Analysis

Arne Vorwerg

September 2023

Preface

This text is essentially a working reference and follows the classical expositions of **Bauer** [1], **Forster** [4], **Hewitt/Stromberg** [2], **Lang**[5] and **Rudin** [8] to develop the foundations of the analysis of functions needed for the research on partial differential equations in probability and physics.The necessary results from set theory and topology can be found in [11] and [9]; the corresponding references are given in the text. For reasons of brevity motivations and proofs for simple definitions and propositions are omitted.

The exposition starts with **measure theory** which is the field of mathematics dedicated to the study of the content or weight of a set expressed by its measure. If the set is defined by a **function** on a certain domain its measure can be written as an **integral**. In this case the function turns out to be the **derivative** of the measure, i.e. it is itself a measure for the rate of change of the given measure depending on the change of the domain. Thus measure theory provides one of the basic methods for the study of functions in **analysis**. Since the measure of a set can be interpreted as the probability for the realization of the events represented by its elements measure theory has proved to be a very useful foundation of **probability theory** and **statistics**.

The first section introduces measurable sets, measures and measurable functions in a pronounced analogy to the open sets, metrics and continuous functions in topology. The concept of integration provides the basis for the extension of measures on product spaces. For the sake of clarity the integral is introduced in the generalized **Bochner** variant for functions with values in **Banach spaces** and later specialized to the usual **Lebesgue integral** so as to profit from the full range of possibilities of **differentiation**. The Lebesgue integral and the associated **product measures** on countable products of measure spaces prove to be a very useful concept for the description of sequences of independent random variables and their mean rep. expected values leading to the **strong law of large numbers**. In analysis they constitute the foundation for the **integral transformations** needed for the solution of **partial differential equations**, e.r. **convolutions**, **distributions** and **fourier transforms**. These integral transformations also provide an easy approach to the **central limit theorem** of probability theory. Mean values resp. Integrals of functions on subsets are themselves measures and the **Lebesgure-Radon-Nikodym theorem** states that in fact every **positive σ-finite measure** can be represented as an **integral** over a suitable second measure. This result provides the foundation for two central theorems in **functional** resp. **real analysis**: **Positive** resp. **bounded** measures on **locally compact vector spaces** prove to be eqiuvalent to the corresponding functionals. Hence the set of all such measures on such a space is the **dual space** of a locally compact vector space. This the content of the **Riesz representation theorem**.The following sections are dedicated to the standard methods of calculus in the setting of the Lebesgue measure, e.g. the **fundamental theorem of calculus** with the **change-of-variables formula** The final sections extend the integral to **manifolds** and **differential forms** culminating in **Stoke's theorem**.

Engelsbrand, September 2023 Arne Vorwerg

Contents

1 Measurable sets 3

2 Pre-measures 4

3 Measures 5

4 Measurable functions 8

5 Integration 12

6 Lebesgue spaces 20

7 Product spaces 24

8 Product measure 26

9 Probability measures 32

10 Measures with densities 35

11 Measures on locally compact spaces 39

12 Differentiation 45

13 The fundamental theorem of calculus 49

14 Differentiation in $\mathbb{R}^n$ 52

15 Some special functions 59

16 Fourier transforms 61

17 The central limit theorem 71

References 77

1 Measurable sets

1.1 Definitions: A family $\mathcal{A} \subset \mathcal{P}(X)$ is an **algebra** iff

1. $\emptyset \in \mathcal{A}$.
2. $A, B \in \mathcal{A} \Rightarrow A \cap B; A \cup B; A \setminus B \in \mathcal{A}$

In the case of

3. $X \in \mathcal{A}$
4. $(A_n)_{n\in\mathbb{N}} \subset \mathcal{A} \wedge (n \neq m \Rightarrow A_n \cap A_m = \emptyset) \Leftrightarrow \mathring{\bigcup}_{n\in\mathbb{N}} A_n \in \mathcal{A}$

we have a **σ-algebra**. The pair $(X; \mathcal{A})$ then is a **measurable space**. Every **σ-algebra** is **closed under arbitrary countable unions and intersections** since for $(A_n)_{n\in\mathbb{N}} \subset \mathcal{A}$ we obtain paiwise disjoint $A'_n := A_n \setminus \bigcup_{1\leq k<n} A_k = \bigcap_{1\leq k<n} (A_n \setminus A_k) \in \mathcal{D}$ whence $\bigcup_{n\in\mathbb{N}} A_n = \mathring{\bigcup}_{n\in\mathbb{N}} A'_n \in \mathcal{D}$ and $\bigcap_{n\in\mathbb{N}} A_n = X \setminus \bigcup_{n\in\mathbb{N}} (X \setminus A_n)$.

1.2 Borel σ-algebrae: For an arbitrary $\mathcal{M} \subset \mathcal{P}(X)$ the intersection $\sigma(\mathcal{M})$ of all σ-algebrae containing $\mathcal{M}$ is again a σ-algebra. It is the σ-algebra **induced** by $\mathcal{M}$ and $\mathcal{M}$ is its **basis**. On a **topological space** $(X; \mathcal{O})$ we have the **Borel σ-algebra** $\mathcal{B}(X) = \sigma(\mathcal{O})$ induced by the topology $\mathcal{O}$. Owing to 1.1 it contains the **open sets** and their **countable intersections**, i.e. the **G$_\delta$-sets** as well as the **closed sets** and their **countable unions**, i.e. the **F$_\sigma$-sets**. The Borel σ-algebra of a **second countable** topological space $\mathcal{O}(\mathcal{E})$ induced by a **countable topogical basis** $\mathcal{E}$ is induced by $\mathcal{E}$ itself, i.e. $\mathcal{B}(X) = \sigma(\mathcal{O}(\mathcal{E})) = \sigma(\mathcal{E})$. In a **Hausdorff space** all **compact** sets are **closed** and hence **Borel measurable**, i.e. measurable with respect to $\mathcal{B}(X)$. For a **locally compact** X which is **countable at infinity** the **Borel σ-algebra** $\mathcal{B}(X) = \sigma(\mathcal{K})$ is induced by the family $\mathcal{K}$ of all compact sets since due to [9, 10.6] every closed set is the countable intersection of compact sets. In a **discrete space** X with $\mathcal{B}(X) = \sigma(\mathcal{O}) = \mathcal{O} = \mathcal{P}(X)$ a set is compact iff it is finite and $\sigma(\mathcal{K})$ is the σ-algebra of all sets $A \subset X$ with countable A or $X \setminus A$. Using **Zorn's lemma** ([11, 14.2.4]) we can infer that $\sigma(\mathcal{K}) = \mathcal{B}(X)$ iff X itself is countable.

1.3 Trace of a σ-algebra: The **trace σ-algebra** $\mathcal{A} \cap B := \{A \cap B : A \in \mathcal{A}\}$ on a subset $B \subset X$ of a measurable space $(X; \mathcal{A})$ simply consists of the **inter sections of measurable** A in X with B. On account of $(O_1 \cap O_2) \cap B = (O_1 \cap B) \cap (O_2 \cap B)$, $(O_1 \cup O_2) \cap B = (O_1 \cap B) \cup (O_2 \cap B)$, $(O_1 \setminus O_2) \cap B = (O_1 \cap B) \setminus (O_2 \cap B)$ and $(\bigcup_{n\in\mathbb{N}} O_n) = \bigcup_{n\in\mathbb{N}} (O_n \cap B)$ the trace $\sigma(\mathcal{O}) \cap B$ of the **Borel σ-algebra** $\mathcal{B}(X) = \sigma(\mathcal{O})$ on a **topological space** $(X; \mathcal{O})$ is identical with the σ-algebra $\sigma(\mathcal{O} \cap B)$ of the trace $\mathcal{O} \cap B$ of the **topology** $\mathcal{O}$ on B.

1.4 Intervals and figures: The **finite** unions of **pairwise disjoint right-open intervals** $\mathcal{I} = \{[a; b[\colon a \leq b \in \mathbb{R}\}$ form the **algebra** $\mathcal{F} = \left\{ \mathring{\bigcup\limits_{0\leq k\leq m}} I_k : I_k \in \mathcal{I}, m \in \mathbb{N} \right\}$ of the **one-dimensional figures** since $\emptyset = [a; a[\in \mathcal{I}$ and for $I, J \in \mathcal{I}$ we have $I \cap J \in \mathcal{I}$, $I \setminus J \in \mathcal{I}$ as well as $I \cup J \in \mathcal{I}$ in the case of $I \cap J \neq \emptyset$ resp. $I \cup J \in \mathcal{F}$ for $I \cap J = \emptyset$. Hence for $F = \mathring{\bigcup\limits_{0\leq k\leq m}} I_k \in \mathcal{F}$ and $G = \mathring{\bigcup\limits_{0\leq l\leq n}} J_l \in \mathcal{F}$ we have $F \cap G = \mathring{\bigcup\limits_{0\leq k\leq n}} \mathring{\bigcup\limits_{0\leq l\leq n}} I_k \cap J_l \in \mathcal{F}$, $F \setminus G = F \setminus (F \cap G) \in \mathcal{F}$ and $F \cup G \in \mathcal{F}$. The **right-open intervals** $[a; b[$ are **G$_\delta$-sets** hence they are Borel-measurable and because of $]a; b[= \bigcup_{k\in\mathbb{N}} \left[a + 2^{-k}; b\right[$ they **induce** the Borel σ-algebra on $\mathbb{R}$ as well as the **algebra** of **figures**: $\mathcal{B} = \sigma(\mathcal{F}) = \sigma(\mathcal{I})$. Alternative basis families are the **closed rays** $[a; \infty[$ since $[a; \infty[= \bigcup_{k\in\mathbb{N}} [a; b + k[$ bzw. $[a; b[= [b; \infty[\setminus [a; \infty[$ as well as $]a; \infty[$, $]-\infty; a[$, $]-\infty; a]$ and $]a; b]$ resp. for $a, b \in \mathbb{R}$ with analogous arguments.

1.5 Dynkin systems: A family $\mathcal{D} \subset \mathcal{P}(X)$ is a **Dynkin system** or **δ-system** iff

1. $\emptyset \in \mathcal{D}$.
2. $A \in \mathcal{D} \Leftrightarrow X \setminus A \in \mathcal{D}$
3. $(A_n)_{n\in\mathbb{N}} \subset \mathcal{D} \wedge (n \neq m \Rightarrow A_n \cap A_m = \emptyset) \Leftrightarrow \mathring{\bigcup}_{n\in\mathbb{N}} A_n \in \mathcal{D}$

1.6 Dynkin δ-π-theorem: The **Dynkin system** $\delta(\mathcal{E})$ generated by a **π-basis** $\mathcal{E} \subset \mathcal{P}(X)$ being **closed under intersections** coincides with the corresponding **σ-algebra** $\sigma(\mathcal{E})$.

Proof: For every $B \subset A$ we have $A \setminus B = X \setminus ((X \setminus A) \mathring{\cup} B)$ whence $(X \setminus A) \cap D = D \setminus (A \cap D) \in \mathcal{D}$ for every $D \in \delta(\mathcal{E})$ and $A \subset X$. Hence the family $\mathcal{D}_D := \{A \subset X : A \cap D \in \delta(\mathcal{E})\}$ is itself a Dynkin system including $\mathcal{E}$ and consequently $\delta(\mathcal{E})$. Hence $\delta(\mathcal{E})$ is closed under **intersection**. On account of $A \cup B = X \setminus ((X \setminus A) \cap (X \setminus B))$, $A \setminus B = A \cap (X \setminus B)$ resp. 1.5.3 it is a σ-algebra, i.e. $\sigma(\mathcal{E}) \subset \delta(\mathcal{E})$ and since every σ-algebra is a Dynkin system we have $\sigma(\mathcal{E}) = \delta(\mathcal{E})$.

2 Pre-measures

2.1 Definition: The inclusion of big sets like $X = \mathbb{C}$ into the **domain** of a measure makes it necessary to include the corresponding value ∞ into its **range**. Since we expect integrals of vanishing functions on sets of infinite measure to have the value zero we define $\infty \cdot 0 := 0 \cdot \infty := 0$. The corresponding extended ranges are denoted as $\mathbb{R} \cup \{\infty\} = \overline{\mathbb{R}}$ resp. $\mathbb{C} \cup \{\infty\} = \overline{\mathbb{C}}$ resp. $[0; \infty[\cup \{\infty\} = [0; \infty]$. A set function $\mu : \mathcal{A} \to [0; \infty]$ on an **algebra** $\mathcal{A} \subset P(X)$ is **finitely additive**, iff $\mu(A \mathring{\cup} B) = \mu(A) + \mu(B)$ for **disjoint** $A, B \in \mathcal{A}$. In the general case with $A \cap B \in A$ follows the **subadditivity** $\mu(A \cup B) \leq \mu(A) + \mu(B)$. If there is an $A \in \mathcal{A}$ with $\mu(A) < \infty$ we have $\mu(\emptyset) = \mu(A \cup \emptyset) - \mu(A) = 0$. Also μ is **monotone**: For $A \subset B$ and $\mu(A) < \infty$ on account of $A \setminus B \in \mathcal{A}$ and $B = A \cup B \setminus A$ we have $\mu(B \setminus A) = \mu(B) - \mu(A)$ and particularly $\mu(A) < \mu(B)$. Note that $\mu(B) = \infty \Rightarrow \mu(B \setminus A) = \infty$ if $\mu(A) < \infty$. In the case of **σ-additivity** with $\mu\left(\mathring{\bigcup}_{n\in\mathbb{N}} A_n\right) = \sum_{n\in\mathbb{N}} \mu(A_n)$ for pairwise disjoint $(A_n)_{n\in\mathbb{N}} \subset \mathcal{A}$ it is a **pre-measure**. The **supremum property** (cf. [9, 14.12]) of the **real numbers** permits the extension of the **subadditivity** to **countable unions**: $\mu\left(\bigcup_{n\in\mathbb{N}} A_n\right) \leq \sum_{n\in\mathbb{N}} \mu(A_n)$.

2.2 Theorem: A **finite and finitely additive** set function $\mu : \mathcal{A} \to [0; \infty[$ on an **algebra** $\mathcal{A} \subset \mathcal{P}(X)$ is a **pre-measure** if one of the following equivalent conditions holds.

1. **σ-Additivi**ty: For a sequence $(A_n)_{n\in\mathbb{N}} \subset \mathcal{A}$ of **pairwise disjoint** measurable sets with $\mathring{\bigcup}_{n\in\mathbb{N}} A_n \in \mathcal{A}$ we have $\mu\left(\mathring{\bigcup}_{n\in\mathbb{N}} A_n\right) = \sum_{n\in\mathbb{N}} \mu(A_n)$.
2. **Continuity from below**: For an **increasing** sequence of measurable sets $A_0 \subset A_1 \subset ...$ with $\bigcup_{n\in\mathbb{N}} A_n \in \mathcal{A}$ we have $\lim\limits_{n\in\mathbb{N}} \mu(A_n) = \mu(\bigcup_{n\in\mathbb{N}} A_n)$.
3. **Continuity from above**: For a **decreasing** sequence of measurable sets $A_0 \supset A_1 \supset ...$ with $\bigcap_{n\in\mathbb{N}} A_n \in \mathcal{A}$ we have $\lim\limits_{n\in\mathbb{N}} \mu(A_n) = \mu\left(\bigcap_{n\in\mathbb{N}} A_n\right)$.
4. **$\emptyset$-Continuity**: For a **decreasing** sequence of measurable sets $A_0 \supset A_1 \supset ...$ with $\bigcap_{n\in\mathbb{N}} A_n = \emptyset$ we have $\lim\limits_{n\in\mathbb{N}} \mu(A_n) = 0$.

Proof:

1. $\Rightarrow$ 2. : With $A'_n := A_n \setminus A_{n-1}$ we obtain a **pairwise disjoint** family $(A'_n)_{n\in\mathbb{N}} \subset \mathcal{A}$ with $\mu(A_n) = \mu\left(\bigcup_{1\leq k\leq n} A'_k\right) = \sum_{1\leq k\leq n} \mu(A'_k)$ such that $\lim\limits_{n\in\mathbb{N}} \mu(A_n) = \sum_{n\in\mathbb{N}} \mu(A'_n) = \mu\left(\bigcup_{n\in\mathbb{N}} A'_n\right) = \mu(\bigcup_{n\in\mathbb{N}} A_n)$.

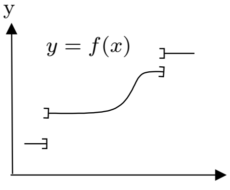
2. $\Rightarrow$ 3. : We apply 2. to the **increasing** sequence $\emptyset = A'_0 \subset A'_1 \subset ...$ of the **complements** $A'_n := A_0 \setminus A_n \in \mathcal{A}$ such that $\lim\limits_{n\in\mathbb{N}} \mu(A_n) = \lim\limits_{n\in\mathbb{N}} \mu(A_0 \setminus A'_n) = \lim\limits_{n\in\mathbb{N}} (\mu(A_0) - \mu(A'_n)) = \mu(A_0) - \lim\limits_{n\in\mathbb{N}} \mu(A'_n)$ $= \mu(A_0) - \mu\left(\bigcup_{n\in\mathbb{N}} A'_n\right) = \mu\left(A_0 \setminus \bigcup_{n\in\mathbb{N}} A'_n\right) = \mu\left(\bigcap_{n\in\mathbb{N}} A_0 \setminus A'_n\right) = \mu\left(\bigcap_{n\in\mathbb{N}} A_n\right)$.

3. $\Rightarrow$ 4. : Obvious.

4. $\Rightarrow$ 1. : With $A'_k := \mathring{\bigcup}_{n>k} A_n$ we obtain a **decreasing** sequence $(A'_k)_{k\in\mathbb{N}}$ with $\bigcap_{k\in\mathbb{N}} A'_k = \emptyset$ and $\mu(A'_k) < \infty$ such that according to 4. we have $0 = \lim\limits_{k\in\mathbb{N}} \mu(A'_k) = \mu\left(\mathring{\bigcup}_{n\in\mathbb{N}} A_n\right) - \lim\limits_{k\in\mathbb{N}} \mu\left(\mathring{\bigcup}_{n\leq k} A_n\right) = \mu\left(\mathring{\bigcup}_{n\in\mathbb{N}} A_n\right) - \sum_{n\in\mathbb{N}} \mu(A_n)$.

2.3 Examples:

1. The **Dirac measure** $\delta_x(A) := \begin{cases} 1, \, x \in A \\ 0, \, x \notin A \end{cases}$ for $A \subset X$ and $x \in X$ is a **pre-measure** on every **ring** on a set X.
2. The **measure** $\mu(A) := \begin{cases} 0 & \text{for countable } A \\ \infty & \text{else} \end{cases}$ on the **algebra** $\mathcal{P}(X)$ of a **discrete** space X according to 1.2.



lower semicontinuous
nondecreasing

2.4 Lebesgue-Borel-Stieltjes pre-measure: The set function $\lambda_f : \mathcal{I} \to [0; \infty[$ defined by $\lambda_f([a;b[) := f(b) - f(a)$ for any **nondecreasing** and **lower semicontinuous** (cf. [9, 3.3]) function $f : \mathbb{R} \to \mathbb{R}$ on the **right-open intervals** $\mathcal{I} = \{[a;b[\colon a \leq b \in \mathbb{R}\}$ can be extended to the **algebra** $\mathcal{F} = \left\{ \overset{\circ}{\bigcup\limits_{0 \leq k \leq m}} I_k : I_k \in \mathcal{I}, m \in \mathbb{N} \right\}$ of the **one-dimensional figures** by $\lambda_f \left(\overset{\circ}{\bigcup\limits_{0 \leq k \leq m}} I_k \right) := \sum\limits_{k=1}^{m} \lambda_f(I_k)$. For every **decreasing** sequence of figures $F_0 \supset F_1 \supset ...$ with $F_n = \overset{\circ}{\bigcup}_{0 \leq k_n \leq m_n} [a_{k_n}; b_{k_n}[\, \in \mathcal{F}$ and $\bigcap_{n \in \mathbb{N}} F_n = \emptyset$ the decreasing character implies that $\forall n > m \; \forall 0 \leq k_n \leq l_n \; \exists 0 \leq k_m \leq l_m$ with $[a_{k_n}; b_{k_n}[\, \subset [a_{k_m}; b_{k_m}[$. The condition $\bigcap_{n \in \mathbb{N}} F_n = \emptyset$ entails that $\forall k \in \mathbb{N} \; \exists n \in \mathbb{N}$ with $[a_{k_n}; b_{k_n}[\, = \emptyset$ since for any decreasing sequence $([a_n; b_n[)_{n \in \mathbb{N}}$ of real intervals due to the **supremum property** [9, 14.12] of the real numbers we have limits $a = \sup\limits_{n \to \infty} a_n$ resp. $b = \inf\limits_{n \to \infty} b_n$ and consequently $[a;b[\, \subset \bigcap\limits_{n \in \mathbb{N}} [a_n; b_n[$. Hence $\lim\limits_{n \to \infty} \lambda_f(F_n) = 0$, i.e. λ_f is a **pre-measure** on $\mathcal{F}$ due to its **$\emptyset$-continuity** 2.2.4 This is the **Lebesgue-Borel-Stieltjes pre-measure** resp. **Lebesgue-Borel pre-measure** λ for the identity $f(x) = x$ which together with the **euclidean metric** provides the foundation of analysis.

3 Measures

3.1 Definition: A pre-measure μ on a **σ-algebra $\mathcal{A}$** is a **measure** and $(X; \mathcal{A}; \mu)$ is a **measure space**. **Probability measures** have the range $[0;1]$ and in that case $(X; \mathcal{A}; \mu)$ is a **probability space**.

3.2 Outer measures: A set function $\tilde{\mu} : P(X) \to [0; \infty]$ is an **outer measure** iff for all $A, B, A_n \in \mathcal{A}, n \in \mathbb{N}$ the following properties hold:

1. $\tilde{\mu}(\emptyset) = 0$
2. $A \subset B \Rightarrow \tilde{\mu}(A) \leq \tilde{\mu}(B)$
3. $\tilde{\mu}\left(\bigcup_{n \in \mathbb{N}} A_n\right) \leq \sum_{n \in \mathbb{N}} \tilde{\mu}(A_n)$

A set $A \subset X$ is **$\tilde{\mu}$-measurable** iff for every $Q \subset X$ we have

4. $\tilde{\mu}(Q) = \tilde{\mu}(Q \cap A) + \tilde{\mu}(Q \setminus A)$.

3.3 Carathéodory's theorem For an outer measure $\tilde{\mu}$ on a set X the system $\mathcal{A}$ of all $\tilde{\mu}$-measurable sets $A \subset X$ is a σ-algebra and the restriction $\tilde{\mu}|_{\mathcal{A}}$ is a measure.

Proof: Obviously we have $\emptyset, X \in \mathcal{A}$ and on account of 3.2.4 every $A \in \mathcal{A}$ has a measurable **complement** $X \setminus A \in \mathcal{A}$. For $A, B \in \mathcal{A}$ the **union** $A \cup B \in \mathcal{A}$ is measurable too since by applying 3.2.4 successively we obtain first an equation (I): $\tilde{\mu}(Q) = \tilde{\mu}(Q \cap A) + \tilde{\mu}(Q \setminus A) = \tilde{\mu}(Q \cap A \cap B) + \tilde{\mu}(Q \cap A \setminus B) + \tilde{\mu}(Q \setminus A \cap B) + \tilde{\mu}(Q \setminus A \setminus B)$ and if we substitute Q with $Q \cap (A \cup B)$ in (I) we arrive at another equatiuon (II): $\tilde{\mu}(Q \cap (A \cup B)) = \tilde{\mu}(Q \cap A \cap B) + \tilde{\mu}(Q \cap A \setminus B) + \tilde{\mu}(Q \setminus A \cap B)$. We can substitute the first three terms in (I) by (II) and hence obtain the measurability of the **union**: $\tilde{\mu}(Q) = \tilde{\mu}(Q \cap (A \cup B)) + \tilde{\mu}(Q \setminus (A \cup B))$. Thuas and because of $A \cap B = X \setminus ((X \setminus A) \cup (X \setminus B))$ and $A \setminus B = A \cap (X \setminus B)$ the family $\mathcal{A}$ is an **algebra**.

For a sequence $(A_n)_{n\in\mathbb{N}} \subset \mathcal{A}$ of **pairwise disjoint measurable** sets $A := \bigcup_{n\in\mathbb{N}} A_n$ equation (II) yields $\tilde{\mu}\left(Q \cap (A_0 \cup A_1)\right) = \tilde{\mu}\left(Q \cap A_0\right) + \tilde{\mu}\left(Q \cap A_1\right)$ resp. by **induction** $\tilde{\mu}\left(Q \cap \bigcup\limits_{k=0}^{n} A_k\right) = \sum\limits_{k=0}^{n} \tilde{\mu}\left(Q \cap A_k\right)$. On account of $\bigcup\limits_{k=0}^{n} A_k \in \mathcal{A}$ and 3.2.2 we conclude that (III): $\tilde{\mu}(Q) = \tilde{\mu}\left(Q \cap \bigcup\limits_{k=0}^{n} A_k\right) + \tilde{\mu}\left(Q \setminus \bigcup\limits_{k=0}^{n} A_k\right)$ $\geq \sum\limits_{k=0}^{n} \tilde{\mu}\left(Q \cap A_k\right) + \tilde{\mu}\left(Q \setminus A\right)$. Since this estimate holds for all $n \in \mathbb{N}$ it extends to $n \to \infty$ such that by 3.2.3 we arrive at the measurability criterion 3.2.4 for A. Due to 1.5.3 the family $\mathcal{A}$ is a **Dynkin system** which is **closed under intersection** and in accordance with 1.6 it is a **$\boldsymbol{\sigma}$-algebra**. If in (III) we substitute $Q = A$ and observe 3.2.3 we obtain the **$\boldsymbol{\sigma}$-additivity** of $\tilde{\mu}$ on $\mathcal{A}$, i.e. $\tilde{\mu}|_{\mathcal{A}}$ is a **measure**.

3.4 Uniqueness theorem Two measures μ_1 and μ_2 on a σ-Algebra $\sigma(\mathcal{E})$ induced by a **π-basis** $\mathcal{E} \subset \mathcal{P}(X)$ are identical iff they coincide on $\mathcal{E}$ and are **σ-finite** on $\mathcal{E}$, i.e. $\exists (E_n)_{n\in\mathbb{N}} \subset \mathcal{E}$ with $\bigcup_{n\in\mathbb{N}} E_n = X$ and $\mu_1(E_n) = \mu_2(E_n) < \infty$ for all $n \in \mathbb{N}$.

Proof: For $E \in \mathcal{E}$ with $\mu_1(E) = \mu_2(E) < \infty$ the family $\mathcal{D}_E := \{D \in \sigma(\mathcal{E}) : \mu_1(E \cap D) = \mu_2(E \cap D)\}$ is a **Dynkin system** since $\emptyset \in \mathcal{D}_E$ and for every $D \in \mathcal{D}_E$ on account of $\mu_1(E \cap X \setminus D) = \mu_1(E) - \mu_1(E \cap D) = \mu_2(E) - \mu_2(E \cap D) = \mu_2(E \cap X \setminus D)$ we also have $X \setminus D \in \mathcal{D}_E$. Criterion 1.5.3 follows from the σ-additivity of μ_1 and μ_2. Since $\mathcal{E}$ is **closed under intersection** we have $\mathcal{E} \subset \mathcal{D}_E$ and since $\mathcal{D}_E$ is a Dynkin system 1.6 entails $\sigma(\mathcal{E}) = \delta(\mathcal{E}) \subset \mathcal{D}_E \subset \sigma(\mathcal{E})$, i.e.$\mathcal{D}_E = \sigma(\mathcal{E})$ resp. $\mu_1(E \cap A) = \mu_2(E \cap A)$ for all $E \in \mathcal{E}$ and $A \in \sigma(\mathcal{E})$.

As in the proof of 2.2.2 we define a sequence of pairwise disjoint sets $E'_n := E_n \setminus \bigcup_{1\leq k<n} E_k \in \sigma(\mathcal{E})$ with $\mathring{\bigcup}_{n\in\mathbb{N}} E'_n = X$ such that for $A \in \sigma(\mathcal{E})$ we have $E'_n \cap A \in \sigma(\mathcal{E})$, hence $\mu_1(E_n \cap E'_n \cap A) = \mu_2(E_n \cap E'_n \cap A)$ and the σ-additivity of μ_1 resp. μ_2 yields $\mu_1(A) = \mu_2(A)$.

3.5 Hahn's extension theorem: Every **σ-finite pre-measure** μ on an **algebra** $\mathcal{A}$ can be extented **in a unique way** to a **measure** μ on $\sigma(\mathcal{A})$.

Proof: For every set $Q \subset X$ let $\mathcal{U}(Q) \neq \emptyset$ be the family of sequences $(A_n)_{n\in\mathbb{N}} \subset \mathcal{A}$ with $Q \subset \bigcup_{n\in\mathbb{N}} A_n$. Then $\tilde{\mu}(Q) := \inf\left\{\sum_{n\in\mathbb{N}} \mu(A_n) : (A_n)_{n\in\mathbb{N}} \in \mathcal{U}(Q)\right\}$ in case of $\mathcal{U}(Q) \neq \emptyset$ and $\tilde{\mu}(Q) := \infty$ else is an **outer measure** since obviously we have $\tilde{\mu}(\emptyset) = 0$ and for $P \subset Q$ follows $\mathcal{U}(P) \supset \mathcal{U}(Q)$ and hence $\tilde{\mu}(P) \leq \tilde{\mu}(Q)$, particularly $\tilde{\mu}(Q) \geq 0 \ \forall Q \subset X$. For every sequence $(Q_n)_{n\in\mathbb{N}} \subset \mathcal{P}(X)$, $\epsilon > 0$ and $n \in \mathbb{N}$ there is a sequence $(A_{nm})_{m\in\mathbb{N}} \subset \mathcal{U}(Q_n) \neq \emptyset$ with $\sum_{m\in\mathbb{N}} \mu(A_{nm}) < \tilde{\mu}(Q_n) + \epsilon \cdot 2^{-n-1}$ and since $(A_{nm})_{n,m\in\mathbb{N}} \subset \mathcal{U}\left(\bigcup_{n\in\mathbb{N}} Q_n\right)$ it follows that $\tilde{\mu}\left(\bigcup_{n\in\mathbb{N}} Q_n\right) \leq \sum_{n,m\in\mathbb{N}} \mu(A_{nm}) < \sum_{n\in\mathbb{N}} \tilde{\mu}(Q_m) + \epsilon$. Since $\epsilon > 0$ is arbitrary condition 3.2.3 is satisfied.

The algebra $\mathcal{A}$ is **$\tilde{\mu}$-measurable** since for every $A \in \mathcal{A}$ and $Q \subset X$ with $(A_n)_{n\in\mathbb{N}} \subset \mathcal{U}(Q)$ we have $(A_n \cap A)_{n\in\mathbb{N}} \subset \mathcal{U}(Q \cap A)$ resp. $(A_n \setminus A)_{n\in\mathbb{N}} \subset \mathcal{U}(Q \setminus A)$ and since $\mu(A_n) = \mu(A_n \cap A) + \mu(A_n \setminus A)$ we obtain $\tilde{\mu}(Q) \geq \tilde{\mu}(Q \cap A) + \tilde{\mu}(Q \setminus A)$ and hence equality on account of 3.2.3. The assertion then follows from 3.3 and 3.4.

3.6 Approximation property: Every set $Q \in \sigma(\mathcal{A})$ with **finite measure** $\mu(Q) < \infty$ on a **σ-algebra** $\sigma(\mathcal{A})$ induced by an **algebra** $\mathcal{A}$ can be approximated **in measure** by a sequence $(C_n)_{n\in\mathbb{N}} \subset \mathcal{A}$ such that $\lim\limits_{n\to\infty} \mu(Q \Delta C_n) = 0$ and particularly $\lim\limits_{n\to\infty} \mu(C_n) = \mu(Q)$.

Proof: As in the proof for 3.5 and since $\mu(Q) < \infty$ for every $\epsilon > 0$ we can find a sequence of w.l.o.g. (cf. proof of 2.2.2) pairwise disjoint sets $(A_k)_{k\in\mathbb{N}} \subset \mathcal{A}$ with $Q \subset \mathring{\bigcup}_{k\in\mathbb{N}} A_k$ and $\mu\left(\mathring{\bigcup}_{k\in\mathbb{N}} A_k\right) - \mu(Q) = \sum_{k\in\mathbb{N}} \mu(A_k) - \mu(Q) < \frac{\epsilon}{2}$. The unions $C_n := \mathring{\bigcup}_{0\leq k\leq n} A_k$ already constitute the desired sequence since owing to $\mu\left(\mathring{\bigcup}_{k\in\mathbb{N}} A_k\right) < \infty$ we can apply 2.2.2 such that there is an $n_0 \in \mathbb{N}$ with $\mu\left(\mathring{\bigcup}_{n\in\mathbb{N}} A_n\right) - \mu(C_{n_0})$ $< \frac{\epsilon}{2}$ and hence $\mu(Q \Delta C_{n_0}) = \mu(Q \setminus C_{n_0}) + \mu(C_{n_0} \setminus Q) \leq \mu\left(\mathring{\bigcup}_{n\in\mathbb{N}} A_n \setminus C_{n_0}\right) + \mu\left(\mathring{\bigcup}_{n\in\mathbb{N}} A_n \setminus Q\right) < \frac{\epsilon}{2} + \frac{\epsilon}{2}$ $= \epsilon$. The second assertion follows from $\mu(C_n) = \mu(Q) + \mu(C_n \setminus C)$ and $\mu(C_n \setminus C) \leq \mu(Q \Delta C_n)$.

3.7 Lebesgue-Borel-Stieltjes measure According to 3.5 we can extend the **Lebesgue-Borel-Stieltjes pre-measure** λ_f from 2.4 on the **algebra of figures** $\mathcal{F}$ from 1.4 to a **Lebesgue-Borel-Stieltjes measure** λ_f on the **Borel σ-algebra** $\mathcal{B} = \sigma(\mathcal{I}) = \sigma(\mathcal{F})$. In section 7 the **measure space** $(\mathbb{R}; \mathcal{B}; \lambda)$ will be extended to finite products of $\mathbb{R}$ and particularly $\mathbb{C}$ whereas the range of λ will be extended to $\mathbb{C}$ in section 8. Owing to 3.4 and since λ_f is **σ-finite** as well as **continuous** such that

$\lambda_f\left([a;b[\right) = \lambda_f\left([a;b]\right) = \lambda_f\left(]a;b]\right) = \lambda_f\left(]a;b[\right) = f\left(b\right) - f\left(a\right)$ for $a \leq b \in \mathbb{R}$ on the **right-open intervals** $\mathcal{I}$ being closed under intersection the Lebesgue-Borel measure is **uniquely determined** by its definition on $\mathcal{I}$. Thus every countable union of single points is a λ_f**-null set**, in particular the **rational numbers**: $\lambda_f\left(\mathbb{Q}\right) = 0$. The **Cantor set** $T := g\left[\{0;2\}^{\mathbb{N}}\right]$ with $g(x) = \sum_{n\geq 1}\frac{x_n}{3^n}$ for any sequence $x = \left(x_n\right)_{n\geq 1}$ with $x_n \in \{0;2\}$ (cf. [9, 2.10] is a λ-null set since $T = \bigcap_{n\in\mathbb{N}} T_n$ with $T_0 = [0;1]$ and T_{n+1} is a union of 2^{n+1} disjoint and closed intervals with **length** resp. **measure** $\frac{1}{3^{n+1}}$ obtained by removing the middle third from the 2^n closed intervals T_n with length $\frac{1}{3^n}$ such that $\lambda\left(T_n\right) = \frac{2^n}{3^n}$ and $\lambda\left(T\right) = \lim_{n\in\mathbb{N}}\lambda\left(T_n\right) = 0$ due to the **continuity from above** (2.2.3). The G_δ-set $U = \bigcap_{n\geq 1} U_n$ with **dense open** sets $U_n = \bigcup_{i\geq 1} B_{n^{-1}.2^{-i-1}}\left(q_i\right)$ based on the enumeration $\mathbb{Q} = \left(q_i\right)_{i\geq 1}$ includes $\mathbb{Q}$ and hence is dense in $\mathbb{R}$. Again due to 2.2.3 and since $\lambda\left(U_n\right) \leq \frac{1}{n}$ it also is a λ-null set: $\lambda\left(U\right) = 0$. The **complements** $\mathbb{R} \setminus U_n$ are **closed** and **nowhere dense** in $\mathbb{R}$ but with measure $\lambda\left(\mathbb{R} \setminus U_n\right) = \infty$ and $\mathbb{R} \setminus U$ is an example for a set of **first category** with measure $\lambda\left(\mathbb{R} \setminus U\right) = \infty$. (cf. [9, 16.1])

3.8 Complete measure: A measure μ is **complete** iff every **subset** of a μ**-null set** is **measurable**.

1. A σ-algebra $\mathcal{A}$ can be **completed** to a σ-Algebra $\mathcal{A}_0 = \{A \cup M : A \in \mathcal{A} \wedge M \subset N \in \mathcal{A} : \mu(N) = 0\}$ by simply adding the requested subset of null sets to the given measurable sets: For $A, B \in \mathcal{A}$ resp. $M_A \subset N_A$, $M_B \subset N_B$ and $\mu\left(N_A\right) = \mu\left(N_B\right) = 0$ we have $\left(A \cup M_A\right) \setminus \left(B \cup M_B\right) = A \setminus \left(B \cup M_B\right) \cup M_A \setminus \left(B \cup M_B\right) = \left(A \setminus B\right) \cap \left(A \setminus N_B\right) \cup \left(N_B \setminus M_B\right) \cup M_A \setminus \left(B \cup M_B\right) \in \mathcal{A}_0$ since $\left(A \setminus B\right) \cap \left(A \setminus N_B\right) \in \mathcal{A}$ and $\left(N_B \setminus M_B\right) \cup M_A \setminus \left(B \cup M_B\right) \subset N_A \cup N_B$ with $\mu\left(N_A \cup N_B\right) = 0$. The σ-additivity is obvious.

2. A set E is $\mathcal{A}_0$-measurable iff there are $A, B \in \mathcal{A}$ with $A \subset E \subset B$ and $\mu\left(B \setminus A\right) = 0$: One the one hand for any $E = A \cup M$ with $M \subset N \in \mathcal{A}$ and $\mu\left(N\right) = 0$ the measurable sets A and $B := A \cup N$ satisfy the criterion. On the other hand for any E and measurable A, B according to the criterion we have $E = A \cup \left(B \setminus A \cap E\right)$ with $B \setminus A \cap E \subset B \setminus A$ and hence $E \in \mathcal{A}_0$.

3. The corresponding **extension** $\mu_0 \supset \mu$ with $\mu_0(A \cup N) := \mu(A)$ for $A \in \mathcal{A}$ and $N \subset M : \mu(M) = 0$ obviously is a complete measure. Thus the **Lebesgue-Borel measure** λ on the σ-algebra $\mathcal{B}$ of the **Borel sets** is extended to the **Lebesgue measure** λ_0 on the completed σ-algebra $\mathcal{B}_0$ of the **Lebesgue sets**. In the following section the index is usually omitted such that the complete Lebesgue space is still denoted as $(X; \mathcal{B}; \lambda)$.

3.9 Almost everywhere existing properties: In **probability theory** the completion is seldom used since it is not generated by the **open** sets any more and hence restricts the choice of possible **measures** resp. **distributions** without granting any gain in information. In **analysis** it is widely adopted though not always necessarily so since a σ-algebra is a family e.g. larger by far than the topology on $\mathbb{R}$ such that it is not a trivial exercise to find non measurable sets at all. In any case we speak of a property $E(x)$ being satisfied μ**-almost everywhere** (μ-a.e.) iff it is satsfied everywhere with the exception of μ**-null sets**, i.e. iff $\mu\left(\neg E\right) = 0$.

3.10 Non measurable sets (Vitali): There is a set $K \subset \mathbb{R}$ which is **not Lebesgue measurable**.

Proof: The **equivalence relation** defined by $xRY \Leftrightarrow x - y \in \mathbb{Q}$ generates a disjoint cover of $\mathbb{R}$ by equivalence classes with the class $\overline{0} = \mathbb{Q}$ and all other classes represented by irrational numbers. Since $\mathbb{Q}$ is dense in $\mathbb{R}$ every class has representants $x \in [0;1]$ and the **axiom of choice** [11, 14.2.1] permits us to choose **exactly one of those for every equivalence class** and thus define a set $K \subset [0;1]$ such that we obtain a **disjoint and countable cover** $\mathbb{R} = \overset{\circ}{\bigcup}_{q\in\mathbb{Q}}\left(q + K\right)$ which due to the σ**-additivity** and the **translation invariance** must satisfy $\infty = \lambda\left(\mathbb{R}\right) = \sum_{q\in\mathbb{Q}}\lambda\left(K\right)$ and hence $\lambda\left(K\right) > 0$. On the other hand we have $\overset{\circ}{\bigcup}_{q\in\mathbb{Q}\cap[0;1]}\left(q + K\right) \subset [0;2]$ and due to the **monotonicity** of the measure $\sum_{q\in\mathbb{Q}}\lambda\left(K\right) \leq \lambda\left([0;2]\right) = 2$ hence $\lambda\left(K\right) = 0$. From this contradiction we must infer that K is **not measurable**.

4 Measurable functions

4.1 Definitions: A mapping $f : (X;\mathcal{A}) \to (Y;\mathcal{B})$ between measurable spaces is **measurable** iff every inverse image $f^{-1}(B)$ of a measurable set $B \in \mathcal{B}$ is again measurable in $(X;\mathcal{A})$, i.e. $f^{-1}(B) \in \mathcal{A}$. Since all necessary set operations transfer to inverse images (cf. [11, 9.2]) it is sufficient that the inverse images of **basis** sets are measurable in X (cf. [9, 3.1]). In analysis the usual basis is the topology $\mathcal{O}$ on Y and the function is **Borel measurable** iff it is measurable with reference to $\mathcal{B} = \sigma(\mathcal{O})$. Hence a function $f : (X;\mathcal{A}) \to (Y;d)$ into a **metric space** is Borel measurable iff $f^{-1}[\mathcal{B}_\epsilon(y)] \in \mathcal{A}$ for every $\epsilon > 0$ and $y \in Y$.

4.2 Real valued Borel measurable functions: According to 2.1 a function $f : X \to \mathbb{R}$ is measurable iff the sets $\{f \geq a\} := f^{-1}[[a;\infty[]$ or the analogously defined $\{f > a\}$, $\{f \leq a\}$ resp. $\{f < a\}$ are measurable in X. In particular for a Borel measurable $f : X \to \mathbb{R}$ the **positive part** $f^+ := \max\{f;0\}$, the **negative part** $f^- := \min\{f;0\}$ are Borel measurable. Since $\mathbb{Q}$ is countable and dense in $\mathbb{R}$ the sets $\{f > g\} = \bigcup_{a\in\mathbb{Q}} (\{f > a\} \cap \{a > g\})$ and $\{f \geq g\} = X \setminus \{f < g\}$ are measurable. Hence the **maximum** $\max\{f;g\}$ and the **minimum** $\min\{f;g\}$ are Borel measurable for any for measurable $f, g : X \to \mathbb{R}$. In the expression for the measure μ of the set of all $x \in X$ for which $A(f(x))$ is true we will often omit not only the argument but also the curly brackets: $\mu(A(f)) = \mu(\{A(f)\}) = \mu(\{x \in X : A(f(x))\})$ as e.g. in $\mu(|f| < \epsilon) = \mu(\{|f| < \epsilon\})$.

4.3 Image of a measure space: The image $f(\mathcal{A}) := \{B \subset Y : f^{-1}[B] \in \mathcal{A}\}$ of a σ-algebra $\mathcal{A}$ on X under $f : X \to Y$ is a σ-algebra on Y and the largest σ-algebra such that f is measurable. The **image of the measure** $f \circ \mu : f(\mathcal{A}) \to [0;\infty]$ with $(f \circ \mu)(B) := \mu(f^{-1}[B])$ resp. $(f \circ \mu)(f[B]) := \mu(B)$ is a measure on $f(\mathcal{A})$ and **transitive** with regard to **composition**: $g \circ f \circ \mu : g \circ f(\mathcal{A}) \to [0;\infty]$ obviously is again a measure. E.g. the **Lebesgue measure** λ is **invariant** under the **translation** $T_c(x) = x + c$ with $(T_c \circ f)([a;b[) = \lambda(T_c^{-1}[[a;b[]) = \lambda([a-c;b-c[) = \lambda([a;b[)$ but not under **dilation** $g(x) = mx$ since $(g \circ \lambda)([a;b[) = \lambda(g^{-1}[[a;b[]) = \lambda\left(\left[\frac{a}{m};\frac{b}{m}\right[\right) = \frac{1}{m}\lambda([a;b[)$.

4.4 Inverse image of a measurable space: The **inverse image** $\sigma(f^{-1}(\mathcal{E})) = f^{-1}(\sigma(\mathcal{E}))$ of the σ-algebra $\sigma(\mathcal{E})$ on Y induced by $\mathcal{E} \subset \mathcal{P}(Y)$ under $f : X \to Y$ is the **smallest** σ-algebra such that f is measurable. The inclusion $\subset$ holds since $f^{-1}(\sigma(\mathcal{E}))$ is a σ-algebra containing $f^{-1}(\mathcal{E})$. The inclusion $\supset$ follows from 4.3 since $f(\sigma(f^{-1}(\mathcal{E})))$ is a σ-algebra on Y including $\mathcal{E}$ and hence $\sigma(\mathcal{E})$.

4.5 Continuous functions: On account of 4.4 a function $f : (X;\mathcal{A}) \to (Y;\sigma(\mathcal{O}_Y))$ into a **topological space** $(Y;\mathcal{O}_Y)$ is **Borel measurable** iff the inverse image of every **open set** in measurable in $(X;\mathcal{A})$: $f^{-1}(\mathcal{O}_Y) \subset \mathcal{A} \Rightarrow f^{-1}(\sigma(\mathcal{O}_Y)) = \sigma(f^{-1}(\mathcal{O}_Y)) \subset \mathcal{A}$. In the case of $\mathcal{A} = \sigma(\mathcal{O}_X)$ also being induced by a topology $\mathcal{O}_X$ on X every **continuous function** is **Borel measurable**. A real function $f : X \to \mathbb{R}$ on a topological space $(X;\mathcal{O})$ is lower resp. upper **semicontinuous** iff $f^{-1}[]a;\infty[] \in \mathcal{O}$ resp. $f^{-1}[]-\infty;b[] \in \mathcal{O}$ for $\forall a, b \in \mathbb{R}$. (cf. [9, 3.3]) According to 1.4 resp. 4.1 these functions are again Borel measurable.

4.6 Compositions: The **composition** $h = g \circ f : X \to Z$ is measurable iff $f : X \to Y$ and $g : Y \to Z$ are measurable. Due to [9, 3.1; 4.2.3 and 10.7]

- the **projections** $\pi_j : \prod_{i\in I} X_i \to X_j$ on a **product space** $\prod_{i\in I} X_i$,
- the **metric** $d : X^2 \to [0;\infty[$ on a **metric space** $(X;d)$,
- the **norm** $|| : X \to [0;\infty[$, the **multiple** $\alpha\cdot : X \to X$ for fixed $\alpha \in \mathbb{C}$ and the **addition** $+ : X^2 \to X$ on a **Banach space** $(X;||)$ (cf. [9, 21.9]) ,
- the **multiplication** $\cdot : X^2 \to X$ on a **Banach algebra** $(X;||)$ (cf. [9, 18.9]) and
- the **multiple** $x \mapsto \alpha\cdot x$ resp. the **powers** $x \mapsto x^\alpha$ for $\alpha \in \mathbb{C}$ as well as in particular the reciprocal $x \mapsto \frac{1}{x}$ on a **field** like $\mathbb{R}$ or $\mathbb{C}$

are **continuous** and hence **Borel measurable**. Hence for Borel measurable $f, g : X \to \mathbb{C}$ the **real part** $\mathrm{Re}f$, **imaginary part** $\mathrm{Im}f$ and **absolute value** $|f|$ are Borel measurable mappings $X \to \mathbb{R}$; likewise the **complex conjugate** $\overline{f}$ as well as $\alpha \cdot f$, f^α, $\frac{1}{f}$, $f + g$ and $f \cdot g$ are Borel measurable mappings $X \to \mathbb{C}$.

4.7 Measurable functions into product spaces: A Borel measurable function $f : (X;\mathcal{A}) \to (\prod_{i\in I} Y_i; \sigma(\bigotimes_{i\in I}\mathcal{O}_i))$ has Borel measurable components $f_i := \pi_i \circ f$. Since the **cylinder sets** $\bigcap_{i\in J}\pi_i^{-1}[O_i]$ with O_i open in Y_i and **finite** $J \subset I$ form a basis for the **product topology** $\bigotimes_{i\in I}\mathcal{O}_i$ (cf. [9, 4.2]) the converse is true if this basis is **countable** (i.e. the product topology is **first countable**, cf. [9, 2.6]) such that the inverse image of every open set in $\prod_{i\in I} Y_i$ is the countable union of inverse images of cylinder sets and hence contained in the σ-algebra $\mathcal{A}$ on X. This condition is satisfied for every **finite product** $\prod\limits_{i=1}^{n} Y_i$ of first countable components Y_i and in particular $\mathbb{C}^n$. Note that the countability condition is not needed for the corresponding statement on **continuous** functions since a **topology** $\mathcal{O}$ on X includes **arbitrary** unions of cylinder sets. Hence $f : X \to \mathbb{C}^n$ is Borel measurable iff every component f_i is Borel measurable.

4.8 Vector spaces of measurable functions: The product Y^2 of two Banach spaces $(Y;||)$ is first countable if Y itself is the **finite product** of first countable spaces, e.g. $\mathbb{C}^n$ or **separable**, e.g. the space $C_c^\infty(\mathbb{C})$ of infinitely derivable functions $f : \mathbb{C} \to \mathbb{C}$ with compact support. In these cases the ordered pair $(f,g) : X \to Y^2$ is Borel measurable if each $f, g : X \to Y$ is Borel measurable and so is their sum $f+g$ such that the Borel measurable functions $f : X \to Y$ into **finite dimensional** or **separable** Banach spaces Y themselves form a **vector space**.

4.9 Pointwise limits of measurable functions: The **pointwise** limit $f = \lim\limits_{n\to\infty} f_n$ of a sequence $(f_n)_{n\geq 1}$ of Borel measurable functions $f_n : X \to Y$ from a **measurable space** $(X;\mathcal{A})$ into a **metric space** (Y,d) is again Borel measurable.

Proof:

For any **open** $U \subset Y$ and $f(x) \in U$ there is an $m \in \mathbb{N}$ with $f_k(x) \in U$ for all $k \geq m$ and hence $f^{-1}[U] \subset \bigcup\limits_{m=1}^{\infty}\bigcap\limits_{k=m}^{\infty} f_k^{-1}[U] \subset \bigcap\limits_{m=1}^{\infty}\bigcup\limits_{k=m}^{\infty} f_k^{-1}[U]$. On the other hand every **closed** $A \subset Y$ containing infinitely many $f_k(x)$ must contain the limit $f(x)$, i.e. $\bigcap\limits_{m=1}^{\infty}\bigcup\limits_{k=m}^{\infty} f_k^{-1}[A] \subset f^{-1}[A]$. For the open sets $V_n = \left\{x \in U : d(x;X\setminus U) < \frac{1}{n}\right\}$ we have $U = \bigcup\limits_{n=1}^{\infty} V_n = \bigcup\limits_{n=1}^{\infty} \overline{V_n}$ and hence $\bigcup\limits_{n=1}^{\infty}\bigcap\limits_{m=1}^{\infty}\bigcup\limits_{k=m}^{\infty} f_k^{-1}\left[\overline{V_n}\right] \subset \bigcup\limits_{n=1}^{\infty} f^{-1}\left[\overline{V_n}\right] = f^{-1}[U] = \bigcup\limits_{n=1}^{\infty} f^{-1}[V_n] \subset \bigcup\limits_{n=1}^{\infty}\bigcap\limits_{m=1}^{\infty}\bigcup\limits_{k=m}^{\infty} f_k^{-1}[V_n]$ whence follows equality since $V_n \subset \overline{V_n}$.

4.10 Convergence in measure and μ-almost everywhere: A sequence $(f_n)_{n\geq 1}$ of Borel measurable functions $f_n : X \to Y$ from a **measure space** $(X;\mathcal{A};\mu)$ into a **Banach space** $(Y,||)$converges to a Borel measurable $f : X \to Y$:

1. **μ-almost everywhere** (**μ-a.e.**) iff one of the following equivalent conditions is satisfied:
    a) $\mu\left(X \setminus \left\{\lim\limits_{n\to\infty}|f_n - f| = 0\right\}\right) = 0$

    b) $\lim\limits_{k\to\infty}\mu\left(\sup\limits_{n\geq k}|f_n - f| \geq \epsilon\right) = \lim\limits_{k\to\infty}\mu\left(\bigcup_{n\geq k}\{|f_n - f| \geq \epsilon\}\right) = 0$ for every $\epsilon > 0$
    $\overset{*}{\Rightarrow} \mu\left(\bigcap_{k\geq 1}\bigcup_{n\geq k}\{|f_n - f| \geq \epsilon\}\right) = 0$ for every $\epsilon > 0$

    c) $\lim\limits_{k\to\infty}\mu\left(\sup\limits_{n\geq k}|f_n - f| \geq \frac{1}{k}\right) = \lim\limits_{k\to\infty}\mu\left(\bigcup_{n\geq k}\left\{|f_n - f| \geq \frac{1}{k}\right\}\right) = 0$
    $\overset{*}{\Rightarrow} \mu\left(\bigcap_{k\geq 1}\bigcup_{n\geq k}\left\{|f_n - f| \geq \frac{1}{k}\right\}\right) = 0.$

2. **in measure** μ iff for every $A \in \mathcal{A}$ with $\mu(A) < \infty$ one of the following equivalent conditions is satisfied
    a) $\lim\limits_{n\to\infty}\mu|_A(|f_n - f| \geq \epsilon) = 0$ for every $\epsilon > 0 \Leftrightarrow$

    b) For every $k \in \mathbb{N}$ there is an $n_k \in \mathbb{N}$ such that $\mu|_A\left(|f_{n_k} - f| \geq 2^{-k}\right) < 2^{-k}$.

Notes:

1. The preceding definition is also known as **local convergence in measure** as opposed to the stronger **global convergence in measure** without the restriction to sets with **finite measure** $\mu(A) < \infty$. For an apriori **finite measure** with $\mu(X) < \infty$ the two definitions obviously coincide. In the case of a **probability measure** the convergence in measure is called **stochastic convergence**.
2. The inclusions $\overset{*}{\Rightarrow}$ become **equivalences** if we can presume the **continuity from above** 2.2.3, i.e. $\mu(X) < \infty$ or at least the existence of a $k \in \mathbb{N}$ such that $\mu\left(\bigcup_{n\geq k}\left\{|f_n - f| \geq \frac{1}{k}\right\}\right) < \infty$. Many of the subsequent convergence theorems also depend heavily on 2.2.3 and hence are restricted to **finite measure spaces** resp. to **local convergence in measure**. In partucular for the **Lebesgue measure** λ they **do not extend to global convergence**.
3. Both convergence criterions imply that the limit function f as well as **finally** (i.e. all except for a finite number) all f_n are **μ-a.e. finite.**

4.11 Lebesgue's convergence theorem: A sequence $(f_n)_{n\geq 1}$ of Borel measurable functions $f_n : X \to Y$ from a **measure space** $(X;\mathcal{A};\mu)$ into a **Banach space** $(Y,||)$ converging **μ-a.e.** to a Borel measurable function $f : X \to Y$ also converges **in measure** to f.

Proof: For every $A \in \mathcal{A}$ with $\mu(A) < \infty$ and $\epsilon > 0$ we have $\inf_{k\geq 1}\sup_{n\geq k}\mu|_A(\{|f_n - f| \geq \epsilon\}) \overset{2.2.2}{=} \inf_{k\geq 1}\mu|_A\left(\bigcup_{n\geq k}\{|f_n - f| \geq \epsilon\}\right) \overset{2.2.3}{=} \mu|_A\left(\bigcap_{k\geq 1}\bigcup_{n\geq k}\{|f_n - f| \geq \epsilon\}\right) = 0.$

Example: The **Lebesgue measure** λ is **not continuous from above**, e.g. $\lambda\left(\bigcap_{n\in\mathbb{N}}\mathbb{R}\setminus B_n(0)\right) = \lambda(\emptyset) = 0$ but $\inf_{n\in\mathbb{N}}\lambda(\mathbb{R}\setminus B_n(0)) = \infty$ since $\lambda(\mathbb{R}\setminus B_n(0)) = \infty$ for every $n \in \mathbb{N}$. Hence in the case of $f_n(x) = \frac{x^2}{n}$ we observe **pointwise convergence** and particularly **λ-a.e. convergence** as well as **compact convergence** to $f(x) = 0$ hence **local convergence in measure** but not **global convergence in measure** since $\lambda(|x| \geq \epsilon) = \lambda(|f_n - f| \geq \sqrt{n\epsilon}) = \infty$ for every $n \in \mathbb{N}$ and $\epsilon > 0$.

4.12 Borel-Cantelli lemma: For every sequence $(A_n)_{n\geq 1}$ of measurable sets $A_n \in \mathcal{A}$ on a **measure space** $(X;\mathcal{A};\mu)$ we have $\sum_{n\geq 1}\mu(A_n) < \infty \Rightarrow \mu\left(\bigcap_{k\geq 1}\bigcup_{n\geq k}A_n\right) = 0$ and in the case of a **probability measure** and **pairwise independent** A_n, i.e. $\mu(A_k \cap A_l) = \mu(A_k)\cdot\mu(A_l)$ for $k \neq l$ the **converse** is also true**:** $\sum_{n\geq 1}\mu(A_n) = \infty \Rightarrow \mu\left(X \setminus \bigcap_{k\geq 1}\bigcup_{n\geq k}A_n\right) = 0$.

Proof: In the first case for every $\epsilon > 0$ there is a $k_\epsilon \geq 1$ with $\sum_{n\geq k_\epsilon}\mu(A_n) < \epsilon$ such that $\mu\left(\bigcap_{k\geq 1}\bigcup_{n\geq k}A_n\right) \leq \mu\left(\bigcup_{n\geq k_\epsilon}A_n\right) \leq \sum_{n\geq k_\epsilon}\mu(A_n) < \epsilon$ and hence the assertion. In the second case with $\mu(X) = 1$ and the **continuity of the exponential function** we have $\mu\left(\bigcap_{k\geq 1}\bigcup_{n\geq k}A_n\right) = 1 - \mu\left(\bigcup_{k\geq 1}\bigcap_{n\geq k}X\setminus A_n\right) \overset{2.2.2}{=} 1 - \sup_{k\geq 1}\mu\left(\bigcap_{n\geq k}X\setminus A_i\right) \overset{2.2.3}{=} 1 - \sup_{k\geq 1}\inf_{n\geq k}\mu\left(\bigcap_{i=k}^{n}X\setminus A_i\right) = 1 - \sup_{k\geq 1}\inf_{n\geq k}\prod_{i=k}^{n}(1 - \mu(A_i)) \geq 1 - \sup_{k\geq 1}\inf_{n\geq k}\prod_{i=k}^{n}\exp(-\mu(A_i)) = 1 - \sup_{k\geq 1}\inf_{n\geq k}\exp\left(-\sum_{n\geq i\geq k}\mu(A_i)\right) = 1.$

4.13 Completeness and **μ-a.e. convergent subsequence for convergence in measure:** For a sequence $(f_n)_{n\geq 1}$ of Borel measurable functions $f_n : X \to Y$ from a **measure space** $(X;\mathcal{A};\mu)$ into a **Banach space** $(Y,||)$ the following statements are equivalent::

1. $(f_n)_{n\geq 1}$ is a **Cauchy sequence in measure,** i.e. $\lim_{k\geq 1}\sup_{n\geq k}\mu|_A(|f_n - f_k| > \epsilon) = 0$ for every $A \in \mathcal{A}$ with $\mu(A) < \infty$ and $\epsilon > 0$.
2. $(f_n)_{n\geq 1}$ converges **in measure** to a Borel measurable function $f : X \to Y$.
3. **Riesz convergence theorem**: Every **subsequence** of $(f_n)_{n\geq 1}$ has another **subsequence converging μ-a.e.** to the same Borel measurable function $f : X \to Y$

Proof: Let $A \in \mathcal{A}$ with $\mu(A) < \infty$.

1. ⇒ 2. : Due to the hypothesis for every $k \geq 1$ there is an $n_k \geq 1$ with $\mu|_A\left(|f_n - f_{n_k}| > 2^{-k}\right) < 2^{-k}$ for all $n \geq n_k$. Hence we have a partial sequence $(f_{n_k})_{k\geq 1}$with w.l.o.g. $n_{k+1} > n_k$ and $B_k = \left\{|f_{n_{k+1}} - f_{n_k}| > 2^{-k}\right\}$ such that $\sum_{k\geq 1}\mu|_A(B_k) < \infty$. According to 4.12 we obtain $\mu|_A\left(\bigcap_{m\geq 1}\bigcup_{k\geq m}(B_k)\right) = \mu(X \setminus B) = 0$ for $B = \bigcup_{m\geq 1}\bigcap_{k\geq m}(X \setminus B_k)$. Hence for every $x \in B$ there is an $m \geq 1$ such that $\sup_{k\geq m}|f_{n_k}(x) - f_{n_m}(x)| \leq \sum_{k\geq m}|f_{n_{k+1}}(x) - f_{n_k}(x)| \leq \sum_{k\geq m} 2^{-k} = 2^{-m+1}$. Thus we have a μ**-a.e. Cauchy sequence** $(f_{n_k})_{k\geq 1}$ which due to the **completeness** of Y and according to 4.7 converges μ-a.e. to a **measurable** $f : B \to Y$. Due to $\mu(A) < \infty$ we can apply 4.11 to find for every $\epsilon > 0$ an $m_\epsilon \geq 1$ such that $\mu|_A\left(|f_{n_m} - f| > \frac{\epsilon}{2}\right) < \frac{\epsilon}{2}$ for every $m \geq m_\epsilon$. Hence for every $n \geq n_m$ with $m \geq \max(m_\epsilon; k)$ and $2^{-k} < \frac{\epsilon}{2}$ we obtain $\mu|_A(|f_n - f| > \epsilon) \leq \mu|_A\left(\{|f_n - f_{n_m}| > \frac{\epsilon}{2}\} \cup \{|f_{n_m} - f| > \frac{\epsilon}{2}\}\right) \leq \mu|_A\left(|f_m - f_{n_m}| > \frac{\epsilon}{2}\right) + \mu|_A\left(|f_{n_m} - f| > \frac{\epsilon}{2}\right) < \epsilon$. This converse-triangle-inequality argument will be repeatedly used in the subsequent proofs.

2. ⇒ 3. : Due to 4.10.2 b) for every $k \geq 1$ there is an $n_k \geq 1$ such that $\mu(B_k) < 2^{-k}$ for $B_k = \left\{|f_{n_k} - f| \geq \frac{1}{k}\right\}$ whence $\mu|_A\left(\bigcup_{k\geq m} B_k\right) \leq 2^{-m+1}$ due to the **subadditivity** 2.2.1 and $\mu|_A\left(\bigcap_{m\geq 1}\bigcup_{k\geq m} B_k\right) = 0$ due to the **continuity from above** 2.2.3. Both properties require $\mu(A) < \infty$. The assertion then follows from 4.10.1 c).

3. ⇒ 1. : Suppose there is an $\epsilon > 0$ such that $\forall n_k \geq 1\ \exists n_{k+1} \geq n_k$ with $\mu|_A\left(|f_{n_{k+1}} - f_{n_k}| > \epsilon\right) > \epsilon$. As above we get $\mu|_A\left(f_{n_k} - f > \frac{\epsilon}{2}\right) + \mu|_A\left(|f_{n_{k+1}} - f| > \frac{\epsilon}{2}\right) \geq \mu|_A\left(|f_{n_k} - f_{n_{k+1}}| > \epsilon\right) > \epsilon$, i.e. either $\mu|_A\left(|f_{n_k} - f| > \frac{\epsilon}{2}\right) \geq \frac{\epsilon}{2}$ or $\mu|_A\left(|f_{n_{k+1}} - f| > \frac{\epsilon}{2}\right) \geq \frac{\epsilon}{2}$. For each $k \in \mathbb{N}$ we choose the f_{n_k} with respectively larger probability $\mu(\ldots)$ of deviation and thus obtain a subsequence $\left(f'_{n_k}\right)_{k\geq 1}$ with $\mu|_A\left(\left|f'_{n_k} - f\right| > \frac{\epsilon}{2}\right) \geq \frac{\epsilon}{2}$ for all $k \geq 1$ such that no part of this subsequence can possibly converge in measure to f and according to 4.11 with $\mu(A) < \infty$ this behaviour transfers to μ-a.e. convergence.

