Measure and Integral

Arne Vorwerg

December 2022

Preface

Measure theory 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**. It has a close relationship to **topology**, i.e. the study of distance between sets resp. functions.

This text is essentially a working reference and follows the classical expositions of **Bauer** [1] **Lang**,[4], **Rudin** [7] and **Hewitt/Stromberg** [2]. The necessary results from set theory and topology can be found in [10] and [8]; the corresponding references are given in the text. For reasons of brevity motivations and proofs for simple definitions and propositions are omitted. Nevertheless these reasonings should be worked out properly and thoroughly in order to facilitate the further understanding of the theory.

The first section introduces measurable sets, measures and measurable functions in a pronounced analogy to the open sets, metrics an continuous functions in topology. The concept of integration provides the basis for the extension of measures on product spaces. **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**. 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** and the **change-of-variables formula**.

Engelsbrand, December 2022 Arne Vorwerg

Contents

1 Measurable sets

1.1 Definitions: A family $\mathcal{A}$ of subsets of a set X is a **ring** iff

1. $\emptyset \in \mathcal{A}$.
2. $A, B \in \mathcal{A} \Leftrightarrow A \setminus B \in \mathcal{A}$
3. $A, B \in \mathcal{A} \Leftrightarrow A \cup B \in \mathcal{A}$.

In the case of

4. $X \in \mathcal{A}$,

we have an **algebra**. The additional condition

3.a) $(A_n)_{n\in N} \subset \mathcal{A} \Rightarrow \bigcup_{n\in\mathbb{N}} A_n \in \mathcal{A}$

provides us with a **σ-algebra**. The pair $(X; \mathcal{A})$ then is a **measurable space**.

On account of $\bigcap_{n\in\mathbb{N}} A_n = X \setminus \bigcup_{n\in\mathbb{N}} (X \setminus A_n)$ an **algebra** resp. a **σ-algebra** is also closed with regard to **finite resp. countably intersections**, i.e. it is a **π-system** resp. a **σ-π-system**

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_δ-sets** as well as the **closed sets** and their **countable unions**, i.e. the **F_σ-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 [8, 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** ([10, 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 $(O_1 \setminus O_2) \cap B = (O_1 \cap B) \setminus O_2 = (O_1 \cap B) \setminus (O_2 \cap B)$ and $\left(\bigcup_{n\in\mathbb{N}} O_n\right) = \bigcup_{n\in\mathbb{N}} (O_n \cap B)$ 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** $\mathcal{I} = \{[a; b[:\ a \le b \in \mathbb{R}\}$ form the **ring** $\mathcal{F} = \left\{ \mathring{\bigcup}_{0\le k\le m} I_k : I_k \in \mathcal{I}, m \in \mathbb{N} \right\}$ 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 $F = \mathring{\bigcup}_{0\le k\le m} I_k \in \mathcal{F}$ and $G = \mathring{\bigcup}_{0\le l\le n} J_l \in \mathcal{F}$ we have $F \cap G = \mathring{\bigcup}_{0\le k\le n} \mathring{\bigcup}_{0\le l\le n} I_k \cap J_l \in \mathcal{F}$, $F \setminus G = F \setminus (F \cap G) \in \mathcal{F}$ and $F \cup G \in \mathcal{F}$. The **right-open intervals** $[a; b[$ are **G_δ-sets**, hence Borel-measurable and because of $]a; b[= \bigcup_{k\in\mathbb{N}} \left[a + 2^{-k}; b\right[$ **induce** the Borel σ-algebra on $\mathbb{R}$ just as well as the **ring** 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 Borel σ-algebrae on the one-point compactification: 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**. In many cases the integration of functions with **singularities** like **meromorphic** or other **almost everywhere** finitely valued functions proves easy without excluding these points from their domain. Again to enjoy this convenience the range has to admit an infinite value ∞. To this end we admit the **one-point compactification**s e.g. $\overline{\mathbb{R}} := \mathbb{R} \cup \{\infty\} :=]-\infty; \infty]$ or $\overline{\mathbb{C}} := \mathbb{C} \cup \{\infty\}$ (cf. [8, 10.4]) as a possible **range** for measures and functions. This simple compactifications offers only limited possibilities in the algebraical as well as in the topological sense: $\infty \cdot a := a \cdot \infty := \infty$ for $a \in \mathbb{R} \setminus \{0\}$ is generally acceptable whereas $\infty \cdot 0 := 0 \cdot \infty := 0$ is a **special definition** in measure theory ensuring the tolerance of the **integral** with regard to single **poles** (cf. 4.5 and 5.7) for the price of **discontinuities** $(0; \infty)$ resp. $(\infty; 0)$ in the multiplication (cf. 4.5). The **difference** $\infty - \infty$ is not defined at all since ∞ has no negative. Furthermore the unique extension of continuous function $f : \mathbb{C} \to Y$ into a compact space Y on the compactification $\overline{\mathbb{C}}$ cannot be guaranteed. Nonetheless for our purposes this construction is sufficient. The **topology** of the one-point compactification is induced by the open sets in $\mathbb{C}$ together with the complements $\overline{\mathbb{C}} \setminus K$ of compact sets K. Due to 1.3 the Borel σ-algebra $\mathcal{B}$ on $\mathbb{C}$ is induced by the open as well as by the compact sets such that on the one hand $\mathcal{B} \subset \overline{\mathcal{B}}$ and on the other hand $A \in \overline{\mathcal{B}} \Rightarrow A \setminus \{\infty\} \in \mathcal{B}$. Since $\{\infty\} = \overline{\mathbb{C}} \setminus \bigcup_{n \in \mathbb{N}} B_n(0)$ is measurable we have $\overline{\mathcal{B}} = \{A; A \cup \{\infty\} : A \in \mathcal{B}\}$. According to 1.3 and 1.4 the Borel σ-algebra $\overline{\mathcal{B}} := \mathcal{B}_{\overline{\mathbb{R}}}$ on the compactificated real numbers $\overline{\mathbb{R}}$ can be generated alternatively by intervals of the form $[a; \infty],]a; \infty],]-\infty; a[$ or $]-\infty; a]$ resp. for $a \in \mathbb{R}$.

1.6 Dynkin-systems: A family $\mathcal{D}$ of subsets of a set X is a **Dynkin-system** or **λ-system** iff

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

Properties of Dynkin systems:

1. For $A, B \in \mathcal{D}$ with $B \subset A$ is $A \setminus B = X \setminus ((X \setminus A) \mathring{\cup} B) \in \mathcal{D}$.
2. A Dynkin-system is a **σ-algebra** iff it is **closed under intersection** since in that case for $A, B \in \mathcal{D}$ we have $A \setminus B = A \setminus (A \cap B) \in \mathcal{D}$ and for $(A_n)_{n \in \mathbb{N}} \subset \mathcal{D}$ we obtain paiwise disjoint $A'_n := A_n \setminus \bigcup_{1 \leq k < n} A_k = \bigcap_{1 \leq k < n} (A_n \setminus A_k) \in \mathcal{D}$ and hence $\bigcup_{n \in \mathbb{N}} A_n = \mathring{\bigcup}_{n \in \mathbb{N}} A'_n \in \mathcal{D}$.

Dynkin λ-π-theorem: The **Dynkin-system** $\lambda(\mathcal{E})$ generated by a **π-basis** $\mathcal{E} \subset P(X)$ coincides with the corresponding **σ-algebra** $\sigma(\mathcal{E})$.

Proof: For every $D \in \lambda(\mathcal{E})$ the family $\mathcal{D}_D := \{A \subset X : A \cap D \in \lambda(\mathcal{E})\}$ on account of $(X \setminus A) \cap D = D \setminus (A \cap D)$ and 1.7.1 is itself a Dynkin-system including $\mathcal{E}$ and hence $\lambda(\mathcal{E})$. Owing to 1.7.2 it is a σ-Algebra, i.e. $\sigma(\mathcal{E}) \subset \lambda(\mathcal{E})$ and since every σ-algebra is a Dynkin-system we have $\sigma(\mathcal{E}) = \lambda(\mathcal{E})$.

2 Pre-measures

2.1 Definition: A set function $\mu : \mathcal{A} \to [0; \infty]$ on a **ring** $\mathcal{A} \subset P(X)$ 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-additiv**ity 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. [8, 14.12]) of the **real numbers** permits the extension of the **subadditivity** to **countable unions**: $\mu(\bigcup_{n \in \mathbb{N}} A_n) \leq \sum_{n \in \mathbb{N}} \mu(A_n)$.

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

1. **σ-Additiv**ity: 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 $\lim_{n \in \mathbb{N}} \mu(A_n) = \mu(\bigcup_{n \in \mathbb{N}} A_n)$.

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 $\mu(A_n) = \mu\left(\bigcup_{1\le k\le n} A'_k\right) = \sum_{1\le k\le n} \mu(A'_k)$ such that $\lim\limits_{n\in\mathbb{N}} \mu(A_n) = \sum_{n\in\mathbb{N}} \mu(A'_n) = \mu\left(\bigcup_{n\in\mathbb{N}} A'_n\right) = \mu\left(\bigcup_{n\in\mathbb{N}} A_n\right)$.

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 $\lim\limits_{n\in\mathbb{N}} \mu(A_n) = \lim\limits_{n\in\mathbb{N}} \mu(A_0 \setminus A'_n) = \lim\limits_{n\in\mathbb{N}} \left(\mu(A_0) - \mu(A'_n)\right) = \mu(A_0) - \lim\limits_{n\in\mathbb{N}} \mu(A'_n)$ $= \mu(A_0) - \mu\left(\bigcup_{n\in\mathbb{N}} A'_n\right) = \mu\left(A_0 \setminus \bigcup_{n\in\mathbb{N}} A'_n\right) = \mu\left(\bigcap_{n\in\mathbb{N}} A_0 \setminus A'_n\right) = \mu\left(\bigcap_{n\in\mathbb{N}} A_n\right)$.

3\. $\Rightarrow$ 4. : Obvious.

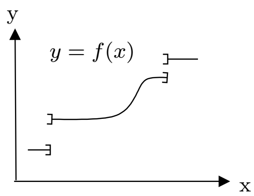
4\. $\Rightarrow$ 1. : With $A'_k := \mathring{\bigcup}_{n>k} A_n$ we obtain a **decreasing** sequence $(A'_k)_{k\in\mathbb{N}}$ with $\bigcap_{k\in\mathbb{N}} A'_k = \emptyset$ and $\mu(A'_k) < \infty$ such that according to 4. we have $0 = \lim\limits_{k\in\mathbb{N}} \mu(A'_k) = \mu\left(\mathring{\bigcup}_{n\in\mathbb{N}}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

2.3 Examples:

1. The **Dirac measure** $\epsilon_x(A) := \begin{cases} 1, & x \in A \\ 0, & x \notin A \end{cases}$ 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.

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. [8, 3.3]) function $f : \mathbb{R} \to \mathbb{R}$ on the **right-open intervals** $\mathcal{I} = \{[a;b[: a \le b \in \mathbb{R}\}$ can be extended to the **ring** $\mathcal{F} = \left\{ \mathring{\bigcup\limits_{0\le k\le m}} I_k : I_k \in \mathcal{I}, m \in \mathbb{N} \right\}$ of the **one-dimensional figures** by $\lambda_f\left(\mathring{\bigcup\limits_{0\le k\le m}} I_k \right) := \sum\limits_{k=1}^{m} \lambda_f(I_k)$. For every **decreasing** sequence of figures $F_0 \supset F_1 \supset ...$ with $F_n = \mathring{\bigcup}_{0\le k_n\le m_n} [a_{k_n}; b_{k_n}[\in \mathcal{F}$ and $\bigcap_{n\in\mathbb{N}} F_n = \emptyset$ the decreasing character implies that $\forall n > m\ \forall 0 \le k_n \le l_n\ \exists 0 \le k_m \le l_m$ 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** [8, 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.



lower continuous
nondecreasing

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. $\tilde{\mu}\left(\bigcup_{n\in\mathbb{N}} A_n\right) \leq \sum_{n\in\mathbb{N}} \tilde{\mu}\left(A_n\right)$

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

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

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): $\tilde{\mu}(Q) = \tilde{\mu}\left(Q \cap A\right) + \tilde{\mu}\left(Q \setminus A\right) = \tilde{\mu}\left(Q \cap A \cap B\right) + \tilde{\mu}\left(Q \cap A \setminus B\right) + \tilde{\mu}\left(Q \setminus A \cap B\right) + \tilde{\mu}\left(Q \setminus A \setminus B\right)$ and if we substitute Q with $Q \cap (A \cup B)$ in (I) we arrive at another equatiuon (II): $\tilde{\mu}\left(Q \cap (A \cup B)\right) = \tilde{\mu}\left(Q \cap A \cap B\right) + \tilde{\mu}\left(Q \cap A \setminus B\right) + \tilde{\mu}\left(Q \setminus A \cap B\right)$. We can substitute the first three terms in (I) by (II) and hence obtain the measurability of the **union**: $\tilde{\mu}(Q) = \tilde{\mu}\left(Q \cap (A \cup B)\right) + \tilde{\mu}\left(Q \setminus (A \cup B)\right)$. Thus $\mathcal{A}$ is an **Algebra**.

For a sequence $\left(A_n\right)_{n\in\mathbb{N}} \subset \mathcal{A}$ 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 1.6.3 the family $\mathcal{A}$ is a **Dynkin system** which is **closed under intersection** and in accordance with 1.7.2 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 $\sigma\left(\mathcal{E}\right)$ induced by a **π-basis** $\mathcal{E} \subset \mathcal{P}\left(X\right)$ are identical iff they coincide on $\mathcal{E}$ and are **$\boldsymbol{\sigma}$-finite** on $\mathcal{E}$, i.e. $\exists \left(E_n\right)_{n\in\mathbb{N}} \subset \mathcal{E}$ with $\bigcup_{n\in\mathbb{N}} E_n = X$ and $\mu_1\left(E_n\right) = \mu_2\left(E_n\right) < \infty$ for all $n \in \mathbb{N}$.

Proof: For $E \in \mathcal{E}$ with $\mu_1\left(E\right) = \mu_2\left(E\right) < \infty$ the family $\mathcal{D}_E := \left\{D \in \sigma\left(\mathcal{E}\right) : \mu_1\left(E \cap D\right) = \mu_2\left(E \cap D\right)\right\}$ is a **Dynkin system** since $\emptyset \in \mathcal{D}_E$ and for every $D \in \mathcal{D}_E$ on account of $\mu_1\left(E \cap X \setminus D\right) = \mu_1(E) - \mu_1\left(E \cap D\right) = \mu_2(E) - \mu_2\left(E \cap D\right) = \mu_2\left(E \cap X \setminus D\right)$ we also have $X \setminus D \in \mathcal{D}_E$. Criterion 1.6.3 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 the **Dynkin λ-π-theorem** 1.8 entails $\sigma\left(\mathcal{E}\right) = \delta\left(\mathcal{E}\right) \subset \mathcal{D}_E \subset \sigma\left(\mathcal{E}\right)$, i.e.$\mathcal{D}_E = \sigma\left(\mathcal{E}\right)$ resp. $\mu_1\left(E \cap A\right) = \mu_2\left(E \cap A\right)$ for all $E \in \mathcal{E}$ and $A \in \sigma\left(\mathcal{E}\right)$.

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

3.5 Extension theorem: Every **$\boldsymbol{\sigma}$-finite pre-measure** μ on a **ring** $\mathcal{R}$ can be extented **in a unique way** to a **measure** μ on $\sigma\left(\mathcal{R}\right)$.

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

The ring $\mathcal{R}$ is **$\tilde{\mu}$-measurable** since for every $A \in \mathcal{R}$ 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 $Q \in \sigma(\mathcal{R})$ with **finite measure** $\mu(Q) < \infty$ on a **σ-algebra** $\sigma(\mathcal{R})$ induced by a **ring** $\mathcal{R}$ can be approximated **in measure** by a sequence $(C_n)_{n\in\mathbb{N}} \subset \mathcal{R}$ 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 $(A_k)_{k\in\mathbb{N}} \subset \mathcal{R}$ 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 **ring 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 $\lambda_f([a;b[) = \lambda_f([a;b]) = \lambda_f(]a;b]) = \lambda_f(]a;b[) = f(b) - f(a)$ 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**: $\lambda_f(\mathbb{Q}) = 0$. 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 $x = (x_n)_{n\geq 1}$ with $x_n \in \{0;2\}$ (cf. [8, 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 $\lambda(T_n) = \frac{2^n}{3^n}$ and $\lambda(T) = \lim\limits_{n\in\mathbb{N}} \lambda(T_n) = 0$ due to the **continuity from above** (2.2.3). The G_δ-set $U = \bigcap_{n\geq 1} U_n$ with **dense open** sets $U_n = \bigcup_{i\geq 1} B_{n^{-1}.2^{-i-1}}(q_i)$ based on the enumeration $\mathbb{Q} = (q_i)_{i\geq 1}$ includes $\mathbb{Q}$ and hence is dense in $\mathbb{R}$. Again due to 2.2.3 and since $\lambda(U_n) \leq \frac{1}{n}$ it also is a λ-null set: $\lambda(U) = 0$. The **complements** $\mathbb{R} \setminus U_n$ are **closed** and **nowhere dense** in $\mathbb{R}$ but with measure $\lambda(\mathbb{R} \setminus U_n) = \infty$ and $\mathbb{R} \setminus U$ is an example for a set of **first category** with measure $\lambda(\mathbb{R} \setminus U) = \infty$. (cf. [8, 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 $\mu(N_A) = \mu(N_B) = 0$ we have $(A \cup M_A) \setminus (B \cup M_B) = A \setminus (B \cup M_B) \cup M_A \setminus (B \cup M_B) = (A \setminus B) \cap (A \setminus N_B) \cup (N_B \setminus M_B) \cup M_A \setminus (B \cup M_B) \in \mathcal{A}_0$ since $(A \setminus B) \cap (A \setminus N_B) \in \mathcal{A}$ and $(N_B \setminus M_B) \cup M_A \setminus (B \cup M_B) \subset N_A \cup N_B$ with $\mu(N_A \cup N_B) = 0$. 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 $\mu(B \setminus A) = 0$: One the one hand for any $E = A \cup M$ with $M \subset N \in \mathcal{A}$ and $\mu(N) = 0$ 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 $E = A \cup (B \setminus A \cap E)$ 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 $\mathcal{B}_0$ of the **Borel sets** is extended to the **Lebesgue measure** λ_0 on the completed σ-algebra $\mathcal{B}_0$ of the **Lebesgue sets**.

