Chapter 10

Measure Theory and Integration

10.1 Measure Theory

Citation from Rudins book, [11, Chapter 1]: Towards the end of the 19th century it became clear to many mathematicians that the Riemann integral should be replaced by some other type of integral, more general and more flexible, better suited for dealing with limit processes. Among the attempts made in this direction, the most notable ones were due to Jordan, Borel, W.H. Young, and Lebesgue. It was Lebesgue's construction which turned out to be the most successful.

In a brief outline, here is the main idea: The Riemann integral of a function f over an interval $[a, b]$ can be approximated by sums of the form

$$\sum_{i=1}^{n} f(t_i)\, m(E_i),$$

where $E_1, \dots, E_n$ are disjoint intervals whose union is $[a, b]$, $m(E_i)$ denotes the length of E_i and $t_i \in E_i$ for $i = 1, \dots, n$. Lebesgue discovered that a completely satisfactory theory of integration results if the sets E_i in the above sum are allowed to belong to a larger class of subsets of the line, the so-called "measurable sets," and if the class of functions under consideration is enlarged to what we call "measurable functions." The crucial set-theoretic properties involved are the following: The union and the intersection of any countable family of measurable sets are measurable; ... the notion of "length" (now called "measure") can be extended to them in such a way that

$$m(E_1 \cup E_2 \cup \cdots) = m(E_1) + m(E_2) + \cdots$$

for any countable collection $\{E_i\}$ of pairwise disjoint measurable sets. This property of m is called *countable additivity*.

The passage from Riemann's theory of integration to that of Lebesgue is a process of completion. It is of the same fundamental importance in analysis as the construction of the real number system from rationals.

The Measure Problem

Lebesgue (1904) states the following problem: We want to associate to each bounded subset E of the real line a positive real number $m(E)$, called measure of E, such that the following properties are satisfied:

(1) Any two congruent sets have the same measure.
(2) The measure is counntable additive.
(3) The measure of the unit interval $[0,1]$ is 1.

He emphasized that he was not able to solve this problem in full detail, but for a certain class of sets which he called measurable. We will see that this restriction to a large family of bounded sets is unavoidable—the measure problem has no solution.

Definition 10.1 Let X be a set. A family (non-empty) family $\mathcal{A}$ of subsets of X is called an *algebra* if

1. $A \in \mathcal{A}$ implies $A^{\mathsf{c}} \in \mathcal{A}$,
2. $A, B \in \mathcal{A}$ implies $A \cup B \in \mathcal{A}$.

An algebra $\mathcal{A}$ is called a σ*-algebra* if for all countable families $\{A_n \mid n \in \mathbb{N}\}$ with $A_n \in \mathcal{A}$ we have $\bigcup\limits_{n\in\mathbb{N}} A_n \in \mathcal{A}$.

Remarks 10.1 (a) Since $A \in \mathcal{A}$ implies $A \cup A^{\mathsf{c}} \in \mathcal{A}$; $X \in \mathcal{A}$ and $\varnothing = X^{\mathsf{c}} \in \mathcal{A}$.
(b) If $\mathcal{A}$ is an algebra, then $A \cap B \in \mathcal{A}$ for all $A, B \in \mathcal{A}$. Indeed, by de Morgan's rule (Lemma 6.9) we have $A \cap B = (A^{\mathsf{c}} \cup B^{\mathsf{c}})^{\mathsf{c}}$, and all the members on the right are in $\mathcal{A}$ by the definition of an algebra.
(c) Let $\mathcal{A}$ be a σ-algebra. Then $\bigcap\limits_{n\in\mathbb{N}} A_n \in \mathcal{A}$ if $A_n \in \mathcal{A}$ for all $n \in \mathbb{N}$. Again by de Morgan's rule

$$\bigcap_{n\in\mathbb{N}} A_n = \left(\bigcup_{n\in\mathbb{N}} A_n^{\mathsf{c}}\right)^{\mathsf{c}}.$$

(d) The family $\mathcal{P}(X)$ of all subsets of X is both an algebra as well as a σ-algebra.
(e) Any σ-algebra is an algebra but there are algebras not being σ-algebras.
(f) The family of finite and cofinite subsets (these are complements of finite sets) of an infinite set form an algebra. Do they form a σ-algebra?

Elementary Sets in $\mathbb{R}^n$

Let $\overline{\mathbb{R}}$ be the real axis together with $\pm\infty$, $\overline{\mathbb{R}} = \mathbb{R} \cup \{+\infty\} \cup \{-\infty\}$. We use the following rules.

1. $-\infty < +\infty$ and $-\infty < x < +\infty$ for all $x \in \mathbb{R}$.
2. $\pm\infty + x = x \pm \infty = \pm\infty$ for all $x \in \mathbb{R}$.

3. If $x > 0$ we have

$$x \cdot (+\infty) = (+\infty) \cdot x = (-\infty) \cdot (-\infty) = +\infty$$
$$x \cdot (-\infty) = (-\infty) \cdot x = (+\infty) \cdot (-\infty) = (-\infty) \cdot (+\infty) = -\infty.$$

4. If $x < 0$ we have

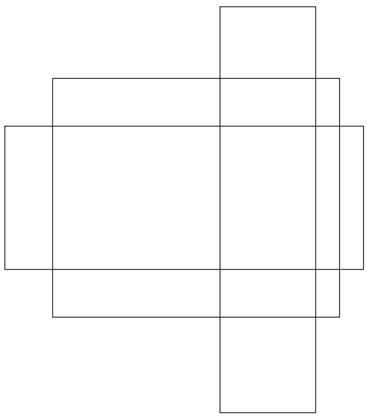
$$x \cdot (+\infty) = (+\infty) \cdot x = -\infty$$
$$x \cdot (-\infty) = (-\infty) \cdot x = +\infty.$$

5. $0 \cdot \pm\infty = \pm\infty \cdot 0 = 0.$

The set

$$I = \{(x_1, \dots, x_n) \in \mathbb{R}^n \mid a_i \preceq x_i \preceq b_i, \quad i = 1, \dots, n\}$$

is called a *rectangle* or a *box* in $\mathbb{R}^n$, where $\preceq$ either stands for $<$ or for $\leq$ where $a_i, b_i \in \overline{\mathbb{R}}$. For example $a_i = -\infty$ and $b_i = +\infty$ yields $I = \mathbb{R}^n$, whereas $a_1 = 2$, $b_1 = 1$ yields $I = \varnothing$. A subset of $\mathbb{R}^n$ is called an *elementary set* if it is the union of a finite number of rectangles in $\mathbb{R}^n$. Let $\mathcal{E}_n$ denote the set of elementary subsets of $\mathbb{R}^n$.



$\mathcal{E}_n$ is an algebra but not a σ-algebra.
Proof. The complement of a finite interval is the union of two intervals, the complement of an infinite interval is an infinite interval. Hence, the complement of a rectangle in $\mathbb{R}^n$ is the finite union of rectangles. The countable (disjoint) union $M = \bigcup_{n\in\mathbb{N}}[n, n + \frac{1}{2}]$ is not an elementary set. ∎

Note that any elementary set is the *disjoint* union of a finite number of rectangles.

Definition 10.2 Let $\mathcal{B}$ be any (nonempty) family of subsets of X. Let $\sigma(\mathcal{B})$ denote the intersection of all σ-algebras containing $\mathcal{B}$, i. e.

$$\sigma(\mathcal{B}) = \bigcap_{i\in I} \mathcal{A}_i,$$

where $\{\mathcal{A}_i \mid i \in I\}$ is the family of all σ-algebras $\mathcal{A}_i$ which contain $\mathcal{B}$, $\mathcal{B} \subseteq \mathcal{A}_i$ for all $i \in I$.

Note that the σ-algebra $\mathcal{P}(X)$ of all subsets is always a member of that family $\{\mathcal{A}_i\}$ such that $\sigma(\mathcal{B})$ exists. We call $\sigma(\mathcal{B})$ the σ-algebra *generated by* $\mathcal{B}$. $\sigma(\mathcal{B})$ is the smallest σ-algebra which contains the sets of $\mathcal{B}$.

