Harmonic Analysis: from Fourier to Haar

María Cristina Pereyra

Lesley A. Ward

Department of Mathematics and Statistics, MSC03 2150, 1 University of New Mexico, Albuquerque, NM 87131-0001, USA
E-mail address: crisp@math.unm.edu

School of Mathematics and Statistics, University of South Australia, Mawson Lakes SA 5095, Australia
E-mail address: Lesley.Ward@unisa.edu.au

2000 *Mathematics Subject Classification.* Primary 42-xx

CHAPTER 2

Interlude

In this chapter we describe some useful classes of functions, sets of measure zero, and various ways (known as *modes of convergence*) in which a sequence of functions can approximate another function. We then describe situations in which we are allowed to interchange limiting operations. Finally we state several density results that appear in different guises throughout the book. A leitmotif in analysis, in particular in harmonic analysis, is that often it is easier to prove results for a class of simpler functions that are dense in a larger class, and then to use limiting processes to pass from the approximating simpler functions to the more complicated limiting function.

This chapter should be read once to get a glimpse of the variety of function spaces and modes of convergence that can be considered. It should be revisited as necessary while the reader gets acquainted with the theory of Fourier series and integrals, and with the other types of time–frequency decompositions described in this book.

2.1. Nested classes of functions on bounded intervals

Here we introduce some classes of functions on bounded intervals I (the intervals can be open, closed, or neither) and on the unit circle $\mathbb{T}$. In Chapters 7 and 8 we introduce analogous function classes on $\mathbb{R}$, and we go beyond functions and introduce *distributions* (generalized functions) in Section 8.2.

2.1.1. Riemann-integrable functions. We review the concept of the Riemann integral, following the exposition in the book by Terence Tao[1] [**Tao1**, Chapter 11]. We highly recommend Tao's book to our readers.

The development of the Riemann integral is in terms of approximation by step functions. We begin with some definitions.

Definition 2.1. Given a bounded interval I, a *partition* P of I is a finite collection of disjoint intervals $\{J_k\}_{k=1,\dots,n}$ such that $I = \bigcup_{k=1}^n J_i$. The intervals J_i are contained in I and are necessarily bounded. $\Diamond$

Definition 2.2. The *characteristic function* $\chi_J(x)$ of an interval J is defined to be the function that takes the value one if x lies in J, and zero otherwise:

$$\chi_J(x) = \begin{cases} 1, & \text{if } x \in J; \\ 0, & \text{if } x \notin J. \end{cases}$$

$\Diamond$

[1]Terence Tao (1975–) is an Australian mathematician. He was awarded the Fields Medal in 2006.

FIGURE 2.1. Graph of the step function $h = 3\chi_{[-1,2)} - \chi_{[2,4]} + 2\chi_{(5,6)}$ in Example 2.4.

The building blocks in the Riemann integration theory are finite linear combinations of characteristic functions of disjoint intervals, called *step functions*.

DEFINITION 2.3. A function h defined on I is a *step function* if there are a partition P of the interval I, and real numbers $\{a_J\}_{J\in P}$, such that

$$h(x) = \sum_{J\in P} a_J \chi_J(x),$$

where $\chi_J(x)$ is the characteristic function of the interval J. ◊

For step functions there is a very natural definition of integral. Let us first consider an example.

EXAMPLE 2.4. The function $h = 3\chi_{[-1,2)} - \chi_{[2,4]} + 2\chi_{(5,6)}$ is a step function defined on the interval $[-1, 6)$. Its graph is shown in Figure 2.1. ◊

If we ask the reader to compute the integral of this function over the interval $[-1, 6)$, she will add (with appropriate signs) the areas of the rectangles in the picture to get $(3 \times 3) + (-1 \times 2) + (2 \times 1) = 9$. More generally, we can define the *integral of a step function h associated to the partition P over the interval I* by

$$\int_I h(x)\,dx := \sum_{J\in P} a_J |J|, \tag{2.1}$$

where $|J|$ denotes the length of the interval J. A given step function could be associated to more than one partition, and one can verify that this definition of the integral is independent of the partition chosen. Notice that the integral of the characteristic function of a subinterval J of I is its length $|J|$.

EXERCISE 2.5. Given an interval $J \subset I$, show that

$$\int_I \chi_J(x)\,dx = |J|.$$

◊

DEFINITION 2.6. A function f is *bounded on I* if there is a constant $M > 0$ such that $|f(x)| \le M$ for all $x \in I$. If so, we say that f is *bounded by M*. ◊

EXERCISE 2.7. Verify that step functions are always bounded functions. ◊

DEFINITION 2.8. A function $f : I \to \mathbb{R}$ is *Riemann integrable* if f is bounded on I, and if the supremum of the integrals of all step functions whose graphs lie below the graph of f is equal to the infimum of the integrals of all step functions whose graphs lie above that of f. In mathematical notation,

$$\sup_{h_1\le f} \int_I h_1(x)\,dx = \inf_{h_2\ge f} \int_I h_2(x)\,dx < \infty,$$

FIGURE 2.2. This figure illustrates the approximation of a function f from below and from above by step functions h_1 and h_2 respectively, as in the definition of the Riemann integral.

where the supremum and infimum are taken over step functions $h_1(x) \leq f(x) \leq h_2(x)$ for all $x \in I$. We denote the common value by

$$\int_I f(x)\,dx, \quad \int_a^b f(x)\,dx, \quad \text{or simply} \quad \int_I f,$$

where a and b ($a \leq b$) are the endpoints of the interval I. We denote the class of Riemann-integrable functions on I by $\mathcal{R}(I)$. See Figure 2.2 illustrating the approximation of f from below and from above by step functions. $\Diamond$

Riemann integrability is a property preserved under linear transformations, taking maximums and minimums, absolute values, and products. It also preserves order, meaning that if f and g are Riemann-integrable functions over I and if $f < g$ then $\int_I f < \int_I g$. The *triangle inequality*

$$\left|\int_I f\right| \leq \int_I |f|$$

holds for Riemann-integrable functions. See [**Tao1**, Section 11.4].

EXERCISE 2.9. (*The Riemann Integral is Linear*) Given two step functions h_1 and h_2 defined on I, verify that for c, $d \in \mathbb{R}$, the linear combination $ch_1 + dh_2$ is again a step function. Also check that

$$\int_I \big(ch_1(x) + dh_2(x)\big)\,dx = c\int_I h_1(x)\,dx + d\int_I h_2(x)\,dx.$$

Warning: The partitions P_1 and P_2 associated to the step functions h_1 and h_2 might be different. Try to find a partition that works for both (a *common refinement*).

Use Definition 2.8 to show that this linearity property is inherited by the Riemann-integrable functions. $\Diamond$

One can deduce from Definition 2.8 the following very useful proposition.

PROPOSITION 2.10. *A bounded function f is Riemann integrable over I if and only if for each $\varepsilon > 0$ there are step functions h_1 and h_2 defined on I below and above f, that is $h_1(x) \leq f(x) \leq h_2(x)$ for all $x \in I$, such that the integral of their difference is bounded by ε. More precisely,*

$$0 \leq \int_I \big(h_2(x) - h_1(x)\big)\,dx < \varepsilon.$$

In particular, if f is Riemann-integrable then there exists a step function h such that

$$\int_I |f(x) - h(x)|\,dx < \varepsilon.$$

FIGURE 2.3. The graph illustrates the integral of the difference of the two step functions in Figure 2.2.

See Figure 2.3 illustrating the integral of the difference of the two step functions in Figure 2.2.

ASIDE 2.11. *Readers familiar with the definition of the Riemann integral via Riemann sums may like to observe that the integral of a step function that is below (respectively above) a given function f is always below a corresponding* lower *(respectively above a corresponding* upper*)* Riemann sum[2] *for f.*

Our first observation is that if f is Riemann integrable on $[-\pi, \pi)$, then the zeroth Fourier coefficient of f is well defined:

$$a_0 = \widehat{f}(0) = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(\theta)\, d\theta.$$

We say that a *complex-valued function* $f : I \to \mathbb{C}$ is *Riemann integrable* if its real and imaginary parts are Riemann integrable. If f is Riemann integrable, then the complex-valued functions $f(\theta)e^{-in\theta}$ are Riemann integrable for each $n \in \mathbb{Z}$, and all the Fourier coefficients $a_n = \widehat{f}(n)$, $n \neq 0$, are also well defined.

EXERCISE 2.12. Verify that if $f : [-\pi, \pi) \to \mathbb{C}$ is Riemann integrable then the function $g : [-\pi, \pi) \to \mathbb{C}$ defined by $g(\theta) = f(\theta)e^{-in\theta}$ is Riemann integrable. $\Diamond$

Here are some examples of familiar functions that are Riemann integrable and for which we can define Fourier coefficients. See Theorem 2.29 for a complete characterization of Riemann-integrable functions.

EXAMPLE 2.13. Uniformly continuous functions on intervals are Riemann integrable. In particular continuous functions on closed intervals are Riemann integrable; see [**Tao1**, Section 11.5]. Monotone[3] bounded functions on intervals are Riemann integrable; see [**Tao1**, Section 11.6]. $\Diamond$

Here is an example of a function that is not Riemann integrable.

