2 Lebesgue Measure and General Measure Theory

2.1 The theory of measure prior to Lebesgue, and preliminaries

For the historical development of measure theory prior to LEBESGUE we refer to ARTHUR SCHOENFLIES' work on set theory (1900) and [226].

One of the first people to recognize the connection between an integration theory and a precise theory for measuring the "length" of sets was GIUSEPPE PEANO, and his work on these matters is contained in his book *Applicazione geometriche del calcolo infinitesimale*, published in 1887 in Torino. Taking a nonnegative, bounded function $f : [a, b] \to \mathbb{R}$, he let E be the point set bounded by the graph of f and the lines $x = a$, $x = b$, $y = 0$, and he then defined

$$c_i(E) = \sup \{a(A) : A \subseteq E, A \text{ a polygonal region}\},$$
$$c_e(E) = \inf \{a(A) : A \supseteq E, A \text{ a polygonal region}\},$$

where $a(A)$ is the area of A. In these terms he observed that *f is Riemann integrable if and only if E is "measurable", i.e., $c_e(E) = c_i(E)$.*

When c_e and c_i are equal we write c for their common value, and PEANO actually developed the theory of finitely additive set functions for such c.

Mind you, there was a great deal of activity in the general problem of measuring sets prior to PEANO. For example, using the fact that a closed set $E \subseteq \mathbb{R}$ can be written as

$$E = \left(\bigcup_{j=1}^{\infty} I_j \right)^{\sim},$$

where $\{I_j : j = 1, \ldots\}$ is a disjoint family of open intervals and $m(I_j)$ denotes the length of the interval I_j, CANTOR and IVAR BENDIXSON were led to define

$$\text{“}m(C)'' = 1 - \sum_{j=1}^{\infty} m(I_j),$$

whose value, as we showed in Example 1.2.7, is 0.

Five years after PEANO's work, CAMILLE JORDAN also developed the theory of finitely additive measures, and, although there are no references to

J.J. Benedetto and W. Czaja, *Integration and Modern Analysis*, Birkhäuser Advanced Texts,
DOI 10.1007/978-0-8176-4656-1_2,

PEANO, it seems likely that JORDAN knew of his work. In any case, mathematically, JORDAN carried finitely additive measure theory very far, and the relation between measure and integrability was quite clearly explicated in JORDAN's theory.

The next major step in the evolution of the fundamental ideas leading to the present notions of integral and measure was taken by ÉMILE BOREL from the time of his doctorate in 1894. To discuss his work we need to define the notion of a *σ-algebra*.

Definition 2.1.1. Rings and algebras
Let X be any set. The power set $\mathcal{P}(X)$ is the set of all subsets of X and it was initially defined in Problem 1.3.

a. We say that a collection of sets $\mathcal{R} \subseteq \mathcal{P}(X)$ is a *ring* if, for all $A, B \in \mathcal{R}$, the following properties are satisfied:

$$A \cap B \in \mathcal{R},$$
$$A \cup B \in \mathcal{R},$$
$$A \setminus B \in \mathcal{R}.$$

A ring $\mathcal{R}$ is a *σ-ring* if, for any sequence $\{A_n : n = 1, \ldots\} \subseteq \mathcal{R}$,

$$\bigcap_{n=1}^{\infty} A_n \in \mathcal{R}$$

and

$$\bigcup_{n=1}^{\infty} A_n \in \mathcal{R}.$$

b. We say that a ring (respectively, σ-ring) $\mathcal{A}$ is an *algebra* (respectively, *σ-algebra*) if

$$X \in \mathcal{A}.$$

In this case, $\emptyset \in \mathcal{A}$, and $A \in \mathcal{A}$ implies $A^{\sim} \in \mathcal{A}$.

Proposition 2.1.2. *For each collection $\mathcal{C} \subseteq \mathcal{P}(X)$ there exists the smallest ring (respectively, σ-ring, algebra, and σ-algebra) that contains $\mathcal{C}$.*

Proof. We shall only show the existence of the smallest ring containing $\mathcal{C}$. Let $\mathcal{F}$ be the family of all rings in $\mathcal{P}(X)$ that contain $\mathcal{C}$, and let $\mathcal{R} = \bigcap\{\mathcal{D} : \mathcal{D} \in \mathcal{F}\}$. Clearly $\mathcal{C} \subseteq \mathcal{R}$. Moreover, $\mathcal{R}$ is a ring. In fact, let $A, B \in \mathcal{R}$. Then, for each $\mathcal{D} \in \mathcal{F}$, $A, B \in \mathcal{D}$. Since $\mathcal{D}$ is a ring, $A \cup B, A \cap B, A \setminus B \in \mathcal{D}$. This is true for each $\mathcal{D} \in \mathcal{F}$, and so $A \cup B, A \cap B, A \setminus B \in \mathcal{R}$. By definition, $\mathcal{R}$ is the smallest ring containing $\mathcal{C}$. $\square$

We say that the smallest ring (σ-ring, algebra, σ-algebra) that contains a given family $\mathcal{C}$ is *generated* by $\mathcal{C}$.

In light of the importance of open and closed sets we make the following definition.

Definition 2.1.3. Borel sets
The collection $\mathcal{B} = \mathcal{B}(X)$ of *Borel sets* in $X \subseteq \mathbb{R}$ is the smallest σ-algebra in $\mathcal{P}(X)$ containing the open sets in X.

If X is any topological space, defined in Appendix A.1, we define the collection $\mathcal{B} = \mathcal{B}(X)$ of *Borel sets* in X in the same way; $\mathcal{B}$ is the *Borel algebra.*

Example 2.1.4. Borel sets
a. We have defined $\mathcal{F}_\sigma$, G_δ, and $F_{\sigma\delta}$ sets in Section 1.3.2. These sets are all Borel sets, as are

$$\mathcal{G}_{\delta\sigma},\ \mathcal{G}_{\delta\sigma\delta},\ \ldots, \mathcal{F}_{\sigma\delta\sigma},\ \mathcal{F}_{\sigma\delta\sigma\delta},\ \ldots.$$

b. There are Borel sets $B \subseteq \mathbb{R}$ that are not of the form $\mathcal{F}_n$ or $\mathcal{G}_m$, where, for example,

$$\mathcal{F}_n = \mathcal{F}_{\sigma\delta\sigma\ldots}.$$

In fact, set

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

where A_n is an $\mathcal{F}_n$ but not an $\mathcal{F}_{n-1}$, e.g., [224], page 182.

c. There are Borel sets $B_1, B_2 \in \mathcal{B}(\mathbb{R})$ such that $B_1 - B_2 = \{x - y : x \in B_1, y \in B_2\} \notin \mathcal{B}(\mathbb{R})$ [401].

d. There is precisely a continuum of Borel sets in $\mathbb{R}$; see, e.g., [307].

Besides first proving the Heine–Borel theorem in the form of Theorem 1.2.10, Borel gave a reasonable definition of measure 0 and stressed countable additivity for his measures. In fact, by dividing $[0,1]$ into equal parts Jordan had reached the conclusion that "$m(\mathbb{Q} \cap [0,1])$"$= 1$. Borel, on the other hand, attached to each $r_n \in \mathbb{Q} \cap [0,1]$ the segment of length ε/n^2. Consequently, he concluded that "$m(\mathbb{Q} \cap [0,1])$"$< \varepsilon \sum 1/n^2$ for each ε; and so "$m(\mathbb{Q} \cap [0,1])$"$= 0$. This example led Borel into his study of measure.

We shall see how countable additivity and integration are related very soon. The countable additivity of Borel, as opposed to the finite additivity of Peano–Jordan, was crucial in Lebesgue's theory for attaining many fundamental results.

Definition 2.1.5. Finitely additive and σ-additive set functions and measures
a. Let $\mathcal{A} \subseteq \mathcal{P}(X)$ and let μ be a set function

$$\mu : \mathcal{A} \to \mathbb{R}^+ \cup \{\infty\}.$$

Then μ is *finitely additive* on $\mathcal{A}$ if for each finite sequence $\{A_n : n = 1, \ldots, N\} \subseteq \mathcal{A}$ of mutually disjoint sets, we have

$$\mu\left(\bigcup_{n=1}^{N} A_n\right) = \sum_{n=1}^{N} \mu(A_n).$$

b. Let $\mathcal{A} \subseteq \mathcal{P}(X)$ and let μ be a set function

$$\mu : \mathcal{A} \to \mathbb{R}^+ \cup \{\infty\}.$$

Then μ is *σ-additive* or *countably additive* on $\mathcal{A}$ if, for each sequence $\{A_n : n = 1, \ldots\} \subseteq \mathcal{A}$ of mutually disjoint sets, we have

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

c. Let $\mathcal{A}$ be a σ-algebra on X. A nonnegative, σ-additive set function $\mu : \mathcal{A} \to \mathbb{R}^+ \cup \{\infty\}$ is called a *countably additive measure* or *σ-additive measure* or, just simply, a *measure*.

It is not a priori clear that a nontrivial measure exists. LEBESGUE solved this existence problem in a very strong way by constructing a σ-algebra $\mathcal{M}(\mathbb{R}) \supseteq \mathcal{B}(\mathbb{R})$ and a measure m on $\mathcal{M}(\mathbb{R})$ with the reasonable properties that

$$\forall\, x \in \mathbb{R} \text{ and } \forall\, A \in \mathcal{M}(\mathbb{R}), \quad m(A + x) = m(A), \tag{2.1}$$

and

$$\forall\, I \subseteq \mathbb{R},\ I \text{ an interval, } m(I) \text{ is the length of } I; \tag{2.2}$$

see Section 2.2 for a proof. We call m *Lebesgue measure*, and conditions (2.1) and (2.2) are specific to m.

2.2 The construction of Lebesgue measure on $\mathbb{R}$

We start our venture into the business of constructing Lebesgue measure by first considering the special case of the real line $\mathbb{R}$. The difference between this case and the general case of $\mathbb{R}^d$ is that the ordering of $\mathbb{R}$ enables us to make simple arguments for comparison between Lebesgue measure on $\mathbb{R}$, which we shall construct, and the notion of length of intervals. This, in turn, allows us to use the family of open sets of $\mathbb{R}$ as our starting point.

As we have indicated, if I is an interval then $m(I)$ will designate its length. For any $A \subseteq \mathbb{R}$ the *Lebesgue outer measure* of A is

$$\begin{aligned} m^*(A) = \inf \Big\{ \sum_{n=1}^{\infty} m(I_n) : A \subseteq \bigcup_{n=1}^{\infty} I_n \text{ and } \{I_n : n = 1, \ldots\} \\ \text{is a countable family of open intervals } \Big\}. \end{aligned} \tag{2.3}$$

Clearly,

i. if $A \subseteq B$ *then* $m^*(A) \leq m^*(B)$,

ii. $m^*(\emptyset) = 0$.

Proposition 2.2.1. *Let $I \subseteq \mathbb{R}$ be an interval. Then $m^*(I) = m(I)$.*

Proof. Let $I = [a, b]$, $-\infty < a \leq b < \infty$. For each $\varepsilon > 0$, $I \subseteq (a - \varepsilon, b + \varepsilon)$, and so

$$m^*(I) \leq m((a - \varepsilon, b + \varepsilon)) = b - a + 2\varepsilon. \tag{2.4}$$

Since (2.4) is true for each $\varepsilon > 0$,

$$m^*(I) \leq b - a.$$

To prove $m^*(I) \geq b - a$, we shall show that, for any sequence $\{I_n : n = 1, \ldots\}$ of open intervals,

$$I \subseteq \bigcup_{n=1}^{\infty} I_n \Longrightarrow \sum_{n=1}^{\infty} m(I_n) \geq b - a. \tag{2.5}$$

By the Heine–Borel theorem each such sequence $\{I_n : n = 1, \ldots\}$, for which $I \subseteq \bigcup I_n$, has a finite subcollection $I_{n_1}, \ldots, I_{n_k}$ covering I, and, in particular,

$$\sum_{n=1}^{\infty} m(I_n) \geq \sum_{j=1}^{k} m(I_{n_j}).$$

Consequently, we need only prove (2.5) for finite covers $\{J_1, \ldots, J_k\}$. Let $a \in \bigcup_{j=1}^{k} J_j$. Without loss of generality, we suppose $a \in J_1 = (a_1, b_1)$. Thus, if $b_1 \geq b$,

$$\sum_{j=1}^{k} m(J_j) \geq b_1 - a_1 > b - a,$$

in which case we are finished. If $a < b_1 < b$, then since $b_1 \in I$ there is $J_2 = (a_2, b_2)$ in our finite subcollection such that $a_2 < b_1 < b_2$. Again, if $b_2 \geq b$ we are finished. If $b_2 < b$ we choose $J_3 = (a_3, b_3)$ from our finite collection, etc. Since the collection is finite, this process ends with $J_n = (a_n, b_n)$, $n \leq k$; and by hypothesis and construction we have $a_n < b < b_n$. Hence,

$$\begin{aligned} \sum_{j=1}^{k} m(J_j) &\geq \sum_{j=1}^{n} (b_j - a_j) \\ &= b_n - (a_n - b_{n-1}) - (a_{n-1} - b_{n-2}) - \cdots - (a_2 - b_1) - a_1. \end{aligned} \tag{2.6}$$

Now $a_j < b_{j-1}$ (recall $a_2 < b_1 < b_2$, etc.), and so, from (2.6),

$$\sum_{j=1}^{k} m(J_j) > b_n - a_1 > b - a;$$

this gives (2.5).

For the case of a bounded interval I, take any $\varepsilon > 0$ and choose a closed interval $J \subseteq I$ for which

$$m(J) > m(I) - \varepsilon.$$

Hence,

$$m(I) - \varepsilon < m(J) = m^*(J) \le m^*(I) \le m^*(\overline{I}) = m(\overline{I}) = m(I);$$

that is, for each $\varepsilon > 0$,

$$m(I) - \varepsilon < m^*(I) \le m(I),$$

and so $m(I) = m^*(I)$.

If I is an infinite interval then for each $r > 0$ there is a closed interval $J \subseteq I$ such that $m(J) = r$. Therefore, $m^*(I) \ge m^*(J) = m(J) = r$, and so, "letting $r \to \infty$", we have $m^*(I) = +\infty$. □

Proposition 2.2.2. *Let* $\{A_n : n = 1, \ldots\} \subseteq \mathcal{P}(\mathbb{R})$. *Then*

$$m^*\left(\bigcup_{n=1}^{\infty} A_n\right) \le \sum_{n=1}^{\infty} m^*(A_n). \tag{2.7}$$

Proof. If $m^*(A_n) = \infty$ for some n we are finished. Therefore, assume $m^*(A_n) < \infty$ for each $n \in \mathbb{N}$.

For a given $\varepsilon > 0$ and n, there is a sequence $\{I_{n,i} : n = 1, \ldots, i = 1, \ldots\}$ of open intervals such that

$$A_n \subseteq \bigcup_{i=1}^{\infty} I_{n,i} \quad \text{and} \quad \sum_{i=1}^{\infty} m(I_{n,i}) < m^*(A_n) + 2^{-n}\varepsilon,$$

by the definition of m^*. Now, card $\{I_{n,i} : i, n\} \le \aleph_0$ and $\bigcup_n A_n \subseteq \bigcup_{n,i} I_{n,i}$. Therefore,

$$\begin{aligned} m^*\left(\bigcup_{n=1}^{\infty} A_n\right) &\le \sum_{n,i=1}^{\infty} m(I_{n,i}) = \sum_{n=1}^{\infty}\sum_{i=1}^{\infty} m(I_{n,i}) \\ &\le \sum_{n=1}^{\infty} m^*(A_n) + \varepsilon \sum_{n=1}^{\infty} \frac{1}{2^n} = \sum_{n=1}^{\infty} m^*(A_n) + \varepsilon. \end{aligned}$$

Since this is true for each ε, we have (2.7). □

Taking A_n to be a one-point set in Proposition 2.2.2 we have that

i. if A is countable then $m^*(A) = 0$.

This fact, combined with Proposition 2.2.1, gives an ingenious though complicated proof that

ii. card $[0, 1] > \aleph_0$.

The following result is straightforward to prove and the proof is left as an exercise (Problem 2.1).

Proposition 2.2.3. *Let $A \subseteq \mathbb{R}$.*

a. *For each $\varepsilon > 0$ there is an open set $U \subseteq \mathbb{R}$ such that*

$$A \subseteq U \quad \text{and} \quad m^*(U) \leq m^*(A) + \varepsilon.$$

b. *There is a $\mathcal{G}_\delta$ set G such that*

$$A \subseteq G \quad \text{and} \quad m^*(A) = m^*(G).$$

We use the Constantin Carathéodory approach (1914) to define Lebesgue measure. A set $A \subseteq \mathbb{R}$ is *Lebesgue measurable* if

$$\forall\, E \subseteq \mathbb{R},\ \ m^*(E) = m^*(E \cap A) + m^*(E \cap A^{\sim}).$$

Thus, A is Lebesgue measurable if "no matter how you cut it" (with any E), m^* is nicely additive just as we hoped it would be. It is clear that

i. if $m^*(A) = 0$ then A is Lebesgue measurable;

ii. A is Lebesgue measurable if and only if $A^{\sim}$ is Lebesgue measurable; and

iii. if A_1, A_2 are Lebesgue measurable then $A_1 \cup A_2$ is Lebesgue measurable.

We shall use $\mathcal{M}(\mathbb{R})$ to denote the collection of all Lebesgue measurable subsets of $\mathbb{R}$. We have the following result by part *iii.*

Proposition 2.2.4. *$\mathcal{M}(\mathbb{R})$ is an algebra.*

More important is the following theorem.

Theorem 2.2.5. Lebesgue and Borel $\boldsymbol{\sigma}$-algebras on $\mathbb{R}$

a. *$\mathcal{M}(\mathbb{R})$ is a σ-algebra.*

b. *$\mathcal{B}(\mathbb{R}) \subseteq \mathcal{M}(\mathbb{R})$.*

Proof. ***a.****i.* We first show that if $E \subseteq \mathbb{R}$ and $\{A_1, \ldots, A_n\} \subseteq \mathcal{M}(\mathbb{R})$ is a disjoint family then

$$m^*\left(E \cap \left(\bigcup_{j=1}^{n} A_j\right)\right) = \sum_{j=1}^{n} m^*(E \cap A_j). \tag{2.8}$$

We shall prove (2.8) by induction, noting that it is obvious for $n = 1$. Assume that (2.8) is true for $A_1, \ldots, A_{n-1}$ and take $A_1, \ldots, A_n$. Then

$$E \cap \left(\bigcup_{j=1}^{n} A_j\right) \cap A_n = E \cap A_n \tag{2.9}$$

and

$$E \cap \left(\bigcup_{j=1}^{n} A_j\right) \cap A_n^{\sim} = E \cap \left(\bigcup_{j=1}^{n-1} A_j\right). \tag{2.10}$$

Equation (2.9) is clear, and (2.10) follows since $\left(\bigcup_{j=1}^{n} A_j\right) \cap A_n^{\sim} = \bigcup_{j=1}^{n} (A_j \cap A_j^{\sim}) = \left(\bigcup_{j=1}^{n-1} (A_j \cap A_n^{\sim})\right) \cup (A_n \cap A_n^{\sim}) = \bigcup_{j=1}^{n-1} A_j$, where we have used the fact that $\{A_1, \ldots, A_n\}$ is disjoint in the last step. Thus, $A_n \in \mathcal{M}(\mathbb{R})$ implies, from (2.9) and (2.10), that

$$\begin{aligned} m^*\left(E \cap \left(\bigcup_{j=1}^{n} A_j\right)\right) &= m^*(E \cap A_n) + m^*\left(E \cap \left(\bigcup_{j=1}^{n-1} A_j\right)\right) \\ &= m^*(E \cap A_n) + \sum_{j=1}^{n-1} m^*(E \cap A_j) \end{aligned}$$

by the induction hypothesis. Consequently, (2.8) holds.

***a.**ii.* We now prove that $\mathcal{M}(\mathbb{R})$ is a σ-algebra. Hence, given $\{B_j : j = 1, \ldots\} \subseteq \mathcal{M}(\mathbb{R})$ we must show that $A = \bigcup_{j=1}^{\infty} B_j \in \mathcal{M}(\mathbb{R})$. Since $\mathcal{M}(\mathbb{R})$ is an algebra, there is a disjoint family $\{A_j : j = 1, \ldots\} \subseteq \mathcal{M}(\mathbb{R})$ such that

$$A = \bigcup_{j=1}^{\infty} B_j = \bigcup_{j=1}^{\infty} A_j.$$

In fact, let $A_1 = B_1$, $A_2 = B_2 \setminus B_1$, $A_3 = B_3 \setminus (B_1 \cup B_2), \ldots$.

