CHAPTER 2

Lebesgue Measure on $\mathbb{R}^n$

Our goal is to construct a notion of the volume, or Lebesgue measure, of rather general subsets of $\mathbb{R}^n$ that reduces to the usual volume of elementary geometrical sets such as cubes or rectangles.

If $\mathcal{L}(\mathbb{R}^n)$ denotes the collection of Lebesgue measurable sets and

$$\mu : \mathcal{L}(\mathbb{R}^n) \to [0, \infty]$$

denotes Lebesgue measure, then we want $\mathcal{L}(\mathbb{R}^n)$ to contain all n-dimensional rectangles and $\mu(R)$ should be the usual volume of a rectangle R. Moreover, we want μ to be countably additive. That is, if

$$\{A_i \in \mathcal{L}(\mathbb{R}^n) : i \in \mathbb{N}\}$$

is a countable collection of disjoint measurable sets, then their union should be measurable and

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

The reason for requiring countable additivity is that finite additivity is too weak a property to allow the justification of any limiting processes, while uncountable additivity is too strong; for example, it would imply that if the measure of a set consisting of a single point is zero, then the measure of every subset of $\mathbb{R}^n$ would be zero.

It is not possible to define the Lebesgue measure of all subsets of $\mathbb{R}^n$ in a geometrically reasonable way. Hausdorff (1914) showed that for any dimension $n \geq 1$, there is no countably additive measure defined on all subsets of $\mathbb{R}^n$ that is invariant under isometries (translations and rotations) and assigns measure one to the unit cube. He further showed that if $n \geq 3$, there is no such finitely additive measure. This result is dramatized by the Banach-Tarski 'paradox': Banach and Tarski (1924) showed that if $n \geq 3$, one can cut up a ball in $\mathbb{R}^n$ into a finite number of pieces and use isometries to reassemble the pieces into a ball of any desired volume *e.g.* reassemble a pea into the sun. The 'construction' of these pieces requires the axiom of choice.[1] Banach (1923) also showed that if $n = 1$ or $n = 2$ there are finitely additive, isometrically invariant extensions of Lebesgue measure on $\mathbb{R}^n$ that are defined on all subsets of $\mathbb{R}^n$, but these extensions are not countably additive. For a detailed discussion of the Banach-Tarski paradox and related issues, see [**10**].

The moral of these results is that some subsets of $\mathbb{R}^n$ are too irregular to define their Lebesgue measure in a way that preserves countable additivity (or even finite additivity in $n \geq 3$ dimensions) together with the invariance of the measure under

[1] Solovay (1970) proved that one has to use the axiom of choice to obtain non-Lebesgue measurable sets.

isometries. We will show, however, that such a measure can be defined on a σ-algebra $\mathcal{L}(\mathbb{R}^n)$ of Lebesgue measurable sets which is large enough to include all set of 'practical' importance in analysis. Moreover, as we will see, it is possible to define an isometrically-invariant, countably *sub*-additive outer measure on all subsets of $\mathbb{R}^n$.

There are many ways to construct Lebesgue measure, all of which lead to the same result. We will follow an approach due to Carathéodory, which generalizes to other measures: We first construct an outer measure on all subsets of $\mathbb{R}^n$ by approximating them from the outside by countable unions of rectangles; we then restrict this outer measure to a σ-algebra of measurable subsets on which it is countably additive. This approach is somewhat asymmetrical in that we approximate sets (and their complements) from the outside by elementary sets, but we do not approximate them directly from the inside.

Jones [**5**], Stein and Shakarchi [**8**], and Wheeler and Zygmund [**11**] give detailed introductions to Lebesgue measure on $\mathbb{R}^n$. Cohn [**2**] gives a similar development to the one here, and Evans and Gariepy [**3**] discuss more advanced topics.

2.1. Lebesgue outer measure

We use rectangles as our elementary sets, defined as follows.

Definition 2.1. An n-dimensional, closed rectangle with sides oriented parallel to the coordinate axes, or rectangle for short, is a subset $R \subset \mathbb{R}^n$ of the form

$$R = [a_1, b_1] \times [a_2, b_2] \times \cdots \times [a_n, b_n]$$

where $-\infty < a_i \leq b_i < \infty$ for $i = 1, \ldots, n$. The volume $\mu(R)$ of R is

$$\mu(R) = (b_1 - a_1)(b_2 - a_2) \ldots (b_n - a_n).$$

If $n = 1$ or $n = 2$, the volume of a rectangle is its length or area, respectively. We also consider the empty set to be a rectangle with $\mu(\varnothing) = 0$. We denote the collection of all n-dimensional rectangles by $\mathcal{R}(\mathbb{R}^n)$, or $\mathcal{R}$ when n is understood, and then $R \mapsto \mu(R)$ defines a map

$$\mu : \mathcal{R}(\mathbb{R}^n) \to [0, \infty).$$

The use of this particular class of elementary sets is for convenience. We could equally well use open or half-open rectangles, cubes, balls, or other suitable elementary sets; the result would be the same.

Definition 2.2. The outer Lebesgue measure $\mu^*(E)$ of a subset $E \subset \mathbb{R}^n$, or outer measure for short, is

$$\mu^*(E) = \inf \left\{ \sum_{i=1}^{\infty} \mu(R_i) : E \subset \textstyle\bigcup_{i=1}^{\infty} R_i,\ R_i \in \mathcal{R}(\mathbb{R}^n) \right\} \tag{2.1}$$

where the infimum is taken over all countable collections of rectangles whose union contains E. The map

$$\mu^* : \mathcal{P}(\mathbb{R}^n) \to [0, \infty], \qquad \mu^* : E \mapsto \mu^*(E)$$

is called outer Lebesgue measure.

In this definition, a sum $\sum_{i=1}^{\infty} \mu(R_i)$ and $\mu^*(E)$ may take the value ∞. We do not require that the rectangles R_i are disjoint, so the same volume may contribute to multiple terms in the sum on the right-hand side of (2.1); this does not affect the value of the infimum.

Example 2.3. Let $E = \mathbb{Q} \cap [0,1]$ be the set of rational numbers between 0 and 1. Then E has outer measure zero. To prove this, let $\{q_i : i \in \mathbb{N}\}$ be an enumeration of the points in E. Given $\epsilon > 0$, let R_i be an interval of length $\epsilon/2^i$ which contains q_i. Then $E \subset \bigcup_{i=1}^{\infty} \mu(R_i)$ so

$$0 \leq \mu^*(E) \leq \sum_{i=1}^{\infty} \mu(R_i) = \epsilon.$$

Hence $\mu^*(E) = 0$ since $\epsilon > 0$ is arbitrary. The same argument shows that any countable set has outer measure zero. Note that if we cover E by a *finite* collection of intervals, then the union of the intervals would have to contain $[0,1]$ since E is dense in $[0,1]$ so their lengths sum to at least one.

The previous example illustrates why we need to use countably infinite collections of rectangles, not just finite collections, to define the outer measure.[2] The 'countable ϵ-trick' used in the example appears in various forms throughout measure theory.

Next, we prove that μ^* is an outer measure in the sense of Definition 1.2.

Theorem 2.4. *Lebesgue outer measure μ^* has the following properties.*

(a) $\mu^*(\varnothing) = 0$;
(b) if $E \subset F$, *then* $\mu^*(E) \leq \mu^*(F)$;
(c) if $\{E_i \subset \mathbb{R}^n : i \in \mathbb{N}\}$ *is a countable collection of subsets of* $\mathbb{R}^n$, *then*

$$\mu^*\left(\bigcup_{i=1}^{\infty} E_i\right) \leq \sum_{i=1}^{\infty} \mu^*(E_i).$$

Proof. It follows immediately from Definition 2.2 that $\mu^*(\varnothing) = 0$, since every collection of rectangles covers $\varnothing$, and that $\mu^*(E) \leq \mu^*(F)$ if $E \subset F$ since any cover of F covers E.

