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 that 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 cannot in general expect to measure the size of *all* subsets of X – some subsets are just so irregular that we cannot 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.

7.1. 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 σ-algebra. More succinctly:

Definition 7.1.1. *Assume that X is a nonempty set. A collection $\mathcal{A}$ of subsets of X is called a σ-*algebra *if the following conditions are satisfied:*

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

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

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

The sets in $\mathcal{A}$ are called measurable *if it is clear which σ-algebra we have in mind, and $\mathcal{A}$-measurable if the σ-algebra needs to be specified. If $\mathcal{A}$ is a σ-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 7.1.2. *Assume that $\mathcal{A}$ is a σ-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 ∞ to indicate that sets have infinite size. This does not mean that we think of ∞ as a number; it is just a symbol to indicate that something has size bigger than can be specified by a number.

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

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

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

Definition 7.1.3. *Assume that $(X, \mathcal{A})$ is a measurable space. A* measure *on $(X, \mathcal{A})$ is a function $\mu\colon \mathcal{A} \to \overline{\mathbb{R}}_+$ 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*

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

(We treat infinite terms in the obvious way: If some of the terms $\mu(A_n)$ in the sum equal ∞, then the sum itself also equals ∞.)

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 $A \subseteq X$, let

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

Then μ 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 $A \subseteq X$, 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 $A \subseteq X$, let

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

where we put $\mu(A) = \infty$ if A has infinitely many elements. Again μ 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 μ is called the *point measure* or *Dirac 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 cannot 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 σ-algebra containing all open and closed sets (we shall describe it more precisely later), and μ 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 in the next chapter is 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 σ-algebra containing all open and closed sets, and μ 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 probability 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 μ be given by

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

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

In probability theory it is usual to call the underlying space Ω (instead of X) and the measure P (instead of μ), 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 ω_i is either H or T. We let Ω be the collection of all such sequences.

To describe the σ-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 Ω, a σ-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 7.1.3 and derive some simple, but extremely useful consequences. Note that all these properties are properties we would expect of a measure.

Proposition 7.1.4. *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 $B \subseteq A$, then $\mu(B) \leq \mu(A)$.*

c) *If $A, B \in \mathcal{A}$, $B \subseteq A$, 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 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 7.1.3.

b) We write $A = B \cup (A \setminus B)$. By Proposition 7.1.2d), $A \setminus B \in \mathcal{A}$, 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:

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

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 7.1.3 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 7.1.5 (Continuity of Measure)**.** *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., $A_n \subseteq A_{n+1}$ 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., $A_n \supseteq A_{n+1}$ 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 μ 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. $\square$

Remark: The finiteness condition in part b) 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 7.1

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 μ and ν are two measures on the same measurable space $(X, \mathcal{A})$, then for all positive numbers $\alpha, \beta \in \mathbb{R}$, the function $\lambda \colon \mathcal{A} \to \mathbb{R}_+$ 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 $\mu_A \colon \mathcal{A} \to \overline{\mathbb{R}}_+$ by
$$\mu_A(B) = \mu(A \cap B) \qquad \text{for all } B \in \mathcal{A}.$$
Show that μ_A is a measure.
9. Let X be an uncountable set, and define
$$\mathcal{A} = \{A \subseteq X \mid A \text{ or } A^c \text{ is countable}\}.$$
Show that $\mathcal{A}$ is a σ-algebra. Define $\mu \colon \mathcal{A} \to \mathbb{R}_+$ by
$$\mu(A) = \begin{cases} 0 & \text{if } A \text{ is countable} \\ \\ 1 & \text{if } A^c \text{ is countable.} \end{cases}$$
Show that μ is a measure.

10. Assume that $(X, \mathcal{A})$ is a measurable space, and let $f\colon X \to Y$ be any function from X to a set Y. Show that
$$\mathcal{B} = \{B \subseteq Y \mid f^{-1}(B) \in \mathcal{A}\}$$
is a σ-algebra.
11. Assume that $(X, \mathcal{A})$ is a measurable space, and let $f\colon Y \to X$ be any function from a set Y to X. Show that
$$\mathcal{B} = \{f^{-1}(A) \mid A \in \mathcal{A}\}$$
is a σ-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 σ-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 μ 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 sets 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 μ 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 μ 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 nonempty 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
   $$\mathcal{A} = \{A \subseteq \mathbb{N} \mid A \text{ or } A^c \text{ is finite}\}.$$
   Show that $\mathcal{A}$ is an algebra, but not a σ-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 σ-algebra.
18. Let X be a nonempty set. We look at a family $\mathcal{D}$ of subsets of X satisfying the following conditions:
   (i) $\emptyset \in \mathcal{D}$.
   (ii) If $A \in \mathcal{D}$, then $A^c \in \mathcal{D}$.
   (iii) If $\{B_n\}$ is a pairwise disjoint sequence of sets in $\mathcal{D}$ (i.e., $B_i \cap B_j = \emptyset$ for $i \neq j$), then $\bigcup_{n\in\mathbb{N}} B_n \in \mathcal{D}$.

