2 Measures and integration

In Definition 1.30 we introduced a uniform metric d on the space $C[a,b]$ and noted in Theorem 1.32 that the metric space $(C[a,b],d)$ is complete. There are, however, other useful metrics on $C[a,b]$ which are defined by integrals.

To begin with, let $\int_a^b f(x)\,dx$ denote the usual Riemann integral of a function $f \in C[a,b]$. Then, for $1 \leq p < \infty$, the function $d_p : C[a,b] \times C[a,b] \longrightarrow \mathbb{R}$ defined by

$$d_p(f,g) = \left(\int_a^b |f(x) - g(x)|^p \, dx \right)^{1/p}$$

is a metric on $C[a,b]$. (The triangle inequality for d_p follows by Minkowski's inequality, which will be proved later.) Unfortunately, the metric space $(C[a,b],d_p)$ is not complete. Since complete metric spaces are much more important than non-complete ones, the situation is undesirable.

Heuristically, the problem is that Cauchy sequences in $(C[a,b],d_p)$ "coverge" to functions which do not belong to $C[a,b]$, and hence may not be Riemann integrable. This is a weakness of the Riemann integral and the remedy is to replace it with the *Lebesgue integral*.

The Lebesgue integral is a more powerful theory of integration which enables a wider class of functions to be integrated. The whole idea is based on the theory of measures. Following the presentation given in *Royden:* Real Analysis, we shall begin with the measures and outer measures in $\mathbb{R}$.

Outer measure in $\mathbb{R}$

Fundamental to the theory of Lebesgue integration is the idea of the size or "measure" of a given set. For instance, for any bounded interval $I = [a,b]$ we say that I has length $l(I) = b - a$.

To define the Lebesgue integral on $\mathbb{R}$ it is necessary to be able to assign a "length" or "measure" to a much more broader class of sets than simply intervals. If m denotes such a measure, we would like it to possess the following properties:

(1) $m(E)$ is defined for every set $E \subset \mathbb{R}$.

(2) For every interval $I \subset \mathbb{R}$, we have $m(I) = l(I)$.

(3) If $\{E_n\}$ is a sequence of disjoint sets in $\mathbb{R}$ for which the measure m is defined, then

$$m\left(\bigcup_n E_n\right) = \sum_n m(E_n),$$

that is, m is *countably additive*.

(4) m is translation invariant, that is, if $E \subset \mathbb{R}$ is a set for which the measure m is defined and if, for any $a \in \mathbb{R}$,

$$a + E = \{a + x \mid x \in E\},$$

then $m(a+E) = m(E)$.

Unfortunately, in general it is not possible to construct a useful definition of "measure" which satisfies all of the above four properties at the same time, see e. g. *Royden,* Chapter 3.4. One of the requirements has to be dropped.

To begin with, we define an outer measure in $\mathbb{R}$.

Definition 2.1 The Lebesgue outer measure of $A \subset \mathbb{R}$ is given by

$$m^*(A) = \inf \left\{ \sum_n l(I_n) \mid I_n\text{'s open intervals such that } A \subset \bigcup_n I_n \right\},$$

where $l(I_n)$ is the usual geometric length of I_n.

Remarks. (1) Since for every $A \subset \mathbb{R}$ we have

$$A \subset \bigcup_{n=1}^{\infty}]-n, n[= \mathbb{R},$$

the outer measure $m^*(A)$ is defined for every $A \subset \mathbb{R}$.

(2) For any $A \subset \mathbb{R}$, $0 \leq m^*(A) \leq \infty$.

Example. Let $A \subset \mathbb{R}$ and $B \subset \mathbb{R}$ such that $A \subset B$. Show that $m^*(A) \leq m^*(B)$.

Theorem 2.2 *The outer measure of a real interval is its length.*

Theorem 2.3 *Let $\{E_n\}$ be a sequence of subsets of $\mathbb{R}$. Then*

$$m^* \left(\bigcup_n A_n \right) \leq \sum_n m^*(A_n),$$

that is, m^* is subadditive.

