

# From Calculus to Modern Analysis

Erik van Erp  
University of Pennsylvania

Fall 2007



# Chapter 1

## Lebesgue Measure

### 1.1 Overview

Measure theory aims to quantify the size of sets of points. It generalizes the notion of length, area, and volume of classical geometric shapes to more general types of sets that appear in the analysis of functions. Besides squares and spheres and cylinders, we want to be able to measure Cantor sets and fractals and other sets that do not have smooth boundaries.

The most important measure for the purposes of analysis is *Lebesgue measure*, introduced by Henri Lebesgue in his doctoral thesis of 1902, as part of the foundation for his new technique of integration. In these notes we give a careful construction of the Lebesgue measure in Euclidean space  $\mathbb{R}^n$ . Most constructions of Lebesgue measure take as their point of departure the familiar notion of the volume of a rectangular box. From there, one gradually extends to more and more complicated sets.

In section 2, we study how to measure sets that can be divided into rectangular cells. This is one way to analyze the area inside a circle, for example: by dividing it up as an infinite collection of smaller and smaller squares. It turns out that all open sets can be measured in this way, but that not all closed sets are amenable to this approach.

Examples like the Cantor set show that closed sets may have no interior, and therefore no rectangle at all will fit inside. To deal with such sets, in section 3 we define outer and inner measures, which are defined for arbitrary sets. Now open and closed sets end up playing complementary roles.

We are now ready to define Lebesgue measure, which is done in section 4. Essentially, to measure a set we squeeze it between a compact set (from inside) and an open set (from outside). If the difference in measure between

the compact and open set can be made arbitrarily small, the inner and outer measures of the set are the same, and the set is Lebesgue measurable. The key property of Lebesgue measure, proven in section 4, is that it is additive for countable disjoint unions of measurable sets.

Finally, in section 5 we give a characterization of measurable sets, and will see that virtually all sets that play a role in analysis are Lebesgue measurable.

## 1.2 Content of Cellular Sets

As usual  $[a, b)$  denotes the half-open interval of numbers  $x \in \mathbb{R}$  with  $a \leq x < b$ . We allow  $a \leq b$ . The *length* of a half-open interval  $[a, b)$  is simply  $b - a$ ). Because there are no numbers  $x$  with  $a \leq x < a$ , we have  $[a, a) = \emptyset$ . We see that the length of  $\emptyset = [a, a)$  is zero.

A *cell* in  $\mathbb{R}^n$  is a product of half-open intervals

$$S = [a_1, b_1) \times [a_2, b_2) \times \cdots \times [a_n, b_n),$$

where  $a_i \leq b_i$ . The intervals  $[a_i, b_i)$  are called the *sides* of  $S$ . The *content* of  $S$  is the product of the length of its sides,

$$\nu(S) = (b_1 - a_1)(b_2 - a_2) \cdots (b_n - a_n).$$

Again,  $\emptyset = [a_1, a_1) \times \cdots \times [a_n, a_n)$  is a cell with content zero. With this convention the intersection of any two cells is again a cell (possibly empty).

A *partition* of an interval  $[a, b) \subset \mathbb{R}$  is a finite collection of disjoint intervals  $[x_i, x_{i+1})$  whose union is  $[a, b)$ . Such a partition corresponds to a finite strictly increasing sequence,

$$a = x_0 < x_1 < \cdots < x_{n-1} < x_n = b.$$

A partition of a cell  $S$  in  $\mathbb{R}^n$  is obtained by partitioning each of the sides of  $S$ ,

$$a_i = x_1^i < x_2^i < \cdots < x_{N_i}^i = b_i,$$

resulting in a collection  $P$  of  $N = N_1 N_2 \cdots N_n$  disjoint cells of the form

$$[x_{j_1}^1, x_{j_1+1}^1) \times [x_{j_2}^2, x_{j_2+1}^2) \times \cdots \times [x_{j_n}^n, x_{j_n+1}^n).$$

If we refer to “a partition  $P$  of  $S$ ”, then  $P$  denotes the set of subcells of  $S$  in the partition, and we will often assume that those subcells are numbered, at random,

$$P = \{S_1, S_2, \dots, S_N\}.$$

The construction of Lebesgue measure in Euclidean space  $\mathbb{R}^n$  has as its starting point the idea that the measure of a cell  $S$  must be equal to its content  $\nu(S)$ . We will first prove that content is well-behaved for disjoint unions of cells.

**Lemma 1** *Let the cell  $S = S_1 \cup S_2 \cup \dots \cup S_N$  be the disjoint union of finitely many subcells (not necessarily a partition). Then*

$$\nu(S) = \nu(S_1) + \nu(S_2) + \dots + \nu(S_N).$$

**Proof.** If the union of cells  $S_i$  is a *partition* of  $S$ , then this follows immediately from the distributive property  $a(b + c) = ab + ac$ .

In the general case, for  $i = 1, 2, \dots, N$ , let

$$S_i = [c_1^{(i)}, d_1^{(i)}] \times \dots \times [c_n^{(i)}, d_n^{(i)}].$$

By putting the various boundaries of sides with the same coordinate

$$\{c_j^{(1)}, d_j^{(1)}, \dots, c_j^{(N)}, d_j^{(N)}\}$$

in ascending order (removing duplicates), we obtain a partition of each  $x_j$ -coordinate in  $\mathbb{R}^n$ , and a corresponding partition  $P$  of  $S$ .

The crucial property of this partition is that each of the cells in  $P$  is entirely contained in one of the sets  $S_i$ . This proves the lemma, because the content  $\nu(S)$  is equal to the content of all the cells in  $P$ , while the content  $\nu(S_i)$  is equal to the sum of the contents of the cells in the partition  $P$  that are contained in  $S_i$ . □

From this simple fact, a much more interesting result is derived.