Let $C_n = \bigcup_{j=1}^{n} A_j$, so that $C_n \in \mathcal{M}(\mathbb{R})$, again using the fact that $\mathcal{M}(\mathbb{R})$ is an algebra. Next note that $A^{\sim} \subseteq C_n^{\sim}$ because

$$C_n = \bigcup_{j=1}^{n} A_j \subseteq A.$$

Taking any $E \subseteq \mathbb{R}$ we calculate

$$m^*(E) = m^*(E \cap C_n) + m^*(E \cap C_n^{\sim}) \geq m^*(E \cap C_n) + m^*(E \cap A^{\sim}). \tag{2.11}$$

From part *a.i* and the disjointness of $\{A_j : j = 1, \ldots\}$ we obtain

$$m^*(E \cap C_n) = m^*\left(E \cap \left(\bigcup_{j=1}^{n} A_j\right)\right) = \sum_{j=1}^{n} m^*(E \cap A_j). \tag{2.12}$$

Combining (2.11) and (2.12) gives

$$m^*(E) \geq \sum_{j=1}^{n} m^*(E \cap A_j) + m^*(E \cap A^{\sim}),$$

and, since the left-hand side is independent of n,

$$m^*(E) \geq \sum_{j=1}^{\infty} m^*(E \cap A_j) + m^*(E \cap A^{\sim}).$$

Thus, by the subadditivity of m^* (Proposition 2.2.2),

$$m^*(E) \geq m^*\left(E \cap \left(\bigcup_{j=1}^{\infty} A_j\right)\right) + m^*(E \cap A^{\sim}). \tag{2.13}$$

The opposite inequality to (2.13) is always true, and so

$$A \in \mathcal{M}(\mathbb{R}).$$

b.*i.* We first show that $(a, \infty) \in \mathcal{M}(\mathbb{R})$. Take $E \subseteq \mathbb{R}$ and set

$$E_1 = E \cap (a, \infty) \quad \text{and} \quad E_2 = E \cap (-\infty, a].$$

It is sufficient to prove that $\varepsilon + m^*(E) \geq m^*(E_1) + m^*(E_2)$ for each $\varepsilon > 0$, and without loss of generality (obviously) we take $m^*(E) < \infty$. Fix $\varepsilon > 0$. Since $m^*(E) < \infty$ there is a sequence $\{I_n : n = 1, \ldots\}$ of open intervals covering E such that

$$\sum_{n=1}^{\infty} m(I_n) \leq \varepsilon + m^*(E).$$

We observe that $I'_n = I_n \cap (a, \infty)$ and $I''_n = I_n \cap (-\infty, a]$ are intervals (possibly empty), and

$$m(I_n) = m(I'_n) + m(I''_n) = m^*(I'_n) + m^*(I''_n)$$

by Proposition 2.2.1. Clearly, $E_1 \subseteq \bigcup I'_n$ and $E_2 \subseteq \bigcup I''_n$. Consequently,

$$m^*(E_1) \leq m^*\left(\bigcup_{n=1}^{\infty} I'_n\right) \leq \sum_{n=1}^{\infty} m^*(I'_n)$$

and

$$m^*(E_2) \leq m^*\left(\bigcup_{n=1}^{\infty} I''_n\right) \leq \sum_{n=1}^{\infty} m^*(I''_n)$$

from Proposition 2.2.2. Therefore,

$$m^*(E_1) + m^*(E_2) \leq \sum_{n=1}^{\infty} (m^*(I'_n) + m^*(I''_n)) = \sum_{n=1}^{\infty} m^*(I_n) \leq \varepsilon + m^*(E).$$

b.*ii.* $\mathcal{B}(\mathbb{R})$ is the smallest σ-algebra containing each (a, ∞); but we have just proved that $\mathcal{M}(\mathbb{R})$ is a σ-algebra containing each (a, ∞). □

The following theorem is elementary to prove, and is the content of Problem 2.3.

Theorem 2.2.6. Properties of sequences of Lebesgue measurable sets
Let $\{A_n : n = 1, \ldots\} \subseteq \mathcal{M}(\mathbb{R})$.

a. $m\left(\bigcup_{n=1}^{\infty} A_n\right) \leq \sum_{n=1}^{\infty} m(A_n)$.

b. *If* $\{A_n : n = 1, \ldots\}$ *is a disjoint family, then* $m\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} m(A_n)$.

c. *If* $A_1 \supseteq A_2 \supseteq A_3 \supseteq \cdots \supseteq A_n \supseteq \cdots$ *and* $m(A_1) < \infty$, *then* $m\left(\bigcap_{n=1}^{\infty} A_n\right) = \lim_{n\to\infty} m(A_n)$.

d. *If* $A_1 \subseteq A_2 \subseteq A_3 \subseteq \cdots \subseteq A_n \subseteq \cdots$, *then* $m\left(\bigcup_{n=1}^{\infty} A_n\right) = \lim_{n\to\infty} m(A_n)$.

We shall prove Theorem 2.2.6*c,d* in a more general setting in Theorem 2.4.3.

Definition 2.2.7. Lebesgue measure
The set function $m : \mathcal{M}(\mathbb{R}) \to \mathbb{R}^+ \cup \{\infty\}$, defined by

$$\forall\, A \in \mathcal{M}(\mathbb{R}), \quad m(A) = m^*(A),$$

is *Lebesgue measure* on $\mathbb{R}$.

The following result is a consequence of the definition of Lebesgue measure. However, its usefulness in problem-solving cannot be overemphasized; and so we also state it as a theorem.

Theorem 2.2.8. Measure zero
A set $A \in \mathcal{M}(\mathbb{R})$ *has Lebesgue measure zero if and only if*

$$\forall\, \varepsilon > 0,\ \exists\ \{I_n : I_n \text{ an open interval and } n = 1, \ldots\} \text{ such that}$$
$$A \subseteq \bigcup_{n=1}^{\infty} I_n \ \text{ and } \ \sum_{n=1}^{\infty} m(I_n) < \varepsilon.$$

Norbert Wiener (AMS semicentennial) has made an historical case for the definition of measure 0 based on considerations from statistical mechanics; cf. [93] by Lennart Carleson.

The following result is interesting in light of the ruler function discussed in Section 1.3. We have only outlined its proof; cf. Problem 2.9.

Proposition 2.2.9. *There are a function* $f : [0,1] \to \mathbb{R}$ *and a set* $D \subseteq [0,1]$ *that is dense in* $[0,1]$ *such that* $D \in \mathcal{M}(\mathbb{R})$, f *is continuous on* D, f *is discontinuous on* $[0,1] \setminus D$, *and* $m(D) = 0$.

Proof. Let $\{r_n : n = 1, \ldots\}$ be an enumeration of $\mathbb{Q} \cap [0,1]$, denote the open interval about r_j of radius $1/(k2^j)$ by I_{jk}, and set

$$D = \bigcap_{k=3}^{\infty} U_k, \ \text{ where } U_k = \bigcup_{j=1}^{\infty} I_{jk}.$$

Then D is not of first category by an application of the Baire category theorem, and $m(D) = 0$, e.g., Theorem 2.2.6c. Define $f = \sum_{k \geq 3} f_k$, where

$$f_k(x) = \begin{cases} 0, & \text{if } x \in U_k, \\ 1/2^k, & \text{if } x \in [0,1] \setminus U_k. \end{cases}$$

□

Example 2.2.10. Measure of perfect symmetric sets
Because of Theorem 2.2.6c, if E is a perfect symmetric set then the definition of $m(E)$ that we gave in Chapter 1 is precisely the Lebesgue measure of E.

Proposition 2.2.11. *Let $A \subseteq \mathbb{R}$ have the property that there is $q \in (0,1)$ such that for all $(a,b) \subseteq \mathbb{R}$ there exist intervals $I_n = (a_n, b_n)$, $n = 1, \ldots,$ such that*

$$A \cap (a,b) \subseteq \bigcup_{n=1}^{\infty} I_n \quad \text{and} \quad \sum_{n=1}^{\infty} m(I_n) \leq q(b-a).$$

Then $m(A) = 0$.

Proof. It is sufficient to prove that $m(A \cap (a,b)) = 0$, since we have $m(A) = m(A \cap (\bigcup J_i)) \leq \sum m(A \cap J_i)$, where $\{J_i : i = 1, \ldots\}$ is a cover of $\mathbb{R}$ by open intervals. Now cover $A \cap (a,b)$ by intervals (a_n, b_n) having the property that $\sum (b_n - a_n) \leq q(b-a)$. Then cover each $A \cap (a_n, b_n)$ by a countable collection of intervals having total length $q(b_n - a_n)$. Thus, we have covered A by open intervals of total length L, where

$$L \leq q(b_1 - a_1) + q(b_2 - a_2) + \cdots = q((b_1 - a_1) + \cdots) \leq q(b-a).$$

Repeating this process we obtain the inequality $m(A) \leq q^n(b-a)$, and so $m(A) = 0$, since $0 < q < 1$. □

With regard to Proposition 2.2.3 we have the following theorem.

Theorem 2.2.12. Properties of Lebesgue outer measure
The following are equivalent.

- ***a.*** $A \in \mathcal{M}(\mathbb{R})$.
- ***b.*** $\forall\, \varepsilon > 0,\ \exists\, U \supseteq A$, *open, such that* $m^*(U \setminus A) < \varepsilon$.
- ***c.*** $\forall\, \varepsilon > 0,\ \exists\, F \subseteq A$, *closed, such that* $m^*(A \setminus F) < \varepsilon$.
- ***d.*** $\exists\, G$, *a* $\mathcal{G}_\delta$ *set, such that* $A \subseteq G$ *and* $m^*(G \setminus A) = 0$.
- ***e.*** $\exists\, F$, *an* $\mathcal{F}_\sigma$ *set, such that* $F \subseteq A$ *and* $m^*(A \setminus F) = 0$.

The straightforward proof of Theorem 2.2.12 is left as an exercise (Problem 2.11).

Let $A \in \mathcal{M}(\mathbb{R})$. Suppose $G \subseteq A$ is Lebesgue measurable, so that $A \setminus G \in \mathcal{M}(\mathbb{R})$. Then

$$A = G \cup (A \setminus G).$$

Because of this and Theorem 2.2.12 we obtain the following important fact.

Theorem 2.2.13. Lebesgue measurable sets and Borel sets
Let $A \in \mathcal{M}(\mathbb{R})$. Then,

$$\exists\, B \in \mathcal{B}(\mathbb{R}) \text{ and } \exists\, E \in \mathcal{M}(\mathbb{R}),\ m(E) = 0, \text{ such that}$$
$$A = B \cup E,\ B \cap E = \emptyset, \text{ and } m(A) = m(B),$$

i.e., every Lebesgue measurable set is a Borel set up to a set of measure zero.

Example 2.2.14. Closed uncountable sets of irrationals
In Problem 1.4a we wanted to find a closed uncountable subset of the irrationals. This is most easily done by taking an open interval of radius $\varepsilon/2^n$ about the nth rational and then looking at the complement of the union of these intervals. Theorem 2.2.12c gives another proof, since $m^*([0,1] \cap (\mathbb{R} \setminus \mathbb{Q})) = 1$, i.e., take $A = [0,1] \cap (\mathbb{R} \setminus \mathbb{Q})$ and use the facts that its outer measure is positive and that it is measurable.

With regard to Problem 1.4b, if $X \in \mathcal{M}(\mathbb{R})$ and $m(X) > 0$, then we can find a closed uncountable set $F \subseteq X$ by the same reasoning, i.e., using Theorem 2.2.12c, since $m(X \setminus F) = m(X) - m(F)$.

Example 2.2.15. A set of measure 0
Let A be the set of those $x \in [0,1]$ whose decimal expansion consists of not more than 9 distinct digits. Then, card $A > \aleph_0$ and $m(A) = 0$. To prove this, let $S_j \subseteq A$ consist of those elements whose decimal expansions do not contain j. Thus, $\bigcup_0^9 S_j = A$. Now S_j is equivalent, bijectively, to $[0,1]$ when we view $x \in [0,1]$ as having an expansion to the base 9. Thus, card $S_j > \aleph_0$ and so card $A > \aleph_0$.

The proof that $m(A) = 0$ reduces to showing that for each j, $m(S_j) = 0$. Fix j and consider the ten intervals $[k/10, (k+1)/10)$, $k = 0, \ldots, 9$. Then throw away the $(j+1)$st interval and so dispense with all decimals whose first term is j; note that the length of this interval is $1/10$. Next we divide each of the remaining 9 intervals into 10 intervals, each of length $1/10^2$. From each of these divisions we throw away the $(j+1)$st interval and thus we have removed all decimals whose second term is j; at this stage, then, we dispense with lengths totaling $9/10^2$. Continuing this process we end up with precisely S_j; on the other hand, we have thrown away, altogether, lengths totaling

$$\frac{1}{10} + \frac{9}{10^2} + \frac{9^2}{10^3} + \cdots = \frac{1}{10}\left(\frac{1}{1 - (9/10)}\right) = 1,$$

and so $m(S_j) = 0$.

We have indicated that there is a continuum of Borel subsets of $\mathbb{R}$ (Example 2.1.4). Assuming this for the moment, we can assert that $\mathcal{M}(\mathbb{R}) \setminus \mathcal{B}(\mathbb{R}) \neq \emptyset$. In fact, all subsets of the ternary Cantor set C are measurable. Since the Cantor set is uncountable, the number of subsets of C (i.e., the cardinality of the power set $\mathcal{P}(C)$) is strictly greater than a continuum, and so there exist measurable subsets of C that are not Borel. See Example 2.4.14 for another proof that $\mathcal{M}(\mathbb{R}) \setminus \mathcal{B}(\mathbb{R}) \neq \emptyset$.

In 1904, LEBESGUE [313] posed the *problem of measure*: *Does there exist a nontrivial σ-additive set function on $\mathcal{P}(\mathbb{R})$ that is translation-invariant on $\mathcal{P}(\mathbb{R})$ and satisfies* (2.2)? GIUSEPPE VITALI settled this question in the negative in 1905 [483]. In particular, this means that $\mathcal{P}(\mathbb{R}) \setminus \mathcal{M}(\mathbb{R}) \neq \emptyset$.

Example 2.2.16. Nonmeasurable sets

a. The following is our translation of VITALI's ingenious original argument of 1905 [483]. We include it, since the publication is difficult to obtain. Subsequently, there have been many other published proofs, e.g., [438], frequently similar, and sometimes more complicated.

On the problem of the measure of subsets of a line by G. VITALI.

The problem of [measuring all bounded] subsets $A \subseteq \mathbb{R}$ is that of determining, for every such A, a positive number $\mu(A)$ that could be called the *measure* of A. This measure should satisfy the following properties (see [313], page 103):

i. $\forall\, x \in \mathbb{R}$ and $\forall\, A \subseteq \mathbb{R}$, bounded,

$$\mu(A) = \mu(A + x);$$

ii. $\forall\, \{A_n \subseteq \mathbb{R} : A_n \text{ bounded}\}$, a disjoint sequence,

$$\sum_{n=1}^{\infty} \mu(A_n) = \mu\left(\bigcup_{n=1}^{\infty} A_n\right);$$

iii. $\mu((0,1)) = 1$.

Let $x \in \mathbb{R}$. The points of $\mathbb{R}$ that differ from x by some element of $\mathbb{Q}$ form a denumerable set $A_x \subseteq \mathbb{R}$. If A_{x_1} and A_{x_2} are two such sets then

$$A_{x_1} \cap A_{x_2} = \emptyset \quad \text{or} \quad A_{x_1} = A_{x_2}. \tag{2.14}$$

Let us consider the various sets A_x, well defined by (2.14), and let $\mathcal{G}$ be the family of these sets. If $r \in \mathbb{R}$, then there is a unique element $A \in \mathcal{G}$, i.e., A is some A_x, for which $r \in A$. For $A \in \mathcal{G}$, choose a point $r_A \in (0, 1/2)$ for which $r_A \in A$; and let us denote by V_0 the set of such points r_A. If $q \in \mathbb{Q}$, we denote by V_q the set of points $r_A + q$, i.e.,

$$V_q = \{r_A + q \in (0, 1/2) + q : A \in \mathcal{G}\}.$$

The sets V_q corresponding to various $q \in \mathbb{Q}$ are pairwise disjoint, there are countably infinitely many of them, and by the translation-invariance [property *i*] they all have the same measure.

The sets

$$V_0, V_{1/2}, V_{1/3}, V_{1/4}, \ldots \subseteq (0, 1);$$

and therefore their union has measure $m \leq 1$ [by properties *ii* and *iii*, if, in fact, the V_q are "measurable"].

However,

$$m = \mu(V_0) + \sum_{n=2}^{\infty} \mu(V_{1/n}) = \lim_{n\to\infty} n\mu(V_0)$$

must also hold, and therefore $\mu(V_0) = 0$. But then the union of all the V_q, $q \in \mathbb{Q}$, must likewise have measure 0. However, this union is $\mathbb{R}$ and therefore must have infinite measure. Thus, we may conclude that *it is impossible to define* $\mu : \mathcal{P}(\mathbb{R}) \to \mathbb{R}^+$ *satisfying the properties i, ii, iii.*

b. Next we include the VITALI example written in more modern language.

For $x, y \in \mathbb{R}$ we write $x \sim y$ if $x - y \in \mathbb{Q}$. Clearly, "$\sim$" is an *equivalence relation*, i.e., it is

i. Reflexive: $\forall\, x \in \mathbb{R},\ x \sim x$,

ii. Symmetric: $\forall\, x, y \in \mathbb{R},\ x \sim y \Longrightarrow y \sim x$,

iii. Transitive: $\forall\, x, y, z \in \mathbb{R},\ x \sim y$ and $y \sim z \Longrightarrow x \sim z$.

The *equivalence class* corresponding to $x \in \mathbb{R}$ is $\{y \in \mathbb{R} : y \sim x\}$.

We now make explicit use of the axiom of choice and define A to be a subset of $(0, 1)$ that contains exactly one point from each equivalence class, noting, for each $x \in \mathbb{R}$, that an equivalence class is of the form $A_x = x + \mathbb{Q}$. We shall prove that $A \notin \mathcal{M}(\mathbb{R})$. Clearly,

$$\text{if } x \in (0,1) \text{ then } \exists\, r \in \mathbb{Q} \cap (-1,1) \text{ such that } x \in A + r;$$

also, by the definition of A we argue by contradiction to prove that

$$\text{if } r, s \in \mathbb{Q},\ r \neq s,\ \text{then } (A + r) \cap (A + s) = \emptyset.$$

Assume $A \in \mathcal{M}(\mathbb{R})$ and define

$$S = \bigcup \{r + A :\ r \in \mathbb{Q} \cap (-1,1)\}.$$

From the translation invariance and the fact that $(A + r) \cap (A + s) = \emptyset$ for $r \neq s$, where $r, s \in \mathbb{Q}$, we have that $S \in \mathcal{M}(\mathbb{R})$, and

$$m(S) = \infty \quad \text{or} \quad m(S) = 0.$$

Since $A \subseteq (0, 1)$, $S \subseteq (-1, 2)$, and so $m(S) \leq 3$. Hence, $m(S) = 0$. On the other hand, $(0, 1) \subseteq S$; and so $m(S) \geq 1$, the desired contradiction. Therefore, $A \notin \mathcal{M}(\mathbb{R})$, i.e., $\mathcal{M}(\mathbb{R})$ cannot be all of $\mathcal{P}(\mathbb{R})$ if m is to satisfy (2.2) as well as being translation invariant.

The following result is true in any measure space (see Section 2.4) having nonmeasurable sets, and indicates how difficult it is to approximate nonmeasurable sets with measurable ones.

Proposition 2.2.17. *Let* $E \subseteq \mathbb{R}$ *be nonmeasurable. There is* $\varepsilon > 0$ *such that if* $E \subseteq A$ *and* $E^{\sim} \subseteq B$*, where* A *and* B *are measurable, then* $m(A \cap B) \geq \varepsilon$*.*

Proof. Assume that the result is false. Then for each n there are sets G_n, $D_n \in \mathcal{M}(\mathbb{R})$ such that

$$E \subseteq G_n,\ E^\sim \subseteq D_n,\ \text{and}\ m(G_n \cap D_n) < 1/n.$$

The sets $G = \bigcap G_n$ and $D = \bigcap D_n$ are Lebesgue measurable, $E \subseteq G, E^\sim \subseteq D$, and $m(G \cap D) = 0$. Since $G \in \mathcal{M}(\mathbb{R})$, $m^*(S) = m^*(S \cap G) + m^*(S \cap G^\sim)$ for each $S \subseteq \mathbb{R}$. Now, $D \subseteq (G \cap D) \cup G^\sim$, and so

$$S \cap D \subseteq (S \cap G \cap D) \cup (S \cap G^\sim).$$

This implies that

$$\begin{aligned} m^*(S \cap D) &\leq m^*(S \cap G^\sim) + m^*((S \cap G) \cap D) \\ &\leq m^*(S \cap G^\sim) + m^*(G \cap D) = m^*(S \cap G^\sim). \end{aligned}$$

Consequently, $m^*(S \cap D) + m^*(S \cap G) \leq m^*(S \cap G^\sim) + m^*(S \cap G) = m^*(S)$. Observe that

$$E \subseteq G \Longrightarrow m^*(S \cap E) \leq m^*(S \cap G)$$

and

$$E^\sim \subseteq D \Longrightarrow m^*(S \cap E^\sim) \leq m^*(S \cap D).$$

Thus, for all $S \subseteq \mathbb{R}$,

$$m^*(S \cap E) + m^*(S \cap E^\sim) \leq m^*(S),$$

and so $E \in \mathcal{M}(\mathbb{R})$, a contradiction. □

2.3 The existence of Lebesgue measure on $\mathbb{R}^d$

In the previous section we have established the existence of Lebesgue measure m on $\mathbb{R}$. This measure is translation-invariant, it measures all Borel sets, and the Lebesgue measure of an interval is equal to its length.

In this section we shall construct an analogue of Lebesgue measure on $\mathbb{R}$ for $\mathbb{R}^d$ for any $d \geq 1$. The procedure is general in the sense that we shall actually give a method of construction of a family of measures on $\mathbb{R}^d$. Then, a choice of a specific ring and a specific set function on that ring will yield the desired analogue of Lebesgue measure. In this context, note that the family of open sets in $\mathbb{R}^d$ does not form a ring of subsets of $\mathbb{R}^d$.

Remark. We shall show that it is possible to extend an arbitrary σ-additive function on a ring $\mathcal{R}$ to a σ-additive function on a σ-ring generated by $\mathcal{R}$. Thus, formally, our construction is not a construction of an arbitrary general measure, which by definition is a σ-additive function on a σ-algebra. However, this approach combined with the properties of the standard volume function on $\mathbb{R}^d$ will yield Lebesgue measure on $\mathbb{R}^d$.

Let $\mathcal{R} \subseteq \mathcal{P}(\mathbb{R}^d)$ be a ring and let $\mu : \mathcal{R} \to \mathbb{R}^+ \cup \{\infty\}$ be a σ-additive set function. Our goal is to extend the function μ from the ring $\mathcal{R}$ to a σ-ring $\mathcal{R}_\sigma$ that contains $\mathcal{R}$. As before, for $A \subseteq \mathbb{R}^d$ we define the *outer measure* of A as

$$\mu^*(A) = \inf_{\{I_n\}} \sum_{n=1}^{\infty} \mu(I_n), \tag{2.15}$$

where the infimum is taken over all countable families $\{I_n\}$ of elements of the ring $\mathcal{R}$, which form a cover of the set A, i.e., $A \subseteq \bigcup_n I_n$ and $I_n \in \mathcal{R}$, $n = 1, \ldots$; cf. (2.3).

Example 2.3.1. Note that depending on the choice of the set X and the ring $\mathcal{R}$, it may happen that it will not be possible to define the outer measure on all of $\mathcal{P}(X)$. Consider X to be any uncountable set and let

$$\mathcal{R} = \{A \subseteq X : \operatorname{card} A < \infty\},$$

the ring of all finite sets, and let $\mu : \mathcal{R} \to \mathbb{R}^+$ be defined as

$$\mu(A) = \operatorname{card} A.$$

Then the outer measure μ^* can be extended only to the collection of countable subsets of X.