3.9 Almost everywhere existing properties: In **probability theory** the completion $\mathcal{B}_0$ 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 $\mu(\neg E) = 0$.

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** [10, 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** $\mathbb{R} = \mathring{\bigcup}_{q\in\mathbb{Q}} (q + K)$ which due to the σ**-additivity** and the **translation invariance** must satisfy $\infty = \lambda_0(\mathbb{R}) = \sum_{q\in\mathbb{Q}} \lambda_0(K)$ and hence $\lambda_0(K) > 0$. On the other hand we have $\mathring{\bigcup}_{q\in\mathbb{Q}\cap[0;1]} (q + K) \subset [0;2]$ and due to the **monotonicity** of the measure $\sum_{q\in\mathbb{Q}} \lambda_0(K) \leq \lambda_0([0;2]) = 2$ hence $\lambda_0(K) = 0$. 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. [10, 9.2]) it is sufficient if the inverse images of **basis** sets are measurable in X (cf. [8, 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 $\mathcal{B} = \sigma(\mathcal{O})$. According to 2.1 a function $f : X \to \overline{\mathbb{R}}$ 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. Since $\mathbb{Q}$ is countable and dense in $\mathbb{R}$ for measurable $f, g : X \to \overline{\mathbb{R}}$ the sets $\{f > g\} = \bigcup_{a\in\mathbb{Q}} (\{f > a\} \cap \{a > g\})$ and $\{f \geq g\} = X \setminus \{f < g\}$ are measurable. 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 $\mu(|f| < \epsilon) = \mu(\{|f| < \epsilon\})$.

4.2 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**: $g \circ f \circ \mu : g \circ f(\mathcal{A}) \to [0;\infty]$ oviously 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 $(g \circ \lambda)([a;b[) = \lambda(g^{-1}[[a;b[]) = \lambda\left(\left[\frac{a}{m};\frac{b}{m}\right[\right) = \frac{1}{m}\lambda([a;b[)$.

4.3 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 4.2 since $f(\sigma(f^{-1}(\mathcal{E})))$ is a σ-algebra on Y including $\mathcal{E}$ and hence $\sigma(\mathcal{E})$.

4.4 Continuous functions: On account of 4.3 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**.

4.5 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 [8, 3.1; 4.2.3 and 10.7]

- the **projections** $(x;y) \mapsto x$ resp. $(x;y) \mapsto y$ and the **euclidean norm** $(x;y) \mapsto \sqrt{x^2 + y^2}$ for $\overline{\mathbb{C}} \simeq \overline{\mathbb{R}^2} \to \overline{\mathbb{R}}$
- the **multiple** $x \mapsto \alpha \cdot x$ and the **powers** $x \mapsto x^\alpha$ for $\alpha \in \mathbb{C}$ as well as in particular the reciprocal $x \mapsto \frac{1}{x}$ for $\overline{\mathbb{C}} \to \overline{\mathbb{C}}$

- the **addition** $(x;y) \mapsto x+y$ for $\overline{\mathbb{C}}^2 \to \overline{\mathbb{C}}$ and the **multiplication** $(x;y) \mapsto x \cdot y$ for $\overline{\mathbb{C}}^2 \setminus \{(0;\infty);(\infty;0)\} \to \overline{\mathbb{C}}$

are **continuous** and hence **Borel measurable**. According to 1.5 the Borel measurability of the multiplication can be extended to the complete compactification $\overline{\mathbb{C}}^2$ since with the definition $\infty \cdot 0 := 0 \cdot \infty := 0$ such that the Borel measurable set $\{(0;\infty);(\infty;0)\} \subset \overline{\mathbb{C}}^2$ of discontinuities of the multiplication is simply added to the null set $\left\{(x;y) \in \overline{\mathbb{C}}^2 : x \cdot y = 0\right\} = \{(0;\infty);(\infty;0)\} \cup \left\{(x;y) \in \mathbb{C}^2 : x \cdot y = 0\right\}$ which is still Borel measurable. For Borel measurable $f, g : X \to \overline{\mathbb{C}}$ the **real part** $\mathrm{Re}f$, **imaginary part** $\mathrm{Im}f$ and **absolute value** $\|f\|$ are Borel measurable mappings $X \to \overline{\mathbb{R}}$; likewise $\alpha \cdot f$, f^α, $\frac{1}{f}$, $f+g$ and $f \cdot g$ are Borel measurable mappings $X \to \overline{\mathbb{C}}$.

4.6 Semicontinuous functions: 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. [8, 3.3]) According to 1.4 resp. 4.1 these functions are again Borel measurable.

4.7 Sequences of functions: For a sequence $(f_n)_{n\in\mathbb{N}}$ of Borel measurable functions $f_n : X \to \overline{\mathbb{R}}$ on a measurable space $(X;\mathcal{A})$ we have $\left\{\inf\limits_{n\in\mathbb{N}} f_n < a\right\} = \bigcup_{n\in\mathbb{N}} \{f_n < a\}$ and $\left\{\sup\limits_{n\in\mathbb{N}} f_n > a\right\} = \bigcup_{n\in\mathbb{N}} \{f_n > a\}$ such that according to 4.1 the limits $\inf\limits_{n\geq 0} f_n$, $\sup\limits_{n\geq 0} f_n$, $\limsup\limits_{n\in\mathbb{N}} f_n := \inf\limits_{n\geq 0}\left(\sup\limits_{k\geq n} f_k\right)$ and $\liminf\limits_{n\in\mathbb{N}} f_n := \sup\limits_{n\geq 0}\left(\inf\limits_{k\geq n} f_k\right)$ are all measurable. In particular for a Borel measurable $f : X \to \overline{\mathbb{R}}$ the **positive part** $f^+ := \max\{f;0\}$, the **negative part** $f^- := \min\{f;0\}$ as well as for every other Borel measurable $g : X \to \overline{\mathbb{R}}$ the **maximum** $\max\{f;g\}$ and the **minimum** $\min\{f;g\}$ are Borel measurable. Owing to 4.5 a **complex valued** function $f : X \to \overline{\mathbb{C}}$ is Borel measurable iff this is true for the **real part** and the **imaginary part** $\mathrm{Re}f, \mathrm{Im}f : X \to \overline{\mathbb{R}}$. Hence the **limit** $\lim\limits_{n\geq 0} f_n$ of a pointwise converging sequence $(f_n)_{n\in\mathbb{N}}$ of Borel measurable functions $f_n : X \to \overline{\mathbb{C}}$ is again Borel measurable.

4.8 Elementary 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** ϵ_x from 2.3.1 albeit with interchanged roles for x and A. The **elementary functions** $\mathcal{E}(X) = \left\{\sum_{0\leq i\leq m} \alpha_i \chi_{A_i} : m \in \mathbb{N}, \alpha_i \in \mathbb{R}_+, A_i \in \mathcal{A}\right\}$ form a real **algebra** (cf. [8, 18.9]) of Borel measurable functions with each f and g containing also $\max\{f;g\}$ and $\min\{f;g\}$. The closedness under linear operations, multiplication as well as taking of maxima and minima is obvious from observing two elementary functions $f = \sum_{0\leq i\leq m} \alpha_i \chi_{A_i}$ and $g = \sum_{0\leq j\leq n} \beta_j \chi_{Bj}$ with w.l.o.g. pairwise disjoint supports A_i and B_j, choosing representations $f = \sum_{0\leq i\leq m} \sum_{0\leq j\leq n} \alpha_i \chi_{A_i \cap B_j}$ resp. $g = \sum_{0\leq i\leq m} \sum_{0\leq j\leq n} \beta_j \chi_{A_i \cap B_j}$ with common and again pairwise disjoint supports $A_i \cap B_j$ and finally considering the equation $\chi_A \cdot \chi_B = \chi_{A\cap B}$.

4.9 The algebra of elementary functions: Every Borel measurable function $f : X \to [0;\infty]$ on a measurable space $(X;\mathcal{A})$ is the supremum of an increasing sequence of elementary functions. The **algebra** of the elementary functions is **dense** in the set of **positive** and **finite** Borel measurable functions with reference to the topology of **pointwise convergence**.

Proof: Every Borel measurable function $f : X \to [0;\infty]$ obviously is the supremum of the increasing sequence $(s_n \circ f)_{n\geq 0}$ of elementary functions defined by $s_n = \sum_{0\leq k\leq n\cdot 2^n} k \cdot 2^{-n} \chi_{A_{nk}}$ with $A_{nk} :=]k \cdot 2^{-n}; (k+1) \cdot 2^{-n}]$ for $0 \leq k < n \cdot 2^n$ resp. $A_{n;n\cdot 2^n} :=]n;\infty]$. For **finite** Borel measurable function $f : X \to [0;\infty[$ and every $x \in X$ there is an $n_0 \in \mathbb{N}$ with $f(x) < n_0$ such that for every $\epsilon > 0$ there is an $n \geq n_0$ with $2^{-n} < \epsilon$ as well as a $k \leq n \cdot 2^n$ with $k \cdot 2^{-n} < f(x) \leq (k+1) \cdot 2^{-n}$ hence $(s_n \circ f)(x) = k \cdot 2^{-n} < f(x)$ and $|(s_n \circ f)(x) - f(x)| < \epsilon$, i.e. $\lim\limits_{n\geq 0} (s_n \circ f)(x) = f(x)$. On account of 4.7 every limit function of a sequence of elementary functions is measurable.

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** $(Y, \|\|)$ (cf. [8, 21.9]) 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 measurable sets $A \in \mathcal{A}$ with **finite measure** $\mu(A) < \infty$. For an apriori **finite measure** with $\mu(X) < \infty$ the two definitions obviously coincide. In tthe most important 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 $\inf\limits_{k\geq 1}\sup\limits_{n\geq k} \mu|_A (\{\|f_n - f\| \geq \epsilon\}) \overset{2.2.2}{=} \inf\limits_{k\geq 1} \mu|_A \left(\bigcup_{n\geq k} \{\|f_n - f\| \geq \epsilon\}\right) \overset{2.2.3}{=} \mu|_A \left(\bigcap_{k\geq 1} \bigcup_{n\geq k} \{\|f_n - f\| \geq \epsilon\}\right) = 0.$

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 $\inf\limits_{n\in\mathbb{N}} \lambda(\mathbb{R} \setminus B_n(0)) = \infty$ since $\lambda(\mathbb{R} \setminus B_n(0)) = \infty$ for every $n \in \mathbb{N}$. Hence the Lebesgue convergence theorem **fails** for the Lebesgue measure: 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 $\lambda(\|f_n(x) - f(x)\| \geq \epsilon) = \lambda(\|x\| \geq \sqrt{n\epsilon}) = \infty$ 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 in case of 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. $\Rightarrow$ 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 **$\mu$-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 $\mu|_A (\|f_{n_m} - f\| > \frac{\epsilon}{2}) < \frac{\epsilon}{2}$ 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 $\mu|_A (\|f_n - f\| > \epsilon) \leq \mu|_A (\{\|f_n - f_{n_m}\| > \frac{\epsilon}{2}\} \cup \{\|f_{n_m} - f\| > \frac{\epsilon}{2}\}) \leq \mu|_A (\|f_m - f_{n_m}\| > \frac{\epsilon}{2}) + \mu|_A (\|f_{n_m} - f\| > \frac{\epsilon}{2}) < \epsilon$. This converse-triangle-inequality argument will be repeatedly used in the subsequent proofs.

2. $\Rightarrow$ 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. $\Rightarrow$ 1. : Suppose there is an $\epsilon > 0$ such that $\forall n_k \geq 1\ \exists n_{k+1} \geq n_k$ with $\mu|_A (\|f_{n_{k+1}} - f_{n_k}\| > \epsilon) > \epsilon$. As above we get $\mu|_A (\|f_{n_k} - f\| > \frac{\epsilon}{2}) + \mu|_A (\|f_{n_{k+1}} - f\| > \frac{\epsilon}{2}) \geq \mu|_A (\|f_{n_k} - f_{n_{k+1}}\| > \epsilon) > \epsilon$, i.e. either $\mu|_A (\|f_{n_k} - f\| > \frac{\epsilon}{2}) \geq \frac{\epsilon}{2}$ or $\mu|_A (\|f_{n_{k+1}} - f\| > \frac{\epsilon}{2}) \geq \frac{\epsilon}{2}$. 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:

$\Rightarrow$: 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).

$\Leftarrow$: 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

5.1 Definition: A complex **functional** $I:\mathcal{L}(X)\to\overline{\mathbb{C}}$ on a **vector space** $\mathcal{L}(X)\subset\overline{\mathbb{C}}^X$ of **complex functions** on a set X is an **integral** iff for $\alpha,\beta\in I[\mathcal{L}(X)]$ and $f,g\in\mathcal{L}(X)$ the following properties hold:

1. $I(\alpha f+\beta g)=\alpha I(f)+\beta I(g)$ (**linearity**)
2. $\|f\|\leq\|g\|\Rightarrow\|I(f)\|\leq\|I(g)\|$ (**monotonicity**)

$\mathcal{L}^1(X)=\{f\in\mathcal{L}(X): I(\|f\|)<\infty\}$is the vector subspace of the **integrable** functions.

5.2 Integral for elementary functions: The expression $\int f d\mu:=\sum_{0\leq i\leq m}\alpha_i\mu(A_i)$ for $f=\sum_{0\leq i\leq m}\alpha_i\chi_{A_i}$ defines an **integral** on the **algebra** $\mathcal{E}(X)$ of the **elementary functions** on a measure space$(X;\mathcal{A};\mu)$. **Uniqueness**, **linearity** and **monotonicity** are obvious if we consider representations with **common** and **pairwise disjunct supports** $A_i\cap B_j$ for two elementary functions f und g as in 4.8 and observe the **additivity** 2.2.1 of the measure.

5.3 Theorem: For every **increasing** sequence $(f_n)_{n\in\mathbb{N}} \subset \mathcal{E}(X)$ and $f \in \mathcal{E}(X)$ with $f \leq \sup\limits_{n\in\mathbb{N}} f_n$ we have $\int f d\mu \leq \sup\limits_{n\in\mathbb{N}} \int f_n d\mu$.

Proof: Due to 4.1 for every $\alpha \in]0;1[$ the sets $B_n = \{f_n \geq \alpha f\}$ form an increasing sequence of measurable sets with $\bigcup_{n\in\mathbb{N}} B_n = X$. For an $f = \sum_{0\leq i\leq m} \alpha_i \chi_{A_i}$ and because of $\chi_A \cdot \chi_B = \chi_{A\cap B}$ resp. the **continuity from below** 2.2.2 (holding even in the case of some $\mu(A_i \cap B_n) = \infty$) we have $\sup\limits_{n\in\mathbb{N}} \int f \chi_{B_n} d\mu = \sum_{0\leq k\leq m} \alpha_i \sup\limits_{n\in\mathbb{N}} \mu(A_i \cap B_n) = \sum_{0\leq i\leq m} \alpha_i \mu\left(\bigcup_{n\in\mathbb{N}} (A_i \cap B_n)\right) = \sum_{0\leq i\leq m} \alpha_i \mu(A_i)$ $= \int f d\mu$. On account of $\alpha f \chi_{B_n} \leq f_n$ resp. due to 5.1.1 we also have $\alpha \int f \chi_{B_n} d\mu \leq \int f_n d\mu$ and since both sides of the inequality are increasing with $n \to \infty$ we infer $\alpha \int f d\mu = \alpha \int f d\mu = \sup\limits_{n\in\mathbb{N}} \alpha \int f \chi_{B_n} d\mu$ $\leq \sup\limits_{n\in\mathbb{N}} \int f_n d\mu$ and hence the assertion.