**Definition 2** *A subset  $E \subseteq \mathbb{R}^n$  is called cellular if it can be divided as a disjoint countable union of cells.*

**Proposition 3** *Let  $E$  be a cellular set, and let  $R_1, R_2, R_3, \dots$  be a sequence of cells (not necessarily disjoint) that covers  $E$ , i.e.,*

$$E \subseteq R_1 \cup R_2 \cup R_3 \cup \dots .$$

*If  $E = S_1 \cup S_2 \cup S_3 \cup \dots$  is any disjoint union of cells equal to  $E$ , then*

$$\sum_{i=1}^{\infty} \nu(S_i) \leq \sum_{j=1}^{\infty} \nu(R_j).$$

**Remark.** We allow for the possibility that the sum  $\sum^\infty \nu(S_i)$  diverges. In that case the inequality in the proposition states that  $\sum^\infty \nu(R_j)$  also diverges.

**Proof.** We first prove that the statement is true if  $E = S$  is just a single set.

Let  $A$  be a cell which is slightly smaller than  $E = S$ , with  $\nu(A) = (1 - \varepsilon)\nu(S)$ , and such that the compact closure  $\bar{A}$  is contained in  $S$ . And let  $B_j$  be a cell which is slightly larger than  $R_j$ , with  $\nu(B_j) = (1 + \varepsilon)\nu(R_j)$ , and such that the interior  $B_j^\circ$  contains  $R_j$ .

Then the open cover  $\{B_j^\circ\}$  of the compact set  $\bar{A}$  has a finite subcover, and we find that

$$B_1 \cup \cdots \cup B_M \supset B_1^\circ \cup \cdots \cup B_M^\circ \supset \bar{A} \supset A.$$

Because we only have a finite number of cells now, a partition argument shows that

$$\nu(B_1) + \cdots + \nu(B_M) \geq \nu(A),$$

and so for every  $\varepsilon > 0$ , we have

$$(1 + \varepsilon) \sum_{j=1}^{\infty} \nu(R_j) \geq (1 - \varepsilon)\nu(S),$$

which gives the desired inequality.

For the general case  $E = S_1 \cup S_2 \cup \cdots$ , we argue by contradiction. If  $\sum^\infty \nu(R_j) = \infty$ , there is nothing to prove. So assume that this sum converges, and that

$$\sum_{i=1}^{\infty} \nu(S_i) > \sum_{j=1}^{\infty} \nu(R_j).$$

Then, by definition of the infinite sum

$$\sum_{i=1}^{\infty} \nu(S_i) = \lim_{N \rightarrow \infty} \sum_{i=1}^N \nu(S_i),$$

and because  $\sum^\infty \nu(R_j)$  is finite, there is an integer  $N$  so that the partial sum satisfies

$$\sum_{i=1}^N \nu(S_i) > \sum_{j=1}^{\infty} \nu(R_j).$$

Define new cells (possibly empty)

$$C_{ij} = S_i \cap R_j.$$

Because the cells  $S_i$  are disjoint, the cells

$$C_{1j}, C_{2j}, \dots, C_{Nj}$$

are also disjoint. All of the cells  $C_{ij}$  with a fixed value for  $j$  are contained in the cell  $R_j$ . Therefore the previous lemma implies that

$$\nu(R_j) \geq \sum_{i=1}^N \nu(C_{ij}).$$

Therefore

$$\sum_{i=1}^N \nu(S_i) > \sum_{j=1}^{\infty} \nu(R_j) \geq \sum_{j=1}^{\infty} \left( \sum_{i=1}^N \nu(C_{ij}) \right) = \sum_{i=1}^N \left( \sum_{j=1}^{\infty} \nu(C_{ij}) \right).$$

This implies that for at least one value of  $i = 1, \dots, N$  we must have

$$\nu(S_i) > \sum_{j=1}^{\infty} \nu(C_{ij}).$$

This inequality contradicts what we proved before if we can see that the cells  $C_{ij}$  with fixed  $i$  cover the cell  $S_i$ . This is true, because  $S_i$  is covered by the cells  $R_j$ . Every point  $x \in S_i$  is contained in at least one  $R_j$ , but then also  $x \in C_{ij} = S_i \cap R_j$ . Therefore

$$S_i = \bigcup_{j=1}^{\infty} C_{ij}.$$

□

This result has important implications.

**Corollary 4** *The total content of a cellular set  $E = S_1 \cup S_2 \cup \dots$ , defined as*

$$\nu(E) = \nu(S_1) + \nu(S_2) + \dots,$$

*is independent of the way  $E$  is divided into disjoint cells.*

**Proof.** If  $E = R_1 \cup R_2 \cup \dots$  is another disjoint union of cells, then  $\bigcup S_i$  is covered by  $\bigcup R_j$ , and so  $\sum \nu(S_i) \leq \sum \nu(R_j)$ . But now  $\bigcup R_j$  is also covered by  $\bigcup S_i$ , which gives  $\sum \nu(R_j) \leq \sum \nu(S_i)$ . □

Thus, we have a well-defined notion of the content of cellular sets.

**Corollary 5** *If  $E$  is a cellular set covered by a countable union  $E_1, E_2, \dots$  of cellular sets, then*

$$\nu(E) \leq \sum_{i=1}^{\infty} \nu(E_i).$$

*If the cellular set  $E$  is a countable disjoint union of cellular sets  $E_1, E_2, \dots$ , then*

$$\nu(E) = \sum_{i=1}^{\infty} \nu(E_i).$$

**Proof.** Just write the sets  $E$  and  $E_i$  as disjoint unions of cells, and apply the proposition. □

A difficulty in developing content into a measure for arbitrary sets arises from the fact that the set-difference of two cellular sets is not necessarily cellular. In general, open sets are cellular, but compact sets are not.

