STUDENT MATHEMATICAL LIBRARY
Volume 78

A User-Friendly Introduction to Lebesgue Measure and Integration

Gail S. Nelson

AMS
AMERICAN MATHEMATICAL SOCIETY

http://dx.doi.org/10.1090/stml/078

STUDENT MATHEMATICAL LIBRARY
Volume 78

A User-Friendly Introduction to Lebesgue Measure and Integration

Gail S. Nelson

AMS

American Mathematical Society
Providence, Rhode Island

2010 *Mathematics Subject Classification.* Primary 26-XX, 28-XX.

For additional information and updates on this book, visit
www.ams.org/bookpages/stml-78

Library of Congress Cataloging-in-Publication Data

Nelson, Gail Susan.
A user-friendly introduction to Lebesgue measure and integration / Gail S. Nelson.
pages cm. – (Student mathematical library ; volume 78)
Includes bibliographical references and index.
ISBN 978-1-4704-2199-1 (alk. paper)
1. Measure theory. 2. Lebesgue integral. 3. Integration, Functional. I. Title.

QC20.7.M43N45 2015
515′.83–dc23 2015022834

10 9 8 7 6 5 4 3 2 1 20 19 18 17 16 15

To all of the students I've had. Thanks for teaching me. Most of all, thanks to my parents, the best teachers I have had.

Contents

Preface

When I first had the chance to teach a second course in real analysis I did the usual thing; I searched for a textbook that would fit the course as I envisioned it. I wanted to show my students that real analysis was more than just ϵ-δ proofs. I also hoped this course would provide a bit of a head start to those students heading off to graduate school. However, I could not find any text that suited the needs of my target audience, undergraduate students who had seen the basics of sequences and series up through and including the introduction to Riemann integration. So I started teaching the course from scratch, creating my own notes as the course progressed. The feedback from my students was that, while they liked the course in general, they missed having a textbook. Based on this feedback, each of the next couple of times I taught the course, part of the work assigned to the students was to "write" their book. The students took turns carefully rewriting their notes which were then collected in a binder in a central location. The resulting "textbooks" generated in this fashion formed the skeleton of this book.

The prerequisite for this course is a standard undergraduate first course in real analysis. Students need to be familiar with basic limit definitions, and how these definitions are used in sequences and in defining continuity and differentiation. The properties of a supremum (or least upper bound) and infimum (or greatest lower bound) are used repeatedly. The definition of compactness via open coverings is used in this text, but primarily for $\mathbb{R}^n$. I also assume students have seen sequences and series of functions and understand pointwise and uniform convergence. Since a major focus of this text is Lebesgue integration, it is also assumed that students have studied Riemann integration in their first real analysis course. Chapter 0 briefly covers Riemann integration with the approach that is later mimicked in defining the Lebesgue integral, that is, the use of upper and lower sums. (I do realize there are other approaches to the Riemann integral. The approach which uses step functions is the one used in Chapter 4 when the subjects of general measure and integration are introduced.) However, Chapter 0 exists primarily as a source of review and can be omitted.

One of the standard topics in the first analysis course that I teach is the completeness of the set of real numbers. The students often see this first in terms of every nonempty bounded set having a least upper bound. Later they are introduced to the Cauchy criterion and shown that in the real number system all Cauchy sequences converge. My experience has been that this Cauchy criterion is not fully appreciated by my students. In this second course in real analysis completeness via Cauchy sequences is a recurring theme; we first revisit the completeness of $\mathbb{R}$, then L^1; and more generally L^p.

I want to keep my course as "real" as possible. Instead of introducing measure via the Carathéodory definition, I opt to introduce Lebesgue measure through the more "concrete" definition using outer measure. In this way, Lebesgue measure is a natural extension of the concept of length, or area, or volume, depending on dimension.

So, here is my course. I start with a review of Riemann integration. I tend to keep this review to a minimum since most of the main theorems have their Lebesgue counterparts later in Chapter 2. As soon as possible, we move into Chapter 1 which covers Lebesgue outer measure and Lebesgue measure. It should be noted that Section 1.3 contains the classic construction of a nonmeasurable set which assumes knowledge of countability and familiarity with the Axiom of Choice. This section is not needed for the later chapters and can be omitted, although it justifies the difference between measure and outer measure and is referenced in Remark 4.1.14. The Lebesgue integral is defined in Chapter 2. Chapter 3 is where I introduce L^p spaces and use these as examples of Banach spaces. Later in the chapter L^2 is shown to be an example of a Hilbert space. My goal for a one-semester course is to end somewhere in Chapter 4, usually around Section 4.3. Sections 4.4 and 4.5 are independent of each other and at various time I have ended with one or the other.

I have also included an appendix entitled "Ideas for Projects". Most of these are topics that I had at one time considered including as part of my course. Instead, I reserved them for student presentations. My students typically work on these in pairs. In the past I just assigned the topic with a pointer to a possible source. However, here I have included sketches of how one might proceed.

Thanks to the members of the Carleton College mathematics department for their support. I also owe a large debt to all of the students who have been a part of this ongoing project. Without them, this book would never have been created.

Please direct comments and corrections to:

`gnelson@carleton.edu`

Enjoy exploring the wide world of real analysis!

Gail S. Nelson

Northfield, Minnesota

http://dx.doi.org/10.1090/stml/078/01

Chapter 0

Review of Riemann Integration

The main goals of this text are to provide introductions to Lebesgue measure, Lebesgue integration, and general measure theory. It is assumed that the reader has studied Riemann integration. Therefore, it is possible to omit this chapter altogether and start with Chapter 1. However, our development of the Lebesgue integral follows very closely the approach used by Darboux. Therefore, we present this optional chapter for those who would like a brief review of this approach to the Riemann integral. This chapter is not a complete treatment of the Riemann integral. For example, it does not include improper integrals, although we will later include unbounded functions in our coverage of the Lebesgue integral. However, it does include the basic development of the Riemann integral that will be imitated in Chapter 2.

0.1. Basic Definitions

To begin, we will be working with the set of bounded functions on a closed interval. We will denote this set by $B[a,b]$, that is,

$$B[a,b] = \left\{ f \;\middle|\; \text{for some } M \in \mathbb{R}, |f(x)| \leq M \text{ for all } x \in [a,b] \right\}.$$

In particular, if $f \in B[a, b]$, then $f(x)$ is defined for all $x \in [a, b]$. The constant M can be thought of as an upper bound for $|f|$, but it is not unique.

Example 0.1.1. Let $f(x) = \frac{1}{x}$. Then f is not in $B[0, 1]$ since f is not defined for $x = 0$. However, $f \in B[1, 2]$ since $|f(x)| \leq 1$ for all $x \in [1, 2]$.

Example 0.1.2. For $x \in [0, 1]$, let

$$f(x) = \begin{cases} \frac{1}{x} & \text{if } x \neq 0, \\ 0 & \text{if } x = 0. \end{cases}$$

This time f is defined for $x = 0$, but f is still not in $B[0, 1]$. Although one can see this by thinking of the graph of f, we give a more formal argument here. For every $M > 1$, $\frac{1}{M+1} \in [0, 1]$ and $f(\frac{1}{M+1}) = M + 1 > M$. Thus f is not in $B[0, 1]$ because no constant M will serve as an upper bound for $|f|$.

Example 0.1.3. Let

$$f(x) = \begin{cases} 1 & \text{if } x \in \mathbb{Q}, \\ 0 & \text{if } x \notin \mathbb{Q}. \end{cases}$$

This function is known as the *Dirichlet function* and is often denoted by $\mathcal{X}_{\mathbb{Q}}(x)$. Because $|\mathcal{X}_{\mathbb{Q}}(x)| \leq 1$ for all x, $\mathcal{X}_{\mathbb{Q}} \in B[0, 1]$.

The Dirichlet function shows that not all functions in $B[a, b]$ are continuous. On the other hand, if f is continuous on $[a, b]$, then $f \in B[a, b]$ by the Extreme Value Theorem, a result found in most texts used for a first real analysis course. For example, see Abbott [**1**].

In order to define the Riemann integral, we start by defining upper sums and lower sums.

Definition 0.1.4. A **partition** P of the interval $[a, b]$ is a finite ordered set

$$P = \{x_0, x_1, x_2, \ldots, x_n\},$$

where

$$a = x_0 < x_1 < x_2 < \ldots < x_n = b\,.$$

Let $f \in B[a, b]$. For each subinterval $[x_{i-1}, x_i]$ of $[a, b]$, set

$$m_i = \inf_{x \in [x_{i-1}, x_i]} f(x) = \inf\{f(x) \mid x \in [x_{i-1}, x_i]\}$$

and

$$M_i = \sup_{x\in[x_{i-1},x_i]} f(x) = \sup\{f(x) \mid x \in [x_{i-1}, x_i]\}\,.$$

Definition 0.1.5. Let $f \in B[a, b]$.

(i) The **lower sum** of f with respect to the partition P, written $L(f, P)$, is

$$L(f, P) = \sum_{i=1}^{n} m_i(x_i - x_{i-1})\,.$$

(ii) The **upper sum** of f with respect to the partition P, written $U(f, P)$, is

$$U(f, P) = \sum_{i=1}^{n} M_i(x_i - x_{i-1})\,.$$

Proposition 0.1.6. *Let $f \in B[a, b]$. For any partition P of $[a, b]$,*

$$m(b - a) \le L(f, P) \le U(f, P) \le M(b - a),$$

where

$$m = \inf_{x\in[a,b]} f(x) \qquad \textit{and} \qquad M = \sup_{x\in[a,b]} f(x)\,.$$

The proof of this proposition is straightforward and is left to the reader as an exercise (See Exercise 1).

From Proposition 0.1.6, the set $\{L(f, P) \mid P \text{ is a partition of } [a, b]\}$ and the set $\{U(f, P) \mid P \text{ is a partition of } [a, b]\}$ are both bounded sets. This allows us to make the following definitions.

Definition 0.1.7. Let $f \in B[a, b]$.

i) The **lower integral** of f is

$$\underline{\int_a^b} f(x)\,dx = \sup_P L(f, P) = \sup\{L(f, P) \mid P \text{ is a partition of } [a, b]\}\,.$$

ii) The **upper integral** of f is

$$\overline{\int_a^b} f(x)\,dx = \inf_P U(f, P) = \inf\{U(f, P) \mid P \text{ is a partition of } [a, b]\}\,.$$

The lower integral and the upper integral need not be equal. The Dirichlet function provides us with such an example.

Example 0.1.8. Consider the Dirichlet function $\mathcal{X}_{\mathbb{Q}}(x)$ as described in Example 0.1.3. For any partition P of $[0,1]$,

$$L(\mathcal{X}_{\mathbb{Q}}, P) = 0 \qquad \text{and} \qquad U(\mathcal{X}_{\mathbb{Q}}, P) = 1\,.$$

Therefore,

$$\underline{\int_0^1} \mathcal{X}_{\mathbb{Q}}(x)\,dx = 0 \qquad \text{and} \qquad \overline{\int_0^1} \mathcal{X}_{\mathbb{Q}}(x)\,dx = 1\,.$$

Finally, we define the Riemann integral for those functions where the lower integral does in fact equal the upper integral.

Definition 0.1.9. Let $f \in B[a,b]$. f is said to be **Riemann integrable** on $[a,b]$ if

$$\underline{\int_a^b} f(x)\,dx = \overline{\int_a^b} f(x)\,dx\,.$$

If f is Riemann integrable on $[a,b]$, we define the **Riemann integral**, denoted $\displaystyle\int_a^b f(x)\,dx$, as

$$\int_a^b f(x)\,dx = \underline{\int_a^b} f(x)\,dx = \overline{\int_a^b} f(x)\,dx\,.$$

This procedure for constructing the Riemann integral is exactly the same as what we will see for the Lebesgue integral; only what we will consider to be a partition will be changed.

Example 0.1.10. In Example 0.1.8 we found that

$$\underline{\int_0^1} \mathcal{X}_{\mathbb{Q}}(x)\,dx = 0 \qquad \text{and} \qquad \overline{\int_0^1} \mathcal{X}_{\mathbb{Q}}(x)\,dx = 1\,.$$

Consequently, the Dirichlet function is not Riemann integrable on $[0,1]$.

Finally, we denote the set of all Riemann integrable functions by $R[a,b]$. That is,

$$R[a,b] = \{f \in B[a,b] \mid f \text{ is Riemann integrable on } [a,b]\}.$$

One of the goals of Lebesgue integration is to generalize integration so that the Dirichlet function will be integrable. We also want to ensure that any function that is Riemann integrable remains integrable and the value of the integral does not change.

0.2. Criteria for Riemann Integrability

These conditions are typically discussed in texts used for a first course in real analysis. They are presented here to give us a framework on how to proceed once we encounter Lebesgue integration. More details can be found in Abbott [**1**] or Lay [**9**].

One of our first goals is to show that the lower integral is always less than or equal to the upper integral. (This may seem obvious to the reader, but one must always be able to provide a proof of the obvious. After all, mathematics occasionally has surprises in store for us!) In order to establish this claim, we will be comparing lower sums and upper sums. Although we know from Proposition 0.1.6, given a partition, that the lower sum is always less than or equal to the upper sum, we have not yet compared the lower sum for one partition with the upper sum for a possibly different partition. To do this, we need the notion of a refinement.

Definition 0.2.1. The partition P^* of $[a, b]$ is a **refinement** of the partition P if each point in P is also in P^*. P^* is a **common refinement** of the partitions P_1 and P_2 if P^* is a refinement of each of P_1 and P_2.

Given any two partitions P_1 and P_2 of $[a, b]$ we can take the points in $P_1 \cup P_2$, put them in increasing order, and form a new partition of $[a, b]$. This new partition will then be a common refinement of P_1 and P_2. Hence, given any two partitions, there always exists a common refinement.

Lemma 0.2.2. *Let $f \in B[a, b]$.*

i) *If P^* is a refinement of the partition P of $[a, b]$, then*
$$L(f, P) \leq L(f, P^*) \leq U(f, P^*) \leq U(f, P)\,.$$

ii) *If P_1 and P_2 are any two partitions of $[a, b]$, then*
$$L(f, P_1) \leq U(f, P_2)\,.$$

Proof. Let $f \in B[a,b]$. We will only give a sketch of the argument and leave some of the details to the reader.

i) The main step for this part is to show the result is true if P^* is just P with one additional point. To this end, let P be the partition

$$P = \{x_0, x_1, x_2, \ldots, x_n\},$$

where

$$a = x_0 < x_1 < x_2 < \ldots < x_n = b\,.$$

Let P^* be P with one additional point. That is,

$$P^* = \{x_0, x_1, \ldots, x_{j-1}, x^*, x_j, \ldots, x_n\},$$

where

$$x_{j-1} < x^* < x_j\,.$$

Most of the terms in $L(f, P^*)$ also appear in $L(f, P)$. To make this comparison precise, set

$$m_i = \inf_{x \in [x_{i-1}, x_i]} f(x)\,,$$

$$m_1^* = \inf_{x \in [x_{j-1}, x^*]} f(x)\,, \qquad \text{and} \qquad m_2^* = \sup_{x \in [x^*, x_j]} f(x)\,.$$

Then $m_j \le m_1^*$ and $m_j \le m_2^*$. Therefore

$$\begin{aligned}
L(f,P) &= \sum_{i=1}^{n} m_i(x_i - x_{i-1}) \\
&= m_j(x_j - x_{j-1}) + \sum_{\substack{i=1 \\ i \neq j}}^{n} m_i(x_i - x_{i-1}) \\
&= m_j(x^* - x_{j-1}) + m_j(x_j - x^*) + \sum_{\substack{i=1 \\ i \neq j}}^{n} m_i(x_i - x_{i-1}) \\
&\le m_1^*(x^* - x_{j-1}) + m_2^*(x_j - x^*) + \sum_{\substack{i=1 \\ i \neq j}}^{n} m_i(x_i - x_{i-1}) \\
&= L(f, P^*)\,.
\end{aligned}$$

By a similar argument, $U(f, P^*) \le U(f, P)$.

The general result follows by induction on the number of points added to P to get P^*.

ii) Let P^* be a common refinement of both P_1 and P_2. By part i),
$$L(f, P_1) \leq L(f, P^*) \leq U(f, P^*)$$
and
$$L(f, P^*) \leq U(f, P^*) \leq U(f, P_2)\,.$$
The result then follows. □

This lemma allows us to conclude the following corollary (something you probably believed to be true all along).

Corollary 0.2.3. *Let $f \in B[a,b]$. Then*
$$\underline{\int_a^b} f(x)\,dx \leq \overline{\int_a^b} f(x)\,dx\,.$$

Proof. By Lemma 0.2.2, given any two partitions P_1 and P_2 of $[a,b]$,
$$L(f, P_1) \leq U(f, P_2)\,.$$
Hence, $U(f, P_2)$ is an upper bound for $\{L(f,P) \mid P \text{ is a partition of } [a,b]\}$. Therefore,
$$\underline{\int_a^b} f(x)\,dx = \sup_P L(f,P) \leq U(f, P_2)\,.$$
P_2 was an arbitrary partition of $[a,b]$. Thus, $\underline{\int_a^b} f(x)\,dx$ is a lower bound for the set $\{U(f,P) \mid P \text{ is a partition of } [a,b]\}$. Consequently,
$$\underline{\int_a^b} f(x)\,dx \leq \inf_P U(f,P) = \overline{\int_a^b} f(x)\,dx,$$
as claimed. □

Finally, we prove our criteria for Riemann integrability.

Theorem 0.2.4. *Let $f \in B[a,b]$. $f \in R[a,b]$ if and only if for every $\epsilon > 0$ there is a partition P of $[a,b]$ such that*
$$U(f,P) - L(f,P) < \epsilon\,.$$

(Sooner or later you knew we had to run into that standard analysis phrase "$\epsilon > 0$"!)

Proof. Assume first f is Riemann integrable on $[a,b]$. Thus,

$$\int_a^b f(x)\,dx = \underline{\int_a^b} f(x)\,dx = \overline{\int_a^b} f(x)\,dx\,.$$

Let $\epsilon > 0$ be given. Since $\underline{\int_a^b} f(x)\,dx = \sup_P L(f,P)$, there is a partition P_1 with

$$\underline{\int_a^b} f(x)\,dx - \frac{\epsilon}{2} < L(f,P_1) \leq \underline{\int_a^b} f(x)\,dx\,.$$

In other words,

$$\int_a^b f(x)\,dx - \frac{\epsilon}{2} < L(f,P_1) \leq \int_a^b f(x)\,dx\,.$$

Similarly, there is a partition P_2 with

$$\int_a^b f(x)\,dx \leq U(f,P_2) < \int_a^b f(x)\,dx + \frac{\epsilon}{2}\,.$$

Let P be a common refinement of P_1 and P_2. Then by Lemma 0.2.2,

$$\begin{aligned} U(f,P) - L(f,P) &\leq U(f,P_2) - L(f,P_1) \\ &< \left(\int_a^b f(x)\,dx + \frac{\epsilon}{2}\right) - \left(\int_a^b f(x)\,dx - \frac{\epsilon}{2}\right) \\ &= \epsilon\,. \end{aligned}$$

Next we prove the converse. To show f is Riemann integrable, we must show that the upper integral and the lower integral are equal. Let $\epsilon > 0$. By assumption, there is a partition P of $[a,b]$ such that

$$U(f,P) - L(f,P) < \epsilon\,.$$

Thus,

$$\begin{aligned} \overline{\int_a^b} f(x)\,dx &\leq U(f,P) \\ &< L(f,P) + \epsilon \\ &\leq \underline{\int_a^b} f(x)\,dx + \epsilon\,. \end{aligned}$$

However, ϵ is arbitrary, hence,

$$\overline{\int_a^b} f(x)\,dx \leq \underline{\int_a^b} f(x)\,dx\,.$$

By Corollary 0.2.3, we also have the inequality

$$\underline{\int_a^b} f(x)\,dx \leq \overline{\int_a^b} f(x)\,dx\,.$$

Therefore, the two quantities are, in fact, equal. That is, the upper integral and lower integral are equal and f is Riemann integrable. $\square$

Example 0.2.5. Let

$$f(x) = \begin{cases} 2 & \text{if } x \neq \frac{1}{2}, \\ 5 & \text{if } x = \frac{1}{2}. \end{cases}$$

We will first show f is Riemann integrable. One option is to compute the lower integral and upper integral by comparing all lower sums and all upper sums. Instead, we will use Theorem 0.2.4. Let $\epsilon > 0$ be small. Let P_ϵ be the partition

$$P_\epsilon = \{0, \frac{1}{2} - \frac{\epsilon}{7}, \frac{1}{2} + \frac{\epsilon}{7}, 1\}\,.$$

Then

$$L(f, P_\epsilon) = 2\left(\frac{1}{2} - \frac{\epsilon}{7}\right) + 2\left(\frac{2\epsilon}{7}\right) + 2\left(\frac{1}{2} - \frac{\epsilon}{7}\right) = 2$$

and

$$U(f, P_\epsilon) = 2\left(\frac{1}{2} - \frac{\epsilon}{7}\right) + 5\left(\frac{2\epsilon}{7}\right) + 2\left(\frac{1}{2} - \frac{\epsilon}{7}\right) = 2 + \frac{6\epsilon}{7}\,.$$

Since $U(f, P_\epsilon) - L(f, P_\epsilon) = \frac{6\epsilon}{7} < \epsilon$, f is Riemann integrable.

Now that we know f is Riemann integrable, we can choose to compute either the upper integral or the lower integral. For this example, the latter is the easier computation since $L(f, P) = 2$ for any partition P. Therefore

$$\int_a^b f(x)\,dx = 2\,.$$

Another standard result concerning Riemann integration is the following. Notice that this is the first time we assume the function f is continuous.

Theorem 0.2.6. *Let f be continuous on $[a, b]$. Then $f \in R[a, b]$.*

Proof. We will use Theorem 0.2.4. Let $\epsilon > 0$ be given. Since f is continuous on the closed and bounded interval $[a, b]$, f must be uniformly continuous on $[a, b]$. Thus, there is a $\delta > 0$ so that

$$|f(x) - f(y)| < \frac{\epsilon}{b-a} \qquad \text{whenever} \qquad |x - y| < \delta\,.$$

Let $P = \{a = x_0, x_1, x_2, \ldots, x_n = b\}$ be a partition of the interval $[a, b]$ with $|x_j - x_{j-1}| < \delta$ for $j = 1, 2, \ldots, n$. It is left to the reader to verify that for this partition,

$$U(f, P) - L(f, P) < \epsilon\,.$$

(See Exercise 12.) Therefore, by Theorem 0.2.4, f is Riemann integrable. □

0.3. Properties of the Riemann Integral

The reader should be aware that at this point we have not proved the Fundamental Theorem of Calculus. In other words, we have no justification for using "antiderivatives". Moreover, the Fundamental Theorem of Calculus only applies to continuous functions. The goal of the Lebesgue integral is to apply the notion of integration to a wider range of functions. Hence, we will not cover the Fundamental Theorem in this text. But to see if you really can handle Riemann integration from just these basics, you should try proving these next results based just on the material presented thus far (without the Fundamental Theorem, of course). The first of these theorems does appear in the exercises, but the others also follow from careful consideration of upper sums and lower sums.

Theorem 0.3.1. *Let $f, g \in R[a, b]$ and $k \in \mathbb{R}$.*

i) *$f + g$ is Riemann integrable and*

$$\int_a^b (f+g)(x)\,dx = \int_a^b f(x)\,dx + \int_a^b g(x)\,dx\,.$$

ii) *kf is Riemann integrable and*

$$\int_a^b (kf)(x)\,dx = k\int_a^b f(x)\,dx\,.$$

Theorem 0.3.2. *Let $f, g \in R[a,b]$. If $f(x) \le g(x)$ for all $x \in [a,b]$, then*

$$\int_a^b f(x)\,dx \le \int_a^b g(x)\,dx\,.$$

Theorem 0.3.3. *Let f be Riemann integrable on $[a,b]$. If $[c,d] \subseteq [a,b]$, then f is Riemann integrable on $[c,d]$.*

Theorem 0.3.4. *Let f be Riemann integrable on $[a,b]$. If $c \in [a,b]$, then*

$$\int_a^b f(x)\,dx = \int_a^c f(x)\,dx + \int_c^b f(x)\,dx\,.$$

Now we have reached the point where we can prove the Fundamental Theorem of Calculus. But our goal is to move towards Lebesgue integration. The interested reader can find a proof of the Fundamental Theorem in many introductory analysis texts such as Abbott [**1**] and Lay [**9**].

0.4. Exercises

(1) Prove Proposition 0.1.6.

(2) Let

$$f(x) = \begin{cases} 1 & \text{if } 0 \le x \le 2, \\ 2 & \text{if } 2 < x \le 3. \end{cases}$$

a) Show $f \in R[0,3]$.

b) Compute $\int_0^3 f(x)\,dx$ using the definition of the Riemann integral.

(3) Let $f, g \in R[a,b]$ with $f(x) \le g(x)$ for all $x \in [a,b]$. Prove that

$$\int_a^b f(x)\,dx \le \int_a^b g(x)\,dx\,.$$

(4) Give an example of a function f with $f \notin R[0,1]$ but $f^2 \in R[0,1]$.

(5) Assume $f \in R[a,b]$.

a) Let $c \in [a,b]$. Suppose g is defined on $[a,b]$ and $g(x) = f(x)$ for all $x \ne c$. Show $g \in R[a,b]$.

b) Suppose g differs from f at a finite number of points. Show $g \in R[a, b]$.
c) Does this extend to the case where g and f differ at a countable number of points? Prove or give a counterexample.

(6) Let $\{f_n\}$ be a sequence of functions with $f_n \in R[a, b]$ for each n. Suppose the sequence $\{f_n\}$ converges uniformly to f on $[a, b]$. Show that $f \in R[a, b]$.

(7) Prove or modify and then prove: Let $f \in B[a, b]$. Define

$$f^+(x) = \begin{cases} f(x) & \text{if } f(x) \geq 0, \\ 0 & \text{otherwise.} \end{cases}$$

$$f^-(x) = \begin{cases} 0 & \text{if } f(x) \geq 0, \\ -f(x) & \text{otherwise.} \end{cases}$$

Then $f \in R[a, b]$ if and only if both $f^+ \in R[a, b]$ and $f^- \in R[a, b]$.

(8) Prove or modify and then prove: Let $f \in R[a, b]$ and $[c, d] \subseteq [a, b]$. Then

$$\int_c^d f(x)\,dx \leq \int_a^b f(x)\,dx\,.$$

(9) Prove or give a counterexample: Suppose $f \in R[a, b]$ and there exists $k > 0$ such that $f(x) \geq k$ for all $x \in [a, b]$. Then $1/f \in R[a, b]$.

(10) Prove or give a counterexample: Let $f \in R[a, b]$ and $g \in R[a, b]$. If h is a function such that $f(x) \leq h(x) \leq g(x)$ for all $x \in [a, b]$, then $h \in R[a, b]$.

(11) Let $\{r_1, r_2, \ldots, r_n, \ldots\}$ be a counting of the rational numbers in the interval $[0, 1]$. For each natural number k, define the function f_k by

$$f_k(x) = \begin{cases} 1 & \text{if } x \in \{r_1, r_2, \ldots, r_k\}, \\ 0 & \text{otherwise.} \end{cases}$$

a) Find f, the pointwise limit of the sequence $\{f_k\}$.
b) Show that $f_k \in R[0, 1]$ for each k.

c) In general, if $\{f_n\}$ is a sequence of Riemann integrable functions which converge pointwise to f, is f Riemann integrable?

(12) Verify that $U(f,P) - L(f,P) < \epsilon$ for the partition P described in Theorem 0.2.6.

(13) Let $f, g \in B[a,b]$. Show that for any partition P of $[a,b]$,

$$\begin{aligned} L(f,P) + L(g,P) &\leq L(f+g,P) \qquad \text{and} \\ U(f+g,P) &\leq L(f,P) + L(g,P)\,. \end{aligned}$$

(14) Prove part i) of Theorem 0.3.1. That is, let $f, g \in R[a,b]$. Prove that $f + g \in R[a,b]$ and

$$\int_a^b (f(x) + g(x))\,dx = \int_a^b f(x)\,dx + \int_a^b g(x)\,dx\,.$$

Suggestion: the previous exercise might be helpful.

(15) Prove part ii) of Theorem 0.3.1.

http://dx.doi.org/10.1090/stml/078/02

Chapter 1

Lebesgue Measure

There are different ways one can look at the size of a set. For example, one could look at the set $A = \{0, 2, 3, 5\}$ and say that A has four elements. From a set theoretic standpoint, this is looking at the cardinality of the set A. On the other hand, if one thinks of the members of A as points on the number line, the set A is miniscule in comparison to the real line. Think of coloring these four points blue and the rest of the line purple (or pick your two favorite colors). How much blue would you see when looking at this colored real line? Would you see anything other than purple? A single point takes up no real width on the real line. In fact, if one were asked for the length of A, the natural answer would probably be zero.

Our goal is to generalize the Riemann integral, which has its origins in the notions of length and area. We will be taking the second point of view when looking for the size of a set. Our first task then is to generalize the notion of area (and length, and volume, etc.).

1.1. Lebesgue Outer Measure

We will start our process by considering a very basic set. In $\mathbb{R}^n$ we define a **closed interval** to be a closed rectangle I where

$$I = \{x = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n \mid a_i \leq x_i \leq b_i \text{ for } i = 1, 2, \ldots, n\}.$$

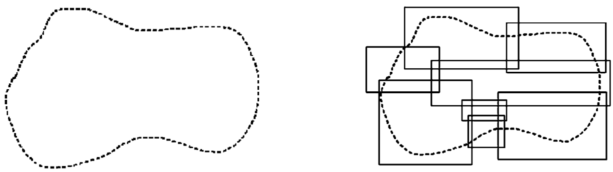

Figure 1.1. Example of a set covered by intervals.

For example,

$$I = \{(x_1, x_2) \in \mathbb{R}^2 \mid 1 \leq x_1 \leq 3, -2 \leq x_2 \leq 4\}$$

is a closed interval in $\mathbb{R}^2$, while

$$J = \{x \in \mathbb{R}^1 \mid 2 \leq x \leq 6\}$$

is a closed interval in $\mathbb{R}^1$. Notice that in $\mathbb{R}^1$ these closed intervals are, in fact, our usual closed intervals. The **volume of a closed interval** in $\mathbb{R}^n$ is

$$v(I) = \prod_{i=1}^{n} (b_i - a_i)\,.$$

Two conventions about closed intervals that will be used throughout this text are that the constants a_i and b_i are all finite, and that $a_i < b_i$ for all i.

Our strategy will be to cover a set $A \subseteq \mathbb{R}^n$ with closed intervals and add the volumes of these intervals. In order to make sense of this, we will cover A with a countable (either a finite or a countably infinite) number of intervals. This should give us an estimate (probably on the large side, but be careful not to assume this) of the volume of A. Then we use this to take the best possible estimate of the volume.

More precisely, let $S = \{I_k\}$ be a countable (finite or countably infinite) collection of closed intervals in $\mathbb{R}^n$. We say S **is a covering of A by closed intervals** if $A \subseteq \bigcup I_k$. Set $\sigma(S) = \sum v(I_k)$. If the series $\sum v(I_k)$ diverges, set $\sigma(S) = +\infty$. This idea is illustrated in Figure 1.1. In general, if S is a covering of A by intervals, we expect $\sigma(S)$ to be one of our overestimates of the volume of A. Here we have used the notation $\sum v(I_k)$, lacking upper and lower limits in the summation, to denote a countable (finite or countably infinite) sum. That is, this sum is either a finite sum or a countably infinite series.

Notice that we can always create a cover A by a countably infinite collection of intervals each with volume 1. If so, for such a covering, call it S, $\sigma(S) = +\infty$. So the set

$$\{\sigma(S) \mid S \text{ is a covering of } A \text{ by closed intervals}\}$$

always includes $+\infty$. We can still make sense of the infimum of this set if we use the convention that $s < +\infty$ for every $s \in \mathbb{R}$ so that $s \leq +\infty$ for all $s \in \mathbb{R}$ or $s = +\infty$. With this convention in mind, generalizing our usual definition of the infimum (or greatest lower bound) makes sense. That is,

$$\alpha = \inf\{\sigma(S) \mid S \text{ is a covering of } A \text{ by closed intervals}\}$$

if and only if $\alpha \leq \sigma(S)$ for every such covering S (that is, α is a lower bound) and $\beta \leq \alpha$ for any other lower bound. The difference now is that this infimum might actually equal $+\infty$. This happens when $\sigma(S) = +\infty$ for every covering S of A by closed intervals.

Also, for every S, a covering of A by closed intervals, $\sigma(S) \geq 0$. This makes 0 a lower bound for the set

$$\{\sigma(S) \mid S \text{ is a covering of } A \text{ by closed intervals}\}.$$

We finally officially define the Lebesgue outer measure of a set A.

Definition 1.1.1. Let $A \subseteq \mathbb{R}^n$. The **Lebesgue outer measure** of A is

$$m^*(A) = \inf\{\sigma(S) \mid S \text{ is a covering of } A \text{ by closed intervals}\}.$$

By definition, $m^*(A)$ is always greater than or equal to 0. Also, it follows that if S is any covering of A by closed intervals, then $m^*(A) \leq \sigma(S) \leq +\infty$. In other words, for any set $A \subseteq \mathbb{R}^n$, $0 \leq m^*(A) \leq +\infty$.

Now that we have introduced $+\infty$ as a possible value for the Lebesgue outer measure of a set, it might be worth pointing out a few things about the arithmetic of $\mathbb{R} \cup \{+\infty\}$. We can make sense of addition in that $c + (+\infty) = (+\infty) + c = +\infty$ for any real number c. Also, is it consistent if we define $(+\infty) + (+\infty) = +\infty$. But we will need to avoid any statements involving $+\infty$ and subtraction. (Think about what subtraction means: $5 - 3 = 2$ because $2 + 3 = 5$.

But $c + (+\infty) = +\infty$ for any real number c, so how does one find $+\infty - (+\infty)$?)

Remark 1.1.2. An important feature of the definition of Lebesgue outer measure is that given any set A and any given $\epsilon > 0$, there is a covering S of A by closed intervals such that

$$\sigma(S) \leq m^*(A) + \epsilon\,.$$

This is easily the case if $m^*(A) = +\infty$. In the case that $m^*(A)$ is finite, this follows by observing that $m^*(A) + \epsilon$ can no longer be a lower bound for

$$\{\sigma(S) \mid S \text{ is a covering of } A \text{ by closed intervals}\}\,.$$

Moreover, if $m^*(A)$ is finite, we can make the inequality a strict inequality. We will be using this property time and time again.

Example 1.1.3. We will compute the Lebesgue outer measure of $A = \{3\}$.

Let $\epsilon > 0$. Set $S = \{[3 - \epsilon, 3 + \epsilon]\}$. Thus,

$$0 \leq m^*(A) \leq \sigma(S) = 2\epsilon\,.$$

Since ϵ was arbitrary, it follows that $m^*(A) = 0$.

Example 1.1.4. The Lebesgue outer measure of $\emptyset$ is 0. To see this, let $\epsilon > 0$ be given. Then $S = \{[-\epsilon, \epsilon]\}$ is a covering of $\emptyset$ by closed intervals. Therefore,

$$m^*(\emptyset) \leq \sigma(S) = 2\epsilon\,.$$

Since ϵ was arbitrary, it follows that $m^*(\emptyset) = 0$.

Example 1.1.5. Let $A = [0, 1]$. The Lebesgue outer measure of A is 1. This should come as no surprise. After all, the length of this interval is 1. In fact, $S = \{[0, 1]\}$ is a covering of A by a single closed interval. Therefore,

$$m^*(A) \leq \sigma(S) = 1\,.$$

However, it is not an easy matter to prove that if S is a random covering of A by closed intervals, then $\sigma(S) \geq 1$. It is only your intuition that tells you that if we cover A by closed intervals, then the sum of the lengths of the intervals must be greater than the length of A. (Don't get me wrong—I'm not trying to tell you that your intuition is incorrect. I am only pointing out that this has not been proved. Try writing a rigorous proof. It is not easy.) We will prove in Proposition 1.1.11 that $m^*(A) = 1$.

Example 1.1.6. Let $E = \{(x, 0) \in \mathbb{R}^2 \bigm| 0 \le x \le 1\}$. Let $\epsilon > 0$ be given. Set

$$I_\epsilon = \{(x, y) \in \mathbb{R}^2 \bigm| 0 \le x \le 1, -\epsilon \le y \le \epsilon\} .$$

Then $S = \{I_\epsilon\}$ is a covering of E by closed intervals. Therefore,

$$m^*(E) \le \sigma(S) = v(I_\epsilon) = 2\epsilon .$$

Since ϵ was arbitrary, $m^*(E) = 0$.

Compare Example 1.1.5 and Example 1.1.6. We think of both as line segments of length 1. However, the first of these is a subset of $\mathbb{R}^1$ while the second is a subset of $\mathbb{R}^2$. It is this difference in dimension that accounts for the difference in the outer measure of these two seemingly similar sets. More generally, for constants $a_1, a_2, \dots, a_n$, $b_1, b_2, \dots, b_n$, $c \in \mathbb{R}$, and fixed k, the set

$$A = \{x \in \mathbb{R}^n \mid a_i \le x_i \le b_i \text{ for } i = 1, 2, \dots, n \text{ for } x \ne k \text{ and } x_k = c\}$$

has Lebesgue outer measure 0.

Example 1.1.7. We will now compute the Lebesgue outer measure of what is known as the Cantor set, or the Cantor middle-third set. Just to make sure we are all thinking of the same set we will start with a description of the Cantor set.

Set

$$C_0 = [0, 1] .$$

The next set in our construction is formed by deleting the open middle third from C_0. In other words,

$$C_1 = [0, \tfrac{1}{3}] \cup [\tfrac{2}{3}, 1] .$$

C_1 consists of two intervals of length $\frac{1}{3}$. C_2 is formed by removing the open middle third from each of these intervals. Hence,

$$C_2 = [0, \tfrac{1}{9}] \cup [\tfrac{2}{9}, \tfrac{1}{3}] \cup [\tfrac{2}{3}, \tfrac{7}{9}] \cup [\tfrac{8}{9}, 1] .$$

Continue with this process. In general, for each positive integer n, C_n consists of 2^n intervals of length $1/3^n$. C_{n+1} is formed from C_n by

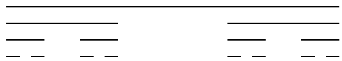

Figure 1.2. The first stages in the construction of the Cantor set.

deleting the open middle third of each of these intervals. The first few stages in this construction are illustrated in Figure 1.2. The Cantor set is what is left in the end. More precisely, the Cantor set is $\mathcal{C}$, where

$$\mathcal{C} = \bigcap_{n=0}^{\infty} C_n \,.$$

The Cantor set has many remarkable features. For example, it is a closed, uncountable set that contains no intervals.

To see that the Cantor set is uncountable, note that every $x \in [0,1]$ has a binary expansion. That is,

$$x = \sum_{n=1}^{\infty} \frac{b_n}{2^n} = \cdot_{(2)} b_1 b_2 b_3 \ldots, \qquad \text{where } b_n = 0 \text{ or } 1\,.$$

For example,

$$\begin{aligned} \frac{1}{5} &= \cdot_{(2)}001100110011\ldots \qquad \text{and} \\ \frac{1}{2} &= \cdot_{(2)}10000000\ldots = \cdot_{(2)}01111111\ldots\,. \end{aligned}$$

Some numbers, such as $\frac{1}{2}$, have more than one binary expansion. For the purpose of this example, when given a choice we will always choose the expansion that does not have a finite number of 1's. For example, we would choose $\frac{1}{2} = \cdot_{(2)}01111\ldots$.

Similarly, every $x \in [0,1]$ has a ternary expansion. That is,

$$x = \sum_{n=1}^{\infty} \frac{a_n}{3^n} = \cdot_{(3)} a_1 a_2 a_3 \ldots, \qquad \text{where } a_n = 0, 1, \text{ or } 2\,.$$

For example,

$$\begin{aligned} \frac{1}{2} &= \cdot_{(3)}11111111\ldots, \\ \frac{1}{3} &= \cdot_{(3)}10000000\ldots = \cdot_{(3)}0222222\ldots. \end{aligned}$$

Again, given a choice of more than one ternary expansion, we will choose the one that does not have a finite number of nonzero digits. Moreover, x is in the Cantor set if and only if x has a ternary expansion where none of the digits are 1.

Finally, we define a function $f\colon [0,1] \to \mathcal{C}$ as follows. For $x \in [0,1]$ write x in its binary form,

$$x = \sum_{n=1}^{\infty} \frac{b_n}{2^n},$$

choosing the expansion that does not have a finite number of 1's when given a choice. Set $f(x)$ to be

$$f(x) = \sum_{n=1}^{\infty} \frac{2b_n}{3^n}.$$

Here $f(x)$ will be a ternary expansion where every digit is 0 or 2. Thus, $f(x)$ will be in the Cantor set. The function f is known as the **Cantor function**. The Cantor function is one-to-one, which gives a one-to-one correspondence between the interval $[0,1]$ and a subset of $\mathcal{C}$. Hence, the Cantor set is uncountable.

But we are here to talk about Lebesgue outer measure. To compute the outer measure of the Cantor set, note that for every n, C_n provides us with a natural covering of C by closed intervals. As noted above, C_n consists of 2^n intervals of length $\frac{1}{3^n}$. Thus,

$$m^*(\mathcal{C}) \leq 2^n \frac{1}{3^n} = \left(\frac{2}{3}\right)^n.$$

The only way this can hold for every positive integer n is for $m^*(\mathcal{C}) = 0$.

The Cantor set shows us that it is possible for an uncountable set to have Lebesgue outer measure 0.

Our goal is to generalize the notion of volume (or area, depending on dimension). One of the properties of volume that we want to retain is that if one set is contained in a second set, then the volume of the first should be less than or equal to the volume of the second. The next proposition shows this to be true for Lebesgue outer measure as well.

Proposition 1.1.8. *If $A \subseteq B \subseteq \mathbb{R}^n$, then $m^*(A) \leq m^*(B)$.*

Proof. Let S be a covering of B by closed intervals. It follows that S is also a covering of A by closed intervals. Thus,

$$m^*(A) \leq \sigma(S),$$

where S is any covering of B by closed intervals. Hence,

$$m^*(A) \leq \inf\{\sigma(S) \mid S \text{ is a covering of } B \text{ by closed intervals}\}.$$

Therefore, $m^*(A) \leq m^*(B)$, as claimed. □

Another feature of volume we wish to retain is that the volume of the union of two sets is less than or equal to the sum of the volumes of the two sets. The next proposition asserts this to be true for outer measure. In addition, the result extends to a countable union of sets.

Proposition 1.1.9. *The following additivity properties hold for Lebesgue outer measure:*

(i) *For any two sets A and B,*

$$m^*(A \cup B) \leq m^*(A) + m^*(B).$$

(ii) *For any countable collection of sets $\{A_n\}$,*

$$m^*\left(\bigcup A_n\right) \leq \sum m^*(A_n).$$

Since we are working with a countable collection of sets, $\{A_n\}$ may be either a finite collection or a countably infinite collection. To avoid having to make separate cases, we simply write $\bigcup A_n$ to indicate that this is either the union of a finite collection of sets or the union of a countably infinite collection of sets. We use a similar convention with the summation $\sum$. We just need to be sure that any assertions we make are true in both cases.

One way to proceed with the proof of (i) is to observe that if S is any covering of A by closed intervals and T is any covering of B by closed intervals, then $S \cup T$ forms a covering of $A \cup B$ by closed intervals. It follows that

$$m^* (A \cup B) \leq \sigma(S \cup T) .$$

On the other hand, we can take the closed intervals in $S \cup T$ and look at those that came from S and those that came from T. It follows that

$$m^* (A \cup B) \leq \sigma(S \cup T) \leq \sigma(S) + \sigma(T) ,$$

keeping in mind that some intervals might appear in both S and T. We can then take first the infimum over all possible coverings of A and then the infimum over all possible coverings of B to obtain the desired result. This argument will generalize to any finite union of sets by using the principle of mathematical induction. However, in (ii) we wish to allow a countably infinite union of sets so induction will no longer apply. Hence, we will demonstrate i) in a manner that can be generalized to verify ii).

Proof. (i) If either $m^*(A)$ or $m^*(B)$ is infinite, then $m^*(A \cup B)$ is also infinite by Proposition 1.1.8. Thus, in this case the result is true by the convention that $+\infty \leq +\infty$. So, assume both $m^*(A)$ and $m^*(B)$ are finite.

Let $\epsilon > 0$ be given. By Remark 1.1.2 there exists S, a covering of A by closed intervals, with

$$m^*(A) \leq \sigma(S) < m^*(A) + \frac{\epsilon}{2} .$$

Similarly, there exists T, a covering of B by closed intervals, with

$$m^*(B) \leq \sigma(T) < m^*(B) + \frac{\epsilon}{2} .$$

Thus,

$$m^* (A \cup B) \leq \sigma (S \cup T) \leq \sigma(S) + \sigma(T) \leq m^*(A) + m^*(B) + \epsilon .$$

Since ϵ is arbitrary, it follows that

$$m^*(A \cup B) \leq m^*(A) + m^*(B),$$

as claimed.

(ii) This is proved in the same spirit as (i). As in the earlier case, if $m^*(A_n)$ is infinite for some n, the result holds. Also, if $\sum m^*(A_n)$ is $+\infty$, the Lebesgue outer measure of any set is less than or equal to $+\infty$, so again the result holds. Therefore, we will assume that all of these quantities are finite.

Let $\epsilon > 0$ be given. For each n take S_n to be a covering of A_n by closed intervals with

$$m^*(A_n) \le \sigma(S_n) < m^*(A_n) + \frac{\epsilon}{2^n}.$$

Thus,

$$\begin{aligned} m^*\left(\bigcup A_n\right) &\le \sigma\left(\bigcup S_n\right) \\ &\le \sum \sigma(S_n) \\ &\le \sum \left(m^*(A_n) + \frac{\epsilon}{2^n}\right) \\ &= \sum m^*(A_n) + \epsilon. \end{aligned}$$

As before, ϵ is arbitrary. Therefore,

$$m^*\left(\bigcup A_n\right) \le \sum m^*(A_n),$$

as claimed. □

For those who are interested in the fine details, in the last part of this proof the reason we can only claim $\sigma\left(\bigcup S_n\right) \le \sum \sigma(S_n)$ is that for distinct i and j, S_i and S_j might contain the same closed interval. This interval would only be counted once in $\sigma\left(\bigcup S_n\right)$ but would be counted more than once in $\sum \sigma(S_n)$.

Corollary 1.1.10. *If $A \subseteq B \subseteq \mathbb{R}^n$ and $m^*(B)$ is finite, then*

$$m^*(B) - m^*(A) \le m^*(B \setminus A).$$

Proof. This is Exercise 7. □

One of the drawbacks of Lebesgue outer measures is that even if sets A and B are disjoint, it is possible (if one assumes the Axiom of Choice) for $m^*(A \cup B) < m^*(A) + m^*(B)$. An example of this will be discussed later in Example 1.3.6. This is not exactly a desirable result. The remedy for this is to define Lebesgue measure, but that is a topic for a later section. In the meantime, we will establish more properties of Lebesgue outer measure.

Proposition 1.1.11. *For any closed interval $I \subseteq \mathbb{R}^n$, $m^*(I) = v(I)$.*

As mentioned in Example 1.1.5, at first glance one might not fully appreciate the reason why we need a proof of Proposition 1.1.11. After all, $S = \{I\}$ is a perfectly acceptable covering of I by closed intervals. This guarantees that $m^*(I) \leq v(I)$. It is establishing the reverse inequality that is trickier, that is, we need to show that $v(I) \leq m^*(I)$. We can accomplish this by showing that if S is any covering of I by closed intervals, then $v(I) \leq \sigma(S)$. If I is covered by a countable collection of closed intervals S, it may seem obvious that the volume of I ought to be less than or equal to $\sigma(S)$, the sum of the volumes of the intervals in S; but have you ever proved this? Remember you must be able to distinguish what you believe ought to be true versus what has or can be established. Think about writing a careful proof of this. In fact, proving this in the case that S is a *finite* collection of closed intervals is not exactly straightforward. The proof for this in $\mathbb{R}^2$ is outlined in Exercise 26 and Exercise 27. The strategies suggested can also be adapted to higher dimensions. The reason for doing these exercises is to show that the proof of this "obviously" requires careful bookkeeping and that doing them is not really necessary in order to proceed. However, one of the big messages in real analysis is that we cannot take it for granted that what may work in a finite case will also work in an infinite case. But in order to keep from getting too bogged down in the details, we will assume the results of Exercise 26 and Exercise 27. That is, we will proceed by assuming that if I is covered by a *finite* collection of closed intervals S, then $v(I) \leq \sigma(S)$. To use this to recover the intuitively clear result for the case when S is a countably infinite covering of I, we will take advantage of the compactness of I. Here, then, is a proof of Proposition 1.1.11.

Proof. We need to show that $v(I) \leq m^*(I)$. Because of the above discussion, we only need to prove that if $S = \{I_k\}$ is a countably infinite covering of I by closed intervals, then $v(I) \leq \sigma(S)$. Given

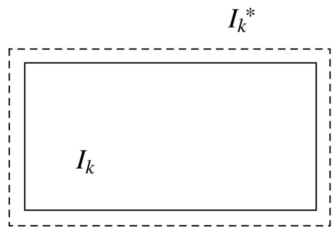

Figure 1.3. I_k^* is an expanded version of I_k.

$\epsilon > 0$, let I_k^* be an expanded version of I_k such that

$$I_k \subseteq \operatorname{int}(I_k^*)$$

and

$$v(I_k^*) \leq (1+\epsilon)v(I_k) \ .$$

This is illustrated in Figure 1.3. Then $\{\operatorname{int}(I_k^*)\}$ is an open cover of I. That is,

$$I \subseteq \bigcup_{k=1}^{\infty} \operatorname{int}(I_k^*) \ .$$

Since I is compact (I is closed and bounded), I can be covered by a finite subcover, say

$$I \subseteq \bigcup_{k=1}^{M} \operatorname{int}(I_k^*) \subseteq \bigcup_{k=1}^{M} I_k^* \, .$$

This means that $S' = \{I_k^*\}_{k=1}^{M}$ is a covering of I by a finite number of closed intervals. (It is at this point where we will assume the intuitive result concerning covering I by a finite number of closed intervals or the result of Exercise 27.) Hence,

$$\begin{aligned} v(I) &\leq \sum_{k=1}^{M} v\left(I_k^*\right) \\ &\leq (1+\epsilon)\sum_{k=1}^{M} v\left(I_k\right) \\ &\leq (1+\epsilon)\sum_{k=1}^{\infty} v\left(I_k\right) \\ &= (1+\epsilon)\sigma(S)\,. \end{aligned}$$

As ϵ is arbitrary, $v(I) \leq \sigma(S)$. Consequently,

$$v(I) \leq \inf\{\sigma(S) \mid S \text{ is a covering of } I \text{ by closed intervals}\}\,.$$

Therefore, we obtain the necessary inequality $v(I) \leq m^*(I)$. □

Example 1.1.12. We can now compute the Lebesgue outer measure of $B = [-1, 2] \cup \{3\}$. By Proposition 1.1.9,

$$m^*(B) \leq m^*([-1, 2]) + m^*(\{3\})\,.$$

By Proposition 1.1.11 and Example 1.1.3,

$$m^*([-1, 2]) = 2 - (-1) = 3 \qquad \text{and} \qquad m^*(\{3\}) = 0\,.$$

Thus,

$$m^*(B) \leq 3\,.$$

On the other hand, $[-1, 2] \subseteq B$. Hence,

$$3 = m^*([-1, 2]) \leq m^*(B)\,.$$

Consequently, $m^*(B) = 3$.

The following theorem (note this is a theorem, not just a proposition) states that any set with finite Lebesgue outer measure is contained in some open set with arbitrarily close outer measure. This may not seem like such a great feature right now. But it tells us that instead of dealing with our original set, we can use an open set with almost the same outer measure. The advantage is that we know some useful properties of open sets.

