# Chapter 5

# Measure and integration

In calculus you have learned how to calculate the size of different kinds of sets: the length of a curve, the area of a region or a surface, the volume or mass of a solid. In probability theory and statistics you have learned how to compute the size of other kinds of sets: the probability that certain events happen or do not happen.

In this chapter we shall develop a general theory for the size of sets, a theory that covers all the examples above and many more. Just as the concept of a metric space gave us a general setting for discussing the notion of distance, the concept of a measure space will provide us with a general setting for discussing the notion of size.

In calculus we use integration to calculate the size of sets. In this chapter we turn the situation around: We first develop a theory of size and then use it to define integrals of a new and more general kind. As we shall sometimes wish to compare the two theories, we shall refer to integration as taught in calculus as *Riemann-integration* in honor of the German mathematician Bernhard Riemann (1826-1866) and the new theory developed here as *Lebesgue integration* in honor of the French mathematician Henri Lebesgue (1875-1941).

Let us begin by taking a look at what we might wish for in a theory of size. Assume what we want to measure the size of subsets of a set $X$ (if you need something concrete to concentrate on, you may let $X = \mathbb{R}^2$ and think of the area of subsets of $\mathbb{R}^2$, or let $X = \mathbb{R}^3$ and think of the volume of subsets of $\mathbb{R}^3$). What properties do we want such a measure to have?

Well, if $\mu(A)$ denotes the size of a subset $A$ of $X$, we would expect

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

as nothing can be smaller than the empty set. In addition, it seems reasonable to expect:

(ii) If $A_1, A_2, A_3 \ldots$ is a disjoint sequence of sets, then

$$\mu(\bigcup_{n\in\mathbb{N}} A_n) = \sum_{n=1}^{\infty} \mu(A_n)$$

These two conditions are, in fact, all we need to develop a reasonable theory of size, except for one complication: It turns out that we can not in general expect to measure the size of *all* subsets of $X$ – some subsets are just so irregular that we can not assign a size to them in a meaningful way. This means that before we impose conditions (i) and (ii) above, we need to decide which properties the *measurable sets* (those we are able to assign a size to) should have. If we call the collection of all measurable sets $\mathcal{A}$, the statement $A \in \mathcal{A}$ is just a shorthand for "$A$ is measurable".

The first condition is simple; since we have already agreed that $\mu(\emptyset) = 0$, we must surely want to impose

(iii) $\emptyset \in \mathcal{A}$

For the next condition, assume that $A \in \mathcal{A}$. Intuitively, this means that we should be able to assign a size $\mu(A)$ to $A$. If the size $\mu(X)$ of the entire space is finite, we ought to have $\mu(A^c) = \mu(X) - \mu(A)$, and hence $A^c$ should be measurable. We shall impose this condition even when $X$ has infinite size:

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

For the third and last condition, assume that $\{A_n\}$ is a sequence of disjoint sets in $\mathcal{A}$. In view of condition (ii), it is natural to assume that $\bigcup_{n\in\mathbb{N}} A_n$ is in $\mathcal{A}$. We shall impose this condition even when the sequence is not disjoint (there are arguments for this that I don't want to get involved in at the moment):

(v) If $\{A_n\}_{n\in\mathbb{N}}$ is a sequence of sets in $\mathcal{A}$, then $\bigcup_{n\in\mathbb{N}} A_n \in \mathcal{A}$.

When we now begin to develop the theory systematically, we shall take the five conditions above as our starting point.

## 5.1 Measure spaces

Assume that $X$ is a nonempty set. A collection $\mathcal{A}$ of subsets of $X$ that satisfies conditions (iii)-(v) above, is called a $\sigma$-algebra. More succinctly:

**Definition 5.1.1** *Assume that $X$ is a nonempty set. A collection $\mathcal{A}$ of subsets of $X$ is called a $\sigma$-*algebra *if the following conditions are satisfied:*

*(i)* $\emptyset \in \mathcal{A}$

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

*(ii) 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 of subsets of* $X$*, we call the pair* $(X, \mathcal{A})$ *a* measurable space*.*

As already mentioned, the intuitive idea is that the sets in $\mathcal{A}$ are those that are so regular that we can measure their size.

Before we introduce measures, we take a look at some simple consequences of the definition above:

**Proposition 5.1.2** *Assume that* $\mathcal{A}$ *is a* $\sigma$*-algebra on* $X$*. Then*

*a)* $X \in \mathcal{A}$*.*

*b) If* $\{A_n\}_{n\in\mathbb{N}}$ *is a sequence of sets in* $\mathcal{A}$*, then* $\bigcap_{n\in\mathbb{N}} A_n \in \mathcal{A}$*.*

*c) If* $A_1, A_2, \ldots, A_n \in \mathcal{A}$*, then* $A_1 \cup A_2 \cup \ldots \cup A_n \in \mathcal{A}$ *and* $A_1 \cap A_2 \cap \ldots \cap A_n \in \mathcal{A}$*.*

*d) If* $A, B \in \mathcal{A}$*, then* $A \setminus B \in \mathcal{A}$*.*

*Proof:* a) By conditions (i) and (ii) in the definition, $X = \emptyset^c \in \mathcal{A}$.

b) By condition (ii), each $A_n^c$ is in $\mathcal{A}$, and hence $\bigcup_{n\in\mathbb{N}} A_n^c \in \mathcal{A}$ by condition (iii). By one of De Morgan's laws,

$$\Big(\bigcap_{n\in\mathbb{N}} A_n\Big)^c = \bigcup_{n\in\mathbb{N}} A_n^c$$

and hence $\big(\bigcap_{n\in\mathbb{N}} A_n\big)^c$ is in $\mathcal{A}$. Using condition (ii) again, we see that $\bigcap_{n\in\mathbb{N}} A_n$ is in $\mathcal{A}$.

c) If we extend the finite sequence $A_1, A_2, \ldots, A_n$ to an infinite one $A_1, A_2, \ldots, A_n, \emptyset, \emptyset, \ldots$, we see that

$$A_1 \cup A_2 \cup \ldots \cup A_n = \bigcup_{n\in\mathbb{N}} A_n \in \mathcal{A}$$

by condition (iii). A similar trick works for intersections, but we have to extend the sequence $A_1, A_2, \ldots, A_n$ to $A_1, A_2, \ldots, A_n, X, X, \ldots$ instead of $A_1, A_2, \ldots, A_n, \emptyset, \emptyset, \ldots$. The details are left to the reader.

d) We have $A \setminus B = A \cap B^c$, which is in $\mathcal{A}$ by condition (ii) and c) above. $\square$

It is time to turn to measures. Before we look at the definition, there is a small detail we have to take care of. As you know from calculus, there are sets of infinite size – curves of infinite length, surfaces of infinite area, solids of infinite volume. We shall use the symbol $\infty$ to indicate that sets have infinite size. This does not mean that we think of $\infty$ as a number; it is just a symbol to indicate that something has size bigger than can be specified by a number.

A measure $\mu$ assigns a value $\mu(A)$ ("the size of $A$") to each set $A$ in the $\sigma$-algebra $\mathcal{A}$. The value is either $\infty$ or a nonnegative number. If we let

$$\overline{\mathbb{R}}_+ = [0, \infty) \cup \{\infty\}$$

be the set of *extended, nonnegative real numbers*, $\mu$ is a function from $\mathcal{A}$ to $\overline{\mathbb{R}}_+$. In addition, $\mu$ has to satisfy conditions (i) and (ii) above, i.e.:

**Definition 5.1.3** *Assume that $(X, \mathcal{A})$ is a measurable space. A* measure *on $(X, \mathcal{A})$ is a function $\mu : \mathcal{A} \to \overline{\mathbb{R}}_+$ such that*

*(i)* $\mu(\emptyset) = 0$

*(ii) (Countable additivity) If $A_1, A_2, A_3 \ldots$ is a disjoint sequence of sets from $\mathcal{A}$, then*

$$\mu(\bigcup_{n=1}^{\infty} A_N) = \sum_{n=1}^{\infty} \mu(A_n)$$

*The triple $(X, \mathcal{A}, \mu)$ is then called a* measure space.

Let us take a look at some examples.

**Example 1:** Let $X = \{x_1, x_2, \ldots, x_n\}$ be a finite set, and let $\mathcal{A}$ be the collection of *all* subsets of $X$. For each set $A \subset X$, let

$$\mu(A) = |A| = \text{the number of elements in } A$$

Then $\mu$ is called the *counting measure* on $X$, and $(X, \mathcal{A}, \mu)$ is a measure space. ♣

The next two examples show two simple modifications of counting measures.

**Example 2:** Let $X$ and $\mathcal{A}$ be as in Example 1. For each element $x \in X$, let $m(x)$ be a nonnegative, real number (the *weight* of $x$). For $A \subset X$, let

$$\mu(A) = \sum_{x \in A} m_x$$

Then $(X, \mathcal{A}, \mu)$ is a measure space. ♣

**Example 3:** Let $X = \{x_1, x_2, \ldots, x_n, \ldots\}$ be a countable set, and let $\mathcal{A}$ be the collection of *all* subsets of $X$. For each set $A \subset X$, let

$$\mu(A) = \text{the number of elements in } A$$

where we put $\mu(A) = \infty$ if $A$ has infinitely many elements. Again $\mu$ is called the *counting measure* on $X$, and $(X, \mathcal{A}, \mu)$ is a measure space. ♣

The next example is also important, but rather special.

**Example 4:** Let $X$ be a any set, and let $\mathcal{A}$ be the collection of *all* subsets of $X$. Choose an element $a \in X$, and define

$$\mu(A) = \begin{cases} 1 & \text{if } a \in A \\ \\ 0 & \text{if } a \notin A \end{cases}$$

Then $(X, \mathcal{A}, \mu)$ is a measure space, and $\mu$ is called the *point measure at* $a$. ♣

The examples we have looked at so far are important special cases, but rather untypical of the theory – they are too simple to really need the full power of measure theory. The next examples are much more typical, but at this stage we can not define them precisely, only give an intuitive description of their most important properties.

**Example 5:** In this example $X = \mathbb{R}$, $\mathcal{A}$ is a $\sigma$-algebra containing all open and closed sets (we shall describe it more precisely later), and $\mu$ is a measure on $(X, \mathcal{A})$ such that

$$\mu([a, b]) = b - a$$

whenever $a \leq b$. This measure is called the *Lebesgue measure* on $\mathbb{R}$, and we can think of it as an extension of the notion of length to more general sets. The sets in $\mathcal{A}$ are those that can be assigned a generalized "length" $\mu(A)$ in a systematic way. ♣

Originally, measure theory was the theory of the Lebesgue measure, and it remains one of the most important examples. It is not at all obvious that such a measure exists, and one of our main tasks later in the next chapter will be to show that it does.

Lebesgue measure can be extended to higher dimensions:

**Example 6:** In this example $X = \mathbb{R}^2$, $\mathcal{A}$ is a $\sigma$-algebra containing all open and closed sets, and $\mu$ is a measure on $(X, \mathcal{A})$ such that

$$\mu([a, b] \times [c, d]) = (b - a)(d - c)$$

whenever $a \leq b$ and $c \leq d$ (this just means that the measure of a rectangle equals its area). This measure is called the *Lebesgue measure* on $\mathbb{R}^2$, and we can think of it as an extension of the notion of area to more general sets. The sets in $\mathcal{A}$ are those that can be assigned a generalized "area" $\mu(A)$ in a systematic way.

There are obvious extensions of this example to higher dimensions: The *three dimensional Lebesgue measure* assigns value

$$\mu([a,b] \times [c,d] \times [e,f]) = (b-a)(d-c)(f-e)$$

to all rectangular boxes and is a generalization of the notion of volume. The *$n$-dimensional Lebesgue measure* assigns value

$$\mu([a_1,b_1] \times [a_2,b_2] \times \cdots \times [a_n,b_n]) = (b_1-a_1)(b_2-a_2) \cdot \ldots \cdot (b_n-a_n)$$

to all $n$-dimensional, rectangular boxes and represents $n$-dimensional volume. ♣

Although we have not yet constructed the Lebesgue measures, we shall feel free to use them in examples and exercises. Let us finally take a look at two examples from probabilty theory.

**Example 7:** Assume we want to study coin tossing, and that we plan to toss the coin $N$ times. If we let H denote "heads" and T "tails", the possible outcomes can be represented as all sequences of H's and T's of length $N$. If the coin is fair, all such sequences have probability $\frac{1}{2^n}$.

To fit this into the framework of measure theory, let $X$ be the set of all sequences of H's and T's of length $N$, let $\mathcal{A}$ be the collection of all subsets of $X$, and let $\mu$ be given by

$$\mu(A) = \frac{|A|}{2^n}$$

where $|A|$ is the number of elements in $A$. Hence $\mu$ is the probability of the event $A$. It is easy to check that $\mu$ is a measure on $(X, \mathcal{A})$. ♣

In probability theory it is usual to call the underlying space $\Omega$ (instead of $X$) and the measure $P$ (instead of $\mu$), and we shall often refer to probability spaces as $(\Omega, \mathcal{A}, P)$.

**Example 8:** We are still studying coin tosses, but this time we don't know beforehand how many tosses we are going to make, and hence we have to consider all sequences of H's and T's of *infinite* length, that is all sequences

$$\omega = \omega_1, \omega_2, \omega_3, \ldots, \omega_n, \ldots$$

where each $\omega_i$ is either H or T. We let $\Omega$ be the collection of all such sequences.

To describe the $\sigma$-algebra and the measure, we first need to introduce the so-called *cylinder sets*: If $\mathbf{a} = a_1, a_2, \ldots, a_n$ is a *finite* sequence of H's and T's, we let

$$\mathcal{C}_{\mathbf{a}} = \{\omega \in \Omega \,|\, \omega_1 = a_1, \omega_2 = a_2, \ldots, \omega_n = a_n\}$$

and call it the *cylinder set* generated by $\mathbf{a}$. Note that $\mathcal{C}_{\mathbf{a}}$ consists of all sequences of coin tosses beginning with the sequence $a_1, a_2, \ldots, a_n$. Since the probability of starting a sequence of coin tosses with $a_1, a_2, \ldots, a_n$ is $\frac{1}{2^n}$, we want a measure such that $P(\mathcal{C}_{\mathbf{a}}) = \frac{1}{2^n}$.

The measure space $(\Omega, \mathcal{A}, P)$ of infinite coin tossing consists of $\Omega$, a $\sigma$-algebra $\mathcal{A}$ containing all cylinder sets, and a measure $P$ such that $P(\mathcal{C}_{\mathbf{a}}) = \frac{1}{2^n}$ for all cylinder sets of length $n$. It is not at all obvious that such a measure space exists, but it does (as we shall prove in the next chapter), and it is the right setting for the study of coin tossing of unrestricted length. ♣

Let us return to Definition 5.1.3 and derive some simple, but extremely useful consequences. Note that all these properties are properties we would expect of a measure.