4.14 Completeness of μ-a.e. convergence: A sequence $(f_n)_{n\geq 1}$ of Borel measurable functions $f_n : X \to Y$ from a **measure space** $(X; \mathcal{A}; \mu)$ into a **Banach space** $(Y, ||)$ **converges μ-a.e.** to a Borel measurable $f : X \to Y$ iff $\lim_{k\to\infty}\mu|_A\left(\sup_{n\geq k}|f_k - f_n| > \epsilon\right) = 0$ for every $\epsilon > 0$**.**

Proof:

⇒: Applying the converse-triangle-inequality argument to suprema we obtain

$$\mu|_A\left(\sup_{n\geq k}|f_k - f_n| > \epsilon\right) \leq \mu|_A\left(|f_k - f| > \frac{\epsilon}{2}\right) + \mu|_A\left(\sup_{n\geq k}|f - f_n| > \frac{\epsilon}{2}\right)$$

The assertion follows from the **convergence in measure** due to 4.11 presuming $\mu(A) < \infty$ resp. the μ**-a.e. convergence** due to 4.10.1 b).

⇐: Due to the **continuity from below** 2.2.2 we obtain

$$\sup_{n\geq k}\mu|_A\left(|f_k - f_n| > \epsilon\right) \leq \mu|_A\left(\bigcup_{n\geq k}|f_k - f_n| > \epsilon\right) = \mu|_A\left(\sup_{n\geq k}|f_k - f_n| > \epsilon\right),$$

i.e. $(f_n)_{n\geq 1}$ **converges in measure** to f . Using again the converse-triangle-inequality we get

$$\mu|_A\left(\sup_{n\geq k}|f - f_n| > \epsilon\right) \leq \mu|_A\left(|f - f_k| > \frac{\epsilon}{2}\right) + \mu|_A\left(\sup_{n\geq k}|f_k - f_n| > \frac{\epsilon}{2}\right)$$

and hence the μ**-a.e. convergence** to f due to 4.10.1 b).

4.15 Egorov's convergence theorem: For every sequence $(f_n)_{n\geq 1}$ of Borel measurable functions $f_n : X \to Y$ from a **finite** measure space $(X;\mathcal{A};\mu)$ into a **Banach space** $(Y,||)$ **converging** μ**-a.e.** to a Borel measurable $f : X \to Y$ and every $\epsilon > 0$ there is a set $A_\epsilon \in \mathcal{A}$ with $\mu(A_\epsilon) < \epsilon$ such that $(f_n)_{n\geq 1}$ **uniformly converges** to f on $X \setminus A_\epsilon$.

Proof: Follows directly from 4.10.1 b) since for $\epsilon > 0$ there is a $k_\epsilon \geq 1$ such that we have $\mu(A_\epsilon) < \epsilon$ for $A_\epsilon := \bigcup_{n\geq k_\epsilon}\left\{|f_n(x) - f(x)| \geq \frac{1}{n}\right\}$ and $(f_n)_{n\geq 1}$ obviously converges **uniformly** to f on $X \setminus A_\epsilon$.

4.16 Examples:

1. The **function** sequence $(f_n)_{n\geq 1}$ with $f_n = \chi_{\left[\frac{j}{2^k};\frac{j+1}{2^k}\right]}$ for $n = 2^k + j$, $0 \leq j < 2^k$ and $k \geq 1$ on $\left([0;1];\mathcal{B}_{[0;1]};\lambda_{[0;1]}\right)$ converges **globally in measure** λ to $f = 0$ but the **point** sequences $(f_n(x))_{n\geq 1}$ converge for **no** $x \in [0;1]$ hence $(f_n)_{n\geq 1}$ converges **not** λ**-a.e.**
2. The **function** sequence $(f_n)_{n\geq 1}$ with $f_n = \chi_{[n;n+1]}$ for $n \geq 1$ on $(\mathbb{R};\mathcal{B};\lambda)$ **converges for every** $x \in \mathbb{R}$ **hence** λ**-a.e.** to $f = 0$ and hence **locally in measure but not globally so** since for $\epsilon < 1$ there is no $k \geq 1$ such that $\lambda\left(\bigcup_{n\geq k}\{|f_n - f| \geq \epsilon\}\right) < \infty$: The **continuity from above** 2.2.3 resp. theorem 4.12 do not apply.

5 Integration

Throughout this section and if not specified otherwise any function from X to Y is **Borel measurable** from a measure space $(X;\mathcal{A};\mu)$ with **positive measure** $\mu : A \to [0;\infty]$ into a Banach space $(Y,||)$ over a **field** K.

5.1 Step functions: The **characteristic functions** $\chi_A : X \to \{0;1\}$ for a measurable **support** $A \in \mathcal{A}$ with $\chi_A(x) = \begin{cases} 1, x \in A \\ 0, x \notin A \end{cases}$ are the most simple measurable functions on a measurable space $(X;\mathcal{A})$. They are identical with the **Dirac measure** δ_x from 2.3.1 albeit with interchanged roles for x and A. The family $\mathcal{S}(X;Y)$ denotes the **step functions** of the form $\sum\limits_{i=0}^{m} y_i\chi_{A_i}$ with $m \in \mathbb{N}$ such that $\bigcup\limits_{i=0}^{m} A_i = X$ with values $y_i \in Y$ and $\mu(A_i) < \infty$ for $1 \leq i \leq m$ but **vanishing** outside of these sets, i.e. $\alpha_0 = 0$. The step functions form a **vector space** of Borel measurable functions and according to 4.9 their **closure** $\overline{\mathcal{S}(X;Y)}$ with regard to pointwise convergence includes Borel measurable maps with **separable range** and **vanishing outside of a countable union of sets with finite measure**. Countable unions of sets with finite measure are called σ**-finite** with the most prominent example represented by $\mathbb{C}^n$ which is also separable. The following theorem shows that under these two conditions $\overline{\mathcal{S}(X;Y)}$ already contains **all** Borel measurable functions modulo null sets, i.e. $\mathcal{S}(X;Y)$ is **dense** in the **quotient space** of the Borel measurable functions with regard to the **equivalence relation** $f \sim g \Leftrightarrow f = g$ μ-a.e.

5.2 Theorem: For every Borel measurable function $f : X \to Y$ from a σ**-finite measure space** $(X;\mathcal{A};\mu)$ into a **separable Banach space** $(Y,||)$ there is a sequence $(\varphi_n)_{n\in\mathbb{N}} \subset \mathcal{S}(X;Y)$ of step functions converging μ-a.e. to f. Also for every set A of **finite measure** $\mu(A) < \infty$ and $\epsilon > 0$ there is a set $Z_\epsilon \subset X$ with measure $\mu(Z_\epsilon) < \epsilon$ such that $(\varphi_n)_{n\in\mathbb{N}}$ converges **uniformly** on $A \setminus Z_\epsilon$.

Note: With 4.9 we obtain a necessary and sufficient condition for measurability: A function $f : X \to Y$ from a σ**-finite measure space** $(X;\mathcal{A};\mu)$ into a **separable Banach space** $(Y,||)$ is **Borel measurable** iff there is a sequence $(\varphi_n)_{n\in\mathbb{N}} \subset \mathcal{S}(X;Y)$ of step functions converging μ-a.e. to f.

Proof: The image $f[A]$ of a set with finite measure $\mu(A) < \infty$ includes a dense subset $(y_l)_{l\geq 1}$ such that for every $n \geq 1$ we have $f[A] \subset \bigcup\limits_{l=1}^{\infty} B_{1/n}(y_l)$ resp. $A \subset \bigcup\limits_{l=1}^{\infty} C_{l,n}$ with $C_{l,n} = f^{-1}\left[B_{1/n}(y_l)\right]$ and consequently there is an $L_n \in \mathbb{N}$ with $\mu\left(A \setminus \bigcup\limits_{l=1}^{L_n} C_{l,n}\right) < 2^{-n}$.

Then the step functions $\varphi_n = \sum\limits_{l=1}^{L_n} y_l \chi_{D_l}$ with $D_l = C_{l,n} \setminus \bigcup\limits_{i=1}^{l-1} C_{i,n}$ converge to f

- uniformly on every $A \cap \bigcup\limits_{l=1}^{L_n} C_{l,n}$ with $\mu(Z_n) < 2^{-n}$ for $Z_n = A \setminus \bigcup\limits_{l=1}^{L_n} C_{l,n}$ and $n \geq 1$
- pointwise on $A \cap \bigcup\limits_{n=1}^{\infty} \bigcup\limits_{l=1}^{L_n} C_{l,n}$ with $\mu(Z) = 0$ for $Z = A \setminus \bigcup\limits_{n=1}^{\infty} \bigcup\limits_{l=1}^{L_n} C_{l,n}$.

For $X = \bigcup\limits_{k=1}^{\infty} A_k$ with w.l.o.g. pairwise disjoint A_k and $\mu(A_k) < \infty$ for every $k \geq 1$ there is a sequence $(\varphi_{k;j})_{j \geq 1}$ of step functions converging to f

- uniformly on $A_k \setminus Z_{k,n}$ with $\mu(Z_{k,n}) \leq 2^{-n}$.
- pointwise on $A_k \setminus Z_k$ with $\mu(Z_k) = 0$.

Then for every set $A \subset \bigcup\limits_{k=1}^{m} A_k$ with $m \geq 1$ the step functions $\psi_n(x) = \begin{cases} \varphi_{k;n}(x) & \text{if}\, x \in A_k; 1 \leq k \leq n \\ 0 & \text{if}\, x \in X \setminus \bigcup\limits_{k=1}^{n} A_k \end{cases}$ converges to f

- uniformly on $A \setminus \bigcup\limits_{k=1}^{\infty} Z_{k,n+m}$ with $\mu\left(\bigcup\limits_{k=1}^{\infty} Z_{k,n+m}\right) < 2^{-m}$ and
- pointwise on $X \setminus \bigcup\limits_{k=1}^{\infty} Z_k$ with $\mu\left(\bigcup\limits_{k=1}^{\infty} Z_k\right) = 0$.

5.3 Integral for step functions: For any **step function** $\varphi = \sum_{0 \leq i \leq m} y_i \chi_{A_i}$ with $y_i \in Y$ and $A_i \in \mathcal{A}$ the **integral** is defined by $\int \varphi d\mu := \sum_{0 \leq i \leq m} y_i \mu(A_i)$. **Uniqueness** and **linearity** $\int (\alpha\varphi + \beta\psi)\, d\mu = \alpha \int \varphi d\mu + \beta \int \psi d\mu$ for $\alpha, \beta \in K$ are obvious if we consider representations with **common** and **pairwise disjoint supports** $A_i \cap B_j$ for two elementary functions f und g as in 5.1 and observe the **additivity** 2.2.1 of the measure. Also we define integrals on **measurable subsets** as $\int_A \varphi d\mu := \int \varphi|_A d\mu$. On account of $\varphi|_{A \mathring{\cup} B} = \varphi|_A + \varphi|_B$ we have $\int_{A \mathring{\cup} B} \varphi d\mu = \int_A \varphi d\mu + \int_B \varphi d\mu$. For **positive** integrands φ with $\varphi[X] \subset [0; \infty[$ we have **monotonicity** in the form $\varphi < \psi \Rightarrow \int \varphi d\mu < \int \psi d\mu$. In general Banach spaces we still have $\left|\int_A \varphi d\mu\right| \leq \int_A |\varphi|\, d\mu \leq \|\varphi\|_\infty \mu(A)$ with the **supremum norm** $\|\varphi\|_\infty = \sup\limits_{x \in X} |\varphi(x)|$. The expression $\|\varphi\|_1 := \int |\varphi|\, d\mu$ defines the **$\mathcal{L}^1$- pseudonorm** (c.f. [9, 1.3]) on $\mathcal{S}(X;Y)$ with obvious **linearity** $\|\alpha\varphi + \beta\psi\|_1 = |\alpha| \cdot \|\varphi\|_1 + |\beta| \cdot \|\psi\|_1$ and the **triangle inequality** $\|\varphi + \psi\|_1 \leq \|\varphi\|_1 + \|\psi\|_1$. The latter follows from an application of the triangle inequality $|y_\varphi + y_\psi|_K \leq |x_\varphi|_K + |y_\psi|_K$ on the field K to representations with **common** as well as **pairwise disjoint supports** $A_i \cap B_j$ and invoking the **monotonicity of the integral** for the positive integrand $|\varphi|$. A sequence $(\varphi_n)_{n \in \mathbb{N}}$ of step functions **converges in mean** or with respect to $\mathcal{L}^1$ to a step function φ iff $\lim\limits_{n \to \infty} \|\varphi_n - \varphi\|_1 = 0$.

5.4 Convergence of step functions: For any $\mathcal{L}^1$- Cauchy sequence $(\varphi_n)_{n \in \mathbb{N}}$ of **step functions** $\varphi_n : X \to Y$ there exists a subsequence $(\varphi_{n_k})_{k \in \mathbb{N}}$ and for every $\epsilon > 0$ a set $Z_\epsilon \subset X$ with measure $\mu(Z_\epsilon) < \epsilon$ such that $(\varphi_{n_k})_{k \in \mathbb{N}}$ converges **absolutely** and **uniformly** on $X \setminus Z_\epsilon$ as well as **μ-a.e.** on X.

Proof: For every $k \geq 1$ there is an $n_k \geq n_{k-1} \in \mathbb{N}$ such that $\|\varphi_n - \varphi_{n_k}\|_1 \leq \frac{1}{2^{2k}}$ for every $n \geq n_k$. Then for $Y_k = \left\{|\psi_{k+1} - \psi_k| \geq \frac{1}{2^k}\right\}$ with $\psi_k := \varphi_{n_k}$ we have $\frac{1}{2^k}\mu(Y_k) = \int_{Y_n} \frac{1}{2^k} \leq \int_X |\psi_{k+1} - \psi_k|\, d\mu \leq \frac{1}{2^{2k}}$ whence $\mu(Y_k) \leq \frac{1}{2^k}$. Hence $\mu(Z_m) \leq \frac{1}{2^{m-1}}$ for $Z_m = \bigcup\limits_{k=m}^{\infty} Y_k$ and $|\psi_{k+1}(x) - \psi_k(x)| < \frac{1}{2^k}$ for every $x \in X \setminus Z_m$ resp. $k \geq m$ such that $\sum\limits_{k=m}^{\infty} (\psi_{k+1} - \psi_k)$ converges **absolutely** and **uniformly** on $X \setminus Z_m$. Hence $(\varphi_{n_k})_{k \geq m}$ converges **absolutely** and **uniformly** on $X \setminus Z_m$ resp. **pointwise** on $X \setminus \bigcap\limits_{m=1}^{\infty} Z_m$. Due to the **continuity from above** 2.2.3 we have $\mu\left(\bigcap\limits_{m=1}^{\infty} Z_m\right) = 0$.

5.5 Integrable functions: : The **Bochner integral** $\int fd\mu := \lim_{n\to\infty} \int \varphi_n d\mu < \infty$ is **well defined** and **finite** for every function $f : X \to Y$ with an **approximating sequence** $(\varphi_n)_{n\in\mathbb{N}}$, i.e. an $\mathcal{L}^1$**-Cauchy** sequence of **step functions** converging μ**-a.e.** to f. The **vector space** $\mathcal{B}(X;Y)$ of these **integrable functions** is the **Bochner space** whereas $\mathcal{L}^1(X;Y) = \{f : X \to Y : \|f\|_1 < \infty\} \subset \mathcal{B}(X;Y)$ of **Lebesgue integrable** functions is called the **Lebesgue space**. Hence the integral is a **linear functional** $I : \mathcal{B} \to K$. According to 4.9 every integrable $f : X \to Y$ from a σ**-finite measure space** $(X;\mathcal{A};\mu)$ into a **separable Banach space** $(Y, ||)$ is measurable.

In order to prove that the definition is independent of the approximating sequence we show: For two $\mathcal{L}^1$**- Cauchy** sequences $(\varphi_n)_{n\in\mathbb{N}}$ and $(\psi_n)_{n\in\mathbb{N}}$ of **step functions** converging μ**-a.e**. to the same function $f : X \to Y$ we have $\lim_{n\to\infty} \int \varphi_n d\mu = \lim_{n\to\infty} \int \psi_n d\mu < \infty$ as well as $\lim_{n\to\infty} \|\varphi_n - \psi_n\|_1 = 0$.

Proof: The existence of the limits is a consequence of the **completeness** of Y since $\left|\int (\varphi_n - \varphi_m) d\mu\right| \leq \|\varphi_n - \varphi_m\|_1$ such that $\left(\int \varphi_n d\mu\right)_{n\in\mathbb{N}}$ and likewise $\left(\int \psi_n d\mu\right)_{n\in\mathbb{N}}$ are again Cauchy sequences in Y. The differences $\gamma_n = \varphi_n - \psi_n$ also are $\mathcal{L}^1$-Cauchy and converge μ-a.e. to 0 such that for every $\epsilon > 0$ there is an $N \in \mathbb{N}$with $\|\gamma_m - \gamma_n\|_1 < \epsilon$ for all $m, n \geq N$. According to definition 5.1 there is a set A with $\mu(A) < \infty$ and $X \setminus A \subset \{\gamma_N = 0\}$ such that $\int_{X\setminus A} |\gamma_n| d\mu = \int_{X\setminus A} |\gamma_n - \gamma_N| d\mu \leq \|\gamma_n - \gamma_N\|_1 < \epsilon$. By the preceding lemma 5.4 there exists a subset $Z \subset A$ with $\mu(Z) < \frac{\epsilon}{1+\|f_N\|_\infty}$ and a subsequence converging to 0 uniformly on $A \setminus Z$ such that there is an $M \geq N$ with $\int_{A\setminus Z} |\gamma_n| d\mu < \epsilon$ for all $n \geq M$. Finally for $n \geq N$ we have $\int_Z |\gamma_n| d\mu \leq \int_Z |\gamma_n - \gamma_N| d\mu + \int_Z |\gamma_N| d\mu \leq \|\gamma_n - \gamma_N\|_1 + \|f_N\|_\infty \cdot \mu(Z) < 2\epsilon$. In sum we arrive at $\int_{X\setminus A} |\gamma_n| d\mu + \int_{A\setminus Z} |\gamma_n| d\mu + \int_Z |\gamma_n| d\mu < 4\epsilon$ which proves the assertion.

5.6 μ**-a.e. properties of integrable functions**:

1. Due to 5.1 integrable functions with approximating sequence $(\varphi_n)_{n\in\mathbb{N}}$ **vanish outside of the** σ**-finite set** $\bigcup_{n\in\mathbb{N}} \{\varphi_n \neq 0\}$.
2. According to 5.4 the integrable functions are μ**-a.e. finite** and **bounded** outside of a set of **finite measure**: Since the φ_n converge uniformly outside of a set Z_ϵ with $\mu(Z_\epsilon) < \epsilon$ for any $c \geq 0$ there is an $n \in \mathbb{N}$ such that $\{|f| \geq c\} \backslash Z_e \subset \{|\varphi_n| \geq \frac{c}{2}\}$ and hence $\mu(|f| \geq c) < \mu\left(|\varphi_n| \geq \frac{c}{2}\right) < \infty$.
3. In the case of **positive integrands** we have $\int fd\mu = 0 \Rightarrow f = 0$ μ-a.e. since for $A_n = \{f > 0\}$ the estimate $\frac{1}{n}\mu(A_n) \leq \int_{A_n} fd\mu \leq \int fd\mu = 0$ yields $\mu(f > 0) = \mu\left(\bigcup_{n\in\mathbb{N}} A_n\right) = 0$ on account of the **continuity form above** 2.2.3. In particular for positive integrable $f, g \in \mathcal{L}^1(X;\mathbb{R})$ with $f \leq g$ we have $\int fd\mu = \int gd\mu \Rightarrow f = g$ μ-a.e.

5.7 Special cases: For every integrable f with approximating sequence $(\varphi_n)_{n\in\mathbb{N}}$ the **restriction** $f|_A$ on any measurable subset is again integrable with the approximating sequence $(\varphi_n|_A)_{n\in\mathbb{N}}$. Hence we can define the **integral on measurable subsets** $\int_A fd\mu := \int f|_A d\mu$ with **additivity** extending to domains by $\int_{A\mathring{\cup}B} fd\mu = \int f|_{A\mathring{\cup}B} d\mu = \int (f|_A + f|_B) d\mu = \int_A fd\mu + \int_B fd\mu$. Likewise the **components** of functions in **finite dimensional Banach spaces** can be integrated separately since for every **continuous** $g : Y \to Z$ into another Banach space Z we have an approximating sequence $(g \circ \varphi_n)_{n\in\mathbb{N}}$ for $g \circ f$ with $\lim_{n\to\infty} (g \circ \varphi_n) = g \circ \lim_{n\to\infty} \varphi_n$ and in the case of **continuous and linear** g we even have $\int g \circ fd\mu = g \circ \int fd\mu$. For $Y = Y_1 \times Y_2$ and the continuous as well as linear **projections** $g = \pi_1 : Y \to Y_1$ resp. $g = \pi_2 : Y \to Y_2$ we obtain $\int (f_1, f_2) d\mu = \left(\int f_1 d\mu, \int f_2 d\mu\right)$. In particular f is integrable iff each of its **components** is integrable or in the case of $Y = \mathbb{C}$ iff $\mathrm{Re}f$ and $\mathrm{Im}f$ are integrable with $\int (\mathrm{Re}f + i\mathrm{Im}f) d\mu = \int \mathrm{Re}f d\mu + i \int \mathrm{Im}f d\mu$. For $Z = \mathbb{R}$ and the **continuous but not linear Banach norm** $g = ||$ we see that for every $f \in \mathcal{B}(X;Y)$ its Banach norm $|f| \in \mathcal{L}^1(X;\mathbb{R})$ is also integrable with approximating sequence $(|\varphi_n|)_{n\in\mathbb{N}}$. Note that in particular $(|\varphi_n|)_{n\in\mathbb{N}}$ is $\mathcal{L}_1$**-Cauchy** since $\||\varphi_n| - |\varphi_m|\|_1 \leq \|\varphi_n - \varphi_m\|_1$. The **converse statement** $|f| \in \mathcal{L}^1(X;\mathbb{R}) \Rightarrow f \in \mathcal{L}^1(X;Y)$ is only true for σ**-finite** $(X;\mathcal{A};\mu)$ and **separable** $(Y, ||)$. (cf. 5.15). The **well ordering** of the real numbers provides the space $\mathcal{L}^1(X;\mathbb{R})$ with additional properties: For $f, g \in \mathcal{L}^1(X;\mathbb{R})$ we have $\sup\{f;g\} = \frac{1}{2}(f + g + |f - g|) \in \mathcal{L}^1(X;\mathbb{R})$ and $\inf\{f;g\} = \frac{1}{2}(f + g - |f - g|) \in \mathcal{L}^1(X;\mathbb{R})$. $f = f^+ - f^- \in \mathcal{L}^1(X;\mathbb{R})$ iff its **positive part** $f^+ = \sup\{f;0\} \in \mathcal{L}^1(X;\mathbb{R})$ and its **negative part** $f^- = \inf\{f;0\} \in \mathcal{L}^1(X;\mathbb{R})$. Also for real valued functions the integral is **monotone**, i.e. $f \leq g \Rightarrow \int fd\mu \leq \int gd\mu$ which for positive integrands $f \geq 0$ extends to the domain in the form $A \subset B \Rightarrow \int_A fd\mu \leq \int_A fd\mu$.

5.8 Pseudonorm for Lebesgue integrable functions: According to 5.7 for every $f \in \mathcal{L}^1(X;Y)$ the integral $\|f\|_1 := \int |f| \, d\mu = \lim\limits_{n\to\infty} \|\varphi_n\|_1$ is well defined and a **pseudonorm** on $\mathcal{L}^1(X;Y)$: For $f, g \in \mathcal{L}^1(X;Y)$ with approximating sequences $(\varphi_n)_{n\in\mathbb{N}}, (\psi_n)_{n\in\mathbb{N}} \in \mathcal{S}(X;Y)$ we have $|f+g| \in \mathcal{L}^1(X;Y)$ with approximating sequence $(|\varphi_n + \psi_n|)_{n\in\mathbb{N}}$ and by **continuity of the addition** we obtain the **triangle inequality** $\|f+g\|_1 = \lim\limits_{n\to\infty} \|\varphi_n + \psi_n\|_1 \le \lim\limits_{n\to\infty} (\|\varphi_n\|_1 + \|\psi_n\|_1) = \lim\limits_{n\to\infty} \|\varphi_n\|_1 + \lim\limits_{n\to\infty} \|\psi_n\|_1 = \|f\|_1 + \|g\|_1$. Likewise the **continuity of the absolute value** extends the **continuity of the integral** from $\mathcal{S}(X;Y)$ to $\mathcal{L}^1(X;Y)$: $\left|\int f d\mu\right| = \left|\lim\limits_{n\to\infty} \int \varphi_n d\mu\right| = \lim\limits_{n\to\infty} \left|\int \varphi_n d\mu\right| \le \lim\limits_{n\to\infty} \int |\varphi_n| \, d\mu = \int |f| \, d\mu = \|f\|_1$.

5.9 Completeness of $\mathcal{L}^1$: The space $\left(\mathcal{L}^1(X;Y); \|\|_1\right)$ of Lebesgue integrable functions is **complete**.

Proof: For an $\mathcal{L}^1$-Cauchy sequence $(f_n)_{n\in\mathbb{N}} \subset \mathcal{L}^1(X;Y)$ there is a $\varphi_n \in \mathcal{S}(X;Y)$ with $\|f_n - \varphi_n\|_1 < \frac{1}{n}$. Hence for every $\epsilon > 0$ there is an $N \in \mathbb{N}$ such that for every $n, m \ge N$ we have $\|f_n - f_m\|_1 < \frac{\epsilon}{3}$ and consequently $\|\varphi_n - \varphi_m\|_1 \le \|\varphi_n - f_n\|_1 + \|f_n - f_m\|_1 + \|f_m - \varphi_m\|_1 \le \frac{1}{n} + \frac{\epsilon}{3} + \frac{1}{m} \le \epsilon$ for $n, m \ge \max\left\{N; \frac{3}{\epsilon}\right\}$, i.e. $(\varphi_n)_{n\in\mathbb{N}}$ is $\mathcal{L}^1$-Cauchy. Due to the 5.4 a subsequence $(\varphi_{n_k})_{k\in\mathbb{N}}$ converges μ-a.e. to an $f \in \mathcal{L}^1(X;Y)$, whence 5.5 yields $\int f d\mu = \lim\limits_{k\to\infty} \int \varphi_{n_k} d\mu$ and furthermore $\|f\|_1 = \lim\limits_{l\to\infty} \|\varphi_{n_l}\|_1$ and particularly $\|f - \varphi_{n_k}\|_1 = \lim\limits_{l\to\infty} \|\varphi_{n_l} - \varphi_{n_k}\|_1$ for every $k \in \mathbb{N}$ with 5.7. Since $(\varphi_k)_{k\in\mathbb{N}}$ is $\mathcal{L}^1$-Cauchy for every $\epsilon > 0$ there is an $k \in \mathbb{N}$ with $n_k \ge \frac{3}{\epsilon}$ such that on the one hand $\|\varphi_{n_l} - \varphi_{n_k}\|_1 < \frac{\epsilon}{3}$ and on the other hand $\left|\|f - \varphi_{n_k}\|_1 - \|\varphi_{n_l} - \varphi_{n_k}\|_1\right| < \frac{\epsilon}{3}$ for every $l \ge k$ whence $\|f - f_{n_k}\|_1 \le \|f - \varphi_{n_k}\|_1 + \|\varphi_{n_k} - f_{n_k}\|_1 \le \|\varphi_{n_l} - \varphi_{n_k}\|_1 + \frac{\epsilon}{3} + \frac{1}{n_k} \le \epsilon$. Hence $(\varphi_{n_k})_{k\in\mathbb{N}}$ is $\mathcal{L}^1$-convergent to f and due to its $\mathcal{L}^1$-Cauchy property the convergence extends to the complete sequence $(\varphi_n)_{n\in\mathbb{N}}$.

5.10 Convergence in mean and μ-a.e: For any $\mathcal{L}^1$-Cauchy sequence $(f_n)_{n\in\mathbb{N}} \subset \mathcal{L}^1(X;Y)$ of **Lebesgue integrable functions** $f_n : X \to Y$ there exists a subsequence $(f_{n_k})_{k\in\mathbb{N}}$ and for every $\epsilon > 0$ a set $Z_\epsilon \subset X$ with measure $\mu(Z_\epsilon) < \epsilon$ such that $(f_{n_k})_{k\in\mathbb{N}}$ converges **absolutely** and **uniformly** on $X \setminus Z_\epsilon$ as well as **μ-a.e.** and **in mean** on X to an integrable $f \in \mathcal{L}^1(X;Y)$.

Proof: According to the preceding theorem $(f_n)_{n\in\mathbb{N}}$ converges **in mean** to an $f \in \mathcal{L}^1(X;Y)$ such that for every $k \ge 1$ there is an $n_k \ge n_{k-1} \in \mathbb{N}$ with $\|f - f_{n_k}\|_1 \le \frac{1}{2^{2k}}$. Then for $Y_k = \left\{|f - f_{n_k}| \ge \frac{1}{2^k}\right\}$ we have $\frac{1}{2^k}\mu(Y_k) = \int_{Y_n} \frac{1}{2^k} d\mu \le \int_X |f - f_{n_k}| \, d\mu \le \frac{1}{2^{2k}}$ whence $\mu(Y_k) \le \frac{1}{2^k}$. Hence $\mu(Z_m) \le \frac{1}{2^{m-1}}$ for $Z_m = \bigcup\limits_{k=m}^{\infty} Y_k$ and $|f(x) - f_{n_k}(x)| < \frac{1}{2^k}$ for every $x \in X \setminus Z_m$ resp. $k \ge m$, i.e. $(f_{n_k})_{k\ge m}$ converges to f **absolutely** and **uniformly** on $X \setminus Z_m$ as well as **pointwise** on $X \setminus \bigcap\limits_{m=1}^{\infty} Z_m$ with $\mu\left(\bigcap\limits_{m=1}^{\infty} Z_m\right) = 0$.

5.11 Norm for Lebesgue integrable functions: Lebesgue integrable functions with common approximatig sequences are μ-a.e. equal and partition $\mathcal{L}^1(X;Y)$ into **equivalence classes** (c.f. 5.1). The corresponding **quotient space** is equally called a **Lebesgue space** and denoted as $L^1(X;Y)$. On this quotient space $\|f\|_1$ is **positive definite** and hence a **norm**, since for $\|f\|_1 = 0$ the null sequence $(0)_{n\in\mathbb{N}}$converges in mean to f and due to the preceding paragraph it also converges μ-a.e. to f whence μ-a.e. $f = 0$. Note that $\left(L^1(X;Y); \|\|_1\right)$ is a **Banach space**, but there is no topology on $\mathcal{L}^1(X;Y)$ corresponding to **μ-a.e. convergence**. (cf. [6])

5.12 Levi's monotone convergence theorem: For every **monotone** sequence $(f_n)_{n\in\mathbb{N}} \in L^1(X;\mathbb{R})$ of **real valued** $f_n : X \to \mathbb{R}$ we have $\int \lim\limits_{n\in\mathbb{N}} f_n d\mu = \lim\limits_{n\in\mathbb{N}} \int f_n d\mu$. In the case of $\lim\limits_{n\in\mathbb{N}} \left|\int f_n d\mu\right| < \infty$ the sequence converges **both in mean and μ-a.e.** to $f = \lim\limits_{n\in\mathbb{N}} f_n \in L^1(X;\mathbb{R})$.

Proof: Due to the **monotonicity of the integral** 5.7 in the case of an **increasing sequence** we have $\sup\limits_{n\in\mathbb{N}} \int f_n d\mu \le \int \sup\limits_{n\in\mathbb{N}} f_n d\mu$ which proves the assertion in the case of $\sup\limits_{n\in\mathbb{N}} \int f_n d\mu = \infty$. For $\sup\limits_{n\in\mathbb{N}} \int f_n d\mu < \infty$ and $n \ge m$ we have $\|f_n - f_m\|_1 = \int (f_n - f_m) \, d\mu = \int f_n d\mu - \int f_m d\mu$ whence follows that $(f_n)_{n\in\mathbb{N}}$ is L^1-Cauchy. According to 5.10 a subsequence converges μ-a.e. and in mean to an $f = \lim\limits_{n\in\mathbb{N}} f_n \in L^1(X;\mathbb{R})$ and due to the increasing character this must be true for the complete sequence. Finally for every $\epsilon > 0$ there is an $n \in \mathbb{N}$ with $\int (f - f_n) \, d\mu < \epsilon$ and hence $\int f d\mu = \int (f - f_n) \, d\mu + \int f_n d\mu = \epsilon + \int f_n d\mu$ which proves $\int \sup\limits_{n\in\mathbb{N}} f_n d\mu = \sup\limits_{n\in\mathbb{N}} \int f_n d\mu$. In the case of a decreasing sequence apply the proof to $(-f_n)_{n\in\mathbb{N}}$.

5.13 Fatou's lemma: Fo every sequence $(f_n)_{n\in\mathbb{N}} \in L^1\left(X;\mathbb{R}_0^+\right)$ of **positive Borel measurable** functions with $\lim\limits_{k\to\infty}\inf\limits_{k\le n}\int f_n d\mu < \infty$ we have $f = \lim\limits_{k\to\infty}\inf\limits_{k\le n} f_n \in L^1\left(X;\mathbb{R}_0^+\right)$ with $\int \lim\limits_{k\to\infty}\inf\limits_{k\le n} f_n d\mu \le \lim\limits_{k\to\infty}\inf\limits_{k\le n}\int f_n d\mu$.

Proof: For every $k \in \mathbb{N}$ the **decreasing** sequence $\left(\inf\limits_{k\le n\le m} f_n\right)_{m\in\mathbb{N}}$ converges μ-a.e. to $\inf\limits_{k\le n} f_n$ such that due to the preceding theorem we have $\int \inf\limits_{k\le n} f_n d\mu = \lim\limits_{m\to\infty}\int \inf\limits_{k\le n\le m} f_n d\mu \le \lim\limits_{m\to\infty}\inf\limits_{k\le n\le m}\int f_n d\mu = \inf\limits_{k\le n}\int f_n d\mu \le \lim\limits_{k\to\infty}\inf\limits_{k\le n}\int f_n d\mu$. Now we apply the monotone convergence theorem a second time to the **increasing** sequence $\left(\inf\limits_{k\le n} f_n\right)_{k\in\mathbb{N}}$ and obtain $\int \lim\limits_{k\to\infty}\inf\limits_{k\le n} f_n d\mu = \lim\limits_{k\to\infty}\int \inf\limits_{k\le n} f_n d\mu \le \lim\limits_{k\to\infty}\inf\limits_{k\le n}\int f_n d\mu$.

5.14 Lebesgue's dominated convergence theorem: A sequence $(f_n)_{n\in\mathbb{N}} \subset L^1(X;Y)$ **converging μ-a. e.** to some f **converges in mean** to f with $f \in L^1(X;Y)$ iff there is an **Lebesgue integrable majorant** $g \in L^1(X;\mathbb{R}_0^+)$ such that for every $n \in \mathbb{N}$ and μ-a.e. we have $|f_n| \le g$.

Proof: For every $K \in \mathbb{N}$ the **increasing** sequence $\left(\sup\limits_{k\le m;n\le l}|f_n - f_m|\right)_{l>k}$ is μ-a.e. bounded by $|f_n - f_m| \le 2g$ and hence has bounded integrals $\int\left(\sup\limits_{k\le m;n\le l}|f_n - f_m|\right)d\mu \le 2\int g d\mu$. According to the **monotone convergence theorem** 5.12 we conclude $\int\left(\sup\limits_{k\le m;n}|f_n - f_m|\right)d\mu \le 2\int g d\mu$ for every $k \in \mathbb{N}$. Hence we can apply the monotone convergence theorem a second time to the **decreasing** sequence $\left(\sup\limits_{k\le m;n}|f_n - f_m|\right)_{k\ge 1}$ converging μ-a.e. to 0 to obtain $\lim\limits_{k\to\infty}\int\left(\sup\limits_{k\le m;n}|f_n - f_m|\right)d\mu = 0$. Hence $(f_n)_{n\in\mathbb{N}}$ is L^1-Cauchy and due to the **completeness** 5.9 of $L^1(X;Y)$ it converges **in mean** to an $f^{\#} \in L^1(X;Y)$ coninciding μ-a.e. with f according to 5.10.

5.15 Absolute value of integrable functions: A Borel measurable function $f : X \to Y$ from a **σ-finite measure space** $(X;\mathcal{A};\mu)$ into a **separable Banach space** $(Y,||)$ is **integrable** with $\int f d\mu \le \int |f|\, d\mu \le \int g d\mu$ if there is a $g \in L^1(X;\mathbb{R})$ with μ-a.e. $|f| \le g$. In particular f is integrable if its **absolute value** $|f|$ is integrable. The inequality is a trivial consequence of the **continuity of the integral** according to 5.8. The converse is true for the subset of the **Lebesgue-integrable** functions but neither for the **Bochner integral** nor for the **improper Riemann integral** which is not included in the Lebesgue integral (cf. 5.25). E.g. $f(x) = \frac{\sin(x)}{x}$ is integrable with **Bochner** and **Riemann** but not with **Lebesgue**.

Proof: According to 5.2 there is a sequence $(\varphi_n)_{n\in\mathbb{N}} \subset S(X;Y)$ of step functions converging μ-a.e. to f. Due to 5.5 the function g is Borel measurable. Hence the sets $\{|\varphi_n| \le 2g\}$ are measurable and by $\psi_n(x) = \begin{cases} \varphi_n(x) & \text{if } |\varphi_n(x)| \le 2g(x) \\ 0 & \text{if } |\varphi_n(x)| > 2g(x) \end{cases}$ we have a sequence of **integrable** step functions **bounded** by g and **converging μ-a.e.** to f . Due to 5.14 the convergence is also in mean and consequently f is integrable.

5.16 Dominated convergence for series: For a sequence $(f_n)_{n\in\mathbb{N}}$ of functions $f_n : X \to Y$ from a **σ-finite measure space** $(X;\mathcal{A};\mu)$ into a **separable Banach space** $(Y,||)$ with $\sum\limits_{n\in\mathbb{N}}\int |f_n|\, d\mu < \infty$ the series $\sum\limits_{n\in\mathbb{N}} f_n := f$ converges **μ-a.e.** as well as **in mean:** $\sum\limits_{n\in\mathbb{N}}\int f_n d\mu = \int f d\mu$.

Proof: Owing the triangle inequality we have $\left|\int f d\mu\right| \overset{5.8}{\le} \int\left|\sum\limits_{n\in\mathbb{N}} f_n\right| d\mu \overset{5.15}{\le} \int\sum\limits_{n\in\mathbb{N}}|f_n|\, d\mu \overset{5.12}{=} \sum\limits_{n\in\mathbb{N}}\int |f_n|\, d\mu$ $< \infty$, i.e. $f \in L^1(X;Y)$ and also μ-a.e. **finite** resp**. convergent** due to **monotone convergence** 5.12. The convergence **in mean** follows by **dominated convergence** 5.14 with the **majorant** $g := \sum\limits_{n\in\mathbb{N}}|f_n|$.

5.17 Sequences with bounded norms: For the μ-**a. e.** limit $f : X \to Y$ of a sequence $(f_n)_{n\in\mathbb{N}} \subset L^1(X;Y)$ of **Lebesgue integrable** functions from a **σ-finite measure space** $(X;\mathcal{A};\mu)$ into a **separable Banach space** $(Y, ||)$ with **bounded norms** $\|f_n\|_1 \le C$ for some $C \ge 0$ and every $n \in \mathbb{N}$ we have $\|f\|_1 \le C$ and in particular $f \in L^1(X;Y)$.

Proof: The f_n are measurable due to 5.5 and so is f according to 4.9. Because of $\lim\limits_{n\to\infty} |f_n| = |f|$ we can apply first **Fatou's lemma** 5.13 to obtain $\|f\|_1 \le C$ and then 5.15 to infer $f \in L^1(X;Y)$.

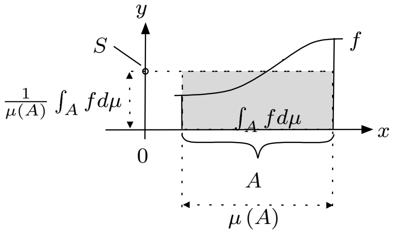
Note: Due to the **missing bound** for the absolute values $|f_n|$ **we can not assert convergence in mean**. E.g. for the sequence $(\varphi_n)_{n\ge 1}$ with $\varphi_n = n \cdot \chi_{[0;\frac{1}{n}]}$ we have $\lim\limits_{n\to\infty} |\varphi_n| = 0$ but $\lim\limits_{n\to\infty} \|\varphi_n\|_1 = 1$.

5.18 Product of Lebesgue integrable and bounded functions: For **σ-finite measure space** $(X;\mathcal{A};\mu)$ and a **separable Banach space** $(Y, ||)$ the product fg of an **(Lebesgue) integrable** $f : X \to Y$ and a **bounded measurable** $g : X \to K$ into the **normed, complete and separable field** K is **(Lebesgue) integrable**.

Proof: Due to 5.2, 5.5 and 5.9 there are sequences $(\varphi_n)_{n\in\mathbb{N}} \subset \mathcal{S}(X;Y)$ and $(\psi_n)_{n\in\mathbb{N}} \subset \mathcal{S}(X;K)$ of step functions with $(\varphi_n)_{n\in\mathbb{N}}$ converging both in **mean** and μ-a.e. to f and $(\psi_n)_{n\in\mathbb{N}}$ converging μ-a.e. to g. Then $(\varphi_n \cdot \psi_n)_{n\in\mathbb{N}} \subset \mathcal{S}(X;Y)$ is a **L^1-Cauchy** sequence converging μ-a.e. and according to 5.4 also **in** mean to fg which is hence integrable with $|fg| \le |f| \cdot \|g\|_\infty < \infty$. In the case of $f \in L^1(X;Y)$ we have $|f| \in L^1(X;Y)$ whence $|f| \cdot \|g\|_\infty \in L^1(X;Y)$ and hence $|fg| \in L^1(X;Y)$ due to 5.15.

5.19 Mean value theorem for integration:: For every **integrable** $f \in \mathcal{B}(X;Y)$ from a **σ-finite measure space** $(X;\mathcal{A};\mu)$ into a **separable Banach space** $(Y, ||)$ with the mean value $\frac{1}{\mu(A)} \int_A f d\mu \in S$ for some **closed** subset $S \subset Y$ and every $A \in \mathcal{A}$ with $0 < \mu(A) < \infty$ we have $\mu(f \notin S) = 0$.

Proof: In the case of $\mu(X) < \infty$ for any closed disk $\overline{B}_r(z) \subset Y \setminus S$ with $\mu(A) > 0$ for $A = f^{-1}\left[\overline{B}_r(z)\right]$ we have $\left|\frac{1}{\mu(A)} \int_A f d\mu - z\right| = \left|\frac{1}{\mu(A)} \int_A (f - z)\, d\mu\right| \le \frac{1}{\mu(A)} \int_A |f - z|\, d\mu \le r$ contrary to $\frac{1}{\mu(A)} \int_A f d\mu \in S$. Therefore we must assume $\mu\left(f \in \overline{B}_r(z)\right) = 0$ and since $Y \setminus S$ is a countable union of such disks the assertion follows from the σ-additivity of μ. Hence if we assume the hypothesis for every $A \cap X_n$ with $A \in \mathcal{A}$, $\mu(X_n) < \infty$ and $X = \bigcup_{n\in\mathbb{N}} X_n$ we obtain $f(x) \in S$ for every $x \in X_n \setminus Z_n$ with $\mu(Z_n) = 0$ and hence for $X \setminus \bigcup_{n\in\mathbb{N}} Z_n$ with $\mu\left(\bigcup_{n\in\mathbb{N}} Z_n\right) = 0$.



The following theorem asserts that step functions on arbitrary measurable sets can be approximated by step functions on an algebra of sets with **finite measures**, e.g. the algebra $\mathcal{F}$ of **figures** in $\mathbb{R}^n$. This step is necessary to identify the **Lebesgue integral** as special case of the **Bochner integral**. The theorem will be prepared by two lemmata:

5.20 Lemma: For every **algebra** $\mathcal{F} \subset \mathcal{A}$ of sets of **finite measure** in a measure space $(X;\mathcal{A};\mu)$ and every $F \in \mathcal{F}$ we consider the **vector space** $\mathcal{S}(\mathcal{F}_F;\mathbb{R})$ of step functions on the **trace algebra** $\mathcal{F}_F$ being of the form $\sum\limits_{i=0}^{m} y_i \chi_{F_i}$ with $m \in \mathbb{N}$ on sets $F_0 = X \setminus F$ resp. $F_i \in \mathcal{F}_F$ with $\bigcup\limits_{i=1}^{m} F_i = F$ and with values $y_0 = 0$ resp. $y_i \in \mathbb{R}$ for $1 \le i \le m$. Then for every $F \in \mathcal{F}$ and the family $\mathcal{N}_F = \left\{A \in \mathcal{A}_F : \chi_A \in \overline{(\mathcal{S}(\mathcal{F}_F;\mathbb{R});\|\|_1)}\right\} \subset \mathcal{A}_F$ is a σ-algebra on the set F.

Proof: Note that every $A \in \mathcal{N}_F$ must be of finite measure but not necessarily be an element of the algebra $\mathcal{F}_F$. Since for $\varphi, \psi \in \mathcal{S}(\mathcal{F}_F;\mathbb{R})$ we obviously have $\sup\{\varphi;\psi\}, \inf\{\varphi;\psi\} \in \mathcal{S}(\mathcal{F}_F;\mathbb{R})$ the closure $\overline{(\mathcal{S}(\mathcal{F}_F;\mathbb{R});\|\|_1)}$ is again a **vector space closed** with respect to **sup** and **inf**. $\mathcal{N}_F$ is an **algebra** since obviously $\emptyset \in \mathcal{N}_F$ and for every $A, B \in \mathcal{N}_F$ the characteristic functions $\chi_{A\cup B} = \sup\{\chi_A;\chi_B\}$, $\chi_{A\cap B} = \inf\{\chi_A;\chi_B\}$ as well as $\chi_{A\setminus B} = \chi_A - \chi_B$ are all in $\overline{(\mathcal{S}(\mathcal{F}_F;\mathbb{R});\|\|_1)}$and consequently their supports $A \cup B$, $A \cap B$ resp. $A \setminus B$ are in $\mathcal{N}_F$. It is a **σ-algebra** since $F \in \mathcal{N}_F$ and for every paiwise disjoint sequence $(A_n)_{n\in\mathbb{N}} \subset \mathcal{N}_F$ with union $A = \mathring{\bigcup}_{n\in\mathbb{N}} A_n$ and every $\epsilon > 0$ due to the **continuity from below** 2.2.2 we have an $N \in \mathbb{N}$ with $\mu\left(\bigcup_{k>N} A_k\right) < \epsilon$ and approximating step functions $(\varphi_n)_{n\in\mathbb{N}} \subset \mathcal{S}(\mathcal{F}_F;\mathbb{R})$ such that $\|\chi_{A_n} - \varphi_n\|_1 < \frac{\epsilon}{2^n}$ for $n \in \mathbb{N}$ whence $\left\|\chi_A - \sum\limits_{k=0}^{n} \varphi_n\right\|_1 \le \left\|\chi_A - \chi_{\bigcup_{0\le k\le n} A_k}\right\|_1 +$

$\left\|\chi_{\bigcup_{0\leq k\leq n}A_k}-\sum\limits_{k=0}^{n}\varphi_n\right\|_1=\left\|\chi_{\bigcup_{k>N}A_k}\right\|_1+\left\|\sum\limits_{k=0}^{n}\chi_{A_k}-\sum\limits_{k=0}^{n}\varphi_n\right\|_1\leq\mu\left(\bigcup_{k>N}A_k\right)+\sum\limits_{k=0}^{n}\left\|\chi_{A_n}-\varphi_n\right\|_1<$ 2ϵ.

5.21 Lemma: Let $\mathcal{F}\subset\mathcal{A}$ be an **algebra** of sets with **finite measure** in a measure space $(X;\mathcal{A};\mu)$ and $F_{n\in\mathbb{N}}\in\mathcal{F}$ with $X=\bigcup_{n\in\mathbb{N}}F_n$ with σ-algebrae $\mathcal{N}_n\subset\mathcal{A}_{F_n}$ according to 5.20. Then the family $\mathcal{N}$ $=\{A\subset X:A\cap F_n\in\mathcal{N}_n\forall n\in\mathbb{N}\}$ also is a σ-algebra on X.

Proof: For every $A\in\mathcal{N}$ we have $X\setminus A\cap F_n\in\mathcal{N}_n$ whence $X\setminus A\in\mathcal{N}$. For every $A,B\in\mathcal{N}$ we have $(A\cap B)\cap F_n=(A\cap F_n)\cap(B\cap F_n)\in\mathcal{N}_n$whence $A\cap B\in\mathcal{N}$. Finally for $(A_m)_{m\in\mathbb{N}}\subset\mathcal{N}$ the equality $\left(\bigcup_{m\in\mathbb{N}}A_m\right)\cap F_n=\mathring{\bigcup}_{m\in\mathbb{N}}(A_m\cap F_n)$ shows that $\bigcup_{m\in\mathbb{N}}A_m\in\mathcal{N}$.

5.22 Theorem: For every **algebra** $\mathcal{F}\subset\mathcal{A}$ of sets with **finite measure generating** $\mathcal{A}=\sigma(\mathcal{F})$ on a **σ-finite measure space** $(X;A;\mu)$ we have $\overline{(\mathcal{S}(\mathcal{F};Y);\|\|_1)}=\mathcal{B}(X;Y)$.

Proof:

According to the hypothesis there is a sequence $(F_n)_{n\geq 1}\subset\mathcal{F}$ of w.l.o.g. pairwise disjoint sets with finite measure $\mu(F_n)<\infty$ and $\mathring{\bigcup}_{n\geq 1}F_n=X$. By lemma 5.21 $\mathcal{N}_{F_n}\subset\mathcal{A}_F$ is a σ-algebra and by lemma 5.21 the family $\mathcal{N}$ is a σ-algebra containing $\mathcal{F}$ and hence $\mathcal{A}=\sigma(\mathcal{F})$ such that for every measurable set $A\in\mathcal{A}$ with finite measure $\mu(A)<\infty$ we have $A\cap F_n\in\mathcal{N}_{F_n}$, i.e. for every $\epsilon>0$ there is a $\varphi_n\in\mathcal{S}(\mathcal{F}_{F_n};\mathbb{R})$ with $\|\chi_{A\cap F_n}-\varphi_n\|_1<\frac{\epsilon}{2^n}$. Due to the **continuity from above** 2.2.3 there is an $N\in\mathbb{N}$ such that $\left\|\chi_A-\sum\limits_{n=1}^{N}\chi_{A\cap F_n}\right\|_1=\mu\left(A-\bigcup\limits_{n=1}^{N}(A\cap F_n)\right)<\epsilon$ whence $\left\|\chi_A-\sum\limits_{n=1}^{N}\varphi_n\right\|_1\leq$ $\left\|\chi_A-\sum\limits_{n=1}^{N}\chi_{A\cap F_n}\right\|_1+\left\|\sum\limits_{n=1}^{N}\chi_{A\cap F_n}-\sum\limits_{n=1}^{N}\varphi_n\right\|_1<2\epsilon$. Thus for every step map $\psi=\sum\limits_{i=0}^{m}y_i\chi_{A_i}\in\mathcal{S}(\mathcal{A};Y)$ with $m\in\mathbb{N}$ such that $\bigcup\limits_{i=0}^{m}A_i=X$ with values $y_i\in Y$ and $\mu(A_i)<\infty$ for $1\leq i\leq m$ and $\alpha_0=0$ there are step maps $\varphi_i=\sum\limits_{n=1}^{N}\varphi_{i,n}\in\mathcal{S}(\mathcal{F};\mathbb{R})$ with $\|\chi_{A_i}-\varphi_i\|_1<\frac{\epsilon}{m\cdot|y_i|}$ such that $\left\|\sum\limits_{i=0}^{m}y_i\chi_{A_i}-\sum\limits_{i=0}^{m}y_i\varphi_i\right\|_1=$ $\|\psi-\varphi\|_1<\epsilon$ with $\varphi=\sum\limits_{i=0}^{m}y_i\varphi_i\in\mathcal{S}(\mathcal{F};Y)$. The assertion now follows from the definition 5.5 of integrable functions since $\overline{(\mathcal{S}(\mathcal{A};Y);\|\|_1)}=\mathcal{B}(X;Y)$.

5.23 Theorem: For every **algebra** $\mathcal{F}\subset\mathcal{A}$ of sets with **finite measure generating** $\mathcal{A}=\sigma(\mathcal{F})$ on a **σ-finite measure space** $(X;A;\mu)$ and every **integrable** $f\in\mathcal{B}(X;Y)$ into a **separable Banach space** $(Y,||)$ the following propositions hold:

1. If $\int_F fd\mu=0$ for every $F\in\mathcal{F}$ then $f=0$ μ-a.e.
2. If $\int f\varphi d\mu=0$ for every $\varphi\in\mathcal{S}(\mathcal{F};\mathbb{R})$ then $f=0$ μ-a.e.
3. If $\int_F fd\mu\leq c\cdot\mu(F)$ for some $c\geq 0$ and every $F\in\mathcal{F}$ with $\mu(F)>0$ then $|f|\leq c$ μ-a.e.

Proof: According to 5.5 and 5.22 f or every measurable set $A\in\mathcal{A}$ with finite measure $\mu(A)<\infty$ there exists a sequence $(\varphi_n)_{n\in\mathbb{N}}\subset\mathcal{S}(\mathcal{F};Y)$ converging in mean as well as μ-a.e. to χ_A. Taking $\sup\{\varphi_n;0\}$ resp. $\inf\{\varphi_n;1\}$ we can w.l.o.g. assume $0\leq\varphi_n\leq 1$. Then we have $|\varphi_n f|\leq|f|$ for every $n\in\mathbb{N}$ and $(\varphi_n f)_{n\in\mathbb{N}}$ converges μ-a.e. to $\chi_A f$. By **dominated convergence** 5.14 and since $\int\varphi_n fd\mu=0$ we conclude $\int\chi_A fd\mu=0$. Now every measurable set is a countable union of w.l.o.g. pairwise disjoint sets of finite measure such that a second instance of **dominated convergence** yields $\int\chi_A fd\mu=0$ for every **measurable** $A\in\mathcal{A}$. **Proposition 1.** now follows from the **mean value** theorem for integrals 5.19 applied to $S=\{0\}$. **Proposition 2.** is obtained from 1. by taking $\varphi=\chi_F$. Finally we derive **Proposition 3.** from 5.19 applied to $S_n=\overline{B}_{c+1/n}(0)$ for $n\geq 1$ and considering $\{|f|\leq c\}=\bigcap_{n\in\mathbb{N}}\{f\in S_n\}$.

5.24 Theorem: A function f from a **σ-finite measure space** $(X;A;\mu)$ into a **separable Banach space** $(Y,||)$ is **integrable** iff there is an **increasing** sequence $(A_n)_{n\in\mathbb{N}} \subset \mathcal{A}$ with $\bigcup_{n\in\mathbb{N}} A_n = X$ and $\lim\limits_{n\to\infty} \int_{A_n} f d\mu \in Y$ exists. In that case we have $\int f d\mu = \lim\limits_{n\to\infty} \int_{A_n} f d\mu$.

Proof: $\Rightarrow$: Take $A_n = X$ for $n \in \mathbb{N}$. $\Leftarrow$: Due to the hypothesis for every $m \geq 1$ there is an $n(m) \in \mathbb{N}$ such that $\left|S - \int_{A_{n(m)}} f d\mu\right| < \frac{1}{2m}$ with $S = \lim\limits_{n\to\infty} \int_{A_n} f d\mu$. Also due to 5.4 there is a $\varphi_{n(m)} \in S(X;Y)$ with $\int \left|f \cdot \chi_{A_{n(m)}} - \varphi_{n(m)}\right| d\mu < \frac{1}{2m}$ and $\left|f(x) - \varphi_{n(m)}(x)\right| < \frac{1}{m}$ for every $x \in A_{n(m)} \setminus Z_{n(m)}$ with $\mu\left(Z_{n(m)}\right) < \frac{1}{m}$. Hence we have $\left|S - \int \varphi_{n(m)} d\mu\right| \leq \left|S - \int_{A_{n(m)}} f d\mu\right| + \left|\int_{A_{n(m)}} f d\mu - \int \varphi_{n(m)} d\mu\right| \leq \frac{1}{2m} + \int \left|f \cdot \chi_{A_{n(m)}} - \varphi_{n(m)}\right| d\mu < \frac{1}{m}$. Furthermore $\lim\limits_{m\to\infty} \left(\varphi_{n(m)}(x)\right) = f(x)$ for every $x \in \bigcup_{m\geq 1} \left(A_{n(m)} \setminus Z_{n(m)}\right) = \bigcup_{m\geq 1} A_{n(m)} \setminus \bigcap_{m\geq 1} Z_{n(m)} = X \setminus \bigcap_{m\geq 1} Z_{n(m)}$ with $\mu\left(\bigcap_{m\geq 1} Z_{n(m)}\right) = 0$. Hence $\left(\varphi_{n(m)}\right)_{m\geq 1} \subset S(X;Y)$ is an approximating sequence for f and we have $S = \lim\limits_{n\to\infty} \int_{A_n} f d\mu = \lim\limits_{m\to\infty} \int \varphi_{n(m)} d\mu = \int f d\mu$.

5.25 Comparison with the Riemann integral:

1. Every **Riemann integrable** function $f : [a;b] \to \mathbb{R}$ is **integrable** and the two integrals are equal: $\int_a^b f(x)dx = \int_{[a;b]} f d\lambda$.
2. $f : \mathbb{R} \to \mathbb{R}$ is **integrable** on $\mathbb{R}$ iff the **improper Riemann integral** exists and in this case the two integrals again coincide: $\lim\limits_{n\in\mathbb{N}} \int_{-n}^{n} f(x)dx = \int_{\mathbb{R}} f d\lambda$.

Proofs:

1. For every partition $z_n := (a = a_0 \leq a_1 \leq ... \leq a_n = b)$ of the interval $[a;b]$ we can compare the **lower Darboux sum** $L_{z_n} := \sum\limits_{i=0}^{n} \overline{\gamma}_i (a_i - a_{i-1}) \leq \int_{[a;b]} l_{z_n} d\lambda$ with $\overline{\gamma}_i := \inf f\,[[a_{i-1};a_i]]$ resp. the **upper Darboux sum** $U_{z_n} := \sum\limits_{i=0}^{n} \overline{\Gamma}_i (a_i - a_{i-1}) \geq \int_{[a;b]} u_{z_n} d\lambda$ with $\overline{\Gamma}_i := \sup f\,[[a_{i-1};a_i]]$ to the **integrals** of the **corresponding step functions** $l_{z_n} := \sum\limits_{i=0}^{n} \gamma_i \chi_{[a_{i-1};a_i[}$ with $\gamma_i := \inf f\,[[a_{i-1};a_i[] \geq \overline{\gamma}_i$ resp. $u_{z_n} := \sum\limits_{i=0}^{n} \Gamma_i \chi_{[a_{i-1};a_i[}$ with $\Gamma_i := \sup f\,[[a_{i-1};a_i[] \leq \overline{\Gamma}_i$. According to the hypothesis there are sequences $(z_n)_{n\in\mathbb{N}}$ of partitions such that z_{n+1} is a refinement of z_n such that due to the **monotonicity of the integral** 5.7 we obtain $\int_a^b f(x)dx = \lim\limits_{n\to\infty} L_{z_n} \leq \lim\limits_{n\to\infty} \int_{[a;b]} l_{z_n} d\lambda \leq \lim\limits_{n\to\infty} \int_{[a;b]} u_{z_n} d\lambda \leq \lim\limits_{n\to\infty} U_{z_n} = \int_a^b f(x)dx$ whence $\int_a^b f(x)dx = \lim\limits_{n\to\infty} \int_{[a;b]} l_{z_n} d\lambda = \lim\limits_{n\to\infty} \int_{[a;b]} u_{z_n} d\lambda$. Since $(u_{z_n})_{n\in\mathbb{N}}$ decreases, $(l_{z_n})_{n\in\mathbb{N}}$ increases, $(u_{z_n} - l_{z_n})_{n\in\mathbb{N}}$ is a decreasing sequence bounded below by 0 such that due to the **completeness** of the **real numbers** there must be a limit $\lim\limits_{n\in\mathbb{N}} (u_{z_n} - l_{z_n}) \geq 0$. According to 4.9 this limit function is **measurable** and from 5.13 follows $0 \leq \int \lim\limits_{n\in\mathbb{N}} (u_{z_n} - l_{z_n}) \leq \liminf\limits_{n\in\mathbb{N}} (U_{z_n} - L_{z_n}) = 0$ whence λ-a.e. $\lim\limits_{n\in\mathbb{N}} (u_{z_n} - l_{z_n}) = 0$ due to 5.6.3. Since λ-a.e. $l_{z_n} \leq f \leq u_{z_n}$ we infer that λ-a.e. $\lim\limits_{n\in\mathbb{N}} l_{z_n} = f$. Finally we apply **dominated convergence** 5.14 with the majorant u_{z_0} and obtain $\int_a^b f(x)dx = \lim\limits_{n\in\mathbb{N}} \int_{[a;b]} l_{z_n} d\lambda = \int_{[a;b]} \left(\lim\limits_{n\in\mathbb{N}} l_{z_n}\right) d\lambda = \int_{[a;b]} f d\lambda$.
2. Follows directly from the preceding theorem 5.24.

Note: In essential, 5.12, 5.13 and 5.14 assert the **continuity of the Bochner and Lebesgue integral**s regarding **pointwise** esp. **μ-a.e. convergence** whereas the **Riemann integral** is only continuous with reference to **uniform convergence** (cf.[7, Th 7.16]).

The classical definition of the **Lebesgue** integral is restricted to **positive** functions such that the Lebesgue integral of real functions requires **separate computing of positive and negative parts** entailing the failure of this method in the case of certain integrands with **alternating** signs like e.g. $\int \frac{\sin(x)}{x} dx = \lim\limits_{n\to\infty} \int_{-n}^{n} \frac{\sin(x)}{x} dx = \pi$. (cf. 16.17.2). Theorem 5.24 does not work with the Lebesgue integral.

6 Lebesgue spaces

6.1 Convex functions: A real function $f : \mathbb{R} \to \mathbb{R}$ is **convex** on the open interval $]a;b[$ iff $f(s) \leq f(r) + (s-r) \cdot \frac{f(t)-f(r)}{t-r} = f(t) - (t-s) \cdot \frac{f(t)-f(r)}{t-r}$ resp. $\frac{f(t)-f(s)}{t-s} \geq \frac{f(t)-f(r)}{t-r} \geq \frac{f(s)-f(r)}{s-r}$ for every $a < r < s < t < b$. Every convex function is **continuous** and in particular **Borel-measurable** since for $s \in]a;b[$ and w.l.o.g. $\min\{1; b-s\} > \epsilon > 0$ we have $|f(r) - f(s)| < |r-s| \cdot \frac{|f(s+\epsilon)-f(s)|}{\epsilon} < \epsilon$ for every $|r-s| < \delta := \frac{\epsilon^2}{\max\{1;|f(s+\epsilon)-f(s)|\}}$.

6.2 Jensen's inequality: For every integrable $g : A \to]a;b[\subset \mathbb{R}$ with $A \subset X$ and $\mu(A) < \infty$ on a measure space $(X; \mathcal{A}, \mu)$ and every convex $f :]a;b[\to \mathbb{R}$ we have $f\left(\frac{1}{\mu(A)} \int_A g d\mu\right) \leq \frac{1}{\mu(A)} \int_A (f \circ g)\, d\mu$.

Proof: For $s := \frac{1}{\mu(A)} \int_A g d\mu$ we have $a < s < b$ and due to 6.1 also $\beta := \sup\limits_{a<r<s} \frac{f(s)-f(r)}{s-r} \leq \frac{f(t)-f(s)}{t-s}$ for all $s < t < b$, hence $f(s) + \beta(t-s) \leq f(t)$ resp. $f(s) + \beta(g(x) - s) \leq f(g(x))$. All summands of this inequality are integrable over A such that on account of the monotonicity of the integral we can infer $\mu(A) \cdot f(s) \leq \int_A (f \circ g)\, d\mu$ and hence the assertion.

6.3 Applications: Choosing $A = \{p_1; ...; p_n\} \subset [0; \infty[$ and $\mu(\{p_i\}) = \alpha_i$ with $\mu(A) = \sum\limits_{i=1}^{n} \alpha_i = 1$ as well as $g(p_i) = \ln(x_i)$ and $f(x) = \exp(x)$ Jensen's inequality yields the following very useful special cases:

1. $x_1^{\alpha_1} \cdot ... \cdot x_n^{\alpha_n} \leq \alpha_1 x_1 + ... + \alpha_n x_n$
2. $(x_1 \cdot ... \cdot x_n)^{\frac{1}{n}} \leq \frac{1}{n}(x_1 + ... + x_n)$ (**geometric** and **arithmetic mean** for $\alpha_i := \frac{1}{n}$)
3. $F \cdot G \leq \frac{1}{p} F^p + \frac{1}{q} G^q$ for $\frac{1}{p} + \frac{1}{q} = 1$ with **equality iff** $F^p = G^q$ for $\alpha_1 = \frac{1}{p}; \alpha_2 = \frac{1}{q}; x_1 = F^p; x_2 = G^q$.