However, it is not difficult to check that the collection

$$\mathcal{H}(\mathcal{R}) = \left\{ A \subseteq \mathbb{R}^d : \ A \subseteq \bigcup_{n=1}^{\infty} I_n, \ I_n \in \mathcal{R} \right\}$$

is a σ-ring. Thus, μ^* is defined on the σ-ring $\mathcal{H}(\mathcal{R}) \subseteq \mathcal{P}(\mathbb{R}^d)$.

The σ-ring $\mathcal{H}(\mathcal{R})$ has the following property: if $A \subseteq B$ and $B \in \mathcal{H}(\mathcal{R})$ then $A \in \mathcal{H}(\mathcal{R})$. This explains the use of letter "h", as in "heritage". If $\mathcal{R} \subseteq \mathcal{P}(\mathbb{R}^d)$ is a ring and $\mathcal{R}_\sigma$ is the σ-ring generated by $\mathcal{R}$, then $\mathcal{R}_\sigma \subseteq \mathcal{H}(\mathcal{R})$.

Remark. We note that if the ring $\mathcal{R}$ contains all open subsets of $\mathbb{R}^d$, then $\mathcal{H}(\mathcal{R}) = \mathcal{P}(\mathbb{R}^d)$. In this case, the simpler arguments of Section 2.2 can be used to prove the existence of Lebesgue measure on $\mathbb{R}^d$. However, in the case of extensions of a ring to a σ-algebra for arbitrary measures in $\mathbb{R}^d$, as well as for the general measure theory developed in Section 2.4, we need the notion of the ring $\mathcal{H}(\mathcal{R})$.

Proposition 2.3.2. *Let $\mu : \mathcal{R} \to \mathbb{R}^+ \cup \{+\infty\}$ be a σ-additive set function on a ring $\mathcal{R} \subseteq \mathcal{P}(\mathbb{R}^d)$ and let $\{A_n : n = 1, \ldots\} \subseteq \mathcal{H}(\mathcal{R})$. Then*

$$\mu^*\left(\bigcup_{n=1}^{\infty} A_n \right) \le \sum_{n=1}^{\infty} \mu^*(A_n). \tag{2.16}$$

The proof of Proposition 2.3.2 follows the proof of Proposition 2.2.2.

Proposition 2.3.3. *Let $\mu : \mathcal{R} \to \mathbb{R}^+ \cup \{+\infty\}$ be a σ-additive set function on a ring $\mathcal{R} \subseteq \mathcal{P}(\mathbb{R}^d)$. If $A \in \mathcal{R}$ then $\mu^*(A) = \mu(A)$.*

Proof. Clearly $A \subseteq A$ and thus $\mu^*(A) \leq \mu(A)$. To prove the opposite inequality, consider a collection $\{I_n : n = 1, \ldots\} \subseteq \mathcal{R}$ such that $A \subseteq \bigcup_{n=1}^{\infty} I_n$. From the σ-additivity of the set function μ on $\mathcal{R}$ it follows that, for any such collection $\{I_n\}$,

$$\mu(A) \leq \sum_{n=1}^{\infty} \mu(I_n);$$

see Problem 2.14. Therefore, by the definition of outer measure we have $\mu(A) \leq \mu^*(A)$. □

We again use Carathéodory's approach to define measurable sets. Let $\mu : \mathcal{R} \to \mathbb{R}^+ \cup \{+\infty\}$ be a σ-additive set function on a ring $\mathcal{R} \subseteq \mathcal{P}(\mathbb{R}^d)$ and let $A \in \mathcal{H}(\mathcal{R})$. Then A is *measurable* with respect to μ and $\mathcal{R}$ if

$$\forall\, E \in \mathcal{H}(\mathcal{R}),\ \mu^*(E) = \mu^*(E \cap A) + \mu^*(E \cap A^{\sim}).$$

Let $\mathcal{A}$ denote the family of all sets measurable with respect to μ and $\mathcal{R}$.

By the definition of the outer measure μ^*,

$$A, B \in \mathcal{H}(\mathcal{R}) \text{ and } A \subseteq B \implies \mu^*(A) \leq \mu^*(B).$$

Proposition 2.3.4. *Let $\mu : \mathcal{R} \to \mathbb{R}^+ \cup \{+\infty\}$ be a σ-additive set function on a ring $\mathcal{R} \subseteq \mathcal{P}(\mathbb{R}^d)$. The elements of $\mathcal{R}$ are measurable with respect to μ and $\mathcal{R}$, i.e., $\mathcal{R} \subseteq \mathcal{A}$.*

Proof. From Proposition 2.3.2 it is clear that $\mu^*(E) \leq \mu^*(E \cap A) + \mu^*(E \cap A^{\sim})$. To prove the opposite inequality, fix $\varepsilon > 0$ and take a collection $\{I_n : n = 1, \ldots\} \subseteq \mathcal{R}$ such that $E \subseteq \bigcup_{n=1}^{\infty} I_n$ and

$$\sum_{n=1}^{\infty} \mu(I_n) < \mu^*(E) + \varepsilon.$$

Since $E \cap A \subseteq \bigcup (I_n \cap A)$ and $E \cap A^{\sim} \subseteq \bigcup (I_n \setminus A)$, we obtain

$$\begin{aligned}\mu^*(E) &> \sum_{n=1}^{\infty} \mu(I_n) - \varepsilon = \sum_{n=1}^{\infty} \left(\mu(I_n \cap A) + \mu(I_n \setminus A)\right) - \varepsilon \\ &\geq \mu^*(E \cap A) + \mu^*(E \cap A^{\sim}) - \varepsilon.\end{aligned}$$

This is true for each ε, and so $\mu^*(E) \geq \mu^*(E \cap A) + \mu^*(E \cap A^{\sim})$. □

Theorem 2.3.5. $\mathcal{A}$ is a ring

Let $\mu : \mathcal{R} \to \mathbb{R}^+ \cup \{+\infty\}$ be a σ-additive set function on a ring $\mathcal{R} \subseteq \mathcal{P}(\mathbb{R}^d)$. Then $\mathcal{A}$ is a ring.

Proof. We need to show only that, for any two measurable sets A, $B \in \mathcal{A}$, we have $A \cup B, A \setminus B \in \mathcal{A}$.

For any $E \in \mathcal{H}(\mathcal{R})$,

$$\begin{aligned}\mu^*(E) &= \mu^*(E \cap A) + \mu^*(E \cap A^\sim)\\ &= \mu^*(E \cap A \cap B) + \mu^*(E \cap A \cap B^\sim) + \mu^*(E \cap A^\sim)\\ &\geq \mu^*(E \cap A \cap B^\sim) + \mu^*(E \cap (A \cap B) \cup (E \cap A^\sim))\\ &= \mu^*(E \cap (A \setminus B)) + \mu^*(E \cap (A \setminus B)^\sim).\end{aligned}$$

Since the reverse inequality is always true, it follows that $A \setminus B \in \mathcal{A}$.

In a similar manner we obtain

$$\mu^*(E) = \mu^*(E\cap A\cap B)+\mu^*(E\cap A\cap B^\sim)+\mu^*(E\cap A^\sim\cap B)+\mu^*(E\cap A^\sim\cap B^\sim),$$

and, replacing E with $E \cap (A \cup B)$, we have that

$$\mu^*(E \cap (A \cup B)) = \mu^*(E \cap A \cap B) + \mu^*(E \cap A \cap B^\sim) + \mu^*(E \cap A^\sim \cap B).$$

Thus,

$$\mu^*(E) = \mu^*(E \cap (A \cup B)) + \mu^*(E \cap (A \cup B)^\sim),$$

and so $A \cup B \in \mathcal{A}$. □

To finish the construction of a measure on $\mathbb{R}^d$, we need the following observation: if $A_1, \ldots, A_n \in \mathcal{A}$ are pairwise disjoint and if $E \in \mathcal{H}(\mathcal{R})$, then

$$\mu^*\left(E \cap \left(\bigcup_{j=1}^{n} A_j\right)\right) = \sum_{j=1}^{n} \mu^*(E \cap A_j). \tag{2.17}$$

We leave the proof as an exercise (Problem 2.15).

Theorem 2.3.6. $\boldsymbol{\mathcal{A}}$ is a $\boldsymbol{\sigma}$-ring
Let $\mu : \mathcal{R} \to \mathbb{R}^+ \cup \{+\infty\}$ be a σ-additive set function on a ring $\mathcal{R} \subseteq \mathcal{P}(\mathbb{R}^d)$. Then $\mathcal{A}$ is a σ-ring.

Proof. In view of Theorem 2.3.5 we need to show only that $\mathcal{A}$ is closed under countable disjoint unions; see Problem 2.16.

Let $E \in \mathcal{H}(\mathcal{R})$, let $\{A_n : n = 1, \ldots\} \subseteq \mathcal{A}$ be a sequence of pairwise disjoint measurable sets, and let $A = \bigcup_{n=1}^{\infty} A_n$. We may assume without loss of generality that $\mu^*(E\cap A) < \infty$. Indeed, if $\mu^*(E\cap A) = \infty$ then $\mu^*(E) = \infty$, and A is measurable.

From Proposition 2.3.2, we have

$$\mu^*(E \cap A) \leq \sum_{n=1}^{\infty} \mu^*(E \cap A_n).$$

Now it follows from (2.17) that for every $\varepsilon > 0$ there exists N such that

$$\mu^*(E \cap A) \le \sum_{n=1}^{N} \mu^*(E \cap A_n) + \varepsilon = \mu^*\left(E \cap \left(\bigcup_{n=1}^{N} A_n\right)\right) + \varepsilon.$$

Because of this and the fact that $\mathcal{A}$ is a ring, in particular $\bigcup_{n=1}^{N} A_n \in \mathcal{A}$, we have

$$\begin{aligned} &\mu^*(E \cap A) + \mu^*(E \cap A^{\sim}) \\ &\quad \le \mu^*\left(E \cap \left(\bigcup_{n=1}^{N} A_n\right)\right) + \varepsilon + \mu^*\left(E \cap \left(\bigcup_{n=1}^{N} A_n\right)^{\sim}\right) \\ &\quad = \mu^*(E) + \varepsilon. \end{aligned}$$

Since this is true for each ε, $\mu^*(E) \ge \mu^*(E \cap A) + \mu^*(E \cap A^{\sim})$. The opposite inequality is easily verified, and so $A = \bigcup_{n=1}^{\infty} A_n \in \mathcal{A}$. □

Theorem 2.3.7. The σ-additive function μ^* on the σ-ring $\mathcal{A}$
Let $\mu : \mathcal{R} \to \mathbb{R}^+ \cup \{+\infty\}$ be a σ-additive set function on a ring $\mathcal{R} \subseteq \mathcal{P}(\mathbb{R}^d)$. Then μ^* is σ-additive on the σ-ring $\mathcal{A}$.

Proof. We shall prove that μ^* is σ-additive on $\mathcal{A}$. Let $A = \bigcup_{n=1}^{\infty} A_n$, where each $A_n \in \mathcal{A}$ and the A_ns are pairwise disjoint. We shall show that for any $E \in \mathcal{H}(\mathcal{R})$ we have

$$\mu^*(E \cap A) = \sum_{n=1}^{\infty} \mu^*(E \cap A_n). \tag{2.18}$$

Then, substituting A for E, we shall have the desired σ-additivity,

$$\mu^*(A) = \sum_{n=1}^{\infty} \mu^*(A_n).$$

Since $E \cap A = E \cap (\bigcup_{n=1}^{\infty} A_n)$, Proposition 2.3.2 implies

$$\mu^*(E \cap A) \le \sum_{n=1}^{\infty} \mu^*(E \cap A_n),$$

and we have one direction of (2.18). For the opposite inequality, note that, for every N,

$$\sum_{n=1}^{N} \mu^*(E \cap A_n) = \mu^*\left(E \cap \left(\bigcup_{n=1}^{N} A_n\right)\right) \le \mu^*(E \cap A).$$

Thus,

$$\sum_{n=1}^{\infty} \mu^*(E \cap A_n) \le \mu^*(E \cap A).$$

□

Therefore, we have extended a nonnegative, σ-additive set function μ from a ring $\mathcal{R}$ to a σ-ring $\mathcal{A}$. This process is in fact parallel to our construction of Lebesgue measure on $\mathbb{R}$, where the set function was defined as the length of an interval.

Example 2.3.8. Volumes of parallelepipeds
Let $\mathcal{Q}$ be the collection of parallelepipeds in $\mathbb{R}^d$, i.e., sets of the form $Q = (a_1, b_1] \times \cdots \times (a_d, b_d]$, for any real $a_1 < b_1, \ldots, a_d < b_d$. Let $\mathcal{A}_{\mathcal{Q}}$ be the family of all finite unions of pairwise disjoint collections of such parallelepipeds. Then $\mathcal{A}_{\mathcal{Q}}$ is a ring.

Let $m^d : \mathcal{A}_{\mathcal{Q}} \to \mathbb{R}^+$ be defined as

$$m^d\left(\bigcup_{j=1}^n Q_j\right) = \sum_{j=1}^n \text{volume } (Q_j) = \sum_{j=1}^n \prod_{k=1}^d (b_k^{(j)} - a_k^{(j)}),$$

where $Q_j = (a_1^{(j)}, b_1^{(j)}] \times \cdots \times (a_d^{(j)}, b_d^{(j)}]$, $j = 1, \ldots, n$, are disjoint and where $\bigcup_{j=1}^n Q_j \in \mathcal{A}_{\mathcal{Q}}$. It is not difficult to see that m^d is a σ-additive set function on $\mathcal{A}_{\mathcal{Q}}$. Thus, we define the *Lebesgue outer measure* m^{d*}, according to (2.15). Then m^{d*} is defined on the corresponding σ-ring $\mathcal{H}(\mathcal{A}_{\mathcal{Q}})$, which in this case is equal to $\mathcal{P}(\mathbb{R}^d)$. Therefore, $\mathcal{H}(\mathcal{A}_{\mathcal{Q}})$ is a σ-algebra. In particular, when $d = 1$, the definition of the outer measure m^{1*} coincides with the definition of the Lebesgue outer measure on $\mathbb{R}$ in Section 2.1.

We now apply Theorem 2.3.7 to the setting of Example 2.3.8. In particular, Lebesgue outer measure m^{d*} is σ-additive on a σ-ring $\mathcal{M}(\mathbb{R}^d)$. Since $\mathbb{R}^d \in \mathcal{H}(\mathcal{A}_{\mathcal{Q}}) = \mathcal{P}(\mathbb{R}^d)$, we can write

$$\forall\, E \in \mathcal{H}(\mathcal{A}_{\mathcal{Q}}), \quad m^{d*}(E \cap \mathbb{R}^d) + m^{d*}(E \setminus \mathbb{R}^d) = m^{d*}(E) + m^{d*}(\emptyset).$$

Clearly $m^{d*}(\emptyset) = 0$ and so $\mathbb{R}^d \in \mathcal{M}(\mathbb{R}^d)$, which implies that $\mathcal{M}(\mathbb{R}^d)$ *is in fact a σ-algebra.* Therefore, we have the following result.

Theorem 2.3.9. Lebesgue measure on $\mathbb{R}^d$
There exist a σ-algebra $\mathcal{M}(\mathbb{R}^d) \supseteq \mathcal{B}(\mathbb{R}^d)$ and a measure m^d on $\mathcal{M}(\mathbb{R}^d)$ such that

$$\forall\, x \in \mathbb{R}^d \text{ and } \forall\, A \in \mathcal{M}(\mathbb{R}^d), \quad m^d(A + x) = m^d(A),$$

and

$$\forall\, Q \subseteq \mathbb{R}^d,\ Q \text{ a cube, } m^d(Q) \text{ is the volume of } Q.$$

A measure that extends the set function m^d from Example 2.3.8, defined on $\mathcal{A}_{\mathcal{Q}}$, to the σ-algebra $\mathcal{M}(\mathbb{R}^d)$ of sets measurable with respect to $\mathcal{A}_{\mathcal{Q}}$ is unique and is called *Lebesgue measure* on $\mathbb{R}^d$. It is obviously translation invariant, and the σ-algebra $\mathcal{M}(\mathbb{R}^d)$ contains all Borel subsets of $\mathbb{R}^d$. The uniqueness follows from Problem 2.20; cf. the more general uniqueness result asserted in Problem 2.19.

Naturally, for $d = 1$, $m^1 = m$, where m is Lebesgue measure as defined in Section 2.2. In fact, our construction of the Lebesgue measure m^d is analogous to our treatment of measure for $\mathbb{R}$, differing only in the fact that we were using d-dimensional "cubes" instead of open intervals (which do not form a ring of subsets of $\mathbb{R}$).

Example 2.3.10. Lebesgue–Stieltjes measure associated with f
Let $f : \mathbb{R} \to \mathbb{R}$ be an increasing and right-continuous function. Let $Q = (a_1, b_1] \times \cdots \times (a_d, b_d] \subseteq \mathbb{R}^d$, and define the set function

$$\mu_f(Q) = (f(b_1) - f(a_1)) \cdots (f(b_d) - f(a_d)).$$

It is not difficult to see that μ_f is a σ-additive set function on the ring $\mathcal{Q}$ of disjoint unions of parallelepipeds in $\mathbb{R}^d$ (Problem 2.21). The procedure culminating in Theorem 2.3.7 allows us to extend μ_f to a measure on a σ-algebra of measurable sets generated by $\mathcal{A}_{\mathcal{Q}}$. This measure, also denoted by μ_f, is the *Lebesgue–Stieltjes measure* associated with f (see Section 3.5), and *Lebesgue measure is the Lebesgue–Stieltjes measure associated with* $f = 1$. Lebesgue–Stieltjes measures will be identified with Radon measures through the Riesz representation theorem in Section 7.5.

Systematic treatments of measure constructions on $\mathbb{R}^d$, from various points of view, are given in the books on Lebesgue integration by Edgar Asplund and Lutz Bungart [15], Jacques Dixmier [142] (course at the Sorbonne), Wendell H. Fleming [179], pages 136–205, B. L. Gurevich and Georgii E. Shilov [208], Frank Jones [264], Elliot H. Lieb and Michael Loss [321], Walter Rudin [405], pages 49–52, Kennan T. Smith [445], Richard L. Wheeden and Antoni Zygmund [502], and John H. Williamson [512]. Also, see [135], [180], [287], [302].

Example 2.3.11. Lebesgue measure of the unit ball
The measure of the unit ball

$$B_d = \left\{ x \in \mathbb{R}^d : \|x\| = \left(\sum_{j=1}^{d} x_j^2 \right)^{1/2} \leq 1 \right\}$$

is

$$m^d(B_d) = \frac{\pi^{d/2}}{(d/2)\Gamma(d/2)}, \;\; d = 1, 2, \ldots,$$

where the *Euler gamma function* is defined as

$$\Gamma(x) = \int_0^\infty t^{x-1} e^{-t} dt, \;\; x > 0.$$

The proof is by induction. First we observe that $m(B_1) = 2$. Next, we calculate

$$\frac{m^d(B_d)}{m^{d-1}(B_{d-1})} = 2\int_0^1 (1-x^2)^{(d-1)/2}dx.$$

Using a change of variables we obtain

$$\frac{m^d(B_d)}{m^{d-1}(B_{d-1})} = \int_0^1 y^{-1/2}(1-y)^{(d-1)/2}dy = \beta\left(\frac{1}{2}, \frac{d+1}{2}\right),$$

where the *Euler beta function* is

$$\beta(x,y) = 2\int_0^{\pi/2} \cos^{2x-1}(\theta)\sin^{2y-1}(\theta)\ d\theta = \frac{\Gamma(x)\Gamma(y)}{\Gamma(x+y)}.$$

Thus,

$$m^d(B_d) = m^{d-1}(B_{d-1})\frac{\Gamma(\frac{1}{2})\Gamma(\frac{d+1}{2})}{\Gamma(\frac{d}{2}+1)}.$$

In order to finish the proof we need to observe that

$$\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$$

and

$$\Gamma(x+1) = x\Gamma(x);$$

see, e.g., [179].

2.4 General measure theory

In the previous two sections we presented methods of constructing σ-additive set functions on σ-rings of subsets of $\mathbb{R}^d$. A similar approach may be used to extend finitely additive functions on more general spaces X. As we shall later see, σ-additivity is crucial for proving general and good theorems about taking limits under integral signs and switching iterated limits. Thus, as an efficiency move, we now proceed directly to this setting.

Let X be a set and let $\mathcal{A}$ be a *σ-algebra* in $\mathcal{P}(X)$, i.e., $\emptyset \in \mathcal{A}$ and $\mathcal{A}$ is closed with respect to taking complements and countable unions. Then $(X, \mathcal{A})$ is *measurable space*. A *measure*, respectively, *finitely additive measure*, μ on $(X, \mathcal{A})$ is a function $\mu : \mathcal{A} \to \mathbb{R}^+ \cup \{+\infty\}$ such that

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

and

$$\mu\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} \mu(A_n),$$

where $\{A_n : n = 1, \ldots\} \subseteq \mathcal{A}$ is a countable, respectively, finite, disjoint family. A *measure space* $(X, \mathcal{A}, \mu)$ is a measurable space $(X, \mathcal{A})$ and a measure

μ. In this case the elements of $\mathcal{A}$ are *μ-measurable sets* or, more simply, *measurable sets*.

The techniques and results of Section 2.3 generalize to this more abstract setting, and we can state the following theorem, which allows us to construct measures on many spaces.

Theorem 2.4.1. Carathéodory theorem
Let X be a set, let $\mathcal{R} \subseteq \mathcal{P}(X)$ be an algebra, and let $\mu : \mathcal{R} \to \mathbb{R}^+ \cup \{+\infty\}$ be a σ-additive set function. Then μ extends to a measure on the σ-algebra of measurable sets $\mathcal{A}$, and so also to a measure on the σ-algebra generated by $\mathcal{R}$.

We leave the details of the proof of Theorem 2.4.1 to Problem 2.20, where the issue of uniqueness of such extensions is also discussed.

If X is a topological space (see Appendix A.1) and $\mathcal{A}$ contains the Borel sets then a measure μ on $(X, \mathcal{A})$ is a *Borel measure*.