Unfortunately, outer measure is not countably additive (see *Royden*, page 56), that is, it does not satisfy the property (3). However, Theorem 2.3 has the following two consequences.

Corollary 2.4 *If $A \subset \mathbb{R}$ is countable (numeroituva), then $m^*(A) = 0$.*

Corollary 2.5 *Any real interval is uncountable (ylinumeroituva).*

Finally, we see that the outer measure satisfies the property (4).

Theorem 2.6 *Outer measure m^* is translation invariant.*

Measurable sets and Lebesgue measure

While outer measure has the advantage that it is defined for all sets, it is not countably additive. It becomes countably additive, however, if we suitably reduce the family of sets on which it is defined.

Definition 2.7 A set $E \subset \mathbb{R}$ is said to be (Lebesgue) measurable if for each set $A \subset \mathbb{R}$ we have

$$m^*(A) = m^*(A \cap E) + m^*(A \setminus E).$$

The set of all (Lebesgue) measurable sets is denoted by $\mathcal{M}$. The restriction of an outer measure to $\mathcal{M}$ is called a (Lebesgue) measure and it is denoted by m.

Remarks. (1) A measure m is a mapping $m : \mathcal{M} \longrightarrow \mathbb{R}_+ \cup \{\infty\}$.

(2) If $E \subset \mathbb{R}$ is measurable, then $m(E) = m^*(E)$.

(3) Measures are defined only for measurable sets: if a given set is not measurable, we cannot discuss about its measure.

(4) Since the outer measure is subadditive by Theorem 2.3, we have for all sets $A \subset \mathbb{R}$ that

$$m^*(A) \leq m^*(A \cap E) + m^*(A \setminus E).$$

Hence, in proving that a given set $E \subset \mathbb{R}$ is measurable, it is enough to show that

$$m^*(A) \geq m^*(A \cap E) + m^*(A \setminus E) \tag{2.1}$$

holds for all $A \subset \mathbb{R}$ satisfying $m^*(A) < \infty$. For if $m^*(A) = \infty$, then (2.1) clearly holds.

Theorem 2.8 *If $E \subset \mathbb{R}$ is measurable, then $\mathbb{R} \setminus E$ is measurable.*

Theorem 2.9 *Let $E \subset \mathbb{R}$. If $m^*(E) = 0$, then E is measurable.*

Theorem 2.9 has the following direct consequence.

Corollary 2.10 *Countable sets of real numbers are measurable and their measure is equal to zero.*

As we already noted, $E \in \mathcal{M}$ implies $m(E) = m^*(E)$. Therefore, the following result (property (3) for m) follows from Theorem 2.6.

Corollary 2.11 *The measure m is translation invariant.*

Going back to properties (1)–(4), we know that m satisfies (4) but does not satisfy (1). To see that m satisfies (2) and (3), we state the following two results. Their proofs are rather long and therefore omitted. The interested reader may look at the book of *Royden* for details.

Theorem 2.12 *If $\{E_n\}$ is a sequence of measurable sets, then*

$$\bigcup_n E_n \qquad \text{and} \qquad \bigcap_n E_n$$

are measurable. Further,

$$m\left(\bigcup_n E_n\right) \leq \sum_n m(E_n),$$

and if all sets E_n are pairwise disjoint (pistevieraita), we have

$$m\left(\bigcup_n E_n\right) = \sum_n m(E_n).$$

Theorem 2.13 *All real intervals are measurable.*

We next characterize a concept which appears to be very useful.

Definition 2.14 The triple $(\mathbb{R}, \mathcal{M}, m)$ is called a measure space (mitta-avaruus).

There exist, of course, other measure spaces than just $(\mathbb{R}, \mathcal{M}, m)$.

