formation of finite intersections (use that fact that $\cap_{i=1}^{n} A_i = (\cup_{i=1}^{n} A_i^c)^c$). Thus we could have defined an algebra using only conditions (a), (b), and (c). A similar argument shows that we could have used only conditions (a), (b), and (d).

Again let X be an arbitrary set. A collection $\mathscr{A}$ of subsets of X is a *σ-algebra*[1] on X if

(a) $X \in \mathscr{A}$,
(b) for each set A that belongs to $\mathscr{A}$, the set A^c belongs to $\mathscr{A}$,
(c) for each infinite sequence $\{A_i\}$ of sets that belong to $\mathscr{A}$, the set $\cup_{i=1}^{\infty} A_i$ belongs to $\mathscr{A}$, and
(d) for each infinite sequence $\{A_i\}$ of sets that belong to $\mathscr{A}$, the set $\cap_{i=1}^{\infty} A_i$ belongs to $\mathscr{A}$.

Thus a σ-algebra on X is a family of subsets of X that contains X and is closed under complementation, under the formation of countable unions, and under the formation of countable intersections. Note that, as in the case of algebras, we could have used only conditions (a), (b), and (c), or only conditions (a), (b), and (d), in our definition.

Each σ-algebra on X is an algebra on X since, for example, the union of the finite sequence $A_1, A_2, \ldots, A_n$ is the same as the union of the infinite sequence $A_1, A_2, \ldots, A_n, A_n, A_n, \ldots$.

If X is a set and $\mathscr{A}$ is a family of subsets of X that is closed under complementation, then X belongs to $\mathscr{A}$ if and only if $\varnothing$ belongs to $\mathscr{A}$. Thus in the definitions of algebras and σ-algebras given above, we can replace condition (a) with the requirement that $\varnothing$ be a member of $\mathscr{A}$. Furthermore, if $\mathscr{A}$ is a family of subsets of X that is nonempty, closed under complementation, and closed under the formation of finite or countable unions, then $\mathscr{A}$ must contain X: if the set A belongs to $\mathscr{A}$, then X, since it is the union of A and A^c, must also belong to $\mathscr{A}$. Thus in our definitions of algebras and σ-algebras, we can replace condition (a) with the requirement that $\mathscr{A}$ be nonempty.

If $\mathscr{A}$ is a σ-algebra on the set X, it is sometimes convenient to call a subset of X *$\mathscr{A}$-measurable* if it belongs to $\mathscr{A}$.

Examples 1.1.1 (Some Families of Sets That Are Algebras or σ-algebras, and Some That Are Not).

(a) Let X be a set, and let $\mathscr{A}$ be the collection of all subsets of X. Then $\mathscr{A}$ is a σ-algebra on X.
(b) Let X be a set, and let $\mathscr{A} = \{\varnothing, X\}$. Then $\mathscr{A}$ is a σ-algebra on X.
(c) Let X be an infinite set, and let $\mathscr{A}$ be the collection of all finite subsets of X. Then $\mathscr{A}$ does not contain X and is not closed under complementation; hence it is not an algebra (or a σ-algebra) on X.

---

[1] The terms *field* and *σ-field* are sometimes used in place of algebra and σ-algebra.

(d) Let X be an infinite set, and let $\mathscr{A}$ be the collection of all subsets A of X such that either A or A^c is finite. Then $\mathscr{A}$ is an algebra on X (check this) but is not closed under the formation of countable unions; hence it is not a σ-algebra.
(e) Let X be an uncountable set, and let $\mathscr{A}$ be the collection of all countable (i.e., finite or countably infinite) subsets of X. Then $\mathscr{A}$ does not contain X and is not closed under complementation; hence it is not an algebra.
(f) Let X be a set, and let $\mathscr{A}$ be the collection of all subsets A of X such that either A or A^c is countable. Then $\mathscr{A}$ is a σ-algebra.
(g) Let $\mathscr{A}$ be the collection of all subsets of $\mathbb{R}$ that are unions of finitely many intervals of the form $(a,b]$, $(a,+\infty)$, or $(-\infty,b]$. It is easy to check that each set that belongs to $\mathscr{A}$ is the union of a finite disjoint collection of intervals of the types listed above, and then to check that $\mathscr{A}$ is an algebra on $\mathbb{R}$ (the empty set belongs to $\mathscr{A}$, since it is the union of the empty, and hence finite, collection of intervals). The algebra $\mathscr{A}$ is not a σ-algebra; for example, the bounded open subintervals of $\mathbb{R}$ are unions of sequences of sets in $\mathscr{A}$ but do not themselves belong to $\mathscr{A}$. □

Next we consider ways of constructing σ-algebras.

Proposition 1.1.2. *Let X be a set. Then the intersection of an arbitrary nonempty collection of σ-algebras on X is a σ-algebra on X.*

Proof. Let $\mathscr{C}$ be a nonempty collection of σ-algebras on X, and let $\mathscr{A}$ be the intersection of the σ-algebras that belong to $\mathscr{C}$. It is enough to check that $\mathscr{A}$ contains X, is closed under complementation, and is closed under the formation of countable unions. The set X belongs to $\mathscr{A}$, since it belongs to each σ-algebra that belongs to $\mathscr{C}$. Now suppose that $A \in \mathscr{A}$. Each σ-algebra that belongs to $\mathscr{C}$ contains A and so contains A^c; thus A^c belongs to the intersection $\mathscr{A}$ of these σ-algebras. Finally, suppose that $\{A_i\}$ is a sequence of sets that belong to $\mathscr{A}$ and hence to each σ-algebra in $\mathscr{C}$. Then $\cup_i A_i$ belongs to each σ-algebra in $\mathscr{C}$ and so to $\mathscr{A}$. □

The reader should note that the union of a family of σ-algebras can fail to be a σ-algebra (see Exercise 5).

Proposition 1.1.2 implies the following result, which is a basic tool for the construction of σ-algebras.

Corollary 1.1.3. *Let X be a set, and let $\mathscr{F}$ be a family of subsets of X. Then there is a smallest σ-algebra on X that includes $\mathscr{F}$.*

Of course, to say that $\mathscr{A}$ is the smallest σ-algebra on X that includes $\mathscr{F}$ is to say that $\mathscr{A}$ is a σ-algebra on X that includes $\mathscr{F}$ and that every σ-algebra on X that includes $\mathscr{F}$ also includes $\mathscr{A}$. If $\mathscr{A}_1$ and $\mathscr{A}_2$ are both smallest σ-algebras that include $\mathscr{F}$, then $\mathscr{A}_1 \subseteq \mathscr{A}_2$ and $\mathscr{A}_2 \subseteq \mathscr{A}_1$, and so $\mathscr{A}_1 = \mathscr{A}_2$; thus the smallest σ-algebra on X that includes $\mathscr{F}$ is unique. The smallest σ-algebra is called the σ-algebra *generated* by $\mathscr{F}$ and is often denoted by $\sigma(\mathscr{F})$.

Proof. Let $\mathscr{C}$ be the collection of all σ-algebras on X that include $\mathscr{F}$. Then $\mathscr{C}$ is nonempty, since it contains the σ-algebra that consists of all subsets of

X. The intersection of the σ-algebras that belong to $\mathscr{C}$ is, according to Proposition 1.1.2, a σ-algebra; it includes $\mathscr{F}$ and is included in every σ-algebra in $\mathscr{C}$—that is, it is included in every σ-algebra on X that includes $\mathscr{F}$. □

We now use the preceding corollary to define an important family of σ-algebras. The *Borel σ-algebra* on $\mathbb{R}^d$ is the σ-algebra on $\mathbb{R}^d$ generated by the collection of open subsets of $\mathbb{R}^d$; it is denoted by $\mathscr{B}(\mathbb{R}^d)$. The *Borel subsets* of $\mathbb{R}^d$ are those that belong to $\mathscr{B}(\mathbb{R}^d)$. In case $d = 1$, one generally writes $\mathscr{B}(\mathbb{R})$ in place of $\mathscr{B}(\mathbb{R}^1)$.

Proposition 1.1.4. *The σ-algebra $\mathscr{B}(\mathbb{R})$ of Borel subsets of $\mathbb{R}$ is generated by each of the following collections of sets:*

(a) *the collection of all closed subsets of $\mathbb{R}$;*
(b) *the collection of all subintervals of $\mathbb{R}$ of the form $(-\infty, b]$;*
(c) *the collection of all subintervals of $\mathbb{R}$ of the form $(a, b]$.*

Proof. Let $\mathscr{B}_1$, $\mathscr{B}_2$, and $\mathscr{B}_3$ be the σ-algebras generated by the collections of sets in parts (a), (b), and (c) of the proposition. We will show that $\mathscr{B}(\mathbb{R}) \supseteq \mathscr{B}_1 \supseteq \mathscr{B}_2 \supseteq \mathscr{B}_3$ and then that $\mathscr{B}_3 \supseteq \mathscr{B}(\mathbb{R})$; this will establish the proposition. Since $\mathscr{B}(\mathbb{R})$ includes the family of open subsets of $\mathbb{R}$ and is closed under complementation, it includes the family of closed subsets of $\mathbb{R}$; thus it includes the σ-algebra generated by the closed subsets of $\mathbb{R}$, namely $\mathscr{B}_1$. The sets of the form $(-\infty, b]$ are closed and so belong to $\mathscr{B}_1$; consequently $\mathscr{B}_1 \supseteq \mathscr{B}_2$. Since $(a, b] = (-\infty, b] \cap (-\infty, a]^c$, each set of the form $(a, b]$ belongs to $\mathscr{B}_2$; thus $\mathscr{B}_2 \supseteq \mathscr{B}_3$. Finally, note that each open subinterval of $\mathbb{R}$ is the union of a sequence of sets of the form $(a, b]$ and that each open subset of $\mathbb{R}$ is the union of a sequence of open intervals (see Proposition C.4). Thus each open subset of $\mathbb{R}$ belongs to $\mathscr{B}_3$, and so $\mathscr{B}_3 \supseteq \mathscr{B}(\mathbb{R})$. □

As we proceed, the reader should note the following properties of the σ-algebra $\mathscr{B}(\mathbb{R})$:

(a) It contains virtually[2] every subset of $\mathbb{R}$ that is of interest in analysis.
(b) It is small enough that it can be dealt with in a fairly constructive manner.

It is largely these properties that explain the importance of $\mathscr{B}(\mathbb{R})$.

Proposition 1.1.5. *The σ-algebra $\mathscr{B}(\mathbb{R}^d)$ of Borel subsets of $\mathbb{R}^d$ is generated by each of the following collections of sets:*

(a) *the collection of all closed subsets of $\mathbb{R}^d$;*
(b) *the collection of all closed half-spaces in $\mathbb{R}^d$ that have the form $\{(x_1, \dots, x_d) : x_i \le b\}$ for some index i and some b in $\mathbb{R}$;*
(c) *the collection of all rectangles in $\mathbb{R}^d$ that have the form*

$$\{(x_1, \dots, x_d) : a_i < x_i \le b_i \text{ for } i = 1, \dots, d\}.$$

[2] See Chap. 8 for some interesting and useful sets that are not Borel sets.

Proof. This proposition can be proved with essentially the argument that was used for Proposition 1.1.4, and so most of the proof is omitted. To see that the σ-algebra generated by the rectangles of part (c) is included in the σ-algebra generated by the half-spaces of part (b), note that each strip that has the form

$$\{(x_1,\ldots,x_d) : a < x_i \le b\}$$

for some i is the difference of two of the half-spaces in part (b) and that each of the rectangles in part (c) is the intersection of d such strips. □

Let us look in more detail at some of the sets in $\mathscr{B}(\mathbb{R}^d)$. Let $\mathscr{G}$ be the family of all open subsets of $\mathbb{R}^d$, and let $\mathscr{F}$ be the family of all closed subsets of $\mathbb{R}^d$. (Of course $\mathscr{G}$ and $\mathscr{F}$ depend on the dimension d, and it would have been more precise to write $\mathscr{G}(\mathbb{R}^d)$ and $\mathscr{F}(\mathbb{R}^d)$.) Let $\mathscr{G}_\delta$ be the collection of all intersections of sequences of sets in $\mathscr{G}$, and let $\mathscr{F}_\sigma$ be the collection of all unions of sequences of sets in $\mathscr{F}$. Sets in $\mathscr{G}_\delta$ are often called G_δ's, and sets in $\mathscr{F}_\sigma$ are often called F_σ's. The letters G and F presumably stand for the German word *Gebiet* and the French word *fermé*, and the letters σ and δ for the German words *Summe* and *Durchschnitt*.

Proposition 1.1.6. *Each closed subset of $\mathbb{R}^d$ is a G_δ, and each open subset of $\mathbb{R}^d$ is an F_σ.*

Proof. Suppose that F is a closed subset of $\mathbb{R}^d$. We need to construct a sequence $\{U_n\}$ of open subsets of $\mathbb{R}^d$ such that $F = \cap_n U_n$. For this define U_n by

$$U_n = \{x \in \mathbb{R}^d : \|x-y\| < 1/n \text{ for some } y \text{ in } F\}.$$

(Note that U_n is empty if F is empty.) It is clear that each U_n is open and that $F \subseteq \cap_n U_n$. The reverse inclusion follows from the fact that F is closed (note that each point in $\cap_n U_n$ is the limit of a sequence of points in F). Hence each closed subset of $\mathbb{R}^d$ is a G_δ.

If U is open, then U^c is closed and so is a G_δ. Thus there is a sequence $\{U_n\}$ of open sets such that $U^c = \cap_n U_n$. The sets U_n^c are then closed, and $U = \cup_n U_n^c$; hence U is an F_σ. □

For an arbitrary family $\mathscr{S}$ of sets, let $\mathscr{S}_\sigma$ be the collection of all unions of sequences of sets in $\mathscr{S}$, and let $\mathscr{S}_\delta$ be the collection of all intersections of sequences of sets in $\mathscr{S}$. We can iterate the operations represented by σ and δ, obtaining from the class $\mathscr{G}$ the classes $\mathscr{G}_\delta$, $\mathscr{G}_{\delta\sigma}$, $\mathscr{G}_{\delta\sigma\delta}$, …, and from the class $\mathscr{F}$ the classes $\mathscr{F}_\sigma$, $\mathscr{F}_{\sigma\delta}$, $\mathscr{F}_{\sigma\delta\sigma}$, …. (Note that $\mathscr{G} = \mathscr{G}_\sigma$ and $\mathscr{F} = \mathscr{F}_\delta$. Note also that $\mathscr{G}_{\delta\delta} = \mathscr{G}_\delta$, that $\mathscr{F}_{\sigma\sigma} = \mathscr{F}_\sigma$, and so on.) It now follows (see Proposition 1.1.6) that all the inclusions in Fig. 1.1 below are valid.

It turns out that no two of these classes of sets are equal and that there are Borel sets that belong to none of them (see Exercises 7 and 9 in Sect. 8.2).

A sequence $\{A_i\}$ of sets is called *increasing* if $A_i \subseteq A_{i+1}$ holds for each i and *decreasing* if $A_i \supseteq A_{i+1}$ holds for each i.

$$\begin{tikzcd}[column sep=small]
\mathscr{G} \arrow[r, phantom, "\subset"] \arrow[dr, hook] & \mathscr{G}_\delta \arrow[r, phantom, "\subset"] \arrow[dr, hook] & \mathscr{G}_{\delta\sigma} \arrow[r, phantom, "\subset"] \arrow[dr, hook] & \mathscr{G}_{\delta\sigma\delta} \arrow[r, phantom, "\subset"] \arrow[dr, hook] & \cdots \\
\mathscr{F} \arrow[r, phantom, "\subset"] \arrow[ur, hook] & \mathscr{F}_\sigma \arrow[r, phantom, "\subset"] \arrow[ur, hook] & \mathscr{F}_{\sigma\delta} \arrow[r, phantom, "\subset"] \arrow[ur, hook] & \mathscr{F}_{\sigma\delta\sigma} \arrow[r, phantom, "\subset"] \arrow[ur, hook] & \cdots
\end{tikzcd}$$

Fig. 1.1

Proposition 1.1.7. *Let X be a set, and let $\mathscr{A}$ be an algebra on X. Then $\mathscr{A}$ is a σ-algebra if either*

(a) *$\mathscr{A}$ is closed under the formation of unions of increasing sequences of sets, or*
(b) *$\mathscr{A}$ is closed under the formation of intersections of decreasing sequences of sets.*

Proof. First suppose that condition (a) holds. Since $\mathscr{A}$ is an algebra, we can check that it is a σ-algebra by verifying that it is closed under the formation of countable unions. Suppose that $\{A_i\}$ is a sequence of sets that belong to $\mathscr{A}$. For each n let $B_n = \cup_{i=1}^n A_i$. The sequence $\{B_n\}$ is increasing, and, since $\mathscr{A}$ is an algebra, each B_n belongs to $\mathscr{A}$; thus assumption (a) implies that $\cup_n B_n$ belongs to $\mathscr{A}$. However, $\cup_i A_i$ is equal to $\cup_n B_n$ and so belongs to $\mathscr{A}$. Thus $\mathscr{A}$ is closed under the formation of countable unions and so is a σ-algebra.