EXAMPLE 2.14. (*Dirichlet's Example*) The function

$$f(x) = \begin{cases} 1, & \text{if } x \in [a,b] \cap \mathbb{Q}; \\ 0, & \text{if } x \in [a,b] \setminus \mathbb{Q} \end{cases}$$

---

[2]A *Riemann sum* associated to a bounded function $f : I \to \mathbb{R}$ and to a given partition P of the interval I is a real number $R(f, P)$ defined by $R(f, P) = \sum_{J \in P} f(x_J)\, |J|$, where x_J is a point in J for each $J \in P$. The *upper (resp. lower) Riemman sum*, $U(f, P)$ (resp. $L(f, P)$ associated to f and P is obtained similarly replacing $f(x_J)$ by $\sup_{x \in J} f(x)$ (resp. $\inf_{x \in J} f(x)$), Note that since f is bounded both the infimum and supremum over J exist.

[3]A function $f : I \to \mathbb{R}$ is monotone if it is either increasing on the whole of I, or decreasing on the whole of I; that is, either for all x, $y \in I$, $x \leq y$, we have $f(x) \leq f(y)$, or for all x, $y \in I$, $x \leq y$, we have $f(x) \geq f(y)$.

is bounded (by one), but it is not Riemann integrable on $[0,1]$. Notice also that it is discontinuous everywhere. $\Diamond$

EXERCISE 2.15. Show that the Dirichlet function is not Riemann integrable. $\Diamond$

2.1.2. Lebesgue integrable functions and L^p-spaces. There are other more general notions of integral; see for example the book by Robert G. Bartle [**Bart2**]. We will consider the *Lebesgue integral*.

The Lebesgue integral is studied in depth in courses on *measure theory*. For a brief introduction see [**Tao2**, Chapters 18, 19]. For more in-depth presentations at a level intermediate between advanced undergraduate and graduate students see [**SS2**] or [**Bart1**]. Classical graduate textbooks include the books by Gerald B. Folland [**Fol**] and by Halsey L. Royden [**Roy**].

In this section we simply aim to give some ideas about what the Lebesgue integral is and what the L^p-spaces are, and we skim over the many subtleties of measure theory.

Every Riemann-integrable function on a bounded interval I is also Lebesgue-integrable on I, and the values of its Riemann and Lebesgue integrals are the same. The Dirichlet function in Example 2.14 is Lebesgue integrable but not Riemann integrable, with Lebesgue integral equal to zero; see Remark 2.33 at the end of Section 2.1.4. In practice when you see an integral, it will usually suffice to think of Riemann integrals even if sometimes we really mean Lebesgue integrals.

In parallel to the definition of the Riemann integral just presented, we sketch the definition of the Lebesgue integral over the bounded interval I. We use as scaffolding the so-called simple functions instead of step functions.

Simple functions are finite linear combinations of characteristic functions of disjoint *measurable sets*. The *characteristic*[4] *function* of a set $A \subset \mathbb{R}$ is defined, similarly to the characteristic function of an interval, to be

$$\chi_A(x) = \begin{cases} 1, & \text{if } x \in A; \\ 0, & \text{otherwise.} \end{cases} \tag{2.2}$$

Intervals are measurable sets, and we can assign to them a *measure*, namely their length. However there are sets that are non-measurable, and we need to exclude such sets. Once we assign a measure to each measurable set, we can define the Lebesgue integral of the characteristic function of a measurable set to be the measure of the set. The Lebesgue integral of a simple function is then defined to be the corresponding finite linear combination of the measures of the underlying sets, exactly as in (2.1), except that now the sets J are disjoint measurable subsets of the measurable set I, and the length $|J|$ is replaced by the measure $m(J)$ of the set J. One can then define Lebesgue integrability of a function $f : I \to \mathbb{R}$ in analogy to Definition 2.8, replacing the boundedness condition by requiring the function to be measurable on I, and replacing step functions by simple functions. A function $f : I \to \mathbb{R}$ is *measurable* if the pre-image under f of each measurable subset of $\mathbb{R}$ is a measurable subset of I. It all boils down to understanding measurable sets and their measures.

[4]In statistics and probability the characteristic function of a set A is called the *indicator function* of the set, and denoted $\mathbb{I}_A$. The term "characteristic function" refers to the Fourier transform of a probability density (so-called because it "characterizes" the probability distribution).

In this book we do not define measurable sets, or their measure, except for sets of measure zero in $\mathbb{R}$ in Section 2.1.4, and for intervals or countable unions of disjoint intervals, whose measure is given, quite naturally, by the sum of the lengths of the intervals.

A measurable function is said to belong to $L^1(I)$ if $\int_I |f(x)|\,dx < \infty$, where the integral is in the sense of Lebesgue. In fact, there is a continuum of Lebesgue spaces, the *L^p-spaces*, defined for each real number p such that $1 \leq p < \infty$ by

$$L^p(I) := \left\{ f : I \to \mathbb{C} : \int_I |f(x)|^p\,dx < \infty \right\},$$

and for $p = \infty$ by

$$L^\infty(I) := \{ f : I \to \mathbb{C} : f \text{ is essentially bounded}\}.$$

The L in L^p and L^∞ stands for Lebesgue.

All the functions involved are assumed to be measurable. In Definition 2.30 we state precisely what *essentially bounded* means. For now, just think of L^∞ functions as *bounded* functions on I. Let us denote by $B(I)$ the space of bounded functions on I:

$$B(I) := \{ f : I \to \mathbb{C} : f \text{ is bounded}\}.$$

There are essentially bounded functions that are not bounded, so $B(I) \subset L^\infty(I)$, but $L^\infty(I)$ is larger.

These spaces are *normed spaces*[5] (after defining properly what it means for a function to be zero in L^p, see Remark 2.33) with the L^p-norm defined to be

$$\|f\|_{L^p(I)} := \left(\int_I |f(x)|^p\,dx \right)^{1/p}, \quad 1 \leq p < \infty, \tag{2.3}$$

$$\|f\|_{L^\infty(I)} := \operatorname*{ess\,sup}_{x \in I} |f(x)|, \quad \|f\|_{B(I)} := \sup_{x \in I} |f(x)|. \tag{2.4}$$

In Definition 2.31 we state precisely what *essential supremum of f* means.

The *L^p-metric* or *L^p-distance* d_p is the metric induced by the L^p-norm, for each p with $1 \leq p \leq \infty$. The L^p-distance between two functions f, $g \in L^p(I)$ is given by the norm of their difference:

$$d_p(f, g) := \|f - g\|_{L^p(I)}. \tag{2.5}$$

A property of normed spaces that we use repeatedly is the *triangle inequality.* In our setting it states that for all f, $g \in L^p(I)$, $1 \leq p \leq \infty$,

$$\|f + g\|_{L^p(I)} \leq \|f\|_{L^p(I)} + \|g\|_{L^p(I)}. \tag{2.6}$$

The Riemann-integrable functions are not only contained in all L^p-spaces for $1 \leq p \leq \infty$, but they are also *dense* in each $L^p(I)$ in its L^p-metric; see Definition 2.62 below. In other words, given any function $f \in L^p(I)$ we can find a Riemann-integrable function arbitrarily close to f in the L^p-metric. Furthermore, for each p the space $L^p(I)$ is the *completion* of $\mathcal{R}(I)$ in the L^p-metric. Being the completion of $\mathcal{R}(I)$ automatically implies that $\mathcal{R}(I)$ is dense in $L^p(I)$, but is a stronger statement. It also implies that the limit of each convergent sequence in

[5]See Appendix B for the definition.

$L^p(I)$ must be a function in $L^p(I)$. Finally, we can think of a *Cauchy*[6] *sequence* as being a sequence that has the right to converge. Every Cauchy sequence in $L^p(I)$ does converge, and its limit is a function in $L^p(I)$.

Informally, the *completion* Z of a subspace Y of a metric space X in the metric of X is the smallest subspace of X that contains Y and that also contains all limits of Cauchy sequences in Z.

ASIDE 2.16. *The reader may be familiar with the construction of the real numbers $\mathbb{R}$ as the completion of the rational numbers $\mathbb{Q}$ in the Euclidean metric $d(x,y) := |x-y|$. One can think of $\mathcal{R}(I)$ as playing the role of $\mathbb{Q}$ and $L^p(I)$ as playing the role of $\mathbb{R}$. There is a procedure that allows the completion of any metric space, mirroring the procedure used in the construction of the real numbers. See for example* [**Tao2**, Exercise 12.4.8].

In fact, we could have *defined* the Lebesgue spaces $L^p(I)$ as the completion of $\mathcal{R}(I)$ in the L^p-metric, instead of attempting to define the Lebesgue integral via measure theory. The important thing is that both processes lead to the same spaces. This fact is so important that we state it (without proof) as a theorem, for future reference.