Measurable functions

We denote $\widehat{\mathbb{R}} = \mathbb{R} \cup \{\infty\} \cup \{-\infty\}$. To define measurable functions, we need

Theorem 2.15 *Let $f : E \longrightarrow \widehat{\mathbb{R}}$ be a function, where $E \in \mathcal{M}$. Then the following conditions are equivalent:*

(a) $\{x \in E \mid f(x) > \alpha\} \in \mathcal{M}$ *for all* $\alpha \in \mathbb{R}$,

(b) $\{x \in E \mid f(x) \geq \alpha\} \in \mathcal{M}$ *for all* $\alpha \in \mathbb{R}$,

(c) $\{x \in E \mid f(x) < \alpha\} \in \mathcal{M}$ *for all* $\alpha \in \mathbb{R}$,

(d) $\{x \in E \mid f(x) \leq \alpha\} \in \mathcal{M}$ *for all* $\alpha \in \mathbb{R}$.

Definition 2.16 A function $f : E \longrightarrow \widehat{\mathbb{R}}$ is measurable, if $E \in \mathcal{M}$ and if one of the conditions (a)–(d) in Theorem 2.15 holds.

The above definition may sound obscure, but if one of the conditions (a)–(d) in Theorem 2.15 holds, then all of them hold.

Example. Show that a constant function $f : \mathbb{R} \longrightarrow \{a\}$, $a \in \mathbb{R}$, is measurable.

Theorem 2.17 *Let $f : E \longrightarrow \widehat{\mathbb{R}}$ and $g : E \longrightarrow \widehat{\mathbb{R}}$ be measurable and let $c \in \mathbb{R}$. Then the functions*

$$c + f, \; cf, \; f + g \text{ and } fg$$

are measurable. Further, function $\frac{1}{f}$ is measurable, provided $f(x) \neq 0$ for all $x \in E$.

Definition 2.18 A given property is said to hold almost everywhere (melkein kaikkialla), briefly a.e. (m.k.), if the set of points where it fails to hold is a set of measure zero.

Example. If $f = g$ a.e. in $E \subset \mathbb{R}$, then $m(\{x \in E \mid f(x) \neq g(x)\}) = 0$.

Theorem 2.19 *If f is a measurable function in $E \subset \mathbb{R}$ and $f = g$ a.e. in E, then g is measurable in E.*

Proof. Define $A = \{x \in E \mid f(x) \neq g(x)\}$. By the assumption, $m(A) = 0$. Now

$$\begin{aligned} &\{x \in E \mid g(x) < \alpha\} \\ &\quad = (\{x \in E \mid f(x) < \alpha\} \cup \{x \in A \mid g(x) < \alpha\}) \setminus \{x \in A \mid g(x) \geq \alpha\}. \end{aligned}$$

$\{x \in E \mid f(x) < \alpha\} \in \mathcal{M}$ since f is measurable. $\{x \in A \mid g(x) < \alpha\} \in \mathcal{M}$ and $\{x \in A \mid g(x) \geq \alpha\} \in \mathcal{M}$ as subsets of a measurable set A of measure zero, see Theorem 2.9. Hence, by Theorems 2.8 and 2.12, $\{x \in E \mid g(x) < \alpha\} \in \mathcal{M}$ for all $\alpha \in \mathbb{R}$ and so g is measurable. □

We next give definitions for a *characteristic function* of a given set and for a *simple function.* These functions are also measurable.

Definition 2.20 The characteristic function of $E \subset \mathbb{R}$ is given by

$$\chi_E(x) = \begin{cases} 1, \ x \in E, \\ 0, \ x \in \mathbb{R} \setminus E. \end{cases}$$

Example. Show that the characteristic function of $E \in \mathcal{M}$ is measurable.

Definition 2.21 A real-valued function ϕ is called simple, if it is measurable and assumes only finitely many values.