6.4 Hölder and **Minkowski inequalities**: For any positive Borel measurable $f, g : X \to Y$ from a measure space $(X; \mathcal{A}, \mu)$ into a Banach space $(Y; |\,|)$ and $\frac{1}{p} + \frac{1}{q} = 1$ resp. $p + q = p \cdot q$ with $\|f\|_p := \left(\int |f|^p \, d\mu\right)^{\frac{1}{p}}$ we have

1. $\|fg\| \leq \|f\|_p \cdot \|g\|_q$ (**Hölder** resp. **Schwarz** for $p = q = 2$) with equality iff μ-a.e. $\frac{f(x)}{\|f\|_p} = \frac{g(x)}{\|g\|_q}$.
2. $\|f + g\| \leq \|f\|_p + \|g\|_p$ (**Minkowski**) with equality iff μ-a.e. $\frac{f(x)}{\|f\|_p} = \frac{g(x)}{\|g\|_p} = \frac{f(x)+g(x)}{\|f+g\|_p}$.

Proof: The integrand is measurable on account of 4.6. For one of the integrals disappearing 5.11 tells us that the integrands $f \cdot g, f + g, f$ and g will disappear μ-a.e. too such that we have equality in this case. Therefore we can assume all integrals > 0 in the following proof.

1. With $F := \frac{|f|}{\|f\|_p}$ resp. $G := \frac{|g|}{\|g\|_q}$ in 6.3.3 an integration yields $\int (F \cdot G)\, d\mu \leq \frac{1}{p} + \frac{1}{q} = 1$ and hence the assertion. In particular $f \cdot g$ is integrable if f^p and g^q are integrable.
2. Applying 1. twice to $(f+g)^p = f \cdot (f+g)^{p-1} + g \cdot (f+g)^{p-1}$ and observing $q(p-1) = p$ we obtain $\|f+g\|_p^p \leq \|f\|_p \cdot \left\|(f+g)^{p-1}\right\|_q + \|g\|_p \cdot \left\|(f+g)^{p-1}\right\|_q = \left(\|f\|_p + \|g\|_p\right) \cdot \|f+g\|_p^{\frac{q}{p}}$. Substituting $p - \frac{p}{q} = 1$ yields the assertion. The convexity of t^p provides the inequality $\left(\frac{f+g}{2}\right)^p \leq \frac{f^p+g^p}{2}$, i.e. the integrability of f^p and g^p entails the integrability of $(f+g)^p$.

6.5 L^p-spaces: For $1 \leq p < \infty$ and any $f : X \to Y$ from a measure space $(X; \mathcal{A}; \mu)$ into a Banach space $(Y; |\,|)$ the expressions $\|f\|_p := \left(\int |f|^p d\mu\right)^{\frac{1}{p}}$ resp. $\|f\|_\infty := \inf\{0 < \alpha < \infty : \mu(|f| > \alpha) = 0\}$ define a **pseudonorm** (cf. [9, 21.1]) on the **vector space** $\mathcal{L}^p(\mu) := \left\{f : X \to Y : \|f\|_p :< \infty\right\}$. The **absolute homogeneity** follows from the **linearity** 5.5 whereas the **triangle inequation** is provided by the **Hölder inequality** 6.4.2. $\mathcal{L}^1(\mu)$ contains the **Lebesgue integrable** functions and $\mathcal{L}^\infty(\mu)$ is the set of all μ-a.e. **bounded** and measurable functions furnished with the **supremum norm** $\|\|_\infty$. Analogously to 5.1 resp. 5.11 the contraction to the **quotient space** $L^p(\mu) := \mathcal{L}^p / \sim$ defined by the **equivalence relation** $f \sim g \Leftrightarrow \mu(f \neq g) = 0$ makes $\|\|_p$ a **norm**. Convergence with respect to $\|\|_p$ is called in the **p-th mean**. On account of 5.6 all $f \in \mathcal{L}^p$ are μ-a.e. finite for $1 \leq p \leq \infty$.

6.6 Relations between L^p-spaces: For $1 \leq p, q \leq \infty$ we have

1. For μ **bounded above** , i.e. $\mu(A) < \alpha \forall A \in \mathcal{A}$ we have $p < q \Rightarrow L^p \supset L^q$.
2. For μ **bounded below**, i.e. $\mu(A) > \alpha \forall A \in \mathcal{A}$ we have $p < q \Rightarrow L^p \subset L^q$.

Note : The Lebesgue measure $\mu = \lambda^n$ satisfies none of the above requested conditions such that $L^p(\lambda^n)$ cannot be linearly ordered by inclusion. E.g. owing to 5.21.2 on the one hand for $g_n(x) := \min\{1; |x|^{-n}\}$ we have $g_n \in L^p \Leftrightarrow n > \frac{1}{p}$ but in the other hand fro $h_n(x) := \max\{1; |x|^{-n}\}$ the relation $g_n \in L^p \Leftrightarrow n < \frac{1}{p}$ holds.

Proof:

1. With $p = \frac{r}{s} \geq 1$, $f = h^s$ and $g = 1$ Hölder 6.4.1 yields $\int |h|^s d\mu \leq \left(\int |h|^r d\mu\right)^{\frac{s}{r}} \cdot \left(\int 1 d\mu\right)^{\frac{s-s}{r}}$ resp. $\|h\|_s = \left(\int |h|^s d\mu\right)^{\frac{1}{s}} \leq \left(\int |h|^r d\mu\right)^{\frac{1}{r}} \cdot (\mu(X))^{\frac{1}{s}-\frac{1}{r}} = \|h\|_r \cdot (\mu(X))^{\frac{1}{s}-\frac{1}{r}}$ and hence the assertion.
2. On account of **Zorn's lemma** ([9, 14.2.4]) the set $\{|f| \geq 1\}$ possesses a **maximal cover** of measurable sets referring to **inclusion** resp. refinement and since $\mathcal{A}$ is closed under intersection this mus be a **partition**. Due to $\int |f|^p d\mu < \infty$ we have $\mu(f \geq 1) < \infty$ and since μ is **bounded below** this maximal partition consists of $n := \frac{\mu(f \geq 1)}{\alpha} + 1$ sets $(A_i)_{1 \leq i \leq n}$ with $\mu(A_i) > \alpha$. Owing to 5.3 for every $\epsilon > 0$ there is an elementary function $e = \sum_{i=1}^{n} \alpha_i \chi_{A_i} \leq f$ with $\int_{\{|f| \geq 1\}} e d\mu = \sum_{i=1}^{n} \alpha_i \mu(A_i) \geq \int_{\{|f| \geq 1\}} |f|^p d\mu - \epsilon \cdot \alpha$. Hence on the one hand for every $x \in A_i$ with $1 \leq i \leq n$ we have $|f|^p(x) \geq \alpha_i \Leftrightarrow |f|^q(x) \geq \alpha_i^{\frac{q}{p}}$ and on the other hand for every $1 \leq i \leq n$ there is an $x_i \in A_i$ with $\alpha_i \geq |f^p(x_i)| - \epsilon \Leftrightarrow \alpha_i^{\frac{q}{p}} \geq (|f^p(x_i)| - \epsilon)^{\frac{q}{p}} \geq |f^q(xxi)| - \epsilon \cdot \frac{q}{p} \cdot (|f^p(x_i)| - \epsilon)^{\frac{q}{p}-1} \geq |f^q(x_i)| - \epsilon \cdot \frac{q}{p} \cdot |f^{q-p}(x_i)|$ since the tangent $t(x + \epsilon) = x^{\frac{q}{p}} + \epsilon \cdot \frac{q}{p} \cdot x^{\frac{q}{p}-1}$ on the convex function $g(x) = x^{\frac{q}{p}}$ always runs below the curve, i.e. $g(x + \epsilon) = (x + \epsilon)^{\frac{q}{p}}$. Thus follows $\int_{\{|f| \geq 1\}} |f|^q d\mu < \sum_{i=1}^{n} \left(\alpha_i^{\frac{q}{p}} + \epsilon \cdot \frac{q}{p} \cdot |f^{q-p}(x_i)\right) \chi_{A_i} < \infty$ and also on the whole set $\int |f|^q d\mu = \int_{\{|f|<1\}} |f|^q d\mu + \int_{\{|f| \geq 1\}} |f|^q d\mu \leq \int_{\{|f|<1\}} |f|^p d\mu + \int_{\{|f| \geq 1\}} |f|^q d\mu < \infty$.

6.7 Completeness: Every L^p-**Cauchy sequence**$(f_n)_{n \in \mathbb{N}} \subset L^p(\mu)$ with $1 \leq p \leq \infty$ converges in the **p-th mean** to a $f \in L^p(\mu)$. Hence $L^p(\mu)$ is a **Banach space**.

Proof: For a Cauchy sequence $(f_n)_{n \in \mathbb{N}} \subset L^p(\mu)$ with $p < \infty$ exists a partial sequence $(f_{n_i})_{i \in \mathbb{N}}$ with $\|f_{n_{i+1}} - f_{n_i}\|_p < \frac{1}{2^{i+1}}$ which entails $\left\|\sum_{i=0}^{k} |f_{n_{i+1}} - f_{n_i}|\right\|_p \leq 1$ due to 6.4.2, hence $\left\|\sum_{i=0}^{\infty} |f_{n_{i+1}} - f_{n_i}|\right\|_p \leq 1$ owing to 5.12 and finally **μ-a.e.** $g := \sum_{i=0}^{\infty} |f_{n_{i+1}} - f_{n_i}| < \infty$ according to 5.6.2. Since Y is **complete** the sequence $(f_{n_i})_{i \in \mathbb{N}} = \sum_{k=1}^{i} (f_{n_k} - f_{n_{k+1}})$ **μ-a.e. converges** to an $f = \lim_{i \to \infty} f_{n_i} = \sum_{i=0}^{\infty} (f_{n_{i+1}} - f_{n_i})$ with $|f| < g$. On account of the **completeness** of μ (cf. 3.9) we can define $f(x) = 0$ on the remaining null set $\{|f| = \infty\}$. According to the hypothesis for every $\epsilon > 0$ there is a $j \in \mathbb{N}$ with $\left\|f_m - f_{n_j}\right\|_p < \epsilon$ for all $m \geq n_j$ whence **Fatou's lemma** 5.13 yields $\left(\liminf_{m \geq n_j} \left|f_m - f_{n_j}\right|\right)^p = \liminf_{m \geq n_j} \left|f_m - f_{n_j}\right|^p \in L_1(X; \mathbb{R})$ with $\int \left(\liminf_{m \geq n_j} \left|f_m - f_{n_j}\right|^p\right) d\mu \leq \liminf_{m \geq n_j} \int \left|f_m - f_{n_j}\right|^p d\mu < \epsilon^p$ Since μ-a.e. $f = \lim_{i \to \infty} f_{n_i}$ we have μ-a.e $\liminf_{m \geq n_j} \left|f_m - f_{n_j}\right| = \left|f - f_{n_j}\right|$ so that $\left\|f - f_{n_j}\right\|_p = \left\|\liminf_{m \geq n_j} \left|f_m - f_{n_j}\right|\right\|_p < \epsilon$, i.e. the subsequence $(f_{n_i})_{i \in \mathbb{N}}$ and hence the entire Cauchy sequence $(f_n)_{n \in \mathbb{N}}$ (cf. [9, 14.1.2]) converges in the p-th mean to f. On account of $\|f\|_p \leq \|f - f_n\|_p + \|f_n\|_p < \infty$ we have $f \subset L^p(\mu)$.

For $p = \infty$ let $A := \bigcup_{m,n \in \mathbb{N}} \left(\{|f_m - f_n| > \|f_m - f_n\|_\infty\} \cup \{|f_m| > \|f_m\|_\infty\}\right)$. Then we have $\mu(A) = 0$ and $(f_n)_{n \in \mathbb{N}}$ is a **Cauchy sequence** on $X \setminus A$ referring to the **supremum norm.** Due to the **completeness** of Y it converges uniformly and in particular with reference to $\|\|_\infty$ to a bounded function $|f| < \lim_{n \to \infty} \|f_n\|_\infty$. Again we define $f(x) = 0$ for $x \in A$ and finally obtain $f \subset L^\infty(\mu)$.

6.8 Notes:

1. The sequence $(f_n)_{n\in\mathbb{N}} \subset L^p(\lambda)$ with $f_n := \chi_{A_n}$ for $A_n := \left[\frac{n}{2^k}; \frac{n+1}{2^k}\right]$ with $k(n) = \min\left\{k : n < 2^k\right\}$ shows that in general the μ-a.e. convergence cannot be extended to the entire sequence:: $\lim\limits_{n\to\infty} \|f_n\|_p = \lim\limits_{n\to\infty} \left(\lambda\left(\left[\frac{n}{2^k}; \frac{n+1}{2^k}\right]\right)\right)^{\frac{1}{p}} = \lim\limits_{n\to\infty} 2^{-\frac{k(n)}{p}} = \lim\limits_{n\to\infty} \frac{1}{n^{1/p}} = 0$ but for every $x \in \left[\frac{1}{2}; 1\right]$ and $k \geq 1$ there is an $n \in \mathbb{N}$ with $x \in \left[\frac{n}{2^k}; \frac{n+1}{2^k}\right]$ such that $(f_n)_{n\in\mathbb{N}}$ does not converge for any $x \in \left[\frac{1}{2}; 1\right]$ whereas the partial sequence $(f_{2^k})_{k\in\mathbb{N}}$ converges for every $x \neq \frac{1}{2}$..
2. $L^2(\mu)$ is a **Hilbert spac**e with the **inner product** $\langle f, g\rangle := \int f\overline{g}d\mu$ and the **norm** $\|f\| := \langle f, g\rangle^{\frac{1}{2}} := \left(\int f\overline{f}d\mu\right)^{\frac{1}{2}} = \left(\int |f|^2 d\mu\right)^{\frac{1}{2}}$.

6.9 Convergence in the p-th mean, in measure and μ-a.e: Every sequence $(f_n)_{n\in\mathbb{N}} \subset L^p(\mu)$ with $1 \leq p < \infty$ converging **in the p-th mean** to an $f \in L^p(\mu)$ converges in measure to f. Also there exists a **subsequence** $(f_{n_k})_{k\in\mathbb{N}}$ converging **μ-a.e.** to $f$and for every $\epsilon > 0$ there is a set $Z_\epsilon \subset X$ with measure $\mu(Z_\epsilon) < \epsilon$ such that $(f_{n_k})_{k\in\mathbb{N}}$ converges **absolutely** and **uniformly** on $X \setminus Z_\epsilon$.

Proof: The **convergence in measure** follows at once from $\epsilon{\cdot}\mu(|f - f_n| \geq \epsilon) = \epsilon^p{\cdot}\mu(|f - f_n|^p \geq \epsilon^p) \leq \int |f - f_n|^p d\mu$ According to the hypothesis for every $k \geq 1$ there is an $n_k \geq n_{k-1} \in \mathbb{N}$ such that $\|f - f_{n_k}\|_p \leq \frac{1}{2^{2k}}$. Then for $Y_k = \left\{|f - f_{n_k}|^p \geq \frac{1}{2^k}\right\}$ we have $\frac{1}{2^k}\mu(Y_k) = \int_{Y_n} \frac{1}{2^k} d\mu \leq \int_X |f - f_{n_k}|^p d\mu \leq \frac{1}{2^{2k}}$ whence $\mu(Y_k) \leq \frac{1}{2^k}$. Hence $\mu(Z_m) \leq \frac{1}{2^{m-1}}$ for $Z_m = \bigcup\limits_{k=m}^{\infty} Y_k$ and $|f(x) - f_{n_k}(x)|^p < \frac{1}{2^k}$ for every $x \in X \setminus Z_m$ resp. $k \geq m$ such that $(f_{n_k})_{k\geq m}$ converges to f **absolutely** and **uniformly** on on $X \setminus Z_m$ as well as **pointwise** on $X \setminus \bigcap\limits_{m=1}^{\infty} Z_m$ with $\mu\left(\bigcap\limits_{m=1}^{\infty} Z_m\right) = 0$.

6.10 Lebesgue's dominated convergence theorem for L^p-spaces: A sequence $(f_n)_{n\in\mathbb{N}} \subset L^p(X; Y)$ **converging μ-a. e.** to some f **converges in the p-th mean** to $f \in L^p(X; Y)$ iff there is an **integrable majorant** $g \in L^p(X; \mathbb{R}_0^+)$ such that for every $n \in \mathbb{N}$ and μ-a.e. we have $|f_n| \leq g$.

Proof: For every $K \in \mathbb{N}$ the **increasing** sequence $\left(\sup\limits_{k\leq m;n\leq l} |f_n - f_m|^p\right)_{l>k}$ is bounded by $|f_n - f_m|^p \leq 2^p g^p$ and hence has bounded integrals $\int \left(\sup\limits_{k\leq m;n\leq l} |f_n - f_m|^p\right) d\mu \leq 2^p \int g^p d\mu = 2^p \|g\|_p^p$. According to the **monotone convergence theorem** 5.12 we conclude $\int \left(\sup\limits_{k\leq m;n} |f_n - f_m|^p\right) d\mu \leq 2^p \int g^p d\mu$ for every $k \in \mathbb{N}$. Hence we can apply the monotone convergence theorem a second time to the **decreasing** sequence $\left(\sup\limits_{k\leq m;n} |f_n - f_m|^p\right)_{k\geq 1}$ converging μ-a.e. to 0 to obtain $\lim\limits_{k\to\infty} \int \left(\sup\limits_{k\leq m;n} |f_n - f_m|^p\right) d\mu = 0$. Hence $(f_n)_{n\in\mathbb{N}}$ is L^p-Cauchy and due to the **completeness** 6.7 of $L^p(X; Y)$ it converges **in the p-th mean** to an $f^{\#} \in L^p(X; Y)$ coninciding μ-a.e. with f according to 5.10.

Note: The proofs of the preceding two theorems is completely analogous to those of the corresponding statements 5.10 resp. 5.13 for L^1 with the small but essential difference that the generalized theorems 6.9 resp. 6.10 require the completeness 6.7 of L^p which like the dominated convergence for L^1 is based in the completeness5.9 of L^1. Alas the proof of this latter property depends on an **elementary approximation by step functions and cannot be duplicated for L^p**.

6.11 Theorem:

1. For $1 \leq p < \infty$ we have $\overline{\left(\mathcal{S}(\mathcal{A}; Y); \|\|_p\right)} = L^p(X; Y)$
2. For a **finite** measure space $(X; A; \mu)$ and a **finite dimensional Banach** space $(K^n; |\,|)$ we have $\overline{(\mathcal{S}(\mathcal{A}; K^n); \|\|_\infty)} = L^\infty(X; K^n)$.

Proof:

1. According to 5.5 for every $f \in L^p(\mu)$ resp. $f^p \in L^1(\mu)$ and $p < \infty$ there is a sequence $(\varphi_n)_{n\in\mathbb{N}} \subset \mathcal{S}(\mathcal{A};Y)$ of **step functions** converging μ-a.e. to f. The truncated version $\psi_n(x) = \left\{\begin{smallmatrix} |\varphi_n(x)| & \text{for } |\varphi_n(x)| \leq 2|f(x)| \\ 0 & \text{else} \end{smallmatrix}\right.$ still converges μ-a-e- to f and satisfies the hypothesis for 6.10 with the majorant $2|f| \in L^p(X;\mathbb{R}_0^+)$ which yields the convergence in tne p-th mean and hence the assertion.
2. For$|f| \leq N$; $M \geq 1$ and $\mathbf{k} = (k_1;...;k_n) \in K_M = [-NM;NM[^n \subset \mathbb{Z}^n$ we define $\varphi_M = \sum_{\boldsymbol{k}\in K_M} \frac{\boldsymbol{k}}{M}\chi_{A_{\boldsymbol{k},M}} \in \mathcal{S}(\mathcal{A};K^n)$ with $A_{\boldsymbol{k},M} = f^{-1}\left[\prod_{i=1}^{n}\left[\frac{k_i}{M};\frac{k_i+1}{M}\right[\right] \in \mathcal{A}$ and $\mu(A_{\boldsymbol{k},M})$ such that $\|f-\varphi_M\|_\infty \leq \frac{\sqrt{n}}{M}$.

6.12 Lemma: For an **integrable** $f \in L^p(X;Y)$ and every $\epsilon > 0$ there is a $\delta > 0$ such that for every $E \in \mathcal{A}$ with $\mu(E) < \delta$ we have $\int_E |f|\,d\mu < \epsilon$.

Proof: The sequence $(\varphi_n)_{n\geq 1}$ with $\varphi_n(x) = \left\{\begin{smallmatrix} |f(x)|, & \text{for } |f(x)| \leq n \\ n & \text{else} \end{smallmatrix}\right.$ satisfies the conditions for **monotone convergence** 5.12 such that $\lim_{n\to\infty}\int \varphi_n d\mu = \int |f|\,d\mu$. Hence for $\epsilon > 0$ there is an $n_0 \geq 1$ such that $\int (|f| - \varphi_n)\,d\mu < \frac{\epsilon}{2}$. Since for $\delta = \frac{\epsilon}{2n}$ and every $E \in \mathcal{A}$ with $\mu(E) < \delta$ we have $\int_E \varphi_n d\mu \leq n\cdot\mu(E) = \frac{\epsilon}{2}$ it follows that $\int_E |f|\,d\mu \leq \int_E (|f| - \varphi_n)\,d\mu + \int_E \varphi_n d\mu \leq \epsilon$.

6.13 Vitali's **convergence theorem**: A sequence $(f_n)_{n\geq 1} \subset L^p(\mu)$ **converging μ-a.e.** for $1 \leq p < \infty$ to some f **also converges in the p-th mean** to $f \in L^p(\mu)$ **iff** for every $\epsilon > 0$

1. there is an $A_\epsilon \in \mathcal{A}$ with $\mu(A_\epsilon) < \infty$ and $\int_{X\setminus A_\epsilon} |f_n|^p\,d\mu < \epsilon$ for all $n \geq 1$.
2. there is a $\delta > 0$ such that for every $E \in \mathcal{A}$ with $\mu(E) < \delta$ we have $\int_E |f_n|^p\,d\mu < \epsilon$ for all $n \geq 1$.

Proof:

$\Rightarrow$: 1.: Due to the hypothesis for $\epsilon > 0$ there is an $n_0 \geq 1$ such that $\int |f_n - f|^p\,d\mu < \epsilon$ for all $n \geq n_0$. Owing to 5.12 with $|f|^p = \sup_{m\geq 1} |f|^p \cdot \chi_{\{|f|^p > \frac{1}{m}\}}$ and $f \in L^p(\mu)$ there is an $m_0 \geq 1$ with $\int |f|^p \cdot \chi_{\{|f|^p \leq \frac{1}{m}\}} d\mu = \int |f|^p\,d\mu - \int |f|^p \cdot \chi_{\{|f|^p > \frac{1}{m}\}} d\mu < \epsilon$ and $\mu\left(|f|^p \leq \frac{1}{m}\right) \leq \mu\left(|f|^p \leq \frac{1}{m}\right) \leq \int |f|^p\,d\mu < \infty$ for all $m \geq m_0$. For those f_n with $1 \leq n \leq n_0$ we use the same reasoning as above to find an $m_1 \geq m_0$ such that the sets $B_\epsilon = \left\{|f|^p > \frac{1}{m_1}\right\} \in \mathcal{A}$ resp. $C_\epsilon = \left\{\max_{1\leq n<n_0} |f_n|^p > \frac{1}{m_1}\right\} \in \mathcal{A}$ with $\mu(X \setminus B_\epsilon), \mu(X \setminus C_\epsilon) < \infty$ satisfy $\int_{X\setminus B_\epsilon} |f|^p\,d\mu < \epsilon$ resp. $\int_{X\setminus C_\epsilon} |f_n|^p\,d\mu < \epsilon$ for all $1 \leq n < n_0$. For $A_\epsilon = B_\epsilon \cup C_\epsilon$ **Minkowski's inequality** 6.4.2 yields $\left(\int_{X\setminus A_\epsilon} |f_n|^p\,d\mu\right)^{\frac{1}{p}} \leq \left(\int_{X\setminus A_\epsilon} |f_n - f|^p\,d\mu\right)^{\frac{1}{p}} + \left(\int_{X\setminus A_\epsilon} |f|^p\,d\mu\right)^{\frac{1}{p}} < \epsilon^{\frac{1}{p}} + \epsilon^{\frac{1}{p}}$ resp. $\int_{X\setminus A_\epsilon} |f_n|^p\,d\mu < 2^p\epsilon$ for all $n \geq 1$.

2.: For a given $\epsilon > 0$ choose $n_0 \geq 1$ as in 1. such that $\int |f_n - f|^p\,d\mu < \epsilon$ for all $n \geq n_0$. According to the preceding lemma 6.12 there is a $\delta > 0$ such that for all $E \in \mathcal{A}$ with $\mu(E) < \delta$ we have $\int_E |f|^p\,d\mu < \epsilon$ resp. $\int_E |f_n|^p\,d\mu < \epsilon$ for all $1 \leq n < n_0$. As in 1. **Minkowski's inequality** 6.4.2 yields the desired estimate $\int_E |f_n|^p\,d\mu < 2^p\epsilon$ for the remaining $n \geq n_0$.

$\Leftarrow$: According to 1. for $\epsilon > 0$ there is an $A_\epsilon \in \mathcal{A}$ with $\mu(A_\epsilon) < \infty$ such that $\int_{X\setminus A_\epsilon} |f_n|^p\,d\mu < \epsilon$ for all $n \geq 1$ which allows the recourse to **Fatou's lemma** 5.13 to yield the estimate $\int_{X\setminus A_\epsilon} |f|^p\,d\mu \leq \liminf_{n\geq 1} \int_{X\setminus A_\epsilon} |f_n|^p\,d\mu < \epsilon$. As above we use **Minkowski's inequality** 6.4.2 to obtain $\left(\int_{X\setminus A_\epsilon} |f - f_n|^p\,d\mu\right)^{\frac{1}{p}} \leq \left(\int_{X\setminus A_\epsilon} |f_n|^p\,d\mu\right)^{\frac{1}{p}} + \left(\int_{X\setminus A_\epsilon} |f|^p\,d\mu\right)^{\frac{1}{p}} < 2\epsilon^{\frac{1}{p}}$. According to 2. resp. **Egorov's theorem** 4.15 for every $\delta > 0$ there is a $B_\delta \in \mathcal{A}$ as well as an $n_0 \geq 1$ with $\mu(B_\delta) < \delta$ such that $|f(x) - f_n(x)|^p < \epsilon$ for every $x \in A_\epsilon \setminus B_\delta$ and hence $\left(\int_{A_\epsilon\setminus B_\delta} |f - f_n|^p\,d\mu\right)^{\frac{1}{p}} < \epsilon^{\frac{1}{p}}$ for every $n \geq n_0$. On the set B_δ we follow the reasoning for $X \setminus A_\epsilon$ from above to find $\int_{B_\delta} |f|^p\,d\mu \leq \liminf_{n\geq 1} \int_{B_\delta} |f_n|^p\,d\mu < \epsilon$ with **Fatou** and finally $\left(\int_{B_\delta} |f - f_n|^p\,d\mu\right)^{\frac{1}{p}} \leq \left(\int_{B_\delta} |f_n|^p\,d\mu\right)^{\frac{1}{p}} + \left(\int_{B_\delta} |f|^p\,d\mu\right)^{\frac{1}{p}} < 2\epsilon^{\frac{1}{p}}$ with **Minkowski**. Combining our results over $X \setminus A_\epsilon$, $A_\epsilon \setminus B_\delta$ and B_δ we obtain $\left(\int_X |f - f_n|^p \, d\mu\right)^{\frac{1}{p}} < 5\epsilon^{\frac{1}{p}}$ for $n \geq n_0$ and hence the assertion.

7 Product spaces

7.1 Initial σ-algebra : The **initial σ-algebra** $\sigma(f_i : i \in I) := \sigma\left(\bigcup_{i\in I} f_i^{-1}(\mathcal{A}_i)\right)$ on a set X referring to the functions $f_i : X \to (Y_i; \mathcal{A}_i)$ with $i \in I$ is the smallest σ-algebra on X such that all f_i are **measurable**. This concept is closely related to that of the **initial topology**, cf. [**?**, 4.1].

7.2 Trace of a measure space: The **trace σ-algebra** $\mathcal{A}_B = \sigma(i)$ on a subset $B \subset X$ of a measure space $(X; \mathcal{A}; \mu)$ ist the **initial σ-algebra** with reference to the **canonical injection** $i : B \to X$. On account of $i^{-1}[A] = A \cap B$ the measurable sets in B simply are the **intersections of the measurable sets in** A in X with B. The **trace of the measure** μ is its **restriction** $\mu|_B$.

7.3 Product-σ-algebra The **product-σ-algebra** $\mathcal{A}_I = \bigotimes_{i\in I} \mathcal{A}_i = \sigma(\pi_i : i \in I)$ on the productt $X_I = \prod_{i\in I} X_i$ of the measurable spaces $(X_i; \mathcal{A}_i)_{i\in I}$ is the initial σ-Algebra with reference to the **projections** $\pi_i : X_I \to X_i$. A mapping $f : Y \to X_I$ iis measurable iff the inverse images $f^{-1}\left[\pi_i^{-1}[A_i]\right] = (\pi_i \circ f)^{-1}[A_i]$ if measurable sets in X_i are measurable in $(Y; \mathcal{A})$. Hence f is measurable iff every **component** $\pi_i \circ f : (Y; \mathcal{A}) \to (X_i; \mathcal{A}_i)$ is measurable. Due to 4.4 the product σ-algebra induced by the families $\mathcal{E}_i \subset \mathcal{P}(X_i)$ with $i \in I$ is $\bigotimes_{i\in I} \sigma(\mathcal{E}_i) = \sigma\left(\bigcup_{i\in I} \pi_i^{-1}(\sigma(\mathcal{E}_i))\right) = \sigma\left(\bigcup_{i\in I} \sigma\left(\pi_i^{-1}(\mathcal{E}_i)\right)\right) = \sigma\left(\bigcup_{i\in I} \pi_i^{-1}(\mathcal{E}_i)\right)$.

7.4 Measurable rectangles and cylinder sets:

1. The family $\mathcal{S}_I = \left\{\bigcap_{j\in J} \pi_j^{-1}[A_j] = \prod_{j\in J} A_i \times \prod_{i\in I\setminus J} X_i : A_j \in \mathcal{A}_j,\ j \in J \subset I \wedge J\,\text{finite}\right\}$ of **measurable rectangles is closed under intersections** and a **basis** for the product-σ-algebra $\mathcal{A}_I = \sigma(\mathcal{S}_I)$.
2. For $J \subset K \subset I$ the **projections** $\pi_K^J : (X_J; \mathcal{A}_J) \to (X_K; \mathcal{A}_K)$ are measurable and for $J \cap K = \emptyset$ we have $\mathcal{A}_{J\cup K} = \mathcal{A}_J \otimes \mathcal{A}_K$.
3. The **algebra** $\mathcal{Z}_I = \left\{\pi_J^{-1}[A_J] = A_J \times \prod_{i\in I\setminus J} X_i : A_J \in \mathcal{A}_J, J \subset I \wedge J\,\text{finite}\right\}$ of **cylinder sets** also is a **π-basis** for the product-σ-algebra: $\mathcal{A}_I = \sigma(\mathcal{Z}_I)$. The cylinder sets $\mathcal{Z}_J = \sigma(\mathcal{S}_J)$ themselves are **σ-algebrae** with $\mathcal{Z}_J \subset \mathcal{Z}_K$ for $J \subset K$.
4. The family $\mathcal{A}_Z = \left\{\pi_J^{-1}[A_J] = A_J \times \prod_{i\in I\setminus J} X_i : A_J \in \mathcal{A}_J, J \subset I \wedge J\,\text{countable}\right\}$ of **countable cylinder sets** is a **σ-algebra** and **identical with** the **product-σ-algebra**: $\mathcal{A}_I = \mathcal{A}_Z$. Every measurable set A of a product-σ-algebra may depend from a **countable set of coordinates** in contrast to the **product topology** whose open sets are defined by **finitely many coordinates** (cf. [9, 4.2]).

Proof:

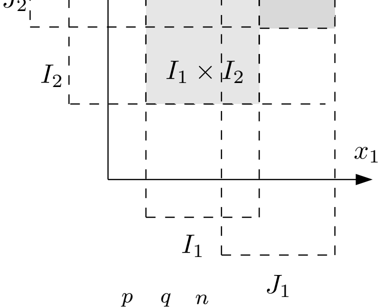
1. $\mathcal{S}_I$ is **closed under intersection** since for **finite** $J, K \subset I$ and $A_j \in \mathcal{A}_j$ with $j \in J$ resp. $B_k \in \mathcal{A}_k$ with $k \in K$ we have $\left(\bigcap_{j\in J} \pi_j^{-1}[A_j]\right) \cap \left(\bigcap_{k\in K} \pi_k^{-1}[B_k]\right) = \left(\bigcap_{j\in J\setminus K} \pi_j^{-1}[A_j]\right) \cap \left(\bigcap_{l\in J\cap K} \pi_l^{-1}[A_l \cap B_l]\right) \cap \left(\bigcap_{k\in K\setminus J} \pi_k^{-1}[B_k]\right) \in \mathcal{S}_I$ with $A_l \cap B_l \in \mathcal{A}_l$ for $l \in J \cap K$. Due to $\left\{\pi_i^{-1}[A_i] : A \in \mathcal{A}_i, i \in I\right\} \subset \mathcal{S}_I$ we have $\mathcal{A}_I = \sigma\left(\left\{\pi_i^{-1}[A_i] : A_i \in \mathcal{A}_i, i \in I\right\}\right) \subset \sigma(\mathcal{S}_I)$ and on account of $\mathcal{S}_I \subset \mathcal{A}_I$ the converse follows: $\sigma(\mathcal{S}_I) \subset \mathcal{A}_I$.
2. The **projections** are measurable since with $\bigcap_{k\in K}\left(\pi_k^K\right)^{-1}[A_k] \in \mathcal{S}_K$ for $A_k \in \mathcal{A}_k$ and $k \in K$ we have $\left(\pi_K^J\right)^{-1}\left(\bigcap_{k\in K}\left(\pi_k^K\right)^{-1}[A_k]\right) = \bigcap_{k\in K}\left(\pi_K^J\right)^{-1}\left(\left(\pi_k^K\right)^{-1}[A_k]\right) = \bigcap_{k\in K}\left(\pi_k^J\right)^{-1}[A_k] \in \mathcal{A}_J$ and hence with 1. follows the assertion. The measurability of $\pi_J^{J\cup K}$ resp. $\pi_K^{J\cup K}$ entails $\mathcal{A}_{J\cup K} \supset \mathcal{A}_J \otimes \mathcal{A}_K$ and from 1. resp. $\mathcal{S}_{J\cup K} \subset \mathcal{A}_J \otimes \mathcal{A}_K$ follows the converse $\mathcal{A}_{J\cup K} = \sigma(\mathcal{S}_{J\cup K}) \subset \mathcal{A}_J \otimes \mathcal{A}_K$.

3. $\mathcal{Z}_I$ is an **algebra** since obviously $\emptyset, X \in \mathcal{A}_Z$ and for $\pi_J^{-1}[A_J], \pi_K^{-1}[A_K] \in \mathcal{Z}_I$ with $A_J \in \mathcal{A}_J, B_K \in \mathcal{A}_K$ and **finite** $J, K \subset I$ owing to 2. we have $\left(\pi_J^{J\cup K}\right)^{-1}[A_J]$, $\left(\pi_K^{J\cup K}\right)^{-1}[B_K] \in \mathcal{A}_{J\cup K}$. Hence the **intersection** $\left(\pi_J^{-1}[A_J]\right)\cap\left(\pi_K^{-1}[B_K]\right) = \pi_{J\cup K}^{-1}\left(\left(\left(\pi_J^{J\cup K}\right)^{-1}[A_J]\right)\cap\left(\pi_K^{J\cup K}\right)^{-1}[B_K]\right)$ $\in \mathcal{Z}_I$ and likewise the **union** are contained in $\mathcal{Z}_I$. Concerning the **complements** we consult e.g. [**?**, 9.2.3] to obtain $X_I \setminus \pi_J^{-1}[A_J] = \left(\pi_J^{-1}[X_J]\right)\setminus\left(\pi_J^{-1}[A_J]\right) = \pi_J^{-1}[X_J \setminus A_J] \in \mathcal{Z}_I$ since $X_J \setminus A_J \in \mathcal{A}_J$. On the one hand we have $\sigma(\mathcal{Z}_I) \subset \mathcal{A}_I$ since according to 2. we have $\mathcal{Z}_I \subset \mathcal{A}_I$. On the other hand 1. yields $\mathcal{A}_I = \sigma(\mathcal{S}_I) \subset \sigma(\mathcal{Z}_I)$ since $\mathcal{S}_I \subset \mathcal{Z}_I$. Again on account of 2. the families $\mathcal{Z}_J = \pi_J^{-1}(\mathcal{A}_J)$ are **σ-algebrae** whereas the linear order by inclusion on the family of cylinder sets follows from $\left(\pi_J^K\right)^{-1}(\mathcal{A}_J) \subset \mathcal{A}_K$ by application of π_K^{-1}. **Note**: The properties of a σ-algebra as well as the linear ordering by inclusion obviously extend to arbitrary index sets, notable countable ones, as shown below:

4. The family $\mathcal{A}_Z$ is again an **algebra** since the reasoning from 3. can be transferred to countable index sets. It is a **σ-algebra** since $\bigcup_{n\in\mathbb{N}} \pi_{J_n}^{-1}[A_{J_n}] = \pi_J^{-1}\left(\bigcup_{n\in\mathbb{N}}\left(\left(\pi_{J_n}^J\right)^{-1}[A_{J_n}]\right)\right) \in \mathcal{A}_Z$ with $\left(\pi_{J_n}^J\right)^{-1}[A_{J_n}] \in \mathcal{A}_J$ and countable $J = \bigcup_{n\in\mathbb{N}} J_n$. In particular we have $\mathcal{A}_Z \subset \sigma(\mathcal{Z}_I) = \mathcal{A}_I$. Conversely from $\mathcal{A}_Z \supset \mathcal{Z}_I$ and 3. follows the inclusion $\mathcal{A}_Z \supset \sigma(\mathcal{Z}_I) = \mathcal{A}_I$.

7.5 Product of Borel σ-algebrae and Borel σ-algebra of a product: The **product** $\mathcal{B}_I := \bigotimes_{i\in I}\sigma(\mathcal{O}_i)$ of the **Borel σ-algebrae** $\mathcal{B}_i$ of the topological spaces $(X_i;\mathcal{O}_i)_{i\in I}$ is the smallest σ-Algebra on $X = \prod_{i\in I} X_i$ or **initial σ-algebra** such that all **projections** $\pi_i : (X;\mathcal{B}_I) \to (X_i\mathcal{B}_i)$ are **measurable**. The π_i are **continuous** with reference to the **product topology** $\mathcal{O} = \bigotimes_{i\in I}\mathcal{O}_i$ (cf. [9, 4.2]) and hence due to 4.4 **measurable** with regard to the Borel σ-algebra $\mathcal{B} = \sigma\left(\bigotimes_{i\in I}\mathcal{O}_i\right)$, i.e. $\mathcal{B}_I = \bigotimes_{i\in I}\sigma(\mathcal{O}_i) \subset \sigma\left(\bigotimes_{i\in I}\mathcal{O}_i\right) = \mathcal{B}$. For **countable** I and **second countable** $\mathcal{O}_i$ the converse inclusion is also true since with **countable bases** $\mathcal{E}_i$ of $\mathcal{O}_i$ the basis $\mathcal{E} = \left\{\pi_i^{-1}(E_i) : E_i \in \mathcal{E}_i, i \in I\right\}$ of the **topology** is again countable and hence also generates the **Borel σ-algebra** $\mathcal{B} = \sigma(\mathcal{O}(\mathcal{E})) = \sigma(\mathcal{E})$ due to 1.2 such that from $\mathcal{E} \subset \mathcal{B}_I$ follows $\mathcal{B} = \sigma(\mathcal{E}) \subset \mathcal{B}_I$. Especially on **polish spaces** the two σ-algebrae coincide: $\mathcal{B} = \mathcal{B}_I$. For **Hausdorff** components according to [9, 7.10] the **separation axiom** T_2 extends to the product space and owing to **Tychonoff's theorem** (cf. [9, 9.9]) any **product** of **compact** sets is again **compact** and hence **Borel measurable** due to 1.2.

7.6 Finite product-σ-algebrae: If every basis $\mathcal{E}_j$ for $1 \leq j \leq m$ includes a **countable cover** $(E_{jn})_{n\in\mathbb{N}} \subset \mathcal{E}_j$ with $\bigcup_{n\in\mathbb{N}} E_{jn} = X_j$ the **product** $\bigotimes\limits_{j=1}^{m}\sigma(\mathcal{E}_j)$ is generated by the **intersections** $\bigcap\limits_{j=1}^{m}\pi_j^{-1}[E_j] = \prod\limits_{j=1}^{m}\mathcal{E}_j$ for all possible $E_j \in \mathcal{E}_j$: $\bigotimes\limits_{j=1}^{m}\sigma(\mathcal{E}_j) = \sigma\left(\prod\limits_{j=1}^{m}\mathcal{E}_j\right)$. Due to 7.4.1 on the one hand we have $\sigma\left(\prod\limits_{j=1}^{m}\mathcal{E}_j\right) \subset \bigotimes\limits_{j=1}^{m}\sigma(\mathcal{E}_j)$ and on the other hand $\pi_i^{-1}[E_i] = \bigcup_{n\in\mathbb{N}}\left(\prod\limits_{j=1}^{m}\pi_j^{-1}[E_{jn}] \cap \pi_i^{-1}[E_i]\right)$ $\in \sigma\left(\left\{\bigcap\limits_{j=1}^{m}\pi_j^{-1}[E_j] : E_j \in \mathcal{E}_j\right\}\right) = \sigma\left(\prod\limits_{j=1}^{m}\mathcal{E}_i\right)$ whence $\bigotimes\limits_{j=1}^{m}\sigma(\mathcal{E}_j) \subset \sigma\left(\prod\limits_{j=1}^{m}\mathcal{E}_j\right)$ on account of 4.1.



7.7 Finite products of Borel σ-algebrae: Analogously to the one dimensional case dealt with in 1.4 and according to 7.5 the **n-dimensional intervals** $\mathcal{I}^n := \left\{\prod\limits_{i=1}^{n}[a_i;b_i[\,:\, a_i \leq b_i \in \mathbb{R}\right\} \subset \mathbb{R}^n$ are are G_δ and hence $\mathcal{B}^n$-measurable. Also on account of $\left(\prod\limits_{i=1}^{n} I_i\right) \cap \left(\prod\limits_{i=1}^{n} J_i\right) = \prod\limits_{i=1}^{n}(I_i \cap J_i)$ with **intervals** $I_i; J_i \subset \mathbb{R}$ they are **closed under intersection**. Their **finite unions** form the **algebra** $\mathcal{F}^n$ of the **n-dimensional figures**: For $F = \bigcup\limits_{k=1}^{p}\prod\limits_{i=1}^{n} I_{k,i}; G = \bigcup\limits_{l=1}^{q}\prod\limits_{i=1}^{n} J_{l,i} \in \mathcal{F}^n$ we obviously have $F \cup G \in \mathcal{F}^n$; $F \cap G = \bigcup\limits_{k=1}^{p}\bigcup\limits_{l=1}^{q}\prod\limits_{i=1}^{n}(I_{k,i} \cap J_{l,i}) \in \mathcal{F}^n$ and $F \setminus G = \bigcup\limits_{k=1}^{p}\bigcup\limits_{l=1}^{q}\prod\limits_{i=1}^{n}(I_{k,i} \setminus J_{l,i}) \in$

$\mathcal{F}^n$. On account of $\prod\limits_{i=1}^{n}]a_i;b_i[= \bigcup_{k\in\mathbb{N}} \prod\limits_{i=1}^{n} \left[a_i+\frac{1}{k};b_i\right[$ oth the intervals and the figures generate the Borel σ-algebra: $\mathcal{B}^n := \bigotimes\limits_{i=1}^{n} \mathcal{B}_i = \sigma(\mathcal{I}^n) = \sigma(\mathcal{F}^n)$. Moreover due to 1.1.2 the Borel σ-algebra is generated by the **products** $\prod\limits_{i=1}^{n}]a_i;\infty[$.

8 Product measure

8.1 Lemma: For two measure spaces $(X_i;\mathcal{A}_i;\mu_i)$ with $i\in\{1;2\}$, every $A\subset\mathcal{A}_1\otimes\mathcal{A}_2$ and $x_1\in X_1, x_2\in X_2$ the **cuts** $A_{x_1} := \{x_2\in X_2 : (x_1;x_2)\in A\}$ resp. A_{x_2} are measurable with respect to $\mathcal{A}_2$ resp. $\mathcal{A}_1$.

Proof: Due to $(X\setminus Q)_{x_1} = X_2\setminus Q_{x_1}$ and $\left(\bigcup_{n\in\mathbb{N}} Q_n\right)_{x_1} = \bigcup_{n\in\mathbb{N}} (Q_n)_{x_1}$ the family of all sets $Q\subset X_1\times X_2$ with measurable cuts $Q_{x_1}\in\mathcal{A}_2$ is a **σ-algebra** containing all **measurable rectangles** $A_1\times A_2$ with $A_1\in\mathcal{A}_1$ resp. $A_2\in\mathcal{A}_2$ since $(A_1\times A_2)_{x_1} = \left\{\begin{smallmatrix} A_2, & x_1\in A_1 \\ \emptyset, & x_1\notin A_1\end{smallmatrix}\right.$. Hence according to 7.4.3 it includes the σ-algebra $\mathcal{A}_1\otimes\mathcal{A}_2$ generated by these sets.

8.2 Lemma: For two **σ-finite** measure spaces $(X_i;\mathcal{A}_i;\mu_i)$ with $i\in\{1;2\}$ and every $A\in\mathcal{A}_1\otimes\mathcal{A}_2$ the mappings $s_{1A}: X_2\to[0;\infty]$ with $s_{1A}(x_2) = \mu_1(A_{x_2})$ resp. $s_{2A}: X_1\to[0;\infty]$ with $s_{2A}(x_1) = \mu_2(A_{x_1})$ are measurable.

Proof: Preliminarily so as to have access to **complements** we confine ourselves to $s_{1nA}(x_2) := \mu_1|_{A_n}(A_{x_2})$ with the restriction $\mu_1|_{A_n}$ on one of the μ_1-finite sets A_{1n} from the w.l.o.g. **increasing** cover $\bigcup_{n\in\mathbb{N}} A_{1n} = X_1$. The family $\mathcal{D}$ of subsets $D\subset X_1\times X_2$ with a measurable s_{1nD} is a **Dynkin system** since the constant function $s_{1n\emptyset} = 0$ is measurable, for every measurable s_{1nA} the **complement** function $s_{1n(X_1\times X_2)\setminus A}(x_2) = \mu_1|_{A_n}\left(((X_1\times X_2)\setminus A)_{x_2}\right) = \mu_1|_{A_n}\left((X_1\times X_2)_{x_2}\setminus A_{x_2}\right) = \mu_1|_{A_n}\left((X_1\times X_2)_{x_2}\right) - \mu_1|_{A_n}(A_{x_2}) = \mu_1|_{A_n}(X_1) - s_{1nA}(x_2)$ is measurable and so is the **summation** function $s_{1n\mathring{\bigcup} D_m} = \sum_{m\in\mathbb{N}} s_{1nD_m}$ with $(s_{1nD_m})_{m\in\mathbb{N}}$ for pairwise disjoint sets $(D_m)_{m\in\mathbb{N}}$owing to 4.9. Furthermore $s_{1n(A_1\times A_2)}(x_2) = \mu_1|_{A_n}\left((A_1\times A_2)_{x_2}\right) = \mu_1|_{A_n}(A_1)\cdot\chi_{A_2}(x_2)$ is measurable for every **measurable rectangle** $A_1\times A_2$ with $A_1\in\mathcal{A}_1$ resp. $A_2\in\mathcal{A}_2$. Hence the system $\mathcal{A}_1\times\mathcal{A}_2$ of measurable rectangles is included in $\mathcal{D}$ and since it is **closed under intersection** we can apply the **Dynkin δ-π-theorem** 1.6 resp. 7.4.3 to obtain $\sigma(\mathcal{A}_1\times\mathcal{A}_2) = \mathcal{A}_1\otimes\mathcal{A}_2\subset\mathcal{D}$. According to the **continuity from below** 2.2.2 and 4.9 the measurability of the s_{1nA} extends to $\sup\limits_{n\in\mathbb{N}} s_{1nA}(x_2) = \sup\limits_{n\in\mathbb{N}} \mu_1|_{A_n}(A_{x_2}) = \sup\limits_{n\in\mathbb{N}} \mu_1(A_n\cap A_{x_2}) = \mu_1\left(\bigcup_{n\in\mathbb{N}} A_n\cap A_{x_2}\right) = \mu_1(A_{x_2}) = s_{1A}(x_2)$. The proof for s_{2A} is of course analogous.

8.3 Product measure: On the **product** $(X_1\times X_2;\mathcal{A}_1\otimes\mathcal{A}_2)$ of **two σ-finite measure spaces** $(X_i;\mathcal{A}_i;\mu_i)$ with $i\in\{1;2\}$ the expression $(\mu_1\otimes\mu_2)(A) := \int \mu_1(A_{x_2})\,d\mu_2 = \int \mu_2(A_{x_1})\,d\mu_1$ for $A\in\mathcal{A}_1\otimes\mathcal{A}_2$ defines a **σ-finite** measure **uniquely determined** by its **multiplicity** $(\mu_1\otimes\mu_2)(A_1\times A_2) = \mu_1(A_1)\cdot\mu_2(A_2)$ for every $A_1\times A_2\in\mathcal{A}_1\times\mathcal{A}_2$.

Proof: On account of $\mu_1\left((A_1\times A_2)_{x_2}\right) = \mu_1(A_1)\cdot\chi_{A_2}(x_2)$ and vice versa the two integrals coincide and the set function $\mu_1\otimes\mu_2$ is **well defined** and obviously **uniquely determined** by its multiplicity on the family $\mathcal{A}_1\times\mathcal{A}_2$ of all **cylinder sets**. Due to 8.2 both integrals are **well defined** on $\mathcal{A}_1\otimes\mathcal{A}_2 = \sigma(\mathcal{A}_1\times\mathcal{A}_2)$. The **first integral** is **σ-additiv**e on $\mathcal{A}_1\otimes\mathcal{A}_2$ since for every sequence $(A_n)_{n\in\mathbb{N}}\subset\mathcal{A}_1\times\mathcal{A}_2$ of pairwise disjoint measurable sets the **σ-additivity** of μ_1 and **monotone convergence** 5.12 applied to μ_2 yield $(\mu_1\otimes\mu_2)\left(\mathring{\bigcup}_{n\in\mathbb{N}} A_n\right) = \int \mu_1\left(\left(\mathring{\bigcup}_{n\in\mathbb{N}} A_n\right)_{x_2}\right) d\mu_2 = \int\left(\sum_{n\in\mathbb{N}} \mu_1(A_n)_{x_2}\right) d\mu_2 = \sum_{n\in\mathbb{N}} \int \mu_1(A_{x_2})\,d\mu_2 = \sum_{n\in\mathbb{N}} (\mu_1\otimes\mu_2)(A_n)$ in the case of the latter series converging to a finite limit. In the case of a diverging series $\sum_{n\in\mathbb{N}} \int \mu_1(A_{x_2})\,d\mu_2 = \infty$ there is an $N\in\mathbb{N}$ with $\int\left(\sum\limits_{n=0}^{N} \mu_1(A_n)_{x_2}\right) d\mu_2 = \sum\limits_{n=0}^{N} \int \mu_1(A_{x_2})\,d\mu_2 \geq C$ for every $C>0$ and hence

$(\mu_1 \otimes \mu_2)\left(\mathring{\bigcup}_{n\in\mathbb{N}} A_n\right) = \sum_{n\in\mathbb{N}} (\mu_1 \otimes \mu_2)(A_n) = \infty$. The same argument of course appplies to the **second integral** such that both are measures on $\mathcal{A}_1 \otimes \mathcal{A}_2$ coinciding on the **π-basis** $\mathcal{A}_1 \times \mathcal{A}_2$ and hence on all of $\mathcal{A}_1 \otimes \mathcal{A}_2$ due to 3.4. $\mu_1 \otimes \mu_2$ is **σ-finite** since for a cover $(A_{in})_{n\in\mathbb{N}} \subset \mathcal{A}_i$ of μ_i-sets A_{in} with $i \in \{1;2\}$ the sequence $(A_{1n} \times A_{2n})_{n\in\mathbb{N}} \subset \mathcal{A}_1 \otimes \mathcal{A}_2$ is a cover of $X_1 \times X_2$ from $\mu_1 \otimes \mu_2$-finite sets $A_{1n} \times A_{2n}$.

8.4 Lemma: **Almost all cuts** Z_{x_1} of a $\mu_1\otimes\mu_2$**-null set** $Z \in \mathcal{A}_1\otimes\mathcal{A}_2$are μ_2**-null sets**: $(\mu_1 \otimes \mu_2)(Z) = 0 \Rightarrow \mu_2(Z_{x_1}) = 0$ for every $x_1 \in X_1 \setminus Z_1$ with $\mu_1(Z_1) = 0$ and analogously for Z_{x_2}.

Proof: By the **approximation property** 3.6 for every $\epsilon > 0$ there exists a sequence $(A_n)_{n\in\mathbb{N}} \subset \mathcal{A}_1 \times \mathcal{A}_2$ of **cylinder sets** with $Z \subset \bigcup_{n\in\mathbb{N}} A_n$ and $\sum_{n\in\mathbb{N}} (\mu_1 \otimes \mu_2)(A_n) \leq \frac{\epsilon}{n\cdot 2^n}$. Hence $Z_{x_1} \subset \bigcup_{n\in\mathbb{N}} A_{n,x_1}$ and for $T_n = \left\{x_1 \in X_1 : \sum_{n\in\mathbb{N}} \mu_2(A_{n,x_1}) \geq \frac{1}{n}\right\}$ we have $\frac{1}{n}\mu_1(T_n) \leq \int \sum_{n\in\mathbb{N}} \mu_2(A_{n,x_1})\, d\mu_1 \overset{5.12}{=} \sum_{n\in\mathbb{N}} \int \mu_2(A_{n,x_1})\, d\mu_1 \overset{8.3}{=} \sum_{n\in\mathbb{N}} (\mu_1 \otimes \mu_2)(A_n) \leq \frac{\epsilon}{n\cdot 2^n}$ whence $\mu_1(\bigcup_{n\in\mathbb{N}} T_n) \leq \sum_{n\in\mathbb{N}} \mu_1(T_n) \leq \epsilon$ and finally $\mu_1(\mu_2(Z_{x_1}) > 0) \leq \mu_1(\mu_2(\bigcup_{n\in\mathbb{N}} A_{n,x_1}) > 0) < \epsilon$ which proves the assertion for $Z_1 = \{x_1 \in X_1 : \mu_2(\bigcup_{n\in\mathbb{N}} A_{n,x_1}) > 0\}$.

8.5 Fubini's theorem: For two **σ-finite** measure spaces $(X_i; \mathcal{A}_i; \mu_i)$ with $i \in \{1;2\}$ every $\mathcal{A}_1 \otimes \mathcal{A}_2$**-measurable** function $f : X_1 \times X_2 \to Y$ into a **separable Banach space** $(Y; |\,|)$ is $\mu_1 \otimes \mu_2$-integrable iff either $f_{x_1} : X_2 \to Y$ with $f_{x_1}(x_2) = f(x_1;x_2)$ is $\mu_2$integrable for μ_1-a.e. $x_1 \in X_1$ and $\int \left(\int f_{x_1} d\mu_2\right) d\mu_1 < \infty$ or vice versa and in that case we have $\int f d(\mu_1 \otimes \mu_2) = \int \left(\int f_{x_1} d\mu_2\right) d\mu_1 = \int \left(\int f_{x_2} d\mu_1\right) d\mu_2$.

Proof:

Step I: The function $f_{x_1} : X_2 \to Y$ with $f_{x_1}(x_2) = f(x_1;x_2)$ is $\mathcal{A}_2$-measurable since due to 8.1 for every Borel measurable set $B \subset Y$ we have $f_{x_1}^{-1}[B] = \{(x_1;\xi_2) : f(x_1;\xi_2) \in B\} = \left(f^{-1}[B]\right)_{x_1} \in \mathcal{A}_2$. For **step functions** $\varphi = \sum_{i=1}^{n} \alpha_i \chi_{F_{1,i}\times F_{2,i}} = \sum_{i=1}^{n} \alpha_i \chi_{F_{1,i}} \cdot \chi_{F_{2,i}} \in \mathcal{S}(\mathcal{F}_1 \times \mathcal{F}_2; Y)$ with $\alpha_i \in Y$ and **w.l.o.g. pairwise disjoint cylinder sets** $F_{1,i} \times F_{2,i} \in \mathcal{F}_1 \times \mathcal{F}_2$ for the algebrae $\mathcal{F}_j$ of μ_j**-finite sets** such that $\mathcal{A}_j = \sigma(\mathcal{F}_j)$ with $j \in \{1;2\}$ the **step function** $\varphi_{x_1} = \sum_{i=1}^{n} \alpha_i \chi_{F_{1,i}}(x_1) \cdot \chi_{F_{2,i}} \in \mathcal{S}(\mathcal{F}_2; Y)$ are obviously $\mathcal{A}_2$-measurable. On account of 8.3 the integration formula holds for these step functions since $\int \varphi d(\mu_1 \otimes \mu_2) = \sum_{i=1}^{n} \alpha_i \cdot \mu_1(F_{1,i}) \cdot \mu_2(F_{2,i}) = \int \left(\int \varphi_{x_1} d\mu_2\right) d\mu_1$. Assuming $f \in L^1(X_1 \times X_2; Y)$ by 7.7 resp. 5.22 there is a sequence $(\varphi_n)_{n\in\mathbb{N}} \subset \mathcal{S}(\mathcal{F}_1 \times \mathcal{F}_2; Y)$ with $\lim_{n\to\infty} \int |\varphi_n - f|\, d(\mu_1 \otimes \mu_2) = 0$ and in particular $\lim_{n\to\infty} \int \left(\int \varphi_{n,x_1} d\mu_2\right) d\mu_1 = \lim_{n\to\infty} \int \varphi_n d(\mu_1 \otimes \mu_2) = \int f d(\mu_1 \otimes \mu_2)$.

Step II: By 5.10 and w.l.o.g. transferring to a subsequence there is a $\mu_1 \otimes \mu_2$-null set $Z \in \mathcal{A}_1 \otimes \mathcal{A}_2$ with $\lim_{n\to\infty} \varphi_n(x_1;x_2) = f(x_1;x_2)$ for every $(x_1;x_2) \in (X_1 \times X_2) \setminus Z$. Hence due to 8.4 we have $\lim_{n\to\infty} \varphi_{n,x_1}(x_2) = f_{x_1}(x_2)$ for every $x_2 \in X_2 \setminus Z_{x_1}$ with $\mu_2(Z_{x_1}) = 0$ and $x_1 \in X_1 \setminus Z_1$ for a μ_1-null set Z_1. The sequence $(\Phi_n)_{n\in\mathbb{N}} \subset \mathcal{S}(X_1; \mathcal{S}(\mathcal{F}_2; Y))$ with $\Phi_n(x_1) = \varphi_{n,x_1}$ is $L^1(\mu_1)$-Cauchy since $\|\Phi_n - \Phi_m\|_1 = \int |\Phi_n - \Phi_m|\, d\mu_1 = \int \left(\int |\varphi_{n,x_1} - \varphi_{m,x_1}|\, d\mu_2\right) d\mu_1 = \int |\varphi_n - \varphi_m|\, d(\mu_1 \otimes \mu_2)$. By 5.10 and w.l.o.g. retreating to a subsequence there is a a $\Phi \in \overline{(\mathcal{S}(X_1; \mathcal{S}(\mathcal{F}_2; Y)); \|\|_1)} = L^1(X_1; \mathcal{S}(\mathcal{F}_2; Y))$ and a μ_1-null set W_1 such that $\lim_{n\to\infty} \|\Phi_n(x_1) - \Phi(x_1)\|_1 = 0$ for every $x_1 \in X_1 \setminus (Z_1 \cup W_1)$. In particular $\|\varphi_{n,x_1}\|_1 = \|\Phi_n(x_1)\|_1 \leq \|\Phi_n(x_1) - \Phi(x_1)\|_1 + \|\Phi(x_1)\|_1 \leq 2\|\Phi(x_1)\|_1$ for n large enough whence $\Phi(x_1) \in L^1(X_2; Y)$ by 5.17. A third instance of 5.10 verifies that μ_2-a.e. and for $x_1 \in X_1 \setminus (Z_1 \cup W_1)$ we have $\Phi(x_1) = f_{x_1}$, i.e. $\lim_{n\to\infty} \|\Phi_n(x_1) - \Phi(x_1)\|_1 = \lim_{n\to\infty} \int |\varphi_{n,x_1} - f_{x_1}|\, d\mu_2 = 0$ and hence $\lim_{n\to\infty} \int \varphi_{n,x_1} d\mu_2 = \lim_{n\to\infty} \int f_{x_1} d\mu_2$.

Step III: Due to 4.9 the function $x_1 \mapsto \int f_{x_1} d\mu_2$ is $\mathcal{A}_1$-measurable. The step functions $(\Psi_n)_{n\in\mathbb{N}} \subset \mathcal{S}(X_1; Y)$ with $\Psi_n(x_1) = \int \varphi_{n,x_1} d\mu_2 = \sum_{i=1}^{n} \alpha_i \chi_{F_{1,i}}(x_1) \cdot \mu_2(F_{2,i})$ are again $L^1(\mu_1)$-Cauchy since $\|\Psi_n - \Psi_m\|_1 = \int |\Psi_n - \Psi_m|\, d\mu_1 = \int \left(\int |\varphi_{n,x_1} - \varphi_{m,x_1}|\, d\mu_2\right) d\mu_1 = \int |\varphi_n - \varphi_m|\, d(\mu_1 \otimes \mu_2)$. Since by **step II** we have $\lim_{n\to\infty} \Psi_n(x_1) = \lim_{n\to\infty} \int \varphi_{n,x_1} d\mu_2 = \int f_{x_1} d\mu_2$ for every $x_1 \in X_1 \setminus (Z_1 \cup W_1)$ by 5.10 we conclude $\lim_{n\to\infty} \int \left|\int \varphi_{n,x_1} d\mu_2 - \int f_{x_1} d\mu_2\right| d\mu_1 = \lim_{n\to\infty} \left\|\Psi_n - \int f_{x_1} d\mu_2\right\|_1 = 0$. In particular by **step I** we have shown that $\int \left(\int f_{x_1} d\mu_2\right) d\mu_1 = \lim_{n\to\infty} \int \left(\int \varphi_{n,x_1} d\mu_2\right) d\mu_1 = \lim_{n\to\infty} \int \varphi_n d(\mu_1 \otimes \mu_2) = \int f d(\mu_1 \otimes \mu_2)$.

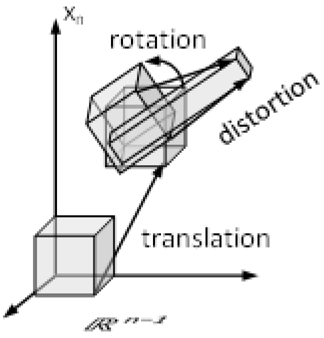
Step IV: By 5.15 we may assume $|f|_{x_1} \in L^1\left(\mu_2; \mathbb{R}_0^+\right)$ for every $x_1 \in X_1 \setminus V_1$ with $\mu_1(V_1) = 0$ resp. $\int \left(\int |f|_{x_1} d\mu_2\right) d\mu_1 < \infty$ and by steps I - III it suffices to show that $|f| \in L^1\left(\mu_1 \otimes \mu_2; \mathbb{R}_0^+\right)$. Since $|f| : X_1 \times X_2 \to \mathbb{R}_0^+$ is measurable 5.2 provides an w.l.o.g. increasing sequence $(\varphi_n)_{n\in\mathbb{N}} \subset \mathcal{S}(X_1 \times X_2; \mathbb{R})$ converging outside of a $\mu_1 \otimes \mu_2$-null set Z to f. As above resp. according to 8.4 we have $\lim\limits_{n\to\infty} \varphi_{n,x_1}(x_2) = |f|_{x_1}(x_2)$ for every $x_2 \in X_2 \setminus Z_{x_1}$ with $\mu_2(Z_{x_1}) = 0$ and $x_1 \in X_1 \setminus Z_1$ for a μ_1-null set Z_1. **Monotone convergence** 5.12 then yields $\lim\limits_{n\to\infty} \int \varphi_{n,x_1} d\mu_2 = \int |f|_{x_1} d\mu_2$ for every $x_1 \in X_1 \setminus (Z_1 \cup V_1)$. By definition 5.1 every step function $\varphi_n \in L^1(X_1 \times X_2; \mathbb{R})$ is integrable so that steps I - III yield $\int \varphi_n d(\mu_1 \otimes \mu_2) = \int \left(\int \varphi_{n,x_1} d\mu_2\right) d\mu_1$. Since the sequence $\left(\int \varphi_{n,x_1} d\mu_2\right)_{n\in\mathbb{N}}$ is increasing we may invoke **monotone convergence** a second time to obtain $\lim\limits_{n\to\infty} \int \left(\int \varphi_{n,x_1} d\mu_2\right) d\mu_1 = \int \left(\int |f|_{x_1} d\mu_2\right) d\mu_1$. A third instance of the **monotone convergence** theorem applied to $(\varphi_n)_{n\in\mathbb{N}} \subset L^1(X_1 \times X_2; \mathbb{R})$ delivers $\lim\limits_{n\to\infty} \int \varphi_n d(\mu_1 \otimes \mu_2) = \int |f| \, d(\mu_1 \otimes \mu_2)$ and hence the assertion.

