum_{i=1}^{n} \alpha_i \chi_{[a_i, a_{i+1}]}(x)$$

where for $A \subset \mathbb{R}$ the *characteristic function* of A is defined:

$$\chi_A(x) = \begin{cases} 1 & x \in A \\ 0 & x \notin A \end{cases}$$

The integral of a stair function is:

$$\int g(x)dx = \sum_{i=1}^{n} \alpha_i (a_{i+1} - a_i)$$

2. Idea of the Lebesgue integral: Decompose the range of the function into intervals $[\alpha_i, \alpha_{i+1})$ and approximate by "simple functions"

$$g(x) = \sum_{i=1}^{n} \alpha_i \chi_{A_i}(x)$$

e.g. $A_i = f^{-1}([\alpha_i, \alpha_{i+1}))$ (not interval in general).
The integral of a simple function is given by:

$$\int g(x)\, dx = \sum_{i=1}^{n} \alpha \lambda(A_i)$$

where $\lambda(A_i)$ is the "length" of A_i (area, volume, measure).

Example 5.2. $f(x) = \chi_{\mathbb{Q} \cap [0,1]}(x)$ is not Rieman integrable, but it is Lebesgue integrable:

$$\int_0^1 f(x)dx = 1 \cdot \lambda(\mathbb{Q} \cap [0,1]) + 0 \cdot \lambda([0,1] \backslash \mathbb{Q}) = 0$$

Remark 5.3. Two advantages of the Lebesgue integral:

1. There are more integrable functions, meaning spaces of Lebesgue integrable functions are complete.
2. The Lebesgue integral can be defined on all spaces where one can define a measure λ (not only on $\mathbb{R}$ or $\mathbb{R}^n$).

5.1 Basic notions of measure theory

In 1924 Banach and Tarski managed to prove that there exists no volume map $\mathrm{vol} : \mathcal{P}(\mathbb{R}^3) \to [0, \infty)$ such that

1. $\mathrm{vol}(\emptyset) = 0$, $\mathrm{vol}([0,1]^3) = 1$
2. $X_1, ..., X_k \in \mathcal{P}(\mathbb{R}^3)$ pairwise disjoint, then
$$\mathrm{vol}\Big(\bigcup_{i=1}^{k} X_i) = \sum_{i=1}^{k} \mathrm{vol}(X_i)$$
3. Invariant under transformations. Let $v \in \mathbb{R}^3$, $A \in O(3)$, $X \in \mathbb{R}^3$, then
$$\mathrm{vol}(\{Ax + v \ : \ x \in X\}) =: \mathrm{vol}(A \cdot X + v) = \mathrm{vol}(X)$$

To circumvent this problem σ-algebras and measure theory was created.

Definition 5.4 (σ-algebra)**.** A family $\mathcal{A} \subset \mathcal{P}(X)$ of subsets of a set X is called *σ-algebra*, if

1. $\emptyset \in \mathcal{A}$
2. $A \in \mathcal{A} \quad \Rightarrow \quad A^C \in \mathcal{A}$
3. $A_k \in A \text{ for } k \in \mathbb{N} \quad \Rightarrow \quad \bigcup_{k=1}^{\infty} A_k \in \mathcal{A}$

The elements of $\mathcal{A}$ are called the $\mathcal{A}$-*measurable sets*.

Proposition 5.5. *Let $\mathcal{A}$ be a σ-algebra on X. Then*

1. $X \in \mathcal{A}$
2. $A_k \in \mathcal{A}$ *for* $k \in \mathbb{N} \quad \Rightarrow \quad \bigcap_{k=1}^{\infty} A_k \in \mathcal{A}$
3. $A, B \in \mathcal{A} \quad \Rightarrow \quad A \cup B \in \mathcal{A}$, $A \cap B \in \mathcal{A}$, *and* $A \backslash B \in \mathcal{A}$

Exercise 5.1. Proof Proposition 5.5.

Example 5.6. 1. $\mathcal{P}(X)$ and $\{\emptyset, X\}$ are σ-algebras on X

2. If $A_j, j \in I$ are, σ-algebras on X, so is $\bigcap_{j\in I} A_j$

Theorem 5.7 (Generating system)**.** *Let $\mathcal{F} \subset \mathcal{P}(X)$. Then the σ-algebra generated by $\mathcal{F}$ is:*

$$\mathcal{A}_{\mathcal{F}} = \bigcap_{\substack{\mathcal{B} \text{ is } \sigma\text{-alg.} \\ \mathcal{F} \subset \mathcal{B}}} \mathcal{B}$$

Any $\mathcal{F} \subset \mathcal{P}(X)$ that generates $\mathcal{A}$ is called generating system for $\mathcal{A}$.

Definition 5.8 (Borel σ-algebra)**.** Let $(X, \mathcal{T})$ be a topological space. Then

$$\mathcal{A}_{\mathcal{T}} = \mathcal{B}$$

is called the *Borel σ-algebra* on X.

Definition 5.9 (Measure)**.** Let $\mathcal{A} \subset \mathcal{P}(X)$ be a σ-algebra. A map $\mu : \mathcal{A} \to [0, \infty]$ is called a *measure*, if

1. $\mu(\emptyset) = 0$
2. For pairwise disjoint sets $A_k \in \mathcal{A}$, $k \in \mathbb{N}$,

$$\mu\Big(\bigcup_{k=1}^{\infty} A_k\Big) = \sum_{k=1}^{\infty} \mu(A_k) \qquad (\sigma\text{-additivity})$$

We further call μ

1. a *finite* measure, $\mu(X) < \infty$,
2. a *σ-finite* measure, if there exists a decomposition $X = \bigcup_{k=1}^{\infty} A_k$ such that $\mu(A_k) < \infty\ \forall k$.

The pair $(X, \mathcal{A})$ is called a *measurable space*, the triple $(X, \mathcal{A}, \mu)$ is called a *measure space*.

Example 5.10. Let X be a set and $x_0 \in X$. Then

$$v : \mathcal{P}(X) \to [0, \infty], \quad A \mapsto v(A) = \begin{cases} |A| & \text{if } A \text{ is finite} \\ \infty & \text{otherwise} \end{cases} \qquad \text{"counting measure"}$$

and

$$\delta_{x_0} : \mathcal{P}(X) \to [0, \infty], \quad A \mapsto \delta_{x_0}(A) = \begin{cases} 1 & \text{if } x_0 \in A \\ 0 & \text{otherwise} \end{cases} \qquad \text{"Dirac measure at } x_0\text{"}$$

are measures.

Proposition 5.11. *Let μ be a measure on $(X, \mathcal{A})$ and $A, B \in \mathcal{A}$. Then*

$$\mu(A \cup B) + \mu(A \cap B) = \mu(A) + \mu(B)$$

and if $A \subset B$

$$\mu(B) = \mu(A) + \mu(B \backslash A) \quad \Rightarrow \mu(A) \leq \mu(B). \qquad \textit{monotony}$$

For $A_j \in \mathcal{A}$, $j \in \mathbb{N}$,

$$\mu\Big(\bigcup_{j=1}^{\infty} A_j\Big) \leq \sum_{j=1}^{\infty} \mu(A_j) \qquad \textit{sub-additivity}$$

and if $A_j \subset A_{j+1}$ $\forall j$, then

$$\lim_{j \to \infty} \mu(A_j) = \mu\Big(\bigcup_{j=1}^{\infty} A_j\Big)$$

Definition 5.12 (Measurable function and the push-forward of a measure)**.** Let $(X, \mathcal{A})$ and $(Y, \mathcal{C})$ be measure spaces. A map $f : X \to Y$ is called *$\mathcal{A}$-$\mathcal{C}$-measurable*, if

$$C \in \mathcal{C} \quad \Rightarrow \quad f^{-1}(C) \in \mathcal{A}.$$

If μ is a measure on $(X, \mathcal{A})$ then

$$f^*\mu : \mathcal{C} \to [0, \infty], \quad C \mapsto f^*\mu(C) = \mu\left(f^{-1}(C)\right)$$

is called its *push-forward* under f.