Now suppose that condition (b) holds. It is enough to check that condition (a) holds. If $\{A_i\}$ is an increasing sequence of sets that belong to $\mathscr{A}$, then $\{A_i^c\}$ is a decreasing sequence of sets that belong to $\mathscr{A}$, and so condition (b) implies that $\cap_i A_i^c$ belongs to $\mathscr{A}$. Since $\cup_i A_i = (\cap_i A_i^c)^c$, it follows that $\cup_i A_i$ belongs to $\mathscr{A}$. Thus condition (a) follows from condition (b), and the proof is complete. □

Exercises

1. Find the σ-algebra on $\mathbb{R}$ that is generated by the collection of all one-point subsets of $\mathbb{R}$.
2. Show that $\mathscr{B}(\mathbb{R})$ is generated by the collection of intervals $(-\infty, b]$ for which the endpoint b is a rational number.
3. Show that $\mathscr{B}(\mathbb{R})$ is generated by the collection of all compact subsets of $\mathbb{R}$.
4. Show that if $\mathscr{A}$ is an algebra of sets, and if $\cup_n A_n$ belongs to $\mathscr{A}$ whenever $\{A_n\}$ is a sequence of disjoint sets in $\mathscr{A}$, then $\mathscr{A}$ is a σ-algebra.
5. Show by example that the union of a collection of σ-algebras on a set X can fail to be a σ-algebra on X. (Hint: There are examples in which X is a small finite set.)
6. Find an infinite collection of subsets of $\mathbb{R}$ that contains $\mathbb{R}$, is closed under the formation of countable unions, and is closed under the formation of countable intersections, but is not a σ-algebra.

7. Let $\mathscr{S}$ be a collection of subsets of the set X. Show that for each A in $\sigma(\mathscr{S})$, there is a countable subfamily $\mathscr{C}_0$ of $\mathscr{S}$ such that $A \in \sigma(\mathscr{C}_0)$. (Hint: Let $\mathscr{A}$ be the union of the σ-algebras $\sigma(\mathscr{C})$, where $\mathscr{C}$ ranges over the countable subfamilies of $\mathscr{S}$, and show that $\mathscr{A}$ is a σ-algebra that satisfies $\mathscr{S} \subseteq \mathscr{A} \subseteq \sigma(\mathscr{S})$ and hence is equal to $\sigma(\mathscr{S})$.)
8. Find all σ-algebras on $\mathbb{N}$.
9. (a) Show that $\mathbb{Q}$ is an F_σ, but not a G_δ, in $\mathbb{R}$. (Hint: Use the Baire category theorem, Theorem D.37.)
   (b) Find a subset of $\mathbb{R}$ that is neither an F_σ nor a G_δ.

1.2 Measures

Let X be a set, and let $\mathscr{A}$ be a σ-algebra on X. A function μ whose domain is the σ-algebra $\mathscr{A}$ and whose values belong to the extended half-line $[0, +\infty]$ is said to be *countably additive* if it satisfies

$$\mu(\cup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} \mu(A_i)$$

for each infinite sequence $\{A_i\}$ of disjoint sets that belong to $\mathscr{A}$. (Since $\mu(A_i)$ is nonnegative for each i, the sum $\sum_{i=1}^{\infty} \mu(A_i)$ always exists, either as a real number or as $+\infty$; see Appendix B.) A *measure* (or a *countably additive measure*) on $\mathscr{A}$ is a function $\mu : \mathscr{A} \to [0, +\infty]$ that satisfies $\mu(\varnothing) = 0$ and is countably additive.

We should note a related concept which is sometimes of interest. Let $\mathscr{A}$ be an algebra (not necessarily a σ-algebra) on the set X. A function μ whose domain is $\mathscr{A}$ and whose values belong to $[0, +\infty]$ is *finitely additive* if it satisfies

$$\mu(\cup_{i=1}^{n} A_i) = \sum_{i=1}^{n} \mu(A_i)$$

for each finite sequence $A_1, \ldots, A_n$ of disjoint sets that belong to $\mathscr{A}$. A *finitely additive measure* on the algebra $\mathscr{A}$ is a function $\mu : \mathscr{A} \to [0, +\infty]$ that satisfies $\mu(\varnothing) = 0$ and is finitely additive.

It is easy to check that every countably additive measure is finitely additive: simply extend the finite sequence $A_1, \ldots, A_n$ to an infinite sequence $\{A_i\}$ by letting $A_i = \varnothing$ if $i > n$, and then use the fact that $\mu(\varnothing) = 0$. There are, however, finitely additive measures that are not countably additive (see Example 1.2.1(d) and Exercise 8 in Sect. 3.5).

Finite additivity might at first seem to be a more natural property than countable additivity. However, countably additive measures on the one hand seem to be sufficient for almost all applications and, on the other hand, support a much more powerful theory of integration than do finitely additive measures. Thus we will follow the usual practice and devote almost all of our attention to countably additive measures.

We should emphasize that in this book the word "measure" (without modifiers) will always denote a countably additive measure. The expression "finitely additive measure" will always be written out in full.

If X is a set, if $\mathscr{A}$ is a σ-algebra on X, and if μ is a measure on $\mathscr{A}$, then the triplet $(X, \mathscr{A}, \mu)$ is often called a *measure space*. Likewise, if X is a set and if $\mathscr{A}$ is a σ-algebra on X, then the pair $(X, \mathscr{A})$ is often called a *measurable space*. If $(X, \mathscr{A}, \mu)$ is a measure space, then one often says that μ is a *measure on* $(X, \mathscr{A})$, or, if the σ-algebra $\mathscr{A}$ is clear from context, a *measure on* X.

Examples 1.2.1.

(a) Let X be an arbitrary set, and let $\mathscr{A}$ be a σ-algebra on X. Define a function $\mu : \mathscr{A} \to [0, +\infty]$ by letting $\mu(A)$ be n if A is a finite set with n elements and letting $\mu(A)$ be $+\infty$ if A is an infinite set. Then μ is a measure; it is often called *counting measure* on $(X, \mathscr{A})$.

(b) Let X be a nonempty set, and let $\mathscr{A}$ be a σ-algebra on X. Let x be a member of X. Define a function $\delta_x : \mathscr{A} \to [0, +\infty]$ by letting $\delta_x(A)$ be 1 if $x \in A$ and letting $\delta_x(A)$ be 0 if $x \notin A$. Then δ_x is a measure; it is called a *point mass* concentrated at x.

(c) Consider the set $\mathbb{R}$ of all real numbers and the σ-algebra $\mathscr{B}(\mathbb{R})$ of Borel subsets of $\mathbb{R}$. In Sect. 1.3 we will construct a measure on $\mathscr{B}(\mathbb{R})$ that assigns to each subinterval of $\mathbb{R}$ its length; this measure is known as Lebesgue measure and will be denoted by λ in this book.

(d) Let X be the set of all positive integers, and let $\mathscr{A}$ be the collection of all subsets A of X such that either A or A^c is finite. Then $\mathscr{A}$ is an algebra, but not a σ-algebra (see Example 1.1.1(d)). Define a function $\mu : \mathscr{A} \to [0, +\infty]$ by letting $\mu(A)$ be 1 if A is infinite and letting $\mu(A)$ be 0 if A is finite. It is easy to check that μ is a finitely additive measure; however, it is impossible to extend μ to a countably additive measure on the σ-algebra generated by $\mathscr{A}$ (if $A_k = \{k\}$ for each k, then $\mu(\cup_{k=1}^{\infty} A_k) = \mu(X) = 1$, while $\sum_{k=1}^{\infty} \mu(A_k) = 0$).

(e) Let X be an arbitrary set, and let $\mathscr{A}$ be an arbitrary σ-algebra on X. Define a function $\mu : \mathscr{A} \to [0, +\infty]$ by letting $\mu(A)$ be $+\infty$ if $A \neq \varnothing$, and letting $\mu(A)$ be 0 if $A = \varnothing$. Then μ is a measure.

(f) Let X be a set that has at least two members, and let $\mathscr{A}$ be the σ-algebra consisting of all subsets of X. Define a function $\mu : \mathscr{A} \to [0, +\infty]$ by letting $\mu(A)$ be 1 if $A \neq \varnothing$ and letting $\mu(A)$ be 0 if $A = \varnothing$. Then μ is not a measure, nor even a finitely additive measure, for if A_1 and A_2 are disjoint nonempty subsets of X, then $\mu(A_1 \cup A_2) = 1$, while $\mu(A_1) + \mu(A_2) = 2$. □

Proposition 1.2.2. *Let $(X, \mathscr{A}, \mu)$ be a measure space, and let A and B be subsets of X that belong to $\mathscr{A}$ and satisfy $A \subseteq B$. Then $\mu(A) \le \mu(B)$. If in addition A satisfies $\mu(A) < +\infty$, then $\mu(B - A) = \mu(B) - \mu(A)$.*

Proof. The sets A and $B - A$ are disjoint and satisfy $B = A \cup (B - A)$; thus the additivity of μ implies that

$$\mu(B) = \mu(A) + \mu(B - A).$$

Since $\mu(B-A) \geq 0$, it follows that $\mu(A) \leq \mu(B)$. In case $\mu(A) < +\infty$, the relation $\mu(B) - \mu(A) = \mu(B-A)$ also follows. □

Let μ be a measure on a measurable space $(X, \mathscr{A})$. Then μ is a *finite* measure if $\mu(X) < +\infty$ and is a σ-*finite* measure if X is the union of a sequence $A_1, A_2, \dots$ of sets that belong to $\mathscr{A}$ and satisfy $\mu(A_i) < +\infty$ for each i. More generally, a set in $\mathscr{A}$ is σ-finite under μ if it is the union of a sequence of sets that belong to $\mathscr{A}$ and have finite measure under μ. The measure space $(X, \mathscr{A}, \mu)$ is also called *finite* or σ-*finite* if μ is finite or σ-finite. Most of the constructions and basic properties that we will consider are valid for all measures. For a few important theorems, however, we will need to assume that the measures involved are finite or σ-finite.

If the measure space $(X, \mathscr{A}, \mu)$ is σ-finite, then X is the union of a sequence $\{B_i\}$ of disjoint sets that belong to $\mathscr{A}$ and have finite measure under μ; such a sequence $\{B_i\}$ can be formed by choosing a sequence $\{A_i\}$ as in the definition of σ-finiteness, and then letting $B_1 = A_1$ and $B_i = A_i - (\cup_{j=1}^{i-1} A_j)$ if $i > 1$.

Examples 1.2.3 (Dealing with σ-Finiteness). Note that the measure defined in Example 1.2.1(a) is finite if and only if the set X is finite and is σ-finite if and only if the set X is the union of a sequence of finite sets that belong to $\mathscr{A}$.[3] The measure defined in Example 1.2.1(b) is finite. Lebesgue measure, described in Example 1.2.1(c), is σ-finite, since $\mathbb{R}$ is the union of a sequence of bounded intervals. See also Exercises 2 and 7 below. □

The following propositions give some elementary but useful properties of measures.

Proposition 1.2.4. *Let $(X, \mathscr{A}, \mu)$ be a measure space. If $\{A_k\}$ is an arbitrary sequence of sets that belong to $\mathscr{A}$, then*

$$\mu(\cup_{k=1}^{\infty} A_k) \leq \sum_{k=1}^{\infty} \mu(A_k).$$

Proof. Define a sequence $\{B_k\}$ of subsets of X by letting $B_1 = A_1$ and letting $B_k = A_k - (\cup_{i=1}^{k-1} A_i)$ if $k > 1$. Then each B_k belongs to $\mathscr{A}$ and is a subset of the corresponding A_k, and so satisfies $\mu(B_k) \leq \mu(A_k)$. Since in addition the sets B_k are disjoint and satisfy $\cup_k B_k = \cup_k A_k$, it follows that

$$\mu(\cup_k A_k) = \mu(\cup_k B_k) = \sum_k \mu(B_k) \leq \sum_k \mu(A_k). \qquad \square$$

In other words, the countable additivity of μ implies the *countable subadditivity* of μ.

---

[3] If in Example 1.2.1(a) the σ-algebra $\mathscr{A}$ contains all the subsets of X, then μ is σ-finite if and only if X is at most countably infinite.

Proposition 1.2.5. *Let $(X, \mathscr{A}, \mu)$ be a measure space.*

(a) *If $\{A_k\}$ is an increasing sequence of sets that belong to $\mathscr{A}$, then $\mu(\cup_k A_k) = \lim_k \mu(A_k)$.*
(b) *If $\{A_k\}$ is a decreasing sequence of sets that belong to $\mathscr{A}$, and if $\mu(A_n) < +\infty$ holds for some n, then $\mu(\cap_k A_k) = \lim_k \mu(A_k)$.*

Proof. First suppose that $\{A_k\}$ is an increasing sequence of sets that belong to $\mathscr{A}$, and define a sequence $\{B_i\}$ of sets by letting $B_1 = A_1$ and letting $B_i = A_i - A_{i-1}$ if $i > 1$. The sets just constructed are disjoint, belong to $\mathscr{A}$, and satisfy $A_k = \cup_{i=1}^k B_i$ for each k. It follows that $\cup_k A_k = \cup_i B_i$ and hence that

$$\mu(\cup_k A_k) = \sum_i \mu(B_i) = \lim_k \sum_{i=1}^k \mu(B_i) = \lim_k \mu(\cup_{i=1}^k B_i) = \lim_k \mu(A_k).$$

This completes the proof of (a).

Now suppose that $\{A_k\}$ is a decreasing sequence of sets that belong to $\mathscr{A}$ and that $\mu(A_n) < +\infty$ holds for some n. We can assume that $n = 1$. For each k let $C_k = A_1 - A_k$. Then $\{C_k\}$ is an increasing sequence of sets that belong to $\mathscr{A}$ and satisfy

$$\cup_k C_k = A_1 - (\cap_k A_k).$$

It follows from part (a) that $\mu(\cup_k C_k) = \lim_k \mu(C_k)$ and hence that

$$\mu(A_1 - (\cap_k A_k)) = \mu(\cup_k C_k) = \lim_k \mu(C_k) = \lim_k \mu(A_1 - A_k).$$

In view of Proposition 1.2.2 and the assumption that $\mu(A_1) < +\infty$, this implies that $\mu(\cap_k A_k) = \lim_k \mu(A_k)$. □

The preceding proposition has the following partial converse, which is sometimes useful for checking that a finitely additive measure is in fact countably additive.

Proposition 1.2.6. *Let $(X, \mathscr{A})$ be a measurable space, and let μ be a finitely additive measure on $(X, \mathscr{A})$. Then μ is a measure if either*

(a) $\lim_k \mu(A_k) = \mu(\cup_k A_k)$ *holds for each increasing sequence $\{A_k\}$ of sets that belong to $\mathscr{A}$, or*
(b) $\lim_k \mu(A_k) = 0$ *holds for each decreasing sequence $\{A_k\}$ of sets that belong to $\mathscr{A}$ and satisfy $\cap_k A_k = \varnothing$.*

Proof. We need to verify the countable additivity of μ. Let $\{B_j\}$ be a sequence of disjoint sets that belong to $\mathscr{A}$; we will prove that $\mu(\cup_j B_j) = \sum_j \mu(B_j)$.

First assume that condition (a) holds, and for each k let $A_k = \cup_{j=1}^k B_j$. Then the finite additivity of μ implies that $\mu(A_k) = \sum_{j=1}^k \mu(B_j)$, while condition (a) implies that $\mu(\cup_{k=1}^\infty A_k) = \lim_k \mu(A_k)$; since $\cup_{j=1}^\infty B_j = \cup_{k=1}^\infty A_k$, it follows that

$$\mu(\cup_{j=1}^\infty B_j) = \mu(\cup_{k=1}^\infty A_k) = \lim_k \mu(A_k) = \sum_{j=1}^\infty \mu(B_j).$$

Now assume that condition (b) holds, and for each k let $A_k = \cup_{j=k}^{\infty} B_j$. Then the finite additivity of μ implies that

$$\mu(\cup_{j=1}^{\infty} B_j) = \sum_{j=1}^{k} \mu(B_j) + \mu(A_{k+1}),$$

while condition (b) implies that $\lim_k \mu(A_{k+1}) = 0$; hence $\mu(\cup_{j=1}^{\infty} B_j) = \sum_{j=1}^{\infty} \mu(B_j)$. $\square$

Let us close this section by introducing a bit of terminology. A measure on $(\mathbb{R}^d, \mathscr{B}(\mathbb{R}^d))$ is often called a *Borel measure* on $\mathbb{R}^d$. More generally, if X is a Borel subset of $\mathbb{R}^d$ and if $\mathscr{A}$ is the σ-algebra consisting of those Borel subsets of $\mathbb{R}^d$ that are included in X, then a measure on $(X, \mathscr{A})$ is called a *Borel measure* on X.

Now suppose that $(X, \mathscr{A})$ is a measurable space such that for each x in X the set $\{x\}$ belongs to $\mathscr{A}$. A finite or σ-finite measure μ on $(X, \mathscr{A})$ is *continuous* if $\mu(\{x\}) = 0$ holds for each x in X and is *discrete* if there is a countable subset D of X such that $\mu(D^c) = 0$. (More elaborate definitions are needed if $\mathscr{A}$ does not contain each $\{x\}$ or if μ is not σ-finite. We will, however, not need to consider such matters.)

