Analysis III:
Measure and Integration Theory
of Several Variables

Peter Philip*

Lecture Notes
Created for the Class of Winter Semester 2016/2017 at LMU Munich

February 14, 2020

Contents

1 Measure Theory ........................................ 4
  1.1 Motivation ........................................... 4
  1.2 Measure Spaces and Generators .................... 5
  1.3 Extension ............................................ 10
    1.3.1 Semirings, Rings, Algebras ................... 10
    1.3.2 Contents, Premeasures ........................ 13
    1.3.3 Lebesgue Premeasure on \( \mathbb{R}^n \) .......... 20
    1.3.4 Outer Measures, Carathéodory Extension Theorem . . . 24
    1.3.5 Dynkin Systems, Uniqueness of Extensions ........ 28
    1.3.6 Completion ................................... 33
  1.4 Inverse Image, Trace .................................. 35
  1.5 Lebesgue Measure on \( \mathbb{R}^n \) .................. 37
    1.5.1 Intervals and Countable Sets ................... 37
    1.5.2 Approximation ................................ 38
    1.5.3 Cantor Set .................................. 39
  1.6 Measurable Maps .................................... 42
    1.6.1 Definition, Composition ......................... 42
    1.6.2 Pushforward Measure ........................... 43
    1.6.3 Invariance of Lebesgue Measure under Euclidean Isometries . . 44

*E-Mail: philip@math.lmu.de
## Contents

1.6.4 Nonmeasurable Sets ........................................ 48
1.6.5 \( \mathbb{R} \)-Valued Measurable Maps ....................... 50
1.6.6 Products, Projections, and Borel Sets ...................... 59

### 2 Integration

2.1 Integration of \( \mathbb{R} \)-Valued Measurable Maps .......... 62
   2.1.1 Simple Functions ........................................ 62
   2.1.2 Nonnegative Functions .................................... 64
   2.1.3 Integrable Functions ...................................... 67
2.2 Null Sets, Properties Holding Almost Everywhere ........... 72
2.3 Convergence Theorems ......................................... 74
   2.3.1 Fatou’s Lemma and Dominated Convergence ............... 74
   2.3.2 Parameter-Dependent Integrals: Continuity, Differentiation ... 76
2.4 Riemann versus Lebesgue ..................................... 78
2.5 Products ..................................................... 84
   2.5.1 Product Measure ......................................... 84
   2.5.2 Theorems of Tonelli and Fubini ......................... 91
2.6 \( L^p \)-Spaces ................................................ 97
   2.6.1 Definition, Norm, Completeness ........................... 97
   2.6.2 Integration with Respect to Pushforward Measures ....... 104
   2.6.3 Dense Subsets, Separability ................................ 105
   2.6.4 Convolution and Fourier Transform ....................... 110
2.7 Change of Variables .......................................... 125

### 3 Integration Over Submanifolds of \( \mathbb{R}^n \)

3.1 Submanifolds of \( \mathbb{R}^n \) .................................... 130
3.2 Metric Tensor, Gram Determinant ................................ 137
3.3 Integration Over Submanifolds ................................ 138
3.4 Tangent Space and Normal Space ................................ 147
3.5 Gauss-Green Theorem and Green’s Identities ................ 150

### A Order on, Arithmetic in, and Topology of \( \mathbb{R} \)

### B Algebraic Structure on Rings of Subsets

### C Initial and Final \( \sigma \)-Algebras
CONTENTS

D Riemann Integral on Intervals in $\mathbb{R}^n$ 166

E Wallis Product 173

F Improper Riemann Integral of the Sinc Function 175

G Topological and Measure-Theoretic Supplements 177
   G.1 Transitivity of Dense Subsets .............................. 177
   G.2 Inverse Image and Trace for Semirings .................. 177
   G.3 Measurability with Respect to Completions ................ 178
   G.4 Interchanging Integrals with Uniform Limits ............... 178

H Polar Coordinates 179

References 180
1 Measure Theory

1.1 Motivation

In Analysis I, we have already measured the size of intervals $I := [a, b] \subseteq \mathbb{R}$ by defining $|I| := |a - b|$. Moreover, we defined the Riemann integral of Riemann-integrable functions $f : I \rightarrow \mathbb{R}$ on such intervals. We also mentioned that, for a Riemann-integrable $f : I \rightarrow \mathbb{R}^+_0$, the Riemann integral $\int_I f$ can be interpreted as the size of the area between the graph of $f$ and the horizontal axis. Major goals of this class include measuring the size of subsets of $\mathbb{R}^n$ and integrating functions over subsets of $\mathbb{R}^n$.

While for $n$-dimensional intervals $I = \prod_{j=1}^n I_j \subseteq \mathbb{R}^n$ with one-dimensional intervals $I_j \subseteq \mathbb{R}$, an obvious choice to assign a volume is $|I| := \prod_{j=1}^n |I_j|$, for general subsets of $\mathbb{R}^n$, it is far from obvious what a reasonable volume should be. For certain sets that have a simple relation with intervals, this relation can be used to assign a volume (e.g., triangles can be decomposed into halves of rectangles, which leads to the formula for its area that is taught in high school). Another common sense rule is that the volume $v$ of the union of $N$ disjoint sets with known volume $v_j$ should be the sum of the volumes, $v = \sum_{j=1}^N v_j$. And it makes sense to extend this rule to countable disjoint unions: For example, if

$$\forall j \in \mathbb{N} \ I_j := \left[\frac{1}{2^j}, \frac{1}{2^{j+1}} \right],$$

then the $I_j$ are clearly disjoint,

$$I := [0, 1] = \{0\} \cup \bigcup_{j=1}^{\infty} I_j,$$

and, due to the formula for the sum of the geometric series,

$$|I| = 1 - \frac{1}{2} = 1 + \sum_{j=1}^{\infty} 2^{-j} = |[0, 0]| + \sum_{j=1}^{\infty} |I_j|.$$

For example, such a rule then allows to compute the area of a circle by decomposing it into countably many triangles (however, the details are a bit cumbersome and we will not pursue this here).

Another common sense property that a volume function should satisfy is translation invariance, i.e., $v(A + x) = v(A)$ should hold for each $A \subseteq \mathbb{R}^n$ and each $x \in \mathbb{R}^n$, where $A + x := \{a + x : a \in A\}$.

**Definition 1.1.** Let $n \in \mathbb{N}$. We call a function $v : \mathcal{P}(\mathbb{R}^n) \rightarrow [0, \infty]$ an **ideal $n$-dimensional volume** if, and only if, the following conditions (i) – (iii) are satisfied:

(a) If $a, b \in \mathbb{R}^n$, $a \leq b$, $I := [a, b]$, then

$$v(I) = \prod_{j=1}^n |b_j - a_j|.$$
(b) \( v \) is translation invariant, i.e., for each \( A \in \mathcal{P}(\mathbb{R}^n) \) and each \( x \in \mathbb{R}^n \),
\[
v(A + x) = v(A), \quad A + x := \{a + x : a \in A\}.
\]

(c) \( v \) satisfies countable additivity, i.e. if \( A, A_j \in \mathcal{P}(\mathbb{R}^n), j \in \mathbb{N} \), and if \( A \) is the disjoint union of the \( A_j \), then
\[
v(A) = \sum_{j=1}^{\infty} v(A_j).
\]

Unfortunately, using the axiom of choice, we will be able to show that no ideal volume exists (see Sec. 1.6.4 and, in particular, Th. 1.74(b)). However, we will see that there exist volume functions (notably so-called Lebesgue-Borel and Lebesgue measure) that satisfy Def. 1.1(a)--(c), except that they are not defined on all of \( \mathcal{P}(\mathbb{R}^n) \) (Lebesgue-Borel measure is defined on a strictly smaller set than Lebesgue measure and is the restriction of Lebesgue measure to this set, cf. Ex. 1.47).

1.2 Measure Spaces and Generators

Now that we have discussed some desirable properties of volume functions, which we will usually call measures \( \mu \) from now on, we will do the mathematicians’ job and condense these properties into an abstract definition, then proceeding to studying some general consequences. As we have already indicated that it would be too restrictive to require measures to be defined on the entire power set \( \mathcal{P}(X) \) of a set \( X \), we start by considering subsets \( \mathcal{A} \subseteq \mathcal{P}(X) \) that are suitable as the domain of a measure. From our discussion in the previous section, it makes sense to require \( \mathcal{A} \) to be closed under countable unions. To be useful, \( \mathcal{A} \) should at least contain \( \emptyset \) and \( X \). A finite measure \( \mu \) should also have the property \( \mu(X \setminus A) = \mu(X) - \mu(A) \). We will see that this actually follows from additivity as long as \( \mathcal{A} \) is also closed under taking complements. With translation invariance in mind, one might also want \( \mathcal{A} \) to be closed under translations (i.e. one might want \( A + x \in \mathcal{A} \) for each \( A \in \mathcal{A} \) and \( x \in X \)), but, as \( X \) might not come with an addition defined on it, we will not include such a requirement for the domain of a measure. Altogether, this leads to the following definition:

**Definition 1.2.** Let \( X \) be a set. A collection of subsets \( \mathcal{A} \subseteq \mathcal{P}(X) \) is called a \( \sigma \)-algebra on \( X \) if, and only if, it satisfies the following three conditions:

(a) \( \emptyset \in \mathcal{A} \).

(b) If \( A \in \mathcal{A} \), then \( A^c = X \setminus A \in \mathcal{A} \).

(c) If \( (A_n)_{n \in \mathbb{N}} \) is a sequence of sets in \( \mathcal{A} \), then \( \bigcup_{n \in \mathbb{N}} A_n \in \mathcal{A} \).
If $\mathcal{A}$ is a $\sigma$-algebra on $X$, then the pair $(X, \mathcal{A})$ is called a measurable space. The sets $A \in \mathcal{A}$ are called $\mathcal{A}$-measurable (or merely measurable if $\mathcal{A}$ is understood).

Measures will be defined on $\sigma$-algebras (see Def. 1.10 below). Note that there are similarities between the definition of a $\sigma$-algebra and the definition of a topology: For example, both topologies and $\sigma$-algebras are always subsets of the power set of some given set of interest; they also both contain at least the empty set and the whole space. Of course, in general, a topology is not a $\sigma$-algebra and a $\sigma$-algebra is not a topology. Still, we will notice repeatedly, that similar concepts and techniques are useful for both measure spaces and topological spaces (there is a branch of mathematics called Category Theory that systematically investigates such aspects – one finds that the category of topological spaces has a lot in common with the category of measurable spaces).

**Lemma 1.3.** Let $X$ be a set and let $\mathcal{A} \subseteq \mathcal{P}(X)$ be a $\sigma$-algebra on $X$.

(a) $\mathcal{A}$ is closed under finite unions: Let $n \in \mathbb{N}$. Then

$$A_1, \ldots, A_n \in \mathcal{A} \implies \bigcup_{j=1}^{n} A_j \in \mathcal{A}.$$ 

(b) $\mathcal{A}$ is closed under countable (both finite and infinite) intersections: Let $I$ be a nonempty countable index set and $A_j \in \mathcal{A}$, $j \in I$. Then

$$\bigcap_{j \in I} A_j \in \mathcal{A}.$$ 

**Proof.** (a) follows from Def. 1.2(a),(c), as

$$\bigcup_{j=1}^{n} A_j = A_1 \cup \cdots \cup A_n \cup \emptyset \cup \emptyset \ldots$$

(b) follows from Def. 1.2(b),(c) and (a), since

$$\left(\bigcap_{j \in I} A_j \right)^c = \bigcup_{j \in I} A_j^c,$$

completing the proof of the lemma.

**Example 1.4.** Let $X$ be set.

(a) Clearly, $\{\emptyset, X\}$ and $\mathcal{P}(X)$ are $\sigma$-algebras on $X$, where $\{\emptyset, X\}$ sometimes called the trivial $\sigma$-algebra on $X$. 
(b) \( A := \{ A \subseteq X : A \text{ countable or } A^c \text{ countable} \} \) constitutes a \( \sigma \)-algebra on \( X \) (clearly, \( \emptyset \in A \); if \( A \in A \), then \( A^c \in A \); if \( A_n \in A \), \( n \in \mathbb{N} \), then \( A := \bigcup_{n \in \mathbb{N}} A_n \in A \): If at least one \( A^c_{n_0}, n_0 \in \mathbb{N} \), is countable, then \( A^c = \bigcap_{n \in \mathbb{N}} A^c_n \subseteq A^c_{n_0} \) is countable; if each \( A^c_n \) is uncountable, then each \( A_n \) must be countable, implying \( A \) to be countable). 

The next result will allow us to generate an abundance of \( \sigma \)-algebras:

**Proposition 1.5.** Arbitrary intersections of \( \sigma \)-algebras yield again \( \sigma \)-algebras: Let \( X \) be a set, let \( I \) be a nonempty index set, and let \( (A_i)_{i \in I} \) be a family of \( \sigma \)-algebras on \( X \). Then

\[
A := \bigcap_{i \in I} A_i
\]

is a again a \( \sigma \)-algebra on \( X \).

**Proof.** Since \( \emptyset \) is in each \( A_i \), it is also in \( A \). If \( A \in A \), then \( A \in A_i \) for each \( i \in I \), i.e. \( A^c \in A_i \) for each \( i \in I \), implying \( A^c \in A \). Similarly, if \( A_n \in A \) for each \( n \in \mathbb{N} \), then \( A_n \in A_i \) for each \( i \in I \) and each \( n \in \mathbb{N} \). Thus, \( A := \bigcup_{n \in \mathbb{N}} A_n \in A_i \) for each \( i \in I \), showing \( A \in A \). \( \blacksquare \)

**Caveat 1.6.** The union of (even two) \( \sigma \)-algebras is, in general, not a \( \sigma \)-algebra: For example, let \( X := \{0, 1, 2\} \), and consider

\[
A_1 := \{\emptyset, \{0\}, \{1, 2\}, X\},
A_2 := \{\emptyset, \{1\}, \{0, 2\}, X\},
B := A_1 \cup A_2.
\]

Then, clearly, \( A_1 \) and \( A_2 \) are \( \sigma \)-algebras on \( X \), but \( B \) is not (for example \( \{0\}, \{1\} \in B \), but \( \{0, 1\} \notin B \)).

**Definition and Remark 1.7.** Let \( X \) be a set. If \( \mathcal{E} \) is a collection of subsets of \( X \), i.e. \( \mathcal{E} \subseteq \mathcal{P}(X) \), then let \( \sigma(\mathcal{E}) := \sigma_X(\mathcal{E}) \) denote the intersection of all \( \sigma \)-algebras on \( X \) that are supersets of \( \mathcal{E} \) (i.e. that contain all the sets in \( \mathcal{E} \)). According to Prop. 1.5, \( \sigma(\mathcal{E}) \) is a \( \sigma \)-algebra on \( X \). Obviously, it is the smallest \( \sigma \)-algebra on \( X \) containing \( \mathcal{E} \). It is, thus, called the \( \sigma \)-algebra generated by \( \mathcal{E} \); \( \mathcal{E} \) is called a generator of \( \sigma(\mathcal{E}) \).

**Example 1.8.** Let \( X \) be a set.

(a) Both \( \{\emptyset\} \) and \( \{X\} \) are generators of the trivial \( \sigma \)-algebra on \( X \). The set \( \mathcal{E} := \{\{x\} : x \in X\} \) generates the \( \sigma \)-algebra defined in Ex. 1.4(b) above.

(b) Let \( \mathcal{T} \) be a topology on \( X \). Then \( \mathcal{B} := \sigma(\mathcal{T}) \), i.e. the \( \sigma \)-algebra generated by the open sets of \( X \), is called the Borel \( \sigma \)-algebra on \( (X, \mathcal{T}) \) (or on \( X \) if the topology \( \mathcal{T} \) is understood). The elements of \( \mathcal{B} \) are called Borel sets.
1 MEASURE THEORY

(c) Let \( n \in \mathbb{N} \). It is often useful to know that the Borel \( \sigma \)-algebra \( \mathcal{B}^n \) on \( \mathbb{R}^n \) (with respect to the norm topology on \( \mathbb{R}^n \)) is also generated by each of the following sets (in each case, it is a simple consequence of the following Lem. 1.9):

\[
\begin{align*}
C^n &= \{ A \subseteq \mathbb{R}^n : A \text{ closed} \}, \\
K^n &= \{ A \subseteq \mathbb{R}^n : A \text{ compact} \}, \\
T^c_n &= \{ [a, b] : a, b \in \mathbb{R}^n, a \leq b \} \cup \{ \emptyset \}, \\
T^c_{c, Q} &= \{ [a, b] : a, b \in \mathbb{Q}^n, a \leq b \} \cup \{ \emptyset \}, \\
T^o_n &= \{ [a, b] : a, b \in \mathbb{R}^n, a \leq b \}, \\
T^o_{c, Q} &= \{ [a, b] : a, b \in \mathbb{Q}^n, a \leq b \}, \\
T^o &= \{ \bigcup_{k=1}^n I_k : N \in \mathbb{N}, I_1, \ldots, I_N \in \mathcal{T}^n \text{ disjoint} \}, \\
T^o_{c, Q} &= \{ \bigcup_{k=1}^n I_k : N \in \mathbb{N}, I_1, \ldots, I_N \in \mathcal{T}^n \text{ disjoint} \}.
\end{align*}
\]

**Lemma 1.9.** Let \((X, \mathcal{T})\) be a topological space and let \( \mathcal{B} := \sigma(\mathcal{T}) \) be the corresponding set of Borel sets. Moreover, let \( \mathcal{C} \) be the set of closed subsets of \( X \), and let \( \mathcal{K} \) be the set of compact subsets of \( X \). Then the following holds true:

(a) \( \mathcal{B} = \sigma(\mathcal{C}) \).

(b) If \((X, \mathcal{T})\) is Hausdorff (i.e. a \( T_2 \) space) and there exists a sequence \((K_n)_{n\in\mathbb{N}}\) of compact sets such that \( X = \bigcup_{n=1}^{\infty} K_n \), then \( \mathcal{B} = \sigma(\mathcal{K}) \).

(c) If \( \mathcal{E} \subseteq \mathcal{B} \) is such that every open set \( O \in \mathcal{T} \) is a countable union of sets from \( \mathcal{E} \), then \( \mathcal{B} = \sigma(\mathcal{E}) \).

**Proof.**

(a): Since \( \mathcal{C} = \{ A \subseteq X : A^c \in \mathcal{T} \} \), one has \( \mathcal{C} \subseteq \mathcal{B} \) and, thus, \( \sigma(\mathcal{C}) \subseteq \mathcal{B} \) (since \( \mathcal{B} \) is a \( \sigma \)-algebra). Analogously, one also has \( \mathcal{T} \subseteq \sigma(\mathcal{C}) \) and \( \mathcal{B} \subseteq \sigma(\mathcal{C}) \), proving (a).

(b): According to [Phi16b, Prop. 3.9(b)], every compact subset of \( X \) is closed, i.e. \( \mathcal{K} \subseteq \mathcal{C} \), implying \( \sigma(\mathcal{K}) \subseteq \sigma(\mathcal{C}) = \mathcal{B} \). On the other hand, if \( A \in \mathcal{C} \), then

\[
A = A \cap X = A \cap \bigcup_{n=1}^{\infty} K_n = \bigcup_{n=1}^{\infty} (A \cap K_n),
\]

i.e. \( A \) is a countable union of sets from \( \mathcal{K} \), showing \( \mathcal{C} \subseteq \sigma(\mathcal{K}) \) and \( \mathcal{B} \subseteq \sigma(\mathcal{K}) \), proving (b).

(c): The hypothesis of (c) implies \( \mathcal{T} \subseteq \sigma(\mathcal{E}) \) and \( \mathcal{B} \subseteq \sigma(\mathcal{E}) \). As \( \mathcal{E} \subseteq \mathcal{B} \) is assumed, this already proves (c). \( \blacksquare \)

Keeping in mind our considerations of the previous section, we are now ready to define what we mean by a measure:

**Definition 1.10.** Let \((X, \mathcal{A})\) be a measurable space.
(a) A map \( \mu : \mathcal{A} \rightarrow [0, \infty] \) is called a measure on \((X, \mathcal{A})\) (or on \(X\) if \(\mathcal{A}\) is understood) if, and only if, \(\mu\) satisfies the following conditions (i) and (ii):

(i) \(\mu(\emptyset) = 0\).

(ii) \(\mu\) is countably additive (also called \(\sigma\)-additive), i.e., if \((A_i)_{i \in \mathbb{N}}\) is a sequence in \(\mathcal{A}\) consisting of (pairwise) disjoint sets, then

\[
\mu \left( \bigcup_{i=1}^{\infty} A_i \right) = \sum_{i=1}^{\infty} \mu(A_i). \tag{1.1}
\]

If \(\mu\) is a measure on \((X, \mathcal{A})\), then the triple \((X, \mathcal{A}, \mu)\) is called a measure space. In the context of measure spaces, the sets \(A \in \mathcal{A}\) are sometimes called \(\mu\)-measurable instead of \(A\)-measurable.

(b) Let \((X, \mathcal{A}, \mu)\) be a measure space. The measure \(\mu\) is called finite or bounded if, and only if, \(\mu(X) < \infty\) (for \(\mu(X) = 1\), \(\mu\) is called a probability measure, \((X, \mathcal{A}, \mu)\) is called a probability space); it is called \(\sigma\)-finite if, and only if, there exists a sequence \((A_i)_{i \in \mathbb{N}}\) in \(\mathcal{A}\) such that \(X = \bigcup_{i=1}^{\infty} A_i\) and \(\mu(A_i) < \infty\) for each \(i \in \mathbb{N}\).

Remark 1.11. In (1.1), we need to extend addition to \([0, \infty]\). A countable sum of summands from \([0, \infty]\) is defined to be \(\infty\) if at least one of the summands equals \(\infty\); if all summands are finite, then the partial sums either converge to some \(s \in \mathbb{R}_0^+\) or they diverge to \(s = \infty\). In both cases, we define \(s\) to be the value of the sum. Throughout the class, we will sometimes need arithmetic in the so-called extended real numbers \(\mathbb{R} := \mathbb{R} \cup \{-\infty, \infty\}\). It is mostly straightforward to extend addition, subtraction, and multiplication from \(\mathbb{R}\) to \(\mathbb{R}\) – the details are provided in Appendix A for future reference.

Example 1.12. Let \((X, \mathcal{A})\) be a measurable space.

(a) The zero measure \(\mu_0 : \mathcal{A} \rightarrow [0, \infty], \mu_0 \equiv 0\), is clearly a measure, and so is

\[
\mu_\infty : \mathcal{A} \rightarrow [0, \infty], \quad \mu_\infty(A) := \begin{cases} 0 & \text{if } A \text{ is countable,} \\ \infty & \text{if } A \text{ is uncountable.} \end{cases}
\]

Then \(\mu_\infty\) is not \(\sigma\)-finite if \(X\) is uncountable (if \(X\) is countable, then \(\mu_\infty = \mu_0\)). For \(X \neq \emptyset\), \(A = \{\emptyset, X\}\), a variant is the measure \(\mu\) defined by \(\mu(\emptyset) = 0\), \(\mu(X) = \infty\), which is never \(\sigma\)-finite.

(b) Another obvious measure is the so-called counting measure, defined by

\[
\mu_c : \mathcal{A} \rightarrow [0, \infty], \quad \mu_c(A) := \begin{cases} \#A & \text{if } A \text{ is finite,} \\ \infty & \text{if } A \text{ is infinite.} \end{cases}
\]

If \(\mathcal{A} = \mathcal{P}(X)\), then counting measure is \(\sigma\)-finite if, and only if, \(X\) is countable.
(c) If $X \neq \emptyset$ and $a \in X$, then

$$
\delta_a : A \rightarrow [0, \infty], \quad \delta_a (A) := \begin{cases} 1 & \text{if } a \in A, \\ 0 & \text{if } a \notin A, \end{cases}
$$

is called the Dirac measure concentrated in $a$ (clearly, it is a measure).

(d) Let $X$ be nonempty and finite, $\#X = n \in \mathbb{N}$. Let $m : X \rightarrow \mathbb{R}_0^+$ be some function. Then

$$
\mu_m : \mathcal{P}(X) \rightarrow \mathbb{R}_0^+, \quad \mu_m (A) := \sum_{x \in A} m(x),
$$

defines a measure on $X$. Moreover, if $M := \mu_m (X) \neq 0$, then $P := M^{-1} \mu_m$ defines a probability measure on $X$.

1.3 Extension

The examples at the end of the previous section show that there are many measures that are simple to define and simple to recognize as measures. However, to accomplish our original goal of defining a measure on the Borel subsets of $\mathbb{R}_n$ that yields the expected geometric size of intervals, is not so simple. We will be able to do this using the technique of extension, which is also useful beyond the application we have in mind. The idea is to first define a suitable $[0, \infty]$-valued function on some set $\mathcal{E}$ that is not a $\sigma$-algebra (e.g. the set of intervals) and then to extend this function to a measure on the $\sigma$-algebra generated by $\mathcal{E}$. We will extend our functions from so-called semirings to $\sigma$-algebras. Useful structures in between are rings and algebras.

1.3.1 Semirings, Rings, Algebras

Definition 1.13. Let $X$ be a set, $\mathcal{E} \subseteq \mathcal{P}(X)$.

(a) $\mathcal{E}$ is called a semiring on $X$ if, and only if, it satisfies the following three conditions:

(i) $\emptyset \in \mathcal{E}$.

(ii) If $A, B \in \mathcal{E}$, then $A \cap B \in \mathcal{E}$.

(iii) If $A, B \in \mathcal{E}$, then there exist disjoint $E_1, \ldots, E_n \in \mathcal{E}$, $n \in \mathbb{N}$, such that

$$
A \setminus B = \bigcup_{i=1}^{n} E_i. \tag{1.2}
$$

(b) $\mathcal{E}$ is called a ring on $X$ if, and only if, it satisfies the following three conditions:

(i) $\emptyset \in \mathcal{E}$.

(ii) If $A, B \in \mathcal{E}$, then $A \cup B \in \mathcal{E}$.
The sets $E$ imply Def. 1.13(a)(ii). However, if everything is immediate from the respective definitions, except that Def. 1.13(b) yields again rings, arbitrary intersections of algebras yield again algebras: Let $A, B \in \mathcal{E}$ be a nonempty index set, and let $(\mathcal{R}_i)_{i \in I}$ be a family of rings on $X$, $\mathcal{R} := \bigcap_{i \in I} \mathcal{R}_i$. Since $\emptyset$ is in each $\mathcal{R}_i$, it is also in $\mathcal{R}$. If $A, B \in \mathcal{R}$, then $A, B \in \mathcal{R}_i$ for each $i \in I$, i.e. $A \setminus B \in \mathcal{R}_i$ and $A \cup B \in \mathcal{R}_i$ for each $i \in I$, implying $A \setminus B \in \mathcal{R}$ and $A \cup B \in \mathcal{A}$, showing $\mathcal{R}$ to be a ring on $X$. Now let $(\mathcal{A}_i)_{i \in I}$ be a family of algebras on $X$, $\mathcal{A} := \bigcap_{i \in I} \mathcal{A}_i$. Since $X$ is in each $\mathcal{A}_i$, it is also in $\mathcal{A}$. Since each $\mathcal{A}_i$ is a ring on $X$, so is $\mathcal{A}$, showing $\mathcal{A}$ is an algebra on $X$. Thus, if $X$ is a set and $\mathcal{E} \subseteq \mathcal{P}(X)$, then we can let $\rho(\mathcal{E}) := \rho_X(\mathcal{E})$ denote the intersection of all rings on $X$ that are supersets of $\mathcal{E}$, and we can let $\alpha(\mathcal{E}) := \alpha_X(\mathcal{E})$ denote the intersection of all algebras on $X$ that are supersets of $\mathcal{E}$. We then know $\rho(\mathcal{E})$ to be a ring and $\alpha(\mathcal{E})$ to be an algebra, namely the smallest ring (resp. algebra) on $X$ containing $\mathcal{E}$. Thus, we call $\rho(\mathcal{E})$ the ring (and $\alpha(\mathcal{E})$ the algebra) generated by $\mathcal{E}$. We point out that there exists a (rough) conceptual analogy between the way a $\sigma$-algebra can be built from a semiring and the way a topology can be built from a subbase.

**Remark 1.14.** Let $X$ be a set, $\mathcal{E} \subseteq \mathcal{P}(X)$. Then we have the following implications:

$$\mathcal{E} \sigma\text{-algebra} \Rightarrow \mathcal{E} \text{ algebra} \Rightarrow \mathcal{E} \text{ ring} \Rightarrow \mathcal{E} \text{ semiring},$$

where everything is immediate from the respective definitions, except that Def. 1.13(b) implies Def. 1.13(a)(ii). However, if $\mathcal{E}$ is a ring and $A, B \in \mathcal{E}$, then

$$A \cap B = A \setminus (A \setminus B) \in \mathcal{E}. \quad (1.3)$$

We also note that $\mathcal{E}$ is a ring in the sense of Def. 1.13(b) if, and only if, $\mathcal{E}$ is a ring in the sense of algebra, provided one uses the symmetric difference $A \Delta B := (A \setminus B) \cup (B \setminus A)$ as addition and $\cap$ as multiplication (see Appendix B).

**Remark and Definition 1.15.** Analogous to Prop. 1.5, arbitrary intersections of rings yield again rings, arbitrary intersections of algebras yield again algebras: Let $X$ be a set, let $I$ be a nonempty index set, and let $(\mathcal{R}_i)_{i \in I}$ be a family of rings on $X$, $\mathcal{R} := \bigcap_{i \in I} \mathcal{R}_i$. Since $\emptyset$ is in each $\mathcal{R}_i$, it is also in $\mathcal{R}$. If $A, B \in \mathcal{R}$, then $A, B \in \mathcal{R}_i$ for each $i \in I$, i.e. $A \setminus B \in \mathcal{R}_i$ and $A \cup B \in \mathcal{R}_i$ for each $i \in I$, implying $A \setminus B \in \mathcal{R}$ and $A \cup B \in \mathcal{A}$, showing $\mathcal{R}$ to be a ring on $X$. Now let $(\mathcal{A}_i)_{i \in I}$ be a family of algebras on $X$, $\mathcal{A} := \bigcap_{i \in I} \mathcal{A}_i$. Since $X$ is in each $\mathcal{A}_i$, it is also in $\mathcal{A}$. Since each $\mathcal{A}_i$ is a ring on $X$, so is $\mathcal{A}$, showing $\mathcal{A}$ is an algebra on $X$. Thus, if $X$ is a set and $\mathcal{E} \subseteq \mathcal{P}(X)$, then we can let $\rho(\mathcal{E}) := \rho_X(\mathcal{E})$ denote the intersection of all rings on $X$ that are supersets of $\mathcal{E}$, and we can let $\alpha(\mathcal{E}) := \alpha_X(\mathcal{E})$ denote the intersection of all algebras on $X$ that are supersets of $\mathcal{E}$. We then know $\rho(\mathcal{E})$ to be a ring and $\alpha(\mathcal{E})$ to be an algebra, namely the smallest ring (resp. algebra) on $X$ containing $\mathcal{E}$. Thus, we call $\rho(\mathcal{E})$ the ring (and $\alpha(\mathcal{E})$ the algebra) generated by $\mathcal{E}$. We point out that there exists a (rough) conceptual analogy between the way a $\sigma$-algebra can be built from a semiring and the way a topology can be built from a subbase.

**Example 1.16. (a)** Let $X$ be a set. Then $\mathcal{S} := \{\emptyset\} \cup \{\{x\} : x \in X\}$ is a semiring on $X$. Moreover, $\rho(\mathcal{S}) = \{A \subseteq X : A \text{ finite}\}$, $\alpha(\mathcal{S}) = \{A \subseteq X : A \text{ finite or } A^c \text{ finite}\}$, and $\sigma(\mathcal{S})$ is the $\sigma$-algebra of Ex. 1.4(b).

**Example 1.16. (b)** The sets $\mathcal{I}_1 = \{[a, b] : a, b \in \mathbb{R}, a \leq b\}$ and $\mathcal{I}_1^\mathbb{Q} = \{[a, b] : a, b \in \mathbb{Q}, a \leq b\}$ of Ex. 1.8(c) are semirings on $\mathbb{R}$, since, for $a, b, c, d \in \mathbb{R}$ with $a \leq b$ and $c \leq d$, letting $\alpha := \max\{a, c\}$, $\beta := \min\{b, d\}$, one has

$$[a, b] \cap [c, d] = \begin{cases} [\alpha, \beta] & \text{if } \alpha < \beta, \\ \emptyset & \text{otherwise}, \end{cases}$$

as well as

$$[a, b] \setminus [c, d] = [a, c] \cup \begin{cases} [d, b] & \text{for } c \leq a \\ \emptyset & \text{for } b \leq d \end{cases}.$$
If \( n > 1 \), then \( \mathcal{I}^n \) and \( \mathcal{I}_Q^n \) are still semirings on \( \mathbb{R}^n \):

**Proposition 1.17. (a)** If \( X, Y \) are sets, \( \mathcal{S} \) is a semiring on \( X \) and \( \mathcal{T} \) is a semiring on \( Y \), then

\[
\mathcal{U} := \mathcal{S} \ast \mathcal{T} := \{ A \times B : A \in \mathcal{S}, B \in \mathcal{T} \}
\]

is a semiring on \( X \times Y \).

(b) For each \( n \in \mathbb{N} \), the sets \( \mathcal{I}^n = \{ [a, b] : a, b \in \mathbb{R}^n, a \leq b \} \) and \( \mathcal{I}_Q^n = \{ [a, b] : a, b \in \mathbb{Q}^n, a \leq b \} \) of Ex. 1.8(c) are semirings on \( \mathbb{R}^n \).

**Proof.** (a): It is \( \emptyset = \emptyset \times \emptyset \in \mathcal{U} \). Now let \( A_1, A_2 \in \mathcal{S} \) and \( B_1, B_2 \in \mathcal{T} \). Then

\[
(A_1 \times B_1) \cap (A_2 \times B_2) = (A_1 \cap A_2) \times (B_1 \cap B_2) \in \mathcal{U}.
\]

Moreover,

\[
(A_1 \times B_1) \setminus (A_2 \times B_2) = ((A_1 \setminus A_2) \times B_1) \cup (A_1 \times (B_1 \setminus B_2))
\]

\[
= ((A_1 \setminus A_2) \times B_1) \cup ((A_1 \cap A_2) \times (B_1 \setminus B_2)).
\]

As \( \mathcal{S} \) and \( \mathcal{T} \) are semirings, there are \( M, N \in \mathbb{N} \), disjoint \( C_1, \ldots, C_M \in \mathcal{S} \), and disjoint \( D_1, \ldots, D_N \in \mathcal{T} \) such that

\[
A_1 \setminus A_2 = \bigcup_{i=1}^{M} C_i, \quad B_1 \setminus B_2 = \bigcup_{i=1}^{N} D_i.
\]

Using this in (1.4), we can write \( (A_1 \times B_1) \setminus (A_2 \times B_2) \) as a disjoint union of sets in \( \mathcal{U} \), namely

\[
(A_1 \times B_1) \setminus (A_2 \times B_2) = \left( \bigcup_{i=1}^{M} (C_i \times B_1) \right) \cup \left( \bigcup_{i=1}^{N} ((A_1 \cap A_2) \times D_i) \right),
\]

which completes the proof that \( \mathcal{U} \) is a semiring.

(b) follows by induction on \( n \): The base case \( (n = 1) \) is given by Ex. 1.16(b), and the induction step follows from (a), since, for \( n > 1 \),

\[
[a_1, \ldots, a_n], \quad (b_1, \ldots, b_n] = [a_1, \ldots, a_{n-1}], \quad (b_1, \ldots, b_{n-1})] \times ]a_n, b_n[
\]

for each \( a, b \in \mathbb{R}^n \) with \( a \leq b \). \( \blacksquare \)

**Proposition 1.18. (a)** Let \( \mathcal{S} \) be a semiring on \( X \) and \( A, B_1, \ldots, B_N \in \mathcal{S} \), \( N \in \mathbb{N} \). Then there exists \( M \in \mathbb{N} \) and disjoint \( C_1, \ldots, C_M \in \mathcal{S} \) such that

\[
A \setminus \bigcup_{i=1}^{N} B_i = \bigcup_{i=1}^{M} C_i.
\]

(b) If \( \mathcal{S} \) is a semiring on \( X \), then

\[
\rho(\mathcal{S}) = \mathcal{R} := \{ \bigcup_{k=1}^{N} A_k : N \in \mathbb{N}, A_1, \ldots, A_N \in \mathcal{S} \text{ disjoint} \}.
\]
(c) For each \( n \in \mathbb{N} \), one has \( \rho(\mathcal{I}_n) = \mathcal{I}_{\Sigma}^n \) and \( \rho(\mathcal{I}_Q^n) = \mathcal{I}_{\Sigma,Q}^n \), where \( \mathcal{I}_n, \mathcal{I}_{\Sigma}, \mathcal{I}_Q, \) and \( \mathcal{I}_{\Sigma,Q} \) are as defined in Ex. 1.8(c).

Proof. (a) is proved via induction on \( N \). The base case \( (N = 1) \) holds, since \( S \) is a semiring. For the induction step, let \( N \in \mathbb{N}, A, B_1, \ldots, B_N, B_{N+1} \in S \). By the induction hypothesis, there exist disjoint \( C_1, \ldots, C_M \in S \) such that (1.5) holds true. Then

\[
A \setminus \bigcup_{i=1}^{N+1} B_i = \left( A \setminus \bigcup_{i=1}^N B_i \right) \setminus B_{N+1} = \left( \bigcup_{i=1}^M C_i \right) \setminus B_{N+1} = \bigcup_{i=1}^M (C_i \setminus B_{N+1}).
\]

Since each of the \( C_i \setminus B_{N+1} \) is a disjoint finite union of sets from \( S \), the induction is complete.

(b): Since rings are closed under finite unions, \( R \subseteq \rho(S) \) is clear. For the remaining inclusion, it suffices to show that \( R \) is a ring. Since \( \emptyset \in S, \emptyset \in R \). Let \( M, N \in \mathbb{N}, A_1, \ldots, A_M \in S \) disjoint, \( B_1, \ldots, B_N \in S \) disjoint, \( A := \bigcup_{i=1}^M A_i, B := \bigcup_{i=1}^N B_i \). We have to show \( A \setminus B \in R \) and \( A \cup B \in R \). It is

\[
A \setminus B = \bigcup_{i=1}^M \left( A_i \setminus \bigcup_{j=1}^N B_j \right).
\]

Since, by (a), each of the \( A_i \setminus \bigcup_{j=1}^N B_j \) is a disjoint finite union of sets from \( S \), we have \( A \setminus B \in R \). We also have

\[
A \cap B = \bigcup_{i=1}^M \bigcup_{j=1}^N A_i \cap B_j \in R,
\]

since each \( A_i \cap B_j \in S \) by Def. 1.13(a)(ii). Thus, \( A \cup B = (A \setminus B) \cup (B \setminus A) \cup (A \cap B) \in R \), completing the proof that \( R \) is a ring and the proof of (b).

(c) follows by combining Prop. 1.17(b) with (b).

1.3.2 Contents, Premeasures

Definition 1.19. Let \( X \) be a set and let \( S \) be a semiring on \( X \).

(a) A map \( \mu : S \rightarrow [0, \infty) \) is called a content if, and only if, \( \mu \) satisfies the following conditions (i) and (ii):

(i) \( \mu(\emptyset) = 0 \).

(ii) \( \mu \) is finitely additive, i.e., if \( n \in \mathbb{N} \) and \( A_1, \ldots, A_n \in S \) are disjoint sets such that \( \bigcup_{i=1}^n A_i \in S \), then

\[
\mu \left( \bigcup_{i=1}^n A_i \right) = \sum_{i=1}^n \mu(A_i).
\]

(1.7)
1 MEASURE THEORY

The content \( \mu \) is called finite or bounded if, and only if, \( \mu(A) < \infty \) for each \( A \in \mathcal{S} \); it is called \( \sigma \)-finite if, and only if, there exists a sequence \( (A_i)_{i \in \mathbb{N}} \) in \( \mathcal{S} \) such that \( X = \bigcup_{i=1}^{\infty} A_i \) and \( \mu(A_i) < \infty \) for each \( i \in \mathbb{N} \).

(b) A content \( \mu : \mathcal{S} \rightarrow [0, \infty] \) is called a premeasure if, and only if, it is countably additive (also called \( \sigma \)-additive), i.e., if \( (A_i)_{i \in \mathbb{N}} \) is a sequence in \( \mathcal{S} \) consisting of disjoint sets such that \( \bigcup_{i=1}^{\infty} A_i \in \mathcal{S} \), then

\[
\mu \left( \bigcup_{i=1}^{\infty} A_i \right) = \sum_{i=1}^{\infty} \mu(A_i).
\]

Example 1.20. (a) A measure is a premeasure that is defined on a \( \sigma \)-algebra.

(b) Among the most important premeasures is so-called Lebesgue premeasure \( \lambda^n \), which, for \( n \in \mathbb{N} \), is defined on the semiring \( \mathcal{I}^n \) on \( \mathbb{R}^n \) from Ex. 1.16(c) by letting

\[
\lambda^n : \mathcal{I}^n \rightarrow \mathbb{R}_0^+, \quad \lambda^n(I) := \begin{cases} 
\prod_{j=1}^{n} (b_j - a_j) & \text{for } I = [a, b], a, b \in \mathbb{R}^n, a < b, \\
0 & \text{for } I = \emptyset.
\end{cases}
\]

We will study it in detail in Sec. 1.3.3 below, where we will show in Th. 1.31 that it is, indeed, a premeasure.

(c) The following example shows that not every content is a premeasure: Consider \( \mathcal{S} := \{ [a, b] : a, b \in \mathbb{R}, 0 \leq a \leq b \leq 1 \} \). Then \( \mathcal{S} \) is clearly a semiring on \( [0, 1] \) (cf. Ex. 1.16(b)). Define

\[
\mu : \mathcal{S} \rightarrow [0, \infty], \quad \mu(I) := \begin{cases} 
0 & \text{for } I = \emptyset, \\
b - a & \text{for } I = [a, b], 0 < a < b \leq 1, \\
\infty & \text{for } I = [0, b], 0 < b \leq 1.
\end{cases}
\]

Then \( \mu \) is a content, since

\[
\forall n \in \mathbb{N} \quad 0 < s_0 < s_1 < \cdots < s_n \leq 1 \quad \sum_{i=1}^{n} \mu([s_{i-1}, s_i]) = \sum_{i=1}^{n} (s_i - s_{i-1}) = s_n - s_0 = \mu([s_0, s_n])
\]

and

\[
\forall n \in \mathbb{N} \quad 0 < s_0 < s_1 < \cdots < s_n \leq 1 \quad \mu([0, s_0]) + \sum_{i=1}^{n} \mu([s_{i-1}, s_i]) = \infty + \sum_{i=1}^{n} (s_i - s_{i-1}) = \infty = \mu([0, s_n]).
\]

However, \( \mu \) is not a premeasure, since, letting, for each \( n \in \mathbb{N} \), \( I_n := \left[ \frac{1}{n+1}, \frac{1}{n} \right] \), one has \( [0, 1] = \bigcup_{n=1}^{\infty} I_n \), but

\[
\sum_{n=1}^{\infty} \mu(I_n) = \sum_{n=1}^{\infty} \left( \frac{1}{n} - \frac{1}{n+1} \right) = \lim_{n \to \infty} \left( 1 - \frac{1}{n+1} \right) = 1 \neq \mu([0, 1]) = \infty.
\]
Before compiling important properties of contents and premeasures in Th. 1.22 and Th. 1.24 below, we show in the following Prop. 1.21 that each content on a semiring extends in a unique way to a content on the generated ring.

**Proposition 1.21.** Let $S$ be a semiring on the set $X$ and let $\mu : S \rightarrow [0, \infty]$ be a content. Furthermore, let $R := \rho(S)$. Then there is a unique extension $\nu : R \rightarrow [0, \infty]$ of $\mu$ such that $\nu$ is a content. Moreover $\nu$ is a premeasure if, and only if, $\mu$ is a premeasure.

**Proof.** Uniqueness: Assume $\nu$ is a content on $R$ such that $\nu|_S = \mu$. If $A \in R$, then, by Prop. 1.18(b), there exist disjoint $A_1, \ldots, A_N \in S, N \in \mathbb{N}$, such that $A = \bigcup_{i=1}^N A_i$. Thus, as $\nu$ is finitely additive, $\nu(A) = \sum_{i=1}^N \mu(A_i)$, showing that $\nu$ is uniquely determined by $\mu$.

Existence: As we have seen that $\nu(A) = \sum_{i=1}^N \mu(A_i)$ must hold if $A$ and $A_1, \ldots, A_N$ are as above, we use this formula to define the function $\nu$ on $R$. We now need to verify that $\nu$ is well-defined, i.e. that, if we pick, possibly different, disjoint $B_1, \ldots, B_M \in S$ with $A = \bigcup_{i=1}^M B_i$, then $\sum_{i=1}^N \mu(A_i) = \sum_{i=1}^M \mu(B_i)$. Indeed, we have, since $\mu$ is a content on $S$,

$$
\sum_{i=1}^N \mu(A_i) = \sum_{i=1}^N \mu \left( A_i \cap \bigcup_{j=1}^M B_j \right) = \sum_{i=1}^N \mu \left( \bigcup_{j=1}^M (A_i \cap B_j) \right) = \sum_{i=1}^N \sum_{j=1}^M \mu(A_i \cap B_j) = \sum_{j=1}^M \mu \left( \bigcup_{i=1}^N (A_i \cap B_j) \right) = \sum_{j=1}^M \mu(B_j),
$$

proving that $\nu$ is well-defined. Then $\nu|_S = \mu$ is also clear. To prove that $\nu$ is a content, we still need to check that it is finitely additive. To this end, let $A, B \in R$ be disjoint. Then there are disjoint $A_1, \ldots, A_N, B_1, \ldots, B_M \in S$; $M, N \in \mathbb{N}$, such that $A = \bigcup_{i=1}^N A_i$ and $B = \bigcup_{i=1}^M B_i$. In consequence,

$$
\nu(A \cup B) = \sum_{i=1}^N \mu(A_i) + \sum_{i=1}^M \mu(B_i) = \nu(A) + \nu(B),
$$

as needed.

Since $\nu|_S = \mu$, if $\nu$ is a premeasure, so is $\mu$. For the converse, assume $\mu$ to be a premeasure and let $(A_i)_{i \in \mathbb{N}}$ be a sequence of disjoint sets in $R$ such that $A := \bigcup_{i=1}^\infty A_i \in R$. Then there are disjoint sets $B_1, \ldots, B_N \in S$, $N \in \mathbb{N}$, with $A = \bigcup_{i=1}^N B_i$, and, for each $i \in \mathbb{N}$, there are disjoint sets $C_{i1}, \ldots, C_{iN_i} \in S$, $i \in \mathbb{N}$, with $A_i = \bigcup_{j=1}^{N_i} C_{ij}$. Then

$$
\forall \ i \in \{1, \ldots, N\} \quad B_i = B_i \cap A = \bigcup_{j=1}^\infty (B_i \cap A_j) = \bigcup_{j=1}^\infty \bigcup_{k=1}^{N_j} (B_i \cap C_{jk}) \in S.
$$

Thus, as $\mu$ is a premeasure,

$$
\forall \ i \in \{1, \ldots, N\} \quad \mu(B_i) = \sum_{j=1}^{N_j} \sum_{k=1}^\infty \mu(B_i \cap C_{jk}) = \sum_{j=1}^\infty \nu(B_i \cap A_j),
$$

as needed.
implying
\[ \nu(A) = \sum_{i=1}^{N} \mu(B_i) = \sum_{i=1}^{N} \sum_{j=1}^{\infty} \nu(B_i \cap A_j) = \sum_{j=1}^{\infty} \nu(A \cap A_j) = \sum_{j=1}^{\infty} \nu(A_j), \]
proving \( \nu \) to be a premeasure.

During the construction of a content \( \mu \), it can be helpful to first define and study it on a semiring. However, Prop. 1.21 tells us that we may assume \( \mu \) to be defined on a ring when considering a content’s general properties, as in the following Th. 1.22 (one still has to use some care, though, cf. Ex. 1.25(c) below).

**Theorem 1.22.** Let \( \mathcal{R} \) be a ring on the set \( X \) and let \( \mu : \mathcal{R} \rightarrow [0, \infty] \) be a content. Then the following rules hold, where we assume \( A, B \in \mathcal{R} \) as well as \( A_i \in \mathcal{R} \) for each \( i \in \mathbb{N} \).

(a) **Monotonicity:** If \( B \subseteq A \), then \( \mu(B) \leq \mu(A) \).

(b) **Subtractivity:** If \( B \subseteq A \) and \( \mu(B) < \infty \), then \( \mu(A \setminus B) = \mu(A) - \mu(B) \).

(c) \( \mu(A) + \mu(B) = \mu(A \cup B) + \mu(A \cap B) \).

(d) If the \( A_i \) are disjoint and \( \bigcup_{i=1}^{\infty} A_i \subseteq A \), then \( \sum_{i=1}^{\infty} \mu(A_i) \leq \mu(A) \).

(e) **Subadditivity:** For each \( n \in \mathbb{N} \), one has \( \mu(\bigcup_{i=1}^{n} A_i) \leq \sum_{i=1}^{n} \mu(A_i) \).

(f) If \( \mu \) is a premeasure, then one also has \( \sigma \)-subadditivity: If \( A \subseteq \bigcup_{i=1}^{\infty} A_i \), then \( \mu(A) \leq \sum_{i=1}^{\infty} \mu(A_i) \).

**Proof.** (a): If \( B \subseteq A \), then \( A = B \cup (A \setminus B) \), implying
\[ \mu(A) = \mu(B) + \mu(A \setminus B). \] (1.8)
Since \( \mu(A \setminus B) \in [0, \infty] \), this shows \( \mu(B) \leq \mu(A) \).

(b) is immediate from (1.8), as \( \mu(B) \) is assumed finite.

(c): Finite additivity yields
\[ \mu(B) = \mu(B \setminus A) + \mu(A \cap B), \]
\[ \mu(A \cup B) = \mu(A) + \mu(B \setminus A). \]
If \( \mu(A \cup B) < \infty \), then we add \( \mu(A) = \mu(A \cup B) - \mu(B \setminus A) \) to the first equation to obtain (c). If \( \mu(A \cup B) = \mu(A) = \infty \), then (c) also holds. In the remaining case, \( \mu(A \cup B) = \infty \) and \( \mu(A) < \infty \), i.e. \( \mu(B \setminus A) = \infty \), implying \( \mu(B) = \infty \), i.e. (c) holds once again.

(d) holds, since, by finite additivity and monotonicity, \( \sum_{i=1}^{n} \mu(A_i) \leq \mu(A) \) holds for each \( n \in \mathbb{N} \).
(e): Since \( \bigcup_{i=1}^{n} A_i = \bigcup_{i=1}^{n} (A_i \setminus \bigcup_{j=1}^{n-1} A_j) \), we have

\[
\mu \left( \bigcup_{i=1}^{n} A_i \right) = \sum_{i=1}^{n} \mu \left( A_i \setminus \bigcup_{j=1}^{n-1} A_j \right) \leq \sum_{i=1}^{n} \mu(A_i).
\]

(f) follows with the same kind of trick as in (e): Since \( A = \bigcup_{i=1}^{n} (A \cap (A_i \setminus \bigcup_{j=1}^{n-1} A_j)) \), we have

\[
\mu(A) = \sum_{i=1}^{\infty} \mu \left( A \cap \left( A_i \setminus \bigcup_{j=1}^{n-1} A_j \right) \right) \leq \sum_{i=1}^{\infty} \mu(A_i),
\]

completing the proof of (f) and the theorem. \(\blacksquare\)

**Notation 1.23.** Let \( A \) and \( A_n, n \in \mathbb{N} \), be sets. We introduce the following notation:

\[
A_n \uparrow A :\Leftrightarrow \left( A = \bigcup_{n=1}^{\infty} A_n \land A_1 \subseteq A_2 \subseteq \ldots \right),
\]

\[
A_n \downarrow A :\Leftrightarrow \left( A = \bigcap_{n=1}^{\infty} A_n \land A_1 \supseteq A_2 \supseteq \ldots \right).
\]

**Theorem 1.24.** Let \( \mathcal{R} \) be a ring on the set \( X \), let \( \mu : \mathcal{R} \to [0, \infty] \) be a content, and consider the following statements:

\begin{enumerate}
  \item \( \mu \) is a premeasure.
  \item \( \mu \) is continuous from below, i.e. if \( A \in \mathcal{R} \) and \( (A_n)_{n \in \mathbb{N}} \) is a sequence in \( \mathcal{R} \), then
    \[
    A_n \uparrow A \implies \lim_{n \to \infty} \mu(A_n) = \mu(A). \tag{1.9}
    \]
  \item \( \mu \) is continuous from above, i.e. if \( A \in \mathcal{R} \) and \( (A_n)_{n \in \mathbb{N}} \) is a sequence in \( \mathcal{R} \), then
    \[
    \left( \mu(A_1) < \infty \land A_n \downarrow A \right) \implies \lim_{n \to \infty} \mu(A_n) = \mu(A). \tag{1.10}
    \]
  \item If \( (A_n)_{n \in \mathbb{N}} \) is a sequence in \( \mathcal{R} \), then
    \[
    \left( \mu(A_1) < \infty \land A_n \downarrow \emptyset \right) \implies \lim_{n \to \infty} \mu(A_n) = 0. \tag{1.11}
    \]
\end{enumerate}

Then the following implications hold:

\[
(i) \iff (ii) \implies (iii) \iff (iv).
\]

Under the additional assumption that \( \mu \) is finite, (i) – (iv) are equivalent.
1 MEASURE THEORY

Proof. (i) ⇒ (ii): Let \( A \in \mathcal{R} \) and let \((A_n)_{n \in \mathbb{N}}\) be a sequence in \( \mathcal{R} \) such that \( A_n \uparrow A \). Then \( A = A_1 \cup \bigcup_{i=2}^{\infty} (A_i \setminus A_{i-1}) \), implying

\[
\mu(A) = \mu(A_1) + \sum_{i=2}^{\infty} \mu(A_i \setminus A_{i-1}) = \lim_{n \to \infty} \left( \mu(A_1) + \sum_{i=2}^{n} \mu(A_i \setminus A_{i-1}) \right) = \lim_{n \to \infty} \mu(A_n),
\]
as desired.

(ii) ⇒ (i): If \( A \in \mathcal{R} \) and \((B_i)_{i \in \mathbb{N}}\) is a sequence of disjoint sets in \( \mathcal{R} \), satisfying \( A = \bigcup_{i=1}^{\infty} B_i \), then, letting, for each \( n \in \mathbb{N} \), \( A_n := \bigcup_{i=1}^{n} B_i \in \mathcal{R} \), one has \( A_n \uparrow A \). Then (ii) yields

\[
\mu(A) = \lim_{n \to \infty} \mu(A_n) = \lim_{n \to \infty} \sum_{i=1}^{n} \mu(B_i) = \sum_{i=1}^{\infty} \mu(B_i),
\]
showing \( \mu \) to be a premeasure.

(iii) ⇒ (ii): Let \( A \in \mathcal{R} \) and let \((A_n)_{n \in \mathbb{N}}\) be a sequence in \( \mathcal{R} \) such that \( A_n \downarrow A, \mu(A_1) < \infty \). Then \( \mu(A) < \infty \) and \( \mu(A_n) < \infty \) for each \( n \in \mathbb{N} \). Moreover,

\[
A_n \downarrow A \quad \Rightarrow \quad \left( A = \bigcap_{n=1}^{\infty} A_n \land A_1 \supset A_2 \supset \ldots \right)
\]

\[
\Rightarrow \quad \left( A_1 \setminus A = \bigcup_{n=1}^{\infty} (A_1 \setminus A_n) \land A_1 \supset A_1 \setminus A_2 \subseteq \ldots \right)
\]

\[
\Rightarrow \quad A_1 \setminus A_n \uparrow A_1 \setminus A.
\]

Thus, we can apply Th. 1.22(b) and (ii) to obtain

\[
\lim_{n \to \infty} \left( \mu(A_1) - \mu(A_n) \right) = \lim_{n \to \infty} \mu(A_1 \setminus A_n) = \mu(A_1 \setminus A) = \mu(A_1) - \mu(A),
\]
proving (iii).

(iii) ⇒ (iv) is immediate by setting \( A := \emptyset \).

(iv) ⇒ (iii): Let \( A \in \mathcal{R} \) and let \((A_n)_{n \in \mathbb{N}}\) be a sequence in \( \mathcal{R} \) such that \( A_n \downarrow A, \mu(A_1) < \infty \). For each \( n \in \mathbb{N} \), set \( B_n := A_n \setminus A \in \mathcal{R} \). Then \( B_n \downarrow \emptyset \) and \( \mu(B_1) \leq \mu(A_1) < \infty \), i.e. we can apply (iv) to obtain

\[
\lim_{n \to \infty} \left( \mu(A_n) - \mu(A) \right) = \lim_{n \to \infty} \mu(B_n) = 0,
\]
proving (iii).

Finally, assume \( \mu \) to be finite and (iv). Let \( A \in \mathcal{R} \) and let \((A_n)_{n \in \mathbb{N}}\) be a sequence in \( \mathcal{R} \) such that \( A_n \uparrow A \). Then \( A \setminus A_n \downarrow \emptyset \) and, since \( \mu(A \setminus A_1) < \infty \), (iv) yields

\[
\lim_{n \to \infty} \left( \mu(A) - \mu(A_n) \right) = \lim_{n \to \infty} \mu(A \setminus A_n) = 0,
\]
proving (ii) and the theorem.
The following Ex. 1.25(a)–(c) provides some counterexamples related to Th. 1.24 above.

**Example 1.25. (a)** The following example shows that one cannot omit the condition \( \mu(A_1) < \infty \) in (iii) and (iv) of Th. 1.24: Consider the counting measure of Ex. 1.12(b) with \( X := \mathbb{N} \). If, for each \( n \in \mathbb{N} \), \( A_n := \{ k \in \mathbb{N} : k \geq n \} \). Then \( A_n \downarrow \emptyset \), but \( \mu(A_n) = \infty \) for each \( n \in \mathbb{N} \).

(b) The following example shows statements (iii) and (iv) of Th. 1.24 are, in general, not equivalent to statements (i) and (ii): Let \( X \) be a countable infinite set (e.g., \( X = \mathbb{N} \)). Then \( \mathcal{A} := \{ A \subseteq X : A \text{ finite or } A^c \text{ finite} \} \) is clearly an algebra on \( X \) and

\[
\mu : \mathcal{A} \longrightarrow [0, \infty], \quad \mu(A) := \begin{cases} 0 & \text{for } A \text{ finite,} \\ \infty & \text{for } A \text{ infinite,} \end{cases}
\]

clearly defines a content on \( \mathcal{A} \). Moreover, \( \mu \) satisfies (iii) and (iv) of Th. 1.24, since, if \( A \in \mathcal{A} \) and \( (A_n)_{n \in \mathbb{N}} \) is a sequence in \( \mathcal{A} \) with \( A_n \downarrow A \), then \( \mu(A_1) < \infty \) implies that all \( A_n \) and \( A \) are finite with \( \mu(A_n) = \mu(A) = 0 \). However, \( \mu \) is not a premeasure, since \( \sum_{x \in X} \mu(\{x\}) = 0 \neq \infty = \mu(X) \) (where we used that \( X \) is the countable disjoint union of the \( \{x\} \)).

(c) The following example shows Th. 1.24 does, in general, not hold if the ring \( \mathcal{R} \) is replaced with a semiring \( \mathcal{S} \): Let \( X := \mathbb{Q} \) and

\[
\mathcal{S} := \mathcal{I}_1|\mathbb{Q} := \{ [a, b] \cap \mathbb{Q} : a, b \in \mathbb{R}, a \leq b \}.
\]

Then, clearly, \( \mathcal{S} \) is a semiring on \( \mathbb{Q} \). We show

\[
\mu : \mathcal{S} \longrightarrow \mathbb{R}_0^+, \quad \mu([a, b] \cap \mathbb{Q}) := b - a \quad (a \leq b)
\]

defines a finite content, satisfying Th. 1.24(ii) on \( \mathcal{S} \) (and, thus, also Th. 1.24(iii),(iv) on \( \mathcal{S} \), since the corresponding proofs of Th. 1.24 remain valid), but (the extension of) \( \mu \) does not satisfy Th. 1.24(ii) on \( \mathcal{R} := \rho(\mathcal{S}) \) (in particular, \( \mu \) is not a premeasure): Since

\[
\forall n \in \mathbb{N} \quad \forall s_0 < s_1 < \ldots < s_n \quad \sum_{i=1}^{n} \mu([s_{i-1}, s_i] \cap \mathbb{Q}) = \sum_{i=1}^{n} (s_i - s_{i-1}) = s_n - s_0 = \mu([s_0, s_n] \cap \mathbb{Q}),
\]

\( \mu \) is a content. We now show \( \mu \) satisfies Th. 1.24(ii) on \( \mathcal{S} \): Let \( A = [a, b] \cap \mathbb{Q} \in \mathcal{S} \), \( a, b \in \mathbb{R}, a < b \), and \( (A_n)_{n \in \mathbb{N}} \) a sequence in \( \mathcal{S} \) with \( A_n \uparrow A \). Then, for each \( n \in \mathbb{N} \), \( A_n = [a_n, b_n] \cap \mathbb{Q} \), \( a_n, b_n \in \mathbb{R}, a_n < b_n \), with \( (a_n)_{n \in \mathbb{N}} \) decreasing, \( (b_n)_{n \in \mathbb{N}} \) increasing,

\[
\lim_{n \to \infty} a_n = a, \quad \lim_{n \to \infty} b_n = b.
\]

Then

\[
\lim_{n \to \infty} \mu(A_n) = \lim_{n \to \infty} (b_n - a_n) = b - a = \mu(A),
\]

showing Th. 1.24(ii) holds on \( \mathcal{S} \). Next, we show that (the extension of) \( \mu \) does not satisfy Th. 1.24(ii) on \( \mathcal{R} := \rho(\mathcal{S}) \): Let \( A := [0, 1] \cap \mathbb{Q} \) and let \( q_1, q_2, \ldots \) be an enumeration of the (countable) set \( A \), \( 0 < \epsilon < 1 \). We recursively construct a sequence \( (A_n)_{n \in \mathbb{N}} \), \( A_n = [a_n, b_n] \cap \mathbb{Q} \), of disjoint sets in \( \mathcal{S} \), and a sequence \( (B_n)_{n \in \mathbb{N}} \)
in \( \mathcal{R} \), \( B_n := \bigcup_{i=1}^{n} A_i \), by letting \( a_1 := \max\{0, q_1 - \epsilon 2^{-1}\} \), \( b_1 := q_1 \) (note \( \mu(B_1) = \mu(A_1) \leq \epsilon 2^{-1} \)) and, for \( n > 1 \), let \( N_n := \min\{k \in \mathbb{N} : q_k \notin B_{n-1}\} \), \( b_n := q_{N_n} \), \( a_n := \max\{t \in B_{n-1} : t < b_n\} \cup \{b_n - \epsilon 2^{-n}\} \), noting \( \mu(A_n) \leq \epsilon 2^{-n} \) and
\[
\mu(B_n) = \sum_{i=1}^{n} \mu(A_i) < \epsilon \sum_{i=1}^{\infty} 2^{-i} = \epsilon < 1.
\] (1.12)

We point out that \( N_n \) above is well-defined, since, if \( B_{n-1} = \bigcup_{i=1}^{n-1} A_i = A \), then \( \sum_{i=1}^{n-1} \mu(A_i) < \epsilon < 1 = \mu(A) \) would mean that \( \mu \) were not a content. Moreover, since \( B_n \uparrow A \) (note that every \( q_k \) must be in precisely one of the \( A_i \)), (1.12) shows that \( \mu \) does not satisfy Th. 1.24(ii) on \( \mathcal{R} \).

1.3.3 Lebesgue Premeasure on \( \mathbb{R}^n \)

At this point, we are in a position to perform the first step of the process described at the beginning of Sec. 1.3: We can define a premeasure on a semiring that, when extended to all Borel sets in \( \mathbb{R}^n \), will satisfy the conditions of Def. 1.1 (cf. Ex. 1.47).

**Definition 1.26.** Let \( n \in \mathbb{N} \) and recall the definition of the semiring \( \mathcal{I}^n \) on \( \mathbb{R}^n \) from Ex. 1.8(c). The function
\[
\lambda^n : \mathcal{I}^n \rightarrow \mathbb{R}_0^+, \quad \lambda^n(I) := \begin{cases} 
\prod_{j=1}^{n} (b_j - a_j) & \text{for } I = [a, b], \ a, b \in \mathbb{R}^n, \ a < b, \\
0 & \text{for } I = \emptyset,
\end{cases}
\] (1.13)
is called the \( (n\text{-dimensional}) \) *Lebesgue premeasure* on \( \mathcal{I}^n \) (or, somewhat less precisely, on \( \mathbb{R}^n \), as in the section title above).

The main objective of this section is to show that \( \lambda^n \) is, indeed, a premeasure. It turns out not to be as easy as one might have thought. We will first prove it to be a content in Prop. 1.28 and then to be a premeasure in Th. 1.31. We introduce the following notions that will be useful to prove Prop. 1.28:

**Definition 1.27.** Let \( n \in \mathbb{N} \).

(a) Given \( k \in \{1, \ldots, n\} \), \( \alpha \in \mathbb{R} \), we define the following subsets of \( \mathbb{R}^n \):
\[
H_{k,\alpha} := \{x \in \mathbb{R}^n : x_k = \alpha\}, \quad H^+_{k,\alpha} := \{x \in \mathbb{R}^n : x_k \geq \alpha\}, \quad \overline{H}^+_{k,\alpha} := \{x \in \mathbb{R}^n : x_k \leq \alpha\},
\] (1.14a)
\[
\overline{H}_{k,\alpha} := \{x \in \mathbb{R}^n : x_k = \alpha\}, \quad \overline{H}^-_{k,\alpha} := \{x \in \mathbb{R}^n : x_k \leq \alpha\},
\] (1.14b)
where \( H_{k,\alpha} \) is called the *hyperplane* and \( H^+_{k,\alpha} \) (resp. \( \overline{H}^-_{k,\alpha} \)) is called the (closed) positive (resp. negative) *halfspace* corresponding to \((k, \alpha)\).
Given a half-open interval \( I = ]a, b[ \in \mathcal{I}^n \), \( a, b \in \mathbb{R}^n \) with \( a < b \), we call the set
\[
C(I) := \{(k, \alpha) \in \{1, \ldots, n\} \times \mathbb{R} : a_k = \alpha \lor b_k = \alpha\}
\]
the coordinate set of \( I \). We also define \( C(\emptyset) := \emptyset \).

Now let \( I, I_1, \ldots, I_N \in \mathcal{I}^n \) be half-open intervals such that \( I = ]a, b[ = \bigcup_{j=1}^N I_j \), \( a, b \in \mathbb{R}^n \) with \( a < b \). Moreover, let \((k_1, \alpha_1), \ldots, (k_M, \alpha_M) \in \{1, \ldots, n\} \times \mathbb{R}, M \in \mathbb{N}\). We say that the hyperplanes \( H_{k_1, \alpha_1}, \ldots, H_{k_M, \alpha_M} \) cut the interval \( I \) into the intervals \( I_1, \ldots, I_N \) if, and only if, the following two conditions hold:
\[
\forall i \in \{1, \ldots, N\} \quad \forall j \in \{1, \ldots, M\} \quad \left( I_i \subseteq \overline{H_{k_j, \alpha_j}} \lor I_i \subseteq \overline{H_{k_j, \alpha_j}} \right), \quad (1.15a)
\]
\[
\left\{(k_j, \alpha_j) : j \in \{1, \ldots, M\}\right\} = \left\{(k, \alpha) \in \bigcup_{j=1}^N C(I_j) : a_k < \alpha < b_k\right\} \quad (1.15b)
\]
(condition \(1.15a\)) says that each small interval \( I_j \) has to lie on precisely one side of each of the cutting halfspaces; condition \(1.15b\) says that all halfspaces intersect the interior of \( I \) and that the cutting halfspaces extend all the faces of the \( I_j \) that intersect the interior of \( I \) to the boundary of \( I \).

**Proposition 1.28.** For each \( n \in \mathbb{N} \), Lebesgue premeasure \( \lambda^n \) as defined in Def. 1.26 constitutes a content on the semiring \( \mathcal{I}^n \).

**Proof.** We conduct the proof in several steps. First, we consider an interval \( I = ]a, b[ \), \( a, b \in \mathbb{R}^n \), \( a < b \), that is cut by a hyperplane \( H_{k, \alpha} \), \( k \in \{1, \ldots, n\} \) and \( \alpha \in ]a_k, b_k[ \) into intervals \( I_1, I_2 \in \mathcal{I}^n \) in the sense of Def. 1.27. Then, clearly \( I_1 \cup I_2 \), where \( I_1 = ]a^1, b^1[ \), \( I_2 = ]a^2, b^2[ \), with
\[
a^1 := a, \quad b^1 := (b_1, \ldots, b_{k-1}, \alpha, b_{k+1}, \ldots, b_n),
\]
\[
a^2 := (a_1, \ldots, a_{k-1}, \alpha, a_{k+1}, \ldots, a_n), \quad b^2 := b.
\]
Moreover,
\[
\lambda^n(I) = \prod_{j=1}^n (b_j - a_j) = \left( \prod_{j=1}^{k-1} (b_j - a_j) \right) (\alpha - a_k) \left( \prod_{j=k+1}^n (b_j - a_j) \right) + \left( \prod_{j=1}^{k-1} (b_j - a_j) \right) (b_k - \alpha) \left( \prod_{j=k+1}^n (b_j - a_j) \right) = \lambda^n(I_1) + \lambda^n(I_2).
\]
We can now use an induction on \( M \in \mathbb{N} \) to show that, if \( I \in \mathcal{I}^n \) is cut by \( M \) hyperplanes \( H_{k_1, \alpha_1}, \ldots, H_{k_M, \alpha_M} \), \( (k_1, \alpha_1), \ldots, (k_M, \alpha_M) \in \{1, \ldots, n\} \times \mathbb{R} \), into intervals \( I_1, \ldots, I_N \in \mathcal{I}^n \), then \( \lambda^n(I) = \sum_{i=1}^N \lambda^n(I_i) \): The case \( M = 1 \) has already been carried out above. Thus, let \( M \in \mathbb{N} \). By induction, the \( M \) hyperplanes \( H_{k_1, \alpha_1}, \ldots, H_{k_M, \alpha_M} \), \( (k_1, \alpha_1), \ldots, (k_M, \alpha_M) \in \{1, \ldots, n\} \times \mathbb{R} \) cut \( I \) into \( N \) intervals \( I_1, \ldots, I_N \in \mathcal{I}^n \) and \( \lambda^n(I) = \sum_{i=1}^N \lambda^n(I_i) \). Now consider \( (k_{M+1}, \alpha_{M+1}) \in \{1, \ldots, n\} \times \mathbb{R} \). Then, using the
case $M = 1$, $H_{k_{M+1}, a_{M+1}}$ cuts each $I_i$ either not at all or into two intervals $I_{(i,1)}, I_{(i,2)} \in \mathcal{I}^n$ such that $\lambda^n(I_i) = \lambda^n(I_{(i,1)}) + \lambda^n(I_{(i,2)})$. Letting
\[
J^{M+1,1} := \{ i \in \{1, \ldots, N\} : I_i \text{ not cut by } H_{k_{M+1}, a_{M+1}} \},
\]
\[
J^{M+1,2} := \{ i \in \{1, \ldots, N\} : I_i \text{ cut into } I_{(i,1)}, I_{(i,2)} \text{ by } H_{k_{M+1}, a_{M+1}} \},
\]
\[
J_i := \{ i \} \text{ for } i \in J^{M+1,1}, \quad J_i := \{ (i, 1), (i, 2) \} \text{ for } i \in J^{M+1,2},
\]
we have
\[
\lambda^n(I) = \sum_{i=1}^{N} \lambda^n(I_i) = \sum_{i=1}^{N} \sum_{k \in J_i} \lambda^n(I_k),
\]
completing the induction. Finally, given $I = [a, b] \in \mathcal{I}^n$ ($a, b \in \mathbb{R}^n$ with $a < b$), let $K \in \mathbb{N}$, and let $I_1, \ldots, I_K \in \mathcal{I}^n$ be disjoint intervals such that $I = \bigcup_{j=1}^{K} I_j$. Set
\[
J := \left\{ (k, \alpha) \in \bigcup_{j=1}^{K} C(I_j) : a_k < \alpha < b_k \right\},
\]
We cut $I$ by the $M := \# J \in \mathbb{N}$ hyperplanes in $\{H_{(k, \alpha)} : (k, \alpha) \in J\}$ (i.e. we extend all the faces of the $I_j$ that intersect the interior of $I$ to the boundary of $I$). Then we know that these hyperplanes cut $I$ into $N$ disjoint intervals $A_1, \ldots, A_N \in \mathcal{I}^n$, $N \in \mathbb{N}$, and $\lambda^n(I) = \sum_{i=1}^{N} \lambda^n(A_i)$. Moreover, we define
\[
\forall j \in \{1, \ldots, K\} \quad J_j := \{ (k, \alpha) \in J : I_j \text{ cut by } H_{(k, \alpha)} \}. \quad \text{If } J_j \neq \emptyset, \text{ then the } M_j \text{ hyperplanes in } \{H_{(k, \alpha)} : (k, \alpha) \in J_j\} \text{ (where } 1 \leq M_j := \# J_j \in \mathbb{N}) \text{ cut } I_j \text{ into } N_j \text{ disjoint intervals } A_{i_1}^j, \ldots, A_{i_{N_j}}^j \in \mathcal{I}^n, N_j \in \mathbb{N}, \text{ and } \lambda^n(I_j) = \sum_{i=1}^{N_j} \lambda^n(A_i^j). \quad \text{For } J_j = \emptyset, \text{ set } A_{i_1}^j := I_j, N_j := 1. \text{ The definition of } J \text{ yields}
\[
\forall i \in \{1, \ldots, N\} \quad \exists j \in \{1, \ldots, K\} \quad A_i \subseteq I_j
\]
and, thus,
\[
\{ A_i : i \in \{1, \ldots, N\} \} = \bigcup_{j=1}^{K} \{ A_{i}^j : i \in \{1, \ldots, N_j\} \}.
\]
In consequence, altogether, we obtain
\[
\lambda^n(I) = \sum_{i=1}^{N} \lambda^n(A_i) = \sum_{j=1}^{K} \sum_{i=1}^{N_j} \lambda^n(A_i^j) = \sum_{j=1}^{K} \lambda^n(I_j),
\]
completing the proof that $\lambda^n$ is a content on $\mathcal{I}^n$. \hfill \qed

To show that $\lambda^n$ is a premeasure, it remains to verify it is $\sigma$-additive. In the literature, one finds many different proofs. Here, we choose to use a method that makes use of $\lambda^n$ being \textit{regular} in the sense of the following Def. 1.29 with respect to the norm topology on $\mathbb{R}^n$. Regular measures on topological spaces play an important role throughout analysis and, while we will not have time in this class to study them in depth, this gives us the opportunity to at least briefly touch on the subject.
1 MEASURE THEORY

Definition 1.29. Let \((X, \mathcal{T})\) be a topological space, let \(\mathcal{S}\) be a semiring on \(X\), and let \(\mu : \mathcal{S} \to \mathbb{R}_0^+\) be a finite content. Then \(\mu\) is called inner regular (with respect to \(\mathcal{T}\)) if, and only if,

\[
\forall \epsilon \in \mathbb{R}^+ \quad \forall A \in \mathcal{S} \quad \exists \ \overline{K} \text{ compact} \land \overline{K} \subseteq A \land \mu(A) \leq \mu(K) + \epsilon.
\]

(1.16)

Proposition 1.30. Let \((X, \mathcal{T})\) be a topological space, let \(\mathcal{S}\) be a semiring on \(X\), and let \(\mu : \mathcal{S} \to \mathbb{R}_0^+\) be an inner regular content. Then the following assertions hold true:

(a) If \(\mathcal{R} := \rho(\mathcal{S})\) and \(\mu : \mathcal{R} \to \mathbb{R}_0^+\) denotes the extension of \(\mu\) to a content on \(\mathcal{R}\), then \(\mu\) is still inner regular.

(b) \(\mu\) is a premeasure.

Proof. (a): According to Prop. 1.18(b), \(\mathcal{R}\) consists of disjoint finite unions of sets from \(\mathcal{S}\). Thus, if \(\mu\) is finite on \(\mathcal{S}\), then \(\mu\) is also finite on \(\mathcal{R}\). Now let \(A_1, \ldots, A_n \in \mathcal{S}\) be disjoint, \(n \in \mathbb{N}\), \(A := \bigcup_{i=1}^n A_i\). Given \(\epsilon \in \mathbb{R}^+\), for each \(i \in \{1, \ldots, n\}\), let \(K_i \in \mathcal{S}\) be such that \(\overline{K_i}\) is compact, \(\overline{K_i} \subseteq A_i\), and \(\mu(A_i) \leq \mu(K_i) + \frac{\epsilon}{i}\). Let \(K := \bigcup_{i=1}^n K_i\). Then \(K \in \mathcal{R}\), \(\overline{K} = \bigcup_{i=1}^n \overline{K_i}\) is compact, \(\overline{K} \subseteq A\), and

\[
\mu(A) = \sum_{i=1}^n \mu(A_i) \leq \sum_{i=1}^n \left( \mu(K_i) + \frac{\epsilon}{i} \right) = \mu(K) + \epsilon,
\]

establishing \(\mu\) to be inner regular on \(\mathcal{R}\).

(b): Using (a), it suffices to show that \(\mu\) on \(\mathcal{R}\) satisfies Th. 1.24(iv). Proceeding by contraposition, we let \((A_n)_{n \in \mathbb{N}}\) be a sequence in \(\mathcal{R}\) such that \(A_1 \supseteq A_2 \supseteq \ldots\) and

\[
\epsilon := \lim_{n \to \infty} \mu(A_n) > 0,
\]

and show that this implies \(\bigcap_{n=1}^\infty A_n \neq \emptyset\). To this end, for each \(n \in \mathbb{N}\), let \(K_n \in \mathcal{R}\) be such that \(\overline{K_n}\) is compact, \(\overline{K_n} \subseteq A_n\), and \(\mu(A_n) \leq \mu(K_n) + 2^{-n} \epsilon\). Also let \(H_n := \bigcap_{i=1}^n K_n\). Then \(H_n \in \mathcal{R}\). Moreover, \(\overline{H_n} \subseteq \overline{K_n}\), i.e. \(\overline{H_n}\) is a closed subset of the compact set \(\overline{K_n}\) and, thus, itself compact. If we can show

\[
\forall n \in \mathbb{N} \quad H_n \neq \emptyset,
\]

(1.17)

then \(\overline{H_n} = \bigcap_{i=1}^n \overline{H_i} \neq \emptyset\) for each \(n \in \mathbb{N}\) and, by the finite intersection property of the compact set \(\overline{H_1}, \bigcap_{n=1}^\infty A_n \supseteq \bigcap_{n=1}^\infty \overline{H_n} \neq \emptyset\) as desired. It remains to show (1.17). We accomplish this via proving, for each \(n \in \mathbb{N}\),

\[
\mu(H_n) \geq \mu(A_n) - \epsilon(1 - 2^{-n}) \geq \epsilon - \epsilon + 2^{-n} = 2^{-n} > 0
\]

(1.18)

by induction on \(n\). For \(n = 1\), \(H_1 = K_1\) and (1.18) holds, since \(\mu(A_1) \leq \mu(K_1) + 2^{-1} \epsilon\) by the choice of \(K_1\). For the induction step, fix \(n \in \mathbb{N}\) and estimate

\[
\mu(H_{n+1}) = \mu(H_n \cap K_{n+1}) \overset{\text{Th. 1.22(c)}}{=} \mu(K_{n+1}) + \mu(H_n) - \mu(H_n \cup K_{n+1}) \overset{\text{ind.hyp}}{\geq} \mu(K_{n+1}) + \mu(A_n) - \epsilon(1 - 2^{-n}) - \mu(H_n \cup K_{n+1}).
\]

(1.19)
Due to the choice of $K_{n+1}$, we have $\mu(K_{n+1}) \geq \mu(A_{n+1}) - 2^{-(n+1)}\epsilon$ as well as $K_{n+1} \cup H_n \subseteq A_{n+1} \cup A_n = A_n$ and, thus, $\mu(K_{n+1} \cup H_n) \leq \mu(A_n)$, which we use in (1.19) to obtain

$$\mu(H_{n+1}) \geq \mu(A_{n+1}) - 2^{-(n+1)}\epsilon \geq \epsilon(1 - 2^{-n}) = \mu(A_{n+1}) - \epsilon(1 - 2^{-(n+1)}),$$

completing the induction as well as the proof of the proposition.

\[\square\]

**Theorem 1.31.** For each $n \in \mathbb{N}$, Lebesgue premeasure $\lambda^n$ as defined in Def. 1.26 constitutes a premeasure on the semiring $\mathcal{I}^n$.

**Proof.** We know $\lambda^n$ to be a content from Prop. 1.28. Thus, the assertion will follow from Prop. 1.30, if we can show $\lambda_n$ to be inner regular with respect to the norm topology on $\mathbb{R}^n$. To this end, let $a, b, c \in \mathbb{R}^n$, $a < c < b$, $A = [a, b]$, $K = [c, b]$. Then $K \subseteq A$ is compact and, letting $M_n := \max\{b_j - a_j : j \in \{1, \ldots, n\}\}$,

$$\lambda^n(A) \leq \lambda^n(K) + M_{n-1}^{n-1} \sum_{j=1}^{n} (c_j - a_j). \quad (1.20)$$

We prove (1.20) by induction on $n \in \mathbb{N}$: For $n = 1$, (1.20) follows from $\lambda^1(A) = b_1 - a_1 = b_1 - c_1 + c_1 - a_1 = \lambda^1(K) + M_1^0(c_1 - a_1)$. For $n \in \mathbb{N}$, we have

$$\lambda^{n+1}(A \times [a_{n+1}, b_{n+1}]) = (b_{n+1} - c_{n+1} + c_{n+1} - a_{n+1}) \lambda^n(A)$$

\begin{align*}
\underset{\text{ind. hyp.}}{\leq} (b_{n+1} - c_{n+1}) \left( \lambda^n(K) + M_{n-1}^{n-1} \sum_{j=1}^{n} (c_j - a_j) \right) + (c_{n+1} - a_{n+1}) \lambda^n(A) \\
\leq \lambda^{n+1}(K \times [c_{n+1}, b_{n+1}]) + M_{n+1}^{n+1} \sum_{j=1}^{n+1} (c_j - a_j) + M_{n+1}^{n+1} (c_{n+1} - a_{n+1}) \\
= \lambda^{n+1}(K \times [c_{n+1}, b_{n+1}]) + M_{n+1}^{n+1} \sum_{j=1}^{n+1} (c_j - a_j),
\end{align*}

as needed to conclude the induction. Now fix $n \in \mathbb{N}$, $\epsilon > 0$, and let $A$ and $M := M_n$ be as above. For each $j \in \{1, \ldots, n\}$, let $c_j := \min\{\frac{b_j + a_j}{2}, a_j + \frac{\epsilon}{nM^{n-1}}\}$, $K := [c, b] \times [c, b]$ (as before). Then

$$\lambda^n(A) \overset{(1.20)}{\leq} \lambda^n(K) + M_{n-1}^{n-1} \sum_{j=1}^{n} (c_j - a_j) \leq \lambda^n(K) + M_{n-1}^{n-1} n \frac{\epsilon}{nM^{n-1}} = \lambda^n(K) + \epsilon,$$

proving $\lambda^n$ to be inner regular and, thus, a premeasure.

\[\square\]

### 1.3.4 Outer Measures, Carathéodory Extension Theorem

We would now like to extend premeasures to measures. An important tool to accomplish this, are so-called outer measures.

**Definition 1.32.** Let $X$ be a set. A map $\eta : \mathcal{P}(X) \rightarrow [0, \infty]$ is called an outer measure on $X$ if, and only if, $\eta$ satisfies the following conditions (i) – (iii):

1. **Non-negativity:** $\eta(A) \geq 0$ for all $A \subseteq X$.
2. **Null set:** $\eta(\emptyset) = 0$.
3. **Monotonicity:** If $A \subseteq B \subseteq X$, then $\eta(A) \leq \eta(B)$.
4. **Subadditivity:** If $\{A_i : i \in I\}$ is a collection of subsets of $X$, then $\eta\left(\bigcup_{i \in I} A_i\right) \leq \sum_{i \in I} \eta(A_i)$.

**Theorem 1.33.** Let $\eta : \mathcal{P}(X) \rightarrow [0, \infty]$ be an outer measure on $X$. Then $\mu : \mathcal{M}(X) \rightarrow [0, \infty]$ defined by

$$\mu(A) := \inf \{\eta(B) : B \supseteq A\}$$

is a measure on $\mathcal{M}(X)$.

**Proof.** We need to verify that $\mu$ is a measure on $\mathcal{M}(X)$. We already know from the construction that $\mu(A) \geq 0$ for all $A \subseteq X$ and that $\mu(\emptyset) = 0$. We need to check subadditivity and countable additivity:

1. **Subadditivity:** Let $\{A_i : i \in I\}$ be a collection of subsets of $X$. Then

$$\mu\left(\bigcup_{i \in I} A_i\right) \leq \sum_{i \in I} \mu(A_i),$$

since $\eta(B) \geq \sum_{i \in I} \eta(A_i)$ for any $B \supseteq \bigcup_{i \in I} A_i$.