THEOREM 2.17. *For each p with $1 \leq p \leq \infty$, the space $L^p(I)$ is the completion of the Riemann-integrable functions $\mathcal{R}(I)$ in the L^p-metric. In particular, every Cauchy sequence in $L^p(I)$ converges to a function in $L^p(I)$.*

Therefore the L^p-spaces with the L^p-norms are *complete normed spaces*. Such spaces are so common and important that they have a special name, *Banach*[7] *spaces*. See Appendix B for the definition of a complete normed space, and see Section 2.4 for some results about density of one space in another. One learns about Banach spaces in a course on functional analysis. Classical graduate textbooks include [**Fol**] and the book by Martin Schechter [**Sch**].

Among the L^p-spaces, both L^1 and L^∞ play special roles. The space

$$L^2(I) := \left\{ f : I \to \mathbb{C} : \int_I |f(x)|^2 \, dx < \infty \right\}$$

of *square-integrable* functions is also very important. The term *square-integrable* emphasizes that the integral of the square of (the absolute value of) $f \in L^2(I)$ is finite. Functions in $L^2(I)$ are also said to have *finite energy*.

The L^2-norm is induced by the *inner product*

$$\langle f, g \rangle_{L^2(I)} := \int_I f(x)\overline{g(x)} \, dx, \tag{2.7}$$

meaning that $\|f\|_{L^2(I)} = \sqrt{\langle f, f \rangle_{L^2(I)}}$.

The space $L^2(I)$ of square-integrable functions over I is a *complete inner-product vector space* (a *Hilbert*[8] *space*), and so it has geometric properties very similar to those of n-dimensional Euclidean space $\mathbb{R}^n$. In particular the notion

[6] The idea of a Cauchy sequence is that the terms must get arbitrarily close to each other as $n \to \infty$. In our setting, the sequence $\{f_n\} \subset L^p(I)$ is a Cauchy sequence if for each $\varepsilon > 0$ there is an $N > 0$ such that for all n, $m > N$, $\|f_n - f_m\|_{L^p(I)} < \varepsilon$. Every convergent sequence is Cauchy, but the converse does not hold unless the space is complete.

[7] Named after the Polish mathematician Stefan Banach (1892–1945).

[8] Named after the German mathematician David Hilbert (1862–1943).

of *orthogonality* is very important. See Appendix B for a more detailed description of inner-product vector spaces and Hilbert spaces. Another reference is [**SS2**, Chapters 4, 5].

In Chapter 5 we will be especially interested in the Hilbert space $L^2(\mathbb{T})$ of square-integrable functions on $\mathbb{T}$, with the L^2-norm

$$\|f\|_{L^2(\mathbb{T})} := \left(\frac{1}{2\pi}\int_{-\pi}^{\pi} |f(\theta)|^2\, d\theta\right)^{1/2}$$

induced by the inner product

$$\langle f, g\rangle := \frac{1}{2\pi}\int_{-\pi}^{\pi} f(\theta)\overline{g(\theta)}\, d\theta.$$

The factor $1/(2\pi)$ is just a normalization constant that forces the trigonometric functions $e_n(\theta) := e^{in\theta}$ to have norm one in $L^2(\mathbb{T})$.

Exercise 2.18. Verify that

$$\langle e_n, e_m\rangle = \delta_{n,m}.$$

In particular, $\|e_n\|_{L^2(\mathbb{T})} = 1$. In the language of Chapter 5, the trigonometric system is an *orthonormal* system. $\Diamond$

2.1.3. Ladder of function spaces on $\mathbb{T}$. So far in this chapter, we have been concerned with function spaces defined by integrability properties. Earlier we encountered function spaces defined by differentiability properties. It turns out that for functions defined on a given closed bounded interval, all these function spaces are nicely nested. The inclusions are summarized in Figure 2.4 below.

Continuous functions on a closed bounded interval $I = [a, b]$ are Riemann integrable, and hence Lebesgue integrable. The collection of continuous functions on $I = [a, b]$ is denoted by $C([a, b])$. Next, $C^k([a, b])$ denotes the collection of k times continuously differentiable functions on $[a, b]$, in other words functions whose k^{th} derivatives exist and are continuous. For convenience we sometimes use the expression 'f is C^k' to mean '$f \in C^k$'. Finally, $C^\infty([a, b])$ denotes the collection of functions on $[a, b]$ that are differentiable infinitely many times. (At the endpoints a and b, we mean continuous or differentiable from the right and from the left respectively.) Notice that there exist functions that are differentiable but not continuously differentiable; the canonical example being $f(x) = x^2 \sin(/x)$ for $x \neq 0$, $f(0) = 0$.

Exercise 2.19. Verify that $f(x) = x^2 \sin(1/x)$ for $x \neq 0$, $f(0) = 0$, is differentiable everywhere but its derivative is not continuous at $x = 0$.

Recall that in Section 1.3.2 we discussed periodic functions on the unit circle $\mathbb{T} = [-\pi, \pi)$, as well as functions that are continuously differentiable on $\mathbb{T}$. In that context we meant that $f : \mathbb{T} \to \mathbb{C}$ is C^k if f is k times continuously differentiable on $\mathbb{T}$ and if, when f is extended periodically, $f^{(j)}(-\pi) = f^{(j)}(\pi)$ for all $0 \leq j \leq k$. This boundary condition implies that when we restrict f to the closed interval $[-\pi, \pi]$ it is C^k, in particular it is Riemann integrable on $[-\pi, \pi)$.

Define $L^p(\mathbb{T})$ to be $L^p([-\pi, \pi))$.

For complex-valued functions defined on $\mathbb{T}$, or on any closed bounded interval $I = [a, b]$, all these spaces are nicely nested. There is a ladder of function spaces in

$\mathbb{T}$, running from the small space $C^\infty(\mathbb{T})$ up to the large space $L^1(\mathbb{T})$:

$$\begin{aligned} C^\infty(\mathbb{T}) \subset C^k(\mathbb{T}) \subset C^2(\mathbb{T}) \subset C^1(\mathbb{T}) \subset C(\mathbb{T}) \\ \subset \mathcal{R}(\mathbb{T}) \subset B(\mathbb{T}) \subset L^\infty(\mathbb{T}) \subset L^{p_1}(\mathbb{T}) \subset L^{p_2}(\mathbb{T}) \subset L^1(\mathbb{T}), \end{aligned} \tag{2.8}$$

for integers $k \geq 2$ and real numbers p_1 and p_2 such that $p_2 < p_1$. In particular, notice that

$$L^2(\mathbb{T}) \subset L^1(\mathbb{T}).$$

Warning: For functions defined on the real line, the L^p-spaces are no longer nested, nor are the continuous functions necessarily bounded or integrable.

Earlier we encountered a different family of function spaces on $\mathbb{T}$, namely the collections of 2π-periodic trigonometric polynomials of degree at most N, for $N = 0$, 1, 2, ... We denote these spaces by $\mathcal{P}_N(\mathbb{T})$. More precisely,

$$\mathcal{P}_N(\mathbb{T}) := \Big\{ \sum_{|n| \leq N} a_n e^{in\theta} : a_n \in \mathbb{C},\ |n| \leq N \Big\}.$$

These spaces are clearly nested and increasing as N increases, and they are certainly subsets of $C^\infty(\mathbb{T})$. Thus we can extend the above ladder to one starting at $\mathcal{P}_0(\mathbb{T})$ (the constant functions) and climbing all the way to $C^\infty(\mathbb{T})$, by including the nested spaces

$$\mathcal{P}_0(\mathbb{T}) \subset \mathcal{P}_{N_1}(\mathbb{T}) \subset \mathcal{P}_{N_2}(\mathbb{T}) \subset C^\infty(\mathbb{T}),$$

for all N_1, N_2 with $0 \leq N_1 \leq N_2$.

The step functions defined in Section 2.1.1 are Riemann integrable, although most of them have discontinuities. Our ladder has a branch:

$$\{ \text{ step functions on } I\} \subset \mathcal{R}(I).$$

The simple functions mentioned in Section 2.1.2 are bounded functions that are not necessarily Riemann integrable. However, they are Lebesgue integrable, and in fact simple functions are in $L^p(I)$ for each p such that $1 \leq p \leq \infty$. Our ladder has a second branch:

$$\{ \text{ simple functions on } I\} \subset B(I).$$

Exercise 2.20. Compute the L^p-norm of a step function defined on an interval I. Compute the L^p-norm of a simple function defined on an interval I. Use the notation $m(A)$ for the measure of a measurable set $A \subset I$. ◊

We will encounter other classes of functions on $\mathbb{T}$, $[a, b]$, or $\mathbb{R}$, such as functions that are *piecewise continuous* or *piecewise smooth*, *monotone functions*, functions of *bounded variation*, *Lipschitz*[9] *functions*, and *Hölder*[10]*-continuous functions*. We define what we mean as they appear in the text. As you encounter these spaces, you are encouraged to try to fit them in to the ladder of spaces in Figure 2.4, which will start looking more like a tree with branches and leaves.

The results of the following exercises justify some of the inclusions in the ladder of function spaces. Operate with the Lebesgue integrals as you would operate with Riemann integrals. Try to specify which properties of the integral you are using.