Lemma 10.1 *For any non-empty family $\mathcal{B}$ of subsets of X, $\sigma(\mathcal{B})$ is a σ-algebra over X. We have $\sigma(\sigma(\mathcal{B})) = \sigma(\mathcal{B})$. If $\mathcal{B}_1 \subset \mathcal{B}_2$ then $\sigma(\mathcal{B}_1) \subset \sigma(\mathcal{B}_2)$.*

Proof. (a) Let $A_n, A \in \sigma(\mathcal{B})$ for all $n \in \mathbb{N}$, that is $A_n, A \in \mathcal{A}_i$ for all $i \in I$. Since $\mathcal{A}_i$ is a σ-algebra, $A^c \in \mathcal{A}_i$ and $\bigcup_{n\in\mathbb{N}} A_n \in \mathcal{A}_i$ for all i. Thus, $A^c \in \sigma(\mathcal{B})$ and $\bigcup_{n\in\mathbb{N}} A_n \in \sigma(\mathcal{B})$. This completes the proof of the first part.
(b) This property is obvious since $\sigma(\mathcal{A}) = \mathcal{A}$ for any σ-algebra $\mathcal{A}$: the smallest σ-algebra, that contains the given σ-algebra $\mathcal{A}$ is of course $\mathcal{A}$ itself. The property $\sigma^2 = \sigma$ is called the *idempotence* of σ. The same facts are true for taking the closure.
(c)

$$\mathcal{B}_1 \subset \mathcal{B}_2 \quad \text{implies} \quad \sigma(\mathcal{B}_1) \subset \sigma(\mathcal{B}_2)$$

In fact, any σ-algebra $\mathcal{A}_i^{(2)}$ that contains $\mathcal{B}_2$ contains $\mathcal{B}_1$ as well. However, among the family of σ-algebras $\{\mathcal{A}_j^{(1)}\}$ containing $\mathcal{B}_1$ there may be more. Hence, the intersection over larger family $\{\mathcal{A}_j^{(1)}\}$ becomes smaller. ∎

Recall that $\mathcal{E}_n$ is the algebra of elementary sets in $\mathbb{R}^n$.

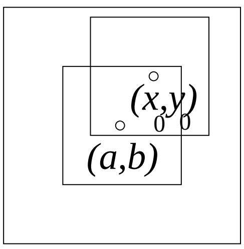
Definition 10.3 The *Borel algebra* in $\mathbb{R}^n$ is the σ-algebra $\sigma(\mathcal{E}_n)$ generated by the elementary sets $\mathcal{E}_n$. Its elements are called *Borel sets.*

The Borel algebra is the smallest σ-algebra which contains all boxes in $\mathbb{R}^n$ We will see that the Borel algebra is a huge family of subsets of $\mathbb{R}^n$ which contains "all sets we are ever interested in." Later, we will construct a non-Borel set.

Proposition 10.2 *Open and closed subsets of* $\mathbb{R}^n$ *are Borel sets.*

Proof. We give the proof in case of $\mathbb{R}^2$. Let $I_\varepsilon(x_0, y_0) = (x_0 - \varepsilon, x_0 + \varepsilon) \times (y_0 - \varepsilon, y_0 + \varepsilon)$ denote the open square of size 2ε by 2ε with midpoint (x_0, y_0). Then $I_{\frac{1}{n+1}} \subseteq I_{\frac{1}{n}}$ for $n \in \mathbb{N}$. Let $M \subset \mathbb{R}^2$ be open. To every point (x_0, y_0) with rational coordinates x_0, y_0 we choose the largest square $I_{1/n}(x_0, y_0) \subseteq M$ in M and denote it by $J(x_0, y_0)$.
We show that

$$M = \bigcup_{(x_0,y_0)\in M,\ \text{rational}} J(x_0, y_0).$$



Since the number of rational points in M is at least countable, the right side is in $\sigma(\mathcal{E}_2)$. Now let $(a, b) \in M$ arbitrary. Since M is open, there exists $n \in \mathbb{N}$ such that $I_{2/n}(a, b) \subseteq M$. Since the rational points are dense in $\mathbb{R}^2$, there is rational point (x_0, y_0) which is contained in $I_{1/n}(a, b)$. Then we have

$$I_{\frac{1}{n}}(x_0, y_0) \subseteq I_{\frac{2}{n}}(a, b) \subseteq M.$$

Since $(a, b) \in I_{\frac{1}{n}}(x_0, y_0) \subseteq J(x_0, y_0)$, we have shown that M is the union of the countable family of sets J. Since closed sets are the complements of open sets and complements are again in the σ-algebra, the assertion follows for closed sets. ∎

Remarks 10.2 (a) We have proved that any open subset M of $\mathbb{R}^n$ is the countable union of rectangles $I \subseteq M$.
(b) The Borel sets in $\mathbb{R}^n$ are exactly the sets which can be built by countable unions or intersections of open or closed sets.
(c) The Borel algebra $\sigma(\mathcal{E}_n)$ is equal to the σ-algebra generated by the family $\mathcal{G}$ of open sets in $\mathbb{R}^n$, cf. Homework 34.2.

Let us look in more detail at some of the sets in $\sigma(\mathcal{E}_n)$. Let $\mathcal{G}$ and $\mathcal{F}$ be the families of all open and closed subsets of $\mathbb{R}^n$, respectively. Let $\mathcal{G}_\delta$ be the collection of all intersection of sequences of open sets (from $\mathcal{G}$), and let $\mathcal{F}_\sigma$ be the collection of all unions of sequences of sets of $\mathcal{F}$. One can prove that $\mathcal{F} \subset \mathcal{G}_\delta$ and $\mathcal{G} \subset \mathcal{F}_\sigma$. These inclusions are strict. Since countable intersection and unions of countable intersections and union are still countable operations, $\mathcal{G}_\delta, \mathcal{F}_\sigma \subset \sigma(\mathcal{E}_n)$
For an arbitrary family $\mathcal{S}$ of sets let $\mathcal{S}_\sigma$ be the collection of all unions of sequences of sets in $\mathcal{S}$, and let $\mathcal{S}_\delta$ be the collection of all unions of sequences of sets in $\mathcal{S}$. We can iterate the operations represented by σ and δ, obtaining from the class $\mathcal{G}$ the classes $\mathcal{G}_\delta$, $\mathcal{G}_{\delta\sigma}$, $\mathcal{G}_{\delta\sigma\delta}$, ... and from $\mathcal{F}$ the classes $\mathcal{F}_\sigma$, $\mathcal{F}_{\sigma\delta}$, ... It turns out that we have inclusions

$$\mathcal{G} \subset \mathcal{G}_\delta \subset \mathcal{G}_{\delta\sigma} \subset \cdots \subset \sigma(\mathcal{E}_n)$$
$$\mathcal{F} \subset \mathcal{F}_\sigma \subset \mathcal{F}_{\sigma\delta} \subset \cdots \subset \sigma(\mathcal{E}_n).$$

No two of these classes are equal. There are Borel sets that belong to none of them.

10.1.1 Additive Functions and Measures

Definition 10.4 (a) Let $\mathcal{A}$ be an algebra over X. An *additive function* or *content* μ on $\mathcal{A}$ is a function $\mu\colon \mathcal{A} \to [0, +\infty]$ such that

(i) $\mu(\varnothing) = 0$,
(ii) $\mu(A \cup B) = \mu(A) + \mu(B)$ for all $A, B \in \mathcal{A}$ with $A \cap B = \varnothing$.

(b) An additive function μ is called *countably additive* (or *σ-additive* in the German literature) on $\mathcal{A}$ if for any disjoint family $\{A_n \mid A_n \in \mathcal{A},\, n \in \mathbb{N}\}$, that is $A_i \cap A_j = \varnothing$ for all $i \neq j$, with $\bigcup_{n\in\mathbb{N}} A_n \in \mathcal{A}$ we have

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

(c) A *measure* is a countably additive function on a σ-algebra $\mathcal{A}$.

If X is a set, $\mathcal{A}$ a σ-algebra on X and μ a measure on $\mathcal{A}$, then the tripel $(X, \mathcal{A}, \mu)$ is called a *measure space*. Likewise, if X is a set and $\mathcal{A}$ a σ-algebra on X, the pair $(X, \mathcal{A})$ is called a *measurable space*.

Notation. We write $\sum\limits_{n\in\mathbb{N}} A_n$ in place of $\bigcup\limits_{n\in\mathbb{N}} A_n$ if $\{A_n\}$ is a disjoint family of subsets. The countable additivity reads as follows

$$\mu(A_1 \cup A_2 \cup \cdots) = \mu(A_1) + \mu(A_2) + \cdots ,$$

$$\mu\left(\{x \in X \mid \ \exists\, n \in \mathbb{N} : x \in A_n\}\right) = \lim_{n\to\infty} \sum_{k=1}^{n} \mu(A_k).$$

We call the additive function μ *finite* if $\mu(X) < \infty$.