2. **Countable additivity:** Let $\{A_i : i \in \mathbb{N}\}$ be a countable collection of disjoint subsets of $X$. Then

$$\mu\left(\bigcup_{i \in \mathbb{N}} A_i\right) = \mu\left(\bigcup_{i \in \mathbb{N}} A_i\right) \leq \sum_{i \in \mathbb{N}} \mu(A_i).$$

Thus, $\mu$ is a measure on $\mathcal{M}(X)$.

\[\square\]
1 MEASURE THEORY

(i) \( \eta(\emptyset) = 0 \).

(ii) **Monotonicity:**
\[
\forall A \subseteq B \subseteq X \quad \eta(A) \leq \eta(B).
\]
\[ (1.21) \]

(iii) \( \eta \) is countably subadditive (also called \( \sigma \)-subadditive), i.e., if \( (A_i)_{i \in \mathbb{N}} \) is a sequence in \( \mathcal{P}(X) \), then
\[
\eta \left( \bigcup_{i=1}^{\infty} A_i \right) \leq \sum_{i=1}^{\infty} \eta(A_i).
\]
\[ (1.22) \]

**Remark 1.33.** Since we know measures to be monotone and \( \sigma \)-subadditive, every measure on \( \mathcal{P}(X) \) is an outer measure. However, the converse is not true in general: For example,
\[
\eta : \mathcal{P}(X) \rightarrow [0, \infty], \quad \eta(A) := \begin{cases} 0 & \text{if } A = \emptyset, \\ 1 & \text{if } A \neq \emptyset, \end{cases}
\]
defines an outer measure that is not a measure if \( X \) has more than one element.

**Definition 1.34.** Let \( X \) be a set and let \( \eta : \mathcal{P}(X) \rightarrow [0, \infty] \) be an outer measure on \( X \). We call \( A \subseteq X \) \( \eta \)-measurable if, and only if,
\[
\forall Q \subseteq X \quad \eta(Q) = \eta(Q \cap A) + \eta(Q \cap A^c).
\]
\[ (1.23) \]

**Lemma 1.35.** Let \( X \) be a set and let \( \eta : \mathcal{P}(X) \rightarrow [0, \infty] \) be an outer measure on \( X \), \( A \subseteq X \).

(a) If \( \eta(A) = 0 \) or \( \eta(A^c) = 0 \), then \( A \) is \( \eta \)-measurable.

(b) \( A \) is \( \eta \)-measurable if, and only if,
\[
\forall Q \subseteq X, \quad \eta(Q) \geq \eta(Q \cap A) + \eta(Q \cap A^c).
\]
\[ (1.23) \]

(c) \( A \) is \( \eta \)-measurable if, and only if,
\[
\forall Q \subseteq X, \quad \eta(Q) = \eta(Q \cap A) + \eta(Q \cap A^c).
\]
\[ (1.24) \]

**Proof.** (a) If \( \eta(A) = 0 \), then the monotonicity of \( \eta \) implies, for each \( Q \subseteq X \), \( \eta(Q \cap A) = 0 \) and, thus \( \eta(Q \cap A) + \eta(Q \cap A^c) = \eta(Q) \leq \eta(Q) \), showing \( A \) to be \( \eta \)-measurable. Analogously, we see \( A \) to be \( \eta \)-measurable for \( \eta(A^c) = 0 \).

(b) is clear, since (1.23) always holds for \( \eta(Q) = \infty \).

(c) is also clear since \( \leq \) always holds in (1.23) by the subadditivity of \( \eta \). \[ \square \]

**Proposition 1.36.** Let \( X \) be a set and let \( \eta : \mathcal{P}(X) \rightarrow [0, \infty] \) be an outer measure on \( X \). Define
\[
\mathcal{A}_\eta := \{ A \subseteq X : A \text{ is } \eta \text{-measurable} \}.
\]
\[ (1.25) \]

Then \( \mathcal{A}_\eta \) is a \( \sigma \)-algebra and \( \eta|_{\mathcal{A}_\eta} \) is a measure.
Proof. We first show \( A_\eta \) to be an algebra: Let \( Q \subseteq X \). Since \( \eta(Q) = \eta(Q \cap X) + \eta(Q \cap \emptyset) \), \( \emptyset \in A_\eta \) and \( X \in A_\eta \). It is immediate from (1.23) that \( A \in A_\eta \) implies \( A^c \in A_\eta \). Now let \( A, B \in A_\eta \). Then

\[
\begin{align*}
\eta(Q) &\geq \eta(Q \cap A) + \eta(Q \cap A^c) \\
&\geq \eta(Q \cap A) + \eta(Q \cap A^c \cap B) + \eta(Q \cap A^c \cap B^c) \\
&\geq \eta((Q \cap A) \cup (Q \cap A^c \cap B)) + \eta(Q \cap (A \cup B)^c) \\
&= \eta(Q \cap (A \cup B)) + \eta(Q \cap (A \cup B)^c),
\end{align*}
\]

proving \( A \cup B \in A_\eta \). Now \( A \setminus B = A \cap B^c = (A^c \cup B)^c \in A_\eta \), establishing \( A_\eta \) to be an algebra.

Now let \( (A_i)_{i \in \mathbb{N}} \) be a sequence of disjoint sets in \( A_\eta \) and let \( A := \bigcup_{i=1}^{\infty} A_i \). We claim that \( A \in A_\eta \) and

\[
\eta(A) = \sum_{i=1}^{\infty} \eta(A_i). \tag{1.27}
\]

To prove the claim, fix \( Q \in \mathcal{P}(X) \). We first show

\[
\forall n \in \mathbb{N} \quad \eta \left( Q \cap \bigcup_{i=1}^{n} A_i \right) = \sum_{i=1}^{n} \eta(Q \cap A_i) \tag{1.28}
\]

via induction on \( n \): For \( n = 1 \), there is nothing to prove. For \( n > 1 \), we apply (1.24) with \( Q \cap \bigcup_{i=1}^{n-1} A_i \) instead of \( Q \) and \( \bigcup_{i=1}^{n-1} A_i \) instead of \( A \) (we know \( \bigcup_{i=1}^{n-1} A_i \in A_\eta \), as \( A_\eta \) is an algebra) to obtain

\[
\eta \left( Q \cap \bigcup_{i=1}^{n} A_i \right) = \eta \left( Q \cap \bigcup_{i=1}^{n-1} A_i \right) + \eta(Q \cap A_n) \overset{\text{ind. hyp.}}{=} \sum_{i=1}^{n} \eta(Q \cap A_i)
\]

as needed. Now we prove \( A \in A_\eta \) and (1.27): For each \( n \in \mathbb{N} \), since \( \bigcup_{i=1}^{n} A_i \in A_\eta \), we can apply (1.28) to estimate

\[
\eta(Q) \geq \eta \left( Q \cap \bigcup_{i=1}^{n} A_i \right) + \eta \left( Q \cap \left( \bigcup_{i=1}^{n} A_i \right)^c \right) \overset{\text{1.23, Def. 1.32(ii)}}{\geq} \sum_{i=1}^{n} \eta(Q \cap A_i) + \eta(Q \cap A^c)
\]

implying

\[
\eta(Q) \geq \sum_{i=1}^{\infty} \eta(Q \cap A_i) + \eta(Q \cap A^c) \overset{\text{Def. 1.32(iii)}}{\geq} \eta(Q \cap A) + \eta(Q \cap A^c) \overset{\text{Def. 1.32(iii)}}{\geq} \eta(Q). \tag{1.29}
\]

Thus, all terms in (1.29) must be equal, proving \( A \in A_\eta \) and, for \( Q := A \), (1.27).
Due to (1.27), \(\eta|_{\mathcal{A}_\eta}\) is a measure, provided \(\mathcal{A}_\eta\) is a \(\sigma\)-algebra. So, finally, let \((A_i)_{i \in \mathbb{N}}\) be a sequence in \(\mathcal{A}_\eta\) (now the \(A_i\) do not have to be disjoint) and let \(A := \bigcup_{i=1}^{\infty} A_i\). Then
\[
A = \bigcup_{i=1}^{\infty} \left( A_i \setminus \bigcup_{k=1}^{i-1} A_k \right) \in \mathcal{A}_\eta
\]
showing \(\mathcal{A}_\eta\) to be a \(\sigma\)-algebra, completing the proof of the proposition.

The following Prop. 1.37 yields a useful construction method for outer measures:

**Proposition 1.37.** Let \(X\) be a set, \(\mathcal{S} \subseteq \mathcal{P}(X)\) with \(\emptyset \in \mathcal{S}\), and \(\mu : \mathcal{S} \to [0, \infty)\) with \(\mu(\emptyset) = 0\). Then
\[
\eta : \mathcal{P}(X) \to [0, \infty], \quad \eta(A) := \inf \left\{ \sum_{i=1}^{\infty} \mu(A_i) : (A_i)_{i \in \mathbb{N}} \text{ sequence in } \mathcal{S}, A \subseteq \bigcup_{i=1}^{\infty} A_i \right\}, \quad \inf \emptyset := \infty. \quad (1.30)
\]
defines an outer measure on \(X\). This outer measure is called the outer measure induced by \(\mu\).

**Proof.** Setting \(A_i := \emptyset\) for each \(i \in \mathbb{N}\) in (1.30) shows \(\eta(\emptyset) = 0\). If \(A \subseteq B \subseteq X\) and \((A_i)_{i \in \mathbb{N}}\) is a sequence in \(\mathcal{S}\) with \(B \subseteq \bigcup_{i=1}^{\infty} A_i\), then \(A \subseteq \bigcup_{i=1}^{\infty} A_i\), showing \(\eta(A) \leq \eta(B)\).

Only the \(\sigma\)-subadditivity of \(\eta\) requires slightly more effort: Let \((A_i)_{i \in \mathbb{N}}\) be a sequence in \(\mathcal{P}(X)\) and \(A := \bigcup_{i=1}^{\infty} A_i\). We need to show
\[
\eta(A) \leq \sum_{i=1}^{\infty} \eta(A_i). \quad (1.31)
\]
Since (1.31) trivially holds if, for at least one \(i \in \mathbb{N}\), \(\eta(A_i) = \infty\), we may assume \(\eta(A_i) < \infty\) for each \(i \in \mathbb{N}\). Let \(\epsilon \in \mathbb{R}^{+}\). Then, for each \(i \in \mathbb{N}\), there exists a sequence \((B_{ik})_{k \in \mathbb{N}}\) in \(\mathcal{S}\) such that
\[
A_i \subseteq \bigcup_{k=1}^{\infty} B_{ik} \quad \land \quad \sum_{k=1}^{\infty} \mu(B_{ik}) < \eta(A_i) + \epsilon \cdot 2^{-i}.
\]
Since \((B_{ik})_{(i,k) \in \mathbb{N}^2}\) is a countable family (i.e. a sequence) in \(\mathcal{S}\) with \(A \subseteq \bigcup_{(i,k) \in \mathbb{N}^2} B_{ik}\),
\[
\eta(A) \leq \sum_{i=1}^{\infty} \sum_{k=1}^{\infty} \mu(B_{ik}) \leq \sum_{i=1}^{\infty} \left( \eta(A_i) + \epsilon \cdot 2^{-i} \right) = \epsilon + \sum_{i=1}^{\infty} \eta(A_i)
\]
implies (1.31) and establishes the case.

**Theorem 1.38** (Carathéodory Extension Theorem). Let \(\mathcal{S}\) be a semiring on the set \(X\) and let \(\mu : \mathcal{S} \to [0, \infty]\) be a content. If \(\eta\) is defined by (1.30), then the following holds:
(a) \( \eta \) is an outer measure and each \( A \in S \) is \( \eta \)-measurable (i.e. \( S \subseteq \sigma(S) \subseteq A_\eta \)).

(b) If \( \mu \) is a premeasure, then \( \eta|_S = \mu \) (thus, in this case, \( \eta|_{A_\eta} \) is a measure that extends the premeasure \( \mu \) to a \( \sigma \)-algebra containing \( \sigma(S) \)).

(c) If \( \mu \) is not a premeasure, then

\[
\exists_{A \in S} \eta(A) < \mu(A). \tag{1.32}
\]

Proof. (a): According to Prop. 1.37, \( \eta \) is an outer measure. It remains to show that each \( A \in S \) is \( \eta \)-measurable, i.e. we have to verify (1.23), given some \( A \in S \). It will be convenient to extend the content \( \mu \) to a content on \( R := \rho(S) \) via Prop. 1.21 (we still call the extension \( \mu \)). Then (1.30) and Prop. 1.18(b) imply

\[
\forall_{Q \in \mathcal{P}(X)} \eta(Q) = \inf \left\{ \sum_{i=1}^{\infty} \mu(A_i) : (A_i)_{i \in \mathbb{N}} \text{ sequence in } R, Q \subseteq \bigcup_{i=1}^{\infty} A_i \right\}. \tag{1.33}
\]

To verify (1.23), let \( Q \subseteq X \). As (1.23) always holds if \( \eta(Q) = \infty \), we assume \( \eta(Q) < \infty \). For each sequence \((A_i)_{i \in \mathbb{N}}\) in \( S \) with \( Q \subseteq \bigcup_{i=1}^{\infty} A_i \) (where \( \eta(Q) < \infty \) guarantees such a sequence exists), the following must hold (using that \( \mu \) is a content on \( R \)):

\[
\sum_{i=1}^{\infty} \mu(A_i) = \sum_{i=1}^{\infty} \mu(A_i \cap A) + \sum_{i=1}^{\infty} \mu(A_i \setminus A) \geq \eta(Q \cap A) + \eta(Q \cap A^c). \tag{1.34}
\]

Now taking the infimum in (1.34) over all admissible sequences \((A_i)_{i \in \mathbb{N}}\) in \( S \) according to (1.30), we obtain \( \eta(Q) \geq \eta(Q \cap A) + \eta(Q \cap A^c) \), proving \( S \subseteq A_\eta \) as claimed.

(b): The inequality \( \eta|_S \leq \mu \) is immediate from (1.30). It remains to show \( \mu \leq \eta|_S \).

We know from Prop. 1.21 that, if \( \mu \) is a premeasure on \( S \), then the extended \( \mu \) is a premeasure on \( R \). Thus, if \((A_i)_{i \in \mathbb{N}}\) is a sequence in \( R \) such that \( Q \subseteq \bigcup_{i=1}^{\infty} A_i \), \( Q \in S \), then, by Th. 1.22(f), \( \mu(Q) \leq \sum_{i=1}^{\infty} \mu(A_i) \). Taking the infimum over all admissible sequences according to (1.33) yields \( \mu(Q) \leq \eta(Q) \) and \( \mu \leq \eta|_S \) as needed.

(c): If \( \mu \) is not a premeasure on \( S \), then there exists \( A \in S \) and a sequence \((A_i)_{i \in \mathbb{N}}\) of disjoint sets in \( S \) such that \( A = \bigcup_{i=1}^{\infty} A_i \) and \( \mu(A) \neq \sum_{i=1}^{\infty} \mu(A_i) \). Due to Th. 1.22(d) (applied to \( \mu \) on \( R \)), \( \sum_{i=1}^{\infty} \mu(A_i) \leq \mu(A) \). Thus,

\[
\eta(A) \leq \sum_{i=1}^{\infty} \mu(A_i) < \mu(A),
\]

which establishes the case.

\[\blacksquare\]

1.3.5 Dynkin Systems, Uniqueness of Extensions

The following simple Ex. 1.39 shows that, in general, one can not expect the extension of a premeasure \( \mu \) on a semiring \( S \) to a measure on \( \sigma(S) \) to be unique:
Example 1.39. Let $X$ be a nonempty set and $\mathcal{R} := \{\emptyset\}$. Then $\mathcal{R}$ is a semiring (and even a ring) on $X$, and $\mu : \mathcal{R} \rightarrow [0, \infty]$, $\mu(\emptyset) := 0$, defines a premeasure. Then $A := \sigma(\mathcal{R}) = \{\emptyset, X\}$ and

$$\forall \alpha \in [0, \infty] \quad \mu_\alpha : A \rightarrow [0, \infty], \quad \mu_\alpha(A) := \begin{cases} 0 & \text{for } A = \emptyset, \\ \alpha & \text{for } A = X, \end{cases}$$

constitutes an extension of $\mu$ to a measure on $A$.

The problem in Ex. 1.39, as it turns out, is that the premeasure $\mu$ on the semiring $\mathcal{S}$ is not $\sigma$-finite. In preparation for the proof of the uniqueness Th. 1.45, we develop some additional technical tools that are also of measure-theoretic use beyond the application in the present section.

Definition 1.40. Let $X$ be a set, $\mathcal{E} \subseteq \mathcal{P}(X)$.

(a) $\mathcal{E}$ is called $\cap$-stable if, and only if,

$$\forall A, B \in \mathcal{E} \quad A \cap B \in \mathcal{E}.$$

(b) $\mathcal{E}$ is called a monotone class on $X$ if, and only if, it satisfies the following two conditions:

(i) If $A \subseteq X$ and $(E_n)_{n \in \mathbb{N}}$ is a sequence in $\mathcal{E}$ with $E_n \uparrow A$, then $A \in \mathcal{E}$.

(ii) If $A \subseteq X$ and $(E_n)_{n \in \mathbb{N}}$ is a sequence in $\mathcal{E}$ with $E_n \downarrow A$, then $A \in \mathcal{E}$.

(c) $\mathcal{E}$ is called a Dynkin system on $X$ if, and only if, it satisfies the following three conditions:

(i) $X \in \mathcal{E}$.

(ii) If $A \in \mathcal{E}$, then $A^c \in \mathcal{E}$.

(iii) If $(E_n)_{n \in \mathbb{N}}$ is a sequence of disjoint sets in $\mathcal{E}$, then $\bigcup_{n \in \mathbb{N}} E_n \in \mathcal{E}$.

Remark and Definition 1.41. Arbitrary intersections of monotone classes yield again monotone classes, arbitrary intersections of Dynkin systems yield again Dynkin systems: Let $X$ be a set, let $I$ be a nonempty index set, and let $(\mathcal{M}_i)_{i \in I}$ be a family of monotone classes on $X$, $\mathcal{M} := \bigcap_{i \in I} \mathcal{M}_i$. If $A \subseteq X$ and $(E_n)_{n \in \mathbb{N}}$ is a sequence in $\mathcal{M}$ with $E_n \uparrow A$ (resp. $E_n \downarrow A$), then, for each $i \in I$, $(E_n)_{n \in \mathbb{N}}$ is a sequence in $\mathcal{M}_i$, i.e. $A \in \mathcal{M}_i$ for each $i \in I$, proving $\mathcal{M}$ to be a monotone class on $X$. Now let $(\mathcal{D}_i)_{i \in I}$ be a family of Dynkin systems on $X$, $\mathcal{D} := \bigcap_{i \in I} \mathcal{D}_i$. Since $X$ is in each $\mathcal{D}_i$, it is also in $\mathcal{D}$. If $A \in \mathcal{D}$, then $A \in \mathcal{D}_i$ for each $i \in I$, i.e. $A^c \in \mathcal{D}_i$ for each $i \in I$, implying $A^c \in \mathcal{D}$. If $(E_n)_{n \in \mathbb{N}}$ is a sequence of disjoint sets in $\mathcal{D}$ then, for each $i \in I$, $(E_n)_{n \in \mathbb{N}}$ is a sequence of disjoint sets in $\mathcal{D}_i$. Thus, $\bigcup_{n \in \mathbb{N}} E_n \in \mathcal{D}$, showing $\mathcal{D}$ is a Dynkin system on $X$. Thus, if $X$ is a set and $\mathcal{E} \subseteq \mathcal{P}(X)$, then we can let $\mu(\mathcal{E}) := \mu_X(\mathcal{E})$ denote the intersection of all
monotone classes on $X$ that are supersets of $\mathcal{E}$, and we can let $\delta(\mathcal{E}) := \delta_X(\mathcal{E})$ denote the intersection of all Dynkin systems on $X$ that are supersets of $\mathcal{E}$. We then know $\mu(\mathcal{E})$ to be a monotone class and $\delta(\mathcal{E})$ to be a Dynkin system, namely the smallest monotone class (resp. Dynkin system) on $X$ containing $\mathcal{E}$. Thus, we call $\mu(\mathcal{E})$ the monotone class (and $\delta(\mathcal{E})$ the Dynkin system) generated by $\mathcal{E}$.

**Proposition 1.42.** Let $X$ be a set, $\mathcal{E} \subseteq \mathcal{P}(X)$. Then $\mathcal{E}$ is a Dynkin system on $X$ if, and only if, the following conditions (i) – (iii) hold:

(i) $X \in \mathcal{E}$.

(ii) For each $A, B \in \mathcal{E}$, $B \subseteq A$ implies $A \setminus B \in \mathcal{E}$.

(iii) $\mathcal{E}$ is a monotone class.

**Proof.** Assume (i) – (iii) hold. We have to show $\mathcal{E}$ is closed under complements and under countable unions of disjoint sets. If $B \in \mathcal{E}$, then, choosing $A := X$ in (ii) shows $B^c = X \setminus B \in \mathcal{E}$. If $(E_i)_{i \in \mathbb{N}}$ is a sequence of disjoint sets in $\mathcal{E}$, then $E_1 \cup E_2 = (E_1 \setminus E_2)^c \in \mathcal{E}$ (if $E_1, E_2$ are disjoint, then $E_2 \subseteq E_1^c$ and $E_1^c \setminus E_2 \in \mathcal{E}$ by (ii)). Now an induction over $n \in \mathbb{N}$ shows $A_n := \bigcup_{i=1}^n E_i \in \mathcal{E}$ for each $n \in \mathbb{N}$. Since $A_n \uparrow A := \bigcup_{i=1}^\infty E_i$, (iii) implies $A \in \mathcal{E}$, showing $\mathcal{E}$ to be a Dynkin system.

Conversely, assume $\mathcal{E}$ to be a Dynkin system. We have to show that (ii) and (iii) hold. If $A, B \in \mathcal{E}$ with $B \subseteq A$, then $A^c, B, \emptyset$ are disjoint sets in $\mathcal{E}$, yielding

$$C := A^c \cup B \cup \emptyset \cup \cdots \in \mathcal{E}, \quad A \setminus B = C^c \in \mathcal{E},$$

showing $\mathcal{E}$ satisfies (ii). Now let $A \subseteq X$ and let $(E_i)_{i \in \mathbb{N}}$ be a sequence of sets in $\mathcal{E}$. If $E_i \uparrow A$, then $A = E_1 \cup \bigcup_{i=2}^\infty (E_i \setminus E_{i-1}) \in \mathcal{E}$. If $E_i \downarrow A$, then $A^c = \bigcup_{i=1}^\infty E_i^c$, i.e. $E_i^c \uparrow A^c \in \mathcal{E}$. Thus, $A \in \mathcal{E}$ as well and $\mathcal{E}$ is a monotone class, which completes the proof. 

**Proposition 1.43.** Let $X$ be a set and let $\mathcal{E} \subseteq \mathcal{P}(X)$ be a Dynkin system on $X$. Then the following statements are equivalent:

(i) $\mathcal{E}$ is $\cap$-stable.

(ii) $\mathcal{E}$ is a $\sigma$-algebra.

**Proof.** As every $\sigma$-algebra is $\cap$-stable, it only remains to show (i) implies (ii). If $A, B \in \mathcal{E}$ and $\mathcal{E}$ is $\cap$-stable, then $A \cup B = (A \cap B^c) \cup (B \cap A^c) \cup (A \cap B) \in \mathcal{E}$. Let $(E_i)_{i \in \mathbb{N}}$ be a sequence of sets in $\mathcal{E}$. An induction over $n \in \mathbb{N}$ shows $A_n := \bigcup_{i=1}^n E_i \in \mathcal{E}$ for each $n \in \mathbb{N}$. Since $A_n \uparrow A := \bigcup_{i=1}^\infty E_i$ and the Dynkin system $\mathcal{E}$ is a monotone class, $A \in \mathcal{E}$, showing $\mathcal{E}$ to be a $\sigma$-algebra.

**Proposition 1.44.** Let $X$ be a set and let $\mathcal{E} \subseteq \mathcal{P}(X)$ be $\cap$-stable. Then $\delta(\mathcal{E}) = \sigma(\mathcal{E})$. 

Proof. Since every \(\sigma\)-algebra is a Dynkin system, we have \(\mathcal{D} := \delta(\mathcal{E}) \subseteq \sigma(\mathcal{E})\). So it only remains to show that \(\mathcal{D}\) is a \(\sigma\)-algebra. Moreover, according to Prop. 1.43, it suffices to show \(\mathcal{D}\) is \(\cap\)-stable. We will achieve this by showing
\[
\forall D \in \mathcal{D} \quad D \subseteq \mathcal{D}(D) := \{A \subseteq X : A \cap D \in D\}. \tag{1.35}
\]

While this strategy might seem strange at first, in various variants, it turns out to be quite powerful in measure theory, and we will use it again in the proof of Th. 1.45 below. We note that, for each \(D \in \mathcal{D}\), \(\mathcal{D}(D)\) is a Dynkin system:
\[
X \cap D = D \in \mathcal{D}. \quad \text{If } A \in \mathcal{D}(D), \text{ then } A^c \in \mathcal{D}(D), \quad \text{as } A^c \cap D = D \setminus (A \cap D) \in \mathcal{D} \text{ by Prop. 1.42(ii).}
\]

If \((A_i)_{i \in \mathbb{N}}\) is a sequence of disjoint sets in \(\mathcal{D}(D)\), then
\[
A := \bigcup_{i=1}^{\infty} A_i \in \mathcal{D}(D),
\]

since
\[
A \cap D = \bigcup_{i=1}^{\infty} (A_i \cap D) \in \mathcal{D}.
\]

Since \(\mathcal{E}\) is \(\cap\)-stable, we have \(\mathcal{E} \subseteq \mathcal{D}(E)\) for each \(E \in \mathcal{E}\). Thus, \(\mathcal{D} = \delta(\mathcal{E}) \subseteq \sigma(\mathcal{E})\) for each \(E \in \mathcal{E}\), since \(\mathcal{D}(E)\) is a Dynkin system. Thus, if \(E \in \mathcal{E}\) and \(D \in \mathcal{D}\), then \(D \in \mathcal{D}(E)\), i.e. \(D \cap E \in \mathcal{D}\). But this means also \(E \in \mathcal{D}(D)\), showing \(\mathcal{E} \subseteq \mathcal{D}(D)\), which proves (1.35) and the proposition.

**Theorem 1.45.** Let \((X, \mathcal{A})\) be a measurable space and let \(\mu, \nu : \mathcal{A} \to [0, \infty]\) be measures. Moreover, let \(\mathcal{E} \subseteq \mathcal{P}(X)\) have the following properties:

(i) \(\sigma(\mathcal{E}) = \mathcal{A}\).

(ii) \(\mu|_\mathcal{E} = \nu|_\mathcal{E}\).

(iii) \(\mathcal{E}\) is \(\cap\)-stable.

(iv) There exists a sequence \((E_i)_{i \in \mathbb{N}}\) of sets in \(\mathcal{E}\) such that
\[
\left(\forall i \in \mathbb{N} \quad \mu(E_i) = \nu(E_i) < \infty\right) \quad \text{and} \quad X = \bigcup_{i=1}^{\infty} E_i.
\]

Then \(\mu = \nu\).

Proof. For each \(E \in \mathcal{E}\) with \(\mu(E) = \nu(E) < \infty\), we show
\[
\mathcal{A} = \mathcal{D}(E) := \{A \in \mathcal{A} : \mu(A \cap E) = \nu(A \cap E)\}. \tag{1.36}
\]

First, we note that, for each \(E \in \mathcal{E}\) with \(\mu(E) = \nu(E) < \infty\), \(\mathcal{D}(E)\) is a Dynkin system: \(X \in \mathcal{D}(E)\), since \(X \cap E = E \in \mathcal{E}\). If \(A \in \mathcal{D}(E)\), then \(A^c \in \mathcal{D}(E)\), since
\[
\mu(A^c \cap E) = \mu(E \setminus (A \cap E)) = \mu(E) - \mu(A \cap E) = \nu(E) - \nu(A \cap E) = \nu(E \setminus (A \cap E)) = \nu(A^c \cap E).
\]
If \((A_i)_{i \in \mathbb{N}}\) is a sequence of disjoint sets in \(\mathcal{D}(E)\), then \(A := \bigcup_{i=1}^{\infty} A_i \in \mathcal{D}(E)\), since
\[
\mu(A \cap E) = \mu \left( \bigcup_{i=1}^{\infty} (A_i \cap E) \right) = \sum_{i=1}^{\infty} \mu(A_i \cap E) = \sum_{i=1}^{\infty} \nu(A_i \cap E) = \nu \left( \bigcup_{i=1}^{\infty} (A_i \cap E) \right) = \nu(A \cap E).
\]
Moreover, since (iii) implies \(\mathcal{E} \subseteq \mathcal{D}(E)\), one obtains
\[
\mathcal{A} = \sigma(\mathcal{E}) \overset{(iii), \text{Prop. } 1.44}{=} \delta(\mathcal{E}) \subseteq \mathcal{D}(E),
\]
proving (1.36). Let \((E_i)_{i \in \mathbb{N}}\) be as in (iv) and define, \(F_0 := \emptyset\) as well as, for each \(n \in \mathbb{N}\), \(F_n := \bigcup_{i=1}^{n} E_i\). Then \(F_n \uparrow X\). For each \(n \in \mathbb{N}\), we have \(F_n = \bigcup_{i=1}^{n} (E_i \cap F_i^c)\), implying
\[
\forall_{n \in \mathbb{N}} \forall_{A \in \mathcal{A}} \mu(A \cap F_n) = \sum_{i=1}^{n} \mu(A \cap F_i^c \cap E_i) = \sum_{i=1}^{n} \nu(A \cap F_i^c \cap E_i) = \nu(A \cap F_n)
\]
and
\[
\forall_{A \in \mathcal{A}} \mu(A) = \lim_{n \to \infty} \mu(A \cap F_n) = \lim_{n \to \infty} \nu(A \cap F_n) = \nu(A),
\]
which proves the theorem.

**Corollary 1.46.** Let \(X\) be a set and let \(\mathcal{S}\) be a semiring on \(X\).

(a) If \(\mu : \mathcal{S} \to [0, \infty]\) is a \(\sigma\)-finite premeasure, then there exists a unique measure \(\nu : \sigma(\mathcal{S}) \to [0, \infty]\) that extends \(\mu\).

(b) One can strengthen (a) to the following comparison result: If \(\alpha, \beta : \sigma(\mathcal{S}) \to [0, \infty]\) are measures that satisfy \(\alpha|_{\mathcal{S}} \leq \beta|_{\mathcal{S}}\), where \(\beta|_{\mathcal{S}}\) is \(\sigma\)-finite, then \(\alpha \leq \beta\) holds on all of \(\sigma(\mathcal{S})\).

**Proof.** (a): Existence was shown in Th. 1.38(b), whereas uniqueness is due Th. 1.45, which applies with \(\mathcal{E} := \mathcal{S}\), since the semiring \(\mathcal{S}\) is \(\cap\)-stable.

(b): Since \(\beta|_{\mathcal{S}}\) is \(\sigma\)-finite and \(\alpha|_{\mathcal{S}} \leq \beta|_{\mathcal{S}}\), \(\alpha|_{\mathcal{S}}\) must be \(\sigma\)-finite as well. Let \(\alpha^*, \beta^* : \mathcal{P}(X) \to [0, \infty]\) be the outer measures induced by \(\alpha|_{\mathcal{S}}\) and \(\beta|_{\mathcal{S}}\), respectively (as defined in Prop. 1.37). Then the Carathéodory Extension Th. 1.38 combined with (a) yields \(\alpha = \alpha^*|_{\sigma(\mathcal{S})}\) and \(\beta = \beta^*|_{\sigma(\mathcal{S})}\). In consequence, \(\alpha|_{\mathcal{S}} \leq \beta|_{\mathcal{S}}\) combined with (1.30) proves \(\alpha \leq \beta\).

**Example 1.47.** Let \(n \in \mathbb{N}\). From Th. 1.31, we know the Lebesgue premeasure \(\lambda^n\) to be a premeasure on the semiring \(\mathcal{T}^n\) on \(\mathbb{R}^n\). Clearly, \(\lambda^n\) is also \(\sigma\)-finite. Thus, according to Cor. 1.46(a), \(\lambda^n\) can be uniquely extended to a measure \(\beta^n\) on \(\mathcal{B}^n = \sigma(\mathcal{T}^n)\), called Lebesgue-Borel measure, where \(\mathcal{B}^n\) denotes the \(\sigma\)-algebra of Borel sets in \(\mathbb{R}^n\) (see Ex. 1.8(c)). According to the Carathéodory Extension Th. 1.38, one obtains \(\beta^n\) by first extending \(\lambda^n\) to the induced outer measure \(\lambda^n : \mathcal{P}(\mathbb{R}^n) \to [0, \infty]\) (also called the Lebesgue outer measure) and restricting it, first to \(\mathcal{L}^n := \mathcal{A}_{\lambda^n}\) and then to \(\mathcal{B}^n\). The
measure $\lambda^n : \mathcal{L}^n \to [0, \infty]$ is called Lebesgue measure. The notation $\beta^n$ will be used when it is to be emphasized that $\lambda^n$ is considered only on $\mathcal{B}^n$. The sets in $\mathcal{B}^n$ are called $\beta^n$-measurable or Borel-measurable, the sets in $\mathcal{L}^n$ are called $\lambda^n$-measurable or Lebesgue-measurable. According to Cor. 1.52 below, Lebesgue measure is the so-called completion of Lebesgue-Borel measure (see Def. 1.48 and Def. 1.50) and one can show that $\mathcal{B}^n \subsetneq \mathcal{L}^n$, where $\mathcal{B}^n$ has the same cardinality as $\mathbb{R}$ and $\mathcal{L}^n$ has the same cardinality as $\mathcal{P}(\mathbb{R})$ (see [Els07, Sec. 2.8.3] and the last paragraph of Sec. 1.5.3 below). We will reencounter the important measures $\beta^n$ and $\lambda^n$ throughout this class. In particular, they will be seen to satisfy Def. 1.1, except that they are not defined on all of $\mathcal{P}(\mathbb{R}^n)$.

1.3.6 Completion

Sets of measure zero play a special role in measure theory, giving rise to the following definitions.

**Definition 1.48.** Let $(X, \mathcal{A}, \mu)$ be a measure space.

(a) A set $N \in \mathcal{A}$ is called a $\mu$-null set (or just a null set if $\mu$ is understood) if, and only if, $\mu(N) = 0$.

(b) $(X, \mathcal{A}, \mu)$ and $\mu$ are called complete if, and only if, each subset of a $\mu$-null set is $\mu$-measurable, i.e. if, and only if,

$$\forall A \in \mathcal{A} \quad \forall B \subseteq A \quad \left( \mu(A) = 0 \Rightarrow B \in \mathcal{A} \right).$$

If $(X, \mathcal{A}, \mu)$ is a measure space, then $\mu$ can be used as the content of the Carathéodory Extension Th. 1.38, which yields the extension $\eta|_{\mathcal{A}_0}$. According to Lem. 1.35(a), $\eta|_{\mathcal{A}_0}$ is complete. Thus, every measure can be extended to a complete measure. This can also be seen, using a simpler and more natural construction that does not make use of outer measures, according to the following Th. 1.49. If $\mu$ is $\sigma$-finite, then both constructions yield the same complete extension (the so-called completion of $\mu$), see Th. 1.51 below.

**Theorem 1.49.** Let $(X, \mathcal{A}, \mu)$ be a measure space. Define $\mathcal{N}$ to be the set consisting of all subsets of $\mu$-null sets, i.e.

$$\mathcal{N} := \left\{ N \in \mathcal{P}(X) : \exists A \in \mathcal{A} \left( N \subseteq A \land \mu(A) = 0 \right) \right\}. \quad (1.37a)$$

Also define

$$\hat{A} := \{ A \cup N : A \in \mathcal{A}, N \in \mathcal{N} \}, \quad (1.37b)$$

$$\hat{\mu} : \hat{A} \to [0, \infty], \quad \hat{\mu}(A \cup N) := \mu(A) \quad \text{for } A \in \mathcal{A}, N \in \mathcal{N}. \quad (1.37c)$$

Then the following holds:
Each complete measure \( \tilde{\mu} \) exists \( (a) \): We verify \( \tilde{\mu} \) and subadditivity of \( \mu \), \( (b) \): Let \(( \tilde{\mu}, \tilde{\rho} )\) be a content on \( \tilde{\rho} \) to be well-defined: If \( \tilde{\rho} \) is \( \mu \)-additive and, thus, a measure. To see that \( \tilde{\mu} \) is well-defined by \( (1.37c) \) proves \( \tilde{\mu} \) to be an extension of \( \mu \). To verify \( \sigma \)-additivity of \( \tilde{\mu} \), let \(( A_k )_{k \in \mathbb{N}} \) be a sequence of disjoint sets in \( \mathcal{A} \) and \( ( N_k )_{k \in \mathbb{N}} \) a sequence of disjoint sets in \( \mathcal{N} \). Then

\[
\tilde{\mu} \left( \bigcup_{k=1}^{\infty} ( A_k \cup N_k ) \right) = \tilde{\mu} \left( \left( \bigcup_{k=1}^{\infty} A_k \right) \cup \left( \bigcup_{k=1}^{\infty} N_k \right) \right) = \mu \left( \bigcup_{k=1}^{\infty} A_k \right) = \sum_{k=1}^{\infty} \mu(A_k) = \sum_{k=1}^{\infty} \mu(A_k \cup N_k),
\]

i.e. \( \tilde{\mu} \) is \( \sigma \)-additive and, thus, a measure. To see that \( \tilde{\mu} \) is complete, let \( C \subseteq A \cup N \) with \( A \in \mathcal{A}, N \in \mathcal{N}, \mu(A) = 0 \). Then there exists \( B \in \mathcal{A} \) such that \( \mu(B) = 0 \). Then \( C \subseteq A \cup B \in \mathcal{A} \) with \( \mu(A \cup B) = 0 \). Thus, \( C \in \mathcal{N} \subseteq \mathcal{A} \), proving \( \tilde{\mu} \) to be complete. Let \( \rho \) be a content on \( \tilde{\rho} \) that is an extension of \( \mu \). Then \( \rho(A) = \mu(A) \) for each \( A \in \mathcal{A} \) and, thus, \( \rho(N) = 0 \) for each \( N \in \mathcal{N} \). Thus, if \( \mathcal{A} \subseteq \mathcal{N} \subseteq \mathcal{A} \), then, by monotonicity and subadditivity of \( \rho \), \( \mu(A) = \rho(A) \leq \rho(A \cup N) \leq \rho(A) \cup \rho(N) = \mu(A) \), showing \( \rho(A \cup N) = \mu(A) \) and \( \rho = \tilde{\mu} \) as claimed.

(b): Let \(( X, \mathcal{B}, \nu \) a complete measure space such that \( \mathcal{A} \subseteq \mathcal{B} \) and \( \nu \upharpoonright \mathcal{A} = \mu \). As \( \nu \) is complete, one has \( \mathcal{N} \subseteq \mathcal{B} \). As \( \nu \) extends \( \mu \), one has \( \nu(N) = 0 \) for each \( N \in \mathcal{N} \). Then \( \tilde{\nu} \subseteq \mathcal{B} \) and \( \nu \upharpoonright \mathcal{A} = \tilde{\nu} \) follows from the last part of (a).

**Definition 1.50.** Let \(( X, \mathcal{A}, \mu \) be a measure space. Then the measure space \(( \tilde{\mathcal{A}}, \tilde{\mu} \) where \( \tilde{\mathcal{A}} \) and \( \tilde{\mu} \) are defined as in \( (1.37) \), is called the **completion** of \(( X, \mathcal{A}, \mu \) ; \( \tilde{\mu} \) is called the **completion** of \( \mu \). According to Th. 1.49(a),(b), the completion of \( \mu \) is the smallest complete measure that extends \( \mu \).

**Theorem 1.51.** Let \( \mathcal{S} \) be a semiring on the set \( X \), let \( \mu : \mathcal{S} \to [0, \infty] \) be a \( \sigma \)-finite premeasure and \( \eta : \mathcal{P}(X) \to [0, \infty] \) the induced outer measure defined by \( (1.30) \). Then
\( \eta \mid _{A_\eta} \) is the completion of \( \eta \mid _{\sigma(S)} \) and, in particular, this is the unique extension of \( \mu \) to a measure on \( A_\eta \).

**Proof.** As we know \( \eta \mid _{A_\eta} \) to be complete, it remains to prove \( A_\eta \subseteq \overline{\sigma(S)} \). To this end, let \( B \in A_\eta \). First, we assume \( \eta(B) < \infty \). Then, for each \( n \in \mathbb{N} \), there exists a sequence \( (A_{nk})_{k \in \mathbb{N}} \) of sets in \( S \) such that \( B \subseteq \bigcup_{k=1}^{\infty} A_{nk} \) and \( \sum_{k=1}^{\infty} \mu(A_{nk}) \leq \eta(B) + \frac{1}{n} \). Defining

\[
A := \bigcap_{n=1}^{\infty} \bigcup_{k=1}^{\infty} A_{nk},
\]

we obtain

\[
A \in \sigma(S), \quad B \subseteq A, \quad \forall n \in \mathbb{N} \quad \eta(A) \leq \eta(B) + \frac{1}{n}.
\]

Thus, \( \eta(A) = \eta(B) \). We can now apply what we have just shown to \( A \setminus B \) instead of \( B \), proving the existence of \( C \in \sigma(S) \) such that \( A \setminus B \subseteq C \) and \( \eta(C) = \eta(A \setminus B) = \eta(A) - \eta(B) = 0 \). Then \( B = (A \setminus C) \cup (B \cap C) \in \sigma(S) \), since \( A \setminus C \in \sigma(S) \) and \( B \cap C \) is a subset of the \( (\eta \mid _{\sigma(S)}) \)-null set \( C \). It remains to consider the case \( B \in A_\eta \) with \( \eta(B) = \infty \). Since \( \mu \) is assumed to be \( \sigma \)-finite, there exists a sequence \( (E_n)_{n \in \mathbb{N}} \) in \( S \) such that \( X = \bigcup_{n=1}^{\infty} E_n \) and \( \mu(E_n) < \infty \) for each \( n \in \mathbb{N} \). Since, for each \( n \in \mathbb{N} \), \( \eta(B \cap E_n) < \infty \) implies \( B \cap E_n \in \sigma(S) \), we obtain \( B = \bigcup_{n=1}^{\infty} (B \cap E_n) \in \sigma(S) \) as needed. According to Th. 1.49(a), \( \eta \mid _{A_\eta} \) is the only measure on \( A_\eta \) that extends \( \eta \mid _{\sigma(S)} \). Since, by Cor. 1.46(a), \( \eta \mid _{\sigma(S)} \) is the only measure on \( \sigma(S) \) that extends \( \mu \), \( \eta \mid _{A_\eta} \) is the only measure on \( A_\eta \) that extends \( \mu \).  

**Corollary 1.52.** Let \( n \in \mathbb{N} \). Then Lebesgue measure \( \lambda^n : \mathcal{L}^n \rightarrow [0, \infty] \) is the completion of Lebesgue-Borel measure \( \beta^n : \mathcal{B}^n \rightarrow [0, \infty] \). Moreover, Lebesgue measure is the unique extension of the premeasure \( \lambda^n : \mathcal{T}^n \rightarrow [0, \infty] \) to a measure on \( \mathcal{L}^n \).

**Proof.** Since \( \lambda^n : \mathcal{T}^n \rightarrow [0, \infty] \) is a \( \sigma \)-finite premeasure, everything is immediate from Th. 1.51.

### 1.4 Inverse Image, Trace

We investigate the behavior of \( \sigma \)-algebras when forming inverse images and subsets.

**Notation 1.53.** Let \( X, Y \) be sets \( f : X \rightarrow Y \). For \( E \subseteq \mathcal{P}(Y) \), we introduce the notation

\[
f^{-1}(E) := \{ f^{-1}(E) : E \in E \} \subseteq \mathcal{P}(X).
\]

**Definition and Remark 1.54.** Let \( X \) be a set, and let \( E \subseteq \mathcal{P}(X) \) be a collection of subsets. If \( B \subseteq X \), then let \( E|B := \{ A \cap B : A \in E \} \) denote the trace of \( E \) on \( B \), also known as the restriction of \( E \) to \( B \). Consider the canonical inclusion map \( f : B \rightarrow X \), \( f(a) = a \). Then,

\[
f^{-1}(E) = \{ f^{-1}(A) : A \in E \} = \{ A \cap B : A \in E \} = E|B.
\]
Proposition 1.55. Consider sets $X, Y$ and a map $f : X \to Y$.

(a) If $\mathcal{A}$ is a $\sigma$-algebra on $X$, then $\mathcal{B} := \{ B \subseteq Y : f^{-1}(B) \in \mathcal{A} \}$ is a $\sigma$-algebra on $Y$.

(b) If $\mathcal{B}$ is a $\sigma$-algebra on $Y$, then $f^{-1}(\mathcal{B})$ is a $\sigma$-algebra on $X$.

(c) Consider $B \subseteq X$. If $\mathcal{A}$ is a $\sigma$-algebra on $X$, then the trace $\mathcal{A}|B$ is a $\sigma$-algebra on $B$.

Proof. (a): Since $\emptyset = f^{-1}(\emptyset) \in \mathcal{A}$, we have $\emptyset \in \mathcal{B}$. If $B \in \mathcal{B}$, then $Y \setminus B \in \mathcal{B}$, since $f^{-1}(Y \setminus B) = X \setminus f^{-1}(B) \in \mathcal{A}$, due to $\mathcal{A}$ being a $\sigma$-algebra. If $(B_n)_{n \in \mathbb{N}}$ is a sequence of sets in $\mathcal{B}$, then

$$f^{-1}\left(\bigcup_{n=1}^{\infty} B_n\right) = \bigcup_{n=1}^{\infty} f^{-1}(B_n) \in \mathcal{A},$$

as $\mathcal{A}$ is a $\sigma$-algebra. Thus, $\bigcup_{n=1}^{\infty} B_n \in \mathcal{B}$, which completes the proof that $\mathcal{B}$ is a $\sigma$-algebra on $Y$.

(b): As $\emptyset = f^{-1}(\emptyset)$, $\emptyset \in f^{-1}(\mathcal{B})$. If $B \in \mathcal{B}$, then $X \setminus f^{-1}(B) = f^{-1}(Y \setminus B) \in f^{-1}(\mathcal{B})$ since $Y \setminus B \in \mathcal{B}$ due to $\mathcal{B}$ being a $\sigma$-algebra. If $(B_n)_{n \in \mathbb{N}}$ is a sequence of sets in $\mathcal{B}$, then, as $\mathcal{B}$ is a $\sigma$-algebra, $\bigcup_{n=1}^{\infty} B_n \in \mathcal{B}$. Then

$$\bigcup_{n=1}^{\infty} f^{-1}(B_n) = f^{-1}\left(\bigcup_{n=1}^{\infty} B_n\right) \in f^{-1}(\mathcal{B}),$$

completing the proof that $f^{-1}(\mathcal{B})$ is a $\sigma$-algebra on $X$.

(c): Due to (1.39), (c) follows immediately from (b).

The following Th. 1.56 can be seen as a generalization of Prop. 1.55(b),(c).

Theorem 1.56. Consider sets $X, Y$ and a map $f : X \to Y$.

(a) The forming of generated $\sigma$-algebras commutes with the forming of inverse images: If $\mathcal{E} \subseteq \mathcal{P}(Y)$, then

$$\sigma_X(f^{-1}(\mathcal{E})) = f^{-1}(\sigma_Y(\mathcal{E})).$$

(1.40)

(b) Let $\mathcal{A}$ be a $\sigma$-algebra on $X$, $B \subseteq X$, and $\mathcal{E} \subseteq \mathcal{A}$. If $\mathcal{A} = \sigma_X(\mathcal{E})$, then $\mathcal{A}|B = \sigma_B(\mathcal{E}|B)$.

(c) Let $\mathcal{T}$ be a topology on $X$ and let $\mathcal{B}$ denote the corresponding Borel $\sigma$-algebra on $X$. If $B \subseteq X$, then $\mathcal{B}|B$ is the Borel $\sigma$-algebra on $B$ with respect to the relative topology $\mathcal{T}_B = \mathcal{T}|B$ on $B$, i.e. $\mathcal{B}|B = \sigma_B(\mathcal{T}_B)$.

Proof. (a): Since $\mathcal{E} \subseteq \sigma_Y(\mathcal{E})$, one has $f^{-1}(\mathcal{E}) \subseteq f^{-1}(\sigma_Y(\mathcal{E}))$, which implies the $\subseteq$-part of (1.40), as $f^{-1}(\sigma_Y(\mathcal{E}))$ is a $\sigma$-algebra by Prop. 1.55(b). To prove the $\supseteq$-part of (1.40), define

$$\mathcal{B} := \left\{ B \subseteq Y : f^{-1}(B) \in \sigma_X(f^{-1}(\mathcal{E})) \right\}.$$
Then $\mathcal{E} \subseteq \mathcal{B}$, $\mathcal{B}$ is a $\sigma$-algebra by Prop. 1.55(a), implying $\sigma_Y(\mathcal{E}) \subseteq \mathcal{B}$. This, in turn, yields $f^{-1}(\sigma_Y(\mathcal{E})) \subseteq f^{-1}(\mathcal{B}) \subseteq \sigma_X(f^{-1}(\mathcal{E}))$, which is precisely the $\supseteq$-part of (1.40), completing the proof.

(b): One merely has to apply (a) to the canonical inclusion map $f : B \rightarrow X, f(a) = a$:

$$A|\mathcal{B} \overset{(1.39)}{=} f^{-1}(A) = f^{-1}(\sigma_X(\mathcal{E})) \overset{(1.40)}{=} \sigma_B(f^{-1}(\mathcal{E})) \overset{(1.39)}{=} \sigma_B(\mathcal{E}|\mathcal{B}),$$

establishing the case.

(c): Since $\mathcal{B} = \sigma_X(\mathcal{T})$, the statement is a special case of (b).

Please note the analogy between the restriction of a $\sigma$-algebra to a subset and the forming of the relative topology on a subset.

**Proposition 1.57.** If $(X, \mathcal{A}, \mu)$ is a measure space and $A \in \mathcal{A}$, then the restriction $\nu := \mu|_{A|A}$ is a measure on $(A, A|A)$.

**Proof.** According to Prop. 1.55(c), $A|A$ is a $\sigma$-algebra on $A$, and $A \in \mathcal{A}$ implies $A|A \subseteq \sigma$ such that $\nu$ is actually defined for each $A' \in A|A$. Moreover, $\nu(\emptyset) = \mu(\emptyset) = 0$, and, as $\nu$ is the restriction of $\mu$, $\nu$ inherits the $\sigma$-additivity from $\mu$, proving that $\nu$ is a measure. ■

**Example 1.58.** Let $n \in \mathbb{N}$, $A \subseteq \mathbb{R}^n$. Then we can restrict the $\sigma$-algebras $\mathcal{B}^n$ and $\mathcal{L}^n$ to $A$ to obtain $\sigma$-algebras $\mathcal{B}^n|A$ and $\mathcal{L}^n|A$. According to Prop. 1.57, if $A \in \mathcal{B}^n$, then $\beta^n|_{\mathcal{B}^n|A}$ is a measure and, if $A \in \mathcal{L}^n$, then $\lambda^n|_{\mathcal{L}^n|A}$ is a measure. By a slight abuse of notation, we also write $(A, \mathcal{B}^n, \beta^n)$ (for $A \in \mathcal{B}^n$) and $(A, \mathcal{L}^n, \lambda^n)$ (for $A \in \mathcal{L}^n$) to denote the measure spaces $(A, \mathcal{B}^n|A, \beta^n|_{\mathcal{B}^n|A})$ and $(A, \mathcal{L}^n|A, \lambda^n|_{\mathcal{L}^n|A})$, respectively.

### 1.5 Lebesgue Measure on $\mathbb{R}^n$

#### 1.5.1 Intervals and Countable Sets

**Proposition 1.59.** Let $n \in \mathbb{N}$.

(a) If $a, b \in \mathbb{R}^n$, $a \leq b$, $I := [a, b]$, then $\lambda^n(I) = \beta^n(I) = \prod_{j=1}^n |b_j - a_j|.$

(b) If $A \subseteq \mathbb{R}^n$ is countable, then $\lambda^n(A) = \beta^n(A) = 0.$

**Proof.** First note that the statements make sense, since all intervals and all countable subsets of $\mathbb{R}^n$ are Borel-measurable.

(a): For each $k \in \mathbb{N}$, let $a^k := (a_1 - \frac{1}{k}, \ldots, a_n - \frac{1}{k})$, $I_k := [a^k, b] \in \mathcal{I}^n$. Then $I_k \downarrow I$ and one computes

$$\lambda^n(I) \overset{(1.10)}{=} \lim_{k \to \infty} \lambda^n(I_k) \overset{\text{Ex. 1.20(b)}}{=} \lim_{k \to \infty} \prod_{j=1}^n (b_j - a_j + \frac{1}{k}) = \prod_{j=1}^n (b_j - a_j),$$

proving (a).
Theorem 1.60. Let \( n \in \mathbb{N} \).

(a) For each \( A \in \mathcal{L}^n \) and each \( \epsilon \in \mathbb{R}^+ \), there exist \( O \subseteq \mathbb{R}^n \) open and \( F \subseteq \mathbb{R}^n \) closed such that

\[
F \subseteq A \subseteq O \quad \land \quad \lambda^n(O \setminus A) < \epsilon \quad \land \quad \lambda^n(A \setminus F) < \epsilon.
\]

Thus, for each \( A \in \mathcal{L}^n \),

\[
\lambda^n(A) = \inf \{ \lambda^n(O) : A \subseteq O, \ O \text{ open} \} = \sup \{ \lambda^n(F) : F \subseteq A, \ F \text{ closed} \} = \sup \{ \lambda^n(C) : C \subseteq A, \ C \text{ compact} \}.
\]

(1.41)

(b) A set \( A \subseteq \mathbb{R}^n \) is \( \lambda^n \)-measurable if, and only if, for each \( \epsilon \in \mathbb{R}^+ \), there exist \( O \subseteq \mathbb{R}^n \) open and \( F \subseteq \mathbb{R}^n \) closed such that

\[
F \subseteq A \subseteq O \quad \land \quad \lambda^n(O \setminus F) < \epsilon.
\]

Proof. (a): Fix \( \epsilon \in \mathbb{R}^+ \). First, assume \( \lambda^n(A) < \infty \). Then, due to (1.30), there exists a sequence \( (I_k)_{k \in \mathbb{N}} \) in \( \mathcal{I}^n \) such that \( A \subseteq \bigcup_{k=1}^{\infty} I_k \) and \( \sum_{k=1}^{\infty} \lambda^n(I_k) < \lambda^n(A) + \frac{\epsilon}{2} \). For each \( k \in \mathbb{N} \), we enlarge \( I_k \) slightly to obtain some \( J_k \in \mathcal{I}^n \) satisfying \( I_k \subseteq J_k \) and \( \lambda^n(J_k) \leq \lambda^n(I_k) + \epsilon \cdot 2^{-k-1} \). Then \( O := \bigcup_{k=1}^{\infty} J_k \) is open, \( A \subseteq O \), and

\[
\lambda^n(O \setminus A) = \lambda^n(O) - \lambda^n(A) \leq \sum_{k=1}^{\infty} \lambda^n(I_k) + \epsilon \sum_{k=1}^{\infty} 2^{-k-1} - \lambda^n(A) < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon.
\]

If \( \lambda^n(A) = \infty \), then, for each \( k \in \mathbb{N} \), set \( A_k := A \cap [-k,k]^n \). Using what we have already shown, for each \( k \in \mathbb{N} \), there exists an open \( O_k \subseteq \mathbb{R}^n \) such that \( A_k \subseteq O_k \) and \( \lambda^n(O_k \setminus A_k) < \epsilon \cdot 2^{-k} \). Letting \( O := \bigcup_{k=1}^{\infty} O_k \), we obtain \( O \) open with \( A \subseteq O \) and

\[
\lambda^n(O \setminus A) \leq \sum_{k=1}^{\infty} \lambda^n(O_k \setminus A) \leq \sum_{k=1}^{\infty} \lambda^n(O_k \setminus A_k) < \epsilon \sum_{k=1}^{\infty} 2^{-k} = \epsilon,
\]
as desired. One now obtains the closed set $F$ by taking complements: To $A^c$, there exists an open set $U \subset \mathbb{R}^n$ such that $A^c \subseteq U$ and $\lambda^n(U \setminus A^c) < \epsilon$. Then $F := U^c$ is closed, $F \subseteq A$, and $\lambda^n(A \setminus F) = \lambda^n(A \cap U) = \lambda^n(U \setminus A^c) < \epsilon$, as needed. The first two equalities of (1.41) are a direct consequence of what we have proved above. It remains to show the third equality of (1.41). To this end, let $0 < \alpha < \lambda^n(A)$. Then there exists a closed $F \subseteq A$ such that $\lambda^n(F) > \alpha$. For each $k \in \mathbb{N}$, the set $C_k := F \cap [-k,k]^n$ is compact and $C_k \uparrow F$. Thus $\lim_{k \to \infty} \lambda^n(C_k) = \lambda^n(F) > \alpha$, i.e. $\lambda^n(C_k) > \alpha$ must hold for almost all of the $C_k$.

(b): First, assume $A \in \mathcal{L}^n$. Given $\epsilon > 0$, choose $O$ and $F$ as in (a), except with $\epsilon$ replaced by $\epsilon/2$. Then

$$\lambda^n(O \setminus F) = \lambda^n(O \setminus A) + \lambda^n(A \setminus F) < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon$$

proves the first implication. For the converse, given $A \subseteq \mathbb{R}^n$, for each $k \in \mathbb{N}$, choose a closed $F_k$ and an open $O_k$ such that $F_k \subseteq A \subseteq O_k$ and $\lambda^n(O_k \setminus F_k) < \frac{\epsilon}{2}$. Then $B := \bigcup_{k=1}^{\infty} F_k \in \mathcal{B}^n$, $C := \bigcap_{k=1}^{\infty} O_k \in \mathcal{B}^n$, $B \subseteq A \subseteq C$, and $\lambda^n(C \setminus B) = 0$. Thus, $A = B \cup (A \setminus B)$ is the union of the Borel set $B$ and the set $A \setminus B$, which is a subset of the $\beta^n$-null set $C \setminus B$. As $\lambda^n$ is the completion of $\beta^n$, $A \in \mathcal{L}^n$. ■

### 1.5.3 Cantor Set

According to Prop. 1.59(b), $\lambda^n(A) = 0$ for each countable $A \subseteq \mathbb{R}^n$. For $n > 1$, Prop. 1.59(a) immediately provides examples of uncountable sets $A \subseteq \mathbb{R}^n$ with $\lambda^n(A) = 0$. Finding uncountable subsets of $\mathbb{R}$ that are $\lambda^1$-null sets is not as simple. The Cantor set $C$ is an example. The Cantor set occurs frequently in the literature due to its many interesting (e.g. topological) properties. The Cantor set $C$ is what is left of the unit interval $[0,1]$ after removing, in the initial step, the open middle third $]0, \frac{1}{3}[$ and, then, in each successive step, the open middle third of each remaining interval (after the first step, one has left $[0, \frac{1}{3}] \cup [\frac{2}{3},1]$, after the second step, one has left $[0, \frac{1}{3}] \cup [\frac{2}{3}, \frac{1}{3}] \cup \ldots$, etc.). We now give the precise definition:

**Definition and Remark 1.61.** Inductively, we define two sequences $(\mathcal{I}_n)_{n \in \mathbb{N}_0}$ and $(\mathcal{K}_n)_{n \in \mathbb{N}_0}$ of subsets of $[0,1]$ such that, for each $n \in \mathbb{N}_0$, the following statements hold true:

(i) $\mathcal{I}_n$ consists of $2^n$ disjoint open intervals $I_{n,1}, \ldots, I_{n,2^n}$.

(ii) $\mathcal{K}_n$ consists of $2^{n+1}$ disjoint compact intervals $K_{n,1}, \ldots, K_{n,2^{n+1}}$, where

$$\forall i \in \{1, \ldots, 2^{n+1}\} \quad (K_{n,i} = [\alpha_{n,i}, \beta_{n,i}], \quad \alpha_{n,i}, \beta_{n,i} \in [0,1], \quad \beta_{n,i} = \alpha_{n,i} + 3^{-n-1}).$$

(iii) $\lambda^1(I) = \lambda^1(K) = 3^{-n-1}$ for each $I \in \mathcal{I}_n$ and each $K \in \mathcal{K}_n$.

(iv) $[0,1]$ is the disjoint union of all elements of $\mathcal{K}_n \cup \bigcup_{m=0}^{n} \mathcal{I}_m$, i.e.

$$[0,1] = \bigcup_{i=1}^{2^{n+1}} K_{n,i} \cup \bigcup_{m=0}^{n} \bigcup_{i=1}^{2^m} I_{m,i}. $$
(v) If \( n > 0 \), then,
\[
\forall i \in \{1,\ldots,2^n\} \quad K_{n-1,i} = K_{n,2i-1} \cup I_{n,i} \cup K_{n,2i}.
\]

Initially, we set \( I_{0,1} := [\frac{1}{3}, \frac{2}{3}] \), \( K_{0,1} := [0, \frac{1}{3}] \), \( K_{0,2} := [\frac{2}{3}, 1] \), clearly satisfying the above conditions for \( n = 0 \). Now let \( n \in \mathbb{N}_0 \) and assume \( I_0,\ldots,I_n \) as well as \( K_0,\ldots,K_n \) to be already defined such that (i) – (v) hold. Define, for each \( i \in \{1,\ldots,2^{n+1}\} \),
\[
I_{n+1,i} := \alpha_{n,i} + 3^{-n-2}, \alpha_{n,i} + 2 \cdot 3^{-n-2}, K_{n+1,2i-1} := [\alpha_{n,i}, \alpha_{n,i} + 3^{-n-2}], \quad K_{n+1,2i} := [\alpha_{n,i} + 2 \cdot 3^{-n-2}, \alpha_{n,i} + 3^{-n-1}].
\]

Then (ii) and (iii) for \( n+1 \) are immediately clear as well as (v) for \( n+1 \). Together with (iv) for \( n \), this implies (iv) for \( n+1 \). Finally, (v) for \( n+1 \) together with (ii) for \( n \) implies (i) for \( n+1 \), completing the inductive definition of \((I_n)_{n \in \mathbb{N}_0}\) and \((K_n)_{n \in \mathbb{N}_0}\). We are now in a position to define the Cantor set as
\[
C := \bigcap_{n=0}^{\infty} 2^{n+1} \bigcup_{i=1}^{2^{n+1}} K_{n,i} = \bigcap_{n=0}^{\infty} \left( [0,1] \setminus \bigcup_{m=0}^{n} \bigcup_{i=1}^{2^m} I_{m,i} \right) = [0,1] \setminus \bigcup_{m=0}^{\infty} \bigcup_{i=1}^{2^m} I_{m,i}. \quad (1.42)
\]

**Proposition 1.62.** Let \( C \subseteq [0,1] \) be the Cantor set as defined in (1.42). Then the following holds:

(a) \( C \) is compact.

(b) \( C \) does not contain any nontrivial intervals, i.e. no intervals of positive length.

(c) \( \lambda^1(C) = 0 \).

**Proof.** (a): As the intersection of compact subsets of \( \mathbb{R}^n \), \( C \) is compact.

(b): Using the notation of Def. and Rem. 1.61, define, for each \( n \in \mathbb{N} \), \( K_n := \bigcup_{i=1}^{2^{n+1}} K_{n,i} \). Let \( I \subseteq \mathbb{R} \) be a nontrivial interval. Then \( \alpha := \lambda^1(I) \) \( > 0 \). If \( n \in \mathbb{N} \) is such that \( 3^{-n-1} < \alpha \), then \( I \not\subseteq K_n \) (since \( K_n \) is the disjoint union of intervals of length \( 3^{-n-1} \)). Since \( C \subseteq K_n \), \( I \not\subseteq C \).

(c): Due to (1.42), we have
\[
\lambda^1(C) = 1 - \sum_{m=0}^{\infty} \sum_{i=1}^{2^m} \lambda^1(I_{m,i}) = 1 - \sum_{m=0}^{\infty} 2^m 3^{-m-1} = 1 - \frac{1}{3} \frac{1}{1-\frac{2}{3}} = 0,
\]
proving (c). \( \blacksquare \)

It is not difficult to see directly that all the endpoints \( \alpha_{n,i} \) and \( \beta_{n,i} \) of the \( K_{n,i} \) are elements of \( C \), showing \( C \) must be infinite. However, there are only countably many \( \alpha_{n,i} \) and \( \beta_{n,i} \), and, thus, to see that \( C \) is uncountable (Th. 1.64(b) below), we first have to learn a bit more about the elements of \( C \).
Lemma 1.63. Using the notation of Def. and Rem. 1.61, the left endpoints \( \alpha_{n,i} \) of \( K_{n,i} \), \( n \in \mathbb{N}_0 \), \( i \in \{1, \ldots, 2^{n+1}\} \), are precisely the numbers \( x \in \mathbb{R} \) that have a finite triadic expansion, where all digits are either 0 or 2: Let \( n \in \mathbb{N}_0 \), \( x \in \mathbb{R} \). Then

\[
\left( \forall_{i \in \{1, \ldots, 2^{n+1}\}} \ x = \alpha_{n,i} \right) \iff \left( \exists_{d_1, \ldots, d_{n+1} \in \{0,2\}} \ x = \sum_{k=1}^{n+1} d_k \cdot 3^{-k} \right).
\]

Similarly, the right endpoints \( \beta_{n,i} \) of \( K_{n,i} \), \( n \in \mathbb{N}_0 \), \( i \in \{1, \ldots, 2^{n+1}\} \), are precisely the numbers \( x \in \mathbb{R} \) that have a periodic triadic expansion, where the period is 2 and all digits are either 0 or 2: Let \( n \in \mathbb{N}_0 \), \( x \in \mathbb{R} \). Then

\[
\left( \exists_{i \in \{1, \ldots, 2^{n+1}\}} \ x = \beta_{n,i} \right) \iff \left( \exists_{d_1, \ldots, d_{n+1} \in \{0,2\}} \ x = \sum_{k=1}^{n+1} d_k \cdot 3^{-k} + \sum_{k=n+2}^{\infty} 2 \cdot 3^{-k} \right).
\]

Proof. Since \( \beta_{n,i} = \alpha_{n,i} + 3^{-n-1} \) and \( \sum_{k=n+2}^{\infty} 2 \cdot 3^{-k} = 2 \cdot 3^{-n-2} \cdot \frac{3}{2} = 3^{-n-1} \), it suffices to prove the statement for the \( \alpha_{n,i} \). The statement for the \( \alpha_{n,i} \) is proved via induction on \( n \in \mathbb{N}_0 \): For \( n = 0 \), \( x = \alpha_{0,1} \) if, and only if, \( x = 0 = 0 \cdot 3^{-1} \); \( x = \alpha_{0,2} \) if, and only if, \( x = \frac{2}{3} = 2 \cdot 3^{-1} \). Thus, the statement holds for \( n = 0 \). For the induction step, assume the statement to hold for some fixed \( n \in \mathbb{N}_0 \). Then \( x = \alpha_{n+1,2i-1} \), \( i \in \{1, \ldots, 2^{n+1}\} \), is equivalent to \( x = \alpha_{n,i} \), and, by induction, this is equivalent to \( x = \sum_{k=1}^{n+1} d_k \cdot 3^{-k} + 0 \cdot 3^{-(n+2)} \) with \( d_1, \ldots, d_{n+1} \in \{0,2\} \). Similarly, \( x = \alpha_{n+1,2i} \), \( i \in \{1, \ldots, 2^{n+1}\} \), is equivalent to \( x = \alpha_{n,i} + 2 \cdot 3^{-n-2} \) and, by induction, this is equivalent to \( x = \sum_{k=1}^{n+1} d_k \cdot 3^{-k} + 2 \cdot 3^{-(n+2)} \) with \( d_1, \ldots, d_{n+1} \in \{0,2\} \), completing the induction and the proof of the lemma.

Theorem 1.64. Let \( C \) be the Cantor set defined in Def. and Rem. 1.61.

(a) \( C \) consists of all \( x \in \{0,1\} \) that have a triadic expansion such that all digits are either 0 or 2.

(b) \( C \) has the same cardinality as \( \mathbb{R} \); in particular, \( C \) is uncountable.

Proof. (a): Let \( x \in \{0,1\} \). According to the definition of \( C \) in (1.42), we have to show

\[
\left( \forall_{n \in \mathbb{N}_0} \ x \in \bigcup_{i=1}^{2^{n+1}} K_{n,i} \right) \iff \left( \exists_{(d_k)_{k \in \mathbb{N}}} \ x = \sum_{k=1}^{\infty} d_k \cdot 3^{-k} \right).
\]

“\( \Leftarrow \)” : Suppose \( x = \sum_{k=1}^{\infty} d_k \cdot 3^{-k} \) with \( d_k \in \{0,2\} \) for each \( k \in \mathbb{N} \). Fix \( n \in \mathbb{N}_0 \). According to Lem. 1.63, there exists \( i \in \{1, \ldots, 2^{n+1}\} \) such that \( \alpha_{n,i} = \sum_{k=1}^{n+1} d_k \cdot 3^{-k} \). Then \( \alpha_{n,i} \leq x \leq \alpha_{n,i} + \sum_{k=n+2}^{\infty} 2 \cdot 3^{-k} = \beta_{n,i} \), showing \( x \in K_{n,i} \).

“\( \Rightarrow \)” : Suppose \( x \in \bigcup_{i=1}^{2^{n+1}} K_{n,i} \) for each \( n \in \mathbb{N}_0 \). If \( x \) is a right endpoint \( \beta_{n,i} \), then \( x \) has a triadic expansion of the required form by Lem. 1.63. If \( x \) is not a right endpoint \( \beta_{n,i} \), then, for each \( n \in \mathbb{N}_0 \), there exists \( i \in \{1, \ldots, 2^{n+1}\} \) such that \( \alpha_{n,i} \leq x < \alpha_{n,i} + 3^{-n-1} \). Thus, \( x - \alpha_{n,i} < 3^{-n-1} \), i.e. the coefficients of \( 3^{-1}, \ldots, 3^{-(n+1)} \) in the triadic expansions of \( \alpha_{n,i} \) and \( x \) must be identical. Then Lem. 1.63 yields that \( x = \sum_{k=1}^{\infty} d_k \cdot 3^{-k} \) with \( d_k \in \{0,2\} \) for each \( k \in \mathbb{N} \).
(b): Consider the map \( \phi : \{0, 2\}^\mathbb{N} \to C \), \( \phi((d_k)_{k \in \mathbb{N}}) := \sum_{k=1}^{\infty} d_k \cdot 3^{-k} \). According to (a), \( \phi \) is well-defined and surjective. It is also injective by [Phi16a, Th. 7.99], proving \( \#\{0, 2\}^\mathbb{N} = \#C \). Thus
\[
\#C = \#\{0, 2\}^\mathbb{N} = \#\{0, 1\}^\mathbb{N} \overset{\text{[Phi16a, Th. F.2]}}{=} \#\mathbb{R},
\]
completing the proof of (b).

As mentioned at the beginning of this section, for \( n > 1 \), Prop. 1.59(a) provides examples of sets \( A \subseteq \mathbb{R}^n \) with \( \#A = \#\mathbb{R} \) and \( \lambda^n(A) = 0 \). We have now also seen that the Cantor set is a set \( A \subseteq \mathbb{R} \) with \( \#A = \#\mathbb{R} \) and \( \lambda^1(A) = 0 \). Since each \( \lambda^n \) is complete, this implies \( \#L^n = \#\mathcal{P}(\mathbb{R}) \) for each \( n \in \mathbb{N} \). As \( \#B^n = \#\mathbb{R} \) according to [Els07, Sec. 2.8.3], this also shows \( B^n \subset L^n \).

### 1.6 Measurable Maps

#### 1.6.1 Definition, Composition

Measurable maps are for measure theory what continuous maps are for topology: Recall that a map \( f : (X, \mathcal{S}) \to (Y, \mathcal{T}) \) is continuous if, and only if, \( f^{-1}(B) \in \mathcal{S} \) for each \( B \in \mathcal{T} \). If one replaces each topology with a \( \sigma \)-algebra, then one obtains the concept of a measurable map:

**Definition 1.65.** Let \((X, \mathcal{A})\) and \((Y, \mathcal{B})\) be measurable spaces, \( f : X \to Y \).

(a) \( f \) is called \( \mathcal{A}-\mathcal{B} \)-measurable (or often merely measurable) if, and only if, \( f^{-1}(B) \in \mathcal{A} \) for each \( B \in \mathcal{B} \).

(b) If \( \mathcal{S} \) and \( \mathcal{T} \) are topologies on \( X \) and \( Y \), respectively, then \( f \) is called Borel measurable if, and only if, it is \( \sigma(\mathcal{S})-\sigma(\mathcal{T}) \)-measurable.

---

According to [Phi16b, Th. 2.7(iii)], to check if a map \( f : (X, \mathcal{S}) \to (Y, \mathcal{T}) \) is continuous, one only needs to verify if all preimages of sets in a subbase of \( \mathcal{T} \) are open in \( X \). The following Prop. 1.66 is the analogous result for measurable maps:

**Proposition 1.66.** Let \((X, \mathcal{A})\) and \((Y, \mathcal{B})\) be measurable spaces. If \( \mathcal{E} \subseteq \mathcal{P}(Y) \) is a generator of \( \mathcal{B} \) (i.e. \( \sigma_Y(\mathcal{E}) = \mathcal{B} \)), then a map \( f : X \to Y \) is \( \mathcal{A}-\mathcal{B} \)-measurable if, and only if,
\[
f^{-1}(B) \in \mathcal{A} \quad \text{for each } B \in \mathcal{E}.
\] (1.43)

**Proof.** Since \( \mathcal{E} \subseteq \mathcal{B} \), \( \mathcal{A}-\mathcal{B} \)-measurability implies (1.43). To prove the converse, note that \( f^{-1}(\mathcal{E}) \subseteq \mathcal{A} \) implies
\[
f^{-1}(\mathcal{B}) = f^{-1}(\sigma_Y(\mathcal{E})) \overset{\text{Th. 1.56(a)}}{=} \sigma_X(f^{-1}(\mathcal{E})) \subseteq \mathcal{A},
\]
completing the proof that \( f \) is \( \mathcal{A}-\mathcal{B} \)-measurable.

---
Example 1.67. (a) Constant maps are always measurable. More precisely, if \((X, \mathcal{A})\) and \((Y, \mathcal{B})\) are measurable spaces, \(f : X \to Y, \; f \equiv c, \; c \in Y\), then, for each \(B \in \mathcal{B}\), there are precisely two possibilities: \(c \in B\), implying \(f^{-1}(B) = X \in \mathcal{A}\); \(c \notin B\), implying \(f^{-1}(B) = \emptyset \in \mathcal{A}\). Thus, \(f\) is \(\mathcal{A}\)-\(\mathcal{B}\)-measurable.

(b) Continuous maps are always Borel measurable: Let \((X, \mathcal{S})\) and \((Y, \mathcal{T})\) be topological spaces with corresponding Borel \(\sigma\)-algebras \(\sigma(\mathcal{S})\) and \(\sigma(\mathcal{T})\). If \(f : X \to Y\) is continuous, then \(f^{-1}(U) \in \mathcal{S} \subseteq \sigma(\mathcal{S})\) for each \(U \in \mathcal{T}\) and Prop. 1.66 implies \(f\) is \(\sigma(\mathcal{S})\)-\(\sigma(\mathcal{T})\) measurable. In particular, if \(M \subseteq \mathbb{R}^n\) and \(f : M \to \mathbb{R}^m\) is continuous \((n, m \in \mathbb{N})\), then \(f\) is Borel measurable, i.e. \(\mathcal{B}^n\)-\(\mathcal{B}^m\)-measurable (cf. Ex. 1.58 and Th. 1.56(c)).

(c) Let \(X\) be a set and \((Y, \mathcal{B})\) a measurable space, \(f : X \to Y\). Then \(\mathcal{A} := f^{-1}(\mathcal{B})\) is the smallest \(\sigma\)-algebra on \(X\) that makes \(f\) measurable (i.e. \(\mathcal{A} = \bigcap\{\mathcal{C} \in \mathcal{P}(X) : \mathcal{C}\ \sigma\)-algebra on \(X\) and \(f\) \(\mathcal{C}\)-\(\mathcal{B}\)-measurable\})).

Proposition 1.68. The composition of measurable maps is measurable – more precisely, if \((X, \mathcal{A}), (Y, \mathcal{B})\), and \((Z, \mathcal{C})\) are measurable spaces, \(f : X \to Y\) is \(\mathcal{A}\)-\(\mathcal{B}\)-measurable, and \(g : Y \to Z\) is \(\mathcal{B}\)-\(\mathcal{C}\)-measurable, then \(g \circ f\) is \(\mathcal{A}\)-\(\mathcal{C}\)-measurable.

Proof. For each \(C \in \mathcal{C}\), we have \(A := (g \circ f)^{-1}(C) = f^{-1}(g^{-1}(C))\). Then \(g^{-1}(C) \in \mathcal{B}\), since \(g\) is \(\mathcal{B}\)-\(\mathcal{C}\)-measurable. In consequence, \(A \in \mathcal{A}\), since \(f\) is \(\mathcal{A}\)-\(\mathcal{B}\)-measurable, proving the \(\mathcal{A}\)-\(\mathcal{C}\) measurability of \(g \circ f\). \(\blacksquare\)

1.6.2 Pushforward Measure

Theorem 1.69. Let \((X, \mathcal{A}, \mu)\) be a measure space, let \((Y, \mathcal{B})\) be a measurable space, and let \(f : X \to Y\) be \(\mathcal{A}\)-\(\mathcal{B}\)-measurable.

(a) The function
\[
\mu_f : \mathcal{B} \to [0, \infty], \quad \mu_f(B) := \mu(f^{-1}(B)),
\]
(1.44)
defines a measure on \(\mathcal{B}\), called a pushforward measure – the measure \(\mu\) is pushed forward from \(\mathcal{A}\) to \(\mathcal{B}\) by \(f\). Alternatively to \(\mu_f\), one also finds the notation \(f(\mu) := \mu_f\).

(b) The forming of push forward measures commutes with the composition of maps: If \((Z, \mathcal{C})\) is another measurable space, where \(g : Y \to Z\) is \(\mathcal{B}\)-\(\mathcal{C}\)-measurable, then
\[
\mu_{gf} = (\mu_f)_g \quad \text{or} \quad (g \circ f)(\mu) = g(f(\mu)).
\]
(1.45)

Proof. (a): One has \(\mu_f(\emptyset) = \mu(\emptyset) = 0\) and, if \((B_n)_{n \in \mathbb{N}}\) is a sequence of disjoint sets in
Let Proposition 1.70.

(b): First note that the hypotheses imply $\sigma$-additivity of $\mu_f$ and completing the proof of (a).

(b): If $\nu = f(\mu)$, then $\tilde{\nu} = f(\tilde{\mu})$.

Proof. (a): Let $\tilde{B} \in \tilde{\mathcal{B}}$. Then there exist $B, C \in \mathcal{B}$ and $N \subseteq C$ such that $\nu(C) = 0$ and $\tilde{B} = B \cup N$. Then

$$\tilde{A} := f^{-1}(\tilde{B}) = f^{-1}(B) \cup f^{-1}(N),$$

where $f^{-1}(B) \in \mathcal{A}$ and $f^{-1}(N) \subseteq f^{-1}(C)$ with $\mu(f^{-1}(C)) = 0$, showing $\tilde{A} \in \tilde{\mathcal{A}}$, i.e. $f$ is $\tilde{\mathcal{A}}\tilde{\mathcal{B}}$-measurable.

(b): If $\nu = f(\mu)$, and $C \in \mathcal{B}$ is a $\nu$-null set, then $\mu(f^{-1}(C)) = \nu(C) = 0$. Thus, (a) applies such that $f$ is $\tilde{\mathcal{A}}\tilde{\mathcal{B}}$-measurable and $f(\tilde{\mu})$ is well-defined. If $\tilde{B}, B, C, N$ are as in the proof of (a), then

$$(f(\tilde{\mu}))(\tilde{B}) = \tilde{\mu}(f^{-1}(B) \cup f^{-1}(N)) = \mu(f^{-1}(B)) = \nu(B) = \tilde{\nu}(\tilde{B}),$$

which establishes the case.

1.6.3 Invariance of Lebesgue Measure under Euclidean Isometries

In the present section, we will show that the measures $\beta^n$ and $\lambda^n$ satisfy Def. 1.1(b), i.e. they are translation invariant. Indeed, in Th. 1.73 below, we will obtain the change of $\beta^n$ and $\lambda^n$ under arbitrary bijective affine maps, which will include the invariance of $\beta^n$ and $\lambda^n$ under Euclidean isometries, in particular translations, as a special case (also see the paragraph before the statement of Th. 1.73 for a brief review of the relation between affine maps, Euclidean isometries, and translations). For $A \in \mathcal{P}(\mathbb{R}^n)$, $x \in \mathbb{R}^n$, $n \in \mathbb{N}$, we recall the notation

$$A + x = \{a + x : a \in A\}.$$
Definition 1.71. Let $n \in \mathbb{N}$.

(a) If $a \in \mathbb{R}^n$, then the affine map

$$T_a : \mathbb{R}^n \to \mathbb{R}^n, \quad T_a(x) := x + a,$$

is called a translation, namely, the translation by $a$.

(b) Let $(\mathbb{R}^n, \mathcal{A}, \mu)$ be a measure. Then $\mu$ is called translation invariant if, and only if, for each $a \in \mathbb{R}^n$, the translation $T_a$ is $\mathcal{A}$-$\mathcal{A}$-measurable and $T_a(\mu) = \mu$, i.e. if, and only if,

$$\forall a \in \mathbb{R}^n \quad \forall A \in \mathcal{A} \quad \mu(A + a) = (T_a^{-1}(\mu))(A) = (T_a(\mu))(A) = \mu(A). \quad (1.46)$$

Theorem 1.72. Let $n \in \mathbb{N}$.

(a) $\beta^n$ (resp. $\lambda^n$) is translation invariant and, moreover, it is the only translation invariant measure $\mu$ on $\mathcal{B}^n$ (resp. $\mathcal{L}^n$) satisfying $\mu([0,1]^n) = 1$.

(b) If $\mu$ is a translation invariant measure on $\mathcal{B}^n$ (resp. $\mathcal{L}^n$) such that $\mu([0,1]^n) = \alpha < \infty$, then $\mu = \alpha \beta^n$ (resp. $\mu = \alpha \lambda^n$).

Proof. (a): Let $a \in \mathbb{R}^n$. Since the translation $T_a$ is continuous, it is $\mathcal{B}^n$-$\mathcal{B}^n$-measurable. Furthermore, if $c, d \in \mathbb{R}^n$ with $c \leq d$, then $[c, d] + a = [c + a, d + a]$ and

$$\beta^n([c, d] + a) = \prod_{j=1}^{n} (d_j + a_j - c_j - a_j) = \prod_{j=1}^{n} (d_j - c_j) = \beta^n([c, d]).$$

Thus, the $\sigma$-finite measures $\beta^n$ and $T_a(\beta^n)$ agree on the semiring $\mathcal{T}^n$. By Cor. 1.46(a), they must also agree on $\mathcal{B}^n$, proving the translation invariance of $\beta^n$. Then, according to Prop. 1.70(a), $T_a$ is $\mathcal{B}^n$-$\mathcal{L}^n$-measurable. Now apply Prop. 1.70(b) with $\mu := \beta^n$ and $\nu := T_a(\beta^n) = \beta^n$, obtaining

$$T_a(\lambda^n) = T_a(\mu) = \nu = \mu = \lambda^n,$$

proving the translation invariance of $\lambda^n$. Conversely, assume $\mu : \mathcal{B}^n \to [0, \infty]$ to be a translation invariant measure. In a first step, we assume $\mu([0,1]^n) = 1$. Given $g := (n_1, \ldots, n_n) \in \mathbb{N}^n$, we can decompose $[0,1]^n$ into $n_1 \cdots n_n$ translative copies of the interval $I_g := \prod_{i=1}^{n} [0, \frac{1}{n_i}]$: If

$$K_g := \left\{ \left( \frac{k_1}{n_1}, \ldots, \frac{k_n}{n_n} \right) \in \mathbb{Q}^n : (k_1, \ldots, k_n) \in (\mathbb{N}_0)^n, \quad \forall i \in \{1, \ldots, n\} \quad 0 \leq k_i < n_i \right\},$$

then one has the disjoint union

$$[0,1]^n = \bigcup_{k \in K_g} (I_g + k).$$
Due to the translation invariance of \( \mu \), this yields
\[
1 = \mu([0, 1]^n) = n_1 \cdots n_n \cdot \mu(I_g) \quad \Rightarrow \quad \mu(I_g) = \frac{1}{n_1 \cdots n_n} = \beta^n(I_g).
\]
If \( \emptyset \neq I \in \mathcal{I}_Q^n \), then there exist \( g := (n_1, \ldots, n_n) \in \mathbb{N}^n \) and \( a, b \in \mathbb{Z}^n \) with \( a < b \), such that
\[
I = \prod_{i=1}^{n} \left[ \frac{a_i}{n_i}, \frac{b_i}{n_i} \right] = \bigcup_{k \in K_{g,a,b}} (I_g + k),
\]
\[
K_{g,a,b} := \left\{ \left( \frac{k_1}{n_1}, \ldots, \frac{k_n}{n_n} \right) \in \mathbb{Q}^n : (k_1, \ldots, k_n) \in \mathbb{Z}^n, \ \forall \ a_i \leq k_i < b_i \right\}.
\]
In this case, the translation invariance of \( \mu \) yields
\[
\mu(I) = \prod_{i=1}^{n} \frac{b_i - a_i}{n_i} = \beta^n(I).
\]
Thus, the \( \sigma \)-finite measures \( \mu \) and \( \beta^n \) agree on \( \mathcal{I}_Q^n \), i.e., they must agree on \( \mathcal{B}^n \) by Cor. 1.46(a). If \( \mu \) extends to a measure on \( \mathcal{L}^n \), then \( \mu = \lambda^n \) by Th. 1.51. It remains to consider a translation invariant measure \( \mu \) on \( \mathcal{B}^n \) (resp. \( \mathcal{L}^n \)) with \( \mu([0, 1]^n) = 1 \). Set
\[
\alpha := \mu([0, 1]^n).
\]
As \( \mu \) is translation invariant, we obtain
\[
1 = \mu([0, 1]^n) \leq \mu([-1, 1]^n) = 2^n \cdot \alpha \quad \Rightarrow \quad \alpha \geq 2^{-n} > 0.
\]
Thus, \( \nu := \alpha^{-1} \mu \) is a translation invariant measure with \( \nu([0, 1]^n) = 1 \), implying \( \nu = \beta^n \) (resp. \( \nu = \lambda^n \)). Since \( \alpha^{-1} = \alpha^{-1} \mu([0, 1]^n) = \nu([0, 1]^n) = \beta^n([0, 1]^n) = 1 \), it is \( \mu = \beta^n \) (resp. \( \mu = \lambda^n \)), completing the proof of (a).