5.4 Integral for positive Borel measurable functions: For every positive Borel measurable function $f = \sup\limits_{n\in\mathbb{N}} f_n : X \to [0;\infty]$ with an increasing sequence $(f_n)_{n\in\mathbb{N}} \subset \mathcal{E}(X)$ according to 4.9 the expression $\int f d\mu := \sup\limits_{n\in\mathbb{N}} \int f_n d\mu$ defines an **integral** and in the case of $\int f d\mu < \infty$ the function f is **integrable**. The **uniqueness** of the integral results from 5.3 since for two increasing sequences $(f_n)_{n\in\mathbb{N}}, (g_n)_{n\in\mathbb{N}} \subset \mathcal{E}(X)$ with $\sup\limits_{n\in\mathbb{N}} f_n = \sup\limits_{n\in\mathbb{N}} g_n$ we have $\int f_n d\mu \leq \sup\limits_{n\in\mathbb{N}} \int g_n d\mu$ and vice versa for every $n \in \mathbb{N}$. The same argument applied only once delivers the **monotonicity** and on account of the integral being independent of the choice of the approximating sequence and the linearity of the supremum for **positive real** coefficients α resp. β we also have **linearity** since $\int (\alpha f + \beta g) d\mu =$ $\sup\limits_{n\in\mathbb{N}} \int (\alpha f_n + \beta g_n) d\mu = \sup\limits_{n\in\mathbb{N}} \left(\alpha \int f_n d\mu + \beta \int g_n d\mu\right) = \alpha \sup\limits_{n\in\mathbb{N}} \int f_n d\mu + \beta \sup\limits_{n\in\mathbb{N}} \int g_n d\mu = \alpha \int f d\mu + \beta \int g d\mu$. Note that this definition as well as the subseqent theorems depend heavily on the **supremum property** of the real numbers, i.e. their **completeness**. (cf.[8, 14.12])

5.5 Levi's monotone convergence theorem: For every **increasing** sequence $(f_n)_{n\in\mathbb{N}}$ **of positive Borel measurable** $f_n : (X;\mathcal{A};\mu) \to \overline{\mathbb{R}}$ we have $\int \sup\limits_{n\in\mathbb{N}} f_n d\mu = \sup\limits_{n\in\mathbb{N}} \int f_n d\mu$.

Proof: On account of the **monotonicity of the integral** we have $\int \sup\limits_{n\in\mathbb{N}} f_n d\mu \geq \sup\limits_{n\in\mathbb{N}} \int f_n d\mu$. To prove the reciprocal relation for every $n \in \mathbb{N}$ we choose an increasing sequence $(e_{nk})_{k\in\mathbb{N}} \subset \mathcal{E}(X)$ with $\sup\limits_{k\in\mathbb{N}} e_{nk} = f_n$ such that the functions $\hat{e}_n := \max\{e_{kk} : 0 \leq k \leq n\} \in \mathcal{E}(X)$ form again an increasing sequence with $\hat{e}_n \leq f_n$, hence $\int \hat{e}_n d\mu \leq \int f_n d\mu$ for all $n \in \mathbb{N}$. But since $\sup\limits_{n\in\mathbb{N}} \hat{e}_n = \sup\limits_{n\in\mathbb{N}} f_n$ definition 5.4 shows that $\int \sup\limits_{n\in\mathbb{N}} f_n d\mu = \sup\limits_{n\in\mathbb{N}} \int \hat{e}_n d\mu \leq \sup\limits_{n\in\mathbb{N}} \int f_n d\mu$.

5.6 Fatou's lemma: Fo every sequence $(f_n)_{n\in\mathbb{N}}$ of **positive Borel measurable** functions on a measure space $(X;\mathcal{A},\mu)$ we have $\int \liminf\limits_{n\in\mathbb{N}} f_n d\mu \leq \liminf\limits_{n\in\mathbb{N}} \int f_n d\mu$.

Proof: Due to 4.5 and 5.5 we have $\int \liminf\limits_{n\in\mathbb{N}} f_n d\mu = \int \sup\limits_{n\in\mathbb{N}} \left(\inf\limits_{k\geq n} f_k\right) d\mu = \sup\limits_{n\in\mathbb{N}} \int \left(\inf\limits_{k\geq n} f_k\right) d\mu$. Due to the **monotonicity of te integral** we also have $\int \left(\inf\limits_{k\geq n} f_k\right) d\mu \leq \inf\limits_{k\geq n} \int f_k d\mu$ and hence the assertion.

5.7 μ-a.e. properties of integrable functions: For any **positive**, **Borel-measurable and integrable** function f we have

1. f is **μ-a.e. finite** since due to the **monotonicity of the integral** and $n\chi_{\{f=\infty\}} < f$ we have $n \cdot \mu(\{f = \infty\}) < \int f d\mu < \infty$ for all $n \geq 1$ and hence $\mu(\{f = \infty\}) = 0$. The tolerance of the integral with regard to singularities on μ-null sets is a dierct consequence of the measure theoretic special definition $0 \cdot \infty = \infty \cdot 0 = 0$ (cf. 1.5).
2. The set $\{f \neq 0\}$ is **σ-finite** since due to $\chi_{\{nf\geq 1\}} < nf$ we have $\mu(\{nf \geq 1\}) < n \int f d\mu < \infty$ and hence $\bigcup_{n\geq 1} \{nf \geq 1\} = \bigcup_{n\geq 1} \left\{f \geq \frac{1}{n}\right\} = \{f \neq 0\}$.
3. Conversely in the case of $\int f d\mu = 0$ for $A_n = \left\{f > \frac{1}{n}\right\}$ we have $\frac{1}{n}\mu(A_n) \leq \int_{A_n} f\, d\mu \leq \int f d\mu = 0$ hence $\mu(f > 0) = \mu\left(\bigcup_{n\in\mathbb{N}} A_n\right) = 0$ on account of the **continuity from above** 2.2.3 and hence

μ-a.e. $f = 0$. In particular for two **positive**, **Borel-measurable and integrable** functions $f \leq g$ with equal integrals $\int f d\mu = \int g d\mu$ follows μ-a.e $f = g$.

5.8 Lebesgue's dominated convergence theorem: A sequence $(f_n)_{n\in\mathbb{N}}$ of positive and Borel-measurable functions **converging μ-a. e.** to $\lim\limits_{n\in\mathbb{N}} f_n := f$ with an **integrable majorant** $g \in \mathcal{L}^1(\mu)$ such that $f_n \leq g \forall n \in \mathbb{N}$ **converges in mean** to f: $\lim\limits_{n\in\mathbb{N}} \int |f_n - f| \, d\mu = 0$.

Proof: Due to 5.6 with $\liminf\limits_{n\in\mathbb{N}} (g - |f - f_n|) = g$ resp. $\int (g - |f - f_n|) \, d\mu + \int |f - f_n| \, d\mu = \int g d\mu$ we have $\int g d\mu \leq \liminf\limits_{n\in\mathbb{N}} \int (g - |f - f_n|) \, d\mu = \liminf\limits_{n\in\mathbb{N}} \left(\int g d\mu - \int |f - f_n| \, d\mu\right) = \int g d\mu - \limsup\limits_{n\in\mathbb{N}} \int |f - f_n| \, d\mu$ $\Leftrightarrow \limsup\limits_{n\in\mathbb{N}} \int |f - f_n| \, d\mu \leq 0$.

5.9 Convergence in mean and μ-a.e: For every sequence $(f_n)_{n\geq 1}$ of Borel measurable functions $f_n : X \to Y$ into a **Banach space** $(Y, \|\|)$ **converging in mean** to an integrable $f : X \to Y$ and every $\epsilon > 0$ there is a set $A_\epsilon \in \mathcal{A}$ with $\mu(A_\epsilon) < \epsilon$ such that a **subsequence** $(f_{n_k})_{k\geq 1}$ **uniformly converges** to f on $X \setminus A_\epsilon$ and μ-a.e. on X.

Proof: According to the hypothesis for $k \in \mathbb{N}$ there is an $n_k \in \mathbb{N}$ with $\int |f_{n_k} - f| \, d\mu \leq 2^{-2k}$ for every $k \geq 1$. Hence for $B_k := \left\{|f_{n_k} - f| \geq 2^{-k}\right\}$ we have $\mu(B_k) \leq 2^k \int_{Y_k} |f_{n_k} - f| \, d\mu \leq 2^k \int |f_{n_k} - f| \, d\mu \leq 2^{-k}$ whence $\mu(A_m) \leq 2^{1-m}$ for $A_m := \bigcup\limits_{k\geq m} B_k$ with $(f_{n_k})_{k\geq 1}$ converging **uniformly** on every $X \setminus A_m$ and **pointwise** on $X \setminus \bigcap\limits_{m\geq 1} A_m$ with $\mu\left(\bigcap\limits_{m\geq 1} A_m\right) = 0$ on account of the **continuity from above** 2.2.3.

5.10 Comparison with the Riemann integral:<

1. Every **Riemann integrable** function $f : [a; b] \to [0; \infty[$ is **Lebesgue integrable** and the two integrals are equal: $\int_a^b f(x)dx = \int_{[a;b]} f d\lambda$.
2. $f : \mathbb{R} \to [0; \infty[$ is Lebesgue 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 $[a; b]$ the **lower Darboux sum** $U_{z_n} := \sum\limits_{i=0}^n \gamma_i (a_i - a_{i-1}) = \int_{[a;b]} u_{z_n} d\lambda$ with $\gamma_i := \inf f[[a_{i-1}; a_i]]$ resp. **upper Daboux sum** $O_{z_n} := \sum\limits_{i=0}^n \Gamma_i (a_i - a_{i-1}) = \int_{[a;b]} o_{z_n} d\lambda$ with $\Gamma_i := \sup f[[a_{i-1}; a_i]]$ can be regarded as **Lebesgue integrals** of the elementary functions $u_{z_n} \sum\limits_{i=0}^n \gamma_i \chi_{[a_{i-1};a_i]}$ resp. $o_{z_n} := \sum\limits_{i=0}^n \Gamma_i \chi_{[a_{i-1};a_i]}$. Due 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 and $\lim\limits_{n\to\infty} U_{z_n} = \lim\limits_{n\to\infty} O_{z_n} = \int_a^b f(x)dx$. Since $(o_{z_n})_{n\in\mathbb{N}}$ decreases, $(u_{z_n})_{n\in\mathbb{N}}$ increases, $(o_{z_n} - u_{z_n})_{n\in\mathbb{N}}$ is a decreasing sequence bounded below by 0 such that due to the **completeness** of the **real numbers** there must a limit $\lim\limits_{n\in\mathbb{N}} (o_{z_n} - u_{z_n}) \geq 0$. Owing to 5.6 follows $0 \leq \int \lim\limits_{n\in\mathbb{N}} (o_{z_n} - u_{z_n})$ $\leq \liminf\limits_{n\in\mathbb{N}} (O_{z_n} - U_{z_n}) = 0$ and due to 5.7 we have λ-a.e. $\lim\limits_{n\in\mathbb{N}} (o_{z_n} - u_{z_n}) = 0$. Since λ-a.e. $u_{z_n} \leq f \leq o_{z_n}$ we infer that λ-a.e. $\lim\limits_{n\in\mathbb{N}} u_{z_n} = f$. Finally we apply 5.8 with the majorant o_{z_0} and obtain $\int_a^b f(x)dx = \lim\limits_{n\in\mathbb{N}} \int_{[a;b]} u_{z_n} d\lambda = \int_{[a;b]} \left(\lim\limits_{n\in\mathbb{N}} u_{z_n}\right) d\lambda = \int_{[a;b]} f d\lambda$.
2. According to 1. and 5.5 we have $\lim\limits_{n\in\mathbb{N}} \int_{-n}^n f(x)dx = \lim\limits_{n\in\mathbb{N}} \int_{[-n;n]} f d\lambda = \sup\limits_{n\in\mathbb{N}} \int f \cdot \chi_{[-n;n]} d\lambda = \int \sup\limits_{n\in\mathbb{N}} f \cdot \chi_{[-n;n]} d\lambda = \int_{\mathbb{R}} f d\lambda$.

Note: In essential, 5.5, 5.6 and 5.8 assert the **continuity of the Lebesgue integral** regarding **pointwise** esp. **μ-a.e. convergence** whereas the **Riemann integral** is only continuous with reference to **uniform convergence** (cf.[5, Th 7.16]).

5.11 Integral on subsets: The integral $\int_A f d\mu := \int f|_A d\mu = \int f d\mu|_A$ on the subset $A \subset X$ of a measure space $(X; \mathcal{A}; \mu)$ is well defined since with $\int e|_A d\mu = \int e \cdot \chi_A d\mu = \int \left(\sum_{0 \le k \le n} a_i \chi_{A_k} \cdot \chi_A\right) d\mu$ $= \int \left(\sum_{0 \le k \le n} a_i \chi_{A_k \cap A}\right) d\mu = \sum_{0 \le k \le n} a_i \cdot \mu\left(A_k \cap A\right) = \sum_{0 \le k \le n} a_i \cdot \mu|_A \left(A_k\right) = \int e d\mu|_A$ we obtain the right hand identity for elementary resp. with 5.4 for integrable functions and since $\sup_{n \in \mathbb{N}} e_n = f$ we have $\sup_{n \in \mathbb{N}} e_n|_A = f|_A$ and hence the left hand identity $\int f|_A d\mu = \sup_{n \in \mathbb{N}} \int e_n|_A d\mu = \sup_{n \in \mathbb{N}} \int e_n d\mu|_A = \int f d\mu|_A$. I_A with $I_A f = \int_A f d\mu$ is the restriction of the Functional I to the vector subspace of the A-measurable functions and hence is again **linear** and **monotone**. For **disjoint subsets** $A, B \subset X$ with $A \cap B = \emptyset$ we obviously have $\int_{A \mathring{\cup} B} f d\mu = \int_A f d\mu + \int_b f d\mu$.

5.12 Integral for real valued functions: For a Borel measurable function $f : X \to \overline{\mathbb{R}}$ we define $\int f d\mu := \int f^+ d\mu - \int f^- d\mu$ and f is **integrable** iff the **positive part** $f^+ = \max\{0; f\}$ and the **negative part** $f^- = -\min\{0; f\}$ are integrable. **Linearity** for $\alpha \ge 0$ is proven as in 5.4 and in the case of $\alpha = -|\alpha| < 0$ the integral is decomposed into its positive and negative parts $\int \alpha f d\mu = \int \left(-|\alpha| f\right) d\mu = \int \left(-|\alpha| f\right)^+ d\mu - \int \left(-|\alpha| f\right)^- d\mu = \int |\alpha| f^- d\mu - \int |\alpha| f^+ d\mu = -|\alpha| \left(\int f^+ d\mu - \int f^- d\mu\right) = \alpha \int f d\mu$. Likewise the **additivity** is shown by thew separate observation of positive and negative parts. The **monotonicity** directly results from the respective property of the intergal for positive functions. On the one hand the **separate computing of positive and negative parts** provides the basis for **dominated convergence** which does not hold for the **Riemann integral.** On the other hand it entails the failure of the Lebesgue integral in the case of certain integrands with alternating signs like e.g. $\int_0^\infty \frac{\sin(x)}{x} dx = \frac{\pi}{2}$. (cf. [6, p.115 ex. 22])

5.13 Integral for complex-valued functions: For a Borel-measurable function $f : X \to \overline{\mathbb{C}}$ we define $\int f d\mu := \int \mathrm{Re} f d\mu + i \int \mathrm{Im} f d\mu$ and f is **integrable** iff the **real part** $\mathrm{Re} f$ and the **imaginary part** $\mathrm{Im} f$ are integrable. Because of $\max\{|\mathrm{Re} f|; |\mathrm{Im} f|\} \le \|f\| \le |\mathrm{Re} f| + |\mathrm{Im} f|$ the function f is integrable iff its **absolute value** $\|f\|$ is integrable. The linearity for **real** $\alpha \in \mathbb{R}$ is proven as in 5.12 whereas for **imaginary** $i\beta$ with $\beta \in \mathbb{R}$ is follows from $\int i\beta f d\mu = \int \mathrm{Re}\left(i\beta f\right) d\mu + i \int \mathrm{Im}\left(i\beta f\right) d\mu = \int \beta \mathrm{Im}\left(-f\right) d\mu + i \int \beta \mathrm{Re} f d\mu = -\beta \int \mathrm{Im} f d\mu + i\beta \int \mathrm{Re} f d\mu = i\beta \int f d\mu$ und and finally for complex $\alpha + i\beta$ from the **distributive law** and the **additivity** which itself is a consequence of separate computation of the real and imaginary parts. This functional is **not montone any more** e.g. for $f = 1$ and $g(t) = 2e^{it}$ on $X = [0; 2\pi]$ on the one hand we have $1 = \|f\| < \|g\| = 2$ but on the other hand $2\pi = \|\int f d\mu\| > \|\int g d\mu\| = 0$. But as the following theorems show almost all other properties are preserved such that this semi-extension of the integral is still denoted as an **integral**.

5.14 Absolute value: For every Borel-measurable $f : X \to \overline{\mathbb{C}}$ we have $\|\int f d\mu\| \le \int \|f\| d\mu$.

Proof: From $\int f d\mu = e^{i\alpha} \cdot \|\int f d\mu\|$ follows $\|\int f d\mu\| = e^{-i\alpha} \cdot \int f d\mu = \int \left(e^{-i\alpha} \cdot f\right) d\mu = \int \left(\mathrm{Re}\left(e^{-i\alpha} \cdot f\right)\right) d\mu$ $\le \int \|f\| d\mu$. In the case of equality $\|\int f d\mu\| = \int \|f\| d\mu$ according to 3 we have μ-a.e. $e^{-i\alpha} \cdot f = \|f\|$ resp. $f = e^{i\beta} \cdot \|f\|$.

5.15 Mean value property: For every Borel measurable $f : X \to \overline{\mathbb{C}}$ on a measure space $(X; \mathcal{A}; \mu)$ with $\int_B f d\mu = \int f d\mu|_B < \infty$ for some $B \in \mathcal{A}$ and the mean value $\frac{1}{\mu|_B(A)} \int_A f d\mu|_B \in S$ for some **closed** subset $S \subset \mathbb{C}$ and every$A \in \mathcal{A}$ with $\mu|_B(A) > 0$ we have $\mu|_B$-a.e. $f(x) \in S$.