8.6 Finite products of measure spaces: On the finite product $\left(\prod_{i\in J} X_i; \bigotimes_{i\in J} \mathcal{A}_i\right)$ of the σ-finite measure spaces $(X_i; \mathcal{A}_i; \mu_i)$ with a finite index set $J = \{1, ..., n\}$ the **product measure** $\bigotimes_{i\in J} \mu_i$ is **uniquely determined** by the **multiplicity** condition $\mu\left(\prod_{i\in J} A_i\right) = \prod_{i\in J} \mu_i(A_i)$ and is **constructed inductively** according to 8.3 by means of $\bigotimes_{1\leq j\leq i} \mu_j := \left(\bigotimes_{1\leq j<i} \mu_j\right) \otimes \mu_i$. The resulting product of measure spaces is denoted as $\bigotimes_{i\in J}(X_i; \mathcal{A}_i; \mu_i) := \left(\prod_{i\in J} X_i; \bigotimes_{i\in J} \mathcal{A}_i; \bigotimes_{i\in J} \mu_i\right)$. For a Borel measurable function $f : \prod_{i\in J} X_i \to Y$ with finite integrals $\int \left(... \left(\int f_{x_{j(2)}...x_{j(n)}} d\mu_{j(1)}\right) ...\right) d\mu_{j(k)}$ for every $1 \leq k \leq n$ and **some permutation** $j : J \to J$ we have $\int f d\mu = \int \left(... \left(\int f_{x_{j(2)}...x_{j(n)}} d\mu_{j(1)}\right) ...\right) d\mu_{j(n)}$ for **every permutation**. Hence the **convergence for one particular order of integration grants the integrability of all permutations**.

8.7 Completion of λ^n: The product $\lambda^n = \bigotimes_{1\leq i\leq n} \lambda$ of the **complete** Lebesgue measures λ on the product $\mathcal{B}^n = \bigotimes_{1\leq i\leq n} \mathcal{B}$ of the Lebesgue σ-algebrae $\mathcal{B}$ on $\mathbb{R}$ is **not complete** any more since for any λ-null set $A \in \mathcal{B}$ we have $\lambda^2(A \times \mathbb{R}) = 0$ and for any non Lebesgue measurable $B \notin \mathcal{B}$ (cf. 3.10) evidently $A \times B \subset A \times \mathbb{R}$ holds but $A \times B \notin \mathcal{B}^2$. The completion of the product according to 3.8 will be included without change of notation in the extension obtained by means of the **Riesz representation theorem** 11.13 to the **Lebesgue measure** λ^n on the **Lebesgue σ-algebra** $\mathcal{B}^n$.

8.8 Translation invariance of the Lebesgue-Borel measure on $\mathbb{R}^n$: The **Lebesgue-Borel** measure λ^n on the **Borel** σ-algebra $\mathcal{B}^n$ on $\mathbb{R}^n$ is **uniquely determined** by its **translation invariance** on the π-**basis** of the **n-dimensional intervals** $\mathcal{I}^n$: For every translation $T_{\mathbf{c}} : \mathbb{R}^n \to \mathbb{R}^n$ with $T_{\mathbf{c}}(\mathbf{x}) = \mathbf{x} + \mathbf{c}$ for a $\mathbf{c} \in \mathbb{R}^\mathbf{n}$ and every interval $[\mathbf{a}; \mathbf{b}[:= \prod\limits_{i=1}^{n} [a_i; b_i[\in \mathcal{I}^n$ with $a_i \leq b_i \in \mathbb{R}$ due to 4.3 and 8.3 we have $T_{\mathbf{c}}(\lambda^n)([\mathbf{a}; \mathbf{b}[) = \lambda^n\left(T_{\mathbf{c}}^{-1}([\mathbf{a}; \mathbf{b}[)\right) = \lambda^n([\mathbf{a} - \mathbf{c}; \mathbf{b} - \mathbf{c}[) = \lambda^n([\mathbf{a}; \mathbf{b}[) = \prod\limits_{i=1}^{n} [b_i - a_i[$, i.e. the σ**-finite** measures $T_{\mathbf{c}}(\lambda^n)$ and λ^n coincide on the π-basis $\mathcal{I}^n$ and hence on $\sigma(\mathcal{I}^n) = \mathcal{B}^n$ due to 3.4.

8.9 Transformation formula: The image of the **Lebesgue-Borel measure** λ^n under a **homomorphism** $T \in GL(n; \mathbb{R})$ is $T \circ \lambda^n = \frac{\lambda^n}{|\det T|}$ such that $\lambda^n(T[A]) = |\det T| \cdot \lambda^n(A)$ for every Borel-measurable $A \in \mathcal{B}^n$. Special cases are:



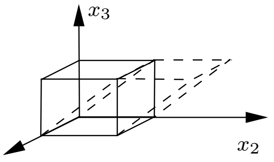
1. A **proprtional distortion** along the axes by $T(e_i) = r_i \cdot e_i$ for $r_i \in \mathbb{R}$ and $0 \leq i \leq n$ results in $\lambda^n(T(A)) = \left|\prod\limits_{i=1}^{n} r_i \cdot \lambda^n(A)\right| \cdot \lambda^n(A)$and particularly a simple **scaling** of the set A by the **scaling factor** $r \in \mathbb{R}$ yields the volume $\lambda^n(rA) = |r^n| \cdot \lambda^n(A)$.
2. A **rotation** by an **orthogonal matrix** $T \in O(n; \mathbb{R})$ leaves the volume unacffected: $\lambda^n(T[A]) = |\det T| \cdot \lambda^n(A) = \lambda^n(A)$.

3. In the **three dimensions** of $\mathbb{R}^3$ the homomorphism T may be represented by a matrix with three linearly independent column vectors $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3 \in \mathbb{R}^3$ generating a **parallelepiped** $T[Q] = \left\{\sum_{1\leq i\leq 3} x_i \mathbf{a}_i : 0 \leq x_i < 1\right\}$ which is the image of the **unit cube** $C_3 = \left\{\sum_{1\leq i\leq 3} x_i \mathbf{e}_i : 0 \leq x_i < 1\right\}$. Its volume is then $\lambda^3(T[Q]) = |\det T| \cdot \lambda^3(Q) = \det(\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3) \cdot 1 = \mathbf{a}_1 * \mathbf{a}_2 \times \mathbf{a}_3$.

Proof: According to e.g. [3, 2.6.3 Satz A] every **homomorphism** resp. every **invertible matrix** is the product of **elementary transformations** resp. **elementary matrices** of the two following types:

$$\begin{array}{cc} & \begin{array}{ccccccc} 1 & \cdots & k & \cdots & l & \cdots & n \end{array} \\ E_{kl} = & \left(\begin{array}{ccccccc} 1 &&&&&& \\ & \ddots &&&&& \\ && 1 && 1 && \\ &&& \ddots &&& \\ &&&& 1 && \\ &&&&& \ddots & \\ &&&&&& 1 \end{array}\right) \begin{array}{c} 1 \\ \vdots \\ k \\ \vdots \\ l \\ \vdots \\ n \end{array} \end{array}$$

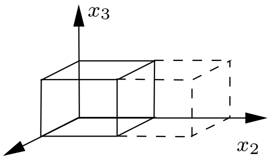
$$E_{23} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$$



$E_{2;3}(H_3)$

$$\begin{array}{cc} & \begin{array}{ccccc} 1 & \cdots & k & \cdots & n \end{array} \\ E_{k\alpha} = & \left(\begin{array}{ccccc} 1 &&&& \\ & \ddots &&& \\ && \alpha && \\ &&& \ddots & \\ &&&& 1 \end{array}\right) \begin{array}{c} 1 \\ \vdots \\ k \\ \vdots \\ n \end{array} \end{array}$$

$$E_{22} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$



$E_{2;2}(H_3)$

Multiplication with E_{kl} results in an addition of the l-th row to the k-th row, i.e. a **shearing** so that the image of the **unit cube** $Q := [\mathbf{0}; \mathbf{1}[$ generated by the **basis vectors** $\mathbf{e}_1, ..., \mathbf{e}_n$ with the **measure** $\lambda^n(Q) = (1-0)^n = 1$ is $E_{kl}[Q] = \left\{\sum_{1\leq i\leq n} x_i \mathbf{e}_i : 0 \leq x_i \leq 1; i \neq k \wedge x_l \leq x_k < x_l + 1\right\}$. This **parallelepiped** can be split into two disjoint halves $L = \{\mathbf{x} \in E_{kl}[Q] : x_l \leq x_k < 1\}$ and $R = \{\mathbf{x} \in E_{kl}[Q] : 1 \leq x_k < x_l + 1\}$ such that $E_{kl}[Q] = L \mathring{\cup} R$ but also $Q = (R - \mathbf{e}_k) \mathring{\cup} L$ and due to the **translation invariance** of λ^n we obtain $\lambda^n(E_{kl}[Q]) = \lambda^n(K) + \lambda^n(L) = \lambda^n(K) + \lambda^n(L - \mathbf{e}_k) = \lambda^n(Q) = 1 \cdot \lambda^n(Q) = |\det E_{kl}| \cdot \lambda^n(Q)$.

Multiplication with $E_{k\alpha}$ results in a multiplication of the k th row with the factor $\alpha \in \mathbb{R}$ resulting in the image $E_{k\alpha}[Q] = \left\{\sum_{1\leq i\leq n} x_i \mathbf{e}_i : 0 \leq x_i < 1; i \neq k \wedge 0 \leq x_k < \alpha\right\}$ with measure $\lambda^n(E_{k\alpha}(H)) = (1-0)^{n-1} \cdot (\alpha - 0) = \alpha = |\det E_{k\alpha}| \cdot \lambda^n(Q)$.

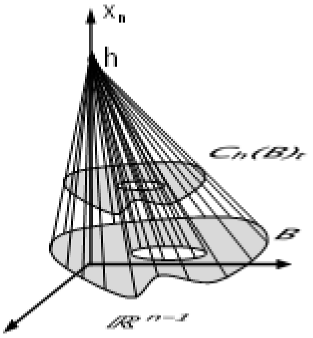
The assertion then follows from the **multiplicity of the determinant**: $|\det(A \cdot B)| = |\det(A)| \cdot |\det(B)|$.

8.10 Cavalieri's principle: For a **compact** $K \subset \mathbb{R}^n$ and any **cut** $K_t = \{\boldsymbol{x} \in \mathbb{R}^{n-1} : (\boldsymbol{x}; t) \in K\}$ with $t \in \mathbb{R}$ we have $\lambda^n(K) = \int_{\mathbb{R}} \lambda^{n-1}(K_t)\, dt$.

Proof: Due to **Fubini's theorem** 8.5 we have $\lambda^n(K) = \int_{\mathbb{R}^n} \chi_K(\boldsymbol{x})\, d\boldsymbol{x} = \int \left(\int_{\mathbb{R}^{n-1}} \chi_K(\boldsymbol{x}; t)\, d\boldsymbol{x}\right) dt = \int \left(\int_{\mathbb{R}^{n-1}} \chi_{K_t}(\boldsymbol{x})\, d\boldsymbol{x}\right) dt = \int_{\mathbb{R}} \lambda^{n-1}(K_t)\, dt$.

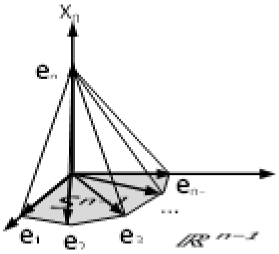
8.11 The cone $C_h(B) = \{((1-\lambda)\boldsymbol{\xi}, \lambda h) \in \mathbb{R}^n : \boldsymbol{\xi} \in B; 0 \le \lambda \le 1\}$ with **compact base** $B \subset \mathbb{R}^{n-1}$and **height** $h > 0$ has the **volume** $\lambda^n(C_h(B)) = \frac{h}{n} \cdot \lambda^{n-1}(B)$.

Proof: According to **Cavalieri's principle**8.10 and by the condition $\lambda h = t$ we obtain the **cuts** $C_h(B)_t = \begin{cases} \left(1 - \frac{t}{h}\right) B & \text{for } 0 \le t \le h \\ 0 & \text{else} \end{cases}$ with $\lambda^{n-1}\left(C_h(B)_t\right) = \left(1 - \frac{t}{h}\right)^{n-1} \cdot \lambda^{n-1}(B)$ due to 8.9 whence $\lambda^n(C_h(B)) = \int_{\mathbb{R}} \lambda^{n-1}\left(C_h(B)_t\right) dt = \lambda^{n-1}(B) \cdot \int_0^h = \left(1 - \frac{t}{h}\right)^{n-1} dt = \frac{h}{n} \cdot \lambda^{n-1}(B)$.



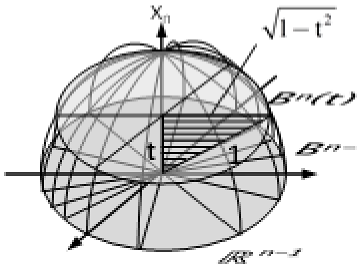
8.12 The unit simplex $S_1^n = \left\{ \sum\limits_{i=1}^{n} \lambda_i \boldsymbol{e}_i : \sum\limits_{i=1}^{n} \lambda_i = 1 \right\} \subset \mathbb{R}^n$ has the volume $\lambda^n(S_1^n) = \frac{1}{n!}$.

Proof: By **induction** over n we start with $\lambda^1(S_1^1) = \lambda^1([0;1]) = 1$ and proceed from $n-1$ to n by 8.11 with $\lambda^n(S_1^n) = \frac{1}{n} \cdot \lambda^{n-1}\left(S_1^{n-1}\right) = \frac{1}{n} \cdot \frac{1}{(n-1)!} = \frac{1}{n!}$.



8.13 The unit sphere B_1^n has the volume $\tau_n = \lambda^n(B_1^n) = \frac{\pi^{n/2}}{\Gamma\left(\frac{n}{2}+1\right)}$.

Proof: As above we procced by **induction** over n starting with $\lambda^1(B_1^1) = \lambda^1([-1;1]) = 2$ and proceed from $n-1$ to n by **Cavalieri's principle** 8.10 with $\lambda^n(B_1^n) = \int_{\mathbb{R}} \lambda^{n-1}\left(\left(B_1^{n-1}\right)_t\right) dt = \int_{\mathbb{R}} \lambda^{n-1}\left(\left(B^{n-1}_{\sqrt{1-t^2}}\right)\right) dt = \lambda^{n-1}\left(B_1^{n-1}\right) \cdot \int_{[-1;1]} \left(1-t^2\right)^{(n-1)/2} dt = \lambda^{n-1}\left(B_1^{n-1}\right) \cdot c_n$. By **substitution** and **integration by parts** we can simplify $c_n = \int_{-1}^{1} \left(1-t^2\right)^{(n-1)/2} dt = 2\int_0^{\pi/2} \sin^n(\alpha) d\alpha = 2(n-1)\int_0^{\pi/2} \cos^2(\alpha) \cdot \sin^{(n-2)}(\alpha) d\alpha$. By expanding this expression to $(1-n)\int_0^{\pi/2} \left(\sin^2(\alpha) + \cos^2(\alpha)\right) \cdot \sin^{(n-2)}(\alpha) da + n\int_0^{\pi/2} \sin^n(\alpha) d\alpha = 0$ we can use **Pythagoras** to obtain



$$c_n = 2 \cdot \frac{n-1}{n} \int_0^{\pi/2} \sin^{n-2}(\alpha) d\alpha$$

$$= 2 \cdot \frac{n-1}{n} \cdot \frac{n-3}{n-2} \cdot \ldots \cdot \begin{cases} \frac{1}{2} \cdot \int_0^{\pi/2} 1 d\alpha = \frac{\pi}{4} & \text{for } n \text{ even} \\ \frac{2}{3} \cdot \int_0^{\pi/2} \sin(\alpha) d\alpha = \frac{2}{3} & \text{for } n \text{ odd} \end{cases}$$

$$= 2 \cdot \begin{cases} \frac{n-1}{n} \cdot \frac{n-3}{n-2} \cdot \ldots \cdot \frac{3}{4} \cdot \frac{1}{2} \cdot \frac{\pi}{2} & \text{for } n \text{ even} \\ \frac{n-1}{n} \cdot \frac{n-3}{n-2} \cdot \ldots \cdot \frac{4}{5} \cdot \frac{2}{3} & \text{for } n \text{ odd} \end{cases}$$

Hence we have $c_n \cdot c_{n-1} = \frac{2\pi}{n}$ so that with $\lambda^2(B_1^2) = \lambda^1(B_1^1) \cdot c_2 = \pi$ follows

$$\lambda^n(B_1^n) = \frac{2\pi}{n} \cdot \lambda^{n-2}\left(B_1^{n-2}\right) = \begin{cases} \frac{2\pi}{n} \cdot \frac{2\pi}{n-2} \cdot \ldots \cdot \frac{2\pi}{4} \cdot \pi = \frac{\pi}{n/2} \cdot \frac{\pi}{n/2-1} \cdot \ldots \cdot \frac{\pi}{2} \cdot \frac{\pi}{1} & \text{for } n \text{ even} \\ \frac{2\pi}{n} \cdot \frac{2\pi}{n-2} \cdot \ldots \cdot \frac{2\pi}{3} \cdot 2 = \frac{\pi}{n/2} \cdot \frac{\pi}{n/2-1} \cdot \ldots \cdot \frac{\pi}{3/2} \cdot \frac{\sqrt{\pi}}{1/2} \cdot \frac{1}{\sqrt{\pi}} & \text{for } n \text{ odd} \end{cases}$$

A comparison with the **Gamma function** (cf. 15.1 resp. 15.3.3) with the **functional equation** $\Gamma(x+1) = x \cdot \Gamma(x)$ for $0 < x < \infty$ and initial values $\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi} \Rightarrow \Gamma\left(\frac{3}{2}\right) = \frac{1}{2} \cdot \sqrt{\pi} \Rightarrow \Gamma\left(\frac{5}{2}\right) = \frac{3}{2} \cdot \frac{1}{2} \cdot \sqrt{\pi} \Rightarrow ...$ resp. $\Gamma(1) = 1 \Rightarrow \Gamma(2) = 1 \Rightarrow \Gamma(3) = 1 \cdot 2 \Rightarrow ...$ yields the desired formula.

8.14 Probability measures on function spaces: On the product $(X_I;\mathcal{A}_I)$ of **probability spaces** $(X_i;\mathcal{A}_i;\mu_i)_{i\in I}$ with **arbitrary index set** I exists a **probability measure** μ_I **uniquely determined** by its **multiplicity** $\mu_I|_{\mathcal{Z}_J}=\mu_J:=\bigotimes_{\in J}\mu_i$ for all **finite** $J\subset I$, i.e. on **cylinder sets** $\pi_J^{-1}(A)\in\mathcal{Z}_J$ with $A\in A_J=\bigotimes_{i\in J}\mathcal{A}_i=\pi_J(\mathcal{A}_I)=\pi_J(\mathcal{Z}_J)$ (cf. 7.4.3) it coincides with the corresponding **finite product measure** $\mu_J=\bigotimes_{\in J}\mu_i$ on the **finite product-σ-algebrae** $\mathcal{A}_J$. The elements $x_I\in X_I$ with $x_I:I\to X_I$ are the **sample path**s or **realizations** of the **stochastic process** $(X_I;\mathcal{A}_I;\mu_I)$

Proof: The function $\mu_I:\mathcal{S}_I\to[0;1]$ given by $\mu_I\left(\pi_J^{-1}\left(\prod_{i\in J}A_i\right)\right):=\prod_{i\in J}\mu_i(A_i)$ for $A_j\in\mathcal{A}_j$ and **finite** $J\subset I$ is **well defined** and in particular independent of the representation of the **measurable rectangle** $S=\prod_{l\in L}S_l=\left(\pi_J^L\right)^{-1}\left(\prod_{j\in J}A_j\right)=\left(\pi_K^L\right)^{-1}\left(\prod_{k\in K}B_k\right)\in\mathcal{S}_L$ with $A_j\in\mathcal{A}_j$, $j\in J$ and $B_k\in\mathcal{A}_k$, $k\in K$ for **finite** $J,K,L\subset I$ with $J\cup K\subset L$. By the equality of the two representations we have $S_j=A_j=B_j$ for $j\in J\cap K$, $Z_j=A_j=X_j$ for $j\in J\setminus K$, $S_j=B_j=X_j$ for $j\in K\setminus J$ and finally $S_l=X_l$ for $l\in L\setminus(J\cup K)$. Hence the **multiplicity condition** with $\mu_i(X_i)=1$ for all $i\in I$ yields $\mu_L(S)=\mu_J\left(\prod_{j\in J}A_j\right)=\prod_{j\in J\cap K}\mu_j(A_j)=\prod_{j\in J\cap K}\mu_j(B_j)=\mu_K\left(\prod_{k\in K}B_k\right)$. According to 8.6 for every **finite** $J\subset I$ there is a **uniquely determined product measure** $\mu_J=\bigotimes_{\in J}\mu_i$ on the **finite product-σ-algebra** $\mathcal{A}_J$ with $\mu_I\left(\prod_{i\in J}A_i\right):=\prod_{i\in J}\mu_i(A_i)$ for $A_j\in\mathcal{A}_j$. Hence the extension $\mu_I:\mathcal{Z}_I\to[0;1]$ given by $\mu_I(Z):=\mu_J(A_J)$ for $Z=\pi_J^{-1}(A_J)$ and $A_J\in\mathcal{A}_J$ with **finite** $J\subset I$ on the **algebra** $\mathcal{Z}_I$ is **well defined** and in particular independent of the representation of the **cylinder set** $Z=\pi_J^{-1}(A_J)=\pi_K^{-1}(B_K)$ with $A_J\in\mathcal{A}_J$ and $B_K\in\mathcal{A}_K$ for **finite** $J,K\subset I$. We now prove that μ_I is **$\emptyset$-continuous** on the algebra of cylinder sets.

To this end for a given **path** $x_J\in X_J$ and a given K-**cylinder set** $Z\in\mathcal{Z}_K$ with **finite** $J\subset K\subset I$ we examine the Z-**extensions** $Z^{x_J}=\left\{\xi_I\in X_I:\left(x_J;\pi_{K\setminus J}(\xi_I)\right)\in Z\right\}=\pi_{K\setminus J}^{-1}(A_{x_J})\in\mathcal{Z}_K$ for $A=\pi_K(Z)\in\mathcal{A}_K=\mathcal{A}_J\otimes\mathcal{A}_{K\setminus J}$ and the **cuts** A_{x_J} of $A\in A_K$ being $\mathcal{A}_{K\setminus J}$-measurable due to 8.1. Hence he family Z^{x_J} consists of all **measurable extensions** $\xi_I\in X_I$ of the given path x_J with an **arbitrary course** during J (!) and **passing through** Z during $K\setminus J$. (cf. the set of all **paths** passing a given **tree** in [9, 15.5]). Owing to 8.3 we have $\mu_I(Z)=\mu_{I\setminus K}\left(\pi_{I\setminus K}(Z)\right)\cdot\mu_K(\pi_K(Z))=1\cdot\mu_K(A)=\int\mu_{K\setminus J}(A_{x_J})\,d\mu_J=\int\mu_I(Z^{x_J})\,d\mu_J$.

Now let $(Z_n)_{n\geq1}\subset\mathcal{Z}_I$ be a **decreasing** sequence of **cylinder sets** $Z_n=\pi_{J_n}^{-1}(A_n)$ with $A_n\in\mathcal{A}_{J_n}$ for finite $J_{n+1}\supset J_n$ and $Z_{n+1}\subset Z_n$ as well as $\mu_I(Z_n)\geq\alpha>0$ for $n\geq1$ such that $\inf\limits_{n\geq1}\mu_I(Z_n)\geq\alpha$. In order to show the **$\emptyset$-continuity** we have to prove that $\bigcap_{n\geq1}Z_n\neq\emptyset$, i.e. we must find a path $x\in\bigcap_{n\geq1}Z_n$. We start on the interval J_1 with a section x_{J_1} and proceed by **induction** to extend it to $\left(x_{J_1};x_{J_2\setminus J_1};...\right)$:

Due to 8.2 the mapping $x_{J_1}\mapsto\mu_I\left(Z_n^{x_{J_1}}\right)=\pi_{J_n\setminus J_1}^{-1}\left((A_n)_{x_{J_1}}\right)$ is measurable and hence the set $Q_n^{J_1}=\left(x_{J_1}\in X_{J_1}:\mu_I\left(Z_n^{x_{J_1}}\right)\geq\frac{\alpha}{2}\right)\in\mathcal{A}_{J_1}$ of all paths $x_{J_1}\in X_{J_1}$ which can be extended with a probability of at least $\frac{\alpha}{2}$ on Z_n is $\mathcal{A}_{J_1}$-measurable. According to the preceding paragraph we obtain the estimate $\alpha\leq\mu_I(Z_n)\leq\int_{Q_n^{J_1}}\mu_I\left(Z_n^{x_{J_1}}\right)d\mu_{J_1}+\int_{X_I\setminus Q_n^{J_1}}\mu_I\left(Z_n^{x_{J_1}}\right)d\mu_{J_1}\leq\mu_{J_1}\left(Q_n^{J_1}\right)+\frac{\alpha}{2}$ and hence $\mu_{J_1}\left(Q_n^{J_1}\right)\geq\frac{\alpha}{2}$ for all $n\geq1$. Since μ_{J_1} is continuous from above and $Q_{n+1}^{J_1}\subset Q_n^{J_1}$ for all $n\geq1$ there is an $x_{J_1}\in\bigcap_{n\geq1}Q_n^{J_1}\neq\emptyset$, i.e. $\mu_I\left(Z_n^{x_{J_1}}\right)\geq\frac{\alpha}{2}$ for all $n\geq1$.

We now extend the path x_{J_1}inductively with $Z_n^{x_{J_k}}$ taking the place of Z_n: Assuming there is an $x_{J_k}\in X_{J_k}$ with $\mu_I\left(Z_n^{x_{J_k}}\right)\geq\frac{\alpha}{2^k}$ for all $n\geq1$ we have

$$Q_n^{J_{k+1}}=\left(x_{J_{k+1}\setminus J_k}\in X_{J_{k+1}\setminus J_k}:\mu_I\left(\left(Z_n^{x_{J_k}}\right)^{x_{J_{k+1}}\setminus x_{J_k}}\right)\geq\frac{\alpha}{2^{k+1}}\right)\in\mathcal{A}_{J_{k+1}}$$

$$\begin{aligned}\text{whence }\frac{\alpha}{2^k}&\leq\mu_I\left(Z_n^{x_{J_k}}\right)\\&\leq\int_{Q_n^{J_{k+1}}}\mu_I\left(\left(Z_n^{x_{J_k}}\right)^{x_{J_{k+1}}\setminus x_{J_k}}\right)d\mu_{J_{k+1}}+\int_{X_I\setminus Q_n^{J_{k+1}}}\mu_I\left(\left(Z_n^{x_{J_k}}\right)^{x_{J_{k+1}}\setminus x_{J_k}}\right)d\mu_{J_{k+1}}\\&\leq\mu_{J_{k+1}}\left(Q_n^{J_{k+1}}\right)+\frac{\alpha}{2^{k+1}}\end{aligned}$$

such that $\mu_{J_{k+1}}\left(Q_n^{J_{k+1}}\right) \geq \frac{\alpha}{2^{k+1}}$ for all $n \geq 1$. Consequently there must exist an extension $x_{J_{k+1} \setminus J_k} \in \bigcap_{n\geq 1} Q_n^{J_{k+1}} \neq \emptyset$, i.e. $\mu_I\left(\left(Z_n^{x_{J_k}}\right)^{x_{J_{k+1}} \setminus x_{J_k}}\right) \geq \frac{\alpha}{2^{k+1}}$ for all $n \geq 1$. If we add the new section to x_{J_k} we obtain $x_{J_{k+1}} := \left(x_{J_k}; x_{J_{k+1}\setminus J_k}\right) \in X_{J_{k+1}}$ with $Z_n^{x_{J_{k+1}}} = \left(Z_n^{x_{J_k}}\right)^{x_{J_{k+1}} \setminus x_{J_k}}$, particularly $\pi_{J_k}^{J_{k+1}}(x_{k+1}) = x_k$ and $\mu_I\left(Z_n^{x_{J_{k+1}}}\right) \geq \frac{\alpha}{2^{k+1}}$ for all $n \geq 1$. Thus we have found a path $x' = \left(x_{J_1}; x_{J_2 \setminus J_1}; ...\right) \in \pi_{\bigcup_{n\geq 1} J_n}\left(\bigcap_{n\geq 1} Z_n\right) \subset X_{\bigcup_{n\geq 1} J_n}$ and by an arbitrary extension on the remaining time $I \setminus \bigcup_{n\geq 1} J_n$ we get the desired $x \in \bigcap_{n\geq 1} Z_n \neq \emptyset$ with $\pi_{\bigcup_{n\geq 1} J_n}(x) = x'$.

Hence μ_I is **$\emptyset$-continuous** and since due to 8.6 it is **finitely additive** as well as **bounded** according to 2.2.4 its **σ-additivity** follows. Due to the **extension theorem** 3.5 the **pre-measure** μ_I on the **algebra** $\mathcal{Z}_I$ of the **cylinder sets** can be extended in a unique way to a **measure** μ_I on the **σ-algebra** $\sigma(Z_I) = A_I$. This completes the proof.

9 Probability measures

9.1 Independence: A family $(A_i)_{i\in I} \subset \mathcal{A}$ of measurable sets on a **probability space** $(X; \mathcal{A}; \mu)$ is **independent**, if $\mu\left(\bigcap_{i\in F} A_i\right) = \prod_{i\in F} \mu(A_i)$ for every finite subset $F \subset I$. A family $(\mathcal{E}_i)_{i\in I}$ of set systems $\mathcal{E}_i \subset \mathcal{A}$ with $i \in I$ is independent if the families $\left(A_{i_f}\right)_{i_f \in F}$ are independent with $A_{i_f} \in \mathcal{E}_{i_f}$ for $i_f \in F$ and every nonempty and finite subset $F \subset I$. For two independent systems $\mathcal{E}, \mathcal{D} \subset \mathcal{A}$ on a probability space$(X; \mathcal{A}; \mu)$ the corresponding **Dynkin-systems** $\delta(\mathcal{E})$ and $\delta(\mathcal{D})$ are independent too since the family $\mathcal{I}(\mathcal{D}) := \{A \in \mathcal{A} : \mu(A \cap D) = \mu(A) \cdot \mu(D)\, \forall D \in \mathcal{D}\}$ already is a Dynkin-system: Obviously we have $X \in \mathcal{I}(\mathcal{D})$ and for $A \in \mathcal{I}(\mathcal{D})$ and $D \in \mathcal{D}$ we have $\mu((X \setminus A) \cap D) = \mu(D \setminus (A \cap D)) = \mu(D) - \mu(A \cap D) = \mu(D) - \mu(A) \cdot \mu(D) = \mu(D) \cdot (1 - \mu(A)) = \mu(X \setminus A) \cdot \mu(D)$ such that $X \setminus A \in \mathcal{I}(\mathcal{D})$. For pairwise disjoint $(A_n)_{n\in\mathbb{N}} \subset \mathcal{I}(\mathcal{D})$ we have $\mu\left(\left(\mathring{\bigcup}_{n\in\mathbb{N}} A_n\right) \cap D\right) = \mu\left(\mathring{\bigcup}_{n\in\mathbb{N}} (A_n \cap D)\right) = \sum_{n\in\mathbb{N}} \mu(A_n \cap D) = \sum_{n\in\mathbb{N}} \mu(A_n) \cdot \mu(D) = \mu(D) \cdot \sum_{n\in\mathbb{N}} \mu(A_n) = \mu(D) \cdot \mu\left(\mathring{\bigcup}_{n\in\mathbb{N}} A_n\right)$ and hence $\mathring{\bigcup}_{n\in\mathbb{N}} A_n \subset \mathcal{I}(\mathcal{D})$. On account of $\mathcal{E} \subset \mathcal{I}(\mathcal{D})$ follows $\delta(\mathcal{E}) \subset \mathcal{I}(\mathcal{D})$ and hence the assertion. Since independence refers to finite subfamilies this property extends to arbitrary independent families $(\mathcal{E}_i)_{i\in I}$ and their **Dynkin-systems** $(\delta(\mathcal{E}_i))_{i\in I}$ and with 1.6 even to their **σ-algebrae** $(\sigma(\mathcal{E}_i))_{i\in I} = (\delta(\mathcal{E}_i))_{i\in I}$ if the $(\mathcal{E}_i)_{i\in I}$ are **closed** with respect to **intersections**. Applying this property to the σ-algebrae $\sigma(\{A\}) = \{\emptyset; A; X \setminus A; X\}$ resp. $\sigma(\{B\})$ generated by two independents sets A and B shows the **independence of the complements**.

9.2 Borel's zero-one-law: For an **independent** sequence $(A_n)_{n\geq 1}$ of measurable sets $A_n \in \mathcal{A}$ on a probability space $(X; \mathcal{A}; \mu)$ we have $\mu\left(\bigcap_{n\geq 1} \bigcup_{k\geq n} A_k\right) \in \{0; 1\}$.

Proof: Due to 9.1 for every $n \geq 1$ the σ-algebrae $\mathcal{T}_{n+1} = \sigma\left(\left\{\bigcap_{m=0}^{j} A_{k_m} : k_m \geq n+1; 0 \leq m \leq j \in \mathbb{N}\right\}\right)$ and $\mathcal{A}_n = \sigma\left(\left\{\bigcap_{m=0}^{j} A_{k_m} : k_m \leq n; 0 \leq m \leq j \in \mathbb{N}\right\}\right)$ are **independent**. Also for every $n \geq 1$ we have $T = \bigcap_{n\geq 1} \bigcup_{k\geq n} A_k \in \mathcal{T}_n$ and hence $\mathcal{A}_n \in \mathcal{I}(T) := \{A \in \mathcal{A} : \mu(A \cap T) = \mu(A) \cdot \mu(T)\}$ as well as $\mathcal{T}_n \in \sigma(\mathcal{A})$ with $\mathcal{A} = \bigcup_{n\geq 1} \mathcal{A}_n$. Since $\mathcal{I}(T)$ is a **Dynkin-system** including the **π-system** $\mathcal{A}$ and consequently $\sigma(\mathcal{A}) = \delta(\mathcal{A}) \subset \mathcal{I}(T)$ follows $T \in \mathcal{I}(T)$, i.e. T is **independent of itself** and hence $\mu(T) = \mu(T \cap T) = \mu(T) \cdot \mu(T) \in \{0; 1\}$.

9.3 Chebyshev's inequality: For every function $f : X \to \mathbb{R}^+$ on a probability space $(X; \mathcal{A}; \mu)$ and every $\alpha > 0$ we have $\alpha \cdot \mu(\{f \geq \alpha\}) \leq \int f d\mu$.

Proof: $\alpha \cdot \mu(\{f \geq \alpha\}) \leq \int_{\{f\geq\alpha\}} f d\mu \leq \int f d\mu$.

9.4 Random variables: Measurable mappings $f : X \to Y$ on probability spaces $(X; \mathcal{A}; \mu)$ habitually are denoted as random variables with their **expected value** $E(f) := \int f d\mu$ and **probability**

distribution $\mu_f := f(\mu)$. The random variables $(f_i)_{i\in I}$ with $f_i : (X;\mathcal{A};\mu) \to (Y_i;\mathcal{A}_i)$ are **independent** if the σ-algebrae $\left(f_i^{-1}(\mathcal{A}_i)\right)_{i\in I}$ with $f_i^{-1}(\mathcal{A}_i) \subset \mathcal{A}$ are independent, i.e. if for all $i, j \in I$ and $A_i \in \mathcal{A}_i, A_j \in \mathcal{A}_j$ holds $\mu\left(f_i^{-1}[A_i] \cap f_j^{-1}[A_j]\right) = \mu_{f_i}(A_i) \cdot \mu_{f_j}(A_j)$. For **real-valued** random variables $f : X \to \mathbb{R}$ we have $0 \le E\left((f - E(f))^2\right) = E(f^2) - (E(f))^2$ and hence $E(f^2) \ge (E(f))^2$. Their **standard deviation** $\sigma(f) := \|f - E(f)\|_2 = \sqrt{E(f^2) - (E(f))^2} = \sigma(f - E(f))$ is independent of the expected value and hence is preserved if we examine the **centered random variable** $f - E(f)$.

9.5 Expected values of products of independent random variables: For **independent** and **real** random variables $f, g \in \mathcal{B}(X;\mathbb{R})$ we have $E(f \cdot g) = E(f) \cdot E(g)$.

Proof: On account of $E(\chi_A \cdot \chi_B) = E(\chi_{A\cap B}) = \mu(A \cap B) = \mu(A) \cdot \mu(B) = E(\chi_A) \cdot E(\chi_B)$ the proposition holds for **characteristic** functions and due to the linearity of the integral also for **step** functions $\varphi, \psi \in \mathcal{S}(X;\mathbb{R})$. For **integrable** functions $f, g \in \mathcal{B}(X;\mathbb{R})$ with μ-a.e. $f = \lim\limits_{n\to\infty} \varphi_n$ resp. $g = \lim\limits_{n\to\infty} \psi_n$ for sequences $(\varphi_n)_{n\in\mathbb{N}}, (\psi_n)_{n\in\mathbb{N}} \subset \mathcal{S}(X;\mathbb{R})$ according to 5.5 we have μ-a.e. $f \cdot g = \lim\limits_{n\to\infty} (\varphi_n \cdot \psi_n)$. According to the hypothesis $E(\varphi_n \cdot \psi_n) = E(\varphi_n) \cdot E(\psi_n) \le 2E(f) \cdot E(g) < \infty$ holds for $n \ge N$ and some $N \in \mathbb{N}$ so that we can apply **monotone convergence** 5.12 to obtain $E(f \cdot g) = \lim\limits_{n\to\infty} E(\varphi_n \cdot \psi_n) = \lim\limits_{n\to\infty} (E(\varphi_n) \cdot E(\psi_n)) = \lim\limits_{n\to\infty} E(\varphi_n) \cdot \lim\limits_{n\to\infty} E(\psi_n) = E(f) \cdot E(g)$.

9.6 Median: The real number $m(f)$ is a **median** of the random variable $f : X \to \mathbb{R}$ iff $\mu(f \le m(f)) \ge \frac{1}{2} \le \mu(f \ge m(f))$. Obviously for two medians $m_1(f) < m_2(f)$ every intermediate value $m_1(f) < \alpha < m_2(f)$ is a median too. The **minimal median** is $m_{min}(f) = \inf\left\{\lambda \in \mathbb{R} : \mu(f \le \lambda) \ge \frac{1}{2}\right\} = \inf\left\{\lambda \in \mathbb{R} : \mu(f > \lambda) \le \frac{1}{2}\right\}$ since due to the **continuity from above** 2.2.3 on the one hand we have $\mu(f \le m_{min}(f)) = \mu\left(\bigcap_{n\ge1}\left\{f \le m_{min}(f) + \frac{1}{n}\right\}\right) = \inf\limits_{n\ge1} \mu\left(f \le m_{min}(f) + \frac{1}{n}\right) \ge \frac{1}{2}$ and on the other hand $\mu(f \ge m_{min}(f)) = \mu\left(\bigcap_{n\ge1}\left\{f \ge m_{min}(f) - \frac{1}{n}\right\}\right) = \inf\limits_{n\ge1} \mu\left(f \ge m_{min}(f) - \frac{1}{n}\right) = 1 - \sup\limits_{n\ge1} \mu\left(f < m_{min}(f) - \frac{1}{n}\right) \ge \frac{1}{2}$, i.e. $m_{min}(f)$ is itself a **median** and since for every $\epsilon > 0$ holds $\mu(f \le m_{min}(f) - \epsilon) < \frac{1}{2}$ it is the **minimal median**. Correspondingly the **maximal median** is $m_{max}(f) = \sup\left\{\lambda \in \mathbb{R} : \mu(f \ge \lambda) \ge \frac{1}{2}\right\} = \sup\left\{\lambda \in \mathbb{R} : \mu(f < \lambda) \le \frac{1}{2}\right\}$. The relation $m_{min}(f) \le m_{max}(f)$ holds since otherwise we had $\sup\limits_{n\ge1} \mu\left(f \ge m_{max}(f) + \frac{1}{n}\right) = \mu\left(\bigcup_{n\ge1}\left\{f \ge m_{max}(f) + \frac{1}{n}\right\}\right)$ $= \mu(f > m_{max}(f)) > \frac{1}{2}$, i.e. there existed a $\lambda = m_{max}(f) + \frac{1}{n}$ with $\mu(f \ge \lambda) \ge \frac{1}{2}$ contrary to the definition of $m_{max}(f)$. Obviously we have **linearity** in the form $c \cdot m(f) = m(c \cdot f)$ and $m(f) + c = m(f + c)$ for every $c \in \mathbb{R}$.

9.7 Lévy's inequality: For **independent** and **real** random variables $f_i : (X, A, \mu) \to \mathbb{R}$, $1 \le i \le n$ with sums $F_m := \sum\limits_{i=1}^{m} f_i$ and every $\epsilon > 0$ we have $\mu\left(\max\limits_{1\le i\le n} |F_i + m(F_n - F_i)| \ge \epsilon\right) \le 2\mu(|F_n| \ge \epsilon)$.

Note: This inequality allows us to obtain an estimate for the maximal deviation $|F_i + m(F_n - F_i)|$ of **all** partial sums F_i given the measure of the deviation $|F_n|$ of the **single** sum F_n.

Proof: For $F_0 := 0$ and $T = \min\limits_{1\le i\le m} \{|F_i + m(F_n - F_i)| \ge \epsilon\}$ if such an i exists and $T := n + 1$ otherwise the **pairwise disjoint** sets $A_i := \{T = i\} \in \sigma(f_1, ..., f_i)$ are **independent** of $B_i = \{F_n - F_i \ge m(F_n - F_i)\} \in \sigma(f_i, ..., f_n)$. Hence from $\mu(B_i) \ge \frac{1}{2}$ follows $\mu(F_n \ge \epsilon) \ge \mu\left(\bigcup\limits_{i=1}^{n} A_i \cap B_i\right)$ $= \sum\limits_{i=1}^{n} \mu(A_i \cap B_i) = \sum\limits_{i=1}^{n} \mu(A_i) \cdot \mu(B_i) \ge \frac{1}{2}\mu(1 \le T \le n) = \frac{1}{2}\mu\left(\max\limits_{1\le i\le n} F_i + m(F_n - F_i) \ge \epsilon\right)$. Since the same inequality holds for $-f_i$ resp. $-F_i$ with $m(-F_n + F_i) = -m(F_n - F_i)$ and all corresponding sets are **disjoint** we can use the **additivity** of μ and simply **add** the two inequalities to obtain the assertion.

9.8 Lévy's convergence theorem: For the sequence $(F_n)_{n\ge1}$ of the sums $F_n := \sum\limits_{i=1}^{n} f_i$ of **real** and **independent** random variables $(f_i)_{i\ge1}$ the **μ-a-e- convergence** is **equivalent** to the **convergence in measure.**

Proof:

⇒: **Lebesgue's convergence theorem** 4.11.

⇐: **Riesz' convergence theorem** 4.13.3 provides for every $\frac{1}{4} > \epsilon > 0$ an $n_\epsilon \geq 1$ with $\mu\left(|F_n - F_m| \geq \epsilon\right) < \epsilon$ for all $n > m \geq n_\epsilon$. In particular we have $\mu\left(|F_n - F_m| \geq \epsilon\right) < \frac{1}{2}$ and hence $|m\left(F_n - F_m\right)| \leq \epsilon$ for $n > m \geq n_\epsilon$. The preceding inequality yields $\mu\left(\max\limits_{m<i\leq n} |F_i - F_m| \geq 2\epsilon\right) \leq 2\mu\left(|F_n - F_m| \geq \epsilon\right) < 2\epsilon$. For $n \to \infty$ follows $\mu\left(\sup\limits_{m<i} |F_i - F_m| \geq 2\epsilon\right) \leq 2\epsilon$ and due to the **completeness** 4.14 of the μ-a-e-convergence we obtain the assertion.

9.9 Abel's partial summation:

1. For two **real** sequences $(a_i)_{i\geq 0}, (b_i)_{i\geq 0} \subset \mathbb{R}$ and $A_n = \sum\limits_{i=0}^{n} a_i$ we have
   $\sum\limits_{i=1}^{n} a_i b_i = A_n b_n - A_0 b_1 - \sum\limits_{i=1}^{n-1} A_i\left(b_{i+1} - b_i\right)$ for $n \geq 1$.
2. If also $\lim\limits_{n\to\infty} A_n = A_0^* < \infty$ with $A_n^* = \sum_{i>n} a_i$ holds we have
   $\sum\limits_{i=1}^{n} a_i b_i = A_0^* b_1 - A_n^* b_n + \sum\limits_{i=1}^{n-1} A_i^*\left(b_{i+1} - b_i\right)$ für $n \geq 1$.
3. If additionally $a_i \geq 0$ and $b_{i+1} \geq b_i \geq 0$ for all $i \geq 0$ is satisfied we have
   $\sum\limits_{i=1}^{n} a_i b_i = A_0^* b_1 + \sum\limits_{i=1}^{n-1} A_i^*\left(b_{i+1} - b_i\right)$ for $n \geq 1$.

Proof:

1. $\sum\limits_{i=1}^{n} a_i b_i = \sum\limits_{i=0}^{n-1}\left(A_{i+1} - A_i\right) b_{i+1} = A_n b_n - \sum\limits_{i=1}^{n-1} A_i\left(b_{i+1} - b_i\right) - A_0 b_1$.
2. Follows from 1. with $a_0 = -\sum\limits_{i=1}^{\infty} a_i = -A_0^*$.
3. In the case of $\lim\limits_{n\to\infty} A_n^* b_n > 0$ with $\sum_{i>n} a_i b_i \geq A_n^* b_n$ and 2. we have $A_0^* b_1 + \sum_{i\geq 1} A_i^*\left(b_{i+1} - b_i\right) \geq \sum_{i>1} a_i b_i = \infty$ and hence the assertion. For $\lim\limits_{n\to\infty} A_n^* b_n = 0$ it directly follows from 2. with $n \to \infty$.

9.10 Kronecker's lemma: For a **positive real** and **increasing** sequence $(b_i)_{i\geq 1}$ with $\lim\limits_{i\to\infty} \frac{1}{b_i} = 0$ and a further **real** sequence $(a_i)_{i\geq 1}$ with $\sum_{i\geq 1} \frac{a_i}{b_i} < \infty$ we have $\lim\limits_{n\to\infty} \frac{1}{b_n} \sum\limits_{i=1}^{n} a_i = 0$.

Proof: From 9.9.2 with $c_i = \frac{a_i}{b_i}$ and $\lim\limits_{n\to\infty} C_n = C_0^* = \sum_{i\geq 1} \frac{a_i}{b_i} < \infty$ resp. $\lim\limits_{n\to\infty} C_n^* = 0$ we have the decomposition $\frac{1}{b_n} \sum\limits_{i=1}^{n} a_i = \frac{1}{b_n} \sum\limits_{i=1}^{n} c_i b_i = \frac{1}{b_n} C_0^* b_1 + C_n^* + \frac{1}{b_n} \sum\limits_{i=1}^{n-1} C_i^*\left(b_{i+1} - b_i\right)$. For $n \to \infty$ the first two summands converge to zero. This also holds for the third summand since for every $\epsilon > 0$ there is an $m \geq 1$ with $|C_i^*| < \epsilon$ for all $i \geq m$ such that on the one hand $\left|\frac{1}{b_n} \sum\limits_{i=m}^{n-1} C_i^*\left(b_{i+1} - b_i\right)\right| < \epsilon \frac{1}{b_n} \sum\limits_{i=m}^{n-1} \left(b_{i+1} - b_i\right) = \epsilon\left(1 - \frac{b_m}{b_n}\right) < \epsilon$ and on the other hand $\left|\frac{1}{b_n} \sum\limits_{i=1}^{m-1} C_i^*\left(b_{i+1} - b_i\right)\right| < \epsilon$ for a sufficiently large $n \geq 1$.

9.11 Khintchin-Kolmogorov convergence theorem For every sequence $(f_n)_{n\geq 1}$ **of independent and centered random variables** $f_n \in L^2(\mu)$ wih $\sum_{n\geq 1} E\left(f_n^2\right) < \infty$ the **sums** $F_m := \sum\limits_{n=1}^{m} f_n$ **converge μ-a.e.** and in **quadratic mean** to a $F = \lim\limits_{m\to\infty} F_m \in L^2(\mu)$ with $E(F)^2 = \sum_{n\geq 1} E\left(f_n^2\right)$.

Proof : Owing to 9.5, $E(f_n) = 0$ for all n≥ 1 and by the hypothesis we have $\lim\limits_{k\to\infty} \sup\limits_{m\geq k} E\left(F_m - F_k\right)^2 = \lim\limits_{k\to\infty} \sup\limits_{m\geq k} \sum\limits_{i=k}^{m} E\left(f_i^2\right) = 0$ such that due to 6.7 there is an $F = \lim\limits_{k\to\infty} F_{m(k)} \in L^2(\mu)$ with a **μ-a.e.**

convergent partial sequence $\left(F_{m(k)}\right)_{k\geq 1}$ as well as convergence of the complete sequence in the **quadratic mean**: $\lim\limits_{m\to\infty} E\left(F-F_m\right)^2 = 0$. Owing to 6.9 we can infer the convergence **in measure** and due to **Lévy's theorem** 9.8 **μ-a.e.** convergence of the **complete series.** Due to 9.5 and $E\left(f_n\right)=0$ we also obtain $E\left(F\right)^2 = \lim\limits_{m\to\infty} E\left(F_m\right)^2 = \sum_{n\geq 1} E\left(f_n^2\right)$.

9.12 Kolmogorov's strong law of large numbers: For every sequence $\left(f_n\right)_{n\geq 1}$ of **independent, identically distributed** and **integrable random variables** the **mean values** $\frac{1}{m}F_m := \frac{1}{m}\sum\limits_{n=1}^{m} f_n$ **converge μ-a.e.** to the **common mean** $E\left(f_1\right) = \lim\limits_{m\to\infty}\frac{1}{m}F_m$.

Note: The strong law of large numbers provides a mathematical basis for the principle of learning from experience and every statistical method in science. From the mean results $\frac{1}{m}F_m$ of independent trials executed under similar conditions in the **past** we infer the expected outcome $E\left(f_1\right)$ in the **future**.

Proof: At first we prove the proposition for **truncated random variables** $g_n = \frac{1}{n}\cdot f_n\cdot\chi_{\{|f_n|\leq n\}}$. Subsequently we show that the deviations from f_n vanish μ-a.e. for $n\to\infty$:

With the sets $A_m = \{m-1<|f_1|\leq m\}$ we obtain $\sum_{n\geq 1} E\left(|g_n|^2\right) = \sum_{n\geq 1}\sum_{n\geq m\geq 1} n^{-2}\int_{A_m}|f_1|^2\,d\mu$ $= \sum_{m\geq 1}\sum_{n\geq m} n^{-2}\int_{A_m}|f_1|^2\,d\mu \leq \sum_{m\geq 1}\frac{2}{m}\int_{A_m}|f_1|^2\,d\mu \leq 2\sum_{m\geq 1}\int_{A_m}|f_1|\,d\mu \leq 2E\left(|f_n|\right) < \infty$ such that due to **Khintchin - Kolmogorov** 9.11 we have μ-a.e. $\sum_{n\geq 1}\left(g_n - E\left(g_n\right)\right) < \infty$. The deviations have the measure $\sum_{n\geq 1}\mu\left(\frac{1}{n}f_n \neq g_n\right) = \sum_{n\geq 1}\mu\left(|f_1|>n\right) \leq \sum_{n\geq 1}\sum_{m\geq n}\mu\left(m+1\geq|f_1|>m\right) \leq$ $\sum_{m\geq 1}\sum_{m\geq n\geq 1}\mu\left(m+1\geq|f_1|>m\right) = \sum_{m\geq 1}\left(m+1\right)\cdot\mu\left(m+1\geq|f_1|>m\right) \leq E\left(|f_1|\right) < \infty$ such that according to **Borel-Cantelli** 4.12 follows $\mu\left(\bigcap_{m\geq 1}\bigcup_{n\geq m}\left\{\frac{1}{n}f_n \neq g_n\right\}\right) = 0$ and with the result from the first estimate above we obtain μ-a.e. $\sum_{n\geq 1}\frac{1}{n}\left(f_n - E\left(n\cdot g_n\right)\right) = \sum_{n\geq 1}\left(\frac{1}{n}\cdot f_n - E\left(g_n\right)\right) < \infty$. On account of $\lim\limits_{m\to\infty}\frac{1}{m}\sum\limits_{n=1}^{m} E\left(n\cdot g_n\right) = \lim\limits_{m\to\infty}\frac{1}{m}\sum\limits_{n=1}^{m} E\left(f_1\cdot\chi_{\{|f_1|\leq n\}}\right) = \lim\limits_{m\to\infty} E\left(f_1\cdot\chi_{\{|f_1|\leq n\}}\right) = E\left(f_1\right)$ and **Kronecker** 9.10 follows $\lim\limits_{m\to\infty}\frac{1}{m}F_m - E\left(f_1\right) = \lim\limits_{m\to\infty}\frac{1}{m}\sum\limits_{n=1}^{m}\left(f_n - E\left(n\cdot g_n\right)\right) = 0$.

10 Measures with densities

10.1 Complex measure and total variation: A **complex measure** is a **complex** and **σ-additive** set function $\mu:\mathcal{A}\to\mathbb{C}$ on a measurable space $\left(X;\mathcal{A}\right)$. Contrary to the **positive measure** $\mu:\mathcal{A}:\to[0;\infty]$ defined in 3.1 the complex measure is **finite**. According to the **theorem of Lévy und Steinitz** ([10, 3.9]) the **σ-additivity** $\mu\left(\mathring{\bigcup}_{n\in\mathbb{N}}A_n\right) = \sum_{n\in\mathbb{N}}\mu\left(A_n\right) < \infty$ resp. the **interchangeability of the union** entail the **absolute convergence** of the series.

So its **total variation** $|\mu|:\mathcal{A}:\to\mathbb{R}$ with $|\mu|\left(A\right) := \sup\left\{\sum_{n\in\mathbb{N}}|\mu\left(A_n\right)| : \left(A_n\right)_{n\in\mathbb{N}}\subset\mathcal{A} : \mathring{\bigcup}_{n\in\mathbb{N}}A_n = A\right\}$ is well defined as well as **σ-additive**: On the one hand for every $A_m\in\mathcal{A}$ and $\epsilon>0$ there is a partition $\left(A_{mn}\right)_{n\in\mathbb{N}}\subset\mathcal{A}$ with $|\mu|\left(A_m\right) - \epsilon\cdot 2^{-m-1} < \sum_{n\in\mathbb{N}}|\mu\left(A_{mn}\right)| \leq |\mu|\left(A_m\right)$ such that $\sum_{m\in\mathbb{N}}|\mu|\left(A_m\right) - \epsilon <$ $\sum_{m,n\in\mathbb{N}}|\mu\left(A_{mn}\right)| \leq \sum_{m\in\mathbb{N}}|\mu|\left(A_m\right)$ and hence $\sum_{m\in\mathbb{N}}|\mu|\left(A_m\right) \leq |\mu|\left(\mathring{\bigcup}_{m\in\mathbb{N}}A_m\right)$. On the other hand for **every** partition $\left(B_n\right)_{n\in\mathbb{N}}\subset\mathcal{A}$ with $\mathring{\bigcup}_{n\in\mathbb{N}}B_n = \mathring{\bigcup}_{m\in\mathbb{N}}A_m$ the intersections $\left(B_n\cap A_m\right)_{n\in\mathbb{N}}$ partition A_m while the intersections $\left(B_n\cap A_m\right)_{m\in\mathbb{N}}$ partition B_n such that due to the σ-additivity of μ holds $\sum_{n\in\mathbb{N}}|\mu\left(B_n\right)| \leq \sum_{m,n\in\mathbb{N}}|\mu\left(A_m\cap B_n\right)| \leq \sum_{m\in\mathbb{N}}|\mu|\left(A_m\right)$. This estimate carries over to the suprema such that $|\mu|\left(\mathring{\bigcup}_{m\in\mathbb{N}}A_m\right) \leq \sum_{m\in\mathbb{N}}|\mu|\left(A_m\right)$. Hence $|\mu|$ is a **measure**.

10.2 Lemma: For any n complex $z_1,...,z_n$ there is a subset $S\subset\{1;...;n\}$ with $|\sum_{k\in S}z_k| \geq \frac{1}{\pi}\sum\limits_{i=1}^{n}|z_i|$.

Proof: For $z_i = |z_i|\cdot e^{i\alpha_i}$ and $-\pi\leq\vartheta\leq\pi$ let $S\left(\vartheta\right) := \{1\leq k\leq n : \cos\left(\alpha_k-\vartheta\right)>0\}$. Then for every such ϑ we have $|\sum_{k\in S}z_k| = \left|\sum_{k\in S}e^{-i\vartheta}\cdot z_k\right| \geq \mathrm{Re}\left(\sum_{k\in S}e^{-i\vartheta}\cdot z_k\right) = \sum_{k\in S}|z_k|\cdot\cos\left(\alpha_k-\vartheta\right) \geq \sum\limits_{i=1}^{n}$ $|z_i|\cdot\cos^+\left(\alpha_k-\vartheta\right)$ and the maximal value of the sum on the right hand side attained for say $\vartheta=\vartheta_0$ is not less than the average $\frac{1}{2\pi}\int_{-\pi}^{\pi}\left(\sum\limits_{i=1}^{n}|z_i|\cdot\cos^{+}(\alpha_k-\vartheta)\right)d\vartheta=\frac{1}{\pi}\sum\limits_{i=1}^{n}|z_i|$ which proves the lemma for $S:=S(\vartheta_0)$.

10.3 Theorem: The total variation $|\mu|$ of a complex measure μ is **finite**.

Proof: Assuming $|\mu|(X)=\infty$ there must be a partition $(A_i)_{i\in\mathbb{N}}\subset\mathcal{A}$ of X and an $n\in\mathbb{N}$ with $\frac{1}{\pi}\sum\limits_{i=1}^{n}|\mu(A_i)|>|\mu(X)|+1$. Due to 10.2 there is a subset $S\subset\{1;...;n\}$ such that for $B_1:=\bigcup_{k\in S}A_i$ on the one hand $|\mu(B_1)|=\left|\sum_{k\in S}\mu(A_i)\right|>|\mu(X)|+1\geq 1$ and on the other hand $|\mu(X\setminus B_1)|=|\mu(X)-\mu(B_1)|\geq|\mu(B_1)|-|\mu(X)|\geq 1$. According to the hypothesis we have either $|\mu|(B_1)=\infty$ or $|\mu|(X\setminus B_1)=\infty$ and assuming this being the case for $X\setminus B_1$ we can repeat the argument from above to split off a subset $B_2\subset X\setminus B_1$ with $|\mu|(X\setminus(B_1\cup B_2))=\infty$ and $|\mu(B_2)|\geq 1$. Hence by induction we obtain a sequence $(B_n)_{n\in\mathbb{N}}\subset\mathcal{A}$ of paiwise disjoint sets B_n with $|\mu(B_n)|\geq 1\forall n\in\mathbb{N}$ and consequently $\left|\mu\left(\mathring{\bigcup}_{n\in\mathbb{N}}B_n\right)\right|=\left|\sum_{n\in\mathbb{N}}\mu(B_n)\right|=\infty$ contrary tor the finite character of μ according to definition 10.1.

10.4 Theorem: The set $\mathcal{M}(\mathcal{A},\mathbb{C})$ of complex measures on a measurable space $(X;\mathcal{A})$ with the operations $(\lambda+\mu)(A):=\lambda(A)+\mu(A)$ resp. $(c\cdot\lambda)(A):=c\cdot\lambda(A)$ for $A\in\mathcal{A}$, $c\in\mathbb{C}$, $\lambda,\mu\in\mathcal{M}$ and the **norm** $\|\mu\|:=|\mu|(X)$ is a **Banach space**.

Proof: The vector space axioms are clearly satisfied. The **positive definiteness** $\|\mu\|=0\Rightarrow\mu=0$ follows from the **monotonicity** $A\subset B\Rightarrow|\mu|(A)\leq|\mu|(B)$ of the total variation. With regard to the **completeness** for every Cauchy sequence $(\mu_n)_{n\in\mathbb{N}}\subset\mathcal{M}(\mathcal{A},\mathbb{C})$ and every measurable set $A\in\mathcal{A}$ we have $|\mu_n(A)-\mu_m(A)|=|(\mu_n-\mu_m)(A)|\leq|\mu_n-\mu_m|(A)\leq|\mu_n-\mu_m|(X)=\|\mu_n-\mu_m\|$ such that the corresponding Cauchy sequence $(\mu_n(A))_{n\in\mathbb{N}}\subset\mathbb{C}$ converges to a complex number $\mu(A)$ hence defining a complex set function $\mu:\mathcal{A}\to\mathbb{C}$. For a sequence of disjoint measurable sets $(A_k)_{k\in\mathbb{N}}\subset\mathcal{A}$ and every $k\in\mathbb{N}$ there is an $n_k\in\mathbb{N}$ with $|\mu_n(A_k)-\mu(A_k)|\leq\epsilon 2^{-k}$ for every $n\geq n_k$ such that for every $N\geq\max\{n_k:k\leq m\}$ and $\sum\limits_{k=0}^{m}\mu_N(A_k)=\mu_N\left(\bigcup\limits_{k=0}^{m}A_k\right)$ we have $\left|\sum\limits_{k=0}^{m}\mu(A_k)-\mu\left(\bigcup\limits_{k=0}^{m}A_k\right)\right|=\left|\sum\limits_{k=0}^{m}\mu(A_k)-\sum\limits_{k=0}^{m}\mu_N(A_k)+\mu_N\left(\bigcup\limits_{k=0}^{m}A_k\right)-\mu\left(\bigcup\limits_{k=0}^{m}A_k\right)\right|\leq\epsilon 2^{-m+1}+\left|\mu_N\left(\bigcup\limits_{k=0}^{m}A_k\right)-\mu\left(\bigcup\limits_{k=0}^{m}A_k\right)\right|\leq\epsilon 2^{-m+2}$ for a suitably large N. Since ϵ and m are arbitrary we have shown the σ-additivity $\sum\limits_{k=0}^{\infty}\mu(A_k)=\mu\left(\bigcup\limits_{k=0}^{\infty}A_k\right)$, i.e. $\mu\in\mathcal{M}$. Assuming there is an $\epsilon>0$ with $\|\mu-\mu_n\|=\sup\left\{\sum_{k\in\mathbb{N}}|(\mu-\mu_n)(A_k)|:(A_k)_{k\in\mathbb{N}}\subset\mathcal{A}:\mathring{\bigcup}_{k\in\mathbb{N}}A_k=X\right\}\leq\epsilon$ for every $n\in\mathbb{N}$ we find an $B_n=\bigcup\limits_{k=0}^{K_n}A_k\in\mathcal{A}$ with $|(\mu-\mu_n)(B_n)|\geq\frac{\epsilon}{2}$ whence $|(\mu-\mu_n)(B)|\geq\frac{\epsilon}{2}$ for $B=\bigcup_{n\in\mathbb{N}}B_n$ and every $n\in\mathbb{N}$ contrary to $(\mu_n(B))$ converging to $\mu(B)$. Hence $\lim\limits_{n\to\infty}\|\mu-\mu_n\|=0$

10.5 Continuous and singular measures: A complex or positive measure μ is **λ-absolutely continuous** with respect to the **positive** measure λ on the same measurable space $(X,\mathcal{A})$ with the notation $\mu\lessdot\lambda$ iff $\lambda(A)=0\Rightarrow\mu(A)=0\forall A\in\mathcal{A}$. The measure μ is **concentrated** on the set $A\in\mathcal{A}$ iff $\lambda(B)=\mu(B\cap A)\forall B\in\mathcal{A}$ resp. $\mu(B)=0\Leftrightarrow A\cap B=\emptyset$. The measures μ and λ are **mutually singular** with the notation $\mu\perp\lambda$ iff μ and λ are concentrated on two disjoint sets. These relations have the following properties:

1. If μ is concentrated on A the so is $|\mu|$ since for every partition $(E_m)_{m\in\mathbb{N}}$ of the set $E\in\mathcal{A}$ with $E\cap A=\emptyset$ we have $\mu(E_m)=0\,\forall m\in\mathbb{N}$.
2. $\mu\perp\lambda\Rightarrow|\mu|\perp|\lambda|$ due to 1.
3. $\mu\lessdot\lambda\Rightarrow|\mu|\lessdot\lambda$ since from $\lambda(A)=0$ for every partition $(A_m)_{m\in\mathbb{N}}$ of A follows $\mu(A_m)=\lambda(A_m)=0\,\forall m\in\mathbb{N}$.
4. $\mu\perp\lambda\wedge\mu\lessdot\lambda\Rightarrow\mu=0$ is obvious.
5. $\mu_1\perp\lambda\wedge\mu_2\perp\lambda\Rightarrow\mu_1+\mu_2\perp\lambda$ since if μ_1, μ_2 and λ are concentrated on A_1, A_2 resp. B with $A_1\cap B=A_2\cap B=\emptyset$ the measure $\mu_1+\mu_2$ is concentrated on $A_1\cup A_2$ with $(A_1\cup A_2)\cap B=\emptyset$.