Theorem 1.1.13. *Let $A \subseteq \mathbb{R}^n$ be a set with finite outer measure. For every $\epsilon > 0$ there is an open set G such that $A \subseteq G$ and*

$$m^*(G) < m^*(A) + \epsilon.$$

Proof. Given $\epsilon > 0$ there is a covering of A by closed intervals $S = \{I_k\}$ such that

$$\sigma(S) = \sum v(I_k) < m^*(A) + \frac{\epsilon}{2}\,.$$

(Here we are using the assumption that $m^*(A)$ is finite to obtain the strict inequality.) Because S can consist of either a finite collection of intervals or a countably infinite collection, we are using the convention mentioned after the statement of Proposition 1.1.9 and are not indicating whether the summation consists of a finite number of terms or an infinite number of terms.

For each k let I_k^* be an expanded version of I_k such that

$$I_k \subseteq \operatorname{int}(I_k^*)$$

and

$$v(I_k^*) \le v(I_k) + \frac{\epsilon}{2^{k+1}}\,.$$

Set

$$G = \bigcup \operatorname{int}(I_k^*)\,.$$

By construction G is an open set. Moreover, by Proposition 1.1.9 and Proposition 1.1.11,

$$\begin{aligned} m^*(G) &\le \sum m^*(I_k^*) \\ &= \sum v(I_k^*) \\ &\le \sum \left(v(I_k) + \frac{\epsilon}{2^{k+1}} \right) \\ &\le \sigma(S) + \frac{\epsilon}{2} \\ &< m^*(A) + \epsilon\,. \end{aligned}$$

□

Corollary 1.1.14. *Let $A \subseteq \mathbb{R}^n$. For every $\epsilon > 0$ there is an open set G such that $A \subseteq G$ and*

$$m^*(G) \le m^*(A) + \epsilon.$$

Proof. If $m^*(A)$ is finite, we can use the open set G from the previous theorem. In the case that $m^*(A)$ is infinite, use $G = \mathbb{R}^n$. $\square$

One of the more tempting traps at this time is to believe that Theorem 1.1.13 and Corollary 1.1.14 tell us something about $m^*(G \setminus A)$. For example, Theorem 1.1.13 does tell us that

$$m^*(G) - m^*(A) < \epsilon\,.$$

Corollary 1.1.10 tells us that

$$m^*(G) - m^*(A) \leq m^*(G \setminus A)\,.$$

Unfortunately, this last inequality goes in the wrong direction. We are unable to make any claims about $m^*(G \setminus A)$. Trust me—in the future it might be very tempting to make such a claim, but it isn't always true.

1.2. Lebesgue Measure

As mentioned before, one of the drawbacks of outer measure is that it may be possible for $m^*(A \cup B) < m^*(A) + m^*(B)$, even when A and B are disjoint sets. This is the idea illustrated by Example 1.3.6. The way we will avoid this is to place a restriction on which subsets of $\mathbb{R}^n$ we will call measurable.

Definition 1.2.1. A set $E \subseteq \mathbb{R}^n$ is **Lebesgue measurable** if for every $\epsilon > 0$ there is an open set G so that $E \subseteq G$ and

$$m^*(G \setminus E) < \epsilon\,.$$

In this case we define the **Lebesgue measure** of E, denoted $m(E)$, to be

$$m(E) = m^*(E)\,.$$

In Chapter 4 we will encounter other measures and outer measures. Until that point, however, any time we say outer measure and measure, we are referring to Lebesgue outer measure and Lebesgue measure, respectively.

Example 1.2.2. We will show that $E = \{3\}$ is Lebesgue measurable.

Given $\epsilon > 0$ let $G = \left(3 - \frac{\epsilon}{3}, 3 + \frac{\epsilon}{3}\right)$. By Propostion 1.1.8 and by Proposition 1.1.11,

$$\begin{aligned} m^*(G \setminus E) &= m^*\left((3 - \tfrac{\epsilon}{3}, 3) \cup (3, 3 + \tfrac{\epsilon}{3})\right) \\ &\le m^*\left([3 - \tfrac{\epsilon}{3}, 3 + \tfrac{\epsilon}{3}]\right) \\ &= \frac{2\epsilon}{3} < \epsilon\,. \end{aligned}$$

Example 1.2.3. If G is an open set, then $m^*(G \setminus G) = m^*(\emptyset) = 0 < \epsilon$ for every $\epsilon > 0$. Consequently, every open set in $\mathbb{R}^n$ is Lebesgue measurable.

Example 1.2.4. Every set with Lebesgue outer measure 0 is measurable. To verify this, suppose $E \subseteq \mathbb{R}^n$ is a set with $m^*(E) = 0$. Given $\epsilon > 0$, by Theorem 1.1.13, there is an open set G containing E with

$$m^*(G) < m^*(E) + \epsilon = \epsilon\,.$$

By Proposition 1.1.8,

$$m^*(G \setminus E) \le m^*(G) < \epsilon\,.$$

Hence, E is Lebesgue measurable.

Theorem 1.2.5. *Let $\{E_k\}$ be a countable collection of Lebesgue measurable sets. Then*

$$E = \bigcup E_k$$

is Lebesgue measurable and

$$m(E) \le \sum m\left(E_k\right)\,.$$

Proof. Let $\epsilon > 0$ be given. We must show there exists an open set G containing $E = \bigcup E_k$ such that $m^*(G \setminus E) < \epsilon$.

For each k there exists an open set G_k containing E_k such that

$$m^*\left(G_k \setminus E_k\right) < \frac{\epsilon}{2^k}\,.$$

Set

$$G = \bigcup G_k\,.$$

It follows that

$$\bigcup G_k \setminus \bigcup E_k \subseteq \bigcup \left(G_k \setminus E_k\right)\,.$$

Thus, G is an open set containing E and

$$\begin{aligned} m^*(G \setminus E) &= m^*\left(\bigcup G_k \setminus \bigcup E_k\right) \\ &\leq m^*\left(\bigcup (G_k \setminus E_k)\right) \\ &\leq \sum m^*(G_k \setminus E_k) \\ &< \sum \frac{\epsilon}{2^k} \leq \epsilon\,. \end{aligned}$$

The assertion that

$$m(E) \leq \sum m(E_k)$$

follows from the definition of Lebesgue measure and Proposition 1.1.9.

□

We will use this to show that some basic sets, namely intervals, are Lebesgue measurable.

Example 1.2.6. Let $I \subseteq \mathbb{R}^n$ be a closed interval in $\mathbb{R}^n$. Then I is the union of its interior, which is an open set, and its sides. The open interior is measurable by Example 1.2.3. The sides are subsets of hyperplanes, which have Lebesgue outer measure 0. Thus, the sides have Lebesgue outer measure 0 and are Lebesgue measurable by Example 1.2.4. Consequently, I is the countable union of measurable sets. By Theorem 1.2.5, I is Lebesgue measurable. To be a little more precise about this, we write

$$\begin{aligned} I &= \{x = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n \mid a_i \leq x_i \leq b_i \text{ for } i = 1, 2, \ldots, n\} \\ &= \{x \in \mathbb{R}^n \mid a_i < x_i < b_i\} \\ &\qquad \cup \bigcup_{k=1}^{n} \{x \in I \mid x_k = a_k\} \cup \bigcup_{k=1}^{n} \{x \in I \mid x_k = b_k\}\,. \end{aligned}$$

Furthermore, since I is Lebesgue measurable,

$$m(I) = m^*(I) = v(I)$$

by Proposition 1.1.11.

We are building towards our goal of showing that Lebesgue measure has the feature we desire, that is, the measure of the union of disjoint Lebesgue measurable sets is the sum of the measures. There

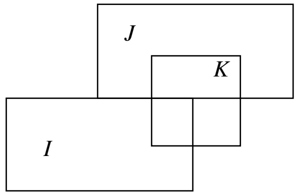

Figure 1.4. I and J are nonoverlapping, while I and K overlap, as do J and K.

are situations close to this. What if two sets abut or are adjacent to each other? To be more precise, we will consider nonoverlapping intervals.

Definition 1.2.7. Let I and J be two closed intervals in $\mathbb{R}^n$. I and J are said to be **nonoverlapping** if I and J have disjoint interiors.

In other words, two closed intervals I and j are nonoverlapping if $I \cap J$ is either empty or consists of only points that are on the boundary of both I and J.

Example 1.2.8. Let

$$\begin{aligned} I &= \{(x, y) \mid 0 \leq x \leq 4, 0 \leq y \leq 2\}, \\ J &= \{(x, y) \mid 2 \leq x \leq 6, 2 \leq y \leq 4\}, \\ K &= \{(x, y) \mid 3 \leq x \leq 5, 1 \leq y \leq 3\}. \end{aligned}$$

Then I and J are nonoverlapping. Neither I and K nor J and K are nonoverlapping.

Lemma 1.2.9. *Let $\{I_n\}_{n=1}^M$ be a finite collection of pairwise nonoverlapping closed intervals. Then*

$$m\left(\bigcup_{n=1}^{M} I_n\right) = \sum_{n=1}^{M} v(I_n).$$

Proof. It follows from Example 1.2.6 and Theorem 1.2.5 that $\bigcup_{n=1}^{M} I_n$ is measurable and

$$m\left(\bigcup_{n=1}^{M} I_n\right) \leq \sum_{n=1}^{M} v(I_n)\,.$$

We need only establish the reverse inequality.

As in Proposition 1.1.11, if $S = \{J_l\}$ is a covering of $\bigcup_{n=1}^{M} I_n$ by closed intervals, our intuition tells us that

$$\sum_{n=1}^{M} v(I_n) \leq \sum v(J_l) = \sigma(S)\,.$$

As before, actually writing a proof of this in just the finite case is quite involved. But one can carry this out by making appropriate adjustments to Exercise 26 and Exercise 27.

Therefore, similar to the proof of Proposition 1.1.11, we will assume the desired inequality is true if we cover $\bigcup_{n=1}^{M} I_n$ by a finite collection of closed intervals and use this to proceed in showing that the inequality remains true if S is a countably infinite collection of intervals. Let $S = \{J_l\}$ be a covering of $\bigcup_{n=1}^{M} I_n$ by closed intervals. Let $\epsilon > 0$. Let J_l^* be an expanded version of J_l such that

$$J_l \subseteq \operatorname{int}(J_l^*)$$

and

$$v(J_l^*) \leq (1+\epsilon) v(J_l)\,.$$

Then $\{\operatorname{int}(J_l^*)\}$ is an open cover of the compact set $\bigcup_{n=1}^{M} I_n$. Thus, for some integer N,

$$\bigcup_{k=1}^{M} I_k \subseteq \bigcup_{l=1}^{N} \operatorname{int}(J_l^*) \subseteq \bigcup_{l=1}^{N} J_l^*\,.$$

We have now covered $\bigcup_{n=1}^{M} I_n$ by a finite collection of closed intervals, so

$$\begin{aligned}\sum_{k=1}^{M} v\left(I_k\right) &\leq \sum_{l=1}^{N} v\left(J_l^*\right) \\ &\leq (1+\epsilon)\sum_{l=1}^{N} v\left(J_l\right) \\ &\leq (1+\epsilon)\sigma(S)\,.\end{aligned}$$

Since ϵ was arbitrary, it follows that

$$\sum_{k=1}^{M} v\left(I_k\right) \leq \sigma(S)$$

for any covering S of $\bigcup_{n=1}^{M} I_n$ by closed intervals. Therefore,

$$\sum_{k=1}^{M} v\left(I_k\right) \leq m\left(\bigcup_{n=1}^{M} I_n\right),$$

as required. □

So far we have shown that not only is it the case that the finite union of nonoverlapping intervals is Lebesgue measurable, we actually can find the Lebesgue measure of such a set by adding the volumes of the intervals. We will next show that any nonempty open set is the countably infinite union of nonoverlapping intervals, and we actually can find its measure by adding the volumes of the intervals.

Lemma 1.2.10. *Every nonempty open set $G \subseteq \mathbb{R}^n$ can be written as the countable union of pairwise nonoverlapping closed intervals.*

Proof. Let $G \subseteq \mathbb{R}^n$ be an open set. Divide $\mathbb{R}^n$ into nonoverlapping intervals along the hyperplanes $x_i = k$, where $k \in \mathbb{Z}$, thus creating a countable collection of closed intervals. Set aside those closed intervals which are completely contained in G. This is our first layer; we have set aside a countable number of closed intervals. For the second step, subdivide each remaining closed interval into subintervals along the hyperplanes $x_i = k/2$, where $k \in \mathbb{Z}$. This takes each

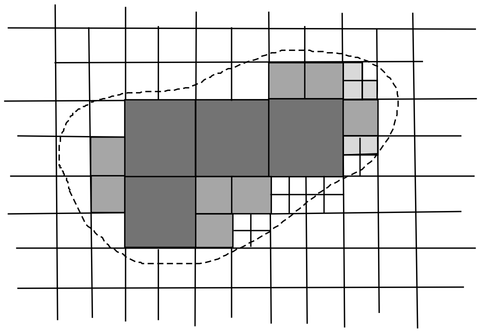

Figure 1.5. Each of the intervals contained in G are retained; the rest are subdivided.

remaining closed interval and creates 2^n nonoverlapping closed subintervals. Once again we have a countable collection of closed intervals. From these, set aside those closed intervals contained in G, again a countable number. Next, subdivide each remaining closed interval into subintervals along the hyperplanes $x_i = k/2^2$, where $k \in \mathbb{Z}$. This takes each remaining closed interval and creates 2^n nonoverlapping closed subintervals. Once again we have a countable collection of closed intervals. From these, set aside those closed intervals contained in G, again a countable number. This is the third step of the process. (See Figure 1.5.) Repeat this process ad infinitum.

We now have set aside a countable collection of nonoverlapping closed intervals $\{I_k\}$ each contained in G. It follows immediately that

$$\bigcup I_k \subseteq G\,.$$

We will show the reverse inclusion.

Let $x \in G$. Suppose to the contrary $x \notin \bigcup I_k$. Since G is open, there is an open ball B centered at x contained in G. Eventually in our process of subdividing intervals, the closed intervals under consideration will have diameters smaller than the radius of this ball. At the first such stage, x must be in one of the closed intervals which in turn will be contained in B. But B is a subset of G, so this closed interval is now contained in G. Therefore, this interval will now be placed in our collection $\{I_k\}$. But this contradicts the assumption that $x \notin \bigcup I_k$. Hence, $x \in \bigcup I_k$, so

$$G \subseteq \bigcup I_k .$$

Therefore,

$$G = \bigcup I_k ;$$

that is, G can be written as a countable union of nonoverlapping closed intervals. $\square$

Note that each of the intervals I_k is closed but G is open. If we only required a finite union of nonoverlapping closed intervals, this would mean G is a closed set. The only nonempty subset of $\mathbb{R}^n$ that is both open and closed is $\mathbb{R}^n$ itself. But $\mathbb{R}^n$ cannot be written as a finite union of closed intervals because each interval is bounded. Therefore, it must be the case that the countable union guaranteed by the previous theorem must be a countably infinite union. We can go a little further with this and say something about the measure of the open set G and the volumes of the intervals created by this lemma.

Corollary 1.2.11. *Every open set $G \subseteq \mathbb{R}^n$ can be written as a countably infinite union of nonoverlapping closed intervals $G = \bigcup_{k=1}^{\infty} I_k$ with*

$$m(G) = \sum_{k=1}^{\infty} v(I_k) .$$

Proof. By Lemma 1.2.10,

$$G = \bigcup_{k=1}^{\infty} I_k ,$$

where $\{I_k\}$ is a countable collection of nonoverlapping closed intervals. By Proposition 1.1.9 and Proposition 1.1.11,

$$m(G) \leq \sum_{k=1}^{\infty} v(I_k) .$$

Therefore, it suffices to establish the reverse inequality.

By Lemma 1.2.9, for each integer $M \in \mathbb{N}$,

$$\sum_{k=1}^{M} v(I_k) = m\left(\bigcup_{k=1}^{M} I_k\right) \leq m(G) .$$

Taking the limit as M approaches infinity establishes the reverse inequality. $\square$

We are accustomed to thinking about the distance between two points in $\mathbb{R}^n$. That is, if $x = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$ and $y = (y_1, y_2, \ldots, y_n) \in \mathbb{R}^n$ then the distance between x and y is

$$d(x, y) = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + \ldots + (x_n - y_n)^2}\,.$$

We can also define the distance between nonempty subsets of $\mathbb{R}^n$.

Definition 1.2.12. Let E_1 and E_2 be nonempty subsets of $\mathbb{R}^n$. The **distance between E_1 and E_2**, denoted $d\,(E_1, E_2)$, is defined as

$$d\,(E_1, E_2) = \inf\left\{d(x, y) \;\middle|\; x \in E_1, y \in E_2\right\}.$$

Notice that it is possible for $d\,(E_1, E_2) = 0$ even if E_1 and E_2 are disjoint.

Example 1.2.13. Consider the following two subsets of $\mathbb{R}^1$, $E_1 = [0, 1)$, and $E_2 = (1, 2]$. Then E_1 and E_2 are disjoint, but $d\,(E_1, E_2) = 0$.

It is also possible for $d\,(E_1, E_2) = 0$ even if E_1 and E_2 are disjoint closed sets.

Example 1.2.14. Let

$$E_1 = \{(x, y) \in \mathbb{R}^2 \mid x > 0, y \geq \frac{1}{x}\},$$

$$E_2 = \{(x, y) \in \mathbb{R}^2 \mid y \leq 0\}.$$

Then once again E_1 and E_2 are disjoint, but $d\,(E_1, E_2) = 0$.

Remark 1.2.15. It is a fact that if E_1 and E_2 are disjoint compact sets, then $d\,(E_1, E_2) > 0$.

The previous remark is actually a theorem resulting from the definition of compactness. Although we will not prove it here, the interested reader may find the proof an interesting exercise. The reason for stating this remark is that we will use it in Theorem 1.2.17.

We will now show that if there is a positive distance between two sets, the outer measure of the union is the sum of the outer measures of the two sets.

Lemma 1.2.16. *If* $d(E_1, E_2) > 0$*, then*

$$m^*(E_1 \cup E_2) = m^*(E_1) + m^*(E_2).$$

Proof. If one of $m^*(E_1)$ or $m^*(E_2)$ is infinite, then $m^*(E_1 \cup E_2)$ will also be infinite by Proposition 1.1.8, and the result is true. So we will assume both of these quantities are finite.

By Proposition 1.1.9 we know that

$$m^*(E_1 \cup E_2) \leq m^*(E_1) + m^*(E_2).$$

Hence, we need to verify the reverse inequality.

Let $\epsilon > 0$ be given. There is a covering $S = \{I_k\}$ of $E_1 \cup E_2$ by closed intervals so that

$$\sigma(S) = \sum v(I_k) < m^*(E_1 \cup E_2) + \epsilon.$$

Take these intervals and subdivide them, if necessary, into nonoverlapping closed subintervals with diameter smaller than $\frac{1}{2}d(E_1, E_2)$. Call this new covering of $E_1 \cup E_2$ S^*. By construction, $\sigma(S^*) = \sigma(S)$. All of the intervals in S^* have diameter less than $d(E_1, E_2)$, so none will overlap both E_1 and E_2. Let

$$\begin{aligned} S_1 &= \{J_l \in S^* \mid J_l \cap E_1 \neq \emptyset\}, \\ S_2 &= \{J_l \in S^* \mid J_l \cap E_2 \neq \emptyset\}, \\ S_3 &= \{J_l \in S^* \mid J_l \cap (E_1 \cup E_2) = \emptyset\}. \end{aligned}$$

In other words, we have taken our new covering of $E_1 \cup E_2$ by closed intervals and sorted the intervals by whether they touch E_1, or E_2, or neither. Thus, S_1 is a covering of E_1 by closed intervals, S_2 is a covering of E_2 by closed intervals, and

$$\begin{aligned} m^*(E_1) + m^*(E_2) &\leq \sigma(S_1) + \sigma(S_2) \\ &\leq \sigma(S_1) + \sigma(S_2) + \sigma(S_3) \\ &= \sigma(S^*) \\ &< m^*(E_1 \cup E_2) + \epsilon. \end{aligned}$$

Since ϵ is arbitrary, we have

$$m^*(E_1) + m^*(E_2) \leq m^*(E_1 \cup E_2).$$

Therefore,

$$m^*(E_1 \cup E_2) = m^*(E_1) + m^*(E_2),$$

as claimed. □

Although our goal is to show that the Lebesgue measure of the union of disjoint sets is the sum of the measures, we still don't have many measurable sets at our disposal. We know that sets with zero outer measure, open sets, intervals, and finite union of the above are measurable. What other sets are measurable? Are closed sets measurable? Given a specific example of a closed set, chances are that one could use what we already know to show that a specific set is measurable. However, we will cover all closed sets with the following theorem.

Theorem 1.2.17. *Every closed subset of* $\mathbb{R}^n$ *is Lebesgue measurable.*

Proof. Let $F \subseteq \mathbb{R}^n$ be a closed set. We will consider two cases.

(i) First assume that F is a bounded set. Hence, F is a compact set and $m^*(F)$ is finite. Let $\epsilon > 0$. By Theorem 1.1.13, there is an open set G containing F with

$$m^*(G) < m^*(F) + \epsilon\,.$$

Remember the caution after Corollary 1.1.14. This alone does not tell us the result we want, that $m^*(G \setminus F) < \epsilon$. It takes a surprising amount of effort to reach this conclusion.

The set $G \setminus F$ is an open set. Thus, by Lemma 1.2.10, $G \setminus F$ can be written as a countable union of nonoverlapping closed intervals, say

$$G \setminus F = \bigcup I_k\,.$$

For each positive integer N, $\bigcup_{k=1}^N I_k$ is a closed and bounded set, and so is compact. Moreover, $F \cap \bigcup_{k=1}^N I_k = \emptyset$. By Remark 1.2.15,

$$d\left(F, \bigcup_{k=1}^{N} I_k\right) > 0\,.$$

For each positive integer N,

$$\begin{aligned}\sum_{k=1}^{N} v(I_k) &= m^*\left(\bigcup_{k=1}^{N} I_k\right) \\ &= m^*\left(F \cup \bigcup_{k=1}^{N} I_k\right) - m^*(F) \\ &\leq m^*(G) - m^*(F) < \epsilon\,.\end{aligned}$$

By taking the limit as N goes to ∞,

$$\sum_{k=1}^{\infty} v(I_k) \leq \epsilon\,.$$

Therefore,

$$\begin{aligned} m^*(G \setminus F) &= m^*\left(\bigcup I_k\right) \\ &\leq \sum_{k=1}^{\infty} v(I_k) \leq \epsilon\,.\end{aligned}$$

Hence, F is measurable.

(ii) Assume F is unbounded. Set

$$\overline{B_R} = \{x \in \mathbb{R}^n \mid |x| \leq R\} \qquad \text{and}$$

$$F_N = F \cap \overline{B_N}\,.$$

Thus, for each integer N, F_N is a closed and bounded set. By (i), F_N is a measurable set for each integer N. Moreover,

$$F = \bigcup_{N=1}^{\infty} F_N$$

is a countable union of measurable sets. Therefore, F is measurable by Theorem 1.2.5. □

In this last proof we tackled the case where F is unbounded by writing F as the countable union of bounded sets. This is a common strategy; to deal with an unbounded set we simply write it as the countable union of bounded sets. We will see this technique used again.

Theorem 1.2.18. *Let $E \subseteq \mathbb{R}^n$. If E is Lebesgue measurable, then*

$$E^c = \{x = (x_1, x_2, \dots, x_n) \in \mathbb{R}^n \mid x \notin E\}$$

is measurable.

Proof. Assume $E \subseteq \mathbb{R}^n$ is measurable. Then for every positive integer k there exists an open set G_k containing E such that

$$m^* (G_k \setminus E) < \frac{1}{k} .$$

For every k,

$$G_k^c \subseteq E^c,$$

hence

$$\bigcup_{k=1}^{\infty} G_k^c \subseteq E^c .$$

Let $Z = E^c \setminus \bigcup_{k=1}^{\infty} G_k^c$ so that

$$E^c = Z \cup \bigcup_{k=1}^{\infty} G_k^c .$$

For each k, G_k^c is a closed set and hence is measurable by Theorem 1.2.17. It remains to show that Z is measurable.

We will show that Z has Lebesgue outer measure 0. Then by Example 1.2.4, Z will be Lebesgue measurable. For each integer k,

$$Z \subseteq E^c \setminus G_k^c = G_k \setminus E\,.$$

Thus,

$$m^*(Z) \leq m^*(G_k \setminus E) < \frac{1}{k}$$

for each positive integer k. It follows that $m^*(Z) = 0$.

Therefore, E^c may be written as

$$E^c = Z \cup \bigcup_{k=1}^{\infty} G_k^c,$$

the union of measurable sets. Thus, E^c is measurable. □

Now we have shown that open sets, closed sets, countable unions of measurable sets, and complements of measurable sets are measurable. One might wonder if the intersection of measurable sets is also measurable. This is indeed the case. We state the following proposition and leave the proof to the reader as an exercise.

Proposition 1.2.19. *Let $\{A_j\}$ be a countable collection of Lebesgue measurable subsets of $\mathbb{R}^n$. Then the set*

$$A = \bigcap A_j$$

is Lebesgue measurable.

Proof. This is Exercise 11. □

We have now shown that the collection of Lebesgue measurable sets contains the empty set, is closed under set complement, and is closed under countable unions. Such a collection of sets is known as a σ-algebra. We will encounter σ-algebras in Chapter 4. For now, we make the observation that since all open sets are measurable and the collection of measurable sets is closed under countable intersections, a set that is the intersection of a countable collection of open sets must be measurable. Similarly, all closed sets are measurable. Thus, a set that is the union of a countable collection of closed sets is also measurable.

Definition 1.2.20. A set H is of *type* G_δ if H is the intersection of a countable collection of open sets. A set H is of *type* F_σ if H is the union of a countable collection of closed sets.

We have already shown that all sets of type G_δ or of type F_σ are Lebesgue measurable. But to get a better feel for these sets we will look at some examples.

Example 1.2.21. The half open interval $(0, 1]$ in $\mathbb{R}^1$ is of type G_δ since

$$(0,1] = \bigcap_{n=1}^{\infty} \left(0, 1 + \tfrac{1}{n}\right) .$$

Example 1.2.22. The set

$$A = \{(x, y) \in \mathbb{R}^2 \mid 1 \le x < 2 \text{ and } 3 < y \le 5\}$$

in $\mathbb{R}^2$ is of type F_σ since

$$A = \bigcup_{n=1}^{\infty} \left\{(x, y) \in \mathbb{R}^2 \mid 1 \le x \le 2 - \tfrac{1}{n} \text{ and } 3 + \tfrac{1}{n} \le y \le 5\right\} .$$

Your next step should be to write down an example of a set and determine if it is of type G_δ or of type F_σ. Once you do so, you will discover that there are many examples of sets that are of type G_δ or of type F_σ. In fact, it is hard to imagine a set that is not one of these two types. In Chapter 4 we will show (assuming the Axiom of Choice) that there are indeed sets that are not one of these two types.

Our definition for a set to be Lebesgue measurable required that the set be contained in an open set where the excess has arbitrarily small outer measure. One can also require that the set contain a closed set where the excess has arbitrarily small measure.

Proposition 1.2.23. *Let $E \subseteq \mathbb{R}^n$ be a set. E is Lebesgue measurable if and only if for every $\epsilon > 0$ there is a closed set F with $F \subseteq E$ and $m^*(E \setminus F) < \epsilon$.*

Proof. This is Exercise 15. □

Finally, we arrive at the advantage that Lebesgue measure has over outer measure. That is, the measure of the union of a countable collection of disjoint measurable sets is the sum of the measures.

Theorem 1.2.24. *Let $\{E_k\}$ be a countable collection of pairwise disjoint Lebesgue measurable subsets of $\mathbb{R}^n$. Then*

$$m\left(\bigcup E_k\right) = \sum m\left(E_k\right) .$$

Proof. First we observe that this result clearly holds if for some $k \in \mathbb{N}$, $m(E_k)$ is infinite. Thus, we may assume that for every $k \in \mathbb{N}$, $m(E_k)$ is finite.

(i) First consider the case where each E_k is a bounded set. By Proposition 1.1.9 we have the inequality

$$m\left(\bigcup E_k\right) \leq \sum m\left(E_k\right) .$$

We need to show the reverse inequality.

Let $\epsilon > 0$ be given. By Proposition 1.2.23 (proved in Exercise 15), for each $k \in \mathbb{N}$ there is a closed set $F_k \subseteq E_k$ with

$$m^*(E_k \setminus F_k) < \frac{\epsilon}{2^k} .$$

In this case, $m(E_k) = m^*(E_k)$ is finite for each k. By Corollary 1.1.10,

$$m^*(E_k) - m^*(F_k) \leq m^*(E_k \setminus F_k) < \frac{\epsilon}{2^k}$$

for each integer k. Since $\{E_k\}$ is a countable collection of pairwise disjoint bounded sets, $\{F_k\}$ is a collection of pairwise disjoint closed sets. Moreover, since each E_k is bounded, each F_k is bounded as well. Hence, $\{F_k\}$ is a collection of pairwise disjoint compact sets. Consequently, pairwise there is a positive distance between these sets. By induction and Lemma 1.2.16, for every positive integer M,

$$m\left(\bigcup_{k=1}^{M} F_k\right) = \sum_{k=1}^{M} m\left(F_k\right) .$$

Therefore, for every positive integer M,

$$\sum_{k=1}^{M} m\left(F_k\right) \leq m\left(\bigcup E_k\right) .$$

Thus, for every M,

$$\sum_{k=1}^{M}\left(m\left(E_k\right)-\frac{\epsilon}{2^k}\right)\leq m\left(\bigcup E_k\right),$$

which implies

$$\sum_{k=1}^{M} m\left(E_k\right)\leq m\left(\bigcup E_k\right)+\epsilon\,.$$

Taking the limit as M approaches infinity yields

$$\sum m\left(E_k\right)\leq m\left(\bigcup E_k\right)+\epsilon\,.$$

Since ϵ was arbitrary, we have established the desired inequality.

(ii) We now consider the situation where it is not the case that each E_k is bounded. As in Theorem 1.2.17, we will reduce this case to our earlier bounded case. Set

$$E_{k,1}=E_k\cap\overline{B_1}$$

and

$$E_{k,j}=E_k\cap\left(\overline{B_j}\setminus\overline{B_{j-1}}\right)$$

for $j=2,3,\ldots$. Now $\{E_{k,j}\}$ is a countable collection of pairwise disjoint, bounded, measurable sets. By part (i),

$$\sum_j m(E_{k,j})=m\left(\bigcup_j E_{k,j}\right)=m(E_k)\,.$$

Also by part (i),

$$m\left(\bigcup_{k,j} E_{k,j}\right)=\sum_{k,j} m\left(E_{k,j}\right)\,.$$

Therefore,

$$\begin{aligned} m\left(\bigcup_k E_k\right) &= m\left(\bigcup_k \bigcup_j E_{k,j}\right) \\ &= \sum_k \sum_j m\left(E_{k,j}\right) \\ &= \sum_k \left(\sum_j m\left(E_{k,j}\right)\right) \\ &= \sum_k m\left(E_k\right) . \end{aligned}$$

Thus,

$$m\left(\bigcup E_k\right) = \sum m\left(E_k\right),$$

as claimed. □

1.3. A Nonmeasurable Set

In this section we will show the existence of a nonmeasurable set in $\mathbb{R}^1$. The proof relies on the Axiom of Choice and can be generalized to $\mathbb{R}^n$. The only place in later chapters that we will use the material from this section is in Remark 4.1.14 so this section can be omitted. On the other hand, to see that we needed to make the seemingly awkward definition of measurability in order to prove something like Theorem 1.2.24 is interesting in its own right.

The main tool to show the existance of a nonmeasurable set is the following lemma.

Lemma 1.3.1. *Let $E \subseteq \mathbb{R}^1$ be a measurable set. If $m(E) > 0$, including infinite measure, then the set of all arithmetic differences*

$$D_E = \{x - y \mid x, y \in E\}$$

contains an interval centered at 0*.*

The proof of this lemma is somewhat long and fairly technical. Our real goal is to show the existance of a nonmeasurable set. In order to keep from getting bogged down in the details of the lemma, we will defer its proof until later. First, we will illustrate this lemma.

Example 1.3.2. Let $E = \{-1\} \cup [2,3) \cup (4,6]$. Then $m(E) = 3$ and

$$D_E = [-7,-5) \cup [-4,4] \cup (5,7],$$

which contains an interval centered at 0.

Example 1.3.3. Although the Cantor set C has measure 0, as we will show, the corresponding set of arithmetic differences is

$$D_C = [-1,1]\,.$$

Since $C \subseteq [0,1]$, it must be the case that $D_C \subseteq [-1,1]$. We will show that the reverse inclusion holds as well.

Let $\alpha \in [-1,1]$. Then $\frac{1}{2}(\alpha+1) \in [0,1]$ has a ternary expansion, say

$$\tfrac{1}{2}(\alpha+1) = ._{(3)}c_1c_2c_3\ldots\,, \qquad \text{where } c_i = 0, 1, \text{ or } 2\,.$$

Set

$$\begin{aligned} x &= ._{(3)}a_1a_2a_3\ldots, \\ y &= ._{(3)}b_1b_2b_3\ldots, \end{aligned}$$

where

$$a_i = \begin{cases} 0 & \text{if } c_i = 0 \text{ or } 1, \\ 2 & \text{if } c_i = 2 \end{cases} \qquad \text{and} \qquad b_i = \begin{cases} 0 & \text{if } c_i = 0, \\ 2 & \text{if } c_i = 1 \text{ or } 2. \end{cases}$$

Thus, x and y are both in the Cantor set (they each have a ternary expansion consisting of only 0's and 2's). By symmetry, $(1-y)$ is also in the Cantor set. Also, $a_i + b_i = 2c_i$ for each i. Therefore,

$$x + y = 2\left(\tfrac{1}{2}(\alpha+1)\right) = \alpha + 1\,.$$

Consequently,

$$\alpha = x - (1-y)\,.$$

We have now shown that α is the difference of two members of the Cantor set.

Observe that in this case, Lemma 1.3.1 does not apply because $m(C) = 0$. Even so, the corresponding set of arithmetic differences does contain an interval centered at the origin. Take a moment to think about what this means. At first glance the Cantor set seems almost sparse. Yet the corresponding set of differences is an interval of length 2!

Example 1.3.4. Let $A = \{2, 6\}$. The corresponding set of arithmetic differences is

$$D_A = \{-4, 0, 4\}\,.$$

This set does not contain an interval. However, this does not contradict Lemma 1.3.1 since $m(A) = 0$.

The theorem that gives us a nonmeasurable set is due to Vitali.

Theorem 1.3.5. *There exists a nonmeasurable subset of* $\mathbb{R}^1$.

Proof. Define the equivalence relation $\sim$ on $\mathbb{R}$ by

$$x \sim y \qquad \text{if and only if} \qquad x - y \in \mathbb{Q}\,.$$

This partitions $\mathbb{R}$ into equivalence classes. For example, the equivalence class of 3, denoted $[3]_\sim$, is

$$\begin{aligned} [3]_\sim &= \{x \in \mathbb{R} \mid x \sim 3\} \\ &= \{x \in \mathbb{R} \mid x - 3 \in \mathbb{Q}\} \\ &= \mathbb{Q}, \end{aligned}$$

while

$$\begin{aligned} [\pi]_\sim &= \{x \in \mathbb{R} \mid x \sim \pi\} \\ &= \{\pi + q \mid q \in \mathbb{Q}\}\,. \end{aligned}$$

Two of these equivalence classes are either the same or disjoint. In fact,

$$\begin{aligned} [x]_\sim = [y]_\sim &\qquad \text{if and only if} \qquad x \sim y, \\ [x]_\sim \cap [y]_\sim = \emptyset &\qquad \text{if and only if} \qquad x \not\sim y\,. \end{aligned}$$

For example, $\left[\sqrt{2} + \frac{2}{3}\right]_\sim = [\sqrt{2}]_\sim$, while $[\pi]_\sim \cap \left[\frac{\pi}{2}\right]_\sim = \emptyset$. Moreover, there are an uncountable number of these equivalence classes.

It is here that we employ the Axiom of Choice. Form a set A by picking exactly one element from each equivalence class. We will show that A must be nonmeasurable. To the contrary, assume that A is measurable. Then either (i) $m(A) > 0$ or (ii) $m(A) = 0$.

(i) Assume A is measurable and $m(A) > 0$. By Lemma 1.3.1, the set of arithmetic differences D_A contains an interval centered at 0. However, if x and y are in different equivalence classes, then $x - y \notin \mathbb{Q}$. Hence, the only rational number in D_A is 0, contradicting Lemma 1.3.1. Therefore, it cannot be the case that $m(A) > 0$.

(ii) Assume A is measurable and $m(A) = 0$. The set of rational numbers is countable. So there exists $\{r_k\}_{k=1}^{\infty}$, a counting of $\mathbb{Q}$. That is, $\mathbb{Q} = \{r_k\}_{k=1}^{\infty}$. For each $k \in \mathbb{N}$ let

$$A_k = \{a + r_k \bigm| a \in A\} .$$

By Exercise 8, $m(A_k) = m(A) = 0$.

If $x \in \mathbb{R}$, then $x \sim a$ for some $a \in A$. After all, $x \in [x]_\sim$ and A contains exactly one element from $[x]_\sim$. Thus, $x = a + q$ for some $q \in \mathbb{Q}$. Therefore,

$$\bigcup_{k=1}^{\infty} A_k = \mathbb{R} .$$

On the other hand, if $k \neq j$, then $A_k \neq A_j$, so $\{A_k\}$ is a countable collection of pairwise disjoint measurable sets. Therefore,

$$m(\mathbb{R}) = m\left(\bigcup_{k=1}^{\infty} A_k\right) = \sum_{k=1}^{\infty} m(A_k) = 0,$$

a contradiction.

Therefore, A must be a nonmeasurable set. □

Now that we have seen how Lemma 1.3.1 is used to show the existence of a nonmeasurable set, we will turn to its proof.

Proof. Assume E is a measurable subset of $\mathbb{R}$ with positive measure. Our goal is to show that the set of arithmetic differences D_E contains an interval centered at 0. If E is not bounded, for $n \in \mathbb{N}$ set $E_n = E \cap [-n, n]$. Then $E = \bigcup E_n$ and

$$m(E) \leq \sum_{n=1}^{\infty} m(E_n) .$$

Thus, $m(E_n) > 0$ for some n. Also,

$$D_{E_n} \subseteq D_E .$$

If D_{E_n} contains an interval centered at 0, then D_E will as well. Hence, without loss of generality, we may assume that E is bounded, for if not, we simply work with the E_n with positive measure.

By Theorem 1.1.13, given any $\epsilon > 0$ there is an open set G containing E with

$$m(G) < m(E) + \epsilon\,.$$

In particular, this is the case for $\epsilon = \frac{1}{3}m(E) > 0$. That is, there is an open set G containing E with

$$m(G) < \frac{4}{3}m(E)\,.$$

By Lemma 1.2.10, the open set G is the union of countably many nonoverlapping closed intervals, say

$$G = \bigcup_{k=1}^{\infty} I_k\,.$$

Also, since $E \subseteq G$,

$$E = \bigcup_{k=1}^{\infty} (E \cap I_k)\ .$$

Moreover, by Corollary 1.2.11,

$$m(G) = \sum_{k=1}^{\infty} m(I_k)\,.$$

Next, we claim that for some k, $m(I_k) \leq \frac{4}{3}m(E \cap I_k)$. If, to the contrary, $m(I_k) > \frac{4}{3}m(E \cap I_k)$ for every k, then

$$\begin{aligned} \frac{4}{3}m(E) &= \frac{4}{3}m\left(\bigcup_{k=1}^{\infty}(E \cap I_k)\right) \\ &\leq \frac{4}{3}\sum m(E \cap I_k) \\ &< \sum m(I_k) = m(G), \end{aligned}$$

contradicting our choice of G.

We now know that $m(I_k) \leq \frac{4}{3}m(E \cap I_k)$ for some k. Denote $\mathcal{I} = I_k$ and $\mathcal{E} = E \cap I_k$. Thus $m(\mathcal{I}) \leq \frac{4}{3}m(\mathcal{E})$. Note that $D_{\mathcal{E}} \subseteq D_E$ since $\mathcal{E} \subseteq E$. We will show that $D_{\mathcal{E}}$ contains an interval centered at the origin.

We showed the existence of the interval $\mathcal{I}$ with $m(\mathcal{I}) \leq \frac{4}{3}m(\mathcal{E})$ because we know what the intervals look like, whereas we have no such knowledge about E or $\mathcal{E}$. Let $d \in \mathbb{R}$ with $|d| < \frac{1}{2}v(\mathcal{I})$. Set

$$\mathcal{I} + d = \{x + d \mid x \in \mathcal{I}\},$$

$$\mathcal{E} + d = \{x + d \mid x \in \mathcal{E}\}.$$

$\mathcal{I} + d$ is merely the interval $\mathcal{I}$ shifted by less than half of the length of $\mathcal{I}$ and will overlap $\mathcal{I}$. In fact, by our choice of d, $m\Big((\mathcal{I}+d) \cup \mathcal{I}\Big) < \frac{3}{2}m(\mathcal{I})$.

We will show $(\mathcal{E} + d) \cap \mathcal{E} \neq \emptyset$ for $d \in \mathbb{R}$. To see this, assume the contrary. By Exercise 8, $m(\mathcal{E} + d) = m(\mathcal{E})$ and $m(\mathcal{I} + d) = m(\mathcal{I})$. If $(\mathcal{E} + d) \cap \mathcal{E} = \emptyset$, then by Theorem 1.2.24,

$$\begin{aligned} 2m(\mathcal{E}) &= m(\mathcal{E} + d) + m(\mathcal{E}) \\ &= m\left((\mathcal{E} + d) \cup \mathcal{E}\right) \\ &\leq m\left((\mathcal{I} + d) \cup \mathcal{I}\right) \\ &< \frac{3}{2}m(\mathcal{I}). \end{aligned}$$

This leads to $\frac{4}{3}m(\mathcal{E}) < m(\mathcal{I})$, a contradiction.

We have established that if $d \in \mathbb{R}$ with $|d| < \frac{1}{2}v(\mathcal{I})$, then $(\mathcal{E} + d) \cap \mathcal{E} \neq \emptyset$. In other words, for some real number x, $x \in (\mathcal{E} + d) \cap \mathcal{E}$. In particular, $x \in \mathcal{E}$ and

$$x = y + d$$

for some $y \in \mathcal{E}$. Hence, $d = x - y$, where both x and y are in $\mathcal{E}$. Thus, $d \in D_{\mathcal{E}}$.

Let $\delta = \frac{1}{2}v(\mathcal{I})$. Whenever $|d| < \delta$, then $d \in D_{\mathcal{E}}$. Therefore,

$$(-\delta, \delta) \subseteq D_{\mathcal{E}} \subseteq D_E.$$

Consequently, D_E contains an interval centered at 0. $\square$

Assuming the Axiom of Choice and the existence of a nonmeasurable set, we will show that there are disjoint sets where the outer measure of the union is strictly less than the sum of the outer measures.

Example 1.3.6. By Exercise 25, there is a nonmeasurable subset A of $[0,1]$. If $m^*(A) = 0$, then A would be a measurable set by Example 1.2.4. Therefore

$$0 < m^*(A) \leq 1\,.$$

Let $\delta = m^*(A)$. The set of rational numbers in the interval $[0,1]$ is a countable set, say $\mathbb{Q} \cap [0,1] = \{r_k\}$. Hence $\{A + r_k\}$ is a countable collection of pairwise disjoint sets with $A + r_k \subseteq [0,2]$ for each k. Thus, for every N,

$$m^*\left(\bigcup_{k=1}^{N}(A + r_k)\right) \leq m^*([0,2]) = 2\,.$$

If it were the case that the outer measure of this union equalled the sum of the outer measures, then

$$N\delta = \sum_{k=1}^{N} m^*(A + r_k) = m^*\left(\bigcup_{k=1}^{N}(A + r_k)\right) \leq 2,$$

a contradiction when N is large.

1.4. Exercises

(1) Let A be a finite set of real numbers. Use the definition of outer measure to show $m^*(A) = 0$.

(2) Let A be a countable set of real numbers. Use the definition of outer measure to show $m^*(A) = 0$.

(3) Let S and T be coverings of a set A by intervals.
  a) Explain why $S \cup T$ is also a covering of A by intervals.
  b) Show that $\sigma(S \cup T) \leq \sigma(S) + \sigma(T)$.

(4) Show that for $c \in \mathbb{R}$ and fixed k, the set (known as a hyperplane in $\mathbb{R}^n$)

$$A = \{x = (x_1, x_2, \ldots, x_k, \ldots, x_n) \in \mathbb{R}^n \mid x_k = c\}$$

has Lebesgue outer measure 0.

(5) Suppose A and B are both Lebesgue measurable. Prove that if both A and B have measure zero, then $A \cup B$ is Lebesgue measurable and $m(A \cup B) = 0$.
  a) Do this directly from Definition 1.2.1.

b) Give a shorter proof by using Theorem 1.2.5.

(6) Suppose A has Lebesgue measure zero and $B \subseteq A$. Prove B is Lebesgue measurable and $m(B) = 0$.

(7) Prove Corollary 1.1.10. Give an example to show that the result does not necessarily hold if $m^*(B)$ is not finite.

(8) Let A be a subset of $\mathbb{R}$ and $c \in \mathbb{R}$. Define $A + c$ to be the set

$$A + c = \{x + c \mid x \in A\}.$$

a) Prove $m^*(A + c) = m^*(A)$.
b) Prove that $A + c$ is Lebesgue measurable if and only if A is Lebesgue measurable.

(9) Generalize the previous exercise to $\mathbb{R}^n$.

(10) Let $c > 0$. For a set $A \subseteq \mathbb{R}$, define cA by

$$cA = \{y \in \mathbb{R} \mid y = c\,x \text{ for some } x \in A\}.$$

Prove that $m^*(cA) = c\,m^*(A)$. What happens in $\mathbb{R}^n$?

(11) Prove Proposition 1.2.19.

(12) Let $Z \subseteq \mathbb{R}$ be a set with $m(Z) = 0$. Let $I = [0, 1]$. Show that $Z \times I$ is a measurable subset of $\mathbb{R}^2$ with Lebesgue measure 0.

(13) Let $Z \subseteq \mathbb{R}$ with $m(Z) = 0$. Set

$$E = \{x^2 \mid x \in Z\}.$$

a) Suppose Z is bounded, that is, $Z \subseteq [-n, n]$ for some integer n. Show that E is Lebesgue measurable and $m(E) = 0$.
b) What if Z is not bounded? Hint:

$$Z = \bigcup_{n=1}^{\infty} (Z \cap [-n, n]).$$

(14) Show that if $m^*(A) = 0$, then for any set B,

$$m^*(A \cup B) = m^*(B).$$

(15) Prove Proposition 1.2.23.

(16) Let E be a measurable subset of $\mathbb{R}^n$. Show that given $\epsilon > 0$ there is a closed set F and an open set G with $F \subseteq E \subseteq G$ and $m(G \setminus E) < \epsilon$.

(17) A measurable set $A \subseteq \mathbb{R}$ is said to have *density* d *at* x if the limit

$$\lim_{h \to 0^+} \frac{m(A \cap [x-h, x+h])}{2h}$$

exists and is equal to d. If $d = 1$, then x is called a *point of density of* A, and if $d = 0$, then x is called a *point of dispersion of* A. Find, with justification, the points of density and the points of dispersion of $A = (-1, 0) \cup (0, 1) \cup \{2\}$. What is the density at other points? Again, justify your answers. Note: x need not be an element of A.

(18) Let $\mathbb{Q}_1 = \mathbb{Q} \cap [0, 1] = \{x \in [0, 1] \mid x \text{ is rational}\}$.
 a) What is $m^*(\mathbb{Q}_1)$? Is $\mathbb{Q}_1$ Lebesgue measurable?
 b) Let $A = \{(x, y) \in \mathbb{R}^2 \;\big|\; x \in \mathbb{Q}_1, 0 \le y \le 1\}$. What is $m^*(A)$? Is A Lebesgue measurable?

(19) Let $\{E_k\}$ be a sequence of Lebesgue measurable sets with

$$E_1 \supseteq E_2 \supseteq E_3 \supseteq \dots.$$

Define the set E to be

$$E = \bigcap_{k=1}^{\infty} E_k .$$

If $m(E_1) < \infty$, show that

$$m(E) = \lim_{k \to \infty} m(E_k) .$$

Show by example that this need not be the case if we remove the assumption that $m(E_1) < \infty$.

(20) Let $\{E_k\}$ be a sequence of Lebesgue measurable sets for which the series $\sum_{k=1}^{\infty} m(E_k)$ converges. Show that

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

(21) Use the previous exercise to prove the **Borel-Cantelli Lemma**: Let $\{E_k\}$ be a sequence of Lebesgue measurable

subsets of $\mathbb{R}$ such that $\sum_{k=1}^{\infty} m(E_k)$ converges. Then almost all $x \in \mathbb{R}$ belong to at most finitely many of the E_k's.

(22) Construct a subset of $[0, 1]$ in the same manner as the Cantor set, except that at the kth stage each open interval removed has length $\delta 3^{-k}$, where δ is a fixed number strictly between 0 and 1. Show that the resulting set is Lebesgue measurable. Find, with justification, the Lebesgue measure of this "fat" Cantor set by computing the measure of its complement in $[0, 1]$.

(23) Construct a 2-dimensional Cantor set in the unit square $[0, 1] \times [0, 1]$ as follows: Subdivide the square into nine congruent subsquares and keep only the four closed corner squares, removing the cross-shaped region. Repeat this process on the four corner squares, etc. Show that the remaining set is $C \times C$, where

$$C \times C = \{(x, y) \in \mathbb{R}^2 \mid x \in C \text{ and } y \in C\}.$$

Here C is the usual Cantor set. Find, with justification, the measure of this 2-dimensional Cantor set.

(24) Let A be a subset of $\mathbb{R}^n$. Show that there is a set H of type G_δ so that

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

(25) Use a process similar to the proof of Theorem 1.3.5 to show (assuming the Axiom of Choice) there exists a nonmeasurable subset of $[0, 1]$.

(26) Let

$$I = \{(x, y) \in \mathbb{R}^2 \mid a \le x \le b, c \le y \le b\}$$

be a closed interval in $\mathbb{R}^2$. Let

$$\begin{aligned} a &= a_0 < a_1 < \ldots < a_m = b \qquad \text{and} \\ c &= c_0 < c_1 < \ldots < c_n = d. \end{aligned}$$

For $i = 1, 2, \ldots, m$ and $j = 1, 2, \ldots, n$, define the rectangle I_{ij} by

$$I_{ij} = \{(x, y) \in \mathbb{R}^2 \mid a_{i-1} \le x \le a_i, c_{j-1} \le y \le c_j\}.$$

(This can be thought of as subdividing I into subrectangles along the vertical lines $x = a_1, x = a_2, \ldots, x = a_{m-1}$ and the horizontal lines $y = c_1, y = c_2, \ldots, y = c_{n-1}$.) Using the definition of volume, prove

$$\sum_{i=1}^{m} \sum_{j=1}^{n} v(I_{ij}) = v(I)\,.$$

(The ambitious reader can generalize this to higher dimensions.)

(27) Let

$$I = \{(x, y) \in \mathbb{R}^2 \mid a \leq x \leq b, c \leq y \leq b\}$$

be a closed interval in $\mathbb{R}^2$. Let $J_1, J_2, \ldots, J_n$ be a finite collection of closed intervals that cover I. That is,

$$I \subseteq \bigcup_{k=1}^{n} J_k\,.$$

By carefully subdividing I and the J_k's into subrectangles, use the previous exercise to show that

$$v(I) \leq \sum_{k=1}^{n} v(J_k)\,.$$

http://dx.doi.org/10.1090/stml/078/03

Chapter 2

Lebesgue Integration

We know that the function $f(x) = x + 1$ for $x \in [0, 2]$ is Riemann integrable because f is a continuous function. One of the goals of Lebesgue integration over closed intervals is to extend the notion of integration to other functions while still including those functions that are Riemann integrable. For example, the Dirichlet function from Example 0.1.3,

$$\mathcal{X}_{\mathbb{Q}}(x) = \begin{cases} 1 & \text{if } x \in \mathbb{Q}, \\ 0 & \text{otherwise,} \end{cases}$$

is a function that is not Riemann integrable on the interval $[0, 1]$, but, as we shall see, is Lebesgue integrable.