(b): If \( \alpha > 0 \), then \( \alpha^{-1} \mu = \beta^n \) (resp. \( \alpha^{-1} \mu = \lambda^n \)) according to (a), proving \( \mu = \alpha \beta^n \) (resp. \( \mu = \alpha \lambda^n \)). If \( \alpha = 0 \), then \( \mu([0, 1]^n) = 0 \) and the translation invariance of \( \mu \) yields
\[
\mu(\mathbb{R}^n) = \mu \left( \bigcup_{a \in \mathbb{Z}^n} [0, 1]^n + a \right) = \sum_{a \in \mathbb{Z}^n} \mu([0, 1]^n) = 0,
\]
showing \( \mu \equiv 0 \equiv 0 \cdot \beta^n \) (resp. \( \mu \equiv 0 \equiv 0 \cdot \lambda^n \)).

Note that Th. 1.72(b) does not hold without the condition \( \alpha < \infty \), as counting measure is a translation invariant measure on both \( \mathcal{B}^n \) and \( \mathcal{L}^n \) that is not a multiple of \( \beta^n \) (resp. \( \lambda^n \)).

In generalization of the above result on the translation invariance of \( \beta^n \) and \( \lambda^n \), we will now investigate these measures’ behavior under general bijective affine maps. Recall from Linear Algebra that a map \( f : \mathbb{R}^n \rightarrow \mathbb{R}^n \) is called affine if, and only if, there exists a linear map \( L : \mathbb{R}^n \rightarrow \mathbb{R}^n \) and a vector \( a \in \mathbb{R}^n \) such that \( f(x) = L(x) + a \) for each \( x \in \mathbb{R}^n \). Thus, using Def. 1.71(a), \( f : \mathbb{R}^n \rightarrow \mathbb{R}^n \) is affine if, and only if, \( f = T_a \circ L \) with a linear \( L \) and a vector \( a \) as above. Clearly, in that case, \( a \) and \( L \) are uniquely determined by the affine \( f \) (\( a = f(0), \ L = f - a \)), and one defines \( \det f := \det L \). In consequence, one has the equivalences
\[
f \text{ bijective} \iff L \text{ bijective} \iff \det L \neq 0 \iff \det f \neq 0.
\]
1 MEASURE THEORY

Theorem 1.73. Let \( n \in \mathbb{N}, f : \mathbb{R}^n \rightarrow \mathbb{R}^n \).

(a) If \( f \) is affine and bijective, then \( f \) is \( B^n \)-\( B^n \)-measurable as well as \( L^\alpha \)-\( L^\alpha \)-measurable and

\[
\begin{align*}
f(\beta^n) &= |\det f|^{-1} \beta^n, \\
f(\lambda^n) &= |\det f|^{-1} \lambda^n. \tag{1.47a}
\end{align*}
\]

Furthermore, if \( A \in B^n \) (resp. \( A \in L^n \)), then \( f(A) \in B^n \) (resp. \( f(A) \in L^n \)) and

\[
\beta^n(f(A)) = |\det f| \beta^n(A), \quad \lambda^n(f(A)) = |\det f| \lambda^n(A). \tag{1.47b}
\]

(b) If \( f \) is a Euclidean isometry (i.e. \( \|f(x) - f(y)\|_2 = \|x - y\|_2 \)), then \( f \) is \( B^n \)-\( B^n \)-measurable as well as \( L^\alpha \)-\( L^\alpha \)-measurable and

\[
\begin{align*}
f(\beta^n) &= \beta^n, \\
f(\lambda^n) &= \lambda^n. \tag{1.48a}
\end{align*}
\]

Furthermore, if \( A \in B^n \) (resp. \( A \in L^n \)), then \( f(A) \in B^n \) (resp. \( f(A) \in L^n \)) and

\[
\beta^n(f(A)) = \beta^n(A), \quad \lambda^n(f(A)) = \lambda^n(A). \tag{1.48b}
\]

Proof. Since we know from Linear Algebra that \( f \) is a Euclidean isometry if, and only if, \( f = L + a \) with a linear map \( L : \mathbb{R}^n \rightarrow \mathbb{R}^n \) that has \( |\det L| = 1 \) (i.e. an orthogonal linear map \( L \)) and \( a \in \mathbb{R}^n \) (see, e.g., [Koe03, Sec. 5.5.1]), it suffices to prove (a).

Since the affine map \( f \) is continuous, it is \( B^n \)-\( B^n \)-measurable. As \( f \) is affine and bijective, there exists a linear map \( L : \mathbb{R}^n \rightarrow \mathbb{R}^n \) with \( \det L \neq 0 \) and some \( a \in \mathbb{R}^n \) such that \( f = T_a \circ L \). The main work consists of showing

\[
L(\beta^n) = |\det L|^{-1} \beta^n. \tag{1.49}
\]

Let \( \mu : B^n \rightarrow [0, \infty] \) be some arbitrary translation invariant measure on \( B^n \). We claim \( L(\mu) \) to be translation invariant as well: Indeed, for each \( A \in B^n \) and each \( v \in \mathbb{R}^n \), using the linearity of \( L \) and the translation invariance of \( \mu \),

\[
(L(\mu))(A + v) = \mu(L^{-1}(A + v)) = \mu(L^{-1}(A) + L^{-1}(v)) = \mu(L^{-1}(A)) = (L(\mu))(A).
\]

We introduce the abbreviation \( I := [0, 1]^n \) and, for the rest of the proof, we make the additional assumption that \( \alpha_\mu := \mu(I) < \infty \). Then \( \mu = \alpha_\mu \beta^n \) by Th. 1.72(b), and

\[
\alpha_{L, \mu} := (L(\mu))(I) = \mu(L^{-1}(I)) = \alpha_\mu \beta^n(L^{-1}(I)) < \infty
\]

(as the continuous image \( L^{-1}(I) \) of the compact set \( I \) is compact), implying

\[
L(\mu) = \alpha_{L, \mu} \beta^n.
\]

We need to show \( \alpha_{L, \beta^n} = |\det L|^{-1} \). Suppose \( L \) is orthogonal (i.e. \( |\det L| = 1 \)) and let \( B := \{ x \in \mathbb{R}^n : \|x\|_2 \leq 1 \} \). Then, since \( L \) is a Euclidean isometry, \( L^{-1}(B) = B \) and, thus,

\[
\alpha_{L, \mu} \beta^n(B) = (L(\mu))(B) = \mu(L^{-1}(B)) = \mu(B) = \alpha_\mu \beta^n(B),
\]
1 MEASURE THEORY

implying \( \alpha_{L, \mu} = \alpha_\mu \), thereby proving (1.49) for orthogonal \( L \). Next, assume \( L \) is represented by a diagonal matrix with diagonal entries \( d_1, \ldots, d_n \in \mathbb{R}^+ \) with respect to the standard basis \((e_1, \ldots, e_n)\) of \( \mathbb{R}^n \). Then

\[
\alpha_{L, \beta^n} = \beta^n (L^{-1}(I)) = \beta^n ([(0, \ldots, 0), (d_1^{-1}, \ldots, d_n^{-1})]) = \prod_{j=1}^{n} d_j^{-1} = |\det L|^{-1},
\]

proving (1.49) also for this case. Now let \( L \) be a general linear map with \( \det L \neq 0 \). We obtain this general case from the above special cases, using a bit more linear algebra: According to [Koe03, Sec. 6.3.2], \( P := LL^t \) is positive definite. Using [Koe03, Sec. 6.3.2], \( P \) can be written as \( P = L_1 D L_1^t \), where \( L_1 \) is orthogonal and \( D \) is represented by a diagonal matrix with diagonal entries \( d_1, \ldots, d_n \in \mathbb{R}^+ \) as above. If \( S \) is represented by the diagonal matrix with diagonal entries \( \sqrt{d_1}, \ldots, \sqrt{d_n} \), then \( S^2 = D \) and \( K := S^{-1} L_1 L \) is orthogonal:

\[
KK^t = S^{-1} L_1^t LL^t L_1 S^{-1} = S^{-1} L_1^t L_1 S^2 L_1^t L_1 S^{-1} = \text{Id}.
\]

Moreover, \( L_1 S K = L_1 S S^{-1} L_1^t L = L \) and, thus, \( |\det L| = |\det L_1| \cdot |\det S| \cdot |\det K| = 1 \cdot |\det S \cdot 1 = \det S \), implying

\[
\alpha_{L, \beta^n} = (L(\beta^n))(I) = ((L_1 \circ S \circ K)(\beta^n))(I) = (L_1 (S(K(\beta^n))))(I) = (L_1 ((\det S)^{-1} \beta^n))(I) = (|\det L|^{-1} \beta^n)(I) = |\det L|^{-1},
\]

completing the proof of (1.49). We now come back to our bijective affine map \( f = T_\alpha \circ L \). Since \( L(\beta^n) \) is translation invariant by (1.49), \( f(\beta^n) = T_\alpha (L(\beta^n)) = L(\beta^n) = |\det f|^{-1} \beta^n \) as claimed. Then Prop. 1.70(b) yield with \( \mu := \beta^n \) and \( \nu := f(\beta^n) = |\det f|^{-1} \beta^n \) and obtain

\[
f(\lambda^n) = f(\mu) = \nu (1.37c) |\det f|^{-1} \mu = |\det f|^{-1} \lambda^n,
\]

completing the proof of (1.47a). Finally, (1.47b) follows by applying (1.47a) to \( f^{-1} \):

\[
\forall A \in \mathcal{L}^n \quad \lambda^n (f(A)) = (f^{-1}(\lambda^n))(A) = |\det f| \lambda^n(A),
\]

finishing the proof of the theorem.

1.6.4 Nonmeasurable Sets

We are now in a position to show the existence of subsets \( M \) of \( \mathbb{R}^n \) that are not \( \lambda^n \)-measurable and, more generally, that no ideal \( n \)-dimensional volume as defined in Def. 1.1 can exist (see Th. 1.74 below). The proof hinges on the validity of the axiom of choice AC. Without the axiom of choice the situation becomes subtle and difficult: It has been shown that there exist models of ZF set theory (without AC), where every
subset of $\mathbb{R}^n$ is $\lambda^n$-measurable. However, these constructions assume the existence of so-called inaccessible cardinals, and if the existence of such cardinals is consistent with ZF remains unknown (see the comments and references at the end of [Els07, Sec. III.3]).

In preparation for Th. 1.74, let $G \subseteq \mathbb{R}^n$, $n \in \mathbb{N}$, be an additive subgroup of $\mathbb{R}^n$ and recall from Linear Algebra that the factor group (or quotient group) $\mathbb{R}^n/G$ consists of all cosets

$$G + a = \{g + a : g \in G\}, \quad a \in \mathbb{R}^n.$$ 

These are precisely the equivalence classes of the equivalence relation defined by $x \sim y :\iff x - y \in G$. By AC, there exists a set $M_G \subseteq \mathbb{R}^n$ of representatives of $\mathbb{R}^n/G$, containing precisely one element from each coset $G + a \in \mathbb{R}^n/G$.

**Theorem 1.74.** Let $n \in \mathbb{N}$.

(a) Let $(\mathbb{R}^n, \mathcal{A}, \mu)$ be a measure space such that $\mu$ extends $\beta^n$ (i.e. $\mathcal{B}^n \subseteq \mathcal{A}$ and $\mu|_{\mathcal{B}^n} = \beta^n$). Moreover, let $G := \mathbb{Q}^n \subseteq \mathbb{R}^n$. We consider $G$ as an additive subgroup of $\mathbb{R}^n$ and let $M := M_G$ be a set of representatives of $\mathbb{R}^n/G$. If $\mu$ is translation invariant with respect to $G$ (i.e. $A + g \in \mathcal{A}$ and $\mu(A + g) = \mu(A)$ hold for each $A \in \mathcal{A}$ and each $g \in G$), then $M \notin \mathcal{A}$ and $\mu(A) = 0$ for each $A \in \mathcal{A}$ with $A \subseteq M$ (in particular, choosing $\mathcal{A} := \mathcal{L}^n$, $\mu := \lambda^n$, yields the existence of a set $M \subseteq \mathbb{R}^n$ (by use of AC) that is not $\lambda^n$-measurable).

(b) There does not exist an ideal $n$-dimensional volume as defined in Def. 1.1.

(c) Let $\eta^n : \mathcal{P}(\mathbb{R}^n) \rightarrow [0, \infty]$ denote the outer measure induced by $\lambda^n$. Then each $A \subseteq \mathbb{R}^n$ with $\eta^n(A) > 0$ contains a set $B$ that is not Lebesgue measurable (i.e. $B \notin \mathcal{L}^n$).

**Proof.** (a): Seeking a contradiction, assume $M \in \mathcal{A}$. Define $I := ]0, 1]^n$ and

$$M_0 := \bigcup_{g \in \mathbb{Z}^n} \left( -g + (M \cap (g + I)) \right).$$

Clearly, $M_0 \subseteq I$. Moreover, $M_0$ is still a set of representatives of $\mathbb{R}^n/G$ (since $\mathbb{R}^n = \bigcup_{g \in \mathbb{Z}^n} (g + I)$ and adding the element $-g \in G$ to a coset does not change the coset). Moreover, the assumed translation invariance of $\mu$ yields $-g + (M \cap (g + I)) \in \mathcal{A}$ for each $g \in \mathbb{Z}^n$ and, thus, $M_0 \in \mathcal{A}$. Using the translation invariance of $\mu$ once again together with the countability of $G$, we obtain

$$\infty = \beta^n(\mathbb{R}^n) = \mu(\mathbb{R}^n) = \mu\left( \bigcup_{g \in G} (g + M_0) \right) = \sum_{g \in G} \mu(g + M_0) = \sum_{g \in G} \mu(M_0),$$

implying $\mu(M_0) > 0$. On the other hand, since $G \cap I$ is infinite (and countable),

$$\sum_{g \in G \cap I} \mu(M_0) = \sum_{g \in G \cap I} \mu(g + M_0) = \mu\left( \bigcup_{g \in G \cap I} (g + M_0) \right) \leq \mu([0, 2]^n) = \beta^n([0, 2]^n) = 2^n < \infty,$$
implying \(\mu(M_0) = 0\), in contradiction to \(\mu(M_0) > 0\). Now, if \(A \in \mathcal{A}\) with \(A \subseteq M\), then one can repeat the above argument with \(M\) replaced by \(A\), except that one no longer obtains \(\sum_{g \in G} \mu(M_0) = \infty\), i.e., one obtains \(\mu(A) = 0\) without any contradiction.

(b): An ideal \(n\)-dimensional volume \(\mu\) would satisfy all the hypotheses of (a) with \(\mathcal{A} = \mathcal{P}(\mathbb{R}^n)\), in contradiction to the existence of some \(M \in \mathcal{P}(\mathbb{R}^n)\) with \(M \notin \mathcal{A}\).

(c): Let \(G := \mathbb{Q}^n\) and \(M\) be as in (a). Then \(\mathbb{R}^n = \bigcup_{g \in G} (g + M)\) and the \(\sigma\)-subadditivity of \(\eta^n\) imply

\[
\eta^n(A) \leq \sum_{g \in G} \eta^n\left(A \cap (g + M)\right).
\]

(1.50)

Since, for each \(g \in G\), \(g + M\) is still a set of representatives of \(\mathbb{R}^n/G\), (a) implies \(\eta^n\left(A \cap (g + M)\right) = \lambda^n\left(A \cap (g + M)\right) = 0\) if \(A \cap (g + M) \in \mathcal{L}^n\). Thus, if \(A \cap (g + M) \in \mathcal{L}^n\) for each \(g \in G\), then (1.50) implies \(\eta^n(A) = 0\).

\[\blacksquare\]

1.6.5 \(\mathbb{R}\)-Valued Measurable Maps

Recall that, in [Phi16b, Ex. 1.52(b)], we extended the total order on \(\mathbb{R}\) to a total order on \(\mathbb{R} = \mathbb{R} \cup \{-\infty, \infty\}\). Then the resulting order topology \(\mathcal{T}\) on \(\mathbb{R}\) also extended the standard topology on \(\mathbb{R}\), where a base for \(\mathcal{T}\) is given by the set of open intervals

\[
\mathcal{T}_b := \{\mathbb{R}\} \cup \{|a, b|: a, b \in \mathbb{R}, a < b\}
\]

\[
\bigcup \{-\infty, a]\}: a \in \mathbb{R} \cup \{\infty\}\bigcup \{|a, \infty| : a \in \mathbb{R} \cup \{-\infty\}\}
\]

(1.51)

(Also see Def. A.2(b) and (A.5) in the Appendix).

**Notation 1.75.** Let \(\mathcal{B}\) denote the Borel sets on \(\mathbb{R}\), i.e. the Borel sets with respect to the order topology \(\mathcal{T}\) on \(\mathbb{R}\).

**Lemma 1.76.** For the Borel sets on \(\mathbb{R}\), we have the following identities:

\[
\mathcal{B} = \{B \cup E : B \in \mathcal{B}^1, E \subseteq \{-\infty, \infty\}\},
\]

(1.52a)

\[
\mathcal{B} = \sigma(\{-\infty, \alpha[: \alpha \in \mathbb{R}\}) = \sigma(\{-\infty, \alpha\} : \alpha \in \mathbb{R})
\]

\[
= \sigma(\{\alpha, \infty| : \alpha \in \mathbb{R}\}) = \sigma(\{\alpha, \infty] : \alpha \in \mathbb{R}\}),
\]

(1.52b)

\[
\mathcal{B}|\mathbb{R} = \mathcal{B}^1.
\]

(1.52c)

**Proof.** (1.52a): Let \(\mathcal{A}\) denote the right-hand side of (1.52a). Since \(\{-\infty\}\) and \(\{\infty\}\) are closed sets in \(\mathbb{R}\), every \(E \subseteq \{-\infty, \infty\}\) is a Borel set, implying \(\mathcal{A} \subseteq \mathcal{B}\). To verify the remaining inclusion, note \(\sigma(\mathcal{I}_0) = \mathcal{B}\), \(\mathcal{I}_0 \subseteq \mathcal{A}\), and \(\mathcal{A}\) is a \(\sigma\)-algebra.

(1.52b): Clearly, all four generator sets are subsets of \(\mathcal{B}\). Let \(\mathcal{E}_1 := \{[-\infty, \alpha[: \alpha \in \mathbb{R}\}, \mathcal{E}_2 := \{[-\infty, \alpha] : \alpha \in \mathbb{R}\}\). We first show \(\sigma(\mathcal{E}_2) = \mathcal{B}\): If \(\alpha, \beta \in \mathbb{R}\), then \(\{\alpha, \beta\} = [-\infty, \beta] \setminus [-\infty, \alpha]\). Thus, \(\mathcal{I}_1 \subseteq \sigma(\mathcal{E}_2)\), implying \(\mathcal{B}^1 \subseteq \sigma(\mathcal{E}_2)\). Since also \(\{-\infty\} = \bigcap_{n \in \mathbb{N}} [-\infty, -n] \in \sigma(\mathcal{E}_2)\) and \(\{\infty\} = \bigcap_{n \in \mathbb{N}} [n, \infty] \in \sigma(\mathcal{E}_2)\), \(\sigma(\mathcal{E}_2) = \mathcal{B}\) now follows from (1.52a). If \(\alpha \in \mathbb{R}\), then \([-\infty, \alpha] = \bigcup_{n \in \mathbb{N}} [-\infty, \alpha - \frac{1}{n}] \in \sigma(\mathcal{E}_1)\) implies \(\sigma(\mathcal{E}_1) = \mathcal{B}\). Since the elements of \(\{\alpha, \infty| : \alpha \in \mathbb{R}\}\) are precisely the complements of the elements of \(\mathcal{E}_2\)
and the elements of \{[\alpha, \infty] : \alpha \in \mathbb{R}\} are precisely the complements of the elements of \mathcal{E}_1, the proof of (1.52b) is complete.

(1.52c): Since the topology on \mathbb{R} is the relative topology inherited from \mathbb{R}, (1.52c) follows from Th. 1.56(c). Alternatively, (1.52c) is immediate from (1.52a).

\begin{definition}
Let \((X, \mathcal{A})\) be a measurable space.
\begin{enumerate}[(a)]
\item A function \(f : X \to \mathbb{R}\) is called measurable if, and only if, it is \(\mathcal{A}-\mathcal{B}\)-measurable.
\item A function \(f : X \to \mathbb{C}\) is called measurable if, and only if, it is \(\mathcal{A}-\mathcal{B}^2\)-measurable, where \(\mathcal{B}^2\) are the Borel sets on \(\mathbb{C}\) with respect to the standard topology on \(\mathbb{C}\) (the notation \(\mathcal{B}^2\) is reasonable, since \(\mathbb{C} = \mathbb{R}^2\) and \(\mathcal{B}^2\) is precisely the set of Borel sets on \(\mathbb{R}^2\)).
\item A function \(f : X \to \mathbb{R}^n, n \in \mathbb{N}\), is called measurable if, and only if, it is \(\mathcal{A}-\mathcal{B}^n\)-measurable.
\item A function \(f : \mathbb{R} \to \mathbb{R}\) is called measurable if, and only if, it is \(\mathcal{L}^1-\mathcal{B}\)-measurable.
\end{enumerate}
\end{definition}

\begin{notation}
In the present context, it is useful to introduce a shorthand notation for certain subsets of \(X\): Let \(\mathcal{F} := (f_i)_{i \in I}\) be a family of functions defined on \(X\). Let \(P(\mathcal{F})\) denote a sentence (i.e. a sequence of symbols) and let \(P(\mathcal{F})(x)\) denote the corresponding sentence, where each occurrence of \(f_i\) in \(P(\mathcal{F})\) was replaced by \(f_i(x)\). If \(P(\mathcal{F})\) is such that \(P(\mathcal{F})(x)\) constitutes a statement (i.e. is either true or false) for each \(x \in X\), then we define
\[
\{P(\mathcal{F})\} := \{x \in X : P(\mathcal{F})(x) \text{ holds true}\}.
\]
For example, consider \(f, g : X \to \overline{\mathbb{R}}\). If \(\mathcal{F} = (f)\) and \(\alpha \in \mathbb{R}\), then, e.g.,
\[
\{f < \alpha\}_{P(\mathcal{F})} = \{x \in X : f(x) < \alpha\} = f^{-1}([-\infty, \alpha])
\]
and, if \(\mathcal{F} = (f, g)\), then, e.g.,
\[
\{f = g\}_{P(\mathcal{F})} = \{x \in X : f(x) = g(x)\}.
\]
\end{notation}

\begin{theorem}
Let \((X, \mathcal{A})\) be a measurable space and \(f : X \to \overline{\mathbb{R}}\). Then the following statements are equivalent:
\begin{enumerate}[(i)]
\item \(f\) is measurable.
\item \(\{f < \alpha\} \in \mathcal{A}\) for each \(\alpha \in \mathbb{R}\).
\item \(\{f \leq \alpha\} \in \mathcal{A}\) for each \(\alpha \in \mathbb{R}\).
\item \(\{f > \alpha\} \in \mathcal{A}\) for each \(\alpha \in \mathbb{R}\).
\end{enumerate}
\end{theorem}
(v) \( \{ f \geq \alpha \} \in A \) for each \( \alpha \in \mathbb{R} \).

**Proof.** One merely has to combine Prop. 1.66 with (1.52b).

**Notation 1.80.** Let \( X \) be a set and let \( (f_i)_{i \in I} \) be a family of functions \( f_i : X \rightarrow \mathbb{R} \). Then define

\[
\inf_{i \in I} f_i : X \rightarrow \mathbb{R}, \quad \left( \inf_{i \in I} f_i \right)(x) := \inf \{ f_i(x) : i \in I \}, \tag{1.55a}
\]

\[
\sup_{i \in I} f_i : X \rightarrow \mathbb{R}, \quad \left( \sup_{i \in I} f_i \right)(x) := \sup \{ f_i(x) : i \in I \}. \tag{1.55b}
\]

For \( I = \mathbb{N} \), we also define

\[
\lim_{i \rightarrow \infty} f_i : X \rightarrow \mathbb{R}, \quad \left( \lim_{i \rightarrow \infty} f_i \right)(x) := \lim \inf_{i \rightarrow \infty} f_i(x), \tag{1.55c}
\]

\[
\limsup_{i \rightarrow \infty} f_i : X \rightarrow \mathbb{R}, \quad \left( \limsup_{i \rightarrow \infty} f_i \right)(x) := \lim \sup_{i \rightarrow \infty} f_i(x), \tag{1.55d}
\]

\[
\liminf_{i \rightarrow \infty} f_i : X \rightarrow \mathbb{R}, \quad \left( \liminf_{i \rightarrow \infty} f_i \right)(x) := \lim \inf_{i \rightarrow \infty} f_i(x), \tag{1.55e}
\]

where (1.55e) is defined, provided each limit \( \lim_{i \rightarrow \infty} f_i(x) \) exists in \( \mathbb{R} \) (of course, this is the pointwise limit of the \( f_i \), which we have already studied previously). The following notation will also turn out to be useful: If \( f : X \rightarrow \mathbb{R} \), then define

\[
f_i \uparrow f : \iff \left( f = \lim_{i \rightarrow \infty} f_i \land \forall x \in X \ f_1(x) \leq f_2(x) \leq \ldots \right), \tag{1.56a}
\]

\[
f_i \downarrow f : \iff \left( f = \lim_{i \rightarrow \infty} f_i \land \forall x \in X \ f_1(x) \geq f_2(x) \geq \ldots \right). \tag{1.56b}
\]

In the case of (1.55e) (and, in particular, (1.56)), we call the convergence **uniform** if, and only if,

\[
\forall \epsilon \in \mathbb{R}^+ \exists N \in \mathbb{N} \forall i \geq N \forall x \in X \ |f_i(x) - f(x)| < \epsilon. \tag{1.57}
\]

**Lemma 1.81.** Let \( X \) be a set and let \( (f_k)_{k \in \mathbb{N}} \) be a sequence of functions \( f_k : X \rightarrow \mathbb{R} \). Then

\[
\forall \alpha \in \mathbb{R} \left( \left\{ \inf_{k \in \mathbb{N}} f_k \geq \alpha \right\} = \bigcap_{k=1}^{\infty} \{ f_k \geq \alpha \} \right) \land \left( \left\{ \sup_{k \in \mathbb{N}} f_k \leq \alpha \right\} = \bigcap_{k=1}^{\infty} \{ f_k \leq \alpha \} \right), \tag{1.58a}
\]

\[
\lim_{k \rightarrow \infty} f_k = \sup_{k \in \mathbb{N}} \left( \inf_{i \geq k} f_i \right) \land \limsup_{k \rightarrow \infty} f_k = \inf_{k \in \mathbb{N}} \left( \sup_{i \geq k} f_i \right). \tag{1.58b}
\]

If \( \lim_{k \rightarrow \infty} f_k \) exists, then also

\[
\lim_{k \rightarrow \infty} f_k = \lim_{k \rightarrow \infty} f_k = \lim_{k \rightarrow \infty} f_k. \tag{1.58c}
\]
Proof. Let \( \alpha \in \mathbb{R}, \ x \in X \). Then the two assertions in (1.58a) are precisely the equivalences
\[
\inf \{ f_k(x) : k \in \mathbb{N} \} \geq \alpha \quad \Leftrightarrow \quad \forall k \in \mathbb{N} \ f_k(x) \geq \alpha,
\]
\[
\sup \{ f_k(x) : k \in \mathbb{N} \} \leq \alpha \quad \Leftrightarrow \quad \forall k \in \mathbb{N} \ f_k(x) \leq \alpha.
\]
We prove (1.58b) next. Let \( x \in X \). We have to show
\[
\alpha := \lim inf_{k \to \infty} f_k(x) = \sup \left\{ \inf \{ f_i(x) : i \geq k \} : k \in \mathbb{N} \right\} =: \beta.
\]
For each \( k \in \mathbb{N} \), let \( r_k := \inf \{ f_i(x) : i \geq k \} \). Then \( \alpha \geq r_k \): If there exists \( i_0 \geq k \) such that \( \alpha \geq f_{i_0}(x) \), then \( \alpha \geq f_{i_0}(x) \geq r_k \), since \( r_k \) is a lower bound for the \( f_i(x), i \geq k \).
If \( \alpha < f_i(x) \) for all \( i \geq k \), then \( \alpha < \infty \) and \( \alpha = r_k = \inf \{ f_i(x) : i \geq k \} \), since, for \( \alpha > -\infty \), each \( \alpha, \alpha + \epsilon, \epsilon > 0 \), must contain some \( f_i(x), i \geq k \), and, for \( \alpha = -\infty \), each \( -\infty, -\epsilon, \epsilon > 0 \), must contain some \( f_i(x) \) (as \( \alpha \) is a cluster point). Now \( \alpha \geq r_k \) for each \( k \in \mathbb{N} \) implies \( \beta \leq \alpha \), since \( \beta \) is the smallest upper bound of the \( r_k \). It remains to show \( \alpha \leq \beta \). To this end, let \( \gamma \in \mathbb{R} \) be an arbitrary upper bound of \( R := \{ r_k : k \in \mathbb{N} \} \). As \( \beta \) is an upper bound of \( R \), it suffices to show \( \alpha \leq \gamma \). Seeking a contradiction, assume \( \gamma < \alpha \). For each \( k \in \mathbb{N} \), since \( r_k \leq \gamma \) and \( r_k = \inf \{ f_i(x) : i \geq k \} \), there exists \( i \geq k \) such that \( f_i(x) < \beta \), showing that \( [\alpha, r_k] \) must contain \( f_j(x) \) for infinitely many \( j \in \mathbb{N} \), in contradiction to \( \alpha \) being the smallest cluster point of \( (f_j(x))_{j \in \mathbb{N}} \). Thus, we have shown \( \alpha = \beta \), as desired. One can conclude the second equation of (1.58b) from the first by applying it to the reversed order on \( \mathbb{R} \). Alternatively, one can conduct an analogous proof as for the first equation: Let \( x \in X \). We have to show
\[
\alpha := \lim sup_{k \to \infty} f_k(x) = \inf \left\{ \sup \{ f_i(x) : i \geq k \} : k \in \mathbb{N} \right\} =: \beta.
\]
For each \( k \in \mathbb{N} \), let \( r_k := \sup \{ f_i(x) : i \geq k \} \). Then \( \alpha \leq r_k \): If there exists \( i_0 \geq k \) such that \( \alpha \leq f_{i_0}(x) \), then \( \alpha \leq f_{i_0}(x) \leq r_k \), since \( r_k \) is an upper bound for the \( f_i(x), i \geq k \).
If \( \alpha > f_i(x) \) for all \( i \geq k \), then \( \alpha > -\infty \) and \( \alpha = r_k = \sup \{ f_i(x) : i \geq k \} \), since, for \( \alpha < \infty \), each \( \alpha - \epsilon, \epsilon > 0 \), must contain some \( f_i(x), i \geq k \), and, for \( \alpha = \infty \), each \( \infty, \epsilon > 0 \), must contain some \( f_i(x) \) (as \( \alpha \) is a cluster point). Now \( \alpha \leq r_k \) for each \( k \in \mathbb{N} \) implies \( \beta \geq \alpha \), since \( \beta \) is the largest lower bound of the \( r_k \). It remains to show \( \alpha \geq \beta \). To this end, let \( \gamma \in \mathbb{R} \) be an arbitrary lower bound of \( R = \{ r_k : k \in \mathbb{N} \} \). As \( \beta \) is a lower bound of \( R \), it suffices to show \( \alpha \geq \gamma \). Seeking a contradiction, assume \( \gamma > \alpha \). For each \( k \in \mathbb{N} \), since \( r_k \geq \gamma \) and \( r_k = \sup \{ f_i(x) : i \geq k \} \), there exists \( i \geq k \) such that \( f_i(x) > \beta \), showing that \( [r_k, \alpha] \) must contain \( f_j(x) \) for infinitely many \( j \in \mathbb{N} \), in contradiction to \( \alpha \) being the largest cluster point of \( (f_j(x))_{j \in \mathbb{N}} \). Thus, we have shown \( \alpha = \beta \), as desired, finishing the proof of (1.58b). If the limit of the sequence \( (f_k(x))_{k \in \mathbb{N}} \), \( x \in X \), exists, then it is the only cluster point of the sequence, proving (1.58c).

\[\text{Theorem 1.82.} \quad \text{Let} \ (X, \mathcal{A}) \text{ be a measurable space, let} \ (f_k)_{k \in \mathbb{N}} \text{ be a sequence of measurable functions,} \ f_k : X \to \mathbb{R}.
\]
\[\begin{align*}
(a) \ & \text{Each of the functions} \ \inf_{k \in \mathbb{N}} f_k, \ \sup_{k \in \mathbb{N}} f_k, \ \lim_{k \to \infty} f_k, \ \limsup_{k \to \infty} f_k, \ \liminf_{k \to \infty} f_k \text{ is measurable. If} \\
\ & f = \lim_{k \to \infty} f_k, \ f : X \to \mathbb{R}, \text{ then} \ f \text{ is measurable as well.}
\end{align*}\]
(b) For each \( n \in \mathbb{N} \), \( \min(f_1, \ldots, f_n) \) and \( \max(f_1, \ldots, f_n) \) are measurable.

Proof. (a): If \( f_k \) is measurable, then \( \{ f_k \geq \alpha \} \in \mathcal{A} \) for each \( \alpha \in \mathbb{R} \). Thus, \( \inf_{k \in \mathbb{N}} f_k \) is measurable by (1.58a) and Th. 1.79(v); \( \sup_{k \in \mathbb{N}} f_k \) is measurable by (1.58a) and Th. 1.79(iii). In consequence, \( \lim_{k \to \infty} f_k \) and \( \liminf_{k \to \infty} f_k \) are measurable by (1.58b), implying \( \lim_{k \to \infty} f_k \) to be measurable by (1.58c).

(b) follows by applying the inf and sup parts of (a) to the sequence of measurable functions \( (f_1, \ldots, f_n, f_n, f_n, \ldots) \).

\[ \tag*{\text{Theorem 1.83}} \]

Let \((X, \mathcal{A})\) be a measurable space and \( n \in \mathbb{N} \).

(a) A function \( f : X \to \mathbb{R}^n \) is measurable if, and only if, all coordinate functions \( f_1, \ldots, f_n : X \to \mathbb{R} \) are measurable.

(b) A function \( f : X \to \mathbb{C} \) is measurable if, and only if, the real-valued functions \( \text{Re} f \) and \( \text{Im} f \) are both measurable.

Proof. (a): Each projection \( \pi_j : \mathbb{R}^n \to \mathbb{R}, j \in \{1, \ldots, n\} \), is continuous and, thus, \( \mathcal{B}^n - \mathcal{B}^1 \)-measurable. Thus, if \( f \) is \( \mathcal{A} - \mathcal{B}^n \)-measurable, then \( f_j = \pi_j \circ f \) is \( \mathcal{A} - \mathcal{B}^1 \)-measurable. Conversely, if each \( f_j, j \in \{1, \ldots, n\} \), is measurable, then, for each \( I := [a, b] \in \mathcal{I}^n \), \( a, b \in \mathbb{R}^n, a < b \), one has \( f^{-1}(I) = \bigcap_{j=1}^{n} f_j^{-1}([a_j, b_j]) \in \mathcal{A} \). Since \( \sigma(\mathcal{I}^n) = \mathcal{B}^n \), this proves the \( \mathcal{A} - \mathcal{B}^n \)-measurability of \( f \).

(b) is an immediate consequence of (a), as \( \mathbb{C} = \mathbb{R}^2 \) with \( \text{Re} f = f_1, \text{Im} f = f_2 \).

\[ \tag*{\text{Theorem 1.84}} \]

Let \((X, \mathcal{A})\) be a measurable space.

(a) If \( f, g : X \to \mathbb{R} \) are measurable and \( \alpha \in \mathbb{R} \), then \( f + g, \alpha f, fg, \) and \( |f| \) are all measurable, and so is each of the sets \( \{ f < g \}, \{ f \leq g \}, \{ f = g \}, \{ f \neq g \} \). If \( g \neq 0, -\infty, \infty \), then \( f/g \) is measurable as well.

(b) If \( f, g : X \to \mathbb{C} \) are measurable and \( \alpha \in \mathbb{C} \), then \( f + g, \alpha f, fg, \) and \( |f| \) are all measurable, and so are the sets \( \{ f = g \}, \{ f \neq g \} \). If \( g \neq 0 \), then \( f/g \) is measurable as well.

Proof. (a): Let \( f, g \) be measurable. Since \( \mathbb{Q} \) is countable, we obtain

\[
\begin{align*}
\{ f < g \} &= \bigcup_{q \in \mathbb{Q}} (\{ f < q \} \cap \{ q < g \}) \in \mathcal{A}, \\
\{ f \leq g \} &= \{ f < g \} \cup \{ f = g \} \in \mathcal{A}, \\
\{ f = g \} &= \{ f \leq g \} \cap \{ g \leq f \} \in \mathcal{A}, \\
\{ f \neq g \} &= \{ f = g \}^c \in \mathcal{A}.
\end{align*}
\]

If \( \alpha \in \mathbb{R} \), then \( \{ f < \alpha \} = \{ -f > -\alpha \} \), showing \( -f \) to be measurable by Th. 1.79. Now assume \( f, g \) to be \( \mathbb{R} \)-valued. Then \( h := (f, g) : X \to \mathbb{R}^2 \) is measurable by Th. 1.83(a). Since the continuous functions \( s, p : \mathbb{R}^2 \to \mathbb{R}, s(x, y) := x + y, p(x, y) := xy \)
are $\mathcal{B}^2, \mathcal{B}^1$-measurable, $f + g = s \circ h$ and $fg = p \circ h$ are measurable. If $f, g$ are $\mathbb{R}$-valued, then we have

\[
A_0 := \{f + g = 0\} = \{f = -g\} \in \mathcal{A},
\]
\[
A_1 := \{f + g = \infty\} = \left(\{f = \infty\} \cap \{g \neq -\infty\}\right) \cup \left(\{g = \infty\} \cap \{f \neq -\infty\}\right) \in \mathcal{A},
\]
\[
A_2 := \{f + g = -\infty\} = \left(\{f = -\infty\} \cap \{g \neq \infty\}\right) \cup \left(\{g = -\infty\} \cap \{f \neq \infty\}\right) \in \mathcal{A}.
\]

Let $C := X \setminus (A_0 \cup A_1 \cup A_2)$. Then $f|_C$ and $g|_C$ are $\mathbb{R}$-valued and $\mathcal{A}|C - \mathcal{B}$-measurable. As $C \in \mathcal{A}$, they are also $\mathcal{A}-\mathcal{B}$-measurable. Thus, $(f + g)|_C$ is $\mathcal{A}-\mathcal{B}$-measurable as well. Now, if $B \in \mathcal{B}$, then

\[
(f + g)^{-1}(B) = (f + g)^{-1}(B \cap \{0, -\infty, \infty\}) \cup (f + g)^{-1}(B \cap \{0, -\infty, \infty\}^c),
\]

where the first set is in $\mathcal{A}$, since $A_0, A_1, A_2 \in \mathcal{A}$, and the second set is in $\mathcal{A}$, since $(f + g)|_C$ is $\mathcal{A}-\mathcal{B}$-measurable. This shows $f + g$ to be measurable. Now $fg$ follows to be measurable in an analogous manner, except this time we let

\[
A_0 := \{fg = 0\} = \{f = 0\} \cup \{g = 0\} \in \mathcal{A},
\]
\[
A_1 := \{fg = \infty\} = \left(\{f = \infty\} \cap \{g > 0\}\right) \cup \left(\{g = \infty\} \cap \{f > 0\}\right)
\]
\[
\quad \cup \left(\{f = -\infty\} \cap \{g < 0\}\right) \cup \left(\{g = -\infty\} \cap \{f < 0\}\right) \in \mathcal{A},
\]
\[
A_2 := \{fg = -\infty\} = \left(\{f = \infty\} \cap \{g < 0\}\right) \cup \left(\{g = \infty\} \cap \{f < 0\}\right)
\]
\[
\quad \cup \left(\{f = -\infty\} \cap \{g > 0\}\right) \cup \left(\{g = -\infty\} \cap \{f > 0\}\right) \in \mathcal{A}.
\]

In particular, $\alpha f$ is measurable, as the constant function with value $\alpha$ is measurable. Next, we note $|f| = \max(f, -f)$ to be measurable. If $g \neq 0, -\infty, \infty$, then $g^{-1}$ is well-defined and measurable (since $g^{-1} = i \circ g$, where $i : \mathbb{R} \setminus \{0\} \to \mathbb{R}$, $i(x) := x^{-1}$, is continuous).

(b) now follows from (a) combined with Th. 1.83(b), where, for $|f|$, one also uses the continuity of the square root function.

**Corollary 1.85.** Let $(X, \mathcal{A})$ be a measurable space.

(a) $f : X \to \overline{\mathbb{R}}$ is measurable if, and only if, both $f^+$ and $f^-$ are measurable.

(b) $f : X \to \mathbb{C}$ is measurable if, and only if, $\text{Re} f^+, \text{Re} f^-, \text{Im} f^+$, $\text{Im} f^-$ all are measurable.

**Proof.** (a): If $f$ is measurable, then $f^+ = \max(f, 0)$ and $f^- = \max(-f, 0)$ are measurable. If $f^+$ and $f^-$ are measurable, then so is $f = f^+ - f^-$. (b) follows by combining (a) with Th. 1.83(b).  ■
Caveat 1.86. One has to use caution with the statement “The composition of measurable functions is measurable.”: While it is correct if it is meant in the sense of Prop. 1.68, it is false if applied to functions $f, g : \mathbb{R} \to \mathbb{R}$, that are measurable in the sense of Def. 1.77(d): If $f$ and $g$ are $\mathcal{L}^1$-$\mathcal{B}^1$-measurable, then $g \circ f$ is not necessarily $\mathcal{L}^1$-$\mathcal{B}^1$-measurable (see, e.g., [RF10, Sec. 3.1, p. 57] for a counterexample). Of course, $g \circ f$ is $\mathcal{L}^1$-$\mathcal{B}^1$-measurable, if $g$ is $\mathcal{B}^1$-$\mathcal{B}^1$-measurable (in particular, if $g$ is continuous).

In the rest of the present section, we study so-called simple or step functions (Def. 1.87). Their importance lies in the fact that every $\mathbb{R}$-valued measurable map is the pointwise limit of a suitable sequence of simple functions (see Th. 1.89 below). We will make use of this result, when developing the integration theory of $\mathbb{R}$-valued measurable maps (see Sections 2.1.1 and 2.1.2).

Definition 1.87. Let $(X, \mathcal{A})$ be a measurable space. Then a real-valued measurable function $f : X \to \mathbb{R}$ is called an $\mathcal{A}$-simple function or an $\mathcal{A}$-step function (often, $\mathcal{A}$ is understood and $f$ is merely called a simple function or a step function) if, and only if, $f$ takes only finitely many different values, i.e. if, and only if, $\#f(X) < \infty$. Moreover, we introduce the following notation:

$$S := S(\mathcal{A}) := \{(f : X \to \mathbb{R}) : f \text{ is } \mathcal{A}\text{-simple}\},$$
$$S^+ := S^+(\mathcal{A}) := \{f \in S : f \geq 0\},$$
$$\mathcal{M} := \mathcal{M}(\mathcal{A}) := \{(f : X \to \mathbb{R}) : f \text{ is measurable}\},$$
$$\mathcal{M}^+ := \mathcal{M}^+(\mathcal{A}) := \{f \in \mathcal{M} : f \geq 0\}.$$ (1.59a)

Caveat: The functions in $S$ and $S^+$ are $\mathbb{R}$-valued, but the functions in $\mathcal{M}$ and $\mathcal{M}^+$ are allowed to be $\mathbb{R}$-valued.

We compile some basic properties of simple functions in the following proposition. In preparation, recall the notion of the characteristic function $\chi_A$ of a subset $A$ of a set $X$:

$$\chi_A : X \to \mathbb{R}, \quad \chi_A(x) := \begin{cases} 1 & \text{if } x \in A, \\ 0 & \text{if } x \notin A. \end{cases}$$

Proposition 1.88. Let $(X, \mathcal{A})$ be a measurable space.

(a) If $f, g \in S$ and $\alpha \in \mathbb{R}$, then $f + g \in S$, $\alpha f \in S$ (i.e. $S$ is a vector space over $\mathbb{R}$), $fg \in S$, $\max(f, g) \in S$, $\min(f, g) \in S$, and $|f| \in S$. Moreover, $f/g \in S$ if $g \neq 0$. Everything remains true if $S$ is replaced by $S^+$ and $\alpha \in \mathbb{R}_0^+$ (of course, $S^+$ is no longer a vector space (if $X \neq \emptyset$)).

(b) If $f : X \to \mathbb{R}$ is a simple function, then there exist distinct numbers $\alpha_1, \ldots, \alpha_n \in \mathbb{R}$, $n \in \mathbb{N}$, and disjoint measurable sets $A_1, \ldots, A_n \in \mathcal{A}$ such that

$$f = \sum_{i=1}^{n} \alpha_i \chi_{A_i}.$$ (1.60)
More precisely, if \( f(X) = \{\alpha_1, \ldots, \alpha_n\} \) with distinct \( \alpha_1, \ldots, \alpha_n \in \mathbb{R} \), then one obtains the representation (1.60) with disjoint \( A_1, \ldots, A_n \in \mathcal{A} \) by setting \( A_i := f^{-1}(\{\alpha_i\}) \).

(c) A function \( f : X \to \mathbb{R} \) is a simple function if, and only if, it is a linear combination of characteristic functions of measurable subsets of \( X \).

Proof. (a): If \( f, g \) are measurable, then we already know \( f + g, \alpha f, fg, \max(f, g), \min(f, g), |f|, \) and \( f/g \) (for \( g \neq 0 \)) to be measurable. Moreover, if \( \#f(X) = m, \#g(X) = n \), with \( m, n \in \mathbb{N} \), \( f(X) = \{\alpha_1, \ldots, \alpha_m\}, g(X) = \{\beta_1, \ldots, \beta_n\} \), then

\[
\#(f + g)(X) = \#\{\alpha_i + \beta_j : i \in \{1, \ldots, m\}, j \in \{1, \ldots, n\}\} \leq m \cdot n < \infty,
\]

\[
\#(\alpha f)(X) = \#\{\alpha \alpha_i : i \in \{1, \ldots, m\}\} \leq m < \infty,
\]

\[
\#(fg)(X) = \#\{\alpha_i \beta_j : i \in \{1, \ldots, m\}, j \in \{1, \ldots, n\}\} \leq m \cdot n < \infty,
\]

\[
\#(\max(f, g))(X) = \#\{\max(\alpha_i, \beta_j) : i \in \{1, \ldots, m\}, j \in \{1, \ldots, n\}\} \leq m + n < \infty,
\]

\[
\#(\min(f, g))(X) = \#\{\min(\alpha_i, \beta_j) : i \in \{1, \ldots, m\}, j \in \{1, \ldots, n\}\} \leq m + n < \infty,
\]

\[
\#|f|(X) = \#\{|\alpha_i| : i \in \{1, \ldots, m\}\} \leq m < \infty.
\]

For \( \beta_1, \ldots, \beta_n \neq 0 \), also

\[
\#(f/g)(X) = \#\{\alpha_i/\beta_j : i \in \{1, \ldots, m\}, j \in \{1, \ldots, n\}\} \leq m \cdot n < \infty.
\]

Clearly, if \( f, g \geq 0 \), then \( f + g \geq 0, \alpha f \geq 0 \) for \( \alpha \geq 0, fg \geq 0, \max(f, g) \geq 0, |f| \geq 0 \) (holds even without the hypothesis \( f \geq 0 \)) and \( f/g \geq 0 \) for \( g \neq 0 \).

(b): If \( f \) is a simple function, then it takes finitely many distinct values \( \alpha_1, \ldots, \alpha_n \in \mathbb{R}, n \in \mathbb{N} \). Thus, if we let \( A_i := f^{-1}(\{\alpha_i\}) \) for \( i \in \{1, \ldots, n\} \), then each \( A_i \) is measurable (as \( f \) is measurable) and the validity of (1.60) is clear.

(c) is an immediate consequence of (a) and (b), since \( \chi_A \in \mathcal{S} \) for \( A \in \mathcal{A} \).

Theorem 1.89. Let \( (X, \mathcal{A}) \) be a measurable space and \( f : X \to \mathbb{R} \).

(a) \( f \in \mathcal{M}^+ \) if, and only if, there exists a sequence \((\phi_k)_{k \in \mathbb{N}} \in \mathcal{S}^+ \) with \( \phi_k \uparrow f \).

(b) \( f \in \mathcal{M} \) with \( f \) bounded from below by \( \alpha \in \mathbb{R} \) (and bounded from above by \( \beta \in \mathbb{R} \)) if, and only if, there exists a sequence \((\phi_k)_{k \in \mathbb{N}} \in \mathcal{S} \) with all \( \phi_k \geq \alpha \) (all \( \phi_k \leq \beta \)) and \( \phi_k \uparrow f \).

(c) \( f \in \mathcal{M} \) if, and only if, there exists a sequence \((\phi_k)_{k \in \mathbb{N}} \in \mathcal{S} \) with \( f = \lim_{k \to \infty} \phi_k \).

Additionally, in each case, if \( f \) is also bounded, then one can obtain the \( \phi_k \) to converge uniformly to \( f \).

Proof. Note that (a) is a special case of (b) (to obtain (a) from (b), choose \( \alpha := 0 \) and, in the bounded case, also choose \( \beta \geq 0 \) to be some upper bound of \( f \) in (b)). We, thus, proceed to prove (b): If \((\phi_k)_{k \in \mathbb{N}} \) is a sequence in \( \mathcal{S} \) that converges pointwise to \( f \), then,
as the simple functions \( \phi_k \) are measurable, \( f \) is measurable by Th. 1.82(a). Moreover, if, for \( x \in X \), \( \alpha \leq \phi_k(x) \) (resp. \( \phi_k(x) \leq \beta \)) for each \( k \in \mathbb{N} \), then, \( \alpha \leq f(x) \) (resp. \( f(x) \leq \beta \)) as well. For the converse, given \( f \), we have to construct suitable simple functions \( \phi_k \). To this end, we define the \( \mathcal{A} \)-measurable sets

\[
\forall k \in \mathbb{N} \quad \forall j \in \{0, 1, \ldots, k \cdot 2^k \} 
A_{k,j} := \begin{cases} 
\{ \alpha + \frac{j}{2^k} \leq f < \alpha + \frac{j+1}{2^k} \} & \text{for } 0 \leq j \leq k \cdot 2^k - 1, \\
\{ f \geq \alpha + k \} & \text{for } j = k \cdot 2^k
\end{cases}
\]

(as \( f \) is assumed to be measurable, the \( A_{k,j} \) are \( \mathcal{A} \)-measurable by Th. 1.79). Clearly, for each fixed \( k \in \mathbb{N} \), the \( A_{k,j} \) form a decomposition of \( X \):

\[
\forall k \in \mathbb{N} \quad X = \bigcup_{j=0}^{k \cdot 2^k} A_{k,j}.
\]

(1.61)

For each \( k \in \mathbb{N} \), we define the simple function

\[
\phi_k : X \rightarrow \mathbb{R}, \quad \phi_k := \sum_{j=0}^{k \cdot 2^k} \left( \alpha + \frac{j}{2^k} \right) \chi_{A_{k,j}}.
\]

Since \( X \) is the disjoint union of the \( A_{k,j} \) by (1.61), this immediately implies

\[
\forall k \in \mathbb{N} \quad \alpha \leq \phi_k \leq \alpha + k \cdot 2^k.
\]

We next verify that the sequence \( (\phi_k)_{k \in \mathbb{N}} \) is increasing: Note that, for each \( k \in \mathbb{N} \), we have the disjoint unions

\[
A_{k,j} = A_{k + 1, 2j} \cup A_{k + 1, 2j + 1} \quad \text{for } 0 \leq j \leq k \cdot 2^k - 1,
\]

\[
A_{k,k \cdot 2^k} = \bigcup_{l=k \cdot 2^k+1}^{(k+1)2^{k+1}} A_{k+1,l}.
\]

Since \( \frac{j}{2^k} \leq \frac{l}{2^{k+1}} \) for \( l \in \{2j, 2j+1\} \), and \( \frac{k \cdot 2^k}{2^k} \leq \frac{l}{2^{k+1}} \) for \( l \in \{k \cdot 2^{k+1}, \ldots, (k+1) \cdot 2^{k+1}\} \), we have \( \phi_k(x) \leq \phi_{k+1}(x) \) for each \( x \in X \). We now fix \( x \in X \) and show \( \lim_{k \to \infty} \phi_k(x) = f(x) \): For each \( k \in \mathbb{N} \), there exists a unique \( j = j(k, x) \) such that \( x \in A_{k,j} \). Then \( \phi_k(x) = \alpha + \frac{j}{2^k} \leq f(x) \), showing \( f(x) \) to be an upper bound for the \( \phi_k(x) \). If \( f(x) = \infty \), then, for each \( k \in \mathbb{N} \), \( \phi_k(x) = \alpha + k \to \infty \). On the other hand, if \( f(x) < \infty \), then, given \( \epsilon > 0 \), we can choose \( k_0 \in \mathbb{N} \) such that \( \frac{1}{2^{k_0}} < \epsilon \) and \( f(x) < \alpha + k_0 \). If \( k \geq k_0 \), then

\[
f(x) - \phi_k(x) < 2^{-k} \leq 2^{-k_0} < \epsilon,
\]

proving \( \lim_{k \to \infty} \phi_k(x) = f(x) \). Under the additional hypothesis \( f \leq \beta \in \mathbb{R} \), one can choose \( \alpha + k_0 \geq \beta \), independently of \( x \), showing uniform convergence. Moreover, in this case, the above observation \( \phi_k(x) \leq f(x) \) for each \( k \in \mathbb{N} \) and each \( x \in X \) also yields \( \phi_k(x) \leq \beta \) for each \( k \in \mathbb{N} \), \( x \in X \).

(c): As before, if \( f = \lim_{k \to \infty} \phi_k \) with \( (\phi_k)_{k \in \mathbb{N}} \) in \( S \), then \( f \in \mathcal{M} \) by Th. 1.82(a). For the converse, we write \( f \in \mathcal{M} \) as \( f = f^+ - f^- \) and, by (a), choose \( (\phi_k)_{k \in \mathbb{N}} \) in \( S^+ \) with \( \phi_k \uparrow f^+ \) and \( (\psi_k)_{k \in \mathbb{N}} \) in \( S^+ \) with \( \psi_k \uparrow f^- \) (with uniform convergence in each case for \( f \) bounded). Then \( f = \lim_{k \to \infty} (\phi_k - \psi_k) \) (with uniform convergence for \( f \) bounded). ■
1.6.6 Products, Projections, and Borel Sets

In [Phi16b, Ex. 1.53(a)], we constructed the product topology on a Cartesian product of topological spaces as the smallest topology that makes all projections continuous. Analogously, we are now going to construct the product $\sigma$-algebra on a Cartesian product of measurable spaces as the smallest $\sigma$-algebra that makes all projections measurable (cf. Th. 1.93(b) below). As in topology, the construction is a special case of so-called initial constructions (cf. Sec. C in the Appendix).

**Definition 1.90.** Let $I$ be a nonempty index set and, for each $i \in I$, let $(X_i, A_i)$ be a measurable space. We consider the Cartesian product $X := \prod_{i \in I} X_i$ and, for each $j \in I$, we recall the projection

$$\pi_j : X \to X_j, \quad \pi_j((x_i)_{i \in I}) := x_j.$$ 

Consider the set

$$E_p := \left\{ \prod_{i \in I} A_i : \left( \forall_{i \in I} A_i \in A_i \right) \land \#\{i \in I : A_i \neq X_i\} \leq 1 \right\}$$

$$= \left\{ \pi_i^{-1}(A_i) : i \in I, A_i \in A_i \right\}.$$ 

Then $A_p := \bigotimes_{i \in I} A_i := \sigma(E_p)$ (the $\sigma$-algebra on $X$ generated by $E_p$), is called the product $\sigma$-algebra on $X$.

**Proposition 1.91.** In the situation of Def. 1.90 above, for each $i \in I$, let $E_i$ be a generator of $A_i$.

(a) The set

$$E_p^* := \left\{ \prod_{i \in I} E_i : \left( \forall_{i \in I} E_i \in E_i \cup \{ X_i \} \right) \land \#\{i \in I : E_i \neq X_i\} \leq 1 \right\}$$

$$= \left\{ \pi_i^{-1}(E_i) : i \in I, E_i \in E_i \right\} \cup \{ X \}$$

is a generator of the product $\sigma$-algebra on $X$, i.e. $A_p = \sigma(E_p^*)$.

(b) If $I$ is finite and, for each $i \in I$, there exists a sequence $(E_{i,k})_{k \in \mathbb{N}}$ in $E_i$ such that

$$X_i = \bigcup_{k=1}^{\infty} E_{i,k}, \quad (1.62)$$

then the set

$$F_p := \left\{ \prod_{i \in I} E_i : \forall_{i \in I} E_i \in E_i \right\} = \left\{ \bigcap_{i \in I} \pi_i^{-1}(E_i) : \forall_{i \in I} E_i \in E_i \right\}$$

is a generator of the product $\sigma$-algebra on $X$, i.e. $A_p = \sigma(F_p)$. 

Proof. (a): As $\mathcal{E}_p^\ast \subseteq \mathcal{E}_p$, we have $\sigma(\mathcal{E}_p^\ast) \subseteq \sigma(\mathcal{E}_p)$. For the remaining inclusion, we note that, by the definition of $\mathcal{E}_p^\ast$, one has $\pi_i^{-1}(E_i) \subseteq \mathcal{E}_p^\ast$ for each $i \in I$. Moreover, by Th. 1.56(a), $\sigma(\pi_i^{-1}(E_i)) = \pi_i^{-1}(\sigma(E_i)) = \pi_i^{-1}(A_i)$, implying $\sigma(\mathcal{E}_p^\ast) \subseteq \sigma(\mathcal{E}_p^\ast)$.

(b): From (a), we know $\mathcal{A}_p = \sigma(\mathcal{E}_p^\ast)$. If $i \in I$ and $E_i \in \mathcal{E}_i$, then $\pi_i^{-1}(E_i) \in \mathcal{E}_p^\ast$ by the definition of $\mathcal{E}_p^\ast$. Thus, if $E_i \in \mathcal{E}_i$ for each $i \in I$, then $\bigcap_{i \in I} \pi_i^{-1}(E_i) \in \sigma(\mathcal{E}_p^\ast)$, since $I$ is finite. This already yields $\sigma(\mathcal{F}_p) \subseteq \sigma(\mathcal{E}_p^\ast)$ (for this inclusion, we actually did not use (1.62)). For the remaining inclusion, we use the finiteness of $I$ together with (1.62) to write every element of $\mathcal{E}_p^\ast$ as a countable union of elements from $\mathcal{F}_p$.

$$\forall i \in I \quad \forall E_i \in \mathcal{E}_i \cup \{X_i\} \quad \pi_i^{-1}(E_i) = \bigcup_{(\kappa_\alpha)_{\alpha \in I \setminus \{i\}} \in \mathbb{N}^{I \setminus \{i\}}} \left( E_i \times \prod_{\alpha \in I \setminus \{i\}} E_{\alpha,\kappa_\alpha} \right) \in \sigma(\mathcal{F}_p).$$

Thus, $\sigma(\mathcal{E}_p^\ast) \subseteq \sigma(\mathcal{F}_p)$, completing the proof of (b). \hfill \blacksquare

**Example 1.92.** Let $p, q \in \mathbb{N}$. For the Borel sets on $\mathbb{R}^p$ and $\mathbb{R}^q$, we obtain $\mathcal{B}^{p+q} = \mathcal{B}^p \otimes \mathcal{B}^q$ (also see Th. 1.94(c) below): We know from Ex. 1.8(c), that the Borel sets are generated by the compact intervals: $\mathcal{B}^p = \sigma(\mathcal{I}_c^p)$, $\mathcal{B}^q = \sigma(\mathcal{I}_c^q)$, and $\mathcal{B}^{p+q} = \sigma(\mathcal{I}_c^{p+q})$. Thus, we can apply Prop. 1.91(b) with $\mathcal{E}_1 = \mathcal{I}_c^p$, $\mathcal{E}_2 = \mathcal{I}_c^q$, and $\mathcal{F}_p = \mathcal{I}_c^{p+q}$ to obtain

$$\mathcal{B}^{p+q} = \sigma(\mathcal{I}_c^{p+q}) = \sigma(\mathcal{F}_p) = \mathcal{B}^p \otimes \mathcal{B}^q.$$ 

**Theorem 1.93.** Let $I$ be a nonempty index set and, for each $i \in I$, let $(X_i, \mathcal{A}_i)$ be a measurable space. We consider the Cartesian product $X := \prod_{i \in I} X_i$ with the product $\sigma$-algebra $\mathcal{A}_p = \sigma(\mathcal{E}_p)$ according to Def. 1.90.

(a) Each projection $\pi_i : X \rightarrow X_i, i \in I$, is $\mathcal{A}_p$-$\mathcal{A}_i$-measurable.

(b) $\mathcal{A}_p$ is the smallest $\sigma$-algebra $\mathcal{A}$ on $X$ such that all projections $\pi_i, i \in I$, are $\mathcal{A}$-$\mathcal{A}_i$-measurable.

(c) If $(Y, \mathcal{B})$ is a measurable space and $f : Y \rightarrow X$, then the following statements are equivalent:

(i) $f$ is $\mathcal{B}$-$\mathcal{A}_p$-measurable.

(ii) For each $i \in I$, the coordinate function $f_i = \pi_i \circ f$ is $\mathcal{B}$-$\mathcal{A}_i$-measurable.

Proof. (a): If $i \in I$ and $A \in \mathcal{A}_i$, then $\pi_i^{-1}(A) \in \mathcal{E}_p \subseteq \mathcal{A}_p$, showing $\pi_i$ to be $\mathcal{A}_p$-$\mathcal{A}_i$-measurable.

(b) is immediate, since each $\sigma$-algebra $\mathcal{A}$ on $X$ such that all projections $\pi_i, i \in I$, are $\mathcal{A}$-$\mathcal{A}_i$-measurable must contain $\mathcal{E}_p$.

(c): If $f$ is $\mathcal{B}$-$\mathcal{A}_p$-measurable, then each $f_i, i \in I$, is $\mathcal{B}$-$\mathcal{A}_i$-measurable by (a) and Prop. 1.68. For the converse, assume each $f_i, i \in I$, to be $\mathcal{B}$-$\mathcal{A}_i$-measurable. According to Prop. 1.66, it suffices to show $f^{-1}(\pi_i^{-1}(A)) \in \mathcal{B}$ for each $i \in I$ and each $A \in \mathcal{A}_i$. Since $f^{-1}(\pi_i^{-1}(A)) = f_i^{-1}(A)$ and $f_i$ is measurable, the proof is, thus, complete. \hfill \blacksquare
When considering the product \((X, \mathcal{T})\) of topological spaces \((X_i, \mathcal{T}_i), i \in I\), where \(\mathcal{T}\) is the corresponding product topology, then one can obtain a \(\sigma\)-algebra \(\mathcal{A}_\mathcal{T}\) on \(X\) by taking the Borel sets with respect to \(\mathcal{T}\). Alternatively, one can take the Borel \(\sigma\)-algebras \(\mathcal{A}_i\) on \(X_i\) and then form the resulting product \(\sigma\)-algebra \(\mathcal{A}_p := \bigotimes_{i \in I} \mathcal{A}_i\) on \(X\). An obvious question is, whether the results are the same. Unfortunately, in general, they are not: While one always has \(\mathcal{A}_p \subseteq \mathcal{A}_\mathcal{T}\) (see Th. 1.94(a) below), the inclusion can be strict (even for the product of just two spaces, see, e.g., [Els07, Rem. III.5.16, Problem III.5.3]). A sufficient condition for \(\mathcal{A}_p = \mathcal{A}_\mathcal{T}\) will be given in Th. 1.94(b).

**Theorem 1.94.** Let \(I\) be a nonempty index set and, for each \(i \in I\), let \((X_i, \mathcal{T}_i)\) be a topological space. Let \(\mathcal{A}_i\) denote the corresponding Borel \(\sigma\)-algebra on \(X_i\). Moreover, let \(X := \prod_{i \in I} X_i\), let \(\mathcal{T}\) denote the product topology on \(X\), let \(\mathcal{A}_{\mathcal{T}}\) be the resulting Borel \(\sigma\)-algebra on \(X\), and let \(\mathcal{A}_p := \bigotimes_{i \in I} \mathcal{A}_i\) be the product \(\sigma\)-algebra of the \(\mathcal{A}_i\).

(a) One always has \(\mathcal{A}_p \subseteq \mathcal{A}_{\mathcal{T}}\).

(b) If \(I\) is countable and each \((X_i, \mathcal{T}_i)\) is a \(C_2\) space (i.e. each \(\mathcal{T}_i\) has a countable base), then \(\mathcal{A}_p = \mathcal{A}_{\mathcal{T}}\).

(c) Let \(p, q \in \mathbb{N}\), \(X \subseteq \mathbb{R}^p\), \(Y \subseteq \mathbb{R}^q\). Then

\[
\mathcal{B}^p|X \otimes \mathcal{B}^q|Y = \mathcal{B}^{p+q}|(X \times Y). \tag{1.63}
\]

In particular, \(\mathcal{B}^{p+q} = \mathcal{B}^p \otimes \mathcal{B}^q\) and \(\mathcal{B}^p = \otimes_{k=1}^{p} B^1\).

**Proof.** (a): Since each projection \(\pi_i, i \in I\), is continuous, it is \(\pi_i^{-1}(A) \in \mathcal{T}\) for each \(A \in \mathcal{T}_i\), implying it to be \(\mathcal{A}_i\)-\(\mathcal{A}_{\mathcal{T}}\)-measurable. Then \(\mathcal{A}_p \subseteq \mathcal{A}_{\mathcal{T}}\) follows from Th. 1.93(b).

(b): Due to (a), it remains to show \(\mathcal{A}_{\mathcal{T}} \subseteq \mathcal{A}_p\). For each \(i \in I\), let \(\mathcal{B}_i\) be a countable base of \(\mathcal{T}_i\). According to [Phi16b, Ex. 1.53(a)], a base for \(\mathcal{T}\) is given by

\[
\mathcal{B} := \left\{ \bigcap_{j \in J} \pi_j^{-1}(B_j) : J \subseteq I, 0 < \#J < \infty, \forall j \in J, B_j \in \mathcal{B}_j \right\}.
\]

Since \(I\) and each \(\mathcal{B}_i\) are countable, so is \(\mathcal{B}\). Since \(\mathcal{B}_i \subseteq \mathcal{A}_i\), we have \(\mathcal{B} \subseteq \mathcal{A}_p\). Now every \(O \in \mathcal{T}\) is a countable union of sets from \(\mathcal{B}\) (since \(\mathcal{B}\) is countable), showing \(\mathcal{T} \subseteq \sigma(\mathcal{B}) \subseteq \mathcal{A}_p\) and, thus, \(\mathcal{A}_T = \sigma(\mathcal{T}) \subseteq \mathcal{A}_p\), as desired.

(c): Let \(\mathcal{T}_p\), \(\mathcal{T}_q\), and \(\mathcal{T}\) denote the norm topologies on \(\mathbb{R}^p\), \(\mathbb{R}^q\), and \(\mathbb{R}^{p+q}\), respectively. As we know the norm topologies to be \(C_2\), we can apply (b) to obtain

\[
\mathcal{B}^p|X \otimes \mathcal{B}^q|Y = \sigma(\mathcal{T}_p)|X \otimes \sigma(\mathcal{T}_q)|Y \tag{1.56(b)} = \sigma_X(\mathcal{T}_p|X) \otimes \sigma_Y(\mathcal{T}_q|Y) \\
= \sigma_X((\mathcal{T}_p|X) \otimes \sigma_Y((\mathcal{T}_q|Y)) \tag{b} = \sigma_X(X) \otimes \sigma_Y(Y) \\
= \sigma_X \otimes \sigma_Y(\mathcal{T}|(X \times Y)) \tag{Th. 1.56(b)} = \sigma(\mathcal{T})(X \times Y)
\]

as claimed. \(\blacksquare\)
Caveat 1.95. While we know from [Phi16b, Ex. 2.12(b)] that every projection is an open map with respect to the product topology (i.e. projections map open sets to open sets), the analogous statement is, in general, not true for product $\sigma$-algebras: For example, it can occur that the projection of a Borel set $B \subseteq [0,1]^2$ is not in $\mathcal{B}^1|[0,1]$ (projections of such Borel sets are known as Suslin sets – the construction of a Suslin set that is not a Borel set needs some work and can be found in [Beh87, Appendix II]).

2 Integration

2.1 Integration of $\mathbb{R}$-Valued Measurable Maps

We define the so-called Lebesgue integral first for nonnegative simple functions in Def. 2.1, then for nonnegative measurable functions in Def. 2.4, and then for so-called integrable functions in Def. 2.11.

2.1.1 Simple Functions

The definition of the Lebesgue integral for nonnegative simple functions makes use of representations of the integrands, where Lem. 2.2 then states that the value of the integral does actually not depend on the representation of the integrated function.

**Definition 2.1.** Let $(X, \mathcal{A}, \mu)$ be a measure space. Moreover, let $f : X \rightarrow \mathbb{R}^+_0$ be a simple function (i.e. $f \in \mathcal{S}^+(\mathcal{A})$), where

$$f = \sum_{i=1}^{N} \alpha_i \chi_{A_i}, \quad (2.1)$$

with $N \in \mathbb{N}; \alpha_1, \ldots, \alpha_N \geq 0$; and $A_1, \ldots, A_N \in \mathcal{A}$. Then

$$\int_X f \, d\mu := \int_X f(x) \, d\mu(x) := \sum_{i=1}^{N} \alpha_i \mu(A_i) \in [0, \infty] \quad (2.2)$$

is called the Lebesgue integral of $f$ over $X$ with respect to $\mu$.

**Lemma 2.2.** Let $(X, \mathcal{A}, \mu)$ be a measure space. Then the value of the integral in (2.2) does not depend on the representation of the simple function $f$: If $M, N \in \mathbb{N}; \alpha_1, \ldots, \alpha_N \geq 0; \beta_1, \ldots, \beta_M \geq 0; A_1, \ldots, A_N \in \mathcal{A}$; and $B_1, \ldots, B_M \in \mathcal{A}$ are such that

$$f = \sum_{i=1}^{N} \alpha_i \chi_{A_i} = \sum_{i=1}^{M} \beta_i \chi_{B_i}, \quad (2.3a)$$

then

$$\sum_{i=1}^{N} \alpha_i \mu(A_i) = \sum_{i=1}^{M} \beta_i \mu(B_i). \quad (2.3b)$$
Proof. We add the list of $B_i$ to the list of $A_i$ by letting $A_{N+1} := B_1, \ldots, A_{N+M} := B_M$, and define the set of intersections

$$\mathcal{D} := \left\{ \bigcap_{i=1}^{N+M} M_i : \forall_{i \in \{1, \ldots, N+M\}} M_i \in \{A_i, A_i^c\} \right\}.$$ 

Note that two distinct elements of $\mathcal{D}$ are always disjoint: Indeed, if $A, B \in \mathcal{D}$ with $A \neq B$, then there exists $i_0 \in \{1, \ldots, N+M\}$ such that $A \subseteq A_{i_0}$ and $B \subseteq A_{i_0}^c$. Next, we consider $i \in \{1, \ldots, N+M\}$ and $x \in A_i$. Then, for each $j \in \{1, \ldots, N+M\}$, either $x \in A_j$ or $x \in A_j^c$, implying

$$A_i = \bigcup \left\{ \bigcap_{j=1}^{N+M} M_j : M_j = A_i \land \forall_{j \in \{1, \ldots, N+M\}} M_j \in \{A_j, A_j^c\} \right\}.$$ 

Thus, if $C_1, \ldots, C_K, K \in \mathbb{N}$, is an enumeration of the distinct elements of $\mathcal{D}$, then we have the disjoint union

$$A_i = \bigcup_{j \in \{1, \ldots, K\}: C_j \subseteq A_i} C_j,$$

implying

$$f = \sum_{i=1}^{K} \gamma_i \chi_{C_i}, \quad \text{where} \quad \gamma_i = \sum_{j \in \{1, \ldots, N\}: C_j \subseteq A_i} \alpha_j = \sum_{j \in \{1, \ldots, M\}: C_j \subseteq B_i} \beta_j$$

and

$$\sum_{i=1}^{N} \alpha_i \mu(A_i) = \sum_{i=1}^{N} \alpha_i \sum_{j \in \{1, \ldots, K\}: C_j \subseteq A_i} \mu(C_j) = \sum_{j=1}^{K} \gamma_j \mu(C_j) = \sum_{i=1}^{M} \beta_i \sum_{j \in \{1, \ldots, K\}: C_j \subseteq B_i} \mu(C_j) = \sum_{i=1}^{M} \beta_i \mu(B_i),$$

proving (2.3b). □

Lemma 2.3. Let $(X, A, \mu)$ be a measure space.

(a) For each $A \in A$, one has $\int_A \chi_A \, d\mu = \mu(A)$.

(b) For each $f, g \in \mathcal{S}^+(A)$ and each $\alpha, \beta \in \mathbb{R}_0^+$, one has

$$\int_X (\alpha f + \beta g) \, d\mu = \alpha \int_X f \, d\mu + \beta \int_X g \, d\mu.$$ 

(c) For each $f, g \in \mathcal{S}^+(A)$, one has

$$f \leq g \quad \Rightarrow \quad \int_X f \, d\mu \leq \int_X g \, d\mu.$$
Proof. (a) is immediate from Def. 2.1.

(b): If \( f = \sum_{i=1}^{N} \alpha_i \chi_{A_i} \) and \( g = \sum_{i=1}^{M} \beta_i \chi_{B_i} \) with \( N, M \in \mathbb{N}; \alpha_1, \ldots, \alpha_N, \beta_1, \ldots, \beta_M \geq 0; \) and \( A_1, \ldots, A_N, B_1, \ldots, B_M \in \mathcal{A} \), then

\[
\alpha f + \beta g = \sum_{i=1}^{N} \alpha_i \chi_{A_i} + \sum_{i=1}^{M} \beta_i \chi_{B_i},
\]

implying

\[
\alpha \int_X f \, d\mu + \beta \int_X g \, d\mu = \sum_{i=1}^{N} \alpha_i \mu(A_i) + \sum_{i=1}^{M} \beta_i \mu(B_i) = \int_X (\alpha f + \beta g) \, d\mu,
\]

proving (b).

(c): If \( f \leq g \), then \( g - f \in \mathcal{S}^+(\mathcal{A}) \). Thus, (b) applies to \( g = f + (g - f) \), yielding

\[
\int_X g \, d\mu = \int_X f \, d\mu + \int_X (g - f) \, d\mu \geq \int_X f \, d\mu,
\]

as desired. \( \blacksquare \)

Note that, so far, we have not used the \( \sigma \)-additivity of the measure \( \mu \) – it would have sufficed to take \( \mu \) to be a content on an algebra. However, this will change in the following section.

2.1.2 Nonnegative Functions

As for nonnegative simple functions above, the definition of the Lebesgue integral for nonnegative measurable functions also makes use of representations of the integrand, where, once again, a subsequent lemma (Lem. 2.5) then states that the value of the integral does actually not depend on the representation of the integrated function.

Definition 2.4. Let \((X, \mathcal{A}, \mu)\) be a measure space. Moreover, let \( f : X \rightarrow [0, \infty] \) be measurable (i.e. \( f \in \mathcal{M}^+(\mathcal{A}) \)), where \((\phi_i)_{i \in \mathbb{N}}\) is a sequence in \( \mathcal{S}^+(\mathcal{A}) \) such that \( \phi_i \uparrow f \). Then

\[
\int_X f \, d\mu := \int_X f(x) \, d\mu(x) := \lim_{i \to \infty} \int_X \phi_i \, d\mu \in [0, \infty] \tag{2.4}
\]

is called the Lebesgue integral of \( f \) over \( X \) with respect to \( \mu \) (note that the limit in (2.4) exists, since \( \int_X \phi \, d\mu \geq \int_X \psi \, d\mu \) for simple functions \( \phi \geq \psi \) due to Lem. 2.3(c)).

Lemma 2.5. Let \((X, \mathcal{A}, \mu)\) be a measure space.

(a) If \((u_i)_{i \in \mathbb{N}}\) is an increasing sequence in \( \mathcal{S}^+(\mathcal{A}) \) and \( v \in \mathcal{S}^+(\mathcal{A}) \), then

\[
v \leq \lim_{i \to \infty} u_i \Rightarrow \int_X v \, d\mu \leq \lim_{i \to \infty} \int_X u_i \, d\mu. \tag{2.5}
\]
(b) The value of the integral in (2.4) does not depend on the representation of the non-negative measurable function \( f \in M^+ (A) \): If \((\phi_i)_{i \in \mathbb{N}}\) and \((\psi_i)_{i \in \mathbb{N}}\) both are sequences in \( S^+ (A) \) with \( \phi_i \uparrow f \) and \( \psi_i \uparrow f \), then

\[
\lim_{i \to \infty} \int_X \phi_i \, d\mu = \lim_{i \to \infty} \int_X \psi_i \, d\mu.
\] (2.6)

In particular, (2.2) and (2.4) are consistently defined.

Proof. (a): Since \((u_i)_{i \in \mathbb{N}}\) is increasing, \( \lim_{i \to \infty} \int_X u_i \, d\mu \) exists in \([0, \infty]\) due to Lem. 2.3(c). Fix \( \beta \in [1, \infty[ \) and, for each \( n \in \mathbb{N} \), define \( B_n := \{ \beta u_n \geq v \} \). The measurability of \( v \) and \( u_n \) implies \( B_n \in A \). Let \( x \in X \). If \( v(x) = 0 \), then \( x \in B_n \) for each \( n \in \mathbb{N} \). If \( v(x) > 0 \), then \( v(x) \leq \lim_{i \to \infty} u_i (x) \) and \( \beta > 1 \) yield the existence of \( N \in \mathbb{N} \) such that \( x \in B_n \) for each \( n \geq N \). Thus, \( B_n \uparrow X \) (since \((u_i)_{i \in \mathbb{N}}\) is increasing). We now write \( v = \sum_{i=1}^N \alpha_i \chi_{A_i} \) with \( \alpha_i \in \mathbb{R}^+ \), \( A_i \in A \), \( N \in \mathbb{N} \), and use \( \mu \) being continuous from below to compute

\[
\int_X v \, d\mu = \sum_{i=1}^N \alpha_i \mu(A_i) = \lim_{n \to \infty} \sum_{i=1}^N \alpha_i \mu(A_i \cap B_n) = \lim_{n \to \infty} \int_X v \cdot \chi_{B_n} \, d\mu
\]

\[
\leq \lim_{n \to \infty} \int_X \beta u_n \cdot \chi_{B_n} \, d\mu \leq \beta \lim_{n \to \infty} \int_X u_n \, d\mu.
\] (2.7)

As (2.7) holds for each \( \beta > 1 \), it must still hold for \( \beta = 1 \), proving (2.5).

(b): For each \( k \in \mathbb{N} \), we have \( \phi_k \leq f = \lim_{i \to \infty} \psi_i \). Thus, by (a),

\[
\forall \ k \in \mathbb{N} \quad \int_X \phi_k \, d\mu \leq \lim_{i \to \infty} \int_X \psi_i \, d\mu,
\]

implying

\[
\lim_{k \to \infty} \int_X \phi_k \, d\mu \leq \lim_{i \to \infty} \int_X \psi_i \, d\mu.
\] (2.8)

Interchanging the roles of \( \phi_i \) and \( \psi_i \) above, yields (2.8) with the inequality reversed, proving (2.6).

Lemma 2.6. Let \((X, A, \mu)\) be a measure space.

(a) For each \( f \in M^+ (A) \), one has

\[
\int_X f \, d\mu = 0 \iff \mu(\{ f > 0 \}) = 0.
\]

(b) For each \( f, g \in M^+ (A) \) and each \( \alpha, \beta \in [0, \infty] \), one has

\[
\int_X (\alpha f + \beta g) \, d\mu = \alpha \int_X f \, d\mu + \beta \int_X g \, d\mu.
\]
(c) For each \( f, g \in \mathcal{M}^+(A) \), one has
\[
f \leq g \Rightarrow \int_X f \, d\mu \leq \int_X g \, d\mu.
\]

(d) For each \( f \in \mathcal{M}^+(A) \), one has
\[
\int_X f \, d\mu = \sup \left\{ \int_X \phi \, d\mu : \phi \in \mathcal{S}^+(A), \phi \leq f \right\}.
\]

**Proof.** (a): Choose a sequence \( (\phi_n)_{n \in \mathbb{N}} \) in \( \mathcal{S}^+(A) \) such that \( \phi_n \uparrow f \). Set \( A := \{ f > 0 \} \). As \( f \) is measurable, we have \( A \in A \). If \( \mu(A) > 0 \), then, letting \( A_n := \{ f > \frac{1}{n} \} \), we have \( A_n \uparrow A \) and, thus, \( \mu(A_{n_0}) > 0 \) for some \( n_0 \in \mathbb{N} \). In consequence,
\[
\int_X f \, d\mu = \lim_{n \to \infty} \int_X \phi_n \, d\mu \overset{\text{Lem. 2.5(a)}}{\geq} \lim_{n \to \infty} \int_X \frac{1}{n} \chi_{A_n} \, d\mu \geq \frac{1}{n_0} \mu(A_{n_0}) > 0.
\]
showing \( \infty \cdot \int_X f \, d\mu = \infty = \int_X \infty \cdot f \, d\mu \). If \( \mu(A) = 0 \), then \( \mu(\{ \phi_n > 0 \}) \leq \mu(A) = 0 \) for each \( n \in \mathbb{N} \), implying \( \int_X f \, d\mu = \lim_{n \to \infty} \int_X \phi_n \, d\mu = 0 \).

(b): Choose sequences \( (\phi_i)_{i \in \mathbb{N}} \) and \( (\psi_i)_{i \in \mathbb{N}} \) in \( \mathcal{S}^+(A) \) such that \( \phi_i \uparrow f \) and \( \psi_i \uparrow f \). Then \( (\phi_i + \psi_i) \uparrow (f + g) \), implying
\[
\int_X (f + g) \, d\mu = \lim_{i \to \infty} \int_X (\phi_i + \psi_i) \, d\mu \overset{\text{Lem. 2.3(b)}}{=} \lim_{i \to \infty} \int_X \phi_i \, d\mu + \lim_{i \to \infty} \int_X \psi_i \, d\mu \\
= \int_X f \, d\mu + \int_X g \, d\mu.
\]

Similarly, if \( \alpha \in \mathbb{R}_0^+ \), then \( \alpha \phi_i \in \mathcal{S}^+(A) \) for each \( i \in \mathbb{N} \), \( (\alpha \phi_i) \uparrow (\alpha f) \), and
\[
\int_X \alpha f \, d\mu = \lim_{i \to \infty} \int_X \alpha \phi_i \, d\mu \overset{\text{Lem. 2.3(b)}}{=} \alpha \lim_{i \to \infty} \int_X \phi_i \, d\mu = \alpha \int_X f \, d\mu.
\]
It remains to consider the case \( \alpha = \infty \). As in the proof of (a), set \( A := \{ f > 0 \} \). Then \( n \cdot \chi_A \uparrow \infty \cdot \chi_A \), implying
\[
\int_X \infty \cdot f \, d\mu = \lim_{n \to \infty} (n \cdot \mu(A)) = \begin{cases} \infty & \text{for } \mu(A) > 0, \\
0 & \text{for } \mu(A) = 0. \end{cases}
\]
If \( \mu(A) > 0 \), then \( \int_X f \, d\mu > 0 \) by (a), showing \( \infty \cdot \int_X f \, d\mu = \infty = \int_X \infty \cdot f \, d\mu \). If \( \mu(A) = 0 \), then \( \int_X f \, d\mu = 0 \) by (a). Thus, using (A.2e), \( \infty \cdot \int_X f \, d\mu = 0 = \int_X \infty \cdot f \, d\mu \).