Proof: For any closed disk $\overline{B}_r(z) \subset \mathbb{C} \backslash S$ with $\mu|_B \left(\left\{f \in \overline{B}_r(z)\right\}\right) > 0$ we have $\left|\frac{1}{\mu|_B(A)} \int_A f d\mu|_B - z\right| = \left|\frac{1}{\mu|_B(A)} \int_A \left(f - z\right) d\mu|_B\right| \le \frac{1}{\mu|_B(A)} \int_A |f - z| d\mu|_B \le r$ contrary to $\frac{1}{\mu(A)} \int_A f d\mu \in S$. Therefore we must assume $\mu\left(\left\{f \in \overline{B}_r(z)\right\}\right) = 0$ and since $\mathbb{C} \setminus S$ is a countable union of such disks the assertion follows from the σ-additivity of μ.

5.16 Extension of dominated convergence: Theorem 5.8 can be transferred to functions $f : X \to \overline{\mathbb{R}}$ by replacing g with $2g$ to compensate for opposite signs in f and f_n. In the last step of the proof we can apply 5.14 to extend the proposition further to functions $f : X \to \overline{\mathbb{C}}$ with the majorant of course referring to the **absolute value**: $\|g\| \in \mathcal{L}^1(\mu)$ mit $\|f_n\| \le \|g\| \, \forall n \in \mathbb{N}$.

5.17 Dominated convergence for series of complex functions: For a sequence $(f_n)_{n\in\mathbb{N}}$ of μ-**almost everywhere defined complex-valued** functions 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: The function $\sum\limits_{n\in\mathbb{N}} f_n$ is defined on the intersection $\bigcap\limits_{n\in\mathbb{N}} D_n := D$ of the domains D_n of the f_n with $\mu\left(X\setminus D\right)=\mu\left(\bigcup\limits_{n\in\mathbb{N}}\left(X\setminus D_n\right)\right)=0$. Owing to5.13 resp. the triangle inequality resp. 5.5 we have $\left\|\int f d\mu\right\|\le\int\|f\|\,d\mu\le\int\sum\limits_{n\in\mathbb{N}}\|f_n\|\,d\mu=\sum\limits_{n\in\mathbb{N}}\int\|f_n\|\,d\mu<\infty$, i.e. $f\in\mathcal{L}^1(\mu)$ and hence μ-a.e. **finite** resp**. convergent** due to 5.7.1. The convergence **in mean** follows from 5.16 with the **majorant** $g:=\sum\limits_{n\in\mathbb{N}}\|f_n\|$.

6 L^p-spaces

6.1 Convex functions: A real function $f:\overline{\mathbb{R}}\to\overline{\mathbb{R}}$ is **convex** on the open interval $]a;b[$ iff $f(s)\le f(r)+(s-r)\cdot\frac{f(t)-f(r)}{t-r}=f(t)-(t-s)\cdot\frac{f(t)-f(r)}{t-r}$ resp. $\frac{f(t)-f(s)}{t-s}\ge\frac{f(t)-f(r)}{t-r}\ge\frac{f(s)-f(r)}{s-r}$ 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 $f:]a;b[\to\overline{\mathbb{R}}$ we have $f\left(\frac{1}{\mu(A)}\int_A g d\mu\right)\le\frac{1}{\mu(A)}\int_A(f\circ g)\,d\mu$.

Proof: For $s:=\frac{1}{\mu(A)}\int_A g d\mu$ we have $a<s<b$ and due to 6.1 also $\beta:=\sup\limits_{a<r<s}\frac{f(s)-f(r)}{s-r}\le\frac{f(t)-f(s)}{t-s}$ for all $s<t<b$, hence $f(s)+\beta(t-s)\le f(t)$ resp. $f(s)+\beta(g(x)-s)\le f\left(g(x)\right)$. All summands of this inequality are integrable over A such that on account of the monotonicity of the integral we can infer $\mu(A)\cdot f(s)\le\int_A(f\circ g)\,d\mu$ 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. $x_1^{\alpha_1}\cdot...\cdot x_n^{\alpha_n}\le\alpha_1x_1+...+\alpha_nx_n$
2. $(x_1\cdot...\cdot x_n)^{\frac{1}{n}}\le\frac{1}{n}(x_1+...+x_n)$ (**geometric** and **arithmetic mean** for $\alpha_i:=\frac{1}{n}$)
3. $F\cdot G\le\frac{1}{p}F^p+\frac{1}{q}G^q$ 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 $f,g:X\to[0;\infty]$ on a measure space $(X;\mathcal{A},\mu)$ 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. $\|fg\|\le\|f\|_p\cdot\|g\|_q$ (**Hölder** resp. **Schwarz** for $p=q=2$) with equality iff μ-a.e. $\frac{f(x)}{\|f\|_p}=\frac{g(x)}{\|g\|_q}$.
2. $\|f+g\|\le\|f\|_p+\|g\|_p$ (**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 4.5. For one of the integrals disappearing 5.7.1 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 $F(x):=\frac{f(x)}{\|f\|_p}$ resp. $G(x):=\frac{g(x)}{\|g\|_q}$ in 6.3.3 an integration yields $\int(F\cdot G)\,d\mu\le\frac{1}{p}+\frac{1}{q}=1$ 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 $\|f+g\|_p^p\le\|f\|_p\cdot\left\|(f+g)^{p-1}\right\|_q+\|g\|_p\cdot\left\|(f+g)^{p-1}\right\|_q=\left(\|f\|_p+\|g\|_p\right)\cdot\|f+g\|_p^{\frac{p}{q}}$. Substituting

$p-\frac{p}{q}=1$ yields the assertion. The convexity of t^p provides the inequality $\left(\frac{f+g}{2}\right)^p \le \frac{f^p+g^p}{2}$, i.e. the integrability of f^p and g^p entails the integrability of $(f+g)^p$.

6.5 L^p**-spaces**: For $1 \le p < \infty$ the above introduced expression $\|f\|_p := \left(\int |f|^p d\mu\right)^{\frac{1}{p}}$ resp. $\|f\|_\infty := \inf\{0<\alpha<\infty : \mu(\{|f|>\alpha\})=0\}$ define a **pseudonorm** (cf. [8, 21.1]) on the **vector space** $\mathcal{L}^p(\mu) := \left\{f : (X;\mathcal{A};\mu) \to \left(\overline{\mathbb{C}};\overline{\mathcal{B}};\lambda\right) : \|f\|_p :< \infty\right\}$. The **absolute homogeneity** follows from the **linearity** 5.1.1 whereas the **triangle inequation** is provided by the **Hölder inequality** 6.4.2. $\mathcal{L}^1(\mu)$ contains the **integrable** functions and $\mathcal{L}^\infty(\mu)$ is the set of all μ-a.e. **bounded** and measurable functions furnished with the **supremum norm** $\|\|_\infty$. The contraction to the **quotient space** $L^p(\mu) := \mathcal{L}^p / \sim$ defined by the **equivalence relation** $f \sim g \Leftrightarrow \mu(\{f \ne g\}) = 0$ makes $\|\|_p$ a **norm**. Convergence with respect to $\|\|_p$ is called in the p**-th mean**. On account of $\mu(\{|f| > \|f\|_\infty\}) = \mu\left(\bigcup_{n\ge 1}\left\{|f| > \|f\|_\infty + \frac{1}{n}\right\}\right) = 0$ and 5.7.1 we have $\mu(\{f=\infty\}) = 0$ for all $f \in \mathcal{L}^p$ and $1 \le p \le \infty$. Since every equivalence class contains representants with only finite values we can restrict our observations to the range $\mathbb{C}$. All subsequent results refer to measurable functions $f : (X;\mathcal{A};\mu) \to \left(\overline{\mathbb{C}};\overline{\mathcal{B}};\lambda\right)$ on a measure space $(X;\mathcal{A};\mu)$.

6.6 Relations between L^p**-spaces:** For $1 \le p, q \le \infty$ 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 induction. E.g. owing to 5.9.2 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 $p = \frac{r}{s} \ge 1$, $f = h^s$ and $g = 1$ Hölder 6.4.1 yields $\int |h|^s d\mu \le \left(\int |h|^r d\mu\right)^{\frac{s}{r}} \cdot \left(\int 1 d\mu\right)^{\frac{s-s}{r}}$ resp. $\|h\|_s = \left(\int |h|^s d\mu\right)^{\frac{1}{s}} \le \left(\int |h|^r d\mu\right)^{\frac{1}{r}} \cdot (\mu(X))^{\frac{1}{s}-\frac{1}{r}} = \|h\|_r \cdot (\mu(X))^{\frac{1}{s}-\frac{1}{r}}$ and hence the assertion.
2. On account of **Zorn's lemma** ([8, 14.2.4]) the set $\{|f| \ge 1\}$ 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 $\mu(f \ge 1) < \infty$ and since μ is **bounded below** this maximal partition consists of $n := \frac{\mu(f\ge 1)}{\alpha} + 1$ sets $(A_i)_{1\le i\le n}$ with $\mu(A_i) > \alpha$. Owing to 5.3 for every $\epsilon > 0$ there is an elementary function $e = \sum_{i=1}^{n} \alpha_i \chi_{A_i} \le f$ with $\int_{\{|f|\ge 1\}} e d\mu = \sum_{i=1}^{n} \alpha_i \mu(A_i) \ge \int_{\{|f|\ge 1\}} |f|^p d\mu - \epsilon \cdot \alpha$. Hence on the one hand for every $x \in A_i$ with $1 \le i \le n$ we have $|f|^p(x) \ge \alpha_i \Leftrightarrow |f|^q(x) \ge \alpha_i^{\frac{q}{p}}$ and on the other hand for every $1 \le i \le n$ there is an $x_i \in A_i$ with $\alpha_i \ge |f^p(x_i)| - \epsilon \Leftrightarrow \alpha_i^{\frac{q}{p}} \ge (|f^p(x_i)| - \epsilon)^{\frac{q}{p}} \ge |f^q(xxi)| - \epsilon \cdot \frac{q}{p} \cdot (|f^p(x_i)| - \epsilon)^{\frac{q}{p}-1} \ge |f^q(x_i)| - \epsilon \cdot \frac{q}{p} \cdot |f^{q-p}(x_i)|$ 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 $\int_{\{|f|\ge 1\}} |f|^q d\mu < \sum_{i=1}^{n} \left(\alpha_i^{\frac{q}{p}} + \epsilon \cdot \frac{q}{p} \cdot |f^{q-p}(x_i)|\right) \chi_{A_i} < \infty$ and also on the whole set $\int |f|^q d\mu = \int_{\{|f|<1\}} |f|^q d\mu + \int_{\{|f|\ge 1\}} |f|^q d\mu \le \int_{\{|f|<1\}} |f|^p d\mu + \int_{\{|f|\ge 1\}} |f|^q d\mu < \infty$.

6.7 Completeness: For $1 \le p \le \infty$ every **Cauchy sequence**$(f_n)_{n\in\mathbb{N}} \subset L^p(\mu)$ converges in the p**-th mean** to a $f \in L^p(\mu)$. Moreover there is a **subsequence** $(f_{n_k})_{k\in\mathbb{N}}$ converging μ**-a.e.** to f. 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 $\left\|\sum_{i=0}^{k} |f_{n_{i+1}} - f_{n_i}|\right\|_p \le 1$ due to 6.4.2, hence $\left\|\sum_{i=0}^{\infty} |f_{n_{i+1}} - f_{n_i}|\right\|_p \le 1$ owing to 5.5 and finally μ**-a.e.** $g := \sum_{i=0}^{\infty} |f_{n_{i+1}} - f_{n_i}| < \infty$ according to 5.7.1. Hence the complex-valued sequence $(f_{n_i})_{i\in\mathbb{N}} = \sum_{k=1}^{i}\left(f_{n_k} - f_{n_{k+1}}\right)$ converges **absolutely** and **μ-a.e.** to a μ-a.e. bounded $f = \lim_{i\to\infty} f_{n_i} = \sum_{i=0}^{\infty}\left(f_{n_{i+1}} - f_{n_i}\right)$ with $|f| < g$. On account of the completeness of μ (cf. 3.9) we can simply 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 $m \geq n_j$ and from 5.6 follows $\left\|f - f_{n_j}\right\|_p \leq \liminf_{m\geq n_j}\left\|f_m - f_{n_j}\right\|_p < \epsilon^p$, i.e. the subsequence $(f_{n_i})_{i\in\mathbb{N}}$ and hence the entire Cauchy sequence $(f_n)_{n\in\mathbb{N}}$ (cf. [8, 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 $\mathbb{C}$ 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:: $\lim_{n\to\infty}\|f_n\|_p = \lim_{n\to\infty}\left(\lambda\left(\left[\frac{n}{2^k}; \frac{n+1}{2^k}\right]\right)\right)^{\frac{1}{p}} = \lim_{n\to\infty} 2^{-\frac{k(n)}{p}} = \lim_{n\to\infty}\frac{1}{n^{1/p}} = 0$ 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. As in 5.4 this proof requires not only the **completeness** but also the **supremum property** resp. the **linear order** of $\mathbb{R}$.
3. $L^2(\mu)$ 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 Elementary functions: The **algebra** $\mathcal{E}(X) = \left\{\sum_{0\leq i\leq m}\alpha_i\chi_{A_i} : m \in \mathbb{N}, \alpha_i \in \mathbb{C}, A_i \in \mathcal{A}\right\}$ of the **complex-valued elementary functions** for $1 \leq p \leq \infty$ is **dense** in $L^p(\mu)$. (cf. 4.8)

Proof: The following argument repeatedly makes use of the **pointwise continuity** of the **power function** $x \to x^\alpha$ for $\alpha > 0$: According to 4.9 in the case of $1 \leq p < \infty$ for every **real-valued and positive** $f \in L^p(\mu)$ resp. $f^p \in L^1(\mu)$ there is an increasing sequence $(e_n)_{n\in\mathbb{N}} \subset \mathcal{E}(X)$ of **real-valued and positive elementary functions** with μ-a.e. $\lim_{n\to\infty}\left(f - e_n^{1/p}\right)^p = \lim_{n\to\infty}\left(f - e_n^{1/p}\right) = \lim_{n\to\infty}(f^p - e_n) = 0$ and the **majorant** f^p for all $\left(f - e_n^{1/p}\right)^p \in L^1(\mu)$ and hence $\lim_{n\to\infty}\left\|f - e_n^{1/p}\right\|_p = \lim_{n\to\infty}\left\|\left(f - e_n^{1/p}\right)^p\right\| = 0$ according to 5.8. In the case $p = \infty$ for every **real-valued and positive** $f \in L^\infty(\mu)$ the sequence $(s_n \circ f)_{n\geq 0}$ of elementary functions from 4.9 **uniformly converges** to f, i.e. $\lim_{n\to\infty}\|s_n \circ f - f\|_\infty = 0$ since $n_0 = \|f\|_\infty$ now is independent from x. The extension to complex-valued functions follows from 5.13.

6.10 Convergence in the p-th mean and in measure: If a sequence $(f_n)_{n\geq 1} \subset L^p(\mu)$ for a $p \in [1; \infty[$ **converges in the p-th mean** to a $f \in L^p(\mu)$ it also **converges in measure** to f.

Proof: Due to the hypothesis for $\epsilon > 0$ there is an $n_0 \geq 1$ such that for all $n \geq n_0$ we have $\|f_n - f\|_p < \epsilon^{1+\frac{1}{p}} \Rightarrow \mu(|f_n - f| > \epsilon) = \mu(|f_n - f|^p > \epsilon^p) \leq \frac{1}{\epsilon^p}\int_{\{|f_n - f|^p > \epsilon^p\}}|f_n - f|^p d\mu \leq \frac{1}{\epsilon^p}\int |f_n - f|^p d\mu < \frac{1}{\epsilon^p}\epsilon^{p+1} = \epsilon$.

6.11 Lemma: For an **integrable** $f : (X; \mathcal{A}; \mu) \to \left(\overline{\mathbb{C}}; \overline{\mathcal{B}}; \lambda\right)$ 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 $\varphi_n(x) = \begin{cases}|f(x)|, & \text{for } |f(x)| \leq n\\ n & \text{else}\end{cases}$ satisfies the conditions for 5.5 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 $\int_E |f| \, d\mu \leq \int_E (|f| - \varphi_n) \, d\mu + \int_E \varphi_n d\mu \leq \epsilon$.