Exercises

1. Suppose that μ is a finite measure on $(X, \mathscr{A})$.
   (a) Show that if A and B belong to $\mathscr{A}$, then
   $$\mu(A \cup B) = \mu(A) + \mu(B) - \mu(A \cap B).$$
   (b) Show that if A, B, and C belong to $\mathscr{A}$, then
   $$\begin{aligned} \mu(A \cup B \cup C) = &\mu(A) + \mu(B) + \mu(C) \\ &- \mu(A \cap B) - \mu(A \cap C) - \mu(B \cap C) \\ &+ \mu(A \cap B \cap C). \end{aligned}$$
   (c) Find and prove a corresponding formula for the measure of the union of n sets.
2. Define μ on $(\mathbb{R}, \mathscr{B}(\mathbb{R}))$ by letting $\mu(A)$ be the number of rational numbers in A (of course $\mu(A) = +\infty$ if there are infinitely many rational numbers in A). Show that μ is a σ-finite measure under which each open subinterval of $\mathbb{R}$ has infinite measure.
3. Let $\mathscr{A}$ be the σ-algebra of all subsets of $\mathbb{N}$, and let μ be counting measure on $(\mathbb{N}, \mathscr{A})$. Give a decreasing sequence $\{A_k\}$ of sets in $\mathscr{A}$ such that $\mu(\cap_k A_k) \neq \lim_k \mu(A_k)$. Hence the finiteness assumption cannot be removed from part (b) of Proposition 1.2.5.

4. Let $(X, \mathscr{A})$ be a measurable space.
   (a) Suppose that μ is a nonnegative countably additive function on $\mathscr{A}$. Show that if $\mu(A)$ is finite for some A in $\mathscr{A}$, then $\mu(\varnothing) = 0$. (Thus μ is a measure.)
   (b) Show by example that in general the condition $\mu(\varnothing) = 0$ does not follow from the remaining parts of the definition of a measure.
5. Let $(X, \mathscr{A})$ be a measurable space, and let x and y belong to X. Show that the point masses δ_x and δ_y are equal if and only if x and y belong to exactly the same sets in $\mathscr{A}$.
6. Let $(X, \mathscr{A})$ be a measurable space.
   (a) Show that if $\{\mu_n\}$ is an increasing sequence of measures on $(X, \mathscr{A})$ (here "increasing" means that $\mu_n(A) \le \mu_{n+1}(A)$ holds for each A and each n), then the formula $\mu(A) = \lim_n \mu_n(A)$ defines a measure on $(X, \mathscr{A})$.
   (b) Show that if $\{\mu_n\}$ is an arbitrary sequence of measures on $(X, \mathscr{A})$, then the formula $\mu(A) = \sum_n \mu_n(A)$ defines a measure on $(X, \mathscr{A})$.
7. Let $\{x_n\}$ be a sequence of real numbers, and define a measure μ on $(\mathbb{R}, \mathscr{B}(\mathbb{R}))$ by $\mu = \sum_n \delta_{x_n}$ (see Exercise 6).
   (a) Show that μ assigns finite values to the bounded subintervals of $\mathbb{R}$ if and only if $\lim_n |x_n| = +\infty$.
   (b) For which sequences $\{x_n\}$ is the measure μ σ-finite?
8. Let $(X, \mathscr{A}, \mu)$ be a measure space, and define $\mu^\bullet \colon \mathscr{A} \to [0, +\infty]$ by

$$\mu^\bullet(A) = \sup\{\mu(B) : B \subseteq A,\, B \in \mathscr{A}, \text{ and } \mu(B) < +\infty\}.$$

   (a) Show that $\mu^\bullet$ is a measure on $(X, \mathscr{A})$.
   (b) Show that if μ is σ-finite, then $\mu^\bullet = \mu$.
   (c) Find $\mu^\bullet$ if X is nonempty and μ is the measure defined by

$$\mu(A) = \begin{cases} +\infty & \text{if } A \in \mathscr{A} \text{ and } A \ne \varnothing, \text{ and} \\ 0 & \text{if } A = \varnothing. \end{cases}$$

9. Let μ be a measure on $(X, \mathscr{A})$, and let $\{A_k\}$ be a sequence of sets in $\mathscr{A}$ such that $\sum_k \mu(A_k) < +\infty$. Show that the set of points that belong to A_k for infinitely many values of k has measure zero under μ. (Hint: Consider the set $\cap_{n=1}^\infty \cup_{k=n}^\infty A_k$, and note that $\mu(\cap_{n=1}^\infty \cup_{k=n}^\infty A_k) \le \mu(\cup_{k=p}^\infty A_k)$ holds for each p.)

1.3 Outer Measures

In this section we develop one of the standard techniques for constructing measures; then we use it to construct Lebesgue measure on $\mathbb{R}^d$.

Let X be a set, and let $\mathscr{P}(X)$ be the collection of all subsets of X. An *outer measure* on X is a function $\mu^* \colon \mathscr{P}(X) \to [0, +\infty]$ such that

(a) $\mu^*(\varnothing) = 0$,
(b) if $A \subseteq B \subseteq X$, then $\mu^*(A) \le \mu^*(B)$, and
(c) if $\{A_n\}$ is an infinite sequence of subsets of X, then $\mu^*(\cup_n A_n) \le \sum_n \mu^*(A_n)$.

Thus an outer measure on X is a *monotone* and *countably subadditive* function from $\mathscr{P}(X)$ to $[0,+\infty]$ whose value at $\varnothing$ is 0.

Note that a measure can fail to be an outer measure; in fact, a measure on X is an outer measure if and only if its domain is $\mathscr{P}(X)$ (see Propositions 1.2.2 and 1.2.4). On the other hand, an outer measure generally fails to be countably additive and so fails to be a measure.

In Theorem 1.3.6, we will prove that for each outer measure μ^* on X there is a relatively natural σ-algebra $\mathscr{M}_{\mu^*}$ on X such that the restriction of μ^* to $\mathscr{M}_{\mu^*}$ is countably additive, and hence a measure. Many important measures can be derived from outer measures in this way.

Examples 1.3.1.

(a) Let X be an arbitrary set, and define μ^* on $\mathscr{P}(X)$ by $\mu^*(A) = 0$ if $A = \varnothing$ and $\mu^*(A) = 1$ otherwise. Then μ^* is an outer measure.
(b) Let X be an arbitrary set, and define μ^* on $\mathscr{P}(X)$ by $\mu^*(A) = 0$ if A is countable, and $\mu^*(A) = 1$ if A is uncountable. Then μ^* is an outer measure.
(c) Let X be an infinite set, and define μ^* on $\mathscr{P}(X)$ by $\mu^*(A) = 0$ if A is finite, and $\mu^*(A) = 1$ if A is infinite. Then μ^* fails to be countably subadditive and so is not an outer measure.
(d) *Lebesgue outer measure* on $\mathbb{R}$, which we will denote by λ^*, is defined as follows. For each subset A of $\mathbb{R}$, let $\mathscr{C}_A$ be the set of all infinite sequences $\{(a_i,b_i)\}$ of bounded open intervals such that $A \subseteq \cup_i (a_i,b_i)$. Then $\lambda^* : \mathscr{P}(\mathbb{R}) \to [0,+\infty]$ is defined by

$$\lambda^*(A) = \inf\left\{ \sum_i (b_i - a_i) : \{(a_i,b_i)\} \in \mathscr{C}_A \right\}.$$

(Note that the set of sums involved here is nonempty and that the infimum of the set consisting of $+\infty$ alone is $+\infty$. We check in the following proposition that λ^* is indeed an outer measure.) □

Proposition 1.3.2. *Lebesgue outer measure on $\mathbb{R}$ is an outer measure, and it assigns to each subinterval of $\mathbb{R}$ its length.*

Proof. We begin by verifying that λ^* is an outer measure. The relation $\lambda^*(\varnothing) = 0$ holds, since for each positive number ε there is a sequence $\{(a_i,b_i)\}$ of open intervals (whose union necessarily includes $\varnothing$) such that $\sum_i (b_i - a_i) < \varepsilon$. For the monotonicity of λ^*, note that if $A \subseteq B$, then each sequence of open intervals that covers B also covers A, and so $\lambda^*(A) \le \lambda^*(B)$. Now consider the countable subadditivity of λ^*. Let $\{A_n\}_{n=1}^{\infty}$ be an arbitrary sequence of subsets of $\mathbb{R}$. If $\sum_n \lambda^*(A_n) = +\infty$, then $\lambda^*(\cup_n A_n) \le \sum_n \lambda^*(A_n)$ certainly holds. So suppose that

$\sum_n \lambda^*(A_n) < +\infty$, and let ε be an arbitrary positive number. For each n choose a sequence $\{(a_{n,i}, b_{n,i})\}_{i=1}^{\infty}$ that covers A_n and satisfies

$$\sum_{i=1}^{\infty}(b_{n,i} - a_{n,i}) < \lambda^*(A_n) + \varepsilon/2^n.$$

If we combine these sequences into one sequence $\{(a_j, b_j)\}$ (see, for example, the construction in the last paragraph of A.6), then the combined sequence satisfies

$$\cup_n A_n \subseteq \cup_j (a_j, b_j)$$

and

$$\sum_j (b_j - a_j) < \sum_n (\lambda^*(A_n) + \varepsilon/2^n) = \sum_n \lambda^*(A_n) + \varepsilon.$$

These relations, together with the fact that ε is arbitrary, imply that $\lambda^*(\cup_n A_n) \leq \sum_n \lambda^*(A_n)$. Thus λ^* is an outer measure.

Now we compute the outer measure of the subintervals of $\mathbb{R}$. First consider a closed bounded interval $[a,b]$. It is easy to see that $\lambda^*([a,b]) \leq b-a$ (cover $[a,b]$ with sequences of open intervals in which the first interval is barely larger than $[a,b]$, and the sum of the lengths of the other intervals is very small). We turn to the reverse inequality. Let $\{(a_i, b_i)\}$ be a sequence of bounded open intervals whose union includes $[a,b]$. Since $[a,b]$ is compact, there is a positive integer n such that $[a,b] \subseteq \cup_{i=1}^{n}(a_i, b_i)$. It is easy to check that $b-a \leq \sum_{i=1}^{n}(b_i - a_i)$ (use induction on n) and hence that $b-a \leq \sum_{i=1}^{\infty}(b_i - a_i)$. Since $\{(a_i, b_i)\}$ was an arbitrary sequence whose union includes $[a,b]$, it follows that $b-a \leq \lambda^*([a,b])$. Thus $\lambda^*([a,b]) = b-a$.

The outer measure of an arbitrary bounded interval is its length, since such an interval I includes and is included in closed bounded intervals of length arbitrarily close to the length of I. Finally, an unbounded interval has infinite outer measure, since it includes arbitrarily long closed bounded intervals. □

Let us look at another basic example.

Example 1.3.3. *Lebesgue outer measure* on $\mathbb{R}^d$, which we will denote by λ^* (or, if necessary in order to avoid ambiguity, by λ_d^*) is defined as follows. A *d-dimensional interval* is a subset of $\mathbb{R}^d$ of the form $I_1 \times \cdots \times I_d$, where $I_1, \ldots, I_d$ are subintervals of $\mathbb{R}$ and $I_1 \times \cdots \times I_d$ is given by

$$I_1 \times \cdots \times I_d = \{(x_i, \ldots, x_d) : x_i \in I_i \text{ for } i = 1, \ldots, d\}.$$

Note that the intervals $I_1, \ldots, I_d$, and hence the d-dimensional interval $I_1 \times \cdots \times I_d$, can be open, closed, or neither open nor closed. The *volume* of the d-dimensional interval $I_1 \times \cdots \times I_d$ is the product of the lengths of the intervals $I_1, \ldots, I_d$, and will be denoted by $\text{vol}(I_1 \times \cdots \times I_d)$. For each subset A of $\mathbb{R}^d$ let $\mathscr{C}_A$ be the set of all sequences $\{R_i\}$ of bounded and open d-dimensional intervals for which $A \subseteq \cup_{i=1}^{\infty} R_i$. Then $\lambda^*(A)$, the outer measure of A, is the infimum of the set

$$\left\{ \sum_{i=1}^{\infty} \operatorname{vol}(R_i) : \{R_i\} \in \mathscr{C}_A \right\}. \qquad \square$$

We note the following analogue of Proposition 1.3.2.

Proposition 1.3.4. *Lebesgue outer measure on $\mathbb{R}^d$ is an outer measure, and it assigns to each d-dimensional interval its volume.*

Proof. Most of the details are omitted, since they are very similar to those in the proof of Proposition 1.3.2. Note, however, that if K is a compact d-dimensional interval and if $\{R_i\}_{i=1}^{\infty}$ is a sequence of bounded and open d-dimensional intervals for which $K \subseteq \cup_{i=1}^{\infty} R_i$, then there is a positive integer n such that $K \subseteq \cup_{i=1}^{n} R_i$, and K can be decomposed into a finite collection $\{K_j\}$ of d-dimensional intervals that overlap only on their boundaries and are such that for each j the interior of K_j is included in some R_i (where $i \le n$). From this it follows that

$$\operatorname{vol}(K) = \sum_j \operatorname{vol}(K_j) \le \sum_i \operatorname{vol}(R_i)$$

and hence that $\operatorname{vol}(K) \le \lambda^*(K)$. The remaining modifications needed to convert our proof of Proposition 1.3.2 into a proof of the present result are straightforward. $\square$

Let X be a set, and let μ^* be an outer measure on X. A subset B of X is μ^*-*measurable* (or *measurable with respect to* μ^*) if

$$\mu^*(A) = \mu^*(A \cap B) + \mu^*(A \cap B^c)$$

holds for *every* subset A of X. Thus a μ^*-measurable subset of X is one that divides each subset of X in such a way that the sizes (as measured by μ^*) of the pieces add properly. A *Lebesgue measurable* subset of $\mathbb{R}$ or of $\mathbb{R}^d$ is of course one that is measurable with respect to Lebesgue outer measure.

Note that the subadditivity of the outer measure μ^* implies that

$$\mu^*(A) \le \mu^*(A \cap B) + \mu^*(A \cap B^c)$$

holds for all subsets A and B of X. Thus to check that a subset B of X is μ^*-measurable, we need only check that

$$\mu^*(A) \ge \mu^*(A \cap B) + \mu^*(A \cap B^c) \tag{1}$$

holds for each subset A of X. Note also that inequality (1) certainly holds if $\mu^*(A) = +\infty$. Thus the μ^*-measurability of B can be verified by checking that (1) holds for each A that satisfies $\mu^*(A) < +\infty$.

Proposition 1.3.5. *Let X be a set, and let μ^* be an outer measure on X. Then each subset B of X that satisfies $\mu^*(B) = 0$ or that satisfies $\mu^*(B^c) = 0$ is μ^*-measurable.*

Proof. Assume that $\mu^*(B) = 0$ or that $\mu^*(B^c) = 0$. According to the remarks above, we need only check that each subset A of X satisfies

$$\mu^*(A) \geq \mu^*(A \cap B) + \mu^*(A \cap B^c).$$

However our assumption about B and the monotonicity of μ^* imply that one of the terms on the right-hand side of this inequality vanishes and that the other is at most $\mu^*(A)$; thus the required inequality follows. □

It follows that the sets $\varnothing$ and X are measurable for every outer measure on X.

The following theorem is the fundamental fact about outer measures; it will be the key to many of our constructions of measures.

Theorem 1.3.6. *Let X be a set, let μ^* be an outer measure on X, and let $\mathscr{M}_{\mu^*}$ be the collection of all μ^*-measurable subsets of X. Then*

(a) *$\mathscr{M}_{\mu^*}$ is a σ-algebra, and*
(b) *the restriction of μ^* to $\mathscr{M}_{\mu^*}$ is a measure on $\mathscr{M}_{\mu^*}$.*

Proof. We begin by showing that $\mathscr{M}_{\mu^*}$ is an algebra of sets. First note that Proposition 1.3.5 implies that X belongs to $\mathscr{M}_{\mu^*}$. Note also that the equation

$$\mu^*(A) = \mu^*(A \cap B) + \mu^*(A \cap B^c)$$

is not changed if the sets B and B^c are interchanged; thus the μ^*-measurability of B implies that of B^c, and so $\mathscr{M}_{\mu^*}$ is closed under complementation. Now suppose that B_1 and B_2 are μ^*-measurable subsets of X; we will show that $B_1 \cup B_2$ is μ^*-measurable. For this, let A be an arbitrary subset of X. The μ^*-measurability of B_1 implies

$$\begin{aligned}\mu^*(A \cap (B_1 \cup B_2)) &= \mu^*(A \cap (B_1 \cup B_2) \cap B_1) + \mu^*(A \cap (B_1 \cup B_2) \cap B_1^c) \\ &= \mu^*(A \cap B_1) + \mu^*(A \cap B_1^c \cap B_2).\end{aligned}$$

If we use this identity and the fact that $(B_1 \cup B_2)^c = B_1^c \cap B_2^c$, and then simplify the resulting expression by appealing first to the measurability of B_2 and then to the measurability of B_1, we find

$$\begin{aligned}&\mu^*(A \cap (B_1 \cup B_2)) + \mu^*(A \cap (B_1 \cup B_2)^c) \\ &\quad = \mu^*(A \cap B_1) + \mu^*(A \cap B_1^c \cap B_2) + \mu^*(A \cap B_1^c \cap B_2^c) \\ &\quad = \mu^*(A \cap B_1) + \mu^*(A \cap B_1^c) \\ &\quad = \mu^*(A).\end{aligned}$$

Since A was an arbitrary subset of X, the set $B_1 \cup B_2$ must be measurable. Thus $\mathscr{M}_{\mu^*}$ is an algebra.