2.1. Measurable Functions

Continuous functions work well with Riemann integration. If a function is continuous on an interval $[a, b]$, it is Riemann integrable over that interval. But not all Riemann integrable functions are continuous. However, Riemann integration was created with continuous functions in mind. For Lebesgue integration, we will work with a different, larger set of functions.

Definition 2.1.1. Let f be defined on $I = [a, b]$. We say f **is Lebesgue measurable on** I if for every $s \in \mathbb{R}$ the set

$$\{x \in I \mid f(x) > s\}$$

is a Lebesgue measurable set.

Note: This definition can also be extended to the case where $I = [a, +\infty)$, $I = (-\infty, b]$, or $I = (-\infty, +\infty)$.

Although we will discuss the more general concept of measurable function in Chapter 4, until that time whenever we say "measurable function", we are referring to a Lebesgue measurable function. We also use the term "measurable" to describe sets as well as functions. This definition of a measurable function does involve measurable sets, but a measurable function is not the same thing as a measurable set. When we describe something as measurable, it should be clear from the context whether we mean a measurable set or a measurable function.

Example 2.1.2. Let $f(x) = x^2$ on the interval $[-1, 5]$. Let $s \in \mathbb{R}$.

(i) If $s \geq 25$, then $\{x \in I \mid f(x) > s\} = \emptyset$, which is a Lebesgue measurable set.

(ii) If $s < 0$, then $\{x \in I \mid f(x) > s\} = [-1, 5]$, which is a Lebesgue measurable set.

(iii) If $0 \leq s < 1$, then $\{x \in I \mid f(x) > s\} = [-1, -\sqrt{s}) \cup (\sqrt{s}, 5]$, which is Lebesgue measurable set.

(iv) If $1 \leq s < 25$, then $\{x \in I \mid f(x) > s\} = (\sqrt{s}, 5]$, which is a Lebesgue measurable set.

Therefore, f is a Lebesgue measurable function on the interval $[-1, 5]$.

Definition 2.1.3. Let A be a set. The **characteristic function of** A, denoted $\mathcal{X}_A$, is the function defined by

$$\mathcal{X}_A(x) = \begin{cases} 1 & \text{if } x \in A, \\ 0 & \text{otherwise.} \end{cases}$$

By Exercise 1, if E is a subset of $[a, b]$, then the characteristic function $\mathcal{X}_E$ is a measurable function if and only if E is a measurable set.

Lebesgue measurable functions will play an important role in Lebesgue integration. Before we proceed further, we will show that our definition of a Lebesgue measurable function is equivalent to several other possible definitions.

Theorem 2.1.4. *Let f be defined on the interval I. The following four statements are equivalent:*

(i) *f is a Lebesgue measurable function.*

(ii) *For every $s \in \mathbb{R}$, the set $\{x \in I \mid f(x) \leq s\}$ is a Lebesgue measurable set.*

(iii) *For every $s \in \mathbb{R}$, the set $\{x \in I \mid f(x) < s\}$ is a Lebesgue measurable set.*

(iv) *For every $s \in \mathbb{R}$, the set $\{x \in I \mid f(x) \geq s\}$ is a Lebesgue measurable set.*

Proof. We will show that

$$\text{(i)} \Rightarrow \text{(ii)} \Rightarrow \text{(iii)} \Rightarrow \text{(iv)} \Rightarrow \text{(i)}.$$

For every $s \in \mathbb{R}$,

$$\{x \in I \mid f(x) \leq s\} = I \setminus \{x \in I \mid f(x) > s\}.$$

Thus, (i) $\Rightarrow$ (ii) by Corollary 1.1.10.

Also,

$$\{x \in I \mid f(x) < s\} = \bigcup_{k=1}^{\infty} \{x \in I \mid f(x) \leq s - \tfrac{1}{k}\}.$$

Hence, if (ii) is true, $\{x \in I \mid f(x) \leq s - \frac{1}{k}\}$ is a measurable set for every k. Thus, by Theorem 1.2.5, $\{x \in I \mid f(x) < s\}$ is a measurable set. Consequently, (ii) $\Rightarrow$ (iii).

The proof that (iii) $\Rightarrow$ (iv) is similar to that of (i) $\Rightarrow$ (ii). Likewise, the proof that (iv) $\Rightarrow$ (i) is similar to the proof of (ii) $\Rightarrow$ (iii). $\square$

Just like the set of continuous functions, the set of Lebesgue measurable functions is a vector space. Actually, it is more than closed under addition and scalar multiplication; it is also closed under multiplication and division by nonzero measurable functions. We will first deal with addition and multiplication by constants.

Theorem 2.1.5. *Suppose f is a Lebesgue measurable function on the interval I. Let $c \in \mathbb{R}$. The following two statements are true:*

(i) *The function $f(x) + c$ is a Lebesgue measurable function on I.*

(ii) *The function $cf(x)$ is a Lebesgue measurable function on I.*

Proof. Let $c \in \mathbb{R}$. Both of these statements are trivial in the case that $c = 0$. Thus, we will assume $c \neq 0$.

To see that $f(x) + c$ is a measurable function, let $s \in \mathbb{R}$. Then

$$\{x \in I \mid f(x) + c > s\} = \{x \in I \mid f(x) > s - c\}.$$

Since f is a measurable function, this must be a measurable set. Hence, $\{x \in I \mid f(x) + c > s\}$ is a measurable set so $f(x) + c$ is a measurable function.

(Note: This last sentence could have been shortened by simply saying "Hence, $\{x \in I \mid f(x) + c > s\}$ is measurable so $f(x) + c$ is measurable." It would then be left to the reader to understand that the first use of "measurable" refers to a set, while the second use of "measurable" refers to a function. This is an example of using the context to distinguish the meaning of the term "measurable".)

To show that $cf(x)$ is a measurable function is similar. The only special consideration is whether c is positive or negative. If $c > 0$, then

$$\{x \in I \mid cf(x) > s\} = \{x \in I \mid f(x) > \tfrac{s}{c}\},$$

which is a measurable set since f is a measurable function. If $c < 0$, then

$$\{x \in I \mid cf(x) > s\} = \{x \in I \mid f(x) < \tfrac{s}{c}\}.$$

In this case, by Theorem 2.1.4 this is a measurable set. Therefore, cf is a measurable function. $\square$

Part (ii) of the previous theorem shows that the collection of measurable functions on an interval is closed under scalar multiplication. The proof that this set is also closed under addition is less straightforward. While it is true that

$$\{x \in I \mid f(x) + g(x) > s\} = \{x \in I \mid f(x) > s - g(x)\},$$

this statement does not immediately show that $f + g$ is a measurable function. After all, the definition of a measurable function only guarantees that for every constant α, the set

$$\{x \in I \mid f(x) > \alpha\}$$

is measurable. The number α is some fixed real number, a constant, and cannot be a function of x. Nonetheless, as asserted earlier the set of measurable functions is closed under both addition and multiplication.

Theorem 2.1.6. *Let f and g be Lebesgue measurable functions on I. The following statements hold:*

(i) *The function $f(x) + g(x)$ is Lebesgue measurable on I.*

(ii) *The function $f(x)\, g(x)$ is Lebesgue measurable on I.*

(iii) *If $g(x) \neq 0$ for all $x \in I$, the function $\frac{f(x)}{g(x)}$ is Lebesgue measurable on I.*

Proof. To show (i), let $s \in \mathbb{R}$. We will use our earlier observation that

$$\{x \in I \mid f(x) + g(x) > s\} = \{x \in I \mid f(x) > s - g(x)\}\,.$$

To avoid the difficulty described above, let $\mathbb{Q} = \{r_k\}$ be a counting of the set of rational numbers. Then for every k,

$$\{x \in I \mid f(x) > r_k\} \qquad \text{and} \qquad \{x \in I \mid g(x) > s - r_k\}$$

are measurable sets. Hence,

$$\begin{aligned} &\{x \in I \mid f(x) > r_k\} \bigcap \{x \in I \mid g(x) > s - r_k\} \\ &\qquad = \{x \in I \mid f(x) > r_k\} \bigcap \{x \in I \mid r_k > s - g(x)\} \end{aligned}$$

is a measurable set for every k. Finally,

$$\begin{aligned} \bigcup_{k=1}^{\infty} \Big(&\{x \in I \mid f(x) > r_k\} \bigcap \{x \in I \mid r_k > s - g(x)\} \Big) \\ &\qquad\qquad = \{x \in I \mid f(x) > s - g(x)\}, \end{aligned}$$

so $f + g$ is a measurable function.

From part (i) and Theorem 2.1.5, $f - g$ is a measurable function. Also, by Exercise 9 $(f+g)^2$ and $(f-g)^2$ are measurable functions. Therefore, by Theorem 2.1.5 and Theorem 2.1.6,

$$fg = \frac{1}{4}(f+g)^2 - \frac{1}{4}(f-g)^2$$

is a Lebesgue measurable function on I. This proves part (ii).

Part (iii) is proved in Exercise 10. □

A type of function that will be useful in the future is a simple function. In Lebesgue integration such functions are the analogue of step functions in Riemann integration.

Definition 2.1.7. A **simple function** is a function φ defined on the interval I of the form

$$\varphi(x) = \sum_{k=1}^{n} a_k \mathcal{X}_{E_k}(x),$$

where a_k are constants and $\{E_k\}$ are pairwise disjoint measurable subsets of I.

The proof that a simple function is measurable is left to the reader in Exercise 15. The Dirichlet function from Example 0.1.3 is an example of a simple function.

In Section 2.2 we will define the Lebesgue integral on the set of bounded measurable functions. But thinking of Riemann integration, we know that the two functions

$$f(x) = \begin{cases} 1 & \text{if } 0 \le x < 1, \\ 2 & \text{if } 1 \le x \le 2 \end{cases} \qquad \text{and} \qquad g(x) = \begin{cases} 1 & \text{if } 0 \le x \le 1, \\ 2 & \text{if } 1 < x \le 2 \end{cases}$$

are different functions since $f(1) \neq g(1)$. Yet

$$\int_0^2 f(x)\,dx = \int_0^2 g(x)\,dx\,.$$

This is because the two functions differ only at a single point. In fact, two Riemann integrable functions will have the same Riemann integral if they differ at only a finite number of points. How far can we take this? The answer is, this is as far as we can go, at least for Riemann integration. For example, the Dirichlet function from Example 0.1.3,

$$\mathcal{X}_{\mathbb{Q}}(x) = \begin{cases} 1 & \text{if } x \in \mathbb{Q}, \\ 0 & \text{otherwise,} \end{cases}$$

is a function that differs from the zero function at only a countable number of points, yet it is not Riemann integrable.

The story is better when it comes to Lebesgue integration. It turns out that in Lebesgue integration, when comparing two functions it is not the *number* of points where they differ that matters, but the *measure* of the set where the two functions differ.

Definition 2.1.8. Let f and g be two functions defined on the interval I.

(i) We say f **equals** g **almost everywhere** on I, written

$$f(x) = g(x) \text{ a.e.} \qquad \text{or} \qquad f = g \text{ a.e.},$$

if the set $\{x \in I \mid f(x) \neq g(x)\}$ has Lebesgue measure 0.

(ii) We say f **is less than or equal to** g **almost everywhere** on I, written

$$f(x) \leq g(x) \text{ a.e.} \qquad \text{or} \qquad f \leq g \text{ a.e.},$$

if the set $\{x \in I \mid f(x) > g(x)\}$ has Lebesgue measure 0.

In a similar fashion we can define $f \geq g$ a.e., $f < g$ a.e., etc. For the most part, the results of our propositions and theorems will be true when two functions are equal almost everywhere. In other words, most of the time equal almost everywhere is almost always good enough!

Proposition 2.1.9. *Suppose f and g are two functions defined on the interval I. If f is Lebesgue measurable on I and $f = g$ a.e. on I, then g is Lebesgue measurable on I.*

Proof. Let $Z = \{x \in I \mid f(x) \neq g(x)\}$. Then Z has measure 0. Moreover, every subset of Z is a measurable set with measure 0. Given $s \in \mathbb{R}$, in order for $g(x) > s$, either $x \notin Z$ (so that $g(x) = f(x)$) and

$f(x) > s$, or $x \in Z$ and $g(x) > s$. Therefore,

$$\begin{aligned}&\{x \in I \mid g(x) > s\}\\ &\qquad = (\{x \in I \mid f(x) > s\} \setminus Z) \cup \{x \in Z \mid g(x) > s\},\end{aligned}$$

which is a combination of measurable sets. Therefore, g is a measurable function. □

If a sequence of Riemann integrable functions converges uniformly to a function, then the limit function is Riemann integrable. Moreover, the integral of the limit is the limit of the integrals. However, it is not always the case that two limit operations, such as integration and the pointwise limit of a sequence of functions, can be interchanged. When we finally arrive at Lebesgue integration, we will again be seeking sufficient conditions for the interchange of limit operations. Instead of just considering the pointwise limit of a function, we will also consider the limit superior and limit inferior, commonly known as the lim sup and lim inf, respectively.

Definition 2.1.10. Let $\{f_n\}$ be a pointwise bounded sequence of functions defined on I. That is, $\{f_n(x)\}$ is a bounded sequence of real numbers for every $x \in I$.

(i) The **lim sup** of the sequence, written $\limsup_{n\to\infty} f_n$ or denoted by f^*, is defined by

$$\limsup_{n\to\infty} f_n(x) = f^*(x) = \lim_{n\to\infty} \left(\sup\{f_n(x), f_{n+1}(x), f_{n+2}(x), \ldots\}\right) .$$

(ii) The **lim inf** of the sequence, written $\liminf_{n\to\infty} f_n$ or denoted by f_*, is defined by

$$\liminf_{n\to\infty} f_n(x) = f_*(x) = \lim_{n\to\infty} \left(\inf\{f_n(x), f_{n+1}(x), f_{n+2}(x), \ldots\}\right) .$$

Note that for each $x \in I$ and $n \in \mathbb{N}$, $\{f_n(x), f_{n+1}(x), f_{n+2}(x), \ldots\}$ is a bounded set of real numbers. Thus, both

$$\begin{aligned} M_n(x) &= \sup\{f_n(x), f_{n+1}(x), f_{n+2}(x), \ldots\},\\ m_n(x) &= \inf\{f_n(x), f_{n+1}(x), f_{n+2}(x), \ldots\}\end{aligned}$$

are defined and finite for every $x \in I$. Also, for each $x \in I$, $\{M_n(x)\}$ is a bounded decreasing sequence while $\{m_n(x)\}$ is a bounded increasing sequence. Hence, both $\limsup_{n\to\infty} f_n = \lim_{n\to\infty} M_n(x)$ and $\liminf_{n\to\infty} f_n =$

$\lim_{n\to\infty} m_n(x)$ are well defined. Moreover, it easily follows from the definition that

$$f_*(x) \leq f^*(x)$$

for all $x \in I$.

Remark 2.1.11. Another useful fact is that $f_*(x) = f^*(x)$ if and only if $\lim_{n\to\infty} f_n(x)$ exists. To prove this we use standard techniques from a first real analysis course. If you feel the need for additional practice with ϵ-N-type proofs, here is a golden opportunity. We will have occasion to use this fact but will not prove it here.

Theorem 2.1.12. *Let $\{f_n\}$ be a pointwise bounded sequence of Lebesgue measurable functions on an interval I. Then both f^* and f_* are Lebesgue measurable functions on I.*

Proof. Let

$$\begin{aligned} M_n(x) &= \sup\{f_n(x), f_{n+1}(x), f_{n+2}(x), \ldots\}, \\ m_n(x) &= \inf\{f_n(x), f_{n+1}(x), f_{n+2}(x), \ldots\}. \end{aligned}$$

The first step in this proof will be to show that for every $n \in \mathbb{N}$, both $M_n(x)$ and $m_n(x)$ are measurable functions.

Fix $n \in \mathbb{N}$ and let $s \in \mathbb{R}$. We will show that $\{x \in I \mid M_n(x) > s\}$ is a measurable set. Note that $M_n(x) > s$ if and only if $f_k(x) > s$ for some $k \geq n$. Therefore,

$$\{x \in I \mid M_n(x) > s\} = \bigcup_{k=n}^{\infty} \{x \in I \mid f_k(x) > s\},$$

a union of measurable sets. Thus, M_n is a measurable function on I.

In a similar fashion we will show that $\{x \in I \mid m_n(x) < s\}$ is a measurable set and use Theorem 2.1.4. Note that $m_n(x) < s$ if and only if $f_k(x) < s$ for some $k \geq n$. Therefore,

$$\{x \in I \mid m_n(x) < s\} = \bigcup_{k=n}^{\infty} \{x \in I \mid f_k(x) < s\},$$

a union of measurable sets. Therefore, m_n is a measurable function on I.

To complete the proof, we observe that for each $x \in I$, the sequence $\{M_n(x)\}$ is a nonincreasing bounded sequence while the sequence $\{m_n(x)\}$ is a nondecreasing bounded sequence. Therefore,

$$\begin{aligned} f^*(x) &= \lim_{n\to\infty} M_n(x) = \inf\{M_n(x)\}, \\ f_*(x) &= \lim_{n\to\infty} m_n(x) = \sup\{m_n(x)\} \end{aligned}$$

are measurable functions by the preceding argument. □

Corollary 2.1.13. *Let $\{f_n\}$ be a sequence of Lebesgue measurable functions on I that converges pointwise to f. Then the function f is Lebesgue measurable on I.*

Proof. In this case, $f^* = f_* = f$. Therefore, f is a measurable function. □

The following corollary is a typical example of where almost everywhere is good enough.

Corollary 2.1.14. *Let $\{f_n\}$ be a sequence of Lebesgue measurable functions on I. If f is a function defined on I with $\lim_{n\to\infty} f_n(x) = f(x)$ a.e., then f is Lebesgue measurable on I.*

Of course, you figured out that $\lim_{n\to\infty} f_n(x) = f(x)$ a.e. means

$$m\left(\{x \in I \mid \lim_{n\to\infty} f_n(x) \neq f(x)\}\right) = 0\,,$$

didn't you?

Proof. Let

$$Z = \{x \in I \mid \lim_{n\to\infty} f_n(x) \neq f(x)\}\,.$$

For each $n \in \mathbb{N}$, set

$$g_n(x) = \begin{cases} f_n(x) & \text{if } x \notin Z, \\ 0 & \text{if } x \in Z \end{cases}$$

and define

$$g(x) = \begin{cases} f(x) & \text{if } x \notin Z, \\ 0 & \text{if } x \in Z. \end{cases}$$

By definition, $g_n = f_n$ a.e., and hence is measurable on I by Corollary 2.1.13. Also, by construction $g(x) = \lim_{n\to\infty} g_n(x)$ for all $x \in I$ and so is measurable on I, again by Corollary 2.1.13. But $g = f$ a.e., and hence f is measurable on I by Proposition 2.1.9. □

2.2. The Lebesgue Integral

As mentioned in the preface, there are several routes one can take to define the Lebesgue integral. One option is to essentially define the Lebesgue integral for nonnegative functions as the measure of the region bounded above by the graph of the function and below by the x-axis. Another route is to approach the integral via simple functions, an approach we will use in Chapter 4. We have chosen an approach that mimics our outline of the Riemann integral. Naturally, in the end these differing approaches arrive at the same mathematical idea.

We will start by investigating this integral on bounded functions on a closed interval,

$$B[a,b] = \{f \mid f \text{ is a bounded function on } [a,b]\}.$$

The difference between Riemann integration and Lebesgue integration starts with a more general partition of the interval $[a,b]$ known as a measurable partition.

Definition 2.2.1. A **measurable partition** of $[a,b]$ is $P = \{E_j\}_{j=1}^n$, a finite collection of subsets of $[a,b]$, such that

(i) E_j is a measurable set for each j,

(ii) $\displaystyle\bigcup_{j=1}^n E_j = [a,b]$, and

(iii) $m(E_i \cap E_j) = 0$ if $i \neq j$.

Instead of partitioning the interval $[a,b]$ into subintervals, as we do in Riemann integration, we are partitioning $[a,b]$ into measurable sets. Condition (iii) is the analog of requiring nonoverlapping intervals.

Example 2.2.2. Let

$$P = \{C, [0,1], [1,2]\},$$

where C denotes the Cantor set. It is a straightforward exercise to check that P is a measurable partition of the interval $[0,2]$.

Remark 2.2.3. Suppose $P = \{E_j\}_{j=1}^n$ is a measurable partition of $[a,b]$. Let

$$\begin{aligned} F_1 &= E_1 \quad \text{and} \\ F_k &= E_k \setminus \bigcup_{i=1}^{k-1} E_i \qquad \text{for } k = 2, 3, \ldots, n\,. \end{aligned}$$

Then $\{F_k\}_{k=1}^n$ is a measurable partition of $[a,b]$ consisting of pairwise disjoint sets. Consequently,

$$\sum_{j=1}^n m(F_j) = m([a,b])\,.$$

But $m(F_j) = m(E_j)$. Therefore,

$$\sum_{j=1}^n m(E_j) = m([a,b]) = b - a\,.$$

The following definitions parallel those relating to Riemann integration.

Definition 2.2.4. Let $f \in B[a,b]$ and $P = \{E_j\}_{j=1}^n$ be a measurable partition of $[a,b]$.

(i) The **upper sum** $U[f,P]$ is

$$U[f,P] = \sum_{j=1}^n M_j\, m(E_j),$$

where $M_j = \sup_{x \in E_j} f(x)$.

(ii) The **lower sum** $L[f,P]$ is

$$L[f,P] = \sum_{j=1}^n m_j\, m(E_j),$$

where $m_j = \inf_{x \in E_j} f(x)$.

We are using only a slight change in notation to help distinguish a lower sum associated with a Riemann-type partition, $L(f,P)$, and the one encountered here, $L[f,P]$. Generally, it should be clear from context whether we are in a Riemann setting or a Lebesgue setting.

In the case that there might be confusion, we are using parentheses for the former and square brackets for the latter.

Next we turn to the notion of upper integral and lower integral. The idea is that each upper sum is greater than or equal to the desired result, while the lower integral is less than or equal to the desired result. So in some sense, we want to find a minimum upper sum or maximum lower sum, but these don't necessarily exist. Instead, we will look at the set of all possible upper sums and take the infimum (or greatest lower bound) of that set. We will treat the set of all possible lower sums in a similar fashion by taking the supremum (or least upper bound). This is what we describe in the next definition. We use $\mathcal{P}$ to denote the collection of all possible measurable partitions, or $\mathcal{P} = \{P \mid P \text{ is a measurable partition of } [a, b]\}$. This means that writing $P \in \mathcal{P}$ is another way of saying that P is a measurable partition of $[a, b]$.

Definition 2.2.5. Let $f \in B[a, b]$.

(i) The **upper integral**, written $\overline{\int_a^b} f$, is

$$\overline{\int_a^b} f = \inf_{P \in \mathcal{P}} U[f, P]\,.$$

(ii) The **lower integral**, written $\underline{\int_a^b} f$, is

$$\underline{\int_a^b} f = \sup_{P \in \mathcal{P}} L[f, P]\,.$$

Here $\inf_{P \in \mathcal{P}} U[f, P]$ means $\inf\{U[f, P] \mid P \in \mathcal{P}\}$; similarly for $\sup_{P \in \mathcal{P}} L[f, P]$.

(iii) If $\overline{\int_a^b} f = \underline{\int_a^b} f$, we say f **is Lebesgue integrable** and write the Lebesgue integral as

$$\int_a^b f = \overline{\int_a^b} f = \underline{\int_a^b} f\,.$$

The computation of the upper sum and lower sum for a given partition is straightforward.

Example 2.2.6. For $x \in [0,1]$, let

$$\mathcal{X}_{\mathbb{Q}} = \begin{cases} 1 & \text{if } x \in \mathbb{Q}, \\ 0 & \text{otherwise} \end{cases}$$

and P be the partition $P = \{\mathbb{Q} \cap [0,1], [0,1] \setminus \mathbb{Q}\}$. Then P is a measurable partition of $[0,1]$. Setting $E_1 = \mathbb{Q} \cap [0,1]$ and $E_2 = [0,1] \setminus \mathbb{Q}$, we have

$$\begin{aligned} M_1 = \sup_{x \in E_1} \mathcal{X}_{\mathbb{Q}}(x) &= 1 = \inf_{x \in E_1} \mathcal{X}_{\mathbb{Q}}(x) = m_1, \\ M_2 = \sup_{x \in E_2} \mathcal{X}_{\mathbb{Q}}(x) &= 0 = \inf_{x \in E_2} \mathcal{X}_{\mathbb{Q}}(x) = m_2 . \end{aligned}$$

Therefore,

$$\begin{aligned} U[\mathcal{X}_{\mathbb{Q}}, P] &= \sum_{j=1}^{2} M_j \, m(E_j) = 0, \\ L[\mathcal{X}_{\mathbb{Q}}, P] &= \sum_{j=1}^{2} m_j \, m(E_j) = 0 . \end{aligned}$$

On the other hand, if P^* is the partition $P^* = \{[0, \frac{1}{2}], [\frac{1}{2}, 1]\}$, then P^* is a measurable partition of $[0,1]$. With $E_1 = [0, \frac{1}{2}]$ and $E_2 = [\frac{1}{2}, 1]$, we have

$$\begin{aligned} M_1 = \sup_{x \in E_1} \mathcal{X}_{\mathbb{Q}}(x) &= 1 = \sup_{x \in E_2} \mathcal{X}_{\mathbb{Q}}(x) = M_2, \\ m_1 = \inf_{x \in E_2} \mathcal{X}_{\mathbb{Q}}(x) &= 0 = \inf_{x \in E_2} \mathcal{X}_{\mathbb{Q}}(x) = m_2 . \end{aligned}$$

This time

$$\begin{aligned} U[\mathcal{X}_{\mathbb{Q}}, P] &= \sum_{j=1}^{2} M_j \, m(E_j) = 1, \\ L[\mathcal{X}_{\mathbb{Q}}, P] &= \sum_{j=1}^{2} m_j \, m(E_j) = 0 . \end{aligned}$$

But to find the Lebesgue integral directly from the definition is a somewhat long and tedious process. In this example, we only compared the upper and lower sums for two partitions. We have yet to compare the values of all possible upper sums in order to find the upper integral. Although it follows immediately from the definition that for every partition P, $L[f,P] \le U[f,P]$, it is not so immediate that $\underline{\int_a^b} f \le \overline{\int_a^b} f$. We need to be able to compare the lower sum with respect to one partition with the upper sum with respect to a possibly different partition.

Definition 2.2.7. Let $P = \{E_j\}_{j=1}^n$ and $P^* = \{F_k\}_{k=1}^m$ be two measurable partitions of $[a,b]$. We say P^* **is a refinement of** P if for every k, there is a j such that $F_k \subseteq E_j$. If P_1, P_2, and P^* are measurable partitions of $[a,b]$ and P^* is a refinement of both P_1 and P_2, we say P^* **is a common refinement of** P_1 **and** P_2.

Given any two measurable partitions $P_1 = \{E_j\}_{j=1}^n$ and $P_2 = \{F_k\}_{k=1}^m$, there always exists a common refinement. For example,

$$P^* = \{E_j \cap F_k \mid E_j \in P_1 \text{ and } F_k \in P_2\}$$

is a common refinement of P_1 and P_2. We will use this notion of a common refinement to establish the next lemma.

Lemma 2.2.8. *Let $f \in B[a,b]$.*

(i) *For any two measurable partitions P_1 and P_2 of $[a,b]$,*

$$L[f,P_1] \le U[f,P_2]\,.$$

(ii) *Consequently,*

$$\underline{\int_a^b} f \le \overline{\int_a^b} f\,.$$

Proof. First we will establish (i). Let P^* be a common refinement of P_1 and P_2. Since P^* is a refinement of both P_1 and P_2, by Exercise 12,

$$L[f,P_1] \le L[f,P^*] \qquad \text{and} \qquad U[f,P^*] \le U[f,P_2]\,.$$

However, by our earlier observation,

$$L[f,P^*] \le U[f,P^*]\,.$$

Therefore,

$$L[f,P_1] \le U[f,P_2],$$

as claimed.

Next, let P' be a measurable partition of $[a,b]$. By part (i), then,

$$\sup_{P\in\mathcal{P}} L[f,P] \le U[f,P'] \,.$$

Hence,

$$\underline{\int_a^b} f \le U[f,P']$$

for any measurable partition P' of the interval $[a,b]$. Therefore

$$\underline{\int_a^b} f \le \inf_{P\in\mathcal{P}} U[f,P] = \overline{\int_a^b} f,$$

as claimed. □

We will use this result to find the Lebesgue integral of $\mathcal{X}_{\mathbb{Q}}$ on the interval $[0,1]$.

Example 2.2.9. As in Example 2.2.6 for $x\in[0,1]$, let

$$\mathcal{X}_{\mathbb{Q}} = \begin{cases} 1 & \text{if } x\in\mathbb{Q}, \\ 0 & \text{otherwise.} \end{cases}$$

In Example 2.2.6 we found a measurable partition P where

$$L[\mathcal{X}_{\mathbb{Q}},P] = U[\mathcal{X}_{\mathbb{Q}},P] = 0\,.$$

Therefore,

$$\sup_{P'\in\mathcal{P}} L[\mathcal{X}_{\mathbb{Q}},P'] \ge 0\,.$$

On the other hand, by Lemma 2.2.8,

$$\sup_{P'\in\mathcal{P}} L[\mathcal{X}_{\mathbb{Q}},P'] \le U[\mathcal{X}_{\mathbb{Q}},P] = 0\,.$$

Thus,

$$\underline{\int_0^1} \mathcal{X}_{\mathbb{Q}} \le 0\,.$$

Combining our two inequalities we see that

$$\underline{\int_0^1} \mathcal{X}_{\mathbb{Q}} = 0\,.$$

In a similar fashion,

$$\overline{\int_0^1} \mathcal{X}_{\mathbb{Q}} = 0\,.$$

As a result, $\mathcal{X}_{\mathbb{Q}}$ is Lebesgue integrable on the interval $[0,1]$ and

$$\int_0^1 \mathcal{X}_{\mathbb{Q}} = 0\,.$$

We will now compare Lebesgue integration with Riemann integration. In particular, we will show that if f is Riemann integrable on $[a,b]$, then f is Lebesgue integrable on $[a,b]$.

Proposition 2.2.10. *Let $f \in B[a,b]$. If f is Riemann integrable on $[a,b]$, then f is Lebesgue integrable on $[a,b]$.*

Proof. For any Riemann partition of $[a,b]$,

$$P_R = \{a = x_0 < x_1 < x_2 < \ldots < x_n = b\}\,,$$

we form a corresponding measurable partition of $[a,b]$ by setting

$$P_L = \{[x_0,x_1],[x_1,x_2],\ldots,[x_{n-1},x_n]\}\,.$$

Note that

$$U(f,P_R) = \sum_{i=1}^{n} M_i(x_i - x_{i-1}) = \sum_{i=1}^{n} M_i m([x_{i-1},x_i]) = U[f,P_L],$$

where $M_i = \sup\limits_{x\in[x_{i-1},x_i]} f(x)$. These partitions describe the partitions under consideration in Riemann integration. However, for Lebesgue integration, there are many more measurable partitions to consider. Thus,

$$\inf_{P_R} U(f,P_R) \geq \inf_{P\in\mathcal{P}} U[f,P]$$

and, consequently,

$$\overline{\int_a^b} f(x)\,dx \geq \overline{\int_a^b} f\,.$$

Here the first integral denotes the upper Riemann integral, while the second is our upper Lebesgue integral.

It can be shown in a similar fashion that

$$\underline{\int_a^b} f(x)\,dx \leq \underline{\int_a^b} f\,.$$

Combining these inequalities, we see that

$$\underline{\int_a^b} f(x)\,dx \le \underline{\int_a^b} f \le \overline{\int_a^b} f \le \overline{\int_a^b} f(x)\,dx\,.$$

Hence, if f is Riemann integrable, that is, $\underline{\int_a^b} f(x)\,dx = \overline{\int_a^b} f(x)\,dx$, it must be the case that $\underline{\int_a^b} f = \overline{\int_a^b} f$. Therefore, if f is Riemann integrable on $[a,b]$, then f is Lebesgue integrable on $[a,b]$. Moreover,

$$\int_a^b f = \int_a^b f(x)\,dx,$$

where the second integral is the Riemann integral. $\square$

The next result echoes Theorem 0.2.4. Although we are now dealing with Lebesgue integration, the proof is much the same as the corresponding result for Riemann integration.

Lemma 2.2.11. *Let $f \in B[a,b]$. Then f is Lebesgue integrable if and only if for every $\epsilon > 0$ there is a measurable partition P such that*

$$U[f,P] - L[f,P] < \epsilon\,.$$

Proof. Assume first that f is Lebesgue integrable on $[a,b]$. Let $\epsilon > 0$ be given. By the definition of the lower integral, there is a measurable partition P_1 of $[a,b]$ such that

$$\underline{\int_a^b} f - \tfrac{\epsilon}{2} < L[f,P_1]\,.$$

Likewise, there is a measurable partition P_2 with

$$\overline{\int_a^b} f + \tfrac{\epsilon}{2} > U[f,P_2]\,.$$

Let P be a common refinement of P_1 and P_2. Then

$$\begin{aligned} U[f,P] - L[f,P] &\leq U[f,P_2] - L[f,P_1] \\ &< \left(\overline{\int_a^b} f + \tfrac{\epsilon}{2}\right) - \left(\underline{\int_a^b} f - \tfrac{\epsilon}{2}\right) \\ &= \left(\overline{\int_a^b} f - \underline{\int_a^b} f\right) + \epsilon \,. \end{aligned}$$

But f is Lebesgue integrable, and hence $\overline{\int_a^b} f = \underline{\int_a^b} f$. Therefore,

$$U[f,P] - L[f,P] < \epsilon \,.$$

Now we will assume that for every $\epsilon > 0$ there is a measurable partition P such that $U[f,P] - L[f,P] < \epsilon$ and show that $\overline{\int_a^b} f = \underline{\int_a^b} f$. Let $\epsilon > 0$. There is a partition P with

$$U[f,P] - L[f,P] < \epsilon \,.$$

So, for this partition

$$U[f,P] < L[f,P] + \epsilon \,.$$

Since $\overline{\int_a^b} f \leq U[f,P]$ and $\underline{\int_a^b} f \geq L[f,P]$ for every measurable partition P, it follows that

$$\overline{\int_a^b} f \leq \underline{\int_a^b} f + \epsilon \,.$$

Since ϵ was arbitrary, this proves that

$$\overline{\int_a^b} f \leq \underline{\int_a^b} f \,.$$

By Lemma 2.2.8, $\underline{\int_a^b} f \leq \overline{\int_a^b} f$, and hence

$$\overline{\int_a^b} f = \underline{\int_a^b} f \,.$$

Therefore, f is Lebesgue integrable. $\square$

We will now show that every bounded measurable function is integrable. Compare this theorem to Theorem 0.2.6.

Theorem 2.2.12. *Let $f \in B[a,b]$. If f is measurable on $[a,b]$, then f is Lebesgue integrable on $[a,b]$.*

Proof. Assume f is a bounded, measurable function on $[a,b]$ and let $\epsilon > 0$. Because f is bounded, there is a positive number M so that $|f(x)| < M$ for all $x \in [a,b]$.

We will now form a measurable partition of $[a,b]$ that satisfies the conditions of Lemma 2.2.11. To do this, let $-M = y_0 < y_1 < y_2 < \ldots < y_n = M$, where $y_1, y_2, \ldots, y_n$ are chosen so that $y_i - y_{i-1} < \frac{\epsilon}{b-a}$ for $i = 1, 2, \ldots, n$. Set

$$E_i = \{x \in [a,b] \mid y_{i-1} \leq f(x) < y_i\} = f^{-1}\left([y_{i-1}, y_i)\right)$$

for $i = 1, 2, \ldots, n$. Since f is measurable and

$$\begin{aligned} E_i &= \{x \in [a,b] \mid y_{i-1} \leq f(x) < y_i\} \\ &= \{x \in [a,b] \mid f(x) < y_i\} \setminus \{x \in [a,b] \mid f(x) < y_{i-1}\}, \end{aligned}$$

E_i is measurable for $i = 1, 2, \ldots, n$. Also,

$$\bigcup_{i=1}^{n} E_i = [a,b]\,.$$

Hence $P = \{E_1, E_2, \ldots, E_n\}$ is a measurable partition of $[a,b]$. Moreover, $\{E_1, E_2, \ldots, E_n\}$ is a pairwise disjoint collection of measurable sets, and hence

$$\sum_{i=1}^{n} m(E_i) = m\left(\bigcup_{i=1}^{n} E_i\right) = m([a,b]) = b - a\,.$$

Finally,

$$\begin{aligned} U[f,P]-L[f,P] &\leq \sum_{i=1}^{n} y_i m(E_i) - \sum_{i=1}^{n} y_{i-1} m(E_i) \\ &= \sum_{i=1}^{n} (y_i - y_{i-1})\, m(E_i) \\ &< \frac{\epsilon}{b-a} \sum_{i=1}^{n} m(E_i) \\ &= \frac{\epsilon}{b-a}(b-a) = \epsilon. \end{aligned}$$

Therefore, by Lemma 2.2.11, f is Lebesgue integrable on $[a,b]$. □

Notice that in the proof of Theorem 0.2.6 we used continuity to make the partition fine enough so that the function did not change too much in each subinterval. On the other hand, in Theorem 2.2.12 we used measurability to divide the range into small intervals and took the inverse images of this division to form our partiton of the interval $[a,b]$. In this way, Theorem 0.2.6 and Theorem 2.2.12 illustrate the difference between Riemann integration and Lebesgue integration. If one thinks of a row of stacks of pennies, not necessarily of the same height, as analogous to the area under the graph of a function, the Riemann approach to counting the coins is to move from left to right counting each stack and keeping a running total until reaching the end. The Lebesgue approach is to keep a running total vertically, that is, count how many stacks have a single penny, add to that two times the number of stacks with two pennies, add to that three times the number of stacks with three pennies, etc., until all stacks have been counted.

We know that there are discontinuous functions that are Riemann integrable. One might wonder if it is possible to find a bounded function that is Lebesgue integrable but is not a measurable function. However, unlike Riemann integration, the converse of Theorem 2.2.12 is true.

Theorem 2.2.13. *Let $f \in B[a,b]$. If f is Lebesgue integrable on $[a,b]$, then f is measurable on $[a,b]$.*

Before proving this theorem, we need a lemma.

Lemma 2.2.14. *Let $f \in B[a,b]$. Suppose f is measurable with $f \geq 0$ a.e. in $[a,b]$ and that $\int_a^b f = 0$. Then $f = 0$ a.e. in $[a,b]$.*

Proof. Set
$$g(x) = \begin{cases} f(x) & \text{if } f(x) \geq 0, \\ 0 & \text{otherwise} \end{cases}$$
so that $(g - f) = 0$ a.e. on $[a,b]$. It is easy to show that $g \in B[a,b]$. By Proposition 2.1.9, g is measurable. By Exercise 14, $\int_a^b (g-f) = 0$. Thus, Exercise 13 implies
$$\int_a^b g = \int_a^b (g-f) + \int_a^b f = 0\,.$$
So g has the properties that $g(x) \geq 0$ for all $x \in [a,b]$, $g = f$ a.e. on $[a,b]$, and $\int_a^b g = 0$. Therefore, without loss of generality, we may assume that $f(x) \geq 0$ for all $x \in [a,b]$.

By Exercise 8 if the set
$$E = \{x \in [a,b] \mid f(x) > 0\}$$
has positive measure, then for some positive integer n, the set
$$E_n = \{x \in [a,b] \mid f(x) > \tfrac{1}{n}\}$$
has positive measure. Let P be the partition $P = \{E_n, [a,b] \setminus E_n\}$. Then
$$L[f,P] \geq \tfrac{1}{n} m(E_n) > 0\,.$$
This contradicts the assumption that $\int_a^b f = 0$. Therefore,
$$E = \{x \in [a,b] \mid f(x) > 0\} = \{x \in [a,b] \mid f(x) \neq 0\}$$
has measure 0. Thus, $f = 0$ a.e. on $[a,b]$, as claimed. $\square$

We will now prove Theorem 2.2.13.

Proof. Assume f is Lebesgue integrable on $[a, b]$. For each $k \in \mathbb{N}$ there exists a partition $P_k = \{E_j^k\}_{j=1}^{n_k}$ of $[a, b]$ such that

$$U[f, P_k] - L[f, P_k] < \frac{1}{k}.$$

As observed in Remark 2.2.3, we can always refine a partition to one that consists of pairwise disjoint sets. Moreover, we can replace P_2 by a common refinement of P_1 and P_2, replace P_3 by a common refinement of this new P_2 and P_3, etc. In other words, without loss of generality, we may assume each partition consists of pairwise disjoint sets and P_{k+1} is a refinement of P_k for each k.

Define a sequence of functions $\{g_k\}$ by

$$g_k(x) = \sum_{j=1}^{n_k} m_j^k \mathcal{X}_{E_j^k}(x), \qquad \text{where} \qquad m_j^k = \inf_{x \in E_j^k} f(x).$$

By Exercise 15, g_k is a measurable function for each k and

$$\int_a^b g_k = \sum_{j=1}^{n_k} m_j^k m(E_j^k) = L[f, P_k].$$

Also, for each $x \in [a, b]$, the sequence $\{g_k(x)\}$ is an increasing sequence that is bounded above by $f(x)$, and hence converges. Define $g(x)$ as $g(x) = \lim\limits_{k \to \infty} g_k(x)$. Then g is a bounded measurable function by Corollary 2.1.13. Moreover, for every k, $g_k(x) \le g(x) \le f(x)$ for all $x \in [a, b]$. Thus,

$$\int_a^b g_k \le \int_a^b g \le \int_a^b f$$

so that

$$L[f, P_k] \le \int_a^b g \le \int_a^b f.$$

In a similar fashion define the sequence of functions $\{h_k\}$ as

$$h_k(x) = \sum_{j=1}^{n_k} M_j^k \mathcal{X}_{E_j^k}(x), \qquad \text{where} \qquad M_j^k = \inf_{x \in E_j^k} f(x).$$

This pointwise decreasing sequence of functions will converge to a measurable function $h \geq f$ with h. Thus,

$$\int_a^b f \leq \int_a^b h \leq \int_a^b h_k$$

so that

$$\int_a^b f \leq \int_a^b h \leq U[f, P_k]\,.$$

This means that $g(x) \leq f(x) \leq h(x)$ for all $x \in [a,b]$ and

$$0 \leq \int_a^b (h-g) = \int_a^b h - \int_a^b g \leq U[f,P_k] - L[f,P_k] < \frac{1}{k}$$

for every positive integer k. Hence, $\displaystyle\int_a^b (h-g) = 0$. Thus, by Lemma 2.2.14, $g = h$ a.e. However, $g(x) \leq f(x) \leq h(x)$ for all $x \in [a,b]$. Therefore, $f = g$ a.e. in $[a,b]$. By Proposition 2.1.9, f is a measurable function on $[a,b]$. $\square$

2.3. Properties of the Lebesgue Integral

Up to this point, we have only considered bounded functions. We will now define the Lebesgue integral for unbounded functions.

Definition 2.3.1. Let f be an unbounded function defined on $[a,b]$.

(i) Suppose $f(x) \geq 0$ for all $x \in [a,b]$. For $N > 0$ define

$$^N f(x) = \begin{cases} f(x) & \text{if } f(x) \leq N, \\ N & \text{otherwise.} \end{cases}$$

We say f **is Lebesgue integrable on** $[a,b]$ if f^N is Lebesgue integrable for all $N > 0$ and $\displaystyle\lim_{N\to+\infty}\left(\int_a^b {}^N f\right)$ is finite. In this case $\displaystyle\int_a^b f$ is defined to be

$$\int_a^b f = \lim_{N\to+\infty}\left(\int_a^b {}^N f\right).$$

(ii) Suppose $f(x) < 0$ for some $x \in [a,b]$. Set

$$f^+(x)=\begin{cases} f(x) & \text{if } f(x) \geq 0, \\ 0 & \text{otherwise} \end{cases} \quad \text{and} \quad f^-(x)=\begin{cases} -f(x) & \text{if } f(x) < 0, \\ 0 & \text{otherwise.} \end{cases}$$

We say f **is Lebesgue integrable on** $[a,b]$ if both f^+ and f^- are Lebesgue integrable on $[a,b]$. In this case we define $\int_a^b f$ as

$$\int_a^b f = \int_a^b f^+ - \int_a^b f^- .$$

The function f^+ is called the positive part of f, and f^- is known as the negative part of f. We define **the set of Lebesgue integrable functions on** $[a,b]$, denoted $\mathcal{L}[a,b]$, by

$$\mathcal{L}[a,b] = \{f \mid f \text{ is Lebesgue integrable on } [a,b]\} .$$

In other words, for a bounded function, we define the Lebesgue integral by our earlier definition. For a positive unbounded function f, we define a capped version of this function, f^N. For each N, f^N is now a bounded function and we return to our earlier definition of Lebesgue integration. For f to be Lebesgue integrable, we need to be able to "lift the cap" and obtain a finite limit. More generally, for an unbounded function f we consider the positive part f^+ and negative part f^- separately. Both of these must be Lebesgue integrable in order for f to be considered Lebesgue integrable. Finally, since $f = f^+ - f^-$, it is natural to define the Lebesgue integral of f as

$$\int_a^b f = \int_a^b f^+ - \int_a^b f^- .$$

Example 2.3.2. Let

$$f(x) = \begin{cases} \frac{1}{x} & \text{if } x \neq 0, \\ 0 & \text{if } x = 0. \end{cases}$$

Our goal is to determine if f is in $\mathcal{L}[0,1]$. For $N > 1$, on this interval

$$^N f(x) = \begin{cases} N & \text{if } 0 < x \leq \frac{1}{N}, \\ \frac{1}{x} & \text{if } \frac{1}{N} < x \leq 1, \\ 0 & \text{if } x = 0, \end{cases}$$

so

$$\int_0^1 {}^N f = \int_0^{\frac{1}{N}} N + \int_{\frac{1}{N}}^1 \frac{1}{x} = 1 + \ln(N)\,.$$

Notice that we were able to evaluate the last two integrals by observing that the integrands are Riemann integrable functions. Hence,

$$\lim_{N\to\infty} \int_0^1 {}^N f = +\infty$$

and $f \notin \mathcal{L}[0,1]$.

Example 2.3.3. Let

$$g(x) = \begin{cases} \frac{1}{\sqrt{x}} & \text{if } x \neq 0, \\ 0 & \text{if } x = 0. \end{cases}$$

We will determine if g is in $\mathcal{L}[0,1]$. For $N > 1$, on this interval

$${}^N g(x) = \begin{cases} N & \text{if } 0 < x \leq \frac{1}{N^2}, \\ \frac{1}{\sqrt{x}} & \text{if } \frac{1}{N^2} < x \leq 1, \\ 0 & \text{if } x = 0, \end{cases}$$

so

$$\int_0^1 {}^N g = \int_0^{\frac{1}{N^2}} N + \int_{\frac{1}{N^2}}^1 \frac{1}{\sqrt{x}} = \frac{1}{N} + \left(2 - \frac{2}{N}\right)\,.$$

Again, we now have Riemann integrable functions. Hence,

$$\lim_{N\to\infty} \int_0^1 {}^N g = 2\,.$$

Therefore $g \in \mathcal{L}[0,1]$ and $\displaystyle\int_0^1 g = 2$.

The previous example also illustrates a difference between Lebesgue integration and Riemann integration. In the latter, $\displaystyle\int_0^1 \frac{1}{\sqrt{x}}\,dx$ is considered as an improper integral and would be evaluated as follows:

$$\int_0^1 \frac{1}{\sqrt{x}}\,dx = \lim_{a\to 0^+} \int_a^1 \frac{1}{\sqrt{x}}\,dx = \lim_{a\to 0^+} \left(2 - 2\sqrt{a}\right) = 2\,.$$

Although the answer is the same, the process is different. For Lebesgue integration the strategy was to truncate the range. For Riemann integration the improper integral is evaluated by first restricting the domain.

The following properties are true for $\mathcal{L}[a, b]$.

Theorem 2.3.4. *Let $f \in \mathcal{L}[a, b]$. If $a < c < b$, then $f \in \mathcal{L}[a, c]$ and $f \in \mathcal{L}[c, b]$, and*

$$\int_a^b f = \int_a^c f + \int_c^b f\,.$$

Theorem 2.3.5. *Let $f, g \in \mathcal{L}[a, b]$ and $c \in \mathbb{R}$.*

(i) *Then $(f + g) \in \mathcal{L}[a, b]$ and*

$$\int_a^b (f + g) = \int_a^b f + \int_a^b g\,.$$

(ii) *$cf \in \mathcal{L}[a, b]$ and*

$$\int_a^b (cf) = c \int_a^b f\,.$$

In other words, the two theorems stated above assert that $\mathcal{L}[a, b]$ is a vector space.

Theorem 2.3.6. *Let $f, g \in \mathcal{L}[a, b]$ with $f(x) \le g(x)$ for all $x \in [a, b]$. Then*

$$\int_a^b f \le \int_a^b g\,.$$

Although the above results may seem obviously true, the proofs are not exactly as straightforward as one might expect. This is due to the fact that we now have to deal with the possibility that the function or functions involved are unbounded. This adds at least one extra case in each proof. Although we are omitting the proofs of these theorems, we will illustrate this added twist by proving the next theorem.

Theorem 2.3.7. *Let $f \in B[a, b]$. Suppose $f \in \mathcal{L}[a, b]$ and $f = g$ a.e. on $[a, b]$. Then $g \in \mathcal{L}[a, b]$ and*

$$\int_a^b g = \int_a^b f\,.$$

Proof. We will show that $g-f$ is in $\mathcal{L}[a,b]$ and $\int_a^b (g-f) = 0$. The result follows by the linearity of the integral.

(i) Assume $(g-f) \in B[a,b]$. Then $(g-f) = 0$ a.e. on $[a,b]$. By Exercise 14,

$$\int_a^b (g-f) = 0\,.$$

(ii) Assume $(g-f)$ is not bounded on $[a,b]$, but $(g-f)(x) \geq 0$ for all $x \in [a,b]$. For each $N > 0$ we set

$$^N(g-f)(x) = \begin{cases} (g-f)(x) & \text{if } (g-f)(x) \leq N, \\ N & \text{otherwise.} \end{cases}$$

For each N, $^N(g-f) = 0$ a.e. in $[a,b]$. By part (i),

$$\int_a^b {}^N(g-f) = 0$$

for each N. Therefore,

$$\int_a^b (g-f) = \lim_{N\to\infty} \int_a^b {}^N(g-f) = 0\,.$$

As a result, $(g-f) \in \mathcal{L}[a,b]$ and $\int_a^b (g-f) = 0$.

(iii) Assume $(g-f)$ is unbounded on $[a,b]$. The result follows by applying the previous parts to $(g-f)^+$ and $(g-f)^-$.

□

Here are more examples of results that seem fairly obvious but where the proofs require consideration of the cases. The first is actually a special case of Theorem 2.3.6.

Theorem 2.3.8. *Let $f \in \mathcal{L}[a,b]$. Suppose $f(x) \geq 0$ a.e. on $[a,b]$. Then*

$$\int_a^b f \geq 0\,.$$

Proof. (i) Assume $f \in B[a,b]$. Consider f^+, the positive part of f. Then $f = f^+$ a.e. on $[a,b]$. By Theorem 2.3.7, $\int_a^b f^+ = \int_a^b f$, but by Theorem 2.3.6,

$$\int_a^b f^+ \geq \int_a^b 0 = 0\,.$$

Therefore $\int_a^b f \geq 0$.