The main property to prove is the countable subadditivity of μ^*. If $\mu^*(E_i) = \infty$ for some $i \in \mathbb{N}$, there is nothing to prove, so we may assume that $\mu^*(E_i)$ is finite for every $i \in \mathbb{N}$. If $\epsilon > 0$, there is a countable covering $\{R_{ij} : j \in \mathbb{N}\}$ of E_i by rectangles R_{ij} such that

$$\sum_{j=1}^{\infty} \mu(R_{ij}) \leq \mu^*(E_i) + \frac{\epsilon}{2^i}, \qquad E_i \subset \bigcup_{j=1}^{\infty} R_{ij}.$$

Then $\{R_{ij} : i, j \in \mathbb{N}\}$ is a countable covering of

$$E = \bigcup_{i=1}^{\infty} E_i$$

---

[2] The use of finitely many intervals leads to the notion of the Jordan content of a set, introduced by Peano (1887) and Jordan (1892), which is closely related to the Riemann integral; Borel (1898) and Lebesgue (1902) generalized Jordan's approach to allow for countably many intervals, leading to Lebesgue measure and the Lebesgue integral.

and therefore

$$\mu^*(E) \leq \sum_{i,j=1}^{\infty} \mu(R_{ij}) \leq \sum_{i=1}^{\infty} \left\{ \mu^*(E_i) + \frac{\epsilon}{2^i} \right\} = \sum_{i=1}^{\infty} \mu^*(E_i) + \epsilon.$$

Since $\epsilon > 0$ is arbitrary, it follows that

$$\mu^*(E) \leq \sum_{i=1}^{\infty} \mu^*(E_i)$$

which proves the result. □

2.2. Outer measure of rectangles

In this section, we prove the geometrically obvious, but not entirely trivial, fact that the outer measure of a rectangle is equal to its volume. The main point is to show that the volumes of a countable collection of rectangles that cover a rectangle R cannot sum to less than the volume of R.[3]

We begin with some combinatorial facts about finite covers of rectangles [**8**]. We denote the interior of a rectangle R by R°, and we say that rectangles R, S are almost disjoint if $R^\circ \cap S^\circ = \varnothing$, meaning that they intersect at most along their boundaries. The proofs of the following results are cumbersome to write out in detail (it's easier to draw a picture) but we briefly explain the argument.

Lemma 2.5. *Suppose that*

$$R = I_1 \times I_2 \times \cdots \times I_n$$

is an n-dimensional rectangle where each closed, bounded interval $I_i \subset \mathbb{R}$ is an almost disjoint union of closed, bounded intervals $\{I_{i,j} \subset \mathbb{R} : j = 1, \ldots, N_i\}$,

$$I_i = \bigcup_{j=1}^{N_i} I_{i,j}.$$

Define the rectangles

$$S_{j_1 j_2 \ldots j_n} = I_{1,j_1} \times I_{1,j_2} \times \cdots \times I_{j_n}. \tag{2.2}$$

Then

$$\mu(R) = \sum_{j_1=1}^{N_1} \cdots \sum_{j_n=1}^{N_n} \mu\left(S_{j_1 j_2 \ldots j_n}\right).$$

Proof. Denoting the length of an interval I by $|I|$, using the fact that

$$|I_i| = \sum_{j=1}^{N_i} |I_{i,j}|,$$

[3] As a partial justification of the need to prove this fact, note that it would not be true if we allowed uncountable covers, since we could cover any rectangle by an uncountable collection of points all of whose volumes are zero.

and expanding the resulting product, we get that

$$\begin{aligned}
\mu(R) &= |I_1||I_2|\dots|I_n| \\
&= \left(\sum_{j_1=1}^{N_1}|I_{1,j_1}|\right)\left(\sum_{j_2=1}^{N_2}|I_{2,j_2}|\right)\dots\left(\sum_{j_n=1}^{N_n}|I_{n,j_n}|\right) \\
&= \sum_{j_1=1}^{N_1}\sum_{j_2=1}^{N_2}\dots\sum_{j_n=1}^{N_n}|I_{1,j_1}||I_{2,j_2}|\dots|I_{n,j_n}| \\
&= \sum_{j_1=1}^{N_1}\sum_{j_2=1}^{N_2}\dots\sum_{j_n=1}^{N_n}\mu\left(S_{j_1j_2\dots j_n}\right).
\end{aligned}$$

□

Proposition 2.6. *If a rectangle R is an almost disjoint, finite union of rectangles $\{R_1, R_2, \dots, R_N\}$, then*

$$\mu(R) = \sum_{i=1}^{N}\mu(R_i). \tag{2.3}$$

If R is covered by rectangles $\{R_1, R_2, \dots, R_N\}$, which need not be disjoint, then

$$\mu(R) \le \sum_{i=1}^{N}\mu(R_i). \tag{2.4}$$

Proof. Suppose that

$$R = [a_1, b_1] \times [a_2, b_2] \times \dots \times [a_n, b_n]$$

is an almost disjoint union of the rectangles $\{R_1, R_2, \dots, R_N\}$. Then by 'extending the sides' of the R_i, we may decompose R into an almost disjoint collection of rectangles

$$\{S_{j_1j_2\dots j_n} : 1 \le j_i \le N_i \text{ for } 1 \le i \le n\}$$

that is obtained by taking products of subintervals of partitions of the coordinate intervals $[a_i, b_i]$ into unions of almost disjoint, closed subintervals. Explicitly, we partition $[a_i, b_i]$ into

$$a_i = c_{i,0} \le c_{i,1} \le \dots \le c_{i,N_i} = b_i, \qquad I_{i,j} = [c_{i,j-1}, c_{i,j}].$$

where the $c_{i,j}$ are obtained by ordering the left and right ith coordinates of all faces of rectangles in the collection $\{R_1, R_2, \dots, R_N\}$, and define rectangles $S_{j_1j_2\dots j_n}$ as in (2.2).

Each rectangle R_i in the collection is an almost disjoint union of rectangles $S_{j_1j_2\dots j_n}$, and their union contains all such products exactly once, so by applying Lemma 2.5 to each R_i and summing the results we see that

$$\sum_{i=1}^{N}\mu(R_i) = \sum_{j_1=1}^{N_1}\dots\sum_{j_n=1}^{N_n}\mu\left(S_{j_1j_2\dots j_n}\right).$$

Similarly, R is an almost disjoint union of all the rectangles $S_{j_1j_2\dots j_n}$, so Lemma 2.5 implies that

$$\mu(R) = \sum_{j_1=1}^{N_1}\dots\sum_{j_n=1}^{N_n}\mu\left(S_{j_1j_2\dots j_n}\right),$$

and (2.3) follows.

If a finite collection of rectangles $\{R_1, R_2, \dots, R_N\}$ covers R, then there is a almost disjoint, finite collection of rectangles $\{S_1, S_2, \dots, S_M\}$ such that

$$R = \bigcup_{i=1}^{M} S_i, \qquad \sum_{i=1}^{M} \mu(S_i) \le \sum_{i=1}^{N} \mu(R_i).$$

To obtain the S_i, we replace R_i by the rectangle $R \cap R_i$, and then decompose these possibly non-disjoint rectangles into an almost disjoint, finite collection of sub-rectangles with the same union; we discard 'overlaps' which can only reduce the sum of the volumes. Then, using (2.3), we get

$$\mu(R) = \sum_{i=1}^{M} \mu(S_i) \le \sum_{i=1}^{N} \mu(R_i),$$

which proves (2.4). □

The outer measure of a rectangle is defined in terms of countable covers. We reduce these to finite covers by using the topological properties of $\mathbb{R}^n$.

Proposition 2.7. *If R is a rectangle in $\mathbb{R}^n$, then $\mu^*(R) = \mu(R)$.*

Proof. Since $\{R\}$ covers R, we have $\mu^*(R) \le \mu(R)$, so we only need to prove the reverse inequality.

Suppose that $\{R_i : i \in \mathbb{N}\}$ is a countably infinite collection of rectangles that covers R. By enlarging R_i slightly we may obtain a rectangle S_i whose interior S_i° contains R_i such that

$$\mu(S_i) \le \mu(R_i) + \frac{\epsilon}{2^i}.$$

Then $\{S_i^\circ : i \in \mathbb{N}\}$ is an open cover of the compact set R, so it contains a finite subcover, which we may label as $\{S_1^\circ, S_2^\circ, \dots, S_N^\circ\}$. Then $\{S_1, S_2, \dots, S_N\}$ covers R and, using (2.4), we find that

$$\mu(R) \le \sum_{i=1}^{N} \mu(S_i) \le \sum_{i=1}^{N} \left\{\mu(R_i) + \frac{\epsilon}{2^i}\right\} \le \sum_{i=1}^{\infty} \mu(R_i) + \epsilon.$$

Since $\epsilon > 0$ is arbitrary, we have

$$\mu(R) \le \sum_{i=1}^{\infty} \mu(R_i)$$

and it follows that $\mu(R) \le \mu^*(R)$. □

2.3. Carathéodory measurability

We will obtain Lebesgue measure as the restriction of Lebesgue outer measure to Lebesgue measurable sets. The construction, due to Carathéodory, works for any outer measure, as given in Definition 1.2, so we temporarily consider general outer measures. We will return to Lebesgue measure on $\mathbb{R}^n$ at the end of this section.

The following is the Carathéodory definition of measurability.