Next suppose that $\{B_i\}$ is an infinite sequence of disjoint μ^*-measurable sets; we will show by induction that

$$\mu^*(A) = \sum_{i=1}^{n} \mu^*(A \cap B_i) + \mu^*(A \cap (\cap_{i=1}^{n} B_i^c)) \tag{2}$$

holds for each subset A of X and each positive integer n. Equation (2) is, in the case where $n = 1$, simply a restatement of the measurability of B_1. As to the induction step, note that the μ^*-measurability of B_{n+1} and the disjointness of the sequence $\{B_i\}$ imply that

$$\begin{aligned}
&\mu^*(A \cap (\cap_{i=1}^n B_i^c)) \\
&\quad = \mu^*(A \cap (\cap_{i=1}^n B_i^c) \cap B_{n+1}) + \mu^*(A \cap (\cap_{i=1}^n B_i^c) \cap B_{n+1}^c) \\
&\quad = \mu^*(A \cap B_{n+1}) + \mu^*(A \cap (\cap_{i=1}^{n+1} B_i^c)).
\end{aligned}$$

With this (2) is proved.

Note that we do not increase the right-hand side of Eq. (2) if we replace $\mu^*(A \cap (\cap_{i=1}^n B_i^c))$ with $\mu^*(A \cap (\cap_{i=1}^\infty B_i^c))$, and thus with $\mu^*(A \cap (\cup_{i=1}^\infty B_i)^c)$; by letting the n in the sum in the resulting inequality approach infinity, we find

$$\mu^*(A) \geq \sum_{i=1}^\infty \mu^*(A \cap B_i) + \mu^*(A \cap (\cup_{i=1}^\infty B_i)^c). \tag{3}$$

This and the countable subadditivity of μ^* imply that

$$\begin{aligned}
\mu^*(A) &\geq \sum_{i=1}^\infty \mu^*(A \cap B_i) + \mu^*(A \cap (\cup_{i=1}^\infty B_i)^c) \\
&\geq \mu^*(A \cap (\cup_{i=1}^\infty B_i)) + \mu^*(A \cap (\cup_{i=1}^\infty B_i)^c) \\
&\geq \mu^*(A);
\end{aligned}$$

it follows that each inequality in the preceding calculation must in fact be an equality and hence that $\cup_{i=1}^\infty B_i$ is μ^*-measurable. Thus $\mathscr{M}_{\mu^*}$ is closed under the formation of unions of disjoint sequences of sets. Since the union of an arbitrary sequence $\{B_i\}$ of sets in $\mathscr{M}_{\mu^*}$ is the union of a disjoint sequence of sets in $\mathscr{M}_{\mu^*}$, namely of the sequence

$$B_1,\, B_1^c \cap B_2,\, \ldots,\, B_1^c \cap B_2^c \cap \cdots \cap B_{n-1}^c \cap B_n,\, \ldots,$$

the algebra $\mathscr{M}_{\mu^*}$ is closed under the formation of countable unions. With this we have proved that $\mathscr{M}_{\mu^*}$ is a σ-algebra.

To show that the restriction of μ^* to $\mathscr{M}_{\mu^*}$ is a measure, we need to verify its countable additivity. If $\{B_i\}$ is a sequence of disjoint sets in $\mathscr{M}_{\mu^*}$, then replacing A with $\cup_{i=1}^\infty B_i$ in inequality (3) yields

$$\mu^*(\cup_{i=1}^\infty B_i) \geq \sum_{i=1}^\infty \mu^*(B_i) + 0;$$

since the reverse inequality is automatic, the countable additivity of the restriction of μ^* to $\mathscr{M}_{\mu^*}$ follows. □

We turn to applications of Theorem 1.3.6 and begin with Lebesgue measure. We will denote the *collection of Lebesgue measurable subsets* of $\mathbb{R}$ by $\mathscr{M}_{\lambda^*}$.

Proposition 1.3.7. *Every Borel subset of* $\mathbb{R}$ *is Lebesgue measurable.*

Proof. We begin by checking that every interval of the form $(-\infty, b]$ is Lebesgue measurable. Let B be such an interval. According to the remarks made just before the statement of Proposition 1.3.5, we need only check that

$$\lambda^*(A) \geq \lambda^*(A \cap B) + \lambda^*(A \cap B^c) \tag{4}$$

holds for each subset A of $\mathbb{R}$ for which $\lambda^*(A) < +\infty$. Let A be such a set, let ε be an arbitrary positive number, and let $\{(a_n, b_n)\}$ be a sequence of open intervals that covers A and satisfies $\sum_{n=1}^{\infty}(b_n - a_n) < \lambda^*(A) + \varepsilon$. Then for each n the sets $(a_n, b_n) \cap B$ and $(a_n, b_n) \cap B^c$ are disjoint intervals (one of which may instead be the empty set) whose union is (a_n, b_n), and so

$$b_n - a_n = \lambda^*((a_n, b_n)) = \lambda^*((a_n, b_n) \cap B) + \lambda^*((a_n, b_n) \cap B^c) \tag{5}$$

(see Proposition 1.3.2). Since the sequence $\{(a_n, b_n) \cap B\}$ covers $A \cap B$ and the sequence $\{(a_n, b_n) \cap B^c\}$ covers $A \cap B^c$, we have from Eq. (5) and the countable subadditivity of λ^* that

$$\begin{aligned}
\lambda^*(A \cap B) + \lambda^*(A \cap B^c) &\leq \sum_n \lambda^*((a_n, b_n) \cap B) + \sum_n \lambda^*((a_n, b_n) \cap B^c) \\
&= \sum_n (b_n - a_n) < \lambda^*(A) + \varepsilon.
\end{aligned}$$

However, ε was arbitrary, and so inequality (4) and the Lebesgue measurability of B follow.

Thus the collection $\mathscr{M}_{\lambda^*}$ of Lebesgue measurable sets is a σ-algebra on $\mathbb{R}$ (Theorem 1.3.6) that contains each interval of the form $(-\infty, b]$. However $\mathscr{B}(\mathbb{R})$ is the smallest σ-algebra on $\mathbb{R}$ that contains all these intervals (Proposition 1.1.4), and so $\mathscr{B}(\mathbb{R}) \subseteq \mathscr{M}_{\lambda^*}$. □

We will also use $\mathscr{M}_{\lambda^*}$ to denote the *collection of Lebesgue measurable subsets* of $\mathbb{R}^d$.

Proposition 1.3.8. *Every Borel subset of* $\mathbb{R}^d$ *is Lebesgue measurable.*

Proof. It is easy to give a proof of Proposition 1.3.8 by modifying that of Proposition 1.3.7; the details are left to the reader. □

The restriction of Lebesgue outer measure on $\mathbb{R}$ (or on $\mathbb{R}^d$) to the collection $\mathscr{M}_{\lambda^*}$ of Lebesgue measurable subsets of $\mathbb{R}$ (or of $\mathbb{R}^d$) is called *Lebesgue measure* and will be denoted by λ or by λ_d. The restriction of Lebesgue outer measure to $\mathscr{B}(\mathbb{R})$ or to $\mathscr{B}(\mathbb{R}^d)$ is also called *Lebesgue measure*, and it too will be denoted by λ or by λ_d. We can specify which version of Lebesgue measure we intend by referring, for example, to Lebesgue measure on $(\mathbb{R}, \mathscr{B}(\mathbb{R}))$ or to Lebesgue measure on $(\mathbb{R}, \mathscr{M}_{\lambda^*})$. We will deal most often with Lebesgue measure on the Borel sets; its relation to the other version of Lebesgue measure is treated in Sect. 1.5.

Two questions arise immediately. Is every subset of $\mathbb{R}$ Lebesgue measurable? Is every Lebesgue measurable set a Borel set? The answer to each of these questions is no; see Sects. 1.4 and 2.1 for details.

We close this section with a technique for constructing and representing all finite measures on $(\mathbb{R},\mathscr{B}(\mathbb{R}))$. We begin with the following elementary fact.

Proposition 1.3.9. *Let μ be a finite measure on $(\mathbb{R},\mathscr{B}(\mathbb{R}))$, and let $F_\mu : \mathbb{R} \to \mathbb{R}$ be defined by $F_\mu(x) = \mu((-\infty,x])$. Then F_μ is bounded, nondecreasing, and right-continuous, and satisfies $\lim_{x\to-\infty} F_\mu(x) = 0$.*

Proof. It follows from Proposition 1.2.2 that $0 \le \mu((-\infty,x]) \le \mu(\mathbb{R})$ holds for all x in $\mathbb{R}$ and that $\mu((-\infty,x]) \le \mu((-\infty,y])$ holds for all x and y in $\mathbb{R}$ such that $x \le y$; hence F_μ is bounded and nondecreasing. Next suppose that $x \in \mathbb{R}$ and that $\{x_n\}$ is the sequence defined by $x_n = x + 1/n$. Then $(-\infty,x] = \cap_{n=1}^\infty(-\infty,x_n]$, and so Proposition 1.2.5 implies that $F_\mu(x) = \lim_n F_\mu(x_n)$. The right continuity of F_μ follows (note that if $x < y < x_n$, then, since F_μ is nondecreasing, $|F_\mu(y) - F_\mu(x)| \le |F_\mu(x_n) - F_\mu(x)|$). A similar argument shows that $\lim_{x\to-\infty} F_\mu(x) = 0$. □

Let μ and F_μ be as in Proposition 1.3.9. The interval $(a,b]$ is the difference of the intervals $(-\infty,b]$ and $(-\infty,a]$, and so Proposition 1.2.2 implies that

$$\mu((a,b]) = F_\mu(b) - F_\mu(a). \tag{6}$$

Since F_μ is bounded and nondecreasing, the limit of $F_\mu(t)$ as t approaches x from the left exists for each x in $\mathbb{R}$; this limit is equal to $\sup\{F_\mu(t) : t < x\}$ and will be denoted by $F_\mu(x-)$. Now let $\{a_n\}$ be a sequence that increases to the real number b; if we apply Eq. (6) to each interval $(a_n,b]$ and then use Proposition 1.2.5, we find that

$$\mu(\{b\}) = F_\mu(b) - F_\mu(b-). \tag{7}$$

Consequently F_μ is continuous at b if $\mu(\{b\}) = 0$, and is discontinuous there, with a jump of size $\mu(\{b\})$ in its graph, if $\mu(\{b\}) \ne 0$. Thus the measure μ is continuous (see Sect. 1.2) if and only if the function F_μ is continuous.

Equations (6) and (7) allow one to use F_μ to recover the measure under μ of certain subsets of $\mathbb{R}$ (see also Exercise 4); however, the following proposition allows us to say more, namely that the measure under μ of *every* Borel subset of $\mathbb{R}$ is in fact determined by F_μ.

Proposition 1.3.10. *For each bounded, nondecreasing, and right-continuous function $F : \mathbb{R} \to \mathbb{R}$ that satisfies $\lim_{x\to-\infty} F(x) = 0$, there is a unique finite measure μ on $(\mathbb{R},\mathscr{B}(\mathbb{R}))$ such that $F(x) = \mu((-\infty,x])$ holds at each x in $\mathbb{R}$.*

Proof. Let F be as in the statement of the proposition. We begin by constructing the required measure μ. Define a function $\mu^* : \mathscr{P}(\mathbb{R}) \to [0,+\infty]$ by letting $\mu^*(A)$ be the infimum of the set of sums $\sum_{n=1}^\infty (F(b_n) - F(a_n))$, where $\{(a_n,b_n]\}$ ranges over the set of sequences of half-open intervals that cover A, in the sense that $A \subseteq \cup_{n=1}^{\infty}(a_n,b_n]$. Then μ^* is an outer measure on $\mathbb{R}$; the reader can check this by modifying some of the arguments used in the proof of Proposition 1.3.2.

Next we verify that $\mu^*((-\infty,x]) = F(x)$ holds for each x in $\mathbb{R}$. The inequality $\mu^*((-\infty,x]) \le F(x)$ holds, since $(-\infty,x]$ can be covered by the intervals in the sequence $\{(x-n,x-n+1]\}_{n=1}^{\infty}$, for which we have $\sum_{n=1}^{\infty}(F(x-n+1)-F(x-n)) = F(x)$. We turn to the reverse inequality. Let $\{(a_n,b_n]\}$ be a sequence that covers $(-\infty,x]$, and let ε be a positive number. Use the fact that $\lim_{t\to-\infty} F(t) = 0$ to choose a number t such that $t < x$ and $F(t) < \varepsilon$, and for each n use the right continuity of F to choose a positive number δ_n such that $F(b_n+\delta_n) < F(b_n)+\varepsilon/2^n$. Then the interval $[t,x]$ is compact, each interval $(a_n,b_n+\delta_n)$ is open, $[t,x] \subseteq \cup_{n=1}^{\infty}(a_n,b_n+\delta_n)$, and $\sum_n(F(b_n+\delta_n)-F(a_n)) \le \sum_n(F(b_n)-F(a_n))+\varepsilon$. The compactness of $[t,x]$ implies that there is a positive integer N such that $[t,x] \subseteq \cup_{n=1}^{N}(a_n,b_n+\delta_n)$. It follows that $(t,x]$ is the union of a finite collection of disjoint intervals $(c_j,d_j]$, each of which is included in some $(a_n,b_n+\delta_n]$. Consequently

$$F(x)-F(t) = \sum_j (F(d_j)-F(c_j)) \le \sum_{n=1}^{\infty}(F(b_n+\delta_n)-F(a_n)),$$

and so

$$F(x)-\varepsilon \le \sum_{n=1}^{\infty}(F(b_n)-F(a_n))+\varepsilon.$$

Since ε and the sequence $\{(a_n,b_n]\}$ are arbitrary, the inequality $F(x) \le \mu^*((-\infty,x])$ follows. With this we have shown that $F(x) = \mu^*((-\infty,x])$.

The reader should check that the proof of Proposition 1.3.7 can be modified so as to show that each interval $(-\infty,b]$ is μ^*-measurable and then that each Borel subset of $\mathbb{R}$ is μ^*-measurable.

Let μ be the restriction of μ^* to $\mathscr{B}(\mathbb{R})$. The preceding steps of our proof, together with Theorem 1.3.6, show that μ is a measure and that it satisfies $\mu((-\infty,x]) = F(x)$ at each x in $\mathbb{R}$. Since F is bounded, while $\mu(\mathbb{R}) = \lim_{n\to\infty}\mu((-\infty,n]) = \lim_{n\to\infty}F(n)$ (Proposition 1.2.5), the measure μ is finite.

Finally we check the uniqueness of μ. Let μ be as constructed above, and let ν be a possibly different measure such that $\nu((-\infty,x]) = F(x)$ holds for each x in $\mathbb{R}$. We first show that

$$\nu(A) \le \mu(A) \tag{8}$$

is true for each Borel subset A of $\mathbb{R}$. To see this, note that if A is a Borel set and if $\{(a_n,b_n]\}$ is a sequence such that $A \subseteq \cup_n(a_n,b_n]$, then (according to (6), applied to ν)

$$\nu(A) \le \sum_n \nu((a_n,b_n]) = \sum_n (F(b_n)-F(a_n)). \tag{9}$$

Since $\mu^*(A)$ was defined to be the infimum of the set of values that can occur as sums on the right side of (9), inequality (8) follows. If we apply inequality (8) to A and to A^c, we find

$$\nu(\mathbb{R}) = \nu(A) + \nu(A^c) \leq \mu(A) + \mu(A^c) = \mu(\mathbb{R}).$$

Since $\nu(\mathbb{R}) = \mu(\mathbb{R}) < +\infty$, it follows that $\nu(A)$ and $\nu(A^c)$ are equal to $\mu(A)$ and $\mu(A^c)$, respectively. With this the proof that $\nu = \mu$ is complete. $\square$

The uniqueness assertion on Proposition 1.3.10 can also be proved by means of other standard techniques; see, for example, the discussion following the proof of Corollary 1.6.3.

