Real analysis for graduate students: measure and integration theory

Richard F. Bass

ISBN-13: 978-1466391574
ISBN-10: 146639157X

To the memory of my parents

Jay and Betty Bass

Contents

Preface

Nearly every Ph.D. student in mathematics needs to pass a preliminary or qualifying examination in real analysis. The purpose of this book is to teach the material necessary to pass such an examination.

I had three main goals in writing this text:

(1) present a very clear exposition;
(2) provide a large collection of useful exercises;
(3) make the text affordable.

Let me discuss each of these in more detail.

(1) There are a large number of real analysis texts already in existence. Why write another? In my opinion, none of the existing texts are ideally suited to the beginning graduate student who needs to pass a "prelim" or "qual." They are either too hard, too advanced, too encyclopedic, omit too many important topics, or take a nonstandard approach to some of the basic theorems.

Students who are starting their graduate mathematics education are often still developing their mathematical sophistication and find that the more details that are provided, the better (within reason). I have tried to make the writing as clear as possible and to provide the details. For the sake of clarity, I present the theorems and results as they will be tested, not in the absolutely most general abstract context. On the other hand, a look at the index will show that no topics that might appear on a preliminary or qualifying examination are omitted.

All the proofs are "plain vanilla." I avoid any clever tricks, sneaky proofs, unusual approaches, and the like. These are the proofs and methods that most experts grew up on.

(2) There are over 220 exercises. I tried to make them interesting and useful and to avoid problems that are overly technical. Many are routine, many are of moderate difficulty, but some are quite challenging. A substantial number are taken from preliminary examinations given in the past at universities with which I have been associated.

I thought long and hard as to whether to provide hints to the exercises. When I teach the real analysis course, I give hints to the harder questions. But some instructors will want a more challenging course than I give and some a less challenging one. I leave it to the individual instructor to decide how many hints to give.

(3) I have on my bookshelf several books that I bought in the early 1970's that have the prices stamped in them: $10-$12. These very same books now sell for $100-$200. The cost of living has gone up in the last 40 years, but only by a factor of 5 or 6, not a factor of 10. Why do publishers make textbooks so expensive? This is particularly troublesome when one considers that nowadays authors do their own typesetting and frequently their own page layout.

My aim was to make the soft cover version of this text cost less than $20 and to provide a version in .pdf format for free. To do that, I am self-publishing the text.

At this point I should tell you a little bit about the subject matter of real analysis. For an interval contained in the real line or a nice region in the plane, the length of the interval or the area of the region give an idea of the size. We want to extend the notion of size to as large a class of sets as possible. Doing this for subsets of the real line gives rise to Lebesgue measure. Chapters 2–4 discuss classes of sets, the definition of measures, and the construction of measures, of which one example is Lebesgue measure on the line. (Chapter 1 is a summary of the notation that is used and the background material that is required.)

Once we have measures, we proceed to the Lebesgue integral. We talk about measurable functions, define the Lebesgue integral, prove the monotone and dominated convergence theorems, look at some simple properties of the Lebesgue integral, compare it to the Riemann integral, and discuss some of the various ways a sequence of functions can converge. This material is the subject of Chapters 5–10.

Closely tied with measures and integration are the subjects of product measures, signed measures, the Radon-Nikodym theorem, the differentiation of functions on the line, and L^p spaces. These are covered in Chapters 11–15.

Many courses in real analysis stop at this point. Others also include some or all of the following topics: the Fourier transform, the Riesz representation theorem, Banach spaces, and Hilbert spaces. We present these in Chapters 16–19.

The prerequisites to this text are a solid background in undergraduate mathematics. An acquaintance with metric spaces is assumed, but no other topology. A summary of what you need to know is in Chapter 1. All the necessary background material can be learned from many sources; one good place is the book [3].

At some universities preliminary or qualifying examinations in real analysis are combined with those in undergraduate analysis or topology or complex analysis. If that is the case at your university, you will have to supplement this text with texts in those subjects.

Further reading is always useful. I have found the books [1], [2], and [4] helpful.

I would like to thank A. Baldenko, I. Ben-Ari, K. Bharath, K. Burdzy, D. Ferrone, E. Giné, M. Gordina, E. Hsu, G. Lawler, L. Lu, K. Marinelli, J. Pitman, M. Poehlitz, H. Ren, L. Rogers, and A. Teplyaev for some very useful suggestions.

If you have comments, suggestions, corrections, etc., I would be glad to hear from you: **r.bass@uconn.edu**. I cannot, however, provide hints or solutions to the exercises.

Good luck with your exam!

Chapter 1

Preliminaries

In this short chapter we summarize some of the notation and terminology we will use and recall a few definitions and results from undergraduate mathematics.

1.1 Notation and terminology

We use A^c, read "A complement," for the set of points not in A. To avoid some of the paradoxes of set theory, we assume all our sets are subsets of some given set X, and to be precise, define

$$A^c = \{x \in X : x \notin A\}.$$

We write

$$A - B = A \cap B^c$$

(it is common to also see $A \setminus B$) and

$$A \Delta B = (A - B) \cup (B - A).$$

The set $A\Delta B$ is called the *symmetric difference* of A and B and is the set of points that are in one of the sets but not the other. If I is some index set, a collection of subsets $\{A_\alpha\}_{\alpha \in I}$ is disjoint if $A_\alpha \cap A_\beta = \emptyset$ whenever $\alpha \neq \beta$.

We write $A_i \uparrow$ if $A_1 \subset A_2 \subset \cdots$ and write $A_i \uparrow A$ if in addition $A = \cup_{i=1}^\infty A_i$. Similarly $A_i \downarrow$ means $A_1 \supset A_2 \supset \cdots$ and $A_i \downarrow A$ means that in addition $A = \cap_{i=1}^\infty A_i$.

We use $\mathbb{Q}$ to denote the set of rational numbers, $\mathbb{R}$ the set of real numbers, and $\mathbb{C}$ the set of complex numbers. We use

$$x \vee y = \max(x, y) \qquad \text{and} \qquad x \wedge y = \min(x, y).$$

We can write a real number x in terms of its positive and negative parts: $x = x^+ - x^-$, where

$$x^+ = x \vee 0 \qquad \text{and} \qquad x^- = (-x) \vee 0.$$

If z is a complex number, then $\overline{z}$ is the complex conjugate of z.

If f is a function whose domain is the reals or a subset of the reals, then $\lim_{y\to x+} f(y)$ and $\lim_{y\to x-} f(y)$ are the right and left hand limits of f at x, resp.

We say a function $f : \mathbb{R} \to \mathbb{R}$ is *increasing* if $x < y$ implies $f(x) \le f(y)$ and f is *strictly increasing* if $x < y$ implies $f(x) < f(y)$. (Some authors use "nondecreasing" for the former and "increasing" for the latter.) We define *decreasing* and *strictly decreasing* similarly. A function is *monotone* if f is either increasing or decreasing.

Given a sequence $\{a_n\}$ of real numbers,

$$\limsup_{n\to\infty} a_n = \inf_n \sup_{m\ge n} a_m,$$

$$\liminf_{n\to\infty} a_n = \sup_n \inf_{m\ge n} a_m.$$

For example, if

$$a_n = \begin{cases} 1, & n \text{ even;} \\ -1/n, & n \text{ odd,} \end{cases}$$

then $\limsup_{n\to\infty} a_n = 1$ and $\liminf_{n\to\infty} a_n = 0$. The sequence $\{a_n\}$ has a limit if and only if $\limsup_{n\to\infty} a_n = \liminf_{n\to\infty} a_n$ and both are finite. We use analogous definitions when we take a limit along the real numbers. For example,

$$\limsup_{y\to x} f(y) = \inf_{\delta>0} \sup_{|y-x|<\delta} f(y).$$

1.2 Some undergraduate mathematics

We recall some definitions and facts from undergraduate topology, algebra, and analysis. The proofs and more details can be found in many places. A good source is [3].

A set X is a *metric space* if there exists a metric $d : X \times X \to \mathbb{R}$ such that
(1) $d(x, y) = d(y, x)$ for all $x, y \in X$;
(2) $d(x, y) \geq 0$ for all $x, y \in X$ and $d(x, y) = 0$ if and only if $x = y$;
(3) $d(x, z) \leq d(x, y) + d(y, z)$ for all $x, y, z \in X$.

Condition (3) is called the triangle inequality.

Given a metric space X, let

$$B(x, r) = \{y \in X : d(x, y) < r\}$$

be the *open ball* of radius r centered at x. If $A \subset X$, the *interior* of A, denoted A^o, is the set of x such that there exists $r_x > 0$ with $B(x, r_x) \subset A$. The *closure* of A, denoted $\overline{A}$, is the set of $x \in X$ such that every open ball centered at x contains at least one point of A. A set A is *open* if $A = A^o$, *closed* if $A = \overline{A}$. If $f : X \to \mathbb{R}$, the *support* of f is the closure of the set $\{x : f(x) \neq 0\}$. f is continuous at a point x if given $\varepsilon > 0$, there exists $\delta > 0$ such that $|f(x) - f(y)| < \varepsilon$ whenever $d(x, y) < \delta$. f is continuous if it is continuous at every point of its domain. One property of continuous functions is that $f^{-1}(F)$ is closed and $f^{-1}(G)$ is open if f is continuous, F is closed, and G is open.

A sequence $\{x_n\} \subset X$ converges to a point $x \in X$ if for each $\varepsilon > 0$ there exists N such that $d(x_n, x) < x$ whenever $n \geq N$. A sequence is a *Cauchy sequence* if for each $\varepsilon > 0$ there exists N such that $d(x_m, x_n) < \varepsilon$ whenever $m, n \geq N$. If every Cauchy sequence in X converges to a point in X, we say X is *complete*.

An *open cover* of a subset K of X is a collection $\{G_\alpha\}_{\alpha \in I}$ of open sets such that $K \subset \cup_{\alpha \in I} G_\alpha$. The index set I can be finite or infinite. A set K is *compact* of every open cover contains a finite subcover, i.e., there exists $G_1, \ldots, G_n \in \{G_\alpha\}_{\alpha \in I}$ such that $K \subset \cup_{i=1}^n G_i$.

We have the following two facts about compact sets.

Proposition 1.1 *If K is compact, $F \subset K$, and F is closed, then F is compact.*

Proposition 1.2 *If K is compact and f is continuous on K, then there exist x_1 and x_2 such that $f(x_1) = \inf_{x \in K} f(x)$ and $f(x_2) = \sup_{x \in K} f(x)$. In other words, f takes on its maximum and minimum values.*

Remark 1.3 If $x \neq y$, let $r = d(x, y)$ and note that $d(x, r/2)$ and $d(y, r/2)$ are disjoint open sets containing x and y, resp. Therefore metric spaces are also what are called Hausdorff spaces.

Let F be either $\mathbb{R}$ or $\mathbb{C}$. X is a *vector space* or *linear space* if there exist two operations, addition (+) and scalar multiplication, such that
(1) $x + y = y + x$ for all $x, y \in X$;
(2) $(x + y) + z = x + (y + z)$ for all $x, y, z \in X$;
(3) there exists an element $0 \in X$ such that $0 + x = x$ for all $x \in X$;
(4) for each x in X there exists an element $-x \in X$ such that $x + (-x) = 0$;
(5) $c(x + y) = cx + cy$ for all $x, y \in X$, $c \in F$;
(6) $(c + d)x = cx + dx$ for all $x \in X$, $c, d \in F$;
(7) $c(dx) = (cd)x$ for all $x \in X$, $c, d \in F$;
(8) $1x = x$ for all $x \in X$.

We use the usual notation, e.g., $x - y = x + (-y)$.

X is a *normed linear space* if there exists a map $x \to \|x\|$ such that
(1) $\|x\| \geq 0$ for all $x \in X$ and $\|x\| = 0$ if and only if $x = 0$;
(2) $\|cx\| = |c| \, \|x\|$ for all $c \in F$ and $x \in X$;
(3) $\|x + y\| \leq \|x\| + \|y\|$ for all $x, y \in X$.

Given a normed linear space X, we can make X into a metric space by setting $d(x, y) = \|x - y\|$.

A set X has an *equivalence relationship* "$\sim$" if
(1) $x \sim x$ for all $x \in X$;
(2) if $x \sim y$, then $y \sim x$;
(3) if $x \sim y$ and $y \sim z$, then $x \sim z$.

Given an equivalence relationship, X can be written as the union of disjoint equivalence classes. x and y are in the same equivalence class if and only if $x \sim y$. For an example, let $X = \mathbb{R}$ and say $x \sim y$ if $x - y$ is a rational number.

A set X has a *partial order* "$\leq$" if
(1) $x \leq x$ for all $x \in X$;
(2) if $x \leq y$ and $y \leq x$, then $x = y$.

Note that given $x, y \in X$, it is not necessarily true that $x \leq y$ or $y \leq x$. For an example, let Y be a set, let X be the collection of all subsets of Y, and say $A \leq B$ if $A, B \in X$ and $A \subset B$.

We need the following three facts about the real line.

Proposition 1.4 *Suppose $K \subset \mathbb{R}$, K is closed, and K is contained in a finite interval. Then K is compact.*

Proposition 1.5 *Suppose $G \subset \mathbb{R}$ is open. Then G can be written as the countable union of disjoint open intervals.*

Proposition 1.6 *Suppose $f : \mathbb{R} \to \mathbb{R}$ is an increasing function. Then both $\lim_{y \to x+} f(y)$ and $\lim_{y \to x-} f(y)$ exist for every x. Moreover the set of x where f is not continuous is countable.*

For an application of Hilbert space techniques to Fourier series, which is the last section of Chapter 19, we will use the *Stone-Weierstrass theorem.* The particular version we will use is the following.

Theorem 1.7 *Let X be a compact metric space and let $\mathcal{A}$ be a collection of continuous complex-valued functions on X with the following properties:*
(1) If $f, g \in \mathcal{A}$ and $c \in \mathbb{C}$, then $f + g$, fg, and cf are in $\mathcal{A}$;
(2) If $f \in \mathcal{A}$, then $\overline{f} \in \mathcal{A}$, where $\overline{f}$ is the complex conjugate of f;
(3) If $x \in X$, there exists $f \in \mathcal{A}$ such that $f(x) \neq 0$;
(4) If $x, y \in X$ with $x \neq y$, there exists $f \in \mathcal{A}$ such that $f(x) \neq f(y)$. Then the closure of $\mathcal{A}$ with respect to the supremum norm is the collection of continuous complex-valued functions on X.

The conclusion can be rephrased as saying that given f continuous on X and $\varepsilon > 0$, there exists $g \in \mathcal{A}$ such that

$$\sup_{x \in X} |f(x) - g(x)| < \varepsilon.$$

When (3) holds, $\mathcal{A}$ is said to *vanish at no point* of X. When (4) holds, $\mathcal{A}$ is said to *separate points.* In (3) and (4), the function f depends on x and on x and y, resp.

For a proof of Theorem 1.7 see [1] or [3].

Chapter 2

Families of sets

2.1 Algebras and σ-algebras

When we turn to constructing measures in Chapter 4, we will see that we cannot in general define the measure of an arbitrary set. We will have to restrict the class of sets we consider. The class of sets that we will want to use are σ-algebras (read "sigma algebras").

Let X be a set.

Definition 2.1 An *algebra* is a collection $\mathcal{A}$ of subsets of X such that
(1) $\emptyset \in \mathcal{A}$ and $X \in \mathcal{A}$;
(2) if $A \in \mathcal{A}$, then $A^c \in \mathcal{A}$;
(3) if $A_1, \dots, A_n \in \mathcal{A}$, then $\cup_{i=1}^n A_i$ and $\cap_{i=1}^n A_i$ are in $\mathcal{A}$.

$\mathcal{A}$ is a σ-*algebra* if in addition
(4) whenever $A_1, A_2, \dots$ are in $\mathcal{A}$, then $\cup_{i=1}^\infty A_i$ and $\cap_{i=1}^\infty A_i$ are in $\mathcal{A}$.

In (4) we allow countable unions and intersections only; we do not allow uncountable unions and intersections. Since $\cap_{i=1}^\infty A_i = (\cup_{i=1}^\infty A_i^c)^c$, the requirement that $\cap_{i=1}^\infty A_i$ be in $\mathcal{A}$ is redundant.

The pair $(X, \mathcal{A})$ is called a *measurable space*. A set A is *measurable* if $A \in \mathcal{A}$.

Example 2.2 Let $X = \mathbb{R}$, the set of real numbers, and let $\mathcal{A}$ be the collection of all subsets of $\mathbb{R}$. Then $\mathcal{A}$ is a σ-algebra.

Example 2.3 Let $X = \mathbb{R}$ and let

$$\mathcal{A} = \{A \subset \mathbb{R} : A \text{ is countable or } A^c \text{ is countable}\}.$$

Verifying parts (1) and (2) of the definition is easy. Suppose $A_1, A_2, \ldots$ are each in $\mathcal{A}$. If each of the A_i are countable, then $\cup_i A_i$ is countable, and so is in $\mathcal{A}$. If $A_{i_0}^c$ is countable for some i_0, then

$$(\cup A_i)^c = \cap_i A_i^c \subset A_{i_0}^c$$

is countable, and again $\cup_i A_i$ is in $\mathcal{A}$. Since $\cap A_i = (\cup_i A_i^c)^c$, then the countable intersection of sets in $\mathcal{A}$ is again in $\mathcal{A}$.

Example 2.4 Let $X = [0,1]$ and let $\mathcal{A} = \{\emptyset, X, [0, \frac{1}{2}], (\frac{1}{2}, 1]\}$. Then $\mathcal{A}$ is a σ-algebra.

Example 2.5 Let $X = \{1,2,3\}$ and let $\mathcal{A} = \{X, \emptyset, \{1\}, \{2,3\}\}$. Then $\mathcal{A}$ is a σ-algebra.

Example 2.6 Let $X = [0,1]$, and let $B_1, \ldots, B_8$ be subsets of X which are pairwise disjoint and whose union is all of X. Let $\mathcal{A}$ be the collection of all finite unions of the B_i's as well as the empty set. (Thus $\mathcal{A}$ consists of 2^8 elements.) Then $\mathcal{A}$ is a σ-algebra.

Lemma 2.7 *If $\mathcal{A}_\alpha$ is a σ-algebra for each α in some index set I, then $\cap_{\alpha \in I} \mathcal{A}_\alpha$ is a σ-algebra.*

Proof. This follows immediately from the definition. □

If we have a collection $\mathcal{C}$ of subsets of X, define

$$\sigma(\mathcal{C}) = \cap\{\mathcal{A}_\alpha : \mathcal{A}_\alpha \text{ is a } \sigma\text{-algebra}, \mathcal{C} \subset \mathcal{A}\},$$

the intersection of all σ-algebras containing $\mathcal{C}$. Since there is at least one σ-algebra containing $\mathcal{C}$, namely, the one consisting of all subsets of X, we are never taking the intersection over an empty class of σ-algebras. In view of Lemma 2.7, $\mathcal{C}$ is a σ-algebra. We call $\sigma(\mathcal{C})$ the *σ-algebra generated by the collection $\mathcal{C}$*, or say that

$\mathcal{C}$ generates the σ-algebra $\sigma(\mathcal{C})$. It is clear that if $\mathcal{C}_1 \subset \mathcal{C}_2$, then $\sigma(\mathcal{C}_1) \subset \sigma(\mathcal{C}_2)$. Since $\sigma(\mathcal{C})$ is a σ-algebra, then $\sigma(\sigma(\mathcal{C})) = \sigma(\mathcal{C})$.

If X has some additional structure, say, it is a metric space, then we can talk about open sets. If $\mathcal{G}$ is the collection of open subsets of X, then we call $\sigma(\mathcal{G})$ the *Borel σ-algebra* on X, and this is often denoted $\mathcal{B}$. Elements of $\mathcal{B}$ are called *Borel sets* and are said to be *Borel measurable*. We will see later that when X is the real line, $\mathcal{B}$ is *not* equal to the collection of all subsets of X.

We end this section with the following proposition.

Proposition 2.8 *If $X = \mathbb{R}$, then the Borel σ-algebra $\mathcal{B}$ is generated by each of the following collection of sets:*
(1) $\mathcal{C}_1 = \{(a,b) : a, b \in \mathbb{R}\}$;
(2) $\mathcal{C}_2 = \{[a,b] : a, b \in \mathbb{R}\}$;
(3) $\mathcal{C}_3 = \{(a,b] : a, b \in \mathbb{R}\}$;
(4) $\mathcal{C}_4 = \{(a,\infty) : a \in \mathbb{R}\}$.

Proof. (1) Let $\mathcal{G}$ be the collection of open sets. By definition, $\sigma(\mathcal{G})$ is the Borel σ-algebra. Since every element of $\mathcal{C}_1$ is open, then $\mathcal{C}_1 \subset \mathcal{G}$, and consequently $\sigma(\mathcal{C}_1) \subset \sigma(\mathcal{G}) = \mathcal{B}$.

To get the reverse inclusion, if G is open, it is the countable union of open intervals by Proposition 1.5. Every finite open interval is in $\mathcal{C}_1$. Since $(a,\infty) = \cup_{n=1}^\infty (a, a+n)$, then $(a,\infty) \in \sigma(\mathcal{C}_1)$ if $a \in \mathbb{R}$ and similarly $(-\infty, a) \in \sigma(\mathcal{C}_1)$ if $a \in \mathbb{R}$. Hence if G is open, then $G \in \sigma(\mathcal{C}_1)$. This says $\mathcal{G} \subset \sigma(\mathcal{C}_1)$, and then $\mathcal{B} = \sigma(\mathcal{G}) \subset \sigma(\sigma(\mathcal{C}_1)) = \sigma(\mathcal{C}_1)$.

(2) If $[a,b] \in \mathcal{C}_2$, then $[a,b] = \cap_{n=1}^\infty (a - \frac{1}{n}, b + \frac{1}{n}) \in \sigma(\mathcal{G})$. Therefore $\mathcal{C}_2 \subset \sigma(\mathcal{G})$, and hence $\sigma(\mathcal{C}_2) \subset \sigma(\sigma(\mathcal{G})) = \sigma(\mathcal{G}) = \mathcal{B}$.

If $(a,b) \in \mathcal{C}_1$, choose $n_0 \ge 2/(b-a)$ and note

$$(a,b) = \cup_{n=n_0}^\infty \Big[a + \frac{1}{n}, b - \frac{1}{n}\Big] \in \sigma(\mathcal{C}_2).$$

Therefore $\mathcal{C}_1 \subset \sigma(\mathcal{C}_2)$, from which it follows that $\mathcal{B} = \sigma(\mathcal{C}_1) \subset \sigma(\sigma(\mathcal{C}_2)) = \sigma(\mathcal{C}_2)$.

(3) Using $(a,b] = \cap_{n=1}^\infty (a, b + \frac{1}{n})$, we see that $\mathcal{C}_3 \subset \sigma(\mathcal{C}_1)$, and as above we conclude that $\sigma(\mathcal{C}_3) \subset \sigma(\mathcal{C}_1) = \mathcal{B}$. Using $(a,b) = \cup_{n=n_0}^\infty (a, b - \frac{1}{n}]$, provided n_0 is taken large enough, $\mathcal{C}_1 \subset \sigma(\mathcal{C}_3)$, and as above we argue that $\mathcal{B} = \sigma(\mathcal{C}_1) \subset \sigma(\mathcal{C}_3)$.

(4) Because $(a,b] = (a,\infty) - (b,\infty)$, then $\mathcal{C}_1 \subset \sigma(\mathcal{C}_4)$. Since $(a,\infty) = \cup_{n=1}^\infty (a, a+n]$, then $\mathcal{C}_4 \subset \sigma(\mathcal{C}_3)$. As above, this is enough to imply that $\sigma(\mathcal{C}_4) = \mathcal{B}$. $\square$

2.2 The monotone class theorem

This section will be used in Chapter 11.

Definition 2.9 A monotone class is a collection of subsets $\mathcal{M}$ of X such that
(1) if $A_i \uparrow A$ and each $A_i \in \mathcal{M}$, then $A \in \mathcal{M}$;
(2) if $A_i \downarrow A$ and each $A_i \in \mathcal{M}$, then $A \in \mathcal{M}$.

The intersection of monotone classes is a monotone class, and the intersection of all monotone classes containing a given collection of sets is the smallest monotone class containing that collection.

The next theorem, the *monotone class theorem*, is rather technical, but very useful.

Theorem 2.10 *Suppose $\mathcal{A}_0$ is an algebra, $\mathcal{A}$ is the smallest σ-algebra containing $\mathcal{A}_0$, and $\mathcal{M}$ is the smallest monotone class containing $\mathcal{A}_0$. Then $\mathcal{M} = \mathcal{A}$.*

Proof. A σ-algebra is clearly a monotone class, so $\mathcal{M} \subset \mathcal{A}$. We must show $\mathcal{A} \subset \mathcal{M}$.

Let $\mathcal{N}_1 = \{A \in \mathcal{M} : A^c \in \mathcal{M}\}$. Note $\mathcal{N}_1$ is contained in $\mathcal{M}$ and contains $\mathcal{A}_0$. If $A_i \uparrow A$ and each $A_i \in \mathcal{N}_1$, then each $A_i^c \in \mathcal{M}$ and $A_i^c \downarrow A^c$. Since $\mathcal{M}$ is a monotone class, $A^c \in \mathcal{M}$, and so $A \in \mathcal{N}_1$. Similarly, if $A_i \downarrow A$ and each $A_i \in \mathcal{N}_1$, then $A \in \mathcal{N}_1$. Therefore $\mathcal{N}_1$ is a monotone class. Hence $\mathcal{N}_1 = \mathcal{M}$, and we conclude $\mathcal{M}$ is closed under the operation of taking complements.

Let $\mathcal{N}_2 = \{A \in \mathcal{M} : A \cap B \in \mathcal{M}$ for all $B \in \mathcal{A}_0\}$. Note the following: $\mathcal{N}_2$ is contained in $\mathcal{M}$ and $\mathcal{N}_2$ contains $\mathcal{A}_0$ because $\mathcal{A}_0$ is an algebra. If $A_i \uparrow A$, each $A_i \in \mathcal{N}_2$, and $B \in \mathcal{A}_0$, then $A \cap B = \cup_{i=1}^\infty (A_i \cap B)$. Because $\mathcal{M}$ is a monotone class, $A \cap B \in \mathcal{M}$, which implies $A \in \mathcal{N}_2$. We use a similar argument when $A_i \downarrow A$.

Therefore $\mathcal{N}_2$ is a monotone class, and we conclude $\mathcal{N}_2 = \mathcal{M}$. In other words, if $B \in \mathcal{A}_0$ and $A \in \mathcal{M}$, then $A \cap B \in \mathcal{M}$.

Let $\mathcal{N}_3 = \{A \in \mathcal{M} : A \cap B \in \mathcal{M} \text{ for all } B \in \mathcal{M}\}$. As in the preceding paragraph, $\mathcal{N}_3$ is a monotone class contained in $\mathcal{M}$. By the last sentence of the preceding paragraph, $\mathcal{N}_3$ contains $\mathcal{A}_0$. Hence $\mathcal{N}_3 = \mathcal{M}$.

We thus have that $\mathcal{M}$ is a monotone class closed under the operations of taking complements and taking intersections. This shows $\mathcal{M}$ is a σ-algebra, and so $\mathcal{A} \subset \mathcal{M}$. □

2.3 Exercises

Exercise 2.1 Find an example of a set X and a monotone class $\mathcal{M}$ consisting of subsets of X such that $\emptyset \in \mathcal{M}$, $X \in \mathcal{M}$, but $\mathcal{M}$ is not a σ-algebra.

Exercise 2.2 Find an example of a set X and two σ-algebras $\mathcal{A}_1$ and $\mathcal{A}_2$, each consisting of subsets of X, such that $\mathcal{A}_1 \cup \mathcal{A}_2$ is not a σ-algebra.

Exercise 2.3 Suppose $\mathcal{A}_1 \subset \mathcal{A}_2 \subset \cdots$ are σ-algebras consisting of subsets of a set X. Is $\cup_{i=1}^\infty \mathcal{A}_i$ necessarily a σ-algebra? If not, give a counterexample.

Exercise 2.4 Suppose $\mathcal{M}_1 \subset \mathcal{M}_2 \subset \cdots$ are monotone classes. Let $\mathcal{M} = \cup_{n=1}^\infty \mathcal{M}_n$. Suppose $A_j \uparrow A$ and each $A_j \in \mathcal{M}$. Is A necessarily in $\mathcal{M}$? If not, give a counterexample.

Exercise 2.5 Let $(Y, \mathcal{A})$ be a measurable space and let f map X onto Y, but do not assume that f is one-to-one. Define $\mathcal{B} = \{f^{-1}(A) : A \in \mathcal{A}\}$. Prove that $\mathcal{B}$ is a σ-algebra of subsets of X.

Exercise 2.6 Suppose $\mathcal{A}$ is a σ-algebra with the property that whenever $A \in \mathcal{A}$, there exist $B, C \in \mathcal{A}$ with $B \cap C = \emptyset$, $B \cup C = A$, and neither B nor C is empty. Prove that $\mathcal{A}$ is uncountable.

Exercise 2.7 Suppose $\mathcal{F}$ is a collection of real-valued functions on X such that the constant functions are in $\mathcal{F}$ and $f + g$, fg, and cf

are in $\mathcal{F}$ whenever $f, g \in \mathcal{F}$ and $c \in \mathbb{R}$. Suppose $f \in \mathcal{F}$ whenever $f_n \to f$ and each $f_n \in \mathcal{F}$. Define the function

$$\chi_A(x) = \begin{cases} 1, & x \in A; \\ 0, & x \notin A. \end{cases}$$

Prove that $\mathcal{A} = \{A \subset X : \chi_A \in \mathcal{F}\}$ is a σ-algebra.

Exercise 2.8 Does there exist a σ-algebra which has countably many elements, but not finitely many?

Chapter 3

Measures

In this chapter we give the definition of a measure, some examples, and some of the simplest properties of measures. Constructing measures is often quite difficult and we defer the construction of the most important one, Lebesgue measure, until Chapter 4

3.1 Definitions and examples

Definition 3.1 Let X be a set and $\mathcal{A}$ a σ-algebra consisting of subsets of X. A *measure* on $(X, \mathcal{A})$ is a function $\mu : \mathcal{A} \to [0, \infty]$ such that
(1) $\mu(\emptyset) = 0$;
(2) if $A_i \in \mathcal{A}$, $i = 1, 2, \ldots$, are pairwise disjoint, then

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

Saying the A_i are pairwise disjoint means that $A_i \cap A_j = \emptyset$ if $i \neq j$.

Definition 3.1(2) is known as *countable additivity*. We say a set function is *finitely additive* if $\mu(\cup_{i=1}^{n} A_i) = \sum_{i=1}^{n} \mu(A_i)$ whenever $A_1, \ldots, A_n$ are in $\mathcal{A}$ and are pairwise disjoint.

The triple $(X, \mathcal{A}, \mu)$ is called a *measure space*.

Example 3.2 Let X be any set, $\mathcal{A}$ the collection of all subsets of X, and $\mu(A)$ the number of elements in A. μ is called *counting measure.*

Example 3.3 Let $X = \mathbb{R}$, $\mathcal{A}$ the collection of all subsets of $\mathbb{R}$, $x_1, x_2, \ldots \in \mathbb{R}$, and $a_1, a_2, \ldots \geq 0$. Set

$$\mu(A) = \sum_{\{i : x_i \in A\}} a_i.$$

Counting measure is a particular case of this is if $x_i = i$ and all the $a_i = 1$.

Example 3.4 Let $\delta_x(A) = 1$ if $x \in A$ and 0 otherwise. This measure is called *point mass* at x.

Proposition 3.5 *The following hold:*
(1) If $A, B \in \mathcal{A}$ with $A \subset B$, then $\mu(A) \leq \mu(B)$.
(2) If $A_i \in \mathcal{A}$ and $A = \cup_{i=1}^\infty A_i$, then $\mu(A) \leq \sum_{i=1}^\infty \mu(A_i)$.
(3) Suppose $A_i \in \mathcal{A}$ and $A_i \uparrow A$. Then $\mu(A) = \lim_{n\to\infty} \mu(A_n)$.
(4) Suppose $A_i \in \mathcal{A}$ and $A_i \downarrow A$. If $\mu(A_1) < \infty$, then we have $\mu(A) = \lim_{n\to\infty} \mu(A_n)$.

Proof. (1) Let $A_1 = A$, $A_2 = B - A$, and $A_3 = A_4 = \cdots = \emptyset$. Now use part (2) of the definition of measure to write

$$\mu(B) = \mu(A) + \mu(B - A) + 0 + 0 + \cdots \geq \mu(A).$$

(2) Let $B_1 = A_1$, $B_2 = A_2 - A_1$, $B_3 = A_3 - (A_1 \cup A_2)$, $B_4 = A_4 - (A_1 \cup A_2 \cup A_3)$, and in general $B_i = A_i - (\cup_{j=1}^{i-1} A_j)$. The B_i are pairwise disjoint, $B_i \subset A_i$ for each i, and $\cup_{i=1}^\infty B_i = \cup_{i=1}^\infty A_i$. Hence

$$\mu(A) = \mu(\cup_{i=1}^\infty B_i) = \sum_{i=1}^\infty \mu(B_i) \leq \sum_{i=1}^\infty \mu(A_i).$$

(3) Define the B_i as in (2). Since $\cup_{i=1}^n B_i = \cup_{i=1}^n A_i$, then

$$\begin{aligned}\mu(A) = \mu(\cup_{i=1}^\infty A_i) = \mu(\cup_{i=1}^\infty B_i) &= \sum_{i=1}^\infty \mu(B_i)\\ &= \lim_{n\to\infty} \sum_{i=1}^n \mu(B_i) = \lim_{n\to\infty} \mu(\cup_{i=1}^n B_i) = \lim_{n\to\infty} \mu(\cup_{i=1}^n A_i).\end{aligned}$$

(4) Apply (3) to the sets $A_1 - A_i$, $i = 1, 2, \ldots$. The sets $A_1 - A_i$ increase to $A_1 - A$, and so

$$\begin{aligned}\mu(A_1) - \mu(A) = \mu(A_1 - A) &= \lim_{n\to\infty} \mu(A_1 - A_n)\\ &= \lim_{n\to\infty} [\mu(A_1) - \mu(A_n)].\end{aligned}$$

Now subtract $\mu(A_1)$ from both sides and then multiply both sides by -1. □

Example 3.6 To see that $\mu(A_1) < \infty$ is necessary in Proposition 3.5, let X be the positive integers, μ counting measure, and $A_i = \{i, i+1, \ldots\}$. Then the A_i decrease, $\mu(A_i) = \infty$ for all i, but $\mu(\cap_i A_i) = \mu(\emptyset) = 0$.

Definition 3.7 A measure μ is a *finite measure* if $\mu(X) < \infty$. A measure μ is *σ-finite* if there exist sets $E_i \in \mathcal{A}$ for $i = 1, 2, \ldots$ such that $\mu(E_i) < \infty$ for each i and $X = \cup_{i=1}^\infty E_i$. If μ is a finite measure, then $(X, \mathcal{A}, \mu)$ is called a finite measure space, and similarly, if μ is a σ-finite measure, then $(X, \mathcal{A}, \mu)$ is called a σ-finite measure space.

If we let $F_n = \cup_{i=1}^n E_i$, then $\mu(F_n) < \infty$ for each n and $F_n \uparrow X$. Therefore there is no loss of generality in supposing the sets E_i in Definition 3.7 are increasing.

Let $(X, \mathcal{A}, \mu)$ be a measure space. A subset $A \subset X$ is a *null set* if there exists a set $B \in \mathcal{A}$ with $A \subset B$ and $\mu(B) = 0$. We do not require A to be in $\mathcal{A}$. If $\mathcal{A}$ contains all the null sets, then $(X, \mathcal{A}, \mu)$ is said to be a *complete measure space*. The *completion* of $\mathcal{A}$ is the smallest σ-algebra $\overline{\mathcal{A}}$ containing $\mathcal{A}$ such that $(X, \overline{\mathcal{A}}, \mu)$ is complete. Sometimes one just says that $\mathcal{A}$ is complete or that μ is complete when $(X, \mathcal{A}, \mu)$ is complete.

A *probability* or *probability measure* is a measure μ such that $\mu(X) = 1$. In this case we usually write $(\Omega, \mathcal{F}, \mathbb{P})$ instead of $(X, \mathcal{A}, \mu)$, and $\mathcal{F}$ is called a *σ-field*, which is the same thing as a σ-algebra.

3.2 Exercises

Exercise 3.1 Suppose $(X, \mathcal{A})$ is a measurable space and μ is a non-negative set function that is finitely additive and such that $\mu(\emptyset) = 0$. Suppose that whenever A_i is an increasing sequence of sets in $\mathcal{A}$, then $\mu(\cup_i A_i) = \lim_{i\to\infty} \mu(A_i)$. Show that μ is a measure.

Exercise 3.2 Let X be an uncountable set and let $\mathcal{A}$ be the collection of subsets A of X such that either A or A^c is countable. Define $\mu(A) = 0$ if A is countable and $\mu(A) = 1$ if A is uncountable. Prove that μ is a measure.

Exercise 3.3 Suppose $(X, \mathcal{A}, \mu)$ is a measure space and $A, B \in \mathcal{A}$. Prove that

$$\mu(A) + \mu(B) = \mu(A \cup B) + \mu(A \cap B).$$

Exercise 3.4 3. Prove that if $\mu_1, \mu_2, \ldots$ are measures on a measurable space and $a_1, a_2, \ldots \in [0, \infty)$, then $\sum_{n=1}^\infty a_n \mu_n$ is also a measure.

Exercise 3.5 3. Prove that if $(X, \mathcal{A}, \mu)$ is a measure space, $B \in \mathcal{A}$, and we define $\nu(A) = \mu(A \cap B)$ for $A \in \mathcal{A}$, then ν is a measure.

Exercise 3.6 Suppose $\mu_1, \mu_2, \ldots$ are measures on a measurable space $(X, \mathcal{A})$ and $\mu_n(A) \uparrow$ for each $A \in \mathcal{A}$. Define

$$\mu(A) = \lim_{n\to\infty} \mu_n(A).$$

Is μ necessarily a measure? If not, give a counterexample. What if $\mu_n(A) \downarrow$ for each $A \in \mathcal{A}$ and $\mu_1(X) < \infty$?

Exercise 3.7 Let $(X, \mathcal{A}, \mu)$ be a measure space, let $\mathcal{N}$ be the collection of null sets with respect to $\mathcal{A}$, and let $\mathcal{B} = \sigma(\mathcal{A} \cup \mathcal{N})$. Show that $(X, \mathcal{B}, \mu)$ is complete and is the completion of $(X, \mathcal{A}, \mu)$.

Exercise 3.8 Suppose X is the set of real numbers, $\mathcal{B}$ is the Borel σ-algebra, and m and n are two measures on $(X, \mathcal{B})$ such that $m((a, b)) = n((a, b))$ whenever $a < b$. Prove that $m(A) = n(A)$ whenever $A \in \mathcal{B}$.

Exercise 3.9 Suppose $(X, \mathcal{A})$ is a measurable space and $\mathcal{C}$ is an arbitrary subset of $\mathcal{A}$. Suppose m and n are two σ-finite measures on $(X, \mathcal{A})$ such that $m(A) = n(A)$ for all $A \in \mathcal{C}$. Is it true that $m(A) = n(A)$ for all $A \in \sigma(\mathcal{C})$? What if m and n are finite measures?

Chapter 4

Construction of measures

Our goal in this chapter is give a method for constructing measures. This is a complicated procedure, and involves the concept of outer measure, which we introduce in Section 4.1.

Our most important example will be one-dimensional Lebesgue measure, which we consider in Section 4.2. Further results and some examples related to Lebesgue measure are given in Section 4.3.

One cannot define the Lebesgue measure of every subset of the reals. This is shown in Section 4.4.

The methods used to construct measures via outer measures have other applications besides the construction of Lebesgue measure. The Carathéodory extension theorem is a tool developed in Section 4.5 that can be used in constructing measures.

Let us present some of the ideas used in the construction of Lebesgue measure on the line. We want the measure m of an open interval to be the length of the interval. Since every open subset of the reals is the countable union of disjoint open intervals (see Proposition 1.5), if $G = \cup_{i=1}^\infty (a_i, b_i)$, where the intervals (a_i, b_i) are pairwise disjoint, we must have

$$m(G) = \sum_{i=1}^\infty (b_i - a_i).$$

We then set

$$m(E) = \inf\{m(G) : G \text{ open}, A \subset G\}$$

for arbitrary subsets $E \subset \mathbb{R}$. The difficulty is that m is not a measure on the σ-algebra consisting of all subsets of the reals; this is proved in Section 4.4. We resolve this by considering a strictly smaller σ-algebra. This is the essential idea behind the construction of Lebesgue measure, but it is technically simpler to work with intervals of the form $(a, b]$ rather than open intervals.

4.1 Outer measures

We begin with the notion of outer measure.

Definition 4.1 Let X be a set. An *outer measure* is a function μ^* defined on the collection of all subsets of X satisfying
(1) $\mu^*(\emptyset) = 0$;
(2) if $A \subset B$, then $\mu^*(A) \le \mu^*(B)$;
(3) $\mu^*(\cup_{i=1}^\infty A_i) \le \sum_{i=1}^\infty \mu^*(A_i)$ whenever $A_1, A_2, \ldots$ are subsets of X.

A set N is a *null set* with respect to μ^* if $\mu^*(N) = 0$.

A common way to generate outer measures is as follows.

Proposition 4.2 *Suppose $\mathcal{C}$ is a collection of subsets of X such that $\emptyset$ and X are both in $\mathcal{C}$. Suppose $\ell : \mathcal{C} \to [0, \infty]$ with $\ell(\emptyset) = 0$. Define*

$$\mu^*(E) = \inf\Big\{\sum_{i=1}^\infty \ell(A_i) : A_i \in \mathcal{C} \text{ for each } i \text{ and } E \subset \cup_{i=1}^\infty A_i\Big\}. \tag{4.1}$$

Then μ^* is an outer measure.

Proof. (1) and (2) of the definition of outer measure are obvious. To prove (3), let $A_1, A_2, \ldots$ be subsets of X and let $\varepsilon > 0$. For each i there exist $C_{i1}, C_{i2}, \ldots \in \mathcal{C}$ such that $A_i \subset \cup_{j=1}^\infty C_{ij}$ and

$\sum_j \ell(C_{ij}) \le \mu^*(A_i) + \varepsilon/2^i$. Then $\cup_{i=1}^\infty A_i \subset \cup_i \cup_j C_{ij}$ and

$$\begin{aligned}\mu^*(\cup_{i=1}^\infty A_i) &\le \sum_{i,j} \ell(C_{ij}) = \sum_i \Big(\sum_j \ell(C_{ij}) \Big)\\ &\le \sum_{i=1}^\infty \mu^*(A_i) + \sum_{i=1}^\infty \varepsilon/2^i\\ &= \sum_{i=1}^\infty \mu^*(A_i) + \varepsilon.\end{aligned}$$

Since ε is arbitrary, $\mu^*(\cup_{i=1}^\infty A_i) \le \sum_{i=1}^\infty \mu^*(A_i)$. □

Example 4.3 Let $X = \mathbb{R}$ and let $\mathcal{C}$ be the collection of intervals of the form $(a, b]$, that is, intervals that are open on the left and closed on the right. Let $\ell(I) = b - a$ if $I = (a, b]$. Define μ^* by (4.1). Proposition 4.2 shows that μ^* is an outer measure, but we will see in Section 4.4 that μ^* is not a measure on the collection of all subsets of $\mathbb{R}$. We will also see, however, that if we restrict μ^* to a σ-algebra $\mathcal{L}$ which is strictly smaller than the collection of all subsets of $\mathbb{R}$, then μ^* will be a measure on $\mathcal{L}$. That measure is what is known as *Lebesgue measure*. The σ-algebra $\mathcal{L}$ is called the *Lebesgue σ-algebra*.

Example 4.4 Let $X = \mathbb{R}$ and let $\mathcal{C}$ be the collection of intervals of the form $(a, b]$ as in the previous example. Let $\alpha : \mathbb{R} \to \mathbb{R}$ be an increasing right continuous function on $\mathbb{R}$. Thus $\alpha(x) = \lim_{y\to x+} \alpha(y)$ for each x and $\alpha(x) \le \alpha(y)$ if $x < y$. Let $\ell(I) = \alpha(b) - \alpha(a)$ if $I = (a, b]$. Again define μ^* by (4.1). Again Proposition 4.2 shows that μ^* is an outer measure. Restricting μ^* to a smaller σ-algebra gives us what is known as *Lebesgue-Stieltjes measure* corresponding to α. The special case where $\alpha(x) = x$ for all x is Lebesgue measure.

In general we need to restrict μ^* to a strictly smaller σ-algebra than the collection of all subsets of $\mathbb{R}$, but not always. For example, if $\alpha(x) = 0$ for $x < 0$ and 1 for $x \ge 1$, then the corresponding Lebesgue-Stieltjes measure is point mass at 0 (defined in Example 3.4), and the corresponding σ-algebra is the collection of all subsets of $\mathbb{R}$.

Definition 4.5 Let μ^* be an outer measure. A set $A \subset X$ is *μ^*-measurable* if

$$\mu^*(E) = \mu^*(E \cap A) + \mu^*(E \cap A^c) \tag{4.2}$$

for all $E \subset X$.

Theorem 4.6 *If μ^* is an outer measure on X, then the collection $\mathcal{A}$ of μ^*-measurable sets is a σ-algebra. If μ is the restriction of μ^* to $\mathcal{A}$, then μ is a measure. Moreover, $\mathcal{A}$ contains all the null sets.*

This is sometimes known as *Carathéodory's theorem*, but do not confuse this with the Carathéodory extension theorem in Section 4.5.

Proof. By Definition 4.1,

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

for all $E \subset X$. Thus to check (4.2) it is enough to show

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

This will be trivial in the case $\mu^*(E) = \infty$.