(ii) Assume f is unbounded on $[a,b]$. We will consider the positive and negative parts of f. Since $f^- = 0$ a.e. in $[a,b]$, Theorem 2.3.7 implies $\int_a^b f^- = 0$. Thus, without loss of generality, we may assume $f(x) \geq 0$ for all $x \in [a,b]$.

For each $N > 0$ set

$${}^N f = \begin{cases} f(x) & \text{if } f(x) \leq N, \\ N & \text{otherwise.} \end{cases}$$

Then ${}^N f \geq 0$ a.e. in $[a,b]$. By part (i),

$$\int_a^b {}^N f \geq 0$$

for every $N > 0$. Therefore,

$$\int_a^b f = \lim_{N\to\infty} \int_a^b {}^N f \geq 0\,. \qquad \square$$

Theorem 2.3.9. *Let $f \in \mathcal{L}[a,b]$. If $f(x) \geq 0$ a.e. on $[a,b]$ and $\int_a^b f = 0$, then $f = 0$ a.e. on $[a,b]$.*

Proof. Without loss of generality, we may assume that $f(x) \geq 0$ for all $x \in [a,b]$.

(i) Assume f is bounded. This is covered by Lemma 2.2.14.

(ii) Assume f is unbounded. Then

$$0 = \int_a^b f = \lim_{N\to\infty} \int_a^b {}^N f\,.$$

By definition, $\int_a^b {}^N f$ increases with N. Also, ${}^N f$ is always nonnegative, and hence $\int_a^b {}^N f$ is nonnegative. Therefore, for each $N > 0$, $\int_a^b {}^N f = 0$. By part (i), ${}^N f(x) = 0$ a.e. on $[a, b]$. Since $f(x) \geq 0$ for all $x \in [a, b]$, ${}^N f(x) = 0$ if and only if $f(x) = 0$; therefore, $f(x) = 0$ a.e. as claimed. $\square$

So far, we have limited ourselves to integrating over closed bounded intervals. We can generalize this to integrating over bounded measurable sets.

Definition 2.3.10. Let E be a measurable subset of $[a, b]$. We define $\int_E f$ as

$$\int_E f = \int_a^b (f\, \mathcal{X}_E)\,.$$

As one might expect, if we integrate an integrable function over a small set, the result is small. This is made more precise in the next lemma.

Lemma 2.3.11. *Let $f \in \mathcal{L}[a, b]$. Given any $\epsilon > 0$, there exists a $\delta > 0$ so that*

$$\textit{if } m(E) < \delta, \ \textit{ then } \left| \int_E f \right| < \epsilon\,.$$

Proof. As usual with $\mathcal{L}[a, b]$, we will break the proof down into cases.

(i) Assume $f \in B[a, b]$. Then there is an $M > 0$ so that

$$-M \leq f(x) \leq M$$

for all $x \in [a, b]$. Hence, for any measurable set E,

$$-M\, \mathcal{X}_E(x) \leq f(x)\, \mathcal{X}_E(x) \leq M\, \mathcal{X}_E(x)$$

for all $x \in [a, b]$. Thus,

$$\int_a^b (-M\, \mathcal{X}_E) \leq \int_a^b f \leq \int_a^b (M\, \mathcal{X}_E)$$

so that

$$-M\, m(E) \leq \int_a^b f \leq M\, m(E)\,.$$

Therefore,

$$\left| \int_E f \right| \leq M\, m(E)\,.$$

Given $\epsilon > 0$, choosing $\delta < \frac{\epsilon}{M}$ suffices.

(ii) Assume $f(x) \geq 0$ for all $x \in [a, b]$, but f is unbounded. Let $\epsilon > 0$ be given. Since

$$\lim_{N\to\infty} \int_a^b {}^N f = \int_a^b f\,,$$

there is a K so that if $N > K$, then

$$\int_a^b (f - {}^N f) = \left| \int_a^b (f - {}^N f) \right| = \left| \int_a^b f - \int_a^b {}^N f \right| < \frac{\epsilon}{2}\,.$$

Pick $M > K$. By part (i) there is a $\delta > 0$ such that if $m(E) < \delta$, then $\left| \int_E {}^M f \right| < \frac{\epsilon}{2}$. Therefore, if $m(E) < \delta$, then

$$\begin{aligned}
\left| \int_E f \right| &= \left| \int_a^b (f\, \mathcal{X}_E) \right| \\
&= \left| \int_a^b \Big((f - {}^M f)\, \mathcal{X}_E \Big) + \int_a^b ({}^M f \mathcal{X}_E) \right| \\
&\leq \int_a^b \Big((f - {}^M f)\, \mathcal{X}_E \Big) + \left| \int_a^b ({}^M f \mathcal{X}_E) \right| \\
&\leq \int_a^b (f - {}^M f) + \left| \int_E {}^M f \right| \\
&< \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon\,.
\end{aligned}$$

(iii) In the most general case we write $f = f^+ - f^-$ and observe that

$$\begin{aligned}
\left|\int_E f\right| &= \left|\int_a^b (f\,\mathcal{X}_E)\right| = \left|\int_a^b \left((f^+ - f^-)\,\mathcal{X}_E\right)\right| \\
&= \left|\int_a^b (f^+\,\mathcal{X}_E) - \int_a^b (f^-\,\mathcal{X}_E)\right| \\
&\leq \left|\int_a^b (f^+\,\mathcal{X}_E)\right| + \left|\int_a^b (f^-\,\mathcal{X}_E)\right| \\
&= \left|\int_E f^+\right| + \left|\int_E f^-\right|.
\end{aligned}$$

Both f^+ and f^- are covered by part (ii). □

2.4. The Lebesgue Dominated Convergence Theorem

Some of the basic questions in real analysis concern interchanging "limit-type operations". For example, if one takes a sequence of Riemann integrable functions that converges, will they converge to a Riemann integrable function? In Riemann integration, uniform convergence was useful. What about in $\mathcal{L}[a, b]$? Do we need uniform convergence on something else?

There are three main results in this section, the Lebesgue Dominated Convergence Theorem, Fatou's Lemma, and the Monotone Covergence Theorem. They all concern sequences of functions in $\mathcal{L}[a, b]$. They actually are equivalent (not shown in this text), but we will prove them in the order stated above. To motivate them, let's consider some examples.

Example 2.4.1. For $x \in [0, 1]$ and positive integer n, let $f_n(x) = x^n$. Then $\lim_{n\to\infty} f_n(x) = f(x),$, where

$$f(x) = \begin{cases} 0 & \text{if } 0 \leq x < 1, \\ 1 & \text{if } x = 1. \end{cases}$$

This example shows that the pointwise limit of a sequence of continuous functions need not be continuous.

Example 2.4.2. Let $\mathbb{Q} \cap [0,1] = \{r_1, r_2, \ldots\}$ be a counting of the rationals in the interval $[0,1]$. For $x \in [0,1]$ and $n \in \mathbb{N}$ define f_n by

$$f_n(x) = \begin{cases} 1 & \text{if } x = r_1, r_2, \ldots, r_n, \\ 0 & \text{otherwise.} \end{cases}$$

For each n, f_n is Riemann integrable since f_n has a finite number of discontinuities. But $\lim\limits_{n \to \infty} f_n(x) = f(x)$, where

$$f(x) = \begin{cases} 1 & \text{if } x \in \mathbb{Q} \cap [0,1], \\ 0 & \text{otherwise.} \end{cases}$$

Here is an example where the pointwise limit of a sequence of Riemann integrable functions need not be Riemann integrable. (This is actually Problem 11 of Chapter 0.)

This leads to the question as to whether or not the pointwise limit of a sequence of Lebesgue integrable functions will be Lebesgue integrable. We give one final example before we begin our buildup to the statement and proof of the Lebesgue Dominated Convergence Theorem.

Example 2.4.3. Define f_n by

$$f_n(x) = n\,\mathcal{X}_{(0,\frac{1}{n}]}(x) = \begin{cases} n & \text{if } 0 < x \leq \frac{1}{n}, \\ 0 & \text{otherwise.} \end{cases}$$

In this case $\lim\limits_{n \to \infty} f_n(x) = 0$ for all $x \in [0,1]$. So here is an example where the pointwise limit of a sequence of Lebesgue integrable functions is Lebesgue integrable. The odd thing is that $\displaystyle\int_0^1 f_n = 1$ for every n, but $\displaystyle\int_0^1 0 = 0$. In other words, here is an example where

$$\lim_{n \to \infty} \int_a^b f_n \neq \int_a^b \left(\lim_{n \to \infty} f_n \right).$$

What we are seeking are conditions that allow us to interchange two limit-type operations, namely, integration (Lebesgue integration in this case) and the limit of a sequence of functions. Our goal is a theorem that addresses this issue, the Lebesgue Dominated Convergence Theorem. Before the statement and proof of this theorem, we need two lemmas.

Lemma 2.4.4. *Let $f \in \mathcal{L}[a,b]$. Suppose $\{A_k\}_{k=1}^{\infty}$ is a countable collection of measurable subsets of $[a,b]$ with*

$$A_1 \subseteq A_2 \subseteq A_3 \subseteq \dots$$

and

$$\bigcup_{k=1}^{\infty} A_k = [a,b] \, .$$

Then

$$\lim_{k\to\infty} \int_{A_k} f = \int_a^b f \, .$$

Note: Intuitively, we want to say that this lemma is asserting that

$$\lim_{k\to\infty} \int_{A_k} f = \int_{\lim_{k\to\infty} A_k} f,$$

but there is a potential problem with this statement. The left-hand side of this equation is a sequence of numbers. However, the careful reader will see that the expression on the right contains what appears to be the limit of a sequence of sets. Although it may seem tempting to try to do so, it is extremely difficult to define $\lim\limits_{k\to\infty} A_k$, the limit of a sequence of sets. So, in our proof we will avoid taking "the limit of a sequence of sets".

Proof. Let $\epsilon > 0$ be given. Our goal is to find N so that if $k > N$, then

$$\left| \int_a^b f - \int_{A_k} f \right| < \epsilon \, .$$

Let $E_k = [a,b] \setminus A_k$. Then

$$\begin{aligned} \int_a^b f - \int_{A_k} f &= \int_a^b f - \int_a^b f \, \mathcal{X}_{A_k} \\ &= \int_a^b f \, \left(1 - \mathcal{X}_{A_k}\right) \\ &= \int_a^b f \, \mathcal{X}_{E_k} \\ &= \int_{E_k} f \, . \end{aligned}$$

Also,

$$E_1 \supseteq E_2 \supseteq E_3 \supseteq \dots$$

and

$$\bigcap_{k=1}^{\infty} E_k = \emptyset\,.$$

By Exercise 19 of Chapter 1, $\lim_{k\to\infty} m(E_k) = 0$. (Notice how this is a limit of a sequence of *numbers*, not *sets*.)

Now we will put the pieces together. By Lemma 2.3.11, there is a $\delta > 0$ so that $\left|\int_E f\right| < \epsilon$ if $m(E) < \delta$. Since $\lim_{k\to\infty} m(E_k) = 0$, there exists an N such that if $k > N$, then $m(E_k) < \delta$. So if $k > N$, then

$$\left|\int_a^b f - \int_{A_k} f\right| = \left|\int_{E_k} f\right| < \epsilon\,. \qquad \square$$

Not only is this next lemma used in the proof of the Lebesgue Dominated Convergence Theorem, it is frequently used in other situations, often without an actual reference. In essence, it says that if a measurable function is bounded by a Lebesgue integrable function, it must also be integrable.

Lemma 2.4.5. *Let $g \in \mathcal{L}[a,b]$. Suppose f is measurable and $|f(x)| \le g(x)$ almost everywhere in $[a,b]$. Then $f \in \mathcal{L}[a,b]$.*

Proof. Without loss of generality, we may assume that $|f(x)| \le g(x)$ for all $x \in [a,b]$. We must show that f^+ and f^- are Lebesgue integrable. Since $|f(x)| \le g(x)$, both $0 \le f^+(x) \le g(x)$ and $0 \le f^-(x) \le g(x)$ for all $x \in [a,b]$. Therefore, it suffices to show that if f is measurable on $[a,b]$ and $0 \le f(x) \le g(x)$ for all $x \in [a,b]$, then $f \in \mathcal{L}[a,b]$.

Since $0 \le f(x) \le g(x)$,

$$0 \le {}^N f(x) \le {}^N g(x)$$

for each N. Thus,

$$0 \le \int_a^b {}^N f(x) \le \int_a^b {}^N g(x) \le \int_a^b g$$

for every N. But $\displaystyle\int_a^b {}^N f(x)$ increases with N. The above inequalities show that $\displaystyle\int_a^b {}^N f(x)$ is bounded. Therefore $\displaystyle\lim_{N\to\infty}\int_a^b {}^N f(x)$ exists and $f \in \mathcal{L}[a,b]$. □

We now come to the Lebesgue Dominated Convergence Theorem, a frequently cited theorem. Some people refer to this result as the Dominated Convergence Theorem, while others simply abbreviate it by LDC. Before continuing with this theorem, the reader is advised to review a solution to Exercise 6 from Chapter 0.

Theorem 2.4.6 (Lebesgue Dominated Convergence Theorem)**.** *Let $\{f_n\}$ be a sequence of measurable functions on $[a,b]$ such that*

$$\lim_{n\to\infty} f_n(x) = f(x) \ \textit{a.e.}$$

on $[a,b]$. Suppose there exists $g \in \mathcal{L}[a,b]$ with

$$|f_n(x)| \le g(x) \ \textit{a.e.}$$

for every n. Then, $f_n \in \mathcal{L}[a,b]$ for every n, $f \in \mathcal{L}[a,b]$, and

$$\lim_{n\to\infty}\int_a^b f_n = \int_a^b f\,.$$

Note: The remarkable feature of this theorem is that the existence of the function g, called a *dominating function*, guarantees that it is safe to interchange the integral and the limit, hence the name Lebesgue Dominated Convergence Theorem.

Proof. Since $\lim_{n\to\infty} f_n(x) = f(x)$ a.e. and $|f_n(x)| \le g(x)$ a.e. on $[a,b]$,

$$|f(x)| \le g(x) \text{ a.e.}$$

on $[a,b]$. Hence by Lemma 2.4.5, $f_n \in \mathcal{L}[a,b]$ for every n and $f \in \mathcal{L}[a,b]$. It remains to show that

$$\lim_{n\to\infty}\int_a^b f_n = \int_a^b f\,.$$

Without loss of generality we may assume that $\lim_{n\to\infty} f_n(x) = f(x)$ for all $x \in [a,b]$. (Make sure you understand why this is true!) Let

$\epsilon > 0$ be given. We must show there exists an N such that if $n > N$, then

$$\left| \int_a^b f_n - \int_a^b f \right| < \epsilon\,.$$

The strategy is to create subsets of $[a, b]$ where the sequence of function is almost converging uniformly. To do this, for each positive integer n set

$$A_n = \left\{ x \in [a, b] \;\middle|\; |f_k(x) - f(x)| < \frac{\epsilon}{2(b-a)} \text{ for all } k \geq n \right\}.$$

(Can you see the idea of uniform convergence being used here?) By definition,

$$A_1 \subseteq A_2 \subseteq A_3 \subseteq \ldots$$

and

$$\bigcup_{k=1}^{\infty} A_k \subseteq [a, b]\,.$$

But for each $x \in [a, b]$, $\lim_{n\to\infty} f_n(x) = f(x)$ (pointwise). Therefore, for each x, there is an N so that whenever $k \geq N$, then $|f_k(x) - f(x)| < \frac{\epsilon}{2(b-a)}$. In other words, for each $x \in [a, b]$ there is an N so that $x \in A_N$. Hence,

$$\bigcup_{k=1}^{\infty} A_k = [a, b]\,.$$

Also,

$$\int_{A_n} |f_n - f| \leq \int_{A_n} \frac{\epsilon}{2(b-a)} = m(A_n)\frac{\epsilon}{2(b-a)} \leq \frac{\epsilon}{2}\,.$$

As in Lemma 2.4.4, set $E_n = [a, b] \setminus A_n$ so that

$$E_1 \supseteq E_2 \supseteq E_3 \supseteq \ldots$$

and

$$\bigcap_{n=1}^{\infty} E_n = \emptyset\,.$$

Thus, $\lim_{n\to\infty} m(E_n) = 0$.

So far, for each n we can control the difference between $\int f_n$ and $\int f$ when we are integrating over the specially designed set A_n. To control the difference between $\int f_n$ and $\int f$ when we are integrating over the rest of the interval $[a, b]$, we will use the dominating function. By Lemma 2.3.11, there exists a $\delta > 0$ so that $\int_E g < \frac{\epsilon}{4}$ whenever $m(E) < \delta$. Since $|f_n(x)| \le g(x)$ and $|f(x)| \le g(x)$, if $m(E) < \delta$, it will also be the case that

$$\int_E |f_n| < \frac{\epsilon}{4} \quad \text{for all } n, \text{ and} \qquad \int_E |f| < \frac{\epsilon}{4}\,.$$

Since $\lim_{n\to\infty} m(E_n) = 0$, there exists N so that if $n > N$, then $m(E_n) < \delta$. Thus, if $n > N$,

$$\begin{aligned}
\left| \int_a^b f_n - \int_a^b f \right| &= \left| \int_a^b (f_n - f) \right| \\
&= \left| \int_{A_n} (f_n - f) \right| + \left| \int_{E_n} (f_n - f) \right| \\
&\le \int_{A_n} |f_n - f| + \int_{E_n} |f_n - f| \\
&\le \int_{A_n} \frac{\epsilon}{2(b-a)} + \int_{E_n} |f_n| + \int_{E_n} |f| \\
&< \frac{\epsilon}{2} + \frac{\epsilon}{4} + \frac{\epsilon}{4} = \epsilon\,.
\end{aligned}$$

□

Example 2.4.7. For each positive integer n and $x \in [0, 2]$ define $f_n(x)$ to be

$$f_n(x) = \begin{cases} 0 & \text{if } 0 \le x < \frac{1}{n}, \\ \sqrt{n} & \text{if } \frac{1}{n} \le x \le \frac{2}{n}, \\ 0 & \text{if } \frac{2}{n} < x \le 2. \end{cases}$$

It is easy to verify $\lim_{n\to\infty} f_n(x) = 0$ for all $x \in [0, 2]$. Let

$$g(x) = \begin{cases} \frac{\sqrt{2}}{\sqrt{x}} & \text{if } x \ne 0, \\ 0 & \text{if } x = 0. \end{cases}$$

Then $g \in \mathcal{L}[0, 2]$ and $|f_n(x)| \le g(x)$ for all $x \in [0, 2]$. Therefore, the Lebesgue Dominated Convergence Theorem guarantees that

$$\lim_{n\to\infty} \int_0^2 f_n = \int_0^2 0 = 0\,.$$

This can also be verfied by directly computing $\lim_{n\to\infty} \int_0^2 f_n$ for each n.

In the above example, we could directly compute $\int_a^b f_n$ for each n. This will not be the case in the next two examples.

Example 2.4.8. For each positive integer n and $x \in [0,1]$ define $f_n(x)$ to be

$$f_n(x) = \begin{cases} 0 & \text{if } x = 0, \\ \left(1 - e^{-\frac{x^2}{n}}\right) \frac{1}{\sqrt{x}} & \text{if } 0 < x \leq 1. \end{cases}$$

This is a case where it is not so easy to compute $\int_a^b f_n$ for each n, yet it is easy to compute the pointwise limit of this sequence of functions, which is the function that is identically 0 on $[0,1]$. If we define g by

$$g(x) = \begin{cases} 0 & \text{if } x = 0, \\ \frac{1}{\sqrt{x}} & \text{if } 0 < x \leq 1. \end{cases}$$

then $g \in \mathcal{L}[0,1]$ and $|f_n(x)| \leq g(x)$ for all $x \in [0,1]$. Therefore

$$\lim_{n\to\infty} \int_0^1 f_n = \int_0^1 0 = 0\,.$$

Without the Lebesgue Dominated Convergence Theorem, it would be difficult to verify this limit.

In case the reader begins to believe that $\int_a^b f_n$ will always equal 0, we have the next example.

Example 2.4.9. For each positive integer n and $x \in [0,1]$ define $f_n(x)$ to be

$$f_n(x) = \frac{n \sin x}{1 + n^2\sqrt{x}} + 2e^{x/n}\,.$$

The pointwise limit is $\lim_{n\to\infty} f_n(x) = 2$. In this case, for $x \neq 0$,

$$|f_n(x)| \leq \frac{n}{1 + n^2\sqrt{x}} + 2 \leq \frac{1}{n\sqrt{x}} + 2 \leq \frac{1}{\sqrt{x}} + 2\,.$$

Therefore, we may use the dominating function g where

$$g(x) = \begin{cases} 2 & \text{if } x = 0, \\ \frac{1}{\sqrt{x}} + 2 & \text{if } 0 < x \leq 1. \end{cases}$$

Since $g \in \mathcal{L}[0,1]$,

$$\lim_{n\to\infty} \int_0^1 f_n = \int_0^1 2 = 2\,.$$

There are two other important results concerning the interchange of integration with limits that are corollaries of the Lebesgue Dominated Convergence Theorem. These are known as Fatou's Lemma and the Monotone Convergence Theorem. We will start with a preliminary version of Fatou's Lemma.

Theorem 2.4.10 (Fatou's Lemma, preliminary version)**.** *Let $\{f_n\}$ be a sequence of nonnegative functions in $\mathcal{L}[a,b]$. Suppose $\lim_{n\to\infty} f_n(x) = f(x)$ a.e. in $[a,b]$.*

(i) *If $f \in \mathcal{L}[a,b]$, then $\displaystyle\int_a^b f \le \liminf_{n\to\infty} \int_a^b f_n$.*

(ii) *If $f \notin \mathcal{L}[a,b]$, then $\displaystyle\liminf_{n\to\infty} \int_a^b f_n = +\infty$.*

Proof. Without loss of generality, we may assume $f(x) \ge 0$ for all $x \in [a,b]$. It follows that for every N,

$$\lim_{n\to\infty} {}^N f_n(x) = {}^N f(x) \le N\,.$$

(Remember, we might be dealing with unbounded functions here.) Using $g(x) = N$ as the dominating function in the Lebesgue Dominated Convergence Theorem,

$$\lim_{n\to\infty} \int_a^b {}^N f_n = \int_a^b {}^N f$$

for each N.

So far, we have been dealing the the "capped" versions of the original functions. Now we will work on "raising the caps". For each n and N,

$$\int_a^b {}^N f_n \le \int_a^b f_n\,,$$

and hence

$$\inf_{k\ge n} \int_a^b {}^N f_k \le \inf_{k\ge n} \int_a^b f_k\,.$$

Thus,

$$\liminf_{n\to\infty} \int_a^b {}^N f_n \le \liminf_{n\to\infty} \int_a^b f_n\,.$$

Note that this includes the possibility that $\liminf_{n\to\infty} \int_a^b f_n = +\infty$.

However, $\lim_{n\to\infty} \int_a^b {}^N f_n = \liminf_{n\to\infty} \int_a^b {}^N f_n = \int_a^b {}^N f$. Therefore we have

$$\int_a^b {}^N f \le \liminf_{n\to\infty} \int_a^b f_n\,.$$

If $f \in \mathcal{L}[a,b]$,

$$\int_a^b f = \lim_{N\to\infty} \int_a^b {}^N f \le \liminf_{n\to\infty} \int_a^b f_n,$$

and we have established (i). On the other hand, if $f \notin \mathcal{L}[a,b]$, then it must be the case that $\lim_{N\to\infty} \int_a^b {}^N f = +\infty$. This establishes part (ii). $\square$

Example 2.4.11. For $x \in [0,1]$ and positive integer n, let

$$f_n(x) = \begin{cases} 2n^2 x & \text{if } 0 \le x \le \frac{1}{2n}, \\ -2n^2\left(x - \frac{1}{n}\right) & \text{if } \frac{1}{2n} < x \le \frac{1}{n}, \\ 0 & \text{otherwise.} \end{cases}$$

For each n, f_n is piecewise linear, connecting the origin with $(\frac{1}{2n}, n)$, which connects to $(\frac{1}{n}, 0)$, forming a triangle. In this case, the pointwise limit of the sequence of functions is the identically 0 function. On the other hand, for each n,

$$\int_0^1 f_n = \frac{1}{2}\,.$$

This example demonstrates that the strict inequality can hold in this preliminary version of Fatou's Lemma.

What is usually stated as Fatou's Lemma actually is a corollary of this preliminary version.

Corollary 2.4.12 (Fatou's Lemma)**.** *Let $\{f_n\}$ be a sequence of non-negative functions in $\mathcal{L}[a,b]$. Suppose $\liminf_{n\to\infty} f_n(x) = f(x)$ a.e. in $[a,b]$.*

(i) *If* $f \in \mathcal{L}[a,b]$, *then* $\displaystyle\int_a^b f \leq \liminf_{n\to\infty} \int_a^b f_n$.

(ii) *If* $f \notin \mathcal{L}[a,b]$, *then* $\displaystyle\liminf_{n\to\infty} \int_a^b f_n = +\infty$.

Proof. For each positive integer n, let $g_n(x) = \inf_{k\geq n} f_k(x)$. Then g_n is nonnegative and measurable for each n. Also, by Lemma 2.4.5, for each n, $g_n \in \mathcal{L}[a,b]$ since $g_n(x) \leq f_n(x)$. Thus,

$$\int_a^b g_n \leq \int_a^b f_n$$

for each n. Consequently,

$$\liminf_{n\to\infty} \int_a^b g_n \leq \liminf_{n\to\infty} \int_a^b f_n \, .$$

On the other hand, $\lim_{n\to\infty} g_n(x) = \liminf_{n\to\infty} f_n(x) = f(x)$ a.e. Therefore, by our preliminary version of Fatou's Lemma, Theorem 2.4.10, if $f \in \mathcal{L}[a,b]$, then

$$\int_a^b f \leq \liminf_{n\to\infty} \int_a^b g_n \leq \liminf_{n\to\infty} \int_a^b f_n \, .$$

Thus, we have established part (i). If $f \notin \mathcal{L}[a,b]$, then

$$\liminf_{n\to\infty} \int_a^b g_n = +\infty \, .$$

In this case,

$$\liminf_{n\to\infty} \int_a^b f_n = +\infty,$$

and we have part (ii). □

The next main result stemming from the Lebesgue Dominated Convergence Theorem is known as the Monotone Convergence Theorem.

Theorem 2.4.13 (Monotone Convergence Theorem)**.** *Let* $\{f_n\}$ *be a sequence of nonnegative functions in* $\mathcal{L}[a,b]$. *Suppose* $\{f_n(x)\}$ *is an increasing sequence for almost every* $x \in [a,b]$ *and* $\lim_{n\to\infty} f_n(x) = f(x)$ *a.e. in* $[a,b]$.

(i) *If* $f \in \mathcal{L}[a, b]$, *then* $\lim_{n\to\infty} \int_a^b f_n = \int_a^b f$.

(ii) *If* $f \notin \mathcal{L}[a, b]$, *then* $\lim_{n\to\infty} \int_a^b f_n = +\infty$.

Proof. For each n, $f_n(x) \le f_{n+1}(x)$ almost everywhere in $[a, b]$. Consequently $f_n \le f$ a.e. in $[a, b]$. Therefore, if $f \in \mathcal{L}[a, b]$, we may use $|f|$ as the dominating function in the Lebesgue Dominated Convergence Theorem to establish part (i).

To prove part (ii), note that

$$\int_a^b f_n \le \int_a^b f_{n+1}$$

for all n. Thus, $\left\{\int_a^b f_n\right\}$ is an increasing sequence of numbers. Moreover,

$$\inf_{k\ge n} \int_a^b f_k = \int_a^b f_n\,.$$

Hence,

$$\liminf_{n\to\infty} \int_a^b f_n = \lim_{n\to\infty} \int_a^b f_n\,.$$

Therefore, part (ii) follows from part (ii) of Fatou's Lemma. □

Although we have chosen to first prove the Lebesgue Dominated Convergence Theorem and then Fatou's Lemma and the Monotone Convergence Theorem, in actuality these three major results are equivalent. There are texts that choose to start with Fatou's Lemma before establishing the other results. Still other texts start with the Monotone Convergence Theorem. It is an interesting exercise to verify for yourself how one of these results can be used to show the other two.

2.5. Further Notes on Integration

We have defined the Lebesgue integral on intervals and measurable subsets of intervals in $\mathbb{R}^1$. Suppose we want to integrate a function of more than one variable, that is, $f : \mathbb{R}^n \to \mathbb{R}^1$. The process would be much the same. First, our definition of measurable function remains much the same.

Definition 2.5.1. Let f be defined on I, a closed interval in $\mathbb{R}^n$, or $I = \mathbb{R}^n$. We say f **is measurable on** I if for every $s \in \mathbb{R}^1$ the set $\{x \in I \mid f(x) > s\}$ is a measurable subset of $\mathbb{R}^n$.

The first step is to define the integral of bounded, measurable functions. We simply extend the definition of a measurable partition in the obvious way. The definitions of upper sum and lower sum then follow naturally. Finally, we reach the definition of the upper integral and the lower integral, and then the Lebesgue integral for bounded functions. This actually parallels the manner in which we define the Riemann integral for functions of several variables. In order to integrate unbounded functions, we follow the same process as before. We consider a nonnegative function and put a "cap" on it. Then we see if there is a finite limit when raising the "cap". Finally, for a general measurable function on I, we consider the positive and negative parts.

However, this is not the procedure used in calculus to evaluate such integrals. Most often, we turn to what is known as an iterated integral. The following example illustrates the difference.

Example 2.5.2. Let $I = \{(x, y) \in \mathbb{R}^2 \mid 0 \le x \le 1 \text{ and } 0 \le y \le 1\}$ be the unit square in $\mathbb{R}^2$. Subdivide I into four congruent squares. Let I_1 denote the lower left corner square. Take the upper right corner square and subdivide it into four congruent squares. Let I_2 denote the lower left corner square in this subdivision. Next take the upper right corner square of this subdivision and subdivide it. Label the resulting lower left corner square as I_3. Continue this process. By construction, $m(I_1) = \frac{1}{4}$, $m(I_2) = \frac{1}{16}$, $m(I_3) = \frac{1}{64}$, and, more generally, $m(I_n) = \frac{1}{4^n}$.

Next, for each n, subdivide I_n into four congruent squares. Label these I_n^1, I_n^2, I_n^3, and I_n^4, starting with the lower left square and moving in a counterclockwise fashion. Next, we will define $f : \mathbb{R}^2 \to \mathbb{R}^1$ by

$$f(x, y) = \begin{cases} 4^n & \text{if } (x, y) \in \operatorname{int}(I_n^1) \text{ or } (x, y) \in \operatorname{int}(I_n^3), \\ -4^n & \text{if } (x, y) \in \operatorname{int}(I_n^2) \text{ or } (x, y) \in \operatorname{int}(I_n^4), \\ 0 & \text{otherwise.} \end{cases}$$

A straightforward check confirms that f is measurable. We now turn to the question of Lebesgue integrability.

Since f is unbounded, we must consider f^+ and f^- separately. In this case,

$$f^+(x,y) = \begin{cases} 4^n & \text{if } (x,y) \in \operatorname{int}(I_n^1) \text{ or } (x,y) \in \operatorname{int}(I_n^3), \\ 0 & \text{otherwise.} \end{cases}$$

Thus,

$$\begin{aligned} \int\!\!\int_I f^+ &= 2\cdot 4\left(\frac{1}{2}m(I_1)\right) + 2\cdot 16\left(\frac{1}{2}m(I_2)\right) \\ &\quad + 2\cdot 64\left(\frac{1}{2}m(I_3)\right) + \dots \\ &= \sum_{n=1}^{\infty} 2\cdot 4^n\left(\frac{1}{2}m(I_n)\right) \\ &= \sum_{n=1}^{\infty} 1 = +\infty\,. \end{aligned}$$

(This can be made more precise by actually considering the "capped" version of f.) Therefore, $f \notin \mathcal{L}(I)$.

On the other hand, for every fixed $x \in [0,1]$, $f(x,y)$ is Lebesgue integrable as a function of y. Moreover,

$$\int_0^1 f(x,y)\,dy = 0\,.$$

Hence, the iterated integral $\int_0^1 \int_0^1 f(x,y)\,dydx$ exists and

$$\int_0^1 \int_0^1 f(x,y)\,dydx = \int_0^1 0\,dx = 0\,.$$

This example shows that the double integral over a rectangle need not equal the corresponding iterated integral. A theorem that addresses this is known as **Fubini's Theorem**. We will not be covering this theorem in this text, but it is a possible project; see Project number 7 in the section Ideas for Projects. This may also be found in [**13**].

Another aspect of Lebesgue integration that we have not covered is integration over unbounded intervals. For example, we define $\mathcal{L}[a,+\infty)$ in the following manner.

(i) If $f \geq 0$, then $f \in \mathcal{L}[a, +\infty)$ if and only if

$$\lim_{N\to\infty} \int_a^N f$$

exists and is finite.

(ii) More generally, $f \in \mathcal{L}[a, +\infty)$ if and only if both $f^+ \in \mathcal{L}[a, +\infty)$ and $f^- \in \mathcal{L}[a, +\infty)$.

At first glance, this looks the same as improper Riemann integrals. However, the next example highlights the difference.

Example 2.5.3. Define f on $[0, +\infty)$ as the piecewise linear function connecting the points $(0, 0)$, $(1, 1)$, $(2, 0)$, $(3, -\frac{1}{2})$, $(4, 0)$, $(5, \frac{1}{3})$, $(6, 0)$, $(7, -\frac{1}{4})$, ... More explicitly, these points are $(n, g(n))$, where

$$g(n) = \begin{cases} 0 & \text{if } n = 2, 4, 6, \ldots, \\ \frac{2(-1)^{\frac{n-1}{2}}}{n+1} & \text{if } n = 1, 3, 5, \ldots. \end{cases}$$

The improper Riemann integral of f is

$$\int_0^\infty f(x)\,dx = \lim_{N\to\infty} \int_0^N f(x)\,dx = \sum_{n=1}^\infty \frac{(-1)^{n+1}}{n}.$$

This is the alternating harmonic series, which converges. Therefore, this improper Riemann integral exists.

On the other hand, to compute the Lebesgue integral we must first consider f^+:

$$\int_0^\infty f^+ = \lim_{N\to\infty} \int_0^N f^+ = \sum_{n=1}^\infty \frac{1}{2n-1}.$$

Therefore, $f \notin \mathcal{L}[0, +\infty)$.

2.6. Exercises

(1) Let $E \subseteq [a, b]$ and let $\mathcal{X}_E$ be the characteristic function of E. Prove that $\mathcal{X}_E(x)$ is a measurable function if and only if E is a measurable set.

(2) Suppose $f\colon [a, b] \to \mathbb{R}$ is a measurable function.

a) Show that the inverse image of a closed interval is a measurable set. Recall: the inverse image of a set C is

$$f^{-1}(C) = \{x \in [a,b] \mid f(x) \in C\}\,.$$

b) Show that the inverse image of an open set in $\mathbb{R}$ is a measurable set.

(3) Let $[c,d] \subseteq [a,b]$. Show that if f is measurable on $[a,b]$, then f is measurable on $[c,d]$.

(4) Find an example of a pointwise bounded sequence of measurable functions $\{f_n\}$ on $[0,1]$ such that each $f_n(x)$ is a bounded function but $f^*(x) = \limsup_{n\to\infty} f_n(x)$ is not a bounded function.

(5) Suppose $f\colon [a,b] \to \mathbb{R}$ is a strictly increasing function. Prove that f is a measurable function.

(6) Let f and g be measurable functions on an interval I.

a) Show that $\{x \in I \mid f(x) > g(x)\}$ is a measurable set.

b) Define

$$h(x) = \max\{f(x), g(x)\} = \begin{cases} f(x) & \text{if } f(x) \geq g(x), \\ g(x) & \text{otherwise.} \end{cases}$$

Show that $h(x)$ is a measurable function.

(7) Show that the function $f\colon [0,1] \to \mathcal{C}$ described in Example 1.1.7 is a measurable function.

(8) Suppose f is measurable on $I = [a,b]$ and $f(x) \geq 0$ a.e. on I. Prove that if the set $\{x \in I \mid f(x) > 0\}$ has positive measure, then for some positive integer n the set

$$E_n = \{x \in I \mid f(x) > \frac{1}{n}\}$$

has positive measure.

(9) Let f be a measurable function on the interval I. Show that f^2 is measurable on the interval I.

(10) Let f and g be measurable on an interval I. If $g(x) \neq 0$ on I, show that $\dfrac{f}{g}$ is a measurable function. (In other words, prove the third part of Theorem 2.1.6.)

(11) Suppose f and g are bounded functions on $[a, b]$. Let P be a measurable partition of $[a, b]$. State and prove a comparison between $U[f+g, P]$ and $U[f, P] + U[g, P]$. Do the same for lower sums.

(12) Let f be measurable on $[a, b]$ and let P_1 and P_2 be measurable partitions of $[a, b]$. If P_2 is a refinement of P_1, show that

$$L[f, P_1] \leq L[f, P_2] \qquad \text{and} \qquad U[f, P_2] \leq U[f, P_1] .$$

(13) Let f and g be bounded, Lebesgue integrable functions on $[a, b]$. Show that $f+g$ is Lebesgue integrable on $[a, b]$ and

$$\int_a^b (f+g) = \int_a^b f + \int_a^b g .$$

(Exercise 11 might be useful.)

(14) Let h be a bounded function that is zero a.e. in $[a, b]$. Show that h is Lebesgue integrable on $[a, b]$ and

$$\int_a^b h = 0 .$$

(15) Let φ be a simple function defined on $[a, b]$.
  a) Show that φ is measurable on $[a, b]$.
  b) Show that φ is Lebesgue integrable on $[a, b]$. Use the definition of the Lebesgue integral to compute

$$\int_a^b \varphi.$$

(16) Let $f \in \mathcal{L}[a, b]$. Show that if g is a bounded measurable function, then $fg \in \mathcal{L}[a, b]$.

(17) Prove or give a counterexample: If $f, g \in \mathcal{L}[a, b]$, then $fg \in \mathcal{L}[a, b]$.

(18) Let f be a differentiable function on $[a, b]$. Prove that f' is a measurable function. Hint: consider

$$\lim_{n \to \infty} \frac{f(x + \frac{1}{n}) - f(x)}{\frac{1}{n}} .$$

(19) Let $f \in \mathcal{L}[a, b]$ and A and B be measurable subsets of $[a, b]$.

a) If $A \cap B = \emptyset$, show that

$$\int_{A\cup B} f = \int_A f + \int_B f\,.$$

b) State and prove a result for the case that $A \cap B \neq \emptyset$.

c) What can you conclude if $A = [a, c]$ and $B = [c, b]$ for some $c \in (a, b)$?

(20) Let $f \in \mathcal{L}[a, b]$. For $x \in (a, b)$, define $G(x)$ by

$$G(x) = \int_a^x f\,.$$

Prove or give a counterexample: G is continuous on (a, b).

(21) Let f be the function defined on $[0, 1]$ by

$$f(x) = \begin{cases} n(-1)^n & \text{if } \frac{1}{n+1} < x \leq \frac{1}{n}, \\ 0 & \text{otherwise.} \end{cases}$$

a) Is f a bounded function?

b) Is $f \in \mathcal{L}[0, 1]$?

c) Does the improper Riemann integral $\displaystyle\int_0^1 f(x)\,dx$ exist?

(22) Let $f \in \mathcal{L}[a, b]$ and $\{A_k\}$ be a countable collection of pairwise disjoint measurable subsets of $[a, b]$. Explain why the statement

$$\int_{\cup_k A_k} f = \sum_k \int_{A_k} f$$

involves interchanging limit operations. Then prove it is true.

(23) Let $\{f_n\}$ be a sequence of funtions in $\mathcal{L}[a, b]$. Suppose that there exists a function $g \in \mathcal{L}[a, b]$ with $|f_n(x)| \leq g(x)$ a.e. for each n. Show that if

$$\lim_{n\to\infty} f_n(x) = f(x) \text{ a.e.}$$

and $h(x)$ is a bounded measurable function, then

$$\lim_{n\to\infty} \int_a^b f_n h = \int_a^b f h\,.$$

(24) Let (f_n) be a sequence of functions in $\mathcal{L}[a,b]$. Suppose $f \in \mathcal{L}[a,b]$ and

$$\lim_{n\to\infty} \int_a^b |f_n - f| = 0\,.$$

If the sequence (f_n) converges pointwise almost everywhere on $[a,b]$ to the function g, show that $f = g$ a.e. on $[a,b]$.

Suggestion: Consider the sequence $(|f - f_n|)$ and Fatou's Lemma.

http://dx.doi.org/10.1090/stml/078/04

Chapter 3

L^p spaces

We have already seen that $\mathcal{L}[a,b]$ is a vector space by Theorem 2.3.5. In more familiar vector spaces, $\mathbb{R}^n$ for example, we have additional features. These include the length of a vector and the dot product of two vectors. In this chapter we will generalize these ideas and, in doing so, explore new vector spaces known as L^p-spaces.

3.1. $L^1[a,b]$

Our first task is to define some sort of length function or norm on $\mathcal{L}[a,b]$. The goal is to mimic the structure of $\mathbb{R}$. That is, we would like to create a complete metric space. First we describe what is meant by a norm.

Definition 3.1.1. A **norm** on a vector space V is a function $\|\cdot\| : V \to \mathbb{R}$ that satisfies the following four properties for all $v, w \in V$ and $c \in \mathbb{R}$:

(i) $\|v\| \geq 0$.

(ii) $\|v\| = 0$ if and only if $v = \overline{0}$. Here $\overline{0}$ denotes the zero vector in V.

(iii) $\|c\,v\| = |c|\,\|v\|$.

(iv) $\|v + w\| \leq \|v\| + \|w\|$.

It doesn't take long to verify that an example of a norm on a vector space is the usual absolute value $|\cdot|$ on $\mathbb{R}$. Also, in the vector space of real numbers, the distance between two numbers $x, y \in \mathbb{R}$ is given by $|x-y|$. In a similar manner, if $\|\cdot\|$ is a norm on a vector space V, then $\|u-v\|$ can be viewed as the "distance" between u and v. By thinking in terms of distance, condition (iv) of a norm is often called the triangle inequality.

A candidate for a norm on $\mathcal{L}[a,b]$ is $\displaystyle\int_a^b |f|$. After all, it certainly is true that

$$\int_a^b |f| \geq 0$$

for all $f \in \mathcal{L}[a,b]$. Also,

$$\int_a^b |c\, f| = |c| \int_a^b |f|\,.$$

In addition,

$$\int_a^b |f+g| \leq \int_a^b (|f|+|g|) = \int_a^b |f| + \int_a^b |g|$$

whenever $f, g \in \mathcal{L}[a,b]$. However, if $\displaystyle\int_a^b |f| = 0$, we can only conclude that $f = 0$ a.e. in $[a,b]$. That leaves us with a lot of possibilities other than the identically 0 function. We are really close, though, to having a norm. To remedy this last problem we will start with the following proposition.

Proposition 3.1.2. *Define $\sim$ on $\mathcal{L}[a,b]$ by*

$$f \sim g \qquad \textit{if and only if} \qquad f = g \textit{ a.e.}$$

Then $\sim$ is an equivalence relation on $\mathcal{L}[a,b]$. That is, the following three conditions hold:

(i) *For all $f \in \mathcal{L}[a,b]$, $f \sim f$.*

(ii) *For all $f, g \in \mathcal{L}[a,b]$, if $f \sim g$, then $g \sim f$.*

(iii) *For all $f, g, h \in \mathcal{L}[a,b]$, if $f \sim g$ and $g \sim h$, then $f \sim h$*

Proof. This is proved in Exercise 3. $\square$

Our remedy is to deal with equivalence classes. We simply group together all equivalent functions and work with representatives of the resulting equivalence classes.

Definition 3.1.3. $L^1[a,b]$ is defined to be $\mathcal{L}[a,b]$ modulo the equivalence relation $\sim$.

This means that when we refer to a function $f \in L^1[a,b]$ we are using f to represent all functions that are equivalent to f. Ordinarily for two functions f and g to be considered the same function, $f(x) = g(x)$ for all $x \in [a,b]$. But we consider two functions f and g to be equal in $L^1[a,b]$ if $f, g \in \mathcal{L}[a,b]$ and $f = g$ a.e. on $[a,b]$. In other words, the difference between $\mathcal{L}[a,b]$ and $L^1[a,b]$ is that "equal almost everywhere" really is good enough!

Now we can tackle the issue of a norm on this new perspective of our Lebesgue integrable functions.

Definition 3.1.4. For $f \in L^1[a,b]$, we define the **L^1-norm of** f, written $\|f\|_1$, to be

$$\|f\|_1 = \int_a^b |f|\,.$$

We must justify the terminology L^1-norm by verifying that we do indeed have a norm.

Proposition 3.1.5. *$\|\cdot\|_1$ is a norm on $L^1[a,b]$.*

Proof. The only requirement for a norm that needs to be checked is part (ii) of Definition 3.1.1. If $\|f\|_1 = 0$, then

$$\int_a^b |f| = 0\,,$$

as mentioned earlier. By Theorem 2.3.9, $f = 0$ a.e. in $[a,b]$. Therefore, in $L^1[a,b]$, f is the zero vector. $\square$

We return to this idea of mimicking the structure of $\mathbb{R}$. The difference between $\mathbb{R}$, the set of real numbers, and $\mathbb{Q}$, the set of rational numbers, can be described by the behavior of sequences. It is possible for a sequence of rational numbers to converge to an irrational number (think of approximations to π or $\sqrt{2}$). On the other hand, if a sequence of real numbers converges, it must converge to a real number. This is described by a result usually covered in a first analysis course asserting that every Cauchy sequence of real numbers converges to a real number. We are going to generalize this idea to a vector space with a norm.

Definition 3.1.6. Let V be a vector space with norm $\|\cdot\|$. V is said to be **complete with respect to** $\|\cdot\|$ if every sequence that is Cauchy with respect to the norm $\|\cdot\|$ converges to some vector $v \in V$. In other words, whenever $\{v_n\}$ is a sequence in V with the property that given any $\epsilon > 0$ there is an N such that

$$\|v_n - v_m\| < \epsilon$$

whenever $m, n > N$, (that is, $\{v_n\}$ is Cauchy with respect to $\|\cdot\|$), then there exists a $v \in V$ with

$$\lim_{n\to\infty} \|v_n - v\| = 0\,.$$

As pointed out earlier, $\mathbb{R}$ is complete with respect to the norm $|\cdot|$, the familiar absolute value, whereas $\mathbb{Q}$ is not complete with respect to absolute value. Our goal is to show that $L^1[a,b]$ is complete with respect to $\|\cdot\|_1$. This is actually part of a more general concept.

Definition 3.1.7. A **Banach space** is a vector space V equipped with a norm $\|\cdot\|$ such that V is complete with respect to the norm $\|\cdot\|$.

As stated above, a basic example of a Banach space is $\mathbb{R}$ with norm $|\cdot|$, the familiar absolute value. Although the proof that $\mathbb{R}$ is complete with respect to this norm is usually covered in a first analysis course, we will once again prove this fact in order to motivate the proof that $L^1[a,b]$ is a Banach space. The proof we give here is probably not the proof you might have seen earlier. For this proof that $\mathbb{R}$ is a Banach space, the assumptions we will make are:

(i) every bounded increasing sequence of real numbers must converge to a real number,

(ii) every bounded sequence has a convergent subsequence, and

(iii) every absolutely convergent infinite series converges.

Using these assumptions we will now show that $\mathbb{R}$ is indeed a Banach space.

Theorem 3.1.8. $\mathbb{R}$ *is a Banach space with respect to the norm* $|\cdot|$.

Proof. Let $\{a_n\}$ be a sequence in $\mathbb{R}$ that is Cauchy with respect to $|\cdot|$. Thus, for every $\epsilon > 0$ there is an N so that

$$|a_n - a_m| < \epsilon \qquad \text{whenever} \qquad n, m > N\,.$$

In order to show that this sequence converges, we must first somehow come up with a reasonable "target". Our strategy will be to create a subsequence of $\{a_n\}$, show that this subsequence converges to some real number α, and then use this as our target for the original sequence.

Because $\{a_n\}$ is Cauchy with respect to $|\cdot|$, there is an N_1 so that

$$|a_n - a_m| < \frac{1}{2} \qquad \text{whenever} \qquad n, m > N_1\,.$$

Pick n_1 so that $n_1 > N_1$. The first number in our subsequence will be a_{n_1}. Next, there is an N_2 so that

$$|a_n - a_m| < \frac{1}{4} = \frac{1}{2^2} \qquad \text{whenever} \qquad n, m > N_2\,.$$

Pick n_2 so that $n_2 > N_2$ and $n_2 > n_1$. This gives us a_{n_2}, the next term in our sequence, with the added information that $|a_{n_1} - a_{n_2}| < \frac{1}{2}$. Next, there is an N_3 so that

$$|a_n - a_m| < \frac{1}{2^3} \qquad \text{whenever} \qquad n, m > N_3\,.$$

Pick n_3 so that $n_3 > N_3$ and $n_3 > n_2$. Hence, $|a_{n_2} - a_{n_3}| < \frac{1}{4}$.

More generally, assume $a_{n_1}, a_{n_2}, \ldots, a_{n_k}$ have been chosen in this fashion. There is an N_{k+1} so that

$$|a_n - a_m| < \frac{1}{2^{k+1}} \qquad \text{whenever} \qquad n, m > N_{k+1}\,.$$

Pick n_{k+1} so that $n_{k+1} > N_{k+1}$ and $n_{k+1} > n_k$. Hence we have $|a_{n_k} - a_{n_{k+1}}| < \frac{1}{2^k}$.

We have now created the subsequence $\{a_{n_k}\}_{k=1}^{\infty}$. Our next task is to show that this sequence converges to some real number α.

The geometric series $\sum_{k=1}^{\infty} \frac{1}{2^k}$ converges and

$$\sum_{k=1}^{\infty} \frac{1}{2^k} = 1\,.$$

Hence, for every M,

$$\sum_{k=1}^{M} |a_{n_k} - a_{n_{k+1}}| \le \sum_{k=1}^{M} \frac{1}{2^k} \le 1\,.$$

Therefore, the partial sums of the series $\sum_{k=1}^{\infty} |a_{n_k} - a_{n_{k+1}}|$ form a bounded increasing sequence. By the first of our assumptions this sequence of partial sums must converge. In other words, $\sum_{k=1}^{\infty}(a_{n_k} - a_{n_{k+1}})$ converges absolutely. Now the third assumption asserts that there exists a real number a so that

$$\sum_{k=1}^{\infty}(a_{n_k} - a_{n_{k+1}}) = a\,.$$

Thus,

$$\lim_{M\to\infty}\left(\sum_{k=1}^{M}(a_{n_k} - a_{n_{k+1}})\right) = \lim_{M\to\infty}(a_{n_1} - a_{n_{M+1}}) = a\,.$$

Consequently,

$$\lim_{M\to\infty} a_{n_{M+1}} = a_{n_1} - a\,.$$

We have now established that our subsequence $\{a_{n_k}\}_{k=1}^{\infty}$ converges to α where $\alpha = a_{n_1} - a$. The fact that our subsequence converges to α does not by itself guarantee that the original sequence converges to α. (In fact, if you think about it, you probably can come up with many examples of sequences that do not converge but have convergent subsequences.) Thus our final task is to return to the original sequence and show $\lim_{n\to\infty} |a_n - \alpha| = 0$.

Let $\epsilon > 0$. Because $\{a_n\}$ is Cauchy with respect to $|\cdot|$, there is an N so that

$$|a_n - a_m| < \frac{\epsilon}{2} \qquad \text{whenever} \qquad n, m > N\,.$$

Because $\{a_{n_k}\}_{k=1}^{\infty}$ converges to α there is a K so that

$$|a_{n_k} - \alpha| < \frac{\epsilon}{2} \qquad \text{whenever} \qquad k > K\,.$$

For $n > N$, choose k so that $n_k > N$ and $k > K$. Therefore, if $n > N$,

$$\begin{aligned} |a_n - \alpha| &\leq |a_n - a_{n_k}| + |a_{n_k} - \alpha| \\ &< \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon\,. \end{aligned}$$

Therefore, there is a real number α with $\lim\limits_{n\to\infty} |a_n - \alpha| = 0$. $\square$

As mentioned earlier, this is not the standard method of showing that $\mathbb{R}$ is complete with respect to $|\cdot|$, the usual norm. However, the basic structure of the proof is similar to the method we will use to show that $L^1[a, b]$ is a Banach space.