**Proposition 6** *Every open set in  $\mathbb{R}^n$  is cellular.*

**Proof** Let  $\mathbb{P}$  be the countable collection of cells of the form

$$S = \left[ \frac{m_1}{2^k}, \frac{m_1 + 1}{2^k} \right) \times \dots \times \left[ \frac{m_n}{2^k}, \frac{m_n + 1}{2^k} \right),$$

where  $k$  is a positive integer, and  $m_i \in \mathbb{Z}, i = 1, \dots, n$ . This collection  $\mathbb{P}$  has the property that for two cells  $S_1, S_2$  in  $\mathbb{P}$ , either  $S_1 \subseteq S_2$ , or  $S_2 \subseteq S_1$ , or else  $S_1 \cap S_2 = \emptyset$ .

Let  $U$  be an open set. Every  $x \in U$  is contained in an open ball  $B_x \subseteq U$ . For every  $x \in U$  there exists a cell  $S \in \mathbb{P}$  with  $x \in S$  such that  $S$  is small enough to fit inside  $B_x$ . Then  $S \subseteq U$ . Let  $S_x$  be the largest cell from the collection  $\mathbb{P}$  that contains  $x$ , and that is contained in  $U$ .

Then for two points  $x, y \in U$ , either  $S_x = S_y$ , or else  $S_x \cap S_y = \emptyset$ . Thus, the cells  $S_x$  form a disjoint collection. It is a countable collection, because  $\mathbb{P}$  is countable. Since every  $S_x \subseteq U$ , and every  $x \in U$  is contained in some  $S_x$ , the union of all  $S_x$  is precisely  $U$ . □

**Proposition 7** *No compact set in  $\mathbb{R}^n$  is cellular.*

**Proof.** We argue by contradiction, assuming that the compact set  $K$  is a disjoint union  $K = \bigcup S_i$  of cells.

Consider the  $x_1$  coordinate as a continuous function

$$f(x) = f(x_1, x_2, \dots, x_n) = x_1.$$

If  $K$  is a compact set, there is a point  $p \in K$  where  $f(p) = p_1$  is maximal. By assumption, there exists a cell in  $K$

$$S_i = [a_1, b_1) \times \cdots \times [a_n, b_n)$$

that contains  $p$ . From  $p \in S_i$  we get  $p_1 \in [a_1, b_1)$ , and so in particular  $p_1 < b_1$ . But that means there are points in  $S_i$ , and hence in  $K$ , with larger  $x_1$ -coordinate than  $p$ , contradicting the assumption. □

**Exercise 1.** Prove that the interior of a cell has the same content as the cell itself.

**Exercise 2.** Let  $C \subset [0, 1]$  be a Cantor set, constructed as follows. Let  $C_0 = [0, 1]$ , and let  $U_0$  be an open interval of length  $\varepsilon$  contained in  $C_0$ , and  $C_1 = C_0 \setminus U_0$ . Then  $C_1$  has two connected components. Let  $U_1^1, U_1^2$  be two open intervals of length  $\varepsilon^2$  each, one contained in each component of  $C_1$ . Write  $U_1 = U_1^1 \cup U_1^2$ , and  $C_2 = C_1 \setminus U_1$ . Then  $C_2$  has four connected components. In general, let  $U_k^1, \dots, U_k^{2^k}$  be intervals of length  $\varepsilon^{k+1}$  each, one contained in every connected component of  $C_k$ . Write  $U_k = U_k^1 \cup \cdots \cup U_k^{2^k}$ , and  $C_{k+1} = C_k \setminus U_k$ .

The intersection

$$C = \bigcap_{i=1}^{\infty} C_i,$$

is a “fat” Cantor set. Show that the complement  $[0, 1] \setminus C$  is cellular, with content strictly less than 1.

**Exercise 3.** Prove the following proposition.

**Proposition 8** *Any countable union (not necessarily disjoint) of cellular sets is cellular.*

**Exercise 4.** Prove the following proposition.

**Proposition 9** *Every finite intersection of cellular cells is cellular.*

**Exercise 5.** Let  $q_1, q_2, q_3, \dots$  be an arbitrary sequence of numbers in  $[0, 1]$ . Consider the union

$$E = \bigcup_{i=1}^{\infty} \left[ q_i, q_i + \frac{1}{3^i} \right).$$

Prove that there exists a number  $x \in [0, 1]$  that is not contained in the set  $E$ .

Hint: Use Proposition 8 and Proposition 3 to show that  $E$  cannot contain all points in  $[0, 1]$ .

### 1.3 Outer and Inner Measure

A convenient way to quantify the size of arbitrary sets is by comparison with cellular sets.

**Definition 10** *Let  $E \subseteq \mathbb{R}^n$ . The outer measure  $\mu^*(E)$  is the infimum*

$$\mu^*(E) = \inf \{ \nu(U) \mid E \subseteq U, U \text{ open} \}.$$

**Proposition 11** *If  $E$  is cellular, then*

$$\mu^*(E) = \nu(E).$$

**Proof.** Since all open sets are cellular,  $E \subseteq U$  implies  $\nu(E) \leq \nu(U)$ , and therefore  $\nu(E) \leq \mu^*(E)$ . We must prove the reverse inequality.

Let  $E = S_1 \cup S_2 \cup \dots$  be a disjoint union of cells. Then

$$\nu(E) = \sum \nu(S_i).$$

For each cell  $S_i$  choose a slightly larger cell  $R_i$ , such that  $\nu(R_i) = (1 + \varepsilon)\nu(S_i)$  and  $S_i$  is contained in the interior  $R_i^o$  of  $R_i$ .

Let  $U_\varepsilon = \bigcup R_i^o$ . Then  $E \subseteq U_\varepsilon$ , and

$$\nu(U_\varepsilon) \leq \sum \nu(R_i) \leq (1 + \varepsilon) \sum \nu(S_i) \leq (1 + \varepsilon)\nu(E).$$

This can be done for any  $\varepsilon > 0$ , so

$$\mu^*(E) \leq \nu(E).$$

□

**Definition 12** Let  $E \subseteq \mathbb{R}^n$ . The inner measure  $\mu_*(E)$  is the supremum

$$\mu_*(E) = \sup \{ \mu^*(K) \mid K \subseteq E, K \text{ compact} \}.$$

If  $K \subseteq E \subseteq U$ , with  $K$  compact and  $U$  open, then  $\mu^*(K) \leq \mu^*(U) = \nu(U)$ . This implies the inequality,

$$\mu_*(E) \leq \mu^*(E).$$

Another obvious fact is that for compact sets,

$$\mu_*(K) = \mu^*(K).$$

We now show that for cells, all defined measures agree.