6.12 Vitali's **convergence theorem**: A sequence $(f_n)_{n\geq 1} \subset L^p(\mu)$ **converging** μ**-a.e.** for a $p \in [1;\infty[$ to a μ**-a.e. finite and measurable** f **also converges in the** p**-th mean** to f and we have $f \in L^p(\mu)$ **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 5.5 with $|f|^p = \sup_{m\geq 1} |f|^p \cdot \chi_{\left\{|f|^p > \frac{1}{m}\right\}}$ and $f \in L^p(\mu)$ there is an $m_0 \geq 1$ with $\int |f|^p \cdot \chi_{\left\{|f|^p \leq \frac{1}{m}\right\}} d\mu = \int |f|^p \, d\mu - \int |f|^p \cdot \chi_{\left\{|f|^p > \frac{1}{m}\right\}} d\mu < \epsilon$ 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 $1 \leq n < n_0$. Hence for $A_\epsilon = B_\epsilon \cup C_\epsilon$ and with 6.4.2 we obtain the estimate $\left(\int_{X\setminus A_\epsilon} |f_n|^p \, d\mu\right)^{\frac{1}{p}} \leq \left(\int_{X\setminus A_\epsilon} |f_n - f|^p \, d\mu\right)^{\frac{1}{p}} + \left(\int_{X\setminus A_\epsilon} |f|^p \, d\mu\right)^{\frac{1}{p}} < \epsilon^{\frac{1}{p}} + \epsilon^{\frac{1}{p}}$ 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 6.11 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** 5.6 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** 5.9 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 4.3 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. [8, 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

$\left(\pi^J_{J_n}\right)^{-1}[A_{J_n}] \in \mathcal{A}_J$ 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. [8, 4.2]) and hence due to 4.3 **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 [8, 7.10] the **separation axiom** T_2 extends to the product space and owing to **Tychonoff's theorem** (cf. [8, 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 $1 \le j \le m$ 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: In $\mathbb{R}^n$ resp. $\mathbb{C}^n \simeq \mathbb{R}^{2n}$ the **n-dimensional intervals** $\mathcal{I}^n := \left\{\prod\limits_{i=1}^{n}[a_i;b_i[: a_i \le b_i \in \mathbb{R}\right\}$ are measurable since they are G_δ. Their **finite unions** form the **ring** of the **n-dimensional figures** (cf. 1.2) and on account of $\prod\limits_{i=1}^{n}]a_i;b_i[= \bigcup_{k\in\mathbb{N}}\prod\limits_{i=1}^{n}\left[a_i + \frac{1}{k};b_i\right[$ generate the Borel σ-algebra: $\mathcal{B}^n := \bigotimes\limits_{i=1}^{n}\mathcal{B}_i = \sigma(\mathcal{I}^n) = \sigma(\mathcal{F}^n)$ (cf. 1.3). Hence resp. due to 1.1.2 the Borel σ-algebra is generated by the **products** $\prod\limits_{i=1}^{n}]a_i;\infty[$. The **one-point compactifications** $\overline{\mathbb{R}^n} = \mathbb{R}^n \cup \{\infty\}$ and particularly $\overline{\mathbb{C}}$ carry the Borel σ-algebrae $\overline{\mathcal{B}^n} = \{A; A\cup\{\infty\} : A \in \mathcal{B}^n\}$ 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 Definition: A measure μ is **σ-finite** on a measure space $(X;\mathcal{A};\mu)$ if there is a cover $(A_n)_{n\in\mathbb{N}} \subset \mathcal{A}$ of μ-finite sets A_n: $\bigcup_{n\in\mathbb{N}} A_n = X$ and $\mu(A_n) < \infty \forall n \in \mathbb{N}$.

8.3 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 4.8. 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 λ-π-theorem** 1.8 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 $\sup_{n\in\mathbb{N}} s_{1nA}(x_2) = \sup_{n\in\mathbb{N}} \mu_1|_{A_n}(A_{x_2}) = \sup_{n\in\mathbb{N}} \mu_1(A_n \cap A_{x_2}) = \mu_1\left(\bigcup_{n\in\mathbb{N}} A_n \cap A_{x_2}\right) = \mu_1(A_{x_2}) = s_{1A}(x_2)$. The proof for s_{2A} is of course analogous.

8.4 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 8.3 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 $(A_n)_{n\in\mathbb{N}} \subset \mathcal{A}_1 \otimes \mathcal{A}_2$ of pairwise disjoint measurable sets the **σ-additivity** of μ_1 and **monotone convergence** 5.5 applied to μ_2 yield $(\mu_1 \otimes \mu_2)\left(\mathring{\bigcup}_{n\in\mathbb{N}} A_n\right) = \int \mu_1\left(\left(\mathring{\bigcup}_{n\in\mathbb{N}} A_n\right)_{x_2}\right) d\mu_2 = \int \left(\sum_{n\in\mathbb{N}} \mu_1(A_n)_{x_2}\right) d\mu_2 = \sum_{n\in\mathbb{N}} \int \mu_1(A_{x_2})\, d\mu_2 = \sum_{n\in\mathbb{N}} (\mu_1 \otimes \mu_2)(A_n)$. 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 $(A_{1n} \times A_{2n})_{n\in\mathbb{N}} \subset \mathcal{A}_1 \times \mathcal{A}_2$ is a cover of $X_1 \times X_2$ from $\mu_1 \otimes \mu_2$-finite sets $A_{1n} \times A_{2n}$.

8.5 Fubini's theorem: For two **σ-finite** measure spaces $(X_i; \mathcal{A}_i; \mu_i)$ with $i \in \{1; 2\}$ and every **$\mathcal{A}_1 \otimes \mathcal{A}_2$-measurable** function $f : X_1 \times X_2 \to \mathbb{C}$ the mappings $x_1 \mapsto f_{x_2}(x_1) = f(x_1; x_2)$ resp. $x_2 \mapsto f_{x_1}(x_2) = f(x_1; x_2)$ as well as $x_1 \mapsto \varphi_1(x_1) = \int f_{x_1} d\mu_2$ resp. $x_2 \mapsto \varphi_2(x_2) = \int f_{x_2} d\mu_1$ are **measurable** with reference to $\mathcal{A}_1$ resp. $\mathcal{A}_2$ and the integral on the product space is computed by means of $\int f d(\mu_1 \otimes \mu_2) = \int \left(\int f_{x_1} d\mu_2\right) d\mu_1 = \int \left(\int f_{x_2} d\mu_1\right) d\mu_2$. If $|f_{x_2}| \in L^1(\mu_1)$ and $\int \left(\int |f_{x_2}|\, d\mu_1\right) d\mu_2 < \infty$ then $f \in L^1(\mu_1 \otimes \mu_2)$. If $f \in L^1(\mu_1 \otimes \mu_2)$ then $f_{x_2} \in L^1(\mu_1)$ for μ_2-a.e. $x_2 \in X_2$ and $\varphi_2 \in L^1(\mu_2)$ and analoguos results for f_{x_1} resp. φ_1.

Proof: As in 5.12 resp. 5.13 we decompose the integral into a sum $\int f d\mu := \int (\mathrm{Re} f)^+ d\mu - \int (\mathrm{Re} f)^- d\mu + i \int (\mathrm{Im} f)^+ d\mu - i \int (\mathrm{Im} f)^- d\mu$ of four integrals of positive functions. Hence we can assume $f : X_1 \times X_2 \to [0; \infty]$ and apply the result four times: The mapping $x_1 \mapsto f_{x_2}(x_1)$ is $\mathcal{A}_1$-measurable since for $A \in \overline{\mathcal{B}}_+$ we have $f_{x_2}^{-1}[A] = \{(\xi_1; x_2) : f(\xi_1; x_2) \in A\} = \left(f^{-1}[A]\right)_{x_2} \in \mathcal{A}_1$ due to 8.1. For **elementary functions** $f = \sum_{i=1}^{n} \alpha_i \chi_{A_i} \in \mathcal{E}(X_1 \times X_2)$ with $A_i \in \mathcal{A}_1 \otimes \mathcal{A}_2$ the mapping $x_1 \mapsto \int f_{x_1} d\mu_2 = \sum_{i=1}^{n} \alpha_i \int (\chi_{A_i})_{x_1} d\mu_2 = \sum_{i=1}^{n} \alpha_i \int \chi_{(A_i)_{x_1}} d\mu_2 = \sum_{i=1}^{n} \alpha_i \mu_2\left((A_i)_{x_1}\right)$ is $\mathcal{A}_1$-measurable on account of 8.3 such that $\int f d(\mu_1 \otimes \mu_2) = \sum_{i=1}^{n} \alpha_i (\mu_1 \otimes \mu_2)(A_i) = \sum_{i=1}^{n} \alpha_i \left(\int \mu_2\left((A_i)_{x_1}\right) d\mu_1\right) = \int \left(\sum_{i=1}^{n} \alpha_i \mu_2\left((A_i)_{x_1}\right)\right) d\mu_1 = \int \left(\int f_{x_1} d\mu_2\right) d\mu_1$ due to 8.4. Finally the integral formulae are transferred from elementary functions to numerical functions by applying **monotone convergence** 5.5.

If $|f_{x_2}| \in L^1(\mu_1)$ and $\int \left(\int |f_{x_2}|\, d\mu_1\right) d\mu_2 < \infty$ then obviously all four integrals are finite, i.e. $f \in L^1(\mu_1 \otimes \mu_2)$.

If $f \in L^1(\mu_1 \otimes \mu_2)$ the all four integrands are smaller then $|f|$ such that e.g. for $(\mathrm{Re}f)^+$ the function $\varphi_2$with $\varphi_2(x_2) = \int (\mathrm{Re}f)^+_{x_2} d\mu_1$ is μ_2- integrable and hence μ_2-a.e. finite which in turn means that $(\mathrm{Re}f)^+_{x_2} \in L^1(\mu_1)$ for μ_2-almost every $x_2 \in X_2$.

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 measure spaces $(X_i; \mathcal{A}_i; \mu_i)$ with a finite index set J 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 8.4 by means of $\bigotimes_{1\leq j<i} \mu_j := \left(\bigotimes_{1\leq j<i} \mu_j\right) \otimes \mu_i$. The resulting product of measure spaces is denoted as $\bigotimes_{i\in j}(X_i; \mathcal{A}_i; \mu_i) := \left(\prod_{i\in J} X_i; \bigotimes_{i\in j} \mathcal{A}_i; \bigotimes_{i\in J} \mu_i\right)$.

8.7 Integral on finite product spaces: For a Borel measurable function $f : \prod_{i\in J} X_i \to \overline{\mathbb{C}}$ on the product $\bigotimes_{i\in J}(X_i; \mathcal{A}_i; \mu_i)$ with $\mu := \bigotimes_{i\in J} \mu_i$ for $J = \{1, ..., n\}$ the expression $\int f d\mu := \int (\mathrm{Re}f)^+ d\mu - \int (\mathrm{Re}f)^- d\mu + i \int (\mathrm{Im}f)^+ d\mu - i \int (\mathrm{Im}f)^- d\mu$ defines an **integral** according to 5.1 if the four **inductively** and independently of the **permutation** $j : J \to J$ computed integrals $\int (\mathrm{Re}f)^+ d\mu := \int \left(... \left(\int (\mathrm{Re}f)^+ d\mu_{j(1)}\right) ...\right) d\mu_{j(n)}$ and correspondingly $\int (\mathrm{Re}f)^- d\mu$, $\int (\mathbf{Im}f)^+ d\mu$ as well as $\int (\mathbf{Im}f)^- d\mu$ are **finite**. In this case f is **integrable** and hence μ-a.e. **finite** according to 5.7.1.

8.8 Completion of λ^n: The product $\lambda^n = \bigotimes_{1\leq i\leq n} \lambda_0$ of the **complete** Lebesgue measures λ_0 on the product $\mathcal{B}^n = \bigotimes_{1\leq i\leq n} \mathcal{B}_0$ of the Lebesgue σ-algebrae $\mathcal{B}_0$ on $\mathbb{R}$ is **not complete** since for any λ-null set $A \in \mathcal{B}_0$ we have $\lambda^2(A \times \mathbb{R}) = 0$ and for any non Lebesgue measurable $B \notin \mathcal{B}_0$ (cf. 3.10) evidently $A \times B \subset A \times \mathbb{R}$ holds but $A \times B \notin \mathcal{B}_0^2$. 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** $\mathcal{B}^n$.

8.9 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 4.2 and 8.4 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 **$\sigma$-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.

8.10 Convolutions: For any pair of **integrable** functions $f, g : \mathbb{R} \to \mathbb{C}$ their **convolution** $h : \mathbb{R} \to \mathbb{C}$ defined by $h(x) = \int f(x - y) g(y) dy$ is also **integrable** with $\|h\|_1 \leq \|f\|_1 \cdot \|g\|_1$.

Proof: According to 4.5 the functions $\varphi, \psi : \mathbb{R}^2 \to \mathbb{R}$ defined by $\varphi(x; y) = x - y$ resp. $\psi(x; y) = y$ are **Borel** measurable and so is the **product** of the **compositions** $(f \circ \varphi) \cdot (g \circ \psi)$. as well as its absolute value $|(f \circ \varphi) \cdot (g \circ \psi)|$. Hence we may apply **Fubini's theorem** 8.5 and since the **translation invariance** 8.9 of λ yields $\int |f(x - y)| dx = \|f\|_1$ for every $y \in \mathbb{R}$ we can write

$$\begin{aligned}
\|h\|_1 &= \int \left| \int f(x - y) g(y) dy \right| dx \\
&\leq \int \int |f(x - y) g(y)| dy dy \\
&= \int |(f \circ \varphi) \cdot (g \circ \psi)| d\lambda^2 \\
&= \int \int |f(x - y) g(y)| dx dy \\
&= \|f\|_1 \cdot \int |g(y)| dx \\
&= \|f\|_1 \cdot \|g\|_1 .
\end{aligned}$$

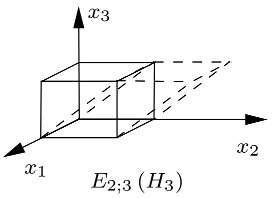
8.11 Transformation formula: The image of the **Lebesgue-Borel measure** λ^n under a **homomorphism** $T \in GL(n;\mathbb{R})$ is $T \circ \lambda^n = \frac{\lambda^n}{|\det T|}$ such that $\lambda^n(T[A]) = |\det T| \cdot \lambda^n(A)$ for every Borel-measurable $A \in \mathcal{B}^n$. In $\mathbb{R}^3$ 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** $Q = \left\{\sum_{1\leq i\leq 3} x_i\mathbf{e}_i : 0 \leq x_i < 1\right\}$. The volume of the parallelepiped is then $\lambda^3(T[Q]) = |\det T| \cdot \lambda^3(Q) = \det(\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3) \cdot 1 = \mathbf{a}_1 * \mathbf{a}_2 \times \mathbf{a}_3$. The **volume of the ball** $B_r(0) = S[B_1(0)]$ with $S(\mathbf{e}_i) = r\mathbf{e}_i$ for $0 \leq i \leq n$ is $\lambda^n(B_r(0)) = \lambda^n(S[B_1(0)]) = |\det S| \cdot \lambda^n(B_1(0)) = r^n \cdot \lambda^n(B_1(0))$. The volume $\lambda^n(B_1(0))$ of the **unit ball** itself will have to wait until we can use the **change-of-variables theorem** 14.6.

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:

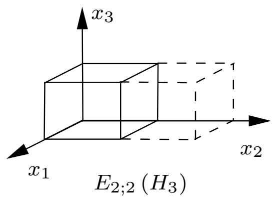
$$E_{kl} = \begin{array}{cc} \begin{array}{ccccccc} 1 & \cdots & k & \cdots & l & \cdots & n \end{array} & \\ \begin{pmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 1 & & 1 & & \\ & & & \ddots & & & \\ & & & & 1 & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{pmatrix} & \begin{array}{c} 1 \\ \vdots \\ k \\ \vdots \\ l \\ \vdots \\ n \end{array} \end{array}
\qquad
E_{k\alpha} = \begin{array}{cc} \begin{array}{ccccc} 1 & \cdots & k & \cdots & n \end{array} & \\ \begin{pmatrix} 1 & & & & \\ & \ddots & & & \\ & & \alpha & & \\ & & & \ddots & \\ & & & & 1 \end{pmatrix} & \begin{array}{c} 1 \\ \vdots \\ k \\ \vdots \\ n \end{array} \end{array}$$

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

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



$E_{2;3}(H_3)$



$E_{2;2}(H_3)$

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 **mulitplicity of the determinant**: $|\det(A \cdot B)| = |\det(A)| \cdot |\det(B)|$**.**

8.12 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 [8, 15.5]). Owing to 8.4 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 $(Z_n)_{n\geq 1} \subset \mathcal{Z}_I$ 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 $n \geq 1$ such that $\inf_{n\geq 1} \mu_I(Z_n) \geq \alpha$. In order to show the **$\emptyset$-continuity** we have to prove that $\bigcap_{n\geq 1} Z_n \neq \emptyset$, i.e. we must find a path $x \in \bigcap_{n\geq 1} Z_n$. 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 8.3 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 $n \geq 1$. Since μ_{J_1} is continuous from above and $Q_{n+1}^{J_1} \subset Q_n^{J_1}$ for all $n \geq 1$ there is an $x_{J_1} \in \bigcap_{n\geq 1} Q_n^{J_1} \neq \emptyset$, i.e. $\mu_I\left(Z_n^{x_{J_1}}\right) \geq \frac{\alpha}{2}$ for all $n \geq 1$.

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 $n \geq 1$ we have

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

and hence

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

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 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 1.8 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 $f : X \to [0; \infty]$ 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 **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-valued** random variables $f, g \in L^1(\mu)$ we have $E(f \cdot g) = E(f) \cdot E(g)$ and and particularly $f \cdot g \in L^1(\mu)$.

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 5.1.1 also for **elementary** functions $\varphi, \gamma \in \mathcal{E}(X)$. **Positive and integrable** functions $f, g : X \to [0; \infty]$ with $f = \sup_{n\in\mathbb{N}} f_n \in L^1(\mu)$ resp. $g = \sup_{n\in\mathbb{N}} g_n \in L^1(\mu)$ for increasing sequences $(f_n)_{n\in\mathbb{N}}, (g_n)_{n\in\mathbb{N}} \subset \mathcal{E}(X)$ according to 4.9 are μ-a.e. **finite** due to 5.7.1 such that the product must be μ-a.e. finite too with $f \cdot g = \sup_{n\in\mathbb{N}} f_n \cdot g_n$ and owing to 5.3 we obtain $E(f \cdot g) = \sup_{n\in\mathbb{N}} E(f_n \cdot g_n) = \sup_{n\in\mathbb{N}} (E(f_n) \cdot E(g_n)) = \sup_{n\in\mathbb{N}} E(f_n) \cdot \sup_{n\in\mathbb{N}} E(f_n) = E(f) \cdot E(g)$. Finally we can extend the proposition to **real-valued and integrable** function in general since with the linearity of the integral for $f = f^+ - f^-, g = g^+ - g^- \in L^1(\mu)$ with positive f^+, f^-, g^+, g^- the assertion follows in the form $E(f \cdot g) = E(f^+ \cdot g^+ - f^+ \cdot g^- - f^- \cdot g^+ + f^- \cdot g^-) = E(f^+) \cdot E(g^+) - E(f^+) \cdot E(g^-) - E(f^-) \cdot E(g^+) + E(f^-) \cdot E(g^-) = E(f) \cdot E(g)$.