Let's consider some examples before proceeding. The goal of these examples is to illustrate the difference between pointwise convergence and convergence in $L^1[a, b]$.

Example 3.1.9. For $x \in [0, 1]$ let

$$f(x) = \begin{cases} 0 & \text{if } x = 0, \\ \frac{1}{x} & \text{if } x \neq 0 \end{cases}$$

and let $f_n(x) = f^n(x)$. (Here f^n is the capped version of f.) For each n, $f_n \in L^1[0, 1]$. Also, $\lim\limits_{n\to\infty} f_n(x) = f(x)$ for each $x \in [0, 1]$. However, $f \notin L^1[0, 1]$.

Hence, the pointwise limit of a sequence of L^1 functions is not necessarily an L^1 function.

Example 3.1.10. For $x \in [0, 1]$ and positive integer n, let

$$f_n(x) = \begin{cases} 2n^2 x & \text{if } 0 \leq x \leq \frac{1}{2n}, \\ -2n^2\left(x - \frac{1}{n}\right) & \text{if } \frac{1}{2n} < x \leq \frac{1}{n}, \\ 0 & \text{otherwise.} \end{cases}$$

This is the same sequence of functions discussed in Example 2.4.11. The pointwise limit of this sequence of functions is $f = 0$, which is certainly in $L^1[0, 1]$. However,

$$\lim_{n\to\infty} \|f_n - f\|_1 = \lim_{n\to\infty} \int_0^1 |f_n - f| = \lim_{n\to\infty} \int_0^1 f^n = \lim_{n\to\infty} \frac{1}{2} = \frac{1}{2}\,.$$

This example demonstrates that even if the pointwise limit of a sequence of L^1 functions is an L^1 function, the sequence of functions might not converge to this function in L^1-norm.

We now turn to the main tool needed to prove that $L^1[a,b]$ is a Banach space. It is the analog of the property of real numbers used earlier: absolutely convergent series converge.

Theorem 3.1.11 (Beppo Levi)**.** *Let $\{g_k\}$ be a sequence of functions in $L^1[a,b]$ such that $\sum_{k=1}^{\infty} \|g_k\|_1$ converges. Then there exists $g \in L^1[a,b]$ such that*

(i) $\sum_{k=1}^{\infty} g_k(x)$ *converges almost everywhere to $g(x)$ and*

(ii) $\sum_{k=1}^{\infty} \int_a^b g_k = \int_a^b g.$

Proof. As with Theorem 3.1.8, the first step is to describe a "target". Set

$$\sigma_n(x) = \sum_{k=1}^{n} |g_k(x)| \,.$$

Since σ_n is the finite sum of functions in $L^1[a,b]$, $\sigma_n \in L^1[a,b]$. Notice how for each $x \in [a,b]$ we are really considering absolute convergence of the series $\sum_{k=1}^{\infty} g_k(x)$. And, for each n,

$$\int_a^b \sigma_n = \int_a^b \sum_{k=1}^{n} |g_k| = \sum_{k=1}^{n} \int_a^b |g_k| = \sum_{k=1}^{n} \|g_k\|_1 \,.$$

Here, the interchange of integration and summation is not an issue because this is a finite sum. Pointwise, $\{\sigma_n\}$ is an increasing sequence. (Here, *pointwise* means to first pick a random $x \in [a,b]$ and then look at the resulting sequence of numbers $\{\sigma_n(x)\}$.) So, for each $x \in [a,b]$ either

$$\lim_{n\to\infty} \sigma_n(x) \text{ is finite or } \lim_{n\to\infty} \sigma_n(x) = +\infty \,.$$

Our first task is to show that $\lim_{n\to\infty} \sigma_n(x)$ is finite almost everywhere.

To see this, let $E = \{x \in [a,b] \mid \lim_{n\to\infty} \sigma_n(x) = +\infty\}$. We wish to show that $m(E) = 0$. For $N > 0$ set

$$\sigma^N(x) = \begin{cases} \lim_{n\to\infty} \sigma_n(x) & \text{if } \lim_{n\to\infty} \sigma_n(x) \leq N, \\ N & \text{otherwise.} \end{cases}$$

Essentially, σ^N is the capped version of $\lim_{n\to\infty} \sigma_n(x)$. In fact, for each fixed $N > 0$,

$$\lim_{n\to\infty} \sigma_n^N(x) = \sigma^N(x)$$

for all $x \in [a,b]$. Consequently, the sequence $\{\sigma_n^N\}$ satisfies the conditions of Theorem 2.4.13, the Monotone Convergence Theorem. Therefore,

$$\begin{aligned} \int_a^b \sigma^N &= \lim_{n\to\infty} \int_a^b \sigma_n^N \\ &\leq \lim_{n\to\infty} \int_a^b \sigma_n \\ &= \lim_{n\to\infty} \sum_{k=1}^{n} \|g_k\|_1 \\ &= \sum_{k=1}^{\infty} \|g_k\|_1 . \end{aligned}$$

Also,

$$Nm(E) \leq \int_a^b N\mathcal{X}_E \leq \int_a^b \sigma^N$$

for each $N > 0$. Combining these yields

$$m(E) \leq \frac{1}{N} \sum_{k=1}^{\infty} \|g_k\|_1 .$$

But this is true for each $N > 0$. By taking the limit as N goes to infinity we find that $m(E) = 0$.

We are getting ready to describe the target function. Set

$$\sigma(x) = \begin{cases} \lim_{n\to\infty} \sigma_n(x) & \text{if } x \notin E, \\ 0 & \text{if } x \in E. \end{cases}$$

By definition, $\lim_{n\to\infty} \sigma_n(x) = \sigma(x)$ a.e. in $[a,b]$. We need to show that the function $\sigma \in L^1[a,b]$. By design, $\{\sigma_n\}$ is a pointwise increasing sequence of nonnegative functions which converges almost everywhere to σ. We again have satisfied the hypotheses of Theorem 2.4.13, the Monotone Convergence Theorem. If to the contrary $\sigma \notin L^1[a,b]$, Theorem 2.4.13 would imply $\lim_{n\to\infty} \int_a^b \sigma_n = +\infty$. However,

$$\lim_{n\to\infty} \int_a^b \sigma_n = \lim_{n\to\infty} \sum_{k=1}^{n} \|g_k\|_1 = \sum_{k=1}^{\infty} \|g_k\|_1,$$

which is finite. Therefore $\sigma \in L^1[a,b]$.

This means that for almost every $x \in [a,b]$, $\sum_{k=1}^{\infty} g_k(x)$ is an absolutely convergent series. Thus, for almost every $x \in [a,b]$, $\sum_{k=1}^{\infty} g_k(x)$ converges. If we let

$$s_n(x) = \sum_{k=1}^{n} g_k(x)\,,$$

there exists a measurable function $g(x)$ such that $\lim_{n\to\infty} s_n(x) = g(x)$ a.e. in $[a,b]$. Since s_n is measurable for every n, g is measurable. Also,

$$\begin{aligned} |s_n(x)| &= \left| \sum_{k=1}^{n} g_k(x) \right| \\ &\le \sum_{k=1}^{n} |g_k(x)| \\ &= \sigma_n(x) \le \sigma(x) \end{aligned}$$

for almost every $x \in [a,b]$. Therefore, $|g(x)| \le \sigma(x)$ a.e. in $[a,b]$. Hence by Lemma 2.4.5, $g \in L^1[a,b]$.

Finally, by Theorem 2.4.6, the Lebesgue Dominated Convergence Theorem, using $\sigma(x)$ as the dominating function,

$$\lim_{n\to\infty} \int_a^b s_n = \int_a^b g\,.$$

In other words,

$$\begin{aligned}\sum_{k=1}^{\infty}\int_a^b g_k &= \lim_{n\to\infty}\left(\sum_{k=1}^{n}\int_a^b g_k\right)\\ &= \lim_{n\to\infty}\int_a^b\left(\sum_{k=1}^{n} g_k\right)\\ &= \lim_{n\to\infty}\int_a^b s_n = \int_a^b g\,.\end{aligned}$$

□

In the following example we will use this theorem to evaluate $\displaystyle\int_0^1\left(\frac{\ln x}{1-x}\right)^2 dx$. Before looking ahead, though, you should first think about tackling this integral by hand.

Example 3.1.12. For $0 \le x < 1$,

$$\frac{1}{(1-x)^2} = \sum_{k=1}^{\infty} k\,x^{k-1}\,.$$

Hence,

$$\left(\frac{\ln x}{1-x}\right)^2 = \frac{1}{(1-x)^2}(\ln x)^2 = \sum_{k=1}^{\infty} k\,x^{k-1}(\ln x)^2$$

for almost every $x \in [0,1]$. Setting $g_k(x) = x^{k-1}(\ln x)^2$ and noting integration by parts twice yields

$$\int_0^1 kx^{k-1}(\ln x)^2 = \frac{2}{k^2}\,.$$

By Theorem 3.1.11,

$$\int_0^1\left(\frac{\ln x}{1-x}\right)^2 dx = \sum_{k=1}^{\infty}\int_0^1 kx^{k-1}(\ln x)^2 = \sum_{k=1}^{\infty}\frac{2}{k^2} = \frac{\pi^2}{3}\,.$$

We now turn to the completeness of $L^1[a,b]$ with respect to the L^1-norm.

Theorem 3.1.13. *The space $L^1[a,b]$ is complete with respect to the norm $\|\cdot\|_1$, the L^1-norm.*

Proof. We must show that if $\{f_n\}$ is a sequence functions in $L^1[a,b]$ which is Cauchy with respect to the L^1-norm, then there exists a function $f \in L^1[a,b]$ such that

$$\lim_{n\to\infty} \|f_n - f\|_1 = 0\,.$$

Let $\{f_n\}$ be a sequence in $L^1[a,b]$ that is Cauchy with respect to the L^1-norm. Thus, for every $\epsilon > 0$ there is an N so that

$$\|f_n - f_m\|_1 < \epsilon \qquad \text{whenever} \qquad n, m > N\,.$$

As with our proof of the completeness of the real numbers, our strategy will be to create a candidate for a "target" function by judiciously choosing a subsequence of $\{f_n\}$, show that this subsequence converges to some function $f \in L^1[a,b]$, and then show that the original sequence converges to f in L^1-norm. The following should feel somewhat familiar by now.

Because $\{f_n\}$ is a Cauchy sequence, there is an N_1 so that

$$\|f_n - f_m\|_1 < \frac{1}{2} \qquad \text{whenever} \qquad n, m > N_1\,.$$

Pick n_1 so that $n_1 > N_1$. Next, there is an N_2 so that

$$\|f_n - f_m\|_1 < \frac{1}{4} = \frac{1}{2^2} \qquad \text{whenever} \qquad n, m > N_2\,.$$

Pick n_2 so that $n_2 > N_2$ and $n_2 > n_1$. Hence, $\|f_{n_1} - f_{n_2}\|_1 < \frac{1}{2}$. Next, there is an N_3 so that

$$\|f_n - f_m\|_1 < \frac{1}{2^3} \qquad \text{whenever} \qquad n, m > N_3\,.$$

Pick n_3 so that $n_3 > N_3$ and $n_3 > n_2$. Hence, $\|f_{n_2} - f_{n_3}\|_1 < \frac{1}{4}$.

More generally, assume $f_{n_1}, f_{n_2}, \ldots, f_{n_k}$ have been chosen in this fashion. There is an N_{k+1} so that

$$\|f_n - f_m\|_1 < \frac{1}{2^{k+1}} \qquad \text{whenever} \qquad n, m > N_{k+1}\,.$$

Pick n_{k+1} so that $n_{k+1} > N_{k+1}$ and $n_{k+1} > n_k$. Hence, $\|f_{n_k} - f_{n_{k+1}}\|_1 < \frac{1}{2^k}$.

We have now created the subsequence $\{f_{n_k}\}_{k=1}^{\infty}$. Our next task is to show that this subsequence converges to some $f \in L^1[a,b]$. Also, keep in mind that there are different ways a sequence of functions can converge. Two examples are pointwise and uniformly. But we seek a particular kind of convergence here. We need

$$\lim_{n\to\infty} \|f_n - f\|_1 = 0\,.$$

However, first things first. We must create our "target".

The geometric series $\displaystyle\sum_{k=1}^{\infty} \frac{1}{2^k}$ converges and

$$\sum_{k=1}^{\infty} \frac{1}{2^k} = 1\,.$$

Hence,

$$\sum_{k=1}^{M} \|f_{n_k} - f_{n_{k+1}}\|_1 \leq \sum_{k=1}^{M} \frac{1}{2^k} \leq 1$$

for every M. Therefore, $\displaystyle\sum_{k=1}^{\infty} \|f_{n_k} - f_{n_{k+1}}\|_1$ converges. By Theorem 3.1.11, using $g_k = f_{n_k} - f_{n_{k+1}}$, there exists $g \in L^1[a,b]$ with

$$\sum_{k=1}^{\infty} f_{n_k}(x) - f_{n_{k+1}}(x) = g(x) \text{ a.e. in } [a,b]\,.$$

But for almost every $x \in [a,b]$,

$$\lim_{M\to\infty}\left(\sum_{k=1}^{M}\big(f_{n_k}(x) - f_{n_{k+1}}(x)\big)\right) = \lim_{M\to\infty}\big(f_{n_1}(x) - f_{n_{M+1}}(x)\big) = g(x);$$

hence,

$$\lim_{M\to\infty} f_{n_{M+1}}(x) = f_{n_1}(x) - g(x) \text{ a.e. in } [a,b]\,.$$

We have now established that our subsequence $\{f_{n_k}(x)\}_{k=1}^{\infty}$ converges pointwise to $f \in L^1[a,b]$, where $f(x) = f_{n_1}(x) - g(x)$. Our final task is to show $\displaystyle\lim_{n\to\infty} \|f_n - f\|_1 = 0$. For every n,

$$\lim_{k\to\infty} |f_n(x) - f_{n_k}(x)| = |f_n(x) - f(x)| \text{ a.e. in } [a,b]\,.$$

Remember that f_{n_k} is from our subsequence, whereas f_n is from the original sequence. By Corollary 2.4.12, Fatou's Lemma, for each fixed n,

$$\int_a^b |f_n - f| \leq \liminf_{k\to\infty} \int_a^b |f_n - f_{n_k}|\,.$$

In other words,

$$\|f_n - f\|_1 \le \liminf_{k\to\infty} \|f_n - f_{n_k}\|_1 .$$

Let $\epsilon > 0$. Because $\{f_n\}$ is Cauchy with respect to $\|\cdot\|_1$, there is an N so that

$$\|f_n - f_m\|_1 < \frac{\epsilon}{2} \qquad \text{whenever} \qquad n, m > N .$$

Hence, for any $n > N$, $\liminf_{k\to\infty} \|f_n - f_{n_k}\|_1 \le \frac{\epsilon}{2} < \epsilon$. Consequently, if $n > N$, then

$$\|f_n - f\|_1 \le \liminf_{k\to\infty} \|f_n - f_{n_k}\|_1 < \epsilon .$$

Therefore,

$$\lim_{n\to\infty} \|f_n - f\|_1 = 0$$

and $L^1[a,b]$ is a Banach space. □

3.2. L^p Spaces

$L^1[a,b]$ is not the only Banach space that arises from Lebesgue integration. In fact, $L^1[a,b]$ is just one of a family of Banach spaces known as L^p spaces.

Definition 3.2.1. For $p \ge 1$ we define $L^p[a,b]$ to be

$$L^p[a,b] = \{f \mid f \text{ is measurable} \\ \text{and } |f|^p \text{ is Lebesgue integrable on } [a,b]\} .$$

(As with $L^1[a,b]$, it is assumed we are again really working with equivalence classes.)

Example 3.2.2. By Exercise 6, $f(x) = x^{-\frac{1}{3}} \in L^2[0,1]$ but $g(x) = x^{-\frac{2}{3}} \notin L^2[0,1]$.

In the previous example the careful reader might have observed that 0 is not necessarily in the domain of the function discussed. We have actually defined the function almost everywhere. If one is bothered by this, it is easy to merely define the functions to take on the value 0 when $x = 0$ and the results do not change. Here is another example of where almost everywhere is (almost always) good enough!

Proposition 3.2.3. *Let $f, g \in L^p[a,b]$, and $c \in \mathbb{R}$.*

(i) $cf \in L^p[a,b]$.

(ii) $f + g \in L^p[a,b]$.

Proof. (i) If $f \in L^p[a,b]$, then $|f|^p$ is Lebesgue integrable. As a consequence, $|c|^p|f|^p = |cf|^p$ is Lebesgue integrable. Hence, $cf \in L^p[a,b]$.

(ii) Since both f and g are measurable, $f+g$ must be measurable. We will show that $|f+g|^p$ is bounded by a Lebesgue integrable function and apply Lemma 2.4.5. For every $x \in [a,b]$,

$$\begin{aligned} |f(x)+g(x)|^p &\leq (|f(x)|+|g(x)|)^p \\ &\leq (2\,\max\{|f(x)|,|g(x)|\})^p \\ &= 2^p\left(\max\{|f(x)|^p,|g(x)|^p\}\right) \\ &\leq 2^p\left(|f(x)|^p+|g(x)|^p\right)\,. \end{aligned}$$

Since $2^p\left(|f(x)|^p+|g(x)|^p\right)$ is Lebesgue integrable, $f+g \in L^p[a,b]$. □

The previous proposition verifies that $L^p[a,b]$ is a vector space. Our goal is to show that $L^p[a,b]$ is a Banach space. To do this, we first need to define a norm on $L^p[a,b]$.

Definition 3.2.4. Let $f \in L^p[a,b]$. The ***L^p*-norm of** f, written $\|f\|_p$, is

$$\|f\|_p = \left(\int_a^b |f|^p\right)^{\frac{1}{p}}\,.$$

Although we have called $\|f\|_p$ a norm, we need to verify that it satisfies the definition of a norm. Most of the properties follow quite easily.

(i) We must show $\|f\|_p \geq 0$. Since $|f(x)|^p \geq 0$ for all $x \in [a,b]$,

$$\left(\int_a^b |f|^p\right)^{\frac{1}{p}} \geq 0\,.$$

Consequently, $\|f\|_p \geq 0$.

(ii) Next we will show that $\|f\|_p = 0$ if and only if $f = 0$ a.e. in $[a, b]$:

$$\begin{aligned}
\|f\|_p = 0 \quad & \text{if and only if} \quad \left(\int_a^b |f|^p\right)^{\frac{1}{p}} = 0 \\
& \text{if and only if} \quad \int_a^b |f|^p = 0 \\
& \text{if and only if} \quad |f|^p = 0 \text{ a.e.} \\
& \text{if and only if} \quad f = 0 \text{ a.e.}
\end{aligned}$$

(iii) To see that $\|cf\|_p = |c| \, \|f\|_p$,

$$\begin{aligned}
\|cf\|_p &= \left(\int_a^b |cf|^p\right)^{\frac{1}{p}} = \left(\int_a^b |c|^p \, |f|^p\right)^{\frac{1}{p}} \\
&= \left(|c|^p \int_a^b |f|^p\right)^{\frac{1}{p}} = |c| \left(\int_a^b |f|^p\right)^{\frac{1}{p}} = |c| \, \|f\|_p \, .
\end{aligned}$$

(The reason for the power of "$\frac{1}{p}$" should now be apparent.)

(iv) The final property we need to verify is that $\|f + g\|_p \leq \|f\|_p + \|g\|_p$. Unlike the first three, this property isn't as easy to verify at this time. We will return to this later.

One of the tools needed to verify this fourth property is known as Hölder's Inequality.

Theorem 3.2.5 (Hölder's Inequality)**.** *Let $f \in L^p[a, b]$ and $g \in L^q[a, b]$, where $p > 1$ and $\frac{1}{p} + \frac{1}{q} = 1$. Then $fg \in L^1[a, b]$ and $\|fg\|_1 \leq \|f\|_p \|g\|_q$.*

Before proving this theorem, let's examine the condition $\frac{1}{p} + \frac{1}{q} = 1$. This means there is a required connection between p and q. For example, if $p = \frac{3}{2}$, then q must be equal to 3. On the other hand, if $p = 9$, then $q = \frac{9}{8}$. If $p = 2$, then $q = 2$. More generally, $q = \frac{p}{p-1}$.

Here is a lemma we need to prove Hölder's Inequality.

Lemma 3.2.6. *Let $\alpha, \beta \in (0, 1)$ with $\alpha + \beta = 1$. Then for any nonnegative numbers a and b,*

$$ab \leq \alpha a^{\frac{1}{\alpha}} + \beta b^{\frac{1}{\beta}} \, .$$

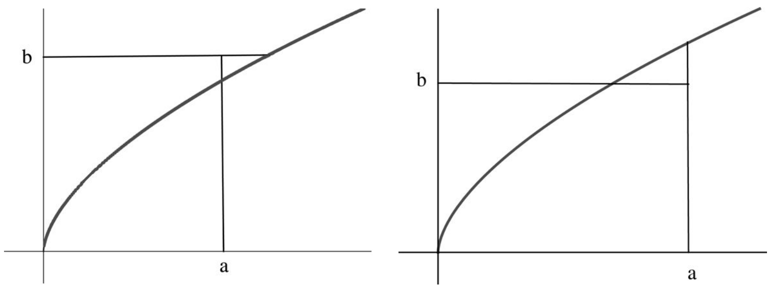

Figure 3.1. Graph of $y = x^\delta$ under the first two possibilities.

Proof. Let $\delta > 0$. (We will specify δ later.) Consider the graph of $y = x^\delta$ for $x \geq 0$. We will look at the three possibilities: $b < a^\delta$, $a^\delta < b$, and $a^\delta = b$.

(i) Suppose $b < a^\delta$. In this case, the horizontal line $y = b$ intersects the graph of $y = x^\delta$ to the left of the vertical line $x = a$. (See the left example in Figure 3.1.) Thus the area of the rectangle formed by the axes and the lines $y = b$ and $x = a$, which equals ab, is less than $\int_0^a x^\delta \, dx$ (the area under the curve $y = x^\delta$) plus $\int_0^b y^{\frac{1}{\delta}} \, dy$ (the remaining area inside the rectangle but above the curve $y = x^\delta$, integrating with respect to y). That is,

$$ab \leq \int_0^a x^\delta \, dx + \int_0^b y^{\frac{1}{\delta}} \, dy \,.$$

(ii) Suppose $a^\delta < b$. This time the horizontal line $y = b$ intersects the graph of $y = x^\delta$ to the right of the vertical line $x = a$. (See the right example in Figure 3.1.) Thus the area of the rectangle formed by the axes and the lines $y = b$ and $x = a$, which equals ab, is less than $\int_0^b y^{\frac{1}{\delta}} \, dy$ (the area of the region in the first quadrant bounded above by the line $y = b$ and below by the curve $y = x^\delta$, integrating with respect to y) plus $\int_0^a x^\delta \, dx$ (the remaining area inside the rectangle

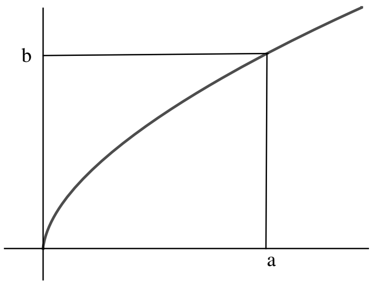

Figure 3.2. Graph of $y = x^\delta$ with $a^\delta = b$.

but below the curve $y = x^\delta$, integrating with respect to y). That is, once again

$$ab \le \int_0^a x^\delta \, dx + \int_0^b y^{\frac{1}{\delta}} \, dy \, .$$

(iii) Suppose $a^\delta = b$ as illustrated in Figure 3.2. In this final case, the two lines $x = a$ and $y = b$ intersect at the point (a, b), which is on the curve $y = x^\delta$. Thus the area of the rectangle formed by the axes and the lines $y = b$ and $x = a$, which equals ab, equals $\int_0^a x^\delta \, dx$ (the area under the curve $y = x^\delta$) plus $\int_0^b y^{\frac{1}{\delta}} \, dy$ (the remaining area inside the rectangle but above the curve $y = x^\delta$, integrating with respect to y). That is,

$$ab = \int_0^a x^\delta \, dx + \int_0^b y^{\frac{1}{\delta}} \, dy \, .$$

In all cases, then,

$$\begin{aligned} ab &\le \int_0^a x^\delta \, dx + \int_0^b y^{\frac{1}{\delta}} \, dy \\ &= \frac{a^{\delta+1}}{\delta+1} + \frac{b^{\frac{1}{\delta}+1}}{\frac{1}{\delta}+1} \\ &= \left(\frac{1}{\delta+1}\right) a^{\delta+1} + \left(\frac{\delta}{\delta+1}\right) b^{\frac{\delta+1}{\delta}} \, . \end{aligned}$$

The lemma follows once we choose δ so that $\frac{1}{\delta+1} = \alpha$ and $\frac{\delta}{\delta+1} = \beta$. $\square$

Now we will prove Theorem 3.2.5, Hölder's Inequality.

Proof. Suppose $f \in L^p[a,b]$ and $g \in L^q[a,b]$, where $\frac{1}{p} + \frac{1}{q} = 1$. By Lemma 3.2.6 with $\alpha = \frac{1}{p}$ and $\beta = \frac{1}{q}$,

$$|f(x)\, g(x)| \leq \frac{1}{p}|f(x)|^p + \frac{1}{q}|g(x)|^q$$

for every $x \in [a,b]$. By Lemma 2.4.5, $fg \in L^1[a,b]$ since $|f|^p$ and $|g|^q$ are both Lebesgue integrable.

To prove the inequality

$$\|fg\|_1 \leq \|f\|_p \|g\|_q\,,$$

we will first observe that this is easily true if either $\|f\|_p = 0$ (that is, $f = 0$ a.e.) or $\|g\|_q = 0$. Therefore, we will assume $\|f\|_p > 0$ and $\|g\|_q > 0$.

We will first look at the special case where $\|f\|_p = \|g\|_q = 1$. As noted above, Lemma 3.2.6 guarantees that

$$|f(x)\, g(x)| \leq \frac{1}{p}|f(x)|^p + \frac{1}{q}|g(x)|^q\,;$$

therefore

$$\int_a^b |fg| \leq \frac{1}{p}\int_a^b |f|^p + \frac{1}{q}\int_a^b |g|^q\,.$$

In other words,

$$\begin{aligned} \|fg\|_1 &\leq \frac{1}{p}\left(\|f\|_p\right)^p + \frac{1}{q}\left(\|g\|_q\right)^q \\ &= \frac{1}{p} + \frac{1}{q} = 1 = \|f\|_p \|g\|_q \end{aligned}$$

(using the assumption that $\|f\|_p = \|g\|_q = 1$), and we are done in this case.

The more general case follows by setting

$$\tilde{f}(x) = \frac{f(x)}{\|f\|_p} \qquad \text{and} \qquad \tilde{g}(x) = \frac{g(x)}{\|g\|_q}\,.$$

Then $\|\tilde{f}\|_p = 1$ and $\|\tilde{g}\|_q = 1$. Therefore, from our previous special case, $\|\tilde{f}\tilde{g}\|_1 \leq 1$. Hence,

$$\int_a^b \frac{|fg|}{\|f\|_p \|g\|_q} \leq 1$$

or

$$\int_a^b |fg| \leq \|f\|_p \|g\|_q \,.$$

In other words, $\|fg\|_1 \leq \|f\|_p \|g\|_q$, as claimed. □

The theorem that actually completes the final requirement for showing that $\| \cdot \|_p$ is a norm on $L^p[a,b]$ is the following.

Theorem 3.2.7 (Minkowski's Inequality). *Let $p \geq 1$. If $f, g \in L^p[a,b]$, then*

$$\|f + g\|_p \leq \|f\|_p + \|g\|_p \,.$$

Proof. We already have this result for the case $p = 1$, so assume $p > 1$. This result is trivially true if $|f + g| = 0$ a.e. in $[a, b]$. (Make sure you understand why this is deemed "trivial".) Hence, we will assume $|f + g| > 0$.

In Proposition 3.2.3 we showed that $|f + g|^p$ is Lebesgue integrable. We will look at this further. Let $q = \frac{p}{p-1}$ (remember, $p > 1$). Then $\frac{1}{p} + \frac{1}{q} = 1$. Note that

$$\begin{aligned}
(\|f + g\|_p)^{p-1} &= \left(\int_a^b |f + g|^p \right)^{\frac{p-1}{p}} \\
&= \left(\int_a^b \left(|f + g|^{p-1}\right)^{\frac{p}{p-1}} \right)^{\frac{p-1}{p}} \\
&= \|(|f + g|^{p-1})\|_q \,.
\end{aligned}$$

Therefore, $|f + g|^{p-1} \in L^q[a,b]$ and $\|(|f + g|^{p-1})\|_q = (\|f + g\|_p)^{p-1}$.

Also,

$$\begin{aligned}
|f(x) + g(x)|^p &= |(f(x) + g(x))^p| \\
&= \left| f(x)\,(f(x) + g(x))^{p-1} + g(x)\,(f(x) + g(x))^{p-1} \right| \\
&\leq |f(x)|\,|f(x) + g(x)|^{p-1} + |g(x)|\,|f(x) + g(x)|^{p-1} \,.
\end{aligned}$$

Thus,

$$\begin{aligned}(\|f+g\|_p)^p &= \int_a^b |f+g|^p \\ &\le \int_a^b |f(x)|\,|f(x)+g(x)|^{p-1} + \int_a^b |g(x)|\,|f(x)+g(x)|^{p-1} \\ &\le \|f\|_p \|(|f+g|^{p-1})\|_q + \|g\|_p \|(|f+g|^{p-1})\|_q\end{aligned}$$

using Hölder's Inequality for the last line. Therefore,

$$(\|f+g\|_p)^p \le \|f\|_p (\|f+g\|_p)^{p-1} + \|g\|_p (\|f+g\|_p)^{p-1}$$

so that

$$\|f+g\|_p \le \|f\|_p + \|g\|_p\,.$$

□

Now that we have completed the last step in showing that $\|\cdot\|_p$ is a norm on $L^p[a,b]$, we will now verify that $L^p[a,b]$ is a Banach space for $p > 1$. The proof is very similar to the proof of Theorem 3.1.13.

Theorem 3.2.8. *For $p > 1$, the space $L^p[a,b]$ is complete with respect to the norm $\|\cdot\|_p$, the L^p-norm.*

Proof. Let $p > 1$ and suppose $\{f_n\}$ is a sequence functions in $L^p[a,b]$ which is Cauchy with respect to the L^p-norm. We must show there exists a function $f \in L^p[a,b]$ such that

$$\lim_{n\to\infty} \|f_n - f\|_p = 0\,.$$

(Hopefully, you have already guessed that our strategy will be to create a subsequence of $\{f_n\}$, show that this subsequence converges to some function $f \in L^p[a,b]$, and then show that the original sequence converges to f in the L^p-norm. You also should be able to guess how we will start.)

Because $\{f_n\}$ is a Cauchy sequence, there is an N_1 so that

$$\|f_n - f_m\|_p < \frac{1}{2} \qquad \text{whenever} \qquad n, m > N_1\,.$$

Pick n_1 so that $n_1 > N_1$. Next, there is an N_2 so that

$$\|f_n - f_m\|_p < \frac{1}{2^2} \qquad \text{whenever} \qquad n, m > N_2\,.$$

Pick n_2 so that $n_2 > N_2$ and $n_2 > n_1$. Hence, $\|f_{n_1} - f_{n_2}\|_p < \frac{1}{2}$. Next, there is an N_3 so that

$$\|f_n - f_m\|_p < \frac{1}{2^3} \qquad \text{whenever} \qquad n, m > N_3 \,.$$

Pick n_3 so that $n_3 > N_3$ and $n_3 > n_2$. Hence, $\|f_{n_2} - f_{n_3}\|_p < \frac{1}{4}$.

More generally, assume $f_{n_1}, f_{n_2}, \ldots, f_{n_k}$ have been chosen in this fashion. There is an N_{k+1} so that

$$\|f_n - f_m\|_p < \frac{1}{2^{k+1}} \qquad \text{whenever} \qquad n, m > N_{k+1} \,.$$

Pick n_{k+1} so that $n_{k+1} > N_{k+1}$ and $n_{k+1} > n_k$. Hence, $\|f_{n_k} - f_{n_{k+1}}\|_p < \frac{1}{2^k}$.

We now have our subsequence $\{f_{n_k}\}_{k=1}^{\infty}$. Our next task is to show that this sequence converges to some function $f \in L^p[a,b]$.

Let

$$g_m(x) = \sum_{k=1}^{m} \left| f_{n_k}(x) - f_{n_{k+1}}(x) \right| \,.$$

Then by Minkowski's Inequality (Theorem 3.2.7),

$$\|g_m\|_p \le \sum_{k=1}^{M} \|f_{n_k} - f_{n_{k+1}}\|_p \le \sum_{k=1}^{M} \frac{1}{2^k} \le 1 \,.$$

Also, $0 \le g_m(x) \le g_{m+1}(x)$ for all m. In an argument similar to that in Theorem 3.1.11, we can show there is a function $g(x)$ with

$$\lim_{m\to\infty} g_m(x) = g(x) \text{ a.e.}$$

(Note: this is not a short argument. It means essentially proving an L^p version of Theorem 3.1.11.) Moreover, by the Monotone Convergence Theorem (Theorem 2.4.13) applied to $\{(g_m)^p\}$, $g^p \in L^1[a,b]$ and

$$\lim_{m\to\infty} \int_a^b |g_m|^p = \int_a^b |g|^p \,.$$

Hence, the series $\displaystyle\sum_{k=1}^{\infty} \big(f_{n_k}(x) - f_{n_{k+1}}(x)\big)$ converges absolutely for almost every x in $[a,b]$. That is, there is a function $\tilde{f}$ with

$$\lim_{m\to\infty} \sum_{k=1}^{m} \big(f_{n_k}(x) - f_{n_{k+1}}(x)\big) = \lim_{m\to\infty} \big(f_{n_1}(x) - f_{n_{m+1}}(x)\big) = \tilde{f}(x) \text{ a.e.}$$

Moreover, $|\tilde{f}(x)| \le g(x)$ a.e.; thus $|\tilde{f}|^p \in L^1[a,b]$ by Lemma 2.4.5. In other words, $\tilde{f} \in L^p[a,b]$. Set

$$f = f_{n_1} - \tilde{f}\,.$$

We now have our target $f \in L^p[a,b]$.

Our final task is to show $\lim_{n\to\infty} \|f_n - f\|_p = 0$. For every n,

$$\lim_{k\to\infty} |f_n(x) - f_{n_k}(x)|^p = |f_n(x) - f(x)|^p \text{ a.e. in } [a,b]\,.$$

By Corollary 2.4.12, Fatou's Lemma,

$$\int_a^b |f_n - f|^p \le \liminf_{k\to\infty} \int_a^b |f_n - f_{n_k}|^p$$

or

$$\left(\int_a^b |f_n - f|^p\right)^{\frac{1}{p}} \le \liminf_{k\to\infty} \left(\int_a^b |f_n - f_{n_k}|^p\right)^{\frac{1}{p}}\,.$$

That is,

$$\|f_n - f\|_p \le \liminf_{k\to\infty} \|f_n - f_{n_k}\|_p\,.$$

Let $\epsilon > 0$. Because $\{f_n\}$ is Cauchy with respect to $\|\cdot\|_p$, there is an N so that

$$\|f_n - f_m\|_p < \frac{\epsilon}{2} \qquad \text{whenever} \qquad n, m > N\,.$$

Hence, for any $n > N$, $\liminf_{k\to\infty} \|f_n - f_{n_k}\|_p \le \frac{\epsilon}{2} < \epsilon$. Consequently, if $n > N$, then

$$\|f_n - f\|_p \le \liminf_{k\to\infty} \|f_n - f_{n_k}\|_p < \epsilon\,.$$

Therefore,

$$\lim_{n\to\infty} \|f_n - f\|_p = 0$$

and $L^p[a,b]$ is a Banach space. □

We have defined $L^p[a,b]$ for $p \ge 1$. We will also define the space $L^\infty[a,b]$. We first need a preliminary definition.

Definition 3.2.9. Let $f : [a,b] \to \mathbb{R}$ be a measurable function.

(i) The **essential supremum** of f is

$$\operatorname*{ess\,sup}_{x\in[a,b]} f = \inf\{\alpha \mid f(x) \le \alpha \text{ a.e.}\}\,.$$

(ii) The **essential infimum** of f is

$$\operatorname*{ess\,inf}_{x\in[a,b]} f = \sup\{\beta \mid f(x) \geq \beta \text{ a.e.}\}\,.$$

Example 3.2.10. For $x \in [0,1]$ let

$$f(x) = \begin{cases} 2x & \text{if } x \notin \mathbb{Q}, \\ q & \text{if } x = \frac{p}{q}. \end{cases}$$

Here we are assuming $\frac{p}{q}$ is in the lowest terms. In this case,

$$\sup_{x\in[0,1]} f = +\infty,$$

while

$$\operatorname*{ess\,sup}_{x\in[0,1]} f = 2\,.$$

Example 3.2.11. For $x \in [0,1]$ let

$$g(x) = \begin{cases} \frac{1}{x} & \text{if } x \notin \mathbb{Q}, \\ 0 & \text{if } x = 0, \\ q & \text{if } x \neq 0 \text{ and } x = \frac{p}{q}. \end{cases}$$

Again, we are assuming $\frac{p}{q}$ is in the lowest terms. In this case,

$$\operatorname*{ess\,sup}_{x\in[0,1]} g = +\infty\,.$$

Definition 3.2.12. The space $L^\infty[a,b]$ is defined as

$$L^\infty[a,b] = \left\{ f \;\middle|\; f \text{ is measurable and } \operatorname*{ess\,sup}_{x\in[a,b]} |f| \text{ is finite} \right\}.$$

For $f \in L^\infty[a,b]$ we define $\|f\|_\infty$ as

$$\|f\|_\infty = \operatorname*{ess\,sup}_{x\in[a,b]} |f|\,.$$

As shown in Exercise 12, the space $L^\infty[a,b]$ is a vector space. Furthermore, $\|\cdot\|_\infty$ is a norm on this space in the next proposition as long as we work with the same sort of equivalence classes as we did with $L^p[a,b]$ for $1 \leq p < +\infty$.

Proposition 3.2.13. *$\|\cdot\|_\infty$ is a norm on the space $L^\infty[a,b]$.*

Proof. This is Exercise 13. ☐

The space $L^\infty[a,b]$ is also a Banach space. Again this is left to the reader as an exercise.

3.3. Approximations in $L^p[a,b]$

Since the set of rational numbers is countable and the set of irrational numbers is uncountable, we know that there are far more irrationals than rationals. Yet, we often approximate irrationals by rationals. For example, $\pi \approx 3.14$ or $e \approx 2.718$. We do this because we are more familiar with the rational numbers. In calculus, the vast majority of functions one encounters are continuous functions. Hence, we are more familiar with continuous functions. However, most of the functions in $L^p[a,b]$ are not continuous. In this section we will show that every L^p function can be approximated by a continuous function. In other words, we will show that the space of continuous functions is dense in $L^p[a,b]$. That is, given $f \in L^p[a,b]$ and $\epsilon > 0$, there is a continuous function g such that $\|f - g\|_p < \epsilon$. We will do this in stages.

Lemma 3.3.1. *Let $p \geq 1$ and $f \in L^p[a,b]$. Given $\epsilon > 0$ there is a bounded function g such that $\|f - g\|_p < \epsilon$.*

Proof. If f is bounded, we are done, so we will assume f is unbounded. For $N > 0$ set

$$g^N(x) = \begin{cases} f(x) & \text{if } |f(x)| \leq N, \\ N & \text{if } f(x) > N, \\ -N & \text{if } f(x) < -N. \end{cases}$$

For each N, g^N is a bounded function. Moreover, for every $x \in [a,b]$,

$$\lim_{N\to\infty} |f(x) - g^N(x)|^p = 0$$

and

$$|f(x) - g^N(x)|^p \leq \left(|f(x)| + |g^N(x)|\right)^p \leq (2|f(x)|)^p = 2^p|f(x)|^p \,.$$

By Theorem 2.4.6, the Lebesgue Dominated Convergence Theorem, using $2^p|f(x)|^p$ as the dominating function,

$$\lim_{N\to\infty} \int_a^b |f - g^N|^p = \int_a^b 0 = 0 \,.$$

Therefore,

$$\lim_{N\to\infty} \|f - g^N\|_p = 0\,.$$

Finally, given $\epsilon > 0$, choose N sufficiently large so that $\|f - g^N\|_p < \epsilon$ and set $g = g^N$. □

This lemma shows that we can approximate functions in $L^p[a, b]$ by bounded functions. Remember our goal in this section is to show that all functions in $L^p[a, b]$ can be approximated in $L^p[a, b]$ by continuous functions. The next stage in this process is to show that bounded functions can be approximated in $L^p[a, b]$ by simple functions.

Lemma 3.3.2. *Let $p \geq 1$ and $f \in L^p[a, b]$ be a bounded function. Then there exists a simple function*

$$\phi = \sum_{i=1}^{n} a_i \mathcal{X}_{A_i}$$

such that $\|f - \phi\|_p < \epsilon$.

Proof. This proof is reminiscent of Theorem 2.2.12. Since f is bounded there exists an $M > 0$ with $-M < f(x) < M$ for all $x \in [a, b]$. Let

$$-M = y_0 < y_1 < y_2 < \ldots < y_n = M$$

where $y_1, y_2, \ldots, y_n$ are chosen so that $y_i - y_{i-1} < \frac{\epsilon}{(b-a)^{\frac{1}{p}}}$ for $i = 1, 2, \ldots, n$. Set

$$A_i = \{x \in [a, b] \mid y_{i-1} \leq f(x) < y_i\}$$

and $a_i = y_{i-1}$ for $i = 1, 2, \ldots, n$. Define ϕ to be

$$\phi = \sum_{i=1}^{n} a_i \mathcal{X}_{A_i}\,.$$

If $x \in A_i$ then $|f(x) - \phi(x)| < \dfrac{\epsilon}{(b-a)^{\frac{1}{p}}}$. Additionally, each $x \in [a,b]$ is in A_i for exactly one i. Therefore,

$$\begin{aligned} \|f-\phi\|_p &= \left(\int_a^b |f-\phi|^p\right)^{\frac{1}{p}} \\ &\le \left(\int_a^b \left(\frac{\epsilon}{(b-a)^{\frac{1}{p}}}\right)^p\right)^{\frac{1}{p}} \\ &= \left((b-a)\frac{\epsilon^p}{(b-a)}\right)^{\frac{1}{p}} = \epsilon\,. \end{aligned}$$ □

Thus, every function in $L^p[a,b]$ can be approximated by a simple function. Our final step will be to show that every simple function can be approximated in $L^p[a,b]$ by a continuous function. In fact, it suffices to show that any characteristic function of a measurable set can by approximated in $L^p[a,b]$ by a continuous function.

Lemma 3.3.3. *Let $E \subseteq [a,b]$ be a measurable set. Given $\epsilon > 0$ and $p \ge 1$ there is a continous function h so that $\|\mathcal{X}_E - h\|_p < \epsilon$.*

Proof. Let $\epsilon > 0$ be given. By the definition of a measurable set, there is an open set G containing E such that $m(G \backslash E) < \epsilon^p/2$. There is also a closed set F contained in E such that $m(E \backslash F) < \epsilon^p/2$. Thus, $F \subseteq E \subseteq G$ and $F \cap G^c = \emptyset$. Also, $m(G \setminus F) < \epsilon^p$.

Since F and G^c are two disjoint closed sets, we may define the function

$$h(x) = \frac{d(x,G^c)}{d(x,F) + d(x,G^c)}\,.$$

The denominator is never 0, so this function is always defined. Also, as functions of x, $d(x,F)$ and $d(x,G^c)$ are continuous. Consequently, we have a continuous function with $0 \le h(x) \le 1$ for all $x \in [a,b]$.

Now, if $x \in G^c$ then $h(x) = 0 = \mathcal{X}_E(x)$. Similarly, if $x \in F$, then $h(x) = 1 = \mathcal{X}_E(x)$. Thus, $|\mathcal{X}_E(x) - h(x)| = 0$ if $x \in F$ or $x \in G^c$.

In addition, if $x \in G \setminus F$, then $\mathcal{X}_E(x) = 0$ or 1 and $0 \le h(x) \le 1$. As a result, $|\mathcal{X}_E(x) - h(x)| \le 1 = \mathcal{X}(G \setminus F)$ if $x \in G \setminus F$. Therefore,

$$|\mathcal{X}_E(x) - h(x)|^p \le \big(\mathcal{X}_{G\setminus F}(x)\big)^p = \mathcal{X}_{G\setminus F}(x)$$

and

$$\begin{aligned}
\|\mathcal{X}_E - h\|_p &= \left(\int_a^b |\mathcal{X}_E - h|^p\right)^{\frac{1}{p}} \\
&\leq \left(\int_a^b (\mathcal{X}_{G\setminus F})^p\right)^{\frac{1}{p}} \\
&= \left(\int_a^b \mathcal{X}_{G\setminus F}\right)^{\frac{1}{p}} \\
&= (m(G \setminus F))^{\frac{1}{p}} \\
&< (\epsilon^p)^{\frac{1}{p}} = \epsilon\,. \qquad \square
\end{aligned}$$

Corollary 3.3.4. *Let $p \geq 1$ and ϕ be a simple function*

$$\phi = \sum_{i=1}^{n} a_i \mathcal{X}_{A_i},$$

where A_i is a measurable subset of $[a,b]$ for each i. Then for every $\epsilon > 0$ there exists a continuous function g such that $\|\phi - g\|_p < \epsilon$.

Proof. This is Exercise 23. $\square$

Theorem 3.3.5. *Let $p \geq 1$ and $f \in L^p[a,b]$. For every $\epsilon > 0$ there is a continuous function g such that $\|f - g\|_p \leq \epsilon$.*

Proof. Let $\epsilon > 0$ be given. By Lemma 3.3.1, there is a bounded function g^N with $\|f - g^N\|_p < \frac{\epsilon}{2}$. By Lemma 3.3.2, there is a simple function ϕ such that $\|g^N - \phi\|_p < \frac{\epsilon}{4}$. Finally, by Corollary 3.3.4, there is a continuous function g with $\|\phi - g\|_p < \frac{\epsilon}{4}$. Therefore,

$$\begin{aligned}
\|f - g\|_p &\leq \|f - g^N\|_p + \|g^N - \phi\|_p + \|\phi - g\|_p \\
&< \frac{\epsilon}{2} + \frac{\epsilon}{4} + \frac{\epsilon}{4} = \epsilon\,. \qquad \square
\end{aligned}$$

3.4. $L^2[a,b]$

$L^2[a,b]$ holds a special place among the L^p-spaces. It is the only one of the L^p-spaces that is a Hilbert space. A Hilbert space, as we shall see, is a Banach space whose norm comes from an inner product.

Before we officially make this definition, we need the definition of an inner product. This is really extending the notion of the dot product on $\mathbb{R}^n$.

Definition 3.4.1. Let V be a vector space. An **inner product** on V is a function $\langle \cdot, \cdot \rangle$ from $V \times V$ to $\mathbb{R}$ such that for every $u, v, w \in V$ and $\alpha, \beta \in \mathbb{R}$:

(i) $\langle \alpha u + \beta v, w \rangle = \alpha \langle u, w \rangle + \beta \langle v, w \rangle$,

(ii) $\langle v, w \rangle = \langle w, v \rangle$,

(iii) $\langle v, v \rangle \geq 0$, and

(iv) $\langle v, v \rangle = 0$ if and only if $v = \vec{0}$, where $\vec{0}$ denotes the zero vector in V.

V together with an inner product $\langle \cdot, \cdot \rangle$ is called an **inner product space**.

We can (and sometimes do) define a complex-valued inner product (for example, see [**11**]), but in this text we will stick with the real-valued inner product. Notice that by using $\alpha = \beta = 0$ in condition (i), it follows that $\langle \vec{0}, v \rangle = 0$ for all $v \in V$.

Example 3.4.2. For $\vec{x}, \vec{y} \in \mathbb{R}^3$ with $\vec{x} = (x_1, x_2, x_3)$ and $\vec{y} = (y_1, y_2, y_3)$ we have the usual dot product,

$$\vec{x} \cdot \vec{y} = x_1 y_1 + x_2 y_2 + x_3 y_3 \,.$$

As expected, $\mathbb{R}^3$ with the dot product is an inner product space.

Of course, the above example extends to $\mathbb{R}^n$. We also have encountered another inner product space.

Example 3.4.3. For $f, g \in L^2[a, b]$ define $\langle f, g \rangle$ to be

$$\langle f, g \rangle = \int_a^b fg \,.$$

One of the first steps we need to do in order to show that we have created an inner product is to verify that we have defined a function from $L^2[a, b] \times L^2[a, b]$ to $\mathbb{R}$. In other words, if f and g are in $L^2[a, b]$, we need to verify that $\int_a^b fg$ is finite. Hölder's Inequality, Theorem 3.2.5, with $p = q = 2$ guarantees that the product fg is in $L^1[a, b]$, which is precisely what we need. (On the other hand, if $p \neq 2$, we cannot use Hölder's inequality to guarantee that $\int_a^b fg$ always produces a finite number.) The four properties of inner products are easily verified.

In $\mathbb{R}^n$ we have the relationship between the dot product and magnitude

$$|v| = \sqrt{v \cdot v}\,.$$

In an inner product space V we will define the *induced norm* $\|\cdot\|$ to be

$$\|v\| = \sqrt{\langle v, v\rangle}\,.$$

At this point, we are using the term "norm", but we have not yet justified that, what we call the "induced norm", satisfies Definition 3.1.1. But we will.

Example 3.4.4. In $L^2[a,b]$ the induced norm is

$$\begin{aligned} \|f\| &= \sqrt{\langle f, f\rangle} \\ &= \left(\int_a^b f^2\right)^{\frac{1}{2}} = \|f\|_2\,. \end{aligned}$$

In the case of $L^2[a,b]$, the induced norm is our familiar L^2-norm.

Now consider showing that, in general, what an induced norm does is a norm. The first three properties of a norm are easily checked. It is the fourth property that is less straightforward. A classic result known as the Cauchy-Schwarz Inequality is needed for this.

Proposition 3.4.5 (Cauchy-Schwarz Inequality)**.** *Let V be an inner product space with inner product $\langle\cdot,\cdot\rangle$. For every $v, w \in V$,*

$$|\langle v, w\rangle| \leq \|v\|\,\|w\|\,.$$

Proof. The result is easily true if either v or w is the zero vector. Hence, we will assume neither v nor w is the zero vector.

By property (iii) of a norm,

$$\langle tv - w, tv - w\rangle \geq 0$$

for every real number t. Therefore,

$$t^2\langle v, v\rangle - 2t\langle v, w\rangle + \langle w, w\rangle \geq 0\,.$$

Here we have used property (i) of a norm to expand $\langle tv - w, tv - w\rangle$. The expression on the left is quadratic in t. Since the quantity on the left must always be greater than or equal to zero, the quadratic in t must have at most one real root. Thinking of the quadratic formula, the quantity inside the square root cannot be positive. In other words,

$$4\left(\langle v, w\rangle\right)^2 - 4\langle v, v\rangle \langle w, w\rangle \leq 0$$

or

$$\left(\langle v, w\rangle\right)^2 \leq \langle v, v\rangle \langle w, w\rangle .$$

Thus

$$|\langle v, w\rangle| \leq \sqrt{\langle v, v\rangle}\sqrt{\langle w, w\rangle} = \|v\| \, \|w\|,$$

as claimed. □

We will now show that the induced norm on an inner product space does satisfy the triangle inequality. Notice that this is a statement which applies to any inner product space, not just $L^2[a,b]$.