**Lemma 13** Let  $S \in \mathbb{R}^n$  be a cell, with interior  $S^o$  and closure  $\bar{S}$ .

Then the inner and outer measures of  $S$ ,  $S^o$ ,  $\bar{S}$  are all equal to the content of  $S$ .

**Proof.** Because  $S^o \subset S \subset \bar{S}$ , we have inequalities

$$\mu^*(S^o) \leq \mu^*(S) \leq \mu^*(\bar{S}).$$

If we can show that  $\mu^*(\bar{S}) \leq \mu^*(S^o)$ , then the three outer measures are equal. Take a cell  $R$  which is slightly larger than  $S$ , say  $\nu(R) = \nu(S) + \varepsilon$ , such that  $\bar{S} \subset R^o$ . We get  $\mu^*(\bar{S}) \leq \mu^*(R^o)$ . The interior  $R^o$  is an open set, and therefore a cellular set. Because  $R^o$  is cellular we have  $\mu^*(R^o) = \nu(R^o)$ , and by Exercise 1 we know that  $\nu(R^o) = \nu(R)$ . Therefore

$$\mu^*(\bar{S}) \leq \mu^*(R^o) = \nu(R^o) = \nu(R) = \nu(S) + \varepsilon = \nu(S^o) + \varepsilon = \mu^*(S^o) + \varepsilon.$$

This establishes

$$\nu(S) = \mu^*(S^o) = \mu^*(S) = \mu^*(\bar{S}).$$

Next we must consider the inner measures. Again, we can start with

$$\mu_*(S^o) \leq \mu_*(S) \leq \mu_*(\bar{S}).$$

This time take a cell  $R$  slightly smaller than  $S$ , with  $\nu(R) = \nu(S) - \varepsilon$ , such that  $\bar{R} \subset S^o$ . Then, by definition,  $\mu^*(\bar{R}) \leq \mu_*(S^o)$ . Because of what we proved about outer measures, we get

$$\mu^*(\bar{S}) = \nu(S) = \nu(R) + \varepsilon \leq \mu_*(S^o) + \varepsilon,$$

which proves  $\mu^*(\bar{S}) \leq \mu_*(S^o)$ . Finally, because the closure  $\bar{S}$  is a compact set, we have  $\mu_*(\bar{S}) = \mu^*(\bar{S})$ . So we have  $\mu_*(\bar{S}) \leq \mu_*(S^o)$ , which implies

$$\mu_*(S^o) = \mu_*(S) = \mu_*(\bar{S}) = \mu^*(\bar{S}).$$

□

## 1.4 Finite Measurable Sets

The content of open sets was defined by means of a cellular decomposition. An outer measure for arbitrary sets was defined by means of the content of open sets. Finally, an inner measure for arbitrary sets was defined by means of the outer measure of compact sets. It turns out that for virtually all sets, inner and outer measure agree.

**Definition 14** A set  $E \subseteq \mathbb{R}^n$  with finite outer measure  $\mu^*(E)$  is called a finite measurable set, if its inner measure is equal to its outer measure. In this case we define

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

and call  $\mu(E)$  the Lebesgue measure of  $E$ .

All we know so far is that compact sets and single cells are measurable. To get results about other sets, our aim in this section is to show how Lebesgue measure behaves with respect to countable unions and set differences.

We first prove *subadditivity* of outer measure. (Outer measure is not countably additive for general sets.)

**Proposition 15** Let  $E = E_1 \cup E_2 \cup E_3 \cup \dots$ , be a union (not necessarily disjoint) of subsets of  $\mathbb{R}^n$ . Then

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

**Proof.** Let  $U_i$  be an open set such that  $E_i \subseteq U_i$ , and

$$\mu^*(E_i) > \mu(U_i) - \varepsilon 2^{-i}.$$

Then

$$\sum_{i=1}^{\infty} \mu^*(E_i) > \sum_{i=1}^{\infty} \mu(U_i) - \varepsilon.$$

Let  $U = \bigcup_{i=1}^{\infty} U_i$ , then  $U$  is an open set that covers  $E$ . From the theory of cellular sets, we know that

$$\sum_{i=1}^{\infty} \mu(U_i) \geq \mu(U),$$

and, of course,  $\mu(U) \geq \mu^*(E)$ . In summary, for every  $\varepsilon > 0$

$$\sum_{i=1}^{\infty} \mu^*(E_i) > \mu^*(E) - \varepsilon.$$

□

As a corollary of this Proposition, we see that all *cellular sets*—and therefore, in particular, all open sets—are measurable.

**Lemma 16** *If  $K_1, K_2$  are two disjoint compact sets in  $\mathbb{R}^n$ , then*

$$\mu(K_1 \cup K_2) = \mu(K_1) + \mu(K_2).$$

**Proof.** By definition of the outer measure  $\mu^*$ , there exists an open set  $U$  such that  $K_1 \cup K_2 \subset U$ , and

$$\mu(K_1 \cup K_2) > \mu(U) - \varepsilon.$$

A standard separation property says that for two disjoint compact sets  $K_1, K_2$  there exists disjoint open sets  $V_1, V_2$  such that  $K_i \subset V_i$ . Then let  $U_i = U \cap V_i$ , so that  $U_1, U_2$  are disjoint open sets with  $U_1 \cup U_2 \subseteq U$  and  $K_i \subset U_i$ . Then

$$\mu(K_1 \cup K_2) > \mu(U) - \varepsilon > \mu(U_1 \cup U_2) - \varepsilon.$$

But the measure of open sets is additive (because they are cellular sets), and we obtain

$$\mu(K_1 \cup K_2) > \mu(U_1) + \mu(U_2) - \varepsilon > \mu(K_1) + \mu(K_2) - \varepsilon.$$

Since this holds for every  $\varepsilon > 0$ , we get the desired result.

