Christopher Heil

Introduction to Real Analysis

Guide and Commentary for Instructors (and Students)

Last Updated: May 18, 2024

©2024 by Christopher Heil

One-Semester Weekly Course Schedule

As discussed in the Preface to the text, I use this text as the basis for a one-semester course that covers most of Chapters 2 through 8. Not every theorem and proof is presented in class. Below is an outline of a weekly schedule that I followed in a recent semester at Georgia Tech. A detailed, annotated list of the definitions, lemmas, theorems, etc. that were presented appears later in this guide.

ONE-SEMESTER WEEKLY OUTLINE: CHAPTERS 2–8

This outline is based on the standard U.S. system of three 50-minute lectures or two 75-minute lectures per week. Below, "n.n" refers to "Section n.n" of the main text.

Week 1: 2.1, 2.2 through closure under countable unions

Week 2: Finish 2.2, begin 2.3 through continuity from above and below

Week 3: Finish 2.3, cover 2.4 (basics of existence of nonmeasurable sets only)

Week 4: 3.1, 3.2, 3.3

Week 5: 3.4, 3.5 (omit 3.6), begin 4.1

Week 6: Finish 4.1, cover 4.2, 4.3

Week 7: 4.4, 4.5

Week 8: 4.6, 5.1

Week 9: 5.2, briefly cover 5.3–5.5 as indicated in detailed guide

Week 10: 6.1, 6.2, 6.3

Week 11: 6.4 (omit 6.5, 6.6), 7.1

Week 12: 7.2, 7.3, 7.4, 8.1

Week 13: 8.2, 8.3, 8.4

A pretty fast pace is required to cover all of the material above (though, as noted in the detailed guide that follows, I do expect students to *read* the text, and so I do not present every single item from the text in class).

Many variations are possible. For example, if an instructor prefers to present Chapter 1 or the online Alternative Chapter 1 in class, then some of the later chapters could be omitted. Alternatively, by covering only Chapters 2 through 6 or 2 through 7, the material could be presented at a slower pace or with more details than are indicated in this schedule.

Guide for a One-Semester Course

This guide shows what I usually cover when I use this text for a one-semester course. Additionally, the guide contains extra comments on the results, along with some alternative proofs and extra problems. Most of these comments are aimed at the instructor, but they may also be useful to students in the course or for readers who are using the text for self-study.

For the instructor, the outline below is only offered as a rough guide. You should adjust, add, subtract, deviate, substitute, or completely ignore my comments as you see fit, depending on your feelings for what is important and taking into account the background and preparation of your particular audience. I often make adjustments to this outline, especially since I usually assign different problems for homework each time I teach it. Topics that students have seen in homework might not need to be covered in class—or perhaps once the students have had a taste of a topic in the homework then it may actually be better to spend *more* time on it in class, perhaps delving into further details or related results.

It is not possible to present every result in the text in class. The instructor will need to make some choices as to what should be stated and proved in class versus what can be stated without proof or simply skipped. I give my suggestions for these choices, but each instructor will no doubt have their own preferences.

For the student, it is not necessary to read this guide, but the extra discussion, notes, and extra problems may provide some help (or hopefully at least a little amusement).

Exercises and Problems. Many exercises and problems appear in each section of the main text. The Exercises are directly incorporated into the development of the theory in each section, while the Problems given at the end of each section provide further practice and opportunities to develop understanding. This guide also contains some Extra Problems that are not included in the main text.

Textbook website. This guide and additional online resources for the text are available at the author's webpage:

`http://people.math.gatech.edu/~heil/real`

Extra resources include:

- Handout: "A Short Review of Cardinality."
- **Chapter 0**, containing a greatly expanded discussion of the material that is briefly covered in the **Preliminaries** section of the text.
- **Alternative Chapter 1**, containing a greatly expanded discussion of the material that is covered in **Chapter 1** of the text.
- **Chapter 10**, a bonus chapter covering abstract measure theory.
- **Selected Solutions for Students**, containing a worked solution to roughly one problem or exercise from each section of the text.
- An **Errata List** for the main text.

Disclaimer. The material in this guide has not been proofread as carefully as has the text proper, so there is a greater probability of errors here than in the main text. Please send comments and corrections to me at the email address "`heil@math.gatech.edu`". I will periodically update and correct this online guide. Unfortunately I cannot update the printed text, but I will maintain an errata list on my website.

NOTES. Within these comments, there are some "asides" prefixed by "*Note*:". These are usually just side remarks for the instructor about something that I find interesting but which is not part of the mainstream development of the course. Some of the remarks are extensive and may give alternative proofs or interesting extra results, but most of the extra material is not meant to be presented in class.

EXTRA PROBLEMS. In the outline below, after the comments for a given section I sometimes list some extra problems for that section that did not make it into the text. There is no particular rhyme or reason to these problems—some were too easy and a few too difficult to include, some just did not seem to fit (especially given that there are a lot of problems in the text already), and some are problems that I only recently came across (and may not even have worked out myself yet!). Some of these extra problems are quite nice, but some are less interesting. I would certainly like to hear your suggestions of more good problems that I could include in this guide (email me at `heil@math.gatech.edu`).

A Short Remark. The name "Heil" is pronounced "Hi-Ell" (rhymes with "mile", "style", "pile", and "smile").

PRELIMINARIES

The unnumbered Preliminaries chapter in the main text is background material that I do not present in class. Students should already be familiar with most of this material, with some exceptions noted below. For the benefit of those readers who would appreciate more detail on the preliminary material, an expanded version of the preliminaries is available in the online **Chapter 0**, which is posted on my website for this text.

The Extended Real Line. Students may not have seen the extended real line $[-\infty, \infty]$ before, but they should not have a problem with it as long as they remember that indeterminate forms need special care.

Extended real values arise naturally in the discussion of measure and integral. For example, even just considering the measures of sets, there are sets whose measure is infinite (for example, the entire Euclidean space $\mathbb{R}^d$ has infinite Lebesgue measure). Another situation that we often encounter is an infinite series of nonnegative functions. If we have functions $f_n \geq 0$ (meaning $f_n(t) \geq 0$ for every n and t), then the series $\sum\limits_{n=1}^{\infty} f_n(t)$ converges for every input t *in the extended real sense*—it could be infinite, but it either converges to a finite real number or it diverges to ∞. Hence $\sum\limits_{n=1}^{\infty} f_n(t)$ is well-defined at every point if we allow it to take extended real values.

The notation $\overline{\mathbf{F}}$. In many circumstances in analysis, we want to be able to use either the real line $\mathbb{R}$ or the complex plane $\mathbb{C}$ as our *scalar field*. In these situations, it is not uncommon to use a symbol such as $\mathbb{F}$ to denote a choice of $\mathbb{R}$ or $\mathbb{C}$. For example, this notation is used in both [Heil11] and [Heil18]. However, in this text the natural choice is between the extended real line $[-\infty, \infty]$ (which some authors denote by $\overline{\mathbb{R}}$) and the complex plane $\mathbb{C}$. In particular, we want to consider both extended real-valued functions and complex-valued functions. The extended real line is not even a group under addition, but it is related to the field $\mathbb{R}$. Hence fields are still the issue, and this is reason for the choice of the letter "F" in this context. In this text, we let $\overline{\mathbf{F}}$ denote a choice of either the extended real line $\overline{\mathbb{R}} = [-\infty, \infty]$ or the complex plane $\mathbb{C}$.

Note: There is another notion, useful in topological contexts, of the *one-point compactification* of $\mathbb{R}$ or of $\mathbb{C}$. This is a distinct concept that will not be used in this text. In particular, for our purposes in analysis, it is not useful to try to define a "complex infinity" in any way.

Countability. Countability is an important concept that is used throughout the course. Most students will likely have been introduced to countability in an undergraduate real analysis class. However, I often see non-math graduate students in the class who are well-prepared in most of the other background material but have not had much experience with countability. If your class contains a large number of such students, then it might be a good idea to give a quick review of countability and uncountability. A short handout devoted solely to cardinality is available at my website for this text, and the same material is also available in the online **Chapter 0**.

Sups and Limsups. Students need to be familiar with suprema, infima, and convergent sequences. Limsup and liminf also play a very important role in the course. Not every student is as comfortable with infs, sups, liminfs, and limsups as I would like, but if they are going to succeed, then they need to be able to review this on their own (there is some discussion of this material in the online **Chapter 0**). Students do not need to be intimately familiar with limsup/liminf right from the beginning of the course, but they will need to review the ideas along the way.

Pretest. The next page has a short pretest over the preliminary material. This may help students determine if they are sufficiently prepared to take this course. All students should be able to complete this pretest, although the limsup question may be challenging for some.

Analysis Preliminaries Pretest

After reading the Preliminaries section in the text, give detailed, rigorous proofs of the following statements.

1. If $f\colon X \to Y$ is a function, $B \subseteq Y$, and $\{B_i\}_{i \in I}$ is a family of subsets of Y, then

$$f^{-1}\biggl(\bigcup_{i\in I} B_i\biggr) = \bigcup_{i\in I} f^{-1}(B_i), \qquad f^{-1}\biggl(\bigcap_{i\in I} B_i\biggr) = \bigcap_{i\in I} f^{-1}(B_i),$$

and $f^{-1}\bigl(B^{\mathrm{C}}\bigr) = (f^{-1}(B))^{\mathrm{C}}$.

2. If $(x_n)_{n\in\mathbb{N}}$ and $(y_n)_{n\in\mathbb{N}}$ are sequences of real numbers, then

$$\inf_{n\in\mathbb{N}} x_n + \inf_{n\in\mathbb{N}} y_n \le \inf_{n\in\mathbb{N}} (x_n + y_n) \le \sup_{n\in\mathbb{N}} (x_n + y_n) \le \sup_{n\in\mathbb{N}} x_n + \sup_{n\in\mathbb{N}} y_n.$$

Show by example that any of the inequalities on the preceding line can be strict.

3. If $(x_n)_{n\in\mathbb{N}}$ is a sequence of real numbers, then

$$\begin{array}{c}(x_n)_{n\in\mathbb{N}} \text{ converges in} \\ \text{the extended real sense}\end{array} \quad \Longleftrightarrow \quad \liminf_{n\to\infty} x_n = \limsup_{n\to\infty} x_n,$$

4. If $c_n \ge 0$ for every n and $\sum\limits_{n=1}^{\infty} c_n < \infty$, then

$$\lim_{N\to\infty} \biggl(\sum_{n=N}^{\infty} c_n\biggr) = 0.$$

CHAPTER 1: METRIC AND NORMED SPACES

Metrics and norms appear throughout the text—occasionally in Chapters 2 through 6, and frequently in Chapters 7, 8, and 9. Metrics and norms are reviewed in Chapter 1. This material is usually covered, at least to some extent, in an undergraduate real analysis class, although some students may only have seen these ideas in the setting of the Euclidean space $\mathbb{R}^d$, rather than in abstract metric spaces and normed spaces.

One option (especially for better-prepared classes) is to assign Chapter 1 as background reading, and then just briefly discuss issues related to metrics or norms when they actually come up in the text. Alternatively, it may make sense to use this background on metric spaces and normed spaces as the introduction to the course. In that case, Chapter 1 covers enough background material that the rest of the text is essentially self-contained.

However, the presentation in Chapter 1 is very compressed. It summarizes the needed material, but does not explain or motivate it, and only a few problems are included. Therefore, an **Alternative Chapter 1** is available in the online resources for this text. It provides much more detailed coverage. This alternative chapter provides a comprehensive review of normed spaces and related topics, and it does so with discussion, motivation, examples, and problems. Thus, if the instructor decides to present this material in class, it may be more beneficial to base the presentation on **Alternative Chapter 1**.

Remarks on Section 1.4. Hölder continuity will play only a limited role in the text, but Lipschitz continuity will appear more often, especially in Chapters 5 and 6. Lipschitz continuity is briefly reviewed in the main text in the places where it comes up.

Pretest. For instructors that are assuming at least some familiarity with Chapter 1, the next page has a potential short pretest on that material.

Metrics and Norms Pretest

After reading Chapter 1 in the text, give detailed, rigorous proofs of the following statements.

1. If E is a subset of a metric space X, then E is closed if and only if the following statement holds:

$$\text{If } x_n \in E \text{ and } x_n \to x \in X, \text{ then } x \in E.$$

2. If $\{x_n\}_{n\in\mathbb{N}}$ is a sequence in a metric space X, then $x_n \to x$ if and only if for every subsequence $\{y_n\}_{n\in\mathbb{N}}$ of $\{x_n\}_{n\in\mathbb{N}}$ there exists a subsequence $\{z_n\}_{n\in\mathbb{N}}$ of $\{y_n\}_{n\in\mathbb{N}}$ such that $z_n \to x$.

3. If X is a normed space, then each open ball $B_r(x)$ is an open, convex set in X.

4. Determine, with proof, whether the following statements are true or false.

(a) $f(x) = x^2$ is uniformly continuous on $\mathbb{R}$.

(b) $f(x) = \sin x^2$ is uniformly continuous on $\mathbb{R}$.

(c) If $f_n(x) = x/n$ then $\{f_n\}_{n\in\mathbb{N}}$ converges uniformly on $\mathbb{R}$.

(d) If $f \in C_0(\mathbb{R})$ then f is uniformly continuous.

5. (a) Given functions $f_n \in C_b(\mathbb{R})$, explain *precisely* what the statement

$$\sum_{n=1}^{\infty} f_n \text{ converges in } C_b(\mathbb{R})$$

means.

(b) Exhibit functions $f_n \in C_b(\mathbb{R})$ such that $\sum_{n=1}^{\infty} f_n$ converges in $C_b(\mathbb{R})$.

(c) Exhibit functions $f_n \in C_b(\mathbb{R})$ such that $\sum_{n=1}^{\infty} f_n$ does not converge in $C_b(\mathbb{R})$.

CHAPTER 2: LEBESGUE MEASURE

This is where the course really begins. The notes below outline what I usually try to cover in class for a one-semester course. You should adjust as appropriate for you and your class. I cover most (though certainly not all) of Chapters 2–8 in one semester, but unfortunately must usually leave Chapter 9 as "bonus material" that I hope the students will read on their own.

For presentation in a single semester the pace needs to be fast (perhaps not so much for the instructor who has taught this material for years, but it is certainly fast for the student who has never seen it). As noted in the outline below, I do not present every fact from the text or every proof in class. Often I prove only one direction of an if and only if theorem in class, or only certain parts of a theorem that has multiple parts. That does not mean that I just breeze through the material—rather, I try to focus my time on the concepts that especially need my explanation in class, knowing that the text has additional details and discussion. I expect the student to be diligent and to read the text for further discussion, examples, proofs, problems, and so forth.

Comments on Chapter 2's Introduction

One point that I'm trying to make in the introduction to Chapter 2 is that as we progress through the course we will see many "obvious" facts that we will prove are true, but nearly as many "obvious" facts that we will prove are false. Here, it seems obvious that we should be able to construct a measure that satisfies all of properties (i)–(iv) for all subsets of $\mathbb{R}^d$, yet the Axiom of Choice (which seems to me to be a perfectly reasonable axiom!) implies that we *cannot* do so.

Aside from the existence of nonmeasurable sets, the Axiom of Choice is rarely mentioned in the remainder of the text. I usually say a few words in class about what the Axiom of Choice is and why it is "reasonable", but I keep it brief. There is some discussion of the Axiom of Choice in the text at the beginning of Section 2.4, just before the proof of the existence of a nonmeasurable set.

Note: For one "obvious" fact that is still an *open problem* in analysis, see the discussion of the *HRT Conjecture* in the comments following the instructor's guide material relating to Section 8.4.

Section 2.1: Exterior Lebesgue Measure

2.1.1 Boxes

I have to admit that Subsections 2.1.1 and 2.1.2 are not particularly exciting. Unfortunately, before we can get to the interesting material, we do have to establish some machinery involving boxes. Most of the results presented here are "obvious" facts that are indeed true, such as the fact that if we subdivide a box into finitely many smaller boxes, then the sum of the volumes of the small boxes equals the volume of the original box (see Lemma 2.1.6). Moreover, the proofs of these facts are essentially straightforward, albeit often tedious, calculations. Therefore I only state these beginning results in class (as detailed below), and tell the students that they should read the proofs in the text, or work them out on their own. The more interesting stuff is coming soon! As soon as we get to Subsection 2.1.3 (the definition of exterior measure), I start giving proofs in class.

Definition 2.1.1 (Boxes). State.

Note: We could have allowed degenerate boxes (where $a_i = b_i$ for some i) if we like; in that case we would have to make corresponding changes throughout the text but would end up with the same theorems. Personally I prefer my boxes to be nondegenerate.

Definition 2.1.2 (Nonoverlapping Boxes). State.

Notation 2.1.3 (Countable Collections of Boxes). State.

Note: A countable collection of infinitely many nonoverlapping boxes $\{Q_k\}_{k\in\mathbb{N}}$ indexed by the natural numbers cannot always be put into "increasing order," where we always have Q_j to the left of Q_k when $j < k$. If we *could* do this, then we could write $Q_k = [a_k, b_k]$ where

$$a_1 < b_1 \le a_2 < b_2 \le a_3 < b_3 \le \cdots . \tag{2.A}$$

For example, consider the collection $\bigl\{[a_k, b_k]\bigr\}_{k\in\mathbb{N}}$ where $[a_1, b_1] = [1, 2]$ and

$$[a_2, b_2] = \bigl[0, \tfrac{1}{2}\bigr], \quad [a_3, b_3] = \bigl[\tfrac{1}{2}, \tfrac{2}{3}\bigr], \quad [a_4, b_4] = \bigl[\tfrac{2}{3}, \tfrac{3}{4}\bigr],$$

and so forth. These boxes are nonoverlapping, and equation (2.A) does not hold. As long as we require the index set to be the natural numbers, there is no way to relabel the boxes so that equation (2.A) does hold.

Definition 2.1.4. State.

2.1.2 Some Facts about Boxes

Lemma 2.1.5. This result is often proved in undergraduate real analysis classes, so I only included the proof in the text for the sake of completeness.

I usually state the lemma and give some quick motivation for its proof via a picture on the board, but I assign the proof itself for reading.

Lemma 2.1.6. The basic philosophy here is that anything that should be true about *finitely many boxes* probably is true, especially if it only involves *volume*. When it comes to *measure* or *infinitely many boxes* then the situation may very well be more complicated! However, boxes and their volumes are "straightforward" (although the calculations may be tedious). State this lemma, but assign the proof as reading.

Exercise 2.1.7. This exercise will be used later, so it is worth stating. However, it again fits the philosophy that *volumes of* ***finitely many*** *boxes are "well-behaved,"* and the reader should not get bogged down trying to prove this exercise. The interesting things start with the *next* definition, which introduces exterior Lebesgue measure.

TYPO in the text: the word "nonoverlapping" should be removed from the hypotheses of Exercise 2.1.7. The result is true even if the boxes overlap, and in fact we do need this more general fact when we apply the exercise in the proof of Theorem 2.1.17.

2.1.3 Exterior Lebesgue Measure

Definition 2.1.8 (Exterior Lebesgue Measure). Motivate and state.

Note: What happens if we use covers by *finitely many* boxes instead of countably many boxes? This is addressed below in this Instructor's Guide; see Extra Problem 3 below in Section 2.2.

Lemma 2.1.9. State. There isn't a proof to give here, as these are simply immediate consequences of the definition of an infimum. This is a good opportunity to advise the students that if these facts about an inf are not clear to them (and hopefully *immediately* clear), then they will have considerable difficulty in the rest of the course.

Example 2.1.10. Omit. There isn't time to cover every detail or remark in class; students should be reading the text in addition to listening to lectures.

Lemma 2.1.11. State. No proofs have been done so far, it's time to start. I usually prove part (d), and perhaps give a short sketch of the ideas behind the proof of parts (a) or (b).

Remark 2.1.12. Omit.

Theorem 2.1.13 (Countable Subadditivity). State and prove.

Definition 2.1.14 (Limsup and Liminf of Sets). State. Although this definition may seem somewhat esoteric at first glance, the characterization given in Exercise 2.1.15 is what makes it useful.

Note: By applying De Morgan's Laws, it follows that

$$\Bigl(\limsup_{k\to\infty} E_k\Bigr)^{\mathrm{C}} = \biggl(\bigcap_{j=1}^{\infty}\biggl(\bigcup_{k=j}^{\infty} E_k\biggr)\biggr)^{\mathrm{C}} = \bigcup_{j=1}^{\infty}\biggl(\bigcap_{k=j}^{\infty} E_k^{\mathrm{C}}\biggr) = \liminf_{k\to\infty} E_k^{\mathrm{C}}.$$

Exercise 2.1.15. State. We will use this exercise several times in coming proofs.

Exercise 2.1.16 (Borel–Cantelli Lemma). State. The proof is a nice exercise for the student. We will use this result in later proofs, although not as often as the characterizations given in Exercise 2.1.15.

2.1.4 The Exterior Measure of a Box

Theorem 2.1.17 (Consistency with Volume). State and prove. The idea of the proof is that we want to reduce to a situation that involves *volumes* (not measures) of *finitely many boxes*, because we can then use the machinery developed earlier regarding volumes of finitely many boxes. To reduce to finitely many we make use of the fact that a box is compact, so if we cover it with open sets then that cover can be reduced to a covering by finitely many of those sets. We have a covering by closed boxes, so to get open sets we fatten up each of the boxes a little to get a covering by open boxes.

This is a good place to (again) observe that students do need to have adequate background knowledge of undergraduate real analysis. In particular, they need to know the *topological definition* of a compact set, not just that a subset of $\mathbb{R}^d$ is compact if and only if it is closed and bounded.

Remark 2.1.18. Omit.

Corollary 2.1.19. I like to prove this, but it's not strictly necessary.

Note: The fact that the measure of $\mathbb{R}^d$ is infinite is not that easy to to prove *directly* from the definition of exterior measure. We easily get the inequality $|Q_k|_e \le \mathrm{vol}(Q_k)$ from the definition of exterior measure, and the inequality $|Q_k|_e \le |\mathbb{R}^d|_e$ from monotonicity. However, these facts alone do not give us any information about the value of $|\mathbb{R}^d|_e$. With Theorem 2.1.17 we see that $|Q_k|_e = \mathrm{vol}(Q_k)$, and *then* we can see that $|\mathbb{R}^d|_e = \infty$.

Exercise 2.1.20. State.

Lemma 2.1.21. State and prove.

Corollary 2.1.22. Omit.

2.1.5 The Cantor Set

Example 2.1.23 (The Cantor Set). Many students will have seen the Cantor set in an undergraduate analysis class, but not all, so it is worth giving the construction. It is a remarkable set!

Exercise 2.1.24. State.

Note: If x is in $(0, 1/3)$, then x belongs to F_1 and we have $d_1 = 0$ (see Figure 2.A). If $x \in (1/3, 2/3)$, then x does not belong to F_1 and $d_1 = 1$. If $x \in (2/3, 1)$, then x belongs to F_1 and $d_1 = 2$. The point $x = 1/3$ has two ternary expansions, one of which is $0.0222\ldots$, which has $d_1 = 0$. The point $x = 2/3$ has two ternary expansions, one of which is $0.2000\ldots$, which has $d_1 = 2$.

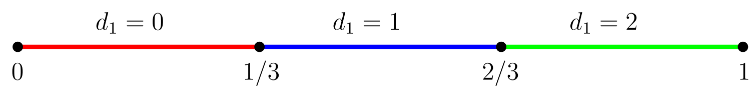

Fig. 2.A The first digit d_1 of the ternary expansion of x is $d_1 = 0$ if x lies in the red interval, $d_1 = 1$ if x lies in the blue interval, and $d_1 = 2$ if x lies in the green interval.

Note: Some non-math students may not be very familiar with the concept of uncountability, even though they may have sufficient familiarity with the other background material. A handout on countability is available on my webpage for this text, and additionally there is a section on countability in the online **Chapter 0** (which provides an expanded version of the notation and preliminaries for the text). These are available at

`http://people.math.gatech.edu/~heil/books/real`

Exercise 2.1.25. State.

Example 2.1.26 (The Fat Cantor Set). This is a fun example, so I usually sketch the construction.

Students often assume that any set that has positive measure must contain an open ball, but this just isn't true:

There exist sets that have positive measure but contain no open balls!

There is an easy counterexample, namely the set of irrationals in $[0, 1]$. This set has measure 1 and contains no intervals. On the other hand, the set of irrationals is not a closed set, and the big surprise provided by the fat Cantor set is that *there exist* ***closed*** *sets that have positive measure and contain no intervals*!

Note: The Cantor set corresponds to the choice $a_n = 3^{-n}$. If we take $a_n = 2^{-3n}$, then $|P|_e = 2/3$.

2.1.6 Regularity of Exterior Measure

Theorem 2.1.27. State and prove.

Corollary 2.1.28. Omit.

Extra Problems for Section 2.1

1. Let S be the set of irrationals in $[0,1]$. Prove that S has positive exterior Lebesgue measure but S contains no open intervals.

2. In the construction of the Cantor set, each set F_n is a union of finitely many closed intervals, so F_n has only finitely many boundary points. Taking all n into account, it follows that only countably many of the points in the Cantor set are boundary points of some F_n. Does the point $x = \frac{1}{4}$ belong to the Cantor set? Is $x = \frac{1}{4}$ a boundary point of any of the sets F_n?

3. Let $S = (0,1)^2$ be the open unit square in $\mathbb{R}^2$. Prove that there does not exist a collection of disjoint open balls $\{B_i\}_{i\in I}$ whose union is S.

Hint: Fix $j \in I$, and prove that B_j and $S \setminus B_j$ are both open and nonempty. This contradicts the fact that S is *connected*. Note: Except for the fact that intervals in $\mathbb{R}$ are connected, connectedness is not otherwise used in the main text.

4. Prove that if a set $E \subseteq \mathbb{R}^d$ has nonempty interior, then $|E|_e > 0$.

Note: The converse does not hold in general, see Problem 2.2.42, or Extra Problem 1 above.

5. Suppose that U is an open, bounded subset of the real line, and let $\overline{U}$ be its closure. Must $\overline{U} \setminus U$ be countable?

6. Show that the definition of exterior Lebesgue measure is unchanged if we replace coverings by countably many *boxes* with coverings by countably many *cubes* (still with edges parallel to the coordinate axes).

7. Show that the definition of exterior Lebesgue measure is unchanged if we replace coverings by countably many *boxes* with coverings by countably many *open boxes* (where an open box is the interior of a box).

8. Suppose that a set $A \subseteq \mathbb{R}^d$ satisfies $|A \cap Q|_e \le |Q|/2$ for every box Q in $\mathbb{R}^d$. Prove that $|A|_e = 0$.

9. Let $\{E_k\}_{k\in\mathbb{N}}$ be a sequence of subsets of $\mathbb{R}^d$, and let

$$E \;=\; \limsup_{k\to\infty} E_k \;=\; \bigcap_{j=1}^{\infty} \biggl(\bigcup_{k=j}^{\infty} E_k\biggr).$$

Prove that

$$\limsup_{k\to\infty} \chi_{E_k}(x) \;=\; \chi_E(x), \qquad \text{for all } x \in \mathbb{R}^d.$$

Does the analogous statement for liminf hold?

10. (From Stein and Shakarchi). For each irrational number x, it can be shown that there exist infinitely many fractions p/q with p and q relatively prime such that $\bigl|x - p/q\bigr| \le 1/q^2$. Let E be the set of all $x \in \mathbb{R}$ for which there exist infinitely many fractions p/q with p and q relatively prime such that

$$\Bigl|x - \frac{p}{q}\Bigr| \;\le\; \frac{1}{q^3}. \qquad \text{(A)}$$

Prove that $|E|_e = 0$. Show also that the exponent 3 can be replaced by $2+\varepsilon$.

Hint: Borel–Cantelli.

11. Let $(\phi_k)_{k\in\mathbb{N}}$ be a sequence of positive integers such that $\sum_{k=1}^{\infty} k^2\,\phi_k^2 < \infty$. Let A be the set of all points $(x,y)\in\mathbb{R}^2$ for which there exist infinitely many $k\in\mathbb{N}$ such that $|(x,y)-(j/k,\ell/k)| < \phi_k$ for some $j,\ell\in\mathbb{Z}$. Prove that $|A| = 0$.

Hint: Fix $N\in\mathbb{N}$, and let A_N be the set of all points $(x,y)\in[-N,N]^2$ for which there exist infinitely many $k\in\mathbb{N}$ such that $|(x,y)-(j/k,\ell/k)| < \phi_k$ for some $j,\ell\in\mathbb{Z}$.

Section 2.2: Lebesgue Measure

2.2.1 Definition and Basic Properties

First paragraph TYPOS: The first paragraph of Section 2.2 has some minor typos. A better wording of the paragraph is as follows.

> To motivate the definition of measurability, suppose that U is an open set that contains a set E and satisfies $|U|_e \le |E|_e + \varepsilon$. As we observed above, we do not know whether $|U|_e$ and $|E|_e + |U\setminus E|_e$ will be equal. *If it were* the case that these quantities were equal, then $|E|_e + |U\setminus E|_e = |U|_e \le |E|_e + \varepsilon$. As long as E has finite measure, this implies that $|U\setminus E|_e \le \varepsilon$. The "measurable sets" are precisely the sets for which this inequality can be achieved. Here is the explicit definition.

Definition 2.2.1 (Lebesgue Measure). Motivate and state. I often draw a picture on the board. Figures 2.B and 2.C give sample one-dimensional and two-dimensional illustrations, respectively.

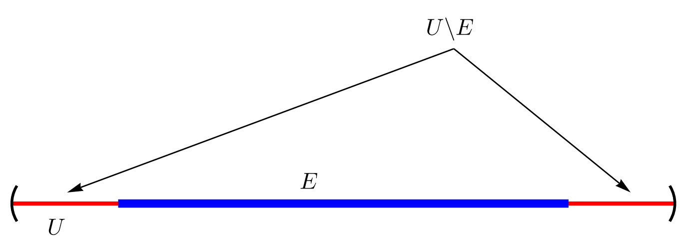

Fig. 2.B One-dimensional picture: If a set E (blue) is measurable, then for every $\varepsilon > 0$ there exists an open set U (black) that contains E such that the "annulus" $U\setminus E$ satisfies $|U\setminus E|_e \le \varepsilon$.

Note: When writing on the board, a standard abbreviation for "measurable" is ⓜ, the letter "m" with a circle around it.

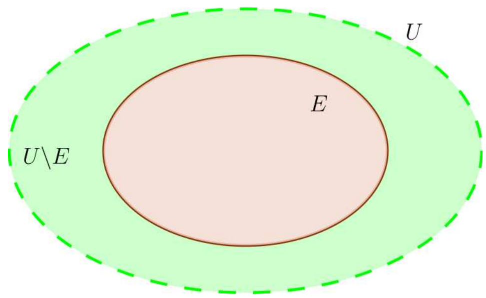

Fig. 2.C Two-dimensional picture: A set E (red) is measurable if for every $\varepsilon > 0$ there is an open set U (green) that contains E and satisfies $|U \setminus E|_e < \varepsilon$. The set E pictured here is a nice connected set, but in general E could be any measurable set.

Note: The following argument, which shows that Lebesgue measure is not additive on arbitrary sets, will be presented in Section 2.4, but it could be appropriate to go ahead and give it now.

Suppose that $E \subseteq \mathbb{R}^d$ is not measurable. Then there exists an $\varepsilon > 0$ such that no matter what open set $U \supseteq E$ that we choose, we will have $|U \setminus E|_e > \varepsilon$. On the other hand, Theorem 2.1.27 implies that there is at least one open set $U \supseteq E$ that satisfies $|U|_e \le |E| + \varepsilon$. By subadditivity,

$$|E \cup (U \setminus E)|_e = |U|_e \le |E| + \varepsilon.$$

Now, E and $U \setminus E$ are disjoint. Yet, since $|U \setminus E|_e > \varepsilon$, we have

$$|E|_e + |U \setminus E|_e > |E|_e + \varepsilon.$$

Therefore

$$|E \cup (U \setminus E)|_e \neq |E|_e + |U \setminus E|_e,$$

even though E and $U \setminus E$ are disjoint! Consequently, if nonmeasurable sets exist (which in Section 2.4 we will see is a consequence of the Axiom of Choice), then external Lebesgue measure is not *additive* on disjoint sets in general.

Notation 2.2.2. State.

Lemma 2.2.3 (Open Sets are Measurable). State and prove.

Lemma 2.2.4 (Null Sets are Measurable). State and prove.

Theorem 2.2.5 (Closure Under Countable Unions). State and prove.

2.2.2 Towards Countable Additivity and Closure under Complements

Corollary 2.2.6 (Boxes are Measurable). State and prove. The point here is that we can write $Q = Q^\circ \cup \partial Q$, and we know that Q° is measurable because it is open and ∂Q is measurable because it has measure zero—but we cannot extend this argument to arbitrary closed sets because the boundary of a closed set *need not* have measure zero in general (consider the fat Cantor set!).

Lemma 2.2.7. (*Separated Sets.*) State and prove. I find this result quite attractive. It states that exterior measure is additive for *arbitrary separated sets*. Hence additivity can fail only when the distance between two sets is zero. Even then, we will prove later that additivity holds if the two disjoint sets are *both measurable*.

Note: In the proof, some of the boxes intersect A, and those become part of the subsequence $\{Q_k^A\}$, and likewise some intersect B. There can also be boxes that do not intersect either of A or B, which is why

$$\sum_k |Q_k^A| \,+\, \sum_k |Q_k^B| \,\le\, \sum_k |Q_k|$$

is an inequality instead of an equality.

Note: Figure 2.4 in the text gives a two-dimensional illustration that if the sidelengths of the boxes Q_k are small enough, then each box can intersect at most one of A or B. A one-dimensional version of this idea is shown in Figure 2.D below.

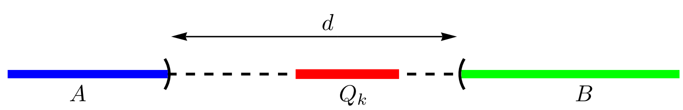

Fig. 2.D Two sets A and B in $\mathbb{R}$ are separated by a positive distance d. A box Q_k whose length is less than d cannot intersect both A and B.

Corollary 2.2.8. State. Perhaps briefly discuss the fact that disjoint compact sets are separated, while the distance between two disjoint closed sets can be zero (compare Problem 2.2.31).

Theorem 2.2.9 (Compact Sets are Measurable). State and prove. A picture on the board is helpful for motivating the proof. The boxes Q_k, and hence the sets R_N, are contained in the "annulus" $U \setminus F$. Figure 2.E gives a sample illustration for the one-dimensional setting (although I usually draw a two-dimensional version on the board).

Corollary 2.2.10 (Closed Sets are Measurable). State and prove.

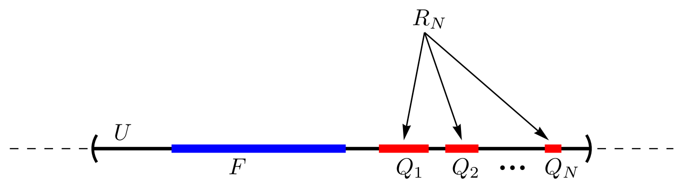

Fig. 2.E An open set U (black) surrounds a compact set F (blue), and nonoverlapping boxes $Q_1, \dots, Q_N$ (red) are contained in $U \setminus F$. The compact set R_N is the union of $Q_1, \dots, Q_N$.

Theorem 2.2.11(Closure Under Complements). State and prove. I usually draw an illustrative figure on the board, showing E, E^{C}, U_k, and $F_k = U_k^{\mathrm{C}}$. The fact that $E^{\mathrm{C}} \setminus U_j^{\mathrm{C}} = U_j \setminus E$ can be illustrated by a Venn diagram similar to the one that appears in Figure 2.F.

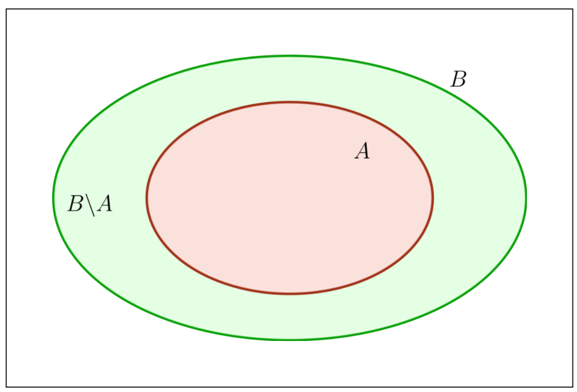

Fig. 2.F Venn diagram: The shaded green region is $B \setminus A$, which equals $A^{\mathrm{C}} \setminus B^{\mathrm{C}}$.

Corollary 2.2.12 (Closure Under Countable Intersections). State. This is an immediate consequence of De Morgan's Laws.

Corollary 2.2.13 (Closure Under Relative Complements). State.

Definition 2.2.14 (Sigma Algebra). State. The similarities and differences between topologies and σ-algebras could be mentioned here.

Note: The assumption that Σ is not empty implies that some subset E of X belongs to Σ. Consequently $E^{\mathrm{C}} \in \Sigma$ since Σ is closed under complements. Therefore $X = E \cup E^{\mathrm{C}} \in \Sigma$, and also $\varnothing = X^{\mathrm{C}} \in \Sigma$.

2.2.3 Countable Additivity

Lemma 2.2.15. State and prove. A picture is useful for this proof as well, see Figure 2.G.

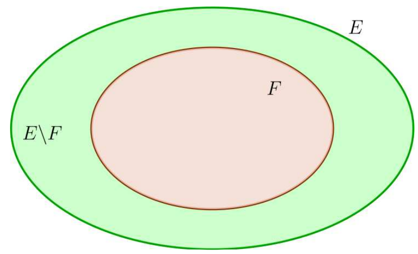

Fig. 2.G $E \backslash F = F^C \backslash E^C$ is the "annulus" between F and E.

Note: Problem 2.2.43 discusses the related concept of the *inner measure* of a set. In particular, that problem shows that if a set A has finite exterior measure, then A is measurable if and only if its exterior measure and inner measure are equal (but this can fail for sets whose measure is infinite).

Theorem 2.2.16 (Countable Additivity). State and prove.

The idea of Step 1 of the proof is that if our countably many disjoint sets E_k are all measurable and bounded, then inside each E_k we can find a compact F_k that is "almost" all of E_k. We proved earlier that Lebesgue measure is additive on *finitely many disjoint compact sets*, so

$$|F_1 \cup \cdots \cup F_N| = |F_1| + \cdots + |F_N|.$$

Letting N go to infinity, we see that the measure of $\cup F_k$ (which is "almost $\cup E_k$") is $\sum |F_k|$ (which is "almost $\sum |E_k|$"). The proof given in the text makes these "almosts" precise.

Step 2 uses the σ-finiteness of $\mathbb{R}^d$ to extend the proof from bounded sets to arbitrary sets. This is the first time this type of argument appears in the text, so I usually present the full proof of Step 2 in class, but after this I usually assign this type of extension as an exercise for the students to fill in on their own.

Note: In Step 2, the set E_k^j is measurable because it equals $E_k \cap B_1(0)$ when $j = 1$, and equals $E_k \cap \big(B_j(0) \setminus B_{j-1}(0)\big)$ when $j > 1$.

Note: In case the student is worried about the double sum that appears in the proof of Step 2, note that all terms in the summation are nonnegative, so Tonelli's Theorem for Series applies and allows us to interchange or reorder the summations as we like. Tonelli's Theorem for Series is not explicitly stated until Chapter 4, because it is related to Tonelli's Theorem for Lebesgue integrals (see Problem 4.6.23). However, the proof can be done directly without reference to measure theory. Therefore you can simply state that the proof that a double sum of nonnegative terms can be reordered is a fact from undergraduate real analysis that they could try to prove on their own. Similar results for integrals will be proved in Chapter 4.

Note: Here is an alternative wording of the same proof of Step 1.

Step 1. Assume first that each set E_k is bounded. By subadditivity we have

$$\left|\bigcup_{k=1}^{\infty} E_k\right| \leq \sum_{k=1}^{\infty} |E_k|,$$

so our task is to prove the opposite inequality.

Fix $\varepsilon > 0$. By Lemma 2.2.15, there exists a closed set $F_k \subseteq E_k$ such that

$$|E_k \setminus F_k| \leq \frac{\varepsilon}{2^k}.$$

Since E_k is bounded, F_k is compact. Hence $\{F_k\}_{k\in\mathbb{N}}$ is a collection of disjoint compact sets. Let $E = \cup E_k$ and $F = \cup F_k$. Then we compute that

$$\begin{aligned}
|E| \geq |F| &= \left|\bigcup_{k=1}^{\infty} F_k\right| \\
&= \lim_{N\to\infty} \left|\bigcup_{k=1}^{N} F_k\right| \qquad \text{(continuity from below)} \\
&= \lim_{N\to\infty} \sum_{k=1}^{N} |F_k| \qquad \text{(Corollary 2.2.8)} \\
&= \sum_{k=1}^{\infty} |F_k| \\
&\geq \sum_{k=1}^{\infty} \left(|E_k| - \frac{\varepsilon}{2^k}\right) = \left(\sum_{k=1}^{\infty} |E_k|\right) - \varepsilon.
\end{aligned}$$

Since ε is arbitrary, equation (2.14) follows.

Corollary 2.2.17. State, but omit the proof.

Note: I don't mention this in the text and usually don't talk about it in class either, but here is an application of Corollary 2.2.17. Suppose that E is a subset of $\mathbb{R}^d$ and we can find nonoverlapping boxes Q_k and nonoverlapping boxes R_k such that $\cup Q_k \subseteq E \subseteq \cup R_k$. Then by applying Corollary 2.2.17 and monotonicity it follows that

$$\sum_k |Q_k| \leq |E|_e \leq \sum_k |R_k|.$$

If we can compute these sums, then by making better and better choices of boxes we may be able to find the exact measure of E. For example, we could do this for sets like a triangle,

$$A = \{(x,y) : 0 \leq x \leq 1,\, 0 \leq y \leq x\}$$

or the region under the graph of x^2:

$$B = \{(x,y) : 0 \leq x \leq 1,\, 0 \leq y \leq x^2\}.$$

Essentially, the idea here is that we can compute the measure of the region under the graph of $y = x$ or $y = x^2$ by forming Riemann sums, as each Riemann sum is the Lebesgue measure of a union of finitely many nonoverlapping boxes. We could continue with this idea and try to relate Riemann integrals to the Lebesgue measure of the region under the graph of other functions. However, we will develop a better approach to integration in Chapter 4 when we define the Lebesgue integral. Further, the exact relation between the Lebesgue integral and the Riemann integral will be explored in Section 4.5.5.

2.2.4 Equivalent Formulations of Measurability

Definition 2.2.18 (G_δ-Sets and F_σ-Sets). State.

Example 2.2.19. I usually present the half-open interval example that precedes this example, but otherwise content myself with encouraging the students to read Example 2.2.19 on their own because it has a neat application of the Baire Category Theorem. (I usually prove Baire Category in the second semester of this course, so this gives the students something to look forward to.)

Lemma 2.2.20. State and prove.

Lemma 2.2.21. State, and prove the implication (a) $\Rightarrow$ (b).

TYPO in the paragraph preceding Exercise 2.2.22: The phrase "open subset of $\mathbb{R}^n$ under f is an open subset of $\mathbb{R}^m$" should be replaced with "open subset of $\mathbb{R}^m$ under f is an open subset of $\mathbb{R}^n$".

Exercise 2.2.22. State the exercise. Part (a) is a nice application of the definition of a compact set in terms of open covers, and if your students are not comfortable with the abstract definition of compact sets, then it might be worth working out this part in class.

Note: Part (a) is also stated in Chapter 1 as Exercise 1.1.14.

2.2.5 Carathéodory's Criterion

Carathéodory's Criterion is an important equivalent characterization of measurability. Its statement does not explicitly involve topology, although topology is still involved, hidden in the fact that the definition of Lebesgue measure involves boxes, which are very special closed sets (and every open set is a union of countably many boxes).

Note: Figure 2.5 in the text illustrates Carathéodory's Criterion; a color version of that figure is given below in Figure 2.H.

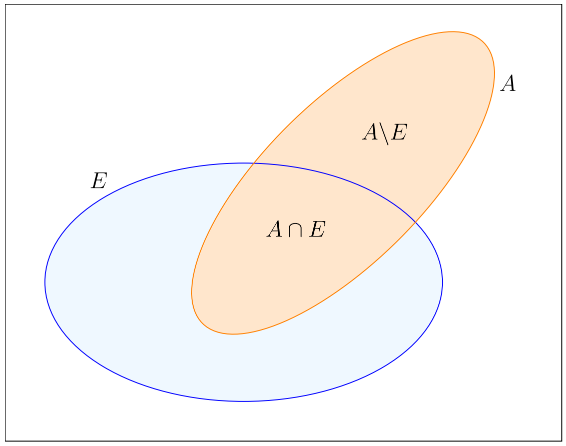

Fig. 2.H A Venn diagram illustrating Carathéodory's Criterion: In order for E to be measurable, $|A \cap E|_e$ and $|A \setminus E|_e$ must sum to $|A|_e$ for *every* set A.

Note: Personally, I feel that the definition of measurability given in Definition 2.2.1 is more intuitive than Carathéodory's Criterion, but I have had a number of discussions with various instructors who feel that Carathéodory's Criterion is more intuitive, and so they prefer to use it as the definition of measurability. I don't think that there is a "correct" view here—both are correct, and it just shows that beauty is in the eye of the beholder.

Note: When we deal with *abstract exterior measures* (covered in the extra online **Chapter 10**, available on the website for this text), there need not be any topology to appeal to, and so in that setting Carathéodory's Criterion is *the* definition of measurable sets.

Theorem 2.2.23 (Carathéodory's Criterion). State and prove.