Remark 5.13 (Terminology from probability theory). A measure space $(X, \mathcal{A}, \mu)$ with $\mu(X) = 1$ is called a *probability space*. Then the elements $A \in \mathcal{A}$ are called *events* and $\mu(A)$ the probability of the event. Measurable functions $f : X \to Y$, $(Y, \mathcal{C})$ a measurable space, are called *random variables* and the probability measure $f^*\mu$ is called the *distribution* of f.

Theorem 5.14 (Lebesgue measure)**.** *There is a unique measure λ on $(\mathbb{R}^n, \mathcal{B})$ that is translation invariant (i.e. $\lambda(A + x) = \lambda(A)$, $\forall A \in \mathcal{B} \forall x \in \mathbb{R}^n$) and normalised to $\lambda((0,1)^n) = 1$. It is called the Lebesgue-Borel measure and its completion is called the Lebesgue measure.*

Exercise 5.2. Show that $\lambda(\mathbb{Q}) = 0$.

5.2 Basic notions of integration theory

Definition 5.15 (Simple function)**.**
A function $g : X \to \overline{\mathbb{R}} = \mathbb{R} \cup \{-\infty, +\infty\}$ is called simple, if $g(X) = \{\alpha_1, \ldots, \alpha_k\}$ is finite, i.e.

$$g(x) = \sum_{j=1}^{k} \alpha_j \chi_{A_j}(x) \qquad \text{with } A_j \cap A_i = \emptyset \text{ for } i \neq j$$

Definition 5.16 (Integral of non-negative measurable functions)**.**
Let $(X, \mathcal{A}, \mu)$ be a measure space and $g : X \to [0, \infty]$ a simple and measurable, then

$$\int\limits_X g \, d\mu = \sum_{j=1}^{k} \alpha_j \mu(A_j)$$

For a measurable function $f : X \to [0, \infty]$

$$\int\limits_X f \, d\mu = \sup \left\{ \int\limits_X g \, d\mu \;\middle|\; g : X \to [0, \infty] \text{ simple, measurable and } g \leq f \right\}$$

Definition 5.17 (Integral of measurable functions)**.**
A measurable function $f : X \to \overline{\mathbb{R}}$ is *integrable*, if for $f_+ = \max\{f, 0\}$ and $f_- = \max\{-f, 0\}$ it holds that

$$\int f_+ \, d\mu < \infty \qquad\qquad \int f_- \, d\mu < \infty.$$

Then

$$\int f \, d\mu = \int f_+ \, d\mu - \int f_- \, d\mu$$

Proposition 5.18. *Let $f, g : X \to \mathbb{R}$ be measurable and integrable and $\alpha \in \mathbb{R}$. Then*

1. $\int \alpha f \, d\mu = \alpha \int f \, d\mu$
2. $\int (f + g) \, d\mu = \int f \, d\mu + \int g \, d\mu$
3. $f \leq g \quad \Rightarrow \int f \, d\mu \leq \int g \, d\mu$

Theorem 5.19 (Beppo-Levi, Monotone convergence)**.** *Let $f_n : X \to [0, \infty]$ measurable and $f_n \leq f_{n+1}$ for all $n \in \mathbb{N}$. Let $f := \lim\limits_{n\to\infty} f_n$ (pointwise), then*

$$\lim_{n\to\infty} \int f_n \, d\mu = \int f \, d\mu$$

Corollary 5.20. *Let $f_n : X \to [0, \infty]$ be measurable. Then*

$$\int \liminf_{n\to\infty} f_n \, d\mu \leq \liminf_{n\to\infty} \int f_n \, d\mu$$

Definition 5.21 (Almost everywhere)**.**
We say that a property of a point $x \in X$ holds *almost surely* or *almost everywhere* with respect to a measure μ on X, if it holds for $x \in A \subset X$ and

$$\mu(X \backslash A) = 0,$$

i.e. if it fails to hold a *null set* only.

Example 5.22. 1. A real number is almost surely irrational with respect to Lebesgue's measure on $\mathbb{R}$.

2. Let $f : X \to [0, \infty]$ be measurable. Then

$$\int_X f \, d\mu = 0 \quad \Leftrightarrow \quad f = 0 \text{ almost everywhere}$$

3. Changing an integrable function f on a null set does not change $\int f \, d\mu$.

4. For integrable functions we do not include $\pm\infty$ into the range anymore.

Remark 5.23. 1. Every Rieman integrable function $f : [a, b] \to \mathbb{R}$ is also Lebesgue integrable and the integrals coincide.

2. A function $f : X \to \mathbb{C}$ is integrable, if $|f|$ is integrable and

$$\int f \, d\mu = \int Re\, f \, d\mu + i \int Im\, f \, d\mu$$

3. Analogously for $f : X \to W$ (W-finite dimensional).

4. For $f : X \to W$, W a Banach space, the generalisation is called the Bochner integral.

Definition 5.24 (L^p-spaces)**.**
Let $(X, \mathcal{A}, \mu)$ be a measure space and $1 \leq p < \infty$. Then

$$\mathcal{L}^p(X, \mu) = \{f : X \to \mathbb{C} \mid f \text{ is measurable and } |f|^p \text{ is integrable}\}$$

and for $f \in \mathcal{L}^p(X, \mu)$

$$\|f\|_{L^p} = \left(\int |f|^p \, d\mu \right)^{\frac{1}{p}} < \infty \,.$$

Moreover, $L^p(X, \mu) = \mathcal{L}^p(X, \mu)/\sim$ with respect to the equivalence relation

$$f \sim g \quad \Leftrightarrow \quad f = g \text{ almost everywhere}\,.$$

Theorem 5.25 (Completeness of L^p-spaces)**.** *Let $(X, \mathcal{A}, \mu)$ be a measure space and $1 \leq p \leq \infty$. Then $(L^p(X, \mu), \|\cdot\|_{L_p})$ is a Banach space.*

Theorem 5.26 (Dominated convergence)**.** *Let $f_n : X \to \mathbb{C}$ be measurable, $n \in \mathbb{N}$, and assume that $f(x) = \lim_{n\to\infty} f_n(x)$ exists for almost all $x \in X$. If for some $g \in L^p(X, \mu)$, $1 \leq p < \infty$ it holds that $|f_n| \leq |g|$ almost everywhere and for all $n \in \mathbb{N}$ then $f_n, f \in L^p(X, \mu)$ and*

$$\lim_{n\to\infty} \|f_n - f\|_{L^p} = 0$$

i.e. $f_n \to f$ in $L^p(X, \mu)$.