□

This was just a special case that we need to prove the following more general statement.

**Proposition 17** *If  $E_1, E_2, E_3, \dots$  are disjoint finite measurable sets, such that the sum  $\sum_{i=1}^{\infty} \mu(E_i)$  is finite, then the countable union  $E = E_1 \cup E_2 \cup E_3 \dots$  is measurable with finite measure*

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

**Proof.** By definition of the inner measure  $\mu_*(E_i)$ , there exists a compact set  $K_i \subseteq E_i$  such that

$$\mu(K_i) > \mu(E_i) - 2^{-i}\varepsilon.$$

The previous lemma implies that

$$\mu(K_1 \cup \dots \cup K_N) = \sum_{i=1}^N \mu(K_i).$$

Then, by definition of inner measure,

$$\mu_*(E) \geq \sup_N \mu(K_1 \cup \dots \cup K_N) = \sum_{i=1}^{\infty} \mu(K_i) > \sum_{i=1}^{\infty} \mu(E_i) - \varepsilon.$$

Therefore

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

□

We see how the proof of countable additivity of Lebesgue measure relies on an interesting back-and-forth between inner and outer measures, and compact and open sets. Another example of this phenomenon is the proof that Lebesgue measure is subtractive.

**Lemma 18** *Let  $E \subseteq \mathbb{R}^n$  be a finite measurable set. For every  $\varepsilon > 0$  there exist compact  $K$  and open  $U$  such that,*

$$K \subseteq E \subseteq U$$

and

$$\mu(U \setminus K) < \varepsilon.$$

**Proof.** By definition of measurability, and of inner and outer measure, we can choose  $K, U$  such that  $K \subseteq E \subseteq U$  while  $\mu_*(E) - \mu(K) < \varepsilon$ , and  $\mu(U) - \mu^*(E) < \varepsilon$ . If  $\mu^*(E) = \mu_*(E)$ , then

$$\mu(U) - \mu(K) < 2\varepsilon.$$

But the set  $U \setminus K$  is open, hence measurable, and we have  $\mu(U \setminus K) + \mu(K) = \mu(U)$ , or

$$\mu(U \setminus K) = \mu(U) - \mu(K) < 2\varepsilon.$$

□

**Proposition 19** *If  $E_1, E_2$  are finite measurable sets, then so is the set difference  $E_1 \setminus E_2$ , and*

$$\mu(E_1 \setminus E_2) = \mu(E_1) - \mu(E_2).$$

**Proof.** Choose  $K_i \subseteq E_i \subseteq U_i$  with  $\mu(U_i \setminus K_i) < \varepsilon$ . Because of the disjoint union

$$U_1 \setminus K_2 \subseteq (U_1 \setminus K_1) \cup (K_1 \setminus U_2) \cup (U_2 \setminus K_2),$$

we have

$$\mu^*(E_1 \setminus E_2) \leq \mu^*(U_1 \setminus K_2) \leq \mu^*(K_1 \setminus U_2) + 2\varepsilon.$$

Now  $K_1 \setminus U_2$  is a compact subset of  $E_1 \setminus E_2$ , and therefore

$$\mu^*(E_1 \setminus E_2) \leq \mu_*(E_1 \setminus E_2) + 2\varepsilon.$$

Since this holds for all  $\varepsilon > 0$ , it follows that  $E_1 \setminus E_2$  is measurable.

The formula for the measure is a consequence of the disjoint union

$$E_1 = E_2 \cup (E_1 \setminus E_2).$$

□

## 1.5 Sets of Infinite Measure

To complete the construction of Lebesgue measure, we discuss the distinction between sets with *infinite measure* and sets that are *not measurable*. This is analogous to the difference between a series like  $\sum_{i=1}^{\infty} i = \infty$ , which can meaningfully be said to converge to infinity, and the divergent series  $\sum_{i=1}^{\infty} (-1)^i$ , which cannot.

**Definition 20** *A measurable set is any countable union of finite measurable sets.*

The following proposition shows that there can be no confusion of terminology when we speak of ‘finite measurable set’ in the sense of Definition 14, or a ‘measurable set’ in the sense of Definition 20 that has finite (inner or outer) measure.

**Proposition 21** *If  $E \subseteq \mathbb{R}^n$  is a measurable set, then the following are equivalent: (1)  $E$  is a finite measurable set, in the sense of Definition 14, (2)  $\mu^*(E)$  is finite, (3)  $\mu_*(E)$  is finite.*

**Proof.** That (1) implies (2) is true by definition, and that (2) implies (3) is also obvious.

To see that (3) implies (1), we argue by contradiction. Let  $E = E_1 \cup E_2 \cup \dots$  be a disjoint union, with  $E_i$  finite measurable, and  $\sum \mu(E_i) = \infty$ . Choose disjoint compact sets  $K_i \subseteq E_i$  with  $\mu(K_i) = (1 - \varepsilon)\mu(E_i)$ . Then  $K_1 \cup \dots \cup K_N$  is a compact set with measure  $(1 - \varepsilon)\sum_{i=1}^N \mu(E_i)$ , which converges to  $\infty$  as  $N \rightarrow \infty$ . □

Observe that it is possible that  $\mu_*(E) = \mu^*(E)$ , while  $E$  is *not* measurable.

**Definition 22** *If  $E$  is measurable, and  $\mu_*(E) = \mu^*(E) = \infty$ , then we say that  $E$  has infinite measure, and write  $\mu(E) = \infty$ .*

We can extend Proposition 17, about countable additivity of Lebesgue measure, to include infinite measurable sets.