6. $\mu_1 \lessdot \lambda \wedge \mu_2 \lessdot \lambda \Rightarrow \mu_1 + \mu_2 \lessdot \lambda$ is obvious.
7. $\mu_1 \perp \lambda \wedge \mu_2 \lessdot \lambda \Rightarrow \mu_1 \perp \mu_2$ since if μ_1 is concentrated on A we have $\mu_1(A) \neq 0$ and hence $\mu_2(A) = \lambda(A) = 0$, i.e. μ_2 is concentrated on $X \setminus A$.

10.6 ϵ-δ-definition of absolute contiuity: A **complex** measure μ is **absolutely continuous** with respect to the **positive** measure λ iff for every $\epsilon > 0$ exists a $\delta > 0$ such that for every $A \in \mathcal{A}$ holds: $\lambda(A) < \delta \Rightarrow |\mu|(A) < \epsilon$.

Proof:

$\Rightarrow$: Assuming an $\epsilon > 0$ and a sequence $(A_n)_{n\in\mathbb{N}} \subset \mathcal{A}$ with $\lambda(A_n) < 2^{-n}$ but $|\mu(A_n)| \geq \epsilon$ then $(B_m)_{m\in\mathbb{N}} \subset \mathcal{A}$ with $B_m = \bigcup_{n\geq m} A_n$ is a decreasing sequence of measurable sets with $\lambda(B_m) < 2^{-m+1}$ and $\lambda\left(\bigcap_{m\in\mathbb{N}} B_m\right) = 0$ on account of the **continuity from above** 2.2.3. But the measure $|\mu|$ is also continuous from above such that $|\mu|\left(\bigcap_{m\in\mathbb{N}} B_m\right) = \lim\limits_{m\to\infty} |\mu|(B_m) \geq \inf\limits_{m\in\mathbb{N}} |\mu|(A_m) \geq \epsilon$ contrary to the hypothesis $|\mu| \lessdot \lambda$ resp. 10.5.3.

$\Leftarrow$: $\lambda(A) = 0 \Rightarrow |\mu|(A) < \epsilon \forall \epsilon > 0 \Rightarrow |\mu(A)| \leq |\mu|(A) = 0$.

10.7 Jordan decomposition of signed measures: The real and complex parts of complex measures are **finite** and are called **signed measures** to distinguish them from the **positive measures**. The **Jordan decomposition** $\mu = \mu^+ - \mu^-$ resp. $|\mu| = \mu^+ + \mu^-$ of a signed measure μ splits it into its **positive and negative variations** $\mu^+ = \frac{1}{2}(|\mu| + \mu)$ resp. $\mu^- = \frac{1}{2}(|\mu| - \mu)$ both being **finite** and **positive**. On account of the **σ-additivity** the **total variation** of a **positive signed measure** coincides with the measure itself: $|\mu^+| = \mu^+$ bzw. $|\mu^-| = \mu^-$.

10.8 Theorem of Lebesgue Radon-Nikodym: For a **positive, σ-finite measure** $\lambda : \mathcal{A} \to [0;\infty]$ and a **complex measure** $\mu : \mathcal{A} \to \mathbb{C}$ on a common measurable space $(X;\mathcal{A})$ exist:

1. a uniquely determined **Lebesgue decomposition** of $\mu = \mu_a + \mu_s$ with respect to λ into two **complex** measures μ_a and μ_s such that $\mu_a \lessdot \lambda$ and $\mu_s \perp \lambda$.
2. a uniquely determined **Radon-Nikodym density** or **derivative** $\frac{d\mu_a}{d\lambda} \in L^1(\lambda)$ with $\mu_a(A) = \int_A \frac{d\mu_a}{d\lambda}\, d\lambda$ for every $A \in \mathcal{A}$.

Proof: The **Lebesgue decomposition** is uniquely determined since for every other decomposition μ'_a and μ'_s we have $\mu'_a - \mu_a \overset{10.4.6}{\lessdot} \sum \lambda$ bzw. $\mu_s - \mu'_s \overset{10.4.5}{\perp} \lambda$ and hence $\mu'_a - \mu_a \overset{10.4.4}{=} \mu_s - \mu'_s = 0$. The uniqueness of the **Radon-Nikodym density** follows from 5.6.3 resp. 10.6.

We start the **construction of the decomposition** with $w = \sum_{n\in\mathbb{N}} \frac{\chi_{A_n}}{2^{n+1}\cdot(1+\lambda(A_n))} : X \to]0;1[$ for a countable cover $(A_n)_{n\in\mathbb{N}} \subset \mathcal{A}$ of X with $\lambda(A_n) < \infty \forall n \in \mathbb{N}$ such that the measure ν with $\nu(A) := \int_A w\, d\lambda$ is **finite** and due to $w > 0$ possesses the same **null sets** as λ. Then $\varphi = |\mu| + \nu$ is again a **positive** and **finite** measure with $\int f\, d\varphi = \int f\, d|\mu| + \int fw\, d\lambda$ for every **step function** f and due to 5.4 for **positive measurable** f. Applying 10.5.1, the **Schwarz inequality** 6.4.1 and the finite character of φ for every $f \in L^2(\varphi)$ we obtain $\left|\int f\, d|\mu|\right| \leq \int |f|\, d\mu \leq \int |f|\, d\varphi \leq \left(\int |f|^2\, d\varphi\right)^{\frac{1}{2}} \cdot (\varphi(X))^{\frac{1}{2}} < \infty$. In particular for every null sequence $(f_n) \subset L^2(\varphi)$ with $(\|f_n\|_2)_n \to 0$ we have $\left(\left|\int f_n\, d|\mu|\right|\right)_n \to 0$, i.e. the **linear functional** $I_\mu : L^2(\varphi) \to [0;\infty[$ with $I_\mu f = \int f\, d|\mu|$ is **continuous at the origin**. According to [9, 20.11] it is also **bounded** resp. **uniformly continuous** and hence a member of the **dual space** $(L^2(\varphi))^*$. Due to [8, p 308 Th 12.5] I_μ possesses a φ-a.e. uniquely determined representant $g \in L^2(\varphi)$ with respect to the **inner product** $\int f\, d|\mu| = I_\mu f = \langle f, g\rangle = \int fg\, d\varphi$ resp. $\int (1-g) f\, d|\mu| = \int fgw\, d\lambda$ for every positive measurable f. We keep this result in mind as equation (X). Choosing $f = \chi_A$ for every $A \in \mathcal{A}$ with $\varphi(A) > 0$ we obtain $0 \leq \int_A g\, d\varphi = |\mu|(A) \leq \varphi(A)$ and hence φ-a.e. $0 \leq g \leq 1$. The **Lebesgue decomposition** of the **total variation** $|\mu| = \mu_a + \mu_s$ can now be given by $\mu_a = |\mu|_{\{g<1\}}$ and $\mu_s = |\mu|_{\{g=1\}}$: Substituting $f = \chi_{\{g=1\}}$ in equation (X) yields $0 = \int_{\{g=1\}} w\, d\lambda$ such that on account of $w(x) > 0$ follows $\lambda(\{g = 1\}) = 0$ and hence $\mu_s \perp \lambda$. The **Radon-Nikodym density** is $\frac{d\mu_a}{d\lambda} = w \sum\limits_{n=1}^{\infty} g^n$ such that $\frac{d\mu_a}{d\lambda}(x) = \frac{w(x)\cdot g(x)}{1-g(x)}$ in the case of $g(x) < 1$ and $\frac{d\mu_a}{d\lambda}(x) = \infty$ else: Substituting $f = \chi_A \cdot \sum\limits_{n=0}^{m} g^n$ in equation (X) we obtain $\int_A \left(1 - g^{m+1}\right) d|\mu| = \int_A w \cdot \sum\limits_{n=1}^{m+1} g^n\, d\lambda$ and taking recourse to **monotone convergence** 5.12 for $m \to \infty$ leads to $\mu_a(A) = \int_A \frac{d\mu_a}{d\lambda}\, d\lambda$ which also yields $\mu_a \lessdot \lambda$. The boundedness of $|\mu|$ transfers to μ_a such that $\frac{d\mu_a}{d\lambda} \in L^1(\lambda)$. The Lebesgue decomposition for the **complex** measure $\mu = \mathrm{Re}\mu + i\mathrm{Im}\mu = (\mathrm{Re}\mu)^+ - (\mathrm{Re}\mu)^- + i\left((\mathrm{Im}\mu)^+ - (\mathrm{Im}\mu)^-\right)$ is accomplished by applying the above construction four times to the positive resp. negative variation of the rea resp. imaginary part of μ.

10.9 Polar representation of complex measures: For every **complex** measure μ exists a measurable complex function $\frac{d\mu}{d|\mu|} : X \to \mathbb{C}$ with $\left|\frac{d\mu}{d|\mu|}\right| = 1$ and $d\mu = \frac{d\mu}{d|\mu|}\, d|\mu|$.

Proof: According to10.8 and on account of $\mu \lessdot |\mu|$ there is a $\frac{d\mu}{d|\mu|} \in L^1$ with $d\mu = \frac{d\mu}{d|\mu|}\, d|\mu|$ which only has to be adapted to the absolute value $\left|\frac{d\mu}{d|\mu|}\right| = 1$: For a partition $(A_n)_{n\in\mathbb{N}}$ of the set $A = \left\{\left|\frac{d\mu}{d|\mu|}\right| < r\right\}$ holds $|\mu|(A) \le \sum_{n\in\mathbb{N}} |\mu(A_n)| = \sum_{n\in\mathbb{N}} \left|\int_{A_n} \frac{d\mu}{d|\mu|}\, d|\mu|\right| \le \sum_{n\in\mathbb{N}} r \cdot |\mu|(A_n) = r \cdot |\mu|(A)$, i.e. for $r < 1$ we have $|\mu|(A) = 0$ resp. μ-a.e. $\left|\frac{d\mu}{d|\mu|}\right| \ge 1$. On the other hand for every $A \in \mathcal{A}$ with $|\mu|(A) > 0$ holds $\left|\frac{1}{|\mu|(A)} \int_A \frac{d\mu}{d|\mu|}\, d|\mu|\right| = \frac{|\mu(A)|}{|\mu|(A)} \le 1$ so that we can apply the **mean value theorem** 5.19 with $S = \overline{B}_1(0)$ to obtain μ-a.e. $\left|\frac{d\mu}{d|\mu|}\right| \le 1$. Hence the assertion holds μ-a.e. and by redefining $\frac{d\mu}{d|\mu|} := 1$ on the μ-null set $\left\{\frac{d\mu}{d|\mu|} \neq 1\right\}$ we obtain the desired absolute value for every $x \in X$.

10.10 Corollary: For a **positive** measure λ and $h \in L^1(\lambda)$ with $d\mu = \frac{d\mu}{d\lambda}\, d\lambda$ we have $d|\mu| = \left|\frac{d\mu}{d\lambda}\right| d\lambda$.

Proof: Owing to 10.9 there is a $\frac{d\mu}{d|\mu|}$ with $\left|\frac{d\mu}{d|\mu|}\right| = 1$ so that $d\mu = \frac{d\mu}{d|\mu|}\, d|\mu|$ and hence $\frac{d\mu}{d|\mu|}\, d|\mu| = \frac{d\mu}{d\lambda}\, d\lambda$ resp. $d\,|\mu| = \overline{\frac{d\mu}{d|\mu|}} \frac{d\mu}{d\lambda}\, d\lambda$. From $|\mu| \ge 0$ and $\lambda \ge 0$ follows λ-a.e. $\overline{\frac{d\mu}{d|\mu|}} \frac{d\mu}{d\lambda} \ge 0$ and hence $\overline{\frac{d\mu}{d|\mu|}} \frac{d\mu}{d\lambda} = \left|\frac{d\mu}{d\lambda}\right|$.

10.11 Decomposition of complex measures: Every **complex** measure μ can be decomposed into four **positive** and **finite** measures according to $\mu = \mathrm{Re}\mu^+ - \mathrm{Re}\mu^- + i\left(\mathrm{Im}\mu^+ - \mathrm{Im}\mu^-\right)$.

Proof: Owing to 10.9 and the additivity of the integral for every measurable A we have $\mu(A) = \int \chi_A (\mathrm{Re}h)^+\, d|\mu| - \int \chi_A (\mathrm{Re}h)^-\, d|\mu| + i\left(\int \chi_A (\mathrm{Im}h)^+\, d|\mu| - \int \chi_A (\mathrm{Im}h)^-\, d|\mu|\right)$. Each of the four summands is a positive and finite measure with the σ-additivity resulting from the **monotone convergence** 5.12 in the form of $\mu\left(\mathring{\bigcup}_{n\in\mathbb{N}} A_n\right) = \int \left(\sum_{n\in\mathbb{N}} \chi_A\right) g d|\mu| = \sum_{n\in\mathbb{N}} \int \chi_A g d|\mu| = \sum_{n\in\mathbb{N}} \mu(A_n)$ for every **positive** and **real measurable** g.

10.12 Hahn decomposition for signed measures: The **Jordan decomposition** of a **signed** measure $\mu = \mu^+ - \mu^-$ extends to the measure space $(X; \mathcal{A}; \mu)$: There is a **Hahn decomposition** of X into two disjoint subsets $M^+ \cup M^- = X$ with $M^+ \cap M^- = \emptyset$ and $\mu^+(A) = \mu(A \cap M^+)$ resp. $\mu^-(A) = \mu(A \cap M^-)$ for every $A \in \mathcal{A}$.

Proof: Due to 10.10 there is a measurable $\frac{d\mu}{d|\mu|} : X \to \{-1; 1\}$ with $d\mu = \frac{d\mu}{d|\mu|}\, d|\mu|$ such that $M^+ := \left\{\frac{d\mu}{d|\mu|} = 1\right\}$ and $M^- := \left\{\frac{d\mu}{d|\mu|} = -1\right\}$ are measurable . On account of $\frac{1}{2}\left(1 + \frac{d\mu}{d|\mu|}\right) = \chi_{M^+}$ follows $\mu^+(A) = \frac{1}{2}\left(|\mu|(A) + \mu(A)\right) = \int_A \frac{1}{2}\left(1 + \frac{d\mu}{d|\mu|}\right) d|\mu| = \mu(A \cap M^+)$ resp. $\mu^-(A) = \mu(A \cap M^-)$.

10.13 Dual space of $L^p(\lambda)$: For every **σ-finite** and **positive** measure λ and $1 < p < \infty$ the **bounded linear functional** $M : L^p(\lambda) \to \mathbb{C}$ can be expressed **uniquely** as an **integral** $Mf = \int f \frac{d\mu}{d\lambda}\, d\lambda$ for $f \in L^p(\lambda)$ with the **Radon-Nikodym density** of the measure μ defined by $\mu(A) = M\chi_A$ with respect to λ. Furthermore we have $\frac{d\mu}{d\lambda} \in L^q(\lambda)$ for $\frac{1}{p} + \frac{1}{q} = 1$ and the **norm** $\|M\|^* = \sup\left\{\left|M\left(\frac{f}{\|f\|_p}\right)\right| : f \in L^p(\lambda)\right\}$ of the linear functional satisfies $\|M\|^* = \left\|\frac{d\mu}{d\lambda}\right\|_q$, i.e. the dual space $(L^p(\lambda))^*$ is **isometric** and hence **isomorphic** to $L^q(\lambda)$.

Proof: The λ-a.e. **uniqueness** of the representant $\frac{d\mu}{d\lambda} = g$ follows from the comparison of two possible candidates g and g' with $f_1 = \chi_{\{g<g'\}}$ resp. $f_2 = \chi_{\{g>g'\}}$ by means of $\int f_1 g'\, d\lambda = \int f_1 g\, d\lambda$ and $\int f_2 g'\, d\lambda = \int f_2 g\, d\lambda$ from 5.6.3.

Before we can use 10.8 we have to show that μ is a complex measure and absolutely continuous with respect to λ. Since we need the **continuity from above** 2.2.3 in this **first part** of the proof we have to restrict our reasoning to the case $\lambda(X) < \infty$. In a **second part** we will adapt the case $\lambda(X) = \infty$ to the first part making use of the σ-finiteness of λ:

For a sequence $(A_k)_{k\in\mathbb{N}} \subset \mathcal{A}$ of pairwise disjoint measurable sets with $B_n = \mathring{\bigcup}_{0\le k\le n} A_k$ and $B = \mathring{\bigcup}_{k\in\mathbb{N}} A_k$ the **continuity from above** 2.2.3 of the measure λ yields $\lim\limits_{n\to\infty} \|\chi_B - \chi_{B_n}\|_p = \lim\limits_{n\to\infty} \left\|\chi_{B\setminus B_n}\right\|_p$ $= \lim\limits_{n\to\infty} (\lambda(B \setminus B_n))^{\frac{1}{p}} = 0$ whence from the **continuity of the functional** M follows $\lim\limits_{n\to\infty} \mu(B_n) = \mu(B)$. Hence μ is **σ-additive** and thus a **complex measure**. For a λ-null set E we have $\|\chi_E\|_p = 0$ and since $M0 = 0$ the continuity of M implies $\mu(E) = 0$, i.e. $\mu \lessdot \lambda$. Hence 10.8 provides $\frac{d\mu}{d\lambda} \in L^1(\lambda)$ with $M\chi_A = \int \chi_A \frac{d\mu}{d\lambda}\, d\lambda$ for all $A \in \mathcal{A}$. The linearity of M guarantees $M\varphi = \int \varphi \frac{d\mu}{d\lambda}\, d\lambda$ for **step functions** $\varphi \in \mathcal{S}(X;\mathbb{C})$. According to 6.11 the **step functions** $\mathcal{S}(X;\mathbb{C})$ are dense in $L^p(\lambda)$ for every $1 \le p \le \infty$ and $\lambda(X) < \infty$. For now we apply only the case $p = \infty$, i.e. we extend the proposition to $f \in L^\infty(\lambda)$: On the left hand side a λ-a.e. bounded $f \in L^\infty(\lambda)$ is a limit of a uniformly convergent sequence $(\varphi_n)_{n\in\mathbb{N}} \subset \mathcal{S}(X)$ converging also in the p-th mean on account of $\|f\|_p \le \|f\|_\infty \cdot (\lambda(X))^{\frac{1}{p}}$ whence follows the convergence of $(M\varphi_n)_{n\in\mathbb{N}}$. On the right hand side the uniform convergence directly entails the convergence of the intgeral on due to $\left|\int f \frac{d\mu}{d\lambda}\, d\lambda\right| \le \|f\|_\infty \cdot \|g\|_1$. In order to extend the validity of the proposition to $f \in L^p(\lambda)$ we show that $g := \frac{d\mu}{d\lambda} \in L^q(\lambda)$: Let $E_n = \{|g| \ge n\}$ for $n \in \mathbb{N}$ and $f = \frac{|g|^q}{g} \cdot \chi_{E_n} \in L^\infty(\lambda)$ for $n \in \mathbb{N}$ such that $|f|^p \cdot \chi_E = |g|^{(q-1)p} \cdot \chi_E = |g|^q \cdot \chi_E = fg$. Hence we have $\int_{E_n} |g|^q\, d\lambda = \int fg\, d\lambda = \Lambda(f) \le \|\Lambda\|^* \cdot \|f\|_p = \|\Lambda\|^* \cdot \left(\int_{E_n} |g|^q\, d\lambda\right)^{\frac{1}{p}} \Leftrightarrow \left(\int_{E_n} |g|^q\, d\lambda\right)^{1-\frac{1}{p}} \le \|\Lambda\|^* \Leftrightarrow$ $\int_{E_n} |g|^q\, d\lambda \le \|\Lambda\|^{*q}$ such that with **monotone convergence** 5.12 we obtain $\|g\|_q \le \|\Lambda\|^* < \infty$ and in particular $g = \frac{d\mu}{d\lambda} \in L^q(\lambda)$. The **Hölder inequality** 6.4.1 combined with $\left\|\frac{d\mu}{d\lambda}\right\|_q < \infty$ asserts the continuity of the mapping $f \mapsto \int f \frac{d\mu}{d\lambda}\, d\lambda$ on $L^p(\lambda)$ and since it coincides on the dense subset $\mathcal{E}(X) \subset L^p(\lambda)$ with the continuous mapping M the assertion follows for $\lambda(X) < \infty$. Another look at **Hölder** yields $\|M\|^* \le \left\|\frac{d\mu}{d\lambda}\right\|_q$ and hence the second assertion $\|M\|^* = \left\|\frac{d\mu}{d\lambda}\right\|_q$.

In the case of $\lambda(X) = \infty$ as in the proof of 10.8 we define $w = \sum_{n\in\mathbb{N}} \frac{\chi_{A_n}}{2^n\cdot(1+\lambda(A_n))} : X \to]0;1[$ for a countable cover $(A_n)_{n\in\mathbb{N}} \subset \mathcal{A}$ of X with $\lambda(A_n) < \infty\, \forall n \in \mathbb{N}$ such that the measure ν with $\nu(A) := \int_A w\, d\lambda$ is **finite** and on account of $w > 0$ has the same null sets as λ. Then the bijection $\omega_p : L^p(\lambda) \to L^p(\nu)$ with $\omega_p(f) = w^{-\frac{1}{p}} \cdot f$ is a **linear isometry** and $M \circ \omega_p^{-1} : L^p(\nu) \to \mathbb{C}$ is a bounded linear functional with $\left\|M \circ \omega_p^{-1}\right\|^* = \sup\left\{\left|M\left(\frac{w^{\frac{1}{p}}\cdot\omega_p(f)}{\left(\int|\omega(f)|_p^p\cdot w d\lambda\right)^{\frac{1}{p}}}\right)\right| : \omega_p(f) \in L^p(\nu)\right\} =$ $\sup\left\{\left|M\left(\frac{f}{\left(\int|f|^p d\lambda\right)^{\frac{1}{p}}}\right)\right| : f \in L^p(\lambda)\right\} = \|M\|^*$. According to the **first part** of the proof there is an $\omega_q\left(\frac{d\mu}{d\lambda}\right) \in L^q(\nu)$ with $\left(M \circ \omega_p^{-1}\right)(\omega_p(f)) = \int \omega_p(f) \cdot \omega_q\left(\frac{d\mu}{d\lambda}\right) w d\lambda$ for all $\omega_p(f) \in L^p(\nu)$ resp. $Mf = \int fg\, d\lambda$ for all $f \in L^p(\lambda)$.

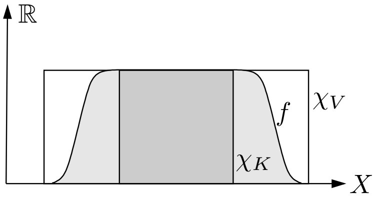
10.14 Note: The special case of the **Hilbert space** with $p = q = 2$ is the central pivot in the proof of the **Lebesgue-Radon-Nikodym theorem** 10.8 where [8, p 308 Th 12.5] is used to find a uniquely determined representant $g \in L^2(\varphi)$ with $Mf = \langle f, g\rangle = \int fg\, d\varphi$ for the bounded functional $M \in (L^2(\varphi))^*$ with $Mf = \int f\, d|\lambda|$. Alas the **isometry** of the two spaces is **not** an issue in this proof.

11 Measures on locally compact spaces

11.1 Linear functionals and measures on locally compact spaces: In this section we examine the **dual space** $(C_c(X,\mathbb{C}))^*$ of the **complex linear functionals** $\Lambda : C_c(X,\mathbb{C}) \to \mathbb{C}$ on the **Banach space** $C_c(X,\mathbb{C})$ of **complex continuous functions** $f : X \to \mathbb{C}$ **with compact support** under the **supremum norm** $\|\|$ on a **locally compact** space X furnished with the **Borel** σ-algebra $\mathcal{B}(X) = \sigma(\mathcal{O})$ induced by its topology $\mathcal{O}$. **Hence in this section and without further mentioning X will always be locally compact**. The $\mathbb{C}$-linearity of a complex functional Λ

implies $\Lambda(\mathrm{Re}f + i\mathrm{Im}f) = \Lambda\mathrm{Re}f + i\Lambda\mathrm{Im}f$ such that it suffices to examine **complex linear functionals** $\Lambda : C_c(X,\mathbb{R}) \to \mathbb{C}$ with **real valued** arguments as e.g. in the case of $\Lambda f = \mathrm{Re}\Lambda f + i\mathrm{Im}\Lambda f = \int f d(\mathrm{Re}\mu) + i\int f d(\mathrm{Im}\mu) = \int f d\mu$ with a **complex measure** $\mu = \mathrm{Re}\mu + i\mathrm{Im}\mu$ according to 10.11. A **complex linear functional** $\Lambda : C_c(X,\mathbb{C}) \to \mathbb{C}$ is **positive** iff for **positive** f the value Λf also is **positive**, the directly available example being the integral $\Lambda f = \int f d\lambda$ with a **positive** measure λ. In fact these examples already cover the range of possibilities: Our aim is to prove the **Riesz representation theorem** which states that the **Banach space** $(C_c(X,\mathbb{C}))^*$ under the **norm** $\|\|^*$ with $\|\Lambda\|^* = \sup\left\{\left|\Lambda\left(\frac{f}{\|f\|_\infty}\right)\right| : f \in C_c(X,\mathbb{C})\right\} = \sup\left\{\left|\Lambda\left(\frac{f}{\|f\|_\infty}\right)\right| : f \in C_c(X,\mathbb{R})\right\}$ (cf. 10.13 and consider $f \in C_c(X,\mathbb{C}) \Rightarrow |f| \in C_c(X,\mathbb{R})$) is **isometric** and **isomorphic** to the Banach space $M_0(\mathcal{B}(X);\mathbb{C})$ of **complex regular Borel measures** on X under the **norm** $\|\|$ with $\|\mu\| = |\mu|(X)$ (cf. 10.1). A **positive measure** μ on a Borel σ-algebra is a **Borel measure** iff every **compact** set K has a **finite** measure $\mu(K) < \infty$. It is **outer regular** iff $\mu(A) = \inf\{\mu(O) : O \text{ open with } A \subset O\}$ and **inner regular** iff $\mu(A) = \sup\{\mu(K) : K \text{ compact with } K \subset A\}$ respectively for every **measurable** $A \in \mathcal{B}(X)$. It is **regular** iff both conditions hold for every **measurable** set A and **σ-regular** if the latter condition holds for measurable sets which are either **open** or **σ-finite**. A set is **σ-finite** iff it is a countable union of sets with finite measure. Hence a **σ-regular** measure on a **σ-finite** space X is already **regular**. A **complex Borel measure** is regular, iff its **variation** $|\mu|$ is regular.

The following two results show the close relationship between **continuous** and **measurable** functions on these spaces. Again we introduce two notations for approximative behaviour: For **real** $f \in C_c(X,\mathbb{R})$, **open** $V \subset X$ and **compact** $K \subset X$ we write $K \prec f$ iff $\chi_K \leq f \leq 1$ and $f \prec V$ iff $0 \leq f \leq \chi_V$. In these terms the **separation property** [9, 10.5] of locally compact spaces simply states that for every **compact** K and **open** $V \supset K$ there is an $f \in C_c(X,\mathbb{R})$ with $K \prec f \prec V$.



Since in a locally compact space the compact neighbourhoods form a **neighbourhood basis** we can strengthen this proposition to $\chi_K = \sup\limits_{f \prec V} f$.

Examples:

1. The **Dirac measure** $\epsilon_x(A) = \chi_A(x)$ for any point $x \in X$ of a **Hausdorff** space X and a **Borel** set $A \in \mathcal{B}(X)$ is **regular**.
2. The measure $\mu(A) := \begin{cases} 0 & \text{for } A \text{ countable} \\ \infty & \text{else} \end{cases}$ defined in 2.3.2 on the σ-algebra $\mathcal{B}(X) = \sigma(\mathcal{O}) = \mathcal{O} = \mathcal{P}(X)$ of a **discrete** space X is a **locally finite** and **outer regular Borel** measure. It is **inner regular** iff X is **countable**.
3. The **Lebesgue measure** $\lambda^n := \bigotimes_{1 \leq i \leq n} \lambda$ on the Borel σ-algebra $\mathcal{B}^n$ of $\mathbb{R}^n$ is a **σ-finite Borel measure** owing to 7.7 resp. the **Heine-Borel theorem** [9, 9.10]. Its **regularity** is a consequence of the **locally compact** character of $\mathbb{R}^n$ and follows from the **Riesz representation theorem** 11.10 applied to the positive functional Λ with $\Lambda f = \int f d\lambda^n$ for $f \in C_c(\mathbb{R}^n,\mathbb{R})$.

11.2 Theorem: For a **positive σ-regular Borel** measure λ and $1 \leq p < \infty$ the space $C_c(X,\mathbb{C})$ is **dense** in $L^p(\lambda)$.

Proof: According to 6.11.1 it suffices to find for every measurable set A with $\lambda(A) < \infty$ a function $g \in C_c(X,\mathbb{R})$ such that $\|\chi_A - g\|_p = \|i\chi_A - ig\|_p < \epsilon$. Since λ is σ-regular and $\lambda(A) < \infty$ there is a compact K and an open V with $K \subset A \subset V$ and $\lambda(K) < \lambda(V) + \epsilon$ as well as a $g \in C_c(X,\mathbb{R})$ with $K \prec g \prec V$ such that $\lambda(K) \leq \int g d\lambda \leq \lambda(V)$ whence $\|\chi_A - g\|_p \leq \|\chi_A - \chi_K\|_p + \|\chi_K - g\|_p < \epsilon^{1/p} + \epsilon^{1/p}$.

11.3 Lusin's Theorem: For every **complex** measurable function f with $\lambda(f \neq 0) < \infty$ with a **positive σ-regular Borel** measure λ and every $\epsilon > 0$ there exists a $g \in C_c(X,\mathbb{C})$ such that $\lambda(f \neq g) < \epsilon$ and $\|g\| \leq \|f\|$.

Proof: Due to $\bigcap_{n\geq 1}\{|f|\geq n\}=\emptyset$ and the continuity of λ from above there is an $n_\epsilon\in\mathbb{N}$ with $\lambda(A_1)<\frac{\epsilon}{4}$ for $A_1=\{|f|\geq n_\epsilon\}$ and hence $f\in L^1(\lambda')$ with $\lambda'=\lambda|_{X\setminus A_1}$. According to 11.2 there is a sequence $(f_n)_{n\in\mathbb{N}}\subset C_c(X\setminus A_1;\mathbb{C})$ converging **in mean** to f and according to 5.10 we have a subsequence **uniformly** converging on $X\setminus(A_1\cup A_2)$ with $\lambda(A_2)<\frac{\epsilon}{4}$ to f and consequently $f\in C(X\setminus(A_1\cup A_2);\mathbb{C})$. By the σ-regularity we find a **compact** $K\subset\{f\neq 0\}\setminus(A_1\cup A_2)$ with $\lambda(A_3)<\frac{\epsilon}{4}$ for $A_3=\{f\neq 0\}\setminus(K\cup A_1\cup A_2)$ and $f\in C(K,\mathbb{C})$. Since in a locally compact space the compact neighbourhoods form a **neighbourhood basis** we find an **open** set $V\subset K$ with **compact closure** $\overline{V}$ which due to the **outer regularity** of λ we can choose such that w.l.o.g. $\lambda(A_4)<\frac{\epsilon}{4}$ for $A_4=V\setminus K$. The compact set $\overline{V}$ is also **normal** such that we can apply **Tietze's extension theorem** [9, 8.5] to find Reg^* resp. Im$g^*\in C\left(\overline{V},\mathbb{R}\right)$ coinciding with Ref resp. Imf on K and vanishing on the closed boundary $\overline{V}\setminus V$. Extending $g^*=\mathrm{Re}g^*+i\mathrm{Im}g^*$ to X by assigning the value 0 outside $\overline{V}$ we obtain a $g\in C_c(X,\mathbb{C})$ coinciding with f on $X\setminus A_\epsilon\subset K\cup X\setminus(A_1\cup A_2\cup\{f\neq 0\}\cup V)$ with $A_\epsilon=A_1\cup A_2\cup A_3\cup A_4$ and $\lambda(A_\epsilon)<\epsilon$. In order to **scale** g according to $\|g\|\leq\|f\|$ we define a continuous $h:\mathbb{C}\to\mathbb{C}$ by $h(z)=\begin{cases} z & \text{if } |z|\leq\|f\| \\ \|f\|\cdot\frac{z}{|z|} & \text{if } |z|>\|f\|\end{cases}$ such that $\|h\circ g\|\leq\|f\|$.

11.4 Lemma: Every **positive** functional $\Lambda: C_c(X,\mathbb{C})\to\mathbb{C}$ is **bounded** on $C_K(X,\mathbb{C})$ for every **compact** K.

Proof: Due to the **separation property** [9, 10.5] of locally compact spaces already cited in 11.1 there is a continuous $g:X\to[0;1]$ with $g^{-1}(\{1\})=K$ and compact support. Then for $f\in C_K(X,\mathbb{C})$ we have $\|\mathrm{Re}f\|\cdot g\pm\mathrm{Re}f\geq 0$ whence $\Lambda(\|\mathrm{Re}f\|\cdot g)\pm\Lambda(\mathrm{Re}f)=\|\mathrm{Re}f\|\cdot\Lambda g\pm\Lambda(\mathrm{Re}f)\geq 0$, i.e. $\Lambda\left(\frac{\mathrm{Re}f}{\|f\|}\right)\leq\Lambda\left(\frac{\mathrm{Re}f}{\|\mathrm{Re}f\|}\right)\leq\Lambda g$ and since the same is true for Imf we obtain $\left|\Lambda\left(\frac{f}{\|f\|}\right)\right|=\left|\Lambda\left(\frac{\mathrm{Re}f}{\|f\|}\right)+i\Lambda\left(\frac{\mathrm{Im}f}{\|f\|}\right)\right|\leq\sqrt{2}\cdot\Lambda g<\infty$.

11.5 Theorem: Every **bounded real** functional $\Lambda\in(C_c(X,\mathbb{R}))^*$ has a **decomposition** $\Lambda=\Lambda^+-\Lambda^-$ with **positive real and bounded** $\Lambda^+;\Lambda^-\in(C_c(X,\mathbb{R}))^*$.

Proof: For **positive** $f\in C_c(X,\mathbb{R})$ define $\Lambda^+f:=\sup\{\Lambda g: g\in C_c(X,\mathbb{R});0\leq g\leq f\}$ such that $0\leq\Lambda^+f\leq\|\Lambda\|^*\|f\|$, i.e. $\Lambda^+$is **positive and bounded**. For **positive** c$\in\mathbb{R}$ we have $g\leq cf\Leftrightarrow g=cg':g'\leq f$ for any positive $g;g'\in C_c(X,\mathbb{R})$ such that $\Lambda^+(cf)=c\Lambda^+f$ thus establishing conformity with **scalar muplitplication**. With regard to additivity we take any positive $f_1;f_2;g_1;g_2;g\in C_c(X,\mathbb{R})$ with $g_1\leq f_1$, $g_2\leq f_2$ resp. $g\leq f_1+f_2$ in order to note that $\Lambda^+f_1+\Lambda^+f_2=\sup\Lambda^+g_1+\sup\Lambda^+g_2=\sup(\Lambda^+g_1+\Lambda^+g_2)=\sup\Lambda^+(g_1+g_2)\leq\sup\Lambda^+g=\Lambda^+(f_1+f_2)$ and conversely $\inf(g;f_1)\leq f_1$ resp. $g-\inf(g;f_1)\leq f_2$ hence $\Lambda^+g\leq\Lambda^+f_1+\Lambda^+f_2$,i.e. $\Lambda^+(f_1+f_2)=\sup\Lambda^+g\leq\Lambda^+f_1+\Lambda^+f_2$ thus demonstrating **additivity**. We extend Λ^+ to **real** $f\in C_c(X,\mathbb{R})$ with decomposition $f=f^+-f^-$with **positive** $f^+;f^-\in C_c(X,\mathbb{R})$ by means of $\Lambda^+f:=\Lambda^+f^+-\Lambda^+f^-$ being independent of the choice of the decomposition and hence **well defined** as well as **linear** on account of the linearity of the components. The same is true for $\Lambda^-:=\Lambda-\Lambda^+$which completes the proof.

11.6 Corollary: Every **complex** functional $\Lambda\in(C_c(X,\mathbb{R}))^*$ allows the decomposition into four **positive bounded** functionals ReΛ^+;ReΛ^-;ImΛ^+;Im$\Lambda^-\in(C_c(X,\mathbb{R}))^*$ such that $\Lambda f=\mathrm{Re}\Lambda^+f-\mathrm{Re}\Lambda^-f+i(\mathrm{Im}\Lambda^+f+\mathrm{Im}\Lambda^-f)$.

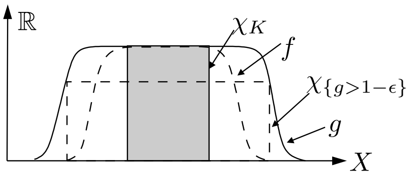
11.7 Lemma: For every **positive** functional $\Lambda: C_c(X,\mathbb{C})\to\mathbb{C}$ the set function $\mu: P(X)\to[0;\infty]$ defined by $\mu(V)=\sup\{\Lambda g: g\prec V\}$ for **open** V and $\mu(A)=\inf\{\mu(V):A\subset V\text{open}\}$ is an **outer measure** according to 3.2 with the **additional regularity property** $\mu(K)\leq\Lambda g\leq\mu(V)$ for any **compact** K and **open** V with $K\prec g\prec V$.

Proof: Obviously we have $\mu(\emptyset)=0$ and $\mu(A)\leq\mu(B)$ if $A\subset B$. The **subadditivity** requires more attention. We start with $\mu(U\cup V)\leq\mu(U)+\mu(V)$ for **open** U and V: Let $f\prec U\cup V$ and $\Phi=\{\sup(g;h):g\prec U;h\prec V\}$ and $\Phi_f=\left\{\inf\left(f;\overline{f}\right):\overline{f}\in\Phi\right\}$. Then $f=\sup\Phi_f\leq\sup\Phi=\chi_{U\cup V}$ such that on account of **Dini's theorem** [9, 9.12] and the **continuity** of Λ we have

$$\begin{aligned}\Lambda f &= \Lambda \sup \mathit{\Phi}_f \\ &= \sup \Lambda \mathit{\Phi}_f \\ &= \sup \left\{\Lambda \left(\inf \left(f; \sup \left(g; h\right)\right)\right) : g \prec U; h \prec V\right\} \\ &\leq \sup \left\{\Lambda \left(\inf \left(f; g\right) + \inf \left(f; h\right)\right) : g \prec U; h \prec V\right\} \\ &\leq \sup \left\{\Lambda \left(g + h\right) : g \prec U; h \prec V\right\} \leq \mu \left(U\right) + \mu \left(V\right).\end{aligned}$$

Since this estimate holds for every $f \prec U \cup V$ we obtain the subadditivity for open sets. In order to show the σ-subadditivity 3.2.3 we take a sequence $\left(A_n\right)_{n\in\mathbb{N}}$ of arbitrary subsets with $A = \bigcup_{n\in\mathbb{N}} A_n$, open sets V_n with $A_n \subset V_n$ and $\mu\left(A_n\right) \leq \mu\left(V_n\right) + \epsilon 2^{-n}$ such that $A \subset V = \bigcup_{n\in\mathbb{N}} V_n$. Since any $g \prec V$ has a compact support there is an $n \in \mathbb{N}$ with $g \prec \bigcup_{k\leq n} V$ and hence $\Lambda g \leq \mu\left(\bigcup_{k\leq n} V_n\right) \leq \sum_{k\leq n} \mu\left(V_n\right)$ due to the subadditivity inductively extended to finite unions. Again we use the validity of this estimate for every $g \prec V$ to infer $\mu\left(A\right) \leq \mu\left(V\right) \leq \sum_{n\in\mathbb{N}} \mu\left(A_n\right) + \epsilon$ thus proving the main assertion.

Concerning the **additional regularity property** we only have to show the left inequality: For any $\epsilon > 0$ we have $K \subset \left\{g > 1 - \epsilon\right\}$ and hence a $f \in C_c\left(X\right)$ with on the one hand $K \prec f \prec \left\{g > 1 - \epsilon\right\}$ such that $\left(1 - \epsilon\right) f \leq g$, i.e. $\left(1 - \epsilon\right) \Lambda f \leq \Lambda g$ and on the other hand $\Lambda f \geq \mu\left(\left\{g > 1 - \epsilon\right\}\right) - \epsilon \geq \mu\left(K\right) - \epsilon$ whence $\left(\mu\left(K\right) - \epsilon\right)\left(1 - \epsilon\right) \leq \Lambda g$ which proves the assertion.



11.8 Lemma: The **outer measure** μ determined by Λ according to the preceding lemma 11.7 is **σ-additive** and hence a **pre-measure** on the **algebra** $\mathcal{A}\left(X\right)$ of all sets A with **finite measure** and $\mu\left(A\right) = \sup\left\{\mu\left(K\right) : A \supset K \text{ compact}\right\}$. Furthermore $\mathcal{A}\left(X\right)$contains all open sets.

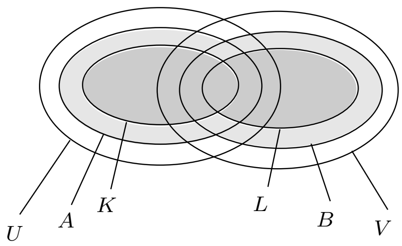
Proof: For brevity in this proof we omit the argument and write $\mathcal{A}$ for $\mathcal{A}\left(X\right)$.

Step I. Every **compact** set K has a **finite measure** and hence belongs to $\mathcal{A}$: There is an open $V \subset K$ with compact closure $\overline{V}$ such that the separation property of locally compact spaces ensures the existence of $f, g \in C_c\left(X\right)$ with $K \prec f \prec V$ resp. $\overline{V} \prec g \prec X$ hence $g - f \geq 0 \Rightarrow \Lambda\left(g - f\right) \geq 0 \Rightarrow \Lambda f \leq \Lambda g < \infty$ due to the positiv and linear character of Λ. Furthermore we can choose f such that $\mu\left(V\right) \leq \Lambda f + \epsilon$ whence $\mu\left(K\right) \leq \mu\left(V\right) \leq \Lambda f + \epsilon \leq \Lambda g + \epsilon < \infty$.

Step II. $\mathcal{A}$ contains every **open** set V: In the case of $\mu\left(V\right) = 0$ the definition of μ immediately yields $\mu\left(K\right) = \inf\left\{\mu\left(V\right) : K \subset V \text{open}\right\} = 0$ for every compact $K \subset V$. Hence we can assume $\mu\left(V\right) > 0$ and for every $\epsilon > 0$ the existence of an $f \prec V$ with $\mu\left(V\right) - \epsilon < \Lambda f < \mu\left(V\right)$ and compact support $K = \overline{\left\{f > 0\right\}}$. For every open $W \supset K$ we have $f \prec W$ and hence $\Lambda f \leq \mu\left(K\right)$ and consequently $\mu\left(V\right) - \epsilon < \Lambda f \leq \mu\left(K\right) < \mu\left(V\right) < \infty$ on account of $K \subset V$ and 11.4.

Step III. μ is **finitely additive** for **compact** sets: For **disjoint** and **compact** sets K, L and $\epsilon > 0$ according to the **separation property** [9, 10.5] of locally compact spaces choose **disjoint** and **open** $U \supset K$, $V \supset L$ and an open $W \supset K \cup L$ with $\mu\left(W\right) < \mu\left(K \cup L\right) + \epsilon$ as well as $f \prec U \cap W$ resp. $g \prec V \cap W$ with $\Lambda f > \mu\left(U \cap W\right) - \epsilon$ resp. $\Lambda g > \mu\left(V \cap W\right) - \epsilon$. We then have $\mu\left(K\right) + \mu\left(L\right) \leq \mu\left(W \cap U\right) + \mu\left(W \cap V\right) \leq \Lambda f + \Lambda g + 2\epsilon = \Lambda\left(f + g\right) + 2\epsilon \leq \mu\left(W\right) + 2\epsilon \leq \mu\left(K \cup L\right) + 3\epsilon$. Since the reverse inequality follows from the montonicity of μ we have proved the asssertion.

Step IV. μ is **σ-additive** on $\mathcal{A}$: For a sequence $\left(A_n\right)_{n\in\mathbb{N}} \subset \mathcal{A}$ with $A = \bigcup_{n\in\mathbb{N}} A_n$ there are compact $K_n \subset A_n$ with $\mu\left(A_n\right) \leq \mu\left(K_n\right) + \epsilon 2^{-n}$ whence $\sum\limits_{k=1}^{n} \mu\left(A_k\right) \leq \sum\limits_{k=1}^{n} \mu\left(K_k\right) + \epsilon = \mu\left(\bigcup\limits_{k=1}^{n} K_k\right) + \epsilon \leq \mu\left(A\right) + \epsilon$. Since this estimate remains valid for $n \to \infty$ and $\epsilon \to 0$ we obtain $\sum_{n\in\mathbb{N}} \mu\left(A_n\right) \leq \mu\left(A\right)$ and with the reverse inequality following from property 3.2.3 of the outer measure we have proved the assertion. Furthermore we note that for every sequence $\left(A_n\right)_{n\in\mathbb{N}} \subset \mathcal{A}$ the union $A = \bigcup_{n\in\mathbb{N}} A_n$ also belongs to $\mathcal{A}$ if it has finite measure, i.e. for **finite** μ the algebra $\mathcal{A}$ is a **σ-algebra**. This will be used of in the subsequent lemma to construct the actual σ-algebra $\mathcal{M}$ carrying the measure μ determined by Λ.



Step V. $\mathcal{A}$ is an **algebra**: Clearly $\emptyset \in \mathcal{A}$. For $A, B \in \mathcal{A}$ we can find compact K, L and open U, V such that $K \subset A \subset V$ resp. $L \subset B \subset V$ and $\mu(K) \leq \mu(A) \leq \mu(U) < \mu(K) + \epsilon$ resp. $\mu(L) \leq \mu(B) \leq \mu(V) < \mu(L) + \epsilon$. By the finite additivity of μ follows $\mu(U \setminus K), \mu(V \setminus L) < \epsilon$ and with $(U \cup V) \setminus (K \cup L) \subset (U \setminus K) \cup (V \setminus L)$ we get $\mu(A \cup B) < \mu(K \cup L) + 2\epsilon$ and hence $A \cup B \in \mathcal{A}$. Regarding the intersection we note that $K \setminus V \subset A \setminus B \subset V \setminus L$ and the two outer sets are **open** with $(V \setminus L) \setminus (K \setminus V) \subset (U \setminus K) \cup (V \setminus L)$ so that $A \setminus B \in \mathcal{A}$ and finally $A \cap B = B \setminus (B \setminus A) \in \mathcal{A}$.

11.9 Lemma: The outer measure μ determined by Λ according to lemma 11.7 is **σ-additive** and hence a **measure** on the σ-algebra $\mathcal{L}(X) = \bigcap_{K \text{ compact}} \mathcal{L}_K(X)$ with $\mathcal{L}_K(X) = \{A \subset X : A \cap K \in \mathcal{A}(X)\}$ including the **Borel σ-algebra** $\mathcal{B}(X)$ as well as the **algebra** $\mathcal{A}(X)$ of sets of finite measure introduced in the preceding lemma 11.8. $\mathcal{A}(X)$ consists precisely of all sets of **finite measure** in $\mathcal{L}(X)$. In particular μ is **complete and σ-regular** on $\mathcal{L}(X)$.

Proof: Again we abbreviate $\mathcal{A} = \mathcal{A}(X)$ etc. Obviously we have $\mathcal{A} \subset \mathcal{L}$. According to the **step IV** of the proof of the preceding lemma the families $\mathcal{L}_K$ are σ-algebrae and so is $\mathcal{L}$. Every $\mathcal{L}_K$ contains all **closed** sets (cf. [9, 9.4] and hence $\mathcal{B}(X) \subset \mathcal{L}$. Every **$\mu$-null set** $A \subset X$ with $\mu(A) = 0$ is either empty or contains a point $x \in A \subset X$ and hence a compact set $\{x\} \subset A$ which must have the measure $\mu(\{x\}) = 0$ due to the **monotonicity** of μ. Hence $A \in \mathcal{L}$ and in particular μ is **complete.**

For $A \in \mathcal{L}$ with $\mu(A) < \infty$ there is an **open** $V \supset A$ with $\mu(V) < \infty$. Furthermore according to **step II** in the proof of 11.8 we can find a **compact** $K \subset V$ such that $\mu(V) < \mu(K) + \epsilon$. Since $A \cap K \in \mathcal{A}$ there is a **compact** $K_A \subset A \cap K$ such that $\mu(A \cap K) < \mu(K_A) + \epsilon$. With $A \subset (A \cap K) \cup V \setminus K$ we obtain $\mu(A) \leq \mu(A \cap K) + \mu(V \setminus K) \leq \mu(K_A) + 2\epsilon$ and since ϵ was arbitrary we have $\mu(A) = \sup\{\mu(K) : A \supset K \text{ compact}\}$ whence follows $A \in \mathcal{A}$. Finally the **σ-additivity** of μ extends from $\mathcal{A}$ to $\mathcal{L}$ since for a disjoint sequence $(A_n)_{n\in\mathbb{N}} \subset \mathcal{L}$ with $\mu(A_n) < \infty$ for all $n \in \mathbb{N}$ we have $(A_n)_{n\in\mathbb{N}} \subset \mathcal{A}$ such that the preceding lemma applies. In the case of $\mu(A_n) = \infty$ for an $n \in \mathbb{N}$ the σ-additivity follows from the **monotonicity** of μ. Due to its definition in 11.7 μ is **outer regular** on $\mathcal{L}$. According to 11.8 it is **inner regular** for all sets **open** or with **finite measure**. For $\epsilon > 0$ and a **σ-finite** set $A = \bigcup_{n\in\mathbb{N}} A_n$ with $\mu(A_n) < \infty$ and w.l.o.g $A_n \subset A_{n+1}$ for $n \in \mathbb{N}$ we find compact $K_n \subset A_n$ with $\mu(K_n) \geq \mu(A_n) - \frac{\epsilon}{2}$ for $n \in \mathbb{N}$. In the case of $\mu(A) < \infty$ there is an $m \in \mathbb{N}$ with $\mu(A_m) \geq \mu(A) - \frac{\epsilon}{2}$ and hence $\mu(K_m) \geq \mu(A) - \epsilon$. In the case of $\mu(A) < \infty$ for every $N \in \mathbb{N}$ there is an $m \in \mathbb{N}$ with $\mu(A_m) \geq N + \frac{\epsilon}{2}$ and hence $\mu(K_m) \geq N$. Hence we have shown that $\mu(A) = \sup\{\mu(K) : K \text{ compact with } K \subset A\}$.

11.10 Riesz representation theorem for positive functionals: Every **positive** functional $\Lambda \in (C_c(X, \mathbb{C}))_+^*$ can be represented as an **integral** $\Lambda = I_\mu$ with $I_\mu f = \int f d\mu$ for $f \in C_c(X, \mathbb{C})$ with respect to the **complete** and **σ-regular positive Borel** measure μ on a σ-algebra $\mathcal{L}(X)$ including the **Borel σ-algebra** $\mathcal{B}(X) \subset \mathcal{L}(X)$.

Proof:

Uniqueness: Assuming that the integrals I_{μ_1} and I_{μ_2} of two σ-regular and complete positive Borel measures μ_1 and μ_2 coincide on the set of continuous complex functions with compact support, i.e. $\int f d\mu_1 = \int f d\mu_2$ for all $f \in C_c(X, \mathbb{C})$, let $\epsilon > 0$, K be compact, $V \supset K$ open with $\mu_2(V) < \mu_2(K) + \epsilon$ and $f \in C_c(X, \mathbb{R})$ with $K \prec f \prec V$, i.e. $\chi_K \leq f \leq \chi_V$ whence $\mu_1(K) \leq \int f d\mu_1 = \int f d\mu_2 \leq \mu_2(V) \leq \mu_2(K) + \epsilon$. and vice versa. Hence the two measures coincide on the **compact** sets and due to their regularity this identity extends first to the **open** sets and finally to all **measurable** sets.

Existence: Since the $f \in C_c(X, \mathbb{C})$ are **continuous** and in particular **Borel measurable** we can restrict the measure μ determined by Λ according to lemma 11.7 on the σ-algebra $\mathcal{L}(X)$ from 11.9 to the **Borel σ-algebra** $\mathcal{B}(X) \subset \mathcal{L}(X)$. On account of $\text{Re}\Lambda f = \Lambda\text{Re}f$ resp. $\text{Im}\Lambda f = \Lambda\text{Im}f$ for positive functionals it suffices to show the equation for real f. Since $f \in C_c(X) \Leftrightarrow -f \in C_c(X)$ we only have to show $\Lambda f \leq \int f d\mu$ for **every** $f \in C_c(X, \mathbb{R})$. Since the **step functions** defining the integral are **not continuous** we have to take recourse to a corresponding **partition of unity** consisting of

continuous functions of compact support being amenable to Λ and providing a result which can be compared to the integral. Furthermore the general case only provides for **pointwise convergence** so that we need the **compactness** of the support $K = \overline{\{f \neq 0\}}$ in order to find elementary functions **uniformly converging** to f: For $\epsilon > 0$ let $A_k = \{k\epsilon \leq f < (k+1)\epsilon\}$ with $-n \leq k \leq n = \left[\frac{\|f\|}{\epsilon}\right]$ such that $(A_k)_{|k|\leq n}$ is a partition of the compact support $K = \bigcup\limits_{k=-n}^{n} A_k$ and $e = \sum\limits_{k=-n}^{n} k\epsilon\chi_{A_k} \in \mathcal{S}(X)$ according to 5.2 and 5.4 such that $e \leq f \leq e + \epsilon$ whence $\int ed\mu\leq \int fd\mu \leq \int ed\mu + \epsilon \cdot \mu(K)$. Due to 11.7 for every $|k| \leq n$ there is an open V_k with $A_k \subset V_k \subset \{f < e + \epsilon\}$ and $\mu(V_k) \leq \mu(A_k) + \frac{\epsilon}{n\|f\|}$. On account of [9, 8.9, 9.5 and 10.5] we can find a partition of unity $(h_k)_{|k|\leq n} \subset C_c(X, \mathbb{R})$ subordinate to $(V_k)_{|k|\leq n}$ with $fh_k \prec V_k$ and $fh_k \leq (k+1)\epsilon h_k$ as well as $K \prec \sum\limits_{k=-n}^{n} h_k$ such that $\mu(K) \leq \sum\limits_{k=-n}^{n} \Lambda h_k$.

Thus we have

$$
\begin{aligned}
\Lambda f &= \sum_{k=-n}^{n} \Lambda f h_k \\
&\leq \sum_{k=-n}^{n} (k+1)\epsilon\Lambda h_k \\
&= \sum_{k=-n}^{n} (k\epsilon + \epsilon + \|f\|)\Lambda h_k - \|f\| \sum_{k=-n}^{n} \Lambda h_k \\
&\leq \sum_{k=-n}^{n} (k\epsilon + \epsilon + \|f\|)\mu(V_k) - \|f\|\mu(K) \\
&\leq \sum_{k=-n}^{n} (k\epsilon + \epsilon + \|f\|)\left(\mu(A_k) + \frac{\epsilon}{n\|f\|}\right) - \|f\|\mu(K) \\
&\leq \int ed\mu + \epsilon\mu(K) + \|f\|\mu(K) + 2(n(n+1)\epsilon + 2n\|f\|)\frac{\epsilon}{n\|f\|} - \|f\|\mu(K) \\
&= \int ed\mu + \epsilon\mu(K) + \frac{2(n+1)\epsilon^2}{\|f\|} + 2\epsilon \\
&\leq \int fd\mu + \epsilon\mu(K) + 6\epsilon.
\end{aligned}
$$

11.11 Riesz representation theorem for complex functionals: There is an **isometric isomorphism** I: $\mathcal{M}_0^*(\mathcal{L}(X);\mathbb{C}) \to (C_c(X,\mathbb{C}))^*$ with $I : \mu \to I_\mu$ defined by $I_\mu f = \int fd\mu = \int f\frac{d\mu}{d|\mu|}d|\mu|$ (cf. 10.9) between the **Banach space** of the **complete and regular complex Borel measures** on X under the **norm** $\|\|$ with $\|\mu\| := |\mu|(X)$ (cf. 10.4) and the **Banach space** $(C_c(X,\mathbb{C}))^*$ under the **norm** $\|\|^*$ with $\|\Lambda\|^* = \sup\left\{\left|\Lambda\left(\frac{f}{\|f\|}\right)\right| : f \in C_c(X,\mathbb{R})\right\}$ (cf. 10.13) .

Proof:

$\mathcal{M}_0^*(\mathcal{L}(X);\mathbb{C})$ is a **Banach** space since it is a **closed** vector subspace (cf. [9, 20.6.6]) of the **Banach** space $\mathcal{M}^*(\mathcal{L}(X);\mathbb{C})$ (cf. 10.4).

The mapping I is **well defined** and **$\mathbb{C}$-linear**: The **complete** and **regular complex measure** $\mu = \mathrm{Re}\mu^+ - \mathrm{Re}\mu^- + i(\mathrm{Im}\mu^+ - \mathrm{Im}\mu^-)$ with

$$\mu(A) = \int \chi_A \mathrm{Re}h^+ d|\mu| - \int \chi_A \mathrm{Re}h^- d|\mu| + i\left(\int \chi_A \mathrm{Im}h^+ d|\mu| - \int \chi_A \mathrm{Im}h^- d|\mu|\right)$$

represented by four **complete** and **regular positive measures** according to 10.11 is mapped to the **complex** functional Λ with

$$\Lambda f = \int fd\mu = \int f\mathrm{Re}h^+ d|\mu| - \int f\mathrm{Re}h^- d|\mu| + i\left(\int f\mathrm{Im}h^+ d|\mu| - \int f\mathrm{Im}h^- d|\mu|\right)$$

constructed of four **positive bounded functionals** matching the four summands in the decomposition of Λ in 11.6. Since the range of μ resp. Λ has been extended to $\mathbb{C}$ the mapping is now completely $\mathbb{C}$-linear.

The mapping I is **surjective**: For every complex functional $\Lambda = \mathrm{Re}\Lambda^{+} - \mathrm{Re}\Lambda^{-} + i\left(\mathrm{Im}\Lambda^{+} + \mathrm{Im}\Lambda^{-}\right)$ each **positive bounded functional** of the decomposition according to 11.6 is represented by an integral, e.g. $\mathrm{Re}\Lambda^{+} f = \int f d\left(\mathrm{Re}\mu^{+}\right)$ for every $f \in C_c\left(X, \mathbb{C}\right)$ resp. a **complete** and **σ-regular positive Borel** measure $\mathrm{Re}\mu^{+}$ etc. due to the preceding version 11.10 of the **Riesz representation** theorem such that $\mu = \mathrm{Re}\mu^{+} - \mathrm{Re}\mu^{-} + i\left(\mathrm{Im}\mu^{+} - \mathrm{Im}\mu^{-}\right)$ is the uniquely determined **complete** and **σ-regular complex Borel** measure with $\Lambda f = \int f d\mu$ for every $f \in C_c\left(X, \mathbb{C}\right)$. For any **complete** and **σ-regular positive Borel** measure λ determined by a **positive bounded** functional Γ, every compact K and $f \in C_c\left(X, [0;1]\right)$ with $K \prec f$ according to 11.10 we have $\lambda\left(K\right) \leq \int f d\lambda \overset{11.9}{=} \Gamma f \overset{11.1}{\leq} \|\Gamma\|^{*} \cdot \|f\| = \|\Gamma\|^{*}$ and on account of the **regularity condition** follows $\|\lambda\| = \mu\left(X\right) = \sup\left\{\lambda\left(K\right) : K \text{ compact}\right\} \leq \|\Gamma\|^{*}$. Hence every component of μ is **finite** and since this condition transfers to μ itself it is also **regular**.

The mapping I is **injective**: Assuming $\Lambda = 0$, i.e. $\Lambda f = \int f h d|\mu| = 0$ for every $f \in C_c\left(X, \mathbb{C}\right)$. Since according to 11.2 the space $C_c\left(X, \mathbb{C}\right)$ is **dense** in $L^1\left(|\mu|\right)$ this implies $\int \chi_A h d|\mu| = \int_A h d|\mu| = 0$ for every measurable A and hence $|\mu|$-a.e. $h = 0$. But on the other hand we have $|h| = 1$ which only leaves $|\mu|\left(X\right) = 0$, i.e. $\mu = 0$. Thus $\ker I = \{0\}$ which implies the assertion.

The mapping I is **isometric**: On the one hand we have $\|\Lambda\|^{*} = \sup\left\{\left|\frac{\int f h d|\mu|}{\sup|f|}\right| : f \in C_c\left(X, \mathbb{R}\right)\right\} = \sup\left\{\left|\int f h d|\mu|\right| : f \in C_c\left(X, \mathbb{R}^{+}\right), \sup f = 1\right\} \leq |\mu|\left(X\right) = \|\mu\|$. On the other hand according to **Lusin's theorem** 11.3 for every $\epsilon > 0$ there exists a $g \in C_c\left(X, \mathbb{C}\right)$ such that $|\mu|\left(\overline{h} \neq g\right) < \epsilon$ and $\|g\| \leq 1$ such that $\|\Lambda\|^{*} \geq \int_X g h d|\mu| \geq |\mu|\left(X \setminus \left\{\overline{h} \neq g\right\}\right) - |\mu|\left(\overline{h} \neq g\right) \geq |\mu|\left(X\right) - 2\epsilon$, hence $\|\Lambda\|^{*} \geq \|\mu\|$.

11.12 Theorem: If every **open** set is **σ-compact** (cf. [9, 10.6]) and λ is a **positive Borel measure** such that $\lambda\left(K\right) < \infty$ for every **compact** K then λ is **σ-finite** and **regular**.

Proof: The **σ-finiteness** directly follows from the hypotheses. We have to show that the measure μ on $\mathcal{L}\left(X\right)$ determined by Λ with $\Lambda f = \int f d\lambda = \int f d\mu$ for $f \in C_c\left(X; \mathbb{R}\right)$ coincides with λ on $\mathcal{B}\left(X\right) \subset \mathcal{L}\left(X\right)$. For any open $V = \bigcup_{n \in \mathbb{N}} K_n$ with compact K_n due to **Urysohn's lemma** (cf. [9, 10.5]) there are $f_n \in C_c\left(X; \mathbb{R}\right)$ with $K_n \prec f_n \prec V$. Then $g_n = \max\left\{f_0; ...; f_n\right\} \in C_c\left(X; \mathbb{R}\right)$ and $g_n\left(x\right)$ increases to $\chi_{V(x)}$ at every $x \in X$. By the **monotone convergence theorem** 5.12 we conclude that $\lambda\left(V\right) = \sup\limits_{n \in \mathbb{N}} \int g_n d\lambda = \sup\limits_{n \in \mathbb{N}} \int g_n d\mu = \mu\left(V\right)$. Since μ is **regular** for any $A \in \mathcal{L}\left(X\right)$ and $\epsilon > 0$ exist closed B and open V such that $B \subset A \subset V$ and $\mu\left(V \setminus B\right) < \epsilon$. Hence $\mu\left(V\right) \leq \mu\left(B\right) + \epsilon \leq \mu\left(A\right) + \epsilon$. Since $V \setminus B$ is open the first part of the proof yields $\lambda\left(V \setminus B\right) < \epsilon$ resp. $\lambda\left(V\right) \leq \lambda\left(A\right) + \epsilon$. Consequently we find that on the one hand $\lambda\left(A\right) \leq \lambda\left(V\right) = \lambda\left(V\right) \leq \mu\left(A\right) + \epsilon$ and on the other hand $\mu\left(A\right) \leq \mu\left(V\right) = \lambda\left(V\right) \leq \lambda\left(A\right) + \epsilon$ so that $\left|\lambda\left(A\right) - \mu\left(A\right)\right| < \epsilon$. Hence we conclude that $\lambda\left(A\right) = \mu\left(A\right)$.

11.13 Lebesgue measure: Since $\mathbb{R}^n$ is **σ-compact** we can apply the preceding theorem to the Lebesgue-Borel measure λ^n and obtain its **σ-finite**, **regular** and **complete** extension, the **Lebesgue measure** λ^n on the extended σ-algebra $\mathcal{L}\left(\mathbb{R}^n\right)$ of the **Lebesgue measurable** sets. A set A is **Lebesgue measurable** iff there are an F_σ-set F and a G_δ-set G such that $F \subset A \subset G$ and $\lambda^n\left(G \setminus F\right) = 0$. This follows from 11.8 resp. 11.9 and the σ-compactness of $\mathbb{R}^n$ together with the observation that for **any** σ-compact set A with $A = \bigcup_{n \in \mathbb{N}} K_n$ for a sequence of compact K_n and any other given compact K the intersection $A \cap K \in \mathcal{A}\left(X\right)$ since $\lambda^n\left(A \cap K\right) = \sup\left\{\lambda^n\left(K_n \cap K\right)\right\} < \infty$. Consequently **every Lebesgue set is the union of a Borel measuralbel G_δ-set and a λ^n-null set**. Thus every **Lebesgue measurable function** f coincides λ^n-a.e. with a Borel measurable function f_0 and identical integral $\int_A f d = \int_A f_0 d$ for every Lebesgue measurable A. The **translation invariance** 8.8 as well as the **transformation formula** 8.9 extend from $\mathcal{B}\left(X\right)$ to $\mathcal{L}\left(X\right)$due to the **regularity** of λ^n.