Johann Radon was the first to define a general measure space (in 1913). The idea was certainly in the air because of Vitali's and Lebesgue's work to extend the fundamental theorem of calculus to $\mathbb{R}^d$; in fact, in such a setting it became important to consider set functions; cf. Section 5.6.3. The key result of Radon's paper was the first form of the now crucial Radon–Nikodym theorem, which is a natural generalization of the fundamental theorem of calculus.

Remark. Let X be a metric space (Appendix A.1), which is also a measure space. Two Borel subsets of a measure space X are *congruent* if there exists an isometry of one set onto the other. One of the simplest examples of such isometry and congruence is translation and translation invariance, which is required in defining Lebesgue measure. Banach and Ulam asked whether, in compact metric spaces, one can always define a finitely additive measure for which all congruent Borel sets have the same measure; see [344], Problem 2. The answer is positive if one makes the additional assumption that X is countable; see [124]. On the other hand, it follows from works of Mycielski [352], [353] that, for each compact metric space X, there exists a Borel measure with the property that all open congruent sets have the same measure.

Example 2.4.2. Measure spaces
We now give some examples of measure spaces.

a. The Lebesgue measure space $(\mathbb{R}^d, \mathcal{M}(\mathbb{R}^d), m^d)$.

b. For any set X and any σ-algebra $\mathcal{A} \subseteq \mathcal{P}(X)$, fix x and define

$$\forall\, A \in \mathcal{A}, \quad \delta_x(A) = \begin{cases} 0, & \text{if } x \notin A, \\ 1, & \text{if } x \in A. \end{cases}$$

Then $(X, \mathcal{A}, \delta_x)$ is a measure space. Here δ_x is the *Dirac measure* at x associated with $\mathcal{A}$.

c. $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d), m^d)$ is a measure space.

d. For any set X and any σ-algebra $\mathcal{A} \subseteq \mathcal{P}(X)$, define

$$\forall\, A \in \mathcal{A}, \quad c(A) = \begin{cases} \operatorname{card} A, & \text{if } \operatorname{card} A < \aleph_0, \\ \infty, & \text{if } \operatorname{card} A \geq \aleph_0. \end{cases}$$

Then $(X, \mathcal{A}, c)$ is a measure space and c is called *counting* measure. For $\mathcal{A} = \mathcal{P}(\mathbb{R})$, c is translation-invariant, i.e., c satisfies (2.1) on $\mathcal{P}(\mathbb{R})$; compare this with Example 2.2.16.

Theorem 2.4.3*c,d* are the results we mentioned after Theorem 2.2.6.

Theorem 2.4.3. The measure of unions and intersections
Let $(X, \mathcal{A}, \mu)$ be a measure space.

a. *If $A, B \in \mathcal{A}$, $A \subseteq B$, then $\mu(A) \leq \mu(B)$.*

b. *For each sequence $\{A_n : n = 1, \ldots\} \subseteq \mathcal{A}$,*

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

c. *If $\{A_n : n = 1, \ldots\} \subseteq \mathcal{A}$ satisfies the conditions that $\mu(A_1) < \infty$ and $A_n \subseteq A_{n-1}$ for each $n \geq 2$, then*

$$\mu\left(\bigcap_{n=1}^{\infty} A_n\right) = \lim_{n\to\infty} \mu(A_n).$$

d. *If $\{A_n : n = 1, \ldots\} \subseteq \mathcal{A}$ is a sequence with the property that $A_n \subseteq A_{n+1}$ for each ≥ 1, then*

$$\mu\left(\bigcup_{n=1}^{\infty} A_n\right) = \lim_{n\to\infty} \mu(A_n).$$

Proof. ***a.*** $B = A \cup (B \setminus A)$, and so $\mu(B) = \mu(A) + \mu(B \setminus A)$. Since $\mu(B \setminus A) \geq 0$ we have $\mu(A) + \mu(B \setminus A) \geq \mu(A)$.

b. Set $G_n = A_n \setminus \left(\bigcup_{j=1}^{n-1} A_j\right)$. Thus, $\{G_n : n = 1, \ldots\}$ is a disjoint family and $\bigcup_{n=1}^{\infty} G_n = \bigcup_{n=1}^{\infty} A_n$. Also, $G_n \subseteq A_n$, and so

$$\mu\left(\bigcup_{n=1}^{\infty} A_n\right) = \mu\left(\bigcup_{n=1}^{\infty} G_n\right) = \sum_{n=1}^{\infty} \mu(G_n) \leq \sum_{n=1}^{\infty} \mu(A_n).$$

c. Let $A = \bigcap_{n=1}^{\infty} A_n$, so that A_1 is the disjoint union

$$A_1 = A \cup \left(\bigcup_{n=1}^{\infty} (A_n \setminus A_{n+1})\right);$$

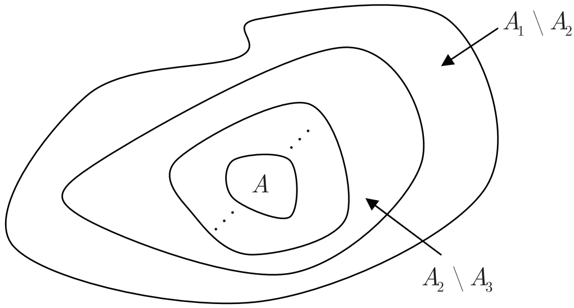

Fig. 2.1. Measure of intersections.

see Figure 2.1. Consequently,

$$\mu(A_1) = \mu(A) + \sum_{n=1}^{\infty} \mu(A_n \setminus A_{n+1}).$$

Also, $A_n = A_{n+1} \cup (A_n \setminus A_{n+1})$ is a disjoint union, from which we conclude that

$$\mu(A_n) - \mu(A_{n+1}) = \mu(A_n \setminus A_{n+1}),$$

since each of the terms $\mu(A_n)$ is finite. By our hypothesis on A_1, $\sum \mu(A_n \setminus A_{n+1})$ converges and $\mu(A) < \infty$; thus

$$\begin{aligned} \mu(A_1) &= \mu(A) + \sum_{n=1}^{\infty} (\mu(A_n) - \mu(A_{n+1})) \\ &= \mu(A) + \lim_{m\to\infty} \sum_{n=1}^{m} (\mu(A_n) - \mu(A_{n+1})) \\ &= \mu(A) + \mu(A_1) - \lim_{m\to\infty} \mu(A_{m+1}). \end{aligned}$$

d. We write

$$\bigcup_{n=1}^{\infty} A_n = A_1 \cup \left(\bigcup_{n=1}^{\infty} (A_{n+1} \setminus A_n) \right),$$

where all sets on the right-hand side of the above equation are mutually disjoint. Thus,

$$\begin{aligned} \mu\left(\bigcup_{n=1}^{\infty} A_n \right) &= \sum_{n=1}^{\infty} \mu(A_{n+1} \setminus A_n) + \mu(A_1) \\ &= \lim_{m\to\infty} \left(\sum_{n=1}^{m} \mu(A_{n+1} \setminus A_n) + \mu(A_1) \right) = \lim_{m\to\infty} \mu(A_m). \quad \square \end{aligned}$$

Theorem 2.4.4. First Borel–Cantelli lemma
Let $(X, \mathcal{A}, \mu)$ be a measure space. If $\{A_n : n = 1, \ldots\} \subseteq \mathcal{A}$ satisfies the condition $\sum_{n=1}^{\infty} \mu(A_n) < \infty$, then

$$\mu\left(\bigcap_{m=1}^{\infty} \bigcup_{n=m}^{\infty} A_n\right) = 0.$$

Proof. From Theorem 2.4.3*b* it follows that for each m the set $\bigcup_{n=m}^{\infty} A_n$ is μ-measurable, and $\mu\left(\bigcup_{n=1}^{\infty} A_n\right) < \infty$. Thus, $\{\bigcup_{n=m}^{\infty} A_n\}$ satisfies the assumptions of Theorem 2.4.3*c* and we may conclude that

$$\mu\left(\bigcap_{m=1}^{\infty} \bigcup_{n=m}^{\infty} A_n\right) = \lim_{m\to\infty} \mu\left(\bigcup_{n=m}^{\infty} A_n\right) \leq \lim_{m\to\infty} \sum_{n=m}^{\infty} \mu(A_n) = 0,$$

since the series is convergent. □

The conclusion of the First Borel–Cantelli lemma may also be understood as follows: *The collection of all those $x \in X$ that belong to infinitely many sets A_n has measure* 0.

Remark. For a given sequence $\mathcal{A} = \{A_n : n = 1, \ldots\} \subseteq [0, 1]$ of sets, Ulam asked for necessary and sufficient conditions in order to define a σ-additive measure on the Borel algebra generated by $\mathcal{A}$, with the property that the measure of the union $\bigcup A_n$ is 1 and single points are of measure 0; see [344], Problem 145.

This question was answered by Banach [20], who introduced the notion of *atoms* associated with the *characteristic function of the sequence* $\mathcal{A} = \{A_n : n = 1, \ldots\}$:

$$\mathbb{1}_{\mathcal{A}}(x) = 2 \sum_{n=1}^{\infty} \frac{1}{3^n} \mathbb{1}_{A_n}(x).$$

Banach's solution is in terms of a condition for the set of values of $\mathbb{1}_{\mathcal{A}}$.

Let $(X, \mathcal{A}, \mu)$ be a measure space. We say that $\{A_n : n = 1, \ldots, N\}$ is a finite collection of *independent sets* if

$$\mu(A_{n_1} \cap \ldots \cap A_{n_k}) = \mu(A_{n_1}) \cdots \mu(A_{n_k}),$$

for all $k \leq N$ and $n_1 < \ldots < n_k \leq N$. An infinite collection of measurable sets is *independent* if each of its finite subcollections is independent.

Theorem 2.4.5. Second Borel–Cantelli lemma
Let $(X, \mathcal{A}, \mu)$ be a measure space such that $\mu(X) = 1$. If $\{A_n : n = 1, \ldots\} \subseteq \mathcal{A}$ is a sequence of independent sets and $\sum_{n=1}^{\infty} \mu(A_n) = \infty$, then

$$\mu\left(\bigcap_{m=1}^{\infty} \bigcup_{n=m}^{\infty} A_n\right) = 1.$$

Proof. The statement of the theorem is equivalent to

$$\mu\left(\bigcup_{m=1}^{\infty}\bigcap_{n=m}^{\infty} A_n^{\sim}\right) = 0.$$

Thus, it is enough to show that $\mu\left(\bigcap_{n=m}^{\infty} A_n^{\sim}\right) = 0$ for all m. Because of the independence and since $1 - x \leq e^{-x}$,

$$\mu\left(\bigcap_{n=m}^{m+k} A_n^{\sim}\right) = \prod_{n=m}^{m+k}(1-\mu(A_n)) \leq \prod_{n=m}^{m+k} e^{-\mu(A_n)};$$

and this last expression converges to 0 as $k \to \infty$ because of our assumption that $\sum_{n=1}^{\infty}\mu(A_n) = \infty$. $\square$

Measure spaces satisfying the property $\mu(X) = 1$ are often called *probability spaces* and μ is called a *probability measure*. The Borel–Cantelli lemmas are important and useful results in probability theory, as is the following result.

Theorem 2.4.6. Kolmogorov zero–one law
Let $(X, \mathcal{A}, \mu)$ be a probability space and let $\{A_n : n = 1, \ldots\} \subseteq \mathcal{A}$ be a sequence of independent sets. Let $\mathcal{A}_m$ denote the σ-algebra generated by $\{A_n : n = m, \ldots\}$. For each $A \in \bigcap_{m=1}^{\infty}\mathcal{A}_m$, either $\mu(A) = 0$ or $\mu(A) = 1$.

For the proof of this result and many other interesting applications of measure theory in probability we refer the interested reader to [61], [266].

Definition 2.4.7. Finite, σ-finite, and complete measure spaces
a. Let $(X, \mathcal{A}, \mu)$ be a measure space. We say that $(X, \mathcal{A}, \mu)$ or X is a *finite measure space* and μ is a *bounded measure* if $\mu(X) < \infty$. We say that X or μ is *σ-finite* if

$$\exists\, \{A_n : n = 1, \ldots\} \subseteq \mathcal{A} \text{ such that } \forall\, n, \quad \mu(A_n) < \infty \quad \text{and} \quad \bigcup_{n=1}^{\infty} A_n = X.$$

Clearly, we can always take $\{A_n : n = 1, \ldots\}$ to be a disjoint family.

b. A measure space $(X, \mathcal{A}, \mu)$ is *complete* if

$$\forall\, A \in \mathcal{A}, \text{ for which } \mu(A) = 0, \text{ and } \forall\, B \subseteq A, \text{ we have } B \in \mathcal{A},$$

and thus $\mu(B) = 0$.

It follows from the definition of Lebesgue measure that $(\mathbb{R}, \mathcal{M}(\mathbb{R}), m)$ is complete. In Example 2.4.14 we shall "construct" a set $L \in \mathcal{M}(\mathbb{R}) \setminus \mathcal{B}(\mathbb{R})$ such that $L \subseteq C$, where C is the ternary Cantor set (Example 1.2.7). Thus, $(\mathbb{R}, \mathcal{B}(\mathbb{R}), m)$ is not a complete measure space.

Note that a complete measure space $(\mathbb{R}, \mathcal{M}(\mathbb{R}), m)$ is formed by "adding" all of the sets A having Lebesgue measure $m(A) = 0$ to $\mathcal{B}(\mathbb{R})$. This phenomenon is general in the following sense.

Theorem 2.4.8. Completeness theorem
Let $(X, \mathcal{A}, \mu)$ be a measure space. There is a measure space $(X, \mathcal{A}_0, \mu_0)$ such that

i. $\mathcal{A} \subseteq \mathcal{A}_0$,
ii. $\mu = \mu_0$ *on* $\mathcal{A}$,
iii. $A \in \mathcal{A}_0 \iff A = B \cup E$, *where* $B \in \mathcal{A}$ *and* $E \subseteq D$, *for some* $D \in \mathcal{A}$ *that satisfies* $\mu(D) = 0$,
iv. if $A \in \mathcal{A}_0$, $\mu_0(A) = 0$, *and* $S \subseteq A$, *then* $S \in \mathcal{A}_0$ *and* $\mu_0(S) = 0$.

Proof. ***a.*** We first show that $\mathcal{A}_0$, defined by *iii*, is a σ-algebra. First note that $\mathcal{A} \subseteq \mathcal{A}_0$ by *iii*, and so we have *i*. Let $A_n = B_n \cup E_n$, where $B_n \in \mathcal{A}$ and $E_n \subseteq D_n$, for some $D_n \in \mathcal{A}$ that satisfies $\mu(D_n) = 0$.

We have $\bigcup A_n = \bigcup (B_n \cup E_n) = (\bigcup B_n) \cup (\bigcup E_n)$ and $\bigcup B_n \in \mathcal{A}$, since $\mathcal{A}$ is a σ-algebra. Clearly, $\bigcup E_n \subseteq \bigcup D_n$, and $\bigcup D_n \in \mathcal{A}$, since $D_n \in \mathcal{A}$; also $\mu(\bigcup D_n) \le \sum \mu(D_n) = 0$ by Theorem 2.4.3*b*. Consequently, $\mathcal{A}_0$ is closed under countable unions. Obviously, $\emptyset \in \mathcal{A}_0$.

Finally, we must show that $A^\sim \in \mathcal{A}_0$ if $A \in \mathcal{A}_0$. Note that $A^\sim = B^\sim \cap E^\sim = (B^\sim \cap D^\sim) \cup (D \cap E^\sim \cap B^\sim)$, $B^\sim \cap D^\sim \in \mathcal{A}$, $D \cap E^\sim \cap B^\sim \subseteq D \in \mathcal{A}$, and $\mu(D) = 0$. Thus, $A^\sim \in A_0$ by the definition of $\mathcal{A}_0$. Therefore, $\mathcal{A}_0$ is a σ-algebra.

b. If $A = B \cup E$, with notation as in *iii*, we define $\mu_0(A) = \mu(B)$. In particular, we have *ii*. We must check that μ_0 is well defined. Letting $A = B_1 \cup E_1 = B_2 \cup E_2$, it is sufficient to prove that $\mu(B_1) = \mu(B_2)$. Clearly, $B_1 \subseteq B_2 \cup E_2 \subseteq B_2 \cup D_2$, and so $\mu(B_1) \le \mu(B_2) + \mu(D) = \mu(B_2)$. Similarly, we compute $\mu(B_2) \le \mu(B_1)$, and hence μ_0 is well defined.

c. Next we show that μ_0 is a measure. From part *b*, $\mu_0 \ge 0$. Now consider $\bigcup A_n$, where $A_n = B_n \cup E_n$ is decomposed as in *iii* and $\{A_n : n = 1, \ldots\}$ is a disjoint collection. Then

$$\mu_0\left(\bigcup_{n=1}^\infty A_n\right) = \mu_0\left(\left(\bigcup_{n=1}^\infty B_n\right) \cup \left(\bigcup_{n=1}^\infty E_n\right)\right) = \mu\left(\bigcup_{n=1}^\infty B_n\right)$$
$$= \sum_{n=1}^\infty \mu(B_n) = \sum_{n=1}^\infty \mu_0(A_n).$$

d. Finally, we verify *iv*. Let $\mu_0(A) = 0$ for a given $A \in \mathcal{A}$ and take $S \subseteq A$. Assume that $A = B \cup E$ is decomposed as in *iii*. Then $\mu(B) = 0$, since $\mu_0(A) = 0$. We write $A = \emptyset \cup (B \cup E)$, noting that

$$B \cup E \subseteq B \cup D \in \mathcal{A} \quad \text{and} \quad \mu(B \cup D) \le \mu(B) + \mu(D) = 0.$$

Then, $S = \emptyset \cup S$ and $S \subseteq A = B \cup E \subseteq B \cup D \in \mathcal{A}$, where $\mu(B \cup D) = 0$. Thus, $S \in \mathcal{A}_0$ and $\mu_0(S) = \mu(\emptyset) = 0$. $\square$

We call $(X, \mathcal{A}_0, \mu_0)$ the *complete measure space corresponding to* $(X, \mathcal{A}, \mu)$.

Theorem 2.4.9. Induced measure spaces
a. *Let* $\{(X_\alpha, \mathcal{A}_\alpha, \mu_\alpha)\}$ *be a collection of measure spaces. Define the triple* $(X, \mathcal{A}, \mu)$ *as* $X = \bigcup X_\alpha$,

$$\mathcal{A} = \{A \subseteq X : \forall\, \alpha,\ A \cap X_\alpha \in \mathcal{A}_\alpha\} \quad \text{and} \quad \forall\, A \in \mathcal{A},\ \mu(A) = \sum \mu_\alpha(A \cap X_\alpha).$$

Then $(X, \mathcal{A}, \mu)$ *is a measure space; and it is* σ*-finite if and only if all but a countable number of the* μ_α *are zero and the remainder are* σ*-finite.*

b. *Let* $(X, \mathcal{A}, \mu)$ *be a measure space and let* $Y \in \mathcal{A}$. *Set*

$$\mathcal{A}_Y = \{A \in \mathcal{A} : A \subseteq Y\}.$$

Define $\mu_Y(A) = \mu(A)$ *for* $A \in \mathcal{A}_Y$. *Then* $(Y, \mathcal{A}_Y, \mu_Y)$ *is a measure space.*

The proof of Theorem 2.4.9 is left as an exercise (Problem 2.30). In light of Theorem 2.4.9*b* we shall make the following notational convention. If we wish to consider Lebesgue measure m^d restricted to the Lebesgue measurable sets contained in a fixed set $X \subseteq \mathbb{R}^d$, we shall write $(X, \mathcal{M}(X), m^d)$ for the corresponding measure space. A similar remark applies to a measure μ defined on the Borel sets of a topological space X, in which case we shall write $(X, \mathcal{B}(X), \mu)$.

Let $\mathbb{R}^* = \mathbb{R} \cup \{\pm\infty\}$. By convention, $\infty + \infty = \infty$, $-\infty - \infty = -\infty$, $\infty \cdot (\pm\infty) = \pm\infty$, $-\infty \cdot (\pm\infty) = \mp\infty$, and $\mp\infty \pm \infty$ is undefined.

Proposition 2.4.10. *Let* $(X, \mathcal{A}, \mu)$ *be a measure space and let* $f : X \to \mathbb{R}^*$ *be a function. The following are equivalent:*

a. $\forall\, \alpha \in \mathbb{R},\ \{x : f(x) > \alpha\} \in \mathcal{A}$,
b. $\forall\, \alpha \in \mathbb{R},\ \{x : f(x) \geq \alpha\} \in \mathcal{A}$,
c. $\forall\, \alpha \in \mathbb{R},\ \{x : f(x) < \alpha\} \in \mathcal{A}$,
d. $\forall\, \alpha \in \mathbb{R},\ \{x : f(x) \leq \alpha\} \in \mathcal{A}$,
e. $\forall\, U \subseteq \mathbb{R}$, U *open*, $f^{-1}(U) \in \mathcal{A}$, *and*

$$f^{-1}(\pm\infty) \in \mathcal{A}.$$

Proof. $a \Longrightarrow d$. Clearly, $\{x : f(x) \leq \alpha\} = X \setminus \{x : f(x) > \alpha\}$, so that since $\{x : f(x) > \alpha\} \in \mathcal{A}$ and $\mathcal{A}$ is an algebra, we have $\{x : f(x) \leq \alpha\} \in \mathcal{A}$.

Similarly, $d \Longrightarrow a$ and $b \iff c$.

$a \Longrightarrow b$. We have $\{x : f(x) \geq \alpha\} = \bigcap \{x : f(x) > \alpha - (1/n)\}$, so that we have the required implication, since $\mathcal{A}$ is a σ-algebra.

Similarly, $b \Longrightarrow a$, and parts a through d are equivalent.