Example 10.1 (a) Let X be a set, $x_0 \in X$ and $\mathcal{A} = \mathcal{P}(X)$. Then

$$\mu(A) = \begin{cases} 1, & x_0 \in A, \\ 0, & x_0 \not\in A \end{cases}$$

defines a finite measure on $\mathcal{A}$. μ is called the *point mass concentrated at* x_0.
(b) Let X be a set and $\mathcal{A} = \mathcal{P}(X)$. Put

$$\mu(A) = \begin{cases} +\infty, & \text{if } A \text{ has infinitely many elements} \\ n, & \text{if } A \text{ has } n \text{ elements.} \end{cases}$$

μ is a measure on $\mathcal{A}$, the so called *counting measure.*
(c) $X = \mathbb{R}^n$, $\mathcal{A} = \mathcal{E}_n$ is the algebra of elementary sets of $\mathbb{R}^n$. Every $A \in \mathcal{E}_n$ is the finite disjoint union of rectangles $A = \sum_{k=1}^m I_k$. We set $\mu(A) = \sum_{k=1}^m \mu(I_k)$ where

$$\mu(I) = (b_1 - a_1)\cdots(b_n - a_n),$$

if $I = \{(x_1,\dots,x_n) \in \mathbb{R}^n \mid a_i \preceq x_i \preceq b_i,\ i = 1,\dots,n\}$ and $a_i \preceq b_i$; $\mu(\varnothing) = 0$. Then μ is an additive function on $\mathcal{E}_n$. It is called the *Lebesgue content* on $\mathbb{R}^n$. Note that μ is not a measure (since $\mathcal{A}$ is not a σ-algebra and μ is not yet shown to be countably additive). However, we will see in Proposition 10.5 below that μ is even *countably additive.* By definition, $\mu(\text{line in } \mathbb{R}^2) = 0$ and $\mu(\text{plane in } \mathbb{R}^3) = 0$.
(d) Let $X = \mathbb{R}$, $\mathcal{A} = \mathcal{E}_1$, and α an increasing function on $\mathbb{R}$. For a, b in $\overline{\mathbb{R}}$ with $a < b$ set

$$\begin{aligned} \mu_\alpha([a,b]) &= \alpha(b+0) - \alpha(a-0), \\ \mu_\alpha([a,b)) &= \alpha(b-0) - \alpha(a-0), \\ \mu_\alpha((a,b]) &= \alpha(b+0) - \alpha(a+0), \\ \mu_\alpha((a,b)) &= \alpha(b-0) - \alpha(a+0). \end{aligned}$$

Then μ_α is an additive function on $\mathcal{E}_1$ if we set

$$\mu_\alpha(A) = \sum_{i=1}^{n} \mu\alpha(I_i), \quad \text{if} \quad A = \sum_{i=1}^{n} I_i.$$

We call μ_α the *Lebesgue–Stieltjes content.*
On the other hand, if $\mu\colon \mathcal{E}_1 \to \overline{\mathbb{R}}$ is an additive function, then $\alpha\colon \mathbb{R} \to \mathbb{R}$ defined by

$$\alpha(x) = \begin{cases} \mu((0,x]), & x \ge 0, \\ -\mu((x,0]), & x < 0, \end{cases}$$

defines an increasing function α on $\mathbb{R}$ such that $\mu = \mu_\alpha$ if α is continuous from the right.

Properties of Additive Functions

Proposition 10.3 *Let $\mathcal{A}$ be an algebra over X and μ an additive function on $\mathcal{A}$. Then*

(a) $\mu\left(\sum_{k=1}^{n} A_k\right) = \sum_{k=1}^{n} \mu(A_k)$ *if $A_i \in \mathcal{A}$ form a disjoint family of subsets.*

(b) $\mu(A \cup B) + \mu(A \cap B) = \mu(A) + \mu(B)$, $A, B \in \mathcal{A}$.

(c) $A \subseteq B$ *implies* $\mu(A) \subseteq \mu(B)$ (μ is monotone).

(d) *If $A \subseteq B$, $A, B \in \mathcal{A}$, and $\mu(A) < +\infty$, then $\mu(B \setminus A) = \mu(B) - \mu(A)$,* ($\mu$ is subtractive).

(e)

$$\mu\left(\bigcup_{k=1}^{n} A_k\right) \le \sum_{k=1}^{n} \mu(A_k),$$

if $A_k \in \mathcal{A}$, $k = 1, \dots, n$, (μ is finitely subadditive).

(f) *If $\{A_k \mid k \in \mathbb{N}\}$ is a disjoint family in $\mathcal{A}$ and $\sum_{k=1}^{\infty} A_k \in \mathcal{A}$. Then*

$$\sum_{k=1}^{\infty} \mu(A_k) \le \mu\left(\sum_{k=1}^{\infty} A_k\right).$$

Proof. (a) is by induction. (d), (c), and (b) are easy (cf. Homework 34.4).
(e) We can write $\bigcup_{k=1}^{n} A_k$ as the finite *disjoint* union of n sets of $\mathcal{A}$:

$$\bigcup_{k=1}^{n} A_k = A_1 \cup \big(A_2 \setminus A_1\big) \cup \big(A_3 \setminus (A_1 \cup A_2)\big) \cup \dots \cup \big(A_n \setminus (A_1 \cup \dots \cup A_{n-1})\big).$$

Since μ is additive,

$$\mu\left(\bigcup_{k=1}^{n} A_k\right) = \sum_{k=1}^{n} \mu\left(A_k \setminus (A_1 \cup \cdots A_{k-1})\right) \le \sum_{k=1}^{n} \mu(A_k),$$

where we used $\mu(B \setminus A) \le \mu(B)$ (from (d)).
(f) Since μ is additive, and monotone

$$\sum_{k=1}^{n} \mu(A_k) = \mu\left(\sum_{k=1}^{n} A_k\right) \le \mu\left(\sum_{k=1}^{\infty} A_k\right).$$

Taking the supremum on the left gives the assertion. ∎

Proposition 10.4 *Let μ be an additive function on the algebra $\mathcal{A}$. Consider the following statements*

(a) *μ is countably additive.*

(b) *For any increasing sequence $A_n \subseteq A_{n+1}$, $A_n \in \mathcal{A}$, with $\bigcup_{n=1}^{\infty} A_n = A \in \mathcal{A}$ we have $\lim_{n\to\infty} \mu(A_n) = \mu(A)$.*

(c) *For any decreasing sequence $A_n \supseteq A_{n+1}$, $A_n \in \mathcal{A}$, with $\bigcap_{n=1}^{\infty} A_n \in \mathcal{A}$ and $\mu(A_n) < \infty$ we have $\lim_{n\to\infty} \mu(A_n) = \mu(A)$.*

(d) *Statement* (c) *with $A = \varnothing$ only.*

We have (a) $\leftrightarrow$ (b) $\to$ (c) $\to$ (d). *In case $\mu(X) < \infty$ (μ is finite), all statements are equivalent.*

Proof. (a) $\to$ (b). Without loss of generality $A_1 = \varnothing$. Put $B_n = A_n \setminus A_{n-1}$ for $n = 2, 3, \dots$. Then $\{B_n\}$ is a disjoint family with $A_n = B_2 \cup B_3 \cup \dots \cup B_n$ and $A = \bigcup_{n=1}^{\infty} B_n$. Hence, by countable additivity of μ,

$$\mu(A) = \sum_{k=2}^{\infty} \mu(B_k) = \lim_{n\to\infty} \sum_{k=2}^{n} \mu(B_k) = \lim_{n\to\infty} \mu\left(\sum_{k=2}^{n} B_k\right) = \lim_{n\to\infty} \mu(A_n).$$

(b) $\to$ (a). Let $\{A_n\}$ be a family of disjoint sets in $\mathcal{A}$ with $\bigcup A_n = A \in \mathcal{A}$; put $B_k = A_1 \cup \dots \cup A_k$. Then B_k is an increasing to A sequence. By (b)

$$\mu(B_n) = \mu\left(\sum_{k=1}^{n} A_k\right) \underset{\mu \text{ is additive}}{=} \sum_{k=1}^{n} \mu(A_k) \underset{n\to\infty}{\longrightarrow} \mu(A) = \mu\left(\sum_{k=1}^{\infty} A_k\right);$$

Thus, $\displaystyle\sum_{k=1}^{\infty} \mu(A_k) = \mu\left(\sum_{k=1}^{\infty} A_k\right)$.

(b) $\to$ (c). Since A_n is decreasing to A, $A_1 \setminus A_n$ is increasing to $A_1 \setminus A$. By (b)

$$\mu(A_1 \setminus A_n) \underset{n\to\infty}{\longrightarrow} \mu(A_1 \setminus A),$$

hence $\mu(A_1) - \mu(A_n) \underset{n\to\infty}{\longrightarrow} \mu(A_1) - \mu(A)$ which implies the assertion.

(c) $\to$ (d) is trivial.