Step 1. First we show $\mathcal{A}$ is an algebra. If $A \in \mathcal{A}$, then $A^c \in \mathcal{A}$ by symmetry and the definition of $\mathcal{A}$. Suppose $A, B \in \mathcal{A}$ and $E \subset X$. Then

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

The second equality follows from the definition of $\mathcal{A}$ with E first replaced by $E \cap A$ and then by $E \cap A^c$. The first three summands on the right of the second equals sign have a sum greater than or equal to $\mu^*(E \cap (A \cup B))$ because $A \cup B \subset (A \cap B) \cup (A \cap B^c) \cup (A^c \cap B)$. Since $A^c \cap B^c = (A \cup B)^c$, then

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

which shows $A \cup B \in \mathcal{A}$. Therefore $\mathcal{A}$ is an algebra.

Step 2. Next we show $\mathcal{A}$ is a σ-algebra. Let A_i be pairwise disjoint sets in $\mathcal{A}$, let $B_n = \cup_{i=1}^n A_i$, and $B = \cup_{i=1}^\infty A_i$. If $E \subset X$,

$$\begin{aligned}\mu^*(E \cap B_n) &= \mu^*(E \cap B_n \cap A_n) + \mu^*(E \cap B_n \cap A_n^c)\\ &= \mu^*(E \cap A_n) + \mu^*(E \cap B_{n-1}).\end{aligned}$$

Similarly, $\mu^*(E \cap B_{n-1}) = \mu^*(E \cap A_{n-1}) + \mu^*(E \cap B_{n-2})$, and continuing, we obtain

$$\mu^*(E \cap B_n) \ge \sum_{i=1}^n \mu^*(E \cap A_i).$$

Since $B_n \in \mathcal{A}$, then

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

Let $n \to \infty$. Recalling that μ^* is an outer measure,

$$\begin{aligned}\mu^*(E) &\ge \sum_{i=1}^\infty \mu^*(E \cap A_i) + \mu^*(E \cap B^c) \qquad (4.3)\\ &\ge \mu^*(\cup_{i=1}^\infty (E \cap A_i)) + \mu^*(E \cap B^c)\\ &= \mu^*(E \cap B) + \mu^*(E \cap B^c)\\ &\ge \mu^*(E).\end{aligned}$$

This shows $B \in \mathcal{A}$.

Now if $C_1, C_2, \ldots$ are sets in $\mathcal{A}$, let $A_1 = C_1$, $A_2 = C_2 - A_1$, $A_3 = C_3 - (A_1 \cup A_2)$, and in general $A_i = C_i - (\cup_{j=1}^{i-1} A_j)$. Since each $C_i \in \mathcal{A}$ and $\mathcal{A}$ is an algebra, then $A_i = C_i \cap C_{i-1}^c \in \mathcal{A}$. The A_i are pairwise disjoint, so from the previous paragraph,

$$\cup_{i=1}^\infty C_i = \cup_{i=1}^\infty A_i \in \mathcal{A}.$$

Also, $\cap_{i=1}^\infty C_i = (\cup_{i=1}^\infty C_i^c)^c \in \mathcal{A}$, and therefore $\mathcal{A}$ is a σ-algebra.

Step 3. We now show μ^* restricted to $\mathcal{A}$ is a measure. The only way (4.3) can hold is if all the inequalities there are actually equalities, and in particular,

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

Taking $E = B$, we obtain

$$\mu^*(B) = \sum_{i=1}^\infty \mu^*(A_i),$$

or μ^* is countably additive on $\mathcal{A}$.

Step 4. Finally, if $\mu^*(A) = 0$ and $E \subset X$, then

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

which shows $\mathcal{A}$ contains all the null sets. □

4.2 Lebesgue-Stieltjes measures

Let $X = \mathbb{R}$ and let $\mathcal{C}$ be the collection of intervals of the form $(a, b]$, that is, intervals that are open on the left and closed on the right. Let $\alpha(x)$ be an increasing right continuous function. This means that $\alpha(x) \le \alpha(y)$ if $x < y$ and $\lim_{z \to x+} \alpha(z) = \alpha(x)$ for all x. We do not require α to be strictly increasing. Define

$$\ell((a, b]) = \alpha(b) - \alpha(a).$$

Define

$$m^*(E) = \inf \Big\{ \sum_{i=1}^\infty \ell(A_i) : A_i \in \mathcal{C} \text{ for each } i \text{ and } E \subset \cup_{i=1}^\infty A_i \Big\}.$$

(In this book we usually use m instead of μ when we are talking about Lebesgue-Stieltjes measures.) We use Proposition 4.2 to tell us that m^* is an outer measure. We then use Theorem 4.6 to show that m^* is a measure on the collection of m^*-measurable sets. Note that if K and L are adjacent intervals, that is, if $K = (a, b]$ and $L = (b, c]$, then $K \cup L = (a, c]$ and

$$\begin{aligned} \ell(K) + \ell(L) &= [\alpha(b) - \alpha(a)] + [\alpha(c) - \alpha(b)] \\ &= \alpha(c) - \alpha(a) = \ell(K \cup L) \end{aligned} \tag{4.4}$$

by the definition of ℓ.

The next step in the construction of Lebesgue-Stieltjes measure corresponding to α is the following.

Proposition 4.7 *Every set in the Borel σ-algebra on $\mathbb{R}$ is m^*-measurable.*

Proof. Since the collection of m^*-measurable sets is a σ-algebra, it suffices to show that every interval J of the form $(c, d]$ is m^*-measurable. Let E be any set with $m^*(E) < \infty$; we need to show

$$m^*(E) \geq m^*(E \cap J) + m^*(E \cap J^c). \tag{4.5}$$

Choose $I_1, I_2, \ldots$, each of the form $(a_i, b_i]$, such that $E \subset \cup_i I_i$ and

$$m^*(E) \geq \sum_i [\alpha(b_i) - \alpha(a_i)] - \varepsilon.$$

Since $E \subset \cup_i I_i$, we have

$$m^*(E \cap J) \leq \sum_i m^*(I_i \cap J)$$

and

$$m^*(E \cap J^c) \leq \sum_i m^*(I_i \cap J^c).$$

Adding we have

$$m^*(E \cap J) + m^*(E \cap J^c) \leq \sum_i [m^*(I_i \cap J) + m^*(I_i \cap J^c)].$$

Now $I_i \cap J$ is an interval that is open on the left and closed on the right, and $I_i \cap J^c$ is the union of zero, one, or two such intervals, depending on the relative locations of I_i and J. Using (4.4) either zero, one, or two times, we see that

$$m^*(I_i \cap J) + m^*(I_i \cap J^c) = m^*(I_i).$$

Thus

$$m^*(E \cap J) + m^*(E \cap J^c) \leq \sum_i m^*(I_i) \leq m^*(E) + \varepsilon.$$

Since ε is arbitrary, this proves (4.5). □

After all this work, it would be upsetting if the measure of a half-open interval $(e, f]$ were not what it is supposed to be. Fortunately, everything is fine, due to Proposition 4.9. First we need the following lemma.

Lemma 4.8 *Let $J_k = (a_k, b_k)$, $k = 1, \ldots, n$, be a finite collection of finite open intervals covering a finite closed interval $[C, D]$. Then*

$$\sum_{k=1}^{n} [\alpha(b_k) - \alpha(a_k)] \geq \alpha(D) - \alpha(C). \tag{4.6}$$

Proof. Since $\{J_k\}$ is a cover of $[C, D]$, there exists at least one interval, say, J_{k_1}, such that $C \in J_{k_1}$. If J_{k_1} covers $[C, D]$, we stop. Otherwise, $b_{k_1} \leq D$, and there must be at least one interval, say, J_{k_2}, such that $b_{k_1} \in J_{k_2}$. If $[C, D] \subset J_{k_1} \cup J_{k_2}$, we stop. If not, then $b_{k_1} < b_{k_2} \leq D$, and there must be at least one interval, say, J_{k_3} that contains b_{k_2}. At each stage we choose J_{k_j} so that $b_{k_{j-1}} \in J_{k_j}$. We continue until we have covered $[C, D]$ with intervals $J_{k_1}, \ldots, J_{k_m}$. Since $\{J_k\}$ is a finite cover, we will stop for some $m \leq n$.

By our construction we have

$$a_{k_1} \leq C < b_{k_1}, \qquad a_{k_m} < D < b_{k_m},$$

and for $2 \leq j \leq m$,

$$a_{k_j} < b_{k_{j-1}} < b_{k_j}.$$

Then

$$\begin{aligned}
\alpha(D) - \alpha(C) &\leq \alpha(b_{k_m}) - \alpha(a_{k_1}) \\
&= [\alpha(b_{k_m}) - \alpha(b_{k_{m-1}})] + [\alpha(b_{k_{m-1}}) - \alpha(b_{k_{m-2}})] + \cdots \\
&\qquad + [\alpha(b_{k_2}) - \alpha(b_{k_1})] + [\alpha(b_{k_1}) - \alpha(a_{k_1})] \\
&\leq [\alpha(b_{k_m}) - \alpha(a_{k_m})] + [\alpha(b_{k_{m-1}}) - \alpha(a_{k_{m-1}})] + \cdots \\
&\qquad + [\alpha(b_{k_2}) - \alpha(a_{k_2})] + [\alpha(b_{k_1}) - \alpha(a_{k_1})].
\end{aligned}$$

Since $\{J_{k_1}, \ldots, J_{k_m}\} \subset \{J_1, \ldots, J_n\}$, this proves (4.6). □

Proposition 4.9 *If e and f are finite and $I = (e, f]$, then $m^*(I) = \ell(I)$.*

Proof. First we show $m^*(I) \leq \ell(I)$. This is easy. Let $A_1 = I$ and $A_2 = A_3 = \cdots = \emptyset$. Then $I \subset \cup_{i=1}^{\infty} A_i$, hence

$$m^*(I) \leq \sum_{i=1}^{\infty} \ell(A_i) = \ell(A_1) = \ell(I).$$

For the other direction, suppose $I \subset \cup_{i=1}^{\infty} A_i$, where $A_i = (c_i, d_i]$. Let $\varepsilon > 0$ and choose $C \in (e, f)$ such that $\alpha(C) - \alpha(e) < \varepsilon/2$. This is possible by the right continuity of α. Let $D = f$. For each i, choose $d_i' > d_i$ such that $\alpha(d_i') - \alpha(d_i) < \varepsilon/2^{i+1}$ and let $B_i = (c_i, d_i')$.

Then $[C, D]$ is compact and $\{B_i\}$ is an open cover for $[C, D]$. Use compactness to choose a finite subcover $\{J_1, \ldots, J_n\}$ of $\{B_i\}$. We now apply Lemma 4.8. We conclude that

$$\ell(I) \le \alpha(D) - \alpha(C) + \varepsilon/2 \le \sum_{k=1}^{n} (\alpha(d_k') - \alpha(c_k)) + \varepsilon/2 \le \sum_{i=1}^{\infty} \ell(A_i) + \varepsilon.$$

Taking the infimum over all countable collections $\{A_i\}$ that cover I, we obtain

$$\ell(I) \le m^*(I) + \varepsilon.$$

Since ε is arbitrary, $\ell(I) \le m^*(I)$. □

We now drop the asterisks from m^* and call m *Lebesgue-Stieltjes measure.* In the special case where $\alpha(x) = x$, m is *Lebesgue measure.* In the special case of Lebesgue measure, the collection of m^*-measurable sets is called the *Lebesgue σ-algebra.* A set is *Lebesgue measurable* if it is in the Lebesgue σ-algebra.

Given a measure μ on $\mathbb{R}$ such that $\mu(K) < \infty$ whenever K is compact, define $\alpha(x) = \mu((0, x])$ if $x \ge 0$ and $\alpha(x) = -\mu((x, 0])$ if $x < 0$. Then α is increasing, right continuous, and Exercise 4.1 asks you to show that μ is Lebesgue-Stieltjes measure corresponding to α.

4.3 Examples and related results

Example 4.10 Let m be Lebesgue measure. If $x \in \mathbb{R}$, then $\{x\}$ is a closed set and hence is Borel measurable. Moreover

$$m(\{x\}) = \lim_{n\to\infty} m((x - (1/n), x]) = \lim_{n\to\infty} [x - (x - (1/n))] = 0.$$

We then conclude

$$m([a, b]) = m((a, b]) + m(\{a\}) = b - a + 0 = b - a$$

and

$$m((a,b)) = m((a,b]) - m(\{b\}) = b - a - 0 = b - a.$$

Since σ-algebras are closed under the operation of countable unions, then countable sets are Borel measurable. Adding 0 to itself countably many times is still 0, so the Lebesgue measure of a countable set is 0.

However there are uncountable sets which have Lebesgue measure 0. See the next example.

Example 4.11 Recall from undergraduate analysis that the *Cantor set* is constructed as follows. Let F_0 be the interval $[0,1]$. We let F_1 be what remains if we remove the middle third, that is,

$$F_1 = F_0 - (\tfrac{1}{3}, \tfrac{2}{3}).$$

F_1 consists of two intervals of length $\frac{1}{3}$ each. We remove the middle third of each of these two intervals and let

$$F_2 = F_1 - [(\tfrac{1}{9}, \tfrac{2}{9}) \cup (\tfrac{7}{9}, \tfrac{8}{9})].$$

We continue removing middle thirds, and the Cantor set F is $\cap_n F_n$. Recall that the Cantor set is closed, uncountable, and every point is a limit point. Moreover, it contains no intervals.

The measure of F_1 is $2(\frac{1}{3})$, the measure of F_2 is $4(\frac{1}{9})$, and the measure of F_n is $(\frac{2}{3})^n$. Since the Cantor set C is the intersection of all these sets, the Lebesgue measure of C is 0.

Suppose we define f_0 to be $\frac{1}{2}$ on the interval $(\frac{1}{3}, \frac{2}{3})$, to be $\frac{1}{4}$ on the interval $(\frac{1}{9}, \frac{2}{9})$, to be $\frac{3}{4}$ on the interval $(\frac{7}{9}, \frac{8}{9})$, and so on. Define $f(x) = \inf\{f_0(y) : y \geq x\}$ for $x < 1$. Define $f(1) = 1$. Notice $f = f_0$ on the complement of the Cantor set. f is increasing, so it has only jump discontinuities; see Proposition 1.6. But if it has a jump continuity, there is a rational of the form $k/2^n$ with $k \leq 2^n$ that is not in the range of f. On the other hand, by the construction, each of the values $\{k/2^n : n \geq 0, k \leq 2^n\}$ is taken by f_0 for some point in the complement of C, and so is taken by f. The only way this can happen is if f is continuous. This function f is called the *Cantor-Lebesgue function* or sometimes simply the *Cantor function*. We will use it in examples later on. For now, we note that it a function that increases only on the Cantor set, which is a set of Lebesgue measure 0, yet f is continuous.

Example 4.12 Let $q_1, q_2, \dots$ be an enumeration of the rationals, let $\varepsilon > 0$, and let I_i be the interval $(q_i - \varepsilon/2^i, q_i + \varepsilon/2^i)$. Then the measure of I_i is $\varepsilon/2^{i-1}$, so the measure of $\cup_i I_i$ is at most 2ε. (It is not equal to that because there is a lot of overlap.) Therefore the measure of $A = [0,1] - \cup_i I_i$ is larger than $1 - 2\varepsilon$. But A contains no rational numbers.

Example 4.13 Let us follow the construction of the Cantor set, with this difference. Instead of removing the middle third at the first stage, remove the middle fourth, i.e., remove $(\frac{3}{8}, \frac{5}{8})$. On each of the two intervals that remain, remove the middle sixteenths. On each of the four intervals that remain, remove the middle interval of length $\frac{1}{64}$, and so on. The total that we removed is

$$\tfrac{1}{4} + 2(\tfrac{1}{16}) + 4(\tfrac{1}{64}) + \cdots = \tfrac{1}{2}.$$

The set that remains contains no intervals, is closed, every point is a limit point, is uncountable, and has measure $1/2$. Such a set is called a *generalized Cantor set*. Of course, other choices than $\frac{1}{4}$, $\frac{1}{16}$, etc. are possible.

Let $A \subset [0,1]$ be a Borel measurable set. We will show that A is "almost equal" to the countable intersection of open sets and "almost equal" to the countable union of closed sets. (A similar argument to what follows is possible for subsets of $\mathbb{R}$ that have infinite measure; see Exercise 4.2.)

Proposition 4.14 *Suppose $A \subset [0,1]$ is a Borel measurable set. Let m be Lebesgue measure.*

(1) Given $\varepsilon > 0$, there exists an open set G so that $m(G-A) < \varepsilon$ and $A \subset G$.

(2) Given $\varepsilon > 0$, there exists a closed set F so that $m(A-F) < \varepsilon$ and $F \subset A$.

(3) There exists a set H which contains A that is the countable intersection of a decreasing sequence of open sets and $m(H - A) = 0$.

(4) There exists a set F which is contained in A that is the countable union of an increasing sequence of closed sets which is contained in A and $m(A - F) = 0$.

Proof. (1) There exists a set of the form $E = \cup_{j=1}^\infty (a_j, b_j]$ such that $A \subset E$ and $m(E - A) < \varepsilon/2$. Let $G = \cup_{j=1}^\infty (a_j, b_j + \varepsilon 2^{-j-1})$. Then G is open and contains A and

$$m(G - E) < \sum_{j=1}^\infty \varepsilon 2^{-j-1} = \varepsilon/2.$$

Therefore

$$m(G - A) \le m(G - E) + m(E - A) < \varepsilon.$$

(2) Find an open set G such that $A' \subset G$ and $m(G - A') < \varepsilon$, where $A' = [0, 1] - A$. Let $F = [0, 1] - G$. Then F is closed, $F \subset A$, and $m(A - F) \le m(G - A') < \varepsilon$.

(3) By (1), for each i, there is an open set G_i that contains A and such that $m(G_i - A) < 2^{-i}$. Then $H_i = \cap_{j=1}^i G_j$ will contain A, is open, and since it is contained in G_i, then $m(H_i - A) < 2^{-i}$. Let $H = \cap_{i=1}^\infty H_i$. H need not be open, but it is the intersection of countably many open sets. The set H is a Borel set, contains A, and $m(H - A) \le m(H_i - A) < 2^{-i}$ for each i, hence $m(H - A) = 0$.

(4) If $A' = [0, 1] - A$, apply (3) to A' to find a set H containing A' that is the countable intersection of a decreasing sequence of open sets and such that $m(H - A') = 0$. Let $J = [0, 1] - H$. It is left to the reader to verify that J has the desired properties. □

The countable intersections of open sets are sometimes called G_δ sets; the G is for *geoffnet*, the German word for "open" and the δ for *Durchschnitt*, the German word for "intersection." The countable unions of closed sets are called F_σ sets, the F coming from *fermé*, the French word for "closed," and the σ coming from *Summe*, the German word for "union."

Therefore, when trying to understand Lebesgue measure, we can look at G_δ or F_σ sets, which are not so bad, and at null sets, which can be quite bad but don't have positive measure.

4.4 Nonmeasurable sets

Theorem 4.15 *Let m^* be defined by* (4.1)*, where $\mathcal{C}$ is the collection of intervals that are open on the left and closed on the right* *and $\ell((a,b]) = b - a$. m^* is not a measure on the collection of all subsets of $\mathbb{R}$.*

Proof. Suppose m^* is a measure. Define $x \sim y$ if $x - y$ is rational. It is easy to see that this is an equivalence relationship on $[0,1]$. For each equivalence class, pick an element out of that class (we need to use the axiom of choice to do this). Call the collection of such points A. Given a set B, define $B + x = \{y + x : y \in B\}$. Note that $\ell((a+q, b+q]) = b - a = \ell((a,b])$ for each a, b, and q, and so by the definition of m^*, we have $m^*(A+q) = m^*(A)$ for each set A and each q. Moreover, the sets $A + q$ are disjoint for different rationals q.

Now

$$[0,1] \subset \cup_{q \in [-1,1] \cap \mathbb{Q}} (A + q),$$

where the sum is only over rational q, so

$$1 \le \sum_{q \in [-1,1], q \in \mathbb{Q}} m^*(A+q),$$

and therefore $m^*(A) > 0$. But

$$\cup_{q \in [-1,1] \cap \mathbb{Q}} (A+q) \subset [-1,2],$$

where again the sum is only over rational q, so if m^* is a measure, then

$$3 \ge \sum_{q \in [0,1], q \in \mathbb{Q}} m^*(A+q),$$

which implies $m^*(A) = 0$, a contradiction. □

4.5 The Carathéodory extension theorem

We prove the Carathéodory extension theorem in this section. This theorem abstracts some of the techniques used above to give a tool for constructing measures in a variety of contexts.

Let $\mathcal{A}_0$ be an algebra but not necessarily a σ-algebra. Saying ℓ is a measure on $\mathcal{A}_0$ means the following: (1) of Definition 3.1

holds and if $A_1, A_2, \ldots$ are pairwise disjoint elements of $\mathcal{A}_0$ and also $\cup_{i=1}^\infty A_i \in \mathcal{A}_0$, then $\ell(\cup_{i=1}^\infty A_i) = \sum_{i=1}^\infty \ell(A_i)$. Sometimes one calls a measure on an algebra a *premeasure*. Recall $\sigma(\mathcal{A}_0)$ is the σ-algebra generated by $\mathcal{A}_0$.

Theorem 4.16 *Suppose $\mathcal{A}_0$ is an algebra and $\ell : \mathcal{A}_0 \to [0,\infty]$ is a measure on $\mathcal{A}_0$. Define*

$$\mu^*(E) = \inf\Big\{\sum_{i=1}^\infty \ell(A_i) : \text{each } A_i \in \mathcal{A}_0, E \subset \cup_{i=1}^\infty A_i\Big\}$$

for $E \subset X$. Then
(1) μ^* is an outer measure;
(2) $\mu^*(A) = \ell(A)$ if $A \in \mathcal{A}_0$;
(3) every set in $\mathcal{A}_0$ is μ^*-measurable;
(4) if ℓ is σ-finite, then there is a unique extension to $\sigma(\mathcal{A}_0)$.

Proof. (1) is Proposition 4.2. We turn to (2). Suppose $E \in \mathcal{A}_0$. We know $\mu^*(E) \le \ell(E)$ since we can take $A_1 = E$ and $A_2, A_3, \ldots$ empty in the definition of μ^*. If $E \subset \cup_{i=1}^\infty A_i$ with $A_i \in \mathcal{A}_0$, let $B_n = E \cap (A_n - (\cup_{i=1}^{n-1} A_i))$. Then the B_n are pairwise disjoint, they are each in $\mathcal{A}_0$, and their union is E. Therefore

$$\ell(E) = \sum_{i=1}^\infty \ell(B_i) \le \sum_{i=1}^\infty \ell(A_i).$$

Taking the infimum over all such sequences $A_1, A_2, \ldots$ shows that $\ell(E) \le \mu^*(E)$.

Next we look at (3). Suppose $A \in \mathcal{A}_0$. Let $\varepsilon > 0$ and let $E \subset X$. Pick $B_1, B_2, \ldots \in \mathcal{A}_0$ such that $E \subset \cup_{i=1}^\infty B_i$ and $\sum_i \ell(B_i) \le \mu^*(E) + \varepsilon$. Then

$$\begin{aligned}\mu^*(E) + \varepsilon &\ge \sum_{i=1}^\infty \ell(B_i) = \sum_{i=1}^\infty \ell(B_i \cap A) + \sum_{i=1}^\infty \ell(B_i \cap A^c)\\ &\ge \mu^*(E \cap A) + \mu^*(E \cap A^c).\end{aligned}$$

Since ε is arbitrary, $\mu^*(E) \ge \mu^*(E \cap A) + \mu^*(E \cap A^c)$. Thus A is μ^*-measurable.

Finally, we look at (4). Suppose we have two extensions to $\sigma(\mathcal{A}_0)$, the smallest σ-algebra containing $\mathcal{A}_0$. One is μ^* and let the other extension be called ν. We will show that if E is in $\sigma(\mathcal{A}_0)$, then $\mu^*(E) = \nu(E)$.

Let us first assume that μ^* is a finite measure. The μ^*-measurable sets form a σ-algebra containing $\mathcal{A}_0$. Because $E \in \mathcal{A}_0$, E must be μ^*-measurable and

$$\mu^*(E) = \inf \Big\{ \sum_{i=1}^\infty \ell(A_i) : E \subset \cup_{i=1}^\infty A_i, \text{ each } A_i \in \mathcal{A}_0 \Big\}.$$

But $\ell = \nu$ on $\mathcal{A}_0$, so $\sum_i \ell(A_i) = \sum_i \nu(A_i)$. Therefore if $E \subset \cup_{i=1}^\infty A_i$ with each $A_i \in \mathcal{A}_0$, then

$$\nu(E) \le \sum_i \nu(A_i) = \sum_i \ell(A_i),$$

which implies

$$\nu(E) \le \mu^*(E). \tag{4.7}$$

Since we do not know that ν is constructed via an outer measure, we must use a different argument to get the reverse inequality. Let $\varepsilon > 0$ and choose $A_i \in \mathcal{A}_0$ such that $\mu^*(E) + \varepsilon \ge \sum_i \ell(A_i)$ and $E \subset \cup_i A_i$. Let $A = \cup_{i=1}^\infty A_i$ and $B_k = \cup_{i=1}^k A_i$. Observe

$$\mu^*(E) + \varepsilon \ge \sum_i \ell(A_i) = \sum_i \mu^*(A_i) \ge \mu^*(\cup_i A_i) = \mu^*(A),$$

hence $\mu^*(A - E) \le \varepsilon$. We have

$$\mu^*(A) = \lim_{k\to\infty} \mu^*(B_k) = \lim_{k\to\infty} \nu(B_k) = \nu(A).$$

Then

$$\begin{aligned}\mu^*(E) &\le \mu^*(A) = \nu(A) = \nu(E) + \nu(A - E)\\ &\le \nu(E) + \mu^*(A - E) \le \nu(E) + \varepsilon,\end{aligned}$$

using (4.7) in the next to last inequality. Since ε is arbitrary, this completes the proof when ℓ is finite.

It remains to consider the case when ℓ is σ-finite. Write $X = \cup_i K_i$, where $K_i \uparrow X$ and $\ell(K_i) < \infty$ for each i. By the preceding paragraph we have uniqueness for the measure ℓ_i defined by $\ell_i(A) = \ell(A \cap K_i)$. If μ and ν are two extensions of ℓ and $A \in \sigma(\mathcal{A}_0)$, then

$$\mu(A) = \lim_{i\to\infty} \mu(A \cap K_i) = \lim_{i\to\infty} \ell_i(A) = \lim_{i\to\infty} \nu(A \cap K_i) = \nu(A),$$

which proves $\mu = \nu$. □

4.6 Exercises

Exercise 4.1 Let μ be a measure on the Borel σ-algebra of $\mathbb{R}$ such that $\mu(K) < \infty$ whenever K is compact, define $\alpha(x) = \mu((0,x])$ if $x \geq 0$ and $\alpha(x) = -\mu((x,0])$ if $x < 0$. Show that μ is the Lebesgue-Stieltjes measure corresponding to α.

Exercise 4.2 Let m be Lebesgue measure and A a Lebesgue measurable subset of $\mathbb{R}$ with $m(A) < \infty$. Let $\varepsilon > 0$. Show there exist G open and F closed such that $F \subset A \subset G$ and $m(G - F) < \varepsilon$.

Exercise 4.3 If $(X, \mathcal{A}, \mu)$ is a measure space, define

$$\mu^*(A) = \inf\{\mu(B) : A \subset B, B \in \mathcal{A}\}$$

for all subsets A of X. Show that μ^* is an outer measure. Show that each set in $\mathcal{A}$ is μ^*-measurable and μ^* agrees with the measure μ on $\mathcal{A}$.

Exercise 4.4 Let m be Lebesgue-Stieltjes measure corresponding to a right continuous increasing function α. Show that for each x,

$$m(\{x\}) = \alpha(x) - \lim_{y \to x-} \alpha(y).$$

Exercise 4.5 Suppose m is Lebesgue measure. Define $x + A = \{x + y : y \in A\}$ and $cA = \{cy : y \in A\}$. Show that for all Lebesgue measurable sets A we have $m(x+A) = m(A)$ and $m(cA) = |c|m(A)$.

Exercise 4.6 Let m be Lebesgue measure. Suppose for each n, A_n is a Lebesgue measurable subset of $[0,1]$. Let B consist of those points x that are in infinitely many of the A_n.
(1) Show B is Lebesgue measurable.
(2) If $m(A_n) > \delta > 0$ for each n, show $m(B) \geq \delta$.
(3) If $\sum_{n=1}^\infty m(A_n) < \infty$, prove that $m(B) = 0$.
(4) Give an example where $\sum_{n=1}^\infty m(A_n) = \infty$, but $m(B) = 0$.

Exercise 4.7 Suppose $\varepsilon \in (0,1)$ and m is Lebesgue measure. Find a measurable set $E \subset [0,1]$ such that the closure of E is $[0,1]$ and $m(E) = \varepsilon$.

Exercise 4.8 If X is a metric space, $\mathcal{B}$ is the Borel σ-algebra, and μ is a measure on $(X, \mathcal{B})$, then the *support* of μ is the smallest closed set F such that $\mu(F^c) = 0$. Show that if F is a closed subset of $[0, 1]$, then there exists a finite measure on $[0, 1]$ whose support is F.

Exercise 4.9 Let m be Lebesgue measure. Find an example of Lebesgue measurable subsets $A_1, A_2, \ldots$ of $[0, 1]$ such that $m(A_n) > 0$ for each n, $m(A_n \Delta A_m) > 0$ if $n \neq m$, and $m(A_n \cap A_m) = m(A_n)m(A_m)$ if $n \neq m$.

Exercise 4.10 Let $\varepsilon \in (0, 1)$, let m be Lebesgue measure, and suppose A is a Borel measurable subset of $\mathbb{R}$. Prove that if

$$m(A \cap I) \leq (1 - \varepsilon)m(I)$$

for every interval I, then $m(A) = 0$.

Exercise 4.11 Suppose m is Lebesgue measure and A is a Borel measurable subset of $\mathbb{R}$ with $m(A) > 0$. Prove that if

$$B = \{x - y : x, y \in A\},$$

then B contains a nonempty open interval centered at the origin.

Exercise 4.12 Let m be Lebesgue measure. Construct a Borel subset A of $\mathbb{R}$ such that $0 < m(A \cap I) < m(I)$ for every open interval I.

Exercise 4.13 Let N be the nonmeasurable set defined in Section 4.4. Prove that if $A \subset N$ and A is Lebesgue measurable, then $m(A) = 0$.

Exercise 4.14 Let m be Lebesgue measure. Prove that if A is a Lebesgue measurable subset of $\mathbb{R}$ and $m(A) > 0$, then there is a subset of A that is nonmeasurable.

Exercise 4.15 Let X be a set and $\mathcal{A}$ a collection of subsets of X that form an algebra of sets. Suppose ℓ is a measure on $\mathcal{A}$ such that $\ell(X) < \infty$. Define μ^* using ℓ as in (4.1). Prove that a set A is μ^*-measurable if and only if

$$\mu^*(A) = \ell(X) - \mu^*(A^c).$$

Exercise 4.16 Suppose μ^* is an outer measure. Show that if $A_n \uparrow A$, then $\mu^*(A_n) \uparrow \mu^*(A)$. Given an example to show that even if μ^* is finite, $A_n \downarrow A$ does not necessarily imply $\mu^*(A_n) \downarrow \mu^*(A)$.

Exercise 4.17 Suppose A is a Lebesgue measurable subset of $\mathbb{R}$ and

$$B = \cup_{x\in A}[x-1, x+1].$$

Prove that B is Lebesgue measurable.

Chapter 5

Measurable functions

We are now ready to move from sets to functions.

5.1 Measurability

Suppose we have a measurable space $(X, \mathcal{A})$.

Definition 5.1 A function $f : X \to \mathbb{R}$ is *measurable* or $\mathcal{A}$ *measurable* if $\{x : f(x) > a\} \in \mathcal{A}$ for all $a \in \mathbb{R}$. A complex-valued function is measurable if both its real and complex parts are measurable.

Example 5.2 Suppose f is real-valued and identically constant. Then the set $\{x : f(x) > a\}$ is either empty or all of X, so f is measurable.

Example 5.3 Suppose $f(x) = 1$ if $x \in A$ and 0 otherwise. Then the set $\{x : f(x) > a\}$ is either $\emptyset$, A, or X. Hence f is measurable if and only if A is in $\mathcal{A}$.

Example 5.4 Suppose X is the real line with the Borel σ-algebra and $f(x) = x$. Then $\{x : f(x) > a\} = (a, \infty)$, and so f is measurable.

Proposition 5.5 *Suppose f is real-valued. The following conditions are equivalent.*

(1) $\{x : f(x) > a\} \in \mathcal{A}$ *for all* $a \in \mathbb{R}$*;*
(2) $\{x : f(x) \leq a\} \in \mathcal{A}$ *for all* $a \in \mathbb{R}$*;*
(3) $\{x : f(x) < a\} \in \mathcal{A}$ *for all* $a \in \mathbb{R}$*;*
(4) $\{x : f(x) \geq a\} \in \mathcal{A}$ *for all* $a \in \mathbb{R}$*.*

Proof. The equivalence of (1) and (2) and of (3) and (4) follow from taking complements, e.g., $\{x : f(x) \leq a\} = \{x : f(x) > a\}^c$. If f is measurable, then

$$\{x : f(x) \geq a\} = \cap_{n=1}^{\infty}\{x : f(x) > a - 1/n\}$$

shows that (4) holds if (1) does. If (4) holds, then (1) holds by using the equality

$$\{x : f(x) > a\} = \cup_{n=1}^{\infty}\{x : f(x) \geq a + 1/n\}.$$

This completes the proof. □

Proposition 5.6 *If* X *is a metric space,* $\mathcal{A}$ *contains all the open sets, and* $f : X \to \mathbb{R}$ *is continuous, then* f *is measurable.*

Proof. Note that $\{x : f(x) > a\} = f^{-1}((a, \infty))$ is open, and hence in $\mathcal{A}$. □

Proposition 5.7 *Let* $c \in \mathbb{R}$*. If* f *and* g *are measurable real-valued functions, then so are* $f + g$*,* $-f$*,* cf*,* fg*,* $\max(f, g)$*, and* $\min(f, g)$*.*

Proof. If $f(x) + g(x) < a$, then $f(x) < a - g(x)$, and there exists a rational r such that $f(x) < r < a - g(x)$. Hence

$$\{x : f(x) + g(x) < a\} = \cup_{r \in \mathbb{Q}}(\{x : f(x) < r\} \cap \{x : g(x) < a - r\}).$$

This proves $f + g$ is measurable.

Since $\{x : -f(x) > a\} = \{x : f(x) < -a\}$, then $-f$ is measurable using Proposition 5.5.

If $c > 0$, then $\{x : cf(x) > a\} = \{x : f(x) > a/c\}$ shows cf is measurable. When $c = 0$, cf is measurable by Example 5.2. When

$c < 0$, write $cf = -(|c|f)$, which is measurable by what we have already proved.

f^2 is measurable since for $a < 0$, $\{x : f(x) > a\} = X$, while for $a \geq 0$,

$$\{x : f(x)^2 > a) = \{x : f(x) > \sqrt{a}\} \cup \{x : f(x) < -\sqrt{a}\}.$$

The measurability of fg follows since

$$fg = \tfrac{1}{2}[(f+g)^2 - f^2 - g^2].$$

The equality

$$\{x : \max(f(x), g(x)) > a\} = \{x : f(x) > a\} \cup \{x : g(x) > a\}$$

shows $\max(f, g)$ is measurable, and the result for $\min(f, g)$ follows from $\min(f, g) = -\max(-f, -g)$. □

Proposition 5.8 *If f_i is a measurable real-valued function for each i, then so are* $\sup_i f_i$, $\inf_i f_i$, $\limsup_{i\to\infty} f_i$, *and* $\liminf_{i\to\infty} f_i$.

Proof. The result will follow for lim sup and lim inf once we have the result for the sup and inf by using the definitions since $\limsup_i f_i = \inf_j \sup_{i\geq j} f_j$ and similarly for the lim inf. We have $\{x : \sup_i f_i > a\} = \cap_{i=1}^{\infty}\{x : f_i(x) > a\}$, so $\sup_i f_i$ is measurable, and the proof for $\inf f_i$ is similar. □

Definition 5.9 We say $f = g$ *almost everywhere*, written $f = g$ a.e., if $\{x : f(x) \neq g(x)\}$ has measure zero. Similarly, we say $f_i \to f$ a.e. if the set of x where $f_i(x)$ does not converge to $f(x)$ has measure zero.

If X is a metric space, $\mathcal{B}$ is the Borel σ-algebra, and $f : X \to \mathbb{R}$ is measurable with respect to $\mathcal{B}$, we say f is Borel measurable. If $f : \mathbb{R} \to \mathbb{R}$ is measurable with respect to the Lebesgue σ-algebra, we say f is Lebesgue measurable.

We saw in Proposition 5.6 that all continuous functions are Borel measurable. The same is true for increasing functions on the real line.

Proposition 5.10 *If $f : \mathbb{R} \to \mathbb{R}$ is monotone, then f is Borel measurable.*

Proof. Let us suppose f is increasing, for otherwise we look at $-f$. Given $a \in \mathbb{R}$, let $x_0 = \sup\{y : f(y) \le a\}$. If $f(x_0) \le a$, then $\{x : f(x) > a\} = (x_0, \infty)$. If $f(x_0) > a$, then $\{x : f(x) > a\} = [x_0, \infty)$. In either case $\{x : f(x) > a\}$ is a Borel set. □

Proposition 5.11 *Let $(X, \mathcal{A})$ be a measurable space and let $f : X \to \mathbb{R}$ be a $\mathcal{A}$ measurable function. If A is in the Borel σ-algebra on $\mathbb{R}$, then $f^{-1}(A) \in \mathcal{A}$.*

Proof. Let $\mathcal{B}$ be the Borel σ-algebra on $\mathbb{R}$ and $\mathcal{C} = \{A \in \mathcal{B} : f^{-1}(A) \in \mathcal{A}\}$. If $A_1, A_2, \ldots \in \mathcal{C}$, then since

$$f^{-1}(\cup_i A_i) = \cup_i f^{-1}(A_i) \in \mathcal{A},$$

we have that $\mathcal{C}$ is closed under countable unions. Similarly $\mathcal{C}$ is closed under countable intersections and complements, so $\mathcal{C}$ is a σ-algebra. Since f is measurable, $\mathcal{C}$ contains (a, ∞) for every real a, hence $\mathcal{C}$ contains the σ-algebra generated by these intervals, that is, $\mathcal{C}$ contains $\mathcal{B}$. □

Example 5.12 Let us construct a set that is Lebesgue measurable, but not Borel measurable. Recall the Lebesgue measurable sets were constructed in Chapter 4 and include the completion of the Borel σ-algebra.

Let f be the Cantor-Lebesgue function of Example 4.11 and define

$$F(x) = \inf\{y : f(y) \ge x\}.$$

Although F is not continuous, observe that F is strictly increasing (hence one-to-one) and maps $[0, 1]$ into C, the Cantor set. Since F is increasing, F^{-1} maps Borel measurable sets to Borel measurable sets.

Let m be Lebesgue measure and let A be the non-measurable set we constructed in Proposition 4.15. Let $B = F(A)$. Since $F(A) \subset C$ and $m(C) = 0$, then $F(A)$ is a null set, hence is Lebesgue measurable. On the other hand, $F(A)$ is not Borel measurable, because if it were, then $A = F^{-1}(F(A))$ would be Borel measurable, a contradiction.

5.2 Approximation of functions

Definition 5.13 Let $(X, \mathcal{A})$ be a measurable space. If $E \in \mathcal{A}$, define the *characteristic function* of E by

$$\chi_E(x) = \begin{cases} 1 & x \in E; \\ 0 & x \notin E. \end{cases}$$

A *simple function* s is a function of the form

$$s(x) = \sum_{i=1}^{n} a_i \chi_{E_i}(x)$$

for real numbers a_i and measurable sets E_i.

Proposition 5.14 *Suppose f is a non-negative and measurable function. Then there exists a sequence of non-negative measurable simple functions increasing to f.*

Proof. Let

$$E_{ni} = \{x : (i-1)/2^n \le f(x) < i/2^n\}$$

and

$$F_n = \{x : f(x) \ge n\}$$

for $n = 1, 2, \ldots$ and $i = 1, 2, \ldots, n2^n$. Then define

$$s_n = \sum_{i=1}^{n2^n} \frac{i-1}{2^n} \chi_{E_{ni}} + n\chi_{F_n}.$$

In words, $f_n(x) = n$ if $f(x) \ge n$. If $f(x)$ is between $(i-1)/2^n$ and $i/2^n$ for $i/2^n \le n$, we let $f_n(x) = (i-1)/2^n$.

It is easy to see that s_n has the desired properties. □

5.3 Lusin's theorem

The following theorem is known as Lusin's theorem. It is very pretty but usually other methods are better for solving problems.

Example 5.16 will illustrate why this is a less useful theorem than at first glance.

We use m for Lebesgue measure. Recall that the support of a function f is the closure of the set $\{x : f(x) \neq 0\}$.

Theorem 5.15 *Suppose $f : [0,1] \to \mathbb{R}$ is Borel measurable, m is Lebesgue measure, and $\varepsilon > 0$. There exists a closed set $F \subset [0,1]$ such that $m([0,1] - F) < \varepsilon$ and the restriction of f to F is a continuous function on F.*

This theorem can be loosely interpreted as saying every measurable function is "almost continuous."

Proof. First let us suppose that $f = \chi_A$, where A is a Borel measurable subset of $[0,1]$. By Proposition 4.14 we can find E closed and G open such that $E \subset A \subset G$ and $m(G - A) < \varepsilon/2$ and $m(A - E) < \varepsilon/2$. Let $\delta = \inf\{|x - y| : x \in E, y \in G^c\}$. Since $E \subset A \subset [0,1]$, E is compact and $\delta > 0$. Letting

$$g(x) = \Big(1 - \frac{d(x,E)}{\delta}\Big)^+,$$

where $y^+ = \max(y,0)$ and $d(x,E) = \inf\{|x - y| : y \in E\}$, we see that g is continuous, takes values in $[0,1]$, is equal to 1 on E, and equal to 0 on G^c. Take $F = (E \cup G^c) \cap [0,1]$. Then $m([0,1] - F) \le m(G - E) < \varepsilon$, and $f = g$ on F.

Next suppose $f = \sum_{i=1}^M a_i \chi_{A_i}$ is simple, where each A_i is a measurable subset of $[0,1]$ and each $a_i \ge 0$. Choose F_i closed such that $m([0,1] - F_i) < \varepsilon/M$ and χ_{A_i} restricted to F_i is continuous. If we let $F = \cap_{i=1}^M F_i$, then F is closed, $m([0,1] - F) < \varepsilon$, and f restricted to F is continuous.

Now suppose f is non-negative, bounded by K, and has support in $[0,1]$. Let

$$A_{in} = \{x : (i-1)/2^n \le f(x) < i/2^n\}.$$

Then

$$f_n(x) = \sum_{i=1}^{K2^n+1} \frac{i}{2^n} \chi_{A_{in}}(x)$$

are simple functions increasing to f. Note that

$$h_n(x) = f_{n+1}(x) - f_n(x)$$

is also a simple function and is bounded by 2^{-n}. Choose F_0 closed such that $m([0,1]-F_0) < \varepsilon/2$ and f_0 restricted to F_0 is continuous. For $n \geq 1$, choose F_n closed such that $m([0,1]-F_n) < \varepsilon/2^{n+1}$ and h_n restricted to F_n is continuous. Let $F = \cap_{n=0}^{\infty} F_n$. Then F, being the intersection of closed sets, will be closed, and $m([0,1]-F) \leq \sum_{n=0}^{\infty} m([0,1]-F_n) < \varepsilon$. On the set F, we have $f_0(x)+\sum_{n=1}^{\infty} h_n(x)$ converges uniformly to $f(x)$ because each h_n is bounded by 2^{-n}. The uniform limit of continuous functions is continuous, hence f is continuous on F.

If $f \geq 0$, let $B_K = \{x : f(x) \leq K\}$. Since f is everywhere finite, $B_K \uparrow [0,1]$ as $K \to \infty$, hence $m(B_K) > 1 - \varepsilon/3$ if K is sufficiently large. Choose $D \subset B_K$ such that D is closed and $m(B_K - D) < \varepsilon/3$. Now choose $E \subset [0,1]$ closed such that $f \cdot \chi_D$ restricted to E is continuous and $m([0,1]-E) < \varepsilon/3$. Then $F = D \cap E$ is closed, $m([0,1]-F) < \varepsilon$, and f restricted to F is continuous.

Finally, for arbitrary measurable f write $f = f^+ - f^-$ and find F^+ and F^- closed such that $m([0,1]-F^+) < \varepsilon/2$, $m([0,1]-F^-) < \varepsilon/2$, and f^+ restricted to F^+ is continuous and similarly for f^-. Then $F = F^+ \cap F^-$ is the desired set. □

Example 5.16 Suppose $f = \chi_B$, where B consists of the irrationals in $[0,1]$. f is Borel measurable because $[0,1]-B$ is countable, hence the union of countably many points, and thus the union of countably many closed sets. Every point of $[0,1]$ is a point of discontinuity of f because for any $x \in [0,1]$, there are both rationals and irrationals in every neighborhood of x, hence f takes the values 0 and 1 in every neighborhood of x.

Recall Example 4.12. f restricted to the set A there is identically one, hence f restricted to A is a continuous function. A is closed because it is equal to the interval $[0,1]$ minus the union of open intervals.

This does not contradict Lusin's theorem. No claim is made that the function f is continuous at most points of $[0,1]$. What is asserted is that there is a closed set F with large measure so that f restricted to F is continuous when viewed as a function from F to $\mathbb{R}$.

5.4 Exercises

Exercise 5.1 Suppose $(X, \mathcal{A})$ is a measurable space, f is a real-valued functions, and $\{x : f(x) > r\} \in \mathcal{A}$ for each rational number r. Prove that f is measurable.

Exercise 5.2 Let $f : (0,1) \to \mathbb{R}$ be such that for every $x \in (0,1)$ there exist $r > 0$ and a Borel measurable function g, both depending on x, such that f and g agree on $B(x,r)$. Prove that f is Borel measurable.