**Proposition 5.1.4** *Assume that $(X, \mathcal{A}, \mu)$ is a measure space.*

*a) (Finite additivity) If $A_1, A_2, \ldots, A_m$ are disjoint sets in $\mathcal{A}$, then*

$$\mu(A_1 \cup A_2 \cup \ldots \cup A_m) = \mu(A_1) + \mu(A_2) + \ldots + \mu(A_m)$$

*b) (Monotonicity) If $A, B \in \mathcal{A}$ and $B \subset A$, then $\mu(B) \leq \mu(A)$.*

*c) If $A, B \in \mathcal{A}$, $B \subset A$, and $\mu(A) < \infty$, then $\mu(A \setminus B) = \mu(A) - \mu(B)$.*

*d) (Countable subadditivity) If $A_1, A_2, \ldots, A_n, \ldots$ is a (not necessarily disjoint) sequence of sets from $\mathcal{A}$, then*

$$\mu(\bigcup_{n \in \mathbb{N}} A_n) \leq \sum_{n=1}^{\infty} \mu(A_n)$$

*Proof:* a) We fill out the the sequence with empty sets to get an infinite sequence

$$A_1, A_2, \ldots, A_m, A_{m+1}, A_{m+2} \ldots$$

where $A_n = \emptyset$ for $n > m$. Then clearly

$$\mu(A_1 \cup A_2 \cup \ldots \cup A_m) = \mu(\bigcup_{n \in \mathbb{N}} A_n) = \sum_{n=1}^{\infty} \mu(A_n) = \mu(A_1) + \mu(A_2) + \ldots + \mu(A_n)$$

where we have used the two parts of definition 5.1.3.

b) We write $A = B \cup (A \setminus B)$. By Proposition 5.1.2d), $A \setminus B \in \mathcal{A}$, and hence by part a) above

$$\mu(A) = \mu(B) + \mu(A \setminus B) \geq \mu(B)$$

c) By the argument in part b),

$$\mu(A) = \mu(B) + \mu(A \setminus B)$$

Since $\mu(A)$ is finite, so is $\mu(B)$, and we may subtract $\mu(B)$ on both sides of the equation to get the result.

d) Define a new, *disjoint* sequence of sets $B_1, B_2, \ldots$ by:

$$B_1 = A_1, \quad B_2 = A_2 \backslash A_1, \quad B_3 = A_3 \backslash (A_1 \cup A_2), \quad B_4 = A_4 \backslash (A_1 \cup A_2 \cup A_3), \ldots$$

Note that $\bigcup_{n\in\mathbb{N}} B_n = \bigcup_{n\in\mathbb{N}} A_n$ (make a drawing). Hence

$$\mu(\bigcup_{n\in\mathbb{N}} A_n) = \mu(\bigcup_{n\in\mathbb{N}} B_n) = \sum_{n=1}^{\infty} \mu(B_n) \leq \sum_{n=1}^{\infty} \mu(A_n)$$

where we have applied part (ii) of Definition 5.1.3 to the disjoint sequence $\{B_n\}$ and in addition used that $\mu(B_n) \leq \mu(A_n)$ by part b) above. □

The next properties are a little more complicated, but not unexpected. They are often referred to as *continuity of measures*:

**Proposition 5.1.5** *Let $\{A_n\}_{n\in\mathbb{N}}$ be a sequence of measurable sets in a measure space $(X, \mathcal{A}, \mu)$.*

*a) If the sequence is increasing (i.e. $A_n \subset A_{n+1}$ for all $n$), then*

$$\mu(\bigcup_{n\in\mathbb{N}} A_n) = \lim_{n\to\infty} \mu(A_n)$$

*b) If the sequence is decreasing (i.e. $A_n \supset A_{n+1}$ for all $n$), and $\mu(A_1)$ is finite, then*

$$\mu(\bigcap_{n\in\mathbb{N}} A_n) = \lim_{n\to\infty} \mu(A_n)$$

*Proof:* a) If we put $E_1 = A_1$ and $E_n = A_n \setminus A_{n-1}$ for n> 1, the sequence $\{E_n\}$ is disjoint, and $\bigcup_{k=1}^{n} E_k = A_n$ for all $N$ (make a drawing). Hence

$$\mu(\bigcup_{n\in\mathbb{N}} A_n) = \mu(\bigcup_{n\in\mathbb{N}} E_n) = \sum_{n=1}^{\infty} \mu(E_n) =$$

$$= \lim_{n\to\infty} \sum_{k=1}^{n} \mu(E_k) = \lim_{n\to\infty} \mu(\bigcup_{k=1}^{n} E_k) = \lim_{n\to\infty} \mu(A_n)$$

where we have used the additivity of $\mu$ twice.

b) We first observe that $\{A_1 \setminus A_n\}_{n\in\mathbb{N}}$ is an increasing sequence of sets with union $A_1 \setminus \bigcap_{n\in\mathbb{N}} A_n$. By part a), we thus have

$$\mu(A_1 \setminus \bigcap_{n\in\mathbb{N}} A_n) = \lim_{n\to\infty} \mu(A_1 \setminus A_n)$$

Applying part c) of the previous proposition on both sides, we get

$$\mu(A_1) - \mu(\bigcap_{n\in\mathbb{N}} A_n) = \lim_{n\to\infty} (\mu(A_1) - \mu(A_n)) = \mu(A_1) - \lim_{n\to\infty} \mu(A_n)$$

Since $\mu(A_1)$ is finite, we get $\mu(\bigcap_{n\in\mathbb{N}} A_n) = \lim_{n\to\infty} \mu(A_n)$, as we set out to prove. □

**Remark:** The finiteness condition in part (ii) may look like an unnecessary consequence of a clumsy proof, but it is actually needed as the following example shows: Let $X = \mathbb{N}$, let $\mathcal{A}$ be the set of all subsets of $A$, and let $\mu(A) = |A|$ (the number of elements in $A$). If $A_n = \{n, n+1, \ldots\}$, then $\mu(A_n) = \infty$ for all $n$, but $\mu(\bigcap_{n\in\mathbb{N}} A_n) = \mu(\emptyset) = 0$. Hence $\lim_{n\to\infty} \mu(A_n) \neq \mu(\bigcap_{n\in\mathbb{N}} A_n)$.

The properties we have proved in this section are the basic tools we need to handle measures. The next section will take care of a more technical issue.

## Exercises for Section 5.1

1. Verify that the space $(X, \mathcal{A}, \mu)$ in Example 1 is a measure space.
2. Verify that the space $(X, \mathcal{A}, \mu)$ in Example 2 is a measure space.
3. Verify that the space $(X, \mathcal{A}, \mu)$ in Example 3 is a measure space.
4. Verify that the space $(X, \mathcal{A}, \mu)$ in Example 4 is a measure space.
5. Verify that the space $(X, \mathcal{A}, \mu)$ in Example 7 is a measure space.
6. Describe a measure space that is suitable for modeling tossing a die $N$ times.
7. Show that if $\mu$ and $\nu$ are two measures on the same measurable space $(X, \mathcal{A})$, then for all positive numbers $\alpha, \beta \in \mathbb{R}$, the function $\lambda : \mathcal{A} \to \mathbb{R}_+$ given by

$$\lambda(A) = \alpha\mu(A) + \beta\nu(A)$$

is a measure.

8. Assume that $(X, \mathcal{A}, \mu)$ is a measure space and that $A \in \mathcal{A}$. Define $\mu_A : \mathcal{A} \to \overline{\mathbb{R}}_+$ by
$$\mu_A(B) = \mu(A \cap B) \qquad \text{for all } B \in \mathcal{A}$$
Show that $\mu_A$ is a measure.

9. Let $X$ be an uncountable set, and define
$$\mathcal{A} = \{A \subset X \mid A \text{ or } A^c \text{ is countable}\}$$
Show that $\mathcal{A}$ is a $\sigma$-algebra. Define $\mu : \mathcal{A} \to \mathbb{R}_+$ by
$$\mu(A) = \begin{cases} 0 & \text{if } A \text{ is countable} \\ 1 & \text{if } A^c \text{ is countable} \end{cases}$$
Show that $\mu$ is a measure.

10. Assume that $(X, \mathcal{A})$ is a measurable space, and let $f : X \to Y$ be any function from $X$ to a set $Y$. Show that
$$\mathcal{B} = \{B \subset Y \mid f^{-1}(B) \in \mathcal{A}\}$$
is a $\sigma$-algebra.

11. Assume that $(X, \mathcal{A})$ is a measurable space, and let $f : Y \to X$ be any function from a set $Y$ to $X$. Show that
$$\mathcal{B} = \{f^{-1}(A) \mid A \in \mathcal{A}\}$$
is a $\sigma$-algebra.

12. Let $X$ be a set and $\mathcal{A}$ a collection of subsets of $X$ such that:
    a) $\emptyset \in \mathcal{A}$
    b) If $A \in \mathcal{A}$, then $A^c \in \mathcal{A}$
    c) If $\{A_n\}_{n\in\mathbb{N}}$ is a sequence of sets from $\mathcal{A}$, then $\bigcap_{n\in\mathbb{N}} A_n \in \mathcal{A}$.

    Show that $\mathcal{A}$ is a $\sigma$-algebra.

13. A measure space $(X, \mathcal{A}, \mu)$ is called *atomless* if $\mu(\{x\}) = 0$ for all $x \in X$. Show that in an atomless space, all countable sets have measure 0.

14. Assume that $\mu$ is a measure on $\mathbb{R}$ such that $\mu([-\frac{1}{n}, \frac{1}{n}]) = 1 + \frac{2}{n}$ for each $n \in \mathbb{N}$. Show that $\mu(\{0\}) = 1$.

15. Assume that a measure space $(X, \mathcal{A}, \mu)$ contains set of arbitrarily large finite measure, i.e. for each $N \in \mathbb{N}$, there is a set $A \in \mathcal{A}$ such that $N \leq \mu(A) < \infty$. Show that there is a set $B \in \mathcal{A}$ such that $\mu(B) = \infty$.

16. Assume that $\mu$ is a measure on $\mathbb{R}$ such that $\mu([a, b]) = b - a$ for all closed intervals $[a, b]$, $a < b$. Show that $\mu((a, b)) = b - a$ for all open intervals. Conversely, show that if $\mu$ is a measure on $\mathbb{R}$ such that $\mu((a, b)) = b - a$ for all open intervals $[a, b]$, then $\mu([a, b]) = b - a$ for all closed intervals.

17. Let $X$ be a set. An *algebra* is a collection $\mathcal{A}$ of subset of $X$ such that

(i) $\emptyset \in \mathcal{A}$

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

(iii) If $A, B \in \mathcal{A}$, then $A \cup B \in \mathcal{A}$.

Show that if $\mathcal{A}$ is an algebra, then:

a) If $A_1, A_2, \ldots, A_n \in \mathcal{A}$, then $A_1 \cup A_2 \cup \ldots \cup A_n \in \mathcal{A}$ (use induction on $n$).

b) If $A_1, A_2, \ldots, A_n \in \mathcal{A}$, then $A_1 \cap A_2 \cap \ldots \cap A_n \in \mathcal{A}$.

c) Put $X = \mathbb{N}$ and define $\mathcal{A}$ by

$$\mathcal{A} = \{A \subset \mathbb{N} \,|\, A \text{ or } A^c \text{ is finite}\}$$

Show that $\mathcal{A}$ is an algebra, but not a $\sigma$-algebra.

d) Assume that $\mathcal{A}$ is an algebra closed under disjoint, countable unions (i.e., $\bigcup_{n\in\mathbb{N}} A_n \in \mathcal{A}$ for all *disjoint* sequences $\{A_n\}$ of sets from $\mathcal{A}$). Show that $\mathcal{A}$ is a $\sigma$-algebra.

18. Let $(X, \mathcal{A}, \mu)$ be a measure space and assume that $\{A_n\}$ is a sequence of sets from $\mathcal{A}$ such that $\sum_{n=1}^{\infty} \mu(A_n) < \infty$. Let

$$A = \{x \in X \,|\, x \text{ belongs to infinitely many of the sets } A_n\}$$

Show that $A \in \mathcal{A}$ and that $\mu(A) = 0$.

## 5.2 Complete measures

Assume that $(X, \mathcal{A}, \mu)$ is a measure space, and that $A \in \mathcal{A}$ with $\mu(A) = 0$. It is natural to think that if $N \subset A$, then $N$ must also be measurable and have measure 0, but there is nothing in the definition of a measure that says so, and, in fact, it is not difficult to find measure spaces where this property does not hold. This is often a nuisance, and we shall now see how it can be cured.

First some definitions:

**Definition 5.2.1** *Assume that $(X, \mathcal{A}, \mu)$ is a measure space. A set $N \subset X$ is called a* null set *if $N \subset A$ for some $A \in \mathcal{A}$ with $\mu(A) = 0$. The collection of all null sets is denoted by $\mathcal{N}$. If all null sets belong to $\mathcal{A}$, we say that the measure space is* complete.

Note that if $N$ is a null set that happens to belong to $\mathcal{A}$, then $\mu(N) = 0$ by Proposition 5.1.4b).

Our purpose in this section is to show that any measure space $(X, \mathcal{A}, \mu)$ can be extended to a complete space (i.e. we can find a complete measure space $(X, \bar{\mathcal{A}}, \bar{\mu})$ such that $\mathcal{A} \subset \bar{\mathcal{A}}$ and $\bar{\mu}(A) = \mu(A)$ for all $A \in \mathcal{A}$).

We begin with a simple observation:

**Lemma 5.2.2** *If $N_1, N_2, \ldots$ are null sets, then $\bigcup_{n\in\mathbb{N}} N_n$ is a null set.*

*Proof:* For each $n$, there is a set $A_n \in \mathcal{A}$ such that $\mu(A_n) = 0$ and $N_n \subset A_n$. Since $\bigcup_{n\in\mathbb{N}} N_n \subset \bigcup_{n\in\mathbb{N}} A_n$ and

$$\mu(\bigcup_{n\in\mathbb{N}} A_n) \leq \sum_{n=1}^{\infty} \mu(A_n) = 0$$

by Proposition 5.1.4d), $\bigcup_{n\in\mathbb{N}} N_n$ is a null set. □

The next lemma tells us how we can extend a $\sigma$-algebra to include the null sets.

**Lemma 5.2.3** *If $(X, \mathcal{A}, \mu)$ is a measure space, then*

$$\bar{\mathcal{A}} = \{A \cup N \mid A \in \mathcal{A} \text{ and } N \in \mathcal{N}\}$$

*is the smallest $\sigma$-algebra containing $\mathcal{A}$ and $\mathcal{N}$ (in the sense that if $\mathcal{B}$ is any other $\sigma$-algebra containing $\mathcal{A}$ and $\mathcal{N}$, then $\mathcal{A} \subset \mathcal{B}$).*

*Proof:* If we can only prove that $\bar{\mathcal{A}}$ is a $\sigma$-algebra, the rest will be easy: Any $\sigma$-algebra $\mathcal{B}$ containing $\mathcal{A}$ and $\mathcal{N}$, must necessarily contain all sets of the form $A \cup N$ and hence be larger than $\bar{\mathcal{A}}$, and since $\emptyset$ belongs to both $\mathcal{A}$ and $\mathcal{N}$, we have $\mathcal{A} \subset \bar{\mathcal{A}}$ and $\mathcal{N} \subset \bar{\mathcal{A}}$.

To prove that $\bar{\mathcal{A}}$ is a $\sigma$-algebra, we need to check the three conditions in Definition 5.1.1. Since $\emptyset$ belongs to both $\mathcal{A}$ and $\mathcal{N}$, condition (i) is obviously satisfied, and condition (iii) follows from the identity

$$\bigcup_{n\in\mathbb{N}} (A_n \cup N_n) = \bigcup_{n\in\mathbb{N}} A_n \cup \bigcup_{n\in\mathbb{N}} N_n$$

and the preceeding lemma.

It remains to prove condition (ii), and this is the tricky part. Given a set $A \cup N \in \bar{\mathcal{A}}$, we must prove that $(A \cup N)^c \in \bar{\mathcal{A}}$. Observe first that we can assume that $A$ and $N$ are disjoint; if not, we just replace $N$ by $N \setminus A$. Next observe that since $N$ is a null set, there is a set $B \in \mathcal{A}$ such that $N \subset B$ and $\mu(B) = 0$. We may also assume that $A$ and $B$ are disjoint; if not, we just replace $B$ by $B \setminus A$. Since

$$(A \cup N)^c = (A \cup B)^c \cup (B \setminus N)$$

(see Figure 1), where $(A \cup B)^c \in \mathcal{A}$ and $B \setminus N \in \mathcal{N}$, we see that $(A \cup N)^c \in \bar{\mathcal{A}}$ and the lemma is proved. □

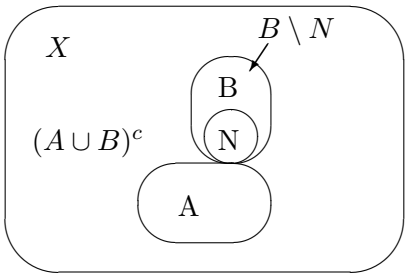


Figure 1: $(A \cup N)^c = (A \cup B)^c \cup (B \setminus N)$

The next step is to extend $\mu$ to a measure on $\bar{\mathcal{A}}$. Here is the key observation:

**Lemma 5.2.4** *If $A_1, A_2 \in \mathcal{A}$ and $N_1, N_2 \in \mathcal{N}$ are such that $A_1 \cup N_1 = A_2 \cup N_2$, then $\mu(A_1) = \mu(A_2)$.*

*Proof:* Let $B_2$ be a set in $\mathcal{A}$ such that $N_2 \subset B_2$ and $\mu(B_2) = 0$. Then $A_1 \subset A_1 \cup N_1 = A_2 \cup N_2 \subset A_2 \cup B_2$, and hence

$$\mu(A_1) \leq \mu(A_1 \cup B_2) \leq \mu(A_2) + \mu(B_2) = \mu(A_2)$$

Interchanging the rôles of $A_1$ and $A_2$, we get the opposite inequality $\mu(A_2) \leq \mu(A_1)$, and hence we must have $\mu(A_1) = \mu(A_2)$. □

We are now ready for the main result. It shows that we can always extend a measure space to a complete measure space in a controlled manner. The measure space $(X, \bar{\mathcal{A}}, \bar{\mu})$ in the theorem below is called the *completion* of the original measure space $(X, \mathcal{A}, \mu)$.