Definition 5.27 (L^∞ and the essential supremum)**.**
Let $(X, \mathcal{A}, \mu)$ be a measure space. For measurable $f : X \to \mathbb{C}$ ($|f| : X \to [0, \infty]$) we define

$$\|f\|_{L^\infty} = \inf\left\{0 \leq \lambda \leq \infty \mid \mu\left(|f|^{-1}((\lambda, \infty])\right) = 0\right\} = \operatorname{ess\,sup}|f|.$$

Using this definition one can define

$$\mathcal{L}^\infty(X) = \{f : X \to \mathbb{C} \mid f \text{ measurable and } \|f\|_{L^\infty} < \infty\}$$

and

$$L^\infty(X) = \mathcal{L}^\infty(X)/\sim$$

Example 5.28. 1. If $\mu(X) < \infty$ and $f \in L^\infty(X)$, then

$$\int_X |f|\, d\mu \leq \int_X \underbrace{\|f\|_{L^\infty}}_{=\|f\|_\infty} d\mu = \|f\|_{L^\infty} \cdot \mu(X)$$

In particular $L^\infty(X) \subset L^1(X)$ in this case. Actually, $L^p(X) \subset L^q(X)$ if $p > q$ and $\mu(X) < \infty$.

2. $X = \mathbb{R}^n$, $\mu = \lambda^n$, then $\mu(X) = \infty$ and Item 1 does not apply. We prove this by the following: Let $\alpha \in \mathbb{R}$

$$f_\alpha : \mathbb{R}^n \setminus \{0\} \to \mathbb{R}, \quad x \mapsto f_\alpha(x) = \frac{1}{\|x\|^\alpha}$$

(a) $A = B_1(0) \subset \mathbb{R}^n$. Then for $\alpha > 0$,

$$\begin{aligned}\int_{B_1(0)} |f_\alpha|^p\, d\lambda &= \int_{B_1(0)} \frac{1}{\|x\|^{\alpha+p}}\, d\lambda = C_n \int_0^1 \frac{1}{r^{\alpha p}} r^{n-1}\, dr \\ &= C_n \int_0^1 \frac{1}{r^{\alpha p+1-n}}\, dr = \begin{cases} < \infty & \text{if } \alpha < \frac{n}{p} \\ = \infty & \text{if } \alpha \geq \frac{n}{p}\end{cases}\end{aligned}$$

We also could have used $L^p(B_1(0)) \subsetneq L^q(B_1(0))$ for $p \geq q$.

(b) $A = \mathbb{R}^n \setminus B_1(0)$. Then

$$\begin{aligned}\int_A |f_\alpha|^p\, d\lambda &= \int_A \frac{1}{\|x\|^{\alpha+p}}\, d\lambda = C_n \int_1^\infty \frac{1}{r^{\alpha p}} r^{n-1}\, dr \\ &= C_n \int_1^\infty \frac{1}{r^{\alpha p+1-n}}\, dr = \begin{cases} < \infty & \text{if } \alpha > \frac{n}{p} \\ = \infty & \text{if } \alpha \leq \frac{n}{p}\end{cases}\end{aligned}$$

Putting both together we can conclude that neither $L^p(\mathbb{R}^n) \subset L^q(\mathbb{R}^n)$ nor $L^q(\mathbb{R}^n) \subset L^p(\mathbb{R}^n)$ for $p \neq q$.

3. At last we want to show that pointwise convergence does not imply convergence in the L^p-norm. Let $f_n : \mathbb{R} \to \mathbb{R}$, $x \mapsto \chi_{[n,n+1]}(x)$. Then for all $p \geq 1$ $f_n \in L^p$ with $\|f_n\|_p = 1$ and $f_n \xrightarrow{\text{p.w.}} f = 0$, but

$$\|f_n - f\|_{L^p} = \|f_n\|_{L^p} = 1\,.$$

This is because there exists no dominating function.

Theorem 5.29 (Hölder inequality)**.** *Let $f, g : X \to \mathbb{C}$ be measurable and $1 \leq p, q \leq \infty$ such that $\frac{1}{p} + \frac{1}{q} = 1$ (conjugated exponents) where $\frac{1}{\infty} = 0$. Then*

$$\|fg\|_{L^1} \leq \|f\|_{L^p} \cdot \|g\|_{L^q}$$

Remark 5.30. For $p = q = 2$ this is the *Cauchy-Schwarz inequality* on the Hilbert space L^2. Hence for $f, g \in L^2 \Rightarrow \overline{f} g \in L^1$, since

$$\underbrace{\left| \int \overline{f} g \, d\mu \right|}_{=|\langle f,g|f,g\rangle_{L^2}|} \leq \int |\overline{f} g| \, d\mu \leq \|f\|_{L^2} \cdot \|g\|_{L^2}\,.$$

Theorem 5.31 (Minkowski inequality)**.** *Let $f, g : X \to \mathbb{C}$ be measurable and $1 \leq p \leq \infty$. Then*

$$\|f + g\|_p \leq \|f\|_p + \|g\|_p\,.$$

5.3 Product measures and Fubini's theorem

Definition 5.32 (Product σ-algebras)**.**
Let $(X_1, \mathcal{A}_1)$ and $(X_2, \mathcal{A}_2)$ be measurable spaces. Then $\mathcal{A}_1 \otimes \mathcal{A}_2$ denotes the product σ-algebra on $X_1 \times X_2$ generated by sets of the form $A_1 \times A_2 \subset X_1 \times X_2$ with $A_1 \in \mathcal{A}_1$ and $A_2 \in \mathcal{A}_2$, the so called *product σ-algebra*.

Example 5.33. Let $\mathcal{B}^n \subset \mathcal{P}(\mathbb{R}^n)$ be the Borel-σ-algebra. Then $\mathcal{B}^n = \mathcal{B}^1 \otimes \ldots \otimes \mathcal{B}^1$.

Theorem 5.34. *Let $(X_1, \mathcal{A}_1, \mu_1)$ and $(X_2, \mathcal{A}_2, \mu_2)$ be σ-finite measure spaces. There exists a unique measure μ on $\mathcal{A}_1 \otimes \mathcal{A}_2$ such that for all $A_1 \in \mathcal{A}_1$, $A_2 \in \mathcal{A}_2$*

$$\mu(A_1 \times A_2) = \mu_1(A_1) \cdot \mu_2(A_2),$$

called the product measure and denoted by $\mu = \mu_1 \otimes \mu_2$.

Example 5.35. The Lebesgue-Borel measure

$$\lambda^n = \lambda^1 \otimes \ldots \otimes \lambda^1.$$

Theorem 5.36 (Tonelli). *Let $(X_1, \mathcal{A}_1, \mu_1)$ and $(X_2, \mathcal{A}_2, \mu_2)$ be σ-finite measure spaces. Let $f : X_1 \times X_2 \to [0, \infty]$ be $\mathcal{A}_1 \otimes \mathcal{A}_2$-mensurable. Then*

$$\int\limits_{X_1 \times X_2} f \, d(\mu_1 \otimes \mu_2) = \int\limits_{X_1} \Big(\underbrace{\int\limits_{X_2} f_{x_1} \, d\mu_2}_{fct.\ of\ x_1} \Big) d\mu_1 = \int\limits_{X_2} \Big(\int\limits_{X_1} f_{x_2} \, d\mu_1 \Big) d\mu_2$$

where $f_{x_1} : X_2 \to \overline{\mathbb{R}}, \quad x_2 \mapsto f(x_1, x_2)$.