If $\phi : \mathbb{R} \longrightarrow \mathbb{R}$ is simple and has values $a_1, \ldots, a_n \in \mathbb{R} \setminus \{0\}$, then we may define pairwise disjoint sets $E_j = \{x \in \mathbb{R} \mid f(x) = a_j\}$ for $j = 1, \ldots, n$. Further, the function ϕ can be represented in the form

$$\phi(x) = \sum_{j=1}^{n} a_j \chi_{E_j}(x). \tag{2.2}$$

A simple function restricted on a closed interval is called a *step function*:

Definition 2.22 A real-valued function φ on an interval $[a, b]$ is called a step function if there exists a partition $a = x_0 < x_1 < \cdots < x_n = b$ such that for each $j \in \{0, \ldots, n\}$ the function φ assumes only one value in the interval (x_j, x_{j+1}).

Defining the Lebesgue integral

We first define the Lebesgue integral for simple functions.

Definition 2.23 Let $a_1, \ldots, a_n \in \mathbb{R}$ and let $E_1, \ldots, E_n$ be pairwise disjoint subsets of $\mathbb{R}$. Let $\phi : \mathbb{R} \longrightarrow \mathbb{R}$,

$$\phi(x) = \sum_{j=1}^{n} a_j \chi_{E_j}(x),$$

be simple. The integral of ϕ is given by

$$\int \phi \, dm = \sum_{j=1}^{n} a_j m(E_j), \tag{2.3}$$

if the sum in (2.3) is defined. Further, ϕ is integrable, if

$$-\infty < \int \phi \, dm < +\infty,$$

that is, if the sum in (2.3) is finite.

Example. Find examples when the sum in (2.3) is not defined.

Example. Let $\phi : \mathbb{R} \longrightarrow \mathbb{R}$ be simple,

$$\phi(x) = \begin{cases} 4,5, & -56 < x < -32, \\ 4\pi, & -10 \le x < 1, \\ -12, & 1 \le x < 100, \\ 0, & \text{otherwise.} \end{cases}$$

Find the representation for ϕ as in (2.2) and calculate $\int \phi \, dm$, if it exists.

We next generalize the concept of an integral.

Definition 2.24 Let $f : \mathbb{R} \longrightarrow [0, \infty]$ be measurable. Then

$$\int f \, dm = \sup \left\{ \int \phi \, dm \mid \phi \text{ simple and } \phi \le f \right\}.$$

Definition 2.25 Let $f : \mathbb{R} \longrightarrow \widehat{\mathbb{R}}$ be measurable and let

$$f^+ = \max\{f, 0\} \quad \text{and} \quad f^- = \max\{-f, 0\}.$$

Then

$$\int f \, dm = \int f^+ \, dm - \int f^- \, dm, \tag{2.4}$$

if the right-hand side of (2.4) is defined. The integral of f over a measurable set E is

$$\int_E f \, dm = \int f \chi_E \, dm.$$

f is integrable, if

$$-\infty < \int f \, dm < +\infty.$$

Our next result shows some basic properties of the Lebesgue integral.

Theorem 2.26 *Let $f : \mathbb{R} \longrightarrow \widehat{\mathbb{R}}$ and $g : \mathbb{R} \longrightarrow \widehat{\mathbb{R}}$ be measurable and let $\alpha \in \mathbb{R}$ be a constant.*

(a) If $\int f \, dm$ exists, then

$$\int \alpha f \, dm = \alpha \int f \, dm.$$

(b) If $f \ge g$ and $\int g \, dm$ exists and $\int g \, dm > -\infty$, then $\int f \, dm$ exists and

$$\int f \, dm \ge \int g \, dm.$$

(b') If $f \ge g$ and $\int f \, dm$ exists and $\int f \, dm < \infty$, then $\int g \, dm$ exists and

$$\int f \, dm \ge \int g \, dm.$$

(c) If $\int f \, dm$ exists, then

$$\left| \int f \, dm \right| \le \int |f| \, dm.$$

(d) Let $f \geq 0$. Then

$$\int_E f\,dm = \sup\left\{\int_E \phi\,dm \mid \phi \text{ simple and } \phi \leq f\right\}.$$

(e) If $\int f\,dm$ exists, then $\int_E f\,dm$ exists for every $E \in \mathcal{M}$.