**Theorem 5.2.5** *Assume that $(X, \mathcal{A}, \mu)$ is a measure space, let*

$$\bar{\mathcal{A}} = \{A \cup N \mid A \in \mathcal{A} \text{ and } N \in \mathcal{N}\}$$

*and define $\bar{\mu} : \bar{\mathcal{A}} \to \overline{\mathbb{R}}_+$ by*

$$\bar{\mu}(A \cup N) = \mu(A)$$

*for all $A \in \mathcal{A}$ and all $N \in \mathcal{N}$. Then $(X, \bar{\mathcal{A}}, \bar{\mu})$ is a complete measure space, and $\bar{\mu}$ is an extension of $\mu$, i.e. $\bar{\mu}(A) = \mu(A)$ for all $A \in \mathcal{A}$.*

*Proof:* We already know that $\bar{\mathcal{A}}$ is a $\sigma$-algebra, and by the lemma above, the definition

$$\bar{\mu}(A \cup N) = \mu(A)$$

is legitimate (i.e. it only depends on the set $A \cup N$ and not on the sets $A \in \mathcal{A}$, $N \in \mathcal{N}$ we use to represent it). Also, we clearly have $\bar{\mu}(A) = \mu(A)$ for all $A \in \mathcal{A}$.

To prove that $\bar{\mu}$ is a measure, observe that since obviously $\bar{\mu}(\emptyset) = 0$, we just need to check that if $\{B_n\}$ is a disjoint sequence of sets in $\bar{\mathcal{A}}$, then

$$\bar{\mu}(\bigcup_{n\in\mathbb{N}} B_n) = \sum_{n=1}^{\infty} \bar{\mu}(B_n)$$

For each $n$, pick sets $A_n \in \mathcal{A}$, $N_n \in \mathcal{N}$ such that $B_n = A_n \cup N_n$. Then the $A_n$'s are clearly disjoint since the $B_n$'s are, and since $\bigcup_{n\in\mathbb{N}} B_n = \bigcup_{n\in\mathbb{N}} A_n \cup \bigcup_{n\in\mathbb{N}} N_n$, we get

$$\bar{\mu}(\bigcup_{n\in\mathbb{N}} B_n) = \mu(\bigcup_{n\in\mathbb{N}} A_n) = \sum_{n=1}^{\infty} \mu(A_n) = \sum_{n=1}^{\infty} \bar{\mu}(B_n)$$

It remains to check that $\bar{\mu}$ is complete. Assume that $C \subset D$, where $\bar{\mu}(D) = 0$; we must show that $C \in \bar{\mathcal{A}}$. Since $\bar{\mu}(D) = 0$, $D$ is of the form $D = A \cup N$, where $A$ is in $\mathcal{A}$ with $\mu(A) = 0$, and $N$ is in $\mathcal{N}$. By definition of $\mathcal{N}$, there is a $B \in \mathcal{A}$ such that $N \subset B$ and $\mu(B) = 0$. But then $C \subset A \cup B$, where $\mu(A \cup B) = 0$, and hence $C$ is in $\mathcal{N}$ and hence in $\bar{\mathcal{A}}$. □

In Lemma 5.2.3 we proved that $\bar{\mathcal{A}}$ is the smallest $\sigma$-algebra containing $\mathcal{A}$ and $\mathcal{N}$. This an instance of a more general phenomenon: Given a collection $\mathcal{B}$ of subsets of $X$, there is always a smallest $\sigma$-algebra $\mathcal{A}$ containing $\mathcal{B}$. It is called the *$\sigma$-algebra generated by $\mathcal{B}$* and is often designated by $\sigma(\mathcal{B})$. The proof that $\sigma(\mathcal{B})$ exists is not difficult, but quite abstract:

**Proposition 5.2.6** *Let $X$ be a non-empty set and $\mathcal{B}$ a collection of subsets of $X$. Then there exists a smallest $\sigma$-algebra $\sigma(\mathcal{B})$ containing $\mathcal{B}$ (in the sense that if $\mathcal{C}$ is any other $\sigma$-algebra containing $\mathcal{B}$, then $\sigma(\mathcal{B}) \subset \mathcal{C}$).*

*Proof:* Observe that there is at least one $\sigma$-algebra containing $\mathcal{B}$, namely the $\sigma$-algebra of *all* subsets of $X$. This guarantees that the following definition makes sense:

$$\sigma(\mathcal{B}) = \{A \subset X \mid A \text{ belongs to } \textit{all } \sigma\text{-algebras containing } \mathcal{B}\}$$

It suffices to show that $\sigma(\mathcal{B})$ is a $\sigma$-algebra as it then clearly must be the smallest $\sigma$-algebra containing $\mathcal{B}$.

We must check the three conditions in Definition 5.1.1. For (i), just observe that since $\emptyset$ belongs to all $\sigma$-algebras, it belongs to $\sigma(\mathcal{B})$. For (ii), observe that if $A \in \sigma(\mathcal{B})$, then $A$ belongs to all $\sigma$-algebras containing $\mathcal{B}$. Since $\sigma$-algebras are closed under complements, $A^c$ belongs to the same $\sigma$-algebras, and hence to $\sigma(\mathcal{B})$. The argument for (iii) is similar: Assume that

the sets $A_n$, $n \in \mathbb{N}$, belong to $\sigma(\mathcal{B})$. Then they belong to all $\sigma$-algebras containing $\mathcal{B}$, and since $\sigma$-algebras are closed under countable unions, the union $\bigcup_{n\in\mathbb{N}} A_n$ belongs to the same $\sigma$-algebras and hence to $\sigma(\mathcal{B})$. $\square$

In many applications, the underlying set $X$ is also a metric space (e.g., $X = \mathbb{R}^d$ for the Lebesgue measure). In this case the $\sigma$-algebra $\sigma(\mathcal{G})$ generated by the collection $\mathcal{G}$ of open sets is called the *Borel $\sigma$-algebra*, a measure defined on $\sigma(\mathcal{G})$ is called a *Borel measure*, and the sets in $\sigma(\mathcal{G})$ are called *Borel sets*. Most useful measures on metric spaces are either Borel measures or completions of Borel measures.

We can now use the results and terminology of this section to give a more detailed descriptions of the Lebesgue measure on $\mathbb{R}^d$. It turns out (as we shall prove in the next chapter) that there is a unique measure on the Borel $\sigma$-algebra $\sigma(\mathcal{G})$ such that

$$\mu([a_1, b_1] \times [a_2, b_2] \times \cdots \times [a_d, b_d]) = (b_1 - a_1)(b_2 - a_2) \cdot \ldots \cdot (b_d - a_d)$$

whenever $a_1 < b_1$, $a_2 < b_2$,..., $a_d < b_d$ (i.e. $\mu$ assigns the "right" value to all rectangular boxes). The completion of this measure is the Lebesgue measure on $\mathbb{R}^d$.

We can give a similar description of the space of all infinite series of coin tosses in Example 8 of section 5.1. In this setting one can prove that there is a unique measure on the $\sigma$-algebra $\sigma(\mathcal{C})$ generated by the cylinder sets, and the completion of this measure is the one used to model coin tossing.

## Exercises to Section 5.2

1. Let $X = \{0, 1, 2\}$ and let $\mathcal{A} = \{\emptyset, \{0, 1\}, \{2\}, X\}$.
   a) Show that $\mathcal{A}$ is a $\sigma$-algebra.
   b) Define $\mu : \mathcal{A} \to \mathbb{R}_+$ by: $\mu(\emptyset) = \mu(\{0, 1\}) = 0$, $\mu(\{2\}) = \mu(X) = 1$. Show that $\mu$ is a measure.
   c) Show that $\mu$ is *not* complete, and describe the completion $(X, \bar{\mathcal{A}}, \bar{\mu})$ of $(X, \mathcal{A}, \mu)$.
2. Redo Problem 1 for $X = \{0, 1, 2, 3\}$ and $\mathcal{A} = \{\emptyset, \{0, 1\}, \{2, 3\}, X\}$.
3. Let $(X, \mathcal{A}, \mu)$ be a complete measure space. Assume that $A, B \in \mathcal{A}$ with $\mu(A) = \mu(B) < \infty$. Show that if $A \subset C \subset B$, then $C \in \mathcal{A}$.
4. Let $\mathcal{A}$ and $\mathcal{B}$ be two collections of subsets of $X$. Assume that any set in $\mathcal{A}$ belongs to $\sigma(\mathcal{B})$ and that any set in $\mathcal{B}$ belongs to $\sigma(\mathcal{A})$. Show that $\sigma(\mathcal{A}) = \sigma(\mathcal{B})$.
5. Assume that $X$ is a metric space, and let $\mathcal{G}$ be the collection of all open sets and $\mathcal{F}$ the collection of all closed sets. Show that $\sigma(\mathcal{G}) = \sigma(\mathcal{F})$.

6. Let $X$ be a set. An *algebra* is a collection $\mathcal{A}$ of subset of $X$ such that

   (i) $\emptyset \in \mathcal{A}$

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

   (iii) If $A, B \in \mathcal{A}$, then $A \cup B \in \mathcal{A}$.

   Show that if $\mathcal{B}$ is a collection of subsets of $X$, there is a smallest algebra $\mathcal{A}$ containing $\mathcal{B}$.

7. Let $X$ be a set. A *monotone class* is a collection $\mathcal{M}$ of subset of $X$ such that

   (i) If $\{A_n\}$ is an increasing sequence of sets from $\mathcal{M}$, then $\bigcup_{n\in\mathbb{N}} A_n \in \mathcal{M}$.

   (ii) If $\{A_n\}$ is a decreasing sequence of sets from $\mathcal{M}$, then $\bigcap_{n\in\mathbb{N}} A_n \in \mathcal{M}$.

   Show that if $\mathcal{B}$ is a collection of subsets of $X$, there is a smallest monotone class $\mathcal{M}$ containing $\mathcal{B}$.

## 5.3 Measurable functions

One of the main purposes of measure theory is to provide a richer and more flexible foundation for integration theory, but before we turn to integration, we need to look at the functions we hope to integrate, the *measurable* functions. As functions taking the values $\infty$ and $-\infty$ will occur naturally as limits of sequences of ordinary functions, we choose to include them from the beginning; hence we shall study functions

$$f : X \to \overline{\mathbb{R}}$$

where $(X, \mathcal{A}, \mu)$ is a measure space and $\overline{\mathbb{R}} = \mathbb{R} \cup \{-\infty, \infty\}$ is the set of *extended real numbers*. Don't spend too much effort on trying to figure out what $-\infty$ and $\infty$ "really" are — they are just convenient symbols for describing divergence.

To some extent we may extend ordinary algebra to $\overline{\mathbb{R}}$, e.g., we shall let

$$\infty + \infty = \infty, \quad -\infty - \infty = -\infty$$

and

$$\infty \cdot \infty = \infty, \quad (-\infty) \cdot \infty = -\infty, \quad (-\infty) \cdot (-\infty) = \infty.$$

If $r \in \mathbb{R}$, we similarly let

$$\infty + r = \infty, \quad -\infty + r = -\infty$$

For products, we have to take the sign of $r$ into account, hence

$$\infty \cdot r = \begin{cases} \infty & \text{if } r > 0 \\ -\infty & \text{if } r < 0 \end{cases}$$

and similarly for $(-\infty) \cdot r$.

All the rules above are natural and intuitive. Expressions that do not have an intuitive interpretation, are usually left undefined, e.g. is $\infty - \infty$ *not* defined. There is one exception to this rule; it turns out that in measure theory (but not in other parts of mathematics!) it is convenient to define $0 \cdot \infty = \infty \cdot 0 = 0$.

Since algebraic expressions with extended real numbers are not always defined, we need to be careful and always check that our expressions make sense.

We are now ready to define measurable functions:

**Definition 5.3.1** *Let $(X, \mathcal{A}, \mu)$ be a measure space. A function $f : X \to \overline{\mathbb{R}}$ is* measurable (with respect to $\mathcal{A}$) *if*

$$f^{-1}([-\infty, r)) \in \mathcal{A}$$

*for all $r \in \mathbb{R}$. In other words, the set*

$$\{x \in X \,:\, f(x) < r\}$$

*must be measurable for all $r \in \mathbb{R}$.*

The half-open intervals in the definition are just a convenient starting point for showing that the inverse images of open and closed sets are measurable, but to prove this, we need a little lemma:

**Lemma 5.3.2** *Any non-empty, open set $G$ in $\mathbb{R}$ is a countable union of open intervals.*

*Proof:* Call an open interval $(a, b)$ *rational* if the endpoints $a, b$ are rational numbers. As there are only countably many rational numbers, there are only countably many rational intervals. It is not hard to check that $G$ is the union of those rational intervals that are contained in $G$. □

**Proposition 5.3.3** *If $f : X \to \overline{\mathbb{R}}$ is measurable, then $f^{-1}(I) \in \mathcal{A}$ for all intervals $I = (s, r), I = (s, r], I = [s, r), I = [s, r]$ where $s, r \in \overline{\mathbb{R}}$. Indeed, $f^{-1}(A) \in \mathcal{A}$ for all open and closed sets $A$.*

*Proof:* We use that inverse images commute with intersections, unions and complements. First observe that for any $r \in \mathbb{R}$

$$f^{-1}\big([-\infty, r]\big) = f^{-1}\big(\bigcap_{n \in \mathbb{N}} [-\infty, r + \frac{1}{n})\big) = \bigcap_{n \in \mathbb{N}} f^{-1}\big([-\infty, r + \frac{1}{n})\big) \in \mathcal{A}$$

which shows that the inverse images of closed intervals $[-\infty, r]$ are measurable. Taking complements, we see that the inverse images of intervals of the form $[s, \infty]$ and $(s, \infty]$ are measurable:

$$f^{-1}([s,\infty]) = f^{-1}([-\infty, s)^c) = \big(f^{-1}([-\infty, s))\big)^c \in \mathcal{A}$$

and

$$f^{-1}((s,\infty]) = f^{-1}([-\infty, s]^c) = \big(f^{-1}([-\infty, s])\big)^c \in \mathcal{A}$$

To show that the inverse images of finite intervals are measurable, we just take intersections, e.g.,

$$f^{-1}((s,r)) = f^{-1}([-\infty, r) \cap (s, \infty]) = f^{-1}([-\infty, r)) \cap f^{-1}((s, \infty]) \in \mathcal{A}$$

If $A$ is open, we know from the lemma above that it is a countable union $A = \bigcup_{n\in\mathbb{N}} I_n$ of open intervals. Hence

$$f^{-1}(A) = f^{-1}\big(\bigcup_{n\in\mathbb{N}} I_n\big) = \bigcup_{n\in\mathbb{N}} f^{-1}(I_n) \in \mathcal{A}$$

Finally, if $A$ is closed, we use that its complement is open to get

$$f^{-1}(A) = \big(f^{-1}(A^c)\big)^c \in \mathcal{A}$$

□

It is sometimes convenient to use other kinds of intervals than those in the definition to check that a function is measurable:

**Proposition 5.3.4** *Let $(X, \mathcal{A}, \mu)$ be a measure space and consider a function $f : X \to \overline{\mathbb{R}}$. If either*

*(i) $f^{-1}([-\infty, r]) \in \mathcal{A}$ for all $r \in \mathbb{R}$, or*

*(ii) $f^{-1}([r, \infty]) \in \mathcal{A}$ for all $r \in \mathbb{R}$, or*

*(iii) $f^{-1}((r, \infty]) \in \mathcal{A}$ for all $r \in \mathbb{R}$,*

*then $f$ is measurable.*

*Proof:* In either case we just have to check that $f^{-1}([-\infty, r)) \in \mathcal{A}$ for all $r \in \mathbb{R}$. This can be done by the techniques in the previous proof. The details are left to the reader. □

The next result tells us that there are many measurable functions. Recall the definition of Borel measures and completed Borel measures from the end of Section 5.2.

**Proposition 5.3.5** *Let $(X, d)$ be a metric space and let $\mu$ be a Borel or a completed Borel measure on $X$. Then all continuous functions $f : X \to \mathbb{R}$ are measurable.*

*Proof:* Since $f$ is continuous and takes values in $\mathbb{R}$,

$$f^{-1}([-\infty, r)) = f^{-1}((-\infty, r))$$

is an open set by Proposition 2.3.9 and measurable since the Borel $\sigma$-algebra is generated by the open sets. □

We shall now prove a series of results showing how we can obtain new measurable functions from old ones. These results are not very exciting, but they are necessary for the rest of the theory. Note that the functions in the next two propositions take values in $\mathbb{R}$ and not $\overline{\mathbb{R}}$.