The following less cluttered argument can be used for the "$\Leftarrow$" direction of the proof if E is bounded. Often, I just present the argument below for bounded sets, and then say that the proof can be extended to unbounded sets by making use of the σ-finiteness of Lebesgue measure, and students should refer to the text for the full proof.

$\Leftarrow$. Assume that E is a *bounded* set that satisfies equation (2.21). Fix $\varepsilon > 0$, and let U be an open set that contains E and satisfies

$$|E|_e \le |U| \le |E|_e + \varepsilon.$$

Using equation (2.21) and the fact that $E = U \cap E$, we compute that

$$|E|_e + |U \setminus E|_e = |U \cap E|_e + |U \setminus E|_e = |U| \le |E|_e + \varepsilon.$$

Since $|E|_e < \infty$, we can subtract it from both sides to obtain $|U \setminus E|_e < \varepsilon$. Therefore E is measurable.

Here is a variation on the preceding proof for bounded sets, using G_δ-sets instead of open sets.

$\Leftarrow$. Assume that E is a *bounded* set that satisfies equation (2.21). By Lemma 2.2.20, there exists a G_δ-set $H \supseteq E$ such that $|H| = |E|_e$. Applying equation (2.21) with $A = H$, we see that

$$|E|_e = |H| = |H \cap E|_e + |H \setminus E|_e = |E|_e + |H \setminus E|_e.$$

Since $|E|_e < \infty$, it follows that $Z = H \setminus E$ has zero exterior measure and hence is measurable. Therefore $E = H \setminus Z$ is measurable as well.

Here is the extension of this G_δ-set approach to arbitrary sets sets E.

$\Leftarrow$. Let E be any subset of $\mathbb{R}^d$ that satisfies equation (2.21). For each $k \in \mathbb{N}$, set

$$E_k = \{x \in E : |x| \le k\},$$

and let $H_k \supseteq E_k$ be a G_δ-set that satisfies $|H_k| = |E_k|_e$. Applying equation (2.21) with $A = H_k$, we have

$$|E_k|_e = |H_k| = |H_k \cap E|_e + |H_k \setminus E|_e \ge |E_k|_e + |H_k \setminus E|_e.$$

Since $|E_k|_e < \infty$, it follows that $Z_k = H_k \setminus E$ has exterior measure zero and hence is measurable. The set $H = \cup H_k$ is measurable, and $Z = H \setminus E$ satisfies

$$Z = H \setminus E = \Bigl(\bigcup_k H_k\Bigr) \setminus E = \bigcup_k (H_k \setminus E) = \bigcup_k Z_k.$$

Therefore Z has measure zero, hence is measurable, so $E = H \setminus Z$ is measurable as well.

2.2.6 Almost Everywhere and the Essential Supremum

Notation 2.2.24 (Almost Everywhere). State.

Example 2.2.25. Discuss.

Definition 2.2.26 (Essential Supremum). State.

Note: I find it amusing that the abbreviation "esssup" has three s's in a row (although a small space is usually inserted between "ess" and "sup"). Clearly I am easily amused.

Example 2.2.27. State.

Lemma 2.2.28 and Corollary 2.2.29. State, but assign the proofs as reading.

Exercise 2.2.30. State.

Problems. *Note on Problem* 2.2.51 in the text: $\Sigma_1 \cap \Sigma_2$ is not formed by intersecting the elements of Σ_1 with those of Σ_2. That is, $\Sigma_1 \cap \Sigma_2$ does not mean $\{A \cap B : A \in \Sigma_1,\, B \in \Sigma_2\}$. Rather, $\Sigma_1 \cap \Sigma_2$ is the collection of sets common to both Σ_1 and Σ_2. In other words,

$$\Sigma_1 \cap \Sigma_2 \,=\, \bigl\{A \subseteq X : A \in \Sigma_1 \text{ and } A \in \Sigma_2\bigr\}.$$

Extra Problems for Section 2.2

1. The *Heaviside function* is $H = \chi_{[0,\infty)}$. Prove that H is continuous at almost every point $x \in \mathbb{R}$, but there is no continuous function g such that $H = g$ a.e.

Note: The first time I encountered the Heaviside function, I thought it was called the "heavy side function," because it is "heavy on one side." Later I learned that it is named after Oliver Heaviside (1850–1925).

2. Using only the results obtained so far in the text, find the Lebesgue measures of the following subsets of $\mathbb{R}^2$.

(a) $A = \{(x,y) : 0 \le x \le 1,\, 0 \le y \le x\}$.

(b) $B = \{(x,y) : 0 \le x \le 1,\, 0 \le y \le x^2\}$.

3. Let $E = \mathbb{Q} \cap [0,1]$, so $|E|_e = 0$. Show that

$$\inf\Biggl\{\sum_{k=1}^{N} |Q_k| : E \subseteq \bigcup_{k=1}^{N} Q_k\Biggr\} \,=\, 1,$$

where we implicitly take the infimum over all *finite* collections of boxes Q_k that cover E. Note that a box in $\mathbb{R}$ is simply a closed finite interval.

Hint: Suppose $Q_1, \dots, Q_N$ cover E, and consider $(0,1)\backslash(Q_1 \cup \cdots \cup Q_N)$.

Note: This shows that we cannot replace *countable* coverings by boxes in the definition of exterior measure with *finite* coverings by boxes.

4. Show that the exterior measure of a set $E \subseteq \mathbb{R}^d$ is unchanged if we replace countable coverings by boxes in the definition of exterior measure with countable coverings by *nonoverlapping* boxes.

5. Assume that $A \subseteq [0,1]$. Prove that A is Lebesgue measurable if and only if $|A|_e + \big|[0,1]\setminus A\big|_e = 1$.

6. (a) *Thomae's function* f is defined as follows. Set $f(x)=0$ if x is irrational, and also set $f(0)=1$. If x is a nonzero rational number, write $x=p/q$ in lowest terms and define $f(x)=1/q$. Show that f is continuous at each irrational point, and discontinuous at each rational point.

(b) Let $\{r_n\}_{n\in\mathbb{N}}$ be an enumeration of the points in $\mathbb{Q}\cap(0,1)$. Define

$$g(x) \,=\, \sum_{n=1}^{\infty} 2^{-n}\,\chi_{(r_n,\infty)}(x), \qquad \text{for } x\in\mathbb{R}.$$

Show that g is monotone increasing on $\mathbb{R}$, continuous at each irrational point $x\in(0,1)$, and discontinuous at each rational point $x\in(0,1)$.

(c) Does there exist a function $f\colon\mathbb{R}\to\mathbb{R}$ that is continuous at each rational point and discontinuous at each irrational?

Hint: Consider Problem 2.2.47.

7. Prove there is no measurable set $E\subseteq\mathbb{R}$ that satisfies $|E\cap(a,b)| = \frac{b-a}{2}$ for all $a<b$.

8. Let E be a measurable subset of $\mathbb{R}^d$ that has finite measure. Suppose that $\{A_k\}_{k\in\mathbb{N}}$ are disjoint measurable subsets of E that all have the same measure. What is $|A_k|$?

9. Suppose that E is a measurable subset of $\mathbb{R}^d$ and $|E\cap(E+t)|=0$ for every nonzero $t\in\mathbb{R}^d$. Prove that $|E|=0$.

10. Let A be any subset of $\mathbb{R}^d$ with $|A|_e<\infty$. Prove that if there exists an F_σ-set $F\subseteq A$ such that $|F|=|A|_e$, then A is measurable.

11. Assume that A and B are measurable subsets of $\mathbb{R}^d$ with finite measure. Prove that $|A\,\triangle\,B|=0$ if and only if $|A\cap B| = \big(|A|+|B|\big)/2$.

12. Prove that statements (a) and (b) of Problem 2.2.38 in the text are also equivalent to the following statement:

(c) For each $\varepsilon>0$ there exists a set S that is a finite union of boxes and satisfies $|E\,\triangle\,S|_e<\varepsilon$, where $E\,\triangle\,S = (E\setminus S)\cup(S\setminus E)$ is the symmetric difference of E and S.

13. Let A be a subset of $\mathbb{R}^d$. Prove that A is measurable if and only if for every $\varepsilon>0$ there exists a measurable set $E\subseteq\mathbb{R}^d$ such that $|A\,\triangle\,E|_e<\varepsilon$.

14.* Show that there exists a measurable set $E\subseteq[0,1]$ with $|E|>0$ that has the property that $|E\cap I|<|I|$ for every open interval $I\subseteq[0,1]$.

Hint: See the sketch of the construction that appears in [Rud83] (W. Rudin, Well-distributed measurable sets, *Amer. Math. Monthly*, **90** (1983), pp. 41–42). The important point is that, by Problem 2.2.42, for each interval $[a,b]$

there exists a "fat Cantor set" $P_{a,b} \subseteq [a,b]$ that is closed, has positive measure, and contains no intervals.

15. Let us say that $f\colon [0,1] \to \mathbb{R}$ is "pretty continuous" if the set of points where f is continuous is a dense subset of $[0,1]$. Show that the sum of two pretty continuous functions is pretty continuous.

Hint: Problem 2.2.46 shows that the set of continuities of f is a G_δ-set. The Baire Category Theorem implies that the intersection of countably many open dense subsets of a complete metric space (such as $[0,1]$) is still dense; see [Heil18, Cor. 2.11.5].

Section 2.3: More Properties of Lebesgue Measure

2.3.1 Continuity from Above and Below

Lemma 2.3.1. State and prove (easy, the main observation is that you can subtract if it doesn't result in an indeterminate form).

Theorem 2.3.2 (Continuity from Below). State and prove. An illustration of nested increasing sets is shown in Figure 2.I.

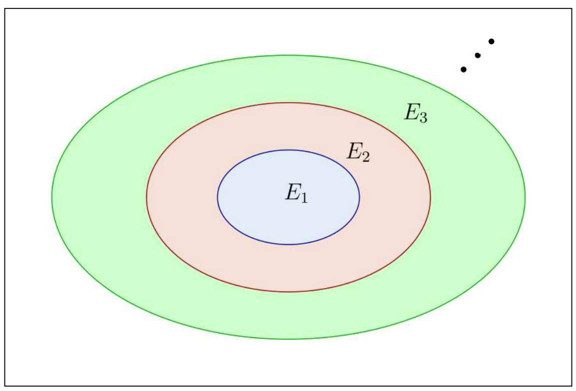

Fig. 2.I Venn diagram showing nested sets $E_1 \subseteq E_2 \subseteq E_3 \subseteq \cdots$. Observe that $\cup E_k$ is the union of the disjoint sets E_1, $E_2 \setminus E_1$, $E_3 \setminus E_2$, ...

Note: Here is an alternative (but essentially similar) proof of Theorem 2.3.2.

Suppose that $E_1 \subseteq E_2 \subseteq \cdots$ are measurable. If we set $E_0 = \varnothing$, then

$$\bigcup_{k=1}^{\infty} E_k \;=\; \bigcup_{j=1}^{\infty} (E_j \setminus E_{j-1}),$$

and the sets on the right-hand side above are disjoint. Therefore, by countable additivity,

$$\begin{aligned}
\left|\bigcup_{k=1}^{\infty} E_k\right| = \left|\bigcup_{j=1}^{\infty} (E_j \setminus E_{j-1})\right| &= \sum_{j=1}^{\infty} |E_j \setminus E_{j-1}| \\
&= \lim_{N\to\infty} \sum_{j=1}^{N} |E_j \setminus E_{j-1}| \\
&= \lim_{N\to\infty} \left|\bigcup_{j=1}^{N} (E_j \setminus E_{j-1})\right| \\
&= \lim_{N\to\infty} |E_N|.
\end{aligned}$$

Example 2.3.3. State.

Note: A one-dimensional example may be easier to visualize, e.g., consider the nested decreasing intervals $E_k = [k, \infty)$ for $k \in \mathbb{N}$.

Theorem 2.3.4 (Continuity from Above). State, but assign the proof for reading—it is similar to the proof for continuity from above.

Corollary 2.3.5. State, assign proof for reading.

2.3.2 Cartesian Products

Exercise 2.3.6 and **Theorem 2.3.7 (Cartesian Products)**. State the theorem, discuss the exercise.

This is a very nice exercise for the students to work through, because it applies continuity from above, the use of G_δ-sets, and other properties of Lebesgue measure.

Depending on your students, you might consider giving some details on why it's not easy to derive the equality $|E \times F| = |E|\,|F|$ directly from the definition of measure. As discussed in the introduction to Section 2.3.2, it is fairly easy to derive the $|E \times F| \le |E|\,|F|$. However, the opposite inequality is not so clear. You might challenge the students to try to find an "easy" proof that $|E \times F| \ge |E|\,|F|$.

The problem is that if $\{Q_k\}_{k\in K}$ is a covering of $E\times F$ by boxes $Q_k \subseteq \mathbb{R}^{m+n}$ for indices k in some countable index set K, then it is *not true* that we can always write

$$\{Q_k\}_{k\in K} = \{R_j \times S_\ell\}_{j\in J,\,\ell\in L}, \qquad \text{(A)}$$

where the R_j are boxes in $\mathbb{R}^m$ and the S_ℓ are boxes in $\mathbb{R}^n$. For a two-dimensional proof by picture, consider the covering of $E \times F = [1,2] \times [0,3]$ by the two boxes $Q_1 = [0,2] \times [0,2]$ and $Q_2 = [1,3] \times [1,3]$. Let $K = \{1,2\}$.

Can we write $\{Q_k\}_{k\in K}$ as $\{R_j \times S_\ell\}_{j\in J,\ell\in L}$? Keep in mind that we must include *every product* $R_j \times S_\ell$ for *all possible choices of* $j \in J$ *and* $\ell \in L$, not just *some* of the choices.

It is true that *every individual box* Q_k in $\mathbb{R}^{m+n}$ is a Cartesian product, so it is true that we can write

$$\{Q_k\}_{k\in K} = \{R_k \times S_k\}_{k\in K}, \qquad \text{(B)}$$

but this doesn't help because it does not have the form given in equation (A). **When equation (A) holds**, we have

$$\begin{aligned}
|E \times F|_e &\le \sum_{k\in\mathbb{N}} \mathrm{vol}(Q_k) \\
&= \sum_{j\in\mathbb{N}}\sum_{\ell\in\mathbb{N}} \mathrm{vol}(R_j \times S_\ell) \\
&= \biggl(\sum_{j\in\mathbb{N}} \mathrm{vol}(R_j) \sum_{\ell\in\mathbb{N}} \mathrm{vol}(S_\ell)\biggr) \\
&= \biggl(\sum_{j\in\mathbb{N}} \mathrm{vol}(R_j)\biggr) \biggl(\sum_{\ell\in\mathbb{N}} \mathrm{vol}(S_\ell)\biggr),
\end{aligned}$$

and then we can take an infimum over all coverings to get $|E\times F|_e \le |E|_e\,|F|_e$ (although we should be careful about cases that involve zero or infinite measure). However, if all that we have is **equation (B)**, then

$$\sum_{k\in\mathbb{N}} \mathrm{vol}(Q_k) = \sum_{k\in\mathbb{N}} \mathrm{vol}(R_k \times S_k) = \sum_{k\in\mathbb{N}} \mathrm{vol}(R_k)\,\mathrm{vol}(S_k),$$

and this is *not equal* in general to

$$\biggl(\sum_{k\in\mathbb{N}} \mathrm{vol}(R_k)\biggr) \biggl(\sum_{k\in\mathbb{N}} \mathrm{vol}(S_k)\biggr),$$

which leaves us stuck.

2.3.3 Linear Changes of Variable

Example 2.3.8. State that it is not true that a continuous function must map measurable sets to measurable sets, and assign this example for reading.

Lemma 2.3.9. State and prove. This gives the missing piece: If a continuous function maps sets with measure zero to sets with measure zero, then it maps measurable sets to measurable sets.

Definition 2.3.10 (Lipschitz Function). State.

Note: The Lipschitz constant is not unique, for if K is a Lipschitz constant then so is any $K' > K$. The smallest Lipschitz constant is sometimes called the optimal Lipschitz constant. The optimal Lipschitz constant defines a seminorm on the space of Lipschitz functions (it is only a seminorm because zero is a Lipschitz constant for any constant function). However, by including an appropriate normalization, the optimal Lipschitz constant can be used to define a norm on the space of Lipschitz functions which makes that space into a Banach space (see Problem 1.4.5).

Lemma 2.3.11. State. Assign the proof for reading.

Note: The easy way to show the inequality that is identified as an "exercise" in the multi-line displayed equation in the proof of this lemma is to use the Cauchy–Schwarz Inequality $|x \cdot y| \le \|x\|\,\|y\|$ for the dot product of vectors in $\mathbb{R}^n$, specifically taking $y = (1, \dots, 1)$.

Exercise 2.3.12. State (note that f need not be linear in this exercise). I usually write out the proof for $d = 1$ since it is so easy: If K is a Lipschitz constant for f, then the image of an interval $[a, b]$ must be contained in an interval of length at most $K(b - a)$.

Suggest that the students determine on their own what goes wrong with the exercise if $f\colon \mathbb{R}^n \to \mathbb{R}^m$ is Lipschitz but $m \ne n$.

Note: **TYPO** in the statement of this exercise: Since we do not know yet that $f(Q)$ is measurable, replace "$|f(Q)|$" with "$|f(Q)|_e$".

Note: We will consider Lipschitz functions in depth in Chapter 5, but you could consider giving a few more details about them now. For example, by using the Mean-Value Theorem we can show that if I is an interval in $\mathbb{R}$ and $f\colon I \to \mathbb{R}$ is a real-valued function such that f is differentiable everywhere on I and f' is bounded on I, then f is Lipschitz (the same is true if f is complex-valued, but the proof is not as straightforward since the MVT need not hold for a complex-valued function, see Problem 1.4.2). However, not all Lipschitz functions are differentiable everywhere, e.g., consider $f(x) = |x|$ on the domain $I = (-1, 1)$.

Note: In Chapter 6 we will explore issues related to the question of when a function maps sets of measure zero to sets of measure zero. We will see that, for functions on an interval $[a, b]$, the *absolutely continuous* functions have this property, and these are precisely the functions that satisfy the Fundamental Theorem of Calculus. All Lipschitz functions are absolutely continuous, but not all absolutely continuous functions are Lipschitz.

Theorem 2.3.13. State and prove.

Note: **TYPO** in the displayed equation in the proof of this theorem. Since we do not know yet that $f(Q_k)$ is measurable, replace "$|f(Q_k)|$" with "$|f(Q_k)|_e$".

Corollary 2.3.14. State.

Note: The case $d = 1$ is trivial, since any linear function $L\colon \mathbb{R} \to \mathbb{R}$ has the form $L(x) = cx$ where c is a constant.

Theorem 2.3.15 (Linear Change of Variables) and Exercise 2.3.16. State the theorem. A proof is given in Exercise 2.3.16 (which is a very nice exercise). It may be worth discussing the exercise briefly, but I wouldn't take the time to write out every step on the board.

The proof sketched in the exercise is based on the singular value decomposition (SVD) of a matrix, so it may be appropriate to briefly discuss the SVD in class The SVD exists for general $m \times n$ matrices, although all we will need here is the SVD for square matrices. For a square matrix, the SVD states that a $d \times d$ matrix can be factored into a product of a rotation or rotation with flip (the orthogonal matrix V^{T}), a dilation (the diagonal matrix Δ, with different axis directions stretched by different amounts), and another rotation or rotation with flip (the orthogonal matrix W). The diagonal part is especially easy to deal with because diagonal matrices map boxes to boxes. The rotations do not map boxes to boxes (because we require a "box" to have sides that are parallel to the coordinate axes), but a rotation does map the unit ball $B_1(0)$ onto itself, which turns out to be very convenient.

I don't believe that I have seen this SVD approach in other texts. Other proofs of Theorem 2.3.15 that I have seen typically factor L into a product of shears. However, this is rather unpleasant because the unit cube $Q_0 = [0,1]^d$ is transformed into a parallelepiped that need not be a box in our sense (because $L(Q_0)$ need not be a *rectangular* parallelepiped, and even if it is, its sides need not be parallel to the coordinate axes). Dealing with such sets is not much fun—just try to give an easy direct proof that the Lebesgue measure of such a set equals its volume in the usual sense.

Note: An orthogonal matrix is also known as a *unitary matrix*, although that terminology is more common when dealind with matrices that have complex entries.

Note: Problem 2.1.35(c) shows that the measure of any proper subspace of $\mathbb{R}^d$ is zero. If $L\colon \mathbb{R}^d \to \mathbb{R}^d$ is a singular (not invertible) linear transformation, then range(L) is a proper subspace of $\mathbb{R}^d$, so Problem 2.1.35(c) can be used to give a simple proof for case where L is singular. However, if that problem has not been worked, then a suitable modification of the steps used in the exercise for the nonsingular case can be made to work for the singular case. Alternatively, once the nonsingular case has been established, then we know that rotations are nonsingular linear transformations that preserve measure, so an arbitrary subspace can be rotated to one that is "parallel to the coordinate axes," and it is easy to show directly from the definition of exterior measure that any proper subspace of that form has measure zero (this is Problem 2.1.36 in the text).

Problems. *Note on Problem* 2.3.19 in the text: This problem is an application of continuity from above and below. However, there is a technicality.

Continuity from above is stated for sequences of sets indexed by the natural numbers, not for sequences indexed by a continuous parameter. Problem 1.1.23 can be used to circumvent this. For example, if we have a set E_t for each real number $t > 0$, then Problem 1.1.23 tells us that

$$\lim_{t \to 0^+} |E_t| = L$$

if and only if for each countable sequence of positive real numbers $t_k \to 0$ we have

$$\lim_{k \to \infty} |E_{t_k}| = L.$$

Continuity from above or below can be applied to each sequence $\{t_k\}_{k \in \mathbb{N}}$ independently. This type of technicality arises at a few other points in the text, and is specifically commented on in Lemma 4.4.9 and Remark 5.2.10.

Extra: Characterizations of Measurability

We collect here most of the characterizations of measurable sets that have been listed in the text.

Theorem. If $E \subseteq \mathbb{R}^d$, then the following statements are equivalent.

(a) E is Lebesgue measurable. That is,

$$\forall\, \varepsilon > 0, \quad \exists \text{ open } U \supseteq E \ \text{ such that } \ |U \setminus E|_e \le \varepsilon.$$

(b) For every $\varepsilon > 0$ there exists a closed set $F \subseteq E$ such that $|E \setminus F|_e < \varepsilon$.

(c) $E = H \setminus Z$ where H is a G_δ-set and $|Z| = 0$.

(d) $E = H \cup Z$ where H is an F_σ-set and $|Z| = 0$.

(e) Carathéodory's Criterion: For every set $A \subseteq \mathbb{R}^d$, we have

$$|A|_e = |A \cap E|_e + |A \setminus E|_e.$$

(f) For every $\varepsilon > 0$ there exists an open set U and a closed set F such that $F \subseteq E \subseteq U$ and $|U \setminus F| < \varepsilon$.

(g) There exists a G_δ-set G and an F_σ-set H such that $H \subseteq E \subseteq G$ and $|G \setminus H| = 0$.

(h) For every box Q we have $|Q| = |Q \cap E|_e + |Q \setminus E|_e$.

If $|E|_e < \infty$, then the statements above are also equivalent to the following statements.

(i) For each $\varepsilon > 0$ we can write $E = (S \cup A) \setminus B$ where S is a union of finitely many nonoverlapping boxes and $|A|_e$, $|B|_e < \varepsilon$.

(j) For each $\varepsilon > 0$ there exists a set S that is a finite union of boxes and satisfies $|E \triangle S|_e < \varepsilon$, where $E \triangle S = (E \setminus S) \cup (S \setminus E)$ is the symmetric difference of E and S.

(k) $|A|_e = |A|_i$, where

$$|A|_i = \sup\bigl\{|F| : F \text{ is closed and } F \subseteq A\bigr\}.$$

Extra Problems for Section 2.3

1. Exhibit a compact subset of $\mathbb{R} \setminus \mathbb{Q}$ that has positive measure.

2. Extra part for Problem 2.3.24.

(g) If $A \subseteq B$ are compact subsets of $\mathbb{R}^d$ and $|A| < t < |B|$, then there exists a compact set K such that $A \subseteq K \subseteq B$ and $|K| = t$.

3. Let A be a measurable subset of $\mathbb{R}^d$ with $|A| < \infty$. Fix $\alpha > 0$, and suppose that $\{E_n\}_{n \in \mathbb{N}}$ is a sequence of measurable subsets of A such that $|E_n| \geq \alpha$ for every n. Let E be the set of all points in $\mathbb{R}^d$ that belong to infinitely many of the E_n. Prove that $|E| \geq \alpha$.

4. Given measurable sets $E_n \subseteq \mathbb{R}^d$ for $n \in \mathbb{N}$, let $\liminf\limits_{n\to\infty} E_n$ and $\limsup\limits_{n\to\infty} E_n$ be as in Exercise 2.1.15.

(a) Prove that

$$\Bigl|\liminf_{n\to\infty} E_n\Bigr| \leq \liminf_{n\to\infty} |E_n|.$$

Show that strict inequality can hold.

(b) Prove that if $\cup E_n$ has finite measure then

$$\Bigl|\limsup_{n\to\infty} E_n\Bigr| \geq \limsup_{n\to\infty} |E_n|.$$

Show that strict inequality can hold, and show that the inequality can fail if $\bigl|\cup E_n\bigr| = \infty$.

5. Assume that $E \subseteq \mathbb{R}^d$ is measurable, with $0 < |E| < \infty$. Set

$$f(t) = |E \cap (E + t)|, \qquad \text{for } t \in \mathbb{R}^d.$$

Show that $f(t) \to 0$ as $\|t\| \to \infty$.

6. This problem gives an alternative approach to proving that Lebesgue measure is invariant under rotations.

(a) Given $E \subseteq \mathbb{R}^d$, define

$$|E|_b = \inf\Bigl\{\sum_k |B_k|\Bigr\},$$

where the infimum is taken over all countable collections of open balls B_k such that $E \subseteq \cup B_k$. Prove that $|E|_b = |E|_e$.

Note: One inequality is easy, the other is not. It is possible to use the Vitali Covering Lemma (Theorem 5.3.3) to prove the opposite inequality, but can you give a direct proof? What about an "easy" direct proof? (I don't know one, but that doesn't mean that a simple proof doesn't exist.)

(b) Prove that Lebesgue measure is invariant under rotations.

7. Let $\mathcal{M}$ be a nonempty collection of subsets of a set X. We say that $\mathcal{M}$ is a *monotone class* if whenever we have sets $E_n, F_n \in \mathcal{M}$ such that $E_1 \subseteq E_2 \subseteq \cdots$ and $F_1 \supseteq F_2 \supseteq \cdots$, then $\cup E_n \in \mathcal{M}$ and $\cap F_n \in \mathcal{M}$. Prove the following statements.

(a) Every σ-algebra on X is a monotone class.

(b) If $\mathcal{A}$ is a nonempty collection of subsets of X, then there is a smallest monotone class $\mathcal{M}$ that contains $\mathcal{A}$.

(c) If $\mathcal{A}$ is a nonempty collection of subsets of X, if $\mathcal{M}$ is the smallest monotone class that contains $\mathcal{A}$, and if Σ is the σ-algebra generated by $\mathcal{A}$, then $\mathcal{A} \subseteq \mathcal{M} \subseteq \Sigma$.

(d) The inclusions in part (c) can be proper.

Section 2.4: Nonmeasurable Sets

2.4.1 The Axiom of Choice

Axioms 2.4.1 or 2.4.2 (Axiom of Choice). I usually state Axiom 2.4.2.

Some discussion may be in order. On the one hand, the statement of the Axiom of Choice is quite reasonable, but on the other hand it also has many surprising implications, such as the non-additivity of Lebesgue exterior measure and the *Banach–Tarski paradox*.

Note: Roughly, Banach–Tarski states the following: It is possible to subdivide the unit ball in $\mathbb{R}^3$ into finitely many disjoint subsets such that after we translate, rotate, and union these sets, we obtain *two* disjoint unit balls.

Note: What happens if we do not accept the Axiom of Choice? Well, other bad things happen, but I leave it to the reader to research this (try doing a web search on "Axiom of Choice").

Note: Some students may not have encountered equivalence relations before, but they are very intuitive. You might give a simple example; my favorite is to take the set of all people and declare that two people are equivalent if they have the same birthday. With this relation, the equivalence class of $x =$ you is $[x] =$ the set of all people that share your birthday. Because of Leap Years, there are 366 distinct equivalence classes for this relation.

2.4.2 Existence of a Nonmeasurable Set

Theorem 2.4.3 (Steinhaus Theorem). State. The proof is interesting and is a nice application of Problem 2.2.39, but in view of time constraints I usually assign the proof as reading.

As mentioned in the text, an elegant way to prove the Steinhaus Theorem is to use the fact that the *convolution* of a bounded function with an integrable function is continuous. In particular, if E is measurable and $0 < |E| < \infty$, then the convolution $f = \chi_E * \chi_{-E}$ is continuous, and its value at the origin is $|E|$, which is strictly positive. Hence f is positive on some interval around 0, and the Steinhaus Theorem follows easily from that (see Problem 4.6.29). If we want to do this without reference to convolution, it amounts to showing that $f(x) = |E \cap (E - x)|$ is a continuous function of x. This approach is spelled out in Problem 2.4.14, although it is not so easy to do this *directly*—that is, using only the results obtained so far in this chapter.

Note: Here is a detailed explanation of the final claim in the proof of Theorem 2.4.3 that $F - F$ contains the interval $\bigl(-\frac{1}{2}|I|,\, \frac{1}{2}|I|\bigr)$.

We have shown that

$$|t| < \frac{1}{2}\,|I| \qquad \Longrightarrow \qquad F \cap (F + t) \neq \varnothing.$$

Now fix any $t \in \bigl(-\frac{1}{2}|I|,\, \frac{1}{2}|I|\bigr)$. Then $F \cap (F + t) \neq \varnothing$, so there exists at least one point x in $F \cap (F + t)$. That is, $x \in F$ and $x \in F + t$, so $x = y + t$ for some $y \in F$. Consequently $t = x - y \in F - F$. This shows that $F - F$ contains the interval $\bigl(-\frac{1}{2}|I|,\, \frac{1}{2}|I|\bigr)$. Since $F \subseteq E$, it follows that $E - E$ must also contain this interval.

Theorem 2.4.4. State and prove.

2.4.3 Further Results

Theorem 2.4.5. Since this is the result that opened the chapter, I state it. The proof is interesting, but I assign it as reading.

Note: Here are the details of the claim made in the proof of Theorem 2.4.5 that

$$[0, 1) \subseteq \bigcup_{k=1}^{\infty} (M + r_k) \subseteq [-1, 2].$$

The second inclusion is easy because $M \subseteq [0,1)$ and each scalar r_k belongs to $[-1,1]$. To prove the first inclusion, choose any point $x \in [0,1)$. Then x belongs to some equivalence class of the relation $\sim$, so there exists some point $y \in M$ such that $x \sim y$. Hence $x = y + r$ where r is rational, and since both x and y belong to $[0,1)$, we must have $r \in [-1,1]$. Hence $r = r_k$ for some k, and therefore $x \in M + r_k$ for that k.

Corollary 2.4.6. Omit. Note, however, that there is a very surprising consequence—we state this below after a remark and the boxed proof of that remark.

Note: Here are the details of the claim made in the proof of Corollary 2.4.6 that the set M is nonmeasurable.

Let M be the set constructed in the proof of Theorem 2.4.5, and let $\{r_k\}_{k\in\mathbb{N}}$ be an enumeration of $\mathbb{Q} \cap [-1,1]$. Then, just as in the proof of Theorem 2.4.5, we have that the sets $M_k = M + r_k$ are disjoint and satisfy

$$[0,1) \subseteq \bigcup_{k=1}^{\infty} M_k \subseteq [-1,2). \qquad \text{(A)}$$

That is, we have infinitely many disjoint sets M_k that all have the same exterior Lebesgue measure (because exterior Lebesgue measure is translation-invariant), and these sets are contained in $[-1,2)$, which is a set that has finite measure. If M was measurable, then each M_k would be measurable, and therefore we could apply countable additivity and translation-invariance and conclude that

$$\biggl|\bigcup_{k=1}^{\infty} M_k\biggr| = \sum_{k=1}^{\infty} |M_k|_e = \sum_{k=1}^{\infty} |M|_e = \begin{cases} 0, & \text{if } |M|_e = 0, \\ \infty, & \text{if } |M|_e > 0. \end{cases}$$

But equation (A) implies that

$$1 \le \biggl|\bigcup_{k=1}^{\infty} M_k\biggr| \le 3,$$

so this is a contradiction. Therefore M cannot be measurable.

Note that we have shown something very counterintuitive in this argument. Since M is nonmeasurable, it cannot have measure zero. We don't know exactly what $|M|_e$ is, but it is some strictly positive, finite real number. Since exterior Lebesgue measure is translation-invariant, we have $|M_k|_e = |M|_e$ for every k. Therefore:

the finite interval $[-1,2)$ *contains infinitely many disjoint sets* M_k *that each have exactly the same strictly positive exterior measure*!

Example 2.4.7. Discuss (if it has not been done already).

Extra Problems for Section 2.4

1. Suppose that $E \subseteq A \subseteq F \subseteq \mathbb{R}^d$, where E and F are measurable and $|E| = |F| < \infty$. Prove that A is measurable. Show by example that this can fail if $|E| = |F| = \infty$.

2. Let $\mathcal{L}$ be the set of all Lebesgue measurable subsets of $\mathbb{R}^d$, and let $\mathcal{N} = \mathcal{P}(\mathbb{R}^d)\backslash\mathcal{L}$ be the set of all nonmeasurable subsets of $\mathbb{R}^d$. For each of the following statements, either prove that the statement is true or exhibit a counterexample.

(a) If $N \in \mathcal{N}$, then $N^{\mathrm{C}} \in \mathcal{N}$.

(b) If $M, N \in \mathcal{N}$, then $M \cap N \in \mathcal{N}$.

(c) If $M, N \in \mathcal{N}$, then $M \cup N \in \mathcal{N}$.

(d) If $E \in \mathcal{L}$ and $N \in \mathcal{N}$, then $E \cap N \in \mathcal{N}$.

(e) If $E \in \mathcal{L}$ and $N \in \mathcal{N}$, then $E \cup N \in \mathcal{N}$.

(f) If $E \in \mathcal{L}$ and $N \in \mathcal{N}$ are disjoint, then $E \cup N \in \mathcal{N}$.

3. Prove the following higher-dimensional version of the Steinhaus Theorem: If $E \subseteq \mathbb{R}^d$ is Lebesgue measurable and $|E| > 0$, then the set of differences

$$E - E = \{x - y : x, y \in E\}$$

contains an open ball $B_r(0)$ for some $r > 0$.

Note: This problem is helpful for solving Problem 2.4.9 in the main text.

CHAPTER 3: MEASURABLE FUNCTIONS

Section 3.1: Measurable Functions

Notation 3.1.1 (Scalars and the Symbol $\overline{\mathbf{F}}$). State. The class of extended real-valued functions is not a subset of the class of complex-valued functions, nor conversely, so these are two different cases that we encounter frequently. Using the symbol $\overline{\mathbf{F}}$ to stand for a choice of either $[-\infty, \infty]$ or $\mathbb{C}$ will allow for some flexibility and conciseness, although there will still be times where we will want to be specific about the type of functions we are dealing with. Note that even if we take $\overline{\mathbf{F}} = [-\infty, \infty]$, the word *scalar* always refers to an actual *number*; $\pm\infty$ are not scalars.

Remark 3.1.2. State.

3.1.1 Extended Real-Valued Functions

Definition 3.1.3 (Extended Real-Valued Measurable Functions). State.

Note: When writing on the board, a standard abbreviation for "measurable" is ⓜ, the letter "m" with a circle around it.

Example 3.1.4. Omit.

Lemma 3.1.5. State, assign proof for reading.

TYPO in the text: We must assume in Lemma 3.1.5 that the domain E is a measurable set.

Note: A corollary of Lemma 3.1.5 is that if f is a measurable function, then $\{f = a\} = \{f \geq a\} \cap \{f \leq a\}$ is a measurable set for every $a \in \mathbb{R}$. However, the converse fails. That is, there exists a function f such that $\{f = a\}$ is measurable for every a yet f is nonmeasurable (see Problem 3.1.17).

Lemma 3.1.6 and the remarks that precede it. State, prove, and discuss.

Lemma 3.1.7. State and prove; this is illustrated in Figure 3.A below.

Corollary 3.1.8. State.

Remark 3.1.9. Briefly discuss. One example is $f(x) = 1/x$ on $[-1, 1]$. If we don't care about sets of measure zero, then we can simply leave $f(0)$ undefined. Regardless of whether we leave $f(0)$ undefined or we assign it a value, f is measurable.

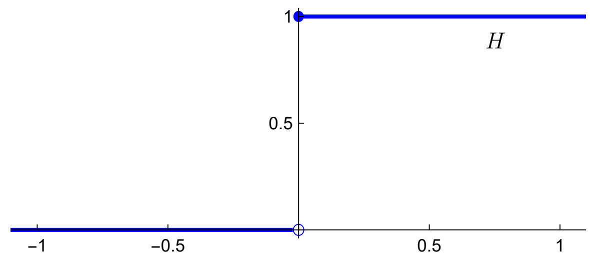

Fig. 3.A The Heaviside function $H = \chi_{[0,\infty)}$. There is no continuous function that equals H at almost every point.

Definition 3.1.10 (Positive and Negative Parts). State. We will encounter f^+ and f^- many times in the remainder of the text (see the illustration in Figure 3.B below).

Note: $f^+(x) - f^-(x)$ can never take an indeterminate form. It can be $\infty - 0$ or $0 - \infty$, but it can never be $\infty - \infty$.

3.1.2 Complex-Valued Functions

Definition 3.1.11 (Complex-Valued Measurable Functions). State.

Note: By Problem 3.1.18, a function $f\colon \mathbb{R}^d \to \mathbb{C}$ is measurable if and only if $f^{-1}(U)$ is measurable for every open $U \subseteq \mathbb{C}$. In this sense measurability is a generalization of continuity.

Lemmas 3.1.12 and 3.1.13. Omit, or a brief remark is enough.

Extra Problems for Section 3.1

1. Let E be a measurable subset of $\mathbb{R}^d$, let $f\colon E \to \overline{\mathbf{F}}$ be given, and let A be a dense subset of $\mathbb{R}$. Prove that f is measurable if and only if $\{f > a\}$ is measurable for every $a \in A$.

2. Show that there exists a nonmeasurable function $f\colon \mathbb{R} \to \mathbb{R}$ such that $|f|$ is measurable on $\mathbb{R}$.

3. Given a point $x \in [0,1]$ with decimal expansion $x = 0.d_1d_2d_3\dots$, set

$$f(x) \;=\; \max_{k\in\mathbb{N}} d_k.$$

Every point in $[0,1]$, except for certain rational points, has a unique decimal expansion, so this uniquely defines $f(x)$ for all but countably many x (specifically, those rational points that have decimal expansions ending in infinitely many 0's or infinitely many 9's). Prove that f is measurable on $[0,1]$ and f is constant a.e. (that is, there exists some scalar $c \in \mathbb{R}$ such that $f(x) = c$ for a.e. $x \in [0,1]$).

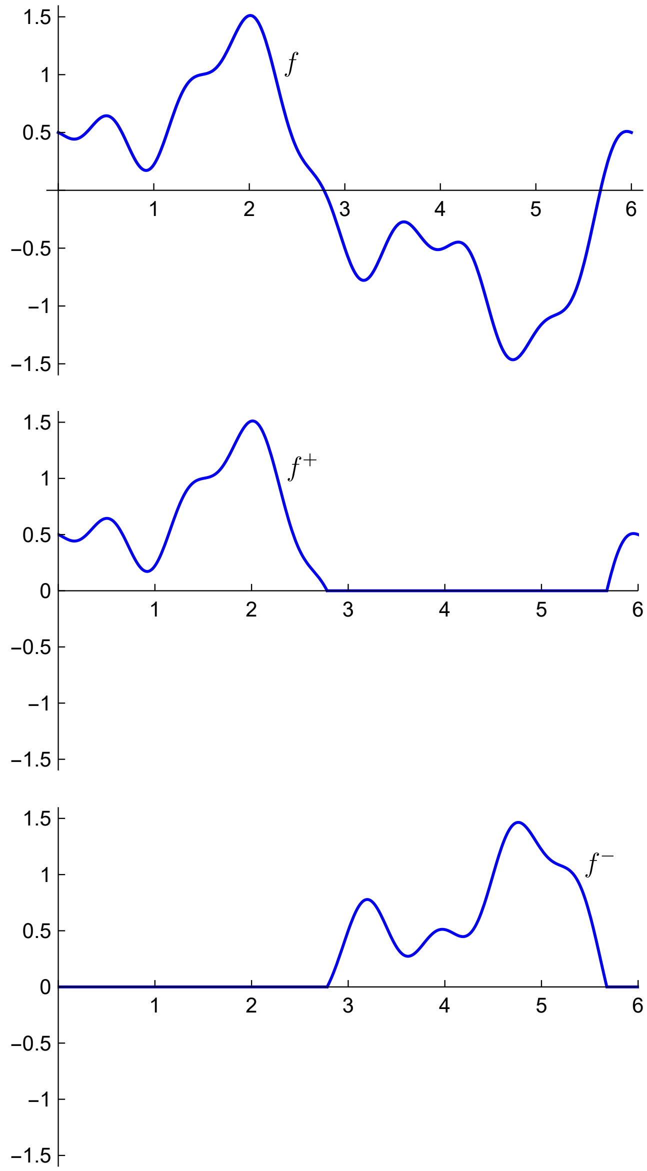

Fig. 3.B A function f (top), its positive part f^+ (middle), and its negative part f^- (bottom).

4. Prove that every upper semicontinuous (usc) function $f\colon \mathbb{R}^d \to \mathbb{R}$ is measurable.

5. Assume $E \subseteq \mathbb{R}^d$ is measurable, and suppose that $\{f_n\}_{n\in\mathbb{N}}$ is a sequence of measurable functions on E such that

$$\sum_{n=1}^{\infty} \bigl|\{|f_n| > 1\}\bigr| \,<\, \infty.$$

Prove that $\limsup_{n\to\infty} |f_n(x)| \le 1$ for a.e. $x \in E$.

6. Let $E \subseteq \mathbb{R}^d$ be measurable, and assume that $\{f_n\}_{n\in\mathbb{N}}$ is a sequence of measurable functions on E such that

$$\sum_{n=1}^{\infty} \bigl|\bigl\{|n^2 f_n| > 1\bigr\}\bigr| \,<\, \infty.$$

Prove that $\sum f_n(x)$ converges for almost every $x \in E$. Hint: Borel–Cantelli.

7. (From Stein and Shakarchi). Let E be a measurable subset of $\mathbb{R}^d$ with $|E| < \infty$. Suppose that functions f_n are measurable and finite a.e. on E. Show that there exist positive numbers $c_n > 0$ such that $f_n/c_n \to 0$ a.e.

Hint: Choose c_n such that $|E_n| < 2^{-n}$, where $E_n = \bigl\{|f|/c_n > 1/n\bigr\}$, and apply Borel–Cantelli.

Section 3.2: Operations on Functions

3.2.1 Sums and Products

Lemma 3.2.1. State and prove.

Lemma 3.2.2. State, leave proof for reading.

Note: Lemma 3.2.2 is what we will use most often when dealing with the sum of two functions. However, although it is not stated in the text, the following result is actually needed in small number of places in the text, most notably in the proof of Egorov's Theorem (Theorem 3.4.2). A related result can be found in Problem 3.2.16.

Theorem 3.2.A. *Let $E \subseteq \mathbb{R}$ be a Lebesgue measurable set and assume that $f, g\colon E \to [-\infty, \infty]$ are measurable functions such that at least one of f or g is finite a.e. Then $f+g$ and $f-g$ are measurable functions.*

Proof. Assume that f is measurable, and g is both measurable and finite a.e. (the case where f is finite a.e. being entirely symmetric). Define

$$A^+ \,=\, \{f = \infty\}, \qquad A^- \,=\, \{f = -\infty\}, \qquad Z \,=\, \{g = \pm\infty\}.$$

The quantity $f(x) + g(x)$ can only take the form $\infty - \infty$ or $-\infty + \infty$ when x is in Z, which is a set of measure zero. Therefore $f(x) + g(x)$ is defined almost everywhere.

Each of A^+, A^-, and Z are measurable sets. Also, A^+ and A^- are disjoint, and Z has measure zero. If we set

$$B = \bigl(A^+ \cup A^- \cup Z\bigr)^{\mathrm{C}},$$

then B is measurable and $f(x)$ and $g(x)$ are both finite for $x \in B$. Consequently

$$f_1 = f \cdot \chi_B \qquad \text{and} \qquad g_1 = g \cdot \chi_B$$

are measurable, and they are finite at *every point* $x \in E$. Lemma 3.2.1 therefore implies that $f_1 + g_1$ is measurable. Now define

$$h(x) = \begin{cases} f_1(x) + g_1(x), & \text{if } t \in B \cup Z, \\ \infty, & \text{if } x \in A^+ \backslash Z, \\ -\infty, & \text{if } x \in A^- \backslash Z. \end{cases}$$

If we fix $a \in \mathbb{R}$, then $h(x) > a$ when $x \in B \cup Z$ and $f_1(x) + g_1(x) > a$, or when $x \in A^+ \backslash Z$. Hence

$$\{h > a\} = \Bigl(\{f_1 + g_1 > a\} \cap (B \cup Z)\Bigr) \cup (A^+ \backslash Z).$$

Since A^+, B, and Z are measurable sets and $f_1 + g_1$ is a measurable function, it follows that $\{h > a\}$ is a measurable set. Since this is true for every $a \in \mathbb{R}$, we conclude that h is measurable. We will show that $f(x) + g(x) = h(x)$ for almost every $x \in E$.