Example 5.37. $X_1 = X_2 = [0, 1], \mu_1 = \lambda^1, \mu_2 = v$ the counting measure. Hence (X_2, μ_2) is not σ-finite.

$$f : X_1 \times X_2 \to [0, \infty]$$
$$(x_1, x_2) \mapsto \delta_{x_1, x_2} := \begin{cases} 1 & \text{if } x_1 = x_2 \\ 0 & \text{otherwise} \end{cases}$$

Now the results for the integrals. For the first integral, we calculate

$$\int \underbrace{f(x_1, x_2)}_{= f_{x_2}(x_1)} \, d\mu_1 = 0$$

and obtain

$$\int \Big(\int f_{x_2}(x_1) \, d\mu_1 \Big) d\mu_2 = 0$$

For the second integral, we calculate

$$\int \underbrace{f(x_1, x_2)}_{f_{x_1}(x_2)} \, d\mu_2 = 0 \cdot \mu_2(\chi_{\{f=0\}}) + 1 \cdot \mu_2(\chi_{\{f=1\}}) = 1$$

hence obtain

$$\int \Big(\int f_{x_1}(x_2) \, d\mu_2 \Big) \, d\mu_1 = 1.$$

I.e. the two integral do not agree.

Theorem 5.38 (Fubini). *Let $(X_1, \mathcal{A}_1, \mu_1)$ and $(X_2, \mathcal{A}_2, \mu_2)$ be σ-finite measure spaces and $f : X_1 \times X_2 \to \mathbb{C}$ measurable. Then the following two statements hold.*

1. *We find the following equivalence:*

$$\int\limits_{X_1} \Big(\int\limits_{X_2} |f_{x_1}| \, d\mu_2 \Big) d\mu_1 < \infty \quad \textit{or} \quad \int\limits_{X_2} \Big(\int\limits_{X_1} |f_{x_2}| \, d\mu_1 \Big) d\mu_2 < \infty$$

if and only if

$$f \in L^1(X_1 \times X_2, \mu_1 \otimes \mu_2) \, .$$

2. If $f \in L^1(X_1 \times X_2, \mu_1 \otimes \mu_2)$*, then*

$$\int\limits_{X_1 \times X_2} f\, d(\mu_1 \otimes \mu_2) = \int\limits_{X_1} \Big(\int\limits_{X_2} f_{x_1} d\mu_2 \Big) d\mu_1$$

$$= \int\limits_{X_2} \Big(\int\limits_{X_1} f_{x_2} d\mu_1 \Big) d\mu_2$$

Example 5.39. $X = \mathbb{N}, \mathcal{A} = \mathcal{P}(\mathbb{N}), \mu = v$ counting measure. $f : X \to \mathbb{R}$, $n \mapsto f(n)$ (real sequences). The integral is then defined as:

$$\int\limits_X f\, d\mu = \sum_{n=1}^{\infty} f(n)$$

Now we consider $g : X \times X \to \mathbb{R}$, $(n, m) \mapsto g(n, m)$. There appears the question of whether it is possible to change the order of summation. Fubini's theorem allows us to answer yes to that question in the case of absolute convergence. Hence if $\sum\limits_m \sum\limits_n |g(n, m)| < \infty$, or equivalently with the summation order changed, we can change the summation.

Chapter 6

Physics preface

A physical theory is a (mathematical) model for how (parts of) the physical world could work. Physics is about

1. inventing or discovering good "theories",
2. collecting empirical data (experiments),
3. comparing the empirical facts about our world with our theoretical worlds.

Quote from Democrit (400 B.C.):

"Apparently there is colour, apparently sweetness, apparently bitterness, while in reality there are only atoms and empty space."

Modern physics uses mathematical models of the world, because (at least in good models) empirical predictions can be deduced by mathematical methods (computing, simulating, proving) in an unambiguous way.

6.1 Classical mechanics of point particles

6.1.1 Newtonian mechanics

Newton established a mathematical model for the motion of N particles (apples, planets, atoms, bullets, ...) in physical space and time. The Newtonian representation of the *physical space* is E^3 (three-dimensional euclidean space), $\mathbb{R}^3$ after the choice of an origin and an orthonormal basis. The *time* in Newtonian mechanics is described by E^1 (euclidean line), $\mathbb{R}$ after the choice of an origin.

Definition 6.1 (Configuration space of N point particles)**.**

$$q \in \mathbb{R}^{3N} = \underbrace{\mathbb{R}^3 \times \ldots \times \mathbb{R}^3}_{N\text{-copies}} = \text{configuration space.}$$

$q = (q_1, q_2, \dots, q_N)$, $q_j \in \mathbb{R}^3$ position of the j^{th} particle.

In this model, the "world" is completely specified by the position of all particles at all times, i.e. by a curve

$$\gamma : \mathbb{R} \to \mathbb{R}^{3N}, \quad t \mapsto \gamma(t)$$

in configuration space.

The physical "law" is a second-order ODE for γ, Newton's law:

$$\ddot{\gamma} = M^{-1} \cdot F(t, \gamma, \dot{\gamma})$$

with M being the mass matrix, $F(t, \gamma, \dot{\gamma})$ the force field and $\ddot{\gamma}$ the acceleration. Only the solutions to this ODE are possible worlds, according to Newtonian mechanics. Assuming sufficient regularity of F, a unique solution is determined by specifying the positions $\gamma(t_0)$ and the velocities $\dot{\gamma}(t_0)$ at some time $t_0 \in \mathbb{R} \Rightarrow$ predictions of the theory.
The explicit specification of M and F is also part of the law.

Example 6.2 (Gravitating bodies)**.**

$$M = \begin{pmatrix} m_1 & & 0 \\ & \ddots & \\ 0 & & m_N \end{pmatrix}$$

with m_i being the mass of the i^{th} body,

$$F_j(t, q, v) = F_j(q) = G \sum_{i \neq j} \frac{m_i m_j (q_i - q_j)}{\|q_i - q_j\|^3} \tag{6.1}$$

For $N = 2$ one finds Kepler's ellipses as special solutions, meaning Kepler's laws follow from Newtonian gravitation.

The gravitational force is an example of a *conservative* force field, i.e. a force $F : \mathbb{R}^{3N} \to \mathbb{R}^{3N}$ that is the negative gradient of a scalar function $V : \mathbb{R}^{3N} \to \mathbb{R}$, the so-called *potential*, $F = -\nabla V$.

Exercise 6.1. Check that a Newtonian potential for Eq. (6.1) is given by $V(q) = -\frac{G}{2} \sum\limits_{i \neq j} \frac{m_j m_i}{\|q_i - q_j\|}$.

For conservative forces, Newtonian mechanics display *"conservation of energy"*. This means the function

$$E : \mathbb{R}^{3N} \times \mathbb{R}^{3N} \to \mathbb{R}$$
$$E(q, v) = \sum_{j=1}^{N} \frac{m_j}{2} \|v_j\|^2 + V(q)$$

is constant along solutions of $\ddot{\gamma} = -M^{-1}\nabla V(\gamma)$, i.e.

$$E(\gamma(t), \dot{\gamma}(t)) = E(\gamma(t_0), \dot{\gamma}(t_0)) \qquad \forall t \in \mathbb{R} \tag{6.2}$$

Exercise 6.2. Show the above assertion, i.e. if $\ddot{\gamma} = -M^{-1}\nabla V(\gamma)$, then Eq. (6.2) holds.

In other words, the solutions of the Newtonian evolution stay on level sets of the function E! If V is translation invariant, i.e.

$$V(q_1 + a, q_2 + a, \ldots, q_N + a) = V(q_1, \ldots, q_N) \qquad \forall a \in \mathbb{R}^3$$

then also the *total momentum*:

$$P(q, v) = P(v) = \sum_{j=1}^{N} m_j v_j \in \mathbb{R}^3$$

is conserved. If V is invariant under rotations of $\mathbb{R}^3$, i.e.

$$V(Rq_1, \ldots, Rq_N) = V(q_1, \ldots, q_N)$$

then *angular momentum*

$$L(q, v) = \sum_{j=1}^{N} m_j q_j \times v_j$$

is conserved. This observation of symmetries leading to the conservation of functions in q and p is more than by accident but follows the so-called conservation laws. As is the case for any 2nd-order ODE, one can write Newton's equation as a first-order ODE on $\mathbb{R}^{6N}$ leading to the concept of *Hamiltonian mechanics*.