Now let μ be finite, in particular, $\mu(B) < \infty$ for all $B \in \mathcal{A}$. We show (d) $\to$ (b). Let (A_n) be an increasing to A sequence of subsets $A, A_n \in \mathcal{A}$. Then $(A \setminus A_n)$ is a decreasing to $\varnothing$ sequence. By (d), $\mu(A \setminus A_n) \underset{n\to\infty}{\longrightarrow} 0$. Since μ is subtractive (Proposition 10.3 (d)) and all values are finite, $\mu(A_n) \underset{n\to\infty}{\longrightarrow} \mu(A)$. ■

Proposition 10.5 *Let α be increasing and μ_α the corresponding Lebesgue–Stieltjes content on $\mathcal{E}_1$. Then μ_α is countably additive if and only if α is continuous from the right.*

Proof. Let μ_α be countably additive and $a \in \mathbb{R}$. Fix $a > 0$. By Proposition 10.4 (c) for any decreasing to a sequence (b_n) we have

$$\alpha(b_n) - \alpha(a) = \mu_\alpha((0, b_n]) - \mu_\alpha((0, a]) = \mu_\alpha((a, b_n]) \underset{n\to\infty}{\longrightarrow} \mu_\alpha(\varnothing) = 0;$$

hence, α is continuous from the right.
Now let α be continuous from the right, i.e. $\mu((a,b]) = \alpha(b) - \alpha(a)$. The proof of countable additivity of $\mu = \mu_\alpha$ is the hard part of the proposition. We will do this in the case

$$(a,b] = \bigcup_{k=1}^{\infty} (a_k, b_k]$$

with a disjoint family $[a_k, b_k)$ of intervals. By Proposition 10.3 (f) we already know

$$\mu\left((a,b]\right) \geq \sum_{k=1}^{\infty} \mu((a_k, b_k]). \tag{10.1}$$

We prove the opposite direction. Let $\varepsilon > 0$. Since α is continuous from the right at a, there exists $a_0 \in [a,b)$ such that $\alpha(a_0) - \alpha(a) < \varepsilon$ and, similarly, for every $k \in \mathbb{N}$ there exists $c_k > b_k$ such that $\alpha(c_k) - \alpha(b_k) < \varepsilon/2^k$. Hence,

$$[a_0, b] \subset \bigcup_{k=1}^{\infty} (a_k, b_k] \subset \bigcup_{k=1}^{\infty} (a_k, c_k)$$

is an open covering of a compact set. By Heine–Borel (Definition 6.11) there exists a finite subcover

$$[a_0, b] \subset \bigcup_{k=1}^{N} (a_k, c_k), \quad \text{hence} \quad (a_0, b] \subset \bigcup_{k=1}^{N} (a_k, c_k],$$

such that by Proposition 10.3 (e)

$$\mu((a_0, b]) \leq \sum_{k=1}^{N} \mu((a_k, c_k]).$$

By the choice of a_0 and c_k,

$$\mu((a_k, c_k]) = \mu((a_k, b_k]) + \alpha(c_k) - \alpha(b_k) \leq \mu((a_k, b_k]) + \frac{\varepsilon}{2^k}.$$

Similarly, $\mu((a,b]) = \mu((a, a_0]) + \mu((a_0, b])$ such that

$$\begin{aligned}
\mu((a,b]) &\leq \mu((a_0, b]) + \varepsilon \leq \sum_{k=1}^{N} \left(\mu((a_k, b_k]) + \frac{\varepsilon}{2^k}\right) + \varepsilon \\
&\leq \sum_{k=1}^{N} \mu((a_k, b_k]) + 2\varepsilon \leq \sum_{k=1}^{\infty} \mu((a_k, b_k]) + 2\varepsilon
\end{aligned}$$

Since ε was arbitrary,

$$\mu((a,b]) \leq \sum_{k=1}^{\infty} \mu((a_k, b_k]).$$

In view of (10.1), μ_α is countably additive. ∎

Corollary 10.6 *The correspondence $\mu \mapsto \alpha(\mu)$ from Example 10.1 (d) defines a bijection between countably additive functions μ on $\mathcal{E}_1$ and the monotonically increasing, right continuous functions α on $\mathbb{R}$ (up to constant functions, i. e. α and $\alpha + c$ define the same additive function).*

Historical Note. It was the great achievment of Émile Borel (1871–1956) that he really *proved* the countable additivity of the Lebesgue measure. He realized that the countable additivity of μ is a serious mathematical problem far from being evident.

10.1.2 Extension of Countably Additive Functions

Here we must stop the rigorous treatment of measure theory. Up to now, we know only two trivial examples of measures (Example 10.1 (a) and (b)). We give an outline of the steps toward the construction of the Lebesgue measure.

- Construction of an *outer measure* μ^* on $\mathcal{P}(X)$ from a countably additive function μ on $\mathcal{A}$.
- Construction of the σ-algebra $\mathcal{A}_\mu$ of *measurable sets.*

Theorem 10.7 (Extension and Uniqueness) *Let μ be a countably additive function on the algebra $\mathcal{A}$.*
(a) *There exists an extension of μ to a measure on the σ-algebra $\sigma(\mathcal{A})$ which coincides with μ on $\mathcal{A}$. We denote the measure on $\sigma(\mathcal{A})$ also by μ.*
(b) *This extension is unique, if $X = \sum_{n=1}^{\infty} A_n$, $A_n \in \mathcal{A}$, and $\mu(A_n) < \infty$ for all $n \in \mathbb{N}$.*

(Without proof)

Remarks 10.3 (a) The extension theory is due to Carathéodory (1914). For a detailed treatment, see [5, Section II.4].
(b) The property (b) is called *σ-finiteness* of μ. For the algebra of elementary sets $\mathcal{E}_n$ of $\mathbb{R}^n$ and the Lebesgue content it is obvioulsy satisfied since $\mathbb{R}^n$ is the countable disjoint union of bounded boxes: for example $\mathbb{R} = \sum_{n=1}^{\infty}((-n-1, -n] \cup (n, n+1])$.

10.1.3 The Lebesgue Measure on $\mathbb{R}^n$

Using the facts from the previous subsection we conclude that for any increasing, right continuous function α on $\mathbb{R}$ there exists a measure μ_α on the σ-algebra of Borel sets. We call this measure the *Lebesgue–Stieltjes measure on* $\mathbb{R}$. In case $\alpha(x) = x$ we call it the *Lebesgue measure.* Extending the Lebesgue content on elementary sets of $\mathbb{R}^n$ to the Borel algebra $\sigma(\mathcal{E}_n)$, we obtain the *n-dimensional Lebesgue measure* λ_n on $\mathbb{R}^n$.

Completeness

A measure $\mu\colon \mathcal{A} \to \overline{\mathbb{R}}_+$ on a σ-algebra $\mathcal{A}$ is said to be *complete* if $A \in \mathcal{A}$, $\mu(A) = 0$, and $B \subset A$ implies $B \in \mathcal{A}$. It turns out that the Lebesgue measure λ_n on the Borel sets of

$\mathbb{R}^n$ is not complete. Adjoining to $\sigma(\mathcal{E}_n)$ the subsets of measure-zero-sets, we obtain the σ-algebra $\mathcal{A}_{\lambda_n}$ of Legesgue measurable sets $\mathcal{A}_{\lambda_n}$:

$$\mathcal{A}_{\lambda_n} = \sigma(\mathcal{F}_n), \quad \mathcal{F}_n = \mathcal{E}_n \cup \{X \subset \mathbb{R}^n \mid X \subset E, \quad E \in \sigma(\mathcal{E}_n), \quad \lambda_n(E) = 0\}.$$

The Lebesgue measure λ_n on $\mathcal{A}_{\lambda_n}$ is now complete.

Remarks 10.4 (a) The Lebesgue measure is invariant under the *motion group* of $\mathbb{R}^n$. More precisely, let $\mathrm{O}(n) = \{T \in \mathrm{M}(n \times n, \mathbb{R}) \mid T^{\mathrm{t}}T = TT^{\mathrm{t}} = E_n\}$ be the group of real orthogonal $n \times n$-matrices ("motions"), then

$$\lambda_n(T(A)) = \lambda_n(A), \quad A \in \mathcal{A}_{\lambda_n}, \quad T \in \mathrm{O}(n).$$

(b) λ_n is *translation invariant*, i. e. $\lambda_n(A) = \lambda_n(x + A)$ for all $x \in \mathbb{R}^n$. Moreover, the invariance of λ_n under translations uniquely characterizes the Lebesgue measure λ_n: If λ is a translation invariant measure on $\sigma(\mathcal{E}_n)$, then $\lambda = c\lambda_n$ for some $c \in \mathbb{R}_+$.
(c) There exist non-measurable subsets in $\mathbb{R}^n$. We construct a subset E of $\mathbb{R}$ that is not Lebesgue measurable.
We write $x \sim y$ if $x - y$ is rational. This is an equivalence relation since $x \sim x$ for all $x \in \mathbb{R}$, $x \sim y$ implies $y \sim x$ for all x and y, and $x \sim y$ and $y \sim z$ implies $x \sim z$. Let E be a subset of $(0, 1)$ that contains exactly one point in every equivalence class. (the assertion that there is such a set E is a direct application of the *axiom of choice*). We claim that E is not measurable. Let $E + r = \{x + r \mid x \in E\}$. We need the following two properties of E:

(a) If $x \in (0, 1)$, then $x \in E + r$ for some rational $r \in (-1, 1)$.
(b) If r and s are distinct rationals, then $(E + r) \cap (E + s) = \varnothing$.

To prove (a), note that for every $x \in (0, 1)$ there exists $y \in E$ with $x \sim y$. If $r = x - y$, then $x = y + r \in E + r$.
To prove (b), suppose that $x \in (E + r) \cap (E + s)$. Then $x = y + r = z + s$ for some $y, z \in E$. Since $y - z = s - r \neq 0$, we have $y \sim z$, and E contains two equivalent points, in contradiction to our choice of E.
Now assume that E is Lebesgue measurable with $\lambda(E) = \alpha$. Define $S = \bigcup(E + r)$ where the union is over all rational $r \in (-1, 1)$. By (b), the sets $E + r$ are pairwise disjoint; since λ is translation invariant, $\lambda(E + r) = \lambda(E) = \alpha$ for all r. Since $S \subset (-1, 2)$, $\lambda(S) \leq 3$. The countable additivity of λ now forces $\alpha = 0$ and hence $\lambda(S) = 0$. But (a) implies $(0, 1) \subset S$, hence $1 \leq \lambda(S)$, and we have a contradiction.
(d) Any countable set has Lebesgue measure zero. Indeed, every single point is a box with edges of length 0; hence $\lambda(\{\mathrm{pt}\}) = 0$. Since λ is countably additive,

$$\lambda(\{x_1, x_2, \dots, x_n, \dots\}) = \sum_{n=1}^{\infty} \lambda(\{x_n\}) = 0.$$

In particular, the rational numbers have Lebesgue measure 0, $\lambda(\mathbb{Q}) = 0$.