   Such a family $\mathcal{D}$ is called a *Dynkin system*.
   a) Show that for all sets $A, B \subseteq X$, we have $A \setminus B = (A^c \cup B)^c$.
   b) Show that if $A, B \in \mathcal{D}$ and $B \subseteq A$, then $A \setminus B \in \mathcal{D}$. (*Hint:* You may find part a) helpful.)

c) Show that if $\{A_n\}$ is an increasing sequence of sets in $\mathcal{D}$ (i.e., $A_n \subseteq A_{n+1}$ for all $n \in \mathbb{N}$), then $\bigcup_{n\in\mathbb{N}} A_n \in \mathcal{D}$.

19. 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

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

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

7.2. 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 $N \subseteq A$, 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 (see Exercise 1). This is often a nuisance, and we shall now see how it can be cured.

First, some definitions:

Definition 7.2.1. *Assume that $(X, \mathcal{A}, \mu)$ is a measure space. A set $N \subseteq X$ is called a* null set *if $N \subseteq A$ 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 7.1.4b).

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 $\mathcal{A} \subseteq \bar{\mathcal{A}}$ and $\bar{\mu}(A) = \mu(A)$ for all $A \in \mathcal{A}$).

We begin with a simple observation:

Lemma 7.2.2. *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 $N_n \subseteq A_n$. Since $\bigcup_{n\in\mathbb{N}} N_n \subseteq \bigcup_{n\in\mathbb{N}} A_n$ and

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

by Proposition 7.1.4d), $\bigcup_{n\in\mathbb{N}} N_n$ is a null set. □

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

Lemma 7.2.3. *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 σ-algebra containing $\mathcal{A}$ and $\mathcal{N}$ (in the sense that if $\mathcal{B}$ is any other σ-algebra containing $\mathcal{A}$ and $\mathcal{N}$, then $\bar{\mathcal{A}} \subseteq \mathcal{B}$).

Proof. If we can only prove that $\bar{\mathcal{A}}$ is a σ-algebra, the rest will be easy: Any σ-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 $\mathcal{A} \subseteq \bar{\mathcal{A}}$ and $\mathcal{N} \subseteq \bar{\mathcal{A}}$.

To prove that $\bar{\mathcal{A}}$ is a σ-algebra, we need to check the three conditions in Definition 7.1.1. 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 preceding 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 $N \subseteq B$ 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 7.2.1), 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. $\square$

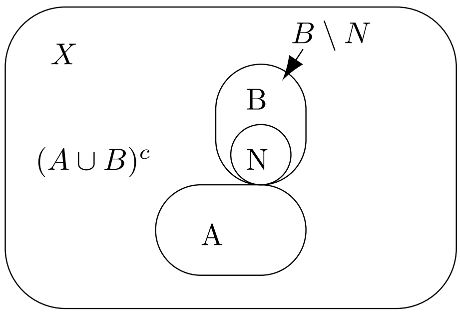

Figure 7.2.1. $(A \cup N)^c = (A \cup B)^c \cup (B \setminus N)$

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

Lemma 7.2.4. *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 $N_2 \subseteq B_2$ and $\mu(B_2) = 0$. Then $A_1 \subseteq A_1 \cup N_1 = A_2 \cup N_2 \subseteq A_2 \cup B_2$, and hence

$$\mu(A_1) \le \mu(A_1 \cup B_2) \le \mu(A_2) + \mu(B_2) = \mu(A_2).$$

Interchanging the roles of A_1 and A_2, we get the opposite inequality $\mu(A_2) \le \mu(A_1)$, and hence we must have $\mu(A_1) = \mu(A_2)$. $\square$

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 7.2.5. *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 $\bar{\mu}\colon \bar{\mathcal{A}} \to \overline{\mathbb{R}}_+$ 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 μ, i.e., $\bar{\mu}(A) = \mu(A)$ for all $A \in \mathcal{A}$.