Definition 2.8. Let μ^* be an outer measure on a set X. A subset $A \subset X$ is Carathéodory measurable with respect to μ^*, or measurable for short, if

$$\mu^*(E) = \mu^*(E \cap A) + \mu^*(E \cap A^c) \tag{2.5}$$

for every subset $E \subset X$.

We also write $E \cap A^c$ as $E \setminus A$. Thus, a measurable set A splits any set E into disjoint pieces whose outer measures add up to the outer measure of E. Heuristically, this condition means that a set is measurable if it divides other sets in a 'nice' way. The regularity of the set E being divided is not important here.

Since μ^* is subadditive, we always have that

$$\mu^*(E) \leq \mu^*(E \cap A) + \mu^*(E \cap A^c).$$

Thus, in order to prove that $A \subset X$ is measurable, it is sufficient to show that

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

for every $E \subset X$, and then we have equality as in (2.5).

Definition 2.8 is perhaps not the most intuitive way to define the measurability of sets, but it leads directly to the following key result.

Theorem 2.9. *The collection of Carathéodory measurable sets with respect to an outer measure μ^* is a σ-algebra, and the restriction of μ^* to the measurable sets is a measure.*

Proof. It follows immediately from (2.5) that $\varnothing$ is measurable and the complement of a measurable set is measurable, so to prove that the collection of measurable sets is a σ-algebra, we only need to show that it is closed under countable unions. We will prove at the same time that μ^* is countably additive on measurable sets; since $\mu^*(\varnothing) = 0$, this will prove that the restriction of μ^* to the measurable sets is a measure.

First, we prove that the union of measurable sets is measurable. Suppose that A, B are measurable and $E \subset X$. The measurability of A and B implies that

$$\begin{aligned} \mu^*(E) &= \mu^*(E \cap A) + \mu^*(E \cap A^c) \\ &= \mu^*(E \cap A \cap B) + \mu^*(E \cap A \cap B^c) \\ &\qquad + \mu^*(E \cap A^c \cap B) + \mu^*(E \cap A^c \cap B^c). \end{aligned} \tag{2.6}$$

Since $A \cup B = (A \cap B) \cup (A \cap B^c) \cup (A^c \cap B)$ and μ^* is subadditive, we have

$$\mu^*(E \cap (A \cup B)) \leq \mu^*(E \cap A \cap B) + \mu^*(E \cap A \cap B^c) + \mu^*(E \cap A^c \cap B).$$

The use of this inequality and the relation $A^c \cap B^c = (A \cup B)^c$ in (2.6) implies that

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

so $A \cup B$ is measurable.

Moreover, if A is measurable and $A \cap B = \varnothing$, then by taking $E = A \cup B$ in (2.5), we see that

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

Thus, the outer measure of the union of disjoint, measurable sets is the sum of their outer measures. The repeated application of this result implies that the finite union of measurable sets is measurable and μ^* is finitely additive on the collection of measurable sets.

Next, we we want to show that the countable union of measurable sets is measurable. It is sufficient to consider disjoint unions. To see this, note that if

$\{A_i : i \in \mathbb{N}\}$ is a countably infinite collection of measurable sets, then

$$B_j = \bigcup_{i=1}^{j} A_i, \qquad \text{for } j \geq 1$$

form an increasing sequence of measurable sets, and

$$C_j = B_j \setminus B_{j-1} \quad \text{for } j \geq 2, \qquad C_1 = B_1$$

form a disjoint measurable collection of sets. Moreover

$$\bigcup_{i=1}^{\infty} A_i = \bigcup_{j=1}^{\infty} C_j.$$

Suppose that $\{A_i : i \in \mathbb{N}\}$ is a countably infinite, disjoint collection of measurable sets, and define

$$B_j = \bigcup_{i=1}^{j} A_i, \qquad B = \bigcup_{i=1}^{\infty} A_i.$$

Let $E \subset X$. Since A_j is measurable and $B_j = A_j \cup B_{j-1}$ is a disjoint union (for $j \geq 2$),

$$\begin{aligned}
\mu^*(E \cap B_j) &= \mu^*(E \cap B_j \cap A_j) + \mu^*(E \cap B_j \cap A_j^c), \qquad . \\
&= \mu^*(E \cap A_j) + \mu^*(E \cap B_{j-1}).
\end{aligned}$$

Also $\mu^*(E \cap B_1) = \mu^*(E \cap A_1)$. It follows by induction that

$$\mu^*(E \cap B_j) = \sum_{i=1}^{j} \mu^*(E \cap A_i).$$

Since B_j is a finite union of measurable sets, it is measurable, so

$$\mu^*(E) = \mu^*(E \cap B_j) + \mu^*(E \cap B_j^c),$$

and since $B_j^c \supset B^c$, we have

$$\mu^*(E \cap B_j^c) \geq \mu^*(E \cap B^c).$$

It follows that

$$\mu^*(E) \geq \sum_{i=1}^{j} \mu^*(E \cap A_i) + \mu^*(E \cap B^c).$$

Taking the limit of this inequality as $j \to \infty$ and using the subadditivity of μ^*, we get

$$\begin{aligned}
\mu^*(E) &\geq \sum_{i=1}^{\infty} \mu^*(E \cap A_i) + \mu^*(E \cap B^c) \\
&\geq \mu^*\left(\bigcup_{i=1}^{\infty} E \cap A_i\right) + \mu^*(E \cap B^c) \qquad (2.7) \\
&\geq \mu^*(E \cap B) + \mu^*(E \cap B^c) \\
&\geq \mu^*(E).
\end{aligned}$$

Therefore, we must have equality in (2.7), which shows that $B = \bigcup_{i=1}^\infty A_i$ is measurable. Moreover,

$$\mu^*\left(\bigcup_{i=1}^\infty E \cap A_i\right) = \sum_{i=1}^\infty \mu^*(E \cap A_i),$$

so taking $E = X$, we see that μ^* is countably additive on the σ-algebra of measurable sets. □

Returning to Lebesgue measure on $\mathbb{R}^n$, the preceding theorem shows that we get a measure on $\mathbb{R}^n$ by restricting Lebesgue outer measure to its Carathéodory-measurable sets, which are the Lebesgue measurable subsets of $\mathbb{R}^n$.

Definition 2.10. A subset $A \subset \mathbb{R}^n$ is Lebesgue measurable if

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

for every subset $E \subset \mathbb{R}^n$. If $\mathcal{L}(\mathbb{R}^n)$ denotes the σ-algebra of Lebesgue measurable sets, the restriction of Lebesgue outer measure μ^* to the Lebesgue measurable sets

$$\mu : \mathcal{L}(\mathbb{R}^n) \to [0, \infty], \qquad \mu = \mu^*|_{\mathcal{L}(\mathbb{R}^n)}$$

is called Lebesgue measure.

From Proposition 2.7, this notation is consistent with our previous use of μ to denote the volume of a rectangle. If $E \subset \mathbb{R}^n$ is any measurable subset of $\mathbb{R}^n$, then we define Lebesgue measure on E by restricting Lebesgue measure on $\mathbb{R}^n$ to E, as in Definition 1.10, and denote the corresponding σ-algebra of Lebesgue measurable subsets of E by $\mathcal{L}(E)$.

Next, we prove that all rectangles are measurable; this implies that $\mathcal{L}(\mathbb{R}^n)$ is a 'large' collection of subsets of $\mathbb{R}^n$. Not all subsets of $\mathbb{R}^n$ are Lebesgue measurable, however; *e.g.* see Example 2.17 below.

Proposition 2.11. *Every rectangle is Lebesgue measurable.*

Proof. Let R be an n-dimensional rectangle and $E \subset \mathbb{R}^n$. Given $\epsilon > 0$, there is a cover $\{R_i : i \in \mathbb{N}\}$ of E by rectangles R_i such that

$$\mu^*(E) + \epsilon \geq \sum_{i=1}^\infty \mu(R_i).$$

We can decompose R_i into an almost disjoint, finite union of rectangles

$$\{\tilde{R}_i, S_{i,1}, \dots, S_{i,N}\}$$

such that

$$R_i = \tilde{R}_i + \bigcup_{j=1}^N S_{i,j}, \qquad \tilde{R}_i = R_i \cap R \subset R, \quad S_{i,j} \subset \overline{R^c}.$$

From (2.3),

$$\mu(R_i) = \mu(\tilde{R}_i) + \sum_{j=1}^N \mu(S_{i,j}).$$

Using this result in the previous sum, relabeling the $S_{i,j}$ as S_i, and rearranging the resulting sum, we get that

$$\mu^*(E) + \epsilon \geq \sum_{i=1}^\infty \mu(\tilde{R}_i) + \sum_{i=1}^\infty \mu(S_i).$$