[9]Named after the German mathematician Rudolf Otto Sigismund Lipschitz (1832–1903).
[10]Named after the German mathematician Otto Ludwig Hölder (1859–1937).

EXERCISE 2.21. Let I be a bounded interval. Show that if $f : I \to \mathbb{C}$ is bounded by $M > 0$ and $f \in L^p(I)$, then $f \in L^q(I)$ for all q such that $p \leq q < \infty$. Furthermore,

$$\int_I |f(x)|^q\,dx \leq M^{q-p} \int_I |f(x)|^p\,dx < \infty,$$

and so in this situation the L^q-norm is controlled by the L^p-norm. In the language of norms, the preceding inequality can be written as

$$\|f\|^q_{L^q(I)} \leq \|f\|^{q-p}_{L^\infty(I)}\,\|f\|^p_{L^p(I)},$$

since $\|f\|_{L^\infty(I)}$ is the infimum of the numbers M that bound f. ◊

For the previous exercise you may find *Hölder's inequality* useful.

LEMMA 2.22 (Hölder's Inequality). *Suppose I is a bounded interval, $1 \leq s \leq \infty$, $\dfrac{1}{s} + \dfrac{1}{t} = 1$, $f \in L^s(I)$, and $g \in L^t(I)$. Then*

$$\left|\int_I f(x)g(x)\,dx\right| \leq \|f\|_{L^s(I)}\|g\|_{L^t(I)}; \tag{2.9}$$

in other words the integral of the product of f and g is smaller in absolute value than the product of the L^s-norm of f and the L^t-norm of g.

Hölder's inequality reappears in later chapters, and it is discussed and proved, together with other very important inequalities, at the end of the book. The special case when $s = t = 2$ is the celebrated *Cauchy–Schwarz*[11] *inequality*. Inequalities are inherent to the way of analysis, and we will see them over and over again throughout the book.

EXERCISE 2.23. Let I be a bounded interval. Show that if $1 \leq p \leq q < \infty$ and $h \in L^q(I)$ then $h \in L^p(I)$. Moreover, show that if $h \in L^q(I)$ then

$$\|h\|_{L^p(I)} \leq |I|^{1/p-1/q}\|h\|_{L^q(I)}.$$

Thus for those exponents, $L^q(I) \subset L^p(I)$. **Hint:** Apply Hölder's inequality on I, with $s = q/p > 1$, $f = |h|^p$, $g = 1$, and $h \in L^q(I)$. ◊

2.1.4. Sets of measure zero. For many purposes we often need only that a function has a given property except on an exceptional set of points x. Such exceptional sets are small. We now introduce the appropriate notion of smallness, given by the concept of a *set of measure zero*.

DEFINITION 2.24. A set $E \subset \mathbb{R}$ has *measure zero* if, for each $\varepsilon > 0$, there are open intervals I_j such that E is covered by the collection $\{I_j\}_{j\in\mathbb{N}}$ of intervals, and the total length of the intervals is less than ε:

$$\text{(i)} \quad E \subset \bigcup_{j=1}^{\infty} I_j, \quad \text{and} \quad \text{(ii)} \quad \sum_{j=1}^{\infty} |I_j| < \varepsilon.$$

[11]This inequality is named after the French mathematician Augustin Louis Cauchy (1789–1857), and the German mathematician Hermann Amandus Schwarz (1843–1921). The Cauchy–Schwarz inequality was discovered 25 years earlier by the Ukranian mathematician Viktor Yakovlevich Bunyakovskii (1804–1889), and it is also known as the Cauchy–Bunyakovskii–Schwarz inequality.

We say that a given property holds *almost everywhere* if the set of points where the property does not hold has measure zero. We use the abbreviation *a.e.* for *almost everywhere*[12]. $\Diamond$

In particular, every countable set has measure zero. There are uncountable sets that have measure zero. The most famous example is the *Cantor*[13] *set*, discussed in Example 2.26.

EXERCISE 2.25. Show that finite subsets of $\mathbb{R}$ have measure zero. Show that countable subsets of $\mathbb{R}$ have measure zero. In particular the set $\mathbb{Q}$ of rational numbers has measure zero. $\Diamond$

EXAMPLE 2.26. (*The Cantor set*) The Cantor set C is obtained from the closed interval $[0, 1]$ by a sequence of successive deletions of open intervals called the *middle thirds*, as follows. Remove the middle-third interval $(1/3, 2/3)$ of $[0, 1]$. We are left with two closed intervals: $[0, 1/3]$ and $[2/3, 1]$. Remove the two middle-third intervals $(1/9, 2/9)$ and $(7/9, 8/9)$ of the remaining closed intervals. We are left with four closed intervals: $[0, 1/9]$, $[2/9, 1/3]$, $[2/3, 7/9]$, and $[8/9, 1]$. Delete the four middle-third intervals $(1/27, 2/27)$, $(7/27, 8/27)$, $(19/27, 20/27)$, and $(25/27, 26/27)$ of the remaining closed intervals.

Continuing in this way, at the n^{th} stage we are left with 2^n disjoint closed intervals each of length 3^{n-1}. Removing the 2^n middle-third intervals, each of length $1/3^n$, we are left with 2^{n+1} disjoint closed intervals of length $1/3^n$. This process can be continued indefinitely, so that the set A of all points removed from the closed interval $[0, 1]$ is the union of a collection of disjoint open intervals. Thus $A \subset [0, 1]$ is an open set.

The Cantor set C is defined to be the set of points that remain after A is deleted: $C := [0, 1] \setminus A$.

The Cantor set can be seen to coincide with the set of numbers in $[0, 1]$ whose *ternary expansion* (base three) uses only the digits 0 and 2. Once this identification is made, one can also verify that C is in a one-to-one correspondence (except for certain points for which the correspondence is two-to-one) with the interval $[0, 1]$. Therefore C has the same cardinality as $\mathbb{R}$, and so C is an uncountable set. However, C has measure zero, because its complement A has *measure one*[14]: the sum of the lengths of the disjoint intervals removed is

$$m(A) := 1 \times \frac{1}{3} + 2 \times \frac{1}{9} + 4 \times \frac{1}{27} + \cdots = \sum_{n=1}^{\infty} \frac{2^{n-1}}{3^n} = 1.$$

The last equality holds because we are dealing with a geometric sum with ratio $r = 2/3 < 1$. $\Diamond$

EXERCISE 2.27. Verify that the Cantor set C is closed. Use Definition 2.24 to verify that C has measure zero. $\Diamond$

[12] In probability theory it is used instead *a.s.* for *almost surely*, in French *p.p.* for *presque partout*, in Spanish *c.s.* for *casi siempre*.

[13] Named after the German mathematician Georg Ferdinand Ludwig Philipp Cantor (1845–1918).

[14] Recall that a natural way of defining the length or measure of a subset A of $\mathbb{R}$ which is a countable union of disjoint intervals $\{I_j\}$ is to declare the measure of A to be the sum of the lengths of the intervals I_j.

Aside 2.28. *We assume that the reader is familiar with basic notions of point set topology on* $\mathbb{R}$*, and in particular with the concepts of* open *and* closed *sets in* $\mathbb{R}$*. To review these ideas, see for example* [**Tao2**, Chapter 12].

Here is an important characterization of Riemann-integrable functions; it appeared in Lebesgue's doctoral dissertation. A proof can be found in Stephen Abbott's book [**Abb**, Section 7.6].

Theorem 2.29 (Lebesgue's Theorem, 1901). *A bounded function* $f : I \to \mathbb{C}$ *is Riemann integrable if and only if* f *is continuous almost everywhere.*

We can now define what the terms essentially bounded and essential supremum mean.

Definition 2.30. A function $f : \mathbb{T} \to \mathbb{C}$ is *essentially bounded* if it is bounded except on a set of measure zero. More precisely, there exists a constant $M > 0$ such that $|f(x)| \le M$ a.e. in I. $\lozenge$

Definition 2.31. Let $f : I \to \mathbb{C}$ be an essentially bounded function. The *essential supremum* $\operatorname{ess\,sup}_{x\in I} |f(x)|$ of f is defined to be the smallest constant M such that $|f(x)| \le M$ a.e. on the interval I. In other words,

$$\operatorname*{ess\,sup}_{x\in I} |f(x)| := \inf\{M \ge 0 : |f(x)| \le M \ \text{ a.e. in } I\}.$$

$\lozenge$

Exercise 2.32. Verify that if f is bounded on I then f is essentially bounded. In other words, the space $\mathcal{B}(I)$ of bounded functions on I is a subset of $L^\infty(I)$. Show that it is a proper subset. Furthermore, verify that if f is bounded then $\operatorname{ess\,sup}_{x\in I} |f(x)| = \sup_{x\in I} |f(x)|$. $\lozenge$

Remark 2.33. For the purposes of Lebesgue integration theory, functions that are equal almost everywhere are regarded as being the same. To be really precise, one should talk about equivalence classes of L^p-integrable functions: f is equivalent to g if and only if $f = g$ a.e. In that case $\int_I |f - g|^p = 0$, and if $f \in L^p(I)$ then so is g. We say $f = g$ in $L^p(I)$ if $f = g$ a.e. In that case, $\int_I |f|^p = \int_I |g|^p$. In particular the Dirichlet function g in Example 2.14 is equal a.e. to the zero function, which is in $L^p([0,1])$. Hence the Dirichlet function g is in $L^p([0,1])$, its integral in the Lebesgue sense vanishes, and its L^p-norm is zero for all $1 \le p \le \infty$.