12 Differentiation

12.1 Dini derivatives: In this section we will prove **Lebesgue's differentiation theorem** which states that a monotone function has a finite derivative almost everywhere. To this end we study the **Dini derivatives** of a function $f : [a; b] \to \mathbb{R}$ for $a < x < b$, i.e. the **lower right derivate**

$(D_+f)(x) = \underline{\lim}\limits_{h\downarrow 0}\frac{f(x+h)-f(x)}{h}$ and the **upper right derivate** $(D^+f)(x) = \overline{\lim}\limits_{h\downarrow 0}\frac{f(x+h)-f(x)}{h}$ resp. **lower left derivate** $(D_-f)(x) = \underline{\lim}\limits_{h\uparrow 0}\frac{f(x+h)-f(x)}{h}$ and the **upper left derivate** $(D^-f)(x) = \overline{\lim}\limits_{h\uparrow 0}\frac{f(x+h)-f(x)}{h}$. These definitions contain the usual notations for e.g. the **lower right limit** $\underline{\lim}\limits_{h\downarrow x}\varphi(h) = \sup\limits_{\delta>0}\inf\limits_{x<h<x+\delta}\varphi(h)$ and the **upper right limit** $\overline{\lim}\limits_{h\downarrow x}\varphi(h) = \inf\limits_{\delta>0}\sup\limits_{x<h<x+\delta}\varphi(h)$ resp. the **lower left limit** $\underline{\lim}\limits_{h\uparrow x}\varphi(h) = \sup\limits_{\delta>0}\inf\limits_{x-\delta<h<x}\varphi(h)$ and the **upper left limit** $\overline{\lim}\limits_{h\uparrow x}\varphi(h) = \inf\limits_{\delta>0}\sup\limits_{x-\delta<h<x}\varphi(h)$. Obviously we have $(D_+f)(x) \le (D^+f)(x)$ and $(D_-f)(x) \le (D^-f)(x)$. In the case of equality we obtain the **right derivative** $D_+^+f(x) = (D_+f)(x) = (D^+f)(x)$ resp. the **left derivative** $D_-^-(x) = (D_-f)(x) = (D^-f)(x)$. Finally if these two also conincide we arrive at the **derivative** $\frac{df}{dy}(x) = D_+^+(x) = D_-^-(x)$ and consequently f is **differentiable** at x. The **derivative of a real function needs not be finite**, e.g. $\frac{df}{dy}(0) = \infty$ for $f(x) = x^{\frac{1}{3}}$. This is in contrast to the **complex case** where we define the derivative as a sum $\frac{df}{dy}(x) = \frac{d\mathrm{Re}f}{dy}(x) + i\frac{d\mathrm{Im}f}{dy}(x)$ requiring **finite** summands $\frac{d\mathrm{Re}f}{dy}(x)$ and $\frac{d\mathrm{Im}f}{dy}(x)$ such that we can write $\frac{df}{dy}(x) = \lim\limits_{|h|\to 0}\frac{f(x+h)-f(x)}{h}$ since both limits from every direction exist and are finite. Obviously every function with finite derivative at x is **continuous** at x whereas the converse is not true at all: the set of all functions on $[0;1]$ which have at least one infinite right derivative at every point in $[0;1]$ is evern **dense** in $C([0;1])$(c.f. [2, th 17.8]).

12.2 Theorem: For every **real** $f : \mathbb{R} \to \mathbb{R}$ the set $\left\{D_+f = D^+f = D_+^+f \neq D_-^-f = D_-f = D^-f\right\}$ is **countable**. **Note**: The cases $D_+f(x) = D^+f(x) = \pm\infty$ resp. $D_-f(x) = D^-f(x) = \pm\infty$ are included in this set. Figuratively it contains every well behaved **jump** and **corner** point but excludes points where the **oscillation** of f prevents the existence of a right resp. left limit of the differential quotients.

Proof: For each $x \in A = \left\{D_+f = D^+f = D_+^+f < D_-^-f = D_-f = D^-f\right\}$ there are $r_x; s_x; t_x \in \mathbb{Q}$ such that $D_+^+f(x) < r_x < D_-^-f(x)$ and $s_x < x < t_x$ such that $\frac{f(y)-f(x)}{y-x} < r_x \Leftrightarrow f(y) - f(x) > r_x(y-x)$ for $x < y < t_x$ and $\frac{f(y)-f(x)}{y-x} > r_x \Leftrightarrow f(y) - f(x) > r_x(y-x)$ for $s_x < y < x$, i.e. the respective second inequality holds for every $s_x < y < t_x$ with $y \neq x$. The mapping $\varphi : A \to \mathbb{Q}^3$ with $\varphi(x) = (r_x; s_x; t_x)$ is **injective** since for $\varphi(x) = (r_x; s_x; t_x) = \varphi(y)$ we have $s_x < x; y < t_x$ and assuming $x \neq y$ we obtain $f(y) - f(x) > r_x(y-x)$ but also with reversed roles $f(x) - f(y) > r_x(x-y)$ yielding the contradiction $0 > 0$. Thus from the countable character of $\mathbb{Q}^3$ we can deduce the countability of A and likewise that of the complementary set $\left\{D_+f = D^+f = D_+^+f > D_-^-f = D_-f = D^-f\right\}$ whence follows the assertion.

12.3 Vitali's covering theorem: Let $A \subset \mathbb{R}$ be an arbitrary set of real numbers and $\mathcal{V}$ a **Vitali cover** of A, i.e. consisting of closed intervals I of positive length $\lambda(I) > 0$ such that for every $\epsilon > 0$ every $x \in A$ is contained in an interval $x \in I \in \mathcal{V}$ with $\lambda(I) < \epsilon$. Then $\mathcal{V}$ has a pairwise disjoint **countable** subset $(I_n)_{n\in\mathbb{N}} \subset V$ **covering** A **almost everywhere**, i.e. $\lambda(A \setminus \bigcup_{n\in\mathbb{N}} I_n) = 0$. In the case of $\lambda(A) < \infty$ we can even state an approximation criterion, i.e. for every $\epsilon > 0$ there is an $m \in \mathbb{N}$ such that $\lambda\left(A \setminus \bigcup\limits_{n=0}^{m} I_n\right) = 0$.

Proof:

Case I: $\lambda(A) < \infty$: Due to the **regularity** 11.1.3 of the Lebesgue measure we can find an **open** $V \supset A$ with $\lambda(V) < \infty$. Since in locally compact spaces the closed neighbourhoods resp. intervals form a **neighbourhood basis** of every $x \in \mathbb{R}$ the subfamily $\mathcal{V}_0 = \{I \in \mathcal{V} : I \subset V\}$ is again a Vitali cover of A. Choose an $I_0 \in \mathcal{V}_0$ and proceed by **induction** as follows: Assume having already chosen pairwise disjoint closed intervals $I_0; ..; I_n \in \mathcal{V}_0$ with $S_n = \bigcup\limits_{k=0}^{n} I_k$ and further assume $A \setminus S_n \neq \emptyset$; otherwise the construction is complete. Then $U_n = V \setminus S_n \neq \emptyset$ is open with $A \setminus S_n \neq \emptyset$ so that we can find an $I_{n+1} \in \mathcal{V}_0$ with $I_{n+1} \subset U_n$ and $\lambda(I_{n+1}) > \frac{\delta_n}{2}$ with $\delta_n = \sup\{\lambda(I) : I \subset U_n \wedge I \in \mathcal{V}_0\}$. We must show that $\lambda(A \setminus S) = 0$ with $S = \bigcup_{n\in\mathbb{N}} I_n$. For each $I_n = [a_n; b_n]$, let $J_n = [3a_n - 2b_n; 3b_n - 2a_n]$ be the closed interval with the same midpoint as I_n and $\lambda(J_n) = 5\lambda(I_n)$ and in particular $\lambda\left(\bigcup\limits_{n=0}^{\infty} J_n\right) \leq \sum\limits_{n=0}^{\infty} \lambda(J_n) = 5 \sum\limits_{n=0}^{\infty} \lambda(I_n) = 5\lambda(S) < 5\lambda(V) < \infty$ such that $\lim\limits_{k\to\infty} \lambda\left(\bigcup\limits_{n=k}^{\infty} J_n\right) = 0$ due to the **σ-additivity** resp. **continuity** of λ. Hence it suffices to show that $A \setminus S \subset \bigcup\limits_{n=k}^{\infty} J_n$ for every $k \in \mathbb{N}$. To this end let $x \in A \setminus S$ and $k \in \mathbb{N}$ such that $x \in A \setminus S_k \subset U_k$ whence there is an $I \in \mathcal{V}_0$ with $x \in I \in U_k$. On account of $\delta_n < 2\lambda(I_{n+1})$ and $\lim\limits_{n\to\infty} \lambda(I_n) = 0$ there is an $m \in \mathbb{N}$ with $\delta_n < \lambda(I)$ resp. $I \subsetneq U_n$ for every $n \geq m$. Due to the well-ordering of the natural numbers there is a smallest $l \in \mathbb{N}$ with $I \subsetneq U_l$ and obviously we have $l > k$, hence $I \cap S_l \neq \emptyset$ and $I \cap S_{l-1} = \emptyset$ such that $I \cap I_l \neq \emptyset$. On account of $I \subset U_{l-1}$ we have $\lambda(I) \leq \delta_{l-1} < 2\lambda(I_l)$ and with $\lambda(J_l) = 5\lambda(I_l)$ follows $x \in I \subset J_l \subset \bigcup\limits_{n=l}^{\infty} J_n$. Since this is true for every $k \in \mathbb{N}$ we have shown that $\lambda(A \setminus S) = 0$. Concerning the approximation criterion let $\epsilon > 0$ and choose a $p \in \mathbb{N}$ large enough so that $\sum\limits_{n=p+1}^{\infty} \lambda(I_n) < \epsilon$. Then we have $A \setminus S_p \subset A \setminus S_p \cup \bigcup\limits_{n=p+1}^{\infty} I_n$ and consequently $\lambda(A \setminus S_p) \leq 0 + \lambda\left(\bigcup\limits_{n=p+1}^{\infty} I_n\right) < \epsilon$.

Case II: $\lambda(A) = \infty$. For every $n \in \mathbb{Z}$ the subfamily $\mathcal{V}_n = \{I \in \mathcal{V} : I \subset]n; n+1[\}$ is a Vitali cover of the section $A_n = A \cap]n; n+1[$ so that we can apply case I to obtain a pairwise disjoint countable selection $\mathcal{W}_n \subset \mathcal{V}_n$ with $\lambda(A_n \setminus \bigcup \mathcal{W}_n) = 0$. The union $\mathcal{W} = \bigcup_{n\in\mathbb{Z}} \mathcal{W}_n$ still is a paiwise disjoint countable family with $\lambda(A \setminus \bigcup \mathcal{W}) = \lambda(\mathbb{Z} \cup \bigcup_{n\in\mathbb{Z}} (A_n \setminus \bigcup \mathcal{W}_n)) \leq \lambda(\mathbb{Z}) + \sum_{n\in\mathbb{Z}} \lambda(A_n \setminus \bigcup \mathcal{W}_n) = 0$.

12.4 Lebesgue's differentiation theorem: Every **real-valued monotone** function f on a **closed** interval $[a; b] \subset \mathbb{R}$ has a **finite derivative** λ-a.e. on $[a; b]$.

Proof: W.l.o.g. suppose f is **nondecreasing**. We first prove that $\lambda(A) = 0$ for $A = \{D_+ f < D^+ f\} \cap [a; b[$. We decompose this set into countable sections $A_{u;v} = \{x \in A : D_+ f(x) < u < v < D^+ f(x)\}$ for every $u; v \in \mathbb{Q}$ with $A = \bigcup \{A_{u;v} : 0 < u < v \in \mathbb{Q}\}$ such that it suffices to show that $\lambda(A_{u;v}) = 0$ for every $0 < u < v \in \mathbb{Q}$. Assume that there is a pair $0 < u < v \in \mathbb{Q}$ with $\lambda(A_{u;v}) = \alpha > 0$. Due to the regularity of λ for any $\epsilon > 0$ there is an open $U \supset A_{u;v}$ such that $\lambda(U) < \alpha + \epsilon$ (1). According to the assumption for each $x \in A_{u;v}$ there are arbitrarily small $h > 0$ with $[x; x+h] \subset U \cap [a; b]$ and $\frac{f(x+h)-f(x)}{h} < u$ (2). Hence the family $\mathcal{V}$ of all such closed intervals is a **Vitali cover** of $A_{u;v}$ whence due to the preceding theorem there is a **finite and pairwise disjoint subfamily** $([x_i; x_i + h_i])_{i=0}^{m}$ such that $\bigcup\limits_{i=0}^{m} [x_i; x_i + h_i] \subset U$ (3), $\lambda\left(A_{u;v} \setminus \bigcup\limits_{i=0}^{m} [x_i; x_i + h_i]\right) < \epsilon$ and $\sum\limits_{i=0}^{m} (f(x_i + h_i) - f(x_i)) \overset{(2)}{<} u \sum\limits_{i=0}^{m} h_i \overset{(3)}{\leq} u \cdot \lambda(U) \overset{(1)}{<} u(\alpha + \epsilon)$ (4). Having identified sets with **small inferior limits** of the differential quotient $\frac{f(x+h)-f(x)}{h}$ we now look for subsets of those same intervals with **large superior limits** of the differential quotient. The comparison of the lengths h of these intervals depending on α and ϵ on the one hand and the corresponding increases $f(x+h) - f(x)$ restricted by the **monotone character** of f on the other hand will result in a delicate contradiction to the assumption $\alpha > 0$: In order to ensure the necessary margins for the differential quotients we remove the boundaries of the closed intervals from above and proceed with $V = \bigcup\limits_{i=0}^{m}]x_i; x_i + h_i[$. We still have $\lambda(A_{u;v} \setminus V) < \epsilon$ and for all $y \in A_{u;v} \cap V$ there are arbitratrily small $k > 0$ with $[y; y+k] \subset V$ and $\frac{f(y+k)-f(y)}{k} > v$ (5). These intervals constitute a Vitali cover of $A_{u;v} \cap V$ such that we find a **finite and pairwise disjoint subfamily** $([y_j; y_j + k_j])_{j=0}^{n}$ with $\lambda\left((A_{u;v} \cap V) \setminus \bigcup\limits_{j=0}^{n} [y_j; y_j + k_j]\right) < \epsilon$ and hence $\alpha \leq \lambda(A_{u;v} \setminus V) + \lambda(A_{u;v} \cap V) < \epsilon + \left(\epsilon + \sum\limits_{j=0}^{n} k_j\right)$. Substituting this inequality in (5) yields $v(\alpha - 2\epsilon) < v \sum\limits_{j=0}^{n} k_j < \sum\limits_{j=0}^{n} (f(y_j + k_j) - f(y_j))$ (6). Since $\bigcup\limits_{j=0}^{n} [y_j; y_j + k_j] \subset \bigcup\limits_{i=0}^{m} [x_i; x_i + h_i]$ and f is **nondecreasing** we also have $\sum\limits_{j=0}^{n} (f(y_j + k_j) - f(y_j)) \leq \sum\limits_{i=0}^{m} (f(x_i + h_i) - f(x_i))$. Substituting (4) and (6) in this estimate results in $v(\alpha - 2\epsilon) < u(\alpha + \epsilon)$. But that implies $\epsilon > \frac{\alpha(v-u)}{u+2v}$ in contradiction to the regularity of λ

which let us find an open $U \supset A_{u;v}$ such that $\lambda(U) < \alpha + \epsilon$ for **any** $\epsilon > 0$. Thus $\lambda(A) = 0$ and so $D_+^+ f(x)$exists λ-a.e. on $[a;b]$. Analogously we can show the same result for $D_-^- f(x)$. The assertion then follows from 12.2.

For every $x \in F = \{f' = \infty\} \cap]a;b[$ and every $n \in \mathbb{N}$ there exist arbitrarily small $h > 0$ such that $[x;x+h] \subset]a;b[$ and $\frac{f(x+h)-f(x)}{h} > n$. Again we invoke Vitali's theorem to obtain a **countable, pairwise disjoint family** $([x_k;x_k+h_k])_{k\in\mathbb{N}}$ such that $\lambda\left(F \setminus \bigcup_{k\in\mathbb{N}}[x_k;x_k+h_k]\right) = 0$. Hence we have $n\lambda(F) \le n\sum_{k\in\mathbb{N}} h_k < \sum_{k\in\mathbb{N}}(f(x_k+h_k) - f(x_k)) \le f(b) - f(a)$ for every $n \in \mathbb{N}$ and since f is supposed to be real-valued we conclude $\lambda(F) = 0$.

12.5 Total variation: The total variation of a **complex** function $f : [a;b] \to \mathbb{C}$ is defined as $V_a^b f = \sup\left\{\sum_{k=1}^{n} |f(x_k) - f(x_{k-1})| : a = x_0 < ... < x_n = b; n \ge 1\right\}$. It is **finite** iff the total variations $V_a^b \mathrm{Re} f$ and $V_a^b \mathrm{Im} f$ are finite. For $a < b < c$ we have $V_a^b f + V_b^c f = V_a^c f$ and the function $x \mapsto V_a^x f$ is **nondecreasing**. The set $\mathcal{D}([a;b])$ of all complex functions $f : [a;b] \to \mathbb{C}$ with $f(a) = 0$ and $V_a^b f < \infty$ with the **norm** $\|f\|_V = V_a^b f$ is a **Banach space** since on account of $\|f\|_\infty \le \|f\|_V$ every $\|\|_V$-Cauchy sequence $(f_n)_{n\in\mathbb{N}} \subset \mathcal{D}([a;b])$ also is a $\|\|_\infty$-Cauchy sequence converging **uniformly** to an $f : [a;b] \to \mathbb{C}$ which therefore satisfies $f(a) = 0$ and $V_a^b f < \infty$.

12.6 Jordan decomposition of functions: Every **real** function f with **finite variation** can be represented as the difference of two **nondecreasing** functions.

Proof: Define $V_a^a f = 0$ for some $a \in \mathbb{R}$ and write $f(x) = V_a^x f - (V_a^x f - f(x))$. The latter part is nondecreasing since for real f and $x < x'$ we always have $V_x^{x'} f \ge f(x') - f(x)$.

12.7 Lebesgue differentiation theorem for complex functions: Every **complex** function with **finite variation** has a **finite derivative** λ-a.e.

Proof: Apply 12.6, 12.4 and 12.1.

12.8 Fubini's differentiable series theorem: For a sequence $(f_n)_{n\in\mathbb{N}}$ of **monotone** functions $f_n : [a;b] \to \mathbb{R}$ such that $s(x) = \sum_{n\in\mathbb{N}} f_n(x)$ is **finite** for every $x \in [a;b]$ the **derivative** exists λ-a.e. in $]a;b[$ and coincides with the limit of the derivatives of the partial sums: $\frac{ds}{dy}(x) = \sum_{n\in\mathbb{N}} \frac{df_n}{dy}(x)$.

Proof: W.l.o.g. we assume positive and nondecreasing f_n such that s is also positive and nondecreasing. Hence according to 12.4 all f_n as well as the partial sums $s_n = \sum_{k=0}^{n} f_k$ and the limit s have λ-a.e. finite derivatives. Since all f_n are nondecreasing we have $\frac{s(x+h)-s(x)}{h} \ge \frac{s_{n+1}(x+h)-s_{n+1}(x)}{h} \ge \frac{s_n(x+h)-s_n(x)}{h} > 0$ for every $x, x+h \in]a;b[$ and hence λ-a.e. $\frac{ds_n}{dy}(x) \le \frac{ds_{n+1}}{dy}(x) \le \frac{ds}{dy}(x)$ for every $n \in \mathbb{N}$. Consequently $\lim\limits_{n\to\infty} \frac{ds_n}{dy}(x) = \sum_{n\in\mathbb{N}} \frac{df_n}{dy}(x) < \infty$ λ-a.e. and it remains to show that $\lim\limits_{n\to\infty} \frac{ds_n}{dy}(x) = \frac{ds}{dy}(x)$ λ-a.e. To this end we investigate the right boundary b and choose an increasing subsequence $(n_k)_{k\in\mathbb{N}} \subset \mathbb{N}$ such that $s(b) - s_{n_k}(b) < \frac{1}{k^2}$ and hence $\sum_{k\in\mathbb{N}}(s(b) - s_{n_k}(b)) < \infty$. Since s and all s_{n_k} are monotone we infer $\sum_{k\in\mathbb{N}}(s(x) - s_{n_k}(x)) < \infty$ for all $x < b$ and since all $s - s_{n_k}$ are again monotone with finite derivatives we can apply the inequality with the differential quotients form above to conclude that $\sum_{k\in\mathbb{N}}\left(\frac{ds}{dy}(x) - \frac{ds_{n_k}}{dy}(x)\right) < \infty$ and hence $\lim\limits_{k\to\infty} \frac{ds_{n_k}}{dy}(x) = \frac{ds}{dy}(x)$. Since all $\frac{df_n}{dy}$ are positive and hence $\left(\frac{ds_n}{dy}(x)\right)_{n\in\mathbb{N}}$ is nondecreasing we infer $\lim\limits_{n\to\infty} \frac{ds_n}{dy}(x) = \frac{ds}{dy}(x)$.

12.9 Mean value theorem: For every **differentiable** function $f : [a;b] \to \mathbb{R}$ there is an $x \in [a;b]$ with $\frac{df}{dy}(x) = \frac{f(b)-f(a)}{b-a}$.

Proof: In a first step we prove **Rolle's theorem** which covers the case $f(a) = f(b)$: Since the continuous image $f[[a;b]]$ of the **compact** interval is again compact and particularly closed it has a minimum $f_{min} = \inf\limits_{a\le x\le b} f(x) \in f[[a;b]]$ and a maximum $f_{max} = \sup\limits_{a\le x\le b} f(x) \in f[[a;b]]$, i.e. there are $a \le x_{min}; x_{max} \le b$ with $f(x_{min}) = f_{min}$ resp. $f(x_{max}) = f_{max}$. Since $f(a) = f(b)$ and $f_{min} \le f_{max}$ one of them must ly inside the interval and w.l.o.g. we can assume $a < x_{max} < b$. Hence according to 12.1 we have $0 \le D_-^-(x_{max}) \le \frac{df}{dy}(x_{max}) \le D_+^+(x_{max}) \le 0$ and hence $\frac{df}{dy}(x_{max}) = 0$.

We apply this result to the general case via the function $g\left(x\right)=f\left(x\right)-\frac{f(b)-f(a)}{b-a}\cdot\left(x-a\right)$ with $\frac{dg}{dy}\left(x\right)=\frac{dg}{dy}\left(x\right)-\frac{f(b)-f(a)}{b-a}$.

12.10 Corollary: Every **differentiable** function $f:[a;b]\to\mathbb{R}$ with vanishing derivative $\frac{df}{dy}\left(x\right)=0$ for every $a\leq x\leq b$ is constant. In particulat every differentiable functions is determined by its derivative up to an additive constant: $\forall a\leq x\leq b:\frac{df}{dy}\left(x\right)=\frac{dg}{dy}\left(x\right)\Rightarrow\exists c\in\mathbb{R}:f=g+c$.

13 The fundamental theorem of calculus

In 2.4 and 3.7 we saw that **every lower semicontinuous** and **nondecreasing** function $f:\mathbb{R}\to\mathbb{R}$ generates a **σ-finite Lebesgue-Borel-Stieltjes measure** $\lambda_f:\mathcal{B}\to[0;\infty]$ on the Borel σ-algebra $\mathcal{B}=\sigma\left(\mathcal{I}\right)=\sigma\left(\mathcal{F}\right)$ induced by the **right-open intervals** $\mathcal{I}=\{[a;b[:\ a\leq b\in\mathbb{R}\}$ resp. the **algebta** $\mathcal{F}=\left\{\overset{\circ}{\bigcup_{0\leq k\leq m}}I_k:I_k\in\mathcal{I},m\in\mathbb{N}\right\}$ of the **one-dimensional figures** by $\lambda_f\left([a;b[\right)=f\left(b\right)-f\left(a\right)$. The relation is **bijective** if we restrict the domain of the functions and the range of the measures slightly:

13.1 Theorem: The mapping $f\mapsto\lambda_f$ with $\lambda_f\left([a;b[\right)=f\left(b\right)-f\left(a\right)$ between the **lower semicontinuous** and **nondecreasing** functions $f:\mathbb{R}\to\mathbb{R}$ **vanishing at** 0 and the **σ-finite** and **regular Borel measures** $\lambda_f:\mathcal{B}\left(\mathbb{R}\right)\to[0;\infty]$ is **bijective** with the **inversion** $\lambda\mapsto f_\lambda$ defined by $f_\lambda\left(t\right)=\begin{cases}\lambda\left([0;t[\right) & :t\geq 0\\ -\lambda\left([t;0[\right) & :t<0\end{cases}$. Furthermore **every σ-finite** Borel measure $\lambda:\mathcal{B}\left(\mathbb{R}\right)\to[0;\infty]$ is **regular** and has a **complete** extension**.**

Proof: λ_f is **regular** since it is the uniquely determined **σ-finite** and **σ-regular positive** measure defined by the **positive functional** $I_{\lambda_f}\in\left(C_c\left(\mathbb{R};\mathbb{R}\right)\right)^*_+$ with $I_{\lambda_f}g=\int gd\lambda_f$ according to the **Riesz representation theorem** 11.10 resp. theorem 11.12 on the **σ-algebra** $\mathcal{L}\left(\mathbb{R}\right)\supset\mathcal{B}$ on the **σ-compact** set $\mathbb{R}$. Note that the **σ-finiteness** of λ_f results from f being **finite** on $\mathbb{R}$. Since the mapping is obviously **injective** it remains to show that it is **surjective**: Let λ be any regular Borel measure and f_λ defined as above. Then f_λ is **nondecreasing** since for $0\leq s<t$ we have $f_\lambda\left(t\right)-f_\lambda\left(s\right)=\lambda\left([0;t[\right)-\lambda\left([0;s[\right)=\lambda\left([0;t[\backslash[0;s[\right)=\lambda\left([s;t[\right)>0$ and similarly for the other cases $s<0<t$ resp. $s<t\leq 0$. Also f_λ is **lower semicontinuous** since for every sequence $\left(\epsilon_n\right)_{n\in\mathbb{N}}\subset]0;1[$ with $\lim\limits_{n\to\infty}\epsilon_n=0$ and $t\geq 0$ we have $\lim\limits_{n\to\infty}f_\lambda\left(t-\epsilon_n\right)=\lim\limits_{n\to\infty}\lambda\left([0;t-\epsilon_n[\right)\overset{2.2.2}{=}\lambda\left(\bigcup_{n\in\mathbb{N}}[0;t-\epsilon_n[\right)=\lambda\left([0;t[\right)=f_\lambda\left(t\right)$ and analogously for $t<0$. The mapping $\lambda\mapsto f_\lambda$ is the **inverse** of the mapping $f\mapsto\lambda_f$ since for $t\geq 0$ we have $g_{\lambda_f}\left(t\right)=\lambda_f\left([0;t[\right)=f\left(t\right)-f\left(0\right)=f\left(t\right)$ and similarly for $t<0$, hence $g_{\lambda_f}=f$. Redundantly we also may prove the converse, i.e. $\mu_{f_\lambda}=\lambda$ by comparing $\mu_{f_\lambda}\left([s;t[\right)=f_\lambda\left(t\right)-f_\lambda\left(s\right)=\lambda\left([s;t[\right)$ for $0\leq s<t$ as shown above whence follows the identity of the two measures on $\mathcal{B}\left(\mathbb{R}\right)=\sigma\left(\mathcal{I}\right)$ according to the **uniqueness theorem** 3.4. Since the regularity is nowhere used in the proof we infer that in fact **every σ**-finite Borel measure is a **Riesz representation** and hence **regular** resp**. complete** on the extended σ-algebra $\mathcal{L}\left(\mathbb{R}\right)$ according to 11.8. Note that the **completeness** is achieved on **non Borel measurable** null sets in $\mathcal{L}\left(\mathbb{R}\right)\backslash\mathcal{B}\left(\mathbb{R}\right)$ so that it is useless to speak of a complete Borel measure.

13.2 Definitions: Our aim is the adaptation of the **fundamental theorem of calculus** and the formulae for **change of variables** resp. **integration by parts** to the **Lebesgue measure** λ^n on the Lebesgue measurable sets $\mathcal{L}\left(\mathbb{R}^n\right)$ according to theorem 11.13. In order to avoid unnecessary cluttering in this section we write λ for λ^n. We will make use of the **Lebesgue-Radon-Nikodym theorem** 10.8 and hence start with the **symmetric derivative** $\frac{\tilde{d\mu}}{d\lambda}:\mathbb{R}^n\to[0;\infty]$ of a **complex Borel measure** $\mu:\mathcal{B}\left(\mathbb{R}^n\right)\to\mathbb{C}$ defned by $\frac{\tilde{d\mu}}{d\lambda}\left(\mathbf{x}\right)=\lim\limits_{r\to 0}\frac{\mu(B_r(\mathbf{x}))}{\lambda(B_r(\mathbf{x}))}$ for $\mathbf{x}\in\mathbb{R}^n$ and the associated **Hardy-Littlewood maximal function** $M\mu\left(\mathbf{x}\right)=\sup\limits_{0<r<\infty}\frac{|\mu|(B_r(\mathbf{x}))}{\lambda(B_r(\mathbf{x}))}$ using the **total variation** $|\mu|$ of μ. Since for $n\geq 2$ the vector space $\mathbb{R}^n$ is not **linearly ordered** any more the symmetric derivative **has lost its information on direction** and provides only the **absolute value of the rate of change** in x. The maximal function $M\mu:\mathbb{R}^n\to[0;\infty]$ is **lower semicontinuous** and hence **measurable** since for $\mathbf{x}\in\{M\mu>\alpha\}$ there are $r>0$ resp. $\epsilon>0$ such that $|\mu|\left(B_r\left(\mathbf{x}\right)\right)>\left(\alpha+\epsilon\right)\lambda\left(B_r\left(\mathbf{x}\right)\right)$ and

$\mathbf{y} \in B_\delta(\mathbf{x})$ we have $|\mu|(B_{r+\delta}(\mathbf{y})) > |\mu|(B_r(\mathbf{x})) > (\alpha+\epsilon)\lambda(B_r(\mathbf{x})) = \alpha\lambda(B_{r'}(\mathbf{x})) > \alpha\lambda(B_\delta(\mathbf{y}))$ for $r' = r \cdot \sqrt[n]{1+\frac{\epsilon}{\alpha}}$ and $\delta < r' - r$ hence $B_\delta(\mathbf{x}) \subset \{M\mu > \alpha\}$.

13.3 Lemma: For any complex Borel measure μ on $\mathbb{R}^n$ and $\alpha > 0$ we have $\lambda(M\mu > \alpha) \le 3^n \frac{\|\mu\|}{\alpha}$ with the norm $\|\mu\| = |\mu|(\mathbb{R}^n)$.

Proof: Any **compact** $K \subset \{M\mu > \alpha\}$ is covered by **finitely** many $B_i = B_{r_i}(\mathbf{x}_i)$ with $|\mu|(B_i) > \alpha\lambda(B_i)$ for $1 \le i \le N$. Ordering the B_i by **decreasing radius**, starting with the largest one and subsequently discarding all remaining balls intersecting the current one we arrive at a **pairwise disjoint** subset $(B_j)_{j\in S}$ with $S \subset \{1;...;N\}$ such that $K \subset \bigcup_{j\in S} B_{3r_j}(x_j)$. Hence $\lambda(K) \le 3^n \cdot \sum_{j\in S}\lambda(B_j) \le \frac{3^n}{\alpha}\sum_{j\in S}|\mu|(B_i) \le 3^n\frac{\|\mu\|}{\alpha}$.

13.4 Lebesgue points: The **Hardy-Littlewood maximal function** Mf of an **integrable function** $f \in L^1(\lambda)$ is defined as the maximal function $M\mu$ associated to the measure μ with $d\mu = fd\lambda$, i.e. $(Mf)(\mathbf{x}) = \sup\limits_{0<r<\infty}\frac{1}{\lambda(B_r)}\int_{B_r(\mathbf{x})}|f|\,d\lambda = \sup\limits_{0<r<\infty}\frac{1}{\lambda(B_r)}\int_{B_r(\mathbf{x})}|f(\mathbf{y})|\,d\mathbf{y}$ using the familiar notation $\int_A f(\mathbf{y})\,d\mathbf{y} = \int_A fd\lambda$ with $\mathbf{y} \in \mathbb{R}^n$ according to 8.8 to denote the **Lebesgue integral**. Correspondingly in Lebesgue integrals we may use the suggestive notation $\frac{d\mu}{d\lambda} = \frac{d\mu}{d\mathbf{y}}$ for **Radon-Nikodym** resp. **symmetric derivatives**. An $\mathbf{x} \in \mathbb{R}^n$ is denoted a **Lebesgue point** of the integrable function $f \in L^1(\lambda)$ iff $\lim\limits_{r\to 0}\frac{1}{\lambda(B_r)}\int_{B_r(\mathbf{x})}|f(\mathbf{y}) - f(\mathbf{x})|\,d\mathbf{y} = 0$, which in particular means $f(\mathbf{x}) = \lim\limits_{r\to 0}\frac{1}{\lambda(B_r)}\int_{B_r(\mathbf{x})} f(\mathbf{y})\,d\mathbf{y}$. Obviously every point of continuity is a Lebesgue point but the converse is not true since the Lebesgue point restricts only the **average** oscillation of f in a neighbourhood of $\mathbf{x}$.

13.5 Theorem: For every Lebesgue integrable $f \in L^1(\lambda)$ **almost every** $\mathbf{x} \in \mathbb{R}^n$ is a **Lebesgue point**.

Proof: We show that $(Tf)(\mathbf{x}) = \limsup\limits_{r\to 0}(T_r f)(\mathbf{x}) = 0$ for $(T_r f)(\mathbf{x}) = \frac{1}{\lambda(B_r)}\int_{B_r(\mathbf{x})}|f(\mathbf{y}) - f(\mathbf{x})|\,d\mathbf{y}$ holds λ-a.e.: According to 11.2 for every $m \in \mathbb{N}$ there is a $g \in C_c(\mathbb{R}^n)$ so that $\|f-g\|_1 < \frac{1}{n}$. Since g is **continuous** we have $Tg = 0$. With $(T_r h)(\mathbf{x}) \le \frac{1}{\lambda(B_r)}\int_{B_r(\mathbf{x})}|h(\mathbf{y})|\,d\mathbf{y} + |h(\mathbf{x})|$ for $h = f - g$ we have $Tf \le Th + Tg = Th \le Mh + |h|$ and hence $\{Tf > 2\epsilon\} \subset \{Mh > \epsilon\} \cup \{|h| > \epsilon\}$ for every $\epsilon > 0$. On the one hand we have $\lambda(|h| > \epsilon) \le \frac{\|h\|_1}{\epsilon} \le \frac{1}{n\epsilon}$ and on the other hand lemma 13.3 yields $\lambda(Mh > \epsilon) \le 3^n\frac{\|h\|_1}{\epsilon} \le \frac{3^n}{n\epsilon}$ such that we arrive at $\lambda(\{Mh > \epsilon\} \cup \{|h| > \epsilon\}) \le \frac{3^n+1}{n\epsilon}$ for every $n \in \mathbb{N}$ and hence $\lambda(\{Mh > \epsilon\} \cup \{|h| > \epsilon\}) = 0$. Since λ is **complete** according to 11.13 we may infer $\lambda(Tf > 2\epsilon) = 0$.

13.6 Theorem: The **symmetric derivative** $\frac{\tilde{d\mu}}{d\lambda}$ of a **complex Borel measure** μ coincides λ-a.e. with the **Radon-Nikodym derivative** $\frac{d\mu}{d\lambda} \in L^1(\lambda)$ of its **λ-absolute continuous Lebesgue component** such that $\frac{\tilde{d\mu}}{d\lambda} = \begin{cases}\frac{d\mu}{d\lambda} & \text{if } \mu \ll \lambda \\ 0 & \text{if } \mu \perp \lambda\end{cases}$ and for **λ-absolutely continuous** μ and every Borel set $A \subset \mathcal{B}(\mathbb{R}^n)$ we have $\mu(A) = \int_A \frac{d\mu}{d\lambda}d\lambda = \int_A \frac{\tilde{d\mu}}{d\lambda}d\lambda$.

Proof:

Case I: $\mu \ll \lambda$. According to 10.8.2 at any Lebesgue point x of $\frac{d\mu}{d\lambda}$ we have $\frac{d\mu}{d\lambda}(\mathbf{x}) = \lim\limits_{r\to 0}\frac{1}{\lambda(B_r)}\int_{B_r(\mathbf{x})}\frac{d\mu}{d\lambda}d\lambda$ $= \lim\limits_{r\to 0}\frac{\mu(B_r(\mathbf{x}))}{\lambda(B_r)} = \frac{\tilde{d\mu}}{d\lambda}(\mathbf{x})$. Hence the assertion follows from 13.5.

Case II: $\mu \perp \lambda$. Like the **Hardy-Littlewood maximal function** $M\mu(\mathbf{x}) = \sup\limits_{0<r<\infty}\frac{|\mu|(B_r(\mathbf{x}))}{\lambda(B_r(\mathbf{x}))}$ from 13.2 the **upper derivative** $D_u\mu(\mathbf{x}) = \inf\limits_{n\to\infty}\sup\limits_{0<r<1/n}\frac{|\mu|(B_r(\mathbf{x}))}{\lambda(B_r(\mathbf{x}))}$ is Borel measurable since $\sup\limits_{0<r<1/n}\frac{|\mu|(B_r(\mathbf{x}))}{\lambda(B_r(\mathbf{x}))}$ is **lower semicontinuous**. Due to 11.1.3 resp. 11.11 μ is **regular** and since $\mu \perp \lambda$ for every $\epsilon > 0$ there is **compact** set K with $\lambda(K) = 0$ but $\mu(K) \ge \|\mu\| - \epsilon$. Then we have $\left\|\mu_{X\setminus K}\right\| < \epsilon$ and for every $x \in X \setminus K$ holds $D_u\mu(\mathbf{x}) = D_u\mu_{X\setminus K}(\mathbf{x}) \le M\mu_{X\setminus K}(\mathbf{x})$ and hence $\{D_u\mu > \alpha\} \subset K \cup \left\{M\mu_{X\setminus K} > \alpha\right\}$ for every $\alpha > 0$. According to 13.3 follows $\lambda(D_u\mu > \alpha) \le 0 + 3^n\frac{\left\|\mu_{X\setminus K}\right\|}{\alpha} < 3^n\frac{\epsilon}{\alpha}$. Since for a given

$\alpha > 0$ for every $\epsilon > 0$ we can find a K such that this inequality holds we infer $\lambda\left(D_u\mu > \alpha\right) = 0$ and since this is true for every $\alpha > 0$ the assertion follows.

13.7 Absolute continuity and total variation: A **complex** function $f : [a; b] \to \mathbb{C}$ is **absolutely continuous** on $I = [a; b]$ if for every $\epsilon > 0$ there is a $\delta > 0$ such that for any disjoint collection $(]\alpha_i; \beta_i[)_{1\leq i\leq n}$ of segments with overall length $\sum\limits_{i=1}^{n} (\beta_i - \alpha_i) < \delta$ we have $\sum\limits_{i=1}^{n} |f(\beta_i) - f(\alpha_i)| < \epsilon$. Its **total variation function** $V_a f : [a; b] \to [0; \infty]$ with $V_a f(x) = V_a^x f$ from 12.5 is also **absolutely continuous** since with $V_{\alpha_i}^{\beta_i} f = \sup\left\{\sum\limits_{k=1}^{n} |f(x_k) - f(x_{k-1})| : \alpha_i = x_0 < ... < x_n = \beta_i; n \geq 1\right\}$ for segments $(]\alpha_i; \beta_i[)_{1\leq i\leq n}$ chosen as above we have $\sum\limits_{i=1}^{n} |V_a f(\beta_i) - V_a f(\alpha_i)| = \sum\limits_{i=1}^{n} V_{\alpha_i}^{\beta_i} f < \epsilon$.

13.8 Lemma: For every **nondecreasing** and **absolutely continuous** function $f : [a; b] \to \mathbb{R}$ on an interval $I = [a; b]$ and every **measurable** $A \in \mathcal{L}([a; b])$ (c.f. 11.13) we have $\lambda(A) = 0 \Rightarrow \lambda(f[A]) = 0$.

Proof: For $A \in \mathcal{S}([a; b])$ with $\lambda(A) = 0$ and every $\epsilon > 0$ there is a $\delta > 0$ such that **any** disjoint collection $(]\alpha_i; \beta_i[)_{1\leq i\leq n}$ of segments whose union $S = \bigcup\limits_{i=1}^{n}]\alpha_i; \beta_i[$ has measure $\lambda(S) = \sum\limits_{i=1}^{n} \lambda(]\alpha_i; \beta_i[) < \delta$ satisfies $\lambda(f[S]) = \sum\limits_{i=1}^{n} \lambda(f\,]\alpha_i; \beta_i[) = \sum\limits_{i=1}^{n} (f(\beta_i) - f(\alpha_i)) < \epsilon$. Due to the **regularity** of λ there is an open set $V \supset A$ with $\lambda(V) < \delta$ and the **second countability** of $\mathbb{R}$ implies the existence of a countable decomposition $V = \overset{\circ}{\bigcup}_{i\leq i<\infty}]\alpha_i; \beta_i[$. Since the condition of absolute continuity holds for **any** partial sum with $n \in \mathbb{N}$ it extends to the limit of the series. Hence we obtain $\lambda(f[A]) \leq \lambda(f[V]) < \epsilon$ for every $\epsilon > 0$, i.e. the assertion.

13.9 Lemma: For every **Lebesgue intgrable** $g : \mathbb{R} \to \mathbb{R}$ and $f(x) = \int_{-\infty}^{x} g(y)\, dy$ for $x \in \mathbb{R}$ we have $g(x) = \frac{df}{dy}(x)$ at every **Lebesgue point** x of g, hence λ-a.e.

Proof: Due to the hypotheses and 13.4 we have $\left|D_+^+ f(x) - g(x)\right| = \lim\limits_{h\downarrow 0} \frac{1}{h} \int_x^{x+h} |g(y) - g(x)|\, dy \leq \lim\limits_{h\downarrow 0} \frac{2}{2h} \int_{x-h}^{x+h} |g(y) - g(x)|\, dy = \lim\limits_{h\downarrow 0} \frac{2}{\lambda(B_h)} \int_{B_h(x)} |g(y) - g(x)|\, dy = 0$ and likewise for $D_-^- f(x)$.

13.10 Fundamental theorem of calculus: For every function $\boldsymbol{f} : [a; b] \to \mathbb{C}^n$ with components $f_i : [a; b] \to \mathbb{C}$ on a compact interval $[a; b] \subset \mathbb{R}$ we have the implications 1. $\Rightarrow$ 2. $\Rightarrow$ 3. $\Rightarrow$ 4. with

1. Every f_i has a **continuous derivative** $\frac{df_i}{dy} : [a; b] \to \mathbb{C}$.
2. Every f_i is **absolutely continuous**.
3. Every f_i is **λ-a.e. differentiable** with $\frac{df_i}{dy} \in L^1(\lambda)$.
4. For every $x \in [a; b]$ we have $\boldsymbol{f}(x) - \boldsymbol{f}(a) = \int_a^x \frac{d\boldsymbol{f}}{dy} dy$.

Proof:

1. $\Rightarrow$ 2.: We actually prove the formula $\boldsymbol{f}(x) - \boldsymbol{f}(a) = \int_a^x \frac{d\boldsymbol{f}}{dy} dy$ independently of 3. in order to show absolute coninuity: Since the continuous derivatives $\frac{df_i}{dy}$ are **uniformly continuous** on the **compact** interval $[a; b]$ the **step functions** $\varphi_n = \sum\limits_{j=0}^{n-1} \frac{df_i}{dy}(a_j)\, \chi_{[a_j; b_j[}$ with $a_j = a \cdot \left(1 - \frac{i}{n}\right) + b \cdot \frac{i}{n}$ and $b_j = a_j + \frac{1}{n}$ form an **approximating sequence** according to 5.5 for the $\frac{df_i}{dy}$ which are hence integrable. Also for every $\epsilon > 0$ there is an $h > 0$ with $\left|\frac{df_i}{dy}(\xi) - \frac{df_i}{dy}(\eta)\right| < \epsilon$ for all $a \leq \xi, \eta \leq b$ with $|\xi - \eta| < h$. Hence for the **integral functions** $\Phi_i(x) = \int_a^x \frac{df_i}{dy} dy$ we have $\left|\Phi_i(x + h) - \Phi_i(x) - h \cdot \frac{df_i}{dy}(x)\right| = \left|\int_{[x; x+h]} \left(\frac{df_i}{dy} - \frac{df_i}{dy}(x)\right) dy\right| \leq h \cdot \sup\limits_{x<z<x+h} \left|\frac{df_i}{dy}(t) - \frac{df_i}{dy}(x)\right| \leq h \cdot \epsilon$ whence $\lim\limits_{h\to 0} \frac{\Phi_i(x+h) - \Phi_i(x)}{h} = \frac{df_i}{dy}(x)$, i.e. $\frac{d\Phi_i}{dy} = \frac{df_i}{dy}$. Due to 12.10 we conclude $\Phi_i = f_i + c$ for some $c \in \mathbb{R}$ whence $f_i(b) - f_i(a) = \int_a^b \frac{df_i}{dy} dy$ for every $1 \leq i \leq n$. Since the continuous derivatives $\frac{df_i}{dy}$ are **bounded** on the **compact interval** $[a; b]$ we obtain the estimate $|f_i(\beta) - f_i(\alpha)| \leq \sup\limits_{a\leq x\leq b} f_i(x) \cdot |\beta - \alpha|$for arbitrary $a \leq \alpha < \beta \leq b$ and hence absolute continuity.

2. ⇒ 3.: Differentiation and integration are executed componentwise so that we can apply the **Jordan decomposition** 12.6 to split every component $f_i = V_a\mathrm{Re}f_i - (V_a\mathrm{Re}f_i - \mathrm{Re}f_i) + iV_a\mathrm{Im}f_i - i\,(V_a\mathrm{Im}f_i - \mathrm{Im}f_i)$ into four **nondecreasing** subcomponents so as to preserve the **information on direction** which will be lost in the multidimensional case. Due to 13.7 the **absolute continuity** extends from f_i to each of the four subcomponents. Hence we can assume a **nondecreasing and absolutely continuous** $f : [a;b] \to \mathbb{R}$ The plan is to apply the **Radon-Nikodym theorem** to the measure μ on the extended σ-algebra $\mathcal{L}([a;b])$ (cf. 11.13) defined by $\mu(A) = \lambda(g[A])$ with the now **strictly increasing** hence **injective** and still **absolutely continuous** function $g(x) = f(x)+x$. This measure is well defined since due to 11.13 there are a F_σ-set $F \in \mathcal{L}([a;b])$ and a G_δ-set $V \in \mathcal{L}([a;b])$ with $F \subset A \subset V$ and $\lambda(A \setminus F) \le \lambda(V \setminus F) = 0$ hence $\lambda(g[A \setminus F]) = 0$ due to 13.8 and consequently $A \setminus F \in \mathcal{L}([a;b])$. On the other hand due to g being **continuous** resp. [11, 9.2.1] the image $g[F] = g[\bigcup_{i\in\mathbb{N}} K_i] = \bigcup_{i\in\mathbb{N}} g[K_i]$ with compact K_i is again a countable union of compact sets and hence $\mathcal{L}([a;b])$-measurable due to 1.2. Thus $g[A] = g[A \setminus F] \cup g[F] \in \mathcal{L}([a;b])$. Due to g being **injective** the image $g[\bigcup_{i\in\mathbb{N}} A_i] = \bigcup_{i\in\mathbb{N}} g[A_i]$ of a sequence of measurable **disjoint** sets is still **disjoint** such that the **σ-additivity** of λ transfers to μ and we have obtained a **positive bounded** measure $\mu \lessdot \lambda$ on $\mathcal{L}([a;b])$. Now the **Radon-Nikodym theorem** 10.8.2 provides a derivative $\frac{d\mu}{d\lambda} \in L^1(\lambda)$ with $\mu(A) = \int_A \frac{d\mu}{d\lambda}\, d\lambda$ for all $A \in \mathcal{L}([a;b])$.

3. ⇒ 4.: According to the hypothesis for $A = [a;x]$ we have $f(x) - f(a) = g(x) - g(a) - (x-a) = \lambda(g[A]) - \lambda(A) = \mu(A) - \lambda(A) = \int_A \left(\frac{d\mu}{d\lambda} - 1\right) d\lambda = \int_a^x \frac{df}{dy}dy$ due to 13.9.

13.11 Mean value theorem for vector valued functions:: For every **differentiable** function $\boldsymbol{f} : [a;b] \to \mathbb{C}^n$ there is a $a \le t \le b$ with $\boldsymbol{f}(b) - \boldsymbol{f}(a) = \frac{d\mathbf{f}}{dy}(t) \cdot (b-a)$

Proof: Immediately follows from the **mean value theorem for integration** 5.19 with $S = \left\{\frac{1}{b-a}\int_a^x \frac{d\boldsymbol{f}}{dy}dy\right\}$ combined with the preceding **fundamental theorem of calculus** 13.10 .

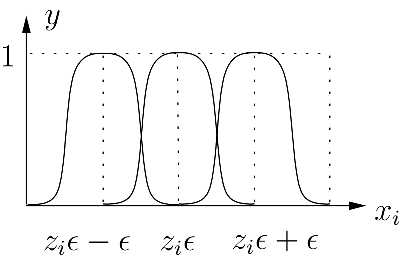
14 Differentiation in $\mathbb{R}^n$

14.1 Lemma: For $B = B_1(\mathbf{0})$ and every continuous $\mathbf{g} : \overline{B} \to \mathbb{R}^n$ with $|\mathbf{g}(\mathbf{x}) - \mathbf{x}| < \epsilon$ for all $\mathbf{x} \in \delta B$ and $0 < \epsilon < 1$ we have $B_{1-\epsilon}(\mathbf{0}) \subset \mathbf{g}[B]$.

Proof: Assuming there exists a $\mathbf{y} \in B_{1-\epsilon}(\mathbf{0}) \setminus \mathbf{g}[B]$ we also have $\mathbf{y} \notin \mathbf{g}[\delta B]$ and hence $\mathbf{y} \notin \mathbf{g}\left[\overline{B}\right]$ such that we can define a **continuous** $\mathbf{h} : \overline{B} \to \delta B$ by $\mathbf{h}(\mathbf{x}) = \frac{\mathbf{y}-\mathbf{g}(\mathbf{x})}{|\mathbf{y}-\mathbf{g}(\mathbf{x})|}$. For $\mathbf{x} \in \delta B$ we have $\mathbf{x} \cdot (\mathbf{y} - \mathbf{g}(\mathbf{x})) = \mathbf{x} \cdot \mathbf{y} + \mathbf{x} \cdot (\mathbf{x} - \mathbf{g}(\mathbf{x})) - \mathbf{x} \cdot \mathbf{x} < |\mathbf{y}| + \epsilon - 1 < 0$ and hence $\mathbf{x} \cdot \mathbf{h}(\mathbf{x}) < 0$, particularly $\mathbf{x} \neq \mathbf{h}(\mathbf{x})$ in contradiction to **Brouwer's fixed point theorem** [9, 22.11].

14.2 Differentiable functions: A function $\mathbf{f} : V \to \mathbb{R}^n$ on some **open** $U \subset \mathbb{R}^m$ is **differentiable** at $\mathbf{x} \in V$ iff there exists a **linear** $\frac{d\mathbf{f}}{d\mathbf{y}}(\mathbf{x}) : \mathbb{R}^m \to \mathbb{R}^n$ with $\lim_{\mathbf{h}\to 0} \frac{1}{|\mathbf{h}|}\left|\mathbf{f}(\mathbf{x}+\mathbf{h}) - \mathbf{f}(\mathbf{x}) - \frac{d\mathbf{f}}{d\mathbf{y}}(\mathbf{x})\,\mathbf{h}\right| = 0$. The linear map $\frac{d\mathbf{f}}{d\mathbf{y}}(\mathbf{x})$ is represented by the **Jacobian matrix** whose elements $\left(\frac{d\mathbf{f}}{d\mathbf{y}}(\mathbf{x})\right)_{ij} = \frac{\delta f_i}{\delta y_j}(\mathbf{x}) = \lim_{h\to 0} \frac{f_i(\mathbf{x}+h\mathbf{e}_j) - f_i(\mathbf{x})}{h}$ are the **partial derivatives** of the **components** in $\mathbf{f} = \sum_{i=1}^{n} f_i \mathbf{e}_i$ according to definition 12.1. The family $\mathcal{C}(U)$ contains every **continuous** $f : U \to \mathbb{R}$ and according to 11.1 the members $\mathcal{C}_c(U) \subset \mathcal{C}(U)$ have a **compact support** in U. For every $k \in \mathbb{N}$ the functions in $\mathcal{C}^k(U) \subset \mathcal{C}(U)$ have **continuous** k-th inductively defined **partial derivatives** $\frac{\delta^k f}{\delta y^k}$. Finally we may combine these properties in $\mathcal{C}_c^k(U) = \mathcal{C}_c(U) \cap \mathcal{C}^k(U)$ resp. extend to the space of **infinitely differentiable** or **smooth** functions $\mathcal{C}^\infty(U) = \bigcap_{k\in\mathbb{N}} \mathcal{C}^k(U)$.

14.3 Differentiable partition of unity: For every $\epsilon > 0$ and every $\boldsymbol{z} \in \mathbb{Z}^n$ there are $\alpha_{\epsilon\boldsymbol{z}} \in C_c^\infty(\mathbb{R}^n;\mathbb{R})$ with $\sum_{\boldsymbol{z}\in\mathbb{Z}^n} \alpha_{\epsilon\boldsymbol{z}} = 1$ and supp$(\alpha_{\epsilon\boldsymbol{z}}) = C_\epsilon(\epsilon\boldsymbol{z})$ meaning the **cube** $C_\epsilon(\epsilon\boldsymbol{z}) = \prod_{i=1}^{n} \epsilon\,[z_i - 1; z_i + 1]$.

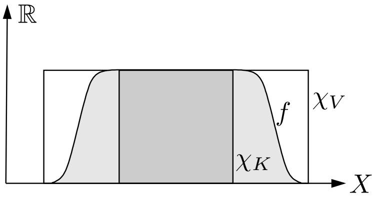


Proof: First we show by induction that the function $g : \to \mathbb{R} \to \mathbb{R}$ with $g(x) = \begin{cases} \exp\left(\frac{1}{x^2-1}\right) & \text{for } |x| < 1 \\ 0 & \text{for } |x| \geq 1 \end{cases}$ is **infinitely differentiable** with $\frac{\delta^k g}{\delta y^k}(x) = \begin{cases} \left(p_k\left(\frac{1}{x-1}\right) + q_k\left(\frac{1}{x+1}\right)\right) \cdot \exp\left(\frac{1}{x^2-1}\right) & \text{for } |x| < 1 \\ 0 & \text{for } |x| \geq 1 \end{cases}$ and **polynomials** p_k, q_k: The start $k = 0$ is trivial whence for $k \geq 0$ and $|x| < 1$ we obtain

$$\begin{aligned}\frac{\delta^{k+1}g}{\delta y^{k+1}}(x) &= \frac{\delta}{\delta y}\left(\left(p_k\left(\frac{1}{x-1}\right) + q_k\left(\frac{1}{x+1}\right)\right) \cdot \exp\left(\frac{1}{x^2-1}\right)\right) \\ &= \frac{2x}{(x-1)^{2\cdot}(x+1)^2} \cdot \left(\frac{1}{(x-1)^2} \cdot \frac{\delta p_k}{\delta y}\left(\frac{1}{x-1}\right) + \frac{1}{(x+1)^2} \cdot \frac{\delta q_k}{\delta y}\left(\frac{1}{x+1}\right)\right) \cdot \exp\left(\frac{1}{x^2-1}\right) \\ &= \left(p_{k+1}\left(\frac{1}{x-1}\right) + q_{k+1}\left(\frac{1}{x+1}\right)\right) \cdot \exp\left(\frac{1}{x^2-1}\right)\end{aligned}$$

since by **partial fraction decomposition** we have $\frac{h(x)}{(x-1)^p \cdot (x+1)^q} = \frac{h(1)}{2^q(x-1)^p} + \frac{h(-1)}{2^p(x+1)^q}$. Due to the symmetry for $|x| = 1$ we may assume $x = 1$ so as to get $\frac{\delta^{k+1}g}{\delta y^{k+1}}(1) = \lim_{h \to 0} \frac{1}{h}\left(\frac{\delta^k g}{\delta y^k}(1-h) - 0\right) = \lim_{h \to 0} \frac{1}{h}\left(p_k\left(-\frac{1}{h}\right) + q_k\left(\frac{1}{2-h}\right)\right) \cdot \exp\left(\frac{1}{h^2-2h}\right) = 0$ since $\lim_{x \to \infty} p(x) \cdot \exp(-x) = 0$ for every polynomial p. Then $h(x) = \frac{g(x)}{G(x)}$ with $G(x) = \sum_{z \in \mathbb{Z}} g(x-z)$ is infinitely differentiable with compact $\operatorname{supp}(h) = [-1;1]$ and $\sum_{z \in \mathbb{Z}} h(x-z) = 1$ for every $x \in \mathbb{R}$. Hence $\alpha_{\epsilon \boldsymbol{z}}(\boldsymbol{x}) = \prod_{i=1}^{n} h\left(\frac{x_i}{\epsilon} - z_i\right)$ is the desired smooth partition of unity.

14.4 Separation property for $\mathbb{R}^n$: For **compact** resp. **open** $K \subset U \subset \mathbb{R}^n$ we can improve on 11.1 and provide a **smooth separating function** $f \in C_c^\infty(\mathbb{R}^n;\mathbb{R})$ with $K \prec f \prec U$ by $f_{K;U} = \sum_{\boldsymbol{z} \in \mathbb{Z}^n} \chi_K(\epsilon \boldsymbol{z}) \cdot \alpha_{\epsilon \boldsymbol{z}}$ with $\epsilon = \frac{1}{M}$ and the smallest common denominator $M \in \mathbb{N}$ of the **rational** $\epsilon_j \in \mathbb{Q}$ such that $K \subset \bigcup_{j=1}^{R} C_{\epsilon_j}(\epsilon_j \boldsymbol{z}) \subset U$ and every $C_{\epsilon_j}(\epsilon_j \boldsymbol{z})$ is the union of **finitely** many $C_\epsilon(\epsilon \boldsymbol{z})$.



14.5 Theorem:

1. $C_c^\infty(\mathbb{R}^n;\mathbb{C})$ is **dense** in $\mathcal{C}_c(\mathbb{R}^n;\mathbb{C})$ with regard to the **supremum norm**.
2. $C_c^\infty(\mathbb{R}^n;\mathbb{C})$ is dense in $L^p(\mathbb{R}^n;\mathbb{C})$ with regard to the L^p**- norm**.

Proof:

1. Since f is **uniformly continnuous** on K or every $f \in \mathcal{C}_c(\mathbb{R}^n;\mathbb{C})$ and $\eta > 0$ there is an $\epsilon > 0$ such that $|f(\boldsymbol{x}) - f(\epsilon \boldsymbol{z})| < \frac{\eta}{2^n}$ for every $\boldsymbol{x} \in C_\epsilon(\epsilon \boldsymbol{z})$. Any $x \in \mathbb{R}^n$ meets at most 2^n adjacent cubes $C_\epsilon(\epsilon \boldsymbol{z}) = \operatorname{supp}(\alpha_{\epsilon z})$ such that
$$\begin{aligned}|f(\boldsymbol{x}) - f_\epsilon(\boldsymbol{x})| &= \left|\sum_{\boldsymbol{z} \in \mathbb{Z}^n} f(\boldsymbol{x}) \cdot \alpha_{\epsilon \boldsymbol{z}}(\boldsymbol{x}) - \sum_{\boldsymbol{z} \in \mathbb{Z}^n} f(\epsilon \boldsymbol{z}) \cdot \alpha_{\epsilon \boldsymbol{z}}(\boldsymbol{x})\right| \\ &= \left|\sum_{\boldsymbol{z} \in J} f(\boldsymbol{x}) \cdot \alpha_{\epsilon \boldsymbol{z}}(\boldsymbol{x}) - \sum_{\boldsymbol{z} \in J} f(\epsilon \boldsymbol{z}) \cdot \alpha_{\epsilon \boldsymbol{z}}(\boldsymbol{x})\right| \\ &\leq \sum_{\boldsymbol{z} \in J} \alpha_{\epsilon \boldsymbol{z}}(\boldsymbol{x}) \cdot |f(\boldsymbol{x}) - f(\epsilon \boldsymbol{z})| \\ &< \eta\end{aligned}$$
with $f_\epsilon = \sum_{\boldsymbol{z} \in \mathbb{Z}^n} f(\epsilon \boldsymbol{z}) \cdot \alpha_{\epsilon \boldsymbol{z}}$ and $J = \{\boldsymbol{z} \in \mathbb{Z}^n : x \in C_\epsilon(\epsilon \boldsymbol{z})\}$.
2. Follows from 11.2 resp. 1. since for any $f \in \mathcal{C}_c(\mathbb{R}^n;\mathbb{C})$ with compact support K we have $\|f\|_p \leq \lambda(K) \cdot \|f\|_\infty$.

14.6 Product rule: For **differentiable** $\mathbf{f}, \boldsymbol{g} : U \to \mathbb{R}^n$ on some **open** $U \subset \mathbb{R}^m$ with a **continuous bilinear** map $\cdot : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^p$ the **product** $\boldsymbol{f} \cdot \boldsymbol{g}$ is again differentiable with $\frac{d\mathbf{f}\cdot\boldsymbol{g}}{d\mathbf{y}}(\mathrm{x}) = \frac{d\mathbf{f}}{d\mathbf{y}}(\mathrm{x}) \cdot \boldsymbol{g}(\boldsymbol{x}) + \boldsymbol{f}(\boldsymbol{x}) \cdot \frac{d\mathbf{g}}{d\mathbf{y}}(\mathrm{x})$ meaning the **linear map** $\frac{d\mathbf{f}\cdot\boldsymbol{g}}{d\mathbf{y}}(\mathrm{x}) : \mathbb{R}^m \to \mathbb{R}^n$ with $\frac{d\mathbf{f}\cdot\boldsymbol{g}}{d\mathbf{y}}(\mathrm{x})\,\boldsymbol{y} = \frac{d\mathbf{f}}{d\mathbf{y}}(\mathrm{x})\,\boldsymbol{y} \cdot \boldsymbol{g}(\boldsymbol{x}) + \boldsymbol{f}(\boldsymbol{x}) \cdot \frac{d\mathbf{g}}{d\mathbf{y}}(\mathrm{x})\,\boldsymbol{y}$.

Proof:

$$\boldsymbol{f}(\boldsymbol{x}+\boldsymbol{h}) \cdot \boldsymbol{g}(\boldsymbol{x}+\boldsymbol{h}) - \boldsymbol{f}(\boldsymbol{x}) \cdot \boldsymbol{g}(\boldsymbol{x})$$

$$= \boldsymbol{f}(\boldsymbol{x}+\boldsymbol{h}) \cdot \boldsymbol{g}(\boldsymbol{x}+\boldsymbol{h}) - \boldsymbol{f}(\boldsymbol{x}) \cdot \boldsymbol{g}(\boldsymbol{x}+\boldsymbol{h}) + \boldsymbol{f}(\boldsymbol{x}) \cdot \boldsymbol{g}(\boldsymbol{x}+\boldsymbol{h}) - \boldsymbol{f}(\boldsymbol{x}) \cdot \boldsymbol{g}(\boldsymbol{x})$$

$$= [\boldsymbol{f}(\boldsymbol{x}+\boldsymbol{h}) - \boldsymbol{f}(\boldsymbol{x})] \cdot \boldsymbol{g}(\boldsymbol{x}+\boldsymbol{h}) + \boldsymbol{f}(\boldsymbol{x}) \cdot [\boldsymbol{g}(\boldsymbol{x}+\boldsymbol{h}) - \boldsymbol{g}(\boldsymbol{x})]$$

$$= \left[\frac{d\mathbf{f}}{d\mathbf{y}}(\mathrm{x})\,\boldsymbol{h} + |\boldsymbol{h}|\,\varphi_f(\boldsymbol{h})\right] \cdot \boldsymbol{g}(\boldsymbol{x}+\boldsymbol{h}) + \boldsymbol{f}(\boldsymbol{x}) \cdot \left[\frac{d\mathbf{g}}{d\mathbf{y}}(\mathrm{x})\,\boldsymbol{h} + |\boldsymbol{h}|\,\varphi_g(\boldsymbol{h})\right]$$

$$= \frac{d\mathbf{f}}{d\mathbf{y}}(\mathrm{x})\,\boldsymbol{h} \cdot \boldsymbol{g}(\boldsymbol{x}) + \boldsymbol{f}(\boldsymbol{x}) \cdot \frac{d\mathbf{g}}{d\mathbf{y}}(\mathrm{x})\,\boldsymbol{h} + |\boldsymbol{h}| \left[\frac{d\mathbf{f}}{d\mathbf{y}}(\mathrm{x})\,\frac{\boldsymbol{h}}{|\boldsymbol{h}|} \cdot [\boldsymbol{g}(\boldsymbol{x}+\boldsymbol{h}) - \boldsymbol{g}(\boldsymbol{x})] + \varphi_f(\boldsymbol{h}) \cdot \boldsymbol{g}(\boldsymbol{x}+\boldsymbol{h}) + \boldsymbol{f}(\boldsymbol{x}) \cdot \varphi_g(\boldsymbol{h})\right].$$

with $\varphi_f = \mathbf{f}(\mathbf{x}+\mathbf{h}) - \mathbf{f}(\mathbf{x}) - \frac{d\mathbf{f}}{d\mathbf{y}}(\mathrm{x})\,\mathbf{h}$ resp. $\varphi_g = \mathbf{g}(\mathbf{x}+\mathbf{h}) - \mathbf{g}(\mathbf{x}) - \frac{d\mathbf{g}}{d\mathbf{y}}(\mathrm{x})\,\mathbf{h}$.