6.1.2 Lagrangian mechanics

Another very popular and useful formalism is the *Lagrangian formulation* of classical mechanics as a variational problem: A Lagrangian function is a function

$$\mathcal{L} : \mathbb{R}^{3N} \times \mathbb{R}^{3N} \to \mathbb{R}, \quad (q, v) \mapsto \mathcal{L}(q, v)$$

(e.g. $\mathcal{L}(q, v) = \sum_{j=1}^{N} \frac{m_j}{2} \|v_j\|^2 - V(q)$). Let

$$\Gamma = \{\gamma : C^2([0, T], \mathbb{R}^{3N})\}$$

the space of C^2-paths in configuration-space on time interval $[0, T]$. The action of such a path is

$$S(\gamma) = \int_0^T \mathcal{L}(\gamma(t), \dot{\gamma}(t))\, dt$$

$$S : \Gamma \to \mathbb{R}$$

Then the principle of least action asserts that the physically possible paths are those for which S (when adding appropriate constraints) is critical, i.e.

$$D(S - \lambda \cdot H)|_\gamma = 0 \qquad \text{Euler-Lagrange equation} \tag{6.3}$$

As

$$\begin{aligned} DS|_\gamma h &= D_v\mathcal{L}|_{(\gamma(T),\dot\gamma(T))} \cdot h(T) - D_v\mathcal{L}|_{(\gamma(0),\dot\gamma(0))} \cdot h(0) \\ &+ \int_0^T \left\{ D_q\mathcal{L}|_{(\gamma(t),\dot\gamma(t))} - \left(\frac{d}{dt} D_v\mathcal{L}_{(\gamma(t),\dot\gamma(t))}\right)\right\} h(t)\, dt \end{aligned}$$

a part of Eq. (6.3) is often (when h is only contained at single points)

$$D_q\mathcal{L}|_{(\gamma(t),\dot\gamma(t))} - \frac{d}{dt} D_v\mathcal{L}|_{(\gamma(t),\dot\gamma(t)} = 0 \qquad \forall t$$

For $\mathcal{L} = \sum \frac{m_j}{2}\|v_j\|^2 - V(q)$ these are exactly Newton's equations.

6.1.3 Hamiltonian mechanics

Another approach is the one of *Hamiltonian mechanics*. The phase space of N particles in $\mathbb{R}^3$ is

$$P = \mathbb{R}^{6N},\ x \in P\,,$$

where

$$x = (\underbrace{q_1, \dots, q_N}_{\text{positions}}, \underbrace{p_1, \dots, p_N}_{\text{momenta}})$$

(in general, P is a symplectic space or manifold). The canonical *symplectic form* on $P = \mathbb{R}^{6N}$ is

$$J : \mathbb{R}^{6N} \times \mathbb{R}^{6N} \to \mathbb{R}, \quad (x_1, x_2) \mapsto \langle x_1 | I x_2 \rangle$$

with

$$I = \begin{pmatrix} 0 & id_{\mathbb{R}^{3N}} \\ -id_{\mathbb{R}^{3N}} & 0 \end{pmatrix}, \quad I^T = -I$$

The law of motion is the first-order ODE on P where the vector field is the symplectic gradient of a function. $H : P \to \mathbb{R}$, the *Hamiltonian*:

$$\dot\alpha = I\nabla H(\alpha), \qquad \alpha : \mathbb{R} \to P = \mathbb{R}^{6N}$$

With $\alpha(t) = (Q(t), P(t))$ this reads

$$\begin{aligned} \begin{pmatrix} \dot Q(t) \\ \dot P(t) \end{pmatrix} &= \begin{pmatrix} 0 & id \\ -id & 0 \end{pmatrix} \cdot \begin{pmatrix} \nabla_q H(Q(t), P(t)) \\ \nabla_p H(Q(t), P(t)) \end{pmatrix} \\ &= \begin{pmatrix} \nabla_p H(Q(t), P(t)) \\ -\nabla_q H(Q(t), P(t)) \end{pmatrix} \end{aligned}$$

For $H(q,p) = \sum_{j=1}^{N} \frac{1}{2m_j} \|p_j\|^2 + V(q)$ one finds again Newton's equation.
Let $\Phi : \mathbb{R} \times P \to P, \quad (t,x) \mapsto \alpha_x(t)$ be the flow of a Hamiltonian system. Then one has

1. conservation of energy: $H \circ \Phi_t = H \quad \forall t \in \mathbb{R}$
2. conservation of phase space volume (Liouville's theorem):

$$\Phi_t^* \lambda = \lambda \qquad \text{(i.e. } \lambda(\Phi_t(A)) = \lambda(A) \quad \forall A \in \mathcal{B}(P))$$

   with λ the Lebesgue measure, respectively Liouville measure.

6.2 Non-relativistic quantum mechanics

We would hope for a mathematical model for the motion of N particles (electrons, nuclei, atoms, ...) in physical space and time. However, quantum mechanics fall short of this expectation. Strictly speaking, quantum mechanics is a mathematical formalism that allows making empirical predictions about such systems, which are confirmed by experiments very well. The physical nature of those particles moving in space and time (also when we do not perform experiments or "observations" on them) is an ongoing debate for almost 100 years.

6.2.1 Quantum mechanics of N interacting spin-less point particles

1. The *state* of the system at time $t \in \mathbb{R}$ is completely described by the *wave-function*

$$\psi(t, \cdot) : \mathbb{R}^{3N} \to \mathbb{C},$$

   where

$$\|\psi(t,\cdot)\|^2_{L^2(\mathbb{R}^{3N})} = \int_{\mathbb{R}^{3N}} |\psi(t,q)|^2 \, d^{3N}q = 1$$

   is assumed. The physical meaning of $\psi(t,\cdot)$ is that

$$\rho(t,q) = |\psi(t,q)|^2$$

   is a *probability density*: The probability that the configuration $Q(t)$ is (or is found to be when someone measures) in a subset $\Lambda \subset \mathbb{R}^{3N}$ of configuration space is given by

$$Prob(Q(t) \in \Lambda) = \mathbb{P}^{\psi_t}(\Lambda) = \int_{\Lambda} |\psi(t,q)|^2 \, d^{3N}q.$$

Thus, the wave-function $\psi_t(\cdot) = \psi(t, \cdot)$ defines a *probability measure* $\mathbb{P}^{\psi_t}$ on configuration space. It should be noted that the wave function does not provide a mass- or charge density. Quantum mechanics is about point particles, not about smeared-out "stuff"!

2. The dynamical law specifying the time-evolution of the state $\psi(t, \cdot)$ is the *Schrödinger equation*

$$\begin{aligned} i\hbar\partial_t\psi(t,q) &= -\hbar^2 \sum_{j=1}^{N} \frac{1}{2m_j}\Delta_{q_j}\psi(t,q) + V(q)\psi(t,q) \\ &= (H\psi)(t,q) \end{aligned}$$

where the linear operator (to be defined on suitable function spaces)

$$H = -\hbar^2 \sum_{j=1}^{N} \frac{1}{2m_j}\Delta_{q_j} + V(q)$$

is called the *Hamiltonian*. The Schrödinger equation is a linear partial differential equation (PDE) for a function on configuration space $\mathbb{R}^{3N}$.