Assume parts a–d and let $U = \bigcup I_j$, where I_j is an open interval. Then $f^{-1}(U) = \bigcup f^{-1}(I_j)$ and $f^{-1}(I_j) \in \mathcal{A}$ because of a–d. Thus, $f^{-1}(U) \in \mathcal{A}$, since $\mathcal{A}$, is a σ-algebra. Conversely, suppose we assume part e and take $U = (\alpha, \infty]$ or $U = (\alpha, \infty)$; then $f^{-1}(U) \in \mathcal{A}$, and we have part a. $\square$

Definition 2.4.11. Measurable functions
a. An *extended real-valued function* $f : X \to \mathbb{R}^*$, defined on a measure space $(X, \mathcal{A}, \mu)$, is *measurable*, more precisely, $\mathcal{A}$*-measurable* or μ*-measurable*, if any of the conditions *a–e* in Proposition 2.4.10 holds, cf., Problem 2.36. When we introduce the notion of "almost everywhere" (*a.e.*) in Definition 2.4.15, we shall assume that measurable functions are finite μ-*a.e.*

b. A complex-valued function $f : X \to \mathbb{C}$ is *measurable* if its real and imaginary parts, $\mathrm{Re}\,(f)$ and $\mathrm{Im}\,(f)$, are measurable.

c. If $\mathcal{A}$ is $\mathcal{M}(\mathbb{R}^d)$ or $\mathcal{B}(\mathbb{R}^d)$, the corresponding measurable function is *Lebesgue* or *Borel measurable*, respectively.

d. We extend parts *a–c* of this definition in the following way. Let $(X, \mathcal{A})$ and $(Y, \mathcal{E})$ be measurable spaces. A function $f : X \to Y$ is *measurable* if

$$\forall\, E \in \mathcal{E}, \quad f^{-1}(E) \in \mathcal{A}.$$

Measurable mappings are a staple in several important subjects such as ergodic theory and the topics of Chapter 9. Perhaps, more surprisingly, they are essential in the underlying geometry of wavelet sets [48].

Note that when X is a topological space (see Appendix A.1), a real-valued function f on X is *continuous* if $f^{-1}(U)$ is open for all open sets $U \subseteq \mathbb{R}$, a fact that we proved in Section 1.3.2 using the metric definition of continuity. Whenever X is a topological space and $(X, \mathcal{A}, \mu)$ is a measure space we shall assume that $\mathcal{B}(X) \subseteq \mathcal{A}$; hence continuous functions $f : (X, \mathcal{A}, \mu) \to \mathbb{R}$ are $\mathcal{A}$-measurable. Of course, every continuous function $f : X \to \mathbb{R}$ is $\mathcal{B}(X)$-measurable.

Also, $A \in \mathcal{A}$ if and only if $\mathbb{1}_A$ is a measurable function.

Proposition 2.4.12. *Let f and g be real-valued measurable functions on a measure space $(X, \mathcal{A}, \mu)$. Then, $f \pm g$, fg, $f + c$, and cf are measurable, where $c \in \mathbb{R}$.*

Proof. We first outline the proof that $f + g$ is measurable. Let $S = \{x : f(x) + g(x) < \alpha\}$. Thus, if $x \in S$ there is $r \in \mathbb{Q}$ such that $f(x) < r < \alpha - g(x)$. Hence,

$$S = \bigcup_{r \in \mathbb{Q}} \left(\{x : f(x) < r\} \cap \{x : g(x) < \alpha - r\}\right).$$

Since f and g are measurable and the union is countable, $f + g$ is measurable.

To show that fg is measurable note that we need only prove that f^2 is measurable, since $fg = (1/2)[(f+g)^2 - f^2 - g^2]$. For $\alpha \geq 0$, $\{x : f^2(x) > \alpha\} = \{x : f(x) > \sqrt{\alpha}\} \cup \{x : f(x) < -\sqrt{\alpha}\} \in \mathcal{A}$ implies that f^2 is measurable, and for $\alpha < 0$, $\{x : f^2(x) > \alpha\}$ is the domain of f. $\square$

Example 2.4.13. Nonmeasurable subsets of perfect symmetric sets
We shall prove that there is a perfect symmetric set $E \subseteq [0, 1]$ having positive Lebesgue measure and with a subset $N \notin \mathcal{M}(\mathbb{R})$. The proof is by

contradiction. From Example 2.2.16, choose $S \notin \mathcal{M}(\mathbb{R})$, where $S \subseteq [0,1]$; and let $\{E_n : n = 1, \ldots\}$ be a sequence of perfect symmetric nowhere-dense sets as defined in Example 1.2.8 such that

$$\forall\, n,\;\; 1 \geq m(E_n) \geq 1 - \frac{1}{n}.$$

If $S \cap E_n \in \mathcal{M}(\mathbb{R})$ for each n then

$$\bigcup_{n=1}^{\infty} (S \cap E_n) \in \mathcal{M}(\mathbb{R}).$$

Let $[0,1] \cap (\bigcup E_n)^{\sim} = F \subseteq [0,1]$. Then $m(F) = 0$ and

$$S = S \cap \left(F \cup \left(\bigcup_{n=1}^{\infty} E_n \right) \right).$$

Now $S \cap F \subseteq F$ implies $m(S \cap F) = 0$, since $(\mathbb{R}, \mathcal{M}(\mathbb{R}), m)$ is complete; and so $(S \cap F) \cup (\bigcup (S \cap E_n)) \in \mathcal{M}(\mathbb{R})$, i.e., $S \in \mathcal{M}(\mathbb{R})$, the desired contradiction. Thus, $S \cap E_n \notin \mathcal{M}(\mathbb{R})$ for some $n \in \mathbb{N}$; and we let $E = E_n$ and $N = S \cap E_n$.

Example 2.4.14. $(\mathbb{R}, \mathcal{B}(\mathbb{R}), m)$ is not complete
We shall find a set $L \in \mathcal{M}(\mathbb{R}) \setminus \mathcal{B}(\mathbb{R})$ such that $L \subseteq C$, where C is the ternary Cantor set. Since $C \in \mathcal{B}(\mathbb{R})$ and $m(C) = 0$, we can conclude that $(\mathbb{R}, \mathcal{B}(\mathbb{R}), m)$ is not a complete measure space.

i. Take the set E of Example 2.4.13 and put the contiguous intervals of C in one-to-one correspondence with those of E by the "almost linear" mapping g depicted in Figure 2.2. Thus, g is defined on $[0,1] \setminus E$, it is increasing, and it maps $[0,1] \setminus E$ onto $[0,1] \setminus C$. By the monotonicity, $g(x\pm)$ exist for all $x \in [0,1]$; and since C is nowhere dense, g can be extended to a continuous increasing surjection $g : [0,1] \to [0,1]$. Thus, g is a Borel measurable function. Take $N \subseteq E$ as in Example 2.4.13 and define $L = g(N)$. Hence $L \subseteq C$, and, since $m(C) = 0$, we conclude that $L \in \mathcal{M}(\mathbb{R})$ and $m(L) = 0$. Also $g^{-1}(L) = N$ because g is injective.

ii. Finally, observe that $L \notin \mathcal{B}(\mathbb{R})$, for, by a routine property of Borel measurable functions f, $f^{-1}(\mathcal{B}(\mathbb{R})) \subseteq \mathcal{B}(\mathbb{R})$ (Problem 2.35); cf. Problem 2.36; and this would imply that N is a Borel set if in fact L were a Borel set.

Definition 2.4.15. Almost everywhere
a. Let $(X, \mathcal{A}, \mu)$ be a measure space and let $S(x)$ be a statement about a point $x \in X$. For example, for a given function $f : X \to \mathbb{R}$, $S(x)$ could be the statement $f(x) > 0$. A statement $S(x)$ is valid *almost everywhere* if there is a set $N \in \mathcal{A}$, for which $\mu(N) = 0$, such that

$$\forall\, x \in X \setminus N,\;\; S(x) \;\text{ is true.}$$

In this case we write S μ*-a.e.*

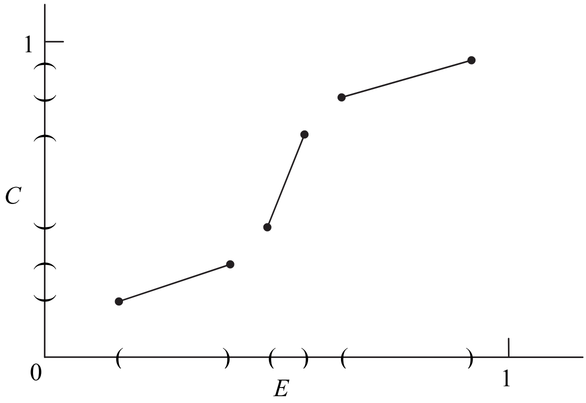

Fig. 2.2. A general Cantor function.

b. For two $\mathbb{C}$- or $\mathbb{R}^*$-valued measurable functions f and g on X, $f = g$ μ-*a.e.* signifies that $\mu(\{x : f(x) \neq g(x)\}) = 0$. Suppose we use the notation $f \sim g$ to mean $f = g$ μ-*a.e.* It is elementary to verify that $\sim$ is a well-defined *equivalence relation* on the set of $\mathbb{C}$- or $\mathbb{R}^*$-valued measurable functions on X.

c. We shall *always* assume that our measurable functions are finite μ-*a.e.*

Proposition 2.4.16. *Let $(X, \mathcal{A}, \mu)$ be a complete measure space. If f is measurable and $f = g$ μ-a.e., then g is measurable.*

Proof. Let $E = \{x : f(x) \neq g(x)\}$; and so $\mu(E) = 0$. Observe that

$$\{x : g(x) > \alpha\} = (\{x : f(x) > \alpha\} \cup \{x \in E : g(x) > \alpha\}) \setminus \{x \in E : g(x) \leq \alpha\}.$$

Since $(X, \mathcal{A}, \mu)$ is complete and $\mu(E) = 0$,

$$\mu(\{x \in E : g(x) > \alpha\}) = 0 = \mu(\{x \in E : g(x) \leq \alpha\}).$$

Thus $\{x : g(x) > \alpha\} \in \mathcal{A}$, because f is measurable. $\square$

Example 2.4.17. Noncompleteness and nonmeasurable functions
For noncomplete measure spaces Proposition 2.4.16 does not hold. Let $L \in \mathcal{M}(\mathbb{R}) \setminus \mathcal{B}(\mathbb{R})$ have Lebesgue measure $m(L) = 0$, and let $G \supseteq L$ be a $\mathcal{G}_\delta$ set for which $m(G) = m(L)$. Define $f = \mathbb{1}_{\mathbb{R}}$ and

$$g(x) = \begin{cases} 1, & \text{if } x \in \mathbb{R} \setminus G, \\ \frac{1}{2}, & \text{if } x \in G \setminus L, \\ 0, & \text{if } x \in L. \end{cases}$$

Then, since $G \in \mathcal{B}(\mathbb{R})$, we have $f = g$ m-*a.e.* in the measure space $(\mathbb{R}, \mathcal{B}(\mathbb{R}), m)$; but $g^{-1}(0) \notin \mathcal{B}(\mathbb{R})$ and so g is not measurable on $(\mathbb{R}, \mathcal{B}(\mathbb{R}), m)$.

Remark. Recall that if $A \in \mathcal{M}(\mathbb{R})$ then $A = B \cup E$, where $B \in \mathcal{B}(\mathbb{R})$ and $m(E) = 0$. On the other hand, there are measure spaces $(\mathbb{R}, \mathcal{A}, \mu)$, for which $\mathcal{B}(\mathbb{R}) \subseteq \mathcal{A}$, where this decomposition is not true; see Problem 2.28.

Remark. Wacław Sierpiński [435] has given an example of a set $A \notin \mathcal{M}(\mathbb{R}^2)$ that has at most two points on each straight line. Using this set A one can find a function $f : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ whose projections are Borel measurable functions but that itself is not Lebesgue measurable. There are positive results in the opposite direction, dating from Lebesgue in 1905, which are important for the finer study of the Fubini theorem. The basic statement of the Fubini theorem is given in Section 3.7.

We shall end this section with the following observation. We have been studying extensions of certain set functions into σ-additive set functions on σ-algebras (and such extensions are called measures). However, if we wish to construct a measure, without any prerequisites, there is another way to do so. This method can be easily extracted from our construction of Lebesgue measure on $\mathbb{R}$. To this end, recall that we defined outer measures in terms of set functions in Sections 2.2 and 2.3. Now we proceed axiomatically.

Definition 2.4.18. Outer measure
An *outer measure* is a nonnegative set function $\mu^* : \mathcal{P}(X) \to \mathbb{R}^+ \cup \{\infty\}$ with the following properties:

i. $A \subseteq B \Longrightarrow \mu^*(A) \leq \mu^*(B)$,
ii. $\mu^*(\emptyset) = 0$,
iii. $\mu^*\left(\bigcup_{n=1}^{\infty} A_n\right) \leq \sum_{n=1}^{\infty} \mu^*(A_n)$.

Theorem 2.4.19. Measures in terms of outer measures
Let $\mu^* : \mathcal{P}(X) \to \mathbb{R}^+ \cup \{\infty\}$ be an outer measure. There exist a σ-algebra $\mathcal{A} \subseteq \mathcal{P}(X)$ and a nonnegative σ-additive set function μ on $\mathcal{A}$ that is the restriction of μ^* to $\mathcal{A}$.

Proof. We use Carathéodory's approach, see [87], one more time, and define $\mathcal{A}$ to be all of the sets $A \subseteq X$ with the following property:

$$\forall\, E \subseteq X, \quad \mu^*(E) = \mu^*(E \cap A) + \mu^*(E \cap A^{\sim}).$$

We also define μ to be μ^* restricted to $\mathcal{A}$. We observe that in order to complete the proof we need the analogues of Theorem 2.2.5a and Theorem 2.2.6b. Thus, we are done because the only properties of μ^* that were used in the proofs of these two results are the analogues of the properties *i*, *ii*, *iii* in Definition 2.4.18 and the nonnegativity of μ^*. $\square$

Finally, the process of obtaining μ^* from μ and vice versa can be summarized in the following way.

Let $(X, \mathcal{A}, \mu)$ be a measure space. Following the constructions in Sections 2.2 and 2.3, define the set function $\mu^* : \mathcal{P}(X) \to \mathbb{R}^*$ to be

$$\forall\, E \in \mathcal{P}(X), \quad \mu^*(E) = \inf\left\{\sum_{j=1}^{\infty} \mu(A_j)\right\}, \tag{2.19}$$

where the infimum is taken over all collections $\{A_j : j = 1, \ldots\} \subseteq \mathcal{A}$ that cover E, see Problem 2.20. It is not difficult to verify that μ^* is an outer measure. Moreover, the restriction of μ^* to the σ-algebra $\mathcal{A}$ coincides with the measure μ. We say that μ^* is the *outer measure associated with* μ.

If, on the other hand, we are given an outer measure μ^*, we can use Theorem 2.4.19 to generate a σ-algebra $\mathcal{A}$ and a measure μ on $\mathcal{A}$ that is a restriction of μ^*. The outer measure μ^o associated with μ does not necessarily coincide with μ^* on $\mathcal{P}(X)$. However, we have

$$\forall\, A \subseteq X, \quad \mu^*(A) \leq \mu^o(A).$$

2.5 Approximation theorems for measurable functions

Recall that our measurable functions are finite *a.e.*

Proposition 2.5.1. *Let $\{f_n : n = 1, \ldots\}$ be a sequence of $\mathbb{R}^*$-valued measurable functions on the measure space $(X, \mathcal{A}, \mu)$. The operations in the following assertions are pointwise for each $x \in X$.*

a. *For each n, $g = \sup\,\{f_1, \ldots, f_n\}$ and $h = \inf\{f_1, \ldots, f_n\}$ are measurable functions.*

b. *$g = \sup_{n\in\mathbb{N}}\,\{f_n\}$ and $h = \inf_{n\in\mathbb{N}}\,\{f_n\}$ are measurable functions if they are finite μ-a.e.*

c. *$g = \overline{\lim}\ f_n$ and $h = \underline{\lim}\ f_n$ are measurable functions if they are finite μ-a.e.*

Proof. We prove one case for each part.

a. $\{x : g(x) > \alpha\} = \bigcup_{j=1}^{n}\{x : f_j(x) > \alpha\} \in \mathcal{A}$, and so g is measurable.

b. A routine argument shows that

$$\{x : g(x) > \alpha\} = \bigcup_{j=1}^{\infty}\{x : f_j(x) > \alpha\} \in \mathcal{A}.$$

c. $\overline{\lim}\ f_n = \inf_n \sup_{k\geq n}\ f_k$, so that g is measurable because of part *b*.

□

The following result is a consequence of Proposition 2.5.1*c*.

Proposition 2.5.2. *Let $\{f_n : n = 1, \ldots\}$ be a sequence of $\mathbb{C}$- or $\mathbb{R}^*$-valued measurable functions on the measure space $(X, \mathcal{A}, \mu)$, and assume that $\lim f_n(x) = f(x)$ exists pointwise for each $x \in X$. Then f is a measurable function.*

Example 2.5.3. Nonmeasurable *a.e.* limits
Naturally we would like to have Proposition 2.5.2 be true with the weaker hypothesis that $f_n \to f$ μ-*a.e.* Unfortunately, such is not the case generally unless $(X, \mathcal{A}, \mu)$ is complete, e.g., Theorem 2.5.4. As a counterexample let $(X, \mathcal{A}, \mu) = (\mathbb{R}, \mathcal{B}(\mathbb{R}), m)$, set $f_n = 0$ on $\mathbb{R}$, and set $f = \mathbb{1}_L$, where $L \in \mathcal{M}(\mathbb{R}) \setminus \mathcal{B}(\mathbb{R})$ and $m(L) = 0$. Then f is not Borel measurable, whereas $f_n \to f$ m-*a.e.* in $(\mathbb{R}, \mathcal{B}(\mathbb{R}), m)$.

The proof of the following theorem is left as an exercise; see Problem 2.34.

Theorem 2.5.4. Measurable functions as limits *a.e.*
Let $(X, \mathcal{A}, \mu)$ be a complete measure space and let $\{f_n : n = 1, \ldots\}$ be a sequence of $\mathbb{C}$- or $\mathbb{R}^*$-valued measurable functions each defined μ-a.e. on X. If f is defined μ-a.e. and $\lim f_n = f$ pointwise μ-a.e., then f is measurable.

For a given measure space $(X, \mathcal{A}, \mu)$, a function $f : X \to \mathbb{R}$ is *simple* if it can be written in the form

$$f = \sum_{j=1}^{n} a_j \mathbb{1}_{A_j}, \quad \text{where } a_j \in \mathbb{R} \text{ and } A_j \in \mathcal{A}.$$

A function $f : X \to \mathbb{C}$ is *simple* if the analogous definition holds for $a_j \in \mathbb{C}$. Clearly, simple functions are measurable. We can approximate measurable functions by simple functions in the following way.

Theorem 2.5.5. Measurable functions as limits of simple functions
Let f be a $\mathbb{C}$- or $\mathbb{R}^*$-valued measurable function on $(X, \mathcal{A}, \mu)$. There is a sequence $\{f_n : n = 1, \ldots\}$ of simple functions such that

i. $\forall\, j$ *and* $\forall\, x \in X$, $|f_j(x)| \le |f_{j+1}(x)|$

and

ii. $\forall\, x \in X$, $\lim_{j \to \infty} f_j(x) = f(x)$.

If there exists $K > 0$ such that for all $x \in X$, $|f(x)| \le K$, i.e., f is a bounded function on X, then the convergence in part ii is uniform.

Proof. ***a.*** Assume $f \ge 0$. For each $n = 1, \ldots$ and $1 \le k \le n2^n$ define the measurable sets,

$$A_{n,k} = \left\{ x \in X : \frac{k-1}{2^n} \le f(x) < \frac{k}{2^n} \right\}$$

and

$$B_n = \{x \in X : f(x) \ge n\}.$$

Set

$$f_n(x) = \sum_{k=1}^{n2^n} \frac{k-1}{2^n} \mathbb{1}_{A_{n,k}}(x) + n \mathbb{1}_{B_n}(x).$$

Clearly, f_n is a simple function for which $0 \le f_n \le n$. We also have

$$\forall\, x \in X \text{ and } \forall\, j = 1, \ldots, \quad 0 \le f_j(x) \le f_{j+1}(x) \le f(x).$$

These inequalities are a consequence of the dyadic partition of each $A_{n,k}$:

$$A_{n,k} = A_{n+1,2k-1} \cup A_{n+1,2k}.$$

Thus, if $x \in A_{n,k}$ then $x \in A_{n+1,2k-1}$ or $x \in A_{n+1,2k}$. In the former case we have $f_n(x) = f_{n+1}(x)$, and, in the latter, we have $f_{n+1}(x) > f_n(x)$.

If $f(x) = \infty$ then $x \in B_n$ for each $n = 1, \ldots$. Therefore, $f_n(x) = n$, for each n, and, so $f_n(x) \to \infty$. If, on the other hand, $f(x)$ is finite, then there is m such that, for all $n \ge m$, $x \in B_n^{\sim}$, i.e., $f(x) < n$; and thus $|f(x) - f_n(x)| < 1/2^n$ for all such n. Consequently, part *ii* is proved.

If $0 \le f \le K$ on X, then

$$\forall\, n \ge K, \quad \sup_{x \in X} |f(x) - f_n(x)| < \frac{1}{2^n},$$

and this yields the desired uniform convergence.

b If f is $\mathbb{R}^*$-valued, we define the nonnegative functions $f^+(x) = \max\{f(x), 0\}$ and $f^-(x) = \max\{(-f)(x), 0\}$, so that $f = f^+ - f^-$. (The same procedure is used in Definition 3.2.3.) Similarly, if f is $\mathbb{C}$-valued we may write $f = (f_1 - f_2) + i(f_3 - f_4)$, where each $f_j \ge 0$. We can then apply the technique of part *a* to obtain the result. □

Remark. The idea of the above proof is to make a finer and finer grid of $X \times [0, \infty]$ as $n \to \infty$, and then to draw the appropriate simple function on the grid. The use of dyadic partitions (or similar ideas) is more than a technical convenience; see Definition 8.6.11 and Theorem 8.6.14. Such partitions arise in any number of places. We mention three. First, the primordial constructions of wavelet theory in the 1980s, going back to Alfréd Haar's thesis (1909–1910), use such partitions, e.g., [118], [348]. Georg Zimmermann's English translation of Haar's thesis appears in [228]. Second, the deep Littlewood–Paley theory from the 1930s makes use of dyadic partitions in a fundamental way; see [183]. Third, there is a dyadic counterpart to Nikolai N. Luzin's problem. In Section 3.8.3 and 4.7.5, we shall discuss Luzin's problem, which deals with the convergence *a.e.* of Fourier series. In 1924, Andrei N. Kolmogorov [289] solved a dyadic version of this problem, and the dyadic partitions inherent to his work inspired constructions of wavelet sets 75 years later [48].