We next state a result dealing with a problem when the sum of given two measurable functions can be integrated. The proof is based on the Monotone Convergence Theorem below and therefore omitted.

Theorem 2.27 *Let f and g be measurable and suppose that $f+g$ is well-defined. If $\int f\,dm$ and $\int g\,dm$ exist and $\int f\,dm + \int g\,dm$ is well-defined (not of the forms $\infty - \infty$ or $-\infty + \infty$), then*

$$\int (f+g)\,dm = \int f\,dm + \int g\,dm.$$

In particular, if f and g are integrable, so is $f+g$.

Classical convergence theorems

We state three classical convergence theorems. Proofs can be found almost in any real analysis book containing Lebesgue integration.

Monotone Convergence Theorem: *Let $\{f_n\}$ be an increasing sequence of nonnegative measurable functions and let*

$$\lim_{n\to\infty} f_n = f.$$

Then

$$\int f\,dm = \lim_{n\to\infty} \int f_n\,dm.$$

For the next result we need to define limit supremum and limit infimum as follows: If $\{f_n\}$ is a sequence, then

$$\limsup_{n\to\infty} f_n = \inf_{n\in\mathbb{N}} \sup_{k\geq n} f_k$$

and

$$\liminf_{n\to\infty} f_n = \sup_{n\in\mathbb{N}} \inf_{k\geq n} f_k.$$

Note that limit supremum and limit infimum always exist, but limit $n \longrightarrow \infty$ does not have to.

Fatou Lemma: *Let $\{f_n\}$ be a sequence of measurable functions.*

(a) *If $f_n \geq f$ for all $n \in \mathbb{N}$, $\int f\,dm$ exists and if $\int f\,dm > -\infty$, then*

$$\liminf_{n\to\infty} \int f_n\,dm \geq \int \liminf_{n\to\infty} f_n\,dm.$$

(b) *If* $f_n \leq f$ *for all* $n \in \mathbb{N}$, $\int f\,dm$ *exists and if* $\int f\,dm < +\infty$, *then*

$$\limsup_{n\to\infty} \int f_n\,dm \leq \int \limsup_{n\to\infty} f_n\,dm.$$

Lebesgue's Dominated Covergence Theorem: *Let* $\{f_n\}$ *be a sequence of measurable functions and let* f *be a function such that*

$$\lim_{n\to\infty} f_n = f \ \ a.e.$$

Then, if there exists an integrable function g *such that* $|f_n| \leq g$ *for all* $n \in \mathbb{N}$, *we have*

$$\lim_{n\to\infty} \int f_n\,dm = \int f\,dm.$$

Riemann versus Lebesgue

Until so far we have seen somewhat strong results (such as the convergence theorems) concerning Lebesgue integrals that do not hold for Riemann integrals in general. The problem is that we cannot evaluate Lebesgue integrals for any other functions than simple ones.

To this end, let $f : [a, b] \longrightarrow \mathbb{R}$ be a Riemann integrable function. Divide $[a, b]$ into n-partition. The upper and lower Riemann sums are Lebesgue integrals of the corresponding step functions (which are, of course, simple). Making the partition more dense (i.e. letting $n \longrightarrow \infty$), f is Riemann integrable, provided that the upper and lower Riemann sums coincide. At the same time the Lebesgue integrals of the step functions corresponding to the two Riemann sums give the Lebesgue integral of f.

We therefore give the following loose statement:

Statement 1: *If* f *is Riemann integrable on a given interval (bounded, unbounded, whatever) then* f *is Lebesgue integrable and the two integrals coincide.*

When evaluating a given complicated Riemann integral, the idea is to consider the corresponding Lebesgue integral, apply the needed tools valid for Lebesgue integrals (such as the measure theory and the convergence theorems), and then to go back to the original Riemann integral. Therefore, in many cases, it does not matter what happens in a set of Lebesgue measure zero.

Statement 2: *Let* $f : [a, b] \longrightarrow \mathbb{R}$ *be bounded. Then* f *is Riemann integrable if and only if* f *is continuous a.e. in* $[a, b]$.