Case 1. If $x \in B$ then $f_1(x) = f(x)$ and $g_1(x) = g(x)$, while the definition of h tells us that $h(x) = f_1(x) + g_1(x)$. Consequently $h(x) = f(x) + g(x)$ for $x \in B$.

Case 2. If $x \in A^+ \backslash Z$ then $f(x) = \infty$ while $g(x)$ is finite and $h(x) = \infty$. Therefore $h(x) = f(x) + g(x)$ for $x \in A^+ \backslash Z$.

Case 3. If $x \in A^- \backslash Z$ then $f(x) = -\infty$ while $g(x)$ is finite and $h(x) = -\infty$. Therefore $h(x) = f(x) + g(x)$ for $x \in A^- \backslash Z$.

The combination of Cases 1, 2, and 3 implies that $h(x) = f(x) + g(x)$ for all x except possibly those in Z (see the Venn diagram in Figure 3.C). But Z has measure zero, so have shown that $h = f + g$ a.e. Since h is measurable, Lemma 3.1.7 therefore implies that $f + g$ is measurable as well. Finally, by replacing g with $-g$, wee see that $f - g = f + (-g)$ is also measurable. □

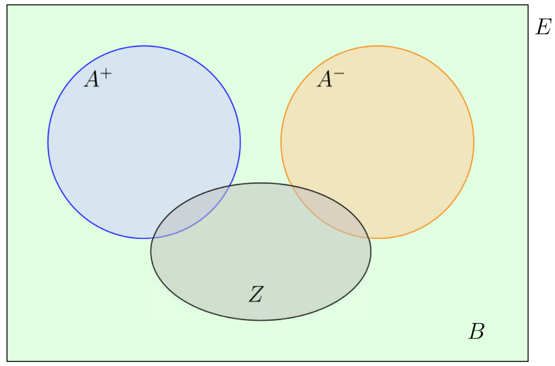

Fig. 3.C Venn diagram: The sets A^+ (blue) and A^- (orange) are disjoint, Z (gray) has measure zero, and B (green) is the complement of $A^+ \cup A^- \cup Z$.

Lemma 3.2.3. State and prove.

Lemma 3.2.4. Just remark that measurability is preserved under quotients if division by zero is avoided, leave the details for reading.

Note: A minor technical point in the proof is that, as long as it does not take an indeterminate form, $f(x)/g(x)$ is the same as $f(x) \cdot 1/g(x)$. However, while $f(x)/g(x)$ is undefined at any points where both $f(x)$ and $g(x)$ are $\pm\infty$, the quantity $f(x) \cdot 1/g(x)$ is defined, in the extended real sense, at every point where $g(x) \neq 0$.

3.2.2 Compositions

Lemma 3.2.5. State and prove.

Lemma 3.2.6. Omit, or state but leave the proof for reading.

3.2.3 Suprema and Limits

Lemma 3.2.7. State. Prove that $\sup_{n \in \mathbb{N}} f_n$ is measurable, assign the other parts for reading.

Note: A **MODIFICATION** is needed in the statement of this Lemma. The lemma remains valid if the hypotheses that the functions f_n are finite a.e. is removed. This is needed in a few places in the text, such as to justify the measurability of functions that appear in Theorems 4.2.1, 4.2.7, 4.3.7, and 4.3.8.

Notation 3.2.8. State.

Exercise 3.2.9. State. A picture may be helpful.

Exercise 3.2.10. Just say that the results for sums, products, etc., in the complex case are similar to those for the extended-real case.

3.2.4 Simple Functions

Definition 3.2.11 (Simple Function). Motivate and state.

Lemma 3.2.12. State, there's really not much to say about the proof except "by inspection."

Definition 3.2.13 (Standard Representation). State.

Theorem 3.2.14. State and discuss. I usually draw the picture and explain the idea, but do not write down the explicit formula for the functions f_n that is given in equation (3.3).

Here is a more detailed discussion, and Figure 3.D below gives a color illustration that improves on Figure 3.1 in the text.

Let $f\colon E \to [0,\infty]$ be an arbitrary measurable, nonnegative function on a measurable domain E. The function f could actually take the value ∞ at some points, but even if it is finite at every point it could still be unbounded (for example, consider $f(t) = 1/t$ on the domain $E = (0,1)$). We want to define $\phi_n(t)$ by taking $f(t)$ and rounding it down to the nearest integer multiple of 2^{-n}. However, this could end up giving ϕ_n infinitely many different values, in which case ϕ_n would not be a simple function. Therefore we modify this idea by stopping the rounding-down process at some finite height (though we let that height increase with n). For example, we could define ϕ_n by

$$\phi_n(t) \;=\; \begin{cases} \dfrac{j}{2^n}, & \text{if } f(t) \le n \text{ and } \dfrac{j}{2^n} \le f(t) < \dfrac{j+1}{2^n} \text{ (where } j \ge 0\text{)},\\[2ex] n, & \text{if } f(t) > n. \end{cases}$$

Figure 3.1 illustrates ϕ_1 and ϕ_2 for a particular function f.

We consider some possibilities.

- **Case 1: $f(t)$ is finite at a point** $t \in E$. In this case $f(t)$ is a nonnegative real number, and for those nonnegative integers n that satisfy $n \le f(t)$ we have $\phi(t) = n$. However, we have $n \le f(t)$ for only finitely many n. For integers $n > f(t)$ we obtain $\phi_n(t)$ by rounding $f(t)$ down to the nearest integer multiple of 2^{-n}. Hence, if $n > f(t)$ then the difference between $f(t)$ and $\phi_n(t)$ is no more than 2^{-n}, so $\phi_n(t)$ converges to $f(t)$ as $n \to \infty$. Further,

$$\phi_0(t) \;\le\; \phi_1(t) \;\le\; \cdots,$$

  so ϕ_n increases monotonically to f.

- **Case 2: $f(t) = \infty$ at a point** $t \in E$. In this case $\phi_n(t) = n$ for every n, so again $\phi_n(t)$ increases monotonically to $f(t)$ as n increases.

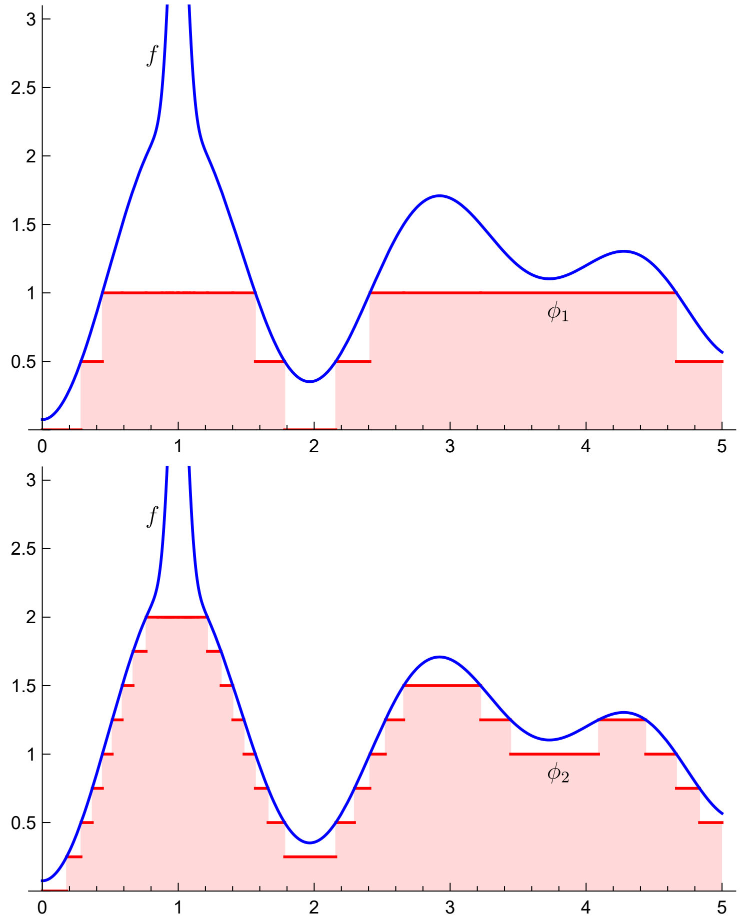

Fig. 3.D Illustration of a function f (blue) and approximating simple functions ϕ_1 and ϕ_2 (red). The function ϕ_1 rounds f down to the nearest integer multiple of $1/2$, but stops the rounding process at height 1. The function ϕ_2 rounds f down to the nearest integer multiple of $1/4$, but stops the rounding process at height 2. The function f pictured here is piecewise continuous, but in general it could be any nonnegative measurable function.

- **Case 3: f is bounded on a subset** A. In general, ϕ_n need not converge uniformly to f on all of E. However, suppose f is bounded on some subset $A \subseteq E$; say $0 \le f(t) \le M$ for $t \in A$. Then for any $n > M$, we have for all $t \in A$ that $\phi_n(t)$ differs from $f(t)$ by no more than 2^{-n}, and therefore

$$\sup_{t \in A} |f(t) - \phi_n(t)| \le 2^{-n}, \qquad \text{for all } n > M.$$

This implies that ϕ_n converges uniformly to f *on the subset* A.

Note: We could just as well use 2^n or any other increasing sequence instead of n as the cutoff heights for the functions ϕ_n.

Note: Explicitly,

$$\phi_1(x) = \begin{cases} 0, & \text{if } 0 \le f(x) < 1, \\ 1, & \text{if } f(x) \ge 1. \end{cases}$$

For ϕ_2 we round down to the nearest integer multiple of $\frac{1}{2}$, except we never exceed height 2, so

$$\phi_2(x) = \begin{cases} 0, & \text{if } 0 \le f(x) < \frac{1}{2}, \\ \frac{1}{2}, & \text{if } \frac{1}{2} \le f(x) < 1, \\ 1, & \text{if } 1 \le f(x) < \frac{3}{2}, \\ \frac{3}{2}, & \text{if } \frac{3}{2} \le f(x) < 2, \\ 2, & \text{if } f(x) \ge 2. \end{cases}$$

Note that if $f(x) \le 2$, then $f(x)$ and $\phi_2(x)$ differ by at most $\frac{1}{2}$ units.

Corollary 3.2.15. Assign for reading.

Extra Problems for Section 3.2

1. Let $E \subseteq \mathbb{R}$ be measurable, and assume that $f\colon E \to [-\infty, \infty]$ is measurable. Prove directly that

$$\alpha(x) = \operatorname{sign}\bigl(f(x)\bigr) = \begin{cases} 1, & \text{if } f(x) > 0, \\ 0, & \text{if } f(x) = 0, \\ -1, & \text{if } f(x) < 0, \end{cases}$$

is measurable on E.

2. Given the same hypotheses as Lemma 3.2.7, prove the following extensions of parts (b) and (c) of that lemma. Note that *converges* means that *the limit exists and is a scalar* (so it specifically excludes the case of divergence to $\pm\infty$).

(b') If $f(x) = \lim\limits_{n\to\infty} f_n(x)$ converges for a.e. $x \in E$, then f is measurable.

(c') If $f(x) = \sum\limits_{n=1}^{\infty} f_n(x)$ converges for a.e. $x \in E$, then f is measurable.

3. Let $\{f_n\}_{n\in\mathbb{N}}$ be a sequence of measurable, real-valued or complex-valued functions whose domain is a measurable set $E \subseteq \mathbb{R}^d$. Show that the set

$$L = \Big\{x \in E : \lim_{n\to\infty} f_n(x) \text{ converges to a scalar}\Big\}$$

is equal to

$$\bigcap_{k=1}^{\infty} \bigcup_{N=1}^{\infty} \bigcap_{n=N}^{\infty} \bigcap_{m=N}^{\infty} \big\{|f_n - f_m| \le \tfrac{1}{k}\big\},$$

and hence is a measurable subset of E.

4. Assume that $f\colon \mathbb{R} \to \mathbb{R}$ is measurable, and prove that

$$B = \big\{(x,y) \in \mathbb{R}^2 : f(x) \ge f(y)\big\}$$

is a measurable subset of $\mathbb{R}^2$.

5. (This problem is related to the definition of the convolution of functions, and it is discussed explicitly in the text in Section 4.6.3. However, all of the tools needed to solve it are available now.)

Suppose that f, $g\colon \mathbb{R}^d \to \overline{\mathbf{F}}$ are measurable functions. Prove that the functions F, $G\colon \mathbb{R}^{2d} \to \overline{\mathbf{F}}$ given by

$$F(x,y) = f(x)\,g(y) \qquad \text{and} \qquad G(x,y) = f(y)\,g(x-y)$$

are measurable on $\mathbb{R}^{2d}$.

6. (From Benedetto and Czaja. This is a nice application of the Steinhaus Theorem.) Assume that $f\colon \mathbb{R} \to \mathbb{R}$ is measurable, $g\colon \mathbb{R}^2 \to \mathbb{R}$ is continuous, and

$$\forall\, x, y \in \mathbb{R}, \quad |f(x+y)| \le g\big(f(x), f(y)\big).$$

Prove the following statements.

(a) There exists a measurable set $E \subseteq \mathbb{R}$ with positive measure such that

$$\sup_{x,y\in E} |f(x-y)| < \infty.$$

(b) There exists a $\delta > 0$ such that f is bounded on $(-\delta, \delta)$.

(c) If n is a positive integer, then f is bounded on $(-n\delta, n\delta)$.

(d) f maps bounded sets to bounded sets, i.e., if $A \subseteq \mathbb{R}$ is bounded, then $f(A)$ is bounded.

7. This problem uses the definition of *Borel sets* introduced in Problem 2.3.25.

(a) Let $f\colon \mathbb{R}^d \to \mathbb{R}$ be Lebesgue measurable. Suppose that $g\colon \mathbb{R} \to \mathbb{R}$ is *Borel measurable*, i.e., $g^{-1}(U)$ is a Borel set for every open $U \subseteq \mathbb{R}$. Show that the composition $g \circ f$ is Lebesgue measurable. Generalize to the case of complex-valued f, g.

(b) Define a square root function $Sz = z^{1/2}$ on $\mathbb{C}$ by

$$S(re^{i\theta})^{1/2} = r^{1/2} e^{i\theta/2}, \qquad \text{for } r > 0,\ 0 \le \theta < 2\pi.$$

Show that S is Borel measurable, and conclude that if f is measurable on $\mathbb{R}^d$ then so is $f^{1/2} = Sf$.

8. Assume that $f\colon \mathbb{R}^2 \to \mathbb{R}$ is continuous, and $g, h\colon [a,b] \to \mathbb{R}$ are measurable. Set $F(x) = f\big(g(x), h(x)\big)$, and prove that $F\colon [a,b] \to \mathbb{R}$ is measurable.

Section 3.3 The Lebesgue Space $L^\infty(E)$

Depending on your audience, at this point you may wish to review in more or less detail some of the concepts related to norms, convergence, and completeness that are covered in Section 1.2. Alternatively, some instructors may prefer to delay all discussion of the L^∞-norm or the L^1-norm until Chapter 7, which discusses the L^p-norms for $1 \le p \le \infty$ in detail.

Remark 3.3.1 and comments preceding. Define the L^∞-norm and the uniform norm.

Note: A common mistake is to assume that "finite a.e." and "essentially bounded" are synonyms. They are not; for example, $f(x) = 1/x$ is finite at almost point of $\mathbb{R}$, but it is not essentially bounded.

Lemma 3.3.2. State; this is a nice contrast between the L^∞-norm and the uniform norm (for the uniform norm, $|f(x)| \le \|f\|_{\mathrm{u}}$ for every x, while for $\|\cdot\|_\infty$ this only holds for *almost every* x).

Definition 3.3.3 (The Lebesgue Space $L^\infty(E)$). State.

Exercise 3.3.4. State. I usually briefly discuss the fact that the "L^∞-norm" is only a seminorm, and briefly mention the fact that we usually just "identify" functions that are equal a.e., in which case it becomes a norm. (This identification of functions that are equal a.e. will be discussed in detail when we present the L^p spaces in Section 7.2.2.)

3.3.1 Convergence and Completeness in $L^\infty(E)$

Definition 3.3.5 (Convergence in L^∞-Norm). State.

Remark 3.3.6. Omit.

Lemma 3.3.7. State, and perhaps sketch the proof. Use this opportunity to briefly discuss completeness and why it is important.

Note: I usually do not present this in class, but if time and interest permit, a discussion of the fact that $L^\infty(\mathbb{R})$ is nonseparable could fit in here. This is discussed in detail in the text in Section 7.4, and Problem 3.3.9 presents the relevant fact that $L^\infty(\mathbb{R})$ contains an uncountable collection of functions in which each element is separated by one unit from every other element.

Extra Problems for Section 3.3

1. Let E be any measurable subset of $\mathbb{R}^d$ such that $|E| > 0$. Exhibit a measurable function $f\colon E \to \mathbb{R}$ that is finite a.e. but does not belong to $L^\infty(E)$.

2. Let $E \subseteq \mathbb{R}^d$ be measurable with $|E| > 0$. Exhibit a set $\{f_i\}_{i\in I}$ of uncountably many functions in $L^\infty(E)$ such that $\|f_i - f_j\|_\infty = 1$ whenever $i \neq j$.

Section 3.4: Egorov's Theorem

Example 3.4.1 (Shrinking Triangles). Discuss. This is an important "standard counterexample."

Theorem 3.4.2 (Egorov's Theorem). State and prove. (This proof is adapted from Wheeden and Zygmund.)

***Note*:** Technically, there is an **omission** in the proof of Theorem 3.4.2, because the measurability of the difference $f - f_n$ is not justified. In particular, although f and f_n are both required to be measurable, only f is assumed to be finite a.e. Consequently Lemma 3.2.2 is not applicable (nor is Problem 3.2.16). However, Theorem 3.2.A (from earlier in this Instructor's Guide) shows that $f - g$ is measurable when f and g are measurable and at least one is finite a.e. Therefore that result justifies why $f - f_n$ is measurable in this proof.

Note: "Egorov" is a transliteration from the Russian. Another common English spelling is "Egoroff." Egorov's Theorem is named for the Russian mathematician Dmitri Egorov (1869–1931). The theorem was also proved independently by Carlo Severini (1872–1951).

Note: In the proof, if $x \in Z$ then all that we know is that $f_n(x) \not\to f(x)$. However, if $x \in Z_k$ then we have much more *quantitative* information, because

$$\begin{aligned} Z_k &= \limsup_{n\to\infty} \Big\{|f - f_n| \ge \tfrac{1}{k}\Big\} \\ &= \Big\{x \in E : |f(x) - f_n(x)| \ge \tfrac{1}{k} \text{ for infinitely many } n\Big\}. \end{aligned}$$

Thus, if $x \in Z_k$ then we have knowledge about the actual distance between $f(x)$ and $f_n(x)$, not for all n but at least for infinitely many n. Since $|f(x) - f_n(x)| \ge \frac{1}{k}$ for infinitely many n we certainly have $f_n(x) \not\to f(x)$, and therefore $Z_k \subseteq Z$, but knowing that $x \in Z_k$ gives more *precise* information than just knowing that $x \in Z$.

Note: For several years after receiving my Ph.D. I never found much use for Egorov's Theorem, then suddenly it seemed that I needed it in one paper after another. I guess this just goes to show that you never know when knowledge will be useful.

Note: We use the uniform norm in the statement of part (b) instead of the L^∞-norm because nothing is gained by using $\|\cdot\|_\infty$. In particular, suppose that A is such that

$$\lim_{n\to\infty}\left(\operatorname*{ess\,sup}_{x\notin A}|f(x)-f_n(x)|\right) \;=\; 0.$$

Then the sets

$$W_n \;=\; \bigl\{|f-f_n| > \|f-f_n\|_\infty\bigr\}, \qquad \text{for } n\in\mathbb{N},$$

each have measure zero, the set

$$B \;=\; A \,\cup\, \left(\bigcup_{n=1}^{\infty} W_n\right)$$

has exactly the same measure as A, and $f_n \to f$ uniformly on $E\setminus B$. So instead of just getting that $f_n \to f$ in L^∞-norm on $E\setminus A$, we get $f_n \to f$ uniformly on $E\setminus B$, and the two sets A and B only differ by a set of measure zero.

Note: In the proof of Egorov's Theorem, we say that $\bigl|A_{n_k}(k)\bigr| < \varepsilon/2^k$, and therefore $A = \bigcup_{k=1}^\infty A_{n_k}(k)$ has measure $|A| < \varepsilon$. One way to see why we get stictly less than here instead of just less than or equal to is by splitting off the first term in the computation:

$$\begin{aligned}
|A| \;=\; \left|\bigcup_{k=1}^{\infty} A_{n_k}(k)\right| \;&=\; \left|A_{n_1}(1) \;\cup\; \bigcup_{k=2}^{\infty} A_{n_k}(k)\right| \\
&\le\; |A_{n_1}(1)| \;+\; \left|\bigcup_{k=2}^{\infty} A_{n_k}(k)\right| \\
&\le\; |A_{n_1}(1)| \;+\; \sum_{k=2}^{\infty} |A_{n_k}(k)| \\
&<\; \frac{\varepsilon}{2} \;+\; \sum_{k=2}^{\infty} |A_{n_k}(k)| \;\le\; \frac{\varepsilon}{2} \;+\; \sum_{k=2}^{\infty} \frac{\varepsilon}{2^k} \;=\; \varepsilon.
\end{aligned}$$

Definition 3.4.3 (Almost Uniform Convergence). State briefly.

Note: Observe that "almost" in this definition does not mean "except for a set of measure zero," as in does in the definition of "almost everywhere." Instead, here it means "except for a set of measure ε." I think it would be advisable to change the name, say to *nearly uniform convergence*. In fact, I've adopted this terminology in my more recent books.

Exercise 3.4.4. State briefly.

Note: **TYPO** in this exercise. Add the hypothesis that f and f_n are finite a.e.

Extra Problems for Section 3.4

1. Suppose that there were a norm $\|\cdot\|$ on $C[0,1]$ such that $\|f - f_n\| \to 0$ if and only if $f_n \to f$ pointwise. Let $\{f_n\}_{n \in \mathbb{N}}$ be the sequence of Shrinking Triangles from Example 3.4.1.

(a) Explain why $\|f_n\| \neq 0$.

(b) Let $g_n = f_n/\|f_n\|$. Show that there exists a function g such that $g_n \to g$ pointwise.

(c) What is $\|g_n\|$? What is $\|g\|$? Why is this a contradiction?

Conclude that no such norm exists. (For this reason, we say that pointwise convergence of functions on $[0,1]$ is not a *normable* convergence criterion.)

2. Prove that the conclusions of Egorov's Theorem can be improved to say that there is a *closed set* $F \subseteq E$ such that:

(a) $|E \backslash F| < \varepsilon$, and

(b) f_n converges uniformly to f on F.

Section 3.5: Convergence in Measure

Definition 3.5.1 (Convergence in Measure). State.

Examples 3.5.2–3.5.5 (Shrinking Boxes I, Boxes Marching to Infinity, and Boxes Marching in Circles). Discuss. These are important "standard counterexamples." The Boxes Marching in Circles are illustrated in Figure 3.E a few pages later in this guide.

Note: The domain for the "Boxes Marching in Circles" is the interval $[0,1]$, but this is topologically isomorphic to a circle if we identify the endpoints 0 and 1.

Lemma 3.5.6. State and prove.

Note: Here are the details of how to construct the indices n_k at the beginning of the proof of Lemma 3.5.6.

Since $f_n \xrightarrow{\text{m}} f$, we know that

$$\lim_{n\to\infty} \bigl|\{|f - f_n| > 1\}\bigr| \,=\, 0.$$

Therefore there exists some positive integer n_1 such that

$$n \geq n_1 \implies \bigl|\{|f - f_n| > 1\}\bigr| \leq \frac{1}{2}.$$

Likewise, we have

$$\lim_{n\to\infty} \bigl|\{|f - f_n| > \tfrac{1}{2}\}\bigr| = 0,$$

so there exists some positive integer n_2 such that

$$n \geq n_2 \implies \bigl|\{|f - f_n| > \tfrac{1}{2}\}\bigr| \leq \frac{1}{4}. \tag{3.B}$$

In fact, this holds for all sufficiently large integers n_2. Therefore, by choosing a larger value for n_2 if necessary, we can fix n_2 so that equation (3.B) holds *and* we have $n_2 > n_1$. Continuing in this way, we obtain strictly increasing indices $0 < n_1 < n_2 < \cdots$ such that

$$\bigl|\{|f - f_n| > \tfrac{1}{k}\}\bigr| \leq 2^{-k}, \qquad \text{for all } n \geq n_k.$$

Exercise 3.5.7 and Corollary 3.5.8. Mention briefly if time permits.

Definition 3.5.9 (Cauchy in Measure). It may suffice to simply say that most convergence criteria have an associated Cauchy criterion, and this is the case for convergence in measure, with details given in the text.

Theorem 3.5.10. Theorem 3.5.10 will be used in Chapter 7 to prove that $L^p(E)$ is complete (see Exercise 7.3.5).

Note: Here are the details of how to construct the indices n_k at the beginning of the proof of Theorem 3.5.10.

Assume that $\{f_k\}_{k\in\mathbb{N}}$ is Cauchy in measure. Then there exists an integer $N_1 > 0$ such that

$$m, n \geq N_1 \implies \bigl|\bigl\{|f_m - f_n| > \tfrac{1}{2}\bigr\}\bigr| < \frac{1}{2}.$$

Set $n_1 = N_1$. Then there exists an integer $N_2 > 0$ such that

$$m, n \geq N_2 \implies \bigl|\bigl\{|f_m - f_n| > \tfrac{1}{2^2}\bigr\}\bigr| < \frac{1}{2^2}.$$

Without loss of generality, we may assume that $N_2 > N_1$. Let $n_2 = N_2$. Then we have n_1, $n_2 \geq N_1$, so

$$\big|\{|f_{n_2} - f_{n_1}| > \tfrac{1}{2}\}\big| < \frac{1}{2}.$$

Continuing in this way, we construct indices $n_1 < n_2 < \cdots$ such that

$$\big|\{|f_{n_{k+1}} - f_{n_k}| > 2^{-k}\}\big| < \frac{1}{2^k}, \qquad \text{for every } k \in \mathbb{N}.$$

Note: Convergence in measure is a *metrizable criterion*. That is, if we let $\mathcal{M}(E)$ be the set of all measurable functions on E and we identify functions that are equal almost everywhere, then there exists a metric d on $\mathcal{M}(E)$ such that $f_n \overset{\mathrm{m}}{\to} f$ if and only if $\mathrm{d}(f, f_n) \to 0$. This is shown in Problems 7.3.26 for sets E that have finite measure. It is also true that a sequence is Cauchy in measure if and only if it is Cauchy with respect to this metric. However, in practice it is usually easier to deal directly with the definition of convergence in measure than to try to work with the metric that induces that convergence criterion. On the other hand, the fact that an underlying metric exists is important, because it means that when dealing with convergence in measure we can use our intuition and experience from metric spaces (in particular, it suffices to use ordinary sequences indexed by the natural numbers $\mathbb{N}$ rather than needing to employ the generalizations of sequences known as "nets"). See Problems 3.5.14 and 3.5.16 for examples of results about convergence in measure that are analogous to results that hold in arbitrary metric spaces.

Extra Problems for Section 3.5

1. Assume that $E \subseteq \mathbb{R}^d$ and $f_n \colon E \to \overline{\mathbf{F}}$ are measurable and

$$\sum_{n=1}^{\infty} \big|\{|f_n| > \varepsilon\}\big| < \infty$$

for each $\varepsilon > 0$. Prove that $f_n \to 0$ pointwise a.e.

2. Let E be a measurable subset of $\mathbb{R}^d$, and suppose that $\{f_n\}_{n\in\mathbb{N}}$ is a sequence of measurable functions on E. Suppose that $f_n \to f$ pointwise a.e. and $f_n \overset{\mathrm{m}}{\to} g$. Prove that $f = g$ a.e.

Section 3.6: Luzin's Theorem

Luzin's Theorem is a very interesting result, and the proof of the generalization given in Problem 3.6.2 is a nice application of the Tietze Extension Theorem. However, Luzin's Theorem plays no further role in this volume. In my life as a professional mathematician, I have never had occasion to use

Luzin's Theorem in any proofs in my research papers (in contrast, Egorov's Theorem has surprised me by appearing multiple times). For these reasons, I usually (albeit with reluctance) only mention Luzin's Theorem in class, and assign this section as reading for the students. Naturally you may feel differently about whether to explicitly discuss Luzin's Theorem in class.

Luzin's Theorem is certainly an illuminating result. In fact, it is one of Littlewood's widely quoted *Three Principles*, originally stated in his 1944 text "Lectures on the Theory of Functions" (Oxford University Press, 1944).

> The extent of knowledge required is nothing like so great as is sometime supposed. There are three principles, roughly expressible in the following terms: Every [measurable] set is nearly a finite union of intervals; Every [measurable] function is nearly continuous; Every convergent sequence of [measurable] functions is nearly uniformly convergent. Most of the results of [the theory] are fairly intuitive applications of these ideas, and the student armed with them should be equal to most occasions when real variable theory is called for. If one of the principles would be the obvious means to settle the problem if it were 'quite' true, it is natural to ask if the 'nearly' is near enough, and for a problem that is actually solvable it generally is.

The first of Littlewood's principles has no name, but corresponds to Problem 2.2.38 in our text (and is 'nearly' our definition of the Lebesgue measure of a subset of the real line). The second of Littlewood's principles is Luzin's Theorem (Theorem 3.6.1), and the third principle is Egorov's Theorem (Theorem 3.4.2).

Note: "Luzin" is a transliteration from the Russian, the other common spelling in English is "Lusin."

Extra Problems for Section 3.6

1. Let E be a measurable subset of $\mathbb{R}^d$, and assume that $f\colon E \to \mathbb{C}$ is measurable. Show that there exist continuous functions $g_n\colon E \to \mathbb{C}$ such that $g_n \to f$ pointwise a.e.

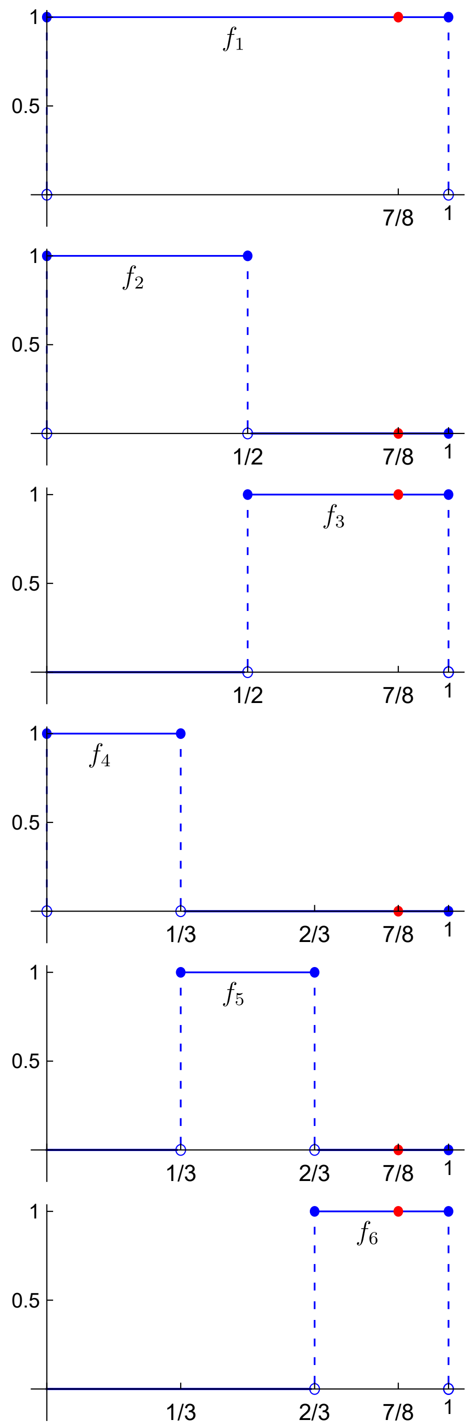

Fig. 3.E The Boxes Marching in Circles from Example 3.5.5. The value of $f_n(7/8)$, shown as a red dot, is 0 for infinitely many n, and 1 for infinitely many other n.

CHAPTER 4: THE LEBESGUE INTEGRAL

Section 4.1: The Lebesgue Integral of Nonnegative Functions

4.1.1 Integration of Nonnegative Simple Functions

Definition 4.1.1 (Integral of a Nonnegative Simple Function). State.

Note: Even though $0 \le c_k < \infty$ for every k, if $|E_k| = \infty$ for some k then we will have $\int_E \phi = \infty$.

Lemma 4.1.2. State and prove. As mentioned in the text, the point of statement (b) is that whenever we write a simple function as $\phi = \sum\limits_{k=1}^{N} c_k \chi_{E_k}$, whether this is the standard representation or not, then the integral of ϕ equals

$$\int_E \phi = \sum_{k=1}^{N} c_k |E_k|.$$

Exercise 4.1.3. State. Sometimes I write out the proof of part (d) in class.

Remark 4.1.4. Indeed, Problem 4.5.33 shows that $\mu(A) = \int_A f(x)\,dx$ (for measurable $A \subseteq E$) defines a signed or complex measure on E, and therefore this measure satisfies countable additivity and continuity from below.

4.1.2 Integration of Nonnegative Functions

Definition 4.1.5 (Lebesgue Integral of a Nonnegative Function). Motivate and state.

Notation 4.1.6. Omit.

Lemma 4.1.7. State but assign proof for reading.

Note: **TYPO** in the statement of the Lemma. Change "If ϕ is a simple function" to "If ϕ is a nonnegative simple function".

Lemma 4.1.8. State but assign proof for reading (some parts of the proof are exercises for the reader).

Theorem 4.1.9 (Tchebyshev's Inequality). State and prove.

Exercise 4.1.10. State. The proof is a fun (albeit simple) application of Tchebyshev's Inequality.

Note: "Tchebyshev" is a transliteration from the Russian, and there are number of different English spellings, including Chebyshev, Tchebysheff, and Chebysheff (with Chebyshev seeming to currently be the preferred spelling).

Extra Problems for Section 4.1

1. Add the following "continuity from above" part to Exercise 4.1.3.

(f) If $A_1 \supseteq A_2 \supseteq \cdots$ are nested measurable subsets of E, $A = \bigcap A_n$ and $\int_{A_1} \phi < \infty$, then $\int_A \phi = \lim_{n\to\infty} \int_{A_n} \phi$.

Section 4.2: The Monotone Convergence Theorem and Fatou's Lemma

4.2.1 The Monotone Convergence Theorem

Theorem 4.2.1 (Monotone Convergence Theorem). State and prove. An almost everywhere version of the MCT is given later, in Theorem 4.3.7.

The Monotone Convergence Theorem, Fatou's Lemma, and the Dominated Convergence Theorem are the three main theorems that deal with the convergence of integrals when the integrands converge pointwise a.e. Not only are these three theorems extremely important and useful, but they are also elegant and their *proofs* are short, attractive, and enlightening.

Note: The Monotone Convergence Theorem is also known as the "Beppo Levi Theorem," or (more rarely as far as I can tell) just "Levi's Theorem." When I first learned this, I thought it was the Beppo–Levi Theorem, proved by two mathematicians whose last names were Beppo and Levi. Only later did I learn that "Beppo Levi" is the full name of an Italian mathematician (1875–1961). I've always wondered why Levi's result is usually referred to by his full name, whereas most theorems or definitions named after mathematicians use only their surnames, e.g., Hilbert space, Young's Inequality, or the Hahn–Banach Theorem (named after Hahn and Banach, not someone whose first name was Hahn and last name Banach).

Remark 4.2.2. Briefly discuss if it seems interesting.

Note: If f_n is Riemann integrable, $0 \le f_n \nearrow f$, *and f is Riemann integrable*, then f_n and f are Lebesgue integrable, their Lebesgue integrals coincide with their Riemann integrals, and therefore the MCT implies that $\int f_n \nearrow \int f$. However, this can fail if f is not Riemann integrable.

Theorem 4.2.3. State and prove.

Corollary 4.2.4. State but assign proof for reading.

Exercise 4.2.5. It may be enough to just say that the integral of nonnegative functions satisfies countable additivity and continuity from below, and "see the text for an exercise that gives the precise formulation of these results."

4.2.2 Fatou's Lemma

Example 4.2.6 (Shrinking Boxes II). State. This is an important "standard counterexample."

Theorem 4.2.7 (Fatou's Lemma). State and prove. An almost everywhere version of Fatou's Lemma appears in in Theorem 4.3.8.

Fatou's Lemma is surprisingly useful. If you have nonnegative functions but they are not monotone increasing, then Fatou's Lemma at least gives you an inequality. I like to joke that, on average, about half the time that inequality will be going in the direction that you want, and then you're done. The other half of the time you're out of luck, but if you work hard you might be able to get it by some other means (probably the Dominated Convergence Theorem, which comes up soon).

Note: I've always wondered why Fatou's Lemma is a lemma but the Monotone Convergence Theorem is a theorem (especially since the two results are equivalent, see Problem 4.2.8).

Problems. **TYPO** *in Problem* 4.2.9 in the text: We should assume in this problem that the functions f_n are finite a.e.

Note on Problem 4.2.17 in the text: Wheeden and Zygmund [WZ77] take the measure of the region under the graph of a nonnegative measurable function f to be their *definition* of the Lebesgue integral of f. There are several advantages to this, e.g., the Monotone Convergence Theorem then follows easily from the continuity from below property of Lebesgue measure. One property that becomes considerably more difficult to prove with this approach is linearity of the integral, as it is not obvious how to relate the measures of the regions under the graphs of f and g to the measure of the region under the graph of $f+g$, in order to prove that $\int f + \int g = \int (f+g)$.

Extra Problems for Section 4.2

1. Let $g(x) = \dfrac{1}{\sqrt{|\sin 2\pi x|}}$, and consider the function $G(x) = \displaystyle\sum_{k=1}^{\infty} \frac{g(kx)}{k^2}$.

(a) Prove that $G(x) = \infty$ on a dense subset of $\mathbb{R}$.

(b) Prove that the series defining $G(x)$ converges to a finite number at almost every point $x \in \mathbb{R}$.

2. Use Problem 4.2.16 to give another proof that $\int_E (f+g) = \int_E f + \int_E g$ for all nonnegative measurable functions f and g on a measurable set E.

Section 4.3: The Lebesgue Integral of Measurable Functions

4.3.1 Extended Real-Valued Functions

Definition 4.3.1 (Lebesgue Integral of an Extended Real-Valued Function). State.

Note: $f^-(x)$ is zero whenever $f^+(x)$ is nonzero, and conversely. In particular, $f^+(x) - f^-(x)$ is never an indeterminate form.

Example 4.3.2. Discuss briefly.

Lemma 4.3.3. State and prove.

4.3.2 Complex-Valued Functions

Definition 4.3.4 (Lebesgue Integral of a Complex-Valued Function). State.

Lemma 4.3.5. State and prove.

There are certain proof strategies that are "stupidly simple but extremely useful." One of those, used in this proof, is helpful when you run into a situation where you have a complex number z but wish that you had its absolute value $|z|$ instead. If $z \neq 0$ and you write z in polar form as $z = re^{i\theta}$, then by taking $\alpha = e^{-i\theta}$ we have

$$|\alpha| = 1 \qquad \text{and} \qquad \alpha z \,=\, e^{-i\theta} z \,=\, r \,=\, |z|.$$

So, at the cost of multiplying by a scalar with unit modulus, you can turn z into $|z|$. You can do this for $z = 0$ too—in this case you can take $\alpha = e^{i\theta}$ with any value of θ that you like.

In this proof the complex number in question is $z = \int_E f$. We do have the inequality

$$|z| \,=\, \Bigl|\int_E f\Bigr| \,\leq\, \int_E |f|,$$

but what we really need in this proof is an equality, not an inequality. So we choose α with unit modulus so that $\alpha z = \bigl|\int_E f\bigr|$, and work with this instead. Simple, but useful.

Note: The statement of this lemma is slightly different than that of Lemma 4.3.3, which is the analgous result for extended real-valued functions. Also, while the proof of Lemma 4.3.5 is not as straightforward as that of Lemma 4.3.3, it is more interesting.

4.3.3 Properties of the Integral

Exercise 4.3.6. This type of exercise has appeared before. So, it may not be necessary to state the exercise precisely—it may suffice just to say in class that the properties like linearity and continuity from below that we proved in Exercise 4.1.3 for simple functions and in Exercise 4.1.8 for nonnegative functions have analogues for extended real-valued functions and complex-valued functions; see this exercise in the text for the precise statements.

ERRATA: However, the careful reader may observe that there is an embarassing **order-of-logic issue** with the proof of part (e), which is supposed to establish that $\int cf = c\int f$. There is no problem with the proof for the case of extended real-valued functions. But there is an issue with the complex case. We would like to argue as follows (but there is an issue with the line in red!): If the integral of $f\colon E \to \mathbb{C}$ exists and $c = a + ib$ where a and b are real, then

$$\begin{aligned}
&\int_E (cf) \\
&= \int_E (a+ib)\,(f_r + if_i) && \text{(substitute)} \\
&= \int_E \Big((af_r - bf_i) \,+\, i\,(bf_r + af_i)\Big) && \text{(complex number arithmetic)} \\
&= \int_E (af_r - bf_i) \,+\, i\int_E (bf_r + af_i) && \text{(def. of complex integrals)} \\
&= a\int_E f_r - b\int_E f_i + ib\int_E f_r + ia\int_E f_i && \textbf{(linearity of real integrals??)} \\
&= (a+ib)\left(\int_E f_r + i\int_E f_i\right) && \text{(complex number arithmetic)} \\
&= c\int_E f && \text{(def. of complex integrals)}.
\end{aligned}$$

Unfortunately, we only know at this point that ***part*** of "linearity of real integrals" holds. While we do know that $\int cf = c\int f$ for extended real-valued functions and real scalars, we have not yet shown that $\int (f+g) = \int f + \int g$ for integrable extended real-valued functions—that is not proved until Theorem 4.4.10. Therefore the conclusion stated in the red line in the calculations above is not yet justified based on what we have done to this point. Observe that this is only an issue for complex-valued functions.

One way to view this may be that I tried to be too clever in attempting to develop the theory of the integral simultaneously for extended real-valued and complex-valued functions. One solution to this issue would be to pause the development of the complex case at this point and proceed with the extended real-valued case alone until Theorem 4.4.10 is reached. At that point, one could return and extend all of the results from Exercise 4.3.6 onwards to the complex-case, until the two cases rejoin at Theorem 4.4.10.

In summary, this is not an unsolvable problem, but rather more of a technical annoyance. I suggest pointing out the error in logic to the class, and perhaps asking them if there is a more elegant solution than what I proposed above, which requires us to split the development into two cases (real and complex) from here until Theorem 4.4.10. Aside from making a new edition of the text, please let me know if you have a suggestion for a better solution!

Theorem 4.3.7 (Monotone Convergence Theorem) and Theorem 4.3.8 (Fatou's Lemma). These are the almost everywhere versions of these two theorems. In class, I usually just say that these are applications of Exercise 4.3.6, and assign them for reading.

Extra Problems for Section 4.3

1. Let f be a complex-valued measurable function defined on a measurable set $E \subseteq \mathbb{R}^d$. Show that if $\int_E f$ exists, then the integral of the complex conjugate of f is the complex conjugate of the integral of f: $\int_E \overline{f(x)}\,dx = \overline{\int_E f(x)\,dx}$.

2. Let $\phi(x) = e^{-\pi x^2}$ be the Gaussian function, and for each $\lambda > 0$ define $\phi_\lambda(x) = \lambda\phi(\lambda x)$. Set

$$f(x) = \lim_{\lambda \to 0^+} \phi_\lambda(x) \qquad \text{and} \qquad g(x) = \lim_{\lambda \to \infty} \phi_\lambda(x).$$

Use the fact that $\int \phi(x)\,dx = 1$ to determine whether $\int \phi_\lambda \to \int f$ as $\lambda \to 0^+$, or $\int \phi_\lambda \to \int g$ as $\lambda \to \infty$, where the integrals are taken over $\mathbb{R}$.

3. Let $E \subseteq \mathbb{R}^d$ be measurable and assume functions $f_n \colon E \to \overline{\mathbf{F}}$ are measurable and nonnegative a.e. Show that if $f_n \overset{\mathrm{m}}{\to} f$, then $\int_E f \le \liminf_{n\to\infty} \int_E f_n$.

4. (a) Let $E \subseteq \mathbb{R}^d$ and $f \colon E \to [\infty, \infty]$ be measurable. Suppose that f^n is integrable for every $n \in \mathbb{N}$ and $I = \lim_{n\to\infty} \int_E f(x)^n\,dx$ exists and is finite. Prove that $|f| \le 1$ a.e.

(b) Does part (a) hold if we only assume that I exists in the extended real sense (that is, if I is $\pm\infty$)?

(c) Does part (a) hold if we assume that f is complex-valued? Either prove that it does or exhibit a counterexample.

5. (a) Let $E \subseteq \mathbb{R}^d$ and $f \colon E \to [-\infty, \infty]$ be measurable. Suppose that f^n is integrable for every $n \in \mathbb{N}$ and there is a scalar c such that $\int_E f(x)^n\,dx = c$ for every n. Prove that there is a measurable set $A \subseteq E$ such that $f = \chi_A$ a.e.

(b) Does part (a) hold if we assume that f is complex-valued? Either prove that it does or exhibit a counterexample.

Section 4.4: Integrable Functions and $L^1(E)$

Definition 4.4.1 (L^1-Norm and Integrable Functions). State.