Since $\boldsymbol{g}$ is **continuous** for $|\boldsymbol{h}| \to 0$ the last bracket vanishes and the formula is proven.

14.7 Integration by parts: For $f \in \mathcal{C}^1(U)$, $g \in \mathcal{C}^1_c(U)$ and $1 \le i \le n$ we have $\int_U \frac{\delta f}{\delta x_i}(\boldsymbol{x}) \cdot g(\boldsymbol{x})\,d\boldsymbol{x} = -\int_U f(\boldsymbol{x}) \cdot \frac{\delta g}{\delta x_i}(\boldsymbol{x})\,d\boldsymbol{x}$.

Proof: Every $h \in \mathcal{C}^1_c(U)$ can be extended to $\mathbb{R}^n$ by $h(x) = 0$ for every $x \in \mathbb{R}^n \setminus \operatorname{supp}(h)$. Due to the **Heine Borel theorem** [9, 9.10] there is an $R > 0$ with $\operatorname{supp}(h) \subset [-R; R]^n$ so that the **fundamental theorem of calculus** 13.10 yields $\int_{-R}^{R} \frac{\delta h}{\delta x_i}(\boldsymbol{x})\,dx_i = h(\boldsymbol{x} + \boldsymbol{e}_i(R - x_i)) - h(\boldsymbol{x} - \boldsymbol{e}_i(R + x_i)) = 0 - 0 = 0$ and with **Fubini's theorem** 8.5 we infer $\int_U \frac{\delta h}{\delta x_i}(\boldsymbol{x})\,d\boldsymbol{x} = \int_{[-R;R]} \frac{\delta h}{\delta x_i}(\boldsymbol{x})\,d\boldsymbol{x} = 0$. The **product rule** 14.6 applied to $h = f \cdot g \in \mathcal{C}^1_c(U)$ provides $\frac{\delta h}{\delta x_i}(\boldsymbol{x}) = \frac{\delta f}{\delta x_i}(\boldsymbol{x}) \cdot g(\boldsymbol{x}) + f(\boldsymbol{x}) \cdot \frac{\delta g}{\delta x_i}(\boldsymbol{x})$ and hence the assertion.

14.8 Chain rule: For **differentiable** $\mathbf{g} : U \to V$ resp. $\mathbf{f} : V \to \mathbb{R}^p$ on some **open** $U \subset \mathbb{R}^m, V \subset \mathbb{R}^n$ the **composition** $\boldsymbol{f} \circ \boldsymbol{g}$ is again differentiable with $\frac{d\mathbf{f}\circ\mathbf{g}}{d\mathbf{y}}(\mathrm{x}) = \frac{d\mathbf{f}}{d\mathbf{g}}(\boldsymbol{g}(\boldsymbol{x})) \circ \frac{d\mathbf{g}}{d\mathbf{y}}(\mathrm{x})$

Proof: $\boldsymbol{f}(\boldsymbol{g}(\boldsymbol{x}+\boldsymbol{h})) - \boldsymbol{f}(\boldsymbol{g}(\boldsymbol{x}))$

$$= \frac{d\mathbf{f}}{d\mathbf{g}}(\boldsymbol{g}(\boldsymbol{x}))\,\boldsymbol{k}(\boldsymbol{h}) + |\boldsymbol{k}(\boldsymbol{h})|\,\varphi_f(\boldsymbol{k}(\boldsymbol{h}))$$

$$= \frac{d\mathbf{f}}{d\mathbf{g}}(\boldsymbol{g}(\boldsymbol{x})) \circ \frac{d\mathbf{g}}{d\mathbf{y}}(\mathrm{x})\,\boldsymbol{h} + \frac{d\mathbf{f}}{d\mathbf{g}}(\boldsymbol{g}(\boldsymbol{x}))\,|\boldsymbol{h}|\,\varphi_g(\boldsymbol{h}) + |\boldsymbol{k}(\boldsymbol{h})|\,\varphi_f(\boldsymbol{k}(\boldsymbol{h}))$$

$$= \frac{d\mathbf{f}}{d\mathbf{g}}(\boldsymbol{g}(\boldsymbol{x})) \circ \frac{d\mathbf{g}}{d\mathbf{y}}(\mathrm{x})\,\boldsymbol{h} + |\boldsymbol{h}| \left[\frac{d\mathbf{f}}{d\mathbf{g}}(\boldsymbol{g}(\boldsymbol{x}))\,\varphi_g(\boldsymbol{h}) + \left(\frac{d\mathbf{g}}{d\mathbf{y}}(\mathrm{x})\,\frac{\boldsymbol{h}}{|\boldsymbol{h}|} + \varphi_g(\boldsymbol{h})\right)\varphi_f(\boldsymbol{k}(\boldsymbol{h}))\right]$$

with $\boldsymbol{k}(\boldsymbol{h}) = \boldsymbol{g}(\boldsymbol{x}+\boldsymbol{h}) - \boldsymbol{g}(\boldsymbol{x}) = \frac{d\mathbf{g}}{d\mathbf{y}}(\mathrm{x})\,\boldsymbol{h} + |\boldsymbol{h}|\,\varphi_g(\boldsymbol{h}) = |\boldsymbol{h}|\left(\frac{d\mathbf{g}}{d\mathbf{y}}(\mathrm{x})\,\frac{\boldsymbol{h}}{|\boldsymbol{h}|} + \varphi_g(\boldsymbol{h})\right)$**.**

For $|\boldsymbol{h}| \to 0$ we have $\varphi_g(\boldsymbol{h}) \to 0$ resp. $\boldsymbol{k}(\boldsymbol{h}) \to 0$. Since for $|\boldsymbol{k}| \to 0$ we have $\varphi_f(\boldsymbol{k}) \to 0$ the bracket vanishes and the formula is proven.

14.9 Mean value theorem for vector spaces: For every **continuously differentiable** $\boldsymbol{f} : U \to \mathbb{R}^n$ on an **open** and **convex** set $U \subset \mathbb{R}^m$ such that the straight path $\{(1-t) \cdot \boldsymbol{a} + t \cdot \boldsymbol{b} : 0 \le t \le 1\} \subset U$ there is a $0 \le t \le 1$ with $\boldsymbol{f}(\boldsymbol{b}) - \boldsymbol{f}(\boldsymbol{a}) = \frac{d\boldsymbol{f}}{d\boldsymbol{y}}((1-t) \cdot \boldsymbol{a} + t \cdot \boldsymbol{b}) \cdot (\boldsymbol{b} - \boldsymbol{a})$.

Proof: We combine the **mean value theorem** 13.11 for vector valued functions with the **chain rule** 14.8 for $\boldsymbol{g} : [0; 1] \to \mathbb{R}^m$ with $g(t) = (1-t) \cdot \boldsymbol{a} + t \cdot \boldsymbol{b}$ to obtain $\boldsymbol{f}(\boldsymbol{b}) - \boldsymbol{f}(\boldsymbol{a}) = \boldsymbol{f}(\boldsymbol{g}(1)) - \boldsymbol{f}(\boldsymbol{g}(0)) = \frac{d\mathbf{f}\circ\boldsymbol{g}}{dy}(t) = \frac{d\mathbf{f}}{d\mathbf{g}}(\boldsymbol{g}(t)) \circ \frac{d\mathbf{g}}{dy}(t) = \frac{d\boldsymbol{f}}{d\boldsymbol{y}}((1-t) \cdot \boldsymbol{a} + t \cdot \boldsymbol{b}) \cdot (\boldsymbol{b} - \boldsymbol{a})$for some $0 \le t \le 1$.

14.10 Definition and theorem: If for some **open** $V \subset \mathbb{R}^n$ the function $\mathbf{g} : V \to \mathbb{R}^n$ is **differentiable** at $\mathbf{x} \in V$ then $\lim\limits_{r \to 0} \frac{\lambda^n(\mathbf{g}[B_r(\mathbf{x})])}{\lambda^n(B_r(\mathbf{x}))} = \left|\det\left(\frac{d\mathbf{g}}{d\mathbf{y}}(\mathbf{x})\right)\right|$.

Proof:

Case I: $\frac{d\mathbf{g}}{d\mathbf{y}}(\mathrm{x}) \in GL(n; \mathbb{R})$. W.l.o.g. we can assume $\mathbf{g}(\mathbf{x}) = \mathbf{x} = \mathbf{0}$. The function $\mathbf{h} = \left(\frac{d\mathbf{g}}{d\mathbf{y}}(\mathbf{0})\right)^{-1} * \mathbf{g}$ obviously is differentiable with $\mathbf{h}(\mathbf{0}) = \mathbf{0}$ and $\frac{d\mathbf{h}}{d\mathbf{y}}(\mathbf{0}) = \mathrm{id}$ such that for $\epsilon > 0$ there is a $\delta > 0$

with $|\mathbf{g}(\mathbf{y})-\mathbf{y}| < \epsilon|\mathbf{y}|$ for every $\mathbf{y} \in B_\delta(\mathbf{0})$. The preceding lemma applied to $\mathbf{h}_r : \overline{B} \to \mathbb{R}^n$ with $\mathbf{h}_r(\mathbf{y}) = \frac{1}{r}\mathbf{h}(r \cdot \mathbf{y})$ with $|\mathbf{h}(\mathbf{y})-\mathbf{y}| < \epsilon r$ for all $\mathbf{y} \in \delta B_r(0)$ with $0 < r < \delta$ resp. $\left|\frac{1}{r}\mathbf{h}\left(r \cdot \frac{\mathbf{y}}{r}\right) - \frac{\mathbf{y}}{r}\right| < \epsilon$ for all all $\mathbf{y} \in \delta B$ implies $B_{1-\epsilon}(\mathbf{0}) \subset \frac{1}{r}\mathbf{h}[r \cdot B]$. Hence $B_{(1-\epsilon)r}(\mathbf{0}) \subset \mathbf{h}[B_r(\mathbf{0})] \subset B_{(1+\epsilon)r}(\mathbf{0})$ whence with 8.9 follows $(1-\epsilon)^n \leq \frac{\lambda^n(\mathbf{h}[B_r(\mathbf{0})])}{\lambda^n(B_r(\mathbf{0}))} \leq (1+\epsilon)^n$. Thus we have $\lim\limits_{r\to 0}\frac{\lambda^n(\mathbf{h}[B_r(\mathbf{0})])}{\lambda^n(B_r(\mathbf{0}))} = 1$ and the assertion follows from 8.9 with $\lambda^n(\mathbf{h}[B_r(\mathbf{0})]) = \lambda^n\left(\left(\frac{d\mathbf{g}}{d\mathbf{y}}(\mathbf{0})\right)^{-1} * \mathbf{g}[B_r(\mathbf{0})]\right) = \frac{\lambda^n(\mathbf{g}[B_r(\mathbf{0})])}{\left|\det\left(\frac{d\mathbf{g}}{d\mathbf{y}}(\mathbf{0})\right)\right|}$.

Case II: $\frac{d\mathbf{f}}{d\mathbf{y}}(\mathbf{x}) \notin GL(n;\mathbb{R})$. Again we can assume $\mathbf{f}(\mathbf{x}) = \mathbf{x} = \mathbf{0}$. Since $\dim\left(\frac{d\mathbf{f}}{d\mathbf{y}}(\mathbf{0}) * V\right) < n$ we have $\lambda^n\left(\frac{d\mathbf{f}}{d\mathbf{y}}(\mathbf{0}) * V\right) = 0$ and due to the **continuity from above** 2.2.3 for every $\epsilon > 0$ there is a $\delta > 0$ such that $\lambda^n\left(U_\delta\left(\frac{d\mathbf{f}}{d\mathbf{y}}(\mathbf{0}) * B_1(\mathbf{0})\right)\right) < \epsilon$ for the δ-neighbourhood $U_\delta(A) = \{\mathbf{y} \in V : d(\mathbf{y};A) < \delta\}$. (cf. [9, 11.8]). Furthermore there is an $\eta > 0$ such that $\left|\mathbf{f}(\mathbf{y}) - \frac{d\mathbf{f}}{d\mathbf{y}}(\mathbf{0}) * \mathbf{y}\right| < \delta|\mathbf{y}|$ for all $\mathbf{y} \in B_\eta(0)$ and hence $\mathbf{f}[B_r(\mathbf{0})] \subset U_{\delta r}\left(\frac{d\mathbf{f}}{d\mathbf{y}}(\mathbf{0}) * B_r(\mathbf{0})\right)$ for every $0 < r < \eta$. Thus we have $\lambda^n(\mathbf{f}[B_r(\mathbf{0})]) \leq \lambda^n\left(U_{\delta r}\left(\frac{d\mathbf{f}}{d\mathbf{y}}(\mathbf{0}) * B_r(\mathbf{0})\right)\right) = \lambda^n\left(S * U_\delta\left(\frac{d\mathbf{f}}{d\mathbf{y}}(\mathbf{0}) * B_1(\mathbf{0})\right)\right) < \epsilon r^n$ with the **dilation** S defined by $S * \mathbf{e}_i = r\mathbf{e}_i$ for $0 \leq i \leq n$ according to 8.9. Since $r^n = \lambda^n(B_r(\mathbf{0}))$ we conclude that $\lim\limits_{r\to 0}\frac{\lambda^n(\mathbf{f}[B_r(\mathbf{0})])}{\lambda^n(B_r(\mathbf{0}))} = 0$.

14.11 Lemma: For every **nonempty open** $V \subset \mathbb{R}^n$ exists

1. a **decomposition** $V = \mathring{\bigcup}_{i\in\mathbb{N}}[\mathbf{a}_i;\mathbf{b}_i[$ of **disjoint intervals** $I_i = [\mathbf{a}_i;\mathbf{b}_i[= \prod\limits_{k=1}^{n}[a_{ik};b_{ik}[\in \mathcal{I}^n$.
2. a **covering** $V \subset \bigcup_{i\in\mathbb{N}} B_{r_i}(\mathbf{x}_i)$ of **open** balls such that $\sum_{i\in\mathbb{N}}\lambda^n(B_{r_i}(\mathbf{x}_i)) < n^{n/2} \cdot \lambda^n(V)$.

Proof:

1. For every $m \geq 1$ let
$$P_m = \left\{\mathbf{x} = \sum_{k=1}^{n} x_k\mathbf{e}_k \in \mathbb{R}^n : x_k = \frac{z}{2^m}; z \in \mathbb{Z}; 1 \leq k \leq n\right\}$$
and
$$\Omega_m = \left\{\left[\mathbf{a}_i;\mathbf{a}_i + \sum_{k=1}^{n}\frac{\mathbf{e}_k}{2^m}\right[: \mathbf{a}_i \in P_m\right\}$$
Obviously the intervals in Ω_m form a **disjoint decomposition** of $\mathbb{R}^n = \mathring{\bigcup}\Omega_m$ and for $m < r$ with $I_m \in \Omega_m$; $I_r \in \Omega_r$ and $I_1 \cap I_2 \neq \emptyset$ we have $I_m \subset I_r$. These properties allow us to apply the following selection process in order to find the desired decomposition: Since every $x \in V$ is contained in an open ball lying in V there is an $m \geq 1$ and an interval $I \in \Omega_m$ with $x \in I \subset V$ such that $V = \bigcup\left(\bigcup_{m\geq 1}\Omega_m^V\right)$ with $\Omega_m^V = \{I \in \Omega_m : I \subset V\}$. From this collection of intervals starting with Ω_1 subsequently for $m \geq 1$ remove all intervals $J \in \bigcup_{l>m}^V$ which are included in any $I \in \Omega_m^V$. The remaining intervals form the desired disjoint decomposition.
2. For every $x \in \mathbb{R}^n$ and $r > 0$ we have
$$\left[\mathbf{x} - \frac{r}{\sqrt{n}}\sum_{k=1}^{n}\mathbf{e}_k;\mathbf{x} + \frac{r}{\sqrt{n}}\sum_{k=1}^{n}\mathbf{e}_k\right[\subset B_r(\mathbf{x}) \subset \left[\mathbf{x} - r\sum_{k=1}^{n}\mathbf{e}_k;\mathbf{x} + r\sum_{k=1}^{n}\mathbf{e}_k\right[$$
and in particular $\left(\frac{2r}{\sqrt{n}}\right)^n < \lambda^n(B_r(\mathbf{x})) < (2r)^n$ such that every $I = \left[\mathbf{a};\mathbf{a} + \sum\limits_{k=1}^{n} 2r\mathbf{e}_k\right[\in \Omega_m$ there is an open ball $B = B_r\left(\mathbf{a} + \sum\limits_{k=1}^{n} r\mathbf{e}_k\right)$ with $I \subset B$ and $\lambda^n(B) \leq n^{n/2}\lambda^n(I)$ whence follows the assertion.

14.12 Lemma: For a function $\mathbf{g} : V \to \mathbb{R}^n$ being **differentiable** at every $\mathbf{x} \in V \subset \mathbb{R}^n$ for an open V for every $A \subset V$ we have $\lambda^n(A) = 0 \Rightarrow \lambda^n(\mathbf{g}[A]) = 0$.

Proof: For $m;p \geq 1$ let $C_{m;p} = \left\{\mathbf{x} \in A : \frac{|\mathbf{g}(\mathbf{y})-\mathbf{g}(\mathbf{x})|}{|\mathbf{y}-\mathbf{x}|} \leq m \,\forall \mathbf{y} \in B_{1/p}(\mathbf{x}) \cap A\right\}$. Since $\lambda^n(C_{m;p}) = 0$ and λ^n is **regular** for every $\epsilon > 0$ there is an open $V \supset C_{m;p}$ with $\lambda^n(V) < n^{-n/2}\epsilon$. Due to the preceding lemma V and especially $C_{m;p}$ can be covered by open balls $\bigcup_{i\in\mathbb{N}} B_{r_i}(\mathbf{x}_i)$ with $r_i < \frac{1}{p}$ and $\sum_{i\in\mathbb{N}} \lambda^n (B_{r_i}(\mathbf{x}_i)) < \epsilon$. For $\mathbf{x} \in C_{m;p} \cap B_{r_i}(\mathbf{x}_i)$ follows $|\mathbf{x}_i - \mathbf{x}| < r_i < \frac{1}{p}$ and $\mathbf{x}_i \in C_{m;p}$. Hence $|\mathbf{g}(\mathbf{x}_i) - \mathbf{g}(\mathbf{x})| \leq m \cdot |\mathbf{x}_i - \mathbf{x}| < mr_i$ so that $\mathbf{g}[C_{m;p} \cap B_{r_i}(\mathbf{x}_i)] \subset B_{mr_i}(\mathbf{g}(\mathbf{x}_i))$. Therefore we have $\mathbf{g}[C_{m;p}] \subset \bigcup_{i\in\mathbb{N}} B_{mr_i}(\mathbf{g}(\mathbf{x}_i))$ and hence $\lambda^n(\mathbf{g}[C_{m;p}]) \leq \sum_{i\in\mathbb{N}} \lambda^n(B_{mr_i}(\mathbf{g}(\mathbf{x}_i))) = m^n \cdot \sum_{i\in\mathbb{N}} \lambda^n(B_{r_i}(\mathbf{x}_i)) < m^n \cdot \epsilon$. Since λ^n is **complete** on $\mathcal{L}(V)$ and ϵ was arbitrary, $\mathbf{g}[C_{m;p}]$ is **Lebesgue measurable** and $\lambda^n(\mathbf{g}[C_{m;p}]) = 0$. The assertion now follows from $A = \bigcup_{m;p\geq 1} C_{m;p}$ and λ^n being **continuous from below 2.2.2.**

14.13 Change-of-variables theorem: For every

- **continuous** $\mathbf{g} : V \to \mathbb{R}^n$ on an **open** $V \subset \mathbb{R}^n$ being **bijective** and **differentiable** on a **Lebesgue measurable** subset $A \in \mathcal{B}(V)$ with $\lambda^n(\mathbf{g}[V \setminus A]) = 0$ and
- **integrable** $\boldsymbol{f} : \boldsymbol{g}[A] \to Y$ into a **Banach space** $(Y; |\,|)$ we have

$$\int_{\mathbf{g}[A]} \boldsymbol{f}(\mathbf{g})\, d\mathbf{g} = \int_A \boldsymbol{f}(\mathbf{g}(\mathbf{y})) \cdot \left|\det\left(\frac{d\mathbf{g}}{d\mathbf{y}}\right)\right| d\mathbf{y}$$

with both $d\mathbf{y}$ resp $d\mathbf{g}$ denoting the **same** Lebesgue measure λ^n_V applied to the **original** set A resp. to the **image** $\mathbf{g}[A]$. This intuitive notation suggests the **chain rule** from differentiation which actually is a special case of this theorem.

Proof: The **restriction** to the Lebesgue measurable subset A is for practical reasons because coordinate transformations like the subsequent examples 14.14 may only be defined on a correspondingly smaller domain. **Step I** depends on the preceding Lemma 14.12 and hence requires the **Lebesgue sets** $\mathcal{L}(V)$ on V: The image $\mathbf{g}[B]$ of a Borel set B is **Lebesgue** measurable but it may not be **Borel** measurable. In the final **step IV** Theorem 5.2 justifies the extension of the formula from **step functions** to **Borel** measurable functions but not to **Lebesgue** measurable ones. Hence the integration on the left side of a **Borel** measurable function f over a **Lebesgue** measurable set $\mathbf{g}[A]$ is only possible since according to 11.13 every **Lebesgue** measurable set in $\mathcal{L}(V)$ is the union of a **Borel** measurable set and a λ^n_V**-null set**. This fact is also utilized in **steps I** and **III.**

Step I. For every $B \in \mathcal{L}(V)$ we have $\mathbf{g}[B] \in \mathcal{L}(V)$: For every λ^n_V**-null set** B_0 the hypothesis yields $\lambda^n_V(\mathbf{g}[B_0 \setminus A]) \leq \lambda^n_V(\mathbf{g}[V \setminus A]) = 0$ whereas from 14.12 we can infer $\lambda^n_V(\mathbf{g}[B_0 \cap A]) = 0$ and since $B_0 \subset (V \setminus A) \cup (B_0 \cap A)$ we conclude $\mathbf{g}[B_0] \in \mathcal{L}(V)$ since λ^n_V is **complete** on $\mathcal{L}(V)$. For a σ**-compact** $B_1 \in \mathcal{L}(V)$ its **continuous** and **bijective** image $\mathbf{g}[B_1]$ is again σ**-compact** such that $\mathbf{g}[B_1] \in \mathcal{L}(V)$. The assertion now follows from 11.13 since every $B \in \mathcal{L}(V)$ is the union of a λ^n_V- null set B_0 and an F_σ-set B_1.

Step II. For every $B \in \mathcal{L}(V)$ we have $\lambda^n_V(\mathbf{g}(A \cap B)) = \int_{A\cap B} \left|\det\left(\frac{d\mathbf{g}}{d\mathbf{y}}\right)\right| d\mathbf{y}$: Because of **step I** for $B \in \mathcal{L}(V)$ the expression $\mu_m(B) = \lambda^n_V(\mathbf{g}(A_m \cap B))$ with $A_m = \{\mathbf{x} \in A : \|\mathbf{g}(\mathbf{x})\| < m\}$ is well defined and σ**-additive** since g is **bijective** on $A_m \subset A$ for every $m \geq 1$. Hence μ_m is a **bounded** and due to 14.12 λ^n_V**-absolutely continuous** measure on $\mathcal{L}(V)$. The **Radon-Nikodym theorem** 10.8.2 assures that $\frac{d\mu_m}{d\lambda^n} \in L^1(\lambda^n_V)$ exists λ^n_V-a.e. with $\mu_m(B) = \int_B \frac{d\mu_m}{d\lambda^n_V} d\lambda^n_V$. For every $\mathbf{x} \in A_m$ there is an $r > 0$ with $B_r(\mathbf{x}) \subset V_m = \{\mathbf{x} \in V : \|\mathbf{g}(\mathbf{x})\| < m\}$. Since $V_m \setminus A_m \subset V \setminus A$ we have

$$\begin{aligned}\mu_m(B_r(\mathbf{x})) &= \lambda^n_V(\mathbf{g}(B_r(\mathbf{x}) \cap A_m)) \\ &= \lambda^n_V(\mathbf{g}(B_r(\mathbf{x}) \cap V_m)) - \lambda^n_V(\mathbf{g}(B_r(\mathbf{x}) \cap (V_m \setminus A_m))) \\ &= \lambda^n_V(\mathbf{g}(B_r(\mathbf{x}))) - 0\end{aligned}$$

according to 8.8. From 14.10 follows $\frac{\mu_m(B_r(\mathbf{x}))}{\lambda^n(B_r(\mathbf{x}))} = \frac{\lambda^n(\mathbf{g}(B_r(\mathbf{x})))}{\lambda^n(B_r(\mathbf{x}))} = \left|\det\left(\frac{d\mathbf{g}}{d\mathbf{y}}\right)(\mathbf{x})\right|$ and with $r \to 0$ we conclude that $\frac{d\mu_m}{d\lambda^n}(\mathbf{x}) = \left|\det\left(\frac{d\mathbf{g}}{d\mathbf{y}}\right)(\mathbf{x})\right|$ due to 13.2 and 13.6. The definition of μ_m implies $\mu_m(A_m \cap B) = \mu_m(B)$ and hence $\mu_m(A_m \cap B) = \int_{A_m\cap B} \frac{d\mu_m}{d\lambda^n_V} d\lambda^n_V = \int_{A_m\cap B} \left|\det\left(\frac{d\mathbf{g}}{d\mathbf{y}}\right)\right| d\mathbf{y}$ for every $m \geq 1$. Finally for $m \to \infty$ the **continuity of** λ^n_V **from below** resp. the **bijectivity** of $\mathbf{g}$ on the left hand side resp. the **monotone convergence theorem** 5.12 on the right hand side yield

$$\begin{aligned}\lambda_V^n\left(\mathbf{g}\left(A\cap B\right)\right) &= \lambda_V^n\left(\mathbf{g}\left(\left(\bigcup_{m\geq 1} A_m\right)\cap B\right)\right)\\ &= \lambda_V^n\left(\bigcup_{m\geq 1}\mathbf{g}\left(A_m\cap B\right)\right)\\ &= \sup_{m\geq 1}\lambda_V^n\left(\mathbf{g}\left(A_m\cap B\right)\right)\\ &= \sup_{m\geq 1}\int\limits_{A_m\cap B}\left|\det\left(\frac{d\mathbf{g}}{d\mathbf{y}}\right)\right|d\mathbf{y}\\ &= \int\limits_B \sup_{m\geq 1}\chi_{A_m}\left|\det\left(\frac{d\mathbf{g}}{d\mathbf{y}}\right)\right|d\mathbf{y}\\ &= \int\limits_{A\cap B}\left|\det\left(\frac{d\mathbf{g}}{d\mathbf{y}}\right)\right|d\mathbf{y}.\end{aligned}$$

Step III: For every $B \in \mathcal{L}(V)$ we have $\int_{g[A]}\chi_B d\mathbf{y} = \int_A (\chi_B \circ \mathbf{g})\left|\det\left(\frac{d\mathbf{g}}{d\mathbf{y}}\right)\right| d\mathbf{y}$: On the one hand for every **Lebesgue** set $B \in \mathcal{L}(V)$ its inverse image $\mathbf{g}^{-1}[B] \in \mathcal{L}(V)$ is **Lebesgue** measurable and $\mathbf{g}\left[A\cap \mathbf{g}^{-1}[B]\right] = \mathbf{g}[A]\cap B$ since $\mathbf{g}$ is **bijective** on A and **continuous** on V. Hence **step II** with $d\mathbf{g} = d\mathbf{y} = d\lambda_V^n$ implies

$$\begin{aligned}\int\limits_{g[A]}\chi_B d\mathbf{g} &= \int\limits_{\mathbf{g}[A]\cap B} d\mathbf{g}\\ &= \lambda_V^n\left(\mathbf{g}[A]\cap B\right)\\ &= \lambda_V^n\left(\mathbf{g}\left[A\cap\mathbf{g}^{-1}[B]\right]\right)\\ &= \int_{A\cap\mathbf{g}^{-1}[B]}\left|\det\left(\frac{d\mathbf{g}}{d\mathbf{y}}\right)\right|d\mathbf{y}\\ &= \int_A \chi_{\mathbf{g}^{-1}[B]}\left|\det\left(\frac{d\mathbf{g}}{d\mathbf{y}}\right)\right|d\mathbf{y}\\ &= \int_A \chi_B\circ\mathbf{g}\left|\det\left(\frac{d\mathbf{g}}{d\mathbf{y}}\right)\right|d\mathbf{y}\end{aligned}$$

On the other hand for every λ_V^n**-null set** N there is a **Borel** set $B \supset N$ with $\lambda_V^n(B) = 0$ such that both sides of the above equation vanish for B and consequently for N since λ_V^n is **complete** on $\mathcal{L}(V)$. Owing to 11.13 every Lebesgue measurable set is the disjoint union of a Borel set and a λ^n-null set so that due to the additivity of the integral the proposition holds for every $B \in \mathcal{L}(V)$. In the final step we only need **Borel** measurable $B \in \mathcal{B}(V)$ but the image $\mathbf{g}[A]$ may not be **Borel** measurable so that we have to extend the argument to **Lebesgue** sets $\mathcal{L}(V)$.

Step IV: The assertion from **step III** obviously extends to every **step function** and by theorem 5.5 to every **integrable** $\boldsymbol{f} : \boldsymbol{g}[A] \to Y$.

Note: As already mentioned in 13.2 and 13.10 the absolute value in the formula indicates the **loss of information on direction** in integral on sets in vector spaces devoid of any **linear order**.

14.14 The closed cylinder $C_{h;R} = \left\{(x_1; x_{2;} x_3) \in \mathbb{R}^n : x_1^2 + x_2^2 \leq R^2 \wedge 0 \leq x_3 \leq R\right\} \subset \mathbb{R}^n$ of radius $R \geq 0$ and height $h \geq 0$ described by **cartesian coordinates** $(x_1; x_{2;} x_3)$ is a more conveniently desribed as a **neither open nor closed cuboid** $\mathbf{g}^{-1}[C_{h;R}] = [0; R] \times [0; 2\pi[\times [0; h]$ defined by **cylinder coordinates** $(r; \varphi; z)$ with the **bijective** and **continuous** transformation $\mathbf{g} : \mathbf{g}^{-1}[C_{h;R}] \to C_{h;R}$ defined by $\mathbf{g}(r; \varphi; z) = (r \cdot \cos(\varphi); r \cdot \sin(\varphi); z)$ with

$$\frac{d\mathbf{g}}{d\mathbf{y}}(r;\varphi;z)=\begin{pmatrix}\cos(\varphi) & -r\cdot\sin(\varphi) & 0\\ \sin(\varphi) & r\cdot\cos(\varphi) & 0\\ 0 & 0 & 1\end{pmatrix}$$

and $\det\left(\frac{d\mathbf{g}}{d\mathbf{y}}\right)(r;\varphi;z)=r$ such that we may apply **Fubini's theorem** 8.5 to compute the **volume**

$$\begin{aligned}\lambda^3(C_{h;R}) &= \int_{C_{h;R}} 1d\mathbf{g}\\ &= \int_{\mathbf{g}^{-1}[C_{h;R}]}(1\circ\mathbf{g})\cdot\left|\det\left(\frac{d\mathbf{g}}{d\mathbf{y}}\right)\right|d\mathbf{y}\\ &= \int_{[0;R]\times[0;2\pi[\times[0;h]} 1\cdot rd\mathbf{y}\\ &= \int_0^R\left(\int_0^{2\pi}\left(\int_0^h rdz\right)d\varphi\right)dr\\ &= \pi r^2\cdot h.\end{aligned}$$

14.15 The closed sphere $B_R=\left\{(x_1;x_{2;}x_3)\in\mathbb{R}^n : x_1^2+x_2^2+x_3^2\leq R^2\right\}\subset\mathbb{R}^n$ of radius $R\geq 0$ is a more conveniently desribed as a **neither open nor closed cuboid** $\mathbf{g}^{-1}[B_R]=[0;R]\times[0;2\pi[\times[0;\pi]$ defined by **spherical coordinates** $(r;\varphi;\psi)$ with the **bijective** and **continuous** transformation $\mathbf{g}:\mathbf{g}^{-1}[B_R]\to B_R$ defined by $\mathbf{g}(r;\varphi;\psi)=(r\cdot\cos(\varphi)\cdot\sin(\psi);r\cdot\sin(\varphi)\cdot\sin(\psi);r\cdot\cos(\psi))$ with

$$\frac{d\mathbf{g}}{d\mathbf{y}}(r;\varphi;\psi)=\begin{pmatrix}\cos(\varphi)\cdot\sin(\psi) & -r\cdot\sin(\varphi)\cdot\sin(\psi) & r\cdot\cos(\varphi)\cdot\cos(\psi)\\ \sin(\varphi)\cdot\sin(\psi) & r\cdot\cos(\varphi)\cdot\sin(\psi) & r\cdot\sin(\varphi)\cdot\cos(\psi)\\ \cos(\psi) & 0 & -r\cdot\sin(\psi)\end{pmatrix}$$

and

$$\det\left(\frac{d\mathbf{g}}{d\mathbf{y}}\right)(r;\varphi;z)=\cos(\psi)\left(-r^2\cdot\cos^2(\psi)\cdot\sin(\psi)\right)-r\cdot\sin(\psi)\cdot r\cdot\sin^2(\psi)=-r^2\sin(\psi)$$

resulting in

$$\begin{aligned}\lambda^3(B_R) &= \int_0^R\left(\int_0^{2\pi}\left(\int_0^{\pi}\left|-r^2\sin(\psi)\right|d\psi\right)d\varphi\right)dr\\ &= \int_0^R\left(\int_0^{2\pi}\left(2r^2\right)d\varphi\right)dr\\ &= \int_0^R\left(4\pi r^2\right)dr\\ &= \frac{4}{3}\pi Rr.\end{aligned}$$

14.16 Special cases:

1. Applying the theorem to $(\boldsymbol{f}\circ\mathbf{g})\circ\mathbf{g}^{-1}$ yields the **inverse variant** $\int_{\mathbf{g}^{-1}[\mathbf{g}[A]]}(\boldsymbol{f}\circ\mathbf{g})\,dy=\int_{\mathbf{g}[A]}(\boldsymbol{f}\circ\mathbf{g})\circ g^{-1}\cdot\left|\det\left(\frac{d\mathbf{g}^{-1}}{d\mathbf{g}}\right)\right|d\mathbf{g}$ resp.

$$\int_A\boldsymbol{f}(\mathbf{g}(\mathbf{y}))\,d\mathbf{y}=\int_{\mathbf{g}[A]}\boldsymbol{f}(\mathbf{g})\cdot\left|\det\left(\frac{d\mathbf{y}}{d\mathbf{g}}\right)\right|d\mathbf{g}.$$

2. In the case of a **monotone and differentiable** $g:[a;b]\to\mathbb{R}$ and $f=1\in\mathbb{R}$ we have $g(b)-g(a)>0\Leftrightarrow\frac{dg_i}{dy}>0$ so that we recover the **fundamental theorem of calculus** :

$$g(b)-g(a)=\int_{g[[a;b]]}dg=\int_{[a;b]}\left|\frac{dg}{dy}\right|dy=\int_a^b\frac{dg}{dy}dy$$

.

3. In the case of a **monotone** and **differentiable** $g : A \to \mathbb{R}$ on a **one-dimensional interval** $A = [a; b] \subset \mathbb{R}$ and an **integrable** $\boldsymbol{f} = (f_1; ...; f_n) : g[A] \to Y$ into a **Banach** space $(Y; |\ |)$ for **each component** f_i we have $\int_{g[A]} f_i(g)\, dg = \int_A f_i(g(y)) \cdot \left|\frac{dg}{dy}\right| dy$. As above for **strictly increasing** g we obtain $\int_{[g(a);g(b)]} f_i(g)\, dg = \int_{[a;b]} f_i(g(y)) \cdot \frac{dg}{dy} dy$ whereas for **strictly decreasing** g it is $\int_{[g(b);g(a)]} f_i(g)\, dg = -\int_{[a;b]} f_i(g(y)) \cdot \frac{dg}{dy} dy$ such that we may apply the **fundamental theorem of calculus** 13.10 to combine the two cases from each component to obtain the **integration by substitution**
$$\int_{g(a)}^{g(b)} \boldsymbol{f}(g)\, dg = \boldsymbol{F}(g(b)) - \boldsymbol{F}(g(a)) = \int_a^b \boldsymbol{f}(g(y)) \cdot \frac{dg}{dy} dy$$

4. For $A = [a; x]$ we can write $\int_{g(a)}^{g(x)} \boldsymbol{f}(g)\, dg = \boldsymbol{F}(g(x)) - \boldsymbol{F}(g(a)) = \int_a^x \boldsymbol{f}(g(y)) \cdot \frac{dg}{dy} dy$ with derivative $\frac{d\boldsymbol{F}}{dg}(g(x)) = \boldsymbol{f}(g(x))$ of the left side and $\frac{d(\boldsymbol{F} \circ g)}{dy}(x) = \boldsymbol{f}(g(x)) \cdot \left|\frac{dg}{dy}(x)\right|$ on the right side which results in the **chain rule**
$$\frac{d(\boldsymbol{F} \circ g)}{dy}(x) = \frac{d\boldsymbol{F}}{dg}(g(x)) \cdot \frac{dg}{dy}(x)$$

15 Some special functions

15.1 Bohr-Mollerup theorem: The **Gamma function** $\Gamma : \mathbb{R}^+ \to \mathbb{R}^+$ with $\Gamma(x) = \int_0^\infty t^{x-1} e^{-t} dt = \lim\limits_{n\to\infty} \frac{n! n^x}{x \cdot (x+1) \cdot ... \cdot (x+n)}$ is well defined and **uniquely determined** by the following three properties:

1. $\Gamma(x+1) = x \cdot \Gamma(x)$
2. $\Gamma(1) = 1$
3. $\ln \Gamma$ is **convex.**

Proof:

The three porperties determine a unique function $\Gamma : \mathbb{R}^+ \to \mathbb{R}^+$ with $\Gamma(x) = \frac{n! n^x}{x \cdot (x+1) \cdot ... \cdot (x+n)}$: Because of 1. we can restrict the argument to $0 < x < 1$. We examine $\ln f$ with $\ln f(1) = 0$ and $\ln f(x+1) = \ln x + \ln f(x)$ resp. $\ln f(n+1+x) = \ln f(x) + \ln[x \cdot (x+1) \cdot ... \cdot (x+n)]$ (*). Due to the convexity of $\ln f$ the **difference quotients** $\frac{\ln f(x+h) - \ln f(x)}{h}$ are nondecreasing, whence $\ln n \leq \frac{\ln f(n+1+x) - \ln f(n+1)}{x} \leq \ln n + 1$. Substituting (*) yields $0 \leq \ln f(x) - \ln\left[\frac{n! n^x}{x \cdot (x+1) \cdot ... \cdot (x+n)}\right] \leq x \ln\left(1 + \frac{1}{n}\right)$. Since the last expression tends to 0 as $n \to \infty$ the Function f is uniquely determined by the expression in the middle.

The integral $I(x) = \int_0^\infty t^{x-1} e^{-t} dt$ is well defined for $0 < x < \infty$: We split the domain as well as the range of the integral into $[0; 1]$ and $[1; \infty]$: In the case of $x < 1$ we consider $\int_0^1 t^{x-1} e^{-t} dt \overset{5.7}{\leq} \int \lim\limits_{n\to\infty} \chi_{[1/n;0]} \cdot t^{x-1} \cdot e \cdot dt \overset{5.12}{=} \lim\limits_{n\to\infty} \int_{1/n}^1 e \cdot t^{x-1} dt = \lim\limits_{n\to\infty} \frac{e}{x}\left(1 - \frac{1}{n^x}\right) = \frac{e}{x} < \infty$ whereas for $x \geq 1$ we have $\int_0^1 t^{x-1} e^{-t} dt \overset{5.7}{\leq} \int_0^1 e^{-t} dt = 1 - \frac{1}{e} < \infty$. The remainder converges in any case since $\int_1^\infty t^{x-1} e^{-t} dt \overset{5.7}{\leq} \int_1^\infty \frac{[x+1]!}{t^2} dt = \frac{[x+1]!}{3} < \infty$.

An **integration by parts** 14.7 delivers the **functional equation** $I(x+1) = x \cdot I(x)$ for $0 < x < \infty$. Since $I(1) = 1$ we have $I(n+1) = n!$ for $n \geq 1$. Finally **Hölder's inequality** 6.4.1 with $f(t) = \left(t^{x-1} e^{-t}\right)^{1/p}$ resp. $g(t) = \left(t^{x-1} e^{-t}\right)^{1/q}$ delivers $I\left(\frac{x}{p} + \frac{y}{q}\right) \leq I^{1/p}(x) \cdot I^{1/q}(y)$ for $\frac{1}{p} + \frac{1}{q} = 1$ such that $\ln I$ is **convex** for $0 < x < \infty$. Owing to the first part we have $I = \Gamma$.

15.2 Theorem: For $x, y > 0$ the **Beta function** is $B(x; y) = \int_0^1 t^{x-1} (1-t)^{y-1} dt = \frac{\Gamma(x) \cdot \Gamma(y)}{\Gamma(x+y)}$.

Proof: As above for each fixed y we find $B\left(\frac{x_1}{p} + \frac{x_2}{q}, y\right) \leq B^{1/p}(x_1) \cdot B^{1/q}(x_2)$ for $\frac{1}{p} + \frac{1}{q} = 1$, i.e. the function $\ln B(x, y)$ is **convex** in x. Also we have $B(1, y) = \frac{1}{y}$ and $B(x+1, y) = \int_0^1 \left(\frac{t}{1-t}\right)^x (1-t)^{x+y-1} dt = \frac{x}{x+y} \int_0^1 \left(\frac{1}{1-t}\right)^2 \left(\frac{t}{1-t}\right)^{x-1} (1-t)^{x+y} dt = \frac{x}{x+y} B(x, y)$. Hence the function

$f\left(x\right)=\frac{\Gamma(x+y)}{\Gamma(y)}\cdot B\left(x,y\right)$ satisfies the three characteristic properties of Γ and we can apply the **Bohr-Mollerup theorem** 15.1 to conclude that $f=\Gamma$.

15.3 Corollary: We have

1. $\frac{\Gamma(x)\cdot\Gamma(y)}{\Gamma(x+y)}\overset{t=\sin^2\vartheta}{=}2\int_0^{\pi/2}\left(\sin\vartheta\right)^{2x-1}\cdot\left(\cos\vartheta\right)^{2y-1}d\vartheta$ which in the case of $x=y=\frac{1}{2}$ gives
2. $\Gamma\left(\frac{1}{2}\right)=\sqrt{\pi}$ and by another application of the Bohr-Mollerup theorem
3. $\Gamma\left(x\right)=\frac{2^{x-1}}{\sqrt{\pi}}\Gamma\left(\frac{x}{2}\right)\cdot\Gamma\left(\frac{x+1}{2}\right)$. The original integral yields the identity
4. $\Gamma\left(x\right)\overset{t=s^2}{=}2\int_0^\infty s^{2x-1}e^{-s^2}ds$ which in the case of $x=\frac{1}{2}$ gives
5. $\int e^{-s^2}ds=\sqrt{\pi}$ resp $\int e^{-s^2/2}ds=\sqrt{2\pi}$.

15.4 Stirling's formula: $\lim\limits_{x\to\infty}\frac{\Gamma(x+1)}{\sqrt{2\pi x}\left(\frac{x}{e}\right)^x}=1$

Proof: The substitution $t=x+s\sqrt{2x}$ gives

$\Gamma\left(x+1\right)$

$$=\int_0^\infty t^xe^{-t}dt=\sqrt{2x}\left(\frac{x}{e}\right)^x=\int_{-\sqrt{x/2}}^\infty\left(1+s\sqrt{2/x}\right)^xe^{-s\sqrt{2x}}ds=\sqrt{2x}\left(\frac{x}{e}\right)^x\int_{-\sqrt{x/2}}^\infty e^{-s^2\cdot h_x(s)}ds$$

with $h_x\left(s\right)=\frac{1}{s^2}\left(s\sqrt{2x}-x\ln\left(1+s\sqrt{2/x}\right)\right)=\frac{2}{\left(s\sqrt{2/x}\right)^2}\left(s\sqrt{2/x}-\ln\left(1+s\sqrt{2/x}\right)\right)$. Thus we have $\lim\limits_{x\to\infty}h_x\left(s\right)=\lim\limits_{u\to0}\frac{2}{u^2}\left(u-\ln\left(1+u\right)\right)=\lim\limits_{u\to0}\frac{2}{u^2}\sum\limits_{k=2}^\infty\frac{u^k}{k}=1$. Furthermore $h_x\left(s\right)>0$ for every $u>-1$ resp. $s>-\sqrt{x/2}$ and $h_x\left(s\right)>1$ for every $u>0$ resp. $s>0$. Hence we can apply **dominated convergence** 5.14 with the **Lebesgue integrable majorant** $m_x\left(s\right)=\begin{cases}0 & \text{for } s\leq-\sqrt{x/2}\\ 1 & \text{for } -\sqrt{x/2}<s<0\\ e^{-s^2} & \text{for } s\geq0\end{cases}$ to obtain $\lim\limits_{x\to\infty}\int_{-\sqrt{x/2}}^\infty e^{-s^2\cdot h_x(s)}ds=\int\lim\limits_{x\to\infty}\chi_{\left[-\sqrt{x/2};\infty\right[}e^{-s^2\cdot h_x(s)}ds=\int e^{-s^2}ds=\sqrt{\pi}$ and hence the assertion.

15.5 Rotational symmetric functions: For a **compact sphere** $K=\{\boldsymbol{x}\in\mathbb{R}^n:\rho\leq|\boldsymbol{x}|\leq R\}\subset\mathbb{R}^n$ with $0\leq\rho<R$ and every **continuous** $f\in L^1\left([\rho;R];\mathbb{R}\right)$ we have $\int_Kf\left(|\boldsymbol{x}|\right)d\boldsymbol{x}=n\cdot\tau_n\cdot\int_\rho^Rf\left(r\right)\cdot r^{n-1}dr$ with the **volume** $\tau_n=\frac{\pi^{n/2}}{\Gamma\left(\frac{n}{2}+1\right)}$ **of the unit sphere** (cf 8.13) .

Proof: For $N\geq2$ and $0\leq k\leq N$ we define $r_k=\rho\frac{k}{N}\left(R-\rho\right)$ and corresponding **partitions** of the interval $[\rho;R[$ by $A_k=\{x\in\mathbb{R}^n:r_{k-1}\leq|x|<r_k\}$ for $1\leq k\leq N$. Due to 8.13 we have $\lambda^n\left(A_k\right)=\lambda^n\left(B_{r_k}\right)-\lambda^n\left(B_{r_{k-1}}\right)=\tau_n\left(r_k^n-r_{k-1}^n\right)$ and according to the **mean value theorem** 12.9 there is a $\xi_k\in]r_{k-1};r_k[$ such that $\lambda^n\left(A_k\right)=n\cdot\tau_n\cdot\xi_k^{n-1}\cdot\left(r_k-r_{k-1}\right)=n\cdot\tau_n\cdot\xi_k^{n-1}\cdot\lambda\left([r_k;r_{k-1}[\right)$. For the **step functions** $\psi_N=\sum\limits_{k=1}^Nf\left(\xi_k\right)\cdot\chi_{A_k}$ resp. $\varphi_N=n\cdot\tau_n\cdot\sum\limits_{k=1}^Nf\left(\xi_k\right)\cdot\xi_k^{n-1}\cdot\chi_{[r_k;r_{k-1}[}$ and considering the λ^n-null set $\{|\boldsymbol{x}|=R\}$ we then have $\int_K\psi_Nd\lambda^n=\int_{[\rho;R]}\varphi_Nd\lambda$. Furthermore since $f\left(|\boldsymbol{x}|\right)$ resp. $f\left(r\right)\cdot r^{n-1}$ are **uniformly continuous** on the **compact** interval $[\rho;R]$ we have **pointwise** $\lim\limits_{N\to\infty}\psi_N\left(|\boldsymbol{x}|\right)=f\left(|\boldsymbol{x}|\right)$ resp. $\lim\limits_{N\to\infty}\varphi_N\left(r\right)=n\cdot\tau_n\cdot f\left(r\right)\cdot r^{n-1}$. Since all concerned functions are bounded on the compact set K resp. $[\rho;R]$ we can apply the **dominated convergence theorem** 5.14 twice to obtain $\int_Kf\left(|\boldsymbol{x}|\right)d\boldsymbol{x}=\int_K\lim\limits_{N\to\infty}\psi_N\left(|\boldsymbol{x}|\right)d\boldsymbol{x}=\lim\limits_{N\to\infty}\int_K\psi_N\left(|\boldsymbol{x}|\right)d\boldsymbol{x}=\lim\limits_{N\to\infty}\int_{[\rho;R]}\varphi_N\left(r\right)dr=\int_{[\rho;R]}\lim\limits_{N\to\infty}\varphi_N\left(r\right)dr=n\cdot\tau_n\cdot\int_\rho^Rf\left(r\right)\cdot r^{n-1}dr$.

15.6: **Moments of inertia**: We compute some moments of inertia $\Theta_{x_1}=\int_K\left(x_2^2+x_3^2\right)\cdot\rho\cdot d\boldsymbol{x}$ of **compact** bodies $K\subset\mathbb{R}^3$ with constant **densities** $\rho>o$ for rotation around the x_1-axis and in relation thier **mass** $m_K=\int_K\rho\cdot d\boldsymbol{x}$:

1. For the **hollow sphere** $K = \{\boldsymbol{x} \in \mathbb{R}^3 : \rho \leq |\boldsymbol{x}| \leq R\}$ with mass $m_K = \rho\tau_3\left(R^3 - \rho^3\right)$ we have $\Theta_{x_1} = \int_K \left(x_2^2 + x_3^2\right) \cdot \rho \cdot d\boldsymbol{x} = \frac{2}{3}\int_K \left(x_1^2 + x_2^2 + x_3^2\right) \cdot \rho \cdot d\boldsymbol{x} = \frac{2}{3}\int_K |\boldsymbol{x}|^2 \cdot \rho \cdot d\boldsymbol{x} \overset{15.5}{=} 2\rho\tau_3 \int_\rho^R r^4 dr = \frac{2}{5}m_K \frac{R^5-\rho^5}{R^3-\rho^3}$.
2. For the **ellipsoid** $K = \left\{\boldsymbol{x} \in \mathbb{R}^3 : \left(\frac{x_1}{a_1}\right)^2 + \left(\frac{x_2}{a_2}\right)^2 + \left(\frac{x_3}{a_3}\right)^2 \leq 1\right\}$ with mass $m_K = \int_K \rho \cdot d\boldsymbol{x} \overset{x_i'=x_i/a_i}{=} \rho\int_{|\boldsymbol{x}'|\leq 1} a_1a_2a_3 d\boldsymbol{x}' = a_1a_2a_3\rho\tau_3$ we have
$\Theta_{x_1} = \int_K \left(x_2^2 + x_3^2\right) \cdot \rho \cdot d\boldsymbol{x} = a_1a_2a_3\rho\int_{|\boldsymbol{x}'|\leq 1} \left(a_2^2{x_2'}^2 + a_3^2{x_3'}^2\right) d\boldsymbol{x}' = a_1a_2a_3\rho\frac{a_2^2+a_3^2}{3}\int_{|\boldsymbol{x}'|\leq 1} |\boldsymbol{x}'|^2 d\boldsymbol{x}'$ $\overset{15.5}{=} a_1a_2a_3\rho\tau_3\left(a_2^2 + a_3^2\right)\int_0^1 r^4 dr = \frac{1}{5}m_K\left(a_2^2 + a_3^2\right)$.
3. For the **tube** $K = \{\boldsymbol{x} \in \mathbb{R}^3 : 0 \leq x_1 \leq a; \rho \leq x_2^2 + x_3^2 \leq R\}$ with mass $m_K = \int_K \rho \cdot d\boldsymbol{x} \overset{14.14}{=} a\rho\pi\left(R^2 - \rho^2\right)$ we have $\Theta_{x_1} = \int_K \left(x_2^2 + x_3^2\right) \cdot \rho \cdot d\boldsymbol{x} \overset{14.14}{=} \int_0^a \int_0^{2\pi} \int_\rho^R r^2 \rho dr d\varphi dx_1 = \frac{2}{3}a\rho\pi\left(R^3 - \rho^3\right) = \frac{2}{5}m_K\frac{R^3-\rho^3}{R^2-\rho^2}$.
4. **Steiner's theorem**: For an axis L' with distance d to the axis L through the **center $\boldsymbol{s}$ of gravity** we have $\Theta_{L'} = \Theta_L + md^2$: W.l.o.g we set the center of mass at the **origin** such that $\boldsymbol{s} = \int_K \boldsymbol{x} d\boldsymbol{x} = \boldsymbol{0}$ and particularly $\int_{K_{x_2x_3}} x_1 dx_1 = 0$ on every **cut** $K_{x_2x_3} = \{\boldsymbol{y} \in K : y_2 = x_2; y_3 = x_3\}$.
Also we choose $L = \{x_2 = x_3 = 0\}$ resp. $L' = \{x_2 = d_2; x_3 = d_3\}$ such that $\Theta_{L'} \overset{x_i'=x_i+d_i}{=} \int_K \left((x_2 + d_2)^2 + (x_3 + d_3)^2\right) \cdot \rho \cdot d\boldsymbol{x} = \int_K \left(x_2^2 + x_3^2\right) \cdot \rho \cdot d\boldsymbol{x} + 2\int_K \left(d_2x_2 + d_3x_3\right) \cdot \rho \cdot d\boldsymbol{x} + \int_K \left(d_2^2 + d_3^2\right) \cdot \rho \cdot d\boldsymbol{x} = \Theta_L + 0 + md^2$.

15.7 The Gauss integral: By 15.5 and **monotone convergence** 5.12 we have $\int_{\mathbb{R}^n} e^{-|\boldsymbol{x}|^2} d\boldsymbol{x} \overset{5.12}{=} \lim_{R\to\infty} \int_{|\boldsymbol{x}|\leq R} e^{-|\boldsymbol{x}|^2} d\boldsymbol{x} \overset{15.5}{=} n \cdot \tau_n \cdot \lim_{R\to\infty} \int_0^R e^{-|\boldsymbol{x}|^2} r^{n-1} dr \overset{t=r^2}{=} \frac{n\cdot}{2} \cdot \tau_n \cdot 2^{n/2} \cdot \lim_{R\to\infty} \int_0^{R^2/2} e^{-t} t^{n/2-1} dt = \frac{n}{2} \cdot \tau_n \cdot \Gamma\left(\frac{n}{2}\right) = \pi^{n/2}$ with the definition resp. functional equation 15.1 of the **Gamma function** as well as the volume $\tau_n = \frac{\pi^{n/2}}{\Gamma\left(\frac{n}{2}+1\right)} = \frac{\pi^{n/2}}{\frac{n}{2}\cdot\Gamma\left(\frac{n}{2}\right)}$ of the unit sphere according to 8.13.

16 Fourier transforms

In this section L^p;C_C^∞, etc. stands for $L^p\left(\mathbb{R}^n; \mathbb{C}\right)$, etc. if not specified otherwise.

16.1 Convolutions: For every $f \in L^1$ and $g \in L^p$ with $1 \leq p \leq \infty$ their **convolution** $f * g : \mathbb{R}^n \to \mathbb{C}$ defined by $(f * g)(\boldsymbol{x}) = \int f(\boldsymbol{x} - \boldsymbol{y}) g(\boldsymbol{y}) d\boldsymbol{y}$ is also p- **integrable** with $\|f * g\|_p \leq \|f\|_1 \cdot \|g\|_p$. Since the convolution is **associative**, **commutative** and **bilinear** the space $L^1\left(\mathbb{R}^n; \mathbb{C}; +; *\right)$ is a **Banach algebra**.

Proof: The algebraic properties of the convolution are obvious resp. (in the case of associativity) tedious. According to the hypothesis resp. the **translation invariance** 8.8 of λ for every $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$ the functions $y \mapsto |g(\boldsymbol{y} - \boldsymbol{x})|^p$ with $\int |g(\boldsymbol{y} - \boldsymbol{x})|^p dy = \|g\|_p^p$ and $\boldsymbol{x} \mapsto \|g\|_p^p \cdot f(\boldsymbol{x})$ are **Lebesgue integrable**. Hence we may apply **Hölder's inequality** 6.4.1 with $\frac{1}{p} + \frac{1}{q} = 1$ and **Fubini's theorem** 8.5 to obtain

$$\begin{aligned}
\|f * g\|_p^p &= \int \left|\int |f(\boldsymbol{x})|^{\frac{1}{p}} \cdot |g(\boldsymbol{y} - \boldsymbol{x})| \cdot |f(\boldsymbol{x})|^{\frac{1}{q}} d\boldsymbol{x}\right|^p d\boldsymbol{y} \\
&= \int \left|\int f(\boldsymbol{x}) \cdot g(\boldsymbol{y} - \boldsymbol{x}) dx\right|^p d\boldsymbol{y} \\
&\overset{6.4.1}{\leq} \int \left(\int |f(\boldsymbol{x})| \cdot |g(\boldsymbol{y} - \boldsymbol{x})|^p d\boldsymbol{x} \cdot \left(\int |f(\boldsymbol{x})| d\boldsymbol{x}\right)^{\frac{p}{q}}\right) d\boldsymbol{y} \\
&\overset{8.5}{=} \int \left(\int |g(\boldsymbol{y} - \boldsymbol{x})|^p d\boldsymbol{y}\right) \cdot |f(\boldsymbol{x})| d\boldsymbol{x} \cdot \|f\|^{\frac{p}{q}} \\
&= \|g\|_p^p \cdot \|f\|_1 \cdot \|f\|^{\frac{p}{q}} \\
&= \|g\|_p^p \cdot \|f\|_1^p.
\end{aligned}$$

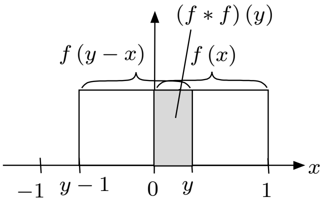
16.2 Examples: We examine the effect of a convolution on $f = \chi_{[0;1]}$: With $\chi_{[0;1]}(x-y) = \chi_{[x-1;x]}(y)$ we compute

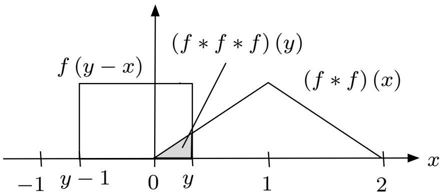
$$(f * f)(x) = \int \chi_{[x-1;x]}(y) \cdot \chi_{[0;1]}(y) dy$$

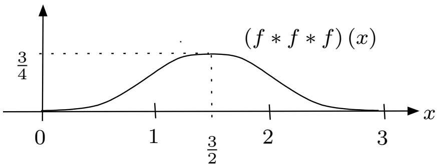
$$= \int \chi_{[x-1;x]\cap[0;1]}(y)\, dy$$

$$= \begin{cases} 0 & : x < 0 \\ \int \chi_{[0;x]}(y)\, dy = x & : 0 \leq x < 1 \\ \int \chi_{[x-1;1]}(y) dy = 2 - x & : 1 \leq x < 2 \\ 0 & : 2 \leq x \end{cases}$$

$$= \sup\{0; 2 - |x|\}$$







$$(f * f * f)(x) = \begin{cases} 0 & : x < 0 \\ \int \chi_{[x-1;x]\cap[0;1]}(y) \cdot y \cdot dy & : 0 \leq x < 1 \\ \int \chi_{[x-1;x]\cap[1;2]}(y) \cdot (2-y) \cdot dy & : 1 \leq x < 2 \\ 0 & : 2 \leq x \end{cases}$$

$$= \begin{cases} 0 & : x < 0 \\ \int \chi_{[0;x]}(y) \cdot y \cdot dy = \frac{1}{2}x^2 & : 0 \leq x < 1 \\ \int \chi_{[x-1;1]}(y) \cdot y \cdot dy + \int \chi_{[1;x]}(2-y) \cdot dy = -\left(x - \frac{3}{2}\right)^2 + \frac{3}{4} & : 1 \leq x < 2 \\ \int \chi_{[x-1;2]}(2-y) \cdot dy = \frac{1}{2}(x-3)^2 & : 2 \leq x < 3 \\ 0 & : 3 \leq x \end{cases}$$

16.3 Continuity under the integral sign: Let $f : \mathbb{R}^n \times I \to \mathbb{R}$ be a function and $I \subset \mathbb{R}$ an interval such that

1. $t \mapsto f(\boldsymbol{x}; t)$ is **continuous** for every $\boldsymbol{x} \in \mathbb{R}^n$
2. $x \mapsto f(\boldsymbol{x}; t)$ is **integrable** for every $t \in I$
3. There exists a **Lebesgue integrable majorant** $F \in L^1(\mathbb{R}^n; \mathbb{R})$ with $|f(\boldsymbol{x}; t)| \leq F(\boldsymbol{x})$ for every $(\boldsymbol{x}; t) \in \mathbb{R}^n \times I$.

Then $g : I \to \mathbb{R}$ with $g(t) = \int f(\boldsymbol{x}; t)\, d\boldsymbol{x}$ is **continuous** on I.

Proof: For every sequence $(t_n)_{n\in\mathbb{N}}$ with $\lim\limits_{n\to\infty} t_n = t$ we have $\lim\limits_{n\to\infty} f(\boldsymbol{x}; t_n) = f(\boldsymbol{x}; t)$ on account of 1. and due to 2. and 3. the **dominated convergence** theorem 5.14 yields

$$\lim_{n\to\infty} g(t_n) = \lim_{n\to\infty} \int f(\boldsymbol{x}; t_n)\, dx = \int \lim_{n\to\infty} f(\boldsymbol{x}; t_n)\, dx = \int f(\boldsymbol{x}; t)\, dx = g(t)$$

and hence the assertion due to [9, 3.2].

16.4 Differentiability under the integral sign: Let $f : \mathbb{R}^n \times I \to \mathbb{R}$ be a function and $I \subset \mathbb{R}$ an interval such that

1. $t \mapsto f(\boldsymbol{x}; t)$ is **differentiable** for every $\boldsymbol{x} \in \mathbb{R}^n$
2. $x \mapsto f(\boldsymbol{x}; t)$ is **integrable** for every $t \in I$
3. There exists a **Lebesgue integrable majorant** $F \in L^1(\mathbb{R}^n; \mathbb{R})$ with $\left|\frac{\delta f}{\delta t}(\boldsymbol{x}; t)\right| \leq F(\boldsymbol{x})$ **uniformly in** t for every $(\boldsymbol{x}; t) \in \mathbb{R}^n \times I$.

Then $g : I \to \mathbb{R}$ with $g(t) = \int f(\boldsymbol{x}; t)\, d\boldsymbol{x}$ is **differentiable** on I with $g'(t) = \int \frac{\delta f}{\delta t}(\boldsymbol{x}; t)\, dx$

Proof: For every sequence $(t_n)_{n\in\mathbb{N}}$ with $\lim\limits_{n\to\infty} t_n = t$ we have $\lim\limits_{n\to\infty} \frac{f(\boldsymbol{x};t)-f(\boldsymbol{x};t_n)}{t-t_n} = \frac{\delta f}{\delta t}(\boldsymbol{x}; t)$ on account of 1. and a sequence $(\tau_n)_{n\in\mathbb{N}}$ with $\tau_n \in [t_n; t]$ resp. $[t; t_n]$ and in particular $\lim\limits_{n\to\infty} \tau_n = t$ such that $\frac{f(\boldsymbol{x};t)-f(\boldsymbol{x};t_n)}{t-t_n} = \frac{\delta f}{\delta t}(\boldsymbol{x}; \tau_n)$ according to the **mean value theorem** 13.11. Owing to 3. we can apply the **dominated convergence** theorem 5.14 to obtain $g'(t) = \lim\limits_{n\to\infty} \frac{g(\boldsymbol{x};t)-g(\boldsymbol{x};t_n)}{t-t_n} = \lim\limits_{n\to\infty} \int \frac{f(\boldsymbol{x};t)-f(\boldsymbol{x};t_n)}{t-t_n} d\boldsymbol{x} = \lim\limits_{n\to\infty} \int \frac{\delta f}{\delta t}(\boldsymbol{x}; \tau_n)\, d\boldsymbol{x} = \int \lim\limits_{n\to\infty} \frac{\delta f}{\delta t}(\boldsymbol{x}; \tau_n)\, d\boldsymbol{x} = \int \frac{\delta f}{\delta t}(\boldsymbol{x}; t)\, d\boldsymbol{x}$ and hence the assertion.