Proposition 3.4.6. *Let V be an inner product space with inner product $\langle \cdot, \cdot\rangle$ and induced norm $\| \cdot \|$. Then for all $v, w \in V$,*

$$\|v + w\| \leq \|v\| + \|w\| .$$

Proof. For every $v, w \in V$, by the definition of the induced norm, properties of the inner product, and the Cauchy-Schwarz Inequality (Proposition 3.4.5),

$$\begin{aligned} \|v + w\|^2 &= \langle v + w, v + w\rangle \\ &= \langle v, v\rangle + 2\langle v, w\rangle + \langle w, w\rangle \\ &\leq \|v\|^2 + 2\|v\| \, \|w\| + \|w\|^2 . \end{aligned}$$

Therefore,

$$\|v + w\|^2 \leq \left(\|v\| + \|w\|\right)^2$$

or

$$\|v + w\| \leq \|v\| + \|w\|,$$

as claimed. □

An inner product space is not necessarily a Banach space. The missing ingredient is completeness. In other words, in an inner product space there is no guarantee that Cauchy sequences converge. When this extra quality occurs, we call the space a Hilbert space.

Definition 3.4.7. A **Hilbert space** is an inner product space that is a Banach space with respect to the induced norm.

In other words, in order for a vector space V to be a Hilbert space, it must have an inner product and it must be complete with respect to the induced norm.

Example 3.4.8. $L^2[a,b]$ is a Hilbert space. The norm induced by the inner product is the same as the L^2-norm, and we have shown that $L^2[a,b]$ is complete with respect to the L^2-norm.

Example 3.4.9. Let $C[a,b]$ denote the space of functions that are continuous on the interval $[a,b]$. We can define an inner product on this space in the same fashion as on $L^2[a,b]$, that is,

$$\langle f, g\rangle = \int_a^b fg\,.$$

The induced norm is identical to the L^2-norm. However, the space $C[a,b]$ is not the same as $L^2[a,b]$. The function

$$f(x) = \begin{cases} 0 & \text{if } x < \frac{1}{2}, \\ 1 & \text{if } x \geq \frac{1}{2} \end{cases}$$

is in $L^2[0,1]$ but not in $C[0,1]$. By Theorem 3.3.5 we can approximate f by continuous functions. That is, we can find a sequence of continuous functions $\{g_n\}$ that converge to f in the L^2-norm. But this means that $\{g_n\}$ must be a Cauchy sequence that does not converge to a continuous function with respect to this induced norm. Therefore, $C[0,1]$ with the norm induced by this inner product is not complete with respect to this norm, and hence is not a Hilbert space.

In general, given a vector space V there might be many choices for a norm on V. Are all of these norms induced by different inner products? We know that given an inner product on V, it induces a norm on V. What about the other way around? To help us make that determination, we have the next proposition, known as the parallelogram law.

Proposition 3.4.10. *Let V be an inner product space with induced norm $\|v\| = \sqrt{\langle v, v\rangle}$. Then for every $v, w \in V$,*

$$\|v+w\|^2 + \|v-w\|^2 = 2\|v\|^2 + 2\|w\|^2 .$$

Proof. This is Exercise 20. □

The next example illustrates how this proposition can be used to show that a norm is not induced by some inner product.

Example 3.4.11. Consider $L^1[0,1]$ with the L^1-norm. Both $f(x) = 1 - x$ and $g(x) = x$ are in $L^1[0,1]$. Thus

$$\|f+g\|_1 = \int_0^1 1 = 1\,,$$

$$\|f-g\|_1 = \int_0^1 |1-2x| = \frac{1}{2}\,,$$

$$\|f\|_1 = \int_0^1 |1-x| = \frac{1}{2}\,, \qquad \text{and}$$

$$\|g\|_1 = \int_0^1 |x| = \frac{1}{2}\,.$$

However,

$$\begin{aligned} \left(\|f+g\|_1\right)^2 + \left(\|f-g\|_1\right)^2 &= 1 + \frac{1}{4} \\ \neq 2\frac{1}{4} + 2\frac{1}{4} &= 2\left(\|f\|_1\right)^2 + 2\left(\|g\|_1\right)^2 . \end{aligned}$$

Therefore, by Proposition 3.4.10, the L^1-norm is not induced by an inner product.

3.5. L^2 Theory of Fourier Series

We conclude this chapter with a brief discussion of Fourier series. This is intended to provide an example where the features of the Hilbert space $L^2[-\pi, \pi]$ come into play. What follows is by no means a complete investigation of Fourier series. Much more information about Fourier series can be found in other texts such as Brown and

Churchill [**4**]. This section is not necessary for later material and may be omitted.

To get started, we need to define a Fourier series. Suppose f is a function defined on $[-\pi, \pi]$ and

$$f(x) = \frac{a_0}{2} + \sum_{k=1}^{\infty} \left(a_k \cos(kx) + b_k \sin(kx)\right) .$$

There are two big assumptions being made here. The first is that the infinite series converges. The second is that the function f is equal to such a series. For now we will ignore any possible issues and figure out what the coefficients $a_0, a_1, b_1, a_2, b_2, \ldots$ ought to be. Assuming we can interchange integration with summation (which is not always the case),

$$\begin{aligned} \int_{-\pi}^{\pi} f(x) &= \pi a_0 + \sum_{k=1}^{\infty} \left(a_k \int_{-\pi}^{\pi} \cos(kx)\, dx + b_k \int_{-\pi}^{\pi} \sin(kx)\, dx \right) \\ &= \pi a_0 . \end{aligned}$$

Hence, we expect

$$a_0 = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) .$$

Continuing, for a fixed positive integer n,

$$\begin{aligned} \int_{-\pi}^{\pi} f(x) \cos(nx) = \frac{a_0}{2} \int_{-\pi}^{\pi} \cos(nx) \\ + \sum_{k=1}^{\infty} \left(a_k \int_{-\pi}^{\pi} \cos(kx) \cos(nx)\, dx + b_k \int_{-\pi}^{\pi} \sin(kx) \cos(nx)\, dx \right) \\ = \pi a_n . \end{aligned}$$

Here we have used the results of Exercise 25, which is essentially a calculus exercise. It makes sense to expect

$$a_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \cos(nx) .$$

Similarly, continuing with our assumptions,

$$\int_{-\pi}^{\pi} f(x) \sin(nx) = \pi b_n$$

so that

$$b_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \sin(nx) .$$

Throughout this section we will assume $f \in L^2[-\pi, \pi]$. We know that $L^2[-\pi, \pi]$ is a Hilbert space. Using the inner product $\langle \cdot, \cdot \rangle$, our computations can be expressed as

$$a_0 = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) = \frac{\langle f, 1 \rangle}{\langle 1, 1 \rangle},$$

$$a_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \cos(nx) = \frac{\langle f, \cos(nx) \rangle}{\langle \cos(nx), \cos(nx) \rangle},$$

$$b_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \sin(nx) = \frac{\langle f, \sin(nx) \rangle}{\langle \sin(nx), \sin(nx) \rangle}.$$

In other words, underlying our computations is the natural inner product on $L^2[-\pi, \pi]$.

For $f \in L^2[-\pi, \pi]$, the $a_0, a_1, b_1, a_2, b_2, \ldots$ defined above are always finite and are known as the **Fourier coefficients** for f. We now define Fourier series for f as

$$f(x) \sim \frac{a_0}{2} + \sum_{k=1}^{\infty} \left(a_k \cos(kx) + b_k \sin(kx) \right),$$

where

$$a_0 = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x),$$

$$a_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \cos(nx),$$

$$b_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \sin(nx).$$

We use the symbol $\sim$ instead of an equal sign because we do not know if the series converges to f. For that matter, we do not know if the series converges.

Now let's move on to the issue of convergence. As usual, let $s_n(x)$ denote the nth partial sum of the Fourier series. That is,

$$s_n(x) = \frac{a_0}{2} + \sum_{k=1}^{n} \left(a_k \cos(kx) + b_k \sin(kx) \right).$$

Our goal is to show that for f in $L^2[-\pi, \pi]$,

$$\lim_{n \to \infty} \|f - s_n\|_2 = 0.$$

In other words, we will show that the Fourier series of an L^2-function converges to that function with respect to the L^2-norm.

As a first step, we can claim the following.

Proposition 3.5.1. *Let $f \in L^2[-\pi,\pi]$. For each positive integer n,*

$$\|f - s_n\|_2^2 = \|f\|_2^2 - \left(\frac{\pi a_0^2}{2} + \pi\sum_{k=1}^{n}(a_k^2 + b_k^2)\right),$$

where $a_0, a_1, b_1, a_2, b_2, \ldots$ are the Fourier coefficients defined above and $s_n(x)$ is the nth partial sum of the Fourier series for f.

Proof. This is Exercise 26. □

As a corollary, we can prove what is known as Bessel's inequality.

Corollary 3.5.2. *Let $f \in L^2[-\pi,\pi]$. Let $a_0, a_1, b_1, a_2, b_2, \ldots$ be the Fourier coefficients for f. Then $\sum_{k=1}^{\infty}(a_k^2 + b_k^2)$ converges and*

$$\frac{\pi a_0^2}{2} + \pi\sum_{k=1}^{\infty}(a_k^2 + b_k^2) \le \|f\|_2^2 .$$

Proof. For each n, by Proposition 3.5.1

$$\frac{\pi a_0^2}{2} + \pi\sum_{k=1}^{n}(a_k^2 + b_k^2) \le \|f\|_2^2 .$$

The result follows by taking n to infinity. □

For a positive integer n, a **trigonometric polynomial** of degree n is a function T_n of the form

$$T_n(x) = A_0 + \sum_{k=1}^{n}\left(A_k\cos(kx) + B_k\sin(kx)\right),$$

where $A_0, A_1, B_1, \ldots, B_n$ are real numbers. Given $f \in L^2[-\pi,\pi]$, the nth partial sum of the Fourier series for f is an example of a trigonometric polynomial of degree n. This next theorem asserts that s_n is the trigonometric polynomial of degree n that is the closest to f in the L^2-norm.

Theorem 3.5.3. *Let $f \in L^2[-\pi,\pi]$ and let T_n be a trigonometric polynomial of degree n. Then*

$$\|f - T_n\|_2 \geq \|f - s_n\|_2,$$

where $a_0, a_1, b_1, \ldots, b_n$ are the Fourier coefficients of f and

$$s_n(x) = \frac{a_0}{2} + \sum_{k=1}^{n} \left(a_k \cos(kx) + b_k \sin(kx)\right) .$$

Proof. Let T_n be a trigonometric polynomial of degree n, say

$$T_n(x) = A_0 + \sum_{k=1}^{n} \left(A_k \cos(kx) + B_k \sin(kx)\right) .$$

We will show that $\|f - T_n\|_2^2 - \|f - s_n\|_2^2 \geq 0$:

$$\begin{aligned} \|f - T_n\|_2^2 &= \int_{-\pi}^{\pi} \left(f(x) - T_n(x)\right)^2 \\ &= \int_{-\pi}^{\pi} f^2 - 2\int_{-\pi}^{\pi} fT_n + \int_{-\pi}^{\pi} (T_n)^2 , \end{aligned}$$

$$\begin{aligned} \int_{-\pi}^{\pi} fT_n &= \int_{-\pi}^{\pi} f \left(A_0 + \sum_{k=1}^{n} \left(A_k \cos(kx) + B_k \sin(kx)\right) \right) \\ &= A_0 \int_{-\pi}^{\pi} f + \sum_{k=1}^{n} \left(A_k \int_{-\pi}^{\pi} f \cos(kx) + B_k \int_{-\pi}^{\pi} f \sin(kx) \right) \\ &= 2\pi A_0 a_0 + \pi \sum_{k=1}^{n} \left(A_k a_k + B_k b_k\right) . \end{aligned}$$

Note that we can switch the integration with the summation because we have a finite sum. Next, using the results of Exercise 25,

$$\int_{-\pi}^{\pi} (T_n)^2 = 2\pi A_0^2 + \pi \sum_{k=1}^{n} \left(A_k^2 + B_k^2\right) .$$

Collecting these results and Proposition 3.5.1, we have

$$\begin{aligned}
&\|f - T_n\|_2^2 - \|f - s_n\|_2^2 \\
&\qquad = -2\left(2\pi A_0 a_0 + \pi \sum_{k=1}^{n} (A_k a_k + B_k b_k)\right) \\
&\qquad\quad + 2\pi A_0^2 + \pi \sum_{k=1}^{n} \left(A_k^2 + B_k^2\right) + \frac{\pi a_0^2}{2} + \pi \sum_{k=1}^{n} (a_k^2 + b_k^2) \\
&\qquad = 2\pi(A_0^2 - 2A_0 a_0 + a_0^2) + \pi \sum_{k=1}^{n} \left(A_k^2 - 2A_k a_k + a_k^2\right) \\
&\qquad\quad + \pi \sum_{k=1}^{n} \left(A_k^2 - 2A_k a_k + a_k^2\right) \\
&\qquad = 2\pi(A_0 - a_0)^2 + \pi \sum_{k=1}^{n} (A_k - a_k)^2 + \pi \sum_{k=1}^{n} (B_k - b_k)^2 \\
&\qquad \geq 0\,.
\end{aligned}$$

□

We will return to the issue of convergence, but in a special case. Before doing so, we need to introduce the Dirichlet kernel.

Definition 3.5.4. For positive integer n, the nth **Dirichlet kernel** is

$$D_n(t) = \frac{1}{2} + \sum_{k=1}^{n} \cos(kt)\,.$$

The convention is to also define $D_0(t) = \frac{1}{2}$.

By definition it is easy to see that every $D_n(t)$ is an even function, periodic with period 2π, and

$$\int_0^\pi D_n(t)\,dt = \frac{\pi}{2}\,.$$

Another kernel we will need is the Fejér kernel.

Definition 3.5.5. For positive integer n, the nth **Fejér kernel** is

$$K_n(t) = \frac{1}{n} \sum_{k=0}^{n-1} D_n(t)\,.$$

Again, $K_n(t)$ is periodic with period 2π and

$$\int_0^\pi K_n(t)\,dt = \frac{\pi}{2}\,.$$

Here, now, is our first result concerning convergence.

Theorem 3.5.6. *Let f be continuous on $[-\pi,\pi]$ with $f(-\pi)=f(\pi)$. For each positive integer n, define $\sigma_n(x)$ to be*

$$\sigma_n(x) = \frac{1}{n}\sum_{k=0}^{n-1} s_n(x),$$

where $s_n(x)$ denotes the nth partial sum of the Fourier series for f. Then the sequence of functions $\sigma_n(x)$ converges uniformly to $f(x)$ on $[-\pi,\pi]$.

Proof. Since f is continuous on $[-\pi,\pi]$, f is uniformly continuous on that interval. By setting $f(x+2\pi)=f(x)$, we may extend f to a function that is defined for all real numbers and is periodic with period 2π. Thus, without loss of generality, we will assume f is uniformly continuous on $(-\infty,\infty)$ and periodic with period 2π.

By Exercise 29 followed by a change of variables,

$$\begin{aligned} s_n(x) &= \frac{1}{\pi}\int_{-\pi}^{\pi} f(t)\,D_n(x-t)\,dt \\ &= \frac{1}{\pi}\int_{x-\pi}^{x+\pi} f(x-u)\,D_n(u)\,du\,. \end{aligned}$$

The interval of integration $[x-\pi,x+\pi]$ has length 2π. The integrand is periodic with period 2π. We can integrate over any interval of length 2π and obtain the same result. Thus,

$$\begin{aligned}
s_n(x) &= \frac{1}{\pi}\int_{-\pi}^{\pi} f(x-u)\,D_n(u)\,du \\
&= \frac{1}{\pi}\int_{-\pi}^{0} f(x-u)\,D_n(u)\,du + \frac{1}{\pi}\int_{0}^{\pi} f(x-u)\,D_n(u)\,du \\
&= \frac{1}{\pi}\int_{0}^{\pi} f(x+t)\,D_n(-t)\,du + \frac{1}{\pi}\int_{0}^{\pi} f(x-t)\,D_n(t)\,dt \\
&\quad \text{(using another change of variables)} \\
&= \frac{1}{\pi}\int_{0}^{\pi} f(x+t)\,D_n(t)\,du + \frac{1}{\pi}\int_{0}^{\pi} f(x-t)\,D_n(t)\,dt \\
&\quad (D_n(t) \text{ is an even function}) \\
&= \frac{1}{\pi}\int_{0}^{\pi} \left(f(x+t)+f(x-t)\right)\,D_n(t)\,dt \\
&= \frac{2}{\pi}\int_{0}^{\pi} \frac{(f(x+t)+f(x-t))}{2}\,D_n(t)\,dt\,.
\end{aligned}$$

Consequently,

$$\sigma_n(x) = \frac{2}{\pi}\int_{0}^{\pi} \frac{(f(x+t)+f(x-t))}{2}\,K_n(t)\,dt\,.$$

But

$$\int_0^{\pi} K_n(t)\,dt = \frac{\pi}{2}\,,$$

and hence

$$\sigma_n(x) - f(x) = \frac{2}{\pi}\int_{0}^{\pi} \left(\frac{(f(x+t)+f(x-t))}{2} - f(x)\right)\,K_n(t)\,dt\,.$$

Our goal is to show that the sequence of functions $\sigma_n(x)$ converges uniformly to $f(x)$ on $[-\pi,\pi]$. Let $\epsilon > 0$ be given. We will now take advantage of the uniform continuity of f. There exists a $\delta > 0$ so that if $|u-v| < \delta$, then $|f(u)-f(v)| < \frac{\epsilon}{2}$. Hence,

$$\begin{aligned}
|\sigma_n(x) - f(x)| \;&\le\; \frac{2}{\pi}\int_0^{\delta} \left|\frac{(f(x+t)+f(x-t))}{2} - f(x)\right| K_n(t)\,dt \\
&\quad + \frac{2}{\pi}\int_{\delta}^{\pi} \left|\frac{(f(x+t)+f(x-t))}{2} - f(x)\right| K_n(t)\,dt\,.
\end{aligned}$$

In the first of these two integrals

$$\left| \frac{(f(x+t)+f(x-t))}{2} - f(x) \right| \leq \frac{\epsilon}{2}$$

by uniform continuity and the choice of δ. Thus,

$$\begin{aligned} \frac{2}{\pi}\int_0^\delta \left| \frac{(f(x+t)+f(x-t))}{2} - f(x) \right| K_n(t)\,dt \\ < \frac{2}{\pi}\int_0^\delta \frac{\epsilon}{2} K_n(t)\,dt \\ \leq \frac{2}{\pi}\int_0^\pi \frac{\epsilon}{2} K_n(t)\,dt \quad \leq \quad \frac{\epsilon}{2}\,. \end{aligned}$$

We now need to deal with the second of the two integrals from above. Our choice of δ was independent of both x and of n. The function f is periodic and continuous on $[-\pi, \pi]$ and so is bounded, say by M. Using this and Exercise 28,

$$\begin{aligned} &\frac{2}{\pi}\int_\delta^\pi \left| \frac{(f(x+t)+f(x-t))}{2} - f(x) \right| K_n(t)\,dt \\ &\quad \leq \frac{2}{\pi}\int_\delta^\pi 2M\,K_n(t)\,dt \leq \frac{2}{\pi}\int_\delta^\pi 2M\,\frac{\sin^2(nt/2)}{2n\sin^2(t/2)}\,dt \\ &\quad \leq \frac{1}{n}\left(\frac{2M}{\pi}\int_\delta^\pi \frac{1}{\sin^2(t/2)}\,dt\right). \end{aligned}$$

The last integral is finite because we are integrating on the interval $[\delta, \pi]$. To complete the proof, choose N large enough so that if $n \geq N$, this last quantity is smaller than $\frac{\epsilon}{2}$. □

As a corollary we have the following.

Corollary 3.5.7. *Let g be continuous on $[-\pi, \pi]$ with $g(-\pi) = g(\pi)$. For each positive integer n, define $\sigma_n(x)$ to be*

$$\sigma_n(x) = \frac{1}{n}\sum_{k=0}^{n-1} s_n(x),$$

where $s_n(x)$ denotes the nth partial sum of the Fourier series for g. Then

$$\lim_{n\to\infty} \|\sigma_n - g\|_2 = 0\,.$$

Proof. Uniform convergence allows us to conclude that

$$\lim_{n\to\infty} \|\sigma_n - g\|_2^2 = \int_{-\pi}^{\pi} \lim_{n\to\infty} \left(|\sigma_n - g|^2\right) = 0\,. \qquad \square$$

We finally turn to the main result of this section.

Theorem 3.5.8. *Let $f \in L^2[-\pi,\pi]$. Let $s_n(x)$ equal the nth partial sum of the Fourier series for f. Then the sequence s_n converges to f with respect to the L^2-norm. That is,*

$$\lim_{n\to\infty} \|s_n - f\|_2 = 0\,.$$

Proof. Given $\epsilon > 0$, by Exercise 24 there is a continuous function g defined on $[-\pi,\pi]$ with $g(-\pi) = g(\pi)$ and

$$\|f - g\|_2 < \frac{\epsilon}{2}\,.$$

By Corollary 3.5.7, there is a positive integer N so that if $n \geq N$,

$$\|\sigma_n - g\|_2 \leq \frac{\epsilon}{2},$$

where the σ_n's correspond to the continuous function g. Let $n \geq N$. Note that σ_{n+1} is a trigonometric polynomial of degree n. By Theorem 3.5.3,

$$\|f - s_n\|_2 \leq \|f - \sigma_{n+1}\|_2 \leq \|f - g\|_2 + \|g - \sigma_{n+1}\|_2 < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon\,.$$

$\square$

In conclusion, we can establish Parseval's equation.

Corollary 3.5.9. *Let $f \in L^2[-\pi,\pi]$. Let $a_0, a_1, b_1, a_2, b_2, \ldots$ be the Fourier coefficients for f. Then*

$$\frac{\pi a_0^2}{2} + \pi \sum_{k=1}^{\infty} (a_k^2 + b_k^2) = \|f\|_2^2\,.$$

Proof. For each n, by Proposition 3.5.1,

$$\|f - s_n\|_2^2 = \|f\|_2^2 - \left(\frac{\pi a_0^2}{2} + \pi \sum_{k=1}^{n} (a_k^2 + b_k^2)\right).$$

The result follows by taking n to infinity and applying Theorem 3.5.8.

$\square$

It is important to note that we have only shown that the Fourier series of an L^2-function converges with respect to the L^2-norm. We did not make any claims about pointwise convergence of the Fourier series. There are results concerning this. See, for example Brown and Churchill [**4**].

3.6. Exercises

(1) Let $C[a,b]$ be the set of functions that are continuous on $[a,b]$. Prove or disprove:

$$\|f\|_{\sup} = \sup\{|f(x)| \mid x \in [a,b]\}$$

is a norm on $C[a,b]$.

(2) Show that $C[0,1] \subseteq L^1[0,1]$. Compare $\|f\|_{\sup}$ and $\|f\|_1$. Justify your answer. Generalize this to the interval $[a,b]$.

(3) Prove Proposition 3.1.2.

(4) Determine whether or not each of the following is a Cauchy sequence in $L^1[0,1]$:

a) $f_n(x) = n\mathcal{X}_{(\frac{1}{n+1},\frac{1}{n}]}(x)$.

b) $f_n(x) = \dfrac{1}{x}\mathcal{X}_{[\frac{1}{n+1},1]}(x)$.

c) $f_n(x) = \dfrac{1}{\sqrt{x}}\mathcal{X}_{[\frac{1}{n+1},1]}(x)$.

(5) Let $p > 1$. Show that $C[0,1] \subseteq L^p[0,1]$. Compare $\|f\|_{\sup}$ and $\|f\|_p$. Justify your answer. Generalize this to the interval $[a,b]$.

(6) Show that $f(x) = x^{-\frac{1}{3}} \in L^2[0,1]$ but $g(x) = x^{-\frac{2}{3}} \notin L^2[0,1]$.

(7) Let $f(x) = x^\alpha$ on $[0,1]$. Show that $f \in L^p[0,1]$ if and only if $\alpha > -\frac{1}{p}$.

(8) Let $p > 1$. Show that $L^p[a,b] \subseteq L^1[a,b]$. Show that $L^1[a,b] \not\subseteq L^p[a,b]$.

(9) Let $f, g : [a,b] \to \mathbb{R}$ be measurable functions. Show that if $f \le g$ a.e., then

$$\operatorname*{ess\,sup}_{x\in[a,b]} f \le \operatorname*{ess\,sup}_{x\in[a,b]} g\,.$$

(10) Show $f \in L^\infty[a,b]$ if and only if $|f| \in L^\infty[a,b]$.

(11) Let $p \geq 1$. Show that $L^\infty[a,b] \subseteq L^p[a,b]$.

(12) Let $f, g \in L^\infty[a,b]$ and $c \in \mathbb{R}$.
  a) Show that $cf \in L^\infty[a,b]$.
  b) Show that $f + g \in L^\infty[a,b]$.

(13) Show that $\|\cdot\|_\infty$ is a norm on the space $L^\infty[a,b]$.

(14) Let $f, g \in L^2[a,b]$. Prove

$$\|f+g\|_2^2 + \|f-g\|_2^2 = 2\|f\|_2^2 + 2\|g\|_2^2 \,.$$

(15) Determine whether or not each of the following is a Cauchy sequence in $L^2[0,\infty)$:
  a) $f_n(x) = \mathcal{X}_{[n,n+1]}(x)$.
  b) $f_n(x) = \dfrac{1}{x}\mathcal{X}_{[1,n]}(x)$.
  c) $f_n(x) = \dfrac{1}{x^2}\mathcal{X}_{[1,n]}(x)$.

(16) Let $1 \leq p < \infty$ and let $\{f_n\}$ be a sequence of functions in $L^p[a,b]$. Suppose there is a function $f \in L^p[a,b]$ with

$$\lim_{n\to\infty} \|f_n - f\|_p = 0 \,.$$

Prove that $\{f_n\}$ is a Cauchy sequence in $L^p[a,b]$.

(17) Prove that $L^\infty[a,b]$ is complete with respect to the norm $\|\cdot\|_\infty$.

(18) Let $f \in L^1[a,b]$ and $\alpha > 0$. Prove that

$$m\left(\{x \in [a,b] \mid |f(x)| > \alpha\}\right) \leq \frac{1}{\alpha}\|f\|_1 \,.$$

This is known as *Tchebychev's Inequality.*

Hint: Let $A = \{x \in [a,b] \mid |f(x)| > \alpha\}$ and compare the quantities

$$\int_A 1, \qquad \int_a^b \mathcal{X}_A, \qquad \text{and} \qquad \int_a^b \frac{|f|}{\alpha} \,.$$

(19) Here is another form of Tchebychev's Inequality. Let $1 < p < \infty$, $f \in L^p[a,b]$ and $\alpha > 0$. Prove that

$$m\left(\{x \in [a,b] \mid |f(x)| > \alpha\}\right) \leq \frac{1}{\alpha^p}\|f\|_p^p \,.$$

(20) Prove Proposition 3.4.10.

(21) Let $f \in L^2[a,b]$. If

$$\int_a^b fg = 0$$

for every $g \in C[a,b]$, show $f = 0$ a.e. in $[a,b]$.

(22) We denote by ℓ^∞ the space of all bounded sequences $(a_n)_{n=1}^\infty$. For example,

$$(1, -2, 1, -2, 1, -2, \ldots) \in \ell^\infty .$$

Define addition and scalar multiplication by

$$(a_n)_{n=1}^\infty + (b_n)_{n=1}^\infty = (a_n + b_n)_{n=1}^\infty \text{ and}$$

$$c(a_n)_{n=1}^\infty = (ca_n)_{n=1}^\infty .$$

a) Let $\|(a_n)_{n=1}^\infty\| = \sup_n |a_n|$. Show that $\|\cdot\|$ is a norm on ℓ^∞.

b) Show that ℓ^∞ is complete with respect to this norm. In other words, prove ℓ^∞ is a Banach space.

(23) Prove Corollary 3.3.4.

(24) Let $f \in L^p[a,b]$ for $p \geq 1$. Given two real numbers, A and B, and $\epsilon > 0$, show that there is a continuous function g defined on $[a,b]$ with $g(a) = A$, $g(b) = B$, and

$$\|f - g\|_p < \epsilon .$$

(25) Let k and n be positive integers. Prove each of the following:

a) $\displaystyle\int_{-\pi}^{\pi} \sin^2(nx)\,dx = \int_{-\pi}^{\pi} \cos^2(nx)\,dx = \pi$.

b) $\displaystyle\int_{-\pi}^{\pi} \sin(kx)\cos(nx)\,dx = 0$.

c) $\displaystyle\int_{-\pi}^{\pi} \sin(kx)\sin(nx)\,dx = \int_{-\pi}^{\pi} \cos(mx)\cos(nx)\,dx = 0$ if $m \neq n$.

(26) Prove Proposition 3.5.1.

(27) Let $D_n(t)$ denote the nth Dirichlet kernel. Show that if t is not a multiple of 2π, then

$$D_n(t) = \frac{\sin(n + \frac{1}{2})t}{2\sin(t/2)} .$$

(28) Let $K_n(t)$ denote the nth Fejér kernel. Use the previous exercise to show that if t is not a multiple of 2π, then

$$K_n(t) = \frac{\sin^2(nt/2)}{2n\sin^2(t/2)} .$$

(29) Let $f \in L^2[-\pi, \pi]$. Let $s_n(x)$ denote the nth partial sum of the Fourier series for f. Show that

$$s_n(x) = \frac{1}{\pi}\int_{-\pi}^{\pi} f(t)\, D_n(x-t)\, dt,$$

where $D_n(t)$ denotes the nth Dirichlet kernel.

(30) Let $f(x) = x$ for $x \in [-\pi, \pi]$. Find the Fourier series for f. Use Parseval's equation, Corollary 3.5.9, to evaluate $\displaystyle\sum_{k=1}^{\infty} \frac{1}{k^2}$.

http://dx.doi.org/10.1090/stml/078/05

Chapter 4

General Measure Theory

Lebesgue measure on $\mathbb{R}^n$ is an extension of our notions about length, area, or volume, depending on the dimension n. Now we are about to generalize the notion of measure. We should think about what properties we want in a measure. For each set in a given collection of sets we want a measure to identify with that set a nonnegative real number or plus infinity, the idea being that a measure should indicate something about the size of that set. For example, the measure of the empty set should be 0. Also, if $A \subseteq B$ it should be the case that the measure of A is smaller than or equal to the measure of B. Finally, we want the measure of a countable union of pairwise disjoint sets to be the sum of the measures of the set. All of these will be addressed when we define a measure.

4.1. Measure Spaces

When we look to generalize measure, we might want to measure very general sets, sets other than subsets of $\mathbb{R}^n$ (although most of our examples will be subsets of $\mathbb{R}^n$). An important consideration is what properties such a collection of sets needs to have for our desired qualities of a measure to make sense. For example, one property we know about Lebesgue measure is that the intersection of two measurable sets is also measurable. This is certainly a minimal requirement we would like to retain. Think about this and the other properties of the collection of Lebesgue measurable subsets when considering the following definitions.

Definition 4.1.1. Let X be a nonempty set. A collection of subsets of X, $\mathcal{A}$, is called an **algebra of sets on** X if

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

(ii) if $A, B \in \mathcal{A}$, then $A \cup B \in \mathcal{A}$, and

(iii) if $A \in \mathcal{A}$, then $A^c = X \setminus A \in \mathcal{A}$.

Property (ii) is often stated as "$\mathcal{A}$ is closed under union". Similarly, property (iii) is often stated as "$\mathcal{A}$ is closed under set complement". Keep in mind that, in this context, the term "closed" has nothing to do with a set being open or closed as in an open or closed interval.

Example 4.1.2. Let $X = \{a, b, c, d, e\}$.

(i) Let $\mathcal{A}_1 = \{\emptyset, X, \{a, c\}, \{b, d\}, \{a, b, c, d\}, \{b, d, e\}, \{a, c, e\}, \{e\}\}$.

(ii) Let $\mathcal{A}_2 = \{\emptyset, X, \{a, c\}, \{b, d, e\}\}$.

(iii) Let $\mathcal{A}_3 = \{\emptyset, X\}$.

(iv) Let $\mathcal{A}_4 = \mathcal{P}(X) = \{C \bigm| C \subseteq X\}$.

$\mathcal{A}_1$, $\mathcal{A}_2$, $\mathcal{A}_3$, and $\mathcal{A}_4$ are each an example of an algebra of sets on X.

We stated that one of the desirable features of a measure is to have the intersection of two measurable sets be a measurable set. Although this is not explicitly stated, it does follow from the definition as demonstrated by the next proposition.

Proposition 4.1.3. *Let $\mathcal{A}$ be an algebra of sets on X. If $A, B \in \mathcal{A}$, then $A \cap B \in \mathcal{A}$.*

Proof. By de Morgan's laws,

$$A \cap B = (A^c \cup B^c)^c .$$

Therefore, this proposition follows from properties (ii) and (iii) of an algebra of sets. □

Thus, an algebra of sets is also closed under intersection. But this is not quite enough for our goal with general measures. After all, we also know that in the case of Lebesgue measure the countable union of measurable sets is again measurable. By induction an algebra of sets is closed under finite union. But this does not stretch into closure under countable unions. We actually need something a little stronger than just an algebra of sets.

Definition 4.1.4. Let X be a nonempty set. A collection $\mathcal{B}$ of subsets of X is called a **σ-algebra of sets on X** if

(i) $\emptyset \in \mathcal{B}$,

(ii) if $\{A_n\}$ is a countable collection of sets in $\mathcal{B}$, then $\bigcup A_n \in \mathcal{B}$, and

(iii) if $A \in \mathcal{B}$, then $A^c \in \mathcal{B}$.

As with the definition of an algebra of sets, properties (ii) and (iii) are summarized by saying that $\mathcal{B}$ is closed under countable unions and set complement.

Example 4.1.5. Let X be a set. Let $\mathcal{P}(X)$ denote the power set of X, that is, $\mathcal{P}(X) = \{C \mid C \subseteq X\}$.

The preceding example is not really that interesting because we are simply collecting all possible subsets of X. Similarly, the σ-algebra $\{\emptyset, X\}$ is also somewhat trivial.

Example 4.1.6. Let $\mathcal{M}$ be the collection of Lebesgue measurable subsets of $\mathbb{R}^n$. By Example 1.2.4, $\emptyset \in \mathcal{M}$. By Theorem 1.2.5, $\mathcal{M}$ is closed under countable unions. Finally, by Theorem 1.2.18, $\mathcal{M}$ is closed under set complement. Therefore, $\mathcal{M}$ is a σ-algebra on $\mathbb{R}^n$.

Clearly, every σ-algebra on X is an algebra of sets. Each of the algebras described in Example 4.1.2 is a σ-algebra. The change in requirement (ii) might appear to be a small one. Is there a difference? The reason the algebras in Example 4.1.2 turned out to be σ-algebras is that we were dealing with a finite set. Consequently, every countable union is in fact a finite union. This is not always the case when dealing with infinite sets. In other words, not every algebra of sets is a σ-algebra.

Example 4.1.7. Let $\mathcal{A} = \{A \subseteq \mathbb{R} \mid A \text{ is finite or } A^c \text{ is finite}\}$. By Exercise 1, $\mathcal{A}$ is an algebra. However, $\mathcal{A}$ is not a σ-algebra. For example, since $\mathbb{Q}$ is countable, it is the countable union of finite sets. However, $\mathbb{Q} \notin \mathcal{A}$. This provides us with an example of an algebra that is not a σ-algebra.

The added requirement that a σ-algebra is closed under countable union carries over to countable intersections.

Proposition 4.1.8. *Let $\mathcal{B}$ be a σ-algebra of sets. If $\{A_n\}$ is a countable collection of sets in $\mathcal{B}$, then $\bigcap A_n \in \mathcal{B}$.*

Proof. This is Exercise 2. □

Given a set X, there may be many different possible σ-algebras on X. In Example 4.1.2 we saw four different σ-algebras on the set $X = \{a, b, c, d, e\}$. Intuitively, it should be clear what is meant when we say that $\mathcal{A}_3$ is smaller than $\mathcal{A}_4$. We will make the following more precise definition of what is meant by comparing σ-algebras.

Definition 4.1.9. Let $\mathcal{B}_1$ and $\mathcal{B}_2$ be two σ-algebras on a set X. We say $\mathcal{B}_1$ **is contained in** $\mathcal{B}_2$ if A is in $\mathcal{B}_2$ whenever A is in $\mathcal{B}_1$.

In Example 4.1.2, $\mathcal{A}_2$ is contained in $\mathcal{A}_1$. Also, $\mathcal{A}_3$ is contained in every σ-algebra of sets on X, whereas $\mathcal{A}_4$ contains every σ-algebra of sets on X.

Sometimes we have a collection $\mathcal{B}$ of subsets of a set X which is not necessarily a σ-algebra on X. Then we might be in the position of seeking a σ-algebra on X that includes everything in $\mathcal{B}$ in some sort of minimal fashion.

Definition 4.1.10. Let $\mathcal{B}$ be a collection of subsets of X. $\mathcal{A}$ **is the σ-algebra generated by** $\mathcal{B}$ if $\mathcal{A}$ is a σ-algebra that contains $\mathcal{B}$ and $\mathcal{A}$ is contained in any σ-algebra that contains $\mathcal{B}$.

In other words, $\mathcal{A}$ is the σ-algebra generated by $\mathcal{B}$ if $\mathcal{A}$ is the smallest σ-algebra that contains $\mathcal{B}$. To see that such an algebra always exists, first note that given a collection of subsets of X, it is always true that $\mathcal{P}(X)$ is a σ-algebra that contains $\mathcal{B}$. In Exercise 4, it is established that the intersection of two sigma algebras is a sigma algebra. More generally it is the case that the intersection of any collection of sigma algebras will be a sigma algebra. Therefore, given a collection $\mathcal{B}$ of subsets of X, one finds the sigma algebra generated by $\mathcal{B}$ by taking the intersection of all sigma algebras that contain $\mathcal{B}$. Hence, there will always exist a σ-algebra generated by $\mathcal{B}$.

Example 4.1.11. Let $\mathcal{B} = \{C \subseteq \mathbb{R} \mid C \text{ is finite}\}$. Then $\mathcal{A}$, the σ-algebra generated by $\mathcal{B}$, must contain all finite sets and their complements. Additionally, $\mathcal{A}$ must contain countable unions of finite sets. In other words, $\mathcal{A}$ must contain countable sets as well as their complements. It remains to show that $\{C \subseteq \mathbb{R} \mid C \text{ is countable, or } C^c \text{ is countable}\}$ is a σ-algebra. This is left as an exercise (see Exercise 3). Therefore, the σ-algebra generated by $\mathcal{B}$ is

$$\{C \subseteq \mathbb{R} \mid C \text{ is countable, or } C^c \text{ is countable}\}.$$

Suppose we start with the collection of open sets in $\mathbb{R}^n$ and use these to generate a σ-algebra. This collection of sets is already closed under countable unions. But a σ-algebra is also closed under countable intersections. We have encountered this before. Recall the defintion of type G_δ from Definition 1.2.20: a set H is of type $G_\delta$if H is the intersection of a countable collection of open sets. In other words, if we use the collection of open sets to generate a σ-algebra, the resulting σ-algebra must contain all sets of type G_δ. In addition, the σ-algebra must be closed under set complement and so must contain all closed sets. As a result, the σ-algebra generated by the collection of open sets also contains all sets that are of type F_σ. (Again from Definition 1.2.20 H is of type F_σ if H is the union of a countable collection of closed sets.) This σ-algebra is an important one known as the collection of Borel sets.

Definition 4.1.12. The collection of **Borel sets** is the σ-algebra $\mathbb{B}$ generated by open subsets of $\mathbb{R}^n$.

Before arriving at our destination of general measures, we will define what is meant by a measurable space.

Definition 4.1.13. A **measurable space** is a pair $(X, \mathcal{B})$ consisting of a set X with a σ-algebra $\mathcal{B}$. A subset A of X is called *measurable* if $A \in \mathcal{B}$.

Remark 4.1.14. Two examples of measurable spaces that we have already encountered are $(\mathbb{R}^n, \mathbb{B})$ and $(\mathbb{R}^n, \mathcal{M})$. Are these two examples really different? It turns out that although most subsets of $\mathbb{R}^n$ that come to mind are Borel sets, using the Axiom of Choice we can show there are Lebesgue measurable sets which are not Borel sets. To do this, recall the Cantor set $\mathcal{C}$ discussed in Example 1.1.7 and the one-to-one function $f\colon [0,1] \to \mathcal{C}$ used to show that the Cantor set is uncountable. Now, by construction $\mathcal{C}$ is a Borel set with Lebesgue measure 0. Hence, every subset of $\mathcal{C}$ is a Lebesgue measurable set with Lebesgue measure 0. Let A be a nonmeasurable subset of $[0,1]$ (such a set exists by Exercise 25 of Chapter 1). Then

$$f(A) = \{f(x) \mid x \in A\}$$

is a subset of $\mathcal{C}$; therefore $f(A)$ must have Lebesgue outer measure 0, and hence is Lebesgue measurable. On the other hand, if $f(A)$ were a Borel set, by Exercise 5, A would be a measurable set, a contradiction. Therefore, $f(A)$ cannot be a Borel set.

Now that we have in mind what is needed to be a measurable space, we can finally define a measure.

Definition 4.1.15. A **measure** on a measurable space $(X, \mathcal{B})$ is a function $\mu\colon \mathcal{B} \to [0, +\infty]$ such that

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

(ii) for any countable collection $\{E_j\}$ of pairwise disjoint sets in $\mathcal{B}$,

$$\mu\left(\bigcup_j E_j\right) = \sum_j \mu(E_j)\ .$$

The triple $(X, \mathcal{B}, \mu)$ is called a *measure space.*

Two examples of measure spaces are $(\mathbb{R}^n, \mathcal{M}, m)$ and $(\mathbb{R}^n, \mathbb{B}, m)$. There are many others.

Example 4.1.16. Define μ on $\mathcal{P}(\mathbb{R})$, the set of subsets of $\mathbb{R}$, by

$$\mu(A) = \begin{cases} 1 & \text{if } \pi \in A, \\ 0 & \text{otherwise.} \end{cases}$$

We will show that $(\mathbb{R}, \mathcal{P}(\mathbb{R}), \mu)$ is a measure space. First of all, $\mathcal{P}(\mathbb{R})$ is a σ-algebra on $\mathbb{R}$ and hence $(\mathbb{R}, \mathcal{P}(\mathbb{R}))$ is a measurable space. Furthermore, by definition $\mu(\emptyset) = 0$. Finally, if $\{E_j\}$ is a countable collection of pairwise disjoint sets in $\mathcal{P}(\mathbb{R})$, then either $\pi \in E_i$ for exactly one i or $\pi \notin E_j$ for all j. In the first case

$$\mu\left(\bigcup_j E_j\right) = 1 = \sum_j \mu\left(E_j\right),$$

and in the latter case

$$\mu\left(\bigcup_j E_j\right) = 0 = \sum_j \mu\left(E_j\right).$$

Thus, $(\mathbb{R}, \mathcal{P}(\mathbb{R}), \mu)$ is a measure space.

In the beginning of this chapter, we stated that one desirable property of a measure was that if $A \subseteq B$, then the measure of B should be greater than or equal to the measure of A. We will show that this follows from our definition of measure.

Proposition 4.1.17. *Let $(X, \mathcal{B}, \mu)$ be a measure space. If $A, B \in \mathcal{B}$ and $A \subseteq B$, then*

$$\mu(A) \leq \mu(B).$$

In addition, if $\mu(A)$ is finite, then

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

Proof. Since $\mathcal{B}$ is a σ-algebra, $B \setminus A = B \cap A^c \in \mathcal{B}$. By definition of a measure, $\mu(B \setminus A) \geq 0$. Also, A and $B \setminus A$ are disjoint. Therefore, by property (ii) of a measure

$$\begin{aligned} \mu(A) &\leq \mu(A) + \mu(B \setminus A) \\ &= \mu\left(A \cup (B \setminus A)\right) = \mu(B). \end{aligned}$$

This proves the first part of the proposition. If $\mu(A)$ is finite, we may subtract this quantity from both sides of the last equality to obtain

$$\mu(B \setminus A) = \mu(B) - \mu(A),$$

as claimed in the second part of the proposition. □

Property (ii) of the definition of measure deals with the measure of a countable union of pairwise disjoint measurable sets. The next theorem concerns the countable union of sets which are not necessarily pairwise disjoint.

Theorem 4.1.18. *Let $(X, \mathcal{B}, \mu)$ be a measure space. If $\{E_j\}$ is a countable collection of sets in $\mathcal{B}$, then*

$$\mu\left(\bigcup_j E_j\right) \le \sum_j \mu\left(E_j\right).$$

Proof. Set $G_1 = E_1$. For each $j > 1$ let

$$G_j = E_j \setminus \bigcup_{i=1}^{j-1} E_i .$$

Then for each j, $G_j \in \mathcal{B}$ and $G_j \subseteq E_j$, and hence $\mu(G_j) \le \mu(E_j)$ by Proposition 4.1.17. But $\{G_j\}$ is a countable collection of pairwise disjoint sets in $\mathcal{B}$; therefore,

$$\begin{aligned} \mu\left(\bigcup_j E_j\right) &= \mu\left(\bigcup_j G_j\right) \\ &= \sum_j \mu\left(G_j\right) \\ &\le \sum_j \mu\left(E_j\right), \end{aligned}$$

as claimed. □

Corollary 4.1.19. *Let $(X, \mathcal{B}, \mu)$ be a measure space. If $B, C \in \mathcal{B}$ and $\mu(C) = 0$, then $\mu(B \cup C) = \mu(B)$.*

Proof. The result follows from the inequalities

$$\mu(B) \le \mu(B \cup C) \le \mu(B) + \mu(C) = \mu(B).$$ □

Suppose $\{A_j\}$ is a countably infinite collection of sets in some measure space with

$$A_1 \subseteq A_2 \subseteq A_3 \subseteq \ldots \subseteq A_j \subseteq \ldots .$$

We know that $A = \bigcup_{j=1}^{\infty} A_j$ will be a measurable set since it is the countable union of measurable sets. It seems reasonable to expect $\mu(A) = \lim_{n\to\infty} \mu(A_n)$. In other words, what we really would like to say is that

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

However, we have no definition of a limit of a sequence of sets. Nonetheless, our intuition is correct in this case, as we show in the next lemma.

Lemma 4.1.20. *Let $(X, \mathcal{B}, \mu)$ be a measure space. If $\{A_j\}$ is a countably infinite collection of sets in $\mathcal{B}$ with*

$$A_1 \subseteq A_2 \subseteq A_3 \subseteq \ldots \subseteq A_j \subseteq \ldots ,$$

then

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

Proof. If $\mu(A_j) = \infty$ for some j, then it must be the case that

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

On the other hand, $A_j \subseteq A_n$ for all $n \geq j$, and hence by Proposition 4.1.17 $\mu(A_n) = \infty$ for all $n \geq j$. Therefore, in this case,

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

So we will assume that $\mu(A_n)$ is finite for all n:

$$\bigcup_{j=1}^{\infty} A_j = A_1 \cup \bigcup_{j=1}^{\infty} \left(A_{j+1} \setminus A_j\right) .$$

Because $\mu(A_j)$ is finite for all j, $\mu\left(A_{j+1} \setminus A_j\right) = \mu\left(A_{j+1}\right) - \mu\left(A_j\right)$ by Proposition 4.1.17. Thus,

$$\begin{aligned}
\mu\left(\bigcup_{j=1}^{\infty} A_j\right) &= \mu\left(A_1 \cup \bigcup_{j=1}^{\infty} \left(A_{j+1} \setminus A_j\right)\right) \\
&= \mu(A_1) + \sum_{j=1}^{\infty} \mu\left(A_{j+1} \setminus A_j\right) \\
&= \mu(A_1) + \sum_{j=1}^{\infty} \left(\mu(A_{j+1}) - \mu(A_j)\right) \\
&= \mu(A_1) + \lim_{n\to\infty} \sum_{j=1}^{n} \left(\mu(A_{j+1}) - \mu(A_j)\right) \\
&= \mu(A_1) + \lim_{n\to\infty} \left(\mu(A_{n+1}) - \mu(A_1)\right) \\
&= \lim_{n\to\infty} \mu(A_n),
\end{aligned}$$

as claimed. □

In Remark 4.1.14 we found a subset of the Cantor set that is not a Borel set (assuming the Axiom of Choice). In other words, if we consider the measure space $(\mathbb{R}, \mathbb{B}, m)$, there is a subset of the Cantor set, a set of Lebesgue measure 0, which is not included in the measurable space $(\mathbb{R}, \mathbb{B})$. Yet if we consider the measure space $(\mathbb{R}, \mathcal{M}, m)$, every subset of a set with measure 0 is included in the measurable space $(\mathbb{R}, \mathcal{M})$, as shown in Exercise 6 of Chapter 1. Such a measure space is known as a complete measure space.

Definition 4.1.21. A measure space $(X, \mathcal{B}, \mu)$ is a **complete measure space** if whenever $C \in \mathcal{B}$ with $\mu(C) = 0$ and $A \subseteq C$, then $A \in \mathcal{B}$.

Example 4.1.22. The measure space $(\mathbb{R}, \mathcal{M}, m)$ is a complete measure space.

Given a measure space it is always possible to create a complete measure space that is a natural extension of the given space. In other words, sets that are considered measurable with the original measure will be measurable under the extension measure. We prove this in the following theorem.

Theorem 4.1.23. *Let $(X, \mathcal{B}, \mu)$ be a measure space. There exists a complete measure space $(X, \mathcal{B}_0, \mu_0)$ such that*

(i) $\mathcal{B} \subseteq \mathcal{B}_0$,

(ii) *if $B \in \mathcal{B}$, then $\mu(B) = \mu_0(B)$, and*

(iii) *$E \in \mathcal{B}_0$ if and only if $E = A \cup B$, where $B \in \mathcal{B}$ and $A \subseteq C$ for some $C \in \mathcal{B}$ with $\mu(C) = 0$.*

Proof. Let

$$\mathcal{B}_0 = \{E = A \cup B \quad \Big| \quad B \in \mathcal{B} \text{ and } A \subseteq C, \text{ where } C \in \mathcal{B} \text{ and } \mu(C) = 0\}.$$

For every $B \in \mathcal{B}$, $B = \emptyset \cup B$. Hence, $\mathcal{B} \subseteq \mathcal{B}_0$.

Our first task is to show that $(X, \mathcal{B}_0)$ is a measurable space. In other words, we must show that $\mathcal{B}_0$ is a σ-algebra. The first requirement of a σ-algebra is straightforward as $\emptyset = \emptyset \cup \emptyset$, $\emptyset \in \mathcal{B}$, and $\mu(\emptyset) = 0$.

The second requirement is also straightforward. Assume $\{E_j\}$ is a countable collection of sets in $\mathcal{B}_0$. Then $E_j = A_j \cup B_j$, where $B_j \in \mathcal{B}$ and $A_j \subseteq C_j$ for some $C_j \in \mathcal{B}$ with $\mu(C_j) = 0$. Thus,

$$\bigcup_j E_j = \left(\bigcup_j A_j\right) \cup \left(\bigcup_j B_j\right).$$

Since $\mathcal{B}$ is a σ-algebra,

$$\bigcup_j B_j \in \mathcal{B}.$$

Also,

$$\bigcup_j A_j \subseteq \bigcup_j C_j$$

and

$$\mu\left(\bigcup_j C_j\right) \leq \sum_j \mu(C_j) = 0.$$

Therefore $\bigcup E_j \in \mathcal{B}_0$.

For the third requirement, we observe that if $E \in \mathcal{B}_0$, then $E = A \cup B$, where $B \in \mathcal{B}$ and $A \subseteq C$ for some $C \in \mathcal{B}$ with $\mu(C) = 0$.