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 $\mu(f \le m_{min}(f)) = \mu\left(\bigcap_{n\ge1}\left\{f \le m_{min}(f) + \frac{1}{n}\right\}\right) = \inf_{n\ge1} \mu\left(f \le m_{min}(f) + \frac{1}{n}\right) \ge \frac{1}{2}$ and on the other hand $\mu(f \ge m_{min}(f)) = \mu\left(\bigcap_{n\ge1}\left\{f \ge m_{min}(f) - \frac{1}{n}\right\}\right) = \inf_{n\ge1} \mu\left(f \ge m_{min}(f) - \frac{1}{n}\right) = 1 - \sup_{n\ge1} \mu\left(f < m_{min}(f) - \frac{1}{n}\right) \ge \frac{1}{2}$, 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 $\sup_{n\ge1} \mu\left(f \ge m_{max}(f) + \frac{1}{n}\right) = \mu\left(\bigcup_{n\ge1}\left\{f \ge m_{max}(f) + \frac{1}{n}\right\}\right)$ $= \mu(f > m_{max}(f)) > \frac{1}{2}$, 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 $F_m := \sum_{i=1}^{m} f_i$ and every $\epsilon > 0$ we have $\mu\left(\max_{1\le i\le n} |F_i + m(F_n - F_i)| \ge \epsilon\right) \le 2\mu(|F_n| \ge \epsilon)$.

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 $T = \min_{1\le i\le m} \{|F_i + m(F_n - F_i)| \ge \epsilon\}$ 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 $\mu(F_n \ge \epsilon) \ge \mu\left(\bigcup_{i=1}^{n} A_i \cap B_i\right)$ $= \sum_{i=1}^{n} \mu(A_i \cap B_i) = \sum_{i=1}^{n} \mu(A_i) \cdot \mu(B_i) \ge \frac{1}{2}\mu(1 \le T \le n) = \frac{1}{2}\mu\left(\max_{1\le i\le n} F_i + m(F_n - F_i) \ge \epsilon\right)$. 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 $\mu$and simply **add** the two inequalities to obtain the assertion.

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

Proof:

$\Rightarrow$: **Lebesgue's convergence theorem** 4.11.

$\Leftarrow$: **Riesz' convergence theorem** 4.13.3 provides for every $\frac{1}{4} > \epsilon > 0$ an $n_\epsilon \geq 1$ with $\mu(|F_n - F_m| \geq \epsilon) < \epsilon$ for all $n > m \geq n_\epsilon$. In particular we have $\mu(|F_n - F_m| \geq \epsilon) < \frac{1}{2}$ and hence $|m(F_n - F_m)| \leq \epsilon$ for $n > m \geq n_\epsilon$. The preceding inequality yields $\mu\left(\max_{m<i\leq n} |F_i - F_m| \geq 2\epsilon\right) \leq 2\mu(|F_n - F_m| \geq \epsilon) < 2\epsilon$. For $n \to \infty$ follows $\mu\left(\sup_{m<i} |F_i - F_m| \geq 2\epsilon\right) \leq 2\epsilon$ 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 $A_n = \sum_{i=0}^{n} a_i$ we have $\sum_{i=1}^{n} a_i b_i = A_n b_n - A_0 b_1 - \sum_{i=1}^{n-1} A_i (b_{i+1} - b_i)$ for $n \geq 1$.
2. If also $\lim_{n\to\infty} A_n = A_0^* < \infty$ with $A_n^* = \sum_{i>n} a_i$ holds we have $\sum_{i=1}^{n} a_i b_i = A_0^* b_1 - A_n^* b_n + \sum_{i=1}^{n-1} A_i^* (b_{i+1} - b_i)$ 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 $\sum_{i=1}^{n} a_i b_i = A_0^* b_1 + \sum_{i=1}^{n-1} A_i^* (b_{i+1} - b_i)$ for $n \geq 1$.

Proof:

1. $\sum_{i=1}^{n} a_i b_i = \sum_{i=0}^{n-1} (A_{i+1} - A_i) b_{i+1} = A_n b_n - \sum_{i=1}^{n-1} A_i (b_{i+1} - b_i) - A_0 b_1$.
2. Follows from 1. with $a_0 = -\sum_{i=1}^{\infty} a_i = -A_0^*$.
3. In the case of $\lim_{n\to\infty} A_n^* b_n > 0$ with $\sum_{i>n} a_i b_i \geq A_n^* b_n$ and 2. we have $A_0^* b_1 + \sum_{i\geq 1} A_i^* (b_{i+1} - b_i) \geq \sum_{i>1} a_i b_i = \infty$ and hence the assertion. For $\lim_{n\to\infty} A_n^* b_n = 0$ 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 $\lim_{i\to\infty} \frac{1}{b_i} = 0$ and a further **real** sequence $(a_i)_{i\geq 1}$ with $\sum_{i\geq 1} \frac{a_i}{b_i} < \infty$ we have $\lim_{n\to\infty} \frac{1}{b_n} \sum_{i=1}^{n} a_i = 0$.

Proof: From 9.9.2 with $c_i = \frac{a_i}{b_i}$ and $\lim_{n\to\infty} C_n = C_0^* = \sum_{i\geq 1} \frac{a_i}{b_i} < \infty$ resp. $\lim_{n\to\infty} C_n^* = 0$ we have the decomposition $\frac{1}{b_n} \sum_{i=1}^{n} a_i = \frac{1}{b_n} \sum_{i=1}^{n} c_i b_i = \frac{1}{b_n} C_0^* b_1 + C_n^* + \frac{1}{b_n} \sum_{i=1}^{n-1} C_i^* (b_{i+1} - b_i)$. 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 $\left|\frac{1}{b_n} \sum_{i=m}^{n-1} C_i^* (b_{i+1} - b_i)\right| < \epsilon \frac{1}{b_n} \sum_{i=m}^{n-1} (b_{i+1} - b_i) = \epsilon\left(1 - \frac{b_m}{b_n}\right) < \epsilon$ and on the other hand $\left|\frac{1}{b_n} \sum_{i=1}^{m-1} C_i^* (b_{i+1} - b_i)\right| < \epsilon$ 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 $\sum_{n\geq 1} E(f_n^2) < \infty$ the **sums** $F_m := \sum_{n=1}^{m} f_n$ **converge** μ**-a.e.** and in **quadratic mean** to a $F = \lim_{m\to\infty} F_m \in L^2(\mu)$ with $E(F)^2 = \sum_{n\geq 1} E(f_n^2)$.

Proof : Owing to 9.5, $E(f_n) = 0$ for all n≥ 1 and the hypothesis we have $\lim\limits_{k\to\infty}\sup\limits_{m\geq k} E(F_m - F_k)^2 = \lim\limits_{k\to\infty}\sup\limits_{m\geq k} \sum\limits_{i=k}^{m} E(f_i^2) = 0$ 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**: $\lim\limits_{m\to\infty} E(F - F_m)^2 = 0$. Owing to 6.10 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 $E(f_n) = 0$ we also obtain $E(F)^2 = \lim\limits_{m\to\infty} E(F_m)^2 = \sum_{n\geq 1} E(f_n^2)$.

9.12 Kolmogorov's strong law of large numbers: For every sequence $(f_n)_{n\geq 1}$ 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** $E(f_1) = \lim\limits_{m\to\infty} \frac{1}{m}F_m$.

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

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 $(X; \mathcal{A})$. 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** ([9, 3.9]) the **σ-additivity** $\mu\left(\mathring{\bigcup}_{n\in\mathbb{N}} A_n\right) = \sum_{n\in\mathbb{N}}\mu(A_n) < \infty$ resp. the **interchangeability of the union** entail the **absolute convergence** of the series.

So its **total variation** $|\mu| : \mathcal{A} :\to \mathbb{R}$ with $|\mu|(A) := \sup\left\{\sum_{n\in\mathbb{N}}|\mu(A_n)| : (A_n)_{n\in\mathbb{N}} \subset \mathcal{A} : \mathring{\bigcup}_{n\in\mathbb{N}} A_n = A\right\}$ 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 $(A_{mn})_{n\in\mathbb{N}} \subset \mathcal{A}$ with $|\mu|(A_m) - \epsilon\cdot 2^{-m-1} < \sum_{n\in\mathbb{N}}|\mu(A_{mn})| \leq |\mu|(A_m)$ such that $\sum_{m\in\mathbb{N}}|\mu|(A_m) - \epsilon < \sum_{m,n\in\mathbb{N}}|\mu(A_{mn})| \leq \sum_{m\in\mathbb{N}}|\mu|(A_m)$ and hence $\sum_{m\in\mathbb{N}}|\mu|(A_m) \leq |\mu|\left(\mathring{\bigcup}_{m\in\mathbb{N}} A_m\right)$. On the other hand for **every** partition $(B_n)_{n\in\mathbb{N}} \subset \mathcal{A}$ with $\mathring{\bigcup}_{n\in\mathbb{N}} B_n = \mathring{\bigcup}_{m\in\mathbb{N}} A_m$ the intersections $(B_n \cap A_m)_{n\in\mathbb{N}}$ partition A_m while the intersections $(B_n \cap A_m)_{m\in\mathbb{N}}$ partition B_n such that due to the σ-additivity of μ holds $\sum_{n\in\mathbb{N}}|\mu(B_n)| \leq \sum_{m,n\in\mathbb{N}}|\mu(A_m \cap B_n)| \leq \sum_{m\in\mathbb{N}}|\mu|(A_m)$. This estimate carries over to the suprema such that $|\mu|\left(\mathring{\bigcup}_{m\in\mathbb{N}} A_m\right) \leq \sum_{m\in\mathbb{N}}|\mu|(A_m)$. 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 $\left|\sum_{k\in S} z_k\right| \geq \frac{1}{\pi}\sum\limits_{i=1}^{n}|z_i|$.

Proof: For $z_i = |z_i| \cdot e^{i\alpha_i}$ and $-\pi \leq \vartheta \leq \pi$ let $S(\vartheta) := \{1 \leq k \leq n : \cos(\alpha_k - \vartheta) > 0\}$. Then for every such ϑ we have $\left|\sum_{k\in S} z_k\right| = \left|\sum_{k\in S} e^{-i\vartheta} \cdot z_k\right| \geq \mathrm{Re}\left(\sum_{k\in S} e^{-i\vartheta} \cdot z_k\right) = \sum_{k\in S} |z_k| \cdot \cos(\alpha_k - \vartheta) \geq \sum_{i=1}^{n} |z_i| \cdot \cos^+(\alpha_k - \vartheta)$ and the maximal value of the sum on the right hand side attained for say $\vartheta = \vartheta_0$ is not less than the average $\frac{1}{2\pi}\int_{-\pi}^{\pi}\left(\sum_{i=1}^{n} |z_i| \cdot \cos^+(\alpha_k - \vartheta)\right) d\vartheta = \frac{1}{\pi}\sum_{i=1}^{n} |z_i|$ 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 $\frac{1}{\pi}\sum_{i=1}^{n} |\mu(A_i)| > |\mu(X)| + 1$. 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 $\sum_{k=0}^{m} \mu_N(A_k) = \mu_N\left(\bigcup_{k=0}^{m} A_k\right)$ we have $\left|\sum_{k=0}^{m} \mu(A_k) - \mu\left(\bigcup_{k=0}^{m} A_k\right)\right| = \left|\sum_{k=0}^{m} \mu(A_k) - \sum_{k=0}^{m} \mu_N(A_k) + \mu_N\left(\bigcup_{k=0}^{m} A_k\right) - \mu\left(\bigcup_{k=0}^{m} A_k\right)\right| \leq \epsilon 2^{-m+1} + \left|\mu_N\left(\bigcup_{k=0}^{m} A_k\right) - \mu\left(\bigcup_{k=0}^{m} A_k\right)\right| \leq \epsilon 2^{-m+2}$ for a suitably large N. Since ϵ and m are arbitrary we have shown the σ-additivity $\sum_{k=0}^{\infty} \mu(A_k) = \mu\left(\bigcup_{k=0}^{\infty} A_k\right)$, 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 $B_n = \bigcup_{k=0}^{K_n} A_k \in \mathcal{A}$ 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 $\lim_{n\to\infty} \|\mu - \mu_n\| = 0$

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 $|\mu|\left(\bigcap_{m\in\mathbb{N}} B_m\right) = \lim_{m\to\infty} |\mu|(B_m) \geq \inf_{m\in\mathbb{N}} |\mu|(A_m) \geq \epsilon$ 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** $\lambda : \mathcal{A}(X) \to [0;\infty]$ and a **complex measure** $\mu : \mathcal{A}(X) \to \mathbb{C}$ on a common σ-algebra $\mathcal{A}(X)$ 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 5.7.3 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 elementary 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 [8, 20.11] it is also **bounded** resp. **uniformly continuous** and hence a member of the **dual space** $\left(L^2(\varphi)\right)^*$. Due to [7, 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 $\int_A \left(1 - g^{m+1}\right) d|\mu| = \int_A w \cdot \sum\limits_{n=1}^{m} g^n \, d\lambda$ and taking recourse to **monotone convergence** 5.5 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 $|\mu|(A) \leq \sum_{n\in\mathbb{N}} |\mu(A_n)| = \sum_{n\in\mathbb{N}} \left|\int_{A_n} \frac{d\mu}{d|\mu|} \, d|\mu|\right| \leq \sum_{n\in\mathbb{N}} r \cdot |\mu|(A_n) = r \cdot |\mu|(A)$, i.e. for $r < 1$ we have $|\mu|(A) = 0$ resp. μ-a.e. $\left|\frac{d\mu}{d|\mu|}\right| \geq 1$. On the other hand for every $A \in \mathcal{A}$ with $|\mu|(A) > 0$ holds $\left|\frac{1}{|\mu|(A)} \int_A \frac{d\mu}{d|\mu|} \, d|\mu|\right| = \frac{|\mu(A)|}{|\mu|(A)} \leq 1$ so that we can apply the **mean value property** 5.15 with $S = \overline{B}_1(0)$ to obtain μ-a.e. $\left|\frac{d\mu}{d|\mu|}\right| \leq 1$. 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 $|\mu| \geq 0$ and $\lambda \geq 0$ follows λ-a.e. $\overline{\frac{d\mu}{d|\mu|}} \frac{d\mu}{d\lambda} \geq 0$ 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** 5.5 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 $\mu^+(A) = \mu\left(A \cap M^+\right)$ resp. $\mu^-(A) = \mu\left(A \cap M^-\right)$ 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 $\mu^+(A) = \frac{1}{2}\left(|\mu|(A) + \mu(A)\right) = \int_A \frac{1}{2}\left(1 + \frac{d\mu}{d|\mu|}\right) d|\mu| = \mu\left(A \cap M^+\right)$ resp. $\mu^-(A) = \mu\left(A \cap M^-\right)$.

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 5.7.3.

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 $B_n = \mathring{\bigcup}_{0\leq k\leq n} A_k$ 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 $Mf = \int f \frac{d\mu}{d\lambda}\, d\lambda$ for **elementary functions** $f \in \mathcal{E}(X)$. According to 6.9 the elementary functions $\mathcal{E}(X)$ are dense in $L^p(\lambda)$ for every $1 \leq p \leq \infty$. 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 $(f_n)_{n\in\mathbb{N}} \subset \mathcal{E}(X)$ converging also in the p-th mean on account of $\|f\|_p \leq \|f\|_\infty \cdot (\lambda(X))^{\frac{1}{p}}$ whence follows the convergence of $(Mf_n)_{n\in\mathbb{N}}$. On the right hand side the uniform convergence directly entails the convergence of the intgeral on due to $\left|\int f \frac{d\mu}{d\lambda}\, d\lambda\right| \leq \|f\|_\infty \cdot \|g\|_1$. 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 $E_n = \{|g| \geq n\}$ 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 $\int_{E_n} |g|^q\, d\lambda = \int fg\, d\lambda = \Lambda(f) \leq \|\Lambda\|^* \cdot \|f\|_p = \|\Lambda\|^* \cdot \left(\int_{E_n} |g|^q\, d\lambda\right)^{\frac{1}{p}} \Leftrightarrow \left(\int_{E_n} |g|^q\, d\lambda\right)^{1-\frac{1}{p}} \leq \|\Lambda\|^* \Leftrightarrow$ $\int_{E_n} |g|^q\, d\lambda \leq \|\Lambda\|^{*q}$ such that with **monotone convergence** 5.5 we obtain $\|g\|_q \leq \|\Lambda\|^* < \infty$ 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 $\|M\|^* \leq \left\|\frac{d\mu}{d\lambda}\right\|_q$ and hence the second assertion $\|M\|^* = \left\|\frac{d\mu}{d\lambda}\right\|_q$.