**Proposition 5.3.6** *Let $(X, \mathcal{A}, \mu)$ be a measure space. If $f : X \to \mathbb{R}$ is measurable, then $\phi \circ f$ is measurable for all continuous functions $\phi : \mathbb{R} \to \mathbb{R}$. In particular, $f^2$ is measurable.*

*Proof:* We have to check that

$$(\phi \circ f)^{-1}((-\infty, r)) = f^{-1}(\phi^{-1}((-\infty, r)))$$

is measurable. Since $\phi$ is contiunuous, $\phi^{-1}((-\infty, r))$ is open, and consequently $f^{-1}(\phi^{-1}((-\infty, r)))$ is measurable by Proposition 5.3.3. To see that $f^2$ is measurable, apply the first part of the theorem to the function $\phi(x) = x^2$. □

**Proposition 5.3.7** *Let $(X, \mathcal{A}, \mu)$ be a measure space. If the functions $f, g : X \to \mathbb{R}$ are measurable, so are $f + g$, $f - g$, and $fg$.*

*Proof:* To prove that $f + g$ is measurable, observe first that $f + g < r$ means that $f < r - g$. Since the rational numbers are dense, it follows that there is a rational number $q$ such that $f < q < r - g$. Hence

$$(f + g)^{-1}([-\infty, r)) = \{x \in X \mid (f + g) < r) =$$

$$\bigcup_{q \in \mathbb{Q}} (\{x \in X \mid f(x) < q\} \cap \{x \in X \mid g < r - q\})$$

which is measurable since $\mathbb{Q}$ is countable and a countabe union of measurable sets is measurable. A similar argument proves that $f - g$ is measurable.

To prove that $fg$ is measurable, note that by Proposition 5.3.6 and what we have already proved, $f^2$, $g^2$, and $(f + g)^2$ are measurable, and hence

$$fg = \frac{1}{2}\left((f + g)^2 - f^2 - g^2\right)$$

is measurable (check the details). □

We would often like to apply the result above to functions taking values in the extended real numbers, but the problem is that the expressions need not make sense. As we shall mainly be interested in functions that are finite except on a set of measure zero, there is a way out of the problem. Let us start with the terminology.

**Definition 5.3.8** *Let $(X, \mathcal{A}, \mu)$ be a measure space. We say that a measurable function $f : X \to \overline{\mathbb{R}}$ is* finite almost everywhere *if the set $\{x \in X : f(x) = \pm\infty\}$ has measure zero. We say that two measurable functions $f, g : X \to \overline{\mathbb{R}}$ are* equal almost everywhere *if the set $\{x \in X : f(x) \neq g(x)\}$ has measure zero. We usually abbreviate "almost everywhere" by "a.e.".*

If the measurable functions $f$ and $g$ are finite a.e., we can modify them to get measurable functions $f'$ and $g'$ which take values in $\mathbb{R}$ and are equal a.e. to $f$ and $g$, respectively (see exercise 13). By the proposition above, $f' + g'$, $f' - g'$ and $f'g'$ are measurable, and for many purposes they are good representatives for $f + g$, $f - g$ and $fg$.

Let us finally see what happens to limits of sequences.

**Proposition 5.3.9** *Let $(X, \mathcal{A}, \mu)$ be a measure space. If $\{f_n\}$ is a sequence of measurable functions $f_n : X \to \overline{\mathbb{R}}$, then $\sup_{n\in\mathbb{N}} f_n(x)$, $\inf_{n\in\mathbb{N}} f_n(x)$, $\limsup_{n\to\infty} f_n(x)$ and $\liminf_{n\to\infty} f_n(x)$ are measurable. If the sequence converges pointwise, then $\lim_{n\to\infty} f_n(x)$ is a measurable function.*

*Proof:* To see that $f(x) = \sup_{n\in\mathbb{N}} f_n(x)$ is measurable, we use Proposition 5.3.4(iii). For any $r \in \mathbb{R}$

$$f^{-1}((r, \infty)]) = \{x \in X : \sup_{n\in\mathbb{N}} f_n(x) > r\} =$$

$$= \bigcup_{n\in\mathbb{N}} \{\mathbf{x} \in X : f_n(x) > r\} = \bigcup_{n\in\mathbb{N}} f_n^{-1}((r, \infty]) \in \mathcal{A}$$

and hence $f$ is measurable by Propostion 5.3.4(iii). The argument for $\inf_{n\in\mathbb{N}} f_n(x)$ is similar.

To show that $\limsup_{n\to\infty} f_n(x)$ is measurable, first observe that the functions

$$g_k(x) = \sup_{n\geq k} f_n(x)$$

are measurable by what we have already shown. Since

$$\limsup_{n\to\infty} f_n(x) = \inf_{k\in\mathbb{N}} g_k(x),$$

the measurability of $\limsup_{n\to\infty} f_n(x)$ follows. A similar argument holds for $\liminf_{n\to\infty} f_n(x)$. If the sequence converges pointwise, then $\lim_{n\to\infty} f_n(x) =$

$\limsup_{n\to\infty} f_n(x)$ and is hence measurable. $\square$

The results above are extremely important. Mathematical analysis abounds with limit arguments, and knowing that the limit function is measurable, is often a key ingredient in these arguments.

## Exercises for Section 5.3

1. Show that if $f : X \to \overline{\mathbb{R}}$ is measurable, the sets $f^{-1}(\{\infty\})$ and $f^{-1}(\{-\infty\})$ are measurable.
2. Complete the proof of Proposition 5.3.3 by showing that $f^{-1}$ of the intervals $(-\infty, r)$, $(-\infty, r]$, $[r, \infty)$, $(r, \infty)$, $(-\infty, \infty)$, where $r \in \mathbb{R}$, are measurable.
3. Prove Proposition 5.3.4.
4. Fill in the details in the proof of Lemma 5.3.2. Explain in particular why there is only a countable number of rational intervals and why the open set $G$ is the union of the rational intervals contained in it.
5. Show that if $f_1, f_2, \ldots, f_n$ are measurable functions with values in $\mathbb{R}$, then $f_1 + f_2 + \cdots + f_n$ and $f_1 f_2 \cdot \ldots \cdot f_n$ are measurable.
6. The *indicator function* of a set $A \subset X$ is defined by

$$\mathbf{1}_A(x) = \begin{cases} 1 & \text{if } x \in A \\ 0 & \text{otherwise} \end{cases}$$

   a) Show that $\mathbf{1}_A$ is a measurable function if and only if $A \in \mathcal{A}$.
   b) A *simple* function is a function $f : X \to \mathbb{R}$ of the form

$$f(x) = \sum_{i=1}^{n} a_i \mathbf{1}_{A_i}(x)$$

   where $a_1, a_2, \ldots, a_n \in \mathbb{R}$ and $A_1, A_2, \ldots, A_n \in \mathcal{A}$. Show that all simple functions are measurable.
7. Show that if $f$ is measurable, then $f^{-1}(B) \in \mathcal{A}$ for all Borel sets $B$ (it may help to take a look at Exercise 5.1.8).
8. Let $\{E_n\}$ be a disjoint sequence of measurable sets such that $\bigcup_{n=1}^{\infty} E_n = X$, and let $\{f_n\}$ be a sequence of measurable functions. Show that the function defined by

$$f(x) = f_n(x) \text{ when } x \in E_n$$

   is measurable.
9. Fill in the details of the proof of the $fg$ part of Proposition 5.3.7. You may want to prove first that if $h : X \to \mathbb{R}$ is measurable, then so is $\frac{h}{2}$.
10. Prove the inf- and the $\liminf$-part of Proposition 5.3.9.
11. Let us write $f \sim g$ to denote that $f$ and $g$ are two measurable functions which are equal a.e.. Show that $\sim$ is an equivalence relation, i.e.:

(i) $f \sim f$

(ii) If $f \sim g$, then $g \sim f$.

(iii) If $f \sim g$ and $g \sim h$, then $f \sim h$.

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

    a) Assume that the measure space is complete. Show that if $f : X \to \overline{\mathbb{R}}$ is measurable and $g : X \to \overline{\mathbb{R}}$ equals $f$ almost everywhere, then $g$ is measurable.

    b) Show by example that the result in a) does not hold without the completeness condition. You may, e.g., use the measure space in Exercise 5.2.1.

13. Assume that the measurable function $f : X \to \mathbb{R}$ is finite a.e. Define a new function $f' : X \to \mathbb{R}$ by

$$f'(x) = \begin{cases} f(x) & \text{if } f(x) \text{ is finite} \\ 0 & \text{otherwise} \end{cases}$$

    Show that $f'$ is measurable and equal to $f$ a.e.

14. A sequence $\{f_n\}$ of measurable functions is said to *converge almost everywhere* to $f$ if there is a set $A$ of measure 0 such that $f_n(x) \to f(x)$ for all $x \notin A$.

    a) Show that if the measure space is complete, then $f$ is necessarily measurable.

    b) Show by example that the result in a) doesn't hold without the completeness assumption (take a look at Problem 12 above).

15. Let $X$ be a set and $\mathcal{F}$ a collection of functions $f : X \to \mathbb{R}$. Show that there is a smallest $\sigma$-algebra $\mathcal{A}$ on $X$ such that all the functions $f \in \mathcal{F}$ are measurable with respect to $\mathcal{A}$ (this is called the *$\sigma$-algebra generated by $\mathcal{F}$*). Show that if $X$ is a metric space and all the functions in $\mathcal{F}$ are continuous, then $\mathcal{A} \subset \mathcal{B}$, where $\mathcal{B}$ is the Borel $\sigma$-algebra.

## 5.4 Integration of simple functions

We are now ready to look at integration. The integrals we shall work with are of the form $\int f \, d\mu$ where $f$ is a measurable function and $\mu$ is a measure, and the theory is at the same time a *refinement* and a *generalization* of the classical theory Riemann integration that you know from calculus.

It is a *refinement* because if we choose $\mu$ to be the one-dimensional Lebesgue measure, the new integral $\int f \, d\mu$ equals the traditional Riemann integral $\int f(x) \, dx$ for all Riemann integrable functions, but is defined for many more functions. The same holds in higher dimensions: If $\mu$ is $n$-dimensional Lebesgue measure, then $\int f \, d\mu$ equals the Riemann integral

$\int f(x_1, \ldots, x_n)\, dx_1 \ldots dx_n$ for all Riemann integrable functions, but is defined for many more functions. The theory is also a vast *generalization* of the old one as it will allow us to integrate functions on all measure spaces and not only on $\mathbb{R}^n$.

One of the advantages of the new (Lebesgue) theory is that it will allow us to interchange limits and integrals:

$$\lim_{n\to\infty} \int f_n \, d\mu = \int \lim_{n\to\infty} f_n \, d\mu$$

in much greater generality than before. Such interchanges are of great importance in many arguments, but are problematic for the Riemann integral as there is in general no reason why the limit function $\lim_{n\to\infty} f_n$ should be Riemann integrable even when the individual functions $f_n$ are. According to Proposition 5.3.9, $\lim_{n\to\infty} f_n$ is measurable whenever the $f_n$'s are, and this makes it much easier to establish limit theorems for the new kind of integrals.

We shall develop integration theory in three steps: In this section we shall look at integrals of so-called *simple functions* which are generalizations of step functions; in the next section we shall introduce integrals of nonnegative mesurable functions; and in section 5.6 we shall extend the theory to functions taking both positive and negative values.

Thoughout this section we shall be working with a measure space $(X, \mathcal{A}, \mu)$. If $A$ is a subset of $X$, we define its *indicator function* by

$$\mathbf{1}_A(x) = \begin{cases} 1 & \text{if } x \in A \\ \\ 0 & \text{otherwise} \end{cases}$$

The indicator function is measurable if and only if $A$ is measurable.

A measurable function $f : X \to \mathbb{R}$ is called a *simple* function if it takes only finitely many different values $a_1, a_2, \ldots, a_n$. We may then write

$$f(x) = \sum_{i=1}^{n} a_1 \mathbf{1}_{A_i}(x)$$

where the sets $A_i = \{x \in X \mid f(x) = a_i\}$ are disjoint and measurable. Note that if one of the $a_i$'s is zero, the term does not contribute to the sum, and it is occasionally convenient to drop it.

If we instead start with measurable sets $B_1, B_2, \ldots, B_m$ and real numbers $b_1, b_2, \ldots, b_m$, then

$$g(x) = \sum_{i=1}^{m} b_i \mathbf{1}_{B_i}(x)$$

is measurable and takes only finitely many values, and hence is a simple function. The difference between $f$ and $g$ is that the sets $A_1, A_2, \ldots, A_n$ in

$f$ are disjoint with union $X$, and that the numbers $a_1, a_2, \ldots, a_n$ are distinct. The same need not be the case for $g$. We say that the simple function $f$ is on *standard form*, while $g$ is not (unless, of course, the $b_i$'s happen to be distinct and the sets $B_i$ are disjoint and make up all of $X$).

You may think of a simple function as a generalized step function. The difference is that step functions are constant on intervals (in $\mathbb{R}$), rectangles (in $\mathbb{R}^2$), or boxes (in higher dimensions), while a simple function need only be constant on much more complicated (but still measurable) sets.

We can now define the integral of a nonnegative simple function.

**Definition 5.4.1** *Assume that*

$$f(x) = \sum_{i=1}^{n} a_i \mathbf{1}_{A_i}(x)$$

*is a nonnegative simple function on standard form. Then the* integral of $f$ with respect to $\mu$ *is defined by*

$$\int f \, d\mu = \sum_{i=1}^{n} a_i \mu(A_i)$$

*Recall that we are using the convention that* $0 \cdot \infty = 0$*, and hence* $a_i \mu(A_i) = 0$ *if* $a_i = 0$ *and* $\mu(A_i) = \infty$*.*

Note that the integral of an indicator function is

$$\int \mathbf{1}_A \, d\mu = \mu(A)$$

To see that the definition is reasonable, assume that you are in $\mathbb{R}^2$. Since $\mu(A_i)$ measures the area of the set $A_i$, the product $a_i \mu(A_i)$ measures in an intuitive way the volume of the solid with base $A_i$ and height $a_i$.

We need to know that the formula in the definition also holds when the simple function is not on standard form. The first step is the following, simple lemma:

**Lemma 5.4.2** *If*

$$g(x) = \sum_{j=1}^{m} b_j \mathbf{1}_{B_j}(x)$$

*is a nonnegative simple function where the* $B_j$*'s are disjoint and* $X = \bigcup_{j=1}^{m} B_j$*, then*

$$\int g \, d\mu = \sum_{j=1}^{n} b_j \mu(B_j)$$

*Proof:* The problem is that the values $b_1, b_2, \ldots, b_m$ need not be distinct, but this is easily fixed: If $c_1, c_2, \ldots, c_k$ are the distinct values taken by $g$, let $b_{i_1}$, $b_{i_2}$,...,$b_{i_{n_i}}$ be the $b_j$'s that are equal to $c_i$, and let $C_i = B_{i_1} \cup B_{i_2} \cup \ldots \cup B_{i_{n_i}}$. Then $\mu(C_i) = \mu(B_{i_1}) + \mu(B_{i_2}) + \ldots + \mu(B_{i_{n_i}})$, and hence

$$\sum_{j=1}^{m} b_j \mu(B_j) = \sum_{i=1}^{k} c_i \mu(C_i)$$

Since $g(x) = \sum_{i=1}^{k} c_i \mathbf{1}_{C_i}(x)$ is the standard form representation of $g$, we have

$$\int g \, d\mu = \sum_{i=1}^{k} c_i \mu(C_i)$$

and the lemma is proved □

The next step is also easy:

**Proposition 5.4.3** *Assume that $f$ and $g$ are two nonnegative simple functions, and let $c$ be a nonnnegative, real number. Then*

*(i)* $\int cf \, d\mu = c \int f \, d\mu$

*(ii)* $\int (f+g) \, d\mu = \int f \, d\mu + \int g \, d\mu$

*Proof:* (i) is left to the reader. To prove (ii), let

$$f(x) = \sum_{i=1}^{n} a_i \mathbf{1}_{A_i}(x)$$

$$g(x) = \sum_{j=1}^{n} b_j \mathbf{1}_{B_j}(x)$$

be standard form representations of $f$ and $g$, and define $C_{i,j} = A_i \cap B_j$. By the lemma above

$$\int f \, d\mu = \sum_{i,j} a_i \mu(C_{i,j})$$

and

$$\int g \, d\mu = \sum_{i,j} b_j \mu(C_{i,j})$$

and also

$$\int (f+g) \, d\mu = \sum_{i,j} (a_i + b_j) \mu(C_{i,j})$$

since the value of $f+g$ on $C_{i,j}$ is $a_i + b_j$ □

**Remark:** Using induction, we can extend part (ii) above to longer sums:

$$\int (f_1 + f_2 + \cdots + f_n)\, d\mu = \int f_1\, d\mu + \int f_2\, d\mu + \ldots + \int f_n\, d\mu$$

for all nonnegative, simple functions $f_1, f_2, \ldots, f_n$,

We can now prove that the formula in Definition 5.4.1 holds for all representations of simple functions, and not only the standard ones:

**Corollary 5.4.4** *If $f(x) = \sum_{n=1} a_i \mathbf{1}_{A_i}(x)$ is a step function with $a_i \geq 0$ for all $i$, then*

$$\int f\, d\mu = \sum_{i=1}^{n} a_i \mu(A_i)$$

*Proof:* By the results above

$$\int f\, d\mu = \int \sum_{i=1}^{n} a_i \mathbf{1}_{A_i}\, d\mu = \sum_{i=1}^{n} \int a_i \mathbf{1}_{A_i}\, d\mu = \sum_{i=1}^{n} a_i \int \mathbf{1}_{A_i}\, d\mu = \sum_{i=1}^{n} a_i \mu(A_i)$$

□

We need to prove yet another almost obvious result. We write $g \leq f$ to say that $g(x) \leq f(x)$ for all $x$.