(c): If \( f \leq g \), then \( g - f \in \mathcal{M}^+(A) \). Thus, (b) applies to \( g = f + (g - f) \), yielding
\[
\int_X g \, d\mu = \int_X f \, d\mu + \int_X (g - f) \, d\mu \geq \int_X f \, d\mu,
\]
as desired.

(d): While “\( \leq \)” is immediate from Def. 2.4, “\( \geq \)” is due to (c).
Theorem 2.7 (Monotone Convergence). If $(X, A, \mu)$ is a measure space and $(f_i)_{i \in \mathbb{N}}$ is an increasing sequence in $\mathcal{M}^+(A)$, then

$$\int_X \left( \lim_{i \to \infty} f_i \right) \, d\mu = \lim_{i \to \infty} \int_X f_i \, d\mu. \quad (2.9)$$

Proof. Letting $f := \lim_{i \to \infty} f_i$, we have $f_i \uparrow f$ and $f \in \mathcal{M}^+(A)$ by Th. 1.82(a). Now “$\geq$” in (2.9) is due to Lem. 2.6(c). To prove the remaining inequality, consider $\phi \in S^+(A)$, $\phi \leq f$, and $\beta \in ]1, \infty]$. Then, for each $n \in \mathbb{N}$, we obtain $B_n := \{\beta f_n \geq \phi\} \in A$ and $\phi_n := \phi \cdot \chi_{B_n} \in S^+(A)$. Also, clearly, $B_n \uparrow X$ and $\phi_n \uparrow \phi$. Then Lem. 2.5(a) yields

$$\int_X \phi \, d\mu \leq \lim_{n \to \infty} \int_X \phi \cdot \chi_{B_n} \, d\mu \leq \beta \lim_{n \to \infty} \int_X f_n \, d\mu$$

and $\int_X \phi \, d\mu \leq \lim_{n \to \infty} \int_X f_n \, d\mu$, since $\beta > 1$ was arbitrary. Now Lem. 2.6(d) yields “$\leq$” in (2.9) and completes the proof. ■

Corollary 2.8. If $(X, A, \mu)$ is a measure space and $(g_i)_{i \in \mathbb{N}}$ is a sequence in $\mathcal{M}^+(A)$, then

$$\int_X \left( \sum_{i=1}^{\infty} g_i \right) \, d\mu = \sum_{i=1}^{\infty} \int_X g_i \, d\mu. \quad (2.10)$$

Proof. Letting, for each $n \in \mathbb{N}$, $f_n := \sum_{i=1}^{n} g_i$, (2.10) is precisely (2.9). ■

Example 2.9. (a) Theorem 2.7 does, in general, not hold without the monotonicity assumption of the $f_n$. Consider $(X, A, \mu) := (\mathbb{R}, B^1, \beta^1)$ with $f_n := \frac{1}{n} \chi_{[0, n]}$. Then $f_n \to 0$ (even uniformly), but

$$\int_X f_n \, d\mu = \frac{1}{n} \cdot n = 1 \not\to 0 = \int_X 0 \, d\mu.$$  

(b) Let $X := \mathbb{N}$, $A := \mathcal{P}(X)$, and let $\mu$ be counting measure. Then $\mathcal{M}^+(A) = \mathcal{F}(\mathbb{N}, [0, \infty]) = [0, \infty]^{\mathbb{N}}$ (all functions are measurable), and we claim that

$$\forall f \in \mathcal{M}^+(A) \quad \int_X f \, d\mu = \sum_{n=1}^{\infty} f(n) :$$

Indeed, letting $g_n := f(n) \cdot \chi_{\{n\}}$ and using Cor. 2.8, we have

$$\int_X f \, d\mu = \sum_{n=1}^{\infty} \int_X g_n \, d\mu = \sum_{n=1}^{\infty} (f(n) \cdot \mu(\{n\})) = \sum_{n=1}^{\infty} f(n).$$

2.1.3 Integrable Functions

We are already familiar with the notation $\mathbb{K}$ to denote either $\mathbb{R}$ or $\mathbb{C}$. Similarly, it will sometimes be useful to treat $\overline{\mathbb{R}}$ and $\mathbb{C}$ in parallel, giving rise to the following notation:
Notation 2.10. We use the symbol $\hat{K}$ to denote either $\mathbb{R}$ or $\mathbb{C}$; we write $(\hat{K}, \hat{B})$ to denote either $(\mathbb{R}, \mathcal{B})$ or $(\mathbb{C}, \mathcal{B}^2)$.

Definition 2.11. Let $(X, \mathcal{A}, \mu)$ be a measure space and let $f : X \to \hat{K}$ be measurable.

(a) $f$ is called integrable (or, more precisely, $\mu$-integrable) if, and only if,
\[
\int_X (\text{Re } f)^+ \, d\mu < \infty, \quad \int_X (\text{Re } f)^- \, d\mu < \infty, \\
\int_X (\text{Im } f)^+ \, d\mu < \infty, \quad \int_X (\text{Im } f)^- \, d\mu < \infty.
\] (2.11)

For integrable functions $f$,
\[
\int_X f \, d\mu := \int_X f(x) \, d\mu(x) \\
:= \int_X (\text{Re } f)^+ \, d\mu - \int_X (\text{Re } f)^- \, d\mu \\
+ i \int_X (\text{Im } f)^+ \, d\mu - i \int_X (\text{Im } f)^- \, d\mu \in \hat{K}
\] (2.12)
is called the Lebesgue integral of $f$ over $X$ with respect to $\mu$.

(b) Let $A \in \mathcal{A}$. If $f \in \mathcal{M}^+(\mathcal{A})$ or $f$ is integrable, then define
\[
\int_A f \, d\mu := \int_A f(x) \, d\mu(x) := \int_X f \chi_A \, d\mu.
\] (2.13)

Remark 2.12. Let $(X, \mathcal{A}, \mu)$ be a measure space.

(a) A measurable function $f : X \to \mathbb{R}$ is integrable if, and only if, both $f^+$ and $f^-$ are integrable and, in that case,
\[
\int_X f \, d\mu = \int_X f^+ \, d\mu - \int_X f^- \, d\mu.
\]

(b) A measurable function $f : X \to \mathbb{C}$ is integrable if, and only if, both $\text{Re } f$ and $\text{Im } f$ are integrable and, in that case,
\[
\int_X f \, d\mu = \int_X \text{Re } f \, d\mu + i \int_X \text{Im } f \, d\mu,
\]
\[
\text{Re} \left( \int_X f \, d\mu \right) = \int_X \text{Re } f \, d\mu, \quad \text{Im} \left( \int_X f \, d\mu \right) = \int_X \text{Im } f \, d\mu.
\]

Lemma 2.13. Let $(X, \mathcal{A}, \mu)$ be a measure space, let $A \in \mathcal{A}$, and let $f : X \to \hat{K}$.

(a) $f \cdot \chi_A$ is $\mathcal{A}$-measurable if, and only if, $f|_A$ is $(\mathcal{A}|A)$-measurable.
(b) It is $f \cdot \chi_A \in \mathcal{M}^+(A)$ if, and only if, $f|_A \in \mathcal{M}^+(A|A)$, and, in that case,
\[
\int_A f \, d\mu = \int_X f \cdot \chi_A \, d\mu = \int_A f|_A \, d(\mu|_{A|A}) \tag{2.14}
\]
(note that we can restrict $\mu$ to $A|A$, since $A \in \mathcal{A}$).

(c) It is $f \cdot \chi_A \mu$-integrable if, and only if, $f|_A$ is $(\mu|_{A|A})$-integrable, and, in that case, (2.14) holds.

**Proof.** As $A \in \mathcal{A}$, we have $\mathcal{A}|A = \{ C \in \mathcal{A} : C \subseteq A \}$.

(a): If $B \in \hat{\mathcal{B}}$, then

\[
(f|_A)^{-1}(B) = A \cap f^{-1}(B), \quad (f \cdot \chi_A)^{-1}(B) = \begin{cases} 
A^c \cup (A \cap f^{-1}(B)) & \text{if } 0 \notin B, \\
A \cap f^{-1}(B) & \text{if } 0 \in B.
\end{cases}
\]

Since $A^c \in \mathcal{A}$, we obtain $(f \cdot \chi_A)^{-1}(B) \in \mathcal{A}$ if, and only if, $(f|_A)^{-1}(B) \in \mathcal{A}|A$, proving (a).

(b): Since $f \cdot \chi_A \equiv 0$ on $A^c$, the equivalence claimed in (b) is clear. If $f \cdot \chi_A \in \mathcal{M}^+(A)$, then let $(\phi_n)$ be a sequence in $\mathcal{S}^+(A)$ with $\phi_n \uparrow (f \cdot \chi_A)$. Then $(\phi_n|_A) \uparrow (f|_A)$ (with $\phi_n|_A \in \mathcal{S}^+(A|A)$) and

\[
\int_A f \, d\mu = \int_X f \cdot \chi_A \, d\mu = \lim_{n \to \infty} \int_X \phi_n \, d\mu = \lim_{n \to \infty} \int_A (\phi_n|_A) \, d(\mu|_{A|A}) = \int_A f|_A \, d(\mu|_{A|A}),
\]

proving (b).

(c) is a direct consequence of (b). 

**Theorem 2.14.** Let $(X, \mathcal{A}, \mu)$ be a measure space, $f : X \to \hat{K}$. Then the following statements are equivalent:

(i) $f$ is integrable.

(ii) $\text{Re} \, f$ and $\text{Im} \, f$ are both integrable.

(iii) $(\text{Re} \, f)^+, (\text{Re} \, f)^-, (\text{Im} \, f)^+, (\text{Im} \, f)^-$ all are integrable.

(iv) There exist integrable functions $p, q, r, s \in \mathcal{M}^+(\mathcal{A})$ such that $f = p - q + i (r - s)$.

(v) $f$ is measurable and there exists an integrable $g \in \mathcal{M}^+(\mathcal{A})$ such that $|f| \leq g$ (i.e. $|f|$ is dominated by $g$).

(vi) $f$ is measurable and $|f|$ is integrable.
Proof. The equivalences of (i) – (iii) are immediate from Def. 2.11(a).

(iii) implies (iv) by setting $p := (\text{Re } f)^+, \ q := (\text{Re } f)^-, \ r := (\text{Im } f)^+, \ s := (\text{Im } f)^-.$

(iv) ⇒ (v): If $f = p - q + i(r - s)$ with integrable $p, q, r, s \in \mathcal{M}^+(\mathcal{A})$, then $|f| \leq g := p + q + r + s$, where $g$ is integrable by Lem. 2.6(b).

(v) implies (vi), since $\int_X |f| \, d\mu \leq \int_X g \, d\mu < \infty$.

(vi) implies (iii): Let $h \in \{(\text{Re } f)^+, (\text{Re } f)^-, (\text{Im } f)^+, (\text{Im } f)^-\}$. Then $h$ is measurable, since $f$ is measurable, and $h$ is integrable, since $\int_X h \, d\mu \leq \int_X |f| \, d\mu < \infty$.

Example 2.15. It can happen that $|f|$ is integrable, but $f$ is not: Let $(X, \mathcal{A}, \mu) := ([0, 1], \mathcal{B}^1, \beta^1), \ A \subseteq X, \ A \notin \mathcal{B}^1$. Then $f := \chi_A - \chi_{A^c}$ is not measurable (in particular, not integrable), but $|f| \equiv 1$ is integrable.

Theorem 2.16. Let $(X, \mathcal{A}, \mu)$ be a measure space, let $f, g : X \to \hat{K}$ be integrable, let $A, B \in \mathcal{A}$, and let $\alpha, \beta \in \mathbb{K}$.

(a) If $f : X \to \hat{R}$ is integrable, then

$$\mu\{|f| = \infty\} = 0.$$  

(b) The Lebesgue integral is linear: $\alpha f + \beta g$ is integrable and

$$\int_A (\alpha f + \beta g) \, d\mu = \alpha \int_A f \, d\mu + \beta \int_A g \, d\mu. \quad (2.15a)$$

(c) If $A, B$ are disjoint, then

$$\int_{A \cup B} f \, d\mu = \int_A f \, d\mu + \int_B f \, d\mu \quad (2.15b)$$

(this also holds for nonintegrable $f \in \mathcal{M}^+(\mathcal{A})$).

(d) The Lebesgue integral is isotone (here, let $f, g$ be $\hat{R}$-valued):

$$f \leq g \quad \Rightarrow \quad \int_A f \, d\mu \leq \int_A g \, d\mu. \quad (2.15c)$$

(e) The Lebesgue integral satisfies the triangle inequality:

$$\left| \int_A f \, d\mu \right| \leq \int_A |f| \, d\mu. \quad (2.15d)$$

(f) Mean Value Theorem for Integration: Let $f$ be $\hat{R}$-valued and assume there exist numbers $m, M \in \mathbb{R}$ such that $m \leq f \leq M$ on $A$. Moreover, let $p : X \to [0, \infty]$ be such that $\chi_A p$ and $\chi_A fp$ are integrable. Then

$$m \int_A p \, d\mu \leq \int_A fp \, d\mu \leq M \int_A p \, d\mu. \quad (2.15e)$$
In particular, if \( f(A) = [m, M] \) (e.g., if \( f \) is \( \mathbb{R} \)-valued and continuous, \( A \) is compact and connected, \( m = \min f(A) \), \( M = \max f(A) \)), then

\[
\exists_{\xi \in A} \int_A f_{p \, d\mu} = f(\xi) \int_A p \, d\mu . \quad (2.15f)
\]

Returning to a general integrable \( f \), if \( p \equiv 1 \) and \( \mu(A) < \infty \), then we obtain the theorem’s classical form:

\[
m \mu(A) \leq \int_A f \, d\mu \leq M \mu(A) , \quad (2.15g)
\]

where the theorem’s name comes from the fact that, for \( 0 < \mu(A) < \infty \),

\[
\int_A f \, d\mu := \mu(A)^{-1} \int_A f \, d\mu \quad (2.15h)
\]

is sometimes referred to as the mean value of \( f \) on \( A \).

**Proof.** (a): Letting \( A_\infty := \{|f| = \infty\} \), we have \( A_\infty \in \mathcal{A} \) and \( h := \infty \cdot \chi_{A_\infty} \leq |f| \). Since \( h, |f| \in \mathcal{M}^+(\mathcal{A}) \), we apply Lem. 2.6(c) to obtain \( \infty \cdot \mu(A_\infty) = \int_X h \, d\mu \leq \int_X |f| \, d\mu < \infty \), implying \( \mu(A_\infty) = 0 \).

(b): It suffices to consider \( A = X \) due to Lem. 2.13. Since \( |\alpha f + \beta g| \leq |\alpha||f| + |\beta||g| \), \( \alpha f + \beta g \) is integrable by Th. 2.14(v). If \( f, g \) are \( \mathbb{R} \)-valued, then \( h := f + g = f^+ - f^- + g^+ - g^- = h^- + f^+ + g^+ \). Thus, due to Lem. 2.6(b),

\[
\int_X h^+ \, d\mu + \int_X f^- \, d\mu + \int_X g^- \, d\mu = \int_X h^- \, d\mu + \int_X f^+ \, d\mu + \int_X g^+ \, d\mu . \quad (2.16)
\]

Due to integrability, all integrals in (2.16) are finite, implying

\[
\int_X (f + g) \, d\mu = \int_X h^+ \, d\mu - \int_X h^- \, d\mu
\]

\[
= \int_X f^+ \, d\mu + \int_X g^+ \, d\mu - \int_X f^- \, d\mu - \int_X g^- \, d\mu = \int_X f \, d\mu + \int_X g \, d\mu .
\]

If \( f, g \) are \( \mathbb{C} \)-valued, then, since \( \text{Re}(f + g) = \text{Re} f + \text{Re} g \) and \( \text{Im}(f + g) = \text{Im} f + \text{Im} g \),

\[
\int_X (f + g) \, d\mu = \int_X f \, d\mu + \int_X g \, d\mu \quad \text{follows from the} \quad \mathbb{R} \quad \text{-valued case}.
\]

If \( f \) is \( \mathbb{R} \)-valued and \( \alpha \geq 0 \), then \( (\alpha f)^+ = \alpha f^+ \), \( (\alpha f)^- = \alpha f^- \) and we obtain \( \int_X \alpha f \, d\mu = \alpha \int_X f \, d\mu \) due to Lem. 2.6(b). If \( \alpha < 0 \), then \( (\alpha f)^+ = |\alpha| f^- \), \( (\alpha f)^- = |\alpha| f^+ \), and one computes

\[
\int_X \alpha f \, d\mu = |\alpha| \left( \int_X f^- \, d\mu - \int_X f^+ \, d\mu \right) = \alpha \int_X f \, d\mu .
\]
Finally, using the $\mathbb{R}$-valued case, one computes for $\mathbb{C}$-valued $f$ and $\alpha \in \mathbb{C}$,
\[
\int_X (\alpha f) = \int_X (\operatorname{Re} \alpha \operatorname{Re} f - \operatorname{Im} \alpha \operatorname{Im} f) \, d\mu + i \int_X (\operatorname{Re} \alpha \operatorname{Im} f + \operatorname{Im} \alpha \operatorname{Re} f) \, d\mu \\
= \operatorname{Re} \alpha \int_X \operatorname{Re} f \, d\mu - \operatorname{Im} \alpha \int_X \operatorname{Im} f \, d\mu + i \operatorname{Re} \alpha \int_X \operatorname{Im} f \, d\mu + i \operatorname{Im} \alpha \int_X \operatorname{Re} f \, d\mu \\
= \alpha \int_X f \, d\mu,
\]
proving (b).

(c): Since $f \chi_{A \cup B} = f \chi_A + f \chi_B$, the case $f \in \mathcal{M}^+(A)$ follows from Lem. 2.6(b) and the integrable case follows from (b).

(d): If $f \leq g$, then $g - f \geq 0$. Thus, (b) applies to $g = f + (g - f)$, yielding
\[
\int_A g \, d\mu = \int_A f \, d\mu + \int_A (g - f) \, d\mu \geq \int_A f \, d\mu,
\]
as desired.

(e) follows from (d): Due to the existence of polar coordinates for complex numbers, there exists $\zeta \in \mathbb{K}$ with $|\zeta| = 1$ and
\[
\left| \int_A f \, d\mu \right| = \zeta \int_A f \, d\mu = \int_A \zeta f \, d\mu.
\]
In the above equality, all terms are real numbers, implying
\[
\left| \int_A f \, d\mu \right| = \operatorname{Re} \left( \int_A \zeta f \, d\mu \right) = \int_A \operatorname{Re}(\zeta f) \, d\mu \leq \int_A |f| \, d\mu,
\]
proving (e).

(f): Since $m p \leq f p \leq M p$, we compute
\[
m \int_A p \, d\mu \overset{(d)}{\leq} \int_A f p \, d\mu \overset{(d)}{\leq} M \int_A p \, d\mu.
\]
If $\int_A p \, d\mu = 0$, then (2.15e) implies $\int_A f p \, d\mu = 0$ and (2.15f) is immediate. It remains to consider that $\int_A p \, d\mu > 0$. If $f(A) = [m, M]$, then (2.15e) shows there is $\xi \in A$ such that $f(\xi) = \frac{\int_A f p \, d\mu}{\int_A p \, d\mu}$, i.e. (2.15f) holds.

(g) follows from Th. 2.14(v), since $\max(f, g)$ and $\min(f, g)$ are measurable and both $|\max(f, g)|$ and $|\min(f, g)|$ are dominated by $|f| + |g|$.

\section*{2.2 Null Sets, Properties Holding Almost Everywhere}

**Definition 2.17.** Let $(X, \mathcal{A}, \mu)$ be a measure space. Let $P(x)$ be a predicate (informally, a property) of $x \in X$, i.e. a sentence that, for each $x \in X$, is either true or false. Then we
say that \( P \) holds \( \mu \)-almost everywhere (abbreviated \( \mu \)-a.e.) if, and only if, there exists a \( \mu \)-null set \( N \in A \) such that \( P(x) \) is true for each \( x \in N^c \). One writes just a.e. instead of \( \mu \)-a.e. if the measure \( \mu \) is understood.

Here are some examples of potentially useful properties that might hold \( \mu \)-almost everywhere: Functions \( f, g : X \to Y \) are equal \( \mu \)-a.e. if, and only if, there exists a \( \mu \)-null set \( N \in A \) such that \( f|_{N^c} = g|_{N^c} \). A function \( f : X \to \mathbb{R} \) is finite \( \mu \)-a.e. if, and only if, there exists a \( \mu \)-null set \( N \in A \) such that \( f(N^c) \subseteq \mathbb{R} \). The sequence \( (f_n)_{n \in \mathbb{N}} \) of functions \( f_n : X \to \hat{K} \) converges a.e. to \( f : X \to \hat{K} \) if, and only if, there exists a \( \mu \)-null set \( N \) such that \( \lim_{n \to \infty} f_n|_{N^c} = f|_{N^c} \).

Note that “\( P \) holds \( \mu \)-a.e.” does not imply that \( M := \{ x \in X : P(x) \) is false\} \( \in A \). It merely implies that there exists a \( \mu \)-null set \( N \in A \) such that \( M \subseteq N \) (of course, if \( \mu \) is complete, then we do know \( M \in A \)).

**Theorem 2.18.** Let \((X, A, \mu)\) be a measure space.

(a) If \( f, g : X \to \mathbb{R} \) are integrable or nonnegative measurable, then \( f \leq g \) \( \mu \)-a.e. implies

\[
\int_X f \, d\mu \leq \int_X g \, d\mu .
\]

In particular, if \( f = g \) \( \mu \)-a.e., then

\[
\int_X f \, d\mu = \int_X g \, d\mu .
\]

(b) If \( f, g : X \to \hat{K} \) are measurable, \( g \) is integrable, and \( f = g \) \( \mu \)-a.e., then \( f \) is integrable and \((2.17b)\) holds.

(c) If \( f : X \to \hat{K} \) is measurable and there exists an integrable \( g \in \mathcal{M}^+(A) \) such that \( |f| \leq g \) \( \mu \)-a.e., then \( f \) is integrable.

(d) If \( f, g : X \to \mathbb{R} \) are integrable and

\[
\forall_{A \in A} \int_A f \, d\mu \leq \int_A g \, d\mu ,
\]

then \( f \leq g \) \( \mu \)-a.e. In particular, if \( (2.18) \) holds with equality, then \( f = g \) \( \mu \)-a.e.

**Proof.** (a): As \( f, g \) are both measurable, we have \( N := \{ f > g \} \in A \) with \( \mu(N) = 0 \). Thus,

\[
\int_X f^+ \, d\mu = \int_N f^+ \, d\mu + \int_{N^c} f^+ \, d\mu \leq 0 + \int_X g^+ \, d\mu .
\]

Analogously, we obtain \( \int_X f^- \, d\mu \geq \int_X g^- \, d\mu \), proving (a).
(b): One applies (a) to the positive and negative parts of the real and imaginary parts of \( f \) and \( g \). If \( g \) is integrable, then \((\text{Re } g)^+\) is integrable, implying \( \int_X (\text{Re } f)^+ \, d\mu = \int_X (\text{Re } g)^+ \, d\mu < \infty \) by (a), i.e. \((\text{Re } f)^+\) is integrable etc. Now (2.17b) also follows.

(c): As in the proof of (b), one obtains \((\text{Re } f)^+, (\text{Re } f)^-, (\text{Im } f)^+, (\text{Im } f)^-\) to be integrable: If \( g \) is integrable, then \( \int_X (\text{Re } f)^+ \, d\mu \leq \int_X |f| \, d\mu \leq \int_X g \, d\mu < \infty \) by (a), i.e. \((\text{Re } f)^+\) is integrable etc.

(d): The measurability of \( f, g \) implies \( M := \{ f > g \} \in \mathcal{A} \) as well as \( M_n := \{ f > g + \frac{1}{n} \} \in \mathcal{A} \) for each \( n \in \mathbb{N} \). We estimate

\[
\int_{M_n} \frac{f}{n} \, d\mu \geq \int_{M_n} \left( g + \frac{1}{n} \right) \, d\mu = \int_{M_n} g \, d\mu + \frac{\mu(M_n)}{n} \geq \int_{M_n} f \, d\mu + \frac{\mu(M_n)}{n},
\]

implying \( \mu(M_n) = 0 \) for each \( n \in \mathbb{N} \). Since \( M_n \uparrow M \), we obtain \( \mu(M) = 0 \) as claimed. ■

2.3 Convergence Theorems

2.3.1 Fatou’s Lemma and Dominated Convergence

**Proposition 2.19** (Fatou’s Lemma). Let \((X, \mathcal{A}, \mu)\) be a measure space. Then, for each sequence \((f_n)_{n \in \mathbb{N}} \in \mathcal{M}^+(\mathcal{A})\), one has

\[
\int_X \lim_{n \to \infty} f_n \, d\mu \leq \liminf_{n \to \infty} \int_X f_n \, d\mu. \tag{2.19}
\]

**Proof.** Letting \( f := \lim_{n \to \infty} f_n \) and, for each \( n \in \mathbb{N} \), \( g_n := \inf_{k \geq n} f_k \), we obtain \( f, g_n \in \mathcal{M}^+(\mathcal{A}) \) by Th. 1.82(a). Clearly, \( g_n \uparrow f \), such that the monotone convergence Th. 2.7 applies, yielding

\[
\int_X f \, d\mu = \lim_{n \to \infty} \int_X g_n \, d\mu.
\]

Since

\[
\left( \forall k \geq n \quad g_n \leq f_k \right) \Rightarrow \int_X g_n \, d\mu \leq \inf_{k \geq n} \int_X f_k \, d\mu,
\]

we obtain

\[
\int_X f \, d\mu \leq \liminf_{n \to \infty} \int_X f_n \, d\mu = \liminf_{n \to \infty} \int_X f_n \, d\mu,
\]

which establishes the case. ■

**Theorem 2.20** (Dominated Convergence). Let \((X, \mathcal{A}, \mu)\) be a measure space. Let \( f, f_n : X \to \mathbb{R} \) be measurable functions, \( n \in \mathbb{N} \), such that \( f = \lim_{n \to \infty} f_n \) \( \mu \)-a.e. Moreover, assume the \( f_n \) to be dominated by an integrable function, i.e. assume there exists an integrable \( g \in \mathcal{M}^+(\mathcal{A}) \) such that, for each \( n \in \mathbb{N} \), \( |f_n| \leq g \) holds \( \mu \)-a.e. Then all \( f_n \) and \( f \) are integrable with

\[
\lim_{n \to \infty} \int_X f_n \, d\mu = \int_X f \, d\mu \tag{2.20a}
\]

and

\[
\lim_{n \to \infty} \int_X |f_n - f| \, d\mu = 0. \tag{2.20b}
\]
Proof. The $f_n$ are integrable by Th. 2.18(c) and then $|f| = \lim_{n \to \infty} |f_n| \leq g \mu$-a.e. implies $f$ to be integrable as well. Then we know that outside of a $\mu$-null set $N \in \mathcal{A}$, all the functions $g, f, f_n, n \in \mathbb{N}$, must be $\mathbb{K}$-valued and $\lim_{n \to \infty} f_n = f$. Since all the involved integrals over $N$ equal 0, we may assume, without loss of generality, that $N = \emptyset$. Then, for each $n \in \mathbb{N}$, $g_n := |f| + g - |f_n - f| \in \mathcal{M}^+(\mathcal{A})$, and Fatou (Prop. 2.19) yields

$$
\int_X (|f| + g) \, d\mu = \int_X \lim_{n \to \infty} g_n \, d\mu \leq \liminf_{n \to \infty} \int_X g_n \, d\mu
$$

$$
= \int_X (|f| + g) \, d\mu - \limsup_{n \to \infty} \int_X |f_n - f| \, d\mu.
$$

Since $\int_X (|f| + g) \, d\mu \in \mathbb{R}_0^+$, we obtain (2.20b). As

$$
\forall n \in \mathbb{N} \left| \int_X f_n \, d\mu - \int_X f \, d\mu \right| \leq \int_X |f_n - f| \, d\mu,
$$

(2.20a) is also proved. ■

Note that we have already seen in Ex. 2.9(a), in the context of the monotone convergence theorem, that one can not expect (2.20a) to hold without additional hypotheses (such as monotonicity or domination).

**Corollary 2.21.** Let $(X, \mathcal{A}, \mu)$ be a measure space.

(a) Let $f, f_n : X \to \mathbb{K}$ be measurable functions, $n \in \mathbb{N}$, such that $f = \sum_{n=1}^{\infty} f_n$ $\mu$-a.e. Moreover, assume there exists an integrable $g \in \mathcal{M}^+(\mathcal{A})$ such that, for each $n \in \mathbb{N}$, $|\sum_{k=1}^{n} f_k| \leq g$ holds $\mu$-a.e. Then all $f_n$ and $f$ are integrable with

$$
\int_X f \, d\mu = \sum_{n=1}^{\infty} \int_X f_n \, d\mu.
$$

(b) Let $f : X \to \mathbb{K}$ be integrable and $A, A_n \in \mathcal{A}$ such that $A = \bigcup_{n=1}^{\infty} A_n$ and, for $k \neq l$, $\mu(A_k \cap A_l) = 0$. Then

$$
\int_A f \, d\mu = \sum_{n=1}^{\infty} \int_{A_n} f \, d\mu.
$$

Proof. (a): The integrability of the $f_n$ follows since $f_n = \sum_{k=1}^{n} f_k - \sum_{k=1}^{n-1} f_k$, where both sums on the right-hand side are integrable by hypothesis. Now the claimed formula for $\int_X f \, d\mu$ is obtained from applying Th. 2.20 with $f_n$ replaced by $\sum_{k=1}^{n} f_k$.

(b): Since $\{|f| = \infty\}$ and $\bigcup_{(k,l) \in \mathbb{N}^2, k \neq l}(A_k \cap A_l)$ are both $\mu$-null sets, we may assume, without loss of generality, that $A$ is the disjoint union of the $A_n$ and that $f \cdot \chi_A$ is $\mathbb{K}$-valued. Then $f \cdot \chi_A = \sum_{n=1}^{\infty} f \cdot \chi_{A_n}$ with $|\sum_{k=1}^{n} f \cdot \chi_{A_k}| \leq \sum_{k=1}^{n} |f| \cdot \chi_{A_k} \leq |f| \cdot \chi_A$, i.e. (b) now follows from (a). ■
2.3.2 Parameter-Dependent Integrals: Continuity, Differentiation

**Theorem 2.22.** Let \((Z, d)\) be a metric space and let \((X, \mathcal{A}, \mu)\) be a measure space. Let \(f : Z \times X \rightarrow \hat{K}\). Fix \(z_0 \in Z\) and assume \(f\) has the following properties:

(a) For each \(z \in Z\), the function \(x \mapsto f(z, x)\) is integrable.

(b) For \(\mu\)-a.e. \(x \in X\), the function \(z \mapsto f(z, x)\) is continuous in \(z_0\).

(c) There exists \(\epsilon > 0\) and an integrable \(h : X \rightarrow [0, \infty]\) such that

\[
\forall z \in B_{\epsilon}(z_0) \quad |f(z, x)| \leq h(x) \quad \text{for } \mu\text{-a.e. } x \in X
\]  

(i.e. \(f\) is uniformly dominated by an integrable \(h\) in some neighborhood of \(z_0\)).

Then the function

\[
\phi : Z \rightarrow \hat{K}, \quad \phi(z) := \int_X f(z, x) \, d\mu(x),
\]  

is continuous in \(z_0\).

**Proof.** Let \((z_m)_{m \in \mathbb{N}}\) be a sequence in \(B_{\epsilon}(z_0)\) such that \(\lim_{m \to \infty} z_m = z_0\). According to (a), this gives rise to a sequence \((g_m)_{m \in \mathbb{N}}\) of integrable functions

\[
g_m : X \rightarrow \hat{K}, \quad g_m(x) := f(z_m, x),
\]

such that, using (b), \(\lim_{m \to \infty} g_m = g\) \(\mu\)-a.e., where

\[
g : X \rightarrow \hat{K}, \quad g(x) := f(z_0, x).
\]

Since, by (c), all \(g_m\) are dominated by the integrable function \(h\), we can apply the dominated convergence Th. 2.20 to obtain

\[
\lim_{m \to \infty} \phi(z_m) = \lim_{m \to \infty} \int_X f(z_m, x) \, d\mu(x) = \lim_{m \to \infty} \int_X g_m(x) \, d\mu(x) \stackrel{\text{Th. 2.20}}{=} \int_X g(x) \, d\mu(x)
\]

\[
= \int_X f(z_0, x) \, d\mu(x) = \phi(z_0),
\]

proving \(\phi\) is continuous at \(z_0\).

**Theorem 2.23.** Let \(I \subseteq \mathbb{R}\) be a nontrivial interval and let \((X, \mathcal{A}, \mu)\) be a measure space. Let \(f : I \times X \rightarrow \mathbb{K}\). Fix \(t_0 \in I\) and assume \(f\) has the following properties:

(a) For each \(t \in I\), the function \(x \mapsto f(t, x)\) is integrable.

(b) For \(\mu\)-a.e. \(x \in X\), the function \(f_x : I \rightarrow \mathbb{K}, f_x(t) := f(t, x)\) is differentiable at \(t_0\).

(c) There exists an integrable \(h : X \rightarrow [0, \infty]\) such that

\[
\forall t \in I \setminus \{t_0\} \quad \left| \frac{f(t, x) - f(t_0, x)}{t - t_0} \right| \leq h(x) \quad \text{for } \mu\text{-a.e. } x \in X.
\]  

\[ (2.23) \]
Then the function
\[ \phi : I \to K, \quad \phi(t) := \int_X f(t, x) \, d\mu(x), \]
(2.24)
is differentiable at \( t_0 \) (this means the existence of the one-sided derivative if \( t_0 \in \partial I \)), 
\( x \mapsto f'(t_0, x) \) (with value 0 on the \( \mu \)-null set \( N \subseteq X \), where \( f'(t_0, x) \) does not exist), is integrable, and
\[ \phi'(t_0) = \int_X f'(t_0, x) \, d\mu(x). \]
(2.25)

Proof. Let \( (t_m)_{m \in \mathbb{N}} \) be a sequence in \( I \setminus \{t_0\} \) such that \( \lim_{m \to \infty} t_m = t_0 \). According to (a), this gives rise to a sequence \( (g_m)_{m \in \mathbb{N}} \) of integrable functions
\[ g_m : X \to K, \quad g_m(x) := \frac{f(t_m, x) - f(t_0, x)}{t_m - t_0}, \]
such that, using (b), \( \lim_{m \to \infty} g_m = g \) \( \mu \)-a.e., where
\[ g : X \to K, \quad g(x) := \begin{cases} f'(t_0, x) & \text{if } x \notin N, \\ 0 & \text{if } x \in N. \end{cases} \]
Since, by (c), all \( g_m \) are dominated by the integrable function \( h \), we can apply the dominated convergence Th. 2.20 to obtain the integrability of the \( g_m \) and \( g \), as well as
\[ \lim_{m \to \infty} \frac{\phi(t_m, x) - \phi(t_0, x)}{t_m - t_0} = \lim_{m \to \infty} \int_X g_m(x) \, d\mu(x) \stackrel{\text{Th. 2.20}}{=} \int_X g(x) \, d\mu(x) = \int_X f'(t_0, x) \, d\mu(x), \]
proving \( \phi \) is differentiable at \( t_0 \) with derivative according to (2.25).

\[ \square \]

Corollary 2.24. The statement of Th. 2.23 remains true if (b) and (c) are replaced by
(b\textsuperscript{'} ) For \( \mu \)-a.e. \( x \in X \), the function \( f_x : I \to K, f_x(t) := f(t, x) \), is differentiable on \( I \).
(c\textsuperscript{'} ) There exists an integrable \( h : X \to [0, \infty] \) such that, for \( \mu \)-a.e. \( x \in X \),
\[ \forall t \in I \quad |f'_x(t)| \leq h(x) \]
(note that, this time, the null set, where the condition does not hold, must not depend on \( t \)).

Proof. Since (b\textsuperscript{'}) immediately implies (b), it remains to verify that (c\textsuperscript{'}) implies (c). To this end, let \( N \) be a \( \mu \)-null set such that, for each \( x \in X \setminus N \), \( |f'_x(t)| \leq h(x) \) holds for each \( t \in I \). Then, for each \( x \in X \setminus N \), by the mean value theorem of differential calculus [Phi16a, Th. 9.18], for each \( t \in I \setminus \{t_0\} \), there exists \( \tau(t, x) \in I \) such that
\[ \left| \frac{f(t, x) - f(t_0, x)}{t - t_0} \right| = \left| f'_x(\tau(t, x)) \right| \leq h(x), \]
implying the validity of (2.23).

\[ \square \]
2.4 Riemann versus Lebesgue

In the present section, we investigate the obvious question regarding the relationship between the Riemann integral of [Phi16a, Sec. 10] and the Lebesgue integral for $L^1$-measurable functions $f : [a, b] \rightarrow \mathbb{K}$.

**Definition 2.25.** Let $n \in \mathbb{N}$, $A \in \mathcal{L}^n$. We call $f : A \rightarrow \mathbb{K}$ Lebesgue integrable if, and only if, $f$ is $(\lambda^n |_A)$-integrable.

We will now see in Th. 2.26(a) that every Riemann integrable function is also Lebesgue integrable with identical values for the integral. Thus, the Lebesgue integral can be seen as a generalization of the Riemann integral. In [Phi16a, Sec. 10], we only studied the Riemann integral for functions on one-dimensional intervals. However, an analogous construction can be performed for functions on intervals in $\mathbb{R}^n$, $n \in \mathbb{N}$. It is provided in Sec. D of the Appendix. While the following Th. 2.26 is provided for arbitrary $n \in \mathbb{N}$, one can understand its contents and structure already from the case $n = 1$ (in particular, without having studied Sec. D). The main difference between $n = 1$ and $n > 1$ is that the notation gets somewhat more complicated.

As a consequence of Th. 2.26(a), for functions on $\mathbb{R}$, one will, in most cases, still compute Lebesgue integrals as Riemann integrals, using the techniques introduced in [Phi16a, Sec. 10]. The most important technique for computing integrals on (subsets of) $\mathbb{R}^n$ is the Fubini Th. 2.36(b) below, which, if it applies, allows to compute an integral of a function on $\mathbb{R}^n$ by evaluating $n$ one-dimensional integrals.

**Theorem 2.26.** Let $n \in \mathbb{N}$ and $I := [a, b] \subseteq \mathbb{R}^n$, where $a, b \in \mathbb{R}^n$ with $a < b$. Moreover, let $f : I \rightarrow \mathbb{C}$ be bounded.

(a) If $f$ is Riemann integrable, then $f$ is Lebesgue integrable and

$$\int_I f \, d\lambda^n = \text{R-} \int_I f,$$

(2.26)

where R- $f$ denotes the Riemann integral.

(b) $f$ is Riemann integrable if, and only if, the set of points where $f$ is discontinuous is a $\lambda^n$-null set (i.e. if, and only if, $f$ is continuous $\lambda^n$-a.e.).

**Proof.** We begin with some general preparations: As we may consider $\text{Re } f$ and $\text{Im } f$, we consider $f$ to be real-valued without loss of generality. For each $N \in \mathbb{N}$ and each $k \in \{1, \ldots, n\}$, consider the partition $\Delta_k = (x_{k,0}^N, \ldots, x_{k,2^N})$ of $[a_k, b_k]$, $x_{k,j}^N := a_k + j \cdot (b_k - a_k) 2^{-N}$ for each $j \in \{0, \ldots, 2^N\}$, into the disjoint intervals $I_{k,1}^N := [x_{k,0}^N, x_{k,1}^N]$, $I_{k,2}^N := [x_{k,1}^N, x_{k,2}^N]$, $\ldots$, $I_{k,2^N}^N := [x_{k,2^N-1}^N, x_{k,2^N}^N]$, and the resulting partition $\Delta^N$ of $I$ (cf. Def. D.2). Analogous to Def. D.2, for each $(k_1, \ldots, k_n) \in P(\Delta^N) = \{1, \ldots, 2^N\}^n$, define

$$I_{(j_1, \ldots, j_n)}^N := \prod_{k=1}^n I_{k,j_k}^N,$$
which is as in (D.4), except that we are using the (disjoint) halfopen intervals here. Then, analogous to Rem. D.3, we obtain the (now disjoint) unions

\[ I = \bigcup_{p \in P(\Delta^N)} I_p^N. \]

According to Def. D.4, if we let, for each \( p \in P(\Delta^N) \),

\[ m_p^N := \inf \{ f(x) : x \in \text{cl}(I_p^N) \}, \quad M_p^N := \sup \{ f(x) : x \in \text{cl}(I_p^N) \}, \]

then we obtain the lower and upper Riemann sums

\[ r(\Delta^N, f) = \sum_{p \in P(\Delta^N)} m_p^N \cdot \lambda^n(I_p^N), \quad R(\Delta^N, f) = \sum_{p \in P(\Delta^N)} M_p^N \cdot \lambda^n(I_p^N). \]

Moreover, defining the simple functions

\[ g_N : I \to \mathbb{R}, \quad g_N := \sum_{p \in P(\Delta^N)} m_p^N \chi_{I_p^N}; \]
\[ h_N : I \to \mathbb{R}, \quad h_N := \sum_{p \in P(\Delta^N)} M_p^N \chi_{I_p^N}, \]

we obtain

\[ r(\Delta^N, f) = \int_I g_N \, d\lambda^n, \quad R(\Delta^N, f) = \int_I h_N \, d\lambda^n. \]

Moreover, clearly, \( g_N \leq f \leq h_N \), \( (g_N)_{N \in \mathbb{N}} \) is increasing, and \( (h_N)_{N \in \mathbb{N}} \) is decreasing. Thus, the pointwise limits \( g := \lim_{N \to \infty} g_N \) and \( h := \lim_{N \to \infty} h_N \) exist, are \( \beta^n \)-measurable, bounded, and, hence, Lebesgue integrable. They also satisfy \( g \leq f \leq h \).

(a): Suppose \( f \) is Riemann integrable. Since

\[ \forall_{N \in \mathbb{N}} \quad |g_N|, |h_N| \leq |g_1| + |h_1|, \]

we can apply the dominated convergence Th. 2.20 to compute

\[ \int_I g \, d\lambda^n = \lim_{N \to \infty} \int_I g_N \, d\lambda^n = \lim_{N \to \infty} R(\Delta^N, f) = R^- \int_I f \]

\[ = \lim_{N \to \infty} R(\Delta^N, f) = \lim_{N \to \infty} \int_I h_N \, d\lambda^n = \int_I h \, d\lambda^n. \]

Thus, \( \int_I (h - g) \, d\lambda^n = 0 \), such that \( h = g \lambda^n \)-a.e. on \( I \) by Lem. 2.6(a). Since \( g \leq f \leq h \), this also yields \( f = g \lambda^n \)-a.e. on \( I \). In consequence, there exists a \( \lambda^n \)-null set \( M \in \mathcal{L}^n | I \) such that \( f = g \) on \( M^c \). If \( B \in \mathcal{B}^1 \), then

\[ f^{-1}(B) = (f^{-1}(B) \cap M^c) \cup (f^{-1}(B) \cap M) = (g^{-1}(B) \cap M^c) \cup (f^{-1}(B) \cap M) \in \mathcal{L}^n | I, \]

since \( g \) is measurable and \( \lambda^n \) is complete. Thus, \( f \) is measurable as well. Moreover,

\[ \int_I f \, d\lambda^n = \int_{M^c} f \, d\lambda^n + \int_{M} f \, d\lambda^n = \int_{M^c} g \, d\lambda^n + 0 = \int_I g \, d\lambda^n = R^- \int_I f, \]
proving (a).

(b): Let \( D := \{ x \in I : f \text{ not continuous at } x \} \) and let

\[
R := \bigcup_{N \in \mathbb{N}} \bigcup_{p \in P(\Delta^N)} \partial I^N_p
\]

be the set of boundary points occurring in any of the considered partitions of \( I \), which, as a countable union of \( \lambda^n \)-null sets is itself a \( \lambda^n \)-null set. We show that \( D \subseteq R \cup \{ g < h \} \) by proving that \((R \cup \{ g < h \})^c = R^c \cap \{ g = h \} \subseteq D^c\): Let \( x \in R^c \cap \{ g = h \} \) and let \( \epsilon > 0 \). Note \( f(x) = g(x) = h(x) \). Choose \( N \in \mathbb{N} \) such that \( f(x) - \epsilon < g_N(x) < h_N(x) < f(x) + \epsilon \).

If \( x \notin R \), there exists \( p_N \in P(\Delta^N) \) such that \( x \in \text{int}(I^N_{p_N}) \). Set \( \delta := \text{dist}(x, \partial I_{p_N}) \). Then \( \delta > 0 \) and, for each \( y \in \mathbb{R}^n \) with \( \|y - x\| < \delta \), one has \( y \in I^N_{p_N} \), implying

\[
f(x) - \epsilon < g_N(x) = m^N_{p_N} \leq f(y) \leq M^N_{p_N} = h_N(x) < f(x) + \epsilon,
\]

showing \( |f(y) - f(x)| < \epsilon \) and the continuity of \( f \) at \( x \), i.e. \( x \in D^c \) as claimed. Now, if \( f \) is Riemann integrable, then, as shown in the proof of (a), \( g = h \ \lambda^n\text{-a.e.}, \) i.e. \( \lambda^n(\{ g < h \}) = 0 \). Since \( D \subseteq R \cup \{ g < h \} \), this implies \( \lambda^n(D) = 0 \). For the converse, we show \( \{ g < h \} \subseteq D \): Indeed, for \( x \in I \) with \( g(x) < h(x) \) let \( \delta := (h(x) - g(x))/3 \). Then, for each \( \epsilon > 0 \), there exist \( x_1, x_2 \in B_\epsilon(x) \) such that \( f(x_1) \leq g(x) + \delta \) and \( f(x_2) \geq h(x) - \delta \), showing \( x \in D \) and \( \{ g < h \} \subseteq D \). Thus, if \( D \in L^n \) with \( \lambda^n(D) = 0 \), then \( g = h \ \lambda^n\text{-a.e.}, \) yielding

\[
\lim_{N \to \infty} r(\Delta^N, f) = \lim_{N \to \infty} \int_I g_N \, d\lambda^n = \int_I g \, d\lambda^n
\]

\[
= \int_I h \, d\lambda^n = \lim_{N \to \infty} \int_I h_N \, d\lambda^n = \lim_{N \to \infty} R(\Delta^N, f),
\]

thereby proving \( f \) to be Riemann integrable by Th. D.11.

**Example 2.27. (a)** Let \( I = [a, b] \subseteq \mathbb{R}^n \) be an interval, \( a, b \in \mathbb{R}^n \), \( n \in \mathbb{N} \), \( a < b \). The **Dirichlet function**

\[
f : I \to \mathbb{R}, \quad f(x) := \begin{cases} 
0 & \text{for } x \in I \setminus \mathbb{Q}^n, \\
1 & \text{for } x \in I \cap \mathbb{Q}^n,
\end{cases}
\]

is everywhere discontinuous and, thus, not Riemann integrable (cf. Ex. D.7(b) and [Phi16a, Ex. 10.7(b)]). However, \( f \) is the characteristic function of the \( \beta^n \)-null set \( I \cap \mathbb{Q}^n \), i.e. \( f \) is Lebesgue integrable with \( \int_I f \, d\lambda^n = 0 \). Now let \( (r_1, r_2, \ldots) \) be an enumeration of \( I \cap \mathbb{Q}^n \) and define

\[
\forall_{k \in \mathbb{N}} \phi_k : I \to \mathbb{R}, \quad \phi_k(x) := \begin{cases} 
1 & \text{for } x \in \{r_1, \ldots, r_k\}, \\
0 & \text{otherwise}.
\end{cases}
\]

Then each \( \phi_k \) is Riemann integrable, as it has only finitely many discontinuities. Since, clearly, \( \phi_k \uparrow f \), this shows that there does not exist a “Riemann measure”, i.e. no measure space \((I, A, \rho)\) such that a function \( g : I \to \mathbb{R} \) is Riemann integrable if, and only if, \( g \) is \( \rho \)-integrable (otherwise, we had \( \phi_k \in S^+(\mathcal{A}) \) for each \( k \in \mathbb{N} \), and \( f \) had to be \( \rho \)-integrable as well).
(b) The following example shows that there exist Riemann integrable functions that are not Borel measurable: Let $C \subset I := [0,1]$ be the Cantor set of Def. and Rem. 1.61 and let $A \subset C$ be such that $A \notin B^1$ (such an $A$ exists since $\#\mathcal{P}(C) = \#\mathcal{P}(\mathbb{R})$, but $\#B^1 = \#\mathbb{R}$). Then $f : I \rightarrow \mathbb{R}$, $f := \chi_A$ is continuous on the open set $I \setminus C$. Since $\lambda^1(C) = 0$, $f$ is Riemann integrable with $\int_I f = 0$.

While we saw in Th. 2.26(a) that every Riemann integrable function is Lebesgue integrable, for improperly Riemann integrable functions this only remains true if the improper Riemann integral converges absolutely:

**Theorem 2.28.** Let $a, b, c \in \mathbb{R}$, $a < c < b$. Let $I \subseteq [a, b]$ be one of the following three kinds of intervals: $I = [c, b]$, $I = [a, c]$, or $I = [a, b]$. Moreover, assume $f : I \rightarrow \mathbb{R}$ to be locally Riemann integrable over each compact subinterval of $I$. Then $f$ is Lebesgue integrable if, and only if, $|f|$ is improperly Riemann integrable. In that case, both integrals agree, i.e. (2.26) holds, where $R-\int$ now denotes the improper Riemann integral.

**Proof.** Let $m = \inf I$, $M = \sup I$. Moreover, let $(a_n)_{n \in \mathbb{N}}$ and $(b_n)_{n \in \mathbb{N}}$ be sequences in $I$ such that $m < a_n < b_n < M$ and $a_n \downarrow m$ as well as $b_n \uparrow M$. If, for each $n \in \mathbb{N}$, $f_n := f \cdot \chi_{[a_n, b_n]}$, then $f_n \uparrow f$ on $]m, M[$. Since $f$ is locally Riemann integrable, each $f_n$ is Lebesgue integrable by Th. 2.26(a), implying $f$ to be Lebesgue measurable. Thus, the monotone convergence Th. 2.7 yields

$$\lim_{n \rightarrow \infty} R-\int_{a_n}^{b_n} |f| = \lim_{n \rightarrow \infty} \int_{[a_n, b_n]} |f| \, d\lambda^1 = \int_I |f| \, d\lambda^1.$$ 

If $|f|$ is improperly Riemann integrable, then the left-hand side is finite, i.e. $|f|$ (and, thus, $f$) is Lebesgue integrable. Conversely, if $f$ is Lebesgue integrable, then so is $|f|$, i.e. the right-hand side is finite, implying $|f|$ to be improperly Riemann integrable. Finally, if $|f|$ is improperly Riemann integrable, then so is $f$ by [Phi16a, Prop. 10.39(b)], implying, together with Th. 2.26(a) and the dominated convergence Th. 2.20,

$$R-\int_{m}^{M} f = \lim_{n \rightarrow \infty} R-\int_{a_n}^{b_n} f = \lim_{n \rightarrow \infty} \int_{[a_n, b_n]} f \, d\lambda^1 = \int_I f \, d\lambda^1,$$

proving the validity of (2.26).

In [Phi16a, Ex. 10.40(c)], we saw that the function

$$f : [0, \infty[ \rightarrow \mathbb{R}, \quad f(t) := \begin{cases} (-1)^{n+1} & \text{for } n \leq t \leq n + \frac{1}{n}, \; n \in \mathbb{N}, \\ 0 & \text{otherwise,} \end{cases} \quad (2.27)$$

is improperly Riemann integrable, but that the improper Riemann integral does not converge absolutely. Therefore, by Th. 2.28, this function is not Lebesgue integrable, even though it is improperly Riemann integrable.

In the next example, we introduce the important **gamma function**, which can be seen as a continuous extension of the factorial function (cf. (2.29b) below).
Example 2.29. The gamma function is defined by

$$\Gamma : \mathbb{R}^+ \to \mathbb{R}^+, \quad \Gamma(x) := \int_0^\infty e^{-t}t^{x-1} \, dt. \quad (2.28)$$

(a) The gamma function is well-defined by (2.28), i.e. the integral exists as an improper Riemann integral and as a Lebesgue integral with values in $\mathbb{R}^+$: Since, for each $t, x > 0$, one has $e^{-t}t^{x-1} > 0$, and $f_x : \mathbb{R}^+ \to \mathbb{R}^+$, $f_x(t) := e^{-t}t^{x-1}$, is continuous, it suffices to show $f_x$ is dominated by some improperly Riemann integrable function $g_x$. In anticipation of (c) below, it will actually be useful to estimate $f_x$ uniformly for each $x$ in some bounded interval. Thus, let $\alpha, \beta \in \mathbb{R}^+$ such that $0 < \alpha < \beta$. Then, for each $x \in [\alpha, \beta]$,

$$\forall 0 < t \leq 1 \quad 0 < f_x(t) \leq t^{x-1} \leq t^{\alpha-1},$$

$$\forall t \geq 1 \quad 0 < f_x(t) = t^{x-1}e^{-t/2} \leq M_{\beta}e^{-t/2},$$

where $M_{\beta} \in \mathbb{R}^+$ is an upper bound for the bounded function $\phi : [1, \infty[ \to \mathbb{R}^+$, $\phi(t) := t^{\beta-1}e^{-t/2}$ (note $\lim_{t \to \infty} \phi(t) = 0$). Thus, the function

$$g_{\alpha, \beta} : \mathbb{R}^+ \to \mathbb{R}^+, \quad g_{\alpha, \beta}(t) := \begin{cases} t^{\alpha-1} & \text{for } t \leq 1, \\ M_{\beta}e^{-t/2} & \text{for } t > 1, \end{cases}$$

donates $f_x$ for each $x \in [\alpha, \beta]$. It is also improperly Riemann integrable, since

$$\int_1^\infty t^{\alpha-1} \, dt = \frac{1}{\alpha} \quad \text{by [Phi16a, Ex. 10.35(a)]},$$

and, if $(t_k)_{k \in \mathbb{N}}$ is a sequence in $]1, \infty[\text{ such that } \lim_{k \to \infty} t_k = \infty$, then

$$\int_1^\infty e^{-t/2} \, dt = \lim_{k \to \infty} \int_1^{t_k} e^{-t/2} \, dt = \lim_{k \to \infty} \left[ -2e^{-t/2} \right]_1^{t_k} = \lim_{k \to \infty} (2e^{-1/2} - 2e^{-t_k/2}) = 2e^{-1/2}.$$

(b) The gamma function satisfies

$$\forall x \in \mathbb{R}^+ \quad \Gamma(x+1) = x \Gamma(x), \quad (2.29a)$$

$$\forall n \in \mathbb{N}_0 \quad \Gamma(n+1) = n!: \quad (2.29b)$$

To prove (2.29a), we fix $x \in \mathbb{R}^+$ and integrate by parts to obtain

$$\Gamma(x+1) = \int_0^\infty e^{-t}t^x \, dt = \left[ -e^{-t}t^x \right]_0^\infty + x \int_0^\infty e^{-t}t^{x-1} \, dt = -0 + x \Gamma(x),$$

establishing (2.29a). Since, using [Phi16a, Ex. 10.35(c)],

$$\Gamma(1) = \int_0^\infty e^{-t} \, dt = 1, \quad (2.30)$$

(2.29a) implies (2.29b) via induction on $n \in \mathbb{N}_0$. 


(c) The gamma function is $C^\infty$ and
\[
\forall \ k \in \mathbb{N}_0 \quad \Gamma^{(k)} : \mathbb{R}^+ \rightarrow \mathbb{R}, \quad \Gamma^{(k)}(x) = \int_0^\infty (\ln t)^k e^{-tx} \, dt : \quad (2.31)
\]
Using [Phi16a, Ex. 9.5(b)], we differentiate the integrand of (2.28) $k$ times with respect to $x$ to obtain, for each $x \in \mathbb{R}^+$,
\[
\forall \ k \in \mathbb{N} \quad \partial_x^k(e^{-tx^{-1}}) = (\ln t)^k e^{-tx}.
\]
Thus, according to Cor. 2.24, (2.31) is proved if we can estimate each integrand of (2.31) by an integrable function, uniformly for $x \in [\alpha, \beta]$, where $0 < \alpha < \beta$ as in (a). For each $k \in \mathbb{N}$ and $x \in [\alpha, \beta]$, we estimate
\[
\forall \ x \geq 0 \quad 0 < |\ln t|^k e^{-tx} \leq |\ln t|^k t^{\alpha/2} e^{-tx} \leq C_{\alpha,k} t^{\alpha/2},
\]
\[
\forall \ t \geq 0 \quad 0 < |\ln t|^k e^{-tx} \leq |\ln t|^k t^{-2} e^{-tx} \leq M_{\beta,k} e^{-tx/2},
\]
where $C_{\alpha,k} \in \mathbb{R}^+$ is an upper bound for the bounded function $\psi_k : [0, 1] \rightarrow \mathbb{R}^+$, $\psi(t) = |\ln t|^k t^{\alpha/2}$ (note $\lim_{t \rightarrow 0} \psi_k(t) = 0$), and $M_{\beta,k} \in \mathbb{R}^+$ is an upper bound for the bounded function $\phi_k : [1, \infty[ \rightarrow \mathbb{R}^+$, $\phi(t) = |\ln t|^k t^{\beta-1} e^{-t/2}$ (note $\lim_{t \rightarrow \infty} \phi_k(t) = 0$). Thus, the function
\[
h_{\alpha,\beta,k} : \mathbb{R}^+ \rightarrow \mathbb{R}^+, \quad h_{\alpha,\beta,k}(t) := \begin{cases} C_{\alpha,k} t^{\alpha/2} & \text{for } t \leq 1, \\ M_{\beta,k} e^{-t/2} & \text{for } t > 1, \end{cases}
\]
dominates $|\ln t|^k e^{-tx}$ for each $x \in [\alpha, \beta]$. That it is improperly Riemann integrable follows analogously to the integrability of $g_{\alpha,\beta}$ in (a).

(d) One has the identities
\[
\forall \ x \in \mathbb{R}^+ \quad \Gamma(x) = \lim_{n \rightarrow \infty} \int_0^n \frac{(1 - t/n)^n}{x(x+1)\cdots(x+n)} t^{x-1} \, dt = \lim_{n \rightarrow \infty} \frac{n^x n!}{x(x+1)\cdots(x+n)} \quad (2.32)
\]
and
\[
\int_{-\infty}^\infty e^{-x^2} \, dx = \Gamma \left( \frac{1}{2} \right) = \sqrt{\pi} : \quad (2.33)
\]
For each $x \in \mathbb{R}$, one has $1 + x \leq \sum_{t=0}^{\infty} x^t = e^x$. Thus, for each $x \in [1, \infty[$, applying $\ln$ yields $\ln(1 + x) \leq x$. If $0 \leq t < n \in \mathbb{N}$, then $-1 < -\frac{t}{n}$ and we can apply $\ln(1 + x) \leq x$ with $x := -\frac{t}{n}$ to obtain $(1 - t/n)^n \leq e^{-t}$. In consequence, we can apply the dominated convergence Th. 2.20 to obtain, for each $x \in \mathbb{R}^+$,
\[
\Gamma(x) = \int_0^\infty e^{-tx} \, dt = \lim_{n \rightarrow \infty} \int_0^n \frac{(1 - t/n)^n}{x(x+1)\cdots(x+n)} t^{x-1} \, dt = \lim_{n \rightarrow \infty} \int_0^n \frac{(1 - t/n)^n}{x(x+1)\cdots(x+n)} t^{x-1} \, dt,
\]
proving the first identity of (2.32). To obtain the second identity of (2.32), for each
\(n \in \mathbb{N}\), we now use integration by parts \(n\) times:

\[
\int_0^n \left(1 - \frac{t}{n}\right)^n t^{x-1} dt = \left[\left(1 - \frac{t}{n}\right)^n \frac{t^x}{x}\right]_0^n + \frac{n}{nx} \int_0^n \left(1 - \frac{t}{n}\right)^{n-1} t^x dt
\]

\[= \ldots \]

\[= 0 + \frac{n!}{n^n x(x+1) \cdots (x+n-1)} \int_0^n \left(1 - \frac{t}{n}\right)^{n-1} t^x dt
\]

\[= \frac{n!}{n^n x(x+1) \cdots (x+n)} \left[t^{x+n}\right]_0^n = \frac{n^n n!}{x(x+1) \cdots (x+n)}.\]

The first identity in (2.33) can be proved using the change of variables \(t := x^2\):

\[
\int_0^\infty e^{-x^2} dx = \int_0^\infty \frac{e^{-t}}{2} t^{-1/2} dt = \frac{1}{2} \Gamma \left(\frac{1}{2}\right).
\]

The second identity in (2.33) is a consequence of (2.32) and the so-called Wallis product (see Prop. E.1(c) of the Appendix for an elementary proof of the Wallis product)

\[
\frac{\pi}{2} = \prod_{k=1}^{\infty} \left(\frac{2k}{2k-1} \cdot \frac{2k}{2k+1}\right) = \lim_{n \to \infty} \prod_{k=1}^{n} \left(\frac{2k}{2k-1} \cdot \frac{2k}{2k+1}\right) = \frac{2}{1} \cdot \frac{2}{3} \cdot \frac{4}{3} \cdot \frac{4}{5} \cdot \frac{6}{5} \cdot \frac{6}{7} \cdots
\]

(2.34)

\[
\sqrt{\pi} = \lim_{n \to \infty} \sqrt{n!} \frac{2^{2n+1}}{1 \cdot 3 \cdots (2n+1)} = \Gamma \left(\frac{1}{2}\right),
\]

where we also used the continuity of the square root function.

2.5 Products

2.5.1 Product Measure

**Definition 2.30.** Let \(n \in \mathbb{N}\) and let \((X_i, \mathcal{A}_i, \mu_i)_{i=1}^n\) be a finite family of measure spaces, \(X := \prod_{i=1}^n X_i, \mathcal{A} := \bigotimes_{i=1}^n \mathcal{A}_i\). Then a measure \(\mu\) on \((X, \mathcal{A})\) is called a product measure if, and only if,

\[
\forall (A_1, \ldots, A_n) \in \prod_{i=1}^n \mathcal{A}_i, \quad \mu \left(\bigotimes_{i=1}^n A_i\right) = \prod_{i=1}^n \mu_i(A_i).
\]

(2.35)
Lemma 2.31. Let $n \in \mathbb{N}$, $n \geq 2$, and let $(X_i, \mathcal{A}_i)_{i=1}^n$ be a finite family of measurable spaces, $X := \prod_{i=1}^n X_i$, $\mathcal{A} := \otimes_{i=1}^n \mathcal{A}_i$. Define the sections

$$\forall M \subseteq X \quad \forall x_n \in X_n \quad M_{x_n} := \left\{ (x_1, \ldots, x_{n-1}) \in \prod_{i=1}^{n-1} X_i : (x_1, \ldots, x_n) \in M \right\}. \quad (2.36)$$

(a) Forming sections commutes with set-theoretic rules: Let $M \subseteq X$ and let $(M_j)_{j \in J}$ be a family of subsets of $X$, $J \neq \emptyset$. Also let $E_i \subseteq X_i$ for $i \in \{1, \ldots, n\}$. Then, for each $x_n \in X_n$,

$$\left( \bigcap_{j \in J} M_j \right)_{x_n} = \bigcap_{j \in J} (M_j)_{x_n}, \quad (2.37a)$$

$$\left( \bigcup_{j \in J} M_j \right)_{x_n} = \bigcup_{j \in J} (M_j)_{x_n}, \quad (2.37b)$$

$$(M^c)_{x_n} = (M_{x_n})^c, \quad (2.37c)$$

$$\left( \prod_{i=1}^n E_i \right)_{x_n} = \begin{cases} \emptyset & \text{if } x_n \notin E_n, \\ \prod_{i=1}^{n-1} E_i & \text{if } x_n \in E_n. \end{cases} \quad (2.37d)$$

Moreover, if $f : X \rightarrow Y$, $f(\cdot, x_n) : \prod_{i=1}^{n-1} X_i \rightarrow Y$, and $B \subseteq Y$, then

$$f(\cdot, x_n)^{-1}(B) = (f^{-1}(B))_{x_n}. \quad (2.37e)$$

(b) If $M$ is measurable, then so is each section:

$$\forall M \subseteq X \quad \forall x_n \in X_n \quad \left( M \in \mathcal{A} \implies M_{x_n} \in \otimes_{i=1}^{n-1} \mathcal{A}_i \right).$$

(c) If $(Y, \mathcal{B})$ is another measurable space and $f : X \rightarrow Y$ is $\mathcal{A} \cdot \mathcal{B}$-measurable, then, for each $x_n \in X_n$, the map $f(\cdot, x_n) : \prod_{i=1}^{n-1} X_i \rightarrow Y$ is $\otimes_{i=1}^{n-1} \mathcal{A}_i \cdot \mathcal{B}$-measurable.

Proof. Let $Q := \prod_{i=1}^{n-1} X_i$, $Q := \otimes_{i=1}^n \mathcal{A}_i$.

(a): One has

$$q \in \left( \bigcap_{j \in J} M_j \right)_{x_n} \iff (q, x_n) \in \bigcap_{j \in J} M_j \iff \forall j \in J \quad (q, x_n) \in M_j \iff \forall j \in J \quad q \in (M_j)_{x_n}$$

$$\iff q \in \bigcap_{j \in J} (M_j)_{x_n},$$

proving (2.37a). One has

$$q \in \left( \bigcup_{j \in J} M_j \right)_{x_n} \iff (q, x_n) \in \bigcup_{j \in J} M_j \iff \exists j \in J \quad (q, x_n) \in M_j \iff \exists j \in J \quad q \in (M_j)_{x_n}$$

$$\iff q \in \bigcup_{j \in J} (M_j)_{x_n},$$

proving (2.37b).
proving (2.37b). One has

\[ q \in (M^c)_{x_n} \quad \iff \quad (q, x_n) \in M^c \quad \iff \quad (q, x_n) \in X \setminus M \quad \iff \quad q \in Q \setminus M_{x_n} \]

proving (2.37c). One has

\[ q \in \left( \prod_{i=1}^{n} E_i \right)_{x_n} \quad \iff \quad (q, x_n) \in \left( \prod_{i=1}^{n-1} E_i \right) \times E_n, \]

proving (2.37d). One has

\[ q \in f(\cdot, x_n)^{-1}(B) \quad \iff \quad f(q, x_n) \in B \quad \iff \quad (q, x_n) \in f^{-1}(B) \quad \iff \quad q \in (f^{-1}(B))_{x_n}, \]

proving (2.37e).

(b): Let \( \mathcal{M} := \{ M \subseteq X : M_{x_n} \in Q \text{ for each } x_n \in X_n \} \).

We show \( \mathcal{M} \) to e a \( \sigma \)-algebra: It is \( \emptyset \in Q \), since \( \emptyset \in A_n \). Let \( M \in \mathcal{M} \) and \( x_n \in X_n \). Then, by (2.37c), \( (M^c)_{x_n} = (M_{x_n})^c \in Q \), since \( M_{x_n} \in Q \). Thus, \( M^c \in \mathcal{M} \). Let \( (M_k)_{k \in \mathbb{N}} \) be a sequence in \( \mathcal{M} \), \( x_n \in X_n \). Then, by (2.37b), \( \bigcup_{k \in \mathbb{N}} M_k \in \mathcal{M} \), showing \( \mathcal{M} \) to be a \( \sigma \)-algebra. If \( M = \prod_{i=1}^{n} A_i \) with each \( A_i \in A_i \), then (2.37d) implies \( M \in \mathcal{M} \). Thus, \( A \subseteq \mathcal{M} \), since \( \mathcal{M} \) is a \( \sigma \)-algebra. Since \( A \subseteq \mathcal{M} \) is precisely the claim of (b), the proof is done.

(c): If \( B \in B \), then, by (2.37e), \( f(\cdot, x_n)^{-1}(B) = (f^{-1}(B))_{x_n} \in Q \) by (b), since \( f^{-1}(B) \in A \) by the measurability of \( f \).

\[ \blacksquare \]

**Theorem 2.32.** Let \( n \in \mathbb{N} \) and let \( (X_i, A_i, \mu_i)_{i=1}^{n} \) be a finite family of measure spaces, \( X := \prod_{i=1}^{n} X_i, A := \otimes_{i=1}^{n} A_i \).

(a) There always exists at least one product measure on \((X, A)\).

(b) Let \( n = 2 \). If \( \mu_1 \) is \( \sigma \)-finite, then, using the notation of (2.36), one has

\[ \forall M \in A \quad x_2 \mapsto \mu_1(M_{x_2}) \quad \text{is } A_2 \text{-}\overline{B} \text{-measurable} \quad \tag{2.38a} \]

and

\[ \mu : A \longrightarrow [0, \infty], \quad \mu(M) := \int_{X_2} \mu_1(M_{x_2}) \, d\mu_2(x_2), \quad \tag{2.38b} \]

defines a product measure.

(c) Let \( n \geq 2 \). If each measure \( \mu_i \) is \( \sigma \)-finite, then there exists a unique product measure \( \mu \) on \((X, A)\), denoted by \( \otimes_{i=1}^{n} \mu_i := \mu \). Moreover, \( \otimes_{i=1}^{n} \mu_i \) is itself \( \sigma \)-finite and, using the notation of (2.36), one has

\[ \forall M \in A \quad x_n \mapsto \left( \otimes_{i=1}^{n-1} \mu_i \right)(M_{x_n}) \quad \text{is } A_n \text{-}\overline{B} \text{-measurable} \quad \tag{2.39a} \]
and
\[ \forall M \in A \quad (\otimes_{i=1}^{n} \mu_{i})(M) = \int_{X_{n}} (\otimes_{i=1}^{n-1} \mu_{i})(M_{x_{n}}) \, d\mu_{n}(x_{n}). \tag{2.39b} \]

In terms of the canonical identification \((\prod_{i=1}^{n-1} X_{i}) \times X_{n} \cong \prod_{i=1}^{n} X_{i}\), one has
\[ (\otimes_{i=1}^{n-1} \mu_{i}) \otimes \mu_{n} = \otimes_{i=1}^{n} \mu_{i}. \tag{2.39c} \]

**Proof.** (a): For \(n = 1\), there is nothing to prove. We consider the case \(n = 2\), from which the case \(n > 2\) follows by induction. Let \(n = 2\). We know from Prop. 1.17(a) that \(E := A_{1} \ast A_{2} = \{ A_{1} \times A_{2} : A_{i} \in A_{i}, i = 1, 2 \}\) is a semiring. We also know \(A = \sigma(E)\).

We now define \(\mu\) on \(E\) by (2.35), i.e. \(\mu(A_{1} \times A_{2}) := \mu_{1}(A_{1}) \mu_{2}(A_{2})\) for \(A_{1} \in A_{1}, A_{2} \in A_{2}\). We show that this defines a premeasure on \(E\): \(\mu(\emptyset) = 0\) is clear. Let \(A \in A_{1}, B \in A_{2}\), and let \((A_{i})_{i \in \mathbb{N}}\) and \((B_{i})_{i \in \mathbb{N}}\) be sequences in \(A_{1}\) and \(A_{2}\), respectively, such that \(A \times B\) is the disjoint union of the \(A_{i} \times B_{i}\), i.e. \(A \times B = \bigcup_{i \in \mathbb{N}} (A_{i} \times B_{i})\). Then
\[ \mu(A \times B) = \int_{X_{2}} \mu_{1}(A) \chi_{B}(x) \, d\mu_{2}(x) = \int_{X_{2}} \mu_{1}((A \times B)_{x}) \, d\mu_{2}(x) \]
\[ \stackrel{(2.37d)}{=} \int_{X_{2}} \mu_{1} \left( \bigcup_{i \in \mathbb{N}} (A_{i} \times B_{i})_{x} \right) \, d\mu_{2}(x) = \int_{X_{2}} \sum_{i=1}^{\infty} \mu_{1}((A_{i} \times B_{i})_{x}) \, d\mu_{2}(x) \]
\[ \cong 2.8 \sum_{i=1}^{\infty} \int_{X_{2}} \mu_{1}((A_{i} \times B_{i})_{x}) \, d\mu_{2}(x) = \sum_{i=1}^{\infty} \mu(A_{i} \times B_{i}), \]

showing \(\mu\) to be a premeasure. Then, using the Carathéodory extension Th. 1.38, \(\mu\) extends to a measure on \((X, A)\) that constitutes a product measure.

(b): For each \(M \in A = A_{1} \otimes A_{2}\), we introduce the notation \(f_{M} : X_{2} \longrightarrow [0, \infty]\), \(f_{M}(x) := \mu_{1}(M_{x})\), for the map defined in (2.38a). The bulk of the proof will consist of showing the measurability of \(f_{M}\). Defining
\[ \mathcal{M} := \{ M \in A : f_{M} \text{ is measurable} \}, \]
we have to show \(\mathcal{M} = A\). If \(A \in A_{1}, B \in A_{2}\), then
\[ \forall x \in X_{2} \quad f_{A \times B}(x) = \mu_{1}((A \times B)_{x}) \stackrel{(2.37d)}{=} \mu_{1}(A) \chi_{B}(x), \]
showing \(f_{A \times B}\) to be measurable. Thus, \(E \subseteq \mathcal{M}\), where \(E := A_{1} \ast A_{2}\) as in the proof of (a). We now consider the case, where \(\mu_{1}\) is finite, i.e. \(\mu_{1}(X_{1}) < \infty\). We show \(\mathcal{M}\) to be a Dynkin system: \(X \in \mathcal{M}\), since \(X \in E\). If \(M \in \mathcal{M}\), then, due to \(\mu_{1}(X_{1}) < \infty\), we can use subadditivity of \(\mu_{1}\) to obtain
\[ \forall x \in X_{2} \quad f_{M^{c}}(x) = \mu_{1}((M^{c})_{x}) \stackrel{(2.37e)}{=} \mu_{1}(M_{x})^{c} = \mu_{1}(X_{1}) - f_{M}(x), \]
showing \(f_{M^{c}}\) to be measurable and \(M^{c} \in \mathcal{M}\). Now let \((M_{k})_{k \in \mathbb{N}}\) be a sequence of disjoint sets in \(\mathcal{M}\), \(M := \bigcup_{k \in \mathbb{N}} M_{k}\). Then
\[ \forall x \in X_{2} \quad f_{M}(x) = \mu_{1}(M_{x}) \stackrel{(2.37b)}{=} \mu_{1} \left( \bigcup_{k \in \mathbb{N}} (M_{k})_{x} \right) = \sum_{k=1}^{\infty} \mu_{1}((M_{k})_{x}) = \sum_{k=1}^{\infty} f_{M_{k}}(x), \]
showing $f_M$ to be measurable by Th. 1.82(a), i.e. $M \in \mathcal{M}$ and $\mathcal{M}$ is a Dynkin system. Since $\mathcal{E}$ is $\cap$-stable (by [Phi6b, (A.1)]), we have $\delta(\mathcal{E}) = \sigma(\mathcal{E}) = \mathcal{A}$ by Prop. 1.44. Since also $\delta(\mathcal{E}) \subseteq \mathcal{M}$, we have $\mathcal{A} \subseteq \mathcal{M} \subseteq \mathcal{A}$, i.e. $\mathcal{M} = \mathcal{A}$, as desired. Now let $\mu_1$ be $\sigma$-finite and choose a sequence $(A_k)_{k \in \mathbb{N}}$ in $\mathcal{A}$ such that $A_k \uparrow X_1$ and $\mu_1(A_k) < \infty$ for each $k \in \mathbb{N}$. Then, clearly, for each $k \in \mathbb{N}$, $\nu_k : \mathcal{A}_1 \to \mathbb{R}^+_0$, $\nu_k(A) := \mu_1(A_k \cap A)$, defines a finite measure on $X_1$, and, for each $M \in \mathcal{A}$, $k \in \mathbb{N}$, we know

$$f_{k,M} : X_2 \to [0, \infty], \quad f_{k,M}(x) := \nu_k(M_x),$$

to be measurable. Then $f_M = \lim_{k \to \infty} f_{k,M}$ must also be measurable. In consequence, the definition of $\mu$ by (2.38b) makes sense and it merely remains to verify $\mu$ is a product measure. We clearly have $\mu(\emptyset) = 0$ and if $(M_k)_{k \in \mathbb{N}}$ is a sequence of disjoint sets in $\mathcal{A}$, $M := \bigcup_{k \in \mathbb{N}} M_k$, then

$$\mu(M) = \left(\int_{X_2} \mu_1 \left(\bigcup_{k \in \mathbb{N}} (M_k)_x\right) \, d\mu_2(x)\right) \text{Cor. 2.8} = \sum_{k=1}^{\infty} \int_{X_2} \mu_1((M_k)_x) \, d\mu_2(x) = \sum_{k=1}^{\infty} \mu(M_k),$$

showing $\mu$ to be a measure. If $A \in \mathcal{A}_1$, $B \in \mathcal{A}_2$, then

$$\mu(A \times B) = \int_{X_2} \mu_1((A \times B)_x) \, d\mu_2(x) = \int_{X_2} \mu_1(A) \chi_B(x) \, d\mu_2(x) = \mu_1(A) \mu_2(B),$$

showing $\mu$ to be a product measure.

(c): It suffices to consider the case $n = 2$, since, then, the general case $n \geq 2$ follows by induction: For the induction step, one merely applies the case $n = 2$ to the two measure spaces $(\prod_{i=1}^{n-1} X_i, \otimes_{i=1}^{n-1} \mathcal{A}_i, \otimes_{i=1}^{n-1} \mu_i)$ (which is uniquely determined and $\sigma$-finite by the induction hypothesis) and $(X_n, \mathcal{A}_n, \mu_n)$. Thus, let $n = 2$. It now suffices to show that the product measure of $\mu_1$ and $\mu_2$ is unique and $\sigma$-finite, since, then, the representation of (2.39b) is due to (b). In the proof of (a), we defined a premeasure $\mu$ on the semiring $\mathcal{E} = \mathcal{A}_1 * \mathcal{A}_2$ by setting $\mu(A_1 \times A_2) := \mu_1(A_1)\mu_2(A_2)$ for $A_1 \in \mathcal{A}_1$, $A_2 \in \mathcal{A}_2$. In our present situation, we have $\mu_1, \mu_2$ $\sigma$-finite and can choose sequences $(A_k)_{k \in \mathbb{N}}$ and $(B_k)_{k \in \mathbb{N}}$ in $\mathcal{A}_1$ and $\mathcal{A}_2$, respectively, such that $A_k \uparrow X_1$, $B_k \uparrow X_2$, and $\mu_1(A_k) < \infty$, $\mu_2(B_k) < \infty$ for each $k \in \mathbb{N}$. Then $(A_k \times B_k) \uparrow X$ with $\mu(A_k \times B_k) < \infty$, showing $\mu$ to be $\sigma$-finite on $\mathcal{E}$. According to Cor. 1.46(a), $\mu$ uniquely extends to a measure on $\sigma(\mathcal{E}) = \mathcal{A}$, proving the existence of a unique product measure. Since $X \in \mathcal{E}$, the extension is still $\sigma$-finite, completing the proof.

**Caveat 2.33.** If $\mu_1$ in Th. 2.32(b) is not $\sigma$-finite, then it can, indeed, occur that the map of (2.38a) is nonmeasurable so that (2.38b) does not even make sense (see [Beh87, p. 96] for an example). Even if $\mu_1$ is $\sigma$-finite (but $\mu_2$ is not), there can exist several different product measures on $X$: Let $X_1 := X_2 := \mathbb{R}$, $\mathcal{A}_1 := \mathcal{A}_2 := \mathcal{B}^1$, $\mu_1 := \beta^1$ (which is $\sigma$-finite), and let $\mu_2$ be counting measure (which is not $\sigma$-finite). Let $\alpha$ be the product measure constructed according to the proof of Th. 2.32(a), i.e. by defining $\alpha$ on $\mathcal{B}^1 * \mathcal{B}^1$ by letting $\alpha(A \times B) := \mu_1(A)\mu_2(B)$ for $A, B \in \mathcal{B}^1$ and extending it to $\mathcal{B}^2$ via Th. 1.38. Then $\alpha$ on $\mathcal{B}^2$ is the restriction of the corresponding outer measure on $\mathcal{P}(\mathbb{R}^2)$. Set $D := \{(x, x) \in \mathbb{R}^2 : x \in [0, 1]\}$. Then $D \in \mathcal{B}^2$ and we claim $\alpha(D) = \infty$: Seeking a
contradiction, assume $\alpha(D) < \infty$. Then there exist sequences $(A_k)_{k \in \mathbb{N}}$ and $(B_k)_{k \in \mathbb{N}}$ in $\mathcal{B}^1$ such that $D \subseteq \bigcup_{k=1}^{\infty} (A_k \times B_k)$ and $\sum_{k=1}^{\infty} \alpha(A_k \times B_k) = \sum_{k=1}^{\infty} \beta^1(A_k) \mu_2(B_k) < \infty$. This implies
\[
\forall k \in \mathbb{N} \quad \left( \beta^1(A_k) = 0 \lor B_k \text{ finite} \right).
\]
Thus, letting $A := \bigcup \{A_k : \beta^1(A_k) = 0\}$, $B := \bigcup \{B_k : B_k \text{ finite}\}$, we have $D \subseteq (A \times \mathbb{R}) \cup (\mathbb{R} \times B)$. Note $\pi_1(M) = \pi_2(M)$ for each $M \subseteq D$. We obtain $\pi_1(D \setminus (A \times \mathbb{R})) = [0, 1] \setminus A$ and $\pi_1(D \setminus (A \times \mathbb{R})) = \pi_2(D \setminus (A \times \mathbb{R})) \subseteq B$. Since $\beta^1([0, 1] \setminus A) = 1$, but $\beta^1(B) = 0$, this is a contradiction, proving $\alpha(D) = \infty$. Now let $\beta$ be the product measure given by Th. 2.32(b). Then
\[
\beta(D) = \int_{[0,1]} \beta^1([x]) \, d\mu_2(x) = \int_{[0,1]} 0 \, d\mu_2(x) = 0.
\]
It is shown in [Beh87, pp. 94–96] that, even though $\mu_2$ is not $\sigma$-finite, in our considered situation, for each $M \in \mathcal{B}^2$, the map $x \mapsto \mu_2(M_x)$, $M_x := \{y \in \mathbb{R} : (x, y) \in M\}$ is $\lambda^1$-measurable and then
\[
\gamma : \mathcal{B}^2 \longrightarrow [0, \infty], \quad \gamma(M) := \int_{[0,1]} \mu_2(M_x) \, d\lambda^1(x),
\]
again, defines a product measure. Since
\[
\gamma(D) = \int_{[0,1]} \mu_2(D_x) \, d\lambda^1(x) = \int_{[0,1]} \mu_2(\{x\}) \, d\lambda^1(x) = \int_{[0,1]} 1 \, d\lambda^1(x) = 1,
\]
we see that the three product measures $\alpha, \beta, \gamma$ must all be different.

**Example 2.34. (a)** Let $p, q \in \mathbb{N}$. We already know $\mathcal{B}^{p+q} = \mathcal{B}^p \otimes \mathcal{B}^q$ from Th. 1.94(c). This $\sigma$-algebra is generated by the semiring $\mathcal{E} := \mathcal{I}^{p+q} = \mathcal{I}^p \ast \mathcal{I}^q$. Since $\mathcal{B}^p \otimes \mathcal{B}^q$ and $\mathcal{B}^{p+q}$ agree on $\mathcal{E}$, we obtain $\mathcal{B}^p \otimes \mathcal{B}^q = \mathcal{B}^{p+q}$. Now let $X \in \mathcal{B}^p$, $Y \in \mathcal{B}^q$, $\beta_X := \mathcal{B}^p|_{\mathcal{B}^p | X}$, $\beta_Y := \mathcal{B}^q|_{\mathcal{B}^q | Y}$. We already know $\mathcal{B}^p|_{X \otimes \mathcal{B}^q | Y} = \mathcal{B}^{p+q} | (X \times Y)$ from Th. 1.94(c). This $\sigma$-algebra is generated by the semiring $\mathcal{F} := \mathcal{B}^p | X \ast \mathcal{B}^q | Y \subseteq \mathcal{B}^{p+q}$. Since we have $(\beta_X \otimes \beta_Y) |_F = (\beta_X \otimes \beta_Y) |_F = (\beta_X |_X \times \beta_Y |_Y) |_F$, we also obtain $\beta_X \otimes \beta_Y = \beta_{X \times Y}.