**Proposition 23** *Any countable union  $E = E_1 \cup E_2 \cup \dots$  of measurable sets is measurable. Moreover, if any one  $\mu(E_i) = \infty$ , or if  $\sum \mu(E_i) = \infty$ , then  $\mu(E) = \infty$ .*

**Proof.** This follows directly from Definition 20 and Proposition 21. □

The following property of measurable sets is the main reason that we distinguish measurable sets with infinite measure from non-measurable sets.

**Proposition 24** *The complement of a measurable set is measurable.*

**Proof.** Let  $\mathbb{R}^n = S_1 \cup S_2 \cup \dots$  be some disjoint union of cells that fills the entire space  $\mathbb{R}^n$ . Suppose  $E$  is a measurable set, with  $E = E_1 \cup E_2 \cup \dots$ , where  $\mu(E_i) < \infty$ . Let

$$G_j = S_j \setminus \bigcup_{i=1}^{\infty} (E_i \cap S_j).$$

Because  $E_i$  and  $S_i$  are finite measurable, and  $G_j \subset S_j$ , the set  $G_j$  is also finite measurable (Proposition 17). The complement  $E^c$  of  $E$  is the countable union  $E^c = G_1 \cup G_2 \cup \dots$ . So  $E^c$  is measurable. □

An immediate corollary is the following.

**Corollary 25** *Countable unions and countable intersections of measurable sets are measurable.*

**Proof.** This is so because  $\cap E_i = (\cap E_i^c)^c$ . □

We can now extend Lemma 18 to sets with infinite measure.

**Proposition 26** *Let  $E \subseteq \mathbb{R}^n$  be a measurable set. For every  $\varepsilon > 0$  there exists a closed set  $F$  and an open set  $G$  such that,*

$$F \subseteq E \subseteq G$$

and

$$\mu(G \setminus F) < \varepsilon.$$

**Proof.** As in the proof of the previous proposition, let  $\mathbb{R}^n = S_1 \cup S_2 \cup \dots$  with  $\mu(S_i)$  finite. The sets  $E_i = E \cap S_i$  are disjoint, and  $E = \bigcup E_i$ . Each  $E_i$  is measurable (Proposition 26), and of finite measure (Proposition 21, using  $\mu^*(E_i) \leq \mu^*(S_i)$ ).

By Lemma 18 there exists an open set  $U_i$  with  $E_i \subseteq U_i$  and  $\mu(U_i \setminus E_i) < \varepsilon 2^{-i}$ , for arbitrary choice of  $\varepsilon > 0$ . The open set  $U = \bigcup U_i$  contains  $E$ , while  $U \setminus E \subseteq \bigcup (U_i \setminus E_i)$  implies

$$\mu(U \setminus E) < \varepsilon.$$

We can take  $G = U$ . Applying the same idea to the measurable set  $E^c$ , we find an open set  $V$  with  $E^c \subseteq V$  and  $\mu(V \setminus E^c) < \varepsilon$ . Take  $F = V^c$ . □

## 1.6 Borel Sets

The essential properties of the collection of measurable sets can be summed up in a definition.

**Definition 27** *A  $\sigma$ -algebra  $\Sigma$  in a set  $X$  is a collection of subsets of  $X$  that has the following properties,*

- (1) *If  $E$  is in  $\Sigma$ , then the complement  $E^c = X \setminus E$  is in  $\Sigma$ ,*
- (2) *if  $E_1, E_2, \dots$  are in  $\Sigma$  then the union  $E_1 \cup E_2 \cup \dots$  is in  $\Sigma$ .*

Observe that the axioms of a  $\sigma$ -algebra imply that finite or countable intersections  $E_1 \cap E_2 \cap \dots$  are also in  $\Sigma$ , because of the formula

$$\bigcap_{i=1}^{\infty} E_i = \left( \bigcup_{i=1}^{\infty} E_i^c \right)^c.$$

From what we have seen, the collection of measurable sets is a  $\sigma$ -algebra in  $\mathbb{R}^n$ . With the inclusion of sets of infinite measure, all open and closed sets are measurable. Countable intersections of open sets are called  $G_\delta$  sets, while countable unions of closed sets are called  $F_\sigma$  sets. In turn, countable unions of  $G_\delta$  sets are called  $G_{\delta\sigma}$  sets, and countable intersections of  $F_\sigma$  sets are called  $F_{\sigma\delta}$  sets. All sets that can be constructed from open and closed sets by repeated application of these basic set operations are called *Borel sets*. To make precise what we mean by ‘repeated application’ would involve the theory of transfinite ordinals (a fundamental aspect of the theory of infinite sets). But a very straightforward definition of Borel sets is as follows.

**Definition 28** *The collection of Borel sets is the smallest  $\sigma$ -algebra in  $\mathbb{R}^n$  that contains all open sets.*

The following proposition describes the relationship between measurable sets and Borel sets.

**Proposition 29** *A set  $E \subseteq \mathbb{R}^n$  is measurable if and only if there exists an  $F_\sigma$  set  $A$ , and a  $G_\delta$  set  $B$  with  $A \subseteq E \subseteq B$  and such that  $\mu(B \setminus A) = 0$ .*

*Less specifically, a set is measurable if and only if it differs from a Borel set by a set of measure zero.*

**Proof.** Suppose  $E$  is measurable. For  $j = 1, 2, \dots$  choose closed sets  $F_j$  and open sets  $G_j$  with  $F_j \subseteq E \subseteq G_j$  and  $\mu(G_j) - \mu(F_j) < j^{-1}$ . Then  $A = F_1 \cup F_2 \cup \dots$ , and  $B = G_1 \cap G_2 \cap \dots$  satisfy the conditions.

Conversely, suppose  $A \subseteq E \subseteq B$ , as stated. Then, since every  $F_\sigma$  and  $G_\delta$  set is measurable,  $A$  and  $B$  are measurable. It follows that  $\mu^*(E \setminus A) \leq \mu^*(B \setminus A) = 0$ , and therefore  $E \setminus A$  is measurable with measure zero. Then also  $E = A \cup (E \setminus A)$  is measurable. □

The crucial properties of Lebesgue measure are abstracted in the following definition of a general measure.