$\mathbb{R}^n$ as a measure space

As mentioned right after Definition 2.14, the triple $(\mathbb{R}, \mathcal{M}, m)$ is not the only measure space there exists. In this subsection we give an idea how to construct a measure and measurable sets in $\mathbb{R}^n$. Further measure spaces will be considered in Analysis 5.

Following the reasoning above, we first define an outer measure in $\mathbb{R}^n$, which satisfies properties analogous to (1), (2) and (4) on page 10 plus is subadditive. Then we define a measure as a restriction of the outer measure to a family of subsets of $\mathbb{R}^n$ such that the outer measure becomes countably additive. All sets in this family will be called measurable.

To begin with, denote $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$ and consider sets of the form

$$I = \{x \in \mathbb{R}^n \mid x_j \in I_j,\ j = 1, \ldots, n\},$$

where each $I_j \subset \mathbb{R}$ is an interval (open, semi-open, closed) with initial points (päätepisteinä) a_j and b_j, $a_j < b_j$. In cases $n = 1$, $n = 2$ and $n = 3$, I is an interval, a rectangle and an orthogonal parallelogram (suorakulmainen suuntaissärmiö), respectively. In what follows, we shall call I as an interval independently of the dimension n. I is called *open*, if each I_j is open. The number

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

is called the *geometric length* of I.

Definition 2.28 The n-dimensional Lebesgue outer measure of $A \subset \mathbb{R}^n$ is given by

$$m^*(A) = \inf \left\{ \sum_k l(I_k) \mid I_k \subset \mathbb{R}^n \text{ open intervals and } A \subset \bigcup_k I_k \right\},$$

where $l(I_k)$ is a geometric length of $I_k \subset \mathbb{R}^n$.

As in the dimension $n = 1$ (see Definition 2.1), the outer measure defined above is defined for all $A \subset \mathbb{R}^n$ and $0 \leq m^*(A) \leq \infty$. Furthermore, the outer measure of an interval $I \subset \mathbb{R}^n$ is its geometric length, m^* is subadditive and translation invariant under any translation of a given set. The details are omitted here, see *Lehto:* Differentiaali- ja Integraalilaskenta III.

Definition 2.29 A set $E \subset \mathbb{R}^n$ is said to be (Lebesgue) measurable if for each set $A \subset \mathbb{R}^n$ we have

$$m^*(A) = m^*(A \cap E) + m^*(A \setminus E).$$

The set of all (Lebesgue) measurable sets is denoted by $\mathcal{M}$. The restriction of an outer measure to $\mathcal{M}$ is called a (Lebesgue) measure and it is denoted by m.

The whole measure theory now extends from the first dimension to the general n^{th} dimension. The triple $(\mathbb{R}^n, \mathcal{M}, m)$ is again called a *measure space.*

The theory of Lebesgue integration of functions $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ can now be built on the measure space $(\mathbb{R}^n, \mathcal{M}, m)$. We first need to define measurable functions, characteristic function of a given set $E \subset \mathbb{R}^n$, and simple functions in $\mathbb{R}^n$. This theory is in parallel with the case $n = 1$.

As a final example we give an idea how to integrate functions $f : \mathbb{C} \longrightarrow \mathbb{C}$:

$$\int f\, dm = \int \Re f\, dm + i \int \Im f\, dm,$$

where m is the Lebesgue measure in $\mathbb{R}^2$ $(= \mathbb{C})$. Further, the functions $\Re f$ and $\Im f$ are both divided in two parts:

$$\Re f = (\Re f)^+ - (\Re f)^- \quad \text{and} \quad \Im f = (\Im f)^+ - (\Im f)^-.$$

Denoting the variable of f by $z = x + iy = re^{i\theta}$, it is customary to take $dm = dxdy$ or $dm = rdrd\theta$. For functions $f : \mathbb{C} \longrightarrow \mathbb{C}$, the notation $dm = dm(z)$ is often used.