Exercise 5.3 Suppose f_n are measurable functions. Prove that

$$A = \{x : \lim_{n\to\infty} f_n(x) \text{ exists}\}$$

is a measurable set.

Exercise 5.4 If $f : \mathbb{R} \to \mathbb{R}$ is Lebesgue measurable, prove that there exists a Borel measurable function g such that $f = g$ a.e.

Exercise 5.5 Give an example of a collection of measurable non-negative functions $\{f_\alpha\}_{\alpha\in A}$ such that if g is defined by $g(x) = \sup_{\alpha\in A} f_\alpha(x)$, then g is finite for all x but g is nonmeasurable. (A is allowed to be uncountable.)

Exercise 5.6 Suppose $f : X \to \mathbb{R}$ is Lebesgue measurable and $g : \mathbb{R} \to \mathbb{R}$ is continuous. Prove that $g \circ f$ is Lebesgue measurable. Is this true if g is Borel measurable instead of continuous? Is this true if g is Lebesgue measurable instead of continuous?

Exercise 5.7 Suppose $f : \mathbb{R} \to \mathbb{R}$ is Borel measurable. Define $\mathcal{A}$ to be the smallest σ-algebra containing the sets $\{x : f(x) > a\}$ for every $a \in \mathbb{R}$. Suppose $g : \mathbb{R} \to \mathbb{R}$ is measurable with respect to $\mathcal{A}$, which means that $\{x : g(x) > a\} \in \mathcal{A}$ for every $a \in \mathbb{R}$. Prove that there exists a Borel measurable function $h : \mathbb{R} \to \mathbb{R}$ such that $g = h \circ f$.

Exercise 5.8 One can show that there exist discontinuous real-valued functions f such that

$$f(x+y) = f(x) + f(y) \qquad (5.1)$$

for all $x, y \in \mathbb{R}$. (The construction uses Zorn's lemma, which is equivalent to the axiom of choice.) Prove that if f satisfies (5.1) and in addition f is Lebesgue measurable, then f is continuous.

Chapter 6

The Lebesgue integral

In this chapter we define the Lebesgue integral. We only give the definition here and consider the properties of the Lebesgue integral in later chapters.

6.1 Definitions

Definition 6.1 Let $(X, \mathcal{A}, \mu)$ be a measure space. If

$$s = \sum_{i=1}^{n} a_i \chi_{E_i}$$

is a non-negative measurable simple function, define the Lebesgue integral of s to be

$$\int s\,d\mu = \sum_{i=1}^{n} a_i \mu(E_i). \tag{6.1}$$

Here, if $a_i = 0$ and $\mu(E_i) = \infty$, we use the convention that $a_i\mu(E_i) = 0$. If $f \geq 0$ is a measurable function, define

$$\int f\,d\mu = \sup\Big\{\int s\,d\mu : 0 \leq s \leq f, s \text{ simple}\Big\}. \tag{6.2}$$

Let f be measurable and let $f^+ = \max(f, 0)$ and $f^- = \max(-f, 0)$. Provided $\int f^+\,d\mu$ and $\int f^-\,d\mu$ are not both infinite, define

$$\int f\,d\mu = \int f^+\,d\mu - \int f^-\,d\mu. \tag{6.3}$$

Finally, if $f = u + iv$ is complex-valued and $\int (|u| + |v|)\, d\mu$ is finite, define

$$\int f\, d\mu = \int u\, d\mu + i \int v\, d\mu. \tag{6.4}$$

A few remarks are in order. A function s might be written as a simple function in more than one way. For example $s = \chi_{A\cup B} = \chi_A + \chi_B$ if A and B are disjoint. It is not hard to check that the definition of $\int s\, d\mu$ is unaffected by how s is written. If $s = \sum_{i=1}^m a_i \chi_{A_i} = \sum_{j=1}^n b_j \chi_{B_j}$, then we need to show

$$\sum_{i=1}^m a_i \mu(A_i) = \sum_{j=1}^n b_j \mu(B_j). \tag{6.5}$$

We leave the proof of this to the reader as Exercise 6.1.

Secondly, if s is a simple function, one has to think a moment to verify that the definition of $\int s\, d\mu$ by means of (6.1) agrees with its definition by means of (6.2).

Definition 6.2 If f is measurable and $\int |f|\, d\mu < \infty$, we say f is *integrable.*

The proof of the next proposition follows from the definitions.

Proposition 6.3 *(1) If f is a real-valued measurable function with $a \le f(x) \le b$ for all x and $\mu(X) < \infty$, then $a\mu(X) \le \int f\, d\mu \le b\mu(X)$;*

(2) If f and g are measurable, real-valued, and integrable and $f(x) \le g(x)$ for all x, then $\int f\, d\mu \le \int g\, d\mu$.

(3) If f is integrable, then $\int cf\, d\mu = c \int f\, d\mu$ for all complex c.

(4) If $\mu(A) = 0$ and f is integrable, then $\int f \chi_A\, d\mu = 0$.

The integral $\int f\chi_A\, d\mu$ is often written $\int_A f\, d\mu$. Other notation for the integral is to omit the μ and write $\int f$ if it is clear which measure is being used, to write $\int f(x)\, \mu(dx)$, or to write $\int f(x)\, d\mu(x)$.

When we are integrating a function f with respect to Lebesgue measure m, it is usual to write $\int f(x)\, dx$ for $\int f(x)\, m(dx)$ and to

define
$$\int_a^b f(x)\,dx = \int_{[a,b]} f(x)\,m(dx).$$

Proposition 6.4 *If f is integrable,*
$$\Big|\int f\Big| \le \int |f|.$$

Proof. For the real case, this is easy. $f \le |f|$, so $\int f \le \int |f|$. Also $-f \le |f|$, so $-\int f \le \int |f|$. Now combine these two facts.

For the complex case, $\int f$ is a complex number. If it is 0, the inequality is trivial. If it is not, then $\int f = re^{i\theta}$ for some r and θ. Then
$$\Big|\int f\Big| = r = e^{-i\theta}\int f = \int e^{-i\theta} f.$$
From the definition of $\int f$ when f is complex, we have $\mathrm{Re}\,(\int f) = \int \mathrm{Re}\,(f)$. Since $|\int f|$ is real, we have
$$\Big|\int f\Big| = \mathrm{Re}\Big(\int e^{-i\theta} f\Big) = \int \mathrm{Re}\,(e^{-i\theta} f) \le \int |f|.$$
□

We do not yet know that $\int (f+g) = \int f + \int g$. We will see this in Theorem 7.4.

6.2 Exercises

Exercise 6.1 Verify (6.5).

Exercise 6.2 Suppose f is non-negative and measurable and μ is σ-finite. Show there exist simple functions s_n increasing to f at each point such that $\mu(\{x : s_n(x) \ne 0\}) < \infty$ for each n.

Exercise 6.3 Let f be a non-negative measurable function. Prove that
$$\lim_{n\to\infty} \int (f \wedge n) \to \int f.$$

Exercise 6.4 Let $(X, \mathcal{A}, \mu)$ be a measure space and suppose μ is σ-finite. Suppose f is integrable. Prove that given ε there exists δ such that

$$\int_A |f(x)| \, \mu(dx) < \varepsilon$$

whenever $\mu(A) < \delta$.

Exercise 6.5 Suppose $\mu(X) < \infty$ and f_n is a sequence of bounded real-valued measurable functions that converge to f uniformly. Prove that

$$\int f_n \, d\mu \to \int f \, d\mu.$$

This is sometimes called the *bounded convergence theorem.*

Exercise 6.6 If f_n is a sequence of non-negative integrable functions such that $f_n(x)$ decreases to $f(x)$ for every x, prove that $\int f_n \, d\mu \to \int f \, d\mu$.

Exercise 6.7 Let $(X, \mathcal{A}, \mu)$ be a measure space and suppose f is a non-negative, measurable function that is finite at each point of X, but not necessarily integrable. Prove that there exists a continuous increasing function $g : [0, \infty) \to [0, \infty)$ such that $\lim_{x\to\infty} g(x) = \infty$ and $g \circ f$ is integrable.

Chapter 7

Limit theorems

The main reason the Lebesgue integral is so much easier to work with than the Riemann integral is that it behaves nicely when taking limits. In this chapter we prove the monotone convergence theorem, Fatou's lemma, and the dominated convergence theorem. We also prove that the Lebesgue integral is linear.

7.1 Monotone convergence theorem

One of the most important results concerning Lebesgue integration is the *monotone convergence theorem.*

Theorem 7.1 *Suppose f_n is a sequence of non-negative measurable functions with $f_1(x) \le f_2(x) \le \cdots$ for all x and with*

$$\lim_{n\to\infty} f_n(x) = f(x)$$

for all x. Then $\int f_n\, d\mu \to \int f\, d\mu$.

Proof. By Proposition 6.3(2), $\int f_n$ is an increasing sequence of real numbers. Let L be the limit. Since $f_n \le f$ for all n, then $L \le \int f$. We must show $L \ge \int f$.

Let $s = \sum_{i=1}^m a_i \chi_{E_i}$ be any non-negative simple function less than or equal to f and let $c \in (0,1)$. Let $A_n = \{x : f_n(x) \ge cs(x)\}$.

Since $f_n(x)$ increases to $f(x)$ for each x and $c < 1$, then $A_n \uparrow X$. For each n,

$$\begin{aligned}\int f_n \ge \int_{A_n} f_n &\ge c\int_{A_n} s_n \\ &= c\int_{A_n} \sum_{i=1}^m a_i \chi_{E_i} \\ &= c\sum_{i=1}^m a_i \mu(E_i \cap A_n).\end{aligned}$$

If we let $n \to \infty$, by Proposition 3.5(3) the right hand side converges to

$$c\sum_{i=1}^m a_i \mu(E_i) = c\int s.$$

Therefore $L \ge c\int s$. Since c is arbitrary in the interval $(0, 1)$, then $L \ge \int s$. Taking the supremum over all simple $s \le f$, we obtain $L \ge \int f$. □

Example 7.2 Let $X = [0, \infty)$ and $f_n(x) = -1/n$ for all x. Then $\int f_n = -\infty$, but $f_n \uparrow f$ where $f = 0$ and $\int f = 0$. The reason the monotone convergence theorem does not apply here is that the f_n are not non-negative.

Example 7.3 Suppose $f_n = n\chi_{(0,1/n)}$. Then $f_n \ge 0$, $f_n \to 0$ for each x, but $\int f_n = 1$ does not converge to $\int 0 = 0$. The reason the monotone convergence theorem does not apply here is that the f_n do not increase to f for each x.

7.2 Linearity of the integral

Once we have the monotone convergence theorem, we can prove that the Lebesgue integral is linear.

Theorem 7.4 *If f and g are non-negative and measurable or if f and g are integrable, then*

$$\int (f + g)\, d\mu = \int f\, d\mu + \int g\, d\mu.$$

Proof. First suppose f and g are non-negative and simple, say, $f = \sum_{i=1}^m a_i \chi_{A_i}$ and $g = \sum_{j=1}^n b_j \chi_{B_j}$. Without loss of generality we may assume that $A_1, \ldots, A_m$ are pairwise disjoint and that $B_1, \ldots, B_n$ are pairwise disjoint. Since $a_i = 0$ and $b_j = 0$ are permissible, we may also assume $\cup_{i=1}^m A_i = X = \cup_{j=1}^n B_j$. Then

$$f + g = \sum_{i=1}^m \sum_{j=1}^n (a_i + b_j) \chi_{A_i \cap B_j},$$

and we have

$$\begin{aligned} \int (f+g) &= \sum_{i=1}^m \sum_{j=1}^n (a_i + b_j) \mu(A_i \cap B_j) \\ &= \sum_{i=1}^m \sum_{j=1}^n a_i \mu(A_i \cap B_j) + \sum_{i=1}^m \sum_{j=1}^n b_j \mu(A_i \cap B_j) \\ &= \sum_{i=1}^m a_i \mu(A_i) + \sum_{j=1}^n b_j \mu(B_j) \\ &= \int f + \int g. \end{aligned}$$

Thus the theorem holds in this case.

Next suppose f and g are non-negative. Take s_n simple and increasing to f and t_n simple and increasing to g. Then $s_n + t_n$ are simple functions increasing to $f + g$, so the result follows from the monotone convergence theorem and

$$\int (f+g) = \lim_{n\to\infty} \int (s_n + t_n) = \lim_{n\to\infty} \int s_n + \lim_{n\to\infty} t_n = \int f + \int g.$$

Suppose now that f and g are real-valued and integrable but take both positive and negative values. Since

$$\int |f+g| \le \int (|f| + |g|) = \int |f| + \int |g| < \infty,$$

then $f + g$ is integrable. Write

$$(f+g)^+ - (f+g)^- = f + g = f^+ - f^- + g^+ - g^-,$$

so that

$$(f+g)^+ + f^- + g^- = f^+ + g^+ + (f+g)^-.$$

Using the result for non-negative functions,

$$\int (f+g)^+ + \int f^- + \int g^- = \int f^+ + \int g^+ + \int (f+g)^-.$$

Rearranging,

$$\begin{aligned}\int (f+g) &= \int (f+g)^+ - \int (f+g)^- \\ &= \int f^+ - \int f^- + \int g^+ - \int g^- \\ &= \int f + \int g.\end{aligned}$$

If f and g are complex-valued, apply the above to the real and imaginary parts. □

Proposition 7.5 *Suppose f_n are non-negative measurable functions. Then*

$$\int \sum_{n=1}^{\infty} f_n = \sum_{n=1}^{\infty} \int f_n.$$

Proof. Let $F_N = \sum_{n=1}^N f_n$. Since $0 \le F_n(x) \uparrow \sum_{n=1}^\infty f_n(x)$, we can write

$$\begin{aligned}\int \sum_{n=1}^{\infty} f_n &= \int \lim_{N\to\infty} \sum_{n=1}^{N} f_n \\ &= \int \lim_{N\to\infty} F_N = \lim_{N\to\infty} \int F_N \qquad (7.1) \\ &= \lim_{N\to\infty} \sum_{n=1}^{N} \int f_n = \sum_{n=1}^{\infty} \int f_n,\end{aligned}$$

using the monotone convergence theorem and the linearity of the integral. □

7.3 Fatou's lemma

The next theorem is known as *Fatou's lemma.*

Theorem 7.6 *Suppose the f_n are non-negative and measurable. Then*

$$\int \liminf_{n\to\infty} f_n \le \liminf_{n\to\infty} \int f_n.$$

Proof. Let $g_n = \inf_{i\ge n} f_i$. Then the g_n are non-negative and g_n increases to $\liminf_n f_n$. Clearly $g_n \le f_i$ for each $i \ge n$, so $\int g_n \le \int f_i$. Therefore

$$\int g_n \le \inf_{i\ge n} \int f_i. \tag{7.2}$$

If we take the limit in (7.2) as $n \to \infty$, on the left hand side we obtain $\int \liminf_n f_n$ by the monotone convergence theorem, while on the right hand side we obtain $\liminf_n \int f_n$. □

A typical use of Fatou's lemma is the following. Suppose we have $f_n \to f$ and $\sup_n \int |f_n| \le K < \infty$. Then $|f_n| \to |f|$, and by Fatou's lemma, $\int |f| \le K$.

7.4 Dominated convergence theorem

Another very important theorem is the *dominated convergence theorem.*

Theorem 7.7 *Suppose that f_n are measurable real-valued functions and $f_n(x) \to f(x)$ for each x. Suppose there exists a non-negative integrable function g such that $|f_n(x)| \le g(x)$ for all x. Then*

$$\lim_{n\to\infty} \int f_n \, d\mu \to \int f \, d\mu.$$

Proof. Since $f_n + g \ge 0$, by Fatou's lemma,

$$\int f + \int g = \int (f+g) \le \liminf_{n\to\infty} \int (f_n + g) = \liminf_{n\to\infty} f_n + \int g.$$

Since g is integrable,

$$\int f \le \liminf_{n\to\infty} \int f_n. \tag{7.3}$$

Similarly, $g - f_n \geq 0$, so

$$\int g - \int f = \int (g-f) \leq \liminf_{n\to\infty} \int (g - f_n) = \int g + \liminf_{n\to\infty} \int (-f_n),$$

and hence

$$-\int f \leq \liminf_{n\to\infty} \int (-f_n) = -\limsup_{n\to\infty} \int f_n.$$

Therefore

$$\int f \geq \limsup_{n\to\infty} \int f_n,$$

which with (7.3) proves the theorem. □

Exercise 7.1 asks you to prove a version of the dominated convergence theorem for complex-valued functions.

Example 7.3 is an example where the limit of the integrals is not the integral of the limit because there is no dominating function g.

If in the monotone convergence theorem or dominated convergence theorem we have only $f_n(x) \to f(x)$ almost everywhere, the conclusion still holds. For example, if the f_n are measurable, non-negative, and $f_n \uparrow f$ a.e., let $A = \{x : f_n(x) \to f(x)\}$. Then $f_n \chi_A(x) \uparrow f\chi_A(x)$ for each x. Since A^c has measure 0, we see from Proposition 6.3(4) and the monotone convergence theorem that

$$\lim_n \int f_n = \lim_n \int f_n \chi_A = \int f\chi_A = \int f.$$

7.5 Exercises

Exercise 7.1 State and prove a version of the dominated convergence theorem for complex-valued functions.

Exercise 7.2 The following generalized dominated convergence theorem is often useful. Suppose f_n, g_n, f, and g are integrable, $f_n \to f$ a.e., $g_n \to g$ a.e., $|f_n| \leq g_n$ for each n, and $\int g_n \to \int g$. Prove that $\int f_n \to \int f$.

Exercise 7.3 Give an example of a sequence of non-negative functions f_n tending to 0 pointwise such that $\int f_n \to 0$, but there is no integrable function g such that $f_n \leq g$ for all n.

Exercise 7.4 Suppose $(X, \mathcal{A}, \mu)$ is a measure space, each f_n is integrable and non-negative, $f_n \to f$ a.e., and $\int f_n \to \int f$. Prove that for each $A \in \mathcal{A}$

$$\int_A f_n \, d\mu \to \int_A f \, d\mu.$$

Exercise 7.5 Suppose f_n and f are integrable, $f_n \to f$ a.e., and $\int |f_n| \to \int |f|$. Prove that

$$\int |f_n - f| \to 0.$$

Exercise 7.6 Suppose $f : \mathbb{R} \to \mathbb{R}$ is integrable, $a \in \mathbb{R}$, and we define

$$F(x) = \int_a^x f(y) \, dy.$$

Show that F is a continuous function.

Exercise 7.7 Let f_n be a sequence of non-negative Lebesgue measurable functions on $\mathbb{R}$. Is it necessarily true that

$$\limsup_{n\to\infty} \int f_n \, dx \le \int \limsup_{n\to\infty} f_n \, dx?$$

If not, give a counterexample.

Exercise 7.8 Find the limit

$$\lim_{n\to\infty} \int_0^n \Big(1 + \frac{x}{n}\Big)^{-n} \log(2 + \cos(x/n)) \, dx$$

and justify your reasoning.

Exercise 7.9 Find the limit

$$\lim_{n\to\infty} \int_0^n \Big(1 - \frac{x}{n}\Big)^{n} \log(2 + \cos(x/n)) \, dx$$

and justify your reasoning.

Exercise 7.10 Prove that the limit exists and find its value:

$$\lim_{n\to\infty} \int_0^1 \frac{1 + nx^2}{(1 + x^2)^n} \log(2 + \cos(x/n)) \, dx.$$

Exercise 7.11 Prove the limit exists and determine its value:

$$\lim_{n\to\infty} \int_0^\infty ne^{-nx} \sin(1/x)\, dx.$$

Exercise 7.12 Let $g : \mathbb{R} \to \mathbb{R}$ be integrable and let $f : \mathbb{R} \to \mathbb{R}$ be bounded, measurable, and continuous at 1. Prove that

$$\lim_{n\to\infty} \int_{-n}^{n} f\Big(1 + \frac{x}{n^2}\Big) g(x)\, dx$$

exists and determine its value.

Exercise 7.13 Suppose $\mu(X) < \infty$, f_n converges to f uniformly, and each f_n is integrable. Prove that f is integrable and $\int f_n \to \int f$. Is the condition $\mu(X) < \infty$ necessary?

Exercise 7.14 Prove that

$$\sum_{k=1}^{\infty} \frac{1}{(p+k)^2} = -\int_0^1 \frac{x^p}{1-x} \log x\, dx$$

for $p > 0$.

Exercise 7.15 Let $\{f_n\}$ be a sequence of real-valued functions on $[0, 1]$ that is uniformly bounded.
(1) Show that if A is a Borel subset of $[0, 1]$, then there exists a subsequence n_j such that $\int_A f_{n_j}(x)\, dx$ converges.
(2) Show that if $\{A_i\}$ is a countable collection of Borel subsets of $[0, 1]$, then there exists a subsequence n_j such that $\int_{A_i} f_{n_j}(x)\, dx$ converges for each i.
(3) Show that there exists a subsequence n_j such that $\int_A f_{n_j}(x)\, dx$ converges for each Borel subset A of $[0, 1]$.

Exercise 7.16 Let $(X, \mathcal{A}, \mu)$ be a measure space. A family of measurable functions $\{f_n\}$ is *uniformly integrable* if given ε there exists M such that

$$\int_{\{x : |f_n(x)| > M\}} |f_n(x)|\, d\mu < \varepsilon$$

for each n. The sequence is *uniformly absolutely continuous* if given ε there exists δ such that

$$\Big| \int_A f_n\, d\mu \Big| < \varepsilon$$

for each n if $\mu(A) < \delta$.

Suppose μ is a finite measure. Prove that $\{f_n\}$ is uniformly integrable if and only if $\sup_n \int |f_n|\, d\mu < \infty$ and $\{f_n\}$ is uniformly absolutely continuous.

Exercise 7.17 The following is known as the *Vitali convergence theorem*. Suppose μ is a finite measure, $f_n \to f$ a.e., and $\{f_n\}$ is uniformly integrable. Prove that $\int f_n \to \int f$.

Exercise 7.18 Suppose μ is a finite measure, $f_n \to f$ a.e., each f_n is integrable, f is integrable, and $\int f_n \to \int f$. Prove that $\{f_n\}$ is uniformly integrable.

Exercise 7.19 Suppose μ is a finite measure and for some $\varepsilon > 0$

$$\sup_n \int |f_n|^{1+\varepsilon}\, d\mu < \infty.$$

Prove that $\{f_n\}$ is uniformly integrable.

Exercise 7.20 Suppose f_n is a uniformly integrable sequence of functions defined on $[0,1]$. Prove that there is a subsequence n_j such that $\int_0^1 f_{n_j} g\, dx$ converges whenever g is a real-valued bounded measurable function.

Exercise 7.21 Suppose μ_n is a sequence of measures on $(X, \mathcal{A})$ such that $\mu_n(X) = 1$ for all n and $\mu_n(A)$ converges as $n \to \infty$ for each $A \in \mathcal{A}$. Call the limit $\mu(A)$.
(1) Prove that μ is a measure.
(2) Prove that $\int f\, d\mu_n \to \int f\, d\mu$ whenever f is bounded and measurable.
(3) Prove that

$$\int f\, d\mu \le \liminf_{n\to\infty} \int f\, d\mu_n$$

whenever f is non-negative and measurable.

Exercise 7.22 Let $(X, \mathcal{A}, \mu)$ be a measure space and let f be non-negative and integrable. Define ν on $\mathcal{A}$ by

$$\nu(A) = \int_A f\, d\mu.$$

(1) Prove that ν is a measure.
(2) Prove that if g is integrable with respect to ν, then fg is integrable with respect to μ and

$$\int g\,d\nu = \int fg\,d\mu.$$

Exercise 7.23 Suppose μ and ν are positive measures on the Borel σ-algebra on $[0,1]$ such that $\int f\,d\mu = \int f\,d\nu$ whenever f is real-valued and continuous on $[0,1]$. Prove that $\mu = \nu$.

Exercise 7.24 Let $\mathcal{B}$ be the Borel σ-algebra on $[0,1]$. Let μ_n be a sequence of finite measures on $([0,1],\mathcal{B})$ and let μ be another finite measure on $([0,1],\mathcal{B})$. Suppose $\mu_n([0,1]) \to \mu([0,1])$. Prove that the following are equivalent:
(1) $\int f\,d\mu_n \to \int f\,d\mu$ whenever f is a continuous real-valued function on $[0,1]$;
(2) $\limsup_{n\to\infty} \mu_n(F) \le \mu(F)$ for all closed subsets F of $[0,1]$;
(3) $\liminf_{n\to\infty} \mu_n(G) \ge \mu(G)$ for all open subsets G of $[0,1]$;
(4) $\lim_{n\to\infty} \mu_n(A) = \mu(A)$ whenever A is a Borel subset of $[0,1]$ such that $\mu(\partial A) = 0$, where $\partial A = \overline{A} - A^o$ is the boundary of A;
(5) $\lim_{n\to\infty} \mu_n([0,x]) = \mu([0,x])$ for every x such that $\mu(\{x\}) = 0$.

Exercise 7.25 Let $\mathcal{B}$ be the Borel σ-algebra on $[0,1]$. Suppose μ_n are finite measures on $([0,1],\mathcal{B})$ such that $\int f\,d\mu_n \to \int_0^1 f\,dx$ whenever f is a real-valued continuous function on $[0,1]$. Suppose that g is a bounded measurable function such that the set of discontinuities of g has measure 0. Prove that

$$\int g\,d\mu_n \to \int_0^1 g\,dx.$$

Exercise 7.26 Let $\mathcal{B}$ be the Borel σ-algebra on $[0,1]$. Let μ_n be a sequence of finite measures on $([0,1],\mathcal{B})$ with $\sup_n \mu_n([0,1]) < \infty$. Define $\alpha_n(x) = \mu_n([0,x])$.
(1) If r is a rational in $[0,1]$, prove that there exists a subsequence $\{n_j\}$ such that $\alpha_{n_j}(r)$ converges.
(2) Prove that there exists a subsequence $\{n_j\}$ such that $\alpha_n(r)$ converges for every rational in $[0,1]$.
(3) Let $\overline{\alpha}(r) = \lim_{n\to\infty} \alpha_n(r)$ for r rational and define

$$\alpha(x) = \lim_{r\to x, r>x, r\in\mathbb{Q}} \overline{\alpha}(r).$$

This means, since clearly $\overline{\alpha}(r) \le \overline{\alpha}(s)$ if $r < s$, that

$$\alpha(x) = \inf\{\overline{\alpha}(r) : r > x, r \in \mathbb{Q}\}.$$

Let μ be the Lebesgue-Stieltjes measure associated with α. Prove that

$$\int f\, d\mu_n \to \int f\, d\mu$$

whenever f is a continuous real-valued function on $[0, 1]$.

Chapter 8

Properties of Lebesgue integrals

We present some propositions which imply that a function is zero a.e. and we give an approximation result.

8.1 Criteria for a function to be zero a.e.

The following two propositions are very useful.

Proposition 8.1 *Suppose f is real-valued and measurable and for every measurable set A we have $\int_A f\,d\mu = 0$. Then $f = 0$ almost everywhere.*

Proof. Let $A = \{x : f(x) > \varepsilon\}$. Then

$$0 = \int_A f \geq \int_A \varepsilon = \varepsilon\mu(A)$$

since $f\chi_A \geq \varepsilon\chi_A$. Hence $\mu(A) = 0$. We use this argument for $\varepsilon = 1/n$ and $n = 1, 2, \ldots$ to conclude

$$\begin{aligned}\mu\{x : f(x) > 0\} &= \mu(\cup_{n=1}^\infty \{x : f(x) > 1/n\})\\ &\leq \sum_{n=1}^\infty \mu(\{x : f(x) > 1/n\}) = 0.\end{aligned}$$

Similarly $\mu\{x : f(x) < 0\} = 0$. □

Proposition 8.2 *Suppose f is measurable and non-negative and $\int f \, d\mu = 0$. Then $f = 0$ almost everywhere.*

Proof. If f is not equal to 0 almost everywhere, there exists an n such that $\mu(A_n) > 0$ where $A_n = \{x : f(x) > 1/n\}$. But since f is non-negative,

$$0 = \int f \geq \int_{A_n} f \geq \frac{1}{n}\mu(A_n),$$

a contradiction. □

As a corollary to Proposition 8.1 we have the following.

Corollary 8.3 *Let m be Lebesgue measure and $a \in \mathbb{R}$. Suppose $f : \mathbb{R} \to \mathbb{R}$ is integrable and $\int_a^x f(y) \, dy = 0$ for all x. Then $f = 0$ a.e.*

Proof. For any interval $[c, d]$,

$$\int_c^d f = \int_a^d f - \int_a^c f = 0.$$

By linearity, if G is the finite union of disjoint intervals, then $\int_G f = 0$. By dominated convergence and Proposition 1.5, $\int_G f = 0$ for any open set G. Again by dominated convergence, if G_n are open sets decreasing to H, then $\int_H f = \lim_n \int_{G_n} f = 0$.

If E is any Borel measurable set, Proposition 4.14 tells us that there exists a sequence G_n of open sets that decrease to a set H where H differs from E by a null set. Then

$$\int_E f = \int f\chi_E = \int f\chi_H = \int_H f = 0.$$

This with Proposition 8.1 implies f is zero a.e. □

8.2 An approximation result

We give a result on approximating a function on $\mathbb{R}$ by continuous functions.

Theorem 8.4 *Suppose f is a Borel measurable real-valued integrable function on $\mathbb{R}$. Let $\varepsilon > 0$. Then there exists a continuous function g with compact support such that*

$$\int |f - g| < \varepsilon.$$

Proof. If we write $f = f^+ - f^-$, it is enough to find continuous functions g_1 and g_2 with compact support such that $\int |f^+ - g_1| < \varepsilon/2$ and $\int |f^- - g_2| < \varepsilon/2$ and to let $g = g_1 - g_2$. Hence we may assume $f \geq 0$.

By monotone convergence $\int f \cdot \chi_{[-n,n]}$ increases to $\int f$, so by taking n large enough, the difference of the integrals will be less than $\varepsilon/2$. If we find g continuous with compact support such that $\int |f \cdot \chi_{[-n,n]} - g| < \varepsilon/2$, then $\int |f - g| < \varepsilon$. Therefore we may in addition assume that f is 0 outside some bounded interval.

Suppose $f = \chi_A$, where A is a bounded Borel measurable set. We can choose G open and F closed such that $F \subset A \subset G$ and $m(G - F) < \varepsilon$ by Proposition 4.14. Without loss of generality, we may assume G is also a bounded set. Since F is compact, there is a minimum distance between F and G^c, say, δ. Let

$$g(x) = \Big(1 - \frac{\text{dist}\,(x, F)}{\delta}\Big)^+.$$

Then g is continuous, $0 \leq g \leq 1$, g is 1 on F, g is 0 on G^c, and g has compact support. We have

$$|g - \chi_A| \leq \chi_G - \chi_F,$$

so

$$\int |g - \chi_A| \leq \int (\chi_G - \chi_F) = m(G - F) < \varepsilon.$$

Thus our result holds for characteristic functions of bounded sets.

If $f = \sum_{i=1}^p a_i \chi_{A_i}$, where each A_i is contained in a bounded interval and each $a_i > 0$, and we find g_i continuous with compact support such that $\int |\chi_{A_i} - g_i| < \varepsilon/a_i p$, then $g = \sum_{i=1}^{p} a_i g_i$ will be the desired function. Thus our theorem holds for non-negative simple functions with compact support.

If f is non-negative and has compact support, we can find simple functions s_m supported in a bounded interval increasing to f whose integrals increase to $\int f$. Let s_m be a simple function such that $s_m \leq f$ and $\int s_m \geq \int f - \varepsilon/2$. We choose continuous g with compact support such that $\int |s_m - g| < \varepsilon/2$ using the preceding paragraphs, and then $\int |f - g| < \varepsilon$. □

The method of proof, where one proves a result for characteristic functions, then simple functions, then non-negative functions, and then finally integrable functions, is very common.

8.3 Exercises

Exercise 8.1 This exercise gives a change of variables formula in two simple cases. Show that if f is an integrable function on the reals and a is a nonzero real number, then

$$\int_{\mathbb{R}} f(x+a)\, dx = \int_{\mathbb{R}} f(x)\, dx$$

and

$$\int_{\mathbb{R}} f(ax)\, dx = a^{-1} \int_{\mathbb{R}} f(x)\, dx.$$

Exercise 8.2 Let $(X, \mathcal{A}, \mu)$ be a σ-finite measure space. Suppose f is non-negative and integrable. Prove that if $\varepsilon > 0$, there exists $A \in \mathcal{A}$ such that $\mu(A) < \infty$ and

$$\varepsilon + \int_A f\, d\mu > \int f\, d\mu.$$

Exercise 8.3 Suppose A is a Borel measurable subset of $[0,1]$, m is Lebesgue measure, and $\varepsilon \in (0,1)$. Prove that there exists a continuous function $f : [0,1] \to \mathbb{R}$ such that $0 \leq f \leq 1$ and

$$m(\{x : f(x) \neq \chi_A(x)\}) < \varepsilon.$$

Exercise 8.4 Suppose f is a non-negative integrable function on a measure space $(X, \mathcal{A}, \mu)$. Prove that

$$\lim_{t\to\infty} t\mu(\{x : f(x) \ge t\}) = 0.$$

Exercise 8.5 Find a non-negative function f on $[0,1]$ such that

$$\lim_{t\to\infty} tm(\{x : f(x) \ge t\}) = 0$$

but f is not integrable, where m is Lebesgue measure.

Exercise 8.6 Suppose μ is a finite measure. Prove that a measurable non-negative function f is integrable if and only if

$$\sum_{n=1}^{\infty} \mu(\{x : f(x) \ge n\}) < \infty.$$

Exercise 8.7 Let μ be a measure, not necessarily σ-finite, and suppose f is real-valued and integrable with respect to μ. Prove that $A = \{x : f(x) \ne 0\}$ has σ-finite measure, that is, there exists $F_n \uparrow A$ such that $\mu(F_n) < \infty$ for each n.

Exercise 8.8 Recall that a function $f : \mathbb{R} \to \mathbb{R}$ is *convex* if

$$f(\lambda x + (1-\lambda)y) \le \lambda f(x) + (1-\lambda) f(y)$$

whenever $x < y \in \mathbb{R}$ and $\lambda \in [0,1]$.
(1) Prove that if f is convex and $x \in \mathbb{R}$, there exists a real number c such that $f(y) \ge f(x) + c(y-x)$ for all $y \in \mathbb{R}$. Graphically, this says that the graph of f lies above the line with slope c that passes through the point $(x, f(x))$.
(2) Let $(X, \mathcal{A}, \mu)$ be a measure space, suppose $\mu(X) = 1$, and let $f : \mathbb{R} \to \mathbb{R}$ be convex. Let $g : X \to \mathbb{R}$ be integrable. Prove *Jensen's inequality*:

$$f\Big(\int g\,d\mu\Big) \le \int_X f \circ g\,d\mu.$$

Exercise 8.9 Suppose f is a real-valued function on $\mathbb{R}$ such that

$$f\Big(\int_0^1 g(x)\,dx\Big) \le \int_0^1 f(g(x))\,dx$$

whenever g is bounded and measurable. Prove that f is convex.

Exercise 8.10 Suppose $g : [0,1] \to \mathbb{R}$ is bounded and measurable and

$$\int_0^1 f(x)g(x)\,dx = 0$$

whenever f is continuous and $\int_0^1 f(x)\,dx = 0$. Prove that g is equal to a constant a.e.

Chapter 9

Riemann integrals

We compare the Lebesgue integral and the Riemann integral. We show that the Riemann integral of a function exists if and only if the set of discontinuities of the function have Lebesgue measure zero, and in that case the Riemann integral and Lebesgue integral agree.

9.1 Comparison with the Lebesgue integral

We only consider bounded measurable functions from $[a, b]$ into $\mathbb{R}$. If we are looking at the Lebesgue integral, we write $\int f$, while, temporarily, if we are looking at the Riemann integral, we write $R(f)$. Recall that the Riemann integral on $[a, b]$ is defined as follows: if $P = \{x_0, x_1, \ldots, x_n\}$ with $x_0 = a$ and $x_n = b$ is a partition of $[a, b]$, let

$$U(P, f) = \sum_{i=1}^{n} \Big[\sup_{x_{i-1} \le x \le x_i} f(x) \Big] (x_i - x_{i-1})$$

and

$$L(P, f) = \sum_{i=1}^{n} \Big[\inf_{x_{i-1} \le x \le x_i} f(x) \Big] (x_i - x_{i-1}).$$

Set

$$\overline{R}(f) = \inf\{U(P, f) : P \text{ is a partition}\}$$

and

$$\underline{R}(f) = \sup\{L(P, f) : P \text{ is a partition}\}.$$

The Riemann integral exists if $\overline{R}(f) = \underline{R}(f)$, and the common value is the Riemann integral, which we denote $R(f)$.

Theorem 9.1 *A bounded Borel measurable real-valued function f on $[a, b]$ is Riemann integrable if and only if the set of points at which f is discontinuous has Lebesgue measure 0, and in that case, the Riemann integral is equal in value to the Lebesgue integral.*

Proof. *Step 1.* First we show that if f is Riemann integrable, then f is continuous a.e. and $R(f) = \int f$. If P is a partition, define

$$T_P(x) = \sum_{i=1}^{n} [\sup_{x_{i-1} \le y \le x_i} f(y)] \chi_{[x_{i-1}, x_i)}(x),$$

and

$$S_P(x) = \sum_{i=1}^{n} [\inf_{x_{i-1} \le y \le x_i} f(y)] \chi_{[x_{i-1}, x_i)}(x).$$

We observe that $\int T_P = U(P, f)$ and $\int S_P = L(P, f)$.

If f is Riemann integrable, there exists a sequence of partitions Q_i such that $U(Q_i, f) \downarrow R(f)$ and a sequence Q_i' such that $L(Q_i', f) \uparrow R(f)$. It is not hard to check that adding points to a partition increases L and decreases U, so if we let $P_i = \cup_{j \le i}(Q_j \cup Q_j')$, then P_i is an increasing sequence of partitions, $U(P_i, f) \downarrow R(f)$ and $L(P_i, f) \uparrow R(f)$. We see also that $T_{P_i}(x)$ decreases at each point, say, to $T(x)$, and $S_{P_i}(x)$ increases at each point, say, to $S(x)$. Also $T(x) \ge f(x) \ge S(x)$. By dominated convergence (recall that f is bounded)

$$\int (T - S) = \lim_{i \to \infty} \int (T_{P_i} - S_{P_i}) = \lim_{i \to \infty} (U(P_i, f) - L(P_i, f)) = 0.$$

We conclude $T = S = f$ a.e. If x is not in the null set where $T(x) \ne S(x)$ nor in $\cup_i P_i$, which is countable and hence of Lebesgue measure 0, then $T_{P_i}(x) \downarrow f(x)$ and $S_{P_i}(x) \uparrow f(x)$. We claim that f is continuous at such x. To prove the claim, given ε, choose i large enough so that $T_{P_i}(x) - S_{P_i}(x) < \varepsilon$ and then choose δ small enough so that $(x - \delta, x + \delta)$ is contained in the subinterval of P_i that contains x. Finally, since

$$R(f) = \lim_{i \to \infty} U(P_i, f) = \lim_{i \to \infty} \int T_{P_i} = \int f$$

by the dominated convergence theorem, we see that the Riemann integral and Lebesgue integral agree.

Step 2. Now suppose that f is continuous a.e. Let $\varepsilon > 0$. Let P_i be the partition where we divide $[a, b]$ into 2^i equal parts. If x is not in the null set where f is discontinuous, nor in $\cup_{i=1}^{\infty} P_i$, then $T_{P_i}(x) \downarrow f(x)$ and $S_{P_i}(x) \uparrow f(x)$. By dominated convergence,

$$U(P_i, f) = \int T_{P_i} \to \int f$$

and

$$L(P_i, f) = \int S_{P_i} \to \int f.$$

This does it. □

Example 9.2 Let $[a, b] = [0, 1]$ and $f = \chi_A$, where A is the set of irrational numbers in $[0, 1]$. If $x \in [0, 1]$, every neighborhood of x contains both rational and irrational points, so f is continuous at no point of $[0, 1]$. Therefore f is not Riemann integrable.

Example 9.3 Define $f(x)$ on $[0, 1]$ to be 0 if x is irrational and to be $1/q$ if x is rational and equals p/q when in reduced form. f is discontinuous at every rational. If x is irrational and $\varepsilon > 0$, there are only finitely many rationals r for which $f(r) \geq \varepsilon$, so taking δ less than the distance from x to any of this finite collection of rationals shows that $|f(y) - f(x)| < \varepsilon$ if $|y - x| < \delta$. Hence f is continuous at x. Therefore the set of discontinuities is a countable set, hence of measure 0, hence f is Riemann integrable.

9.2 Exercises

Exercise 9.1 Find a measurable function $f : [0, 1] \to \mathbb{R}$ such that $\overline{R}(f) \neq \int_0^1 f(x)\,dx$ and $\underline{R}(f) \neq \int_0^1 f(x)\,dx$.

Exercise 9.2 Find a function $f : (0, 1] \to \mathbb{R}$ that is continuous, is not Lebesgue integrable, but where the improper Riemann integral exists. Thus we want f such that $\int_0^1 |f(x)|\, m(dx) = \infty$ but $\lim_{a\to 0+} R(f\chi_{[a,1]})$ exists.

Exercise 9.3 Suppose $f : [0,1] \to \mathbb{R}$ is integrable, f is bounded on $(a,1]$ for each $a > 0$, and the improper Riemann integral

$$\lim_{a\to 0+} R(f\chi_{(a,1]})$$

exists. Show that the limit is equal to $\int_0^1 f(x)\,dx$.

Exercise 9.4 Divide $[a,b]$ into 2^n equal subintervals and pick a point x_i out of each subinterval. Let μ_n be the measure defined by

$$\mu_n(dx) = 2^{-n}\sum_{i=1}^{2^n}\delta_{x_i}(dx),$$

where δ_y is point mass at y. Note that if f is a bounded measurable real-valued function on $[a,b]$, then

$$\int_a^b f(x)\,\mu_n(dx) = \sum_{i=1}^{2^n} f(x_i)2^{-n} \tag{9.1}$$

is a Riemann sum approximation to $R(f)$.
(1) Prove that $\mu_n([0,x]) \to m([0,x])$ for every $x \in [0,1]$. Conclude by Exercise 7.24 that $\int f\,d\mu_n \to \int_0^1 f\,dx$ whenever f is continuous.
(2) Use Exercise 7.25 to see that if f is a bounded and measurable function on $[a,b]$ whose set of discontinuities has measure 0, then the Riemann sum approximation of f given in (9.1) converges to the Lebesgue integral of f. This provides an alternative proof of Step 2 of Theorem 9.1.

Exercise 9.5 Let f be a bounded, real-valued, and measurable function. Prove that if

$$\overline{f} = \lim_{\delta\to 0}\sup_{|y-x|<\delta, a\le y\le b} f(y),$$

then $\overline{f} = T$ a.e., where we use the notation of Theorem 9.1. Conclude $\overline{f}$ is Lebesgue measurable.

Exercise 9.6 Define $\underline{f} = \lim_{\delta\to 0}\inf_{|y-x|<\delta, a\le y\le b} f(y)$ and let $\overline{f}$ be defined as in Exercise 9.5.
(1) Suppose that the set of discontinuities of a bounded real-valued

measurable function f has positive Lebesgue measure. Prove that there exists $\varepsilon > 0$ such that if

$$A_\varepsilon = \{x \in [a,b] : \overline{f}(x) - \underline{f}(x) > \varepsilon\},$$

then $m(A_\varepsilon) > 0$.
(2) Prove that $U(P,f) - L(P,f) > \varepsilon m(A_\varepsilon)$ for every partition P on $[a,b]$, using the notation of Theorem 9.1. Conclude that f is not Riemann integrable. This provides another proof of Step 1 of Theorem 9.1.

Exercise 9.7 A real-valued function on a metric space is *lower semicontinuous* if $\{x : f(x) > a\}$ is open whenever $a \in \mathbb{R}$ and *upper semicontinuous* if $\{x : f(x) < a\}$ is open whenever $a \in \mathbb{R}$.
(1) Prove that if f_n is a sequence of real-valued continuous functions increasing to f, then f is lower semicontinuous.
(2) Find a bounded lower semicontinuous function $f : [0,1] \to \mathbb{R}$ such that f is continuous everywhere except at $x = 1/2$.
(3) Find a bounded lower semicontinuous real-valued function f defined on $[0,1]$ such that the set of discontinuities of f is equal to the set of rationals in $[0,1]$.
(4) Find a bounded lower semicontinuous function $f : [0,1] \to \mathbb{R}$ such that the set of discontinuities of f has positive measure.
(5) Does there exist a bounded lower semicontinuous function $f : [0,1] \to \mathbb{R}$ such that f is discontinuous a.e.?

Exercise 9.8 Find a sequence f_n of continuous functions mapping $[0,1]$ into $[0,1]$ such that the f_n increase to a bounded function f which is not Riemann integrable. Such an example shows there is no monotone convergence theorem or dominated convergence theorem for Riemann integrals.

Exercise 9.9 Let $M > 0$ and let $\mathcal{B}$ be the σ-algebra on $[-M,M]^2$ generated by the collection of sets of the form $[a,b] \times [c,d]$ with $-M \le a \le b \le M$ and $-M \le c \le d \le M$. Suppose μ is a measure on $([-M,M]^2, \mathcal{B})$ such that

$$\mu([a,b] \times [c,d]) = (b-a)(d-c).$$

(We will construct such a measure μ in Chapter 11.) Prove that if f is continuous with support in $[-M,M]^2$, then the Lebesgue integral of f with respect to μ is equal to the double (Riemann) integral of f and the two multiple (Riemann) integrals of f.

Chapter 10

Types of convergence

There are various ways in which a sequence of functions f_n can converge, and we compare some of them.

10.1 Definitions and examples

Definition 10.1 If μ is a measure, we say a sequence of measurable functions f_n *converges almost everywhere* to f and write $f_n \to f$ a.e. if there is a set of measure 0 such that for x not in this set we have $f_n(x) \to f(x)$.