Note: When writing on the board, a standard abbreviation for "integrable" is ⓘ, i.e., the letter "i" with a circle around it.

Example 4.4.2. Discuss briefly.

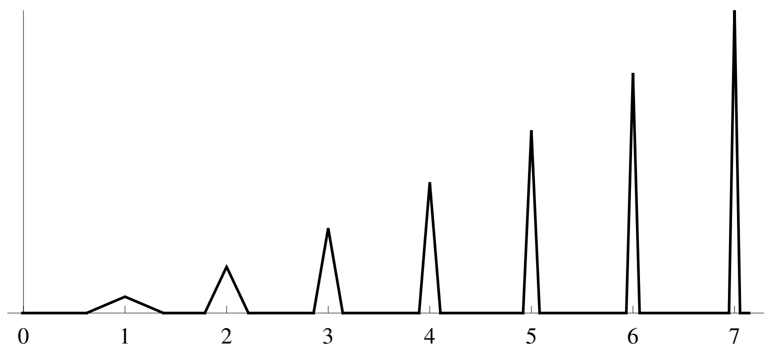

Fig. 4.A A continuous, unbounded function f.

Note: The idea of a continuous and unbounded yet integrable function is illustrated in Figure 4.A. By letting the heights of the triangles grow while their bases shrink quickly, we can make f unbounded yet still have $\int |f| < \infty$.

4.4.1 The Lebesgue Space $L^1(E)$

Definition 4.4.3 (The Lebesgue Space $L^1(E)$). State, and mention that $L^1(E)$ is closed under both addition and scalar multiplication.

Remark 4.4.4. As a harmonic analyst, convolution is very important to me so I usually make a brief remark about it here. But we will discuss convolution in detail in Section 4.6.3 (where it is presented as a nice application of Fubini's and Tonelli's theorems), so there is no reason that it has to be mentioned now.

Exercise 4.4.5. State. This is an easy exercise, but it is quite important, since it shows that $\|\cdot\|_1$ is "almost" a norm on $L^1(E)$. However, it is not a norm, because instead of precisely satisfying the uniqueness requirement we only have that $\|f\|_1 = 0$ when $f = 0$ *almost everywhere*. Still, this makes $L^1(E)$ into a very nice space.

Note: If we "identify" functions that are equal almost everywhere then $\|\cdot\|_1$ becomes a true norm on $L^1(E)$. We will discuss this idea of identification of functions that are equal a.e. in detail when we present the L^p spaces in Section 7.2.2.

4.4.2 Convergence in L^1-Norm

Definition 4.4.6 (Convergence in L^1-Norm). State.

Example 4.4.7. Discuss briefly.

Lemma 4.4.8. State and prove.

Figure 4.3. This diagram shows what I think are the most useful implications among the major forms of convergence criteria that we have encountered to this point.

Lemma 4.4.9. Omit.

4.4.3 Linearity of the Integral for Integrable Functions

Theorem 4.4.10. State and prove.

Lemma 4.4.11. Simple but useful; state and prove.

4.4.4 Inclusions between $L^1(E)$ and $L^\infty(E)$

Figure 4.4. Remark: The two functions that appear in this figure are

$$\begin{aligned} f(x) &= 3e^{-2000(x-1)^2} + 0.05\cos(65x) + 0.07\sin(50x), \\ g(x) &= f(x) + 0.5\sin(2x) + 0.7\cos(3x), \end{aligned}$$

on the domain $[0,4]$.

Lemma 4.4.12. If time is pressing (which it usually seems to be), then I just mention the lemma but assign it and the proof for reading.

Lemma 4.4.13. Also simple but useful; state and prove.

Corollary 4.4.14 (Uniform Convergence Theorem). State.

Extra Problems for Section 4.4

1. Prove that if $f \in L^1(E)$ where $E \subseteq \mathbb{R}^d$ is measurable, then the following statements hold.

(a) If A, $B \subseteq E$ are measurable with $|A \cap B| = 0$, then

$$\int_{A\cup B} f = \int_A f + \int_B f.$$

(b) If A is a measurable subset of E, then $\displaystyle\int_{E\setminus A} f = \int_E f - \int_A f$.

(c) If $A_1 \supseteq A_2 \supseteq \cdots$ are measurable subsets of E and $A = \cap A_n$, then

$$\int_A f = \lim_{n\to\infty} \int_{A_n} f.$$

2. Prove that if $f \in L^1(\mathbb{R})$, then $\lim\limits_{n\to\infty} \dfrac{1}{2n}\displaystyle\int_{-n}^{n} f(x)\,dx = 0$. Show by example that this can fail if f is not integrable.

3. Either prove or exhibit a counterexample: If $f_n \in L^1[0,1]$ and $\|f_n\|_1 \le 1$ for every n, then $f_n/n \to 0$ pointwise a.e. as $n \to \infty$.

4. Prove that if $f \in L^1(\mathbb{R})$, then $\sum\limits_{n=1}^{\infty} \bigl|\{|f| \ge n\}\bigr| \,<\, \infty$.

5. Given $f \in L^1(\mathbb{R})$ and $\alpha > 0$, show that $\lim\limits_{n\to\infty} n^{-\alpha} f(nx) = 0$ for a.e. $x \in \mathbb{R}$.

6. Let $E \subseteq \mathbb{R}^d$ be measurable, and suppose that functions $f_n,\ f \in L^1(E)$ satisfy $\|f - f_n\|_1 \le 1/n^2$ for $n \in \mathbb{N}$. Prove that $f_n \to f$ pointwise a.e.

7. Prove that $f(x) = x^{-1/2}\,\chi_{(0,1]}(x)$ is integrable on $\mathbb{R}$. Then let $\{r_n\}_{n\in\mathbb{N}}$ be an enumeration of the rational numbers $\mathbb{Q}$, and prove that the function $g(x) = \sum\limits_{n=1}^{\infty} 2^{-n} f(x - r_n)$ is integrable on $\mathbb{R}$ even though g *is unbounded on every interval.*

8. Let $E \subseteq \mathbb{R}^d$ be measurable, and let $f\colon E \to [-\infty, \infty]$ be a fixed function on E (note that we are not assuming that f is measurable!). Suppose that for each $\varepsilon > 0$ there exist functions $g,\ h \in L^1(E)$ such that $g(x) \le f(x) \le h(x)$ for a.e. $x \in E$ and $\|h - g\|_1 < \varepsilon$. Prove that f is measurable and $f \in L^1(E)$.

9. Suppose that $g \in L^1[0,1]$ satisfies $\|g\|_1 > 0$ and $f \in L^\infty[0,1]$ is such that $\bigl\{|f| = \|f\|_\infty\bigr\}$ has measure zero. Prove that

$$\left|\int_0^1 f(x)\,g(x)\,dx\right| \,<\, \|g\|_1\,\|f\|_\infty.$$

10. Let $E \subseteq \mathbb{R}^d$ be measurable. Prove that if $f \in L^1(E)$ satisfies $\int_{E\cap U} f \ge 0$ for every open set $U \subseteq \mathbb{R}^d$, then $f \ge 0$ a.e.

11. Suppose that $f \in L^1(\mathbb{R}^d)$ and there are real numbers $a < b$ such that

$$a\,|U| \,\le\, \int_U f \,\le\, b\,|U|, \quad \text{for every open set } U.$$

Prove that $a \le f \le b$ a.e.

12. (a) Prove *Barbălat's Lemma*: If $f \in C^1(a,\infty)$, if $\lim\limits_{t\to\infty} f(t)$ exists and is a scalar, and if f' is uniformly continuous, then $\lim\limits_{t\to\infty} f'(t) = 0$.

(b) How does this relate to Problem 4.4.16(c)?

13. Let $E \subseteq \mathbb{R}^d$ be measurable, and fix $f \in L^1(E)$. Prove that if $\int_E fg$ exists and is a scalar for every $g \in L^1(E)$, then $f \in L^\infty(E)$.

14. Show that if $f \in L^1(\mathbb{R})$, then the series

$$\sum_{k=1}^{\infty} \frac{1}{\sqrt{k}}\, f(x+\sqrt{k})$$

converges absolutely for almost every $x \in \mathbb{R}$.

Section 4.5: The Dominated Convergence Theorem

4.5.1 The Dominated Convergence Theorem

Theorem 4.5.1 (Dominated Convergence Theorem). State and prove.

This is an extremely important theorem. It's not that hard to prove, but in practice it's the first theorem that you turn to whenever you have a pointwise convergent sequence of functions and you want to prove that the integrals of those functions converge. Maybe you're lucky and your functions are monotone increasing, so you use the MCT, or they're nonnegative and you can get by with Fatou's Lemma, but aside from that your only recourse is likely to be the DCT.

Note: Theorem 4.5.1 is also known as the *Lebesgue Dominated Convergence Theorem*. with corresponding acronym LDCT.

Corollary 4.5.2 (Bounded Convergence Theorem). A simple but useful corollary. State; the proof is immediate from the DCT (because a constant function on a domain with finite measure is integrable).

Exercise 4.5.3. This proof of the DCT is much shorter than the one given in the text and may strike the expert as being more elegant, but I find that the longer version is more "enlightening" for the reader encountering the proof for the first time. In any case, this is a very nice exercise for the student to work out—a fun application of Fatou's Lemma.

4.5.2 First Applications of the DCT

Lemma 4.5.4. Even though the proof is very easy, I usually do present it in class as a quick and clear application of the DCT. Then the students can try their hand at applying the DCT in Exercise 4.5.5.

Exercise 4.5.5. State. Part (a) is an application of the DCT, and part (b) can be solved by applying part (a).

4.5.3 Approximation by Continuous Functions

Exercise 4.5.6 State the conclusion that $f(x) = \operatorname{dist}(x, A)$ is uniformly continuous.

Note: Here is a short direct proof of the fact that f is uniformly continuous.

Lemma. Let X be a metric space. If A is a nonempty subset of X, then the function $f\colon X \to \mathbb{R}$ defined by

$$f(x) \,=\, \mathrm{dist}(x,A) \,=\, \inf\bigl\{\mathrm{d}(x,z) : z \in A\bigr\}, \qquad x \in X,$$

is uniformly continuous on X.

Proof. Fix $\varepsilon > 0$, and set $\delta = \varepsilon/2$. Choose any two points $x,\, y \in X$ such that $\mathrm{d}(x,y) < \delta$. By definition of the distance function, there exist points $a,\, b \in A$ such that

$$\mathrm{d}(x,a) \,<\, \mathrm{dist}(x,A) \,+\, \delta \qquad \text{and} \qquad \mathrm{d}(y,b) \,<\, \mathrm{dist}(y,A) \,+\, \delta.$$

Consequently,

$$\begin{aligned} f(y) \,=\, \mathrm{dist}(y,A) &\le \mathrm{d}(y,a) \\ &\le \mathrm{d}(y,x) \,+\, \mathrm{d}(x,a) \\ &< \delta \,+\, \bigl(\mathrm{dist}(x,A) + \delta\bigr) \,=\, f(x) \,+\, \varepsilon. \end{aligned}$$

Interchanging the roles of x and y, we similarly obtain $f(x) < f(y) + \varepsilon$. Therefore $|f(x) - f(y)| < \varepsilon$ whenever $\mathrm{d}(x,y) < \delta$, so f is uniformly continuous on X. □

Theorem 4.5.7 (Urysohn's Lemma) State and prove.

Note: The proof for metric spaces based on Exercise 4.5.6 is very short. This theorem can be generalized to normal topological spaces, although the proof in that setting is more difficult.

Theorem 4.5.8. State and prove. The remarks after the proof talk about why this theorem implies that $C_c(\mathbb{R}^d)$ is a *dense* subset of $L^1(\mathbb{R}^d)$. This terminology will be used throughout the rest of the text, especially in Chapter 7, so it might be appropriate to recall the definition of a dense set and to briefly discuss Theorem 4.5.8 in that context. Dense sets are briefly introduced in Definition 1.1.5, but there is also a more detailed discussion of them in Section $1'$.10 of **Alternative Chapter 1**.

Note: I find the fact that infinite-dimensional spaces can contain proper dense *subspaces* to be quite counterintuitive based on our experience with finite-dimensional spaces such as $\mathbb{R}^d$.

For example, the set of rationals $\mathbb{Q}$ is a dense *subset* of the real line, but it is not a *subspace* because it is not closed under multiplication by arbitrary scalars. Recall that in this text we have specified that the scalar field can only be $\mathbb{R}$ or $\mathbb{C}$, and $\mathbb{Q}$ is not closed under multiplication by scalars in these fields (e.g., 2 is rational and π is a scalar, but 2π is not rational).

Exercise 4.5.9 (Strong Continuity of Translation). This is an important exercise. We will see one application in Theorem 5.5.3, and more in Chapter 9 when we study convolution and the Fourier transform. It is not difficult to solve this exercise, but the solution is nice and is a typical example of how to take advantage of the fact that some "nice" subspace is dense in the space you are working in (in this case, C_c is dense in L^1). Therefore I state the exercise in class (and almost always assign it as a formal homework assignment to be turned in).

The fact that $T_a f \to f$ in L^1-norm almost seems "obvious," so I use the example of the box function $\chi_{[0,1]}$ to show that the analogous result fails when we use the L^∞-norm instead of the L^1-norm. This is a good illustration of how the meaning of "close" changes depending on what norm you use.

Note: Although $T_a f \to f$ in L^1-norm, it is *not true* in general that $T_a f \to f$ pointwise or pointwise almost everywhere (this is *Extra Problem 1* below).

Remark 4.5.10. Omit.

4.5.4 Approximation by Really Simple Functions

Definition 4.5.11 (Really Simple Function). State.

Note: Leib and Loss' text [LL01] is the only one that I'm aware of that uses the terminology "really simple function." When I first encountered this name I thought it was too glib, but it didn't take very long for me to change my mind. After all, these are simple functions (in the usual mathematical sense) that are "extra simple," being finite linear combinations of characteristic functions of *intervals* (instead of characteristic functions of *measurable sets*). The traditional name of "step function" is also quite descriptive, so I often use the two names interchangeably—although, a really simple function as defined in the text is slightly more restrictive because a really simple function is a linear combination of characteristic functions of half-open intervals $[a, b)$, while a "step function" is usually taken to be a linear combination of characteristic functions of any type of interval.

Theorem 4.5.12. State and prove.

4.5.5 Relation to the Riemann Integral

Theorem 4.5.13. State. I usually prove part (a) in full, and then make some remarks about the proof of part (b) but assign the details for reading. The proof of part (b) is more or less similar to that of part (a), but in part (b) we must deal with all possible partitions Γ_k such that $|\Gamma_k| \to 0$, and consequently we cannot assume that the corresponding functions ϕ_k and ψ_k are monotone increasing or decreasing.

It is observed after the proof that an improper Riemann integral need not equal a Lebesgue integral over the same interval. This is not so surprising, since an improper Riemann integral is a limit of Riemann integrals, so it is not itself a Riemann integral.

Note: One way to see the difference between the Riemann and the Lebesgue integral is to look at which axis is subdivided in the definition. In the Riemann integral, we partition the x-axis. We let $a = x_0 < x_1 < \cdots < x_n = b$ be a partition of the interval $[a,b]$ in the x-axis, and form a corresponding Riemann sum

$$R_\Gamma = \sum_{j=1}^{n} f(\xi_j)\,(x_j - x_{j-1}),$$

where $\xi_j \in [x_{j-1}, x_j]$. We then consider appropriate limits of these Riemann sums to form the Riemann integral.

In contrast, one way to view the Lebesgue integral is to consider the approximation by simple functions that we first described in Theorem 3.2.14. There, instead of partitioning the x-axis, we partition the *y-axis* into subintervals of length 2^{-n}, and consider the corresponding simple functions ϕ_n defined in the proof of that theorem. The reader should draw a picture that compares the two approaches (but beware, because for a *continuous function* f whose domain is an interval $[a,b]$ the pictures look somewhat similar—remember that Lebesgue integral applies to a much broader class of functions than does the Riemann integral).

Note: Here is a detailed proof of the somewhat simpler special case of the Riemann integral of a nonnegative continuous function on a closed bounded interval.

Theorem. If f is a nonnegative continuous function on the interval $[a,b]$, then its Riemann integral equals its Lebesgue integral.

Proof. Let f be a nonnegative continuous function f on a finite closed interval, which we take to be $[0,1]$ for convenience. Such a function is measurable and bounded, so its Lebesgue integral $\int_0^1 f(t)\,dt$ exists and is finite. Additionally, f is Riemann integrable, and we can write its Riemann integral as the limit of Riemann sums using regular partitions, rather than needing to deal with arbitrary partitions.

To set the notation, fix an integer $n > 0$, and let $\Delta_n = \frac{1}{n}$. Define points

$$t_k = \frac{k}{n}, \qquad \text{for } k = 0, 1, \dots, n,$$

and note that we are implicitly letting t_k depend on the value of n. Since f is continuous, for each $k = 1, \dots, n$ there is a point $t_k^* \in [t_{k-1}, t_k]$ where f achieves its minimum on that interval (again, t_k^* implicitly depends on n). Then $L_n = \sum_{k=1}^{n} f(t_k^*)\,\Delta_n$ is a *lower Riemann sum* for f.

Since f is Riemann integrable, these lower Riemann sums converge to the Riemann integral, which we will call I:

$$I = \lim_{n\to\infty} L_n = \lim_{n\to\infty} \sum_{k=1}^{n} f(t_k^*)\,\Delta_n. \tag{4.A}$$

Define

$$\phi_n = \sum_{k=1}^{n} f(t_k^*)\,\chi_{[t_{k-1},t_k)}.$$

This is a simple function, and its *Lebesgue integral* $\int_0^1 \phi_n$ equals the lower Riemann sum L_n. Since f is continuous, the functions ϕ_n converge pointwise to f on the interval $[0,1)$. We declare that $\phi_n(1) = f(1)$, so $\phi_n(t)$ converges pointwise to $f(t)$ for every $t \in [0,1]$. Since $\phi_n \le f$,

$$L_n = \int_0^1 \phi_n \le \int_0^1 f. \tag{4.B}$$

Therefore,

$$\begin{aligned}
I = \lim_{n\to\infty} L_n \le \int_0^1 f & \qquad \text{(by equations (4.A) and (4.B))}\\
= \int_0^1 \Bigl(\lim_{n\to\infty} \phi_n\Bigr) & \qquad \text{(since } \phi_n \to f \text{ pointwise)}\\
\le \liminf_{n\to\infty} \int_0^1 \phi_n & \qquad \text{(Fatou's Lemma)}\\
= \liminf_{n\to\infty} L_n = I. & \qquad \text{(by equation (4.A))}
\end{aligned}$$

Hence the Lebesgue integral of f equals the Riemann integral of f. □

Note: Here are the details of why a claim made in the proof of part (b) of Theorem 4.5.13 is true.

Assume that f is discontinuous at a point $x \notin Z \cup S$. In this case $x \in (a,b)$, since a and b are both in S. There must exist some $\varepsilon > 0$ such that for every $\delta > 0$ there is a point $t \in (x-\delta, x+\delta) \cap (a,b)$ such that $|f(x) - f(t)| \ge \varepsilon$. **CLAIM:** It follows from this that

$$\psi_k(x) - \phi_k(x) \ge \varepsilon \quad \text{for every } k \in \mathbb{N}.$$

To see why this claim is true, fix any particular $k \in \mathbb{N}$. Write the partition Γ_k as

$$\Gamma_k = \{a = x_0 < x_1 < \cdots x_n = b\}.$$

Since $x \notin S$, we know that x is not equal to any x_j. Hence $x \in (x_{j-1}, x_j)$ for some j, and since this interval is open there is some $\delta > 0$ such that $(x-\delta, x+\delta) \subseteq (x_{j-1}, x_j)$. Let t be the point specified above.

Now, either $f(x) \geq f(t)$ or $f(t) \leq f(x)$. If we have $f(t) \geq f(x)$, then $f(t) - f(x) \geq \varepsilon$, so

$$\begin{aligned}
\psi_k(x) = M_j &= \sup_{u \in [t_{j-1}, t_j]} f(u) && \text{(definition of } M_j\text{)} \\
&\geq f(t) && \big(\text{since } t \in (x-\delta, x+\delta) \subseteq (t_{j-1}, t_j)\big) \\
&\geq f(x) + \varepsilon \\
&\geq \left(\inf_{u \in [t_{j-1}, t_j]} f(u)\right) + \varepsilon && \big(\text{since } x \in (t_{j-1}, t_j)\big) \\
&= m_j + \varepsilon = \phi_k(x) + \varepsilon.
\end{aligned}$$

Hence $\psi_k(x) - \phi_k(x) \geq \varepsilon$ in this case. A symmetric argument shows that this inequality also holds if $f(x) \geq f(t)$.

An Extra Theorem.

The proof of Theorem 4.5.8 given in the text employs Urysohn's Lemma. Below is another theorem whose proof is a nice application of Urysohn's Lemma. I like this proof, but I didn't include it in the text and usually don't present it in class (although it would be a good reading assignment). The proof is a little long, and shorter proofs can be given using techniques that we will develop in Section 6.4. In fact, this result is Theorem 6.4.7 in the main text, and it is proved there by applying the Weierstrass Approximation Theorem and integration by parts. Additionally, Problem 9.1.32 uses the technique of convolution to give a variation on and improvement to this theorem.

Theorem. If $f \in L^1[a,b]$ satisfies

$$\int_a^b f(x)\, g(x)\, dx = 0, \qquad \text{for all } g \in C[a,b], \tag{A}$$

then $f = 0$ a.e.

Proof. Before beginning the proof, we observe that if we were allowed to take $g \in L^\infty[a,b]$ in equation (A) instead of $g \in C[a,b]$, then the proof would be easy, because we could choose g so that $|g(x)| = 1$ and $f(x)\,g(x) = |f(x)|$ for every x. Unfortunately, such a function g need not be continuous, so we must be more careful.

Case 1: $\overline{\mathbf{F}} = [-\infty,\infty]$. Suppose that $f\colon [a,b] \to [-\infty,\infty]$ is integrable and equation (A) holds for all continuous real-valued functions g.

If $\{f > 0\}$ has positive measure, then there must exist some $\delta > 0$ such that

$$m \;=\; \bigl|\{f > \delta\}\bigr| \;>\; 0.$$

By Exercise 4.5.5(b), there exists an $\varepsilon > 0$ such that for each measurable set $A \subseteq [a,b]$ we have

$$|A| < \varepsilon \quad\implies\quad \int_A |f| \;<\; \frac{\delta m}{2}. \tag{B}$$

By Lemma 2.2.15, there exists a closed set $K \subseteq \{f > \delta\}$ such that

$$|K| > \frac{m}{2} \qquad \text{and} \qquad \bigl|\{f > \delta\}\backslash K\bigr| \;<\; \frac{\varepsilon}{2}.$$

Likewise, there exists an open set $U \supseteq \{f > \delta\}$ such that

$$\bigl|U\backslash\{f > \delta\}\bigr| \;<\; \frac{\varepsilon}{2}.$$

Since K and U^{C} are disjoint closed sets, Urysohn's Lemma implies that there exists a continuous function $\theta\colon \mathbb{R} \to \mathbb{R}$ such that $0 \le \theta \le 1$ everywhere on $\mathbb{R}$, $\theta = 1$ on K, and $\theta = 0$ on $\mathbb{R}\backslash U$.

Let $V = U \cap [a,b]$. Then $|V \backslash K| < \varepsilon$, so by equation (B) we have

$$\biggl|\int_{V\backslash K} f\theta\biggr| \;\le\; \int_{V\backslash K} |f\theta| \;\le\; \int_{V\backslash K} |f| \;<\; \frac{\delta m}{2}.$$

On the other hand, $f > \delta$ and $\theta = 1$ on K, so

$$\int_K f\theta \;\ge\; \int_K \delta \;=\; \delta\,|K| \;>\; \frac{\delta m}{2}.$$

Within the interval $[a,b]$, the product $f\theta$ is zero everywhere outside of the set V. Hence

$$0 \;=\; \int_a^b f\theta \;=\; \int_V f\theta \;=\; \int_K f\theta \;+\; \int_{V\backslash K} f\theta \;>\; \frac{\delta m}{2} \;-\; \frac{\delta m}{2} \;=\; 0.$$

This is a contradiction, so $\{f > 0\}$ must have measure zero. A symmetric argument shows that $|\{f < 0\}| = 0$, and therefore $f = 0$ a.e.

Case 2: $\overline{\mathbf{F}} = \mathbb{C}$. Suppose that f is a complex-valued integrable function such that equation (A) holds for all continuous complex-valued functions g on $[a,b]$. Write $f = f_r + i f_i$ where f_r and f_i are real-valued. Then equation (A) implies that $\int_a^b f_r\, g = 0$ for every continuous real-valued g, so $f_r = 0$ a.e. by Case 1. Similarly, $f_i = 0$ a.e., so it follows that $f = 0$ a.e. □

Extra Discussion: A Side Journey into Abstract Measure Theory

I do not always present this in class, but here is some extra discussion that gives some context to part (b) of Exercise 4.5.5. The way I teach the course, the first semester (corresponding to the material in this textbook) is focused on Lebesgue measure, and abstract measure theory is not covered until the second semester. Sometimes I present this material in class as a small preview of something that is coming in the second semester, However, abstract measures are not needed in the remainder of this volume. A very short version of this material is incorporated into the text in Problem 4.5.33. An extra Chapter 10 on abstract measure theory is available online at the text's website, see `http://people.math.gatech.edu/~heil/real/`

A Side Journey into Abstract Measure Theory

To elaborate on the significance of Exercise 4.5.5(b), consider the following definition from abstract measure theory (for motivation, recall from Section 2.2.2 that the set $\mathcal{L}$ of all Lebesgue measurable subsets of $\mathbb{R}^d$ is a σ-algebra of subsets of $\mathbb{R}^d$).

Definition. (Signed and Complex Measures) Let X be a set, and let Σ be a σ-algebra of subsets of X (see Definition 2.2.14). A function $\nu\colon \Sigma \to [-\infty, \infty]$ is a *signed measure* on (X, Σ) if:

(a) $\nu(\varnothing) = 0$,

(b) $\nu(A)$ takes at most one of the values ∞ and $-\infty$, and

(c) ν is countably additive. That is, if $A_1, A_2, \dots$ are countably many disjoint sets in Σ, then

$$\nu\biggl(\bigcup_k A_k\biggr) \;=\; \sum_k \nu(A_k).$$

A *complex measure* on (X, Σ) is a function $\nu\colon X \to \mathbb{C}$ that satisfies the same three properties (although property (b) is superfluous in this case, since $\nu(A)$ must always be a complex number and therefore can never be $\pm\infty$).

We say that a signed or complex measure ν is a *positive measure* on (X, Σ) if $\nu(A) \geq 0$ for every $A \in \Sigma$. ♢

For example, the results of Chapter 2 show that Lebesgue measure is a positive measure on $(\mathbb{R}^d, \mathcal{L})$. The following lemma shows how to use integrable functions to construct other examples of signed and complex measures on $(\mathbb{R}^d, \mathcal{L})$.

Lemma. Fix $f \in L^1(\mathbb{R}^d)$ and set

$$\nu_f(A) = \int_A f(t)\,dt, \qquad \text{for all measurable } A \subseteq \mathbb{R}^d.$$

Then ν is a signed measure on $(\mathbb{R}^d, \mathcal{L})$ if f is extended real-valued, and it is a complex measure on $(\mathbb{R}^d, \mathcal{L})$ if f is complex-valued.

Proof. We trivially have $\nu_f(\varnothing) = 0$, and if A is a measurable set then

$$|\nu_f(A)| = \left|\int_A f(t)\,dt\right| \leq \int_A |f(t)|\,dt \leq \int_E |f(t)|\,dt = \|f\|_1 < \infty.$$

Hence $\nu_f(A)$ can never be $\pm\infty$. Finally, if $A_1, A_2, \dots$ are disjoint measurable sets, then Exercise 4.3.6(f) implies that

$$\nu\biggl(\bigcup_k A_k\biggr) = \int_{\cup A_k} f = \sum_k \int_{A_k} f = \sum_k \nu(A_k).$$

Therefore properties (a)–(c) in the Definition are all satisfied. □

Essentially, the measure ν_f uses the function f to place a "weighting" on $\mathbb{R}^d$. Instead of taking the measure of a set A to be its Lebesgue measure, ν_f assigns it the value $\int_A f$. We think of $\nu_f(A) = \int_A f$ as being the measure of the set A under the measure ν_f. For example, if f is identically 1, then $\nu_f(A)$ is simply the Lebesgue measure of A. However, in contrast to Lebesgue measure, since f need not be nonnegative, the measure $\nu_f(A)$ of a set A might be negative or complex. Still, ν_f has the countable additivity property that is given in statement (c) of the Definition above.

The Lemma above is not the only way to create signed or complex measures. For example, Problem 4.5.33 constructs two interesting measures known as *counting measure* and the *δ measure*. However, there is an important difference between those two measures and the measures ν_f

constructed in the Lemma above. Rewording Exercise 4.5.5(b) in terms of ν_f, we have that if f is an integrable function, then given any $\varepsilon > 0$ there is a $\delta > 0$ such that

$$|A| < \delta \quad \Longrightarrow \quad |\nu_f(A)| < \varepsilon.$$

In particular, it follows (why?) that

$$|A| = 0 \quad \Longrightarrow \quad \nu_f(A) = 0.$$

Thus, if the Lebesgue measure of a set is zero, then the ν_f-measure of that set is also zero. Using the language of abstract measure theory, we say that the measure ν_f is *absolutely continuous* with respect to Lebesgue measure. In contrast, Problem 4.5.33 shows that counting measure and the δ measure are *not* absolutely continuous with respect to Lebesgue measure.

Although we will not delve further into the meaning of this type of absolute continuity, we mention that there is a relationship to the *absolutely continuous functions* that we will study in Chapter 6. In particular if f is an integrable function on $E = [a,b]$, then its antiderivative

$$F(x) = \int_a^x f(t)\,dt = \nu_f([a,x]), \qquad x \in [a,b],$$

is an absolutely continuous function in the sense of Chapter 6. For more details on abstract measure theory, we refer to texts such as [Rud87], [Fol99], or [BC09].

Extra Problems for Section 4.5

1. Prove that if

$$f(t) = \begin{cases} e^{-|t|}, & \text{if } t \text{ is irrational}, \\ 0, & \text{if } t \text{ is rational}, \end{cases}$$

then the following statements hold.

(a) f is integrable on $\mathbb{R}$, and hence Exercise 4.5.9 implies that $T_a f \to f$ in L^1-norm as $a \to 0$.

(b) There is *no point* $t \in \mathbb{R}$ where $T_a f(t) \to f(t)$ as $a \to 0$.

2. Evaluate the following limits.

(a) $\displaystyle\lim_{n\to\infty} \int_0^1 \frac{ne^x}{1+n^2x^{1/2}}\,dx.$

(b) $\displaystyle\lim_{n\to\infty} \int_0^1 \frac{2x^2+1+x^5/n}{x^{1/2}+\sin(x/n)}\,dx.$

(c) $\displaystyle\sum_{n=0}^{\infty} \frac{1}{n!} \int_1^2 \ln^n x\,dx.$

3. Let f be an integrable function defined on a measurable set $E \subseteq \mathbb{R}^d$. Show that if $\{A_n\}_{n\in\mathbb{N}}$ is a sequence of measurable subsets of E such that $|A_n| \to 0$, then $\int_{A_n} f \to 0$.

4. Prove that if f is monotone increasing on $[a,b]$, then f is Riemann integrable.

5. (a) Prove that if f is continuous and nonnegative on $[a,\infty)$, then its Lebesgue integral and improper Riemann integral on $[a,\infty)$ both exist and are equal:

$$\int_a^\infty f = \lim_{b\to\infty} \int_a^b f.$$

Remark: Since f is nonnegative and measurable, both the improper Riemann integral and the Lebesgue integral exist in the extended real sense, so the issue is to show that they are equal (note that they could be infinite).

(b) Prove that if f is continuous and nonnegative on $(a,b]$ then its Lebesgue integral and improper Riemann integral on $[a,b]$ both exist and are equal:

$$\int_a^b f = \lim_{c\to a^+} \int_c^b f.$$

6. (a) Assume that f is continuous a.e. on $(a,b]$, and is bounded on $[c,b]$ for every $a < c < b$. Prove that if f is Lebesgue integrable on $[a,b]$, then its improper Riemann integral on that interval exists and equals its Lebesgue integral:

$$\int_a^b f = \lim_{c\to a^+} \int_c^b f.$$

(b) Assume that f is continuous a.e. on $[a,\infty)$, and is bounded on $[a,b]$ for every $b > a$. Prove that if f is Lebesgue integrable on $[a,\infty)$, then its improper Riemann integral on that interval exists and equals its Lebesgue integral:

$$\int_a^\infty f = \lim_{b\to\infty} \int_a^b f.$$

7. Prove that if $f \in L^1[0,1]$, then $\displaystyle\lim_{n\to\infty} \int_0^1 x^n f(x)\,dx = 0$. Also show that if $\displaystyle\alpha = \lim_{x\to 1^-} f(x)$ exists, then $\displaystyle\lim_{n\to\infty} n\int_0^1 x^n f(x)\,dx = \alpha$.

8. (From Folland.) Compute the following limit for each of the cases $a > 0$, $a = 0$, and $a < 0$:

$$\lim_{n\to\infty} \int_a^\infty \frac{n}{1+n^2x^2}\,dx.$$

9. Suppose that $g \in L^1(\mathbb{R})$, and let f be a continuous function on $\mathbb{R}$ that has compact support. Evaluate the following limits.

(a) $\displaystyle\lim_{t\to\infty} \int_{-\infty}^{\infty} f(tx)\, g(x)\, dx.$

(b) $\displaystyle\lim_{t\to 0} \int_{-\infty}^{\infty} f(tx)\, g(x)\, dx.$

10. Choose $f \in L^1(\mathbb{R})$, and for each $\lambda > 0$ define $f_\lambda(x) = \lambda f(\lambda x)$. Prove the following statements.

(a) $\|f_\lambda\|_1 = \|f\|_1$ for each $\lambda > 0$ (for this reason, we refer to f_λ as an L^1-*normalized dilation* of f).

(b) The L^1-normalized dilation is *strongly continuous on* $L^1(\mathbb{R})$. That is,

$$\lim_{\lambda\to 1} \|f - f_\lambda\|_1 \,=\, 0.$$

11. Let E be a measurable subset of $\mathbb{R}^d$. A (possibly uncountable) collection of integrable functions $\{f_j\}_{j\in J}$ is said to be *uniformly integrable* on E if for every $\varepsilon > 0$ there exists a $\delta > 0$ such that for every measurable set $A \subseteq E$ we have

$$|A| < \delta \quad \Longrightarrow \quad \int_A |f_j| < \varepsilon \quad \text{for all } j \in J.$$

Prove the following statements.

(a) Any set of finitely many integrable functions is uniformly integrable.

(b) If a sequence $\{f_n\}_{n\in\mathbb{N}}$ of integrable functions is Cauchy in L^1-norm, then $\{f_n\}_{n\in\mathbb{N}}$ is uniformly integrable.

(c) If a sequence $\{f_n\}_{n\in\mathbb{N}}$ of integrable functions converges in L^1-norm, then $\{f_n\}_{n\in\mathbb{N}}$ is uniformly integrable.

(d) If $|E| < \infty$, $\{f_n\}_{n\in\mathbb{N}}$ is uniformly integrable, and $f_n \to f$ a.e., then $f_n \to f$ in L^1-norm.

(e) The assumption in part (d) that E has finite measure is necessary.

12. Given a bounded measurable function $\alpha\colon [a,b] \to \mathbb{C}$, prove that there exist simple functions ϕ_n such that:

(a) $\displaystyle\phi_n = \sum_{j=1}^{M_n} c_j^n\, \chi_{[a_{j-1}^n, a_j^n)}$ where $a = a_0^n < a_1^n < \cdots < a_{M_n}^n = b$,

(b) $|c_n^j| \le \|\alpha\|_\infty$ for every n and j, and

(c) $\phi_n \to \alpha$ pointwise a.e.

13. Fix $f \in L^1(\mathbb{R})$, and for each $n \in \mathbb{N}$ let $\displaystyle g_n(x) = \int_{x-n}^{x+n} f(t)\, dt.$

(a) Prove that g_n is continuous.

(b) Given $x \in \mathbb{R}$, determine whether $\lim\limits_{n\to\infty} g_n(x)$ exists, and if so find it.

(c) Given $n \in \mathbb{N}$, determine whether $\lim\limits_{x\to\infty} g_n(x)$ exists, and if so find it.

14. Let $[a,b]$ be a finite interval.

(a) Prove that $\displaystyle\lim_{n\to\infty} \int_a^b \cos^2 nx\,dx \;=\; \frac{b-a}{2}.$

(b) Suppose that $a_n \cos nx \to 0$ pointwise a.e. on $[a,b]$. Prove that $a_n \to 0$ as $n \to \infty$.

15. Evaluate $\displaystyle\lim_{n\to\infty} \int_0^\infty \frac{x^{n-2}}{1+x^n}\, \sin(\pi x/n)\,dx.$

16. Evaluate $\displaystyle\lim_{N\to\infty} N \int_0^N \frac{1}{t}\, \ln\bigl(1+\tfrac{t}{N}\bigr)\, \frac{dt}{1+t^2}\,dt.$

17. Assume that functions $f_n\colon [0,1] \to [0,\infty]$ are measurable and $f_n \to f$ a.e. for some $f \in L^1[0,1]$.

(a) Prove that the integrals $\int_0^1 \min\bigl\{f_n(x), f(x)\bigr\}\,dx$ exist for each n, and

$$\lim_{n\to\infty} \int_0^1 \min\bigl\{f_n(x), f(x)\bigr\}\,dx \;=\; \int_0^1 f(x)\,dx.$$

(b) Prove that if $\displaystyle\lim_{n\to\infty} \int_0^1 f_n = \int_0^1 f$, then $f_n \to f$ in L^1-norm.

18. Let $\mathcal{P}_e$ be the set of all even polynomials. Taking our domain to be the interval $[-1,1]$, what is the closure of $\mathcal{P}_e$ in $L^1[-1,1]$?

19. Let $E \subseteq \mathbb{R}^d$ be measurable. Assume that functions $f_n \in L^1(E)$ satisfy $\|f_n\|_1 \to 0$, and there is some function $g \in L^1(E)$ such that $|f_n|^2 \le g$ a.e. for every n. Prove that $\int_E |f_n|^2 \to 0$ as $n \to \infty$.

20. Assume that f is continuous on $[0,1]$, g is measurable on $[0,1]$, and $0 \le g(x) \le 1$ a.e. Evaluate $\displaystyle\lim_{n\to\infty} \int_0^1 f\bigl(g(x)^n\bigr)\,dx.$

21. Add this question to Problem 4.5.25: What is the limit if we only assume that K is bounded, instead of assuming that K is compact?

22. Prove that if $f \in L^1(\mathbb{R})$, then $\displaystyle\lim_{n\to\infty} \int_{-\infty}^\infty f(x)\, \sin nx\,dx = 0.$

Hint: First prove the result for really simple functions.

23. Prove that if $f \in L^1[0,1]$, then

$$\lim_{n\to\infty} \int_0^1 f(x)\,|\sin 2\pi nx|\,dx \;=\; \frac{2}{\pi} \int_0^1 f(x)\,dx.$$

Hint: First prove the result for really simple functions.

24. Use Problem 4.5.27 to give another solution to Problem 4.4.24.

25. Let X be a set, and let Σ be a σ-algebra of subsets of X. A function $\nu\colon \Sigma \to \mathbb{C}$ is a *complex measure* on (X, Σ) if: $\nu(\varnothing) = 0$ and ν is countably additive, i.e., if $E_1, E_2, \dots$ are countably many disjoint sets that belong to Σ, then

$$\nu\biggl(\bigcup_k E_k\biggr) \;=\; \sum_k \nu(E_k).$$

Extend the results of Problem 4.5.33 to complex measures.

26. Let $f(x) = \sin x^2$. Prove the following statements.

(a) The Lebesgue integral of f on the interval $[0, \infty)$ does not exist.

(b) The improper Riemann integral $\displaystyle\lim_{b\to\infty} \int_0^b \sin x^2\, dx$ does exist.

Remark: The Riemann integrals $\int_0^b \sin x^2\, dx$ and $\int_0^b \cos x^2\, dx$ are known as *Fresnel integrals*. It can be shown that they converge to $\sqrt{\pi/8}$ as $b \to \infty$.

27. Given $f \in L^1(\mathbb{R})$, compute (with proof): $\displaystyle\lim_{t\to\infty} \int_{-\infty}^{\infty} |f(x-t) - f(x)|\, dx.$

Remark: The limit here is as $t \to \infty$, not $t \to 0$.

Section 4.6: Repeated Integration

The point of this section is that we cannot always interchange the order of repeated integrals. Essentially, there can be difficulties if indeterminate forms arise in the wrong place. The theorems of Fubini and Tonelli give hypotheses which allow us to avoid this situation. Fubini's Theorem avoids indeterminate forms by imposing an integrability hypothesis, so certain integrals are all finite. Tonelli's Theorem allows infinite integrals but requires the integrand to be nonnegative, thereby again avoiding indeterminate forms.

Remark: The proofs in Section 4.6 are directly inspired by the presentation in [WZ77].

4.6.1 Fubini's Theorem

The general form of the proof of Fubini's Theorem follows a path that we have seen many times: Establish the result first for characteristic functions of boxes, then open sets, then sets of measure zero, then G_δ-sets, then arbitrary measurable sets. After that, extend to simple functions and finally to arbitrary integrable functions. However, the details are somewhat technical. I like to present the proof of Lemma 4.6.4 in full because it is a nice application of the Monotone Convergence Theorem. However, because of time constraints I often end up only sketching the proofs of some of the remaining lemmas, such as Lemma 4.6.5.

Theorem 4.6.1 (Fubini's Theorem). Motivate and state. I often use Problem 4.6.12 as motivation, because it's easy to see that the two iterated integrals in that problem are zero, yet $\iint |f| = \infty$. Another nice example is given in *Extra Problem 1* below, because in that problem the three possible integrals are all different.

Note: My instructor for graduate real analysis at the University of Maryland was Prof. Umberto Neri, who was a Ph.D. student of Alberto Calderón (1920–1998), who was a Ph.D. student of Antoni Zygmund (1900–1992) (so it is not surprising that our text was the book by Wheeden and Zygmund). Neri referred to Problem 4.6.12 as "Zygmund's example."

Note: "Fubini" has entered the mathematical lexicon as a verb, e.g., we say that "now we Fubini the integrals" to mean that we apply Fubini's Theorem to interchange the order of an iterated integral.

Lemma 4.6.2 and Lemma 4.6.3. These are easy and "obvious," so I state them but assign the proof as reading.

Lemma 4.6.4. State, prove part (a) and assign the proof of part (b) as reading.

Note: We cannot replace the hypothesis that $f_k \nearrow f$ (monotone increasing at every point) with $f_k \nearrow f$ a.e. We assume that $f_k \nearrow f$ at every point because this implies that for every y we have $f_k^y(x) \nearrow f^y(x)$ for every x. If we only assume that $f_k \nearrow f$ almost everywhere, then it is not easy to show (without using Tonelli's Theorem) that for a.e. y we have $f_k^y(x) \nearrow f^y(x)$ for a.e. x. If there was an easy way to do this then we could simplify the whole proof of Fubini's Theorem, but I don't see any easy way to do this.

Lemma 4.6.5. State. Usually I sketch the idea of the proof but do not give all of the details.

In the proof of this lemma, it may be more enlightening to discuss Step 2 first, as this gives some motivation for why the technical Step 1 is needed.

Note: I refer to the iteration

$$A_1 = Q_1, \qquad A_{k+1} = Q_{k+1} \setminus (Q_1 \cup \cdots \cup Q_k)$$

as the "Disjointization Trick." It's yet another one of those simple but extremely useful techniques that keep popping up.

Lemma 4.6.6. State and prove (or sketch).

Theorem 4.6.7. State and prove (or sketch).

4.6.2 Tonelli's Theorem

Theorem 4.6.8 (Tonelli's Theorem). State and prove.

Corollary 4.6.9. State and prove. In practice, when we want to use Fubini's Theorem we usually have to employ this corollary first in order to establish that the hypotheses of Fubini's Theorem are satisfied.

Lemma 4.6.10. I usually state and discuss this lemma, simply because variations on this result seem to arise with great regularity in daily mathematical life.

4.6.3 Convolution

Theorem 4.6.11. I like to present convolution in some detail. First, it is a beautiful application of Tonelli and Fubini and second, as a harmonic analyst convolution plays an extremely important role in the mathematics that I do every day.

Problems. *Note on Problem* 4.6.18 in the text: There are several other ways to evaluate this particular integral, such as contour integration.

Note on Problem 4.6.19 in the text: The improper integral can also be evaluated by other methods, such as contour integration.

Extra Problems for Section 4.6

1. As illustrated in Figure 4.B, for $x, y \ge 0$, define

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

Prove that

$$\int_0^\infty \int_0^\infty f(x,y)\,dy\,dx \;=\; 0,$$
$$\int_0^\infty \int_0^\infty f(x,y)\,dx\,dy \;=\; 1,$$
$$\iint_{[0,\infty)^2} |f(x,y)|\,(dx\,dy) \;=\; \infty.$$

2 Let E be a measurable subset of $\mathbb{R}$ such that $0 < |E| < \infty$. Prove that $E + E = \{x + y : x, y \in E\}$ contains an open interval centered at the origin.