**Proposition 5.4.5** *Assume that $f$ and $g$ are two nonnegative simple functions. If $g \leq f$, then*

$$\int g\, d\mu \leq \int f\, d\mu$$

*Proof:* We use the same trick as in the proof of Proposition 5.4.3: Let

$$f(x) = \sum_{i=1}^{n} a_i \mathbf{1}_{A_i}(x)$$

$$g(x) = \sum_{j=1}^{m} b_j \mathbf{1}_{B_j}(x)$$

be standard form representations of $f$ and $g$, and define $C_{i,j} = A_i \cap B_j$. Then

$$\int f\, d\mu = \sum_{i,j} a_i \mu(C_{i,j}) \geq \sum_{i,j} b_j \mu(C_{i,j}) = \int g\, d\mu$$

□

We shall end this section with a key result on limits of integrals, but first we need some notation. Observe that if $f = \sum_{i=1}^{n} a_i \mathbf{1}_{A_i}$ is a simple

function and $B$ is a measurable set, then $\mathbf{1}_B f = \sum_{i=1}^n a_i \mathbf{1}_{A_i \cap B}$ is also a simple function. We shall write

$$\int_B f \, d\mu = \int \mathbf{1}_B f \, d\mu$$

and call this the *integral of $f$ over $B$*. The lemma below may seem obvious, but it is the key to many later results.

**Lemma 5.4.6** *Assume that $B$ is a measurable set, $b$ a positive real number, and $\{f_n\}$ an increasing sequence of nonnegative simple functions such that $\lim_{n\to\infty} f_n(x) \geq b$ for all $x \in B$. Then $\lim_{n\to\infty} \int_B f_n \, d\mu \geq b\mu(B)$.*

*Proof:* Let $a$ be any positive number less than $b$, and define

$$A_n = \{x \in B \mid f_n(x) \geq a\}$$

Since $f_n(x) \uparrow b$ for all $x \in B$, we see that the sequence $\{A_n\}$ is increasing and that

$$B = \bigcup_{n=1}^{\infty} A_n$$

By continuity of measure (Proposition 5.1.5(i)), $\mu(B) = \lim_{n\to\infty} \mu(A_n)$, and hence for any positive number $m$ less that $\mu(B)$, we can find an $N \in \mathbb{N}$ such that $\mu(A_n) > m$ when $n \geq N$. Since $f_n \geq a$ on $A_n$, we thus have

$$\int_B f_n \, d\mu \geq \int_{A_n} a \, d\mu = am$$

whenever $n \geq N$. Since this holds for any number $a$ less than $b$ and any number $m$ less than $\mu(B)$, we must have $\lim_{n\to\infty} \int_B f_n \, d\mu \geq b\mu(B)$. □

To get the result we need, we extend the lemma to simple functions:

**Proposition 5.4.7** *Let $g$ be a nonnegative simple function and assume that $\{f_n\}$ is an increasing sequence of nonnegative simple functions such that $\lim_{n\to\infty} f_n(x) \geq g(x)$ for all $x$. Then*

$$\lim_{n\to\infty} \int f_n \, d\mu \geq \int g \, d\mu$$

*Proof:* Let $g(x) = \sum_{i=1}^m b_i \mathbf{1}_{B_1}(x)$ be the standard form of $g$. If any of the $b_i$'s is zero, we may just drop that term in the sum, so that we from now on assume that all the $b_i$'s are nonzero. By Corollary 5.4.3(ii), we have

$$\int_{B_1 \cup B_2 \cup \ldots \cup B_m} f_n \, d\mu = \int (\mathbf{1}_{B_1} + \mathbf{1}_{B_2} + \ldots + \mathbf{1}_{B_m}) \, f_n \, d\mu =$$

$$= \int (\mathbf{1}_{B_1} f_n + \mathbf{1}_{B_2} f_n + \ldots + \mathbf{1}_{B_m} f_n)\, d\mu =$$

$$= \int_{B_1} f_n\, d\mu + \int_{B_2} f_n\, d\mu + \ldots + \int_{B_m} f_n\, d\mu$$

By the lemma, $\lim_{n\to\infty} \int_{B_i} f_n\, d\mu \geq b_i \mu(B_i)$, and hence

$$\lim_{n\to\infty} \int f_n\, d\mu \geq \lim_{n\to\infty} \int_{B_1 \cup B_2 \cup \ldots \cup B_m} f_n\, d\mu \geq \sum_{i=1}^{m} b_i \mu(B_i) = \int g\, d\mu$$

□

We are now ready to extend the integral to all positive, measurable functions. This will be the topic of the next section.

### Exercises for Section 5.4

1. Show that if $f$ is a measurable function, then the *level set*
$$A_a = \{x \in X \mid f(x) = a\}$$
is measurable for all $a \in \overline{\mathbb{R}}$.
2. Check that according to Definition 5.4.1, $\int \mathbf{1}_A\, \mathbf{d}\mu = \mu(A)$ for all $A \in \mathcal{A}$.
3. Prove part (i) of Proposition 5.4.3.
4. Show that if $f_1, f_2, \ldots, f_n$ are simple functions, then so are
$$h(x) = \max\{f_1(x), f_2(x), \ldots, f_n(x)\}$$
and
$$h(x) = \min\{f_1(x), f_2(x), \ldots, f_n(x)\}$$
5. Let $\mu$ be Lebesgue measure, and define $A = \mathbb{Q} \cap [0,1]$. The function $\mathbf{1}_A$ is not integrable in the Riemann sense. What is $\int \mathbf{1}_A\, d\mu$?
6. Let $f$ be a nonnegative, simple function on a measure space $(X, \mathcal{A}, \mu)$. Show that
$$\nu(B) = \int_B f\, d\mu$$
defines a measure $\nu$ on $(X, \mathcal{A})$.

## 5.5 Integrals of nonnegative functions

We are now ready to define the integral of a general, nonnegative, measurable function. Throughout the sextion, $(X, \mathcal{A}, \mu)$ is a measure space.

**Definition 5.5.1** *If $f : X \to \overline{\mathbb{R}}_+$ is measurable, we define*

$$\int f\, d\mu = \sup\{\int g\, d\mu \mid g \text{ is a nonnegative simple function, } g \leq f\}$$

**Remark:** Note that if $f$ is a simple function, we now have two definitions of $\int f\,d\mu$; the original one in Definition 5.4.1 and a new one in the definition above. It follows from Proposition 5.4.5 that the two definitions agree.

The definition above is natural, but also quite abstract, and we shall work toward a reformulation that is often easier to handle.

**Proposition 5.5.2** *Let $f : X \to \overline{\mathbb{R}}_+$ be a measurable function, and assume that $\{h_n\}$ is an increasing sequence of simple functions converging pointwise to $f$. Then*

$$\lim_{n\to\infty} \int h_n\,d\mu = \int f\,d\mu$$

*Proof:* Since the sequence $\{\int h_n\,d\mu\}$ is increasing by Proposition 5.4.5, the limit clearly exists (it may be $\infty$), and since $\int h_n\,d\mu \le \int f\,d\mu$ for all $n$, we must have

$$\lim_{n\to\infty} \int h_n\,d\mu \le \int f\,d\mu$$

To get the opposite inequality, it suffices to show that

$$\lim_{n\to\infty} \int h_n\,d\mu \ge \int g\,d\mu$$

for each simple function $g \le f$, but this follows from Proposition 5.4.7. □

The proposition above would lose much of its power if there weren't any increasing sequences of simple functions converging to $f$. The next result tells us that there always are. Pay attention to the argument; it is a key to why the theory works.

**Proposition 5.5.3** *If $f : X \to \overline{\mathbb{R}}_+$ is measurable, there is an increasing sequence $\{h_n\}$ of simple functions converging pointwise to $f$. Moreover, for each $n$ either $f(x) - \frac{1}{2^n} < h_n(x) \le f(x)$ or $h_n(x) = 2^n$*

*Proof:* To construct the simple function $h_n$, we cut the interval $[0, 2^n)$ into half-open subintervals of length $\frac{1}{2^n}$, i.e. intervals

$$I_k = \left[\frac{k}{2^n}, \frac{k+1}{2^n}\right)$$

where $0 \le k < 2^{2n}$, and then let

$$A_k = f^{-1}(I_k)$$

We now define

$$h_n(x) = \sum_{k=0}^{2^{2n}-1} \frac{k}{2^n}\mathbf{1}_{A_k}(x) + 2^n\mathbf{1}_{\{x\,|\,f(x)\ge 2^n\}}$$

By definition, $h_n$ is a simple function no greater than $f$. Since the intervals get narrower and narrower and cover more and more of $[0, \infty)$, it is easy to see that $h_n$ converges pointwise to $f$. To see why the sequence increases, note that each time we increase $n$ by one, we split each of the former intervals $I_k$ in two, and this will cause the new step function to equal the old one for some $x$'s and jump one step upwards for others (make a drawing).

The last statement follows directly from the construction. □

**Remark:** You should compare the partitions in the proof above to the partitions you have earlier seen in Riemann integration. When we integrate a function of one variable in calculus, we partition an interval $[a, b]$ on the *x-axis* and use this partition to approximate the original function by a step function. In the proof above, we instead partitioned the *y-axis* into intervals and used this partition to approximate the original function by a simple function. The latter approach gives us much better control over what is going one; the partition controls the oscillations of the function. The price we have to pay, it that we get simple functions instead of step functions, and to use simple functions for integration, we need measure theory.

Let us combine the last two results in a handy corollary:

**Corollary 5.5.4** *If $f : X \to \overline{\mathbb{R}}_+$ is measurable, there is an increasing sequence $\{h_n\}$ of simple functions converging pointwise to $f$, and*

$$\int f \, d\mu = \lim_{n\to\infty} \int h_n \, d\mu$$

Let us take a look at some properties of the integral.

**Proposition 5.5.5** *Assume that $f, g : X \to \overline{\mathbb{R}}_+$ are measurable functions and that $c$ is a nonnegative, real number. Then:*

*(i)* $\int cf \, d\mu = c \int f \, d\mu$.

*(ii)* $\int (f + g) \, d\mu = \int f \, d\mu + \int g \, d\mu$.

*(iii) If $g \leq f$, then* $\int g \, d\mu \leq \int f \, d\mu$.

*Proof:* (iii) is immediate from the definition, and (i) is left to the reader. To prove (ii), let $\{f_n\}$ and $\{g_n\}$ be to increasing sequence of simple functions converging to $f$ and $g$, respectively. Then $\{f_n + g_n\}$ is an increasing sequence of simple functions converging to $f + g$, and

$$\int (f + g) \, d\mu = \lim_{n\to\infty} \int (f_n + g_n) \, d\mu = \lim_{n\to\infty} \left( \int f_n \, d\mu + \int g_n \, d\mu \right) =$$

$$= \lim_{n\to\infty} \int f_n \, d\mu + \lim_{n\to\infty} \int g_n \, d\mu = \int f \, d\mu + \int g \, d\mu$$

where we have used Proposition 5.4.3(ii) to go from $\int (f_n+g_n)\,d\mu$ to $\int f_n\,d\mu + \int g_n\,d\mu$. □

One of the great advantages of the Lebesgue integration theory we are now developing is that it is much better behaved with respect to limits than the Riemann theory you are used to. Here is a typical example:

**Theorem 5.5.6 (Monotone Convergence Theorem)** *If $\{f_n\}$ is an increasing sequence of nonnegative, measurable functions such that $f(x) = \lim_{n\to\infty} f_n(x)$ for all $x$, then*

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

*In other words,*

$$\lim_{n\to\infty} \int f_n\,d\mu = \int \lim_{n\to\infty} f_n\,d\mu$$

*Proof:* We know from Proposition 5.3.8 that $f$ is measurable, and hence the integral $\int f\,d\mu$ is defined. Since $f_n \le f$, we have $\int f_n\,d\mu \le \int f\,d\mu$ for all $n$, and hence

$$\lim_{n\to\infty} \int f_n\,d\mu \le \int f\,d\mu$$

To prove the opposite inequality, we approximate each $f_n$ by simple functions as in the proof of Proposition 5.5.3; in fact, let $h_n$ be the $n$-th approximation to $f_n$. Assume that we can prove that the sequence $\{h_n\}$ converges to $f$; then

$$\lim_{n\to\infty} \int h_n\,d\mu = \int f\,d\mu$$

by Proposition 5.5.2. Since $f_n \ge h_n$, this would give us the desired inequality

$$\lim_{n\to\infty} \int f_n\,d\mu \ge \int f\,d\mu$$

It remains to show that $h_n(x) \to f(x)$ for all $x$. From Proposition 5.5.3 we know that for all $n$, either $f_n(x) - \frac{1}{2^n} < h_n(x) \le f_n(x)$ or $h_n(x) = 2^n$. If $h_n(x) = 2^n$ for infinitely many $n$, then $h_n(x)$ goes to $\infty$, and hence to $f(x)$. If $h_n(x)$ is *not* equal to $2^n$ for infinitely many $n$, then we eventually have $f_n(x) - \frac{1}{2^n} < h_n(x) \le f_n(x)$, and hence $h_n(x)$ converges to $f(x)$ since $f_n(x)$ does. □

We would really have liked the formula

$$\lim_{n\to\infty} \int f_n\,d\mu = \int \lim_{n\to\infty} f_n\,d\mu \tag{5.5.1}$$

above to hold in general, but as the following example shows, this is not the case.

**Example 1:** Let $\mu$ be the counting measure on $\mathbb{N}$, and define the sequence $\{f_n\}$ by

$$f_n(x) = \begin{cases} 1 & \text{if } x = n \\ \\ 0 & \text{otherwise} \end{cases}$$

Then $\lim_{n\to\infty} f_n(x) = 0$ for all $x$, but $\int f_n \, d\mu = 1$. Hence

$$\lim_{n\to\infty} \int f_n \, d\mu = 1$$

but

$$\int \lim_{n\to\infty} f_n \, d\mu = 0$$

♣

There are many results in measure theory giving conditions for (5.5.1) to hold, but there is no ultimate theorem covering all others. There is, however, a simple inequality that always holds.

**Theorem 5.5.7 (Fatou's Lemma)** *Assume that $\{f_n\}$ is a sequence of non-negative, measurable functions. Then*

$$\liminf_{n\to\infty} \int f_n \, d\mu \geq \int \liminf_{n\to\infty} f_n \, d\mu$$

*Proof:* Let $g_k(x) = \inf_{k\geq n} f_n(x)$. Then $\{g_k\}$ is an increasing sequence of measurable functions, and by the Monotone Convergence Theorem

$$\lim_{k\to\infty} \int g_k \, d\mu = \int \lim_{k\to\infty} g_k \, d\mu = \int \liminf_{n\to\infty} f_n \, d\mu$$

where we have used the definition of $\liminf$ in the last step. Since $f_k \geq g_k$, we have $\int f_k \, d\mu \geq \int g_k \, d\mu$, and hence

$$\liminf_{k\to\infty} \int f_k \, d\mu \geq \lim_{k\to\infty} \int g_k \, d\mu = \int \liminf_{n\to\infty} f_n \, d\mu$$

and the result is proved. □

Fatou's Lemma is often a useful tool in establishing more sophisticated results, see Exercise 16 for a typical example.