Now that we have given the relevant definitions, we summarize in Figure 2.4 the ladder of function spaces discussed in Section 2.1.3. In Figure 2.4, p_1 and p_2 are real numbers such that $1 < p_1 < 2 < p_2 < \infty$, and N_1 and N_2 are natural numbers such that $N_1 \le N_2$. The blank box appears because there is no commonly used symbol for the class of differentiable functions. Notice that as we move down the ladder into the smaller function classes, the functions become better behaved. For instance, the L^p spaces on $\mathbb{T}$ are nested and decreasing as the real number p increases from 1 to ∞ (Exercise 2.23).

Two warnings are in order. First, when the underlying space is changed from the circle $\mathbb{T}$ to another space X, the L^p spaces need not be nested: for real numbers p, q with $1 \le p \le q \le \infty$, $L^p(X)$ need not contain $L^q(X)$. For example, the L^p spaces on the real line $\mathbb{R}$ are not nested. (By contrast, the C^k classes are always nested: $C^k(X) \supset C^{k+1}(X)$ for all spaces X (in which differentaibility is defined) and for all $k \in \mathbb{N}$.) Second, as mentioned above, there are important function classes that do not fall into the nested chain of classes shown here.

$L^1(\mathbb{T})$	Lebesgue integrable: $\int_{\mathbb{T}} \lvert f(x)\rvert\, dx < \infty$
$\cup$	
$L^{p_1}(\mathbb{T})$	$\int_{\mathbb{T}} \lvert f(x)\rvert^{p_1}\, dx < \infty\,,\quad 1 < p_1 < 2$
$\cup$	
$L^2(\mathbb{T})$	square integrable: $\int_{\mathbb{T}} \lvert f(x)\rvert^2\, dx < \infty$
$\cup$	
$L^{p_2}(\mathbb{T})$	$\int_{\mathbb{T}} \lvert f(x)\rvert^{p_2}\, dx < \infty\,,\quad 2 < p_1 < \infty$
$\cup$	
$L^\infty(\mathbb{T})$	essentially bounded: $\operatorname{ess\,sup}_{x\in\mathbb{T}} \lvert f(x)\rvert < \infty$
$\cup$	
$B(\mathbb{T})$	bounded functions: $\lvert f(x)\rvert \le M$ for all $x \in \mathbb{T}$
$\cup$	
$\mathcal{R}(\mathbb{T})$	Riemann integrable
$\cup$	
$\mathcal{C}(\mathbb{T})$	continuous
$\cup$	
$\boxed{}$	differentiable
$\cup$	
$C^1(\mathbb{T})$	continuously differentiable: f' exists and is continuous
$\cup$	
$C^2(\mathbb{T})$	twice continuously differentiable: $f,\ f'$ exist and are continuous
$\cup$	
$\vdots$	$\vdots$
$\cup$	
$C^k(\mathbb{T})$	k times continuously differentiable: $f,\ f',\ \dots,\ f^{(k)}$ exist and are continuous
$\cup$	
$\vdots$	$\vdots$
$\cup$	
$C^\infty(\mathbb{T})$	infinitely often differentiable: $f,\ f',\ f'',\ \dots,$ exist and are continuous
$\cup$	
$\vdots$	$\vdots$
$\cup$	
$\mathcal{P}_{N_2}(\mathbb{T})$	trigonometric polynomials of degree $N_2 > N_1$
$\cup$	
$\mathcal{P}_{N_1}(\mathbb{T})$	trigonometric polynomials of degree $N_1 > 0$
$\cup$	
$\vdots$	$\vdots$
$\cup$	
$\mathcal{P}_0(\mathbb{T})$	constant functions

FIGURE 2.4. Ladder of nested classes of functions $f : \mathbb{T} \to \mathbb{C}$.

2.2. Modes of convergence

Given a bounded interval I and functions g_n, $g : I \to \mathbb{C}$, what does it mean to say that $g_n \to g$? In this section we discuss seven different ways in which the functions g_n could approximate g: pointwise; almost everywhere; uniformly; uniformly outside a set of measure zero; in the sense of convergence in the mean (convergence in L^1); in the sense of mean-square convergence (convergence in L^2); and in the sense of convergence in L^p. The first four of these modes of convergence are defined in terms of the behavior of the functions g_n and g at individual points x, therefore the itnerval I can be replaced by any subset X of real numbers. The last three of these modes are defined in terms of integrals of powers of the difference $|g_n - g|$, in these cases I can be replaced by any set of real numbers X for which the integral over the set is defined. Some of these modes of convergence are stronger than others, uniform convergence being the strongest. Figure 2.5 illustrates the interrelations between five of these modes of convergence.

In Chapter 3, we will be concerned with pointwise and uniform convergence of Fourier series, and in Chapter 5 we will be concerned with mean-square convergence of Fourier series. At the end of the book, we come full circle and show that convergence in $L^p(\mathbb{T})$ of Fourier series is a consequence of the boundedness in $L^p(\mathbb{T})$ of the *periodic Hilbert transform*.

See [**Bart1**, pp.68–72] for yet more modes of convergence, such as convergence in measure and almost uniform convergence (not to be confused with uniform convergence outside a set of measure zero).

We begin with the modes of convergence that are defined in terms of individual points x. In what follows X is a subset of the real numbers.

Definition 2.34. A sequence of functions $g_n : X \to \mathbb{C}$ *converges pointwise* to $g : X \to \mathbb{C}$ if for each $x \in X$,

$$\lim_{n\to\infty} g_n(x) = g(x).$$

Equivalently, g_n converges pointwise to g if for each $\varepsilon > 0$ and for each $x \in X$ there is an $N = N(\varepsilon, x) > 0$ such that for all $n > N$,

$$|g_n(x) - g(x)| < \varepsilon.$$

$\Diamond$

Almost-everywhere convergence is a little weaker than pointwise convergence.

Definition 2.35. A sequence of functions $g_n : X \to \mathbb{C}$ *converges almost everywhere*, or *pointwise a.e.*, to $g : X \to \mathbb{C}$ if it converges pointwise to g except on a set of measure zero:

$$\lim_{n\to\infty} g_n(x) = g(x) \quad \text{a.e. in } X.$$

Equivalently, g_n converges a.e. to g if there exists a set E of measure zero, $E \subset X$, such that for each $x \in X \setminus E$ and for each $\varepsilon > 0$ there is an $N = N(\varepsilon, x) > 0$ such that for all $n > N$,

$$|g_n(x) - g(x)| < \varepsilon.$$

$\Diamond$

Uniform convergence is stronger than pointwise convergence. Informally, given a fixed ε, for uniform convergence the same N must work for all x, while for pointwise convergence we may use different numbers N for different points x.

Uniform convergence:
$|g_n(x) - g(x)| \to 0$
uniformly in x.
For each $\varepsilon > 0$ there is an N
that works for all $x \in \mathbb{T}$.

$\swarrow \qquad \searrow$

Pointwise convergence:
$|g_n(x) - g(x)| \to 0$
for all $x \in \mathbb{T}$.
For each $\varepsilon > 0$ and each $x \in \mathbb{T}$
there is an $N = N(x, \varepsilon)$ that works.

L^2 *(mean-square) convergence*:
$\int_{\mathbb{T}} |g_n(x) - g(x)|^2 \, dx \to 0.$

$\Downarrow \qquad \Downarrow$

Pointwise a.e. convergence:
$|g_n(x) - g(x)| \to 0$
for almost all $x \in \mathbb{T}$.
For each $\varepsilon > 0$ and almost every $x \in \mathbb{T}$
there is an $N = N(x, \varepsilon)$ that works.

L^1 *(mean) convergence*:
$\int_{\mathbb{T}} |g_n(x) - g(x)| \, dx \to 0.$

FIGURE 2.5. Relations between five of the seven modes of convergence discussed in Section 2.2, for functions g_n, $g : \mathbb{T} \to \mathbb{C}$, showing which modes imply which other modes. All potential implications between these five modes other than those shown here are false.

DEFINITION 2.36. A sequence of bounded functions $g_n : X \to \mathbb{C}$ *converges uniformly* to $g : X \to \mathbb{C}$ if given $\varepsilon > 0$ there is an $N = N(\varepsilon) > 0$ such that for all $n > N$,

$$|g_n(x) - g(x)| < \varepsilon \quad \text{for all } \; x \in X.$$

Equivalently, g_n converges uniformly to g if

$$\lim_{n \to \infty} \sup_{x \in X} |g_n(x) - g(x)| = 0.$$

$\Diamond$

Here is a slightly weaker version of uniform convergence.