(b) Let $p, q \in \mathbb{N}$. Then
\[
\mathcal{L}^p \otimes \mathcal{L}^q \subseteq \mathcal{L}^{p+q} : \quad (2.40)
\]

The inclusion in (2.40) holds: Letting $\mathcal{E} := \mathcal{L}^p \ast \mathcal{L}^q$, we have $\mathcal{E} \subseteq \mathcal{L}^{p+q}$: Indeed, if $E \in \mathcal{E}$, then $E = (A \cup M) \times (B \cup N)$ with $M \subseteq M_1$, $N \subseteq N_1$, where $A, M_1 \in \mathcal{B}^p$, $B, N_1 \in \mathcal{B}^q$, $\beta^p(M_1) = \beta^q(N_1) = 0$. Then $E = (A \times B) \cup P$, where $P = (A \times N) \cup (M \times B) \cup (M \times N) \subseteq P_1 := (A \times N_1) \cup (M_1 \times B) \cup (M_1 \times N_1) \in \mathcal{B}^{p+q}$ with $\beta^{p+q}(P_1) = 0$, showing $E \in \mathcal{L}^{p+q}$. Thus $\mathcal{L}^p \otimes \mathcal{L}^q = \sigma(\mathcal{E}) \subseteq \mathcal{L}^{p+q}$. The inclusion in (2.40) is strict: Let $M \subseteq [0, 1]^p$ such that $M \notin \mathcal{L}^p$ (such a set $M$ exists by Th. 1.74(c)). Also let $y \in \mathbb{R}^q$. Then $N := M \times \{y\} \in \mathcal{L}^{p+q}$, since $N$ is a subset of the $\beta^{p+q}$-null set $[0, 1]^p \times \{y\}$. However, $N \notin \mathcal{L}^p \otimes \mathcal{L}^q$, since, otherwise, the section $N_y = M$ had to be in $\mathcal{L}^p$ by Lem. 2.31(b). In consequence, the complete measure $\lambda^{p+q}$ is a strict extension of the incomplete measure $\lambda^p \otimes \lambda^q$. 

---

2 INTEGRATION

89
(c) Let \( n \in \mathbb{N} \). We compute the volume of \( n \)-dimensional balls with respect to the Euclidean norm: Let \( x \in \mathbb{R}^n, r > 0, B_r(x) := \{ y \in \mathbb{R}^n : \| y - x \|_2 < r \} \).

\[
\text{vol}(B_r(x)) := \beta^n(B_r(x)) = \omega_n r^n, \quad \text{where} \quad \omega_n := \beta^n(B_1(0)) = \frac{\pi^{n/2}}{\Gamma(\frac{n}{2} + 1)},
\]

where \( \Gamma \) denotes the Gamma function from Ex. 2.29: It suffices to prove the formula for \( \omega_n \): Indeed, consider the bijective affine map \( A : \mathbb{R}^n \rightarrow \mathbb{R}^n, A(v) := rv + x \). Then \( A(B_1(0)) = B_r(x), \det A = r^n, \) and \( \beta^n(B_r(x)) = r^n \beta^n(B_1(0)) \) by (1.47b).

We now prove the formula for \( \omega_n \) via induction on \( n \in \mathbb{N} \): For \( n = 1 \), the claim is \( \omega_1 = \beta^1([-1, 1]) = 2 = \sqrt{\pi} \), which holds, as \( \Gamma(\frac{3}{2}) = \frac{1}{2} \Gamma(\frac{1}{2}) = \frac{1}{2} \sqrt{\pi} \) by (2.29a) and (2.33).

For the induction step, let \( n > 1 \), write \( \mathbb{R}^n = \mathbb{R}^{n-1} \times \mathbb{R} \) and use \( \beta^n = \beta^{n-1} \otimes \beta^1 \). For \( x \in \mathbb{R} \), we obtain the sections

\[
(B_1(0))_x = \left\{ y \in \mathbb{R}^{n-1} : x^2 + \sum_{i=1}^{n-1} y_i^2 < 1 \right\} = \left\{ \begin{array}{ll}
\emptyset & \text{for } |x| \geq 1, \\
B_{\sqrt{1-x^2}}^{-1}(0) & \text{for } -1 < x < 1,
\end{array} \right.
\]

where \( B_{\sqrt{1-x^2}}^{-1}(0) \) denotes a ball in \( \mathbb{R}^{n-1} \). Thus,

\[
\omega_n = \int_{-1}^{1} \beta^{-1}(B_1(0))_x \, d\beta^1(x) = \omega_{n-1} \int_{-1}^{1} (1 - x^2)^{(n-1)/2} \, dx.
\]

As we may consider the last integral as a 1-dimensional Riemann integral, we may use the change of variables \( t := \arccos x \) (recall \( \arccos'(x) = -(1-x^2)^{-1/2} \)) to obtain

\[
\omega_n = -\omega_{n-1} \int_{\pi}^{0} (1 - (\cos t)^2)^{n/2} \, dt = \omega_{n-1} \int_{0}^{\pi} (\sin t)^n \, dt.
\]

In particular, for \( n = 2 \), \( \omega_2 = \omega_1 \int_{0}^{\pi} (\sin t)^2 \, dt \) \( \text{(E.3a)} \) \( = 2 \cdot \frac{\pi}{2} = \pi = \frac{\pi}{\Gamma(2)} \), proving (2.41) for \( n = 2 \). For \( n > 2 \),

\[
\omega_n = \omega_{n-2} \int_{0}^{\pi} (\sin t)^{n-1} \, dt \int_{0}^{\pi} (\sin t)^n \, dt.
\]

Finally, using the induction hypothesis and (E.3), we obtain, for \( n \) even

\[
\omega_n = \frac{\pi^{(n-2)/2}}{\Gamma(\frac{n}{2})} \frac{2\pi}{2} \prod_{k=1}^{n/2-1} \frac{2k}{2k+1} \prod_{k=1}^{n/2} \frac{2k-1}{2k} = \frac{2\pi^{n/2}}{n\Gamma(\frac{n}{2})} \quad \text{(2.29a)}
\]

and, for \( n \) odd,

\[
\omega_n = \frac{\pi^{(n-2)/2}}{\Gamma(\frac{n}{2})} \frac{2\pi}{2} \prod_{k=1}^{(n-1)/2} \frac{2k-1}{2k} \prod_{k=1}^{(n-1)/2} \frac{2k}{2k+1} = \frac{2\pi^{n/2}}{n\Gamma(\frac{n}{2})} \quad \text{(2.29a)}
\]

completing the induction.
2.5.2 Theorems of Tonelli and Fubini

For the actual computation of higher-dimensional integrals, the theorems of Tonelli and Fubini, which we provide in Th. 2.36 below, are among the most powerful tools, as they allow, under suitable hypotheses, to compute an $n$-dimensional integral as a sequence of $n$ 1-dimensional integrals.

The literature knows many variants of the theorems of Tonelli and Fubini, where different variants can have subtle differences, and one has to use some care in applications to always verify that the variant one desires to use does, indeed, apply. The most natural variant in our setting is probably the one that considers two $\sigma$-finite measure spaces $(X_1, \mathcal{A}_1, \mu_1)$ and $(X_2, \mathcal{A}_2, \mu_2)$ and a function $f$ that is $(\mu_1 \otimes \mu_2)$-measurable or $(\mu_1 \otimes \mu_2)$-integrable. However, one would also like to apply the theorem to functions that are $\lambda^{p+q}$-measurable or $\lambda^{p+q}$-integrable and, recalling Ex. 2.34(b), $\mathcal{L}^p \otimes \mathcal{L}^q \subset \mathcal{L}^{p+q}$. For this reason, we also provide a version that assumes $f$ to be $\mu$-measurable or $\mu$-integrable, where $\mu$ is the completion of $\mu_1 \otimes \mu_2$.

The following convention is sometimes useful, especially in the context of the theorems of Tonelli and Fubini:

**Convention 2.35.** Let $(X, \mathcal{A}, \mu)$ be a measure space, $N \in \mathcal{A}$ with $\mu(N) = 0$, $f : N^c \to \hat{\mathbb{K}}, z \in f(X)$. Define

$$g : X \to \hat{\mathbb{K}}, \ g(x) := \begin{cases} f(x) & \text{for } x \in N^c, \\ z & \text{for } x \in N. \end{cases}$$

If $f$ is integrable or nonnegative measurable (with respect to $\mathcal{A}$ and $\mu$ restricted to $N^c$), then, clearly, $g$ is $\mu$-integrable or nonnegative $\mu$-measurable. In this case, one defines

$$\int_X f \, d\mu := \int_X g \, d\mu$$

and observes that this definition does not depend on the choice of $z \in f(X)$ (since $N$ is a $\mu$-null set). In practise, one will usually still denote $g$ by $f$ and one often does not even specify $z$ explicitly.

**Theorem 2.36.** Let $n \in \mathbb{N}, n \geq 2$. For each $i \in \{1, \ldots, n\}$, let $(X_i, \mathcal{A}_i, \mu_i)$ be a $\sigma$-finite measure space, $X := \prod_{i=1}^n X_i$. We consider two cases:

(i) $(X, \mathcal{A}, \mu)$ is the corresponding product measure space.

(ii) Each $\mu_i$ is complete and $(X, \mathcal{A}, \mu)$ is the completion of the corresponding product measure space.

In both cases, the following holds:

(a) Tonelli’s Theorem: Let $f : X \to [0, \infty]$ be $\mathcal{A}$-measurable. For $n = 2$, the function $f(x_1, \cdot)$ is $\mathcal{A}_2$-measurable (for every $x_1 \in X_1$ in Case (i), for $\mu_1$-almost every $x_1 \in$
2 INTEGRATION

X_1 in Case (ii)), the function f(·, x_2) is A_1-measurable (for every x_2 ∈ X_2 in Case (i), for μ_2-almost every x_2 ∈ X_2 in Case (ii)), the functions given by

\[ x_1 \mapsto \int_{X_2} f(x_1, x_2) \, d\mu_2(x_2), \quad x_2 \mapsto \int_{X_1} f(x_1, x_2) \, d\mu_1(x_1), \]  

(2.42a)

are A_1-measurable and A_2-measurable, respectively (in Case (ii), they are defined according to Convention 2.35), and

\[ \int_{X_1 \times X_2} f(x_1, x_2) \, d\mu(x_1, x_2) = \int_{X_2} \int_{X_1} f(x_1, x_2) \, d\mu_1(x_1) \, d\mu_2(x_2) = \int_{X_1} \int_{X_2} f(x_1, x_2) \, d\mu_2(x_2) \, d\mu_1(x_1). \]  

(2.42b)

For n ≥ 2, by applying (2.42b) inductively, one obtains

\[ \int_X f \, d\mu = \int_{X_1} \ldots \int_{X_n} f(x_1, \ldots, x_n) \, d\mu_n(x_n) \ldots d\mu_1(x_1), \]  

(2.42c)

where one can also arbitrarily permute the order of the inner integrals without changing the overall value (as above, the inner integrals are defined according to Convention 2.35 in Case (ii)).

(b) Fubini’s Theorem: Let f : X → ˆK be μ-integrable. For n = 2, the function f(x_1, ·) is μ_2-integrable for μ_1-almost every x_1 ∈ X_1 (in both cases), the function f(·, x_2) is μ_1-integrable for μ_2-almost every x_2 ∈ X_2 (in both cases), the functions given by (2.42a) are μ_1-integrable and μ_2-integrable, respectively (in the sense of Convention 2.35), and (2.42b) holds. As in (a), for n ≥ 2, by applying (2.42b) inductively, one obtains the validity of (2.42c), where, once again, one can also arbitrarily permute the order of the inner integrals without changing the overall value (of course, as before, the inner integrals are defined according to Convention 2.35).

(c) If f : X → ˆK is μ-measurable and one of the integrals

\[ \int_X |f| \, d\mu, \]

\[ \int_{X_{\pi(1)}} \ldots \int_{X_{\pi(n)}} |f(x_1, \ldots, x_n)| \, d\mu_{\pi(n)}(x_{\pi(n)}) \ldots d\mu_{\pi(1)}(x_{\pi(1)}), \]  

(2.43)

where π is a permutation of \{1, \ldots, n\}, is finite, then all the integrals (for an arbitrary permutation π of \{1, \ldots, n\}) are finite, f is μ-integrable and, in particular, all the assertions of (b) hold true.

Proof. It suffices to prove the case n = 2 – then the case n > 2 follows by induction (and using that every permutation is a finite composition of transpositions).
(a): Case (i): Let $M \in \mathcal{A} = \mathcal{A}_1 \otimes \mathcal{A}_2$. If $x_2 \in X_2$, then, for the section $M_{x_2} \subseteq X_1$, one has $M_{x_2} \in \mathcal{A}_1$ by Lem. 2.31(b), and $x_2 \mapsto \mu_1(M_{x_2})$ is $\mathcal{A}_2$-$\mathcal{B}$-measurable by Th. 2.32(b).

One computes

\[
\int_X \chi_M(x_1, x_2) \, d\mu(x_1, x_2) = \mu(M) \overset{(2.38b)}{=} \int_{X_2} \mu_1(M_{x_2}) \, d\mu_2(x_2)
\]

\[
= \int_{X_2} \int_{X_1} \chi_{M_{x_2}}(x_1) \, d\mu_1(x_1) \, d\mu_2(x_2) = \int_{X_2} \int_{X_1} \chi_M(x_1, x_2) \, d\mu_1(x_1) \, d\mu_2(x_2)
\]

proving the first equality of (2.42b) for characteristic functions of measurable sets. Now let $f$ be a simple function $f \in \mathcal{S}^+(\mathcal{A})$, i.e., $f = \sum_{k=1}^N \alpha_k \chi_{M_k}$ with $\alpha_1, \ldots, \alpha_N \in \mathbb{R}_0^+$; $M_1, \ldots, M_n \in \mathcal{A}$; $N \in \mathbb{N}$. Then the linearity of the integrals yields

\[
\int_X f \, d\mu = \sum_{k=1}^N \alpha_k \int_X \chi_{M_k}(x_1, x_2) \, d\mu(x_1, x_2)
\]

\[
= \sum_{k=1}^N \alpha_k \int_{X_2} \int_{X_1} \chi_{M_k}(x_1, x_2) \, d\mu_1(x_1) \, d\mu_2(x_2)
\]

\[
= \int_{X_2} \int_{X_1} f(x_1, x_2) \, d\mu_1(x_1) \, d\mu_2(x_2), \quad (2.44)
\]

proving the first equality of (2.42b) for nonnegative simple functions. Next, we consider $f \in \mathcal{M}^+(\mathcal{A})$ and choose a sequence $(\phi_k)_{k \in \mathbb{N}}$ in $\mathcal{S}^+(\mathcal{A})$ such that $\phi_k \uparrow f$. According to Lem. 2.31(c), for each $x_2 \in X_2$, $f(\cdot, x_2) \in \mathcal{M}^+(\mathcal{A}_1)$ and, for each $k \in \mathbb{N}$, $\phi_k(\cdot, x_2) \in \mathcal{S}^+(\mathcal{A}_1)$. Since $\phi_k(\cdot, x_2) \uparrow f(\cdot, x_2)$, we obtain, according to (2.4)

\[
\int_{X_1} \phi_k(x_1, x_2) \, d\mu_1(x_1) \uparrow \int_{X_1} f(x_1, x_2) \, d\mu_1(x_1). \quad (2.45)
\]

Moreover, each function $x_2 \mapsto \int_{X_1} \phi_k(x_1, x_2) \, d\mu_1(x_1)$ on the left-hand side of (2.45) is $\mathcal{A}_2$-measurable, implying the function $x_2 \mapsto \int_{X_1} f(x_1, x_2) \, d\mu_1(x_1)$ on the right-hand side to be $\mathcal{A}_2$-measurable as well. We now compute

\[
\int_X f \, d\mu \overset{(2.4)}{=} \lim_{k \to \infty} \int_X \phi_k \, d\mu = \lim_{k \to \infty} \int_{X_2} \int_{X_1} \phi_k(x_1, x_2) \, d\mu_1(x_1) \, d\mu_2(x_2)
\]

\[
= \int_{X_2} \int_{X_1} \lim_{k \to \infty} \phi_k(x_1, x_2) \, d\mu_1(x_1) \, d\mu_2(x_2)
\]

\[
\overset{\text{MCT}}{=} \int_{X_2} \int_{X_1} \phi_k(x_1, x_2) \, d\mu_1(x_1) \, d\mu_2(x_2)
\]

\[
\overset{(2.4)}{=} \int_{X_2} \int_{X_1} f(x_1, x_2) \, d\mu_1(x_1) \, d\mu_2(x_2), \quad (2.46)
\]

proving the first equality of (2.42b) for each $f \in \mathcal{M}^+(\mathcal{A})$ as claimed. To prove the second equality of (2.42b) as well as all the remaining claims of (a) (in Case (i)), one merely exchanges the roles of $(X_1, \mathcal{A}_1, \mu_1)$ and $(X_2, \mathcal{A}_2, \mu_2)$ in the above argument.

Case (ii): Let $M \in \mathcal{A} = (\mathcal{A}_1 \otimes \mathcal{A}_2)^\sim$. Then there exist $A, C \in \mathcal{A}_1 \otimes \mathcal{A}_2$ and $N \subseteq C$, such that $(\mu_1 \otimes \mu_2)(C) = 0$ and $M = A \cup N$. If $x_2 \in X_2$, then, for the section $M_{x_2} \subseteq X_1$, one
Moreover, for each \( \mu \)-null set is still a \( \mu \)-null set. In other words, \( \int X \chi_M(x_1,x_2) \, d\mu(x_1,x_2) = \mu(M) = \mu(A) \) \((\text{2.38b})\) \(= \int_{X_2} \mu_1(A_{x_2}) \, d\mu_2(x_2) \)

\[
= \int_{X_2} \int_{X_1} \chi_{A_{x_2}}(x_1) \, d\mu_1(x_1) \, d\mu_2(x_2) = \int_{X_2} \int_{X_1} \chi_M(x_1,x_2) \, d\mu_1(x_1) \, d\mu_2(x_2),
\]

proving the first equality of (2.42b) for characteristic functions of measurable sets. If \( f \in \mathcal{S}^+(\mathcal{A}) \), \( f = \sum_{k=1}^N \alpha_k \chi_{M_k} \) as in the proof of Case (i), then the linearity of the integrals, once again, yields the validity of (2.44) proving the first equality of (2.42b) for \( f \in \mathcal{S}^+(\mathcal{A}) \). As in Case (i), we now consider \( f \in \mathcal{M}^+(\mathcal{A}) \) and choose a sequence \((\phi_k)_{k \in \mathbb{N}}\) in \( \mathcal{S}^+(\mathcal{A}) \) such that \( \phi_k \uparrow f \). We have already seen above that, if \( M \in \mathcal{A} \), then \( \chi_M(\cdot,x_2) \) is \( \mu_1 \)-measurable for \( \mu_2 \)-a.e. \( x_2 \in X_2 \). Since a countable union of \( \mu_2 \)-null sets is still a \( \mu_2 \)-null set, there exists a \( \mu_2 \)-null \( N \) in \( \mathcal{A}_2 \) such that all \( \phi_k(\cdot,x_2), k \in \mathbb{N} \), are \( \mu_1 \)-measurable for each \( x_2 \in N^c \). This shows \( f(\cdot,x_2) \) to be \( \mu_1 \)-measurable for each \( x_2 \in N^c \) (i.e. for \( \mu_2 \)-a.e. \( x_2 \in X_2 \)) as well. Thus, for each \( x_2 \in N^c \), \( f(\cdot,x_2) \in \mathcal{M}^+(\mathcal{A}_1) \), \( \phi_k(\cdot,x_2) \in \mathcal{S}^+(\mathcal{A}_1) \), \( \phi_k(\cdot,x_2) \uparrow f(\cdot,x_2) \), and (2.45) holds. As in Case (i), each function \( x_2 \mapsto \int_{X_1} \phi_k(x_1,x_2) \, d\mu_1(x_1) \) on the left-hand side of (2.45) (extended to \( N \) by 0) is \( \mathcal{A}_2 \)-measurable, implying the function \( x_2 \mapsto \int_{X_1} f(x_1,x_2) \, d\mu_1(x_1) \) on the right-hand side to be \( \mathcal{A}_2 \)-measurable as well. Then, in terms of Convention 2.35, (2.46) still holds, proving the first equality of (2.42b) for each \( f \in \mathcal{M}^+(\mathcal{A}) \). Once again, exchanging the roles of \((X_1,\mathcal{A}_1,\mu_1)\) and \((X_2,\mathcal{A}_2,\mu_2)\) proves the remaining claims of (a).

(b): If \( f \) is \( \mu \)-integrable, then so is \(|f|\) and (a) yields

\[
\int_{X_2} \int_{X_1} |f(x_1,x_2)| \, d\mu_1(x_1) \, d\mu_2(x_2) = \int_X |f| \, d\mu < \infty,
\]

implying

\[
\int_{X_1} |f(x_1,x_2)| \, d\mu_1(x_1) < \infty \quad \mu_2\text{-a.e.}
\]

In other words, \( f(\cdot,x_2) \) is \( \mu_1 \)-integrable for each \( x_2 \in N^c \), where \( N \) is a \( \mu_2 \)-null set, i.e., letting \( f_1 := (\text{Re} \, f)^+, f_2 := -(\text{Re} \, f)^-, f_3 := i \, (\text{Im} \, f)^+, f_4 := -i \, (\text{Im} \, f)^- \),

\[
\forall x_2 \in N^c \quad \int_{X_1} f(x_1,x_2) \, d\mu_1(x_1) = \sum_{k=1}^4 \int_{X_1} f_k(x_1,x_2) \, d\mu_1(x_1).
\]

Moreover, for each \( k \in \{1, 2, 3, 4\} \), \( x_2 \mapsto \int_{X_1} f_k(x_1,x_2) \, d\mu_1(x_1) \) (extended to \( N \) by 0) is \( \mu_2 \)-integrable, since

\[
\int_{X_2} \int_{X_1} |f_k(x_1,x_2)| \, d\mu_1(x_1) \, d\mu_2(x_2) \leq \int_{X_2} \int_{X_1} |f(x_1,x_2)| \, d\mu_1(x_1) \, d\mu_2(x_2) < \infty.
\]
We apply \((a)\) to each \(|f_k|\) to obtain
\[
\int_{X_2} \int_{X_1} f(x_1, x_2) \, d\mu_1(x_1) \, d\mu_2(x_2) = \sum_{k=1}^{4} \int_{X_2} \int_{X_1} f_k(x_1, x_2) \, d\mu_1(x_1) \, d\mu_2(x_2) = 4 \sum_{k=1}^{4} \int f_k \, d\mu = \int f \, d\mu,
\]
i.e. the first equality of (2.42b). Exchanging the roles of \((X_1, A_1, \mu_1)\) and \((X_2, A_2, \mu_2)\) completes the proof of \((b)\).

\((c)\) is just \((a)\) combined with \((b)\).  

The technique employed in the proof of Th. 2.36 can often be employed to prove an assertion about integrable functions: One first proves the assertion for characteristic functions of measurable sets, then for simple functions (linear combinations of characteristic functions of measurable sets), then for nonnegative measurable functions (pointwise limits of simple functions), then, finally, for integrable functions \(f\) (whose parts \((\text{Re} f)^\pm\) and \((\text{Im} f)^\pm\) are nonnegative measurable).

**Example 2.37.**  
(a) Let \(I := [1, 2] \times [2, 3] \times [0, 2]\). We use the Fubini theorem to compute \(\int_I f := \int_I f \, d\lambda^3\) for
\[
f : I \longrightarrow \mathbb{R}, \quad f(x, y, z) := \frac{2z}{(x+y)^2};
\]
Since \(f\) is continuous on the compact interval \(I\), it is integrable and Fubini applies, yielding
\[
\int_I f = \int_1^2 \int_2^3 \int_0^{z_2} \frac{2z}{(x+y)^2} \, dz \, dy \, dx = \int_1^2 \int_2^{z_2} \left[ \frac{z^2}{(x+y)^2} \right]_{z_2=0}^{z_2=2} \, dy \, dx \\
= \int_1^2 \int_2^3 \frac{4}{(x+y)^2} \, dy \, dx = \int_1^2 -\frac{4}{x+y} \big|_{y=2}^{y=3} \, dx = \int_1^2 -\frac{4}{x+3} + \frac{4}{x+2} \, dx \\
= 4 \left[ \ln(x+2) - \ln(x+3) \right]_1^2 = 4 \ln\left( \frac{4 \cdot 4 \cdot \frac{1}{3} \cdot \frac{1}{5} }{15} \right) = 4 \ln \frac{16}{15}
\]
(b) Using the Tonelli theorem, we provide another proof for \(\int_{-\infty}^{\infty} e^{-x^2} \, dx = \sqrt{\pi}\) of (2.33): We let \(I := [0, \infty]^2\) and consider
\[
f : I \longrightarrow \mathbb{R}_0^+, \quad f(x, y) := y e^{-(1+x^2)y^2}.
\]
Since \(f\) is continuous, it is measurable. It is also nonnegative, i.e. \(f \in \mathcal{M}^+\) and the Tonelli theorem applies. Thus, we can compute \(\int_I f := \int_I f \, d\lambda^2\) in two different
ways, where, for the second way, we also use the change of variable \( t := xy \),

\[
\int f = \int_0^\infty \int_0^\infty ye^{-(1+x^2)y^2} \, dy \, dx = \int_0^\infty \left[ \frac{e^{-(1+x^2)y^2}}{2(1+x^2)} \right]_{y=0}^{y=\infty} \, dx
\]

\[
= \frac{1}{2} \int_0^\infty \frac{1}{1+x^2} \, dx = \frac{1}{2} [\arctan x]_0^\infty = \frac{\pi}{4},
\]

\[
\int f = \int_0^\infty e^{-y^2} \int_0^\infty ye^{-x^2y^2} \, dx \, dy = \int_0^\infty e^{-y^2} \int_0^\infty e^{-t^2} \, dt \, dy = \left( \int_0^\infty e^{-t^2} \, dt \right)^2,
\]

proving \( \int_{-\infty}^{\infty} e^{-x^2} \, dx = \sqrt{\pi} \).

(c) The following example shows that, in general, the Fubini theorem does not hold for nonintegrable functions: It can happen that all iterated integrals exist, but the result depends on the order in which they are executed: Consider \( X_1 := X_2 := [0, 1[, A_1 := A_2 := \mathcal{B}^1, \mu_1 := \mu_2 := \beta^1 \). Consider the continuous function

\[
f : X \rightarrow \mathbb{R}, \quad f(x, y) := \frac{x^2 - y^2}{(x^2 + y^2)^2}.
\]

We show

\[
\int_0^1 \int_0^1 f(x, y) \, dy \, dx = \frac{\pi}{4} \neq -\frac{\pi}{4} = \int_0^1 \int_0^1 f(x, y) \, dx \, dy.
\]

Observe that, on \( X \),

\[
\partial_y \arctan \frac{x}{y} = -\frac{1}{1 + \frac{x^2}{y^2}}, \quad \frac{x}{y^2} = -\frac{x}{x^2 + y^2},
\]

\[
\partial_x \partial_y \arctan \frac{x}{y} = -\frac{x^2 + y^2 - 2x^2}{(x^2 + y^2)^2} = -\frac{x^2 - y^2}{(x^2 + y^2)^2},
\]

\[
\partial_x \arctan \frac{y}{x} = -\frac{y}{x^2 + y^2},
\]

\[
\partial_y \partial_x \arctan \frac{y}{x} = \frac{y^2 - x^2}{(x^2 + y^2)^2}.
\]

Thus,

\[
\int_0^1 \int_0^1 f(x, y) \, dy \, dx = \int_0^1 \left[ \frac{y}{x^2 + y^2} \right]_{y=0}^{y=1} \, dx = \int_0^1 \frac{1}{1 + x^2} \, dx = [\arctan x]_0^1 = \frac{\pi}{4},
\]

\[
\int_0^1 \int_0^1 f(x, y) \, dx \, dy = \int_0^1 \left[ -\frac{x}{x^2 + y^2} \right]_{x=0}^{x=1} \, dy = -\int_0^1 \frac{1}{1 + y^2} \, dy = -\frac{\pi}{4}.
\]

In particular, \( f \) cannot be \( \beta^2 \)-integrable over \( X \).
2.6 \( L^p \)-Spaces

2.6.1 Definition, Norm, Completeness

**Notation 2.38.** Let \((X, \mathcal{A}, \mu)\) be a measure space. Using the convention
\[
\forall \quad \infty^p := \infty,\quad (2.47)
\]
we define for every measurable \( f : X \to \hat{\mathbb{K}} \):
\[
\forall \quad p \in \mathbb{R}^+, \quad N_p(f) := \left( \int_X |f|^p \, d\mu \right)^{1/p} \in [0, \infty] \quad (2.48a)
\]
and
\[
N_\infty(f) := \inf \left\{ \sup \left\{ |f(x)| : x \in X \setminus N \right\} : N \in \mathcal{A}, \mu(N) = 0 \right\} \in [0, \infty], \quad (2.48b)
\]
where the number \( N_\infty(f) \) is also known as the **essential supremum** of \( f \) (in (2.48b), it is useful to set \( \sup \emptyset := \inf [0, \infty) = 0 \) – this setting is only relevant if \( \mu(X) = 0 \), resulting in \( N_\infty(f) = 0 \)). Note that \( |f| \leq N_\infty(f) \) \( \mu \)-a.e., since, for \( N_\infty(f) < \infty \),
\[
\left\{ |f| > N_\infty(f) \right\} = \bigcup_{n \in \mathbb{N}} \left\{ |f| > N_\infty(f) + \frac{1}{n} \right\}
\]
is a \( \mu \)-null set.

**Theorem 2.39** (Hölder Inequality). Let \((X, \mathcal{A}, \mu)\) be a measure space and \( p, q \in [1, \infty) \) such that \( \frac{1}{p} + \frac{1}{q} = 1 \) (note \( 1/\infty = 0 \)). If \( f, g : X \to \hat{\mathbb{K}} \) are measurable, then
\[
N_1(fg) \leq N_p(f) N_q(g), \quad (2.49)
\]
where (2.49) is known as Hölder inequality. We point out the following special cases:

(a) For \( p = q = 2 \), one obtains the Cauchy-Schwarz inequality. 
(b) If \( \mu \) is counting measure on \( \{1, \ldots, n\} \), \( n \in \mathbb{N} \), then one recovers the Hölder inequality of [Phi16b, Th. 1.12]. 
(c) If \( \mu \) is counting measure on \( \mathbb{N} \), then one obtains the Hölder (and, in particular, Cauchy-Schwarz) inequality for series of real (or complex) numbers.

**Proof.** If \( p = 1, q = \infty, \) then
\[
N_1(fg) = \int_X |fg| \, d\mu \leq N_\infty(g) \int_X |f| \, d\mu
\]
proves (2.49), and the case \( p = \infty, q = 1 \) is analogous. Thus, let \( 1 < p, q < \infty \). If \( N_p(f) = 0 \) or \( N_q(g) = 0 \), then \( fg = 0 \) \( \mu \)-a.e., implying \( N_1(fg) = 0 \) and (2.49). As (2.49) is also immediate for \( N_p(f) N_q(g) = \infty \), it merely remains to consider the case \( 0 < N_p(f), N_q(g) < \infty \). Recall that, according to [Phi16a, Th. 9.35], if \( n \in \mathbb{N}, x_1, \ldots, x_n \geq 0 \)
and \(\lambda_1, \ldots, \lambda_n > 0\) such that \(\lambda_1 + \cdots + \lambda_n = 1\), then \(x_1^{\lambda_1} \cdots x_n^{\lambda_n} \leq \lambda_1 x_1 + \cdots + \lambda_n x_n\).

Using this with \(n = 2\), \(\lambda_1 := \frac{1}{p}, \lambda_2 := \frac{1}{q}, x_1 := (|f|/N_p(f))^p, x_2 := (|g|/N_q(g))^q\), yields

\[
\frac{|f|}{N_p(f)} \frac{|g|}{N_q(g)} \leq \frac{1}{p} \left( \frac{|f|}{N_p(f)} \right)^p + \frac{1}{q} \left( \frac{|g|}{N_q(g)} \right)^q
\]

(which, trivially, also holds for \(|f| = \infty\) or \(|g| = \infty\)). Integrating this inequality over \(X\) proves (2.49). ■

As in the special case of \(\mathbb{K}^n\) in Analysis II, we obtain the Minkowski inequality (triangle inequality for \(N_p\)) from the Hölder inequality:

**Theorem 2.40.** Let \((X, \mathcal{A}, \mu)\) be a measure space and let \(f, g : X \to \mathbb{K}\) be measurable functions.

(a) **Minkowski Inequality:** If \(p \in [1, \infty]\), then

\[
N_p(f + g) \leq N_p(f) + N_p(g).
\]

We point out the following special cases: (a) If \(\mu\) is counting measure on \(\{1, \ldots, n\}, n \in \mathbb{N}\), then one recovers the Minkowski inequality of [Phi16b, Th. 1.13]. (b) If \(\mu\) is counting measure on \(\mathbb{N}\), then one obtains the Minkowski inequality for series of real (or complex) numbers.

(b) If \(p \in [0, 1]\), then one has the following inequalities:

\[
N_p^p(f + g) \leq N_p^p(f) + N_p^p(g),
\]

\[
N_p(f + g) \leq 2^{\frac{1}{p}-1} (N_p(f) + N_p(g)).
\]

**Proof.** (a): If \(N_p(f + g) = 0\) or \(N_p(f) = \infty\) or \(N_p(g) = \infty\), then (2.50) is trivially true. If \(p = 1\), then

\[
N_1(f + g) = \int_X (|f + g|) \, d\mu \leq \int_X |f| \, d\mu + \int_X |g| \, d\mu = N_1(f) + N_1(g).
\]

If \(p = \infty\), then

\[
\sup_{A \subseteq X} \{ |f(x) + g(x)| : x \in A \} \leq \sup \{ |f(x)| : x \in A \} + \sup \{ |g(x)| : x \in A \},
\]

proves \(N_\infty(f + g) \leq N_\infty(f) + N_\infty(g)\). For the remaining case, consider \(1 < p < \infty\) and \(f, g\) such that \(N_p(f + g) > 0\) as well as \(0 \leq N_p(f), N_p(g) < \infty\). Set \(q := \left(1 - \frac{1}{p}\right)^{-1} = \frac{p}{p-1}\)
and apply the Hölder inequality (2.49) to obtain

\[ N_p(f + g) = \int_X |f + g|^p \, d\mu \leq \int_X |f| |f + g|^{p-1} \, d\mu + \int_X |g| |f + g|^{p-1} \, d\mu \]

(2.49)

\[ \leq (N_p(f) + N_p(g)) \left( \int_X |f + g|^{(p-1)q} \, d\mu \right)^{1/q} \]

\[ = \left( N_p(f) + N_p(g) \right)^{p/q} \]

\[ \leq \left( N_p(f) + N_p(g) \right)^{p/q} \]

Next, we use

\[ N(f) = 0, \quad \psi \leq |a|, \quad \psi(t) = a^p + t^p - (a + t)^p, \]

which is, clearly, continuous and, for \( t > 0 \), differentiable. Moreover,

\[ \forall t > 0 \quad \phi'(t) = pt^{p-1} - p(a + t)^{p-1} > 0 \]

(since \( p - 1 < 0 \), it is \( t^{p-1} > (a + t)^{p-1} \)). Thus, \( \phi \) is strictly increasing and, noting \( \phi(0) = 0 \), we have \( a^p + t^p - (a + t)^p \geq 0 \) for each \( t \geq 0 \). We restate our findings as

\[ \forall_{a,b \in \mathbb{R}_0^+} (a + b)^p \leq a^p + b^p. \]

Applying this with \( a := |f| \) and \( b := |g| \) yields (2.51a) after an integration over \( X \). To prove (2.51b), we, once again, fix \( a \in \mathbb{R}^+ \), this time considering the function

\[ \psi : \mathbb{R}_0^+ \rightarrow \mathbb{R}, \quad \psi(t) := (a^{1/p} + t^{1/p}) (a + t)^{-1/p}, \]

which is also, clearly, continuous and, for \( t > 0 \), differentiable, where

\[ \psi'(t) = \frac{1}{p} t^\frac{1}{p} - 1 (a + t)^{-1/p} - \frac{1}{p} (a^{1/p} + t^{1/p}) (a + t)^{-\frac{1}{p} - 1} \]

\[ \leq \psi'(t) = \frac{(a + t)^{-\frac{1}{p} - 1} t^\frac{1}{p} - 1}{p} (t^\frac{1}{p} - 1 (a + t)^{-\frac{1}{p} - 1} - a^\frac{1}{p} (a + t)^{-\frac{1}{p} - 1}) \]

\[ = \frac{a (a + t)^{-\frac{1}{p} - 1} t^\frac{1}{p} - 1}{p} (t^\frac{1}{p} - 1 - a^\frac{1}{p} (a + t)^{-\frac{1}{p} - 1}). \]
proves \( t \) global min at \( t = a, \psi(a) = 2a^{1/p}2^{-1/p}a^{-1/p} = 2^{1-\frac{1}{p}}. \) Taking reciprocals in \( \psi(b) \geq \psi(a) \) proves 

\[
\forall_{a,b \in [0,\infty]} \quad (a+b)^{1/p} \leq 2\frac{1}{p-1}(a^{1/p} + b^{1/p}).
\]

We apply this with \( a := N_p^p(f) \) and \( b := N_p^p(g) \) to obtain

\[
N_p(f+g) \overset{(2.51a)}{=} (N_p^p(f) + N_p^p(g))^{1/p} \leq 2\frac{1}{p-1}(N_p(f) + N_p(g)),
\]

proving (2.51b).

\[ \blacksquare \]

**Definition and Remark 2.41.** Let \( (X, \mathcal{A}, \mu) \) be a measure space and \( p \in ]0,\infty]. \)

(a) Let \( \mathcal{L}^p := \mathcal{L}^p_{\infty} := \mathcal{L}^p_{\infty}(\mu) := \mathcal{L}^p(X, \mathcal{A}, \mu) \) denote the set of all \( \mu \)-measurable functions \( f : X \to \mathbb{K} \) such that \( N_p(f) < \infty \) (i.e. such that \( |f|^p \) is \( \mu \)-integrable (for \( p < \infty \))). For \( p \geq 1 \), it is common to call \( N_p \) the \( p \)-norm (caveat: on \( \mathcal{L}^p \), \( N_p \) is, in general, only a seminorm, see below) and to introduce the notation

\[
\forall_{p \in [1,\infty]} \forall_{f \in \mathcal{L}^p} \quad \|f\|_p := N_p(f).
\]

We also define

\[
\forall_{p \in [1,\infty]} \quad d_p : \mathcal{L}^p \times \mathcal{L}^p \to \mathbb{R}_0^+, \quad d_p(f, g) := \|f - g\|_p,
\]

\[
\forall_{p \in [0,1]} \quad d_p : \mathcal{L}^p \times \mathcal{L}^p \to \mathbb{R}_0^+, \quad d_p(f, g) := N_p^p(f - g).
\]

Clearly, each \( \mathcal{L}^p \) forms a vector space over \( \mathbb{K} \) (here we use that \( \mathcal{L}^p \) contains only \( \mathbb{K} \)-valued functions and not \( \hat{\mathbb{K}} \)-valued functions). Moreover, for \( p \in [1,\infty], N_p \) constitutes a seminorm on \( \mathcal{L}^p \) (i.e. \( N_p(0) = 0, N_p \) is homogeneous of degree 1, and \( N_p \) satisfies the triangle inequality due to the Minkowski inequality (2.50), cf. [Phi16b, Def. E.2]) and, in consequence, \( d_p \) forms a pseudometric on \( \mathcal{L}^p \) (i.e. \( d_p(f, f) = 0 \), \( d_p \) is symmetric, and \( d_p \) satisfies the triangle inequality, cf. [Phi16b, Def. E.1]). For \( p \in ]0,1[ \), \( d_p \) with the modified definition (2.53c) still forms a pseudometric on \( \mathcal{L}^p \): Indeed, \( d_p(f, f) = 0 \) and \( d_p(f, g) = d_p(g, f) \) are clear,

\[
\forall_{p \in [0,1]} \forall_{f,g \in \mathcal{L}^p} \quad d_p(f, g) = N_p^p(f - g) = N_p^p(f - h + h - g) \leq N_p^p(f - h) + N_p^p(h - g) = d_p(f, h) + d_p(h, g)
\]

is due to (2.51a). As described in [Phi16b, Sec. E], pseudometrics induce topologies in precisely the same way that metrics do (the open balls \( B_r(x) \) still form the base of a topology). However, if the pseudometric \( d \) is not a metric (resp., the seminorm is not a norm), i.e. if it is not positive definite and there are \( f \neq g \) such that \( d(f, g) = 0 \), then the induced topology is not Hausdorff (not even \( T_1 \), since there exists no open set separating \( f \) and \( g \)). In the present case, \( N_p \) and \( d_p \) are positive definite if, and only if, there does not exist a nonempty \( \mu \)-null set \( N \) (otherwise, \( f \equiv 0 \) and \( g := \chi_N \) satisfy \( f \neq g \), but \( d_p(f, g) = 0 \)). Due to this reason, one can not be completely happy with \( \mathcal{L}^p \), but needs to proceed to (b):
Let $\mathcal{N}$ denote the set of measurable $f : X \to \mathbb{K}$ that vanish $\mu$-almost everywhere. Clearly, $\mathcal{N}$ is a vector space over $\mathbb{K}$, namely a subspace of $L^p$. Thus, we can define the factor vector space (a.k.a. quotient vector space):

$$L^p := L^p_{\mathbb{K}} := L^p_{\mathbb{K}}(\mu) := L^p_{\mathbb{K}}(X, \mathcal{A}, \mu) := L^p_{\mathbb{K}}(X, \mathcal{A}, \mu)/\mathcal{N}.$$  

If $X = \mathbb{N}$, $\mathcal{A} = \mathcal{P}(\mathbb{N})$, and $\mu$ is counting measure, then it is customary to write $l^p := L^p$ (i.e., $l^p$ is the space of $p$-summable sequences; in this case $\mathcal{N} = \{0\}$ and $l^p = L^p = \ell^p$). Coming back to the general case, the elements of $L^p$ are precisely the equivalence classes of the equivalence relation defined by

$$f \sim g \iff f - g \in \mathcal{N} \iff d_p(f, g) = 0,$$

$L^p = \{[f] : f \in L^p\}$. If $f, g \in L^p$ represent the same element of $L^p$, i.e. $[f] = [g] \in L^p$, then $N_p(f) = N_p(g)$. Thus, it makes sense to define

$$\forall_{p \in [0, \infty]} \quad N_p : L^p \to \mathbb{R}_0^+,$$

$$\forall_{p \in [1, \infty]} \quad \| \cdot \| : L^p \to \mathbb{R}_0^+,$$

$$\forall_{p \in [0, \infty]} \quad d_p : L^p \times L^p \to \mathbb{R}_0^+,$$

$$N_p([f]) := N_p(f),$$

$$\| [f] \|_p := \| f \|_p,$$

$$d_p([f], [g]) := d_p(f, g).$$

It is then straightforward to verify that $\| \cdot \|_p$ constitutes a norm on $L^p$ for each $p \in [1, \infty]$, and $d_p$ constitutes a metric on $L^p$ for each $p \in [0, \infty]$, where the map $\iota : \ell^p \to L^p$, $\iota(f) := [f]$, is surjective and continuous (see [Phi16b, Th. E.8, Th. E.9]). In practise, it is very common not to properly distinguish between elements $[f] \in L^p$ and their representatives $f \in \ell^p$ (subsequently, we will also make use of this practise, where suitable). In most situations, using representatives of $[f]$ does not cause any problems. Some caution is, however, advisable: There are circumstances (e.g., restricting elements of $L^p$ or $\ell^p$ to $\mu$-null sets, typically to obtain traces on boundaries), where one does have to properly distinguish between $f$ and $[f]$.

An obvious question is if, for $p < q$, one of the spaces $\ell^p$ and $\ell^q$ contains the other? The answer is “no” in general (e.g., $f \equiv 1$ is in $\ell^\infty(\lambda^1)$, but not in $\ell^1(\lambda^1)$, $g$ with $g(t) := |t|^{-1/2} \cdot \chi_{[0,1]}(t)$ is in $\ell^1(\lambda^1)$, but not in $\ell^\infty(\lambda^1)$), but the answer is affirmative for spaces with finite measure. The latter result is a simple consequence of the Hölder inequality (2.49) and is provided as Th. 2.42. The complete answer is then given (with proof omitted) in Rem. 2.43 below.

**Theorem 2.42.** Let $(X, \mathcal{A}, \mu)$ be a finite measure space, i.e. $\mu(X) < \infty$. Then, for each $p, q \in [0, \infty]$ with $0 < p < q \leq \infty$, one has $\ell^q \subseteq \ell^p$ and

$$\forall_{f \in \ell^q} \quad N_p(f) \leq \mu(X)^{\frac{1}{p} - \frac{1}{q}} N_q(f).$$
Proof. It suffices to prove the estimate, which, clearly, implies $\mathcal{L}^q \subseteq \mathcal{L}^p$. Thus, let $f \in \mathcal{L}^q$. For $q = \infty$, we compute

$$N_p(f) = \left( \int_X |f|^p \, d\mu \right)^{\frac{1}{p}} \leq \left( \int_X \left( N_\infty(f) \right)^p \, d\mu \right)^{\frac{1}{p}} = N_\infty(f)^{\frac{1}{p}} \mu(X)^{\frac{1}{p}},$$

thereby establishing the case. For $q < \infty$, set $r := \frac{q}{p} > 1$ and $s := \frac{1}{r} = \frac{r - 1}{r} > 1$. Then $\frac{1}{r} + \frac{1}{s} = 1$. Since $|f|^p \in \mathcal{L}^r$ (since $|f|^{|r|} \in \mathcal{L}^1$) and $1 \in \mathcal{L}^s$, we obtain

$$N_p^r(f) = \int_X |f|^r \, 1 \, d\mu \leq N_r(|f|^p) N_s(1) = \left( \int_X |f|^p \, d\mu \right)^{\frac{1}{2}} \mu(X)^{\frac{1}{2}} = N_q^p(f)^{\frac{1}{r}} \mu(X)^{\frac{1}{r}} = N_q^p(f)^{\frac{1}{r}} \mu(X)^{\frac{1}{r}},$$

which, after taking $p$th roots, completes the proof. \hfill \blacksquare

Remark 2.43. Let $(X, \mathcal{A}, \mu)$ be a measure space. In generalization of Th. 2.42, one can show (cf. Problems 6.2.5 and 6.2.6 in [Els07]) the equivalence of the statements

(i) there exist $p, q \in [0, \infty]$ such that $0 < p < q \leq \infty$ and $\mathcal{L}^q \subseteq \mathcal{L}^p$,

(ii) $\sup \{ \mu(A) : A \in \mathcal{A}, \mu(A) < \infty \} < \infty$,

(iii) for each $p, q \in [0, \infty]$ such that $0 < p < q \leq \infty$, one has $\mathcal{L}^q \subseteq \mathcal{L}^p$;

as well as the equivalence of the statements

(i') there exist $p, q \in [0, \infty]$ such that $0 < p < q \leq \infty$ and $\mathcal{L}^p \subseteq \mathcal{L}^q$,

(ii') $\inf \{ \mu(A) : A \in \mathcal{A}, \mu(A) > 0 \} > 0$,

(iii') for each $p, q \in [0, \infty]$ such that $0 < p < q \leq \infty$, one has $\mathcal{L}^p \subseteq \mathcal{L}^q$.

In particular, a space with $\mu(X) = \infty$ can only have the properties (i) – (iii) if it is not $\sigma$-finite. As a concrete simple example, consider $X := \{0, 1\}$, $\mathcal{A} := \mathcal{P}(X)$, $\mu(\emptyset) := \mu(\{0\}) := 0$, $\mu(\{1\}) := \mu(X) := \infty$. Then every function $f : X \to \mathbb{K}$ is measurable, properties (ii) and (ii') are clearly satisfied and, for each $p \in [0, \infty]$, $\mathcal{L}^p = \{ (f : X \to \mathbb{K}) : f(1) = 0 \}$.

The following Th. 2.44(a), regarding the completeness of the $L^p$ spaces, is sometimes known as the Riesz-Fischer theorem. However, part of the literature reserves the name for a different theorem and, thus, we leave Th. 2.44(a) unnamed.

Theorem 2.44. Let $(X, \mathcal{A}, \mu)$ be a measure space and $p \in [0, \infty]$.

(a) All the spaces $L^p_\mathbb{K}(\mu)$ are complete (with respect to the pseudometric $d_p$), where the definition of Cauchy sequence and completeness in pseudometric spaces is precisely the same as in metric spaces. In particular, all $L^p_\mathbb{K}(\mu)$, $p \in [1, \infty]$, are Banach spaces; all $L^p_\mathbb{K}(\mu)$, $p \in [0, 1]$, are complete metric spaces.
(b) If \((f_k)_{k \in \mathbb{N}}\) is a sequence in \(L^p_x(\mu)\), converging to \(g \in L^p_x(\mu)\), then \((f_k)_{k \in \mathbb{N}}\) has a subsequence that converges to \(g\) pointwise \(\mu\)-almost everywhere. For \(p = \infty\), \((f_k)_{k \in \mathbb{N}}\) itself converges to \(g\), even uniformly \(\mu\)-almost everywhere.

Proof. (a): Let \((f_k)_{k \in \mathbb{N}}\) be a Cauchy sequence in \(L^p\). First, consider \(0 < p < \infty\). Then there exists a subsequence \((f_{k_l})_{l \in \mathbb{N}}\) such that
\[
\forall_{l \in \mathbb{N}} \forall_{m > k_l} d_p(f_m, f_{k_l}) < 2^{-l}.
\]
For each \(l \in \mathbb{N}\), set \(g_l := f_{k_l} - f_{k_{l+1}}\). Then
\[
\forall_{n \in \mathbb{N}} \forall_{p \geq 1} N_p \left( \sum_{l=1}^{n} |g_l| \right) \leq \sum_{l=1}^{n} N_p(g_l) = \sum_{l=1}^{n} d_p(f_{k_l}, f_{k_{l+1}}) < \sum_{l=1}^{n} 2^{-l} < 1,
\]
\[
\forall_{n \in \mathbb{N}} \forall_{p < 1} N_p \left( \sum_{l=1}^{n} |g_l| \right) \leq \sum_{l=1}^{n} N_p(g_l) = \sum_{l=1}^{n} d_p(f_{k_l}, f_{k_{l+1}}) < \sum_{l=1}^{n} 2^{-l} < 1.
\]
The monotone convergence Th. 2.7 allows to take the limit for \(n \to \infty\), resulting in \(N_p(\sum_{l=1}^{\infty} |g_l|) \leq 1\). In particular, for \(\mu\)-a.e. \(x \in X\), the series \(\sum_{l=1}^{\infty} g_l(x)\) converges absolutely and, in consequence, converges to some \(h(x) \in \mathbb{K}\). However,
\[
\forall_{x \in X} \forall_{n \in \mathbb{N}} \sum_{l=1}^{n} g_l(x) = f_{k_l}(x) - f_{k_{l+1}}(x),
\]
showing there exists a \(\mu\)-null set \(N\) such that \((f_{k_l})_{l \in \mathbb{N}}\) converges on \(N^c\) (i.e. \(\mu\)-a.e.) to the measurable function
\[
f : X \longrightarrow \mathbb{K}, \quad f(x) := \begin{cases} f_{k_l}(x) - h(x) & \text{for } x \in N^c, \\ 0 & \text{for } x \in N. \end{cases}
\]
We still need to verify \(f \in L^p\) and \(\lim_{k \to \infty} d_p(f_k, f) = 0\). To this end, let \(\epsilon > 0\). Then there exists \(k_0 \in \mathbb{N}\) such that \(N_p^p(f_k - f_m) < \epsilon\) for each \(k, m > k_0\). Fixing \(m > k_0\) and applying Fatou (Prop. 2.19) to the sequence \((|f_{k_l} - f_m|^p)_{l \in \mathbb{N}}\) yields
\[
N_p^p(f_m - f) = \int_X |f - f_m|^p \, d\mu = \int_X \lim_{l \to \infty} |f_{k_l} - f_m|^p \, d\mu \leq \liminf_{l \to \infty} \int_X |f_{k_l} - f_m|^p \, d\mu
\]
\[
= \liminf_{l \to \infty} N_p^p(f_{k_l} - f_m) \leq \epsilon.
\]
Since \(f = f - f_m + f_m\), this shows both \(f \in L^p\) and \(\lim_{k \to \infty} d_p(f_k, f) = 0\). It remains to consider the case \(p = \infty\). In this case, we have \(N \in \mathcal{A}\) with \(\mu(N) = 0\) for the set
\[
N := \bigcup_{(k,l) \in \mathbb{N}^2} \{|f_k - f_l| > N_\infty(f_k - f_l)\}.
\]
If \(x \in N^c\), then \((f_k(x))_{k \in \mathbb{N}}\) is a Cauchy sequence in \(\mathbb{K}\) (since \((f_k)_{k \in \mathbb{N}}\) is a Cauchy sequence in \(L^\infty\)). Thus, there exists \(f(x) \in \mathbb{K}\) such that \(f(x) = \lim_{k \to \infty} f_k(x)\). On \(N^c\), the
However, for no convergence $f_k \to f$ is even uniform: Given $\epsilon > 0$, choose $n_0 \in \mathbb{N}$ such that $N_\infty(f_k - f) < \frac{\epsilon}{2}$ for each $k, l > n_0$. Also choose $m > n_0$ such that $|f_m(x) - f(x)| < \frac{\epsilon}{2}$. Then, for each $k > n_0$ (and independently of $x \in N^c$), $|f_k(x) - f(x)| \leq |f_k(x) - f_m(x)| + |f_m(x) - f(x)| < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon$. Extending $f$ by $0$ to $X$, we obtain $f = \lim_{k \to \infty} f_k \chi_{N^c}$. In particular, $f$ is measurable as the pointwise limit of measurable functions. The uniform convergence $f_k \to f$ on $N^c$ also yields $\lim_{k \to \infty} N_\infty(f_k - f) = 0$ and $N_\infty(f) \leq N_\infty(f - f_k) + N_\infty(f_k) < \infty$, showing $f \in L^\infty$.

(b): If $(f_k)_{k \in \mathbb{N}}$ converges in $L^p$, then it is a Cauchy sequence in $L^p$. Thus, the proof of (a) shows the statement of (b) to be true if $g$ is replaced by the function $f \in L^p$ constructed in the proof of (a). Then $d_p(f, g) \leq d_p(f, f_k) + d_p(f_k, g) \to 0$ for $k \to \infty$, showing $d_p(f, g) = 0$, i.e. $f = g \mu$-a.e., thereby establishing the case.

In Th. 2.44(b), for $p < \infty$, one can, in general, not expect for the sequence $(f_k)_{k \in \mathbb{N}}$ to converge pointwise $\mu$-a.e. to some $f \in L^p$, as is demonstrated by the following example: Consider the measure space $([0, 1], \mathcal{B}, \beta^1)$ and the sequence of intervals
\[(I_k)_{k \in \mathbb{N}} := \left( [0, 1], \left[ 0, \frac{1}{2} \right], \left[ \frac{1}{2}, 1 \right], \left[ 0, \frac{1}{3} \right], \left[ \frac{1}{3}, \frac{2}{3} \right], \left[ \frac{2}{3}, 1 \right], \left[ 0, \frac{1}{4} \right], \ldots \right).\]
If, for each $k \in \mathbb{N}$, $f_k := \chi_{I_k}$, then, for each $p \in [0, \infty]$, $\lim_{k \to \infty} d_p(f_k, 0) = 0$: Indeed,
\[\lim_{k \to \infty} N_p(f_k) = \lim_{k \to \infty} \int_0^1 \chi_{I_k}^p \, d\beta^1 = \lim_{k \to \infty} \beta^1(I_k) = \lim_{k \to \infty} \frac{1}{k} = 0.\]
However, for no $x \in [0, 1]$ does $(f_k(x))_{k \in \mathbb{N}}$ converge.

### 2.6.2 Integration with Respect to Pushforward Measures

Given a measure space $(X, \mathcal{A}, \mu)$, a measurable space $(Y, \mathcal{B})$, and an $\mathcal{A}$-$\mathcal{B}$-measurable $\phi : X \to Y$, recall the pushforward measure
\[\mu_{\phi} : \mathcal{B} \to [0, \infty], \quad \mu_{\phi}(B) = \mu(\phi^{-1}(B)),\]
from Th. 1.69(a).

**Theorem 2.45. (a)** Let $(X, \mathcal{A}, \mu)$ be a measure space, $(Y, \mathcal{B})$ a measurable space, $\phi : X \to Y$ $\mathcal{A}$-$\mathcal{B}$-measurable, and $\mu_{\phi}$ the resulting pushforward measure on $(Y, \mathcal{B})$. Then, for each $f \in \mathcal{M}^+(\mathcal{B})$, one has
\[\int_Y f \, d\mu_{\phi} = \int_X (f \circ \phi) \, d\mu. \quad (2.54)\]
Moreover, a $\mathcal{B}$-measurable function $f : Y \to \hat{\mathbb{K}}$ is $\mu_{\phi}$-integrable if, and only if, $f \circ \phi$ is $\mu$-integrable, and then (2.54) holds.
(b) Let \( n \in \mathbb{N} \) and consider \((\mathbb{R}^n, \mathcal{A}, \mu^n)\), where \((\mathcal{A}, \mu^n)\) is either \((\mathcal{B}^n, \beta^n)\) or \((\mathcal{L}^n, \lambda^n)\). Moreover, let \( \phi : \mathbb{R}^n \to \mathbb{R}^n \) be bijective and affine. If \( f : \mathbb{R}^n \to \mathbb{K} \) is nonnegative \(\mu^n\)-measurable or \(\mu^n\)-integrable, then \( f \circ \phi \) has the same property and

\[
\int_{\mathbb{R}^n} f \, d\mu^n = |\det \phi| \int_{\mathbb{R}^n} (f \circ \phi) \, d\mu^n. \tag{2.55}
\]

Proof. (a): We note that \( f \in \mathcal{M}^+(\mathcal{B}) \) implies \( f \circ \phi \in \mathcal{M}^+(\mathcal{A}) \) and then use the technique summarized after the proof of Th. 2.36: First, let \( f = \chi_B, B \in \mathcal{B} \). Then

\[
\int_Y \chi_B \, d\mu_\phi = \mu_\phi(B) = \mu(\phi^{-1}(B)) = \int_X \chi_{\phi^{-1}(B)} \, d\mu = \int_X (\chi_B \circ \phi) \, d\mu,
\]

such that (2.54) holds. Due to the linearity of the integral, (2.54) then also holds for each \( f \in \mathcal{S}^+(\mathcal{B}) \). Now let \( f \in \mathcal{M}^+(\mathcal{B}) \) and choose a sequence \((u_k)_{k \in \mathbb{N}}\) in \(\mathcal{S}^+(\mathcal{B})\) such that \( u_k \uparrow f \). Then, clearly, \((u_k \circ \phi)_{k \in \mathbb{N}}\) is a sequence in \(\mathcal{S}^+(\mathcal{A})\) such that \((u_k \circ \phi) \uparrow (f \circ \phi)\), implying

\[
\int_Y f \, d\mu_\phi = \lim_{k \to \infty} \int_Y u_k \, d\mu_\phi = \lim_{k \to \infty} \int_X (u_k \circ \phi) \, d\mu = \int_X (f \circ \phi) \, d\mu,
\]

proving (2.54) for each \( f \in \mathcal{M}^+(\mathcal{B}) \). The remaining assertion of (a) follows by applying (2.54) to \((\text{Re} \, f)^+\) and \((\text{Im} \, f)^+\).

(b): According to Th. 1.73(a), \( \phi \) is \(\mu^n\)-\(\mu^n\)-measurable and \((\mu^n)_\phi = |\det \phi|^{-1} \mu^n\). Thus, if \( f : \mathbb{R}^n \to \mathbb{K} \) is \(\mu^n\)-measurable, then so is \( f \circ \phi \). Using (a), if \( f : \mathbb{R}^n \to \mathbb{K} \) is nonnegative \(\mu^n\)-measurable or \(\mu^n\)-integrable, then \( f \circ \phi \) has the same property and

\[
\int_{\mathbb{R}^n} (f \circ \phi) \, d\mu^n \overset{(2.54)}{=} \int_{\mathbb{R}^n} f \, d(\mu^n)_\phi = |\det \phi|^{-1} \int_{\mathbb{R}^n} f \, d\mu^n,
\]

completing the proof of (b).

\[\blacksquare\]

### 2.6.3 Dense Subsets, Separability

Below, we provide some dense subsets of \(L^p\)-spaces (for \( p < \infty \) – \(L^\infty\) turns out to have a less benign structure in most situations). Dense subsets are, e.g., useful, since continuous functions are already uniquely determined by values on dense subsets. Recall that a topological space is called *separable* if, and only if, it has a countable dense subset. Separable spaces are typically much easier to handle than nonseparable spaces. We will provide some sufficient conditions for \(L^p\)-spaces to be separable.

**Remark 2.46.** Let \((X, \mathcal{A}, \mu)\) be a measure space, \( p \in ]0, \infty] \), and let \( \mathcal{N} \) denote the set of measurable \( f : X \to \mathbb{K} \) that vanish \( \mu \)-almost everywhere. Let \( \iota : \mathcal{L}_p^0(\mu) \to L^p_\mu(\mu), \iota(f) := [f], \) be the quotient map. Consider \( \mathcal{F} \subseteq \mathcal{L}^p \) and \( F := \iota(\mathcal{F}) \subseteq \mathcal{L}^p \). Then \( \mathcal{F} \) is dense in \( \mathcal{L}^p \) if, and only if, \( F \) is dense in \( L^p \) (in particular, \( \mathcal{L}^p \) is separable if, and only if, \( L^p \) is separable): Note that \( d_p(f, g) = d_p(\iota(f), \iota(g)) \) for each \( f, g \in \mathcal{L}^p \). Let \( \mathcal{F} \) be dense in \( \mathcal{L}^p \). Given \( \iota(g) \in \mathcal{L}^p \) and \( \epsilon > 0 \), choose \( f \in \mathcal{F} \) such that \( d_p(f, g) < \epsilon \). Then \( \iota(f) \in F \)
and \( d_p(\epsilon(f), \epsilon(g)) < \epsilon \), showing \( F \) to be dense in \( L^p \). Conversely, let \( F \) be dense in \( L^p \). Given \( g \in L^p \) and \( \epsilon > 0 \), choose \( f \in F \) such that \( d_p(\epsilon(f), \epsilon(g)) < \epsilon \). Then \( d_p(f, g) < \epsilon \), showing \( F \) to be dense in \( L^p \).

**Theorem 2.47.** Let \((X, A, \mu)\) be a measure space, \( p \in ]0, \infty[ \).

(a) The vector space over \( \mathbb{K} \), defined by

\[
S_0 := S_0(A, \mu) := \text{span}\{\chi_A : A \in A, \mu(A) < \infty\},
\]

i.e. the vector space of simple functions that are 0 outside a set of finite measure, is dense in \( L^p \): More precisely,

\[
\forall f \in L^p \quad \forall \epsilon \in \mathbb{R}^+ \quad \exists \phi \in S_0 \quad \left( |\phi| \leq |f| \land d_p(f, \phi) < \epsilon \right).
\]

(b) If \( S \subseteq A \) is a semiring on \( X \) such that \( \sigma(S) = A \) and \( \mu |_S \) is \( \sigma \)-finite, then the vector space over \( \mathbb{K} \), defined by

\[
D := \text{span}\{\chi_A : A \in S, \mu(A) < \infty\},
\]

is dense in \( L^p \).

(c) If there exists a countable semiring \( S \) on \( X \) such that \( A = \sigma(S) \) and \( \mu |_S \) is \( \sigma \)-finite, then \( L^p \) is separable.

(d) If there exists a countable semiring \( S \) on \( X \) such that \( E = \sigma(S) \), \( E \subseteq A \subseteq \tilde{E} \), where \((X, \tilde{E}, \tilde{\alpha})\) is the completion of \((X, E, \alpha)\), \( \alpha := \mu |_E \), \( \mu = \tilde{\alpha} |_A \), and \( \mu |_S \) is \( \sigma \)-finite, then \( L^p(\mu) \) is separable.

(e) Let \( n \in \mathbb{N} \) and consider \((\mathbb{R}^n, B, \mu^n)\), where \((B, \mu^n) = (B^n, \beta^n) \) or \((B, \mu^n) = (L^n, \lambda^n)\). If \( B \in B \) and \((X, A, \mu) = (B, B|B, \mu^n |_A)\), then \( L^p(\mu) \) is separable.

**Proof.** (a): Clearly, \( S_0(A, \mu) \subseteq L^p \). Let \( f \in L^p \). First, we also assume \( f \geq 0 \). Then there exists a sequence \((\phi_k)_{k \in \mathbb{N}}\) in \( S^+(A) \) such that \( \phi_k \uparrow f \). Then \( N_p(\phi_k) \leq N_p(f) < \infty \), showing \( \phi_k \in S_0 \). Since \( |f - \phi_k| \leq |f| \), we can apply the dominated convergence Th. 2.20 to obtain,

\[
\lim_{k \to \infty} d_p(f, \phi_k) = \lim_{k \to \infty} N_p(f - \phi_k) = \lim_{k \to \infty} \int_X |f - \phi_k|^p \, d\mu \stackrel{\text{Th. 2.20}}{=} 0,
\]

proving the assertion for \( f \geq 0 \). For a general \( f \in L^p \), we proceed in the usual way, letting \( f_1 := (\text{Re} f)^+ \), \( f_2 := (\text{Re} f)^- \), \( f_3 := (\text{Im} f)^+ \), \( f_4 := (\text{Im} f)^- \). Then each \( f_k \in L^p \) and \( f_k \geq 0 \), \( k = 1, \ldots, 4 \). Given \( \epsilon > 0 \), choose \( \phi_k \in S_0 \) such that \( \phi_k \leq f_k \) and \( d_p(f_k, \phi_k) < \frac{\epsilon}{4} \) and let \( \phi := \phi_1 - \phi_2 + i(\phi_3 - \phi_4) \in S_0 \). Then

\[
|\phi|^2 = (\phi_1 + \phi_2)^2 + (\phi_3 + \phi_4)^2 \leq (f_1 + f_2)^2 + (f_3 + f_4)^2 = |f|^2
\]
2 INTEGRATION

and, for \( p \geq 1 \),

\[
N_p(f - \phi) = N_p(f_1 - \phi_1 - (f_2 - \phi_2) + i((f_3 - \phi_3) - (f_4 - \phi_4))) \leq \sum_{k=1}^{4} N_p(f_k - \phi_k) < \epsilon.
\]

For \( p < 1 \), one replaces \( N_p \) by \( N_p^p \) in the above estimate, thereby completing the proof of (a).

(b): According to (a), it suffices to show that, given \( E \in A \) with \( \mu(E) < \infty \) and \( \epsilon > 0 \), there exist disjoint sets \( A_1, \ldots, A_N \in S \), \( N \in \mathbb{N} \), such that \( \mu(A_k) < \infty \) and

\[
N_p^p \left( \chi_E - \sum_{k=1}^{N} \chi_{A_k} \right) = \int_X \left| \chi_E - \chi_{\bigcup_{k=1}^{N} A_k} \right|^p d\mu
\]

\[
= \mu \left( E \setminus \bigcup_{k=1}^{N} A_k \right) + \mu \left( \left( \bigcup_{k=1}^{N} A_k \right) \setminus E \right) < \epsilon.
\]

To obtain the \( A_k \), we use the Carathéodory extension Th. 1.38 together with the uniqueness result of Cor. 1.46(a) (which holds, since \( \mu|_S \) is \( \sigma \)-finite) to see that \( \mu \) on \( A = \sigma(S) \) must be the restriction of the outer measure on \( P(X) \) induced by \( \mu \) on \( S \). Thus, if \( \mathcal{R} := \rho(S) \) is the generated ring, then

\[
\mu(E) \overset{(1.30)}{=} \inf \left\{ \sum_{k=1}^{\infty} \mu(A_k) : (A_k)_{k \in \mathbb{N}} \text{ sequence in } S, E \subseteq \bigcup_{k=1}^{\infty} A_k \right\}
\]

\[
\overset{\text{Prop. 1.18(b)}}{=} \inf \left\{ \sum_{k=1}^{\infty} \mu(A_k) : (A_k)_{k \in \mathbb{N}} \text{ sequence in } \mathcal{R}, E \subseteq \bigcup_{k=1}^{\infty} A_k \right\}
\]

\[
= \inf \left\{ \sum_{k=1}^{\infty} \mu(A_k) : (A_k)_{k \in \mathbb{N}} \text{ disjoint sequence in } \mathcal{R}, E \subseteq \bigcup_{k=1}^{\infty} A_k \right\}
\]

\[
\overset{\text{Prop. 1.18(a)}}{=} \inf \left\{ \sum_{k=1}^{\infty} \mu(A_k) : (A_k)_{k \in \mathbb{N}} \text{ disjoint sequence in } S, E \subseteq \bigcup_{k=1}^{\infty} A_k \right\}.
\]

In consequence, since \( \mu(E) < \infty \), there exists a sequence of disjoint sets \( (A_k)_{k \in \mathbb{N}} \in S \), \( E \subseteq \bigcup_{k=1}^{\infty} A_k \) and \( \mu \left( \bigcup_{k=1}^{\infty} A_k \right) < \mu(E) + \frac{\epsilon}{2} \). Then

\[
\lim_{N \to \infty} \mu \left( E \setminus \bigcup_{k=1}^{N} A_k \right) = 0 \quad \text{and} \quad \lim_{N \to \infty} \mu \left( \left( \bigcup_{k=1}^{N} A_k \right) \setminus E \right) \leq \mu \left( \bigcup_{k=1}^{\infty} A_k \right) - \mu(E) < \frac{\epsilon}{2}.
\]

Thus, there exists \( N \in \mathbb{N} \) such that

\[
\mu \left( E \setminus \bigcup_{k=1}^{N} A_k \right) + \mu \left( \left( \bigcup_{k=1}^{N} A_k \right) \setminus E \right) < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon,
\]

and the proof of (b) is complete.
Let \( \text{Th. 2.49.} \): According to (b), \( D \subset \mathbb{R} \) is dense in \( Q \). The set span \( L \) is dense in \( L \). If \( f \in L \), then, in general, the space \( \mathbb{R} \) is typically not separable (cf. Rem. 2.55 below).

(d): According to (c), \( L^p(\alpha) \) is separable. Let \( D \) be a countable dense subset of \( L^p(\alpha) \). If \( f \in L^p(\mu) \), then, by Prop. G.3, there exists \( g \in L^p(\alpha) \) such that \( f = g \) \( \alpha \)-a.e. Thus, given \( \epsilon > 0 \), there exists \( \phi \in D \), satisfying \( d_\mu(f, \phi) = d_\mu(g, \phi) < \epsilon \), showing \( D \) to be dense in \( L^p(\mu) \).

(e): First, consider the case \( (B, \mu^n) = (B^n, \beta^n) \). We know that \( B^n \) is generated by the countable semiring \( T_\mathbb{Q}^n \): \( B^n = \sigma(T_\mathbb{Q}^n) \). If \( B \subseteq \mathbb{R}^n \), then \( S := T_\mathbb{Q}^n \cap B \) is a countable semiring on \( B \) by Prop. G.2. For \( B \subseteq B^n \), according to Th. 1.56(b), \( A = B^n \cap B = \sigma_B(S) \) and, for \( \mu = \beta^n \upharpoonright A \), \( L^p(\mu) \) is separable by (c). The case \( (B, \mu^n) = (\mathcal{L}^n, \lambda^n) \) now follows from (d).

It turns out that, in many respects, \( L^\infty \) is typically less benign than \( L^p \) for \( p < \infty \). For example, in general, the space \( S_0 \) of Th. 2.47(a) is not dense in \( L^\infty \) and, in the situations of Th. 2.47(c),(d),(e), \( L^\infty \) is typically not separable (cf. Rem. 2.55 below).

\textbf{Definition 2.48.} (a) Let \( (X, \mathcal{T}) \) be a topological space. Given a function \( f : X \rightarrow \mathbb{K} \), we call the set
\[
\text{supp } f := \{ x \in X : f(x) \neq 0 \}
\]
the \textit{support} of \( f \). Continuous functions whose support is a \textit{compact} subset of \( X \) are of particular importance, giving rise to the notation
\[
C_c(X) := C_c(X, \mathbb{K}) := \{ f \in C(X, \mathbb{K}) : \text{supp } f \text{ is compact} \}.
\]

If \( O \subseteq \mathbb{R}^n \) is open, \( n \in \mathbb{N} \), then we also define
\[
\forall k \in \mathbb{N}_0 \cup \{ \infty \} \quad C^k_c(O) := C^k_c(O, \mathbb{K}) := C_c(O, \mathbb{K}) \cap C^k(O, \mathbb{K}).
\]

(b) If \( O \subseteq \mathbb{R}^n \) is open, \( n \in \mathbb{N} \), then a function \( f : O \rightarrow \mathbb{K} \) is said to \textit{vanish at infinity} if, and only if, for each \( \epsilon \in \mathbb{R}^+ \), there exists a compact set \( K \subseteq O \) such that \( |f(x)| < \epsilon \) for each \( x \in O \setminus K \) (for \( O = \mathbb{R}^n \), this is, clearly, equivalent to
\[
\lim_{x \rightarrow \infty} f(x) = 0, \quad (2.56)
\]
where (2.56) is defined to mean that, for each sequence \( (x_k)_{k \in \mathbb{N}} \) in \( \mathbb{R}^n \) such that \( \lim_{k \rightarrow \infty} \|x_k\| = \infty \), one has \( \lim_{k \rightarrow \infty} f(x_k) = 0 \). Define
\[
C_0(O) := C_0(O, \mathbb{K}) := \{ f \in C(O, \mathbb{K}) : f \text{ vanishes at infinity} \}.
\]
Clearly, \( C_c(O) \subset C_0(O) \) (if \( O = \mathbb{R}^n \), then \( f(x) := e^{-\|x\|^2} \) defines \( f \in C_0(O) \setminus C_c(O) \);
if \( O \neq \mathbb{R}^n \), then \( f(x) := \text{dist}(x, O^c) \cdot e^{-\|x\|^2} \) defines \( f \in C_0(O) \setminus C_c(O) \)).

\textbf{Theorem 2.49.} Let \( n \in \mathbb{N} \) and let \( O \subseteq \mathbb{R}^n \) be open and consider \( (O, \mathcal{A}, \mu^n) \), where \( (\mathcal{A}, \mu^n) \) is either \( (B^n, \beta^n) \) or \( (L^n, \lambda^n) \).
For each \( p \in ]0, \infty[ \), \( C_c(O, \mathbb{K}) \) is dense in \( \mathcal{L}_p^p(\mu^n) \).

(b) For each \( p \in ]0, \infty[ \) and \( f \in \mathcal{L}_p^p(\mathbb{R}^n, \mathcal{A}, \mu^n) \), one has

\[
\lim_{t \to 0} \int_{\mathbb{R}^n} |f(x + t) - f(x)|^p \, d\mu^n(x) = 0. \tag{2.57}
\]

**Proof.** (a): It suffices to consider \((\mathcal{A}, \mu^n) = (\mathcal{L}'^n, \lambda^n)\). Due to Th. 2.47(a), it suffices to show that, for each \( A \in \mathcal{A} \) with \( \lambda^n(A) < \infty \) and each \( \epsilon > 0 \), there exists \( g \in C_c(O) \) such that \( N_p^p(\lambda^A - g) < \epsilon \). Let \( \delta := \min\{\frac{\epsilon}{2}, \frac{\epsilon}{2}\} \). Choose \( k \in \mathbb{N} \) sufficiently large such that, for \( B := A \cap [-k, k]^n \), one has \( \lambda^n(A) \leq \lambda^n(B) + \delta \). Then \( N_p^p(\lambda^A - \lambda^B) < \delta \). According to (1.41), there exists a closed set \( K \subseteq \mathbb{R}^n \) and an open set \( V \subseteq \mathbb{R}^n \) such that \( K \subseteq B \subseteq V \) and \( \lambda^n(V \setminus K) < \delta \). If we let \( U := V \cap O \cap [-k-1, k+1]^n \), then \( U \) is still open with \( B \subseteq U \) and \( \lambda^n(U \setminus K) < \delta \), but \( U \) is also bounded and \( U \subseteq O \). Moreover, since \( K \) is a subset of the bounded set \( B \), \( K \) is compact. If \( K = \emptyset \), then \( \lambda^n(B) < \delta \) and \( \lambda^n(A) < \epsilon \), such that we can choose \( g \equiv 0 \). Thus, consider \( K \neq \emptyset \). Then, using the continuity of \( \text{dist} \) (cf. [Phi16b, Ex. 2.6(b)]), we observe \( \text{dist}(K, U^c) > 0 \) and define the continuous function

\[
g : O \to \mathbb{R}, \quad g(x) := 1 - \min\left\{1, \frac{\text{dist}(x, K)}{\text{dist}(K, U^c)}\right\}. \tag{2.58}
\]

Then, clearly, \( 0 \leq g \leq 1 \), \( g \mid K = 1 \), and \( g \mid U^c = 0 \). The support of \( g \) is a closed subset of the compact set \( \overline{U} \), i.e. supp \( g \) is compact and \( g \in C_c(O) \). Moreover, noting

\[
\begin{align*}
|\chi_B(x) - g(x)| &= 1 - 1 = 1 - 1 = |\chi_U(x) - \chi_K(x)| \quad \text{for} \ x \in K, \\
|\chi_B(x) - g(x)| &= 1 - g(x) \leq 1 - 0 = |\chi_U(x) - \chi_K(x)| \quad \text{for} \ x \in B \setminus K, \\
|\chi_B(x) - g(x)| &= 0 - g(x) \leq 1 - 0 = |\chi_U(x) - \chi_K(x)| \quad \text{for} \ x \in U \setminus B, \\
|\chi_B(x) - g(x)| &= 0 - 0 = 0 - 0 = |\chi_U(x) - \chi_K(x)| \quad \text{for} \ x \in U^c,
\end{align*}
\]

we obtain

\[
N_p^p(\chi_B - g) = \int_O |\chi_B - g|^p \, d\lambda^n \leq \int_O |\chi_U - \chi_K|^p \, d\lambda^n < \delta.
\]

Then, for \( p < 1 \), \( N_p^p(\chi_A - g) \leq N_p^p(\chi_A - \chi_B) + N_p^p(\chi_B - g) < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon \). For \( p \geq 1 \), \( N_p(\chi_A - g) \leq N_p(\chi_A - \chi_B) + N_p(\chi_B - g) < 2\delta^{1/p} \leq 2\frac{\epsilon}{2} = \epsilon \), completing the proof of (a).

(b): Seeking a contradiction, assume there exists \( \epsilon > 0 \) such that for each \( s > 0 \), there exists \( s > t > 0 \) with \( \int_{\mathbb{R}^n} |f(x + t) - f(x)|^p \, d\mu^n(x) \geq \epsilon \). Using (a), choose \( g \in C_c(\mathbb{R}^n, \mathbb{K}) \) such that \( \int_{\mathbb{R}^n} |f(x + t) - g(x)|^p \, d\mu^n(x) < \frac{\epsilon}{3} \). Since \( g \) is continuous and thus, bounded on the compact set supp \( g \), there exists \( M \in \mathbb{R}^+ \) such that \( 2^p |g| \leq M \). Choosing \( r > 0 \) such that supp \( g \subseteq B_r(0) \) and letting \( B := B_r(0), M \chi_B \) is integrable and dominates \( x \mapsto |g(x+t) - g(x)|^p \) for all sufficiently small \( t \geq 0 \). Thus, by the dominated convergence Th. 2.20,

\[
\lim_{t \to 0} \int_{\mathbb{R}^n} |g(x + t) - g(x)|^p \, d\mu^n(x) \overset{\text{Th. 2.20}}{=} \lim_{t \to 0} \int_{\mathbb{R}^n} |g(x + t) - g(x)|^p \, d\mu^n(x) \overset{\text{g cont.}}{=} 0.
\]

In consequence,

\[
\exists \ s > 0 \quad \forall \ 0 \leq t < s \quad \int_{\mathbb{R}^n} |g(x + t) - g(x)|^p \, d\mu^n(x) < \frac{\epsilon}{3}
\]
and, for each $0 \leq t < s$,
\[
\int_{\mathbb{R}^n} |f(x + t) - f(x)|^p \, d\mu^n(x)
\leq \int_{\mathbb{R}^n} |f(x + t) - g(x + t)|^p \, d\mu^n(x) + \int_{\mathbb{R}^n} |g(x + t) - g(x)|^p \, d\mu^n(x)
+ \int_{\mathbb{R}^n} |g(x) - f(x)|^p \, d\mu^n(x) \overset{(2.55)}{=} \frac{\epsilon}{3} + \frac{\epsilon}{3} + \frac{\epsilon}{3} = \epsilon.
\]
This contradiction proves (2.57). \hfill \blacksquare

We will see in Th. 2.54(d) below that even $C_c^{\infty}(O, \mathbb{K})$ is dense in $L_{\mathbb{K}}^p(\mu^n)$ for $p \in [0, \infty[$. One can actually extend Th. 2.49(a) to measure spaces $(X, \mathcal{A}, \mu)$, where $\mu$ is a suitable measure on a locally compact Hausdorff space (see, e.g., [Rud87, Th. 3.14]).

### 2.6.4 Convolution and Fourier Transform

We will now define the so-called convolution of two integrable functions on $\mathbb{R}^n$.

**Remark and Definition 2.50.** Let $n \in \mathbb{N}$ and consider $(\mathbb{R}^n, \mathcal{A}, \mu^n)$, where $(\mathcal{A}, \mu^n)$ is either $(\mathcal{B}^n, \beta^n)$ or $(\mathcal{L}^n, \lambda^n)$. Consider $f, g \in L_{\mathbb{K}}^1(\mu^n)$. Then
\[
h : \mathbb{R}^{2n} \rightarrow \mathbb{K}, \ h(x, y) := f(x - y)g(y),
\]
is $\mu^{2n}$-measurable: Indeed, the function $A : \mathbb{R}^{2n} \rightarrow \mathbb{R}^{2n}, A(x, y) := (x - y, y)$, is linear and bijective and, thus, $\mu^{2n}$-$\mu^{2n}$-measurable. The projections $\pi_1, \pi_2 : \mathbb{R}^{2n} \rightarrow \mathbb{R}^n$, $\pi_1(x, y) = x, \pi_2(x, y) = y$ are $\mu^{2n}$-$\mu^n$-measurable. Thus, as $f, g$ are $\mu^n$-measurable, and $h = (f \circ \pi_1 \circ A) \cdot (g \circ \pi_2)$, $h$ is $\mu^{2n}$-measurable. Next, we compute
\[
\int_{\mathbb{R}^n} \int_{\mathbb{R}^n} |f(x - y)g(y)| \, d\mu^n(x) \, d\mu^n(y) = \int_{\mathbb{R}^n} |g(y)| \int_{\mathbb{R}^n} |f(x - y)| \, d\mu^n(x) \, d\mu^n(y)
\overset{(2.55)}{=} \|f\| \int_{\mathbb{R}^n} |g(y)| \, d\mu^n(y) = \|f\| \|g\| < \infty.
\]
Since $h$ is measurable, we may apply Th. 2.36(c) to conclude that $h$ is $\mu^{2n}$-integrable and there exists a $\mu^n$-null set $N \subseteq \mathbb{R}^n$ such that $y \mapsto h(x, y)$ is $\mu^n$-integrable for each $x \in N^c$. We now define
\[
f \ast g : \mathbb{R}^n \rightarrow \mathbb{K}, \quad (f \ast g)(x) := \begin{cases} \int_{\mathbb{R}^n} f(x - y)g(y) \, d\mu^n(y) & \text{for } x \in N^c, \\ 0 & \text{for } x \in N, \end{cases}
\]
and call this function the convolution of $f$ and $g$. From our considerations above, we then know $f \ast g \in L_{\mathbb{K}}^1(\mu^n)$. If $\tilde{f}, \tilde{g} \in L_{\mathbb{K}}^1(\mu^n)$ with $f = \tilde{f}$ $\mu^n$-a.e. and $g = \tilde{g}$ $\mu^n$-a.e., then, for each $x \in \mathbb{R}^n$, we have $f(x - y)g(y) = \tilde{f}(x - y)\tilde{g}(y)$ for a.e. $y \in \mathbb{R}^n$. Thus, $y \mapsto \tilde{f}(x - y)\tilde{g}(y)$ is $\mu^n$-integrable for each $x \in N^c$, and $f \ast g = \tilde{f} \ast \tilde{g}$. In particular, convolution “$\ast$” is well-defined as a map from $L_{\mathbb{K}}^1(\mu^n) \times L_{\mathbb{K}}^1(\mu^n)$ to $L_{\mathbb{K}}^1(\mu^n)$.
Proposition 2.51. Let $n \in \mathbb{N}$ and consider $(\mathbb{R}^n, \mathcal{A}, \mu^n)$, where $(\mathcal{A}, \mu^n)$ is either $(\mathcal{B}^n, \beta^n)$ or $(\mathcal{L}^n, \lambda^n)$. Let

$$
*: \mathcal{L}^1_k(\mu^n) \times \mathcal{L}^1_k(\mu^n) \rightarrow \mathcal{L}^1_k(\mu^n)
$$

be the convolution map as defined in Rem. and Def. 2.50 above.