Proof. We already know that $\bar{\mathcal{A}}$ is a σ-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 $C \subseteq D$, 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 $N \subseteq B$ and $\mu(B) = 0$. But then $C \subseteq A \cup B$, where $\mu(A \cup B) = 0$, and hence C is in $\mathcal{N}$ and hence in $\bar{\mathcal{A}}$. □

In Lemma 7.2.3 we proved that $\bar{\mathcal{A}}$ is the smallest σ-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 σ-algebra $\mathcal{A}$ containing $\mathcal{B}$. It is called the *σ-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 7.2.6. *Let X be a nonempty set and $\mathcal{B}$ a collection of subsets of X. Then there exists a smallest σ-algebra $\sigma(\mathcal{B})$ containing $\mathcal{B}$ (in the sense that if $\mathcal{C}$ is any other σ-algebra containing $\mathcal{B}$, then $\sigma(\mathcal{B}) \subseteq \mathcal{C}$).*

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

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

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

We must check the three conditions in Definition 7.1.1. For (i), just observe that since $\emptyset$ belongs to all σ-algebras, it belongs to $\sigma(\mathcal{B})$. For (ii), observe that if $A \in \sigma(\mathcal{B})$, then A belongs to all σ-algebras containing $\mathcal{B}$. Since σ-algebras are closed under complements, A^c belongs to the same σ-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 σ-algebras containing $\mathcal{B}$, and since σ-algebras are closed under countable unions, the union $\bigcup_{n\in\mathbb{N}} A_n$ belongs to the same σ-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 σ-algebra $\sigma(\mathcal{G})$ generated by the collection $\mathcal{G}$ of open sets is called the *Borel σ-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 description 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 σ-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 $a_1 < b_1$, $a_2 < b_2, \ldots,$ $a_d < b_d$ (i.e., μ 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 7.1. In this setting one can prove that there is a unique measure on the σ-algebra $\sigma(\mathcal{C})$ generated by the cylinder sets such that $P(\mathcal{C}_{\mathbf{a}}) = \frac{1}{2^n}$ for all cylinder sets of length n. The completion of this measure is the one used to model coin tossing. We shall carry out this construction in Section 8.6.

Exercises to Section 7.2.

1. Let $X = \{0, 1, 2\}$ and let $\mathcal{A} = \{\emptyset, \{0, 1\}, \{2\}, X\}$.
   a) Show that $\mathcal{A}$ is a σ-algebra.
   b) Define $\mu \colon \mathcal{A} \to \mathbb{R}_+$ by: $\mu(\emptyset) = \mu(\{0, 1\}) = 0$, $\mu(\{2\}) = \mu(X) = 1$. Show that μ is a measure.
   c) Show that μ is *not* complete, and describe the completion $(X, \bar{\mathcal{A}}, \bar{\mu})$ of $(X, \mathcal{A}, \mu)$.
2. Redo Problem 1 for $X = \{0, 1, 2, 3\}$, $\mathcal{A} = \{\emptyset, \{0, 1\}, \{2, 3\}, X\}$, and $\mu(\emptyset) = \mu(\{0, 1\}) = 0$, $\mu(\{2, 3\}) = \mu(X) = 1$.
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 $A \subseteq C \subseteq B$, 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 nonempty 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 nonempty 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 $\mathcal{B}$.

8. Let X be a nonempty set. A family $\mathcal{D}$ of subsets of X is called a *Dynkin system* if it satisfies the following conditions:
   (i) $\emptyset \in \mathcal{D}$.
   (ii) If $A \in \mathcal{D}$, then $\bar{A} \in \mathcal{D}$.
   (iii) If $\{B_n\}$ is a pairwise disjoint sequence of sets in $\mathcal{D}$ (i.e., $B_i \cap B_j = \emptyset$ for $i \neq j$), then $\bigcup_{n\in\mathbb{N}} B_n \in \mathcal{D}$.

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

7.3. 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 ∞ 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

$$f\colon X \to \overline{\mathbb{R}},$$

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 ∞ "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 $(-\infty) \cdot r$. We also have a natural ordering of $\overline{\mathbb{R}}$: If $a \in \mathbb{R}$, then

$$-\infty < a < \infty.$$

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 7.3.1. *Let $(X, \mathcal{A}, \mu)$ be a measure space. A function $f\colon X \to \overline{\mathbb{R}}$ 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 7.3.2. *Any nonempty, 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 7.3.3. *If $f\colon X \to \overline{\mathbb{R}}$ 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 $A \subseteq \mathbb{R}$.*

Proof. We use that inverse images commute with intersections, unions, and complements (see Section 1.4) . First observe that for any $r \in \mathbb{R}$,