Exercises

1. Define functions $\mu_1^*, \ldots, \mu_6^*$ on $\mathscr{P}(\mathbb{R})$ by

$$\mu_1^*(A) = \begin{cases} 0 & \text{if } A \text{ is empty,} \\ 1 & \text{if } A \text{ is nonempty,} \end{cases}$$

$$\mu_2^*(A) = \begin{cases} 0 & \text{if } A \text{ is empty,} \\ +\infty & \text{if } A \text{ is nonempty,} \end{cases}$$

$$\mu_3^*(A) = \begin{cases} 0 & \text{if } A \text{ is bounded,} \\ 1 & \text{if } A \text{ is unbounded,} \end{cases}$$

$$\mu_4^*(A) = \begin{cases} 0 & \text{if } A \text{ is empty,} \\ 1 & \text{if } A \text{ is nonempty and bounded,} \\ +\infty & \text{if } A \text{ is unbounded,} \end{cases}$$

$$\mu_5^*(A) = \begin{cases} 0 & \text{if } A \text{ is countable,} \\ 1 & \text{if } A \text{ is uncountable,} \end{cases}$$

$$\mu_6^*(A) = \begin{cases} 0 & \text{if } A \text{ is countable, and} \\ +\infty & \text{if } A \text{ is uncountable.} \end{cases}$$

   (a) Which of μ_2^*, μ_3^*, μ_4^*, and μ_6^* are outer measures? (We noted in Examples 1.3.1(a) and 1.3.1(b) that μ_1^* and μ_5^* are outer measures.)
   (b) For each i such that μ_i^* is an outer measure determine the μ_i^*-measurable subsets of $\mathbb{R}$.
2. Let C be a countable subset of $\mathbb{R}$. Using only the definition of λ^*, show that $\lambda^*(C) = 0$.

3. Show that for each subset A of $\mathbb{R}$ there is a Borel subset B of $\mathbb{R}$ that includes A and satisfies $\lambda(B) = \lambda^*(A)$.
4. Let $F\colon \mathbb{R} \to \mathbb{R}$ be a bounded, nondecreasing, and right-continuous function that satisfies $\lim_{x\to-\infty} F(x) = 0$, and let μ be the measure on $(\mathbb{R}, \mathscr{B}(\mathbb{R}))$ that is associated to F by Proposition 1.3.10. Show that if a and b belong to $\mathbb{R}$ and satisfy $a < b$, then

$$\mu((-\infty,b)) = F(b-),$$
$$\mu((a,b)) = F(b-) - F(a),$$
$$\mu([a,b]) = F(b) - F(a-), \text{ and}$$
$$\mu([a,b)) = F(b-) - F(a-).$$

5. Let X be a set, let $\mathscr{A}$ be an algebra of subsets of X, and let μ be a finitely additive measure on $\mathscr{A}$. For each subset A of X let $\mu^*(A)$ be the infimum of the set of sums $\sum_{k=1}^{\infty} \mu(A_k)$, where $\{A_k\}$ ranges over the sequences of sets in $\mathscr{A}$ for which $A \subseteq \cup_{k=1}^{\infty} A_k$.
   (a) Show that μ^* is an outer measure on X.
   (b) Show that each set in $\mathscr{A}$ is μ^*-measurable.
   (c) Show that if μ is countably additive (in the sense that $\mu(\cup_k A_k) = \sum_k \mu(A_k)$ holds whenever $\{A_k\}$ is a sequence of disjoint sets in $\mathscr{A}$ for which $\cup_k A_k$ belongs to $\mathscr{A}$), then each A in $\mathscr{A}$ satisfies $\mu(A) = \mu^*(A)$.
   (d) Conclude that if μ is a countably additive measure on the algebra $\mathscr{A}$, then there is a countably additive measure on $\sigma(\mathscr{A})$ that agrees with μ on $\mathscr{A}$.
6. (Continuation.) Let X, $\mathscr{A}$, μ, and μ^* be as in Exercise 5, and assume that μ is countably additive.
   (a) Show that if ν is a countably additive measure on $\sigma(\mathscr{A})$ that agrees with μ on $\mathscr{A}$, then $\nu(A) \le \mu^*(A)$ holds for each A in $\sigma(A)$.
   (b) Conclude that if μ is finite (or if X is the union of a sequence of sets that belong to $\mathscr{A}$ and have finite measure under μ), then μ can be extended to a countably additive measure on $\sigma(\mathscr{A})$ in only one way.
7. Show that a subset B of $\mathbb{R}$ is Lebesgue measurable if and only if

$$\lambda^*(I) = \lambda^*(I \cap B) + \lambda^*(I \cap B^c)$$

   holds for each open subinterval I of $\mathbb{R}$.
8. Let I be a bounded subinterval of $\mathbb{R}$. Show that a subset B of I is Lebesgue measurable if and only if it satisfies $\lambda^*(I) = \lambda^*(B) + \lambda^*(I \cap B^c)$.
9. Let λ^* be Lebesgue outer measure on $\mathbb{R}$, and let π be the projection of $\mathbb{R}^2$ onto $\mathbb{R}$ given by $\pi(x,y) = x$. Define a function $\mu^*\colon \mathscr{P}(\mathbb{R}^2) \to [0,+\infty]$ by $\mu^*(A) = \lambda^*(\pi(A))$.
   (a) Show that μ^* is an outer measure on $\mathbb{R}^2$.
   (b) Show that a subset B of $\mathbb{R}^2$ is measurable for the outer measure μ^* defined in this exercise if and only if there are Lebesgue measurable subsets B_0 and B_1 of $\mathbb{R}$ such that $B_0 \subseteq B_1$, $\lambda^*(B_1 - B_0) = 0$, and $B_0 \times \mathbb{R} \subseteq B \subseteq B_1 \times \mathbb{R}$.

1.4 Lebesgue Measure

This section contains a number of the basic properties of Lebesgue measure on $\mathbb{R}^d$. The reader who wants to move quickly on to Chap. 2 might restrict his or her attention to Proposition 1.4.1, Proposition 1.4.4, and Theorem 1.4.9.

Proposition 1.4.1. *Let A be a Lebesgue measurable subset of $\mathbb{R}^d$. Then*

(a) $\lambda(A) = \inf\{\lambda(U) : U$ *is open and* $A \subseteq U\}$, *and*
(b) $\lambda(A) = \sup\{\lambda(K) : K$ *is compact and* $K \subseteq A\}$.

Proposition 1.4.1 can be put more briefly, namely as the assertion that Lebesgue measure is *regular*. In the interest of simplicity, however, we will delay the study and even the definition of regularity until Sect. 1.5 and Chap. 7.

Proof. Note that the monotonicity of λ implies that

$$\lambda(A) \le \inf\{\lambda(U) : U \text{ is open and } A \subseteq U\}$$

and

$$\lambda(A) \ge \sup\{\lambda(K) : K \text{ is compact and } K \subseteq A\}.$$

Hence we need only prove the reverse inequalities.

We begin with part (a). Since the required equality clearly holds if $\lambda(A) = +\infty$, we can assume that $\lambda(A) < +\infty$. Let ε be an arbitrary positive number. Then according to the definition of Lebesgue measure, there is a sequence $\{R_i\}$ of open d-dimensional intervals such that $A \subseteq \cup_i R_i$ and $\sum_i \text{vol}(R_i) < \lambda(A) + \varepsilon$. Let U be the union of these intervals. Then U is open, $A \subseteq U$, and (see Propositions 1.2.4 and 1.3.4)

$$\lambda(U) \le \sum_i \lambda(R_i) = \sum_i \text{vol}(R_i) < \lambda(A) + \varepsilon.$$

Since ε is arbitrary, part (a) is proved.

We turn to part (b) and deal first with the case where A is bounded. Let C be a closed and bounded set that includes A, and let ε be an arbitrary positive number. Use part (a) to choose an open set U that includes $C - A$ and satisfies

$$\lambda(U) < \lambda(C - A) + \varepsilon. \tag{1}$$

Let $K = C - U$. (Drawing a sketch might help the reader.) Then K is a closed and bounded (and hence compact) subset of A; furthermore, $C \subseteq K \cup U$ and so

$$\lambda(C) \le \lambda(K) + \lambda(U). \tag{2}$$

Inequalities (1) and (2) (and the fact that $\lambda(C - A) = \lambda(C) - \lambda(A)$) imply that $\lambda(A) - \varepsilon < \lambda(K)$. Since ε was arbitrary, part (b) is proved in the case where A is bounded.

Finally, consider the case where A is not bounded. Suppose that b is a real number less than $\lambda(A)$; we will produce a compact subset K of A such that $b < \lambda(K)$. Let $\{A_j\}$ be an increasing sequence of bounded measurable subsets of A such that $A = \cup_j A_j$ (for example, we might let A_j be the intersection of A with the closed ball of radius j about the origin). Proposition 1.2.5 implies that $\lambda(A) = \lim_j \lambda(A_j)$, and so we can choose j_0 such that $\lambda(A_{j_0}) > b$. Now apply to A_{j_0} the weakened form of part (b) that was proved in the preceding paragraph; this gives a compact subset K of A_{j_0} (and hence of A) such that $\lambda(K) > b$. Since b was an arbitrary number less than $\lambda(A)$, the proof is complete. □

The following lemma will be needed for the proof of Proposition 1.4.3. In this lemma we will be dealing with a certain collection of half-open cubes, namely with those that have the form

$$\{(x_1, \ldots, x_d) : j_i 2^{-k} \le x_i < (j_i + 1)2^{-k} \text{ for } i = 1, \ldots, d\} \tag{3}$$

for some integers $j_1, \ldots, j_d$ and some positive integer k.

Lemma 1.4.2. *Each open subset of $\mathbb{R}^d$ is the union of a countable disjoint collection of half-open cubes, each of which is of the form given in expression* (3)*.*

Proof. For each positive integer k let $\mathscr{C}_k$ be the collection of all cubes of the form

$$\{(x_1, \ldots, x_d) : j_i 2^{-k} \le x_i < (j_i + 1)2^{-k} \text{ for } i = 1, \ldots, d\},$$

where $j_1, \ldots, j_d$ are arbitrary integers. It is easy to see that

(a) each $\mathscr{C}_k$ is a countable partition of $\mathbb{R}^d$, and
(b) if $k_1 < k_2$, then each cube in $\mathscr{C}_{k_2}$ is included in some cube in $\mathscr{C}_{k_1}$.

The reader should keep these facts about the family $\{\mathscr{C}_k\}$ in mind when checking that the collection $\mathscr{D}$ defined below has the properties claimed for it.

Suppose that U is an open subset of $\mathbb{R}^d$. We construct a collection $\mathscr{D}$ of cubes inductively by letting $\mathscr{D}$ be empty at the start, and then at step k (for $k = 1, 2, \ldots$) adding to $\mathscr{D}$ those cubes in $\mathscr{C}_k$ that are included in U but are disjoint from all the cubes put into $\mathscr{D}$ at earlier steps. It is clear that $\mathscr{D}$ is a countable disjoint collection of cubes whose union is included in U. It remains only to check that its union includes U. Let x be a member of U. Since U is open, the cube in $\mathscr{C}_k$ that contains x is included in U if k is sufficiently large. Let k_0 be the smallest such k. Then the cube in $\mathscr{C}_{k_0}$ that contains x belongs to $\mathscr{D}$, and so x belongs to the union of the cubes in $\mathscr{D}$. □

Proposition 1.4.3. *Lebesgue measure is the only measure on $(\mathbb{R}^d, \mathscr{B}(\mathbb{R}^d))$ that assigns to each d-dimensional interval, or even to each half-open cube of the form given in expression* (3)*, its volume.*

Proof. That Lebesgue measure does assign to each d-dimensional interval its volume was noted in Sect. 1.3. So we need only assume that μ is a measure on $(\mathbb{R}^d, \mathscr{B}(\mathbb{R}^d))$ that assigns to each cube of the form given in expression (3) its volume and prove that $\mu = \lambda$. First suppose that U is an open subset of $\mathbb{R}^d$. Then according to Lemma 1.4.2 there is a disjoint sequence $\{C_j\}$ of half-open cubes that have the form given in expression (3) and whose union is U, and so

$$\mu(U) = \sum_j \mu(C_j) = \sum_j \lambda(C_j) = \lambda(U);$$

hence μ and λ agree on the open subsets of $\mathbb{R}^d$. Next suppose that A is an arbitrary Borel subset of $\mathbb{R}^d$. If U is an open subset of $\mathbb{R}^d$ that includes A, then $\mu(A) \le \mu(U) = \lambda(U)$; it follows that $\mu(A) \le \inf\{\lambda(U) : U \text{ is open and } A \subseteq U\}$. The regularity of λ (Proposition 1.4.1) now implies that

$$\mu(A) \le \lambda(A). \tag{4}$$

We need to show that this inequality can be replaced with an equality. First suppose that A is a bounded Borel subset of $\mathbb{R}^d$ and that V is a bounded open set that includes A. Then inequality (4), applied to the sets A and $V - A$, implies that

$$\mu(V) = \mu(A) + \mu(V - A) \le \lambda(A) + \lambda(V - A) = \lambda(V);$$

since the extreme members of this inequality are equal, and since $\mu(A)$ and $\mu(V - A)$ are no larger than $\lambda(A)$ and $\lambda(V - A)$, respectively, it follows that $\mu(A)$ and $\lambda(A)$ are equal. Finally, an arbitrary Borel subset A of $\mathbb{R}^d$ is the union of a sequence of disjoint bounded Borel sets and so must satisfy $\mu(A) = \lambda(A)$. □

For each element x and subset A of $\mathbb{R}^d$ we will denote by $A + x$ the subset of $\mathbb{R}^d$ defined by

$$A + x = \{y \in \mathbb{R}^d : y = a + x \text{ for some } a \text{ in } A\};$$

the set $A + x$ is called the *translate* of A by x. We turn to the invariance of Lebesgue measure under such translations.

Proposition 1.4.4. *Lebesgue outer measure on $\mathbb{R}^d$ is translation invariant, in the sense that if $x \in \mathbb{R}^d$ and $A \subseteq \mathbb{R}^d$, then $\lambda^*(A) = \lambda^*(A + x)$. Furthermore, a subset B of $\mathbb{R}^d$ is Lebesgue measurable if and only if $B + x$ is Lebesgue measurable.*

Proof. The equality of $\lambda^*(A)$ and $\lambda^*(A + x)$ follows from the definition of λ^* and the fact that the volume of a d-dimensional interval is invariant under translation. The second assertion follows from the first, together with the definition of a Lebesgue measurable set—note that a set B satisfies

$$\lambda^*(A - x) = \lambda^*((A - x) \cap B) + \lambda^*((A - x) \cap B^c)$$

for all sets $A - x$ if and only if $B + x$ satisfies

$$\lambda^*(A) = \lambda^*(A \cap (B + x)) + \lambda^*(A \cap (B + x)^c)$$

for all sets A. □

Lebesgue measure on $(\mathbb{R}^d, \mathscr{B}(\mathbb{R}^d))$ is characterized up to constant multiples by the following result; see Chap. 9 for analogous results that hold in more general situations.

Proposition 1.4.5. *Let μ be a nonzero measure on $(\mathbb{R}^d, \mathscr{B}(\mathbb{R}^d))$ that is finite on the bounded Borel subsets of $\mathbb{R}^d$ and is translation invariant, in the sense that $\mu(A) = \mu(A+x)$ holds for each A in $\mathscr{B}(\mathbb{R}^d)$ and each x in $\mathbb{R}^d$. Then there is a positive number c such that $\mu(A) = c\lambda(A)$ holds for each A in $\mathscr{B}(\mathbb{R}^d)$.*

Note that for the concept of translation invariance for measures on $(\mathbb{R}^d, \mathscr{B}(\mathbb{R}^d))$ to make sense, the Borel σ-algebra on $\mathbb{R}^d$ must be translation invariant, in the sense that if $A \in \mathscr{B}(\mathbb{R}^d)$ and $x \in \mathbb{R}^d$, then $A + x \in \mathscr{B}(\mathbb{R}^d)$. To check this translation invariance of $\mathscr{B}(\mathbb{R}^d)$, note that $\{A \subseteq \mathbb{R}^d : A + x \in \mathscr{B}(\mathbb{R}^d)\}$ is a σ-algebra that contains the open sets and hence includes $\mathscr{B}(\mathbb{R}^d)$.

Proof. Let $C = \{(x_1, \dots, x_d) : 0 \le x_i < 1 \text{ for each } i\}$, and let $c = \mu(C)$. Then c is finite (since μ is finite on the bounded Borel sets) and positive (if it were 0, then $\mathbb{R}^d$, as the union of a sequence of translates of C, would have measure zero under μ). Define a measure ν on $\mathscr{B}(\mathbb{R}^d)$ by letting $\nu(A) = (1/c)\mu(A)$ hold for each A in $\mathscr{B}(\mathbb{R}^d)$. Then ν is translation invariant, and it assigns to the set C defined above its Lebesgue measure, namely 1. If D is a half-open cube that has the form given in expression (3) and whose edges have length 2^{-k}, then C is the union of 2^{dk} translates of D, and so

$$2^{dk}\nu(D) = \nu(C) = \lambda(C) = 2^{dk}\lambda(D);$$

thus ν and λ agree on all such cubes. Proposition 1.4.3 now implies that $\nu = \lambda$ and hence that $\mu = c\lambda$. □

Example 1.4.6 (The Cantor Set). We should note a few facts about the *Cantor set*, a set which turns out to be a useful source of examples. Recall that it is defined as follows. Let K_0 be the interval $[0,1]$. Form K_1 by removing from K_0 the interval $(1/3, 2/3)$. Thus $K_1 = [0, 1/3] \cup [2/3, 1]$. Continue this procedure, forming K_n by removing from K_{n-1} the open middle third of each of the intervals making up K_{n-1}. Thus K_n is the union of 2^n disjoint closed intervals, each of length $(1/3)^n$. The Cantor set (which we will temporarily denote by K) is the set of points that remain; thus $K = \cap_n K_n$.