**Definition 30** *A measure  $\mu$  on a  $\sigma$ -algebra  $\Sigma$  in a set  $X$  is a function*

$$\mu: \Sigma \rightarrow [0, \infty]$$

*that assigns a non-negative real number  $\mu(E)$  (possibly  $\infty$ ) to every set  $E \in \Sigma$ . A measure must be countably additive, which means that if  $E_1, E_2, E_3, \dots$  is a sequence of disjoint sets in  $\Sigma$ , then*

$$\mu(E_1 \cup E_2 \cup \dots) = \sum_{i=1}^{\infty} \mu(E_i).$$

A measure space is a set  $X$ , together with a  $\sigma$ -algebra  $\Sigma$  of subsets in  $X$ , and a measure  $\mu$  defined for sets in  $\Sigma$ .

We have seen that Lebesgue measure satisfies all the necessary properties. Its  $\sigma$ -algebra is the collection of all measurable sets, which includes all Borel sets, and all sets of measure zero. While it is true that there exist subsets of  $\mathbb{R}^n$  that are not Lebesgue measurable, it is not possible to construct an example of such a set by means of explicit operations. Non-measurable sets can only be defined indirectly, by means of a notorious (if not controversial) axiom of set theory called the *axiom of choice*. Therefore, in practice, every set you will ever encounter will be measurable.

In short, Lebesgue measure is a very powerful extension of the classical notion of volume (or length, or area). It applies to virtually every kind of set that plays a role in analysis, while at the same time retaining the important property of (countable) additivity.



## Chapter 2

# Lebesgue Integral

### 2.1 Integral

There are several alternative definitions of the Lebesgue integral. We begin with a definition that is a straightforward generalization of the definition of the Riemann integral. All we need to do is replace partitions of a domain into rectangular cells with a more flexible division into measurable sets.

A *division*  $P$  of a finite measurable set  $A \subseteq \mathbb{R}^n$  is a finite collection  $P = \{R_1, R_2, \dots, R_N\}$  of mutually disjoint measurable sets  $R_i$ , such that

$$A = R_1 \cup R_2 \cup \dots \cup R_N.$$

We call the subsets  $R_i$  the *regions* of the division. Observe that a *partition* (as used in the definition of the Riemann integral) is a special kind of division. But not every division is a partition, nor can it always be refined to a partition. (Think, for example, of the division of the unit interval  $A = [0, 1]$  into the Cantor set and its complement.) Thus, divisions are more flexible than partitions.

We first consider integration of functions  $f: \mathbb{R}^n \rightarrow \mathbb{R}$  that are (1) bounded, i.e.,  $-M \leq f(x) \leq M$ , and (2) have bounded domain, i.e., there exists a bounded set  $A$  such that  $f(x) = 0$  if  $x \notin A$ . We may assume that  $A$  is a measurable set. All Riemann integrable functions are of this type. We now repeat the basic steps of the definition of the Riemann integral.

For every division  $P$  of the domain  $A \subseteq \mathbb{R}^n$ , and every region  $R \in P$ , define

$$M_R = \sup \{f(x) \mid x \in R\}$$
$$m_R = \inf \{f(x) \mid x \in R\}.$$

The *upper* and *lower sums* are

$$U_P = \sum_{R \in P} M_R \cdot \mu(R)$$

$$L_P = \sum_{R \in P} m_R \cdot \mu(R).$$

A division  $P_1$  is called a *refinement* of a division  $P_2$ , if every region  $R \in P_1$  is a subset  $R \subseteq T$  of a region  $T \in P_2$ . If  $P_1$  is a refinement of  $P_2$ , then

$$L_{P_2} \leq L_{P_1} \leq U_{P_1} \leq U_{P_2}.$$

Also, for any two measurable divisions there always exists a common refinement. If  $P_1 = \{R_1, \dots, R_N\}$ , and  $P_2 = \{T_1, \dots, T_M\}$ , then the division  $P_3$  consisting of the sets  $V_{ij} = R_i \cap T_j$  is a common refinement. It follows that all lower sums are  $\leq$  all upper sums, and, just as in the case of the Riemann integral,

$$\sup_P L_P \leq \inf_P U_P,$$

where the supremum and infimum are taken over all divisions  $P$  of  $A \subseteq \mathbb{R}^n$

**Definition 31** A bounded function  $f: \mathbb{R}^n \rightarrow \mathbb{R}$  with bounded domain  $A \subseteq \mathbb{R}^n$  is Lebesgue integrable if

$$\sup_P L_P = \inf_P U_P.$$

If this is the case we define the Lebesgue integral  $\int f = \inf U_P = \sup L_P$ .

It is immediately clear from this definition that every Riemann integrable function is Lebesgue integrable, with the same value for the integral. Because the definition is so similar to that of the Riemann integral, various simple facts about the Riemann integral are still true about the Lebesgue integral. For example, ‘Riemann’s Condition’ is still valid.

**Proposition 32** A bounded function with bounded domain is Lebesgue integrable if and only if for every  $\varepsilon > 0$  there exists a measurable division of the domain for which

$$U_Q - L_Q < \varepsilon.$$

The proof of this fact is the same as in the case of the Riemann integral, and we will not repeat it here.

Our first goal is to see that virtually *all* functions are Lebesgue integrable. As we will see, the criterion for ‘Lebesgue integrability’ is much looser than that for ‘Riemann integrability’. The key concept is that of a ‘measurable function’.

The pre-image  $f^{-1}(B)$  of a set  $B \subseteq \mathbb{R}$  is the set of all points  $x \in \mathbb{R}^n$  for which  $f(x) \in B$ .