Definition 2.37. A sequence of L^∞-functions $g_n : X \to \mathbb{C}$ *converges uniformly outside a set of measure zero*, or *in* L^∞, to $g : X \to \mathbb{C}$ if

$$\lim_{n\to\infty} \|g_n - g\|_{L^\infty(X)} = 0.$$

We write $g_n \to g$ in $L^\infty(X)$. ◊

Remark 2.38. If the functions g_n and g are continuous and bounded on $X \subset \mathbb{R}$, then

$$\|g_n - g\|_{L^\infty(X)} = \sup_{x\in X} |g_n(x) - g(x)|,$$

and in this case convergence in $L^\infty(X)$ and uniform convergence coincide. This is why the L^∞-norm for continuous functions is often called the *uniform norm*. The metric space of continuous functions over a closed bounded interval with the metric induced by the uniform norm is a complete metric space (a Banach space), see [**Tao2**, Section 4.4]. In particular, if the functions g_n in Definition 2.36 are continuous and converge uniformly, then the limiting function is itself continuous; see Theorem 2.55.

We move to the modes of convergence defined via integrals. The two most important are convergence in L^1 and in L^2; these are special cases of convergence in L^p.

The integral used in these definitions is the Lebesgue integral, and by integrable we mean Lebesgue integrable. here we restrict to integrals on bounded intervals I.

Definition 2.39. A sequence of integrable functions $g_n : I \to \mathbb{C}$ *converges in the mean*, or *in* $L^1(I)$, to $g : I \to \mathbb{C}$ if for each $\varepsilon > 0$ there is a number $N > 0$ such that for all $n > N$

$$\int_I |g_n(x) - g(x)|\, dx < \varepsilon.$$

Equivalently, g_n converges in L^1 to g if $\lim_{n\to\infty} \|g_n - g\|_{L^1(I)} = 0$. We write $g_n \to g$ in $L^1(I)$. ◊

Definition 2.40. A sequence of square-integrable functions $g_n : I \to \mathbb{C}$ *converges in mean-square*, or *in* $L^2(I)$, to $g : I \to \mathbb{C}$ if for each $\varepsilon > 0$ there is an $N > 0$ such that for all $n > N$,

$$\int_I |g_n(x) - g(x)|^2\, dx < \varepsilon.$$

Equivalently, g_n converges in L^2 to g if $\lim_{n\to\infty} \|g_n - g\|_{L^2(I)} = 0$. We write $g_n \to g$ in $L^2(I)$. ◊

Convergence in $L^1(I)$ and in $L^2(I)$ are particular cases of convergence in $L^p(I)$.

Definition 2.41. For p such that $1 \le p < \infty$, a sequence of L^p-integrable functions $g_n : I \to \mathbb{C}$ *converges in* $L^p(I)$ to $g : I \to \mathbb{C}$ if given $\varepsilon > 0$ there exists $N > 0$ such that for all $n > N$

$$\int_I |g_n(x) - g(x)|^p\, dx < \varepsilon.$$

Equivalently, g_n converges in L^p to g if $\lim_{n\to\infty} \|g_n - g\|_{L^p(I)} = 0$. We write $g_n \to g$ in $L^p(I)$. ◊

We have already noted that the L^p-spaces are complete normed spaces, also known as Banach spaces. Therefore the limit functions in Definition 2.41 have no choice but to be in $L^p(I)$ themselves. In particular, in Definitions 2.39 and 2.40 the limit functions must be in $L^1(I)$ and $L^2(I)$ respectively.

EXERCISE 2.42. Verify that uniform convergence on a bounded interval I implies the other modes of convergence on I. ◊

EXERCISE 2.43. Show that for $1 \leq q < p \leq \infty$ convergence in $L^p(I)$ implies convergence in $L^q(I)$. **Hint:** use Hölder's Inequality (2.9). ◊

The following examples show that in Figure 2.5 the implications go only in the directions indicated. In these examples we work on the unit interval $[0, 1]$.

EXAMPLE 2.44. (*Pointwise but not uniform*) The sequence of functions defined by $f_n(x) = x^n$ for $x \in [0, 1]$ converges pointwise to the function f defined by $f(x) = 0$ for $0 \leq x < 1$ and $f(1) = 1$, but it does not converge uniformly to f or to any other function. See Figure 2.6. ◊

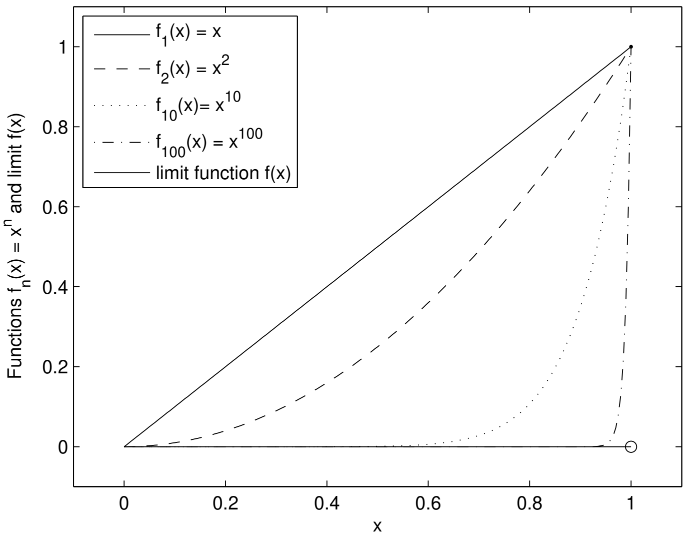

FIGURE 2.6. The functions $f_n(x) = x^n$ converge pointwise, but not uniformly, on $[0, 1]$ to the function $f(x)$ that is 0 for $x \in [0, 1)$ and 1 at $x = 1$.

EXAMPLE 2.45. (*Pointwise but not in the mean*) The sequence of functions $g_n : [0, 1] \to \mathbb{R}$ defined by

$$g_n(x) = \begin{cases} n, & \text{if } 0 < x < 1/n; \\ 0, & \text{otherwise,} \end{cases}$$

converges pointwise everywhere to $g(x) = 0$, but it does not converge in the mean or in L^2, nor does it converge uniformly. See Figure 2.7. ◊

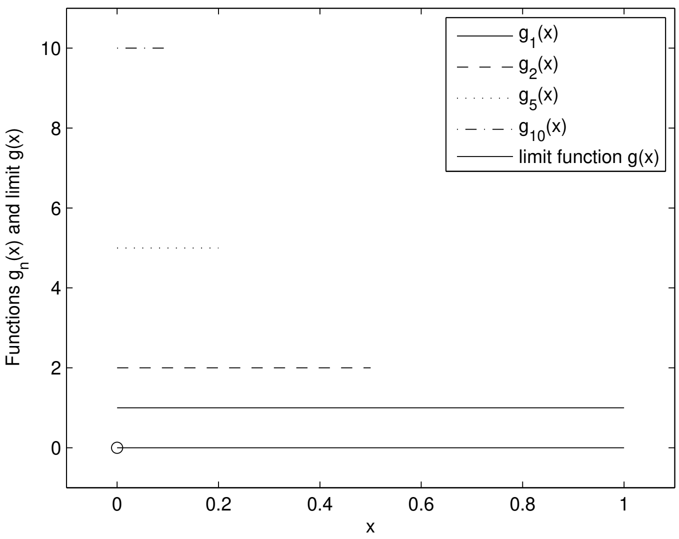

FIGURE 2.7. The functions g_n defined in Example 2.45 converge pointwise, but not in the mean, in L^2, or uniformly.

EXAMPLE 2.46. (*In the mean but not pointwise*) Given a positive integer n, let m, k be the unique non-negative integers such that $n = 2^k + m$ and $0 \leq m < 2^k$. The sequence of functions $h_n(x) : [0, 1] \to \mathbb{R}$ defined by

$$h_n(x) = \begin{cases} 1, & \text{if } \dfrac{m}{2^k} < x < \dfrac{m+1}{2^k}; \\ 0, & \text{otherwise,} \end{cases}$$

converges in the mean to the zero function, but it does not converge pointwise at any point $x \in [0, 1]$. This sequence of functions is sometimes called the *walking function*. See Figure 2.8. $\Diamond$

Despite these examples, there is a connection between pointwise convergence and convergence in L^p. For p such that $1 \leq p < \infty$, if f_n, $f \in L^p(\mathbb{T})$, and if $f_n \to f$ in L^p, then there is a subsequence $\{f_{n_k}\}$ that converges to f a.e. on $\mathbb{T}$. See [**SS2**, Corollary 2.2].