Since the rectangles $\{\tilde{R}_i : i \in \mathbb{N}\}$ cover $E \cap R$ and the rectangles $\{S_i : i \in \mathbb{N}\}$ cover $E \cap R^c$, we have

$$\mu^*(E \cap R) \leq \sum_{i=1}^{\infty} \mu(\tilde{R}_i), \qquad \mu^*(E \cap R^c) \leq \sum_{i=1}^{\infty} \mu(S_i).$$

Hence,

$$\mu^*(E) + \epsilon \geq \mu^*(E \cap R) + \mu^*(E \cap R^c).$$

Since $\epsilon > 0$ is arbitrary, it follows that

$$\mu^*(E) \geq \mu^*(E \cap R) + \mu^*(E \cap R^c),$$

which proves the result. □

An open rectangle R° is a union of an increasing sequence of closed rectangles whose volumes approach $\mu(R)$; for example

$$\begin{aligned}&(a_1, b_1) \times (a_2, b_2) \times \cdots \times (a_n, b_n)\\ &\quad = \bigcup_{k=1}^{\infty} [a_1 + \frac{1}{k}, b_1 - \frac{1}{k}] \times [a_2 + \frac{1}{k}, b_2 - \frac{1}{k}] \times \cdots \times [a_n + \frac{1}{k}, b_n - \frac{1}{k}].\end{aligned}$$

Thus, R° is measurable and, from Proposition 1.12,

$$\mu(R^\circ) = \mu(R).$$

Moreover if $\partial R = R \setminus R^\circ$ denotes the boundary of R, then

$$\mu(\partial R) = \mu(R) - \mu(R^\circ) = 0.$$

2.4. Null sets and completeness

Sets of measure zero play a particularly important role in measure theory and integration. First, we show that all sets with outer Lebesgue measure zero are Lebesgue measurable.

Proposition 2.12. *If $N \subset \mathbb{R}^n$ and $\mu^*(N) = 0$, then N is Lebesgue measurable, and the measure space $(\mathbb{R}^n, \mathcal{L}(\mathbb{R}^n), \mu)$ is complete.*

Proof. If $N \subset \mathbb{R}^n$ has outer Lebesgue measure zero and $E \subset \mathbb{R}^n$, then

$$0 \leq \mu^*(E \cap N) \leq \mu^*(N) = 0,$$

so $\mu^*(E \cap N) = 0$. Therefore, since $E \supset E \cap N^c$,

$$\mu^*(E) \geq \mu^*(E \cap N^c) = \mu^*(E \cap N) + \mu^*(E \cap N^c),$$

which shows that N is measurable. If N is a measurable set with $\mu(N) = 0$ and $M \subset N$, then $\mu^*(M) = 0$, since $\mu^*(M) \leq \mu(N)$. Therefore M is measurable and $(\mathbb{R}^n, \mathcal{L}(\mathbb{R}^n), \mu)$ is complete. □

In view of the importance of sets of measure zero, we formulate their definition explicitly.

Definition 2.13. A subset $N \subset \mathbb{R}^n$ has Lebesgue measure zero if for every $\epsilon > 0$ there exists a countable collection of rectangles $\{R_i : i \in \mathbb{N}\}$ such that

$$N \subset \bigcup_{i=1}^{\infty} R_i, \qquad \sum_{i=1}^{\infty} \mu(R_i) < \epsilon.$$

The argument in Example 2.3 shows that every countable set has Lebesgue measure zero, but sets of measure zero may be uncountable; in fact the fine structure of sets of measure zero is, in general, very intricate.

Example 2.14. The standard Cantor set, obtained by removing 'middle thirds' from $[0,1]$, is an uncountable set of zero one-dimensional Lebesgue measure.

Example 2.15. The x-axis in $\mathbb{R}^2$

$$A = \left\{(x,0) \in \mathbb{R}^2 : x \in \mathbb{R}\right\}$$

has zero two-dimensional Lebesgue measure. More generally, any linear subspace of $\mathbb{R}^n$ with dimension strictly less than n has zero n-dimensional Lebesgue measure.

2.5. Translational invariance

An important geometric property of Lebesgue measure is its translational invariance. If $A \subset \mathbb{R}^n$ and $h \in \mathbb{R}^n$, let

$$A + h = \{x + h : x \in A\}$$

denote the translation of A by h.

Proposition 2.16. *If $A \subset \mathbb{R}^n$ and $h \in \mathbb{R}^n$, then*

$$\mu^*(A + h) = \mu^*(A),$$

and $A + h$ is measurable if and only if A is measurable.

Proof. The invariance of outer measure μ^* result is an immediate consequence of the definition, since $\{R_i + h : i \in \mathbb{N}\}$ is a cover of $A + h$ if and only if $\{R_i : i \in \mathbb{N}\}$ is a cover of A, and $\mu(R + h) = \mu(R)$ for every rectangle R. Moreover, the Carathéodory definition of measurability is invariant under translations since

$$(E + h) \cap (A + h) = (E \cap A) + h.$$

□

The space $\mathbb{R}^n$ is a locally compact topological (abelian) group with respect to translation, which is a continuous operation. More generally, there exists a (left or right) translation-invariant measure, called Haar measure, on any locally compact topological group; this measure is unique up to a scalar factor.

The following is the standard example of a non-Lebesgue measurable set, due to Vitali (1905).

Example 2.17. Define an equivalence relation $\sim$ on $\mathbb{R}$ by $x \sim y$ if $x - y \in \mathbb{Q}$. This relation has uncountably many equivalence classes, each of which contains a countably infinite number of points and is dense in $\mathbb{R}$. Let $E \subset [0,1]$ be a set that contains exactly one element from each equivalence class, so that $\mathbb{R}$ is the disjoint union of the countable collection of rational translates of E. Then we claim that E is not Lebesgue measurable.

To show this, suppose for contradiction that E is measurable. Let $\{q_i : i \in \mathbb{N}\}$ be an enumeration of the rational numbers in the interval $[-1,1]$ and let $E_i = E + q_i$ denote the translation of E by q_i. Then the sets E_i are disjoint and

$$[0,1] \subset \bigcup_{i=1}^{\infty} E_i \subset [-1,2].$$

The translational invariance of Lebesgue measure implies that each E_i is measurable with $\mu(E_i) = \mu(E)$, and the countable additivity of Lebesgue measure implies that

$$1 \leq \sum_{i=1}^{\infty} \mu(E_i) \leq 3.$$

But this is impossible, since $\sum_{i=1}^{\infty} \mu(E_i)$ is either 0 or ∞, depending on whether if $\mu(E) = 0$ or $\mu(E) > 0$.

The above example is geometrically simpler on the circle $\mathbb{T} = \mathbb{R}/\mathbb{Z}$. When reduced modulo one, the sets $\{E_i : i \in \mathbb{N}\}$ partition $\mathbb{T}$ into a countable union of disjoint sets which are translations of each other. If the sets were measurable, their measures would be equal so they must sum to 0 or ∞, but the measure of $\mathbb{T}$ is one.

2.6. Borel sets

The relationship between measure and topology is not a simple one. In this section, we show that all open and closed sets in $\mathbb{R}^n$, and therefore all Borel sets (*i.e.* sets that belong to the σ-algebra generated by the open sets), are Lebesgue measurable.

Let $\mathcal{T}(\mathbb{R}^n) \subset \mathcal{P}(\mathbb{R}^n)$ denote the standard metric topology on $\mathbb{R}^n$ consisting of all open sets. That is, $G \subset \mathbb{R}^n$ belongs to $\mathcal{T}(\mathbb{R}^n)$ if for every $x \in G$ there exists $r > 0$ such that $B_r(x) \subset G$, where

$$B_r(x) = \{y \in \mathbb{R}^n : |x - y| < r\}$$

is the open ball of radius r centered at $x \in \mathbb{R}^n$ and $|\cdot|$ denotes the Euclidean norm.

Definition 2.18. The Borel σ-algebra $\mathcal{B}(\mathbb{R}^n)$ on $\mathbb{R}^n$ is the σ-algebra generated by the open sets, $\mathcal{B}(\mathbb{R}^n) = \sigma(\mathcal{T}(\mathbb{R}^n))$. A set that belongs to the Borel σ-algebra is called a Borel set.

Since σ-algebras are closed under complementation, the Borel σ-algebra is also generated by the closed sets in $\mathbb{R}^n$. Moreover, since $\mathbb{R}^n$ is σ-compact (*i.e.* it is a countable union of compact sets) its Borel σ-algebra is generated by the compact sets.