3. (a) Define f on $\mathbb{R}^2$ by

$$f(x,y) \;=\; \begin{cases} 1/x^2, & \text{if } 0 < y < x < 1, \\ -1/y^2, & \text{if } 0 < x < y < 1, \\ 0, & \text{otherwise.} \end{cases}$$

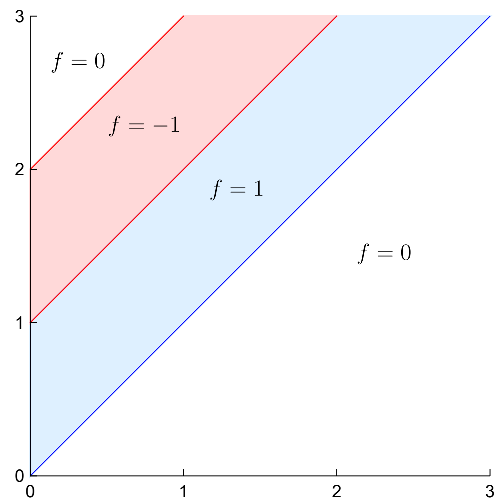

Fig. 4.B The function f from Extra Problem 1.

Compute $\iint |f|$ and the two iterated integrals of f. Do they exist? Are they equal? Are these results consistent with Fubini's Theorem?

(b) Similar to part (a), but on the domain $[-1,1]^2$ consider

$$\int_{-1}^{1}\int_{-1}^{1} \frac{xy}{(x^2+y^2)^2}\,dx\,dy.$$

(c) Similar to part (a), but on the domain $[-1,1]^2$ consider

$$\int_{-1}^{1}\int_{-1}^{1} \frac{x^2-y^2}{(x^2+y^2)^2}\,dx\,dy.$$

4. Let $g\colon \mathbb{R}\to\mathbb{R}$ be a nonmeasurable function, and define $f\colon \mathbb{R}^2\to\mathbb{R}$ by

$$f(x,y) \;=\; \begin{cases} g(x), & y\in\mathbb{Q},\\ e^{-|x|-|y|}, & y\notin\mathbb{Q}. \end{cases}$$

Is f measurable? Is f integrable?

5. Let $f \in L^1(\mathbb{R})$ be given, and define $g(x) = \displaystyle\int_{x-1}^{x} \frac{f(y)}{\sqrt{x-y}}\,dy$ for $x\in\mathbb{R}$. Prove directly (that is, from Fubini/Tonelli, rather than by appealing to Theorem 4.6.11) that g is defined for a.e. x, and g is measurable and integrable on $\mathbb{R}$.

6. Exhibit real scalars a_{mn} such that the following two series exist but satisfy

$$\sum_{m=1}^{\infty}\sum_{n=1}^{\infty} a_{mn} \neq \sum_{n=1}^{\infty}\sum_{m=1}^{\infty} a_{mn}.$$

7. (a) Prove that $f(x) = \displaystyle\int_{-\infty}^{\infty} \frac{1}{1+t^2}\,\frac{\sin(x-t)}{1+(x-t)^2}\,dt$ is continuous on $\mathbb{R}$.

(b) Is $F(x,t) = \dfrac{1}{1+t^2}\,\dfrac{\sin(x-t)}{1+(x-t)^2}$ integrable on $\mathbb{R}^2$?

8. Assume that $f \in L^1(\mathbb{R})$ satisfies $\|f\|_1 \le 1$. Define

$$g(x) = \int_{-\infty}^{\infty} \frac{f(y)}{1+|x-y|^2}\,dy.$$

Prove the following statements.

(a) g is continuous.

(b) $g(x) \to 0$ as $|x| \to \infty$.

(c) There exists a point x with $|x| < 100$ such that $|g(x)| < 1$.

9. Let $W(x) = \max\{1-|x|, 0\}$ be the "hat function" on $[-1,1]$. Given f in $L^1(\mathbb{R})$, let

$$g(y) = \int_{-\infty}^{\infty} f(t)\, e^{-2\pi i y t}\,dt, \qquad y \in \mathbb{R}.$$

Prove that g is bounded on $\mathbb{R}$, and for a.e. x we have

$$\int_{-\infty}^{\infty} f(y)\left(\frac{\sin \pi(x-y)}{\pi(x-y)}\right)^2 dy = \int_{-1}^{1} g(t)\,(1-|t|)\,e^{2\pi i t x}\,dt.$$

Hint: $\displaystyle\int_{-\infty}^{\infty} W(t)\, e^{2\pi i y t}\,dt = \left(\frac{\sin \pi y}{\pi y}\right)^2.$

10. Show that if f is measurable on $[0,1]$ and $f(x) - f(y)$ is integrable on $[0,1]^2$, then $f \in L^1[0,1]$.

11. Compute $\displaystyle\int_0^1 \int_y^1 x^{-3/2} \cos\left(\frac{xy}{2x}\right) dx\,dy.$

12. Let F be a closed subset of $(0,1)$, and set

$$M(x) = \int_0^1 \frac{\operatorname{dist}(y,F)}{|x-y|^2}\,dy, \qquad \text{for } x \in \mathbb{R}.$$

Prove that $M(x) = \infty$ for all $x \in F^{\mathrm{C}}$, but $M(x) < \infty$ for a.e. $x \in F$.

Hint: Consider $\int_F M(x)\,dx$.

13. Assume that $E \subseteq \mathbb{R}^d$ is measurable, $f \in L^1(E)$, and $g \in L^\infty(E)$ is nonnegative. Prove that $\displaystyle\int_E f(x)\,g(x)\,dx = \int_0^\infty \int_{\{g>t\}} f(x)\,dx\,dt.$

14. Let f be a bounded measurable function on a measurable set $E \subseteq \mathbb{R}^d$, and suppose that there exist some constants $C > 0$ and $0 < \alpha < 1$ such that $\bigl|\{|f| > t\}\bigr| \le Ct^{-\alpha}$ for all $t > 0$. Prove that f is integrable.

15. Prove that if $f,\, g \in L^1(\mathbb{R})$ are nonnegative a.e., then $\|f*g\|_1 = \|f\|_1\,\|g\|_1$.

16. Let E and F are measurable subsets of $\mathbb{R}$ with finite measure, and let A be an arbitrary measurable subset of $\mathbb{R}$. Prove the following statements.

(a) $|E \cap (F+x)| = (\chi_E * \chi_{-F})(x)$.

(b) $|E \cap (F+x)|$ is a continuous function of x. Hint: Problem 4.6.27.

(c) $\lim\limits_{x\to 0} |E \cap (F+x)| = |E \cap F|$.

(d) $\lim\limits_{x\to \infty} |E \cap (F+x)| = 0$.

(e) If $|E \cap (F+x)| = 0$ for almost every x, then either $|E| = 0$ or $|F| = 0$. Hint: Extra Problem 2.

(f) If $|A \backslash (A+x)| = 0$ for almost every x, then either $|A| = 0$ or $|A^{\mathrm{C}}| = 0$.

17. Suppose that A and B are Lebesgue measurable subsets of $[0,1]$, and $|A| = |B| = 1/2$. Prove that there exists at least one point $x \in [-1,1]$ such that $|(A+x) \cap B| \ge 1/10$.

18. (a) Given measurable sets $A,\, B \subseteq \mathbb{R}$ with finite measure, prove that

$$\lim_{t\to 0} |A \triangle (A+t)| \;=\; 0,$$

where $A \triangle B = (A\backslash B) \cup (B\backslash A)$ is the symmetric difference of A and B.

(b) Show by example that if $E \subseteq \mathbb{R}$ is measurable but has infinite measure, then we need not have $\lim\limits_{t\to 0} |E \triangle (E+t)| \;=\; 0$.

19. (a) Prove that $\displaystyle\int_0^\infty x^{2n}\, e^{-x^2}\, dx \;=\; \frac{(2n)!}{2^{2n}\, n!}\,\frac{\sqrt{\pi}}{2}$ for integer $n \ge 0$.

Hint: Induction; the base step is Problem 4.6.18.

(b) Prove that if $a > 0$ then $\displaystyle\int_{-\infty}^\infty e^{-x^2} \cos ax\, dx \;=\; \sqrt{\pi}\, e^{-a^2/4}$.

20. Assume $f \in L^1(0,\infty)$, and define

$$g(x) \;=\; \int_0^\infty \frac{f(y)}{x+y}\, dy, \qquad x > 0.$$

Prove that g is differentiable at every point $x > 0$, and if $a > 0$ then g' belongs to $L^1(a,\infty)$. Must g' be integrable on $(0,\infty)$?

CHAPTER 5: DIFFERENTIATION

This chapter contains a higher proportion of highly technical proofs than in most of the other chapters. In particular, this includes theorems in Sections 5.3, 5.4, and 5.5. This requires the instructor to make some difficult choices. In my experience, it is better to assign the less-enlightening proofs as reading than to attempt to present them in detail in class. I indicate my choices in this regard in the notes below; naturally you may feel differently about what is most important to discuss in the classroom.

***Note*:** A streamlined presentation of the main material from Chapters 5 and 6 can be found in the article:

> C. Heil, *Absolute Continuity and the Banach–Zaretsky Theorem*, in: "Excursions in Harmonic Analysis," Volume 6, M. Hirn et al., eds., Birkhäuser, Cham (2021), pp. 27–51.

A recording of a related video lecture on *Absolute Continuity and the Banach–Zaretsky Theorem* can be found at

`https://www.youtube.com/watch?v=YSwNcVhV18w` (5.A)

Section 5.1: The Cantor–Lebesgue Function

Definition 5.1.1 (Cantor–Lebesgue Function). Present the construction of the Cantor–Lebesgue function, culminating with this definition.

Theorem 5.1.2. State and prove.

Definition 5.1.3 (Singular Function). State.

Example 5.1.4. Discuss.

Note: The Hungarian composer Grörgy Ligeti (1923–2006) wrote a series of études for piano. Étude 13 is entitled *L'escalier du diable*, or *The Devil's Staircase*. You can find a variety of performances on the internet. I do not know if Ligeti was inspired by the Cantor–Lebesgue function or not, but when I listen to this piece I am certainly reminded of it!

Note: Here are some interesting facts about the Cantor–Lebesgue function that I do not usually present in class.

Just for fun, we mention that the reflected Cantor–Lebesgue function φ pictured in Figure 5.2 satisfies the following *refinement equation*:

$$\varphi(x) \,=\, \frac{1}{2}\varphi(3x) + \frac{1}{2}\varphi(3x-1) + \varphi(3x-2) + \frac{1}{2}\varphi(3x-3) + \frac{1}{2}\varphi(3x-4).$$

That is, φ equals a finite linear combination of compressed and translated copies of itself, and so its graph exhibits a type of self-similarity. Changing the refinement equation slightly gives some related functions. For example, it can be shown that there exists a continuous, compactly supported function f that satisfies the refinement equation

$$f(x) \,=\, \frac{2}{3}f(3x) + \frac{1}{3}f(3x-1) + f(3x-2) + \frac{1}{3}f(3x-3) + \frac{2}{3}f(3x-4).$$

This function f, which is pictured in Figure 5.A, was (so far as I am aware) first constructed by De Rham (but by a recursive procedure similar to the one that constructs the Cantor–Lebesgue function, not via a refinement equation). It is not differentiable at *any* point in $[0,2]$ (see [DL91], reference given below).

Refinement equations and their solutions (which are called *refinable functions*) play important roles in wavelet theory and in subdivision schemes in computer-aided graphics. For more information on refinement equations, we refer to sources such as [Dau92], [DL91] (reference below), or [Heil11, Sec. 12.5].

[DL91] I. Daubechies and J. C. Lagarias, Two-scale difference equations: I. Existence and global regularity of solutions, *SIAM J. Math. Anal.*, **22** (1991), pp. 1388–1410.

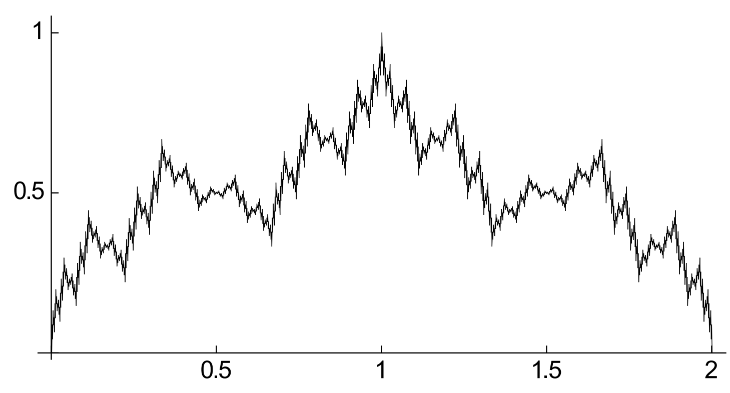

Fig. 5.A De Rham's nowhere differentiable function.

Extra Problems for Section 5.1

1. Use the refinement equations given above to plot the Cantor–Lebesgue function and de Rham's function. More generally, given $\alpha \in \mathbb{R}$ investigate solutions to the refinement equation

$$\begin{aligned} f(x) \;=\; &(\tfrac{1}{2}+\alpha)\, f(3x) \;+\; (\tfrac{1}{2}-\alpha)\, f(3x-1) \;+\; f(3x-2) \\ &+\; (\tfrac{1}{2}-\alpha)\, f(3x-3) + (\tfrac{1}{2}+\alpha)\, f(3x-4). \end{aligned}$$

Hint: We know $f(k)$ for k integer.

2. (a) Show that there exists a compactly supported (but discontinuous) function D_2 that satisfies the refinement equation

$$D_2(x) \;=\; D_2(2x) \;+\; D_2(2x-1).$$

(b) Show that there exists a continuous and compactly supported function D_4 that satisfies the refinement equation

$$D_4(x) = \tfrac{1+\sqrt{3}}{4} D_4(2x) + \tfrac{3+\sqrt{3}}{4} D_4(2x-1) + \tfrac{3-\sqrt{3}}{4} D_4(2x-2) + \tfrac{1-\sqrt{3}}{4} D_4(2x-3).$$

Note: Part (a) is easy; part (b) is challenging. The functions D_2 and D_4 are the first in the sequence of *Daubechies scaling functions*; see [Dau92] or [Heil11, Ch. 12].

3. This problem will give an alternative derivation of the Cantor–Lebesgue function. Let

$$X \;=\; \bigl\{f \in C[0,1] : f(0) = 0 \text{ and } f(1) = 1\bigr\}.$$

Although X is not a subspace of $C[0,1]$, it is a subset, and therefore

$$\mathrm{d_u}(f,g) \;=\; \|f-g\|_{\mathrm{u}} \;=\; \sup_{x\in[0,1]} |f(x)-g(x)|$$

defines a metric on X. Define an operator $T\colon X \to X$ by

$$(Tf)(x) \;=\; \begin{cases} \frac{1}{2} f(3x), & \text{if } 0 \le x \le \frac{1}{3}, \\ \frac{1}{2}, & \text{if } \frac{1}{3} \le x \le \frac{2}{3}, \\ \frac{1}{2} + \frac{1}{2} f(3x-2), & \text{if } \frac{2}{3} \le x \le 1. \end{cases}$$

Prove that T maps X into itself, and it is a contraction in the sense that

$$\|Tf - Tg\|_{\mathrm{u}} \;\le\; \frac{\|f-g\|_{\mathrm{u}}}{2}, \qquad \text{for all } f, g \in X.$$

As a consequence, the Banach Fixed Point Theorem (also known as the Contractive Mapping Theorem, see [Heil18, Prob. 2.9.20]), implies that T has a unique fixed point in X. Prove that the Cantor–Lebesgue function is that fixed point.

4. This problem is about an interesting nonlinear operator.

(a) Let

$$M = \{f \in C[0,1] : f(0) = 0 \text{ and } f(1) = 1\}.$$

Prove that M is a closed subset (but not a subspace) of $C[0,1]$, and therefore M is a complete metric space with respect to the uniform metric.

(b) Given $f \in M$, define Af by

$$Af(x) = \begin{cases} \frac{1}{2} f\left(\frac{x}{1-x}\right), & \text{if } 0 \le x < \frac{1}{2}, \\ 1 - \frac{1}{2} f\left(\frac{1-x}{x}\right), & \text{if } \frac{1}{2} \le x \le 1. \end{cases}$$

Prove that $Af \in M$, and therefore A is an operator that maps M into M.

(c) Show that

$$\|Af - Ag\|_{\mathrm{u}} \le \frac{\|f-g\|_{\mathrm{u}}}{2}, \qquad \text{for all } f, g \in M.$$

Consequently A is Lipschitz on M.

(d) Use the Banach Fixed Point Theorem to prove that there exists a unique function $m \in M$ such that $Am = m$. This function is called the *Minkowski question mark function* or the *slippery Devil's staircase*.

(e) Set $f_0(x) = x$, and define $f_{n+1} = Af_n$ for $n \in \mathbb{N}$. Prove that f_n converges uniformly to m (this fact was used to generate the approximation to m that appears in Figure 5.B).

Note: For more information on the Minkowski question mark function, see

W. Van Assche, Orthogonal polynomials for Minkowski's question mark function, *J. Comput. Appl. Math.*, **284** (2015), pp. 171–183.

Section 5.2: Functions of Bounded Variation

5.2.1 Definition and Examples

Definition 5.2.1 (Bounded Variation). Motivate and state.

Note: The idea of the arc length of a curve is mentioned in the text. This is distinct from the idea of bounded variation, but it is closely related. For details on arc length, the reader should look for articles dealing with *rectifiable curves* (for example, see [WZ77, Sec. 2.2]).

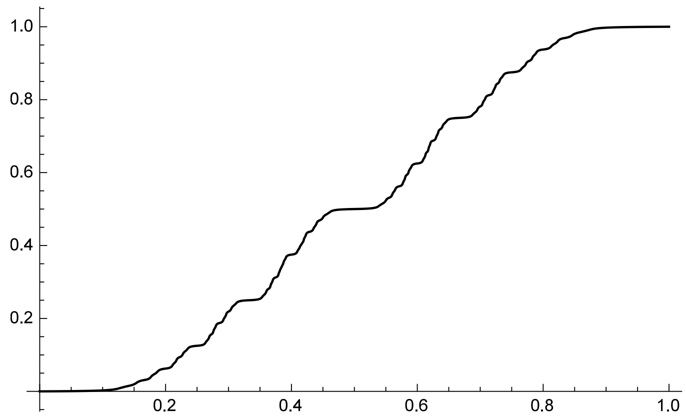

Fig. 5.B The function f_7 from Extra Problem 4, approximating the slippery Devil's staircase.

Note: We usually will write $V[f]$ for the total variation of f over $[a,b]$ when the interval $[a,b]$ is understood, and $V[f;a,b]$ when we need to be more explicit. This notation is inspired by the notation used in [WZ77].

Note: In Definition 5.2.1 we declare $\mathrm{BV}[a,b]$ to be the set of all functions $f\colon [a,b] \to \mathbb{C}$ that have bounded variation. This includes the real-valued functions of bounded variation as a subspace. A function with bounded variation must be finite at every point (that is, it cannot take the values $\pm\infty$), so there is nothing to be gained by considering functions $f\colon [a,b] \to \overline{\mathbf{F}}$ that have bounded variation.

Note: For functions whose domain is the entire real line $\mathbb{R}$ we usually define the following two types of bounded variation.

First, we say that $f\colon \mathbb{R} \to \mathbb{C}$ has *bounded variation on* $\mathbb{R}$ if

$$V[f;\mathbb{R}] \,=\, \sup_{a<b} V[f;a,b] \,<\, \infty,$$

and corresponding let

$$\mathrm{BV}(\mathbb{R}) \,=\, \bigl\{f\colon\mathbb{R}\to\mathbb{C} \,:\, V[f;\mathbb{R}] < \infty\bigr\}$$

be the space of functions with bounded variation on $\mathbb{R}$.

Second, we say that $f\colon \mathbb{R} \to \mathbb{C}$ has *locally bounded variation on* $\mathbb{R}$ if $V[f;a,b] < \infty$ for every finite interval $[a,b]$. The space of functions with locally bounded variation on $\mathbb{R}$ is

$$\mathrm{BV}_{\mathrm{loc}}(\mathbb{R}) \,=\, \bigl\{f\colon\mathbb{R}\to\mathbb{C} \,:\, V[f;a,b] < \infty \text{ for all } a<b\bigr\}.$$

Clearly $\mathrm{BV}(\mathbb{R}) \subseteq \mathrm{BV}_{\mathrm{loc}}(\mathbb{R})$. A function like $f(x) = x$ has locally bounded variation on $\mathbb{R}$, but does not have bounded variation on $\mathbb{R}$.

Defining bounded variation for functions in higher dimensions is a considerably more subtle issue; see the following references.

E. Giusti, *Minimal Surfaces and Functions of Bounded Variation*, Birkhäuser Verlag, Basel, 1984.

P. Góra and A. Boyarsky, On functions of bounded variation in higher dimensions, *Amer. Math. Monthly*, **99** (1992), 159–160.

Remark 5.2.2. Omit.

Example 5.2.3. Mention.

Exercise 5.2.4. Briefly discuss, the important point being that even differentiable functions can have unbounded variation.

5.2.2 Lipschitz and Hölder Continuous Functions

Lemma 5.2.5 and Corollary 5.2.6. State after defining Lipschitz functions. Briefly mention that this lemma and corollary follow from the Mean Value Theorem (but be careful for complex-valued functions; apply the MVT to the real and imaginary parts separately).

Note: Sometimes the terms "Lipschitz continuous" and "Hölder continuous" are used interchangeably. That is, for some authors a "Lipschitz continuous function" is what we consider a "Hölder continuous function" and vice versa.

Lemma 5.2.7. State; the proof is easy.

Note: While every Lipschitz function has bounded variation, if we fix any $0 < \alpha < 1$ then we can construct a function that is Hölder continuous with exponent α yet has unbounded variation (see Problem 5.2.22). On the other hand, some functions that are Hölder continuous but not Lipschitz do have bounded variation; for example, consider the Cantor–Lebesgue function or $x^{1/2}$ on the interval $[0, 1]$.

5.2.3 Indefinite Integrals and Antiderivatives

Exercise 5.2.8 (Simple Version of the FTC). State, ask the students to prove it themselves.

Lemma 5.2.9. State and prove. It's only a lemma because we will later (after considerable work) be able to greatly strengthen the conclusions.

Remark 5.2.10. I mention this briefly during the proof of Lemma 5.2.9.

Questions after Remark 5.2.10. Discuss, the point being that there is still much that we do not know about indefinite integrals, and it will take more work than might be expected at first glance before we can answer these questions.

Note: Absolute continuity is mentioned here and at several other points in this chapter. It is defined precisely in Chapter 6, but it may be appropriate to state the definition here, and to comment that the definition of absolute continuity is similar to but "more stringent" than the definition of uniform continuity.

5.2.4 The Jordan Decomposition

Exercise 5.2.11. State. This exercise, and the next two upcoming lemmas, will be very useful but the proofs are technical computations that are not worth spending classtime on.

Lemma 5.2.12. State, but assign the proof for reading.

Definition 5.2.13 (Positive and Negative Variation). State.

Note: For example, if S_Γ had the form

$$S_\Gamma \,=\, |5| + |-2| + |-8| + |3| \,=\, 18,$$

then we would have

$$S_\Gamma^+ \,=\, 5+0+0+3 = 8 \qquad \text{and} \qquad S_\Gamma^- \,=\, 0+2+8+0 = 10.$$

Lemma 5.2.14. By definition, $S_\Gamma^+ + S_\Gamma^- = S_\Gamma$ and $S_\Gamma^+ - S_\Gamma^- = f(b) - f(a)$. This lemma shows that the same relationships hold for the positive and negative variations. This is completely reasonable and the proof is a technical verification, so I motivate and state the lemma, but assign its proof for reading.

Theorem 5.2.15 (Jordan Decomposition). State and prove.

Corollary 5.2.16. Briefly state.

Extra Problems for Section 5.2

1. Let $f(x) = e^{2\pi i x}$. Directly compute $V[f;0,1]$ and $\int_0^1 |f'|$.

Note: Compare your result to Corollaries 5.4.3 and 6.4.5.

2. Suppose that $f \in \mathrm{BV}[a,b]$ and $f' \ge 0$ a.e. Must f be monotone increasing? What if f is differentiable everywhere on $[a,b]$ and $f'(x) \ge 0$ for every x?

3. Suppose that $f\colon [a,b] \to \mathbb{R}$ is continuous (but we do not assume that f has bounded variation). Given $0 < M < V[f;a,b]$, show that there exists a number $\delta > 0$ such that if Γ is any partition of $[a,b]$ with mesh size $|\Gamma| < \delta$, then $S_\Gamma \ge M$.

4. Prove that if scalars a_n satisfy $\sum\limits_{n=0}^{\infty} |a_n| < \infty$, then $f(x) = \sum\limits_{n=0}^{\infty} a_n x^n$ has bounded variation on $[-1,1]$.

5. Suppose that a function $f \in \mathrm{BV}[a,b]$ is continuous. Prove that $f = g - h$ where g, h are continuous and monotone increasing.

6. (From Benedetto and Czaja [BC09]). Define $f(x) = 0$ if $x = 0$ or x is irrational. If $x = p/q$ where p and q are relatively prime integers, then set $f(x) = 1/(p^2q^2)$. Prove that $f \in \mathrm{BV}[0,1]$ and hence f' exists a.e., yet f' does not exist on a dense subset of $[0,1]$.

7. (This is a more detailed version of part (e) of Problem 5.2.19 in the text.) Prove that if f and g belong to $\mathrm{BV}[a,b]$, then $fg \in \mathrm{BV}[a,b]$, and

$$V[fg;a,b] \;\le\; \|g\|_\infty \, V[f;a,b] \;+\; \|f\|_\infty \, V[g;a,b].$$

8. Taking $f = g$ in part (e) of Problem 5.2.19 in the text, or in the preceding Extra Problem 7, we see that if $f \in \mathrm{BV}[a,b]$ then $f^2 \in \mathrm{BV}[a,b]$, and

$$V[f^2;a,b] \;\le\; 2\,\|f\|_\infty \, V[f;a,b]^2.$$

Clearly equality holds if f is a constant function, for in this case both sides are zero. Equality need not hold if f is not constant, but this problem will show that the constant 2 is the best possible over all nonconstant functions, even if we require that $f(a) = 0$. In particular, it is *not* true that $V[f^2;a,b]$ need be less than or equal to $\|f\|_\infty \, V[f;a,b]^2$.

(a) Fix $0 < \varepsilon < 2$. Exhibit a nonconstant function $f \in \mathrm{BV}[a,b]$ such that

$$(2-\varepsilon)\,\|f\|_\infty \, V[f;a,b]^2 \;\le\; V[f^2;a,b] \;\le\; 2\|f\|_\infty \, V[f;a,b]^2.$$

(b) Fix $0 < \varepsilon < 2$. Exhibit a nonconstant function $f \in \mathrm{BV}[a,b]$ with $f(a) = 0$ such that

$$(2-\varepsilon)\,\|f\|_\infty \, V[f;a,b]^2 \;\le\; V[f^2;a,b] \;\le\; 2\,\|f\|_\infty \, V[f;a,b]^2.$$

(c) Is there a nonconstant function $f \in \mathrm{BV}[a,b]$ for which equality holds? That is, such that $V[f^2;a,b] \;=\; 2\,\|f\|_\infty \, V[f;a,b]^2$? Remark: I think not, but I don't have a proof.

Another remark: Compare the inequality $V[f^2;a,b] \le 2\,\|f\|_\infty \, V[f;a,b]^2$ obtained above with the inequality from Problem 6.4.16 in the text. In terms of variation functions, Problem 6.4.16 states that if $f \in \mathrm{AC}[a,b]$ and $f(a) = 0$ then $V[f^2;a,b] \le V[f;a,b]^2$.

9. Give an example that shows that the conclusion of Problem 5.2.27 in the text can fail if f is not continuous. In fact, it can fail if f has a single removable discontinuity and is continuous at all other points in $[a, b]$.

Note: Part (c) of Problem 5.2.27 is not that difficult for real-valued functions; the hint is to use the Mean-Value Theorem and the fact that f' is Riemann integrable. However, the Mean-Value Theorem does not hold for complex-valued functions, so more care is required for that case.

Section 5.3: Covering Lemmas

5.3.1 The Simple Vitali Lemma

Theorem 5.3.1 (Simple Vitali Lemma). This result is surprising at first glance, and the proof is quite elegant, so I like to state and prove it. The proof in the text is adapted from Folland's text [Fol99].

Note: There are variations on Theorem 5.3.1 that use coverings by closed balls or closed cubes instead of open balls, but the proofs are *not as elegant*. For example, Wheeden and Zygmund (see [WZ77, Lemma 7.4]) prove a result for cubes, but it takes considerably more work (although they do point out that the easy greedy algorithm proof will work for cubes *if we know that the set E is measurable*). It almost seems like you should be able to use the simple greedy approach when the sets are cubes or closed balls—just fatten them up a little to get open sets, which you can then reduce to finitely many sets just as we did in our proof of Theorem 5.3.1. The trouble with this approach seems to be that there's no easy way to correctly undo the fattening in the end.

5.3.2 The Vitali Covering Lemma

If time is short, which is often the case, then I omit discussion of the Vitali Covering Lemma. The proof is nice, but it is quite long and quite technical, and in this text this theorem is only needed for the proof that that monotone increasing functions are differentiable a.e. (which is a proof that I usually do not have time to cover in class).

Definition 5.3.2 (Vitali Cover). State if time permits.

Theorem 5.3.3 (Vitali Covering Lemma). State if time permits, but assign the proof for reading.

Note: By replacing $s_n/2$ with αs_n where $0 < \alpha < 1$, we can replace the constant 5 in the proof of Theorem 5.3.3 with $3 + \varepsilon$.

Note: We only proved that the Simple Vitali Lemma (Theorem 5.3.1) holds for open balls, but *if we accept that it also holds for closed balls* (which does require a nontrivial proof), then the following nice argument from Wheeden and Zygmund [WZ77] shows how to use that fact to give a different proof of Theorem 5.3.3.

Assume the following form of the Simple Vitali Lemma: There exists a constant $0 < \beta < 1$ (depending only on the dimension d) such that if $E \subseteq \mathbb{R}^d$ has finite exterior measure and $\mathcal{B}$ is a covering of E by closed balls, then there exist disjoint balls $B_1, \dots, B_N \in \mathcal{B}$ such that

$$\sum_{k=1}^{N} |B_k| \;>\; \beta\,|E|_e.$$

(Unfortunately, as remarked before, the proof of this fact is *not as easy* as our proof of Theorem 5.3.1.) Assuming this, we will give a proof of Theorem 5.3.3.

Proof of Theorem 5.3.3. Assume that $0 < |E|_e < \infty$ and let $\mathcal{B}$ be a Vitali covering of E. Let β be as given above, and without loss of generality assume that ε is small enough that $\varepsilon < \beta/2$. Let $C = 1-(\beta/2)$, and note that $0 < C < 1$.

Let $U \supseteq E$ be an open set such that $|U| < (1+\varepsilon)\,|E|_e$. Remove all balls from $\mathcal{B}$ that are not contained in U; this still leaves us with a Vitali cover of E. By the Simple Vitali Lemma for closed balls, we can find disjoint closed balls $B_1, \dots, B_{N_1}$ in $\mathcal{B}$ such that

$$\sum_{k=1}^{N_1} |B_k| \;>\; \beta\,|E|_e.$$

Then

$$\begin{aligned}
\biggl|E \setminus \bigcup_{k=1}^{N_1} B_k\biggr|_e \;\le\; \biggl|U \setminus \bigcup_{k=1}^{N_1} B_k\biggr| \;&=\; |U| \;-\; \sum_{k=1}^{N_1} |B_k| \\
&<\; (1+\varepsilon)\,|E|_e \;-\; \beta\,|E|_e \\
&=\; (1+\varepsilon-\beta)\,|E|_e \\
&<\; \left(1-\frac{\beta}{2}\right)|E|_e. \;=\; C\,|E|_e.
\end{aligned}$$

Let $E_1 \;=\; E \setminus \bigcup_{k=1}^{N_1} B_k$. Then

$$\bigl\{B \in \mathcal{B} \,:\, B \text{ is disjoint from } B_1, \dots, B_{N_1}\bigr\}$$

is a Vitali covering of E_1. By the Simple Vitali Lemma, there exist disjoint closed balls $B_{N_1+1}, \dots, B_{N_2}$ that are disjoint from $B_1, \dots, B_{N_1}$ such that

$$\sum_{k=N_1+1}^{N_2} |B_k| \;>\; \beta\,|E_1|_e.$$

Arguing similarly as before, we see that

$$\biggl|E \setminus \bigcup_{k=1}^{N_1} B_k\biggr|_e \;=\; \biggl|E_1 \setminus \bigcup_{k=N_1+1}^{N_2} B_k\biggr|_e \;<\; C\,|E_1|_e \;<\; C^2\,|E|_e.$$

Continuing in this way we obtains disjoint closed balls $B_1, B_2, \dots$ and integers $N_1 < N_2 < \cdots$ such that for each $m \in \mathbb{N}$ we have

$$\biggl|E \setminus \bigcup_{k=1}^{N_m} B_k\biggr|_e \;<\; C^m\,|E|_e. \qquad \square$$

Section 5.4: Differentiability of Monotone Functions

Definition 5.4.1 (Dini Numbers). Assign for reading; this is only relevant for part (b) of Theorem 5.4.2, which is the only part of that theorem that I do not present in class.

Theorem 5.4.2 (Differentiability of Monotone Increasing Functions). I state the theorem and prove parts (a), (c), and (d) in class.

This proof of part (b) relies on the Vitali Covering Lemma. It is an interesting proof, but it is technical and long, so in the interest of time and clarity of understanding I assign the proof of part (b) as reading.

Note: Part (a) of Theorem 5.4.2 is sometimes called the *Darboux–Froda Theorem*.

Note: It's quite fascinating to me that the proof of part (b) is difficult. Part (a) shows that a monotone function has at most countably many discontinuities—so just how complicated could a monotone increasing function be? Shouldn't it be differentiable at all but those countably many discontinuities? Well, no, it's more complicated than that. Things would be easier if the discontinuities were separated, but there's no reason that they have to be. For example, Problem 5.4.7 constructs a monotone increasing function that is discontinuous at every rational point!

Note: I give a streamlined discussion of the material in Chapters 5 and 6, including a discussion of why Theorem 5.4.2(b) is difficult, in the survey paper and video lecture listed in the boxed material in (5.A) at the beginning of the comments for this chapter.

Note: There are several approaches to the proof that monotone functions are differentiable a.e., and the reader may find it interesting to to compare the proofs given in different texts, such as [BBT97, Thm. 7.5], [Fol99, Thm. 3.23], [SS05, Thm. 3.3.14], or [WZ77, Thm. 7.5] (which is the inspiration for our proof). None of these proofs are simple or straightforward, in my opinion.

Note: If f is monotone increasing on $[a,b]$ then $\int_a^b f' \le f(b) - f(a)$, but strict inequality can hold. Problem 6.4.22 will show that if f is monotone increasing and we let E be the set of points in $[a,b]$ where f is differentiable, then $\int_a^b f' = |f(E)|_e$.

Corollary 5.4.3. State and prove.

Lemma 5.4.4. Assign for reading. (This proof is adapted from [WZ77].)

Note: Corollary 5.4.3 shows that if $f \in \mathrm{BV}[a,b]$, then we have the inequality $|f'| \le V'$ a.e. Wheeden and Zygmund (see [WZ77, Thm. 7.24]) use Lemma 5.4.4 to prove that if $f \in \mathrm{BV}[a,b]$ is *real-valued*, then $|f'| = V'$ almost everywhere. I have not been able to determine whether this fact still holds for complex-valued functions $f \in \mathrm{BV}[a,b]$. It is certainly true if f is absolutely continuous (we prove this later in Corollary 6.4.5), but what if a complex-valued f has bounded variation but is not absolutely continuous? I don't present this in class, but following is my exposition of the proof that appears in [WZ77, Thm. 7.24].

Theorem. If $f \in \mathrm{BV}[a,b]$ is *real-valued* and we set $V(x) = V[f;a,x]$ for $x \in [a,b]$, then V is differentiable a.e. and

$$V'(x) \;=\; |f'(x)|, \quad \text{for a.e. } x \in [a,b].$$

Proof. By definition, $V(b) = V[f;a,b]$ is the supremum of all sums S_Γ over all finite partitions Γ of $[a,b]$. Therefore, we can choose a sequence of partitions

$$\Gamma_k \;=\; \bigl\{a = x_0^k < x_1^k < \cdots < x_{m_k}^k = b\bigr\}$$

such that

$$0 \;\le\; V(b) - S_{\Gamma_k} \;<\; 2^{-k}, \qquad \text{for } k \in \mathbb{N},$$

where

$$S_{\Gamma_k} \;=\; \sum_{j=1}^{m_k} |f(x_j^k) - f(x_{j-1}^k)|.$$

Choose scalars c_j^k in such a way that

$$f_k(x) \;=\; \begin{cases} f(x) + c_j^k, & \text{if } x \in [x_{j-1}^k, x_j^k] \text{ and } f(x_j^k) \ge f(x_{j-1}^k), \\ -f(x) + c_j^k, & \text{if } x \in [x_{j-1}^k, x_j^k] \text{ and } f(x_j^k) < f(x_{j-1}^k), \end{cases}$$

is well-defined (takes a single value) at each point x_j^k for $j = 1, \dots, m_k$, and satisfies $f_k(a) = 0$. Then we have for each choice of j and k that

$$f_k(x_j^k) - f_k(x_{j-1}^k) \;=\; |f(x_j^k) - f(x_{j-1}^k)|.$$

Consequently,

$$S_{\Gamma_k} \;=\; \sum_{j=1}^{m_k} \bigl(f_k(x_j^k) - f_k(x_{j-1}^k)\bigr) \;=\; f_k(b) - f_k(a) \;=\; f_k(b).$$

In particular, we have for each $k \in \mathbb{N}$ that

$$0 \;\le\; V(b) - f_k(b) \;=\; V(b) - S_{\Gamma_k} \;<\; 2^{-k}.$$

We claim that for each fixed k, the function $V(x) - f_k(x)$ is an increasing function of x. To see this, suppose that $a \le x < y \le b$. If there is a single j such that $x,\, y \in [x_{j-1}^k, x_j^k]$, then

$$f_k(y) - f_k(x) \;=\; |f(y) - f(x)| \;\le\; V(y) - V(x).$$

On the other hand, if $x \in [x_{j-1}^k, x_j^k]$ and $y \in [x_{\ell-1}^k, x_\ell^k]$ with $j < \ell$, then

$$\begin{aligned}
f_k(y) - f_k(x) \;&=\; \bigl(f_k(y) - f_k(x_{\ell-1}^k)\bigr) \,+ \\
&\qquad \sum_{i=j+1}^{\ell-1} \bigl(f_k(x_i^k) - f_k(x_{i-1}^k)\bigr) \;+\; \bigl(f_k(x_j^k) - f_k(x)\bigr) \\
&\le\; \bigl(V(y) - V(x_{\ell-1}^k)\bigr) \,+ \\
&\qquad \sum_{i=j+1}^{\ell-1} \bigl(V(x_i^k) - V(x_{i-1}^k)\bigr) \;+\; \bigl(V(x_j^k) - V(x)\bigr) \\
&=\; V(y) - V(x).
\end{aligned}$$

In any case, we obtain $f_k(y) - f_k(x) \le V(y) - V(x)$, and therefore

$$V(x) - f_k(x) \;\le\; V(y) - f_k(y).$$

Hence $V(x) - f_k(x)$ is indeed increasing with x.

Therefore, for $x \in [a,b]$ we have that

$$0 \;=\; V(a) - f_k(a) \;\le\; V(x) - f_k(x) \;\le\; V(b) - f_k(b) \;<\; 2^{-k}.$$

Consequently,

$$0 \le \sum_{k=1}^{\infty} \bigl(V(x) - f_k(x)\bigr) \le \sum_{k=1}^{\infty} 2^{-k} < \infty.$$

Lemma 5.4.4 therefore implies that the series

$$\sum_{k=1}^{\infty} \bigl(V'(x) - f_k'(x)\bigr)$$

converges for almost every x. Hence $V(x) - f_k(x) \to 0$, and therefore $f_k'(x) \to V'(x)$, for almost every x. But since V is increasing we have $V'(x) \ge 0$ a.e., so

$$|f'(x)| = |f_k'(x)| \to |V'(x)| = V'(x) \quad \text{a.e.}$$

Thus $|f'(x)| = V'(x)$ a.e. $\square$

Extra Problems for Section 5.4

1. Let $f(x) = x\sin(1/x)$ for $x \ne 0$, and set $f(0) = 0$. Compute the Dini numbers of f at $x = 0$.

2. Given $f\colon [a,b] \to \mathbb{R}$ and $a < x < b$, prove that if all four Dini numbers of f are finite at x, then f is continuous at x. Must f be differentiable at x?

3. Assume $f\colon [a,b] \to \mathbb{R}$ has bounded variation. Suppose that there exist numbers α, $\beta > 0$ such that

$$\bigl|\{D^+ f > \alpha\}\bigr|_e > \beta.$$

Prove that $V[f;a,b] \ge \alpha\beta$.

Section 5.5: The Lebesgue Differentiation Theorem

Some difficult choices may need to be made at this point due to the lack of time. Although the Maximal Theorem is very important, especially in subsequent courses, I feel that it does not have the highest priority for presentation in *this* first course on real analysis. Therefore, if time is short (which it usually is), I significantly compress the remainder of Chapter 5 and cover only the following:

- State Lemma 5.5.1 as motivation.
- State Theorem 5.5.3, and prove if possible (because it is a nice application of convolution).

- State the Lebesgue Differentiation Theorem.

Naturally you may feel differently about priorities for the course. If time permits (which it actually has for me in some semesters), then I state and prove the remaining results of this chapter as indicated in the comments below.

Lemma 5.5.1. Motivate and state, but I usually assign the proof as reading.

Exercise 5.5.2. State.

5.5.1 L^1-Convergence of Averages

Theorem 5.5.3. State and prove. This result, which establishes the L^1-norm convergence of the averages $\widetilde{f}_h$, is motivation for the statement of the Lebesgue Differentiation Theorem, which concerns pointwise a.e. convergence of averages.

I usually briefly discuss the connection between averages and convolution, which is presented in the text after the proof of Theorem 5.5.3.

Note: Since the averages converge in L^1-norm, there is at least a subsequence $\{h_n\}_{n\in\mathbb{N}}$ such that $\widetilde{f}_{h_n} \to f$ pointwise a.e. However, it is usually difficult, if not impossible, to obtain anything stronger than the existence of a pointwise a.e. convergent subsequence from convergence in L^1-norm. This is one reason why the Lebesgue Differentiation Theorem is so surprising, since it says that if f is locally integrable then we have pointwise a.e. convergence of the full sequence of averages.

5.5.2 Locally Integrable Functions

Definition 5.5.4. State.

5.5.3 The Maximal Theorem

Definition 5.5.5 (Hardy–Littlewood Maximal Function). Motivate and state.

Theorem 5.5.6 (Maximal Theorem). Motivate, state, and prove.

5.5.4 The Lebesgue Differentiation Theorem

Theorem 5.5.7 (Lebesgue Differentiation Theorem). State and prove. (This proof is adapted from [Fol99].)

5.5.5 Lebesgue Points

Definition 5.5.8 (Lebesgue Points and the Lebesgue Set). State.

Definition 5.5.9 (Regularly Shrinking Family). State.

Theorem 5.5.10. State and prove.

Corollary 5.5.11. State and prove.

Extra Problems for Section 5.5

1. Add the following to Problem 5.5.20 in the text: How does this problem relate to Extra Problem 12 in the comments to Section 2.2?

2. Suppose that g is locally integrable on $\mathbb{R}$, and for all $r,\, s \in \mathbb{Q}$ with $r \neq 0$ we have that $\int_0^1 g(rx+s)\,dx = 0$. Prove that $g = 0$ a.e.

3. Suppose that $f \in L^1(\mathbb{R})$ is such that $\int_{-\infty}^{\infty} f\phi = 0$ for every integrable simple function ϕ that satisfies $\int_{-\infty}^{\infty} \phi = 0$. Prove that $f = 0$ a.e.

4. Let A be a measurable subset of $[0,1]$.

(a) Prove that if $|A| > 2/3$, then A contains an arithmetic progression of length 3 (that is, there exist $a,\, d \in \mathbb{R}$ such that $a,\, a+d,\, a+2d \in A$.
Hint: Consider $h(x) = \chi_A(x) + \chi_A(x+\frac{1}{3}) + \chi_A(x+\frac{2}{3})$.

(b) Use part (a) to prove that if $|A| > 0$, then A contains an arithmetic progression of length 3.

5. (From [SS05].) Given $f \in L^1_{\mathrm{loc}}(\mathbb{R}^d)$, prove that

$$\bigl|\{Mf > t\}\bigr| \le \frac{2 \cdot 3^d}{t} \int_{\{|f|>t/2\}} |f|, \qquad \text{for all } t > 0.$$

Hint: Apply the Maximal Theorem to the function that is equal to f when $|f| > t/2$ and is zero otherwise.

6. Let $E \subseteq \mathbb{R}$ be measurable with $0 < |E| < \infty$. For each $r > 0$, define $h_r(x) = \bigl|E \cap [x-r, x+r]\bigr|$ for $x \in \mathbb{R}$.

(a) Prove that h_r is continuous.

(b) Prove that there exists some radius $r_0 > 0$ such that if $0 < r < r_0$, then there exists an x such that $h_r(x) = r$. Hint: Consider $\lim\limits_{r \to 0} (1/r)\, h_r(x)$.

CHAPTER 6: ABSOLUTE CONTINUITY AND THE FUNDAMENTAL THEOREM OF CALCULUS

Section 6.1: Absolutely Continuous Functions

Definition 6.1.1 (Absolutely Continuous Function). Motivate and state.