(e) There are uncountable sets with measure zero. The Cantor set (Cantor: 1845–1918, inventor of set theory) is a prominent example:

$$C = \left\{ \sum_{i=1}^{\infty} \frac{a_i}{3^i} \;\middle|\; a_i \in \{0,2\} \right\};$$

Obviously, $C \subset [0,1]$; C is compact and can be written as the intersection of a decreasing sequence (C_n) of closed subsets; $C_1 = [0,1/3] \cup [2/3,1]$, $\lambda(C_1) = 2/3$, and, recursively,

$$C_{n+1} = \frac{1}{3}C_n \cup \left(\frac{2}{3} + \frac{1}{3}C_n\right).$$

Clearly,

$$\lambda(C_{n+1}) = \frac{2}{3}\lambda(C_n) = \cdots = \left(\frac{2}{3}\right)^n \lambda(C_1) = \left(\frac{2}{3}\right)^{n+1}.$$

By Proposition 10.4 (c), $\lambda(C) = \lim\limits_{n\to\infty} \lambda(C_n) = 0$. However, C has the same cardinality as $\{0,2\}^{\mathbb{N}} \cong \{0,1\}^{\mathbb{N}} \cong \mathbb{R}$ which is uncountable.

10.2 Measurable Functions

Let $\mathcal{A}$ be a σ-algebra over X.

Definition 10.5 A real function $f\colon X \to \mathbb{R}$ is called $\mathcal{A}$-*measurable* if for all $a \in \mathbb{R}$ the set $\{x \in X \mid f(x) > a\}$ belongs to $\mathcal{A}$.
A complex function f is said to be $\mathcal{A}$-*measurable* if both $\operatorname{Re} f$ and $\operatorname{Im} f$ are $\mathcal{A}$-measurable.
A function $f\colon U \to \mathbb{R}$, $U \subset \mathbb{R}^n$, is said to be a *Borel function* if f is $\sigma(\mathcal{E}_n)$-measurable, i. e. f is measurable with respect to the Borel algebra on $\mathbb{R}^n$.
A function $f\colon U \to V$, $U \subset \mathbb{R}^n$, $V \subset \mathbb{R}^m$, is called a *Borel function* if $f^{-1}(B)$ is a Borel set for all Borel sets $B \subset V$.

Note that $\{x \in X \mid f(x) > a\} = f^{-1}((a,+\infty))$. From Proposition 10.8 below it becomes clear that the last two notions are consistent. Note that no measure on $(X,\mathcal{A})$ needs to be specified to define a measurable function.

Example 10.2 (a) Any continuous function $f\colon U \to \mathbb{R}$, $U \subset \mathbb{R}^n$, is a Borel function. Indeed, since f is continuous and $(a,+\infty)$ is open, $f^{-1}((a,+\infty))$ is open as well and hence a Borel set (cf. Proposition 10.2).
(b) The characteristic function χ_A is $\mathcal{A}$-measurable if and only if $A \in \mathcal{A}$ (see homework 35.3).
(c) Let $f\colon U \to V$ and $g\colon V \to W$ be Borel functions. Then $g\circ f\colon U \to W$ is a Borel function, too. Indeed, for any Borel set $C \subset W$, $g^{-1}(C)$ is a Borel set in V since g is a Borel function. Since f is a Borel function $(g\circ f)^{-1}(C) = f^{-1}(g^{-1}(C))$ is a Borel subset of U which shows the assertion.

Proposition 10.8 *Let $f\colon X \to \overline{\mathbb{R}}$ be a function. The following are equivalent*

(a) $\{x \mid f(x) > a\} \in \mathcal{A}$ *for all* $a \in \mathbb{R}$ *(i. e.* f *is* $\mathcal{A}$*-measurable).*
(b) $\{x \mid f(x) \geq a\} \in \mathcal{A}$ *for all* $a \in \mathbb{R}$.
(c) $\{x \mid f(x) < a\} \in \mathcal{A}$ *for all* $a \in \mathbb{R}$.
(d) $\{x \mid f(x) \leq a\} \in \mathcal{A}$ *for all* $a \in \mathbb{R}$.
(e) $f^{-1}(B) \in \mathcal{A}$ *for all Borel sets* $B \in \sigma(\mathcal{E}_1)$.

Proof. (a) $\to$ (b) follows from the identity

$$\{x \mid f(x) \geq a\} = \bigcap_{n\in\mathbb{N}} \{x \mid f(x) > a + 1/n\}.$$

Since f is $\mathcal{A}$-measurable and $\mathcal{A}$ is a σ-algebra, the countable intersection on the right is in $\mathcal{A}$.
(a) $\to$ (d) follows from $\{x \mid f(x) \leq a\} = \{x \mid f(x) > a\}^{\mathsf{c}}$. The remaining directions are left to the reader (see also homework 35.5). ∎

Lemma 10.9 *Let* $f, g\colon X \to \overline{\mathbb{R}}$ *be* $\mathcal{A}$*-measurable. Then* $\{x \mid f(x) > g(x)\}$ *and* $\{x \mid f(x) = g(x)\}$ *are in* $\mathcal{A}$.

Proof. Since

$$\{x \mid f(x) < g(x)\} = \bigcup_{q\in\mathbb{Q}} \left(\{x \mid f(x) < q\} \cap \{x \mid q < g(x)\}\right),$$

and all sets $\{f < q\}$ and $\{q < g\}$ the right are in $\mathcal{A}$, and on the right there is a countable union, the right hand side is in $\mathcal{A}$. A similar argument works for $\{f > g\}$. Note that the sets $\{f \geq g\}$ and $\{f \leq g\}$ are the complements of $\{f < g\}$ and $\{f > g\}$, respectively; hence they belong to $\mathcal{A}$ as well. Finally, $\{f = g\} = \{f \geq g\} \cap \{f \leq g\}$. ∎

It is easy to see (cf. homework 35.4) that for any sequence (a_n) of real numbers

$$\overline{\lim_{n\to\infty}}\, a_n = \inf_{n\in\mathbb{N}} \sup_{k\geq n} a_k \quad \textit{and} \quad \underline{\lim_{n\to\infty}}\, a_n = \sup_{n\in\mathbb{N}} \inf_{k\geq n} a_k. \tag{10.2}$$

Proposition 10.10 *Let* (f_n) *be a sequence of* $\mathcal{A}$*-measurable functions on* X*. Then* $\sup\limits_{n\in\mathbb{N}} f_n$, $\inf\limits_{n\in\mathbb{N}} f_n$, $\overline{\lim\limits_{n\to\infty}}\, f_n$, $\underline{\lim\limits_{n\to\infty}}\, f_n$ *are* $\mathcal{A}$*-measurable. In particular* $\lim\limits_{n\to\infty} f_n$ *is measurable if the limit exists.*

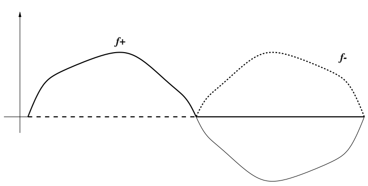
Proof. Note that for all $a \in \mathbb{R}$ we have (cf. homework 35.4)

$$\{\sup f_n \leq a\} = \bigcap_{n\in\mathbb{N}} \{f_n \leq a\}.$$

Since all f_n are measurable, so is $\sup f_n$. A similar proof works for $\inf f_n$. By (10.2), $\overline{\lim}_{n\to\infty} f_n$ and $\underline{\lim}_{n\to\infty} f_n$, are measurable, too. ∎

Proposition 10.11 *Let $f, g\colon X \to \mathbb{R}$ Borel functions on $X \subset \mathbb{R}^n$. Then $\alpha f + \beta g$, $f \cdot g$, and $|\,f\,|$ are Borel functions, too.*

Proof. The function $h(x) = (f(x), g(x))\colon X \to \mathbb{R}^2$ is a Borel function since its coordinate functions are so. Since the sum $s(x, y) = x + y$ and the product $p(x, y) = xy$ are continuous functions, the compositions $s{\circ}h$ and $p{\circ}h$ are Borel functions by Example 10.2 (c). Since the constant functions α and β are Borel, so are αf, βg, and finally $\alpha f + \beta g$. Hence, the Borel functions over X form a linear space, moreover a real algebra. In particular $-f$ is Borel and so is $|\,f\,| = \max\{f, -f\}$. ∎



Let $(X, \mathcal{A}, \mu)$ be a measure space and $f\colon X \to \overline{\mathbb{R}}$ arbitrary. Let $f^+ = \max\{f, 0\}$ and $f^- = \max\{-f, 0\}$ denote the *positive and negative parts of f*. We have $f = f^+ - f^-$ and $|\,f\,| = f^+ + f^-$; moreover f^+, $f^- \geq 0$.