Definition 6.3 (Definition)**.** Square integrable solutions of the time-independent Schrödinger equation

$$(H\psi_E)(q) = E\psi_E(q) \qquad \text{for some } E \in \mathbb{R}$$

are called *eigenstates* of H (or energy eigenstates), and

$$\psi(t,q) = e^{-itE}\psi_E(q)$$

is a *stationary solution* of the time-dependent Schrödinger equation.
Typically, only for a discrete set $\{E_j\} \subset \mathbb{R}$ solutions of the time-independent Schrödinger equation in L^2 exist.

Example 6.4. 1. Free particle in a box:

$$-\frac{d^2}{dx^2}\psi(x) = E\psi(x) \qquad \psi \in L^2([0,1])$$

with Dirichlet boundary conditions $\psi(0) = \psi(1) = 0$.

$$\Rightarrow \psi_n(x) = \sqrt{2}\sin(n \cdot \pi \cdot x) \qquad n \in \mathbb{N}$$

with the "energy eigenstates" $E_n = n^2 \cdot \pi^2$.

2. Hydrogen atom:

$$H = -\frac{1}{2m_e}\Delta_q - \frac{c}{|q|} \qquad \text{on } L^2(\mathbb{R}^3)$$

with the energy eigenstates of $E_n = -\frac{\tilde{c}}{n^2}$ for $n \in \mathbb{N}$. The differences $E_n - E_m$ correspond to energies of spectral lines of hydrogen atoms, i.e. to photons absorbed or emitted by hydrogen. The corresponding eigenfunctions are called *orbitals*.

3. Harmonic oscillator:
$$H = -\frac{h^2}{2m}\frac{d^2}{dx^2} + \frac{m}{2}\omega^2 x^2$$
For the time-independent Schrödinger equation
$$H\psi_n = E_n\psi_n \quad \text{one gets} \quad E_n = \hbar\omega\left(n + \frac{1}{2}\right) \quad n \in \mathbb{N}_0$$
and
$$\psi_n(x) = \frac{1}{\sqrt{2^n n!}}\left(\frac{m\omega}{\pi\hbar}\right)^{1/4} H_n\left(\sqrt{\frac{m\omega}{\hbar}}x\right)e^{-\frac{m\omega x^2}{2\hbar}}$$
with $H_n(\cdot)$ being the Hermite polynomials.

6.2.2 "Axiomatic" formulation of the quantum measurement formalism

1. *State space:* The possible states of a quantum system are described by normalised vectors $\psi \in \mathcal{H}$ in a Hilbert space $\mathcal{H}$.

2. *Observables and operators:* Every physical observable A corresponds to a self-adjoint operator $\hat{A}$ on $\mathcal{H}$.

3. *Measurement process:*

   (a) *Possible outcomes:* The measurement of an observable A yields as an outcome one of the eigenvalues of the corresponding operator $\hat{A}$.

   (b) *Probabilities:* Let A be an observable and a_i an eigenvalue of the associated operator $\hat{A}$ and P_{a_i} the spectral projection on the corresponding eigenspace. The probability for obtaining the result a_i when measuring the observable A on a system in the state ψ is
   $$Prob(A = a_i \,|\, \psi) = \|P_{a_i}\psi\|^2.$$

   (c) *State after the measurement:* If the measurement of the observable A on a quantum system in the state ψ yields the outcome a_i, then after the measurement the state of the system is
   $$\psi_{coll.} = \frac{P_{a_i}\psi}{\|P_{a_i}\psi\|}$$
   "collapse of the wave function".

4. *Dynamic law:* In between measurements the state ψ of the system evolves according to the Schrödinger equation:

$$i\hbar\frac{d}{dt}\psi(t) = \hat{H}\psi(t),$$

where $\hat{H}$ is the Hamilton operator (the operator corresponding to the "energy observable")

5. *Correspondence principle:* Let $A : \mathbb{R}^{6N} \to \mathbb{R}, \quad (q,p) \mapsto A(q,p)$ be a classical observable. Then the corresponding operator $\hat{A}$ is the given

$$\hat{A} = A(q, -i\hbar\nabla_q)$$

acting on $L^2(\mathbb{R}^{3N}) = \mathcal{H}$.

Example 6.5 (Examples of the applied correspondence principle)**.**

1. $\hat{H} = H(q, -i\hbar\nabla_q) = -\hbar^2 \sum\limits_{j=1}^{N} \frac{1}{2m_j}\Delta_{q_j} + V(q)$
2. $\hat{p}_j = -i\hbar\nabla_{q_j}$ the momentum operator
3. $\hat{q}_j = q_j$ the position operator
4. $[\hat{q}_j, \hat{p}_i] = i\hbar\delta_{ij} \leftrightarrow i\hbar\{q_j, p_j\}$, i.e. the Poisson bracket of classical mechanics is replaced by the commutator.

Remark 6.6 (Dirac notation). It has turned out to be quite advantageous in the context of quantum mechanics to use the notation of Paul Dirac (1902 - 1984).

1. Vectors $\psi \in \mathcal{H}$ are written as $|\psi\rangle$ ("ket")
2. The linear functional $J_\psi : \mathcal{H} \to \mathbb{C}, \quad \varphi \mapsto J_\psi(\phi) = \langle\psi|\varphi\rangle_{\mathcal{H}}$ is written as $\langle\psi|$ ("bra")
3. The inner product of $\psi, \varphi \in \mathcal{H}$ then becomes:

$$\langle\varphi|\psi\rangle = \langle\varphi|\psi\rangle_{\mathcal{H}}$$

4. while "ket bra" is a rank one operator:

$$|\psi\rangle\langle\varphi| : \mathcal{H} \to \mathcal{H}, \quad \chi \mapsto |\psi\rangle\langle\varphi|\,\chi = \langle\varphi|\chi\rangle_{\mathcal{H}}\,\psi\,.$$

5. A bounded linear operator $\hat{A} \in \mathcal{L}(\mathcal{H})$ can be written in terms of an ONB (ψ_j) as

$$\hat{A} = \sum_{j,i=1}^{\infty} |\psi_j\rangle \left\langle \psi_j \middle| \hat{A}\psi_i \right\rangle \langle\psi_i| = \sum_{ij} \left\langle \psi_j \middle| \hat{A}\psi_i \right\rangle |\psi_j\rangle\langle\psi_i|\ .$$

Definition 6.7 (Spin)**.** Particles with "spin" are described by $\mathbb{C}^n$-valued wave functions:

1. The wave function for one particle with spin $\frac{m}{2}$ $(m \in \mathbb{N}_0)$

$$\psi : \mathbb{R}^{3N} \to \mathbb{C}^{m+1}$$

(e.g. for electrons $m = 1$, but for nuclei $m = 0$ or $m > 1$ are possibilities).

2. wave function for N particles with spin $\frac{m}{2}$

$$\psi : \mathbb{R}^{3N} \to \mathbb{C}^{(m+1)^N}$$

(e.g. 6 electrons of a carbon atom: $\psi : \mathbb{R}^{18} \to \mathbb{C}^{64}$).

3. The Pauli Hamiltonian for $N = 1$, $m = 1$ is given by

$$H\psi = \frac{1}{2m}\left(-i\nabla_q + eA(q)\right)^2 \psi - c\underbrace{\langle \sigma | B(q) \rangle_{\mathbb{R}^3}}_{\in \mathcal{L}(\mathbb{C}^2)} \psi\,,$$

with σ being Pauli matrices, $A(q)$ the vector potential and $B(q)$ the magnetic field ($B = \operatorname{curl} A$).