EXERCISE 2.47. Verify that the sequences described in Examples 2.44, 2.45, and 2.46 do have the convergence properties described there. Adapt the sequences in the examples to have the same convergence properties on the interval $[a, b]$. $\Diamond$

EXERCISE 2.48. Where would the two missing modes, namely convergence outside a set of measure zero and L^p convergence for a given p with $1 \leq p \leq \infty$, fit in Figure 2.4? Prove the implications that involve these two modes, and devise examples to show that other potential implications involving these two modes do not hold. $\Diamond$

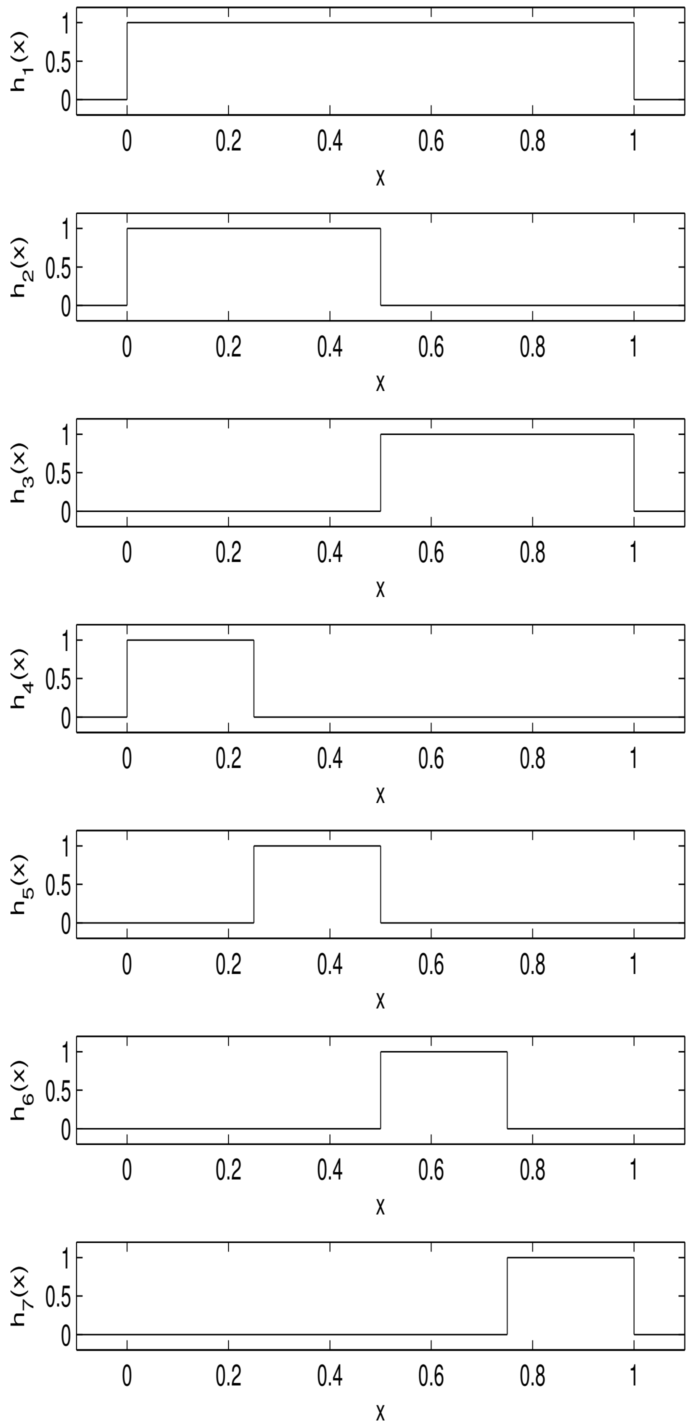

FIGURE 2.8. The functions $h_n(x)$ defined in Example 2.46 converge in mean, but do not converge pointwise at any x.

2.3. Interchanging limit operations

Uniform convergence on closed bounded intervals is a very strong form of convergence which is amenable to interchange with other limiting operations and properties defined by limits, such as integration, differentiation and continuity. We list here without proof several results from advanced calculus that will be used throughout the book. Some results do involve uniform convergence as one of the limit operations, some others involve other limit operations (like interchanging two integrals).

THEOREM 2.49 (Interchange of Limits and Integrals). *If $\{f_n\}_{n\in\mathbb{N}}$ is a sequence of integrable functions that converges uniformly to f in $[a,b]$ then f is integrable on $[a,b]$ and*

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

In other words,

The limit of the integrals of a uniformly convergent sequence of functions on $[a,b]$ is the integral of the limit function on $[a,b]$.

Notice that the functions in Example 2.45 do not converge uniformly, and they provide an example where this interchange fails. In fact, $\int_0^1 g_n = 1$ for all n and $\int_0^1 g = 0$, so

$$\lim_{n\to\infty}\int_0^1 g_n = 1 \neq 0 = \int_0^1 \lim_{n\to\infty} g_n = \int_0^1 g.$$

EXERCISE 2.50. Given a closed bounded interval I and integrable functions $g_n : I \to \mathbb{C}$, $n \in \mathbb{N}$, suppose the partial sums $\sum_{n=0}^{N} g_n(x)$ converge uniformly on I as $N \to \infty$. Show that

$$\int_I \sum_{n=0}^{\infty} g_n(x)\,dx = \sum_{n=0}^{\infty}\int_I g_n(x)\,dx.$$

$\Diamond$

We can interchange infinite summation and integration provided the convergence of the partial sums is uniform.

It is therefore important to know when do partial sums converge uniformly to their series. The Weierstrass M-Test is a very useful technique to deduce uniform convergence of partial sums of functions.

THEOREM 2.51 (Weierstrass M-Test). *Let $\{f_n\}_{n=1}^{\infty}$ be a sequence of bounded real-valued functions on a subset X of the real numbers. Assume there is a sequence of positive real numbers such that $|f_n(x)| \leq a_n$ for all $x \in X$, and $\sum_{n=1}^{\infty} a_n < \infty$. Then the series $\sum_{n=1}^{\infty} f_n$ converges uniformly to a real-valued function on X.*

Under certain conditions, involving uniform convergence of a sequence of functions and their derivatives, one can exchange differentiation and the limit.

THEOREM 2.52 (Interchange of Limits and Derivatives). *Let $f_n : [a,b] \to \mathbb{C}$ be a sequence of C^1 functions converging uniformly on $[a,b]$ to the function f. Assume* *that the derivatives $(f_n)'$ also converge uniformly on $[a,b]$ to some function g. Then f is C^1, and $f' = g$, that is*

$$\lim_{n\to\infty} (f_n)' = \Big(\lim_{n\to\infty} f_n\Big)'.$$

In other words,

> *We can interchange limits and differentiation provided the convergence of the functions and the derivatives is uniform.*

Exercise 2.53. Show that the hypotheses in Theorem 2.52 are necessary. $\Diamond$

Exercise 2.54. What are the necessary hypotheses on a series $\sum_{n=0}^{\infty} g_n(x)$ so that it is legal to interchange sum and differentiation? In other words, when can we guarantee that

$$\left(\sum_{n=0}^{\infty} g_n(x)\right)' = \sum_{n=0}^{\infty} (g_n)'(x)?$$

$\Diamond$

Here is another classical and useful result in the same vein, already mentioned in Remark 2.38.

Theorem 2.55. *The uniform limit of continuous functions on a subset X of the real numbers is continuous on X.*

Exercise 2.56. Prove Theorem 2.55.

Notice that this is the same as saying that if $g_n \to g$ uniformly, and the functions g_n are continuous, then for all $x_0 \in X$ we can interchange the limit as x approaches x_0 with the limit as n tend to infinity,

$$\lim_{x\to x_0} \lim_{n\to\infty} g_n(x_0) = \lim_{n\to\infty} \lim_{x\to x_0} g_n(x). \tag{2.10}$$

Perhaps inserting a few more equalities will make the above statement clear:

$$\begin{aligned}
\lim_{x\to x_0} \lim_{n\to\infty} g_n(x) &= \lim_{x\to x_0} g(x) \qquad (g_n \text{ converges unif. to } g)\\
&= g(x_0) \qquad (g \text{ is continuous at } x_0)\\
&= \lim_{n\to\infty} g_n(x_0) \qquad (g_n \text{ converges unif. to } g)\\
&= \lim_{n\to\infty} \lim_{x\to x_0} g_n(x) \qquad (g_n \text{ is continuous at } x_0).
\end{aligned}$$

Example 2.44 shows a sequence of continuous functions on $[0,1]$ that converge pointwise to a function that is discontinuous (at $x = 1$). Thus the convergence cannot be uniform.

There is a partial converse, called Dini's[15] Theorem, to Theorem 2.55. It says that the pointwise limit of a decreasing sequence of functions is also their uniform limit.

Theorem 2.57 (Dini's Theorem). *Suppose that f, $f_n : I \to \mathbb{R}$ are continuous functions. Assume that (i) f_n converges pointwise to f as $n \to \infty$; and (ii) f_n is a decreasing sequence, namely $f_n(x) \ge f_{n+1}(x)$ for all $x \in I$. Then f_n converges uniformly to f.*

[15] Ulisse Dini, Italian mathematician (1845–1918).

EXERCISE 2.58. Prove Dini's Theorem, and show that hypothesis (ii) is essential. A continuous version of Example 2.45, where the "steps" are replaced by "tents", might help. ◊

In Weiertrass's M-Test, Theorem 2.51, if the functions $\{f_n\}_{n=1}^{\infty}$ being summed up are continuous then the series $\sum_{n=1}^{\infty} f_n(x)$ is a continuous function.

Interchanging integrals is another instance of these interchange-of-limit operations. The theorems that allow for such interchange go by the name Fubini's Theorems[16]. Here is one such theorem, valid for continuous functions.