**Definition 33** *A function  $f: \mathbb{R}^n \rightarrow \mathbb{R}$  is called a measurable function if the pre-image  $f^{-1}([a, b))$  of the half-open interval  $[a, b) \subseteq \mathbb{R}$  is a measurable set in  $\mathbb{R}^n$  for every  $a < b$ .*

*A function  $f: \mathbb{R}^n \rightarrow \mathbb{R}$  is called a Borel function if  $f^{-1}([a, b))$  is a Borel set for every  $a < b$ .*

Because virtually every set is measurable, virtually every function is measurable. And just as you cannot *construct* a non-measurable set (even though they ‘exist’ according to the axiom of choice) you also cannot *construct* a non-measurable function (even though they ‘exist’ in some formal sense). Before exploring the meaning of this concept, we first prove the key property of the Lebesgue integral.

**Proposition 34** *Every measurable bounded function with bounded domain is Lebesgue integrable.*

**Proof.** Let  $f: A \rightarrow \mathbb{R}$  be a measurable function, with bounded domain  $A$  (a finite measurable set), and such that

$$-M < f(x) < M,$$

for all  $x \in A$ . We partition the range  $[-M, M]$  of the function  $f$  into  $N$  equal parts  $\Delta y = 2M/N$ ,

$$y_0 = -M < y_1 = -M + \Delta y < \cdots < y_N = -M + N\Delta y = +M.$$

This partition of the range induces a division  $P = \{A_1, \dots, A_N\}$  of the domain  $A$  as follows,

$$A_i = f^{-1}([y_{i-1}, y_i)), \quad i = 1, 2, \dots, N.$$

Since  $f$  is measurable, each  $A_i$  will be measurable. (In general,  $P$  is not a partition!) For this particular division  $P$  of  $A$ , we have  $M_R - m_R < \Delta y$  for

each region  $R = A_i$  in  $P$ . Therefore

$$U_P - L_P = \sum_{i=1}^N (M_i - m_i) \mu(A_i) \leq \Delta y \sum_{i=1}^N \mu(A_i) = \Delta y \cdot \mu(A).$$

By choosing  $\Delta y$  sufficiently small, we see that the function  $f$  satisfies the condition in Proposition 32, and so  $f$  is Lebesgue integrable.  $\square$

This proposition shows that ‘virtually every’ bounded function on a bounded domain is Lebesgue integrable. The real advantages of the Lebesgue integral become evident when we consider how it relates to limits of sequences of functions. First we will consider some of the features of ‘measurable functions’, to see that, indeed, all functions of interest in analysis are measurable.

## 2.2 Measurable functions

The operation of taking the pre-image  $f^{-1}(A) \subseteq \mathbb{R}^n$  of a set  $A \subset \mathbb{R}$  commutes with the various Boolean set operations. For example,

$$f^{-1}(A \cup B) = f^{-1}(A) \cup f^{-1}(B).$$

To see that this is so, notice that  $x \in f^{-1}(A) \cup f^{-1}(B)$  means that  $x \in f^{-1}(A)$  or  $x \in f^{-1}(B)$ , which is another way of saying that  $f(x) \in A$  or  $f(x) \in B$ , or simply  $f(x) \in A \cup B$ . But that is equivalent to  $x \in f^{-1}(A \cup B)$ , which proves the equality. In a similar way, you can check that

$$\begin{aligned} f^{-1}(A \cap B) &= f^{-1}(A) \cap f^{-1}(B), \\ f^{-1}(A \setminus B) &= f^{-1}(A) \setminus f^{-1}(B), \\ f^{-1}(A^c) &= f^{-1}(A)^c. \end{aligned}$$

The same is true for *infinite* unions and intersections. The following proposition is an immediate corollary of these simple rules.

**Proposition 35** *If  $f$  is a measurable function, then  $f^{-1}(A)$  is a measurable set in  $\mathbb{R}^n$  whenever  $A$  is a Borel set in  $\mathbb{R}$ .*

*If  $f$  is a Borel function, then  $f^{-1}(A)$  is a Borel set in  $\mathbb{R}^n$  whenever  $A$  is a Borel set in  $\mathbb{R}$ .*

**Proof.** Every Borel set in  $\mathbb{R}$  can be constructed from half-open intervals by successive taking countable unions, countable intersections, and complements. For example, let  $A = A_1 \cup A_2 \cup \dots$  be a countable union, and suppose we have proven that each  $A_1, A_2, \dots$  is such that  $f^{-1}(A_i)$  is measurable (or Borel), then it follows that also  $f^{-1}(A)$  is measurable (or Borel), because of the rule

$$f^{-1}(A) = f^{-1}(A_1) \cup f^{-1}(A_2) \cup \dots$$

(We have to be a little vague about the fact that you can repeat such operations an ‘unlimited number’ of times. A rigorous proof requires the notion of ‘transfinite induction’ from the set theory of ordinals.)

□

Virtually all the techniques we have for creating new functions out of old ones preserve measurability. Observe, to begin, that of course every *continuous* function is measurable, and even a Borel function.

**Proposition 36** *If  $f: \mathbb{R}^n \rightarrow \mathbb{R}$  is measurable, and  $\phi: \mathbb{R} \rightarrow \mathbb{R}$  is any Borel function, then the composition*

$$\mathbb{R}^n \rightarrow \mathbb{R} : x \mapsto \phi(f(x))$$

*is measurable.*

*In particular,  $|f|$  and  $cf$  (for arbitrary scalar  $c \in \mathbb{R}$ ) are measurable.*

**Proof.** The pre-image  $\phi^{-1}([a, b] \subseteq \mathbb{R})$  is a Borel set, because  $\phi$  is a Borel function. Therefore  $f^{-1}(\phi^{-1}([a, b])) \subseteq \mathbb{R}^n$  is a measurable set. Of course,  $(f \circ \phi)^{-1} = f^{-1} \circ \phi^{-1}$ .

□

So you can perform any basic operation to a measurable function  $f$ , and obtain a new measurable function.