Of course K is closed and bounded. Furthermore, K has no interior points, since an open interval included in K would for each n be included in one of the intervals making up K_n and so would have length at most $(1/3)^n$. The cardinality of K is that of the continuum: it is easy to check that the map that assigns to a sequence $\{z_n\}$ of 0's and 1's the number $\sum_{n=1}^{\infty} 2z_n/3^n$ is a bijection of the set of all such sequences onto K; hence the cardinality of K is that of the set of all sequences of 0's and 1's and so that of the continuum (see Appendix A). □

Proposition 1.4.7. *The Cantor set is a compact set that has the cardinality of the continuum but has Lebesgue measure zero.*

Proof. We have already noted that the Cantor set (again call it K) is compact and has the cardinality of the continuum. To compute the measure of K, note that for each n it is included in the set K_n constructed above and that $\lambda(K_n) = (2/3)^n$. Thus $\lambda(K) \le (2/3)^n$ holds for each n, and so $\lambda(K)$ must be zero. (For an alternative proof, check that the sum of the measures of the intervals removed from $[0,1]$ during the construction of K is the sum of the geometric series

$$\frac{1}{3} + \frac{2}{3}\cdot\frac{1}{3} + \left(\frac{2}{3}\right)^2\cdot\frac{1}{3} + \left(\frac{2}{3}\right)^3\cdot\frac{1}{3} + \dots,$$

and so is 1.) □

Example 1.4.8 (A Nonmeasurable Set). We now return to one of the promises made in Sect. 1.3 and prove that there is a subset of $\mathbb{R}$ that is not Lebesgue measurable. Note that our proof of this uses the axiom of choice.[4] Whether the use of this axiom is essential was an open question until the mid-1960s, when R.M. Solovay showed that if a certain consistency assumption holds, then the existence of a subset of $\mathbb{R}$ that is not Lebesgue measurable cannot be proved from the axioms of Zermelo–Frankel set theory without the use of the axiom of choice.[5]

Theorem 1.4.9. *There is a subset of* $\mathbb{R}$*, and in fact of the interval* $(0,1)$*, that is not Lebesgue measurable.*

Proof. Define a relation $\sim$ on $\mathbb{R}$ by letting $x \sim y$ hold if and only if $x - y$ is rational. It is easy to check that $\sim$ is an equivalence relation: it is reflexive ($x \sim x$ holds for each x), symmetric ($x \sim y$ implies $y \sim x$), and transitive ($x \sim y$ and $y \sim z$ imply $x \sim z$). Note that each equivalence class under $\sim$ has the form $\mathbb{Q} + x$ for some x and so is dense in $\mathbb{R}$. Since these equivalence classes are disjoint, and since each intersects the interval $(0,1)$, we can use the axiom of choice to form a subset E of $(0,1)$ that contains exactly one element from each equivalence class. We will prove that the set E is not Lebesgue measurable.

Let $\{r_n\}$ be an enumeration of the rational numbers in the interval $(-1,1)$, and for each n let $E_n = E + r_n$. We will check that

(a) the sets E_n are disjoint,
(b) $\cup_n E_n$ is included in the interval $(-1,2)$, and
(c) the interval $(0,1)$ is included in $\cup_n E_n$.

To check (a), note that if $E_m \cap E_n \neq \varnothing$, then there are elements e and e' of E such that $e + r_m = e' + r_n$; it follows that $e \sim e'$ and hence that $e = e'$ and $m = n$. Thus (a) is proved. Assertion (b) follows from the inclusion $E \subseteq (0,1)$ and the fact that each term of the sequence $\{r_n\}$ belongs to $(-1,1)$. Now consider assertion (c). Let x be

---

[4] See items A.12 and A.13 in Appendix A.

[5] For details, see Solovay [110].

an arbitrary member of $(0,1)$, and let e be the member of E that satisfies $x \sim e$. Then $x - e$ is rational and belongs to $(-1,1)$ (recall that both x and e belong to $(0,1)$) and so has the form r_n for some n. Hence $x \in E_n$, and assertion (c) is proved.

Suppose that the set E is Lebesgue measurable. Then for each n the set E_n is measurable (Proposition 1.4.4), and so property (a) above implies that

$$\lambda(\cup_n E_n) = \sum_n \lambda(E_n);$$

furthermore, the translation invariance of λ implies that $\lambda(E_n) = \lambda(E)$ holds for each n. Hence if $\lambda(E) = 0$, then $\lambda(\cup_n E_n) = 0$, contradicting assertion (c) above, while if $\lambda(E) \neq 0$, then $\lambda(\cup_n E_n) = +\infty$, contradicting assertion (b). Thus the assumption that E is measurable leads to a contradiction, and the proof is complete. □

Let A be a subset of $\mathbb{R}$. Then $\mathrm{diff}(A)$ is the subset of $\mathbb{R}$ defined by

$$\mathrm{diff}(A) = \{x - y : x \in A \text{ and } y \in A\}.$$

The following fact about such sets is occasionally useful.

Proposition 1.4.10. *Let A be a Lebesgue measurable subset of $\mathbb{R}$ such that $\lambda(A) > 0$. Then* $\mathrm{diff}(A)$ *includes an open interval that contains* 0.

Proof. According to Proposition 1.4.1, there is a compact subset K of A such that $\lambda(K) > 0$. Since $\mathrm{diff}(K)$ is then included in $\mathrm{diff}(A)$, it is enough to prove that $\mathrm{diff}(K)$ includes an open interval that contains 0. Note that a real number x belongs to $\mathrm{diff}(K)$ if and only if K intersects $x + K$; thus it suffices to prove that if $|x|$ is sufficiently small, then K intersects $x + K$.

Use Proposition 1.4.1 to choose an open set U such that $K \subseteq U$ and $\lambda(U) < 2\lambda(K)$. The distances between the points in K and the points outside U are bounded away from 0 (since the distance from a point x of U to the complement of U is a continuous strictly positive function of x and so has a positive minimum on the compact set K; see D.27 and D.18). Thus there is a positive number ε such that if $|x| < \varepsilon$, then $x + K$ is included in U. Suppose that $|x| < \varepsilon$. If $x + K$ were disjoint from K, then it would follow from the translation invariance of λ and the relation $x + K \subseteq U$ that

$$2\lambda(K) = \lambda(K) + \lambda(x + K) = \lambda(K \cup (x + K)) \leq \lambda(U).$$

However this contradicts the inequality $\lambda(U) < 2\lambda(K)$, and so K and $x + K$ cannot be disjoint. Therefore, $x \in \mathrm{diff}(K)$. Consequently the interval $(-\varepsilon, \varepsilon)$ is included in $\mathrm{diff}(K)$, and thus in $\mathrm{diff}(A)$. □

We can use Proposition 1.4.10, plus a modification of the proof of Theorem 1.4.9, to prove the following rather strong result (see the remark at the end of this section and the one following the proof of Proposition 1.5.4).

Proposition 1.4.11. *There is a subset A of $\mathbb{R}$ such that each Lebesgue measurable set that is included in A or in A^c has Lebesgue measure zero.*

Proof. Define subsets G, G_0, and G_1 of $\mathbb{R}$ by

$$G = \{x : x = r + n\sqrt{2} \text{ for some } r \text{ in } \mathbb{Q} \text{ and } n \text{ in } \mathbb{Z}\},$$

$$G_0 = \{x : x = r + 2n\sqrt{2} \text{ for some } r \text{ in } \mathbb{Q} \text{ and } n \text{ in } \mathbb{Z}\}, \text{ and}$$

$$G_1 = \{x : x = r + (2n+1)\sqrt{2} \text{ for some } r \text{ in } \mathbb{Q} \text{ and } n \text{ in } \mathbb{Z}\}.$$

It is easy to see that G and G_0 are subgroups of $\mathbb{R}$ (under addition), that G_0 and G_1 are disjoint, that $G_1 = G_0 + \sqrt{2}$, and that $G = G_0 \cup G_1$. Define a relation $\sim$ on $\mathbb{R}$ by letting $x \sim y$ hold when $x - y \in G$; the relation $\sim$ is then an equivalence relation on $\mathbb{R}$. Use the axiom of choice to form a subset E of $\mathbb{R}$ that contains exactly one representative of each equivalence class of $\sim$. Let $A = E + G_0$ (that is, let A consist of the points that have the form $e + g_0$ for some e in E and some g_0 in G_0).

We now show that there does not exist a Lebesgue measurable subset B of A such that $\lambda(B) > 0$. For this let us assume that such a set exists; we will derive a contradiction. Proposition 1.4.10 implies that there is an interval $(-\varepsilon, \varepsilon)$ that is included in $\text{diff}(B)$ and hence in $\text{diff}(A)$. Since G_1 is dense in $\mathbb{R}$, it meets the interval $(-\varepsilon, \varepsilon)$ and hence meets $\text{diff}(A)$. This, however, is impossible, since each element of $\text{diff}(A)$ is of the form $e_1 - e_2 + g_0$ (where e_1 and e_2 belong to E and g_0 belongs to G_0) and so cannot belong to G_1 (the relation $e_1 - e_2 + g_0 = g_1$ would imply that $e_1 = e_2$ and $g_0 = g_1$, contradicting the disjointness of G_0 and G_1). This completes our proof that every Lebesgue measurable subset of A must have Lebesgue measure zero.

It is easy to check that $A^c = E + G_1$ and hence that $A^c = A + \sqrt{2}$. It follows that each Lebesgue measurable subset of A^c is of the form $B + \sqrt{2}$ for some Lebesgue measurable subset B of A. Since A has no Lebesgue measurable subsets of positive measure, it follows that A^c also has no such subsets, and with this the proof is complete. □

Note that the set A of Proposition 1.4.11 is not Lebesgue measurable: if it were, then both A and A^c would include (in fact, would be) Lebesgue measurable sets of positive Lebesgue measure. Thus we could have presented Theorem 1.4.9 as a corollary of Proposition 1.4.11. (Of course, the proof of Theorem 1.4.9 presented earlier is simpler than the proofs of Propositions 1.4.10 and 1.4.11 taken together and is in fact a classical and well-known argument; hence it was included.)

Exercises

1. Prove that under Lebesgue measure on $\mathbb{R}^2$
   (a) every straight line has measure zero, and
   (b) every circle has measure zero.

2. Let A be a subset of $\mathbb{R}^d$. Show that the conditions

   (i) A is Lebesgue measurable,
   (ii) A is the union of an F_σ and a set of Lebesgue measure zero, and
   (iii) there is a set B that is an F_σ and satisfies $\lambda^*(A \triangle B) = 0$

   are equivalent.
3. Let T be a rotation of $\mathbb{R}^2$ about the origin (or, more generally, a linear map from $\mathbb{R}^d$ to $\mathbb{R}^d$ that preserves distances).
   (a) Show that a subset A of $\mathbb{R}^2$ (or of $\mathbb{R}^d$) is Borel if and only if $T(A)$ is Borel. (Hint: See the remark following the statement of Proposition 1.4.5.)
   (b) Show that each Borel subset A of $\mathbb{R}^2$ (or of $\mathbb{R}^d$) satisfies $\lambda(A) = \lambda(T(A))$. (Hint: Use Proposition 1.4.5.)
4. Show that for each number α that satisfies $0 < \alpha < 1$ there is a closed subset C of $[0,1]$ that satisfies $\lambda(C) = \alpha$ and includes no nonempty open set. (Hint: Imitate the construction of the Cantor set.)
5. Show that there is a Borel subset A of $\mathbb{R}$ such that $0 < \lambda(I \cap A) < \lambda(I)$ holds whenever I is a bounded open subinterval of $\mathbb{R}$.
6. Show that if B is a subset of $\mathbb{R}$ that satisfies $\lambda^*(B) > 0$, then B includes a set that is not Lebesgue measurable. (Hint: Use Proposition 1.4.11.)
7. Show that there exists a decreasing sequence $\{A_n\}$ of subsets of $[0,1]$ such that $\lambda^*(A_n) = 1$ holds for each n, but for which $\cap_n A_n = \varnothing$. (Hint: Let B be a Hamel basis[6] for $\mathbb{R}$ as a vector space over $\mathbb{Q}$, and let $\{B_n\}$ be a strictly increasing sequence of sets such that $B = \cup_n B_n$. For each n let V_n be the subspace of $\mathbb{R}$ spanned by B_n, and let $A_n = [0,1] \cap V_n^c$. Use Proposition 1.4.10 to show that each Borel subset of V_n has Lebesgue measure zero and hence that $\lambda^*(A_n) = 1$.)

1.5 Completeness and Regularity

Let $(X, \mathscr{A}, \mu)$ be a measure space. The measure μ (or the measure space $(X, \mathscr{A}, \mu)$) is *complete* if the relations $A \in \mathscr{A}$, $\mu(A) = 0$, and $B \subseteq A$ together imply that $B \in \mathscr{A}$. It is sometimes convenient to call a subset B of X *μ-negligible* (or *μ-null*) if there is a subset A of X such that $A \in \mathscr{A}$, $B \subseteq A$, and $\mu(A) = 0$. Thus the measure μ is complete if and only if every μ-negligible subset of X belongs to $\mathscr{A}$.

It follows from Proposition 1.3.5 that if μ^* is an outer measure on the set X and if $\mathscr{M}_{\mu^*}$ is the σ-algebra of all μ^*-measurable subsets of X, then the restriction of μ^* to $\mathscr{M}_{\mu^*}$ is complete. In particular, Lebesgue measure on the σ-algebra of

[6] This means that B spans $\mathbb{R}$ (i.e., that $\mathbb{R}$ is the smallest linear subspace of $\mathbb{R}$ that includes B) and that no proper subset of B spans $\mathbb{R}$. The axiom of choice implies that such a set B exists; see, for example, Lang [80, Section 5 of Chapter III].

Lebesgue measurable subsets of $\mathbb{R}^d$ is complete. On the other hand, as we will soon see, the restriction of Lebesgue measure to the σ-algebra of Borel subsets of $\mathbb{R}$ is not complete.

It is sometimes convenient to be able to deal with arbitrary subsets of sets of measure zero, and at such times complete measures are desirable. In many such situations the following construction proves useful.

Let $(X, \mathscr{A})$ be a measurable space, and let μ be a measure on $\mathscr{A}$. The *completion* of $\mathscr{A}$ under μ is the collection $\mathscr{A}_\mu$ of subsets A of X for which there are sets E and F in $\mathscr{A}$ such that

$$E \subseteq A \subseteq F \tag{1}$$

and

$$\mu(F - E) = 0. \tag{2}$$

A set that belongs to $\mathscr{A}_\mu$ is sometimes said to be μ*-measurable*.

Suppose that A, E, and F are as in the preceding paragraph. It follows immediately that $\mu(E) = \mu(F)$. Furthermore, if B is a subset of A that belongs to $\mathscr{A}$, then

$$\mu(B) \leq \mu(F) = \mu(E).$$

Hence

$$\mu(E) = \sup\{\mu(B) : B \in \mathscr{A} \text{ and } B \subseteq A\},$$

and so the common value of $\mu(E)$ and $\mu(F)$ depends only on the set A (and the measure μ), and not on the choice of sets E and F satisfying (1) and (2). Thus we can define a function $\overline{\mu}\colon \mathscr{A}_\mu \to [0, +\infty]$ by letting $\overline{\mu}(A)$ be the common value of $\mu(E)$ and $\mu(F)$, where E and F belong to $\mathscr{A}$ and satisfy (1) and (2). This function $\overline{\mu}$ is called the *completion* of μ.

Proposition 1.5.1. *Let $(X, \mathscr{A})$ be a measurable space, and let μ be a measure on $\mathscr{A}$. Then $\mathscr{A}_\mu$ is a σ-algebra on X that includes $\mathscr{A}$, and $\overline{\mu}$ is a measure on $\mathscr{A}_\mu$ that is complete and whose restriction to $\mathscr{A}$ is μ.*

Proof. It is clear that $\mathscr{A}_\mu$ includes $\mathscr{A}$ (for A in $\mathscr{A}$ let the sets E and F in (1) and (2) equal A), and in particular that $X \in \mathscr{A}_\mu$. Note that the relations $E \subseteq A \subseteq F$ and $\mu(F - E) = 0$ imply the relations $F^c \subseteq A^c \subseteq E^c$ and $\mu(E^c - F^c) = 0$; thus $\mathscr{A}_\mu$ is closed under complementation. Next suppose that $\{A_n\}$ is a sequence of sets in $\mathscr{A}_\mu$. For each n choose sets E_n and F_n in $\mathscr{A}$ such that $E_n \subseteq A_n \subseteq F_n$ and $\mu(F_n - E_n) = 0$. Then $\cup_n E_n$ and $\cup_n F_n$ belong to $\mathscr{A}$ and satisfy $\cup_n E_n \subseteq \cup_n A_n \subseteq \cup_n F_n$ and

$$\mu(\cup_n F_n - \cup_n E_n) \leq \mu(\cup_n (F_n - E_n)) \leq \sum_n \mu(F_n - E_n) = 0;$$

thus $\cup A_n$ belongs to $\mathscr{A}_\mu$. This completes the proof that $\mathscr{A}_\mu$ is a σ-algebra on X that includes $\mathscr{A}$.