(a) Convolution is commutative:

$$
\forall_{f,g \in \mathcal{L}^1_k(\mu^n)} f \ast g = g \ast f.
$$

(b) Convolution is distributive:

$$
\forall_{f,g,h \in \mathcal{L}^1_k(\mu^n)} (f + g) \ast h = f \ast h + g \ast h \quad \mu^n\text{-a.e.}
$$

(c) Convolution is associative:

$$
\forall_{f,g,h \in \mathcal{L}^1_k(\mu^n)} (f \ast g) \ast h = f \ast (g \ast h) \quad \mu^n\text{-a.e.}
$$

(d) $\text{supp}(f \ast g) \subseteq \text{supp } f + \text{supp } g$.

Proof. (a): Since, for each $x \in \mathbb{R}^n$, the map $\phi_x : \mathbb{R}^n \rightarrow \mathbb{R}^n$, $\phi_x(y) := x - y$, is affine and bijective with determinant $|\det \phi_x| = 1$, (2.55) implies, for each $x \in \mathbb{R}^n$, $y \mapsto f(x - y)g(y)$ to be $\mu^n$-integrable if, and only if, $y \mapsto f(y)g(x - y)$ is $\mu^n$-integrable and, in that case,

$$(f \ast g)(x) = \int_{\mathbb{R}^n} f(x - y)g(y) \, d\mu^n(y) = \int_{\mathbb{R}^n} f(y)g(x - y) \, d\mu^n(y) = (g \ast f)(x).$$

(b) is due to the fact that, for each $x, y \in \mathbb{R}^n$,

$$(f + g)(x - y)h(y) = (f(x - y) + g(x - y))h(y) = f(x - y)h(y) + g(x - y)h(y),$$

which is applied for each $x \in \mathbb{R}^n$ such that $y \mapsto (f + g)(x - y)h(y)$ is $\mu^n$-integrable.

(c): We apply Fubini for each $x \in \mathbb{R}^n$ such that $y \mapsto (f \ast g)(x - y)h(y)$ is $\mu^n$-integrable to obtain

$$
((f \ast g) \ast h)(x) = \int_{\mathbb{R}^n} (f \ast g)(x - y)h(y) \, d\mu^n(y) \stackrel{(2.55)}{=} \int_{\mathbb{R}^n} (f \ast g)(y)h(x - y) \, d\mu^n(y) \stackrel{(a)}{=} \int_{\mathbb{R}^n} \left( \int_{\mathbb{R}^n} f(z)g(y - z) \, d\mu^n(z) \right) h(x - y) \, d\mu^n(y) \stackrel{\text{Fubini}}{=} \int_{\mathbb{R}^n} f(z) \int_{\mathbb{R}^n} g(y - z)h(x - y) \, d\mu^n(y) \, d\mu^n(z) \stackrel{(2.55)}{=} \int_{\mathbb{R}^n} f(z)(g \ast h)(x - z) \, d\mu^n(z) \stackrel{(2.55)}{=} \left( f \ast (g \ast h) \right)(x),
$$
proving associativity.

(d): If \((f * g)(x) \neq 0\), then \(\int_{\mathbb{R}^n} f(x - y)g(y) \, d\mu^n(y) \neq 0\), i.e. there exists \(y \in \text{supp } g\) such that \(x - y \in \text{supp } f\), showing \(x = y + (x - y) \in \text{supp } f + \text{supp } g\) and proving (d). ■

**Definition 2.52.** Let \(n \in \mathbb{N}\). We call a sequence \((\varphi_k)_{k \in \mathbb{N}}\) in \(L^1(\beta^n)\) a Dirac sequence if, and only if, it has the following properties (i) – (iii):

(i) \(\varphi_k \geq 0\) for each \(k \in \mathbb{N}\),

(ii) \(\|\varphi_k\|_1 = 1\) for each \(k \in \mathbb{N}\),

(iii) \(\lim_{k \to \infty} \text{diam } \{0\} \cup \text{supp } \varphi_k = 0\).

A Dirac sequence can be seen as an approximation of the Dirac measure \(\delta_0 : \mathcal{P}(\mathbb{R}^n) \to [0, \infty]\) (cf. Ex. 1.12(c)) which, in turn, constitutes a representation of the so-called Dirac distribution (concentrated at 0).

**Example 2.53.** Let \(n \in \mathbb{N}\).

(a) Clearly,

\[
\varphi_k : \mathbb{R}^n \to \mathbb{K}, \quad \varphi_k(x) := \begin{cases} (2/k)^{-n} & \text{for } x \in [-\frac{1}{k}, \frac{1}{k}]^n, \\ 0 & \text{otherwise}, \end{cases}
\]

defines a (discontinuous) Dirac sequence.

(b) The sequence defined by

\[
\varphi_k : \mathbb{R}^n \to \mathbb{K}, \quad \varphi_k(x) := \prod_{j=1}^{n} \max\{0, k - k^2|x_j|\},
\]

defines a continuous (but not differentiable) Dirac sequence: Clearly, \(\varphi_k \geq 0\), \(\text{supp } \varphi_k = [-\frac{1}{k}, \frac{1}{k}]^n\), and, using Fubini,

\[
\|\varphi_k\|_1 = \int_{\mathbb{R}^n} \varphi_k = \prod_{j=1}^{n} \left(2 \int_0^{1/k} (k - k^2 x_j) \, dx_j\right) = \prod_{j=1}^{n} \left(2 \left[kx_j - \frac{k^2 x_j^2}{2}\right]_0^{1/k}\right) = 1.
\]

(c) We show the sequence defined by

\[
\varphi_k : \mathbb{R}^n \to \mathbb{K}, \quad \varphi_k(x) := \begin{cases} c_k \exp\left(-\frac{1}{k^2 - \|x\|^2}\right) & \text{for } \|x\| < \frac{1}{k}, \\ 0 & \text{otherwise}, \end{cases}
\]

where

\[
c_k := \left(\int_{B} \exp\left(-\frac{1}{k^2 - \|x\|^2}\right) \, dx\right)^{-1}, \quad B := \{x \in \mathbb{R}^n : \|x\| < 1/k\},
\]
2 INTEGRATION

defines a Dirac sequence in \( C^\infty(\mathbb{R}^n) \): Clearly, each \( \varphi_k \) maps into \( \mathbb{R}_0^+ \), \( \| \varphi_k \|_1 = 1 \) by the definition of \( c_k \), and \( \text{supp} \varphi_k = B_{k^{-1}}(0) \), showing \((\varphi_k)_{k \in \mathbb{N}}\) to be a Dirac sequence. It remains to verify that each \( \varphi_k \) is \( C^\infty \). To this end, we show that, for each \( p = (p_1, \ldots, p_m) \in \{1, \ldots, n\}^m, m \in \mathbb{N}, \partial_p \varphi_k \) exists. Let \( x \in \mathbb{R}^m \). For \( \|x\|_2 > \frac{1}{k} \), \( \partial_p \varphi_k(x) = 0 \) is immediate from the definition of \( \varphi_k \). We claim that, for \( \|x\|_2 < \frac{1}{k} \),

\[
\partial_p \varphi_k(x) = c_k \left( \sum_{j=1}^{N_p} l_{p,j} \frac{\partial_j \nu_{p,j}(x)}{\alpha(k,x)^2} \right) \exp \left( -\frac{1}{\alpha(k,x)} \right) ,
\]

where \( \alpha(k,x) := (k^{-2} - \|x\|_2^2), N_p \in \mathbb{N}, \) each \( l_{p,j} \in \mathbb{Z}, \) and each \( \nu_{p,j} \) is a monomial in \( x_1, \ldots, x_n \): To prove (2.60) by induction, fix \( p = (p_1, \ldots, p_m) \in \{1, \ldots, n\}^m, m \in \mathbb{N}, \) assume (2.60) to hold for this \( p, \) and let \( \nu \in \{1, \ldots, n\} \). Then we have to differentiate \( \partial_p \varphi_k \) with respect to \( x_\nu \) using the product rule, resulting in

\[
\partial_\nu \partial_p \varphi_k(x) = c_k \left( \sum_{j=1}^{N_p} \frac{(-j)(-2x_\nu) l_{p,j} \nu_{p,j}(x)}{\alpha(k,x)^3} \right) \exp \left( -\frac{1}{\alpha(k,x)} \right) + \left( \sum_{j=1}^{N_p} \frac{l_{p,j} \nu_{p,j}(x)}{\alpha(k,x)^2} \right) \left( -\frac{2x_\nu}{\alpha(k,x)} \right) \exp \left( -\frac{1}{\alpha(k,x)} \right) + \left( \sum_{j=1}^{N_q} \frac{l_{q,j} \nu_{q,j}(x)}{\alpha(k,x)^2} \right) \exp \left( -\frac{1}{\alpha(k,x)} \right) ,
\]

where \( q := (\nu, p_1, \ldots, p_m), N_q \in \mathbb{N}, \) each \( l_{q,j} \in \mathbb{Z}, \) and each \( \nu_{q,j} \) is a monomial in \( x_1, \ldots, x_n \), completing the induction. Finally, suppose \( \|x\|_2 = k^{-1} \). We prove via induction that, for each \( p = (p_1, \ldots, p_m) \in \{1, \ldots, n\}^m, m \in \mathbb{N}, \partial_p \varphi_k(x) = 0 \): Once again, we assume the claim for a fixed \( p \) and let \( \nu \in \{1, \ldots, n\} \). Let \( (t_\gamma)_{\gamma \in \mathbb{N}} \) be a sequence in \( \mathbb{R} \setminus \{0\} \) such that, with \( x_\gamma := x + t_\gamma e_\nu, (x_\gamma)_{\gamma \in \mathbb{N}} \) is a sequence in \( B_{k^{-1}}(0) \), satisfying \( \lim_{\gamma \to \infty} x_\gamma = x \). Then, for each \( \gamma \in \mathbb{N}, \)

\[
\alpha(k,x_\gamma) = k^{-2} - \|x_\gamma\|_2^2 - \|x_\gamma\|_2^2 = x_\nu^2 - (x_\nu + t_\gamma)^2 = -2x_\nu t_\gamma + t_\gamma^2
\]

and

\[
\lim_{\gamma \to \infty} \alpha(k,x_\gamma) = 0 \quad \Rightarrow \quad \lim_{\gamma \to \infty} \exp \left( -\frac{1}{\alpha(k,x_\gamma)} \right) = 0
\]

\[
\quad \Rightarrow \forall j \in \mathbb{N}, \lim_{\gamma \to \infty} \left( \alpha(k,x_\gamma) \right)^{-j} \exp \left( -\frac{1}{\alpha(k,x_\gamma)} \right) = 0.
\]

This, together with (2.60), \( t_\gamma^{-1} = (-2x_\nu - t_\gamma) (\alpha(k,x_\gamma))^{-1}, \) and \( \lim_{\gamma \to \infty} \nu_{p,j}(x_\gamma) = \nu_{p,j}(x) \in \mathbb{R}, \) implies \( \lim_{\gamma \to \infty} (t_\gamma^{-1} \partial_p \varphi_k(x_\gamma)) = 0 \) and, thus,

\[
\partial_\nu \partial_p \varphi_k(x) = \lim_{\gamma \to \infty} \frac{\partial_p \varphi_k(x_\gamma) - 0}{t_\gamma} = 0.
\]
completing the proof that \( \varphi_k \) is \( C^\infty \).

**Theorem 2.54.** Let \( n \in \mathbb{N} \).

(a) If the sequence \((\varphi_k)_{k \in \mathbb{N}}\) in \( L^1(\mathbb{R}^n) \) is a Dirac sequence according to Def. 2.52, then

\[
\forall f \in L^1(\mathbb{R}^n) \quad \lim_{k \to \infty} \| \varphi_k * f - f \|_1 = 0. \tag{2.61}
\]

(b) If \( f, g \in L^1(\mathbb{R}^n) \) and at least one of the functions \( f, g \) is bounded, then \( f * g \) is uniformly continuous.

(c) If \( f \in L^1(\mathbb{R}^n) \) and \( g \in C^\infty_c(\mathbb{R}^n) \), then \( f * g \in C^\infty(\mathbb{R}^n) \), and, for each \( p = (p_1, \ldots, p_k) \in \{1, \ldots, n\}^k \), \( k \in \mathbb{N} \),

\[
\partial_p (f * g) = \partial_{p_1} \cdots \partial_{p_k} (f * g) = f * (\partial_p g). \tag{2.62}
\]

(d) If \( O \subseteq \mathbb{R}^n \) is open, then \( C^\infty_c(O, \mathbb{R}^n) \) is dense in \( L^p_c(O, L^1(\mathbb{R}^n), \lambda^n) \) for each \( p \in ]0, \infty[ \).

**Proof.** (a): Let \( f \in L^1(\mathbb{R}^n) \) and \( \epsilon > 0 \). According to (2.57),

\[
\exists \delta > 0 \quad \forall t \in B_\delta(0) \quad \int_{\mathbb{R}^n} \left| f(x + t) - f(x) \right| d\lambda^n(x) < \epsilon.
\]

Since \((\varphi_k)_{k \in \mathbb{N}}\) is a Dirac sequence, there exists \( N \in \mathbb{N} \) such that, for each \( k > N \), \( \text{supp} \varphi_k \subseteq B_\delta(0) \). Thus, for each \( k > N \),

\[
\| \varphi_k * f - f \|_1 = \int_{\mathbb{R}^n} \left| \int_{\mathbb{R}^n} \varphi_k(x - y) f(y) d\lambda^n(y) - f(x) \right| d\lambda^n(x) \\
= \int_{\mathbb{R}^n} \left| \int_{\mathbb{R}^n} \varphi_k(y) (f(x - y) - f(x)) d\lambda^n(y) \right| d\lambda^n(x) \\
\overset{\text{Fubini}}{\leq} \int_{\mathbb{R}^n} \varphi_k(y) \int_{\mathbb{R}^n} \left| f(x - y) - f(x) \right| d\lambda^n(x) \ d\lambda^n(y) \\
\leq \int_{\mathbb{R}^n} \varphi_k(y) \epsilon \ d\lambda^n(y) = \epsilon,
\]

proving (a).

(b): Let \( f, g \in L^1(\mathbb{R}^n) \). Since convolution is commutative, it suffices to consider the case, where \( g \) is bounded. If \( g \) is bounded, then, for each \( x \in \mathbb{R}^n \), \( y \mapsto f(y)g(x - y) \) is \( \lambda^n \)-integrable, implying \( y \mapsto f(x - y)g(y) \) to be \( \lambda^n \)-integrable as well. If \( x, h \in \mathbb{R}^n \), then

\[
| (f * g)(x + h) - (f * g)(x) | \leq \int_{\mathbb{R}^n} \left| f(x + h - y) - f(x - y) \right| | g(y) | \ d\mu^n(y) \\
= \int_{\mathbb{R}^n} \left| f(y + h) - f(y) \right| | g(x - y) | \ d\mu^n(y) \\
\leq \| g \|_\infty \int_{\mathbb{R}^n} \left| f(y + h) - f(y) \right| \ d\mu^n(y).
\]
Thus, according to (2.57), for each $\epsilon > 0$, there exists $\delta > 0$ (not depending on $x \in \mathbb{R}^n$), such that $|(f * g)(x + h) - (f * g)(x)| < \epsilon$ for each $h \in B_\delta(0)$, proving $f * g$ to be uniformly continuous.

(c): Let $f \in L^1(\mathbb{R}^n)$, $g \in C_\infty^0(\mathbb{R}^n)$, $j \in \{1, \ldots, n\}$. Clearly, $\partial_j g \in C_\infty^0(\mathbb{R}^n)$ as well, and $\partial_j g$ is uniformly continuous by [Phil16b, Th. 3.20]. Thus, given $\epsilon > 0$, there exists $\delta > 0$, such that $|\partial_j g(u) - \partial_j g(v)| < \epsilon$ for each $u, v \in \mathbb{R}^n$ with $||u - v||_1 < \delta$. Let $e_j$ denote the $j$th standard unit vector of $\mathbb{R}^n$. Then, for each $x \in \mathbb{R}^n$ and each real number $0 \neq t \in ]-\delta, \delta[$, one computes

$$
\left| \frac{(f * g)(x + te_j) - (f * g)(x)}{t} - (f * (\partial_j g))(x) \right|
= \int_{\mathbb{R}^n} f(y) \left( \frac{g(x + te_j - y) - g(x - y)}{t} - \partial_j g(x - y) \right) \, d\lambda^n(y)
= \int_{\mathbb{R}^n} \frac{f(y)}{t} \int_0^t \left( \partial_j g(x + se_j - y) - \partial_j g(x - y) \right) \, d\lambda^1(s) \, d\lambda^n(y)
\leq \frac{||f||_1}{t} \epsilon,
$$
proving $\partial_j(f * g) = f * (\partial_j g)$. Now (2.62) follows by induction, implying $f * g \in C_\infty^0(\mathbb{R}^n)$ as well.

(d): According to Th. 2.49(a), $C_c(O, \mathbb{R})$ is dense in $L^p_c(\mathbb{R}^n)$ and it merely remains to show that $C_\infty^0(O)$ is dense in $C_c(O)$. Thus, given $f \in C_c(O)$ and $\epsilon > 0$, we need to find $g \in C_\infty^0(O)$ such that $N_p(f - g) < \epsilon$. If $\text{supp } f = \emptyset$, then $f \equiv 0$ and one sets $g := f$. Thus, let $\text{supp } f \neq \emptyset$. Extending $f$ by 0, we regard $f$ as an element of $C_\infty^0(\mathbb{R}^n)$. We also make use of the Dirac sequence $(\varphi_k)_{k \in \mathbb{N}}$ from Ex. 2.53(c), which is in $C_\infty^0(\mathbb{R}^n)$. For each $k \in \mathbb{N}$, we have $f * \varphi_k \in C_\infty^0(\mathbb{R}^n)$: Indeed, $f * \varphi_k \in C_\infty^0(\mathbb{R}^n)$ by (c) and, since $f$ and $\varphi_k$ both have compact support, so has $f * \varphi_k$ by Prop. 2.51(d). Let

$$
\alpha := \begin{cases} 
\text{dist}(\text{supp } f, O^c) \in \mathbb{R}^+ & \text{for } O^c \neq \emptyset, \\
\infty & \text{for } O^c = \emptyset.
\end{cases}
$$

Using the uniform continuity of $f$, we choose $\delta > 0$ such that $|f(u) - f(v)| < \epsilon \cdot (\beta^n(\text{supp } f) + 1)^{-1/p}$ for each $u, v \in \mathbb{R}^n$ with $||u - v|| < \delta$. As

$$
\text{supp } f = \bigcap_{\nu \in \mathbb{N}} K_{\nu}, \quad K_{\nu} := \{ x \in \mathbb{R}^n : \text{dist}(x, \text{supp } f) \leq \nu^{-1} \},
$$

we can choose $\nu_0 \in \mathbb{N}$ such that $\beta^n(K_{\nu_0}) < \beta^n(\text{supp } f) + 1$ and $\nu_0^{-1} < \alpha$. Since $(\varphi_k)_{k \in \mathbb{N}}$ is a Dirac sequence, we can choose $k_0 \in \mathbb{N}$, such that $\text{supp } \varphi_{k_0} \subseteq B_\beta(0)$, $\beta := \min\{\delta, \nu_0^{-1}\}$. Then, using $\int_{\mathbb{R}^n} \varphi_{k_0} \, d\lambda^n = 1$,

$$
\forall x \in \mathbb{R}^n
\frac{|(f * \varphi_{k_0})(x) - f(x)|}{|f(x - y) - f(x)|} \varphi_{k_0}(y) \, d\lambda^n(y)
\leq \int_{\text{supp } \varphi_{k_0}} |f(x - y) - f(x)| \varphi_{k_0}(y) \, d\lambda^n(y)
\leq \epsilon \cdot (\beta^n(\text{supp } f) + 1)^{-1/p}.
$$
We note, using Prop. 2.51(d), \(\text{supp}(f * \varphi_{k_0}) \subseteq \text{supp} f + \text{supp} \varphi_{k_0} \subseteq K_{\nu_0} \subseteq O\), since \(\beta \leq \nu_0^{-1} < \alpha\). Thus, the estimate

\[
N_p((f * \varphi_{k_0}) - f) \leq \beta^n(K_{\nu_0}^{-1})^{1/p} \cdot \epsilon \cdot (\beta^n(\text{supp} f) + 1)^{-1/p} \leq (\beta^n(\text{supp} f) + 1)^{1/p} \cdot \epsilon \cdot (\beta^n(\text{supp} f) + 1)^{-1/p} = \epsilon
\]

concludes the proof.

**Remark 2.55.** As mentioned after the proof of Th. 2.47, \(L^\infty\) is typically less benign than \(L^p\) for \(p < \infty\):

(a) If \(\mu(X) = \infty\), then the space \(S_0\) of Th. 2.47(a) is not dense in \(L^\infty\): If \(f \in S_0\), then there exists \(A \in \mathcal{A}\) such that \(\mu(A) < \infty\) and \(f|_A = 0\), i.e. \(\|f - \chi_X\|_\infty = \|f - 1\|_\infty \geq 1\) (note \(\mu(A^c) = \infty\)).

(b) Consider \((\mathbb{R}^n, \mathcal{A}, \mu^n)\), \(n \in \mathbb{N}\), where \((\mathcal{A}, \mu^n)\) is either \((\mathcal{B}^n, \beta^n)\) or \((\mathcal{L}^n, \lambda^n)\). If \(B \in \mathcal{A}\) is such that \(\mu^n(B) > 0\), then \(L^\infty(B, \mathcal{A}|B, \mu^n|\mathcal{A}|B)\) is not separable (see [Alt06, Ex. 2.17(4)]) – whereas the corresponding \(L^p\) with \(p < \infty\) is separable according to Th. 2.47(e).

(c) Let \(O \subseteq \mathbb{R}^n\) be open. In contrast to the case \(p < \infty\) of Th. 2.54(d), \(C_c(O, \mathbb{K})\) is not dense in \(L^\infty_c(O, \mathcal{A}, \mu^n)\) (using the same notation as in (b) above): Indeed, the closure of \(C^\infty_c(O)\) with respect to \(\|\cdot\|_\infty\) turns out to be the space \(C_0(O)\) as defined in Def. 2.48(b) (see [Rud87, Th. 3.17]).

**Definition 2.56.** Consider \((\mathbb{R}^n, \mathcal{L}^n, \lambda^n)\), \(n \in \mathbb{N}\). For each \(f \in L^1_c(\lambda^n)\), define

\[
\hat{f} : \mathbb{R}^n \to \mathbb{C}, \quad \hat{f}(t) := (2\pi)^{-n/2} \int_{\mathbb{R}^n} e^{-i(t \cdot x)} f(x) \, d\lambda^n(x), \tag{2.63a}
\]

\[
\check{f} : \mathbb{R}^n \to \mathbb{C}, \quad \check{f}(t) := \hat{f}(-t) := (2\pi)^{-n/2} \int_{\mathbb{R}^n} e^{i(t \cdot x)} f(x) \, d\lambda^n(x), \tag{2.63b}
\]

where \((\cdot, \cdot)\) denotes the Euclidean scalar product on \(\mathbb{R}^n\) (note that, since \(|e^{-i(t \cdot x)}| = |e^{i(t \cdot x)}| = 1\), the integrability of \(f\) implies the integrability of both integrands in (2.63)). We call \(\hat{f}\) the **Fourier transform** of \(f\) and \(\check{f}\) the **inverse Fourier transform** of \(f\) (the name for \(\check{f}\) is due to Th. 2.59 below). Note that, if \(f = g\ \lambda^n\text{-a.e.},\ then \(\hat{f} = \hat{g}\) and \(\check{f} = \check{g}\), i.e. \(\hat{f}\) and \(\check{f}\) are also well-defined for \(f \in L^1_c(\lambda^n)\).

**Theorem 2.57.** Let \(n \in \mathbb{N}\) and let \(f, g \in L^1_c(\lambda^n)\).

(a) Both Fourier transform and inverse Fourier transform are linear, i.e. if \(\alpha, \beta \in \mathbb{C}\), then \((\alpha f + \beta g) = \alpha \hat{f} + \beta \hat{g}\) and \((\alpha f + \beta g)' = \alpha \hat{f}' + \beta \hat{g}'.

(b) Riemann-Lebesgue Lemma: \(\hat{f} \in C_0(\mathbb{R}^n)\), i.e. \(\hat{f}\) is continuous with \(\lim_{t \to \infty} \hat{f}(t) = 0\). Moreover \(\hat{f}\) is bounded: \(|\hat{f}| \leq (2\pi)^{-n/2} \|f\|_1\).

(c) \(\hat{f} * g = (2\pi)^{n/2} \hat{f} \cdot \hat{g}\).
(d) For each $a \in \mathbb{R}^n$ and $r \in \mathbb{R}^+$, let

$$f^* : \mathbb{R}^n \to \mathbb{C}, \quad f^*(x) := f(-x),$$

$$f_a : \mathbb{R}^n \to \mathbb{C}, \quad f_a(x) := f(a + x),$$

$$(M_r f) : \mathbb{R}^n \to \mathbb{C}, \quad (M_r f)(x) := r^n f(rx).$$

Then, for each $t \in \mathbb{R}^n$, one has

$$\hat{f}^*(t) = \overline{\hat{f}(t)},$$

$$\hat{f}_a(t) = e^{i \langle a, t \rangle} \hat{f}(t),$$

$$(M_r f)(t) = \hat{f} \left( \frac{t}{r} \right),$$

$$(e^{-i \langle a, x \rangle} f)(t) = (\hat{f}_a(t)).$$

(e) Let $p = (p_1, \ldots, p_k) \in \{1, \ldots, n\}^k$, $k \in \mathbb{N}$. Using the notation $x^p := \prod_{j=1}^{k} x_{p_j}$, for each $x \in \mathbb{R}^n$, and assuming $\partial_p f$ exists as well as $x^q f \in \mathcal{L}^1$ for each $q = (p_\alpha, \ldots, p_k)$ with $\alpha \in \{1, \ldots, k\}$, we have

$$\partial_p \hat{f} = (-i)^k (x^p f)^\wedge.$$ 

Proof. (a) is immediate from the linearity of the integral.

(b): One has

$$\forall t \in \mathbb{R}^n \quad (2\pi)^{n/2} |\hat{f}(t)| = \left| \int_{\mathbb{R}^n} e^{-i \langle t, x \rangle} f(x) \, d\lambda^n(x) \right| \leq \int_{\mathbb{R}^n} |f(x)| \, d\lambda^n(x) = \|f\|_1.$$ 

Since, for each $t \in \mathbb{R}^n$, the integrands in (2.63a) are dominated by the integrable $|f|$, $\hat{f}$ is continuous by Th. 2.22. We note that, for each $t, x \in \mathbb{R}^n$, $t \neq 0$,

$$\left\langle t, x + \frac{\pi}{\|t\|_2^2} t \right\rangle = \langle t, x \rangle + \pi,$$

implying (since $e^{-i\pi} = -1$)

$$(2\pi)^{n/2} \hat{f}(t) = \int_{\mathbb{R}^n} e^{-i \langle t, x \rangle} f(x) \, d\lambda^n(x) \overset{(2.55)}{=} - \int_{\mathbb{R}^n} e^{-i \langle t, x \rangle} f \left( x + \frac{\pi}{\|t\|_2^2} t \right) \, d\lambda^n(x)$$

and

$$2 (2\pi)^{n/2} |\hat{f}(t)| \leq \int_{\mathbb{R}^n} \left| f(x) - f \left( x + \frac{\pi}{\|t\|_2^2} t \right) \right| \, d\lambda^n(x) \to 0 \quad \text{for } t \to \infty$$

by (2.57).
(c): For each \( t \in \mathbb{R}^n \), we compute
\[
\int e^{i(t,x)} \, dx = (2\pi)^{-n/2} \int e^{-i(t,x)} \, dx \quad \text{Fubini} \quad (2.55) = (2\pi)^{-n/2} \int e^{i(t,x)} \, dx = \hat{f}(t),
\]
proving (c).

(d): Let \( t \in \mathbb{R}^n \). Then
\[
\hat{f}^* (t) = (2\pi)^{-n/2} \int e^{-i(t,x)} \hat{f}(x) \, dx \quad \text{Fubini} \quad (2.55) = (2\pi)^{-n/2} \int e^{i(t,x)} \hat{f}(x) \, dx = \hat{f}(t),
\]
proving (2.64a);
\[
\hat{f}_a(t) = (2\pi)^{-n/2} \int e^{-i(t,x)} f(a + x) \, dx \quad \text{Fubini} \quad (2.55) = e^{i(a,t)} \hat{f}(t),
\]
proving (2.64b);
\[
\hat{M}_r f(t) = (2\pi)^{-n/2} \int e^{-i(t,x)} r^n f(rx) \, dx \quad \text{Fubini} \quad (2.55) = \hat{f} \left( \frac{t}{r} \right),
\]
proving (2.64c);
\[
(e^{-i(a,x)} f)^\wedge(t) = (2\pi)^{-n/2} \int e^{-i(t,x)} e^{-i(a,x)} f(x) \, dx = \hat{f}(t + a) = (\hat{f})_a(t),
\]
proving (2.64d).

(e): Let \( j \in \{1, \ldots, n\} \). According to the hypothesis, \(|x_j f|\) is integrable and dominates the corresponding derivative of the integrand of (2.63a) and we can apply Cor. 2.24 to obtain
\[
\partial_j \hat{f}(t) = \partial_{i_j} \hat{f}(t) = (2\pi)^{-n/2} (-i) \int e^{-i(t,x)} x_j f(x) \, dx = (-i) (x_j f)^\wedge(t).
\]
From this, the general formula of (e) follows by induction.  

\begin{example}
(a) Let \( a \in \mathbb{R}^+ \) and consider
\[
f : \mathbb{R} \to \mathbb{C}, \quad f := \chi_{[-a,a[}.
\]
Then \( f \in \mathcal{L}^1 \) and, for each \( t \in \mathbb{R} \),
\[
\hat{f}(t) = (2\pi)^{-1/2} \int_{-a}^a e^{-itx} \, dx = (2\pi)^{-1/2} \left[ e^{-itx} \right]_{-a}^a
\]
\[
= (2\pi)^{-1/2} \left( -\frac{\sin(-ta)}{t} + \frac{\sin(ta)}{t} \right) = 2(2\pi)^{-1/2} \frac{\sin(ta)}{t}.
\]
\end{example}
\(\hat{f} \notin \mathcal{L}^1\), since
\[
(2 (2\pi)^{-1/2})^{-1} \int_{\mathbb{R}} |\hat{f}| \geq \int_0^{\infty} \frac{|\sin(ta)|}{t} \, dt = \int_0^{\infty} \frac{|\sin x|}{x} \, dx \\
\geq \sum_{k=1}^{\infty} \left( \frac{1}{k\pi} \int_{(k-1)\pi}^{k\pi} |\sin x| \, dx \right) = \frac{2}{\pi} \sum_{k=1}^{\infty} \frac{1}{k} = \infty.
\]

(b) Consider \(\phi : \mathbb{R} \to \mathbb{C}, \quad \phi(x) := \max\{0, 1 - |x|\}\). (2.65a)
Then \(\phi \in \mathcal{L}^1\) and we show
\[
\hat{\phi} : \mathbb{R} \to \mathbb{C}, \quad \hat{\phi}(t) = (2\pi)^{-1/2} 2 \left(1 - \cos \frac{t}{2}\right)^2.
\]

We compute, for each \(t \in \mathbb{R}\),
\[
\hat{\phi}(t) = (2\pi)^{-1/2} \int_{-1}^{1} e^{-itx} (1 - |x|) \, dx.
\]
According to (a) (with \(a = 1\)),
\[
\int_{-1}^{1} e^{-itx} \, dx = 2 \frac{\sin t}{t}.
\]

Moreover,
\[
\int_{-1}^{1} e^{-itx} |x| \, dx = \int_{-1}^{1} |x| \cos(tx) \, dx - i \int_{-1}^{1} |x| \sin(tx) \, dx = 2 \int_{0}^{1} x \cos(tx) \, dx \\
= 2 \left[ \frac{\cos(tx)}{t^2} + \frac{x \sin(tx)}{t} \right]_{0}^{1} = 2 \left( \frac{\cos t}{t^2} + \frac{\sin t}{t} - \frac{1}{t^2} \right).
\]

Putting everything together, we obtain (2.65b). By (2.65b), \(\hat{\phi} \in \mathcal{L}^1\), since
\[
\int_{\mathbb{R}} |\hat{\phi}| \, d\lambda^1 = 2 \int_{0}^{1} |\hat{\phi}| \, d\lambda^1 + 2 \int_{1}^{\infty} |\hat{\phi}| \, d\lambda^1 < \infty \quad (2.65c)
\]
(the first integral is finite, since \(|\hat{\phi}|\) is continuous on the bounded set \([0, 1]\), the second integral is also finite, as \(\int_{1}^{\infty} |\hat{\phi}| \, d\lambda^1 \leq 4(2\pi)^{-1/2} \int_{1}^{\infty} t^{-2} \, dt = 4(2\pi)^{-1/2}\)). We now show
\[
(\hat{\phi})^{\vee} = \phi : \quad (2.65d)
\]
For each \(x \in \mathbb{R}\), we obtain
\[
\pi (\hat{\phi})^{\vee}(x) = \frac{1}{2} \int_{\mathbb{R}} e^{itx} \left( \frac{\sin(t/2)}{t/2} \right)^2 \, dt = \int_{\mathbb{R}} t^{-2} (1 - \cos t \cos(tx)) \, dt.
\]
Note
\[
\cos(tx + t) + \cos(tx - t) = \cos(tx) \cos t - \sin(tx) \sin t + \cos(tx) \cos t - \sin(tx) \sin(-t) = 2 \cos(tx) \cos t.
\]
Thus,
\[
\pi (\hat{\phi})^\vee (x) = 2 \int_0^\infty t^{-2} \left( \cos(tx) - \frac{1}{2} \left( \cos(t(x+1)) + \cos(t(x-1)) \right) \right) dt
= 2 \left[ -t^{-1}(1 - \cos t) \cos(tx) \right]_0^\infty
- 2 \int_0^\infty t^{-1} \left( x \sin(tx) - \frac{1}{2} \left( (x+1) \sin(t(x+1)) + (x-1) \sin(t(x-1)) \right) \right) dt.
\]
We proceed to evaluate each term:
\[
\left[ -t^{-1}(1 - \cos t) \cos(tx) \right]_0^\infty = \left[ t \frac{1 - \cos t}{t^2} \cos(tx) \right]_0^\infty = 0.
\]
For \(x = 0\), we have \(\int_0^\infty t^{-1} x \sin(tx) = 0\). For \(x > 0\), we have
\[
\int_0^\infty t^{-1} x \sin(tx) = x \int_0^\infty \frac{\sin u}{u} du \overset{(F.2)}{=} x \frac{\pi}{2}.
\]
For \(x < 0\), we have
\[
\int_0^\infty t^{-1} x \sin(tx) = x \int_0^{-\infty} \frac{\sin u}{u} du \overset{(F.2)}{=} -x \frac{\pi}{2}.
\]
Altogether, we obtain
\[
\forall x \leq -1 \quad \pi (\hat{\phi})^\vee (x) = -\frac{\pi}{2} \left( -x + \frac{x+1}{2} + \frac{x-1}{2} \right) = 0,
\]
\[
\forall -1 \leq x \leq 0 \quad \pi (\hat{\phi})^\vee (x) = -\pi \left( -x - \frac{x+1}{2} + \frac{x-1}{2} \right) = \pi (x+1),
\]
\[
\forall 0 \leq x \leq 1 \quad \pi (\hat{\phi})^\vee (x) = -\pi \left( x - \frac{x+1}{2} + \frac{x-1}{2} \right) = \pi (-x+1),
\]
\[
\forall 1 \leq x \quad \pi (\hat{\phi})^\vee (x) = -\frac{\pi}{2} \left( x - \frac{x+1}{2} - \frac{x-1}{2} \right) = 0,
\]
showing (2.65d).

(c) Let \(n \in \mathbb{N}\). We will use the results of (b) to compute the Fourier transforms of the members of the Dirac sequence \((\varphi_k)_{k \in \mathbb{N}}\) of Ex. 2.53(b), and we will show that (2.65d) still holds for \(\phi\) replaced by \(\varphi_k\). Recall that
\[
\varphi_k : \mathbb{R}^n \rightarrow \mathbb{C}, \quad \varphi_k(x) := \prod_{j=1}^n \max\{0, k - k^2|x_j|\}.
\]
Using $\phi$ from (2.65a) and the notation from Th. 2.57(d), we define, for each $k \in \mathbb{N}$,
\[ \phi_k : \mathbb{R} \to \mathbb{C}, \quad \phi_k(x) := (M_k\phi)(x) = k\phi(kx) = \max\{0, k - k^2|x|\}. \]

Then (2.65b) together with (2.64c) yields
\[ \hat{\phi}_k : \mathbb{R} \to \mathbb{C}, \quad \hat{\phi}_k(t) = \hat{\phi}\left(\frac{t}{k}\right) = \left(2\pi\right)^{-1/2} \left(\frac{\sin(t/(2k))}{t/(2k)}\right)^2. \]

In other words, $\hat{\phi}_k = k(M_{1/k}\hat{\phi})$ and we can apply (2.64c) again to obtain, for each $x \in \mathbb{R}$,
\[ (\hat{\phi}_k)^\vee(x) = (\hat{\phi}_k)(-x) = k\hat{\phi}(-kx) = k(\hat{\phi})^\vee(kx) = k\phi(kx) = \phi_k(x), \]
showing
\[ \forall k \in \mathbb{N} \quad (\hat{\phi}_k)^\vee = \phi_k. \]

Finally, for $\varphi_k$, we obtain, for each $t \in \mathbb{R}^n$,
\[ \begin{align*}
\hat{\varphi}_k(t) &= (2\pi)^{-n/2} \int_{\mathbb{R}^n} e^{-i \langle t, x \rangle} \varphi_k(x) \, d\lambda^n(x) \\
&= (2\pi)^{-n/2} \int_{\mathbb{R}^n} \prod_{j=1}^n (e^{-i t_j x_j} \phi_k(x_j)) \, d\lambda^n(x) \\
&= \text{Fubini} \prod_{j=1}^n \hat{\phi}_k(t_j) = (2\pi)^{-n/2} \prod_{j=1}^n \left(\sin\left(t_j/(2k)\right)/t_j/(2k)\right)^2,
\end{align*} \tag{2.66a} \]

and, for each $x \in \mathbb{R}^n$,
\[ \begin{align*}
(\hat{\varphi}_k)^\vee(x) &= (2\pi)^{-n/2} \int_{\mathbb{R}^n} e^{i \langle t, x \rangle} \hat{\varphi}_k(t) \, d\lambda^n(t) \\
&= (2\pi)^{-n/2} \int_{\mathbb{R}^n} \prod_{j=1}^n \left(e^{i t_j x_j} \left(\sin\left(t_j/(2k)\right)/t_j/(2k)\right)^2\right) \, d\lambda^n(t) \\
&= \text{Fubini} \prod_{j=1}^n (\hat{\varphi}_k)^\vee(x_j) = \prod_{j=1}^n \phi_k(x_j) = \varphi_k(x),
\end{align*} \]
proving
\[ \forall k \in \mathbb{N} \quad (\hat{\varphi}_k)^\vee = \varphi_k. \tag{2.66b} \]

Due to Th. 2.57(b), not every function in $L^1$ can be the Fourier transform of an $L^1$ function. On the other hand, we know from Ex. 2.58(a) that the Fourier transform of an $L^1$ function is not necessarily in $L^1$. Still one has the following inversion Th. 2.59(a) for the Fourier transform of $L^1$ functions. But the point is that, due to the mentioned shortcomings of the Fourier transform for $L^1$ functions, one has to assume $f \in L^1$ in Th. 2.59(a). There is a subspace $\mathcal{S} \subseteq L^1 \cap L^2$, the space of so-called Schwartz functions, on which the Fourier transform is, indeed, bijective, but the details are beyond the scope of the present class.
Theorem 2.59 (Inversion Theorem). Let n ∈ \mathbb{N}.

(a) If f ∈ L^1_C(\lambda^n) and \hat{f} ∈ L^1_C(\lambda^n), then

\[ f = (\hat{f})^\vee \quad \text{\lambda}^n\text{-a.e.} \]

(the equality becomes exact for f, \hat{f} ∈ L^1_C(\lambda^n)).

(b) If f, g ∈ L^1_C(\lambda^n), then \hat{f} = \hat{g} implies f = g \text{ \lambda}^n\text{-a.e.} (i.e. the Fourier transform is injective as a map on L^1_C(\lambda^n)).

Proof. (a): We will make use of the fact that, from (2.66b) of Ex. 2.58(c), we already know \varphi_k = (\hat{\varphi}_k)^\vee for each member of the Dirac sequence (\varphi_k)_{k \in \mathbb{N}} of Ex. 2.53(b). Since \hat{\varphi}_k is bounded by Th. 2.57(b) and \hat{f} ∈ L^1, we have \hat{\varphi}_k \hat{f} ∈ L^1 as well. Thus, we obtain, for each \( x \in \mathbb{R}^n \),

\[
(\hat{\varphi}_k \hat{f})^\vee (x) = (2\pi)^{-n/2} \int_{\mathbb{R}^n} e^{i (x,t)} \hat{\varphi}_k(t) \hat{f}(t) \, d\lambda^n(t)
\]

\[
= (2\pi)^{-n/2} (2\pi)^{-n/2} \int_{\mathbb{R}^n} e^{i (x,t)} \hat{\varphi}_k(t) \int_{\mathbb{R}^n} e^{-i (t,z)} f(z) \, d\lambda^n(z) \, d\lambda^n(t)
\]

\[
\text{Fubini} \quad (2\pi)^{-n/2} \int_{\mathbb{R}^n} f(z) \int_{\mathbb{R}^n} e^{i (x-z)} \hat{\varphi}_k(t) \, d\lambda^n(t) \, d\lambda^n(z)
\]

\[
= (2\pi)^{-n/2} \int_{\mathbb{R}^n} f(z) (\hat{\varphi}_k)^\vee (x-z) \, d\lambda^n(z)
\]

\[
\overset{(2.66b)}{=} (2\pi)^{-n/2} \int_{\mathbb{R}^n} f(z) \varphi_k(x-z) \, d\lambda^n(z) = (2\pi)^{-n/2} (f * \varphi_k)(x). \quad (2.67)
\]

Next, we note that \( \lim_{k \to \infty} \hat{\varphi}_k(t) = (2\pi)^{-n/2} \) for each \( t \in \mathbb{R}^n \) due to (2.66a), and

\[
\|\hat{\varphi}_k\|_\infty \leq (2\pi)^{-n/2} \|\varphi_k\|_1 = (2\pi)^{-n/2}
\]

by Th. 2.57(b) and Def. 2.52(ii). Thus, we can apply the dominated convergence Th. 2.20 to obtain

\[
\lim_{k \to \infty} \| (2\pi)^{n/2} \hat{\varphi}_k \hat{f} - \hat{f} \|_1 = \lim_{k \to \infty} \int_{\mathbb{R}^n} \left| (2\pi)^{n/2} \hat{\varphi}_k(t) \hat{f}(t) - \hat{f}(t) \right| d\lambda^n(t) \overset{\text{DCT}}{=} 0.
\]

Recalling that \( \hat{f}(t) = \hat{f}(-t) \), we can apply Th. 2.57(b) again to conclude

\[
\lim_{k \to \infty} \left\| (2\pi)^{n/2} (\hat{\varphi}_k \hat{f})^\vee - (\hat{f})^\vee \right\|_\infty = \lim_{k \to \infty} \left\| (2\pi)^{n/2} (\hat{\varphi}_k \hat{f})^\vee - (\hat{f})^\vee \right\|_\infty \leq (2\pi)^{-n/2} \lim_{k \to \infty} \left\| (2\pi)^{n/2} \hat{\varphi}_k \hat{f} - \hat{f} \right\|_1 = 0,
\]

i.e. uniform convergence \((2\pi)^{n/2} (\hat{\varphi}_k \hat{f})^\vee \to (\hat{f})^\vee\). Applying Prop. G.4 of the Appendix then yields, for each \( r \in \mathbb{R}^+ \),

\[
\int_{B_r(0)} |(\hat{f})^\vee - f| \, d\lambda^n \overset{(G.4a)}{=} \lim_{k \to \infty} \int_{B_r(0)} |(2\pi)^{n/2} (\hat{\varphi}_k \hat{f})^\vee - f| \, d\lambda^n \overset{(2.67)}{=} \lim_{k \to \infty} \int_{B_r(0)} |(f * \varphi_k) - f| \, d\lambda^n \leq \lim_{k \to \infty} \| \varphi_k * f - f \|_1 \overset{(2.61)}{=} 0,
\]
proving \((\hat{f})^\vee = f\) \(\lambda^n\text{-a.e.}\) as claimed.

(b): If \(\hat{f} = \hat{g}\), then \((f - g)^\vee = 0\). Then, according to (a), \(f - g = 0\) \(\lambda^n\text{-a.e.}\), proving (b).

\[\tag{2.10}\]

Theorem 2.60 (Plancherel). Let \(n \in \mathbb{N}\).

(a) Fourier transform is an \(L^2\)-isometry. More precisely, if \(f \in L^1_\mathcal{E}(\lambda^n) \cap L^2_\mathcal{E}(\lambda^n)\), then \(\hat{f} \in L^2_\mathcal{E}(\lambda^n)\) and

\[
\int_{\mathbb{R}^n} |f|^2 \, d\lambda^n = \int_{\mathbb{R}^n} |\hat{f}|^2 \, d\lambda^n.
\]

(b) If \(f, g \in L^1_\mathcal{E}(\lambda^n) \cap L^2_\mathcal{E}(\lambda^n)\), then

\[
\int_{\mathbb{R}^n} f \overline{g} \, d\lambda^n = \int_{\mathbb{R}^n} \hat{f} \overline{\hat{g}} \, d\lambda^n.
\]

Proof. (a): Consider \(f \in L^1 \cap L^2\). According to Th. 2.57(d), if we let \(f^* : \mathbb{R}^n \to \mathbb{C}\), \(f^*(x) := \overline{f(-x)}\), then \(\mathcal{F}f^* = \overline{\mathcal{F}f}\). Since \(f \in L^1\), we have \(f^* \in L^1\), and we can form \(g := f \ast f^* \in L^1\). For \(a, b \in \mathbb{R}\), we know \((a - b)^2 = a^2 - 2ab + b^2 \geq 0\), which we can apply with \(a := |f(x - y)|, b := |f(-y)|\) to obtain

\[
\forall x, y \in \mathbb{R}^n \quad |f(x - y) \overline{f(-y)}| \leq \frac{1}{2} \left(|f(x - y)|^2 + |f(-y)|^2\right).
\]

Since \(f \in L^2\), this implies \(y \mapsto f(x - y) \overline{f(-y)}\) to be integrable for each \(x \in \mathbb{R}^n\) and, thus,

\[
\forall x \in \mathbb{R}^n \quad g(x) = \int_{\mathbb{R}^n} f(x - y) \overline{f(-y)} \, d\lambda^n(y) = \int_{\mathbb{R}^n} f(x + y) \overline{f(y)} \, d\lambda^n(y).
\]

We use this equation for \(g\) together with the Cauchy-Schwarz inequality (i.e. (2.49) with \(p = q = 2\)) to estimate, for each \(x, x' \in \mathbb{R}^n\),

\[
|g(x) - g(x')|^2 \leq \left( \int_{\mathbb{R}^n} |f(x + y) - f(x' + y)| \, |f(y)| \, d\lambda^n(y) \right)^2 \leq \left( \int_{\mathbb{R}^n} |f(x + y) - f(x' + y)|^2 \, d\lambda^n(y) \right) \cdot \left( \int_{\mathbb{R}^n} |f(y)|^2 \, d\lambda^n(y) \right),
\]

where the right-hand side converges to 0 for \(x \to x'\) (due to Th. 2.49(b) and \(f \in L^2\)). In other words, the function \(g\) is continuous and

\[
\forall x \in \mathbb{R}^n \quad \exists \varepsilon \in \mathbb{R}^n \quad \forall x \in B_\varepsilon(0) \quad |g(x) - g(0)| < \varepsilon.
\]

Once again, we consider the Dirac sequence \((\varphi_k)_{k \in \mathbb{N}}\) of Ex. 2.53(b). We choose \(N \in \mathbb{N}\) such that \(\text{supp} \varphi_k \subseteq B_\delta(0)\) for each \(k \geq N\). Then,

\[
\forall k \geq N \quad \left(\left|\left(g \ast \varphi_k\right)(x) - g(0)\right| \varphi_k(-x) = \varphi_k(x) \right) \int_{\mathbb{R}^n} \left|g(x) - g(0)\right| \varphi_k(x) \, d\lambda^n(x) < \varepsilon.
\]
and we have shown
\[ \lim_{k \to \infty} (g * \varphi_k)(0) = g(0) = \int_{\mathbb{R}^n} |f|^2 \, d\lambda^n. \]

Since \( g * \varphi_k \in L^1 \), we may form its Fourier transform and use Th. 2.57(c) to conclude
\[ \hat{g} * \hat{\varphi}_k = (2\pi)^n/2 \hat{g} \cdot \hat{\varphi}_k \in L^1 \]
(since \( \hat{g} \) is bounded by Th. 2.57(b) and \( \hat{\varphi}_k \in L^1 \)). Thus, according to the inversion Th. 2.59(a),
\[ \left((2\pi)^{n/2} \hat{g} \cdot \hat{\varphi}_k\right)^\vee = (g * \varphi_k)^\vee = g * \varphi_k, \]
where the equality even holds exactly (not merely \( \lambda^n \text{-a.e.} \)), since \( g * \varphi_k \) is continuous by Th. 2.54(b). Next, we apply Fatou (Prop. 2.19) together with \( \lim_{k \to \infty} \hat{\varphi}_k(t) = (2\pi)^{-n/2} \)
and \( \hat{g} = \int \hat{f} \, d\lambda^n = (2\pi)^{n/2} \int |\hat{f}|^2 \, d\lambda^n \) to obtain
\[ \int_{\mathbb{R}^n} |\hat{f}|^2 \, d\lambda^n = (2\pi)^{n/2} \liminf_{k \to \infty} \int_{\mathbb{R}^n} \hat{\varphi}_k \hat{g} \, d\lambda^n \]
\[ = (2\pi)^{n/2} \liminf_{k \to \infty} (\hat{\varphi}_k \hat{g})(0) = \liminf_{k \to \infty} (g * \varphi_k)(0) = \int_{\mathbb{R}^n} |f|^2 \, d\lambda^n. \] (2.68)

In (2.68), the right-hand side is finite, showing \( \hat{f} \in L^2 \). Since \( ||\hat{\varphi}||_\infty \leq 1 \), this also shows that, instead of applying Fatou, one can actually use the dominated convergence Th. 2.20 to obtain (2.68) with limits instead of limit inferiors and with equality instead of the estimate, thereby proving completing the proof of (a).

(b): Since, for each \( u, v \in \mathbb{C} \), one has
\[ u\overline{v} = \frac{1}{4} \left( (u + v)(\overline{u} + \overline{v}) - (u - v)(\overline{u} - \overline{v}) + i(u + iv)(\overline{u} - i\overline{v}) - i(u - iv)(\overline{u} + i\overline{v}) \right) \]
\[ = \frac{1}{4} \left( |u + v|^2 - |u - v|^2 + i|u + iv|^2 - i|u - iv|^2 \right), \]
(b) is a direct consequence of (a).

\[ \square \]

**Example 2.61.** The Plancherel Th. 2.60 can sometimes be used to evaluate integrals in a simple manner: According to Ex. 2.58(a), for
\[ f_a : \mathbb{R} \to \mathbb{C}, \quad f_a := \chi_{[-a,a]}, \quad a \in \mathbb{R}^+, \]
we have
\[ \hat{f}_a : \mathbb{R} \to \mathbb{C}, \quad \hat{f}_a(t) = 2(2\pi)^{-1/2} \frac{\sin(ta)}{t}. \]

Thus, for each \( a, b \in \mathbb{R}^+ \),
\[ \int_{-\infty}^{\infty} \frac{\sin(at) \sin(bt)}{t^2} \, dt = \frac{2\pi}{4} \int_{\mathbb{R}} \hat{f}_a \hat{f}_b \, d\lambda^1 \]
\[ = \frac{\pi}{2} \int_{\mathbb{R}} f_a \overline{f}_b \, d\lambda^1 = \pi \min\{a, b\}. \]

Here \( f_a, f_b \in L^1 \cap L^2 \) and, even though \( \hat{f}_a, \hat{f}_b \) are not integrable (according to Ex. 2.58(a)), \( \hat{f}_a \hat{f}_b \) turns out to be integrable.
2.7 Change of Variables

Our main goal in this section is to prove the change of variables theorem for (subsets of) \( \mathbb{R}^n \) (Th. 2.66 below). It generalizes Th. 1.73 and (2.55) regarding the transformation of \( \beta^n \) and \( \lambda^n \) (and corresponding integrals) with respect to affine bijective transformations \( \phi \) to general diffeomorphisms \( \phi \) (see Def. and Rem. 2.64). It also partially generalizes the one-dimensional version of [Phi16a, Th. 10.25]. The \( n \)-dimensional version is significantly harder to prove and we will need some preparation. Still, for simplicity, the hypothesis in Th. 2.66 will be stronger than necessary. The hypothesis that the transformation is a diffeomorphism can be relaxed, see, e.g., [Els07, Th. V.4.10] and [Rud87, Th. 7.26]. As shown in [Els07], one obtains [Els07, Th. V.4.10] from Th. 2.66 with not too much extra work. However, the version [Rud87, Th. 7.26] makes use of the Brouwer fixed-point theorem of algebraic topology (see, e.g., [Oss09, 5.6.10]). We begin our preparations:

**Proposition 2.62.** If \( (X, \mathcal{A}, \mu) \) is a measure space and \( f : X \to [0, \infty] \) is measurable, then
\[
(f \mu)(A) := \int_A f \, d\mu,
\]
(2.69)
defines a measure on \( (X, \mathcal{A}) \).

**Proof.** We have \((f \mu)(\emptyset) = \int_\emptyset f \, d\mu = 0 \) and, if \((A_i)_{i \in \mathbb{N}}\) is a sequence of disjoint sets in \( \mathcal{A} \), then
\[
(f \mu) \left( \bigcup_{i=1}^\infty A_i \right) \quad = \quad \int_X \left( f \sum_{i=1}^\infty \chi_{A_i} \right) \, d\mu = \sum_{i=1}^\infty \int_X f \chi_{A_i} \, d\mu
\]
\[
= \sum_{i=1}^\infty \int_{A_i} f \, d\mu = \sum_{i=1}^\infty (f \mu)(A_i),
\]
(2.70)
thereby establishing the case.\[\square\]

**Definition 2.63.** If \( \mu, \nu \) are measures on the measurable space \( (X, \mathcal{A}) \), then a measurable function \( f : X \to [0, \infty] \) is called a density of \( \nu \) with respect to \( \mu \) if, and only if, \( \nu = f \mu \) with \( f \mu \) as in (2.69).

**Definition and Remark 2.64.** Let \( n \in \mathbb{N} \) and let \( O, U \subseteq \mathbb{R}^n \) be open. A bijective map \( \phi : O \to U \) is called a \( C^1 \)-diffeomorphism if, and only if, both \( \phi \) and \( \phi^{-1} \) are continuously differentiable. If \( \phi \) and \( \phi^{-1} \) are differentiable, then, by the chain rule,
\[
\forall x \in O \quad D(\phi^{-1})(\phi(x)) \circ D\phi(x) = \text{Id},
\]
showing all \( D\phi(x) \) and all \( D(\phi^{-1})(\phi(x)) \) to be invertible (i.e. they have nonzero determinants). Thus, by the inverse function theorem ([Phi16b, Th. 4.50]), a bijective continuously differentiable map \( \phi : O \to U \) is a \( C^1 \)-diffeomorphism if, and only if, \( \det D\phi \) has no zeros.
**Definition and Remark 2.65.** Let \( n \in \mathbb{N} \). Given a measure space \((\mathbb{R}^n, \mathcal{A}, \mu)\) and \( A \in \mathcal{A} \), we introduce the abbreviations

\[
\mathcal{A}_A := \mathcal{A}|_A, \quad \mu_A := \mu|_{\mathcal{A}_A}.
\]

Let \( O, U \subseteq \mathbb{R}^n \) be open. If \( \phi : O \to U \) is a homeomorphism (in particular, if \( \phi \) is a \( C^1 \)-diffeomorphism), then

\[
\mathcal{B}_U^n = \{ \phi(A) : A \in \mathcal{B}_O^n \}:
\]

Indeed, both \( \phi \) and \( \phi^{-1} : U \to O \) are continuous and, thus, Borel-measurable. Thus, if \( A \in \mathcal{B}_O^n \), then \( \phi(A) = (\phi^{-1})^{-1}(A) \in \mathcal{B}_U^n \). On the other hand, if \( B \in \mathcal{B}_U^n \), then \( A := \phi^{-1}(B) \in \mathcal{B}_O^n \), implying \( B = \phi(A) \).

**Theorem 2.66 (Change of Variables).** Let \( n \in \mathbb{N} \). Consider \((\mathbb{R}^n, \mathcal{A}, \mu^n)\), \( n \in \mathbb{N} \), where \((\mathcal{A}, \mu^n)\) is either \((\mathcal{B}^n, \beta^n)\) or \((\mathcal{L}^n, \lambda^n)\). Let \( O, U \subseteq \mathbb{R}^n \) be open and let \( \phi : O \to U \) be a \( C^1 \)-diffeomorphism.

(a) \(|\det(D\phi)|^{-1}\) is a density of \( \phi\mu^n_O \) with respect to \( \mu^n_U \), i.e.

\[
\phi\mu^n_O = |\det(D\phi)|^{-1}\mu^n_U,
\]

\[
\forall B \in \mathcal{A}_U \quad \mu^n\left(\phi^{-1}(B)\right) = \int_B |\det(D\phi)|^{-1} d\mu^n
\]

\[
\forall A \in \mathcal{A}_O \quad \mu^n\left(\phi(A)\right) = \int_A |\det(D\phi)| d\mu^n.
\]

(b) \( f \in \mathcal{M}^+(U, \mathcal{A}_U) \) or \( f : U \to \mathbb{K} \) is \( \mu^n \)-integrable if, and only if, \((f \circ \phi)|\det(D\phi)|\)

has the corresponding property with respect to \( O \) and, in that case,

\[
\int_U f d\mu^n = \int_O (f \circ \phi)|\det(D\phi)| d\mu^n.
\]

(c) Let \( A \in \mathcal{A}_O \). If \( f \in \mathcal{M}^+(\phi(A), \mathcal{A}_{\phi(A)}) \) or \( f : \phi(A) \to \mathbb{K} \) is \( \mu^n \)-integrable, then

(2.72) holds with \((O, U)\) replaced by \((A, \phi(A))\).

**Proof.** We consider the countable semiring

\[
\mathcal{S} := \left\{ [a, b] : a, b \in \bigcup_{k=1}^{\infty} 2^{-k}\mathbb{Z}^n, a \leq b, [a, b] \subseteq O \right\}
\]

(\(\mathcal{S}\) is countable as the set of admissible interval endpoints is countable; that \(\mathcal{S}\) is a semiring follows, for \( n = 1 \), in the same manner as in Ex. 1.16(b) and, then, for \( n > 1 \), from Prop. 1.17(a)). Since every open subset of \( O \) is the (countable) union of sets from \( \mathcal{S} \), we have \( \sigma(\mathcal{S}) = \mathcal{B}_O^n \). We devide the proof into several steps: The first step consists of showing that (2.71c) holds with \( \leq \) for each \( A \in \mathcal{S} \). This already contains the main work of the proof.
Claim 1.
\[
\forall I \in S \quad \beta^n(\phi(I)) \leq \int_I |\operatorname{det}(D\phi)| \, d\beta^n.
\]

Proof. Let \(\epsilon > 0\). The key idea is to write \(I \in S\) as the union of sufficiently small intervals \(I_\nu\) and, on each \(I_\nu\), to approximate \(\phi\) by a suitable affine map \(x_\nu + L_\nu\). It will be convenient to use the max-norm \(\|\cdot\|_\infty = \|\cdot\|_\infty\) on \(\mathbb{R}^n\). Recalling \(\operatorname{det}(D\phi(x)), \operatorname{det}(D(\phi^{-1})(x)) \in \mathbb{R}^n\), it would also be convenient to use the corresponding operator norm \(\|\cdot\|_\infty\) on \(\mathbb{R}^n\), since it satisfies \(\|T(x)\|_\infty \leq \|T\|_\infty \|x\|_\infty\) for each \(T \in \mathbb{R}^n\) and each \(x \in \mathbb{R}^n\). But if one is not familiar with the concept of an operator norm, one can avoid it, one can recall from [Phi16b, Lem. 5.4] that there exists a norm \(\|\cdot\|\) on \(\mathbb{R}^n\), by the equivalence of norms, there exists a constant \(c_0 \in \mathbb{R}^+\) such that
\[
\forall T \in \mathbb{R}^{n^2} \quad \forall x \in \mathbb{R}^n \quad \|T(x)\|_\infty \leq c_0 \|T\| \|x\|_\infty.
\]
In other words, using balls with respect to \(\|\cdot\|_\infty\),
\[
\forall T \in \mathbb{R}^{n^2} \quad \forall \rho > 0 \quad T(B_\rho(0)) \subseteq B_{c_0\rho\|T\|}(0).
\]
(2.73)

For \(a \in \bar{T}\), consider the map
\[
h_a : O \to \mathbb{R}^n, \quad h_a(x) := \phi(x) - (D\phi(a))(x),
\]
\[
Dh_a : O \to \mathbb{R}^{n^2}, \quad Dh_a(x) := D\phi(x) - D\phi(a).
\]
Due to [Phi16b, Th. 4.38], there exists a constant \(c_1 \in \mathbb{R}^+\) such that, given an open convex set \(C \subseteq O\) and \(\alpha \in \mathbb{R}^+\) with
\[
\forall a \in \bar{T} \quad \forall x \in C \quad \|Dh_a(x)\| \leq \alpha,
\]
each \(h_a\) is \((c_1\alpha)\)-Lipschitz on \(C\) with respect to \(\|\cdot\|_\infty\), i.e.
\[
\forall a \in \bar{T} \quad \forall x, y \in C \quad \|h_a(x) - h_a(y)\|_\infty \leq c_1 \alpha \|x - y\|_\infty.
\]
(2.74)

On the compact set \(\bar{T} \subseteq O\), the continuous map \(x \mapsto \|D(\phi(x))^{-1}\|\) attains its max
\[
M := \max \{\|D(\phi(x))^{-1}\| : x \in \bar{T}\} \in \mathbb{R}^+.
\]
(2.75)

Since \(\bar{T} \subseteq O\), we have \(\operatorname{dist}(\bar{T}, O^c) > 0\) and we may choose \(s > 0\) such that
\[
K_s := \{x \in \mathbb{R}^n : \operatorname{dist}(x, \bar{T}) \leq s\} \subseteq O
\]
(as we compute dist with respect to \(\|\cdot\|_\infty\), \(K_s\) is even an interval). Since \(K_s\) is a closed and bounded subset of \(\mathbb{R}^n\), \(K_s\) is compact. Since the continuous map \(D\phi : O \to \mathbb{R}^{n^2}, x \mapsto D\phi(x)\), is uniformly continuous on the compact set \(K_s\) (cf. [Phi16b, Th. 3.20]), there exists \(0 < r \leq s\) such that
\[
\forall a \in \bar{T} \quad \forall x \in B_r(a) \quad \|D\phi(x) - D\phi(a)\| \leq \frac{\epsilon}{c_1 M}.
\]
(2.76)
Since \( r \leq s \), we have \( \overline{B}_r(a) \subseteq K_s \) for each \( a \in \mathcal{T} \). According to (2.76),

\[
\forall \ a \in \mathcal{T}, \quad \forall \ x \in \overline{B}_r(a) \quad \|Dh_a(x)\| \leq \frac{\epsilon}{c_1 M}
\]

and (2.74) yields

\[
\forall \ a \in \mathcal{T}, \quad \forall \ x, y \in \overline{B}_r(a) \quad \|h_a(x) - h_a(y)\|_\infty \leq \frac{\epsilon}{M} \|x - y\|_\infty.
\] (2.77)

We now choose \( m \in \mathbb{N} \) sufficiently large such that \( I = [a, b] \) with \( a, b \in 2^{-m} \mathbb{Z}^n \). We can then write

\[
I = \bigcup_{\nu=1}^{N} I_\nu, \quad N \in \mathbb{N},
\]

with each \( I_\nu \in \mathcal{S} \) being a hypercube with side length \( d := 2^{-m} \). Then \( \beta^n(I_\nu) = 2^{-mn} \) and \( \beta^n(I) = N \, d^n \). By possibly enlarging \( m \), we may assume \( \text{diam} \ I_\nu < r \). Since \( \det(D\phi) \) is continuous on the compact set \( \mathcal{T}_\nu \), for each \( \nu \in \{1, \ldots, N\} \), we may choose \( a_\nu \in \mathcal{T}_\nu \) such that

\[
\det(D\phi)(a_\nu) = \min \left\{ \left| \det(D\phi)(x) \right| : x \in \mathcal{T}_\nu \right\}.
\] (2.78)

We apply (2.77) with \( y := a_\nu \) to obtain, for each \( \nu \in \{1, \ldots, N\} \) and each \( x \in I_\nu \) (using \( I_\nu \subseteq B_r(a_\nu) \) due to \( \text{diam} \ I_\nu < r \))

\[
\|\phi(x) - \phi(a_\nu) - (D\phi(a_\nu))(x - a_\nu)\|_\infty = \|h_{a_\nu}(x) - h_{a_\nu}(a_\nu)\|_\infty \leq \frac{\epsilon}{M} \|x - a_\nu\|_\infty \leq \frac{\epsilon d}{M}.
\]

Letting \( T_\nu := D\phi(a_\nu) \) and \( y := \phi(x) - \phi(a_\nu) - T_\nu(x - a_\nu) \in B_{\frac{m}{e}}(0) \), this shows

\[
\phi(x) = \phi(a_\nu) + T_\nu(x - a_\nu) + y \quad \text{and} \quad \phi(I_\nu) \subseteq \phi(a_\nu) + T_\nu(I_\nu - a_\nu) + B_{\frac{m}{e}}(0).
\]

Moreover,

\[
B_{\frac{m}{e}}(0) = T_\nu\left(T_\nu^{-1}(B_{\frac{m}{e}}(0))\right) \subseteq T_\nu(B_{c_0 \epsilon d}(0)),
\]

implying

\[
\phi(I_\nu) \subseteq \phi(a_\nu) + T_\nu(I_\nu - a_\nu + B_{c_0 \epsilon d}(0)).
\]

Now note that the set \( I_\nu - a_\nu + B_{c_0 \epsilon d}(0) \) is contained in a hypercube with side length \( d + 2 c_0 \epsilon d = d(1 + 2 c_0 \epsilon) \). In consequence, we obtain

\[
\beta^n(\phi(I_\nu)) \leq \left| \det T_\nu \right| (1 + 2 c_0 \epsilon)^n = \left| \det T_\nu \right| (1 + 2 c_0 \epsilon)^n \beta^n(I_\nu)
\] (1.47b)

and

\[
\beta^n(\phi(I)) = \sum_{\nu=1}^{N} \beta^n(\phi(I_\nu)) \leq (1 + 2 c_0 \epsilon)^n \sum_{\nu=1}^{N} \left| \det T_\nu \right| \beta^n(I_\nu)
\] \leq (2.78)

\[
(1 + 2 c_0 \epsilon)^n \int_I |\det(D\phi)| \, d\beta^n.
\]

Since \( \epsilon > 0 \) was arbitrary, this concludes the proof of Cl. 1. \( \Box \)
Claim 2.
\[ \forall A \in \mathcal{B}_O^n \quad \beta^n(\phi(A)) \leq \int_A |\det(D\phi)| \, d\beta^n. \]

Proof. The statement of Cl. 2 is precisely that \((\beta^n)_{\phi^{-1}} \leq |\det(D\phi)|\beta^n\) holds for the two measures on \(\mathcal{B}_O^n\). However, from Cl. 1, we already know that the measures’ respective restrictions to the generating semiring \(\mathcal{S}\) satisfy the inequality. Since the measures are also \(\sigma\)-finite on \(\mathcal{S}\), the validity of Cl. 2 is due to Cor. 1.46(b). \(\blacksquare\)

Claim 3.
\[ \forall f \in \mathcal{M}^+(U, \mathcal{B}_O^n) \quad \int_U f \, d\beta^n \leq \int_O (f \circ \phi) |\det(D\phi)| \, d\beta^n. \]

Proof. Let \(f = \chi_B\) for \(B \in \mathcal{B}_O^n, A := \phi^{-1}(B)\). Then
\[ \int_U \chi_B \, d\beta^n = \beta^n(B) = \beta^n(\phi(A)) \leq \int_A |\det(D\phi)| \, d\beta^n = \int_O (\chi_B \circ \phi) |\det(D\phi)| \, d\beta^n, \]
proving Cl. 3 for measurable characteristic functions. It then also holds for each simple function \(f \in \mathcal{S}^+(U, \mathcal{B}_O^n)\) by the linearity of the integral. For a general \(f \in \mathcal{M}^+(U, \mathcal{B}_O^n)\), we choose a sequence \((f_k)_{k \in \mathbb{N}}\) in \(\mathcal{S}^+(U, \mathcal{B}_O^n)\) with \(f_k \uparrow f\) and use the monotone convergence Th. 2.7 to estimate
\[ \int_U f \, d\beta^n = \lim_{k \to \infty} \int_U f_k \, d\beta^n \leq \lim_{k \to \infty} \int_O (f_k \circ \phi) |\det(D\phi)| \, d\beta^n \overset{\text{MCT}}{=} \int_O (f \circ \phi) |\det(D\phi)| \, d\beta^n, \]
concluding the proof of Cl. 3. \(\blacksquare\)

Claim 4.
\[ \forall f \in \mathcal{M}^+(U, \mathcal{B}_O^n) \quad \int_U f \, d\beta^n = \int_O (f \circ \phi) |\det(D\phi)| \, d\beta^n. \]

Proof. We apply Cl. 3 with \(\phi^{-1}\) instead of \(\phi\) and \((f \circ \phi) |\det(D\phi)|\) instead of \(f\) to obtain
\[ \int_O (f \circ \phi) |\det(D\phi)| \, d\beta^n \leq \int_U \left( (f \circ \phi) |\det(D\phi)| \circ \phi^{-1} \right) |\det(D\phi^{-1})| \, d\beta^n = \int_U f \, d\beta^n, \]
since \(((D\phi)(\phi^{-1})) \circ D\phi^{-1} = \text{Id}\) by the chain rule. The inequality just proved, together with the one of Cl. 3, establishes Cl. 4. \(\blacksquare\)

Applying Cl. 4 to \(f = \chi_A\) with \(A \in \mathcal{B}_O^n\), we obtain (2.71c) for \((\mathcal{A}, \mu^n) = (\mathcal{B}^n, \beta^n)\). In particular, (2.71c) then says that \(\beta^n(A) = 0\) if, and only if, \(\beta^n(\phi(A)) = 0\), showing that \(\phi\) provides a bijection between \(\beta^n\)-null sets, also implying \(\phi\) to be \(\lambda^n\)-measurable. In consequence, (2.71c) also holds for \((\mathcal{A}, \mu^n) = (\mathcal{L}^n, \lambda^n)\). We then obtain (2.71b) (which is equivalent to (2.71a)) by applying (2.71c) with \(\phi^{-1}\) instead of \(\phi\), completing the proof of (a). According to Cl. 4, (b) holds for \((\mathcal{A}, \mu^n) = (\mathcal{B}^n, \beta^n)\) and \(f \geq 0\). If \(f \in \mathcal{M}^+(U, \mathcal{B}_O^n)\), then, by Prop. G.3, there exists \(g \in \mathcal{M}^+(U, \mathcal{B}_O^n)\), such that \(g = f \beta^n\)-a.e. Then
\[ \int_U f \, d\lambda^n = \int_U g \, d\beta^n = \int_O (g \circ \phi) |\det(D\phi)| \, d\beta^n = \int_O (f \circ \phi) |\det(D\phi)| \, d\lambda^n, \]
where the last equality makes, again, use of the fact that \( \phi \) maps \( \beta^n \)-null sets to \( \beta^n \)-null sets. We have, thus, proved (2.72) for \( f \in \mathcal{M}^+(U, \mathcal{L}^n_\mu) \). The \( \mu^n \)-integrable case then also follows, in the usual way, by decomposing \( f \) into \((\text{Re} \, f)^\pm\) and \((\text{Im} \, f)^\pm\), completing the proof of (b). Finally, (c) is obtained by applying (b) with \( f \) replaced by \( g \) with

\[
g : U \rightarrow \mathbb{K}, \quad g(x) := \begin{cases} f(x) & \text{for } x \in \phi(A), \\ 0 & \text{otherwise.} \end{cases}
\]

This concludes the proof of the theorem. \( \blacksquare \)

**Example 2.67.** Let

\[
O := \mathbb{R}^+ \times ]0, 2\pi[, \quad U := \mathbb{R}^2 \setminus \{(x, 0) : x \geq 0\}. \tag{2.79}
\]

As shown in Sec. H of the Appendix, the function defining polar coordinates, i.e.

\[
\phi : O \rightarrow U, \quad \phi(r, \varphi) := (r \cos \varphi, r \sin \varphi), \tag{2.80}
\]

constitutes a \( C^1 \)-diffeomorphism with \( \det D\phi(r, \varphi) = r \) for each \((r, \varphi) \in O\). Since \( \{(x, 0) : x \geq 0\}, \{0\} \times [0, 2\pi], \mathbb{R}^+_0 \times \{0\}, \) and \( \mathbb{R}^+_0 \times \{2\pi\} \) all are \( \lambda^2 \)-null sets, we can use Th. 2.66(b) to conclude, for each \( f \in \mathcal{M}^+(\mathbb{R}^2, \mathcal{L}^2)\),

\[
\int_{\mathbb{R}^2} f \, d\lambda^2 = \int_{O} f(r \cos \varphi, r \sin \varphi) \, r \, d\lambda^2(r, \varphi). \tag{2.81}
\]

Moreover, \( f : \mathbb{R}^2 \rightarrow \mathbb{K} \) is \( \lambda^2 \)-integrable if, and only if, \((r, \varphi) \mapsto r f(r \cos \varphi, r \sin \varphi)\) is \( \lambda^2 \)-integrable, and then (2.81) holds. We can use (2.81) in combination with Fubini to give a third proof of

\[
\int_{-\infty}^{\infty} e^{-x^2} \, dx = \sqrt{\pi}, \tag{2.82}
\]

We compute

\[
\left(\int_{-\infty}^{\infty} e^{-x^2} \, dx\right)^2 = \int_{\mathbb{R}^2} e^{-x^2-y^2} \, d\lambda^2(x, y) \stackrel{(2.81)}{=} \int_{O} r \, e^{-r^2} \, d\lambda^2(r, \varphi) \]
\[
= \int_{0}^{\infty} \int_{0}^{2\pi} r \, e^{-r^2} \, d\varphi \, dr = 2\pi \int_{0}^{\infty} r \, e^{-r^2} \, dr = 2\pi \left[ \frac{-1}{2} e^{-r^2} \right]_{0}^{\infty} = \pi,
\]

proving (2.82).

## 3 Integration Over Submanifolds of \( \mathbb{R}^n \)

### 3.1 Submanifolds of \( \mathbb{R}^n \)

Roughly speaking, a \( k \)-dimensional submanifold of \( \mathbb{R}^n \) is a subset that, locally, looks like an open subset of \( \mathbb{R}^k \). There are a number of equivalent ways of making this idea precise, see Def. 3.1 and Th. 3.4 below.
Definition 3.1. Let $k, n \in \mathbb{N}, k < n$, $\alpha \in \mathbb{N} \cup \{\infty\}$. A set $M \subseteq \mathbb{R}^n$ is called a $k$-dimensional submanifold of class $C^\alpha$ if, and only if, for each $a \in M$, there exists an open neighborhood $O_a \subseteq \mathbb{R}^n$ of $a$ and a map $f^a \in C^\alpha(O_a, \mathbb{R}^{n-k})$, satisfying the following two conditions

(a) $M \cap O_a = (f^a)^{-1}(\{0\})$,
(b) $\text{rk } Df^a(a) = n - k$, i.e. the $n - k$ vectors $\nabla f^a_1(a), \ldots, \nabla f^a_{n-k}(a)$ are linearly independent elements of $\mathbb{R}^n$.

Submanifolds of dimension $k = 1$ are called paths, curves or lines; submanifolds of dimension $k = 2$ are called surfaces, submanifolds of dimension $k = n - 1$ are called hypersurfaces.

Example 3.2. Let $n \in \mathbb{N}, n \geq 2$.

(a) Perhaps the most simple submanifolds of $\mathbb{R}^n$ are its affine subspaces: Let $k \in \mathbb{N}, k < n$, and let $V \subseteq \mathbb{R}^n$ be a $k$-dimensional vector space over $\mathbb{R}$. Moreover, let $x_M \in \mathbb{R}^n$ and $M := V + x_M$. Then $M$ is a $k$-dimensional submanifold of $\mathbb{R}^n$ of class $C^\infty$: Let $\{v_1, \ldots, v_n\}$ be a basis of $\mathbb{R}^n$ such that $\{v_1, \ldots, v_k\}$ is a basis of $V$. Let $A : \mathbb{R}^n \longrightarrow \mathbb{R}^{n-k}$ be the linear map defined by

$$A(v_i) := \begin{cases} 0 & \text{for } i \in \{1, \ldots, k\}, \\ v_i & \text{for } i \in \{k+1, \ldots, n\}, \end{cases}$$

and

$$f : \mathbb{R}^n \longrightarrow \mathbb{R}^{n-k}, \quad f(x) := A(x - x_M).$$

We can now let, for each $a \in M$, $O_a := \mathbb{R}^n, f^a := f$. We verify Def. 3.1(a),(b) to be satisfied: First note that $f$ is $C^\infty$, since $f$ is the composition of a translation with a linear map, $f = A \circ T_{-x_M}$. Moreover, for each $x \in \mathbb{R}^n$,

$$f(x) = A(x - x_M) = 0 \iff x - x_M \in V \iff x \in V + x_M = M,$$

proving Def. 3.1(a); and

$$\text{rk } Df(x) = \text{rk } A = n - k,$$

proving Def. 3.1(b).

(b) Each Euclidean sphere is an $(n - 1)$-dimensional submanifold of $\mathbb{R}^n$: Let $x_M \in \mathbb{R}^n$ and $r > 0, M := \{x \in \mathbb{R}^n : \|x - x_M\|_2 = r\}$. As in (a), for each $a \in M$, we may take $O_a := \mathbb{R}^n$ and the same $f^a := f$ for each $a \in M$, where

$$f : \mathbb{R}^n \longrightarrow \mathbb{R}, \quad f(x) := \|x - x_M\|_2^2 - r^2 = \sum_{i=1}^n (x_i - x_{M,i})^2 - r^2.$$
Since $f$ is a polynomial, it is $C^\infty$. Since 
\[ M = \{ x \in \mathbb{R}^n : f(x) = 0 \}, \]
Def. 3.1(a) is satisfied. Since 
\[ \forall x \in \mathbb{R}^n \quad Df(x) = \nabla f(x) = 2(x_1 - x_{M,1}, \ldots, x_n - x_{M,n}), \]
$Df(x) \neq 0$ for each $x \neq x_M$. In particular, $\text{rk } Df(a) = 1$ for each $a \in M$.

**Definition 3.3.** Let $n \in \mathbb{N}$, $\alpha \in \mathbb{N} \cup \{ \infty \}$.

(a) Let $O, U \subseteq \mathbb{R}^n$ be open. A bijective map $F : O \rightarrow U$ is called a $C^\alpha$-diffeomorphism if, and only if, $F \in C^\alpha(O, \mathbb{R}^n)$ and $F^{-1} \in C^\alpha(U, \mathbb{R}^n)$.

(b) Let $k \in \mathbb{N}$, $k \leq n$, and let $O \subseteq \mathbb{R}^k$ be open. A map $\phi \in C^\alpha(O, \mathbb{R}^n)$ is called a $C^\alpha$-immersion if, and only if, 
\[ \forall x \in O \quad \text{rk } D\phi(x) = k \quad (3.1) \]
(i.e., for each $x \in O$, the $k$ vectors $\partial_1 \phi(x), \ldots, \partial_k \phi(x)$ are linearly independent elements of $\mathbb{R}^n$).

(c) Let $\pi \in S_n$ be a permutation of $\{1, \ldots, n\}$. Then, clearly, 
\[ P_\pi : \mathbb{R}^n \rightarrow \mathbb{R}^n, \quad P_\pi(x_1, \ldots, x_n) := (x_{\pi(1)}, \ldots, x_{\pi(n)}), \quad (3.2) \]
the map that reorders the variables according to $\pi$, is a linear isomorphism (in particular, a $C^\infty$-diffeomorphism).

**Theorem 3.4.** Given $k, n \in \mathbb{N}$ with $k < n$, $\alpha \in \mathbb{N} \cup \{ \infty \}$, and $M \subseteq \mathbb{R}^n$, the following statements are equivalent:

(i) $M$ is a $k$-dimensional submanifold of class $C^\alpha$.

(ii) $M$ can, locally, be represented as the graph of a $C^\alpha$-function, depending on $k$ variables, i.e., for each $a \in M$, there exists a permutation $\pi^a \in S_n$ and open neighborhoods $A_a$ and $B_a$ of $\alpha^a$ and $\beta^a$, respectively, where $\alpha^a := (a_{\pi^a(1)}, \ldots, a_{\pi^a(k)})$, $\beta^a := (a_{\pi^a(k+1)}, \ldots, a_{\pi^a(n)})$, i.e.
\[ \alpha^a \in A_a \subseteq \mathbb{R}^k, \quad \beta^a \in B_a \subseteq \mathbb{R}^{n-k}, \]
plus a $C^\alpha$-map $g^a : A_a \rightarrow B_a$, satisfying 
\[ M \cap P_{\pi^a}(A_a \times B_a) = P_{\pi^a}(\{(x, y) \in A_a \times B_a : y = g^a(x)\}), \quad \tau^a := (\pi^a)^{-1}. \quad (3.3) \]

(iii) For each $a \in M$, there exists an open neighborhood $V_a \subseteq \mathbb{R}^n$, an open set $U_a \subseteq \mathbb{R}^n$, and a $C^\alpha$-diffeomorphism $F^a : V_a \rightarrow U_a$, satisfying 
\[ F^a(M \cap V_a) = E_k \cap U_a, \quad (3.4) \]
where 
\[ E_k := \{ x \in \mathbb{R}^n : x_{k+1} = \cdots = x_n = 0 \}. \quad (3.5) \]
(iv) For each \( a \in M \), there exists an \( M \)-open neighborhood \( M_a \subseteq M \) (open with respect to the relative topology on \( M \) ), an open set \( W_a \subseteq \mathbb{R}^k \), and a \( C^\infty \)-immersion \( \phi^a : W_a \to \mathbb{R}^n \), such that \( M_a = \phi^a(W_a) \) and \( \phi^a : W_a \to M_a \) is a homeomorphism. The maps \( \phi^a \) are called (local) charts or (local) parametrizations or (local) coordinates of \( M \).

**Proof.** (i) \( \Rightarrow \) (ii): Given \( a \in M \), let \( f^a : O_a \to \mathbb{R}^{n-k} \) be according to Def. 3.1. Since \( \text{rk} \, D f^a(a) = n - k \), there exists \( \pi^a \in S_n \) such that

\[
\begin{vmatrix}
\det \left( \partial_{\pi^a(n)} f_1^a & \cdots & \partial_{\pi^a(n)} f_{n-k}^a \\
\cdots & \cdots & \cdots \\
\partial_{\pi^a(n)} f_{n-k}^a & \cdots & \partial_{\pi^a(n)} f_{n-k}^a \\
\end{vmatrix}(a) \neq 0.
\]

Let \( G_a := P_{\pi^a}(O_a) \). Then \( G_a \subseteq \mathbb{R}^n \) is open by [Phi16b, Cor. 4.51]. Let \( \tau^a := (\pi^a)^{-1} \). Then \( P_{\tau^a}(G_a) = O^a, \ P_{\tau^a}(\alpha^a, \beta^a) = a \). If we let

\[
h^a : G_a \to \mathbb{R}^{n-k}, \quad h^a(x, y) := f^a(P_{\tau^a}(x, y)), \quad x \in \mathbb{R}^k, \ y \in \mathbb{R}^{n-k},
\]

then, by the chain rule,

\[
D h^a(\alpha^a, \beta^a) = D f^a(P_{\tau^a}(\alpha^a, \beta^a)) P_{\tau^a} = D f^a(a) P_{\tau^a}
\]

(by matrix form, \( P_{\tau^a} \) is a permutation matrix that such that right multiplication of \( A \) by \( P_{\tau^a} \) permutes the columns of \( A \) according to \( (\tau^a)^{-1} = \pi^a \). Thus, \( h^a(\alpha^a, \beta^a) = f^a(a) = 0 \), det \( D h^a(\alpha^a, \beta^a) \neq 0 \), and the implicit function theorem [Phi16b, Th. 4.49] provides open neighborhoods \( A_a \subseteq \mathbb{R}^k \) of \( \alpha^a \) and \( B_a \subseteq \mathbb{R}^{n-k} \) of \( \beta^a \), and a \( C^\infty \)-map \( g^a : A_a \to B_a \) such that

\[
(A_a \times B_a) \cap (h^a)^{-1}\{0\} = \{(x, y) \in A_a \times B_a : y = g^a(x)\}.
\]

Then

\[
z \in M \cap P_{\tau^a}(A_a \times B_a) \iff z \in (f^a)^{-1}\{0\} \cap P_{(\pi^a)^{-1}}(A_a \times B_a)
\]

\[
\iff P_{\pi^a}(z) \in (h^a)^{-1}\{0\} \cap (A_a \times B_a)
\]

\[
\iff P_{\pi^a}(z) \in \{(x, y) \in A_a \times B_a : y = g^a(x)\}
\]

\[
\iff z \in P_{\tau^a}\{(x, y) \in A_a \times B_a : y = g^a(x)\},
\]

proving (ii).