Definition 6.8 (Bosonic and fermionic wave functions)**.** Let us consider N identical particles, i.e. $q = (q_1, \ldots, q_N)$. We then distinguish two cases.

1. m even $\to$ bosons $\to$ $\psi : \mathbb{R}^{3N} \to \mathbb{C}^{(m+1)^N}$ symmetric under permutation of arguments,

$$\psi(q_{\pi(1)}, \ldots, q_{\pi(n)}) = U_\pi \psi(q_1, \ldots, q_N) \quad \forall \pi \in S_N$$

2. m odd $\to$ fermions $\to$ $\psi : \mathbb{R}^{3N} \to \mathbb{C}^{(m+1)^N}$ anti-symmetric under permutation of arguments

$$\psi(q_{\pi(1)}, \ldots, q_{\pi(N)}) = sgn(\pi) U_\pi \psi(q_1, \ldots, q_N) \quad \forall \pi \in S_N$$

Chapter 7

Math of (finite dimensional) quantum system

We have discussed in the previous chapter that a quantum system is described using a normalised vector in a Hilbert space. Further, we have seen that using the correspondence principle we can turn a classical observable into a quantum one, by just replacing $q \mapsto \hat{q}$ and $p \mapsto -i\hbar\nabla_q$ in the function that is an observable. Up to this point, it is, however, completely unclear what we mean by the function of operators on a Hilbert space, including also differential operators. All inconsistencies and hand-wavy arguments given so far can be moulded into a solid mathematical framework which you will come to know under the name functional calculus. In this chapter, we will develop the simplest version of a functional calculus namely it being defined on a finite-dimensional Hilbert space. Note that in this section, although we try to be as abstract and general as possible, we will always assume that $\mathcal{H}$ is finite-dimensional.

Definition 7.1 (Inner product space)**.** Let $\mathcal{H}$ be a complex vector space. We define the following

1. A *sesqui-linear form* on $\mathcal{H}$ is a map
$$B : \mathcal{H} \times \mathcal{H} \to \mathbb{C}$$
that is
   (a) linear in the second argument, i.e. $\forall \psi, \phi, \chi \in \mathcal{H}$, $\lambda \in \mathbb{C}$:
$$B(\psi, \lambda\phi + \chi) = \lambda B(\psi, \phi) + B(\psi, \chi)\,,$$
   (b) and conjugated symmetric, i.e. satisfies $B(\psi, \phi) = \overline{B(\phi, \psi)}$ for all $\psi, \phi \in \mathcal{H}$.
2. A sesqui-linear form
$$\langle \cdot | \cdot \rangle : \mathcal{H} \to \mathcal{H} \to \mathbb{C}$$

that is positive definite, i.e. $\langle\psi|\psi\rangle > 0$ for $\psi \in \mathcal{H}\setminus\{0\}$, is called *inner product* or *scalar product*. The pair $(\mathcal{H}, \langle\cdot|\cdot\rangle)$ is called *inner product space* or *pre-Hilbert space*. We say that two vectors $\phi, \psi \in \mathcal{H}$ are orthogonal, if and only if $\langle\psi|\phi\rangle = 0$.

Theorem 7.2 (Norm induced by the inner product)**.** *Let $(\mathcal{H}, \langle\cdot|\cdot\rangle)$ be an inner product space. The map*

$$\|\cdot\| : \mathcal{H} \to [0, \infty), \quad \psi \mapsto \|\psi\| := \sqrt{\langle\psi|\psi\rangle}$$

is a norm on $\mathcal{H}$.

Definition 7.3 (Hilbert space)**.** We call an inner product space $(\mathcal{H}, \|\cdot\|)$ a *Hilbert space* if it is complete with respect to the norm induced by the inner product.

Remark 7.4. Remember that finite dimensional normed spaces are complete. Thereby every finite-dimensional inner product space is a Hilbert space.

Definition 7.5 (Dual of a Hilbert space)**.** Let $(\mathcal{H}, \langle\cdot|\cdot\rangle)$ a Hilbert space. We call the space $\mathcal{L}(\mathcal{H}, \mathbb{C})$ (i.e. the set of bounded linear maps from $\mathcal{H} \to \mathbb{C}$) the space of *linear functionals* on $\mathcal{H}$ or *dual space* of $\mathcal{H}$ and denote it by $\mathcal{H}'$.

Theorem 7.6 (Riesz representation theorem)**.** *Let $\mathcal{H}$ be a Hilbert space. The map*

$$J : \mathcal{H} \to \mathcal{H}', \quad \psi \mapsto J(\psi) \quad \textit{with} \quad J(\psi)(\varphi) := \langle\psi|\varphi\rangle_{\mathcal{H}} \, ,$$

is an anti-linear isometric isomorphism.

Definition 7.7 (Adjoint operator)**.** Let $A \in \mathcal{B}(\mathcal{H})$, then the *adjoint* $A^* \in \mathcal{B}(\mathcal{H})$ of A is implicitly defined by

$$\langle\psi|A\varphi\rangle = \langle A^*\psi|\varphi\rangle \qquad \forall\psi, \varphi \in \mathcal{H} \; .$$

We call $A \in \mathcal{B}(\mathcal{H})$ *self adjoint*, if $A^* = A$.

Remark 7.8. In the above definition, we have swiped many things under the rock. Indeed it is neither clear that the adjoint operator is well-defined, linear and bounded and one needs to prove those assertions. To prove those we can immediately apply the Riesz-representation theorem.

Proof of Definition 7.7. For better understanding, we indicate in the subindex of the norm, the space they act on.

1. Well defined: For $\psi \in \mathcal{H}$ the map $\varphi \mapsto \langle\psi|A\varphi\rangle$ is a bounded linear functional, since using Cauchy-Schwarz

$$\langle\psi|A\varphi\rangle \le \|\psi\|_{\mathcal{H}}\|A\varphi\|_{\mathcal{H}} \le \underbrace{\|\psi\|_{\mathcal{H}}\|A\|_{\mathcal{B}(\mathcal{H})}}_{=C}\|\varphi\|_{\mathcal{H}} \, . \tag{7.1}$$

   This means, however, that there exists a unique $\tilde{\psi} = J^{-1}(\langle\psi|A\cdot\rangle)$ and we can just define $A^*\psi := \tilde{\psi}$.

2. Linear: Using the representation from above, for $\psi_1, \psi_2 \in \mathcal{H}$, $\alpha \in \mathbb{C}$, we find that

$$\begin{aligned} A^*(\psi_1 + \alpha\psi_2) &= J^{-1}(\langle \psi_1 + \alpha\psi_2 | A\cdot\rangle) = J^{-1}(\langle \psi_1 | A\cdot\rangle + \overline{\alpha}\,\langle \psi_2 | A\cdot\rangle) \\ &= J^{-1}(\langle \psi_1 | A\cdot\rangle) + \alpha J^{-1}(\langle \psi_2 | A\cdot\rangle) \\ &= A^*\psi_1 + \alpha A^*\psi_2 \end{aligned}$$

where we used the anti-linearity of the inner product and the anti-linearity of J^{-1}.