Just as for simple functions, we define integrals over measurable subsets $A$ of $X$ by the formula

$$\int_A f\,d\mu = \int \mathbf{1}_A f\,d\mu$$

So far we have allowed our integrals to be infinite, but we are mainly interested in situations where $\int f\,d\mu$ is finite:

**Definition 5.5.8** *A function $f : X \to [0, \infty]$ is said to be* integrable *if it is measurable and $\int f\,d\mu < \infty$.*

## Comparison with Riemann integration

We shall end this section with a quick comparison between the integral we have now developed and the Riemann integral you learned in calculus. Let us begin with a quick review of the Riemann integral[1].

Assume that $[a, b]$ is a closed and bounded interval, and let $f : [a, b] \to \mathbb{R}$ be a nonnegative, bounded function. Recall that a *partition* $\mathcal{P}$ of the interval $[a, b]$ is a finite set $\{x_0, x_1, \ldots, x_n\}$ such that

$$a = x_0 < x_1 < x_2 < \ldots < x_n = b$$

The lower and upper values of $f$ over the interval $(x_{i-1}, x_i]$ are

$$m_i = \inf\{f(x) \,|\, x \in (x_{i-1}, x_i]\}$$

and

$$M_i = \sup\{f(x) \,|\, x \in (x_{i-1}, x_i]\}$$

respectively, and the lower and upper sums of the partition $\mathcal{P}$ are

$$L(\mathcal{P}) = \sum_{i=1}^{n} m_i(x_i - x_{i-1})$$

and

$$U(\mathcal{P}) = \sum_{i=1}^{n} M_i(x_i - x_{i-1})$$

The function $f$ is *Riemann integrable* if the lower integral

$$\underline{\int_a^b} f(x)\,dx = \sup\{L(\mathcal{P}) \,|\, \mathcal{P} \text{ is a partition of } [a, b]\}$$

and the upper integral

$$\overline{\int_a^b} f(x)\,dx = \inf\{U(\mathcal{P}) \,|\, \mathcal{P} \text{ is a partition of } [a, b]\}$$

[1] The approach to Riemann integration that I describe here is actually due to the French mathematician Gaston Darboux (1842-1917).

coincide, in which case we define the *Riemann integral* $\int_a^b f(x)\,dx$ to be the common value.

We are now ready to compare the Riemann integral $\int_a^b f(x)\,dx$ and the Lebesgue integral $\int_{[a,b]} f\,d\mu$ ($\mu$ is now the Lebesgue measure). Observe first that if we define simple functions

$$\phi_{\mathcal{P}} = \sum_{i=1}^{n} m_i \mathbf{1}_{(x_{i-1},x_i]}$$

and

$$\Phi_{\mathcal{P}} = \sum_{i=1}^{n} M_i \mathbf{1}_{(x_{i-1},x_i]}$$

we have

$$\int \phi_{\mathcal{P}}\,d\mu = \sum_{i=1}^{n} m_i(x_i - x_{i-1}) = N(\mathcal{P})$$

and

$$\int \Phi_{\mathcal{P}}\,d\mu = \sum_{i=1}^{n} M_i(x_i - x_{i-1}) = U(\mathcal{P})$$

**Theorem 5.5.9** *Assume that $f : [a,b] \to [0,\infty)$ is a bounded, Riemann integrable function on $[a,b]$. Then $f$ is measurable and the Riemann and the Lebesgue integral coincide:*

$$\int_a^b f(x)\,dx = \int_{[a.b]} f\,d\mu$$

*Proof:* Since $f$ is Riemann integrable, we can pick a sequence $\{\mathcal{P}_n\}$ of partitions such that the sequences $\{\phi(\mathcal{P}_n)\}$ of lower step functions is increasing, the sequence $\{\Phi(\mathcal{P}_n)\}$ of upper step functions is decreasing, and

$$\lim_{n\to\infty} L(\mathcal{P}_n) = \lim_{n\to\infty} U(\mathcal{P}_n) = \int_a^b f(x)\,dx$$

(see Exercise 10 for help), or in other words

$$\lim_{n\to\infty} \int \phi_{\mathcal{P}_n}\,d\mu = \lim_{n\to\infty} \int \Phi_{\mathcal{P}_n}\,d\mu = \int_a^b f(x)\,dx$$

This means that

$$\lim_{n\to\infty} \int (\Phi_{\mathcal{P}_n} - \phi_{\mathcal{P}_n})\,d\mu = 0$$

and by Fatou's lemma, we have

$$\int \lim_{n\to\infty} (\Phi_{\mathcal{P}_n} - \phi_{\mathcal{P}_n})\,d\mu = 0$$

(the limits exists since the sequence $\Phi_{\mathcal{P}_n} - \phi_{\mathcal{P}_n}$ is decreasing). This means that $\lim_{n\to\infty} \phi_{\mathcal{P}_n} = \lim_{n\to\infty} \Phi_{\mathcal{P}_n}$ a.e., and since

$$\lim_{n\to\infty} \phi_{\mathcal{P}_n} \leq f \leq \lim_{n\to\infty} \Phi_{\mathcal{P}_n},$$

$f$ must be measurable as it squeezed between two almost equal, measurable functions. Also, since $f = \lim_{n\to\infty} \phi_{\mathcal{P}_n}$ a.s., the Monotone Convergence Theorem (we are actually using the slightly extended version in Exercise 13) tells us that

$$\int_{[a,b]} f\, d\mu = \lim_{n\to\infty} \int \phi_{\mathcal{P}_n}\, d\mu = \lim_{n\to\infty} U(\mathcal{P}_n) = \int_a^b f(x)\, dx$$

□

The theorem above can be extended in many directions. Exactly the same proof works for Riemann integrals over rectangular boxes in $\mathbb{R}^d$, and once we have introduced integrals of functions taking both positive and negative values in the next section, it easy to extend the theorem above to that situation. There are some subtleties concerning improper integrals, but we shall not touch on these here. Our basic message is: Lebesgue integration is just like Riemann integration, only better (because more functions are integrable and we can integrate in completely new contexts — all we need is a measure)!

## Exercises for Section 5.5

1. Assume $f : X \to [0, \infty]$ is a nonnegative simple function. Show that the two definitions of $\int f\, d\mu$ given in Definitions 5.4.1 and 5.5.1 coincide.
2. Prove Proposition 5.5.5(i).
3. Show that if $f : X \to [0, \infty]$ is measurable, then

$$\mu(\{x \in X \mid f(x) \geq a\}) \leq \frac{1}{a} \int f\, d\mu$$

   for all positive, real numbers $a$.
4. In this problem, $f, g : X \to [0, \infty]$ are measurable functions.
   a) Show that $\int f\, d\mu = 0$ if and only if $f = 0$ a.e.
   b) Show that if $f = g$ a.e., then $\int f\, d\mu = \int g\, d\mu$.
   c) Show that if $\int_E f\, d\mu = \int_E g\, d\mu$ for all measurable sets $E$, then $f = g$ a.e.
5. Assume that $(X, \mathcal{A}, \mu)$ is a measure space and that $f : X \to [0, \infty]$ is a nonnegative, measurable function
   a) Show that if $A$, $B$ are measurable sets with $A \subset B$, then $\int_A f\, d\mu \leq \int_B f\, d\mu$

b) Show that if $A, B$ are disjoint, measurable sets, then $\int_{A\cup B} f\,d\mu = \int_A f\,d\mu + \int_B f\,d\mu$.

c) Define $\nu : \mathcal{A} \to \overline{\mathbb{R}}$ by

$$\nu(A) = \int_A f\,d\mu$$

Show that $\nu$ is a measure.

6. Show that if $f : X \to [0, \infty]$ is integrable, then $f$ is finite a.e.

7. Let $\mu$ be Lebesgue measure on $\mathbb{R}$ and assume that $f : \mathbb{R} \to \overline{\mathbb{R}}_+$ is a nonnegative, measurable function. Show that

$$\lim_{n\to\infty} \int_{[-n,n]} f\,d\mu = \int f\,d\mu$$

8. Let $\mu$ be Lebesgue measure on $\mathbb{R}$. Show that for all measurable sets $A \subset \mathbb{R}$

$$\lim_{n\to\infty} \int_A \sum_{k=1}^{n} \frac{x^{2k}}{k!}\,d\mu = \int_A e^{x^2}\,d\mu$$

9. Let $f : \mathbb{R} \to \mathbb{R}$ be the function

$$f(x) = \begin{cases} 1 & \text{if } x \text{ is rational} \\ 0 & \text{otherwise} \end{cases}$$

and for each $n \in \mathbb{N}$, let $f_n : \mathbb{R} \to \mathbb{R}$ be the function

$$f_n(x) = \begin{cases} 1 & \text{if } x = \frac{p}{q} \text{ where } p \in \mathbb{Z},\ q \in \mathbb{N}, q \le n \\ 0 & \text{otherwise} \end{cases}$$

a) Show that $\{f_n(x)\}$ is an increasing sequence converging to $f(x)$ for all $x \in \mathbb{R}$.

b) Show that each $f_n$ is Riemann integrable over $[0, 1]$ with $\int_0^1 f_n(x)\,dx = 0$ (this is integration as taught in calculus courses).

c) Show that $f$ is *not* Riemann integrable over $[0, 1]$.

d) Show that the one-dimensional Lebesgue integral $\int_{[0,1]} f\,d\mu$ exists and find its value.

10. In this problem we shall sketch how one may construct the sequence $\{\mathcal{P}_n\}$ of partitions in the proof of Theorem 5.5.9.

a) Call a partition $\mathcal{P}$ of $[a, b]$ *finer* than another partition $\mathcal{Q}$ if $\mathcal{Q} \subset \mathcal{P}$, and show that if $\mathcal{P}$ is finer than $\mathcal{Q}$, then $\phi_{\mathcal{P}} \ge \phi_{\mathcal{Q}}$ and $\Phi_{\mathcal{P}} \le \Phi_{\mathcal{Q}}$.

b) Show that if $f$ is as in Theorem 5.5.9, there are sequences of partitions $\{\mathcal{Q}_n\}$ and $\{\mathcal{R}_n\}$ such that

$$\lim_{n\to\infty} L(\mathcal{Q}_n) = \int_a^b f(x)\,dx$$

and

$$\lim_{n\to\infty} U(\mathcal{R}_n) = \int_a^b f(x)\,dx$$

c) For each $n$, let $\mathcal{P}_n$ be the common refinement of all partitions $\mathcal{Q}_k$ and $\mathcal{R}_k$, $k \leq n$, i.e.

$$\mathcal{P}_n = \bigcup_{k=1}^{n} (\mathcal{Q}_k \cup \mathcal{R}_k)$$

Show that $\{\mathcal{P}_n\}$ satisfies the requirements in the proof of Theorem 5.5.9.

11. a) Let $\{u_n\}$ be a sequence of positive, measurable functions. Show that

$$\int \sum_{n=1}^{\infty} u_n \, d\mu = \sum_{n=1}^{\infty} \int u_n \, d\mu$$

b) Assume that $f$ is a nonnnegative, measurable function and that $\{B_n\}$ is a disjoint sequence of measurable sets with union $B$. Show that

$$\int_B f \, d\mu = \sum_{n=1}^{\infty} \int_{B_n} f \, d\mu$$

12. Assume that $f$ is a nonnegative, measurable function and that $\{A_n\}$ is an increasing sequence of measurable sets with union $A$. Show that

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

13. Show the following generalization of the Monotone Convergence Theorem: If $\{f_n\}$ is an increasing sequence of nonnegative, measurable functions such that $f(x) = \lim_{n\to\infty} f_n(x)$ *almost everywhere.* (i.e. for all $x$ outside a set $N$ of measure zero), then

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

14. Find a decreasing sequence $\{f_n\}$ of measurable functions $f_n : \mathbb{R} \to [0, \infty)$ converging pointwise to zero such that $\lim_{n\to\infty} \int f_n \, d\mu \neq 0$

15. Assume that $f : X \to [0, \infty]$ is a measurable function, and that $\{f_n\}$ is a sequence of measurable functions converging pointwise to $f$. Show that if $f_n \leq f$ for all $n$,

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

16. Assume that $\{f_n\}$ is a sequence of nonnegative functions converging pointwise to $f$. Show that if

$$\lim_{n\to\infty} \int f_n \, d\mu = \int f \, d\mu < \infty,$$

then

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

for all measurable $E \subset X$.

17. Assume that $g : X \to [0, \infty]$ is an *integrable* function, and that $\{f_n\}$ is a sequence of nonnegative, measurable functions converging pointwise to a function $f$. Show that if $f_n \leq g$ for all $n$, then

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

*Hint:* Apply Fatou's Lemma to both sequences $\{f_n\}$ and $\{g - f_n\}$.

18. Let $(X, \mathcal{A})$ be a measurable space, and let $\mathcal{M}^+$ be the set of all non-negative, measurable functions $f : X \to \overline{\mathbb{R}}_+$. Assume that $I : \mathcal{M}^+ \to \overline{\mathbb{R}}_+$ satisfies the following three conditions:

    (i) $I(\alpha f) = \alpha I(f)$ for all $\alpha \in [0, \infty)$ and all $f \in \mathcal{M}^+$.

    (ii) $I(f + g) = I(f) + I(g)$ for all $f, g \in \mathcal{M}^+$.

    (iii) If $\{f_n\}$ is an increasing sequence from $\mathcal{M}^+$ converging pointwise to $f$, then $\lim_{n\to\infty} I(f_n) = I(f)$.

    a) Show that $I(f_1 + f_2 + \cdots + f_n) = I(f_1) + I(f_2) + \cdots + I(f_n)$ for all $n \in \mathbb{N}$ and all $f_1, f_2, \ldots, f_n \in \mathcal{M}^+$.

    b) Show that if $f, g \in \mathcal{M}^+$ and $f(x) \leq g(x)$ for all $x \in X$, then $I(f) \leq I(g)$.

    c) Show that

    $$\mu(E) = I(\mathbf{1}_E) \quad \text{for } E \in \mathcal{A}$$

    defines a measure on $(X, \mathcal{A})$.

    d) Show that $I(f) = \int f \, d\mu$ for all non-negative simple functions $f$.

    e) Show that $I(f) = \int f \, d\mu$ for all $f \in \mathcal{M}^+$.

## 5.6 Integrable functions

So far we only know how to integrate nonnegative functions, but it is not difficult to extend the theory to general functions. We have, however, to be a little more careful with the size of the functions we integrate: If a nonnegative function $f$ is too big, we may just put the integral $\int f \, d\mu$ equal to $\infty$, but if the function can take negative values as well as positive, there may be infinite contributions of opposite signs that are difficult to balance. For this reason we shall only define the integral for a class of *integrable* functions where this problem does not occur.

Given a function $f : X \to \overline{\mathbb{R}}$, we first observe that $f = f_+ - f_-$, where $f_+$ and $f_-$ are the nonnegative functions

$$f_+(x) = \begin{cases} f(x) & \text{if } f(x) > 0 \\ 0 & \text{otherwise} \end{cases}$$

and

$$f_-(x) = \begin{cases} -f(x) & \text{if } f(x) < 0 \\ 0 & \text{otherwise} \end{cases}$$

Note that $f_+$ and $f_-$ are measurable if $f$ is.

Recall that a nonnegative, measurable function $f$ is integrable if $\int f\,d\mu < \infty$.

**Definition 5.6.1** *A function $f : X \to \overline{\mathbb{R}}$ is called* integrable *if it is measurable, and $f+$ and $f_-$ are integrable. We define the* integral *of $f$ by*

$$\int f\,d\mu = \int f_+\,d\mu - \int f_-\,d\mu$$

The definition illustrates our point above: If both $\int f_+\,d\mu$ and $\int f_-\,d\mu$ are infinite, there is no natural way to define the difference $\int f_+\,d\mu - \int f_-\,d\mu$.

The next lemma gives a useful characterization of integrable functions.

**Lemma 5.6.2** *A measurable function $f$ is integrable if and only if its absolute value $|f|$ is integrable, i.e. if and only if $\int |f|\,d\mu < \infty$.*

*Proof:* Note that $|f| = f_+ + f_-$. Hence

$$\int |f|\,d\mu = \int f_+\,d\mu + \int f_-\,d\mu$$

by Proposition 5.5.5(ii), and we see that $\int |f|\,d\mu$ is finite if and only if both $\int f_+\,d\mu$ and $\int f_-\,d\mu$ are finite. □

The next lemma is another useful technical tool. It tells us that if we split $f$ as a difference $f = g - h$ of two nonnegative, integrable functions, we always get $\int f\,d\mu = \int g\,d\mu - \int h\,d\mu$ (so far we only know this for $g = f_+$ and $h = f_-$).