(ii) \( \Rightarrow \) (iii): Given \( a \in M \), let \( \pi^a \in S_n, \ (\alpha^a, \beta^a) = P_{\pi^a}(a), \ A_a \subseteq \mathbb{R}^k, \ B_a \subseteq \mathbb{R}^{n-k}, \) and \( g^a : A_a \to B_a \) be according to (ii). Define \( \tau^a := (\pi^a)^{-1}, \ V_a := P_{\tau^a}(A_a \times B_a) \). Then \( V_a \) is an open neighborhood of \( a \) and we define

\[
F^a : V_a \to \mathbb{R}^n, \quad F^a(z) := (x, y - g^a(x)), \quad \text{where} \ (x, y) := P_{\tau^a}(z).
\]
Then $g^a$ being $C^\alpha$ implies $F^a$ to be $C^\alpha$. Moreover, $\det F^a \equiv 1$, since

$$DF^a = \begin{pmatrix} \text{Id}_k & 0 \\ -Dg^a & \text{Id}_{n-k} \end{pmatrix} P_{\pi^n}.$$ 

Then $U_a := F^a(V_a)$ is open by [Phi16b, Cor. 4.51] and $F^a : V^a \to U_a$ constitutes a $C^\alpha$-diffeomorphism. Finally,

$$z \in F^a(M \cap V_a) \iff z \in F^a(M \cap P_\tau(A_a \times B_a)) \quad (3.3)$$

$$\iff z \in F^a(P_\tau\{ (x, y) \in A_a \times B_a : y = g^a(x) \})$$

$$\iff z \in E_k \cap U_a,$$

proving (iii).

(iii) $\Rightarrow$ (i): For $a \in M$, let $F^a : V_a \to U_a$ be given by (iii) such that (3.4) holds. We set $O_a := V_a$ and define

$$f^a : O_a \to \mathbb{R}^{n-k}, \quad \forall_{j \in \{1, \ldots, n-k\}} f^a_j := F^a_{k+j}. \quad (3.6)$$

Then $f^a$ is $C^\alpha$, since $F^a$ is $C^\alpha$, $\rk Df^a(a) = n - k$, since $\rk DF^a(a) = n$ (using that $F^a$ is a diffeomorphism). Moreover, for each $x \in \mathbb{R}^n$,

$$f^a(x) = 0 \iff F^a_{k+1}(x) = \cdots = F^a_{n}(x) = 0 \iff F(x) \in E_k \cap U_a \iff x \in M \cap O_a,$$

proving $M \cap O_a = (f^a)^{-1}(\{0\})$ and, therefore, that $M$ is a $k$-dimensional submanifold of class $C^\alpha$.

(ii) $\Rightarrow$ (iv): Given $a \in M$, let $\pi^n \in S_n$, $\alpha(a, \beta^a) = P_{\pi^n}(a)$, $A_a \subseteq \mathbb{R}^k$, $B_a \subseteq \mathbb{R}^{n-k}$, and $g^a : A_a \to B_a$ be according to (ii). Define $\tau^a := (\pi^n)^{-1}$, $W_a := A_a$, and

$$\phi^a : W_a \to \mathbb{R}^n, \quad \phi^a(x) := P_{\tau^a}(x, g^a(x)).$$

Then $\phi^a$ is $C^\alpha$, since $g^a$ is $C^\alpha$. As, clearly, $\rk D\phi^a = k$, $\phi^a$ is a $C^\alpha$-immersion. Set $M_a := \phi^a(W_a)$. Then $\phi^a(\alpha^a) = P_{\tau^a}(\alpha^a, \beta^a) = a$, shows $a \in M_a$. Due to (3.3), we have

$$M_a = P_{\tau^a}\{ (x, y) \in A_a \times B_a : y = g^a(x) \} = M \cap P_{\tau^a}(A_a \times B_a),$$

showing $M_a$ to be $M$-open. Since

$$(\phi^a)^{-1} : M_a \to W_a, \quad (\phi^a)^{-1}(z) = x, \quad \text{where} \quad (x, y) := P_{\pi^n}(z),$$

is, clearly, continuous, $\phi^a : W_a \to M^a$ is a homeomorphism, proving (iv).
(iv) \(\Rightarrow\) (iii): Given \(a \in M\), let \(W_a \subseteq \mathbb{R}^k\), \(\phi^a : W_a \to \mathbb{R}^n\), \(M_a = \phi^a(W_a)\) be according to (iv). Since \(M_a\) is \(M\)-open, there exists \(O_a \subseteq \mathbb{R}^n\) such that \(M_a = M \cap O_a\). Let \(c \in W_a\) be such that \(\phi^a(c) = a\). Since \(\operatorname{rk} D\phi^a(c) = k\), there exists \(\pi^a \in S_n\) such that
\[
\det \begin{pmatrix}
\partial_1 \phi^a_{\pi^a(1)} & \cdots & \partial_k \phi^a_{\pi^a(1)} \\
\vdots & \ddots & \vdots \\
\partial_1 \phi^a_{\pi^a(k)} & \cdots & \partial_k \phi^a_{\pi^a(k)}
\end{pmatrix}(c) \neq 0.
\]
According to the inverse function theorem [Phi16b, Th. 4.50] there exist open sets \(A_a, B_a \subseteq \mathbb{R}^k\) such that \(c \in A_a \subseteq W_a\) and
\[
\varphi^a : A_a \to B_a, \quad \varphi^a(x) := (\phi^a_{\pi^a(1)}(x), \ldots, \phi^a_{\pi^a(k)}(x)),
\]
is a \(C^\alpha\)-diffeomorphism. Let \(U_a := A_a \times \mathbb{R}^{n-k}\) and define, with \((x, y) \in A_a \times \mathbb{R}^{n-k}\),
\[
\varphi^a : U_a \to B_a \times \mathbb{R}^{n-k},
\]
\[
\varphi^a_j(x, y) := \begin{cases}
\phi^a_{\pi^a(j)}(x) & \text{for } 1 \leq j \leq k, \\
\phi^a_{\pi^a(j)}(x) + y_i & \text{for } j = k + i, 1 \leq i \leq n - k,
\end{cases}
\]
\[
\tau^a := (\pi^a)^{-1},
\]
and
\[
G^a : U_a \to \mathbb{R}^n, \quad G^a_j(z) := \varphi^a_{\pi^a(1)}(z).
\]
Then \(G^a\) is \(C^\alpha\), since \(\varphi^a\) is \(C^\alpha\). Moreover, \(\det DG^a = \det D\varphi^a \neq 0\), since
\[
DG^a = P_{\tau^a} \begin{pmatrix}
D\varphi^a & 0 \\
D_x(\varphi^a_{k+1}, \ldots, \varphi^a_n) & \operatorname{Id}_{n-k}
\end{pmatrix}.
\]
In particular, \(V_a := G^a(U_a)\) is open by [Phi16b, Cor. 4.51] and \(G^a : U_a \to V_a\) is a \(C^\alpha\)-diffeomorphism. Then \(F^a := (G^a)^{-1} : V_a \to U_a\) is a \(C^\alpha\)-diffeomorphism as well. Since \(G^a(c, 0) = \varphi^a(c) = a\), it only remains to show (3.4), i.e.
\[
G^a(E_k \cap U_a) = M \cap V_a.
\]
To prove this equality, note
\[
z \in G^a(E_k \cap U_a) \iff z \in G^a(A_a \times \{0\}) \iff z \in \phi^a(W_a) \cap V_a \iff z \in M \cap V_a,
\]
showing (iii) and completing the proof of the theorem.

\begin{proof}
\end{proof}

**Theorem 3.5 (Chart Transition).** Let \(k, n \in \mathbb{N}\) with \(k < n\), \(\alpha \in \mathbb{N} \cup \{\infty\}\), and let \(M \subseteq \mathbb{R}^n\) be a \(k\)-dimensional submanifold of class \(C^\alpha\). For \(j \in \{1, 2\}\), let \(\phi_j : W_j \to M_j \subseteq M\) be local charts according to Th. 3.4(iv), i.e. \(W_j \subseteq \mathbb{R}^k\) open, the \(M_j\) are \(M\)-open, and both \(\phi_j\) are \(C^\alpha\)-immersions as well as homeomorphisms. Suppose \(V := M_1 \cap M_2 \neq \emptyset\). Then \(O_j := \phi_j^{-1}(V) \subseteq W_j\) are open and
\[
\tau := \phi_2^{-1} \circ \phi_1 : O_1 \to O_2
\]
is a \(C^\alpha\)-diffeomorphism. The map \(\tau\) is then also called a transition between the charts \(\phi_1\) and \(\phi_2\) (note \(\phi_1 = \phi_2 \circ \tau\) on \(O_1\)).
Proof. Since \( V \) is \( M \)-open and \( \phi_j \) are continuous, both \( O_j \) are open. Clearly, \( \tau \) is bijective (injective, since the \( \phi_j \) are injective, surjective by the choice of the \( O_j \)). Since \( \phi_2^{-1} \) is not a \( C^\alpha \)-map defined on an open subset of \( \mathbb{R}^n \), we can not directly apply the chain rule, but have to work slightly harder. Fix an arbitrary \( c_1 \in O_1 \) and set
\[
a := \phi_1(c_1), \quad c_2 := \phi_2^{-1}(a) = \tau(c_1).
\]
Since \( V \) is \( M \)-open, there exists \( \tilde{U} \subseteq \mathbb{R}^n \) open such that \( V = \tilde{U} \cap M \). Since \( a \in V \), by Th. 3.4(iii), there exist open sets \( V_a, U_a \subseteq \mathbb{R}^n \) such that \( a \in V_a \subseteq \tilde{U} \) and there exists a \( C^\alpha \)-diffeomorphism \( F^a : V_a \rightarrow U_a \), satisfying
\[
F^a(M \cap V_a) = E_k \cap U_a.
\]
We also have \( M \cap V_a \subseteq M \cap \tilde{U} = V \). Let \( \tilde{W}_j := \phi_j^{-1}(M \cap V_a) \) and define
\[
g : \tilde{W}_1 \rightarrow \mathbb{R}^k, \quad g(x) := ((F^a \circ \phi_1)_1, \ldots, (F^a \circ \phi_1)_k)(x),
\]
\[
h : \tilde{W}_2 \rightarrow \mathbb{R}^k, \quad h(x) := ((F^a \circ \phi_2)_1, \ldots, (F^a \circ \phi_2)_k)(x).
\]
Now we are in a situation, where the chain rule applies (to \( g \) and to \( h \)), yielding
\[
\forall x \in \tilde{W}_1 \quad Dg(x) = P \circ DF^a(\phi_1(x)) \circ D\phi_1(x),
\]
\[
\forall x \in \tilde{W}_2 \quad Dh(x) = P \circ DF^a(\phi_2(x)) \circ D\phi_2(x),
\]
where \( P \) denotes the projection from \( \mathbb{R}^n \) to \( \mathbb{R}^k \). Since \( \text{rk} \, D\phi_j(x) = k \), \( D\phi_j \) must be injective. Since \( F^a \) is a diffeomorphism, \( D(F^a \circ \phi_j) \) must still be injective. Since \( \phi_j \) maps into \( M \cap V_a \) and \( F^a \) maps \( M \cap V_a \) into \( E_k \) (i.e. the last \( n - k \) components of \( F^a \circ \phi_j \) are identically 0 on \( \tilde{W}_1 \)), both \( Dg(x) \) and \( Dh(x) \) must be surjective as well as injective, showing \( g : \tilde{W}_1 \rightarrow g(\tilde{W}_1) \) and \( h : \tilde{W}_2 \rightarrow g(\tilde{W}_2) \) to be \( C^\alpha \)-diffeomorphisms. Thus,
\[
\tau|_{\tilde{W}_1} = (\phi_2^{-1} \circ \phi_1)|_{\tilde{W}_1} = h^{-1} \circ g
\]
must be a \( C^\alpha \)-diffeomorphism as well. We have shown that, for each \( c_1 \in O_1 \), there exists an open neighborhood \( \tilde{W}_1 \) of \( c_1 \) such that \( \tau \) is a \( C^\alpha \)-diffeomorphism on \( \tilde{W}_1 \). Since \( \tau \) is also bijective on all of \( O_1 \), it must be a \( C^\alpha \)-diffeomorphism on all of \( O_1 \) by Def. and Rem. 2.64 (\( \det D\tau \) has no zeros on \( O_1 \)).

**Theorem 3.6.** Let \( k, n \in \mathbb{N} \) with \( k < n \) and let \( M \subseteq \mathbb{R}^n \) be a \( k \)-dimensional submanifold of class \( C^1 \). Then the following holds:

(a) \( M \) is the union of countably many compact sets.

(b) \( M \in \mathcal{B}^\beta \), i.e. \( M \) is \( \beta^n \)-measurable (one can also show that \( M \) is even a \( \beta^n \)-null set).

**Proof.** (a): Recall that each open set \( O \subseteq \mathbb{R}^n \) is the countable union of compact intervals, for example of intervals from the set \( \mathcal{C} := \mathcal{I}_{c, \mathbb{Q}}^n \) (defined in Ex. 1.8(c)). Let \( a \in M \) and let
Let $f^a \in C^1(O_a, \mathbb{R}^{n-k})$ be according to Def. 3.1. Then, there exists a sequence $(I_i^a)_{i \in \mathbb{N}}$ in $C$ such that $O_a = \bigcup_{i \in \mathbb{N}} I_i^a$. Thus,

$$M \cap O_a = M \cap O_a \cap \bigcup_{i \in \mathbb{N}} I_i^a.$$ 

Since $C$ is countable, so is $C_M := \{ I_i^a : a \in M, i \in \mathbb{N} \}$. Let $(I_i)_{i \in \mathbb{N}}$ be an enumeration of all elements in $C_M$. For each $i \in \mathbb{N}$, choose $a_i \in M$ such that $I_i = I_i^{a_i}$. Then

$$M = M \cap \bigcup_{i \in \mathbb{N}} I_i^a = \bigcup_{i \in \mathbb{N}} (M \cap I_i^{a_i}) = \bigcup_{i \in \mathbb{N}} (M \cap O_{a_i} \cap I_i^{a_i}) \overset{\text{Def. 3.1}[a]}{=} \bigcup_{i \in \mathbb{N}} (f^a)^{-1}(\{0\}) \cap I_i^{a_i}.$$ 

Since $f^a$ is continuous, $(f^a)^{-1}(\{0\})$ is closed. As $I_i^{a_i}$ is compact, $(f^a)^{-1}(\{0\}) \cap I_i^{a_i}$ is compact as well, proving (a).

(b) follows from (a), as each compact set is in $B^n$.

### 3.2 Metric Tensor, Gram Determinant

Before we can integrate over submanifolds, we still need to introduce the metric tensor and the Gram determinant as technical tools.

**Definition 3.7.** Let $k, n \in \mathbb{N}$ with $k < n$ and let $M \subseteq \mathbb{R}^n$ be a $k$-dimensional submanifold of class $C^1$. Moreover, let $\phi : W \rightarrow V \subseteq M$ be a local chart according to Th. 3.4(iv), i.e. $W \subseteq \mathbb{R}^k$ open, $V$ is $M$-open, and $\phi$ is a $C^1$-immersion as well as a homeomorphism. Then the matrix-valued function

$$G := G(\phi) : W \rightarrow \mathbb{R}^{k^2}, \quad G(x) := (D\phi(x))^t \, D\phi(x), \quad (3.8)$$

is called the chart’s *Gram matrix* or its *metric tensor*. If we denote the columns of $D\phi$ by $\partial_i \phi$, then we can write the components $g_{ij}$ of $G$ as follows:

$$\forall x \in W, \quad \forall i,j \in \{1, \ldots, k\}, \quad g_{ij}(x) = \langle \partial_i \phi(x), \partial_j \phi(x) \rangle = \sum_{\nu=1}^n \partial_i \phi_\nu(x) \partial_j \phi_\nu(x). \quad (3.9)$$

The map

$$\gamma := \gamma(\phi) : W \rightarrow \mathbb{R}, \quad \gamma(x) := \det G(x), \quad (3.10)$$

is called the chart’s *Gram determinant*.

**Theorem 3.8.** Let $k, n \in \mathbb{N}$ with $k < n$ and let $M \subseteq \mathbb{R}^n$ be a $k$-dimensional submanifold of class $C^1$.

(a) Let $\phi : W \rightarrow V \subseteq M$ be a local chart according to Th. 3.4(iv), i.e. $W \subseteq \mathbb{R}^k$ open, $V$ is $M$-open, and $\phi$ is a $C^1$-immersion as well as a homeomorphism. Then its Gram matrix $G = G(\phi) : W \rightarrow \mathbb{R}^{k^2}$ is symmetric positive definite and, in consequence, its Gram determinant $\gamma = \gamma(\phi) : W \rightarrow \mathbb{R}$ is actually $\mathbb{R}^+$-valued.
3. Integration Over Submanifolds of $\mathbb{R}^N$

(b) For $j \in \{1, 2\}$, let $\phi_j : O_j \rightarrow V \subseteq M$ be local charts, i.e. $O_j \subseteq \mathbb{R}^k$ open, $V$ is $M$-open, and both $\phi_j$ are $C^1$-immersions as well as homeomorphisms. Let $\gamma_j := \gamma_j(\phi_j) : O_j \rightarrow \mathbb{R}^+$ be the corresponding Gram determinants. Moreover, let $\tau = \phi_2^{-1} \circ \phi_1 : O_1 \rightarrow O_2$ be the corresponding chart transition according to Th. 3.5 (i.e. $\phi_1 = \phi_2 \circ \tau$). Then the following formula for the transition of the Gram determinant holds:

$$\forall x \in O_1 \quad \gamma_1(x) = |\det D\tau(x)|^2 \gamma_2(\tau(x)).$$  

(3.11)

Proof. (a): Let $x \in W$ and set $A := D\phi(x)$. Then $G(x) = A^tA$, i.e. $G(x)^t = (A^tA)^t = A^tA = G(x)$, showing $G(x)$ to be symmetric. If $0 \neq y \in \mathbb{R}^k$, then $y^tG(x)y = y^tA^tAy = \|Ay\|^2 > 0$, since $y \neq 0$ and $A$ injective (since $\text{rk} A = k$) imply $Ay \neq 0$. Thus, $G(x)$ is positive definite. In particular, $\gamma(x) = \det G(x) > 0$.

(b): For each $x \in O_1$, we have, by the chain rule,

$$D\phi_1(x) = D\phi_2(\tau(x))D\tau(x).$$

Thus,

$$G(\phi_1)(x) = (D\phi_2(\tau(x))D\tau(x))^t(D\phi_2(\tau(x))D\tau(x)) = (D\tau(x))^tG(\phi_2)(\tau(x))D\tau(x),$$

implying (3.11).  

3.3 Integration Over Submanifolds

If $M \subseteq \mathbb{R}^n$ is a $k$-dimensional submanifold, then the goal is to integrate $f : M \rightarrow \mathbb{R}$ over $M$. Since, locally, $M$ “looks like” an open set of $\mathbb{R}^k$, we would like to integrate with respect to a measure that, locally, “looks like” $\beta^k$ (or $\lambda^k$) on $\mathbb{R}^k$. The natural idea is to transport the measure from $\mathbb{R}^k$ to $M$ via local charts. While this does work, some subtleties arise: One has to make sure that the measure does not depend on the choice of local chart. To achieve this, the Gram determinant of the previous section comes into play. The other issue one has to deal with are overlapping charts ($M$ might not be coverable with just one chart), and here one makes use of a so-called partition of unity, see Def. 3.9(b) below. For simplicity, we will only consider submanifolds that are coverable by finitely many local charts. The treatment in [Kön04, Sec. 11.4, 11.5] shows that this assumption is not really necessary.

**Definition 3.9.** Let $k, n \in \mathbb{N}$ with $k < n$ and let $M \subseteq \mathbb{R}^n$ be a $k$-dimensional submanifold of class $C^1$.

(a) A family $(\phi_i)_{i \in I}$ of local charts $\phi_i : W_i \rightarrow M_i \subseteq M$ is called a (finite) cover of $M$ if, and only if, $M = \bigcup_{i \in I} M_i$ (with $I$ finite). Moreover, $M$ is called *finitely coverable* if, and only if, there exists a finite cover of local charts for $M$. 

Given a finite cover $\mathcal{C} := (\phi_1, \ldots, \phi_N)$, $N \in \mathbb{N}$, of local charts for $M$, a family of functions $(\eta_1, \ldots, \eta_N)$, $\eta_i : M \rightarrow \mathbb{R}$, is called a partition of unity subordinate to $\mathcal{C}$ if, and only if, the following conditions are satisfied:

(i) $0 \leq \eta_i \leq 1$ for each $i \in \{1, \ldots, N\}$.

(ii) $\eta_i|_{M \setminus M_i} \equiv 0$ for each $i \in \{1, \ldots, N\}$.

(iii) $\sum_{i=1}^{N} \eta_i \equiv 1$.

We call the partition of unity measurable if, and only if, each function $\eta_i \circ \phi$ is $\lambda^k$-measurable.

Let $\mathcal{C} := (\phi_1, \ldots, \phi_N)$, $N \in \mathbb{N}$, be a finite cover of $M$ via local charts. A function $f : M \rightarrow \mathbb{K}$ is called $\sigma$-measurable (or merely measurable if the context is clear) if, and only if, $(f \circ \phi_i)$ is $\lambda^k$-measurable for each $i \in \{1, \ldots, N\}$. If $f : M \rightarrow [0, \infty]$ is measurable and $(\eta_1, \ldots, \eta_N)$ constitutes a measurable partition of unity subordinate to $\mathcal{C}$, then we define the integral

$$\int_{M} f \, d\sigma := \sum_{i=1}^{N} \int_{W_i} ((\eta_i f) \circ \phi_i) \sqrt{\gamma(\phi_i)} \, d\lambda^k,$$

where $\gamma(\phi_i)$ denotes the Gram determinant of $\phi_i$ (recall that $\gamma(\phi_i)$ is $\mathbb{R}^+\text{-valued}$). If $A \subseteq M$ and $\chi_A$ is $\sigma$-measurable, then we call $A$ $\sigma$-measurable and define

$$\sigma(A) := \int_{M} \chi_A \, d\sigma.$$

We call $f : M \rightarrow \mathbb{K}$ $\sigma$-integrable if, and only if, $f$ is $\sigma$-measurable and $\int_{M} g \, d\sigma < \infty$ for each $g \in \{(\text{Re } f)^+, (\text{Re } f)^-, (\text{Im } f)^+, (\text{Im } f)^-\}$. In that case, we define

$$\int_{M} f \, d\sigma := \int_{M} (\text{Re } f)^+ \, d\sigma - \int_{M} (\text{Re } f)^- \, d\sigma + i \int_{M} (\text{Im } f)^+ \, d\sigma - i \int_{M} (\text{Im } f)^- \, d\sigma.$$

**Example 3.10.** Let $k, n \in \mathbb{N}$ with $k < n$ and let $M \subseteq \mathbb{R}^n$ be a $k$-dimensional submanifold of class $C^1$.

(a) If $M$ is the homeomorphic image of a $C^1$-immersion $\phi : W \rightarrow M$, $W \subseteq \mathbb{R}^k$, then $M$ is finitely coverable (coverable by the single local chart $\phi$). Special cases are now considered in (b) and (c).

(b) Let $I \subseteq \mathbb{R}$ be a nonempty open interval and let $\phi : I \rightarrow \mathbb{R}^n$, $n \geq 2$, be $C^1$ with $D\phi(x) \neq 0$ for each $x \in I$. If $\phi : I \rightarrow \phi(I) = M$ is a homeomorphism, then $M$ is a $1$-dimensional submanifold (i.e. a curve) coverable by the single local chart $\phi$. Here, since $D\phi(x)$ is represented by a column vector, the corresponding Gram matrix is $1 \times 1$ and its entry as well as its determinant is given by the squared Euclidean norm of $D\phi(x)$:

$$\gamma(\phi) : I \rightarrow \mathbb{R}^+, \quad \gamma(\phi)(x) = \|D\phi(x)\|_2^2 = \|\phi'(x)\|_2^2 = \sum_{i=1}^{n} |\phi_i'(x)|^2.$$
(c) Graphs of $C^1$-functions are submanifolds coverable by a single local chart: Let $A \subseteq \mathbb{R}^k$ be open and let $g : A \rightarrow \mathbb{R}^{n-k}$ be $C^1$. Then the $k$-dimensional submanifold

$$M = \{(x, y) \in A \times \mathbb{R}^{n-k} : y = g(x)\}$$

is coverable by the single local chart

$$\phi : A \rightarrow M, \quad \phi(x) := (x, g(x)).$$

Thus, the Gram matrix is

$$G(\phi) : A \rightarrow \mathbb{R}^{k^2}, \quad G(\phi)(x) = \text{Id}_k + (Dg(x))^t Dg(x).$$

For $k = n - 1$, this simplifies to

$$G(\phi) : A \rightarrow \mathbb{R}^{k^2}, \quad G(\phi)(x) = \text{Id}_k + (\nabla g(x))^t \nabla g(x),$$

and one can show (exercise) that, in this case, the Gram determinant is

$$\gamma(\phi) : A \rightarrow \mathbb{R}^+, \quad \gamma(\phi)(x) = 1 + \| \nabla g(x) \|^2_2 = 1 + \sum_{i=1}^{n-1} |\partial_i g(x)|^2. \quad (3.16)$$

(d) If $M$ is compact, then $M$ is finitely coverable: Indeed, for each $a \in M$, let $\phi^a : W_a \rightarrow \mathbb{R}^n$ be a local chart according to Th. 3.4(iv). Then $(M_a)_{a \in M}$ is an open cover of $M$. As $M$ is compact, it must have a finite subcover, showing $M$ to be finitely coverable.

**Theorem 3.11.** Let $k, n \in \mathbb{N}$ with $k < n$ and let $M \subseteq \mathbb{R}^n$ be a $k$-dimensional submanifold of class $C^1$. Assume $M$ to be finitely coverable.

(a) Given a finite cover $C := (\phi_1, \ldots, \phi_N), N \in \mathbb{N}$, of local charts for $M$, there exists a measurable partition of unity subordinate to $C$.

(b) The measurability of $f : M \rightarrow \mathbb{K}$, as defined in Def. 3.9(c), does not depend on the chosen local charts covering $M$. The integral defined in Def. 3.9(c) is well-defined, i.e. the integrals on the right-hand side of (3.12) all make sense and the value of $\int_M f \, d\sigma$ does neither depend on the choice of covering local charts nor on the choice of partition of unity. In consequence, the integral defined in (3.14) is well-defined as well.

(c) The set $A_\sigma := \{A \subseteq M : A \text{ $\sigma$-measurable}\}$ constitutes a $\sigma$-algebra on $M$; $\sigma : A \rightarrow [0, \infty]$ defined by (3.13) constitutes a measure on $M$, called the surface measure. Then $\sigma$-measurability (resp. $\sigma$-integrability) as defined in Def. 3.9(c) coincides with the usual measurability (resp. integrability) with respect to the measure $\sigma$; the integral defined in Def. 3.9(c) is the usual integral with respect to the measure $\sigma$. 
Proof. (a): If $\phi_i : W_i \rightarrow M_i, i \in \{1, \ldots, N\}$, then define

$$
\forall i \in \{1, \ldots, N\}, \quad \eta_i : M \rightarrow \mathbb{R}, \quad \eta_i := \chi_{B_i}, \quad B_i := M \setminus \bigcup_{j=1}^{i-1} M_j.
$$

(3.17)

Then Def. 3.9(b)(i)–(iii) are, clearly, satisfied ((iii), since $M = \bigcup_{i=1}^{N} M_i$). It remains to show that the $\eta_i$ are measurable, i.e. that the $B_i$ are measurable. As the $M_i$ are $M$-open, there exists $O_i \subseteq \mathbb{R}^n$ open such that $M_i = M \cap O_i$. Since $M \in \mathcal{B}^n$ by Th. 3.6(b), $M_i \in \mathcal{B}^n$, showing $B_i \in \mathcal{B}^n$ as well.

(b): If $f : M \rightarrow [0, \infty]$, then all integrands on the right-hand side of (3.12) are nonnegative. If $f$ is $\sigma$-measurable, then the integrands are $\lambda^k$-measurable as they are the product of the $\lambda^k$-measurable functions $f \circ \phi_i$, $\eta_i \circ \phi_i$, and $x \mapsto \sqrt{\gamma(\phi_i)(x)}$, the last function even being continuous. It remains to show the independence of the chosen charts and partition of unity. Thus, in addition to the chart cover $\mathcal{C} = (\phi_1, \ldots, \phi_N)$, $\phi_i : W_i \rightarrow M_i$ and the $\eta_i$, let $\mathcal{D} := (\psi_1, \ldots, \psi_M)$, $M \in \mathbb{N}$, also be a finite cover of $M$ via local charts, $\psi_i : U_i \rightarrow N_i$, where $(\alpha_1, \ldots, \alpha_M)$ constitutes a measurable partition of unity subordinate to $\mathcal{D}$. Let $V_{ij} := M_i \cap N_j$ and $K := \{(i, j) : V_{ij} \neq \emptyset\}$. According to Th. 3.5, for each $(i, j) \in K$, there exists a $C^1$-diffeomorphism $\tau_{ij} : U_{ij} \rightarrow W_{ij}$ such that $\psi_j = \phi_i \circ \tau_{ij}$ on $U_{ij}$, where $U_{ij} := \psi_j^{-1}(V_{ij})$, $W_{ij} := \phi_j^{-1}(V_{ij})$. Then, for each $f : M \rightarrow \mathbb{K}$ and $j \in \{1, \ldots, M\}$,

$$
f \circ \psi_j = \sum_{i=1}^{N} (\eta_i f) \circ \psi_j = \sum_{i : (i,j) \in K} (\eta_i f) \circ \phi_i \circ \tau_{ij}.
$$

Since each $\tau_{ij}$ is $\lambda^k$-measurable, if $f$ is $\sigma$-measurable, then each summand is $\lambda^k$-measurable, showing $f \circ \psi_j$ to be $\lambda^k$-measurable as well. In consequence, the $\sigma$-measurability of $f$ does not depend on the chosen charts. Now let $f : M \rightarrow [0, \infty]$ be $\sigma$-measurable. One computes

$$
\sum_{j=1}^{M} \int_{U_j} ((\alpha_j f) \circ \psi_j) \sqrt{\gamma(\psi_j)} \, d\lambda^k = \sum_{j=1}^{M} \sum_{i=1}^{N} \int_{U_j} ((\eta_i \alpha_j f) \circ \psi_j) \sqrt{\gamma(\psi_j)} \, d\lambda^k
$$

(3.11)

$$
= \sum_{(i,j) \in K} \int_{U_{ij}} ((\eta_i \alpha_j f) \circ \psi_j) \sqrt{\gamma(\psi_j)} \, d\lambda^k
$$

(3.72)

$$
= \sum_{(i,j) \in K} \int_{W_{ij}} ((\eta_i \alpha_j f) \circ \phi_i) \sqrt{\gamma(\phi_i)} \, d\lambda^k
$$

$$
= \sum_{i=1}^{N} \sum_{j=1}^{M} \int_{W_{ij}} ((\eta_i \alpha_j f) \circ \phi_i) \sqrt{\gamma(\phi_i)} \, d\lambda^k = \sum_{i=1}^{N} \int_{W_{ij}} ((\eta_i f) \circ \phi_i) \sqrt{\gamma(\phi_i)} \, d\lambda^k,
$$

showing the integral in (3.12) to be well-defined.
(c): Let \( \phi_1, \ldots, \phi_N \) be as in the proof of (b). Let \( A \subseteq M \). Then \( A \in \mathcal{A}_\sigma \) if, and only if, \( \phi_i^{-1}(A) \in \mathcal{L}^k \) for each \( i \in \{1, \ldots, N\} \). Thus,

\[
\mathcal{A}_\sigma = \bigoplus_{i=1}^N \mathcal{B}_i, \quad \mathcal{B}_i := \{ B \in M : \phi_i^{-1}(B) \in \mathcal{L}^k \}.
\] (3.18)

Since each \( \mathcal{B}_i \) is a \( \sigma \)-algebra by Prop. 1.55(a), \( \mathcal{A}_\sigma \) is a \( \sigma \)-algebra as well. We show \( \sigma \) to be a measure on \( \mathcal{A}_\sigma \): Clearly, \( \sigma(\emptyset) = 0 \). Let \( (A_i)_{i \in \mathbb{N}} \) be a sequence of disjoint sets in \( \mathcal{A}_\sigma \) and let \( A := \bigcup_{j \in \mathbb{N}} A_j \). Then, for each \( i \in \{1, \ldots, N\} \),

\[
\int_{W_i} ((\eta_k \chi_A) \circ \phi_i) \sqrt{\gamma(\phi_i)} \, d\lambda^k = \int_{\phi_i^{-1}(A)} \eta_k \phi_i \sqrt{\gamma(\phi_i)} \, d\lambda^k
\]

\[= \sum_{j=1}^\infty \int_{\phi_i^{-1}(A_j)} (\eta_k \phi_i) \sqrt{\gamma(\phi_i)} \, d\lambda^k = \sum_{j=1}^\infty \int_{W_i} (\eta_k \chi_{A_j} \circ \phi_i) \sqrt{\gamma(\phi_i)} \, d\lambda^k,
\]

implying

\[
\sigma(A) = \sum_{j=1}^\infty \sigma(A_j),
\]

i.e. \( \sigma \) is \( \sigma \)-additive and a measure. Next, we show \( f : M \rightarrow \mathbb{K} \) to be \( \sigma \)-measurable as defined in Def. 3.9(c) if, and only if, \( f \) is \( \mathcal{A}_\sigma \)-measurable, i.e. if, and only if, \( f^{-1}(B) \in \mathcal{A}_\sigma \) for each \( B \in \mathcal{B} \):

\[
f \sigma \text{-measurable } \iff \forall_{i \in \{1, \ldots, N\}} f \circ \phi_i \text{ } \lambda^k \text{-measurable }
\]

\[
\iff \forall_{B \in \mathcal{B}} \forall_{i \in \{1, \ldots, N\}} (f \circ \phi_i)^{-1}(B) = \phi_i^{-1}(f^{-1}(B)) \in \mathcal{L}^k
\]

\[
\iff \forall_{B \in \mathcal{B}} f^{-1}(B) \in \mathcal{A}_\sigma \iff f \mathcal{A}_\sigma \text{-measurable.}
\]

It is then immediate from Def. 3.9(c) that the respective notions of integrability are also identical. It remains to show that the integral defined in Def. 3.9(c) is the usual integral with respect to the measure \( \sigma \). To this end, for the remainder of the proof, denote the measure \( \sigma \) by \( \lambda^k \) (to notationally distinguish the two integrals – the notation is also reasonable as one may interpret the measure \( \sigma \) as a \( k \)-dimensional Lebesgue measure on \( M \)). Let \( f \in \mathcal{S}^+(\mathcal{A}_\sigma) \) be a simple function. Then there exist \( A_1, \ldots, A_m \in \mathcal{A}_\sigma \) and \( s_1, \ldots, s_m \in \mathbb{R}_0^+ \) such that \( f = \sum_{j=1}^m s_j \chi_{A_j} \). Thus,

\[
\int_M f \, d\sigma = \sum_{i=1}^N \int_{W_i} ((\eta_k f) \circ \phi_i) \sqrt{\gamma(\phi_i)} \, d\lambda^k = \sum_{i=1}^N \sum_{j=1}^m s_j \int_{W_i} ((\eta_k \chi_{A_j}) \circ \phi_i) \sqrt{\gamma(\phi_i)} \, d\lambda^k
\]

\[= \sum_{j=1}^m s_j \int_M \chi_{A_j} \, d\sigma \overset{(3.13)}{=} \sum_{j=1}^m s_j \sigma(A_j) = \sum_{j=1}^m s_j \lambda^k(A_j) = \int_M f \, d\lambda^k.
\]
Now let \( f \in \mathcal{M}^+(A_\sigma) \) and let \((f_j)_{j \in \mathbb{N}}\) be a sequence in \( \mathcal{S}^+(A_\sigma) \) such that \( f_j \uparrow f \). Then, using the monotone convergence Th. 2.7,

\[
\int_M f \, d\sigma = \sum_{i=1}^N \int_{W_i} ((\eta_i f_i) \circ \phi_i) \sqrt{\gamma(\phi_i)} \, d\lambda^k = \lim_{j \to \infty} \sum_{i=1}^N \int_{W_i} ((\eta_i f_j) \circ \phi_i) \sqrt{\gamma(\phi_i)} \, d\lambda^k
= \lim_{j \to \infty} \int_M f_j \, d\lambda^k = \int_M f \, d\lambda^k.
\]

Finally, \( \int_M f \, d\sigma = \int_M f \, d\lambda^k \) for integrable \( f : M \to \mathbb{K} \) follows in the usual way by decomposing \( f \) into \((\text{Re } f)^\pm, (\text{Im } f)^\pm\).

\[\textbf{Remark 3.12.}\] The main benefit of Th. 3.11(c) is that we have identified the integrals \( \int_M f \, d\sigma \) as measure-theoretic integrals in the sense of Sec. 2. In consequence all the abstract results of Sec. 2 regarding such integrals must hold and are available for integrals over submanifolds.

\[\textbf{Example 3.13. (a)}\] As in Ex. 3.10(b), let \( I \subseteq \mathbb{R} \) be a nonempty open interval and let \( \phi : I \to \mathbb{R}^n \), \( n \geq 2 \), be \( C^1 \) with \( D\phi(x) \neq 0 \) for each \( x \in I \) and such that \( \phi : I \to \phi(I) \) is a homeomorphism. We know from Ex. 3.10(b) that \( M := \phi(I) \) is a 1-dimensional submanifold covered by the single chart \( \phi \) and that

\[
\gamma(\phi) : I \to \mathbb{R}^+, \quad \gamma(\phi)(x) = \|\phi'(x)\|^2_2.
\]

Thus, if \( f : M \to \mathbb{K} \) is nonnegative measurable or integrable, then

\[
\int_M f \, d\sigma = \int_I f(\phi(x)) \|\phi'(x)\|_2 \, d\lambda^1(x).
\]

In particular,

\[
l(\phi) := \sigma(M) = \int_M 1 \, d\sigma = \int_I \|\phi'(x)\|_2 \, d\lambda^1(x)
\]

is called the \textit{arc length} of \( \phi \) (or \( M \)). Moreover, \( \phi \) and \( M \) are called \textit{rectifiable} if, and only if, \( l(\phi) < \infty \). We also extend these notions to \( C^1 \)-paths \( \phi : I \to \mathbb{R}^n \) on general (not necessarily open) intervals by setting \( l(\phi) := l(\phi|_{I^o}) \) (where \( I^o \) denotes the interval’s interior). It is then clear from (3.20) that \( C^1 \)-paths defined on \textit{compact} intervals are always rectifiable.

\[\textbf{(b)}\] Let \( x_0 \in [0, 2\pi] \) and consider

\[
\phi : [0, x_0] \to \mathbb{R}^2, \quad \phi(x) := (\cos x, \sin x),
\quad \phi' : [0, x_0] \to \mathbb{R}^2, \quad \phi'(x) = (-\sin x, \cos x) \neq (0, 0).
\]

Then \( \|\phi'(x)\|_2 = \sqrt{(\sin x)^2 + (\cos x)^2} = 1 \), implying

\[
l(\phi) = \int_{[0, x_0]} \|\phi'(x)\|_2 \, d\lambda^1(x) = \int_0^{x_0} 1 \, dx = x_0.
\]
This finally proves that the definition of sine and cosine as given in [Phi16a, Def. and Rem. 8.21] is the same as the high school definition, where \( \cos x \) and \( \sin x \) are defined to be the coordinates of the point \( p = (p_1, p_2) \in \mathbb{R}^2 \) on the unit circle, such that \( x \) is the angle measured in radian (i.e. the arc length of the corresponding circle section as computed above) between the line segment between \((0, 0)\) and \((1, 0)\) and the line segment between \((0, 0)\) and \(p\) (cf. remarks at the beginning of [Phi16a, Sec. 8.4]).

**Example 3.14.** (a) Let \( n \in \mathbb{N}, n \geq 2. \) As in Ex. 3.10(c) (with \( k := n - 1 \)) let \( A \subseteq \mathbb{R}^{n-1} \) be open and let \( g : A \longrightarrow \mathbb{R} \) be \( C^1 \). We know from Ex. 3.10(c) that the \((n-1)\)-dimensional submanifold

\[
M := \{(x, y) \in A \times \mathbb{R} : y = g(x)\}
\]

is coverable by the single local chart

\[
\phi : A \longrightarrow M, \quad \phi(x) := (x, g(x))
\]

and that

\[
\gamma(\phi) : A \longrightarrow \mathbb{R}^+, \quad \gamma(\phi)(x) = 1 + \| \nabla g(x) \|^2 = 1 + \sum_{i=1}^{n-1} |\partial_i g(x)|^2.
\]

Thus, if \( f : M \longrightarrow \hat{K} \) is nonnegative measurable or integrable, then

\[
\int_M f \, d\sigma = \int_A f(x, g(x)) \sqrt{1 + \| \nabla g(x) \|^2} \, d\lambda^{n-1}(x). \tag{3.21}
\]

In particular,

\[
\sigma(M) = \int_M 1 \, d\sigma = \int_A \sqrt{1 + \| \nabla g(x) \|^2} \, d\lambda^{n-1}(x). \tag{3.22}
\]

(b) Let \( n \in \mathbb{N}, n \geq 2. \) As a concrete example of the situation in (a), we consider the \((n-1)\)-dimensional upper hemisphere with center 0 and radius \( r > 0 \):

\[
M := \{x \in \mathbb{R}^n : \|x\|_2 = r, x_n > 0\}. \tag{3.23}
\]

Letting \( A := \{x \in \mathbb{R}^{n-1} : \|x\|_2 < r\} \), we obtain \( M \) as the graph of the \( C^1 \)-function

\[
g : A \longrightarrow \mathbb{R}, \quad g(x_1, \ldots, x_{n-1}) := \sqrt{r^2 - x_1^2 - \cdots - x_{n-1}^2}, \tag{3.24}
\]

where

\[
\forall_{i \in \{1, \ldots, n-1\}} \partial_i g(x) = -\frac{x_i}{g(x)}. \tag{3.25}
\]

We have \( M = \phi(A) \) if \( \phi \) is defined as in (a). Thus, computing \( \gamma(\phi) \) according to (a) yields

\[
\gamma(\phi) : A \longrightarrow \mathbb{R}^+, \quad \gamma(\phi)(x) = 1 + \sum_{i=1}^{n-1} \frac{x_i^2}{(g(x))^2} = 1 + \frac{\|x\|^2}{(g(x))^2} = \frac{r^2}{(g(x))^2}
\]
and, if \( f : M \rightarrow \mathbb{K} \) is nonnegative measurable or integrable, then, using balls with respect to the 2-norm and the (linear) change of variables \( T : \mathbb{R}^{n-1} \rightarrow \mathbb{R}^{n-1}, \)

\[
\int_M f \, d\sigma = \int_A f(x, g(x)) \sqrt{1 + \| \nabla g(x) \|^2} \, d\lambda^{n-1}(x)
\]

\[
= \int_{B_r(0)} f \left( x, \sqrt{r^2 - \|x\|^2} \right) \frac{r}{\sqrt{r^2 - \|x\|^2}} \, d\lambda^{n-1}(x)
\]

\[
= \int_{B_r(0)} f \left( rt, \sqrt{1 - \|t\|^2} \right) \frac{r^{n-1}}{\sqrt{1 - \|t\|^2}} \, d\lambda^{n-1}(t). \quad (3.26)
\]

**Theorem 3.15.** Let \( k, n \in \mathbb{N} \) with \( k < n \), \( \alpha \in \mathbb{N} \cup \{ \infty \} \), and let \( M \subseteq \mathbb{R}^n \) be a \( k \)-dimensional submanifold of class \( C^\alpha \). Given \( r \in \mathbb{R}^+ \), consider the linear isomorphism

\[
T_r : \mathbb{R}^n \rightarrow \mathbb{R}^n, \quad T_r(x) := rx.
\]

Then \( T_r(M) \) is also a \( k \)-dimensional submanifold of class \( C^\alpha \). If \( M \) is finitely coverable, then so is \( T_r(M) \), and \( f : T_r(M) \rightarrow \mathbb{K} \) is nonnegative \( \sigma \)-measurable (resp. \( \sigma \)-integrable) if, and only if, \( (f \circ T_r) : M \rightarrow \mathbb{K} \) has the corresponding property. In that case,

\[
\int_{T_r(M)} f \, d\sigma = \int_M (f \circ T_r) r^k \, d\sigma. \quad (3.28)
\]

**Proof.** If the \( C^\alpha \)-immersion \( \phi : W \rightarrow \mathbb{R}^n, W \subseteq \mathbb{R}^k \), is a local chart for \( M \), then the \( C^\alpha \)-immersion \( \psi := T_r \circ \phi : W \rightarrow \mathbb{R}^n \) is a local chart for \( T_r(M) \). Moreover, if \( \mathcal{C} := (\phi_1, \ldots, \phi_N), N \in \mathbb{N}, \) is a finite cover of \( M \) via local charts and \( (\eta_1, \ldots, \eta_N) \), \( \eta_i : M \rightarrow \mathbb{R} \), is a measurable partition of unity subordinate to \( \mathcal{C} \), then \( \mathcal{D} := (T_r \circ \phi_1, \ldots, T_r \circ \phi_N), \) is a finite cover via local charts of \( T_r(M), (\eta_1 \circ T_r^{-1}, \ldots, \eta_N \circ T_r^{-1}), \eta_i \circ T_r^{-1} : T_r(M) \rightarrow \mathbb{R}, \) is a measurable partition of unity subordinate to \( \mathcal{D} \). The \( \sigma \)-measurability of both \( f \) and \( f \circ T_r \) is then equivalent to the \( \lambda^k \)-measurability of each \( f \circ T_r \circ \phi_\nu \). Let \( g_{\nu ij} \) and \( \tilde{g}_{\nu ij} \) denote the entries of the Gram matrices \( G(\phi_\nu) \) and \( G(T_r \circ \phi_\nu) \), respectively. If \( \phi_\nu : W_\nu \rightarrow \mathbb{R}^n \), then

\[
\forall \ x \in W_\nu \quad \tilde{g}_{\nu ij} = \langle \partial_i(T_r \circ \phi_\nu)(x), \partial_j(T_r \circ \phi_\nu)(x) \rangle = \langle r \partial_i \phi_\nu(x), r \partial_j \phi_\nu(x) \rangle = r^2 g_{\nu ij},
\]

showing

\[
\forall \ x \in W_\nu \quad \gamma(T_r \circ \phi_\nu)(x) = r^{2k} \gamma(\phi_\nu)(x).
\]

Thus, if \( f \) is nonnegative \( \sigma \)-measurable, then

\[
\int_{T_r(M)} f \, d\sigma = \sum_{i=1}^N \int_{W_i} \left( (\eta_i \circ T_r^{-1}) f \circ T_r \circ \phi_i \right) \sqrt{\gamma(T_r \circ \phi_i)} \, d\lambda^k
\]

\[
= \sum_{i=1}^N \int_{W_i} \left( (\eta_i(f \circ T_r)) \circ \phi_i \right) r^k \sqrt{\gamma(\phi_i)} \, d\lambda^k = \int_M (f \circ T_r) r^k \, d\sigma,
\]

proving (3.28). The integrable case now follows in the usual way by decomposition of \( f \) into \( (\text{Re} \, f) \pm \) and \( (\text{Im} \, f) \pm \).
Theorem 3.16. Let \( n \in \mathbb{N} \), \( n \geq 2 \). Let \( f : \mathbb{R}^n \rightarrow \mathbb{H} \) be \( \lambda^n \)-integrable. Then, for \( \lambda^1 \)-a.e. \( r \in \mathbb{R}^+, \) \( f \) is \( \sigma \)-integrable over the sphere \( S_r(0) = \{ x \in \mathbb{R}^n : \| x \|_2 = r \} \) and
\[
\int_{\mathbb{R}^n} f \, d\lambda^n = \int_0^\infty \int_{S_r(0)} f(x) \, d\sigma(x) \, d\lambda^1(r) \quad \text{(3.28)}
\]
\\nThis formula is sometimes called the coarea formula. It is actually a special case of the more general coarea formula of geometric measure theory.

Proof. Denote the (open) upper and lower half spaces by \( H_\pm := \{ x \in \mathbb{R}^n : \pm x_n > 0 \} \) and \( U := \{ x \in \mathbb{R}^{n-1} : \| x \|_2 < 1 \} \). We will apply the change of variables given by the \( C^1 \)-diffeomorphism
\[
\phi : U \times \mathbb{R}^+ \rightarrow H_+, \quad \phi(x_1, \ldots, x_{n-1}, r) := \left( rx_1, \ldots, rx_{n-1}, r \sqrt{1 - \| x \|_2^2} \right),
\]
\[
\phi^{-1} : H_+ \rightarrow U \times \mathbb{R}^+, \quad \phi^{-1}(x_1, \ldots, x_n) := \left( \frac{x_1}{\| x \|_2}, \ldots, \frac{x_{n-1}}{\| x \|_2}, \| x \|_2 \right),
\]
\[
D\phi : U \times \mathbb{R}^+ \rightarrow H_+,
\]
\[
D\phi(x_1, \ldots, x_{n-1}, r) := \begin{pmatrix}
r & 0 & \cdots & 0 & x_1 \\
0 & r & \cdots & 0 & x_2 \\
\vdots & \ddots & \ddots & \vdots & \vdots \\
0 & \cdots & 0 & r & x_{n-1}
\end{pmatrix}
- \frac{rx_1}{\sqrt{1 - \| x \|_2^2}} \quad - \frac{rx_2}{\sqrt{1 - \| x \|_2^2}} \quad \cdots \quad - \frac{rx_{n-1}}{\sqrt{1 - \| x \|_2^2}} \quad \sqrt{1 - \| x \|_2^2}
\]

We compute \( \det D\phi(x, r) \) via eliminating the nondiagonal elements of the last row by successively adding, for \( i = 1, \ldots, n-1 \), the \( i \)-th row multiplied by \(-\frac{x_i}{\sqrt{1 - \| x \|_2^2}}\) to the last row, ending up with an upper triangular matrix, where the upper \( n-1 \) diagonal elements are \( r \) and the last diagonal element is
\[
\sqrt{1 - \| x \|_2^2} - \sum_{i=1}^{n-1} \frac{x_i^2}{\sqrt{1 - \| x \|_2^2}} = \sqrt{1 - \| x \|_2^2} - \frac{\| x \|_2^2}{\sqrt{1 - \| x \|_2^2}} = \frac{1}{\sqrt{1 - \| x \|_2^2}}
\]

Thus,
\[
\det D\phi : U \times \mathbb{R}^+ \rightarrow \mathbb{R}^+, \quad \det D\phi(x, r) = \frac{r^{n-1}}{\sqrt{1 - \| x \|_2^2}}
\]

We now apply the change of variables formula (2.72) followed by Fubini to the integrable function \( f \) to obtain
\[
\int_{H_+} f \, d\lambda^n = \int_0^\infty \int_U f \left( rx, r \sqrt{1 - \| x \|_2^2} \right) \frac{r^{n-1}}{\sqrt{1 - \| x \|_2^2}} \, d\lambda^{n-1}(x) \, d\lambda^1(r)
\]
\[
\overset{(3.26)}{=} \int_0^\infty \int_{H_+ \cap S_r(0)} f \, d\sigma \, d\lambda^1(r). \quad (3.30)
\]
As \( f \) is integrable, Fubini also yields that the inner integral must be finite \( \lambda^1 \)-a.e. as claimed. Next, we note that the result of (3.30) still holds with \( H_{+} \) replaced by \( H_{-} \): For \( H_{-} \), one merely puts a negative sign in front of the last component of \( \phi \), then \( |\det D\phi| \) remains the same as before and in the middle term of (3.30), one has a “\(-\)” in the second argument of \( f \), which is consistent with what one obtains if one does Ex. 3.14(b) for the lower hemisphere (i.e. with \( g \) replaced by \( -g \)). Since \( S_{r}(0) \cap \{ x \in \mathbb{R}^n : x_n = 0 \} \) is a \( \sigma \)-null set (exercise), the proof is complete. \( \square \)

Example 3.17. Let \( n \in \mathbb{N}, n \geq 2 \). We would like to compute the surface measure of the Euclidean sphere \( S_{r}(0) = \{ x \in \mathbb{R}^n : \|x\|_2 = r \}, r \in \mathbb{R}^+ \). We recall the formula for the corresponding Euclidean ball \( B_{r}(0) \) from Ex. 2.34(c):

\[
\beta^n(B_{r}(0)) = \omega_n r^n, \quad \omega_n = \beta^n(B_{1}(0)) = \frac{\pi_{n/2}}{\Gamma\left(\frac{n}{2} + 1\right)}.
\]

Applying (3.29) to \( f := \chi_{B_{1}(0)} \) yields

\[
\omega_n = \beta^n(B_{1}(0)) = \int_{\mathbb{R}^n} \chi_{B_{1}(0)} \, d\lambda^n \overset{(3.29)}{=} \int_{0}^{1} \left( \int_{S_{1}(0)} \right. 1 \, d\sigma \left. \right) r^{n-1} \, d\lambda^1(r) = \frac{\sigma(S_{1}(0))}{n},
\]

implying

\[
\sigma_n := \sigma(S_{1}(0)) = n \omega_n = \frac{n \pi_{n/2}}{\Gamma\left(\frac{n}{2} + 1\right)} = \frac{2 \pi_{n/2}}{\Gamma\left(\frac{n}{2}\right)}.
\]

An application of Th. 3.15 now yields the formula for \( \sigma(S_{r}(0)) = \sigma(T_{r}(S_{1}(0))) \):

\[
\sigma(S_{r}(0)) = \int_{T_{r}(S_{1}(0))} 1 \, d\sigma \overset{(3.28)}{=} \int_{S_{1}(0)} r^{n-1} \, d\sigma = \sigma_n r^{n-1}. \quad (3.32)
\]

Example 3.18. We apply Th. 3.16 to rotationally symmetric functions: Let \( f : \mathbb{R}_0^+ \rightarrow \mathbb{K} \) be such that

\[
g : \mathbb{R}^n \rightarrow \mathbb{K}, \quad g(x) := f(\|x\|_2),
\]

is nonnegative measurable or integrable. Then

\[
\int_{\mathbb{R}^n} g \, d\lambda^n = \int_{\mathbb{R}^n} f(\|x\|_2) \, d\lambda^n (x) = \int_{0}^{\infty} \left( \int_{S_{1}(0)} f(r) \, d\sigma (t) \right) r^{n-1} \, d\lambda^1(r) = \sigma_n \int_{0}^{\infty} f(r) r^{n-1} \, d\lambda^1(r).
\]

\[\overset{(3.31)}{=} \quad (3.33)\]

### 3.4 Tangent Space and Normal Space

A tangent vector to a submanifold \( M \) is a vector that is tangent to a curve in \( M \). At each point \( a \in M \), the set of all tangent vectors forms a vector space, the so-called tangent space \( T_{a}M \). Normal vectors at \( a \) are vectors that are perpendicular to the tangent space \( T_{a}M \):
Definition 3.19. Let \( k, n \in \mathbb{N} \) with \( k < n \), and let \( M \subseteq \mathbb{R}^n \) be a \( k \)-dimensional submanifold of class \( C^1 \), \( a \in M \).

(a) A vector \( \tau \in \mathbb{R}^n \) is called a tangent vector to \( M \) at \( a \) if, and only if, there exists \( \epsilon \in \mathbb{R}^+ \) and a \( C^1 \)-function

\[
\psi : ] - \epsilon, \epsilon [ \rightarrow M,
\]

such that

\[
\psi(0) = a \quad \text{and} \quad \psi'(0) = \tau.
\]

The set (we will see in the next theorem it is, indeed, a \( k \)-dimensional vector space over \( \mathbb{R} \))

\[
T_a M := \{ \tau \in \mathbb{R}^n : \tau \text{ is tangent vector to } M \text{ at } a \} \tag{3.35}
\]

is called the tangent space to \( M \) at \( a \).

(b) A vector \( \nu \in \mathbb{R}^n \) is called a normal vector to \( M \) at \( a \) if, and only if, it is perpendicular to all tangent vectors at \( a \) (with respect to the Euclidean scalar product), i.e. if,

\[
\forall \tau \in T_a M \quad \langle \tau, \nu \rangle = 0. \tag{3.36}
\]

We also define

\[
N_a M := \{ \nu \in \mathbb{R}^n : \nu \text{ is normal vector to } M \text{ at } a \} \tag{3.37}
\]

to be the normal space to \( M \) at \( a \).

Definition and Remark 3.20. Let \( n, k \in \mathbb{N}_0 \) and let \( V \subseteq \mathbb{R}^n \) be a vector subspace of \( \mathbb{R}^n \), \( \dim V = k \). Then

\[
V^\perp := \{ x \in \mathbb{R}^n : \langle x, v \rangle = 0 \text{ for each } v \in V \} \tag{3.38}
\]

is the space perpendicular to \( V \). We then know from Linear Algebra that \( V^\perp \) is a vector subspace of \( \mathbb{R}^n \), \( \mathbb{R}^n = V + V^\perp \), \( V \cap V^\perp = \{0\} \), and, in particular, \( \dim V^\perp = n - k \).

Theorem 3.21. Let \( k, n \in \mathbb{N} \) with \( k < n \), and let \( M \subseteq \mathbb{R}^n \) be a \( k \)-dimensional submanifold of class \( C^1 \), \( a \in M \).

(a) The tangent space \( T_a M \) is a \( k \)-dimensional vector space over \( \mathbb{R} \). Moreover, if \( \phi : W \rightarrow V \subseteq M, W \subseteq \mathbb{R}^k \) open, \( V \) \( M \)-open, is a local chart for \( M \) with \( \phi(c) = a, \)

\[
\{ \partial_1 \phi(c), \ldots, \partial_k \phi(c) \} \subseteq \mathbb{R}^n
\]
forms a basis of \( T_a M \).

(b) The normal space \( N_a M \) is a \( (n - k) \)-dimensional vector space over \( \mathbb{R} \). Moreover, if \( f \in C^1(O, \mathbb{R}^{n-k}) \) with \( a \in O, O \subseteq \mathbb{R}^n \) open, according to Def. 3.1, i.e. such that \( M \cap O = f^{-1}(\{0\}) \) and \( \text{rk} Df(a) = n - k \), then

\[
\{ \nabla f_1(a), \ldots, \nabla f_{n-k}(a) \} \subseteq \mathbb{R}^n
\]
forms a basis of \( N_a M \).
(c) $T_a M \perp N_a M$, i.e. if $\tau \in T_a M$ and $\nu \in N_a M$, then $\langle \tau, \nu \rangle = 0$.

Proof. Set

$$V_1 := \text{span} \{ \partial_1 \phi(c), \ldots, \partial_k \phi(c) \},$$

$$U := \text{span} \{ \nabla f_1(a), \ldots, \nabla f_{n-k}(a) \},$$

$$V_2 := U^\perp.$$

Then $V_1, U, V_2$ all are vector subspaces of $\mathbb{R}^n$. Since $\partial_1 \phi(c), \ldots, \partial_k \phi(c)$ are linearly independent, we have $\dim V_1 = k$. Since $\nabla f_1(a), \ldots, \nabla f_{n-k}(a)$ are linearly independent, we have $\dim U = n - k$. Then we know from Def. and Rem. 3.20 that $\dim V_2 = k$. We will show

$$V_1 \subseteq T_a M \subseteq V_2. \quad (3.39)$$

Then $\dim V_1 = \dim V_2$ implies $V_1 = V_2$ and $V_1 = T_a M = V_2$, proving (a). Then $N_a M = V_2^\perp = (T_a M)^\perp$, proving (b) and (c). It remains to show (3.39).

"$V_1 \subseteq T_a M$": Let $\tau \in V_1$. Then $\exists \alpha_1, \ldots, \alpha_k, \in \mathbb{R} \quad \tau = \sum_{i=1}^{k} \alpha_i \partial_i \phi(c).$

Since $c \in W$ with $W$ open, there exists $\epsilon \in \mathbb{R}^+$ such that

$$\left\{ x \in \mathbb{R}^k : \forall_{i \in \{1, \ldots, k\}} |x_i - c_i| < \epsilon |\alpha_i| \right\} \subseteq W.$$ 

We consider the $C^1$-function

$$\psi : ] - \epsilon, \epsilon [ \rightarrow M, \quad \psi(s) := \phi(c_1 + s\alpha_1, \ldots, c_k + s\alpha_k).$$

Then $\psi(0) = \phi(c) = a$ and, by the chain rule,

$$\psi'(0) = D\phi(c) \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_k \end{pmatrix} = \sum_{i=1}^{k} \alpha_i \partial_i \phi(c) = \tau,$$

showing $\tau \in T_a M$.

"$T_a M \subseteq V_2$": Let $\tau \in T_a M$. Then there exists $\epsilon > 0$ and a $C^1$-function $\psi : ] - \epsilon, \epsilon [ \rightarrow M$ such that $a = \psi(0)$ and $\tau = \psi'(0)$. Since $a \in O$, $O$ open, we may actually assume that $\psi$ maps into $M \cap O = f^{-1}(\{0\})$. Thus,

$$\forall_{i \in \{1, \ldots, n-k\}} f_i \circ \psi \equiv 0,$$

implying, for each $i \in \{1, \ldots, n-k\}$,

$$0 = (f_i \circ \psi)'(0) = \nabla f_i(\psi(0)) \psi'(0) = \langle \nabla f_i(a), \tau \rangle,$$

showing $\tau \in U^\perp = V_2$. \qed
3 INTEGRATION OVER SUBMANIFOLDS OF $\mathbb{R}^N$

3.5 Gauss-Green Theorem and Green’s Identities

The Gauss-Green Th. 3.30 allows, under suitable hypotheses, to transform an integral over an open subset $\Omega$ of $\mathbb{R}^n$ into an integral over the boundary $M := \partial \Omega$, where $M$ is an $(n-1)$-dimensional submanifold. The Gauss-Green Th. 3.30 can be seen as an $n$-dimensional generalization of the fundamental theorem of calculus, [Phi16a, Th. 10.20]. Before we can state and prove Th. 3.30, we still need some preparation.

Definition 3.22. Let $n \in \mathbb{N}$, $n \geq 2$, and let $\Omega \subseteq \mathbb{R}^n$ be open. We define $\Omega$ to have a $C^1$-boundary if, and only if, for each $a \in \partial \Omega$, there exists an open neighborhood $O_a \subseteq \mathbb{R}^n$ of $a$ and a map $f^a \in C^1(O_a, \mathbb{R})$, satisfying the following two conditions

(a) $\overline{\Omega} \cap O_a = \{ x \in O_a : f^a(x) \leq 0 \}$,
(b) $\nabla f^a(x) \neq 0$ for each $x \in O_a$.

Theorem 3.23. Let $n \in \mathbb{N}$, $n \geq 2$, and let $\Omega \subseteq \mathbb{R}^n$ be open, having a $C^1$-boundary. If $a \in \partial \Omega$ and $f^a \in C^1(O_a, \mathbb{R})$ is according to Def. 3.22, then

$$M \cap O_a = (f^a)^{-1}(\{0\}).$$

In particular, $\partial \Omega$ is an $(n-1)$-dimensional submanifold of class $C^1$.

Moreover, there exists $\epsilon \in \mathbb{R}^+$ such that

$$\forall s \in [0, \epsilon[ \quad a + s \nabla f^a(a) \notin \Omega.$$ (3.41)

Proof. “$\subseteq$ of (3.40)”: Recall $\overline{\Omega} = \Omega \cup \partial \Omega$, where the union is disjoint, as $\Omega$ is open. Thus, according to Def. 3.22(a), we have to show that $f^a(x) < 0$ implies $x \notin \partial \Omega$. However, if $x \in \overline{\Omega} \cap O_a$ with $f^a(x) < 0$, then, by the continuity of $f^a$, there exists an open set $U \subseteq O_a$ such that $f^a < 0$ on $U$, showing $x \notin \partial \Omega$.

“$\supseteq$ of (3.40)”: Let $x \in O_a$ with $f^a(x) = 0$. We need to verify $x \in M$. Consider the auxiliary function

$$\psi : ]-\epsilon, \epsilon[ \rightarrow \mathbb{R}, \quad \psi(s) := f^a(x + s \nabla f^a(x)),$$

which, as $x \in O_a$, $O_a$ open, is defined for sufficiently small $\epsilon > 0$. Moreover, $\psi$ is $C^1$, since $f^a$ is $C^1$. We obtain

$$\psi'(0) = \| \nabla f^a(x) \|_2^2 > 0.$$ 

In particular, $\psi'(0) = \| \nabla f^a(x) \|_2^2 > 0$. Thus, as $\psi'$ is continuous, $\psi$ is strictly increasing in an entire neighborhood of 0. Since $\psi(0) = f^a(x) = 0$, this shows every neighborhood of $x$ contains points from $\Omega$ and points from $\Omega^c$, showing $x \in M$. Setting $x = a$, we have also shown (3.41).

That $M = \partial \Omega$ is an $(n-1)$-dimensional submanifold of class $C^1$ is now immediate from Def. 3.1. □
Corollary 3.24 (Existence of Outer Unit Normal). Let \( n \in \mathbb{N}, \ n \geq 2, \) and let \( \Omega \subseteq \mathbb{R}^n \) be open, having a \( C^1 \)-boundary. Then, for each \( a \in \partial \Omega, \) there exists a unique normal vector \( \nu(a) \in N_a \partial \Omega, \) satisfying the following two conditions:

(i) \( \|\nu(a)\|_2 = 1. \)

(ii) There exists \( \epsilon \in \mathbb{R}^+ \) such that

\[
\forall s \in [0, \epsilon[ \quad a + s \nu(a) \notin \Omega.
\]

The vector \( \nu(a) \) is called the outer unit normal to \( \partial \Omega \) at \( a. \) Moreover, the function

\[
\nu : \partial \Omega \rightarrow \mathbb{R}^n, \quad a \mapsto \nu(a),
\]

is continuous.

Proof. Let \( a \in \partial \Omega \) and let \( f^a \in C^1(O_a, \mathbb{R}) \) be according to Def. 3.22. Then \( \nabla f^a \neq 0 \) and we may define

\[
\nu(a) := \frac{\nabla f^a(a)}{\| \nabla f^a(a) \|_2}.
\] (3.42)

Then \( \|\nu(a)\|_2 \) is clear, condition (ii) holds according to (3.41), and \( \nu(a) \in N_a \partial \Omega \) by Th. 3.21(b). To show \( \nu \) as a function on \( \partial \Omega \) is continuous, it suffices to show \( \nu \) is continuous at each \( a \in \partial \Omega. \) However, given \( a \in \partial \Omega, \) on \( O_a \cap \partial \Omega, \) \( \nu \) is given by the continuous function \( x \mapsto \frac{\nabla f^a(x)}{\| \nabla f^a(x) \|_2}. \) \( \square \)

Example 3.25. Let \( n \in \mathbb{N}, \ n \geq 2. \) For \( x_M \in \mathbb{R}^n \) and \( r > 0, \) consider the Euclidean ball \( \Omega := B_r(x_M). \) Then \( \partial \Omega = \{ x \in \mathbb{R}^n : \| x - x_M \|_2 = r \} \) and, according to Ex. 3.2(b), for each \( a \in \partial \Omega, \) we may choose \( f^a := f \) with

\[
f : \mathbb{R}^n \rightarrow \mathbb{R}, \quad f(x) := \| x - x_M \|_2^2 - r^2 = \sum_{i=1}^n (x_i - x_{M,i})^2 - r^2,
\]

\[
\nabla f : \mathbb{R}^n \rightarrow \mathbb{R}^n, \quad \nabla f(x) = 2(x_1 - x_{M,1}, \ldots, x_n - x_{M,n}).
\]

Thus,

\[
\forall a \in \partial \Omega \quad \nu(a) = \frac{\nabla f(a)}{\| \nabla f(a) \|_2} = \frac{a - x_M}{r}.
\] (3.43)

Proposition 3.26. Let \( n \in \mathbb{N}, \ n \geq 2, \) and let \( \Omega \subseteq \mathbb{R}^n \) be open, having a \( C^1 \)-boundary. Then, for each \( a \in M := \partial \Omega, \) there exists a permutation \( \pi^a \in S_n \) and open intervals \( A_a \subseteq \mathbb{R}^{n-1} \) and \( I_a \subseteq \mathbb{R}, \) such that

\[
\alpha^a \in A_a \subseteq \mathbb{R}^{n-1}, \quad \beta^a \in I_a \subseteq \mathbb{R},
\]

where \( \alpha^a := (a_{\pi^a(1)}, \ldots, a_{\pi^a(n-1)}), \) \( \beta^a := a_{\pi^a(n)}, \) and there exists a \( C^1 \)-map \( g^a : A_a \rightarrow I_a, \) satisfying

\[
M \cap P_{\pi^a}(A_a \times I_a) = P_{\pi^a}\left(\{(x, y) \in A_a \times I_a : y = g^a(x)\}\right), \quad \tau^a := (\pi^a)^{-1},
\] (3.44a)
and either
\[ \Omega \cap P_{\pi^a}(A_a \times I_a) = P_{\pi^a}(\{(x, y) \in A_a \times I_a : y < g^a(x)\}) \] (3.44b)
or
\[ \Omega \cap P_{\pi^a}(A_a \times I_a) = P_{\pi^a}(\{(x, y) \in A_a \times I_a : y > g^a(x)\}). \] (3.44c)

Then the outer unit normal \( \nu(a) \) is given by
\[ \nu(a) = \pm \frac{(-\nabla g(\alpha^a), 1) P_{\pi^a}}{\sqrt{1 + \|\nabla g(\alpha^a)\|^2_2}}, \quad (\alpha^a, \beta^a) := P_{\pi^a}(a) \] (3.45)
(with “+” for (3.44b) and “−” for (3.44c)).

Proof. Let \( f^a \in C^1(O_a, \mathbb{R}), \) \( a \in O_a \subseteq \mathbb{R}^n \) open, be according to Def. 3.22. Then, by Def. 3.22 and Th. 3.23, \( f^a \) satisfies the conditions of Def. 3.1. In particular, \( M = \partial \Omega \) is an \((n-1)\)-dimensional submanifold of class \( C^1 \), and we can then choose \( \pi^a, A_a, I_a := B_a, \) and \( g^a \) according to Th. 3.4(ii). If necessary, we make the sets \( A_a, B_a \) of Th. 3.4(ii) smaller to obtain them to be open intervals and we make \( O_a \) smaller to have \( O_a = P_{\pi^a}(A_a \times I_a). \) By possibly shrinking \( A_a \) again, we may assume that \( \text{dist}(g^a(A_a), (I_a)^c) := \epsilon > 0 \) (i.e. \( g^a \) maps \( A_a \) into a compact subinterval of \( I_a \)). We may also assume \( g^a \) to be obtained from \( h^a := f^a \circ P_{\pi^a} \) via the implicit function theorem as in the proof of “(i) \( \Rightarrow \) (ii)” of Th. 3.4(ii). Then, for \( i_n := \tau_n(n) \), we have \( \partial_{i_n} f^a(a) \neq 0. \) By the continuity of \( \partial_{i_n} f^a \), we may assume that \( \partial_{i_n} f^a \) has only one sign (everywhere positive or everywhere negative) on \( O_a = P_{\pi^a}(A_a \times I_a) \) (i.e., on \( O_a \), \( f^a \) is either strictly increasing or strictly decreasing with respect to the \( i_n \)-th variable). Now (3.44a) holds due to Th. 3.4(ii) and, by Def. 3.22 and Th. 3.23, we have
\[ M \cap O_a = (f^a)^{-1}(\{0\}), \]
\[ \Omega \cap O_a = \{x \in O_a : f^a(x) < 0\}. \]

We consider the open sets
\[ U_+ := P_{\pi^a}(\{(x, y) \in A_a \times I_a : 0 < g^a(x) - y\}), \]
\[ U_- := P_{\pi^a}(\{(x, y) \in A_a \times I_a : 0 > g^a(x) - y\}), \]
\[ O_a = U_+ \cup U_- \cup (M \cap O_a). \]

Then \( \Omega \cap O_a \subseteq U_+ \cup U_- \). However \( \Omega \cap O_a \) cannot have nonempty intersection with both \( U_+ \) and \( U_- \): Suppose \( z_1 \in \Omega \cap O_a \cap U_+ \) and \( z_2 \in \Omega \cap O_a \cap U_- \). Let \( (x_1, y_1) := P_{\pi^a}(z_1) \in A_a \times I_a, \)
\( (x_2, y_2) := P_{\pi^a}(z_2) \in A_a \times I_a. \) Then \( g^a(x_1) > y_1, g^a(x_2) < y_2. \) Since \( h^a \) is continuous, \( h^a(x_1, y_1) = f^a(z_1) < 0, \) and \( h^a \) does not have a zero on \( \{(x_1, y) : y_1 \leq y < g^a(x_1)\} \), we can assume \( \epsilon > c := g^a(x_1) - y_1 > 0. \) Then, since \( A_a \) is convex, the map
\[ \psi_1 : [0, 1] \rightarrow A_a \times I_a, \quad \psi_1(s) := (x_1 + s(x_2 - x_1), g^a(x_1 + s(x_2 - x_1)) - c) \]
is well-defined with $\psi_1(0) = (x_1, y_1)$, $\psi_1(1) = (x_2, g^a(x_2) - c)$. Since $h^a \circ \psi_1$ is continuous, $(h^a \circ \psi_1)(0) < 0$ and $h^a \circ \psi_1$ does not have any zeros on $[0, 1]$ (since $c > 0$), $(h^a \circ \psi_1)(1) < 0$, i.e. $h^a(x_2, g^a(x_2) - c) < 0$. However, due to the choice of $i_n$ above, the map

$$\psi_2 : [g^a(x_2) - c, y_2] \rightarrow \mathbb{R}, \quad \psi_2(s) := h^a(x_2, s),$$

either strictly increasing or strictly decreasing. Since $\psi_2(g^a(x_2) - c) = h^a(x_2, g^a(x_2) - c) < 0$, $\psi_2(g^a(x_2)) = h^a(x_2, g^a(x_2)) = 0$, $\psi_2(y_2) = h^a(x_2, y_2) = f^a(z_2) < 0$, we obtain a contradiction, proving $\Omega \cap O_a \subseteq U_+$ (i.e. (3.44b)) or $\Omega \cap O_a \subseteq U_-$ (i.e. (3.44c)). For (3.44b), we now consider

$$f_a : P_{\pi^a}(A_a \times I_a) \rightarrow \mathbb{R}, \quad f_a(z) := x_n - g(x_1, \ldots, x_{n-1}), \quad x := P_{\pi^a}(z)$$

(this might not be the same $f^a$ as before, but it also satisfies Def. 3.22). We then obtain

$$\nu(a) = \frac{\nabla f_a(a)}{\|\nabla f_a(a)\|_2} = \frac{\left(-\nabla g((a^a), 1)\right) P_{\pi^a}}{\sqrt{1 + \|\nabla g((a^a))\|_2^2}}, \quad (\alpha^a, \beta^a) := P_{\pi^a}(a),$$

proving (3.45). For (3.44c), one uses $f_a(z) := x_n + g(x_1, \ldots, x_{n-1}), x := P_{\pi^a}(z)$, again proving (3.45).

**Lemma 3.27.** Let $n \in \mathbb{N}$, $O \subseteq \mathbb{R}^n$ open, and $\eta \in C^1_c(O, \mathbb{R})$. Then

$$\forall \quad \int_O \partial_i \eta \, d\lambda^n = 0. \quad (3.46)$$

**Proof.** Extending $\eta$ by 0, we may assume $O = \mathbb{R}^n$. Then there exists $r \in \mathbb{R}^+$ such that $\text{supp} \eta \subseteq [-r, r]^n$. Then, using Fubini,

$$\int_O \partial_i \eta \, d\lambda^n = \int_{[-r, r]^n} \partial_i \eta \, d\lambda^n$$

$$= \int_{[-r, r]^n-1} \left(\eta(x_1, \ldots, x_i = r, \ldots, x_n) - \eta(x_1, \ldots, x_i = -r, \ldots, x_n)\right) d\lambda^{n-1}(x_1, \ldots, \hat{x_i}, \ldots, x_n) = 0,$$

where $\hat{x}_i$ means that $x_i$ is omitted. $$\Box$$

In preparation for the proof of the Gauss-Green Th. 3.30, we prove a special case in the following lemma:

**Lemma 3.28.** Let $n \in \mathbb{N}$, $n \geq 2$. Let $A \subseteq \mathbb{R}^{n-1}$ be open and let $I := ]\alpha, \beta[$ with $\alpha, \beta \in \mathbb{R}$, $\alpha < \beta$, be an open interval. Given a $C^1$-function

$$g : A \rightarrow I,$$

set

$$\Omega := \{(\hat{x}, x_n) \in A \times I : x_n < g(\hat{x})\},$$
$$M := \{(\hat{x}, x_n) \in A \times I : x_n = g(\hat{x})\} \subseteq \partial \Omega.$$
Then, for each $f \in C^1_c(A \times I, \mathbb{R})$, one has

$$\forall i \in \{1, \ldots, n\} \quad \int_{\Omega} \partial_i f \, d\lambda^n = \int_M f \nu_i \, d\sigma,$$  \hspace{1cm} (3.47)

where

$$\nu : M \to \mathbb{R}^n, \quad \nu(\tilde{x}, x_n) := \left( -\nabla g(\tilde{x}), 1 \right) \sqrt{1 + \| \nabla g(\tilde{x}) \|^2}.$$  \hspace{1cm} (3.48)

Moreover, (3.47) still holds if one has “>” instead of “<” in the definition of $\Omega$ and a “−” in front of the right-hand side of (3.48).

**Proof.** First, let $1 \leq i \leq n - 1$. We define the auxiliary function

$$F : A \times I \to \mathbb{R}, \quad F(\tilde{x}, x_n) := \int_{\alpha}^{x_n} f(\tilde{x}, s) \, ds.$$  

Then

$$\forall (\tilde{x}, x_n) \in A \times I \quad \left( \partial_n F(\tilde{x}, x_n) = f(\tilde{x}, x_n), \quad \partial_i F(\tilde{x}, x_n) = \int_{\alpha}^{x_n} \partial_i f(\tilde{x}, s) \, ds \right)$$

and the chain rule yields, for each $(\tilde{x}, x_n) \in A \times I$,

$$\partial_i \int_{\alpha}^{g(\tilde{x})} f(\tilde{x}, s) \, ds = \partial_i F(\tilde{x}, g(\tilde{x})) = \int_{\alpha}^{g(\tilde{x})} \partial_i f(\tilde{x}, s) \, ds + f(\tilde{x}, g(\tilde{x})) \partial_i g(\tilde{x}).$$  \hspace{1cm} (3.49)

The function $\tilde{x} \mapsto \int_{\alpha}^{g(\tilde{x})} f(\tilde{x}, s) \, ds$ is in $C^1_c(A, \mathbb{R})$ and, thus, Lem. 3.27 yields

$$\int_{A} \partial_i \left( \int_{\alpha}^{g(\tilde{x})} f(\tilde{x}, s) \, ds \right) \, d\lambda^{n-1}(\tilde{x}) = 0.$$  \hspace{1cm} (3.50)

Using Fubini, we obtain

$$\int_{\Omega} \partial_i f \, d\lambda^n = \int_{A} \int_{\alpha}^{g(\tilde{x})} \partial_i f(\tilde{x}, s) \, ds \, d\lambda^{n-1}(\tilde{x})$$

$$\overset{(3.49)}{=} \int_{A} \left( \int_{\alpha}^{g(\tilde{x})} f(\tilde{x}, s) \, ds \right) \frac{\partial_i g(\tilde{x})}{\sqrt{1 + \| \nabla g(\tilde{x}) \|^2}} \sqrt{1 + \| \nabla g(\tilde{x}) \|^2} \, d\lambda^{n-1}(\tilde{x})$$

$$\overset{(3.50)}{=} - \int_{A} f(\tilde{x}, g(\tilde{x})) \frac{\partial_i g(\tilde{x})}{\sqrt{1 + \| \nabla g(\tilde{x}) \|^2}} \, d\lambda^{n-1}(\tilde{x})$$

$$\overset{(3.48),(3.21)}{=} \int_{M} f \nu_i \, d\sigma.$$

It remains to consider the case $i = n$. For each $\tilde{x} \in A$, the function $s \mapsto f(\tilde{x}, s)$ is in $C^1_c(I, \mathbb{R})$, implying

$$\int_{\alpha}^{g(\tilde{x})} \partial_n f(\tilde{x}, s) \, ds = f(\tilde{x}, g(\tilde{x})) - 0 = f(\tilde{x}, g(\tilde{x}))$$
3 INTEGRATION OVER SUBMANIFOLDS OF $\mathbb{R}^N$

and, thus,

$$
\int_{\Omega} \partial_n f \, d\lambda^n = \int_{A} \int_{\alpha} g(\tilde{x}) \, \partial_n f(\tilde{x}, s) \, ds \, d\lambda^{n-1}(\tilde{x}) = \int_{A} f(\tilde{x}, g(\tilde{x})) \, d\lambda^{n-1}(\tilde{x})
$$

(3.48), (3.21) $\equiv \int_{M} f \nu \, d\sigma,

thereby completing the proof of the first case. One now obtains the second case by applying the first to $-\Omega$, $-M$, $-g : -A \to -I$, followed by a change of variables $x \mapsto -x$. 

As a last preparation for the proof of the Gauss-Green Th. 3.30, we construct a $C^\infty$ partition of unity. Here, the partition of unity is subordinate to an open cover of $\mathbb{R}^n$, not to an open cover of a submanifold. For $\epsilon > 0$, we consider the cover given by

$$(B_{2\epsilon}(\epsilon p))_{p \in \mathbb{Z}^n},$$

where the balls are defined with respect to the $\infty$-norm on $\mathbb{R}^n$ (i.e. they are open intervals of side length $4\epsilon$).

Proposition 3.29. Let $n \in \mathbb{N}$, $\epsilon \in \mathbb{R}^+$. Then there exists a family $(\eta_p)_{p \in \mathbb{Z}^n}$ of functions in $C^\infty_c(\mathbb{R}^n, \mathbb{R})$, satisfying the following conditions (i) – (iii) (the balls are defined with respect to the $\infty$-norm):

(i) $0 \leq \eta_p \leq 1$ for each $p \in \mathbb{Z}^n$.

(ii) $\text{supp} \eta_p = \overline{B}_\epsilon(\epsilon p)$ for each $p \in \mathbb{Z}^n$.

(iii) $\sum_{p \in \mathbb{Z}^n} \eta_p \equiv 1$.

Proof. According to Ex. 2.53(c), the function

$$
\varphi : \mathbb{R} \to [0, 1], \quad \varphi(t) := \begin{cases} 
\exp\left(-\frac{1}{t^2-\epsilon^2}\right) & \text{for } |t| < \epsilon, \\
0 & \text{otherwise},
\end{cases}
$$

is in $C^\infty_c(\mathbb{R}, \mathbb{R})$, $\text{supp} \varphi = [-\epsilon, \epsilon]$. Define

$$
G : \mathbb{R} \to \mathbb{R}^+, \quad G(t) := \sum_{p \in \mathbb{Z}} \varphi(t - \epsilon p). \quad (3.51)
$$

Since the support of $t \mapsto \varphi(t - \epsilon p)$ is $[\epsilon p - \epsilon, \epsilon p + \epsilon]$, for each $t \in \mathbb{R}$, at least one and at most two summands in (3.51) are nonzero. In particular, $G$ is well-defined as a function into $\mathbb{R}^+$. Thus,

$$
g : \mathbb{R} \to [0, 1], \quad g(t) := \frac{\varphi(t)}{G(t)},
$$

is well-defined, $g \in C^\infty_c(\mathbb{R}, \mathbb{R})$, $\text{supp} g = [-\epsilon, \epsilon]$. We claim that letting

$$
\forall_{p \in \mathbb{Z}} \eta_p : \mathbb{R} \to [0, 1], \quad \eta_p(t) := g(t - \epsilon p),
$$

...
proves the proposition for \( n = 1 \): Clearly, (i) and (ii) are satisfied. Since, for each \( t \in \mathbb{R} \),
\[
\sum_{p \in \mathbb{Z}} \eta_p(t) = \sum_{p \in \mathbb{Z}} \frac{\varphi(t - \epsilon p)}{G(t - \epsilon p)} = \sum_{p \in \mathbb{Z}} \frac{\varphi(t - \epsilon p)}{G(t)} = 1,
\]
(iii) holds as well. Now, for an arbitrary \( n \in \mathbb{N} \), define
\[
\forall \ p \in \mathbb{Z}^n \eta_p : \mathbb{R}^n \to [0, 1], \quad \eta_p(x) := \prod_{i=1}^{n} g(x_i - \epsilon p_i).
\]
As before, \( \eta_p \in C_c^\infty(\mathbb{R}^n, \mathbb{R}) \) and (i) and (ii) are immediate consequences of the properties of \( g \). To obtain (iii), we carry out and induction, assuming the case \( n - 1 \) by induction hypothesis, obtaining, for each \( x \in \mathbb{R}^n \),
\[
\sum_{p \in \mathbb{Z}^n} \eta_p(x) = \sum_{k \in \mathbb{Z}} \sum_{q \in \mathbb{Z}^{n-1}} g(x_n - \epsilon k) \prod_{i=1}^{n-1} g(x_i - \epsilon q_i) \\
= \sum_{k \in \mathbb{Z}} g(x_n - \epsilon k) \sum_{q \in \mathbb{Z}^{n-1}} \prod_{i=1}^{n-1} g(x_i - \epsilon q_i) \\
= \sum_{k \in \mathbb{Z}} g(x_n - \epsilon k) = 1,
\]
completing the proof.