Remark 2.19. This definition is not constructive, since we start with the power set of $\mathbb{R}^n$ and narrow it down until we obtain the smallest σ-algebra that contains the open sets. It is surprisingly complicated to obtain $\mathcal{B}(\mathbb{R}^n)$ by starting from the open or closed sets and taking successive complements, countable unions, and countable intersections. These operations give sequences of collections of sets in $\mathbb{R}^n$

$$G \subset G_\delta \subset G_{\delta\sigma} \subset G_{\delta\sigma\delta} \subset \dots, \qquad F \subset F_\sigma \subset F_{\sigma\delta} \subset F_{\delta\sigma\delta} \subset \dots, \tag{2.8}$$

where G denotes the open sets, F the closed sets, σ the operation of countable unions, and δ the operation of countable intersections. These collections contain each other; for example, $F_\sigma \supset G$ and $G_\delta \supset F$. This process, however, has to be repeated up to the first uncountable ordinal before we obtain $\mathcal{B}(\mathbb{R}^n)$. This is because if, for example, $\{A_i : i \in \mathbb{N}\}$ is a countable family of sets such that

$$A_1 \in G_\delta \setminus G, \quad A_2 \in G_{\delta\sigma} \setminus G_\delta, \quad A_3 \in G_{\delta\sigma\delta} \setminus G_{\delta\sigma}, \dots$$

and so on, then there is no guarantee that $\bigcup_{i=1}^{\infty} A_i$ or $\bigcap_{i=1}^{\infty} A_i$ belongs to any of the previously constructed families. In general, one only knows that they belong to the $\omega + 1$ iterates $G_{\delta\sigma\delta\dots\sigma}$ or $G_{\delta\sigma\delta\dots\delta}$, respectively, where ω is the ordinal number of $\mathbb{N}$. A similar argument shows that in order to obtain a family which is closed under countable intersections or unions, one has to continue this process until one has constructed an uncountable number of families.

To show that open sets are measurable, we will represent them as countable unions of rectangles. Every open set in $\mathbb{R}$ is a countable disjoint union of open intervals (one-dimensional open rectangles). When $n \geq 2$, it is not true that every open set in $\mathbb{R}^n$ is a countable disjoint union of open rectangles, but we have the following substitute.

Proposition 2.20. *Every open set in $\mathbb{R}^n$ is a countable union of almost disjoint rectangles.*

Proof. Let $G \subset \mathbb{R}^n$ be open. We construct a family of cubes (rectangles of equal sides) as follows. First, we bisect $\mathbb{R}^n$ into almost disjoint cubes $\{Q_i : i \in \mathbb{N}\}$ of side one with integer coordinates. If $Q_i \subset G$, we include Q_i in the family, and if Q_i is disjoint from G, we exclude it. Otherwise, we bisect the sides of Q_i to obtain 2^n almost disjoint cubes of side one-half and repeat the procedure. Iterating this process arbitrarily many times, we obtain a countable family of almost disjoint cubes.

The union of the cubes in this family is contained in G, since we only include cubes that are contained in G. Conversely, if $x \in G$, then since G is open some sufficiently small cube in the bisection procedure that contains x is entirely contained in G, and the largest such cube is included in the family. Hence the union of the family contains G, and is therefore equal to G. □

In fact, the proof shows that every open set is an almost disjoint union of dyadic cubes.

Proposition 2.21. *The Borel algebra $\mathcal{B}(\mathbb{R}^n)$ is generated by the collection of rectangles $\mathcal{R}(\mathbb{R}^n)$. Every Borel set is Lebesgue measurable.*

Proof. Since $\mathcal{R}$ is a subset of the closed sets, we have $\sigma(\mathcal{R}) \subset \mathcal{B}$. Conversely, by the previous proposition, $\sigma(\mathcal{R}) \supset \mathcal{T}$, so $\sigma(\mathcal{R}) \supset \sigma(\mathcal{T}) = \mathcal{B}$, and therefore $\mathcal{B} = \sigma(\mathcal{R})$. From Proposition 2.11, we have $\mathcal{R} \subset \mathcal{L}$. Since $\mathcal{L}$ is a σ-algebra, it follows that $\sigma(\mathcal{R}) \subset \mathcal{L}$, so $\mathcal{B} \subset \mathcal{L}$. □

Note that if

$$G = \bigcup_{i=1}^{\infty} R_i$$

is a decomposition of an open set G into an almost disjoint union of closed rectangles, then

$$G \supset \bigcup_{i=1}^{\infty} R_i^\circ$$

is a disjoint union, and therefore

$$\sum_{i=1}^{\infty} \mu(R_i^\circ) \leq \mu(G) \leq \sum_{i=1}^{\infty} \mu(R_i).$$

Since $\mu(R_i^\circ) = \mu(R_i)$, it follows that

$$\mu(G) = \sum_{i=1}^{\infty} \mu(R_i)$$

for any such decomposition and that the sum is independent of the way in which G is decomposed into almost disjoint rectangles.

The Borel σ-algebra $\mathcal{B}$ is not complete and is strictly smaller than the Lebesgue σ-algebra $\mathcal{L}$. In fact, one can show that the cardinality of $\mathcal{B}$ is equal to the cardinality $\mathfrak{c}$ of the real numbers, whereas the cardinality of $\mathcal{L}$ is equal to $2^{\mathfrak{c}}$. For example, the Cantor set is a set of measure zero with the same cardinality as $\mathbb{R}$ and every subset of the Cantor set is Lebesgue measurable.

We can obtain examples of sets that are Lebesgue measurable but not Borel measurable by considering subsets of sets of measure zero. In the following example of such a set in $\mathbb{R}$, we use some properties of measurable functions which will be proved later.

Example 2.22. Let $f : [0,1] \to [0,1]$ denote the standard Cantor function and define $g : [0,1] \to [0,1]$ by

$$g(y) = \inf \{x \in [0,1] : f(x) = y\}.$$

Then g is an increasing, one-to-one function that maps $[0,1]$ onto the Cantor set C. Since g is increasing it is Borel measurable, and the inverse image of a Borel set under g is Borel. Let $E \subset [0,1]$ be a non-Lebesgue measurable set. Then $F = g(E) \subset C$ is Lebesgue measurable, since it is a subset of a set of measure zero, but F is not Borel measurable, since if it was $E = g^{-1}(F)$ would be Borel.

Other examples of Lebesgue measurable sets that are not Borel sets arise from the theory of product measures in $\mathbb{R}^n$ for $n \geq 2$. For example, let $N = E \times \{0\} \subset \mathbb{R}^2$ where $E \subset \mathbb{R}$ is a non-Lebesgue measurable set in $\mathbb{R}$. Then N is a subset of the x-axis, which has two-dimensional Lebesgue measure zero, so N belongs to $\mathcal{L}(\mathbb{R}^2)$ since Lebesgue measure is complete. One can show, however, that if a set belongs to $\mathcal{B}(\mathbb{R}^2)$ then every section with fixed x or y coordinate, belongs to $\mathcal{B}(\mathbb{R})$; thus, N cannot belong to $\mathcal{B}(\mathbb{R}^2)$ since the $y = 0$ section E is not Borel.

As we show below, $\mathcal{L}(\mathbb{R}^n)$ is the completion of $\mathcal{B}(\mathbb{R}^n)$ with respect to Lebesgue measure, meaning that we get all Lebesgue measurable sets by adjoining all subsets of Borel sets of measure zero to the Borel σ-algebra and taking unions of such sets.

2.7. Borel regularity

Regularity properties of measures refer to the possibility of approximating in measure one class of sets (for example, nonmeasurable sets) by another class of sets (for example, measurable sets). Lebesgue measure is Borel regular in the sense that Lebesgue measurable sets can be approximated in measure from the outside by open sets and from the inside by closed sets, and they can be approximated by Borel sets up to sets of measure zero. Moreover, there is a simple criterion for Lebesgue measurability in terms of open and closed sets.

The following theorem expresses a fundamental approximation property of Lebesgue measurable sets by open and compact sets. Equations (2.9) and (2.10) are called outer and inner regularity, respectively.