We say f_n *converges in measure* to f if for each $\varepsilon > 0$

$$\mu(\{x : |f_n(x) - f(x)| > \varepsilon\}) \to 0$$

as $n \to \infty$.

Let $1 \le p < \infty$. We say f_n *converges in L^p* to f if

$$\int |f_n - f|^p \, d\mu \to 0$$

as $n \to \infty$.

Proposition 10.2 *Suppose μ is a finite measure.*

(1) If $f_n \to f$, a.e., then f_n converges to f in measure.

(2) If $f_n \to f$ in measure, there is a subsequence n_j such that $f_{n_j} \to f$, a.e.

Proof. Let $\varepsilon > 0$ and suppose $f_n \to f$ a.e. If

$$A_n = \{x : |f_n(x) - f(x)| > \varepsilon\},$$

then $\chi_{A_n} \to 0$ a.e., and by dominated convergence,

$$\mu(A_n) = \int \chi_{A_n}(x)\,\mu(dx) \to 0.$$

This proves (1).

To prove (2), suppose $f_n \to f$ in measure, let $n_1 = 1$, and choose $n_j > n_{j-1}$ by induction so that

$$\mu(\{x : |f_{n_j}(x) - f(x)| > 1/j\}) \le 2^{-j}.$$

Let $A_j = \{x : |f_{n_j}(x) - f(x)| > 1/j\}$. Then $\mu(A_j) \le 2^{-j}$, and if we set

$$A = \cap_{k=1}^\infty \cup_{j=k}^\infty A_j,$$

then by Proposition 3.5

$$\mu(A) = \lim_{k\to\infty} \mu(\cup_{j=k}^\infty A_j) \le \lim_{k\to\infty} \sum_{j=k}^\infty \mu(A_j) \le \lim_{k\to\infty} 2^{-k+1} = 0.$$

Therefore A has measure 0. If $x \notin A$, then $x \notin \cup_{j=k}^\infty A_j$ for some k, and so $|f_{n_j}(x) - f(x)| \le 1/j$ for $j \ge k$. This implies $f_{n_j} \to f$ on A^c. $\square$

Example 10.3 Part (1) of the above proposition is not true if $\mu(X) = \infty$. To see this, let $X = \mathbb{R}$ and let $f_n = \chi_{(n,n+1)}$. We have $f_n \to 0$ a.e., but f_n does not converge in measure.

The next proposition compares convergence in L^p to convergence in measure. Before we prove this, we prove an easy preliminary result known as *Chebyshev's inequality.*

Lemma 10.4 *If* $1 \le p < \infty$, *then*

$$\mu(\{x : |f(x)| \ge a\}) \le \frac{\int |f|^p}{a^p}.$$

Proof. Let $A = \{x : |f(x)| \geq a\}$. Since $\chi_A \leq |f|^p \chi_A / a^p$, we have

$$\mu(A) \leq \int_A \frac{|f|^p}{a^p}\, d\mu \leq \frac{1}{a^p} \int |f|^p \, d\mu.$$

This is what we wanted. □

Proposition 10.5 *If f_n converges to f in L^p, then it converges in measure.*

Proof. If $\varepsilon > 0$, by Chebyshev's inequality

$$\mu(\{x : |f_n(x) - f(x)| > \varepsilon\}) = \mu(\{x : |f_n(x) - f(x)|^p > \varepsilon^p\}) \\ \leq \frac{\int |f_n - f|^p}{\varepsilon^p} \to 0$$

as required. □

Example 10.6 Let $f_n = n^2 \chi_{(0,1/n)}$ on $[0,1]$ with Lebesgue measure. This gives an example where f_n converges to 0 a.e. and in measure, but does not converge in L^p for any $p \geq 1$.

Example 10.7 We give an example where $f_n \to f$ in measure and in L^p, but not almost everywhere. Let $S = \{e^{i\theta} : 0 \leq \theta < 2\pi\}$ be the unit circle in the complex plane and define

$$\mu(A) = m(\{\theta \in [0, 2\pi) : e^{i\theta} \in A\})$$

to be arc length measure on S, where m is Lebesgue measure on $[0, 2\pi)$.

Let $X = S$ and let $f_n(x) = \chi_{F_n}(x)$, where

$$F_n = \Big\{ e^{i\theta} : \sum_{j=1}^{n} \frac{1}{j} \leq \theta \leq \sum_{j=1}^{n+1} \frac{1}{j} \Big\}.$$

Let $f(e^{i\theta}) = 0$ for all θ.

Then $\mu(F_n) \leq 1/(n+1) \to 0$, so $f_n \to f$ in measure. Also, since f_n is either 1 or 0,

$$\int |f_n - f|^p \, d\mu = \int \chi_{F_n} \, d\mu = \mu(F_n) \to 0.$$

But because $\sum_{j=1}^\infty 1/j = \infty$, each point of S is in infinitely many F_n, and each point of S is in $S - F_n$ for infinitely many n, so f_n does not converge to f at any point.

The F_n are arcs whose length tends to 0, but such that $\cup_{n\geq m} F_n$ contains S for each m.

The following is known as *Egorov's theorem.*

Theorem 10.8 *Suppose μ is a finite measure, $\varepsilon > 0$, and $f_n \to f$ a.e. Then there exists a measurable set A such that $\mu(A) < \varepsilon$ and $f_n \to f$ uniformly on A^c.*

This type of convergence is sometimes known as *almost uniform convergence.* Egorov's theorem is not as useful for solving problems as one might expect, and students have a tendency to try to use it when other methods work much better.

Proof. Let

$$A_{nk} = \cup_{m=n}^\infty \{x : |f_m(x) - f(x)| > 1/k\}.$$

For fixed k, A_{nk} decreases as n increases. The intersection $\cap_n A_{nk}$ has measure 0 because for almost every x, $|f_m(x) - f(x)| \leq 1/k$ if m is sufficiently large. Therefore $\mu(A_{nk}) \to 0$ as $n \to \infty$. We can thus find an integer n_k such that $\mu(A_{n_k k}) < \varepsilon 2^{-k}$. Let

$$A = \cup_{k=1}^\infty A_{n_k k}.$$

Hence $\mu(A) < \varepsilon$. If $x \notin A$, then $x \notin A_{n_k k}$, and so $|f_n(x) - f(x)| \leq 1/k$ if $n \geq n_k$. Thus $f_n \to f$ uniformly on A^c. □

10.2 Exercises

Exercise 10.1 Suppose that f_n is a sequence that is Cauchy in measure. This means that given ε and $a > 0$, there exists N such that if $m, n \geq N$, then

$$\mu(\{x : |f_n(x) - f_m(x)| > a\}) < \varepsilon.$$

Prove that f_n converges in measure.

Exercise 10.2 Suppose $\mu(X) < \infty$. Define

$$d(f, g) = \int \frac{|f-g|}{1+|f-g|}\, d\mu.$$

Prove that d is a metric on the space of measurable functions, except for the fact that $d(f, g) = 0$ only implies that $f = g$ a.e., not necessarily everywhere. Prove that $f_n \to f$ in measure if and only if $d(f_n, f) \to 0$.

Exercise 10.3 Prove that if $f_n \to f$ in measure and each f_n is non-negative, then

$$\int f \le \liminf_{n\to\infty} \int f_n.$$

Exercise 10.4 Prove that if A_n is measurable and $\mu(A_n) < \infty$ for each n and χ_{A_n} converges to f in measure, then there exists a measurable set A such that $f = \chi_A$ a.e.

Exercise 10.5 Suppose for each ε there exists a measurable set F such that $\mu(F^c) < \varepsilon$ and f_n converges to f uniformly on F. Prove that f_n converges to f a.e.

Exercise 10.6 Suppose that f_n and f are measurable functions such that for each $\varepsilon > 0$ we have

$$\sum_{n=1}^{\infty} \mu(\{x : |f_n(x) - f(x)| > \varepsilon\}) < \infty.$$

Prove that $f_n \to f$ a.e.

Exercise 10.7 Let f_n be a sequence of measurable functions and define

$$g_n(x) = \sup_{m\ge n} |f_m(x) - f_n(x)|.$$

Prove that if g_n converges in measure to 0, then f_n converges a.e.

Exercise 10.8 If $(X, \mathcal{A}, \mu)$ is a measure space and f_n is a sequence of real-valued measurable functions such that $\int f_n g\, d\mu$ converges to 0 for every integrable g, is it necessarily true that f_n converges to 0 in measure? If not, give a counterexample.

Chapter 11

Product measures

We have defined Lebesgue measure on the line. Now we give a method for constructing measures on the plane, in n-dimensional Euclidean spaces, and many other product spaces. The main theorem, the Fubini theorem, which allows one to interchange the order of integration, is one of the most important theorems in real analysis.

11.1 Product σ-algebras

Suppose $(X, \mathcal{A}, \mu)$ and $(Y, \mathcal{B}, \nu)$ are two measure spaces and suppose also that μ and ν are σ-finite measures. A *measurable rectangle* is a set of the form $A \times B$, where $A \in \mathcal{A}$ and $B \in \mathcal{B}$.

Let $\mathcal{C}_0$ be the collection of finite unions of disjoint measurable rectangles. Thus every element of $\mathcal{C}_0$ is of the form $\cup_{i=1}^n (A_i \times B_i)$, where $A_i \in \mathcal{A}$, $B_i \in \mathcal{B}$, and if $i \neq j$, then $(A_i \times B_i) \cap (A_j \times B_j) = \emptyset$. Since $(A \times B)^c = (A \times B^c) \cup (A^c \times Y)$ and the intersection of two measurable rectangles is a measurable rectangle, it is easy to check that $\mathcal{C}_0$ is an algebra of sets. We define the *product σ-algebra*

$$\mathcal{A} \times \mathcal{B} = \sigma(\mathcal{C}_0).$$

If $E \in \mathcal{A} \times \mathcal{B}$, we define the *$x$-section* of E by

$$s_x(E) = \{y \in Y : (x, y) \in E\}$$

and similarly define the *y-section*:

$$t_y(E) = \{x : (x, y) \in E\}.$$

Given a function $f : X \times Y \to \mathbb{R}$ that is $\mathcal{A} \times \mathcal{B}$ measurable, for each x and y we define $S_x f : Y \to \mathbb{R}$ and $T_y f : X \to \mathbb{R}$ by

$$S_x f(y) = f(x, y), \qquad T_y f(x) = f(x, y).$$

Lemma 11.1 *(1) If $E \in \mathcal{A} \times \mathcal{B}$, then $s_x(E) \in \mathcal{B}$ for each x and $t_y(E) \in \mathcal{A}$ for each y.*
(2) If f is $\mathcal{A} \times \mathcal{B}$ measurable, then $S_x f$ is $\mathcal{B}$ measurable for each x and $T_y f$ is $\mathcal{A}$ measurable for each y.

Proof. (1) Let $\mathcal{C}$ be the collection of sets in $\mathcal{A} \times \mathcal{B}$ for which $s_x(E) \in \mathcal{B}$ for each x. We will show that $\mathcal{C}$ is a σ-algebra containing the measurable rectangles, and hence is all of $\mathcal{A} \times \mathcal{B}$.

If $E = A \times B$, then $s_x(E)$ is equal to B if $x \in A$ and equal to $\emptyset$ if $x \notin A$. Hence $s_x(E) \in \mathcal{B}$ for each x when E is a measurable rectangle.

If $E \in \mathcal{C}$, then $y \in s_x(E^c)$ if and only if $(x, y) \in E^c$, which happens if and only if $y \notin s_x(E)$. Therefore $s_x(E^c) = (s_x(E))^c$, and $\mathcal{C}$ is closed under the operation of taking complements. Similarly, it is easy to see that $s_x(\cup_{i=1}^\infty E_i) = \cup_{i=1}^\infty s_x(E_i)$, and so $\mathcal{C}$ is closed under the operation of countable unions.

Therefore $\mathcal{C}$ is a σ-algebra containing the measurable rectangles, and hence is equal to $\mathcal{A} \times \mathcal{B}$. The argument for $t_y(E)$ is the same.

(2) Fix x. If $f = \chi_E$ for $E \in \mathcal{A} \times \mathcal{B}$, note that $S_x f(y) = \chi_{s_x(E)}(y)$, which is $\mathcal{B}$ measurable. By linearity, $S_x f$ is $\mathcal{B}$ measurable when f is a simple function. If f is non-negative, take $\mathcal{A} \times \mathcal{B}$ measurable simple functions r_n increasing to f, and since $S_x r_n \uparrow S_x f$, then $S_x f$ is $\mathcal{B}$ measurable. Writing $f = f^+ - f^-$ and using linearity again shows that $S_x f$ is $\mathcal{B}$ measurable. The argument for $T_y f$ is the same. $\square$

Let $E \in \mathcal{A} \times \mathcal{B}$ and let

$$h(x) = \nu(s_x(E)), \qquad k(y) = \mu(t_y(E)).$$

Proposition 11.2 *(1) h is $\mathcal{A}$ measurable and k is $\mathcal{B}$ measurable.*
(2) We have

$$\int h(x)\,\mu(dx) = \int k(y)\,\nu(dy). \tag{11.1}$$

Since $\chi_{s_x(E)}(y) = S_x\chi_E(y)$ for all x and y, (11.1) could be written as

$$\int \Big[\int S_x\chi_E(y)\,\nu(dy)\Big]\,\mu(dx) = \int \Big[\int T_y\chi_E(x)\,\mu(dx)\Big]\,\nu(dy).$$

We will usually write this as

$$\int\int \chi_E(x,y)\,\nu(dy)\,\mu(dx) = \int\int \chi_E(x,y)\,\mu(dx)\,\nu(dy).$$

Proof. First suppose μ and ν are finite measures. Let $\mathcal{C}$ be the collection of sets in $\mathcal{A}\times\mathcal{B}$ for which (1) and (2) hold. We will prove that $\mathcal{C}$ contains $\mathcal{C}_0$ and is a monotone class. This will prove that $\mathcal{C}$ is the smallest σ-algebra containing $\mathcal{C}_0$ and hence is equal to $\mathcal{A}\times\mathcal{B}$.

If $E = A\times B$, with $A\in\mathcal{A}$ and $B\in\mathcal{B}$, then $h(x) = \chi_A(x)\nu(B)$, which is $\mathcal{A}$ measurable, and $\int h(x)\,\mu(dx) = \mu(A)\nu(B)$. Similarly, $k(y) = \mu(A)\chi_B(y)$ is $\mathcal{B}$ measurable and $\int k(y)\,\nu(dy) = \mu(A)\nu(B)$. Therefore (1) and (2) hold for measurable rectangles.

If $E = \cup_{i=1}^n E_i$, where each E_i is a measurable rectangle and the E_i are disjoint, then $s_x(E) = \cup_{i=1}^n s_x(E_i)$, and since the $s_x(E_i)$ are disjoint, then

$$h(x) = \nu(s_x(E)) = \nu(\cup_{i=1}^n s_x(E_i)) = \sum_{i=1}^n \nu(s_x(E_i)).$$

This shows that h is $\mathcal{A}$ measurable, since it is the sum of $\mathcal{A}$ measurable functions. Similarly $k(y)$ is $\mathcal{B}$ measurable. If we let $h_i(x) = \nu(s_x(E_i))$ and define $k_i(y)$ similarly, then

$$\int h_i(x)\,\mu(dx) = \int k_i(y)\,\nu(dy)$$

by the preceding paragraph, and then (2) holds for E by linearity. Therefore $\mathcal{C}$ contains $\mathcal{C}_0$.

Suppose $E_n \uparrow E$ and each $E_n\in\mathcal{C}$. If we let $h_n(x) = \nu(s_x(E_n))$ and let $k_n(y) = \mu(t_n(E_n))$, then $h_n\uparrow h$ and $k_n\uparrow k$. Therefore h is

$\mathcal{A}$ measurable and k is $\mathcal{B}$ measurable. We have (11.1) holding when h and k are replaced by h_n and k_n, resp. We let $n \to \infty$ and use the monotone convergence theorem to see that (11.1) holds with h and k.

If $E_n \downarrow E$ with each $E_n \in \mathcal{C}$, almost the same argument shows that h and k are measurable with respect to $\mathcal{A}$ and $\mathcal{B}$, and that (11.1) holds. The only difference is that we use the dominated convergence theorem in place of the monotone convergence theorem. This is where we need μ and ν to be finite measures.

We have shown $\mathcal{C}$ is a monotone class containing $\mathcal{C}_0$. By the monotone class theorem (Theorem 2.10), $\mathcal{C}$ is equal to $\sigma(\mathcal{C}_0)$, which is $\mathcal{A} \times \mathcal{B}$.

Finally suppose μ and ν are σ-finite. Then there exist $F_i \uparrow X$ and $G_i \uparrow Y$ such that each F_i is $\mathcal{A}$ measurable and has finite μ measure and each G_i is $\mathcal{B}$ measurable and has finite ν measure. Let $\mu_i(A) = \mu(A \cap F_i)$ for each $A \in \mathcal{A}$ and $\nu_i(A) = \nu(A \cap G_i)$ for each $B \in \mathcal{B}$. Let $h_i(x) = \nu_i(s_x(E)) = \nu(s_x(E) \cap G_i)$ and similarly define $k_i(y)$. By what we have proved above, h_i is $\mathcal{A}$ measurable, k_i is $\mathcal{B}$ measurable, and (11.1) holds if we replace h and k by h_i and k_i, resp. Now $h_i \uparrow h$ and $k_i \uparrow k$, which proves the measurability of h and k. Applying the monotone convergence theorem proves that (11.1) holds with h and k. □

We now define $\mu \times \nu$ by

$$\mu \times \nu(E) = \int h(x)\,\mu(dx) = \int k(y)\,\nu(dy). \tag{11.2}$$

Clearly $\mu \times \nu(\emptyset) = 0$. If $E_1, \ldots, E_n$ are disjoint and in $\mathcal{A} \times \mathcal{B}$ and $E = \cup_{i=1}^n E_i$, then we saw in the proof of Proposition 11.2 that $\nu(s_x(E)) = \sum_{i=1}^n \nu(s_x(E_i))$. We conclude that

$$\begin{aligned}\mu \times \nu(E) &= \int \nu(s_x(E))\,\mu(dx) = \sum_{i=1}^n \int \nu(s_x(E_i))\,\mu(dx)\\ &= \sum_{i=1}^n \mu \times \nu(E_i),\end{aligned}$$

or $\mu \times \nu$ is finitely additive. If $E_n \uparrow E$ with each $E_n \in \mathcal{A} \times \mathcal{B}$ and we let $h_n(x) = \nu(s_x(E_n))$, then $h_n \uparrow h$, and by monotone convergence, $\mu \times \nu(E_n) \uparrow \mu \times \nu(E)$. Therefore $\mu \times \nu$ is a measure.

Note that if $E = A \times B$ is a measurable rectangle, then $h(x) = \chi_A(x)\nu(B)$ and so

$$\mu \times \nu(A \times B) = \mu(A)\nu(B),$$

which is what it should be.

11.2 The Fubini theorem

The main result of this chapter is the *Fubini theorem*, which allows one to interchange the order of integration. This is sometimes called the *Fubini-Tonelli theorem*.

Theorem 11.3 *Suppose $f : X \times Y \to \mathbb{R}$ is measurable with respect to $\mathcal{A} \times \mathcal{B}$. If either*
(a) f is non-negative, or
(b) $\int |f(x,y)|\, d(\mu \times \nu)(x,y) < \infty$,
then
(1) for each x, the function $y \mapsto f(x,y)$ is measurable with respect to $\mathcal{B}$;
(2) for each y, the function $x \mapsto f(x,y)$ is measurable with respect to $\mathcal{A}$;
(3) the function $g(x) = \int f(x,y)\,\nu(dy)$ is measurable with respect to $\mathcal{A}$;
(4) the function $h(y) = \int f(x,y)\,\mu(dx)$ is measurable with respect to $\mathcal{B}$;
(5) we have

$$\begin{aligned}\int f(x,y)\, d(\mu \times \nu)(x,y) &= \int \Big[\int f(x,y)\, d\mu(x) \Big]\, d\nu(y) \qquad (11.3)\\ &= \int \Big[\int f(x,y)\, d\nu(y) \Big]\, \mu(dx).\end{aligned}$$

The last integral in (11.3) should be interpreted as

$$\int \Big[\int S_x f(y)\, \nu(dy) \Big]\, \mu(dx)$$

and similarly for the second integral in (11.3). Since no confusion results, most often the brackets are omitted in (11.3).

Proof. If f is the characteristic function of a set in $\mathcal{A} \times \mathcal{B}$, then (1)–(5) are merely a restatement of Lemma 11.1 and Proposition

11.2. By linearity, (1)–(5) hold if f is a simple function. Since the increasing limit of measurable functions is measurable, then writing a non-negative function as the increasing limit of simple functions and using the monotone convergence theorem, we see that (1)–(5) hold when f is non-negative. In the case where $\int |f|\, d(\mu\times\nu) < \infty$, writing $f = f^+ - f^-$ and using linearity proves (1)–(5) for this case, too. $\square$

Observe that if we know

$$\int\int |f(x,y)|\,\mu(dx)\,\nu(dy) < \infty,$$

then since $|f(x,y)|$ is non-negative the Fubini theorem tells us that

$$\int |f(x,y)|\,d(\mu\times\nu) = \int\int |f(x,y)|\,\mu(dx)\,\nu(dy) < \infty$$

We can then apply the Fubini theorem again to conclude

$$\int f(x,y)\,d(\mu\times\nu) = \int\int f(x,y)\,d\mu\,d\nu = \int\int f(x,y)\,d\nu\,d\mu.$$

Thus in the hypotheses of the Fubini theorem, we could as well assume $\int\int |f(x,y)|\,d\mu\,d\nu < \infty$ or $\int\int |f(x,y)|\,d\nu\,d\mu < \infty$.

When f is measurable with respect to $\mathcal{A}\times\mathcal{B}$, we sometimes say that f is *jointly measurable*.

Even when $(X,\mathcal{A},\mu)$ and $(Y,\mathcal{B},\nu)$ are complete, it will not be the case in general that $(X\times Y,\mathcal{A}\times\mathcal{B},\mu\times\nu)$ is complete. For example, let $(X,\mathcal{A},\mu) = (Y,\mathcal{B},\nu)$ be Lebesgue measure on $[0,1]$ with the Lebesgue σ-algebra. Let A be a non-measurable set in $[0,1]$ and let $E = A\times\{1/2\}$. Then E is not a measurable set with respect to $\mathcal{A}\times\mathcal{B}$, or else $A = t_{1/2}(E)$ would be in $\mathcal{A}$ by Lemma 11.1. On the other hand, $E\subset[0,1]\times\{1/2\}$, which has zero measure with respect to $\mu\times\nu$, so E is a null set.

One can take the completion of $(X\times Y,\mathcal{A}\times\mathcal{B},\mu\times\nu)$ without great difficulty. See [4] for details.

There is no difficulty extending the Fubini theorem to the product of n measures. If we have $\mu_1,\ldots,\mu_n$ all equal to m, Lebesgue measure on $\mathbb{R}$ with the Lebesgue σ-algebra $\mathcal{L}$, then the completion of $(\mathbb{R}^n,\mathcal{L}\times\cdots\times\mathcal{L}, m\times\cdots\times m)$ is called *n-dimensional Lebesgue measure*.

For a general change of variables theorem, see [1].

11.3 Examples

We give two examples to show that the hypotheses of the Fubini theorem are necessary.

Example 11.4 Let $X = Y = [0,1]$ and μ and ν both being Lebesgue measure. Let g_i be continuous functions with support in $(1/(i+1), 1/i)$ such that $\int_0^1 g_i(x)\,dx = 1$, $i = 1, 2, \ldots$. Let

$$f(x,y) = \sum_{i=1}^{\infty} [g_i(x) - g_{i+1}(x)] g_i(y).$$

For each point (x,y) at most two terms in the sum are nonzero, so the sum is actually a finite one. If we first integrate with respect to y, we get

$$\int_0^1 f(x,y)\,dy = \sum_{i=1}^{\infty} [g_i(x) - g_{i+1}(x)].$$

This is a telescoping series, and sums to $g_1(x)$. Therefore

$$\int_0^1 \int_0^1 f(x,y)\,dy\,dx = \int_0^1 g_1(x)\,dx = 1.$$

On the other hand, integrating first with respect to x gives 0, so

$$\int_0^1 \int_0^1 f(x,y)\,dx\,dy = 0.$$

This doesn't contradict the Fubini theorem because

$$\int_0^1 \int_0^1 |f(x,y)|\,dx\,dy = \infty.$$

Example 11.5 For this example, you have to take on faith a bit of set theory. There exists a set X together with a partial order "$\le$" such that X is uncountable but for any $y \in X$, the set $\{x \in X : x \le y\}$ is countable. An example is to let X be the set of countable ordinals. Define μ on X by $\mu(A) = 0$ if A is countable and 1 if A is uncountable. The σ-algebra is the collection of subsets A of X such that either A or A^c is countable. Define f on $X \times X$ by $f(x,y) = 1$ if $x \le y$ and zero otherwise. Then $\int \int f(x,y)\,dy\,dx = 1$ but $\int \int f(x,y)\,dx\,dy = 0$. The reason there is no contradiction is that f is not measurable with respect to the product σ-algebra.

11.4 Exercises

Exercise 11.1 State and prove a version of the Fubini theorem for complex-valued functions.

Exercise 11.2 Let $(X, \mathcal{A})$ and $(Y, \mathcal{B})$ be two measurable spaces and let $f \geq 0$ be measurable with respect to $\mathcal{A} \times \mathcal{B}$. Let $g(x) = \sup_{y \in Y} f(x, y)$ and suppose $g(x) < \infty$ for each x. Is g necessarily measurable with respect to $\mathcal{A}$? If not, find a counterexample.

Exercise 11.3 Prove the equality

$$\int_{-\infty}^{\infty} |f(x)|\, dx = \int_0^{\infty} m(\{x : |f(x)| \geq t\})\, dt,$$

where m is Lebesgue measure.

Exercise 11.4 Let A be a Lebesgue measurable subset of $[0, 1]^2$ with $m_2(A) = 1$, where m_2 is two-dimensional Lebesgue measure. Show that for almost every $x \in [0, 1]$ (with respect to one dimensional Lebesgue measure) the set $s_x(A)$ has one-dimensional Lebesgue measure one.

Exercise 11.5 Let $f : [0, 1]^2 \to \mathbb{R}$ be such that for every $x \in [0, 1]$ the function $y \to f(x, y)$ is Lebesgue measurable on $[0, 1]$ and for every $y \in [0, 1]$ the function $x \to f(x, y)$ is continuous on $[0, 1]$. Prove that f is measurable with respect to the the completion of the product σ-algebra $\mathcal{L} \times \mathcal{L}$ on $[0, 1]^2$. Here $\mathcal{L}$ is the Lebesgue σ-algebra on $[0, 1]$.

Exercise 11.6 Suppose f is real-valued and integrable with respect to two-dimensional Lebesgue measure on $[0, 1]^2$ and

$$\int_0^a \int_0^b f(x, y)\, dy\, dx = 0$$

for all $a \in [0, 1]$ and $b \in [0, 1]$. Prove that $f = 0$ a.e.

Exercise 11.7 Prove that

$$\int_0^1 \int_0^1 \frac{x^2 - y^2}{(x^2 + y^2)^2} \log(4 + \sin x)\, dy\, dx = \int_0^1 \int_0^1 \frac{x^2 - y^2}{(x^2 + y^2)^2} \log(4 + \sin x)\, dx\, dy.$$

Exercise 11.8 Let $X = Y = [0,1]$ and let $\mathcal{B}$ be the Borel σ-algebra. Let m be Lebesgue measure and μ counting measure on $[0,1]$.
(1) If $D = \{(x,y) : x = y\}$, show that D is measurable with respect to $\mathcal{B} \times \mathcal{B}$.
(2) Show that

$$\int_X \int_Y \chi_D(x,y)\,\mu(dy)\,m(dx) \neq \int_Y \int_X \chi_D(x,y)\,m(dx)\,\mu(dy).$$

Why does this not contradict the Fubini theorem?

Exercise 11.9 Let $X = Y = \mathbb{R}$ and let $\mathcal{B}$ be the Borel σ-algebra. Define

$$f(x,y) = \begin{cases} 1, & x \geq 0 \text{ and } x \leq y < x+1; \\ -1, & x \geq 0 \text{ and } x+1 \leq y < x+2; \\ 0, & \text{otherwise.} \end{cases}$$

Show that

$$\int \int f(x,y)\,dy\,dx \neq \int \int f(x,y)\,dx\,dy.$$

Why does this not contradict the Fubini theorem?

Exercise 11.10 Find a real-valued function f that is integrable on $[0,1]^2$ such that

$$\int_0^a \int_0^1 f(x,y)\,dy\,dx = 0, \qquad \int_0^1 \int_0^b f(x,y)\,dy\,dx = 0$$

for every $a, b \in [0,1]$, but f is not zero almost everywhere with respect to 2-dimensional Lebesgue measure.

Exercise 11.11 Let μ be a finite measure on $\mathbb{R}$ and let $\alpha(x) = \mu((-\infty, x])$. Show

$$\int [\alpha(x+c) - \alpha(x)]\,dx = c\mu(\mathbb{R}).$$

Exercise 11.12 Use

$$\frac{1}{x} = \int_0^\infty e^{-xy}\,dy$$

and the Fubini theorem to calculate

$$\int_0^b \int_0^\infty e^{-xy} \sin x \, dy \, dx$$

two different ways. Then prove that

$$\lim_{b\to\infty} \int_0^b \frac{\sin x}{x} \, dx = \frac{\pi}{2}.$$

Recall that

$$\int e^{au} \sin u \, du = \frac{e^{au}(a \sin u - \cos u)}{1+a^2} + C.$$

Exercise 11.13 Let $X = \{1, 2, \ldots\}$ and let μ be counting measure on X. Define $f : X \times X \to \mathbb{R}$ by

$$f(x,y) = \begin{cases} 1, & x = y; \\ -1, & x = y + 1; \\ 0, & \text{otherwise.} \end{cases}$$

Show that

$$\int_X \int_X f(x,y) \, \mu(dx) \, \mu(dy) \neq \int_X \int_X f(x,y) \, \mu(dy) \, \mu(dx).$$

Why is this not a contradiction to the Fubini theorem?

Exercise 11.14 Let $\{a_n\}$ and $\{r_n\}$ be two sequences of real numbers such that $\sum_{n=1}^\infty |a_n| < \infty$. Prove that

$$\sum_{n=1}^\infty \frac{a_n}{\sqrt{|x - r_n|}}$$

converges absolutely for almost every $x \in \mathbb{R}$.

Exercise 11.15 Let $(X, \mathcal{A}, \mu)$ and $(Y, \mathcal{B}, \nu)$ be measure spaces. Prove that if λ is a measure on $\mathcal{A} \times \mathcal{B}$ such that

$$\lambda(A \times B) = \mu(A)\nu(B)$$

whenever $A \in \mathcal{A}$ and $B \in \mathcal{B}$, then $\lambda = \mu \times \nu$ on $\mathcal{A} \times \mathcal{B}$.

Exercise 11.16 Let S be the unit circle $\{e^{i\theta} : 0 \leq \theta < 2\pi\}$ and define a measure μ on S by $\mu(A) = m(\{\theta : e^{i\theta} \in A\})$, where m is Lebesgue measure on $[0, 2\pi)$. Let m_2 be two-dimensional Brownian motion. Show that if A is a Borel subset of S and $R > 0$, then

$$m_2(\{re^{i\theta} : 0 < r < R, e^{i\theta} \in A\}) = \mu(A)R^2/2.$$

Exercise 11.17 Use Exercise 11.16 to prove that if f is a continuous real-valued function with support in the ball $B(0, R) = \{(x, y) : x^2 + y^2 < R^2\}$, then

$$\int\int_{B(0,R)} f(x, y)\, dy\, dx = \int_0^{2\pi} \int_0^R f(r\cos\theta, r\sin\theta)\, r\, dr\, d\theta.$$

Exercise 11.18 Prove that

$$\int_0^\infty e^{-x^2/2}\, dx = \sqrt{\pi/2}$$

by filling in the missing steps and making rigorous the following. If $I = \int_0^\infty e^{-x^2/2}\, dx$, then

$$I^2 = \int_0^\infty \int_0^\infty e^{-(x^2+y^2)/2}\, dy\, dx = \int_0^{\pi/2} \int_0^\infty e^{-r^2/2} r\, dr\, d\theta = \pi/2.$$

Exercise 11.19 If $M = (M_{ij})_{i,j=1}^n$ is a $n \times n$ matrix and $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$, define Mx to be the element of $\mathbb{R}^n$ whose i^{th} coordinate is $\sum_{j=1}^n M_{ij}x_j$. (This is just the usual matrix multiplication of a $n \times n$ matrix and a $n \times 1$ matrix.) If A is a Borel subset of $\mathbb{R}^n$, let $M(A) = \{Mx : x \in A\}$.
(1) If $c \in \mathbb{R}$ and

$$M_{ij} = \begin{cases} c, & i = j = 1; \\ 1, & i = j \neq 1; \\ 0, & i \neq j; \end{cases}$$

show

$$m_n(M(A)) = |c| m_n(A) = |\det M| m_n(A),$$

where we use m_n for n-dimensional Lebesgue measure. (Multiplication by M multiplies the first coordinate by c.)
(2) If $1 \leq k \leq n$ and

$$M_{ij} = \begin{cases} 1, & i = 1 \text{ and } j = k; \\ 1, & j = 1 \text{ and } i = k; \\ 1, & i = j \text{ and neither equals } k; \\ 0, & \text{otherwise}; \end{cases}$$

show

$$m_n(M(A)) = m_n(A) = |\det M| m_n(A).$$

(Multiplication by M interchanges the i^{th} and k^{th} coordinates.)
(3) If $c \in \mathbb{R}$ and

$$M_{ij} = \begin{cases} 1, & i = j; \\ c, & i = 1, j = 2; \\ 0, & \text{otherwise}, \end{cases}$$

show

$$m_n(M(A)) = m_n(A) = |\det M| m_n(A).$$

(Multiplication by M replaces x_1 by $x_1 + cx_2$.)
(4) Since every $n \times n$ matrix can be written as the product of matrices each of which has the form given in (1), (2), or (3), conclude that if M is any $n \times n$ matrix, then

$$m_n(M(A)) = |\det M| m_n(A).$$

(5) If M is an orthogonal matrix, so that M times its transpose is the identity, show $m_n(M(A)) = m_n(A)$. (Multiplication by an orthogonal matrix is a rotation of $\mathbb{R}^n$.)

Chapter 12

Signed measures

Signed measures have the countable additivity property of measures, but are allowed to take negative as well as positive values. We will see shortly that an example of a signed measure is $\nu(A) = \int_A f\,d\mu$, where f is integrable and takes both positive and negative values.

12.1 Positive and negative sets

Definition 12.1 Let $\mathcal{A}$ be a σ-algebra. A *signed measure* is a function $\mu : \mathcal{A} \to (-\infty, \infty]$ such that $\mu(\emptyset) = 0$ and $\mu(\cup_{i=1}^\infty A_i) = \sum_{i=1}^\infty \mu(A_i)$ whenever the A_i are pairwise disjoint and all the A_i are in $\mathcal{A}$.

When we want to emphasize that a measure is defined as in Definition 3.1 and only takes non-negative values, we refer to it as a *positive measure*.

Definition 12.2 Let μ be a signed measure. A set $A \in \mathcal{A}$ is called a *positive set* for μ if $\mu(B) \geq 0$ whenever $B \subset A$ and $B \in \mathcal{A}$. We say $A \in \mathcal{A}$ is a *negative set* if $\mu(B) \leq 0$ whenever $B \subset A$ and $B \in \mathcal{A}$. A *null set* A is one where $\mu(B) = 0$ whenever $B \subset A$ and $B \in \mathcal{A}$.

Note that if μ is a signed measure, then

$$\mu(\cup_{i=1}^\infty A_i) = \lim_{n\to\infty} \mu(\cup_{i=1}^n A_i).$$

The proof is the same as in the case of positive measures.

Example 12.3 Suppose m is Lebesgue measure and

$$\mu(A) = \int_A f\, dm$$

for some integrable f. If we let $P = \{x : f(x) \geq 0\}$, then P is easily seen to be a positive set, and if $N = \{x : f(x) < 0\}$, then N is a negative set. The Hahn decomposition which we give below is a decomposition of our space (in this case $\mathbb{R}$) into the positive and negative sets P and N. This decomposition is unique, except that $C = \{x : f(x) = 0\}$ could be included in N instead of P, or apportioned partially to P and partially to N. Note, however, that C is a null set. The Jordan decomposition below is a decomposition of μ into μ^+ and μ^-, where $\mu^+(A) = \int_A f^+\, dm$ and $\mu^-(A) = \int_A f^-\, dm$.

Proposition 12.4 *Let μ be a signed measure which takes values in $(-\infty, \infty]$. Let E be measurable with $\mu(E) < 0$. Then there exists a measurable subset F of E that is a negative set with $\mu(F) < 0$.*

Proof. If E is a negative set, we are done. If not, there exists a measurable subset with positive measure. Let n_1 be the smallest positive integer such that there exists $E_1 \subset E$ with $\mu(E_1) \geq 1/n_1$. We then define pairwise disjoint measurable sets $E_2, E_3, \ldots$ by induction as follows. Let $k \geq 2$ and suppose $E_1, \ldots, E_{k-1}$ are pairwise disjoint measurable sets with $\mu(E_i) > 0$ for $i = 1, \ldots, k-1$. If $F_k = E - (E_1 \cup \cdots \cup E_{k-1})$ is a negative set, then

$$\mu(F_k) = \mu(E) - \sum_{i=1}^{k-1} \mu(E_i) \leq \mu(E) < 0$$

and F_k is the desired set F. If F_k is not a negative set, let n_k be the smallest positive integer such that there exists $E_k \subset F_k$ with E_k measurable and $\mu(E_k) \geq 1/n_k$.

We stop the construction if there exists k such that F_k is a negative set with $\mu(F_k) < 0$. If not, we continue and let $F = \cap_k F_k = E - (\cup_k E_k)$. Since $0 > \mu(E) > -\infty$ and $\mu(E_k) \geq 0$, then

$$\mu(E) = \mu(F) + \sum_{k=1}^{\infty} \mu(E_k).$$

Then $\mu(F) \leq \mu(E) < 0$, so the sum converges.

It remains to show that F is a negative set. Suppose $G \subset F$ is measurable with $\mu(G) > 0$. Then $\mu(G) \geq 1/N$ for some N. But this contradicts the construction, since for some k, $n_k > N$, and we would have chosen the set G instead of the set E_k at stage k. Therefore F must be a negative set. □

12.2 Hahn decomposition theorem

Recall that we write $A\Delta B$ for $(A - B) \cup (B - A)$. The following is known as the *Hahn decomposition*.

Theorem 12.5 *(1) Let μ be a signed measure taking values in $(-\infty, \infty]$. There exist disjoint measurable sets E and F in $\mathcal{A}$ whose union is X and such that E is a negative set and F is a positive set.*

(2) If E' and F' are another such pair, then $E\Delta E' = F\Delta F'$ is a null set with respect to μ.

(3) If μ is not a positive measure, then $\mu(E) < 0$. If $-\mu$ is not a positive measure, then $\mu(F) > 0$.

Proof. (1) Let $L = \inf\{\mu(A) : A \text{ is a negative set}\}$. Choose negative sets A_n such that $\mu(A_n) \to L$. Let $E = \cup_{n=1}^{\infty} A_n$. Let $B_n = A_n - (B_1 \cup \cdots \cup B_{n-1})$ for each n. Since A_n is a negative set, so is each B_n. Also, the B_n are disjoint and $\cup_n B_n = \cup_n A_n = E$. If $C \subset E$, then

$$\mu(C) = \lim_{n\to\infty} \mu(C \cap (\cup_{i=1}^{n} B_i)) = \lim_{n\to\infty} \sum_{i=1}^{n} \mu(C \cap B_i) \leq 0.$$

Thus E is a negative set.

Since E is a negative set,

$$\mu(E) = \mu(A_n) + \mu(E - A_n) \leq \mu(A_n).$$

Letting $n \to \infty$, we obtain $\mu(E) = L$.

Let $F = E^c$. If F were not a positive set, there would exist $B \subset F$ with $\mu(B) < 0$. By Proposition 12.4 there exists a negative set C contained in B with $\mu(C) < 0$. But then $E \cup C$ would be a negative set with $\mu(E \cup C) < \mu(E) = L$, a contradiction.

(2) To prove uniqueness, if E', F' are another such pair of sets and $A \subset E - E' \subset E$, then $\mu(A) \leq 0$. But $A \subset E - E' = F' - F \subset F'$, so $\mu(A) \geq 0$. Therefore $\mu(A) = 0$. The same argument works if $A \subset E' - E$, and any subset of $E \Delta E'$ can be written as the union of A_1 and A_2, where $A_1 \subset E - E'$ and $A_2 \subset E' - E$.

(3) Suppose μ is not a positive measure but $\mu(E) = 0$. If $A \in \mathcal{A}$, then

$$\mu(A) = \mu(A \cap E) + \mu(A \cap F) \geq \mu(E) + \mu(A \cap F) \geq 0,$$

which says that μ must be a positive measure, a contradiction. A similar argument applies for $-\mu$ and F. □

Let us say two measures μ and ν are *mutually singular* if there exist two disjoint sets E and F in $\mathcal{A}$ whose union is X with $\mu(E) = \nu(F) = 0$. This is often written $\mu \perp \nu$.

Example 12.6 If μ is Lebesgue measure restricted to $[0, 1/2]$, that is, $\mu(A) = m(A \cap [0, 1/2])$, and ν is Lebesgue measure restricted to $[1/2, 1]$, then μ and ν are mutually singular. We let $E = [0, 1/2]$ and $F = (1/2, 1]$. This example works because the Lebesgue measure of $\{1/2\}$ is 0.

Example 12.7 A more interesting example is the following. Let f be the Cantor-Lebesgue function where we define $f(x) = 1$ if $x \geq 1$ and $f(x) = 0$ if $x \leq 0$ and let ν be the Lebesgue-Stieltjes measure associated with f. Let μ be Lebesgue measure. Then $\mu \perp \nu$. To see this, we let $E = C$, where C is the Cantor set, and $F = C^c$. We already know that $m(E) = 0$ and we need to show $\nu(F) = 0$. To do that, we need to show $\nu(I) = 0$ for every open interval contained in F. This will follow if we show $\nu(J) = 0$ for every interval of the form $J = (a, b]$ contained in F. But f is constant on every such interval, so $f(b) = f(a)$, and therefore $\nu(J) = f(b) - f(a) = 0$.

12.3 Jordan decomposition theorem

The following is known as the *Jordan decomposition theorem.*

Theorem 12.8 *If μ is a signed measure on a measurable space $(X, \mathcal{A})$, there exist positive measures μ^+ and μ^- such that $\mu = \mu^+ - \mu^-$ and μ^+ and μ^- are mutually singular. This decomposition is unique.*

Proof. Let E and F be negative and positive sets, resp., for μ so that $X = E \cup F$ and $E \cap F = \emptyset$. Let $\mu^+(A) = \mu(A \cap F)$, $\mu^-(A) = -\mu(A \cap E)$. This gives the desired decomposition.

If $\mu = \nu^+ - \nu^-$ is another such decomposition with ν^+, ν^- mutually singular, let E' be a set such that $\nu^+(E') = 0$ and $\nu^-((E')^c) = 0$. Set $F' = (E')^c$. Hence $X = E' \cup F'$ and $E' \cap F' = \emptyset$. If $A \subset F'$, then $\nu^-(A) \le \nu^-(F') = 0$, and so

$$\mu(A) = \nu^+(A) - \nu^-(A) = \nu^+(A) \ge 0,$$

and consequently F' is a positive set for μ. Similarly, E' is a negative set for μ. Thus E', F' gives another Hahn decomposition of X. By the uniqueness part of the Hahn decomposition theorem, $F \Delta F'$ is a null set with respect to μ. Since $\nu^+(E') = 0$ and $\nu^-(F') = 0$, if $A \in \mathcal{A}$, then

$$\begin{aligned} \nu^+(A) &= \nu^+(A \cap F') = \nu^+(A \cap F') - \nu^-(A \cap F') \\ &= \mu(A \cap F') = \mu(A \cap F) = \mu^+(A), \end{aligned}$$

and similarly $\nu^- = \mu^-$. □

The measure

$$|\mu| = \mu^+ + \mu^- \tag{12.1}$$

is called the *total variation measure* of μ.

12.4 Exercises

Exercise 12.1 Suppose μ is a signed measure. Prove that A is a null set with respect to μ if and only if $|\mu|(A) = 0$.

Exercise 12.2 Let μ be a signed measure. Define

$$\int f\,d\mu = \int f\,d\mu^+ - \int f\,d\mu^-.$$

Prove that

$$\Big|\int f\,d\mu\Big| \le \int |f|\,d|\mu|.$$

Exercise 12.3 Let μ be a signed measure on $(X,\mathcal{A})$. Prove that

$$|\mu(A)| = \sup\Big\{\Big|\int_A f\,d\mu\Big| : f \le 1\Big\}.$$

Exercise 12.4 Let μ be a positive measure and ν a signed measure. Prove that $\nu \ll \mu$ if and only if $\nu^+ \ll \mu$ and $\nu^- \ll \mu$.

Exercise 12.5 Let $(X,\mathcal{A})$ be a measurable space. Suppose $\lambda = \mu - \nu$, where μ and ν are finite positive measures. Prove that $\mu(A) \ge \lambda^+(A)$ and $\nu(A) \ge \lambda^-(A)$ for every $A \in \mathcal{A}$.

Exercise 12.6 Let $(X,\mathcal{A})$ be a measurable space. Prove that if μ and ν are finite signed measures, then $|\mu+\nu|(A) \le |\mu(A)| + |\nu(A)|$ for every $A \in \mathcal{A}$.