3. Bounded: Since J is an isometric isomorphism, i.e. preserves the norm. For $\psi \in \mathcal{H}$

$$\|A^*\psi\|_{\mathcal{H}} = \left\|J^{-1}(\langle \psi | A\cdot\rangle)\right\|_{\mathcal{H}} = \|\langle \psi | A\cdot\rangle\|_{\mathcal{H}'} \leq \|A\|_{\mathcal{B}(\mathcal{H})}\|\psi\|_{\mathcal{H}}\,.$$

To conclude the last inequality, we just used Eq. (7.1).

□

Definition 7.9 (Projection)**.** A bounded linear operator $P \in \mathcal{B}(\mathcal{H})$ on a Hilbert space $\mathcal{H}$, satisfying

$$P^2 = P$$

is called a *projection*. If P in addition is self adjoint, i.e. $P^* = P$ then P is called an *orthogonal projection*.

Theorem 7.10. *Every finite-dimensional Hilbert space of dimension n is isometrically isomorphic to $\mathbb{C}^n$. By fixing a basis, we also can identify the linear maps on $\mathcal{H}$ with matrices from $\mathbb{C}^{n\times n}$.*

Theorem 7.11 (Spectral theorem in finite dimensions)**.** *Let $\mathcal{H}$ be a finite Hilbert space of dimension n, and $A \in \mathcal{B}(\mathcal{H})$ a linear, self-adjoint operator. Then there exist real eigenvalues $a_1, \ldots, a_k$ of A and corresponding eigenspaces $E(a_1), \ldots, E(a_k)$ as well as orthogonal projections $P(a_1), \ldots, P(a_k)$ onto those subspaces with $k \leq n$ satisfying the following properties:*

1. *It holds that $\mathcal{H} = \bigoplus_{i=1}^{k} E(a_i)$.*
2. *The projections are mutually orthogonal $P(a_i)P(a_j) = P(a_j)P(a_i) = \delta_{ij}P(a_i)$ for $i, j = 1, \ldots, k$,*
3. *and sum to one $\sum_{i=1}^{k} P(a_i) = \mathbb{1}$.*
4. *The linear operator can be written as $A = \sum_{i=1}^{k} a_i P(a_i)$.*

Remark 7.12. In other words: For every self-adjoint operator on a finite-dimensional Hilbert space, there exists an orthonormal basis of eigenvectors.

We now come to the part where we sketch a mathematical framework which allows us to insert self-adjoint operators into functions. In this context, some natural requirements appear, that directly relate structures in the function space with structures in the operator space. Those requirements are the defining properties of a functional calculus.

Definition 7.13 (Definition of a functional calculus)**.** Let $(\mathcal{H}, \langle\cdot|\cdot\rangle)$ a seperable[1] Hilbert space. Let $A \in \mathcal{B}(\mathcal{H})$ self adjoint. A mapping that maps each element $f : \mathbb{R} \to \mathbb{C}$ from a sub-algebra $\mathcal{E}$ of

$$\mathfrak{B}(\mathbb{R}) := \{f : \mathbb{R} \to \mathbb{C} \ : \ f \text{ is Borel measurable and bounded, i.e. } \sup_{x\in\mathbb{R}} |f(x)| < \infty\} .$$

to an operator $f(H) \in \mathcal{B}(\mathcal{H})$, i.e.

$$\pi_A : \mathcal{E} \to \mathcal{B}(\mathcal{H}), \qquad f \mapsto \pi_A(f) = f(A) ,$$

is called a *functional calculus* if it satisfies the following properties:

1. π_A is an algebra-homomorphism, i.e.

   $$\pi_A(f + \alpha g) = (f + \alpha g)(A) = f(A) + \alpha g(A) = \pi_A(f) + \alpha\pi_A(g)$$

   and

   $$\pi_A(f \cdot g) = (f \cdot g)(A) = f(A)g(A)$$

   for all $f, g \in \mathcal{E}$, $\alpha \in \mathbb{C}$.

2. $\pi_A(\overline{f}) = \overline{f}(A) = f(A)^* = \pi_A(f)^*$ for all $f \in \mathcal{E}$.

3. $\|\pi_A(f)\| \leq \|f\|_\infty$ for all $f \in \mathcal{E}$.

4. For $z \in \mathbb{C} \setminus \mathbb{R}$ and $r_z(x) := (x - z)^{-1}$ we recover the resolvent, i.e. $r_z(A) = (A - z)^{-1}$.

5. $f \in C_0^\infty(\mathbb{R})$ satisfies $\operatorname{supp} f \cap \sigma(A) = \emptyset$, i.e. the map is zero on the spectrum of A, then $f(A) = 0$.

Theorem 7.14 (An "almost" functional calculus for polynomials)**.** *Let $\mathcal{H}$ be a finite Hilbert space of dimension n, and $A \in \mathcal{B}(\mathcal{H})$ a linear, self-adjoint operator. Let further $p \in \mathbb{C}[x]$ a polynomial, i.e.*

$$p(x) = \sum_{j=1}^{d} c_j x^j \qquad c_j \in \mathbb{C} .$$

[1] Seperable means that the Hilbert space admits for a dense countable subset, i.e. a set D which has countably many elements and $\overline{D} = \mathcal{H}$.

Then

$$p(A) = \sum_{j=1}^{d} c_j A^j = \sum_{i=1}^{k} p(a_i) P(a_i)\,.$$

defines "almost" a functional calculus on the polynomials $\mathbb{C}[x]$ *(without the fulfilment of Item 4 as* $x \mapsto (x-\lambda)^{-1}$ *is not a polynomial).*

Theorem 7.15 (Cauchy's integral formula)**.** *Let* $f : D \to \mathbb{C}$ *analytic with* $D \subset \mathbb{C}$ *a simply connected domain*[2]*. For* $\gamma : [0,1] \to D$ *a closed curve (continuous) with no self-intersections, then*

$$f(z) = \frac{1}{2\pi i} \oint_{\gamma} f(\omega) \frac{1}{\omega - z} d\omega \qquad \forall z \text{ in the interior of } \gamma\,,$$

where γ *is oriented counterclockwise.*

Cauchy's integral formula can now be used to define a proper functional calculus using the resolvent of a self-adjoint bounded operator as a building block.

Theorem 7.16 (Functional calculus for analytic functions)**.** *Let* $\mathcal{H}$ *be a finite-dimensional Hilbert space of dimension* n*, and* $A \in \mathcal{B}(\mathcal{H})$ *a linear, self-adjoint operator. Let further* $f : D \mapsto \mathbb{C}$ *a holomorphic function on the simply connected domain* D *which contains the spectrum* $\sigma(A)$ *of* A*. Then*

$$f(A) = \frac{1}{2\pi i} \oint_{\gamma} f(\omega) \frac{1}{\omega - A} d\omega$$

for any non self-intersecting path going anti-clockwise in $D \backslash \sigma(A)$ *and containing* $\sigma(A)$*, defines a functional calculus. Here* $\frac{1}{\omega - A}$ *is the resolvent of* A*. We can further simplify*

$$f(A) = \sum_{i=1}^{k} \Big(\frac{1}{2\pi i} \oint_{\gamma} f(\omega) \frac{1}{\omega - a_i} d\omega \Big) P(a_i) = \sum_{i=1}^{k} f(a_i) P(a_i)\,.$$

Remark 7.17. The path in the above theorem does not matter and neither does the shape of D. We just need that $\sigma(A) \subset D$ and that the path "encircles" the spectrum whilst not hitting any eigenvalues. Note further that the above definition of a functional calculus agrees on polynomials with Theorem 7.14. Note that under a further assumption on the functional calculus, one can show that it is indeed unique.

---

[2]The set D is open and further every continuous closed path can be shrunk to a point continuously.

Bibliography