Now consider the function $\overline{\mu}$. It is an extension of μ, since for A in $\mathscr{A}$ we can again let E and F equal A. It is clear that $\overline{\mu}$ has nonnegative values and satisfies

$\overline{\mu}(\varnothing) = 0$, and so we need only check its countable additivity. Let $\{A_n\}$ be a sequence of disjoint sets in $\mathscr{A}_\mu$, and for each n again choose sets E_n and F_n in $\mathscr{A}$ that satisfy $E_n \subseteq A_n \subseteq F_n$ and $\mu(F_n - E_n) = 0$. The disjointness of the sets A_n implies the disjointness of the sets E_n, and so we can conclude that

$$\overline{\mu}(\cup_n A_n) = \mu(\cup_n E_n) = \sum_n \mu(E_n) = \sum_n \overline{\mu}(A_n).$$

Thus $\overline{\mu}$ is a measure. It is easy to check that $\overline{\mu}$ is complete. □

We turn to an example.

Proposition 1.5.2. *Lebesgue measure on $(\mathbb{R}^d, \mathscr{M}_{\lambda^*})$ is the completion of Lebesgue measure on $(\mathbb{R}^d, \mathscr{B}(\mathbb{R}^d))$.*

We begin with the following lemma.

Lemma 1.5.3. *Let A be a Lebesgue measurable subset of $\mathbb{R}^d$. Then there exist Borel subsets E and F of $\mathbb{R}^d$ such that $E \subseteq A \subseteq F$ and $\lambda(F - E) = 0$.*

Proof. First suppose that A is a Lebesgue measurable subset of $\mathbb{R}^d$ such that $\lambda(A) < +\infty$. For each positive integer n, use Proposition 1.4.1 to choose a compact set K_n such that $K_n \subseteq A$ and $\lambda(A) - 1/n < \lambda(K_n)$ and an open set U_n such that $A \subseteq U_n$ and $\lambda(U_n) < \lambda(A) + 1/n$. Let $E = \cup_n K_n$ and $F = \cap_n U_n$. Then E and F belong to $\mathscr{B}(\mathbb{R}^d)$ and satisfy $E \subseteq A \subseteq F$. The relation

$$\lambda(F - E) \le \lambda(U_n - K_n) = \lambda(U_n - A) + \lambda(A - K_n) < 2/n$$

holds for each n, and so $\lambda(F - E) = 0$. Thus the lemma is proved in the case where $\lambda(A) < +\infty$.

If A is an arbitrary Lebesgue measurable subset of $\mathbb{R}^d$, then A is the union of a sequence $\{A_n\}$ of Lebesgue measurable sets of finite Lebesgue measure. For each n we can choose Borel sets E_n and F_n such that $E_n \subseteq A_n \subseteq F_n$ and $\lambda(F_n - E_n) = 0$. The sets E and F defined by $E = \cup_n E_n$ and $F = \cup_n F_n$ then satisfy $E \subseteq A \subseteq F$ and $\lambda(F - E) = 0$ (note that $F - E \subseteq \cup_n (F_n - E_n)$). □

Proof of Proposition 1.5.2. Let λ be Lebesgue measure on $(\mathbb{R}^d, \mathscr{B}(\mathbb{R}^d))$, let $\overline{\lambda}$ be the completion of λ, and let λ_m be Lebesgue measure on $(\mathbb{R}^d, \mathscr{M}_{\lambda^*})$. Lemma 1.5.3 implies that $\mathscr{M}_{\lambda^*}$ is included in the completion of $\mathscr{B}(\mathbb{R}^d)$ under λ and that λ_m is the restriction of $\overline{\lambda}$ to $\mathscr{M}_{\lambda^*}$. Thus we need only check that each set A that belongs to the completion of $\mathscr{B}(\mathbb{R}^d)$ under λ is Lebesgue measurable. For such a set A there exist Borel sets E and F such that $E \subseteq A \subseteq F$ and $\lambda(F - E) = 0$. Since $A - E \subseteq F - E$ and $\lambda_m(F - E) = \lambda(F - E) = 0$, the completeness of Lebesgue measure on $\mathscr{M}_{\lambda^*}$ implies that $A - E \in \mathscr{M}_{\lambda^*}$. Thus A, since it is the union of $A - E$ and E, must belong to $\mathscr{M}_{\lambda^*}$. □

We will see in Sect. 2.1 that

(a) there are Lebesgue measurable subsets of $\mathbb{R}$ that are not Borel sets, and
(b) the restriction of Lebesgue measure to $\mathscr{B}(\mathbb{R})$ is not complete.

It should be noted that although replacing a measure space $(X, \mathscr{A}, \mu)$ with its completion $(X, \mathscr{A}_\mu, \overline{\mu})$ enables one to avoid some difficulties, it introduces others. Some difficulties arise because the completed σ-algebra $\mathscr{A}_\mu$ is often more complicated than the original σ-algebra $\mathscr{A}$. Others are caused by the fact that for measures μ and ν defined on a common σ-algebra $\mathscr{A}$, the completions $\mathscr{A}_\mu$ and $\mathscr{A}_\nu$ of $\mathscr{A}$ under μ and ν may not be equal (see Exercise 3). Because of these complications it seems wise whenever possible to avoid arguments that depend on completeness; it turns out that in the basic parts of measure theory this can almost always be done.

Let $(X, \mathscr{A})$ be a measurable space, let μ be a measure on $\mathscr{A}$, and let A be an arbitrary subset of X. Then $\mu^*(A)$, the *outer measure* of A, is defined by

$$\mu^*(A) = \inf\{\mu(B) : A \subseteq B \text{ and } B \in \mathscr{A}\}, \tag{3}$$

and $\mu_*(A)$, the *inner measure* of A, is defined by

$$\mu_*(A) = \sup\{\mu(B) : B \subseteq A \text{ and } B \in \mathscr{A}\}.$$

It is easy to check that $\mu_*(A) \le \mu^*(A)$ holds for each subset A of X.

Proposition 1.5.4. *Let $(X, \mathscr{A})$ be a measurable space, and let μ be a measure on $(X, \mathscr{A})$. Then the function $\mu^* : \mathscr{P}(X) \to [0, +\infty]$ defined by Eq. (3) is an outer measure (as defined in Sect. 1.3) on X.*

Proof. Certainly μ^* satisfies $\mu^*(\varnothing) = 0$ and is monotone. We turn to its subadditivity. Let $\{A_n\}$ be a sequence of subsets of X. The inequality $\mu^*(\cup_n A_n) \le \sum_n \mu^*(A_n)$ is clear if $\sum_n \mu^*(A_n) = +\infty$. So suppose that $\sum_n \mu^*(A_n) < +\infty$. Let ε be an arbitrary positive number, and for each n choose a set B_n that belongs to $\mathscr{A}$, includes A_n, and satisfies $\mu(B_n) \le \mu^*(A_n) + \varepsilon/2^n$. Then the set B defined by $B = \cup_n B_n$ belongs to $\mathscr{A}$, includes $\cup_n A_n$, and satisfies $\mu(B) \le \sum_n \mu^*(A_n) + \varepsilon$ (see Proposition 1.2.4); thus $\mu^*(\cup_n A_n) \le \sum_n \mu^*(A_n) + \varepsilon$. Since ε is arbitrary, the proof is complete. □

Note that Proposition 1.4.11 can now be rephrased: there is a subset A of $\mathbb{R}$ such that $\lambda_*(A) = 0$ and $\lambda_*(A^c) = 0$.

Proposition 1.5.5. *Let $(X, \mathscr{A})$ be a measurable space, let μ be a measure on $\mathscr{A}$, and let A be a subset of X such that $\mu^*(A) < +\infty$. Then A belongs to $\mathscr{A}_\mu$ if and only if $\mu_*(A) = \mu^*(A)$.*

Proof. If A belongs to $\mathscr{A}_\mu$, then there are sets E and F that belong to $\mathscr{A}$ and satisfy $E \subseteq A \subseteq F$ and $\mu(F - E) = 0$. Then

$$\mu(E) \le \mu_*(A) \le \mu^*(A) \le \mu(F),$$

and since $\mu(E) = \mu(F)$, the relation $\mu_*(A) = \mu^*(A)$ follows.

One can obtain a proof that the relation $\mu_*(A) = \mu^*(A) < +\infty$ implies that A belongs to $\mathscr{A}_\mu$ by modifying the first paragraph of the proof of Lemma 1.5.3; the details are left to the reader (replace appeals to Proposition 1.4.1 with appeals to the definitions of μ_* and μ^*). □

In this section we have been dealing with one way of approximating sets from above and from below by measurable sets. We turn to another such approximation.

Let $\mathscr{A}$ be a σ-algebra on $\mathbb{R}^d$ that includes the σ-algebra $\mathscr{B}(\mathbb{R}^d)$ of Borel sets. A measure μ on $(\mathbb{R}^d, \mathscr{A})$ is *regular* if

(a) each compact subset K of $\mathbb{R}^d$ satisfies $\mu(K) < +\infty$,
(b) each set A in $\mathscr{A}$ satisfies

$$\mu(A) = \inf\{\mu(U) : U \text{ is open and } A \subseteq U\}, \text{ and}$$

(c) each open subset U of $\mathbb{R}^d$ satisfies

$$\mu(U) = \sup\{\mu(K) : K \text{ is compact and } K \subseteq U\}.$$

Proposition 1.4.1 implies that Lebesgue measure, whether on $(\mathbb{R}^d, \mathscr{M}_{\lambda^*})$ or on $(\mathbb{R}^d, \mathscr{B}(\mathbb{R}^d))$, is regular. Part (b) of that proposition appears to be stronger than condition (c) in the definition of regularity; however, we will see in Chap. 7 that every regular measure on $(\mathbb{R}^d, \mathscr{B}(\mathbb{R}^d))$ satisfies the analogue of part (b) of Proposition 1.4.1. In Chap. 7 we will also see that on more general spaces, the analogue of condition (c) above, rather than of part (b) of Proposition 1.4.1, is the condition that should be used in the definition of regularity.

Proposition 1.5.6. *Let μ be a finite measure on $(\mathbb{R}^d, \mathscr{B}(\mathbb{R}^d))$. Then μ is regular. Moreover, each Borel subset A of $\mathbb{R}^d$ satisfies*

$$\mu(A) = \sup\{\mu(K) : K \subseteq A \text{ and } K \text{ is compact}\}. \tag{4}$$

Let us first prove the following weakened form of Proposition 1.5.6.

Lemma 1.5.7. *Let μ be a finite measure on $(\mathbb{R}^d, \mathscr{B}(\mathbb{R}^d))$. Then each Borel subset A of $\mathbb{R}^d$ satisfies*

$$\mu(A) = \inf\{\mu(U) : A \subseteq U \text{ and } U \text{ is open}\} \text{ and} \tag{5}$$

$$\mu(A) = \sup\{\mu(C) : C \subseteq A \text{ and } C \text{ is closed}\}. \tag{6}$$

Proof. Let $\mathscr{R}$ be the collection of those Borel subsets A of $\mathbb{R}^d$ that satisfy (5) and (6).

We begin by showing that $\mathscr{R}$ contains the open subsets of $\mathbb{R}^d$. Let V be an open subset of $\mathbb{R}^d$. Of course V satisfies

$$\mu(V) = \inf\{\mu(U) : V \subseteq U \text{ and } U \text{ is open}\}.$$

According to Proposition 1.1.6, there is a sequence $\{C_n\}$ of closed subsets of $\mathbb{R}^d$ such that $V = \cup_n C_n$. We can assume that the sequence $\{C_n\}$ is increasing (replace C_n with $\cup_{i=1}^n C_i$ if necessary). Proposition 1.2.5 implies that $\mu(V) = \lim_n \mu(C_n)$, and so V satisfies

$$\mu(V) = \sup\{\mu(C) : C \subseteq V \text{ and } C \text{ is closed}\}.$$

With this we have proved that $\mathscr{R}$ contains all the open subsets of $\mathbb{R}^d$.

It is easy to check (do so) that $\mathscr{R}$ consists of the Borel sets A that satisfy

$$\begin{aligned}&\text{for each positive } \varepsilon \text{ there exist an open set } U \text{ and a closed set } C\\&\text{such that } C \subseteq A \subseteq U \text{ and } \mu(U - C) < \varepsilon.\end{aligned} \tag{7}$$

We now show that $\mathscr{R}$ is a σ-algebra. If contains $\mathbb{R}^d$, since $\mathbb{R}^d$ is open. If $A \in \mathscr{R}$, if ε is a positive number, and if C and U are, respectively, closed and open and satisfy $C \subseteq A \subseteq U$ and $\mu(U - C) < \varepsilon$, then U^c and C^c are respectively closed and open and satisfy $U^c \subseteq A^c \subseteq C^c$ and $\mu(C^c - U^c) < \varepsilon$; thus it follows (from (7)) that $\mathscr{R}$ is closed under complementation. Now let $\{A_k\}$ be a sequence of sets in $\mathscr{R}$ and let ε be a positive number. For each k choose a closed set C_k and an open set U_k such that $C_k \subseteq A_k \subseteq U_k$ and $\mu(U_k - C_k) < \varepsilon/2^k$. Let $U = \cup_k U_k$ and $C = \cup_k C_k$. Then U and C satisfy the relations $C \subseteq \cup_k A_k \subseteq U$ and

$$\mu(U - C) \le \mu(\cup_k (U_k - C_k)) \le \sum_k (U_k - C_k) < \varepsilon. \tag{8}$$

The set U is open, but the set C can fail to be closed. However, for each n the set $\cup_{k=1}^n C_k$ is closed, and it follows from (8), together with the fact that $\mu(U - C) = \lim_n \mu(U - \cup_{k=1}^n C_k)$ that there is a positive integer n such that $\mu(U - \cup_{k=1}^n C_k) < \varepsilon$. Then U and $\cup_{k=1}^n C_k$ are the sets required in (7), and $\mathscr{R}$ is closed under the formation of countable unions.

We have now shown that $\mathscr{R}$ is a σ-algebra on $\mathbb{R}^d$ that contains the open sets. Since $\mathscr{B}(\mathbb{R}^d)$ is the smallest σ-algebra on $\mathbb{R}^d$ that contains the open sets, it follows that $\mathscr{B}(\mathbb{R}^d) \subseteq \mathscr{R}$. With this Lemma 1.5.7 is proved. □

Proof of Proposition 1.5.6. Condition (a) in the definition of regularity follows from the finiteness of μ, while condition (b) follows from Lemma 1.5.7. We turn to condition (c) and Eq. (4). Let A be a Borel subset of $\mathbb{R}^d$ and let ε be a positive number. Then according to Lemma 1.5.7 there is a closed subset C of A such that $\mu(C) > \mu(A) - \varepsilon$. Choose an increasing sequence $\{C_n\}$ of closed and bounded (hence compact) sets whose union is C (these sets can, for example, be constructed by letting $C_n = C \cap \{x \in \mathbb{R}^d : \|x\| \le n\}$). Proposition 1.2.5 implies that $\mu(C) = \lim_n \mu(C_n)$, and so if n is large enough, then C_n is a compact subset of A such that $\mu(C_n) > \mu(A) - \varepsilon$. Equation (4) and condition (c) follow. □

Exercises

1. Let $(X, \mathscr{A}, \mu)$ be a measure space. Show that $(\mathscr{A}_\mu)_{\overline{\mu}} = \mathscr{A}_\mu$ and $\overline{\overline{\mu}} = \overline{\mu}$.
2. (a) Find the completion of $\mathscr{B}(\mathbb{R})$ under the point mass concentrated at 0.

(b) Let $\mathscr{A}$ be the σ-algebra on $\mathbb{R}^2$ that consists of all unions of (possibly empty) collections of vertical lines. Find the completion of $\mathscr{A}$ under the point mass concentrated at $(0,0)$.