16.5 The Bessel functions $J_p(x) = \frac{1}{\sqrt{\pi}} \frac{\left(\frac{x}{2}\right)^p}{\Gamma\left(p+\frac{1}{2}\right)} \int_0^\pi \sin^{2p} t \cdot e^{-ix\cdot\cos t} dt$ for any $p \geq 0$ are generalized trigonometric functions solving the radial component of the **Laplace equation** in cylindrical coordinates as well as the radial component of the **Helmholtz equation** in spherical coordinates, i.e. the **Bessel differential equation** $\frac{\delta^2 J_p}{\delta r^2}(r) + \frac{1}{r}\frac{\delta J_p}{\delta r}(r) + \left(1 - \frac{p^2}{r^2}\right) J_p(r) = 0$. We have $J_{1/2}(x) = \sqrt{\frac{2}{\pi}} \cdot \frac{\sin x}{\sqrt{x}}$ and $J_{p+1}(x) = -\frac{\delta J_p}{\delta x}(x) + \frac{p}{x} J_p(x)$.

Proof: Preliminarily we examine the functions $f_p(x) = \int_0^\pi \sin^{2p} t \cdot e^{-x\cdot\cos t} dt$ with $\frac{\delta^k f_p}{\delta r^k}(x) = (-i)^k \cdot \int_0^\pi \cos^k t \cdot \sin^{2p} t \cdot e^{-x\cdot\cos t} dt$. With $\cos^2 t = 1 - \sin^2 t$ and omitting the argument x for brevity we obtain $\frac{\delta^2 f_p}{\delta x^2} = f_{p+1} - f_p$ (*). In order to express f_{p+1} in terms of $\frac{\delta f_p}{\delta x}$ we compute

$$\begin{aligned}\frac{\delta f_p}{\delta x} &= -i \cdot \int_0^\pi \cos t \cdot \sin^{2p} t \cdot e^{-x\cdot\cos t} dt \\ &= 0 + i \cdot \int_0^\pi \sin t \cdot \left(\cos t \cdot 2p \cdot \sin^{2p-1} t + \sin^{2p} t \cdot x \cdot \sin t\right) \cdot e^{-x\cdot\cos t} dt \\ &= -2p \cdot \frac{\delta f_p}{\delta x} - x \cdot f_{p+1}\end{aligned}$$

resp. $(2p+1)\frac{\delta f_p}{\delta x} = -x \cdot f_{p+1}$ (**). Hence we can substitute f_{p+1} in equation (*) to obtain $\frac{\delta^2 f_p}{\delta x^2} + \frac{2p+1}{x} \cdot \frac{\delta f_p}{\delta x} + f_p = 0$. Since $J_p = c_p \cdot x^p \cdot f_p$ and dividing by $c_p = \frac{1}{\sqrt{\pi}} \frac{1}{2^p \cdot \Gamma\left(p+\frac{1}{2}\right)}$ we have

$$\begin{aligned}\frac{\delta^2 J_p}{\delta x^2} + \frac{1}{x}\frac{\delta J_p}{\delta x} + \left(1 - \frac{p^2}{x^2}\right) J_p &= \left(p(p-1) \cdot x^{p-2} \cdot f_p + 2p \cdot x^{p-1} \cdot \frac{\delta f_p}{\delta x} + x^p \cdot \frac{\delta^2 f_p}{\delta x^2}\right) \\ &\quad + \frac{1}{x} \cdot \left(p \cdot x^{p-1} \cdot f_p + x^p \cdot \frac{\delta f_p}{\delta x}\right) + \left(1 - \frac{p^2}{x^2}\right) \cdot x^p \cdot f_p \\ &= x^p \cdot \frac{\delta^2 f_p}{\delta x^2} + \left(2p \cdot x^{p-1} + x^{p-1}\right) \cdot \frac{\delta f_p}{\delta x} \\ &\quad + \left(p(p-1) \cdot x^{p-2} + p \cdot x^{p-2} + x^p - p^2 x^{p-2}\right) \cdot f_p \\ &= x^p \cdot \left(\frac{\delta^2 f_p}{\delta x^2} + \frac{2p+1}{x} \cdot \frac{\delta f_p}{\delta x} + f_p\right) \\ &= 0\end{aligned}$$

For $p = \frac{1}{2}$ an elementary integration yields $J_{1/2}(x) = \sqrt{\frac{x}{2\pi}} \cdot \int_0^\pi \sin t \cdot e^{-ix\cdot\cos t} dt = -\sqrt{\frac{x}{2\pi}} \cdot \int_0^\pi e^{-ix\cdot\cos t} d\cos t = -\sqrt{\frac{1}{2\pi x}} \cdot i \left[e^{-ix\cdot\cos t}\right]_0^\pi = -\sqrt{\frac{1}{2\pi x}} \cdot i \left[e^{ix\cdot} - e^{-ix\cdot}\right]_0^\pi = \sqrt{\frac{1}{2\pi x}} \cdot 2\sin x = \sqrt{\frac{2}{\pi}} \cdot \frac{\sin x}{\sqrt{x}}$. The recusive formula is derived from (**) considering the **functional equation of the Gamma function** $\Gamma\left(p+\frac{3}{2}\right) = \left(p+\frac{1}{2}\right) \cdot$

$\Gamma\left(p+\frac{1}{2}\right)$ by substituting $(2p+1)\left(2^p \cdot \Gamma\left(p+\frac{1}{2}\right)\right) \frac{\delta}{\delta x}\left(x^{-p} \cdot J_p\right) = -x^{-p} \cdot \left(2^{p+1} \cdot \Gamma\left(p+\frac{3}{2}\right)\right) J_{p+1}$ whence $(2p+1)\left(-p \cdot x^{-p-1} \cdot J_p + x^{-p} \cdot \frac{\delta J_p}{\delta x}\right) = -x^{-p} \cdot 2\left(p+\frac{1}{2}\right) \cdot J_{p+1}$ resulting in the desired formula.

16.6 Fourier transforms: Due to 5.18 for every **(Lebesgue) integrable** $f : \mathbb{R}^n \to \mathbb{C}$ the product $\boldsymbol{x} \mapsto f(\boldsymbol{x}) \cdot e^{-i\langle \boldsymbol{x},\boldsymbol{\xi}\rangle}$ is again **(Lebesgue) integrable** and the **Fourier transform** $\hat{f}(\boldsymbol{\xi}) = \frac{1}{(2\pi)^{n/2}} \int f(\boldsymbol{x}) \cdot e^{-i\langle \boldsymbol{x},\boldsymbol{\xi}\rangle} d\boldsymbol{x}$ is **continuous** due to 16.3 and in the case of $f \in L^1$ **bounded** since $\left|\hat{f}(\boldsymbol{\xi})\right| = \frac{\|f\|_1}{(2\pi)^{n/2}}$ according to 5.15.

16.7 Examples:

1. For $f(x) = \chi_{[-1;1]}$ we have $\hat{f}(\xi) = \sqrt{\frac{2}{\pi}} \frac{\sin \xi}{\xi}$
2. For $f(x) = e^{-|x|}$ we have $\hat{f}(\xi) = \sqrt{\frac{2}{\pi}} \frac{1}{1+\xi^2}$
3. For $f(x) = e^{-x^2/2}$ we have $\hat{f}(\xi) = f(\xi)$
4. For $f(\boldsymbol{x}) = \chi_{B_1^n(\boldsymbol{0})}(\boldsymbol{x})$ we have $\hat{f}(\boldsymbol{\xi}) = \frac{J_{n/2}(|\boldsymbol{\xi}|)}{|\boldsymbol{\xi}|^{n/2}}$
5. For $f(\boldsymbol{x}) = e^{-|\boldsymbol{x}|^2/2}$ we have $\hat{f}(\boldsymbol{\xi}) = f(\boldsymbol{\xi})$
6. For $f(\boldsymbol{x}) = e^{-\langle \boldsymbol{x}, A\boldsymbol{x}\rangle}$ with any **symmetric** and **positive definite** $A \in \mathbb{R}^{n \times n}$ we have $\hat{f}(\boldsymbol{\xi}) = \prod_{k=1}^{n} \frac{1}{\sqrt{2d_k}} \cdot e^{-x_k^2/4d_k}$ with positive **eigenvalues** $d_k > 0$ of A.

Proof:

1. $\hat{f}(\xi) = \frac{1}{\sqrt{2\pi}} \int \chi_{[-1;1]} \cdot e^{-ix\xi} dx \overset{13.10}{=} \frac{1}{\sqrt{2\pi}} \left[\frac{e^{-ix\xi}}{-i\xi}\right]_{x=-1}^{x=1} = \frac{1}{\sqrt{2\pi}} \cdot \frac{e^{-ix\xi} - e^{-ix\xi}}{i\xi} = \sqrt{\frac{2}{\pi}} \frac{\sin \xi}{\xi}$.
2. By splitting the integral into two parts we get
$$\begin{aligned}
\hat{f}(\xi) &= \frac{1}{\sqrt{2\pi}} \int e^{-|x|} \cdot e^{-ix\xi} dx \\
&= \frac{1}{\sqrt{2\pi}} \int \lim_{R\to\infty} \chi_{[0;R]} \cdot e^{-x} \cdot \left(e^{-ix\xi} + e^{-ix\xi}\right) dx \\
&\overset{5.14}{=} \frac{1}{\sqrt{2\pi}} \lim_{R\to\infty} \int \chi_{[0;R]} \cdot e^{-x} \cdot \left(e^{-ix\xi} + e^{-ix\xi}\right) dx \\
&\overset{13.10}{=} \frac{1}{\sqrt{2\pi}} \lim_{R\to\infty} \left[\frac{e^{-x(1+i\xi)}}{-1-i\xi} + \frac{e^{-x(1-i\xi)}}{-1+i\xi}\right]_{x=0}^{x=R} \\
&= \sqrt{\frac{2}{\pi}} \frac{1}{1+\xi^2}
\end{aligned}$$
3. Due to $\left|\frac{\delta f}{\delta \xi}(x;\xi)\right| = \left|-ix \cdot e^{-x^2/2} \cdot e^{-ix\xi}\right| = |x| \cdot e^{-x^2/2} \in L^1(\mathbb{R};\mathbb{R})$ for every $\xi \in \mathbb{R}$ and according to 16.4 the function $\hat{f}(\xi) = \frac{1}{\sqrt{2\pi}} \int e^{-x^2/2} \cdot e^{-ix\xi} dx$ is differentiable with
$$\begin{aligned}
\frac{\delta \hat{f}}{\delta \xi}(\xi') &= -\frac{i}{\sqrt{2\pi}} \int x \cdot e^{-x^2/2} \cdot e^{-ix\xi} dx \\
&= -\frac{i}{\sqrt{2\pi}} \int \lim_{R\to\infty} \chi_{[-R;R]} \cdot x \cdot e^{-x^2/2} \cdot e^{-ix\xi} dx \\
&\overset{5.14}{=} -\frac{i}{\sqrt{2\pi}} \lim_{R\to\infty} \int_{-R}^{R} x \cdot e^{-x^2/2} \cdot e^{-ix\xi} dx \\
&= 0 - \frac{\xi}{\sqrt{2\pi}} \lim_{R\to\infty} \int_{-R}^{R} e^{-x^2/2} \cdot e^{-ix\xi} dx \\
&= -\xi \hat{f}(\xi)
\end{aligned}$$
making use of **dominated convergence** 5.14 resp. **integration by parts** 14.7. Considering 15.3.5 we have the initial value $\hat{f}(0) = 1$ so that the differential equation is solved by $\hat{f}(\xi) = e^{-\xi^2/2} = f(\xi)$.

4. On account of the **rotational symmetry** of the integrand we can restrict the computation to $\boldsymbol{\xi} = (0, ..., 0, \xi_n)$ such that

$$\begin{aligned}\hat{f}(\boldsymbol{\xi}) &= \frac{1}{(2\pi)^{n/2}} \int \chi_{B_1^n(\mathbf{0})}(\boldsymbol{x}) \cdot e^{-i\langle \boldsymbol{x},\boldsymbol{\xi}\rangle} d\boldsymbol{x} \\ &= \frac{1}{(2\pi)^{n/2}} \int \chi_{B_1^n(\mathbf{0})}(x_1, ..., x_{1n}) \cdot e^{-ix_n\xi_n} d\boldsymbol{x} \\ &= \frac{1}{(2\pi)^{n/2}} \int \lambda^{n-1}\left(\left(B_1^{n-1}\right)_{x_n}\right) \cdot e^{-ix_n\xi_n} dx_n \\ &\overset{8.13}{=} \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}} \lambda^{n-1}\left(\left(B^{n-1}_{\sqrt{1-x_n^2}}\right)\right) \cdot e^{-ix_n\xi_n} dx_n \\ &= \frac{\tau_{n-1}}{(2\pi)^{n/2}} \cdot \int_{-1}^{1} \left(1 - x_n^2\right)^{(n-1)/2} \cdot e^{-ix_n\xi_n} dx_n \\ &\overset{x_n=\cos t}{=} \frac{1}{(2\pi)^{n/2}} \cdot \frac{\pi^{(n-1)/2}}{\Gamma\left(\frac{n}{2}+\frac{1}{2}\right)} \int_0^{\pi} \sin^n t \cdot e^{-i\xi_n \cos t} dt \\ &\overset{16.5}{=} \frac{J_{n/2}(|\boldsymbol{\xi}|)}{|\boldsymbol{\xi}|^{n/2}}\end{aligned}$$

using the **Bessel functions** of 5.16. A direct calculation results in
$\hat{f}(\mathbf{0}) = \frac{1}{(2\pi)^{n/2}} \lambda^n(B_1^n) = \frac{1}{2^{n/2}\Gamma\left(\frac{n}{2}+1\right)}$.

5. $\hat{f}(\boldsymbol{\xi}) = \frac{1}{(2\pi)^{n/2}} \int e^{-|\boldsymbol{x}|^2/2} \cdot e^{-i\langle \boldsymbol{x},\boldsymbol{\xi}\rangle} d\boldsymbol{x} = \prod\limits_{k=1}^{n} \left(\frac{1}{\sqrt{2\pi}} \int e^{-x_k^2/2} \cdot e^{-ix_k\xi_k} dx_k\right) \overset{3.}{=} \prod\limits_{k=1}^{n} e^{-x_k^2/2} = e^{-|\boldsymbol{\xi}|^2/2}$.

6. According to [3, S. 312] we have the decomposition $A = O^{-1}DO$ into an **orthogonal** $O \in \mathcal{O}(n;\mathbb{R})$ with $O^{-1} = O^T$ resp. $|\det O| = 1$ and a **diagonal** $D = (d_k\delta_{ki})_{1\leq k;i\leq n}$ containing the **positive eigenvalues** $d_k > 0$ for $1 \leq k \leq n$. Hence with $\sqrt{D} := (\sqrt{d_k}\delta_{ki})_{1\leq k;i\leq n}$ and $\left(\sqrt{D}\right)^{-1} = \left(\frac{\delta_{ki}}{\sqrt{d_k}}\right)_{1\leq k;i\leq n}$ we obtain

$$\begin{aligned}\hat{f}(\boldsymbol{\xi}) &= \frac{1}{(2\pi)^{n/2}} \int e^{-\langle \boldsymbol{x},A\boldsymbol{x}\rangle} \cdot e^{-i\langle \boldsymbol{x},\boldsymbol{\xi}\rangle} d\boldsymbol{x} \\ &\overset{\boldsymbol{y}=O\boldsymbol{x}}{=} \frac{1}{(2\pi)^{n/2}} \int e^{-\langle \boldsymbol{y},D\boldsymbol{y}\rangle} \cdot e^{-i\langle \boldsymbol{y},O\boldsymbol{\xi}\rangle} d\boldsymbol{y} \\ &\overset{\boldsymbol{z}=\sqrt{2D}\boldsymbol{y}}{=} \frac{1}{(2\pi)^{n/2}} \int e^{-|\boldsymbol{z}|^2/2} \cdot \exp\left(-i\left\langle \boldsymbol{z}, \left(\sqrt{2D}\right)^{-1} O\boldsymbol{\xi}\right\rangle\right) \cdot \frac{1}{\det\sqrt{2D}} d\boldsymbol{z} \\ &\overset{5.}{=} \frac{1}{\det\sqrt{2D}} \cdot \exp\left(-\frac{1}{2}\left|\left(\sqrt{2D}\right)^{-1} O\boldsymbol{\xi}\right|^2\right) \\ &= \frac{1}{\det\sqrt{2D}} \cdot \exp\left(-\frac{1}{4}\left\langle\left(\sqrt{D}\right)^{-1}\boldsymbol{\xi}; \left(\sqrt{D}\right)^{-1}\boldsymbol{\xi}\right\rangle\right) \\ &= \prod_{k=1}^{n} \frac{1}{\sqrt{2d_k}} \cdot e^{-x_k^2/4d_k}.\end{aligned}$$

16.8 Properties of the Fourier transforms: For **integrable** $f, g : \mathbb{R}^n \to \mathbb{C}$ we have

1. $(f \circ A)^{\wedge}(\boldsymbol{\xi}) = \frac{1}{|\det A|} \cdot \hat{f} \circ \left(A^T\right)^{-1}(\boldsymbol{\xi})$ for every **linear transformation** $A(\boldsymbol{x}) = A\boldsymbol{x}$ with $A \in GL(n; \mathbb{R})$
2. $(f \circ \sigma_\alpha)^{\wedge}(\boldsymbol{\xi}) = \frac{1}{\alpha^n} \cdot \hat{f}\left(\frac{\boldsymbol{\xi}}{\alpha}\right)$ for every **homothety** $\sigma_\alpha(\boldsymbol{x}) = \alpha\boldsymbol{x}$ with $\alpha \in \mathbb{R}$
3. $(f \circ \tau_a)^{\wedge}(\boldsymbol{\xi}) = e^{-i\langle \boldsymbol{a}, \boldsymbol{\xi}\rangle} \cdot \hat{f}(\boldsymbol{\xi})$ for every **translation** $\tau_{\boldsymbol{a}}(\boldsymbol{x}) = \boldsymbol{x} - \boldsymbol{a}$ with $\boldsymbol{a} \in \mathbb{R}^n$
4. $(f * g)^{\wedge}(\boldsymbol{\xi}) = (2\pi)^{n/2} \cdot \hat{f}(\boldsymbol{\xi}) \cdot \hat{g}(\boldsymbol{\xi})$
5. $\left(\frac{\delta f}{\delta x_i}\right)^{\wedge}(\boldsymbol{\xi}) = i\xi_i \cdot \hat{f}(\boldsymbol{\xi})$ for $f \in \mathcal{C}_c^1$
6. $(x_i \cdot f)^{\wedge}(\boldsymbol{\xi}) = i\frac{\delta \hat{f}}{\delta \xi_i}(\boldsymbol{\xi})$ for $x_i f \in L^1$
7. $\int \hat{f}(\boldsymbol{x}) \cdot g(\boldsymbol{x})\, d\boldsymbol{x} = \int f(\boldsymbol{y}) \cdot \hat{g}(\boldsymbol{y})\, d\boldsymbol{y}$
8. $\left|\widehat{g} - \widehat{f}\right| \leq \frac{1}{(2\pi)^{n/2}} \|g - f\|_1$, in particular for any sequence $(f_n)_{n\in\mathbb{N}} \subset L^1$ converging **in mean** to an $f \in L^1$ the Fourier transforms $\left(\hat{f}_n\right)_{n\in\mathbb{N}}$ converge **uniformly** to $\hat{f}$.

Proof:

1. Follows from the **change of variable theorem** 14.16.1 with $e^{-i\langle \boldsymbol{x}^T, \boldsymbol{\xi}\rangle} = e^{-i\left\langle \boldsymbol{x}^T A^T, \left(A^T\right)^{-1}\boldsymbol{\xi}\right\rangle} = e^{-i\left\langle A\boldsymbol{x}, \left(A^T\right)^{-1}\boldsymbol{\xi}\right\rangle}$.
2. As in 1. with $\left|\det\left(\frac{d\boldsymbol{x}}{d\alpha}\right)\right| = \frac{1}{\alpha^n}$ and $e^{-i\langle \boldsymbol{x}, \boldsymbol{\xi}\rangle} = e^{-i\langle \sigma_\alpha(\boldsymbol{x}), \boldsymbol{\xi}/\alpha\rangle}$.
3. As in 1. with $\left|\det\left(\frac{d\boldsymbol{x}}{d\tau_{\boldsymbol{a}}}\right)\right| = 1$ and $e^{-i\langle \boldsymbol{x}, \boldsymbol{\xi}\rangle} = e^{-i\langle \boldsymbol{a}, \boldsymbol{\xi}\rangle} \cdot e^{-i\langle \tau_{\boldsymbol{a}}(\boldsymbol{x}), \boldsymbol{\xi}\rangle}$.
4. By **Fubini's theorem** 8.5 resp. 1. we obtain

$$\begin{aligned}
(f * g)^{\wedge}(\boldsymbol{\xi}) &= \frac{1}{(2\pi)^{n/2}} \int \int f(\boldsymbol{x} - \boldsymbol{y})\, g(\boldsymbol{y})\, d\boldsymbol{y} \cdot e^{-i\langle \boldsymbol{x}, \boldsymbol{\xi}\rangle} d\boldsymbol{x} \\
&= \frac{1}{(2\pi)^{n/2}} \int \left(\int f(\boldsymbol{x} - \boldsymbol{y}) \cdot e^{-i\langle \boldsymbol{x} - \boldsymbol{y}, \boldsymbol{\xi}\rangle} d(\boldsymbol{x} - \boldsymbol{y})\right) g(\boldsymbol{y})\, d\boldsymbol{y} \cdot e^{-i\langle \boldsymbol{y}, \boldsymbol{\xi}\rangle} d\boldsymbol{x} \\
&= (2\pi)^{n/2} \cdot \hat{f}(\boldsymbol{\xi}) \cdot \hat{g}(\boldsymbol{\xi})
\end{aligned}$$

5. Since $e^{-i\langle \boldsymbol{x}, \boldsymbol{\xi}\rangle}$ is continuously differentiable we can apply **integration by parts** 14.7 to get $\left(\frac{\delta f}{\delta x_i}\right)^{\wedge}(\boldsymbol{\xi}) = \frac{1}{(2\pi)^{n/2}} \int \frac{\delta f}{\delta x_i}(\boldsymbol{x}) \cdot e^{-i\langle \boldsymbol{x}, \boldsymbol{\xi}\rangle} d\boldsymbol{x} = -\frac{1}{(2\pi)^{n/2}} \int f(\boldsymbol{x}) \cdot \frac{\delta}{\delta x_i} e^{-i\langle \boldsymbol{x}, \boldsymbol{\xi}\rangle} d\boldsymbol{x} = i\xi_i \cdot \hat{f}(\boldsymbol{\xi})$.
6. Since
   $\xi_i \mapsto f(\boldsymbol{x}) \cdot e^{-i\langle \boldsymbol{x}, \boldsymbol{\xi}\rangle}$ is **differentiable** for every $\boldsymbol{x} \in \mathbb{R}^n$
   $\boldsymbol{x} \mapsto f(\boldsymbol{x}) \cdot e^{-i\langle \boldsymbol{x}, \boldsymbol{\xi}\rangle}$ due to 5.18 is **integrable** for every $\xi_i \in \mathbb{R}$
   $\boldsymbol{x} \mapsto \left|f(\boldsymbol{x}) \cdot x_i \cdot e^{-i\langle \boldsymbol{x}, \boldsymbol{\xi}\rangle}\right|$ is a **Lebesgue integrable majorant** for $f(\boldsymbol{x}) \cdot \frac{\delta}{\delta x_i} e^{-i\langle \boldsymbol{x}, \boldsymbol{\xi}\rangle}$
   we may **exchange the order of integration and differentiation** according to 16.4 whence

$$\begin{aligned}
(x_i \cdot f)^{\wedge}(\boldsymbol{\xi}) &= \frac{1}{(2\pi)^{n/2}} \int f(\boldsymbol{x}) \cdot x_i \cdot e^{-i\langle \boldsymbol{x}, \boldsymbol{\xi}\rangle} d\boldsymbol{x} \\
&= \frac{i}{(2\pi)^{n/2}} \int f(\boldsymbol{x}) \cdot \frac{\delta}{\delta x_i} e^{-i\langle \boldsymbol{x}, \boldsymbol{\xi}\rangle} d\boldsymbol{x} \\
&= \frac{i}{(2\pi)^{n/2}} \cdot \frac{\delta}{\delta x_i} \int f(\boldsymbol{x}) \cdot \frac{\delta}{\delta x_i} e^{-i\langle \boldsymbol{x}, \boldsymbol{\xi}\rangle} d\boldsymbol{x} \\
&= i\frac{\delta \hat{f}}{\delta \xi_i}(\boldsymbol{\xi})
\end{aligned}$$

7. By **Fubini's theorem** 8.5 resp. 1. we obtain

$$\begin{aligned}\int \hat{f}(\boldsymbol{x}) \cdot g(\boldsymbol{x})\, d\boldsymbol{x} &= \frac{1}{(2\pi)^{n/2}} \int \left(\int f(\boldsymbol{y}) \cdot e^{-i\langle \boldsymbol{y},\boldsymbol{x}\rangle} d\boldsymbol{y} \right) \cdot g(\boldsymbol{x})\, d\boldsymbol{x} \\ &= \frac{1}{(2\pi)^{n/2}} \int \int f(\boldsymbol{y}) \cdot g(\boldsymbol{x}) \cdot e^{-i\langle \boldsymbol{y},\boldsymbol{x}\rangle} d\boldsymbol{y} d\boldsymbol{x} \\ &= \frac{1}{(2\pi)^{n/2}} \int \left(\int g(\boldsymbol{x}) \cdot e^{-i\langle \boldsymbol{y},\boldsymbol{x}\rangle} d\boldsymbol{x} \right) \cdot f(\boldsymbol{y})\, d\boldsymbol{y} \\ &= \int f(\boldsymbol{y}) \cdot \hat{g}(\boldsymbol{y})\, d\boldsymbol{y}\end{aligned}$$

8. Directly follows from the definition 16.6.

16.9 Hermite functions:

16.9.1: For the **Hermite polynomials** $H_n(x) = (-1)^n \cdot e^{x^2} \cdot \frac{d^n}{dx^n}\left(e^{-x^2}\right)$ with $n \in \mathbb{N}$ we have the **recursive formula** $H_{n+2}(x) = 2xH_{n+1}(x) - (2n+2)H_n(x)$ and the **differential equation** $\frac{d^2}{dx^2}H_n(x) - 2x \cdot \frac{d}{dx}H_n(x) + 2nH_n(x) = 0$.

16.9.2: The **Hermite functions** $h_n(x) = H_n(x) \cdot e^{-x^2/2}$ are the solutions of the differential equation $\frac{d^2}{dx^2}h_n(x) - x^2 h_n(x) + (2n+1)h_n(x) = 0$.

16.9.3: The **Hermite functions** h_n are **othogonal** in L^2.

16.9.4: The **Hermite functions** h_n are **Fourier eigenfunctions**, i.e. $\hat{h}_n(x) = \lambda_n \cdot h_n(x)$ with **eigenvalues** $\lambda_n \in \mathbb{C}$.

Note: By **spectral theory** the **eigenvalues** can be determined by $\lambda_n^4 - 1 = 0$ i.e. $\lambda_n \in \{\pm 1; \pm i\}$.

Proof of 16.9.1:

We first prove the **recursive formula** by **induction**:

$n = 0:$

$H_0(x) \cdot e^{-x^2} = e^{-x^2}; H_1(x) \cdot e^{-x^2} = 2x \cdot e^{-x^2}; H_2(x) \cdot e^{-x^2} = (-2+4x^2) \cdot e^{-x^2} = (2xH_1(x) - 2H_0(x)) \cdot e^{-x^2}$

$n \Rightarrow n+1:$

$$\begin{aligned}H_{n+3}(x) \cdot e^{-x^2} &= -\frac{d}{dx}\left(H_{n+2}(x) \cdot e^{-x^2}\right) \\ &= -\left(\frac{d}{dx}H_{n+2}(x) - 2xH_{n+2}(x)\right) \cdot e^{-x^2} \\ &= -\left(\frac{d}{dx}(2xH_{n+1}(x) - (2n+2)H_n(x)) - 2xH_{n+2}(x)\right) \cdot e^{-x^2} \\ &= \left(2xH_{n+2}(x) - 2H_{n+1}(x) - 2x\frac{d}{dx}H_{n+1}(x) + (2n+2)\frac{d}{dx}H_n(x)\right) \cdot e^{-x^2} \\ &= (2xH_{n+2}(x) - 2H_{n+1}(x) - 2x(2xH_{n+1}(x) - H_{n+2}(x)) + (2n+2)(2xH_n(x) - H_{n+1}(x))) \cdot e^{-x} \\ &= (2xH_{n+2}(x) - (2n+4)H_{n+1}(x) + 2x[H_{n+2}(x) + 2xH_{n+1}(x) - (2n+2)H_n(x)]) \cdot e^{-x^2} \\ &= (2xH_{n+2}(x) - (2n+4)H_{n+1}(x) + 0) \cdot e^{-x^2}\end{aligned}$$

The **differential equation** then follows from $H_{n+1}(x) \cdot e^{-x^2} = -\frac{d}{dx}\left(H_n(x) \cdot e^{-x^2}\right) = -\left(\frac{d}{dx}H_n(x) - 2xH_n(x)\right) \cdot e^{-x^2}$, i.e.

$\frac{d}{dx}H_n(x) = 2xH_n(x) - H_{n+1}(x)$

and

$H_{n+2}(x) \cdot e^{-x^2} = \frac{d^2}{dx^2}\left(H_n(x) \cdot e^{-x^2}\right) = \left(\frac{d^2}{dx^2}H_n(x) - 4x\frac{d}{dx}H_n(x) + (4x^2-2)H_n(x)\right) \cdot e^{-x^2}$, i.e.

$\frac{d^2}{dx^2}H_n(x) = H_{n+2}(x) + 4x\frac{d}{dx}H_n(x) - (4x^2-2)H_n(x) = H_{n+2}(x) - 4xH_{n+1}(x) + (4x^2+2)H_n(x)$
whence by the recursive formula we obtain

$$\begin{aligned}\frac{d^2}{dx^2}H_n(x) - 2x\cdot\frac{d}{dx}H_n(x) + 2nH_n(x) &= H_{n+2}(x) - 4xH_{n+1}(x) + \left(4x^2+2\right)H_n(x) - 4x^2H_n(x) + 2xH_{n+1}(x) + \\ &= H_{n+2}(x) - 2xH_{n+1}(x) + (2n+2)H_n(x) \\ &= 0\end{aligned}$$

Proof of 16.9.2:
With $\frac{d}{dx}h_n(x) = \left(\frac{d}{dx}H_n(x) - x\cdot H_n(x)\right)\cdot e^{-x^2/2}$ and $\frac{d^2}{dx^2}h_n(x) = \left(\frac{d^2}{dx^2}H_n(x) - 2x\frac{d}{dx}H_n(x) + (x^2-1)\cdot H_n(x)\right)\cdot e^{-x^2/2}$ we have

$$\begin{aligned}\frac{d^2}{dx^2}h_n(x) + \left(\lambda_n - x^2\right)h_n(x) &= \left(\left(\frac{d^2}{dx^2}H_n(x) - 2x\frac{d}{dx}H_n(x) + \left(x^2-1\right)\cdot H_n(x)\right) + \left(2n-1-x^2\right)H_n(x)\right)\cdot e^{-x} \\ &= \left(\frac{d^2}{dx^2}H_n(x) - 2x\frac{d}{dx}H_n(x) + 2nH_n(x)\right)\cdot e^{-x^2/2} \\ &= 0\end{aligned}$$

Proof of 16.9.3:

By **integration by parts** we obtain $\langle h_n; h_m\rangle = \int h_n(x)\cdot h_m(x)\,dx = \int H_n(x)\cdot H_m(x)\cdot e^{-x^2}dx = \delta_{nm}$.

Proof of 16.9.4:

By 16.8.5 and 16.8.6 the differential equation $\frac{d^2}{dx^2}h_n(x) - x^2h_n(x) + (2n+1)h_n(x) = 0$ transforms to the identical equation $0 = \left(\frac{d^2}{dx^2}h_n - x^2h_n + (2n+1)h_n\right)^\wedge(\xi) = -\xi^2\hat{h}_n(\xi) + \frac{d^2}{d\xi^2}\hat{h}_n(\xi) + (2n+1)\hat{h}_n(\xi) = 0$ having the same solutions up to a complex constant, i.e.$h_n = \lambda_n\cdot\hat{h}_n$ with $\lambda_n\in\mathbb{C}$.

16.10 Corollary: Let C_0 the vector space of all continuous functions $f:\mathbb{R}^n\to\mathbb{C}$ **vanishing at infinity**, i.e. $\lim\limits_{|\boldsymbol{x}|\to\infty}|f(\boldsymbol{x})| = 0$. Hence we have $C_c\subset C_0\subset C$ and also the following inclusions:

1. For every $f\in\mathcal{C}_c^k$ with $k\in\mathbb{N}$ there exists an $M>0$ such that $\left|\hat{f}(\xi)\right|\le\frac{M}{(1+|\xi|)^k}$ for every $\xi\in\mathbb{R}^n$ and particularly $\hat{f}\in C_0^k\cap L^1$.
2. For every $f\in L^1$ we have $\hat{f}\in C_0$.

Proof:

1. We define **multiindices $\boldsymbol{\alpha}\in\mathbb{N}^n$** with $|\boldsymbol{\alpha}| = \sum\limits_{i=1}^n\alpha_i$ for use in situations as e.g. $\boldsymbol{x}^\alpha = \prod\limits_{i=1}^n x_i^{\alpha_i}$ and $\frac{\delta^{\boldsymbol{\alpha}}f}{\delta\boldsymbol{x}^{\boldsymbol{\alpha}}} = \frac{\delta^{|\boldsymbol{\alpha}|}f}{\delta^{\alpha_1}x_1\cdot\ldots\cdot\delta^{\alpha_n}x_n}$. Hence owing to 16.8.5 for $|\boldsymbol{\alpha}|\le k$ we have $\left(\frac{\delta^{\boldsymbol{\alpha}}f}{\delta\boldsymbol{x}^{\boldsymbol{\alpha}}}\right)^\wedge(\boldsymbol{\xi}) = i^{|\boldsymbol{\alpha}|}\cdot\boldsymbol{\xi}^{\boldsymbol{\alpha}}\cdot\hat{f}(\boldsymbol{\xi})$ whence $\left|\boldsymbol{\xi}^{\boldsymbol{\alpha}}\cdot\hat{f}(\boldsymbol{\xi})\right|\le\frac{1}{(2\pi)^{n/2}}\left\|\frac{\delta^{\boldsymbol{\alpha}}f}{\delta\boldsymbol{x}^{\boldsymbol{\alpha}}}\right\|_1<\infty$ since $\mathcal{C}_c^k\subset L^1$. Consequently there is an $M>0$ such that $(1+|\xi_1|+\ldots+|\xi_n|)^k\left|\hat{f}(\boldsymbol{\xi})\right|\le M$ for all $\xi\in\mathbb{R}^n$ which proves the assertion.

2. According to 14.5.2 there is a $g\in C_c^1$ with $\|f-g\|_1<\epsilon$ whence $\left|\hat{f}(\xi)-\hat{g}(\xi)\right|<\frac{\epsilon}{(2\pi)^{n/2}}$ for every $\xi\in\mathbb{R}^n$ so that the assertion follows from 1.

16.11 Note: According to 16.17.2 the Fourier transform $\hat{f}(\xi) = \frac{\sin\xi}{\xi}$ from 16.7.1 is **integrable** with $\int\frac{\sin\xi}{\xi}d\xi = \pi$ but **not Lebesgue integrable** any more, since for every $k\ge 1$ we have $\int_{k\pi}^{(k+1)\pi}\left|\frac{\sin\xi}{\xi}\right|\ge\frac{2}{k\pi}$ whence $\int_{N\pi}^{(2N+1)\pi}\frac{\sin\xi}{\xi}d\xi\ge\frac{2}{\pi}\sum\limits_{k=N}^{2N}\frac{1}{k} = \frac{2}{\pi N}\sum\limits_{k=1}^{N}\frac{N}{N+k}\ge=\frac{2}{\pi N}\sum\limits_{k=1}^{N}\frac{1}{2}\ge\frac{1}{\pi}$ for every $N\ge 1$.

16.12 Lemma: For every $f,\psi\in L^1$ with $\int\psi(\boldsymbol{x})\,d\boldsymbol{x} = 1$ and $\psi_\alpha(\boldsymbol{\xi}) = \frac{1}{\alpha^n}\psi\left(\frac{\boldsymbol{\xi}}{\alpha}\right)$ we have $\lim\limits_{\alpha\to 0}\|f-f*\psi_\alpha\|_1 = 0$.

Proof: Due to 14.5.2 it suffices to prove the assertion for $f\in\mathcal{C}_c^\infty$. Hence there is an $R>0$ such that $\operatorname{supp}(f)<B_R(\mathbf{0})$ and for every $\epsilon>0$ we can find a $\delta>0$ with $|f(\boldsymbol{x})-f(\boldsymbol{y})|<\frac{\epsilon}{2\cdot\lambda^n(B_{R+1}(\mathbf{0}))\cdot\|\psi\|_1}$ for every $x, y \in \mathbb{R}^n$ with $|\boldsymbol{x} - \boldsymbol{y}| < \delta$. Owing to the **monotone convergence** theorem 5.12 and $|\psi| \in L^1$ exists an $\alpha > 0$ such that $\int_{|\boldsymbol{\xi}|>\delta} |\psi_\alpha(\boldsymbol{\xi})|\, d\boldsymbol{\xi} \overset{\boldsymbol{\xi}=\alpha\boldsymbol{x}}{=} \int_{|\boldsymbol{x}|>\delta/\alpha} |\psi(\boldsymbol{x})|\, d\boldsymbol{x} < \frac{\epsilon}{4\|f\|_1}$ we have

$$\begin{aligned}
\|f - f * \psi_\alpha\|_1 &\leq \int_{|\boldsymbol{x}|\leq R} \left(\int_{|\boldsymbol{\xi}|\leq\delta} |f(\boldsymbol{x}) - f(\boldsymbol{x} - \boldsymbol{\xi})| \cdot |\psi_\alpha(\boldsymbol{\xi})|\, d\boldsymbol{\xi} \right) d\boldsymbol{x} \\
&\quad + \int_{|\boldsymbol{x}|\leq R} \left(\int_{|\boldsymbol{\xi}|>\delta} |f(\boldsymbol{x}) - f(\boldsymbol{x} - \boldsymbol{\xi})| \cdot |\psi_\alpha(\boldsymbol{\xi})|\, d\boldsymbol{\xi} \right) d\boldsymbol{x} \\
&\leq \lambda^n(B_{R+1}(\mathbf{0})) \cdot \sup_{|\boldsymbol{x}-\boldsymbol{y}|<\delta} |f(\boldsymbol{x}) - f(\boldsymbol{y})| \cdot \|\psi\|_1 + 2\|f\|_1 \cdot \int_{|\boldsymbol{\xi}|>\delta} |\psi_\alpha(\boldsymbol{\xi})|\, d\boldsymbol{\xi} \\
&\leq \frac{\epsilon}{2} + \frac{\epsilon}{2} \\
&= \epsilon.
\end{aligned}$$

16.13 Fourier inversion formula: For **Lebesgue integrable** $f, \hat{f} \in L^1$ we have λ^n-a.e. $f(\boldsymbol{x}) = \int \hat{f}(\boldsymbol{\xi}) \cdot e^{i\langle\boldsymbol{\xi};\boldsymbol{x}\rangle} d\boldsymbol{\xi}$, i.e. $f = \hat{\hat{f}} \circ \sigma_{-1}$ resp. $\overline{f} = \overline{\hat{\hat{f}}}$.

Proof: According to 16.10.2 we have $\hat{\hat{f}} \circ \sigma_{-1} \in C_0$. Also for $\psi(\xi) = \frac{1}{(2\pi)^{n/2}} \cdot e^{-|\boldsymbol{\xi}|^2/2}$ holds $\lim_{\alpha\to 0} \psi(\alpha\boldsymbol{\xi}) = 1$ with monotonically increasing $(\psi(\alpha\boldsymbol{\xi}))_{\alpha>0}$ for every $\boldsymbol{\xi} \in \mathbb{R}^n$ such that we can use **monotone convergence** to obtain

$$\begin{aligned}
\int \hat{f}(\boldsymbol{\xi}) \cdot e^{i\langle\boldsymbol{\xi};\boldsymbol{x}\rangle} d\boldsymbol{\xi} &\overset{5.12}{=} \lim_{\alpha\to 0} \left(\int \hat{f}(\boldsymbol{\xi}) \cdot e^{i\langle\boldsymbol{\xi},\boldsymbol{x}\rangle} \cdot \psi(\alpha\boldsymbol{\xi})\, d\boldsymbol{\xi} \right) \\
&\overset{16.8.1}{=} \lim_{\alpha\to 0} \left(\int (f \circ \tau_{-\boldsymbol{x}})^\wedge(\boldsymbol{\xi}) \cdot (\psi \circ \sigma_\alpha)(\boldsymbol{\xi})\, d\boldsymbol{\xi} \right) \\
&\overset{16.8.6}{=} \lim_{\alpha\to 0} \left(\int (f \circ \tau_{-\boldsymbol{x}})(\boldsymbol{\xi}) \cdot (\psi \circ \sigma_\alpha)^\wedge(\boldsymbol{\xi})\, d\boldsymbol{\xi} \right) \\
&\overset{16.8.2}{=} \lim_{\alpha\to 0} \left(\int f(\boldsymbol{\xi} + \boldsymbol{x}) \cdot \frac{1}{\alpha^n} \cdot \hat{\psi}\left(\frac{\boldsymbol{\xi}}{\alpha}\right) d\boldsymbol{\xi} \right) \\
&\overset{16.7.5}{=} \lim_{\alpha\to 0} \left(\int f(\boldsymbol{\xi} + \boldsymbol{x}) \cdot \frac{1}{\alpha^n} \cdot \psi\left(\frac{\boldsymbol{\xi}}{\alpha}\right) d\boldsymbol{\xi} \right) \\
&= \lim_{\alpha\to 0} \left(\int f(\boldsymbol{x} - \boldsymbol{\xi}) \cdot \frac{1}{\alpha^n} \cdot \psi\left(-\frac{\boldsymbol{\xi}}{\alpha}\right) d\boldsymbol{\xi} \right) \\
&= \lim_{\alpha\to 0} (f * \psi_\alpha)(\boldsymbol{x}) \\
&= f(\boldsymbol{x}) \ \lambda^n\text{-a.e.}
\end{aligned}$$

since due to the preceding lemma 16.12 we have $\lim_{\alpha\to 0} \|f - f * \psi_\alpha\|_1 = 0$ such that according to 5.10 a subsequence converges λ^n-a.e. to f.

16.14 Examples: The fourier inversion yields some convenient formulae for integrals of real valued trigonometric functions with d $\operatorname{Re}(e^{ixt}) = \cos(tx)$:

1. According to 16.2, 16.7.1 and 16.8.4 we have $\sup\{0; 2 - |x|\} = \frac{2}{\pi} \int \left(\frac{\sin x}{x}\right)^2 \cos(x\xi)\, dx$ and in particular for $\xi = 0$ we obtain $\int \left(\frac{\sin x}{x}\right)^2 dx = \pi$.
2. From 16.7.2 follows $\int \frac{\cos(tx)}{1+x^2} dx = \pi e^{-|t|}$.

16.15 Lemma: For every $f \in L^1 \cap L^2$ and every $\epsilon > 0$ there is a $\varphi \in \mathcal{C}_C^\infty$ such that $\|f - \varphi\|_1 < \epsilon$ **and** $\|f - \varphi\|_2 < \epsilon$.

Proof: According to 14.5.2 for $\epsilon > 0$ and $h_a(\boldsymbol{x}) = e^{-\alpha|\boldsymbol{x}|^2}$ there is a $\psi \in \mathcal{C}_C^\infty$ with $\|f - \psi\|_2 < \min\left\{\frac{\epsilon}{2}, \frac{\epsilon}{2\|h_a\|_2}\right\}$. Since $h_\alpha \uparrow 1$ for $\alpha \to 0$ we can invoke **monotone convergence** 5.12 to find an $\alpha > 0$ such that $\|f - f \cdot h_\alpha\|_p^p = \int |f(x)|^p \cdot |1 - h_\alpha(x)|^p\, dx < \frac{\epsilon}{2}$. Hence on the one hand we have $\|f \cdot h_\alpha - \psi \cdot h_\alpha\|_2 = \|f - \psi\|_2 \cdot \|h_\alpha\|_2 < \|f - \psi\|_2 < \frac{\epsilon}{2}$ and owing to **Hölder's inequality** 6.4.1 on the other hand there is $\|f \cdot h_\alpha - \psi \cdot h_\alpha\|_1 \leq \|f - \psi\|_2 \cdot \|h_\alpha\|_2 < \frac{\epsilon}{2}$. The assertion now follows with the **triangle inequality** for $p = 1; 2$ for $\varphi = \psi \cdot h_\alpha$.

16.16 Plancherel's theorem: There is a L^2**-norm-preserving isomorphism** $\hat{}: L^2 \to L^2$ with $\|f\|_2 = \left\|\hat{f}\right\|_2$ for $f \in L^2$ and $\hat{f}(\xi) = \frac{1}{\sqrt{2\pi}} \int f(x) \cdot e^{-i\langle x,\xi\rangle} dx$ for $f \in L^1 \cap L^2$.

Proof: According to the preceding lemma 16.15 for every $f \in L^1 \cap L^2$ there is a sequence $(f_k)_{k\in\mathbb{N}} \subset C_C^\infty$ converging in the **first and second mean** to f. Due to 16.10.1 we have $\left(\hat{f}_{k\in\mathbb{N}}\right)_{k\in\mathbb{N}} \subset L^1$and with 16.8.6 follows $\left\|\hat{f}_k\right\|_2^2 = \int \hat{f}_k \overline{\hat{f}_k} d\lambda^n = \int \hat{f}_k \hat{\overline{\hat{f}}}_k d\lambda^n = \int f_k \overline{f}_k d\lambda^n = \|f_k\|_2^2$. Hence $\left(\widehat{f_k}\right)_{k\in\mathbb{N}}$ is L^2-Cauchy, which due to 6.7 converges **in the second mean** to a $g \in L^2$ and according to 6.9 a subsequence converges λ^n-a.e. to g. But owing to 16.8.8 the entire sequence $\left(\widehat{f_k}\right)_{k\in\mathbb{N}}$ **uniformly** converges to $\hat{f}$ which means λ^n-a.e. $g = \hat{f} \in L^2(\mathbb{R}^n)$. Also we have $\left\|\hat{f}\right\|_2 = \lim_{k\to\infty} \left\|\hat{f}_k\right\|_2 = \lim_{k\to\infty} \|f_k\|_2 = \|f\|_2$. Thus we have shown that the Fourier transform$\hat{}: L^1 \cap L^2 \to L^2$ is L^2-norm preserving. Since L^2 is complete and $C_C^\infty \subset L^1 \cap L^2 \subset L^2$ according to 14.5.2 is L^2- dense in L^2 it can be extended to L^2 as usual by assigning to every $f \in L^2$ as Fourier transform the L^2- resp. λ^n-a.e. limit $\hat{f} := \lim_{k\to\infty} \hat{f}_k \in L^2$ of the L^2-Cauchy sequence $\left(\hat{f}_k\right)_{k\in\mathbb{N}} \subset L^2$ of Fourier transforms of any sequence $(f_k)_{k\in\mathbb{N}} \subset L^1 \cap L^2$ converging in the second mean to f. Owing to the L^2-norm preserving character of the Fourier transform and in particular $\left\|\hat{f}_k\right\|_2 = \|f_k\|_2$ resp. the **positive definiteness** of the norm due to 5.8 the Fourier transform $\hat{f}$ is λ^n-a.e. determined and in this sense independent of the approximating sequence. The same argument applies to show that the Fourier transform is **injective** and finally it is **surjective** since the set of all Fourier transforms is closed with respect to the L^2-norm and includes the L^2- dense subset C_C^∞.

Note: In the proof a sequence $(f_k)_{k\in\mathbb{N}} \subset L^1 \cap L^2$ converging in L^2 and λ^n-a-e- to f is used to define the extended Fourier transform $\hat{f} := \lim_{k\to\infty} \hat{f}_k \in L^2$. In the following example we will use the inverse direction and take a sequence $\left(\hat{f}_k\right)_{k\in\mathbb{N}} \subset L^1 \cap L^2$ converging in L^2 and λ^n-a-e- to $\hat{f}$ to find its inverse $f := \lim_{k\to\infty} \hat{\overline{\hat{f}}}_k \in L^2$:

16.17 Examples:

1. $\int_0^\infty \left|J_{n/2}(r)\right|^2 \frac{dr}{r} = \frac{1}{n}$ for every $n \geq 1$ since on the one hand we have $\left\|\hat{\chi}_{B_1^n}\right\|_2^2 \overset{16.7.4}{=} \frac{J_{n/2}^2(|\boldsymbol{\xi}|)}{|\boldsymbol{\xi}|^n} d\boldsymbol{\xi} \overset{15.5}{=} n \cdot \tau_n \cdot \int_0^\infty \left|J_{n/2}(r)\right|^2 \frac{dr}{r}$ and on the other hand $\left\|\chi_{B_1^n}\right\|_2^2 \overset{8.13}{=} \tau_n$ so that the formula follows from **Plancherel's theorem** 16.16.
2. We want to show that $\hat{\overline{\hat{f}}} = f$ and thereby compute the integral $\int \frac{\sin\xi}{\xi} d\xi$: The Fourier transform $\hat{f}(\xi) = \sqrt{\frac{2}{\pi}} \frac{\sin\xi}{\xi} \in L^2$ of $f(x) = \chi_{[-1;1]}(x) \in L^2$ from 16.7.1 is **integrable** due to 5.24 and **square integrable** owing to 16.14.1 but **not Lebesgue integrable**. Hence we cannot directly apply the **inversion formula** 16.13 and have to invoke **Plancherel's theorem** 16.16: The sequence $\left(\hat{f}_k\right)_{k\geq1}^2 \subset L_1$ with $\hat{f}_k = \hat{f} \cdot \chi_{[-k;k]}$ converges to $\hat{f}^2$ **pointwise** and owing to **monotone convergence** 5.12 the sequence $\left(\hat{f}_k\right)_{k\geq1} \subset L_1 \cap L_2$ converges to $\hat{f} \in L^2$ in **quadratic mean**. Due to **Plancherel's theorem** 16.16 the sequence $(f_k)_{k\geq1} = \left(\hat{\overline{\hat{f}}}_k\right)_{k\geq1} \subset L_1 \cap L_2$ converges in quadratic mean and due to 6.9 also λ^n-a.e. to $\chi_{[-1;1]}(x) = f(x) = \lim_{k\to\infty} f_k(x) = \lim_{k\to\infty} \hat{\overline{\hat{f}}}_k(x) = \lim_{k\to\infty} \left(\frac{1}{\sqrt{2\pi}} \int \chi_{[-k;k]}(\xi) \cdot \hat{f}(\xi) \cdot e^{ix\xi} d\xi\right) = \frac{1}{\pi} \cdot \lim_{k\to\infty} \int \chi_{[-k;k]}(\xi) \cdot \frac{\sin\xi}{\xi} \cdot \cos(x\xi)\, d\xi \overset{5.24}{=} \frac{1}{\pi} \cdot \int \frac{\sin\xi}{\xi} \cdot \cos(x\xi)\, d\xi$. For $x = 0$ we obtain $1 = \frac{1}{\pi} \cdot \lim_{k\to\infty} \int \frac{\sin\xi}{\xi} d\xi$ whence $\int \frac{\sin\xi}{\xi} d\xi = \pi$.

17 The central limit theorem

17.1 Characteristic funcions:

Index

$0 \cdot \infty$, 4

Abel's partial summation, 34
absolute convergence, 35
absolute homogeneity, 20
absolute value, 8, 16
absolutely continuous, 51
absolutely continuous measure, 50
addition, 8
additivity, 13, 33
algebra, 3, 24, 25, 31, 42
almost everywhere, 7
approximating sequence, 14, 51
approximation property, 6, 27
arithmetic mean, 20
axiom of choice, 7

Banach algebra, 8, 61
Banach space, 8, 15, 21, 36, 39, 44, 48
base, 30
basis, 8
basis of a σ-algebra, 3, 24
Bessel differential equation, 63
Bessel function, 65
Bessel functions, 63
Beta function, 59
Bochner integral, 14, 16, 17
Bochner space, 14
Bohr-Mollerup theorem, 59, 60
Borel measurable, 3
Borel measurable function, 8, 45
Borel measure, 40
Borel sets, 7
Borel σ-algebra, 3, 6, 43
Borel's zero-one-law, 32
Borel-Cantelli lemma, 10, 35
Borel-measurable, 20
bounded, 20
bounded linear functional, 37, 38
Brouwer's fixed point theorem, 52

canonical injection, 24
Cantor set, 7
Carathéodory's theorem, 5
cartesian coordinates, 57
Cauchy sequence, 10, 11, 13, 21
Cavalieri's principle, 29, 30
center of gravity, 61
centered random variable, 33
Chain rule, 54
chain rule, 54, 56, 59
change of variable, 66
change of variables, 49
Change-of-variables theorem, 56
characteristic function, 33
characteristic functions, 12
Chebyshev's inequality, 32
closed set, 3
compact, 3, 25, 48, 51
compact convergence, 10
compact support, 39, 52
complete measure, 7, 43, 50
complete space, 11, 15
Completeness, 10, 21
completeness, 19, 22, 34
complex conjugate, 8
complex measure, 35
component, 24
components, 14
composition, 8
concentrated measure, 36
cone, 30
continuity form above, 14
continuity from above, 4, 7, 10–12, 33, 38, 39, 55
continuity from below, 4, 11, 17, 26
continuity of the integral, 16
Continuity under the integral sign, 62
continuous, 6, 20, 25, 37, 46
continuous from above, 10
continuous function, 8
contiuity of the integral, 19
convergence in p-th mean, 22
convergence in mean, 13, 16
convergence in measure, 9
convergence in the p-th mean, 20, 23
convergence μ-a.e., 23
convergence μ-almost everywhere, 9
convex, 20, 59
convolution, 61
countable at infinity, 3
cube, 52
cut, 26, 31, 61
cylinder, 57
cylinder coordinates, 57
cylinder set, 24, 26, 31
cylinder sets, 9, 27

Darboux sum, 19
δ-system, 3
dense, 7, 12
density, 60
derivative, 46

derivative of a measure, 49
diagonal matrix, 65
difference quotients, 59
Differentiability under the integral sign, 63
differentiable, 46, 52
Differentiable partition of unity, 52
differential equation, 67
dilation, 8
Dini derivatives, 45
Dini's theorem, 41
Dirac measure, 5, 40
discrete space, 3
domain, 4
dominated convergence, 16, 18, 19, 60, 62–64
dual space, 37–39
Dynkin δ-π-theorem, 4, 26
Dynkin system, 3, 6, 26
Dynkin-system, 32

Egorov's convergence theorem, 12, 23
eigenvalues, 64, 65, 67
elementary matrix, 29
elementary transformation, 29
ellipsoid, 61
$\emptyset$-continuity, 4, 5, 31
equivalence classes, 15
equivalence relation, 7, 12, 20
euclidean metric, 5
expected value, 32
exponential function, 10
extension, 31
extension theorem, 32

Fatou's lemma, 17, 19, 21, 23
Fatou's lemma:, 16
field, 8
figure, 5, 6, 49
figures, 3, 17
finite dimensional, 22
finite measure, 10
finitely additive, 4
first category, 7
first countable, 9
Fourier eigenfunctions, 67
Fourier inversion formula, 69
Fourier transform, 64
Fubini's differentiable series theorem, 48
Fubini's theorem, 27, 29, 54, 58, 61, 66, 67
functional equation, 30, 59
fundamental theorem of calculus, 49, 51, 52, 54, 58

Gamma function, 30, 59, 61
Gauss integral, 61
geometric mean, 20
global convergence in measure, 10

Hahn decomposition, 38
Hahn's extension theorem, 6
Hausdorff, 25
Hausdorff space, 3
height, 30
Heine Borel theorem, 54
Heine-Borel theorem, 40
Helmholtz equation, 63
Hermite functions, 67
Hermite polynomials, 67
Hilbert space, 22, 39
homomorphism, 28
homothety, 66
Hölder inequality, 20, 39
Hölder's inequality, 59, 61, 69

image of a measure, 8
image of a σ-algebra, 8
imaginary part, 8
improper Riemann integral, 16, 19
independent, 32
independent random variables, 33
independent sets, 10
independent σ-algebrae, 33
induction, 6, 28, 30, 46
infinitely differentiable, 52
Initial σ-algebra, 24
initial σ-algebra, 24, 25
initial topology, 24
injective, 70
inner product, 22, 37
inner regular, 40
integrable, 19, 20, 64
integrable functions, 14
integral function, 51
Integration by parts, 54
integration by parts, 49, 59, 64, 66, 68
integration by substitution, 59
invariant, 8
inverse image of a σ-algebra, 8
invertible matrix, 29
isometric, 40, 44
isometry, 39
isomorphic, 40
isomorphism, 44, 70

Jacobian matrix, 52
Jensen's inequality, 20
Jordan decomposition, 37, 38, 52
Jordan decomposition of functions, 48

Khintchin-Kolmogorov convergence theorem, 34, 35
Kolmogorov's strong law of large numbers, 35
Kronecker's lemma, 34, 35

Laplace equation, 63
Lebesgue differentiation theorem for complex functions, 48
Lebesgue integrable, 14
Lebesgue integral, 17, 50
Lebesgue intgrable, 51
Lebesgue measurable, 45
Lebesgue measurable function, 45
Lebesgue measure, 7, 10, 28, 40, 45, 49
Lebesgue point, 50, 51
Lebesgue Radon-Nikodym theorem, 37
Lebesgue sets, 7
Lebesgue σ-algebra, 28
Lebesgue space, 14
Lebesgue's convergence theorem, 10, 34
Lebesgue's differentiation theorem, 45, 47
Lebesgue's dominated convergence theorem:, 16, 22
Lebesgue-Borel measure, 7
Lebesgue-Borel pre-measure, 5
Lebesgue-Borel-Stieltjes measure, 6, 49
Lebesgue-Borel-Stieltjes pre-measure, 5
Lebesgue-integrable, 16
Lebesgue-Radon-Nikodym theorem, 39, 49
left derivative, 46
Levi's monotone convergence theorem, 15
linear functional, 37, 39
linear order, 57
linear transformation, 66
linearity, 20, 33
linearly ordered, 49
local convergence in measure, 10
locally compact, 3
lower semicontinuous, 5, 50
Lusin's Theorem, 40
Lusin's theorem, 45
Lévy und Steinitz, theorem of, 35
Lévy's convergence theorem, 33, 35
Lévy's inequality, 33

mass, 60
F$_\sigma$-set, 3
G$_\delta$-set, 25
G$_\delta$-set, 3
maximal function of a function, 50
maximal function of a measure, 49, 50
maximal median, 33
maximum, 8
Mean value theorem, 48
mean value theorem, 38, 52, 54, 60, 63
mean value theorem for integration, 17, 18, 52
mean value theorem for vector spaces, 54
measurable function, 8
measurable rectangle, 24, 26, 31
measurable set, 5
measurable space, 3, 8
measure, 5
measure space, 5
median, 33
metric, 8
metric space, 8
minimal median, 33
minimum, 8
Minkowski inequality, 20
Minkowski's inequality, 23
Moments of inertia, 60
Monotone convergence, 28
monotone convergence, 16, 22, 23, 26, 33, 38, 61, 69, 70
monotone convergence theorem, 39, 45
monotonicity, 13, 36
monotonicity of a pre-measure, 4
monotonicity of the integral, 14, 15
μ-absolutely continuous measure, 36
mulitplicity of the determinant, 29
multiple, 8
multiplication, 8
multiplicity, 28, 31
multiplicity of a measure, 26
mutually singular measures, 36

n-dimensional figure, 25
n-dimensional interval, 25, 28
negative part, 8, 14
neighbourhood basis, 46
Non measurable sets, 7
norm, 8, 20, 22, 36, 38, 48
norm-preserving, 70
nowhere dense, 7
null set, 7

open, 7
open set, 3, 8
orthogonal matrix, 28, 65
outer measure, 5, 41
outer measure, 6
outer regular, 40

parallelepiped, 29
partial derivatives, 52
partial fraction decomposition, 53
partition, 21, 60
partition of unity, 43

path, 31
paths, 31
permutation, 28
π-basis, 28
π-basis, 6
π-system, 4, 32
Plancherel's theorem, 70
pointwise convergence, 10
polar representation of complex measures, 38
polish space, 25
polynomials, 53
positive definite, 64
positive definiteness, 15, 36, 70
positive functional, 40, 41
positive measure, 35
positive part, 8, 14
positive und negative variations, 37
powers, 8
pre-measure, 4, 32, 42
probability distribution, 33
probability measure, 10, 31
probability measures, 5
probability space, 5, 31
probability theory, 7
product measure, 26
Product rule, 54
product rule, 54
product space, 8
product topology, 9, 24, 25
product-σ-algebra, 24
projection, 8, 24, 25
projections, 14
proprtional distortion, 28
pseudonorm, 13, 15, 20
Pythagoras, 30

quotient space, 12, 15, 20

Radon-Nikodym density, 37, 38
Radon-Nikodym derivative, 37, 50
Radon-Nikodym theorem, 52, 56
random variables, 32
range, 4
rational numbers, 7
ray, 3
real numbers, 4, 19
real part, 8
realization, 31
reciprocal, 8
recursive formula, 67
regular, 40, 45
regular measure, 49, 50
regularity, 46
restriction, 14
restriction of a measure, 24
Riemann integrable, 19
Riemann integral, 19
Riesz convergence theorem, 10, 34
Riesz representation theorem, 28, 40, 43, 44, 49
right derivative, 46
right-open interval, 3, 5, 7, 49
Rolle's theorem, 48
rotation, 28
rotational symmetric functions, 60

sample path, 31
scaling, 28
Schwarz inequality, 37
second countability, 51
second countable, 3, 25
semicontinuous, 8
separable, 9, 12, 14, 16–19
separating function, 53
separation axiom, 25
separation property, 42, 53
shearing, 29
σ-additive, 4, 42, 43
σ-additivity, 35
σ-algebra, 3
σ-compact, 45, 49
σ-finite, 6, 12, 28, 40, 45, 49
σ-finite, 6
σ-finite measure, 38
σ-regular, 40, 49
signed measure, 37
simplex, 30
smooth, 52
spectral theory, 67
sphere, 58, 60, 61
spherical coordinates, 58
standard deviation, 33
Steiner's theorem, 61
step function, 12, 13, 27, 33, 37, 39, 43, 57
step functions, 23, 39, 51, 60
Stirling's formula, 60
stochastic convergence, 10
stochastic process, 31
subadditivity, 4, 11, 41
support, 12
supports, 13
supremum norm, 13, 20, 39
supremum property, 4, 5
surjective, 70
symmetric, 64
symmetric derivative, 50

three dimensions, 29
topogical basis, 3

topological space, 3, 8
total variation, 35, 48, 49, 51
trace algebra, 17
trace σ-algebra, 3, 24
trace topology, 3
transformation formula, 28, 45
translation, 8, 66
translation invariance, 45, 61
translation invariant, 28
tree, 31
triangle inequality, 13, 70
triangle inequation, 20
truncated random variables, 35
tube, 61
Tychonoff's theorem, 25

uniform convergence, 12, 19, 48
uniformly continuous, 37, 51
uniqueness theorem, 6, 49
unit cube, 29
unit sphere, 30, 60
upper derivative, 50

vanishing at infinity, 68
variation of a measure, 40
vector space, 20
Vitali, 7
Vitali cover, 46, 47
Vitali's convergence theorem, 23
Vitali's covering theorem, 46
volume, 30, 58

well ordering, 14

Zorn's lemma, 3, 21

References

[1] Heinz Bauer. *Maß- und Integrationstheorie.* De Gruyter, 1990.

[2] Karl Stromberg Edwin Hewitt. *Real and Abstract Analysis.* Springer, 1965.

[3] Gerd Fischer. *Lineare Algebra.* Vieweg, 2000.

[4] Otto Forster. *Analysis III.* Vieweg, 1984.

[5] Serge Lang. *Real and Functional Analysis.* Springer, 1996.

[6] Gerhard Preuss. Order convergence and convergence almost everywhere revisited. *Refubium . Repositorium der freien Universität Berlin*, 2010.

[7] Walter Rudin. *Principles of mathematical analysis.* McGraw Hill, 1976.

[8] Walter Rudin. *Functional Analysis.* McGraw Hill, 1991.

[9] Arne Vorwerg. Topology. 2020.

[10] Arne Vorwerg. Bedingte konvergenz von vektorsummen. 2021.

[11] Arne Vorwerg. Mengenlehre. 2022.