Corollary 10.12 *Let f is a Borel function if and only if both f^+ and f^- are Borel.*

10.3 The Lebesgue Integral

We define the Lebesgue integral of a complex function in 3 steps; first for positive, simple functions, then for positive measurable functions and finally for arbitrary measurable functions. In this section $(X, \mathcal{A}, \mu)$ is a measure space.

10.3.1 Simple Functions

Definition 10.6 Let $M \subseteq X$ be a subset. The function

$$\chi_M(x) = \begin{cases} 1, & x \in M, \\ 0, & x \notin M, \end{cases}$$

is called *characteristic function* of M.
An $\mathcal{A}$-measurabel function $f\colon X \to \mathbb{R}$ is called *simple* if f takes only finitely many values $c_1, \dots, c_n$.

Clearly, if $c_1, \dots, c_n$ are the distinct values of the simple function f, then

$$f = \sum_{i=1}^{n} c_i \chi_{A_i},$$

where $A_i = \{x \mid f(x) = c_i\}$. It is clear, see homework 35.2, that f measurable if and only if $A_i \in \mathcal{A}$ for all i. Obviously, $\{A_i \mid i = 1, \dots, n\}$ is a disjoint family of subsets of X.
We denote the set of simple functions on $(X, \mathcal{A})$ by $\mathcal{S}$; the set of non-negative simple functions is denoted by $\mathcal{S}_+$. It is easy to see that $f, g \in \mathcal{S}$ implies $\alpha f + \beta g \in \mathcal{S}$, $\max\{f, g\} \in \mathcal{S}$, $\min\{f, g\} \in \mathcal{S}$, and $fg \in \mathcal{S}$.

For $f = \sum_{i=1}^{n} c_i \chi_{A_i} \in \mathcal{S}_+$ define

$$\int_X f \, \mathrm{d}\mu = \sum_{i=1}^{n} c_i \, \mu(A_i). \tag{10.3}$$

The convention $0 \cdot (+\infty)$ is used here; it may happen that $c_i = 0$ for some i and $\mu(A_i) = +\infty$.

Remarks 10.5 (a) Since $c_i \geq 0$ for all i, the right hand side is well-defined in $\overline{\mathbb{R}}$.
(b) Given another presentation of f, say, $f(x) = \sum_{j=1}^{m} d_j \chi_{B_j}$, $\sum_{j=1}^{m} d_j \, \mu(B_j)$ gives the same value as (10.3).

The following properties are easily checked.

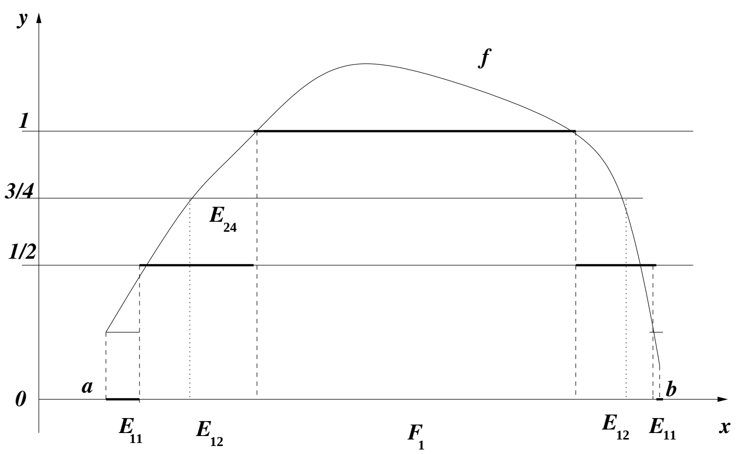
Lemma 10.13 *For $f, g \in \mathcal{S}_+$, $A \in \mathcal{A}$, $c \in \mathbb{R}_+$ we have*

(1) $\int_X \chi_A \, \mathrm{d}\mu = \mu(A)$.
(2) $\int_X cf \, \mathrm{d}\mu = c \int_X f \, \mathrm{d}\mu$.
(3) $\int_X (f+g) \, \mathrm{d}\mu = \int_X f \, \mathrm{d}\mu + \int_X g \, \mathrm{d}\mu$.
(4) $f \leq g$ *implies* $\int_X f \, \mathrm{d}\mu \leq \int_X g \, \mathrm{d}\mu$.

10.3.2 Approximation of Measurable Functions by Simple Functions

Theorem 10.14 *Let $f\colon X \to [0, +\infty]$ be measurable. There exist simple functions s_n, $n \in \mathbb{N}$, on X such that*

(a) $0 \leq s_1 \leq s_2 \leq \cdots \leq f$.
(b) $s_n(x) \xrightarrow[n\to\infty]{} f(x)$, *as $n \to \infty$, for every $x \in X$.*



Example. $X = (a, b)$, $n = 1$, $1 \leq i \leq 2$. Then

$$E_{11} = f^{-1}\left(\left[0, \frac{1}{2}\right)\right),$$
$$E_{12} = f^{-1}\left(\left(\frac{1}{2}, 1\right)\right),$$
$$F_2 = f^{-1}\left([1, +\infty]\right).$$

Proof. For $n \in \mathbb{N}$ and for $1 \leq i \leq n2^n$, define

$$E_{ni} = f^{-1}\left(\left[\frac{i-1}{2^n}, \frac{i}{2^n}\right)\right) \quad \text{and} \quad F_n = f^{-1}([n, \infty])$$

and put

$$s_n = \sum_{i=1}^{n2^n} \frac{i-1}{2^n} \chi_{E_{ni}} + n\chi_{F_n}.$$

Proposition 10.8 shows that E_{ni} and F_n are measurable sets. It is easily seen that the functions s_n satisfy (a). If x is such that $f(x) < +\infty$, then

$$0 \leq f(x) - s_n(x) \leq \frac{1}{2^n} \tag{10.4}$$

as soon as n is large enough, that is, $x \in E_{ni}$ for some $n, i \in \mathbb{N}$ and not $x \in F_n$. If $f(x) = +\infty$, then $s_n(x) = n$; this proves (b). ∎

From (10.4) it follows, that $s_n \rightrightarrows f$ uniformly on X if f is bounded.

Definition 10.7 (Lebesgue Integral) Let $f\colon X \to [0, +\infty]$ be measurable. Let $(s_n(x))$ be an increasing sequence of simple functions s_n converging to $f(x)$ for all $x \in X$, $\lim_{n\to\infty} s_n(x) = \sup_{n\in\mathbb{N}} s_n(x) = f(x)$. Define

$$\int_X f \, \mathrm{d}\mu = \sup_{n\in\mathbb{N}} \int_X s_n \, \mathrm{d}\mu \tag{10.5}$$

and call this number in $[0, +\infty]$ the *Lebesgue integral of $f(x)$ over X with respect to the measure μ* or *μ-integral of f over X*.

The definition of the Lebesgue integral does not depend on the special choice of the increasing functions $s_n \nearrow f$. One can define

$$\int_X f \, \mathrm{d}\mu = \sup\left\{ \int_X s \, \mathrm{d}\mu \mid s \leq f, \quad \text{and } s \text{ is a simple function} \right\}.$$

Observe, that we apparently have two definitions for $\int_X f \, \mathrm{d}\mu$ if f is a simple function. However these assign the same value to the integral since f is the largest simple function greater than or equal to f.

Proposition 10.15 *The properties* (1) *to* (4) *from Lemma 10.13 hold for any non-negative measurable functions $f, g\colon X \to [0, +\infty]$, $c \in \mathbb{R}_+$.*

(Without proof.)

10.3.3 Integration of Arbitrary Complex Functions

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

Definition 10.8 (Lebesgue Integral—Continued) A complex, measurable function $f\colon X \to \mathbb{C}$ is called *μ-integrable* if

$$\int_X |\, f \,| \,\mathrm{d}\mu < \infty.$$

If $f = u + \mathrm{i}v$ is μ-integrable, where $u = \operatorname{Re} f$ and $v = \operatorname{Im} f$ are the real and imaginary parts of f, u and v are real measurable functions on X. Define the *μ-integral of f over X* by

$$\int_X f \,\mathrm{d}\mu = \int_X u^+ \,\mathrm{d}\mu - \int_X u^- \,\mathrm{d}\mu + \mathrm{i}\int_X v^+ \,\mathrm{d}\mu - \mathrm{i}\int_X v^- \,\mathrm{d}\mu. \tag{10.6}$$

These four functions u^+, u^-, v^+, and v^- are measurable, real, and non-negative. Since we have $u^+ \leq |\, u \,| \leq |\, f \,|$ etc., each of these four integrals is finite. Thus, (10.6) defines the integral on the left as a complex number.