Theorem 2.5.5 can be compared with the characterization of regulated functions in Problem 3.21. Also the following corollary is an elementary consequence of Theorem 2.5.5, see Problem 2.40.

Corollary 2.5.6. *Let $(X, \mathcal{A}, \mu)$ be a measure space and let $f : X \to \mathbb{R}^*$ be a function that is finite μ-a.e. or let f be $\mathbb{C}$-valued μ-a.e.*

a. *Then f is measurable if and only if f is an everywhere pointwise limit of simple functions increasing in absolute value to $|f|$.*

b. *Assume that f is a bounded function on X. Then f is measurable if and only if f is the uniform limit of simple functions increasing in absolute value to $|f|$.*

In 1911, Dimitri F. Egorov proved the following result for intervals.

Theorem 2.5.7. Egorov theorem
Let $(X, \mathcal{A}, \mu)$ be a finite measure space and let $\{f_n : n = 1, \ldots\}$ be a sequence of $\mathbb{C}$- or $\mathbb{R}^*$-valued measurable functions each defined μ-a.e. on X. Assume that f is defined and finite μ-a.e. on X and that $f_n \to f$ μ-a.e. Then

$$\forall\, \delta > 0,\ \exists\, A \in \mathcal{A} \text{ such that } \mu(A) < \delta \text{ and } \lim_{n\to\infty} f_n = f \text{ uniformly on } A^{\sim}.$$

Proof. Using the convenience (tradition) of tradition (convenience) we give the proof in the following two parts.

i. Given $\varepsilon > 0$ and $\delta > 0$, we shall find a set $A_{\varepsilon,\delta} \in \mathcal{A}$ and an integer N such that $\mu(A_{\varepsilon,\delta}) < \delta$ and

$$\forall\, x \notin A_{\varepsilon,\delta},\ \forall\, n \geq N, \quad |f_n(x) - f(x)| < \varepsilon.$$

Without loss of generality assume $f_n \to f$ everywhere and set

$$A_m = \{x : \exists\, n \geq m \text{ such that } |f_n(x) - f(x)| \geq \varepsilon\}.$$

Clearly $A_m \supseteq A_{m+1}$, and for $x \in X$ there is an m for which $x \notin A_m$. Therefore, $\bigcap A_m = \emptyset$, and, since $\mu(A_j) < \infty$, we have $\lim \mu(A_m) = 0$. Consequently, choose $A_{\varepsilon,\delta} = A_N$, where N is chosen so that $\mu(A_m) < \delta$ for all $m \geq N$.

ii. Given $\delta > 0$ and m, we use part i to find $B_m \in \mathcal{A}$ and N_m such that $\mu(B_m) < \delta/2^m$ and

$$\forall\, x \notin B_m, \forall\, n \geq N_m, \quad |f_n(x) - f(x)| < \frac{1}{m}.$$

Set $A = \bigcup B_m$. □

Example 2.5.8. Finite measure hypothesis for Egorov theorem
The hypothesis that $\mu(X) < \infty$ is necessary in Theorem 2.5.7. Take $f_n = \mathbb{1}_{[n,n+1]}$ on $(\mathbb{R}, \mathcal{M}(\mathbb{R}), m)$. Clearly, $f_n \to f = 0$ pointwise, whereas, if we take $A \in \mathcal{M}(\mathbb{R})$ for which $m(A) < 1$, then

$$\forall\, n\ , \exists\, x_n \in [n, n+1] \cap A^{\sim} \text{ such that } |f(x_n) - f_n(x_n)| = 1.$$

Definition 2.5.9. $\boldsymbol{L^\infty_\mu(X)}$
a. Let $(X, \mathcal{A}, \mu)$ be a measure space and let $f : X \to \mathbb{C}$ be measurable. Define

$$\|f\|_\infty = \inf\,\{M : \mu(\{x : |f(x)| > M\}) = 0\}.$$

Notationally, we also write

$$\|f\|_\infty = \operatorname*{ess\,sup}_{x \in X}\ |f(x)|.$$

b. $\mathcal{L}^\infty_\mu(X)$ is the set of all measurable functions $f : X \to \mathbb{C}$ such that $\|f\|_\infty < \infty$.

c. Let $L^\infty_\mu(X)$ be the set of equivalence classes defined by the equivalence relation $\sim$ on $\mathcal{L}^\infty_\mu(X)$.

Recall from Definition 2.4.15 that $f \sim g$ signifies that $\mu(\{x : f(x) \neq g(x)\}) = 0$. We shall usually think of $F \in L^\infty_\mu(X)$ in terms of any one of its representatives $f \in \mathcal{L}^\infty_\mu(X)$. In this case, we write $f \in L^\infty_\mu(X)$. No confusion arises, either technically or conceptually, because of the following Remark, which also includes another calculation about $L^\infty_\mu(X)$.

Remark. ***a.*** Note that if $f, g \in \mathcal{L}^\infty_\mu(X)$ then

$$f \sim g \quad\Longrightarrow\quad \|f\|_\infty = \|g\|_\infty < \infty.$$

In fact, if $\|f\|_\infty < \|g\|_\infty$, then we can choose $\|f\|_\infty \leq M < \|g\|_\infty$. Thus,

$$\mu(\{x : |f(x)| > M\}) = 0 \ \text{ and } \ \mu(\{x : |g(x)| > M\}) > 0.$$

This contradicts the hypothesis that $f \sim g$.

b. We shall verify that

$$f \in L^\infty_\mu(X) \quad\Longrightarrow\quad |f(x)| \leq \|f\|_\infty \ \ \mu\text{-}a.e. \tag{2.20}$$

First note that if $f \in \mathcal{L}^\infty_\mu(X)$, then

$$\forall\, K > \|f\|_\infty, \quad \mu(\{x : |f(x)| > K\}) = 0 \tag{2.21}$$

by definition of $\|f\|_\infty < \infty$. Next, let $\{K_n\} \subseteq \mathbb{R}^+$ decrease strictly to $\|f\|_\infty$. Then we have

$$\{x : |f(x)| > \|f\|_\infty\} = \bigcup_{n=1}^{\infty} \{x : |f(x)| > K_n\}. \tag{2.22}$$

Each term of the union on the right side of (2.22) has measure 0 by (2.21); and so $\mu(\{x : |f(x)| > \|f\|_\infty\}) = 0$. If $x \notin \{x : |f(x)| > \|f\|_\infty\}$ then $|f(x)| \leq \|f\|_\infty$, and (2.20) is verified.

c. It is elementary to prove that $\|\ldots\|_\infty$ is a norm and that $L^\infty_\mu(X)$ is a Banach space, i.e., a complete normed vector space. For these facts and more on $L^\infty_\mu(X)$, see Section 5.5.

It is elementary to prove the following result.

Proposition 2.5.10. *Let $\{f, f_n : n = 1, \ldots\} \subseteq L^\infty_\mu(X)$ be a sequence of measurable functions on the σ-finite measure space $(X, \mathcal{A}, \mu)$. Then* $\lim_{n\to\infty} \|f_n - f\|_\infty = 0$ *if and only if there is a set $A \in \mathcal{A}$ such that $\mu(A) = 0$ and $f_n \to f$ uniformly on $A^\sim$.*

After Definition 2.4.11, we noted that when X is both a topological space and a measure space then continuous functions $f : X \to \mathbb{R}$ are also measurable. We close this section by looking more closely at the relation between continuous and measurable functions, especially in light of Theorems 2.5.4. 2.5.5, and 2.5.7.

Recall that $\mathbb{1}_{\mathbb{R}\setminus\mathbb{Q}}$ is not the pointwise limit of continuous functions, e.g., Problem 1.15, although, from Theorem 2.5.5, every measurable function is the pointwise limit of simple functions. On the other hand, we have the following example due to Johann P. G. Lejeune Dirichlet.

Example 2.5.11. Dirichlet example
Set

$$g_m(x) = \lim_{n\to\infty} (1 - \cos^{2n}(m!\pi x)).$$

Clearly,

$$g_m(x) = \begin{cases} 0, & \text{if } m!x \in \mathbb{Z}, \\ 1, & \text{if } m!x \notin \mathbb{Z}, \end{cases}$$

and so

$$\forall\, x \in \mathbb{R}, \quad \lim_{m\to\infty} g_m(x) = \mathbb{1}_{\mathbb{R}\setminus\mathbb{Q}}.$$

Remark. In Problem 1.15b we gave necessary conditions in order that a sequence of continuous functions converge pointwise to a function on $[a, b]$. In fact there are the following necessary and sufficient conditions, e.g., [64], pages 99–102. *A sequence $\{f_n : n = 1, \ldots\}$ of continuous functions $f_n : [a, b] \to \mathbb{R}$ converges pointwise to a function f if and only if for every closed set $P \subseteq [a, b]$ without isolated points*

$$\overline{C(f) \cap P} = P.$$

Also, from Theorem 2.5.7, say, we know that if a topological space X has enough continuous functions defined on it, e.g., if $X = [a, b]$, then the following assertion is true. *Let $f : X \to \mathbb{R}^*$ be measurable, $\mu(X) < \infty$, and let $\varepsilon > 0$; then there is a continuous function $h : X \to \mathbb{R}$ such that $\mu(\{x : |f(x) - h(x)| \geq \varepsilon\}) < \varepsilon$.* A much more powerful result is the *Vitali–Luzin theorem*, commonly called the *Luzin theorem*. We shall give Vitali's original proof from 1905 [484] for two reasons: first, for historical reasons, and, second, since it is an efficient and intuitive proof. Luzin's proof appeared in 1912 [332]; and there are standard proofs due to Sierpiński [436] and

L. Cohen [107]. All the proofs that we have seen have a similar conceptual flavor. Such a theorem, relating topological and measure-theoretic notions, was certainly thought of quite early, and Vitali suggests that Borel and Lebesgue might have known of it before him; in any case, Vitali published the first proof. Naturally the setting for Vitali was $([a,b], \mathcal{M}([a,b]), m)$. We shall state the result in its most convenient present-day setting, and this requires some definitions.

A Hausdorff topological space X is *locally compact* if each point has a neighborhood basis of compact sets (see Appendix A.1). Such spaces are auspicious in measure theory, since they guarantee the existence of nontrivial continuous functions, and we *do* want to integrate continuous functions. If the notion of local compactness is not in your toolkit yet you may think of X as an interval for the time being.

Definition 2.5.12. Regular Borel measures

Let X be a locally compact Hausdorff space and let $(X, \mathcal{A}, \mu)$ be a measure space for which $\mathcal{B}(X) \subseteq \mathcal{A}$. Then μ is a *Borel measure* and $(X, \mathcal{A}, \mu)$ is a *Borel measure space*. The measure μ is a *regular Borel measure* and $(X, \mathcal{A}, \mu)$ is a *regular Borel measure space* if

i. $\forall\, F \subseteq X,$ compact, $\mu(F) < \infty,$

ii. $\forall\, A \in \mathcal{A},$

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

iii. $\forall\, U \subseteq X$, open or $\mu(U) < \infty,$

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

This definition is equivalent to an apparently weaker criterion, where the supremum property over compact F is omitted in the case, $\mu(U) < \infty$; see [235], Theorem 10.30 and page 178.

Theorem 2.5.13. Vitali–Luzin theorem

Let $(X, \mathcal{A}, \mu)$ be a measure space, where X is a locally compact Hausdorff space and μ is regular. Choose $A \in \mathcal{A}$ for which $\mu(A) < \infty$ and take a measurable function $f : X \to \mathbb{R}^*$ that vanishes on $A^\sim$. For each $\varepsilon > 0$ there is a continuous function $g : X \to \mathbb{R}$ that vanishes outside of a compact set such that

$$\mu\left(\{x : f(x) \neq g(x)\}\right) < \varepsilon.$$

Proof. Without loss of generality assume that $\mu(X) < \infty$ and $A = X$. We first prove that

$$\forall\, \varepsilon, \delta > 0,\ \exists\, F \subseteq X, \text{ compact, such that } \mu(F) > \mu(X) - \varepsilon, \tag{2.23}$$

and that

$$\forall\, x \in F,\ \exists\, U_x,\ \text{an open neighborhood of } x,\ \text{such that} \quad \sup_{y \in U_x \cap F} f(y) - \inf_{y \in U_x \cap F} f(y) \le \delta; \tag{2.24}$$

cf. the notion of oscillation $\omega(f, I)$ of f on I defined before Proposition 1.3.6. To do this, first set

$$A_{n,\delta} = \{x : n\delta \le f(x) < (n+1)\delta\},$$

and note that

$$\mu(X) = \sum_{n=-\infty}^{\infty} \mu_{n,\delta},$$

where $\mu_{n,\delta} = \mu(A_{n,\delta})$. Choose $n_1, \ldots, n_k$ such that

$$\mu(X) - \sum_{j=1}^{k} \mu_{n_j,\delta} < \frac{\varepsilon}{2},$$

and let $r_j < \mu_{n_j,\delta}$, $j = 1, \ldots, k$, have the property that $\sum_{j=1}^{k} r_j < \varepsilon/2$. From the regularity of μ there are compact sets $F_j \subseteq A_{n_j,\delta}$, $j = 1, \ldots, k$, for which $\mu(F_j) > \mu_{n_j,\delta} - r_j$. Consequently, $F = \bigcup_{j=1}^{k} F_j$ is compact and

$$\mu(F) = \sum_{j=1}^{k} \mu(F_j) > \sum_{j=1}^{k} (\mu_{n_j,\delta} - r_j) > \mu(X) - \varepsilon.$$

Thus, (2.23) is obtained; and (2.24) follows from the definition of $A_{n,\delta}$.

We now use (2.23) and (2.24) countably often to obtain the result. Let $\varepsilon > 0$. Choose $\varepsilon_j, \delta_j > 0$ such that $\sum_{j=1}^{\infty} \varepsilon_j < \varepsilon$ and $\lim_{j \to \infty} \delta_j = 0$. At the first step let F_1 be compact, with measure $\mu(F_1) > \mu(X) - \varepsilon_1$, and let it also satisfy the oscillation condition (2.24) with $\delta = \delta_1$.

Choose a compact set $F_2 \subseteq F_1$ such that $\mu(F_2) > \mu(F_1) - \varepsilon_2$ and such that the oscillation condition (2.24) is satisfied with $\delta = \delta_2$. Thus, $\mu(\bigcap F_j) > \mu(X) - \varepsilon$ and f is continuous on $\bigcap F_j$. The fact that we can extend f from the compact set $\bigcap F_j$ to a continuous function g on X follows from the Urysohn lemma (Theorem A.1.3). □

Nicolas Bourbaki [69], Chapitre IV.5.1, uses the Vitali–Luzin criterion as the definition of measurable function. In fact, we have the following corollary to Theorem 2.5.13.

Corollary 2.5.14. *Let $(X, \mathcal{A}, \mu)$ be a measure space, where X is a locally compact Hausdorff space and μ is regular. A function $f : X \to \mathbb{R}^*$ that is finite μ-a.e. is measurable if and only if for every compact set $K \subseteq X$ and for every $\varepsilon > 0$,*

$$\exists\, F \subseteq K, \text{ compact, such that } \mu(K \setminus F) < \varepsilon$$

and such that f restricted to F is continuous.

A locally compact Hausdorff space is *σ-compact* if it is the countable union of compact sets.

Corollary 2.5.15. *Let $(X, \mathcal{A}, \mu)$ be a complete measure space, where X is σ-compact and μ is regular. A function $f : X \to \mathbb{R}^*$ that is finite μ-a.e. is measurable if and only if there is a sequence of continuous functions $f_n : X \to \mathbb{R}$ such that $f_n \to f$ μ-a.e.*

Proof. The sufficient conditions follow from Theorem 2.5.4. For the necessary conditions we use Theorem 2.5.13, and for each n we choose a continuous function f_n such that

$$\mu(A_n) = \mu(\{x : f(x) \neq f_n(x)\}) < \frac{1}{2^n}. \tag{2.25}$$

Setting

$$A = \bigcap_{k=1}^{\infty} \bigcup_{n=k}^{\infty} A_n,$$

we have $\mu(A) = 0$, since $\mu(\bigcup_{m=k}^{\infty} A_m) \leq 1/2^{k-1}$; cf. Theorem 2.4.4.

Also, by definition, A is precisely the set of points x that are in infinitely many A_k. Consequently, if $x \notin A$ then x is in at most finitely many A_k, and so $f_n(x) = f(x)$ for all large n. Then $f_n \to f$ μ-*a.e.* $\square$

2.6 Potpourri and titillation

1. The *axiom of choice* (AC) states that *for every family* $\mathbf{A}$ *of disjoint nonempty sets there exists a set B that has exactly one element in common with each set in* $\mathbf{A}$.

AC is equivalent to the following statement KZ: *If in a nonempty, partially ordered set A each linearly ordered subset has a supremum then there exists a maximal element of A.*

KZ is known as Zorn's lemma, the Zorn maximum principle, or the Kuratowski–Zorn lemma. KZ was proved independently by Kazimierz Kuratowski [306] and Max Zorn [523]. The formulations of the two authors were different, but they turned out to be equivalent to each other and to AC; e.g., see [278], [279], [308].

There is an unverifiable anecdote that relates that Kuratowski was once attending a conference at which Zorn was a speaker. During his talk, Zorn referred to Zorn's lemma, and then paused and said "Zorn is me, and over there sits Professor Kuratowski, who proved it first".

For a detailed treatment of AC, see [231].

2. In 1964, Robert M. Solovay proved that a stronger version of Zermelo–Fraenkel set theory (ZF), but one that still does not contain the axiom of choice, is consistent, i.e., there is a model of set theory with the property LM, which states that

$$\text{every } A \subseteq \mathbb{R} \text{ is Lebesgue measurable.}$$

The main paper in which his work appears is [446]; see [104], [494] for more recent expositions. Solovay's theorem does *not* say that

$$\text{if AC fails then LM,} \tag{2.26}$$

i.e., it does not say that AC is necessary to find $A \notin \mathcal{M}(\mathbb{R})$. In fact, in Paul J. Cohen's model for establishing certain results concerning the independence of the axiom of choice fails, and there are nonmeasurable sets. Solovay's result does say that LM is consistent with but independent of the failure of AC. If we wanted to prove that the existence of a nonmeasurable set implies the axiom of choice, then we would have to show that in every model in which AC fails every set is Lebesgue measurable. Solovay has given one such model in which this occurs. He also proves that ZF is consistent with the property that every uncountable set $X \subseteq \mathbb{R}$ contains an uncountable closed set; e.g., Problem 1.4*b*, Example 2.2.14, and Problem 2.10. Related to statement (2.26) he shows that even with AC we cannot produce a definable (from a countable sequence of ordinals) set $A \notin \mathcal{M}(\mathbb{R})$. We mention again the reference [104], which makes all of this understandable to nonexperts like ourselves.

3. Cantor's *continuum hypothesis* is the statement that

$$\forall\, A \subseteq \mathbb{R}, \quad \text{card}\, A > \aleph_0 \implies \text{card}\, A = \text{card}\, \mathbb{R},$$

i.e., card $A > \aleph_0$ implies that there is a bijection $A \to \mathbb{R}$. Explicitly using the continuum hypothesis, Banach and Kuratowski [22] and Ulam [477] proved that there is no nontrivial measure on all of $\mathcal{P}(\mathbb{R})$ that satisfies (2.2); see [440], pages 107–109, and [362], pages 24–26. In fact, the result is that there are no nontrivial continuous measures (such measures are defined in Chapters 4 and 5) defined on $\mathcal{P}(\mathbb{R})$. These are stronger conclusions than what we proved in Example 2.2.16. In 1950, Shizuo Kakutani, Kunihiko Kodaira, and Oxtoby [272], [273] proved a positive result in this area by extending Lebesgue measure to a large σ-algebra $\mathcal{A} \supsetneq \mathcal{M}(\mathbb{R})$ while preserving the translation-invariance property (2.1) on $\mathcal{A}$; see also [247]. It is not difficult to define translation-invariant measures on $\mathcal{P}(\mathbb{R})$ that do not have property (2.2), e.g., Example 2.4.2.

4. With regard to the continuum hypothesis, it can be verified that *if* $F \subseteq \mathbb{R}$ *is closed and uncountable then*

$$\text{card}\; F = \text{card}\; \mathbb{R}.$$

This is a consequence of the following results.

i. If card $F > \aleph_0$*, then* $F = P \cup D$*, where* D *is countable and* P *is a closed set without isolated points.*

ii. If $P \subseteq \mathbb{R}$ *is a nonempty closed set without isolated points then* card $P =$ card $\mathbb{R}$.

Part *i* is the *Cantor–Bendixson theorem.* Part *ii* is proved using CANTOR's theorem, which asserts that *if* $\{F_n : n = 1, \ldots\} \subseteq \mathcal{P}(\mathbb{R})$ *is a nested decreasing sequence of closed sets whose diameters tend to* 0 *then* $\bigcap F_n$ *is a single point.* We shall prove CANTOR's result in the proof of the Baire category theorem (Theorem A.6.1). Compare CANTOR's theorem with Theorem 1.2.10*b*.

Two of the major theorems dealing with the continuum hypothesis are due to KURT GÖDEL (late 1930s) and P. COHEN (1966). Both results assume the consistency of the traditional axioms of set theory, viz., the Zermelo–Fraenkel axioms and the axiom of choice, but conclude with apparently alarmingly diverse conclusions about the continuum hypothesis. A recent analysis of the topic, which placated one of the authors, who studied GÖDEL's Annals of Mathematics Studies in 1960–1961 as an assignment from GLEASON (see page 1), is found in [513], [514].

5. In Example 2.2.16 the countable subadditivity of m was crucial to establishing the existence of an element $A \in \mathcal{P}(\mathbb{R}) \setminus \mathcal{M}(\mathbb{R})$. For $\mathbb{R}^3$ there is the following more general result.

Theorem 2.6.1. Nonexistence of finitely additive set functions in $\mathcal{P}(\mathbb{R}^3)$

There is no nontrivial finitely additive set function $\mu : \mathcal{P}(\mathbb{R}^3) \to \mathbb{R}^+ \cup \{+\infty\}$ *such that* $\mu(B) < \infty$ *for bounded sets* $B \subseteq \mathbb{R}^3$ *and for which* $\mu(A) = \mu(B)$ *if* A *and* B *are isometric.*

(See Appendix A.1 for the definition of isometry.)