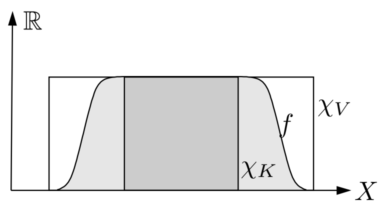
In the case of $\lambda(X) = \infty$ as in the proof for 10.8 we define $w = \sum_{n\in\mathbb{N}} \frac{\chi_{A_n}}{2^n \cdot (1+\lambda(A_n))} : X \to]0;1[$ for a countable cover $(A_n)_{n\in\mathbb{N}} \subset \mathcal{A}$ of X with $\lambda(A_n) < \infty \forall n \in \mathbb{N}$ such that the measure 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 [7, 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 $M \in \left(L^2(\varphi)\right)^*$ 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 possibilies: Our aim is to prove the **Riesz representation theorem** which states that the **Banach space** $(C_c(X,\mathbb{C}))^*$ under the **norm** $\|\|^*$ with $\|\Lambda\|^* = \sup\left\{\left|\Lambda\left(\frac{f}{\|f\|}\right)\right| : f \in C_c(X,\mathbb{C})\right\} = \sup\left\{\left|\Lambda\left(\frac{f}{\|f\|}\right)\right| : f \in C_c(X,\mathbb{R})\right\}$ (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** [8, 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** [8, 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 6.9 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 $\lambda' = \lambda|_{X \setminus A_1}$. Due 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 5.9 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** [8, 8.5] to find $\mathrm{Re}g^*$resp. $\mathrm{Im}g^* \in C\left(\overline{V}, \mathbb{R}\right)$ coinciding with $\mathrm{Re}f$ resp. $\mathrm{Im}f$ 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** [8, 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 $\mathrm{Im}f$ 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 $\mathrm{Re}\Lambda^+; \mathrm{Re}\Lambda^-; \mathrm{Im}\Lambda^+; \mathrm{Im}\Lambda^- \in (C_c(X, \mathbb{R}))^*$ such that $\Lambda f = \mathrm{Re}\Lambda^+ f - \mathrm{Re}\Lambda^- f + i(\mathrm{Im}\Lambda^+ f + \mathrm{Im}\Lambda^- f)$.

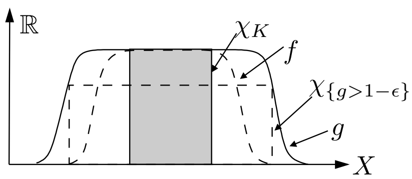
11.7 Lemma: For every **positive** functional $\Lambda : C_c(X, \mathbb{C}) \to \mathbb{C}$ the set function $\mu : P(X) \to [0; \infty]$ defined by $\mu(V) = \sup\{\Lambda g : g \prec V\}$ for **open** V and $\mu(A) = \inf\{\mu(V) : A \subset V\text{open}\}$ is an **outer measure** according to 3.2 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** [8, 9.12] and the **continuity** of Λ we have

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

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 $(A_n)_{n\in\mathbb{N}}$of arbitrary subsets with $A = \bigcup_{n\in\mathbb{N}} A_n$, open sets V_n with $A_n \subset V_n$ and $\mu(A_n) \leq \mu(V_n) + \epsilon 2^{-n}$ 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 $\Lambda g \leq \mu\left(\bigcup_{k\leq n} V_n\right) \leq \sum_{k\leq n} \mu(V_n)$ due to the subadditivity inductively extended to finite unions. Again we use the validity of this estimate for every $g \prec V$ to infer $\mu(A) \leq \mu(V) \leq \sum_{n\in\mathbb{N}} \mu(A_n) + \epsilon$ thus proving the main assertion.

Concerning the **additional regularity property** we only have to show the left inequality: For any $\epsilon > 0$ we have $K \subset \{g > 1-\epsilon\}$ and hence a $f \in C_c(X)$ with on the one hand $K \prec f \prec \{g > 1-\epsilon\}$ such that $(1-\epsilon) f \leq g$, i.e. $(1-\epsilon)\Lambda f \leq \Lambda g$ and on the other hand $\Lambda f \geq \mu(\{g > 1-\epsilon\}) - \epsilon \geq \mu(K) - \epsilon$ whence $(\mu(K) - \epsilon)(1-\epsilon) \leq \Lambda g$ 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** $\mathcal{A}(X)$ of all sets A with **finite measure** and $\mu(A) = \sup\{\mu(K) : A \supset K \text{ compact}\}$. Furthermore $\mathcal{A}(X)$contains all open sets.

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

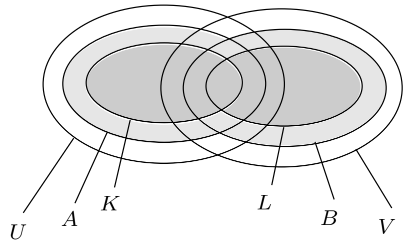
Step I. Every **compact** set K has a **finite measure** and hence belongs to $\mathcal{A}$: There is an open $V \subset K$ with compact closure $\overline{V}$ such that the separation property of locally compact spaces ensures the existence of $f, g \in C_c(X)$ with $K \prec f \prec V$ resp. $\overline{V}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

Step II. $\mathcal{A}$ contains every **open** set V: In the case of $\mu(V) = 0$ the definition of μ immediately yields $\mu(K) = \inf\{\mu(V) : K \subset V\text{open}\} = 0$ for every compact $K \subset V$. Hence we can assume $\mu(V) > 0$ and for every $\epsilon > 0$ the existence of an $f \prec V$ with $\mu(V) - \epsilon < \Lambda f < \mu(V)$ and compact support $K = \overline{\{f > 0\}}$. For every open $W \supset K$ we have $f \prec W$ and hence $\Lambda f \leq \mu(K)$ and consequently $\mu(V) - \epsilon < \Lambda f \leq \mu(K) < \mu(V) < \infty$ 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** [8, 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 $\mu(W) < \mu(K \cup L) + \epsilon$ as well as $f \prec U \cap W$ resp. $g \prec V \cap W$ with $\Lambda f > \mu(U \cap W) - \epsilon$ resp. $\Lambda g > \mu(V \cap W) - \epsilon$. We then have $\mu(K) + \mu(L) \leq \mu(W \cap U) + \mu(W \cap V) \leq \Lambda f + \Lambda g + 2\epsilon = \Lambda(f+g) + 2\epsilon \leq \mu(W) + 2\epsilon \leq \mu(K \cup L) + 3\epsilon$. 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 $(A_n)_{n\in\mathbb{N}} \subset \mathcal{A}$ with $A = \bigcup_{n\in\mathbb{N}} A_n$ there are compact $K_n \subset A_n$ with $\mu(A_n) \leq \mu(K_n) + \epsilon 2^{-n}$ whence $\sum_{k=1}^{n} \mu(A_k) \leq \sum_{k=1}^{n} \mu(K_k) + \epsilon = \mu\left(\bigcup_{k=1}^{n} K_k\right) + \epsilon \leq \mu(A) + \epsilon$. Since this estimate remains valid for $n \to \infty$ and $\epsilon \to 0$ we obtain $\sum_{n\in\mathbb{N}} \mu(A_n) \leq \mu(A)$ 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 $(A_n)_{n\in\mathbb{N}} \subset \mathcal{A}$ 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 we the algebra $\mathcal{A}$ is a **σ-algebra**. This will me made use of in the subsequent lemma to connstruct 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 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. [8, 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 $\mathrm{Re}\Lambda f = \Lambda \mathrm{Re} f$ resp. $\mathrm{Im}\Lambda f = \Lambda \mathrm{Im} f$ 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 **elementary 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 $e = \sum\limits_{k=-n}^{n} k\epsilon\chi_{A_k} \in \mathcal{E}(X)$ such that $e \leq f \leq e+\epsilon$ whence $\int ed\mu \leq \int fd\mu \leq \int ed\mu + \epsilon\cdot\mu(K)$ according to 4.9 and 5.4. 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 [8, 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. [8, 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 $\Lambda = \text{Re}\Lambda^{+} - \text{Re}\Lambda^{-} + i\left(\text{Im}\Lambda^{+} + \text{Im}\Lambda^{-}\right)$ each **positive bounded functional** of the decomposition according to 11.6 is represented by an integral, e.g. $\text{Re}\Lambda^{+} f = \int f d\left(\text{Re}\mu^{+}\right)$ for every $f \in C_c\left(X, \mathbb{C}\right)$ resp. a **complete** and **σ-regular positive Borel** measure $\text{Re}\mu^{+}$ etc. due to the preceding version 11.10 of the **Riesz representation** theorem such that $\mu = \text{Re}\mu^{+} - \text{Re}\mu^{-} + i\left(\text{Im}\mu^{+} - \text{Im}\mu^{-}\right)$ 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 $\lambda\left(K\right) \leq \int f d\lambda \overset{11.9}{=} \Gamma f \overset{11.1}{\leq} \left\|\Gamma\right\|^{*} \cdot \left\|f\right\| = \left\|\Gamma\right\|^{*}$ and on account of the **regularity condition** follows $\left\|\lambda\right\| = \mu\left(X\right) = \sup\left\{\lambda\left(K\right) : K \text{ compact}\right\} \leq \left\|\Gamma\right\|^{*}$. 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 $\left\|\Lambda\right\|^{*} = \sup\left\{\left|\frac{\int f h d|\mu|}{\sup|f|}\right| : f \in C_c\left(X, \mathbb{R}\right)\right\} = \sup\left\{\left|\int f h d|\mu|\right| : f \in C_c\left(X, \mathbb{R}^{+}\right), \sup f = 1\right\} \leq |\mu|\left(X\right) = \left\|\mu\right\|$. 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 $\left\|g\right\| \leq 1$ such that $\left\|\Lambda\right\|^{*} \geq \int_X g h d|\mu| \geq |\mu|\left(X \setminus \left\{\overline{h} \neq g\right\}\right) - |\mu|\left(\overline{h} \neq g\right) \geq |\mu|\left(X\right) - 2\epsilon$, hence $\left\|\Lambda\right\|^{*} \geq \left\|\mu\right\|$.

11.12Theorem: If every **open** set is **σ-compact** (cf. [8, 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 $\Lambda$with $\Lambda f = \int f d\lambda = \int f d\mu$ for $f \in C_c\left(X; \mathbb{R}\right)$ conicides 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. [8, 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 $x \in X$. According to the hypothesis an the **monotone convergence theorem** 5.5 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** 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** 8.9 as well as the **transformation formula** 8.11 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**

$(D_+f)(x) = \underline{\lim}_{h\downarrow 0}\frac{f(x+h)-f(x)}{h}$ and the **upper right derivate** $(D^+f)(x) = \overline{\lim}_{h\downarrow 0}\frac{f(x+h)-f(x)}{h}$ resp. **lower left derivate** $(D_-f)(x) = \underline{\lim}_{h\uparrow 0}\frac{f(x+h)-f(x)}{h}$ and the **upper left derivate** $(D^-f)(x) = \overline{\lim}_{h\uparrow 0}\frac{f(x+h)-f(x)}{h}$. These definitions contain the usual notations for e.g. the **lower right limit** $\underline{\lim}_{h\downarrow x}\varphi(h) = \sup_{\delta>0}\inf_{x<h<x+\delta}\varphi(h)$ and the **upper right limit** $\overline{\lim}_{h\downarrow x}\varphi(h) = \inf_{\delta>0}\sup_{x<h<x+\delta}\varphi(h)$ resp. the **lower left limit** $\underline{\lim}_{h\uparrow x}\varphi(h) = \sup_{\delta>0}\inf_{x-\delta<h<x}\varphi(h)$ and the **upper left limit** $\overline{\lim}_{h\uparrow x}\varphi(h) = \inf_{\delta>0}\sup_{x-\delta<h<x}\varphi(h)$. Obviously we have $(D_+f)(x) \leq (D^+f)(x)$ and $(D_-f)(x) \leq (D^-f)(x)$. 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 need 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 $\frac{df}{dy}(x) = \lim_{|h|\to 0}\frac{f(x+h)-f(x)}{h}$ 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 contability 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 $\lambda\left(A\setminus\bigcup_{n=0}^{m} I_n\right) = 0$.

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 $S_n = \bigcup_{k=0}^{n} I_k$ 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 $\lambda\left(\bigcup_{n=0}^{\infty} J_n\right) \leq \sum_{n=0}^{\infty} \lambda(J_n) = 5\sum_{n=0}^{\infty} \lambda(I_n) = 5\lambda(S) < 5\lambda(V) < \infty$ such that $\lim_{k\to\infty} \lambda\left(\bigcup_{n=k}^{\infty} J_n\right) = 0$ due to the **σ-additivity** resp. **continuity** of λ. Hence it suffices to show that $A \setminus S \subset \bigcup_{n=k}^{\infty} J_n$ for every $k \in \mathbb{N}$. To this end let $x \in A \setminus S$ and $k \in \mathbb{N}$ such that $x \in A \setminus S_k \subset U_k$ and so there must exist an $I \in \mathcal{V}_0$ with $x \in I \in U_k$. On account of $\delta_n < 2\lambda(I_{n+1})$ and $\lim_{n\to\infty} \lambda(I_n) = 0$ 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 $x \in I \subset J_l \subset \bigcup_{n=l}^{\infty} J_n$. 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 $\sum_{n=p+1}^{\infty} \lambda(I_n) < \epsilon$. Then we have $A \setminus S_p \subset A \setminus S_p \cup \bigcup_{n=p+1}^{\infty} I_n$ and consequently $\lambda(A \setminus S_p) \leq 0 + \lambda\left(\bigcup_{n=p+1}^{\infty} I_n\right) < \epsilon$.

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 $A_{u;v}$ and due to the preceding theorem there is a **finite, pairwise disjoint subfamily** $([x_i; x_i + h_i])_{i=0}^{m}$ such that $\bigcup_{i=0}^{m} [x_i; x_i + h_i] \subset U$ (3), $\lambda\left(A_{u;v} \setminus \bigcup_{i=0}^{m} [x_i; x_i + h_i]\right) < \epsilon$ and $\sum_{i=0}^{m} (f(x_i + h_i) - f(x_i)) \overset{(2)}{<} u \sum_{i=0}^{m} h_i \overset{(3)}{\leq} u \cdot \lambda(U) \overset{(1)}{<} u(\alpha + \epsilon)$ (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 $V = \bigcup_{i=0}^{m}]x_i; x_i + h_i[$. 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, pairwise disjoint subfamily** $([y_j; y_j + k_j])_{j=0}^{n}$ with $\lambda\left((A_{u;v} \cap V) \setminus \bigcup_{j=0}^{n} [y_j; y_j + k_j]\right) < \epsilon$ and hence $\alpha \leq \lambda(A_{u;v} \setminus V) + \lambda(A_{u;v} \cap V) < \epsilon + \left(\epsilon + \sum_{j=0}^{n} k_j\right)$. Substituting this inequality in (5) yields $v(\alpha - 2\epsilon) < v \sum_{j=0}^{n} k_j < \sum_{j=0}^{n} (f(y_j + k_j) - f(y_j))$ (6). Since $\bigcup_{j=0}^{n} [y_j; y_j + k_j] \subset \bigcup_{i=0}^{m} [x_i; x_i + h_i]$ and f is **nondecreasing** we also have $\sum_{j=0}^{n} (f(y_j + k_j) - f(y_j)) \leq \sum_{i=0}^{m} (f(x_i + h_i) - f(x_i))$. 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 $\lambda(F \setminus \bigcup_{k\in\mathbb{N}}[x_k;x_k+h_k]) = 0$. 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 $\lim_{n\to\infty} \frac{ds_n}{dy}(x) = \sum_{n\in\mathbb{N}} \frac{df_n}{dy}(x) < \infty$ λ-a.e. and it remains to show that $\lim_{n\to\infty} \frac{ds_n}{dy}(x) = \frac{ds}{dy}(x)$ 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 $\lim_{k\to\infty} \frac{ds_{n_k}}{dy}(x) = \frac{ds}{dy}(x)$. 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 $\lim_{n\to\infty} \frac{ds_n}{dy}(x) = \frac{ds}{dy}(x)$.