3. Let μ and ν be finite measures on a measurable space $(X,\mathscr{A})$.
   (a) Show by example that $\mathscr{A}_\mu$ and $\mathscr{A}_\nu$ need not be equal.
   (b) Prove or disprove: $\mathscr{A}_\mu = \mathscr{A}_\nu$ if and only if μ and ν have exactly the same sets of measure zero.
4. Show that there is a Lebesgue measurable subset of $\mathbb{R}^2$ whose projection on $\mathbb{R}$ under the map $(x,y) \mapsto x$ is not Lebesgue measurable.
5. Let μ be a measure on $(X,\mathscr{A})$. Show that for each subset A of X there are sets A_0 and A_1 that belong to $\mathscr{A}$ and satisfy the conditions $A_0 \subseteq A \subseteq A_1$, $\mu(A_0) = \mu_*(A)$, and $\mu(A_1) = \mu^*(A)$.
6. Show by example that half of Proposition 1.5.5 can fail if the assumption that $\mu^*(A) < +\infty$ is omitted.
7. Suppose that μ is a measure on $(X,\mathscr{A})$. Show that each subset A of X satisfies $\mu^*(A) + \mu_*(A^c) = \mu(X)$.
8. Show that there is a subset A of the interval $[0,1]$ that satisfies $\lambda^*(A) = 1$ and $\lambda_*(A) = 0$. (Hint: Use Proposition 1.4.11.)
9. Let μ be a σ-finite measure on $(X,\mathscr{A})$, and let μ^* be the outer measure defined in formula (3). Show that $\mathscr{A}_\mu$ is equal to the σ-algebra of μ^*-measurable sets and that $\overline{\mu}$ is the restriction of μ^* to $\mathscr{A}_\mu$.
10. Show that if A is a Lebesgue measurable subset of $\mathbb{R}$, then $\{(x,y) \in \mathbb{R}^2 : x \in A\}$ is a Lebesgue measurable subset of $\mathbb{R}^2$.
11. Let $(X,\mathscr{A})$ be a measurable space, and let C be a subset of X (it is not assumed that C belongs to $\mathscr{A}$).
   (a) Show that the collection of subsets of C that have the form $A \cap C$ for some A in $\mathscr{A}$ is a σ-algebra on C. This σ-algebra is sometimes called the *trace* of $\mathscr{A}$ on C and is denoted by $\mathscr{A}_C$.
   (b) Now suppose that μ is a finite measure on $(X,\mathscr{A})$. Let C_1 be a set that belongs to $\mathscr{A}$, includes C, and satisfies $\mu(C_1) = \mu^*(C)$ (see Exercise 5). Show that if A_1 and A_2 belong to $\mathscr{A}$ and satisfy $A_1 \cap C = A_2 \cap C$, then $\mu(A_1 \cap C_1) = \mu(A_2 \cap C_1)$. Thus we can use the formula $\mu_C(A \cap C) = \mu(A \cap C_1)$ to define a function $\mu_C \colon \mathscr{A}_C \to [0,+\infty)$.
   (c) Show that $\mu_C(B) = \mu^*(B)$ holds for each B in $\mathscr{A}_C$. Thus μ_C does not depend on the choice of the set C_1.
   (d) Show that μ_C is a measure on $(C,\mathscr{A}_C)$. The measure μ_C is sometimes called the *trace* of μ on C.
12. Let $(X,\mathscr{A})$ be a measurable space, and let C be a subset of X.
   (a) Show that the sets that belong to $\sigma(\mathscr{A} \cup \{C\})$ are exactly those that have the form $(A_1 \cap C) \cup (A_2 \cap C^c)$ for some A_1 and A_2 in $\mathscr{A}$.
   (b) Now suppose that μ is a finite measure on $(X,\mathscr{A})$. Let C_0 and C_1 be $\mathscr{A}$-measurable subsets of C and C^c that satisfy $\mu(C_0) = \mu_*(C)$ and $\mu(C_1) = \mu_*(C^c)$, and let μ_C and μ_{C^c} be the traces of μ on C and C^c (see Exercises 5 and 11). Show that the formulas

$$\mu_0(A) = \mu(A \cap C_0) + \mu_{C^c}(A \cap C^c)$$

and

$$\mu_1(A) = \mu_C(A \cap C) + \mu(A \cap C_1)$$

define measures μ_0 and μ_1 on $\sigma(\mathscr{A} \cup \{C\})$, that these measures agree with μ on $\mathscr{A}$, and that they satisfy $\mu_0(C) = \mu_*(C)$ and $\mu_1(C) = \mu^*(C)$.

(c) Show that for each α between $\mu_*(C)$ and $\mu^*(C)$ there is a measure ν on $\sigma(\mathscr{A} \cup \{C\})$ that agrees with μ on $\mathscr{A}$ and satisfies $\nu(C) = \alpha$. (Hint: Let $\nu = t\mu_0 + (1-t)\mu_1$ for a suitable t.)

1.6 Dynkin Classes

This section is devoted to a technique that is often useful for verifying the equality of measures and the measurability of functions (measurable functions will be defined in Sect. 2.1). We begin with a basic definition.

Let X be a set. A collection $\mathscr{D}$ of subsets of X is a *d-system* (or a *Dynkin class*) on X if

(a) $X \in \mathscr{D}$,
(b) $A - B \in \mathscr{D}$ whenever $A, B \in \mathscr{D}$ and $A \supseteq B$, and
(c) $\cup_n A_n \in \mathscr{D}$ whenever $\{A_n\}$ is an increasing sequence of sets in $\mathscr{D}$.

A collection of subsets of X is a *π-system* on X if it is closed under the formation of finite intersections.

Example 1.6.1. Suppose that X is a set and that $\mathscr{A}$ is a σ-algebra on X. Then $\mathscr{A}$ is certainly a d-system. Furthermore, if μ and ν are finite measures on $\mathscr{A}$ such that $\mu(X) = \nu(X)$, then the collection $\mathscr{S}$ of all sets A that belong to $\mathscr{A}$ and satisfy $\mu(A) = \nu(A)$ is a d-system; it is easy to show by example that $\mathscr{S}$ is not necessarily a σ-algebra (see Exercise 3). The fact that such families $\mathscr{S}$ are d-systems forms the basis for many of the applications of d-systems. □

Note that the intersection of a nonempty family of d-systems on a set X is a d-system on X and that an arbitrary collection of subsets of X is included in some d-system on X, namely the collection of all subsets of X. Hence if $\mathscr{C}$ is an arbitrary collection of subsets of X, then the intersection of all the d-systems on X that include $\mathscr{C}$ is a d-system on X that includes $\mathscr{C}$; this intersection is the smallest such d-system and is called the d-system *generated* by $\mathscr{C}$. We will sometimes denote this d-system by $d(\mathscr{C})$.

Theorem 1.6.2. *Let X be a set, and let $\mathscr{C}$ be a π-system on X. Then the σ-algebra generated by $\mathscr{C}$ coincides with the d-system generated by $\mathscr{C}$.*

Proof. Let $\mathscr{D}$ be the d-system generated by $\mathscr{C}$, and, as usual, let $\sigma(\mathscr{C})$ be the σ-algebra generated by $\mathscr{C}$. Since every σ-algebra is a d-system, the σ-algebra $\sigma(\mathscr{C})$

is a d-system that includes $\mathscr{C}$; hence $\mathscr{D} \subseteq \sigma(\mathscr{C})$. We can prove the reverse inclusion by showing that $\mathscr{D}$ is a σ-algebra, for then $\mathscr{D}$, as a σ-algebra that includes $\mathscr{C}$, must include the σ-algebra generated by $\mathscr{C}$, namely $\sigma(\mathscr{C})$.

We begin the proof that $\mathscr{D}$ is a σ-algebra by showing that $\mathscr{D}$ is closed under the formation of finite intersections. Define a family $\mathscr{D}_1$ of subsets of X by letting

$$\mathscr{D}_1 = \{A \in \mathscr{D} : A \cap C \in \mathscr{D} \text{ for each } C \text{ in } \mathscr{C}\}.$$

The fact that $\mathscr{C} \subseteq \mathscr{D}$ implies that $X \in \mathscr{D}_1$; furthermore, the identities

$$(A - B) \cap C = (A \cap C) - (B \cap C)$$

and

$$(\cup_n A_n) \cap C = \cup_n (A_n \cap C),$$

together with the fact that $\mathscr{D}$ is a d-system, imply that $\mathscr{D}_1$ is closed under the formation of proper differences and under the formation of unions of increasing sequences of sets. Thus $\mathscr{D}_1$ is a d-system. Since $\mathscr{C}$ is closed under the formation of finite intersections and is included in $\mathscr{D}$, it is included in $\mathscr{D}_1$. Thus $\mathscr{D}_1$ is a d-system that includes $\mathscr{C}$; hence it must include $\mathscr{D}$. With this we have proved that we get a set in $\mathscr{D}$ whenever we take the intersection of a set in $\mathscr{D}$ and a set in $\mathscr{C}$.

Next define $\mathscr{D}_2$ by letting

$$\mathscr{D}_2 = \{B \in \mathscr{D} : A \cap B \in \mathscr{D} \text{ for each } A \text{ in } \mathscr{D}\}.$$

The previous step of this proof shows that $\mathscr{C} \subseteq \mathscr{D}_2$, and a straightforward modification of the argument in the previous step shows that $\mathscr{D}_2$ is a d-system. It follows that $\mathscr{D} \subseteq \mathscr{D}_2$—in other words, that $\mathscr{D}$ is closed under the formation of finite intersections.

It is now easy to complete the proof. Parts (a) and (b) of the definition of a d-system imply that $X \in \mathscr{D}$ and that $\mathscr{D}$ is closed under complementation. As we have just seen, $\mathscr{D}$ is also closed under the formation of finite intersections, and so it is an algebra. Finally $\mathscr{D}$, as a d-system, is closed under the formation of unions of increasing sequences of sets, and so by Proposition 1.1.7 it must be a σ-algebra; with that the proof is complete. □

We turn to some applications of Theorem 1.6.2.

Corollary 1.6.3. *Let $(X, \mathscr{A})$ be a measurable space, and let $\mathscr{C}$ be a π-system on X such that $\mathscr{A} = \sigma(\mathscr{C})$. If μ and ν are finite measures on $\mathscr{A}$ that satisfy $\mu(X) = \nu(X)$ and that satisfy $\mu(C) = \nu(C)$ for each C in $\mathscr{C}$, then $\mu = \nu$.*

Proof. Let $\mathscr{D} = \{A \in \mathscr{A} : \mu(A) = \nu(A)\}$. As we noted above, $\mathscr{D}$ is a d-system. Since $\mathscr{C}$ is a π-system and is included in $\mathscr{D}$, it follows from Theorem 1.6.2 that $\mathscr{D} \supseteq \sigma(\mathscr{C}) = \mathscr{A}$. Thus $\mu(A) = \nu(A)$ holds for each A in $\mathscr{A}$, and the proof is complete. □

Now suppose that μ and ν are finite Borel measures on $\mathbb{R}$ such that $\mu(I) = \nu(I)$ holds for each interval I of the form $(-\infty, b]$. Note that $\mathbb{R}$ is the union of an increasing sequence of intervals of the form $(-\infty, b]$ and hence that $\mu(\mathbb{R}) = \nu(\mathbb{R})$. Since the collection of all intervals of the form $(-\infty, b]$ is a π-system that generates $\mathscr{B}(\mathbb{R})$ (see Proposition 1.1.4), it follows from Corollary 1.6.3 that $\mu = \nu$. With this we have another proof of the uniqueness assertion in Proposition 1.3.10.

The following result is essentially an extension of Corollary 1.6.3 to the case of σ-finite measures. Note that it implies that Lebesgue measure is the only measure on $\mathscr{B}(\mathbb{R}^d)$ that assigns to each d-dimensional interval its volume, and so it provides a second proof of part of Proposition 1.4.3.

Corollary 1.6.4. *Let $(X, \mathscr{A})$ be a measurable space, and let $\mathscr{C}$ be a π-system on X such that $\mathscr{A} = \sigma(\mathscr{C})$. If μ and ν are measures on $(X, \mathscr{A})$ that agree on $\mathscr{C}$, and if there is an increasing sequence $\{C_n\}$ of sets that belong to $\mathscr{C}$, have finite measure under μ and ν, and satisfy $\cup_n C_n = X$, then $\mu = \nu$.*

Proof. Choose an increasing sequence $\{C_n\}$ of sets that belong to $\mathscr{C}$, have finite measure under μ and ν, and satisfy $\cup_n C_n = X$. For each positive integer n define measures μ_n and ν_n on $\mathscr{A}$ by $\mu_n(A) = \mu(A \cap C_n)$ and $\nu_n(A) = \nu(A \cap C_n)$. Corollary 1.6.3 implies that for each n we have $\mu_n = \nu_n$. Since

$$\mu(A) = \lim_n \mu_n(A) = \lim_n \nu_n(A) = \nu(A)$$

holds for each A in $\mathscr{A}$, the measures μ and ν must be equal. $\square$

Exercises

1. Give at least six π-systems on $\mathbb{R}$, each of which generates $\mathscr{B}(\mathbb{R})$.
2. (b) Check that the rectangles of the form considered in part (c) of Proposition 1.1.5, together with the empty set, form a π-system on $\mathbb{R}^d$.
   (b) What is the smallest π-system on $\mathbb{R}^d$ that contains all the half-spaces of the form considered in part (b) of Proposition 1.1.5?
3. Give a measurable space $(X, \mathscr{A})$ and finite measures μ and ν on it that satisfy $\mu(X) = \nu(X)$ but are such that

$$\{A \in \mathscr{A} : \mu(A) = \nu(A)\}$$

   is not a σ-algebra. (Hint: Don't work too hard; X can be a fairly small finite set.)
4. Show by example that Corollary 1.6.3 would be false if the hypothesis that μ and ν are finite were replaced with the hypothesis that they are σ-finite. (See, however, Corollary 1.6.4.)
5. Use Theorem 1.6.2 to give another proof of Proposition 1.5.6. (Hint: Show that the collection consisting of those Borel subsets of $\mathbb{R}^d$ that can be approximated

from below with compact sets and from above with open sets is a d-system, and that this d-system contains each rectangle of the form considered in part (c) of Proposition 1.1.5.)

6. Let X be a set. A collection $\mathscr{C}$ of subsets of X is a *monotone class* on X if it is closed under monotone limits, in the sense that

   (i) if $\{A_n\}$ is an increasing sequence of sets that belong to $\mathscr{C}$, then $\cup_n A_n$ belongs to $\mathscr{C}$, and
   (ii) if $\{A_n\}$ is a decreasing sequence of sets that belong to $\mathscr{C}$, then $\cap_n A_n$ belongs to $\mathscr{C}$.

   (a) Show that if $\mathscr{A}$ is a collection of subsets of X, then there is a smallest monotone class on X that includes $\mathscr{A}$. This smallest monotone class is called the monotone class *generated* by $\mathscr{A}$; let us denote it by $m(\mathscr{A})$.
   (b) Prove the *monotone class theorem*: if $\mathscr{A}$ is an algebra of subsets of X, then $m(\mathscr{A}) = \sigma(\mathscr{A})$. (Hint: Modify the proof of Theorem 1.6.2.)

Notes

Halmos [54] is a standard reference for the theory of measure and integration. The books by Bartle [3], Berberian [7], Billingsley [8], Bruckner, Bruckner, and Thomson [23], Dudley [40], Folland [45], Hewitt and Stromberg [59], Munroe [92], Royden [102], Rudin [105], and Wheeden and Zygmund [127] are also well known and useful. The reader should see Billingsley [8] and Dudley [40] for applications to probability theory, Rudin [105] and Benedetto and Czaja [6] for a great variety of applications to analysis, and Wheeden and Zygmund [127] for applications to harmonic analysis. Gelbaum and Olmsted [48] contains an interesting collection of counterexamples. Bogachev's recent two-volume work [15] and Fremlin's five-volume work [46] are good references. Pap [95] is a collection of survey papers on measure theory. Federer [44], Krantz and Parks [75], Morgan [89], and Rogers [100] treat topics in measure theory that are not touched upon here.

Theorem 1.6.2 is due to Dynkin [43] (see also Blumenthal and Getoor [14]).

See Dudley [40] and Bogachev [15] for very thorough historical notes and bibliographic citations.

Chapter 2
Functions and Integrals

This chapter is devoted to the definition and basic properties of the Lebesgue integral. We first introduce measurable functions—the functions that are simple enough that the integral can be defined for them if their values are not too large (Sect. 2.1). After a brief look in Sect. 2.2 at properties that hold almost everywhere (that is, that may fail on some set of measure zero, as long as they hold everywhere else), we turn to the definition of the Lebesgue integral and to its basic properties (Sects. 2.3 and 2.4). The chapter ends with a sketch of how the Lebesgue integral relates to the Riemann integral (Sect. 2.5) and then with a few more details about measurable functions (Sect. 2.6).

2.1 Measurable Functions

In this section we introduce measurable functions and study some of their basic properties. We begin with the following elementary result.

Proposition 2.1.1. *Let $(X, \mathscr{A})$ be a measurable space, and let A be a subset of X that belongs to $\mathscr{A}$. For a function $f\colon A \to [-\infty, +\infty]$ the conditions*

(a) *for each real number t the set $\{x \in A : f(x) \leq t\}$ belongs to $\mathscr{A}$,*
(b) *for each real number t the set $\{x \in A : f(x) < t\}$ belongs to $\mathscr{A}$,*
(c) *for each real number t the set $\{x \in A : f(x) \geq t\}$ belongs to $\mathscr{A}$, and*
(d) *for each real number t the set $\{x \in A : f(x) > t\}$ belongs to $\mathscr{A}$*

are equivalent.

Proof. The identity

$$\{x \in A : f(x) < t\} = \bigcup_n \{x \in A : f(x) \leq t - 1/n\}$$

D.L. Cohn, *Measure Theory: Second Edition*, Birkhäuser Advanced Texts Basler Lehrbücher, DOI 10.1007/978-1-4614-6956-8_2,