Consequently,

$$\begin{aligned} E^c &= (A \cup B)^c \\ &= (C \setminus (A \cup B)) \cup (C^c \cap (A \cup B)^c) \\ &= (C \setminus (A \cup B)) \cup (C^c \cap A^c \cap B^c)\,. \end{aligned}$$

But $A \subseteq C$, so $C^c \subseteq A^c$. Therefore

$$E^c = (C \setminus (A \cup B)) \cup (C^c \cap B^c)$$

with $C^c \cap B^c \in \mathcal{B}$ and $(C \setminus (A \cup B)) \subseteq C$. Hence $E^c \in \mathcal{B}_0$.

It seems natural to define μ_0 on $\mathcal{B}_0$ as follows: if $E \in \mathcal{B}_0$, then $E = A \cup B$, where $B \in \mathcal{B}$ and $A \subseteq C$ for some $C \in \mathcal{B}$ with $\mu(C) = 0$. Set $\mu_0(E) = \mu(B)$. Here is where we run into a potential problem. Given a set $E \in \mathcal{B}_0$, it might be possible that this set E is in $\mathcal{B}_0$ for multiple reasons. That is, it is possible for $E = A \cup B = A_1 \cup B_1$, where $B, B_1 \in \mathcal{B}$, $A \subseteq C$, and $A_1 \subseteq C_1$, where $\mu(C) = \mu(C_1) = 0$. At this point, it is not yet clear that under these different decompositions of E, we will come up with the same value for $\mu_0(E)$. We need to show that our proposed measure μ_0 is well defined.

Suppose $E = A \cup B = A_1 \cup B_1$, where $B, B_1 \in \mathcal{B}$, $A \subseteq C$, and $A_1 \subseteq C_1$, where $\mu(C) = \mu(C_1) = 0$. Then

$$\begin{aligned} B \cup C \cup C_1 &= B \cup A \cup C \cup C_1 \qquad \text{(since } A \subseteq C\text{)} \\ &= E \cup C \cup C_1 \\ &= B_1 \cup A_1 \cup C \cup C_1 = B_1 \cup C \cup C_1, \end{aligned}$$

and hence

$$\mu(B) = \mu(B \cup C \cup C_1) = \mu(B_1 \cup C \cup C_1) = \mu(B_1)\,.$$

Therefore, μ_0 is well defined.

Finally, we will show that μ_0 is a measure. By our definition,

$$\mu_0(\emptyset) = \mu_0(\emptyset \cup \emptyset) = \mu(\emptyset) = 0\,.$$

The first requirement of a measure is satisfied.

If $\{E_j\}$ is a countable collection of pairwise disjoint sets in $\mathcal{B}_0$, then for each j, $E_j = A_j \cup B_j$, where $B_j \in \mathcal{B}$ and $A_j \subseteq C_j$ for some

$C_j \in \mathcal{B}$ with $\mu(C_j) = 0$. Therefore,

$$\bigcup_j A_j \subseteq \bigcup_j C_j$$

and

$$\begin{aligned}
\mu_0\left(\bigcup_j E_j\right) &= \mu_0\left(\bigcup_j A_j \cup \bigcup_j B_j\right) \\
&= \mu\left(\bigcup_j B_j\right) \\
&= \sum_j \mu(B_j) \\
&= \sum_j \mu_0(E_j)\,.
\end{aligned}$$

As a result, μ_0 is a measure. □

Example 4.1.24. By Exercise 9 the completion of $(\mathbb{R}^n, \mathbb{B}, m)$ is $(\mathbb{R}^n, \mathcal{M}, m)$.

4.2. Measurable Functions

Our next goal is to generalize integration. We will start by assuming $(X, \mathcal{B})$ is a measurable space. As we did with Lebesgue integration, we begin by characterizing the type of function we will integrate.

There is one additional consideration we will make at this time. When defining Lebesgue integration, we started by first considering bounded functions and then generalized to unbounded functions. In fact, for purposes of Lebesgue integration, a function need only be defined as a finite number *almost everywhere*. Those places where a function was not finite turned out to be unimportant as long as there weren't too many of them. We will now extend our notion of function to allow a function to take on either the value $+\infty$ or $-\infty$ at those places we previously said a function was undefined. But we have to be a little careful. Thinking ahead to the addition of such functions, we could run into the situation where we need to decide what to do with $\infty - \infty$. However, if this only happens on a set of measure 0, it won't matter how we define this addition.

In other words, we allow a function to take on the values in $\overline{\mathbb{R}} = \mathbb{R} \cup \{-\infty, +\infty\}$. When we write $f : X \to \overline{\mathbb{R}}$ we mean that f is also finite almost everywhere. Once we make this convention, it turns out that the results in this section are very similar to those encountered in Section 2.1.

Definition 4.2.1. Let $(X, \mathcal{B})$ be a measurable space. A function $f : X \to \overline{\mathbb{R}}$ is **measurable with respect to $\mathcal{B}$** or *measurable* if for every $s \in \mathbb{R}$, the set $\{x \in X \mid f(x) > s\}$ is an element of $\mathcal{B}$.

All that is really needed here is a measurable space $(X, \mathcal{B})$ in order to determine whether or not a function is measurable with respect to $\mathcal{B}$. However, most of the time, we will also have at our disposal a measure μ associated with $(X, \mathcal{B})$; in other words, we will have a measure space $(X, \mathcal{B}, \mu)$. When this is the case, we will say a function $f : X \to \overline{\mathbb{R}}$ is *μ-measurable* if for every $s \in \mathbb{R}$, the set $\{x \in X \mid f(x) > s\}$ is in the domain of μ.

Example 4.2.2. Let $X = \{a, b, c, d, e\}$ and

$$\mathcal{B} = \{\emptyset, X, \{a, c\}, \{b, d, e\}, \{b, d\}, \{a, c, e\}, \{e\}, \{a, b, c, d\}\} .$$

(Of course, one should first start by verifying that $(X, \mathcal{B})$ is a measurable space. We are omitting that step here.) Define $f : X \to \mathbb{R}$ and $g : X \to \mathbb{R}$ by

$$f(a) = 3, f(b) = \pi, f(c) = 3, f(d) = \pi, f(e) = \sqrt{2} \text{ and}$$
$$g(a) = 1, g(b) = \pi, g(c) = 3, g(d) = \pi, g(e) = \sqrt{2}\,.$$

To show that f is measurable is very similar to Example 2.1.2. We will look at the following cases:

(i) If $s \geq \pi$, then $\{x \in I \mid f(x) > s\} = \emptyset$, which is a set in $\mathcal{B}$.

(ii) If $s < \sqrt{2}$, then $\{x \in I \mid f(x) > s\} = X$, which is a set in $\mathcal{B}$.

(iii) If $\sqrt{2} \leq s < 3$, then $\{x \in I \mid f(x) > s\} = \{e\}$, which is a set in $\mathcal{B}$.

(iv) If $3 \le s < \pi$, then $\{x \in I \mid f(x) > s\} = \{a, c, e\}$, which is a set in $\mathcal{B}$.

Although f and g differ only by the value assigned to a, f is measurable with respect to $\mathcal{B}$ but g is not. In particular,

$$\{x \in X \bigm| g(x) > 1\} = \{b, c, d, e\} \notin \mathcal{B}\,.$$

As with Lebesgue measurable functions, there are equivalent definitions.

Theorem 4.2.3. *Let $f : X \to \overline{\mathbb{R}}$. The following statements are equivalent:*

(i) *f is measurable.*

(ii) *For every $s \in \mathbb{R}$, the set $\{x \in X \bigm| f(x) \le s\}$ is a measurable set.*

(iii) *For every $s \in \mathbb{R}$, the set $\{x \in X \bigm| f(x) < s\}$ is a measurable set.*

(iv) *For every $s \in \mathbb{R}$, the set $\{x \in X \bigm| f(x) \ge s\}$ is a measurable set.*

Proof. The proof of this theorem is similar to the proof of Theorem 2.1.4. For example, to show that (ii) implies (iii), note that

$$\{x \in X \bigm| f(x) < s\} = \bigcup_{k=1}^{\infty} \{x \in X \bigm| f(x) \le s - \tfrac{1}{k}\}\,.$$

Since $\{x \in I \bigm| f(x) \le s - \frac{1}{k}\}$ is a measurable set for every k, that is, $\{x \in I \bigm| f(x) \le s - \frac{1}{k}\} \in \mathcal{B}$, and $\mathcal{B}$ is a σ-algebra,

$$\{x \in X \bigm| f(x) < s\} = \bigcup_{k=1}^{\infty} \{x \in X \bigm| f(x) \le s - \tfrac{1}{k}\} \in \mathcal{B}\,. \qquad \square$$

Theorem 4.2.4. *Let $f : X \to \overline{\mathbb{R}}$ be a measurable function and let $c \in \mathbb{R}$. Then the following two statements are true:*

(i) *The function $f(x) + c$ is measurable.*

(ii) *The function $cf(x)$ is measurable.*

Proof. The proof is similar to the proof of Theorem 2.1.5. $\square$

Theorem 4.2.5. *Let $f, g : X \to \overline{\mathbb{R}}$ be measurable functions. Then*

(i) *the function $f(x) + g(x)$ is measurable,*

(ii) *the function $f(x)\, g(x)$ is measurable, and*

(iii) *the function $\frac{f(x)}{g(x)}$ is measurable as long as $g \neq 0$ almost everywhere.*

Proof. The proof is similar to the proof of Theorem 2.1.6. □

Recall that in Definition 2.1.10 we defined $f^* = \limsup_{n\to\infty} f_n$ and $f_* = \liminf_{n\to\infty} f_n$. We make similar definitions in this more general setting. The differences here are that we are in a more general measure space and are allowing our functions to take on the values $\pm\infty$.

Definition 4.2.6. Let $\{f_n\}$ be a sequence of functions defined on X.

(i) The **lim sup** of the sequence, written $\limsup_{n\to\infty} f_n$ or denoted by f^*, is defined by

$$\limsup_{n\to\infty} f_n(x) = f^*(x) = \lim_{n\to\infty} \left(\sup\{f_n(x), f_{n+1}(x), f_{n+2}(x), \ldots\}\right) .$$

(ii) The **lim inf** of the sequence, written $\liminf_{n\to\infty} f_n$ or denoted by f_*, is defined by

$$\liminf_{n\to\infty} f_n(x) = f_*(x) = \lim_{n\to\infty} \left(\inf\{f_n(x), f_{n+1}(x), f_{n+2}(x), \ldots\}\right) .$$

Because we are allowing our functions to take on the values $\pm\infty$, f^* and f_* can also take on these values. Fortunately, measurability is preserved as the next theorem demonstrates.

Theorem 4.2.7. *Let $\{f_n\}$ be a pointwise bounded sequence of measurable functions. Then both f^* and f_* are measurable functions on I.*

Proof. As one might expect, the proof is very similar to that for Lebesgue measurable functions. Let

$$\begin{aligned} M_n(x) &= \sup\{f_n(x), f_{n+1}(x), f_{n+2}(x), \ldots\} \text{ and} \\ m_n(x) &= \inf\{f_n(x), f_{n+1}(x), f_{n+2}(x), \ldots\} . \end{aligned}$$

The first step in this proof will be to show that for every $n \in \mathbb{N}$, both $M_n(x)$ and $m_n(x)$ are measurable functions.

Fix $n \in \mathbb{N}$ and let $s \in \mathbb{R}$. We will show that $\{x \in X \mid M_n(x) > s\}$ is a measurable set. Note that $M_n(x) > s$ if and only if $f_k(x) > s$ for some $k \geq n$. Therefore,

$$\{x \in X \mid M_n(x) > s\} = \bigcup_{k=n}^{\infty} \{x \in X \mid f_k(x) > s\}.$$

Since $\{x \in X \mid f_k(x) > s\} \in \mathcal{B}$ for each k, $\{x \in X \mid M_n(x) > s\} \in \mathcal{B}$ by Proposition 4.1.8. Therefore, M_n is a measurable function.

In a similar fashion we will show that $\{x \in X \mid m_n(x) < s\}$ is a measurable set. Note that $m_n(x) < s$ if and only if $f_k(x) < s$ for some $k \geq n$. Therefore,

$$\{x \in X \mid m_n(x) < s\} = \bigcup_{k=n}^{\infty} \{x \in X \mid f_k(x) < s\}.$$

As with M_n, $\{x \in X \mid m_n(x) < s\} \in \mathcal{B}$ because $\mathcal{B}$ is a σ-algebra. Therefore, m_n is a measurable function.

To complete the proof, we observe that for each $x \in X$, the sequence $\{M_n(x)\}$ is a nonincreasing sequence while the sequence $\{m_n(x)\}$ is a nondecreasing sequence. Therefore,

$$\begin{aligned} f^*(x) &= \lim_{n\to\infty} M_n(x) = \inf\{M_n(x)\}, \\ f_*(x) &= \lim_{n\to\infty} m_n(x) = \sup\{m_n(x)\}, \end{aligned}$$

which are measurable by the preceding argument. $\square$

We next wish to generalize the notion of $f = g$ a.e. in this more abstract setting. The definition is exactly what one might expect.

Definition 4.2.8. Let $(X, \mathcal{B}, \mu)$ be a measure space. We say f **equals g almost everywhere with respect to μ** or μ-almost everywhere, written

$$f(x) = g(x) \text{ a.e. } (\mu) \qquad \text{or} \qquad f = g \text{ a.e. } (\mu),$$

if the set $\{x \in X \mid f(x) \neq g(x)\}$ has μ-measure 0. That is,

$$\mu\left(\{x \in X \mid f(x) \neq g(x)\}\right) = 0.$$

Let's consider the following example.

Example 4.2.9. Let $X = \{a, b, c, d, e\}$ and

$$\mathcal{B} = \{\emptyset, X, \{a, c\}, \{b, d\}, \{a, b, c, d\}, \{b, d, e\}, \{a, c, e\}, \{e\}\}\,.$$

Define $\mu : \mathcal{B} \to \overline{\mathbb{R}}$ by

$$\begin{array}{cccc} \mu(\emptyset)=0\,, & \mu(\{a,c\})=0\,, & \mu(\{b,d\})=\frac{2}{3}\,, & \mu(\{e\})=\frac{1}{3}\,, \\ \mu(\{a,b,c,d\})=\frac{2}{3}\,, & \mu(\{b,d,e\})=1\,, & \mu(\{a,c,e\})=\frac{1}{3}\,, & \mu(X)=1\,. \end{array}$$

(Observe that the first line above completely determines μ.) Define the functions $f, g : X \to \mathbb{R}$ by

$$\begin{array}{ccccc} f(a) = 3\,, & f(b) = \pi\,, & f(c) = 3\,, & f(d) = \pi\,, & f(e) = 5\,, \\ g(a) = 6\,, & g(b) = \pi\,, & g(c) = 2\,, & g(d) = \pi\,, & g(e) = 5\,. \end{array}$$

In this case

$$\mu\left(\{x \in X \mid f(x) \neq g(x)\}\right) = \mu(\{a, c\}) = 0\,.$$

Therefore, $f = g$ a.e. (μ).

The preceding example brings up a potential problem. Notice that the function f is μ-measurable. One should really verify this, but it is a matter of checking a number of cases as in Example 4.2.2. On the other hand, $\{x \in X \mid g(x) > 5\} = \{a\}$. But $\{a\} \notin \mathcal{B}$, so g is not μ-measurable. Thus Propostion 2.1.9 does not directly generalize. Take a moment to think about why this is so. If you do, you should come to the conclusion that the real problem is that although the set $\{a, c\}$ is a μ-measurable set with μ-measure 0, it has a subset $\{a\}$ which is not μ-measurable. Here is where we need a complete measure space.

Proposition 4.2.10. *Let $(X, \mathcal{B}, \mu)$ be a complete measure space. Suppose f and g are two functions defined on X. If f is measurable and $f = g$ a.e. (μ), then g is measurable.*

Proof. The proof of this proposition is exactly the same as the proof of Proposition 2.1.9. Let $Z = \{x \in X \mid f(x) \neq g(x)\}$. Then $\mu(Z) = 0$. Since $(X, \mathcal{B}, \mu)$ is a complete measure space, every subset of Z is measurable (and has measure 0). Given $s \in \mathbb{R}$, in order for $g(x) > s$ either $x \notin Z$ (so that $g(x) = f(x)$) and $f(x) > s$, or $x \in Z$ and

$g(x) > s$. Therefore,

$$\begin{aligned}&\{x \in I \mid g(x) > s\}\\ &\qquad = (\{x \in I \mid f(x) > s\} \setminus Z) \cup \{x \in Z \mid g(x) > s\},\end{aligned}$$

which is a combination of measurable sets. Notice that we need to have a complete measure in order to guarantee that the last set is measurable. Therefore, g is a measurable function. □

Notice how we needed the concept of a complete measure space for the sets $\{x \in Z \bigm| g(x) > s\}$ and $\{x \in Z \bigm| g(x) \leq s\}$, which are subsets of Z, to be measurable.

We now introduce simple functions in this general setting. These will be of importance when we turn to integration with respect to different measures.

Definition 4.2.11. A **simple function** is a function $\phi : X \to \mathbb{R}$ expressible as

$$\phi(x) = \sum_{i=1}^{n} a_i \mathcal{X}_{E_i}(x)$$

where $a_i \in \mathbb{R}$, and the sets E_i are pairwise disjoint measurable sets.

By definition, a simple function is bounded. By Exercise 12 a function is a simple function if and only if it takes on only a finite number of values. Nonetheless, every measurable function can be expressed as the pointwise limit of a sequence of simple functions. Again, the next results are similar to their Lebesgue counterparts.

Theorem 4.2.12. *Let $(X, \mathcal{B}, \mu)$ be a measure space. Given a nonnegative measurable function $f : X \to \mathbb{R} \cup \{+\infty\}$ there exists a sequence of nonnegative simple functions $\{\phi_n\}$ such that*

$$\lim_{n\to\infty} \phi_n(x) = f(x)$$

for all $x \in X$.

Proof. For each n and $1 \leq k \leq 2^{2n} - 1$ set

$$E_k^n = \left\{x \in X \Bigm| \frac{k}{2^n} \leq f(x) < \frac{k+1}{2^n}\right\}$$

and set

$$E^n_{2^{2n}} = \{x \in X \mid f(x) \geq 2^n\}.$$

Since f is measurable, each E^n_k is a measurable set. Define ϕ_n by

$$\phi_n(x) = \sum_{k=1}^{2^{2n}} \frac{k}{2^n} \mathcal{X}_{E^n_k}(x).$$

For all $x \in X$ and positive integers n,

$$0 \leq \phi_n(x) \leq \phi_{n+1}(x) \leq f(x).$$

Thus, for each $x \in X$, the sequence $\{\phi_n(x)\}$ is increasing and, in the case that $f(x)$ is finite, bounded. To show that $\lim\limits_{n\to\infty} \phi_n(x) = f(x)$ we will consider three distinct cases.

The first possibility is that $f(x) = 0$. In this case, $x \notin E^n_k$ for all n and k. Thus, $\phi_n(x) = 0$ for all n. Consequently, $\lim\limits_{n\to\infty} \phi_n(x) = 0 = f(x)$.

The second case is when $f(x) = +\infty$. In this case, $x \in E^n_{2^{2n}}$ for all n, and hence $\phi_n(x) = 2^n$. Therefore, $\lim\limits_{n\to\infty} \phi_n(x) = +\infty = f(x)$.

The third and final case is when $0 < f(x) < +\infty$. In this case the bounded, increasing sequence $\{\phi_n(x)\}$ must converge to some number, say ρ and $\phi_n(x) \leq \rho \leq f(x)$ for all n. If $\rho < f(x)$, then there is a rational number of the form $\frac{j}{2^N}$ with

$$\rho < \frac{j}{2^N} < f(x).$$

But by construction, $\phi_N(x) \geq \frac{j}{2^N}$, a contradiction. Therefore $\rho = f(x)$. □

Corollary 4.2.13. *Let $(X, \mathcal{B}, \mu)$ be a measure space. For any measurable function f, there is a sequence of simple functions $\{\phi_n\}$ such that*

$$\lim_{n\to\infty} \phi_n(x) = f(x)$$

for all $x \in X$.

Proof. Apply the previous theorem to the positive and negative parts of f. □

4.3. Integration

Throughout this section we will assume that $(X, \mathcal{B}, \mu)$ denotes a complete measure space. Now we will define integration over a general measure space. One approach is to follow our development of the Lebesgue integral by first discussing the integral of bounded functions via measurable partitions, etc. However, we will follow another approach which is also sometimes used to develop the Lebesgue integral.

Definition 4.3.1. Let $\phi : X \to \mathbb{R}$ be a nonnegative simple function, that is,

$$\phi(x) = \sum_{i=1}^{n} a_i \mathcal{X}_{E_i}(x),$$

where $a_i \in \mathbb{R}$ and $E_i \in \mathcal{B}$ for each i and the sets E_i are pairwise disjoint. The **integral of ϕ with respect to the measure μ**, written $\displaystyle\int \phi \, d\mu$, is

$$\int_X \phi \, d\mu = \sum_{i=1}^{n} a_i \mu(E_i)\,.$$

The first challenge confronting us when approaching integration through simple functions is to show that the integral of a simple function is well defined. To see why this is the case, consider the following two simple functions:

$$\phi(x) = 2\mathcal{X}_{[0,2]}(x) + 3\mathcal{X}_{(2,3]}(x),$$

$$\psi(x) = 2\mathcal{X}_{[0,1]}(x) + 2\mathcal{X}_{(1,2]}(x) + 3\mathcal{X}_{(2,3]}(x)\,.$$

These both satisfy the definition of a simple function. In fact, these two functions are equal. We need to establish that in such a situation, applying Definition 4.3.1 gives the same result. This is Exercise 14.

Now we define the integral of nonnegative functions.

Definition 4.3.2. Let $f : X \to \mathbb{R} \cup \{+\infty\}$ be a nonnegative measurable function. The **integral of f with respect to the measure μ**, written $\int f \, d\mu$, is

$$\int_X f \, d\mu = \sup \left\{ \int \phi \, d\mu \;\middle|\; \phi \text{ is a simple function with } 0 \le \phi \le f \right\}\,.$$

If this supremum is finite, we say f **is integrable with respect to** μ.

In a sense, this definition is analogous to considering the supremum over all possible lower sums. But with Lebesgue integration, we only considered lower sums when we had a bounded function. In this case, f does not need to be bounded. As mentioned in the opening of Section 2.2, we are using another approach to integration, one that can also be used to define the Lebesgue integral.

Definition 4.3.3. Let $f : X \to \overline{\mathbb{R}}$ be a measurable function. We say f **is integrable with respect to** μ if both f^+ and f^- are integrable with respect to μ. Here f^+ and f^- are the positive and negative parts of f respectively. The **integral of** f **with respect to the measure** μ, written $\displaystyle\int_X f\, d\mu$, is

$$\int_X f\, d\mu = \int_X f^+\, d\mu - \int_X f^-\, d\mu\,.$$

One might wonder why we are once again considering the positive and negative parts of a function. The answer is the same as before: to avoid the situation where we might, in essence, be dealing with an "infinity minus infinity".

Example 4.3.4. Define μ on $\mathcal{P}(\mathbb{R})$, the set of subsets of $\mathbb{R}$, by

$$\mu(A) = \begin{cases} 1 & \text{if } \pi \in A, \\ 0 & \text{otherwise.} \end{cases}$$

As in Example 4.1.16 $(\mathbb{R}, \mathcal{P}(\mathbb{R}), \mu)$ is a measure space. Notice that in this case all functions $f : \mathbb{R} \to \overline{\mathbb{R}}$ are μ-measurable. Let $f \geq 0$. We will determine when f is integrable and, if so, $\displaystyle\int_{\mathbb{R}} f\, d\mu$.

Suppose $\phi(x)$ is a simple function with $0 \leq \phi(x) \leq f(x)$ for all $x \in \mathbb{R}$, say,

$$\phi(x) = \sum_{i=1}^{n} a_i \mathcal{X}_{E_i}(x),$$

where $a_i \geq 0$. Then

$$\begin{aligned} \int_{\mathbb{R}} \phi \, d\mu &= \sum_{i=1}^{n} a_i \mu(E_i) \\ &= \begin{cases} a_i & \text{if for some } i, \pi \in E_i, \\ 0 & \text{otherwise} \end{cases} \\ &= \phi(\pi) \,. \end{aligned}$$

However, $\phi(\pi) \leq f(\pi)$. Therefore,

$$\int_{\mathbb{R}} \phi \, d\mu \leq f(\pi) \,.$$

By the definition of the integral of f, $\displaystyle\int_{\mathbb{R}} f \, d\mu \leq f(\pi)$.

In the case that $f(\pi)$ is finite, we will establish the reverse inequality. All that is needed is the simple function

$$\psi(x) = f(\pi)\mathcal{X}_E(x),$$

where $E = \{\pi\}$. Again by the definition of the integral,

$$f(\pi) = \int_{\mathbb{R}} \psi \, d\mu \leq \int_{\mathbb{R}} f \, d\mu \,.$$

If $f(\pi) = \infty$, consider the sequence of simple functions

$$\psi_n(x) = n\mathcal{X}_E(x),$$

where, again, $E = \{\pi\}$ to see that $n \leq \displaystyle\int_{\mathbb{R}} f \, d\mu$ for all n. Thus, $\displaystyle\int_{\mathbb{R}} f \, d\mu = \infty$.

In both cases, $\displaystyle\int_{\mathbb{R}} f \, d\mu = f(\pi)$.

In the previous example the measure μ had an interesting feature: no matter which function f was integrated with respect to this measure, the integral always returned the answer $f(\pi)$. This is an example of what is often called *point-mass measure*, indicating that all of its weight is carried at a single point. Of course, it is easy to create a point-mass measure concentrated at any point of your choosing.

As with Lebesgue integration, sets of μ-measure 0 do not affect the integral, as the next proposition demonstrates.

Proposition 4.3.5. *Let f and g be two functions that are integrable with respect to μ. If $f(x) = g(x)$ a.e. (μ), then*

$$\int f\,d\mu = \int g\,d\mu\,.$$

Proof. We will first prove this in the case where both f and g are nonnegative. Let $Z = \{x \in X \mid f(x) \neq g(x)\}$. Then $\mu(Z) = 0$. Let ϕ be a simple function with $0 \leq \phi \leq f$, say

$$\phi(x) = \sum_{i=1}^{n} a_i \mathcal{X}_{E_i}(x),$$

where $a_i \geq 0$. Set

$$\psi(x) = \sum_{i=1}^{n} a_i \mathcal{X}_{E_i \setminus Z}(x)\,.$$

Then $0 \leq \psi \leq g$ and

$$\int \psi\,d\mu = \sum_{i=1}^{n} a_i\,\mu(E_i \setminus Z) = \sum_{i=1}^{n} a_i\,\mu(E_i) = \int \phi\,d\mu\,.$$

Thus,

$$\int \phi\,d\mu \leq \int g\,d\mu\,.$$

Taking the supremum over all simple functions $0 \leq \phi \leq f$,

$$\int f\,d\mu \leq \int g\,d\mu\,.$$

The reverse inequality follows in the same manner. Therefore, in this case,

$$\int f\,d\mu = \int g\,d\mu\,.$$

Finally, the general result follows by considering the positive and negative parts of f and g. □

Proposition 4.3.6. *Let f and g be two functions that are integrable with respect to μ. If $0 \leq f(x) \leq g(x)$ a.e. (μ), then*

$$\int f\,d\mu \leq \int g\,d\mu\,.$$

Proof. This is Exercise 15. □

At this point, one might expect that we would next show that the space of functions that are integrable with respect to μ is a vector space. However, we will defer this until proving the general version of the Lebesgue Dominated Convergence Theorem.

In Chapter 2, our approach was to first prove the Lebesgue Dominated Convergence Theorem, then use it to prove Fatou's Lemma, and then followed by the Monotone Convergence Theorem. This time, however, we will change the order (so that the reader can see another approach to the "big three" theorems). First we will prove Fatou's Lemma, then the Monotone Convergence Theorem, and, finally, the Lebesgue Dominated Convergence Theorem in this more general setting.

Theorem 4.3.7 (Fatou's Lemma, preliminary version)**.** *Let $\{f_n\}$ be a sequence of measurable nonnegative functions on the complete measure space $(X, \mathcal{B}, \mu)$ with $\lim_{n\to\infty} f_n(x) = f(x)$ for all $x \in X$. Then*

$$\int f\, d\mu \le \liminf_{n\to\infty} \int f_n\, d\mu\,.$$

Proof. Suppose ϕ is a simple function with $0 \le \phi(x) \le f(x)$ for all $x \in X$, say

$$\phi(x) = \sum_{i=1}^{m} a_i \mathcal{X}_{E_i}(x),$$

where $a_i \ge 0$ and the sets E_i are μ-measurable. Without loss of generality we may assume $a_i > 0$ for all i. To prove this theorem, we need to show that

$$\int \phi\, d\mu \le \liminf_{n\to\infty} \int f_n\, d\mu\,.$$

i) Suppose $\int \phi\, d\mu = +\infty$. (This means $\int f\, d\mu = +\infty$.) It must be the case that $\mu(E_j) = +\infty$ for some j. We will use this to show that there is some n with $\int f_k\, d\mu = +\infty$ for all $k \ge n$, which will in turn make $\liminf_{n\to\infty} \int f_n\, d\mu = +\infty$.

Set $a = a_j/2$, where j is the integer with $\mu(E_j) = +\infty$. As a result,

$$E_j \subseteq \{x \in X \mid \phi(x) > a\}\,.$$

But $\phi(x) \leq f(x)$, and hence

$$E_j \subseteq A = \{x \in X \mid f(x) > a\}\,.$$

Consequently, $\mu(A) = +\infty$. Let

$$A_n = \{x \in X \mid f_k(x) > a \text{ for all } k \geq n\}\,.$$

By definition

$$A_1 \subseteq A_2 \subseteq A_3 \subseteq \dots\,.$$

Therefore, we have created a nested sequence of sets. By our hypotheses $\lim\limits_{n\to\infty} f_n(x) = f(x)$. Thus, if $f(x) > a$, then there is some n for which $f_k(x) > a$ for all $k \geq n$. Consequently,

$$A \subseteq \bigcup_{n=1}^{\infty} A_n\,.$$

By Lemma 4.1.20,

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

Set $\phi_n(x) = a\,\mathcal{X}_{A_n}(x)$. By design, $\phi_n(x)$ is a simple function with $\phi_n(x) \leq f_k(x)$ for all $k \geq n$. As a result,

$$a\mu(A_n) = \int \phi_n \, d\mu \leq \int f_k \, d\mu$$

for all $k \geq n$ or

$$a\mu(A_n) \leq \inf_{k\geq n} \int f_k \, d\mu\,.$$

Finally, taking the limit as n goes to infinity, we have

$$\begin{aligned} \int \phi \, d\mu = \quad +\infty \quad &= \lim_{n\to\infty} a\mu(A_n) \\ &\leq \liminf_{n\to\infty} \int f_n \, d\mu\,. \end{aligned}$$

ii) Suppose $\int \phi \, d\mu$ is finite. Then the set

$$A = \{x \in X \mid \phi(x) > 0\} = \bigcup_{i=1}^{m} E_i$$

has finite measure. By assumption $0 \le \phi(x) \le f(x)$ for all $x \in X$. Let $0 < \epsilon < 1$. If $x \in A$, then

$$0 < (1-\epsilon)\phi(x) < \phi(x) \le f(x) \, .$$

(We are using the ϵ to create a strict inequality.) Using an argument similar to that in part i), for each $x \in A$, $(1-\epsilon)\phi(x) < f_k(x)$ for all sufficiently large k. As before, we set

$$A_n = \{x \in A \mid (1-\epsilon)\phi(x) < f_k(x) \text{ for all } k \ge n\} \, .$$

Then

$$A_1 \subseteq A_2 \subseteq A_3 \subseteq \dots$$

and

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

(In part i) we only used "$\subseteq$". Think about why we have "$=$" this time.) By Lemma 4.1.20,

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

Consequently,

$$\lim_{n\to\infty} \mu(A \setminus A_n) = \lim_{n\to\infty} (\mu(A) - \mu(A_n)) = 0 \, .$$

(It is for this last statement that we need the set A to have finite measure.) Thus, for n sufficiently large, say, for $n \ge N$, $\mu(A \setminus A_n) < \epsilon$.

The function $\phi \, \mathcal{X}_{A_n}$ is a simple function as is $(1-\epsilon)\phi \, \mathcal{X}_{A_n}$. In fact,

$$(1-\epsilon)\phi(x) \, \mathcal{X}_{A_n}(x) = \sum_{i=1}^{m} (1-\epsilon) a_i \, \mathcal{X}_{E_i \cap A_n}(x) \, .$$

If $x \notin A_n$, $(1-\epsilon)\phi \, \mathcal{X}_{A_n}(x) = 0$. If $x \in A_n$, then

$$(1-\epsilon)\phi \, \mathcal{X}_{A_n}(x) = (1-\epsilon)\phi(x) < f_k(x)$$

for all $k \geq n$. Therefore, for all $x \in X$ and $k \geq n$,

$$0 \leq (1-\epsilon)\phi(x)\,\mathcal{X}_{A_n}(x) \leq f_k(x)\,.$$

Let $M = \max\{a_i\}$. Then for $n \geq N$ (hence, $\mu(A \setminus A_n) < \epsilon$) and $k \geq n$,

$$\begin{aligned}
\int f_k\,d\mu &\geq \int (1-\epsilon)\phi\,\mathcal{X}_{A_n}\,d\mu \\
&= \sum_{i=1}^{m}(1-\epsilon)a_i\,\mu(E_i \cap A_n) \\
&= \sum_{i=1}^{m}(1-\epsilon)a_i\,\left(\mu(E_i) - \mu(E_i \setminus A_n)\right) \\
&= \sum_{i=1}^{m}(1-\epsilon)a_i\,\mu(E_i) - \sum_{i=1}^{m}(1-\epsilon)a_i\,\mu(E_i \setminus A_n) \\
&\geq (1-\epsilon)\sum_{i=1}^{m} a_i\,\mu(E_i) - M(1-\epsilon)\sum_{i=1}^{m}\mu(E_i \setminus A_n) \\
&= (1-\epsilon)\int \phi\,d\mu - M(1-\epsilon)\,\mu(A \setminus A_n) \\
&\geq (1-\epsilon)\int \phi\,d\mu - M(1-\epsilon)\epsilon\,.
\end{aligned}$$

(Check this line by line to see where we have used the definition of the integral of a simple function and properties of a measure.) Consequently,

$$\liminf_{n\to\infty}\int f_n\,d\mu \geq (1-\epsilon)\int \phi\,d\mu - M(1-\epsilon)\epsilon\,.$$

Since ϵ is arbitrary,

$$\liminf_{n\to\infty}\int f_n\,d\mu \geq \int \phi\,d\mu,$$

as claimed.

Finally, in both cases we have reached the result

$$\int \phi\,d\mu \leq \liminf_{n\to\infty}\int f_n\,d\mu\,.$$

But we started with an arbitrary simple function ϕ with $0 \le \phi(x) \le f(x)$ for all $x \in X$. By the definition of the integral,

$$\int f\, d\mu \le \liminf_{n\to\infty} \int f_n\, d\mu\,,$$

as claimed. □

It took many steps to prove this first result. But as is often the case, the equivalent results followed with much less effort. For example, the result traditionally known as Fatou's Lemma follows as a corollary. The proof is similar to the proof of Corollary 2.4.12 and will not be repeated here.

Corollary 4.3.8 (Fatou's Lemma)**.** *Let $\{f_n\}$ be sequence of measurable nonnegative functions on the complete measure space $(X, \mathcal{B}, \mu)$ and f a nonnegative function with $\lim_{n\to\infty} f_n(x) = f(x)$ a.e. (μ). Then*

$$\int f\, d\mu \le \liminf_{n\to\infty} \int f_n\, d\mu\,.$$

We next turn to the Monotone Convergence Theorem. Our proof must be different than that of Theorem 2.4.13 because we don't yet have the Lebesgue Dominated Convergence Theorem. Despite this, the proof is relatively straightforward.

Theorem 4.3.9 (Monotone Convergence Theorem)**.** *Let $\{f_n\}$ be a sequence of nonnegative measurable functions with $f_n(x) \le f_{n+1}(x)$ a.e. (μ) for every n. Suppose $\lim_{n\to\infty} f_n(x) = f(x)$ a.e. (μ). Then*

$$\int f\, d\mu = \lim_{n\to\infty} \int f_n\, d\mu\,.$$

Proof. By Fatou's Lemma, Corollary 4.3.8,

$$\int f\, d\mu \le \liminf_{n\to\infty} \int f_n\, d\mu\,.$$

By the definitions of lim inf and lim sup,

$$\liminf_{n\to\infty} \int f_n\, d\mu \le \limsup_{n\to\infty} \int f_n\, d\mu\,.$$

Since the sequence $\{f_n(x)\}$ is a nondecreasing sequence for almost every $x \in X$, $f_n(x) \leq f(x)$ a.e. (μ) for all n. Hence by Proposition 4.3.6,

$$\int f_n \, d\mu \leq \int f \, d\mu \,.$$

Consequently,

$$\limsup_{n\to\infty} \int f_n \, d\mu \leq \int f \, d\mu \,.$$

Putting these all together we have

$$\int f \, d\mu \leq \liminf_{n\to\infty} \int f_n \, d\mu \leq \limsup_{n\to\infty} \int f_n \, d\mu \leq \int f \, d\mu \,.$$

Therefore,

$$\int f \, d\mu = \liminf_{n\to\infty} \int f_n \, d\mu = \limsup_{n\to\infty} \int f_n \, d\mu \,.$$

As a result, $\displaystyle\lim_{n\to\infty} \int f_n \, d\mu$ exists and

$$\int f \, d\mu = \lim_{n\to\infty} \int f_n \, d\mu,$$

as claimed. □

Before moving on to the Lebesgue Dominated Convergence Theorem, we will state and prove the first step in showing that the space of all μ-integrable functions is a vector space.

Proposition 4.3.10. *Let f and g be nonnegative measurable functions. For any nonnegative numbers a and b,*

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

Proof. Although we will not explicitly show it here, the result is true if both f and g are simple functions. This is because the linear combination of simple functions is again a simple function.

Assuming this, we will prove this in the case of more general nonnegative measurable functions.

By Theorem 4.2.12, there are sequences $\{\phi_n\}$ and $\{\psi_n\}$ of nonnegative simple functions with

$$\lim_{n\to\infty} \phi_n(x) = f(x) \qquad \text{and} \qquad \lim_{n\to\infty} \psi_n(x) = g(x)$$

for all $x \in X$. Therefore,

$$\lim_{n\to\infty}(a\,\phi_n(x) + b\,\psi_n(x)) = a\,f(x) + b\,g(x)$$

for all $x \in X$. By construction in the proof of Theorem 4.2.12, for all n, $0 \le \phi_n(x) \le f(x)$ and $0 \le \psi_n(x) \le g(x)$; consequently,

$$0 \le a\,\phi_n(x) + b\,\psi_n(x) \le a\,f(x) + b\,g(x)$$

for all $x \in X$. By Theorem 4.3.9,

$$\begin{aligned}\int (a\,f + b\,g)\,d\mu &= \lim_{n\to\infty}\int (a\,\phi_n + b\,\psi_n)\,d\mu \\ &= \lim_{n\to\infty} a\int \phi_n\,d\mu + \lim_{n\to\infty} b\int \psi_n\,d\mu \\ &= a\int f\,d\mu + b\int g\,d\mu,\end{aligned}$$

as claimed. □

As a corollary of this proposition and the Monotone Convergence Theorem we have the following result.

Corollary 4.3.11. *Let $\{f_n\}$ be a sequence of nonnegative measurable functions. Then*

$$\int\left(\sum_{n=1}^{\infty} f_n\right)d\mu = \sum_{n=1}^{\infty}\left(\int f_n\,d\mu\right).$$

Proof. By Theorem 4.2.7, $\displaystyle\sum_{n=1}^{\infty} f_n = \lim_{N\to\infty}\sum_{n=1}^{N} f_n$ is a measurable function. By Proposition 4.3.10 and the Monotone Convergence Theorem (Theorem 4.3.9),

$$\begin{aligned}\sum_{n=1}^{\infty}\left(\int f_n\,d\mu\right) &= \lim_{N\to\infty}\sum_{n=1}^{N}\left(\int f_n\,d\mu\right) \\ &= \lim_{N\to\infty}\int\left(\sum_{n=1}^{N} f_n\,d\mu\right) \\ &= \int\left(\sum_{n=1}^{\infty} f_n\,d\mu\right).\end{aligned}$$

Note that we used the Monotone Convergence Theorem to obtain the last equality. □

We now prove the general form of the Lebesgue Dominated Convergence Theorem.

Theorem 4.3.12 (Lebesgue Dominated Convergence Theorem)**.** *Let $\{f_n\}$ be a sequence of measurable functions such that $\lim_{n\to\infty} f_n(x)=f(x)$ a.e. (μ). Suppose there exists a μ-integrable function g with*

$$|f_n(x)| \le g(x) \ \text{a.e.}\ (\mu)$$

for every n. Then, f_n is μ-integrable for every n, f is μ-integrable, and

$$\lim_{n\to\infty} \int f_n \, d\mu = \int f \, d\mu \,.$$

Proof. Since $|f_n(x)| \le g(x)$ a.e. (μ) for every n, $|f(x)| \le g(x)$ a.e. (μ). Hence, by Exercise 17 f_n is μ-integrable for every n and f is μ-integrable.

Applying Fatou's Lemma (Corollary 4.3.8) to both $g - f_n$ and $g + f_n$ we obtain

$$\int (g - f) \, d\mu \le \liminf_{n\to\infty} \int (g - f_n) \, d\mu$$

and

$$\int (g + f) \, d\mu \le \liminf_{n\to\infty} \int (g + f_n) \, d\mu \,.$$

By Exercise 16,

$$\begin{aligned} \int g \, d\mu - \int f \, d\mu &\le \liminf_{n\to\infty} \left(\int g \, d\mu - \int f_n \, d\mu \right) \\ &= \int g \, d\mu - \limsup_{n\to\infty} \int f_n \, d\mu \end{aligned}$$

and

$$\begin{aligned} \int g \, d\mu + \int f \, d\mu &\le \liminf_{n\to\infty} \left(\int g \, d\mu + \int f_n \, d\mu \right) \\ &= \int g \, d\mu + \liminf_{n\to\infty} \int f_n \, d\mu \,. \end{aligned}$$

Consequently,
$$\limsup_{n\to\infty} \int f_n \, d\mu \leq \int f \, d\mu$$
and
$$\int f \, d\mu \leq \liminf_{n\to\infty} \int f_n \, d\mu \, .$$
Collecting these inequalities we have
$$\int f \, d\mu \leq \liminf_{n\to\infty} \int f_n \, d\mu \leq \limsup_{n\to\infty} \int f_n \, d\mu \leq \int f \, d\mu \, .$$
Therefore, $\displaystyle\lim_{n\to\infty} \int f_n \, d\mu$ exists and
$$\lim_{n\to\infty} \int f_n \, d\mu = \int f \, d\mu \, .$$
□

4.4. Measures from Outer Measures

How does one construct a measure? In Chapter 1 we constructed Lebesgue measure from Lebesgue outer measure, which in turn came from a very basic notion of the length of an interval. However, Lebesgue outer measure had a defect; it is possible for the outer measure of the union of two disjoint sets to be strictly greater than the sum of the outer measures of the two sets. As a result, Lebesgue measure is not defined for all subsets of $\mathbb{R}^n$. We are going to mimic this process. Given a set X, suppose we had something similar to Lebesgue outer measure, that is, some preliminary notion of the size of subsets of X. What properties should this preliminary measure have? From this preliminary measure, how does one define a measure?

Definition 4.4.1. Let X be a set. An **outer measure** on X is a function $\mu^* : \mathcal{P}(X) \to [0, +\infty]$ such that

i) $\mu^*(\emptyset) = 0$,

ii) if $A \subseteq B$, then $\mu^*(A) \leq \mu^*(B)$, and

iii) if $\{A_j\}$ is a countable collection of subsets of X, then
$$\mu^*\left(\bigcup_j A_j\right) \leq \sum_j \mu^*(A_j) \, .$$

Notice that an outer measure is defined for all subsets of X. Also, some texts modify condition iii) by requiring the collection $\{A_j\}$ to be a countable collection of pairwise disjoint subsets of X. However, this is in fact equivalent to our condition.

Example 4.4.2. Lebesgue outer measure m^* is an outer measure on $\mathbb{R}^n$.

We defined a set E to be Lebesgue measurable if for every $\epsilon > 0$ there exists an open set G containing E such that $m^*(G \setminus E) < \epsilon$. However, in a more abstract setting we won't necessarily have a well-defined notion of "open sets". Therefore, we will use a different definition of measurable set.

Definition 4.4.3. Let μ^* be an outer measure on a set X. A set $E \subseteq X$ is called **μ^*-measurable** if for every set $A \subseteq X$,

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

The condition "$\mu^*(A) = \mu^*(A \cap E) + \mu^*(A \setminus E)$ for any $A \subseteq X$" is due to Carathéodory and is called the Carathéodory condition.

By the definition of outer measure, it will always be the case that

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

Consequently, when we need to show that a set E is μ^*-measurable, the goal will often be to show that for any set A,

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

But we are getting ahead of ourselves. Our first task is to show that we have produced a measurable space and a measure on that space. The candidate for our σ-algebra is

$$\mathcal{B} = \{E \subseteq X \mid E \text{ is } \mu^*\text{-measurable}\}.$$

Theorem 4.4.4. *$(X, \mathcal{B})$ is a measurable space.*

Proof. We must show that $\mathcal{B}$ is a σ-algebra.

i) We must show that $\emptyset \in \mathcal{B}$. By definition, $\mu^*(\emptyset) = 0$. Hence, for any set A,

$$\mu^*(A \cap \emptyset) + \mu^*(A \setminus \emptyset) = \mu^*(\emptyset) + \mu^*(A) = \mu^*(A)$$

or

$$\mu^*(A) = \mu^*(A \cap \emptyset) + \mu^*(A \setminus \emptyset) .$$

Therefore, $\emptyset \in \mathcal{B}$.

ii) Assume $E \in \mathcal{B}$. We must show $E^c \in \mathcal{B}$. For any set A

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

The last equality holds because E is μ^*-measurable. Therefore, E^c is μ^*-measurable.

iii) We will start by considering the union of just two sets. Suppose E_1 and E_2 are both in $\mathcal{B}$. Let A be any subset of X. It suffices to show that

$$\mu^*(A) \geq \mu^*\left(A \cap (E_1 \cup E_2)\right) + \left(A \setminus (E_1 \cup E_2)\right) .$$

Since $E_2 \in \mathcal{B}$

$$\mu^*(A) = \mu^*(A \cap E_2) + \mu^*(A \setminus E_2) .$$

Since $E_1 \in \mathcal{B}$

$$\begin{aligned} \mu^*(A \setminus E_2) &= \mu^*\left((A \setminus E_2) \cap E_1\right) + \mu^*\left((A \setminus E_2) \setminus E_1\right) \\ &= \mu^*\left((A \setminus E_2) \cap E_1\right) + \mu^*\left(A \setminus (E_1 \cup E_2)\right) \end{aligned}$$

using $A \setminus E_2$ in the definition of measurability. Consequently,

$$\mu^*(A) = \mu^*(A \cap E_2) + \mu^*\left((A \setminus E_2) \cap E_1\right) + \mu^*\left(A \setminus (E_1 \cup E_2)\right) .$$

But $(A \cap E_2) \cup ((A \setminus E_2) \cap E_1) = A \cap (E_1 \cup E_2)$, so

$$\mu^*(A \cap E_2) + \mu^*\left((A \setminus E_2) \cap E_1\right) \geq \mu^*\left(A \cap (E_1 \cup E_2)\right) .$$

Therefore,

$$\begin{aligned} \mu^*(A) &= \mu^*(A \cap E_2) + \mu^*\left((A \setminus E_2) \cap E_1\right) + \mu^*\left(A \setminus (E_1 \cup E_2)\right) \\ &\geq \mu^*\left(A \cap (E_1 \cup E_2)\right) + \mu^*\left(A \setminus (E_1 \cup E_2)\right) \\ &\geq \mu^*(A) . \end{aligned}$$

Hence, the above inequalities must be equalities and

$$\mu^*(A) = \mu^*\left(A \cap (E_1 \cup E_2)\right) + \mu^*\left(A \setminus (E_1 \cup E_2)\right)$$

so that the union of two μ^*-measurable sets is μ^*-measurable.

Notice that in addition to showing that the union of two μ^*-measurable sets is μ^*-measurable, we may also conclude that if E_1 and E_2 are disjoint,

$$\mu^*(A) = \mu^*(A \cap E_1) + \mu^*(A \cap E_2) + \mu^*(A \setminus (E_1 \cup E_2))$$

since, in this case, $(A \setminus E_2) \cap E_1 = A \cap E_1$. Further, by using mathematical induction, we can show that the union of n μ^*-measurable sets is μ^*-measurable. Moreover, if $E_1, E_2, \ldots, E_n$ are pairwise disjoint μ^*-measurable sets, and A is any subset of X, then

$$\begin{aligned} \mu^*(A) &= \mu^*(A \cap E_1) + \mu^*(A \cap E_2) + \ldots \\ &\qquad + \mu^*(A \cap E_n) + \mu^*\left(A \setminus \bigcup_{i=1}^{n} E_n\right). \end{aligned}$$

Now we turn our attention to a more general case. Suppose $\{E_j\}$ is a countably infinite collection of pairwise disjoint μ^*-measurable sets. We wish to show that $\bigcup_j E_j$ is μ^*-measurable. Let A be any subset of X. Set

$$G_n = \bigcup_{j=1}^{n} E_j \qquad \text{and} \qquad G = \bigcup_{j=1}^{\infty} E_j .$$

Then, because we have a union of pairwise disjoint sets and $G_n \subseteq G$,

$$\begin{aligned} \mu^*(A) &= \left(\sum_{j=1}^{n} \mu^*(A \cap E_j)\right) + \mu^*(A \setminus G_n) \\ &\geq \left(\sum_{j=1}^{n} \mu^*(A \cap E_j)\right) + \mu^*(A \setminus G). \end{aligned}$$

Taking the limit as n goes to infinity,

$$\begin{aligned}
\mu^*(A) &\geq \left(\sum_{j=1}^{\infty} \mu^*(A \cap E_j)\right) + \mu^*(A \setminus G) \\
&\geq \mu^*\left(\bigcup_{j=1}^{\infty} (A \cap E_j)\right) + \mu^*(A \setminus G) \\
&= \mu^*\left(A \cap \bigcup_{j=1}^{\infty} E_j\right) + \mu^*(A \setminus G) \\
&= \mu^*(A \cap G) + \mu^*(A \setminus G) \\
&\geq \mu^*(A)\,.
\end{aligned}$$

Once again, the above inequalities must be equalities and

$$\mu^*(A) = \mu^*(A \cap G) + \mu^*(A \setminus G)\,.$$

Thus, G is μ^*-measurable. As an added bonus, we also have in this case

$$\mu^*(A) = \left(\sum_{j=1}^{\infty} \mu^*(A \cap E_j)\right) + \mu^*\left(A \setminus \bigcup_{j=1}^{\infty} E_j\right)\,.$$

The most general case where $\{E_j\}$ are not necessarily pairwise disjoint follows by writing

$$\bigcup_{j=1}^{\infty} E_j = E_1 \cup (E_2 \setminus E_1) \cup (E_3 \setminus (E_1 \cup E_2)) \ldots\,.$$

In other words, we have taken the union of a countable collection of μ^*-measurable sets and written it as the union of a countable collection of pairwise disjoint μ^*-measurable sets.

We have now shown that $\mathcal{B}$ is a σ-algebra. Therefore, $(X, \mathcal{B})$ is a measurable space. $\square$

Now that we have a measurable space, there is a natural function to use as a measure on this space, namely μ^*. We next show that this indeed is a measure.