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

where we have used that each set $f^{-1}\big([-\infty, r + \frac{1}{n})\big)$ is in $\mathcal{A}$ by definition, and that $\mathcal{A}$ is closed under countable intersections. This 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, to prove the proposition for closed sets A, we are going to use that the complement (in $\mathbb{R}$) of a closed set is an open set. We have to be a little careful, however, as complements in $\mathbb{R}$ are not the same as complements in $\overline{\mathbb{R}}$. Note that if $O = \mathbb{R} \setminus A$ is the complement of A in $\mathbb{R}$, then O is open, and $A = O^c \cap \mathbb{R}$, where O^c is the complement of O in $\overline{\mathbb{R}}$. Hence

$$f^{-1}(A) = f^{-1}(O^c \cap \mathbb{R}) = f^{-1}(O)^c \cap f^{-1}(\mathbb{R}) \in \mathcal{A}. \qquad \square$$

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

Proposition 7.3.4. *Let $(X, \mathcal{A}, \mu)$ be a measure space and consider a function $f\colon X \to \overline{\mathbb{R}}$. 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. $\square$

The next result tells us that there are many measurable functions. Recall that a Borel measure is a measure defined on the σ-algebra generated by the open sets.

Proposition 7.3.5. *Let (X, d) be a metric space and let μ be a Borel or a completed Borel measure on X. Then all continuous functions $f\colon X \to \mathbb{R}$ 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 3.3.10 and measurable since the Borel σ-algebra is generated by the open sets. $\square$

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 7.3.6. *Let $(X, \mathcal{A}, \mu)$ be a measure space. If $f\colon X \to \mathbb{R}$ is measurable, then $\phi \circ f$ is measurable for all continuous functions $\phi\colon \mathbb{R} \to \mathbb{R}$. 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 ϕ is continuous, $\phi^{-1}((-\infty, r))$ is open, and consequently $f^{-1}(\phi^{-1}((-\infty, r)))$ is measurable by Proposition 7.3.3. To see that f^2 is measurable, apply the first part of the theorem to the function $\phi(x) = x^2$. $\square$

Proposition 7.3.7. *Let $(X, \mathcal{A}, \mu)$ be a measure space. If the functions $f, g\colon X \to \mathbb{R}$ 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

$$\begin{aligned}(f+g)^{-1}([-\infty,r)) &= \{x \in X \mid (f+g)<r) \\ &= \bigcup_{q\in\mathbb{Q}} \left(\{x\in X \mid f(x)<q\} \cap \{x \in X \mid g<r-q\}\right),\end{aligned}$$

which is measurable since $\mathbb{Q}$ is countable, and a countable 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 7.3.6 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 7.3.8. *Let $(X,\mathcal{A},\mu)$ be a measure space. We say that a measurable function $f\colon X\to\overline{\mathbb{R}}$ is* finite almost everywhere *if the set $\{x\in X : f(x)=\pm\infty\}$ has measure zero. We say that two measurable functions $f,g\colon X\to\overline{\mathbb{R}}$ 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.[1]

Proposition 7.3.9. *Let $(X,\mathcal{A},\mu)$ be a measure space. If $\{f_n\}$ is a sequence of measurable functions $f_n\colon X\to\overline{\mathbb{R}}$, 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 7.3.4(iii). For any $r\in\mathbb{R}$,

$$\begin{aligned}f^{-1}((r,\infty]) &= \{x\in X : \sup_{n\in\mathbb{N}} f_n(x) > r\} \\ &= \bigcup_{n\in\mathbb{N}} \{x\in X : f_n(x)>r\} = \bigcup_{n\in\mathbb{N}} f_n^{-1}((r,\infty]) \in \mathcal{A},\end{aligned}$$

and hence f is measurable by Proposition 7.3.4(iii). A similar argument can be used for $\inf_{n\in\mathbb{N}} f_n(x)$.

[1] If you are unfamiliar with the notions of lim inf and lim sup, take a look at Section 2.2.

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

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

(for the last equality, use that the sequence $g_k(x)$ is decreasing) the measurability of $\limsup_{n\to\infty} f_n(x)$ follows. A completely similar proof can be used to prove that $\liminf_{n\to\infty} f_n(x)$ is measurable. Finally, if the sequence converges pointwise, then $\lim_{n\to\infty} f_n(x) = \limsup_{n\to\infty} f_n(x)$ and is hence measurable. □

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

Exercises for Section 7.3.