Theorem 2.23. *If $A \subset \mathbb{R}^n$, then*

$$\mu^*(A) = \inf \{\mu(G) : A \subset G,\ G \text{ open}\}, \tag{2.9}$$

and if A is Lebesgue measurable, then

$$\mu(A) = \sup \{\mu(K) : K \subset A,\ K \text{ compact}\}. \tag{2.10}$$

PROOF. First, we prove (2.9). The result is immediate if $\mu^*(A) = \infty$, so we suppose that $\mu^*(A)$ is finite. If $A \subset G$, then $\mu^*(A) \le \mu(G)$, so

$$\mu^*(A) \le \inf\{\mu(G) : A \subset G, G \text{ open}\},$$

and we just need to prove the reverse inequality,

$$\mu^*(A) \ge \inf\{\mu(G) : A \subset G, G \text{ open}\}. \tag{2.11}$$

Let $\epsilon > 0$. There is a cover $\{R_i : i \in \mathbb{N}\}$ of A by rectangles R_i such that

$$\sum_{i=1}^{\infty} \mu(R_i) \le \mu^*(A) + \frac{\epsilon}{2}.$$

Let S_i be an rectangle whose interior S_i° contains R_i such that

$$\mu(S_i) \le \mu(R_i) + \frac{\epsilon}{2^{i+1}}.$$

Then the collection of open rectangles $\{S_i^\circ : i \in \mathbb{N}\}$ covers A and

$$G = \bigcup_{i=1}^{\infty} S_i^\circ$$

is an open set that contains A. Moreover, since $\{S_i : i \in \mathbb{N}\}$ covers G,

$$\mu(G) \le \sum_{i=1}^{\infty} \mu(S_i) \le \sum_{i=1}^{\infty} \mu(R_i) + \frac{\epsilon}{2},$$

and therefore

$$\mu(G) \le \mu^*(A) + \epsilon. \tag{2.12}$$

It follows that

$$\inf\{\mu(G) : A \subset G, G \text{ open}\} \le \mu^*(A) + \epsilon,$$

which proves (2.11) since $\epsilon > 0$ is arbitrary.

Next, we prove (2.10). If $K \subset A$, then $\mu(K) \le \mu(A)$, so

$$\sup\{\mu(K) : K \subset A, K \text{ compact}\} \le \mu(A).$$

Therefore, we just need to prove the reverse inequality,

$$\mu(A) \le \sup\{\mu(K) : K \subset A, K \text{ compact}\}. \tag{2.13}$$

To do this, we apply the previous result to A^c and use the measurability of A.

First, suppose that A is a bounded measurable set, in which case $\mu(A) < \infty$. Let $F \subset \mathbb{R}^n$ be a compact set that contains A. By the preceding result, for any $\epsilon > 0$, there is an open set $G \supset F \setminus A$ such that

$$\mu(G) \le \mu(F \setminus A) + \epsilon.$$

Then $K = F \setminus G$ is a compact set such that $K \subset A$. Moreover, $F \subset K \cup G$ and $F = A \cup (F \setminus A)$, so

$$\mu(F) \le \mu(K) + \mu(G), \qquad \mu(F) = \mu(A) + \mu(F \setminus A).$$

It follows that

$$\begin{aligned} \mu(A) &= \mu(F) - \mu(F \setminus A) \\ &\le \mu(F) - \mu(G) + \epsilon \\ &\le \mu(K) + \epsilon, \end{aligned}$$

which implies (2.13) and proves the result for bounded, measurable sets.

Now suppose that A is an unbounded measurable set, and define

$$A_k = \{x \in A : |x| \le k\}\,. \tag{2.14}$$

Then $\{A_k : k \in \mathbb{N}\}$ is an increasing sequence of bounded measurable sets whose union is A, so

$$\mu(A_k) \uparrow \mu(A) \qquad \text{as } k \to \infty. \tag{2.15}$$

If $\mu(A) = \infty$, then $\mu(A_k) \to \infty$ as $k \to \infty$. By the previous result, we can find a compact set $K_k \subset A_k \subset A$ such that

$$\mu(K_k) + 1 \ge \mu(A_k)$$

so that $\mu(K_k) \to \infty$. Therefore

$$\sup\{\mu(K) : K \subset A,\, K \text{ compact}\} = \infty,$$

which proves the result in this case.

Finally, suppose that A is unbounded and $\mu(A) < \infty$. From (2.15), for any $\epsilon > 0$ we can choose $k \in \mathbb{N}$ such that

$$\mu(A) \le \mu(A_k) + \frac{\epsilon}{2}.$$

Moreover, since A_k is bounded, there is a compact set $K \subset A_k$ such that

$$\mu(A_k) \le \mu(K) + \frac{\epsilon}{2}.$$

Therefore, for every $\epsilon > 0$ there is a compact set $K \subset A$ such that

$$\mu(A) \le \mu(K) + \epsilon,$$

which gives (2.13), and completes the proof. □

It follows that we may determine the Lebesgue measure of a measurable set in terms of the Lebesgue measure of open or compact sets by approximating the set from the outside by open sets or from the inside by compact sets.

The outer approximation in (2.9) does not require that A is measurable. Thus, for any set $A \subset \mathbb{R}^n$, given $\epsilon > 0$, we can find an open set $G \supset A$ such that $\mu(G) - \mu^*(A) < \epsilon$. If A is measurable, we can strengthen this condition to get that $\mu^*(G \setminus A) < \epsilon$; in fact, this gives a necessary and sufficient condition for measurability.

Theorem 2.24. *A subset $A \subset \mathbb{R}^n$ is Lebesgue measurable if and only if for every $\epsilon > 0$ there is an open set $G \supset A$ such that*

$$\mu^*(G \setminus A) < \epsilon. \tag{2.16}$$

PROOF. First we assume that A is measurable and show that it satisfies the condition given in the theorem.

Suppose that $\mu(A) < \infty$ and let $\epsilon > 0$. From (2.12) there is an open set $G \supset A$ such that $\mu(G) < \mu^*(A) + \epsilon$. Then, since A is measurable,

$$\mu^*(G \setminus A) = \mu^*(G) - \mu^*(G \cap A) = \mu(G) - \mu^*(A) < \epsilon,$$

which proves the result when A has finite measure.

If $\mu(A) = \infty$, define $A_k \subset A$ as in (2.14), and let $\epsilon > 0$. Since A_k is measurable with finite measure, the argument above shows that for each $k \in \mathbb{N}$, there is an open set $G_k \supset A_k$ such that

$$\mu(G_k \setminus A_k) < \frac{\epsilon}{2^k}.$$

Then $G = \bigcup_{k=1}^{\infty} G_k$ is an open set that contains A, and

$$\mu^*(G \setminus A) = \mu^* \left(\bigcup_{k=1}^{\infty} G_k \setminus A \right) \leq \sum_{k=1}^{\infty} \mu^*(G_k \setminus A) \leq \sum_{k=1}^{\infty} \mu^*(G_k \setminus A_k) < \epsilon.$$

Conversely, suppose that $A \subset \mathbb{R}^n$ satisfies the condition in the theorem. Let $\epsilon > 0$, and choose an open set $G \supset A$ such that $\mu^*(G \setminus A) < \epsilon$. If $E \subset \mathbb{R}^n$, we have

$$E \cap A^c = (E \cap G^c) \cup (E \cap (G \setminus A)).$$

Hence, by the subadditivity and monotonicity of μ^* and the measurability of G,

$$\begin{aligned} \mu^*(E \cap A) + \mu^*(E \cap A^c) &\leq \mu^*(E \cap A) + \mu^*(E \cap G^c) + \mu^*(E \cap (G \setminus A)) \\ &\leq \mu^*(E \cap G) + \mu^*(E \cap G^c) + \mu^*(G \setminus A) \\ &< \mu^*(E) + \epsilon. \end{aligned}$$

Since $\epsilon > 0$ is arbitrary, it follows that

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

which proves that A is measurable. □

This theorem states that a set is Lebesgue measurable if and only if it can be approximated from the outside by an open set in such a way that the difference has arbitrarily small outer Lebesgue measure. This condition can be adopted as the definition of Lebesgue measurable sets, rather than the Carathéodory definition which we have used *c.f.* [**5, 8, 11**].

The following theorem gives another characterization of Lebesgue measurable sets, as ones that can be 'squeezed' between open and closed sets.

Theorem 2.25. *A subset $A \subset \mathbb{R}^n$ is Lebesgue measurable if and only if for every $\epsilon > 0$ there is an open set G and a closed set F such that $G \supset A \supset F$ and*

$$\mu(G \setminus F) < \epsilon. \tag{2.17}$$

If $\mu(A) < \infty$, then F may be chosen to be compact.

Proof. If A satisfies the condition in the theorem, then it follows from the monotonicity of μ^* that $\mu^*(G \setminus A) \leq \mu(G \setminus F) < \epsilon$, so A is measurable by Theorem 2.24.