Note: According to this definition, in order for f to be absolutely continuous, no matter how we distribute the subintervals $[a_j, b_j]$ in $[a,b]$, as long as they are nonoverlapping and their total length satisfies $\sum_j (b_j - a_j) < \delta$, then we must have $\sum_j |f(b_j) - f(a_j)| < \varepsilon$. This is much more restrictive than simply requiring that f be continuous on $[a,b]$.

Note: A function $f\colon \mathbb{R} \to \mathbb{C}$ is *locally absolutely continuous on* $\mathbb{R}$ if it is absolutely continuous on every finite interval $[a,b]$. The space of locally absolutely continuous functions is denoted by $\mathrm{AC}_{\mathrm{loc}}(\mathbb{R})$.

Example 6.1.2. Discuss. It may suffice to just sketch the idea with a picture.

Lemma 6.1.3. State and prove.

Example 6.1.4. Discuss.

6.1.1 Differentiability of Absolutely Continuous Functions

Corollary 6.1.5. State.

Lemma 6.1.6. State and prove.

Note: It seems “obvious” that $G' = g$, but we cannot prove this yet!

Extra Problems for Section 6.1

1. Given $a < b$, is the function G constructed in Problem 4.5.31(b) absolutely continuous on the finite interval $[a,b]$?

2. Suppose that $f\colon [a,b] \to [c,d]$ is absolutely continuous and $\varphi\colon [c,d] \to \mathbb{C}$ is Lipschitz. Prove that $\varphi \circ f \in \mathrm{AC}[a,b]$.

3. (a) Show that if $f \in \mathrm{AC}[a,b]$ and $1 \le p < \infty$, then $|f|^p \in \mathrm{AC}[a,b]$.

(b) Fix $0 < p < 1$ and let $f(x) = x \sin x^{-p}$ for $x > 0$ and $f(0) = 0$. Prove that $f \in \mathrm{AC}[0,1]$, but $|f|^p \notin \mathrm{BV}[0,1]$.

Section 6.2: Growth Lemmas

Lemma 6.2.1 (Growth Lemma I). State and prove.

Note: I find the two "Growth Lemmas" proved in this section to be quite elegant, yet they are not covered in most textbook developments of absolute continuity that I have seen. This first Growth Lemma relates the amount that a function f blows up the measure of a set to the size of its derivative f' on that set. This is quite intuitive—the derivative measures the rate of increase of f, so it seems natural that this should be related to how much the size of a set is increased when passed through f.

Note (typed in 2019): This lemma can be found in the paper [Var65] by Varberg. This seems to be the first published proof of this Growth Lemma, although Varberg himself states that he found it given as an exercise in a 1955 text [Nat55] by Natanson! Perhaps the result was known in some circles, but it is interesting to me that this basic result was not published until 1965, much later than what we usually think of as the "development period" for most of the other theorems presented in this text.

***Note Added*:** Recently (2020) I discovered that the text

J. J. Benedetto and W. Czaja, *Integration and Modern Analysis*, Birkhäuser, Boston, 2009.

states that a proof of the Growth Lemmas appears in a text by Saks that appeared in 1937. An English translation is available:

S. Saks, *Theory of the Integral*, Second revised edition, English translation by L. C. Young, Dover, New York, 1964.

Indeed, the two Growth Lemmas are Lemma VII.6.3 and Theorem VII.6.5 in Saks' text. I discuss this in somewhat more detail in the following paper, which gives a streamlined presentation of the main material from Chapters 5 and 6.

C. Heil, *Absolute Continuity and the Banach–Zaretsky Theorem*, in: "Excursions in Harmonic Analysis," Volume 6, M. Hirn et al., eds., Birkhäuser, Cham (2021), pp. 27–51.

A recording of a related video lecture can be found at

`https://www.youtube.com/watch?v=YSwNcVhV18w`

Note: One of the few textbooks that includes results related to the Growth Lemmas is Bruckner, Bruckner, and Thompson [BBT97]. They prove somewhat more general versions of the Growth Lemmas stated in terms of derived numbers instead of derivatives (e.g., see Lemma 7.9 in [BBT97]). Our proof of Growth Lemma I is inspired by the proof given in [BBT97].

Corollary 6.2.2. Mention briefly. The proof is nice but somewhat long, so I assign it as reading. It is only needed for a couple of problems and in Section

6.5, which I usually do not cover in class anyway. The two Growth Lemmas are the important results of this section.

Note: Here are some details of claims made in the proof of Corollary 6.2.2.

First we demonstrate that $D = \cup D_n$. We have $D_n \subseteq D$ by definition, so certainly $\cup D_n \subseteq D$. On the other hand, if we choose $x \in D$, then $f'(x)$ exists and is nonzero. Set $\varepsilon = |f'(x)|$. Since the derivative exists at x, there is a $\delta > 0$ such that

$$0 \,<\, |y-x| \,<\, \delta \implies \left|\frac{f(y)-f(x)}{y-x}\right| \,>\, \frac{\varepsilon}{2}.$$

Choose n large enough that

$$\frac{1}{n} < \delta \qquad \text{and} \qquad \frac{1}{n} < \frac{\varepsilon}{2}.$$

If we choose any y such that $0 < |y-x| < \frac{1}{n}$, then we have $|y-x| < \delta$, and therefore

$$\left|\frac{f(y)-f(x)}{y-x}\right| \,>\, \frac{\varepsilon}{2} \,>\, \frac{1}{n}.$$

Therefore $x \in D_n$. This shows that $D \subseteq \cup D_n$.

Second, we verify that $D_n \cap J = \cup_k A_k$. By definition, $A_k \subseteq D_n \cap J$ for every k. Conversely, choose any point $x \in D_n \cap J$. Since $D_n \subseteq D \subseteq E$, we have $f(x) \in f(E) \subseteq \cup Q_k$, so $f(x) \in Q_k$ for some k and therefore $x \in f^{-1}(Q_k)$ for that k. Therefore $x \in f^{-1}(Q_k) \cap D_n \cap J = A_k$.

Finally, we check that $f(A_k) \subseteq Q_k$. If we choose $y \in f(A_k)$, then $y = f(x)$ for some $x \in A_k$. By definition of A_k, this implies that $x \in f^{-1}(Q_k)$, and this tells us that $f(x) \in Q_k$. Hence $y = f(x) \in Q_k$.

Corollary 6.2.3. Mention briefly.

Lemma 6.2.4 (Growth Lemma II). State and prove.

Note: This proof includes another one of those "stupidly simple" but extremely useful tricks. For this proof it is basic arithmetic: the fact that $k = (k-1) + 1$ is used in the final string of inequalities!

Section 6.3: The Banach–Zaretsky Theorem

The Banach–Zaretsky Theorem is one of my favorite results. I believe that absolute continuity is absolutely fundamental to the study of real analysis, but in many courses the only thing we learn is that AC functions are the ones for which the Fundamental Theorem of Calculus holds. The Banach–Zaretsky Theorem shines another light on the subject: AC functions are essentially the continuous functions that map sets of measure zero to sets of measure zero, and measurable sets to measurable sets. AC functions do not do “strange” things like blowing up a set of measure zero into a set with positive measure.

Theorem 6.3.1 (Banach–Zaretsky Theorem). State and prove. This is a really beautiful theorem.

Note: In my opinion, the Banach–Zaretsky Theorem is *central* to classical analysis, yet it seems to be one of the most overlooked theorems in the standard developments of this field.

Note: Zaretsky is sometimes transliterated as Zarecki.

Note: The hypothesis that f maps sets of measure zero to sets of measure zero is referred to in [BBT97] as *Lusin's Condition*.

Note: If f is complex-valued, then the appropriate hypothesis is not that f maps sets of measure zero to sets of measure zero. The issue is that a subset of $\mathbb{C}$ can have measure zero even though its restriction to the real (or imaginary) axis has positive measure (compare Problem 6.3.5). Instead, the criterion is that both the real and imaginary parts of f should map sets of measure zero to sets of measure zero.

Corollary 6.3.2. State. This is an immediate consequence of results from Chapter 2.

Note: Example 5.1.4 showed that the Cantor–Lebesgue function φ maps a set with measure zero to a set with measure zero. Applying Theorem 6.3.1, this gives us another proof that φ is not absolutely continuous on $[0, 1]$.

Note: Even though an absolutely continuous function must map sets with measure zero to sets with measure zero, the inverse image $f^{-1}(Z)$ of a set with $|Z| = 0$ need not have measure zero. For example, if f is constant on $[a, b]$, say $f = c$, then f is absolutely continuous and $Z = \{c\}$ has measure zero, but $f^{-1}(Z) = [a, b]$ has positive measure.

Note: Even though an absolutely continuous function maps measurable sets to measurable sets, the inverse image $f^{-1}(E)$ of a Lebesgue measurable set under an absolutely continuous function f need not be measurable. For example, see the construction by Spătaru of an absolutely continuous, strictly increasing function $f\colon [0, 1] \to [0, 1]$ such that f^{-1} is not absolutely continuous:

S. Spătaru, An absolutely continuous function whose inverse function is not absolutely continuous, *Note Mat.*, **23** (2004/05), pp. 47–49.

Corollary 6.3.3. State and prove.

Note: We will see an application of Corollary 6.3.3 in the proof of Theorem 9.2.14, which concerns the decay and smoothness properties of the Fourier transform of a function in $L^1(\mathbb{R})$.

Note: Problem 6.3.8 asks for the following extension of Corollary 6.3.3: If f is continuous, is differentiable at all but *countably many* points of $[a, b]$, and $f' \in L^1[a, b]$, then f is absolutely continuous. This is actually easy to prove using the Banach–Zaretsky Theorem. Yet Folland (whose book is a favorite of mine, even though he does not present Banach–Zaretsky), states in his end of chapter notes (p. 110 of [Fol99]) that the result of Problem 6.3.8 "is a highly nontrivial theorem". The Cantor–Lebesgue function φ shows that this statement cannot be extended to functions that are not differentiable at uncountably many points.

Note: The function $|x|^{3/2} \sin \frac{1}{x}$ is differentiable everywhere on $[0, 1]$ and has an integrable but unbounded derivative. The "simpler" function $f(x) = x^{1/2}$ almost satisfies the same requirements. This function is differentiable everywhere on $[0, 1]$ *except* at the point $x = 0$, and its derivative $f'(x) = \frac{1}{2}x^{-1/2}$ is integrable on $[0, 1]$. Problem 6.3.8 therefore implies that f absolutely continuous on $[0, 1]$.

Corollary 6.3.4 (AC + Singular Implies Constant). State and prove.

Note: One standard proof of this corollary uses the Vitali Covering Lemma (details below). I think the Banach–Zaretsky proof is much more enlightening (especially on the first encounter). The Vitali Covering Lemma does still play a role, in that it was needed in Chapter 5 to prove that monotone increasing functions are differentiable a.e., from which it follows that BV and AC functions are differentiable a.e.

Following is a proof of Corollary 6.3.4 based directly on the Vitali Covering Lemma (this exposition is adapted from [WZ77, Thm. 7.28]).

Corollary. If $f\colon [a, b] \to \mathbb{C}$ is both absolutely continuous and singular, then f is constant.

Proof. Suppose that f is both absolutely continuous and singular. We will show that $f(a) = f(b)$. Since the same argument can be applied to any subinterval of $[a, b]$, it follows from this that f is constant.

Since f is singular, $E = \big\{x \in (a, b) : f'(x) = 0\big\}$ is a set of full measure, meaning that $|E| = b - a$.

Suppose that $x \in E$, and fix any $\varepsilon > 0$. Then, since $f'(x) = 0$, we can find a point $y_x > x$ such that we have both $[x, y_x] \subseteq (a, b)$ and

$$x < y < y_x \quad \Longrightarrow \quad \frac{|f(y) - f(x)|}{y - x} < \varepsilon.$$

Then

$$\mathcal{B} = \big\{[x, y] : x \in E \text{ and } x < y < y_x\big\}$$

is a Vitali cover of E by closed intervals (compare Definition 5.3.2).

By hypothesis, f is absolutely continuous. Fix $\varepsilon > 0$, and let δ be the number corresponding to ε that is given by the definition of absolute continuity (Definition 6.1.1). Applying the Vitali Covering Lemma (see Theorem 5.3.3 and the inequalities that follow it), there exist finitely many disjoint intervals $\{[x_j, y_j]\}_{j=1}^N$ belonging to $\mathcal{B}$ such that

$$\sum_{j=1}^N (y_j - x_j) \;>\; (b-a) - \delta. \qquad\qquad \text{(A)}$$

Note that the fact that $[x_j, y_j] \in \mathcal{B}$ implies that

$$\frac{|f(y_j) - f(x_j)|}{y_j - x_j} \;<\; \varepsilon, \qquad \text{for } j = 1, \dots, N. \qquad\qquad \text{(B)}$$

Set $y_0 = a$ and $x_{N+1} = b$. Then we have

$$a \;=\; y_0 \;\le\; x_1 \;<\; y_1 \;<\; x_2 \;<\; \cdots \;<\; y_{N-1} \;<\; x_N \;<\; y_N \;\le\; x_{N+1} \;=\; b.$$

Considering equation (A), we conclude that

$$\sum_{j=0}^N (x_{j+1} - y_j) \;<\; \delta. \qquad\qquad \text{(C)}$$

Since f is absolutely continuous, it follows from equation (C) that

$$\sum_{j=0}^N |f(x_{j+1}) - f(y_j)| \;<\; \varepsilon.$$

On the other hand, equation (B) implies that

$$\sum_{j=1}^N |f(y_j) - f(x_j)| \;<\; \varepsilon \sum_{j=1}^N (y_j - x_j) \;\le\; \varepsilon\,(b-a).$$

Hence

$$\begin{aligned} |f(b) - f(a)| \;&\le\; \sum_{j=0}^N |f(x_{j+1}) - f(y_j)| \;+\; \sum_{j=1}^N |f(y_j) - f(x_j)| \\ &\le\; \varepsilon \;+\; \varepsilon\,(b-a). \end{aligned}$$

Since ε is arbitrary, we conclude that $f(a) = f(b)$. $\square$

Important Extra Remark: Failure of the MVT for AC functions. Absolutely continuous functions have many attractive properties. However, it is not true that every theorem from undergraduate calculus about differentiable functions extends to absolutely continuous functions. In particular, the following example shows that *the Mean Value Theorem can fail for absolutely continuous functions*. I usually mention this in class briefly.

The Mean Value Theorem states that if f is a continuous real-valued function on a closed interval $[a,b]$ and f is differentiable everywhere on (a,b), then there exists a point $c \in (a,b)$ such that

$$f'(c) \;=\; \frac{f(b)-f(a)}{b-a}.$$

We cannot relax the hypotheses to assume just absolute continuity, or even to just Lipschitzness, in this theorem. For example, $f(x) = |x|$ is Lipschitz and therefore absolutely continuous on $[-1,1]$. The slope of the secant line that joins the point $\bigl(-1, f(-1)\bigr)$ to $\bigl(1, f(1)\bigr)$ is zero, but there is no point $c \in (-1,1)$ such that $f'(c) = 0$.

Extra Problems for Section 6.3

1. Suppose that funjctions $f, g \in \mathrm{AC}[a,b]$ are such that $f' = g'$ a.e. Prove that $f = g + c$ for some constant c.

2. Prove that the function g defined in Problem 1.4.4(d) is absolutely continuous on $[0, \frac{1}{2}]$, even though it is not Hölder continuous on that interval for any positive exponent.

3. Assume that $f\colon \mathbb{R} \to \mathbb{R}$ is continuous, and $g\colon [a,b] \to \mathbb{R}$ is absolutely continuous. Prove that $G(x) = \int_0^{g(x)} f(t)\,dt$ is absolutely continuous on $[a,b]$.

(b) Does part (a) still hold if we only assume that f is integrable? Either prove that it does or give a counterexample.

4. Assume $f_0 \in L^1[0,1]$ is nonnegative, and for each integer $n \ge 0$ define

$$f_{n+1}(x) \;=\; \biggl(\int_0^x f_n(t)\,dt\biggr)^{1/2}, \qquad \text{for } x \in [0,1].$$

Assume that $f_1(x) \le f_0(x)$ for every $x \in [0,1]$.

(a) Prove that for each $x \in [0,1]$, the sequence $\{f_n(x)\}_{n\in\mathbb{N}}$ converges monotonically to a nonnegative number $f(x)$.

(b) Prove that f is integrable on $[0,1]$, and $f(x) = \left(\int_0^x f(t)\,dt\right)^{1/2}$ for $x \in [0,1]$.

(c) Prove that if $x \in [0,1]$ and $f(x) > 0$, then f is differentiable at x. What is $f'(x)$?

(d) Assuming $f(x) > 0$ for every $x \in (0,1]$, find an explicit simple formula for f.

Section 6.4: The Fundamental Theorem of Calculus

Lemma 6.4.1. State and prove.

Theorem 6.4.2 (Fundamental Theorem of Calculus). State and prove. Together with the Banach–Zaretsky Theorem, we now see many reasons why absolute continuity is important.

6.4.1 Applications of the FTC

Corollary 6.4.3. State and prove.

Theorem 6.4.4. State and prove.

Corollary 6.4.5. State and prove.

6.4.2 Integration by Parts

Theorem 6.4.6 (Integration by Parts). State and prove (or perhaps just say that it follows by applying absolute continuity to the product rule, and assign the proof for reading).

Theorem 6.4.7. This is a useful theorem and the proof is a nice application of absolute continuity and integration by parts, but time constraints usually require me to assign the theorem and its proof for reading.

Remark: An Extra Theorem. I usually do not prove this in class, but the heart of the proof of the Classical Uncertainty Principle is a nice application of absolute continuity and integration by parts. The mathematical formulation of this result is often stated in terms of the Fourier transform of functions on $L^2(\mathbb{R})$, but here is a version that omits mention of the Fourier transform. The proof does use the Cauchy–Bunyakovski–Schwarz (CBS) Inequality, which is the special case of Hölder's Inequality for $p = 2$. Hölder's Inequality is proved in Chapter 7, and the CBS Inequality is proved for abstract Hilbert spaces in Chapter 8. The L^2-norm of a measurable function f is $\|f\|_2 = \left(\int |f(x)|^2\,dx\right)^{1/2}$, and $L^2(\mathbb{R})$ is the set of measurable functions whose L^2-norm is finite. One consequence of the CBS Inequality is that if g, $h \in L^2[a,b]$ then their product gh is integrable on $[a,b]$. Our convention in this discussion is that $\infty^{1/2} = \infty$. A simpler version of this result where the domain is a finite interval $[a,b]$ is given in Extra Problem 12 below.

Theorem (Heart of the Uncertainty Principle). Assume that $f \in L^2(\mathbb{R})$ is locally absolutely continuous. That is, assume that $f \in \mathrm{AC}[a,b]$ for every finite interval $[a,b]$ (note that this implies that f' exists a.e.). Then

$$\|f\|_2^2 \,\le\, 2\,\|xf(x)\|_2\,\|f'\|_2,$$

where the right-hand side of this inequality could be infinite. Explicitly,

$$\int_{-\infty}^{\infty} |f(x)|^2 \,\le\, 2\biggl(\int_{-\infty}^{\infty} |xf(x)|^2\biggr)^{1/2}\biggl(\int_{-\infty}^{\infty} |f'(x)|^2\biggr)^{1/2}.$$

Proof. If the right-hand side is infinite then there is nothing to prove, so we may assume that $xf(x)$ and f' are both square-integrable. As a consequence, the product $xf(x)\,\overline{f'(x)}$ is integrable.

The product of two absolutely continuous functions is absolutely continuous, so $u(x) = xf(x)$ is absolutely continuous. At any point where f is differentiable, which is almost everywhere, the Product Rule implies that

$$u'(x) \,=\, xf'(x) \,+\, f(x). \qquad \text{(A)}$$

The function $v(x) = \overline{f(x)}$ is also absolutely continuous. Since integration by parts holds for absolutely continuous functions, we therefore compute that for any $a < b$ we have

$$\int_a^b xf(x)\,\overline{f'(x)}\,dx \,=\, \int_a^b u(x)\,v'(x)\,dx$$

$$=\, u(b)\,v(b) \,-\, u(a)\,v(a) \,-\, \int_a^b u'(x)\,v(x)\,dx \qquad \text{(integration by parts)}$$

$$=\, b\,|f(b)|^2 \,-\, a\,|f(a)|^2 \,-\, \int_a^b \bigl(xf'(x) + f(x)\bigr)\,\overline{f(x)}\,dx \qquad \text{(by equation (A))}$$

$$=\, b\,|f(b)|^2 \,-\, a\,|f(a)|^2 \,-\, \int_a^b xf'(x)\,\overline{f(x)}\,dx \,-\, \int_a^b |f(x)|^2\,dx.$$

Rearranging, we see that

$$2\,\mathrm{Re}\biggl(\int_a^b xf(x)\,\overline{f'(x)}\,dx\biggr) \,=\, \int_a^b xf(x)\,f'(x)\,dx \,+\, \int_a^b xf'(x)\,\overline{f(x)}\,dx$$

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

Since $xf(x)\,\overline{f'(x)}$ and $|f|^2$ are both integrable, if we fix a then the following limit exists:

$$\begin{aligned}
\lim_{b\to\infty} b\,|f(b)|^2 \\
&= \lim_{b\to\infty} \biggl(2\,\mathrm{Re}\biggl(\int_a^b xf(x)\,\overline{f'(x)}\,dx\biggr) \,+\, \int_a^b |f(x)|^2\,dx \,+\, a\,|f(a)|^2\biggr) \\
&= 2\,\mathrm{Re}\biggl(\int_a^\infty xf(x)\,\overline{f'(x)}\,dx\biggr) \,+\, \int_a^\infty |f(x)|^2\,dx \,+\, a\,|f(a)|^2.
\end{aligned}$$

But $f \in L^2(\mathbb{R})$, so the only way that the limit of $b\,|f(b)|^2$ can exist is if it is zero. Similarly, the limit of $a\,|f(a)|^2$ as $a \to -\infty$ exists, and therefore it must be zero as well. Consequently, by applying the Cauchy–Bunyakovski–Schwarz Inequality, we see that

$$\begin{aligned}
\|f\|_2^2 &= \int_{-\infty}^{\infty} |f(x)|^2\,dx \\
&= -2\,\mathrm{Re}\biggl(\int_{-\infty}^{\infty} xf(x)\,\overline{f'(x)}\,dx\biggr) \\
&\le 2\int_{-\infty}^{\infty} |xf(x)f'(x)|\,dx && (\mathrm{Re}(-z) \le |z|) \\
&= 2\,\|xf(x)f'(x)\|_1 \\
&\le 2\,\|xf(x)\|_2\,\|f'\|_2 && \text{(CBS Inequality)}. \qquad \square
\end{aligned}$$

Note: The main ideas of the proof just given appear in a 1928 text (in German) by Hermann Weyl (1885–1955); also see the English translation of Appendix 1 of Weyl's book in

H. Weyl, *The Theory of Groups and Quantum Mechanics*, Dover, New York, 1949.

In particular, although he does not justify its use and does not mention absolute continuity, the key point in Weyl's proof is integration of $x\,(f\bar{f})'(x)$ by parts. For more details, see the discussion in my forthcoming text on harmonic analysis.

Extra Problems for Section 6.4

1. Assume that f is Lipschitz on $[a,b]$, $f(a) = f(b) = 0$, and $\int_a^b f(x)^2\,dx = 1$.

(a) Prove that $\operatorname{Re}\biggl(\int_a^b x\,f(x)\,\overline{f'(x)}\,dx\biggr) \;=\; -\frac{1}{2}.$

(b) Use the Cauchy–Bunyakovski–Schwarz (CBS) Inequality to prove that

$$\biggl(\int_a^b |x f(x)|^2\,dx\biggr)\biggl(\int_a^b |f'(x)|^2\,dx\biggr) \;\ge\; \frac{1}{4}.$$

2. Prove that if $f \in \mathrm{AC}[a,b]$ and $f(a) = 0$, then $\|f\|_{\mathrm{u}} \le \|f'\|_1$. Show by example that the hypothesis of absolute continuity is necessary.

3. This problem gives an alternative (and more difficult) proof of integration by parts for absolutely continuous functions, based on Fubini's Theorem.

Fix $f, g \in \mathrm{AC}[a,b]$, and set $E = \bigl\{(x,y) \in [a,b]^2 : x \le y\bigr\}$. Use Fubini's Theorem to show that the two iterated integrals

$$\iint_E f'(x)\,g'(y)\,dx\,dy \;=\; \int_a^b \biggl(\int_a^y f'(x)\,dx\biggr)\,g'(y)\,dy$$

and

$$\iint_E f'(x)\,g'(y)\,dy\,dx \;=\; \int_a^b f'(x)\,\biggl(\int_x^b g'(y)\,dy\biggr)\,dx$$

are equal. Use this to prove that equation (6.12) (that is, the integration by parts formula) holds.

Note: The method of this exercise can also be used to prove a more general version of integration by parts that employs Riemann–Stieltjes integrals. For more details, see [Fol99, Thm. 3.36] or [WZ77, Thm. 7.32].

4. Given a monotone increasing function f on $[a,b]$, prove that the following two statements are equivalent.

(a) $f \in \mathrm{AC}[a,b]$.

(b) For every function $g \in \mathrm{AC}[a,b]$ and every $x \in [a,b]$,

$$\int_a^x f(t)\,g'(t)\,dt \;+\; \int_a^x f'(t)\,g(t)\,dt \;=\; f(x)g(x) - f(a)g(a).$$

5. Assume that f is monotone increasing on $[a,b]$ and $f(a) = 0$. What can you determine about the following functions?

(a) f', (b) $F(x) = \int_a^x f'(t)\,dt$, (c) F', (d) $G(x) = \int_a^x F'(t)\,dt$.

6. Assume that functions $f_n \in \mathrm{AC}[a,b]$ are such that:

- each f_n is monotone increasing on $[a,b]$,
- $f_n(a) = 0$ for every n,
- $\sup_n f_n(b) < \infty$, and

- the sequence $\{f_n'(x)\}_{n\in\mathbb{N}}$ is monotone increasing for a.e. x.

Prove that there exists an absolutely continuous function f on $[a,b]$ such that $f_n \to f$ uniformly.

7. Suppose that $\{f_n\}_{n\in\mathbb{N}}$ is a sequence of absolutely continuous functions on $[a,b]$ such that $f_n(a)=0$ for every n and $\{f_n'\}_{n\in\mathbb{N}}$ is Cauchy in L^1-norm. Prove that there exists an absolutely continuous function f on $[a,b]$ such that $f_n \to f$ uniformly.

8. Suppose that functions $f_n \in \mathrm{AC}[a,b]$ are such that $f_n(a)=0$ for every n and $\sum\limits_{n=1}^{\infty} \|f_n\|_1 < \infty$. Show that:

(a) The series $f(x) = \sum\limits_{n=1}^{\infty} f_n(x)$ converges for every x,

(b) $f \in \mathrm{AC}[a,b]$, and

(c) $f'(x) = \sum\limits_{n=1}^{\infty} f_n'(x)$ a.e.

9. Assume that $f\colon \mathbb{R} \to [0,\infty)$ is measurable, and $\varphi\colon [0,\infty) \to [0,\infty)$ is monotonic increasing and absolutely continuous on every interval $[0,b]$ with $b>0$. Prove that if $\varphi(0)=0$, then

$$\int_{-\infty}^{\infty} (\varphi \circ f)(x)\,dx \;=\; \int_0^{\infty} |\{f>t\}|\,\varphi'(t)\,dt.$$

10. Given $f \in C^1[a,b]$ and $\varepsilon > 0$, prove that there exists a polynomial p such that $\|f-p\|_{C^1} \;=\; \|f-p\|_{\mathrm{u}} + \|f'-p'\|_{\mathrm{u}} \;<\; \varepsilon$.

11. Assume that $f\colon \mathbb{R} \to \mathbb{R}$ is monotone increasing, $\lim\limits_{x\to-\infty} f(x) = 0$, and $\lim\limits_{x\to\infty} f(x) = 1$. Prove that f is absolutely continuous on every finite interval $[a,b]$ if and only if $\int_{-\infty}^{\infty} f'(x)\,dx = 1$.

12. Let X be the set of all absolutely continuous functions $f \in \mathrm{AC}[a,b]$ such that $f' \in L^1[a,b]$.

(a) Prove that

$$\|f\| \;=\; \int_a^b |f(x)|\,dx \;+\; \int_a^b |f'(x)|\,dx$$

is a norm on X, and X is a Banach space with respect to this norm.

(b) Show that

$$|||f||| \;=\; |f(a)| \;+\; \int_a^b |f'(x)|\,dx$$

is a norm on X that defines the same convergence criterion as $\|\cdot\|$; that is, it is a norm and if $f_n, f \in X$ then $\|f - f_n\| \to 0$ if and only if $|||f - f_n||| \to 0$.

(c) Show that $\|\cdot\|$ and $|||\cdot|||$ are *equivalent norms* on X; that is, there exist constants $A, B > 0$ such that $A\,\|f\| \le |||f||| \le B\,\|f\|$ for every $f \in X$.

13. Prove that if $f \in \mathrm{AC}[a,b]$ satisfies $f(a) = 0$, then

$$V[f^2; a, b] \,\le\, V[f; a, b]^2.$$

14. Given a function $f\colon [0,1] \to \mathbb{C}$, prove that f is absolutely continuous on $[0,1]$ if and only if there exist Lipschitz functions f_n on $[0,1]$ such that $V[f - f_n] \to 0$ as $n \to \infty$.

Section 6.5: The Chain Rule and Changes of Variable

The results in this section are elegant applications of absolute continuity, and have practical application to situations that requires a nonlinear change of variable. Unfortunately, there isn't time to cover every nice result, so some choices have to be made. Therefore I usually, although reluctantly, assign the material in this section as reading.

Extra Problems for Section 6.5

1. Assume that $f\colon \mathbb{R} \to \mathbb{R}$ is continuous, and $g\colon [a,b] \to \mathbb{R}$ is absolutely continuous. Prove that

$$G(x) \,=\, \int_0^{g(x)} f(t)\,dt$$

is absolutely continuous on $[a,b]$, and find G'.

(b) Does part (a) still hold if we only assume that f is integrable? Either prove that it does or give a counterexample.

Section 6.6: Convex Functions and Jensen's Inequality

Jensen's Inequality is also an elegant result, but it does not play much of a role in this text. Therefore I assign this section as reading, and do not cover it in class.

Note: Our proof of Theorem 6.6.10 is adapted from the proof in Folland's text [Fol99].

Problems. **TYPO** *in Problem* 6.6.14 in the text: Replace "$f\colon E \to \mathbb{R}$ is measurable" with "$f\colon E \to \mathbb{R}$ is integrable".

Extra Problems for Section 6.6

1. Suppose that $\phi\colon (a,b) \to \mathbb{R}$ is convex. Prove that $\{\phi < \alpha\}$ is a convex set for each $\alpha \in \mathbb{R}$.

2. Prove that ϕ is convex on (a,b) if and only if ϕ is convex on (c,d) for all $a < c < d < b$.

3. Suppose that $\phi\colon \mathbb{R} \to \mathbb{R}$ satisfies

$$\phi\biggl(\int_0^1 f(x)\,dx\biggr) \;\le\; \int_0^1 \phi(f(x))\,dx$$

for every bounded measurable function $f\colon [0,1] \to \mathbb{R}$. Prove that ϕ is convex on $\mathbb{R}$.

4. (a) Prove that $\varphi(x) = x \ln x$ is convex on $(0,\infty)$.

(b) Let E be a measurable subset of $\mathbb{R}^d$ such that $|E| < \infty$. Suppose that $f\colon E \to (0,\infty)$ is a measurable function that satisfies $\dfrac{1}{|E|}\displaystyle\int_E f = 1$. Prove that

$$\frac{1}{|E|}\int_E f(x)\,\ln f(x)\,dx \;\ge\; 0.$$

5. Fix $2 \le p < \infty$, and prove the following statements.

(a) $f(t) = t^{p/2}$ is convex on $(0,1)$.

(b) $a^p + b^p \le (a^2+b^2)^{p/2}$ for all $a, b \ge 0$.

(c) If $0 \le x \le 1$, then

$$\biggl(\frac{1-x}{2}\biggr)^p \;+\; \biggl(\frac{1+x}{2}\biggr)^p \;\le\; \frac{1+x^p}{2}.$$

(d) If $a, b \in \mathbb{R}$, then

$$\biggl(\frac{|a-b|}{2}\biggr)^p \;+\; \biggl(\frac{|a+b|}{2}\biggr)^p \;\le\; \frac{|a|^p+|b|^p}{2}.$$

(e) If $f, g \in L^p(E)$, then

$$\biggl\|\frac{f-g}{2}\biggr\|_p^p \;+\; \biggl\|\frac{f+g}{2}\biggr\|_p^p \;\le\; \frac{1}{2}\,\|f\|_p^p \;+\; \frac{1}{2}\,\|f\|_p^p.$$

(f) If $0 \le x \le 1$, then $\;2\,(1+x^p) \;\le\; (1+x)^p \;+\; (1-x)^p$.

(g) If $f, g \in L^p(E)$, then

$$2\,\|f/2\|_p^p \;+\; 2\,\|g/2\|_p^p \;\le\; \biggl\|\frac{f-g}{2}\biggr\|_p^p \;+\; \biggl\|\frac{f+g}{2}\biggr\|_p^p.$$

CHAPTER 7: The L^p Spaces

Note: The ℓ^p spaces are covered in more detail in the online **Alternative Chapter 1**, and additional problems and exercises can be found there.

Section 7.1: The ℓ^p Spaces

Definition 7.1.1 (p-Summable and Bounded Sequences). State.

Definition 7.1.2 (The ℓ^p Spaces). State.

Note: We often pronounce ℓ^p as "little ℓ^p" (and think of it familiarly as "baby ℓ^p"). This is to emphasize that we are referring to the sequence space ℓ^p rather than the Lebesgue space $L^p(E)$ that will be defined in Section 7.2. In situations where we need to consider ℓ^p and $L^p(E)$ at the same time, we often emphasize the distinction by referring to L^p as "Big L^p." When L^p appears alone, we usually do not say "Big," but somehow it seems natural to say "little ℓ^p" even when the only space we are discussing is ℓ^p.

Remark 7.1.3. Briefly mention.

Lemma 7.1.4. I usually mention this lemma in class but do not take time to write it down.

Note: In statement (b) of the lemma, if $\|x\|_p = \infty$ or $\|y\|_p = \infty$ then the right-hand side of the inequality is ∞, and so there is nothing to prove.

7.1.1 Hölder's Inequality

Define the dual index.

Note: In terms of dual indices, the number $p = 2$ is the natural dividing point of the extended interval $[1, \infty]$, since

$$1 \le p \le 2 \quad \iff \quad 2 \le p' \le \infty.$$

For this reason analysts often consider $p = 2$ to be the "midpoint" of the extended interval $[1, \infty]$. Alternatively, it may be more appropriate to think of the index as being $1/p$ rather than p. The value of $1/p$ ranges through the interval $[0, 1]$, and $1/p = 1/2$ corresponds to the actual midpoint of $[0, 1]$. We will discuss Hilbert spaces and $L^2(E)$ extensively in Chapter 8.

Exercise 7.1.5 and Remark 7.1.6. Usually I explain that equation (7.3) is a generalization of the arithmetic-geometric mean inequality, and say that Exercise 7.1.5 gives one proof of this inequality. Two other proofs are also mentioned in the text following Remark 7.1.6, one based on equation (7.4) (which is a special case of an inequality due to Young), and one based on Jensen's Inequality (given earlier in Problem 6.6.12).

Theorem 7.1.7 (Hölder's Inequality). State and prove for $1 < p < \infty$; the endpoint cases are exercises.

Note: Equation (7.5) holds in the extended real sense for all sequences x and y. If $x \notin \ell^p$ or $y \notin \ell^{p'}$, then the right-hand side of equation (7.5) is infinity, so the equation holds trivially. Of course, in that case we cannot derive any conclusion about whether xy belongs to ℓ^1 or not.

Note: "Hölder" is difficult for me to pronounce correctly, I have trouble getting the German umlaut right. I usually settle for the Americanization "Holder," but ask if there are any German-speaking students in the class who could give us the correct pronunciation.

7.1.2 Minkowski's Inequality

Exercise 7.1.8 and Theorem 7.1.9 (Minkowski's Inequality). State and prove the theorem (which covers $1 < p < \infty$); the exercise is the endpoint cases $p = 1$ and $p = \infty$.

Note: In various proofs, there are sometimes trivial cases that we do not explicitly discuss, because they do not have a significant impact on the proof. For example, at the end of the proof of Theorem 7.1.9, we can only divide by $\|x+y\|_p^{p-1}$ if $x+y$ is not the zero vector. However, if $x+y=0$ then the result is trivial, so we do not bother to mention this case in the text, even though it technically does need to be considered.

Theorem 7.1.10. State.

Corollary 7.1.11. I usually make a brief mention that things carry over to other countably infinite index sets, and that if I is finite then $\ell^p(I)$ is simply $\mathbb{C}^d$, so we have a whole suite of norms for $\mathbb{C}^d$ (and likewise for $\mathbb{R}^d$ if we only use real scalars).

7.1.3 Convergence in the ℓ^p Spaces

Definition 7.1.12 (Convergence in ℓ^p). State. This is convergence in norm in the space ℓ^p.

Definition 7.1.13 (Componentwise Convergence). State. Convergence in ℓ^p-norm and componentwise convergence are each perfectly good convergence criteria on ℓ^p, but one is the convergence notion induced from the norm of ℓ^p, and the other is not.

Lemma 7.1.14. State (the proof should be obvious). Although convergence in ℓ^p-norm implies componentwise convergence, the example after the lemma shows that the converse fails in general. Hence these are not "equivalent" convergence notions.

Note: In contrast, in the finite-dimensional Euclidean spaces $\mathbb{R}^d$ or $\mathbb{C}^d$, convergence with respect to the norm $\|\cdot\|_p$ is equivalent to componentwise convergence.

7.1.4 Completeness of the ℓ^p Spaces

Theorem 7.1.15 (Completeness of ℓ^p). Define Cauchy sequences, state and prove the theorem.

Note: The text gives a proof of equation (7.16) based on Fatou's Lemma for series. Here is a direct (but longer) version of the same argument that avoids appealing to Fatou.

If we fix $\varepsilon > 0$, then there is an $N > 0$ such that $\|x_m - x_n\|_p < \varepsilon$ for all $m, n \geq N$. Choose any particular $n \geq N$, and fix an integer $M > 0$. Then, since M is finite,

$$\begin{aligned}\sum_{k=1}^{M} |x(k) - x_n(k)|^p &= \sum_{k=1}^{M} \lim_{m\to\infty} |x_m(k) - x_n(k)|^p \\ &= \lim_{m\to\infty} \sum_{k=1}^{M} |x_m(k) - x_n(k)|^p \\ &\leq \lim_{m\to\infty} \|x_m - x_n\|_p^p \leq \varepsilon^p.\end{aligned}$$

Since this is true for every M, we conclude that

$$\begin{aligned}\|x - x_n\|_p^p &= \sum_{k=1}^{\infty} |x(k) - x_n(k)|^p \\ &= \lim_{M\to\infty} \sum_{k=1}^{M} |x(k) - x_n(k)|^p \leq \varepsilon^p.\end{aligned}$$

Corollary 7.1.16. Mention.

Note: In fact, all norms on a finite-dimensional vector space X are equivalent, and hence X is complete with respect to each of these norms (for some proofs, see [Con90, Thm. III.3.1] or [Heil18, Thm. 3.7.2]). In contrast, an infinite-dimensional normed space can be incomplete. One example is $C_c(\mathbb{R})$ with respect to the uniform norm, and another is c_{00} with respect to the sup-norm.

7.1.5 ℓ^p for $p < 1$

Exercise 7.1.17 and Theorem 7.1.18. I usually write down the metric for $p < 1$, but then just state that an exercise in the text shows that ℓ^p is a complete metric space when $p < 1$.

Note: If $p < 1$ then $\|\cdot\|_p^p$ satisfies the Triangle Inequality, but it is not a norm because it does not satisfy the homogeneity requirement. Indeed,

$\|cx\|_p^p = |c|^p\,\|x\|_p^p$. Furthermore, since open balls in ℓ^p are not convex when $p < 1$, the metric d_p cannot be induced from any norm. That is, if $p < 1$ then there does not exist any norm $\|\cdot\|$ on ℓ^p such that $\mathrm{d}_p(x,y) = \|x-y\|$. Convexity plays an important role in functional analysis (which I usually cover in the second semester of this course).

7.1.6 c_0 and c_{00}

This section introduces c_0 and c_{00}. These are important spaces, and several of the problems deal with properties of c_0 and c_{00}. Usually I define these spaces and remark that the standard basis is a vector space basis (or *Hamel basis*) for c_{00}, but it is not a basis for c_0 or ℓ^p in the usual vector space sense because Hamel bases are defined using only *finite* linear combinations.

Note: c_{00} is a *dense but proper* subspace of ℓ^p for $1 \le p < \infty$. In a finite-dimensional normed space, every proper subspace is closed. However, c_{00} is not a closed subspace of ℓ^p. Thus, *an infinite-dimensional Banach space can contain subspaces that are not closed*. Even after many years I still find this fact to be rather amazing.

Note: When dealing with infinite-dimensional Banach spaces it is often important to consider generalizations of bases that allow the use of "infinite linear combinations" instead of the *finite linear combinations* that are employed in the usual vector space definition of a basis. This idea is made precise in the notion of a *Schauder basis*, which we will touch on in Chapter 8 when we discuss orthonormal bases for Hilbert spaces. Also, Problem 7.4.11 shows that the standard basis $\mathcal{E} = \{\delta_n\}_{n\in\mathbb{N}}$ is a Schauder basis for c_0 (with respect to the sup-norm), and if $1 \le p < \infty$ then $\mathcal{E}$ is a Schauder basis for ℓ^p (with respect to the ℓ^p-norm).

Extra Problems for Section 7.1

1. Let $w\colon \mathbb{N} \to (0,\infty)$ be a fixed function. For each sequence of scalars $x = (x_k)_{k\in\mathbb{N}}$, set

$$\|x\|_{p,w} = \begin{cases} \displaystyle\biggl(\sum_{k=1}^{\infty} |x_k|^p\, w(k)^p\biggr)^{1/p}, & 0 < p < \infty, \\ \displaystyle\sup_{k\in\mathbb{N}} |x_k|\, w(k), & p = \infty, \end{cases}$$

and define $\ell^p_w = \bigl\{x : \|x\|_{p,w} < \infty\bigr\}$. Prove that this *weighted ℓ^p-space* is a Banach space for each index $1 \le p \le \infty$.

2. Fix $1 < p < \infty$. Given $x, y \in \ell^p$, prove that $\|x+y\|_p = \|x\|_p + \|y\|_p$ if and only if $x = 0$ or y is a positive scalar multiple of x. What happens if $p = 1$ or $p = \infty$?

3. We take $\overline{\mathbf{F}} = \mathbb{R}$ for this problem. Let R be the "closed first quadrant" in ℓ^1:

$$R = \big\{x = (x_k)_{k\in\mathbb{N}} \in \ell^1 : x_k \ge 0 \text{ for every } k\big\}.$$

(a) Determine, with proof, whether R is a closed subset of ℓ^1.

(b) Challenge (it's not what you expect): What is the boundary of R?

4. For each $n \in \mathbb{N}$, define $y_n = (1, \dots, 1, 0, 0, \dots)$, where the 1 is repeated n times. As usual, the norm on c_0 is the sup-norm $\|\cdot\|_\infty$. Part (a) of this problem is a subset of Problem 7.4.11 in the text, but part (b) is extra.

(a) Show that $\{y_n\}_{n\in\mathbb{N}}$ is a Schauder basis for c_0, That is, show that for each $x \in c_0$ there exist unique scalars $c_n(x)$ such that

$$x = \sum_{n=1}^{\infty} c_n(x)\, y_n. \qquad \text{(A)}$$

Note that, by definition, an infinite series converges if the partial sums converge in norm. Hence to prove that the series in equation (A) converges and equals x, you must prove that

$$\lim_{N\to\infty} \bigg\|x - \sum_{n=1}^{N} c_n(x)\, y_n\bigg\|_\infty = 0.$$

(b) Show that $\{y_n\}_{n\in\mathbb{N}}$ is a *conditional* Schauder basis for c_0; that is, there exists some $x \in c_0$ such that the series in equation (A) does not converge unconditionally. To do this, demonstrate that there is some bijection $\sigma\colon \mathbb{N} \to \mathbb{N}$ such that the reordered series $\sum\limits_{n=1}^{\infty} c_{\sigma(n)}(x)\, y_{\sigma(n)}$ does not converge.

Note: In contrast, the standard basis is an unconditional basis for ℓ^p (for finite p) and for c_0 (with respect to $\|\cdot\|_\infty$).

5. This problem will show that any incomplete normed space can be naturally embedded into a larger normed space that is complete.

Let X be a normed space that is not complete. Let $\mathcal{C}$ be the set of all Cauchy sequences in X, and define a relation $\sim$ on $\mathcal{C}$ by declaring that if $\mathcal{F} = \{f_n\}_{n\in\mathbb{N}}$ and $\mathcal{G} = \{g_n\}_{n\in\mathbb{N}}$ are two Cauchy sequences, then $\mathcal{F} \sim \mathcal{G}$ if and only if $\lim\limits_{n\to\infty} \|f_n - g_n\| = 0$.

(a) Prove that $\sim$ is an equivalence relation on $\mathcal{C}$.