**Lemma 5.6.3** *Assume that $g : X \to [0, \infty]$ and $h : X \to [0, \infty]$ are two integrable, nonnegative functions, and that $f(x) = g(x) - h(x)$ at all points where the difference is defined. Then $f$ is integrable and*

$$\int f\,d\mu = \int g\,d\mu - \int h\,d\mu$$

*Proof:* Note that since $g$ and $h$ are integrable, they are finite a.e., and hence $f = g - h$ a.e. Modifying $g$ and $h$ on a set of measure zero (this will not change their integrals), we may assume that $f(x) = g(x) - h(x)$ for all $x$. Since $|f(x)| = |g(x) - h(x)| \leq |g(x)| + |h(x)|$, it follows from the lemma above that $f$ is integrable.

As

$$f(x) = f_+(x) - f_-(x) = g(x) - h(x)$$

we have

$$f_+(x) + h(x) = g(x) + f_-(x)$$

where we on both sides have sums of nonnegative functions. By Proposition 5.5.5(ii), we get

$$\int f_+ \, d\mu + \int h \, d\mu = \int g \, d\mu + \int f_- \, d\mu$$

Rearranging the integrals (they are all finite), we get

$$\int f \, d\mu = \int f_+ \, d\mu - \int f_- \, d\mu = \int g \, d\mu - \int h \, d\mu$$

and the lemma is proved. □

We are now ready to prove that the integral behaves the way we expect:

**Proposition 5.6.4** *Assume that $f, g : X \to \overline{\mathbb{R}}$ are integrable functions, and that $c$ is a constant. Then $f + g$ and $cf$ are integrable, and*

*(i)* $\int cf \, d\mu = c \int f \, d\mu$.

*(ii)* $\int (f + g) \, d\mu = \int f \, d\mu + \int g \, d\mu$.

*(iii) If $g \leq f$, then* $\int g \, d\mu \leq \int f \, d\mu$.

*Proof:* (i) is left to the reader (treat positive and negative $c$'s separately). To prove (ii), first note that since $f$ and $g$ are integrable, the sum $f(x) + g(x)$ is defined a.e., and by changing $f$ and $g$ on a set of measure zero (this doesn't change their integrals), we may assume that $f(x) + g(x)$ is defined everywhere. Since

$$|f(x) + g(x)| \leq |f(x)| + |g(x)|,$$

$f + g$ is integrable. Obviously,

$$f + g = (f_+ - f_-) + (g_+ - g_-) = (f_+ + g_+) - (f_- + g_-)$$

and hence by the lemma above and Proposition 5.5.5(ii)

$$\int (f + g) \, d\mu = \int (f_+ + g_+) \, d\mu - \int (f_- + g_-) \, d\mu =$$

$$= \int f_+ \, d\mu + \int g_+ \, d\mu - \int f_- \, d\mu - \int g_- \, d\mu =$$

$$= \int f_+ \, d\mu - \int f_- \, d\mu + \int g_+ \, d\mu - \int g_- \, d\mu =$$

$$= \int f \, d\mu + \int g \, d\mu$$

To prove (iii), note that $f - g$ is a nonnegative function and hence by (i) and (ii):

$$\int f\,d\mu - \int g\,d\mu = \int f\,d\mu + \int (-1)g\,d\mu = \int (f-g)\,d\mu \geq 0$$

Consequently, $\int f\,d\mu \geq \int g\,d\mu$ and the proposition is proved. □

We can now extend our limit theorems to integrable functions taking both signs. The following result is probably the most useful of all limit theorems for integrals as it is quite strong and at the same time easy to use. It tells us that if a convergent sequence of functions is dominated by an integrable function, then

$$\lim_{n\to\infty} \int f_n\,d\mu = \int \lim_{n\to\infty} f_n\,d\mu$$

**Theorem 5.6.5 (Lebesgue's Dominated Convergence Theorem)** *Assume that $g : X \to \overline{\mathbb{R}}$ is a nonnegative, integrable function and that $\{f_n\}$ is a sequence of measurable functions converging pointwise to $f$. If $|f_n| \leq g$ for all $n$, then*

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

*Proof:* First observe that since $|f| \leq g$, $f$ is integrable. Next note that since $\{g - f_n\}$ and $\{g + f_n\}$ are two sequences of nonnegative measurable functions, Fatou's Lemma gives:

$$\liminf_{n\to\infty} \int (g-f_n)\,d\mu \geq \int \liminf_{n\to\infty}(g-f_n)\,d\mu = \int (g-f)\,d\mu = \int g\,d\mu - \int f\,d\mu$$

and

$$\liminf_{n\to\infty} \int (g+f_n)\,d\mu \geq \int \liminf_{n\to\infty}(g+f_n)\,d\mu = \int (g+f)\,d\mu = \int g\,d\mu + \int f\,d\mu$$

On the other hand,

$$\liminf_{n\to\infty} \int (g-f_n)\,d\mu = \int g\,d\mu - \limsup_{n\to\infty} \int f_n\,d\mu$$

and

$$\liminf_{n\to\infty} \int (g+f_n)\,d\mu = \int g\,d\mu + \liminf_{n\to\infty} \int f_n\,d\mu$$

Combining the two expressions for $\liminf_{n\to\infty} \int (g - f_n)\,d\mu$, we see that

$$\int g\,d\mu - \limsup_{n\to\infty} \int f_n\,d\mu \geq \int g\,d\mu - \int f\,d\mu$$

and hence

$$\limsup_{n\to\infty} \int f_n \, d\mu \leq \int f \, d\mu$$

Combining the two expressions for $\liminf_{n\to\infty} \int (g + f_n) \, d\mu$, we similarly get

$$\liminf_{n\to\infty} \int f_n \, d\mu \geq \int f \, d\mu$$

Hence

$$\limsup_{n\to\infty} \int f_n \, d\mu \leq \int f \, d\mu \leq \liminf_{n\to\infty} \int f_n \, d\mu$$

which means that $\lim_{n\to\infty} \int f_n \, d\mu = \int f \, d\mu$. The theorem is proved. □

**Remark:** It is easy to check that we can relax the conditions above somewhat: If $f_n(x)$ converges to $f(x)$ a.e., and $|f_n(x)| \leq g(x)$ fails on a set of measure zero, the conclusion still holds (see Exercise 7 for the precise statement).

Let us take a look at a typical application of the theorem:

**Proposition 5.6.6** *Let $f : \mathbb{R} \times X \to \mathbb{R}$ be a function which is*

*(i) continuous in the first variable, i.e. for each $y \in X$, the function $x \mapsto f(x, y)$ is continuous*

*(ii) measurable in the second component, i.e. for each $x \in X$, the function $y \mapsto f(x, y)$ is measurable*

*(iii) uniformly bounded by an integrable function, i.e. there is an integrable function $g : \mathbb{R} \to [0, \infty]$ such that $|f(x, y)| \leq g(y)$ for all $x, y \in \mathbb{R}$.*

*Then the function*

$$h(x) = \int f(x, y) \, d\mu(y)$$

*is continuous (the expression $\int f(x, y) \, d\mu(y)$ means that we for each fixed $x$ integrate $f(x, y)$ as a function of $y$).*

*Proof:* According to Proposition 2.2.5 it suffices to prove that if $\{a_n\}$ is a sequence converging to a point $a$, then $h(a_n)$ converges to $h(a)$. Observe that

$$h(a_n) = \int f(a_n, y) \, d\mu(y)$$

and

$$h(a) = \int f(a, y) \, d\mu(y)$$

Observe also that since $f$ is continuous in the first variable, $f(a_n, y) \to f(a, y)$ for all $y$. Hence $\{f(a_n, y)\}$ is a sequence of functions which is dominated by the integrable function $g$ and which converges pointwise to $f(a, y)$. By Lebesgue's Dominated Convergence Theorem,

$$\lim_{n\to\infty} h(a_n) = \lim_{n\to\infty} \int f(a_n, y)\, d\mu = \int f(a, y)\, d\mu = h(a)$$

and the proposition is proved. □

As before, we define $\int_A f\, d\mu = \int f\mathbf{1}_A\, d\mu$ for measurable sets $A$. We say that $f$ *is integrable over* $A$ if $f\mathbf{1}_A$ is integrable.

## Exercises to Section 5.6

1. Show that if $f$ is measurable, so are $f_+$ and $f_-$.
2. Show that if an integrable function $f$ is zero a.e., then $\int f\, d\mu = 0$.
3. Prove Proposition 5.6.4(i). You may want to treat positive and negative $c$'s separately.
4. Assume that $f : X \to \overline{\mathbb{R}}$ is a measurable function.

   a) Show that if $f$ is integrable over a measurable set $A$, and $A_n$ is an increasing sequence of measurable sets with union $A$, then

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

   b) Assume that $\{B_n\}$ is a decreasing sequence of measurable sets with intersection $B$. Show that if $f$ is integrable over $B_1$, then

   $$\lim_{n\to\infty} \int_{B_n} f\, d\mu = \int_B f\, d\mu$$

5. Show that if $f : X \to \mathbb{R}$ is integrable over a measurable set $A$, and $A_n$ is a disjoint sequence of measurable sets with union $A$, then

   $$\int_A f\, d\mu = \sum_{n=1}^{\infty} \int_{A_n} f\, d\mu$$

6. Let $f : \mathbb{R} \to \overline{\mathbb{R}}$ be a measurable function, and define

   $$A_n = \{x \in X \mid f(x) \geq n\}$$

   Show that

   $$\lim_{n\to\infty} \int_{A_n} f\, d\mu = 0$$

7. Prove the following slight extension of the Dominated Convergence Theorem:

   **Theorem:** Assume that $g : X \to \overline{\mathbb{R}}$ is a nonnegative, integrable function and that $\{f_n\}$ is a sequence of measurable functions converging a.e. to $f$. If $|f_n(x)| \leq g(x)$ a.e. for each $n$, then

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

8. Assume that $g : \mathbb{R} \times X \to \mathbb{R}$ is continuous in the first variable and that $y \to g(x, y)$ is integrable for each $x$. Assume also that the partial derivative $\frac{\partial g}{\partial x}(x, y)$ exists for all $x$ and $y$, and that there is an integrable function $h : \mathbb{R} \to [0, \infty]$ such that

$$\left| \frac{\partial g}{\partial x}(x, y) \right| \leq h(y)$$

   for all $x, y$. Show that the function

$$f(x) = \int g(x, y) \, d\mu(y)$$

   is differentiable at all points $x$ and

$$f'(x) = \int \frac{\partial g}{\partial x}(x, y) \, d\mu(y)$$

   This is often referred to as "differentiation under the integral sign".

9. Let $\mu$ be the Lebesgue measure on $\mathbb{R}$. Show that if $a, b \in \mathbb{R}$, $a < b$, and $f : [a, b] \to \mathbb{R}$ is a bounded, Riemann integrable function, then $f$ is integrable over $[a, b]$ and

$$\int_a^b f(x) \, dx = \int_{[a,b]} f \, d\mu$$

   (*Hint:* Since $f$ is bounded, there is a constant $M$ such that $f + M$ is nonnegative. Apply Theorem 5.5.9 to this function.)

## 5.7 $L^1(X, \mathcal{A}, \mu)$ and $L^2(X, \mathcal{A}, \mu)$

In this section we shall connect integration theory to the theory of normed spaces in Chapter 4. Recall from Definition 4.5.2 that a norm on a real vector space $V$ is a function $\| \cdot \| : V \to [0, \infty)$ satisfying

(i) $\|\mathbf{u}\| \geq 0$ with equality if and only if $\mathbf{u} = 0$.

(ii) $\|\alpha \mathbf{u}\| = |\alpha| \|\mathbf{u}\|$ for all $\alpha \in \mathbb{R}$ and all $\mathbf{u} \in V$.

(iii) $\|\mathbf{u} + \mathbf{v}\| \leq \|\mathbf{u}\| + \|\mathbf{v}\|$ for all $\mathbf{u}, \mathbf{v} \in V$.

Let us now put

$$\mathcal{L}^1(X, \mathcal{A}, \mu) = \{f : X \to \overline{\mathbb{R}} \,:\, f \text{ is integrable}\}$$

and define $\|\cdot\|_1 : \mathcal{L}^1(X, \mathcal{A}, \mu) \to [0, \infty)$ by

$$\|f\|_1 = \int |f| \, d\mu$$

It is not hard to see that $\mathcal{L}^1(X)$ is a vector space (see Exercise 1), and that $\|\cdot\|_1$ satisfies the three axioms above with one exception; $\|f\|_1$ may be zero even when $f$ is not zero — actually $\|f\|_1 = 0$ if and only if $f = 0$ a.e.

The usual way to fix this is to consider two functions $f$ and $g$ to be equal if they are equal almost everywhere. To be more precise, let us write $f \sim g$ if $f$ and $g$ are equal a.e., and define the *equivalence class* of $f$ to be the set

$$[f] = \{g \in \mathcal{L}^1(X, \mathcal{A}, \mu) \,|\, g \sim f\}$$

Note that two such equivalence classes $[f]$ and $[g]$ are either equal (if $f$ equals $g$ a.e.) or disjoint (if $f$ is not equal to $g$ a.e.). If we let $L^1(X, \mathcal{A}, \mu)$ be the collection of all equivalence classes, we can organize $L^1(X, \mathcal{A}, \mu)$ as a normed vector space by defining

$$\alpha[f] = [\alpha f] \quad \text{and} \quad [f] + [g] = [f + g] \quad \text{and} \quad |[f]|_1 = \|f\|_1$$

The advantage of the space $(L^1(X), |\cdot|_1)$ compared to $(\mathcal{L}^1(X, \mathcal{A}, \mu), \|\cdot\|_1)$ is that it is a normed space where all the theorems we have proved about such spaces apply — the disadvantage is that the elements are no longer functions, but equivalence classes of functions. In practice, there is very little difference between $(L^1(X), |\cdot|_1)$ and $(\mathcal{L}^1(X, \mathcal{A}, \mu), \|\cdot\|_1)$, and mathematicians tend to blur the distinction between the two spaces: they pretend to work in $L^1(X, \mathcal{A}, \mu)$, but still consider the elements as functions. We shall follow this practice here; it is totally harmless as long as you remember that whenever we talk about an element of $L^1(X, \mathcal{A}, \mu)$ as a function, we are really choosing a representative from an equivalence class (Exercise 3 gives a more thorough and systematic treatment of $L^1(X, \mathcal{A}, \mu)$).

The most important fact about $(L^1(X), |\cdot|_1)$ is that it is complete. In many ways, this is the most impressive success of the theory of measures and integration: We have seen in previous chapters how important completeness is, and it is a great advantage to work with a theory of integration where the space of integrable functions is naturally complete. Before we turn to the proof, you may want to remind yourself of Proposition 4.5.5 which shall be our main tool.

**Theorem 5.7.1** $(L^1(X), |\cdot|_1)$ *is complete.*

*Proof:* Assume that $\{u_n\}$ is a sequence of functions in $L^1(X, \mathcal{A}, \mu)$ such that the series $\sum_{n=1}^{\infty} u_n$ converges absolutely, i.e. that $\sum_{n=1}^{\infty} |u_n|_1 < \infty$. According to Proposition 4.5.5, it suffices to show that the series $\sum_{n=1}^{\infty} u_n(x)$ must converge in $L^1(X, \mathcal{A}, \mu)$.

We first use the absolute convergence to prove that the series $\sum_{n=1}^{\infty} |u_n(x)|$ converges to an integrable function:

$$\int \sum_{n=1}^{\infty} |u_n| \, d\mu = \int \lim_{N\to\infty} \sum_{n=1}^{N} |u_n| \, d\mu = \lim_{N\to\infty} \int \sum_{n=1}^{N} |u_n| \, d\mu$$
$$= \lim_{N\to\infty} \sum_{n=1}^{N} \int |u_n| \, d\mu = \lim_{N\to\infty} \sum_{n=1}^{N} |u_n|_1 = \sum_{n=1}^{\infty} |u_n|_1 < \infty$$

where we have used the Monotone Convergence Theorem to move the limit inside the integral sign. This means that the function

$$g(x) = \sum_{n=1}^{\infty} |u_n(x)|$$

is integrable. We shall use $g$ as the dominating function in the Dominated Convergence Theorem.