Conversely, if A is measurable then A^c is measurable, and by Theorem 2.24 given $\epsilon > 0$, there are open sets $G \supset A$ and $H \supset A^c$ such that

$$\mu^*(G \setminus A) < \frac{\epsilon}{2}, \qquad \mu^*(H \setminus A^c) < \frac{\epsilon}{2}.$$

Then, defining the closed set $F = H^c$, we have $G \supset A \supset F$ and

$$\mu(G \setminus F) \leq \mu^*(G \setminus A) + \mu^*(A \setminus F) = \mu^*(G \setminus A) + \mu^*(H \setminus A^c) < \epsilon.$$

Finally, suppose that $\mu(A) < \infty$ and let $\epsilon > 0$. From Theorem 2.23, since A is measurable, there is a compact set $K \subset A$ such that $\mu(A) < \mu(K) + \epsilon/2$ and

$$\mu(A \setminus K) = \mu(A) - \mu(K) < \frac{\epsilon}{2}.$$

As before, from Theorem 2.24 there is an open set $G \supset A$ such that

$$\mu(G) < \mu(A) + \epsilon/2.$$

It follows that $G \supset A \supset K$ and

$$\mu(G \setminus K) = \mu(G \setminus A) + \mu(A \setminus K) < \epsilon,$$

which shows that we may take $F = K$ compact when A has finite measure. $\square$

From the previous results, we can approximate measurable sets by open or closed sets, up to sets of arbitrarily small but, in general, nonzero measure. By taking countable intersections of open sets or countable unions of closed sets, we can approximate measurable sets by Borel sets, up to sets of measure zero

Definition 2.26. The collection of sets in $\mathbb{R}^n$ that are countable intersections of open sets is denoted by $G_\delta(\mathbb{R}^n)$, and the collection of sets in $\mathbb{R}^n$ that are countable unions of closed sets is denoted by $F_\sigma(\mathbb{R}^n)$.

G_δ and F_σ sets are Borel. Thus, it follows from the next result that every Lebesgue measurable set can be approximated up to a set of measure zero by a Borel set. This is the Borel regularity of Lebesgue measure.

Theorem 2.27. *Suppose that $A \subset \mathbb{R}^n$ is Lebesgue measurable. Then there exist sets $G \in G_\delta(\mathbb{R}^n)$ and $F \in F_\sigma(\mathbb{R}^n)$ such that*

$$G \supset A \supset F, \qquad \mu(G \setminus A) = \mu(A \setminus F) = 0.$$

Proof. For each $k \in \mathbb{N}$, choose an open set G_k and a closed set F_k such that $G_k \supset A \supset F_k$ and

$$\mu(G_k \setminus F_k) \leq \frac{1}{k}$$

Then

$$G = \bigcap_{k=1}^{\infty} G_k, \qquad F = \bigcup_{k=1}^{\infty} F_k$$

are G_δ and F_σ sets with the required properties. $\square$

In particular, since any measurable set can be approximated up to a set of measure zero by a G_δ or an F_σ, the complexity of the transfinite construction of general Borel sets illustrated in (2.8) is 'hidden' inside sets of Lebesgue measure zero.

As a corollary of this result, we get that the Lebesgue σ-algebra is the completion of the Borel σ-algebra with respect to Lebesgue measure.

Theorem 2.28. *The Lebesgue σ-algebra $\mathcal{L}(\mathbb{R}^n)$ is the completion of the Borel σ-algebra $\mathcal{B}(\mathbb{R}^n)$.*

Proof. Lebesgue measure is complete from Proposition 2.12. By the previous theorem, if $A \subset \mathbb{R}^n$ is Lebesgue measurable, then there is a F_σ set $F \subset A$ such that $M = A \setminus F$ has Lebesgue measure zero. It follows by the approximation theorem that there is a Borel set $N \in G_\delta$ with $\mu(N) = 0$ and $M \subset N$. Thus, $A = F \cup M$ where $F \in \mathcal{B}$ and $M \subset N \in \mathcal{B}$ with $\mu(N) = 0$, which proves that $\mathcal{L}(\mathbb{R}^n)$ is the completion of $\mathcal{B}(\mathbb{R}^n)$ as given in Theorem 1.15. $\square$

2.8. Linear transformations

The definition of Lebesgue measure is not rotationally invariant, since we used rectangles whose sides are parallel to the coordinate axes. In this section, we show that the resulting measure does not, in fact, depend upon the direction of the coordinate axes and is invariant under orthogonal transformations. We also show that Lebesgue measure transforms under a linear map by a factor equal to the absolute value of the determinant of the map.

As before, we use μ^* to denote Lebesgue outer measure defined using rectangles whose sides are parallel to the coordinate axes; a set is Lebesgue measurable if it satisfies the Carathéodory criterion (2.8) with respect to this outer measure. If $T : \mathbb{R}^n \to \mathbb{R}^n$ is a linear map and $E \subset \mathbb{R}^n$, we denote the image of E under T by

$$TE = \{Tx \in \mathbb{R}^n : x \in E\}.$$

First, we consider the Lebesgue measure of rectangles whose sides are not parallel to the coordinate axes. We use a tilde to denote such rectangles by $\tilde{R}$; we denote closed rectangles whose sides are parallel to the coordinate axes by R as before. We refer to $\tilde{R}$ and R as oblique and parallel rectangles, respectively. We denote the volume of a rectangle $\tilde{R}$ by $v(\tilde{R})$, *i.e.* the product of the lengths of its sides, to avoid confusion with its Lebesgue measure $\mu(\tilde{R})$. We know that $\mu(R) = v(R)$ for parallel rectangles, and that $\tilde{R}$ is measurable since it is closed, but we have not yet shown that $\mu(\tilde{R}) = v(\tilde{R})$ for oblique rectangles.

More explicitly, we regard $\mathbb{R}^n$ as a Euclidean space equipped with the standard inner product,

$$(x, y) = \sum_{i=1}^{n} x_i y_i, \qquad x = (x_1, x_2, \dots, x_n), \quad y = (y_1, y_2, \dots, y_n).$$

If $\{e_1, e_2, \dots, e_n\}$ is the standard orthonormal basis of $\mathbb{R}^n$,

$$e_1 = (1, 0, \dots, 0), \quad e_2 = (0, 1, \dots, 0), \dots e_n = (0, 0, \dots, 1),$$

and $\{\tilde{e}_1, \tilde{e}_2, \dots, \tilde{e}_n\}$ is another orthonormal basis, then we use R to denote rectangles whose sides are parallel to $\{e_i\}$ and $\tilde{R}$ to denote rectangles whose sides are parallel to $\{\tilde{e}_i\}$. The linear map $Q : \mathbb{R}^n \to \mathbb{R}^n$ defined by $Qe_i = \tilde{e}_i$ is orthogonal, meaning that $Q^T = Q^{-1}$ and

$$(Qx, Qy) = (x, y) \qquad \text{for all } x, y \in \mathbb{R}^n.$$

Since Q preserves lengths and angles, it maps a rectangle R to a rectangle $\tilde{R} = QR$ such that $v(\tilde{R}) = v(R)$.

We will use the following lemma.

Lemma 2.29. *If an oblique rectangle $\tilde{R}$ contains a finite almost disjoint collection of parallel rectangles $\{R_1, R_2, \dots, R_N\}$ then*

$$\sum_{i=1}^{N} v(R_i) \leq v(\tilde{R}).$$

This result is geometrically obvious, but a formal proof seems to require a fuller discussion of the volume function on elementary geometrical sets, which is included in the theory of valuations in convex geometry. We omit the details.

Proposition 2.30. *If $\tilde{R}$ is an oblique rectangle, then given any $\epsilon > 0$ there is a collection of parallel rectangles $\{R_i : i \in \mathbb{N}\}$ that covers $\tilde{R}$ and satisfies*

$$\sum_{i=1}^{\infty} v(R_i) \leq v(\tilde{R}) + \epsilon.$$

PROOF. Let $\tilde{S}$ be an oblique rectangle that contains $\tilde{R}$ in its interior such that

$$v(\tilde{S}) \leq v(\tilde{R}) + \epsilon.$$

Then, from Proposition 2.20, we may decompose the interior of S into an almost disjoint union of parallel rectangles

$$\tilde{S}^{\circ} = \bigcup_{i=1}^{\infty} R_i.$$

It follows from the previous lemma that for every $N \in \mathbb{N}$

$$\sum_{i=1}^{N} v(R_i) \leq v(\tilde{S}),$$

which implies that

$$\sum_{i=1}^{\infty} v(R_i) \leq v(\tilde{S}) \leq v(\tilde{R}) + \epsilon.$$

Moreover, the collection $\{R_i\}$ covers $\tilde{R}$ since its union is $\tilde{S}^{\circ}$, which contains $\tilde{R}$. □

Conversely, by reversing the roles of the axes, we see that if R is a parallel rectangle and $\epsilon > 0$, then there is a cover of R by oblique rectangles $\{\tilde{R}_i : i \in \mathbb{N}\}$ such that

$$\sum_{i=1}^{\infty} v(\tilde{R}_i) \leq v(R) + \epsilon. \tag{2.18}$$

Theorem 2.31. *If $E \subset \mathbb{R}^n$ and $Q : \mathbb{R}^n \to \mathbb{R}^n$ is an orthogonal transformation, then*

$$\mu^*(QE) = \mu^*(E),$$

and E is Lebesgue measurable if an only if QE is Lebesgue measurable.