Exercise 12.7 Suppose that μ is a signed measure on $(X,\mathcal{A})$. Prove that if $A \in \mathcal{A}$, then

$$\mu^+(A) = \sup\{\mu(B) : B \in \mathcal{A}, B \subset A\}$$

and

$$\mu^-(A) = -\inf\{\mu(B) : B \in \mathcal{A}, B \subset A\}.$$

Exercise 12.8 Suppose that μ is a signed measure on $(X,\mathcal{A})$. Prove that if $A \in \mathcal{A}$, then

$$|\mu|(A) = \sup\Big\{\sum_{j=1}^n |\mu(B_j)| : \text{ each } B_j \in \mathcal{A}, \\ \text{the } B_j \text{ are disjoint}, \cup_{j=1}^n B_j = A\Big\}.$$

Chapter 13

The Radon-Nikodym theorem

Suppose f is non-negative and integrable with respect to μ. If we define ν by

$$\nu(A) = \int_A f\,d\mu, \tag{13.1}$$

then ν is a measure. The only part that needs thought is the countable additivity. If A_n are disjoint measurable sets, we have

$$\nu(\cup_n A_n) = \int_{\cup_n A_n} f\,d\mu = \sum_{n=1}^\infty \int_{A_n} f\,d\mu = \sum_{n=1}^\infty \nu(A_n)$$

by using Proposition 7.5. Moreover, $\nu(A)$ is zero whenever $\mu(A)$ is.

In this chapter we consider the converse. If we are given two measures μ and ν, when does there exist f such that (13.1) holds? The Radon-Nikodym theorem answers this question.

13.1 Absolute continuity

Definition 13.1 A measure ν is said to be *absolutely continuous* with respect to a measure μ if $\nu(A) = 0$ whenever $\mu(A) = 0$. We write $\nu \ll \mu$.

Proposition 13.2 *Let ν be a finite measure. Then ν is absolutely continuous with respect to μ if and only if for all ε there exists δ such that $\mu(A) < \delta$ implies $\nu(A) < \varepsilon$.*

Proof. Suppose for each ε, there exists δ such that $\mu(A) < \delta$ implies $\nu(A) < \varepsilon$. If $\mu(A) = 0$, then $\nu(A) < \varepsilon$ for all ε, hence $\nu(A) = 0$, and thus $\nu \ll \mu$.

Suppose now that $\nu \ll \mu$. If there exists an ε for which no corresponding δ exists, then there exists E_k such that $\mu(E_k) < 2^{-k}$ but $\nu(E_k) \geq \varepsilon$. Let $F = \cap_{n=1}^\infty \cup_{k=n}^\infty E_k$. Then

$$\mu(F) = \lim_{n\to\infty} \mu(\cup_{k=n}^\infty E_k) \leq \lim_{n\to\infty} \sum_{k=n}^\infty 2^{-k} = 0,$$

but

$$\nu(F) = \lim_{n\to\infty} \nu(\cup_{k=n}^\infty E_k) \geq \varepsilon;$$

This contradicts the absolute continuity. □

13.2 The main theorem

Lemma 13.3 *Let μ and ν be finite positive measures on a measurable space $(X, \mathcal{A})$. Either $\mu \perp \nu$ or else there exists $\varepsilon > 0$ and $G \in \mathcal{A}$ such that $\mu(G) > 0$ and G is a positive set for $\nu - \varepsilon\mu$.*

Proof. Consider the Hahn decomposition for $\nu - \frac{1}{n}\mu$. Thus there exists a negative set E_n and a positive set F_n for this measure, E_n and F_n are disjoint, and their union is X. Let $F = \cup_n F_n$ and $E = \cap_n E_n$. Note $E^c = \cup_n E_n^c = \cup_n F_n = F$.

For each n, $E \subset E_n$, so

$$\nu(E) \leq \nu(E_n) \leq \tfrac{1}{n}\mu(E_n) \leq \tfrac{1}{n}\mu(X).$$

Since ν is a positive measure, this implies $\nu(E) = 0$.

One possibility is that $\mu(E^c) = 0$, in which case $\mu \perp \nu$. The other possibility is that $\mu(E^c) > 0$. In this case, $\mu(F_n) > 0$ for some n. Let $\varepsilon = 1/n$ and $G = F_n$. Then from the definition of F_n, G is a positive set for $\nu - \varepsilon\mu$. □

We now are ready for the *Radon-Nikodym theorem.*

Theorem 13.4 *Suppose μ is a σ-finite positive measure on a measurable space $(X,\mathcal{A})$ and ν is a finite positive measure on $(X,\mathcal{A})$ such that ν is absolutely continuous with respect to μ. Then there exists a μ-integrable non-negative function f which is measurable with respect to $\mathcal{A}$ such that*

$$\nu(A) = \int_A f\,d\mu$$

for all $A \in \mathcal{A}$. Moreover, if g is another such function, then $f = g$ almost everywhere with respect to μ.

The function f is called the *Radon-Nikodym derivative* of ν with respect to μ or sometimes the *density* of ν with respect to μ, and is written $f = d\nu/d\mu$. Sometimes one writes

$$d\nu = f\,d\mu.$$

The idea of the proof is to look at the set of f such that $\int_A f\,d\mu \le \nu(A)$ for each $A \in \mathcal{A}$, and then to choose the one such that $\int_X f\,d\mu$ is largest.

Proof. *Step 1.* Let us first prove the uniqueness assertion. For every set A we have

$$\int_A (f-g)\,d\mu = \nu(A) - \nu(A) = 0.$$

By Proposition 8.1 we have $f - g = 0$ a.e. with respect to μ.

Step 2. Let us assume μ is a finite measure for now. In this step we define the function f. Define

$$\mathcal{F} = \Big\{g \text{ measurable} : g \ge 0, \int_A g\,d\mu \le \nu(A) \text{ for all } A \in \mathcal{A}\Big\}.$$

$\mathcal{F}$ is not empty because $0 \in \mathcal{F}$. Let $L = \sup\{\int g\,d\mu : g \in \mathcal{F}\}$, and let g_n be a sequence in $\mathcal{F}$ such that $\int g_n\,d\mu \to L$. Let $h_n = \max(g_1,\ldots,g_n)$.

We claim that if g_1 and g_2 are in $\mathcal{F}$, then $h_2 = \max(g_1, g_2)$ is also in $\mathcal{F}$. To see this, let $B = \{x : g_1(x) \geq g_2(x)\}$, and write

$$\begin{aligned}\int_A h_2\, d\mu &= \int_{A\cap B} h_2\, d\mu + \int_{A\cap B^c} h_2\, d\mu \\ &= \int_{A\cap B} g_1\, d\mu + \int_{A\cap B^c} g_2\, d\mu \\ &\leq \nu(A\cap B) + \nu(A\cap B^c) \\ &= \nu(A).\end{aligned}$$

Therefore $h_2 \in \mathcal{F}$.

By an induction argument, h_n is in $\mathcal{F}$.

The h_n increase, say to f. By monotone convergence, $\int f\, d\mu = L$ and

$$\int_A f\, d\mu \leq \nu(A) \tag{13.2}$$

for all A.

Step 3. Next we prove that f is the desired function. Define a measure λ by

$$\lambda(A) = \nu(A) - \int_A f\, d\mu.$$

λ is a positive measure since $f \in \mathcal{F}$.

Suppose λ is not mutually singular to μ. By Lemma 13.3, there exists $\varepsilon > 0$ and G such that G is measurable, $\mu(G) > 0$, and G is a positive set for $\lambda - \varepsilon\mu$. For any $A \in \mathcal{A}$,

$$\nu(A) - \int_A f\, d\mu = \lambda(A) \geq \lambda(A\cap G) \geq \varepsilon\mu(A\cap G) = \int_A \varepsilon\chi_G\, d\mu,$$

or

$$\nu(A) \geq \int_A (f + \varepsilon\chi_G)\, d\mu.$$

Hence $f + \varepsilon\chi_G \in \mathcal{F}$. But

$$\int_X (f + \varepsilon\chi_G)\, d\mu = L + \varepsilon\mu(G) > L,$$

a contradiction to the definition of L.

Therefore $\lambda \perp \mu$. Then there must exist $H \in \mathcal{A}$ such that $\mu(H) = 0$ and $\lambda(H^c) = 0$. Since $\nu \ll \mu$, then $\nu(H) = 0$, and hence

$$\lambda(H) = \nu(H) - \int_H f\, d\mu = 0.$$

This implies $\lambda = 0$, or $\nu(A) = \int_A f\,d\mu$ for all A.

Step 4. We now suppose μ is σ-finite. There exist $F_i \uparrow X$ such that $\mu(F_i) < \infty$ for each i. Let μ_i be the restriction of μ to F_i, that is, $\mu_i(A) = \mu(A \cap F_i)$. Define ν_i, the restriction of ν to F_i, similarly. If $\mu_i(A) = 0$, then $\mu(A \cap F_i) = 0$, hence $\nu(A \cap F_i) = 0$, and thus $\nu_i(A) = 0$. Therefore $\nu_i \ll \mu_i$. If f_i is the function such that $d\nu_i = f_i\,d\mu_i$, the argument of Step 1 shows that $f_i = f_j$ on F_i if $i \le j$. Define f by $f(x) = f_i(x)$ if $x \in F_i$. Then for each $A \in \mathcal{A}$,

$$\nu(A \cap F_i) = \nu_i(A) = \int_A f_i\,d\mu_i = \int_{A \cap F_i} f\,d\mu.$$

Letting $i \to \infty$ shows that f is the desired function. □

13.3 Lebesgue decomposition theorem

The proof of the *Lebesgue decomposition theorem* is almost the same.

Theorem 13.5 *Suppose μ and ν are two finite positive measures. There exist positive measures λ, ρ such that $\nu = \lambda + \rho$, ρ is absolutely continuous with respect to μ, and λ and μ are mutually singular.*

Proof. Define $\mathcal{F}$ and L and construct f as in the proof of the Radon-Nikodym theorem. Let $\rho(A) = \int_A f\,d\mu$ and let $\lambda = \nu - \rho$. Our construction shows that

$$\int_A f\,d\mu \le \nu(A),$$

so $\lambda(A) \ge 0$ for all A. We have $\rho + \lambda = \nu$. We need to show μ and λ are mutually singular.

If not, by Lemma 13.3, there exists $\varepsilon > 0$ and $F \in \mathcal{A}$ such that $\mu(F) > 0$ and F is a positive set for $\lambda - \varepsilon\mu$. We get a contradiction exactly as in the proof of the Radon-Nikodym theorem. We conclude that $\lambda \perp \mu$. □

13.4 Exercises

Exercise 13.1 This exercise asks you to prove the Radon-Nikodym theorem for signed measures. Let $(X, \mathcal{A})$ be a measurable space. Suppose ν is a signed measure, μ is a finite positive measure, and $\nu(A) = 0$ whenever $\mu(A) = 0$ and $A \in \mathcal{A}$. Show there exists an integrable real-valued function f such that $\nu(A) = \int_A f\, d\mu$ for all $A \in \mathcal{A}$.

Exercise 13.2 We define a *complex measure* μ on a measurable space $(X, \mathcal{A})$ to be a map from $\mathcal{A}$ to $\mathbb{C}$ such that $\mu(\emptyset) = 0$ and $\mu(\cup_{i=1}^\infty A_i) = \sum_{i=1}^\infty \mu(A_i)$ whenever the A_i are in $\mathcal{A}$ and are pairwise disjoint. Formulate and prove a Radon-Nikodym theorem for complex measures.

Exercise 13.3 Let $(X, \mathcal{A})$ be a measurable space and let μ and ν be two finite measures. We say μ and ν are *equivalent measures* if $\mu \ll \nu$ and $\nu \ll \mu$. Show that μ and ν are equivalent if and only if there exists a measurable function f that is strictly positive a.e. with respect to μ such that $d\nu = f\, d\mu$.

Exercise 13.4 Suppose μ and ν are two finite measures such that ν is absolutely continuous with respect to μ. Let $\rho = \mu + \nu$. Note that $\mu(A) \le \rho(A)$ and $\nu(A) \le \rho(A)$ for each measurable A. In particular, $\mu \ll \rho$ and $\nu \ll \rho$. Prove that if $f = d\mu/d\rho$ and $g = d\nu/d\rho$, then g is strictly positive for almost every x with respect to μ, $f + g = 1$, and $d\nu = (f/g)\, d\mu$.

Exercise 13.5 If μ is a signed measure on $(X, \mathcal{A})$ and $|\mu|$ is the total variation measure, prove that there exists a real-valued function f that is measurable with respect to $\mathcal{A}$ such that $|f| = 1$ a.e. with respect to μ and $d\mu = f\, d|\mu|$.

Exercise 13.6 Suppose $\nu \ll \mu$ and $\rho \ll \nu$. Prove that $\rho \ll \mu$ and

$$\frac{d\rho}{d\mu} = \frac{d\rho}{d\nu} \cdot \frac{d\nu}{d\mu}.$$

Exercise 13.7 Suppose λ_n is a sequence of positive measures on a measurable space $(X, \mathcal{A})$ with $\sup_n \lambda_n(X) < \infty$ and μ is another finite positive measure on $(X, \mathcal{A})$. Suppose $\lambda_n = f_n\, d\mu + \nu_n$ is the Lebesgue decomposition of λ_n; in particular, $\nu_n \perp \mu$. If $\lambda = \sum_{n=1}^\infty \lambda_n$, show that

$$\lambda = \Big(\sum_{n=1}^\infty f_n\Big)\,d\mu + \sum_{n=1}^\infty \nu_n$$

is the Lebesgue decomposition of λ.

Exercise 13.8 Let $(X, \mathcal{F}, \mu)$ be a measure space, and suppose $\mathcal{E}$ is a sub-σ-algebra of $\mathcal{F}$, that is, $\mathcal{E}$ is itself a σ-algebra and $\mathcal{E} \subset \mathcal{F}$. Suppose f is a non-negative integrable function that is measurable with respect to $\mathcal{F}$. Define $\nu(A) = \int_A f\,d\mu$ for $A \in \mathcal{E}$ and let $\overline{\mu}$ be the restriction of μ to $\mathcal{E}$.
(1) Prove that $\nu \ll \overline{\mu}$.
(2) Since ν and $\overline{\mu}$ are measures on $\mathcal{E}$, then $g = d\nu/d\overline{\mu}$ is measurable with respect to $\mathcal{E}$. Prove that

$$\int_A g\,d\mu = \int_A f\,d\mu \tag{13.3}$$

whenever $A \in \mathcal{E}$. g is called the *conditional expectation* of f with respect to $\mathcal{E}$ and we write $g = \mathbb{E}\,[f \mid \mathcal{E}]$. If f is integrable and real-valued but not necessarily non-negative, we define

$$\mathbb{E}\,[f \mid \mathcal{E}] = \mathbb{E}\,[f^+ \mid \mathcal{E}] - \mathbb{E}\,[f^- \mid \mathcal{E}].$$

(3) Show that $f = g$ if and only if f is measurable with respect to $\mathcal{E}$.
(4) Prove that if h is $\mathcal{E}$ measurable and $\int_A h\,d\mu = \int_A f\,d\mu$ for all $A \in \mathcal{E}$, then $h = g$ a.e. with respect to μ.

Exercise 13.9 Suppose $(X, \mathcal{A}, \mu)$ is a measure space and f is integrable and measurable with respect to $\mathcal{A}$. Suppose in addition that $B_1, B_2, \ldots, B_n$ is a finite sequence of disjoint elements of $\mathcal{A}$ whose union is X and that each B_j has positive μ measure. Let $\mathcal{C} = \sigma(B_1, \ldots, B_n)$. Prove that

$$\mathbb{E}\,[f \mid \mathcal{C}] = \sum_{j=1}^n \frac{\int_{B_j} f\,d\mu}{\mu(B_j)} \chi_{B_j}.$$

Exercise 13.10 Suppose that $(X, \mathcal{F}, \mu)$ is a measure space, $\mathcal{E}$ is a sub-σ-algebra of $\mathcal{F}$, and $\mathcal{D}$ is a sub-σ-algebra of $\mathcal{E}$. Suppose f is integrable, real-valued, and measurable with respect to $\mathcal{F}$. Prove that

$$\mathbb{E}\,[\,\mathbb{E}\,[f \mid \mathcal{E}] \mid \mathcal{D}] = \mathbb{E}\,[f \mid \mathcal{D}]$$

and

$$\mathbb{E}\,[\,\mathbb{E}\,[f \mid \mathcal{D}] \mid \mathcal{E}] = \mathbb{E}\,[f \mid \mathcal{D}].$$

Exercise 13.11 Suppose that $(X, \mathcal{F}, \mu)$ is a measure space and $\mathcal{E}$ is a sub-σ-algebra of $\mathcal{F}$. Suppose that f and fg are integrable real-valued functions, where f is measurable with respect to $\mathcal{F}$ and g is measurable with respect to $\mathcal{E}$. Prove that

$$\mathbb{E}\,[fg \mid \mathcal{E}] = g\mathbb{E}\,[f \mid \mathcal{E}].$$

Chapter 14

Differentiation

In this chapter we want to look at when a function from $\mathbb{R}$ to $\mathbb{R}$ is differentiable and when the fundamental theorem of calculus holds. Briefly, our results are the following.
(1) The derivative of $\int_a^x f(y)\,dy$ is equal to f a.e. if f is integrable (Theorem 14.5);
(2) Functions of bounded variation, in particular monotone functions, are differentiable (Theorem 14.8);
(3) $\int_a^b f'(y)\,dy = f(b) - f(a)$ if f is absolutely continuous (Theorem 14.14).

Our approach uses what are known as maximal functions and uses the Radon-Nikodym theorem and the Lebesgue decomposition theorem. However, some students and instructors prefer a more elementary proof of the results on differentiation. In Sections 14.5, 14.6, and 14.7 we give an alternative approach that avoids the use of the Radon-Nikodym theorem and Lebesgue decomposition theorem.

The definition of derivative is the same as in elementary calculus. A function f is *differentiable* at x if

$$\lim_{h\to 0} \frac{f(x+h) - f(x)}{h}$$

exists, and the limit is called the *derivative* of f at x and is denoted $f'(x)$. If $f : [a, b] \to \mathbb{R}$, we say f is differentiable on $[a, b]$ if the derivative exists for each $x \in (a,b)$ and both

$$\lim_{h\to 0+} \frac{f(a+h)-f(a)}{h} \quad \text{and} \quad \lim_{h\to 0-} \frac{f(b+h)-f(b)}{h}$$

exist.

14.1 Maximal functions

In this section we consider real-valued functions on $\mathbb{R}^n$. Let $B(x,r)$ be the open ball with center x and radius r.

The following is an example of what is known as a *covering lemma.* We use m for Lebesgue measure on $\mathbb{R}^n$ throughout this section.

Proposition 14.1 *Suppose $E \subset \mathbb{R}^n$ is covered by a collection of balls $\{B_\alpha\}$ of bounded diameter. There exists a disjoint sequence $B_1, B_2, \ldots$ of elements of $\{B_\alpha\}$ such that*

$$m(E) \le 5^n \sum_k m(B_k).$$

Proof. Let $d(B_\alpha)$ be the diameter of B_α. Choose B_1 such that

$$d(B_1) \ge \tfrac{1}{2} \sup_\alpha d(B_\alpha).$$

Once $B_1, \ldots, B_k$ are chosen, choose B_{k+1} disjoint from $B_1, \ldots, B_k$ such that

$$d(B_{k+1}) \ge \tfrac{1}{2} \sup\{d(B_\alpha) : B_\alpha \text{ is disjoint from } B_1, \ldots, B_k\}.$$

The procedure might terminate after a finite number of steps or it might not.

If $\sum_k m(B_k) = \infty$, we have our result. Suppose $\sum_k m(B_k) < \infty$. Let B_k^* be a ball with the same center as B_k but 5 times the radius. We claim $E \subset \cup_k B_k^*$. Once we have this,

$$m(E) \le m(\cup_k B_k^*) \le \sum_k m(B_k^*) = 5^n \sum_k m(B_k).$$

To show $E \subset \cup_k B_k^*$, it suffices to show each $B_\alpha \subset \cup_k B_k^*$, since $\{B_\alpha\}$ is a cover of E. Fix α. If B_α is one of the B_k, we are done.

If $\sum_k m(B_k) < \infty$, then $d(B_k) \to 0$. Let k be the smallest integer such that $d(B_{k+1}) < \frac{1}{2} d(B_\alpha)$. B_α must intersect one of $B_1, \ldots, B_k$, or else we would have chosen it instead of B_{k+1}. Therefore B_α intersects B_{j_0} for some $j_0 \le k$. We know $\frac{1}{2} d(B_\alpha) \le d(B_{j_0})$, and some simple geometry shows that $B_\alpha \subset B_{j_0}^*$. In fact, let x_{j_0} be the center of B_{j_0} and y a point in $B_\alpha \cap B_{j_0}$. If $x \in B_\alpha$, then

$$|x - x_{j_0}| \le |x - y| + |y - x_{j_0}| < d(B_\alpha) + d(B_{j_0})/2 \le \tfrac{5}{2} d(B_{j_0}),$$

or $x \in B_{j_0}^*$. Therefore $B_\alpha \subset B_{j_0}$, and the proof is complete. □

We say f is *locally integrable* if $\int_K |f(x)|\, dx$ is finite whenever K is compact. If f is locally integrable, define

$$Mf(x) = \sup_{r>0} \frac{1}{m(B(x,r))} \int_{B(x,r)} |f(y)|\, dy.$$

Note that without the supremum, we are looking at the average of $|f|$ over $B(x,r)$. The function Mf is called the *maximal function* of f.

We now prove a *weak 1-1 inequality*, due to Hardy and Littlewood. It is so named because M does not map integrable functions into integrable functions, but comes close in a certain sense to doing so.

Theorem 14.2 *If f is integrable, then for all $\beta > 0$*

$$m(\{x : Mf(x) > \beta\}) \le \frac{5^n}{\beta} \int |f(x)|\, dx.$$

Proof. Fix β and let $E_\beta = \{x : Mf(x) > \beta\}$. If $x \in E_\beta$, then there exists a ball B_x centered at x such that $\int_{B_x} |f(x)|\, dx > \beta m(B_x)$ by the definition of $Mf(x)$. Then

$$m(B_x) \le \frac{\int |f|}{\beta},$$

so $\{B_x\}$ is a cover of E_β by balls of bounded diameter. Extract a disjoint sequence $B_1, B_2, \ldots$ such that $5^n \sum_k m(B_k) \ge m(E_\beta)$. Then

$$m(E_\beta) \le 5^n \sum_k m(B_k) \le \frac{5^n}{\beta} \sum_k \int_{B_k} |f|$$

$$= \frac{5^n}{\beta} \int_{\cup_k B_k} |f| \le \frac{5^n}{\beta} \int |f|,$$

as desired. □

If we look at the function $f(x) = \chi_B$, where B is the unit ball, note that $Mf(x)$ is approximately a constant times $|x|^{-n}$ for x large, so Mf is not integrable. Hence M does not map the class of integrable functions into the class of integrable functions.

Theorem 14.3 *Let*

$$f_r(x) = \frac{1}{m(B(x,r))} \int_{B(x,r)} f(y)\, dy. \tag{14.1}$$

If f is locally integrable, then $f_r(x) \to f(x)$ a.e. as $r \to 0$.

Proof. It suffices to prove that for each N, $f_r(x) \to f(x)$ for almost every $x \in B(0,N)$. Let $N > 0$. We may suppose without loss of generality that f is 0 outside of $B(0,2N)$, and thus we may suppose f is integrable.

Fix $\beta > 0$. Let $\varepsilon > 0$. Using Theorem 8.4, take g continuous with compact support such that $\int |f-g|\, dm < \varepsilon$. If g_r is defined analogously to f_r using (14.1),

$$\begin{aligned} |g_r(x) - g(x)| &= \Big| \frac{1}{m(B(x,r))} \int_{B(x,r)} [g(y) - g(x)]\, dy \Big| \qquad (14.2)\\ &\le \frac{1}{m(B(x,r))} \int_{B(x,r)} |g(y) - g(x)|\, dy \to 0 \end{aligned}$$

as $r \to 0$ by the continuity of g. We have

$$\begin{aligned} \limsup_{r\to 0} |f_r(x) - f(x)| \le \limsup_{r\to 0} &|f_r(x) - g_r(x)|\\ &+ \limsup_{r\to 0} |g_r(x) - g(x)|\\ &+ |g(x) - f(x)|. \end{aligned}$$

The second term on the right is 0 by (14.2). We now use Theorem 14.2 and Lemma 10.4 to write

$$m(\{x : \limsup_{r\to 0} |f_r(x) - f(x)| > \beta\})$$

$$
\begin{aligned}
&\le m(\{x : \limsup_{r\to 0} |f_r(x) - g_r(x)| > \beta/2\}) \\
&\qquad + m(\{x : |f(x) - g(x)| > \beta/2\}) \\
&\le m(\{x : M(f-g)(x) > \beta/2\}) + \frac{\int |f-g|}{\beta/2} \\
&\le \frac{2(5^n+1)}{\beta} \int |f-g| \\
&< \frac{2(5^n+1)\varepsilon}{\beta},
\end{aligned}
$$

where we use the definition of the maximal function to see that

$$|f_r(x) - g_r(x)| \le M(f-g)(x)$$

for all r. This is true for every ε, so

$$m(\{x : \limsup_{r\to 0} |f_r(x) - g_r(x)| > \beta\}) = 0.$$

We apply this with $\beta = 1/j$ for each positive integer j, and we conclude

$$
\begin{aligned}
&m(\{x : \limsup_{r\to 0} |f_r(x) - f(x)| > 0\}) \\
&\qquad \le \cup_{j=1}^{\infty} m(\{x : \limsup_{r\to 0} |f_r(x) - f(x)| > 1/j\}) = 0.
\end{aligned}
$$

This is the result we seek. □

We can get a stronger statement:

Theorem 14.4 *For almost every x*

$$\frac{1}{m(B(x,r))} \int_{B(x,r)} |f(y) - f(x)|\, dy \to 0$$

as $r \to 0$.

Proof. For each rational c there exists a set N_c of measure 0 such that

$$\frac{1}{m(B(x,r))} \int_{B(x,r)} |f(y) - c|\, dy \to |f(x) - c|$$

for $x \notin N_c$; we see this by applying Theorem 14.3 to the function $|f(x) - c|$. Let $N = \cup_{c \in \mathbb{Q}} N_c$ and suppose $x \notin N$. Let $\varepsilon > 0$ and choose c rational such that $|f(x) - c| < \varepsilon$. Then

$$\begin{aligned}
\frac{1}{m(B(x,r))} \int_{B(x,r)} |f(y) - f(x)|\,dy & \\
\le \frac{1}{m(B(x,r))} \int_{B(x,r)} |f(y) - c|\,dy & \\
+ \frac{1}{m(B(x,r))} \int_{B(x,r)} |f(x) - c|\,dy & \\
= \frac{1}{m(B(x,r))} \int_{B(x,r)} |f(y) - c|\,dy + |f(x) - c| &
\end{aligned}$$

and hence

$$\limsup_{r \to 0} \frac{1}{m(B(x,r))} \int_{B(x,r)} |f(y) - f(x)|\,dy \le 2|f(x) - c| < 2\varepsilon.$$

Since ε is arbitrary, our result follows. □

If we apply the above to the function $f = \chi_E$, then for almost all $x \in E$

$$\frac{1}{m(B(x,r)} \int_{B(x,r)} \chi_E \to 1,$$

or

$$\frac{m(E \cap B(x,r))}{m(B(x,r))} \to 1,$$

and similarly, for almost all $x \notin E$, the ratio tends to 0. The points where the ratio tends to 1 are called *points of density* for E.

14.2 Antiderivatives

For the remainder of the chapter we consider real-valued functions on the real line $\mathbb{R}$. We can use the results on maximal functions to show that the derivative of the antiderivative of an integrable function is the function itself. A ball $B(x,h)$ in $\mathbb{R}$ is merely the interval $(x-h, x+h)$. We use m for Lebesgue measure throughout.

Define the *indefinite integral* or *antiderivative* of an integrable function f by

$$F(x) = \int_a^x f(t)\,dt.$$

Recall by Exercise 7.6 that F is continuous.

Theorem 14.5 *Suppose $f : \mathbb{R} \to \mathbb{R}$ is integrable and $a \in \mathbb{R}$. Define*

$$F(x) = \int_a^x f(y)\, dy.$$

Then F is differentiable almost everywhere and $F'(x) = f(x)$ a.e.

Proof. If $h > 0$, we have

$$F(x+h) - F(x) = \int_x^{x+h} f(y)\, dy,$$

so

$$\begin{aligned}\Big|\frac{F(x+h) - F(x)}{h} - f(x)\Big| &= \frac{1}{h}\Big|\int_x^{x+h} (f(y) - f(x))\, dy\Big| \\ &\le 2\frac{1}{m(B(x,h))}\int_{x-h}^{x+h} |f(y) - f(x)|\, dy.\end{aligned}$$

By Theorem 14.4, the right hand side goes to 0 as $h \to 0$ for almost every x, and we conclude the right hand derivative of F exists and equals f for almost every x. The left hand derivative is handled similarly. □

14.3 Bounded variation

In this section we show that functions of bounded variation are differentiable almost everywhere. We start with right continuous increasing functions.

Lemma 14.6 *Suppose $H : \mathbb{R} \to \mathbb{R}$ is increasing, right continuous, and constant for $x \ge 1$ and $x \le 0$. Let λ be the Lebesgue-Stieltjes measure defined using the function H and suppose λ and m are mutually singular. Then*

$$\lim_{r \to 0} \frac{\lambda(B(x,r))}{m(B(x,r))} = 0$$

for almost every x.

Proof. This is clear if $x < 0$ or $x > 1$. Since $\lambda \perp m$, there exist measurable sets E and F such that $\lambda(F) = 0$, $m(E) = 0$, and $F = E^c$. Let $\varepsilon > 0$.

Step 1. The first step of the proof is to find a bounded open set G such that $F \subset G$ and $\lambda(G) < \varepsilon$. By the definition of Lebesgue-Stieltjes measure, there exist $a_i < b_i$ such that $F \subset \cup_{i=1}^\infty (a_i, b_i]$ and

$$\sum_{i=1}^\infty [H(b_i) - H(a_i)] < \varepsilon/2.$$

Since H is right continuous, for each i there exists $b_i' > b_i$ such that

$$H(b_i') - H(b_i) < \varepsilon/2^{i+1}.$$

If $G' = \cup_{i=1}^\infty (a_i, b_i')$, then G' is open, G' contains F, and

$$\lambda(G') \le \sum_{i=1}^\infty \lambda((a_i, b_i')) \le \sum_{i=1}^\infty \lambda((a_i, b_i']) = \sum_{i=1}^\infty [H(b_i') - H(a_i)] < \varepsilon.$$

Since H is constant on $(-\infty, 0]$ and $[1, \infty)$, we can take G to be the set $G = G' \cap (-1, 2)$.

Step 2. If $\beta > 0$, let

$$A_\beta = \Big\{x \in F \cap [0,1] : \limsup_{r \to 0} \frac{\lambda((x,r))}{m(B(x,r))} > \beta\Big\}.$$

The second step is to show that $m(A_\beta) = 0$. If $x \in A_\beta$, then $x \in F \subset G$, and there exists an open ball B_x centered at x and contained in G such that $\lambda(B_x)/m(B_x) > \beta$. Use Proposition 14.1 to find a disjoint subsequence $B_1, B_2, \ldots$ such that

$$m(A_\beta) \le 5 \sum_{i=1}^\infty m(B_i).$$

Then

$$m(A_\beta) \le 5 \sum_{i=1}^\infty m(B_i) \le \frac{5}{\beta} \sum_{i=1}^\infty \lambda(B_i) \le \frac{5}{\beta} \lambda(G) \le \frac{5}{\beta} \varepsilon.$$

Since ε is arbitrary, and our construction of G did not depend on β, then $m(A_\beta) = 0$.

Since $m(A_{1/k}) = 0$ for each k, then

$$m(\{x \in F \cap [0,1] : \limsup_{r \to 0} \lambda(B(x,r))/m(B(x,r)) > 0\}) = 0.$$

Since $m(E) = 0$, this completes the proof. □

Proposition 14.7 *Let $F : \mathbb{R} \to \mathbb{R}$ be an increasing and right continuous function. Then F' exists a.e. Moreover, F' is locally integrable and for every $a < b$, $\int_a^b F'(x)\,dx \le F(b) - F(a)$.*

Proof. We will show F is differentiable a.e. on $[0,1]$. Once we have that, the same argument can be used to show that F is differentiable a.e. on $[-N, N]$ for each N, and that proves that F is differentiable a.e. on $\mathbb{R}$. If we redefine F so that $F(x) = \lim_{y\to 0+} F(y)$ if $x \le 0$ and $F(x) = F(1)$ if $x > 1$, then F is still right continuous and increasing, and we have not affected the differentiability of F on $[0,1]$ except possibly at the points 0 and 1.

Let ν be the Lebesgue-Stieltjes measure defined in terms of F. By the Lebesgue decomposition theorem, we can write $\nu = \lambda + \rho$, where $\lambda \perp m$ and $\rho \ll m$. Note

$$\rho([0,1]) \le \nu([0,1]) = F(1) - F(0).$$

By the Radon-Nikodym theorem there exists a non-negative integrable function f such that $\rho(A) = \int_A f\,dm$ for each measurable A.

Let

$$H(x) = \lambda((0,x]) = \nu((0,x]) - \rho((0,x]) = F(x) - F(0) - \int_0^x f(y)\,dy.$$

By Exercise 7.6, the function $x \to \int_0^x f(y)\,dy$ is continuous, so H is right continuous, increasing, and λ is the Lebesgue-Stieltjes measure defined in terms of H. By Lemma 14.6,

$$\begin{aligned}\limsup_{h\to 0+} \frac{H(x+h) - H(x)}{h} &\le \limsup_{h\to 0+} \frac{H(x+h) - H(x-h)}{h}\\ &= \limsup_{h\to 0+} \frac{\lambda((x-h, x+h])}{h}\\ &\le 4 \limsup_{h\to 0+} \frac{\lambda(B(x,2h))}{4h} = 0\end{aligned}$$

for almost every x. The same is true for the left hand derivative, so H' exists and equals 0 for almost every x. We saw by Theorem 14.5 that the function $x \to \int_0^x f(y)\,dy$ is differentiable almost everywhere, and we conclude that F is differentiable a.e.

We have shown that $F' = f$ a.e. If $a < b$,

$$\int_a^b F'(x)\,dx = \int_a^b f(x)\,dx = \rho((a,b]) \le \nu((a,b]) = F(b) - F(a).$$

This completes the proof. $\square$

Here is the main theorem on the differentiation of increasing functions.

Theorem 14.8 *If $F : \mathbb{R} \to \mathbb{R}$ is increasing, then F' exists a.e. and*

$$\int_a^b F'(x)\,dx \le F(b) - F(a) \tag{14.3}$$

whenever $a < b$.

Proof. Let $G(x) = \lim_{y \to x+} F(y)$. Since F is increasing, there are at most countably many values of x where F is not continuous, so $F(x) = G(x)$ a.e. Since G is increasing and right continuous, G is differentiable a.e. by Proposition 14.7. We will show that if x is a point where G is differentiable and at the same time $F(x) = G(x)$, then $F'(x)$ exists and is equal to $G'(x)$.

Let x be such a point, let $L = G'(x)$ and let $\varepsilon > 0$. Because F and G are increasing, for any $h > 0$ there exists a point x_h strictly between $x + h$ and $x + (1+\varepsilon)h$ where F and G agree, and so

$$F(x+h) \le F(x_h) = G(x_h) \le G(x + (1+\varepsilon)h).$$

Then

$$\begin{aligned}
\limsup_{h \to 0+} \frac{F(x+h) - F(x)}{h} &\le \limsup_{h \to 0+} \frac{G(x + (1+\varepsilon)h) - G(x)}{h} \\
&= (1+\varepsilon) \limsup_{h \to 0+} \frac{G(x + (1+\varepsilon)h) - G(x)}{(1+\varepsilon)h} \\
&= (1+\varepsilon)L.
\end{aligned}$$

Similarly, $\liminf_{h \to 0+} [F(x+h) - F(x)]/h \ge (1-\varepsilon)L$. Since ε is arbitrary, the right hand derivative of F exists at x and is equal to L. That the left hand derivative equals L is proved similarly.

Since $F' = G'$ a.e., then F' is locally integrable. If $a < b$, take $a_n \downarrow a$ and $b_n \uparrow b$ such that F and G agree on a_n and b_n. Then using Proposition 14.7,

$$\begin{aligned} F(b) - F(a) &\geq F(b_n) - F(a_n) \\ &= G(b_n) - G(a_n) \geq \int_{a_n}^{b_n} G'(x)\,dx \\ &= \int_{a_n}^{b_n} F'(x)\,dx. \end{aligned}$$

Now let $n \to \infty$ and use the monotone convergence theorem. □

Remark 14.9 Note that if F is the Cantor-Lebesgue function, then $F'(x) = 0$ a.e., in fact at every point of C^c, where C is the Cantor set. Thus

$$1 = F(1) - F(0) > 0 = \int_0^1 F'(x)\,dx,$$

and we do not in general have equality in (14.3).

A real-valued function f is of *bounded variation* on $[a, b]$ if

$$\sup\Big\{\sum_{i=1}^k |f(x_i) - f(x_{i-1})|\Big\}$$

is finite, where the supremum is over all partitions $a = x_0 < x_1 < \cdots < x_k = b$ of $[a, b]$.

Lemma 14.10 *If f is of bounded variation on $[a, b]$, then f can be written as $f = f_1 - f_2$, the difference of two increasing functions on $[a, b]$.*

Proof. Define

$$f_1(y) = \sup\Big\{\sum_{i=1}^k [f(x_i) - f(x_{i-1})]^+\Big\}$$

and

$$f_2(y) = \sup\Big\{\sum_{i=1}^k [f(x_i) - f(x_{i-1})]^-\Big\},$$

where the supremum is over all partitions $a = x_0 < x_1 < \cdots < x_k = y$ for $y \in [a, b]$. f_1 and f_2 are measurable since they are both increasing. Since

$$\sum_{i=1}^{k}[f(x_i) - f(x_{i-1})]^+ = \sum_{i=1}^{k}[f(x_i) - f(x_{i-1})]^- + f(y) - f(a),$$

taking the supremum over all partitions of $[a, y]$ yields

$$f_1(y) = f_2(y) + f(y) - f(a).$$

Clearly f_1 and f_2 are increasing in y, and the result follows by solving for $f(y)$. $\square$

Using this lemma and Theorem 14.8, we see that functions of bounded variation are differentiable a.e. Note that the converse is not true: the function $\sin(1/x)$ defined on $(0, 1]$ is differentiable everywhere, but is not of bounded variation.

Remark 14.11 If we write a function f of bounded variation as the difference of two increasing functions f_1 and f_2, then the quantity $(f_1(b) + f_2(b)) - (f_1(a) + f_2(a))$ is called the *total variation* of f on the interval $[a, b]$. We make the observation that if f is of bounded variation on the interval $[a, b]$ and on the interval $[b, c]$, then it is of bounded variation on the interval $[a, c]$.

14.4 Absolutely continuous functions

A real-valued function f is *absolutely continuous* on $[a, b]$ if given ε there exists δ such that $\sum_{i=1}^{k} |f(b_i) - f(a_i)| < \varepsilon$ whenever $\{(a_i, b_i)\}$ is a finite collection of disjoint intervals with $\sum_{i=1}^{k} |b_i - a_i| < \delta$.

It is easy to see that absolutely continuous functions are continuous and that the Cantor-Lebesgue function is not absolutely continuous.

Lemma 14.12 *If f is absolutely continuous, then it is of bounded variation.*

Proof. By the definition of absolutely continuous function with $\varepsilon = 1$, there exists δ such that $\sum_{i=1}^{k} |f(b_i) - f(a_i)| < 1$ whenever

$\sum_{i=1}^k (b_i - a_i) \le \delta$ and the (a_i, b_i) are disjoint open intervals. Hence the total variation of f on $[a + j\delta, a + (j+1)\delta]$ is less than or equal to 1. Using Remark 14.11, we see the total variation of f on $[a, b]$ is finite. □

Lemma 14.13 *Suppose f is of bounded variation and we decompose f as $f = f_1 - f_2$, where f_1 and f_2 are increasing functions. If f is absolutely continuous, then so are f_1 and f_2.*

Proof. Given ε there exists δ such that $\sum_{i=1}^k |f(b_i) - f(a_i)| < \varepsilon$ whenever $\sum_{i=1}^k (b_i - a_i) \le \delta$ and the (a_i, b_i) are disjoint open intervals. Partitioning each interval (a_i, b_i) into subintervals with $a_i = s_{i0} < s_{i1} < \cdots < s_{iJ_i} = b_i$, then

$$\sum_{i=1}^k \sum_{j=0}^{J_i - 1} (s_{i,j+1} - s_{ij}) = \sum_{i=1}^k (b_i - a_i) \le \delta.$$

Hence

$$\sum_{i=1}^k \sum_{j=0}^{J_i - 1} |f(s_{i,j+1}) - f(s_{ij})| \le \varepsilon.$$

Taking the supremum over all such partitions,

$$\sum_{i=1}^k |(f_1 + f_2)(b_i) - (f_1 + f_2)(a_i)| \le \varepsilon,$$

and our conclusion follows. □

Here is the main theorem on absolutely continuous functions.

Theorem 14.14 *If F is absolutely continuous, then F' exists a.e., and*

$$\int_a^b F'(x)\, dx = F(b) - F(a).$$

Proof. By Lemma 14.13 it suffices to suppose F is increasing and absolutely continuous. Let ν be the Lebesgue-Stieltjes measure defined in terms of F. Since F is continuous, $F(d) - F(c) = \nu((c, d))$.

Taking a limit as $k \to \infty$, we see that given ε there exists δ such that $\sum_{i=1}^\infty |F(b_i) - F(a_i)| \le \varepsilon$ whenever $\{(a_i, b_i)\}$ is a collection of disjoint intervals with $\sum_{i=1}^\infty (b_i - a_i) < \delta$. Since any open set G can be written as the union of disjoint intervals $\{(a_i, b_i)\}$, this can be rephrased as saying, given ε there exists δ such that

$$\nu(G) = \sum_{i=1}^\infty \nu((a_i, b_i)) = \sum_{i=1}^\infty (F(b_i) - F(a_i)) \le \varepsilon$$

whenever G is open and $m(G) < \delta$. If $m(A) < \delta$ and A is Borel measurable, then there exists an open set G containing A such that $m(G) < \delta$, and then $\nu(A) \le \nu(G) \le \varepsilon$. We conclude that $\nu \ll m$.

Hence there exists a non-negative integrable function f such that

$$\nu(A) = \int_A f\, dm$$

for all Borel measurable sets A. In particular, for each $x \in [a, b]$,

$$F(x) - F(a) = \nu((a, x)) = \int_a^x f(y)\, dy.$$

By Theorem 14.5, F' exists and is equal to f a.e. Setting $x = b$ we obtain

$$F(b) - F(a) = \int_a^b F'(y)\, dy$$

as desired. □

14.5 Approach 2 – differentiability

In this and the following two sections we give an alternative approach to Theorems 14.5, 14.8, and 14.14 that avoids the use of the Radon-Nikodym theorem, and instead proceeds via a covering lemma due to Vitali.

Let m be Lebesgue measure. Let $E \subset \mathbb{R}$ be a measurable set and let $\mathcal{G}$ be a collection of intervals. We say $\mathcal{G}$ is a *Vitali cover* of E if for each $x \in E$ and each $\varepsilon > 0$ there exists an interval $G \in \mathcal{G}$ containing x whose length is less than ε.

The following is known as the *Vitali covering lemma*, and is a refinement of Proposition 14.1.

Lemma 14.15 *Suppose E has finite measure and let $\mathcal{G}$ be a Vitali cover of E. Given $\varepsilon > 0$ there exists a finite subcollection of disjoint intervals $I_1, \ldots, I_n \in \mathcal{G}$ such that $m(E - \cup_{i=1}^n I_n) < \varepsilon$.*

Proof. We may replace each interval in $\mathcal{G}$ by a closed one, since the set of endpoints of a finite subcollection will have measure 0.

Let G be an open set of finite measure containing E. Since $\mathcal{G}$ is a Vitali cover, we may suppose without loss of generality that each set of $\mathcal{G}$ is contained in G. Let

$$a_0 = \sup\{m(I) : I \in \mathcal{G}\}.$$

Let I_1 be any element of $\mathcal{G}$ with $m(I_1) \geq a_0/2$. Let

$$a_1 = \sup\{m(I) : I \in \mathcal{G}, I \text{ disjoint from } I_1\},$$

and choose $I_2 \in \mathcal{G}$ disjoint from I_1 such that $m(I_2) \geq a_1/2$. Continue in this way, choosing I_{n+1} disjoint from $I_1, \ldots, I_n$ and in $\mathcal{G}$ with length at least one half as large as any other such interval in $\mathcal{G}$ that is disjoint from $I_1, \ldots, I_n$.

If the process stops at some finite stage, we are done. If not, we generate a sequence of disjoint intervals $I_1, I_2, \ldots$ Since they are disjoint and all contained in G, then $\sum_{i=1}^\infty m(I_i) \leq m(G) < \infty$. Therefore there exists N such that $\sum_{i=N+1}^\infty m(I_i) < \varepsilon/5$.

Let $R = E - \cup_{i=1}^N I_i$. We claim $m(R) < \varepsilon$. Let I_n^* be the interval with the same center as I_n but five times the length. Let $x \in R$. Since we supposed each interval in $\mathcal{G}$ was to be modified so as to include its endpoints, then $\cup_{i=1}^n I_i$ is closed. Hence there exists an interval $I \in \mathcal{G}$ containing x with I disjoint from $I_1, \ldots, I_N$. Since $\sum_n m(I_n) < \infty$, then $\sum_n a_n \leq 2 \sum_n m(I_n) < \infty$, and $a_n \to 0$. Hence I must either be one of the I_n for some $n > N$ or at least intersect it, for otherwise we would have chosen I at some stage. Let n be the smallest integer such that I intersects I_n; note $n > N$. We have $m(I) \leq a_{n-1} \leq 2m(I_n)$. Since x is in I and I intersects I_n, the distance from x to the midpoint of I_n is at most $m(I) + m(I_n)/2 \leq (5/2)m(I_n)$. Therefore $x \in I_n^*$.