Theorem 2.6.1 is proved using the axiom of choice in the form of the *Hausdorff paradox*: *Let* S^2 *be the surface of the unit sphere in* $\mathbb{R}^3$*; then there is a countable set* $D \subseteq S^2$ *and there are disjoint subsets* $A, B, C \subseteq S^2 \setminus D$ *such that*

$$S^2 \setminus D = A \cup B \cup C \ \text{ and } \ A \simeq B \simeq C \simeq B \cup C,$$

where "$\simeq$" designates *congruence*, i.e., a surjective isometry. We refer to [441] for an exposition of the equivalence of AC and the Hausdorff paradox. To prove Theorem 2.6.1 from the Hausdorff paradox, assume that such a μ exists and define $\nu : \mathcal{P}(S^2) \to \mathbb{R}^+ \cup \{+\infty\}$ as follows: if $X \subseteq S^2$, $\nu(X)$ is μ of the union of all radii of the unit sphere with endpoints in X. Then, ν satisfies the same conditions as μ and we obtain a contradiction to the hypothesis of finite additivity because of FELIX HAUSDORFF's result.

The *Banach–Tarski paradox*, see [25], is a corollary of the Hausdorff paradox. We shall settle for a statement of the Banach–Tarski result and again refer to [441] for details: *Let* $A, B \subseteq \mathbb{R}^3$ *be disjoint solid spheres having the* *same radius; there exist* $C_1, \dots, C_{41} \subseteq A$ *and* $D_1, \dots, D_{41} \subseteq A \cup B$ *such that* $A = \bigcup C_j$*,* $A \cup B = \bigcup D_j$*,* $C_j \simeq D_j$ *for each* j*, and*

$$C_i \cap C_j = D_i \cap D_j = \emptyset, \ \ 1 \le i < j \le 41.$$

There is also a proof of the Banach–Tarski paradox as a consequence of the Hahn–Banach theorem; see [367]. An explanation of this phenomenon is due to John (Johann) von Neumann, who showed that the Banach–Tarski paradox is impossible in $\mathbb{R}$ and $\mathbb{R}^2$, because only $\mathbb{R}^d$ for $d \ge 3$ contains free nonabelian groups; e.g., see Jean Dieudonné's biography of von Neumann in [197].

As our final remark on these matters, Banach has shown that on $\mathbb{R}$ and $\mathbb{R}^2$ there are nonnegative finitely additive set functions on $\mathcal{P}(\mathbb{R})$ and $\mathcal{P}(\mathbb{R}^2)$ that are finite for bounded sets, translation-invariant, and equal to Lebesgue measure for Lebesgue measurable sets. It is consistent with ZF to extend Banach's result to the σ-additive case.

6. In *The Scottish Book* [344], Steinhaus proposed in the following problem related to finite measures.

Given three sets A_1*,* A_2*,* A_3 *in* $\mathbb{R}^3$*, each having finite Lebesgue measure, does there exist a plane cutting each of the three sets* A_1*,* A_2*,* A_3 *into two parts of equal measure? There is the analogous problem for* d *sets in* $\mathbb{R}^d$*.*

In $\mathbb{R}^3$ this problem was solved by Steinhaus in 1936. His result is called the "ham sandwich theorem"; see William A. Beyer and Andrew Zardecki, Amer. Math. Monthly 111 (2004), 58–61, for an early history of the ham sandwich theorem. The idea is to start by bisecting one set ("ham") by a plane and observing that continuous changes of these bisections, which cut the "two slices of bread", provide an odd mapping, g, i.e., for all $x \in S^2 \subseteq \mathbb{R}^3$ $g(-x) = -g(x)$, from $S^2 \subseteq \mathbb{R}^3$ into $\mathbb{R}^2$. The rest of the proof follows from an application of the *Borsuk–Ulam theorem*, conjectured by Ulam and proved by Karol Borsuk in 1932. In one of its forms the Borsuk–Ulam theorem asserts that *any continuous function from* S^d*, the* d*-dimensional sphere in* $\mathbb{R}^{d+1}$*, to* $\mathbb{R}^d$*, must send some pair of opposite (bipolar) points to the same point*; see Andrew Browder, Amer. Math. Monthly 113 (2006), 935–937, for an elementary proof in dimension 2.

The Borsuk–Ulam theorem itself has led to extensive developments. For example, it implies the *Brouwer fixed-point theorem*; see Francis E. Su, Amer. Math. Monthly 104 (1997), 855–859.

7. In his 1933 classic [293], Kolmogorov defined a *probability space* as a triple $(X, \mathcal{R}, p)$ satisfying the following conditions on the set X, the algebra $\mathcal{R}$ of subsets of X, and the function $p : \mathcal{R} \to \mathbb{R}^+$: $p(X) = 1$, p is finitely additive, and if $\{A_n : n = 1, \dots\} \subseteq \mathcal{R}$ is decreasing and $\bigcap A_n = \emptyset$, then $\lim_{n\to\infty} p(A_n) = 0$. In this case the following theorem can be proved; cf. Theorems 2.2.6*c*, 2.3.7, 2.4.1, and 2.4.3*c*, as well as the uniqueness results in Problems 2.19 and 2.20: *Let* $(X, \mathcal{R}, p)$ *be a probability space; there is a*

unique measure on $\mathcal{A}$, the σ-algebra generated by $\mathcal{R}$, that is an extension of p. This measure is the *probability measure* on $\mathcal{A}$ and is also denoted by p.

Because of this theorem and the results proved in Chapter 3, and as mentioned prior to Theorem 2.4.6, we say that a *probability space* is any triple $(X, \mathcal{A}, p)$, where X is a set, $\mathcal{A}$ is a σ-algebra in $\mathcal{P}(X)$, and p is a measure for which $p(X) = 1$.

The definition of a probability space frequently assumes the completeness of the measure space, and, as we have seen, the Carathéodory extension procedure leads automatically to complete measures. After KOLMOGOROV's short treatise [293] there appeared several other major probability books in the 1940s and 1950s including those of ALEKSANDR Y. KHINCHIN (1948) [284], PAUL LÉVY (1948) [319], WILLIAM FELLER (1950) [171], and JOSEPH DOOB (1953) [144]. Of course, probability theory has been around forever, e.g., the work of GEROLAMO CARDANO (1501–1576) on gambling and the remarkable contribution of probability to arithmetic by BOREL in 1909; and it is an important and multifaceted field.

One such facet with extraordinary implications is prediction theory and the spectral theory of minimal sequences due to KOLMOGOROV himself [292]. (There is an English translation by GOPINATH KALLIANPUR given to one of the authors by the magisterial PESI MASANI.) It is of the KOLMOGOROV paper that HARALD CRAMÉR (1976) wrote, "The fundamental importance of this work of Kolmogorov lies in the fact that he showed how the abstract theory of Hilbert space (as well, of course, as other types of spaces) could be applied to the theory of random variables and stochastic processes". We refer to [38] for an exposition of this area that relates several topics in modern analysis including frames and positive operator-valued measures, e.g., [47].

8. Let $(X, \mathcal{A}, p)$ be a probability space. An $\mathcal{A}$-measurable function $f : X \to \mathbb{R}$ is a *random variable*. In the case $f : X \to \mathbb{C}$ then f is a *random variable* if $\mathrm{Re}\,(f)$ and $\mathrm{Im}\,(f)$ are $\mathcal{A}$-measurable. It is elementary to check that *if $f : X \to \mathbb{R}$ is a random variable and $B \in \mathcal{B}(\mathbb{R})$, then $f^{-1}(B) \in \mathcal{A}$.*

One chooses $\mathcal{B}(\mathbb{R})$ and not $\mathcal{M}(\mathbb{R})$ in the probabilistic interpretation of a random variable for the following reasons. The $\mathcal{B}(\mathbb{R})$-case gives more random variables than the $\mathcal{M}(\mathbb{R})$-case, because $\mathcal{B}(\mathbb{R}) \subseteq \mathcal{M}(\mathbb{R})$; and it is easier to check that one has a random variable f in the $\mathcal{B}(\mathbb{R})$-case, because one need only check that $f^{-1}(I) \in \mathcal{A}$ for intervals I.

This measure-theoretic setting for probability theory and random variables is a natural model for many probabilistic experiments. In particular, X represents a set of *experimental outcomes*, $\mathcal{A}$ is a class of *events*, and p is the probability assigned to these events. For example, in the experiment of rolling a die, X consists of six faces of the die.

2.7 Problems

Some of the more elementary problems in this set are Problems 2.1, 2.3, 2.4, 2.7, 2.9, 2.11, 2.13, 2.14, 2.16, 2.18, 2.27, 2.29, 2.31, 2.33, 2.34, 2.36, 2.38, 2.39.

2.1. Prove Proposition 2.2.3.

2.2. Find a function $f : [0,1] \to \mathbb{R}$ and a set $D \subseteq [0,1]$ such that $D \in \mathcal{M}([0,1])$, $m(D) = 0$, D is uncountable, f is continuous on D, and f is discontinuous on $[0,1] \setminus D$.
[*Hint.* Modify the Cantor function appropriately. Also see the example in the proof of Proposition 2.2.9.]

2.3. ***a.*** Prove Theorem 2.2.6.

b. Give an example to show that the hypothesis is necessary in Theorem 2.2.6*c*.

c. Find two disjoint sets A_1, $A_2 \subseteq [0,1]$ such that

$$m^*(A_1 \cup A_2) < m^*(A_1) + m^*(A_2).$$

d. Find $\{A_j : j = 1, \ldots\} \subseteq \mathcal{P}([0,1])$, pairwise disjoint, such that

$$m^*\left(\bigcup_{j=1}^{\infty} A_j\right) < \sum_{j=1}^{\infty} m^*(A_j).$$

e. Find $\{A_j : j = 1, \ldots\} \subseteq \mathcal{P}([0,1])$ such that $A_{j+1} \subseteq A_j$ and

$$m^*\left(\bigcap_{j=1}^{\infty} A_j\right) < \lim_{j\to\infty} m^*(A_j).$$

Obviously, $m^*(A_j) < \infty$ for each j.

2.4. Let $\mu : \mathcal{A} \to \mathbb{R}^+ \cup \{\infty\}$ be a σ-additive set function on an algebra $\mathcal{A}$. Let μ^* be the associated outer measure. If $\mu^*(A) = 0$ prove that A is measurable.

2.5. Let $(\mathbb{R}^d, \mathcal{A}, \mu)$, $\mathcal{B}(\mathbb{R}^d) \subseteq \mathcal{A}$, be a regular measure space, and let μ^* be the associated outer measure. Generalize Proposition 2.2.3 to this situation.

2.6. Let $(X, \mathcal{A}, \mu)$ be a measure space, where $\mathcal{A}$ is a σ-algebra generated by an algebra $\mathcal{R}$. Prove that μ is σ-finite on $\mathcal{R}$ if and only if μ is σ-finite on $\mathcal{A}$.

2.7. Define $\mathcal{A} \subseteq \mathcal{P}(\mathbb{R})$ by the property that either card A or card $(\mathbb{R} \setminus A)$ is countable for $A \in \mathcal{A}$. The set function $\mu : \mathcal{A} \to \mathbb{R} \cup \{\infty\}$ is given by

$$\mu(A) = \begin{cases} 0, & \text{if card } A \le \aleph_0, \\ \infty, & \text{if card } (\mathbb{R} \setminus A) \le \aleph_0. \end{cases}$$

Prove that $(\mathbb{R}, \mathcal{A}, \mu)$ is a measure space.

2.8. For each $X \subseteq \mathbb{R}$, define

$$X_d = \{x \in (0,1) : \exists\, k \in \mathbb{Z} \text{ such that } kx \in X\}.$$

Prove that

$$\forall\, N > 0 \text{ and } \forall\, \varepsilon > 0,\ \exists\, X \subseteq \mathbb{R} \text{ such that } m(X) > N \text{ and } m(X_d) < \varepsilon.$$

[*Hint.* Begin with the following special case. Take

$$X = \bigcup_{k=1}^{[n/\delta]} \left(\frac{n}{k}, \frac{n+\delta}{k} \right),$$

where $[n/\delta]$ is the greatest integer less than or equal to n/δ. Show that

$$\forall\, n,\ \ m(X) = \sum_{k=1}^{[n/\delta]} \frac{\delta}{k} > \delta \log\left(\frac{n}{\delta}\right)$$

and that for large n

$$m(X_d) = \delta + \sum_{k=n+1}^{[n/\delta]} \frac{\delta}{k} < \delta + 2\delta \log\left(\frac{1}{\delta}\right).$$

Consequently, "$m(X) \to \infty$" and "$m(X_d) \to 0$".]

2.9. ***a.*** Find $X \subseteq [0,1]$ such that $m(X) = 0$ and X is not of first category.
[*Hint.* Consider a perfect symmetric set with perfect symmetric sets in each of its contiguous intervals, etc.; make the measures add up to 1 and look at the complementary set.]

b. For each $\varepsilon > 0$ construct an open set $U \subseteq [0,1]$ such that $\overline{U} = [0,1]$ and $m(U) = \varepsilon$.
[*Hint.* Take intervals of length $\varepsilon/2^k$ about the rationals.]

c. Find $A, B \subseteq [0,1]$ such that $A, B \in \mathcal{M}([0,1])$, $A \cap B = \emptyset$,

$$m(A) = 0 \text{ and } A \text{ is not of first category,}$$

and

$$m(B) = 1 \text{ and } B \text{ is of first category.}$$

[*Hint.* Check Proposition 2.2.9. Can you find a simpler example?]

2.10. ***a.*** Prove that every $B \in \mathcal{B}(\mathbb{R})$ that is not of first category contains a closed uncountable subset.

b. Is part *a* true for uncountable Borel sets of first category?

c. Prove that Felix Bernstein's example, which is what we reference in Problem 1.4*b*, contains uncountable sets of measure 0.

2.11. Prove Theorem 2.2.12.

2.12. Let m^2 be Lebesgue measure on $\mathbb{R}^2$.

a. A *packing problem.* Let U be the open unit disk in $\mathbb{R}^2$ and let $\{U_n : n = 1, \ldots\}$ be a sequence of open disks in $\mathbb{R}^2$ such that

i. $\overline{U}_n \subseteq U$ for each n,
ii. $\{\overline{U}_n : n = 1, \ldots\}$ is pairwise disjoint,
iii. $\sum r_n < \infty$, where r_n is the radius of U_n.

Define $X = U \setminus \bigcup_{n=1}^{\infty} U_n$ and prove $m(X) > 0$; see [499].

b. Show that every set of positive Lebesgue measure in $\mathbb{R}^2$ contains the vertices of an equilateral triangle.

c. Is part *b* true for other polygons in $\mathbb{R}^2$?

Remark. There are many packing problems with many modern applications, e.g., to error-correcting codes [468]. We refer to the treatise by John Conway and Neil J. A. Sloane [114] for a tantalizing and deep treatment.

Related to packing there is the following problem/theorem due to Herman Auerbach, Banach, Mazur, and Ulam. It is called the "sack of potatoes theorem", and it appeared in *The Scottish Book* [344] in the following form: *If* $\{K_n\}_{n=1}^{\infty}$ *is a sequence of convex bodies in* $\mathbb{R}^d$*, each of diameter less than or equal to a, and let the sum of their volumes be less than or equal to b; prove that there exists a cube with diameter* $c = f(a, b)$ *such that one can put all the given bodies in it disjointly.* Recall that $X \subseteq \mathbb{R}^d$ is a *convex set* if, for any $x, y \in \mathbb{R}^d$, the line segment $\{z = (1-t)x + ty : 0 \leq t \leq 1\} \subseteq \mathbb{R}^d$ is contained in X. The *diameter* of X is defined in Definition A.1.4*f*.

The first published proof is in [300].

2.13. ***a.*** Prove that there are no numbers ε and δ, such that $0 < \varepsilon \leq \delta < 1$, that have the following property: if $\{A_n : n = 1, \ldots\} \subseteq \mathcal{P}([0,1])$ is any sequence of Lebesgue measurable sets each satisfying $m(A_n) \geq \delta$ then there is a set A of measure ε for which $A \subseteq A_n$ for infinitely many n.
[*Hint.* Let A_n be the set of $x \in (0,1)$ such that the nth digit in its decimal expansion is nonzero.]

b. Let $\{A_n : n = 1, \ldots\} \subseteq \mathcal{M}([0,1])$, $A_n \subseteq [0,1]$, and assume that 1 is a limit point of $\{m(A_n) : n = 1, \ldots\}$. Prove that there is a subsequence $\{n_k : k = 1, \ldots\}$ for which

$$m\left(\bigcap_{k=1}^{\infty} A_{n_k}\right) > 0. \tag{2.27}$$

[*Hint.* $\sum_{k=1}^{\infty}(1 - m(A_{n_k})) < 1$.]

Remark. If we begin by taking an uncountable collection then (2.27) is always possible. For further results on this type of problem we refer to [195], [196].

2.14. Let $\mu : \mathcal{R} \to \mathbb{R}^+ \cup \{\infty\}$ be a σ-additive set function on a ring $\mathcal{R}$. Prove that μ satisfies the following properties for $E, E_j, F \in \mathcal{R}$:

a. $E \subseteq F \implies \mu(E) \leq \mu(F)$;
b. $E \subseteq F \implies \mu(E) + \mu(F \setminus E) = \mu(F)$;
c. $E \subseteq \bigcup_{j=1}^{\infty} E_j$ and $\mu(E) < \infty \implies \mu(E) \leq \sum_{j=1}^{\infty} \mu(E_j)$;
d. if $\bigcup_{j=1}^{\infty} E_j \subseteq E$ and if the E_j are pairwise disjoint, then

$$\sum_{j=1}^{\infty} \mu(E_j) \leq \mu(E).$$

2.15. Let $\mathcal{R} \subseteq \mathcal{P}(X)$ be a ring and let $\mu : \mathcal{R} \to \mathbb{R}^+ \cup \{+\infty\}$ be a σ-additive set function. Let $E \in \mathcal{H}(\mathcal{R})$ and let $\{A_1, \ldots, A_N\} \subseteq \mathcal{A}$ be a collection of measurable, pairwise disjoint elements of the σ-algebra $\mathcal{A}$ generated by $\mathcal{R}$. Prove the equality

$$\mu^*\left(\bigcup_{n=1}^{N} A_n \cap E\right) = \sum_{n=1}^{N} \mu^*(A_n \cap E).$$

[*Hint.* First prove it for $N = 2$ and use induction.]

2.16. Prove that a ring $\mathcal{R}$ that is closed under taking countable unions of its elements is a σ-ring.

2.17. Prove that a family of sets that is closed under countable increasing unions and under countable decreasing intersections, and that contains an algebra $\mathcal{A}$, also contains the σ-algebra generated by $\mathcal{A}$.

2.18. Let $(X, \mathcal{A}, \mu)$ be a measure space, and let $A, B \in \mathcal{A}$. Show that if $\mu(A \cup B) = \mu(A) + \mu(B)$, then $\mu(A \cap B) = 0$.

2.19. Let $(X, \mathcal{A})$ be a measurable space, and let μ and ν be measures on $(X, \mathcal{A})$. Assume that $\mu = \nu$ on some subset $\mathcal{D} \subseteq \mathcal{A}$. A reasonable problem is to consider nontrivial conditions in order that $\mu = \nu$ on $\mathcal{A}$. More specifically, prove the following result. *Assume that X and $\emptyset$ are elements of $\mathcal{D}$, that $\mathcal{D}$ is closed under finite unions and intersections, and that $\mathcal{A}$ is the smallest σ-algebra containing $\mathcal{D}$; if*

$$\exists\, \{D_n : n = 1, \ldots\} \subseteq \mathcal{D} \textit{ such that } \bigcup_{n=1}^{\infty} D_n = X \text{ and } \forall\, n,\ \mu(D_n) < \infty$$

then $\mu = \nu$ on $\mathcal{A}$; cf. Problem 2.20. Consequently, we see that a measure defined on the Borel subsets of the line is uniquely determined there by its values on the half-open intervals $(a, b]$, including $(-\infty, b]$ and (a, ∞). This uniqueness issue leads to the following problem: Suppose that ν is not necessarily a measure but only finitely additive on $\mathcal{D}$, whereas the other hypotheses for the above problem are satisfied; when can we conclude that $\mu = \nu$ on $\mathcal{A}$?

2.20. Let X be a set and let $\mathcal{R} \subseteq \mathcal{P}(X)$ be an algebra. Assume that the set function $\mu : \mathcal{R} \to \mathbb{R}^+$, for which $\mu(\emptyset) = 0$, satisfies

$$\forall \{A_j\} \subseteq \mathcal{R}, \text{ a disjoint sequence such that } \bigcup_{j=1}^{\infty} A_j \in \mathcal{R},$$

$$\mu\left(\bigcup_{j=1}^{\infty} A_j\right) = \sum_{j=1}^{\infty} \mu(A_j).$$

Define

$$\forall\, E \in \mathcal{P}(X), \quad \mu^*(E) = \inf\left\{\sum_{j=1}^{\infty} \mu(A_j)\right\},$$

where the infimum is taken over all collections $\{A_j : j = 1, \ldots\} \subseteq \mathcal{R}$ that cover E. Define $E \in \mathcal{P}(X)$ to be μ^**-measurable* if

$$\forall\, F \in \mathcal{P}(X), \quad \mu^*(F) \geq \mu^*(F \cap E) + \mu^*(F \cap E^\sim);$$

cf. the development in Section 2.2.

a. Prove the *Carathéodory theorem*: *The μ^*-measurable sets form a σ-algebra $\mathcal{A}$ containing $\mathcal{R}$ such that $(X, \mathcal{A}, \mu^*)$ is a measure space and $\mu^* = \mu$ on $\mathcal{R}$*. Note that $\mathcal{A}$ is not the σ-algebra generated by $\mathcal{R}$, but the σ-algebra of measurable sets with respect to $\mathcal{R}$ and μ^*.

b. Prove that *if μ is σ-finite then μ^* is the unique extension of μ as a measure to the smallest σ-algebra containing $\mathcal{A}$* (this latter part is precisely Problem 2.19).

2.21. Let $f : \mathbb{R} \to \mathbb{R}$ be an increasing, right continuous function. Prove that the set function μ_f defined in Example 2.3.10 is σ-additive on the ring of disjoint unions of parallelepipeds in $\mathbb{R}^d$. The set of parallelepipeds in $\mathbb{R}^d$ is an example of a semiring; see Problem 2.22.