PROOF. Let $\tilde{E} = QE$. Given $\epsilon > 0$ there is a cover of $\tilde{E}$ by parallel rectangles $\{R_i : i \in \mathbb{N}\}$ such that

$$\sum_{i=1}^{\infty} v(R_i) \leq \mu^*(\tilde{E}) + \frac{\epsilon}{2}.$$

From (2.18), for each $i \in \mathbb{N}$ we can choose a cover $\{\tilde{R}_{i,j} : j \in \mathbb{N}\}$ of R_i by oblique rectangles such that

$$\sum_{i=1}^{\infty} v(\tilde{R}_{i,j}) \leq v(R_i) + \frac{\epsilon}{2^{i+1}}.$$

Then $\{\tilde{R}_{i,j} : i, j \in \mathbb{N}\}$ is a countable cover of $\tilde{E}$ by oblique rectangles, and

$$\sum_{i,j=1}^{\infty} v(\tilde{R}_{i,j}) \leq \sum_{i=1}^{\infty} v(R_i) + \frac{\epsilon}{2} \leq \mu^*(\tilde{E}) + \epsilon.$$

If $R_{i,j} = Q^T \tilde{R}_{i,j}$, then $\{R_{i,j} : j \in \mathbb{N}\}$ is a cover of E by parallel rectangles, so

$$\mu^*(E) \le \sum_{i,j=1}^{\infty} v(R_{i,j}).$$

Moreover, since Q is orthogonal, we have $v(R_{i,j}) = v(\tilde{R}_{i,j})$. It follows that

$$\mu^*(E) \le \sum_{i,j=1}^{\infty} v(R_{i,j}) = \sum_{i,j=1}^{\infty} v(\tilde{R}_{i,j}) \le \mu^*(\tilde{E}) + \epsilon,$$

and since $\epsilon > 0$ is arbitrary, we conclude that

$$\mu^*(E) \le \mu^*(\tilde{E}).$$

By applying the same argument to the inverse mapping $E = Q^T \tilde{E}$, we get the reverse inequality, and it follows that $\mu^*(E) = \mu^*(\tilde{E})$.

Since μ^* is invariant under Q, the Carathéodory criterion for measurability is invariant, and E is measurable if and only if QE is measurable. □

It follows from Theorem 2.31 that Lebesgue measure is invariant under rotations and reflections.[4] Since it is also invariant under translations, Lebesgue measure is invariant under all isometries of $\mathbb{R}^n$.

Next, we consider the effect of dilations on Lebesgue measure. Arbitrary linear maps may then be analyzed by decomposing them into rotations and dilations.

Proposition 2.32. *Suppose that $\Lambda : \mathbb{R}^n \to \mathbb{R}^n$ is the linear transformation*

$$\Lambda : (x_1, x_2, \dots, x_n) \mapsto (\lambda_1 x_1, \lambda_2 x_2, \dots, \lambda_n x_n) \tag{2.19}$$

where the $\lambda_i > 0$ are positive constants. Then

$$\mu^*(\Lambda E) = (\det \Lambda)\mu^*(E),$$

and E is Lebesgue measurable if and only if ΛE is Lebesgue measurable.

Proof. The diagonal map Λ does not change the orientation of a rectangle, so it maps a cover of E by parallel rectangles to a cover of ΛE by parallel rectangles, and conversely. Moreover, Λ multiplies the volume of a rectangle by $\det \Lambda = \lambda_1 \dots \lambda_n$, so it immediate from the definition of outer measure that $\mu^*(\Lambda E) = (\det \Lambda)\mu^*(E)$, and E satisfies the Carathéodory criterion for measurability if and only if ΛE does. □

Theorem 2.33. *Suppose that $T : \mathbb{R}^n \to \mathbb{R}^n$ is a linear transformation and $E \subset \mathbb{R}^n$. Then*

$$\mu^*(TE) = |\det T|\, \mu^*(E),$$

and TE is Lebesgue measurable if E is measurable

Proof. If T is singular, then its range is a lower-dimensional subspace of $\mathbb{R}^n$, which has Lebesgue measure zero, and its determinant is zero, so the result holds.[5] We therefore assume that T is nonsingular.

[4] Unlike differential volume forms, Lebesgue measure does not depend on the orientation of $\mathbb{R}^n$; such measures are sometimes referred to as densities in differential geometry.

[5] In this case TE, is always Lebesgue measurable, with measure zero, even if E is not measurable.

In that case, according to the polar decomposition, the map T may be written as a composition

$$T = QU$$

of a positive definite, symmetric map $U = \sqrt{T^T T}$ and an orthogonal map Q. Any positive-definite, symmetric map U may be diagonalized by an orthogonal map O to get

$$U = O^T \Lambda O$$

where $\Lambda : \mathbb{R}^n \to \mathbb{R}^n$ has the form (2.19). From Theorem 2.31, orthogonal mappings leave the Lebesgue measure of a set invariant, so from Proposition 2.32

$$\mu^*(TE) = \mu^*(\Lambda E) = (\det \Lambda)\mu^*(E).$$

Since $|\det Q| = 1$ for any orthogonal map Q, we have $\det \Lambda = |\det T|$, and it follows that $\mu^*(TE) = |\det T|\, \mu^*(E)$.

Finally, it is straightforward to see that TE is measurable if E is measurable. $\square$

2.9. Lebesgue-Stieltjes measures

We briefly consider a generalization of one-dimensional Lebesgue measure, called Lebesgue-Stieltjes measures on $\mathbb{R}$. These measures are obtained from an increasing, right-continuous function $F : \mathbb{R} \to \mathbb{R}$, and assign to a half-open interval $(a, b]$ the measure

$$\mu_F\left((a, b]\right) = F(b) - F(a).$$

The use of half-open intervals is significant here because a Lebesgue-Stieltjes measure may assign nonzero measure to a single point. Thus, unlike Lebesgue measure, we need not have $\mu_F([a, b]) = \mu_F((a, b])$. Half-open intervals are also convenient because the complement of a half-open interval is a finite union of (possibly infinite) half-open intervals of the same type. Thus, the collection of finite unions of half-open intervals forms an algebra.

The right-continuity of F is consistent with the use of intervals that are half-open at the left, since

$$\bigcap_{i=1}^{\infty} (a, a + 1/i] = \varnothing,$$

so, from (1.2), if F is to define a measure we need

$$\lim_{i\to\infty} \mu_F\left((a, a + 1/i]\right) = 0$$

or

$$\lim_{i\to\infty} \left[F(a + 1/i) - F(a)\right] = \lim_{x\to a^+} F(x) - F(a) = 0.$$

Conversely, as we state in the next theorem, any such function F defines a Borel measure on $\mathbb{R}$.

Theorem 2.34. *Suppose that $F : \mathbb{R} \to \mathbb{R}$ is an increasing, right-continuous function. Then there is a unique Borel measure $\mu_F : \mathcal{B}(\mathbb{R}) \to [0, \infty]$ such that*

$$\mu_F\left((a, b]\right) = F(b) - F(a)$$

for every $a < b$.

The construction of μ_F is similar to the construction of Lebesgue measure on $\mathbb{R}^n$. We define an outer measure $\mu_F^* : \mathcal{P}(\mathbb{R}) \to [0, \infty]$ by

$$\mu_F^*(E) = \inf \left\{ \sum_{i=1}^{\infty} [F(b_i) - F(a_i)] : E \subset \bigcup_{i=1}^{\infty} (a_i, b_i] \right\},$$

and restrict μ_F^* to its Carathéodory measurable sets, which include the Borel sets. See *e.g.* Section 1.5 of Folland [**4**] for a detailed proof.

The following examples illustrate the three basic types of Lebesgue-Stieltjes measures.

Example 2.35. If $F(x) = x$, then μ_F is Lebesgue measure on $\mathbb{R}$ with

$$\mu_F\left((a, b]\right) = b - a.$$

Example 2.36. If

$$F(x) = \begin{cases} 1 & \text{if } x \geq 0, \\ 0 & \text{if } x < 0, \end{cases}$$

then μ_F is the δ-measure supported at 0,

$$\mu_F(A) = \begin{cases} 1 & \text{if } 0 \in A, \\ 0 & \text{if } 0 \notin A. \end{cases}$$

Example 2.37. If $F : \mathbb{R} \to \mathbb{R}$ is the Cantor function, then μ_F assigns measure one to the Cantor set, which has Lebesgue measure zero, and measure zero to its complement. Despite the fact that μ_F is supported on a set of Lebesgue measure zero, the μ_F-measure of any countable set is zero.