THEOREM 2.59 (Fubini's Theorem for Continuous Functions). *Let I and J be closed intervals. Let $F : I \times J \to \mathbb{C}$ be a continuous function. Assume*

$$\iint_{I\times J} |F(x,y)|\,dA < \infty,$$

where dA denotes the differential of area. Then

$$\int_I \int_J F(x,y)\,dx\,dy = \int_J \int_I F(x,y)\,dy\,dx = \iint_{I\times J} F(x,y)\,dA.$$

Usually Fubini's Theorem refers to an L^1-version of Theorem 2.59 that is stated and proved in every measure theory book. You can find a two-dimensional version in [**Tao2**, Section 19.5].

Interchanging limit operations is a delicate maneuver that cannot always be accomplished. We illustrate throughout the book settings where this interchange is allowed and settings where it is illegal, such as the ones just described involving uniform convergence. If we insist on carrying on the interchange, false conclusions could be deduced.

Unlike the Riemann integral, the Lebesgue theory allows for the interchange of a pointwise limit and an integral. There are several landmark theorems that you learn in a course on measure theory. We state one such result.

THEOREM 2.60 (Lebesgue Dominated Convergence Theorem). *Consider a sequence of measurable functions f_n defined on the interval I, converging pointwise a.e. to a function f. Suppose there exists a dominating function $g \in L^1(I)$, meaning that $|f_n(x)| \leq g(x)$ a.e. for each $n > 0$. Then*

$$\lim_{n\to\infty} \int_I f_n(x)\,dx = \int_I f(x)\,dx.$$

EXERCISE 2.61. Verify that the functions in Example 2.45 cannot be dominated by an integrable function g. ◊

2.4. Density

Having different ways to measure convergence of functions to other functions provides ways of deciding when we can approximate functions in a given class by functions in another class. In particular we can decide when a subset is dense in a larger set of functions.

[16]Named after the Italian mathematician Guido Fubini (1879–1943).

FIGURE 2.9. The graph shows a step function whose corners have been rounded.

DEFINITION 2.62. Given a normed space X of functions defined on a bounded interval I, with norm denoted by $\|\cdot\|_X$, we say that a subset $A \subset X$ is *dense in* X if given any $f \in X$ there exists a sequence of functions $f_n \in A$ such that

$$\lim_{n\to\infty} \|f_n - f\|_X = 0.$$

Equivalently, A is dense in X if given any $f \in X$ and any $\varepsilon > 0$ there exists a function $h \in A$ such that

$$\|h - f\|_X < \varepsilon.$$

$\Diamond$

One can deduce from the definition of the Riemann-integrable functions on a bounded interval I that the step functions (see Definition 2.3) are dense in $\mathcal{R}(I)$ with respect to the L^1-norm; see Proposition 2.10. We state this result as a theorem, for future reference.

THEOREM 2.63. *Given a Riemann-integrable function on a bounded interval* I*,* $f \in \mathcal{R}(I)$*, and given* $\varepsilon > 0$*, there exists a step function* h *such that*

$$\|h - f\|_{L^1(I)} < \varepsilon.$$

REMARK 2.64. We can choose the step functions that approximate a given Riemann-integrable function to lie entirely below the function, or entirely above it.

In Chapter 4 we will learn that

> *Trigonometric polynomials can approximate uniformly continuous functions on* $\mathbb{T}$.

This is the celebrated Weierstrass theorem (see Theorem 3.4).

In Chapter 9 we will learn that on closed bounded intervals,

> *Step functions can approximate uniformly continuous functions, despite the fact that step functions are not continuous.*

It is not hard to deduce directly from the fact that continuous functions can be approximated uniformly by step functions that they can also be approximated in the L^p-norm by step functions.

> *Step functions can approximate continuous functions in the* L^p*-norm.*

What is perhaps surprising is that

> *Continuous functions can approximate step functions pointwise and in the* L^p*-norm.*

We verify this result in Chapter 4, when we introduce the concepts of convolution and approximations of the identity. If you draw a picture of a step function and then round off the corners, this result seems more reasonable (see Figure 2.4).

Assuming these results, proved in later chapters, we can establish the density of the continuous functions in the class of Riemann-integrable functions and in L^p-spaces.

THEOREM 2.65. *The continuous functions on a closed bounded interval I are dense in $\mathcal{R}(I)$ with respect to the L^p-norm.*

PROOF. Start with $f \in \mathcal{R}(I)$. There is a step function h that is ε-close in the L^p-norm to f, that is $\|f-h\|_{L^p(I)} < \varepsilon$. There is also a continuous function $g \in C(I)$ that is ε-close in the L^p-norm to the step function h, that is $\|h-g\|_{L^p(I)} < \varepsilon$. By the triangle inequality, the continuous function g is 2ε-close in the L^p-norm to our initial function f:

$$\|f-g\|_{L^p(I)} \leq \|f-h\|_{L^p(I)} + \|h-g\|_{L^p(I)} \leq 2\varepsilon.$$

Hence continuous functions can get arbitrarily close in the L^p-norm to any Riemann-integrable function. $\square$

REMARK 2.66. We will use a very precise version of this result for $p = 1$ in Chapter 4 (see Lemma 4.8), where we require the approximating continuous functions to be uniformly bounded by the same bound that controls the Riemann-integrable function. Here is the idea. Given real-valued $f \in \mathcal{R}(I)$ and $\varepsilon > 0$, there exists a real-valued continuous function g on I (by Theorem 2.65) such that $\|g-f\|_{L^1(I)} < \varepsilon$. Let f be bounded by $M > 0$, that is $|f(x)| \leq M$ for all $x \in I$. If g is bounded by M there is nothing to be done. If g is not bounded by M, consider the new function defined on I by $g_0(x) = g(x)$ if $|g(x)| \leq M$, $g(x) = M$ if $g(x) > M$, and $g(x) = -M$ if $g(x) < -M$. One can show that g_0 is continuous, $g_0(x) \leq M$ for all $x \in I$, and g_0 is closer to f than g is. In particular the L^1-norm of the difference is at most ε, and possibly even smaller.

EXERCISE 2.67. Show that the function g_0 defined in Remark 2.66 satisfies the properties claimed there. $\Diamond$

EXERCISE 2.68. Let X be a normed space of functions over a bounded interval I. Suppose A is a dense subset of X, and $A \subset B \subset X$. Then B is a dense subset of X (all with respect to the norm of the ambient space X). $\Diamond$

EXERCISE 2.69. (*Density is a Transitive Property*) Let X be a normed space of functions over a bounded interval I. Suppose A is a dense subset of B, and B is a dense subset of X. Then A is a dense subset of X (all with respect to the norm of the ambient space X). $\Diamond$

Notice that the notion of density is strongly dependent on the norm or metric of the ambient space. For example, when considering nested L^p-spaces on a closed bounded interval, in Section 2.1 we learned that $C(I) \subset L^2(I) \subset L^1(I)$. Suppose we knew that the continuous functions are dense in $L^2(I)$ (with respect to the L^2-norm), and the square-integrable functions are dense in $L^1(I)$ (with respect to the L^1-norm). Even so, we could not use transitivity to conclude that the continuous functions are dense in $L^1(I)$. We need to know more about the relation between the L^1- and the L^2-norms. In fact, to complete the argument, it suffices to know that $\|f\|_{L^1(I)} \leq C\|f\|_{L^2(I)}$ (see Exercise 2.23), because that would imply that if the continuous functions are dense in $L^2(I)$ (with respect to the L^2-norm), then they are also dense with respect to the L^1-norm, and now we can use the transitivity.

Such norm inequality is stronger than the statement that $L^2(I) \subset L^1(I)$. It says that the *identity map* or *embedding* $E : L^2(I) \to L^1(I)$, defined by $Ef = f$ for $f \in L^2(I)$, is a *continuous map* or a *continuous embedding*.

We do not prove the following statements about various subsets of functions being dense in $L^p(I)$, since they rely on Lebesgue integration theory. However, we use these results in subsequent chapters, for example to show that the trigonometric system is a basis for $L^2(\mathbb{T})$.

We know that $L^p(I)$ is the completion of $\mathcal{R}(I)$ with respect to the L^p-norm; see Theorem 2.17. In particular,

> *The Riemann-integrable functions are dense in $L^p(I)$ with respect to the L^p-norm.*

With this fact in mind, we see that any set of functions that is dense in $\mathcal{R}(I)$ with respect to the L^p-norm is also dense in $L^p(I)$ with respect to the L^p-norm, by transitivity. For future reference we state here the specific density of the continuous functions in $L^p(I)$.

THEOREM 2.70. *The continuous functions on a closed bounded interval I are dense in $L^p(I)$, for all p with $1 \le p \le \infty$.*

EXERCISE 2.71. Given $f \in L^p(I)$, $1 \le p < \infty$, and $\varepsilon > 0$, and assuming all the theorems and statements in this section, show that there exists a step function h on I such that

$$\|h(x) - f(x)\|_{L^p(I)} < \varepsilon.$$

$\Diamond$