We define $\mathcal{L}^1(X, \mu)$ to be the collection of all complex μ-integrable functions f on X. It is sometimes desirable to define the integral of a measurable function $f\colon X \to [-\infty, +\infty]$ to be

$$\int_X f \,\mathrm{d}\mu = \int_X f^+ - \int_X f^- \,\mathrm{d}\mu,$$

if *at least one* of the integrals on the right is finite. The left side is then a number in $[-\infty, +\infty]$.
Note that for an integrable functions f, $\int_X f \,\mathrm{d}\mu$ is a finite number.

Proposition 10.16 *Let $f, g\colon X \to \mathbb{C}$ be measurable.*

(a) *f is μ-integrable if and only if $|\, f \,|$ is μ-integrable and we have*

$$\left| \int_X f \,\mathrm{d}\mu \right| \leq \int_X |\, f \,| \,\mathrm{d}\mu.$$

(b) *f is μ-integrable if and only if there exists an integrable function h with $|\, f \,| \leq h$.*

(c) *If f, g are integrable, so is $c_1 f + c_2 g$ where*

$$\int_X (c_1 f + c_2 g) \,\mathrm{d}\mu = c_1 \int_X f \,\mathrm{d}\mu + c_2 \int_X g \,\mathrm{d}\mu.$$

(d) *If $f \leq g$ on X, then $\int_X f \,\mathrm{d}\mu \leq \int_X g \,\mathrm{d}\mu$.*

It follows that the set $\mathcal{L}^1(X, \mu)$ of μ-integrable complex-valued functions on X is a linear space. The Lebesgue integral defines a positive linear functional on $\mathcal{L}^1(X, \mu)$. Note that (b) implies that any measurable and bounded function f on a space X with $\mu(X) < \infty$ is integrable.

Definition 10.9 Let $A \in \mathcal{A}$, $f\colon X \to \overline{\mathbb{R}}$ or $f\colon X \to \mathbb{C}$ measurable. The function f is called *μ-integrable over A* if $\chi_A f$ is μ-integrable over X. In this case we put,

$$\int_A f \,\mathrm{d}\mu = \int_X \chi_A f \,\mathrm{d}\mu.$$

In particular, Lemma 10.13 (1) now reads $\int_A \mathrm{d}\mu = \mu(A)$.

10.4 Some Theorems on Lebesgue Integrals

10.4.1 The Role Played by Measure Zero Sets

Definition 10.10 Let P be a property which a point x may or may not have. For instance, P may be the property "$f(x) > 0$" if f is a given function or "$(f_n(x))$ converges" if (f_n) is a given sequence of functions. If $(X, \mathcal{A}, \mu)$ is a measure space and $A \in \mathcal{A}$, we say *"P holds almost everywhere on A"*, abbreviated by "P holds a. e. on E", if there exists $N \in \mathcal{A}$ such that $\mu(N) = 0$ and P holds for every point $x \in E \setminus N$.

This concept of course strongly depends on the measure μ, and sometimes we write μ-a. e. For example, if f, g are measurable functions, we write $f \sim g$ if $f = g$ μ-a. e. on X. This means $\mu(\{x \mid f(x) \neq g(x)\}) = 0$. This is indeed an equivalence relation. Note that $f = g$ a. e. on X implies

$$\int_X f \,\mathrm{d}\mu = \int_X g \,\mathrm{d}\mu.$$

Indeed, let N denote the zero-set where $f \neq g$. Then

$$\int_X (f-g) \,\mathrm{d}\mu = \int_N (f-g) \,\mathrm{d}\mu + \int_{X \setminus N} (f-g) \,\mathrm{d}\mu \leq \mu(N)(\infty) + \mu(X \setminus N) \cdot 0 = 0.$$

Proposition 10.17 *Let $f\colon X \to [0, +\infty]$ be measurable. Then*

$\int_X f \,\mathrm{d}\mu = 0$ *if and only if* $f = 0$ *a. e. on* X.

(See homework 36.3)

Definition 10.11 Let $(X, \mathcal{A}, \mu)$ be a measure space. For $1 \leq p < \infty$, $\mathcal{L}^p(X, \mu)$ denotes the set of $\mathcal{A}$-measurable functions f on X such that $|\, f\,|^p \in \mathcal{L}^1(X, \mu)$. For $f \in \mathcal{L}^p(X, \mu)$ set

$$\|f\|_p = \left(\int_X |\, f\,|^p \,\mathrm{d}\mu\right)^{\frac{1}{p}}. \tag{10.7}$$

From Proposition 10.17 it follows that $\|\cdot\|_p$ is not a norm, in general, since $\int_X |\, f\,|^p \,\mathrm{d}\mu = 0$ implies only $f = 0$ a. e. However, identifying functions f and g which are equal a. e. , (10.7) defines a norm on those equivalence classes.
Let $\mathcal{N} = \{f\colon X \to \mathbb{C} \mid \ f \text{ is measurable and } \ f = 0 \ \text{a. e. }\}$. Then $\mathcal{N}$ is a linear subspace of $\mathcal{L}^p(X, \mu)$ for all all p, and $f = g$ a. e. if and only if $f - g \in \mathcal{N}$.

Definition 10.12 Let $(X, \mathcal{A}, \mu)$ be a measure space. $\mathrm{L}^p(X, \mu)$ denotes the set of equivalence classes of functions of $\mathcal{L}^p(X, \mu)$ with respect to the equivalence relation $f = g$ a. e. that is,

$$\mathrm{L}^p(X, \mu) = \mathcal{L}^p(X, \mu)/\mathcal{N}$$

is the quotient space. $\mathrm{L}^p(X, \mu)$ is a normed space with the norm (10.7).

Example. We have $\chi_{\mathbb{Q}} = 0$ in $\mathrm{L}^p(\mathbb{R}, \lambda)$ since $\chi_{\mathbb{Q}} = 0$ a. e. on $\mathbb{R}$ with respect to the Lebesgue measure (note that $\mathbb{Q}$ is a set of measure zero).

10.4.2 The Monotone Convergence Theorem

The following theorem about the monotone convergence by Beppo Levi (1875–1961) is one of the most important in the theory of integration. The theorem holds for an *arbitrary* increasing sequence of measurable functions with, possibly, $\int_X f_n \,\mathrm{d}\mu = +\infty$.

Theorem 10.18 (Monotone Convergence Theorem/Beppo Levi) *Let (f_n) be a sequence of measurable functions on X and suppose that*

(1) $0 \le f_1(x) \le f_2(x) \le \cdots \le +\infty$ *for all* $x \in X$,
(2) $f_n(x) \underset{n\to\infty}{\longrightarrow} f(x)$, *for every* $x \in X$.

Then f is measurable, and

$$\lim_{n\to\infty} \int_X f_n \,\mathrm{d}\mu = \int_X f \,\mathrm{d}\mu.$$

(Without proof) Note, that f is measurable is a consequence of Proposition 10.10.

Corollary 10.19 *Let $f_n \colon X \to [0, +\infty]$ be measurable for all $n \in \mathbb{N}$, and $f(x) = \sum\limits_{n=1}^{\infty} f_n(x)$ for $x \in X$. Then*

$$\int_X \sum_{n=1}^{\infty} f_n \,\mathrm{d}\mu = \sum_{n=1}^{\infty} \int_X f_n \,\mathrm{d}\mu.$$

Example 10.3 (a) Let $X = \mathbb{N}$, $\mathcal{A} = \mathcal{P}(\mathbb{N})$ the σ-algebra of all subsets, and μ the counting measure on $\mathbb{N}$. The functions on $\mathbb{N}$ can be identified with the sequences (x_n), $f(n) = x_n$. Trivially, any function is $\mathcal{A}$-measurable.
What is $\int_{\mathbb{N}} x_n \,\mathrm{d}\mu$? First, let $f \ge 0$. For a simple function g_n, given by $g_n = x_n \chi_{\{n\}}$, we obtain $\int g_n \,\mathrm{d}\mu = x_n \mu(\{n\}) = x_n$. Note that $f = \sum\limits_{n=1}^{\infty} g_n$ and $g_n \ge 0$ since $x_n \ge 0$. By Corollary 10.19,

$$\int_{\mathbb{N}} f \,\mathrm{d}\mu = \sum_{n=1}^{\infty} \int_{\mathbb{N}} g_n \,\mathrm{d}\mu = \sum_{n=1}^{\infty} x_n.$$

Now, let f be arbitrary integrable, i. e. $\int_{\mathbb{N}} |\, f \,| \,\mathrm{d}\mu < \infty$; thus $\sum_{n=1}^{\infty} |\, x_n \,| < \infty$. Therefore, $(x_n) \in \mathcal{L}^1(\mathbb{N}, \mu)$ if and only if $\sum x_n$ converges *absolutely*. The space of absolutely convergent series is denoted by ℓ_1 or $\ell_1(\mathbb{N})$.