Thus we have $R \subset \cup_{i=N+1}^\infty I_i^*$, so

$$m(R) \leq \sum_{i=N+1}^\infty m(I_i^*) = 5 \sum_{i=N+1}^\infty m(I_i) < \varepsilon.$$

This completes the proof. □

Given a real-valued function f, we define the *derivates* of f at x by

$$D^+ f(x) = \limsup_{h\to 0+} \frac{f(x+h)-f(x)}{h},$$
$$D^- f(x) = \limsup_{h\to 0-} \frac{f(x+h)-f(x)}{h},$$
$$D_+ f(x) = \liminf_{h\to 0+} \frac{f(x+h)-f(x)}{h},$$
$$D_- f(x) = \liminf_{h\to 0-} \frac{f(x+h)-f(x)}{h}.$$

If all the derivates are equal, then f is differentiable at x and $f'(x)$ is the common value.

Theorem 14.16 *Suppose f is increasing on $[a,b]$. Then f is differentiable almost everywhere, f' is integrable, and*

$$\int_a^b f'(x)\,dx \le f(b) - f(a). \tag{14.4}$$

Proof. We will show that the set where any two derivates are unequal has measure zero. Let us consider the set

$$E = \{x : D^+ f(x) > D_- f(x)\},$$

the other sets being similar. Let

$$E_{uv} = \{x : D^+ f(x) > v > u > D_- f(x)\}.$$

If we show $m(E_{uv}) = 0$, then observing that $E \subset \cup_{u<v,u,v\in\mathbb{Q}} E_{uv}$ will show that $m(E) = 0$.

Let $s = m(E_{uv})$, let $\varepsilon > 0$, and choose an open set G such that $E_{uv} \subset G$ and $m(G) < s + \varepsilon$. For each $x \in E_{uv}$ there exists an arbitrarily small interval $[x-h, x]$ contained in G such that $f(x) - f(x-h) < uh$. Use Lemma 14.15 to choose $I_1, \ldots, I_N$ which are disjoint and whose interiors cover a subset A of E_{uv} of measure greater than $s - \varepsilon$. Write $I_n = [x_n - h_n, x_n]$. Thus

$$A = E_{uv} \cap (\cap_{n=1}^N (x_n - h_n, x_n)).$$

Taking a sum,

$$\sum_{n=1}^{N}[f(x_n) - f(x_n - h_n)] < u\sum_{n=1}^{n} h_n < um(G) < u(s+\varepsilon).$$

Each point y in the subset A is the left endpoint of an arbitrarily small interval $(y, y+k)$ that is contained in some I_n and for which $f(y+k) - f(y) > vk$. Using Lemma 14.15 again, we pick out a finite collection $J_1, \ldots, J_M$ whose union contains a subset of A of measure larger than $s - 2\varepsilon$. Summing over these intervals yields

$$\sum_{i=1}^{M}[f(y_i + k_i) - f(y_i)] > v\sum_{i=1}^{M} k_i > v(s - 2\varepsilon).$$

Each interval J_i is contained in some interval I_n, and if we sum over those i for which $J_i \subset I_n$ we find

$$\sum_{\{i:J_i\subset I_n\}} [f(y_i + k_i) - f(y_i)] \le f(x_n) - f(x_n - h_n),$$

since f is increasing. Thus

$$\sum_{n=1}^{N}[f(x_n) - f(x_n - h_n)] \ge \sum_{i=1}^{M}[f(y_i + k_i) - f(y_i)],$$

and so $u(s+\varepsilon) > v(s - 2\varepsilon)$. This is true for each ε, so $us \ge vs$. Since $v > u$, this implies $s = 0$.

We prove similarly that $D^+ f = D_+ f$ a.e. and $D^- f = D_- f$ a.e. and conclude that

$$g(x) = \lim_{h\to 0} \frac{f(x+h) - f(x)}{h}$$

is defined almost everywhere and that f is differentiable wherever g is finite.

Define $f(x) = f(b)$ if $x \ge b$. Let

$$g_n(x) = n[f(x + 1/n) - f(x)].$$

Then $g_n(x) \to g(x)$ for almost all x, and so g is measurable. Since f is increasing, $g_n \ge 0$. By Fatou's lemma and the fact that f is increasing,

$$\begin{aligned}\int_a^b g &\le \liminf_n \int_a^b g_n \\ &= \liminf_n n \int_a^b [f(x+1/n) - f(x)] \\ &= \liminf_n \Big[n \int_b^{b+1/n} f - n \int_a^{a+1/n} f \Big] \\ &= \liminf_n \Big[f(b) - n \int_a^{a+1/n} f \Big] \\ &\le f(b) - f(a).\end{aligned}$$

We used a change of variables in the second equality. This shows that g is integrable and hence finite almost everywhere. □

We refer the reader to Lemma 14.10 to see that a function of bounded variation is differentiable almost everywhere.

14.6 Approach 2 – antiderivatives

Continuing the alternative approach, we look at when the derivative of

$$F(x) = \int_a^x f(t)\, dt \tag{14.5}$$

is equal to $f(x)$ a.e.

Theorem 14.17 *If f is integrable and F is defined by* (14.5)*, then $F'(x) = f(x)$ for almost every x.*

Proof. By writing $f = f^+ - f^-$, it suffices to consider the case where f is non-negative. In this case F is increasing, and so F' exists a.e. By Exercise 7.6 we know F is continuous.

Suppose for the moment that f is bounded by K. Then

$$\Big| \frac{F(x+1/n) - F(x)}{1/n} \Big| = \Big| n \int_x^{x+1/n} f(t)\, dt \Big|$$

is also bounded by K. By dominated convergence,

$$\begin{aligned}\int_a^c F'(x)\,dx &= \lim_n n\int_a^c [F(x+1/n)-F(x)]\,dx\\ &= \lim_n \Big[n\int_c^{c+1/n} F(x)\,dx - n\int_a^{a+c} F(x)\,dx\Big]\\ &= F(c)-F(a) = \int_a^c f(x)\,dx.\end{aligned}$$

We used a change of variables for the second equality and the fact that F is continuous for the third equality. Therefore

$$\int_a^c [F'(x)-f(x)]\,dx = 0$$

for all c, which implies $F' = f$ a.e. by Corollary 8.3.

We continue to assume f is non-negative but now allow f to be unbounded. Since $f-(f\wedge K)\geq 0$, then

$$G_K(x) = \int_a^x [f-(f\wedge K)]\,dx$$

is increasing, and hence has a derivative almost everywhere. Moreover,

$$G_K'(x) = \lim_{n\to\infty}\frac{G_K(x+1/n)-G_K(x)}{1/n}\geq 0$$

at points x where G' exists since G is increasing. By the preceding paragraph, we know the derivative of

$$H_K(x) = \int_a^x (f\wedge K)\,dx$$

is equal to $f\wedge K$ almost everywhere. Therefore

$$F'(x) = G_K'(x) + H_K'(x) \geq (f\wedge K)(x), \qquad \text{a.e.}$$

Since K is arbitrary, $F'\geq f$ a.e., and so

$$\int_a^b F' \geq \int_a^b f = F(b)-F(a).$$

Combining with (14.4) we conclude that $\int_a^b [F'-f] = 0$. Since $F'-f\geq 0$ a.e., this tells us that $F' = f$ a.e. □

14.7 Approach 2 – absolute continuity

Finally, we continue the alternative approach to look at when $\int_a^b F'(y)\,dy = F(b) - F(a)$.

We refer the reader to Lemma 14.12 for the proof that if f is absolutely continuous, then it is of bounded variation.

Lemma 14.18 *If f is absolutely continuous on $[a,b]$ and $f'(x) = 0$ a.e., then f is constant.*

The Cantor-Lebesgue function is an example to show that we need the absolute continuity.

Proof. Let $c \in [a,b]$, let $E = \{x \in [a,c] : f'(x) = 0\}$, and let $\varepsilon > 0$. Choose δ such that $\sum_{i=1}^K |f(b_i) - f(a_i)| < \varepsilon$ whenever $\sum_{i=1}^K |b_i - a_i| \le \delta$ and the (a_i, b_i) are disjoint intervals. For each point $x \in E \cap [a,c)$ there exist arbitrarily small intervals $[x, x+h] \subset [a,c]$ such that $|f(x+h) - f(x)| < \varepsilon h$. By Lemma 14.15 we can find a finite disjoint collection of such intervals that cover all of E except for a set of measure less than δ. We label the intervals $[a_i, b_i]$ so that $a_i < b_i \le a_{i+1}$. Except for a set of measure less than δ, E is covered by $\cup_i (a_i, b_i)$. This implies that $\cup_i (b_i, a_{i+1})$ has measure less than δ, or $\sum_i |a_{i+1} - b_i| \le \delta$. By our choice of δ and the definition of absolute continuity,

$$\sum_i |f(a_{i+1}) - f(b_i)| < \varepsilon.$$

On the other hand, by our choice of the intervals (a_i, b_i),

$$\sum_i |f(b_i) - f(a_i)| < \varepsilon \sum_i (b_i - a_i) \le \varepsilon(c - a).$$

Adding these two inequalities together,

$$\begin{aligned} |f(c) - f(a)| &= \Big|\sum_i [f(a_{i+1}) - f(b_i)] + \sum_i [f(b_i) - f(a_i)]\Big| \\ &\le \varepsilon + \varepsilon(c - a). \end{aligned}$$

Since ε is arbitrary, then $f(c) = f(a)$, which implies that f is constant. □

Theorem 14.19 *If F is absolutely continuous, then*

$$F(b) - F(a) = \int_a^b F'(y)\,dy.$$

Proof. Suppose F is absolutely continuous on $[a, b]$. Then F is of bounded variation, so $F = F_1 - F_2$ where F_1 and F_2 are increasing, and F' exists a.e. Since $|F'(x)| \le F_1'(x) + F_2'(x)$, then

$$\int |F'(x)|\,dx \le (F_1(b) + F_2(b)) - (F_1(a) - F_2(a)),$$

and hence F' is integrable. If

$$G(x) = \int_a^x F'(t)\,dt,$$

then G is absolutely continuous by Exercise 14.2, and hence $F - G$ is absolutely continuous. Then $(F - G)' = F' - G' = F' - F' = 0$ a.e., using Theorem 14.17 for the second equality. By Lemma 14.18, $F - G$ is constant, and thus $F(x) - G(x) = F(a) - G(a)$. We conclude

$$F(x) = \int_a^x F'(t)\,dt + F(a).$$

If we set $x = b$, we get our result. □

14.8 Exercises

Exercise 14.1 (1) Show that if f and g are absolutely continuous on an interval $[a, b]$, then the product fg is also.
(2) Prove the *integration by parts* formula:

$$fb)g(b) - f(a)g(a) = \int_a^b f(x)g'(x)\,dx + \int_a^b f'(x)g(x)\,dx.$$

Exercise 14.2 If f is integrable and real-valued, $a \in \mathbb{R}$, and

$$F(x) = \int_a^x f(y)\,dy,$$

prove that F is of bounded variation and is absolutely continuous.

Exercise 14.3 Suppose that f is a real-valued continuous function on $[0,1]$ and that $\varepsilon > 0$. Prove that there exists a continuous function g such that $g'(x)$ exists and equals 0 for a.e. x and

$$\sup_{x\in[0,1]} |f(x) - g(x)| < \varepsilon.$$

Exercise 14.4 Suppose f is a real-valued continuous function on $[0,1]$ and f is absolutely continuous on $(a,1]$ for every $a \in (0,1)$. Is f necessarily absolutely continuous on $[0,1]$? If f is also of bounded variation on $[0,1]$, is f absolutely continuous on $[0,1]$? If not, give counterexamples.

Exercise 14.5 A real-valued function f is *Lipschitz* with constant M if

$$|f(x) - f(y)| \leq M|x - y|$$

for all $x, y \in \mathbb{R}$. Prove that f is Lipschitz with constant M if and only if f is absolutely continuous and $|f'| \leq M$ a.e.

Exercise 14.6 Suppose F_n is a sequence of increasing non-negative right continuous functions on $[0,1]$ such that $\sup_n F_n(1) < \infty$. Let $F = \sum_{n=1}^{\infty} F_n$. Prove that

$$F'(x) = \sum_{n=1}^{\infty} F_n'(x)$$

for almost every x.

Exercise 14.7 Suppose f is absolutely continuous on $[0,1]$ and for $A \subset [0,1]$ we let $f(A) = \{f(x) : x \in A\}$. Prove that if A has Lebesgue measure 0, then $f(A)$ has Lebesgue measure 0.

Exercise 14.8 If f is real-valued and differentiable at each point of $[0,1]$, is f necessarily absolutely continuous on $[0,1]$? If not, find a counterexample.

Exercise 14.9 Find an increasing function f such that $f' = 0$ a.e. but f is not constant on any open interval.

Exercise 14.10 If $f : [a,b] \to \mathbb{R}$ is continuous, let $M(y)$ be the number of points x in $[a,b]$ such that $f(x) = y$. $M(y)$ may be finite or infinite. Prove that M is Borel measurable and $\int M(y)\,dy$ equals the total variation of f on $[a,b]$.

Exercise 14.11 Let $\alpha \in (0,1)$. Find a Borel subset E of $[-1,1]$ such that

$$\lim_{r\to 0+} \frac{m(E\cap[-r,r])}{2r} = \alpha.$$

Exercise 14.12 Suppose f is a real-valued continuous function on $[a,b]$ and the derivate D^+f is non-negative on $[a,b]$. Prove that $f(b)\ge f(a)$. What if instead we have that D_+f is non-negative on $[a,b]$?

Exercise 14.13 Let

$$f(x) = \int_{-\infty}^{\infty} \frac{e^{-xy^2}}{1+y^2}\,dy.$$

(1) Find the derivative of f.
(2) Find an ordinary differential equation that f solves. Find the solution to this ordinary differential equation to determine an explicit value for $f(x)$.

Exercise 14.14 Let $(X,\mathcal{A},\mu)$ be a measure space and let f be a real-valued integrable function. Define

$$g(x) = \int |f(y)-x|\,\mu(dy)$$

for $x\in\mathbb{R}$.
(1) Prove that g is absolutely continuous.
(2) Prove that $\lim_{x\to\infty} g(x)=\infty$ and $\lim_{x\to-\infty} g(x)=\infty$.
(3) Find $g'(x)$ and prove that $g(x_0)=\inf_{x\in\mathbb{R}} g(x)$ if and only if

$$\mu(\{y: f(y)>x_0\}) = \mu(\{y: f(y)<x_0\}).$$

Exercise 14.15 Suppose $A\subset[0,1]$ has Lebesgue measure zero. Find an increasing function $f:[0,1]\to\mathbb{R}$ that is absolutely continuous, but

$$\lim_{h\to 0}\frac{f(x+h)-f(x)}{h} = \infty$$

for each $x\in A$.

Exercise 14.16 Suppose that μ is a measure on the Borel σ-algebra on $[0,1]$ and for every f that is real-valued and continuously differentiable we have

$$\Big|\int f'(x)\,\mu(dx)\Big| \le \Big(\int_0^1 f(x)^2\,dx\Big)^{1/2}.$$

(1) Show that μ is absolutely continuous with respect to Lebesgue measure on $[0,1]$.
(2) If g is the Radon-Nikodym derivative of μ with respect to Lebesgue measure, prove that there exists a constant $c > 0$ such that

$$|g(x) - g(y)| \le c|x-y|^{1/2}, \qquad x, y \in [0,1].$$

Exercise 14.17 Let $p > 1$ and $f, g \in L^p(\mathbb{R})$. Define

$$H(t) = \int_{-\infty}^{\infty} |f(x) + tg(x)|^p \, dx$$

for $t \in \mathbb{R}$. Prove that H is a differentiable function and find its derivative.

Chapter 15

L^p spaces

We introduce some spaces of functions, called the L^p spaces. We define the L^p norm of a function, prove completeness of the norm, discuss convolutions, and consider the bounded linear functionals on L^p. We assume throughout this chapter that the measure μ is σ-finite.

15.1 Norms

Let $(X, \mathcal{A}, \mu)$ be a σ-finite measure space. For $1 \le p < \infty$, define the L^p norm of f by

$$\|f\|_p = \Big(\int |f(x)|^p \, d\mu \Big)^{1/p}. \tag{15.1}$$

For $p = \infty$, define the L^∞ norm of f by

$$\|f\|_\infty = \inf\{M : \mu(\{x : |f(x)| \ge M\}) = 0\}. \tag{15.2}$$

Thus the L^∞ norm of a function f is the smallest number M such that $|f| \le M$ a.e.

For $1 \le p \le \infty$ the space L^p is the set $\{f : \|f\|_p < \infty\}$. One can also write $L^p(X)$ or $L^p(\mu)$ if one wants to emphasize the space or the measure. It is clear that $\|f\|_p = 0$ if and only if $f = 0$ a.e.

If $1 < p < \infty$, we define q by

$$\frac{1}{p} + \frac{1}{q} = 1$$

and call q the *conjugate exponent* of p.

Basic to the study of L^p spaces is Hölder's inequality. Note that when $p = q = 2$, this is the Cauchy-Schwarz inequality.

Proposition 15.1 (Hölder's inequality) *If* $1 < p, q < \infty$ *and* $p^{-1} + q^{-1} = 1$, *then*

$$\int |fg| \, d\mu \leq \|f\|_p \|g\|_q.$$

This also holds if $p = \infty$ *and* $q = 1$.

Proof. If $M = \|f\|_\infty$, then $|f| \leq M$ a.e. and $\int fg \leq M \int |g|$. The case $p = \infty$ and $q = 1$ follows.

Now let us assume $1 < p, q < \infty$. If $\|f\|_p = 0$, then $f = 0$ a.e. and $\int |fg| = 0$, so the result is clear if $\|f\|_p = 0$ and similarly if $\|g\|_q = 0$. Let $F(x) = |f(x)|/\|f\|_p$ and $G(x) = |g(x)|/\|g\|_q$. Note $\|F\|_p = 1$ and $\|G\|_q = 1$, and it suffices to show that $\int FG \, d\mu \leq 1$.

The second derivative of the function e^x is again e^x, which is everywhere positive. Any function whose second derivative is everywhere non-negative is convex, so if $0 \leq \lambda \leq 1$, we have

$$e^{\lambda a + (1-\lambda) b} \leq \lambda e^a + (1-\lambda) e^b \qquad (15.3)$$

for every pair of reals $a \leq b$. If $F(x), G(x) \neq 0$, let $a = p \log F(x)$, $b = q \log G(x)$, $\lambda = 1/p$, and $1 - \lambda = 1/q$. We then obtain from (15.3) that

$$F(x)G(x) \leq \frac{F(x)^p}{p} + \frac{G(x)^q}{q}.$$

Clearly this inequality also holds if $F(x) = 0$ or $G(x) = 0$. Integrating,

$$\int FG \, d\mu \leq \frac{\|F\|_p^p}{p} + \frac{\|G\|_q^q}{q} = \frac{1}{p} + \frac{1}{q} = 1.$$

This completes the proof. □

One application of Hölder's inequality is to prove Minkowski's inequality, which is simply the triangle inequality for L^p.

We first need the following lemma:

Lemma 15.2 *If $a, b \geq 0$ and $1 \leq p < \infty$, then*

$$(a+b)^p \leq 2^{p-1}a^p + 2^{p-1}b^p.$$

Proof. The case $a = 0$ is obvious, so we assume $a > 0$. Dividing both sides by a^p, letting $x = b/a$, and setting

$$f(x) = 2^{p-1} + 2^{p-1}x^p - (1+x)^p,$$

the inequality we want to prove is equivalent to showing $f(x) \geq 0$ for $x \geq 0$. Note $f(0) > 0$, $f(1) = 0$, $\lim_{x\to\infty} f(x) = \infty$, and the only solution to $f'(x) = 0$ on $(0, \infty)$ is $x = 1$. We conclude that f takes its minimum at $x = 1$ and hence $f(x) \geq 0$ for $x \geq 0$. □

Proposition 15.3 (Minkowski's inequality) *If $1 \leq p \leq \infty$, then*

$$\|f + g\|_p \leq \|f\|_p + \|g\|_p.$$

Proof. Since $|(f+g)(x)| \leq |f(x)| + |g(x)|$, integrating gives the case when $p = 1$. The case $p = \infty$ is also easy. Now let us suppose $1 < p < \infty$. If $\|f\|_p$ or $\|g\|_p$ is infinite, the result is obvious, so we may assume both are finite. The inequality Lemma 15.2 with $a = |f(x)|$ and $b = |g(x)|$ yields, after an integration,

$$\int |(f+g)(x)|^p \, d\mu \leq 2^{p-1} \int |f(x)|^p \, d\mu + 2^{p-1} \int |g(x)|^p \, d\mu.$$

We therefore have $\|f+g\|_p < \infty$. Clearly we may assume $\|f+g\|_p > 0$.

Now write

$$|f+g|^p \leq |f| \, |f+g|^{p-1} + |g| \, |f+g|^{p-1}$$

and apply Hölder's inequality with $q = (1 - \frac{1}{p})^{-1}$. We obtain

$$\int |f+g|^p \leq \|f\|_p \Big(\int |f+g|^{(p-1)q} \Big)^{1/q} + \|g\|_p \Big(\int |f+g|^{(p-1)q} \Big)^{1/q}.$$

Since $p^{-1} + q^{-1} = 1$, then $(p-1)q = p$, so we have

$$\|f+g\|_p^p \leq \Big(\|f\|_p + \|g\|_p \Big) \|f+g\|_p^{p/q}.$$

Dividing both sides by $\|f+g\|_p^{p/q}$ and using the fact that $p-(p/q) = 1$ gives us our result. □

Recall the definition of normed linear space from Chapter 1. We would like to say that by virtue of Minkowski's inequality, L^p is a normed linear space. This is not quite right. The L^p norm of a function satisfies all the properties of a norm except that $\|f\|_p = 0$ does not imply that f is the zero function, only that $f = 0$ a.e. The procedure we follow to circumvent this is to say two functions are equivalent if they differ on a set of measure 0. This is an equivalence relation for functions. We then define the space L^p to be the set of equivalence classes with respect to this equivalence relation, and define $\|f\|_p$ to be the L^p norm of any function in the same equivalence class as f. We then have that $\|\cdot\|_p$ is a norm on L^p. We henceforth keep this interpretation in the back of our minds when we talk about a function being in L^p; the understanding is that we identify functions that are equal a.e.

Recall Definition 10.1: f_n converges to f in L^p if $\|f_n - f\|_p \to 0$ as $n \to \infty$. In terms of L^p norms, this is equivalent to $\|f_n-f\|_p^p \to 0$ as $n \to \infty$.

Related to the definition of L^∞ is the following terminology. Given a real-valued measurable function f, the *essential supremum* and *essential infimum* are defined by

$$\text{ess sup } f = \inf\{M : \mu(\{x : f(x) > M\}) = 0\}$$

and

$$\text{ess inf } f = \sup\{m : \mu(\{x : f(x) < m\}) = 0\}.$$

15.2 Completeness

We show that the space L^p viewed as a metric space is complete.

Theorem 15.4 *If $1 \le p \le \infty$, then L^p is complete.*

Proof. We will do only the case $p < \infty$ and leave the case $p = \infty$ as Exercise 15.1.

Step 1. Suppose f_n is a Cauchy sequence in L^p. Our first step is to find a certain subsequence. Given $\varepsilon = 2^{-(j+1)}$, there exists n_j

such that if $n, m \geq n_j$, then $\|f_n - f_m\|_p \leq 2^{-(j+1)}$. Without loss of generality we may assume $n_j \geq n_{j-1}$ for each j.

Step 2. Set $n_0 = 0$ and define f_0 to be identically 0. Our candidate for the limit function will be $\sum_j [f_{n_j} - f_{n_{j-1}}]$. In this step we show absolute convergence of this series. If

$$A_j = \{x : |f_{n_j}(x) - f_{n_{j-1}}(x)| > 2^{-j/2}\},$$

then from Lemma 10.4, $\mu(A_j) \leq 2^{-jp/2}$. We have

$$\mu(\cap_{j=1}^\infty \cup_{m=j}^\infty A_m) = \lim_{j\to\infty} \mu(\cup_{m=j}^\infty A_m) \leq \lim_{j\to\infty} \sum_{m=j}^\infty \mu(A_m) = 0.$$

Thus, except for a set of measure 0, for each x there exists a j (depending on x) such that $x \notin \cup_{m=j}^\infty A_m$. Hence for each x (except for those in the null set) there is a j_0 (depending on x) such that if $j \geq j_0$, then $|f_{n_j}(x) - f_{n_{j-1}}(x)| \leq 2^{-j}$.

Set

$$g_j(x) = \sum_{m=1}^j |f_{n_m}(x) - f_{n_{m-1}}(x)|.$$

$g_j(x)$ increases in j for each x. The limit is finite for almost every x because for almost every x, $|f_{n_m}(x) - f_{n_{m-1}}(x)| \leq 2^{-m}$ for m large (how large depends on x).

Step 3. We define our function f. Set

$$f(x) = \sum_{m=1}^\infty [f_{n_m}(x) - f_{n_{m-1}}(x)].$$

We showed in Step 2 that this series is absolutely convergent for almost every x, so f is well defined for a.e. x. Set $f(x) = 0$ for any x where absolute convergence does not hold. We have

$$f(x) = \lim_{K\to\infty} \sum_{m=1}^K [f_{n_m}(x) - f_{n_{m-1}}(x)] = \lim_{K\to\infty} f_{n_K}(x)$$

since we have a telescoping series. By Fatou's lemma,

$$\begin{aligned} \|f - f_{n_j}\|_p^p &= \int |f - f_{n_j}|^p \leq \liminf_{K\to\infty} \int |f_{n_K} - f_{n_j}|^p \\ &= \liminf_{K\to\infty} \|f_{n_K} - f_{n_j}\|_p^p \leq 2^{(-j+1)p}. \end{aligned}$$

Step 4. We have thus shown that $\|f - f_{n_j}\|_p \to 0$ as $j \to \infty$. It is standard that a Cauchy sequence with a convergent subsequence itself converges. Here is the proof in our case. Given $\varepsilon > 0$, there exists N such that $\|f_n - f_m\|_p < \varepsilon$ if $m, n \geq N$. In particular, $\|f_{n_j} - f_m\|_p < \varepsilon$ if j is large enough. By Fatou's lemma,

$$\|f - f_m\|_p^p \leq \liminf_{j\to\infty} \|f_{n_j} - f_m\|_p^p \leq \varepsilon^p$$

if $m \geq N$. This shows that f_m converges to f in L^p norm. □

Next we show:

Proposition 15.5 *The set of continuous functions with compact support is dense in $L^p(\mathbb{R})$.*

Proof. Suppose $f \in L^p$. We have $\int |f - f\chi_{[-n,n]}|^p \to 0$ as $n \to \infty$ by dominated convergence, the dominating function being $|f|^p$. Hence it suffices to approximate functions in L^p that have compact support. By writing $f = f^+ - f^-$ we may suppose $f \geq 0$. Consider simple functions s_m increasing to f; then we have $\int |f - s_m|^p \to 0$ by dominated convergence, so it suffices to approximate simple functions with compact support. By linearity, it suffices to approximate characteristic functions with compact support. Given E, a Borel measurable set contained in a bounded interval, and $\varepsilon > 0$, we showed in Proposition 8.4 that there exists g continuous with compact support and with values in $[0,1]$ such that $\int |g - \chi_E| < \varepsilon$. Since $|g - \chi_E| \leq 1$, then $\int |g - \chi_E|^p \leq \int |g - \chi_E| < \varepsilon$. This completes the proof. □

The same proof shows the following corollary.

Corollary 15.6 *The set of continuous functions on $[a,b]$ are dense in the space $L^2([a,b])$ with respect to $L^2([a,b])$ norm.*

15.3 Convolutions

The *convolution* of two measurable functions f and g is defined by

$$f * g(x) = \int f(x-y)g(y)\,dy,$$

provided the integral exists. By a change of variables, this is the same as $\int f(y)g(x-y)\,dy$, so $f*g = g*f$.

Proposition 15.7 *If $f, g \in L^1$, then $f*g$ is in L^1 and*

$$\|f*g\|_1 \le \|f\|_1\|g\|_1. \tag{15.4}$$

Proof. We have

$$\int |f*g(x)|\,dx \le \int\int |f(x-y)|\,|g(y)|\,dy\,dx. \tag{15.5}$$

Since the integrand on the right is non-negative, we can apply the Fubini theorem to see that the right hand side is equal to

$$\begin{aligned}\int\int |f(x-y)|\,dx\,|g(y)|\,dy &= \int\int |f(x)|\,dx\,|g(y)|\,dy \qquad (15.6)\\ &= \|f\|_1\|g\|_1.\end{aligned}$$

The first equality here follows by a change of variables (see Exercise 8.1). This together with (15.5) proves (15.4). From (15.4) we conclude that $f*g$ is finite a.e. □

15.4 Bounded linear functionals

A *linear functional* on L^p is a map H from L^p to $\mathbb{R}$ satisfying

$$H(f+g) = H(f) + H(g), \qquad H(af) = aH(f)$$

whenever $f, g \in L^p$ and $a \in \mathbb{R}$. (One can also have complex-valued linear functionals, but we do not consider them in this section. See, however, Exercise 15.28.) H is a *bounded linear functional* if

$$\|H\| = \sup\{|Hf| : \|f\|_p \le 1\} \tag{15.7}$$

is finite. The dual space of L^p is the collection of all bounded linear functionals with norm given by (15.7). Our goal in this section is to identify the dual of L^p.

We define the *signum function* or *sign function* by

$$\operatorname{sgn}(x) = \begin{cases} -1, & x < 0; \\ 0, & x = 0; \\ 1, & x > 0. \end{cases}$$

Note $x \operatorname{sgn}(x) = |x|$.

The following is very useful.

Theorem 15.8 *For $1 < p < \infty$ and $p^{-1} + q^{-1} = 1$,*

$$\|f\|_p = \sup\Big\{\int fg\,d\mu : \|g\|_q \le 1\Big\}. \tag{15.8}$$

When $p = 1$, (15.8) holds if we take $q = \infty$, and if $p = \infty$, (15.8) holds if we take $q = 1$.

Proof. The right hand side of (15.8) is less than the left hand side by Hölder's inequality. Thus we need only show that the right hand side is greater than the left hand side.

Case 1: $p = 1$. Take $g(x) = \operatorname{sgn} f(x)$. Then g is bounded by 1 and $fg = |f|$. This takes care of the case $p = 1$.

Case 2: $p = \infty$. If $\|f\|_\infty = 0$, the result is trivial, so suppose $\|f\|_\infty > 0$. Since μ is σ-finite, there exist sets F_n increasing up to X such that $\mu(F_n) < \infty$ for each n. If $M = \|f\|_\infty$, let a be any finite real less than M. By the definition of L^∞ norm, the measure of $A_n = \{x \in F_n : |f(x)| > a\}$ must be positive if n is sufficiently large. Let

$$g_n(x) = \frac{\operatorname{sgn} f(x)\chi_{A_n}(x)}{\mu(A_n)}.$$

Then the L^1 norm of g_n is 1 and $\int fg_n = \int_{A_n} |f|/\mu(A_n) \ge a$. Since a is arbitrary, the supremum on the right hand side of (15.8) must be M.

Case 3: $1 < p < \infty$. We may suppose $\|f\|_p > 0$. Let F_n be measurable sets of finite measure increasing to X, q_n a sequence of non-negative simple functions increasing to f^+, r_n a sequence of non-negative simple functions increasing to f^-, and

$$s_n(x) = (q_n(x) - r_n(x))\chi_{F_n}(x).$$

Then $s_n(x) \to f(x)$ for each x, $|s_n(x)|$ increases to $|f(x)|$ for each x, each s_n is a simple function, and $\|s_n\|_p < \infty$ for each n. Then $\|s_n\|_p \to \|f\|_p$ by monotone convergence, whether or not $\|f\|_p$ is finite. For n sufficiently large, $\|s_n\|_p > 0$.

Let

$$g_n(x) = (\operatorname{sgn} f(x))\frac{|s_n(x)|^{p-1}}{\|s_n\|_p^{p/q}}.$$

g_n is again a simple function. Since $(p-1)q = p$, then

$$\|g_n\|_q = \frac{(\int |s_n|^{(p-1)q})^{1/q}}{\|s_n\|_p^{p/q}} = \frac{\|s_n\|_p^{p/q}}{\|s_n\|_p^{p/q}} = 1.$$

On the other hand, since $|f| \geq |s_n|$,

$$\int f g_n = \frac{\int |f|\,|s_n|^{p-1}}{\|s_n\|_p^{p/q}} \geq \frac{\int |s_n|^p}{\|s_n\|_p^{p/q}} = \|s_n\|_p^{p-(p/q)}.$$

Since $p - (p/q) = 1$, then $\int f g_n \geq \|s_n\|_p$, which tends to $\|f\|_p$. This proves the right hand side of (15.8) is at least as large as the left hand side. □

The proof of Theorem 15.8 also establishes

Corollary 15.9 *For $1 < p < \infty$ and $p^{-1} + q^{-1} = 1$,*

$$\|f\|_p = \sup\Big\{ \int fg : \|g\|_q \leq 1, g \text{ simple}\Big\}.$$

Proposition 15.10 *Suppose $1 < p < \infty$, $p^{-1} + q^{-1} = 1$, and $g \in L^q$. If we define $H(f) = \int fg$ for $f \in L^p$, then H is a bounded linear functional on L^p and $\|H\| = \|g\|_q$.*

Proof. The linearity is obvious. That $\|H\| \leq \|g\|_q$ follows by Hölder's inequality. Using Theorem 15.8 and writing

$$\|H\| = \sup_{\|f\|_p \leq 1} |H(f)| = \sup_{\|f\|_p \leq 1} \Big| \int fg \Big| \geq \sup_{\|f\|_p \leq 1} \int fg = \|g\|_q$$

completes the proof. □

Theorem 15.11 *Suppose $1 < p < \infty$, $p^{-1} + q^{-1} = 1$, and H is a real-valued bounded linear functional on L^p. Then there exists $g \in L^q$ such that $H(f) = \int fg$ and $\|g\|_q = \|H\|$.*

This theorem together with Proposition 15.10 allows us to identify the dual space of L^p with L^q.

Proof. Suppose we are given a bounded linear functional H on L^p. First suppose $\mu(X) < \infty$. Define $\nu(A) = H(\chi_A)$. We will show that ν is a measure, that $\nu \ll \mu$ and that $g = d\nu/d\mu$ is the function we seek.

If A and B are disjoint, then

$$\begin{aligned}\nu(A \cup B) &= H(\chi_{A\cup B}) = H(\chi_A + \chi_B)\\ &= H(\chi_A) + H(\chi_B) = \nu(A) + \nu(B).\end{aligned}$$

To show ν is countably additive, it suffices to show that if $A_n \uparrow A$, then $\nu(A_n) \to \nu(A)$, and then use Exercise 3.1. But if $A_n \uparrow A$, then $\chi_{A_n} \to \chi_A$ in L^p, and so $\nu(A_n) = H(\chi_{A_n}) \to H(\chi_A) = \nu(A)$; we use here the fact that $\mu(X) < \infty$. We conclude that ν is a countably additive signed measure. Moreover, if $\mu(A) = 0$, then $\chi_A = 0$ a.e., hence $\nu(A) = H(\chi_A) = 0$. Using Exercise 13.1, which is the Radon-Nikodym theorem for signed measures, we see there exists a real-valued integrable function g such that $\nu(A) = \int_A g$ for all sets A.

If $s = \sum_i a_i \chi_{A_i}$ is a simple function, by linearity we have

$$H(s) = \sum_i a_i H(\chi_{A_i}) = \sum_i a_i \nu(A_i) = \sum_i a_i \int g\chi_{A_i} = \int gs. \tag{15.9}$$

By Corollary 15.9 and (15.9),

$$\begin{aligned}\|g\|_q &= \sup\Big\{\int gs : \|s\|_p \le 1, s \text{ simple}\Big\}\\ &= \sup\{H(s) : \|s\|_p \le 1, s \text{ simple}\} \le \|H\|.\end{aligned}$$

If s_n are simple functions tending to f in L^p (see Exercise 15.2), then $H(s_n) \to H(f)$, while by Hölder's inequality

$$\Big|\int s_n g - \int fg\Big| = \Big|\int (s_n - f)g\Big| \le \|s_n - f\|_p \|g\|_q \to 0,$$

so $\int s_n g \to \int fg$. We thus have $H(f) = \int fg$ for all $f \in L^p$, and $\|g\|_q \le \|H\|$. By Hölder's inequality, $\|H\| \le \|g\|_q$.

In the case where μ is σ-finite, but not necessarily finite, let $F_n \uparrow X$ so that $\mu(F_n) < \infty$ for each n. Define functionals H_n by $H_n(f) = H(f\chi_{F_n})$. Clearly each H_n is a bounded linear functional on L^p. Applying the above argument, we see there exist g_n such that $H_n(f) = \int fg_n$ and $\|g_n\|_q = \|H_n\| \le \|H\|$. It is easy to see that g_n is 0 if $x \notin F_n$. Moreover, by the uniqueness part of the Radon-Nikodym theorem, if $n > m$, then $g_n = g_m$ on F_m. Define g by setting $g(x) = g_n(x)$ if $x \in F_n$. Then g is well defined. By Fatou's lemma, g is in L^q with a norm bounded by $\|H\|$. Note $f\chi_{F_n} \to f$ in L^p by dominated convergence. Since H is a bounded linear functional on L^p, we have $H_n(f) = H(f\chi_{F_n}) \to H(f)$. On the other hand

$$H_n(f) = \int_{F_n} fg_n = \int_{F_n} fg \to \int fg$$

by dominated convergence. Thus $H(f) = \int fg$. Again by Hölder's inequality $\|H\| \le \|g\|_q$. $\square$

15.5 Exercises

Exercise 15.1 Show that L^∞ is complete.

Exercise 15.2 Prove that the collection of simple functions is dense in L^p.

Exercise 15.3 Prove the equality

$$\int |f(x)|^p \, dx = \int_0^\infty pt^{p-1} \, m(\{x : |f(x)| \ge t\}) \, dt$$

for $p \ge 1$.

Exercise 15.4 Consider the measure space $([0,1], \mathcal{B}, m)$, where $\mathcal{B}$ is the Borel σ-algebra and m is Lebesgue measure, and suppose f is a measurable function. Prove that $\|f\|_p \to \|f\|_\infty$ as $p \to \infty$.

Exercise 15.5 When does equality hold in Hölder's inequality? When does equality hold in the Minkowski inequality?

Exercise 15.6 Give an example to show that $L^p \not\subset L^q$ in general if $1 < p < q < \infty$. Give an example to show that $L^q \not\subset L^p$ in general if $1 < p < q < \infty$.

Exercise 15.7 Define

$$g_n(x) = n\chi_{[0,n^{-3}]}(x).$$

(1) Show that if $f \in L^2([0,1])$, then

$$\int_0^1 f(x)g_n(x)\,dx \to 0$$

as $n \to \infty$.
(2) Show that there exists $f \in L^1([0,1])$ such that $\int_0^1 f(x)g_n(x)\,dx \not\to 0$.

Exercise 15.8 Suppose μ is a finite measure on the Borel subsets of $\mathbb{R}$ such that

$$f(x) = \int_{\mathbb{R}} f(x+t)\,\mu(dt), \qquad \text{a.e.},$$

whenever f is real-valued, bounded, and integrable. Prove that $\mu(\{0\}) = 1$.

Exercise 15.9 Suppose μ is a measure with $\mu(X) = 1$ and $f \in L^r$ for some $r > 0$, where we define L^r for $r < 1$ exactly as in (15.1). Prove that

$$\lim_{p\to 0} \|f\|_p = \exp\Big(\int \log|f|\,d\mu\Big),$$

where we use the convention that $\exp(-\infty) = 0$.

Exercise 15.10 Suppose $1 < p < \infty$ and q is the conjugate exponent to p. Suppose $f_n \to f$ a.e. and $\sup_n \|f_n\|_p < \infty$. Prove that if $g \in L^q$, then

$$\lim_{n\to\infty} \int f_n g = \int fg.$$

Does this extend to the case where $p = 1$ and $q = \infty$? If not, give a counterexample.

Exercise 15.11 If $f \in L^1(\mathbb{R})$ and $g \in L^p(\mathbb{R})$ for some $p \in [1,\infty)$, prove that

$$\|f * g\|_p \le \|f\|_1 \|g\|_p.$$

Exercise 15.12 Suppose $p \in (1,\infty)$ and q is its conjugate exponent. Prove that if $f \in L^p(\mathbb{R})$ and $g \in L^q(\mathbb{R})$, then $f*g$ is uniformly continuous and $f*g(x) \to 0$ as $x \to \infty$ and as $x \to -\infty$.

Exercise 15.13 Show that if f and g are continuous with compact support, then $f * g$ is continuous with compact support.

Exercise 15.14 Suppose $f \in L^\infty(\mathbb{R})$, $f_h(x) = f(x+h)$, and

$$\lim_{h \to 0} \|f_h - f\|_\infty = 0.$$

Prove that there exists a uniformly continuous function g on $\mathbb{R}$ such that $f = g$ a.e.

Exercise 15.15 Let $p \in [1, \infty)$. Prove that $f \in L^p(\mu)$ if and only if

$$\sum_{n=1}^\infty (2^n)^p \mu(\{x : |f(x)| > 2^n\}) < \infty.$$

Exercise 15.16 Suppose $\mu(X) = 1$ and f and g are non-negative functions such that $fg \ge 1$ a.e. Prove that

$$\Big(\int f\, d\mu\Big)\Big(\int g\, d\mu\Big) \ge 1.$$

Exercise 15.17 Suppose $f : [1, \infty) \to \mathbb{R}$, $f(1) = 0$, f' exists and is continuous and bounded, and $f' \in L^2([1, \infty))$. Let $g(x) = f(x)/x$. Show $g \in L^2([1, \infty))$.

Exercise 15.18 Find an example of a measurable $f : [1, \infty) \to \mathbb{R}$ such that $f(1) = 0$, f' exists and is continuous and bounded, $f' \in L^1([1, \infty))$, but the function $g(x) = f(x)/x$ is not in L^1.

Exercise 15.19 Prove the *generalized Minkowski inequality*: If $(X, \mathcal{A}, \mu)$ and $(Y, \mathcal{B}, \nu)$ are measure spaces, f is measurable with respect to $\mathcal{A} \times \mathcal{B}$, and $1 < p < \infty$, then

$$\Big(\int_X \Big(\int_Y |f(x,y)|\, \nu(dy)\Big)^p \mu(dx)\Big)^{1/p} \le \int_Y \Big(\int_X |f(x,y)|^p\, \mu(dx)\Big)^{1/p} \nu(dy).$$

Does this extend to the cases where $p = 1$ or $p = \infty$? If not, give counterexamples.

If $Y = \{1, 2\}$, $\nu(dy) = \delta_1(dy) + \delta_2(dy)$, where δ_1 and δ_2 are point masses at 1 and 2, resp. and we let $g_1(x) = f(x, 1)$, $g_2(x) = f(x, 2)$, we recover the usual Minkowski inequality, Proposition 15.3.

Exercise 15.20 Let $\alpha \in (0,1)$ and $K(x) = |x|^{-\alpha}$ for $x \in \mathbb{R}$. Note that K is not in L^p for any $p \geq 1$. Prove that if f is non-negative, real-valued, and integrable on $\mathbb{R}$ and

$$g(x) = \int f(x-t)K(t)\,dt,$$

then g is finite a.e.

Exercise 15.21 Suppose $p > 1$ and q is its conjugate exponent, f is an absolutely continuous function on $[0,1]$ with $f' \in L^p$, and $f(0) = 0$. Prove that if $g \in L^q$, then

$$\int_0^1 |fg|\,dx \leq \Big(\frac{1}{p}\Big)^{1/p} \|f'\|_p \|g\|_q.$$

Exercise 15.22 Suppose $f : \mathbb{R} \to \mathbb{R}$ is in L^p for some $p > 1$ and also in L^1. Prove there exist constants $c > 0$ and $\alpha \in (0,1)$ such that

$$\int_A |f(x)|\,dx \leq cm(A)^\alpha$$

for every Borel measurable set $A \subset \mathbb{R}$, where m is Lebesgue measure.

Exercise 15.23 Suppose $f : \mathbb{R} \to \mathbb{R}$ is integrable and there exist constants $c > 0$ and $\alpha \in (0,1)$ such that

$$\int_A |f(x)|\,dx \leq cm(A)^\alpha$$

for every Borel measurable set $A \subset \mathbb{R}$, where m is Lebesgue measure. Prove there exists $p > 1$ such that $f \in L^p$.

Exercise 15.24 Suppose $1 < p < \infty$, $f : (0,\infty) \to \mathbb{R}$, and $f \in L^p$ with respect to Lebesgue measure. Define

$$g(x) = \frac{1}{x} \int_0^x f(y)\,dy.$$

Prove that

$$\|g\|_p \leq \frac{p}{p-1} \|f\|_p.$$

This is known as *Hardy's inequality.*

Exercise 15.25 Suppose $(X, \mathcal{A}, \mu)$ is a measure space and suppose $K : X \times X \to \mathbb{R}$ is measurable with respect to $\mathcal{A} \times \mathcal{A}$. Suppose there exists $M < \infty$ such that

$$\int_X |K(x,y)|\,\mu(dy) \le M$$

for each x and

$$\int_X |K(x,y)|\,\mu(dx) \le M$$

for each y. If f is measurable and real-valued, define

$$Tf(x) = \int_X K(x,y) f(y)\,\mu(dy)$$

if the integral exists.
(1) Show that $\|Tf\|_1 \le M\|f\|_1$.
(2) If $1 < p < \infty$, show that $\|Tf\|_p \le M\|f\|_p$.

Exercise 15.26 Suppose A and B are two Borel measurable subsets of $\mathbb{R}$ with finite strictly positive Lebesgue measure. Show that $\chi_A * \chi_B$ is a continuous non-negative function that is not identically equal to 0.