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 $\mathcal{B} = \sigma(\mathcal{I}) = \sigma(\mathcal{F})$ induced by the **right-open intervals** $\mathcal{I} = \{[a;b[: a \le b \in \mathbb{R}\}$ resp. the **ring** $\mathcal{F} = \left\{\mathring{\bigcup}_{0\le k\le m} I_k : I_k \in \mathcal{I}, m \in \mathbb{N}\right\}$ of the **one-dimensional figures** by $\lambda_f([a;b[) = f(b) - f(a)$. 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 $\lambda_f([a;b[) = f(b) - f(a)$ between the **lower semicontinuous** and **nondecreasing** functions $f : \mathbb{R} \to \mathbb{R}$ **vanishing at** 0 and the **σ-finite** and **regular Borel measures** $\lambda_f : \mathcal{B}(\mathbb{R}) \to [0;\infty]$ is **bijective** with the **inversion** $\lambda \mapsto f_\lambda$ defined by

$f_\lambda(t) = \begin{cases} \lambda([0;t[) & : t \geq 0 \\ -\lambda([t;0[) & : t < 0 \end{cases}$. Furthermore **every σ-finite** Borel measure $\lambda : \mathcal{B}(\mathbb{R}) \to [0;\infty]$ 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** $I_{\lambda_f} \in (C_c(\mathbb{R};\mathbb{R}))^*_+$ with $I_{\lambda_f} g = \int g d\lambda_f$ according to the **Riesz representation theorem** 11.10 resp. theorem 11.12 on the **σ-algebra** $\mathcal{L}(\mathbb{R}) \supset \mathcal{B}$ 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 $f_\lambda(t) - f_\lambda(s) = \lambda([0;t[) - \lambda([0;s[) = \lambda([0;t[\backslash[0;s[) = \lambda([s;t[) > 0$ and similarly for the other cases $s < 0 < t$ resp. $s < t \leq 0$. Also f_λ is **lower semicontinuous** since for every sequence $(\epsilon_n)_{n\in\mathbb{N}} \subset]0;1[$ with $\lim\limits_{n\to\infty} \epsilon_n = 0$ and $t \geq 0$ we have $\lim\limits_{n\to\infty} f_\lambda(t-\epsilon_n) = \lim\limits_{n\to\infty} \lambda([0;t-\epsilon_n[) \overset{2.2.2}{=} \lambda(\bigcup_{n\in\mathbb{N}}[0;t-\epsilon_n[) = \lambda([0;t[) = f_\lambda(t)$ 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 $g_{\lambda_f}(t) = \lambda_f([0;t[) = f(t) - f(0) = f(t)$ 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 $\mu_{f_\lambda}([s;t[) = f_\lambda(t) - f_\lambda(s) = \lambda([s;t[)$ for $0 \leq s < t$ as shown above whence follows the identity of the two measures on $\mathcal{B}(\mathbb{R}) = \sigma(\mathcal{I})$ 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 $\mathcal{L}(\mathbb{R})$ according to 11.8. Note that the **completeness** is achieved on **non Borel measurable** null sets in $\mathcal{L}(\mathbb{R})\backslash\mathcal{B}(\mathbb{R})$ 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 $\mathcal{L}(\mathbb{R}^n)$ 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** $\mu : \mathcal{B}(\mathbb{R}^n) \to \mathbb{C}$ defnied by $\frac{\tilde{d\mu}}{d\lambda}(\mathbf{x}) = \lim\limits_{r\to 0} \frac{\mu(B_r(\mathbf{x}))}{\lambda(B_r(\mathbf{x}))}$ for $\mathbf{x} \in \mathbb{R}^n$ and the associated **Hardy-Littlewood maximal function** $M\mu(\mathbf{x}) = \sup\limits_{0<r<\infty} \frac{|\mu|(B_r(\mathbf{x}))}{\lambda(B_r(\mathbf{x}))}$ using the **total variation** $|\mu|$ of μ. 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 $|\mu|(B_r(\mathbf{x})) > (\alpha+\epsilon)\lambda(B_r(\mathbf{x}))$ 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 $\lambda(M\mu > \alpha) \leq 3^n \frac{\|\mu\|}{\alpha}$ 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 $1 \leq i \leq N$. 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 $\lambda(K) \leq 3^n \cdot \sum_{j\in S}\lambda(B_j) \leq \frac{3^n}{\alpha}\sum_{j\in S}|\mu|(B_i) \leq 3^n\frac{\|\mu\|}{\alpha}$.

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 $d\mu = f d\lambda$, 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 $\int_A f(\mathbf{y})\,d\mathbf{y} = \int_A f d\lambda$ with $\mathbf{y} \in \mathbb{R}^n$ according to 8.9 to denote the **Lebesgue intgeral**. 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 integrable $f \in L^1(\lambda)$ **almost every** $\mathbf{x} \in \mathbb{R}^n$ is a **Lebesgue point**.

Proof: We show that $(Tf)(\mathbf{x}) = \limsup_{r\to 0}(T_r f)(\mathbf{x}) = 0$ 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 $(T_r h)(\mathbf{x}) \leq \frac{1}{\lambda(B_r)}\int_{B_r(\mathbf{x})}|h(\mathbf{y})|\,d\mathbf{y} + |h(\mathbf{x})|$ for $h = f - g$ we have $Tf \leq Th + Tg = Th \leq Mh + |h|$ and hence $\{Tf > 2\epsilon\} \subset \{Mh > \epsilon\} \cup \{|h| > \epsilon\}$ for every $\epsilon > 0$. On the one hand we have $\lambda(|h| > \epsilon) \leq \frac{\|h\|_1}{\epsilon} \leq \frac{1}{n\epsilon}$ and on the other hand lemma 13.3 yields $\lambda(Mh > \epsilon) \leq 3^n\frac{\|h\|_1}{\epsilon} \leq \frac{3^n}{n\epsilon}$ such that we arrive at $\lambda(\{Mh > \epsilon\} \cup \{|h| > \epsilon\}) \leq \frac{3^n+1}{n\epsilon}$ 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** 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 $\frac{\tilde{d\mu}}{d\lambda} = \begin{cases}\frac{d\mu}{d\lambda} & \text{if } \mu \lessdot \lambda \\ 0 & \text{if } \mu \perp \lambda\end{cases}$ 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: $\mu \lessdot \lambda$. According to 10.8.2 at any Lebesgue point x of $\frac{d\mu}{d\lambda}$ we have $\frac{d\mu}{d\lambda}(\mathbf{x}) = \lim_{r\to 0}\frac{1}{\lambda(B_r)}\int_{B_r(\mathbf{x})}\frac{d\mu}{d\lambda}d\lambda$ $= \lim_{r\to 0}\frac{\mu(B_r(\mathbf{x}))}{\lambda(B_r)} = \frac{\tilde{d\mu}}{d\lambda}(\mathbf{x})$. Hence the assertion follows from 13.5.

Case II: $\mu \perp \lambda$. Like the **Hardy-Littlewood maximal function** $M\mu(\mathbf{x}) = \sup_{0<r<\infty}\frac{|\mu|(B_r(\mathbf{x}))}{\lambda(B_r(\mathbf{x}))}$ from 13.2 the **upper derivative** $D_u\mu(\mathbf{x}) = \inf_{n\to\infty}\sup_{0<r<1/n}\frac{|\mu|(B_r(\mathbf{x}))}{\lambda(B_r(\mathbf{x}))}$ is Borel measurable since $\sup_{0<r<1/n}\frac{|\mu|(B_r(\mathbf{x}))}{\lambda(B_r(\mathbf{x}))}$ 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 $\mu(K) \geq \|\mu\| - \epsilon$. Then we have $\left\|\mu_{X\setminus K}\right\| < \epsilon$ and for every $x \in X \setminus K$ holds $D_u\mu(\mathbf{x}) = D_u\mu_{X\setminus K}(\mathbf{x}) \leq M\mu_{X\setminus K}(\mathbf{x})$ 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 $\lambda(D_u\mu > \alpha) \leq 0 + 3^n\frac{\|\mu_{X\setminus K}\|}{\alpha} < 3^n\frac{\epsilon}{\alpha}$. Since for a given $\alpha > 0$ for every $\epsilon > 0$ we can find a K such that this inequality holds we infer $\lambda(D_u\mu > \alpha) = 0$ 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 $\sum_{i=1}^{n}(\beta_i - \alpha_i) < \delta$ we have $\sum_{i=1}^{n}|f(\beta_i) - f(\alpha_i)| < \epsilon$. 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 $V_{\alpha_i}^{\beta_i}f = \sup\left\{\sum_{k=1}^{n}|f(x_k) - f(x_{k-1})| : \alpha_i = x_0 < ... < x_n = \beta_i; n \geq 1\right\}$ for segments $(]\alpha_i;\beta_i[)_{1\leq i\leq n}$ chosen as above we have $\sum_{i=1}^{n}|V_a f(\beta_i) - V_a f(\alpha_i)| = \sum_{i=1}^{n}V_{\alpha_i}^{\beta_i}f < \epsilon$.

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 $S = \bigcup_{i=1}^{n}]\alpha_i;\beta_i[$ has measure $\lambda(S) = \sum_{i=1}^{n}\lambda(]\alpha_i;\beta_i[) < \delta$ satisfies $\lambda(f[S]) = \sum_{i=1}^{n}\lambda(f]\alpha_i;\beta_i[) = \sum_{i=1}^{n}(f(\beta_i) - f(\alpha_i)) < \epsilon$. 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 $V = \mathring{\bigcup}_{i\leq i<\infty}]\alpha_i;\beta_i[$. 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 $\left|D_+^+ f(x) - g(x)\right| = \lim_{h\downarrow 0} \frac{1}{h}\int_x^{x+h} |g(y)-g(x)|\,dy \le \lim_{h\downarrow 0} \frac{2}{2h}\int_{x-h}^{x+h} |g(y)-g(x)|\,dy = \lim_{h\downarrow 0} \frac{2}{\lambda(B_h)}\int_{B_h(x)} |g(y)-g(x)|\,dy = 0$ and likewise for $D_-^- f(x)$.

13.10 The Fundamental Theorem of Calculus: Every **absolutely continuous** function $f : [a;b] \to \mathbb{C}$ is λ-a.e. **differentiable** on $I = [a;b]$ with $\frac{df}{dy} \in L^1(\lambda)$ and

$$f(x) - f(a) = \int_a^x \frac{df}{dy} dy$$

Proof:

We apply the **Jordan decomposition** 12.6 to split $f = V_a\mathrm{Re}f - (V_a\mathrm{Re}f - \mathrm{Re}f) + iV_a\mathrm{Im}f - i(V_a\mathrm{Im}f - \mathrm{Im}f)$ into four **nondecreasing** components. Due to 13.7 the **absolute continuity** transfers from f to each of the four components. 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. [10, 9.2.1] the image $g[F] = g\left[\bigcup_{i\in\mathbb{N}} K_i\right] = \bigcup_{i\in\mathbb{N}} g[K_i]$ 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 $g\left[\bigcup_{i\in\mathbb{N}} A_i\right] = \bigcup_{i\in\mathbb{N}} g[A_i]$ 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 $A \in \mathcal{L}([a;b])$. Particularly 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.

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** [8, 22.11].

14.2 Definition: A function $\mathbf{f} : V \to \mathbb{R}^n$ for some **open** $V \subset \mathbb{R}^n$ is **differentiable** at $\mathbf{x} \in V$, i.e. there exists a **linear** $\frac{d\mathbf{f}}{d\mathbf{y}}(\mathbf{x}) : \mathbb{R}^n \to \mathbb{R}^n$ 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.

14.3 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 $\lim_{r\to 0} \frac{\lambda^n(\mathbf{g}[B_r(\mathbf{x})])}{\lambda^n(B_r(\mathbf{x}))} = \left|\det\left(\frac{d\mathbf{g}}{d\mathbf{y}}(\mathbf{x})\right)\right|$.

Proof:

Case I: $\frac{d\mathbf{g}}{d\mathbf{y}}(\mathbf{x}) \in GL(n;\mathbb{R})$. 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 8.11 follows $(1-\epsilon)^n\le\frac{\lambda^n(\mathbf{h}[B_r(\mathbf{0})])}{\lambda^n(B_r(\mathbf{0}))}\le(1+\epsilon)^n$. Thus we have $\lim\limits_{r\to0}\frac{\lambda^n(\mathbf{h}[B_r(\mathbf{0})])}{\lambda^n(B_r(\mathbf{0}))}=1$ and the assertion follows from 8.11 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 $\epsilon>0$ ther 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. [8, 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 $\lambda^n(\mathbf{f}[B_r(\mathbf{0})])\le\lambda^n\left(U_{\delta r}\left(\frac{d\mathbf{f}}{d\mathbf{y}}(\mathbf{0})*B_r(\mathbf{0})\right)\right)=\lambda^n\left(S*U_\delta\left(\frac{d\mathbf{f}}{d\mathbf{y}}(\mathbf{0})*B_1(\mathbf{0})\right)\right)<\epsilon r^n$ with the **dilation** S defined by $S*\mathbf{e}_i=r\mathbf{e}_i$ for $0\le i\le n$ according to 8.11. Since $r^n=\lambda^n(B_r(\mathbf{0}))$ we conclude that $\lim\limits_{r\to0}\frac{\lambda^n(\mathbf{f}[B_r(\mathbf{0})])}{\lambda^n(B_r(\mathbf{0}))}=0$.

14.4 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 $m\ge1$ let
$$P_m=\left\{\mathbf{x}=\sum_{k=1}^{n}x_k\mathbf{e}_k\in\mathbb{R}^n:x_k=\frac{z}{2^m};z\in\mathbb{Z};1\le k\le n\right\}$$
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 $m\ge1$ and an interval $I\in\Omega_m$ with $x\in I\subset V$ such that $V=\bigcup\left(\bigcup_{m\ge1}\Omega_m^V\right)$ with $\Omega_m^V=\{I\in\Omega_m:I\subset V\}$. From this collection of intervals starting with Ω_1 subsequently for $m\ge1$ 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 $\lambda^n(B)\le n^{n/2}\lambda^n(I)$ whence follos the assertion.

14.5 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 $m;p\ge1$ let $C_{m;p}=\left\{\mathbf{x}\in A:\frac{|\mathbf{g}(\mathbf{y})-\mathbf{g}(\mathbf{x})|}{|\mathbf{y}-\mathbf{x}|}\le m\,\forall\mathbf{y}\in B_{1/p}(\mathbf{x})\cap A\right\}$. 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.

14.6 Change-of-variables Theorem: For any **Borel measurable** $f : V \to [0;\infty]$ and **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 $\lambda^n(\mathbf{g}[V \setminus A]) = 0$ we have

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

with both $d\mathbf{y}$ resp $d\mathbf{g}$ denoting the **same** Lebesgue measure λ_V^n 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 open subset V is for practical reasons because coordinate transformations like the subsequent examples 14.7 may only be defined on a correspondingly smaller domain. **Step I** depends on the preceding Lemma 14.5 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 4.9 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_V^n**-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_V^n**-null set** B_0 the hypothesis yields $\lambda_V^n(\mathbf{g}[B_0 \setminus A]) \leq \lambda_V^n(\mathbf{g}[V \setminus A]) = 0$ whereas from 14.5 we can infer $\lambda_V^n(\mathbf{g}[B_0 \cap A]) = 0$ 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_V^n 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_V^n- null set B_0 and an F_σ-set B_1.

Step II. For every $B \in \mathcal{L}(V)$ we have $\lambda_V^n(\mathbf{g}(A \cap B)) = \int_{A\cap B} \left|\det\left(\frac{d\mathbf{g}}{d\mathbf{y}}\right)\right| d\mathbf{y}$: Because of **step I** for $B \in \mathcal{L}(V)$ the expression $\mu_m(B) = \lambda_V^n(\mathbf{g}(A_m \cap B))$ 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 14.5 λ_V^n**-absolutely continuous** measure on $\mathcal{L}(V)$. The **Radon-Nikodym theorem** 10.8.2 assures that $\frac{d\mu_m}{d\lambda^n} \in L^1(\lambda_V^n)$ exists λ_V^n-a.e. with $\mu_m(B) = \int_B \frac{d\mu_m}{d\lambda_V^n} d\lambda_V^n$. 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

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

according to 8.9. From 14.3 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 $\mu_m(A_m \cap B) = \int_{A_m\cap B} \frac{d\mu_m}{d\lambda_V^n} d\lambda_V^n = \int_{A_m\cap B} \left|\det\left(\frac{d\mathbf{g}}{d\mathbf{y}}\right)\right| d\mathbf{y}$ for every $m \geq 1$. Finally for $m \to \infty$ the **continuity of** λ_V^n **from below** resp. the **bijectivity** of $\mathbf{g}$ on the left hand side resp. the **monotone convergence theorem** 5.8 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 $\int_{g[A]}\chi_B d\mathbf{y} = \int_A\left(\chi_B\circ\mathbf{g}\right)\left|\det\left(\frac{d\mathbf{g}}{d\mathbf{y}}\right)\right| d\mathbf{y}$: 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 **elementary function** and by theorem 4.9 resp. another application of **monotone convergence theorem** 5.8 to every **Borel measurable** $f: V\to[0;\infty]$.

14.7 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** $\mathbf{g}^{-1}\left[C_{h;R}\right] = [0;R]\times[0;2\pi[\times[0;h]$ defined by **cylinder coordinates** $(r;\varphi;z)$ with the **bijective** and **continuous** transformation $\mathbf{g}:\mathbf{g}^{-1}\left[C_{h;R}\right]\to C_{h;R}$ 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**

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

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

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

14.9 Special cases:

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

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

2. In the case of $f=f\circ g=1$ on a **one-dimensional interval** $A=[a;b]$ we obtain the **fundamental theorem of Calculus**: $g[A]=\int_{g[A]}dg=\int_A\left|\frac{dg}{dy}\right|dy$ resp. $|g(b)-g(a)|=\int_a^b\left|\frac{dg}{dy}\right|dy$. Because of $g(b)-g(a)>0\Leftrightarrow\frac{dg}{dy}>0$ we arrive at the familiar formula

$$g(b)-g(a)=\int_a^b\frac{dg}{dy}dy$$

.

3. For arbitrary f on a **one-dimensional interval** $A=[a;b]$ we have $\int_{g[A]} f(g)\,dg=\int_A f(g(y))\cdot\left|\frac{dg}{dy}\right|dy$. As above for **strictly increasing** g we obtain $\int_{[g(a);g(b)]} f(g)\,dg=\int_{[a;b]} f(g(y))\cdot\frac{dg}{dy}dy$ whereas for **strictly decreasing** g it is $\int_{[g(b);g(a)]} f(g)\,dg=-\int_{[a;b]} f(g(y))\cdot\frac{dg}{dy}dy$ such that we may apply the **fundamental theorem of calculus** 13.10 to combine the two cases in the **integration by substitution**

$$\int_{g(a)}^{g(b)} f(g)\,dg=F(g(b))-F(g(a))=\int_a^b f(g(y))\cdot\frac{dg}{dy}dy$$

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

$$\frac{d(F\circ g)}{dy}(x)=\frac{dF}{dg}(g(x))\cdot\frac{dg}{dy}(x)$$

Index

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

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

[6] Walter Rudin. *Real and Complex Analysis.* McGraw Hill, 1987.

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

[8] Arne Vorwerg. Topology. 2020.

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

[10] Arne Vorwerg. Mengenlehre. 2022.