(b) Let $a_{nm} \geq 0$ for all $n, m \in \mathbb{N}$. Then

$$\sum_{n=1}^{\infty}\sum_{m=1}^{\infty} a_{mn} = \sum_{m=1}^{\infty}\sum_{n=1}^{\infty} a_{mn}.$$

Proof. Consider the measure space $(\mathbb{N}, \mathcal{P}(\mathbb{N}), \mu)$ from (a). For $n \in \mathbb{N}$ define functions $f_n(m) = a_{mn}$. By Corollary 10.19 we then have

$$\begin{aligned}\int_X \underbrace{\sum_{n=1}^{\infty} f_n(m)}_{f(m)} \,\mathrm{d}\mu &= \int_X f \,\mathrm{d}\mu \underset{(a)}{=} \sum_{m=1}^{\infty} f(m) = \sum_{m=1}^{\infty}\sum_{n=1}^{\infty} a_{mn} \\ &= \sum_{n=1}^{\infty}\int_X f_n(m)\,\mathrm{d}\mu = \sum_{n=1}^{\infty}\sum_{m=1}^{\infty} a_{mn}.\end{aligned}$$

■

Proposition 10.20 *Let $f\colon X \to [0, +\infty]$ be measurable. Then*

$$\varphi(A) = \int_A f \,\mathrm{d}\mu, \quad A \in \mathcal{A}$$

defines a measure φ on $\mathcal{A}$.

(See homework 36.4)

10.4.3 The Dominated Convergence Theorem

Besides the monotone convergence theorem the present theorem is the most important one. It is due to Henry Lebesgue. The great advantage, compared with Theorem7.7, is that $\mu(X) = \infty$ is allowed, that is, non-compact domains X are included. We only need the *pointwise* convergence of (f_n), not the *uniform* convergence. The main assumtion here is the existence of an integrable upper bound for all f_n.

Theorem 10.21 (Dominated Convergence/Lebesgue) *Let $f_n, f\colon X \to \overline{\mathbb{R}}$ or into $\mathbb{C}$, $n \in \mathbb{N}$, be measurable functions such that*

(1) $f_n \to f$ *as* $n \to \infty$ *a. e.*

(2) $|\, f_n \,| \leq g$ *a. e.*

(3) $\displaystyle\int_X g \,\mathrm{d}\mu < +\infty.$

Then $\int_X |\, f \,| \,\mathrm{d}\mu < \infty$ *and*

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

$$\lim_{n\to\infty}\int_X |\, f_n - f \,| \,\mathrm{d}\mu = 0. \tag{10.8}$$

Note, that (10.8) shows that (f_n) converges to f in the normed space $\mathrm{L}^1(X, \mu)$.

Example 10.4 (a) Let $A_n \in \mathcal{A}$, $n \in \mathbb{N}$, $A_1 \subset A_2 \subset \cdots$ be an increasing sequence with $\bigcup\limits_{n\in\mathbb{N}} A_n = A$.
If $f \in \mathcal{L}^1(A, \mu)$, then

$$\lim_{n\to\infty} \int_{A_n} f \,\mathrm{d}\mu = \int_A f \,\mathrm{d}\mu.$$

Indeed, the sequence $(\chi_{A_n} f)$ is pointwise converging to $\chi_A f$ and $|\,\chi_{A_n} f\,| \le |\,\chi_A f\,|$ which is integrable. By Lebesgue's theorem,

$$\lim_{n\to\infty} \int_{A_n} f \,\mathrm{d}\mu = \lim_{n\to\infty} \int_X \chi_{A_n} f = \int_X \chi_A f \,\mathrm{d}\mu = \int_A f \,\mathrm{d}\mu.$$

However, if we do not assume $f \in \mathcal{L}^1(A, \mu)$, the statement is not true (see Remark 10.6 below).
(b) Let $f_n = n\chi_{(0,1/n)}$ be defined on $(0, 1)$. Then $f_n \to 0$ pointwise, however

$$\int_0^1 f_n \,\mathrm{d}\mu = 1 \not\to \int_0^1 0 = 0.$$

Note, that f_n is neither monotonic nor dominated by any integrable function g.

10.4.4 The Riemann and the Lebesgue Integrals

Proposition 10.22 *Let f be a bounded function on the finite interval $[a, b]$.*
(a) *f is Riemann integrable on $[a, b]$ if and only if f is continuous a. e. on $[a, b]$.*
(b) *If f is Riemann integrable on $[a, b]$, then f is Lebesgue integrable, too. Both integrals coincide.*
Let $I \subset \mathbb{R}$ be an interval such that f is Riemann integrable on all compact subintervals of I.
(c) *f is Lebesgue integrable on I if and only if $|\,f\,|$ is improperly Riemann integrable on I (see Section 5.4); both integrals coincide.*

Remarks 10.6 (a) The characteristic function $\chi_{\mathbb{Q}}$ on $[0, 1]$ is Lebesgue but not Riemann integrable; $\chi_{\mathbb{Q}}$ is nowhere continuous on $[0, 1]$.
(b) The (improper) Riemann integral

$$\int_1^\infty \frac{\sin x}{x} \,\mathrm{d}x$$

converges (see Example 5.9); however, the Lebesgue integral does not exist since the integral does not converge absolutely.

10.4.5 Fubini's Theorem

Theorem 10.23 *Let $(X_1, \mathcal{A}_1, \mu_1)$ and $(X_2, \mathcal{A}_2, \mu_2)$ be σ-finite measure spaces, let f be an $\mathcal{A}_1 \otimes \mathcal{A}_2$-measurable function and $X = X_1 \times X_2$.*
(a) *If $f\colon X \to [0, +\infty]$, $\varphi(x_1) = \int\limits_{X_2} f(x_1, x_2)\, \mathrm{d}\mu_2$, $\psi(x_2) = \int\limits_{X_1} f(x_1, x_2)\, \mathrm{d}\mu_1$, then*

$$\int_{X_2} \psi(x_2)\, \mathrm{d}\mu_2 = \int\limits_{X_1 \times X_2} f\, \mathrm{d}(\mu_1 \otimes \mu_2) = \int_{X_1} \varphi(x_1)\, \mathrm{d}\mu_1.$$

(b) *If $f \in \mathcal{L}^1(X, \mu_1 \otimes \mu_2)$ then*

$$\int_{X_1 \times X_2} f\, \mathrm{d}(\mu_1 \otimes \mu_2) = \int_{X_2} \left(\int_{X_1} f(x_1, x_2)\, \mathrm{d}\mu_1 \right) \mathrm{d}\mu_2.$$

Here $\mathcal{A}_1 \otimes \mathcal{A}_2$ denotes the smallest σ-algebra over X, which contains all sets $A \times B$, $A \in \mathcal{A}_1$ and $B \in \mathcal{A}_2$. Define $\mu(A \times B) = \mu_1(A)\mu_2(B)$ and extend μ to a measure $\mu_1 \otimes \mu_2$ on $\mathcal{A}_1 \otimes \mathcal{A}_2$.

Remark 10.7 In (a), as in Levi's theorem, we don't need any assumption on f to change the order of integration since $f \geq 0$. In (b) f is an arbitrary measurable function on $X_1 \times X_2$, however, the integral $\int_X |\, f \,|\, \mathrm{d}\mu$ needs to be finite.

Bibliography

[1] I. Agricola and T. Friedrich, *Globale Analysis (in German)*, Friedr. Vieweg & Sohn, Braunschweig, 2001.

[2] Glen E. Bredon, *Topology and geometry*, Graduate Texts in Mathematics, no. 139, Springer-Verlag, New York, 1997.

[3] R. Courant, *Differential and integral calculus I–II*, Wiley Classics Library, John Wiley & Sons, New York etc., 1988.

[4] J. Dieudonné, *Treatise on analysis. Volume I – IX*, Pure and Applied Mathematics, Academic Press, Boston, 1993.

[5] J. Elstrodt, *Maß- und Integrationstheorie*, third ed., Springer, Berlin, 2002.

[6] O. Forster, *Analysis 1 – 3 (in German)*, Vieweg Studium: Grundkurs Mathematik, Vieweg, Braunschweig, 2001.

[7] K. Königsberger, *Analysis 1 (English)*, Springer-Verlag, Berlin, Heidelberg, New York, 1990.

[8] S. Lang, *Undergraduate Analysis*, second ed., Undergraduate texts in mathematics, Springer, New York-Heidelberg, 1989.

[9] J. Marsden and A. Weinstein, *Calculus. I, II, III*, Undergraduate Texts in Mathematics, Springer-Verlag, New York etc., 1985.

[10] P. V. O'Neil, *Advanced calculus*, Collier Macmillan Publishing Co., London, 1975.

[11] W. Rudin, *Real and Complex Analysis*, International Student Edition, McGraw-Hill Book Co., New York-Toronto, 1966.

[12] ———, *Principles of mathematical analysis*, third ed., International Series in Pure and Applied Mathematics, McGraw-Hill Book Co., New York-Auckland-Düsseldorf, 1976.

[13] M. Spivak, *Calculus on manifolds*, W. A. Benjamin, New York, Amsterdam, 1965.

[14] ———, *Calculus*, Publish or Perish, Inc., Berkeley, California, 1980.

[15] C. von Westenholz, *Differential forms in mathematical physics*, second ed., Studies in Mathematics and its Applications, no. 3, North-Holland Publishing Co., Amsterdam-New York, 1981.