Exercise 15.27 Suppose A and B are two Borel measurable subsets of $\mathbb{R}$ with strictly positive Lebesgue measure. Show that

$$C = \{x + y : x \in A, y \in B\}$$

contains a non-empty open interval.

Exercise 15.28 Suppose $1 < p < \infty$ and q is the conjugate exponent of p. Prove that if H is a bounded complex-valued linear functional on L^p, then there exists a complex-valued measurable function $g \in L^q$ such that $H(f) = \int fg$ for all $f \in L^p$ and $\|H\| = \|g\|_q$.

Chapter 16

Fourier transforms

Fourier transforms give a representation of a function in terms of frequencies. There is a great deal known about them and their applications. We give an introduction here.

16.1 Basic properties

If f is a complex-valued function and $f \in L^1(\mathbb{R}^n)$, define the Fourier transform $\widehat{f}$ to be the function with domain $\mathbb{R}^n$ and range $\mathbb{C}$ given by

$$\widehat{f}(u) = \int_{\mathbb{R}^n} e^{iu\cdot x} f(x)\, dx, \qquad u \in \mathbb{R}^n. \tag{16.1}$$

We are using $u \cdot x$ for the standard inner product in $\mathbb{R}^n$. Various books have slightly different definitions. Some put a negative sign and/or 2π before the $iu \cdot x$, some have a $(2\pi)^{-1}$ or a $(2\pi)^{-1/2}$ in front of the integral. The basic theory is the same in any case.

Some basic properties of the Fourier transform are given by

Proposition 16.1 *Suppose f and g are in L^1. Then*
(1) $\widehat{f}$ is bounded and continuous;
(2) $\widehat{(f+g)}(u) = \widehat{f}(u) + \widehat{g}(u)$;
(3) $\widehat{(af)}(u) = a\widehat{f}(u)$ if $a \in \mathbb{C}$;
(4) if $a \in \mathbb{R}^n$ and $f_a(x) = f(x+a)$, then $\widehat{f_a}(u) = e^{-iu\cdot a}\widehat{f}(u)$;

(5) if $a \in \mathbb{R}^n$ and $g_a(x) = e^{ia\cdot x}g(x)$, then $\widehat{g}_a(u) = \widehat{f}(u+a)$;
(6) if a is a nonzero real number and $h_a(x) = f(ax)$, then $\widehat{h}_a(u) = a^{-n}\widehat{f}(u/a)$.

Proof. (1) $\widehat{f}$ is bounded because $f \in L^1$ and $|e^{iu\cdot x}| = 1$. We have

$$\widehat{f}(u+h) - \widehat{f}(u) = \int \Big(e^{i(u+h)\cdot x} - e^{iu\cdot x}\Big) f(x)\, dx.$$

Then

$$|\widehat{f}(u+h) - \widehat{f}(u)| \le \int \big|e^{iu\cdot x}\big| \cdot \big|e^{ih\cdot x} - 1\big|\, |f(x)|\, dx.$$

The integrand is bounded by $2|f(x)|$, which is integrable, and $e^{ih\cdot x} - 1 \to 0$ as $h \to 0$. Thus the continuity follows by dominated convergence.

(2) and (3) are easy by a change of variables. (4) holds because

$$\widehat{f}_a(u) = \int e^{iu\cdot x} f(x+a)\, dx = \int e^{iu\cdot(x-a)} f(x)\, dx = e^{-iu\cdot a}\widehat{f}(u)$$

by a change of variables. For (5),

$$\widehat{g}_a(u) = \int e^{iu\cdot x}e^{ia\cdot x} f(x)\, dx = \int e^{i(u+a)\cdot x} f(x)\, dx = \widehat{f}(u+a).$$

Finally for (6), by a change of variables,

$$\begin{aligned}\widehat{h}_a(u) &= \int e^{iu\cdot x} f(ax)\, dx = a^{-n}\int e^{iu\cdot(y/a)} f(y)\, dy\\ &= a^{-n}\int e^{i(u/a)\cdot y} f(y)\, dy = a^{-n}\widehat{f}(u/a),\end{aligned}$$

as required. □

One reason for the usefulness of Fourier transforms is that they relate derivatives and multiplication.

Proposition 16.2 *Suppose $f \in L^1$ and $x_j f(x) \in L^1$, where x_j is the j^{th} coordinate of x. Then*

$$\frac{\partial \widehat{f}}{\partial u_j}(u) = i\int e^{iu\cdot x} x_j f(x)\, dx.$$

Proof. Let e_j be the unit vector in the j^{th} direction. Then

$$\frac{\widehat{f}(u+he_j)-\widehat{f}(u)}{h} = \frac{1}{h}\int \Big(e^{i(u+he_j)\cdot x} - e^{iu\cdot x}\Big) f(x)\,dx = \int e^{iu\cdot x}\Big(\frac{e^{ihx_j}-1}{h}\Big) f(x)\,dx.$$

Since

$$\Big|\frac{1}{h}\Big(e^{ihx_j}-1\Big)\Big| \le |x_j|$$

and $x_j f(x) \in L^1$, the right hand side converges to $\int e^{iu\cdot x} i x_j f(x)\,dx$ by dominated convergence. Therefore the left hand side converges. Of course, the limit of the left hand side is $\partial \widehat{f}/\partial u_j$. □

Proposition 16.3 *Suppose $f : \mathbb{R} \to \mathbb{R}$ is integrable, f is absolutely continuous, and f' is integrable. Then the Fourier transform of f' is $-iu\widehat{f}(u)$.*

The higher dimensional version of this is left as Exercise 16.4.

Proof. Since f' is integrable,

$$|f(y) - f(x)| \le \int_x^y |f'(z)|\,dz \to 0$$

as $x, y \to \infty$ by dominated convergence. This implies that $f(y_n)$ is a Cauchy sequence whenever $y_n \to \infty$, and we conclude that $f(y)$ converges as $y \to \infty$. Since f is integrable, the only possible value for the limit is 0. The same is true for the limit as $y \to -\infty$.

By integration by parts (use Exercise 14.1 and a limit argument),

$$\widehat{f'}(u) = \int_{-\infty}^{\infty} e^{iux} f'(x)\,dx = -\int_{-\infty}^{\infty} iue^{iux} f(x)\,dx = -iu\widehat{f}(u),$$

as desired. □

Recall the definition of convolution given in Section 15.3. Recall also (15.6), which says that

$$\int\int |f(x-y)|\,|g(y)|\,dx\,dy = \|f\|_1 \|g\|_1. \qquad (16.2)$$

Proposition 16.4 *If $f, g \in L^1$, then the Fourier transform of $f * g$ is $\widehat{f}(u)\widehat{g}(u)$.*

Proof. We have

$$\begin{aligned}\widehat{f*g}(u) &= \int e^{iu\cdot x} \int f(x-y)g(y)\,dy\,dx\\ &= \int\int e^{iu\cdot(x-y)} f(x-y)\, e^{iu\cdot y} g(y)\,dx\,dy\\ &= \int \widehat{f}(u) e^{iu\cdot y} g(y)\,dy = \widehat{f}(u)\widehat{g}(u).\end{aligned}$$

We applied the Fubini theorem in the second equality; this is valid because as we see from (16.2), the absolute value of the integrand is integrable. We used a change of variables to obtain the third equality. □

16.2 The inversion theorem

We want to give a formula for recovering f from $\widehat{f}$. First we need to calculate the Fourier transform of a particular function.

Proposition 16.5 *(1) Suppose $f_1 : \mathbb{R} \to \mathbb{R}$ is defined by*

$$f_1(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}.$$

Then $\widehat{f}_1(u) = e^{-u^2/2}$.

(2) Suppose $f_n : \mathbb{R}^n \to \mathbb{R}$ is given by

$$f_n(x) = \frac{1}{(2\pi)^{n/2}} e^{-|x|^2/2}.$$

Then $\widehat{f}_n(u) = e^{-|u|^2/2}$.

Proof. (1) may also be proved using contour integration, but let's give a (mostly) real variable proof. Let $g(u) = \int e^{iux} e^{-x^2/2}\,dx$. Differentiate with respect to u. We may differentiate under the integral sign because $(e^{i(u+h)x} - e^{iux})/h$ is bounded in absolute value by $|x|$ and $|x|e^{-x^2/2}$ is integrable; therefore dominated convergence applies. We then obtain

$$g'(u) = i\int e^{iux} x e^{-x^2/2}\,dx.$$

By integration by parts (see Exercise 14.1) this is equal to

$$-u\int e^{iux} e^{-x^2/2}\,dx = -ug(u).$$

Solving the differential equation $g'(u) = -ug(u)$, we have

$$[\log g(u)]' = \frac{g'(u)}{g(u)} = -u,$$

so $\log g(u) = -u^2/2 + c_1$, and then

$$g(u) = c_2 e^{-u^2/2}. \qquad (16.3)$$

By Exercise 11.18, $g(0) = \int e^{-x^2/2}\,dx = \sqrt{2\pi}$, so $c_2 = \sqrt{2\pi}$. Substituting this value of c_2 in (16.3) and dividing both sides by $\sqrt{2\pi}$ proves (1).

For (2), since $f_n(x) = f_1(x_1)\cdots f_1(x_n)$ if $x = (x_1, \ldots, x_n)$, then

$$\begin{aligned}\widehat{f_n}(u) &= \int \cdots \int e^{i\sum_j u_j x_j} f_1(x_1)\cdots f_1(x_n)\,dx_1\cdots dx_n\\ &= \widehat{f_1}(u_1)\cdots \widehat{f_1}(u_n) = e^{-|u|^2/2}.\end{aligned}$$

This completes the proof. □

One more preliminary is needed before proving the inversion theorem.

Proposition 16.6 *Suppose φ is in L^1 and $\int \varphi(x)\,dx = 1$. Let $\varphi_\delta(x) = \delta^{-n}\varphi(x/\delta)$.*

(1) If g is continuous with compact support, then $g*\varphi_\delta$ converges to g pointwise as $\delta \to 0$.

(2) If g is continuous with compact support, then $g*\varphi_\delta$ converges to g in L^1 as $\delta \to 0$.

(3) If $f \in L^1$, then $\|f*\varphi_\delta - f\|_1 \to 0$ as $\delta \to 0$.

Proof. (1) We have by a change of variables (Exercise 8.1) that $\int \varphi_\delta(y)\,dy = 1$. Then

$$\begin{aligned}|g*\varphi_\delta(x)-g(x)| &= \Big|\int (g(x-y)-g(x))\,\varphi_\delta(y)\,dy\Big|\\ &= \Big|\int (g(x-\delta y)-g(x))\,\varphi(y)\,dy\Big|\\ &\le \int |g(x-\delta y)-g(x)|\,|\varphi(y)|\,dy.\end{aligned}$$

Since g is continuous with compact support and hence bounded and φ is integrable, the right hand side goes to zero by dominated convergence, the dominating function being $2\|g\|_\infty\varphi$.

(2) We now use the Fubini theorem to write

$$\begin{aligned}\int |g*\varphi_\delta(x)-g(x)|\,dx &= \int \Big|\int (g(x-y)-g(x))\,\varphi_\delta(y)\,dy\Big|\,dx\\ &= \int \Big|\int (g(x-\delta y)-g(x))\,\varphi(y)\,dy\Big|\,dx\\ &\le \int\int |g(x-\delta y)-g(x)|\,|\varphi(y)|\,dy\,dx\\ &= \int\int |g(x-\delta y)-g(x)|\,dx\,|\varphi(y)|\,dy.\end{aligned}$$

Let

$$G_\delta(y) = \int |g(x-\delta y)-g(x)|\,dx.$$

By dominated convergence, for each y, $G_\delta(y)$ tends to 0 as $\delta \to 0$, since g is continuous with compact support. Moreover G_δ is bounded in absolute value by $2\|g\|_1$. Using dominated convergence again and the fact that φ is integrable, we see that $\int G_\delta(y)\,|\varphi(y)|\,dy$ tends to 0 as $\delta \to 0$.

(3) Let $\varepsilon > 0$. Let g be a continuous function with compact support so that $\|f-g\|_1 < \varepsilon$. Let $h = f - g$. A change of variables shows that $\|\varphi_\delta\|_1 = \|\varphi\|_1$. Observe

$$\|f*\varphi_\delta - f\|_1 \le \|g*\varphi_\delta - g\|_1 + \|h*\varphi_\delta - h\|_1.$$

Also

$$\|h*\varphi_\delta - h\|_1 \le \|h\|_1 + \|h*\varphi_\delta\|_1 \le \|h\|_1 + \|h\|_1\|\varphi_\delta\|_1 < \varepsilon(1+\|\varphi\|_1).$$

Therefore, using (2),

$$\limsup_{\delta\to 0} \|f * \varphi_\delta - f\|_1 \le \limsup_{\delta\to 0} \|h * \varphi_\delta - h\|_1 \le \varepsilon(1 + \|\varphi\|_1).$$

Since ε is arbitrary, we have our conclusion. □

Now we are ready to give the inversion formula. The proof seems longer than one might expect it to be, but there is no avoiding the introduction of the function H_a or some similar function.

Theorem 16.7 *Suppose f and $\widehat{f}$ are both in L^1. Then*

$$f(y) = \frac{1}{(2\pi)^n} \int e^{-iu\cdot y} \widehat{f}(u)\, du, \qquad \text{a.e.}$$

Proof. If $g(x) = a^{-n}k(x/a)$, then the Fourier transform of g is $\widehat{k}(au)$. Hence the Fourier transform of

$$\frac{1}{a^n} \frac{1}{(2\pi)^{n/2}} e^{-x^2/2a^2}$$

is $e^{-a^2u^2/2}$. If we let

$$H_a(x) = \frac{1}{(2\pi)^n} e^{-|x|^2/2a^2},$$

we have

$$\widehat{H}_a(u) = (2\pi)^{-n/2} a^n e^{-a^2|u|^2/2}.$$

We write

$$\begin{aligned}\int \widehat{f}(u) e^{-iu\cdot y} H_a(u)\, du &= \int\int e^{iu\cdot x} f(x) e^{-iu\cdot y} H_a(u)\, dx\, du \\ &= \int\int e^{iu\cdot(x-y)} H_a(u)\, du\, f(x)\, dx \\ &= \int \widehat{H}_a(x-y) f(x)\, dx. \qquad (16.4)\end{aligned}$$

We can interchange the order of integration because

$$\int\int |f(x)|\, |H_a(u)|\, dx\, du < \infty$$

and $|e^{iu\cdot x}| = 1$. The left hand side of the first line of (16.4) converges to $(2\pi)^{-n}\int \widehat{f}(u)e^{-iu\cdot y}\,dy$ as $a \to \infty$ by dominated convergence since $H_a(u) \to (2\pi)^{-n}$ and $\widehat{f} \in L^1$. The last line of (16.4) is equal to

$$\int \widehat{H}_a(y-x)f(x)\,dx = f * \widehat{H}_a(y), \tag{16.5}$$

using that $\widehat{H}_a$ is symmetric. But by Proposition 16.6, setting $\delta = a^{-1}$, we see that $f * \widehat{H}_a$ converges to f in L^1 as $a \to \infty$. □

16.3 The Plancherel theorem

The last topic that we consider is the *Plancherel theorem.*

Theorem 16.8 *Suppose f is continuous with compact support. Then $\widehat{f} \in L^2$ and*

$$\|f\|_2 = (2\pi)^{-n/2}\|\widehat{f}\|_2. \tag{16.6}$$

Proof. First note that if we combine (16.4) and (16.5), take $y = 0$, and use the symmetry of $\widehat{H}_a$, we obtain

$$\int \widehat{f}(u)H_a(u)\,du = f * \widehat{H}_a(0). \tag{16.7}$$

Let $g(x) = \overline{f(-x)}$, where $\overline{a}$ denotes the complex conjugate of a. Since $\overline{ab} = \overline{a}\,\overline{b}$,

$$\begin{aligned}\widehat{g}(u) &= \int e^{iu\cdot x}\overline{f(-x)}\,dx = \overline{\int e^{-iu\cdot x}f(-x)\,dx}\\ &= \overline{\int e^{iu\cdot x}f(x)\,dx} = \overline{\widehat{f}(u)}.\end{aligned}$$

The third equality follows by a change of variables. By (16.7) with f replaced by $f * g$,

$$\int \widehat{f * g}(u)H_a(u)\,du = f * g * \widehat{H}_a(0). \tag{16.8}$$

Since $\widehat{f * g}(u) = \widehat{f}(u)\widehat{g}(u) = |\widehat{f}(u)|^2$, the left hand side of (16.8) converges by monotone convergence to $(2\pi)^{-n}\int |\widehat{f}(u)|^2\,du$ as $a \to$

∞. Since f and g are continuous with compact support, then by Exercise 15.13, $f * g$ is also, and so the right hand side of (16.8) converges to $f * g(0) = \int f(y)g(-y)\,dy = \int |f(y)|^2\,dy$ by Proposition 16.6(2). $\square$

Remark 16.9 We can use Theorem 16.8 to define $\widehat{f}$ when $f \in L^2$ so that (16.6) will continue to hold. The set of continuous functions with compact support is dense in L^2 by Proposition 15.5. Given a function f in L^2, choose a sequence $\{f_m\}$ of continuous functions with compact support such that $f_m \to f$ in L^2. Then $\|f_m - f_n\|_2 \to 0$ as $m, n \to \infty$. By (16.6), $\{\widehat{f_m}\}$ is a Cauchy sequence in L^2, and therefore converges to a function in L^2, which we call $\widehat{f}$.

Let us check that the limit does not depend on the choice of the sequence. If $\{f'_m\}$ is another sequence of continuous functions with compact support converging to f in L^2, then $\{f_m - f'_m\}$ is a sequence of continuous functions with compact support converging to 0 in L^2. By (16.6), $\widehat{f_m} - \widehat{f'_m}$ converges to 0 in L^2, and therefore $\widehat{f'_m}$ has the same limit as $\widehat{f_m}$. Thus $\widehat{f}$ is defined uniquely up to almost everywhere equivalence. By passing to the limit in L^2 on both sides of (16.6), we see that (16.6) holds for $f \in L^2$.

16.4 Exercises

Exercise 16.1 Find the Fourier transform of $\chi_{[a,b]}$ and in particular, find the Fourier transform of $\chi_{[-n,n]}$.

Exercise 16.2 Find a real-valued function $f \in L^1$ such that $\widehat{f} \notin L^1$.

Exercise 16.3 Show that if $f \in L^1$ and f is everywhere strictly positive, then $|\widehat{f}(y)| < \widehat{f}(0)$ for $y \neq 0$.

Exercise 16.4 If f is integrable, real-valued, and all the partial derivatives $f_j = \partial f / \partial x_j$ are integrable, prove that the Fourier transform of f_j is given by $\widehat{f_j}(u) = -iu_j \widehat{f}(u)$.

Exercise 16.5 Let $\mathcal{S}$ be the class of real-valued functions f on $\mathbb{R}$ such that for every $k \geq 0$ and $m \geq 0$, $|x|^m|f^{(k)}(x)| \to 0$ as $|x| \to \infty$, where $f^{(k)}$ is the k^{th} derivative of f when $k \geq 1$ and $f^{(0)} = f$. The collection $\mathcal{S}$ is called the *Schwartz class*. Prove that if $f \in \mathcal{S}$, then $\widehat{f} \in \mathcal{S}$.

Exercise 16.6 The Fourier transform of a finite signed measure μ on $\mathbb{R}^n$ is defined by

$$\widehat{\mu}(u) = \int e^{iu\cdot x}\, \mu(dx).$$

Prove that if μ and ν are two finite signed measures on $\mathbb{R}^n$ (with respect to the completion of $\mathcal{L} \times \mathcal{L}$, where $\mathcal{L}$ is the Lebesgue σ-algebra on $\mathbb{R}$) such that $\widehat{\mu}(u) = \widehat{\nu}(u)$ for all $u \in \mathbb{R}^n$, then $\mu = \nu$.

Exercise 16.7 If f is real-valued and continuously differentiable on $\mathbb{R}$, prove that

$$\Big(\int |f|^2\,dx\Big)^2 \leq 4\Big(\int |xf(x)|^2\,dx\Big)\Big(\int |f'|^2\,dx\Big).$$

Exercise 16.8 Prove *Heisenberg's inequality* (which is very useful in quantum mechanics): there exists $c > 0$ such that if $a, b \in \mathbb{R}$ and f is in L^2, then

$$\Big(\int (x-a)^2|f(x)|^2\,dx\Big)\Big(\int (u-b)^2|\widehat{f}(u)|^2\,du\Big) \geq c\Big(\int |f(x)|^2\,dx\Big)^2.$$

Find the best constant c.

Chapter 17

Riesz representation

In Chapter 4 we constructed measures on $\mathbb{R}$. In this chapter we will discuss how to construct measures on more general topological spaces X.

If X is a topological space, let $\mathcal{B}$ be the Borel σ-algebra and suppose μ is a σ-finite measure on $(X, \mathcal{B})$. Throughout this chapter we will restrict our attention to real-valued functions. If f is continuous on X, let us define

$$L(f) = \int_X f\,d\mu.$$

Clearly L is linear, and if $f \geq 0$, then $L(f) \geq 0$. The main topic of this chapter is to prove a converse, the Riesz representation theorem.

We need more hypotheses on X than just that it is a topological space. For simplicity, throughout this chapter we suppose X is a compact metric space. In fact, with almost no changes in the proof, we could let X be a compact Hausdorff space, and with only relatively minor changes, we could even let X be a locally compact Hausdorff metric space. See Remark 17.1. But here we stick to compact metric spaces.

We let $\mathcal{C}(X)$ be the collection of continuous functions from X to $\mathbb{R}$. Recall that the support of a function f is the closure of $\{x : f(x) \neq 0\}$. We write $\operatorname{supp}(f)$ for the support of f. If G is an open subset of X, we define $\mathcal{F}_G$ by

$$\mathcal{F}_G = \{f \in \mathcal{C}(X) : 0 \le f \le 1, \operatorname{supp}(f) \subset G\}.$$

Observe that if $f \in \mathcal{F}_G$, then $0 \le f \le \chi_G$, but the converse does not hold. For example, if $X = [-2, 2]$, $G = (-1, 1)$, and $f(x) = (1 - x^2)^+$, then $0 \le f \le \chi_G$, but the support of f, which is $[-1, 1]$, is not contained in G.

17.1 Partitions of unity

The reason we take our set X to be a metric space is that if $K \subset G \subset X$, where K is compact and G is open, then there exists $f \in \mathcal{F}_G$ such that f is 1 on K. If we let

$$f(x) = \Big(1 - \frac{d(x, K)}{\delta/2}\Big)^+,$$

where $d(x, K) = \inf\{d(x, y) : y \in K\}$ is the distance from x to K and $\delta = \inf\{d(x, y) : x \in K, y \in G^c\}$, then this f will do the job.

Remark 17.1 If X is a compact Hausdorff space instead of a compact metric one, we can still find such an f, that is, $f \in \mathcal{F}_G$ with $f \ge \chi_K$ when $K \subset G$, K is compact, and G is open. Urysohn's lemma is the result from topology that guarantees such an f exists. (A Hausdorff space X is one where if $x, y \in X$, $x \ne y$, there exist disjoint open sets G_x and G_y with $x \in G_x$ and $y \in G_y$. An example of a compact Hausdorff space that is not a metric space and cannot be made into a metric space is $[0, 1]^{\mathbb{R}}$ with the product topology.) See [1] for details.

We will need the following proposition.

Proposition 17.2 *Suppose K is compact and $K \subset G_1 \cup \cdots \cup G_n$, where the G_i are open sets. There exist $g_i \in \mathcal{F}_{G_i}$ for $i = 1, 2, \ldots, n$ such that $\sum_{i=1}^n g_i(x) = 1$ if $x \in K$.*

The collection $\{g_i\}$ is called a *partition of unity* on K, subordinate to the cover $\{G_i\}$.

Proof. Let $x \in K$. Then x will be in at least one G_i. Single points are always compact, so there exists $h_x \in \mathcal{F}_{G_i}$ such that $h_x(x) = 1$.

Let $N_x = \{y : h_x(y) > 0\}$. Since h_x is continuous, then N_x is open, $x \in N_x$, and $\overline{N_x} \subset G_i$.

The collection $\{N_x\}$ is an open cover for the compact set K, so there exists a finite subcover $\{N_{x_1}, \ldots, N_{x_m}\}$. For each i, let

$$F_i = \cup\{\overline{N_{x_j}} : \overline{N_{x_j}} \subset G_i\}.$$

Each F_i is closed, and since X is compact, F_i is compact. We have $F_i \subset G_i$. Let us choose $f_i \in \mathcal{F}_{G_i}$ such that f_i is 1 on F_i.

Now define

$$\begin{aligned} g_1 &= f_1, \\ g_2 &= (1 - f_1) f_2, \\ &\cdots \\ g_n &= (1 - f_1)(1 - f_2) \cdots (1 - f_{n-1}) f_n. \end{aligned}$$

Clearly $g_i \in \mathcal{F}_{G_i}$. Note $g_1 + g_2 = 1 - (1 - f_1)(1 - f_2)$, and an induction argument shows that

$$g_1 + \cdots + g_n = 1 - (1 - f_1)(1 - f_2) \cdots (1 - f_n).$$

If $x \in K$, then $x \in N_{x_j}$ for some j, so $x \in F_i$ for some i. Then $f_i(x) = 1$, which implies $\sum_{k=1}^n g_k(x) = 1$. □

17.2 The representation theorem

Let L be a linear functional mapping $\mathcal{C}(X)$ to $\mathbb{R}$. Thus $L(f + g) = L(f) + L(g)$ and $L(af) = aL(f)$ if $f, g \in \mathcal{C}(X)$ and $a \in \mathbb{R}$. L is a *positive linear functional* if $L(f) \geq 0$ whenever $f \geq 0$ on X.

Here is the *Riesz representation theorem.* $\mathcal{B}$ is the Borel σ-algebra on X, that is, the smallest σ-algebra that contains all the open subsets of X.

Theorem 17.3 *Let X be a compact metric space and L a positive linear functional on $\mathcal{C}(X)$. Then there exists a measure μ on $(X, \mathcal{B})$ such that*

$$L(f) = \int f(y)\, \mu(dy), \qquad f \in \mathcal{C}(X). \tag{17.1}$$

We often write Lf for $L(f)$. Since X is compact, taking f identically equal to 1 in (17.1) shows that μ is a finite measure.

Proof. If G is open, let

$$\ell(G) = \sup\{Lf : f \in \mathcal{F}_G\}$$

and for $E \subset X$, let

$$\mu^*(E) = \inf\{\ell(G) : E \subset G, G \text{ open}\}.$$

Step 1 of the proof will be to show μ^* is an outer measure. Step 2 is to show that every open set is μ^*-measurable. Step 3 is to apply Theorem 4.6 to obtain a measure μ. Step 4 establishes some regularity of μ and Step 5 shows that (17.1) holds.

Step 1. We show μ^* is an outer measure. The only function in $\mathcal{F}_\emptyset$ is the zero function, so $\ell(\emptyset) = 0$, and therefore $\mu^*(\emptyset) = 0$. Clearly $\mu^*(A) \le \mu^*(B)$ if $A \subset B$.

To show the countable subadditivity of μ^*, first let $G_1, G_2, \ldots$ be open sets. For any open set H we see that $\mu^*(H) = \ell(H)$. Let $G = \cup_i G_i$ and let f be any element of $\mathcal{F}_G$. Let K be the support of f. Then K is compact, $\{G_i\}$ is an open cover for K, and therefore there exists n such that $K \subset \cup_{i=1}^n G_i$. Let $\{g_i\}$ be a partition of unity for K subordinate to $\{G_i\}_{i=1}^n$. Since K is the support of f, we have $f = \sum_{i=1}^n fg_i$. Since $g_i \in \mathcal{F}_{G_i}$ and f is bounded by 1, then $fg_i \in \mathcal{F}_{G_i}$. Therefore

$$Lf = \sum_{i=1}^n L(fg_i) \le \sum_{i=1}^n \mu^*(G_i) \le \sum_{i=1}^\infty \mu^*(G_i).$$

Taking the supremum over $f \in \mathcal{F}_G$,

$$\mu^*(G) = \ell(G) \le \sum_{i=1}^\infty \mu^*(G_i).$$

If $A_1, A_2, \ldots$ are subsets of X, let $\varepsilon > 0$, and choose G_i open such that $\ell(G_i) \le \mu^*(G_i) + \varepsilon 2^{-i}$. Then

$$\mu^*(\cup_{i=1}^\infty A_i) \le \mu^*(\cup_{i=1}^\infty G_i) \le \sum_{i=1}^\infty \mu^*(G_i) \le \sum_{i=1}^\infty \mu^*(A_i) + \varepsilon.$$

Since ε is arbitrary, countable subadditivity is proved, and we conclude that μ^* is an outer measure.

Step 2. We show that every open set is μ^*-measurable. Suppose G is open and $E \subset X$. It suffices to show

$$\mu^*(E) \geq \mu^*(E \cap G) + \mu^*(E \cap G^c), \tag{17.2}$$

since the opposite inequality is true by the countable subadditivity of μ^*.

First suppose E is open. Choose $f \in \mathcal{F}_{E\cap G}$ such that

$$L(f) > \ell(E \cap G) - \varepsilon/2.$$

Let K be the support of f. Since K^c is open, we can choose $g \in \mathcal{F}_{E\cap K^c}$ such that $L(g) > \ell(E \cap K^c) - \varepsilon/2$. Then $f + g \in \mathcal{F}_E$, and

$$\begin{aligned} \ell(E) \geq L(f+g) = Lf + Lg &\geq \ell(E \cap G) + \ell(E \cap K^c) - \varepsilon \\ &= \mu^*(E \cap G) + \mu^*(E \cap K^c) - \varepsilon \\ &\geq \mu^*(E \cap G) + \mu^*(E \cap G^c) - \varepsilon. \end{aligned}$$

Since ε is arbitrary, (17.2) holds when E is open.

If $E \subset X$ is not necessarily open, let $\varepsilon > 0$ and choose H open such that $E \subset H$ and $\ell(H) \leq \mu^*(E) + \varepsilon$. Then

$$\begin{aligned} \mu^*(E) + \varepsilon \geq \ell(H) = \mu^*(H) &\geq \mu^*(H \cap G) + \mu^*(H \cap G^c) \\ &\geq \mu^*(E \cap G) + \mu^*(E \cap G^c). \end{aligned}$$

Since ε is arbitrary, (17.2) holds.

Step 3. Let $\mathcal{B}$ be the Borel σ-algebra on X. By Theorem 4.6, the restriction of μ^* to $\mathcal{B}$, which we call μ, is a measure on $\mathcal{B}$. In particular, if G is open, $\mu(G) = \mu^*(G) = \ell(G)$.

Step 4. In this step we show that if K is compact, $f \in \mathcal{C}(X)$, and $f \geq \chi_K$, then $L(f) \geq \mu(K)$. Let $\varepsilon > 0$ and define

$$G = \{x : f(x) > 1 - \varepsilon\},$$

which is open. If $g \in \mathcal{F}_G$, then $g \leq \chi_G \leq f/(1-\varepsilon)$, so $(1-\varepsilon)^{-1} f - g \geq 0$. Because L is a positive linear functional, $L((1-\varepsilon)^{-1} f - g) \geq 0$, which leads to $Lg \leq Lf/(1-\varepsilon)$. This is true for all $g \in \mathcal{F}_G$, hence

$$\mu(K) \leq \mu(G) \leq \frac{Lf}{1-\varepsilon}.$$

Since ε is arbitrary, $\mu(K) \leq Lf$.

Step 5. We now establish (17.1). By writing $f = f^+ - f^-$ and using the linearity of L, to show (17.1) for continuous functions we may suppose $f \ge 0$. Since X is compact, then f is bounded, and multiplying by a constant and using linearity, we may suppose $0 \le f \le 1$.

Let $n \ge 1$ and let $K_i = \{x : f(x) \ge i/n\}$. Since f is continuous, each K_i is a closed set, hence compact. K_0 is all of X. Define

$$f_i(x) = \begin{cases} 0, & x \in K_{i-1}^c; \\ f(x) - \frac{i-1}{n}, & x \in K_{i-1} - K_i; \\ \frac{1}{n}, & x \in K_i. \end{cases}$$

Note $f = \sum_{i=1}^n f_i$ and $\chi_{K_i} \le n f_i \le \chi_{K_{i-1}}$. Therefore

$$\frac{\mu(K_i)}{n} \le \int f_i \, d\mu \le \frac{\mu(K_{i-1})}{n},$$

and so

$$\frac{1}{n} \sum_{i=1}^n \mu(K_i) \le \int f \, d\mu \le \frac{1}{n} \sum_{i=0}^{n-1} \mu(K_i). \tag{17.3}$$

Let $\varepsilon > 0$ and let G be an open set containing K_{i-1} such that $\mu(G) < \mu(K_{i-1}) + \varepsilon$. Then $n f_i \in \mathcal{F}_G$, so

$$L(n f_i) \le \mu(G) \le \mu(K_{i-1}) + \varepsilon.$$

Since ε is arbitrary, $L(f_i) \le \mu(K_{i-1})/n$. By Step 4, $L(n f_i) \ge \mu(K_i)$, and hence

$$\frac{1}{n} \sum_{i=1}^n \mu(K_i) \le L(f) \le \frac{1}{n} \sum_{i=0}^{n-1} \mu(K_i). \tag{17.4}$$

Comparing (17.3) and (17.4) we see that

$$\Big| L(f) - \int f \, d\mu \Big| \le \frac{\mu(K_0) - \mu(K_n)}{n} \le \frac{\mu(X)}{n}.$$

Since, as we saw above, $\mu(X) = L(1) < \infty$ and n is arbitrary, then (17.1) is established. □

Example 17.4 If f is continuous on $[a, b]$, let $L(f)$ be the Riemann integral of f on the interval $[a, b]$. Then L is a positive linear functional on $\mathcal{C}([a, b])$. In this case, the measure whose existence is given by the Riesz representation theorem is Lebesgue measure.

Remark 17.5 Let X be a metric space, not necessarily compact. A continuous function f *vanishes at infinity* if given $\varepsilon > 0$ there exists a compact set K such that $|f(x)| < \varepsilon$ if $x \notin K$. $\mathcal{C}_0(X)$ is the usual notation for the set of continuous functions vanishing at infinity. There is a version of the Riesz representation theorem for $\mathcal{C}_0(X)$. See [1] for details.

17.3 Regularity

We establish the following regularity property of measures on compact metric spaces.

Proposition 17.6 *Suppose X is a compact measure space, $\mathcal{B}$ is the Borel σ-algebra, and μ is a finite measure on the measurable space $(X, \mathcal{B})$. If $E \in \mathcal{B}$ and $\varepsilon > 0$, there exists $K \subset E \subset G$ such that K is compact, G is open, $\mu(G - E) < \varepsilon$, and $\mu(E - K) < \varepsilon$. (K and G depend on ε as well as on E.)*

Proof. Let us say that a subset $E \in \mathcal{B}$ is approximable if given $\varepsilon > 0$ there exists $K \subset E \subset G$ with K compact, G open, $\mu(G - E) < \varepsilon$, and $\mu(E - F) < \varepsilon$. Let $\mathcal{H}$ be the collection of approximable subsets. We will show $\mathcal{H}$ contains all the compact sets and $\mathcal{H}$ is a σ-algebra, which will prove that $\mathcal{H} = \mathcal{B}$, and thus establish the proposition.

If K is compact, let $G_n = \{x : d(x, K) < 1/n\}$. Then the G_n are open sets decreasing to K, and if n is large enough, $\mu(G_n - K) < \varepsilon$. Thus every compact set is in $\mathcal{H}$.

If E is in $\mathcal{H}$ and $\varepsilon > 0$, then choose $K \subset E \subset G$ with K compact, G open, $\mu(E - K) < \varepsilon$, and $\mu(G - E) < \varepsilon$. Then $G^c \subset E^c \subset K^c$, G^c is closed, hence compact, K^c is open, $\mu(K^c - E^c) = \mu(E - K) < \varepsilon$, and $\mu(E^c - G^c) = \mu(G - E) < \varepsilon$. Therefore $\mathcal{H}$ is closed under the operation of taking complements.

Suppose $E_1, E_2, \ldots \in \mathcal{H}$. For each i choose K_i compact and G_i open such that $K_i \subset E_i \subset G_i$, $\mu(G_i - E_i) < \varepsilon 2^{-i}$, and $\mu(E_i - K_i) < \varepsilon 2^{-(i+1)}$. Then $\cup_{i=1}^\infty G_i$ is open, contains $\cup_{i=1}^\infty E_i$, and

$$\mu(\cup_i G_i - \cup_i E_i) \le \sum_{i=1}^\infty \mu(G_i - E_i) < \varepsilon.$$

We see that $\cup_{i=1}^\infty K_i$ is contained in $\cup_{i=1}^\infty E_i$ and similarly,

$$\mu(\cup_{i=1}^\infty E_i - \cup_{i=1}^\infty K_i) \le \sum_{i=1}^\infty \mu(E_i - K_i) < \varepsilon/2.$$

Since $\cup_{i=1}^n K_i$ increases to $\cup_{i=1}^\infty K_i$, we can choose n large so that

$$\mu(\cup_{i=n+1}^\infty K_i) < \varepsilon/2.$$

Then $\cup_{i=1}^n K_i$, being the finite union of compact sets, is compact, is contained in $\cup_{i=1}^\infty E_i$, and

$$\mu(\cup_{i=1}^\infty E_i - \cup_{i=1}^n K_i) < \varepsilon.$$

This proves that $\cup_i E_i$ is in $\mathcal{H}$.

Since $\mathcal{H}$ is closed under the operations of taking complements and countable unions and $\cap_i E_i = (\cup_i E_i^c)^c$, then $\mathcal{H}$ is also closed under the operation of taking countable intersections. Therefore $\mathcal{H}$ is a σ-algebra. □

A measure is called *regular* if

$$\mu(E) = \inf\{\mu(G) : G \text{ open}, E \subset G\}$$

and

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

for all measurable E. An immediate consequence of what we just proved is that finite measures on $(X, \mathcal{B})$ are regular when X is a compact metric space.

17.4 Exercises

Exercise 17.1 Suppose F is a closed subset of $[0,1]$ and we define

$$L(f) = \int_0^1 f\chi_F \, dx$$

for real-valued continuous functions f on $[0,1]$. Prove that if μ is the measure whose existence is given by the Riesz representation theorem, then $\mu(A) = m(A \cap F)$, where m is Lebesgue measure.

Exercise 17.2 Suppose X is a compact metric space and μ is a finite regular measure on $(X, \mathcal{B})$, where $\mathcal{B}$ is the Borel σ-algebra. Prove that if f is a real-valued measurable function and $\varepsilon > 0$, there exists a closed set F such that $\mu(F^c) < \varepsilon$ and the restriction of f to F is a continuous function on F.

Exercise 17.3 Let $C^1([0,1])$ be the set of functions whose derivative exists and is continuous on $[0,1]$. Suppose L is a linear functional on $C^1([0,1])$ such that

$$|L(f)| \le c_1 \|f'\| + c_2 \|f\|$$

for all $f \in C^1([0,1])$, where c_1 and c_2 are positive constants and the norm is the supremum norm. Show there exists a signed measure μ on the Borel subsets of $[0,1]$ and a constant K such that

$$L(f) = \int f' \, d\mu + K f(0), \qquad f \in C^1([0,1]).$$

Exercise 17.4 Suppose X and Y are compact metric spaces and $F : X \to Y$ is a continuous map from X onto Y. If ν is a finite measure on the Borel sets of Y, prove that there exists a measure μ on the Borel sets of X such that

$$\int_Y f \, d\nu = \int_X f \circ F \, d\mu$$

for all f that are continuous on Y.

Exercise 17.5 Let X be a compact metric space. Prove that $C(X)$ has a countable dense subset.

Exercise 17.6 Let X be a compact metric space and let $\mathcal{B}$ be the Borel σ-algebra on X. Let μ_n be a sequence of finite measures on $(X, \mathcal{B})$ and let μ be another finite measure on $(X, \mathcal{B})$. Suppose $\mu_n(X) \to \mu(X)$. Prove that the following are equivalent:
(1) $\int f \, d\mu_n \to \int f \, d\mu$ whenever f is a continuous real-valued function on X;
(2) $\limsup_{n\to\infty} \mu_n(F) \le \mu(F)$ for all closed subsets F of X;
(3) $\liminf_{n\to\infty} \mu_n(G) \ge \mu(G)$ for all open subsets G of X;
(4) $\lim_{n\to\infty} \mu_n(A) = \mu(A)$ whenever A is a Borel subset of X such that $\mu(\partial A) = 0$, where $\partial A = \overline{A} - A^o$ is the boundary of A.

Exercise 17.7 Let X be a compact metric space and let $\mathcal{B}$ be the Borel σ-algebra on X. Let μ_n be a sequence of finite measures on $(X, \mathcal{B})$ and suppose $\sup_n \mu_n(X) < \infty$.
(1) Prove that if $f \in C(X)$, there is a subsequence $\{n_j\}$ such that $\int f \, d\mu_{n_j}$ converges.
(2) Let A be a countable dense subset of $C(X)$. Prove that there is a subsequence $\{n_j\}$ such that $\int f \, d\mu_{n_j}$ converges for all $f \in A$.
(3) With $\{n_j\}$ as in (2), prove that $\int f \, d\mu_{n_j}$ converges for all $f \in C(X)$.
(4) Let $L(f) = \lim_{n_j \to \infty} \int f \, d\mu_{n_j}$. Prove that $L(f)$ is a positive linear functional on $C(X)$. Conclude that there exists a measure μ such that

$$\int f \, d\mu_{n_j} \to \int f \, d\mu$$

for all $f \in C(X)$.

Chapter 18

Banach spaces

Banach spaces are normed linear spaces that are complete. We will give the definitions, discuss the existence of bounded linear functionals, prove the Baire category theorem, and derive some consequences such as the uniform boundedness theorem and the open mapping theorem.

18.1 Definitions

The definition of normed linear space X over a field of scalars F, where F is either the real numbers or the complex numbers, was given in Chapter 1. Recall that a normed linear space is a metric space if we use the metric $d(x, y) = \|x - y\|$.

Definition 18.1 We define a *Banach space* to be a normed linear space that is complete, that is, where every Cauchy sequence converges.

A *linear map* is a map L from a normed linear space X to a normed linear space Y satisfying $L(x + y) = L(x) + L(y)$ for all $x, y \in X$ and $L(\alpha x) = \alpha L(x)$ for all $x \in X$ and $\alpha \in F$. We will sometimes write Lx for $L(x)$. Since $L(0) = L(0+0) = L(0)+L(0)$, then $L(0) = 0$.

Definition 18.2 A linear map f from X to $\mathbb{R}$ is a *real linear functional*, a linear map from X to $\mathbb{C}$ a *complex linear functional*. f is a *bounded linear functional* if

$$\|f\| = \sup\{|f(x)| : x \in X, \|x\| \leq 1\} < \infty.$$

Proposition 18.3 *The following are equivalent.*
(1) The linear functional f is bounded.
(2) The linear functional f is continuous.
(3) The linear functional f is continuous at 0.

Proof. $|f(x) - f(y)| = |f(x-y)| \leq \|f\| \, |x-y|$, so (1) implies (2). That (2) implies (3) is obvious. To show (3) implies (1), if f is not bounded, there exists a sequence $x_n \in X$ such that $\|x_n\| = 1$ for each n, but $|f(x_n)| \to \infty$. If we let $y_n = x_n/|f(x_n)|$, then $y_n \to 0$ but $|f(x_n)| = 1 \not\to 0$, contradicting (3). □

18.2 The Hahn-Banach theorem

We want to prove that there are plenty of linear functionals, but first we need Zorn's lemma, which is equivalent to the axiom of choice.

If we have a set Y with a partial order "$\leq$," a linear ordered subset $X \subset Y$ is one such that if $x, y \in X$, then either $x \leq y$ or $y \leq x$ (or both) holds. A linearly ordered subset $X \subset Y$ has an upper bound if there exists an element z of Y (but it is not necessary that $z \in X$) such that $x \leq z$ for all $x \in X$. An element z of Y is maximal if $z \leq y$ for $y \in Y$ implies $y = z$.

Here is Zorn's lemma.

Lemma 18.4 *If Y is a partially ordered set and every linearly ordered subset of Y has an upper bound, then Y has a maximal element.*

Now we give the *Hahn-Banach theorem* for real linear functionals.

Theorem 18.5 *If M is a subspace of a normed linear space X and f is a bounded real linear functional on M, then f can be extended to a bounded linear functional F on X such that $\|F\| = \|f\|$.*

Saying that F is an extension of f means that the domain of F contains the domain of f and $F(x) = f(x)$ if x is in the domain of f.

Proof. If $\|f\| = 0$, then we take F to be identically 0, so we may assume that $\|f\| \neq 0$, and then by multiplying by a constant, that $\|f\| = 1$. We first show that we can extend f by at least one dimension.

Choose $x_0 \in X - M$ and let M_1 be the vector space spanned by M and x_0. Thus M_1 consists of all vectors of the form $x + \lambda x_0$, where $x \in X$ and λ is real.

We have for all $x, y \in M$

$$f(x) - f(y) = f(x - y) \leq \|x - y\| \leq \|x - x_0\| + \|y - x_0\|.$$

Hence

$$f(x) - \|x - x_0\| \leq f(y) + \|y - x_0\|$$

for all $x, y \in M$. Choose $\alpha \in \mathbb{R}$ such that

$$f(x) - \|x - x_0\| \leq \alpha \leq f(y) + \|y - x_0\|$$

for all $x, y \in M$. Define $f_1(x + \lambda x_0) = f(x) + \lambda\alpha$. This is clearly an extension of f to M_1.

We need to verify that the norm of f_1 is less than or equal to 1. Let $x \in M$ and $\lambda \in \mathbb{R}$. By our choice of α, $f(x) - \|x - x_0\| \leq \alpha$, or $f(x) - \alpha \leq \|x - x_0\|$, and $\alpha \leq f(x) + \|x - x_0\|$, or $f(x) - \alpha \geq -\|x - x_0\|$. Thus

$$|f(x) - \alpha| \leq \|x - x_0\|.$$

Replacing x by $-x/\lambda$ and multiplying by $|\lambda|$, we get

$$|\lambda| \, | - f(x)/\lambda - \alpha| \leq |\lambda| \, \| - x/\lambda - x_0\|,$$

or

$$|f_1(x + \lambda x_0)| = |f(x) + \lambda\alpha| \leq \|x + \lambda x_0\|,$$

which is what we wanted to prove.