Let us first observe that since $g(x) = \sum_{n=1}^{\infty} |u_n(x)|$ is integrable, the series converges a.e. Hence the sequence $\sum_{n=1}^{\infty} u_n(x)$ (without the absolute values) converges absolutely a.e., and hence it converges a.e. in the ordinary sense. Let $f(x) = \sum_{n=1}^{\infty} u_n(x)$ (put $f(x) = 0$ on the null set where the series does not converge). It remains to prove that the series converges to $f$ in $L^1$-sense, i.e. that $|f - \sum_{n=1}^{N} u_n|_1 \to 0$ as $N \to \infty$. By definition of $f$, we know that $\lim_{N\to\infty} \left( f(x) - \sum_{n=1}^{N} u_n(x) \right) = 0$ a.e. Since $|f(x) - \sum_{n=1}^{N} u_n(x)| = |\sum_{n=N+1}^{\infty} u_n(x)| \leq g(x)$ a.e., it follows from Dominated Convergence Theorem (actually from the slight extension in Exercise 5.7.7) that

$$|f - \sum_{n=1}^{N} u_n|_1 = \int |f - \sum_{n=1}^{N} u_n| \, d\mu \to 0$$

The theorem is proved. □

It turns of that $L^1(X, \mathcal{A}, \mu)$ is just one of infinitely many spaces of the same kind. In fact, for any real number $p \geq 1$, we may let

$$\mathcal{L}^p(X, \mathcal{A}, \mu) = \{f : X \to \overline{\mathbb{R}} \, : \, |f|^p \text{ is integrable}\}$$

and define $\| \cdot \|_p : \mathcal{L}^p(X, \mathcal{A}, \mu) \to [0, \infty)$ by

$$\|f\|_p = \left( \int |f|^p \, d\mu \right)^{\frac{1}{p}}$$

It turns out that $\mathcal{L}^p(X, \mathcal{A}, \mu)$ is a vector space, and that $\|\cdot\|_p$ is a norm on $\mathcal{L}^p(X, \mathcal{A}, \mu)$, except that $\|f\|_p = 0$ if $f = 0$ a.e. If we consider functions as equal if they are equal a.e., we can turn $(\mathcal{L}^p(X), \|\cdot\|_p)$ into a normed space $(L^p(X), |\cdot|_p)$ just as we did with $\mathcal{L}^1(X, \mathcal{A}, \mu)$.

We shall not pursue the general theory of $L^p$-spaces here, but we shall take a closer look at the case $p = 2$, i.e. the space

$$\mathcal{L}^2(X, \mathcal{A}, \mu) = \{f : X \to \overline{\mathbb{R}} \,:\, |f|^2 \text{ is integrable}\}$$

with the norm

$$\|f\|_2 = \left(\int |f|^2 \, d\mu\right)^{\frac{1}{2}}$$

This space is particularly important as it turns out to be an inner product space with inner product

$$\langle f, g\rangle = \int fg \, d\mu$$

But let us begin from the beginning. To prove that $L^2(X, \mathcal{A}, \mu)$ is a vector space, we need a simple lemma:

**Lemma 5.7.2** *For all real numbers $a$, $b$*

$$(a + b)^2 \leq 2a^2 + 2b^2$$

*Proof:*

$$2a^2 + 2b^2 - (a + b)^2 = a^2 + b^2 - 2ab = (a - b)^2 \geq 0$$

□

It is now easy to prove that $\mathcal{L}^2(X, \mathcal{A}, \mu)$ is a vector space:

**Proposition 5.7.3** *$\mathcal{L}^2(X, \mathcal{A}, \mu)$ is a vector space, i.e.*

*(i) If $f \in \mathcal{L}^2(X, \mathcal{A}, \mu)$, then $cf \in \mathcal{L}^2(X, \mathcal{A}, \mu)$ for all $c \in \mathbb{R}$.*

*(ii) If $f, g \in \mathcal{L}^2(X, \mathcal{A}, \mu)$, then $f + g \in \mathcal{L}^2(X, \mathcal{A}, \mu)$.*

*Proof:* Part (i) is easy, and part (ii) follows from the lemma since

$$\int (f + g)^2 \, d\mu \leq \int (2f^2 + 2g^2) \, d\mu = 2\int f^2 \, d\mu + 2\int g^2 \, d\mu$$

□

We are now ready to prove that

$$\langle f, g\rangle = \int fg \, d\mu$$

is almost an inner product on $\mathcal{L}^2(X, \mathcal{A}, \mu)$.

**Proposition 5.7.4** *If $f, g \in \mathcal{L}^2(X, \mathcal{A}, \mu)$, then $fg$ is integrable and*

$$\langle f, g \rangle = \int fg \, d\mu$$

*satisfies*

*(i)* $\langle f, g \rangle = \langle g, f \rangle$ *for all* $f, g \in \mathcal{L}^2(X, \mathcal{A}, \mu)$.

*(ii)* $\langle f + g, h \rangle = \langle f, h \rangle + \langle g, h \rangle$ *for all* $f, g, h \in \mathcal{L}^2(X, \mathcal{A}, \mu)$ .

*(iii)* $\langle cf, g \rangle = c\langle f, g \rangle$ *for all* $c \in \mathbb{R}$, $f, g \in \mathcal{L}^2(X, \mathcal{A}, \mu)$.

*(iv) For all* $f \in \mathcal{L}^2(X, \mathcal{A}, \mu)$, $\langle f, f \rangle \geq 0$ *with equality if and only if* $f = 0$ *a.e.*

*Proof:* To see that $fg$ is integrable, note that

$$fg = \frac{1}{2}\left((f+g)^2 - f^2 - g^2\right)$$

and hence

$$\int |fg| \, d\mu \leq \frac{1}{2}\left(\int (f+g)^2 \, d\mu + \int f^2 \, d\mu + \int g^2 \, d\mu\right) < \infty$$

where we have used the previous proposition.

Properties (i)-(iv) are easy consequences of properties we have already proved and are left to the reader. □

Note that $\langle \cdot, \cdot \rangle$ would have been an inner product if instead of (iv) we had had

(iv)' For all $f \in \mathcal{L}^2(X, \mathcal{A}, \mu)$, $\langle f, f \rangle \geq 0$ with equality if and only if $f(x) = 0$ for all $x \in X$.

To turn $\langle \cdot, \cdot \rangle$ into an inner product, we use the same trick as for $\mathcal{L}^2(X, \mathcal{A}, \mu)$: We say that two functions $f, g \in \mathcal{L}^2(X, \mathcal{A}, \mu)$ are equivalent if they are equal a.e., and we let $L^2(X, \mathcal{A}, \mu)$ be the set of all equivalence classes. As before, we let $[f]$ denote the equivalence class of $f$, and define

$$\langle [f], [g] \rangle = \langle f, g \rangle = \int fg \, d\mu$$

for all $[f], [g] \in L^2(X, \mathcal{A}, \mu)$ (you should check that this definition makes sense; i.e. that it is independent of the representatives $f$ and $g$ we pick from the equivalence classes $[f]$ and $[g]$).

It follows from the proposition above and the theory in section 4.6 that $L^2(X, \mathcal{A}, \mu)$ is an inner product space with norm

$$|\,[f]\,|_2 = \langle [f], [f] \rangle^{\frac{1}{2}} = \left(\int f^2 \, d\mu\right)^{\frac{1}{2}}$$

It is usual to blur the distinction between $L^2(X, \mathcal{A}, \mu)$ and $\mathcal{L}^2(X, \mathcal{A}, \mu)$ just as one blurs the distinction between $L^1(X, \mathcal{A}, \mu)$ and $\mathcal{L}^1(X, \mathcal{A}, \mu)$, and we shall follow this tradition and refer to elements in $L^2(X, \mathcal{A}, \mu)$ as if they were functions and not equivalence classes of functions.

We have the same main result for $L^2(X)$ as for $L^1(X)$:

**Theorem 5.7.5** $(L^2(X), |\cdot|_2)$ *is complete.*

*Proof:* This is almost a copy of the proof that $L^1(X)$ is complete. In fact, once it has been proved that all the $L^p$-norms really are norms, the same argument can be used to prove that all $L^p$-spaces, $p \geq 1$, are complete.

We begin by assuming that $\{u_n\}$ is a sequence of functions in $L^2(X)$ such that the series $\sum_{n=1}^{\infty} u_n$ converges absolutely, i.e. that $\sum_{n=1}^{\infty} |u_n|_2 < \infty$. According to Proposition 4.5.5, it suffices to show that the series $\sum_{n=1}^{\infty} u_n(x)$ converges in $L^2(X)$.

Observe first that by the Monotone Convergence Theorem

$$\int \left( \sum_{n=1}^{\infty} |u_n(x)| \right)^2 d\mu = \int \lim_{N\to\infty} \left( \sum_{n=1}^{N} |u_n(x)| \right)^2 d\mu =$$

$$= \lim_{N\to\infty} \int \left( \sum_{n=1}^{N} |u_n(x)| \right)^2 d\mu$$

Taking square roots, we get $|\sum_{n=1}^{\infty} |u_n(x)|\,|_2 = \lim_{N\to\infty} |\sum_{n=1}^{N} |u_n(x)|\,|_2$

The next step is to use this equality and the absolute convergence to prove that the series $\sum_{n=1}^{\infty} |u_n(x)|$ converges to an $L^2$-function:

$$|\sum_{N=1}^{\infty} |u_n(x)|\,|_2 = \lim_{N\to\infty} |\sum_{n=1}^{N} |u_n(x)|\,|_2 \leq$$

$$\leq \lim_{N\to\infty} \sum_{n=1}^{N} |u_n(x)|_2 = \sum_{n=1}^{\infty} |u_n(x)[_2 < \infty$$

This means that the function

$$g(x) = \sum_{n=1}^{\infty} |u_n(x)|$$

is in $L^2(X)$. We shall use $g$ as the dominating function in the Dominated Convergence Theorem.

Let us first observe that since $g(x) = \sum_{n=1}^{\infty} |u_n(x)|$ is in $L^2(X)$, the series converges a.e. Hence the sequence $\sum_{n=1}^{\infty} u_n(x)$ (without the absolute values) converges absolutely a.e., and hence it converges a.e. in the ordinary sense. Let $f(x) = \sum_{n=1}^{\infty} u_n(x)$ (put $f(x) = 0$ on the null set where the series

does not converge). It remains to prove that the series converges to $f$ in $L^2$-sense, i.e. that $|f - \sum_{n=1}^{N} u_n|_2 \to 0$ as $N \to \infty$. By definition of $f$, we know that $\lim_{N\to\infty} \left(f(x) - \sum_{n=1}^{N} u_n(x)\right) = 0$ a.e. Since $|f(x) - \sum_{n=1}^{N} u_n(x)| = |\sum_{n=N+1}^{\infty} u_n(x)| \leq g(x)$ a.e. and $g \in L^2(X)$, it follows from Dominated Convergence Theorem that

$$|f - \sum_{n=1}^{N} u_n|_2 = \left(\int \left(f - \sum_{n=1}^{N} u_n\right)^2 d\mu\right)^{\frac{1}{2}} \to 0$$

The theorem is proved. □

## Exercises for Section 5.7

1. Show that $\mathcal{L}^1(X, \mathcal{A}, \mu)$ is a vector space. Since the set of *all* functions from $X$ to $\mathbb{R}$ is a vector space, it suffices to show that $\mathcal{L}^1(X, \mathcal{A}, \mu)$ is a subspace, i.e. that $cf$ and $f + g$ are in $\mathcal{L}^1(X, \mathcal{A}, \mu)$ whenever $f, g \in \mathcal{L}^1(X, \mathcal{A}, \mu)$ and $c \in \mathbb{R}$.

2. Show that $\|\cdot\|_1$ satisfies the following conditions:

   (i) $\|f\|_1 \geq 0$ for all $f$, and $\|\mathbf{0}\|_1 = 0$ (here $\mathbf{0}$ is the function that is constant 0).

   (ii) $\|cf\|_1 = |c|\|f\|_1$ for all $f \in \mathcal{L}^1(X, \mathcal{A}, \mu)$ and all $c \in \mathbb{R}$.

   (iii) $\|f + g\|_1 \leq \|f\|_1 + \|g\|_1$ for all $f, g \in \mathcal{L}^1(X, \mathcal{A}, \mu)$

   This means that $\|\cdot\|_1$ is a *seminorm*.

3. If $f, g \in \mathcal{L}^1(X, \mathcal{A}, \mu)$, we write $f \sim g$ if $f = g$ a.e. Recall that the equivalence class $[f]$ of $f$ is defined by

$$[f] = \{g \in \mathcal{L}(X) \,:\, g \sim f\}$$

   a) Show that two equivalence classes $[f]$ and $[g]$ are either equal or disjoint.

   b) Show that if $f \sim f'$ and $g \sim g'$, then $f + g \sim f' + g'$. Show also that $cf \sim cf'$ for all $c \in \mathbb{R}$.

   c) Show that if $f \sim g$, then $\|f - g\|_1 = 0$ and $\|f\|_1 = \|g\|_1$.

   d) Show that the set $L^1(X, \mathcal{A}, \mu)$ of all equivalence classes is a normed space if we define scalar multiplication, addition and norm by:

      (i) $c[f] = [cf]$ for all $c \in \mathbb{R}$, $f \in \mathcal{L}^1(X, \mathcal{A}, \mu)$.

      (ii) $[f] + [g] = [f + g]$ for all $f, g \in \mathcal{L}^1(X, \mathcal{A}, \mu)$

      (iii) $|[f]|_1 = \|f\|_1$ for all $f \in \mathcal{L}^1(X, \mathcal{A}, \mu)$.

      Why do we need to establish the results in (i), (ii), and (iii) before we can make these definitions?

4. Let $X = \{1, 2, 3, \ldots, d\}$, let $\mathcal{A}$ be the collection of all subsets of $X$, and let $\mu$ be the counting measure, i.e. $\mu(\{i\}) = 1$ for all $i$. Show that $|f|_2 = \sum_{i=1}^{d} f(i)^2$, and explain that $L^2(X, \mathcal{A}, \mu)$ is essentially the same as $\mathbb{R}^d$ with the usual metric.

5. Let $X = \mathbb{N}$, let $\mathcal{A}$ be the collection of all subsets of $X$, and let $\mu$ be the counting measure, i.e. $\mu(\{i\}) = 1$ for all $i$. Show that $L^1(X, \mathcal{A}, \mu)$ consists of all functions $f$ such that the series $\sum_{n=1}^{\infty} f(n)$ converges absolutely. Show also that $|f|_1 = \sum_{n=1} |f(n)|$. Give a similar description of $L^2(X, \mathcal{A}, \mu)$ and $|f|_2$.

6. Prove (i)-(iv) in Proposition 5.7.4.

7. In this problem $(X, \mathcal{A}, \mu)$ is a *finite* measure space (i.e. $\mu(X) < \infty$) and all functions are measurable functions from $X$ til $\overline{\mathbb{R}}$. We shall use the abbreviated notation
$$\{f > M\} = \{x \in X \,:\, f(x) > M\}$$

   a) Assume that $f$ is nonnegative. Show that $f$ is integrable if and only if there is a number $M \in \mathbb{R}$ such that
$$\int_{\{f>M\}} f \, d\mu < \infty$$

   b) Assume that $f$ is nonnegative and integrable. Show that
$$\lim_{M\to\infty} \int_{\{f>M\}} f \, d\mu = 0$$

   c) Assume that $\{f_n\}$ is a sequence of nonnegative, integrable functions converging pointwise to $f$. Let $M \in \mathbb{R}$. Show that
$$\liminf_{n\to\infty} \mathbf{1}_{\{f_n>M\}} f_n(x) \geq \mathbf{1}_{\{f>M\}} f(x)$$

   d) Let $\{f_n\}$, $f$ og $M$ be as above. Show that if
$$\int_{\{f_n>M\}} f_n(x) \, d\mu \leq \alpha$$
for all $n$, then
$$\int_{\{f>M\}} f(x) \, d\mu \leq \alpha$$

A squence $\{f_n\}$ of nonnegative functions is called *uniformly integrable* if
$$\lim_{M\to\infty} \left( \sup_{n\in\mathbb{N}} \int_{\{f_n>M\}} f_n \, d\mu \right) = 0$$
(compare this to part b)).

   e) Assume that $\{f_n\}$ is a uniformly integrable sequence of nonnegative functions converging pointwise to $f$. Show that $f$ is integrable.

   f) Let $\{f_n\}$ and $f$ be as in part e). Show that $\{f_n\}$ converges to $f$ i $L^1$-norm, i.e.,
$$\|f - f_n\|_{L^1(\mu)} = \int |f - f_n| \, d\mu \to 0 \quad \text{når } n \to \infty$$