(b) Let $[\mathcal{F}] = \big\{\mathcal{G} : \mathcal{G} \sim \mathcal{F}\big\}$ denote the equivalence class of $\mathcal{F}$ with respect to the relation $\sim$. Let $\widetilde{X}$ be the set of all equivalence classes $[\mathcal{F}]$. Define the norm of an equivalence class to be

$$\big\|\,[\mathcal{F}]\,\big\|_{\widetilde{X}} = \lim_{n\to\infty} \|f_n\|.$$

Prove that $\|\cdot\|_{\widetilde{X}}$ is well-defined and is a norm on $\widetilde{X}$.

(c) Given $f \in X$, let $[f]$ denote the equivalence class of the Cauchy sequence $\{f, f, f, \dots\}$. Prove that $T\colon f \mapsto [f]$ is an isometric map of X into $\widetilde{X}$. Show also that $T(X)$ is a dense subspace of $\widetilde{X}$ (so, in the sense of identifying of X with $T(X)$, we can consider X to be a subspace of $\widetilde{X}$).

(d) Prove that $\widetilde{X}$ is a Banach space with respect to $\|\cdot\|_{\widetilde{X}}$. We call $\widetilde{X}$ the *completion* of X.

(e) Prove that $\widetilde{X}$ is unique in the sense that if Y is a Banach space and $U\colon X \to Y$ is a linear isometry such that $U(X)$ is dense in Y, then there exists a linear isometric bijection $V\colon Y \to \widetilde{X}$.

Remark: A mapping A is an *isometry* is $\|Af\| = \|f\|$ for every f in the domain of A.

Section 7.2: The Lebesgue Space $L^p(E)$

Definition 7.2.1 (The Lebesgue Space $L^p(E)$). State.

Remark 7.2.2. Mention.

7.2.1 Seminorm Properties of $\|\cdot\|_p$

Exercises 7.2.3 and 7.2.4 (Hölder's and Minkowski's Inequalities). These work just like they do for ℓ^p.

Theorem 7.2.5. State, emphasizing the *almost everywhere uniqueness* property of the seminorm $\|\cdot\|_p$.

7.2.2 Identifying Functions that are Equal Almost Everywhere

Notation 7.2.6 (Informal Convention for Elements of $L^p(E)$). Discuss, but keep it brief. It's very easy to get bogged down in an extended discussion of the "correct" definition of the L^p spaces as sets of equivalence classes. This is one case where the informal convention is both more clear and usually "less dangerous" in some sense than the precise formal definition.

Exercise 7.2.7. Assign for reading.

Notation 7.2.8 (Continuity for Elements of $L^p(E)$). Discuss briefly. This convention means that when we write

"the continuous function $f(x) = e^{-|x|}$ belongs to $L^p(\mathbb{R})$,"

we recognize that any function g that equals $e^{-|x|}$ a.e. is the same element of $L^p(\mathbb{R})$, even though g need not be continuous. In the same way, we write

"$C_c(\mathbb{R}) \subseteq L^p(\mathbb{R})$," implicitly recognizing that elements of $C_c(\mathbb{R})$ are functions, while elements of $L^p(\mathbb{R})$ are really equivalence classes of functions that are equal a.e.

7.2.3 $L^p(E)$ *for* $0 < p < 1$

Exercise 7.2.9. Assign for reading.

Remark: Part (b) asks for a proof that $L^p(E)$ is complete when $0 < p < 1$. This part really should have a pointer to the later Exercise 7.3.5, which sketches a proof that $L^p(E)$ is complete when $1 \le p < \infty$. The approach for $0 < p < 1$ is entirely similar.

7.2.4 The Converse of Hölder's Inequality

Theorem 7.2.10 (Converse of Hölder's Inequality). State and prove. Alternatively, the proof of the special case $p = 2$ is somewhat less cluttered (see below); you could just prove it and refer to the text for the general case.

Proof of Theorem 7.2.10 for $p = 2$. If $p = 2$ then $p' = 2$ as well. Hölder's Inequality gives us equation (7.20), so we just need to prove that equality holds. Fix $f \in L^p(E)$. If $f = 0$ a.e., then the result is trivial, so we can assume that f is not the zero vector in $L^p(E)$. By choosing an appropriate representative of f, we can further assume that f is finite at every point. That is, if needed we simply redefine $f(x)$ at any point in the set of measure zero where it takes the value $\pm\infty$ to be some finite value, such as 0.

For each x, let $\alpha(x)$ be a scalar such that $|\alpha(x)| = 1$ and $\alpha(x)\, f(x) = |f(x)|$. Explicitly, we can take

$$\alpha(x) \;=\; \begin{cases} |f(x)|/f(x), & \text{if } f(x) \neq 0, \\ 0, & \text{if } f(x) = 0. \end{cases}$$

This function α is measurable and bounded. Set

$$g(x) \;=\; \frac{\alpha(x)\,|f(x)|}{\|f\|_2}, \qquad \text{for } x \in E.$$

Then

$$\|g\|_2^2 \;=\; \int_E \biggl(\frac{|f(x)|}{\|f\|_2}\biggr)^2 dx \;=\; \frac{\int_E |f(x)|^2\, dx}{\|f\|_2^2} \;=\; 1.$$

Thus g is a unit vector in $L^2(E)$. Also,

$$\int_E fg\,dx \;=\; \int_E f(x)\,\frac{\alpha(x)\,|f(x)|}{\|f\|_2}\,dx$$

$$=\; \frac{\int_E |f(x)|^2\,dx}{\|f\|_2} \;=\; \frac{\|f\|_2^2}{\|f\|_2} \;=\; \|f\|_2.$$

This shows that equality holds in equation (7.21), and that the supremum in that equation is achieved. □

An Extra Theorem. Although I usually do not present this in class, here is a nice theorem (from Folland [Fol99]) that would fit well into the discussion of the Converse to Hölder's Inequality. This theorem is useful for characterizing the dual space of $L^p(E)$ (which is something that I prove in the second semester of the course).

Theorem. Assume that $E \subseteq \mathbb{R}^d$ is measurable, and fix $1 \le p \le \infty$. Let S be the set of simple functions on E that vanish outside a set of finite measure:

$$S \;=\; \Big\{\phi\colon E \to \mathbb{C} \,:\, \phi \text{ is simple and } |\{\phi \ne 0\}| < \infty\Big\}.$$

Suppose that:

(a) $g\colon E \to \overline{\mathbf{F}}$ is Lebesgue measurable,

(b) $\phi\, g \in L^1(E)$ for each $\phi \in S$, and

(c) $M_g \;=\; \sup\biggl\{\biggl|\int_E \phi\, g\biggr| : \phi \in S,\ \|\phi\|_p = 1\biggr\} \;<\; \infty.$

Then $g \in L^{p'}(E)$ and $\|g\|_{p'} = M_g$.

Proof. By Hölder's Inequality, if $\phi \in S$ then

$$\biggl|\int_E \phi\, g\biggr| \;\le\; \|\phi\|_p\, \|g\|_{p'},$$

so we automatically have $M_g \le \|g\|_{p'}$.

Let α be a measurable function with unit modulus such that

$$|g(x)| \;=\; \alpha(x)\, g(x), \qquad \text{for all } x \in E.$$

Step 1. Fix any function $\phi \in S$. Then $\phi \in L^p(E)$ and, by the definition of M_g,

$$\biggl|\int_E \phi\, g\biggr| \;\le\; M_g\, \|\phi\|_p.$$

We will extend this formula from functions $\phi \in S$ to all bounded measurable functions f on E that vanish outside a set of finite measure.

Suppose that f is bounded and measurable, and that f is nonzero only on a set A of finite measure. Then $\chi_A \in S$, so $\chi_A\, g \in L^1(E)$. Since f is bounded and zero outside of A, we therefore have $fg \in L^1(E)$.

By our standard approximation theorems, we know that we can find simple functions ϕ_k such that $\phi_k \to f$ pointwise a.e. and $|\phi_k| \le |f|$ for every k. Each ϕ_k belongs to S, and we have that $\phi_k\, g \to fg$ pointwise a.e. and $|\phi_k\, g| \le |fg| \in L^1(E)$ for every k. Therefore, by the Dominated Convergence Theorem,

$$\left|\int_E fg\right| \;=\; \lim_{k\to\infty} \left|\int_E \phi_k\, g\right| \;\le\; \lim_{k\to\infty} M_g\, \|\phi_k\|_p \;\le\; M_g\, \|f\|_p. \qquad \text{(A)}$$

Step 2. Suppose that $1 < p < \infty$, in which case $1 < p' < \infty$. If g is the zero function (or is zero at almost every point) then there is nothing to prove, so we can assume that g is not zero at almost every point.

Let ϕ_k be simple functions on E such that $0 \le \phi_k \nearrow |g|^{p'}$. If necessary, by replacing ϕ_k with $\phi_k \cdot \chi_{E\cap[-k,k]}$ we may assume that each ϕ_k vanishes outside a set of finite measure. Since each ϕ_k is nonnegative, we can define

$$g_k \;=\; \alpha\, \phi_k^{1/p}.$$

Motivation:

$$|g_k| \;=\; \phi_k^{1/p} \;=\; \phi_k^{1-\frac{1}{p'}} \;\approx\; \bigl(|g|^{p'}\bigr)^{1-\frac{1}{p'}} \;=\; |g|^{p'-1},$$

and therefore

$$|g_k\, g| \;\approx\; |g|^{p'}.$$

More precisely, since $|g_k| = \phi_k^{1/p}$, we have

$$\|g_k\|_p \;=\; \left(\int_E |g_k|^p\right)^{1/p} \;=\; \left|\int_E \phi_k\right|^{1/p} \;=\; \|\phi_k\|_1^{1/p}$$

and

$$\phi_k \;=\; \phi_k^{1/p}\, \phi_k^{1/p'} \;\le\; \phi_k^{1/p}\, |g| \;=\; \phi_k^{1/p}\, \alpha\, g \;=\; g_k\, g.$$

Further, each g_k is bounded and vanishes outside a set of finite measure, so by equation (A) we have

$$\begin{aligned} \|\phi_k\|_1^{1/p}\, \|\phi_k\|_1^{1/p'} \;&=\; \|\phi_k\|_1 \\ &=\; \int_E \phi_k \end{aligned}$$

$$\leq \int_E g_k\, g$$

$$\leq M_g \, \|g_k\|_p \;=\; M_g \, \|\phi_k\|_1^{1/p}.$$

Since g is not zero a.e., it must be the case that ϕ_k is not zero a.e. for all large enough k. Thus $\|\phi_k\|_1 \neq 0$ for large k, so the preceding inequality reduces to

$$\|\phi_k\|_1^{1/p'} \;\leq\; M_g, \qquad \text{(B)}$$

for all large enough k. Using Fatou's Lemma, we therefore compute that

$$\begin{aligned} \|g\|_{p'}^{p'} \;=\; \int_E |g|^{p'} \;&=\; \int_E \liminf_{k\to\infty} |\phi_k| \\ &\leq\; \liminf_{k\to\infty} \int_E |\phi_k| && \text{(Fatou)} \\ &=\; \liminf_{k\to\infty} \|\phi_k\|_1 \;\leq\; M_g^{p'}. && \text{(by equation (B))} \end{aligned}$$

This implies that $g \in L^{p'}(E)$ and $\|g\|_{p'} \leq M_g$.

Step 3. Suppose that $p = 1$, so $p' = \infty$. Fix $\varepsilon > 0$, and let

$$A \;=\; \bigl\{|g| \geq M_g + \varepsilon\bigr\}.$$

We will show that A must have measure zero.

Suppose that $|A| > 0$, and choose a measurable set $B \subseteq A$ with $0 < |B| < \infty$. Let

$$f \;=\; \alpha \chi_B.$$

This function f is bounded and vanishes outside a set of finite measure, so by equation (A) we have that

$$\left| \int_E f g \right| \;\leq\; M_g \, \|f\|_1 \;=\; M_g \, |B|.$$

However,

$$\int_E f g \;=\; \int_B \alpha\, g \;=\; \int_B |g| \;\geq\; (M_g + \varepsilon)\, |B|.$$

Therefore

$$(M_g + \varepsilon)\, |B| \;\leq\; M_g \, |B|,$$

which is a contradiction. Consequently, we must have $|A| = 0$, which implies $|g| \leq M_g + \varepsilon$ almost everywhere. Therefore $\|g\|_\infty \leq M_g + \varepsilon$. Since this is true for every $\varepsilon > 0$, we conclude that $\|g\|_\infty \leq M_g < \infty$.

Step 4. Suppose that $p = \infty$, so $p' = 1$. Set

$$g_k \;=\; \alpha\,\chi_{E\cap[-k,k]}.$$

Then g_k is bounded and vanishes outside a set of finite measure, so

$$\int_{E\cap[-k,k]} |g| \;=\; \int_E g_k\, g \;\le\; M_g\, \|g_k\|_\infty \;=\; M_g.$$

Since this is true for every k, we see that

$$\|g\|_1 \;=\; \int_E |g| \;\le\; M_g \;<\; \infty.$$

Hence $g \in L^1(E)$. $\square$

Extra Problems for Section 7.2

1. Suppose that $1 \le p, q \le \infty$ with $p \ne q$. Find a function f such that $10^6 < \|f\|_p < \infty$ and $\|f\|_q < 10^{-6}$.

2. Prove the following statements.

(a) $C_c(\mathbb{R}) \subsetneq L^p(\mathbb{R})$ for every $0 < p \le \infty$.

(b) $C_0(\mathbb{R})$ is not contained in $L^p(\mathbb{R})$ for any $0 < p < \infty$. In particular, $f(x) = \bigl(1 + \bigl|\ln|x|\bigr|\bigr)^{-1}$ belongs to $C_0(\mathbb{R})$ but $f \notin L^p(\mathbb{R})$ for any $0 < p < \infty$.

(c) $L^p(\mathbb{R})$ is not contained in $C_0(\mathbb{R})$ for any $0 < p < \infty$. In fact, if p is finite then there exists a continuous but unbounded function that belongs to $L^p(\mathbb{R})$.

3. Fix $1 \le p \le \infty$. Prove that the norm $\|\cdot\|_p$ is *translation-invariant* on $L^p(\mathbb{R}^d)$. That is, $\|T_a f\|_p = \|f\|_p$ for all $f \in L^p(\mathbb{R}^d)$ and $a \in \mathbb{R}^d$, where $T_a f(x) = f(x-a)$.

4. Fix $1 \le p \le \infty$, and let $E \subseteq \mathbb{R}^d$ be measurable. Show that if functions f, $g\colon E \to \overline{\mathbf{F}}$ are measurable on E and finite a.e., then $\|f+g\|_p \le \|f\|_p + \|g\|_p$ (note that these quantities might be infinite).

5. Show that if $f \in L^3[-1,1]$, then $\displaystyle\int_{-1}^{1} \frac{f(x)}{\sqrt{|x|}}\,dx$ exists and is a finite scalar.

6. Let $E \subseteq \mathbb{R}^d$ be measurable. Prove that if $f \in L^p(E)$ and $p > 4/3$, then

$$\lim_{t\to 0^+} \int_0^t x^{-1/4} f(x)\,dx \;=\; 0.$$

7. Fix $1 \le p < \infty$ and let $E \subseteq \mathbb{R}^d$ be measurable. Given $f \in L^p(E)$, prove that $\lim_{t\to\infty} t^p \bigl|\{|f| > t\}\bigr| = 0$.

8. (From Wheeden and Zygmund [WZ77]) Assume that $E \subseteq \mathbb{R}^d$ is measurable and $0 < |E| < \infty$. For $1 \le p < \infty$ set

$$N_p[f] \,=\, \biggl(\frac{1}{|E|}\int_E |f|^p\biggr)^{1/p}, \qquad \text{for } f \in L^p(E).$$

Prove the following statements.

(a) If $1 \le p_1 < p_2 < \infty$ then $N_{p_1}[f] \le N_{p_2}[f]$.

(b) $N_p[f+g] \le N_p[f] + N_p[g]$ for $f, g \in L^p(E)$.

(c) $N_1[fg] \le N_p[f]\, N_p[g]$ for $f \in L^p(E)$, $g \in L^{p'}(E)$.

(d) If $f \in L^p(E)$ for some $p < \infty$, then $\lim_{p\to\infty} N_p[f] = \|f\|_\infty$.

Quoting [WZ77]: "Thus, N_p behaves like $\|\cdot\|_p$ but has the advantage of being monotonic in p."

9. Let $E \subseteq \mathbb{R}^d$ be measurable and fix $1 < p < \infty$. Choose functions $f, g \in L^p(E)$ with $g \ne 0$ (that is, g is not the zero vector in $L^p(E)$). Prove that if equality holds in Minkowski's Inequality (so $\|f+g\|_p = \|f\|_p + \|g\|_p$), then f is a scalar multiple of g.

10. (This problem is a special case of Problem 7.2.16, but the point is to work it directly; in fact it may give some insight into the more general result of Problem 7.2.16.) Given $0 < p < q \le \infty$, prove the following statements.

(a) $L^q[0,1] \subsetneq L^p[0,1]$, and $\|f\|_p \le \|f\|_q$ for all $f \in L^p[0,1]$.

(b) $L^p(\mathbb{R})$ is not contained in $L^q(\mathbb{R})$, and $L^q(\mathbb{R})$ is not contained in $L^p(\mathbb{R})$.

11. Formulate and prove an analogue of the Converse to Hölder's Inequality (Theorem 7.2.10) for the ℓ^p spaces.

12. Prove that equation (7.21) in the converse to Hölder's Inequality holds in the extended real sense for all $f \in L^p_{\text{loc}}(E)$, even if $f \notin L^p(E)$. That is, assume $E \subseteq \mathbb{R}^d$ is measurable, fix $1 \le p \le \infty$, and let $f\colon E \to \overline{\mathbf{F}}$ be any measurable function such that $\|f \cdot \chi_{E\cap K}\|_p < \infty$ for every compact $K \subseteq \mathbb{R}^d$. Prove that

$$\sup_{\|g\|_{p'}=1} \biggl|\int_E fg\biggr| \,=\, \|f\|_p.$$

13. This problem, which is a special case of *Hardy's Inequalities*, appears later in the text as Problem 8.1.14, but we include it here because it is related to the next few extra problems. Prove that if $f \in L^2[0,\infty)$, then

$$\left|\int_0^x f(t)\,dt\right|^2 \;\le\; 2x^{1/2}\int_0^x t^{1/2}\,|f(t)|^2\,dt, \qquad \text{for } x \ge 0.$$

Then define

$$F(x) \;=\; \frac{1}{x}\int_0^x f(t)\,dt, \qquad \text{for } x \ge 0,$$

and show that $F \in L^2[0,\infty)$ and $\|F\|_2 \le 2\,\|f\|_2$.

14. This problem will establish a special case of *Hardy's Inequalities*. Given $1 < p < \infty$ and $f \in L^2[0,\infty)$, let

$$F(x) \;=\; \frac{1}{x}\int_0^x f(t)\,dt, \qquad x \ge 0,$$

and show that $F \in L^p[0,\infty)$ and $\|F\|_p \le \dfrac{p}{p-1}\,\|f\|_p$.

15. This problem will establish a version of *Hardy's Inequalities*.

(a) Fix $1 \le p < \infty$. Show that if $\alpha < -1$, then there exists a constant $C(\alpha,p)$ such that for any measurable function $f\colon (0,\infty) \to [0,\infty]$ we have

$$\int_0^\infty \left(\int_0^x f(t)\,dt\right)^p x^\alpha\,dx \;\le\; C(\alpha,p)\int_0^\infty f(t)^p\,t^{\alpha+p}\,dt, \qquad \text{(A)}$$

while if $\alpha > -1$ then the inequality is

$$\int_0^\infty \left(\int_x^\infty f(t)\,dt\right)^p x^\alpha\,dx \;\le\; C(\alpha,p)\int_0^\infty f(t)^p\,t^{\alpha+p}\,dt.$$

(b) Show that if $\alpha = -p < -1$, then the optimal constant in equation (A) is

$$C(-p,p) \;=\; (p')^p \;=\; \left(\frac{p}{p-1}\right)^p.$$

(c) Suppose that $f \in L^p(\mathbb{R})$, where $1 < p < \infty$, and define

$$F(x) \;=\; \frac{1}{x}\int_0^x |f(t)|\,dt, \qquad \text{for } x \in \mathbb{R}.$$

Show that

$$\|F\|_p \;\le\; p'\,\|f\|_p, \qquad \text{(B)}$$

and prove that p' is the best possible constant. Also show that equality holds in equation (B) if and only if $f = 0$ a.e.

16. Let $E \subseteq \mathbb{R}^d$ be measurable. Assume that $0 < p < 1$ and $-\infty < q < 0$ satisfy $\frac{1}{p} + \frac{1}{q} = 1$. Let $f, g\colon E \to \overline{\mathbf{F}}$ be measurable functions such that $|f|^p$, $|g|^q$, and $|fg|$ are all integrable. Prove that

$$\left(\int_E |f|^p\right)^{1/p} \left(\int_E |g|^q\right)^{1/q} \le \int_E |fg|.$$

Hint: Show that $r = 1/p$ and $s = -q/p$ are dual indices, and write $|f|^p = |fg|^p\,|g|^{-p}$.

17. Let S be the set of all nonnegative measurable functions on $[0,\infty)$ that satisfy $\int_0^\infty f(x)^4\,dx \le 1$. Find $\sup_{f\in S} \int_0^\infty f(x)^3\,e^{-x}\,dx$.

18. For $x > 0$ define

$$f(x) = x^{-1/2}\left(1+|\ln x|\right)^{-1} = \frac{1}{x^{1/2}\left(1+|\ln x|\right)}.$$

Prove that $f \in L^p(0,\infty)$ if and only if $p = 2$.

19. Suppose that $f\colon [0,1] \to [0,\infty)$ is measurable and $\int_A f \le |A|^{1/2}$ for every measurable set A in $[0,1]$. Prove that $f \in L^p[0,1]$ for $1 \le p < 2$.

Hint: This is a simpler version of Problem 7.2.22, but try to do it directly by considering $|\{2^k \le f < 2^{k+1}\}|$.

20. Prove that if $f \in L^1(\mathbb{R}) \cap L^4(\mathbb{R})$, then for every $1 \le p \le 4$ we have both

(a) $f \in L^p(\mathbb{R})$, and

(b) $\lim_{k\to\infty} k^p \left|\{x \in \mathbb{R} : |f(x)| > k\}\right| = 0.$

21. Let E be a measurable subset of $\mathbb{R}^d$ with finite measure, and let $\mathcal{L}_E$ be the σ-algebra of all measurable subsets of E. Define the distance between sets $A, B \in \mathcal{L}_E$ to be $\mathrm{d}(A,B) = |A \triangle B|$.

(a) Prove that if we identify sets that differ only by a set of measure zero, then d is a metric on $\mathcal{L}_E$, and $\mathcal{L}_E$ is complete with respect to this metric.

(b) Prove that $\mathcal{L}_E$ is not compact.

Section 7.3: Convergence in L^p-norm

Definition 7.3.1 (Convergence in $L^p(E)$). State.

Remark 7.3.2. State.

Note: Looking at equation (7.24) we see that, for finite p,

$$f \in L^p(E) \quad\Longleftrightarrow\quad |f|^p \in L^1(E),$$

and

$$f_n \to f \text{ in } L^p(E) \quad\Longleftrightarrow\quad |f - f_n|^p \to 0 \text{ in } L^1(E).$$

Using this equivalence we can adapt previous L^1-norm results to the setting of $L^p(E)$. For example, if we want to show that $f_n \to f$ in $L^p(E)$, we could use the Dominated Convergence Theorem to prove that $|f - f_n|^p \to 0$ in $L^1(E)$. For an example of this type of argument, see the proof of Theorem 7.3.9.

Example 7.3.3 and Theorem 7.3.4. Mention briefly.

Note: Here are the details of the use of Tchebyshev's Inequality in the proof of Theorem 7.3.4.

Assume $p < \infty$ and $f_k \to f$ in $L^p(E)$. Choose any $\varepsilon > 0$. Then by Tchebyshev's Inequality,

$$\begin{aligned} \big|\{|f - f_k| > \varepsilon\}\big| &= \big|\{|f - f_k|^p > \varepsilon^p\}\big| \\ &\le \frac{1}{\varepsilon^p} \int_E |f - f_k|^p \\ &= \frac{1}{\varepsilon^p} \, \|f - f_k\|_p^p \to 0 \text{ as } k \to \infty. \end{aligned}$$

Hence $f_k \xrightarrow{\text{m}} f$.

The same result also holds for $p = \infty$. In that case $\big|\{|f - f_k| > \varepsilon\}\big| = 0$ for all k large enough, so we again conclude that $f_k \xrightarrow{\text{m}} f$.

Exercise 7.3.5 and Theorem 7.3.6 ($L^p(E)$ is a Banach Space). State briefly, all of this is similar to the situation for ℓ^p. The proof of the theorem is Exercise 7.3.5.

7.3.1 Dense Subsets of $L^p(E)$

Lemma 7.3.7. Review the definition of density for subsets of metric (or normed) spaces and state the lemma, but encourage students to prove it on their own. Density is briefly defined in Section 1.1.2, but the **Alternative Chapter 1** contains a more extensive discussion and review of density.

Definition 7.3.8 (Compact Support). State.

Note: If E is bounded, then E is contained in some box Q, which is a compact set, and $E \backslash Q$ is empty, so every function on a bounded set E is compactly supported.

Theorem 7.3.9 (Compactly Supported Functions are Dense). State. The proof is easy, but it shows how to use the DCT for L^p-norm convergence.

Exercise 7.3.10, 7.3.11, and 7.3.12. Just mention that several density results that we proved for L^1 have extensions to L^p, including the following.

- The simple functions are dense in $L^p(E)$ for $1 \le p \le \infty$.
- The continuous, compactly supported functions are dense in $L^p(\mathbb{R}^d)$ for finite p.
- The really simple functions are dense in $L^p(\mathbb{R})$ for finite p.

Extra Problems for Section 7.3

1. This problem is a special case of Extra Problem 2, but it may be instructive to try this version first.

(a) Let $E \subseteq \mathbb{R}^d$ be measurable with $|E| < \infty$. Assume that $\{f_n\}_{n\in\mathbb{N}}$ is a bounded sequence in $L^2(E)$ and there exists a function f such that $f_n(x) \to f(x)$ for a.e. $x \in E$. Prove that $\|f - f_n\|_1 \to 0$ as $n \to \infty$.

(b) Show that the conclusion of part (a) can fail if $|E| = \infty$.

2. Let $E \subseteq \mathbb{R}^d$ be measurable with $|E| < \infty$, and let f_n be measurable functions on E such that $f_n \to f$ pointwise a.e. Show that if $\|f_n\|_2 \le 1$ for every n, then $\|f - f_n\|_p \to 0$ for each $1 \le p < 2$.

Hint: Egorov's Theorem and Hölder with indices $p/2$ and $(p/2)'$.

3. Assume that E_n are disjoint measurable sets in $\mathbb{R}^d$, and $f_n \in L^p(E_n)$ for each n. Extend f_n to $\mathbb{R}^d$ by declaring it to be zero outside of E_n. Prove that the series $f = \sum_n f_n$ converges in L^p-norm if and only if $\sum_n \|f_n\|_p^p < \infty$, and in this case we have $\|f\|_p^p = \sum_n \|f_n\|_p^p$.

4. Fix $0 < p < 1$. Using the metric $\mathrm{d}_p(f,g) = \|f - g\|_p^p$ on $L^p(\mathbb{R})$, prove that $C_c(\mathbb{R})$ and the set of all really simple functions are dense in $L^p(\mathbb{R})$.

5. Fix $1 \le p < \infty$, and let E be a measurable subset of $\mathbb{R}^d$ such that $0 < |E| < \infty$. Suppose that $f_n \in L^p(E)$, $f_n \to f$ a.e., and $\sup_n \|f_n\|_p < \infty$. Prove that $f \in L^p(E)$ and $\|f - f_n\|_q \to 0$ for each index $1 \le q < p$. However, show by example that $\|f - f_n\|_p$ need not converge to 0. (This problem has some similarities to Problem 7.3.21.)

6. Let $E \subseteq \mathbb{R}^d$ be measurable with $0 < |E| < \infty$. Assume that:

(a) $f_n \in L^1(E)$ for every n,

(b) there exists a function f such that $f_n \to f$ pointwise a.e., and

(c) there exists some $1 < p \le \infty$ such that $\sup \|f_n\|_p < \infty$.

Prove that $f \in L^1(E)$ and $f_n \to f$ in L^1-norm. Show by example that this conclusion can fail if we allow $p = 1$ in assumption (c). (This problem has some similarities to Problem 7.3.21.)

7. Given $1 \le p < \infty$ and $f_n \in L^p(\mathbb{R}^d)$, prove that $\{f_n\}_{n\in\mathbb{N}}$ is a Cauchy sequence in $L^p(\mathbb{R})$ if and only if the following three conditions hold.

(a) $\{f_n\}_{n\in\mathbb{N}}$ is Cauchy in measure.

(b) For every $\varepsilon > 0$ there exists a $\delta > 0$ such that if E is measurable and $|E| < \delta$ then $\int_E |f_n|^p < \varepsilon$ for every n.

(c) For every $\varepsilon > 0$ there exists a measurable set E with $|E| < \infty$ such that $\int_{E^C} |f_n|^p < \varepsilon$ for every n.

8. (From Folland). Let E be a measurable subset of $\mathbb{R}^d$ and fix indices $1 \le p < q \le \infty$. Prove the following statements.

(a) $\|f\| = \inf\bigl\{\|g\|_p + \|h\|_q : f = g+h \text{ with } g \in L^p(E),\, h \in L^q(E)\bigr\}$ defines a norm on

$$L^p(E) + L^q(E) = \bigl\{f+g : f \in L^p(E),\, g \in L^q(E)\bigr\}.$$

(b) $L^p(E) + L^q(E)$ is a Banach space with respect to this norm.

(c) If $1 \le p < r < q \le \infty$ then $L^r(E) \subseteq L^p(E) + L^q(E)$.

9. Suppose that $\{f_n\}_{n\in\mathbb{N}}$ is a sequence of functions in $L^1(\mathbb{R})$ such that

$$\lim_{n\to\infty} \int_{-\infty}^{\infty} f_n(x)\, g(x)\, dx = g(0) \qquad \text{for every } g \in C_0(\mathbb{R}).$$

Prove that $\{f_n\}_{n\in\mathbb{N}}$ is not Cauchy in $L^1(\mathbb{R})$.

10. Let $\mathcal{S}$ be the set of all infinite sequences $x = (x_k)_{k\in\mathbb{N}}$. This problem will show that componentwise convergence is a metrizable convergence criterion on $\mathcal{S}$. Prove the following statements.

(a) If $k \in \mathbb{N}$ is fixed, then $|||x|||_k = |x_k|$ for $x \in \mathcal{S}$ defines a seminorm on $\mathcal{S}$. However, $|||\cdot|||_k$ is not a norm on $\mathcal{S}$.

(b) The following is a metric on $\mathcal{S}$:

$$\mathrm{d}(x,y) = \sum_{k=1}^{\infty} 2^{-k} \frac{|x_k - y_k|}{1 + |x_k - y_k|} \qquad \text{for } x,\, y \in \mathcal{S}.$$

Hint: Show that if $a,\, b,\, c \ge 0$ and $a \le b + c$, then $\frac{a}{1+a} \le \frac{b}{1+b} + \frac{c}{1+c}$.

(c) Given sequences $x_n,\, x \in \mathcal{S}$, we have that x_n converges componentwise to x if and only if $\mathrm{d}(x, x_n) \to 0$ as $n \to \infty$.

(d) $\mathcal{S}$ is complete with respect to the metric d.

11. This problem will show that $L^1_{\mathrm{loc}}(\mathbb{R})$ has a metrizable topology. Prove the following statements.

(a) If $N \in \mathbb{N}$ is fixed, then $|||f|||_N = \|f\,\chi_{[-N,N]}\|_1$ defines a seminorm on $L^1_{\mathrm{loc}}(\mathbb{R})$. However, $|||\cdot|||_N$ is not a norm on $L^1_{\mathrm{loc}}(\mathbb{R}^d)$, even if we identify functions that are equal a.e.

(b) If we identify functions that are equal a.e., then

$$\mathrm{d}(f,g) \,=\, \sum_{N=1}^{\infty} 2^{-N} \frac{|||f-g|||_N}{1+|||f-g|||_N} \qquad \text{for } f,\, g \in L^1_{\mathrm{loc}}(\mathbb{R}),$$

defines a metric on $L^1_{\mathrm{loc}}(\mathbb{R})$.

Hint: Show that if a, b, $c \ge 0$ and $a \le b+c$, then $\frac{a}{1+a} \le \frac{b}{1+b} + \frac{c}{1+c}$.

(c) $f_k \to f$ with respect to the metric d if and only if for every compact set $K \subseteq \mathbb{R}$ we have $\|(f-f_k)\chi_K\|_1 \to 0$ as $k \to \infty$.

(d) $L^1_{\mathrm{loc}}(\mathbb{R})$ is complete with respect to the metric d.

12. (From Stroock [Str11]) Let E be a measurable subset of $\mathbb{R}^d$ such that $|E| < \infty$. Given measurable functions $f_n \colon E \to \overline{\mathbf{F}}$, prove that

$$f_n \overset{\mathrm{m}}{\to} f \qquad \Longleftrightarrow \qquad \lim_{n\to\infty} \int_E \max\bigl\{|f-f_n|,\, 1\bigr\} \,=\, 0.$$

13. Let $\mathcal{F}(\mathbb{R})$ be the vector space of all functions $f \colon \mathbb{R} \to \mathbb{C}$. Show that pointwise convergence is not a *normable* criterion. That is, there does not exist a norm $\|\cdot\|$ on $\mathcal{F}(\mathbb{R})$ such that

$$f_k \to f \text{ pointwise} \qquad \Longleftrightarrow \qquad \lim_{k\to\infty} \|f-f_k\| \,=\, 0.$$

Hint: Shrinking boxes.

14. Let $\mathcal{M}[0,1]$ be the space of all measurable functions $f \colon [0,1] \to \overline{\mathbf{F}}$ that are finite a.e. (where we identify functions that are equal a.e.). Show that convergence in measure is not a *normable* criterion on $\mathcal{M}[0,1]$, i.e., there does not exist a norm $\|\cdot\|$ on $\mathcal{M}[0,1]$ such that

$$f_k \overset{\mathrm{m}}{\to} f \qquad \Longleftrightarrow \qquad \lim_{k\to\infty} \|f-f_k\| \,=\, 0.$$

Compare this to Problem 7.3.26 or Extra Problem 11 above, which show that convergence in measure is *metrizable*.

15. Let $\mathcal{M}$ be the space of all Lebesgue measurable subsets of $[0,1]$, where we identify any sets A and B such that their symmetric difference $A \triangle B$ has measure zero. Given A, $B \in \mathcal{M}$, set $\mathrm{d}(A,B) = \int_0^1 |\chi_A - \chi_B|$. Prove that d is a metric on $\mathcal{M}$, and $\mathcal{M}$ is complete with respect to this metric.

16. Given $1 \le p \le \infty$ and $1 \le q < \infty$, the *Wiener amalgam space* $W(L^p, \ell^q)$ consists of those functions $f \in L^p_{\mathrm{loc}}(\mathbb{R})$ for which

$$\|f\|_{W(L^p,\ell^q)} \,=\, \biggl(\sum_{k\in\mathbb{Z}} \|f \cdot \chi_{[k,k+1]}\|_{L^p}^q\biggr)^{1/q}$$

is finite. For $q = \infty$ we define

$$\|f\|_{W(L^p,\ell^\infty)} = \sup_{k\in\mathbb{Z}} \|f \cdot \chi_{[k,k+1]}\|_{L^p}.$$

Prove that $\|\cdot\|_{W(L^p,\ell^q)}$ is a norm, and $W(L^p, \ell^q)$ is a Banach space with respect to this norm.

For more details on amalgam spaces, with applications to time-frequency analysis, see [Heil11, Chap. 11] and [Grö01].

Section 7.4: Separability of $L^p(E)$

Definition 7.4.1 (Separable Space). State and discuss.

Theorem 7.4.2 (Separability of $L^p(\mathbb{R})$). State and prove.

Theorem 7.4.3 (Separability of $L^p(E)$). Mention.

Theorem 7.4.4. State and prove. Use this to prove that $L^\infty(\mathbb{R})$ is nonseparable.

Note: **TYPO** in the proof of this theorem. On line 2 of page 287, change "$\|t - x_t\|_\infty < \frac{1}{2}$ to "$\mathrm{d}(t, x_t) < \frac{1}{2}$".

Note: Nonseparable Banach spaces are "unpleasant" in many respects. Unfortunately we will not see why in this semester. I cover basic functional analysis in the second semester of this course, and we see there some of the less pleasant features of L^∞.

Problems: *Note on Problem* 7.4.11 in the text: This problem shows that only a separable Banach space can have a Schauder basis. However, not every separable Banach space has a Schauder basis! A separable (and reflexive) Banach space that does not have a Schauder basis was constructed in the following paper, thereby settling negatively a longstanding open problem known as the *Basis Problem*.

P. Enflo, A counterexample to the approximation problem in Banach spaces, *Acta Math.*, **130** (1973), pp. 309–317.

Extra Problems for Section 7.4

1. We say that a sequence $\{x_n\}_{n\in\mathbb{N}}$ in a Banach space X is *ω-dependent* if there exist scalars c_n, not all zero, such that $\sum_{n=1}^{\infty} c_n x_n = 0$, where the series converges in the norm of X. A sequence is *ω-independent* if it is not ω-dependent.

(a) Prove that every Schauder basis is both complete and ω-independent.

(b) Let α and β be fixed nonzero scalars such that $|\alpha| > |\beta|$. Let $\{\delta_n\}_{n\in\mathbb{N}}$ be the sequence of standard basis vectors, and define

$$x_0 \;=\; \delta_1 \qquad \text{and} \qquad x_n \;=\; \alpha\delta_n + \beta\delta_{n+1},\; n \in \mathbb{N}.$$

Prove that the sequence $\{x_n\}_{n\ge 0}$ is complete and finitely linearly independent in ℓ^2, but it is not ω-independent and therefore is not a Schauder basis for ℓ^2.

2. Fix $0 < p < 1$. Prove that $L^p(\mathbb{R})$ is separable with respect to the metric $\mathrm{d}_p(f,g) = \|f-g\|_p^p$.

CHAPTER 8: Hilbert Spaces and $L^2(E)$

Section 8.1: Inner Products and Hilbert Spaces

In this first section we consider abstract inner product spaces, not just the specific inner product space $L^2(E)$. This simplifies the notation, and also allows us to concentrate on issues related to inner products without having to worry about issues related to integrability. Depending on the preparation of your students you may want to give more or less detail than I indicate below. A review of norms and Banach spaces could be appropriate before beginning this chapter. There is a short review of norms in Chapter 1, and an extended discussion is in the **Alternative Chapter 1**.

8.1.1 The Definition of an Inner Product

Definition 8.1.1 (Semi-Inner Product, Inner Product). State.

Note: A function of two variables that is linear in the first variable and antilinear in the second variable is called a *sesquilinear form* (the prefix "sesqui-" means "one and a half"). In particular, a semi-inner product $\langle \cdot, \cdot \rangle$ is a sesquilinear form. There are many different standard notations for semi-inner products. While our preferred notation is $\langle f, g \rangle$, the notations $[f, g]$, (f, g), and $\langle f | g \rangle$ are also common.

Note: Sometimes an inner product is required to be antilinear in the first variable and linear in the second (this is common in the physics literature). This is also convenient in finite-dimensional linear algebra, since if we define $x \cdot y = \overline{x_1} y_1 + \cdots + \overline{x_d} y_d$, then we can write $x \cdot y$ as the matrix product $x \cdot y = x^{\mathrm{H}} y$, where $x^{\mathrm{H}} = [\overline{x_1} \cdots \overline{x_d}]$ is the Hermitian, or conjugate transpose, of a vector $x = \begin{bmatrix} x_1 \\ \vdots \\ x_d \end{bmatrix} \in \mathbb{C}^d$.

8.1.2 Properties of an Inner Product

Exercise 8.1.2. State.

Theorem 8.1.3 (Cauchy–Bunyakovski–Schwarz Inequality). State and prove.

Lemma 8.1.4. State and prove.

Exercise 8.1.5. State.

8.1.3 Hilbert Spaces

Definition 8.1.6 (Hilbert Space). State and discuss.

Example 8.1.7 (The ℓ^2-Inner Product). State. We refer to an element x of ℓ^2 as a *square-summable sequence*.

Example 8.1.8 (The L^2-Inner Product). State. We refer to an element f of $L^2(E)$ as a *square-integrable function*.

Extra Problems for Section 8.1

1. If $x \in H$, where H is an inner product space, then $\|x\| = \sup_{\|y\|=1} |\langle x, y\rangle|$.

2. Suppose that $\{x_n\}_{n\in\mathbb{N}}$ is a sequence in a Hilbert space H and there exists a constant $B \geq 0$ such that

$$\sum_{n=1}^{\infty} |\langle x, x_n\rangle|^2 \;\leq\; B\,\|x\|^2, \qquad \text{for all } x \in H.$$

(Such a sequence is called a *Bessel sequence* in H.) Prove that if $(c_n)_{n\in\mathbb{N}}$ is a sequence of scalars with at most finitely many nonzero components, then

$$\biggl\|\sum_{n=1}^{\infty} c_n x_n\biggr\|^2 \;\leq\; B \sum_{n=1}^{\infty} |c_n|^2.$$

Hint: Extra Problem 1 and equation (8.3).

3. Prove the following variation on the Polar Identity: If x and y are vectors in an inner product space H, then

$$\|x + iy\|^2 \;=\; \|x\|^2 \;-\; 2\,\mathrm{Im}\langle x, y\rangle \;+\; \|y\|^2.$$

4. Let H be an inner product space.

(a) Characterize the vectors $x,\, y \in H$ such that $\|x+y\|^2 = \|x\|^2 + \|y\|^2$.

(b) Characterize the vectors $x,\, y \in H$ such that $\|x+y\| = \|x\| + \|y\|$.

5. Let H be a Hilbert space. Recall that the Pythagorean Theorem states that if $\langle x, y\rangle = 0$, then $\|x+y\|^2 = \|x\|^2 + \|y\|^2$.

(a) Prove that if H is a real Hilbert space then the converse to the Pythagorean Theorem holds, i.e., if $\|x+y\|^2 = \|x\|^2 + \|y\|^2$ then $\langle x, y\rangle \neq 0$.

(b) Show by example that if H is a complex Hilbert space then there can exist vectors $x, y \in H$ such that $\|x+y\|^2 = \|x\|^2 + \|y\|^2$ but $\langle x, y\rangle \neq 0$.

6. Suppose that functions $f_n \in L^2[0,1]$ satisfy $f_n \to f$ a.e. on $[0,1]$, and for every n we have $|f_n(x)| \leq x^{-1/3}$ a.e. Prove that $\langle f_n, g\rangle \to \langle f, g\rangle$ for every $g \in L^2[0,1]$.

7. Prove that if $f \in L^2[0,1]$, then

$$g(x) \;=\; \int_0^1 \frac{f(t)}{|x-t|^{1/2}}\,dt$$

exists for a.e. x, the function g defined in this way is measurable, belongs to $L^2[0,1]$, and satisfies $\|g\|_2 \le 2^{3/2}\,\|f\|_2$.

8. Let H be the set of all absolutely continuous functions $f \in \mathrm{AC}[a,b]$ such that $f(a)=0$ and $f' \in L^2[a,b]$. Prove that H is a Hilbert space with respect to the inner product $\langle f,g\rangle = \int_a^b f'(x)\,\overline{g'(x)}\,dx$.

9. Show that if $1 \le p \le \infty$ and $p \ne 2$, then the norm on ℓ^p is not induced from any inner product, i.e., there is no inner product $\langle \cdot,\cdot\rangle$ on ℓ^p such that $\langle x,x\rangle = \|x\|_p^2$ for all $x \in \ell^p$.

10. Fix $f \in L^2(0,\infty)$, and let

$$F(x) = \int_0^\infty \frac{f(t)}{1+xt}\,dt, \qquad \text{for } x>0.$$

Prove that F is continuous and differentiable on $(0,\infty)$.

11. Suppose that $f\colon \mathbb{R} \to \mathbb{C}$ is absolutely continuous on every compact interval and $f' \in L^2(\mathbb{R})$. Prove that $\sum\limits_{n\in\mathbb{Z}} |f(n+1)-f(n)|^2 < \infty$.

12. Let $\{x_n\}_{n\in\mathbb{N}}$ be a sequence of vectors in a Hilbert space H such that $\|x_n\|=1$ for every n. Suppose that "$\|x_m+x_n\| \to 2$ as $m,\,n\to\infty$," which means precisely that for every $\varepsilon>0$ there exists some $N>0$ such that

$$m,\,n \,\ge\, N \;\;\Longrightarrow\;\; 2-\varepsilon \,\le\, \|x_m+x_n\| \,\le\, 2+\varepsilon.$$

Prove that there exists some $x \in H$ such that $x_n \to x$.

13. Suppose that X is a complex Banach space whose norm $\|\cdot\|$ satisfies the Parallelogram Law. Prove that

$$\langle f,g\rangle \;=\; \frac{1}{4}\big(\|f+g\|^2 - \|f-g\|^2 + i\|f+ig\|^2 - i\|f-ig\|^2\big)$$

defines an inner product on X, and the norm induced from this inner product is the original norm $\|\cdot\|$ on X, i.e., $\|f\|^2 = \langle f,f\rangle$ for every $f \in X$.

Note: The proof is long and tedious.

Section 8.2: Orthogonality

Definition 8.2.1. State.

Lemma 8.2.2. State and prove. We will use this lemma to prove Theorem 8.2.11. It might be good for students to think about what this lemma means geometrically.

8.2.1 Orthogonal Complements

Definition 8.2.3 (Orthogonal Subsets). State.