The following version of the Gauss-Green theorem is certainly not the most general. For example, it still holds if, instead of a \( C^1 \)-boundary, one just has a so-called Lipschitz boundary (which, e.g., allows corners if they are sufficiently benign, including polyhedral sets), see [Kön64].

**Theorem 3.30 (Gauss-Green).** Let \( n \in \mathbb{N}, n \geq 2 \). Let \( \Omega \subseteq \mathbb{R}^n \) be open and bounded, having a \( C^1 \)-boundary (then \( \overline{\Omega} \) and \( \partial \Omega \) are compact; in particular, \( \partial \Omega \) is a finitely cov erable \((n - 1)\)-dimensional submanifold of class \( C^1 \)). If \( \overline{\Omega} \subseteq U \subseteq \mathbb{R}^n \) with \( U \) open and \( F \in C^1(U, \mathbb{R}^n) \) (in this context, \( F \) is often called a vector field), then
\[
\int_{\Omega} \text{div} \, F \, d\lambda^n = \int_{\partial \Omega} \langle F, \nu \rangle \, d\sigma \tag{3.52}
\]
(as before, \( \nu : \partial \Omega \to \mathbb{R}^n \) denotes the outer unit normal to \( \Omega \), also recall that \( \text{div} \, F = \sum_{i=1}^{n} \partial_i F_i \)).

**Proof.** Without loss of generality, we may assume \( U \) to be bounded, i.e. \( \overline{U} \) to be compact. For each \( a \in \partial \Omega \), we let \( g_a : A_a \to I_a \), and \( \pi^a, \tau^a \in S_n \) be as in Prop. 3.26. Then \( (P_{r^a}(A_a \times I_a))_{a \in \partial \Omega} \) together with \( \Omega \) provide an open cover \( \mathcal{O} := (U_j)_{j \in J} \) of the compact set \( \overline{\Omega} \), where, by making the \( A_a, I_a \) smaller if necessary, we may also assume \( U_j \subseteq U \), i.e. \( V := \bigcup_{j \in J} U_j \) is open and bounded, \( \alpha := \text{dist}(\overline{\Omega}, V^c) > 0 \) (here, and in the following, we consider \( \mathbb{R}^n \) with the \( \infty \)-norm). Moreover, \( \mathcal{O} \) still constitutes an open cover of the
compact set \( C := \{ x \in \mathbb{R}^n : \text{dist}(x, \overline{\Omega} \leq \frac{\epsilon}{2} \right\} \). According to [Phi16b, Th. 3.21], we can choose a Lebesgue number \( \delta > 0 \) for this cover. As we choose it with respect to the \( \infty \)-norm on \( \mathbb{R}^n \), then each hypercube \( I \) with side length \( \delta \), we have \( I \cap \partial \Omega \) is contained in one of the open sets of the cover. We choose \( \epsilon > 0 \) such that \( 2\epsilon < \delta \), and we let \( (\eta_p)_{p \in \mathbb{Z}^n} \) be the corresponding family of functions in \( C_c^\infty(\mathbb{R}^n, \mathbb{R}) \) given by Prop. 3.29. Since, for each \( p \in \mathbb{Z}^n \), \( \text{supp} \eta_p = \overline{B}_\epsilon(p) \), if \( \text{supp} \eta_p \cap \overline{\Omega} \neq \emptyset \), then \( \text{supp} \eta_p \subseteq C \) and, by the choice of \( \delta \), there exists \( j \in J \) such that \( \text{supp} \eta_p \subseteq U_j \). If \( \text{supp} \eta_p \subseteq \Omega \), then

\[
\int_\Omega \text{div}(\eta_p F) \, d\lambda^n = 0 = \int_{\partial \Omega} \langle \eta_p F, \nu \rangle \, d\sigma.
\]

If \( \text{supp} \eta_p \subseteq P_{+a} (A_a \times I_a) \), then we apply Lem. 3.28 to obtain, for each \( i \in \{1, \ldots, n\} \),

\[
\int_\Omega \partial_i (\eta_p F_i) \, d\lambda^n = \int_\Omega \partial_i (\eta_p F_i) \, d\lambda^n = \int_{(P_{+a} \cap \partial \Omega) \cap (A_a \times I_a)} \partial_i (\eta_p F_i) \, d\lambda^n = \int_{(P_{+a} \cap \partial \Omega) \cap (A_a \times I_a)} (\eta_p F_i) \, d\sigma
\]

(3.47)

(3.45)

(3.46)

(3.43)

(3.53)

(3.43)

(3.43)

(3.43)

(3.43)

(3.43)

(3.43)

(3.43)

(3.43)

(3.43)

(3.43)

(3.43)

(3.43)

(3.43)

(3.43)

(3.43)

(3.43)

(3.43)

(3.43)

(3.43)

(3.43)

(3.43)

(3.43)

(3.43)

(3.43)

(3.43)

(3.43)

(3.43)

(3.43)

(3.43)

(3.43)

(3.43)

(3.43)

(3.43)

(3.43)

(3.43)

(3.43)

(3.43)

(3.43)

(3.43)

(3.43)

(3.43)

(3.43)

(3.43)

(3.43)

(3.43)

(3.43)

(3.43)

(3.43)

(3.43)

(3.43)

(3.43)

(3.43)

(3.43)

(3.43)

(3.43)

(3.43)

(3.43)

(3.43)

(3.43)

(3.43)

(3.43)

(3.43)

(3.43)

(3.43)

(3.43)

(3.43)

(3.43)

(3.43)

(3.43)

(3.43)

(3.43)

(3.43)

(3.43)

(3.43)

(3.43)

(3.43)

(3.43)

(3.43)

(3.43)

(3.43)

(3.43)

(3.43)

(3.43)

(3.43)

(3.43)

(3.43)

(3.43)

(3.43)

(3.43)

(3.43)

(3.43)

(3.43)

(3.43)

(3.43)

(3.43)

(3.43)

(3.43)
(b) To evaluate the integral \( \int_{S_1(0)} (x^4 + y^4) \, d\sigma (x, y) \), we apply (3.52) with

\[
F : \mathbb{R}^2 \rightarrow \mathbb{R}^2, \quad F(x, y) = (x^3, y^3),
\]

\[
\text{div} F : \mathbb{R}^2 \rightarrow \mathbb{R}, \quad \text{div} F(x, y) = 3x^2 + 3y^2.
\]

Since \( S_1(0) = \partial B_1(0) \), we can use (3.43) to obtain \( \nu(x, y) = (x, y) \) for \((x, y) \in S_1(0)\).
Thus,

\[
\int_{S_1(0)} (x^4 + y^4) \, d\sigma (x, y) = \int_{S_1(0)} \left( (x^3, y^3), (x, y) \right) \, d\sigma (x, y)
\]

\[
= \int_{B_1(0)} (3x^2 + 3y^2) \, d\lambda^2 (x, y)
\]

\[
= 3 \int_0^1 \int_0^{2\pi} r^2 \, d\varphi \, dr = \frac{3 \cdot 2\pi}{4} = \frac{3\pi}{2},
\]

where we used polar coordinates to evaluate the integral over \( B_1(0) \).

Finally, as another application of Th. 3.30, we will prove Green’s identities, which, for example, are of use in the theory of partial differential equations.

**Theorem 3.32** (Green’s Identities). Let \( n \in \mathbb{N} \), \( n \geq 2 \). Let \( \Omega \subseteq \mathbb{R}^n \) be open and bounded, having a \( C^1 \)-boundary (then \( \overline{\Omega} \) and \( \partial \Omega \) are compact; in particular, \( \partial \Omega \) is a finitely coverable \((n - 1)\)-dimensional submanifold of class \( C^1 \)). If \( \overline{\Omega} \subseteq U \subseteq \mathbb{R}^n \) with \( U \) open and \( f, g \in C^2(U, \mathbb{R}) \), then the following identities hold:

\[
\int_{\Omega} \langle \nabla f, \nabla g \rangle \, d\lambda^n = \int_{\partial \Omega} f \langle \nabla g, \nu \rangle \, d\sigma - \int_{\Omega} f \Delta g \, d\lambda^n, \tag{3.54a}
\]

\[
\int_{\Omega} (f \Delta g - g \Delta f) \, d\lambda^n = \int_{\partial \Omega} \left( f \langle \nabla g, \nu \rangle - g \langle \nabla f, \nu \rangle \right) \, d\sigma, \tag{3.54b}
\]

recalling \( \Delta f = \sum_{i=1}^n \partial_i \partial_i f \).

**Proof.** To prove (3.54a), apply Th. 3.30 to

\[
F : U \rightarrow \mathbb{R}^n, \quad F(x) := f(x) \nabla g(x) :
\]

One has

\[
\text{div} F : U \rightarrow \mathbb{R}, \quad \text{div} F(x) = \langle \nabla f(x), \nabla g(x) \rangle + f(x) \Delta g(x) : \tag{3.55}
\]

Indeed,

\[
\langle \nabla f, \nabla g \rangle + f \Delta g = \sum_{i=1}^n \left( \partial_i f \partial_i g + f \partial_i \partial_i g \right) = \sum_{i=1}^n \partial_i \left( f \partial_i g \right) = \text{div} F(x).
\]
Thus,
\[ \int_{\Omega} (\langle \nabla f, \nabla g \rangle + f \Delta g) \, d\lambda^n = \int_{\Omega} \text{div} \, F \, d\lambda^n \overset{(3.52)}{=} \int_{\partial \Omega} \langle F, \nu \rangle \, d\sigma = \int_{\partial \Omega} f \langle \nabla g, \nu \rangle \, d\sigma, \]
proving (3.54a).

If one interchanges the roles of \( f \) and \( g \) in (3.54a) and subtracts the resulting equation from the original one, then one obtains (3.54b). ■

# A Order on, Arithmetic in, and Topology of \( \mathbb{R} \)

## Notation A.1.
By \( \mathbb{R} := \mathbb{R} \cup \{-\infty, \infty\} \), we denote the set of extended real numbers.

One defines the order on, arithmetic in, and topology on \( \mathbb{R} \) such that they, as far as possible, constitute extensions of the respective (standard) definitions on \( \mathbb{R} \). In [Phi16b, Ex. 1.52(b)], we had actually already considered the order and (resulting order) topology on \( \mathbb{R} \). Here, we include it again for convenience.

**Definition A.2.**

(a) The (total) order on \( \mathbb{R} \) is extended to \( \mathbb{R} \) by setting \( -\infty < a < \infty \) for each \( a \in \mathbb{R} \). The absolute value function is extended from \( \mathbb{R} \) to \( \mathbb{R} \) by defining \( |\infty| := | - \infty| := \infty \).

(b) The set of open intervals \( \mathcal{I}_o \) in \( \mathbb{R} \) is defined to consist of \( \mathbb{R} \) plus all open intervals in \( \mathbb{R} \) plus all intervals of the form \( [ -\infty, a[ \) or \( ]a, \infty[ \), \( a \in \mathbb{R} \):

\[
\mathcal{I}_o := \{ \mathbb{R} \} \cup \{ [a, b[ : a, b \in \mathbb{R}, a < b \} \\
\cup \{ [-\infty, a[ : a \in \mathbb{R} \cup \{ \infty \} \} \cup \{ ]a, \infty[ : a \in \mathbb{R} \cup \{ -\infty \} \}. \quad (A.1)
\]

Then the (standard) topology on \( \mathbb{R} \) is defined by calling a set \( O \subseteq \mathbb{R} \) open if, and only if, each element \( x \in O \) is contained in an open interval \( I \in \mathcal{I}_o \) that is contained in \( O \), i.e. \( x \in I \subseteq O \). In other words, \( \mathcal{I}_o \) is defined to be a base for the topology on \( \mathbb{R} \).

(c) Addition, subtraction, and multiplication are extended from \( \mathbb{R} \) to \( \mathbb{R} \) by defining:

\[
\forall a \in \mathbb{R} \quad a + (\pm \infty) := (\pm \infty) + a := \pm \infty, \quad (A.2a) \\
\forall a \in \mathbb{R} \quad a - (\pm \infty) := -(\pm \infty) + a := \mp \infty, \quad (A.2b) \\
\infty + \infty := \infty, \quad -\infty + (-\infty) := -\infty, \quad -(\pm \infty) := \mp \infty, \quad (A.2c) \\
\forall a \in \mathbb{R} \quad a \cdot (\pm \infty) := (\pm \infty) \cdot a := \begin{cases} \pm \infty & \text{for } a \in [0, \infty], \\ \mp \infty & \text{for } a \in [-\infty, 0[; \end{cases}, \quad (A.2d) \\
0 \cdot (\pm \infty) := (\pm \infty) \cdot 0 := 0, \quad (A.2e) \\
\infty - \infty := -\infty + \infty := 0. \quad (A.2f)
\]
Caveat A.3. Addition is not associative on $\mathbb{R}$:
\[(−\infty + \infty) + 1 = 0 + 1 = 1 \neq 0 = −\infty + \infty = −\infty + (\infty + 1).\] (A.3)

Addition and multiplication are not distributive on $\mathbb{R}$:
\[(-1 + 2) \cdot \infty = 1 \cdot \infty = \infty \neq 0 = −\infty + \infty = (−1) \cdot \infty + 2 \cdot \infty.\] (A.4)

However, the following Lem. A.4 shows that commutativity of addition and multiplication hold on $\mathbb{R}$ as well as associativity of multiplication; associativity of addition also holds if one removes one of the infinities. Moreover, $(\mathbb{R}, +)$ is not a group, as $+$ is not associative; $(\mathbb{R} \setminus \{0\}, -)$ is not a group, as the infinities do not have multiplicative inverses; $(\mathbb{R} \setminus \{\infty\}, +)$ and $(\mathbb{R} \setminus \{-\infty\}, +)$ are not groups, as, in both cases, the remaining infinity does not have an additive inverse.

Lemma A.4. Consider $\mathbb{R}$ with addition and multiplication as defined in Def. A.2(c).

(a) Addition is associative on $\mathbb{R}$: If $a,b,c \in \mathbb{R}$, then
\[(a + b) + c = a + (b + c).\] (A.5)

(b) Addition is commutative on $\mathbb{R}$: If $a, b \in \mathbb{R}$, then
\[a + b = b + a.\] (A.6)

(c) Multiplication is associative on $\mathbb{R}$: If $a, b, c \in \mathbb{R}$, then
\[(ab)c = a(bc).\] (A.7)

(d) Multiplication is commutative on $\mathbb{R}$. If $a, b \in \mathbb{R}$, then
\[ab = ba.\] (A.8)

Proof. (a): If $a, b, c \in [-\infty, \infty]$, then both sides of (A.5) equal $\infty$ if, and only if, $\infty \in \{a,b,c\}$; if $a,b,c \in [-\infty, \infty]$, then both sides of (A.5) equal $-\infty$ if, and only if, $-\infty \in \{a,b,c\}$, proving (a).

(b): Commutativity of addition is immediate from (A.2a) and (A.2f).

(c): If $0 \notin \{a,b,c\}$, then (A.7) holds, since both sides are 0. If $a,b,c \in \mathbb{R}$, then (A.7) holds, since $\cdot$ is associative on $\mathbb{R}$. If $0 \notin \{a,b,c\}$, but $\{-\infty, \infty\} \cap \{a,b,c\} \neq \emptyset$, then $(ab)c = \pm \infty$ and $a(bc) = \pm \infty$. However, letting
\[\text{sgn}(\pm \infty) := \pm 1,\] (A.9)
due to (A.2d), we obtain
\[(ab)c = (\text{sgn} a \cdot \text{sgn} b \cdot \text{sgn} c) \cdot \infty = (\text{sgn} a \cdot (\text{sgn} b \cdot \text{sgn} c) \cdot \infty = a(bc),\]
proving (c).

(d): Commutativity of multiplication is immediate from (A.2d) and (A.2e).
Remark A.5. with the topology defined in Def. A.2(b) constitutes a so-called compactification of \( \mathbb{R} \), i.e. it is a compact topological space such that the topology on \( \mathbb{R} \) is recovered as the relative topology when considering \( \mathbb{R} \) as a subset of \( \overline{\mathbb{R}} \). Actually, it is straightforward to verify \( \overline{\mathbb{R}} \) is homeomorphic to \([−1, 1]\), where

\[
\phi : \overline{\mathbb{R}} \rightarrow [−1, 1], \quad \phi(x) := \begin{cases} 
-1 & \text{for } x = -\infty, \\
\frac{x}{|x|+1} & \text{for } x \in \mathbb{R}, \\
1 & \text{for } x = \infty,
\end{cases}
\]

provides a homeomorphism.

\section*{B Algebraic Structure on Rings of Subsets}

\textbf{Notation B.1.} For sets \( A, B \), we let

\[
A \Delta B := (A \setminus B) \cup (B \setminus A) = (A \cup B) \setminus (A \cap B)
\]

(B.1)

denote the symmetric difference of \( A \) and \( B \).

Let \( X \) be a set. Recall that, for \( A \subseteq X \), we have the characteristic function

\[
\chi_A : X \rightarrow \{0, 1\}, \quad \chi_A(x) := \begin{cases} 
1 & \text{if } x \in A, \\
0 & \text{if } x \notin A.
\end{cases}
\]

(B.2)

Moreover, we know from [Phi16b, Prop. 2.18] that

\[
\chi : \mathcal{P}(X) \rightarrow \{0, 1\}^X, \quad \chi(A) := \chi_A.
\]

(B.3)

is bijective. Also recall from Linear Algebra that, with addition and multiplication modulo 2, \( \mathbb{Z}_2 = \{0, 1\} \) constitutes a field.

\textbf{Proposition B.2.} If \( X \neq \emptyset \), then \((\{0, 1\}^X, +, \cdot)\) constitutes a commutative ring with unity, if \(+\) and \(\cdot\) denote pointwise addition and multiplication, respectively.

\textit{Proof.} First note that \(+\) and \(\cdot\) are associative, commutative, and distributive on \(\{0, 1\}^X\), as they are associative, commutative, and distributive on \(\{0, 1\}\). The neutral element of addition is \(\chi_\emptyset \equiv 0\); the neutral element of multiplication is \(\chi_X \equiv 1\). If \( f \in \{0, 1\}^X \), then \( f \) is its own additive inverse:

\[
\forall x \in X \quad (f + f)(x) = \begin{cases} 
0 + 0 = 0 & \text{if } f(x) = 0, \\
1 + 1 = 0 & \text{if } f(x) = 1,
\end{cases}
\]

completing the proof that \((\{0, 1\}^X, +, \cdot)\) constitutes a commutative ring with unity. \(\blacksquare\)
Proposition B.3. Let \( X \neq \emptyset \) and \( A, B \subseteq X \). Then the following holds:

(a) \( \chi_A + \chi_B = \chi_{A \Delta B} \).

(b) \( \chi_A \cdot \chi_B = \chi_{A \cap B} \).

Proof. (a): Let \( x \in X \). Then

\[
(\chi_A + \chi_B)(x) = 0 \iff (\chi_A(x) = 0 \land \chi_B(x) = 0) \lor (\chi_A(x) = 1 \land \chi_B(x) = 1)
\]

\[
\iff (x \notin A \land x \notin B) \lor (x \in A \land x \in B)
\]

\[
\iff x \notin A \cup B \lor x \in A \cap B
\]

\[
\iff x \notin A \Delta B \iff \chi_{A \Delta B}(x) = 0.
\]

(b): Let \( x \in X \). Then

\[
(\chi_A \cdot \chi_B)(x) = 0 \iff \chi_A(x) = 0 \lor \chi_B(x) = 0 \iff x \notin A \lor x \notin B
\]

\[
\iff x \notin A \cap B \iff \chi_{A \cap B}(x) = 0,
\]

completing the proof. \( \blacksquare \)

Lemma B.4. Let \( A, B, C \) be sets. Then one has the following rules:

(a) The symmetric difference is commutative: \( A \Delta B = B \Delta A \).

(b) The symmetric difference is associative: \((A \Delta B) \Delta C = A \Delta (B \Delta C)\).

(c) The following law of distributivity holds: \((A \Delta B) \cap C = (A \cap C) \Delta (B \cap C)\).

(d) \( A \cup B = (A \Delta B) \Delta (A \cap B) \).

(e) \( A \setminus B = A \Delta (A \cap B) \).

Proof. (a) is immediate from (B.1), since \( \cup \) is commutative. Rules (b) – (d) can be proved concisely by making use of Prop. B.2 and Prop. B.3: If we let \( X := A \cup B \cup C \) then \( A, B, C \in \mathcal{P}(X) \) and we are in the situation of the propositions. Let \( x \in X \). For (b), we compute

\[
x \in (A \Delta B) \Delta C \iff \chi_{(A \Delta B) \Delta C}(x) = 1 \iff ((\chi_A + \chi_B) + \chi_C)(x) = 1
\]

\[
\iff (\chi_A + (\chi_B + \chi_C))(x) = 1 \iff \chi_{A \Delta (B \Delta C)}(x) = 1
\]

\[
\iff x \in A \Delta (B \Delta C).
\]

For (c), we compute

\[
x \in (A \Delta B) \cap C \iff \chi_{(A \Delta B) \cap C}(x) = 1 \iff ((\chi_A + \chi_B) \cdot \chi_C)(x) = 1
\]

\[
\iff ((\chi_A \cdot \chi_C) + (\chi_B \cdot \chi_C))(x) = 1 \iff \chi_{(A \cap C) \Delta (B \cap C)}(x) = 1
\]

\[
\iff x \in (A \cap C) \Delta (B \cap C).
\]
For (d), we compute

\[ x \in (A \Delta B) \Delta (A \cap B) \iff \chi_{(A \Delta B) \Delta (A \cap B)}(x) = 1 \iff ((\chi_A + \chi_B) + \chi_A \cdot \chi_B)(x) = 1 \]
\[ \iff \chi_A(x) = 1 \lor \chi_B(x) = 1 \iff x \in A \cup B. \]

Finally, (e) holds, since

\[ A \Delta (A \cap B) = (A \setminus (A \cap B)) \cup ((A \cap B) \setminus A) = (A \setminus B) \cup \emptyset = A \setminus B, \]

which completes the proof.

Proposition B.5. Let \( X \) be a set, \( \mathcal{E} \subseteq \mathcal{P}(X) \). Then \( (\mathcal{E}, \Delta, \cap) \) is a ring in the sense of Def. 1.13(b) if, and only if, it is a ring in the sense of algebra (i.e. in the sense of [Phi16a, Def. C.7]).

Proof. Let \( (\mathcal{E}, \Delta, \cap) \) be a ring in the sense of Def. 1.13(b). If \( A, B \in \mathcal{E} \), then \( A \Delta B \in \mathcal{E} \) and \( A \cap B \in \mathcal{E} \). By Lem. B.4(a),(b), \( \Delta \) is commutative and associative. The empty set is neutral for \( \Delta \), and \( A \Delta A = \emptyset \) shows that every set is inverse to itself with respect to \( \Delta \). Thus, \( (\mathcal{E}, \Delta) \) is a commutative group. Since \( \cap \) is also associative and the law of distributivity holds by Lem. B.4(c), \( (\mathcal{E}, \Delta, \cap) \) is a ring in the sense of algebra. Conversely, let \( (\mathcal{E}, \Delta, \cap) \) be a ring in the sense of algebra. Then \( \mathcal{E} \neq \emptyset \), i.e. there is \( A \in \mathcal{E} \). Then \( \emptyset = A \Delta A \in \mathcal{E} \). If \( A, B \in \mathcal{E} \), then \( A \cup B = (A \Delta B) \Delta (A \cap B) \in \mathcal{E} \) by Lem. B.4(d) and \( A \setminus B = A \Delta (A \cap B) \) by Lem. B.4(e), showing \( (\mathcal{E}, \Delta, \cap) \) to be a ring in the sense of Def. 1.13(b).

Remark B.6. The reason that one can not replace \( \Delta \) by \( \cup \) in Prop. B.5 lies in the fact of missing inverse elements: While \( \emptyset \) is still neutral for \( \cup \), if \( A \neq \emptyset \), then there does not exist a set \( B \) such that \( A \cup B = \emptyset \).

C Initial and Final \( \sigma \)-Algebras

In [Phi16b, Sec. F], the construction of initial and final topological spaces was introduced. Analogous constructions exist for measurable spaces and we will briefly study them in the present section (analogous to topology, initial \( \sigma \)-algebras include the construction of product \( \sigma \)-algebras and trace \( \sigma \)-algebras (cf. Ex. C.4(a),(b) below)).

Definition C.1. Let \( X \) be a set and let \( (X_i, \mathcal{A}_i)_{i \in I} \) be a family of measurable spaces, \( I \neq \emptyset \).

(a) Given a family of functions \( (f_i)_{i \in I}, f_i : X \rightarrow X_i \), the initial \( \sigma \)-algebra on \( X \) with respect to the family \( (f_i)_{i \in I} \) is the smallest \( \sigma \)-algebra \( \mathcal{A} \) on \( X \) that makes all \( f_i \) \( \mathcal{A} \)-\( \mathcal{A}_i \)-measurable (i.e. \( \mathcal{A} \) is the intersection of all \( \sigma \)-algebras on \( X \) that make all \( f_i \) measurable – this intersection is well-defined, since the \( \sigma \)-algebra \( \mathcal{P}(X) \) on \( X \) always makes all \( f_i \) measurable).
(b) Given a family of functions \((f_i)_{i \in I}, f_i : X_i \rightarrow X\), the final \(\sigma\)-algebra on \(X\) with respect to the family \((f_i)_{i \in I}\) is

\[
\mathcal{A} := \left\{ A \subseteq X : \forall i \in I f_i^{-1}(A) \in \mathcal{A}_i \right\}.
\] (C.1)

We will see in Lem. C.2(b) below (in generalization of Prop. 1.55(a)) that \(\mathcal{A}\) is, indeed, a \(\sigma\)-algebra, and that it is the largest \(\sigma\)-algebra on \(X\) that makes all \(f_i\) measurable.\(^1\)

**Lemma C.2.** Let \(X\) be a set and let \((X_i, \mathcal{A}_i)_{i \in I}\) be a family of measurable spaces, \(I \neq \emptyset\).

(a) Given a family of functions \((f_i)_{i \in I}, f_i : X \rightarrow X_i\), the set

\[
\mathcal{E} := \left\{ f_i^{-1}(A_i) : A_i \in \mathcal{A}_i, i \in I \right\}
\] (C.2)

is a generator of the initial \(\sigma\)-algebra \(\mathcal{A}\) on \(X\) with respect to the family \((f_i)_{i \in I}\).

(b) Given a family of functions \((f_i)_{i \in I}, f_i : X_i \rightarrow X\), the final \(\sigma\)-algebra \(\mathcal{A}\) on \(X\) with respect to the family \((f_i)_{i \in I}\) as defined in (C.1) is, indeed, a \(\sigma\)-algebra on \(X\), and it is the largest \(\sigma\)-algebra on \(X\) that makes all \(f_i\) measurable.

**Proof.** (a) is immediate, since each \(\sigma\)-algebra \(\mathcal{A}\) on \(X\) such that all \(f_i, i \in I\), are \(\mathcal{A}\)-\(\mathcal{A}_i\)-measurable must contain \(\mathcal{E}\).

(b): For each \(i \in I\), since \(\emptyset = f_i^{-1}(\emptyset) \in \mathcal{A}_i\), one has \(\emptyset \in \mathcal{A}\). If \(A \in \mathcal{A}\), then \(X \setminus A \in \mathcal{A}\), since, for each \(i \in I\), \(f_i^{-1}(X \setminus A) = X_i \setminus f_i^{-1}(A) \in \mathcal{A}_i\), due to \(\mathcal{A}_i\) being a \(\sigma\)-algebra. If \((A_n)_{n \in \mathbb{N}}\) is a sequence of sets in \(\mathcal{A}\), then, for each \(i \in I\),

\[
f_i^{-1}\left(\bigcup_{n=1}^{\infty} A_n\right) = \bigcap_{n=1}^{\infty} f_i^{-1}(A_n) \in \mathcal{A}_i,
\]

as \(\mathcal{A}_i\) is a \(\sigma\)-algebra. Thus, \(\bigcup_{n=1}^{\infty} A_n \in \mathcal{A}\), which completes the proof that \(\mathcal{A}\) is a \(\sigma\)-algebra on \(X\). It is immediate from the definition of \(\mathcal{A}\) that each \(f_i, i \in I\), is \(\mathcal{A}_i\)-\(\mathcal{A}\)-measurable. To see that \(\mathcal{A}\) is the largest \(\sigma\)-algebra on \(X\) with this property, we still need to show that every \(\sigma\)-algebra \(\mathcal{B}\) on \(X\) making all \(f_i, \mathcal{A}_i\)-\(\mathcal{B}\)-measurable is contained in \(\mathcal{A}\). To this end, let \(\mathcal{B}\) be such a \(\sigma\)-algebra on \(X\). If \(B \in \mathcal{B}\), then, for each \(i \in I\), \(f_i^{-1}(B) \in \mathcal{A}_i\), i.e. \(B \in \mathcal{A}\), showing \(\mathcal{B} \subseteq \mathcal{A}\).

**Proposition C.3.** Let \(X\) be a set and let \((X_i, \mathcal{A}_i)_{i \in I}\) be a family of measurable spaces, \(I \neq \emptyset\).

(a) Given a family of functions \((f_i)_{i \in I}, f_i : X \rightarrow X_i\), let \(\mathcal{A}\) denote the initial \(\sigma\)-algebra on \(X\) with respect to the family \((f_i)_{i \in I}\). Then \(\mathcal{A}\) has the property that each map \(g : Z \rightarrow X\) from a measurable space \((Z, \mathcal{A}_Z)\) into \(X\) is measurable if, and only if, each map \((f_i \circ g) : Z \rightarrow X_i\) is measurable. Moreover, \(\mathcal{A}\) is the only \(\sigma\)-algebra on \(X\) with this property.

\(^1\)In the language of so-called *Category Theory*, we can say that the category of measurable spaces (analogous to the category of topological spaces) has initial and final objects.
(b) Given a family of functions \((f_i)_{i \in I}\), \(f_i : X_i \rightarrow X\), let \(\mathcal{A}\) denote the final \(\sigma\)-algebra on \(X\) with respect to the family \((f_i)_{i \in I}\). Then \(\mathcal{A}\) has the property that each map \(g : X \rightarrow Z\) from \(X\) into a measurable space \((Z, \mathcal{A}_Z)\) is measurable if, and only if, each map \((g \circ f_i) : X_i \rightarrow Z\) is measurable. Moreover, \(\mathcal{A}\) is the only \(\sigma\)-algebra on \(X\) with this property.

**Proof.** (a): If \(g\) is \(\mathcal{A}_Z\)-\(\mathcal{A}\)-measurable, then each composition \(f_i \circ g\), \(i \in I\), is \(\mathcal{A}_Z\)-\(\mathcal{A}\)-measurable. For the converse, assume each \(f_i \circ g\), \(i \in I\), to be \(\mathcal{A}_Z\)-\(\mathcal{A}\)-measurable. According to Lem. C.2(a), it suffices to show \(g^{-1}(f_i^{-1}(A)) \in \mathcal{A}_Z\) for each \(i \in I\) and each \(A \in \mathcal{A}\). Since \(g^{-1}(f_i^{-1}(A)) = (f_i \circ g)^{-1}(A)\) and \(f_i \circ g\) is measurable, the proof is, thus, complete. Now let \(\mathcal{B}\) be an arbitrary \(\sigma\)-algebra on \(X\) with the property stated in the hypothesis. Letting \((Z, \mathcal{A}_Z) := (X, \mathcal{B})\) and \(g := \text{Id}_X\), we see that each \(f_i = f_i \circ g\) is \(\mathcal{B}\)-\(\mathcal{A}\)-measurable, implying \(\mathcal{A} \subseteq \mathcal{B}\). Now let \(\mathcal{A}'\) be an arbitrary \(\sigma\)-algebra on \(X\) that makes all \(f_i\) \(\mathcal{A}'\)-\(\mathcal{A}\)-measurable. Letting \((Z, \mathcal{A}_Z) := (X, \mathcal{A}')\), we see that \(g := \text{Id}_X\) is \(\mathcal{A}'\)-\(\mathcal{B}\)-measurable (since each \(f_i = \text{Id}_X\circ f_i\) is \(\mathcal{A}'\)-\(\mathcal{A}\)-measurable) i.e., for each \(B \in \mathcal{B}\), we have \(g^{-1}(B) = B \in \mathcal{A}'\), showing \(\mathcal{B} \subseteq \mathcal{A}'\) and \(\mathcal{B} \subseteq \mathcal{A}\), also completing the proof of \(\mathcal{B} = \mathcal{A}\).

(b): If \(g\) is \(\mathcal{A}_Z\)-\(\mathcal{A}\)-measurable, then each composition \(g \circ f_i\), \(i \in I\), is \(\mathcal{A}_Z\)-\(\mathcal{A}\)-measurable. For the converse, assume each \(g \circ f_i\), \(i \in I\), to be \(\mathcal{A}_Z\)-\(\mathcal{A}\)-measurable. If \(A \in \mathcal{A}_Z\), then, for each \(i \in I\), \(f_i^{-1}(g^{-1}(A)) \in \mathcal{A}_i\), showing \(g^{-1}(A) \in \mathcal{A}\) according to (C.1). Thus, \(g\) is \(\mathcal{A}_Z\)-\(\mathcal{A}\)-measurable. Now let \(\mathcal{B}\) be an arbitrary \(\sigma\)-algebra on \(X\) with the property stated in the hypothesis. Letting \((Z, \mathcal{A}_Z) := (X, \mathcal{B})\) and \(g := \text{Id}_X\), we see that each \(f_i = g \circ f_i\) is \(\mathcal{A}_i\)-\(\mathcal{A}\)-measurable, implying \(\mathcal{B} \subseteq \mathcal{A}\). Now let \(\mathcal{A}'\) be an arbitrary topology on \(X\) that makes all \(f_i\) \(\mathcal{A}'\)-\(\mathcal{A}\)-measurable. Letting \((Z, \mathcal{A}_Z) := (X, \mathcal{A}')\), we see that \(g := \text{Id}_X\) is \(\mathcal{B}\)-\(\mathcal{A}'\)-measurable (since each \(f_i = f_i \circ \text{Id}_X\) is \(\mathcal{A}_i\)-\(\mathcal{A}'\)-measurable) i.e., for each \(A \in \mathcal{A}'\), we have \((g^{-1}(A) = A \in \mathcal{B}\), showing \(\mathcal{A}' \subseteq \mathcal{B}\) and \(\mathcal{A} \subseteq \mathcal{B}\), also completing the proof of \(\mathcal{B} = \mathcal{A}\).

**Example C.4.** (a) The product \(\sigma\)-algebra on \(X = \prod_{i \in I} X_i\) (cf. Def. 1.90) is the initial \(\sigma\)-algebra with respect to the projections \((\pi_i)_{i \in I}\), \(\pi_i : X \rightarrow X_i\) (as is clear from Lem. C.2(a)).

(b) The trace \(\sigma\)-algebra \(\mathcal{A}|B\) on \(B \subseteq X\), where \((X, \mathcal{A})\) is a measurable space (cf. Prop. 1.55(c)), is the initial \(\sigma\)-algebra with respect to the identity inclusion map \(\iota : B \rightarrow X\), \(\iota(x) := x\): This is also clear from Lem. C.2(a), since

\[
\mathcal{A}|B = \{ A \cap B : A \in \mathcal{A} \} = \{ \iota^{-1}(A) : A \in \mathcal{A} \}.
\]

(c) In [Phi16b, Ex. F.4(c)], we considered the quotient topology as an example for a final topology. Analogously, we consider the quotient \(\sigma\)-algebra as an example for a final \(\sigma\)-algebra (however, in contrast to quotient topologies, quotient \(\sigma\)-algebras do not seem to play a particularly important role in the literature): Let \((X, \mathcal{A})\) be a measurable space and let \(\sim\) be an equivalence relation on \(X\). Moreover, let \(Y := X/\sim = \{ [x] : x \in X \}\) be the corresponding quotient set (i.e. the set of corresponding equivalence classes). Then the quotient \(\sigma\)-algebra on \(Y\) with respect
to \sim$, denoted $\mathcal{A}/\sim$, is defined as the final \(\sigma\)-algebra with respect to the canonical projection\footnote{\begin{equation}
\pi : X \rightarrow Y, \quad \pi(x) := \lfloor x \rfloor.
\end{equation}}

Thus, by (C.1),

\[ \mathcal{A}/\sim = \{ A \subseteq Y : \pi^{-1}(A) \in \mathcal{A} \}. \]

\section*{D Riemann Integral on Intervals in $\mathbb{R}^n$}

In generalization of [Phi16a, Sec. 10], we will define Riemann integrals for suitable functions $f : I \rightarrow \mathbb{R}$, where $I = [a, b]$ is a subset of $\mathbb{R}^n$, $a, b \in \mathbb{R}^n$, $a \leq b$.

In generalization of [Phi16a, Sec. 10], given a nonnegative function $f : I \rightarrow \mathbb{R}^+_0$, we aim to compute the \((n + 1)\)-dimensional volume $\int_I f$ of the set “under the graph” of $f$, i.e. of the set

\[ \{(x_1, \ldots, x_n, x_{n+1}) \in \mathbb{R}^{n+1} : (x_1, \ldots, x_n) \in I \text{ and } 0 \leq x_{n+1} \leq f(x_1, \ldots, x_n)\}. \]  

(D.1)

Moreover, for functions $f : I \rightarrow \mathbb{R}$ that are not necessarily nonnegative, we would like to count volumes of sets of the form (D.1) (which are below the graph of $f$ and above the set $I \equiv \{(x_1, \ldots, x_n, 0) \in \mathbb{R}^{n+1} : (x_1, \ldots, x_n) \in I\} \subseteq \mathbb{R}^{n+1}$) with a positive sign, whereas we would like to count volumes of sets above the graph of $f$ and below the set $I$ with a negative sign. In other words, making use of the positive and negative parts $f^+ = \max(f, 0)$ and $f^- = \max(-f, 0)$ of $f = f^+ - f^-$, the integral needs to satisfy

\[ \int_I f = \int_I f^+ - \int_I f^- \]  

(D.2)

As in [Phi16a, Sec. 10], we restrict ourselves to bounded functions $f : I \rightarrow \mathbb{R}$.

As in [Phi16a, Sec. 10.1], the basic idea for the definition of the Riemann integral $\int_I f$ is to decompose the interval $I$ into small intervals $I_1, \ldots, I_N$ and approximate $\int_I f$ by the finite sum $\sum_{j=1}^N f(x_j)|I_j|$, where $x_j \in I_j$ and $|I_j|$ denotes the volume of the interval $I_j$.

Define $\int_I f$ as the limit of such sums as the size of the $I_j$ tends to zero (if the limit exists). However, to carry out this idea precisely and rigorously is somewhat cumbersome, for example due to the required notation.

As each $n$-dimensional interval $I$ is a product of one-dimensional intervals, we will obtain our decompositions of $I$ from decompositions of one-dimensional intervals (the sides of $I$).

\textbf{Notation D.1.} In generalization of [Phi16a, Def. 10.2], if $a, b \in \mathbb{R}^n$, $n \in \mathbb{N}$, $a \leq b$, and $I := [a, b] = [a_1, b_1] \times \cdots \times [a_n, b_n]$, then we define

\[ |I| := \prod_{j=1}^n (b_j - a_j) = \prod_{j=1}^n |a_j - b_j| = \prod_{j=1}^n |I_j| \quad (I_j := [a_j, b_j]). \]  

(D.3)
Definition D.2. Recall the notion of partition for a 1-dimensional interval \([a, b] \subseteq \mathbb{R}\) from [Phi16a, Def. 10.3]. We now generalize this notion to \(n\)-dimensional intervals, \(n \in \mathbb{N}\): Given an interval \(I := [a, b] \subseteq \mathbb{R}^n\), \(a, b \in \mathbb{R}^n\), \(a < b\), i.e. \(I = [a, b] = [a_1, b_1] \times \cdots \times [a_n, b_n]\), a partition \(\Delta\) of \(I\) is given by (1-dimensional) partitions \(\Delta_k = (x_{k,0}, \ldots, x_{k,N_k})\) of \([a_k, b_k]\), \(k \in \{1, \ldots, n\}\), \(N_k \in \mathbb{N}\). Given such a partition \(\Delta\) of \(I\), for each \((k_1, \ldots, k_n) \in P(\Delta) := \coprod_{k=1}^{n} \{1, \ldots, N_k\}\), define

\[
I_{(j_1, \ldots, j_n)} := \prod_{k=1}^{n} [x_{k,j_k-1}, x_{k,j_k}] = [x_{1,j_1-1}, x_{1,j_1}] \times \cdots \times [x_{n,j_n-1}, x_{n,j_n}],
\]

(D.4)

The number

\[
|\Delta| := \max \{|\Delta_k| : k \in \{1, \ldots, n\}\}
= \max \{x_{k,l} - x_{k,l-1} : k \in \{1, \ldots, n\}, l \in \{1, \ldots, N_k\}\},
\]

(D.5)

is called the \textit{mesh size} of \(\Delta\).

Moreover, if each \(\Delta_k\) is tagged by \((t_{k,1}, \ldots, t_{k,N_k}) \in \mathbb{R}^{N_k}\) such that \(t_{k,j} \in [x_{k,j-1}, x_{k,j}]\) for each \(j \in \{1, \ldots, N_j\}\), then \(\Delta\) is tagged by \((t_p)_{p \in P(\Delta)}\), where

\[
t_{(j_1, \ldots, j_n)} := (t_{1,j_1}, \ldots, t_{n,j_n}) \in I_{(j_1, \ldots, j_n)} \text{ for each } (j_1, \ldots, j_n) \in P(\Delta).
\]

(D.6)

Remark D.3. If \(\Delta\) is a partition of \(I = [a, b] \subseteq \mathbb{R}^n\), \(n \in \mathbb{N}\), \(a, b \in \mathbb{R}^n\), \(a < b\), as in Def. D.2 above, then

\[
I = \bigcup_{p \in P(\Delta)} I_p
\]

(D.7)

and

\[
|I_p \cap I_q| = 0 \text{ for each } p, q \in P(\Delta) \text{ such that } p \neq q,
\]

(D.8)

since, for \(p \neq q\), \(I_p \cap I_q\) is either empty or it is an interval such that one side consists of precisely one point. Moreover, as a consequence of (D.7) and (D.8):

\[
|I| = \sum_{p \in P(\Delta)} |I_p|.
\]

(D.9)

Definition D.4. Consider an interval \(I := [a, b] \subseteq \mathbb{R}^n\), \(n \in \mathbb{N}\), \(a, b \in \mathbb{R}^n\), \(a < b\), with a partition \(\Delta\) of \(I\) as in Def. D.2. In generalization of [Phi16a, Def. 10.4], given a function \(f : I \rightarrow \mathbb{R}\) that is bounded, define, for each \(p \in P(\Delta)\),

\[
m_p := m_p(f) := \inf \{f(x) : x \in I_p\}, \quad M_p := M_p(f) := \sup \{f(x) : x \in I_p\},
\]

(D.10)

and

\[
r(\Delta, f) := \sum_{p \in P(\Delta)} m_p |I_p|,
R(\Delta, f) := \sum_{p \in P(\Delta)} M_p |I_p|,
\]

(D.11a) (D.11b)
where \( r(\Delta, f) \) is called the lower Riemann sum and \( R(\Delta, f) \) is called the upper Riemann sum associated with \( \Delta \) and \( f \). If \( \Delta \) is tagged by \( \tau := (t_p)_{p \in P(\Delta)} \), then we also define the intermediate Riemann sum

\[
\rho(\Delta, f) := \sum_{p \in P(\Delta)} f(t_p)|I_p|.
\]  

(D.11c)

**Definition D.5.** Let \( I = [a, b] \subseteq \mathbb{R}^n \) be an interval, \( a, b \in \mathbb{R}^n \), \( n \in \mathbb{N} \), \( a < b \), and suppose \( f : I \rightarrow \mathbb{R} \) is bounded.

(a) Define

\[
J_*(f, I) := \sup \{ r(\Delta, f) : \Delta \text{ is a partition of } I \},
\]

(D.12a)

\[
J^*(f, I) := \inf \{ R(\Delta, f) : \Delta \text{ is a partition of } I \}.
\]

(D.12b)

We call \( J_*(f, I) \) the lower Riemann integral of \( f \) over \( I \) and \( J^*(f, I) \) the upper Riemann integral of \( f \) over \( I \).

(b) The function \( f \) is called Riemann integrable over \( I \) if, and only if, \( J_*(f, I) = J^*(f, I) \).

If \( f \) is Riemann integrable over \( I \), then

\[
\int_I f(x) \, dx := \int_I f := J_*(f, I) = J^*(f, I)
\]

(D.13)

is called the Riemann integral of \( f \) over \( I \). The set of all functions \( f : I \rightarrow \mathbb{R} \) that are Riemann integrable over \( I \) is denoted by \( \mathcal{R}(I, \mathbb{R}) \) or just by \( \mathcal{R}(I) \).

(c) The function \( g : I \rightarrow \mathbb{C} \) is called Riemann integrable over \( I \) if, and only if, both \( \text{Re} \, g \) and \( \text{Im} \, g \) are Riemann integrable. The set of all Riemann integrable functions \( g : I \rightarrow \mathbb{C} \) is denoted by \( \mathcal{R}(I, \mathbb{C}) \). If \( g \in \mathcal{R}(I, \mathbb{C}) \), then

\[
\int_I g := \left( \int_I \text{Re} \, g, \int_I \text{Im} \, g \right) = \int_I \text{Re} \, g + i \int_I \text{Im} \, g \in \mathbb{C}
\]

(D.14)

is called the Riemann integral of \( g \) over \( I \).

**Remark D.6.** If \( I \) and \( f \) are as before, then (D.10) implies

\[
m_p(f) = -M_p(-f) \quad \text{and} \quad m_p(-f) = -M_p(f),
\]

(D.15a)

(D.11) implies

\[
r(\Delta, f) = -R(\Delta, -f) \quad \text{and} \quad r(\Delta, -f) = -R(\Delta, f),
\]

(D.15b)

and (D.12) implies

\[
J_*(f, I) = -J^*(-f, I) \quad \text{and} \quad J_*(-f, I) = -J^*(f, I).
\]

(D.15c)

**Example D.7.** Let \( I = [a, b] \subseteq \mathbb{R}^n \) be an interval, \( a, b \in \mathbb{R}^n \), \( n \in \mathbb{N} \), \( a < b \).
D RIEMANN INTEGRAL ON INTERVALS IN $\mathbb{R}^N$

(a) Analogous to [Phi16a, Ex. 10.7(a)], if $f : I \rightarrow \mathbb{R}$ is constant, i.e. $f \equiv c$ with $c \in \mathbb{R}$, then $f \in \mathcal{R}(I)$ and

$$\int_I f = c |I|.$$  \hspace{1cm} (D.16)

We have, for each partition $\Delta$ of $I$,

$$r(\Delta, f) = \sum_{p \in P(\Delta)} m_p |I_p| = c \sum_{p \in P(\Delta)} |I_p| = c |I| = \sum_{p \in P(\Delta)} M_p |I_p| = R(\Delta, f),$$  \hspace{1cm} (D.17)

proving $J^*(f, I) = c |I| = J^*(f, I)$.

(b) An example of a function that is not Riemann integrable is given by the $n$-dimensional version of the Dirichlet function of [Phi16a, Ex. 10.7(b)], i.e. by

$$f : I \rightarrow \mathbb{R}, \quad f(x) := \begin{cases} 0 & \text{for } x \in I \setminus \mathbb{Q}^n, \\ 1 & \text{for } x \in I \cap \mathbb{Q}^n. \end{cases}$$  \hspace{1cm} (D.18)

Since $r(\Delta, f) = 0$ and $R(\Delta, f) = \sum_{p \in P(\Delta)} |I_p| = |I|$ for every partition $\Delta$ of $I$, one obtains $J^*(f, I) = 0 \neq |I| = J^*(f, I)$, showing that $f \notin \mathcal{R}(I)$.

For a general characterization of Riemann integrable functions, see Th. 2.26(b).

**Definition D.8.** Recall the notions of refinement and superposition of partitions of a 1-dimensional interval $[a, b] \subseteq \mathbb{R}$ from [Phi16a, Def. 10.8]. We now generalize both notions to partitions of $n$-dimensional intervals, $n \in \mathbb{N}$:

(a) If $\Delta$ is a partition of $[a, b] \subseteq \mathbb{R}^n$, $n \in \mathbb{N}$, $a, b \in \mathbb{R}^n$, $a < b$, as in Def. D.2, then another partition $\Delta'$ of $[a, b]$ given by partitions $\Delta'_k$ of $[a_k, b_k]$, $k \in \{1, \ldots, n\}$, respectively, is called a refinement of $\Delta$ if, and only if, each $\Delta'_k$ is a (1-dimensional) refinement of $\Delta_k$ in the sense of [Phi16a, Def. 10.8(a)].

(b) If $\Delta$ and $\Delta'$ are partitions of $[a, b] \subseteq \mathbb{R}^n$, $n \in \mathbb{N}$, $a, b \in \mathbb{R}^n$, $a < b$, then the superposition $\Delta + \Delta'$ is given by the (1-dimensional) superpositions $\Delta_k + \Delta'_k$ of the $[a_k, b_k]$ in the sense of [Phi16a, Def. 10.8(b)]. Note that the superposition of $\Delta$ and $\Delta'$ is always a common refinement of $\Delta$ and $\Delta'$.

**Lemma D.9.** Let $n \in \mathbb{N}$, $a, b \in \mathbb{R}^n$, $a < b$, $I := [a, b]$, and suppose $f : I \rightarrow \mathbb{R}$ is bounded with $M := \|f\|_{\sup} \in \mathbb{R}^+_0$. Let $\Delta'$ be a partition of $I$ and define

$$\alpha := \sum_{k=1}^n \#(\nu(\Delta'_k) \setminus \{a_k, b_k\})$$  \hspace{1cm} (D.19)

i.e. $\alpha$ is the number of interior nodes that occur in the $\Delta'_k$. Then, for each partition $\Delta$ of $I$, the following holds:

$$r(\Delta, f) \leq r(\Delta + \Delta', f) \leq r(\Delta, f) + 2 \alpha M\phi(I) |\Delta|,$$  \hspace{1cm} (D.20a)

$$R(\Delta, f) \geq R(\Delta + \Delta', f) \geq R(\Delta, f) - 2 \alpha M\phi(I) |\Delta|.$$  \hspace{1cm} (D.20b)
where
\[ \phi(I) := \max \left\{ \frac{|I|}{b_k - a_k} : k \in \{1, \ldots, n\} \right\} \] (D.21)
is the maximal volume of the \((n-1)\)-dimensional faces of \(I\).

**Proof.** We carry out the proof of (D.20a) – the proof of (D.20b) can be conducted completely analogous. Consider the case \(\alpha = 1\). Then there is a unique \(k_0 \in \{1, \ldots, n\}\) such that there exists \(\xi \in \nu(\Delta'_{k_0})\backslash\{a_{k_0}, b_{k_0}\}\). If \(\xi \in \nu(\Delta_{k_0})\), then \(\Delta + \Delta' = \Delta\), and (D.20a) is trivially true. If \(\xi \notin \nu(\Delta_{k_0})\), then \(x_{k_0,l-1} < \xi < x_{k_0,l}\) for a suitable \(l \in \{1, \ldots, N_{k_0}\}\). Recalling the notation from Def. D.2, we let
\[ P_l(\Delta) := \{(j_1, \ldots, j_n) \in P(\Delta) : j_{k_0} = l\}, \] (D.22)
and define, for each \((j_1, \ldots, j_n) \in P_l(\Delta)\),
\[ I'_{j_1,\ldots,j_n} := \prod_{k=1}^{k_0-1} [x_{k,j_k-1}, x_{k,j_k}] \times [x_{k_0,l-1}, \xi] \times \prod_{k=k_0+1}^{n} [x_{k,j_k-1}, x_{k,j_k}], \] (D.23a)
\[ I''_{j_1,\ldots,j_n} := \prod_{k=1}^{k_0-1} [x_{k,j_k-1}, x_{k,j_k}] \times [\xi, x_{k_0,l}] \times \prod_{k=k_0+1}^{n} [x_{k,j_k-1}, x_{k,j_k}], \] (D.23b)
and
\[ m'_p := \inf \{ f(x) : x \in I'_{p} \}, \quad m''_p := \inf \{ f(x) : x \in I''_{p} \} \quad \text{for each } p \in P_l(\Delta). \] (D.24)
Then we obtain
\[ r(\Delta + \Delta', f) - r(\Delta, f) = \sum_{p \in P_l(\Delta)} (m'_p |I'_p| + m''_p |I''_p| - m_p |I_p|) \]
\[ = \sum_{p \in P_l(\Delta)} (|I'_p| - m'_p |I'_p|) + (|I''_p| - m''_p |I''_p|). \] (D.25)
Together with the observation
\[ 0 \leq m'_p - m_p \leq 2M, \quad 0 \leq m''_p - m_p \leq 2M, \] (D.26)
(D.25) implies
\[ 0 \leq r(\Delta + \Delta', f) - r(\Delta, f) \leq 2M \sum_{p \in P_l(\Delta)} |I_p| \]
\[ = 2M (x_{k_0,l} - x_{k_0,l-1}) \frac{|I|}{b_{k_0} - a_{k_0}} \]
\[ \leq 2M |\Delta_{k_0}| \phi(I) \leq 2M |\Delta| \phi(I). \] (D.27)
The general form of (D.20a) now follows by an induction on \(\alpha\).

**Theorem D.10.** Let \(n \in \mathbb{N}\), \(a, b \in \mathbb{R}^n\), \(a < b\), \(I := [a,b]\), and let \(f : I \to \mathbb{R}\) be bounded.
(a) Suppose $\Delta$ and $\Delta'$ are partitions of $I$ such that $\Delta'$ is a refinement of $\Delta$. Then
\[ r(\Delta, f) \leq r(\Delta', f), \quad R(\Delta, f) \geq R(\Delta', f). \quad \text{(D.28)} \]

(b) For arbitrary partitions $\Delta$ and $\Delta'$, the following holds:
\[ r(\Delta, f) \leq R(\Delta', f). \quad \text{(D.29)} \]

(c) $J_*(f, I) \leq J^*(f, I)$.

(d) For each sequence of partitions $(\Delta^k)_{k \in \mathbb{N}}$ of $I$ such that $\lim_{k \to \infty} |\Delta^k| = 0$, one has
\[ \lim_{k \to \infty} r(\Delta^k, f) = J_*(f, I), \quad \lim_{k \to \infty} R(\Delta^k, f) = J^*(f, I). \quad \text{(D.30)} \]

In particular, if $f \in \mathcal{R}(I)$, then
\[ \lim_{k \to \infty} r(\Delta^k, f) = \lim_{k \to \infty} R(\Delta^k, f) = \int_I f. \quad \text{(D.31a)} \]

and if $f \in \mathcal{R}(I)$ and the $\Delta^k$ are tagged, then also
\[ \lim_{k \to \infty} \rho(\Delta^k, f) = \int_I f. \quad \text{(D.31b)} \]

(e) If there exists $\alpha \in \mathbb{R}$ such that
\[ \alpha = \lim_{k \to \infty} \rho(\Delta^k, f) \quad \text{(D.32)} \]

for each sequence of tagged partitions $(\Delta^k)_{k \in \mathbb{N}}$ of $I$ such that $\lim_{k \to \infty} |\Delta^k| = 0$, then $f \in \mathcal{R}(I)$ and $\alpha = \int_I f$.

Proof. (a): If $\Delta'$ is a refinement of $\Delta$, then $\Delta' = \Delta + \Delta'$. Thus, (D.28) is immediate from (D.20).

(b): This also follows from (D.20):
\[ r(\Delta, f) \leq r(\Delta + \Delta', f) \leq R(\Delta + \Delta', f) \leq R(\Delta', f). \quad \text{(D.33)} \]

(c): One just combines (D.12) with (b).

(d): Let $(\Delta^k)_{k \in \mathbb{N}}$ be a sequence of partitions of $I$ such that $\lim_{k \to \infty} |\Delta^k| = 0$, and let $\Delta'$ be an arbitrary partition of $I$ with numbers $\alpha$, $M$, and $\phi(I)$ defined as in Lem. D.9. Then, according to (D.20a):
\[ r(\Delta^k, f) \leq r(\Delta^k + \Delta', f) \leq r(\Delta^k, f) + 2\alpha M \phi(I) |\Delta^k| \quad \text{for each } k \in \mathbb{N}. \quad \text{(D.34)} \]

From (b), we conclude the sequence $(r(\Delta^k, f))_{k \in \mathbb{N}}$ is bounded. Recall from [Phi16a, Th. 7.27] that each bounded sequence $(t_k)_{k \in \mathbb{N}}$ in $\mathbb{R}$ has a smallest cluster point $t_* \in \mathbb{R}$,
and a largest cluster point $t^* \in \mathbb{R}$. Moreover, by [Phi16a, Prop. 7.26], there exists at least one subsequence converging to $t_*$ and at least one subsequence converging to $t^*$, and, in particular, the sequence converges if, and only if, $t_* = t^* = \lim_{k \to \infty} t_k$. We can apply this to the present situation: Suppose $(r(\Delta^k, f))_{k \in \mathbb{N}}$ is a converging subsequence of $(r(\Delta^k, f))_{k \in \mathbb{N}}$ with

$$
\beta := \lim_{k \to \infty} r(\Delta^k, f).
$$

First note $\beta \leq J_*(f, I)$ due to the definition of $J_*(f, I)$. Moreover, (D.34) implies $\lim_{\Delta \to \Delta/l} r(\Delta^k + \Delta', f) = \beta$. Since $r(\Delta', f) \leq r(\Delta^k + \Delta', f)$ and $\Delta'$ is arbitrary, we obtain $J_*(f, I) \leq \beta$, i.e. $J_*(f, I) = \beta$. Thus, we have shown that every subsequence of $(r(\Delta^k, f))_{k \in \mathbb{N}}$ converges to $\beta$, showing

$$
J_*(f, I) = \beta = \lim_{k \to \infty} r(\Delta^k, f)
$$

as claimed. In the same manner, one conducts the proof of $J^*(f, I) = \lim_{\Delta \to \infty} R(\Delta^k, f)$. Then (D.31a) is immediate from the definition of Riemann integrability, and (D.31b) follows from (D.31a), since (D.11) implies

$$
r(\Delta, f) \leq \rho(\Delta, f) \leq R(\Delta, f)
$$

for each tagged partition $\Delta$ of $I$.

(e): Due to the definition of inf and sup,

$$
\forall \emptyset \neq \Delta \subseteq I \quad \forall \epsilon > 0 \quad \exists \ t_* \in A \quad f(t_*) < \inf\{f(x) : x \in A\} + \epsilon, \quad (D.38a)
$$

$$
\forall \emptyset \neq \Delta \subseteq I \quad \forall \epsilon > 0 \quad \exists \ t^* \in A \quad f(t^*) > \sup\{f(x) : x \in A\} - \epsilon. \quad (D.38b)
$$

In consequence, for each partition $\Delta$ of $I$ and each $\epsilon > 0$, there are tags $\tau_* := (t_*)_{p \in P(\Delta)}$ and $\tau^* := (t^*)_{p \in P(\Delta)}$ such that

$$
\rho(\Delta, \tau_*, f) < r(\Delta, f) + \epsilon \quad \land \quad \rho(\Delta, \tau^*, f) > R(\Delta, f) - \epsilon. \quad (D.39)
$$

Now let $(\Delta^k)_{k \in \mathbb{N}}$ be a sequence of partitions of $I$ such that $\lim_{k \to \infty} |\Delta^k| = 0$. According to the above, for each $\Delta^k$, there are tags $\tau^k_* := (t^k_*)_{p \in P(\Delta^k)}$ and $\tau^k* := (t^k^*)_{p \in P(\Delta^k)}$ such that

$$
\forall \ k \in \mathbb{N} \quad \left(\rho(\Delta^k, \tau^k_*, f) < r(\Delta^k, f) + \frac{1}{k} \quad \land \quad \rho(\Delta^k, \tau^k*, f) > R(\Delta^k, f) - \frac{1}{k}\right). \quad (D.40)
$$

Thus,

$$
J_*(f, I) \overset{(D.30)}{=} \lim_{k \to \infty} r(\Delta^k, f) \overset{(*)}{=} \lim_{k \to \infty} \rho(\Delta^k, \tau^k_*, f) = \alpha = \lim_{k \to \infty} \rho(\Delta^k, \tau^k*, f) \overset{(**)}{=} \lim_{k \to \infty} R(\Delta^k, f) \overset{(D.30)}{=} J^*(f, I),
$$

where, at $(*)$ and $(**)$, we used (D.37), (D.40), and the Sandwich theorem [Phi16a, Th. 7.16]. Since (D.41) establishes both $f \in \mathcal{R}(I)$ and $\alpha = \int_I f$, the proof is complete. □
Theorem D.11 (Riemann’s Integrability Criterion). Let \( I = [a, b] \subseteq \mathbb{R}^n \) be an interval, \( a, b \in \mathbb{R}^n, \ n \in \mathbb{N}, \ a < b \), and suppose \( f : I \rightarrow \mathbb{R} \) is bounded. Then \( f \) is Riemann integrable over \( I \) if, and only if, for each \( \epsilon > 0 \), there exists a partition \( \Delta \) of \( I \) such that

\[
R(\Delta, f) - r(\Delta, f) < \epsilon.
\]

(D.42)

Proof. Suppose, for each \( \epsilon > 0 \), there exists a partition \( \Delta \) of \( I \) such that (D.42) is satisfied. Then

\[
J^*(f, I) - J_*(f, I) \leq R(\Delta, f) - r(\Delta, f) < \epsilon,
\]

(D.43)
showing \( J^*(f, I) \leq J_*(f, I) \). As the opposite inequality always holds, we have \( J^*(f, I) = J_*(f, I) \), i.e. \( f \in \mathcal{R}(I) \) as claimed. Conversely, if \( f \in \mathcal{R}(I) \) and \( (\Delta^k)_{k \in \mathbb{N}} \) is a sequence of partitions of \( I \) with \( \lim_{k \to \infty} |\Delta^k| = 0 \), then (D.31a) implies that, for each \( \epsilon > 0 \), there is \( N \in \mathbb{N} \) such that \( R(\Delta^k, f) - r(\Delta^k, f) < \epsilon \) for each \( k > N \). ■

E Wallis Product

Proposition E.1. For each \( n \in \mathbb{N}_0 \), consider the integral

\[
I(n) := \int_0^\pi (\sin x)^n \, dx.
\]

(E.1)

(a) The integrals of (E.1) satisfy the following recursion:

\[
\forall n \geq 2, \quad I(n) = \frac{n-1}{n} I(n-2).
\]

(E.2)

(b) The integrals of (E.1) have the following product representations for each \( n \in \mathbb{N}_0 \):

\[
I(2n) = \int_0^\pi (\sin x)^{2n} \, dx = \pi \prod_{k=1}^n \frac{2k-1}{2k},
\]

(E.3a)

\[
I(2n+1) = \int_0^\pi (\sin x)^{2n+1} \, dx = 2 \prod_{k=1}^n \frac{2k}{2k+1}.
\]

(E.3b)

(c) One has the following representation of \( \pi \) as an infinite product, known as the Wallis product:

\[
\frac{\pi}{2} = \prod_{k=1}^\infty \left( \frac{2k}{2k-1} \cdot \frac{2k}{2k+1} \right) = \lim_{n \to \infty} \prod_{k=1}^n \left( \frac{2k}{2k-1} \cdot \frac{2k}{2k+1} \right) = \frac{2}{1} \cdot \frac{2}{3} \cdot \frac{4}{3} \cdot \frac{4}{5} \cdot \frac{6}{5} \cdot \frac{6}{7} \cdots.
\]

(E.4)
Proof. (a): We let \( n \geq 2 \) and integrate by parts to obtain

\[
I(n) = \int_0^\pi (\sin x)^n \, dx = \left[ -(\sin x)^{n-1} \cos x \right]_0^\pi + \int_0^\pi (n-1) \cos x (\sin x)^{n-2} \cos x \, dx
\]

\[
= 0 + (n-1) \int_0^\pi (1-(\sin x)^2) (\sin x)^{n-2} \, dx
\]

\[
= (n-1) I(n-2) - (n-1) I(n), \tag{E.5}
\]

implying \( n I(n) = (n-1) I(n-2) \) and (E.2).

(b): We use (a) to prove (E.3) via induction on \( n \in \mathbb{N}_0 \): For (E.3a), we have

\[
I(0) = \int_0^\pi \sin x \, dx = \left[ -\cos x \right]_0^\pi = -(-1) - (-1) = 2 = 2 \prod_{k=1}^n \frac{2k-1}{2k} \tag{E.6}
\]

and, for each \( n \in \mathbb{N} \),

\[
I(2n) = \frac{n-1}{n} I(2n-2) \tag{ind.hyp.} = \frac{2n-1}{2n} \pi \prod_{k=1}^{n-1} \frac{2k-1}{2k} = \pi \prod_{k=1}^n \frac{2k-1}{2k}, \tag{E.7}
\]

as needed. Similarly, for (E.3b), we have

\[
I(1) = \int_0^\pi \sin x \, dx = \left[ -\cos x \right]_0^\pi = -(-1) - (-1) = 2 = 2 \prod_{k=1}^n \frac{2k}{2k+1} \tag{E.8}
\]

and, for each \( n \in \mathbb{N} \),

\[
I(2n+1) = \frac{2n+1-1}{2n+1} I(2n-1) \tag{ind.hyp.} = \frac{2n}{2n+1} \prod_{k=1}^{n-1} \frac{2k}{2k+1} = \frac{2}{2n+1} \prod_{k=1}^n \frac{2k}{2k+1}, \tag{E.9}
\]

completing the induction and the proof of (b).

(c): For each \( x \in [0, 1] \), we have \( 0 \leq \sin x \leq 1 \), implying

\[
\forall \, x \in [0, 1] \quad \forall \, n \in \mathbb{N}_0 \quad 1 \geq (\sin x)^n \geq (\sin x)^{n+1} \geq 0 \tag{E.10}
\]

and

\[
\forall \, n \in \mathbb{N}_0 \quad I(n) \geq I(n+1) \geq 0. \tag{E.11}
\]

Thus,

\[
\forall \, n \in \mathbb{N} \quad 1 \leq \frac{I(2n)}{I(2n+1)} \leq \frac{I(2n-1)}{I(2n+1)} = \frac{2n+1}{2n}, \tag{E.12}
\]

implying

\[
1 = \lim_{n \to \infty} \frac{2n+1}{2n} = \lim_{n \to \infty} \frac{I(2n)}{I(2n+1)} \tag{E.3} \leq \frac{\pi}{2} \lim_{n \to \infty} \prod_{k=1}^n \left( \frac{2k-1}{2k} \cdot \frac{2k+1}{2k} \right), \tag{E.13}
\]

proving (c).
**F Improper Riemann Integral of the Sinc Function**

**Definition F.1.** The function

\[ f : \mathbb{R} \to \mathbb{R}, \quad f(x) := \begin{cases} \frac{\sin x}{x} & \text{for } x \neq 0, \\ 1 & \text{for } x = 0 \end{cases} \]  

is called the sinc function.

The sinc function is, clearly, continuous. We have seen in Ex. 2.58(a) that sinc is not Lebesgue integrable. The goal of this section is to show that sinc is still improperly Riemann integrable with

\[ \int_0^\infty \frac{\sin x}{x} \, dx = \frac{\pi}{2}. \]  

To this end, consider the function

\[ g : \mathbb{R}^+ \to \mathbb{R}, \quad g(t) := \int_0^\infty e^{-tx} \frac{\sin x}{x} \, dx. \]  

For each \( t \in \mathbb{R}^+ \), the integrand and its derivative with respect to \( t \) are Lebesgue integrable, since they are dominated by the Lebesgue integrable function

\[ x \mapsto \begin{cases} \frac{|\sin x|}{x} & \text{for } x < 1, \\ xe^{-tx} & \text{for } x \geq 1. \end{cases} \]  

For each \( \epsilon > 0 \) and each \( t \geq \epsilon \) the integrand in (F.3) is uniformly dominated by

\[ x \mapsto \begin{cases} \frac{|\sin x|}{x} & \text{for } x < 1, \\ xe^{-\epsilon x} & \text{for } x \geq 1. \end{cases} \]  

This shows that Cor. 2.24 applies on \([\epsilon, \infty]\), implying \( g \) to be differentiable (even on \( \mathbb{R}^+ \), since \( \epsilon > 0 \) was arbitrary),

\[ g' : \mathbb{R}^+ \to \mathbb{R}, \quad g'(t) := -\int_0^\infty x e^{-tx} \frac{\sin x}{x} \, dx = -\int_0^\infty e^{-tx} \sin x \, dx. \]  

To compute the integral in (F.6) explicitly, we claim that an antiderivative of \( h(x) := -e^{-tx} \sin x \) is given by

\[ H(x) := \frac{\cos x + t \sin x}{e^{tx}(1 + t^2)}. \]  

Indeed,

\[ H'(x) = \frac{e^{tx}(t \cos x - \sin x) - te^{tx}(\cos x + t \sin x)}{e^{2tx}(1 + t^2)} = \frac{-\sin x - t^2 \sin x}{e^{2tx}(1 + t^2)} = -e^{-tx} \sin x = h(x). \]
Thus, for each \( t \in \mathbb{R}^+ \),
\[
g'(t) = \left[ \frac{\cos x + t \sin x}{e^{tx}(1 + t^2)} \right]_{x=\infty}^{x=0} = 0 - \frac{1}{1 + t^2} = -\frac{1}{1 + t^2}. \tag{F.9}
\]
In consequence, both \( g \) and \((- \arctan)\) are antiderivatives of \( g' \), implying
\[
\exists C \in \mathbb{R} \quad \forall t \in \mathbb{R}^+ \quad g(t) = C - \arctan(t). \tag{F.10}
\]
Using dominated convergence, we obtain
\[
\lim_{t \to \infty} g(t) = \lim_{t \to \infty} \int_0^{\infty} e^{-tx} \frac{\sin x}{x} \, dx = 0 \tag{F.11}
\]
and
\[
C = \lim_{t \to \infty} (g(t) + \arctan(t)) = 0 + \frac{\pi}{2} = \frac{\pi}{2}. \tag{F.12}
\]
that means
\[
\forall t \in \mathbb{R}^+ \quad g(t) = \frac{\pi}{2} - \arctan(t). \tag{F.13}
\]
For each \( r > 0 \), we can apply dominated convergence once again to yield
\[
\int_0^r \frac{\sin x}{x} \, dx = \lim_{t \to 0} \int_0^r e^{-tx} \frac{\sin x}{x} \, dx. \tag{F.14}
\]
For each \( t \geq 0 \) and each \( k \in \mathbb{N} \), let
\[
a_{t,k} := \int_{(k-1)\pi}^{k\pi} e^{-tx} \frac{\sin x}{x} \, dx. \tag{F.15}
\]
According to the Leibniz criterion [Phi16a, Th. 7.85],
\[
\forall t > 0 \quad g(t) = \sum_{k=1}^{\infty} a_{t,k}, \quad \forall j \in \mathbb{N} \quad \left| \sum_{k=j+1}^{\infty} a_{t,k} \right| < |a_{0,j+1}|. \tag{F.16}
\]
Thus, given \( \epsilon > 0 \), we can choose \( R > 0 \), such that
\[
\forall r \geq R \quad \forall t > 0 \quad \left| g(t) - \int_0^r e^{-tx} \frac{\sin x}{x} \, dx \right| = \left| \int_r^{\infty} e^{-tx} \frac{\sin x}{x} \, dx \right| < \frac{\epsilon}{4}. \tag{F.17}
\]
and
\[
\forall r \geq R \quad \left| \int_0^{\infty} \frac{\sin x}{x} \, dx - \int_0^r \frac{\sin x}{x} \, dx \right| = \left| \int_r^{\infty} \frac{\sin x}{x} \, dx \right| < \frac{\epsilon}{4}. \tag{F.18}
\]
Now let \( r \geq R \). Using (F.13) and (F.14), we choose \( \delta > 0 \) such that
\[
\forall 0 < t < \delta \quad \frac{\pi}{2} > \int_0^\infty e^{-tx} \sin x \, dx = g(t) > \frac{\pi}{2} - \frac{\epsilon}{4}. \tag{F.19}
\]
and
\[ \left| \int_0^r \sin \frac{x}{x} \, dx - \int_0^r e^{-tx} \sin \frac{x}{x} \, dx \right| < \frac{\epsilon}{4}. \]  
(F.20)

Altogether, we then obtain, for \( 0 < t < \delta \),
\[ \left| \pi - \int_0^\infty \sin \frac{x}{x} \, dx \right| \leq \left| \pi - g(t) \right| + \left| g(t) - \int_0^r e^{-tx} \sin \frac{x}{x} \, dx \right| \]
\[ + \left| \int_0^r e^{-tx} \sin \frac{x}{x} \, dx - \int_0^r \frac{\sin x}{x} \, dx \right| \]
\[ + \left| \int_0^r \frac{\sin x}{x} \, dx - \int_0^\infty \frac{\sin x}{x} \, dx \right| < 4 \cdot \frac{\epsilon}{4} = \epsilon. \]  
(F.21)

As \( \epsilon > 0 \) was arbitrary, this proves (F.2).

G Topological and Measure-Theoretic Supplements

G.1 Transitivity of Dense Subsets

**Proposition G.1.** Let \((X, T)\) be a topological space, \(B \subseteq A \subseteq X\). If \(B\) is dense in \(A\) and \(A\) is dense in \(X\), then \(B\) is dense in \(X\).

**Proof.** Let \(\emptyset \neq O \in T\). Since \(A\) is dense in \(X\), \(O \cap A \neq \emptyset\). Since \(O \cap A \in T_A\) and \(B\) is dense in \(A\), \(O \cap A \cap B \neq \emptyset\). Thus, \(O \cap B \neq \emptyset\), showing \(B\) to be dense in \(X\). \( \blacksquare \)

G.2 Inverse Image and Trace for Semirings

**Proposition G.2.** Consider sets \(X, Y\) and a map \(f : X \to Y\).

(a) If \(S\) is a semiring on \(Y\), then \(f^{-1}(S)\) is a semiring on \(X\).

(b) Consider \(B \subseteq X\). If \(S\) is a semiring on \(X\), then the trace \(S|B\) is a semiring on \(B\).

**Proof.** (a): As \(\emptyset = f^{-1}(\emptyset)\), \(\emptyset \in f^{-1}(S)\). If \(R, S \in S\), then
\[ f^{-1}(R) \cap f^{-1}(S) = f^{-1}(R \cap S) \in f^{-1}(S), \]
since \(R \cap S \in S\) due to \(S\) being a semiring. Moreover, as \(S\) is a semiring, for \(R, S \in S\), there are finitely many disjoint sets \(T_1, \ldots, T_n\), \(n \in \mathbb{N}\), such that \(T_i \in S\) for each \(i \in \{1, \ldots, n\}\), and \(R \backslash S = \bigcup_{i=1}^n T_i\). Then
\[ f^{-1}(R) \backslash f^{-1}(S) = f^{-1}(R \backslash S) = f^{-1} \left( \bigcup_{i=1}^n T_i \right) = \bigcup_{i=1}^n f^{-1}(T_i), \]
completing the proof that \(f^{-1}(S)\) is a semiring.

(b): Due to (1.39), (b) follows immediately from (a). \( \blacksquare \)
G.3 Measurability with Respect to Completions

Proposition G.3. Let \((X, \mathcal{A}, \mu), (X, \mathcal{B}, \nu)\) be measure spaces such that \(\mathcal{A} \subseteq \mathcal{B} \subseteq \hat{\mathcal{A}}, \mu = \nu|_\mathcal{A}, \nu = \hat{\mu}|_\mathcal{B}\), where \((X, \hat{\mathcal{A}}, \hat{\mu})\) is the completion of \((X, \mathcal{A}, \mu)\). If \(f : X \to K\) is \(\mathcal{B}\)-measurable, then there exists an \(\mathcal{A}\)-measurable function \(g : X \to K\) such that \(f = g\) \(\mu\)-almost everywhere.

**Proof.** First, assume \(f \geq 0\) (i.e. \(f \in M^+(\mathcal{B})\)) and choose a sequence \((\phi_k)_{k \in \mathbb{N}}\) in \(S^+(\mathcal{B})\) such that \(\phi_k \uparrow f\). Then

\[
f = \lim_{N \to \infty} \phi_N = \lim_{N \to \infty} \left( \phi_1 + \sum_{k=2}^{N} (\phi_k - \phi_{k-1}) \right).
\]

Since \(\phi_1 \in S^+(\mathcal{B})\) and each \((\phi_k - \phi_{k-1}) \in S^+(\mathcal{B})\), there exist sequences \((\alpha_k)_{k \in \mathbb{N}}\) in \(\mathbb{R}^+\) and \((B_k)_{k \in \mathbb{N}}\) in \(\mathcal{B}\) such that \(f = \sum_{k=1}^{\infty} \alpha_k \chi_{B_k}\). Since \(B_k \in \mathcal{B} \subseteq \hat{\mathcal{A}}\), there exist \(A_k, N_k \in \mathcal{A}\), \(M_k \subseteq N_k\), such that \(B_k = A_k \cup M_k\), \(\mu(N_k) = 0\). Define

\[
g := \sum_{k=1}^{\infty} \alpha_k \chi_{A_k}.
\]

Then \(g \in M^+(\mathcal{A})\). We claim \(f = g\) \(\mu\)-a.e.: Indeed, if \(x \notin \bigcup_{k=1}^{\infty} (B_k \setminus A_k)\), then \(f(x) = g(x)\) (if \(x \notin \bigcap_{k=1}^{\infty} B_k\), then \(f(x) = g(x) = 0\); if \(x \in B_k\) implies \(x \in A_k\) for each \(k \in \mathbb{N}\), then \(f(x) = g(x) = \sum_{k=1}^{\infty} \alpha_k\)). Since, for each \(k \in \mathbb{N}\), \(B_k \setminus A_k \subseteq M_k \subseteq N_k\),

\[
\mu \left( \bigcup_{k=1}^{\infty} (B_k \setminus A_k) \right) \leq \sum_{k=1}^{\infty} \mu(N_k) = 0,
\]

we have shown \(f = g\) \(\mu\)-a.e. For a general \(\mathcal{B}\)-measurable \(f : X \to K\), we decompose \(f\) into \((\text{Re } f)^\pm\) and \((\text{Im } f)^\pm\) in the usual way, and apply the case \(f \geq 0\) to the functions \((\text{Re } f)^\pm\), \((\text{Im } f)^\pm\).

\[\blacksquare\]

G.4 Interchanging Integrals with Uniform Limits

**Proposition G.4.** Let \((X, \mathcal{A}, \mu)\) be a measure space with \(\mu(X) < \infty\). Let \(f, f_n : X \to K\) be measurable functions, \(n \in \mathbb{N}\), such that each \(f_n\) is \(\mu\)-integrable and \(f = \lim_{n \to \infty} f_n\) uniformly \(\mu\)-a.e. (i.e. there exists a \(\mu\)-null set \(N \in \mathcal{A}\) such that \(f = \lim_{n \to \infty} f_n\) uniformly on \(N^c\)). Then \(f\) is \(\mu\)-integrable with

\[
\lim_{n \to \infty} \int_X f_n \, d\mu = \int_X f \, d\mu \quad (G.4a)
\]

and

\[
\lim_{n \to \infty} \int_X |f_n - f| \, d\mu = 0. \quad (G.4b)
\]
Proof. If \( \mu(X) = 0 \), there is nothing to prove. Otherwise, given \( \epsilon > 0 \), there exists \( N \in \mathbb{N} \) such that, for each \( n > N \), \( \|f_n - f\|_\infty < \epsilon/\mu(X) \). Thus, for each \( n > N \),

\[
\int_X |f_n - f| \, d\mu \leq \frac{\epsilon}{\mu(X)} \cdot \mu(X) = \epsilon,
\]

proving (G.4b). As

\[
\forall \ n \in \mathbb{N} \quad \left| \int_X f_n \, d\mu - \int_X f \, d\mu \right| \leq \int_X |f_n - f| \, d\mu,
\]

(G.4a) is also proved. Since \( \|f\|_1 \leq \|f - f_n\|_1 + \|f_n\|_1 < \infty \), \( f \) is \( \mu \)-integrable. ■

\section*{H Polar Coordinates}

Let

\[ O := \mathbb{R}^+ \times ]0, 2\pi[, \quad U := \mathbb{R}^2 \setminus \{(x, 0) : x \geq 0\}. \]

(H.1)

We verify that the function defining polar coordinates, i.e.

\[ \phi : O \to U, \quad \phi(r, \varphi) := (r \cos \varphi, r \sin \varphi), \]

(H.2)

constitutes a \( C^1 \)-diffeomorphism:

We first show \( \phi \) to be bijective with inverse map

\[ \phi^{-1} : U \to O, \quad \phi^{-1}(x, y) = (\phi_1^{-1}(x, y), \phi_2^{-1}(x, y)), \]

(H.3a)

where, recalling the definition of \( \arccot \) from [Phi16a, Def. and Rem. 8.27],

\[
\phi_1^{-1}(x, y) := \sqrt{x^2 + y^2}, \quad \phi_2^{-1}(x, y) := \begin{cases} \arccot(x/y) & \text{for } y > 0, \\ \pi & \text{for } y = 0, \\ \pi + \arccot(x/y) & \text{for } y < 0. \end{cases}
\]

(H.3b)

(H.3c)

One verifies \( \phi^{-1} \circ \phi = \text{Id}_O \):

\[
\forall \ (r, \varphi) \in O \quad (\phi^{-1} \circ \phi)(r, \varphi) = \phi^{-1}(r \cos \varphi, r \sin \varphi) = (r, \varphi),
\]

since

\[
\phi_1^{-1}(r \cos \varphi, r \sin \varphi) = \sqrt{r^2 \cos^2 \varphi + r^2 \sin^2 \varphi} = r,
\]

\[
\phi_2^{-1}(r \cos \varphi, r \sin \varphi) = \begin{cases} \arccot(\cot \varphi) = \varphi & \text{for } 0 < \varphi < \pi, \\ \pi = \varphi & \text{for } \varphi = \pi, \\ \pi + \arccot(\cot \varphi) = \pi + \varphi - \pi = \varphi & \text{for } \pi < \varphi < 2\pi; \end{cases}
\]
and $\phi \circ \phi^{-1} = \text{Id}_U$:

$$
\forall \ (x,y) \in U \quad (\phi \circ \phi^{-1})(x,y) = \left(\sqrt{x^2 + y^2 \cos \phi^{-1}(x,y)}, \sqrt{x^2 + y^2 \sin \phi^{-1}(x,y)}\right) = (x,y),
$$

since

$$
\cos \phi^{-1}(x,y) = \begin{cases}
\frac{1}{\sqrt{1 + \cot \arccot(x/y)}} = \frac{1}{\sqrt{1 + \cot \arccot(x/y)}} = \frac{x}{\sqrt{x^2 + y^2}} & \text{for } x > 0 \ (\Rightarrow y \neq 0), \\
\cos(\pi/2) = 0 = \frac{x}{\sqrt{x^2 + y^2}} & \text{for } x = 0, y > 0, \\
\cos(3\pi/2) = 0 = \frac{y}{\sqrt{x^2 + y^2}} & \text{for } x = 0, y < 0, \\
\cos \pi = -1 = \frac{x}{\sqrt{x^2 + y^2}} & \text{for } x < 0, y \neq 0,
\end{cases}
$$

$$
\sin \phi^{-1}(x,y) = \begin{cases}
\frac{1}{\sqrt{1 + \cot^2 \arccot(x/y)}} = \frac{1}{\sqrt{1 + \cot^2 \arccot(x/y)}} = \frac{y}{\sqrt{x^2 + y^2}} & \text{for } y > 0, \\
\sin \pi = 0 = \frac{y}{\sqrt{x^2 + y^2}} & \text{for } y = 0, \\
\sqrt{1 + \cot^2 \arccot(x/y)} = \frac{y}{\sqrt{x^2 + y^2}} & \text{for } y < 0.
\end{cases}
$$

Next, we note that $\phi$ has continuous first partials, where

$$
\forall \ (r,\varphi) \in O \quad D\phi(r,\varphi) = \begin{pmatrix}
\cos \varphi & -r \sin \varphi \\
\sin \varphi & r \cos \varphi
\end{pmatrix}, \quad \det D\phi(r,\varphi) = r > 0. \quad (\text{H.4})
$$

Thus, $\phi$ is a $C^1$-diffeomorphism by Def. and Rem. 2.64.

References


REFERENCES