1. Show that if $f\colon X \to \overline{\mathbb{R}}$ is measurable, the sets $f^{-1}(\{\infty\})$ and $f^{-1}(\{-\infty\})$ are measurable.
2. Complete the proof of Proposition 7.3.3 by showing that f^{-1} of the intervals $(-\infty, r)$, $(-\infty, r]$, $[r, \infty)$, (r, ∞), $(-\infty, \infty)$, where $r \in \mathbb{R}$, are measurable.
3. Prove Proposition 7.3.4.
4. Fill in the details in the proof of Lemma 7.3.2. 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 $f_1, f_2, \dots, f_n$ 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 $A \subseteq X$ 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 $f\colon X \to \mathbb{R}$ of the form

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

   where $a_1, a_2, \dots, a_n \in \mathbb{R}$ and $A_1, A_2, \dots, A_n \in \mathcal{A}$. Show that all simple functions are measurable.
7. Show that if $f\colon X \to \mathbb{R}$ is measurable, then $f^{-1}(B) \in \mathcal{A}$ for all Borel sets B (it may help to take a look at Exercise 7.1.10).
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 7.3.7. You may want to prove first that if $h\colon X \to \mathbb{R}$ is measurable, then so is $\frac{h}{2}$.
10. Prove the inf- and the lim inf-part of Proposition 7.3.9.

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 $f\colon X \to \overline{\mathbb{R}}$ is measurable and $g\colon X \to \overline{\mathbb{R}}$ 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 7.2.1.
13. Assume that the measurable function $f\colon X \to \overline{\mathbb{R}}$ is finite a.e. Define a new function $f'\colon X \to \mathbb{R}$ 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 $f\colon X \to \mathbb{R}$. Show that there is a smallest σ-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 *σ-algebra generated by $\mathcal{F}$*). Show that if X is a metric space and all the functions in $\mathcal{F}$ are continuous, then $\mathcal{A} \subseteq \mathcal{B}$, where $\mathcal{B}$ is the Borel σ-algebra.

7.4. 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 μ is a measure, and the theory is at the same time a *refinement* and a *generalization* of the classical theory of Riemann integration that you know from calculus.

It is a *refinement* because if we choose μ 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 μ 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 7.3.9, $\lim_{n\to\infty} f_n$ 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 measurable functions; and in Section 7.6 we shall extend the theory to functions taking both positive and negative values.

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

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

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 7.4.1. *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 μ *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 7.4.2. *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*

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

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 $C_i = B_{i_1} \cup B_{i_2} \cup \ldots \cup B_{i_{n_i}}$ (make a drawing!). 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

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

and the lemma is proved. □

The next step is also easy:

Proposition 7.4.3. *Assume that* f *and* g *are two nonnegative simple functions, and let* c *be a nonnegative, real number. Then*

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

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

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 7.4.1 holds for all representations of simple functions, and not only the standard ones:

Corollary 7.4.4. *If $f(x) = \sum_{i=1} a_i \mathbf{1}_{A_i}(x)$ 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

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

which proves the result. □

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 7.4.5. *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. Since f, g, and $f - g$ are nonnegative simple functions, we have

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

by Proposition 7.4.3(ii). □

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 7.4.6. *Assume that* B *is a measurable set,* b *a nonnegative 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. Observe first that we may assume that $b > 0$ and $\mu(B) > 0$ as otherwise the conclusion obviously holds. 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 7.1.5a)), $\mu(B) = \lim_{n\to\infty} \mu(A_n)$, 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 7.4.7. *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 just drop that term in the sum, so that we from now on assume that all the b_i's are nonzero. By Proposition 7.4.3(ii), we have

$$\begin{aligned}\int_{B_1\cup B_2\cup\ldots\cup B_m} f_n\,d\mu &= \int (\mathbf{1}_{B_1}+\mathbf{1}_{B_2}+\ldots+\mathbf{1}_{B_m})\,f_n\,d\mu \\ &= \int_{B_1} f_n\,d\mu + \int_{B_2} f_n\,d\mu + \ldots + \int_{B_m} f_n\,d\mu = \sum_{i=1}^m \int_{B_i} f_n\,d\mu.\end{aligned}$$

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

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

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 7.4.

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 7.4.1, $\int \mathbf{1}_A\,d\mu = \mu(A)$ for all $A\in\mathcal{A}$.
3. Prove part (i) of Proposition 7.4.3.
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 μ 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 ν on $(X,\mathcal{A})$.

7.5. Integrals of nonnegative functions

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

Definition 7.5.1. *If $f\colon X\to\overline{\mathbb{R}}_+$ is measurable, we define*

$$\int f\,d\mu = \sup\left\{\int g\,d\mu \;\middle|\; g \text{ is a nonnegative simple function, } g\leq f\right\}.$$