2.22. We say that a collection $\mathcal{S} \subseteq \mathcal{P}(X)$ is a *semiring* if

$$\forall\, A,\, B \in \mathcal{S}, \quad A \cap B \in \mathcal{S}$$

and

$$A \setminus B = \bigcup_{j=1}^{n} A_j, \ \text{ for some pairwise disjoint sequence } \{A_1, \ldots, A_n\} \subseteq \mathcal{S}.$$

If $\mathcal{S} \subseteq \mathcal{P}(X)$ is a semiring and $X \in \mathcal{S}$, then we say that $\mathcal{S}$ is a *semialgebra*.

a. If $\mathcal{R}$ is a ring generated by a semiring $\mathcal{S}$, i.e., the smallest ring that contains $\mathcal{S}$, prove that

$$\forall\, A \in \mathcal{R},\ A = \bigcup_{j=1}^{n} A_j, \text{ for some pairwise disjoint sequence } \{A_1, \ldots, A_n\} \subseteq \mathcal{S}.$$

b. Let $\mu : \mathcal{S} \to \mathbb{R}^+ \cup \{\infty\}$ be a σ-additive set function on a semiring $\mathcal{S}$. Prove that there exists a unique extension of μ to a σ-additive set function on the ring generated by $\mathcal{S}$.

Thus, if, in addition, $\mathcal{S}$ is a semialgebra and μ is σ-finite on $\mathcal{S}$, then there exists a unique extension of μ from $\mathcal{S}$ to a σ-additive set function on the σ-algebra generated by $\mathcal{S}$.
[*Hint.* Use part *a*.]

2.23. Let $f : \mathbb{R} \to \mathbb{R}$ be a Lebesgue measurable function and let $G : \mathbb{R}^2 \to \mathbb{R}$ be continuous. Show that if

$$\forall\, x, y \in \mathbb{R}, \;\; |f(x+y)| \le G(f(x), f(y)),$$

then f is bounded on bounded sets.
[*Hint.* $h(x) = |f(x)| + |f(-x)|$ is measurable, and so there is a set $X \subseteq \mathbb{R}$ such that $m(X) > 0$ and h is bounded on X; thus for all $x, y \in X$, $f(x-y)$ is bounded. Since $X - X$ is a neighborhood of 0 (a useful result with an ingenious proof due to STEINHAUS, which we shall give in Problem 3.6), f is bounded on U. By applying induction to the hypothesis we see that $f(nz)$ is bounded, for $z \in U$; and so f is bounded on any bounded set.]

2.24. ***a.*** Let $A \subseteq [0,1]$, and assume $m^*(A) > 0$. Prove that there is a nonmeasurable set $E \subseteq A$. Can you show that for any $\alpha \in (0,1)$ there is a nonmeasurable set $E \subseteq [0,1]$ for which $m^*(E) = \alpha$? For purposes of comparison note Example 2.4.13.

b. For $Y \subseteq \mathbb{R}$ let $\tau_{-\alpha} Y = \{y + \alpha : y \in Y\}$. If $S = \bigcup_{k=-\infty}^{\infty} [2k, 2k+1)$, prove that $\tau_{-\alpha} S = S^{\sim}$ when α is an odd integer.

c. Define $\tau S = \inf\, \{\alpha \in \mathbb{R}, \alpha > 0 : \tau_{-\alpha} S = S^{\sim}\}$; thus τS is 1 for S defined in part *b*. Prove that if $\tau S = 0$, then $S \notin \mathcal{M}(\mathbb{R})$.

d. The set $H \subseteq \mathbb{R}$ is a *Hamel basis* if

$$\forall\, x \in \mathbb{R}, \; \exists\, \{r_\alpha\} \subseteq \mathbb{Q} \;\text{ and }\; \exists\, \{h_\alpha\} \subseteq H, \;\text{ such that } x = \sum r_\alpha h_\alpha,$$

where the sum is finite and the representation is unique. Using Zorn's lemma, which is an equivalent form of the axiom of choice, it is easy to prove that Hamel bases exist using the following argument: let $\mathcal{F}$ be the family of all subsets $S \subseteq \mathbb{R}$ that are linearly independent over $\mathbb{Q}$; then there is a maximal element $H \in \mathcal{F}$. Prove that if H is a Hamel basis and $H \in \mathcal{M}(\mathbb{R})$ then $m(H) = 0$. Thus, if H is a Hamel basis and $m^*(H) > 0$, we conclude that $H \notin \mathcal{M}(\mathbb{R})$.

e. The existence of a Hamel basis with $m^*(H) > 0$ is a corollary of the following fact: *There is a Hamel basis H_B that intersects every closed uncountable set in* $\mathbb{R}$.

This result was first proved by C. BURSTIN in 1916, and a simpler proof of this fact is due to ALEXANDER ABIAN [1]. Assuming BURSTIN's result prove that $m^*(H_B) > 0$.

f. Let $H = \{x\} \cup \{x_\alpha : \alpha \in A,\ \text{ an index set}\} \subseteq \mathbb{R}$ be a Hamel basis, and define X to be the set of elements $y \in (0,1)$ whose Hamel expansions do not use x. Prove that $X \notin \mathcal{M}(\mathbb{R})$, e.g., [2], [434].

Remark. In [440], SIERPIŃSKI defined a set $X \subseteq \mathbb{R}$ to have the *property* S if

$$\forall\, A \subseteq \mathbb{R},\ m(A) = 0 \implies \text{card}\, A \cap X \leq \aleph_0.$$

Thus, X has the property S if and only if every uncountable subset of X is nonmeasurable. He then proved that uncountable sets with the property S exist if the continuum hypothesis is assumed; cf. [145] for a more recent development.

2.25. ***a.*** Let $(X, \mathcal{A}, \mu)$ be a measure space and let $E \in \mathcal{A}$. Suppose $M \in \mathcal{P}(X)$ has the property that the only measurable subsets of $M \cap E$ and $M^\sim \cap E$ are of measure zero. Prove that $\mu^*(M \cap E) = \mu^*(M^\sim \cap E) = \mu(E)$.

b. Let $(\mathbb{R}, \mathcal{M}(\mathbb{R}), m)$ be the Lebesgue measure space. Prove that there exists an m-measurable set $M \in \mathcal{M}(\mathbb{R})$ such that

$$\forall\, E \subseteq \mathcal{M}(\mathbb{R}), \quad m^*(M \cap E) = m^*(M^\sim \cap E) = m^*(E).$$

[*Hint.* Take any $\xi \in \mathbb{R}\backslash\mathbb{Q}$ and define $A = \{n+m\xi : n, m \in \mathbb{Z}\}$, $B = \{n+2m\xi : n, m \in \mathbb{Z}\}$, and $C = \{n+(2m+1)\xi : n, m \in \mathbb{Z}\}$. Similarly to Example 2.2.16, let $x \sim y$ if $x, y \in \mathbb{R}$ and $x - y \in A$. Clearly, "$\sim$" is an equivalence relation. Let S be a set that contains one representative from each equivalence class and define

$$M = S + B = \bigcup_{b \in B} (S + b).$$

Observe that $(M - M) \cap C = \emptyset$, and use STEINHAUS' theorem, mentioned in Problem 2.23 and Problem 3.6, to show that every measurable subset of M must have measure zero. Also, observe that $M^\sim = M + \xi$, and conclude a similar result for $M^\sim$. To finish the problem use part *a.*]

2.26. ***a.*** Consider the situation as in Theorem 2.4.1, i.e., let $\mathcal{R} \subseteq \mathcal{P}(X)$ be an algebra, and let $\mu : \mathcal{R} \to \mathbb{R}^+ \cup \{+\infty\}$ be a σ-additive set function that extends to a measure (also denoted by μ) on the σ-algebra of measurable sets $\mathcal{A}$. Prove that this extension $(X, \mathcal{A}, \mu)$ is a complete measure space.

b. Let $(X, \mathcal{A}, \mu)$ be a σ-finite measure space, where $\mathcal{A}$ is the σ-algebra generated by an algebra $\mathcal{R}$. Define $(X, \mathcal{A}_0, \mu_0)$ to be the complete measure space corresponding to $(X, \mathcal{A}, \mu)$, as in Theorem 2.4.8. Prove that $\mathcal{A}_0 = \mathcal{M}(\mathcal{R})$, the σ-algebra of measurable sets generated by $\mathcal{R}$ and μ.

[*Hint.* Clearly, by definition, $\mathcal{A}_0 \subseteq \mathcal{M}(\mathcal{R})$. To prove the other inclusion, we first note that if $A \in \mathcal{A}_0$ and $\mu^*(A) < \infty$, then $A \in \mathcal{A}_0$. This follows since, for a given $n > 0$, we can find $F_n \in \mathcal{A}$ such that $A \subseteq F_n$ and $\mu^*(F_n \setminus A) < 1/n$; and hence $F = \bigcap F_n \in \mathcal{A}$. The general assertion, that $A \in \mathcal{M}(\mathcal{R})$ implies $A \in \mathcal{A}_0$, is a consequence of the fact that $(X, \mathcal{A}, \mu)$ is σ-finite.]

2.27. With regard to Example 2.3.11, prove that $\lim_{d\to\infty} m^d(B_d) = 0$. In fact, the rate at which $\lim_{d\to\infty} m^d(B_d) = 0$ can be quantified. For example, prove that if $d = 2n$ is even, then, not only is $m^d(B_d) = \pi^n/n!$, but also

$$\frac{\pi^n}{n!} = \frac{1}{\sqrt{2\pi n}} \left(\frac{e\pi}{n}\right)^n u_n,$$

where $u_n \to 1$.
[*Hint.* Use the Stirling formula,

$$\lim_{n\to\infty} \frac{(n/e)^n\sqrt{2\pi n}}{n!} = 1;$$

e.g., [504].]

2.28. Find an example of a measure space $(\mathbb{R}, \mathcal{A}, \mu)$, for which $\mathcal{B}(\mathbb{R}) \subseteq \mathcal{A}$, and where it is not true that if $A \in \mathcal{M}(\mathbb{R})$ then $A = B \cup E$, for $B \in \mathcal{B}(\mathbb{R})$ and $m(E) = 0$.

2.29. Let $\mu : \mathcal{R} \to \mathbb{R}^+ \cup \{\infty\}$ be a σ-additive set function on an algebra $\mathcal{R}$. Prove that the outer measure μ^* is complete on the σ-algebra $\mathcal{A}$ of measurable sets generated by $\mathcal{R}$.

2.30. Prove Theorem 2.4.9.

2.31. Prove that there is a continuous function $f : [0, 1] \to \mathbb{R}$ such that $f(x) \in \mathbb{Q}$ m-*a.e.*, and for which f is not constant on any interval.
[*Hint.* At the first step form a sequence $\{f_{n,1} : n = 1, \ldots\}$ of continuous functions such that $f_{n,1} \to f_1$ uniformly and f_1 has rational values on a set of Lebesgue measure $1/2$.]

2.32. For each $x \in (0, 1]$, let $f(x) = \sum a_j/j$ when $x = .a_1 \ldots$ (2), where we take the expansion of x to have an infinite number of 1s in the case that x has the form $.a_1 \ldots a_n$ (2). Are $f^{-1}(\mathbb{R}^+)$ and $f^{-1}(\infty)$ Lebesgue measurable sets?

2.33. Let f be a decreasing and bounded real-valued function on $[0, 1]$. Show that there is a sequence $\{f_n : n = 1, \ldots\}$ of continuous decreasing functions such that $f_n \to f$ m-*a.e.*

2.34. Prove Theorem 2.5.4.
[*Hint.* Define $A \in \mathcal{A}$ with empty complement in terms of $\{f_n : n = 1, \ldots\}$ and f, set $g_n = f_n \mathbb{1}_A$, $g = f \mathbb{1}_A$, and use Proposition 2.4.16.]

2.35. Let $f : \mathbb{R} \to \mathbb{R}$ be Borel measurable. Prove that $f^{-1}(\mathcal{B}(\mathbb{R})) \subseteq \mathcal{B}(\mathbb{R})$.

2.36. ***a.*** Let $f : [a, b] \to \mathbb{R}$ be a Lebesgue measurable function. Prove that

$$\forall\, B \in \mathcal{B}(\mathbb{R}), \quad f^{-1}(B) \in \mathcal{M}([a, b]).$$

Thus, Lebesgue measurable functions f on $[a, b]$ are characterized by the condition, $f^{-1}(\mathcal{B}(\mathbb{R})) \subseteq \mathcal{M}([a, b])$, as Example 2.4.14 shows.
[*Hint.* Let $\mathcal{C} = \{A \subseteq \mathbb{R} : f^{-1}(A) \in \mathcal{M}([a, b])\}$ and prove that $\mathcal{C}$ is a σ-algebra.]

b. Give an example of a Lebesgue measurable function $f : [a, b] \to \mathbb{R}$ for which $f^{-1}(\mathcal{M}(\mathbb{R}))$ is not contained in $\mathcal{M}([a, b])$; cf. Figure 2.2.

c. Is it true that $f^{-1}(\mathcal{B}(\mathbb{R})) \subseteq \mathcal{B}([a, b])$ for every Lebesgue measurable function $f : [a, b] \to \mathbb{R}$?

2.37. Find an everywhere discontinuous function $f : \mathbb{R} \to \mathbb{R}$ such that f is measurable on $(\mathbb{R}, \mathcal{M}(\mathbb{R}), m)$ but not measurable on $(\mathbb{R}, \mathcal{B}(\mathbb{R}), m)$.
[*Hint.* Use the idea of Example 2.4.17.]

2.38. Let the functions $f, g : (0, 1) \to \mathbb{R}$ have the following property:

$$\forall\, \alpha \in \mathbb{R}, \;\; m(\{x : f(x) > \alpha\}) = m(\{x : g(x) > \alpha\}). \tag{2.28}$$

Prove that if f and g are left continuous for all $x \in (0, 1)$, as well as being decreasing, then $f = g$ on $(0, 1)$.
[*Hint.* Assume $f(x_0) > g(x_0)$, let $\varepsilon = f(x_0) - g(x_0)$, and find $\delta > 0$ such that

$$m(\{x : f(x) \geq f(x_0)\}) = x_0 \quad \text{and} \quad m(\{x : g(x) \geq g(x_0)\}) = x_0 - \delta.]$$

2.39. Find a sequence $\{f_n : n = 1, \ldots\}$ of functions $[0, 1] \to \mathbb{R}$ such that

$$\forall\, x \in [0, 1], \quad \lim_{n \to \infty} f_n(x) = 0,$$

whereas for each $[a, b] \subseteq [0, 1]$, $\{f_n : n = 1, \ldots\}$ does not converge uniformly on $[a, b]$.

2.40. Prove Corollary 2.5.6.

2.41. With respect to Problem 1.13, prove or disprove the following statement: for any fixed n there is a set $X \subseteq \mathbb{R}$, with $m(X) = 0$, such that

$$\forall\, d_1, \ldots, d_n > 0, \; \exists\, x_0, \ldots, x_n \in X \text{ for which } \forall\, j = 1, \ldots, n,$$
$$d_j = x_j - x_{j-1}.$$

2.42. Let $\alpha_n \to 0$. Find a bounded Lebesgue measurable function $f : [0, 1] \to \mathbb{R}$ such that $f(x - \alpha_n)$ does not tend to $f(x)$ m-*a.e.*
[*Hint.* Let $f = \mathbb{1}_E$, where E is a perfect symmetric set of positive Lebesgue measure.]

2.43. Let $(X, \mathcal{A}, \mu)$ be a measure space, let X be a locally compact Hausdorff space, let μ be a Borel measure, and assume that μ is regular on the Borel algebra $\mathcal{B}(X)$. Prove that μ need not be regular on $\mathcal{A}$.

Remark. A refinement of this exercise is found in [212], page 230, exercise (5); the related exercise (10) on page 231 is due to DIEUDONNÉ.

2.44. Let $\mathcal{L}$ be a collection of closed line segments in $\mathbb{R}^d$ and let $E(\mathcal{L})$ be the set of all endpoints of the members of $\mathcal{L}$. Any set $X \subseteq \mathbb{R}^d$ of two or more points has the form $E(\mathcal{L})$ for some $\mathcal{L}$. A set $X \subseteq \mathbb{R}^d$ is an *endset* if $X = E(\mathcal{L})$ for a pairwise disjoint collection $\mathcal{L}$. Prove the following results.

a. In $\mathbb{R}$, if S is an endset then $\operatorname{card} S \leq \aleph_0$.

b. In $\mathbb{R}^2$, if S is a closed bounded endset then $m^2(S) = 0$.

Remark. Part *b* does not extend to $\mathbb{R}^d$ for $d \geq 3$. In fact, for $d \geq 4$ there is a compact endset of positive Lebesgue measure. For this result and a discussion of the difficult case of $d = 3$, we refer to the Amer. Math. Monthly 78 (1971), 516–518, article by ANDREW M. BRUCKNER and J. G. CEDER.

2.45. ***a.*** Let $A \in \mathcal{M}(\mathbb{R})$, where $A \subseteq [0,1]$ and $0 < m(A) < 1$. Prove that

$$\inf\ m(A \cap I)/m(I) = 0 \ \text{ and } \ \sup\ m(A \cap I)/m(I) = 1,$$

where the inf and sup are taken over all nontrivial proper subintervals I of $[0,1]$.

b. Find $A \in \mathcal{M}(\mathbb{R})$, where $A \subseteq [0,1]$, such that for each nontrivial proper subinterval $I \subseteq [0,1]$,

$$m(A \cap I) > 0 \ \text{ and } \ m(A^{\sim} \cap I) > 0.$$

Remark. In BERNSTEIN's example, mentioned in Problems 1.4*b* and 2.10*c*, $\mathbb{R}$ is decomposed as a disjoint union $A \cup B$ where neither A nor B contains uncountable closed sets F but where for each such F, $A \cap F$, $B \cap F \neq \emptyset$. It is easy to check that neither A nor B is measurable; in fact, assume that A is measurable, approximate the measure from within by compact sets, and obtain a contradiction to the decomposition properties. Compare this situation with Problem 2.45*b*. Also, see Example A.6.7.

2.46. Let $(X, \mathcal{B}(X), \mu)$ be a Borel measure space on a separable metric space X; see Appendix A.1 and A.3 for definitions.

a. If μ is a bounded measure, prove that it is also regular on the Borel algebra $\mathcal{B}(X)$.

b. If μ is a σ-finite measure, prove that

$$\forall\ A \in \mathcal{B}(X), \quad \mu(A) = \sup\{\mu(F) : F \subseteq A \text{ and } F \text{ is compact}\}.$$

c. If μ is a σ-finite measure with the property that every compact set has finite measure, prove that μ is regular.

d. Find an example of a σ-finite measure on the measurable space $(X, \mathcal{B}(X))$ that does not satisfy the following condition:

$$\forall\ A \in \mathcal{B}(X), \quad \mu(A) = \inf\{\mu(U) : A \subseteq U \text{ and } U \text{ is open}\}.$$

2.47. A *homeomorphism* $h : [0,1] \to [0,1]$ is a continuous bijection whose inverse is also continuous; cf. the definition in Appendix A.1. Note that the Borel sets are invariant under homeomorphisms $h : [0,1] \to [0,1]$, and that such is not the case for the Lebesgue measurable sets (Example 2.4.14). This is another reason why in probability theory the class of "probabilizable" sets are the Borel sets and not the Lebesgue measurable sets; see Section 2.6.8.

a. Let $E \subseteq [0,1]$ be a closed nowhere dense set. Show that there is a homeomorphism $h : [0,1] \to [0,1]$ such that $m(h(E)) = 0$.
[*Hint.* Define

$$h(x) = m([0,x] \cap ([0,1] \setminus E))/m([0,1] \setminus E).]$$

b. Let $E \subseteq [0,1]$ be a set of first category. Show that there is a homeomorphism $h : [0,1] \to [0,1]$ such that $m(h(E)) = 0$.
[*Hint.* Let $\mathcal{H}$ be the family of homeomorphisms $[0,1] \to [0,1]$ with metric $\rho(f,g) = \sup_{x\in[0,1]} |f(x) - g(x)|$. If $E = \bigcup E_n$, where E_n is nowhere dense, define

$$A_{n,k} = \{h \in \mathcal{H} : m(\overline{E}_n) < 1/k\}.$$

Prove that $A_{n,k}$ is open in $\mathcal{H}$, and, setting $A = \bigcap A_{n,k}$, note that $m(h(E)) = 0$ for each $h \in A$.]

Remark. Actually it is possible to find an uncountable set $F \subseteq [0,1]$ such that

$$\forall\, h \in \mathcal{H},\ \ m(h(F)) = 0. \tag{2.29}$$

This result makes explicit use of the continuum hypothesis and was first proved by Abram S. Besicovitch [54]. For each $h \in \mathcal{H}$ there is a corresponding Hausdorff measure μ_h (we define this notion in Chapter 9), and Besicovitch actually proved $\mu_h(F) = 0$ for each $h \in \mathcal{H}$; (2.29) follows from this in a straightforward manner. Relative to our remark about Borel sets and homeomorphisms at the beginning of this exercise, we know that $F \in \mathcal{M}([0,1]) \setminus \mathcal{B}([0,1])$.

Besicovitch's solution to (2.29) is essentially equivalent to his solution of a famous conjecture made by Borel (1919). In order to state the conjecture we say that a set $F \subseteq [0,1]$ has the *property C* if

$$\forall\,\{a_n : a_n > 0, n = 1,\ldots\},\ \exists\,\{I_n = (c_n,d_n) : d_n - c_n = a_n, n = 1,\ldots\}$$

such that

$$F \subseteq \bigcup_{n=1}^{\infty} I_n.$$

The *Borel conjecture* was that every set with the property C had to be countable. In order to construct his counterexample to the Borel conjecture, Besicovitch defined the notion of a *concentrated set F in the neighborhood of a given countable set H* by the property that

$$\operatorname{card}(F \setminus (U \cap F)) \leq \aleph_0$$

for each open set U containing H. Using the continuum hypothesis, Besicovitch was able to construct such sets F for which card $F > \aleph_0$. These concentrated sets provide the solution to Borel's conjecture as well as to (2.29); see Section 9.3.

http://www.springer.com/978-0-8176-4656-1

Integration and Modern Analysis
Benedetto, J.J.; Czaja, W.
2009, XIX, 575 p. 24 illus.,
ISBN: 978-0-8176-4656-1
A product of Birkhäuser Basel