Theorem 4.4.5. *Let μ^* be an outer measure on X and $\mathcal{B}$ be the collection of all μ^*-measurable sets. Define $\mu : \mathcal{B} \to [0, \infty]$ by $\mu(E) = \mu^*(E)$. Then $(X, \mathcal{B}, \mu)$ is a measure space.*

Proof. We need to show that μ is a measure. Since $\mathcal{B}$ is a σ-algebra and μ^* is an outer measure, $\emptyset \in \mathcal{B}$ and

$$\mu(\emptyset) = \mu^*(\emptyset) = 0\,.$$

Next, suppose $\{E_j\}$ is a countable collection of pairwise disjoint sets in $\mathcal{B}$. In the proof of Theorem 4.4.4, we saw that for any set $A \subseteq X$,

$$\mu^*(A) = \left(\sum_{j=1}^{\infty} \mu^*(A \cap E_j)\right) + \mu^*\left(A \setminus \bigcup_{j=1}^{\infty} E_j\right)\,.$$

This must hold if we take $A = \bigcup_{j=1}^{\infty} E_j$. Hence,

$$\begin{aligned} \mu^*\left(\bigcup_{j=1}^{\infty} E_j\right) &= \left(\sum_{j=1}^{\infty} \mu^*(E_j)\right) + \mu^*(\emptyset) \\ &= \left(\sum_{j=1}^{\infty} \mu^*(E_j)\right)\,. \end{aligned}$$

Therefore,

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

and μ is a measure. □

Let us return to Lebesgue outer measure. The next theorem asserts that the sets we defined as Lebesgue measurable in Chapter 1 are precisely those sets that are measurable under this new way of generating a measure.

Theorem 4.4.6. *Let $E \subseteq \mathbb{R}^n$. E is Lebesgue measurable if and only if for any $A \subseteq \mathbb{R}^n$,*

$$m^*(A) = m^*(A \cap E) + m^*(A \setminus E),$$

where m^* is Lebesgue outer measure.

Proof. Assume first that E is Lebesgue measurable. Let A be a subset of $\mathbb{R}^n$. It will always be the case that

$$m^*(A) \leq m^*(A \cap E) + m^*(A \setminus E)\,.$$

Our task is to establish equality, or at least, establish the reverse inequality.

If $m^*(A) = +\infty$, it must be the case that

$$+\infty = m^*(A) \leq m^*(A \cap E) + m^*(A \setminus E)\,.$$

Hence,

$$m^*(A \cap E) + m^*(A \setminus E) = +\infty = m^*(A),$$

and we are done.

Now suppose $m^*(A) < +\infty$. By Exercise 24 of Chapter 1, there is a set H of type G_δ such that $A \subseteq H$ and $m^*(A) = m^*(H)$. H is Lebesgue measurable, as are the disjoint sets $H \cap E$ and $H \setminus E$. Therefore,

$$\begin{aligned} m^*(A) &= m^*(H) \\ &= m(H) \\ &= m(H \cap E) + m(H \setminus E) \\ &= m^*(H \cap E) + m^*(H \setminus E) \\ &\geq m^*(A \cap E) + m^*(A \setminus E)\,. \end{aligned}$$

Thus, E satisfies the Carathéodory condition.

Now assume that E satisfies the Carathéodory condition. We will show that E is Lebesgue measurable. First consider the case where $m^*(E) < +\infty$. Again by Exercise 24 of Chapter 1, there is a set H of type G_δ such that $E \subseteq H$ and $m^*(E) = m^*(H)$. Since H is of type G_δ, H must be Lebesgue measurable. Set $Z = H \setminus E$. Then $E = H \setminus Z$. If we show that Z is Lebesgue measurable, we are done. Using H in the Carathéodory condition,

$$\begin{aligned} m^*(E) = m^*(H) &= m^*(H \cap E) + m^*(H \setminus E) \\ &= m^*(E) + m^*(Z)\,. \end{aligned}$$

Thus, $m^*(Z) = 0$. But any set with Lebesgue outer measure 0 is measurable. Consequently, E is measurable.

Finally, we consider the case where $m^*(E) = +\infty$ and E satisfies the Carathéordory condition. For each positive integer k let $B_k = \{x \in \mathbb{R}^n \bigm| |x| < k\}$. Then B_k is Lebesgue measurable (it is an open set), so by the first part of this proof, B_k satisfies the Carathéodory condition. The sets satisfying this condition form a σ-algebra by Theorem 4.4.4, thus $E \cap B_k$ satisfies the Carathéodory condition for each k. But $m^*(E \cap B_k)$ is finite. Thus, as shown earlier, $E \cap B_k$ is Lebesgue measurable for each k. Finally,

$$E = \bigcup_{k=1}^{\infty} (E \cap B_k),$$

so E is the countable union of Lebesgue measurable sets. Therefore, E is Lebesgue measurable. □

As an example of using an outer measure to construct a measure we will conclude this section by generating Hausdorff measure. The first step is to define our outer measure. In fact, we will define a whole family of outer measures and measures.

We begin by using the following notation. For a set $A \subseteq \mathbb{R}^n$ let

$$\delta(A) = \sup\{|x - y| \bigm| x, y \in A\}.$$

That is, $\delta(A)$ is the diameter of the set A. Corresponding to each pair $\alpha \geq 0$ and $\epsilon > 0$ we will define an outer measure. To form this outer measure of $E \subseteq \mathbb{R}^n$, we will cover E with a countable collection of sets with small diameter. More precisely, let $\alpha \geq 0$ and $\epsilon > 0$. The H_α^ϵ outer measure of E is

$$H_\alpha^\epsilon(E) = \inf\left\{\sum_k (\delta(A_k))^\alpha \;\middle|\; E \subseteq \bigcup_k A_k \text{ and } \delta(A_k) < \epsilon\right\}.$$

We leave the details of proving that H_α^ϵ is an outer measure as an exercise (see Exercise 20). Some authors choose to to take

$$H_\alpha^\epsilon(E) = \inf\left\{\sum_k \omega_\alpha (\delta(A_k))^\alpha \;\middle|\; E \subseteq \bigcup_k A_k \text{ and } \delta(A_k) < \epsilon\right\},$$

where the constant ω_α is chosen in order to have Hausdorff measure coincide with Lebesgue measure.

The proof that H_α is an outer measure is straightforward and is also left to the reader as an exercise (see Exercise 21). Because this is an outer measure, we can use the Carathéodory condition (see Definition 4.4.3) to determine a σ-algebra of measurable sets. In fact, for every $\alpha \geq 0$, we have a different outer measure which generates a different measure. The next proposition shows a slight connection between these outer measures.

Notice that if $\epsilon_1 < \epsilon_2$, then $H_\alpha^{\epsilon_1}(E) \geq H_\alpha^{\epsilon_2}(E)$. Hence, for any set $E \subseteq \mathbb{R}^n$, $\lim\limits_{\epsilon \to 0^+} H_\alpha^\epsilon(E)$ will either be a nonnegative real number or $+\infty$.

Definition 4.4.7. The **Hausdorff outer measure of dimension** α of a set E, written $H_\alpha(E)$, is

$$H_\alpha(E) = \lim_{\epsilon \to 0^+} H_\alpha^\epsilon(E)\,.$$

Proposition 4.4.8. *If $0 \leq \alpha < \beta$ and $H_\alpha(E) < +\infty$ for some $E \subseteq \mathbb{R}^n$, then $H_\beta(E) = 0$.*

Proof. Let $\epsilon > 0$ be given and suppose $\{A_k\}$ is a covering of E by sets with diameter less than ϵ. Then

$$\sum_k (\delta(A_k))^\beta = \sum_k (\delta(A_k))^{\beta-\alpha} (\delta(A_k))^\alpha \leq \epsilon^{\beta-\alpha} \sum_k (\delta(A_k))^\alpha\,.$$

Therefore, for every $\epsilon > 0$,

$$0 \leq H_\beta^\epsilon(E) \leq \epsilon^{\beta-\alpha} H_\alpha^\epsilon(E)\,.$$

The result follows by taking the limit as ϵ approaches 0. □

Notice that when taking the limit as ϵ approaches 0, it was important for $H_\alpha^\epsilon(E)$ to have a finite limit. We can take another perspective on our argument. If $H_\beta(E)$ happens to be $+\infty$ and $0 \leq \alpha < \beta$, the same reasoning shows that $H_\alpha(E) = +\infty$. We have the following corollary.

Corollary 4.4.9. *If $0 \leq \alpha < \beta$ and $H_\beta(E) = +\infty$ where $E \subseteq \mathbb{R}^n$, then $H_\alpha(E) = +\infty$.*

For each $\alpha \geq 0$ we have constructed an outer measure H_α. The sets that satisfy the Carathéodory condition (see Definition 4.4.3) for the outer measure H_α for all $\alpha \geq 0$ are called Hausdorff measurable sets or simply Hausdorff measurable. Now we will show that most of the sets we encounter are Hausdorff measurable.

Theorem 4.4.10. *Every Borel set in $\mathbb{R}^n$ is Hausdorff measurable.*

The main step here is to establish the following lemma.

Lemma 4.4.11. *Every closed set in $\mathbb{R}^n$ is Hausdorff measurable.*

Proof. Let $\alpha \geq 0$ and let F be a closed set in $\mathbb{R}^n$. By Definition 4.4.3 (the Carathéodory condition) we must show that for any subset $A \subseteq \mathbb{R}^n$,

$$H_\alpha(A) = H_\alpha(A \cap F) + H_\alpha(A \setminus F)\,.$$

But, in fact, we really need to show that

$$H_\alpha(A) \geq H_\alpha(A \cap F) + H_\alpha(A \setminus F)$$

because the reverse inequality is always true. Moreover, it is easy to see that this inequality is true if $H_\alpha(A) = +\infty$. So we assume $H_\alpha(A)$ is finite.

For each $n \in \mathbb{N}$ set

$$B_n = \left\{ x \in A \setminus F \mid d(x, F) \geq \frac{1}{n} \right\}.$$

By Exercise 22, for each n,

$$H_\alpha(A \cap F) + H_\alpha(B_n) = H_\alpha(A \cap F \cup B_n) \leq H_\alpha(A)\,.$$

As a sequence of real numbers, $\{H_\alpha(B_n)\}$ is an increasing sequence. Also, each $B_n \subseteq A \setminus F \subseteq A$ so $\{H_\alpha(B_n)\}$ is a bounded sequence. Therefore, we know that $\lim\limits_{n\to\infty} H_\alpha(B_n)$ exists. However, we don't yet know that this limit equals $H_\alpha(A \setminus F)$. We can only say that $\lim\limits_{n\to\infty} H_\alpha(B_n) \leq H_\alpha(A \setminus F)$. Our proof will be complete once we establish the reverse inequality.

For each n, let $C_n = B_{n+1} \setminus B_n$. If $|i - j| > 1$, not only are C_i and C_j disjoint, the distance between them will be positive. Hence,

$$H_\alpha(C_i \cup C_j) = H_\alpha(C_i) + H_\alpha(C_j)\,,$$

again by Exercise 22. In particular, this will be the case if i and j are both even or both odd. Thus, for any N,

$$\begin{aligned}\sum_{j=1}^{N} H_\alpha(C_{2j}) &= H_\alpha\left(\bigcup_{j=1}^{N} C_{2j}\right) \le H_\alpha(A \setminus F) \qquad \text{and} \\ \sum_{j=1}^{N} H_\alpha(C_{2j+1}) &= H_\alpha\left(\bigcup_{j=1}^{N} C_{2j+1}\right) \le H_\alpha(A \setminus F)\,.\end{aligned}$$

Therefore, the series $\sum_{j=1}^{\infty} H_\alpha(C_j)$ converges (the partial sums are bounded by $2H_\alpha(A \setminus F)$).

Finally, for each n,

$$A \setminus F = B_n \cup \bigcup_{j \ge n} C_j\,.$$

Hence,

$$H_\alpha(A \setminus F) \le H_\alpha(B_n) + \sum_{j \ge n} H_\alpha(C_j)\,.$$

We obtain the desired inequality by taking the limit as n goes to infinity in the above inequality and observing that the last term is simply the "tail" of a convergent series and hence goes to 0 as n goes to infinity.

In conclusion, every closed set F satisfies the Carathéodory condition and is Hausdoff measureable. □

The remaining details of the proof of Theorem 4.4.10 are left as an exercise (Exercise 23).

Proposition 4.4.8 guarantees that for a set A, if you have a non-negative value α where the H_α Hausdorff measure is finite, then for any β larger than α, the H_β Hausdorff measure of A will be 0. On the other hand, the corollary of that proposition states that once one finds a dimension where the Hausdorff measure is infinite, the Hausdorff measure will be infinite for all smaller dimensions. This allows us to define the **Hausdorff dimension** of a set A as

$$\inf\{\alpha \mid H_\alpha(A) = 0\}\,.$$

Equivalently, the Hausdorff dimension of A is

$$\sup\{\alpha \mid H_\alpha(A) = +\infty\}.$$

Example 4.4.12. In $\mathbb{R}^2$, let A be a line segment of length l. A is a Borel set, and so is Hausdorff measurable. Given $\epsilon > 0$, we need roughly $\frac{l}{\epsilon}$ balls of radius ϵ to cover A. For each α, a candidate for $H_\alpha^\epsilon(A)$ is $\frac{l}{\epsilon}(\epsilon)^\alpha$. Consequently, $H_\alpha^\epsilon(A) = +\infty$ if $\alpha < 1$, $H_\alpha^\epsilon(A) = 0$ if $\alpha > 1$, and $H_1^\epsilon(A) = l$. The Hausdorff dimension of a line segment in $\mathbb{R}^2$ is 1.

This example demonstrates a remarkable feature of Hausdorff measure; it does not depend on the dimension of the surrounding space. A line segment of length l will have Hausdorff dimension 1 and Hausdorff measure (of the same dimension) l whether the line segment resides in $\mathbb{R}^1$, $\mathbb{R}^2$, or even $\mathbb{R}^n$. This is not the case with Lebesgue measure.

4.5. Signed Measures

What if we allowed the measure of a set to be negative? First of all, why would we want to do this? We might want to think of the measure of a set as representing the amount of "stuff" in that set, where the "stuff" might be either positive or negative. For example, think of an electric charge. We can extend the definition of a measure, in limited circumstances, and allow some sets to have a negative measure by defining a signed measure.

Throughout this section $(X, \mathcal{B})$ will denote a measurable space.

Definition 4.5.1. Let $(X, \mathcal{B})$ be a measurable space. We say ν **is a signed measure on** $(X, \mathcal{B})$ if $\nu : \mathcal{B} \to [-\infty, +\infty]$, where

i) ν assumes at most one of the values $+\infty, -\infty$,

ii) $\nu(\emptyset) = 0$, and

iii) for any countable collection $\{E_j\}$ of pairwise disjoint sets in $\mathcal{B}$,

$$\nu\left(\bigcup_j E_j\right) = \sum_j \nu(E_j),$$

where equality is taken to mean that the series on the right converges absolutely if $\nu\left(\bigcup_j E_j\right)$ is finite and diverges otherwise.

Before proceeding take a moment to consider why we impose each of these conditions. The first condition appears in order to prevent any chance of ending up with an "infinity minus infinity" situation. Conditions ii) and iii) are certainly desirable of a measure, signed or otherwise.

Example 4.5.2. Let $f \in \mathcal{L}[a,b]$ and $\mathcal{B}$ be the set of all Lebesgue measurable subsets of $[a,b]$. For any set $E \in \mathcal{B}$ define $\nu(E)$ by

$$\nu(E) = \int_E f\,.$$

ν is an example of a signed measure on the measurable space $([a,b],\mathcal{B})$. Verification that ν is a signed measure is left as an exercise (see Exercise 26).

Definition 4.5.3. Let ν be a signed measure on the measurable space $(X,\mathcal{B})$. Then:

i) We say a set $A \in \mathcal{B}$ is **positive** with respect to the signed measure ν if $\nu(A) \geq 0$ for every measurable $E \subseteq A$.

ii) We say a set $A \in \mathcal{B}$ is **negative** with respect to the signed measure ν if $\nu(A) \leq 0$ for every measurable $E \subseteq A$.

iii) We say a set $A \in \mathcal{B}$ is **null** with respect to the signed measure ν if $\nu(A) = 0$ for every measurable $E \subseteq A$.

An easy example of all of the above is the empty set. Additionally, if a set A is both positive and negative with respect to the signed measure ν, then A is null with respect to ν. However, a set may have ν-measure 0 but not be null with respect to ν.

Example 4.5.4. As in Example 4.5.2, let $f \in \mathcal{L}[a,b]$ and $\mathcal{B}$ be the set of all Lebesgue measurable subsets of $[a,b]$ and define $\nu(E)$ by

$$\nu(E) = \int_E f\,.$$

If A is a measurable subset of $[a,b]$ and $f(x) \geq 0$ for all $x \in A$, then A is a positive set.

Lemma 4.5.5. *Let ν be a signed measure on $(X, \mathcal{B})$. Then:*

i) *Each measurable subset of a positive set is itself positive.*

ii) *The countable union of positive sets is a positive set.*

Proof. Part i) follows directly from the definition of a positive set.

To show ii), assume $\{A_j\}$ is a countable collection of positive sets and suppose E is a measurable set with

$$E \subseteq \bigcup_j A_j \,.$$

Set $E_1 = E \cap A_1$. For each $j > 1$ let $E_j = E \cap (A_j \setminus A_{j-1})$. Then $\{E_j\}$ is a countable collection of pairwise disjoint measurable sets and

$$E = \bigcup E_j \,.$$

Since A_j is positive and $E_j \subseteq A_j$, $\nu(E_j) \geq 0$ for each j. Consequently,

$$\nu(E) = \sum_j \nu(E_j) \geq 0 \,.$$

Therefore, $\bigcup_j A_j$ is a positive set. □

Similar statements can be made about negative sets and null sets.

If a set has positive measure, it is not necessarily a positive set. However, every set with finite positive measure contains a subset that is a positive set.

Lemma 4.5.6. *Let ν be a signed measure on $(X, \mathcal{B})$. If $E \in \mathcal{B}$ with*

$$0 < \nu(E) < \infty \,,$$

then there exists a measurable positive subset A of E with $\nu(A) > 0$.

Proof. If E is itself a positive set, we are done.

If E is not a positive set, there must be some measurable set $B \subset E$ with $\nu(B) < 0$. We will first establish that such a set B must have finite measure. If $\nu(B) = -\infty$, then $\nu(E) = \nu(E \setminus B) + \nu(B)$ would imply $\nu(E) = -\infty$, a contradiction (recall that a signed measure can only take on one of the values $+\infty, -\infty$). Therefore, the signed measure of B must be finite. Hence, there is a smallest positive integer n_B,

$$\nu(B) < -\frac{1}{n_B} < 0\,.$$

We now take a slightly different perspective on this. Consider the set of all positive integers that corresponded in this manner to some subset of E. This set contains a smallest integer. Let n_1 be the smallest positive integer such that there exists a measurable subset $E_1 \subset E$ with

$$\nu(E_1) < -\frac{1}{n_1}\,.$$

Since $\nu(E) = \nu(E \setminus E_1) + \nu(E_1)$ and $\nu(E)$ is positive and finite, whereas $\nu(E_1)$ is negative and finite, it must be the case that $\nu(E \setminus E_1)$ is positive and finite.

If the set $E \setminus E_1$ is a positive set, we have found the set claimed. So, assume $E \setminus E_1$ is not a positive set. We will repeat the process again starting with the set $E \setminus E_1$. Consider all subsets of $E \setminus E_1$ with negative measure. Each of these sets corresponds to some positive integer in the manner described above. We will choose n_2 to be the smallest positive integer such that there exists a set $E_2 \subset E \setminus E_1$ with

$$\nu(E_2) < -\frac{1}{n_2}\,.$$

Since it is also the case that $E_2 \subset E$, n_2 would have been under consideration when we chose n_1. Hence, $n_1 \leq n_2$.

If $E \setminus (E_1 \cup E_2)$ is a positive set, we are done. Otherwise, we repeat the process with $E \setminus (E_1 \cup E_2)$ and continue. In general, $E_k \subset E \setminus \bigcup_{j=1}^{k-1}$ with

$$\nu(E_k) < -\frac{1}{n_k} \qquad \text{and}$$
$$n_1 \leq n_2 \leq \ldots \leq n_k\,.$$

If at any point this process terminates, that is, the set $E \setminus \bigcup_{j=1}^{k} E_j$ is a positive set, we are done.

Assume this process continues indefinitely. Set

$$A = E \setminus \bigcup_{k=1}^{\infty} E_k\,.$$

We will (eventually) show that A is the positive set we seek.

The sets E_k are pairwise disjoint and have negative measure, and hence

$$\nu\left(\bigcup_{k=1}^{\infty} E_k\right) = \sum_{k=1}^{\infty} \nu(E_k) < 0\,.$$

Therefore, the set $\bigcup_{k=1}^{\infty} E_k$ was one of the sets considered at the start of the process. This means that the set $\bigcup_{k=1}^{\infty} E_k$ must have finite (yet negative) measure. But

$$\nu\left(\bigcup_{k=1}^{\infty} E_k\right) = \sum_{k=1}^{\infty} \nu(E_k)$$

and $\nu(E_k) < 0$ for each k. Thus, the series $\sum\limits_{k=1}^{\infty} \nu(E_k)$ converges absolutely. Therefore,

$$0 < \sum_{k+1}^{\infty} \frac{1}{n_k} < \sum_{k+1}^{\infty} -\nu(E_k) < \infty\,.$$

In particular,

$$\lim_{k\to\infty} \frac{1}{n_k} = 0 \qquad \text{or} \qquad \lim_{k\to\infty} n_k = +\infty\,.$$

Next,

$$\nu(E) = \nu(A) + \sum_{k=1}^{\infty} \nu(E_k)$$

and the sets E_k all have negative measure. Consequently, $\nu(A) > \nu(E) > 0$. It remains to show that A is a positive set. To this end, suppose B is a measurable subset of A. Then for every k,

$$B \subseteq E \setminus \bigcup_{j=1}^{k-1} E_k\,.$$

If B had negative measure, it would have been a candidate in our process of choosing n_k and E_k. Also, B would have corresponded to some integer n_B. Since $\lim\limits_{k\to\infty} n_k = +\infty$, at some point, B should have been chosen, but it wasn't. Therefore, for each n_k, B did not qualify as a potential corresponding set. In other words,

$$\nu(B) \geq -\frac{1}{n_k}$$

for every k. But then $\nu(B) \geq 0$ and A is a positive set. $\square$

The next proposition asserts that given any signed measure on a measurable space, the space can be decomposed into two disjoint sets, one positive and the other negative.

Proposition 4.5.7 (Hahn Decomposition Theorem)**.** *Let ν be a signed measure on $(X, \mathcal{B})$. Then there exists a positive set A and a negative set B with $X = A \cup B$ and $A \cap B = \emptyset$.*

Proof. By definition, ν can take on at most one of the values $+\infty$, $-\infty$. Without loss of generality, assume ν never takes on the value $+\infty$. Let

$$\lambda = \sup\{\nu(E) \mid E \in \mathcal{B} \text{ and } E \text{ is positive}\}.$$

(In the case that ν does assume the value $+\infty$, we would start this process by considering negative sets.) Since $\emptyset$ is a positive set in $\mathcal{B}$, $\lambda \geq 0$. At this point, it could be the case that λ is $+\infty$. However, we will show that there is a positive set A with $\nu(A) = \lambda$. Since ν is assumed never to take on the value $+\infty$, this will show that λ must be finite.

By the definition of λ, there exists a sequence $\{E_n\}$ of positive sets with

$$\lim_{n\to\infty} \nu(E_n) = \lambda.$$

Let

$$A = \bigcup_{n=1}^{\infty} E_n.$$

By Lemma 4.5.5, A is a positive set. Also, E_n and $A \setminus E_n$ are disjoint sets. Thus,

$$\nu(A) = \nu(E_n) + \nu(A \setminus E_n)$$

for each n. Since A is a positive set and $A \setminus E_n \subseteq A$, $\nu(A \setminus E_n) \geq 0$; hence,

$$\nu(A) \geq \nu(E_n)$$

for each n. Taking the limit as n goes to infinity, $\nu(A) \geq \lambda$. However, by the definition of λ, $\nu(A) \leq \lambda$. Therefore, $\nu(A) = \lambda$ and λ is finite.

Let $B = X \setminus A$. It remains to show that B is a negative set. To the contrary, suppose B is not a negative set. Then B contains a subset E with positive measure. Since ν does not take on the value $+\infty$, $0 < \nu(E) < +\infty$. By Lemma 4.5.6, E must contain a positive set $\tilde{A}$ with $\nu(\tilde{A}) > 0$. Then

$$\nu(\tilde{A}) \subseteq E \subseteq X \setminus A\,.$$

$A \cup \tilde{A}$ is the disjoint union of positive sets. So we now have a positive set with

$$\nu(A \cup \tilde{A}) = \nu(\tilde{A}) + \nu(A) > \nu(A) = \lambda\,,$$

a contradiction. Therefore, B is a negative set. □

The sets A and B guaranteed by this theorem are known as a Hahn decomposition of the space X. Unfortunately, this decomposition is not necessarily unique, as demonstrated by the next example.

Example 4.5.8. As in Example 4.5.2, let $f \in \mathcal{L}[a,b]$ and $\mathcal{B}$ be the set of all Lebesgue measurable subsets of $[a,b]$, and define $\nu(E)$ by

$$\nu(E) = \int_E f\,.$$

Define A and B by

$$\begin{aligned} A &= \{x \in [a,b] \mid f(x) \geq 0\}, \\ B &= \{x \in [a,b] \mid f(x) < 0\}\,. \end{aligned}$$

Then A and B form a Hahn decomposition of $[a,b]$ with respect to the signed measure ν. However, the sets A' and B' defined by

$$\begin{aligned} A' &= \{x \in [a,b] \mid f(x) > 0\}, \\ B' &= \{x \in [a,b] \mid f(x) \leq 0\} \end{aligned}$$

also form a Hahn decomposition of $[a,b]$.

Given a signed measure ν on a measurable space $(X, \mathcal{B})$, we can use the Hahn decomposition to define the positive and negative parts of the measure. For if $X = A \cup B$ is such a decomposition, letting

$$\begin{aligned} \nu^+(E) &= \nu(E \cap A), \\ \nu^-(E) &= -\nu(E \cap B) \end{aligned}$$

gives us two measures defined on $(X, \mathcal{B})$ with $\nu(E) = \nu^+(E) - \nu^-(E)$. There are actually two things to be shown here. First, ν^+ and ν^- are measures, and, second, the positive and negative parts of ν are well defined. These are Exercise 27 and Exercise 28. Armed with this, we then define the measure $|\nu|$ by

$$|\nu|(E) = \nu^+(E) + \nu^-(E)\,.$$

Definition 4.5.9. Let μ_1 and μ_2 be two measures (not signed measures) on the measurable space $(X, \mathcal{B})$. **μ_1 and μ_2 are mutually singular** on $(X, \mathcal{B})$ if there exist disjoint sets $A, B \in \mathcal{B}$ such that $X = A \cup B$ with

$$\mu_1(A) = \mu_2(B) = 0\,.$$

In other words, the Hahn decomposition theorem tells us that any signed measure ν can be decomposed into two measures, the positive and negative parts of ν, and that these two measures are mutually singular.

Example 4.5.10. Let C be the Cantor set. Again, the Cantor set is a Borel set and so is Hausdorff measurable. To find the Hausdorff dimension of the Cantor set, recall that at the kth stage of construction, C_k consisted of 2^k intervals of length $\frac{1}{3^k}$. Hence, a candidate for $H_\alpha^{\frac{1}{k}}(C)$ is $\frac{2^k}{3^{k\alpha}}$. Therefore, to find $H_\alpha(C)$ we will need to take

$$\lim_{k\to\infty} \frac{2^k}{3^{k\alpha}} = \lim_{k\to\infty} \left(\frac{2}{3^\alpha}\right)^k .$$

If $\frac{2}{3^\alpha} < 1$, this limit will be 0. If $\frac{2}{3^\alpha} > 1$, this limit will be $+\infty$. As predicted by Proposition 4.4.8, there is one dimension that is "just right" when $\frac{2}{3^\alpha} = 1$. Therefore, the Hausdorff dimension of the Cantor set is $\alpha = \frac{\ln 2}{\ln 3}$.

4.6. Exercises

(1) Let $\mathcal{A} = \{A \subseteq \mathbb{R} \mid A \text{ is finite or } A^c \text{ is finite}\}$. Show that $\mathcal{A}$ is an algebra on $\mathbb{R}$.

(2) Prove Proposition 4.1.8.

(3) Show that $\{C \subseteq \mathbb{R} \mid C \text{ is countable or } C^c \text{ is countable}\}$ is a σ-algebra on $\mathbb{R}$.

(4) Let $\mathcal{B}_1$ and $\mathcal{B}_2$ be two σ-algebras on a set X. Define $\mathcal{B} = \mathcal{B}_1 \cap \mathcal{B}_2$ by

$$A \in \mathcal{B} \qquad \text{if and only if} \qquad A \in \mathcal{B}_1 \text{ and } A \in \mathcal{B}_2\,.$$

Show that $\mathcal{B}$ is a σ-algebra on X.

(5) Let $f\colon [a,b] \to \mathbb{R}$ be a (Lebesgue) measurable function on $[a,b]$. Prove that the inverse image of every Borel set is a (Lebesgue) measurable set.

(6) Let $(X, \mathcal{B}, \mu)$ be a measure space. Show that if $A, B \in \mathcal{B}$ and $\mu(A \triangle B) = 0$, then

$$\mu(A) = \mu(B)\,.$$

Here $A \triangle B = (A \setminus B) \cup (B \setminus A)$.

(7) Let $(X, \mathcal{B}, \mu)$ be a measure space. Suppose $Y \in \mathcal{B}$. Let $\mathcal{B}_Y$ consist of those sets in $\mathcal{B}$ that are contained in Y and $\mu_Y(E) = \mu(E)$ if $E \in \mathcal{B}_Y$. Show that $(Y, \mathcal{B}_Y, \mu_Y)$ is a measure space.

(8) Let $(X, \mathcal{B}, \mu)$ be a measure space. Prove the following: If $\{E_i\}$ is a countable collection of sets in $\mathcal{B}$ with $\mu(E_1) < \infty$ and $E_i \supseteq E_{i+1}$ for $i = 1, 2, \ldots$, then

$$\mu\left(\bigcap_{i=1}^{\infty} E_i\right) = \lim_{n\to\infty} \mu(E_n)\,.$$

(9) Prove that the completion of the measure space $(\mathbb{R}^n, \mathbb{B}, m)$ is the measure space $(\mathbb{R}^n, \mathcal{M}, m)$.

(10) Let $(X, \mathcal{B})$ be a measurable space and $\{\mu_n\}$ a sequence of measures with the property that for every $E \in \mathcal{B}$,

$$\mu_n(E) \le \mu_{n+1}(E), \qquad n = 1, 2, \ldots\,.$$

Let $\mu(E) = \lim_{n\to\infty} \mu_n(E)$. Show that $(X, \mathcal{B}, \mu)$ is a measure space.

(11) Let $(X, \mathcal{B}, \mu)$ be a complete measure space. Suppose f is measurable and $f = g$ a.e. (μ). Show that g is measurable.

(12) Let $f\colon X \to \mathbb{R}$ be a μ-measurable function. Show that f takes on only a finite number of values if and only if f is equal to a simple function.

(13) Let $(X, \mathcal{B}, \mu)$ be a measure space. Suppose $\phi(x)$ is a simple function defined on X, that is,

$$\phi(x) = \sum_{i=1}^{n} a_i \mathcal{X}_{E_i}(x),$$

where the a_i's are distinct real numbers and the sets E_i are pairwise disjoint. Show that ϕ is a μ-measurable function if and only if each E_i is a μ-measurable set.

(14) Let $(X, \mathcal{B}, \mu)$ be a complete measure space. Show that if ϕ and ψ are two equal nonnegative μ-measurable simple functions, then

$$\int \phi \, d\mu = \int \psi \, d\mu \, .$$

(15) Prove Proposition 4.3.6 without using the linearity of the integral.

(16) Prove the linearity of the integral. That is, let $(X, \mathcal{B}, \mu)$ be a complete measure space. Show that if f and g are both integrable with respect to the measure μ and $a, b \in \mathbb{R}$, then the function $af + bg$ is integrable with respect to μ and

$$\int (af + g) \, d\mu = a \int f \, d\mu + b \int g \, d\mu \, .$$

(17) Let $(X, \mathcal{B}, \mu)$ be a complete measure space. Suppose f, g are measurable and $|f(x)| \le g(x)$ for all $x \in X$. Show that if g is integrable with respect to μ, then f is as well.

(18) Let $(X, \mathcal{B}, \mu)$ be a complete measure space. Let f be integrable with respect to the measure μ. As with Lebesgue integration, for $A \in \mathcal{B}$ we define

$$\int_A f \, d\mu = \int f \mathcal{X}_A \, d\mu \, .$$

Prove: Given $\epsilon > 0$ there is a $\delta > 0$ such that if $A \in \mathcal{B}$ and $\mu(A) < \delta$, then

$$\left| \int_A f \, d\mu \right| < \epsilon \, .$$

(19) Let $(X, \mathcal{B}, \mu)$ be a complete measure space. We say the sequence of functions $\{f_n\}$ *converges in measure* to the function f if for every $\epsilon > 0$ there is an integer N and a measurable set E so that $\mu(E) < \epsilon$, and if $x \notin E$, then

$$|f_n(x) - f(x)| < \epsilon \qquad \text{for all } n \geq N \, .$$

Prove that if $\{f_n\}$ converges in measure to f, then a subsequence of $\{f_n\}$ converges to f almost everywhere.

(20) Let $\alpha \geq 0$ and $\epsilon > 0$. Prove that H_α^ϵ is an outer measure on $\mathbb{R}^n$.

(21) Let $\alpha \geq 0$. Prove that H_α is an outer measure on $\mathbb{R}^n$.

(22) Let $E_1, E_2 \subseteq \mathbb{R}^n$ and $\alpha \geq 0$. Show that if $d(E_1, E_2) > 0$, then

$$H_\alpha(E_1 \cup E_2) = H_\alpha(E_1) + H_\alpha(E_2) \, .$$

(23) Complete the proof of Theorem 4.4.10.

(24) Let μ^* be an outer measure on $\mathbb{R}^n$ with the property that for any two sets, if $d(E_1, E_2) > 0$, then $\mu^*(E_1 \cup E_2) = \mu^*(E_1) + \mu^*(E_2)$. Show that every closed subset of $\mathbb{R}^n$ is μ^*-measurable.

(25) Let μ^* be an outer measure on $\mathbb{R}^n$ with the property that for any two sets, if $d(E_1, E_2) > 0$, then $\mu^*(E_1 \cup E_2) = \mu^*(E_1) + \mu^*(E_2)$. Show that every Borel subset of $\mathbb{R}^n$ is μ^*-measurable.

(26) Verify that the measure described in Example 4.5.2 satisfies the definition of a signed measure.

(27) Let ν be a signed measure on the measurable space $(X, \mathcal{B})$. Suppose A is positive set and define μ^+ to be

$$\mu^+(E) = \mu(E \cap A) \, .$$

Show that μ^+ is a measure on $(X, \mathcal{B})$.

(28) Let ν be a signed measure on the measurable space $(X, \mathcal{B})$. Suppose A and A' are positive sets and B and B' are negative sets with $X = A \cup B = A' \cup B'$ and $A \cap B = A' \cap B' = \emptyset$. Show that for every measurable set E,

$$\begin{aligned} \nu(E \cap A) &= \nu(E \cap A'), \\ \nu(E \cap B) &= \nu(E \cap B') . \end{aligned}$$

http://dx.doi.org/10.1090/stml/078/06

Ideas for Projects

These are some of the topics that I have used for final presentations. In my classes, these presentations are typically around 20 minutes long and are accompanied by a written report (the time allotted for presentations does not always permit demonstration of all the details). Most of the necessary proofs have been outlined here. Further details about these topics can be found in [**1**], [**10**], and [**13**], as well as other sources.

(1) Egorov's Theorem: This theorem shows that if a sequence of functions $\{f_n\}$ converges pointwise to a function f, then, in some sense, the convergence is almost uniform.

Theorem. *Let (f_n) be a sequence of measurable functions on $[a, b]$ that converges pointwise on $[a, b]$ to the function f. Then for every $\epsilon > 0$, there is a closed set $F \subseteq [a, b]$ such that*

$$m([a, b] \setminus F) < \epsilon \text{ and } f_n \to f \text{ uniformly on } F .$$

a) Show that if (f_n) is a sequence of measurable functions on $[a, b]$ that converges pointwise on $[a, b]$ to the function f, then for every $\eta > 0$ and $\delta > 0$ there is a positive integer N and a measurable subset E of $[a, b]$ such that

$$|f_n - f| < \eta \text{ on } E \text{ for all } n \geq N \text{ and } m([a, b] \setminus E) < \delta .$$

To show this you might look at

$$E_n = \{x \in [a,b] \mid |f_k(x) - f(x)| < \eta \text{ for all } k \geq n\}\,.$$

What can be said about $\lim_{n\to\infty} m(E_n)$ and why?

b) Apply the results of the previous steps using $\delta = \epsilon/2^{n+1}$ and $\eta = 1/n$ to create sets $A_n = [a,b] \setminus E_n$. What can be said about

$$A = \bigcap_{n=1}^{\infty} A_n \ ?$$

c) Does the sequence (f_n) converge uniformly to f on A? Use this to justify the existence of the desired closed set F.

d) Illustrate the results of Egorov's Theorem by creating examples.

(2) Convergence in measure: We have several ways of describing how a sequence of functions $\{f_n\}$ converges to a function f, from pointwise to with respect to a norm. Here is another type.

Definition. *Let (f_n) be a sequence of measurable functions on $[a,b]$ and f a measurable function on $[a,b]$. The sequence (f_n) is said to* **converge in measure** *on $[a,b]$ to f provided that for each $\eta > 0$,*

$$\lim_{n\to\infty} m(\{x \in [a,b] \mid |f_n(x) - f(x)| > \eta\}) = 0\,.$$

a) Show that if the sequence (f_n) of measurable functions on $[a,b]$ converges pointwise to f on $[a,b]$, then the sequence converges in measure to f on $[a,b]$. (This uses Egorov's Theorem.)

b) Show that if the sequence (f_n) converges in measure to f on $[a,b]$, then there is a subsequence (f_{n_k}) that converges pointwise a.e. on $[a,b]$ to f. (This is **Riesz's Theorem**.)

c) Show that the Lebesgue Dominated Convergence Theorem remains valid if "pointwise" convergence a.e. is replace by "convergence in measure".

(3) Lebesgue's criterion for Riemann integrability: Lebesgue showed that a bounded function is Riemann integrable if and only if the set of discontinuities is a set of measure zero.

Theorem. *Let f be a bounded function on the closed interval $[a,b]$. Then f is Riemann integrable over $[a,b]$ if and only if the set of points in $[a,b]$ at which f fails to be continuous has measure zero.*

a) Assume f is Riemann integrable. Show that you can take a sequence of (Riemann) partitions, each a refinement of the previous, so that the sequence of upper sums and the sequence of lower sums converge to the integral of f. Use these partitions to define sequences of step functions, (ψ_n) and (φ_n). Where are these functions continuous? Where will they converge to f?

b) To prove the reverse direction, let P_n be partitions where the lengths of the intervals tend to 0 as n goes to infinity. The goal is to show that $(U(f,P_n) - L(f,P_n))$ also goes to 0 as n goes to infinity. (Why will this do the job?) Construct step functions (ψ_n) and (φ_n) based on these partitions. Compare $(U(f,P_n) - L(f,P_n))$ and

$$\int_a^b (\psi(x) - \varphi(x))\, dx\,.$$

If x is never a partition point, and f is continuous at x, show that $(\psi_n(x))$ and $(\varphi_n(x))$ converge to $f(x)$. Use the assumption that f is bounded and one of our convergence of sequences of integrals theorems to complete the result.

(4) Lebesgue's theorem concerning the differentiability of monotone functions: Lebesgue proved that a monotone function must be differentiable almost everywhere.

Theorem. *If the function f is monotone on the open interval (a,b), then it is differentiable almost everywhere on (a,b).*

A self-contained proof of this can be found in [**10**], Royden and Fitzpatrick, pp. 109-112.

(5) Rapidly Cauchy sequences:

Definition. *Let V be a vector space with norm $\| \cdot \|$. A sequence (f_n) in V is said to be* **rapidly Cauchy** *if there is a convergent series of positive numbers $\sum_{k=1}^{\infty} \epsilon_k$ for which*

$$\|f_{k+1} - f_k\| \leq \epsilon_k^2 \text{ for all } k\,.$$

a) Let V be a normed vector space. Show that every rapidly Cauchy sequence in V is a Cauchy sequence. Show that every Cauchy sequence in V has a subsequence that is rapidly Cauchy.

b) Let $p \geq 1$. Suppose (f_n) is a rapidly Cauchy sequence in $L^p[a, b]$ corresponding to the series of positive numbers $\sum_{k=1}^{\infty} \epsilon_k$. Set

$$E_k = \{x \in [a, b] \mid |f_{k+1}(x) - f_k(x)| \geq \epsilon_k\}\,.$$

Use Tchebychev's Inequality to show that $m(E_k) \leq \epsilon_k^p$.

c) Show that there is a subset Z of $[a, b]$ that has measure zero such that for each $x \in [a, b] \backslash Z$ there is a $K = K(x)$ with

$$|f_{k+1}(x) - f_k(x)| < \epsilon_k \text{ for all } k \geq K(x)\,.$$

Suggestion: use the Borel-Cantelli Lemma.

d) For such an x as described in c), and $n \geq K(x)$, show that

$$|f_{k+n}(x) - f_k(x)| < \sum_{j=k}^{k+n-1} |f_{j+1}(x) - f_j(x)|\,.$$

Use this to conclude that for almost every $x \in [a, b]$, $(f_n(x))$ is a Cauchy sequence.

e) Prove:

Theorem. *Let $p \geq 1$. Prove that every rapidly Cauchy sequence in $L^p[a, b]$ converges both with respect to the L^p-norm and pointwise almost everywhere to a function $f \in L^p[a, b]$.*

(6) Convex functions and Jensen's Inequality: If φ is a convex function, this inequality compares integrating the composition of φ and f with evaluating φ at the integral of f. This

uses the following: If f is increasing on (a, b), then f is continuous on (a, b) with the exception of at most a countable number of points.

Definition. *A function φ on (a, b) is said to be* **convex** *if for every pair of points $x_1, x_2 \in (a, b)$ and $t \in [0, 1]$,*

$$\varphi(tx_1 + (1-t)x_2) \leq t\varphi(x_1) + (1-t)\varphi(x_2)\,.$$

Theorem (Jensen's Inequality). *Let φ be a convex function on $(-\infty, \infty)$, f a Lebesgue integrable function over $[0, 1]$, and $\varphi \circ f$ also integrable over $[0, 1]$. Then*

$$\varphi\left(\int_0^1 f\right) \leq \int_0^1 \varphi \circ f\,.$$

a) Show that φ must be continuous.

b) Show that if $a < x_1 < x, x_2 < b$, then

$$\frac{\varphi(x) - \varphi(x_1)}{x - x_1} \leq \frac{\varphi(x_2) - \varphi(x_1)}{x_2 - x_1} \leq \frac{\varphi(x) - \varphi(x_2)}{x - x_2}\,.$$

c) For a function g, define the right-hand derivation $g'(x^+)$ and the left-hand derivative $g(x^-)$. Show that a convex function always has both a right-hand and left-hand derivative at every point in (a, b). Moreover, if $a < u < v < b$, then

$$\varphi'(u^-) \leq \varphi'(u^+) \leq \frac{\varphi(u) - \varphi(v)}{u - v} \leq \varphi'(v^-) \leq \varphi'(v^+)\,.$$

d) Show that if φ is convex and $\alpha \in (a, b)$, then there is a real number m so that

$$m(t - \alpha) + \varphi(\alpha) \leq \varphi(t)$$

for all $t \in (a, b)$.

e) Let $t = f(x)$ and $\alpha = \int_0^1 f$ in the above inequality, and then integrate.

(7) Fubini's Theorem: This is a theorem that looks at when a double integral equals the corresponding iterated integral. This is related to the example discussed in Section 2.5. Some notation: Let I_1 be a closed interval in $\mathbb{R}^n$ and I_2 be a closed

interval in $\mathbb{R}^m$. Then $I = I_1 \times I_2$ is a closed interval in $\mathbb{R}^{n+m}$. Let $\mathbf{x} \in I_1$ and $\mathbf{y} \in I_2$. Then $(\mathbf{x}, \mathbf{y}) \in I$. A function f defined on I will be written as $f(\mathbf{x}, \mathbf{y})$ and its integral over I will be written as $\int \int_I f(\mathbf{x}, \mathbf{y})\, d\mathbf{x} d\mathbf{y}$.

Theorem (Fubini's Theorem). *Let $f \in \mathcal{L}(I)$, where $I = I_1 \times I_2$. Then*

i) *for almost every $\mathbf{x} \in I_1$, as a function of $\mathbf{y}$, $f(\mathbf{x}, \mathbf{y})$ is measurable and integrable on I_2,*
ii) *as a function of $\mathbf{x}$, $\int_{I_2} f(\mathbf{x}, \mathbf{y})\, d\mathbf{y}$ is measurable and integrable on I_1, and*
iii)

$$\int \int_I f(\mathbf{x}, \mathbf{y})\, d\mathbf{x} d\mathbf{y} = \int_{I_1} \left[\int_{I_2} f(\mathbf{x}, \mathbf{y})\, d\mathbf{y} \right] d\mathbf{x}\,.$$

To prove Fubini's Theorem, let $\mathcal{F}$ be the set of functions in $\mathcal{L}(I)$ for which Fubini's Theorem is true.

a) Show that $\mathcal{F}$ is a vector space.
b) Let (f_k) be a sequence of functions in $\mathcal{F}$. If this sequence increases pointwise to a function $f \in \mathcal{L}(I)$, show that $f \in \mathcal{F}$.
c) Show that if E is a measurable subset of I, then $\mathcal{X}_E \in \mathcal{F}$. Do this in steps. First consider the case where E is a closed interval. Then show the result for open sets, sets of type G_δ, sets of measure zero, and finally general measurable sets.
c) Show that if $f \in \mathcal{L}(I)$ is a nonnegative function, then $f \in \mathcal{F}$. The idea is to approximate f by simple functions.
d) Complete the proof of Fubini's Theorem.

(8) Product measures (this assumes material from Section 4.4): Suppose you are given two measure spaces $(X, \mathcal{A}, \mu)$ and $(Y, \mathcal{B}, \nu)$. How could you use this to create a measure on the space $X \times Y$?

a) Let $A \in \mathcal{A}$ and $B \in \mathcal{B}$. What do you expect for the measure of $A \times B$? (Subsets of $X \times Y$ that are of this form will be called *measurable rectangles.*) Not all subsets of $X \times Y$ are of this form. How could you find the "measure" of a set $C \subseteq X \times Y$? Have you described a measure or an outer measure?

b) Let $A \in \mathcal{A}$ and $B \in \mathcal{B}$. Suppose there is a countable collection of measurable rectangles $\{A_k \times B_k\}$ such that

$$A \times B = \bigcup_{k=1}^{\infty} (A_k \times B_k) .$$

Show that

$$\mu(A)\nu(B) = \bigcup_{k=1}^{\infty} (\mu(A_k)\nu(B_k)) .$$

Suggested strategy: Note that if you fix $x \in A$ and consider a possible $y \in B$, the point (x, y) belongs to exactly one $A_k \times B_k$. Explain why

$$B = \bigcup_{\{k | x \in A_k\}} B_k \qquad \text{and}$$

$$\nu(B) = \sum_{\{k | x \in A_k\}} \nu(B_k) .$$

Multiply this last equality by $\mathcal{X}_A(x)$ and integrate in the variable x with respect to the measure μ. (Note where you are using the assumptions that μ and ν are measures. Also, carefully justify interchanging integration and summation.)

c) Use these measurable rectangles to describe an outer measure, λ^*, on $X \times Y$.

d) The measurable subsets of $X \times Y$ are the λ^*-measurable sets. We define the product measure $\mu \times \nu$ of one of these sets as $\mu \times \nu(E) = \lambda^*(E)$. Show that a measurable rectangle is a μ-measurable set.

e) Let m_1 denote Lebesgue measure in $\mathbb{R}$ and m_2 denote Lebesgue measure in $\mathbb{R}^2$. Compare $m_1 \times m_1$, the product measure on $\mathbb{R} \times \mathbb{R}$, with m_2. How does this generalize?

(9) The Riesz representation for the dual of L^p: Let $p > 1$ and $\frac{1}{p} + \frac{1}{q} = 1$. If $g \in L^q[a,b]$, then

$$T(f) = \int_a^b fg$$

is a linear transformation from $L^p[a,b]$ to $\mathbb{R}$ (due to Hölder's inequality). Does this describe all linear transformations from $L^p[a,b]$ to $\mathbb{R}$? In other words, if you have a linear transformation T from $L^p[a,b]$ to $\mathbb{R}$, will there be a $g \in L^q[a,b]$ so that in fact

$$T(f) = \int_a^b fg\,?$$

A proof of this in a slightly more general setting can be found in [**10**], Royden and Fitzpatrick pp. 155-161, *Real Analysis*, Prentice Hall, 4th edition. Adapt the proof to the case where $p > 1$ and $E = [a,b]$.

References

[1] Abbott, Stephen. *Understanding Analysis*. New York: Springer-Verlag, 2001.

[2] Boas, Jr., Ralph P. *A Primer of Real Functions*. Mathematical Association of America, 1996.

[3] Bressoud, David M. *A Radical Approach to Lebesgue's Theory of Integration*. Mathematical Association of America, 2008.

[4] Brown, James W., and Churchill, Ruel V. *Fourier Series and Boundary Value Problems, 7th edition*. McGraw Hill, 2008.

[5] Capiński, Marek, and Kopp, Ekkehard. *Measure, Integral and Probability, 2nd edition*. Springer, 2008.

[6] Edgar, Gerald A. *Integral, Probability, and Fractal Measures*. Springer-Verlag, 1998.

[7] Franks, John. *A (Terese) Introduction to Lebesgue Integration*. American Mathematical Society, 2000.

[8] Kaczor, W. J., and Nowak, M. T. *Problems in Mathematical Analysis III: Integration*. American Mathematical Society, 2001.

[9] Lay, Steven R. *Analysis with an Introduction to Proof, 5th edition*. New York: Pearson, 2013.

[10] Royden, H. L., and Fitzpatrick, P. M. *Real Analysis, 4th edition*. Prentice Hall, 2010.

[11] Rudin, Walter. *Principles of Mathematical Analysis*. McGraw Hill, 1976.

[12] Taylor, Michael E. *Measure Theory and Integration*. American Mathematical Society, 2000.

[13] Wheeden, Richard L., and Zygmund, Antoni. *Measure and Integral: An Introduction to Real Analysis*. CRC Pure and Applied Mathematics, 1977.

[14] Zorn, Paul. *Understanding Real Analysis*. A K Peters, Ltd., 2010.

Index

Photo courtesy of Sara Rubinstein, Carleton College

A User-Friendly Introduction to Lebesgue Measure and Integration provides a bridge between an undergraduate course in Real Analysis and a first graduate-level course in Measure Theory and Integration. The main goal of this book is to prepare students for what they may encounter in graduate school, but will be useful for many beginning graduate students as well. The book starts with the fundamentals of measure theory that are gently approached through the very concrete example of Lebesgue measure. With this approach, Lebesgue integration becomes a natural extension of Riemann integration.

Next, L^p-spaces are defined. Then the book turns to a discussion of limits, the basic idea covered in a first analysis course. The book also discusses in detail such questions as: When does a sequence of Lebesgue integrable functions converge to a Lebesgue integrable function? What does that say about the sequence of integrals? Another core idea from a first analysis course is completeness. Are these L^p-spaces complete? What exactly does that mean in this setting?

This book concludes with a brief overview of General Measures. An appendix contains suggested projects suitable for end-of-course papers or presentations.

The book is written in a very reader-friendly manner, which makes it appropriate for students of varying degrees of preparation, and the only prerequisite is an undergraduate course in Real Analysis.

ISBN: 978-1-4704-2199-1

9 781470 421991

STML/78

For additional information and updates on this book, visit

www.ams.org/bookpages/stml-78