We now establish the existence of an extension of f to all of X. Let $\mathcal{F}$ be the collection of all linear extensions F of f satisfying

$\|F\| \leq 1$. This collection is partially ordered by inclusion. That is, if f_1 is an extension of f to a subspace M_1 and f_2 is an extension of f to a subspace M_2, we say $f_1 \leq f_2$ if $M_1 \subset M_2$. Since the union of any increasing family of subspaces of X is again a subspace, then the union of a linearly ordered subfamily of $\mathcal{F}$ lies in $\mathcal{F}$. By Zorn's lemma, $\mathcal{F}$ has a maximal element, say, F_1. By the construction of the preceding two paragraphs, if the domain of F_1 is not all of X, we can find an extension, which would be a contradiction to F_1 being maximal. Therefore F_1 is the desired extension. □

To get a version for complex valued linear functionals is quite easy. Note that if $f(x) = u(x) + iv(x)$, then the real part of f, namely, $u = \operatorname{Re} f$, is a real valued linear functional. Also, $u(ix) = \operatorname{Re} f(ix) = \operatorname{Re} i f(x) = -v(x)$, so that $v(x) = -u(ix)$, and hence $f(x) = u(x) - iu(ix)$.

Theorem 18.6 *If M is a subspace of a normed linear space X and f is a bounded complex linear functional on M, then f can be extended to a bounded linear functional F on X such that $F = f$.*

Proof. Assume without loss of generality that $\|f\| = 1$. Let $u = \operatorname{Re} f$. Note $|u(x)| \leq |f(x)| \leq \|x\|$. Now use the version of the Hahn-Banach theorem for real linear functionals to find a linear functional U that is an extension of u to X such that $\|U\| \leq 1$. Let $F(x) = U(x) - iU(ix)$.

It only remains to show that the norm of F is at most 1. Fix x, and write $F(x) = re^{i\theta}$. Then

$$|F(x)| = r = e^{-i\theta} F(x) = F(e^{-i\theta} x).$$

Since this quantity is real and non-negative,

$$|F(x)| = U(e^{i\theta} x) \leq \|U\| \, \|e^{-i\theta} x\| \leq \|x\|.$$

This holds for all x, so $\|F\| \leq 1$. □

As an application of the Hahn-Banach theorem, given a subspace M and an element x_0 not in M, we can define $f(x + \lambda x_0) = \lambda x_0$ for $x \in M$, and then extend this linear functional to all of X. Then f will be 0 on M but nonzero at x_0.

Another application is to fix $x_0 \neq 0$, let $f(\lambda x_0) = \lambda \|x_0\|$, and then extend f to all of X. Thus there exists a linear functional f such that $f(x_0) = \|x_0\|$ and $\|f\| = 1$.

18.3 Baire's theorem and consequences

We turn now to the Baire category theorem and some of its consequences. Recall that if A is a set, we use $\overline{A}$ for the closure of A and A^o for the interior of A. A set A is dense in X if $\overline{A} = X$ and A is nowhere dense if $(\overline{A})^o = \emptyset$.

The *Baire category theorem* is the following. Completeness of the metric space is crucial to the proof.

Theorem 18.7 *Let X be a complete metric space.*
(1) If G_n are open sets dense in X, then $\cap_n G_n$ is dense in X.
(2) X cannot be written as the countable union of nowhere dense sets.

Proof. We first show that (1) implies (2). Suppose we can write X as a countable union of nowhere dense sets, that is, $X = \cup_n E_n$ where $(\overline{E_n})^o = \emptyset$. We let $F_n = \overline{E_n}$, which is a closed set, and then $F_n^o = \emptyset$ and $X = \cup_n F_n$. Let $G_n = F_n^c$, which is open. Since $F_n^o = \emptyset$, then $\overline{G_n} = X$. Starting with $X = \cup_n F_n$ and taking complements, we see that $\emptyset = \cap_n G_n$, a contradiction to (1).

We must prove (1). Suppose $G_1, G_2, \ldots$ are open and dense in X. Let H be any nonempty open set in X. We need to show there exists a point in $H \cap (\cap_n G_n)$. We will construct a certain Cauchy sequence $\{x_n\}$ and the limit point, x, will be the point we seek.

Let $B(z, r) = \{y \in X : d(z, y) < r\}$, where d is the metric. Since G_1 is dense in X, $H \cap G_1$ is nonempty and open, and we can find x_1 and r_1 such that $\overline{B(x_1, r_1)} \subset H \cap G_1$ and $0 < r_1 < 1$. Suppose we have chosen x_{n-1} and r_{n-1} for some $n \geq 2$. Since G_n is dense, then $G_n \cap B(x_{n-1}, r_{n-1})$ is open and nonempty, so there exists x_n and r_n such that $\overline{B(x_n, r_n)} \subset G_n \cap B(x_{n-1}, r_{n-1})$ and $0 < r_n < 2^{-n}$. We continue and get a sequence x_n in X. If $m, n > N$, then x_m and x_m both lie on $B(x_N, r_N)$, and so $d(x_m, x_n) < 2r_N < 2^{-N+1}$. Therefore x_n is a Cauchy sequence, and since X is complete, x_n converges to a point $x \in X$.

It remains to show that $x \in H \cap (\cap_n G_n)$. Since x_n lies in $\overline{B(x_N, r_N)}$ if $n > N$, then x lies in each $\overline{B(x_N, r_N)}$, and hence in each G_N. Therefore $x \in \cap_n G_n$. Also,

$$x \in \overline{B(x_n, r_n)} \subset B(x_{n-1}, r_{n-1}) \subset \cdots \subset B(x_1, r_1) \subset H.$$

Thus we have found a point x in $H \cap (\cap_n G_n)$. $\square$

A set $A \subset X$ is called *meager* or of the *first category* if it is the countable union of nowhere dense sets; otherwise it is of the *second category.*

An important application of the Baire category theorem is the *Banach-Steinhaus theorem*, also called the *uniform boundedness theorem.*

Theorem 18.8 *Suppose X is a Banach space and Y is a normed linear space. Let A be an index set and let $\{L_\alpha : \alpha \in A\}$ be a collection of bounded linear maps from X into Y. Then either there exists a $M < \infty$ such that $\|L_\alpha\| \le M$ for all $\alpha \in A$ or else $\sup_\alpha \|L_\alpha x\| = \infty$ for some x.*

Proof. Let $\ell(x) = \sup_{\alpha \in A} \|L_\alpha x\|$. Let $G_n = \{x : \ell(x) > n\}$. We argue that G_n is open. The map $x \to \|L_\alpha x\|$ is a continuous function for each α since L_α is a bounded linear functional. This implies that for each α, the set $\{x : \|L_\alpha x\| > n\}$ is open. Since $x \in G_n$ if and only if for some $\alpha \in A$ we have $\|L_\alpha x\| > n$, we conclude G_n is the union of open sets, hence is open.

Suppose there exists N such that G_N is not dense in X. Then there exists x_0 and r such that $\overline{B(x_0, r)} \cap G_N = \emptyset$. This can be rephrased as saying that if $\|x - x_0\| \le r$, then $\|L_\alpha(x)\| \le N$ for all $\alpha \in A$. If $\|y\| \le r$, we have $y = (x_0 + y) - x_0$. Then $\|(x_0 + y) - x_0\| = \|y\| \le r$, and hence $\|L_\alpha(x_0 + y)\| \le N$ for all α. Also, of course, $\|x_0 - x_0\| = 0 \le r$, and thus $\|L_\alpha(x_0)\| \le N$ for all α. We conclude that if $\|y\| \le r$ and $\alpha \in A$,

$$\|L_\alpha y\| = \|L_\alpha((x_0 + y) - x_0)\| \le \|L_\alpha(x_0 + y)\| + \|L_\alpha x_0\| \le 2N.$$

Consequently, $\sup_\alpha \|L_\alpha\| \le M$ with $M = 2N/r$.

The other possibility, by the Baire category theorem, is that every G_n is dense in X, and in this case $\cap_n G_n$ is dense in X. But $\ell(x) = \infty$ for every $x \in \cap_n G_n$. $\square$

The following theorem is called the *open mapping theorem.* It is important that L be onto. A mapping $L : X \to Y$ is *open* if $L(U)$ is open in Y whenever U is open in X. For a measurable set A, we let $L(A) = \{Lx : x \in A\}$.

Theorem 18.9 *Let X and Y be Banach spaces. A bounded linear map L from X onto Y is open.*

Proof. We need to show that if $B(x,r)\subset X$, then $L(B(x,r))$ contains a ball in Y. We will show $L(B(0,r))$ contains a ball centered at 0 in Y. Then using the linearity of L, $L(B(x,r))$ will contain a ball centered at Lx in Y. By linearity, to show that $L(B(0,r))$ contains a ball centered at 0, it suffices to show that $L(B(0,1))$ contains a ball centered at 0 in Y.

Step 1. We show that there exists r such that $B(0,r2^{-n})\subset \overline{L(B(0,2^{-n}))}$ for each n. Since L is onto, $Y=\cup_{n=1}^\infty L(B(0,n))$. The Baire category theorem tells us that at least one of the sets $L(B(0,n))$ cannot be nowhere dense. Since L is linear, $L(B(0,1))$ cannot be nowhere dense. Thus there exist y_0 and r such that $B(y_0,4r)\subset\overline{L(B(0,1))}$.

Pick $y_1\in L(B(0,1))$ such that $\|y_1-y_0\|<2r$ and let $z_1\in B(0,1)$ be such that $y_1=Lz_1$. Then $B(y_1,2r)\subset B(y_0,4r)\subset\overline{L(B(0,1))}$. Thus if $\|y\|<2r$, then $y+y_1\in B(y_1,2r)$, and so

$$y=-Lz_1+(y+y_1)\in\overline{L(-z_1+B(0,1))}.$$

Since $z_1\in B(0,1)$, then $-z_1+B(0,1)\subset B(0,2)$, hence

$$y\in\overline{L(-z_1+B(0,1))}\subset\overline{L(B(0,2))}.$$

By the linearity of L, if $\|y\|<r$, then $y\in\overline{L(B(0,1))}$. It follows by linearity that if $\|y\|<r2^{-n}$, then $y\in\overline{L(B(0,2^{-n}))}$. This can be rephrased as saying that if $\|y\|<r2^{-n}$ and $\varepsilon>0$, then there exists x such that $\|x\|<2^{-n}$ and $\|y-Lx\|<\varepsilon$.

Step 2. Suppose $\|y\|<r/2$. We will construct a sequence $\{x_j\}$ by induction such that $y=L(\sum_{j=1}^\infty x_j)$. By Step 1 with $\varepsilon=r/4$, we can find $x_1\in B(0,1/2)$ such that $\|y-Lx_1\|<r/4$. Suppose we have chosen $x_1,\ldots,x_{n-1}$ such that

$$\Big\|y-\sum_{j=1}^{n-1}Lx_j\Big\|<r2^{-n}.$$

Let $\varepsilon=r2^{-(n+1)}$. By Step 1, we can find x_n such that $\|x_n\|<2^{-n}$ and

$$\Big\|y-\sum_{j=1}^{n}Lx_j\Big\|=\Big\|\Big(y-\sum_{j=1}^{n-1}Lx_j\Big)-Lx_n\Big\|<r2^{-(n+1)}.$$

We continue by induction to construct the sequence $\{x_j\}$. Let $w_n = \sum_{j=1}^n x_j$. Since $\|x_j\| < 2^{-j}$, then w_n is a Cauchy sequence. Since X is complete, w_n converges, say, to x. But then $\|x\| < \sum_{j=1}^\infty 2^{-j} = 1$, and since L is continuous, $y = Lx$. That is, if $y \in B(0, r/2)$, then $y \in L(B(0,1))$. $\square$

18.4 Exercises

Exercise 18.1 Find a measure space $(X, \mathcal{A}, \mu)$, a subspace Y of $L^1(\mu)$, and a bounded linear functional f on Y with norm 1 such that f has two distinct extensions to $L^1(\mu)$ and each of the extensions has norm equal to 1.

Exercise 18.2 Show that $L^p([0,1])$ is *separable*, that is, there is a countable dense subset, if $1 \le p < \infty$. Show that $L^\infty([0,1])$ is not separable.

Exercise 18.3 For $k \ge 1$ and functions $f : [0,1] \to \mathbb{R}$ that are k times differentiable, define

$$\|f\|_{C^k} = \|f\|_\infty + \|f'\|_\infty + \cdots + \|f^{(k)}\|_\infty,$$

where $f^{(k)}$ is the k^{th} derivative of f. Let $C^k([0,1])$ be the collection of k times differentiable functions f with $\|f\|_{C^k} < \infty$. Is $C^k(0,1])$ complete with respect to the norm $\|\cdot\|_{C^k}$?

Exercise 18.4 Let $\alpha \in (0,1)$. For f a real-valued continuous function on $[0,1]$ define

$$\|f\|_{C^\alpha} = \sup_{x\in[0,1]} |f(x)| + \sup_{x,y\in[0,1], x\ne y} \frac{|f(x) - f(y)|}{|x-y|^\alpha}.$$

Let $C^\alpha([0,1])$ be the set of functions f with $\|f\|_{C^\alpha} < \infty$. Is $C^\alpha([0,1])$ complete with respect to the norm $\|\cdot\|_{C^\alpha}$?

Exercise 18.5 For positive integers n let

$$A_n = \Big\{ f \in L^1([0,1]) : \int_0^1 |f(x)|^2\, dx \le n \Big\}.$$

Show that each A_n is a closed subset of $L^1([0,1])$ with empty interior.

Exercise 18.6 Suppose L is a linear functional on a normed linear space X. Prove that L is a bounded linear functional if and only if the set $\{x \in X : L(x) = 0\}$ is closed.

Exercise 18.7 A set A in a normed linear space is *convex* if

$$\lambda x + (1 - \lambda)y \in A$$

whenever $x, y \in A$ and $\lambda \in [0, 1]$.
(1) Prove that if A is convex, then the closure of A is convex.
(2) Prove that the open unit ball in a normed linear space is convex. (The open unit ball is the set of x such that $\|x\| < 1$.)

Exercise 18.8 The unit ball in a normed linear space X is *strictly convex* if $\|\lambda x + (1 - \lambda)y\| < 1$ whenever $\|f\| = \|g\| = 1$, $f \neq g$, and $\lambda \in (0, 1)$.
(1) Let $(X, \mathcal{A}, \mu)$ be a measure space. Prove that the unit ball in $L^p(\mu)$ is strictly convex.
(2) Prove that the unit balls in $L^1(\mu)$, $L^\infty(\mu)$, and $C(X)$ are not strictly convex provided X consists of more than one point.

Exercise 18.9 Let f_n be a sequence of continuous functions on $\mathbb{R}$ that converge at every point. Prove there exist an interval and a number M such that $\sup_n |f_n|$ is bounded by M on that interval.

Exercise 18.10 Suppose $\|\cdot\|_1$ and $\|\cdot\|_2$ are two norms such that $\|x\|_1 \le \|x\|_2$ for all x in a vector space X, and suppose X is complete with respect to both norms. Prove that there exists a positive constant c such that

$$\|x\|_2 \le c\|x\|_1$$

for all $x \in X$.

Exercise 18.11 Suppose X and Y are Banach spaces.
(1) Let $X \times Y$ be the set of ordered pairs (x, y) with

$$(x_1 + x_2, y_1 + y_2) = (x_1, y_1) + (x_2, y_2)$$

for each $x_1, x_2 \in X$ and $y_1, y_2 \in Y$ and $c(x, y) = (cx, cy)$ if $x \in \mathbb{R}$. Define $\|(x, y)\| = \|x\| + \|y\|$. Prove that $X \times Y$ is a Banach space.
(2) Let L be a linear map from X into Y such that if $x_n \to x$ in

X and $Lx_n \to y$ in Y, then $y = Lx$. Such a map is called a *closed map.* Let G be the *graph* of L, defined by $G = \{(x, y) : y = Lx\}$. Prove that G is a closed subset of $X \times Y$, hence is complete.
(3) Prove that the function $(x, Lx) \to x$ is continuous, one-one, linear, and maps G onto X.
(4) Prove the *closed graph theorem*, which says that if L is a linear map from one Banach space to another that is a closed map, then L is a continuous map.

Exercise 18.12 Let $X = C^1([0, 1])$ (defined in Exercise 18.3) and $Y = C([0, 1])$. Define $D : X \to Y$ by $Df = f'$. Show that D is a closed map but not a bounded one.

Exercise 18.13 Let A be the set of real-valued continuous functions on $[0, 1]$ such that

$$\int_0^{1/2} f(x)\,dx - \int_{1/2}^1 f(x)\,dx = 1.$$

Prove that A is a closed convex subset of $C([0, 1])$, but there does not exist $f \in A$ such that

$$\|f\| = \inf_{g \in A} \|g\|.$$

Exercise 18.14 Let A_n be the subset of the real-valued continuous functions on $[0, 1]$ given by

$$A_n = \{f : \text{there exists } x \in [0,1] \text{ such that } |f(x) - f(y)| \le n|x - y| \text{ for all } y \in [0,1]\}.$$

(1) Prove that A_n is nowhere dense in $C([0, 1])$.
(2) Prove that there exist functions f in $C([0, 1])$ which are *nowhere differentiable* on $[0, 1]$, that is, $f'(x)$ does not exist at any point of $[0, 1]$.

Chapter 19

Hilbert spaces

Hilbert spaces are complete normed linear spaces that have an inner product. This added structure allows one to talk about orthonormal sets. We will give the definitions and basic properties. As an application we briefly discuss Fourier series.

19.1 Inner products

Recall that if a is a complex number, then $\overline{a}$ represents the complex conjugate. When a is real, $\overline{a}$ is just a itself.

Definition 19.1 Let H be a vector space where the set of scalars F is either the real numbers or the complex numbers. H is an *inner product space* if there is a map $\langle \cdot, \cdot \rangle$ from $H \times H$ to F such that
(1) $\langle y, x \rangle = \overline{\langle x, y \rangle}$ for all $x, y \in H$;
(2) $\langle x + y, z \rangle = \langle x, z \rangle + \langle y, z \rangle$ for all $x, y, z \in H$;
(3) $\langle \alpha x, y \rangle = \alpha \langle x, y \rangle$, for $x, y \in H$ and $\alpha \in F$;
(4) $\langle x, x \rangle \geq 0$ for all $x \in H$;
(5) $\langle x, x \rangle = 0$ if and only if $x = 0$.

We define $\|x\| = \langle x, x \rangle^{1/2}$, so that $\langle x, x \rangle = \|x\|^2$. From the definitions it follows easily that $\langle 0, y \rangle = 0$ and $\langle x, \alpha y \rangle = \overline{\alpha} \langle x, y \rangle$.

The following is the *Cauchy-Schwarz inequality*. The proof is the same as the one usually taught in undergraduate linear algebra classes, except for some complications due to the fact that we allow the set of scalars to be the complex numbers.

Theorem 19.2 *For all $x, y \in H$, we have*

$$|\langle x, y\rangle| \le \|x\|\,\|y\|.$$

Proof. Let $A = \|x\|^2$, $B = |\langle x, y\rangle|$, and $C = \|y\|^2$. If $C = 0$, then $y = 0$, hence $\langle x, y\rangle = 0$, and the inequality holds. If $B = 0$, the inequality is obvious. Therefore we will suppose that $C > 0$ and $B \neq 0$.

If $\langle x, y\rangle = Re^{i\theta}$, let $\alpha = e^{i\theta}$, and then $|\alpha| = 1$ and $\alpha\langle y, x\rangle = |\langle x, y\rangle| = B$. Since B is real, we have that $\overline{\alpha}\langle x, y\rangle$ also equals $|\langle x, y\rangle|$.

We have for real r

$$\begin{aligned} 0 &\le \|x - r\alpha y\|^2 \\ &= \langle x - r\alpha y, x - r\alpha y\rangle \\ &= \langle x, x\rangle - r\alpha\langle y, x\rangle - r\overline{\alpha}\langle x, y\rangle + r^2\langle y, y\rangle \\ &= \|x\|^2 - 2r|\langle x, y\rangle| + r^2\|y\|^2. \end{aligned}$$

Therefore

$$A - 2Br + Cr^2 \ge 0$$

for all real numbers r. Since we are supposing that $C > 0$, we may take $r = B/C$, and we obtain $B^2 \le AC$. Taking square roots of both sides gives the inequality we wanted. □

From the Cauchy-Schwarz inequality we get the triangle inequality:

Proposition 19.3 *For all $x, y \in H$ we have*

$$\|x + y\| \le \|x\| + \|y\|.$$

Proof. We write

$$\begin{aligned} \|x + y\|^2 &= \langle x + y, x + y\rangle = \langle x, x\rangle + \langle x, y\rangle + \langle y, x\rangle + \langle y, y\rangle \\ &\le \|x\|^2 + 2\|x\|\,\|y\| + \|y\|^2 = (\|x\| + \|y\|)^2, \end{aligned}$$

as desired. □

The triangle inequality implies

$$\|x-z\| \le \|x-y\| + \|y-z\|.$$

Therefore $\|\cdot\|$ is a norm on H, and so if we define the distance between x and y by $\|x-y\|$, we have a metric space.

Definition 19.4 A *Hilbert space* H is an inner product space that is complete with respect to the metric $d(x,y) = \|x-y\|$.

Example 19.5 Let μ be a positive measure on a set X, let $H = L^2(\mu)$, and define

$$\langle f, g\rangle = \int f\overline{g}\, d\mu.$$

This is easily seen to be a Hilbert space. To show the completeness we use Theorem 15.4.

If we let μ be counting measure on the natural numbers, we get what is known as the space ℓ^2. An element of ℓ^2 is a sequence $a = (a_1, a_2, \ldots)$ such that $\sum_{n=1}^\infty |a_n|^2 < \infty$ and if $b = (b_1, b_2, \ldots)$, then

$$\langle a, b\rangle = \sum_{n=1}^\infty a_n\overline{b_n}.$$

We get another common Hilbert space, n-dimensional Euclidean space, by letting μ be counting measure on $\{1, 2, \ldots, n\}$.

Proposition 19.6 *Let $y \in H$ be fixed. Then the functions $x \to \langle x, y\rangle$ and $x \to \|x\|$ are continuous.*

Proof. By the Cauchy-Schwarz inequality,

$$|\langle x, y\rangle - \langle x', y\rangle| = |\langle x - x', y\rangle| \le \|x - x'\|\,\|y\|,$$

which proves that the function $x \to \langle x, y\rangle$ is continuous. By the triangle inequality, $\|x\| \le \|x - x'\| + \|x'\|$, or

$$\|x\| - \|x'\| \le \|x - x'\|.$$

The same holds with x and x' reversed, so

$$|\,\|x\| - \|x'\|\,| \le \|x - x'\|,$$

and thus the function $x \to \|x\|$ is continuous. □

19.2 Subspaces

Definition 19.7 A subset M of a vector space is a *subspace* if M is itself a vector space with respect to the same operations of addition and scalar multiplication. A *closed subspace* is a subspace that is closed relative to the metric given by $\langle \cdot, \cdot \rangle$.

For an example of a subspace that is not closed, consider ℓ^2 and let M be the collection of sequences for which all but finitely many elements are zero. M is clearly a subspace. Let $x_n = (1, \frac{1}{2}, \ldots, \frac{1}{n}, 0, 0, \ldots)$ and $x = (1, \frac{1}{2}, \frac{1}{3}, \ldots)$. Then each $x_n \in M$, $x \notin M$, and we conclude M is not closed because

$$\|x_n - x\|^2 = \sum_{j=n+1}^{\infty} \frac{1}{j^2} \to 0$$

as $n \to \infty$.

A set $E \subset H$ is *convex* if $\lambda x + (1 - \lambda x) \in E$ whenever $0 \le \lambda \le 1$ and $x, y \in E$.

Proposition 19.8 *Each nonempty closed convex subset E of H has a unique element of smallest norm.*

Proof. Let $\delta = \inf\{\|x\| : x \in E\}$. Since $\|x + y\|^2 = \langle x + y, x + y \rangle$ and similarly for $\|x - y\|^2$, $\|x\|^2$, and $\|y\|^2$, a simple calculation shows that

$$\|x + y\|^2 + \|x - y\|^2 = 2\|x\|^2 + \|y\|^2.$$

Dividing by 4, if $x, y \in E$, then

$$\tfrac{1}{4}\|x - y\|^2 = \tfrac{1}{2}\|x\|^2 + \tfrac{1}{2}\|y\|^2 - \left\|\frac{x + y}{2}\right\|^2.$$

Since E is convex, if $x, y \in E$, then $(x + y)/2 \in E$, and we have

$$\|x - y\|^2 \le 2\|x\|^2 + 2\|y\|^2 - 4\delta^2. \qquad (19.1)$$

Choose $y_n \in E$ such that $\|y_n\| \to \delta$. Applying (19.1) with x replaced by y_n and y replaced by y_m, we see that

$$\|y_n - y_m\|^2 \le 2\|y_n\|^2 + 2\|y_m\|^2 - 4\delta^2,$$

and the right hand side tends to 0 as m and n tend to infinity. Hence y_n is a Cauchy sequence, and since H is complete, it converges to some $y \in H$. Since $y_n \in E$ and E is closed, $y \in E$. Since the norm is a continuous function, $\|y\| = \lim \|y_n\| = \delta$.

If y' is another point with $\|y'\| = \delta$, then by (19.1) with x replaced by y' we have $\|y - y'\| = 0$, and hence $y = y'$. □

We say $x \perp y$, or x is *orthogonal* to y, if $\langle x, y\rangle = 0$. Let $x^\perp$, read "x perp," be the set of all y in X that are orthogonal to x. If M is a subspace, let $M^\perp$ be the set of all y that are orthogonal to all points in M. It is clear from the linearity of the inner product that $x^\perp$ is a subspace of H. The subspace $x^\perp$ is closed because it is the same as the set $f^{-1}(\{0\})$, where $f(x) = \langle x, y\rangle$, which is continuous by Proposition 19.6. Also, it is easy to see that $M^\perp$ is a subspace, and since

$$M^\perp = \cap_{x\in M} x^\perp,$$

$M^\perp$ is closed. We make the observation that if $z \in M \cap M^\perp$, then

$$\|z\|^2 = \langle z, z\rangle = 0,$$

so $z = 0$.

The following is sometimes called the *Riesz representation theorem*, although usually that name is reserved for Theorem 17.3. To motivate the theorem, consider the case where H is n-dimensional Euclidean space. Elements of $\mathbb{R}^n$ can be identified with $n \times 1$ matrices and linear maps from $\mathbb{R}^n$ to $\mathbb{R}^m$ can be represented by multiplication on the left by a $m \times n$ matrix A. For bounded linear functionals on H, $m = 1$, so A is $1 \times n$, and the y of the next theorem is the vector associated with the transpose of A.

Theorem 19.9 *If L is a bounded linear functional on H, then there exists a unique $y \in H$ such that $Lx = \langle x, y\rangle$.*

Proof. The uniqueness is easy. If $Lx = \langle x, y\rangle = \langle x, y'\rangle$, then $\langle x, y - y'\rangle = 0$ for all x, and in particular, when $x = y - y'$.

We now prove existence. If $Lx = 0$ for all x, we take $y = 0$. Otherwise, let $M = \{x : Lx = 0\}$, take $z \neq 0$ in $M^\perp$, and let $y = \alpha z$ where $\alpha = \overline{Lz}/\langle z, z\rangle$. Notice $y \in M^\perp$,

$$Ly = \frac{\overline{Lz}}{\langle z, z\rangle} Lz = |Lz|^2/\langle z, z\rangle = \langle y, y\rangle,$$

and $y \neq 0$.

If $x \in H$ and

$$w = x - \frac{Lx}{\langle y, y\rangle} y,$$

then $Lw = 0$, so $w \in M$, and hence $\langle w, y\rangle = 0$. Then

$$\langle x, y\rangle = \langle x - w, y\rangle = Lx$$

as desired. □

19.3 Orthonormal sets

A subset $\{u_\alpha\}_{\alpha\in A}$ of H is *orthonormal* if $\|u_\alpha\| = 1$ for all α and $\langle u_\alpha, u_\beta\rangle = 0$ whenever $\alpha, \beta \in A$ and $\alpha \neq \beta$.

The Gram-Schmidt procedure from linear algebra also works in infinitely many dimensions. Suppose $\{x_n\}_{n=1}^\infty$ is a linearly independent sequence, i.e., no finite linear combination of the x_n is 0. Let $u_1 = x_1/\|x_1\|$ and define inductively

$$v_N = x_N - \sum_{i=1}^{n-1} \langle x_N, u_i\rangle u_i,$$
$$u_N = v_N/\|v_N\|.$$

We have $\langle v_N, u_i\rangle = 0$ if $i < N$, so $u_1, \ldots, u_N$ are orthonormal.

Proposition 19.10 *If $\{u_\alpha\}_{\alpha\in A}$ is an orthonormal set, then for each $x \in H$,*

$$\sum_{\alpha\in A} |\langle x, u_\alpha\rangle|^2 \leq \|x\|^2. \tag{19.2}$$

This is called *Bessel's inequality.* This inequality implies that only finitely many of the summands on the left hand side of (19.2) can be larger than $1/n$ for each n, hence only countably many of the summands can be nonzero.

Proof. Let F be a finite subset of A. Let

$$y = \sum_{\alpha\in F} \langle x, u_\alpha\rangle u_\alpha.$$

Then
$$0 \le \|x-y\|^2 = \|x\|^2 - \langle x,y\rangle - \langle y,x\rangle + \|y\|^2.$$
Now
$$\langle y,x\rangle = \Big\langle \sum_{\alpha\in F} \langle x,u_\alpha\rangle u_\alpha, x\Big\rangle = \sum_{\alpha\in F} \langle x,u_\alpha\rangle\langle u_\alpha,x\rangle = \sum_{\alpha\in F} |\langle x,u_\alpha\rangle|^2.$$
Since this is real, then $\langle x,y\rangle = \langle y,x\rangle$. Also
$$\begin{aligned}\|y\|^2 = \langle y,y\rangle &= \Big\langle \sum_{\alpha\in F} \langle x,u_\alpha\rangle u_\alpha, \sum_{\beta\in F} \langle x,u_\beta\rangle u_\beta\Big\rangle\\ &= \sum_{\alpha,\beta\in F} \langle x,u_\alpha\rangle\overline{\langle x,u_\beta\rangle}\langle u_\alpha,u_\beta\rangle\\ &= \sum_{\alpha\in F} |\langle x,u_\alpha\rangle|^2,\end{aligned}$$
where we used the fact that $\{u_\alpha\}$ is an orthonormal set and so $\langle u_\alpha,u_\beta\rangle$ equals 0 if $\alpha\ne\beta$ and equals 1 if $\alpha=\beta$. Therefore
$$0 \le \|y-x\|^2 = \|x\|^2 - \sum_{\alpha\in F} |\langle x,u_\alpha\rangle|^2.$$
Rearranging,
$$\sum_{\alpha\in F} |\langle x,u_\alpha\rangle|^2 \le \|x\|^2$$
when F is a finite subset of A. If N is an integer larger than $n\|x\|^2$, it is not possible that $|\langle x,u_\alpha\rangle|^2 > 1/n$ for more than N of the α. Hence $|\langle x,u_\alpha\rangle|^2 \ne 0$ for only countably many α. Label those α's as $\alpha_1,\alpha_2,\ldots$. Then
$$\sum_{\alpha\in A} |\langle x,u_\alpha\rangle|^2 = \sum_{j=1}^\infty |\langle x,u_{\alpha_j}\rangle|^2 = \lim_{J\to\infty} \sum_{j=1}^J |\langle x,u_{\alpha_j}\rangle|^2 \le \|x\|^2,$$
which is what we wanted. □

Proposition 19.11 *Suppose $\{u_\alpha\}_{\alpha\in A}$ is orthonormal. Then the following are equivalent.*
(1) If $\langle x,u_\alpha\rangle = 0$ for each $\alpha\in A$, then $x=0$.
(2) $\|x\|^2 = \sum_{\alpha\in A} |\langle x,u_\alpha\rangle|^2$ for all x.
(3) For each $x\in H$, $x = \sum_{\alpha\in A} \langle x,u_\alpha\rangle u_\alpha$.

We make a few remarks. When (1) holds, we say the orthonormal set is *complete*. (2) is called *Parseval's identity*. In (3) the convergence is with respect to the norm of H and implies that only countably many of the terms on the right hand side are nonzero.

Proof. First we show (1) implies (3). Let $x \in H$. By Bessel's inequality and the remarks following the statement of Proposition 19.10 there can be at most countably many α such that $|\langle x, u_\alpha\rangle|^2 \neq 0$. Let $\alpha_1, \alpha_2, \ldots$ be an enumeration of those α. By Bessel's inequality, the series $\sum_i |\langle x, u_{\alpha_i}\rangle|^2$ converges. Using that $\{u_\alpha\}$ is an orthonormal set,

$$\Big\| \sum_{j=m}^{n} \langle x, u_{\alpha_j}\rangle u_{\alpha_j} \Big\|^2 = \sum_{j,k=m}^{n} \langle x, u_{\alpha_j}\rangle \overline{\langle x, u_{\alpha_k}\rangle} \langle u_{\alpha_j}, u_{\alpha_k}\rangle$$
$$= \sum_{j=m}^{n} |\langle x, u_{\alpha_j}\rangle|^2 \to 0$$

as $m, n \to \infty$. Thus $\sum_{j=1}^{n} \langle x, u_{\alpha_j}\rangle u_{\alpha_j}$ is a Cauchy sequence, and hence converges. Let $z = \sum_{j=1}^{\infty} \langle x, u_{\alpha_j}\rangle u_{\alpha_j}$. Then $\langle z - x, u_{\alpha_j}\rangle = 0$ for each α_j. By (1), this implies $z - x = 0$.

We see that (3) implies (2) because

$$\|x\|^2 - \sum_{j=1}^{n} |\langle x, u_{\alpha_j}\rangle|^2 = \Big\| x - \sum_{j=1}^{n} \langle x, u_{\alpha_j}\rangle u_{\alpha_j} \Big\|^2 \to 0.$$

That (2) implies (1) is clear. □

Example 19.12 Take $H = \ell^2 = \{x = (x_1, x_2, \ldots) : \sum |x_i|^2 < \infty\}$ with $\langle x, y\rangle = \sum_i x_i \overline{y_i}$. Then $\{e_i\}$ is a complete orthonormal system, where $e_i = (0, 0, \ldots, 0, 1, 0, \ldots)$, i.e., the only nonzero coordinate of e_i is the i^{th} one.

A collection of elements $\{e_\alpha\}$ is a *basis* for H if the set of finite linear combinations of the e_α is dense in H.

Proposition 19.13 *Every Hilbert space has an orthonormal basis.*

This means that (3) in Proposition 19.11 holds.

Proof. If $B = \{u_\alpha\}$ is orthonormal, but not a basis, let V be the closure of the linear span of B, that is, the closure with respect to the norm in H of the set of finite linear combinations of elements of B. Choose $x \in V^\perp$, and if we let $B' = B \cup \{x/\|x\|\}$, then B' is a basis that is strictly bigger than B.

It is easy to see that the union of an increasing sequence of orthonormal sets is an orthonormal set, and so there is a maximal one by Zorn's lemma. By the preceding paragraph, this maximal orthonormal set must be a basis, for otherwise we could find a larger basis. □

19.4 Fourier series

An interesting application of Hilbert space techniques is to *Fourier series*, or equivalently, to *trigonometric series*. For our Hilbert space we take $H = L^2([0, 2\pi))$ and let

$$u_n = \frac{1}{\sqrt{2\pi}} e^{inx}$$

for n an integer. (n can be negative.) Recall that

$$\langle f, g\rangle = \int_0^{2\pi} f(x)\overline{g(x)}\, dx$$

and $\|f\|^2 = \int_0^{2\pi} |f(x)|^2\, dx$.

It is easy to see that $\{u_n\}$ is an orthonormal set:

$$\int_0^{2\pi} e^{inx} e^{-imx}\, dx = \int_0^{2\pi} e^{i(n-m)x}\, dx = 0$$

if $n \neq m$ and equals 2π if $n = m$.

Let $\mathcal{F}$ be the set of finite linear combinations of the u_n. We want to show that $\mathcal{F}$ is a dense subset of $L^2([0, 2\pi))$. The first step is to show that the closure of $\mathcal{F}$ with respect to the supremum norm is equal to the set of continuous functions f on $[0, 2\pi)$ with $f(0) = f(2\pi)$. We will accomplish this by using the Stone-Weierstrass theorem, Theorem 1.7.

We identify the set of continuous functions on $[0, \infty)$ that take the same value at 0 and 2π with the continuous functions on the circle. To do this, let $S = \{e^{i\theta} : 0 \le \theta < 2\pi\}$ be the unit circle in $\mathbb{C}$. If f is continuous on $[0, 2\pi)$ with $f(0) = f(2\pi)$, define $\widetilde{f} : S \to \mathbb{C}$ by $\widetilde{f}(e^{i\theta}) = f(\theta)$. Note $\widetilde{u}_n(e^{i\theta}) = e^{in\theta}$.

Let $\widetilde{\mathcal{F}}$ be the set of finite linear combinations of the $\widetilde{u}_n$. S is a compact metric space. Since the complex conjugate of $\widetilde{u}_n$ is $\widetilde{u}_{-n}$, then $\widetilde{\mathcal{F}}$ is closed under the operation of taking complex conjugates. Since $\widetilde{u}_n \cdot \widetilde{u}_m = \widetilde{u}_{n+m}$, it follows that $\mathcal{F}$ is closed under the operation of multiplication. That it is closed under scalar multiplication and addition is obvious. $\widetilde{u}_0$ is identically equal to 1, so $\widetilde{\mathcal{F}}$ vanishes at no point. If $\theta_1, \theta_2 \in S$ and $\theta_1 \ne \theta_2$, then $\theta_1 - \theta_2$ is not an integer multiple of 2π, so

$$\frac{\widetilde{u}_1(\theta_1)}{\widetilde{u}_1(\theta_2)} = e^{i(\theta_1 - \theta_2)} \ne 1,$$

or $\widetilde{u}_1(\theta_1) \ne \widetilde{u}_1(\theta_2)$. Therefore $\mathcal{F}$ separates points. By the Stone-Weierstrass theorem (Theorem 1.7), the closure of $\mathcal{F}$ with respect to the supremum norm is equal to the set of continuous complex-valued functions on S.

If $f \in L^2([0, 2\pi))$, then

$$\int |f - f\chi_{[1/m, 2\pi - 1/m]}|^2 \to 0$$

by dominated convergence as $m \to \infty$. By Corollary 15.6 any function in $L^2([1/m, 2\pi - 1/m])$ can be approximated in L^2 by continuous functions which have support in the interval $[1/m, 2\pi - 1/m]$. By what we showed above, a continuous function with support in $[1/m, 2\pi - 1/m]$ can be approximated uniformly on $[0, 2\pi)$ by elements of $\mathcal{F}$. Finally, if g is continuous on $[0, 2\pi)$ and $g_m \to g$ uniformly on $[0, 2\pi)$, then $g_m \to g$ in $L^2([0, 2\pi))$ by dominated convergence. Putting all this together proves that $\mathcal{F}$ is dense in $L^2([0, 2\pi))$.

It remains to show the completeness of the u_n. If f is orthogonal to each u_n, then it is orthogonal to every finite linear combination, that is, to every element of $\mathcal{F}$. Since $\mathcal{F}$ is dense in $L^2([0, 2\pi))$, we can find $f_n \in \mathcal{F}$ tending to f in L^2. Then

$$\|f\|^2 = |\langle f, \overline{f} \rangle| \le |\langle f - f_n, \overline{f} \rangle| + |\langle f_n, \overline{f} \rangle|.$$

The second term on the right of the inequality sign is 0. The first term on the right of the inequality sign is bounded by $\|f - f_n\| \, \|f\|$

by the Cauchy-Schwarz inequality, and this tends to 0 as $n \to \infty$. Therefore $\|f\|^2 = 0$, or $f = 0$, hence the $\{u_n\}$ are complete. Therefore $\{u_n\}$ is a complete orthonormal system.

Given f in $L^2([0, 2\pi))$, write

$$c_n = \langle f, u_n \rangle = \int_0^{2\pi} f\overline{u_n}\, dx = \frac{1}{\sqrt{2\pi}} \int_0^{2\pi} f(x) e^{-inx}\, dx,$$

the Fourier coefficients of f. Parseval's identity says that

$$\|f\|^2 = \sum_n |c_n|^2.$$

For any f in L^2 we also have

$$\sum_{|n| \le N} c_n u_n \to f$$

as $N \to \infty$ in the sense that

$$\Big\| f - \sum_{|n| \le N} c_n u_n \Big\|_2 \to 0$$

as $N \to \infty$.

Using $e^{inx} = \cos nx + i \sin nx$, we have

$$\sum_{n=-\infty}^{\infty} c_n e^{inx} = A_0 + \sum_{n=1}^{\infty} B_n \cos nx + \sum_{n=1}^{\infty} C_n \sin nx,$$

where $A_0 = c_0$, $B_n = c_n + c_{-n}$, and $C_n = i(c_n - c_{-n})$. Conversely, using $\cos nx = (e^{inx} + e^{-inx})/2$ and $\sin nx = (e^{inx} - e^{-inx})/2i$,

$$A_0 + \sum_{n=1}^{\infty} B_n \cos nx + \sum_{n=1}^{\infty} C_n \sin nx = \sum_{n=-\infty}^{\infty} c_n e^{inx}$$

if we let $c_0 = A_0$, $c_n = B_n/2 + C_n/2i$ for $n > 0$ and $c_n = B_n/2 - C_n/2i$ for $n < 0$. Thus results involving the u_n can be transferred to results for series of sines and cosines and vice versa.

19.5 Exercises

Exercise 19.1 For $f, g \in L^2([0, 1])$, let $\langle f, g \rangle = \int_0^1 f(x)\overline{g(x)}\, dx$. Let $H = C([0, 1])$ be the functions that are continuous on $[0, 1]$. Is H a Hilbert space with respect to the norm defined in terms of the inner product $\langle \cdot, \cdot \rangle$? Justify your answer.

Exercise 19.2 Suppose H is a Hilbert space with a countable basis. Suppose $\|x_n\| \to \|x\|$ as $n \to \infty$ and $\langle x_n, y\rangle \to \langle x, y\rangle$ as $n \to \infty$ for every $y \in H$. Prove that $\|x_n - x\| \to 0$ as $n \to \infty$.

Exercise 19.3 Prove that if M is a closed subset of a Hilbert space H, then $(M^\perp)^\perp = M$. Is this necessarily true if M is not closed? If not, give a counterexample.

Exercise 19.4 Prove that if H is infinite-dimensional, that is, it has no finite basis, then the closed unit ball in H is not compact.

Exercise 19.5 Suppose a_n is a sequence of real numbers such that

$$\sum_{n=1}^\infty a_n b_n < \infty$$

whenever $\sum_{n=1}^\infty b_n^2 < \infty$. Prove that $\sum_{n=1}^\infty a_n^2 < \infty$.

Exercise 19.6 We say $x_n \to x$ *weakly* if $\langle x_n, y\rangle \to \langle x, y\rangle$ for every y in H. Prove that if x_n is a sequence in H with $\sup_n \|x_n\| \le 1$, then there is a subsequence $\{n_j\}$ and an element x of H with $\|x\| \le 1$ such that x_{n_j} converges to x weakly.

Exercise 19.7 If A is a measurable subset of $[0, 2\pi]$, prove that

$$\lim_{n\to\infty} \int_A e^{inx}\, dx = 0.$$

This is known as the *Riemann-Lebesgue lemma.*

Exercise 19.8 The purpose of Exercise 13.4 was to show that in proving the Radon-Nikodym theorem, we can assume that $\nu(A) \le \mu(A)$ for all measurable A. Assume for the current problem that this is the case and that μ and ν are finite measures. We use this to give an alternative proof of the Radon-Nikodym theorem.

For f real-valued and in L^2 with respect to μ, define $L(f) = \int f\, d\nu$.
(1) Show that L is a bounded linear functional on $L^2(\mu)$.
(2) Conclude by Theorem 19.9 that there exists a real-valued measurable function g in $L^2(\mu)$ such that $L(f) = \int fg\, d\mu$ for all $f \in L^2(\mu)$. Prove that $d\nu = g\, d\mu$.

Exercise 19.9 Suppose f is a continuous real-valued function on $\mathbb{R}$ such that $f(x+1) = f(x)$ for every x. Let γ be an irrational number. Prove that

$$\lim_{n\to\infty} \frac{1}{n} \sum_{j=1}^{n} f(j\gamma) = \int_0^1 f(x)\,dx.$$

Exercise 19.10 If M is a closed subspace of a Hilbert space, let $x + M = \{x + y : y \in M\}$.
(1) Prove that $x + M$ is a closed convex subset of H.
(2) Let Qx be the point of $x+M$ of smallest norm and $Px = x - Qx$. P is called the *projection* of x onto M. Prove that P and Q are mappings of H into M and $M^\perp$, respectively.
(3) Prove that P and Q are linear mappings.
(4) Prove that if $x \in M$, then $Px = x$ and $Qx = 0$.
(5) Prove that if $x \in M^\perp$, then $Px = 0$ and $Qx = x$.
(6) Prove that

$$\|x\|^2 = \|Px\|^2 + \|Qx\|^2.$$

Bibliography

[1] G.B. Folland. *Real Analysis: Modern Techniques and their Applications*, 2nd ed. Wiley, New York, 1999.

[2] H. Royden. *Real Analysis*, 3rd ed. Macmillan, New York, 1988.

[3] W. Rudin. *Principles of Mathematical Analysis*, 3rd ed. McGraw-Hill, New York, 1976.

[4] W. Rudin. *Real and Complex Analysis*, 3rd ed. McGraw-Hill, New York, 1987.

Index