Definition 8.2.4 (Orthogonal Complement). State.

Exercise 8.2.5. State. This is a good practice exercise for students.

Lemma 8.2.6. State. Assign the proof as reading.

8.2.2 Orthogonal Projections

Theorem 8.2.7 (Closest Point Theorem). State. I usually just sketch the proof of the existence of a closest point, especially how we get a Cauchy sequence, and assign the details and the proof of uniqueness for reading.

Definition 8.2.8 (Orthogonal Projection). State.

Example 8.2.9. State; this is motivation for the next lemma.

Lemma 8.2.10. State and prove. This lemma suggests that we might expect similar formulas to hold for infinite-dimensional closed subspaces, and we will see that this is the case in Section 8.3.

Note: By using the Gram–Schmidt orthogonalization procedure, we will prove in Theorem 8.3.10 that every finite-dimensional Hilbert space does have an orthonormal basis.

8.2.3 Characterizations of the Orthogonal Projection

Theorem 8.2.11. State and prove the implication (a) $\Rightarrow$ (b).

8.2.4 The Closed Span

Notation 8.2.12 (Closed Span). State. Explain that an upcoming theorem will show that if $A = \{x_n\}_{n\in\mathbb{N}}$ is an *orthonormal sequence* then we will have a very precise and convenient way to represent the elements of the closed span of A (specifically, if $A = \{x_n\}_{n\in\mathbb{N}}$ is an orthonormal sequence and $x \in \overline{\text{span}}(A)$ then x can be written as the "infinite linear combination" $x = \sum_{n=1}^{\infty} \langle x, x_n\rangle\, x_n$). However, if A is a generic sequence, then the most we can usually say is that $\overline{\text{span}}(A)$ is the smallest closed subspace that contains A (equivalently, it is the set of all limits of finite linear combinations of elements of A). In general it is **not** true that every element of $\overline{\text{span}}(A)$ has the form $\sum_{n=1}^{\infty} c_n x_n$ for some scalars c_n (one counterexample is given in Problem 8.4.13).

Exercise 8.2.13 (Smallest Closed Subspace). State.

8.2.5 The Complement of the Complement

Lemma 8.2.14. State and prove.

Exercise 8.2.15. State.

Corollary 8.2.16. State. The proof follows immediately from Exercise 8.2.15 and Lemma 8.2.14.

8.2.6 Complete Sequences

Definition 8.2.17 (Complete Sequence). State.

Corollary 8.2.18. State.

Problems: *Note on Problem* 8.2.24 in the text: Suppose that $\{x_n\}_{n\in\mathbb{N}}$ is a sequence in a Hilbert space H. A basic exercise is to show that the sequence is finitely linearly independent if and only if no vector x_m is a finite linear combination of the vectors x_n with $n \neq m$. That is

$$\{x_n\}_{n\in\mathbb{N}} \textit{ is independent} \quad\Longleftrightarrow\quad x_m \notin \operatorname{span}\{x_n\}_{n\neq m} \text{ for every } m \in \mathbb{N}.$$

However, even if a sequence is linearly independent, it is possible that some vector x_m may be in the *closed span* of the remaining vectors; an example is given in Extra Problem 1 below. Problem 8.2.24 defines a more stringent kind of independence; specifically, we say that

$$\{x_n\}_{n\in\mathbb{N}} \textit{ is minimal} \quad\Longleftrightarrow\quad x_m \notin \overline{\operatorname{span}}\{x_n\}_{n\neq m} \text{ for every } m \in \mathbb{N}.$$

Problem 7.4.11 introduced the notion of a *Schauder basis* for a Banach space. Every Schauder basis for a Hilbert space is both minimal and complete. However, a sequence that is both minimal and complete need not be a Schauder basis. For more details on Schauder bases and minimal sequences we refer to [Heil11].

Extra Problems for Section 8.2

1. Let $\{e_n\}_{n\in\mathbb{N}}$ be an orthonormal sequence in a Hilbert space H. Then the series

$$e_0 = \sum_{n=1}^{\infty} 2^{-n} e_n$$

converges by Theorem 8.3.1. Prove that the sequence $\{e_n\}_{n\geq 0}$ is finitely linearly independent, yet e_0 belongs to the closed span of the remaining vectors:

$$e_0 \in \overline{\text{span}}\{e_n\}_{n>0}.$$

Consequently $\{e_n\}_{n\geq 0}$ is not *minimal* in the sense introduced in Problem 8.2.24, even though it is finitely linearly independent.

2. Let E be the set of all even functions in $L^2(\mathbb{R})$. What is the orthogonal projection of $f \in L^2(\mathbb{R})$ onto E?

3. Let

$$M = \bigl\{x = (x_k)_{k\in\mathbb{N}} \in \ell^2(\mathbb{N}) : x_1 = x_2 = x_3\bigr\}.$$

Prove that M is a closed subspace of $\ell^2(\mathbb{N})$, and find the dimension of $M^\perp$.

4. Prove that

$$M = \biggl\{f \in L^2(\mathbb{R}) : \int_0^2 f(x)\,dx = 0\biggr\}$$

is a closed subspace of $L^2(\mathbb{R})$, and compute $M^\perp$.

5. In the Hilbert space $L^2[-\pi,\pi]$, let $f(x) = x^2$ and set

$$M = \text{span}\{1,\ \cos x,\ \sin x,\ \cos 2x,\ \sin 2x\}.$$

Find $\text{dist}(f, M)$.

6. Suppose that $\{x_n\}_{n\in\mathbb{N}}$ is a sequence in a Hilbert space H, and $y \in H$ is orthogonal to x_n for every n. Prove the following statements.

(a) y is orthogonal to every vector in $\text{span}\{x_n\}_{n\in\mathbb{N}}$.

(b) y is orthogonal to every vector in $\overline{\text{span}}\{x_n\}_{n\in\mathbb{N}}$.

(c) $\bigl(\{x_n\}_{n\in\mathbb{N}}\bigr)^\perp = \bigl(\overline{\text{span}}\{x_n\}_{n\in\mathbb{N}}\bigr)^\perp$.

7. Prove that $C[a,b]$ is a proper, dense subspace of $L^2[a,b]$, and therefore $C[a,b]$ is not complete with respect to the L^2-inner product. Give an explicit example, with proof, of a sequence in $C[a,b]$ that is Cauchy with respect to the L^2-norm but which does not converge to an element of $C[a,b]$.

8. This problem will characterize all of the inner products on $\mathbb{C}^d$. We say that a $d\times d$ matrix A with complex entries is *positive definite* if $Ax\cdot x > 0$ for all nonzero vectors $x \in \mathbb{C}^d$, where $x\cdot y$ denotes the usual dot product of vectors in $\mathbb{C}^d$.

(a) Suppose that S is an invertible $d\times d$ matrix and Λ is a diagonal matrix whose diagonal entries are all positive. Prove that $A = S\Lambda S^{\mathrm{H}}$ is a positive definite matrix, where $S^{\mathrm{H}} = \overline{S^{\mathrm{T}}}$ is the complex conjugate of the transpose of S (usually referred to as the *Hermitian* of S).

Note: In fact, every positive definite matrix has this form, but that is not needed for this problem.

(b) Show that if A is a positive definite $d \times d$ matrix, then

$$\langle x, y \rangle_A \;=\; Ax \cdot y, \qquad \text{for } x, y \in \mathbb{C}^d,$$

defines an inner product on $\mathbb{C}^d$.

(c) Show that if $\langle \cdot, \cdot \rangle$ is an inner product on $\mathbb{C}^d$, then there exists some positive definite $d \times d$ matrix A such that $\langle \cdot, \cdot \rangle = \langle \cdot, \cdot \rangle_A$.

9. Suppose that $f, g \colon [a,b] \to \mathbb{R}$ are such that g is differentiable everywhere on $[a,b]$ and $g'(x) = f(x)^2$. Prove that $f \in L^1[a,b]$.

10. Let H be a Hilbert space, and let U and V be any two subspaces of H (not necessarily closed or orthogonal). The *sum* of U and V is

$$U + V \;=\; \{f + g : f \in U,\, g \in V\}.$$

Prove the following statements.

(a) $U^\perp = \overline{U}^\perp$.

(b) $(U+V)^\perp \;=\; U^\perp \cap V^\perp$.

(c) $\overline{U^\perp + V^\perp} \subseteq (U \cap V)^\perp$.

(d) If U and V are closed subspaces, then $\overline{U^\perp + V^\perp} = (U \cap V)^\perp$.

11. Compute the maximum value of

$$\left| \int_{-1}^{1} x^5 f(x)\, dx \right|$$

over all $f \in L^2[-1,1]$ such that $\|f\|_2 = 1$ and

$$\int_{-1}^{1} f(x)\, dx \;=\; \int_{-1}^{1} x f(x)\, dx \;=\; \int_{-1}^{1} x^2 f(x)\, dx \;=\; 0.$$

12. For this problem we take scalars to be real. Let S be the "closed first quadrant" in ℓ^2. That is, S is the set of all sequences $y = (y_k)_{k \in \mathbb{N}}$ in ℓ^2 such that $y_k \geq 0$ for every k.

(a) Show that S is a closed, convex subset of ℓ^2.

(b) Given $x \in \ell^2$, find the sequence $y \in S$ that is closest to x.

13. Let $\{x_n\}_{n \in \mathbb{N}}$ be a sequence in a Hilbert space H. Using the definitions of a *Schauder basis* (Problem 7.4.11), a *minimal sequence* (Problem 8.2.24), an *ω-independent sequence* (Extra Problems to Section 7.4), and a *linearly independent sequence*, prove the following implications:

$$\text{Schauder basis} \implies \text{minimal} \implies \omega\text{-independent} \implies \text{independent}.$$

Show by example that none of the converses to these implications hold in general.

Remark: For more details on Schauder bases and related topics, see [Heil11].

Section 8.3: Orthonormal Sequences and Orthonormal Bases

8.3.1 Orthonormal Sequences

Theorem 8.3.1. State and prove.

8.3.2 Unconditional Convergence

Definition 8.3.2 through Example 8.3.5. I usually state the definition of unconditional convergence and discuss it briefly. However, this topic can just be assigned for reading, as it is not crucial to any of the proofs.

Note: If $\sum\limits_{n=1}^{\infty} c_n$ is a series of real *scalars* that converges conditionally (that is, it converges but does not converge unconditionally), then given any real number x there exists a permutation $\sigma\colon \mathbb{N} \to \mathbb{N}$ such that $\sum\limits_{n=1}^{\infty} c_{\sigma(n)}$ converges and equals x. There are also permutations such that $\sum\limits_{n=1}^{\infty} c_{\sigma(n)}$ diverges to infinity or does not converge at all; see [Heil11, Lemma 3.3] for details.

8.3.3 Orthogonal Projections Revisited

Theorem 8.3.6. State, prove part (a). Given an orthonormal sequence $\{e_n\}_{n\in\mathbb{N}}$, this theorem characterizes the orthogonal projection onto the closed span of the sequence, and also tells us exactly how to represent the elements of the closed span as "infinite linear combinations" of the vectors e_n.

8.3.4 Orthonormal Bases

Theorem 8.3.7. State, prove a sample implication.

Definition 8.3.8 (Orthonormal Basis). State.

Note: If $\{e_n\}_{n\in\mathbb{N}}$ is an orthonormal basis for H, then each $x \in H$ has a unique representation of the form $x = \sum_{n=1}^{\infty} c_n(x)\, e_n$. Using the terminology of Banach space theory, this says that $\{e_n\}_{n\in\mathbb{N}}$ is a *Schauder basis* for H (compare Problem 7.4.11). In fact, because these series expansions converge unconditionally for every x, the sequence is an *unconditional Schauder basis*; see [Heil11] or [Heil18].

Note: Suppose that $\{x_n\}_{n\in\mathbb{N}}$ is a sequence in H that is both complete and linearly independent. This *does not imply* that $\{x_n\}_{n\in\mathbb{N}}$ is a Schauder basis. Theorem 8.3.7 tells us that *completeness plus orthonormality* implies Schauder. However, *completeness plus independence* is simply not enough on its own to guarantee that we have a Schauder basis. A few counterexamples appear in the problems, e.g., Problem 8.3.27 shows that the set of monomials $\{x^k\}_{k\ge 0}$ is a complete, linearly independent sequence in $L^2[a,b]$, but it is not a Schauder basis for that space. Indeed, that problem shows that there are proper subsets of the monomials that are still complete in $L^2[a,b]$ (no proper subset of a Schauder basis can be complete!).

Note: The *Müntz–Szász Theorem* (see [Heil11, Thm. 5.6] for the precise statement) is a striking result that characterizes the sequences of monomials that are complete in $C[a,b]$. Since $C[a,b]$ is dense in $L^2[a,b]$, any such sequence will be complete in $L^2[a,b]$. For the case $0 < a < b < \infty$, this theorem says that if $0 < n_1 \le n_2 \le \cdots$ is an increasing sequence of integers and $n_k \to \infty$, then $\{x^{n_k}\}_{k\in\mathbb{N}}$ is complete in $C[a,b]$ if and only if

$$\sum_{k=1}^{\infty} \frac{1}{n_k} = \infty.$$

Example 8.3.9. Mention briefly.

8.3.5 Existence of an Orthonormal Basis

Theorem 8.3.10. It may be enough to just say that Gram–Schmidt works in any finite-dimensional Hilbert space.

Theorem 8.3.11. State, perhaps briefly sketch the idea.

The only difficulty in this proof is that we need to know that H contains a complete, linearly independent sequence, so that we have something to apply Gram–Schmidt to and can be sure that the resulting sequence will be complete. One way to do this is to note that if H is separable then it contains a countable dense sequence $\{x_n\}_{n\in\mathbb{N}}$. Such a sequence cannot be independent (it is analogous to the set of rationals in the real line), but we can extract a linearly independent subsequence through a recursive process. Choose the first nonzero vector in the sequence, then proceed down the list until you find the first vector that is not in the span of the previous ones, and repeat.

8.3.6 The Legendre Polynomials

Discuss briefly.

8.3.7 The Haar System

Theorem 8.3.12 (The Haar System). State and prove.

Note: This proof is a nice application of the Lebesgue Differentiation Theorem, but it does not give any insight into the construction of other wavelet orthonormal bases for $L^2(\mathbb{R})$. Wavelets are discussed in detail in Chapter 12 of [Heil11].

8.3.8 Unitary Operators

The results of this subsection will not be used in the rest of the chapter; indeed, they will only be referred to in Section 9.4, where they are used to facilitate the definition of the Fourier transform on $L^2(\mathbb{R})$. Consequently I do not cover this subsection in class.

For a more complete treatment of operator theory see [Heil18].

Note: The operator $U(x) = \big(\langle x, e_n\rangle\big)_{n\in\mathbb{N}}$ defined in Theorem 8.3.17 is called the *analysis operator* associated with the sequence $\{e_n\}_{n\in\mathbb{N}}$. For more details, see [Heil11, Ch. 7] or [Heil18].

Extra Problems for Section 8.3

1. Prove that if a Hilbert space H is separable, then every orthogonal set of nonzero vectors in H is countable.

2. Let $u_1, \dots, u_d$ be the columns of a $d \times d$ matrix U. Prove that $\{u_1, \dots, u_d\}$ is an orthonormal basis for $\mathbb{C}^d$ if and only if $U^{\mathrm{H}}U = I$, where $U^{\mathrm{H}} = \overline{U^{\mathrm{T}}}$ is the Hermitian of U.

3. Does there exist a function $f \in L^2[a,b]$ such that

$$\int_a^b x f(x)\,dx = 1 \qquad \text{and} \qquad \int_a^b x^k f(x)\,dx = 0, \quad \text{for } k = 0, 2, 3, \dots?$$

4. Find real numbers a, b, and c that minimize the quantity

$$\int_{-1}^{1} |x^3 - a - bx - cx^2|^2\,dx.$$

5. Let $\{e_n\}_{n\in\mathbb{N}}$ be an orthonormal basis for a Hilbert space H. Determine whether the following are closed subsets of H:

$$S = \bigg\{x \in H : \sum_{n=1}^{\infty} |\langle x, e_n\rangle| \le 1\bigg\}, \qquad T = \bigg\{x \in H : \sum_{n=1}^{\infty} |\langle x, e_n\rangle| \ge 1\bigg\}.$$

6. Suppose that $\{e_n\}_{n\in\mathbb{N}}$ is an orthonormal basis for a Hilbert space H, and $\{x_n\}_{n\in\mathbb{N}}$ is an orthonormal sequence such that $\sum_{n=1}^{\infty} \|e_n - x_n\|^2 < \infty$. Prove that $\{x_n\}_{n\in\mathbb{N}}$ is an orthonormal basis for H.

Hint: Consider the case $\sum \|e_n - x_n\|^2 < 1$ first.

7. Suppose that $\{e_n\}_{n\in\mathbb{N}}$ is an orthonormal basis for a Hilbert space H, and $\{x_n\}_{n\in\mathbb{N}}$ is any sequence that satisfies $\sum \|e_n - x_n\|^2 < \infty$. Prove that the orthogonal complement of $\overline{\text{span}}\{x_n\}_{n\in\mathbb{N}}$ is finite-dimensional.

8. Show that if $\{x_n\}_{n\in\mathbb{N}}$ is a complete sequence in a Hilbert space H that satisfies

$$\biggl\|\sum_{n=1}^{N} c_n x_n\biggr\|^2 = \sum_{n=1}^{N} |c_n|^2$$

for all $N \in \mathbb{N}$ and scalars $c_1, \dots, c_N$, then $\{x_n\}_{n\in\mathbb{N}}$ is an orthonormal basis for H.

9. Assume that $\{e_n(x)\}_{n\in\mathbb{N}}$ is an orthonormal basis for $L^2[0,1]$, and fix g in $L^\infty[0,1]$. Prove that $\{g(x)\, e_n(x)\}_{n\in\mathbb{N}}$ is an orthonormal basis for $L^2[0,1]$ if and only if $|g(x)| = 1$ a.e. on $[0,1]$.

10. Let H be any infinite-dimensional Hilbert space. Prove that there exist closed subspaces $K_t \subseteq H$ for $t \in \mathbb{R}$ such that if $s < t$ then $K_s \subsetneq K_t$.

11. The definition of weak convergence in a Hilbert space was given in Problem 8.1.10. Give an example of a sequence of functions $\{f_n\}_{n\in\mathbb{N}}$ in $L^2(\mathbb{R})$ such that $\|f_n\|_2 = 1$ for every n, yet f_n converges weakly to 0 as $n \to \infty$.

12. Lebesgue measure on $\mathbb{R}^d$ has the properties that all balls have positive and finite measure, and that measure is invariant under translation. Show that if H is an infinite-dimensional Hilbert space, then there does not exist a measure on H that has analogous properties.

Here, a "measure" is a function μ that is defined on some class of subsets of H that includes all countable unions of open balls, is countably additive, monotonic, translation-invariant on balls, and satisfies $0 < \mu\bigl(B_r(x)\bigr) < \infty$ for every open ball.

Section 8.4: The Trigonometric System

Theorem 8.4.1. State and discuss. Orthonormality is easy, but completeness is nontrivial. In Chapter 9 we will use convolution and approximate identities to give a proof of Theorem 8.4.1. For now we will simply take the completeness of the trigonometric system as given. A different (but still nontrivial) proof of completeness based on the *Stone–Weierstrass Theorem* can be found in [Heil18, Thm. 4.7.2].

Theorem 8.4.2 (Fourier Series for $L^2[0,1]$). State and discuss.

Note: From the viewpoint of a harmonic analyst, the trigonometric system is *the* most important example of an orthonormal basis.

Note: We can view the domain of a 1-periodic function as being the interval $[0,1)$ under the operation of addition modulo 1. This is isomorphic to the circle group $S^1 = \{z \in \mathbb{C} : |z| = 1\}$ (under the operation of multiplication). The circle is the one-dimensional torus, hence our use of the symbol $\mathbb{T}$ to denote the domain of a 1-periodic function.

Exercise 8.4.3. Optional. Gabor systems are central to *time-frequency analysis*, and are discussed in detail in Chapter 11 of [Heil11], and in Gröchenig's text [Grö01].

Extra: The HRT Conjecture

The HRT (Heil–Ramanathan–Topiwala) Conjecture is (as of 2024) an open mathematical problem, set in the Hilbert space $L^2(\mathbb{R})$. It is also known as the *Linear Independence of Time-Frequency Translates Conjecture.*

HRT Conjecture. If $g \in L^2(\mathbb{R})$ is not the zero function and $\{(\alpha_k, \beta_k)\}_{k=1}^N$ is any set of finitely many distinct points in $\mathbb{R}^2$, then $\{e^{2\pi i\beta_k x} g(x-\alpha_k)\}_{k=1}^N$ is a linearly independent set of functions in $L^2(\mathbb{R})$. That is,

$$\sum_{k=1}^{N} c_k \, e^{2\pi i \beta_k x} g(x-\alpha_k) \,=\, 0 \quad \text{a.e.} \qquad \Longrightarrow \qquad c_1 \,=\, \cdots \,=\, c_N \,=\, 0. \qquad \diamondsuit$$

Despite the striking simplicity of the *statement* of the conjecture, it appears to be a surprisingly difficult problem. For expanded discussion of this conjecture and the partial results that are known about it, see the exposition in [Heil11, Sec. 11.9], or the following two survey papers.

C. Heil, Linear independence of finite Gabor systems, in: *Harmonic Analysis and Applications,* Birkhäuser, Boston, 2006, pp. 171–206.

C. Heil and D. Speegle, The HRT Conjecture and the Zero Divisor Conjecture for the Heisenberg group, in: *Excursions in Harmonic Analysis*, Volume 3, R. Balan et al., eds., Birkhäuser, Boston (2015), pp. 159–176.

The conjecture was originally stated in the following paper.

C. Heil, J. Ramanathan, and P. Topiwala, Linear independence of time-frequency translates, *Proc. Amer. Math. Soc.* **124** (1996), pp. 2787–2795.

Various special cases where the conjecture can be proved to be true are known. For example, the conjecture is known to be true if N is 1, 2, or 3. However, the following special case of the conjecture is currently open!

HRT Subconjecture. If $g \in L^2(\mathbb{R})$ is not the zero function then the set of four functions

$$\bigl\{g(x),\;\; g(x-1),\;\; e^{\,2\pi i x} g(x),\;\; e^{2\pi i \sqrt{2} x} g(x-\sqrt{3})\bigr\}$$

is linearly independent in $L^2(\mathbb{R})$. $\diamondsuit$

In fact, this subconjecture is open even if we impose the assumption that $g \in L^2(\mathbb{R})$ is *infinitely differentiable*, or that g belongs to the *Schwartz space* of infinitely differentiable, rapidly decreasing functions.

CHAPTER 9: CONVOLUTION AND THE FOURIER TRANSFORM

Sadly, I am typically unable to fit this chapter into a one-semester course—there usually just isn't enough time in the semester. This is unfortunate both because this material is central to my field of harmonic analysis, and because it contains many beautiful applications of topics from previous chapters. Below I give some suggestions and comments for instructors who are presenting this chapter, or for readers studying the chapter on their own.

Note: I do have a volume on harmonic analysis in preparation. That text will present convolution, the Fourier transform, Fourier series, and other topics in more detail than is found here.

Section 9.1: Convolution

Convolution was introduced and briefly discussed in Section 4.6.3; now we take a more in-depth look at this operation.

Definition 9.1.1 (Convolution). State.

Note: It is important for this definition that $\mathbb{R}$ is a group (under addition in this case). Convolution can be defined more generally on any *locally compact group* (although there is a difference between right-convolution and left-convolution if the group is not abelian).

Exercise 9.1.2. State. Observe that $\mathrm{supp}(\chi * \chi) \subseteq \mathrm{supp}(\chi) + \mathrm{supp}(\chi)$; this is a general feature of convolutions (compare Problem 4.6.28).

Note: The *n*th-*order B-spline function* is $B_n = \chi_{[0,1]} * \cdots * \chi_{[0,1]}$, where there are n factors in the convolution.

Theorem 9.1.3. State. This theorem is a review of properties of convolution that were established in Section 4.6.3.

Exercise 9.1.4. State.

Theorem 9.1.5. State and prove.

Exercise 9.1.6. State.

Subsection 9.1.3 and Figure 9.2. Motivate convolution as an averaging process.

Exercise 9.1.7. Motivate and state.

Definition 9.1.8 (Approximate Identities). Motivate and state.

Exercise 9.1.9. State.

Exercise 9.1.10 (The Fejér Kernel). State.

Theorem 9.1.11. State and prove.

Theorem 9.1.12. State and prove.

Exercise 9.1.13 and Theorem 9.1.14 (Young's Inequality). I would just state the inequality $\|f * g\|_p \le \|f\|_p \, \|g\|_1$, and refer to the details given in the text for methods of proof.

Mention Problem 9.1.21, which establishes the more general form of Young's Inequality, $\|f * g\|_r \le \|f\|_p \, \|g\|_q$ where $\frac{1}{p} + \frac{1}{q} = 1 + \frac{1}{r}$.

Theorem 9.1.15. State and prove.

Exercise 9.1.16. State.

Exercise 9.1.17. State.

Extra Problems for Section 9.1

1. Use approximate identities to prove that if $f \in L^1(\mathbb{R})$ satisfies $\int_{-\infty}^{\infty} fg = 0$ for every $g \in C_c(\mathbb{R})$, then $f = 0$ a.e.

2. Suppose that $g \in L^1(\mathbb{R})$ is such that for every $\varphi \in C_b^1(\mathbb{R})$ we have

$$(g * \varphi')(x) \;=\; \varphi(x+h) - \varphi(x-h) \quad \text{for all } x \in \mathbb{R}.$$

Prove that $g = \chi_{[-h,h]}$ a.e.

3. Define $\gamma(x) = e^{-1/x^2}\, \chi_{(0,\infty)}(x)$, and prove the following statements.

(a) For each $n \in \mathbb{N}$, there exists a polynomial p_n of degree $3n$ such that

$$\gamma^{(n)}(x) \;=\; p_n(x^{-1})\, e^{-x^{-2}}\, \chi_{(0,\infty)}(x).$$

(b) $\gamma \in C^\infty(\mathbb{R})$ and $\gamma^{(n)}(0) = 0$ for every $n \ge 0$.

(c) If $a < b$ then the function $f(x) = \gamma(x-a)\, \gamma(b-x)$ belongs to $C_c^\infty(\mathbb{R})$ and satisfies $f(x) > 0$ for $x \in (a,b)$ and $f(x) = 0$ for $x \notin (a,b)$.

4. (a) Suppose that $f \in L^1(\mathbb{R})$. Prove that $f * f = 0$ a.e. if and only if $f = 0$ a.e.

(b) Use part (a) to prove that if $A \subseteq \mathbb{R}$ has positive and finite measure, then $A + A$ contains an open interval.

(c) Show that if A and B are subsets of $\mathbb{R}$ that each have positive measure, then $A + B$ contains an open interval.

5. Assume $k \in L^1(\mathbb{R})$ is given. Set $r = \int_{-\infty}^{\infty} k$, and define $k_N(x) = Nk(Nx)$.

(a) Prove that if $1 \le p < \infty$, then for each $f \in L^p(\mathbb{R})$ we have that $f * k_N \to rf$ in L^p-norm as $N \to \infty$. Note that this includes the possibility that r may be complex or zero.

(b) Prove that if $f \in C_0(\mathbb{R})$ then $f * k_N \to rf$ uniformly as $N \to \infty$.

6. Suppose that $\{f_n\}_{n\in\mathbb{N}}$ is a sequence of functions in $L^1(\mathbb{R})$ such that

$$\lim_{n\to\infty} \int_{-\infty}^{\infty} f_n(x)\, g(x)\, dx \;=\; g(0), \qquad \text{for every } g \in C_0(\mathbb{R}).$$

Prove that $\{f_n\}_{n\in\mathbb{N}}$ is not Cauchy in $L^1(\mathbb{R})$.

Section 9.2: The Fourier Transform

Definition 9.2.1 (Fourier Transform on $L^1(\mathbb{R})$). State.

Note: When speaking aloud, we usually pronounce the symbols $\widehat{f}$ as "*f-hat*."

Note: A trivial but often useful fact is that

$$\widehat{f}(0) \;=\; \int_{-\infty}^{\infty} e^{-2\pi i 0 x} f(x)\, dx \;=\; \int_{-\infty}^{\infty} f(x)\, dx.$$

Remark 9.2.2. Discuss in as much detail as you feel appropriate.

Lemma 9.2.3. State and prove.

Note: Technically, the Dominated Convergence Theorem applies to sequences indexed by the natural numbers. To rigorously justify its application in the proof of Lemma 9.2.3 (and in other similar instances), we should use the approach of Problem 4.5.30 and consider all possible sequences $\eta_k \to 0$.

Example 9.2.4. State and prove.

Note: The sinc function is also known as the *cardinal sine*, and indeed "sinc" is a contraction of *sinus cardinalis*. The "cardinal" nature of the sinc function is the fact that it is an *interpolating function*, because $\mathrm{sinc}(0) = 1$ while $\mathrm{sinc}(n) = 0$ for all integers $n \neq 0$.

Note: The *Dirichlet function* is

$$\bigl(\chi_{[-1/2\pi,1/2\pi]}\bigr)^{\wedge}(\xi) \;=\; \frac{\sin \xi}{\pi \xi}.$$

Theorem 9.2.5 (Riemann–Lebesgue Lemma). State and prove.

Note: Let $\mathcal{F}$ denote the Fourier transform as an operator. That is, consider the linear mapping $\mathcal{F}(f) = \widehat{f}$ that takes f to $\widehat{f}$. The Riemann–Lebesgue Lemma tells us that $\mathcal{F}\colon L^1(\mathbb{R}) \to C_0(\mathbb{R})$. Further, we saw earlier that

$$\|\mathcal{F}(f)\|_{\mathrm{u}} \;=\; \|\widehat{f}\|_{\mathrm{u}} \;\le\; \|f\|_1.$$

Using the language of operator theory, this inequality says that $\mathcal{F}$ is a *bounded* operator from $L^1(\mathbb{R})$ into $C_0(\mathbb{R})$, because the norm of $\widehat{f}$ is no more than a fixed constant (in this case the constant is 1) times the norm of f. Later we will see the *Uniqueness Theorem*, which tells us that $\mathcal{F}$ is injective. However, $\mathcal{F}$ is not surjective; instead, the remarks after Theorem 9.2.5 say that the range of $\mathcal{F}$ is a dense but proper subspace of $C_0(\mathbb{R})$. A (nontrivial) consequence of these facts is that the inverse mapping from $\mathrm{range}(\mathcal{F}) = A(\mathbb{R})$ back to $L^1(\mathbb{R})$ is *unbounded* (this follows from the *Open Mapping Theorem*, which is usually covered in the second semester of our graduate real analysis sequence).

Exercise 9.2.6. State.

Corollary 9.2.7. State and prove.

Note: A δ-function, if one existed, would be a function $\delta \in L^1(\mathbb{R})$ that is zero at all points except $x = 0$, yet has an "infinite spike" at $x = 0$ with enough "mass" concentrated into the spike that $\int_{-\infty}^{\infty} \delta(x)\, dx = 1$. There is no such function, for even if we did define

$$\delta(x) \;=\; \begin{cases} \infty, & \text{if } x = 0, \\ 0, & \text{if } x \neq 0, \end{cases} \qquad \text{(A)}$$

then δ is zero almost everywhere, and therefore $\int_{-\infty}^{\infty} \delta(x)\, dx = 0$. The function δ defined by equation (A) is *not* an identity for convolution—instead it is a representative of the zero function in $L^1(\mathbb{R})$!

Still, suppose for the moment that there *were* such a delta function. In this case we would have

$$\int_{-\infty}^{\infty} f(x)\, \delta(x)\, dx \;=\; f(0) \qquad \text{(B)}$$

for every f (but note that this equation ***does not hold*** because $\delta = 0$ a.e.!). Such a function δ *does not exist*, but even so it is not uncommon in the literature to see δ defined by equation (A) and then used as in equation (B). What is really happening here is one of the "standard abuses of notation" that occur in mathematics. There does exist a *measure* δ for which the equality $\int_{-\infty}^{\infty} f(x)\, d\delta(x) = f(0)$ does hold. However, δ is not a *function* defined by equation (A); rather it is the *measure of sets* defined by $\delta(E) = 1$ if the set E contains the origin, and $\delta(E) = 0$ otherwise:

$$\delta(E) \;=\; \begin{cases} 1, & \text{if } 0 \in E, \\ 0, & \text{if } 0 \notin E. \end{cases}$$

The integral $\int_{-\infty}^{\infty} f(x)\, d\delta(x)$ is *not* a *Lebesgue integral of the function* $f(x)\,\delta(x)$, but instead is an *integral of the function* f *with respect to the measure* δ. As long as we understand that notation is being abused and δ is a measure rather than a function, then this abuse of notation is not usually problematic. However, problems can arise when we begin to think that δ is an actual function, rather than a measure.

An alternative but essentially equivalent way to rigorously define δ is as a *distribution*, or *generalized function*, instead of a measure.

Definition 9.2.8 (Inverse Fourier Transform on $L^1(\mathbb{R})$). State.

Note: We usually pronounce the symbols $\check{f}$ as "*f-check*."

Theorem 9.2.9 (Inversion Formula). Motivate and state.

Lemma 9.2.10. Motivate and state.

Lemma 9.2.11. Motivate and state.

Proof of Theorem 9.2.9. Restate the theorem and prove.

Corollary 9.2.12. State and prove.

Note: An alternative proof is to apply Lemma 9.2.11 directly. Specifically, if $\widehat{f} = 0$ a.e., then that lemma implies that

$$(f * w_N)(x) \;=\; \int_{-N}^{N} \widehat{f}(\xi)\left(1 - \frac{|\xi|}{N}\right) e^{2\pi i \xi x}\, d\xi \;=\; 0.$$

Since $f * w_N \to f$ in L^1-norm, it follows that $f = 0$ a.e.

Theorem 9.2.13. State and prove.

Theorem 9.2.14. State and prove.

Corollary 9.2.15. State and prove.

Corollary 9.2.16. State and prove.

Note: There is a **TYPO** in Problem 9.2.20 in the text. Change the definition "$\psi = \chi_{[0,\frac{1}{2})} - \chi_{[-\frac{1}{2},0]}$" to "$\psi = \chi_{[-\frac{1}{2},0)} - \chi_{[0,\frac{1}{2})}$".

Section 9.3: Fourier Series

Introduction. Define $L^p(\mathbb{T})$.

Note: We can view the domain of a 1-periodic function as being the interval $[0, 1)$ under the operation of addition modulo 1. This is isomorphic to the circle group $S^1 = \{z \in \mathbb{C} : |z| = 1\}$ (under the operation of multiplication). The circle is the one-dimensional torus, hence our use of the symbol $\mathbb{T}$ to denote the domain of a 1-periodic function.

Remark 9.3.1. Discuss briefly.

Exercise 9.3.2 (Riemann–Lebesgue Lemma). State.

Theorem 9.3.3. State and prove.

Exercise 9.3.4. State.

Definition 9.3.5 (Convolution). State.

Exercise 9.3.6. State.

Definition 9.3.7 (Approximate Identity). State.

Exercise 9.3.8. State.

Exercise 9.3.9. State.

Lemma 9.3.10. Motivate by discussing the Fejér and Dirichlet kernels; state and prove.

Theorem 9.3.11 (Inversion Formula). State and prove.

Corollary 9.3.12 (Uniqueness Theorem). State and prove.

Theorem 9.3.13. State and prove.

Corollary 9.3.14 (The Trigonometric System is an ONB). State and prove.

Example 9.3.15 and the remarks following it. Discuss.

Theorem 9.3.16. State and discuss.

Theorem 9.3.17. State and discuss.

Theorem 9.3.18 (Carleson–Hunt Theorem). State and discuss.

Extra Problems for Section 9.3

1. Define an appropriate "dilation method" for constructing approximate identities on $\mathbb{T}$. For example, let k be any integrable function on $\mathbb{R}$ such that $k(x) = 0$ for all $x \notin [0,1)$ and $\int_{-\infty}^{\infty} k = 1$. For each $N \in \mathbb{N}$, set $g_N(x) = N\,k(Nx)$ for $x \in \mathbb{R}$. Then let k_N be the 1-periodic extension of g_N from $[0,1)$ to the real line, i.e., let $k_N(x+j) = g_N(x)$ for all $x \in [0,1)$ and $j \in \mathbb{Z}$. Prove that $\{k_N\}_{N\in\mathbb{N}}$ is an approproximate identity on $\mathbb{T}$.

2. Assume that f is a measurable function on $[a,b]$. Given any $\delta, \varepsilon > 0$, show that there exists a function $g \in \mathrm{AC}[a,b]$ and a measurable set $A \subseteq [a,b]$ such that $|A| < \delta$ and $\sup\limits_{x\notin A} |f(x) - g(x)| < \varepsilon$.

3. Given $f \in L^2(\mathbb{T})$ and $a \in \mathbb{R}$, define $T_a f(x) = f(x-a)$. Suppose that $\alpha \in \mathbb{R}$ is irrational. Prove that $f = T_\alpha f$ a.e. if and only if f is equal almost everywhere to a constant function.

4. Assume that a 1-periodic function f is absolutely continuous on $[0,1]$, and

$$\int_0^1 f(x)\,dx \;=\; \int_0^1 f(x)\,e^{-2\pi i x}\,dx \;=\; \int_0^1 f(x)\,e^{2\pi i x}\,dx \;=\; 0.$$

Prove that $\|f'\|_2 \ge 4\pi\|f\|_2$.

Section 9.4: The Fourier Transform on $L^2(\mathbb{R})$

Lemma 9.4.2. Motivate, state, and prove.

Definition 9.4.3 (The Fourier Transform on $L^2(\mathbb{R})$). Cover the discussion given before the definition, which shows that the Fourier transform is well-defined on $L^2(\mathbb{R})$. State the definition.

Definition 9.4.3 (The Fourier Transform on $L^2(\mathbb{R})$).

Lemma 9.4.4. State and prove.

Lemma 9.4.5. State and prove.

Theorem 9.4.6. State and prove.

Example 9.4.7. Discuss.

Lemma 9.4.8. Assign for reading.

Extra Problems for Section 9.4

1. The text shows how to extend the Fourier transform from a dense subspace of $L^2(\mathbb{R})$ to all of $L^2(\mathbb{R})$. This is a special case of the following more general problem.

Let X and Z be metric spaces such that Z is complete, and let Y be a proper dense subset of X. Suppose that $f\colon Y \to Z$ is uniformly continuous. Prove that there exists a uniformly continuous function $g : X \to Z$ such that $g(x) = f(x)$ for every $x \in Y$.

Extra: A Nowhere Differentiable Function

A continuous, nowhere differentiable function was constructed in 1830 by B. Balzano but went unnoticed. (Note: Confirmation and references for these statements are needed; if you have knowledge of these historical details please let me know.) The first widely recognized continuous, nowhere differentiable function was constructed by Karl Weierstrass (1815–1897). He showed that

$$g(x) \;=\; \sum_{m=0}^{\infty} a^{-m}\cos(b^m x) \qquad\qquad \text{(A)}$$

is nowhere differentiable if b is an odd integer and $b/a > 1 + (3\pi/2)$. G. H. Hardy (1877–1947) showed that the function g in equation (A) is nowhere differentiable for every choice of real numbers $b \ge a > 1$, as is the function

$$h(x) \;=\; \sum_{m=0}^{\infty} a^{-m}\sin(b^m x)$$

for the same range of parameters.

We will use the Fourier transform to give a construction of a similar continuous function that is not differentiable at any point. Our proof is adapted from:

J. Johnsen, Simple proofs of nowhere-differentiability for Weierstrass's function and cases of slow growth, *J. Fourier Anal. Appl.*, **16** (2010), pp. 17–33.

A Nowhere Differentiable Function. Fix $0 < \alpha \le 1$ and define

$$g(x) \;=\; \sum_{m=0}^{\infty} 2^{-\alpha m}\, e^{2\pi i 2^m x} \qquad x \in \mathbb{R}.$$

The series defining g converges absolutely with respect to the uniform norm, so g is continuous (and 1-periodic, i.e., $g(x+1) = g(x)$ for every x). The real part of the function g corresponding to the choice $\alpha = 1/3$ is pictured in Figure 9.A.

Let K be any infinitely differentiable function such that $K(1) = 1$ and $\mathrm{supp}(K) \subseteq \left[\frac{1}{2}, 2\right]$. Then K belongs to the Schwartz class $\mathcal{S}(\mathbb{R})$ that was introduced in Problem 9.2.32.It is shown in that problem that the Fourier transform maps $\mathcal{S}(\mathbb{R})$ onto itself. Therefore, the function $k = \check{K}$ belongs to $\mathcal{S}(\mathbb{R})$. Hence $\widehat{k} = K$ is infinitely differentiable, $\widehat{k}(1) = 1$, and $\mathrm{supp}(\widehat{k}) \subseteq \left[\frac{1}{2}, 2\right]$. In particular, $\widehat{k}(n) = 0$ for all integers $n \ne 1$. Therefore

$$\int_{-\infty}^{\infty} k \;=\; \widehat{k}(0) \;=\; 0,$$

so if we set $k_\lambda(x) = \lambda k(\lambda x)$ then $\{k_\lambda\}_{\lambda\in\mathbb{N}}$ is *not* an approximate identity in $L^1(\mathbb{R})$. Even so, it will still be useful for our analysis.

The function g is continuous and 1-periodic on $\mathbb{R}$, so its convolution with k_λ is well-defined and belongs to $C_b(\mathbb{R})$. Considering $\lambda = 2^n$, we compute that

$$\begin{aligned}
(g * k_{2^n})(x) &= \int_{-\infty}^{\infty} g(x-y)\, k_{2^n}(y)\, dy \\
&= \sum_{m=0}^{\infty} 2^{-\alpha m}\, e^{2\pi i 2^m x} \int_{-\infty}^{\infty} e^{-2\pi i 2^m y}\, k_{2^n}(y)\, dy \\
&= \sum_{m=0}^{\infty} 2^{-\alpha m}\, e^{2\pi i 2^m x}\, \widehat{k_{2^n}}(2^m) \\
&= \sum_{m=0}^{\infty} 2^{-\alpha m}\, e^{2\pi i 2^m x}\, \widehat{k}(2^{m-n}) \\
&= 2^{-\alpha n}\, e^{2\pi i 2^n x}.
\end{aligned}$$

The interchange of integration and summation in the calculation above can be justified by using Fubini's Theorem.

Next, using the fact that $\int k = 0$ we see that

$$\begin{aligned}
2^{(1-\alpha)n}\, e^{2\pi i 2^n x} &= 2^n\, (g * k_{2^n})(x) \;-\; 2^n\, g(x) \int_{-\infty}^{\infty} k_{2^n}(y)\, dy \\
&= \int_{-\infty}^{\infty} \bigl(g(x-y) \;-\; g(x)\bigr)\; 2^n k(2^n y)\; 2^n dy \\
&= \int_{-\infty}^{\infty} \frac{g(x-y) \;-\; g(x)}{y}\; 2^n y\, k(2^n y)\; 2^n dy \\
&= \int_{-\infty}^{\infty} \frac{g(x-2^{-n}y) \;-\; g(x)}{2^{-n}y}\; y\, k(y)\; dy. \qquad \text{(B)}
\end{aligned}$$

Now, if g is differentiable at x then the function

$$F(h) \;=\; \frac{g(x-h) - g(x)}{h}$$

is bounded for h small, and it is also bounded for h large since g is continuous and bounded. Let C be a constant such that $|F(h)| \le C$ for all $h \in \mathbb{R}$. Then

$$|F(2^{-n}y)\, yk(y)| \;\le\; C\, |yk(y)| \;\in\; L^1(\mathbb{R}).$$

Therefore we can apply the Lebesgue Dominated Convergence Theorem to equation (B) to obtain

$$\begin{aligned}\lim_{n\to\infty} 2^{(1-\alpha)n}\, e^{2\pi i 2^n x} &= \lim_{n\to\infty} \int_{-\infty}^{\infty} \frac{g(x-2^{-n}y) \,-\, g(x)}{2^{-n}y}\; y\, k(y)\; dy \\ &= -g'(x) \int_{-\infty}^{\infty} y\, k(y)\, dy.\end{aligned}$$

However, since $0 < \alpha \le 1$ the quantity $2^{(1-\alpha)n}\, e^{2\pi i 2^n x}$ does not converge as $n \to \infty$. This contradiction implies that g cannot be differentiable at x. Since x is arbitrary, we conclude that g is not differentiable at any point.

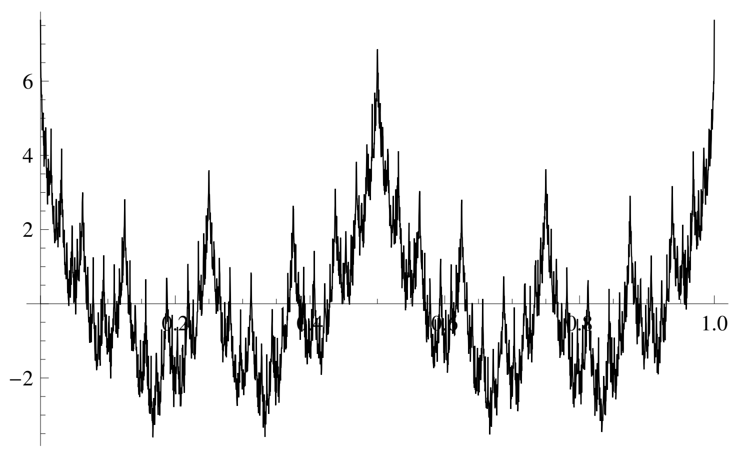

Fig. 9.A Real part of the nowhere differentiable function g corresponding to the choice $\alpha = 1/3$.